imposes $(m+1)m/2$ constraints and the fact that $\|\phi_j\|$ is a constant imposes $n-1$ constraints. Since the sum of the squares of the entries of the matrix is m, $\alpha = m/n$. Given such a tight frame every possible rotation in $\mathbb{R}^m$ of the frame vectors also gives a tight frame (with constant norm frame vectors). Therefore (modulo rotations and reflections about the origin - any vector φ_j can be negated) all such tight frames are characterized by a manifold of dimension $mn - m(m+1)/2 - (n-1) - m(m-1)/2 = (m-1)(n-m-1)$. When $n = m+1$ the there is only one solution (modulo rotations and reflections) as $n - m - 1 = 0$. In this case $\{\varphi_j\}$ is given by the $n = m+1$ vertices of a regular simplex with edges of length $\sqrt{\frac{2n+2}{n}}$ and whose centroid at the origin. For example when $m = 2$ the vectors $\{\varphi_j\}$ are the vertices of an equilateral triangle with edges of length $\sqrt{3}$, and when $m = 3$ they are the vertices of a regular tetrahedron with edges of length $\sqrt{\frac{8}{3}}$. For arbitrary m and n there seems to be no obvious characterization of the solution manifold. However when $m = 2$ solutions for arbitrary n are given by

$$\varphi_j = \begin{bmatrix} \cos\beta_j \\ \sin\beta_j \end{bmatrix}, \quad \text{for} \quad j \in \{1, 2, \ldots, n\} \quad \text{where} \quad \sum_{j=1}^{n} e^{\imath 2\beta_j} = 0.$$

Example 10 (Generalized Frame Pair) Let

$$\begin{bmatrix} \phi_1 & \phi_2 & \phi_3 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \tilde{\phi}_1 & \tilde{\phi}_2 & \tilde{\phi}_3 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/2 & -1/2 & 0 \end{bmatrix}.$$

The pair $\left(\{\phi_j\}, \{\tilde{\phi}_j\}\right)$ forms a generalized frame pair for $\mathcal{H}$ and $\tilde{\mathcal{H}}$. The operator

$$\sum_n \phi_n \rangle\langle \tilde{\phi}_n = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1/3 & 1/2 \\ 0 & 1/3 & -1/2 \\ 0 & 1/3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

is the identity operator on $\mathcal{H}$ since

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

The transpose of this operator is the identity operator on $\tilde{\mathcal{H}}$. This generalized frame pair *is not* a generalized Riesz basis pair because the matrix of inner products $\left\langle \tilde{\phi}_i, \phi_j \right\rangle$ is given by

$$\begin{bmatrix} 0 & 1/3 & 1/2 \\ 0 & 1/3 & -1/2 \\ 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Example 11 (Generalized Riesz Basis Pair) Let

$$\begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \tilde{\phi}_1, \tilde{\phi}_2 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} -1 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

They form a generalized Riesz basis pair for $(\mathcal{H}, \tilde{\mathcal{H}})$. The operator

$$\sum_n \phi_n \rangle \langle \tilde{\phi}_n = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & -1 & 2 \end{bmatrix}$$

is the identity operator on $\mathcal{H}$ since

$$\begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}$$

The transpose is verified to be the identity operator on $\tilde{\mathcal{H}}$. Also note that the matrix of inner products $(\left\langle \tilde{\phi}_i, \phi_j \right\rangle)$ is given by

$$\begin{bmatrix} -1 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If J is a subset of the index set of a Riesz basis, $\{\phi_j\}_{j\in J}$ and $\left\{\tilde{\phi}_j\right\}_{j\in J}$ form a generalized Riesz basis pair.

4.2 Hilbert Space Decomposition/Recomposition Theorems

There is a natural connection between PR filter banks and transmultiplexers and change of bases for separable Hilbert spaces. Let $\{\phi_n\}$ be a set of vectors in a separable Hilbert space $\mathcal{X}$. For FIR sequences $\{h_i\}$ define the vectors

$$\psi_{i,k} = \sum_n h_i(Mk - n)\phi_n. \tag{4.8}$$

If $Span\,\{\phi_n\} = \mathcal{H}$, the sequences $\{h_i\}$ specify $\mathcal{H}_i = Span\,\{\psi_{i,k}\}$ (which are subspaces of $\mathcal{H}$). What are the conditions on $\{h_i\}$ for $\mathcal{H}_i$ to be a *decomposition* of $\mathcal{H}$? When is such a decomposition *orthogonal*? Similarly given FIR sequences $\{g_i\}$ and vectors $\{\psi_{i,k}\} \subset \mathcal{X}$, we may define the vectors

$$\phi_n = \sum_i \sum_k g_i(n - Mk)\psi_{i,k}. \tag{4.9}$$

This specifies $\mathcal{H} = Span\,\{\phi_n\}$ contained in the direct sum of $\mathcal{H}_i$. What are the conditions on $\{g_i\}$ so that $\mathcal{H}$ is a direct sum of $\mathcal{H}_i$? If the spaces $\mathcal{H}_i$ are orthogonal and $\{\psi_{i,k}\}$ is an orthonormal basis for $\mathcal{H}_i$, when is $\{\phi_n\}$ is an orthogonal basis for $\mathcal{H}$? This section answers these and related questions.

Theorem 21 *(Decomposition Theorem)* Let $\{\phi_n\} \subset \mathcal{X}$ and let $Span\,\{\phi_n\} = \mathcal{H}$. Let $\{\psi_{i,k}\}$, $i \in \{0, 1, \ldots, L-1\}$, be defined as in Eqn. 4.8 and let

$\mathcal{H}_i = Span\{\psi_{i,k}\}$. Then $\mathcal{H}_i$ is a decomposition of $\mathcal{H}$ with $\{\phi_n\}$ recoverable from $\{\psi_{i,k}\}$ as in Eqn. 4.9 (for arbitrary $\{\phi_n\}$) iff $\{h_i\}$ and $\{g_i\}$ form analysis/synthesis banks of an FIR PR filter bank.

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \ldots \oplus \mathcal{H}_{L-1}. \tag{4.10}$$

Proof: Eqn. 4.9 follows from the filter bank PR property because

$$\begin{aligned}\sum_i \sum_k g_i(n-Mk)\psi_{i,k} &= \sum_i \sum_k \sum_l g_i(n-Mk)h_i(Mk-l)\phi_l \\ &= \sum_l \left[\sum_i \sum_k g_i(n-Mk)h_i(Mk-l)\right]\phi_l \\ &= \sum_l \delta(n-l)\phi_l = \phi_n.\end{aligned}$$

Conversely Eqn. 4.9 implies

$$\phi_n = \sum_l \left[\sum_i \sum_k g_i(n-Mk)h_i(Mk-l)\right]\phi_l,$$

and by choosing $\{\phi_n\}$ to be an orthonormal system the filter bank PR property (Eqn. 3.8) follows. □

For a decomposition the number of constituent spaces L must be lower bounded by M in Eqn. 4.8.

Theorem 22 *(Recomposition Theorem)* Let $\{\psi_{i,k}\} \subset \mathcal{X}$ and let $Span\{\phi_{i,k}\} = \mathcal{H}_i$. Define $\{\phi_n\}$ as in Eqn. 4.9 and let $\mathcal{H} = Span\{\phi_n\}$. Then $\mathcal{H}$ is a recomposition of $\mathcal{H}_i$ with $\{\psi_{i,k}\}$ recoverable from $\{\phi_n\}$ (for arbitrary $\{\psi_{i,k}\}$) as in Eqn. 4.8 iff $\{h_i\}$ and g_i form analysis/synthesis filters of an FIR PR transmultiplexer.

Proof: Eqn. 4.8 follows from the transmultiplexer PR property because

$$\sum_n h_i(Mk-n)\phi_n = \sum_n \sum_j \sum_l h_i(Mk-n)g_j(n-Ml)\psi_{j,l}$$

$$= \sum_j \sum_l \left[\sum_n h_i(Mk-n) g_j(n-Ml) \right] \psi_{j,l}$$
$$= \sum_j \sum_l \left[\delta(l-k)\delta(i-j) \right] \psi_{j,l} = \psi_{i,k}.$$

Conversely Eqn. 4.8 implies

$$\psi_{i,k} = \sum_j \sum_l \left[\sum_n h_i(Mk-n) g_j(n-Ml) \right] \psi_{j,l},$$

and by choosing $\{\psi_{i,k}\}$ to be an orthonormal system the transmultiplexer PR property (Eqn. 3.8) follows. □

For a recomposition the number of constituent spaces L must be upper bounded by M in Eqn. 4.9.

