CHAPTER

8 Wavelet Packet and Local Cosine Bases

Different types of time-frequency structures are encountered in complex signals such as speech recordings. This motivates the design of bases with time-frequency properties that may be adapted. Wavelet bases are one particular family of bases that represent piecewise smooth signals effectively. Other time-frequency bases are constructed to approximate different types of signals such as audio recordings.

Orthonormal wavelet packet bases are computed with conjugate mirror filters that divide the frequency axis in separate intervals of various sizes. Different conjugate mirror filter banks correspond to different wavelet packet bases. If the signal properties change over time, it is preferable to isolate different time intervals with translated windows. Local cosine bases are constructed by multiplying these windows with cosine functions. Wavelet packets segment the frequency axis and are uniformly translated in time, whereas local cosine bases divide the time axis and are uniformly translated in frequency. Both types of bases are extended in two dimensions for image processing.

8.1 WAVELET PACKETS

8.1.1 Wavelet Packet Tree

Wavelet packets were introduced by Coifman, Meyer, and Wickerhauser [182] by generalizing the link between multiresolution approximations and wavelets. A space $\mathbf{V}_j$ of a multiresolution approximation is decomposed in a lower-resolution space $\mathbf{V}_{j+1}$ plus a detail space $\mathbf{W}_{j+1}$. This is done by dividing the orthogonal basis $\{\phi_j(t - 2^j n)\}_{n\in\mathbb{Z}}$ of $\mathbf{V}_j$ into two new orthogonal bases

$$\{\phi_{j+1}(t - 2^{j+1} n)\}_{n\in\mathbb{Z}} \text{ of } \mathbf{V}_{j+1} \quad \text{and} \quad \{\psi_{j+1}(t - 2^{j+1} n)\}_{n\in\mathbb{Z}} \text{ of } \mathbf{W}_{j+1}.$$

The decompositions (7.107) and (7.109) of ϕ_{j+1} and ψ_{j+1} in the basis $\{\phi_j(t - 2^j n)\}_{n\in\mathbb{Z}}$ are specified by a pair of conjugate mirror filters $h[n]$ and

$$g[n] = (-1)^{1-n} h[1 - n].$$

Theorem 8.1 generalizes this result to any space $\mathbf{U}_j$ that admits an orthogonal basis of functions translated by $n2^j$ for $n \in \mathbb{Z}$.

Theorem 8.1: *Coifman, Meyer, Wickerhauser.* Let $\{\theta_j(t-2^j n)\}_{n\in\mathbb{Z}}$ be an orthonormal basis of a space $\mathbf{U}_j$. Let h and g be a pair of conjugate mirror filters. Define

$$\theta^0_{j+1}(t)=\sum_{n=-\infty}^{+\infty} h[n]\,\theta_j(t-2^j n) \quad \text{and} \quad \theta^1_{j+1}(t)=\sum_{n=-\infty}^{+\infty} g[n]\,\theta_j(t-2^j n). \tag{8.1}$$

The family

$$\left\{\theta^0_{j+1}(t-2^{j+1}n),\ \theta^1_{j+1}(t-2^{j+1}n)\right\}_{n\in\mathbb{Z}}$$

is an orthonormal basis of $\mathbf{U}_j$.

Proof. This proof is very similar to the proof of Theorem 7.3. The main steps are outlined. Theorem 3.4 shows that $\{\theta_j(t-2^j n)\}_{n\in\mathbb{Z}}$ is orthogonal if and only if

$$\frac{1}{2^j}\sum_{k=-\infty}^{+\infty}\left|\hat\theta_j\Big(\omega+\frac{2k\pi}{2^j}\Big)\right|^2=1. \tag{8.2}$$

We derive from (8.1) that the Fourier transform of θ^0_{j+1} is

$$\hat\theta^0_{j+1}(\omega)=\hat\theta_j(\omega)\sum_{n=-\infty}^{+\infty} h[n]\,\exp(-i2^j n\omega)=\hat h(2^j\omega)\,\hat\theta_j(\omega). \tag{8.3}$$

Similarly, the Fourier transform of θ^1_{j+1} is

$$\hat\theta^1_{j+1}(\omega)=\hat g(2^j\omega)\,\hat\theta_j(\omega). \tag{8.4}$$

Proving that $\{\theta^0_{j+1}(t-2^{j+1}n)\}$ and $\{\theta^1_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ are two families of orthogonal vectors is equivalent to showing that for $l=0$ or $l=1$

$$\frac{1}{2^{j+1}}\sum_{k=-\infty}^{+\infty}\left|\hat\theta^l_{j+1}\Big(\omega+\frac{2k\pi}{2^{j+1}}\Big)\right|^2=1. \tag{8.5}$$

These two families of vectors yield orthogonal spaces if and only if

$$\frac{1}{2^{j+1}}\sum_{k=-\infty}^{+\infty}\hat\theta^0_{j+1}\left(\omega+\frac{2k\pi}{2^{j+1}}\right)\hat\theta^{1*}_{j+1}\left(\omega+\frac{2k\pi}{2^{j+1}}\right)=0. \tag{8.6}$$

The relations (8.5) and (8.6) are verified by replacing $\hat\theta^0_{j+1}$ and $\hat\theta^1_{j+1}$ by (8.3) and (8.4), respectively, and by using the orthogonality of the basis (8.2) and the conjugate mirror filter properties

$$\begin{aligned}|\,\hat h(\omega)|^2+|\,\hat h(\omega+\pi)|^2&=2,\\ |\hat g(\omega)|^2+|\hat g(\omega+\pi)|^2&=2,\\ \hat g(\omega)\,\hat h^*(\omega)+\hat g(\omega+\pi)\,\hat h^*(\omega+\pi)&=0.\end{aligned}$$

To prove that the family $\{\theta^0_{j+1}(t-2^{j+1}n), \theta^1_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ generates the same space as $\{\theta_j(t-2^j n)\}_{n\in\mathbb{Z}}$, we must prove that for any $a[n]\in\ell^2(\mathbb{Z})$ there exist $b[n]\in\ell^2(\mathbb{Z})$ and $c[n]\in\ell^2(\mathbb{Z})$ such that

$$\sum_{n=-\infty}^{+\infty} a[n]\,\theta_j(t-2^j n) = \sum_{n=-\infty}^{+\infty} b[n]\,\theta^0_{j+1}(t-2^{j+1}n) + \sum_{n=-\infty}^{+\infty} c[n]\,\theta^1_{j+1}(t-2^{j+1}n). \quad \textbf{(8.7)}$$

To do this, we relate $\hat{b}(\omega)$ and $\hat{c}(\omega)$ to $\hat{a}(\omega)$. The Fourier transform of (8.7) yields

$$\hat{a}(2^j\omega)\,\hat{\theta}_j(\omega) = \hat{b}(2^{j+1}\omega)\,\hat{\theta}^0_{j+1}(\omega) + \hat{c}(2^{j+1}\omega)\,\hat{\theta}^1_{j+1}(\omega). \quad \textbf{(8.8)}$$

One can verify that

$$\hat{b}(2^{j+1}\omega) = \frac{1}{2}\Big(\hat{a}(2^j\omega)\,\hat{h}^*(2^j\omega) + \hat{a}(2^j\omega+\pi)\,\hat{h}^*(2^j\omega+\pi)\Big)$$

and

$$\hat{c}(2^{j+1}\omega) = \frac{1}{2}\Big(\hat{a}(2^j\omega)\,\hat{g}^*(2^j\omega) + \hat{a}(2^j\omega+\pi)\,\hat{g}^*(2^j\omega+\pi)\Big)$$

satisfy (8.8). ■

Theorem 8.1 proves that conjugate mirror filters transform an orthogonal basis $\{\theta_j(t-2^j n)\}_{n\in\mathbb{Z}}$ in two orthogonal families $\{\theta^0_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ and $\{\theta^1_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$. Let $\mathbf{U}^0_{j+1}$ and $\mathbf{U}^1_{j+1}$ be the spaces generated by each of these families. Clearly $\mathbf{U}^0_{j+1}$ and $\mathbf{U}^1_{j+1}$ are orthogonal and

$$\mathbf{U}^0_{j+1}\oplus\mathbf{U}^1_{j+1} = \mathbf{U}_j.$$

Computing the Fourier transform of (8.1) relates the Fourier transforms of θ^0_{j+1} and θ^1_{j+1} to the Fourier transform of θ_j:

$$\hat{\theta}^0_{j+1}(\omega) = \hat{h}(2^j\omega)\,\hat{\theta}_j(\omega), \quad \hat{\theta}^1_{j+1}(\omega) = \hat{g}(2^j\omega)\,\hat{\theta}_j(\omega). \quad \textbf{(8.9)}$$

Since the transfer functions $\hat{h}(2^j\omega)$ and $\hat{g}(2^j\omega)$ have their energy concentrated in different frequency intervals, this transformation can be interpreted as a division of the frequency support of $\hat{\theta}_j$.

Binary Wavelet Packet Tree

Instead of dividing only the approximation spaces $\mathbf{V}_j$ to construct detail spaces $\mathbf{W}_j$ and wavelet bases, Theorem 8.1 proves that we can set $\mathbf{U}_j = \mathbf{W}_j$ and divide these detail spaces to derive new bases. The recursive splitting of vector spaces is represented in a binary tree. If the signals are approximated at the scale 2^L, to the root of the tree we associate the approximation space $\mathbf{V}_L$. This space admits an orthogonal basis of scaling functions $\{\phi_L(t-2^L n)\}_{n\in\mathbb{Z}}$ with $\phi_L(t) = 2^{-L/2}\,\phi(2^{-L}t)$.

Any node of the binary tree is labeled by (j,p), where $j-L\geqslant 0$ is the depth of the node in the tree, and p is the number of nodes that are on its left at the same

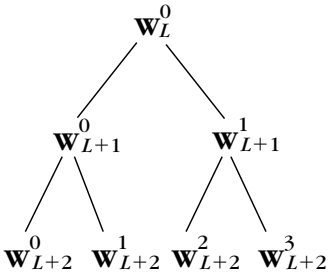

FIGURE 8.1

Binary tree of wavelet packet spaces.

depth $j-L$. Such a tree is illustrated in Figure 8.1. To each node (j,p) we associate a space $\mathbf{W}_j^p$, which admits an orthonormal basis $\{\psi_j^p(t-2^j n)\}_{n\in\mathbb{Z}}$ by going down the tree. At the root we have $\mathbf{W}_L^0=\mathbf{V}_L$ and $\psi_L^0=\phi_L$. Suppose now that we have already constructed $\mathbf{W}_j^p$ and its orthonormal basis $\mathcal{B}_j^p=\{\psi_j^p(t-2^j n)\}_{n\in\mathbb{Z}}$ at the node (j,p). The two wavelet packet orthogonal bases at the children nodes are defined by the splitting relations (8.1):

$$\psi_{j+1}^{2p}(t)=\sum_{n=-\infty}^{+\infty} h[n]\,\psi_j^p(t-2^j n) \tag{8.10}$$

and

$$\psi_{j+1}^{2p+1}(t)=\sum_{n=-\infty}^{+\infty} g[n]\,\psi_j^p(t-2^j n). \tag{8.11}$$

Since $\{\psi_j^p(t-2^j n)\}_{n\in\mathbb{Z}}$ is orthonormal,

$$h[n]=\langle\psi_{j+1}^{2p}(u),\psi_j^p(u-2^j n)\rangle,\quad g[n]=\langle\psi_{j+1}^{2p+1}(u),\psi_j^p(u-2^j n)\rangle. \tag{8.12}$$

Theorem 8.1 proves that $\mathcal{B}_{j+1}^{2p}=\{\psi_{j+1}^{2p}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ and $\mathcal{B}_{j+1}^{2p+1}=\{\psi_{j+1}^{2p+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ are orthonormal bases of two orthogonal spaces $\mathbf{W}_{j+1}^{2p}$ and $\mathbf{W}_{j+1}^{2p+1}$ such that

$$\mathbf{W}_{j+1}^{2p}\oplus\mathbf{W}_{j+1}^{2p+1}=\mathbf{W}_j^p. \tag{8.13}$$

This recursive splitting defines a binary tree of wavelet packet spaces where each parent node is divided in two orthogonal subspaces. Figure 8.2 displays the eight wavelet packets ψ_j^p at the depth $j-L=3$, calculated with a Daubechies 5 filter. These wavelet packets are frequency ordered from left to right, as explained in Section 8.1.2.

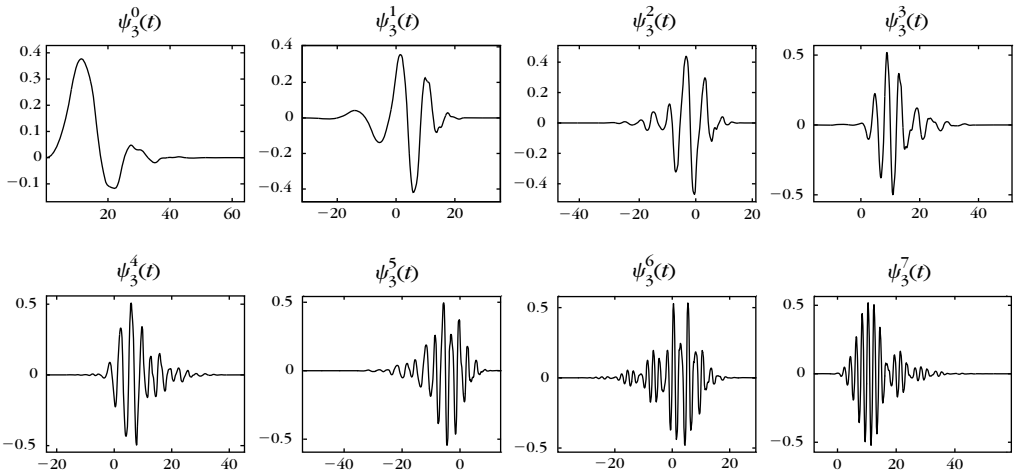

FIGURE 8.2

Wavelet packets computed with a Daubechies 5 filter at the depth $j - L = 3$ of the wavelet packet tree, with $L = 0$. They are ordered from low to high frequencies.

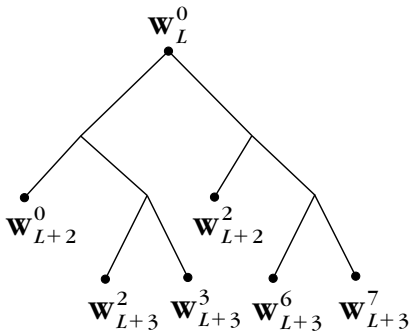

FIGURE 8.3

Example of an admissible wavelet packet binary tree.

Admissible Tree

We call any binary tree where each node has either zero or two children an *admissible tree*, as shown in Figure 8.3. Let $\{j_i, p_i\}_{1 \leqslant i \leqslant I}$ be the leaves of an admissible binary tree. By applying the recursive splitting (8.13) along the branches of an admissible tree, we verify that the spaces $\{\mathbf{W}_{j_i}^{p_i}\}_{1 \leqslant i \leqslant I}$ are mutually orthogonal and add up to $\mathbf{W}_L^0$:

$$\mathbf{W}_L^0 = \oplus_{i=1}^{I} \mathbf{W}_{j_i}^{p_i}. \tag{8.14}$$

The union of the corresponding wavelet packet bases

$$\{\psi_{j_i}^{p_i}(t - 2^{j_i}n)\}_{n\in\mathbb{Z}, 1\leq i\leq I}$$

thus defines an orthogonal basis of $\mathbf{W}_L^0 = \mathbf{V}_L$.

Number of Wavelet Packet Bases

The number of different wavelet packet orthogonal bases of $\mathbf{V}_L$ is equal to the number of different admissible binary trees. Theorem 8.2 proves that there are more than $2^{2^{J-1}}$ different wavelet packet orthonormal bases included in a full wavelet packet binary tree of depth J.

Theorem 8.2. The number B_J of wavelet packet bases in a full wavelet packet binary tree of depth J satisfies

$$2^{2^{J-1}} \leq B_J \leq 2^{\frac{5}{4}2^{J-1}}. \tag{8.15}$$

Proof. This result is proved by induction on the depth J of the wavelet packet tree. The number B_J of different orthonormal bases is equal to the number of different admissible binary trees of depth of at most J, which have nodes with either zero or two children. For $J = 0$, the tree is reduced to its root, so $B_0 = 1$.

Observe that the set of trees of depth of at most $J + 1$ is composed of trees of depth of at least 1 and at most $J + 1$ plus one tree of depth 0 that is reduced to the root. A tree of depth of at least 1 has a left and a right subtree that are admissible trees of depth of at most J. The configuration of these trees is a priori independent and there are B_J admissible trees of depth J, so

$$B_{J+1} = B_J^2 + 1. \tag{8.16}$$

Since $B_1 = 2$ and $B_{J+1} \geq B_J^2$, we prove by induction that $B_J \geq 2^{2^{J-1}}$. Moreover,

$$\log_2 B_{J+1} = 2\log_2 B_J + \log_2(1 + B_J^{-2}).$$

If $J \geq 1$, then $B_J \geq 2$, so

$$\log_2 B_{J+1} \leq 2\log_2 B_J + \frac{1}{4}. \tag{8.17}$$

Since $B_1 = 2$,

$$\log_2 B_{J+1} \leq 2^J + \frac{1}{4}\sum_{j=0}^{J-1} 2^j \leq 2^J + \frac{2^J}{4},$$

so $B_J \leq 2^{\frac{5}{4}2^{J-1}}$. ■

For discrete signals of size N, we shall see that the wavelet packet tree is at most of depth $J = \log_2 N$. This theorem proves that the number of wavelet packet bases satisfies $2^{N/2} \leq B_{\log_2 N} \leq 2^{5N/8}$.

Wavelet Packets on Intervals

To construct wavelet packet bases of $\mathbf{L^2}[0,1]$, we use the border techniques developed in Section 7.5 to design wavelet bases of $\mathbf{L^2}[0,1]$. The simplest approach constructs periodic bases. As in the wavelet case, the coefficients of $f \in \mathbf{L^2}[0,1]$ in a periodic wavelet packet basis are the same as the decomposition coefficients of $f^{\mathrm{per}}(t) = \sum_{k=-\infty}^{+\infty} f(t+k)$ in the original wavelet packet basis of $\mathbf{L^2}(\mathbb{R})$. The periodization of f often creates discontinuities at the borders $t=0$ and $t=1$, which generate large-amplitude wavelet packet coefficients.

Section 7.5.3 describes a more sophisticated technique that modifies the filters h and g in order to construct boundary wavelets that keep their vanishing moments. A generalization to wavelet packets is obtained by using these modified filters in Theorem 8.1. This avoids creating the large-amplitude coefficients at the boundary, which is typical of the periodic case.

Biorthogonal Wavelet Packets

Nonorthogonal wavelet bases are constructed in Section 7.4 with two pairs of perfect reconstruction filters (h, g) and $(\tilde{h}, \tilde{g})$ instead of a single pair of conjugate mirror filters. The orthogonal splitting Theorem 8.1 is extended into a biorthogonal splitting by replacing the conjugate mirror filters with these perfect reconstruction filters. A Riesz basis $\{\theta_j(t-2^j n)\}_{n\in\mathbb{Z}}$ of $\mathbf{U}_j$ is transformed into two Riesz bases $\{\theta^0_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ and $\{\theta^1_{j+1}(t-2^{j+1}n)\}_{n\in\mathbb{Z}}$ of two nonorthogonal spaces $\mathbf{U}^0_{j+1}$ and $\mathbf{U}^1_{j+1}$ such that

$$\mathbf{U}^0_{j+1} \oplus \mathbf{U}^1_{j+1} = \mathbf{U}_j.$$

A binary tree of nonorthogonal wavelet packet Riesz bases can be derived by induction using this vector space division. As in the orthogonal case, the wavelet packets at the leaves of an admissible binary tree define a basis of $\mathbf{W}^0_L$, but this basis is not orthogonal.

The lack of orthogonality is not a problem by itself as long as the basis remains stable. Cohen and Daubechies proved [171] that when the depth $j-L$ increases, the angle between the spaces $\mathbf{W}^p_j$ located at the same depth can become progressively smaller. This indicates that some of the wavelet packet bases constructed from an admissible binary tree become unstable. Thus, we concentrate on orthogonal wavelet packets constructed with conjugate mirror filters.

8.1.2 Time-Frequency Localization

Time Support

If the conjugate mirror filters h and g have a finite impulse response of size K, Theorem 7.5 proves that ϕ has a support size of $K-1$, so $\psi^0_L = \phi_L$ has a support size of $(K-1)2^L$. Since

$$\psi^{2p}_{j+1}(t) = \sum_{n=-\infty}^{+\infty} h[n]\,\psi^p_j(t-2^j n), \quad \psi^{2p+1}_{j+1}(t) = \sum_{n=-\infty}^{+\infty} g[n]\,\psi^p_j(t-2^j n), \qquad \textbf{(8.18)}$$

an induction on j shows that the support size of ψ_j^p is $(K-1)2^j$. Thus, the parameter j specifies the scale 2^j of the support. The wavelet packets in Figure 8.2 are constructed with a Daubechies filter of $K=10$ coefficients with $j=3$ and thus have a support size of $2^3(10-1)=72$.

Frequency Localization

The frequency localization of wavelet packets is more complicated to analyze. The Fourier transform of (8.18) proves that the Fourier transforms of wavelet packet children are related to their parent by

$$\hat{\psi}_{j+1}^{2p}(\omega)=\hat{h}(2^j\omega)\,\hat{\psi}_j^p(\omega), \quad \hat{\psi}_{j+1}^{2p+1}(\omega)=\hat{g}(2^j\omega)\,\hat{\psi}_j^p(\omega). \tag{8.19}$$

The energy of $\hat{\psi}_j^p$ is mostly concentrated over a frequency band and the two filters $\hat{h}(2^j\omega)$ and $\hat{g}(2^j\omega)$ select the lower- or higher-frequency components within this band. To relate the size and position of this frequency band to the indexes (p, j), we consider a simple example.

Shannon Wavelet Packets

Shannon wavelet packets are computed with perfect discrete low-pass and high-pass filters

$$|\hat{h}(\omega)|=\begin{cases}\sqrt{2} & \text{if } \omega\in[-\pi/2+2k\pi,\ \pi/2+2k\pi] \text{ with } k\in\mathbb{Z}\\ 0 & \text{otherwise}\end{cases} \tag{8.20}$$

and

$$|\hat{g}(\omega)|=\begin{cases}\sqrt{2} & \text{if } \omega\in[\pi/2+2k\pi,\ 3\pi/2+2k\pi] \text{ with } k\in\mathbb{Z}\\ 0 & \text{otherwise.}\end{cases} \tag{8.21}$$

In this case, it is relatively simple to calculate the frequency support of the wavelet packets. The Fourier transform of the scaling function is

$$\hat{\psi}_L^0=\hat{\phi}_L=\mathbf{1}_{[-2^{-L}\pi,\,2^{-L}\pi]}. \tag{8.22}$$

Each multiplication with $\hat{h}(2^j\omega)$ or $\hat{g}(2^j\omega)$ divides the frequency support of the wavelet packets in two. The delicate point is to realize that $\hat{h}(2^j\omega)$ does not always play the role of a low-pass filter because of the side lobes that are brought into the interval $[-2^{-L}\pi,\ 2^{-L}\pi]$ by the dilation. At depth $j-L$, Theorem 8.3 proves that $\hat{\psi}_j^p$ is proportional to the indicator function of a pair of frequency intervals, which are labeled I_j^k. The permutation that relates p and k is characterized recursively [71].

Theorem 8.3: *Coifman, Wickerhauser.* For any $j-L>0$ and $0\leq p<2^{j-L}$, there exists $0\leq k<2^{j-L}$ such that

$$|\hat{\psi}_j^p(\omega)|=2^{j/2}\,\mathbf{1}_{I_j^k}(\omega), \tag{8.23}$$

where I_j^k is a symmetric pair of intervals

$$I_j^k = [-(k+1)\pi 2^{-j}, -k\pi 2^{-j}] \cup [k\pi 2^{-j}, (k+1)\pi 2^{-j}]. \tag{8.24}$$

The permutation $k = G[p]$ satisfies for any $0 \leq p < 2^{j-L}$

$$G[2p] = \begin{cases} 2G[p] & \text{if } G[p] \text{ is even} \\ 2G[p]+1 & \text{if } G[p] \text{ is odd} \end{cases} \tag{8.25}$$

$$G[2p+1] = \begin{cases} 2G[p]+1 & \text{if } G[p] \text{ is even} \\ 2G[p] & \text{if } G[p] \text{ is odd.} \end{cases} \tag{8.26}$$

Proof. Equations (8.23), (8.25), and (8.26) are proved by induction on depth $j-L$. For $j-L=0$, (8.22) shows that (8.23) is valid. Suppose that (8.23) is valid for $j=l \geq L$ and any $0 \leq p < 2^{l-L}$. We first prove that (8.25) and (8.26) are verified for $j=l$. From these two equations we then easily carry the induction hypothesis to prove that (8.23) is true for $j=l+1$ and for any $0 \leq p < 2^{l+1-L}$.

Equations (8.20) and (8.21) imply that

$$|\hat{h}(2^l\omega)| = \begin{cases} \sqrt{2} & \text{if } \omega \in [-2^{-l-1}(4m-1)\pi, 2^{-l-1}(4m+1)\pi] \text{ with } m \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \tag{8.27}$$

$$|\hat{g}(2^l\omega)| = \begin{cases} \sqrt{2} & \text{if } \omega \in [-2^{-l-1}(4m+1)\pi, 2^{-l-1}(4m+3)\pi] \text{ with } m \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases} \tag{8.28}$$

Since (8.23) is valid for l, the support of $\hat{\psi}_l^p$ is

$$I_l^k = [-(2k+2)\pi 2^{-l-1}, -2k\pi 2^{-l-1}] \cup [2k\pi 2^{-l-1}, (2k+2)\pi 2^{-l-1}].$$

The two children are defined by

$$\hat{\psi}_{l+1}^{2p}(\omega) = \hat{h}(2^l\omega)\,\hat{\psi}_l^p(\omega), \quad \hat{\psi}_{l+1}^{2p+1}(\omega) = \hat{g}(2^l\omega)\,\hat{\psi}_l^p(\omega).$$

Thus, we derive (8.25) and (8.26) by checking the intersection of I_l^k with the supports of $\hat{h}(2^j\omega)$ and $\hat{g}(2^j\omega)$ specified by (8.27) and (8.28). ■

For Shannon wavelet packets, Theorem 8.3 proves that $\hat{\psi}_j^p$ has a frequency support located over two intervals of size $2^{-j}\pi$, centered at $\pm(k+1/2)\pi 2^{-j}$. The Fourier transform expression (8.23) implies that these Shannon wavelet packets can be written as cosine-modulated windows

$$\psi_j^p(t) = 2^{-j/2+1}\,\theta(2^{-j}t)\,\cos\left[2^{-j}\pi(k+1/2)(t-\tau_{j,p})\right], \tag{8.29}$$

with

$$\theta(t) = \frac{\sin(\pi t/2)}{\pi t},$$

and thus,

$$\hat{\theta}(\omega) = \mathbf{1}_{[-\pi/2,\pi/2]}(\omega).$$

The translation parameter $\tau_{j,p}$ can be calculated from the complex phase of $\hat{\psi}_j^p$.

Frequency Ordering

It is often easier to label ψ_j^k a wavelet packet ψ_j^p that has a Fourier transform centered at $\pm(k+1/2)\pi 2^{-j}$ with $k = G[p]$. This means changing its position in the wavelet packet tree from node p to node k. The resulting wavelet packet tree is frequency ordered. The left child always corresponds to a lower-frequency wavelet packet and the right child to a higher-frequency one.

The permutation $k = G[p]$ is characterized by the recursive equations (8.25) and (8.26). The inverse permutation $p = G^{-1}[k]$ is called a *Gray code* in coding theory. This permutation is implemented on binary strings by deriving the following relations from (8.25) and (8.26). If p_i is the i^{th} binary digit of the integer p and k_i the i^{th} digit of $k = G[p]$, then

$$k_i = \left(\sum_{l=i}^{+\infty} p_l \right) \bmod 2, \tag{8.30}$$

$$p_i = (k_i + k_{i+1}) \bmod 2. \tag{8.31}$$

Compactly Supported Wavelet Packets

Wavelet packets of compact support have a more complicated frequency behavior than Shannon wavelet packets, but the previous analysis provides important insights. If h is a finite impulse response filter, $\hat{h}$ does not have a support restricted to $[-\pi/2, \pi/2]$ over the interval $[-\pi, \pi]$. It is true, however, that the energy of $\hat{h}$ is mostly concentrated in $[-\pi/2, \pi/2]$. Similarly, the energy of $\hat{g}$ is mostly concentrated in $[-\pi, -\pi/2] \cup [\pi/2, \pi]$, for $\omega \in [-\pi, \pi]$. As a consequence, the localization properties of Shannon wavelet packets remain qualitatively valid. The energy of $\hat{\psi}_j^p$ is mostly concentrated over

$$I_j^k = [-(k+1)\pi 2^{-j}, \quad -k\pi 2^{-j}] \cup [k\pi 2^{-j}, \quad (k+1)\pi 2^{-j}],$$

with $k = G[p]$. The larger the proportion of energy of $\hat{h}$ in $[-\pi/2, \pi/2]$, the more concentrated the energy of $\hat{\psi}_j^p$ in I_j^k. The energy concentration of $\hat{h}$ in $[-\pi/2, \pi/2]$ is increased by having more zeroes at π, so that $\hat{h}(\omega)$ remains close to zero in $[-\pi, -\pi/2] \cup [\pi/2, \pi]$. Theorem 7.4 proves that this is equivalent to imposing that the wavelets constructed in the wavelet packet tree have many vanishing moments.

These qualitative statements must be interpreted carefully. The side lobes of $\hat{\psi}_j^p$ beyond the intervals I_j^k are not completely negligible. For example, wavelet packets created with a Haar filter are discontinuous functions. Thus, $|\hat{\psi}_j^p(\omega)|$ decays like $|\omega|^{-1}$ at high frequencies, which indicates the existence of large side lobes outside I_k^p. It is also important to note that contrary to Shannon wavelet packets, compactly supported wavelet packets cannot be written as dilated windows modulated by cosine functions of varying frequency. When the scale increases, wavelet packets generally do not converge to cosine functions. They may have a wild behavior with localized oscillations of considerable amplitude.

Walsh Wavelet Packets

Walsh wavelet packets are generated by the Haar conjugate mirror filter

$$h[n] = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } n = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

They have very different properties from Shannon wavelet packets since the filter h is well localized in time but not in frequency. The corresponding scaling function is $\phi = \mathbf{1}_{[0,1]}$ and the approximation space $\mathbf{V}_L = \mathbf{W}_L^0$ is composed of functions that are constant over the intervals $[2^L n, 2^L(n+1))$ for $n \in \mathbb{Z}$. Since all wavelet packets created with this filter belong to $\mathbf{V}_L$, they are piecewise constant functions. The support size of h is $K = 2$, so Walsh functions ψ_j^p have a support size of 2^j. The wavelet packet recursive relations (8.18) become

$$\psi_{j+1}^{2p}(t) = \frac{1}{\sqrt{2}}\psi_j^p(t) + \frac{1}{\sqrt{2}}\psi_j^p(t - 2^j), \tag{8.32}$$

and

$$\psi_{j+1}^{2p+1}(t) = \frac{1}{\sqrt{2}}\psi_j^p(t) - \frac{1}{\sqrt{2}}\psi_j^p(t - 2^j). \tag{8.33}$$

Since ψ_j^p has a support size of 2^j, it does not intersect the support of $\psi_j^p(t - 2^j)$. Thus, these wavelet packets are constructed by juxtaposing ψ_j^p with its translated version that has a sign that might be changed. Figure 8.4 shows the Walsh functions at depth $j - L = 3$ of the wavelet packet tree. Theorem 8.4 computes the number of oscillations of ψ_j^p.

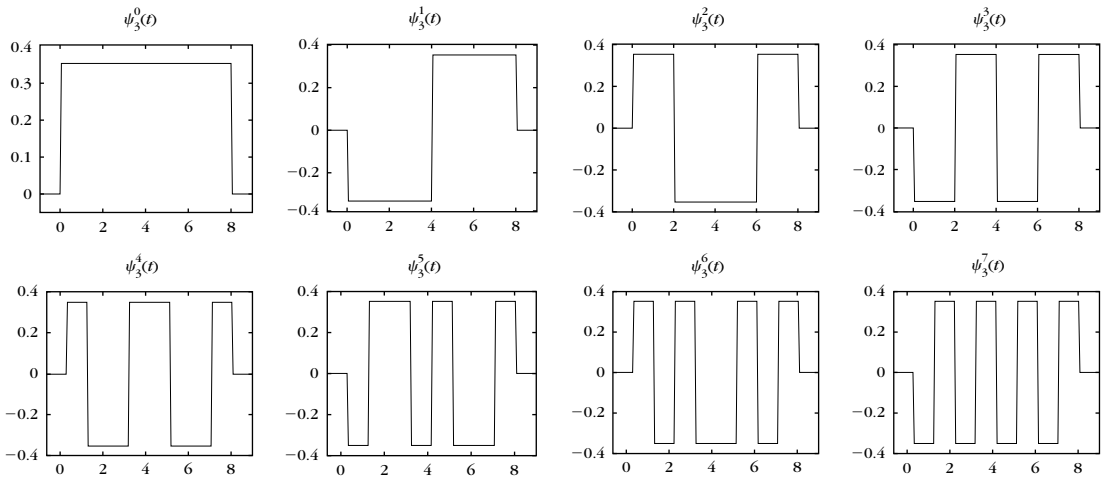

FIGURE 8.4

Frequency-ordered Walsh wavelet packets computed with a Haar filter at depth $j - L = 3$ of the wavelet packet tree, with $L = 0$.

Theorem 8.4. The support of a Walsh wavelet packet ψ_j^p is $[0, 2^j]$. Over its support, $\psi_j^p(t) = \pm 2^{-j/2}$. It changes sign $k = G[p]$ times, where $G[p]$ is the permutation defined by (8.25) and (8.26).

Proof. By induction on j, we derive from (8.32) and (8.33) that the support is $[0, 2^j]$ and that $\psi_j^p(t) = \pm 2^{-j/2}$ over its support. Let k be the number of times that ψ_j^p changes sign. The number of times that ψ_{j+1}^{2p} and ψ_{j+1}^{2p+1} change sign is either $2k$ or $2k+1$ depending on the sign of the first and last nonzero values of ψ_j^p. If k is even, then the sign of the first and last nonzero values of ψ_j^p are the same. Thus, the number of times ψ_{j+1}^{2p} and ψ_{j+1}^{2p+1} change sign is, respectively, $2k$ and $2k+1$. If k is odd, then the sign of the first and last nonzero values of ψ_j^p are different. The number of times ψ_{j+1}^{2p} and ψ_{j+1}^{2p+1} change sign is then $2k+1$ and $2k$. These recursive properties are identical to (8.25) and (8.26). ■

Therefore, a Walsh wavelet packet ψ_j^p is a square wave with $k = G[p]$ oscillations over a support size of 2^j. This result is similar to (8.29), which proves that a Shannon wavelet packet ψ_j^p is a window modulated by a cosine of frequency $2^{-j}k\pi$. In both cases, the oscillation frequency of wavelet packets is proportional to $2^{-j}k$.

Heisenberg Boxes

For display purposes, we associate to any wavelet packet $\psi_j^p(t - 2^j n)$ a Heisenberg rectangle, which indicates the time and frequency domains where the energy of this wavelet packet is mostly concentrated. The time support of the rectangle is set to be the same as the time support of a Walsh wavelet packet $\psi_j^p(t - 2^j n)$, which is equal to $[2^j n, 2^j(n+1)]$. The frequency support of the rectangle is defined as the positive-frequency support $[k\pi 2^{-j}, (k+1)\pi 2^{-j}]$ of Shannon wavelet packets with $k = G[p]$. The scale 2^j modifies the time and frequency elongation of this time-frequency rectangle, but its surface remains constant. The indices n and k give its localization in time and frequency. General wavelet packets—for example, computed with Daubechies filters—have a time and frequency spread that is much wider than this Heisenberg rectangle. However, this convention has the advantage of associating a wavelet packet basis to an exact paving of the time-frequency plane. Figure 8.5 shows an example of such a paving and the corresponding wavelet packet tree.

Figure 8.6 displays the decomposition of a multichirp signal having a spectrogram shown in Figure 4.3. The wavelet packet basis is computed with a Daubechies 10 filter. As expected, the large-amplitude coefficients are along the trajectory of linear and quadratic chirps that appear in Figure 4.3. We also see the trace of the two modulated Gaussian functions located at $t = 0.5$ and $t = 0.87$.

8.1.3 Particular Wavelet Packet Bases

Among the many wavelet packet bases, we describe here the properties of M-band wavelet bases, "local cosine type" bases, and "best" bases. The wavelet packet

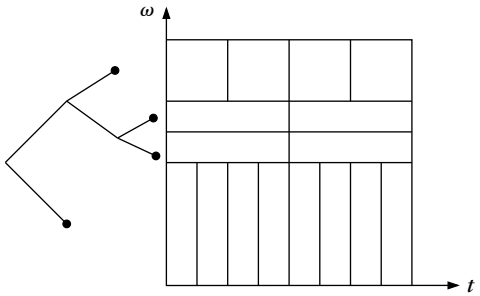

FIGURE 8.5

The wavelet packet tree (*left*) divides the frequency axis in several intervals. The Heisenberg boxes (*right*) of the corresponding wavelet packet basis.

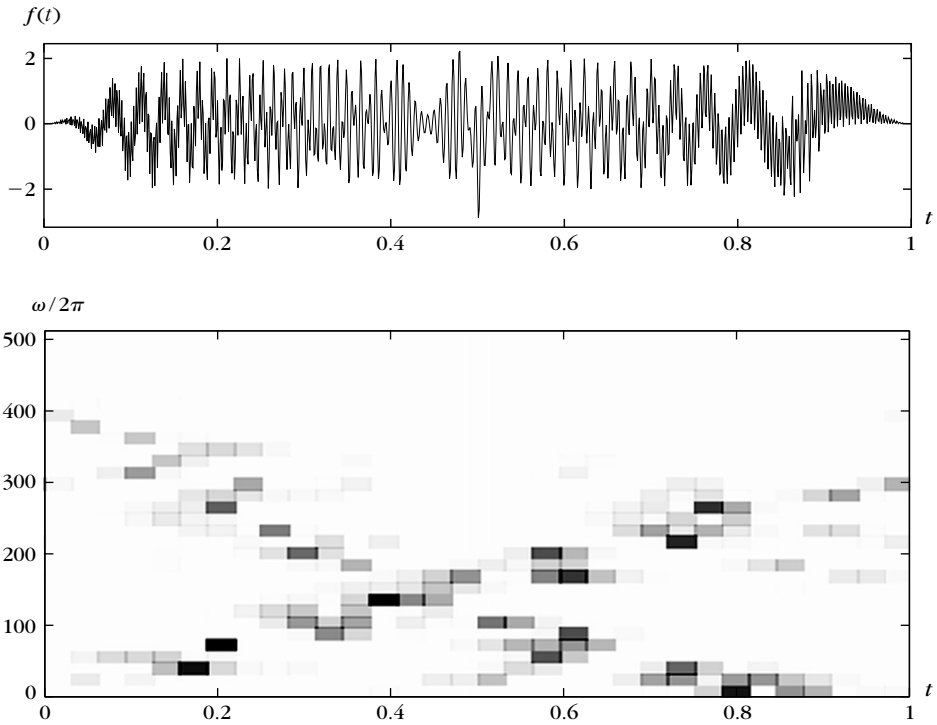

FIGURE 8.6

Wavelet packet decomposition of the multichirp signal the spectrogram of which is shown in Figure 4.3. The darker the gray level of each Heisenberg box, the larger the amplitude $|\langle f, \psi_j^p \rangle|$ of the corresponding wavelet packet coefficient.

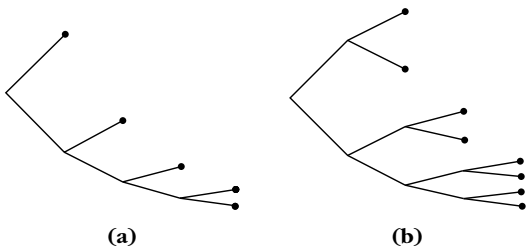

FIGURE 8.7

(a) Wavelet packet tree of a dyadic wavelet basis. **(b)** Wavelet packet tree of an M-band wavelet basis with $M = 2$.

tree is frequency ordered, which means that ψ_j^k has a Fourier transform with an energy essentially concentrated in the interval $[k\pi 2^{-j}, (k+1)\pi 2^{-j}]$ for positive frequencies.

M-Band Wavelets

The standard dyadic wavelet basis is an example of a wavelet packet basis of $\mathbf{V}_L$, obtained by choosing the admissible binary tree shown in Figure 8.7(a). Its leaves are the nodes $k = 1$ at all depth $j - L$ and thus correspond to the wavelet packet basis

$$\{\psi_j^1(t - 2^j n)\}_{n\in\mathbb{Z}, j>L}$$

constructed by dilating a single wavelet ψ^1:

$$\psi_j^1(t) = \frac{1}{\sqrt{2^j}}\,\psi^1\left(\frac{t}{2^j}\right).$$

The energy of $\hat{\psi}^1$ is mostly concentrated in the interval $[-2\pi, -\pi]\cup[\pi, 2\pi]$. The octave bandwidth for positive frequencies is the ratio between the bandwidth of the pass band and its distance to the zero frequency. It is equal to 1 octave. This quantity remains constant by dilation and specifies the frequency resolution of the wavelet transform.

Wavelet packets include other wavelet bases constructed with several wavelets having a better frequency resolution. Let us consider the admissible binary trees of Figure 8.7(b), which have leaves that are indexed by $k = 2$ and $k = 3$ at all depth $j - L$. The resulting wavelet packet basis of $\mathbf{V}_L$ is

$$\{\psi_j^2(t - 2^j n),\ \psi_j^3(t - 2^j n)\}_{n\in\mathbb{Z}, j>L+1}.$$

These wavelet packets can be rewritten as dilations of two elementary wavelets ψ^2 and ψ^3:

$$\psi_j^2(t) = \frac{1}{\sqrt{2^{j-1}}}\,\psi^2\Big(\frac{t}{2^{j-1}}\Big), \quad \psi_j^3(t) = \frac{1}{\sqrt{2^{j-1}}}\,\psi^3\Big(\frac{t}{2^{j-1}}\Big).$$

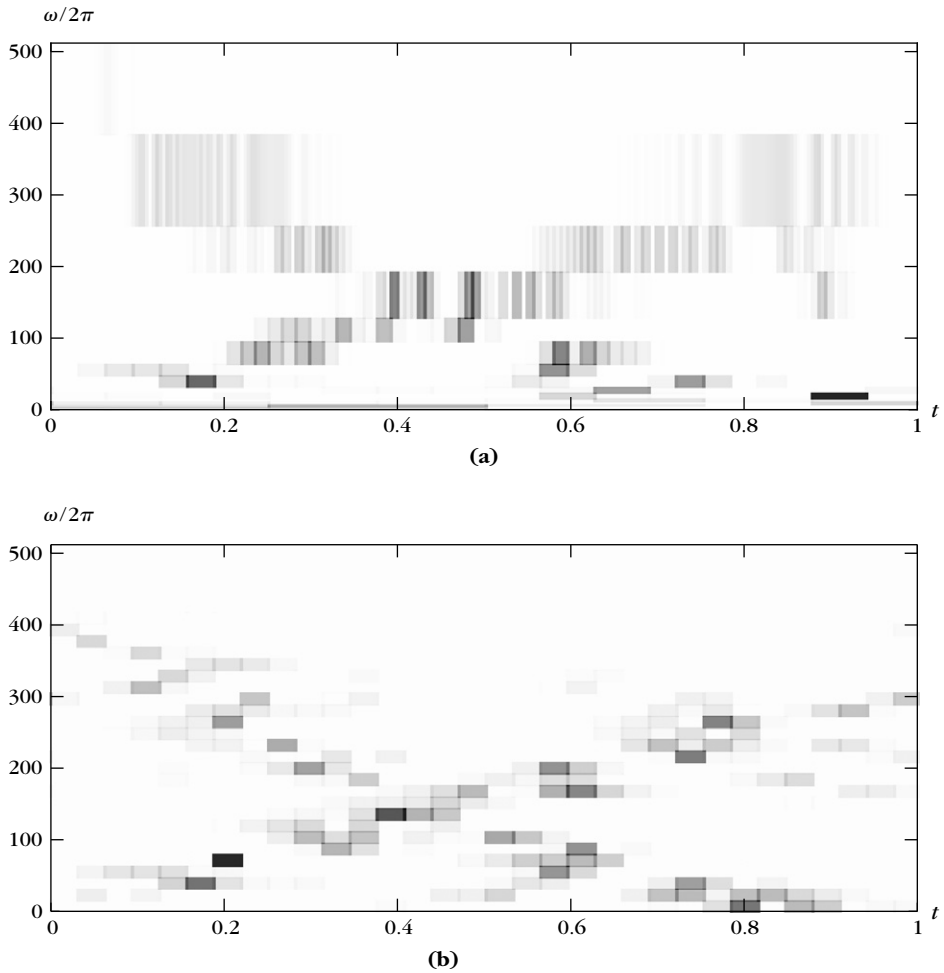

FIGURE 8.8

(a) Heisenberg boxes of a two-band wavelet decomposition of the multichirp signal shown in Figure 8.6. **(b)** Decomposition of the same signal in a pseudo-local cosine wavelet packet basis.

Over positive frequencies, the energy of $\hat{\psi}^2$ and $\hat{\psi}^3$ is mostly concentrated, respectively, in $[\pi, 3\pi/2]$ and $[3\pi/2, 2\pi]$. Thus, the octave bandwidths of $\hat{\psi}^2$ and $\hat{\psi}^3$ are equal to $1/2$ and $1/3$, respectively. Wavelets ψ^2 and ψ^3 have a higher-frequency resolution than ψ^1, but their time support is twice as large. Figure 8.8(a) gives a two-band wavelet decomposition of the multichirp signal shown in Figure 8.6, calculated with a Daubechies 10 filter.

Higher-resolution wavelet bases can be constructed with an arbitrary number of $M = 2^l$ wavelets. In a frequency-ordered wavelet packet tree, we define an admissible

binary tree with leaves indexed by $2^l \leq k < 2^{l+1}$ at the depth $j - L > l$. The resulting wavelet packet basis

$$\{\psi_j^k(t - 2^j n)\}_{M \leq k < 2M,\, j > L+l}$$

can be written as dilations and translations of M elementary wavelets

$$\psi_j^k(t) = \frac{1}{\sqrt{2^{j-l}}}\, \psi^k\Big(\frac{t}{2^{j-l}}\Big).$$

The support size of ψ^k is proportional to $M = 2^l$. For positive frequencies, the energy of $\hat{\psi}^k$ is mostly concentrated in $[k\pi 2^{-l}, (k+1)\pi 2^{-l}]$. The octave bandwidth is therefore $\pi 2^{-l}/(k\pi 2^{-l}) = k^{-1}$ for $M \leq k < 2M$. The M wavelets $\{\psi^k\}_{M \leq k < 2M}$ have an octave bandwidth smaller than M^{-1} but a time support M times larger than the support of ψ^1. Such wavelet bases are called *M-band wavelets*. More general families of M-band wavelets can also be constructed with other M-band filter banks studied in [68].

Pseudo-Local Cosine Bases

Pseudo-local cosine bases are constructed with an admissible binary tree that is a full tree of depth $J - L \geq 0$. The leaves are the nodes indexed by $0 \leq k < 2^{J-L}$ and the resulting wavelet packet basis is

$$\{\psi_J^k(t - 2^J n)\}_{n \in \mathbb{Z},\, 0 \leq k < 2^{J-L}}. \tag{8.34}$$

If these wavelet packets are constructed with a conjugate mirror filter of size K, they have a support of size $(K-1)2^J$. For positive frequencies, the energy of $\hat{\psi}_j^k$ is concentrated in $[k\pi 2^{-J}, (k+1)\pi 2^{-J}]$. Therefore, the bandwidth of all these wavelet packets is approximately constant and equal to $\pi 2^{-J}$. The Heisenberg boxes of these wavelet packets have the same size and divide the time-frequency plane in the rectangular grid illustrated in Figure 8.9.

Shannon wavelet packets ψ_J^k are written in (8.29) as a dilated window θ modulated by cosine functions of frequency $2^{-J}(k + 1/2)\pi$. In this case, the uniform

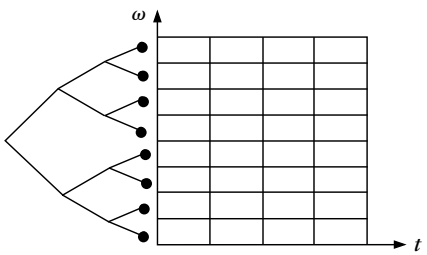

FIGURE 8.9

Admissible tree (*left*) and Heisenberg boxes (*right*) of a wavelet packet pseudo-local cosine basis.

wavelet packet basis (8.34) is a local cosine basis with windows of constant size. This result is not valid for wavelet packets constructed with different conjugate mirror filters. Nevertheless, the time and frequency resolution of uniform wavelet packet bases (8.34) remains constant, like that of local cosine bases constructed with windows of constant size. Figure 8.8(b) gives the decomposition coefficients of a signal in such a uniform wavelet packet basis.

Best Basis

Applications of orthogonal bases often rely on their ability to efficiently approximate signals with only a few nonzero vectors. Choosing a wavelet packet basis that concentrates the signal energy over a few coefficients also reveals its time-frequency structures. Section 12.2.2 describes a fast algorithm that searches for a "best" basis that minimizes a Schur concave cost function among all wavelet packet bases. The wavelet packet basis of Figure 8.6 is calculated with this best basis search.

8.1.4 Wavelet Packet Filter Banks

Wavelet packet coefficients are computed with a filter bank algorithm that generalizes the fast discrete wavelet transform. This algorithm is a straightforward iteration of the two-channel filter bank decomposition presented in Section 7.3.2. Therefore, it was used in signal processing by Croisier, Esteban, and Galand [189] when they introduced the first family of perfect reconstruction filters. The algorithm is presented here from a wavelet packet point of view.

To any discrete signal input $b[n]$ sampled at intervals $N^{-1} = 2^L$, as in (7.111) we associate $f \in \mathbf{V}_L$ with decomposition coefficients $a_L[n] = \langle f, \phi_{L,n} \rangle$ that satisfy

$$b[n] = N^{1/2}\, a_L[n] \approx f(N^{-1}n). \tag{8.35}$$

For any node (j, p) of the wavelet packet tree, we denote the wavelet packet coefficients

$$d_j^p[n] = \langle f(t), \psi_j^p(t - 2^j n) \rangle.$$

At the root of the tree $d_L^0[n] = a_L[n]$ is computed from $b[n]$ with (8.35).

Wavelet Packet Decomposition

We denote $\bar{x}[n] = x[-n]$ and by $\check{x}$ the signal obtained by inserting a zero between each sample of x. Theorem 8.5 generalizes the fast wavelet transform Theorem 7.10.

Theorem 8.5. At the decomposition

$$d_{j+1}^{2p}[k] = d_j^p \star \bar{h}[2k] \quad \text{and} \quad d_{j+1}^{2p+1}[k] = d_j^p \star \bar{g}[2k]. \tag{8.36}$$

At the reconstruction

$$d_j^p[k] = \check{d}_{j+1}^{2p} \star h[k] + \check{d}_{j+1}^{2p+1} \star g[k]. \tag{8.37}$$

The proof of these equations is identical to the proof of Theorem 7.10. The coefficients of wavelet packet children d_{j+1}^{2p} and d_{j+1}^{2p+1} are obtained by subsampling

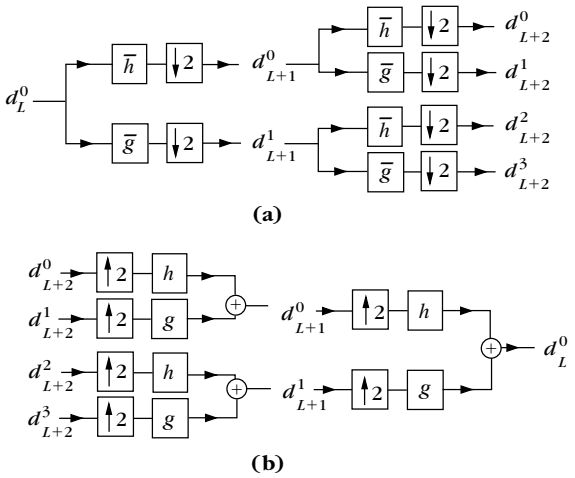

FIGURE 8.10

(a) Wavelet packet filter bank decomposition with successive filterings and subsamplings. **(b)** Reconstruction by inserting zeros and filtering the outputs.

the convolutions of d_j^p with $\bar{h}$ and $\bar{g}$. Iterating these equations along the branches of a wavelet packet tree computes all wavelet packet coefficients, as illustrated by Figure 8.10(a). From the wavelet packet coefficients at the leaves $\{j_i,\ p_i\}_{1 \leq i \leq I}$ of an admissible subtree, we recover d_L^0 at the top of the tree by computing (8.37) for each node inside the tree, as illustrated by Figure 8.10(b). ■

Finite Signals

If a_L is a finite signal of size $2^{-L} = N$, we are facing the same border convolution problems as in a fast discrete wavelet transform. One approach explained in Section 7.5.1 is to periodize the wavelet packet basis. The convolutions (8.36) are then replaced by circular convolutions. To avoid introducing sharp transitions with the periodization, one can also use the border filters described in Section 7.5.3. In either case, d_j^p has 2^{-j} samples. At any depth $j - L$ of the tree, the wavelet packet signals $\{d_j^p\}_{0 \leq p < 2^{j-L}}$ include a total of N coefficients. Since the maximum depth is $\log_2 N$, there are at most $N \log_2 N$ coefficients in a full wavelet packet tree.

In a full wavelet packet tree of depth $\log_2 N$, all coefficients are computed by iterating (8.36) for $L \leq j < 0$. If h and g have K nonzero coefficients, this requires $KN \log_2 N$ additions and multiplications. This is quite spectacular since there are more than $2^{N/2}$ different wavelet packet bases included in this wavelet packet tree.

The computational complexity to recover $a_L = d_L^0$ from the wavelet packet coefficients of an admissible tree increases with the number of inside nodes of the admissible tree. When the admissible tree is the full binary tree of depth $\log_2 N$, the number of operations is maximum and equal to $KN \log_2 N$ multiplications and

additions. If the admissible subtree is a wavelet tree, we need fewer than $2KN$ multiplications and additions.

Discrete Wavelet Packet Bases of $\ell^2(\mathbb{Z})$

The signal decomposition in a conjugate mirror filter bank can also be interpreted as an expansion in discrete wavelet packet bases of $\ell^2(\mathbb{Z})$. This is proved with a result similar to Theorem 8.1.

Theorem 8.6. Let $\{\theta_j[m-2^{j-L}n]\}_{n\in\mathbb{Z}}$ be an orthonormal basis of a space $\mathbf{U}_j$ with $j-L\in\mathbb{N}$. Define

$$\theta^0_{j+1}[m]=\sum_{n=-\infty}^{+\infty} h[n]\,\theta_j[m-2^{j-L}n], \quad \theta^1_{j+1}[m]=\sum_{n=-\infty}^{+\infty} g[n]\,\theta_j[m-2^{j-L}n]. \tag{8.38}$$

The family

$$\left\{\theta^0_{j+1}[m-2^{j+1-L}n],\ \theta^1_{j+1}[m-2^{j+1-L}n]\right\}_{n\in\mathbb{Z}}$$

is an orthonormal basis of $\mathbf{U}_j$.

The proof is similar to the proof of Theorem 8.1. As in the continuous-time case, we derive from this theorem a binary tree of discrete wavelet packets. At the root of the discrete wavelet packet tree is the space $\mathbf{W}^0_L=\ell^2(\mathbb{Z})$ of discrete signals obtained with a sampling interval $N^{-1}=2^L$. It admits a canonical basis of Diracs $\{\psi^0_L[m-n]=\delta[m-n]\}_{n\in\mathbb{Z}}$. The signal $a_L[m]$ is specified by its sample values in this basis. One can verify that the convolutions and subsamplings (8.36) compute

$$d^p_j[n]=\langle a_L[m],\psi^p_j[m-2^{j-L}n]\rangle,$$

where $\{\psi^p_j[m-2^{j-L}n]\}_{n\in\mathbb{Z}}$ is an orthogonal basis of a space $\mathbf{W}^p_j$. These discrete wavelet packets are recursively defined for any $j\geqslant L$ and $0\leqslant p<2^{j-L}$ by

$$\psi^{2p}_{j+1}[m]=\sum_{n=-\infty}^{+\infty} h[n]\,\psi^p_j[m-2^{j-L}n], \quad \psi^{2p+1}_{j+1}[m]=\sum_{n=-\infty}^{+\infty} g[n]\,\psi^p_j[m-2^{j-L}n]. \tag{8.39}$$

8.2 IMAGE WAVELET PACKETS

8.2.1 Wavelet Packet Quad-Tree

We construct wavelet packet bases of $\mathbf{L}^2(\mathbb{R}^2)$ with separable products of wavelet packets $\psi^p_j(x_1-2^jn_1)\,\psi^q_j(x_2-2^jn_2)$ having the same scale along x_1 and x_2. These separable wavelet packet bases are associated to quad-trees, and divide the two-dimensional Fourier plane (ω_1,ω_2) into square regions of varying sizes. Separable wavelet packet bases are extensions of separable wavelet bases.

If images are approximated at the scale 2^L, to the root of the quad-tree we associate the approximation space $\mathbf{V}_L^2 = \mathbf{V}_L \otimes \mathbf{V}_L \subset \mathbf{L^2}(\mathbb{R}^2)$ defined in Section 7.7.1. In Section 8.1.1 we explain how to decompose $\mathbf{V}_L$ with a binary tree of wavelet packet spaces $\mathbf{W}_j^p \subset \mathbf{V}_L$ that admit an orthogonal basis $\{\psi_j^p(t - 2^j n)\}_{n\in\mathbb{Z}}$. The two-dimensional wavelet packet quad-tree is composed of separable wavelet packet spaces. Each node of this quad-tree is labeled by a scale 2^j and two integers $0 \leqslant p < 2^{j-L}$ and $0 \leqslant q < 2^{j-L}$, and corresponds to a separable space

$$\mathbf{W}_j^{p,q} = \mathbf{W}_j^p \otimes \mathbf{W}_j^q. \tag{8.40}$$

The resulting separable wavelet packet for $x = (x_1, x_2)$ is

$$\psi_j^{p,q}(x) = \psi_j^p(x_1)\, \psi_j^q(x_2).$$

Theorem 7.25 proves that an orthogonal basis of $\mathbf{W}_j^{p,q}$ is obtained with a separable product of the wavelet packet bases of $\mathbf{W}_j^p$ and $\mathbf{W}_j^q$, which can be written as

$$\left\{\psi_j^{p,q}(x - 2^j n)\right\}_{n\in\mathbb{Z}^2}.$$

At the root $\mathbf{W}_L^{0,0} = \mathbf{V}_L^2$, the wavelet packet is a two-dimensional scaling function

$$\psi_L^{0,0}(x) = \phi_L^2(x) = \phi_L(x_1)\, \phi_L(x_2).$$

One-dimensional wavelet packet spaces satisfy

$$\mathbf{W}_j^p = \mathbf{W}_{j+1}^{2p} \oplus \mathbf{W}_{j+1}^{2p+1} \quad \text{and} \quad \mathbf{W}_j^q = \mathbf{W}_{j+1}^{2q} \oplus \mathbf{W}_{j+1}^{2q+1}.$$

Inserting these equations in (8.40) proves that $\mathbf{W}_j^{p,q}$ is the direct sum of the four orthogonal subspaces

$$\mathbf{W}_j^{p,q} = \mathbf{W}_{j+1}^{2p,2q} \oplus \mathbf{W}_{j+1}^{2p+1,2q} \oplus \mathbf{W}_{j+1}^{2p,2q+1} \oplus \mathbf{W}_{j+1}^{2p+1,2q+1}. \tag{8.41}$$

These subspaces are located at the four children nodes in the quad-tree, as shown by Figure 8.11. We call any quad-tree that has nodes with either zero or four children an *admissible quad-tree*. Let $\{j_i, p_i, q_i\}_{0\leqslant i\leqslant I}$ be the indices of the nodes at the leaves of an admissible quad-tree. Applying recursively the reconstruction sum (8.41) along the branches of this quad-tree gives an orthogonal decomposition of $\mathbf{W}_L^{0,0}$:

$$\mathbf{W}_L^{0,0} = \bigoplus_{i=1}^{I} \mathbf{W}_{j_i}^{p_i,q_i}.$$

The union of the corresponding wavelet packet bases

$$\left\{\psi_{j_i}^{p_i,q_i}(x - 2^{j_i} n)\right\}_{(n_1,n_2)\in\mathbb{Z}^2,\ 1\leqslant i\leqslant I}$$

is therefore an orthonormal basis of $\mathbf{V}_L^2 = \mathbf{W}_L^{0,0}$.

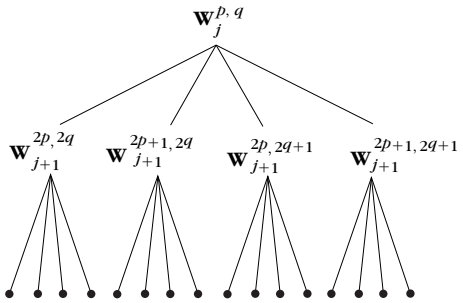

FIGURE 8.11

A wavelet packet quad-tree for images is constructed recursively by decomposing each separable space $\mathbf{W}_j^{p,q}$ in four subspaces.

Number of Wavelet Packet Bases

The number of different bases in a full wavelet packet quad-tree of depth J is equal to the number of admissible subtrees. Theorem 8.7 proves that there are more than $2^{4^{J-1}}$ such bases.

Theorem 8.7. The number B_J of wavelet packet bases in a full wavelet packet quad-tree of depth J satisfies

$$2^{4^{J-1}} \leq B_J \leq 2^{\frac{49}{48}4^{J-1}}.$$

Proof. This result is proved with induction, as in the proof of Theorem 8.7. The reader can verify that B_J satisfies an induction relation similar to (8.16):

$$B_{J+1} = B_J^4 + 1. \tag{8.42}$$

Since $B_0 = 1, B_1 = 2$, and $B_{J+1} \geq B_J^4$, we derive that $B_J \geq 2^{4^{J-1}}$. Moreover, for $J \geq 1$

$$\log_2 B_{J+1} = 4 \log_2 B_J + \log_2(1 + B_J^{-4}) \leq 4 \log_2 B_J + \frac{1}{16} \leq 4^J + \frac{1}{16} \sum_{j=0}^{J-1} 4^j,$$

which implies that $B_J \geq 2^{\frac{49}{48}4^{J-1}}$. ■

For an image of $N = N_1 N_2$, pixels, if $N_1 = N_2$, then the wavelet packet quad-tree has a depth at most $\log_2 N_1 = \log_2 N^{1/2}$. Thus, the number of wavelet packet bases satisfies

$$2^{\frac{N}{4}} \leq B_{\log_2 N} \leq 2^{\frac{49}{48}\frac{N}{4}}. \tag{8.43}$$

Spatial and Frequency Localization

The spatial and frequency localization of two-dimensional wavelet packets is derived from the time-frequency analysis performed in Section 8.1.2. If the conjugate mirror

filter h has K nonzero coefficients, we proved that ψ_j^p has a support size of $2^j(K-1)$, thus $\psi_j^p(x_1)\,\psi_j^q(x_2)$ has a square support of width $2^j(K-1)$.

We showed that the Fourier transform of ψ_j^p has its energy mostly concentrated in

$$[-(k+1)2^{-j}\pi, -k2^{-j}\pi]\cup[k2^{-j}\pi, (k+1)2^{-j}\pi],$$

where $k = G[p]$ is specified by Theorem 8.3. Therefore, the Fourier transform of a two-dimensional wavelet packet $\psi_j^{p,q}$ has its energy mostly concentrated in

$$[k_1 2^{-j}\pi, (k_1+1)2^{-j}\pi]\times[k_2 2^{-j}\pi, (k_2+1)2^{-j}\pi], \tag{8.44}$$

with $k_1 = G[p]$ and $k_2 = G[q]$, and in the three squares that are symmetric with respect to the two axes $\omega_1 = 0$ and $\omega_2 = 0$. An admissible wavelet packet quad-tree decomposes the positive-frequency quadrant into squares of dyadic sizes, as illustrated in Figure 8.12. For example, the leaves of a full wavelet packet quad-tree of depth $j - L$ define a wavelet packet basis that decomposes the positive-frequency quadrant into squares of constant width equal to $2^{-j}\pi$. This wavelet packet basis is similar to a two-dimensional local cosine basis with windows of constant size.

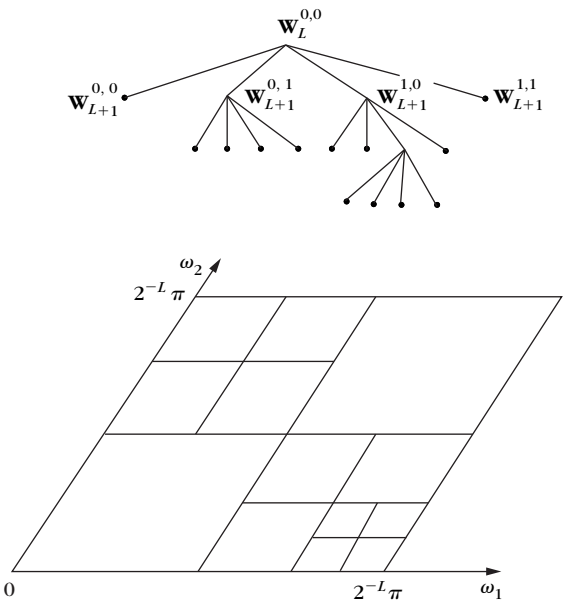

FIGURE 8.12

A wavelet packet quad-tree decomposes the positive-frequency quadrant into squares of progressively smaller sizes as we go down the tree.

8.2.2 Separable Filter Banks

The decomposition coefficients of an image in a separable wavelet packet basis are computed with a separable extension of the filter bank algorithm described in Section 8.1.4. Let $b[n]$ be an input image with pixels at a distance 2^L. We associate $b[n]$ to a function $f \in \mathbf{V}_L^2$ approximated at the scale 2^L, with decomposition coefficients $a_L[n] = \langle f(x),\ \phi_L^2(x - 2^L n)\rangle$ that are defined like in (7.232):

$$b[n] = 2^{-L}\, a_L[n] \approx f(2^L n).$$

The wavelet packet coefficients

$$d_j^{p,q}[n] = \langle f,\, \psi_j^{p,q}(x - 2^j n)\rangle$$

characterize the orthogonal projection of f in $\mathbf{W}_j^{p,q}$. At the root, $d_L^{0,0} = a_L$.

Separable Filter Bank

From the separability of wavelet packet bases and the one-dimensional convolution formula of Theorem (8.5), we derive that for any $n = (n_1, n_2)$,

$$d_{j+1}^{2p,2q}[n] = d_j^{p,q} \star \bar{h}\bar{h}[2n], \quad d_{j+1}^{2p+1,2q}[n] = d_j^{p,q} \star \bar{g}\bar{h}[2n], \tag{8.45}$$

$$d_{j+1}^{2p,2q+1}[n] = d_j^{p,q} \star \bar{h}\bar{g}[2n], \quad d_{j+1}^{2p+1,2q+1}[n] = d_j^{p,q} \star \bar{g}\bar{g}[2n]. \tag{8.46}$$

Thus, the coefficients of a wavelet packet quad-tree are computed by iterating these equations along the quad-tree's branches. The calculations are performed with separable convolutions along the rows and columns of the image, as illustrated in Figure 8.13.

At the reconstruction,

$$\begin{aligned} d_j^{p,q}[n] = {} & \check{d}_{j+1}^{2p,2q} \star hh[n] + \check{d}_{j+1}^{2p+1,2q} \star gh[n] \\ & + \check{d}_{j+1}^{2p,2q+1} \star hg[n] + \check{d}_{j+1}^{2p+1,2q+1} \star gg[n]. \end{aligned} \tag{8.47}$$

The image $a_L = d_L^{0,0}$ is reconstructed from wavelet packet coefficients stored at the leaves of any admissible quad-tree by repeating the partial reconstruction (8.47) in the inside nodes of this quad-tree.

Finite Images

If the image a_L has $N = 2^{-2L}$ pixels, the one-dimensional convolution border problems are solved with one of the two approaches described in Sections 7.5.1 and 7.5.3. Each wavelet packet image $d_j^{p,q}$ includes 2^{-2j} pixels. At depth $j - L$, there are N wavelet packet coefficients in $\{d_j^{p,q}\}_{0 \leqslant p,q < 2^{j-L}}$. Thus, a quad-tree of maximum depth $\log_2 N^{1/2}$ includes $N \log_2 N^{1/2}$ coefficients. If h and g have K nonzero coefficients, the one-dimensional convolutions that implement (8.45) and (8.46) require $2K2^{-2j}$ multiplications and additions. Thus, all wavelet packet coefficients at depth $j + 1 - L$ are computed from wavelet packet coefficients located at depth $j - L$ with

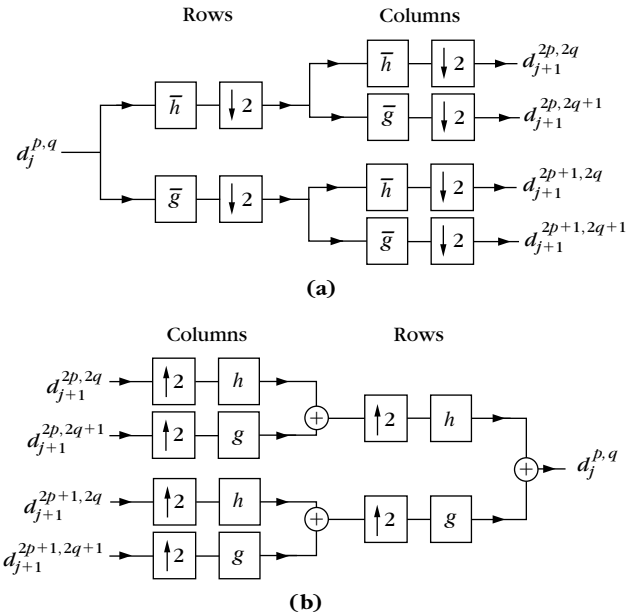

FIGURE 8.13

(a) Wavelet packet decomposition implementing (8.45) and (8.46) with one-dimensional convolutions along the rows and columns of $d_1^{p,q}$. **(b)** Wavelet packet reconstruction implementing (8.47).

$2KN$ calculations. The $N \log_2 N^{1/2}$ wavelet packet coefficients of a full tree of depth $\log_2 N^{1/2}$ are therefore obtained with $KN \log_2 N$ multiplications and additions. The numerical complexity of reconstructing a_L from a wavelet packet basis depends on the number of inside nodes of the corresponding quad-tree. The worst case is a reconstruction from the leaves of a full quad-tree of depth $\log_2 N^{1/2}$, which requires $KN \log_2 N$ multiplications and additions.

8.3 BLOCK TRANSFORMS

Wavelet packet bases are designed by dividing the frequency axis in intervals of varying sizes. These bases are particularly well adapted to decomposing signals that have different behavior in different frequency intervals. If f has properties that vary in time, it is then more appropriate to decompose f in a *block basis* that segments the time axis in intervals with sizes that are adapted to the signal structures. Section 8.3.1 explains how to generate a block basis of $\mathbf{L^2}(\mathbb{R})$ from any basis of $\mathbf{L^2}[0, 1]$. The block cosine bases described in Sections 8.3.2 and 8.3.3 are used for signal and image compression.

8.3.1 Block Bases

Block orthonormal bases are obtained by dividing the time axis in consecutive intervals $[a_p, a_{p+1}]$ with

$$\lim_{p\to-\infty} a_p = -\infty \quad \text{and} \quad \lim_{p\to+\infty} a_p = +\infty.$$

The size $l_p = a_{p+1} - a_p$ of each interval is arbitrary. Let $g = \mathbf{1}_{[0,1]}$. An interval is covered by the dilated rectangular window

$$g_p(t) = \mathbf{1}_{[a_p,a_{p+1}]}(t) = g\left(\frac{t-a_p}{l_p}\right). \tag{8.48}$$

Theorem 8.8 constructs a block orthogonal basis of $\mathbf{L^2}(\mathbb{R})$ from a single orthonormal basis of $\mathbf{L^2}[0, 1]$.

Theorem 8.8. If $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbf{L^2}[0, 1]$, then

$$\left\{g_{p,k}(t) = g_p(t)\,\frac{1}{\sqrt{l_p}}\,e_k\Big(\frac{t-a_p}{l_p}\Big)\right\}_{(p,k)\in\mathbb{Z}} \tag{8.49}$$

is a block orthonormal basis of $\mathbf{L^2}(\mathbb{R})$.

Proof. One can verify that the dilated and translated family

$$\left\{\frac{1}{\sqrt{l_p}}\,e_k\Big(\frac{t-a_p}{l_p}\Big)\right\}_{k\in\mathbb{Z}} \tag{8.50}$$

is an orthonormal basis of $\mathbf{L^2}[a_p, a_{p+1}]$. If $p \neq q$, then $\langle g_{p,k}, g_{q,k}\rangle = 0$ since their supports do not overlap. Thus, the family (8.49) is orthonormal. To expand a signal f in this family, it is decomposed as a sum of separate blocks

$$f(t) = \sum_{p=-\infty}^{+\infty} f(t)\,g_p(t),$$

and each block $f(t)g_p(t)$ is decomposed in the basis (8.50). ■

Block Fourier Basis

A block basis is constructed with the Fourier basis of $\mathbf{L^2}[0, 1]$:

$$\left\{e_k(t) = \exp(i2k\pi t)\right\}_{k\in\mathbb{Z}}.$$

The time support of each block Fourier vector $g_{p,k}$ is $[a_p, a_{p+1}]$ of size l_p. The Fourier transform of $g = \mathbf{1}_{[0,1]}$ is

$$\hat{g}(\omega) = \frac{\sin(\omega/2)}{\omega/2}\,\exp\left(\frac{i\omega}{2}\right)$$

and

$$\hat{g}_{p,k}(\omega) = \sqrt{l_p}\,\hat{g}(l_p\omega - 2k\pi)\,\exp\left(\frac{-i2\pi k a_p}{l_p}\right).$$

It is centered at $2k\pi\, l_p^{-1}$ and has a slow asymptotic decay proportional to $l_p^{-1}|\omega|^{-1}$. Because of this poor frequency localization, even though a signal f is smooth, its decomposition in a block Fourier basis may include large high-frequency coefficients. This can also be interpreted as an effect of periodization.

Discrete Block Bases

For all $p \in \mathbb{Z}$, we suppose that $a_p \in \mathbb{Z}$. Discrete block bases are built with discrete rectangular windows having supports on intervals $[a_p, a_{p+1} - 1]$:

$$g_p[n] = \mathbf{1}_{[a_p, a_{p+1}-1]}(n).$$

Since dilations are not defined in a discrete framework, we generally cannot derive bases of intervals of varying sizes from a single basis. Thus, Theorem 8.9 supposes that we can construct an orthonormal basis of $\mathbb{C}^l$ for any $l > 0$. The proof is straightforward.

Theorem 8.9. Suppose that $\{e_{k,l}\}_{0 \leq k < l}$ is an orthogonal basis of $\mathbb{C}^l$ for any $l > 0$. The family

$$\left\{g_{p,k}[n] = g_p[n]\, e_{k,l_p}[n - a_p]\right\}_{0 \leq k < l_p, p \in \mathbb{Z}} \tag{8.51}$$

is a block orthonormal basis of $\ell^2(\mathbb{Z})$.

A discrete block basis is constructed with discrete Fourier bases

$$\left\{e_{k,l}[n] = \frac{1}{\sqrt{l}}\,\exp\left(\frac{i2\pi k n}{l}\right)\right\}_{0 \leq k < l}.$$

The resulting block Fourier vectors $g_{p,k}$ have sharp transitions at the window border, and thus are not well localized in frequency. As in the continuous case, the decomposition of smooth signals f may produce large-amplitude, high-frequency coefficients because of border effects.

Block Bases of Images

General block bases of images are constructed by partitioning the plane $\mathbb{R}^2$ into rectangles $\{[a_p, b_p] \times [c_p, d_p]\}_{p \in \mathbb{Z}}$ of arbitrary length $l_p = b_p - a_p$ and width $w_p = d_p - c_p$. Let $\{e_k\}_{k \in \mathbb{Z}}$ be an orthonormal basis of $\mathbf{L}^2[0, 1]$ and $g = \mathbf{1}_{[0,1]}$. We denote

$$g_{p,k,j}(x, y) = g\left(\frac{x - a_p}{l_p}\right) g\left(\frac{y - c_p}{w_p}\right) \frac{1}{\sqrt{l_p w_p}}\, e_k\left(\frac{x - a_p}{l_p}\right) e_j\left(\frac{y - c_p}{w_p}\right).$$

The family $\{g_{p,k,j}\}_{(k,j) \in \mathbb{Z}^2}$ is an orthonormal basis of $\mathbf{L}^2([a_p, b_p] \times [c_p, d_p])$, and thus $\left\{g_{p,k,j}\right\}_{(p,k,j) \in \mathbb{Z}^3}$ is an orthonormal basis of $\mathbf{L}^2(\mathbb{R}^2)$.

For discrete images, we define discrete windows that cover each rectangle

$$g_p = \mathbf{1}_{[a_p, b_p - 1] \times [c_p, d_p - 1]}.$$

If $\{e_{k,l}\}_{0 \leq k < l}$ is an orthogonal basis of $\mathbb{C}^l$ for any $l > 0$, then

$$\left\{ g_{p,k,j}[n_1, n_2] = g_p[n_1, n_2]\, e_{k,l_p}[n_1 - a_p]\, e_{j,w_p}[n_2 - c_p] \right\}_{(k,j,p) \in \mathbb{Z}^3}$$

is a block basis of $\ell^2(\mathbb{Z}^2)$.

8.3.2 Cosine Bases

If $f \in \mathbf{L}^2[0, 1]$ and $f(0) \neq f(1)$, even though f might be a smooth function, the Fourier coefficients

$$\langle f(u), \mathrm{e}^{i2k\pi u} \rangle = \int_0^1 f(u)\, \mathrm{e}^{-i2k\pi u}\, du$$

have a relatively large amplitude at high frequencies $2k\pi$. Indeed, the Fourier series expansion

$$f(t) = \sum_{k=-\infty}^{+\infty} \langle f(u), \mathrm{e}^{i2k\pi u} \rangle\, \mathrm{e}^{i2k\pi t}$$

is a function of period 1, equal to f over $[0, 1]$, and that is therefore discontinuous if $f(0) \neq f(1)$. This shows that the restriction of a smooth function to an interval generates large Fourier coefficients. As a consequence, block Fourier bases are rarely used. A cosine I basis reduces this border effect by restoring a periodic extension $\tilde{f}$ of f, which is continuous if f is continuous. Thus, high-frequency cosine I coefficients have a smaller amplitude than Fourier coefficients.

Cosine I Basis

We define $\tilde{f}$ to be the function of period 2 that is symmetric at about 0 and equal to f over $[0, 1]$:

$$\tilde{f}(t) = \begin{cases} f(t) & \text{for } t \in [0, 1] \\ f(-t) & \text{for } t \in (-1, 0). \end{cases} \tag{8.52}$$

If f is continuous over $[0, 1]$, then $\tilde{f}$ is continuous over $\mathbb{R}$, as shown in Figure 8.14. However, if f has a nonzero right derivative at 0 or left derivative at 1, then $\tilde{f}$ is nondifferentiable at integer points.

The Fourier expansion of $\tilde{f}$ over $[0, 2]$ can be written as a sum of sine and cosine terms:

$$\tilde{f}(t) = \sum_{k=0}^{+\infty} a[k] \cos\left(\frac{2\pi k t}{2}\right) + \sum_{k=1}^{+\infty} b[k] \sin\left(\frac{2\pi k t}{2}\right).$$

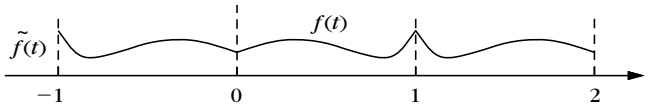

FIGURE 8.14

The function $\tilde{f}(t)$ is an extension of $f(t)$; it is symmetric at about 0 and of period 2.

The sine coefficients $b[k]$ are zero because $\tilde{f}$ is even. Since $f(t) = \tilde{f}(t)$ over $[0, 1]$, this proves that any $f \in \mathbf{L^2}[0, 1]$ can be written as a linear combination of the cosines $\{\cos(k\pi t)\}_{k\in\mathbb{N}}$. One can verify that this family is orthogonal over $[0, 1]$. Therefore, it is an orthogonal basis of $\mathbf{L^2}[0, 1]$, as stated by Theorem 8.10.

Theorem 8.10: *Cosine I.* The family

$$\left\{\lambda_k \sqrt{2}\, \cos(\pi k t)\right\}_{k\in\mathbb{N}} \quad \text{with} \quad \lambda_k = \begin{cases} 2^{-1/2} & \text{if } k = 0 \\ 1 & \text{if } k \neq 0 \end{cases}$$

is an orthonormal basis of $\mathbf{L^2}[0, 1]$. ■

Block Cosine Basis

Let us divide the real line with square windows $g_p = \mathbf{1}_{[a_p, a_{p+1}]}$. Theorem 8.8 proves that

$$\left\{g_{p,k}(t) = g_p(t) \sqrt{\frac{2}{l_p}} \lambda_k \cos\left(\pi k \frac{t - a_p}{l_p}\right)\right\}_{k\in\mathbb{N}, p\in\mathbb{Z}}$$

is a block basis of $\mathbf{L^2}(\mathbb{R})$. The decomposition coefficients of a smooth function have a faster decay at high frequencies in a block cosine basis than in a block Fourier basis, because cosine bases correspond to a smoother signal extension beyond the intervals $[a_p, a_{p+1}]$.

Cosine IV Basis

Other cosine bases are constructed from the Fourier series, with different extensions of f beyond $[0, 1]$. The cosine IV basis appears in fast numerical computations of cosine I coefficients. It is also used to construct local cosine bases with smooth windows in Section 8.4.2.

Any $f \in \mathbf{L^2}[0, 1]$ is extended into a function $\tilde{f}$ of period 4, which is symmetric about 0 and antisymmetric about 1 and -1:

$$\tilde{f}(t) = \begin{cases} f(t) & \text{if } t \in [0, 1] \\ f(-t) & \text{if } t \in (-1, 0) \\ -f(2 - t) & \text{if } t \in [1, 2) \\ -f(2 + t) & \text{if } t \in [-1, -2) \end{cases}$$

If $f(1) \neq 0$, the antisymmetry at 1 creates a function $\tilde{f}$ that is discontinuous at $f(2n + 1)$ for any $n \in \mathbb{Z}$, as shown in Figure 8.15. Therefore, this extension is less regular than the cosine I extension (8.52).

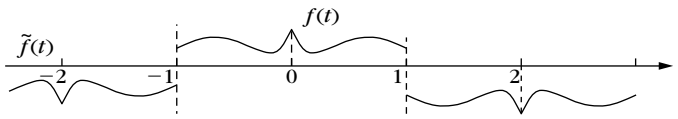

FIGURE 8.15

A cosine IV extends $f(t)$ into a signal $\tilde{f}(t)$ of period 4, which is symmetric with respect to 0 and antisymmetric with respect to 1.

Since $\tilde{f}$ is 4 periodic, it can be decomposed as a sum of sines and cosines of period 4:

$$\tilde{f}(t) = \sum_{k=0}^{+\infty} a[k] \cos\left(\frac{2\pi kt}{4}\right) + \sum_{k=1}^{+\infty} b[k] \sin\left(\frac{2\pi kt}{4}\right).$$

The symmetry about 0 implies that

$$b[k] = \frac{1}{2}\int_{-2}^{2} \tilde{f}(t) \sin\left(\frac{2\pi kt}{4}\right) dt = 0.$$

For even frequencies, the antisymmetry about 1 and -1 yields

$$a[2k] = \frac{1}{2}\int_{-2}^{2} \tilde{f}(t) \cos\left(\frac{2\pi(2k)t}{4}\right) dt = 0.$$

Thus, the only nonzero components are cosines of odd frequencies:

$$\tilde{f}(t) = \sum_{k=0}^{+\infty} a[2k+1] \cos\left(\frac{(2k+1)2\pi t}{4}\right). \tag{8.53}$$

Since $f(t) = \tilde{f}(t)$ over $[0, 1]$, this proves that any $f \in \mathbf{L}^2[0, 1]$ is decomposed as a sum of such cosine functions. One can verify that the restriction of these cosine functions to $[0, 1]$ is orthogonal in $\mathbf{L}^2[0, 1]$, which implies Theorem 8.11.

Theorem 8.11: *Cosine IV.* The family

$$\left\{\sqrt{2} \cos\left[\left(k + \frac{1}{2}\right)\pi t\right]\right\}_{k\in\mathbb{N}}$$

is an orthonormal basis of $\mathbf{L}^2[0, 1]$. ■

The cosine transform IV is not used in block transforms because it has the same drawbacks as a block Fourier basis. Block cosine IV coefficients of a smooth f have a slow decay at high frequencies, because such a decomposition corresponds to a discontinuous extension of f beyond each block. Section 8.4.2 explains how to avoid this issue with smooth windows.

8.3.3 Discrete Cosine Bases

Discrete cosine bases are derived from the discrete Fourier basis with the same approach as in the continuous time case. To simplify notations, the sampling distance is normalized to 1. If the sampling distance was originally N^{-1}, then the frequency indexes that appear in this section must be multiplied by N.

Discrete Cosine I

A signal $f[n]$ defined for $0 \leqslant n < N$ is extended by symmetry with respect to $-1/2$ into a signal $\tilde{f}[n]$ of size $2N$:

$$\tilde{f}[n] = \begin{cases} f[n] & \text{for } 0 \leqslant n < N \\ f[-n-1] & \text{for } -N \leqslant n \leqslant -1. \end{cases} \tag{8.54}$$

The $2N$ discrete Fourier transform of $\tilde{f}$ can be written as a sum of sine and cosine terms:

$$\tilde{f}[n] = \sum_{k=0}^{N-1} a[k] \cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right] + \sum_{k=0}^{N-1} b[k] \sin\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right].$$

Since $\tilde{f}$ is symmetric about $-1/2$, then necessarily $b[k] = 0$ for $0 \leqslant k < N$. Moreover, $f[n] = \tilde{f}[n]$ for $0 \leqslant n < N$, so any signal $f \in \mathbb{C}^N$ can be written as a sum of these cosine functions. The reader can also verify that these discrete cosine signals are orthogonal in $\mathbb{C}^N$. Thus, we obtain Theorem 8.12.

Theorem 8.12: *Cosine I.* The family

$$\left\{\lambda_k \sqrt{\frac{2}{N}} \cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right]\right\}_{0 \leqslant k < N} \quad \text{with} \quad \lambda_k = \begin{cases} 2^{-1/2} & \text{if } k = 0 \\ 1 & \text{otherwise} \end{cases}$$

is an orthonormal basis of $\mathbb{C}^N$. ■

This theorem proves that any $f \in \mathbb{C}^N$ can be decomposed into

$$f[n] = \frac{2}{N} \sum_{k=0}^{N-1} \hat{f}_I[k] \lambda_k \cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right], \tag{8.55}$$

where

$$\hat{f}_I[k] = \left\langle f[n], \lambda_k \cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right]\right\rangle = \lambda_k \sum_{n=0}^{N-1} f[n] \cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right] \tag{8.56}$$

is the discrete cosine transform I (DCT-I) of f. The next section describes a fast discrete cosine transform that computes $\hat{f}_I$ with $O(N \log_2 N)$ operations.

Discrete Block Cosine Transform

Let us divide the integer set $\mathbb{Z}$ with discrete windows $g_p[n] = \mathbf{1}_{[a_p, a_p-1]}(n)$ with $a_p \in \mathbb{Z}$. Theorem 8.9 proves that the corresponding block basis

$$\left\{ g_{p,k}[n] = g_p[n]\,\lambda_k \sqrt{\frac{2}{l_p}} \cos\left[\frac{k\pi}{l_p}\left(n + \frac{1}{2} - a_p\right)\right] \right\}_{0 \leq k < N, p \in \mathbb{Z}}$$

is an orthonormal basis of $\ell^2(\mathbb{Z})$. Over each block of size $l_p = a_{p+1} - a_p$, the fast DCT-I algorithm computes all coefficients with $O(l_p \log_2 l_p)$ operations. Section 10.5.1 describes the JPEG image compression standard, which decomposes images in a separable block cosine basis. A block cosine basis is used as opposed to a block Fourier basis, because it yields smaller-amplitude, high-frequency coefficients, which improves the coding performance.

Discrete Cosine IV

To construct a discrete cosine IV basis, a signal f of N samples is extended into a signal $\tilde{f}$ of period $4N$, which is symmetric with respect to $-1/2$ and antisymmetric with respect to $N - 1/2$ and $-N + 1/2$. As in (8.53), the decomposition of $\tilde{f}$ over a family of sines and cosines of period $4N$ has no sine terms and no cosine terms of even frequency. Since $\tilde{f}[N] = f[n]$ for $0 \leq n < N$, we derive that f can also be written as a linear expansion of these odd-frequency cosines, which are orthogonal in $\mathbb{C}^N$. Thus, we obtain Theorem 8.13.

Theorem 8.13: *Cosine IV.* The family

$$\left\{ \sqrt{\frac{2}{N}} \cos\left[\frac{\pi}{N}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right] \right\}_{0 \leq k < N}$$

is an orthonormal basis of $\mathbb{C}^N$. ■

This theorem proves that any $f \in \mathbb{C}^N$ can be decomposed into

$$f[n] = \frac{2}{N} \sum_{k=0}^{N-1} \hat{f}_{IV}[k] \cos\left[\frac{\pi}{N}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right], \tag{8.57}$$

where

$$\hat{f}_{IV}[k] = \sum_{n=0}^{N-1} f[n] \cos\left[\frac{\pi}{N}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right] \tag{8.58}$$

is the discrete cosine transform IV (DCT-IV) of f.

8.3.4 Fast Discrete Cosine Transforms

The DCT-IV of a signal of size N is related to the discrete Fourier transform (DFT) of a complex signal of size $N/2$ with a formula introduced by Duhamel, Mahieux, and Petit [40, 236]. By computing this DFT with the fast Fourier transform (FFT)

described in Section 3.3.3, we need $O(N \log_2 N)$ operations to compute the DCT-IV. The DCT-I coefficients are then calculated through an induction relation with the DCT-IV, due to Wang [479].

Fast DCT-IV

To clarify the relation between a DCT-IV and a DFT, we split $f[n]$ in two half-size signals of odd and even indices:

$$b[n] = f[2n],$$
$$c[n] = f[N-1-2n].$$

The DCT-IV (8.58) is rewritten as

$$\begin{aligned}\hat{f}_{IV}[k] &= \sum_{n=0}^{N/2-1} b[n] \cos\left[\left(2n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\frac{\pi}{N}\right] + \\ &\quad \sum_{n=0}^{N/2-1} c[n] \cos\left[\left(N-1-2n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\frac{\pi}{N}\right] \\ &= \sum_{n=0}^{N/2-1} b[n] \cos\left[\left(n+\frac{1}{4}\right)\left(k+\frac{1}{2}\right)\frac{2\pi}{N}\right] + \\ &\quad (-1)^k \sum_{n=0}^{N/2-1} c[n] \sin\left[\left(n+\frac{1}{4}\right)\left(k+\frac{1}{2}\right)\frac{2\pi}{N}\right].\end{aligned}$$

Thus, the even-frequency indices can be expressed as a real part,

$$\begin{aligned}&\hat{f}_{IV}[2k] = \\ &\operatorname{Re}\left\{\exp\left[\tfrac{-i\pi k}{N}\right]\textstyle\sum_{n=0}^{N/2-1}(b[n]+ic[n])\exp\left[-i(n+\frac{1}{4})\frac{\pi}{N}\right]\exp\left[\frac{-i2\pi kn}{N/2}\right]\right\},\end{aligned} \tag{8.59}$$

whereas the odd coefficients correspond to an imaginary part,

$$\begin{aligned}&\hat{f}_{IV}[N-2k-1] = \\ &-\operatorname{Im}\left\{\exp\left[\tfrac{-i\pi k}{N}\right]\textstyle\sum_{n=0}^{N/2-1}(b[n]+ic[n])\exp\left[-i(n+\frac{1}{4})\frac{\pi}{N}\right]\exp\left[\frac{-i2\pi kn}{N/2}\right]\right\}.\end{aligned} \tag{8.60}$$

For $0 \leqslant n < N/2$, we denote

$$g[n] = (b[n] + i\,c[n]) \exp\left[-i\left(n+\frac{1}{4}\right)\frac{\pi}{N}\right].$$

The DFT $\hat{g}[k]$ of $g[n]$ is computed with an FFT of size $N/2$. Equations (8.59) and (8.60) prove that

$$\hat{f}_{IV}[2k] = \operatorname{Re}\left\{\exp\left[\frac{-i\pi k}{N}\right]\hat{g}[k]\right\},$$

and

$$\hat{f}_{IV}[N-2k-1] = -\mathrm{Im}\left\{\exp\left[\frac{-i\pi k}{N}\right]\hat{g}[k]\right\}.$$

The DCT-IV coefficients $\hat{f}_{IV}[k]$ are obtained with one FFT of size $N/2$ plus $O(N)$ operations, which makes a total of $O(N \log_2 N)$ operations. To normalize the DCT-IV, the resulting coefficients must be multiplied by $\sqrt{2/N}$. An efficient implementation of the DCT-IV with a split-radix FFT requires [40]

$$\mu_{DCT-IV}(N) = \frac{N}{2}\log_2 N + N \tag{8.61}$$

real multiplications and

$$\alpha_{DCT-IV}(N) = \frac{3N}{2}\log_2 N \tag{8.62}$$

additions.

The inverse DCT-IV of $\hat{f}_{IV}$ is given by (8.57). Up to the proportionality constant $2/N$, this sum is the same as (8.58), where $\hat{f}_{IV}$ and f are interchanged. This proves that the inverse DCT-IV is computed with the same fast algorithm as the forward DCT-IV.

Fast DCT-I

A DCT-I is calculated with an induction relation that involves the DCT-IV. Regrouping the terms $f[n]$ and $f[N-1-n]$ of a DCT-I (8.56) yields

$$\hat{f}_I[2k] = \lambda_k \sum_{n=0}^{N/2-1} (f[n] + f[N-1-n]) \cos\left[\frac{\pi k}{N/2}\left(n+\frac{1}{2}\right)\right], \tag{8.63}$$

$$\hat{f}_I[2k+1] = \sum_{n=0}^{N/2-1} (f[n] - f[N-1-n]) \cos\left[\frac{\pi\,(k+1/2)}{N/2}\left(n+\frac{1}{2}\right)\right]. \tag{8.64}$$

The even index coefficients of the DCT-I are equal to the DCT-I of the signal $f[n] + f[N-1-n]$ of length $N/2$. The odd coefficients are equal to the DCT-IV of the signal $f[n] - f[N-1-n]$ of length $N/2$. Thus, the number of multiplications of a DCT-I is related to the number of multiplications of a DCT-IV by the induction relation

$$\mu_{DCT-I}(N) = \mu_{DCT-I}(N/2) + \mu_{DCT-IV}(N/2), \tag{8.65}$$

while the number of additions is

$$\alpha_{DCT-I}(N) = \alpha_{DCT-I}(N/2) + \alpha_{DCT-IV}(N/2) + N. \tag{8.66}$$

Since the number of multiplications and additions of a DCT-IV is $O(N \log_2 N)$, this induction relation proves that the number of multiplications and additions of this algorithm is also $O(N \log_2 N)$.

If the DCT-IV is implemented with a split-radix FFT, by inserting (8.61) and (8.62) in the recurrence equations (8.65) and (8.66), we derive that the number of multiplications and additions to compute a DCT-I of size N is

$$\mu_{DCT-I}(N) = \frac{N}{2}\log_2 N + 1, \tag{8.67}$$

and

$$\alpha_{DCT-I}(N) = \frac{3N}{2}\log_2 N - N + 1. \tag{8.68}$$

The inverse DCT-I is computed with a similar recursive algorithm. Applied to $\hat{f}_I$, it is obtained by computing the inverse DCT-IV of the odd index coefficients $\hat{f}_I[2k+1]$ with (8.64) and an inverse DCT-I of size $N/2$ applied to the even coefficients $\hat{f}_I[2k]$ with (8.63). From the values $f[n]+f[N-1-n]$ and $f[n]-f[N-1-n]$, we recover $f[n]$ and $f[N-1-n]$. The inverse DCT-IV is identical to the forward DCT-IV up to a multiplicative constant. Thus, the inverse DCT-I requires the same number of operations as the forward DCT-I.

8.4 LAPPED ORTHOGONAL TRANSFORMS

Cosine and Fourier block bases are computed with discontinuous rectangular windows that divide the real line in disjoint intervals. Multiplying a signal with a rectangular window creates discontinuities that produce large-amplitude coefficients at high frequencies. To avoid these discontinuity artifacts, it is necessary to use smooth windows.

The Balian-Low theorem (5.20) proves that for any u_0 and ξ_0, there exists no differentiable window g of compact support such that

$$\left\{ g(t - nu_0)\ \exp(ik\xi_0 t) \right\}_{(n,k)\in\mathbb{Z}^2}$$

is an orthonormal basis of $\mathbf{L^2}(\mathbb{R})$. This negative result discouraged any research in this direction, until Malvar discovered in discrete signal processing that one could create orthogonal bases with smooth windows modulated by a cosine IV basis [368, 369]. This result was independently rediscovered for continuous-time functions by Coifman and Meyer [181] with a different approach that we shall follow here. The roots of these new orthogonal bases are lapped projectors, which split signals in orthogonal components with overlapping supports [43]. Section 8.4.1 introduces these lapped projectors, the construction of continuous time and discrete lapped orthogonal bases are explained in the following sections. The particular case of local cosine bases is studied in more detail.

8.4.1 Lapped Projectors

Block transforms compute the restriction of f to consecutive intervals $[a_p, a_{p+1}]$ and decompose this restriction in an orthogonal basis of $[a_p, a_{p+1}]$. Formally, the

restriction of f to $[a_p, a_{p+1}]$ is an orthogonal projection on space $\mathbf{W}^p$ of functions with a support included in $[a_p, a_{p+1}]$. To avoid the discontinuities introduced by this projection, we introduce new orthogonal projectors that perform a smooth deformation of f.

Projectors on Half Lines

Let us first construct two orthogonal projectors that decompose any $f \in \mathbf{L^2}(\mathbb{R})$ in two orthogonal components P^+f and P^-f with supports that are, respectively, $[-1, +\infty)$ and $(-\infty, 1]$. For this purpose we consider a monotone increasing profile function β such that

$$\beta(t) = \begin{cases} 0 & \text{if } t < -1 \\ 1 & \text{if } t > 1 \end{cases} \tag{8.69}$$

and

$$\forall t \in [-1, 1], \quad \beta^2(t) + \beta^2(-t) = 1. \tag{8.70}$$

A naive definition

$$P^+f(t) = \beta^2(t) f(t) \quad \text{and} \quad P^-f(t) = \beta^2(-t) f(t)$$

satisfies the support conditions but does not define orthogonal functions. Since the supports of $P^+f(t)$ and $P^-f(t)$ overlap only on $[-1, 1]$, the orthogonality is obtained by creating functions having a different symmetry with respect to 0 on $[-1, 1]$:

$$P^+f(t) = \beta(t)\,[\beta(t) f(t) + \beta(-t) f(-t)] = \beta(t)\, p(t), \tag{8.71}$$

and

$$P^-f(t) = \beta(-t)\,[\beta(-t) f(t) - \beta(t) f(-t)] = \beta(-t)\, q(t). \tag{8.72}$$

The functions $p(t)$ and $q(t)$ are, respectively, even and odd, and since $\beta(t)\beta(-t)$ is even, it follows that

$$\langle P^+f, P^-f \rangle = \int_{-1}^{1} \beta(t)\, \beta(-t)\, p(t)\, q^*(t)\, dt = 0. \tag{8.73}$$

Clearly, P^+f belongs to space $\mathbf{W}^+$ of functions $f \in \mathbf{L^2}(\mathbb{R})$ such that there exists $p(t) = p(-t)$ with

$$f(t) = \begin{cases} 0 & \text{if } t < -1 \\ \beta(t)\, p(t) & \text{if } t \in [-1, 1]. \end{cases}$$

Similarly, P^-f is in space $\mathbf{W}^-$ composed of $f \in \mathbf{L^2}(\mathbb{R})$ such that there exists $q(t) = -q(-t)$ with

$$f(t) = \begin{cases} 0 & \text{if } t > 1 \\ \beta(-t)\, q(t) & \text{if } t \in [-1, 1]. \end{cases}$$

Functions in $\mathbf{W}^+$ and $\mathbf{W}^-$ may have an arbitrary behavior on $[1, +\infty)$ and $(-\infty, -1]$, respectively. Theorem 8.14 characterizes P^+ and P^-. We denote the identity operator by Id.

Theorem 8.14: *Coifman, Meyer.* Operators P^+ and P^- are orthogonal projectors on $\mathbf{W}^+$ and $\mathbf{W}^-$, respectively. Spaces $\mathbf{W}^+$ and $\mathbf{W}^-$ are orthogonal and

$$P^+ + P^- = \mathrm{Id}. \tag{8.74}$$

Proof. To verify that P^+ is a projector we show that any $f \in \mathbf{W}^+$ satisfies $P^+ f = f$. If $t < -1$, then $P^+ f(t) = f(t) = 0$, and if $t > 1$, then $P^+ f(t) = f(t) = 1$. If $t \in [-1, 1]$, then $f(t) = \beta(t)\, p_0(t)$ and inserting (8.71) yields

$$P^+ f(t) = \beta(t)\,[\beta^2(t)\, p_0(t) + \beta^2(-t)\, p_0(-t)] = \beta(t)\, p_0(t),$$

because $p_0(t)$ is even and $\beta(t)$ satisfies (8.70). Projector P^+ is proved to be orthogonal by showing that it is self-adjoint:

$$\langle P^+ f, g\rangle = \int_{-1}^{1} \beta^2(t) f(t)\, g^*(t)\, dt + \int_{-1}^{1} \beta(t)\, \beta(-t) f(-t)\, g^*(t)\, dt + \int_{1}^{+\infty} f(t)\, g^*(t)\, dt.$$

A change of variable $t' = -t$ in the second integral verifies that this formula is symmetric in f and g and thus $\langle P^+ f, g\rangle = \langle f, P^+ g\rangle$. Identical derivations prove that P^- is an orthogonal projector on $\mathbf{W}^-$.

The orthogonality of $\mathbf{W}^-$ and $\mathbf{W}^+$ is proved in (8.73). To verify (8.74), for $f \in \mathbf{L^2}(\mathbb{R})$ we compute

$$P^+ f(t) + P^- f(t) = f(t)\,[\beta^2(t) + \beta^2(-t)] = f(t). \qquad \blacksquare$$

These half-line projectors are generalized by decomposing signals in two orthogonal components with supports that are, respectively, $[a - \eta, +\infty)$ and $(-\infty, a + \eta]$. For this purpose, we scale and translate the profile function $\beta(\frac{t-a}{\eta})$ so that it increases from 0 to 1 on $[a - \eta, a + \eta]$, as illustrated in Figure 8.16. The symmetry with respect to 0, which transforms $f(t)$ in $f(-t)$, becomes a symmetry with respect to a, which transforms $f(t)$ in $f(2a - t)$. The resulting projectors are

$$P^+_{a,\eta} f(t) = \beta\Big(\frac{t-a}{\eta}\Big)\left[\beta\Big(\frac{t-a}{\eta}\Big) f(t) + \beta\Big(\frac{a-t}{\eta}\Big) f(2a-t)\right], \tag{8.75}$$

and

$$P^-_{a,\eta} f(t) = \beta\Big(\frac{a-t}{\eta}\Big)\left[\beta\Big(\frac{a-t}{\eta}\Big) f(t) - \beta\Big(\frac{t-a}{\eta}\Big) f(2a-t)\right]. \tag{8.76}$$

A straightforward extension of Theorem 8.14 proves that $P^+_{a,\eta}$ is an orthogonal projector on space $\mathbf{W}^+_{a,\eta}$ of functions $f \in \mathbf{L^2}(\mathbb{R})$ such that there exists $p(t) = p(2a - t)$ with

$$f(t) = \begin{cases} 0 & \text{if } t < a - \eta \\ \beta(\eta^{-1}(t-a))\, p(t) & \text{if } t \in [a-\eta, a+\eta]. \end{cases} \tag{8.77}$$

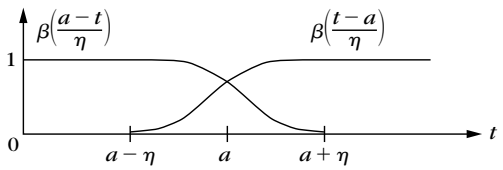

FIGURE 8.16

A multiplication with $\beta(\frac{t-a}{\eta})$ and $\beta(\frac{a-t}{\eta})$ restricts the support of functions to $[a-\eta, +\infty)$ and $(-\infty, a+\eta]$.

