[43] P. P. Vaidyanathan. Quadrature mirror filter banks, m-band extensions and perfect-reconstruction techniques. *IEEE ASSP Magazine*, pages 4–20, 1987.

[44] P. P. Vaidyanathan. Improved Technique for The Design Of Perfect Reconstruction FIR QMF Banks With Lossless Polyphase Matrices. *IEEE Trans. in ASSP*, 37(7):1042–1056, 1989.

[45] P. P. Vaidyanathan and Z. Doganata. The Role Of Lossless Systems in Modern Digital Signal Processing. *IEEE Trans. on Education*, 32:181–197, August 1989.

[46] P. P. Vaidyanathan and Phuong-Quan Hoang. Lattice Structures for Optimal Design and Robust Implementation of Two-Channel Perfect Reconstruction Qmf banks. *IEEE Trans. in ASSP*, 36(1):81–93, 1988.

[47] M. Vetterli and C. Herley. Wavelets And Filter Banks : Theory And Design. *IEEE, Trans. in ASSP*, pages 2207–2232, Sept 1992.

[48] G. Wornell. Karunhen-Loéve Like Expansion for $1/f$ Noise Processes. *IEEE Trans. in ASSP*, 1991.

[23] K. Hoffman. *Banach Spaces of Analytic Functions.* Dover Publications. Inc., 1962.

[24] L. Hormander. *Distribution Theory and Fourier Analysis.* Springer-Verlag, 1990.

[25] I.Daubechies and J. C. Lagarias. Two-Scale Difference Equations : Local Regularity, Infinite Products Of Matrices And Fractals. *SIAM Journal of Math. Anal.*, 24.

[26] T. Kailath. *Linear Systems.* Prentice Hall, Englewood Cliffs, New Jersey, 1980.

[27] W. M. Lawton. Private communication.

[28] W. M. Lawton. Necessary And Sufficient Conditions for Existence Of on Wavelet Bases. *Journal of Math. Physics*, 1990.

[29] W. M. Lawton. Tight Frames Of Compactly Supported Wavelets. *Journal of Mathematical Physics*, 1990.

[30] S. Mallat. Multiresolution approximations and wavelets. Technical report, GRASP Lab, Dept. of Computer and Information Science, University of Pennsylvania, 1987.

[31] S. Mallat. A theory for Multiscale Signal Decomposition : The Wavelet Representation. *IEEE Trans. on Pattern and Machine Intelligence*, 11(7):674–693, 1989.

[32] Y. Meyer. Principe d'incértitude, bases hilbertiennes et algèbres d'opérateurs. *Séminaire Bourbaki*, (662), 1985-1986.

[33] T. Q. Nguyen and P. P. Vaidyanathan. Maximally Decimated Perfect-Reconstruction FIR Filter Banks with Pairwise Mirror-Image Analysis and Synthesis Frequency Responses. *IEEE Trans. in ASSP*, 36(5):693–706, 1988.

[34] A. Oppenheim and R.W.Schafer. *Discrete-Time Signal Processing.* Prentice Hall,, 1989.

[35] T. W. Parks and C. S. Burrus. *Digital Filter Design.* John Wiley and Sons. Inc., 1987.

[36] D. Pollen. $SU_I(z, 1/z)$ for f a subfield of $\mathbb{C}$. *Journal of the Amer. Math. Soc.*, 1990.

[37] M. R. Portnoff. Time frequency distributions of digital systems and signals based on short-time fourier analysis. *IEEE Trans. in ASSP*, ASSP-28:55–69, 1980.

[38] L. Rabiner and D. Crochière. *Multirate Digital Signal Processing.* Prentice-Hall, 1983.

[39] H. L. Resnikoff. Weierstrass functions and compactly supported wavelets. Technical Report AD900810, Aware Inc., Cambridge, MA-02138, 1990.

[40] H.K.Dunn R.Koenig and L .Y. Lacy. The Sound Spectogram. *Journal of the Acoustic Society of America*, 18:19–49, 1946.

[41] W. Rudin. *Real and Complex Analysis.* McGraw-Hill., 3 edition, 1987.

[42] M. J. Smith and T. P. Barnwell. Exact Reconstruction techniques for tree-structured subband coders. *IEEE Trans. in ASSP*, 34:434–441, 1986.

[6] R. Coifman, Y. Meyer, S. Quake, and M. V. Wickerhause. Signal processing and compression with wave packets. preprint.

[7] I. Daubechies. Orthonormal Bases Of Compactly Supported Wavelets. *Communications in Pure and Applied Mathematics*, (41):909–996, 1988.

[8] I. Daubechies. The Wavelet Transform, Time-Frequency Localization and Signal Analysis, 1990.

[9] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992. CBMS Lecture Series.

[10] I. Daubechies, A. Grossman, and Y. Meyer. Painless non-orthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283, 1986.

[11] I. Daubechies and J. C. Lagarias. Two-Scale Difference Equations : Existence And Global Regularity Of Solutions. *SIAM Journal of Math. Anal.*, 22:1388–1410.

[12] D. Esteban and C. Galand. Applications Of Quadrature Mirror Filters To Split Band Voice Coding Schemes. In *Proc. of ICASSP*, pages 191–195, 1977.

[13] J. Flanagan. *Speech Analysis and Synthesis Perception*. Springer-Verlag, New York, 1972.

[14] B. Francis. *A Course in H^∞ Control Theory*. Springer-Verlag, 1987. Lecture Notes in Control And Information Sciences.

[15] B. Friedlander and B. Porat. Detection of transient signals by the Gabor representation. *IEEE Trans. on ASSP*, 37(2), 1989.

[16] Carl-Erik Froberg. *Numerical Mathematics*. Benjamin/Cummings Publishing Co., 1985.

[17] D. Gabor. Theory of communication. *Journal of the Institute for Electrical Engineers*, 93:429–439, 1946.

[18] R. A. Gopinath. Moment Theorem for the Wavelet-Galerkin Method. Technical report, Aware Inc., Cambridge, MA-02138, Cambridge, MA-02138, 1990.

[19] R. A. Gopinath. Wavelet Transforms and Time-Scale Analysis of Signals. Master's thesis, Rice University, Houston, TX-77251, 1990.

[20] R. A. Gopinath and C. S. Burrus. State-space Approach to Multiplicity M Orthonormal Wavelet Bases. Technical Report CML TR91-22, Computational Mathematics Laboratory, Rice University, 1991. Submitted to IEEE Trans. in SP.

[21] R. A. Gopinath, W. M. Lawton, and C. S. Burrus. Wavelet-Galerkin Approximation of Linear Translation Invariant Operators. In *Proceedings of the ICASSP '91, Toronto, Canada*, volume 3, pages 2021–2023. IEEE, May 1991.

[22] A. Grossman and J. Morlet. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM Journal of Mathematical Analysis*, (15):723–736, 1984.

9 Summary

The concept of frequency decompositions gives was used to introduce short-time Fourier analysis and wavelet analysis in a common framework. DWT type frequency decompositions were used to illustrate how naturally filter banks arise in such frequency decompositions. We then gave a concise account of filter bank theory as it relates to wavelet theory. Using results in filter bank theory (purely algebraic) and some analytical constraints on the filter banks, we showed that multiplicity M wavelet frames and wavelet tight frames can be obtained. We showed how multiresolution analyses arise naturally in the context of wavelet frames. Though precise conditions, for wavelet frames to be bases are known, we did not discuss it since most wavelet frames are bases. Since all the computations in wavelet analysis are done in efficient structures for filter banks, we discussed the polyphase approach to general filter banks and the cascade approach to unitary filter banks. We this discussed computational aspects in basic analysis of signals in wavelet frames. Finally we explored the powerful algebraic aspects of filter banks, and how they induce decompositions of abstract separable Hilbert spaces and introduce a notion of *scale* in these abstract spaces. This was used to illustrate how most of the approximate practical computations in *wavelet* bases, turn out to be *exact* calculations in some *smooth* bases, when suitably interpreted. This meant that the regularity of the wavelets and scaling functions are probably not critical for practical applications, contrary to the popular belief. However, the regularity of the scaling vectors and the wavelet vectors, may be important for some applications, and in those cases corresponding scaling function and wavelets turn out to be regular.

Acknowledgments

The authors would like to thank colleagues at Aware.Inc. Cambridge, MA, and members of the Computational Mathematics Laboratory, Rice University, for helpful discussions. Comments and suggestions from Howard Resnikoff and Wayne Lawton in particular at Aware, and Raymond Wells and Jan Erik Odegard in particular at CML, is appreciated.

References

[1] V. Belevitch. *Classical Network Synthesis.* Holden-Day, San Francisco, CA, 1968.

[2] C. K. Chui. Wavelets and spline interpolation. Technical Report CAT-227, Texas A&M University, Center for Approximation Theory, Dept. of Math., Texas A&M, College Station, TX-77843, 1990.

[3] C. K. Chui and J. Z. Wang. On Compactly Supported Spline Wavelets And A Duality Principle. Technical Report CAT-213, Texas A&M University, Center for Approximation Theory, Dept. of Math., Texas A&M, College Station, TX-77843, May 1990.

[4] A. Cohen, I.Daubechies, and J.C. Feauveau. Biorthogonal Bases of Compactly Supported Wavelets. Preprint. To Appear in Comm. in Pure and Applied Math.

[5] L. Cohen. Time-Frequency Distributions - A Review. *Proceedings of the IEEE*, 77(7):941–981, 1989.

of a scale. Given $\mathcal{H}$, all the subspaces obtained by decomposing it with an M channel filter bank are said to be at scale $\frac{1}{M}$. Notice that scale is also a measure of the number of basis elements in $\mathcal{H}_i$ relative to $\mathcal{H}$. And if we recombine then the scale of the resulting space is M times the scale of the constituent spaces. However, we have not been able to come up with a clean rigorous definition of *scale* for these decompositions. The full impact of these general decomposition/recomposition ideas is a subject of great interest. In any case, using the ideas described above we obtain the following result.

Theorem 4 *For every finite rational partition of the the interval* $[0, 1)$ *there is a decomposition of every separable Hilbert space into subspaces whose scales correspond to partitions of the interval.*

Proof : Obvious □

We now use these results to discuss some of the issues concerning the regularity of wavelets in practical computations. It is often assumed, that wavelets have introduced, an underlying set of continuous functions to filter bank analysis. Be that as it may, it is important to realize that digital signal processing had its roots in the Shannon sampling theorem, that allows one to go from continuous time to discrete time. So underlying continuous functions for *analog* interpretation of filter banks is nothing new. Now consider the analysis of a signal wrt an orthonormal wavelet basis $\psi_{i,j,k}(t)$. Assume further that the samples (usually finite) of the signal f are given. Usually, the samples are *assumed* to be $Wf(0, J_f, k)$ and the numbers $Wf(i, j, k)$, for $J_c \leq j < J_f$ are then computed using a filter bank. Notice this is an approximate method to compute the DWT. We now try to give an *exact* interpretation of this process.

The idea is to interpret the samples $f(\frac{n}{T})$ as the the Nyquist rate samples of a bandlimited signal. These turn out to be the expansion coefficients with respect to the *sinc* basis. Notice that the sinc function is infinitely differentiable. There is no problem with the fact that there are only finitely many samples of f, since bandlimited signals could have finitely many Nyquist samples. Let $\mathcal{H}$ denote the space of bandlimited signals in $[0, \pi/T)$. Then $f \in \mathcal{H}$. By the decomposition lemmas discussed earlier, what we are doing is decomposing $\mathcal{H}$, recursively, and the numbers $W(i, j, k)$ obtained are precisely, the expansion coefficients with respect to the bases in the decomposition subspaces. Since all the basis function at any level are linear combinations (finite in the FIR case) of the sinc function, they are infinitely smooth, regardless, of whether the filters in the filter bank give rise to *regular* wavelets. Hence, even assuming we have a filter bank with tight frame wavelets, there is no reason to impose constraints on the sequences h_i so that they give smooth wavelets, because, all practical computations, can be interpreted as being, done with some other smooth basis. The regularity of the filters are important *only* if we are going to have an infinite cascade of filter banks ($J_c \to \infty$), a practical impossibility.

The important implication of the above argument is that, for best frequency localization, the filters $h_i(n)$, need not be the ones that give rise to wavelet bases. Does this mean that, it is unimportant to construct regular wavelet bases? The answer is no. In numerical analysis applications (like the approximation of differential operators), the vanishing moments property of the scaling vector (which causes smoothness of the corresponding $\psi_0(t)$) is important [21]. However, in signal processing applications, usually regularity is unimportant.

Proof : Let $G(z) = G_p(z)$ for some unitary filter bank. Let $H^T = G^{-1}$, and $H_p = H$. Then (H_p, G_p) form a filter bank with filters h_i and g_i.

$$\phi_k = \sum_i \sum_n g_i(k - Mn)\psi_{i,n} \tag{207}$$

$$\tilde{\phi}_k = \sum_i \sum_n h_i(-k + Mn)\tilde{\psi}_{i,n} \tag{208}$$

Then from Lemma. 8

$$\begin{aligned}
\sum_k \phi_k >< \tilde{\phi}_k &= \sum_i \sum_l h_i(-k + Ml)\psi_{i,l} >< \sum_j \sum_m g_i(k - Mm)\tilde{\psi}_{i,m} \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \tilde{\psi}_{i,m} \left[\sum_k h_i(-k + Ml) g_j(k - Mm) \right] \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \tilde{\psi}_{i,m} \left[\sum_k h_i(k) g_j(-k + M(l - m)) \right] \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \tilde{\psi}_{i,m} \left[\delta(i - j)\delta(l - m)) \right] \\
&= \sum_{i,l} \psi_{i,l} >< \tilde{\psi}_{i,l}
\end{aligned} \tag{209}$$

Again, if for all i, $\{\psi_{i,k}\}$ forms a biorthogonal basis then $\{\phi_n\}$ forms a biorthogonal basis from Lemma. 10 □

Remark: In all of the above decompsitions and recompositions discussed, if one has to go from the expansion coefficients in an old basis to the expansion coefficients in the new basis obtained as above, then depending on whether it is decomposition or recomposition, one has to merely run the data through the the corresponding analysis or synthesis bank. Moreover, from the theory discussed before, for tight frames case, the unitary representation is the computationally optimal, while for the frames case, the general polyphase is the best known computational technique.

Let us assume we start with a Hilbert space and decompose it into M subspaces using an filter bank. Now each of the subspaces could themselves be decomposed into a different (not M) number of subspaces using some other filter bank and so on. Since the filter banks could all be different and of different multiplicities we have a rich class of decompositions for the Hilbert space. Moreover, some of the subspaces could also be recombined using filter banks again (as allowed by the recomposition lemmas). Notice also, that the filter banks used for the recomposition may have nothing to do with filter banks used for decomposition.

At this point, there is the need for the concept of a *scale* for these decompositions. To illustrate how it occurs, let us decompose $\mathcal{H}$ into 3 subspaces using a filter bank. Now if we recombine $\mathcal{H}_1$ and $\mathcal{H}_2$ to give $\mathcal{G}$ using some other filter bank, effectively $\mathcal{H}$ has been decomposed into 2 subspaces. That is,

$$\mathcal{H} = \mathcal{G} \oplus \mathcal{H}_3 \tag{210}$$

Clearly this decomposition of $\mathcal{H}$ is quite different from any decomposition obtained with a 2 channel filter bank. In order to distinguish between these decompositions we have to introduce the concept

□ Notice how in both these proofs the dual lemmas, Lemma. 8 and Lemma. 10, play a role, one to to show that we have a frame and the other (the dual) to show that if we have a basis to start with then the decomposition also has a basis.

We now discuss the corresponding recomposition lemmas for filter banks.

Lemma 14 (Tight Frames Recomposition Lemma) *Let $G \in U_{M\times M}$. Then, for a collection of M separable Hilbert spaces $\mathcal{H}_i$, equipped with a orthonormal bases (tight frames) $\{\psi_{i,k}\}$, respectively, G induces a natural orthonormal basis (tight frame) ϕ_k for $\mathcal{H}$ the direct sum of these separable Hilbert spaces. That is,*

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \ldots \oplus \mathcal{H}_{M-1} \tag{203}$$

and ϕ_k is an orthonormal basis (tight frame) for $\mathcal{H}$.

Proof : Take $G(z) = G_p(z)$ for some unitary filter bank. Let $g_i(n)$ denote the corresponding filters in the filter bank. Define,

$$\phi_k = \sum_i \sum_n g_i(k - Mn)\psi_{i,n} \tag{204}$$

Then from Lemma. 7

$$\begin{aligned}
\sum_k \phi_k >< \phi_k &= \sum_i \sum_l g_i(k - Ml)\psi_{i,l} >< \sum_j \sum_m g_j(k - Mm)\psi_{i,m} \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \psi_{i,m} \left[\sum_k g_i(k - Ml)g_j(k - Mm)\right] \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \psi_{i,m} \left[\sum_k g_i(k)g_j(k - M(l - m))\right] \\
&= \sum_{i,j,l,m} \psi_{i,l} >< \psi_{i,m}[\delta(i - j)\delta(l - m))] \\
&= \sum_{i,l} \psi_{i,l} >< \psi_{i,l}.
\end{aligned} \tag{205}$$

Using Lemma. 9 we can now show that, if, for all i $\{\psi_{i,k}\}$ is an orthonormal basis for $\mathcal{H}_i$, then $\{\phi_n\}$ is an orthonormal basis for $\mathcal{H}$. □

Lemma 15 (Frames Recomposition Lemma) *Let $G \in GL^{\infty}_{M\times M}$. Then, for a collection of M separable Hilbert spaces $\mathcal{H}_i$, equipped with a biorthogonal bases (frames), $\{\psi_{i,k}\}$ and dual bases (frames) $\left\{\tilde{\psi}_{i,k}\right\}$, respectively, G induces a natural biorthogonal basis (frame) ϕ_k with dual $\tilde{\phi}_k$ for $\mathcal{H}$ the (external) direct sum of these separable Hilbert spaces. That is, if*

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \ldots \oplus \mathcal{H}_{M-1} \tag{206}$$

then, ϕ_n and $\tilde{\phi}_n$ form dual frames $\mathcal{H}$.

assume that we have a multiplicity M wavelet tight frame. Then the space $V_{0,j}$ associated with it could be decomposed using any unitary matrix H not necessarily, the one from which $V_{0,j}$ was constructed. In particular, if the unitary matrix is the one from which $V_{0,j}$ is constructed we have a wavelet decomposition.