Theorem 23 *(Orthogonal Decomposition/Recomposition)* If $\{\phi_n\}$ is an orthonormal basis for $\mathcal{H}$ then $\{\psi_{i,k}\}$ forms an orthonormal basis for $\mathcal{H}_i$ (as defined in Theorem 21) and $\mathcal{H}_i \perp \mathcal{H}_j, i \neq j$ iff $\{h_i\}$ forms analysis filters of an M channel unitary transmultiplexer. Also, if $\{\psi_{i,k}\}$ forms an orthonormal basis for $\mathcal{H}_i$ and $H_i \perp H_j, i \neq j$, then $\{\phi_n\}$ forms an orthonormal basis for $\mathcal{H}$ (as defined in Theorem 22) iff $\{g_i\}$ forms the synthesis filters of an M channel unitary filter bank.

Proof: Let $\langle \phi_m, \phi_n \rangle = \delta(m-n)$. Then

$$\begin{aligned}
\langle \psi_{i,k}, \psi_{j,l} \rangle &= \left\langle \sum_m h_i(Mk-m)\phi_m, \sum_n h_j(Ml-n)\phi_n \right\rangle \\
&= \sum_m \sum_n h_i(Mk-m) h_j(Ml-n) \langle \phi_m, \phi_n \rangle \\
&= \sum_n \sum_m h_i(Mk-m) h_j(Ml-n) \delta(m-n) \\
&= \delta(k-l)\delta(i-j)
\end{aligned}$$

iff $\{h_i\}$ forms a unitary transmultiplexer. However Theorem 21 implies $\{h_i\}$ and $\{g_i\}$ must form a PR filter bank. This is possible iff $L = M$, and $g_i(n) = h_i(-n)$.

Conversely let $\langle \psi_{i,k}, \psi_{j,l} \rangle = \delta(k-l)\delta(i-j)$. Then

$$\begin{aligned} \langle \phi_m, \phi_n \rangle &= \left\langle \sum_i \sum_k g_i(m-Mk)\psi_{i,k}, \sum_j \sum_l g_j(n-Ml)\psi_{j,l} \right\rangle \\ &= \sum_i \sum_j \sum_k \sum_l g_i(m-Mk) g_j(n-Ml) \left[\langle \psi_{i,k}, \psi_{j,l} \rangle\right] \\ &= \sum_n \sum_m g_i(m-Mk) g_j(n-Ml) \left[\delta(k-l)\delta(i-j)\right] \\ &= \sum_i \sum_k g_i(m-Mk) g_i(n-Mk) = \delta(m-n) \end{aligned}$$

iff $\{g_i\}$ forms the synthesis filters of a unitary filter bank. However, from Theorem 22, $\{g_i\}$ must form synthesis filters of a PR transmultiplexer. Therefore $L = M$ and $h_i(n) = g_i(-n)$. □

In summary, decomposition requires the filter bank PR property, recomposition requires the transmultiplexer PR property, and either in conjunction with orthogonality requires the M channel (i.e $L = M$) unitary PR property.

The PR properties give a more structured decomposition/recomposition of separable Hilbert spaces than discussed until now. Given $\left\{\tilde{\phi}_n\right\}$ let $Span\left\{\tilde{\phi}_n\right\} = \tilde{\mathcal{H}}$. Define

$$\tilde{\psi}_{i,k} = \sum_n g_i(n-Mk)\tilde{\phi}_n. \tag{4.11}$$

Then $\tilde{\mathcal{H}}_i = Span\left\{\tilde{\psi}_{i,k}\right\}$ forms a decomposition of $\tilde{\mathcal{H}}$ iff $\{g_i\}$ forms synthesis filters of a PR filter bank (from Lemma 6 we can interchange analysis and synthesis filters). Moreover

$$\tilde{\phi}_n = \sum_i \sum_k h_i(Mk-n)\tilde{\psi}_{i,k}. \tag{4.12}$$

The analysis filters $\{h_i\}$ of this PR filter bank may be used on another set of functions $\{\phi_n\}$ to give a decomposition of $Span\{\phi_n\} = \mathcal{H}$. These two sets of functions $\{\phi_n\}$ and $\left\{\tilde{\phi}_n\right\}$ could be arbitrary. Assume that they are such that the operators $I = \sum_n \phi_n \rangle\langle \tilde{\phi}_n$ $\tilde{I} = \sum_n \tilde{\phi}_n \rangle\langle \phi_n$ are well defined. I will be called the *projex* operator of the pair $(\left\{\tilde{\phi}_n\right\}, \{\phi_n\})$, and similarly, $\tilde{I}$, the *projex* operator of $(\{\phi_n\}, \left\{\tilde{\phi}_n\right\})$.

Theorem 24 *(Main Decomposition Theorem)* Given $\{\phi_n\}, \left\{\tilde{\phi}_n\right\} \subset \mathcal{X}$, and a PR filter bank with analysis and synthesis filters $\{h_i\}$ and $\{g_i\}$, $i \in \{0, 1, \ldots, L-1\}$, let $\{\psi_{i,k}\}$ and $\left\{\tilde{\psi}_{i,k}\right\}$ be defined as in Eqn. 4.8 and Eqn. 4.11. Then there is a natural decomposition of $Span\{\phi_n\} = \mathcal{H}$, and $Span\left\{\tilde{\phi}_n\right\} = \tilde{\mathcal{H}}$ with a corresponding decomposition of projex operators.

$$I = \sum_n \phi_n \rangle\langle \tilde{\phi}_n = \sum_{i=0}^{L-1}\left[\sum_k \psi_{i,k} \rangle\langle \tilde{\psi}_{i,k}\right] \stackrel{\text{def}}{=} \sum_{i=0}^{L-1} I_i$$

and

$$\tilde{I} = \sum_n \tilde{\phi}_n \rangle\langle \phi_n = \sum_{i=0}^{L-1}\left[\sum_k \tilde{\psi}_{i,k} \rangle\langle \psi_{i,k}\right] \stackrel{\text{def}}{=} \sum_{i=0}^{L-1} \tilde{I}_i$$

Proof: The decomposition of the spaces $\mathcal{H}$ and $\tilde{\mathcal{H}}$ follow immediately from Theorem 21. As for the projex operators from the filter bank PR property

$$\begin{aligned}
\sum_{i=0}^{L-1} I_i &= \sum_{i=0}^{L-1}\sum_k \psi_{i,k} \rangle\langle \tilde{\psi}_{i,k} \\
&= \sum_{i=0}^{L-1}\sum_k\sum_n h_i(Mk-l)\phi_n \rangle\langle \sum_m g_i(m-Mk)\tilde{\phi}_m \\
&= \sum_n\sum_m \phi_n \rangle\langle \tilde{\phi}_m \left[\sum_{i=0}^{L-1}\sum_k h_i(Mk-n)g_i(m-Mk)\right] \\
&= \sum_n\sum_m \phi_n \rangle\langle \tilde{\phi}_m \left[\delta(n-m)\right] = \sum_n \phi_n \rangle\langle \tilde{\phi}_n = I.
\end{aligned}$$

The decomposition of $\tilde{I}$ follows similarly. □

We now give the corresponding recomposition theorem.

Theorem 25 *(Main Recomposition Theorem)* Given $\{\psi_{i,k}\}, \left\{\tilde{\psi}_{i,k}\right\} \subset \mathcal{X}$ for $i \in \{0, 1, \ldots, L-1\}$, and a PR filter bank with analysis and synthesis filters $\{h_i\}$ and $\{g_i\}$, let $\{\phi_n\}$ and $\left\{\tilde{\phi}_n\right\}$ be defined as in Eqn. 4.9 and Eqn. 4.12. Then there is a natural recomposition of $Span\{\psi_{i,k}\} = \mathcal{H}_i$ and

$Span\left\{\tilde{\psi}_{i,k}\right\} = \tilde{\mathcal{H}}_i$ with a corresponding recomposition of projex operators.

$$I = \sum_n \phi_n \rangle\langle \tilde{\phi}_n = \sum_{i=0}^{L-1}\left[\sum_k \psi_{i,k}\rangle\langle \tilde{\psi}_{i,k}\right] \stackrel{\text{def}}{=} \sum_{i=0}^{L-1} I_i$$

and

$$\tilde{I} = \sum_n \tilde{\phi}_n \rangle\langle \phi_n = \sum_{i=0}^{L-1}\left[\sum_k \tilde{\psi}_{i,k}\rangle\langle \psi_{i,k}\right] \stackrel{\text{def}}{=} \sum_{i=0}^{L-1} \tilde{I}_i$$

Proof: The recomposition of spaces follows directly from Theorem 22. Also from the transmultiplexer PR property

$$\begin{aligned}
I &= \sum_n \sum_i \sum_k h_i(Mk-n)\psi_{i,k}\rangle\langle \sum_j \sum_m g_i(n-Mm)\tilde{\psi}_{j,m} \\
&= \sum_{i,j,k,m} \psi_{i,n}\rangle\langle \tilde{\psi}_{j,m}\left[\sum_n h_i(Mk-n)g_j(n-Mm)\right] \\
&= \sum_{i,j,k,m} \psi_{i,n}\rangle\langle \tilde{\psi}_{i,m}\left[\delta(i-j)\delta(n-m)\right] \\
&= \sum_{i=0}^{L-1}\left[\sum_k \psi_{i,k}\rangle\langle \tilde{\psi}_{i,k}\right] = \sum_{i=0}^{L-1} I_i
\end{aligned}$$

□

Theorem 26 *(Main Orthogonal Decomposition/Recomposition)* If $\{\phi_n\}$ and $\left\{\tilde{\phi}_n\right\}$ form a biorthogonal system then the decomposition (described in Theorem 24 and Theorem 25) is biorthogonal iff $\{h_i\}$ and $\{g_i\}$ form analysis and synthesis filters of an M channel filter bank (or transmultiplexer).