Similarly, $P^-_{a,\eta}$ is an orthogonal projector on space $\mathbf{W}^-_{a,\eta}$ composed of $f \in \mathbf{L^2}(\mathbb{R})$ such that there exists $q(t) = -q(2a-t)$ with

$$f(t) = \begin{cases} 0 & \text{if } t < -1 \\ \beta(\eta^{-1}(a-t))\, q(t) & \text{if } t \in [a-\eta, a+\eta]. \end{cases} \tag{8.78}$$

Spaces $\mathbf{W}^+_{a,\eta}$ and $\mathbf{W}^-_{a,\eta}$ are orthogonal and

$$P^+_{a,\eta} + P^-_{a,\eta} = \mathrm{Id}. \tag{8.79}$$

Projectors on Intervals

A lapped projector splits a signal in two orthogonal components that overlap on $[a-\eta, a+\eta]$. Repeating such projections at different locations performs a signal decomposition into orthogonal pieces with supports that overlap. Let us divide the time axis in overlapping intervals:

$$I_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1}]$$

with

$$\lim_{p \to -\infty} a_p = -\infty \quad \text{and} \quad \lim_{p \to +\infty} a_p = +\infty. \tag{8.80}$$

To ensure that I_{p-1} and I_{p+1} do not intersect for any $p \in \mathbb{Z}$, we impose that

$$a_{p+1} - \eta_{p+1} \geqslant a_p + \eta_p,$$

and thus,

$$l_p = a_{p+1} - a_p \geqslant \eta_{p+1} + \eta_p. \tag{8.81}$$

The support of f is restricted to I_p by the operator

$$P_p = P^+_{a_p,\eta_p} P^-_{a_{p+1},\eta_{p+1}}. \tag{8.82}$$

Since $P^+_{a_p,\eta_p}$ and $P^-_{a_{p+1},\eta_{p+1}}$ are orthogonal projections on $\mathbf{W}^+_{a_p,\eta_p}$ and $\mathbf{W}^-_{a_{p+1},\eta_{p+1}}$, it follows that P_p is an orthogonal projector on

$$\mathbf{W}^p = \mathbf{W}^+_{a_p,\eta_p} \cap \mathbf{W}^-_{a_{p+1},\eta_{p+1}}. \tag{8.83}$$

Let us divide I_p in two overlapping intervals O_p and O_{p+1}, and a central interval C_p:

$$I_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1}] = O_p \cup C_p \cup O_{p+1} \tag{8.84}$$

with

$$O_p = [a_p - \eta_p, a_p + \eta_p] \quad \text{and} \quad C_p = [a_p + \eta_p, a_{p+1} - \eta_{p+1}].$$

Space $\mathbf{W}^p$ is characterized by introducing a window g_p support in I_p, and that has a raising profile on O_p and a decaying profile on O_{p+1}:

$$g_p(t) = \begin{cases} 0 & \text{if } t \notin I_p \\ \beta(\eta_p^{-1}(t - a_p)) & \text{if } t \in O_p \\ 1 & \text{if } t \in C_p \\ \beta(\eta_{p+1}^{-1}(a_{p+1} - t)) & \text{if } t \in O_{p+1}. \end{cases} \tag{8.85}$$

This window is illustrated in Figure 8.17. It follows from (8.77), (8.78), and (8.83) that $\mathbf{W}^p$ is the space of functions $f \in \mathbf{L}^2(\mathbb{R})$ that can be written as

$$f(t) = g_p(t)\, h(t) \quad \text{with} \quad h(t) = \begin{cases} h(2a_p - t) & \text{if } t \in O_p \\ -h(2a_{p+1} - t) & \text{if } t \in O_{p+1}. \end{cases} \tag{8.86}$$

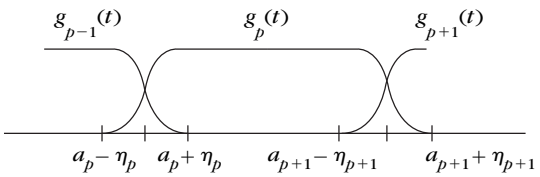

FIGURE 8.17

Each window g_p has a support $[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$ with an increasing profile and a decreasing profile over $[a_p - \eta_p, a_p + \eta_p]$ and $[a_{p+1} - \eta_{p+1}, a_{p+1} + \eta_{p+1}]$.

The function h is symmetric with respect to a_p and antisymmetric with respect to a_{p+1}, with an arbitrary behavior in C_p. Projector P_p on $\mathbf{W}^p$ defined in (8.82) can be rewritten as

$$P_p f(t) = \begin{cases} P^-_{a_p, \eta_p} f(t) & \text{if } t \in O_p \\ f(t) & \text{if } t \in C_p \\ P^+_{a_{p+1}, \eta_{p+1}} f(t) & \text{if } t \in O_{p+1} = g_p(t)\, h_p(t), \end{cases} \tag{8.87}$$

where $h_p(t)$ is calculated by inserting (8.75) and (8.76):

$$h_p(t) = \begin{cases} g_p(t) f(t) + g_p(2a_p - t) f(2a_p - t) & \text{if } t \in O_p \\ f(t) & \text{if } t \in C_p \\ g_p(t) f(t) - g_p(2a_{p+1} - t) f(2a_{p+1} - t) & \text{if } t \in O_{p+1}. \end{cases} \tag{8.88}$$

Theorem 8.15 derives a decomposition of the identity.

Theorem 8.15. Operator P_p is an orthogonal projector on $\mathbf{W}^p$. If $p \neq q$, then $\mathbf{W}^p$ is orthogonal to $\mathbf{W}^q$ and

$$\sum_{p=-\infty}^{+\infty} P_p = \mathrm{Id}. \tag{8.89}$$

Proof. If $p \neq q$ and $|p-q|>1$, then functions in $\mathbf{W}^p$ and $\mathbf{W}^q$ have supports that do not overlap, so these spaces are orthogonal. If $q = p+1$, then

$$\mathbf{W}^p = \mathbf{W}^+_{a_p,\eta_p} \cap \mathbf{W}^-_{a_{p+1},\eta_{p+1}} \quad \text{and} \quad \mathbf{W}^{p+1} = \mathbf{W}^+_{a_{p+1},\eta_{p+1}} \cap \mathbf{W}^-_{a_{p+2},\eta_{p+2}}.$$

Since $\mathbf{W}^-_{a_{p+1},\eta_{p+1}}$ is orthogonal to $\mathbf{W}^+_{a_{p+1},\eta_{p+1}}$, it follows that $\mathbf{W}^p$ is orthogonal to $\mathbf{W}^{p+1}$. To prove (8.89), we first verify that

$$P_p + P_{p+1} = P^+_{a_p,\eta_p} P^-_{a_{p+2},\eta_{p+2}}. \tag{8.90}$$

This is shown by decomposing P_p and P_{p+1} with (8.87) and inserting

$$P^+_{a_{p+1},\eta_{p+1}} + P^-_{a_{p+1},\eta_{p+1}} = \mathrm{Id}.$$

As a consequence,

$$\sum_{p=n}^{m} P_p = P^+_{a_n,\eta_n} P^-_{a_m,\eta_m}. \tag{8.91}$$

For any $f \in \mathbf{L^2}(\mathbb{R})$,

$$\|f - P^+_{a_n,\eta_n} P^-_{a_m,\eta_m} f\|^2 \geqslant \int_{-\infty}^{a_n+\eta_n} |f(t)|^2 dt + \int_{a_m-\eta_m}^{+\infty} |f(t)|^2 dt$$

and inserting (8.80) proves that

$$\lim_{\substack{n\to-\infty \\ m\to+\infty}} \|f - P^+_{a_n,\eta_n} P^-_{a_m,\eta_m} f\|^2 = 0.$$

The summation (8.91) implies (8.89). ■

Discretized Projectors

Projectors P_p that restrict the signal support to $[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$ are easily extended for discrete signals. Suppose that $\{a_p\}_{p\in\mathbb{Z}}$ are half integers, which means that $a_p + 1/2 \in \mathbb{Z}$. The windows $g_p(t)$ defined in (8.85) are uniformly sampled $g_p[n] = g_p(n)$. As in (8.86) we define the space $\mathbf{W}^p \subset \boldsymbol{\ell}^2(\mathbb{Z})$ of discrete signals

$$f[n] = g_p[n]\, h[n] \quad \text{with} \quad h[n] = \begin{cases} h[2a_p - n] & \text{if } n \in O_p \\ -h[2a_{p+1} - n] & \text{if } n \in O_{p+1}. \end{cases} \tag{8.92}$$

The orthogonal projector P_p on $\mathbf{W}^p$ is defined by an expression identical to (8.87) and (8.88):

$$P_p f[n] = g_p[n]\, h_p[n] \tag{8.93}$$

with

$$h_p[n] = \begin{cases} g_p[n]f[n] + g_p[2a_p - n]f[2a_p - n] & \text{if } n \in O_p \\ f[n] & \text{if } n \in C_p \\ g_p[n]f[n] - g_p[2a_{p+1} - n]f[2a_{p+1} - n] & \text{if } n \in O_{p+1}. \end{cases} \tag{8.94}$$

Finally, we prove as in Theorem 8.15 that if $p \neq q$, then $\mathbf{W}^p$ is orthogonal to $\mathbf{W}^q$ and

$$\sum_{p=-\infty}^{+\infty} P_p = \mathrm{Id}. \tag{8.95}$$

8.4.2 Lapped Orthogonal Bases

An orthogonal basis of $\mathbf{L}^2(\mathbb{R})$ is defined from a basis $\{e_k\}_{k\in\mathbb{N}}$ of $\mathbf{L}^2[0, 1]$ by multiplying a translation and dilation of each vector with a smooth window g_p defined in (8.85). A local cosine basis of $\mathbf{L}^2(\mathbb{R})$ is derived from a cosine IV basis of $\mathbf{L}^2[0, 1]$.

The support of g_p is $[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$ with $l_p = a_{p+1} - a_p$, as illustrated in Figure 8.17. The design of these windows also implies symmetry and quadrature properties on overlapping intervals:

$$g_p(t) = g_{p+1}(2a_{p+1} - t) \quad \text{for} \quad t \in [a_{p+1} - \eta_{p+1}, a_{p+1} + \eta_{p+1}], \tag{8.96}$$

and

$$g_p^2(t) + g_{p+1}^2(t) = 1 \quad \text{for} \quad t \in [a_{p+1} - \eta_{p+1}, a_{p+1} + \eta_{p+1}].$$

Each $e_k \in \mathbf{L}^2[0, 1]$ is extended over $\mathbb{R}$ into a function $\tilde{e}_k$ that is symmetric with respect to 0 and antisymmetric with respect to 1. The resulting $\tilde{e}_k$ has period 4 and is defined over $[-2, 2]$ by

$$\tilde{e}_k(t) = \begin{cases} e_k(t) & \text{if } t \in [0, 1] \\ e_k(-t) & \text{if } t \in (-1, 0) \\ -e_k(2 - t) & \text{if } t \in [1, 2) \\ -e_k(2 + t) & \text{if } t \in [-1, -2). \end{cases}$$

Theorem 8.16 derives an orthonormal basis of $\mathbf{L}^2(\mathbb{R})$.

Theorem 8.16: *Coifman, Malvar, Meyer.* Let $\{e_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathbf{L}^2[0, 1]$. The family

$$\left\{ g_{p,k}(t) = g_p(t) \frac{1}{\sqrt{l_p}} \tilde{e}_k\left(\frac{t - a_p}{l_p}\right) \right\}_{k\in\mathbb{N},\, p\in\mathbb{Z}} \tag{8.97}$$

is an orthonormal basis of $\mathbf{L}^2(\mathbb{R})$.

Proof. Since $\tilde{e}_k(l_p^{-1}(t - a_p))$ is symmetric with respect to a_p and antisymmetric with respect to a_{p+1}, it follows from (8.86) that $g_{p,k} \in \mathbf{W}^p$ for all $k \in \mathbb{N}$. Theorem 8.15 proves that

spaces $\mathbf{W}^p$ and $\mathbf{W}^q$ are orthogonal for $p \neq q$ and that $\mathbf{L^2}(\mathbb{R}) = \bigoplus_{p=-\infty}^{+\infty} \mathbf{W}^p$. To prove that (8.97) is an orthonormal basis of $\mathbf{L^2}(\mathbb{R})$, we thus need to show that

$$\left\{ g_{p,k}(t) = g_p(t) \frac{1}{\sqrt{l_p}} \tilde{e}_k \left(\frac{t - a_p}{l_p} \right) \right\}_{k \in \mathbb{N}, p \in \mathbb{Z}} \tag{8.98}$$

is an orthonormal basis of $\mathbf{W}^p$.

Let us prove first that any $f \in \mathbf{W}^p$ can be decomposed over this family. Such a function can be as written as $f(t) = g_p(t)\, h(t)$, where the restriction of h to $[a_p, a_{p+1}]$ is arbitrary, and h is, respectively, symmetric and antisymmetric with respect to a_p and a_{p+1}. Since $\{\tilde{e}_k\}_{k \in \mathbb{N}}$ is an orthonormal basis of $\mathbf{L^2}[0, 1]$, clearly

$$\left\{ \frac{1}{\sqrt{l_p}} \tilde{e}_k \Big(\frac{t - a_p}{l_p} \Big) \right\}_{k \in \mathbb{N}} \tag{8.99}$$

is an orthonormal basis of $\mathbf{L^2}[a_p, a_{p+1}]$. The restriction of h to $[a_p, a_{p+1}]$ can therefore be decomposed in this basis. This decomposition remains valid for all $t \in [a_p - \eta_p, a_{p+1} + \eta_{p+1}]$ since $h(t)$ and the $l_p^{-1/2} \tilde{e}_k(l_p^{-1}(t - a_p))$ have the same symmetry with respect to a_p and a_{p+1}. Therefore, $f(t) = h(t) g_p(t)$ can be decomposed over the family (8.98). Lemma 8.1 finishes the proof by showing that the orthogonality of functions in (8.98) is a consequence of the orthogonality of (8.99) in $\mathbf{L^2}[a_p, a_{p+1}]$.

Lemma 8.1. If $f_b(t) = h_b(t)\, g_p(t) \in \mathbf{W}^p$ and $f_c(t) = h_c(t)\, g_p(t) \in \mathbf{W}^p$, then

$$\langle f_b, f_c \rangle = \int_{a_p - \eta_p}^{a_{p+1} + \eta_{p+1}} f_b(t) f_c^*(t)\, dt = \int_{a_p}^{a_{p+1}} h_b(t)\, h_c^*(t)\, dt. \tag{8.100}$$

Let us evaluate

$$\langle f_b, f_c \rangle = \int_{a_p - \eta_p}^{a_{p+1} + \eta_{p+1}} h_b(t)\, h_c^*(t)\, g_p^2(t)\, dt. \tag{8.101}$$

We know that $h_b(t)$ and $h_c(t)$ are symmetric with respect to a_p, so

$$\int_{a_p - \eta_p}^{a_p + \eta_p} h_b(t)\, h_c^*(t)\, g_p^2(t)\, dt = \int_{a_p}^{a_p + \eta_p} h_b(t)\, h_c^*(t)\, [g_p^2(t) + g_p^2(2a_p - t)]\, dt.$$

Since $g_p^2(t) + g_p^2(2a_{p+1} - t) = 1$ over this interval, we obtain

$$\int_{a_p - \eta_p}^{a_p + \eta_p} h_b(t)\, h_c^*(t)\, g_p^2(t)\, dt = \int_{a_p}^{a_p + \eta_p} h_b(t)\, h_c(t)\, dt. \tag{8.102}$$

The functions $h_b(t)$ and $h_c(t)$ are antisymmetric with respect to a_{p+1}, so $h_b(t) h_c^*(t)$ is symmetric about a_{p+1}. Thus we prove similarly that

$$\int_{a_{p+1} - \eta_{p+1}}^{a_{p+1} + \eta_{p+1}} h_b(t)\, h_c^*(t)\, g_{p+1}^2(t)\, dt = \int_{a_{p+1} - \eta_{p+1}}^{a_{p+1}} h_b(t)\, h_c^*(t)\, dt. \tag{8.103}$$

Since $g_p(t) = 1$ for $t \in [a_p + \eta_p, a_{p+1} - \eta_{p+1}]$, inserting (8.102) and (8.103) in (8.101) proves the lemma property (8.100). ■

Theorem 8.16 is similar to the block basis theorem (8.8) but it has the advantage of using smooth windows g_p as opposed to the rectangular windows that are indicator functions of $[a_p, a_{p+1}]$. It yields smooth functions $g_{p,k}$ only if the extension $\tilde{e}_k$ of e_k is a smooth function. This is the case for the cosine IV basis $\{e_k(t) = \sqrt{2}\cos[(k + 1/2)\pi t]\}_{k\in\mathbb{N}}$ of $\mathbf{L^2}[0, 1]$ defined in Theorem 8.11. Indeed $\cos[(k + 1/2)\pi t]$ has a natural symmetric and antisymmetric extension with respect to 0 and 1 over $\mathbb{R}$. Corollary 8.1 derives a local cosine basis.

Corollary 8.1. The family of local cosine functions

$$\left\{ g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]\right\}_{k\in\mathbb{N},\, p\in\mathbb{Z}} \tag{8.104}$$

is an orthonormal basis of $\mathbf{L^2}(\mathbb{R})$.

Cosine–Sine I Basis

Other bases can be constructed with functions having a different symmetry. To maintain the orthogonality of the windowed basis, we must ensure that consecutive windows g_p and g_{p+1} are multiplied by functions that have an opposite symmetry with respect to a_{p+1}. For example, we can multiply g_{2p} with functions that are symmetric with respect to both ends a_{2p} and a_{2p+1}, and multiply g_{2p+1} with functions that are antisymmetric with respect to a_{2p+1} and a_{2p+2}. Such bases can be constructed with the cosine I basis $\{\sqrt{2}\lambda_k \cos(\pi k t)\}_{k\in\mathbb{Z}}$ defined in Theorem 8.10, with $\lambda_0 = 2^{-1/2}$ and $\lambda_k = 1$ for $k \neq 0$, and with the sine I family $\{\sqrt{2}\sin(\pi k t)\}_{k\in\mathbb{N}^*}$, which is also an orthonormal basis of $\mathbf{L^2}[0, 1]$. The reader can verify that if

$$g_{2p,k}(t) = g_{2p}(t)\sqrt{\frac{2}{l_{2p}}}\lambda_k \cos\left[\pi k\,\frac{t-a_{2p}}{l_{2p}}\right]$$

$$g_{2p+1,k}(t) = g_{2p+1}(t)\sqrt{\frac{2}{l_{2p+1}}}\sin\left[\pi k\,\frac{t-a_{2p+1}}{l_{2p+1}}\right],$$

then $\{g_{p,k}\}_{k\in\mathbb{N},p\in\mathbb{Z}}$ is an orthonormal basis of $\mathbf{L^2}(\mathbb{R})$.

Lapped Transforms in Frequency

Lapped orthogonal projectors can also divide the frequency axis in separate overlapping intervals. This is done by decomposing the Fourier transform $\hat{f}(\omega)$ of $f(t)$ over a local cosine basis defined on the frequency axis $\{g_{p,k}(\omega)\}_{p\in\mathbb{Z},k\in\mathbb{N}}$. This is also equivalent to decomposing $f(t)$ on its inverse Fourier transform $\{\frac{1}{2\pi}\hat{g}_{p,k}(-t)\}_{p\in\mathbb{Z},k\in\mathbb{N}}$. As opposed to wavelet packets, which decompose signals in dyadic-frequency bands, this approach offers complete flexibility on the size of the frequency intervals $[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$.

A signal decomposition in a Meyer wavelet or wavelet packet basis can be calculated with a lapped orthogonal transform applied in the Fourier domain. Indeed, the Fourier transform (7.87) of a Meyer wavelet has a compact support and

$\{|\hat{\psi}(2^j\omega)|\}_{j\in\mathbb{Z}}$ can be considered as a family asymmetric window, with support that only overlaps with adjacent windows with appropriate symmetry properties. These windows cover the whole frequency axis: $\sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j\omega)|^2 = 1$. As a result, the Meyer wavelet transform can be viewed as a lapped orthogonal transform applied in the Fourier domain. Thus, it can be efficiently implemented with the folding algorithm from Section 8.4.4.

8.4.3 Local Cosine Bases

The local cosine basis defined in (8.104) is composed of functions

$$g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right]$$

with a compact support $[a_p-\eta_p, a_{p+1}+\eta_{p+1}]$. The energy of their Fourier transforms is also well concentrated. Let $\hat{g}_p$ be the Fourier transform of g_p,

$$\hat{g}_{p,k}(\omega) = \frac{\exp(-ia_p\xi_{p,k})}{2}\sqrt{\frac{2}{l_p}}\left(\hat{g}_p(\omega-\xi_{p,k})+\hat{g}_p(\omega+\xi_{p,k})\right),$$

where

$$\xi_{p,k} = \frac{\pi(k+1/2)}{l_p}.$$

The bandwidth of $\hat{g}_{p,k}$ around $\xi_{p,k}$ and $-\xi_{p,k}$ is equal to the bandwidth of $\hat{g}_p$. If sizes η_p and η_{p+1} of the variation intervals of g_p are proportional to l_p, then this bandwidth is proportional to l_p^{-1}.

For smooth functions f, we want to guarantee that the inner products $\langle f, g_{p,k}\rangle$ have a fast decay when the center frequency $\xi_{p,k}$ increases. The Parseval formula proves that

$$\langle f, g_{p,k}\rangle = \frac{\exp(ia_p\xi_{p,k})}{2\pi}\sqrt{\frac{2}{l_p}}\int_{-\infty}^{+\infty}\hat{f}(\omega)\left(\hat{g}_p^*(\omega-\xi_{p,k})+\hat{g}_p^*(\omega+\xi_{p,k})\right)d\omega.$$

The smoothness of f implies that $|\hat{f}(\omega)|$ has a fast decay at large frequencies ω. Therefore, this integral will become small when $\xi_{p,k}$ increases if g_p is a smooth window, because $|\hat{g}_p(\omega)|$ has a fast decay.

Window Design

The regularity of g_p depends on the regularity of profile β, as shown by (8.85). This profile must satisfy

$$\beta^2(t)+\beta^2(-t) = 1 \quad \text{for} \quad t\in[-1,1], \tag{8.105}$$

plus $\beta(t)=0$ if $t<-1$ and $\beta(t)=1$ if $t>1$. One example is

$$\beta_0(t) = \sin\left(\frac{\pi}{4}(1+t)\right) \quad \text{for} \quad t\in[-1,1],$$

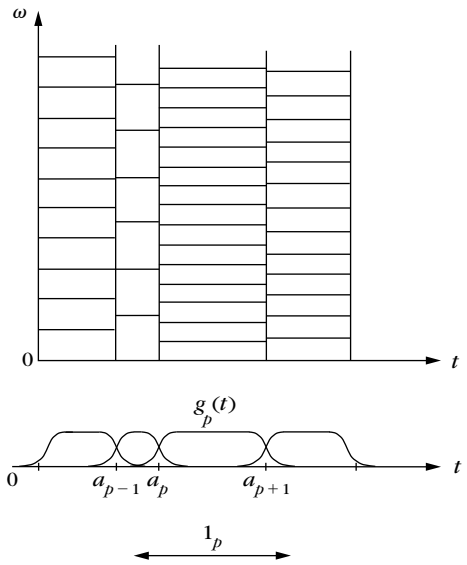

FIGURE 8.18

Heisenberg boxes of local cosine vectors define a regular grid over the time-frequency plane.

but its derivative at $t = \pm 1$ is nonzero, so β is not differentiable at ± 1. Windows of higher regularity are constructed with a profile β_k defined by induction for $k \geq 0$ by

$$\beta_{k+1}(t) = \beta_k\left(\sin \frac{\pi t}{2}\right) \quad \text{for} \quad t \in [-1, 1].$$

For any $k \geq 0$, one can verify that β_k satisfies (8.105) and has $2^k - 1$ vanishing derivatives at $t = \pm 1$. The resulting β and g_p are therefore $2^k - 1$ times continuously differentiable.

Heisenberg Box

A local cosine basis can be symbolically represented as an exact paving of the time-frequency plane. The time and frequency region of high-energy concentration for each local cosine vector $g_{p,k}$ is approximated by a Heisenberg rectangle

$$[a_p, a_{p+1}] \times \left[\xi_{p,k} - \frac{\pi}{2l_p}, \xi_{p,k} + \frac{\pi}{2l_p}\right],$$

as illustrated in Figure 8.18. A local cosine basis $\{g_{p,k}\}_{k \in \mathbb{N}, p \in \mathbb{Z}}$ corresponds to a time-frequency grid which varies in time.

Figure 8.19(a) shows the decomposition of a digital recording of the sound "grea" coming from the word "greasy." The window sizes are adapted to the signal structures with the best basis algorithm described in Section 12.2.2. High-amplitude

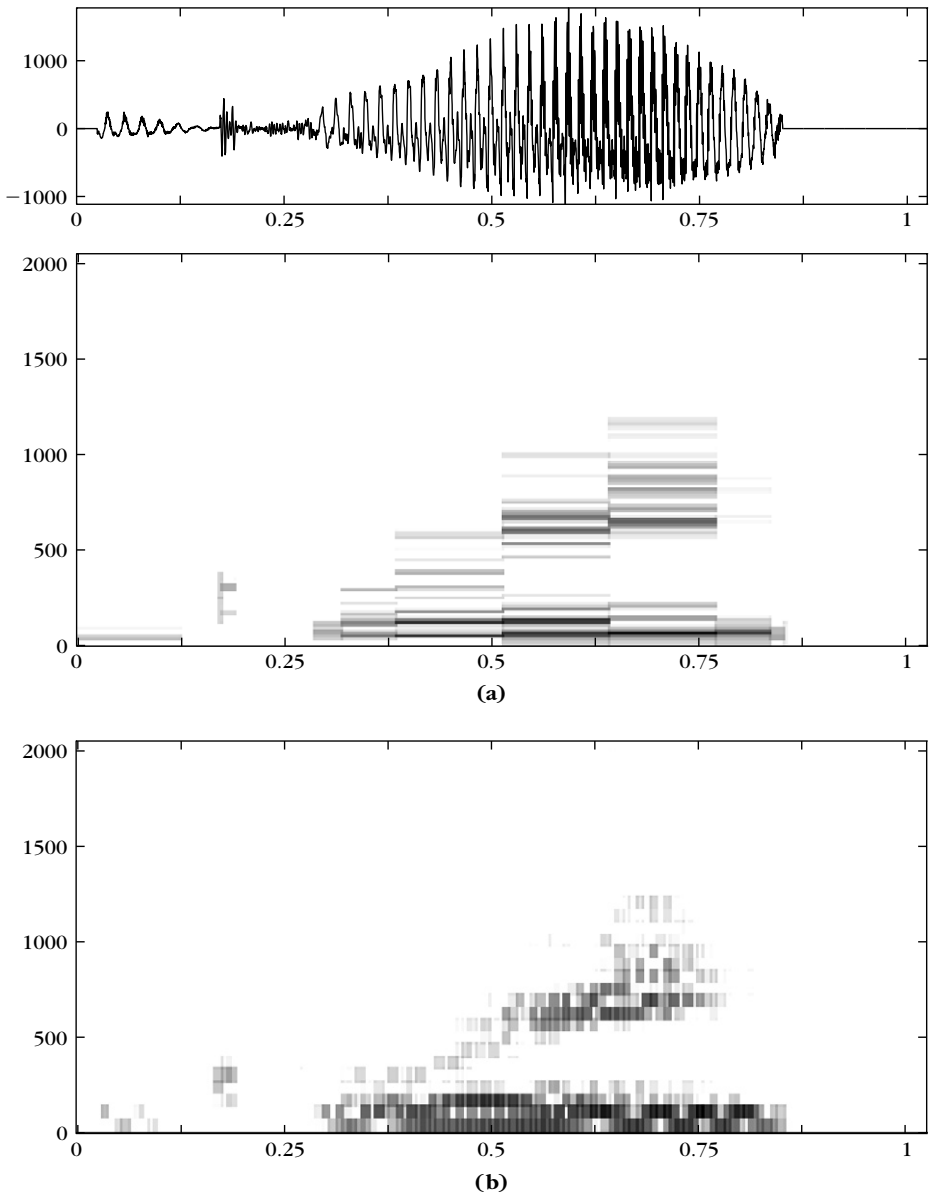

FIGURE 8.19

(a) The signal at the top is a recording of the sound "grea" in the word "greasy." This signal is decomposed in a local cosine basis with windows of varying sizes. The larger the amplitude of $|\langle f, g_{p,k}\rangle|$, the darker the gray level of the Heisenberg box. **(b)** Decomposition in a local cosine basis with small windows of constant size.

coefficients are along spectral lines in the time-frequency plane. Most Heisenberg boxes appear in white, which indicates that the corresponding inner product is nearly zero. Thus, this signal can be approximated with a few nonzero local cosine vectors. Figure 8.19(b) decomposes the same signal in a local cosine basis composed of small windows of constant size. The signal time-frequency structures do not appear as well as in Figure 8.19(a).

Translation and Phase

Cosine modulations as opposed to complex exponentials do not provide easy access to phase information. The translation of a signal can induce important modifications of its decomposition coefficients in a cosine basis. Consider, for example,

$$f(t) = g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}}\cos\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_p}\right].$$

Since the basis is orthogonal, $\langle f, g_{p,k}\rangle = 1$ and all other inner products are zero. After a translation by $\tau = l_p/(2k+1)$,

$$f_\tau(t) = f\left(t-\frac{l_p}{2k+1}\right) = g_p(t)\sqrt{\frac{2}{l_p}}\sin\left[\pi\left(k+\frac{1}{2}\right)\frac{t-a_p}{l_k}\right].$$

The opposite parity of sine and cosine implies that $\langle f_\tau, g_{p,k}\rangle \approx 0$. In contrast, $\langle f_\tau, g_{p,k-1}\rangle$ and $\langle f_\tau, g_{p,k+1}\rangle$ become nonzero. After translation, a signal component initially represented by a cosine of frequency $\pi(k+1/2)/l_p$ is therefore spread over cosine vectors of different frequencies.

This example shows that the local cosine coefficients of a pattern are severely modified by any translation. We are facing the same translation distortions as observed in Section 5.1.5 for wavelets and time-frequency frames. This lack of translation invariance makes it difficult to use these bases for pattern recognition.

8.4.4 Discrete Lapped Transforms

Lapped orthogonal bases are discretized by replacing the orthogonal basis of $\mathbf{L}^2[0, 1]$ with a discrete basis of $\mathbb{C}^N$, and uniformly sampling the windows g_p. Discrete local cosine bases are derived with discrete cosine IV bases.

Let $\{a_p\}_{p\in\mathbb{Z}}$ be a sequence of half integers $a_p + 1/2 \in \mathbb{Z}$ with

$$\lim_{p\to-\infty} a_p = -\infty \quad \text{and} \quad \lim_{p\to+\infty} a_p = +\infty.$$

A discrete lapped orthogonal basis is constructed with the discrete projectors P_p defined in (8.93). These operators are implemented with the sampled windows $g_p[n] = g_p(n)$. Suppose that $\{e_{k,l}[n]\}_{0\leq k<l}$ is an orthogonal basis of signals defined for $0 \leq n < l$. These vectors are extended over $\mathbb{Z}$ with a symmetry with respect to $-1/2$ and an antisymmetry with respect to $l - 1/2$. The resulting extensions have a

period $4l$ and are defined over $[-2l, 2l-1]$ by

$$\tilde{e}_{l,k}[n] = \begin{cases} e_{l,k}[n] & \text{if } n \in [0, l-1] \\ e_{l,k}[-1-n] & \text{if } n \in [-l, -1] \\ -e_k[2l-1-n] & \text{if } n \in [l, 2l-1] \\ -e_k[2l+n] & \text{if } n \in [-2l, -l-1]. \end{cases}$$

Theorem 8.17 proves that multiplying these vectors with the discrete windows $g_p[n]$ yields an orthonormal basis of $\ell^2(\mathbb{Z})$.

Theorem 8.17: *Coifman, Malvar, Meyer.* Suppose that $\{e_{k,l}\}_{0 \leq k < l}$ is an orthogonal basis of $\mathbb{C}^l$ for any $l > 0$. The family

$$\left\{ g_{p,k}[n] = g_p[n]\, \tilde{e}_{k,l_p}[n - a_p] \right\}_{0 \leq k < l_p, p \in \mathbb{Z}} \tag{8.106}$$

is a lapped orthonormal basis of $\ell^2(\mathbb{Z})$.

The proof of this theorem is identical to the proof of Theorem 8.16 since we have a discrete equivalent of the spaces $\mathbf{W}^p$ and their projectors. It is also based on a discrete equivalent of Lemma 8.1, which is verified with the same derivations. Beyond the proof of Theorem 8.17, we shall see that Lemma 8.2 is important for quickly computing the decomposition coefficients $\langle f, g_{p,k} \rangle$. ■

Lemma 8.2. Any $f_b[n] = g_p[n]\, h_b[n] \in \mathbf{W}^p$ and $f_c[n] = g_p[n]\, h_c[n] \in \mathbf{W}^p$ satisfy

$$\langle f_b, f_c \rangle = \sum_{a_p - \eta_p < n < a_{p+1} + \eta_{p+1}} f_b[n] f_c^*[n] = \sum_{a_p < n < a_{p+1}} h_b[n]\, h_c^*[n]. \tag{8.107}$$

Theorem 8.17 is similar to the discrete block basis theorem (8.9) but constructs an orthogonal basis with smooth discrete windows $g_p[n]$. The discrete cosine IV bases

$$\left\{ e_{l,k}[n] = \sqrt{\frac{2}{l}} \cos\left[\frac{\pi}{l} \left(k + \frac{1}{2} \right) \left(n + \frac{1}{2} \right) \right] \right\}_{0 \leq k < l}$$

have the advantage of including vectors that have a natural symmetric and antisymmetric extension with respect to $-1/2$ and $l - 1/2$. This produces a discrete local cosine basis of $\ell^2(\mathbb{Z})$.

Corollary 8.2. The family

$$\left\{ g_{p,k}[n] = g_p[n] \sqrt{\frac{2}{l_p}} \cos\left[\pi \left(k + \frac{1}{2} \right) \frac{n - a_p}{l_p} \right] \right\}_{0 \leq k < l_p, p \in \mathbb{Z}} \tag{8.108}$$

is an orthonormal basis of $\ell^2(\mathbb{Z})$.

Fast Lapped Orthogonal Transform

A fast algorithm introduced by Malvar [40] replaces the calculations of $\langle f, g_{p,k}\rangle$ by a computation of inner products in the original bases $\{e_{l,k}\}_{0\leq k<l}$ with a folding procedure. In a discrete local cosine basis, these inner products are calculated with the fast DCT-IV algorithm.

To simplify notations, as in Section 8.4.1 we decompose $I_p = [a_p - \eta_p, a_{p+1} + \eta_{p+1}]$ into $I_p = O_p \cup C_p \cup O_{p+1}$ with

$$O_p = [a_p - \eta_p, a_p + \eta_p] \quad \text{and} \quad C_p = [a_p + \eta_p, a_{p+1} - \eta_{p+1}].$$

The orthogonal projector P_p on space $\mathbf{W}^p$ generated by $\{g_{p,k}\}_{0\leq k<l_p}$ was calculated in (8.93):

$$P_p f[n] = g_p[n]\, h_p[n],$$

where h_p is a folded version of f:

$$h_p[n] = \begin{cases} g_p[n] f[n] + g_p[2a_p - n] f[2a_p - n] & \text{if } n \in O_p \\ f[n] & \text{if } n \in C_p \\ g_p[n] f[n] - g_p[2a_{p+1} - n] f[2a_{p+1} - n] & \text{if } n \in O_{p+1}. \end{cases} \tag{8.109}$$

Since $g_{p,k} \in \mathbf{W}^p$,

$$\langle f, g_{p,k}\rangle = \langle P_p f, g_{p,k}\rangle = \langle g_p h_p,\, g_p \tilde{e}_{l_p,k}\rangle.$$

Since $\tilde{e}_{l_p,k}[n] = e_{l_p,k}[n]$ for $n \in [a_p, a_{p+1}]$, Lemma 8.2 derives that

$$\langle f, g_{p,k}\rangle = \sum_{a_p < n < a_{p+1}} h_p[n]\, e_{l_p,k}[n] = \langle h_p, e_{l_p,k}\rangle_{[a_p, a_{p+1}]}. \tag{8.110}$$

This proves that the decomposition coefficients $\langle f, g_{p,k}\rangle$ can be calculated by folding f into h_p and computing the inner product with the orthogonal basis $\{e_{l_p,k}\}_{0\leq k<l_p}$ defined over $[a_p, a_{p+1}]$.

For a discrete cosine basis, the DCT-IV coefficients

$$\langle h_p, e_{l_p,k}\rangle_{[a_p,a_{p+1}]} = \sum_{a_p < n < a_{p+1}} h_p[n] \sqrt{\frac{2}{l_p}} \cos\left[\pi\left(k + \frac{1}{2}\right)\frac{n - a_p}{l_p}\right] \tag{8.111}$$

are computed with the fast DCT-IV algorithm from Section 8.3.4, which requires $O(l_p \log_2 l_p)$ operations. The inverse lapped transform recovers $h_p[n]$ over $[a_p, a_{p+1}]$ from the l_p inner products $\{\langle h_p, e_{l_p,k}\rangle_{[a_p,a_{p+1}]}\}_{0\leq k<l_p}$. In a local cosine IV basis, this is done with the fast inverse DCT-IV, which is identical to the forward DCT-IV and requires $O(l_p \log_2 l_p)$ operations. The reconstruction of f is done by applying (8.95), which proves that

$$f[n] = \sum_{p=-\infty}^{+\infty} P_p f[n] = \sum_{p=-\infty}^{+\infty} g_p[n]\, h_p[n]. \tag{8.112}$$

Let us denote $O_p^- = [a_p - \eta_p, a_p]$ and $O_p^+ = [a_p, a_p + \eta_p]$. The restriction of (8.112) to $[a_p, a_{p+1}]$ gives

$$f[n] = \begin{cases} g_p[n]\, h_p[n] + g_{p-1}[n]\, h_{p-1}[n] & \text{if } n \in O_p^+ \\ h_p[n] & \text{if } n \in C_p \\ g_p[n]\, h_p[n] + g_{p+1}[n]\, h_{p+1}[n] & \text{if } n \in O_{p+1}^-. \end{cases}$$

The symmetry of the windows guarantees that $g_{p-1}[n] = g_p[2a_p - n]$ and $g_{p+1}[n] = g_p[2a_{p+1} - n]$. Since $h_{p-1}[n]$ is antisymmetric with respect to a_p and $h_{p+1}[n]$ is symmetric with respect to a_{p+1}, we can recover $f[n]$ on $[a_p, a_{p+1}]$ from the values of $h_{p-1}[n]$, $h_p[n]$, and $h_{p+1}[n]$ computed, respectively, on $[a_{p-1}, a_p]$, $[a_p, a_{p+1}]$, and $[a_{p+1}, a_{p+2}]$:

$$f[n] = \begin{cases} g_p[n]\, h_p[n] - g_p[2a_p - n]\, h_{p-1}[2a_p - n] & \text{if } n \in O_p^+ \\ h_p[n] & \text{if } n \in C_p \\ g_p[n]\, h_p[n] + g_p[2a_{p+1} - n]\, h_{p+1}[2a_{p+1} - n] & \text{if } n \in O_{p+1}^-. \end{cases} \tag{8.113}$$

This unfolding formula is implemented with $O(l_p)$ calculations. Thus, the inverse local cosine transform requires $O(l_p \log_2 l_p)$ operations to recover $f[n]$ on each interval $[a_p, a_{p+1}]$ of length l_p.

Finite Signals

If $f[n]$ is defined for $0 \leqslant n < N$, the extremities of the first and last interval must be $a_0 = -1/2$ and $a_q = N - 1/2$. A fast local cosine algorithm needs $O(l_p \log_2 l_p)$ additions and multiplications to decompose or reconstruct the signal on each interval of length l_p. On the whole signal of length N, it thus needs a total of $O(N \log_2 L)$ operations, where $L = \sup_{0 \leqslant p < q} l_p$.

Since we do not know the values of $f[n]$ for $n < 0$, at the left border we set $\eta_0 = 0$. This means that $g_0[n]$ jumps from 0 to 1 at $n = 0$. The resulting transform on the left boundary is equivalent to a straight DCT-IV. Section 8.3.2 shows that since cosine IV vectors are even on the left boundary, the DCT-IV is equivalent to a symmetric signal extension followed by a discrete Fourier transform. This avoids creating discontinuity artifacts at the left border.

At the right border, we also set $\eta_q = 0$ to limit the support of g_{q-1} to $[0, N-1]$. Section 8.4.4 explains that since cosine IV vectors are odd on the right boundary, the DCT-IV is equivalent to an antisymmetric signal extension. If $f[N-1] \neq 0$, this extension introduces a sharp signal transition that creates artificial high frequencies. To reduce this border effect, we replace the cosine IV modulation

$$g_{q-1,k}[n] = g_{q-1}[n] \sqrt{\frac{2}{l_{q-1}}} \cos\left[\pi\left(k + \frac{1}{2}\right)\frac{n - a_{q-1}}{l_{q-1}}\right]$$

with a cosine I modulation,

$$g_{q-1,k}[n] = g_{q-1}[n] \sqrt{\frac{2}{l_{q-1}}}\, \lambda_k \cos\left[\pi k\, \frac{n - a_{q-1}}{l_{q-1}}\right].$$

The orthogonality with the other elements of the basis is maintained because these cosine I vectors, like cosine IV vectors, are even with respect to a_{q-1}. Since $\cos\left[\pi k(n-a_{q-1})/l_{q-1}\right]$ is also symmetric with respect to $a_q = N - 1/2$, computing a DCT-I is equivalent to performing a symmetric signal extension at the right boundary, which avoids discontinuities. In the fast local cosine transform, we thus compute a DCT-I of the last folded signal h_{q-1} instead of a DCT-IV. The reconstruction algorithm uses an inverse DCT-I to recover h_{q-1} from these coefficients.

8.5 LOCAL COSINE TREES

Corollary 8.1 constructs local cosine bases for any segmentation of the time axis into intervals $[a_p, a_{p+1}]$ of arbitrary lengths. This result is more general than the construction of wavelet packet bases that can only divide the frequency axis into dyadic intervals with a length proportional to a power of 2. However, Coifman and Meyer [181] showed that restricting the intervals to dyadic sizes has the advantage of creating a tree structure similar to a wavelet packet tree. "Best" local cosine bases can then be adaptively chosen with the fast dynamical programming algorithm, described in Section 12.2.2.

8.5.1 Binary Tree of Cosine Bases

A local cosine tree includes orthogonal bases that segment the time axis in dyadic intervals. For any $j \geqslant 0$, the interval [0, 1] is divided in 2^j intervals of length 2^{-j} by setting

$$a_{p,j} = p\,2^{-j} \quad \text{for} \quad 0 \leqslant p \leqslant 2^j.$$

These intervals are covered by windows $g_{p,j}$ defined by (8.85) with a support $[a_{p,j} - \eta, a_{p+1,j} + \eta]$:

$$g_{p,j}(t) = \begin{cases} \beta(\eta^{-1}(t - a_{p,j})) & \text{if } t \in [a_{p,j} - \eta, a_{p,j} + \eta] \\ 1 & \text{if } t \in [a_{p,j} + \eta, a_{p+1,j} - \eta] \\ \beta(\eta^{-1}(a_{p+1,j} - t)) & \text{if } t \in [a_{p+1,j} - \eta, a_{p+1,j} + \eta] \\ 0 & \text{otherwise.} \end{cases} \tag{8.114}$$

To ensure that the support of $g_{p,j}$ is in [0, 1] for $p = 0$ and $p = 2^j - 1$, we modify, respectively, the left and right sides of these windows by setting $g_{0,j}(t) = 1$ if $t \in [0, \eta]$, and $g_{2^j-1,j}(t) = 1$ if $t \in [1 - \eta, 1]$. It follows that $g_{0,0} = \mathbf{1}_{[0,1]}$. The size η of the raising and decaying profiles of $g_{p,j}$ is independent of j. To guarantee that windows overlap only with their two neighbors, length $a_{p+1,j} - a_{p,j} = 2^{-j}$ must be larger than size 2η of the overlapping intervals, and thus

$$\eta \leqslant 2^{-j-1}. \tag{8.115}$$

Similar to wavelet packet trees, a local cosine tree is constructed by recursively dividing spaces built with local cosine bases. A tree node at depth j and position p

is associated to space $\mathbf{W}_j^p$ generated by the local cosine family

$$\mathcal{B}_j^p = \left\{ g_{p,j}(t) \sqrt{\frac{2}{2^{-j}}} \cos\left[\pi\left(k+\frac{1}{2}\right) \frac{t-a_{p,j}}{2^{-j}} \right] \right\}_{k\in\mathbb{Z}}. \tag{8.116}$$

Any $f \in \mathbf{W}_p^j$ has a support in $[a_{p,j}-\eta, a_{p+1,j}+\eta]$ and can be written as $f(t) = g_{p,j}(t)\,h(t)$, where $h(t)$ is symmetric and antisymmetric, respectively, to $a_{p,j}$ and $a_{p+1,j}$. Theorem 8.18 shows that $\mathbf{W}_j^p$ is divided in two orthogonal spaces $\mathbf{W}_{j+1}^{2p}$ and $\mathbf{W}_{j+1}^{2p+1}$ that are built over the two half intervals.

Theorem 8.18: *Coifman, Meyer.* For any $j \geqslant 0$ and $p < 2^j$, spaces $\mathbf{W}_{j+1}^{2p}$ and $\mathbf{W}_{j+1}^{2p+1}$ are orthogonal and

$$\mathbf{W}_j^p = \mathbf{W}_{j+1}^{2p} \oplus \mathbf{W}_{j+1}^{2p+1}. \tag{8.117}$$

Proof. The orthogonality of $\mathbf{W}_{j+1}^{2p}$ and $\mathbf{W}_{j+1}^{2p+1}$ is proved by Theorem 8.15. We denote $P_{p,j}$ as the orthogonal projector on $\mathbf{W}_j^p$. With the notation of Section 8.4.1, this projector is decomposed into two splitting projectors at $a_{p,j}$ and $a_{p+1,j}$:

$$P_{p,j} = P^+_{a_{p,j},\eta} P^-_{a_{p+1,j},\eta}.$$

Equation (8.90) proves that

$$P_{2p,j+1} + P_{2p+1,j+1} = P^+_{a_{2p,j+1},\eta} P^-_{a_{2p+2,j+1},\eta} = P^+_{a_{p,j},\eta} P^-_{a_{p+1,j},\eta} = P_{p,j}.$$

This equality on orthogonal projectors implies (8.117). ■

Space $\mathbf{W}_j^p$ located at node (j,p) of a local cosine tree is therefore the sum of the two spaces $\mathbf{W}_{j+1}^{2p}$ and $\mathbf{W}_{j+1}^{2p+1}$ located at the children nodes. Since $g_{0,0} = \mathbf{1}_{[0,1]}$ it follows that $\mathbf{W}_0^0 = \mathbf{L^2}[0,1]$. The maximum depth J of the binary tree is limited by the support condition $\eta \leqslant 2^{-J-1}$, and thus

$$J \leqslant -\log_2(2\eta). \tag{8.118}$$

Admissible Local Cosine Bases

As in a wavelet packet binary tree, many local cosine orthogonal bases are constructed from this local cosine tree. We call any subtree of the local cosine tree with nodes that have either zero or two children *an admissible binary tree*. Let $\{j_i, p_i\}_{1\leqslant i\leqslant I}$ be the indices at the leaves of a particular admissible binary tree. Applying the splitting property (8.117) along the branches of this subtree proves that

$$\mathbf{L^2}[0,1] = \mathbf{W}_0^0 = \bigoplus_{i=1}^{I} \mathbf{W}_{j_i}^{p_i}.$$

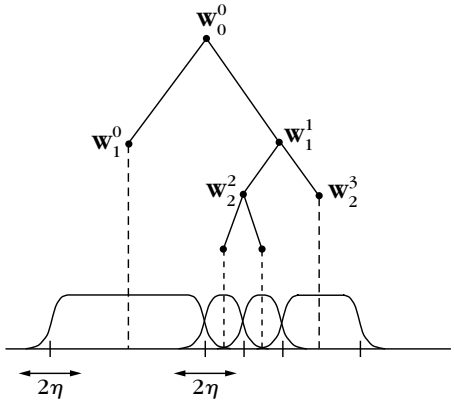

FIGURE 8.20

An admissible binary tree of local cosine spaces divides the time axis in windows of dyadic lengths.

Thus, the union of local cosine bases $\bigcup_{i=1}^{I} \mathcal{B}_{j_i}^{p_i}$ is an orthogonal basis of $\mathbf{L^2}[0, 1]$. This can also be interpreted as a division of the time axis into windows of various length, as illustrated by Figure 8.20.

The number B_J of different dyadic local cosine bases is equal to the number of different admissible subtrees of depth of at most J. For $J = -\log_2(2\eta)$,Theorem 8.2 proves that

$$2^{1/(4\eta)} \leqslant B_J \leqslant 2^{3/(8\eta)}.$$

Figure 8.19 on page 421 shows the decomposition of a sound recording in two dyadic local cosine bases selected from the binary tree. The basis in (a) is calculated with the best basis algorithm of Section 12.2.2.

Choice of η

At all scales 2^j, windows $g_{p,j}$ of a local cosine tree have raising and decaying profiles of the same size η. Thus, these windows can be recombined independently from their scale. If η is small compared to the interval size 2^{-j}, then $g_{p,j}$ has a relatively sharp variation at its borders compared to the size of its support. Since η is not proportional to 2^{-j}, the energy concentration of $\hat{g}_{p,j}$ is not improved when the window size 2^{-j} increases. Even though f may be very smooth over $[a_{p,j}, a_{p+1,j}]$, the border variations of the window create relatively large coefficients up to a frequency of the order of π/η.

To reduce the number of large coefficients we must increase η, but this also increases the minimum window size in the tree, which is $2^{-J} = 2\eta$. The choice of η is therefore the result of a trade-off between window regularity and the maximum resolution of the time subdivision. There is no equivalent limitation in the construction of wavelet packet bases.

8.5.2 Tree of Discrete Bases

For discrete signals of size N, a binary tree of discrete cosine bases is constructed like a binary tree of continuous-time cosine bases. To simplify notations, the sampling distance is normalized to 1. If it is equal to N^{-1}, then frequency parameters must be multiplied by N.

The subdivision points are located at half integers:

$$a_{p,j} = p\,N\,2^{-j} - 1/2 \quad \text{for} \quad 0 \leqslant p \leqslant 2^j.$$

The discrete windows are obtained by sampling the windows $g_p(t)$ defined in (8.114), $g_{p,j}[n] = g_{p,j}(n)$. The same border modification is used to ensure that the support of all $g_{p,j}[n]$ is in $[0, N-1]$.

A node at depth j and position p in the binary tree corresponds to space $\mathbf{W}_j^p$ generated by the discrete local cosine family

$$\mathcal{B}_j^p = \left\{ g_{p,j}[n] \sqrt{\frac{2}{2^{-j}N}} \cos\left[\pi \left(k + \frac{1}{2}\right) \frac{n - a_{p,j}}{2^{-j}N} \right] \right\}_{0 \leqslant k < N 2^{-j}}.$$

Since $g_{0,0} = \mathbf{1}_{[0,N-1]}$, the space $\mathbf{W}_0^0$ at the root of the tree includes any signal defined over $0 \leqslant n < N$, so $\mathbf{W}_0^0 = \mathbb{C}^N$. As in Theorem 8.18 we verify that $\mathbf{W}_j^p$ is orthogonal to $\mathbf{W}_j^q$ for $p \neq q$ and that

$$\mathbf{W}_j^p = \mathbf{W}_{j+1}^{2p} \oplus \mathbf{W}_{j+1}^{2p+1}. \tag{8.119}$$

The splitting property (8.119) implies that the union of local cosine families $\mathcal{B}_j^p$ located at the leaves of an admissible subtree is an orthogonal basis of $\mathbf{W}_0^0 = \mathbb{C}^N$. The minimum window size is limited by $2\eta \leqslant 2^{-j}N$, so the maximum depth of this binary tree is $J = \log_2 \frac{N}{2\eta}$. Thus, one can construct more than $2^{2^{J-1}} = 2^{N/(4\eta)}$ different discrete local cosine bases within this binary tree.

Fast Calculations

The fast local cosine transform algorithm described in Section 8.4.4 requires $O(2^{-j}N \log_2(2^{-j}N))$ operations to compute the inner products of f with the $2^{-j}N$ vectors in the local cosine family $\mathcal{B}_j^p$. The total number of operations to perform these computations at all nodes (j,p) of the tree, for $0 \leqslant p < 2^j$ and $0 \leqslant j \leqslant J$, is therefore $O(N J \log_2 N)$. The local cosine decompositions in Figure 8.19 are calculated with this fast algorithm. To improve the right border treatment, Section 8.4.4 explains that the last DCT-IV should be replaced by a DCT-I at each scale 2^j. The signal f is recovered from the local cosine coefficients at the leaves of any admissible binary tree with the fast local cosine reconstruction algorithm, which needs $O(N \log_2 N)$ operations.

8.5.3 Image Cosine Quad-Tree

A local cosine binary tree is extended in two dimensions into a quad-tree, which recursively divides square image windows into four smaller windows. This separable

approach is similar to the extension of wavelet packet bases in two dimensions, described in Section 8.2.

Let us consider square images of N pixels. A node of the quad-tree is labeled by its depth j and two indices p and q. Let $g_{p,j}[n]$ be the discrete one-dimensional window defined in Section 8.5.2. At depth j, a node (p, q) corresponds to a separable space

$$\mathbf{W}_j^{p,q} = \mathbf{W}_j^p \otimes \mathbf{W}_j^q, \tag{8.120}$$

which is generated by a separable local cosine basis of $2^{-2j}N$ vectors

$$\mathcal{B}_j^{p,q} = \left\{ g_{p,j}[n_1]\, g_{q,j}[n_2]\, \frac{2}{2^{-j}N^{1/2}} \quad \cos\left[\pi\left(k_1 + \frac{1}{2}\right)\frac{n_1 - a_{p,j}}{2^{-j}N^{1/2}}\right] \cos\left[\pi\left(k_2 + \frac{1}{2}\right)\frac{n_2 - a_{q,j}}{2^{-j}N^{1/2}}\right] \right\}_{0 \leqslant k_1, k_2 < 2^{-j}N^{1/2}}$$

We know from (8.119) that

$$\mathbf{W}_j^p = \mathbf{W}_{j+1}^{2p} \oplus \mathbf{W}_{j+1}^{2p+1} \quad \text{and} \quad \mathbf{W}_j^q = \mathbf{W}_{j+1}^{2q} \oplus \mathbf{W}_{j+1}^{2q+1}.$$

Inserting these equations in (8.120) proves that $\mathbf{W}_j^{p,q}$ is the direct sum of four orthogonal subspaces:

$$\mathbf{W}_j^{p,q} = \mathbf{W}_{j+1}^{2p,2q} \oplus \mathbf{W}_{j+1}^{2p+1,2q} \oplus \mathbf{W}_{j+1}^{2p,2q+1} \oplus \mathbf{W}_{j+1}^{2p+1,2q+1}. \tag{8.121}$$

Space $\mathbf{W}_j^{p,q}$ at node (j, p, q) is therefore decomposed in the four subspaces located at the four children nodes of the quad-tree. This decomposition can also be interpreted as a division of the square window $g_{p,j}[n_1]g_{q,j}[n_2]$ into four subwindows of equal sizes, as illustrated in Figure 8.21. The space located at the root of the tree is

$$\mathbf{W}_0^{0,0} = \mathbf{W}_0^0 \otimes \mathbf{W}_0^0. \tag{8.122}$$

It includes all images of N pixels. Size η of the raising and decaying profiles of the one-dimensional windows defines the maximum depth $J = \log_2 \frac{N^{1/2}}{2\eta}$ of the quad-tree.

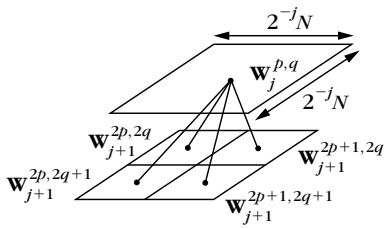

FIGURE 8.21

Functions in $\mathbf{W}_j^{p,q}$ have a support located in a square region of the image. It is divided into four subspaces that cover smaller squares in the image.

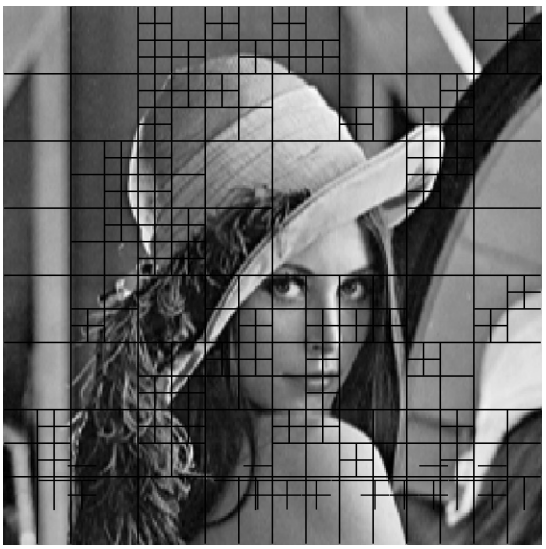

FIGURE 8.22

The grid shows the support of the windows $g_{j,p}[n_1]\,g_{j,q}[n_2]$ of a "best" local cosine basis selected in the local cosine quad-tree.

Admissible Quad-Trees

An admissible subtree of this local cosine quad-tree has nodes that have either zero or four children. Applying the decomposition property (8.121) along the branches of an admissible quad-tree proves that spaces $\mathbf{W}_{j_i}^{p_i,q_i}$ located at the leaves decompose $\mathbf{W}_0^{0,0}$ in orthogonal subspaces. The union of the corresponding two-dimensional local cosine bases $\mathcal{B}_{j_i}^{p_i,q_i}$ is therefore an orthogonal basis of $\mathbf{W}_0^{0,0}$. We proved in (8.42) that there are more than $2^{4^{J-1}} = 2^{N/16\eta^2}$ different admissible trees of maximum depth $J = \log_2 \frac{N^{1/2}}{2\eta}$. These bases divide the image plane into squares of varying sizes. Figure 8.22 gives an example of image decomposition in a local cosine basis corresponding to an admissible quad-tree. This local cosine basis is selected with the best basis algorithm of Section 12.2.2.

Fast Calculations

The decomposition of an image $f[n]$ over a separable local cosine family $\mathcal{B}_j^{p,q}$ requires $O(2^{-2j} N \log_2(2^{-j}N))$ operations with a separable implementation of the fast one-dimensional local cosine transform. For a full local cosine quad-tree of depth J, these calculations are performed for $0 \leqslant p, q < 2^j$ and $0 \leqslant j \leqslant J$, which requires $O(N J \log_2 N)$ multiplications and additions. The original image is recovered from the local cosine coefficients at the leaves of any admissible subtree with $O(N \log_2 N)$ computations.

8.6 EXERCISES

8.1 [2] Prove the discrete splitting Theorem 8.6.

8.2 [2] Meyer wavelet packets are calculated with a Meyer conjugate mirror filter (7.84). Compute the size of the frequency support of $\hat{\psi}_j^p$ as a function of 2^j. Study the convergence of $\psi_{j,n}(t)$ when the scale 2^j goes to $+\infty$.

8.3 [1] Extend the separable wavelet packet tree of Section 8.2.2 for discrete p-dimensional signals. Verify that the wavelet packet tree of a p-dimensional discrete signal of N samples includes $O(N \log_2 N)$ wavelet packet coefficients that are calculated with $O(K N \log_2 N)$ operations if the conjugate mirror filter h has K nonzero coefficients.

8.4 [2] Anisotropic wavelet packets $\psi_j^p[a - 2^{L-j} n_1]\, \psi_l^q[b - 2^{L-l} n_2]$ may have different scales 2^j and 2^l along the rows and columns. A decomposition over such wavelet packets is calculated with a filter bank that filters and subsamples the image rows $j - L$ times, whereas the columns are filtered and subsampled $l - L$ times. For an image $f[n]$ of N pixels, show that a dictionary of anisotropic wavelet packets includes $O(N(\log_2 N)^2)$ different vectors. Compute the number of operations needed to decompose f in this dictionary.

8.5 [2] *Hartley transform*. Let $\mathrm{cas}(t) = \cos(t) + \sin(t)$. We define

$$\mathcal{B} = \left\{ g_k[n] = \frac{1}{\sqrt{N}} \mathrm{cas}\left(\frac{2\pi n k}{N}\right) \right\}_{0 \leq k < N}.$$

(a) Prove that $\mathcal{B}$ is an orthonormal basis of $\mathbb{C}^N$.
(b) For any signal $f[n]$ of size N, find a fast Hartley transform algorithm based on the FFT, which computes $\{\langle f, g_k \rangle\}_{0 \leq k < N}$ with $O(N \log_2 N)$ operations.

8.6 [2] Prove that $\{\sqrt{2} \sin[(k + 1/2)\pi t]\}_{k \in \mathbb{Z}}$ is an orthonormal basis of $\mathbf{L^2}[0, 1]$. Find a corresponding discrete orthonormal basis of $\mathbb{C}^N$.

8.7 [2] Prove that $\{\sqrt{2} \sin(k\pi t)\}_{k \in \mathbb{Z}}$ is an orthonormal basis of $\mathbf{L^2}[0, 1]$. Find a corresponding discrete orthonormal basis of $\mathbb{C}^N$.

8.8 [3] *Lapped Fourier basis*:
(a) Construct a lapped orthogonal basis $\{\tilde{g}_{p,k}\}_{(p,k) \in \mathbb{Z}}$ of $\mathbf{L^2}(\mathbb{R})$ from the Fourier basis $\{\exp(i 2\pi k t)\}_{k \in \mathbb{Z}}$ of $\mathbf{L^2}[0, 1]$.
(b) Explain why this local Fourier basis does not contradict the Balian-Low Theorem 5.20.

(c) Let $f \in \mathbf{L}^2(\mathbb{R})$ be such that $|\hat{f}(\omega)| = O((1+|\omega|^p)^{-1})$ for some $p > 0$. Compute the rate of decay of $|\langle f, \tilde{g}_{p,k}\rangle|$ when the frequency index $|k|$ increases. Compare it with the rate of decay of $|\langle f, g_{p,k}\rangle|$, where $g_{p,k}$ is a local cosine vector (8.104). How do the two bases compare for signal-processing applications?

8.9 [3] Describe a fast algorithm to compute the Meyer orthogonal wavelet transform with a lapped transform applied in the Fourier domain. Calculate the numerical complexity of this algorithm for periodic signals of size N. Compare this result with the numerical complexity of the standard fast wavelet transform algorithm, where the convolutions with Meyer conjugate mirror filters are calculated with an FFT.

8.10 [3] *Mirror wavelets*. Let (h, g) be a pair of conjugate mirror filters. A mirror wavelet packet first decomposes the input signal a_0 into a_1 and d_1 by filtering it with $\bar{h}[n] = h[-n]$ and $\bar{g}[n] = g[-n]$, respectively, and subsampling the output. The signals a_1 and d_1 are subdecomposed with an orthogonal wavelet transform filter bank tree. Decomposing a_1 and d_1 with a cascade of $j-2$ filtering with $\bar{h}$ and a filtering with $\bar{g}$ yields wavelet coefficients $\langle a_0, \psi_{j,m}\rangle$ and wavelet packet coefficients that we write as $\langle a_0, \tilde{\psi}_{j,m}\rangle$. Prove that the discrete Fourier transform of these *mirror wavelets* satisfies $|\widehat{\tilde{\psi}}_{j,m}[k]| = |\widehat{\psi}_{j,m}[N/2-k]|$. Show that if the filters $\bar{h}$ and $\bar{g}$ that decompose d_1 are replaced by h and g, then the resulting mirror wavelets satisfy $\tilde{\psi}_{j,m}[n] = (-1)^{n-1}\psi_{j,m}[1-n]$.

8.11 [2] *Arbitrary Walsh tilings*:

(a) Prove that two Walsh wavelet packets $\psi^p_{j,n}$ and $\psi^{p'}_{j',n'}$ are orthogonal if their Heisenberg boxes, defined in Section 8.1.2, do not intersect in the time-frequency plane [71].

(b) A dyadic tiling of the time-frequency plane is an exact cover $\{[2^j n, 2^j(n+1)] \times [k\pi 2^{-j}, (k+1)\pi 2^{-j}]\}_{(j,n,k)\in I}$, where the index set I is adjusted to guarantee that the time-frequency boxes do not intersect and that they leave no hole. Prove that any such tiling corresponds to a Walsh orthonormal basis of $\mathbf{L}^2(\mathbb{R})$ $\{\psi^p_{j,n}\}_{(p,j,n)\in I}$.

8.12 [3] *Double tree*. We want to construct a dictionary of block wavelet packet bases, which has the freedom to segment both the time and frequency axes. For this purpose, as in a local cosine basis dictionary, we construct a binary tree, which divides $[0, 1]$ in 2^j intervals $[p2^{-j}, (p+1)2^{-j}]$ that correspond to nodes indexed by p at the depth j of the tree. At each of these nodes, we construct another tree of wavelet packet orthonormal bases of $\mathbf{L}^2[p2^{-j}, (p+1)2^{-j}]$ [299].

(a) Define admissible subtrees in this double tree, with leaves corresponding to orthonormal bases of $\mathbf{L}^2[0, 1]$. Give an example of an admissible tree and draw the resulting tiling of the time-frequency plane.

(b) Give a recursive equation that relates the number of admissible subtrees of depth $J+1$ and of depth J. Give an upper bound and a lower bound for the total number of orthogonal bases in this double-tree dictionary.

(c) Can one find a basis in a double tree that is well adapted to implement an efficient transform code for audio signals? Justify your answer.

8.13 [3] An anisotropic local cosine basis for images is constructed with rectangular windows that have a width 2^j that may be different from their height 2^l. Similar to a local cosine tree, such bases are calculated by progressively dividing windows, but the horizontal and vertical divisions of these windows are done independently. Show that a dictionary of anisotropic local cosine bases can be represented as a graph. Implement numerically an algorithm that decomposes images in a graph of anisotropic local cosine bases.