Lemma 13 (Frames Decomposition Lemma) *Let $H(z) \in GL^{\infty}_{M\times M}$. For a separable Hilbert space $\mathcal{H}$ with a biorthogonal basis (frame) $\{\phi_k\}$ and dual basis (frame) $\left\{\tilde{\phi}_k\right\}$, H induces two (dual) frame decompositions*

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \ldots \oplus \mathcal{H}_{M-1} \tag{197}$$

and

$$\mathcal{H} = \tilde{\mathcal{H}}_0 \oplus \tilde{\mathcal{H}}_1 \ldots \tilde{\mathcal{H}}_{M-1} \tag{198}$$

Proof: By assumption, H is invertible on the unit circle. Let $G^T = H^{-1}$. Then, $G^T H = HG^T = I$. Consider an filter bank with $H_p = H$ and $G_p = G$. Define

$$\psi_{i,k} = \sum_n h_i(Mk - n)\phi_n \quad ; \tilde{\psi}_{i,k} = \sum_n g_i(-Mk + n)\tilde{\phi}_n \tag{199}$$

Let $\mathcal{H}_i = Span\{\psi_{i,k}\}$ and let $\tilde{\mathcal{H}}_i = Span\{\tilde{\psi}_{i,k}\}$. Then, $\mathcal{H}_i$ and $\tilde{\mathcal{H}}_i$ give the desired decompositions. For $f \in \mathcal{H}$ consider the following operator whose range is in $\mathcal{H}_i$

$$I_i = \sum_k \psi_{i,k} >< \tilde{\psi}_{i,k}. \tag{200}$$

By application of Lemma. 10

$$\begin{aligned}
\sum_{i=0}^{M-1} I_i &= \sum_{i=0}^{M-1} \sum_k \psi_{i,k} >< \tilde{\psi}_{i,k} \\
&= \sum_{i=0}^{M-1} \sum_k \sum_l h_i(Mk - l)\phi_l >< \sum_m g_i(-Mk + m)\tilde{\phi}_m \\
&= \sum_{l,m} \phi_l >< \tilde{\phi}_m \left[\delta(l - m)\right] \\
&= \sum_k \phi_k >< \tilde{\phi}_k = I,
\end{aligned} \tag{201}$$

the identity operator in $\mathcal{H}$ by assumption. If moreover, $\{\phi_n\}$ and $\left\{\tilde{\phi}_n\right\}$ forms a biorthogonal basis for $\mathcal{H}$, then $\{\psi_{i,k}\}$ and $\left\{\tilde{\psi}_{i,k}\right\}$ forms a biorthogonal basis for $\mathcal{H}_i$. This follows from Lemma. 8, since

$$\begin{aligned}
\left\langle \psi_{i,k}, \tilde{\psi}_{i,l} \right\rangle &= \left\langle \sum_m h_i(Mk - m)\phi_m, \sum_n g_i(-Ml + nn)\tilde{\phi}_n \right\rangle \\
&= \sum_m \sum_n h_i(Mk - m)g_i(-Ml + n)\left\langle \phi_m, \tilde{\phi}_n \right\rangle \\
&= \sum_n h_i(Mk - n)g_i(-Ml + n) \\
&= \sum_n h_i(n)g_i(M(k - l) - n) = \delta(k - l)
\end{aligned} \tag{202}$$

we show that filter banks not only induce decompositions of the spaces $V_{0,j}$ associated with the scaling function, but more generally all separable Hilbert spaces. Recall that when frequency decompositions were introduced in Section. 2, it was mentioned that frequency decompositions can be thought of as a special case of the decomposition of separable a Hilbert space. Filter banks give one such structured, efficient decomposition. Furthermore, filter bank decompositions give a natural concept of *scale* to Hilbert spaces that are so decomposed.

Lemma 12 (Tight Frames Decomposition Lemma) *Let $H \in U_{M\times M}$. For a separable Hilbert space $\mathcal{H}$ equipped with a orthonormal basis (tight frame) $\{\phi_n\}$, H induces a natural tight frame decomposition*

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \ldots \oplus \mathcal{H}_{M-1} \tag{192}$$

Proof : Take $H(z) = H_p(z)$ for some unitary filter bank. Let $h_i(n)$ denote the corresponding filters in the filter bank. Define

$$\psi_{i,k} = \sum_n h_i(Mk - n)\phi_n \tag{193}$$

and let $\mathcal{H}_i = Span\{\psi_{i,k}\}$. Then, $\mathcal{H}_i$ is the desired decomposition. Let I_i be operator from $\mathcal{H}$ into $\mathcal{H}_i$ defined by

$$I_i = \sum_k \psi_{i,k} >< \psi_{i,k}. \tag{194}$$

Then by application of Lemma. 9

$$\begin{aligned}
\sum_{i=0}^{M-1} I_i &= \sum_{i=0}^{M-1}\sum_k \psi_{i,k} >< \psi_{i,k} \\
&= \sum_{i=0}^{M-1}\sum_k\sum_l h_i(Mk-l)\phi_l >< \sum_m h_i(Mk-m)\phi_m \\
&= \sum_{l,m} \phi_l >< \phi_m \left[\delta(l-m)\right] \\
&= \sum_k \phi_k >< \phi_k = I,
\end{aligned} \tag{195}$$

the identity operator on $\mathcal{H}$. Furthermore, if $\{\phi_n\}$ forms an orthonormal basis for $\mathcal{H}$, then $\{\psi_{i,k}\}$ forms an orthonormal basis for $\mathcal{H}_i$. To see this note using Lemma. 7 that,

$$\begin{aligned}
\langle \psi_{i,k}, \psi_{i,l} \rangle &= \left\langle \sum_m h_i(Mk-m)\phi_m, \sum_n h_i(Ml-n)\phi_n \right\rangle \\
&= \sum_m\sum_n h_i(Mk-m)h_i(Ml-n)\left\langle \phi_m, \phi_n \right\rangle \\
&= \sum_n h_i(Mk-n)h_i(Ml-n) = \delta(k-l)
\end{aligned} \tag{196}$$

□

The Lemma is extremely general and depends only on the fact that we have a unitary matrix H (there are no "DC" conditions that were required to construct wavelet tight frames). Let us

and therefore

$$Wf(i, j-1, n) = \sum_k h_i(k) Wf(0, j, Mn+k) \tag{187}$$

and this equation can be implemented with a an analysis filter bank where the analysis filters are the time reversals of $h_i(n)$. The input to the filter bank is $W(0, j, k)$ and the output is the desired numbers.

7.2.3 Synthesis

We are given $Wf(i, j-1, k)$ and we want $Wf(0, j, k)$. Firstly it should be noted that,

$$\psi_{0,j,n} = \sum_{i=0}^{M-1} \sum_k g_i(Mk-n) \psi_{i,j-1,k}. \tag{188}$$

The result can be seen as follows:

$$\begin{aligned}
\sum_{i=0}^{M-1} \sum_k g_i(Mk-n)\psi_{i,j-1,k} &= \sum_{i=0}^{M-1} \sum_k g_i(Mk-n) \sum_l h_i(l) \psi_{0,j,Mk+l} \\
&= \sum_{i=0}^{M-1} \sum_k g_i(Mk-n) \sum_l h_i(-Mk+l) \psi_{0,j,l} \\
&= \sum_l \psi_{0,j,l} \left[\sum_{i=0}^{M-1} \sum_k g_i(Mk-n) h_i(-Mk+l) \right] \\
&= \sum_l \psi_{0,j,l} \delta(l-n) \\
&= \psi_{0,j,n}
\end{aligned} \tag{189}$$

Therefore we have,

$$Wf(0, j, n) = \sum_{i=0}^{M-1} \sum_k \sum_k g_i(Mk-n) Wf(i, j-1, k) \tag{190}$$

and this is just the synthesis equations of a filter bank with the the filters $g_i(n)$ time-reversed.

In summary, for a wavelet frame constructed from a filter bank, the computations are done in another filter bank, whose analysis and synthesis filters are time-reverses of the filters in the original filter bank. If this situation is not aesthetically appealing, then, the following trick fixes the problem : If we define instead

$$Wf(i, j, k) = \langle f, \psi_{i,j,-k} \rangle \tag{191}$$

then the filter bank used in computations is precisely the original filter bank.

8 Hilbert Space Decompositions

We have seen that wavelet frames and tight frames can be constructed from filter banks and consequently wavelet analysis can be done efficiently using discrete representations. In this section

7.2.1 The Projection onto the Finest Scale

In order to compute the DWT, the most efficient way is to start with $W(0, J_f, k)$ and then use the filter bank to compute the DWT. So the first step is the projection onto the finest scale V_{0,J_f}. In most applications $Wf(0, J_f, k)$ cannot be computed exactly, partly because the inner-products involved can seldom be computed analytically and partly because, only the samples of $f(t)$ may be available.

From the samples, one can use polynomial, spline or any other interpolation technique and then compute $Wf(0, J_f, k)$ by numerical quadrature. This procedure is greatly simplified by the fact that the moments of the function $\psi_0(t)$ can be *exactly* computed and hence numerical quadrature reduces to a discrete convolution [19, 18]. Let $\mu(i, k)$ and $m(i, k)$ denote the the k^{th} moment of $\psi_i(t)$ and $h_i(n)$ respectively for all i.

$$\mu(i,k) = \int t^k \psi_i(t)\, dt \tag{182}$$

$$m(i,k) = \sum_n n^k h_i(n) \tag{183}$$

Then from Eqn. 134 and Eqn. 135, it can be shown that

$$\mu(i,k) = \frac{1}{M^{k+\frac{1}{2}}} \sum_{j=0}^{k} \binom{k}{j} m(i,j)\mu(0,k-j). \tag{184}$$

Thus the moments of all the functions $\psi_i(t)$, can be computed from the moments of the sequences $h_i(n)$ associated with them.

The fact that $\mu(0, k)$ can be exactly computed, and that $\psi_0(t)$ is compactly supported, can be combined to give a nice technique to compute $W(0, J_f, k)$ from the samples of f. For every k, if a local polynomial approximation of $f(t)$ around the point $M^{-J_f}k$ is taken in order to compute $W(0, J_f, k)$, then it can be shown that there exists a sequence $e(k)$, that depends on the order of the approximation, such that

$$W(0, J_f, k) = e(k) * f(M^{-J_f}k) \tag{185}$$

Perhaps the most common approach to assume that $f(M^{-J_f}k)$ is given and take $Wf(0, J_f, k) \approx f(M^{-J_f}k)$. The argument for this is that since $\psi_0(t)$ is compactly supported and $\int \psi(t)\, dt = 1$, for large j, $\psi_{0,j,k}$ approaches the Dirac delta measure at $M^{-j}k$. In practical cases, especially with wavelets that are reasonably regular (like the Daubechies' wavelets in [7]), this approximation gives very good results [27, 18] because of certain interesting relationships between the the moments $\mu(0, k)$.

7.2.2 Analysis

We are given $Wf(0, j, k)$ and we want to compute the numbers $Wf(i, j-1, k)$. From Eqn. 145 we have

$$\psi_{i,j-1,n} = \sum_k h_i(k)\psi_{0,j,Mn+k} \tag{186}$$

7.1.2 Infinite Product Method

This method to compute the samples of ψ_i is approximate and works best for for M-adics of the form $\frac{i}{M^J}$ where J is approximately 8 to 10. This technique, essentially due to I.Daubechies [7], is based on the fact that ψ_0^j in the proof of Theorem. 2, is always a staircase function, whose heights can be directly computed using the filter bank associated with the wavelet. These heights are taken to be the approximate samples of ψ_0. A very good approximation is obtained after about 8 or 9 iterations i.e the staircase approximation corresponding to ψ_0^8 or ψ_0^9. From the samples of the scaling function the samples of the wavelet is then then easily obtained.

In the first method the values of ψ_0 at the integers are first obtained exactly. Hence the values at the M-adic rationals are also obtained exactly within limits of numerical errors. In the second method since an infinite product is truncated only an approximation of ψ_i is obtained. An advantage with the infinite product method is that there is no need to solve a system of linear equations. Both techniques are identical in their implementation in that the interpolation starts with the values at the integers and goes through a process while the infinite convolution method starts off with a delta and goes through the same process. Eventually they both converge to the same values.

7.2 Analysis/Synthesis in Wavelet Bases

This section deals with the computational aspects related to analysis and synthesis of signals in wavelet bases. The *wavelet* expansion of a signal with respect to a wavelet frame $\psi_{i,j,k}$ is given by

$$f(t) = \sum_{i=1}^{M-1} \sum_{j,k} \langle \psi_{i,j,k}(t), f \rangle \, \tilde{\psi}_{i,j,k} \tag{178}$$

The DWT of the function with respect to this frame is defined by

$$Wf(i,j,k) = \langle \psi_{i,j,k}(t), f \rangle \tag{179}$$

where $i = 1, 2, \ldots, M-1$. For computational purposes we define the Discrete Scaling Transform by

$$Wf(0,j,k) = \langle \psi_{0,j,k}(t), f \rangle \tag{180}$$

Clearly for any fixed j the DST contains all the information in $V_{0,j}$.

Since only finite sequences can be stored and manipulated, both j and k have to be restricted to be finite. Since most signals are of finite duration it is natural to assume that $f(t)$ is compactly supported, and then if $\psi_i(t)$ are compactly supported, $Wf(i,j,k)$ vanishes outside a finite set.

Since $V_{0,j} \rightarrow L^2(\mathbb{R})$, any $f \in L^2(\mathbb{R})$ can be approximated to any desired accuracy by a function in $V_{0,j}$ for large enough j, say J_f (finest scale). Therefore the DWT is need not be computed for $j \leq J_f$. Similarly some coarsest scale of interest J_c is also determined. Information in coarser scales is retained in the DST at that scale. That is,

$$f(t) = \sum_k Wf(0, J_c, k) \tilde{\psi}_{0,J_c,k}(t) + \sum_{i=1}^{M-1} \sum_{j=J_c}^{J_f} \sum_k Wf(i,j,k) \tilde{\psi}_{i,j,k} \tag{181}$$

7.1.1 Interpolation Method

In this method first $\psi_0(n)$ is obtained by solving the following eigenvalue problem.

$$\psi_0(n) = \sqrt{M} \sum_{k=0}^{N-1} h_i(k)\psi_0(Mn-k) \tag{171}$$

From this information $\psi_0(\frac{i}{M})$ is computed for all i. Using this information $\psi_0(\frac{i}{M^2})$ is computed. This recursive interpolation process is continued until the samples at the required accuracy, namely $\psi_0(\frac{i}{M^J})$ is obtained. Define the N-vector functions $\Psi_i(t), i = 0, 1, \ldots, M-1$ on $[0,1)$, made up of the samples of $\psi_i(t)$ respectively at the integers.

$$[\Psi_i]_j(t) = \psi_i(t+j); \quad 0 \leq j < N \tag{172}$$

Without loss of generality we will assume that ψ_i is supported in $[0, N-1]$. This is equivalent to assuming that the $h_i(n)$ are finite sequences running from $n = 0$ to $n = N-1$. Since ψ_i is supported in $[0, N-1]$, it is clearly completely determined by $\Phi_i(t)$ for $t \in [0,1]$. Now define $N \times N$ matrices $M_{i,j}, i = 0, 1, \ldots, M-1, j = 0, \ldots M-1$ by

$$(M_{i,j})(k,l) = \sqrt{M} h_i(Mk - l + i). \tag{173}$$

Then the recursive equations

$$\psi_i(t) = \sqrt{M} \sum_k h_i(k)\psi_0(Mt-k) \tag{174}$$

can be written in the compact form

$$\Psi_i(t) = M_{i,j}\Phi_0(Mt-j) \text{ for } \frac{j}{M} \leq t \leq \frac{j+1}{M}, i = 0, 1, \ldots, M-1. \tag{175}$$

In particular when $t = 0$ and $i = 0$, we have

$$\Psi_0(0) = M_{0,0}\Psi_0(0) \tag{176}$$

an eigenvalue problem that can be solved for $\Psi_0(0)$, which give the values of $\psi_0(n)$. As an aside note that for any i, $\Psi_0(\frac{i}{M-1}) = M_{0,i}\Psi_0(\frac{i}{M-1})$. For continuous ψ_0, it is completely specified by its values at the M-adic rationals, that is, numbers of the form

$$\sum_{N_1}^{N_2} t_i M^i$$

with $t_i \in \{0, 1, \ldots, M-1\}$. Let t be an M-adic rational in $[0,1)$. Then t has an M-ary expansion $.t_1t_2t_3\ldots t_K$. Then from Eqn. 175 by repeated substitution we obtain

$$\Psi_i(.t_1t_2\ldots t_K) = M_{i,t_1}M_{0,t_2}\ldots M_{0,t_K}\Psi_0(0). \tag{177}$$

Thus given any M-adic rational t, $\psi_i(t)$ can be obtained in a finite number of steps. Though the notation suggests that matrix products are involved in the computation, in reality once $\psi_i(t)$ is obtained at the integers(by solving the system of linear equations in Eqn. 176) the rest of the computation can be carried out using the filter bank associated with the filters h_i [19]. The idea of representing Eqn. 134 as a matrix product was first used by H.Resnikoff to study the relationship between Weierstrass functions and compactly supported wavelets [39], and later used by I.Daubechies to study the differentiability properties of wavelets [11, 25] both in the multiplicity 2 case.