Proof: Let $\left\langle \phi_m, \tilde{\phi}_n \right\rangle = \delta(m-n)$. Then

$$\left\langle \psi_{i,k}, \tilde{\psi}_{j,l} \right\rangle = \left\langle \sum_m h_i(Mk-m)\phi_m, \sum_n g_j(n-Ml)\tilde{\phi}_n \right\rangle$$

$$\begin{aligned} &= \sum_m \sum_n h_i(Mk-m) g_j(-Ml+n) \left\langle \phi_m, \tilde{\phi}_n \right\rangle \\ &= \sum_n h_i(Mk-n) g_j(n-Ml) \\ &= \sum_n h_i(n) g_j(M(k-l)-n) = \delta(i-j)\delta(k-l) \end{aligned}$$

iff $\{h_i\}$ and $\{g_i\}$ form a PR transmultiplexer. Theorem 21, however requires that $\{h_i\}$ and $\{g_i\}$ form a PR filter bank. Therefore they must form an M channel PR filter bank. $\mathcal{H}_i \perp \tilde{\mathcal{H}}_j, i \neq j$, and moreover, $\{\psi_{i,k}\}$ and $\left\{\tilde{\psi}_{i,k}\right\}$ forms a biorthogonal system. Conversely let $\langle \psi_{i,k}, \psi_{j,l} \rangle = \delta(i-j)\delta(k-l)$. Then

$$\begin{aligned} \left\langle \phi_n, \tilde{\phi}_m \right\rangle &= \left\langle \sum_i \sum_l g_i(n-Ml)\psi_{i,l}, \sum_j \sum_k h_j(Mk-m)\tilde{\psi}_{j,k} \right\rangle \\ &= \sum_{i,j,k,l} \left\langle \psi_{i,l}, \tilde{\psi_{j,k}} \right\rangle [g_i(n-Ml)h_j(Mk-m)] \\ &= \sum_i \sum_k g_i(n-Mk)h_i(Mk-m) = \delta(n-m), \end{aligned}$$

iff $\{h_i\}$ and $\{g_i\}$ form a PR transmultiplexer. From Theorem 22 these sequences must form a PR filter bank and hence the result. □

Remarks: Notice that in the above theorems $\{\phi_n\}$ and $\left\{\tilde{\phi}_n\right\}$ could be arbitrary. Therefore, we have a number of consequences that must be noted.

1. Let $(\{\phi_k\}, \left\{\tilde{\phi}_k\right\})$ form a (generalized) frame pair for $(\mathcal{H}, \tilde{\mathcal{H}})$. If $\{h_i\}$ and $\{g_i\}$ forms a PR filter bank then from Theorem 24 (decomposition of projex operators), $(\{\psi_{i,k}\}, \left\{\tilde{\psi}_{i,k}\right\})$ also forms a (generalized) frame, dual frame pair for $(\mathcal{H}, \tilde{\mathcal{H}})$. It may not be the case that $\mathcal{H}_i \perp \tilde{\mathcal{H}}_j, i \neq j$.

2. Let $(\{\phi_n\}, \left\{\tilde{\phi}_n\right\})$ form a (generalized) Riesz basis pair for $\left(\mathcal{H}, \tilde{\mathcal{H}}\right)$. If $\{h_i\}$ and $\{g_i\}$ form an M-channel PR filter bank from Theorem 26 the decomposition gives (generalized) Riesz basis pairs $(\{\psi_{i,k}\}, \left\{\tilde{\psi}_{i,k}\right\})$ for $\left(\mathcal{H}_i, \tilde{\mathcal{H}}_i\right)$ with an additional *biorthogonality* property: $\mathcal{H}_i \perp \tilde{\mathcal{H}}_j, i \neq j$.

3. Let $\{\phi_k\}$ be a tight frame for $\mathcal{H}$. If $\{h_i\}$ forms a *unitary* filter bank, $\{\psi_{i,k}\}$ also forms a tight frame for $\mathcal{H}$. The decomposition may not be orthogonal. One may not have $\mathcal{H}_i \perp \mathcal{H}_j, i \neq j$.

4. Let $\{\phi_n\}$ be an orthonormal basis for $\mathcal{H}$. If $\{h_i\}$ forms an M channel *unitary* filter bank, the decomposition is orthogonal and $\{\psi_{i,k}\}$ forms an orthonormal basis for $\mathcal{H}_i$.

5. Let $(\{\psi_{i,k}\}, \left\{\tilde{\psi}_{i,k}\right\})$ be a (generalized) frame pair for $(\mathcal{H}_i, \tilde{\mathcal{H}}_i)$. If $\{h_i\}$ and $\{g_i\}$ form a PR transmultiplexer then from Theorem 25 (recomposition of projex operators), $(\{\phi_n\}, \left\{\tilde{\phi}_n\right\})$ forms a (generalized) frame pair for $(\mathcal{H}, \tilde{\mathcal{H}})$.

6. If $(\{\psi_{i,k}\}, \left\{\tilde{\psi}_{i,k}\right\})$, is a (generalized) Riesz basis for pair for $(\mathcal{H}_i, \tilde{\mathcal{H}}_i)$, and $\mathcal{H}_i \perp \tilde{\mathcal{H}}_j$, for $i \neq j$, then if $\{h_i\}$ and $\{g_i\}$ form an M-channel PR transmultiplexer (filter bank), then $(\{\phi_n\}, \left\{\tilde{\phi}_n\right\}$ forms a (generalized) Riesz basis pair for $(\mathcal{H}, \tilde{\mathcal{H}})$.

7. If $\{\psi_{i,k}\}$ is a tight frame for $\mathcal{H}_i$ and $\{h_i\}$ forms a *unitary* filter bank, then $\{\phi_n\}$ forms a tight frame for $\mathcal{H}$.

8. Let $\{\psi_{i,k}\}$ be a tight frame for $\mathcal{H}_i$ and let $\mathcal{H}_i \perp \mathcal{H}_j, i \neq j$. If $\{h_i\}$ forms an M channel *unitary* filter bank, then $\{\phi_n\}$ forms an orthonormal basis for $\mathcal{H}$.

The preceding discussion gives a rich family of ways to decompose a separable Hilbert space. *No assumptions are made on the dimension of the analysis or synthesis bank; all results hold in the multidimensional case also.* The only difference is that the functions $\{\phi_n\}$, etc. are indexed by $n \in \mathbf{Z}^d$, rather than $n \in \mathbf{Z}$. Consider an M channel PR unitary filter bank. This gives an orthonormal decomposition of a Hilbert space $\mathcal{H}$ equipped with an orthonormal basis into M orthogonal subspaces each equipped with an orthonormal basis. By changing the filters $\{h_i\}$ one obtains a rich family of decompositions. One could further decompose each of the subspaces $\mathcal{H}_i$ using a variety of unitary FIR filter banks (with probably a different choice of of the factor M). Again, sets of spaces (leaves of this tree decomposition) could be recombined using a unitary synthesis filter bank. Moreover, the synthesis filter bank could be different from any of the analysis filter banks.

