Wavelet Packets and the Best Basis Algorithm.

A. The DWT as a subband coding scheme.

1. The fact that the auxiliary function $m_0(\gamma)$ satisfies $m_0(0) = 1$ and $m_0(1/2) = 0$ identifies m_0 roughly as a "low-pass filter," that is, convolution of a signal $c(n)$ with the scaling filter $h(n)$ (essentially the inverse Fourier transform of m_0) suppresses high frequencies in $c(n)$.

2. The operator H reduces essentially to two steps: (a) convolution of the signal c by h (really by the involution $\underline{h}$, although for example MATLAB does not do this), followed by (b) decimation of the resulting signal. This second step can be thought of as "coding" the convolved signal, the idea being that since high frequencies in the signal have been suppressed then it is sufficient to keep only half of the entries. This is not strictly speaking true however.

3. To see what parts of $c(n)$ have been recorded in the sequence $Hc(n)$, note that the sequence $(H^*Hc)(n)$ is the orthogonal projector onto the span of the collection $\{h(n-2k)\}_{k\in\mathbf{Z}}$, and that as we have seen before
$$(H^*Hc)^\wedge(\gamma) = \widehat{c}(\gamma)\,|m_0(\gamma)|^2 + \widehat{c}(\gamma+1/2)\,m_0(\gamma)\,\overline{m_0(\gamma+1/2)}.$$
The second term in the sum is referred to as "aliasing error".

4. We can make the corresponding observations about $m_1(\gamma)$ and the wavelet filter $g(n)$, namely (a) $m_1(\gamma)$ has the characteristics of a "high-pass" filter, (b) the sequence Gc can be thought of as a coded version of the high frequency parts of c, and (c) more specifically we know that
$$(G^*Gc)^\wedge(\gamma) = \widehat{c}(\gamma)\,|m_1(\gamma)|^2 + \widehat{c}(\gamma+1/2)\,m_1(\gamma)\,\overline{m_1(\gamma+1/2)}$$
where the second term in the sum is aliasing error. The fact that $(H^*H + G^*G)c = c$ is the statement that the aliasing error from the low and high pass filters cancel out.

5. As a useful paradigm, we can use the scaling and wavelet filters for the bandlimited wavelet as these are perfect low and high pass filters with no aliasing. We can use this paradigm to see how the DWT sequences $c_j(n)$ and $d_j(n)$ correspond to coded, band-pass-filtered versions of the signal $c_0(n)$.

B. Wavelet Packet Decompositions.

Definition 0.1 *Given a signal $c_0(n)$, the DWPT of $c_0(n)$ is the collection of sequences $d_j^n = \{d_j^n(k)\}_{k\in\mathbf{Z}}$ for $n \in \mathbf{Z}^+$, $j \in \mathbf{N}$ defined by*
$$d_j^{2n}(k) = H\,d_{j-1}^n(k), \qquad d_j^{2n+1}(k) = G\,d_{j-1}^n(k),$$
where $c_0(n) = d_0^0(n)$.

The DWPT is inverted by means of the formula
$$d_j^n(k) = H^* d_{j+1}^{2n}(k) + G^* d_{j+1}^{2n+1}(k).$$

Each set of coefficients d_j^n codes a particular subband of the original sequence $c_0(n) = d_0^0(n)$. In order to see which subband corresponds to which set of coefficients, let's assume that our scaling and wavelet filters are the perfect low-pass and high-pass filters of the bandlimited MRA. Then the following theorem holds.

Theorem 0.1 *Let $j \geq 0$ be fixed and let $0 \leq n < 2^j - 1$. Suppose that*

$$n = \epsilon_0 + 2\epsilon_1 + 4\epsilon_2 + \cdots + 2^{j-1}\epsilon_{j-1}$$

with $\epsilon_k = 0$ or 1. (In other words, $\epsilon_1\epsilon_2 \cdots \epsilon_{j-1}$ is the binary representation of the integer n.) Then under the assumption that the scaling and wavelet filters are the perfect low and high-pass filters found in the bandlimited MRA, the coefficients d_j^n codes the sequence whose Fourier transform is given by

$$\widehat{c_0}(\gamma) = |m_{\epsilon_0}(2^{j-1}\gamma)|^2\, |m_{\epsilon_1}(2^{j-2}\gamma)|^2 \cdots |m_{\epsilon_{j-1}}(\gamma)|^2.$$

In fact a "full traversal" of the wavelet packet tree corresponds to the expansion coefficients of the signal in an orthonormal basis. This basis corresponds to a particular dyadic partition of the frequency axis $[0, 1/2]$.

C. Wavelet packet bases for $L^2(\mathbf{R})$.

Definition 0.2 *Let $w^0(x)$ be an orthogonal scaling function with corresponding scaling filter $h(n)$. Define the wavelet filter $g(n)$ as usual, and define the sequence $\{w^m(x)\}_{m\in\mathbf{Z}^+}$ of* wavelet packet functions *by*

$$w^{2n}(x) = \sum_k h(k) w_{1,k}^n(x),$$
$$w^{2n+1}(x) = \sum_k g(k)\, w_{1,k}^n(x).$$

Theorem 0.2 *Suppose that for some $N \in \mathbf{N}$, the scaling function $w^0(x)$ is supported in $[0, 2N-1]$ and that the scaling filter $h(k)$ satisfies $h(k) = 0$ for $k < 0$ and $k \geq 2N$. Then for each $n \in \mathbf{N}$, $w^n(x)$ is supported in $[0, 2N-1]$.*

Theorem 0.3 *Let $n \in \mathbf{Z}^+$ be given, and let $\{\epsilon_0, \epsilon_1, \ldots, \epsilon_{k-1}\}$ be the unique sequence such that $\epsilon_i = 0$ or 1 and such that*

$$n = \epsilon_0 + 2\epsilon_1 + 4\epsilon_2 + \cdots + 2^{k-1}\epsilon_{k-1}.$$

(In other words, $\epsilon_1\epsilon_2 \cdots \epsilon_{k-1}$ is the binary representation of the integer n.) Then

$$\widehat{w^n}(\gamma) = m_{\epsilon_0}(\gamma/2)\, m_{\epsilon_1}(\gamma/4) \cdots m_{\epsilon_{k-1}}(\gamma/2^k)\, \widehat{\varphi}(\gamma/2^k).$$

Exercises.

1. Exercise 11.8, *An Introduction to Wavelet Analysis*
2. Exercise 11.9, *An Introduction to Wavelet Analysis*
3. Exercise 11.12, *An Introduction to Wavelet Analysis*
4. Exercise 11.41, *An Introduction to Wavelet Analysis*