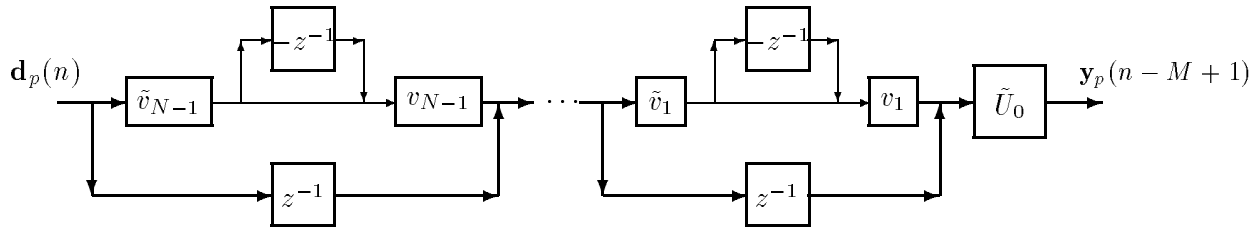

Figure 11: Cascade Implementation of Synthesis Filter Bank - Causal

We now show that the synthesis bank can also be implemented with the same number of delays and complexity as the analysis bank. Moreover, it turns out that the same program that was used for the analysis bank could be used for the synthesis bank. First note that $z^{-(N-1)}G^T(z)$, is a unitary polynomial matrix in z^{-1} and hence by Lemma. 5, there exist unitary vectors w_i and a unitary matrix W_0 such that

$$z^{-(M-1)}G^T(z) = [I - w_{N-1}\tilde{w}_{N-1} + z^{-1}w_{N-1}\tilde{w}_{N-1}]\ldots[I - w_1\tilde{w}_1 + z^{-1}w_1\tilde{v}_1]W_0 \tag{169}$$

Therefore,

$$z^{-(M-1)}\mathbf{y}_p(z) = \left\{[I - w_{N-1}\tilde{w}_{N-1} + z^{-1}w_{N-1}\tilde{w}_{N-1}]\ldots[I - w_1\tilde{w}_1 + z^{-1}w_1\tilde{w}_1]U_0\right\}\mathbf{d}_p(z) \tag{170}$$

and this can be implemented exactly as in Fig. 9. The computational and storage complexity is minimal. However, there is one disadvantage. Now, one has to compute the factorization parameters of both H_p and $z^{M-1}G_p(z)$ and store both of them.

7 Computation in Wavelet Bases

7.1 Generation of ψ_i

Given a multiplicity M wavelet frame or tight frame, the first question for applications is to know how they look like. This section deals with the computation of the samples of ψ_i. It turns out that for a multiplicity M frame, computation of arbitrary samples of ψ_i is relatively difficult, while samples at the M-adic rationals (i.e numbers of the form $\frac{i}{M^J}$ for fixed J). It turns out that M-adic samples of ψ_i can be computed using the filters h_i in the filter bank associated with the frame. We give two efficient methods to compute ψ_i at the M-adics. Notice that the samples don't really give a picture of ψ_i unless the J is large *and* the functions ψ_i are continuous, for they may behave wildly in between samples. Since it only makes sense to sample continuous functions, we will assume that ψ_i are continuous. We will also assume that ψ_i are compactly supported, or equivalently that $h_i(n)$ are FIR.

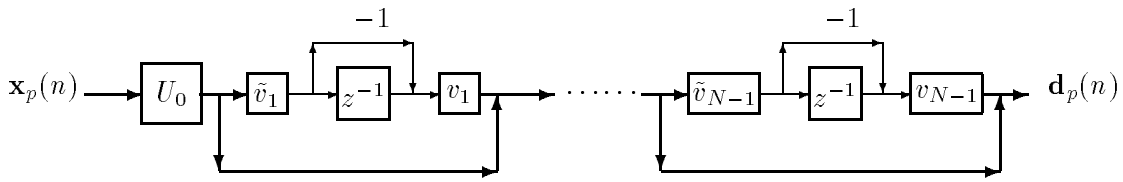

Figure 9: Cascade Implementation of the Analysis Filter Bank

is regular and can be easy be coded. The thick lines in Fig. 9 correspond to vector data, and the thin lines to scalar data. Therefore the number of storage elements required to implement the entire filter bank is $N-1$, the McMillan degree of H_p and hence the realization of the analysis FB is minimal. Let us estimate the computational complexity of the analysis FB. Recall that since each of the filters is of length NM, computation by direct convolution requires NM^3 multiplies and $(NM-1)M^2$ adds (for each vector output $\mathbf{d}_p$), while the polyphase structure requires NM^2 multiplies and $(N-1)M^2$ additions. As for the cascade implementation, there are $N-1$ blocks, followed by a unitary matrix multiplication. Since each block requires $2M$ multiplies and M adds, the entire cascade requires $[(N-1)2M+M^2]$ multiplies and $[(N-1)M+(M-1)M]$ additions. These counts could be moderately improved by taking advantage of the fact that the unitary matrix multiplication can be done more efficiently than a general matrix multiplication. As for storage requirements in a direct form implementation of $H_p(z)$, the obvious convolution implementation requires $(NM-1)M$ scalar storages (delays), the polyphase implementation $(N-1)M$ delays, and the cascade implementation requires just $N-1$ delays (the minimal possible since the McMillan degree of $H_p(z)$ is $N-1$).

The polyphase synthesis bank matrix is given by

$$G_p^T(z) = \tilde{H}_p(z) = \tilde{U}_0[I - v_1\tilde{v}_1 + zv_1\tilde{v}_1]\ldots[I - v_{N-2}\tilde{v}_{N-2} + zv_{N-2}\tilde{v}_{N-2}] \tag{168}$$

The corresponding structure for the synthesis bank is shown in Fig. 10 The computational and

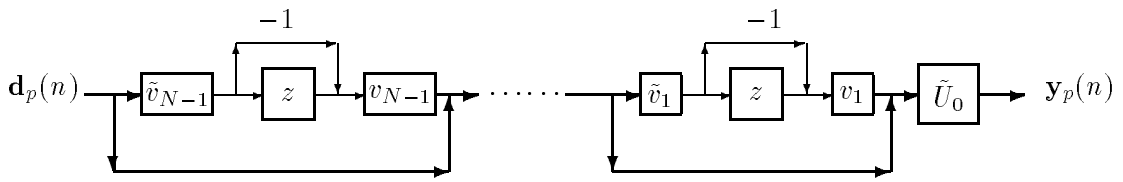

Figure 10: Cascade Implementation of Synthesis Filter Bank

storage complexity of this structure is identical to that of the analysis bank. However, because z-blocks are involved, the implementation is non-causal and cannot be realized. A fix is to reorganize the structure as shown in Fig. 11 This implementation has the same complexity, but requires, $(M+1$ delays per block for a total of $(N-1)(M+1)$ storage elements.

As for the synthesis bank, from Eqn. 58 we have

$$\mathbf{y}_p = G_p^T \mathbf{d}_p \tag{166}$$

and this can be implemented as shown in Fig. 6.1 As in the analysis bank this implementation is

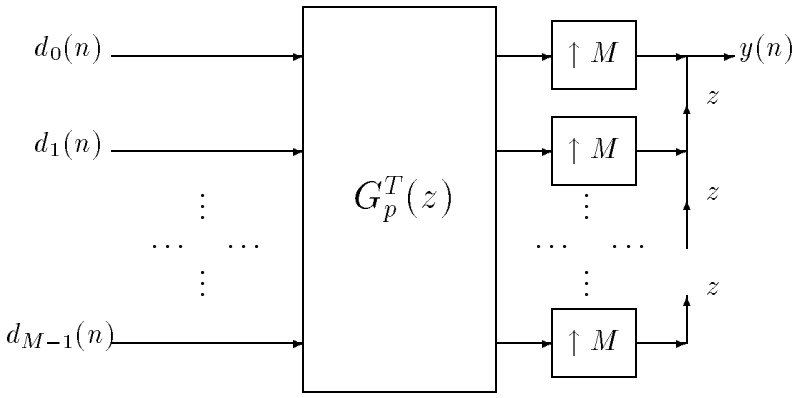

Figure 8: Polyphase Structure for the Synthesis Bank

more efficient by a factor of M.

6.2 The Cascade Structure for Unitary Filter Banks

Given an FB, in general the best possible implementation is the polyphase implementation. However, if the FBs are unitary, the factorization result in Lemma. 4 can be used to further simplify the polyphase structure into smaller computational units. There are many other factorizations [44, 1]. It is useful to recall that that for a constant unitary matrix, there are two well known factorizations, one based on Given's rotations, and the other based on Householder reflections [16]. Analogously, in the case of matrices unitary on the unit circle there is a factorization based on Given's rotations ([44]) and another based on Householder reflections. The factorization discussed in Section. 59 is based on Householder reflections. From a computational perspective, in the two channel case, we find that the Given's rotation based factorizations gives a (moderately) more efficient structure to implement the filter bank. The structure is called the *lattice* structure [46]. In the general M channel case, the Householder reflection method is much more efficient. Factorization approaches are useful only for IIR and FIR rational, stable unitary matrices.

Consider an FIR filter bank with filters $h_i(n)$ all of length NM. For simplicity let us further assume that they vanish for n less that 0 and n greater than $NM-1$. Then the polyphase matrix $H_p(z)$ will be of (McMillan) degree $N-1$. Hence by Lemma. 4 H_p can be parameterized by $N-1$ unitary vectors $v_i, i = 1, 2, \ldots, N-1$ and a unitary matrix U_0.

$$H_p(z) = [I - v_{N-1}\tilde{v}_{N-1} + z^{-1} v_{N-1}\tilde{v}_{N-1}] \ldots [I - v_1\tilde{v}_1 + z^{-1} v_1\tilde{v}_1] U_0 \tag{167}$$

From this, it is clear that the analysis FB can be implemented as shown in Fig. 9. The structure

What is the relationship between, the spaces $V_{i,j}$ and $\tilde{V}_{i,j}$? When we have biorthogonal wavelet bases it turns out that have for all i, $V_{i,j} \perp \tilde{V}_{l,j}$ when $l \neq j$.

Necessary and sufficient conditions so that the frames become a biorthogonal basis can be derived exactly as in the multiplicity 2 case in [4]. Most of the wavelet frames will be biorthogonal bases. At this, point regularity conditions can also be imposed on the wavelets and scaling functions. The function $\tilde{\psi}_0(t)$ is called the dual scaling function and the functions $\tilde{\psi}_i(t)$ are called dual wavelets. Multiplicity 2 biorthogonal wavelet bases have been constructed by Vetterli ([47]) , Chui ([2]) and Cohen ([4]). But for Chui, the others have used filter bank theory as the starting point. Chui starts from a multiresolution analysis with spline functions and constructs spline wavelets. Since spline functions are as smooth as we require, these are regular wavelet bases. As mentioned earlier, the reasons to require the regularity of the functions are usually suspect.

6 Computation in Filter Banks

6.1 Polyphase Structure for FIR Filter Banks

For the general filter bank the most efficient way to implement the analysis and synthesis banks is to use polyphase filter structures [38]. The essential idea is to avoid computation of the filter outputs of the analysis bank at all time instants since only every Mth output is required. From Eqn. 57

$$\mathbf{d}_p = H_p \mathbf{x}_p. \tag{165}$$

This can be represented in a block diagram as in Fig. 6.1 and constitutes the polyphase implementation of the analysis filter bank. It is clear that the computational complexity of this implementation

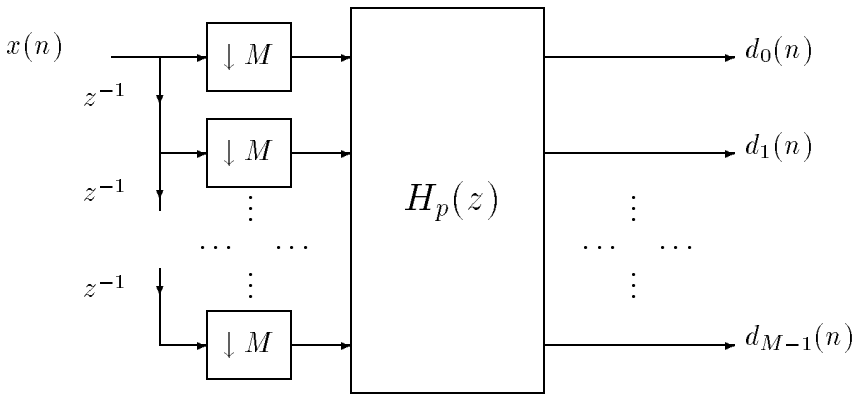

Figure 7: Polyphase Structure for the Analysis Bank

is is less than that of the direct implementation by a factor of M, a substantial gain. And this is not surprising, since only only every Mth output of the filters is being computed. Once again, this brings out the usefulness of considering the analysis as a MIMO linear shift-invariant system.

$$\left\| f - \sum_{|j|\leq L,|k|\leq L} \sum_i I_{i,j} f \right\| \tag{155}$$

$$= \sup_{\|g\|=1} \left| \langle g, f\rangle - \sum_{|j|\leq L,|k|\leq L} \sum_i \langle g, I_{i,j} f\rangle \right|$$

$$\leq \sup_{\|g\|=1} \sum_{i=1}^{M-1} \sum_{|j|>L,|K|>L} \left| \sum_{|j|>L,|k|>L} \sum_i \langle g, I_{i,j} f\rangle \right|$$

$$\leq \sup_{\|g\|=1} \left[\sum_{i=1}^{M-1} \sum_{j,k} |\langle g, \psi_{i,j,k}\rangle| \right]^{\frac{1}{2}} \left[\sum_{j>L,k>L} \left| \left\langle \tilde{\psi}_{i,j,k}, f \right\rangle \right| \right]^{\frac{1}{2}}$$

$$\leq C \left[\sum_{i=1}^{M-1} \sum_{j>L,k>L} \left| \left\langle \tilde{\psi}_{i,j,k}, f \right\rangle \right| \right]^{\frac{1}{2}} \tag{156}$$

□

Wavelet Frames are also associated with a multiresolution analysis, but the structure of the multiresolution analysis is more complex than in the tight frames case. Define the spaces

$$V_{i,j} = Span\left\{\psi_{i,j,k}(t)\right\} \quad \text{and} \quad \tilde{V}_{i,j} = Span\left\{\tilde{\psi}_{i,j,k}(t)\right\} \tag{157}$$

Then for any f in $L^2(\mathbb{R})$, $I_{i,j} f \in \tilde{V}_{i,j}$. For $f \in L^2(\mathbb{R})$, from Eqn. 144

$$I_{0,j} f = \sum_{i=0}^{M-1} I_{i,j-1} f \quad \text{and} \quad \tilde{I}_{0,j} f = \sum_{i=0}^{M-1} \tilde{I}_{i,j-1} f \tag{158}$$

and from this we have a decompositions

$$V_{0,j} = V_{0,j-1} \oplus V_{1,j-1} \ldots \oplus V_{M-1,j-1} \tag{159}$$

$$\tilde{V}_{0,j} = \tilde{V}_{0,j-1} \oplus \tilde{V}_{1,j-1} \ldots \oplus \tilde{V}_{M-1,j-1}. \tag{160}$$

From the fact that we have a frame

$$\lim_{j\to\infty} V_{0,j} = L^2(\mathbb{R}) = \lim_{j\to\infty} \tilde{V}_{0,j} \tag{161}$$

and as before

$$\lim_{j\to-\infty} V_{0,j} = \{0\} = \lim_{j\to-\infty} \tilde{V}_{0,j} \tag{162}$$

Thus we have a double multiresolution analysis of $L^2(\mathbb{R})$, with a chain of closed subspaces

$$\{0\} \subset \ldots V_{0,-1} \subset V_{0,0} \subset V_{0,1} \ldots \subset L^2(\mathbb{R}) \tag{163}$$

and

$$\{0\} \subset \ldots \tilde{V}_{0,-1} \subset \tilde{V}_{0,0} \subset \tilde{V}_{0,1} \ldots \subset L^2(\mathbb{R}) \tag{164}$$

for all f in $L^2(\mathbb{R})$. Since $\psi_{0,0,0}(t)$ and $\tilde{\psi}_{0,0,0}(t)$ may *not* form a partition of unity, we cannot use the approach taken in the tight frames case.

But as in [4], we first show that there is weak convergence and then strengthen it to strong convergence in $L^2(\mathbb{R})$. For weak convergence we must have for any f, g in $L^2(\mathbb{R})$

$$\sum_{i=1}^{M-1} \sum_{j=-\infty}^{\infty} \langle g, I_{i,j} f \rangle = \lim_{j \to \infty} \langle g, I_{0,j} f \rangle = \langle g, f \rangle . \tag{150}$$

Since both $\psi_{0,j,k}$ and $\tilde{\psi}_{0,j,k}$ both approach the Dirac measure at $2^{-j}k$ for large j (recall that $H_0(0) = G_0(0) = \sqrt{M}$ and hence $\hat{\psi}_0(0) = \hat{\tilde{\psi}}_0(0) = 1$), this is expected, and in fact can be proved rigorously by noting that

$$\begin{aligned}
&\langle g, I_{0,j} f \rangle \qquad (151)\\
&= \sum_k \langle g, \psi_{0,j,k} \rangle \left\langle \tilde{\psi}_{0,j,k}, f \right\rangle \\
&= M^j \sum_k \left\langle g, \psi_0(M^j t - k) \right\rangle \left\langle \tilde{\psi}_0(M^j t - k, f \right\rangle \\
&= M^{-j} \sum_k [\frac{1}{2\pi} \int \hat{g}(\omega) \hat{\psi}_0^*(\frac{\omega}{M^j}) e^{i\frac{\omega}{M^j}k} d\omega] \left[\frac{1}{2\pi} \int \hat{f}^*(\lambda) \hat{\tilde{\psi}}_0(\frac{\lambda}{M^j}) e^{-i\frac{\lambda}{M^j}k} d\lambda \right] \\
&= \frac{1}{2\pi} \int \hat{g}(\omega) \hat{\psi}_0(\frac{\omega}{M^j}) \left[\hat{f}^*(\lambda) \hat{\psi}_0(\frac{\omega}{M^j}) \frac{1}{2\pi M^j} \sum_k e^{i\frac{\omega-\lambda}{M^j}k} d\lambda \right] d\omega \\
&= \frac{1}{2\pi} \int \hat{g}(\omega) \hat{\psi}_0(\frac{\omega}{M^j}) \left[\hat{f}^*(\lambda) \hat{\psi}_0(\frac{\lambda}{M^j}) \sum_l \delta(\frac{\omega-\lambda}{M^j} + 2\pi l) d\lambda \right] d\omega \\
&= \sum_l \frac{1}{2\pi} \int \hat{g}(\omega) \hat{\psi}_0(\frac{\omega}{M^j}) \hat{f}^*(\omega + 2\pi M^j l) \hat{\psi}_0(\frac{\omega}{M^j} + 2\pi l) d\omega \qquad (152)
\end{aligned}$$

and all terms in the right hand side are small for $l \neq 0$ and large j and the $l = 0$ term approaches the desired inner product. To strengthen this weak convergence to strong convergence we need the following result.

$$\sum_{i,j,k} |\langle f, \psi_{i,j,k} \rangle|^2 \leq C \|f\|^2 \tag{153}$$

for some $C \geq 0$. Notice that since the sum over i is finite it suffices to show that

$$\sum_{j,k} |\langle f, \psi_{i,j,k} \rangle|^2 \leq C_i \|f\|^2 \tag{154}$$

for some $C_i \geq 0$. This can be proved by elementary means using epsilon-delta arguments. The two important facts required for the proof are that for $i \neq 0$, firstly $\hat{\psi}_i(0) = 0$ and secondly, that $\hat{\psi}(\omega)$ decays faster than $|\omega|^{-\frac{1}{2}}$ for large ω. The first fact follows trivially from the hypothesis that $H_i(0) = G_i(0) = 0$, and the second from the fact that $\psi_i(t)$ is in $L^2(\mathbb{R})$. These two facts can be used to show that for j and k tending to ∞, the numbers $|\langle f, \psi_{i,j,k} \rangle|^2$ decay rapidly enough for the double sum to converge.