4.3 Multiplicity M Wavelet Tight Frames

Wavelet orthonormal bases have been constructed and studied extensively from both mathematical and signal processing points of view [21, 56, 57, 19, 15, 89, 90, 19, 73]. Wavelets overcome some of the shortcomings of Short-Time Fourier decompositions [61, 35, 90, 22, 40, 33], by decomposing a signal into channels that have the same bandwidth on a logarithmic scale. High frequency channels have wide bandwidth and low frequency channels have narrow bandwidth. This is well adapted for the analysis of sharp transients (spikes, impulses etc.,) submerged in low frequency ambient signals. However, if one has to resolve high frequency signals with relatively narrow bandwidth (like a long RF pulse), then the relative narrow bandwidth, makes the wavelet decomposition unsuitable. In this section we propose a solution to this problem, namely multiplicity M wavelet tight frames (WTFs). Multiplicity M wavelet theory is a compromise between the short-time fourier transform, and the multiplicity 2 wavelet transform (octave band frequency decomposition). Multiplicity M wavelets have been constructed independently by Heller, Wells and Resnikoff [43] and Zou and Tewfik [98] (called M-band wavelets). This section derives a complete parameterization of all compactly supported multiplicity M WTFs. Multiplicity M wavelets give the required flexibility to multiplicity 2 wavelet decompositions.

Our development of multiplicity M wavelets will parallel the multiplicity 2 construction of Daubechies [21]. As will be apparent from the development, filter bank theory plays a fundamental role in multiplicity M wavelet theory.

Definition 12 A sequence $h_0(n)$ is called a multiplicity M, unitary *scaling vector* if it satisfies the following linear and quadratic constraints:

$$h_0(k) = \sqrt{M} \tag{4.13}$$

and

$$\sum_k h_0(k)h_0(k+Ml) = \delta(l). \qquad (4.14)$$

Eqn. 4.14 is precisely the unitary transmultiplexer PR property corresponding to $h_0(k)$ (considered as a filter in a unitary transmultiplexer). Moreover, the linear constraint ensures that h_0 corresponds to a *lowpass* filter. This observation leads to a complete parameterization of multiplicity M finite length unitary scaling vectors. Let $h_0(n) = 0$ for $n < 0$ and $h(0) \neq 0$. Also (by padding with zeros if necessary) let length of h_0 be $N = MK$.

Lemma 26 For h_0 of length $N = MK$ let $H_0(z) = \sum_{k=0}^{M-1} z^{-k} H_{0,k}(z^M)$ and let

$$\mathbf{h}_0(z) = \begin{bmatrix} H_{0,1}(z) \\ H_{0,2}(z) \\ \cdots \\ H_{0,M-1}(z) \end{bmatrix} \quad \text{and} \quad \mathbf{e} = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix}.$$

Then there exist $K-1$ unit norm vectors v_i such that

$$\mathbf{h}_0(z) = \left\{ \prod_{i=1}^{K-1} \left[I - v_i v_i^T + z^{-1} v_i v_i^T \right] \right\} \mathbf{e}. \qquad (4.15)$$

Moreover, the following *equipartition property* also holds:

$$\sum_k h_0(Mk+m) = \sqrt{M} \quad \text{for all } m. \qquad (4.16)$$

Proof: From Eqn. 4.14 the unitary scaling vector may be considered to be an analysis filter in a unitary FIR filter bank. Clearly, $H_{0,j}(z)$ is the j^{th} (dual first-orthant) polyphase component of $H_0(z)$. Therefore (from Eqn. 3.8) $\sum_{j=0}^{M-1} H_{0,j}(z^{-1})H_{0,j}(z) = 1$. Now $\mathbf{h}_0(z)$ is a polynomial vector of degree $K-1$ unitary on the unit circle. From Fact 4, $\mathbf{h}_0(z)$ is parameterized by $K-1$ unit norm Householder parameters v_i, $i \in$

$\{1,\ldots,K-1\}$, and the unit vector v_0. Also

$$\mathbf{h}_0(z) = \left\{ \prod_{i=1}^{K-1} \left[I - v_i v_i^T + z^{-1} v_i v_i^T \right] \right\} v_0.$$

From the linear constraint Eqn. 4.13,

$$\sqrt{M} = \sum_k h_0(k) = H_0(z)|_{z=1} = \left[\sum_k z^{-k} H_{0,k}(z) \right]_{z=1} = \sum_k H_{0,k}(z)|_{z=1}. \qquad (4.17)$$

Since, for all i, $\left[I - v_i v_i^T + z^{-1} v_i v_i^T\right]_{z=1} = I$, $\mathbf{h}_0(1) = v_0$ and $\sqrt{M} = \sum_k H_{0,j}(z)|_{z=1} = \sum_k (v_0)_k = 1$, where $(v_0)_k$ is the k^{th} component of $v_0 \in \mathbb{R}^M$. The unique unit vector v_0 that solves this equation is $\mathbf{e}$. □

Remark: Therefore a length $N = MK$ unitary scaling vector is parameterized by $K-1$ unitary vectors $v_i \in \mathbb{R}^M$, and therefore has $(M-1)(K-1)$ degrees of freedom. Also from Eqn. 4.16 all the (generalized) polyphase components of h_0 have the same frequency response at $\omega = 0$.

The importance of unitary scaling vectors in wavelet theory stems from the fact that they uniquely determine the multiresolution analysis that gives rise to a multiplicity M WTF. However, wavelets and hence the WTF, is *not* uniquely determined by the unitary scaling vector. The wavelets are parameterized by precisely $\binom{M-1}{2}$ parameters. We now state and prove this result.

Theorem 27 *(Wavelet Tight Frames Theorem)* Given a length $N = MK$, multiplicity M, unitary scaling vector h_0, there exists a unique, compactly supported *scaling function* $\psi_0(t) \in L^2(\mathbb{R})$ with support in $[0, \frac{N-1}{M-1})$, determined by the *scaling recursion*,

$$\psi_0(t) = \sqrt{M} \sum_k h_i(k) \psi_0(Mt - k). \qquad (4.18)$$

Moreover, there exist $(M-1)$ unitary wavelet vectors $h_i, i \in \{1,\ldots,M-1\}$, all of the same length N, that satisfy the equation,

$$\sum_k h_i(k) h_j(k + Ml) = \delta(l)\delta(i-j). \qquad (4.19)$$

The wavelet vectors are non-unique and parameterized by $\begin{pmatrix} M-1 \\ 2 \end{pmatrix}$ parameters If for each wavelet vector we define the corresponding *wavelet*, $\psi_i(t)$, compactly supported in $[0, \frac{N-1}{M-1}]$, and given by,

$$\psi_i(t) = \sqrt{M} \sum_k h_i(k) \psi_0(Mt - k), \tag{4.20}$$

then, $\{\psi_{i,j,k}(t)\}$ defined by

$$\psi_{i,j,k}(t) = M^{j/2} \psi_i(M^j t - k), \tag{4.21}$$

forms a tight frame for $L^2(\mathbb{R})$. In other words, for all $f \in L^2(\mathbb{R})$,

$$f(t) = \sum_{i=1}^{M-1} \sum_{j,k} \langle f, \psi_{i,j,k}(t) \rangle \, \psi_{i,j,k}(t). \tag{4.22}$$

Also

$$f(t) = \sum_k \langle f, \psi_{0,0,k}(t) \rangle \, \psi_{0,0,k}(t) + \sum_{i=1}^{M-1} \sum_{j=1}^{\infty} \sum_k \langle f, \psi_{i,j,k}(t) \rangle \, \psi_{i,j,k}(t). \tag{4.23}$$

Proof: The proof is similar to that for the multiplicity 2 case in [57]. The difficulty in the multiplicity M case is in the construction of the wavelet vectors h_i. The outline of the proof is as follows:

1. $\psi_0(t)$ is constructed and shown to be in $L^2(\mathbb{R})$.
2. The wavelet vectors are constructed.
3. The functions $\{\psi_{i,j,k}\}$ is shown to be a tight frame.

