The Journal of Fourier Analysis and Applications
Volume 3, Number 1, 1997

Unique Reconstruction of Band-limited Signals by a Mallat-Zhong Wavelet Transform Algorithm

C. J. Kicey and C. J. Lennard

ABSTRACT. We show that uniqueness and existence for signal reconstruction from multiscale edges in the Mallat and Zhong algorithm become possible if we restrict our signals to Paley-Wiener space, band-limit our wavelets, and irregularly sample at the wavelet transform (absolute) maxima—the edges—while possibly including (enough) extra points at each level. We do this in a setting that closely resembles the numerical analysis setting of Mallat and Zhong and that seems to capture something of the essence of their (practical) reconstruction method. Our work builds on a uniqueness result for reconstructing an L^2 signal from irregular sampling of its wavelet transform of Gröchenig and the related work of Benedetto, Heller, Mallat, and Zhong. We show that the rate of convergence for this reconstruction algorithm is geometric and computable in advance. Finally, we consider the effect on the rate of convergence of not sampling enough local maxima.

1. Introduction

We begin by informally discussing the general framework we will use for a version of the Mallat-Zhong signal reconstruction algorithm introduced in Mallat and Zhong [MZ1, MZ2].

We consider a signal $f \in L^2 = L^2(\mathbf{R})$. Next we introduce a *scaling* or *smoothing* function φ, with corresponding *wavelet* ψ, about which we will say more soon. (See Figs. 2 and 3 for our main examples.) Throughout this paper, for any function $u : \mathbf{R} \to \mathbf{C}$, we define $u_s(x) := (1/s)u(x/s)$ for all $x \in \mathbf{R}$, and all $s > 0$. We *smooth* f *at level* 1 by φ to get $f * \varphi_1$ and then decompose f into $J+1$ levels ($J \geq 1$): $f * \psi_2, f * \psi_{2^2}, \ldots, f * \psi_{2^J}, f * \varphi_{2^J}$, which are the details of the signal at J coarser and coarser levels, followed finally by the resulting low resolution signal corresponding to f.

Define $\mathbf{W}f := (f * \psi_2, f * \psi_{2^2}, \ldots, f * \psi_{2^J}, f * \varphi_{2^J})$, for all $f \in L^2(\mathbf{R})$. We call $\mathbf{W}f$ the wavelet transform of f, and $\mathbf{W}$ the wavelet transform. We have

$$\|\mathbf{W}f\|^2_{\ell^2(L^2)} = \|f * \varphi\|^2_{L^2}.$$

Math Subject Classifications. Primary 42C; Secondary 46C
Keywords and Phrases. wavelet transform, band-limited function, Mallat-Zhong algorithm, Paley-Wiener theory, alternating projections, signal reconstruction, frames, Shannon's sampling theorem, Hilbert transform, local maxima
Acknowledgements and Notes. The authors thank Ka-Sing Lau, Juan Manfredi, and the members of the wavelet seminar, especially Ching-Chun Li, Wonkoo Kim, Hyung Jun Kim, and Mingui Sun, for many valuable insights into wavelet theory, Paley-Wiener theory, and the Mallat-Zhong algorithm. We thank Catherine for typing part of the manuscript. We have also benefitted from conversations and communications with Peter Dodds, Jon Borwein, Jerry Kautsky, Radka Turcajova, Heinz Bauschke, Akram Aldroubi, Chris Heil, John Benedetto, Karlheinz Gröchenig, Hans Feichtinger, Brailey Sims, and Susumu Okada. Finally, we thank the referees for their helpful suggestions. The research of C. J. Kicey was partially supported by a University of Pittsburgh Office of Research Small grant. The research of C. J. Lennard was partially supported by a University of Pittsburgh FAS grant.

ISSN 1069-5869

This follows because we may arrange, and will always assume, that φ and ψ satisfy, for almost all $\xi \in \mathbf{R}$,

$$|\widehat{\varphi}(\xi)|^2 = |\widehat{\psi}(2\xi)|^2 + |\widehat{\varphi}(2\xi)|^2.$$

We remark that an orthonormal wavelet pair (φ, ψ) arising from a multiresolution analysis always satisfies this property, although the converse does not hold. We will discuss this again in §5.

The aim of the Mallat-Zhong algorithm is to record a discrete set of data, the local maxima of the functions $|f * \psi_{2^j}|$, which contain edge information of certain dyadic frequency ranges of the signal, and iteratively reconstruct the signal via the method of alternating projections onto convex sets in Hilbert space. Mallat and Zhong [MZ1, MZ2] use various convex sets and settle on the Hilbert space $H^1(\mathbf{R})$ of L^2 signals f whose distributional derivative f' also belongs to L^2. A key feature of their method is the use of dilation equations for φ (with a finite number of nonzero coefficients, or with coefficients that decrease rapidly at ∞ and $-\infty$) to allow fast numerical implementation of the wavelet transform on a discrete and *finite* domain of regularly spaced sample points. They obtain a computationally fast numerical algorithm, which seems to converge geometrically in [MZ1] and [MZ2], with theoretical calculations pointing to the reason for this: A certain family of functions related to the sampling at local maxima should be a frame (see Duffin and Schaeffer [DS], Benedetto [Ben]).

At this point, let us remark that a variation on the theme of the Mallat and Zhong algorithm has recently been introduced and successfully numerically implemented by Çetin and Ansari [CA].

The theoretical method in [MZ1, MZ2], uses a wavelet transform with an infinite number of levels. A counterexample of Meyer [Me2, Me3] shows that the theoretical Mallat and Zhong algorithm need not converge to the original signal; that is, two different signals $f \in H^1$ may have the same local extrema at all levels $f * \psi_{2^j}$, where $j \in \mathbf{Z}$, of their wavelet transform. Berman [Ber1, Ber2] showed that also in the case of the (finite number of levels) numerical Mallat-Zhong algorithm, uniqueness of reconstruction is not generally assured.

None of these counterexamples address band-limited signals with band-limited wavelets; although [MZ2] mentions the work of Gröchenig. In addition, in [G2] Gröchenig suggests that the link between band-limited signals and the method of alternating projections should be explored. The important paper of Gröchenig [G3, Theorem 1] shows that for the wavelet transform of Mallat and Zhong with an infinite number of levels, any $f \in L^2$ is uniquely determined by irregularly sampling every level of the wavelet transform sufficiently often, when ψ is bandlimited. Moreover, [G3, Corollary 1] gives a frame algorithm with a computable geometric rate of convergence for reconstructing f from the irregular samples of $\mathbf{W}f$. And indeed, the generality of [G3] also allows for the irregular sampling of the *levels* of the wavelet transform, as well as including adaptive weights, so that only upper bounds (not lower bounds) of the sampling rates at each level are needed.

Our reasons for choosing our setting are the following. We look only at the wavelet transform with a finite number of levels because this is all that is needed for practical image reconstruction. Our restriction to band-limited signals is because the discrete signals on an evenly spaced grid (e.g., those considered by Mallat and Zhong) may naturally be thought of, from the point of view of image reconstruction, as band-limited. In addition, band-limited signals often arise naturally. We stay with dyadic levels of the (continuous) wavelet transform for simplicity and ease of comparison to [MZ1, MZ2]. We ask for upper and lower bounds on our sampling rates at each level, again, for simplicity but also because once we restrict ourselves to discrete signals (using the pyramid algorithm to compute wavelet and inverse wavelet transforms), we will naturally have a lower bound on our sampling rate at each level. Moreover, gaining estimates on the convergence rate of the reconstruction algorithm in $\ell^2(P_\Omega)$ seems to depend on lower bounds on the sampling rates (see the proof of Theorem 4.2).

Of course, Gröchenig [G3] invites comparison of convergence rates of the frame algorithm versus the Mallat-Zhong alternating projections algorithm for signal reconstruction. Some calculations show that the convergence rates estimates of this paper (see Corollary 4.3 and §6) are much higher than are the actual rates gained in [MZ1, MZ2]. We mention that many other reconstruction algorithms for band-limited functions are introduced or described and numerically compared in Feichtinger and Gröchenig [FG]. We will briefly discuss why the Mallat-Zhong algorithm is not a frame algorithm and related questions later (see Remark 3.2). Let us further remark that if the signal f is not band-limited, and the wavelet ψ is band-limited (as in [G3]), the convergence of the algorithm and the errors introduced in reconstruction (such as aliasing) require further study.

We begin by recapitulating in our setting the essence of Gröchenig [G3, Theorem 1], using theorems of Benedetto and Heller [BH], Heller [He], Benedetto [Ben], and Gröchenig [G1, G2]. We show that reconstruction of a band-limited signal by a Mallat-Zhong algorithm, with both existence and uniqueness, is possible. One should irregularly sample at the (absolute) wavelet transform maxima (*the edges*) while including enough extra points at each level to ensure that sampling is above the Nyquist rate of each level (which halves from one level to the next). Another advantage of this framework is that we are assured that (the real and imaginary parts of) our nonzero signals *have* a discrete family of positions where local extrema occur, possibly accumulating only at ∞ or $-\infty$, since they have an analytic extension to the whole complex plane. (See, e.g., Rudin [R] or Benedetto [Ben], for a discussion covering the Paley-Wiener theory that we will use.)

We give a proof of a computable geometric convergence rate by using the theorems on irregular sampling of bandlimited functions of Benedetto and Heller [BH], Benedetto [Ben], and Gröchenig [G1, G2] and adapting the methods of [MZ2] to our framework.

Next we give a variation on this theme for real-valued L^2 functions involving the Hilbert transform that allows for approximately quartering the previous sampling rates.

Continuing on, we then investigate the discrete sampling case, starting with a signal consisting of time samples on an evenly spaced grid. Our variation on a theme, using the Hilbert transform, allows us to use the pyramid algorithm to easily compute discrete (though infinite) wavelet and inverse wavelet transforms. We show that all projections involved in the reconstruction algorithm may be computed in this setting.

Finally, we show that (again subsampling on an evenly spaced grid) there is an advantage to including the local maxima positions in our sample points. Indeed, by obtaining a lower bound on the convergence rate of the algorithm when started at zero in $\ell^2(P_\Omega)$, we show that not sampling in a way that makes the total sampled energy as large as practicable (compared with the total energy of the signal), may force the convergence to be slow.

2. Preliminaries

The real and complex numbers are denoted by **R** and **C**, respectively; **Z** is the set of all integers; and **N** denotes the set of positive integers. Let $1 \leq p < \infty$. We let $L^p = L^p(\mathbf{R})$ denote the space of all (equivalence classes) of Lebesgue-measurable functions $f : \mathbf{R} \to \mathbf{C}$ for which $\|f\|_p := \left(\int_{\mathbf{R}} |f|^p \, d\lambda\right)^{1/p} < \infty$. Here λ denotes Lebesgue measure on **R**. We have that $(L^p, \|\cdot\|_p)$ is a Banach space, as is $(L^\infty, \|\cdot\|_\infty)$, defined analogously, with $\|f\|_\infty := \text{ess-sup}_{x\in\mathbf{R}}|f(x)|$. For any Lebesgue-measurable subset A of **R**, $L^p(A)$ denotes the analogous space to $L^p(\mathbf{R})$, with the domains of the functions f changed from **R** to A. For $f \in L^1$, we define the Fourier transform $\widehat{f}$ of f by

$$\widehat{f}(\xi) := \int_{\mathbf{R}} f(x)\, e^{-i\xi x}\, dx \quad \text{for all } \xi \in \mathbf{R}$$

and the convolution $f * u$ of $f, u \in L^1$ by

$$(f * u)(x) := \int_{\mathbf{R}} f(y)\, u(x - y)\, dy \quad \text{for all } x \in \mathbf{R}.$$

Also, $u^{\vee}(x) := (1/2\pi)\widehat{u}(-x)$, $x \in \mathbf{R}$, gives the inverse Fourier transform of a function $u \in L^1$. Further, we define $\widetilde{u}(t) := \overline{u(-t)}$, for all $t \in \mathbf{R}$ and all functions $u : \mathbf{R} \to \mathbf{C}$.