We have strong convergence because,

for suitable non-negative constants. And this follows from Bessel inequality by noting that the functions $\left\{\psi_{i,j,(N-1)m+k}/\left\|\psi_i\right\|\right\}$ for fixed i, j and k, with $k = 0, 1, \ldots, N-2$, form an orthonormal family (since they don't overlap) and hence

$$\sum_k |\langle f, \psi_{i,j,k}\rangle|^2 \leq (N-1)\left\|\psi_i\right\|\left\|f\right\|^2 . \tag{143}$$

Similarly for the case of $\tilde{\psi}_{i,j,k}$. For any fixed j,

$$I_{0,j} = \sum_{i=0}^{M-1} I_{i,j-1} \quad \text{and} \quad \tilde{I}_{0,j} = \sum_{i=0}^{M-1} \tilde{I}_{i,j-1} \tag{144}$$

This result can be proved algebraically and only depends on the fact that the sequences h_i and g_i form a filter bank. Firstly notice that for all j,

$$\psi_{i,j-1,k}(t) = \sum_l h_i(l)\psi_{0,j,Mk+l} \tag{145}$$

and

$$\tilde{\psi}_{i,j-1,k}(t) = \sum_l g_i(l)\psi_{0,j,Mk-l} \tag{146}$$

Now we have,

$$\begin{aligned}
\sum_{i=0}^{M-1} I_{i,j-1} &= \sum_{i=0}^{M-1}\sum_k \psi_{i,j-1,k} >< \tilde{\psi}_{i,j-1,k} \\
&= \sum_{i=0}^{M-1}\sum_k\sum_l h_i(l)\psi_{0,j,Mk+l} >< \sum_m g_i(m)\tilde{\psi}_{0,j,Mk-m} \\
&= \sum_{i=0}^{M-1}\sum_k\sum_l h_i(-Mk+l)\psi_{0,j,l} >< \sum_m g_i(Mk-m)\tilde{\psi}_{0,j,m} \\
&= \sum_{l,m}\psi_{0,j,l} >< \tilde{\psi}_{0,j,m}\left[\sum_{i=0}^{M-1}\sum_k h_i(Mk+l)g_i(-Mk-m)\right] \\
&= \sum_{l,m}\psi_{0,j,l} >< \tilde{\psi}_{0,j,m}\left[\delta(l-m)\right] \\
&= \sum_k \psi_{0,j,k} >< \tilde{\psi}_{0,j,k} = I_{0,j}
\end{aligned} \tag{147}$$

Lemma. 10 has been used in the proof. By repeatedly substituting for $I_{0,j-1}$, $I_{0,j-2}$ etc. we conclude that for a fixed J

$$I_{0,J} = \sum_{i=1}^{M-1}\sum_{j=-\infty}^{J-1} I_{i,j} \tag{148}$$

Notice that $\psi_{i,j,k}(t)$ approaches the zero function for all i and k as, $j \to \infty$ and hence the infinite sum over j makes sense. We need to show that

$$\lim_{j\to\infty} I_{0,j} f = f \tag{149}$$

Define the functions

$$\psi_0(\omega) = \prod_{j\geq 1} H_0(\frac{\omega}{M^j}) \tag{132}$$

$$\tilde{\psi}_0(\omega) = \prod_{j\geq 1} G_0(\frac{-\omega}{M^j}) \tag{133}$$

These functions represent Fourier Transforms of distributions in general. Assume further that $\psi_0(t)$ *and* $\tilde{\psi}_0(t)$ *are in* $L^2(\mathbb{R})$ *(notice that this implies restrictions on the* h_i *and* g_i*). Now define*

$$\psi_i(t) = \sqrt{M}\sum_k h_i(k)\psi_0(Mt-k) \;\;;\;\; \psi_{i,j,k}(t) = M^{j/2}\psi_i(M^j t-k) \tag{134}$$

$$\tilde{\psi}_i(t) = \sqrt{M}\sum_k g_i(k)\tilde{\psi}_0(Mt+k) \;\;;\;\; \tilde{\psi}_{i,j,k}(t) = M^{j/2}\tilde{\psi}_i(M^j t-k) \tag{135}$$

then $\{\psi_{i,j,k}\}$ *and* $\left\{\tilde{\psi}_{i,j,k}\right\}$ *form dual frames for* $L^2(\mathbb{R})$*. That is,*

$$f(t) = \sum_{i=1}^{M-1}\sum_{j,k}\left\langle\tilde{\psi}_{i,j,k}(t), f\right\rangle\psi_{i,j,k} = \sum_{i=1}^{M-1}\sum_{j,k}\langle f,\psi_{i,j,k}(t)\rangle\,\tilde{\psi}_{i,j,k}. \tag{136}$$

Also

$$f(t) = \sum_k\left\langle\tilde{\psi}_{0,0,k}(t), f\right\rangle\psi_{0,0,k}(t) + \sum_{i=1}^{M-1}\sum_{j=1}^{\infty}\sum_k\left\langle\tilde{\psi}_{i,j,k}(t), f\right\rangle\psi_{i,j,k}(t) \tag{137}$$

$$f(t) = \sum_k\langle f,\psi_{0,0,k}(t)\rangle\, f\tilde{\psi}_{0,0,k}(t) + \sum_{i=1}^{M-1}\sum_{j=1}^{\infty}\sum_k\langle f,\psi_{i,j,k}(t)\rangle\,\tilde{\psi}_{i,j,k}(t). \tag{138}$$

Proof : The first part of the proof of this theorem parallels the wavelets tight frames theorem. The functions $\psi_i(t)$ and $\tilde{\psi}_i(t)$ are all compactly supported in an interval of length $\frac{N-1}{M-1}$ as in the tight frames case. Define two families of operators

$$\tilde{I}_{i,j} = \sum_k \tilde{\psi}_{i,j,k} >< \psi_{i,j,k} \tag{139}$$

$$I_{i,j} = \sum_k \psi_{i,j,k} >< \tilde{\psi}_{i,j,k} \tag{140}$$

These operators makes sense for all i for an arbitrary $f \in L^2(\mathbb{R})$. To see this, one cannot use Bessel inequality directly as in the tight frames case, but notice that it suffices to show that given f,

$$\sum_k |\langle f,\psi_{i,j,k}\rangle|^2 \leq C_i\,\|f\|^2 \tag{141}$$

$$\sum_k \left|\left\langle f,\tilde{\psi}_{i,j,k}\right\rangle\right|^2 \leq \tilde{C}_i\,\|f\|^2 \tag{142}$$

It is also easily seen that

$$\lim_{j \to -\infty} V_{0,j} = \{0\} \tag{128}$$

Thus we have a multiresolution analysis of $L^2(\mathbb{R})$, with a chain of closed subspaces

$$\{0\} \subset \ldots V_{0,-1} \subset V_{0,0} \subset V_{0,1} \ldots \subset L^2(\mathbb{R}) \tag{129}$$

In general the spaces $V_{i,j}$ for fixed j are not mutually orthogonal. But when we have an orthonormal basis, it is indeed true that for fixed j, and $i \neq l$, $V_{i_1,j} \perp V_{i_2,j}$

In Section. 8 we will further explore these ideas and show that this type of multiresolution analysis have only to do with filter banks, and have little do with the fact that there exist scaling functions and wavelets.

In summary, given a unitary filter bank with a "DC" condition, there exist wavelet tight frames. In the multiplicity M case also necessary and sufficient conditions for orthonormality of the bases can be worked out as in [28, 20]. An important point to note is that most of the tight frames would be orthonormal bases. Once we have an orthonormal basis, we can impose regularity on the functions $\psi_i(t)$, or equivalently on $\psi_0(t)$. Most of the work in [7] is in constructing wavelets with arbitrary large degree of differentiability. As is shown there, the regularity is related to the vanishing of the discrete moments of the sequences $h_0(n)$, or equivalently the flatness of $H_0(\omega)$ at $\omega = 0$. A number of researchers have stressed the importance of regularity of the analysis functions [7, 47, 4, 2]. In Section. 8 we argue that that the regularity of the functions $\psi_i(t)$ may not be of any practical significance per se, but that the flatness of $H_0(\omega)$ may be required in some applications.

5 Wavelet Frames

From unitary FIR FBs with the property that $H_0(\omega) = \sqrt{M}$, we can construct compactly supported wavelet tight frames and orthonormal bases. This raises the following interesting question. Do all FIR FBs modulo some linear conditions (like $H_0(0) = \sqrt{M}$) lead to expansions like Eqn. 95? This section tries to answer this question. For the case $M = 2$ a reasonably complete answer to this question is given in [4]. Recall that in the construction of tight frames, only one sequence $h_0(n)$ was required. The FB and the wavelet tight frame was constructed from it. This was facilitated by the elegant factorization theory of unitary matrices over the Laurent polynomials. However in the general case such an approach is not known. Yet given an FIR FB, with $M + 1$ "DC" constraints, $H_i(1) = \sqrt{M}\delta(i)$ and $G_0(1) = \sqrt{M}$, we can show that we have a wavelet frame (a dual set of functions that behaves like the a biorthogonal basis and its dual as far as expansions and expansion coefficients are concerned).

Theorem 3 (Wavelet Frames Theorem) *Let h_i, and g_i be the analysis and synthesis filters in an FIR FB such that,*

$$H_0(0) = G_0(0) = \sqrt{M} \tag{130}$$

and for all $i = 1, 2, \ldots, M-1$

$$H_i(0) = G_i(0) = 0 \tag{131}$$

continuous f with compact support, if j and k together approach infinity such that $M^{-j}k \to t$, then,

$$\lim_{j,k\to\infty} \langle \psi_{0,j,k}, f(x) \rangle = f(t) \tag{120}$$

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

$$\sum_k \psi_0^0(t+k) = 1. \tag{121}$$

Now by induction, we have for any j,

$$\sum_k \psi_0^j(t+k) = \sqrt{M} \sum_k \sum_l h_0(l)\psi_0^{j-1}(Mt - Mk - l). \tag{122}$$

Changing variables and interchanging the order of summation and invoking Eqn. 98 in Lemma. 11 we get,

$$\begin{aligned} \sum_k \psi_0^j(t+k) &= \sqrt{M} \sum_k \sum_l h_0(Mk-l)\psi_0^{j-1}(Mt+l) \\ &= \sqrt{M} \sum_l [\sum_l h_0(Mk-l)]\psi_0^{j-1}(Mt+l) \\ &= \sum_l \psi_0^{j-1}(Mt+l) \\ &= 1 \end{aligned} \tag{123}$$

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

The function $\psi_0(t)$ is called the scaling function and the the $M-1$ functions, $\psi_i(t), i \neq 1$ are called wavelets. The corresponding sequences are called scaling vectors and wavelet vectors respectively.

Multiplicity M wavelet tight frames have a natural multiresolution analysis structure associated with them. It essentially comes from the filter bank structure that lead to Eqn. 116. To see this, define the spaces,

$$V_{i,j} = Span\{\psi_{i,j,k}\}. \tag{124}$$

For any $f \in L^2(\mathbb{R})$, note that $I_{i,j}f \in V_{i,j}$. In particular, for $f \in V_{0,j}f$,

$$f = I_{0,j}f = \sum_{i=0}^{M-1} I_{i,j-1}f \tag{125}$$

leading to a decomposition

$$V_{0,j} = V_{0,j-1} \oplus V_{1,j-1} \ldots \oplus V_{M-1,j-1}. \tag{126}$$

Also from the fact that $I_{0,j}$ becomes I on $L^2(\mathbb{R})$ we have

$$\lim_{j\to\infty} V_{0,j} = L^2(\mathbb{R}) \tag{127}$$

We now need to show that $\psi_{i,j,k}$ as defined in Eqn. 103 form a tight frame. Firstly, since the sequences, h_i are all of finite length, say N and are supported in $[0, N-1)$ it can be shown using standard Paley-Weiner arguments that the functions $\psi_{i,j}(t)$ are compactly supported in $[0, M^{-j}(\frac{N-1}{M-1})]$ [4]. Now define the operators

$$I_{i,J} = \sum_k \psi_{i,J,k} >< \psi_{i,J,k} \tag{114}$$

The operators make sense for any f in $L^2(\mathbb{R})$. A simple check is to use the fact that $\left\{\psi_{i,j,(N-1)m+k}(t)/\left\|\psi_i\right\|\right\}$ is an orthonormal family for any fixed $k = 0, 1, 2, \ldots, N-2$ (because $\psi_{i,j}$ are compactly supported) and hence by Bessel Inequality [41],

$$\left\|I_{i,j} f\right\| \leq (N-1)\left\|\psi_0\right\|^2 \left\|f\right\|. \tag{115}$$

Notice also that the operator is uniformly bounded in j for all i. Now for any fixed j,

$$I_{0,j} = \sum_{i=0}^{M-1} I_{i,j-1} \tag{116}$$

since

$$\begin{aligned}
\sum_{i=0}^{M-1} I_{i,j-1} &= \sum_{i=0}^{M-1} \sum_k \psi_{i,j-1,k} >< \psi_{i,j-1,k} \\
&= \sum_{i=0}^{M-1} \sum_k \sum_l h_i(l)\psi_{0,j,Mk+l} >< \sum_m h_i(m)\psi_{0,j,Mk+m} \\
&= \sum_{i=0}^{M-1} \sum_k \sum_l h_i(-Mk+l)\psi_{0,j,l} >< \sum_m h_i(-Mk+m)\psi_{0,j,m} \\
&= \sum_{l,m} \psi_{0,j,l} >< \psi_{0,j,m} \left[\sum_{i=0}^{M-1} \sum_k h_i(Mk+l)h_i(Mk+m)\right] \\
&= \sum_{l,m} \psi_{0,j,l} >< \psi_{0,j,m} \left[\delta(l-m)\right] \\
&= \sum_k \psi_{0,j,k} >< \psi_{0,j,k} = I_{0,j}.
\end{aligned} \tag{117}$$

Notice that we used Lemma. 9 to obtain this result. By repeatedly substituting for $I_{0,j-1}$, $I_{0,j-2}$ etc. we conclude for a fixed J,

$$I_{0,J} = \sum_{i=1}^{M-1} \sum_{j=-\infty}^{J-1} I_{i,j}. \tag{118}$$

We will now show that $I_{0,j,k}$ approximates the identity operator in $L^2(\mathbb{R})$. That is for all $f \in L^2(\mathbb{R})$

$$\lim_{j \to \infty} I_{0,j} f = f. \tag{119}$$

This result is to be expected since $\psi_{0,j,0}$ approaches the Dirac measure at the origin as $j \to \infty$. From Eqn. 115 it suffices to show that $I_{0,j} f = f$, only for a dense subset of $L^2(\mathbb{R})$. For any

$$
\begin{aligned}
&= \left\langle \sqrt{M}\sum_n h_0(n)\psi_0^j(Mt-n), \sqrt{M}\sum_l h_0(l)\psi_0^j(Mt-Mk-l) \right\rangle \\
&= \sum_{n,l} h_0(n)h_0(l)[M\int \psi_0^j(Mt-n)\psi_0^j(Mt-Mk-l)\,dt] \\
&= \sum_{n,l} h_0(n)h_0(l)\delta(n-Mk-l) \\
&= \sum_n h_0(n)h_0(n-Mk) = \delta(k). \qquad (108)
\end{aligned}
$$

Therefore for all j, $\left\|\psi_0^j\right\| = 1$.

We now need to show that g does indeed satisfy Eqn. 102. To see that first we show that ψ_0^j converges as a distribution to some function that satisfies Eqn. 102. Since the distribution topology is weaker than the weak topology [24] in $L^2(\mathbb{R})$, this would imply that $g = \psi_0^j$. The Fourier transform of Eqn. 106 gives

$$\hat{\psi}_0^{j+1}(\omega) = \hat{\psi}_0^j(\frac{\omega}{M})\left[\frac{1}{\sqrt{M}}H_0(\frac{\omega}{M})\right]. \qquad (109)$$

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

$$\hat{\psi}_0(\omega) = \prod_{j\geq 1}\left[\frac{1}{\sqrt{M}}H_0(\frac{\omega}{M^j})\right]. \qquad (110)$$

Let $C = \max \hat{\psi}_0^0(\omega)$ for $|\omega| \leq 1$. Let $A = \max\left[\frac{1}{\sqrt{M}}|H_0(\omega)|\right] \leq 1$ by Eqn. 96. For any $\omega, |\omega| \geq 1$,

$$M^{\lfloor log_M(\omega)\rfloor} \leq \omega \leq M^{\lceil log_M(\omega)\rceil} \qquad (111)$$

and therefore from Eqn. 109

$$\hat{\psi}_0^j(\omega) \leq CA^{\lceil \log_M(\omega)\rceil} \qquad (112)$$

thus showing that $\hat{\psi}_0^j$ are bounded by polynomial growth for large ω hence in the time domain, ψ_0^j, converges to a distribution ψ_0 whose Fourier Transform is $\hat{\psi_0}$ defined in Eqn. 110.

From Lemma. 11, we know that $\mathbf{h}_0$ is of the form

$$\mathbf{h}_0(z) = P_L(z)P_{L-1}(z)\ldots P_1(z)\mathbf{e} \qquad (113)$$

for some L. Notice that for any scaling vector the corresponding $\mathbf{h}_0(z)$ will be of the above form and thus all scaling vectors could be generated from the above equation.

Now we need to construct a unitary FIR filter bank, such that h_0 is one of the analysis filters. That is we need to construct $h_i, i = 1, 2, \ldots, M-1$. From Lemma. 4, and Eqn. 113, it is clear that this problem is equivalent to constructing a unitary matrix U_0 whose first column is $\mathbf{e}$. This can be done by a Gram-Schmidt process in $\mathbb{R}^M$. A more efficient procedure using state-space techniques is given in [**?**]. This process in non-unique (unlike the multiplicity 2 case, where h_1 was determined by h_0), and therefore the sequences h_i are not uniquely determined.