From Eqn. 4.18 $\psi_0(t)$ is fixed point of the operator T defined by

$$Tf(t) = \sqrt{M} \sum_k h_0(k) f(Mt - k).$$

Let $\psi_0^0(t)$, the characteristic function of $[0,1)$, initialize the recursion $\psi_0^{j+1} = T\psi_0^j$. This recursion converges to the scaling function. At each stage this recursion preserves both the integral and energy of ψ_0^j:

$$\int_{\mathbf{R}} \psi_0^j(t)\,dt = 1 \quad \text{and} \left\|\psi_0^j\right\|_2 = 1. \tag{4.24}$$

Both results follow by induction. Since $\psi_0^0(t)$ is the characteristic function of $[0,1)$, Eqn. 4.24 is true for $j=0$. By induction on j

$$\begin{aligned}\int_{\mathbf{R}} \psi_0^{j+1}(t)\,dt &= \int_{\mathbf{R}} \sqrt{M}\sum_k h_0(k)\psi_0^j(Mt-k)\,dt \\ &= \frac{1}{\sqrt{M}}\sum_k h_0(k)\left[\int_{\mathbf{R}} \psi_0^j(t-k)\,dt\right] \\ &= \frac{1}{\sqrt{M}}\sum_k h_0(k) = 1\end{aligned}$$

and

$$\begin{aligned}&\left\langle \psi_0^{j+1}(t), \psi_0^{j+1}(t-k)\right\rangle \\ &= \left\langle \sqrt{M}\sum_n h_0(n)\psi_0^j(Mt-n), \sqrt{M}\sum_l h_0(l)\psi_0^j(Mt-Mk-l)\right\rangle \\ &= \sum_{n,l} h_0(n)h_0(l)\left[M\int \psi_0^j(Mt-n)\psi_0^j(Mt-Mk-l)\,dt\right] \\ &= \sum_{n,l} h_0(n)h_0(l)\delta(n-Mk-l) = \sum_n h_0(n)h_0(n-Mk) = \delta(k).\end{aligned}$$

Since $\left\|\psi_0^j\right\| = 1$, ψ_0^j is a sequence in the unit ball in $L^2(\mathbb{R})$. By the weak compactness of the unit ball [28], a subsequence of ψ_0^j converges weakly to some function $g \in L^2(\mathbb{R})$. Since the sequence of Fourier transforms, $\widehat{\psi_0^j}(\omega)$, converges uniformly on compact subsets, $\psi_0^j(t)$, converges uniquely to the tempered distribution $\psi_0(t)$ that satisfies Eqn. 4.18.

$$\hat{\psi}_0^{j+1}(\omega) = \hat{\psi}_0^j\left(\frac{\omega}{M}\right)\left[\frac{1}{\sqrt{M}}H_0\left(\frac{\omega}{M}\right)\right]. \tag{4.25}$$

Since $H_0(\omega)$ is a trigonometric polynomial and $H_0(0) = \sqrt{M}$, the sequence $\hat{\psi}_0^j(\omega)$ converges pointwise uniformly on compact subsets to the function

$$\hat{\psi}_0(\omega) = \prod_{j\geq 1} \left[\frac{1}{\sqrt{M}} H_0(\frac{\omega}{M^j}) \right]. \tag{4.26}$$

Let $C = \max \left| \hat{\psi}_0^0(\omega) \right|$ for $|\omega| \leq 1$. Let $A = \max [\frac{1}{\sqrt{M}} |H_0(\omega)|] \leq 1$ by Eqn. 4.15. For any $\omega, |\omega| \geq 1$,

$$M^{\lfloor log_M(\omega) \rfloor} \leq \omega \leq M^{\lceil log_M(\omega) \rceil}$$

and therefore from Eqn. 4.25

$$\left| \hat{\psi}_0^j(\omega) \right| \leq C A^{\lceil \log_M(\omega) \rceil}$$

thus showing that $\hat{\psi}_0^j$ are bounded by polynomial growth for large ω. In the time domain $\psi_0^j(t)$, converges to a *unique* distribution $\psi_0(t)$ that satisfies Eqn. 4.18. Since the distribution topology is weaker than the weak topology in $L^2(\mathbb{R})$ [46], $\psi_0(t) = g(t)$ and therefore $\psi_0(t) \in L^2(\mathbb{R})$.

We now show that $\psi_0(t)$ is compactly supported in $[0, \frac{N-1}{M-1}]$. If $Supp \left\{ \psi_0^j \right\} = [0, a_j)$ it is easy to see that $a_{j+1} = (a_j + N - 1)/M$ and hence

$$a_j = M^{-j} \left(1 + (N-1) \frac{M^j - 1}{M - 1} \right).$$

Therefore as $j \rightarrow \infty$, $a_j \rightarrow \frac{N-1}{M-1}$.

The scaling vector determines the unitary vector $\mathbf{h}_0$ in Eqn. 4.15. This unitary vector polynomial can be completed to give a unitary matrix polynomial - one just completes the unitary vector $\mathbf{e}$ to an orthogonal matrix V_0 by a Gram-Schmidt process. The resulting unitary matrix may be considered to be the polyphase component matrix of the filters $h_i(n)$ of a unitary transmultiplexer. These filters satisfy Eqn. 4.19 and are the wavelet vectors. By construction all of then have length $N = MK$. The wavelets defined in Eqn. 4.20 also have compact support in $[0, \frac{N-1}{M-1}]$.

For $f \in L^2(\mathbb{R})$, and fixed i and j, let $I_{i,j}$ denote the projex operator corresponding to $(\{\psi_{i,k}\}, \{\psi_{i,k}\})$.

$$I_{i,j} = \sum_k \psi_{i,j,k} \rangle\langle \psi_{i,j,k}.$$

Since $\psi_i(t)$ is compactly supported, these operators are well defined for all $f \in L^2(\mathbb{R})$. Indeed, if $L = \lceil \frac{N-1}{M-1} \rceil$, then, $\{\psi_{i,j,Lm+k}(t)/\|\psi_i\|\}$ is an orthonormal family for any fixed $k = 0, 1, 2, \ldots, L-1$ and hence by Bessel inequality [74],

$$\|I_{i,j} f\| \leq L \|\psi_i\|^2 \|f\|. \tag{4.27}$$

Moreover, the projex operators are uniformly bounded in j. Since $\{h_i\}$ constitutes an M-channel unitary transmultiplexer, it also constitutes an M-channel unitary filter bank and therefore from Theorem 24

$$I_{0,j} = \sum_{i=0}^{M-1} I_{i,j-1}. \tag{4.28}$$

By telescoping one concludes that, for a fixed J,

$$I_{0,J} = \sum_{i=1}^{M-1} \sum_{j=-\infty}^{J-1} I_{i,j}.$$

Now it suffices to show that $I_{0,j}$ approximates the identity operator in $L^2(\mathbb{R})$.

$$\lim_{j \to \infty} I_{0,j} f = f.$$

One expects this since $\psi_{0,j,0}$ approaches the Dirac measure at the origin as $j \to \infty$. From Eqn. 4.27 it suffices to show that $\|I_{0,j} f - f\|_2 = 0$, for a dense subset of $L^2(\mathbb{R})$. For continuous f with compact support, if j and k together approach infinity such that $M^{-j} k \to t$, then,

$$\lim_{j,k \to \infty} \langle \psi_{0,j,k}, f(x) \rangle = f(t)$$

uniformly in x. We will now show that $M^{-j/2} \psi_{0,j,k}(t)$ forms a partition of unity and this would imply that the result is true. In the construction of $\psi_0(t)$, we started with

$\psi_0^0(t)$, the characteristic function of $[0, 1)$. Clearly, $\psi_0^0(t)$ forms a partition of unity since

$$\sum_k \psi_0^0(t+k) = 1. \tag{4.29}$$

Now by induction, we have for any j,

$$\sum_k \psi_0^j(t+k) = \sqrt{M} \sum_k \sum_l h_0(l) \psi_0^{j-1}(Mt - Mk - l). \tag{4.30}$$

Changing variables and interchanging the order of summation and invoking Eqn. 4.16 in Lemma 26 we get

$$\begin{aligned} \sum_k \psi_0^j(t+k) &= \sqrt{M} \sum_k \sum_l h_0(Mk - l) \psi_0^{j-1}(Mt + l) \\ &= \sqrt{M} \sum_l [\sum_l h_0(Mk - l)] \psi_0^{j-1}(Mt + l) \\ &= \sum_l \psi_0^{j-1}(Mt + l) = 1 \end{aligned} \tag{4.31}$$

and therefore $M^{-j/2}\psi_{0,j,k}(t)$ forms a partition of unity. □

Multiplicity M, compactly supported WTFs have a natural multiresolution analysis structure associated with them. It essentially comes from the filter bank structure that leads to Eqn. 4.28. If we define the spaces, $W_{i,j} = Span\{\psi_{i,j,k}\}$, then, for $f \in L^2(\mathbb{R})$, $I_{i,j} f \in W_{i,j}$, and in particular for $f \in W_{0,j}$

$$f = I_{0,j} f = \sum_{i=0}^{M-1} I_{i,j-1} f$$

leading to a decomposition

$$W_{0,j} = W_{0,j-1} \oplus W_{1,j-1} \ldots \oplus W_{M-1,j-1}. \tag{4.32}$$

Also since $I_{0,j}$ approaches the identity operator on $L^2(\mathbb{R})$ $\lim_{j \to \infty} W_{0,j} = L^2(\mathbb{R})$. Similarly $\lim_{j \to -\infty} W_{0,j} = \{0\}$. Hence we have a multiresolution analysis of $L^2(\mathbb{R})$ with a chain of closed subspaces:

$$\{0\} \subset \ldots W_{0,-1} \subset W_{0,0} \subset W_{0,1} \ldots \subset L^2(\mathbb{R}).$$

In general, the spaces $W_{i,j}$ for fixed j are not mutually orthogonal. But when the WTF is an orthonormal basis, then $W_{i,j} \perp W_{k,j}$, $i \neq k$ and $i, k \in \{1, 2, \ldots, M-1\}$. Notice that the multiresolution analysis is determined only by $W_{0,j}$, i.e the scaling function and *not by the wavelets.*

Given a unitary filter bank with a "DC" condition (Eqn. 4.13), there exists a unique WTF. Given a scaling vector (and not the wavelet vectors) there is a unique multiresolution analysis associated with a multiplicity M WTF. The WTF, however is determined by $\begin{pmatrix} M-1 \\ 2 \end{pmatrix}$ degrees of freedom that are required to specify the wavelet vectors (of the same McMillan degree as the scaling vector).