Now, if $f, u \in L^1$, then $f * u \in L^1$ and $(f * u)(\xi) = \widehat{f}(\xi)\, \widehat{u}(\xi)$, for almost all $\xi \in \mathbf{R}$. Also, for all $f \in L^1$, $\widehat{f} \in C_0 = C_0(\mathbf{R})$, the space of all continuous scalar-valued functions on $\mathbf{R}$ that vanish at ∞ and $-\infty$, with the supremum norm $\|\cdot\|_\infty$. In addition, $\|f\|_\infty \leq \|f\|_1$. Moreover, Plancherel's theorem gives us that the Fourier transform extends from the dense subset $L^1 \cap L^2$ of L^2, to a constant multiple of an isometry from L^2 onto L^2. Indeed, for all $f, u \in L^2$,

$$\frac{1}{\sqrt{2\pi}} \|\widehat{f}\|_2 = \|f\|_2 \quad \text{and} \quad \frac{1}{2\pi} \langle \widehat{f}, \widehat{u} \rangle = \langle f, u \rangle.$$

Here, $\langle f, u \rangle := \int_{\mathbf{R}} f(x)\, \overline{u(x)}\, dx$, is the usual inner product on L^2. Note that for all $f, u \in L^2$, $f * u$ defined as above exists in L^∞ and $\|f * u\|_\infty \leq \|f\|_2 \|u\|_2$. Moreover, by Plancherel's theorem,

$$\begin{aligned} (f * u)(x) &= \frac{1}{2\pi} \int_{\xi \in \mathbf{R}} \widehat{f}(\xi)\, \overline{\overline{u(x - \cdot)}^{\wedge}(\xi)}\, d\xi \\ &= \frac{1}{2\pi} \int_{\xi \in \mathbf{R}} \widehat{f}(\xi)\, \widehat{u}(\xi)\, e^{ix\xi}\, d\xi = \left(\widehat{f} \cdot \widehat{u}\right)^{\vee}(x), \end{aligned}$$

since $\widehat{f} \cdot \widehat{u} \in L^1$. Thus we get that $f * u \in C_0$ for all $f, u \in L^2$. Now suppose that we have $f, u \in L^2$ and, moreover, that $\widehat{u} \in L^\infty$. Then $\widehat{f} \cdot \widehat{u} \in L^2$; therefore $\left(\widehat{f} \cdot \widehat{u}\right)^{\vee} \in L^2$. Thus $f * u \in L^2 \cap C_0$, and we get the formula $(f * u)^{\wedge} = \widehat{f} \cdot \widehat{u}$, in complete analogy with the situation for L^1 functions. We think of $f * u$ as a *smoothing of f by u*.

We refer the reader to Rudin [R], Champeney [Cha], and Körner [K], for example, for more details on these and other standard facts about the Fourier transform, convolution, and Fourier analysis in general. We follow the above-mentioned papers of Mallat and Zhong and of Gröchenig in our definition of Fourier transform, but differ from both Rudin and Benedetto.

Given $1 \leq p < \infty$, a set A, and a Banach space $(X, \|\cdot\|_X)$, we define $\ell^p(A, X)$ to be the set of all families (functions) $g = (g_n)_{n \in A}$ with values in X (i.e., each $g_n \in X$), such that

$$\|g\|_{\ell^p(A,X)} := \left(\sum_{n \in A} \|g_n\|_X^p \right)^{1/p} < \infty.$$

This function is a norm, turning $\ell^p(A, X)$ into a Banach space. We make the obvious modification to the definition above to get $\ell^\infty(A, X)$. For us, the index set A usually is simply $\{1, \ldots, N\}$, for some $n \in \mathbf{N}$. In this case, we write $\ell_N^p(X)$ (or just $\ell^p(X)$) instead of $\ell^p(A, X)$. If $(X, \langle \cdot, \cdot \rangle_X)$ is a Hilbert space, then so is $\ell^2(A, X)$, with the inner product given by $\langle g, h \rangle_{\ell^2(A,X)} := \sum_{n \in A} \langle g_n, h_n \rangle_X$, for all $g, h \in \ell^2(A, X)$.

We use the notation $\mathbf{P}_V$ to denote the orthogonal projection (or nearest point map) onto a closed, convex subset V of a Hilbert space. A sequence $(y_n)_n$ in a separable Hilbert space $(X, \langle \cdot, \cdot \rangle_X)$ is called a *frame* if there exist constants $0 < A \leq B < \infty$ such that

$$A\, \|x\|^2 \leq \sum_n |\langle x, y_n \rangle|^2 \leq B\, \|x\|^2 \quad \text{for all } x \in X.$$

The corresponding frame operator $S : X \to X$ is given by $Sx := \sum_n \langle x, y_n \rangle\, y_n$. Let $e_b(x) := e^{ixb}$, for all x and $b \in \mathbf{R}$. If $T > 0$ and $(b_n)_n$ is a sequence in $\mathbf{R}$, then $(e_{b_n})_n$ is called a Fourier frame for $L^2[-T, T]$ precisely when it is a frame for $L^2[-T, T]$.

For any Lebesgue-measurable set $E \subseteq \mathbf{R}$, χ_E is the usual characteristic function of E, which equals 1 at any $t \in E$ and 0 otherwise. We denote by supp$f$ the support of a measurable function $f : \mathbf{R} \to \mathbf{C}$, which is the closed set $K := \bigcap \Gamma$, where Γ is the collection of all closed subsets C of $\mathbf{R}$ satisfying $f(x) = 0$ for almost all $x \in \mathbf{R} \backslash C$. Since the real line $\mathbf{R}$ is second countable, every subset of $\mathbf{R}$ is Lindelöf. Therefore, K equals the intersection of only countably many members of Γ, which implies that K itself belongs to Γ. In particular, $f(x) = 0$ for almost all $x \notin K$. We remark that the support of f coincides with the usual definition of the support of the distribution generated by the locally integrable function f.

Finally, we refer to elements of $f \in L^2$ that are *real-valued* as *signals*. We use $\mathbf{x}$ to denote the identity function on $\mathbf{R}$. Also, $\mathcal{H}$ denotes the Hilbert transform, which as a distribution is simply convolution by $1/\pi\mathbf{x}$ and is given by the usual principal value integral formula. Extended to an isometry on L^2, $\mathcal{H}$ may be given in terms of its Fourier transform as

$$\widehat{\mathcal{H}f}(\xi) := -i \operatorname{sgn}(\xi)\, \widehat{f}(\xi) \quad \text{for almost all } \xi \in \mathbf{R}.$$

Here, $\operatorname{sgn} t := 1$, if $t > 0$; $\operatorname{sgn} t := -1$, if $t < 0$; and $\operatorname{sgn} 0 := 0$.

For any $\Delta > 0$, we define the Paley-Wiener space $PW_\Delta = P_\Delta$ to be the set of all $f \in L^2(\mathbf{R})$ such that $\widehat{f}(\xi) = 0$ for almost all $|\xi| > \Delta$. We remark that from the Paley–Wiener theory (see, e.g., [R] or [Ben]), every $f \in P_\Delta$ agrees almost everywhere with the restriction to the real line of an entire function $F : \mathbf{C} \to \mathbf{C}$ such that $|F(z)| \le C\, e^{\Delta|z|}$, for all $z \in \mathbf{C}$, for some $C \in (0, \infty)$. The converse fact that all such analytic functions, with restriction to the real line in L^2, have the support of their Fourier transform contained in $[-\Delta, \Delta]$, is an important theorem of Paley and Wiener [PW]. With the inner product inherited from L^2, P_Δ is a Hilbert space.

For any Lebesgue-measurable subset B of $\mathbf{R}$, we also denote by $P^{\mathbf{R}}(B)$ the Paley–Wiener space $\{f \in L^2 :\ \operatorname{supp} \widehat{f} \subseteq B \text{ and } f \text{ is real-valued}\}$. Fix $0 < \gamma < \delta < \infty$ and set $\Delta := (\delta - \gamma)/2$. Let $\mu := (\gamma + \delta)/2$ and $B := [-\delta, -\gamma] \cup [\gamma, \delta]$. It is straightforward to verify that, as real subspaces of L^2, $P^{\mathbf{R}}(B)$ and P_Δ are Banach space isomorphic under the mappings Φ and Φ^{-1} given by

$$(\Phi f)(t) := e^{-it\mu}\, (f(t) + i(\mathcal{H}f)(t)) \quad \text{for all } t \in \mathbf{R} \text{ and all } f \in P^{\mathbf{R}}(B),$$

$$(\Phi^{-1}u)(t) := \operatorname{Re}(e^{it\mu}\, u(t)) \quad \text{for all } t \in \mathbf{R} \text{ and all } u \in P_\Delta.$$

Indeed, Φ is an isometry, up to a multiplicative constant that is independent of γ and δ. We finally introduce the modified Paley-Wiener space, $P\mathcal{H}_\Delta$, given by

$$P\mathcal{H}_\Delta := \{f + i\mathcal{H}f : f \in P_\Delta \text{ and } f \text{ is real-valued}\}.$$

$P\mathcal{H}_\Delta$ is a closed (complex) subspace of P_Δ.

3. Uniqueness of Reconstruction in Paley–Wiener Space

Throughout this section we fix $\Omega > 0$. We also fix a function $\varphi \in L^2$ such that $\widehat{\varphi} \in L^\infty$ and $1 \ge |\widehat{\varphi}(\xi)| \ge C > 0$ on $[-\Omega, \Omega]$. We further assume that there are functions G and H in L^∞ such that $\widehat{\varphi}(\xi) = H(\xi/2)\, \widehat{\varphi}(\xi/2)$ and $|G(\xi)|^2 + |H(\xi)|^2 = 1$, for almost all $\xi \in \mathbf{R}$. We then define $v(\xi) := G(\xi/2)\, \widehat{\varphi}(\xi/2)$ for all $\xi \in \mathbf{R}$. Since $v \in L^2 \cap L^\infty$, it follows that $\psi := v^\vee$ belongs to L^2, with $\widehat{\psi} \in L^\infty$ and, moreover, for almost all $\xi \in \mathbf{R}$,

$$|\widehat{\varphi}(\xi)|^2 = |\widehat{\psi}(2\xi)|^2 + |\widehat{\varphi}(2\xi)|^2.$$

We also fix an integer $J \in \mathbf{N}$. Just as in the Introduction, we introduce the (dyadic) wavelet transform $\mathbf{W}$ (of length J) by defining

$$\mathbf{W}f := (f * \psi_2, f * \psi_{2^2}, \ldots, f * \psi_{2^J}, f * \varphi_{2^J}) \quad \text{for all } f \in P_\Omega.$$

We remark that more information on wavelet transforms may be found, for example, in Chui [Chu]. We also note that in electrical engineering terminology, $\mathbf{W}f$ is a decomposition of f with respect to an "octave filter bank."

From §2, for all $f \in P_\Omega$, $f * \varphi \in L^2 \cap C_0$, and $|(f * \varphi)^\wedge(\xi)| = |\widehat{f}(\xi)\widehat{\varphi}(\xi)|$ for all $\xi \in \mathbf{R}$. It follows, via Plancherel, that

$$C\|f\|_2 \le \|\mathbf{W}f\|_{\ell^2(L^2)} = \|f * \varphi\|_2 \le \|f\|_2.$$

Therefore, $\mathbf{W} : P_\Omega \to \ell^2(L^2) = \ell^2_{J+1}(L^2)$ is an isomorphism. Define $\psi^{0,j}$ for $j = 1, \ldots, J$ and $\varphi^{0,J}$ by

$$\widehat{\psi^{0,j}}(\xi) := \frac{\widehat{\psi}(2^j\xi)}{|\widehat{\varphi}(\xi)|^2} \quad \text{if } \xi \in [-\Omega, \Omega], \quad \text{with } \widehat{\psi^{0,j}}(\xi) := 0 \text{ otherwise,}$$

$$\widehat{\varphi^{0,J}}(\xi) := \frac{\widehat{\varphi}(2^J\xi)}{|\widehat{\varphi}(\xi)|^2} \quad \text{if } \xi \in [-\Omega, \Omega], \quad \text{with } \widehat{\varphi^{0,J}}(\xi) := 0 \text{ otherwise.}$$

Next, for all $g = (g_1, \ldots, g_{J+1}) \in \ell^2(L^2)$, define

$$\mathbf{M}g := \sum_{j=1}^{J} g_j * \widetilde{\psi^{0,j}} + g_{J+1} * \widetilde{\varphi^{0,J}}.$$

Fix $f \in P_\Omega$. Then $\mathbf{MW}f = f * h$, where $h := \sum_{j=1}^{J} \widetilde{\psi^{0,j}} * \psi_{2^j} + \widetilde{\varphi^{0,J}} * \varphi_{2^J}$. Now,

$$\widehat{h}(\xi) = \frac{1}{|\widehat{\varphi}(\xi)|^2}\left(\sum_{j=1}^{J} |\widehat{\psi}(2^j\xi)|^2 + |\widehat{\varphi}(2^J\xi)|^2\right) = 1 \quad \text{for all } \xi \in [-\Omega, \Omega],$$

and $\widehat{h}(\xi) = 0$, otherwise. Therefore, $\mathbf{MW}f = f$.