Now for all i,

$$P_i(z)|_{z=1} = [I - v\tilde{v}_i + z^{-1}v\tilde{v}_i]_{z=1} = I$$

and therefore $\mathbf{h}_0(1) = V_0$ and hence

$$\sqrt{M} = \sum_k H_{0,j}(z)|_{z=1} = \sum_k V_{0,k} = 1.$$

The only solution to this equation for V_0, with $\|V_0\| = 1$ is $\mathbf{e}$. □ We are now ready to give the Wavelet Tight Frames Theorem.

Theorem 2 (Wavelet Tight Frames Theorem) *Let h_0 be a scaling vector of length N. Then, there exists a function $\psi_0(t) \in L^2(\mathbb{R})$*

$$\psi_0(t) = \sqrt{M}\sum_k h_i(k)\psi_0(Mt - k) \tag{102}$$

and sequences $h_i, i = 1, 2, \ldots M-1$, such that h_i form the analysis filters of a unitary filter bank. Define

$$\psi_i(t) = \sqrt{M}\sum_k h_i(k)\psi_0(Mt - k) \;\; ; \;\; \psi_{i,j,k}(t) = M^{j/2}\psi_i(M^j t - k) \tag{103}$$

then $\{\psi_{i,j,k}\}$ forms a tight frame for $L^2(\mathbb{R})$. That is,

$$f(t) = \sum_{i=1}^{M-1}\sum_{j,k}\langle f, \psi_{i,j,k}(t)\rangle\,\psi_{i,j,k}. \tag{104}$$

Also

$$f(t) = \sum_k \langle f, \psi_{0,0,k}(t)\rangle\,\psi_{0,0,k}(t) + \sum_{i=1}^{M-1}\sum_{j=1}^{\infty}\sum_k \langle f, \psi_{i,j,k}(t)\rangle\,\psi_{i,j,k}(t). \tag{105}$$

Proof : The proof is similar to the proof in the multiplicity 2 case in [29]. The difficulty in the multiplicity M case is the construction of wavelet sequences (i.e analogues of h_1). The outline of the proof is as follows. We first construct a $\psi_0(t)$, from $h_0(n)$, that is in $L^2(\mathbb{R})$ and satisfies Eqn. 102. We then construct the sequences h_i, and define functions $\psi_i(t) \in L^2(\mathbb{R})$ and $\psi_{i,j,k}$ as in Eqn. 103. We then establish Eqn. 95, and will be done.

Let $\psi_0^0(t)$ denote the characteristic function of $[0, 1)$. Define

$$\psi_0^j(t) = \sqrt{M}\sum_k h_0(k)\psi_0^j(Mt - k). \tag{106}$$

This recursion converges weakly to a function in $L^2(\mathbb{R})$. To see this it suffices to know that for every j, $\left\|\psi_0^j\right\| = 1$ (i.e is in the unit ball in $L^2(\mathbb{R})$). Then by the weak compactness of the unit ball in $L^2(\mathbb{R})$, there exists a g such that $\psi_0^j \rightharpoonup g$. Now for any j, $\left\langle \phi_0^j(t), \phi_0^j(t-k)\right\rangle = \delta_{0k}$ because $\langle \phi_0^0(t), \phi_0^0(t-k)\rangle = \delta(k)$ (recall $\phi_0^0(t)$ is the characteristic function of $[0, 1)$), and by induction

$$\left\langle \psi_0^{j+1}(t), \psi_0^{j+1}(t-k)\right\rangle \tag{107}$$

Proof : See Lawton [29]. □ Even though 95 holds, the functions $\psi_{j,k}$ need not be an orthonormal basis. This can be seen even in finite dimensions. The reason for this is that the functions $\psi_{j,k}$ may not be linearly independent. That is, one of them could be in the closure of the linear span of the others.

In his paper Lawton also proves that for almost all choices of h_0, $\psi_{j,k}$ forms an orthonormal basis. Necessary and sufficient conditions on the sequence h_0 such that $\phi_{j,k}$ give rise to an orthonormal basis have been independently obtained by Lawton and Cohen, but from a practical point of view they do not seem to be important since the numerical algorithms are the same.

We now try to extend the results to the case of multiplicity M DWT type frequency decompositions. We first obtain a parameterization of $h_0(n)$ that satisfy

$$\sum_k h_0(k)h_0(k+Ml) = \delta(l) \text{ and } \sum_k h_0(k) = \sqrt{M} \tag{96}$$

Such sequences are called *multiplicity M scaling vectors* or just *scaling vectors.*

Lemma 11 *Let h_0 be a multiplicity M scaling vector of length N. Write*

$$H_0(z) = \sum_k z^{-k} H_{0,j}(z^M)$$

and define

$$\mathbf{h}_0(z) = [H_{0,1}, H_{0,2}, \ldots, H_{0,M-1}]^T$$
$$\mathbf{e} = [\frac{1}{\sqrt{M}}\frac{1}{\sqrt{M}}, \ldots, \frac{1}{\sqrt{M}}]^T.$$

Then for every h_0 there exists an integer N, such that

$$\mathbf{h}_0(z) = P_N(z)P_{N-1}(z)\ldots P_1(z)\mathbf{e}. \tag{97}$$

From this it easily follows that,

$$\sum_k h_0(Mk+m) = \sqrt{M} \quad \textit{for all } m \tag{98}$$

Proof : By the proof of Lemma. 7,

$$\sum_j H_{0,j}(z^{-1})H_{0,j}(z) = 1 \tag{99}$$

and therefore $\mathbf{h}_0$ is a unitary vector. Hence by Lemma. 5,

$$h_0(z) = P_N(z)P_{N-1}(z)\ldots P_1(z)V_0 \tag{100}$$

where P_i are 1st degree polynomial matrices as in Eqn. 69 and $V_0 \in R^M$ with $\|V_0\| = 1$ Now we are given that $\sum_k h_0(k) = \sqrt{M}$. Hence,

$$\sqrt{M} = \sum_k h_0(k) = H_0(z)|_{z=1} = [\sum_k z^{-k} H_{0,j}(z)]|_{z=1} = \sum_k H_{0,j}(z)|_{z=1}. \tag{101}$$

4 Wavelet Tight Frames

Armed with filter bank theory we are now ready to construct non-overlapping frequency decompositions of the DWT type. By a multiplicity M-DWT type frequency decomposition we mean the decomposition shown in Fig. 5. Such a decomposition is naturally associated with an M-channel filter bank. Therefore, unitary FIR filter bank theory can be used to describe all *compactly supported* multiplicity M orthonormal wavelet bases. Recall that in the 2-channel FIR *unitary* filter in Section. 3 the the analysis filters, and in particular h_0 satisfies (see Lemma. 7)

$$\sum_k h_0(k)h_0(k+2l) = \delta(l) \tag{90}$$

For the non-overlapping DWT type frequency decomposition considered in Section. 2 , this is precisely the equation satisfied by the sequence $h_0(n)$. Furthermore, in that case $h_0(n)$ determined the frequency decomposition completely. For arbitrary h_0 satisfying Eqn. 90, there there cannot be orthonormal wavelet bases since that would imply that the corresponding scaling function ϕ (see Section. 2), would have to satisfy Eqn. 31,

$$\hat{\phi}(\omega) = \hat{\phi}(0)\prod_{i=1}^{\infty}[\frac{1}{\sqrt{2}}H_0(\omega/2^i)] \tag{91}$$

and if $H_0(0) \neq \sqrt{2}$, the infinite product diverges. But if $H_0(0) = \sqrt{2}$ it turns out that, the unitary filter bank does indeed give rise to a tight frame (a set of functions, that practically behave like an orthonormal basis as far as expansions and expansion coefficients are concerned).

In fact more is true. One need not start with $h_0(k)$ from a unitary filter bank. One can just start with a sequence h_0 that satisfies both Eqn. 32 and Eqn. 90. Recall that such sequences are called scaling vectors. Daubechies give a parameterization of all scaling vectors [7]. An independent paramertrization of scaling vectors has been obtained ([36]). Using this parameterization, with some added constraints on the scaling vectors (referred to as regularity conditions) Daubechies constructed orthonormal bases for $L^2(\mathbb{R})$. Based on numerical evidence [29], Lawton realized and proved that all sequences $h_0(n)$ (independent of regularity assumptions) give rise to tight frames. More precisely,

Theorem 1 *Let h_0 be a scaling vector. Then there exists a $\phi(t) \in L^2(\mathbb{R})$ such that*

$$\phi(t) = \sum_k h_0(k)\phi(2t-k). \tag{92}$$

Define

$$h_1(k) = (-1)^k h_0(-k+1) \tag{93}$$

$$\psi(t) = \sum_k h_1(k)\phi(2t-k) \quad \textit{and} \quad \psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{94}$$

Then $\psi_{j,k}$ forms a tight frame for $L^2(\mathbb{R})$ That is, for all $f(t) \in L^2(\mathbb{R})$ we have

$$f(t) = \sum_{j,k} \langle f, \psi_{j,k}\rangle \, \psi_{j,k}(t). \tag{95}$$

Lemma 9 *$(H_i(z), G_i(z))$ form analysis/synthesis filters of a unitary filter bank iff*

$$\sum_i \sum_k h_i(Mk+n_1)h_i(Mk+n_2) = \sum_i \sum_k g_i(Mk+n_1)g_i(Mk+n_2) = \delta(n_1-n_2) \tag{82}$$

Proof : Define

$$f(n_1,n_2) = \sum_i \sum_k h_i(Mk+n_1)h_i(Mk+n_2) \tag{83}$$

Clearly, $f(n_1+M, n_2+M) = f(n_1,n_2)$ and therefore it is necessary only to consider $n_1 = 0,\ldots,M-1$. Let $n_2 = Mn+r$, where $r \in \{0,1,\ldots,M-1\}$. Then, $f(n_1,n_2)$ is of the form

$$\sum_k \sum_i h_i(Mk+n_1)h_i(M(k+n)+r) \tag{84}$$

Taking the $\mathcal{Z}$-transform of this expression as a sequence in n, for fixed n_1 and n_2 we get

$$\sum_i H_{i,n_1}(z)H_{i,r}(z^{-1}) = \delta(n_1-r) \tag{85}$$

and therefore is zero except when $n_1 = r$. When $n_1 = r$, the inverse $\mathcal{Z}$-transform is $\delta(n)$, and therefore the expression is zero except when $n = 0$. That is $n_2 = n_1$. □ The dual result for general filter banks is the following.

Lemma 10 *$(H_i(z), G_i(z))$ form analysis/synthesis filters of a filter bank iff*

$$\sum_i \sum_k h_i(Mk+n_1)g_i(-Mk-n_2) = \delta(n_1-n_2) \tag{86}$$

Proof : The proof is similar to the previous result. Define

$$f(n_1,n_2) = \sum_i \sum_k h_i(Mk+n_1)g_i(-Mk-n_2) \tag{87}$$

Clearly, $f(n_1+M, n_2+M) = f(n_1,n_2)$ and therefore it is necessary only to consider $n_1 = 0,\ldots,M-1$. Now let $n_2 = -Mn+r$, where $r \in \{0,1,\ldots,M-1\}$. Then, $f(n_1,n_2)$ is of the form

$$\sum_k \sum_i h_i(Mk+n_1)g_i(M(n-k)-r) \tag{88}$$

Taking the $\mathcal{Z}$-transform of this expression as a sequence in n, for fixed n_1 and n_2 we get

$$\sum_i H_{i,n_1}(z)G_{i,r}(z) = \delta(n_1-r) \tag{89}$$

and therefore is zero except when $n_1 = r$. When $n_1 = r$, the inverse $\mathcal{Z}$-transform is $\delta(n)$, and therefore the expression is zero except when $n = 0$. That is $n_2 = n_1$. □ .

Lemma. 8 is dual to Lemma. 10 in the sense that the former says that the rows of G_p are orthogonal to all but one of the columns of H_p, while the latter says that, the rows of H_p are orthogonal to all but one of the columns of G_p. In other words one depends on the fact that $G_p^T H_p = I$, while the other depends on the fact that $H_p G_p^T = I$.

it follows that one could swap the analysis and synthesis filters. Since wavelet theory deals with sequences $h_i(n)$ in the filter bank, this *duality* comes out as a non-trivial result [3].

Lemma 7 *The pair $(H_i(z), G_i(z))$ form analysis/synthesis filters of a unitary filter bank iff*

$$\sum_k h_i(k)h_j(k+Ml) = \sum_k g_i(k)g_j(k+Ml) = \delta(l)\delta(i-j) \tag{75}$$

Proof : We only need to show this result for the h_i since by Lemma. 6 the same would be true for g_i. Also note that we have a unitary filter bank iff $G_p(z) = H^T(z^{-1})$, which is equivalent to

$$\sum_k H_{i,k}(z)H_{j,k}(z^{-1}) = \delta(i-j). \tag{76}$$

Now let $\alpha(l) = \sum_k h_i(k)h_j(k+Ml)$. Then from Eqn. 48 and Eqn. 76

$$\begin{aligned} \mathcal{Z}(\alpha(l)) &= [\downarrow M]H_i(z)H_j(z^{-1}) \\ &= [\downarrow M]\sum_{k,l} z^{-k}H_{i,k}(z^M)z^l H_{j,l}(z^{-M}) \\ &= \sum_k H_{i,k}(z)H_{j,k}(z^{-1}) = \delta(i-j) \end{aligned} \tag{77}$$

□ This result has the following interpretation. In a unitary filter bank every filter $h_i(n)$ is orthogonal to the uniform translates of any other filter $h_j(n)$ (including itself) by multiples of M. Notice this result is exactly what we got for the scaling vector and wavelet vector in Section. 2 and that it came from the orthonormality properties of the underlying scaling function and wavelet.

There is a corresponding result for general filter banks.

Lemma 8 *$(H_i(z), G_i(z))$ form analysis/synthesis filters of a filter bank iff*

$$\sum_k h_i(k)g_j(Ml-k) = \delta(l)\delta(i-j) \tag{78}$$

Proof : The fact that $G_p^T(z)H_p(z) = I$ implies that

$$\sum_k H_{i,k}(z)G_{j,k}(z) = \delta(i-j) \tag{79}$$

Define $\alpha(l) = \sum_k h_i(k)g_j(Ml-k)$. Then from Eqn. 48, Eqn. 49 and Eqn. 79

$$\begin{aligned} \mathcal{Z}(a(l)) &= [\downarrow M]H_i(z)G_j(z) \\ &= [\downarrow M]\sum_{k,l} z^{-k}H_{i,k}(z^M)z^l G_{j,l}(z^M) \\ &= \sum_k H_{i,k}(z)G_{j,k}(z) = \delta(i-j) \end{aligned} \tag{80}$$

Therefore, taking the inverse $\mathcal{Z}$-transform on both sides we get

$$\sum_k h_i(k)g_j(Ml-k) = \delta(l)\delta(i-j) \tag{81}$$

□ There is a *dual* time-domain characterization of both unitary and general filter banks which will be required in later sections.

Proof : See Vaidyanathan. [45] □

Another result that we find useful when we discuss wavelet theory is the following representation theorem for *unitary* vectors on the unit circle. An $M \times 1$ vector $V(z)$ is said to be unitary on the unit circle if $\tilde{V}(z)V(z) = 1$. By McMillan degree of a unitary vector $V(z)$, we mean the degree of the least common multiple of the denominators of the entries in of $V(z)$. We have the following factorization theorem for unitary matrices of McMillan degree N analogous to the result for unitary matrices.

Lemma 5 *Let V_N be an arbitrary unitary vector of McMillan degree N. Then, there exists a factorization of the form*

$$V_N(z) = P_N(z)P_{N-1}(z)\dots P_1(z)V_0 \tag{70}$$

where V_0 is a constant vector with $\|V_0\| = 1$ Moreover, if V_N is a polynomial vector in z^{-1} then the factors P_i can can be shown to be of the form

$$P_i(z) = \left[I - v_i\tilde{v}_i + z^{-1}v_i\tilde{v}_i\right]. \tag{71}$$

Proof : See Vaidyanathan [44]. □

Lemma. 4 gives a complete characterization of all rational unitary filter banks. Designing such an filter bank is equivalent to choosing the parameters v_i and a_i, and clearly numerical optimization schemes for choosing these parameters for minimizing certain objective functions, that depend on the analysis filters h_i can be devised [33]. Another advantage of unitary filter banks is that, there is no need to solve for the synthesis bank by algebraic inversion of H_p since the synthesis filters are time-reverses of the analysis filters

Lemma 6 *$(H_i(z), G_i(z))$ form the analysis and synthesis filters of an M-channel unitary filter bank iff (up to delays)*

$$G_i(z) = H_i(z^{-1}) \tag{72}$$

Proof : From Eqn. 49 and Eqn. 48,

$$G_i(z) = \sum_{k=0}^{M-1} z^k G_{i,k}(z) = \sum_{k=0}^{M-1} z^k H_{i,k}(z^{-1}) = H_i(z^{-1}). \tag{73}$$

□ This result implies that in the wavelet basis considered in Section. 2, the scaling vector $h_0(n)$ and wavelet vector $h_1(n)$ form a unitary filter bank.

So far we have considered filter bank theory in the $\mathcal{Z}$-transform domain since it is much easier there than in terms of the sequences $h_i(n)$ and $g_i(n)$. Also recall that the fundamental result in filter bank theory, namely that $G_p^T = H_p^{-1}$ is of a purely algebraic character. It turns out that the algebraic result in the time domain looks relatively complex in that it gives no insight into what is happening. Nevertheless, it is the time domain equations that will be useful in dealing with wavelet theory. To give an example of the clarity of the $\mathcal{Z}$-domain approach consider the following question. Can the analysis and synthesis filters be exchanged? It is not at all obvious that this is true from the time-domain. However, from the trivial fact that

$$G_p^T H_p = I = H_p^T G_p \tag{74}$$

Lemma. 64 is more useful than Lemma. 1 in that given an arbitrary set of FIR filters one can easily check if there exits an FIR synthesis bank. Besides, one can start with an arbitrary polynomial matrix $P(z)$, whose determinant is of the form Kz^k, set $H_p(z) = P(z), G_p^T(z) = P^{-1}(z)$ to construct an FIR filter bank. A special case occurs when $\det H_p(z) = K$, a constant. Such matrices are called *unimodular* matrices and can be constructed as the product of elementary polynomial matrices [26]. For such a $P(z)$, both $H_p(z)$ and $G_p(z)$ are causal.

Finally, notice that given an arbitrary Laurent polynomial matrix that is invertible on the unit circle, with appropriate multiplication by z^{-N} for large N one can also transform it into a polynomial matrix in z^{-1}, then do an allpass/minimum-phase factorization to obtain two other filter banks as before.

All of these approaches to construct filter banks suffer from one problem. It is not easy to impose design constraints on the analysis/synthesis filters. In the case of *allpass/minimum-phase* factorization, we have no control even over the polyphase matrix $H_p(z)$. This problem can be eliminated if we can construct arbitrary minimum phase $H(z)$ directly (without factorization) and can be achieved [**?**]. In the case of unitary filter banks a fairly complete answer to this problem is known from the elegant factorization theory for elements of $U_{M\times M}$ matrices. By the degree (or McMillan degree) of a rational matrix, we mean the smallest number of delays required to implement it as the transfer function of a linear system [26]. It turns out that for a unitary matrix the McMillan degree is simply $deg(\det H_p(z))$. We now summarize the factorization results for rational unitary matrices of McMillan degree N.

Let $v \in \mathbf{C}^M$, with $\|v\| = 1$. Define for some $a \in \mathbf{C}$

$$P(z) = \left[I - v\tilde{v} + \frac{-a^\star + z^{-1}}{1 - az^{-1}} v\tilde{v} \right] \tag{65}$$

where $\tilde{v}$ denotes the transpose-conjugate of v. $P(z)$ is of degree 1 and it can be easily verified that

$$\tilde{P}(z)P(z) = I \tag{66}$$

where $\tilde{P}(z) = P^T(z^{-1})$. and furthermore if

$$P_i(z) = \left[I - v_i\tilde{v_i} + \frac{-a_i^\star + z^{-1}}{1 - az^{-1}} v_i\tilde{v_i} \right]. \tag{67}$$

Then the product $P_N P_{N-1} \ldots P_1$ is of McMillan degree N and is unitary. Up to a unitary constant matrix, it turns out that this is the most general rational, inner matrix of McMillan degree N

Lemma 4 *Let U_N be an arbitrary stable, rational, unitary matrix of McMillan degree N. Then, there exists a factorization of the form*

$$U_N(z) = P_N(z)P_{N-1}(z) \ldots P_1(z)U_0 \tag{68}$$

where U_0 is a constant unitary matrix where the P_i are of the form given in Eqn. 65. In particular, if $U_N(z)$ is a polynomial matrix in z^{-1}, the factors P_i can can be shown to be of the form

$$P_i(z) = \left[I - v_i\tilde{v}_i + z^{-1}v_i\tilde{v}_i \right]. \tag{69}$$

on the unit circle. This minimum phase rational functions with no zeros on the unit circle have causal (minimum-phase) rational inverses. Now let $RH^{\infty}_{M\times M}$ denote the set of $M \times M$ matrices with entries in RH^{∞}. A matrix $P \in RH^{\infty}_{M\times M}$, is said to be *unitary*, if it is unitary on the unit circle (such functions are sometimes called *inner* or *lossless*). We have,

$$\tilde{P}(z)P(z) = I_M = P(z)\tilde{P}(z) \tag{62}$$

where $\tilde{P}(z) = P^T(z^{-1})$. This generalizes the notion of a allpass function. A matrix $Q \in RH^{\infty}_{M\times M}$, is said to be *minimum phase*, if its inverse is also in $RH^{\infty}_{M\times M}$. This generalizes the notion of a minimum-phase function. Matrices in $RH^{\infty}_{M\times M}$ also have an *allpass/minimum-phase* factorization [14, 23].

Lemma 2 *Let $H \in RH^{\infty}_{M\times M}$. Then there exists an all pass $H_{all}(z)$ and minimum phase $H_{min}(z)$ such that*

$$H(z) = H_{all}(z)H_{min}(z) \tag{63}$$

Proof : See Francis. [14] □ Let $U_{M\times M}$ denote the set of $M \times M$ unitary (allpass) functions. Clearly $U_{M\times M} \in GL^{\infty}_{M\times M}$. Therefore, given an arbitrary matrix $P \in U_{M\times M}$ one can construct a filter bank as follows. Take $H_p(z) = P(z)$ and $G_p(z) = P(z^{-1}$. Since $P(z)$ is analytic outside the unit disc, $P(z^{-1})$ is analytic inside the unit disc and therefore the analysis filters are causal and the synthesis filters are anti-causal. However, if the synthesis filters are *finitely non-causal*, then the filter bank is realizable. This occurs precisely when $P(z)$ is a polynomial in z^{-1}. We will shortly, see how to construct all realizable unitary matrices $P(z)$. Also notice that given a minimum-phase $Q(z) \in RH^{\infty}_{M\times M}$, if $\det Q$ has no zeros on the unit circle $Q \in GL^{\infty}_{M\times M}$ and therefore by taking $H_p(z) = Q(z)$ and $G_p(z) = Q^{-1}(z)$ we have a filter bank. In this case since both Q and Q^{-1} are analytic outside the unit disc, we have a *causal*, rational and hence *realizable* filter bank.

From these observations we can construct a number of filter banks with the following recipe. Start with an arbitrary $H \in RH^{\infty}_{M\times M} \cap GL^{\infty}_{M\times M}$ Find its allpass and minimum-phase factors H_{all} and H_{min} and then construct two filter banks, one possibly non-realizable. In order to ensure that both are realizable we only have to start with a polynomial in z^{-1} for $H(z)$.

Another simpler way to impose realizability is to require that the analysis and synthesis filter are FIR. In that case,$H_p(z)$ and $G_p(z)$ are Laurent polynomials. We have the following result that characterized FIR filter banks.

Lemma 3 *h_i and g_i form an FIR filter bank iff one of the following is true*

1. *$H_p(z)$ or $G_p(z)$ is invertible over the Laurent polynomials*
2. *$\det H_p(z)$ or $\det G_p(z)$ is of the form Kz^k for some integer k and constant K*

Proof: From Eqn. 59, $G_p^T(z) = H^{-1}(z)$ and hence

$$G_p^T(z) = \frac{1}{\det H_p(z)}\mathrm{adj}(H_p(z)) \tag{64}$$

Since the adjoint of a Laurent polynomial is a Laurent polynomial, Eqn. 64 holds with $G_p^T(z)$ a Laurent polynomial iff $\det H_p(z) = Kz^k$ for some integer k. □

satisfied these properties). Reinterpreting results from classical inverse scattering theory of two-port networks ([1]) the $\mathcal{Z}$-transform domain, , Vaidyanathan et. al., ([44]), transformed the factorization theorems for *lossless* scattering matrices to factorization theorems for rational matrix functions, unitary on the unit circle. Since then, *unitariness* has played a fundamental role in developing filter banks with desirable properties. In any case, from Eqn. 59, it is obvious that having perfect reconstruction filter banks is equivalent to the inversion of H_p and hence definitely feasible. From now on by a filter bank we always mean a perfect reconstruction filter bank.

We now give the most fundamental result in filter bank theory. An $M \times M$, matrix function of the complex variable z, that has full rank for all z on the unit circle (i.e $z = e^{i\omega}$) is said to be in $GL^{\infty}_{M\times M}$. By the standard abuse of notation, for a $\mathcal{Z}$-transform $H(z)$, we will denote $H(e^{i\omega})$ by $H(\omega)$.

Lemma 1 *Let $h_i(n)$ denote the the analysis filter bank. Assume that for all i, $H_i(\omega) < \infty$. Then, there exists a synthesis filter bank with filters $g_i(n)$ iff, $H_p(\omega) \in GL^{\infty}_{M\times M}$, or equivalently if* $\det H_p(\omega) \neq 0$ *for all ω*

Proof : If $H_p(\omega) \in GL^{\infty}_{M\times M}$, then it has full rank for all ω, is hence invertible and therefore $H_p^{-1}(\omega)$ is well defined. Take $G_p(\omega) = [H_p^{-1}]^T(\omega)$. From $G_p(\omega)$, the numbers $g_{i,j}(n)$ and hence $g_i(n)$ can be obtained. □

How does this result help in the construction of filter banks? Clearly for a given choice of $h_i(n)$, it is no easy task (in general) to check that $H_p(\omega)$ has full rank for all ω. So the trick is to start with an arbitrary $H_p(\omega) \in GL^{\infty}_{M\times M}$, and then construct $h_i(n)$ and $g_i(n)$ from it. One problem with this approach is that there is no explicit control over the filters $h_i(n)$ and $g_i(n)$, only over $h_{i,j}$ and $g_{i,j}$. A complete solution to this problem is known only in the special case of some (*unitary*) FIR filter banks. Nevertheless, this method gives rise to many filter banks and in some cases wavelet bases as we will shortly see. Another problem with this method is that the filters may not be realizable (causal). If the filters $h_{i,j}$ are causal then necessarily the filters h_i are causal. But if the filters $g_{i,j}$ are causal it does not that the filters g_i are as can be seen from Eqn. 49. However, the filters g_i are finitely non-causal and hence realizable. From now on by causality of the filters we mean that $h_{i,j}$ and $g_{i,j}$ are causal.

We now proceed to show how realizability can be imposed as a constraint in the filter bank. We need some preliminary mathematical results. Let RH^{∞}, denote the set of real, causal (stable), rational functions. In other words, real rational functions, analytic in the exterior of the closed unit disc in the complex plane. Given a function $H(z) \in RH^{\infty}$, by reflecting all the zeros of the function outside the unit circle, inside the circle, it is well known that $H(z)$ has an *allpass/minimum-phase* factorization.

$$H(z) = H_{all}(z)H_{min}(z) \tag{60}$$

The function $H_{min}(z)$ has all its poles and zeros inside the unit circle and is called *minimum phase* because for all other functions in RH^{∞} that have the same magnitude on the unit circle, $H_{min}(z)$ has the least phase. The function $H_{all}(z)$ is called allpass because, its magnitude on the unit circle is unity.

$$\left[H_{all}^T(z^{-1})H_{all}(z)\right]_{|z|=1} = 1 \tag{61}$$

The inverse of a minimum-phase function is also in RH^{∞} since all its zeros are inside (or on) the unit circle. If the function has no zeros on the unit circle then it is in $GL^{\infty}_{1\times 1}$ and therefore invertible

Equating like powers of z we have

$$Y_k(z) = \sum_i G_{i,k}(z) D_i(z). \tag{54}$$

To simplify notation and make things transparent, define the *polyphase* component matrices

$$H_p(z) = \begin{bmatrix} H_{0,0}(z) & H_{0,1}(z) & \cdots & H_{0,M-1}(z) \\ H_{1,0}(z) & H_{1,1}(z) & \cdots & H_{1,M-1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ H_{M-1,0}(z) & H_{M-1,1}(z) & \cdots & H_{M-1,M-1}(z) \end{bmatrix} \tag{55}$$

$$G_p(z) = \begin{bmatrix} G_{0,0}(z) & G_{0,1}(z) & \cdots & G_{0,M-1}(z) \\ G_{1,0}(z) & G_{1,1}(z) & \cdots & G_{1,M-1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ G_{M-1,0}(z) & G_{M-1,1}(z) & \cdots & G_{M-1,M-1}(z) \end{bmatrix}. \tag{56}$$

Let

$$\mathbf{x}_p(z) = [X_0(z), X_1(z), \ldots, X_{M-1}(z)]^T$$
$$\mathbf{y}_p(z) = [Y_0(z), Y_1(z), \ldots, Y_{M-1}(z)]^T$$
$$\mathbf{d}(z) = [D_0(z), D_1(z), \ldots, D_{M-1}(z)]^T.$$

Then Eqn. 52 and Eqn. 53 can be written compactly as

$$\mathbf{d}(z) = H_p(z)\mathbf{x}_p(z) \tag{57}$$

$$\mathbf{y}_p(z) = G_p^T(z)\mathbf{d}(z) \tag{58}$$

$$\mathbf{y}_p(z) = G_p^T(z)H_p(z)\mathbf{x}_p(z) \tag{59}$$

These equations are perhaps the most important equations in filter bank theory. It shows that the filter bank could be thought of as the cascade of two multi-input multi-output (MIMO) linear time-invariant systems, that are inverses of each other. Notice that Eqn. 57 and Eqn. 58 hold independent of whether the h_i and g_i are causal, realizable, rational etc. H_p and G_p represent MIMO systems corresponding to the analysis and synthesis filter bank respectively. Moreover given an arbitrary MIMO system, there exists an analysis filter bank and a synthesis filter bank, of which the MIMO system could be considered to be the polyphase representation of. Seen in the MIMO setting the basic results in filter bank theory become extremely simple. To appreciate the simplicity and elegance of the equation, an understanding of the history of filter bank theory is essential. Filter banks (actually two channel case) were first introduced in 1977 by Galand and Esteban [12]. For a long time, no non-trivial examples of *realizable* perfect reconstruction filter banks were known (for a a history of filter banks see [43]). The first successes were in affecting alias cancellation (in which case $y(n) = x(n) * h(n)$, for some sequence $h(n)$ that is not $\delta(n)$), and in obtaining exact magnitude responses ($|Y(\omega)| = |X(\omega)|$) and phase responses. Finally,in 1984 that Smith and Barnwell constructed exact reconstruction filter banks [42]. Later Vaidyanathan noticed the importance of a certain *unitariness* property in filter banks (the Smith-Barnwell examples

$Y(z)$ be rewritten as

$$X(z) = \sum_{k=0}^{M-1} z^k X_k(z^M) \tag{46}$$

$$Y(z) = \sum_{k=0}^{M-1} z^k Y_k(z^M) \tag{47}$$

Let $H_i(z)$ and $G_i(z)$ denote the $\mathcal{Z}$-transforms of the analysis and synthesis filters. Then $H_i(z)$ and $G_i(z)$ (possibly non-causal and irrational) can be written as

$$H_i(z) = \sum_{k=0}^{M-1} z^{-k} H_{i,k}(z^M) \tag{48}$$

$$G_i(z) = \sum_{k=0}^{M-1} z^k G_{i,k}(z^M) \tag{49}$$

where

$$H_{i,k}(z) = \sum_n h_i(Mn+k) z^{-n} \tag{50}$$

$$G_{i,k}(z) = \sum_n g_i(Mn-k) z^{-n}. \tag{51}$$

Now from Fig. 3, we have

$$\begin{aligned} D_i(z) &= [\downarrow M] H_i(z) X(z) \\ &= [\downarrow M] \left[\sum_{k=0}^{M-1} z^{-k} H_{i,k}(z^M) \right] \left[\sum_{j=0}^{M-1} z^j X_j(z^M) \right] \\ &= \sum_{k=0}^{M-1} H_{i,k} X_k. \end{aligned} \tag{52}$$

The last step, follows from the fact that

$$[\downarrow M] z^k = \begin{cases} z^m \text{ for } k = mM \\ 0 \text{ otherwise.} \end{cases}$$

Similarly, for the synthesis bank,

$$\begin{aligned} \sum_i z^k Y_k(z^M) = Y(z) &= \sum_i \{[\uparrow M] D_i(z)\} G_i(z) \\ &= \sum_i D_i(z^M) G_i(z) \\ &= \sum_{i,k} z^k D_i(z^M) G_{i,k}(z^M) \\ &= \sum_k z^k \left[\sum_i D_i(z^M) G_{i,k}(z^M) \right]. \end{aligned} \tag{53}$$

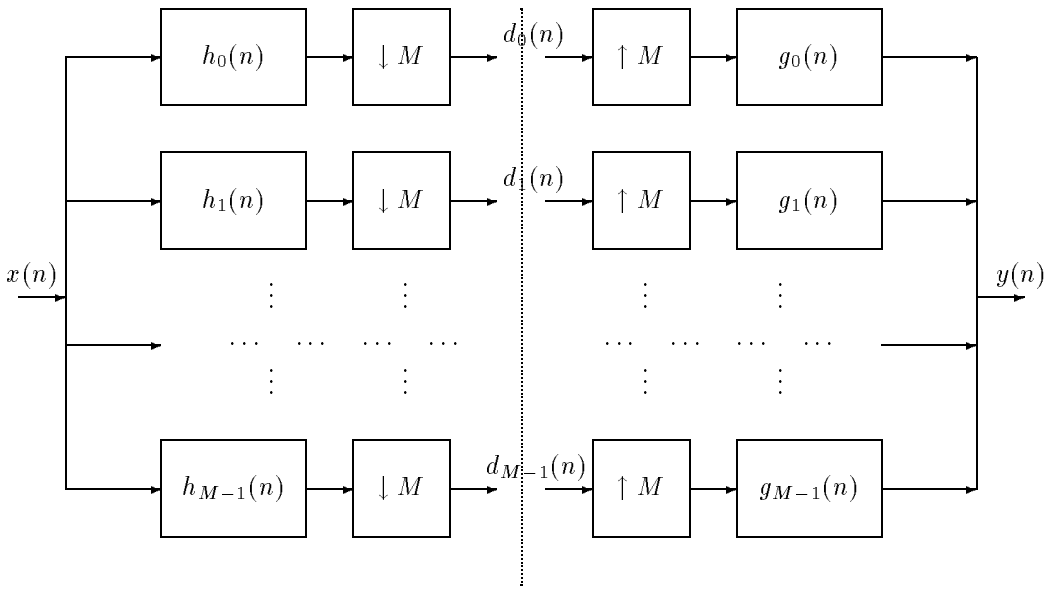

Figure 6: An M Channel Filter Bank

(preferably realizable) such that $y(n) = x(n)$. Clearly for arbitrary realizable $h_i(n)$, there may not exist $g_i(n)$, such that $y(n) = x(n)$. And even if $g_i(n)$ did exist, they may not be realizable. As far as filter banks in wavelet theory are concerned, from Section. 2, it should be clear that realizability of the filters is not an issue in the construction of wavelet bases. Hence for wavelet theory it is perhaps necessary, even to consider non-realizable filter banks. In any case, from a computational perspective, exact signal analysis is not possible in wavelet bases associated with non-realizable filter banks, and hence it is definitely important to know the theory of realizable filter banks. Not surprisingly, most of filter bank theory has dealt with only a special class of *realizable* filters, namely FIR filters.

Most of the literature in filter bank theory specializes to FIR $h_i(n)$ and $g_i(n)$ partly because FIR FIR filters can be easily implemented, and partly because stability issues do not have to be addressed. Surprisingly, FIR filter bank theory includes all orthogonal and biorthogonal wavelet bases with compactly supported wavelets as a special case.

We now review the main results in filter bank theory introducing some new approaches to filter bank theory on the way. The question of existence of perfect reconstruction filter banks can be reduced to the question of invertibility of of multi-input multi-output (MIMO) linear shift-invariant system. If we block $x(n)$ and $y(n)$ into M-vectors, then from Fig. 3 it is obvious that the analysis bank maps an M-vector signal to another M-vector signal ($d_i(n)$), and the synthesis bank maps this M-vector signal yet another M-vector signal. More precisely, let the input $X(z)$ and output

3 Filter Bank Theory

Filter banks are efficient convolution structures that have been extensively used in sub-band coders for speech, and in transmultiplexers (devices to convert time-division multiplexed data to frequency-division multiplexed data and vice-versa). In a filter bank, a data sequence $x(n)$ (see Fig. 3), is decomposed into M channels by convolution with sequences $h_i(n), i = 1, 2, \ldots, M-1$, called the *analysis* filters, down-sampled by M on each channel, upsampled by M on each channel, and then convolved with the sequences $g_i(n), i = 1, 2, \ldots, M-1$, called the *synthesis* filters, and recombined to give $y(n)$. A potential application is in speech coding, where the sequences $d_i(n)$ are encoded with different number of bits depending on the sensitivity of the ear to channel i. A fundamental question is whether there exist perfect reconstruction filter banks (i.e $x(n) = y(n)$). Recall that in Section. 2, the wavelet and scaling vectors give rise to a two channel ($M = 2$) perfect reconstruction filter bank! In that case, we also had that the sequences h_i and g_i were time reverses of each other.

$$h_i(k) = g_i(-k) \qquad H_i(z) = G_i(z^{-1}) \tag{42}$$

In order to analyze the filter bank we introduce some standard signal processing definitions and notation. For a sequence $h(k)$ (also called a filter) the bilateral $\mathcal{Z}$-transform (in its region of convergence) is defined to be

$$H(z) = \sum_k h(k) z^{-k}. \tag{43}$$

Usually the domain of convergence is assumed to include the unit circle so that the Discrete Time Fourier Transform of the $h(n)$ is defined [35, 34]. If $h(n) \in \ell^1(\mathbf{Z})$, then

$$\sum_k |h(k)| < \infty, \tag{44}$$

and the unit circle is in the region of convergence. Most of the filters we consider will be in $\ell^1(\mathbf{Z})$. A filter $h(k)$ is said to be *causal* if $h(n) = 0$ for negative n. This means that $H(z)$ is analytic in the exterior of the closed unit disc since $H(z)$ is a power series in z^{-1}. Every filter has a *causal-anticausal* decomposition

$$H(z) = H_c(z) + zH_a(z^{-1}) \tag{45}$$

where $H_c(z)$ and $H_a(z)$ are causal (that is the corresponding sequences $h_c(n)$ and $h_a(n)$ are causal). A non-causal filter is said to be *finitely non-causal*, if there exists an N such that $h(n-N)$ is causal. An equivalent characterization for finite non-causality is that $H_a(z)$ is a polynomial in z^{-1}. A filter $h(k)$ is said to be FIR, if $H_a(z)$ and $H_c(z)$ are polynomials in z^{-1}. In other words, $H(z)$ is a Laurent polynomial (sum of a polynomial in z and z^{-1}). Clearly all FIR filters are finitely non-causal. A filter $h(k)$ is said to be IIR if it is not FIR, The $\mathcal{Z}$-transform of an IIR filter is a Laurent series that is not a Laurent polynomial. A filter is said to be *rational* if $H_a(z)$ and $H_c(z)$ are rational functions of z. A filter is said to be *realizable* if convolution with the sequence $h(n)$ can be implemented. Only *finitely non-causal*, *rational* filters are realizable. Thus for a realizable filter, $H_a(z)$ is a polynomial, and $H_c(z)$ is a rational. The scaling and wavelet vectors in Section. 2 are neither finitely causal, nor rational and hence are not realizable. The standard problem in filter bank theory is to start with a realizable set of filters $h_i(n)$, with desired frequency behavior, and then to determine $g_i(n)$

out ad infinitum. For example, in the DWT decomposition, the low-frequency channel V_j i ndeed decomposed infinitely many times.

In Section. 8 we will show that when bins are split recursively, as described above, one can do so using other sequences h_0 and h_1 that form CQFs. In fact, we will have a variety of efficient decompositions that have the computational advantages of the wavelet type decomposition, but there will be no underlying *scaling function* and *wavelet*. While the recursive decomposition or *scaling* described above that is at the heart of things, and the existence of a scaling function and wavelet is wavelet is intimately related to whether these decompositions can be carried out ad infinitum.

Another feature in the DWT decomposition in Fig. 2.2 is that V_j is split into *two* subspaces corresponding to the split of $[0, 2^j\pi]$, into $[0, 2^{j-1}\pi]$ and $[2^{j-1}pi, 2^j\pi]$. At each stage the bin corresponding to V_j could have been split into M sub-bins and then we would have a decomposition as shown in Fig. 5. The low-frequency sub-bin is further split into M bins and so on thus introducing the concept of *scale*. Such a decomposition is called a multiplicity M wavelet decomposition. If $W_{i,j}$ denotes the space of functions bandlimited to $[iM^j, (i+1)M^j]$, then it is clear that for all j, we have a decomposition of $W_{0,j}$ into M sub-channels $W_{0,j-1}$ through $W_{M-1,j-1}$. As in the two channel case there would be sequences $h_i(n)$ $i = 0, \ldots, M-1$, relating the basis for $W_{0,j}$ and the basis for $W_{i,j-1}$. These sequences also relate the expansion coefficients as in Eqn. 36 and Eqn. 37. The corresponding *computational* structure relating the expansion coefficients in these bases is an M-channel filter bank associated with the sequences $h_i(n)$.

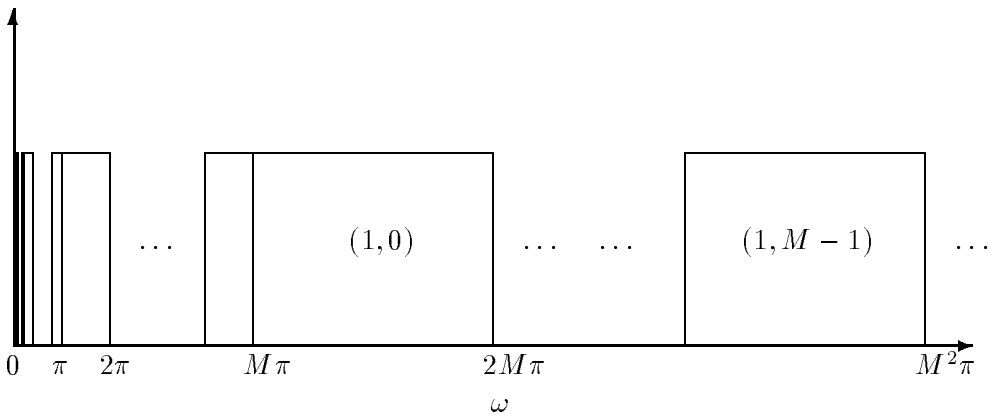

Figure 5: Multiplicity M-DWT decomposition

The wavelet packets idea can be combined with multiplicity M DWT decompositions in order to give a very rich family of computationally efficient decompositions. In the next two sections we describe the essence of filter bank theory in relation to wavelet theory, and describe how computations can be performed in filter banks. The theory developed will be used to introduce, orthonormal and biorthogonal wavelet bases, far-reaching generalizations of DWT type frequency decompositions.

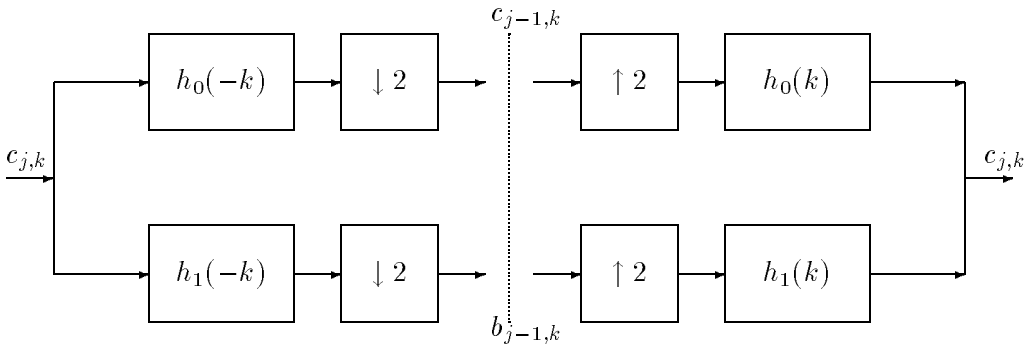

Figure 4: DWT Decomposition and the Filter Bank Structure

1. The coefficients $a_{j,k}$ are real for real signals.

2. There is a natural recursive structure associated with the DWT decomposition that can be used to compute the DWT coefficients recursively. The reason for this is the scaling property of the basis functions.

3. In the example above the DWT coefficients $b_{j,k}$ can be interpreted as the Nyquist rate samples of $f_j(t)$. Though the general DWT decompositions discussed later do not satisfy this, this intuition is useful.

At this point, an apparent disadvantage with the filter bank structure is that since the sequences $h_i(n)$ are of infinite extent they cannot be realized on a digital computer. We will see how filter bank theory gives a solution to this problem.

A natural question at this point is whether STFT type frequency decompositions can be obtained that have the nice computational properties of DWT type decompositions. To answer this question consider Fig. 2.2. The frequency bin $[0, 2^j\pi]$ corresponds to V_j and is split into two bins corresponding to V_{j-1} and W_{j-1} with V_{j-1} corresponding to the low frequency part of V_j and W_{j-1} the high frequency part. The scaling property of the DWT derives from the fact that V_{j-1} is then split into V_{j-2} and W_{j-2}. In order to get uniform bins as in the STFT type frequency decomposition we only need to also split the high frequency part of V_j (namely W_j) into two uniform width bins. Corresponding to these bins there would be two subspaces of W_j with orthonormal bases for each one of them. Clearly there would be sequences relating these bases to the basis for W_j. It turns out that these sequences are precisely the scaling and wavelet vectors. Thus expansion coefficients in these bases can be obtained by starting with $b_{j,k}$, the expansion coefficients in W_j and then using Eqn. 36 and Eqn. 37, with $c_{j,k}$ replaced by $b_{j,k}$ and the left hand side interpreted as the expansion coefficients in the new bases. Now each of the new subspaces can further be decomposed into two bins, and this process can be continued any number of times. It turns out in all these cases the bases are related by the same sequences h_0 and h_1. The bevy of functions thus obtained are sometimes called *coiflets* [9], after Coifman who first discovered them. They are also sometimes referred to as *wavelet packets* [6]. In some cases, these decompositions could be carried

Notice also that

$$\int \psi(t)\, dt = 0 \tag{33}$$

$$\int \phi(t)\, dt = 1. \tag{34}$$

The big breakthrough for DWT type orthonormal frequency decompositions came about when Ingrid Daubechies parameterized all finite length sequences $h_0(k)$ satisfying Eqn. 27 and Eqn. 32, and showed that many of them give rise to overlapping DWT type orthonormal frequency decompositions. Independently of her efforts, in the theory of Filter Banks (where sequences satisfying Eqn. 90 are called Conjugate Quadrature Filters (CQFs) or Quadrature Mirror Filters (QMFs)), all such sequences had also been parameterized [46].

We will now see how the DWT can be computed using a "filter bank". Thinking of the channel W_j as the orthogonal complement of V_j in V_{j+1} will be useful. Let $c_{j,k}$ denote expansion coefficients of $f(t) \in V_j$.

$$c_{j,k} = \langle \phi_{j,k}, f(t) \rangle \tag{35}$$

Then $c_{j,k}$ is related to $c_{j-1,k}$ and $b_{j,k}$ in terms of h_0 and h_1 as,

$$c_{j-1,k} = \sum_n h_0(k-2n) c_{j,n} \tag{36}$$

$$b_{j-1,k} = \sum_n h_1(k-2n) c_{j,n} \tag{37}$$

$$c_{j,n} = \sum_k h_0(k-2n) c_{j-1,k} + \sum_k h_1(k-2n) b_{j-1,k}. \tag{38}$$

These equations are obtained by invoking Eqn. 23 and Eqn. 24. The computations corresponding to these equations can be represented as in Fig. 4, known as a two channel filter bank. The symbols $\downarrow$ and $\uparrow$ correspond to the dual operations of sub-sampling and upsampling. That is,

$$[\downarrow 2]x(n) = x(2n) \tag{39}$$

$$[\uparrow 2]x(n) = \begin{cases} x(\frac{n}{2}) & \text{for } n \text{ even} \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

Given $c_{j,k}$ for fixed j, the filter bank can be used to compute $c_{j-1,k}$ and $b_{j-1,k}$. and hence given $c_{J,k}$, the DWT $b_{j,k}$ can be computed for all j less than J. Now in practical signal analysis, a reasonable assumption to make is that the signals are approximately bandlimited. That is, f can be assumed to reside in some V_J for large enough J. All information about such signals are in $c_{J,k}$ and the DWT coefficients satisfy

$$b_{j,k} = 0 \quad \text{for } j \geq J\ . \tag{41}$$

Once $c_{J,k}$ is obtained (by taking inner-products or one of the methods discussed in Section. 7), $b_{j,k}$ can be obtained by using the filter bank equations. The filter bank structure is more that just a computational tool. Orthonormal and biorthogonal wavelet bases can be developed with filter bank theory as a starting point and this viewpoint is very useful.

In summary, compared to the DTSTFT, the DWT above has a number of attractive features

V_j is trivially contained in V_{j+1}. In particular, $\phi(t) \in V_0 \subset V_1$ and therefore there exists a sequence $h_0(k)$ such that

$$\phi(t) = \sum_k h_0(k)\phi_{1,k}(t) = \sqrt{2}\sum_k h_0(k)\phi(2t-k) \tag{23}$$

Similarly, $\psi(t) \in W_0 \subset V_1$ and therefore there exists a sequence $h_1(k)$ such that

$$\psi(t) = \sum_k h_1(k)\phi_{1,k}(t) = \sqrt{2}\sum_k h_1(k)\phi(2t-k) \tag{24}$$

Because the function $\phi(t)$ satisfies a *dilation* equation (Eqn. 23), it is called the *scaling* function. The sequences h_0 and h_1 are called the *scaling vector* and *wavelet vector* respectively. From the sampling theorem it is easy to show that

$$h_0(k) = \frac{1}{\sqrt{2}}\phi(\frac{k}{2}) = \sqrt{2}\frac{\sin(\frac{k}{2}\pi)}{k\pi} \tag{25}$$

$$h_1(k) = \frac{1}{\sqrt{2}}\psi(\frac{k}{2}) = [\sqrt{2}\phi(k) - \frac{1}{\sqrt{2}}\phi(\frac{k}{2})] = \sqrt{2}\delta(k) - h_0(k) \tag{26}$$

showing that the wavelet vector is determined by the scaling vector. Since $\{\phi(t-k)\} = \phi_{0,k}(t)$ forms an orthonormal basis for V_0, $\{\psi(t-k)\} = \phi_{0,k}(t)$ forms an orthonormal basis for W_0, and $V_0 \perp W_0$, from Eqn. 23, and Eqn. 24

$$\int \phi(t)\phi(t-l)\,dt = \delta(l) = \sum_k h_0(k)h_0(k+2l). \tag{27}$$

$$\int \psi(t)\psi(t-l)\,dt = \delta(l) = \sum_k h_1(k)h_1(k+2l). \tag{28}$$

$$\int \phi(t)\psi(t-l)\,dt = 0 = \sum_k h_0(k)h_1(k+2l) \tag{29}$$

Thus all the orthogonality properties of the scaling function and wavelet are reflected in the scaling and wavelet vectors. In fact more is true. The scaling vector $h_0(n)$, by itself, completely determines the DWT frequency decomposition. To see this take the Fourier transform of Eqn. 23 and let $H_0(\omega)$ denote the Fourier transform of $h_0(n)$.

$$\hat{\phi}(\omega) = \left[\frac{1}{\sqrt{2}}H_0(\frac{\omega}{2})\right]\hat{\phi}(\frac{\omega}{2}) \tag{30}$$

or equivalently,

$$\hat{\phi}(\omega) = \hat{\phi}(0)\prod_{j=1}^{\infty}\left[\frac{1}{\sqrt{2}}H_0(\omega/2^j)\right] \tag{31}$$

Thus if the infinite product converges, by the uniqueness of the Fourier transform ϕ and hence ψ and therefore the DWT frequency decomposition is determined completely by $h_0(k)$. A necessary condition for the convergence of the infinite product is that $H_0(0) = \sqrt{2}$, or equivalently

$$\sum_k h_0(k) = \sqrt{2}. \tag{32}$$

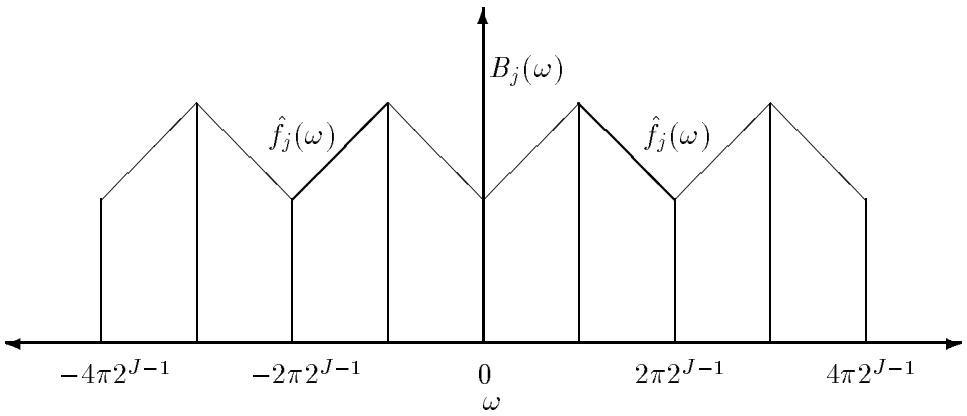

Figure 3: Periodization of $\hat{f}_j(\omega)$

$$\phi(t) = \frac{sin(\pi t)}{\pi t} \tag{19}$$

$$a_{j,k} = \sqrt{2^{-j}} b_{j,k} \tag{20}$$

Eqn. 16 is the *wavelet* expansion of $f(t)$ with respect to the *wavelet* ψ. The numbers $a_{j,k}$ are called the to as the DWT coefficients. Notice that Eqn. 16 is exactly in the form of Eqn. 7 and that the functions form an $\psi_{j,k}$ ON family. The essential difference is that now, the functions $w_{j,k}$ are translates and dilates of the wavelet w. From orthonormality, the DWT coefficients are given by

$$a_{j,k} = \langle f(t), \psi_{j,k} \rangle \,. \tag{21}$$

For this simple wavelet $a_{j,k} = f_j(2^{-j}k)$, and hence the numbers can also be interpreted as the "Nyquist rate" samples of each of the frequency channels. This interpretation is generic for all DWT type frequency decompositions.

So far we have seen that the STFT and the DWT decompositions are very similar in the non-overlapping frequency decomposition case. We now show that, the *scaling* property implies existence of discrete sequences associated with the frequency decomposition that contain all the information about the decomposition. Furthermore, this scaling structure also gives an efficient means to compute $a_{j,k}$.

Notice that the wavelet $\psi(t)$ is the difference of the function ϕ in Eqn. 19 and a scaled version of it. The function ϕ is bandlimited to $[0, \pi]$, and $\psi(t)$ is the difference of ϕ and a scaled version of ϕ. Let V_j denote the space of signals band-limited to $[0, 2^j\pi]$. Define

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k) \tag{22}$$

Then for fixed j, $\{\phi_{j,k}\}$ is an orthonormal basis for V_j (the sinc basis). Also for fixed j, $\{\psi_{j,k}\}$ is an orthonormal basis for W_j the space of bandlimited functions in $[2^j\pi, 2^{j+1}\pi]$. We have already seen that $\psi_{j,k}$ forms an orthonormal basis for W_j. Now from Fig. 2.2, it is clear that any function in V_{j+1} can be decomposed into a function in V_j and a function in W_j. Also any function in W_j or

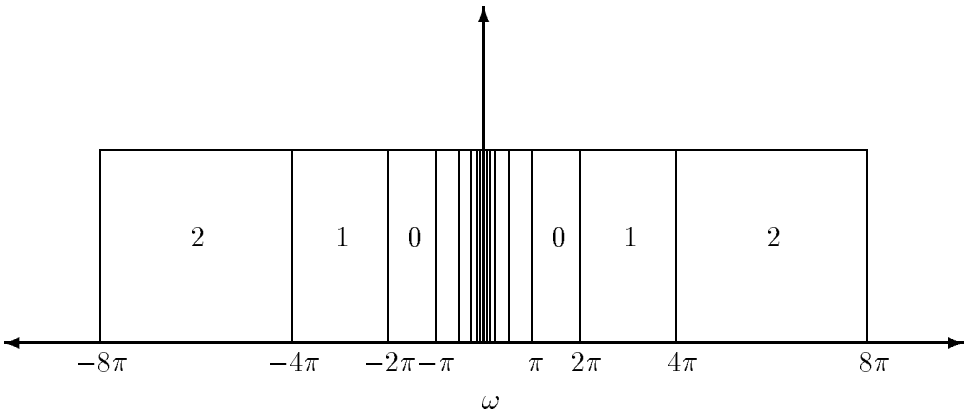

Figure 2: Frequency Channels in a Wavelet Decomposition

indicates that the retinal image is decomposed spatially-oriented frequency channels each having the same bandwidth on a log scale [31, 30]. That is, the retinal system, splits the visual input into information at different scales. Furthermore, in many naturally occurring classes of signals like transients, it has been observed that the larger the center frequency of the signal, the smaller its duration, and hence greater its bandwidth. Such a frequency behavior has been found to be useful in the modeling of nearly $\frac{1}{f}$ noise processes [48]. And moreover, from a practical point of view, unlike the DSTFT type of frequency decomposition, this *wavelet* type frequency decomposition, allows of non-trivial and interesting examples, orthogonal and otherwise. The scaling structure is associated with efficient computational structures. Notice also that the channels are symmetric with respect to the origin, and this will make the DWT representation real for real signals.

As in the STFT case, the restrictions of $\hat{f}(\omega)$ to each of the bins can be expanded as a Fourier series. But then, the coefficients $a_{j,k}$ obtained would in general be complex (for real $f(t)$) and it turns out that one can do better. Let $b_{j,k}$ be the Fourier series coefficients of $B_j(\omega)$, the periodization of $\hat{f}_j(\omega)$ by period $2^j\pi$ (see Fig. 3)

$$B_j(\omega) = \sum_k \hat{f}_j(\omega + 2\pi 2^j k). \tag{13}$$

Notice that by the sampling theorem $b_{j,k} = f_j(2^{-j}k)$. Then,

$$f_j(\omega) = B_j(\omega)(\sqcap(\frac{\omega}{2^{j+1}\pi}) - \sqcap(\frac{\omega}{2^j\pi})) \tag{14}$$

$$f_j(t) = \sum_k b_{j,k}(2^{j+1}\phi(2^{j+1}(t-k)) - 2^j\phi(2^j(t-k))) \tag{15}$$

$$f(t) = \sum_j f_j(t) = \sum_{j,k} a_{j,k}\psi_{j,k}(t) \tag{16}$$

where

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{17}$$

$$\psi(t) = 2\phi(2t) - \phi(t) \tag{18}$$

$w(t)$ [15, 17]. Since the Fourier basis is ON, by choice of non-overlapping bins the functions $w_{j,k}$ are ON. Thus the DSTFT can be computed by taking inner-products with the basis functions.

$$a_{j,k} = \langle f(t), w_{j,k} \rangle \tag{10}$$

In general orthogonal DSTFT type frequency decompositions can be viewed as generalizations of this decomposition where the frequency channels are non-overlapping, yet mutually orthogonal. The essential idea is the window function $w(t)$, such that modulates and translates $(w(t+k)e^{iw\pi jt})$ form a basis. For fixed j, $\{w_{j,k} | k \in \mathbf{Z}\}$ spans the j^{th} frequency channel which is a frequency shift of the 0^{th} channel. This is a distinguishing feature of the DSTFT. It is quite conceivable that one can have an orthogonal frequency decomposition where the bins are approximations to that in Fig. 2.1, and yet the channels are *not* frequency shifts of each other. In other words, there is no window $w(t)$ such that $w_{j,k}$ spans $L^2(\mathbb{R})$. Clearly such a frequency decomposition is very similar to the STFT, in that the channels are of uniform effective bandwidth and hence useful. Later we will see how to construct many such frequency decompositions. Even though orthonormal DSTFT bases can be constructed [8, 32, 10], there is no general technique to obtain such decompositions with windows that have both desirable localization properties and good numerical algorithms associated with them. Usually one starts with a desirable $w(t)$ ($w_{j,k}$ not necessarily orthonormal) [5, 15], and hopes that expansions like Eqn. 7 hold. Since the basis functions $w_{j,k}$ are no longer orthonormal, Eqn. 10 does not hold and hence cannot be used to compute the DSTFT coefficients$a_{j,k}$. However, for most choices $w(t)$ there exists a function $\tilde{w}(t)$, the dual window, such that,

$$a_{j,k} = \langle f(t), w_{j,k} \rangle \,. \tag{11}$$

The set of functions $w_{j,k}$ and $\tilde{w}_{j,k}$ are referred to as a *biorthogonal* basis, which means no more than that a function can be projected onto one set and expanded in the other to recover it.

Incidentally, in the sinc window case, the coefficients,$a_{j,k}$ are seen to be the samples of samples of j^{th} frequency channel and hence there is no need to compute inner products.

$$a_{j,k} = f_j(k) \tag{12}$$

Moreover, for general ON orthogonal or biorthogonal DSTFT type frequency decompositions, the numbers $a_{j,k}$ can be interpreted as generalized samples of the j^{th} frequency channel (up to a constant).

In summary, the DSTFT arises from an frequency decomposition, where the frequency channels are of the same bandwith and are shifts of each other.

2.2 The DWT Decomposition

The idealized choice of frequency bins for the DWT type frequency decomposition is shown in Fig. 2.2. An important feature of this decomposition is that if $f(t)$ is in the j^{th} bin (that is $\hat{f}(\omega)$ is zero outside the j^{th} bin), then $f(2t)$ is in the $j+1^{st}$ bin. Just as channels in the DSTFT case were related by modulates, here the channels are related by *scaling* by factors of 2, (called the multiplicity of the DWT frequency decomposition). Also notice that the bandwith of the jth bin is twice the bandwidth of the $(j-1)^{st}$ bin, and hence bandwidths are constant on a logarithmic scale. Why are such frequency decompositions interesting? The physiology of early mammalian vision

The DSTFT decomposition is associated with a basis made up of translates and modulates of a single function called the *window* function. All translates of a fixed modulation of the window, span a specific channel. The idealized choice of frequency bins for the DSTFT type frequency decomposition is shown in Fig. 2.1. For this choice of frequency bins, we can construct an

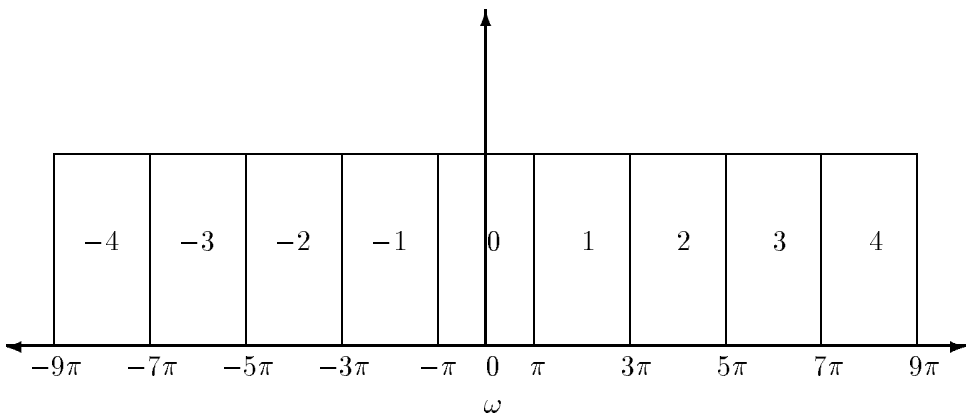

Figure 1: Frequency Decomposition for Short Time Fourier Transform

orthonormal basis for each channel, and thus an orthonormal basis for $L^2(\mathbb{R})$ as follows. For an arbitrary $f(t)$, let its Fourier transform restricted to the j^{th} bin be represented as a Fourier series with coefficients $a_{j,k}$. Then,

$$\hat{f}_j(\omega) = \left[\sum_k a_{j,k} e^{i\omega k}\right] \sqcap \left(\frac{\omega - j2\pi}{2\pi}\right) \tag{4}$$

where

$$\sqcap(\omega) = \begin{cases} 1 \text{ for } -\frac{1}{2} \leq \omega \leq \frac{1}{2} \\ 0 \text{ otherwise.} \end{cases} \tag{5}$$

Then,

$$\hat{f}(\omega) = \sum_{j,k} a_{j,k} e^{i\omega k} \sqcap \left(\frac{\omega - j2\pi}{2\pi}\right). \tag{6}$$

Taking the inverse Fourier transform on both sides

$$f(t) = \sum_{j,k} a_{j,k} w_{j,k}(t) \tag{7}$$

where

$$w(t) = \frac{sin(\pi t)}{\pi t} = sinc(t) \tag{8}$$

$$w_{j,k}(t) = w(t+k) e^{i2\pi jt} \tag{9}$$

Eqn. 7 is precisely, the DSTFT expansion of $f(t)$ with respect to the window $w(t)$, sometimes called the *sinc* window. Then sequence $a_{j,k}$ is referred to as the DSTFT of $f(t)$ with respect to

frequency decomposition splits it into a number of frequency channels, we have, $\hat{f}_j(\omega)$, such that

$$\hat{f}(\omega) = \sum_j \hat{f}_j(\omega) \quad \text{and} \quad f(t) = \sum_j f_j(t). \tag{1}$$

If the channels are mutually orthogonal, then we call it an orthogonal frequency decomposition.

$$\int f_i(t) f_j^*(t)\, dt = \frac{1}{2\pi} \int \hat{f}_i(\omega) \hat{f}_j^\star(t)\, d\omega = \delta_{i-j} \tag{2}$$

Some orthogonal frequency decompositions can be constructed easily. For example, the the frequency axis could be split into *non-overlapping* bins and $\hat{f}_j(\omega)$ could be the restriction of $\hat{f}(\omega)$ to the j^{th} bin. Since each of the channels then, become ideal "bandpass" channels, the decomposition is not realizable practically. Thus the problem is to come up with realizable frequency decompositions with useful properties. If some structure is introduced into the frequency decomposition, then it is possible to construct orthogonal and non-orthogonal frequency decompositions. This section will examine structured frequency decompositions like the Discrete Short Time Fourier Transform (DSTFT) type frequency decomposition and the Discrete Wavelet Transform (DWT) type frequency decomposition. Since the essential features of these frequency decompositions are captured in the non-overlapping frequency channels case, we examine this case only.

Since discrete data is easily stored and manipulated, it is desirable to have discrete representations of all signals in a frequency decomposition. Discrete representations could be obtained by sampling, or by expansion in bases. Given an orthogonal frequency decomposition, we can construct orthonormal (ON) bases for $L^2(\mathbb{R})$, and would give a discrete representation. For example, for fixed j, let $\{\psi_{j,k} | j, k \in \mathbf{Z}\}$, be an orthonormal family that spans the j^{th} channel. This means that there exist numbers $a_{j,k}$ such that

$$f_j(t) = \sum_k a_{j,k} \psi_{j,k}(t) \quad \text{and} \quad \hat{f}_j(\omega) = \sum_k a_{j,k} \hat{\psi}_{j,k}(\omega) \tag{3}$$

Then $\{\psi_{j,k} | j, k \in \mathbf{Z}\}$ is an orthonormal basis for $L^2(R)$. It is not obvious that non-orthogonal frequency decompositions with discrete representations can be constructed. Shortly, while discussing DWT type frequency decompositions we will identify a certain *filter bank* structure, that provides the key to the construction of a rich class of computationally efficient frequency decompositions with discrete representations.

Notice that in Eqns. 1-3, both time and frequency domains have been considered. This is to highlight that frequency decompositions could also be thought of as time decompositions (decomposition of a signal into "time" channels). For example, one could split the time-axis into non-overlapping bins, and one would then have an orthogonal overlapping frequency decomposition and a non-overlapping time decomposition. In both cases, $L^2(\mathbb{R})$, a separable Hilbert space is being decomposed into a possibly infinite sequence of subspaces. This viewpoint will be further explored in Section.8

2.1 The DSTFT decomposition

Short-Time Fourier analysis began essentially to study the non-stationary behavior of speech signals [40] and since then has been used successfully for the analysis of non-stationary signals [13, 37, 17].

1 Introduction

Wavelet theory deals with the study of the time-scale behavior of functions just as the Short-Time Fourier analysis deals with the time-frequency behavior of functions. Typically, in wavelet analysis a *generating* function, the *wavelet* is chosen, and an associated transform gives a time-scale representation of functions (typically finite-energy) called the wavelet representation. In a wavelet representation, the one or both of the time and scale parameters may be discrete or continuous and correspondingly we have the Discrete Wavelet transform (DWT- both parameters are discrete), the Continuous Wavelet Transform (CWT- both parameters are continuous) the Discrete-Scale Wavelet Transform (DSWT - only the time parameter is continuous) and the Discrete-Time Wavelet transform (DTWT - only the scale parameter is continuous) [19, 8, 22]. In this chapter we discuss DWTs only. For example, the orthonormal wavelet bases constructed by Daubechies [7], give rise to a discrete, time-scale representation of finite-energy signals. If the ON wavelet is compactly supported, then finite-power signals can also be localized in time and scale, thus facilitating the study of the time-scale behavior of periodic and non-stationary signals.

The enormous flexibility in the choice of the wavelet (reminiscent of the choice of the window function in Short-Time Fourier analysis), permits the use of optimal wavelets for specific applications. Since compactly supported wavelets are determined by a finite sequence of numbers, one can optimize over these to obtain best suited wavelets for specific applications. Even though, wavelet analysis can be introduced from a multiscale analysis framework viewpoint, we try to introduce it from a frequency decomposition framework which includes the familiar STFT and the wavelet analysis as limiting cases, and besides admits of generalizations that cannot be easily seen from a multiscale framework.

In general wavelet analysis is associated with fast computational techniques. This is because of certain efficient computational structures called *filter banks* associated with them. Filter bank theory not only gives computational algorithms for wavelet analysis, but moreover, acts as a starting point for some of the interesting aspects wavelet theory. Using the frequency decomposition framework we introduce filter banks and then construct multiplicity M wavelet frames and tight frames. We finally, describe how filter banks give a general decomposition of separable Hilbert spaces, and using these results to argue that, regularity of the wavelets may not be as critical as often presupposed (in practical applications).

2 Frequency Decompositions

The essential ideas in Wavelet and Short-Time-Fourier decompositions can be understood in the context of frequency decompositions. Frequency decompositions, overcome some of the shortcomings of Fourier analysis. The Fourier transform decomposes a signal into individual frequency components, but does not tell *when* the frequencies occurred. This is because the complex exponentials, the Fourier basis functions, are infinite in extent. Frequency decompositions generally decompose a signal into a countable number of frequency channels, so that temporal information can be obtained in each frequency channel.

For a function $f(t) \in L^2(\mathbb{R})$, its Fourier transform $\hat{f}(\omega)$ uniquely determines it. Since a

WAVELET TRANSFORMS AND FILTER BANKS

*R.A.Gopinath*and C.S.Burrus*
Department of Electrical and Computer Engineering,
Rice University, Houston, TX-77251

CML TR-91-20
30th September '91

Abstract

Wavelet and short-time Fourier analysis is introduced in the context of frequency decompositions. Wavelet type frequency decompositions are associated with filter banks, and using this fact, filter bank theory is used to construct multiplicity M wavelet frames and tight frames. The way in which filter banks lead to decomposition and recomposition of arbitrary separable Hilbert spaces is also described. Efficient computational structures for both filter banks and wavelets are also discussed.

Contact Address:
Ramesh A. Gopinath
Department of EE, A235
Rice University,
Houston, TX-77251
Phone (713) 527-8750 x3577
email: ramesh@rice.edu

*This work was supported by AFOSR under grant 90-0334 funded by DARPA

Appears in *Wavelets: A Tutorial in Theory and Applications*, ed. C.K.Chui, Academic Press

WAVELETS AND FILTER BANKS

R.A.Gopinath and C.S.Burrus
Department of Electrical and Computer Engineering,
Rice University, Houston, TX-77251

CML TR-91-20
30th September '91