4.3.1 Characterization of Orthonormality for a WTF

When is a WTF an orthonormal basis? Stated differently, what are the conditions on the scaling vector such that the WTF constructed from it forms an ON basis? It is relatively easy to see that *if* the scaling function and its integer translates form an orthonormal system, then the WTF is an ON basis. First notice that

$$\begin{aligned} \int_{\mathrm{R}} \psi_i(t)\psi_j(t-k)\,dt &= \sum_{m,n} h_i(m)h_j(n)\left[M\int_{\mathrm{R}} \psi_0(Mt-m)\psi_0(Mt-n-Mk)\,dt\right] \\ &= \sum_{n} h_i(n+Mk)h_j(n) = \delta(i-j)\delta(k). \end{aligned} \tag{4.33}$$

and therefore $W_{i,0} \perp W_{j,0}$, for $i \neq j$. Moreover $W_{i,0}$ is equipped with ON basis. For any J one readily sees that $W_{i,J} \perp W_{j,J}$, for $i \neq j$ and hence the result follows.

When $M = 2$, Cohen [19] and Lawton [56] have independently obtained characterizations of the scaling vector such that the scaling function and its translates form an orthonormal system. We now extend these results to the multiplicity M case. Let

$$a(n) = \int_{\mathrm{R}} \psi_0(t)\psi_0(t-n)\,dt. \tag{4.34}$$

By taking Fourier transforms on both sides, $a(n) = \delta(n)$ iff

$$\sum_k \left|\hat{\psi}_0(\omega + 2\pi k)\right|^2 = 1. \tag{4.35}$$

Definition 13 A compact set Γ is *congruent* to $[-\pi, \pi](\mathrm{mod} 2\pi)$, if the measure of Γ is 2π and for every point $\omega \in [-\pi, \pi]$, there exists an $n \in \mathbf{Z}$ such that $\omega + 2\pi n \in \Gamma$.

Theorem 28 The following conditions are equivalent:

1. $\psi_0(t)$ and its translates are orthonormal (i.e $a(n) = \delta(n)$).
2. There exists Γ congruent to $[-\pi, \pi]$, containing a neighborhood of zero such for $\omega \in \Gamma$,
$$\hat{\psi}_0(\omega) \geq C > 0 \tag{4.36}$$
3. There exists Γ (as in 2) such that for $\omega \in \Gamma$,
$$\inf_{j>0} \inf_{\omega \in \Gamma} \left| \frac{1}{\sqrt{M}} H_0 \left(\frac{\omega}{M^j} \right) \right| = B > 0. \tag{4.37}$$
4. $a(n) = \delta(n)$ is the unique solution of the equation
$$a(k) = \sum_n a(Mk+n) \left[\sum_m h_0(m) h_0(n+m) \right]. \tag{4.38}$$
5. $A(\omega) = 1$ is the unique solution of the equation
$$A(\omega) = \frac{1}{M} \sum_{k=0}^{M-1} \left| H_i \left(\frac{\omega + 2\pi k}{M} \right) \right|^2 A \left(\frac{\omega + 2\pi k}{M} \right). \tag{4.39}$$
6. There is no non-trivial cycle Π of the map $\omega \mapsto M\omega (\mathrm{mod}\ 2\pi)$, such that $H_0(\omega) = 1$, for all $\omega \in \Pi$.

Proof: 1 implies 2 by exactly the same arguments as the 2-band case in [23, p. 182]. 2 implies 1 follows by using the arguments in [23, p. 184] in conjunction with Eqn. 3.19. 4 and 5 are equivalent via the Fourier transform. 2 implies 3 because for any j, and $\omega \in \Gamma$, from Eqn. 4.36 (since $\left| \frac{H_0(\omega)}{\sqrt{M}} \right| \leq 1$ and $\hat{\psi}_0(\omega) \leq 1$)

$$\left| \frac{1}{\sqrt{M}} H_0 \left(\frac{\omega}{M^j} \right) \right| \geq \prod_{i=1}^{j} \left| \frac{1}{\sqrt{M}} H_0 \left(\frac{\omega}{M^i} \right) \right| = \frac{\hat{\psi}_0(\omega)}{\hat{\psi}_0(\frac{\omega}{M^j})} \geq C > 0.$$

3 implies 2 and can be seen as follows:

$$\left|H_0(\omega) - \sqrt{M}\right| = |H_0(\omega) - H_0(0)| \le \sum_{n=0}^{N-1} |h_0(n)| \left|e^{-\imath\omega n} - 1\right| \le A\,|\omega|$$

for some $A > 0$. Hence for $\omega \in \Gamma$ and k sufficiently large, $\left|\frac{1}{\sqrt{M}} H_0\left(\frac{\omega}{M^k}\right)\right| \ge 1 - A\left|\frac{\omega}{M^k}\right| \ge e^{-A\left|\frac{\omega}{M^j}\right|}$ and therefore

$$\begin{aligned}\left|\hat{\psi}_0(\omega)\right| &\ge B^{k-1} \prod_{j=k}^{\infty} \left[1 - A\left|\frac{\omega}{M^j}\right|\right] \\ &\ge B^{k-1} e^{-A\sum_{j=k}^{\infty}\left|\frac{\omega}{M^j}\right|} \ge C.\end{aligned}$$

4 implies 1 since from Eqn. 4.18

$$a(n) = \int_{\mathbf{R}} \psi_0(t)\psi_0(t+n)\,dt = \sum_n a(Mk+n)\left[\sum_m h_0(m)h_0(n+m)\right],$$

and by hypothesis $a(n) = \delta(n)$ is the only solution.

One proves that 1 implies 5 by contradiction. Let $\{\psi_0(t-k)\}$ be an ON system and let there exist $A(\omega) \ne 1$ that satisfies Eqn. 4.39. We may assume $A(\omega) > 0$ by adding an appropriate constant to it if necessary (Eqn. 4.39 will still be satisfied). Define

$$H_0'(\omega) = H_0(\omega)\sqrt{\frac{A(\omega)}{A(M\omega)}}.$$

Then $H_0'(\omega)$ is also a unitary scaling vector since from Eqn. 4.39

$$\frac{1}{M}\sum_{k=0}^{M-1}\left|H_0'\left(\frac{\omega+2\pi k}{M}\right)\right|^2 = 1.$$

Let $\psi_0'(t)$ be the corresponding scaling function (possibly infinitely supported).

$$\hat{\psi}_0'(\omega) = \prod_1^{\infty}\left[\frac{1}{\sqrt{M}} H_0'\left(\frac{\omega}{M^J}\right)\right] = \hat{\psi}_0(\omega)/\sqrt{A(\omega)}.$$

Since the zero sets of H_0 and H_0' coincide ($A(\omega) > 0$), if $\hat{\psi}_0(\omega)$ is bounded below on a compact set Γ then so is $\hat{\psi}_0{}'(\omega)$. Therefore, $\{\psi_0{}'(t-k)\}$ is also an orthonormal system and

$$1 = \sum_k \left|\hat{\psi}_0'(\omega + 2\pi k)\right|^2 = \sum_k \left|\hat{\psi}_0(\omega + 2\pi k)\right|^2 / A(\omega) = 1.$$

Therefore $A(\omega) = 1$ (recall $A(\omega)$ it is periodic), a contradiction. Equivalence of 5 and 6 can be proved based on ideas for the 2-band case originally developed by Cohen in his PhD thesis [18]. □

The characterizations of orthonormality may be used to show that a particular wavelet basis constructed is orthonormal. Of all the characterizations of orthonormality Eqn. 4.38 is the easiest to verify. It says that given an unitary scaling filter, the corresponding wavelet basis is ON iff the *Lawton matrix* (after Lawton who constructed it for the 2-band case [56]) defined below has a unique eigenvector of eigenvalue 1. If $r(n)$ is the autocorrelation sequence of $h_0(n)$ (of length $N = MK$), Eqn. 4.38 becomes

$$a(n) = \sum_l r(l)a(Mn - l) \tag{4.40}$$

The Lawton matrix Q is defined by

$$q_{i,j} = \begin{cases} r(M(i-1)) \text{ for } j = 1 \\ r(M(i-1)+j-1) + r(M(i-1)-j+1) \text{ for } 2 \le j \le N \end{cases}$$

If $v = \begin{bmatrix} a(0) & a(1)\ldots & a(N-2) \end{bmatrix}^T$, Eqn. 4.40 becomes $Qv = v$. $v = [1, 0, \ldots, 0]^T$ is always an eigenvector of Q with eigenvalue 1. If there exists any other eigenvector for Q, then $\{\psi_0(t-k)\}$ is not an orthonormal system (and the WTF may not be an ON basis).