In the rest of the section, we fix $\Omega_1 > \Omega$ and suppose that $\operatorname{supp}\widehat{\varphi} \subseteq [-\Omega_1, \Omega_1]$. Note that $\widehat{\varphi}$ and $\widehat{\psi}$ both belong to L^2. We know that $\widehat{\psi}(\xi) = G(\xi/2)\,\widehat{\varphi}(\xi/2)$ for almost all $\xi \in \mathbf{R}$ and $\widehat{\varphi}(\xi/2) = 0$ for almost all $|\xi| > 2\Omega_1$. Therefore, $\operatorname{supp}\widehat{\psi} \subseteq [-2\Omega_1, 2\Omega_1]$. Again fix $f \in P_\Omega$. Then $(f * \psi_{2^1})^\wedge(\xi) = \widehat{f}(\xi)\widehat{\psi}(2\xi) = 0$, for almost all $|\xi| > \Omega_1 \wedge \Omega$. Similarly, for all $j = 1, \ldots, J$,

$$(f * \psi_{2^j})^\wedge(\xi) = \widehat{f}(\xi)\widehat{\psi}(2^j\xi) = 0 \quad \text{for almost all } |\xi| > \frac{\Omega_1}{2^{j-1}} \wedge \Omega,$$

while

$$(f * \varphi_{2^J})^\wedge(\xi) = \widehat{f}(\xi)\widehat{\varphi}(2^J\xi) = 0 \quad \text{for all } |\xi| > \frac{\Omega_1}{2^J} \wedge \Omega.$$

From above, if we define $\Sigma_j := (\Omega_1/2^{j-1}) \wedge \Omega$, for all $j \in \{1, \ldots, J+1\}$, then

$$\mathbf{W} : P_\Omega \to \left[P_{\Sigma_1} \oplus P_{\Sigma_2} \oplus \cdots \oplus P_{\Sigma_J} \oplus P_{\Sigma_{J+1}}\right]_2 =: Y \subseteq X := \ell^2(P_\Omega),$$

and $\mathbf{W}$ is a Banach space isomorphism of P_Ω onto $\mathbf{V} := \mathbf{W}(P_\Omega)$. Also, $\mathbf{M} : \mathbf{V} \to P_\Omega$ and $\mathbf{MW}f = f$ for all $f \in P_\Omega$; that is, $\mathbf{M}$ is a left inverse of $\mathbf{W}$. Further, $\mathbf{M}$ is bounded. Indeed, fix $g \in \ell^2(P_\Omega)$. Then

$$\begin{aligned}
\|\mathbf{M}g\|_{L^2}^2 &= \frac{1}{2\pi}\|\widehat{\mathbf{M}g}\|_{L^2}^2 \\
&= \frac{1}{2\pi}\int_{\xi=-\Omega}^{\Omega} \left|\sum_{j=1}^{J} \widehat{g}_j(\xi)\frac{\overline{\widehat{\psi}(2^j\xi)}}{|\widehat{\varphi}(\xi)|^2} + \widehat{g}_{J+1}(\xi)\frac{\overline{\widehat{\varphi}(2^J\xi)}}{|\widehat{\varphi}(\xi)|^2}\right|^2 d\xi \\
&\le \frac{1}{2\pi}\int_{\xi=-\Omega}^{\Omega} \sum_{j=1}^{J+1} |\widehat{g}_j(\xi)|^2 \frac{(\sum_{j=1}^{J} |\widehat{\psi}(2^j\xi)|^2 + |\widehat{\varphi}(2^J\xi)|^2)}{|\widehat{\varphi}(\xi)|^4}\, d\xi \\
&= \frac{1}{2\pi}\int_{\xi=-\Omega}^{\Omega} \sum_{j=1}^{J+1} |\widehat{g}_j(\xi)|^2 |\widehat{\varphi}(\xi)|^{-2}\, d\xi \le \frac{1}{2\pi}\frac{1}{C^2}\int_{\xi=-\Omega}^{\Omega} \sum_{j=1}^{J+1} |\widehat{g}_j(\xi)|^2\, d\xi \\
&= \frac{1}{C^2}\frac{1}{2\pi}\sum_{j=1}^{J+1} \|\widehat{g}_j\|_{L^2}^2 = \frac{1}{C^2}\sum_{j=1}^{J+1} \|g_j\|_{L^2}^2 \le \frac{1}{C^2}\|g\|_{\ell^2(L^2)}^2.
\end{aligned}$$

Since $\mathbf{W}$ takes P_Ω into $\ell^2(P_\Omega)$, while $\mathbf{M}$ maps $\ell^2(P_\Omega)$ into P_Ω, we see that $\mathbf{WM}$ maps $\ell^2(P_\Omega)$ into $\ell^2(P_\Omega)$. We now check that $\mathbf{P} := \mathbf{WM}$ is the orthogonal projection (in the Hilbert space $X = \ell^2(P_\Omega)$) of $\ell^2(P_\Omega)$ onto $\mathbf{V}$; that is, $\mathbf{P} = \mathbf{P_V}$. We denote the Hilbert space adjoint of a bounded linear operator L by $L^\star$. It is easy to see that Range $\mathbf{W}$ = Range $\mathbf{P}$ and that $\mathbf{P}^2 = \mathbf{P}$. Now define $A : P_\Omega \to P_\Omega$ by $Af := f * \varphi * \widetilde{\varphi}$. Then A is an invertible linear operator on P_Ω, A is self-adjoint, and both A and A^{-1} are bounded. Moreover, for all $f \in P_\Omega$ and $g \in \ell^2(P_\Omega)$, $\langle \mathbf{W}f, g\rangle_{\ell^2(P_\Omega)} = \langle f, A\mathbf{M}g\rangle$. Therefore, $\mathbf{W}^\star = A\mathbf{M}$ and $\mathbf{M}^\star = \mathbf{W}A^{-1}$. Hence $\mathbf{P}^\star = (\mathbf{WM})^\star = \mathbf{M}^\star\mathbf{W}^\star = \mathbf{WM} = \mathbf{P}$.

Now let $\Sigma_j := (\Omega_1/2^{j-1}) \wedge \Omega$ for all $j = 1, \ldots, J+1$. Next fix a function $f \in P_\Omega$. Let $t^1 = (t_n^1)_{n\in\mathbf{Z}}$ be a sequence in $\mathbf{R}$ that is strictly increasing, $\lim_{n\to\infty} t_n^1 = +\infty$, $\lim_{n\to-\infty} t_n^1 = -\infty$,

$$0 < m_1 := \inf_{n\in\mathbf{Z}} \left(t_{n+1}^1 - t_n^1\right) \quad \text{and} \quad \sup_{n\in\mathbf{Z}} \left(t_{n+1}^1 - t_n^1\right) =: M_1 < \frac{\pi}{\Sigma_1}.$$

Choose $\Sigma_1' \in (\Sigma_1, \pi/M_1)$. Then, by Benedetto [Ben, Theorem 51], which depends on Gröchenig [G2, Theorem 1], $(e_{-t_n^1})_{n\in\mathbf{Z}}$ is a Fourier frame for $L^2[-\Sigma_1', \Sigma_1']$; therefore, by a result of Benedetto and Heller [BH] (see also [Ben, Theorem 46]), P_{Σ_1} is determined by $(t_n^1)_{n\in\mathbf{Z}}$ and the frame operator S for $(e_{-t_n^1})_{n\in\mathbf{Z}}$. However, $f * \psi_2 \in P_{\Sigma_1} = P_\Omega$. Therefore $q_1 = f * \psi_2$ is determined uniquely among the q in P_{Σ_1} by $(t_n^1)_{n\in\mathbf{Z}}$ and also, via [BH] (see also [Ben, Algorithm 50]), by the sequence of sampled values $(q(t_n^1))_{n\in\mathbf{Z}}$.

Moreover, the work of Heller [He] (see also [Ben, Lemma 47]), tells us that even if $t^1 = (t_n^1)$ is so oversampled that $m_1 = \inf_{n\in\mathbf{Z}}(t_{n+1}^1 - t_n^1) = 0$, we may pass to a (constructable) subsequence of $(t_n^1)_n$, also denoted $(t_n^1)_n$, so that $(e_{-t_n^1})_n$ is a Fourier frame. Therefore, $f * \psi_2$ is still uniquely determined in P_{Σ_1} by $(t_n^1)_{n\in\mathbf{Z}}$ and $((f * \psi_2)(t_n^1))_{n\in\mathbf{Z}}$.

Fix $j \in \{1, \ldots, J\}$. Let $t^j = (t_n^j)_{n\in\mathbf{Z}}$ be a strictly increasing sequence in $\mathbf{R}$, with $\lim_{n\to\pm\infty} t_n^j = \pm\infty$ and

$$\sup_{n\in\mathbf{Z}}(t_{n+1}^j - t_n^j) =: M_j < \frac{\pi}{\Sigma_j}.$$

Choose $\Sigma_j' \in (\Sigma_j, \pi/M_j)$. Then by an argument similar to that for t^1, $f * \psi_{2^j}$ is uniquely determined in P_{Σ_j} by $(t_n^j)_{n\in\mathbf{Z}}$ and $((f * \psi_{2^j})(t_n^j))_{n\in\mathbf{Z}}$. Next, let $t^{J+1} = (t_n^{J+1})_{n\in\mathbf{Z}}$ be another strictly increasing real-valued sequence, such that $\lim_{n\to\pm\infty} t_n^{J+1} = \pm\infty$ and

$$\sup_{n\in\mathbf{Z}}(t_{n+1}^{J+1} - t_n^{J+1}) =: M_{J+1} < \frac{\pi}{\Sigma_{J+1}}.$$

Then we also have that $f * \varphi_{2^J}$ is uniquely determined in $P_{\Sigma_{J+1}}$ by $(t_n^{J+1})_{n\in\mathbf{Z}}$ and $((f * \varphi_{2^J})(t_n^{J+1}))_{n\in\mathbf{Z}}$.

For our fixed function f in P_Ω, we now introduce the set $\Gamma := \Gamma_f \subseteq \ell^2(P_\Omega)$ given by

$$\begin{aligned}\Gamma := \big\{g = (g_j)_{j=1}^{J+1} \in \ell^2(P_\Omega)\text{: } &\text{for all } j \in \{1, \ldots, J\} \\ &g_j(t_n^j) = (f * \psi_{2^j})(t_n^j) =: C_n^j, \ \text{for all } n \in \mathbf{Z} \\ &\text{and } \ g_{J+1}(t_n^{J+1}) = (f * \varphi_{2^J})(t_n^{J+1}) =: C_n^{J+1}, \ \text{for all } n \in \mathbf{Z}\big\}.\end{aligned}$$

It is easy to see that Γ is an affine closed subspace of $\ell^2(P_\Omega)$, which from the definition above satisfies

$$\Gamma \cap \mathbf{V} = \{\mathbf{W}f\}.$$

Using the notation for projections from §2, let us now define $\mathbf{T} : \ell^2(P_\Omega) \to \ell^2(P_\Omega)$ by $\mathbf{T} := \mathbf{P_V}\mathbf{P_\Gamma}$. Then $\mathbf{T}$ is nonexpansive, and by a result of Youla and Webb [YW] (using the work of Opial [O]), for all $g \in \ell^2(P_\Omega)$, the sequence $(\mathbf{T}^n g)_{n\in\mathbf{N}}$ converges weakly in $\ell^2(P_\Omega)$ to a member of $\Gamma \cap \mathbf{V}$ that is, to $\mathbf{W}f$. However, Γ is affine. By a result of von Neumann [N] (or see, e.g., [YW], [Hu]), $(\mathbf{T}^n g)_{n\in\mathbf{N}}$ converges in norm to $\mathbf{W}f$ (Fig. 1).

We summarize the discussion of this section as a theorem.

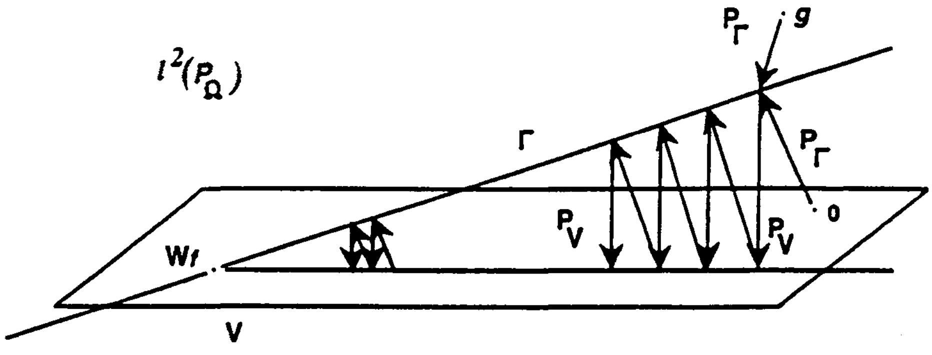

FIGURE 1.

3.1. *Theorem*

Let $\Omega > 0$ and $f \in P_\Omega$. Let $J, \Omega_1, \mathbf{W}$ and $\mathbf{V}$ be as described previously. For each $j \in \{1, \dots, J+1\}$ fix a strictly increasing sequence t^j of real numbers, accumulating at ∞ and $-\infty$ and with $M_j := \sup_{n\in\mathbb{Z}}(t^j_{n+1} - t^j_n) < \pi/\Sigma_j$, where each $\Sigma_j = (\Omega_1/2^{j-1}) \wedge \Omega$. Let the affine set $\Gamma \subseteq \ell^2(P_\Omega)$ be defined in terms of each t^j, as above. Also let $\mathbf{T} := \mathbf{P}_\mathbf{V}\mathbf{P}_\Gamma$. Then, for all $g \in \ell^2(P_\Omega)$. $(\mathbf{T}^n g)_{n\in\mathbb{N}}$ converges in norm to $\mathbf{W}f$.

3.2. Remark. In the limiting case of the main wavelets of this paper, to be introduced in §5 (see Figs. 2 and 3), when $\Omega_1 = \Omega$, (φ, ψ) is the Shannon or sinc wavelet pair, corresponding to the ideal low and high pass filters on the frequency range $[-\Omega, \Omega]$. In this case the formulas for $\mathbf{M}$ and $\mathbf{W}$ in §3 show that $\mathbf{P}_\mathbf{V} g = \mathbf{W}\mathbf{M}g$, for $g \in \ell^2(P_\Omega)$, projects each coordinate g_j of g into the Paley-Wiener subspace P_{Σ_j}. Consequently, the Mallat-Zhong reconstruction algorithm proceeds in parallel on each level. However, when $\Omega_1 > \Omega$, the algorithm introduces partial mixing of information from different (adjacent) levels on each application of $\mathbf{P}_\mathbf{V}$, and therefore each iteration of $\mathbf{T} = \mathbf{P}_\mathbf{V}\mathbf{P}_\Gamma$. Thus each level of $\mathbf{W}f$ is not recovered independently of the others. Moreover, the algorithm is not a frame method, even in the sinc case. The frame method iterates an operator of the form $I - \rho S$, where ρ is a positive constant defined in terms of the frame bounds and S is the frame operator, which is positive and self-adjoint. Consequently, $I - \rho S$ is a self-adjoint, linear operator. On the other hand, the Mallat-Zhong operator $\mathbf{T} = \mathbf{P}_\mathbf{V}\mathbf{P}_\Gamma$, is the product of a linear and an affine projection, which is not generally linear. Even if $0 \in \Gamma$, so that $\mathbf{T}$ is linear, $\mathbf{T}$ can only be self-adjoint if $\mathbf{P}_\mathbf{V}$ and $\mathbf{P}_\Gamma$ commute. This means that $\mathbf{T} = \mathbf{P}_{\mathbf{V}\cap\Gamma}$, which is generally not true. Hence, we can see that the Mallat-Zhong algorithm is not a frame algorithm.

4. Geometric Convergence with a Computable Convergence Rate

We begin by recalling some general facts about Hilbert space geometry. Let $(X, \langle\cdot,\cdot\rangle)$ be a complex Hilbert space with corresponding norm $\|\cdot\|$. Suppose that V and S are closed subspaces of X such that

$$K_{S,V} := \sup_{s\in S, v\in V} \frac{|\langle s, v\rangle|}{\|s\|\,\|v\|} < 1.$$

It is easy to verify that $\|\mathbf{P}_V s\| \le K_{S,V}\|s\|$ and $\|\mathbf{P}_S v\| \le K_{S,V}\|v\|$, for all $s \in S$ and $v \in V$. Thus for $U := \mathbf{P}_V\mathbf{P}_S$, we have that $\|Uv\| \le K^2_{S,V}\,\|v\|$, for all $v \in V$.

Next, suppose Γ is a closed affine subspace of X; that is, Γ is a translation of a closed subspace S of X. Suppose that there exists $q \in \Gamma \cap V$. Then $S = \Gamma - q$ and $\mathbf{P}_\Gamma h = \mathbf{P}_S h + q - \mathbf{P}_S q$, for all $h \in X$.

Let $\mathbf{T} := \mathbf{P}_V \mathbf{P}_\Gamma$. Again assume that $K_{S,V} < 1$. It follows readily that $\|\mathbf{T}^n v - q\| \leq (K_{S,V}^2)^n \|v - q\|$, for all $v \in V$.

The following result is well-known and can be proven in a straightforward manner. Mallat and Zhong [MZ2] use a variation of this result to show that under certain circumstances their algorithm converges geometrically. Our work follows their lead. When combined with the band-limited notions of [G2] and [Ben], we will get that our algorithm always converges geometrically, with a computable estimate for the rate of convergence in terms of the irregular sampling rates at each level of the wavelet transform.

4.1. *Lemma*

Let $(X, \langle \cdot, \cdot \rangle)$ be a complex Hilbert space with corresponding norm $\|\cdot\|$. Suppose that V is a closed subspace of X and Γ is a translation of a closed subspace S of X, with $q \in \Gamma \cap V$. Then the following are equivalent:

1. *There exists $C \in (0, 1]$ such that $\|h - g\| \geq C\, \|q - g\|$, for all $h \in \Gamma$ and $g \in V$.*
2. *$K_{S,V} < 1$.*

Moreover, if C is given as above, then $K_{S,V} \leq \sqrt{1 - C^2}$.

We refer the reader to Bauschke and Borwein [BB] for recent work on von Neumann's alternating projection algorithm. For the remainder of this section we use the setting and notation of §3. Our next result is an analogue of [MZ2, Appendix F].

4.2. *Theorem*

Let $\Omega > 0$ and $f \in P_\Omega$. Let $J, \Omega_1, \mathbf{W}, \mathbf{V}$, and Σ_j be as in Theorem 3.1. Let $q := \mathbf{W}f$. For each $j \in \{1, \ldots, J+1\}$, fix a strictly increasing sequence t^j of real numbers, accumulating at ∞ and $-\infty$, with $M_j := \sup_{n \in \mathbf{Z}} (t_{n+1}^j - t_n^j) < \pi / \Sigma_j$ and $m_j := \inf_{n \in \mathbf{Z}} (t_{n+1}^j - t_n^j) > 0$. Let the affine set $\Gamma \subseteq \ell^2(P_\Omega)$ be defined in terms of each t^j as above. Then there exist positive constants C_1 and C_2 such that for all $h \in \Gamma$ and $g \in V$ we have that

$$\|h - g\|^2 \geq C_1 \sum_{n \in \mathbf{Z}} \sum_{j=1}^{J+1} |C_n^j - g_j(t_n^j)|^2, \tag{1}$$

while

$$\|q - g\|^2 \leq C_2 \sum_{n \in \mathbf{Z}} \sum_{j=1}^{J+1} |C_n^j - g_j(t_n^j)|^2. \tag{2}$$

Proof. Fix $j \in \{1, \ldots, J+1\}$. By a result of Plancherel and Pólya [PP] (see also, e.g., [Y, §2.2.3], [Ben, Lemma 42(b)]), since $h_j - g_j \in P_\Omega$, it follows that

$$\|h_j - g_j\|_2^2 \geq \zeta_j \sum_{n \in \mathbf{Z}} |(h_j - g_j)(t_n^j)|^2,$$

where

$$\zeta_j := \frac{\pi \Omega m_j^2}{8\left(e^{(1/2)\,\Omega\, m_j} - 1\right)}.$$

Therefore,

$$\|h_j - g_j\|_2^2 \geq \zeta_j \sum_{n \in \mathbf{Z}} |C_n^j - g_j(t_n^j)|^2.$$

On the other hand, $q_j - g_j \in P_{\Sigma_j}$. By Benedetto [Ben, Theorem 51], which uses Gröchenig [G2, Theorem 1],

$$\|q_j - g_j\|_2^2 \leq \eta_j \sum_{n \in \mathbf{Z}} |(q_j - g_j)(t_n^j)|^2,$$

where

$$\eta_j := \frac{M_j}{\left(1 - \frac{M_j\,\Sigma_j}{\pi}\right)^2}.$$

Consequently,

$$\|q_j - g_j\|_2^2 \le \eta_j \sum_{n\in\mathbf{Z}} |C_n^j - g_j(t_n^j)|^2.$$

Let $C_1 := \min_{1\le j\le J+1} \zeta_j$ and $C_2 := \max_{1\le j\le J+1} \eta_j$. Then $C_1, C_2 \in (0, \infty)$; and statements (1) and (2) of the theorem now clearly follow. □

The proof of Theorem 4.2, together with Lemma 4.1 and the discussion preceding it yield the following corollary.

4.3. *Corollary*

Under the hypotheses of Theorem 4.2, *with ζ_j and η_j defined as in the proof of Theorem* 4.2, *let $\beta := \min_{1\le j\le J+1}(\zeta_j/\eta_j)$. Then for all $h \in \Gamma$ and $g \in \mathbf{V}$, $\|h - g\|_{\ell^2(P_\Omega)} \ge \sqrt{\beta}\,\|q - g\|_{\ell^2(P_\Omega)}$. Moreover, for $\mathbf{T} := \mathbf{P_V P_\Gamma}$, for all $v \in \mathbf{V}$ and $n \in \mathbf{N}$, we have that*

$$\|\mathbf{T}^n v - q\|_{\ell^2(P_\Omega)} \le (1-\beta)^n \|v - q\|_{\ell^2(P_\Omega)}.$$

We remark that there is only a small loss of generality in starting the alternating projections algorithm in $\mathbf{V}$, rather than in $\ell^2(P_\Omega)$. Indeed, if we start at $g \in \ell^2(P_\Omega)$, then $v = \mathbf{T}g$ belongs to $\mathbf{V}$. Further, since $0 \in \mathbf{V}$, this natural and simple choice is often made. Second, note that the estimates in the proof of Theorem 4.2 resemble those gained by Mallat and Zhong [MZ2, Appendix F], using a Sobolev norm and a calculus of variations argument.

5. A Useful Sharpening of the Reconstruction Theorem

Theorem 3.1 is not sharp enough to enable implementation of a discrete version of the Mallat-Zhong algorithm. We specialize the setting of §3 in this section, in order to do this. First, we remark that a function $u \in L^2$ is real-valued (i.e., is a signal) if and only if $\overline{\widehat{u}(-\xi)} = \widehat{u}(\xi)$ for almost all $\xi \in \mathbf{R}$.

Fix $\Delta > 0$. Note that if $u \in L^2$, then $u \in P_\Delta$ implies that $f = \operatorname{Re} u \in P_\Delta$. On the other hand, if $f \in P_\Delta$ is real-valued, then $\mathcal{H}f \in P_\Delta$ and is real-valued, so that $u := f + i\mathcal{H}f \in P_\Delta$ and $f = \operatorname{Re} u$. Moreover,

$$\widehat{u}(\xi) = 2\widehat{f}(\xi)\,\chi_{[0,\infty)}(\xi) \qquad \text{for almost all } \xi \in \mathbf{R}.$$

Next we fix $\Omega > 0$ and $\Omega_1 > \Omega$ with $\Omega_1 \le 2\Omega$ and $\nu \in \mathbf{N} \cup \{0\}$. We then choose functions c and s mapping $[\Omega, \Omega_1] \to \mathbf{R}$ such that both are ν times continuously differentiable on $[\Omega, \Omega_1]$ (i.e., $c, s \in C^{(\nu)}[\Omega, \Omega_1]$), with all one-sided derivatives at Ω and Ω_1 equal to 0, such that c strictly decreases from 1 at Ω to 0 at Ω_1, whereas s strictly increases from 0 at Ω to 1 at Ω_1, and such that $c(\xi)^2 + s(\xi)^2 = 1$ for all $\xi \in [\Omega, \Omega_1]$.

We continue by defining two functions w and v (using the fact that $\Omega_1 \le 2\Omega$):

$$\begin{aligned} w(\xi) &:= 1 \quad \text{if } |\xi| \in [0, \Omega],\\ w(\xi) &:= c(|\xi|) \quad \text{if } |\xi| \in (\Omega, \Omega_1], \quad \text{and}\\ w(\xi) &:= 0 \quad \text{for all other } \xi \in \mathbf{R}.\end{aligned}$$

$$\begin{aligned} v(\xi) &:= s(|\xi|) \quad \text{if } |\xi| \in [\Omega, \Omega_1],\\ v(\xi) &:= 1 \quad \text{if } |\xi| \in (\Omega_1, 2\Omega],\\ v(\xi) &:= c(|\xi|/2) \quad \text{if } |\xi| \in (2\Omega, 2\Omega_1], \quad \text{and}\\ v(\xi) &:= 0 \quad \text{for all other } \xi \in \mathbf{R}.\end{aligned}$$

We now define $\varphi := w^\vee$ and $\psi := v^\vee$ (Figs. 2 and 3). Since w and v are even, real-valued functions, φ and ψ are also real-valued. Moreover, by our choice of s and c, w and v are $C^{(\nu)}$ functions on $\mathbf{R}$. In particular, they each have ν distributional derivatives in $L^2 \cap L^1$ and therefore $\mathbf{x}^\nu\varphi$ and $\mathbf{x}^\nu\psi$ belong to $L^2 \cap C_0$. We will return to the rapid decay of φ and ψ in §6. We remark that we could also arrange for $\nu = \infty$. The construction above closely resembles that of Meyer [Me1], who produced the first band-limited, orthonormal wavelet basis of L^2, with decay at ∞ and $-\infty$ faster than the reciprocal of any polynomial.

It is easy to check that for almost all $\xi \in \mathbf{R}$,

$$|\widehat{\varphi}(\xi)|^2 = |\widehat{\psi}(2\xi)|^2 + |\widehat{\varphi}(2\xi)|^2.$$

We note that both $\widehat{\varphi}$ and $\widehat{\psi}$ vanish outside of a bounded interval, while $\widehat{\varphi}$ is continuous at 0 with $\widehat{\varphi}(0) = 1$. Consequently, ψ satisfies the identity $\sum_{j\in\mathbf{Z}} |\widehat{\psi}(2^j\xi)|^2 = 1$ for almost all $\xi \in \mathbf{R}$, which is the condition assumed by Mallat and Zhong in [MZ1]. Moreover, $\widehat{\varphi}$ is continuous on $\mathbf{R}$ with $\widehat{\varphi}(0) = 1$, and for all $\Omega_0 \in [\Omega, \Omega_1)$ there exists $0 < A_0 \le B_0 < \infty$ with $A_0 \le \sum_{k\in\mathbf{Z}} |\widehat{\varphi}(\xi + 2k\Omega_0)|^2 \le B_0$ for almost all $\xi \in \mathbf{R}$. By, for example, Daubechies [D, §5.3.2], it follows that the translates of φ by all integer multiples of $T_0 := \pi/\Omega_0$ are a Riesz basis for a closed subspace V_0, which generates a multiresolution analysis of $L^2(\mathbf{R})$. However, due to the fact that $\sum_{k\in\mathbf{Z}} |\widehat{\varphi}(\xi + 2k\Omega_0)|^2$ is not usually

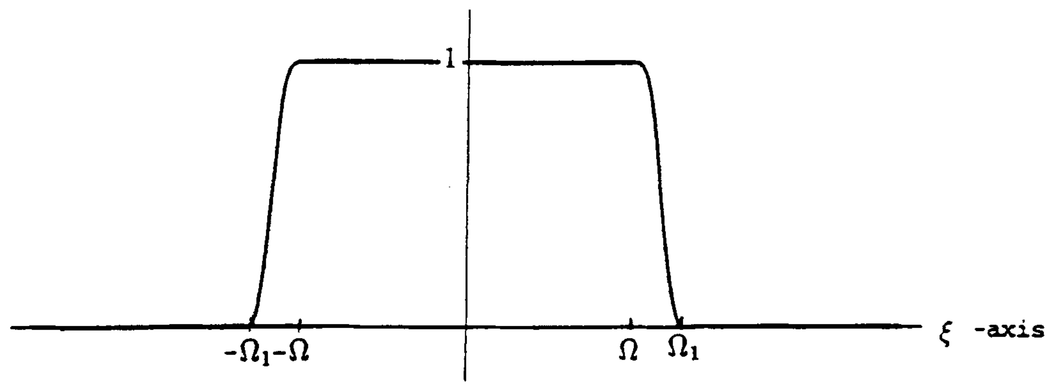

FIGURE 2.

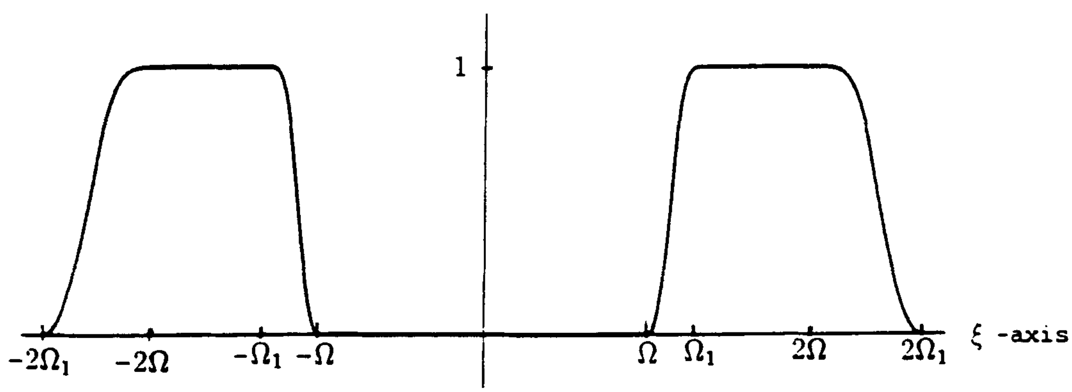

FIGURE 3.

identically equal to 1, (φ, ψ) may not be an orthonormal wavelet pair based on the integer grid $(kT_0)_{k\in\mathbf{Z}}$. On the other hand, it is easy to choose the functions c and s satisfying $c(\xi) = s(\Omega_1+\Omega-\xi)$ for all $\xi \in [\Omega, \Omega_1]$, which gives us $A_0 = B_0 = 1$ for the value $\Omega_0 = (\Omega + \Omega_1)/2$.

Now, let us fix a *signal* $f \in P_\Omega$. We form $u := f + i\mathcal{H}f \in P_\Omega$. We know that $\operatorname{supp} \widehat{\psi} \subseteq [-2\Omega_1, -\Omega] \cup [\Omega, 2\Omega_1]$. We define

$$B_1 := \left[-\Omega, -\frac{\Omega}{2}\right] \cup \left[\frac{\Omega}{2}, \Omega\right],$$

$$B_j := \left[-\frac{\Omega_1}{2^{j-1}}, -\frac{\Omega}{2^j}\right] \cup \left[\frac{\Omega}{2^j}, \frac{\Omega_1}{2^{j-1}}\right] \quad \text{for all } j \in \{2, \dots, J\},$$

$$B_{J+1} := \left[-\frac{\Omega_1}{2^J}, \frac{\Omega_1}{2^J}\right].$$

Thus, $f * \psi_{2^j} \in P^{\mathbf{R}}(B_j)$ for all $j \in \{1, \dots, J\}$, and $f * \varphi_{2^J} \in P^{\mathbf{R}}(B_{J+1})$. Define

$$\Sigma_1 := \frac{\Omega}{4},$$

$$\Sigma_j := \frac{1}{2}\left(\frac{\Omega_1}{2^{j-1}} - \frac{\Omega}{2^j}\right) \quad \text{for all } j \in \{2, \dots, J\},$$

$$\Sigma_{J+1} := \frac{1}{2}\,\frac{\Omega_1}{2^J}.$$

As noted above, each $P^{\mathbf{R}}(B_j)$ is real Banach space isomorphic to P_{Σ_j} via the isomorphism Φ_j given by

$$(\Phi_j k)(t) := e^{-it\mu_j}\,(k(t) + i(\mathcal{H}k)(t)) \quad \text{for all } t \in \mathbf{R} \text{ and all } k \in P^{\mathbf{R}}(B_j).$$

Here,

$$\mu_1 := \frac{1}{2}\left(\frac{\Omega}{2} + \Omega\right),$$

$$\mu_j := \frac{1}{2}\left(\frac{\Omega}{2^j} + \frac{\Omega_1}{2^{j-1}}\right) \quad \text{for } j \in \{2, \dots, J\},\ \mu_{J+1} := \frac{\Omega_1}{2^{J+1}}.$$

In using the Hilbert transform in this context we follow many authors, notably Logan [L]. Fix $j \in \{1, \dots, J+1\}$. Let $t^j = (t_n^j)_{n\in\mathbf{Z}}$ be a strictly increasing sequence in $\mathbf{R}$, with $\lim_{n\to\pm\infty} t_n^j = \pm\infty$ and

$$\sup_{n\in\mathbf{Z}}(t_{n+1}^j - t_n^j) =: M_j < \frac{\pi}{\Sigma_j}.$$

We now consider the modified Paley–Wiener space $P\mathcal{H}_\Omega$. $P\mathcal{H}_\Omega$ and $\ell^2(P\mathcal{H}_\Omega)$ are closed (complex) subspaces of P_Ω and $\ell^2(P_\Omega)$, respectively. Let $\mathbf{Q}$ denote the orthogonal projection onto $P\mathcal{H}_\Omega$ in P_Ω and let $\mathbf{P}$ be the orthogonal projection onto $\ell^2(P\mathcal{H}_\Omega)$ in $\ell^2(P_\Omega)$. Note that $\mathbf{QM} = \mathbf{MP}$. For the φ and ψ of this section, the wavelet transform $\mathbf{W}$, restricted to $P\mathcal{H}_\Omega$, maps $P\mathcal{H}_\Omega$ into $\ell^2(P\mathcal{H}_\Omega)$, and is an isometric Hilbert isomorphism. This is because φ and ψ are real-valued and the Hilbert transform commutes with convolution. Similarly, the inverse wavelet transform $\mathbf{M}$ maps $\ell^2(P\mathcal{H}_\Omega)$ into $P\mathcal{H}_\Omega$. We now redefine $\mathbf{V}$ as

$$\mathbf{V} := \mathbf{W}(P_\Omega) \cap \ell^2(P\mathcal{H}_\Omega) = \mathbf{W}(P\mathcal{H}_\Omega).$$

The calculations of §3 now show us that **WQM** is the orthogonal projection onto **V** in the Hilbert space $\ell^2(P_\Omega)$. Note that u, as defined above, is in $P\mathcal{H}_\Omega$. Now define $\Gamma := \Gamma_u \subseteq \ell^2(P_\Omega)$ given by

$$\Gamma := \Big\{g = (g_j)_{j=1}^{J+1} \in \ell^2(P_\Omega)\colon \text{ for all } j \in \{1, \dots, J\}$$
$$g_j(t_n^j) = (u * \psi_{2^j})(t_n^j) =: C_n^j, \text{ for all } n \in \mathbf{Z}$$
$$\text{and } g_{J+1}(t_n^{J+1}) = (u * \varphi_{2^J})(t_n^{J+1}) =: C_n^{J+1}, \text{ for all } n \in \mathbf{Z}\Big\}.$$

5.1. *Theorem*

Let Ω, J, Ω_1, $\mathbf{W}$, $\mathbf{V}$ and Σ_j be as described previously, with the modifications of §5. Fix a signal $f \in P_\Omega$. Let $u := f + i\mathcal{H}f$. For each $j \in \{1, \dots, J+1\}$, fix a strictly increasing sequence t^j of real numbers, accumulating at ∞ and $-\infty$ and with $M_j := \sup_{n\in\mathbf{Z}}(t_{n+1}^j - t_n^j) < \pi/\Sigma_j$. Let the affine set $\Gamma = \Gamma_u \subseteq \ell^2(P_\Omega)$ be defined in terms of each t^j, as in this section. Also let $\mathbf{T} := \mathbf{P_V P_\Gamma}$. Then for all $g \in \ell^2(P_\Omega)$, $(\mathbf{T}^n g)_{n\in\mathbf{N}}$ converges in norm to $\mathbf{W}u$. Moreover, f may be recovered by the formula $f = \operatorname{Re} u$.

Proof. Uniqueness is the only issue here. We wish to show that $\Gamma \cap \mathbf{V} = \{\mathbf{W}u\}$. Now suppose that $h \in \Gamma \cap \mathbf{V}$. Then $h = \mathbf{W}u_1$, where $u_1 = f_1 + i\mathcal{H}f_1$, $f_1 \in P_\Omega$ and f_1 is real-valued, since $h \in \mathbf{V}$. Now fix $j \in \{1, \dots, J+1\}$. Since $h \in \Gamma$, $((\mathbf{W}u_1)_j(t_n^j))_{n\in\mathbf{Z}} = ((\mathbf{W}u)_j(t_n^j))_{n\in\mathbf{Z}}$. Consequently, $((\mathbf{W}f_1)_j(t_n^j))_{n\in\mathbf{Z}} = ((\mathbf{W}f)_j(t_n^j))_{n\in\mathbf{Z}}$ and $((\mathbf{W}\mathcal{H}f_1)_j(t_n^j))_{n\in\mathbf{Z}} = ((\mathbf{W}\mathcal{H}f)_j\ (t_n^j))_{n\in\mathbf{Z}}$. Hence, from above, $((\Phi_j(\mathbf{W}f_1)_j)(t_n^j))_{n\in\mathbf{Z}} = ((\Phi_j(\mathbf{W}f)_j)(t_n^j))_{n\in\mathbf{Z}}$.

By a similar argument to that in §3, $\Phi_j(\mathbf{W}f)_j$ is uniquely determined in P_{Σ_j} by $(t_n^j)_{n\in\mathbf{Z}}$ and $((\Phi_j(\mathbf{W}f)_j)(t_n^j))_{n\in\mathbf{Z}}$. Thus $\Phi_j(\mathbf{W}f_1)_j = \Phi_j(\mathbf{W}f)_j$. Since Φ_j is an isomorphism, we see that $(\mathbf{W}f_1)_j = (\mathbf{W}f)_j$. And since $j \in \{1, \dots, J+1\}$ is arbitrary, we see that $\mathbf{W}f_1 = \mathbf{W}f$. Therefore, $f_1 = f$. At last we arrive at the fact that $u_1 = u$, or equivalently, $h = \mathbf{W}u$. □

The usefulness of Theorem 5.1, as we shall soon see, derives from the fact that we have roughly quartered the sampling rate used in Theorem 3.1, when Ω_1 is close to Ω.

5.2. Remark. With the modifications to our setting of this section, if in Theorem 4.2 and Corollary 4.3 we replace the original Σ_j's by the new sharper Σ_j's of this section, while changing f everywhere to u, where $u := f + i\mathcal{H}f$ for some given real-valued $f \in P_\Omega$, then we get two new true statements. We will call them Theorems 5.3 and 5.4, respectively. Moreover, the constants ζ_j and η_j are given by the same formulas in terms of m_j and M_j, respectively.

6. The Algorithm for Discrete Signals

We continue with the notation and definitions of §5. We further restrict Ω_1 to be such that $\Omega_1 \leq 4\Omega/3$. Now we define a function H with some of the properties of the H mentioned in §3, and two functions G^R and G^L, which are modifications of the G discussed there. We introduce the notation $y^+ := (1/2)(y + i\mathcal{H}y)$ and $y^- := (1/2)(y - i\mathcal{H}y)$ for all real-valued L^2 functions y. Note that $\widehat{y^+} = \widehat{y} \cdot \chi_{[0,\infty)}$ and $\widehat{y^-} = \widehat{y} \cdot \chi_{(-\infty,0]}$, for all such y.

Recall the functions c, s, φ, and ψ introduced in §5. We suggest that the reader quickly sketch the functions H, G^R, and G^L that are introduced subsequently in terms of c and s, and that are closely related to φ and ψ. Also refer to Figs. 2 and 3 for illustrations of $\widehat{\varphi}$ and $\widehat{\psi}$.

Let $\varepsilon := \Omega - \Omega_1/2$, and note that $\varepsilon > 0$. For all $\xi \in [-\Omega, \Omega]$ define

$$H(\xi) := 1 \quad \text{if } |\xi| \in [0, \Omega/2],$$
$$H(\xi) := c(|\xi|) \quad \text{if } |\xi| \in (\Omega/2, \Omega_1/2], \quad \text{and}$$
$$H(\xi) := 0 \quad \text{for all other } \xi \in [-\Omega, \Omega].$$

Next, extend H by periodicity to a 2Ω-periodic function on $\mathbf{R}$. Note that $\Omega + \varepsilon \geq \Omega_1$, because $\Omega_1 \leq 4\Omega/3$. It follows that $\widehat{\varphi}(\xi) = H(\xi/2)\, \widehat{\varphi}(\xi/2)$ for almost all $\xi \in \mathbf{R}$.

The 2Ω-periodicity of H implies that $H(\xi) = \sum_{k\in\mathbf{Z}} c_k e^{-ikT\xi}$, $\xi \in \mathbf{R}$, with convergence in $L^2[-\Omega, \Omega]$; where $T := \pi/\Omega$ and $c_k = T(H \cdot \chi_{[-\Omega,\Omega]})^\vee(kT)$, $k \in \mathbf{Z}$, is the $\ell^2(\mathbf{Z}, \mathbf{C})$ sequence of Fourier coefficients of H. Applying the inverse Fourier transform to the previous equation for $\widehat{\varphi}(\xi)$, and using the fact that, by its definition, $\sum_{k\in\mathbf{Z}} |\widehat{\varphi}(\xi + 2k\Omega)|$ is uniformly bounded in ξ from above by 2, it follows that φ satisfies the dilation equation

$$\varphi(t) = \sum_{k\in\mathbf{Z}} 2c_k\varphi(2t - kT), \qquad t \in \mathbf{R}, \tag{$\spadesuit$}$$

with convergence in $L^2 \cap C_0$. Also, by our choice of s and c, $H \cdot \chi_{[-\Omega,\Omega]} \in C^{(\nu)}$. Consequently, $(c_k)_{k\in\mathbf{Z}}$ is $o(1/k^\nu)$ as $|k| \to \infty$.

In addition, $\mathcal{H}\varphi$, φ^+, and φ^- each satisfy a dilation equation with precisely the same coefficients c_k. We denote these equations by ($\spadesuit\mathcal{H}$), ($\spadesuit+$), and ($\spadesuit-$), respectively.

Let $\varepsilon_1 := (\Omega_1 - \Omega)/2$. For all $\xi \in [\Omega_1 + \varepsilon_1 - 2\Omega, \Omega_1 + \varepsilon_1]$ define

$$\begin{aligned}
&G^R(\xi) := s(2\xi) \quad \text{if } \xi \in [\Omega/2, \Omega_1/2],\\
&G^R(\xi) := 1 \quad \text{if } \xi \in (\Omega_1/2, \Omega_1],\\
&G^R(\xi) := c(2\xi - 2\Omega_1 + \Omega) \quad \text{if } \xi \in (\Omega_1, \Omega_1 + \varepsilon_1], \quad \text{and}\\
&G^R(\xi) := 0 \quad \text{for all other } \xi \in [\Omega_1 + \varepsilon_1 - 2\Omega, \Omega_1 + \varepsilon_1].
\end{aligned}$$

Next, extend G^R by periodicity to a 2Ω-periodic function on $\mathbf{R}$. Note that $\Omega_1 + \varepsilon_1 - 2\Omega \le 0$, because $\Omega_1 \le 4\Omega/3$. It follows that $\widehat{\psi^+}(\xi) = G^R(\xi/2)\,\widehat{\varphi^+}(\xi/2)$ for almost all $\xi \in \mathbf{R}$.

The 2Ω-periodicity of G^R implies that ψ^+ satisfies the following equation:

$$\psi^+(t) = \sum_{k\in\mathbf{Z}} 2d_k^+\varphi^+(2t - kT), \qquad t \in \mathbf{R}, \tag{$\heartsuit+$}$$

with convergence in $L^2 \cap C_0$. Here $d_k^+ = T(G^R \cdot \chi_{[\Omega_1+\varepsilon_1-2\Omega,\Omega_1+\varepsilon_1]})^\vee(kT)$ for each $k \in \mathbf{Z}$. We also have that $(d_k^+)_{k\in\mathbf{Z}}$ is $o(1/k^\nu)$ as $|k| \to \infty$.

Letting $G^L(\xi) := G^R(-\xi)$, for all $\xi \in \mathbf{R}$, and replacing φ^+, ψ^+, and G^R by φ^-, ψ^-, and G^L (respectively) in the immediately preceding discussion results in an equation for ψ^- with coefficients $d_k^- = d_{-k}^+$, which we label ($\heartsuit-$). Some simple algebra gives

$$\psi(t) = \sum_{k\in\mathbf{Z}}(d_k^+ + d_k^-)\varphi(2t - kT) + i\sum_{k\in\mathbf{Z}}(d_k^+ - d_k^-)(\mathcal{H}\varphi)(2t - kT), \tag{$\heartsuit$}$$

$$(\mathcal{H}\psi)(t) = -i\sum_{k\in\mathbf{Z}}(d_k^+ - d_k^-)\varphi(2t - kT) + \sum_{k\in\mathbf{Z}}(d_k^+ + d_k^-)(\mathcal{H}\varphi)(2t - kT), \tag{$\heartsuit\mathcal{H}$}$$

for all $t \in \mathbf{R}$, with convergence in $L^2 \cap C_0$.

Now, let us recall Shannon's sampling theorem, due to Whittaker [W] and Shannon [S] (see also, e.g., Young [Y], Marks [Ma]). Fix $u \in P_\Omega$, with $T = \pi/\Omega$. Then

$$u = \sum_{n\in\mathbf{Z}} u(nT)\operatorname{sinc}(\Omega(\cdot - nT)),$$

with convergence in $L^2 \cap C_0$. As usual, $\operatorname{sinc}(t) := \sin t/t$ if $t \neq 0$ and $\operatorname{sinc}(0) := 1$. Moreover, $\mathbf{S} : u \to (\sqrt{T}\, u(nT))_{n\in\mathbf{Z}}$ is a Hilbert space (isometric) isomorphism of P_Ω onto $\ell^2(\mathbf{Z}, \mathbf{C})$, and $(\gamma_n := (1/\sqrt{T})\operatorname{sinc}(\Omega(\cdot - nT)))_{n\in\mathbf{Z}}$ is an orthonormal basis for P_Ω.

Thus prepared, we turn to the discrete algorithm. Let us suppose that we have a one-dimensional signal, sampled on an evenly spaced time grid, $\sigma = (\sigma(nT))_{n\in\mathbf{Z}} \in \ell^2(\mathbf{Z}, \mathbf{C})$, where $T > 0$ and $\Omega = \pi/T$. We know that σ interpolates a band-limited signal $f \in P_\Omega$. Therefore, $f(nT) = (1/\sqrt{T})\sigma(nT)$, for all n, and $f = \mathbf{S}^{-1}\sigma$. If we only know the values $\sigma(nT)$ for n in a finite set $E = \{0, \dots, N\}$ of consecutive integers, then we could assume, for simplicity, that $\sigma(nT) = 0$, for all $n \notin E$. Alternatively, we could extend σ to a signal in $\ell^2(\mathbf{Z}, \mathbf{C})$ in ways that would allow $\widehat{f}$ to have more smoothness.

Let $u := f + i\mathcal{H}f$. Now we turn to the construction of $\mathbf{W}u$, or equivalently $\mathbf{W}f$ and $\mathbf{W}\mathcal{H}f$, using a variation on the usual pyramid algorithm. This amounts to calculating all $J + 1$ levels of these transforms at each grid point nT. From Shannon's formula, for all $m \in \mathbf{Z}$,

$$(\mathcal{H}f)(mT) = \sum_{k\in\mathbf{Z}} f(kT)\beta_{m,k}, \tag{$\mathcal{H}$}$$

where

$$\beta_{m,k} = \frac{1-(-1)^{m-k}}{\pi(m-k)} \quad \text{if } k \neq m \text{ and } \beta_{m,m} = 0.$$

By our choice of φ, $f * \varphi = f$ and $\mathcal{H}f * \varphi = \mathcal{H}f$. It follows from equation (♠) that

$$f * \varphi_2(mT) = \sum_{k\in\mathbf{Z}} c_k\, f((m-k)T), \qquad m \in \mathbf{Z}, \tag{$P\varphi 1$}$$

$$\mathcal{H}f * \varphi_2(mT) = \sum_{k\in\mathbf{Z}} c_k\, (\mathcal{H}f)((m-k)T), \qquad m \in \mathbf{Z}. \tag{$P\mathcal{H}\varphi 1$}$$

Using equations (♡) and (♡$\mathcal{H}$), we also see that for all $m \in \mathbf{Z}$,

$$\begin{aligned} f * \psi_2(mT) = &\sum_{k\in\mathbf{Z}} \frac{(d_k^+ + d_k^-)}{2}\, f((m-k)T) \\ &+ i\sum_{k\in\mathbf{Z}} \frac{(d_k^+ - d_k^-)}{2}\, (\mathcal{H}f)((m-k)T), \end{aligned} \tag{$P\psi 1$}$$

$$\begin{aligned} \mathcal{H}f * \psi_2(mT) = &-i\sum_{k\in\mathbf{Z}} \frac{(d_k^+ - d_k^-)}{2} f((m-k)T) \\ &+ \sum_{k\in\mathbf{Z}} \frac{(d_k^+ + d_k^-)}{2} (\mathcal{H}f)((m-k)T). \end{aligned} \tag{$P\mathcal{H}\psi 1$}$$

Continuing inductively, again applying (♠), we deduce that

$$f * \varphi_{2^2}(mT) = \sum_{k\in\mathbf{Z}} c_k\, f * \varphi_2((m-2k)T), \qquad m \in \mathbf{Z}. \tag{$P\varphi 2$}$$

At this second stage we can also similarly write equations for $\mathcal{H}f * \varphi_{2^2}(mT)$, $f * \psi_{2^2}(mT)$, and $\mathcal{H}f * \psi_{2^2}(mT)$, for each $m \in \mathbf{Z}$. Continuing further, by induction, we produce for each level $j \in \{1, \ldots, J\}$, equations for $f * \varphi_{2^j}$, $\mathcal{H}f * \varphi_{2^j}$, $f * \psi_{2^j}$, and $\mathcal{H}f * \varphi_{2^j}$ evaluated at each point mT in terms of $f * \varphi_{2^{j-1}}$, $\mathcal{H}f * \varphi_{2^{j-1}}$, $f * \psi_{2^{j-1}}$, and $\mathcal{H}f * \varphi_{2^{j-1}}$ evaluated at all the points kT. We label these equations $(P\varphi j)$, $(P\mathcal{H}\varphi j)$, $(P\psi j)$, and $(P\mathcal{H}\psi j)$, respectively.

The previous algorithm is the complete modified pyramid algorithm for calculating the wavelet transform of $u = f + i\mathcal{H}f \in P\mathcal{H}_\Omega$ with respect to the *dilation equation solution-wavelet pair* (φ, ψ) on the discrete grid kT. It can also be viewed as a simultaneous calculation of the (φ, ψ)-wavelet transform of f and the $(\mathcal{H}\varphi, \mathcal{H}\psi)$-wavelet transform of f. The calculation of a second transform is the price we pay to ensure convergence of our reconstruction algorithm in the discrete setting, using the lower sampling rates from §5. We choose G^R, instead of the usual G, to ensure that our filter has smaller period, equal to 2Ω, with sufficient $(C^{(\nu)})$ smoothness. Thus we ensure that our coefficients c_k and d_k^+ have rapid decay, and we expect this to facilitate the minimization of numerical errors when the pyramid algorithm is inevitably restricted to a time grid of finite length. We remark that when $\Omega_1 = 4\Omega/3$, $\Omega_1 + \varepsilon_1 - 2\Omega = -\Omega/2$ and the functions c and s may be chosen so that G^R is an even function. In this case, $\widehat{\psi}(\xi) = G^R(\xi/2)\,\widehat{\varphi}(\xi/2)$ for almost all $\xi \in \mathbf{R}$, therefore the pyramid algorithm above for u simplifies to the usual one [MZ2] for each of f and $\mathcal{H}f$.

We now fix $g \in \ell^2(P_\Omega)$ and describe the calculation of $\mathbf{QM}g$ via the pyramid algorithm. In the current setting, which was introduced in §4, $\mathbf{M}g$ may be given by

$$\mathbf{M}g := \sum_{j=1}^{J} g_j * \widetilde{\psi_{2^j}} + g_{J+1} * \widetilde{\varphi_{2^J}}. \qquad (I)$$

Now, $\mathbf{QM}g = \mathbf{MP}g = \mathbf{M}\left(\widetilde{(\mathbf{Q}g_j)_{j=1}^{J+1}}\right)$. Each $\mathbf{Q}g_j \in P\mathcal{H}_\Omega$, and therefore $\mathbf{Q}g_j * \widetilde{\psi_{2^j}} = (\widetilde{\mathbf{Q}g_j} * \psi_{2^j})\tilde{}$ and $\mathbf{Q}g_{J+1} * \widetilde{\varphi_{2^J}} = (\widetilde{\mathbf{Q}g_{J+1}} * \varphi_{2^J})\tilde{}$ may be calculated from the values of $\mathbf{Q}g_j$ at each kT, via the modified pyramid algorithm. Then, invoking equation (I), we can calculate $((\mathbf{MP}g)(mT))_{m\in\mathbf{Z}}$. Here, each

$$\mathbf{Q}g_j = \left(\widehat{g_j}\,\chi_{[0,\infty)}\right)^{\vee},$$

and therefore each grid point evaluation $\mathbf{Q}g_j(mT)$ can be calculated from the values of $g_j(kT)$, via Shannon's theorem, in a manner similar to the calculation of the discrete Hilbert transform described above.

We can now see how to calculate $\mathbf{P_V} = \mathbf{WQM}$ in the discrete setting. And what of $\Gamma = \Gamma_u$ and the projection $\mathbf{P}_\Gamma$? Let us define $M_1' := 3T$, $M_j' := \gamma((6/5)2^j)\,T$ and $M_{J+1}' := \gamma((3/2)2^J)\,T$. Here, $\gamma(t) :=$ the greatest integer less than t, for all real t. Fix $j \in \{1, \dots, J+1\}$. It is easy to check that since $\Omega_1 \le 4\Omega/3$, we have that $M_j' < \pi/\Sigma_j$. We choose a strictly increasing sequence t^j of real numbers from $(mT)_{m\in\mathbf{Z}}$, accumulating at ∞ and $-\infty$ and with $M_j := \sup_{n\in\mathbf{Z}}(t_{n+1}^j - t_n^j) \le M_j'$. Define $\Gamma := \Gamma_u \subseteq \ell^2(P_\Omega)$ given by

$$\begin{aligned}\Gamma := \big\{g = (g_j)_{j=1}^{J+1} \in \ell^2(P_\Omega)\colon\ & \text{for all } j \in \{1, \dots, J\}\\ & g_j(t_n^j) = (u * \psi_{2^j})(t_n^j) =: C_n^j, \text{ for all } n \in \mathbf{Z}\\ & \text{and } g_{J+1}(t_n^{J+1}) = (u * \varphi_{2^J})(t_n^{J+1}) =: C_n^{J+1}, \text{ for all } n \in \mathbf{Z}\big\}.\end{aligned}$$

Also let $\mathbf{T} := \mathbf{P_V P_\Gamma}$. Then, for all $g \in \ell^2(P_\Omega)$, $(\mathbf{T}^n g)_{n\in\mathbf{N}}$ converges in norm to $q = \mathbf{W}u$, with a geometric convergence rate computable from Theorem 5.3 (where we may take $m_j := T$ for each j). For example, we may start the reconstruction algorithm at $g = 0 \in \ell^2(P_\Omega)$.

Now let us calculate the orthogonal projection $\mathbf{P}_\Gamma$. Fix $h \in \ell^2(P_\Omega)$. Next, fix $j \in \{1, \dots, J+1\}$. Let E_j be the set of all $m \in \mathbf{Z}$ such that mT appears in the sequence t^j. Then by Shannon's theorem, the jth coordinate of $\mathbf{P}_\Gamma h$ is given by

$$(\mathbf{P}_\Gamma h)_j = \sum_{n\in\mathbf{Z}} C_n^j \operatorname{sinc}(\Omega(\cdot - t_n^j)) + \sum_{m\notin E_j} \eta_m^j \operatorname{sinc}(\Omega(\cdot - mT)).$$

Here $\eta^j = \eta^j(h) \in \ell^2(\mathbf{Z}, \mathbf{C})$ is chosen subject to the constraints that $\eta_m^j := (\mathbf{W}u)_j(mT) = C_n^j$ whenever $m \in E_j$, so that $mT = t_n^j$ for some $n \in \mathbf{Z}$; and $\zeta = \eta^j$ minimizes

$$\sum_{m\notin E_j} |\zeta_m - h_j(mT)|^2$$

as ζ varies over all $\zeta \in \ell^2(\mathbf{Z}, \mathbf{C})$. It follows that for all $h \in \ell^2(P_\Omega)$ and for each $j \in \{1, \dots, J+1\}$,

$$(\mathbf{P}_\Gamma h)_j = \sum_{n\in\mathbf{Z}} C_n^j \operatorname{sinc}(\Omega(\cdot - t_n^j)) + \sum_{m\notin E_j} h_j(mT) \operatorname{sinc}(\Omega(\cdot - mT)).$$

7. Local Maxima and their Significance

Let us continue with the setting of §§5 and 6. In particular $f \in P_\Omega$ is a signal and our wavelet ψ is such that supp $\widehat{\psi} \subseteq [-2\Omega_1, -\Omega] \cup [\Omega, 2\Omega_1]$. Also, $\widehat{\psi} \in L^\infty$. Under these circumstances,

$$u_1(\xi) := \frac{\widehat{\psi}(\xi)}{i\xi}, \qquad \xi \in \mathbf{R}$$

shares the properties of $\widehat{\psi}$, and $\theta := u_1^\vee \in L^2 \cap C_0$ is such that $(f * \theta)'(t) = f * \psi(t)$, for all $t \in \mathbf{R}$ (since $\psi = \theta'$). Therefore, the local maxima of $|f * \psi|$ occur at those places where the signal f, smoothed by θ, is varying most rapidly. This is similar for $|f * \psi_{2^j}|$, $j \in \{1, \dots, J\}$. Thus we may interpret the positions of the local maxima of $|f * \psi_{2^j}|$ to be the *edges* (points of sharp variation) of the signal f, smoothed by θ_{2^j}, where $f * \theta_{2^j}$ is a smoothing of the signal f, restricted to the frequency range $[-\Omega_1/2^{j-1}, -\Omega/2^j] \cup [\Omega/2^j, \Omega_1/2^{j-1}]$.

We remark that we may think of θ as generating an *integrated wavelet transform* or smoothing transform $\mathbf{I}$, given by

$$\mathbf{I} : u \to (u * \theta_{2^1}, u * \theta_{2^2}, \dots, u * \theta_{2^J}, u * \varphi_{2^J}), \qquad u \in P_\Omega.$$

It is easy to check that $\|\mathbf{I} \cdot\|_{\ell^2(P_\Omega)}$ is equivalent to $\|\cdot\|_2$.

Mallat and Zhong [MZ1, MZ2] only sample $f * \psi_{2^j}$ values at the edges and, indeed, discard all such local maxima values below a threshold. This corresponds to only keeping the sufficiently sharp edge features of the signal at each level j. Mallat and Zhong find that their reconstructed signal is *visually close* to (or hard to distinguish from) the original signal. We discuss one advantage of keeping sample values whose total energy is large (in our one-dimensional setting) in the next section.

8. Not Choosing Local Maxima Slows the Rate of Convergence

As in §4, let $(X, \langle\cdot,\cdot\rangle)$ be a complex Hilbert space with corresponding norm $\|\cdot\|$. Suppose that V and S are closed subspaces of X. Define $U := \mathbf{P}_V\mathbf{P}_S$.

8.1. *Lemma*

Let X, S, V, and U be as defined previously. Then for all $n \in \mathbf{N} \cup \{0\}$,

$$\|U^{2^n}h\| \geq \|h\| \left(\frac{\langle \mathbf{P}_S h, h\rangle}{\|h\|^2}\right)^{2^n} \quad \textit{for all } h \in X\backslash\{0\}.$$

Proof. Fix $h \in X\backslash\{0\}$. From the positive homogeneity in h of the inequality to be shown, we may assume that $\|h\| = 1$. Define the one-dimensional subspace $\Lambda := \{\alpha h : \alpha \in \mathbf{C}\}$ of X. Of course,

$$\mathbf{P}_\Lambda x = \langle x, h\rangle\, h \quad \text{for all } x \in X. \qquad (\ddagger)$$

Note that $\mathbf{P}_\Lambda\mathbf{P}_V = \mathbf{P}_\Lambda$. We proceed via induction on n. Let $n = 0$. By equation ($\ddagger$), for all $x \in X$,

$$\|Ux\| \geq \|\mathbf{P}_\Lambda\mathbf{P}_V\mathbf{P}_S x\| = \|\mathbf{P}_\Lambda\mathbf{P}_S x\| = |\langle \mathbf{P}_S x, h\rangle|. \qquad (\#)$$

Therefore, $\|Uh\| \geq |\langle \mathbf{P}_S h, h\rangle|$. Note that, for all $k \in \mathbf{N}$,

$$\mathbf{P}_V U^k = \mathbf{P}_V\mathbf{P}_V\mathbf{P}_S U^{k-1} = \mathbf{P}_V\mathbf{P}_S U^{k-1} = U^k,$$

$$\mathbf{P}_V[U^*]^{k-1}\mathbf{P}_S = \mathbf{P}_V[(\mathbf{P}_V\mathbf{P}_S)^*]^{k-1}\mathbf{P}_S = \mathbf{P}_V[\mathbf{P}_S\mathbf{P}_V]^{k-1}\mathbf{P}_S = [\mathbf{P}_V\mathbf{P}_S]^k = U^k.$$

Now assume that $n \in \mathbf{N}$ and

$$\|U^{2^{n-1}}h\| \geq |\langle \mathbf{P}_S h, h\rangle|^{2^{n-1}}.$$

Then, from inequality (#) and the identities above, we have that

$$\begin{aligned}\|U^{2^n}h\| &= \|UU^{2^n-1}h\| \geq |\langle \mathbf{P}_S U^{2^n-1}h, h\rangle| \\ &= |\langle \mathbf{P}_S U^{2^{n-1}-1}\mathbf{P}_V U^{2^{n-1}}h, h\rangle| = |\langle U^{2^{n-1}}h, \mathbf{P}_V (U^*)^{2^{n-1}-1}\mathbf{P}_S h\rangle| \\ &= |\langle U^{2^{n-1}}h, U^{2^{n-1}}h\rangle| = \|U^{2^{n-1}}h\|^2 \geq |\langle \mathbf{P}_S h, h\rangle|^{2^n}. \quad \square\end{aligned}$$

8.2. *Lemma*

Let $(X, \langle\cdot,\cdot\rangle)$ be a complex Hilbert space with corresponding norm $\|\cdot\|$. Suppose that V is a closed subspace of X and Γ is a closed affine subspace of X such that there exists $q \in \Gamma \cap V$. Define $\mathbf{T} := \mathbf{P}_V\mathbf{P}_\Gamma$. Then for all $n \in \mathbf{N} \cup \{0\}$,

$$\|\mathbf{T}^{2^n}x - q\| \geq \|x - q\|\ (Z_x)^{2^n},$$

where

$$Z_x := \frac{|\langle \mathbf{P}_\Gamma x - q, x - q\rangle|}{\|x - q\|^2} \quad \textit{for all } x \in X\backslash\{q\}.$$

Proof. Now, Γ is a translation of a closed subspace S of X. Moreover, $S = \Gamma - q$ and $\mathbf{P}_\Gamma\, x = \mathbf{P}_S\, x + q - \mathbf{P}_S\, q$, for all $x \in X$. Let $U := \mathbf{P}_V\mathbf{P}_S$. Applying Lemma 8.1 to this choice of V, S, and U, Lemma 8.2 easily follows. $\square$

We apply Lemma 8.2 to prove our final result concerning our Mallat-Zhong algorithm, in the setting of §§5–7.

We may again consider that our signal $f \in P_\Omega$ is discrete and of the form $(f(kT))_{k\in\mathbf{Z}}$, where $T = \pi/\Omega$, and may apply Shannon's theorem to each level of $\mathbf{W}f$. Let $u := f + i\mathcal{H}f$. Once again, we sample each $(\mathbf{W}u)_j$ at a sequence t^j from $(mT)_{m\in\mathbf{Z}}$ such that t^j has all the properties of §5, for each j. By Theorem 5.4, the Mallat-Zhong alternating projection algorithm converges geometrically in the norm to some point of $\Gamma \cap \mathbf{V}$. The following result gives a lower bound on this geometric convergence rate.

8.3. *Theorem*

With the discrete setting of this section, let f be a signal in P_Ω and $u := f + i\mathcal{H}f$. Let $\Gamma := \Gamma_u \subseteq \ell^2(P_\Omega)$ and set $\mathbf{T} := \mathbf{P}_V\mathbf{P}_\Gamma$. Then for all $n \in \mathbf{N} \cup \{0\}$.

$$\|\mathbf{T}^{2^n}0 - \mathbf{W}u\|_{\ell^2(P_\Omega)} \geq \|\mathbf{W}u\|_{\ell^2(P_\Omega)}(Z_0)^{2^n},$$

where

$$Z_0 := \frac{\displaystyle\sum_{j=1}^{J+1}\sum_{m\notin E_j} |(\mathbf{W}u)_j(mT)|^2}{\displaystyle\sum_{j=1}^{J+1}\sum_{m\in\mathbf{Z}} |(\mathbf{W}u)_j(mT)|^2}.$$

Here, for each $j \in \{1, \ldots, J+1\}$, E_j is the set of all $m \in \mathbf{Z}$ such that mT appears in the sampling sequence t^j.

Proof. Let $q := \mathbf{W}u$. We need only calculate Z_x, as defined in Lemma 8.2, for $x = 0 \in \ell^2(P_\Omega)$. By Shannon's theorem, for all $h \in \ell^2(P_\Omega)$ and for each $j \in \{1, \ldots, J+1\}$,

$$(\mathbf{P}_\Gamma h)_j = \sum_{n\in\mathbf{Z}} C_n^j \operatorname{sinc}(\Omega(\cdot - t_n^j)) + \sum_{m\notin E_j} h_j(mT) \operatorname{sinc}(\Omega(\cdot - mT)).$$

Fix $j \in \{1, \ldots, J+1\}$. Then

$$(\mathbf{P}_\Gamma 0)_j = \sum_{n\in\mathbf{Z}} C_n^j \operatorname{sinc}(\Omega(\cdot - t_n^j)),$$

$$q_j = \sum_{n\in\mathbf{Z}} C_n^j \operatorname{sinc}(\Omega(\cdot - t_n^j)) + \sum_{m\notin E_j} q_j(mT) \operatorname{sinc}(\Omega(\cdot - mT)).$$

However,

$$Z_0 = \frac{|\langle q - \mathbf{P}_\Gamma 0, q\rangle|}{\|q\|^2}.$$

Therefore the result follows from Shannon's theorem and the definition of the norm in $\ell^2(P_\Omega)$. □

8.4. Remark. Theorem 8.3 tells us that for each level j of our wavelet transform, if we sample $(\mathbf{W}u)_j$ on our discrete grid $(mT)_{m\in\mathbf{Z}}$ at a family of points $(mT)_{m\in E_j}$ that allows for unique reconstruction at a geometric rate (as described above), and such that the total energy of all the sampled points at all levels is small compared with the total energy of the signal, then the lower bound on the geometric rate of convergence of our Mallat-Zhong algorithm with initial point 0, is large, that is, close to 1. Choosing a sample point mT inside a neighborhood U of a point kT where $|(\mathbf{W}u)_j(kT)|$ is locally maximum, instead of choosing kT itself (with all other sample points remaining unchanged), is thus certain to increase our lower estimate on the convergence rate. Of course, we can often simply choose two points mT and nT in U so that $|(\mathbf{W}u)_j(nT)| + |(\mathbf{W}u)_j(mT)| > |(\mathbf{W}u)_j(kT)|$; however, the choice of an extra point (or points) increases the total number of sample values and sample points that need to be transmitted to the remote site where reconstruction is to occur. Thus there is some kind of balance going on locally, between sample size and number of samples, that needs to be explored further by numerical and theoretical means.

Theorem 8.3 may therefore be viewed as some theoretical progress in the direction of the well-tested computational facts (see MZ1, MZ2) that sampling at as many local maxima as is practicable gives a better reconstructed image in fewer iterations of the algorithm.

References

[BB] Bauschke, H. H., and Borwein, J. M. (1993). On the convergence of von Neumann's alternating projection algorithm for two sets. *Set-Valued Anal.*, **1**, 185–212.

[Ben] Benedetto, J. (1992). Irregular sampling and frames. In *Wavelets. A Tutorial in Theory and Applications*, C. K. Chui, ed., Academic Press, New York.

[BH] Benedetto, J., and Heller, W. (1990). Irregular sampling and the theory of frames. *Note Mat.* **10(1)**, 103–125.

[Ber1] Berman, Z. (1992). Generalizations and properties of the multiscale maxima and zero-crossings representations. Ph.D. dissertation, University of Maryland, College Park, MD.

[Ber2] ———. (1991). The uniqueness question of discrete wavelet maxima representation. Technical Report TR 91-48r1, System Research Center, University of Maryland, College Park, MD.

[CA] Çetin, A. E., and Ansari, R. (1994). Signal recovery from wavelet transform maxima. *IEEE Trans. Signal Process.*, **42(1)**, 194–196.

[Cha] Champeney, D. C. (1987). *A Handbook of Fourier Theorems*. Cambridge University Press, Cambridge.

[Chu] Chui, C. K. (1992). *An Introduction to Wavelets*. Academic Press, New York.

[D] Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics No. 61, Society for Industrial and Applied Mathematics, Philadelphia, PA.

[DS] Duffin, R. J., and Schaeffer, A. C. (1952). A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.* **72**, 341–366.

[FG] Feichtinger, H. G., and Gröchenig, K. (1993). Theory and practice of irregular sampling. In *Wavelets: Mathematics and Applications*. J. Benedetto and M. Frazier, eds., CRC Press, Boca Raton, FL.

[G1] Gröchenig, K. (1991). Sharp results on random sampling of band-limited functions. *Proc. Nato ASI Workshop Stochastic Processes*. Kluwer Academic Publishers, New York.

[G2] ———. (1992). Reconstruction algorithms in irregular sampling. *Math. Comp.* **59**, 181–194.

[G3] ———. (1993). Irregular sampling of wavelet and short-time Fourier transforms. *Constr. Approx.* **9**, 283–297.

[He] Heller, W. (1991). Frames of exponentials and applications. Ph.D. dissertation, University of Maryland, College Park, MD.

[Hu] Hurt, N. E. (1989). *Phase Retrieval and Zero Crossings.* Kluwer Academic Publishers, New York.

[K] Körner, T. W. (1988). *Fourier Analysis.* Cambridge University Press, Cambridge.

[L] Logan, B. F. (1977). Information in the zero crossings of bandpass signals. *Bell Systems Tech. J.* **56**, 487–510.

[MZ1] Mallat, S., and Zhong, S. (1989). Complete signal representation with multiscale edges. Technical Report No. 483, December 1989. Dept. of Computer Science, Courant Institute of Mathematical Sciences.

[MZ2] Mallat S., and Zhong, S. (1992). Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Mach. Intell.* **14(7)**, 710-732.

[Ma] Marks II, R. J. (1991). *Introduction to Shannon Sampling and Interpolation Theory.* Springer-Verlag New York.

[Me1] Meyer, Y. (1985–1986). Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. Séminaire Bourbaki, No. 662.

[Me2] ———. Un contre-exemple à la conjecture de Marr et à celle de S. Mallat, preprint.

[Me3] ———. (1993). *Wavelets. Algorithms and Applications* (translated and revised by R. D. Ryan), Society for Industrial and Applied Mathematics, Philadelphia, PA.

[N] von Neumann, J. (1950). *Functional Operators Vol. II,* Ann. Math. Studies 22, Princeton, Princeton University Press, Princeton, NJ.

[O] Opial, Z. (1967). Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.* **73**, 591–597.

[PP] Plancherel, M., and Pólya, G. (1938). Fonctions entières et intégrales de Fourier multiples (Seconde partie). *Commentat. Math. Helv.* **10**, 110–163.

[PW] Paley, R., and Wiener, N. (1934). Fourier transforms in the complex domain. *Amer. Math. Soc. Colloq. Publ.* **19**, American Mathematical Society, Providence, RI.

[R] Rudin, W. (1974). *Real and Complex Analysis,* McGraw-Hill, New York.

[S] Shannon, C. E. (1949). Communication in the presence of noise. *Proc. I.R.E.* **37**, 10–21.

[W] Whittaker, E. T. (1915). On the functions which are represented by the expansions of the interpolation theory. *Proc. R. Soc. Edinburgh* **35**, 181–194.

[YW] Youla, D. C., and Webb, H. (1982). Image restoration by the method of convex projections: Part I—Theory. *IEEE Trans. Med. Imaging* **MI-1(2)**, 81–94.

[Y] Young, R. M. (1980). *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York.

Received September 15, 1994

Department of Mathematics and Computer Science, Valdosta State University, Valdosta, Georgia 31698, U.S.A.
e-mail: ckicey@grits.valdosta.peachnet.edu

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, U.S.A.
e-mail: lennard+@pitt.edu