There is a well-known sufficient condition for orthonormality in the multiplicity 2 case due to Mallat [60] which is easy to verify and stated in terms of $H_0(\omega)$. This condition can be generalized immediately to the multiplicity M and we have the following corollary of Theorem 28. The essential idea in this case one can show $\hat{\psi}_0(\omega)$ does not vanish on the compact set $\Gamma = [-\pi, \pi]$.

Corollary 3 If $H_0(\omega)$ does not vanish for $|\omega| \le \frac{\pi}{M}$, then the wavelet basis generated from it is orthonormal.

Proof: If $H_0(\omega)$ does not vanish on $[-\frac{\pi}{M}, \frac{\pi}{M}]$, then $H_0(\omega/M^j)$ is non-zero on $[-\pi, \pi]$ for all $j \geq 1$. Now from Theorem 15.5 in [74] it follows that $\hat{\psi}_0(\omega)$ (see Eqn. 4.26) is non-zero on $[-\pi, \pi]$. Take $\Gamma = [-\pi, \pi]$ in the previous Theorem to obtain the result. $\square$

4.4 State-Space Approach to Orthonormal Wavelet Bases

The previous section gives a complete parameterization of multiplicity M, compactly supported WTFs, and characterizes the conditions under which a WTF is an orthonormal basis. Checking whether a WTF is an orthonormal basis is equivalent to checking whether the Lawton matrix of the unitary scaling vector has a unique eigenvector of eigenvalue one. An important property of multiplicity 2 wavelets in [21] is that arbitrary large number of moments of the wavelets can be made to vanish. This in turn implies that the corresponding scaling functions can represent polynomials of arbitrarily large degree exactly. This property has found a number of applications in the approximation of operators [39, 5]. The vanishing moments property is intimately related to the smoothness or *regularity* of the wavelet [21]. This section generalizes these concepts to the multiplicity M case. Just as in the multiplicity 2 case, the vanishing of the moments of the wavelets is equivalent to the vanishing of the discrete moments of the wavelet vectors. It is our belief that in most practical applications where regularity has been exploited, it is the vanishing moments property, rather than regularity per se, that is crucial. Therefore, we define regularity to be a property of the scaling vector, rather than the scaling function or wavelets.

For regular WTFs and orthonormal bases, this section gives yet another parameterization, using state-space techniques, of all compactly supported regular WTFs. This does *does not rely on the Householder factorization* of unitary matrices on the unit circle. Independently of this work, multiplicity M compactly supported orthonormal wavelet bases have been obtained by several other researchers [98, 43]. Zou and Tewfik [98], impose regularity on the the Householder parameterization, which results in a set of non-linear equations for the Householder parameters of regular unitary scaling vectors. Heller et. al [43], approach the problem in a manner similar to ours, and does not require the solution of non-linear equations. However, neither approach the parameterization of of wavelets from a state-space point of view.

The two main results in this section are

1. Construction of *regular* multiplicity M scaling vectors and scaling functions
2. A state-space approach to the generation of the wavelet vectors and wavelets.

4.4.1 State-Space Description of Rational Matrices

This section briefly review the state-space description of rational matrices, and gives results relevant to the state-space characterization of wavelets. Let $R_{P\times M}$ denote the set of all real rational matrices (most of the results are true over an arbitrary field). Every function in $R_{P\times M}$ has a Laurent series expansion about any point in $\mathbb{C}$. A function in $H(z) \in R_{P\times M}$ is said to be *proper* if it is analytic at ∞.

$$H(z) = \sum_{k=0}^{\infty} W_k z^{-k}, \tag{4.41}$$

Every *proper* real rational function has a *realization* of the form

$$H(z) = C(zI - A)^{-1}B + D \tag{4.42}$$

where A, B, C and D are matrices of appropriate dimensions with A being a square matrix [49]. Realizations are not unique since $[A, B, C, D]$ and $[TAT^{-1}, TB, CT^{-1}, D]$ (for an arbitrary invertible T) give rise to the same rational matrix $H(z)$. Note that

$$\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} B \\ 0 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix}, D$$

and $[A, B, C, D]$ give rise to the same $H(z)$. A realization is said to be *minimal* if the matrix A has the least possible dimension among all realizations of $H(z)$. Minimal realizations are necessarily related by a non-singular transformation T [49]. The dimension of A in a minimal realization is called the McMillan degree of $H(z)$. $H(z)$ is said to be stable if $\|A\| < 1$, for any realization of $H(z)$.

Let $[A, B, C, D]$ be a minimal realization for a stable $H(z)$ of McMillan degree N. For this state-space representation the *controllability* and *observability* gramians are defined respectively as

$$W_c = \sum_k A^k BB^T (A^T)^k \quad \text{and} W_o = \sum_k (A^T)^k C^T C A^k. \tag{4.43}$$

If $\|A\| < 1$ (H is stable) the series converges and W_c and W_o are well defined are seen to be symmetric. An important characterization of minimal realizations is that $[A, B, C, D]$ is minimal iff W_c and W_o are positive definite [49]. By pre-multiplying and post-multiplying the gramians with A or A^T appropriately, W_o and W_c are seen to satisfy the *Lyapunov* equations,

$$W_c = AW_cA^T + BB^T \tag{4.44}$$

$$W_o = A^T W_o A + C^T C \tag{4.45}$$

Lyapunov's theorem says that if $[A, B, C, D]$ satisfy the Lyapunov equations, for some positive definite W_o and W_c, then $H(z)$ is stable. Notice that the Lyapunov's equations are linear in W_o and W_c and hence, given $[A, B, C, D]$, W_c and W_o can be obtained by solving them (rather than computing the infinite sum in Eqn. 4.43). If $[A, B, C, D]$ and $[A_1, B_1, C_1, D_1]$ represent two minimal realizations of $H(z)$ then $\tilde{W}_c = TW_cT^T$ and $\tilde{W}_o = (T^{-1})^T W_o T^{-1}$. Clearly the gramians are not an invariant of $H(z)$. However the product W_cW_o transforms as $\tilde{W}_c\tilde{W}_o = TW_cW_oT^{-1}$ and hence the eigenvalues of W_cW_o are invariants of $H(z)$. If W_c and W_o are diagonal and equal the realization is said to be *balanced* [64, 69]. Such realizations play an important role in approximating $H(z)$ with another matrix of smaller McMillan degree (using Hankel norm approximations [31]).

A rational function is *left unitary* (on the unit circle) if $H^T(z^{-1})H(z) = I$ and *right unitary* if $H(z)H^T(z^{-1}) = I$. $H(z)$ is *unitary* if it is both left and right unitary. Unitary $H(z)$, as we have already seen, plays an important role in unitary filter bank theory and in the theory of wavelet tight frames. Unitary $H(z)$ has the following characterization [31, 41]:

Fact 8 Given a minimal realization $[A, B, C, D]$ (not necessarily stable), the following statements are equivalent:

1. $H(z)$ is unitary on the unit circle.
2. There exists positive definite W_c, and W_o, with $W_c W_o = I$, that satisfy the Lyapunov equations.

Moreover, given W_o and W_c are as above, D satisfies the equations,

$$D^T D + B^T W_o B = I \tag{4.46a}$$

$$DD^T + CW_c C^T = I \tag{4.46b}$$

$$D^T C + B^T W_o A = 0 \tag{4.46c}$$

$$DB^T + CW_c A^T = 0 \tag{4.46d}$$

As stability is not assumed the gramians in Eqn. 4.43. However, if $H(z)$ is stable W_c and W_o in Fact 8 are precisely the gramians. If we have a balanced realization, then $W_c = W_o = I$, and hence the Lyapunov equations and Eqn. 4.46a-4.46d can be written in the compact form $Y^T Y = YY^T = I$ where

$$Y = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{4.47}$$

A similar result has been reported in [83]. Notice in particular, that if we partition Y as $\begin{bmatrix} Y_1 & Y_2 \end{bmatrix}$, then, just the fact that $W_c = I$, ensures that $Y_1 Y_1^T = I$.

4.4.2 K-regular Multiplicity M Unitary Scaling Vectors

In order to obtain *regular* scaling functions in the multiplicity 2 case, I.Daubechies [21], imposes a set of linear conditions on the scaling vector h_0, which essentially amounts to a certain number of the moments of h_0 being zero. We impose similar restrictions in the multiplicity M scaling vectors, referring to them as *regular* unitary scaling vectors.

Definition 14 A multiplicity M unitary scaling vector is said to be K-*regular* if it has a polynomial factor of the form $\left(\frac{1+z^{-1}+\ldots+z^{-(M-1)}}{M}\right)^K$ for maximal possible K. That is

$$H_0(z) = \left[\frac{1+z^{-1}+\ldots+z^{-(M-1)}}{M}\right]^K Q(z) = \frac{1}{M^K z^{MK-1}}\left[\frac{z^M-1}{z-1}\right]^K Q(z). \tag{4.48}$$

If a scaling vector is K-regular, $H_0(z)$ and its first $(K-1)$ derivatives vanish for $z = e^{\imath 2\pi k/M}, k \in \{1, 2, \ldots, M-1\}$.

$$\left(\frac{d}{dz^{-1}}\right)^j H_0(z)\Bigg]_{z=e^{\imath 2\pi k/M}} = 0 \tag{4.49}$$

This is equivalent to a set of $(M-1)(K-1)$ complex linear constraints on the scaling vector. Since the scaling vector is assumed to have real coefficients the zeros occur in complex conjugate pairs. Hence the set of $(M-1)(K-1)$ complex constraints reduce to $(M-1)(K-1)$ real linear constraints on the scaling vector.

It also follows from Definition 14 that every unitary scaling vector is 1-regular. Indeed from the unitariness condition in the Fourier domain (see Eqn. 3.19), it is clear that $H_0(z)$ vanishes for $z = e^{\imath 2\pi k/M}, k \in \{1, 2, \ldots, M-1\}$. The scaling function and wavelets associated with K-regular scaling vectors will be called K-regular scaling function and wavelets respectively.

K-regularity has a number of equivalent characterizations, each of which shows how regularity plays an important role in applications. K-regularity has been used by Daubechies in the 2-band case to ensure that the scaling vector gives rise to multiplicity 2 ON wavelet basis (not a WTF) [21]. Daubechies also shows that that the regularity of the scaling function (measured by the number of continuous derivatives it has - or equivalently its Hölder exponent) increases linearly with the K, the regularity of the scaling vector. If the scaling function is K times differentiable *it is necessary* that the scaling filter is $(K-1)$-regular. K-regularity is equivalent to saying that all polynomials of degree $(K-1)$ are contained in $W_{0,j}$ for all j. This coupled with the compact support of the scaling functions (and wavelets) implies that K-regular scaling functions can be used to capture local polynomial behavior. This feature of K-regular scaling vectors is particularly useful in image processing applications [97]. K-regularity is also useful in numerical analysis applications [55], where one tries to approximate operators in wavelet bases. In these applications the regularity K of the scaling vector is a measure of the approximation order. From a purely signal processing point of view K-regularity says that the magnitude squared Fourier transform of the scaling vector is flat of order $2K$ at zero frequency.

The moments of h_i and $\psi_i(t)$, and the partial moments of h_0 are defined respectively as follows:

$$\mu(i,k) = \int t^k \psi_i(t)\, dt, \quad m(i,k) = \sum_n n^k h_i(n) \quad \text{and} \quad \eta_{k,l} = \sum_n (Mn+l)^k h_0(Mn+l),$$

so that $m(0,k) = \sum_{l=0}^{M-1} \eta_{k,l}$.

Theorem 29 *(Equivalent Characterizations of K-regularity)* A unitary scaling vector is K-regular iff

1. The frequency response of the scaling vector has a zero of order K at the M^{th} roots of unity.
2. The partial moments up to order K of the scaling vector are equal.
3. The magnitude-squared frequency response of the scaling vector is flat of order $2K$ at $\omega = 0$.

4. All polynomial sequences up to degree $(K-1)$ can be expressed as a linear combination of M-shifts of the scaling vector.
5. All moments up to order $(K-1)$ of the wavelet vectors vanish.
6. All moments up to order $(K-1)$ of the wavelets vanish.
7. Polynomials of degree $(K-1)$ or less are contained in $W_{0,j}$ for all j.

Proof: From Eqn. 4.48

$$H_0(\omega) = e^{-\imath(M-1)K\omega/2}\left(\frac{\sin(M\omega/2)}{\sin(\omega/2)}\right)^K Q(\omega),$$

and therefore for small ω, $H_0\left(\omega + \frac{2\pi k}{M}\right) = O(\omega^K)$, $k \in \{1, 2, \ldots, M-1\}$, implying that the derivatives up to order $(K-1)$ vanish at the roots of unity. Equivalently for $k \in \{1, \ldots, M-1\}$, $i \in \{0, \ldots, K-1\}$

$$\begin{aligned}
0 = \left[\frac{d^i}{d\omega^i} H_0(\omega)\right]_{\omega=2\pi k/M} &= \sum_n (-\imath n)^i\, h_0(n) e^{-\imath\frac{2\pi k n}{M}} \\
\Rightarrow \sum_{l=0}^{M-1}\left[\sum_n h_0(Mn+l)(Mn+l)^i\right] e^{-\imath\frac{2\pi k l}{M}} &= \sum_{l=0}^{M-1} \eta_{i,l} e^{-\imath\frac{2\pi k l}{M}} = 0 \\
\Rightarrow \eta_{i,l} \quad &\text{is a constant independent of } l.
\end{aligned}$$

Because of the unitariness of h_0 one also has for small ω

$$|H_0(\omega)|^2 = M - \sum_{k=1}^{M-1}\left|H_0(\frac{\omega + 2\pi k}{M})\right|^2 = M - O(\omega^{2K}).$$

It follows immediately that for $k \leq K-1$, one can express n^k as a linear combination of $h_0(Ml+n)$:

$$n^k = \sum_l \alpha_{i,l} h_0(Ml+n).$$

As a consequence of this representation, the moments of the wavelet vectors vanish up to order $(K-1)$ since

$$m(i,k) = \sum_n n^k h_i(n) = \sum_l \alpha_{i,l}\left[\sum_n h_i(n) h_0(Ml+n)\right] = 0.$$

Now this implies that $\mu(i,k)=0$ since (from Eqn. 4.20) they are related to $m(i,k)$ as

$$\mu(i,k)=\frac{1}{M^{k+\frac{1}{2}}}\sum_{j=0}^{k}\binom{k}{j}m(i,j)\mu(0,k-j).$$

Since the wavelets are compactly supported, the form a basis for $L^2_{\text{loc}}(\mathbb{R})$ and therefore for $k\in\{0,\ldots,K-1\}$,

$$\begin{aligned} t^k &= \sum_k \left\langle t^k,\psi_{0,J,k}(t)\right\rangle \psi_{0,J,k}(t)+\sum_{i=1}^{M-1}\sum_{j=J+1}^{\infty}\sum_{k=-\infty}^{\infty}\langle f,\psi_{i,j,k}(t)\rangle\,\psi_{i,j,k}(t) \\ \Rightarrow\quad t^k &= \sum_k \left\langle t^k,\psi_{0,J,k}(t)\right\rangle \psi_{0,J,k}(t). \end{aligned}$$

Therefore polynomials of degree $(K-1)$ can be effectively expressed as linear combinations of $\{\psi_{0,j,k}\}$ for fixed j (one might loosely say $W_{0,j}$ contains polynomials of degree $(K-1)$, even though polynomials are not in $L^2(\mathbb{R})\supset W_{0,j}$). □

4.4.3 K-regularity and Regularity of Scaling Functions/Wavelets

The precise relationship between K-regularity of the scaling vector and the smoothess of the scaling functions and wavelets is unknown even in the 2-band case. However, using the techniques in [21] it is easy to show that if $Q(\omega)$ is bounded above by an appropriate constant, then the regularity of the scaling function can be estimated. It can be shown that (see [78] for details)

$$\left|\hat{\psi}_0(\omega)\right|\leq C\left[1+|\omega|\right]^{\log_M \sup_\omega Q(\omega)-K-\frac{1}{2}}. \tag{4.50}$$

Therefore if $\sup_\omega Q(\omega)<M^{K-m-\frac{1}{2}}$, then $\psi_0(t)$ associated with a K-regular scaling vector is m times differentiable. However, since $Q\left(\omega\right)_{\omega=0}=\sqrt{M}$, $\psi_0(t)$ can be at most $(K-2)$ times differentiable. The sufficient condition for $\psi_0(t)$ to be m times differentiable is precisely given below [24]: