Discrete Fourier Transform and Wavelet Transforms

Supplementary Class Notes for
Math 357: Topics in Applied Algebra

Rutgers, The State University of New Jersey

Roe Goodman

Revised May 3, 2008

The citation Leon in the notes refers to *Linear Algebra with Applications* (6th edition) by Steven J. Leon (published by Prentice-Hall).

The citation Ripples in the notes refers to *Ripples in Mathematics: The Discrete Wavelet Transform* by A. Jensen and A. la Cour-Harbo (published by Springer-Verlag).

Contents

Chapter 1

Complex Inner Product Spaces

1.1 Complex Numbers

Introductory linear algebra courses consider vector spaces of column vectors and matrices (or functions) with real entries (or real values). For Fourier and Wavelet analysis it is more natural in some situations to use vector spaces with *complex numbers* as scalars. We begin with a brief description of the £eld $\mathbb{C}$ of complex numbers.

De£nition 1.1.1. A *complex number* z is a pair (x, y) of real numbers, denoted by $z = x + y\,\mathrm{i}$ (in engineering texts usually denoted by $z = x + y\,\mathrm{j}$). We call x the *real part* of z and y the *imaginary part* of z. Complex numbers are added by adding the real and imaginary parts separately, just as if they were vectors in $\mathbb{R}^2$. Multiplication is de£ned by

$$(a + b\,\mathrm{i})(x + y\,\mathrm{i}) = (ax - by) + (by + ax)\,\mathrm{i}$$

for $a, b, x, y \in \mathbb{R}$. In particular, $\mathrm{i}^2 = -1$ and $a(x + y\,\mathrm{i}) = ax + ay\,\mathrm{i}$ is the same as multiplying a vector in $\mathbb{R}^2$ by a real scalar. ■

We denote the set of all complex numbers by $\mathbb{C}$. It is clear from the de£nition that addition and multiplication of complex numbers is commutative: $(a + b\,\mathrm{i})(x + y\,\mathrm{i}) = (x + y\,\mathrm{i})(a + b\,\mathrm{i})$. The distributive law $u(v+w) = uv+uw$ is also obvious. An easy but slightly tedious calculation shows that multiplication of complex numbers is associative:

$$u(vw) = (uv)w \quad \text{for } u, v, w \in \mathbb{C}. \tag{1.1}$$

(This property can be deduced from the associativity of matrix multiplication–see the exercises). For a complex number $c = a + \mathrm{i}\,b$ we de£ne the *complex conjugate* $\bar{c} = a - b\,\mathrm{i}$ and the *modulus* $|c|$ by

$$|c| = \sqrt{\bar{c}c} = \sqrt{a^2 + b^2}.$$

Note that if $c \neq 0$ then $|c| > 0$ and $(|c|^{-2}\bar{c})c = 1$. Hence every nonzero complex number c has a *multiplicative inverse*

$$c^{-1} = |c|^{-2}\bar{c} = \left(\frac{a}{a^2 + b^2}\right) - \left(\frac{b}{a^2 + b^2}\right)\mathrm{i}$$

It is easy to check that $\overline{zw} = \bar{z}\bar{w}$ for every $z, w \in \mathbb{C}$. Hence

$$|zw| = |z|\,|w| \quad \text{and} \quad |z^n| = |z|^n \quad \text{for all integers } n \text{ if } z \neq 0.$$

We view the set $\mathbb{R}$ of real numbers as a subset of $\mathbb{C}$ by identifying $x \in \mathbb{R}$ with the complex number $x + 0\,\mathrm{i}$. The algebraic operations in $\mathbb{C}$ just de£ned then reduce to the usual algebraic operation on real numbers under this identi£cation.

As an algebraic system, the set of complex numbers satis£es the axioms of a *£eld* (relative to addition and multiplication), just like the real numbers. There is no *order relation* $a < b$ for complex numbers, unlike the case of the real numbers. However, there is a *polar decomposition* of the complex number $z = x + y\,\mathrm{i}$:

$$z = r\mathrm{e}^{\theta\,\mathrm{i}} = r\cos(\theta) + r\sin(\theta)\,\mathrm{i}\,, \quad \text{where } r = |z| \text{ and } \theta = \arctan(y/x). \tag{1.2}$$

This is just the formula for polar coordinates in $\mathbb{R}^2$ written in terms of complex numbers. We call θ the *argument* of z. Since $\mathrm{e}^{(\theta+2\pi m)\,\mathrm{i}} = \mathrm{e}^{\theta\,\mathrm{i}}$ for any integer m, the argument of z is only determined up to the addition of integer multiples of 2π. In particular, for any $\theta \in \mathbb{R}$, we have $|\mathrm{e}^{\theta\,\mathrm{i}}| = 1$ and the complex number $w = \mathrm{e}^{\theta\,\mathrm{i}}$ is a point on the unit circle in $\mathbb{R}^2$.

The complex exponential function in equation (1.2) can be de£ned directly in terms of its Taylor series centered at zero:

$$\mathrm{e}^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!} + \cdots$$

(here t can be any complex number). Since $|t^n/n!| = |t|^n/n!$, it follows from the ratio test that the partial sums of this series converge absolutely and uniformly in every disc $|t| \leq R$ for any value of R. Here *convergence* of a sequence of complex numbers means convergence of the real and imaginary parts of the sequence. The exponential function satis£es the *law of exponents*

$$\mathrm{e}^{s+t} = \mathrm{e}^s\mathrm{e}^t \quad \text{for all } s, t \in \mathbb{C}$$

(this is easy to verify from the binomial formula and rearrangement of the exponential series).

For any positive integer N the complex number $w = \mathrm{e}^{-2\pi\,\mathrm{i}/N} = \cos(2\pi/N) - \mathrm{i}\,\sin(2\pi/N)$ has absolute value one and satis£es $w^N = 1$ (since $(\mathrm{e}^{-2\pi\,\mathrm{i}/N})^N = \mathrm{e}^{2\pi\,\mathrm{i}} = 1$ by the law of exponents). The number w is a *primitive* Nth root of unity, since every other Nth root of 1 is of the form w^k for some integer k.

See the appendix to *Calculus* by James Stewart for a more details about complex numbers.

1.2 Real and Complex Vector Spaces

Now that we have enlarged the £eld of scalars from $\mathbb{R}$ to $\mathbb{C}$, we can carry out all the constructions of linear algebra using complex numbers. We de£ne $\mathbb{C}^n$ to be the set of all column vectors

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \quad \text{where } z_1, \ldots, z_n \in \mathbb{C}.$$

We add complex column vectors and multiply column vectors by complex scalars in terms of their components, just as in the real case, but now using addition and multiplication of complex numbers.

More generally, a *complex vector space* V is a set of vectors with an addition operation and multiplication by complex scalars, satisfying the same commutative, associative, and distributive properties as in the real case (see Leon, page 129).

Example 1.2.1. We denote by $M_{m\times n}(\mathbb{C})$ the set of $m \times n$ matrices whose entries are complex numbers. We add matrices of the same size and multiply matrices by complex scalars just as in the case of real matrices. This makes $M_{m\times n}(\mathbb{C})$ into a complex vector space. We de£ne *matrix multiplication* as in the real case, but using the formula for multiplying complex numbers. Since multiplication of complex numbers is associative, it follows that matrix multiplication is also associative: $A(BC) = (AB)C$ when A, B, and C are complex matrices of compatible sizes. ■

Example 1.2.2. Let V be the collection of all complex-valued functions de£ned on a set X. We make V into a vector space by *pointwise* operations:

$$(f+g)(x) = f(x) + g(x), \quad (\alpha f)(x) = \alpha f(x) \quad \text{for } f, g \in V,\ \alpha \in \mathbb{C}, \text{ and } x \in X.$$

These operations on V satisfy the vector space axioms (see Leon, page 130). In particular, when $X = \{1, 2, \ldots, n\}$ we can identify V with $\mathbb{C}^n$ by viewing the components of a column vector as a function on X. ■

All the de£nitions and results of linear algebra, such as *subspace, linear independence, basis, dimension, linear transformation, null space, range*, and *rank*, apply to complex vector spaces with no modi£cation except the use of complex scalars. This is true because the proofs of these results only use the vector space axioms and fact that the scalars satisfy the £eld axioms (see Leon, Chapters 3 and 4). Note, however, that if we view $\mathbb{R}^n$ as the subset of real column vectors in $\mathbb{C}^n$, then $\mathbb{R}^n$ is *not* a subspace of $\mathbb{C}^n$, since $\mathrm{i}\,\mathbb{R}^n$ is not contained in $\mathbb{R}^n$.

If V and W are £nite-dimensional complex vector spaces and $T : V \longrightarrow W$ is a complex linear transformation, then T can be represented by a matrix with complex entries, relative to choices of bases for V and W. Composition of linear transformations then corresponds to multiplication of the matrices for the transformations, just as in the real case (see Leon, Chapter 4.2).

1.3 Linear Transformations and Matrices

Let V and W be vector spaces with scalars $\mathbb{F}$ (real or complex numbers). Let $L : V \longrightarrow W$ be a function. We say that L is a *linear transformation* if

$$L(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 L\mathbf{v}_1 + \alpha_2 L\mathbf{v}_2$$

for all vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and all scalars α_2 and α_2.

1.4 Inner Products and Unitary Transformations

The formula de£ning the usual inner product and norm on $\mathbb{R}^n$ (see Leon, Section 5.4) needs to be modi£ed when we de£ne an inner product on $\mathbb{C}^n$. For a nonzero real number x we always have

$x^2 > 0$. But this is not true for complex numbers, since $\mathrm{i}^2 = -1$. The way around this difficulty is to use the fact that $\bar{z}z > 0$ if z is a nonzero complex number. Thus we define the *standard inner product* on $\mathbb{C}^n$ to be

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^{n} u_k \bar{v}_k \quad \text{for } \mathbf{u}, \mathbf{v} \in \mathbb{C}^n.$$

Just as in the real case we can write the inner product in terms of matrix multiplication of a row vector ($1 \times n$ matrix) and a column vector ($n \times 1$ matrix). For this we define the *Hermitian transpose* $\mathbf{v}^{\mathrm{H}} = \bar{\mathbf{v}}^{\mathrm{T}}$. Likewise, if A is an $m \times n$ matrix, we write $A^{\mathrm{H}} = \bar{A}^{\mathrm{T}}$. (Note that in MATLAB all matrices are automatically assumed to have complex entries, and `A'` gives the Hermitian transpose of a matrix `A`.) Then we can express

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^{\mathrm{H}}\, \mathbf{u} \quad \text{for } \mathbf{u}, \mathbf{v} \in \mathbb{C}^n.$$

With this definition we have

$$\langle \mathbf{u}, \mathbf{u} \rangle = \sum_{k=1}^{n} |u_k|^2 = \mathbf{u}^{\mathrm{H}}\, \mathbf{u},$$

which is positive (unless $\mathbf{u} = 0$, when it is zero). Thus we can define the *norm* $||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ which measures the total size of a vector with complex components.

Definition 1.4.1. Let V be a complex vector space. An *inner product* on V is a complex-valued function $\langle \mathbf{u}, \mathbf{v} \rangle$ defined for all $\mathbf{u}, \mathbf{v} \in V$ that satisfies the following conditions:

(Positivity) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ with equality if and only if $\mathbf{u} = 0$.

(Conjugate Symmetry) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$.

(Linearity) $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and complex numbers α and β.

■

When $V = \mathbb{C}^n$ then the standard inner product defined above satisfies these conditions. Here is another important example.

Example 1.4.2. Consider the complex vector space V of all complex-valued continuous functions on a finite interval $[a,\, b]$. Let $w(x)$ be any continuous function on $[a,\, b]$ that is strictly positive. (For example, $w(x) = (1 + x^2)^p$ for some fixed real number p.) Given two functions f and g in V, define

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\, w(x)dx. \tag{1.3}$$

To verify that this is an inner product, note that $f(x)\overline{f(x)} \geq 0$, so we have $\langle f, f \rangle \geq 0$. If $\langle f, f \rangle = 0$ then the same argument as in Leon, Section 5.4 (p. 261) shows that $f(x) = 0$ for all $a \leq x \leq b$. The conjugate symmetry and linearity are obvious. ■

Let V be a complex vector space with a £xed inner product. The *norm* associated with an inner product is $||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$, just as in the case of $\mathbb{C}^n$. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, and we write $\mathbf{u} \perp \mathbf{v}$. For orthogonal vectors we have the *Pythagorean Law* (complex version):

$$||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 \quad \text{when } \mathbf{u} \perp \mathbf{v}$$

with the same proof as in the real case (see Leon, page 262). For any pair of vectors $\mathbf{u}$, $\mathbf{v}$ with $\mathbf{v} \neq 0$, the *vector projection* of $\mathbf{u}$ onto $\mathbf{v}$ is given by

$$\mathbf{p} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$$

just as in the real case. Since $(\mathbf{u} - \mathbf{p}) \perp \mathbf{p}$ and $\mathbf{u} = (\mathbf{u} - \mathbf{p}) + \mathbf{p}$, the Pythagorean Law gives

$$||\mathbf{u}||^2 = ||\mathbf{u} - \mathbf{p}||^2 + ||\mathbf{p}||^2.$$

Using this equation we obtain the *Cauchy-Schwarz inequality*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}|| \, ||\mathbf{v}|| \tag{1.4}$$

by the same argument as in Leon, page 265. From the Cauchy-Schwarz inequality we obtain the *triangle inequality*

$$||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}|| \quad \text{for all vectors } \mathbf{u}, \mathbf{v} \in V.$$

(see Leon, page 266).

De£nition 1.4.3. A set of nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots,$ in an inner product space V is called *orthogonal* if $\mathbf{v}_j \perp \mathbf{v}_k$ for all $j \neq k$. If the set is orthogonal and each vector satis£es $||\mathbf{v}_j|| = 1$ then the set is called *orthonormal*. ■

An orthonormal set of vectors is always linearly independent (see Leon, page 270). Assume that $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a £nite orthonormal set. Let U be the subspace of V spanned by this set of vectors. Then $\dim U = n$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an *orthonormal basis* for W. Every vector $\mathbf{u} \in W$ can be expressed in terms of this basis as

$$\mathbf{u} = c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n, \quad \text{where } c_j = \langle \mathbf{w}, \mathbf{u}_j \rangle.$$

(The formula for the coef£cient c_j follows by taking the inner product of $\mathbf{u}$ with $\mathbf{u}_j$ and using orthonormality.) Then

$$||\mathbf{u}||^2 = |c_1|^2 + \cdots + |c_n|^2 \qquad \text{(Parseval's Formula)}$$

If $\mathbf{v} = d_1 \mathbf{u}_1 + \cdots + d_n \mathbf{u}_n$ is another vector in U, then

$$\langle \mathbf{u}, \mathbf{v} \rangle = c_1 \bar{d}_1 + \cdots + c_n \bar{d}_n$$

(see Leon, page 272). For any vector $\mathbf{v} \in V$ we de£ne

$$P\mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n.$$

Then $P\mathbf{v} \in U$, since it is a linear combination of the vectors $\mathbf{u}_k$. Furthermore, $\mathbf{v} - P\mathbf{v} \perp U$ since

$$\begin{aligned}\langle \mathbf{v} - P\mathbf{v}, \mathbf{u}_j \rangle &= \langle \mathbf{v}, \mathbf{u}_j \rangle - \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{u}_k \rangle \langle \mathbf{u}_k, \mathbf{u}_j \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v}, \mathbf{u}_j \rangle \langle \mathbf{u}_j, \mathbf{u}_j \rangle = 0\end{aligned}$$

by orthogonality and the fact that $||\mathbf{u}_j|| = 1$. We call $P\mathbf{v}$ the *orthogonal projection of* $\mathbf{v}$ *onto the subspace* U. By the Pythagorean Law,

$$||\mathbf{v}||^2 = ||P\mathbf{v}||^2 + ||\mathbf{v} - P\mathbf{v}||^2 \tag{1.5}$$

This implies that $P\mathbf{v}$ is the vector in U that is closest to $\mathbf{v}$ (see Leon, Theorem 5.5.8).

Example 1.4.4 (Fourier Series)**.** Let V be the complex vector space of piecewise continuous complex-valued functions $f(x)$ on the interval $0 \le x \le 2\pi$. De£ne an inner product on V by

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{g(x)}\, dx.$$

The functions $\phi_n(x) = \mathrm{e}^{\mathrm{i}\,nx}$ for $n \in \mathbb{Z}$ (the set of all integers) are in V, and they are an orthonormal set of functions. To see this, let $k \neq n$ and calculate

$$\langle \phi_n, \phi_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\,(n-k)x}\, dx = \left. \frac{\mathrm{e}^{\mathrm{i}\,(n-k)x}}{2\pi\,\mathrm{i}\,(n-k)} \right|_{x=0}^{x=2\pi} = 0$$

because $\mathrm{e}^{2m\pi\,\mathrm{i}} = 1$ for all integers m. Thus $\phi_n \perp \phi_k$. Furthermore, since $\mathrm{e}^0 = 1$, we have

$$\langle \phi_n, \phi_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} 1\, dx = 1$$

Thus $||\phi_n|| = 1$ and $\{\phi_k \,:\, k \in \mathbb{Z}\}$ is an orthonormal set in V.

If $f \in V$ then the complex numbers

$$c_k = \langle f, \phi_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\mathrm{e}^{-\mathrm{i}\,kx}\, dx \tag{1.6}$$

are called the *Fourier coef£cients* of f. It is an important result of Fourier analysis that f can be represented by its *Fourier series*:

$$f = \sum_{k\in\mathbb{Z}} c_k\, \phi_k \tag{1.7}$$

This is analogous to the representation of a vector in $\mathbb{C}^n$ in terms of an orthonormal basis for $\mathbb{C}^n$ (since $c_k = \langle f, \phi_k \rangle$). Furthermore, the *in£nite series* version of Parseval's formula is valid:

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2\, dx = \sum_{k\in\mathbb{Z}} |c_k|^2 \tag{1.8}$$

In particular, the in£nite series on the right side of (1.8) converges. (The convergence properties of the series (1.7) and the proof of (1.8) require results from advanced calculus that will not be discussed in this course).

Let $\mathcal{TP}_n$ be the linear span of the set of functions $\{\phi_k(x) : |k| \le n\}$. If $f(x) \in \mathcal{TP}_n$ then $c_k = 0$ for $|k| > n$ and the right side of (1.7) is a *trigonometric polynomial*

$$f(x) = \sum_{-n \le k \le n} c_k \, \phi_k(x)$$

When $f(x)$ is real-valued its Fourier coefficients c_k have the property

$$\overline{c_k} = c_{-k},$$

since $\overline{\phi_k(x)} = \phi_{-k}(x)$. For example, the formulas

$$\sin(nx) = \frac{1}{2\,\mathrm{i}} \mathrm{e}^{nx\,\mathrm{i}} - \frac{1}{2\,\mathrm{i}} \mathrm{e}^{-nx\,\mathrm{i}}, \qquad \cos(nx) = \frac{1}{2} \mathrm{e}^{nx\,\mathrm{i}} + \frac{1}{2} \mathrm{e}^{-nx\,\mathrm{i}}$$

show that the real-valued functions $f_n(x) = \sin(nx)$ and $g_n(x) = \cos(nx)$ are in $\mathcal{TP}_n$ and have Fourier series

$$f_n(x) = \frac{1}{2\,\mathrm{i}} \phi_n(x) - \frac{1}{2\,\mathrm{i}} \phi_{-n}(x)\,, \qquad g_n(x) = \frac{1}{2} \phi_n(x) + \frac{1}{2} \phi_{-n}(x) \tag{1.9}$$

For any function $f \in V$ and positive integer n, the trigonometric polynomial

$$\psi_n(x) = \sum_{-n \le k \le n} \langle f, \phi_k \rangle \phi_k(x)$$

is the projection of $f(x)$ onto the subspace $\mathcal{TP}_n$, since $\{\phi_k(x) : -n \le k \le n\}$ is an orthonormal basis for $\mathcal{TP}_n$. The function $\psi_n(x)$ gives the *best approximation* to $f(x)$ (in the sense of the norm $|| \cdot ||$), since we minimize the norm $||f - \psi||$, where ψ is a trigonometric polynomial in $\mathcal{TP}_n$, by taking $\psi = \psi_n$. ■

Definition 1.4.5. An $n \times n$ matrix U is said to be a *unitary matrix* if the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of columns of U is orthonormal. ■

The matrix U is unitary if and only if

$$\langle U\mathbf{v}, U\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n. \tag{1.10}$$

To prove this, use the linearity of the inner product in each variable to see that (1.10) is satisfied for all vectors $\mathbf{v}, \mathbf{w}$ if and only if it is satisfied when $\mathbf{v} = \mathbf{e}_j$ and $\mathbf{w} = \mathbf{e}_k$ (the standard basis vectors for $\mathbb{C}^n$). Since the jth column of U is $U\mathbf{e}_j$ and the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is orthonormal, it follows that (1.10) is equivalent to the statement that the columns of U are an orthonormal set.

An alternate characterization of unitary matrices is that $U^{\mathrm{H}}U = I$, where U^{H} denotes the conjugate transpose matrix (the proof is the same as for real orthogonal matrices–see Leon, page 273). Hence a unitary matrix is invertible, with inverse $U^{-1} = U^{\mathrm{H}}$.

Now let V and W be finite-dimensional complex inner product spaces of the same dimension, and let T be a linear transformation from V to W. We say that T is a *unitary transformation* if

$$\langle T\mathbf{u}, T\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \quad \text{for all vectors } \mathbf{u}, \mathbf{v} \in V. \tag{1.11}$$

Note that in equation (1.11) the inner product on the left is for the space W, while the inner product on the right is for the space V. Taking $\mathbf{u} = \mathbf{v}$, we see that $||T\mathbf{u}|| = ||\mathbf{u}||$ for all $\mathbf{u}$. Hence the null space of T is 0. Since V and W have the same dimension, T is represented by a square matrix (relative to a choice of bases for V and W). This matrix has no null space, so it is invertible. Thus every unitary transformation is invertible.

Example 1.4.6. Let $V = W = \mathbb{C}^n$, and let the linear transformation T have matrix

$$U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$$

relative to the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{C}^n$ (where $\mathbf{e}_j$ has 1 in the jth entry and zero elsewhere). Since the standard basis is orthonormal, we see from (1.11) that T is a unitary transformation if and only if U is a unitary matrix. ■

Example 1.4.7. Let $V = \mathcal{TP}_2$ be the space of trigonometric polynomials of degree at most 2 with the inner product (1.3). Then the set of functions $\{\phi_{-2},\, \phi_{-1},\, \phi_0,\, \phi_1,\, \phi_2\}$ is an orthonormal basis for V. If $f \in V$, de£ne

$$Tf = \begin{bmatrix} c_{-2} \\ c_{-1} \\ c_0 \\ c_1 \\ c_2 \end{bmatrix} \tag{1.12}$$

where c_k are the Fourier coef£cients of $f(x)$ de£ned by equation (1.6). Notice that there is no variable x displayed in formula (1.12); Tf is a vector in $\mathbb{C}^5$ with £ve numerical components, whereas the continuous function $f(x)$ is considered as a vector in the space V.

Since the Fourier coef£cients depend linearly on f, it is clear that T is a linear transformation from V to $\mathbb{C}^5$. The basis function $\phi_k(x)$ for $\mathcal{TP}_2$ is tranformed by T into the standard basis vector $\mathbf{e}_{3+k}$ for $k = -2, \ldots, 2$, hence T is unitary. From equation (1.9) we see that the functions $f_2(x) = \sin(2x)$ and $g_2(x) = \cos(2x)$ have transforms

$$Tf_2 = \frac{1}{2\,\mathrm{i}} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \qquad Tg_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Since T is unitary, it follows that

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} \sin(2x)\cos(2x)\,dx &= \langle f_2, g_2\rangle = \langle Tf_2, Tg_2\rangle = 0 \\ \frac{1}{2\pi}\int_0^{2\pi} \sin^2(2x)\,dx &= \langle f_2, f_2\rangle = \langle Tf_2, Tf_2\rangle = \frac{1}{2} \end{aligned}$$

The two integrals on the left can be evaluated by double-angle formulas, of course, but this is not necessary because we already know that T is unitary. ■

1.5 Exercises

1. Given a complex number $z = x + \mathrm{i}\,y$, let $T(z) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$.

   (a) Show that for any complex numbers z and w, the following matrix equalities hold:

   $$T(z + w) = T(z) + T(w) \quad \text{and} \quad T(zw) = T(z)T(w)$$

(matrix addition and multiplication on the right sides of the equalities). Since matrix multiplication is associative, it follows that multiplication of complex numbers is also associative.

(b) Express $T(\bar{z})$ and $|z|^2$ in terms of the matrix $T(z)$.

2. Consider the set $\mathbf{H}$ of all complex 2×2 matrices of the form $A = \begin{bmatrix} w & -\bar{z} \\ z & \bar{w} \end{bmatrix}$, where w and z are arbitrary complex numbers.

   (a) Show that the sum and the product of two matrices in $\mathbf{H}$ are again in $\mathbf{H}$, and any *real* multiple of a matrix in $\mathbf{H}$ is in $\mathbf{H}$.

   (b) From (a) it follows that $\mathbf{H}$ is a real vector space. Find a basis for it. (HINT: Consider the special cases where one of z or w is either 1 or i, and the other is zero.)

   (c) Which matrices in $\mathbf{H}$ are invertible? (HINT: Calculate $\det A$.)

   (d) If a matrix in $\mathbf{H}$ is invertible, is the inverse matrix also in $\mathbf{H}$?

   (e) Is $AB = BA$ for all $A, B \in \mathbf{H}$? (HINT: Calculate the products of the basis matrices from (b).)

3. Let $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Find unit vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}^2$ such that $R\mathbf{u} = \mathrm{i}\,\mathbf{u}$ and $R\mathbf{v} = -\mathrm{i}\,\mathbf{v}$. Prove that $\mathbf{u} \perp \mathbf{v}$ and construct a unitary matrix U such that $U^H R U = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}$.

4. Let $V = C[-1, 1]$ (the continuous real-valued functions on the interval $-1 \leq x \leq 1$). Let U be the set of all functions $f \in V$ such that $f(-1) = 2f(1)$. *True* or *False*: U is a subspace. Justify your answer.

5. Let P_3 be the space of all polynomials of degree less than 3. Let $f_1(x) = x+2$, $f_2(x) = x+3$, and $f_3(x) = x^2 + x$.

   (a) *True* or *False*: $\{f_1, f_2, f_3\}$ spans P_3. Justify your answer.

   (b) *True* or *False*: $\{f_1, f_2, f_3\}$ is a linearly independent set. Justify your answer.

6. Let P_4 be the space of all polynomials of degree less than 4. Let V be the subspace of P_4 consisting of all polynomials $f(x)$ such that $f(0) = 0$ and $f(1) = 0$. Give a basis for V. Justify your answer by showing that your basis set satis£es the two conditions needed for a basis. (HINT: Every function in V is divisible by $x(x-1)$.)

7. Let P_2 be the space of all polynomials of degree less than 2 and P_3 the space of all polynomials of degree less than 3. Let L be the linear transformation from P_2 to P_3 given by

$$Lf(x) = 2f'(x) + (3x + 4)f(x).$$

   Calculate the action of L on the ordered basis $\{x, 1\}$ for P_2 in terms of the ordered basis $\{x^2, x, 1\}$ for P_3. Use this to £nd the matrix A for L relative to these ordered bases.

8. Consider the complex vector space $\mathbb{C}^2$ with the standard inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^H \mathbf{u}$. Let $\mathbf{u} = \begin{bmatrix} 2 \\ 3 + 4\,\mathrm{i} \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 5\,\mathrm{i} \\ 1 + \mathrm{i} \end{bmatrix}$. Calculate $\langle \mathbf{u}, \mathbf{v} \rangle$.

9. Consider the vector space $C[0,1]$ of continuous real-valued functions $f(x)$ on $0 \le x \le 1$, with inner product
$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx.$$
Let $f(x) = x^2$ and $g(x) = x$.

   (a) Calculate $\langle f, g \rangle$.

   (b) Calculate $||f||$ and $||g||$.

   (c) Illustrate the *Cauchy-Schwarz inequality* with the functions f and g.

   (d) Calculate the *vector projection* p of f onto g.

   (e) Verify that the function p you found satis£es $(f - p) \perp g$.

10. Consider the vector space $V = C[0,1]$ of continuous real-valued functions $f(x)$ on $0 \le x \le 1$, with inner product
$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx.$$
Let U be the subspace of V with basis $\{u_1, u_2\}$, where $u_1(x) = 1$ and $u_2(x) = \sqrt{3}(2x - 1)$.

   (a) Show that $\{u_1, u_2\}$ is an orthonormal set of functions.

   (b) Let $f(x) = x^2$. Calculate the projection $p(x)$ of $f(x)$ onto U.

   (c) Let $f(x) = x^2$ as in (b). Suppose $u(x)$ is any function in U. What choice of u gives the minimum value for $||f - u||$?

Chapter 2

Discrete Fourier Transform

2.1 Finite Fourier Transform

We shall call a piecewise continuous complex-valued function $s(t)$ of the real variable t an *analog signal* (think of t as *time* and $s(t)$ as measuring the intensity of a sound). We assume that $s(t)$ is of *£nite duration*, so that is zero outside some interval $a \leq t \leq b$. We shift and rescale the variable t to make $a = 0$ and $b = 2\pi$. Next, we choose integers $m < n$ and replace $s(t)$ by the best approximation to $s(t)$ by trigonometric polynomials with frequencies in the range $m \leq k < n$:

$$q(t) = \sum_{m \leq k < n} c_k \mathrm{e}^{\,\mathrm{i}\,kt}$$

The Fourier coef£cients c_k are obtained by integration:

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} s(t) \mathrm{e}^{-\,\mathrm{i}\,kt}\, dt$$

The *mean square approximation error* is

$$||s - q||^2 = \frac{1}{2\pi} \int_0^{2\pi} |s(t) - q(t)|^2\, dt = \sum_{k<m} |c_k|^2 + \sum_{k \geq n} |c_k|^2 \tag{2.1}$$

by Parseval's equality (1.8). The right side of (2.1) is the tail of a convergent series, so $q(t)$ will be a good approximation to $s(t)$ (on average) if the frequency band $m \leq k < n$ is chosen suf£ciently wide.

For a given signal $s(t)$ we £x a *frequency band* $m \leq k < n$ so that the approximation error (2.1) is small. Let $N = n - m$. We replace the functions $s(t)$ and $q(t)$ by

$$f(t) = \mathrm{e}^{-\,\mathrm{i}\,mt} s(t) \quad \text{and} \quad p(t) = \mathrm{e}^{-\,\mathrm{i}\,mt} q(t)$$

This frequency shift doesn't change the approximation error (2.1), since $|\mathrm{e}^{-\,\mathrm{i}\,mt}| = 1$. Since $\mathrm{e}^{-\,\mathrm{i}\,mt}\mathrm{e}^{\,\mathrm{i}\,kt} = \mathrm{e}^{\,\mathrm{i}\,(k-m)t}$, the trigonometric polynomial $p(t)$ has frequencies $0 \leq k < N$:

$$p(t) = \sum_{0 \leq k < N} d_k\, \mathrm{e}^{\,\mathrm{i}\,kt}$$

Here the Fourier coef£cients are

$$d_k = \frac{1}{2\pi}\int_0^{2\pi} f(t)\mathrm{e}^{-\mathrm{i}\,kt}\,dt \tag{2.2}$$

In signal processing applications there is no formula for $f(t)$, so the integrals (2.2) must be approximated using some numerical method. The simplest way to do this is to convert f into a *digital signal* $\mathbf{y} \in \mathbb{C}^N$ by sampling f at the N equal-spaced t values $t_j = 2\pi j/N$, for $j = 0, 1, \ldots, N-1$:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[1] \\ \vdots \\ \mathbf{y}[N-1] \end{bmatrix} \quad \text{where } \mathbf{y}[j] = f(t_j) \text{ for } j = 0, 1, \ldots, N-1. \tag{2.3}$$

Here we are following the notation in Ripples for the value $\mathbf{y}[k]$ of the digital signal $\mathbf{y}$ at discrete time k (note that the indexing of the components in $\mathbf{y}$ is different than the usual MATLAB indexing, which would go from 1 to N). We call N the *sampling rate*; the choice of this sampling rate is determined by the number of Fourier coef£cients that we need to get a good representation of the signal (more coef£cients require a higher sampling rate). With this choice we have

$$\Delta t = t_j - t_{j-1} = 2\pi/N, \qquad \frac{\Delta t}{2\pi} = \frac{1}{N}.$$

Hence we can approximate the integral (2.2) by the Riemann sum

$$d_k \approx \frac{1}{N}\sum_{j=0}^{N-1} f(t_j)\,\mathrm{e}^{-\mathrm{i}\,kt_j} = \frac{1}{N}\sum_{j=0}^{N-1} y[j]\,w^{-jk} \tag{2.4}$$

where $w = \mathrm{e}^{2\pi\,\mathrm{i}/N} = \cos(2\pi/N) + \mathrm{i}\,\sin(2\pi/N)$ is a primitive Nth root of unity.

De£nition 2.1.1 (Fourier Matrix)**.** Let F_N be the $N \times N$ matrix with (j,k) entry $w^{-(j-1)(k-1)}$, where $w = \mathrm{e}^{2\pi\,\mathrm{i}/N}$. The entries in the £rst column of F_N are all 1. The second column consists of the powers of w^{-1} from 0 to $N-1$, the third column consists of the powers of w^{-2} from 0 to $N-1$, and so on. Since F_N is symmetric, the same description applies to the rows. ■

For example, since $\mathrm{e}^{2\pi\,\mathrm{i}/2} = -1$ the 2×2 Fourier matrix is

$$F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{2.5}$$

For $N = 4$ we have $w = \mathrm{e}^{2\pi\,\mathrm{i}/4} = \mathrm{i}$ and $w^{-1} = -\mathrm{i}$. Hence the 4×4 Fourier matrix is

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & (-\mathrm{i})^2 & (-\mathrm{i})^3 \\ 1 & (-\mathrm{i})^2 & (\mathrm{i})^4 & (-\mathrm{i})^6 \\ 1 & (-\mathrm{i})^3 & (-\mathrm{i})^6 & (-\mathrm{i})^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix}. \tag{2.6}$$

Let $\mathbf{d} \in \mathbb{C}^N$ be the vector with components $d_0, d_1, \ldots, d_{N-1}$ given by (2.4). Then

$$\mathbf{d} = \frac{1}{N}F_N\mathbf{y} \tag{2.7}$$

Theorem 2.1.2. *The matrix $(1/\sqrt{N})F_N$ is unitary. Hence the matrix $(1/N)F_N$ has inverse $\overline{F}_N$, and the digital signal vector $\mathbf{y}$ can be reconstructed from the sampled Fourier coef£cient vector $\mathbf{d}$ by $\mathbf{y} = \overline{F}_N\mathbf{d}$.*

Proof. To simplify the notation we label the columns of F_N from 0 to $N-1$. The kth column of F_N is then

$$\mathbf{h}_k = \begin{bmatrix} 1 & w^{-k} & w^{-2k} & \cdots & w^{-(N-1)k} \end{bmatrix}^{\mathrm{T}}$$

Hence the inner product of the jth and kth columns of F_N is

$$\langle \mathbf{h}_j, \mathbf{h}_k \rangle = \mathbf{h}_k^{\mathrm{H}}\,\mathbf{h}_j = 1 + w^{k-j} + w^{2(k-j)} + \cdots + w^{(N-1)(k-j)} \tag{2.8}$$

since $\overline{w} = w^{-1}$. For $j = k$ this gives $\langle \mathbf{h}_j, \mathbf{h}_j \rangle = N$. Now suppose $j \neq k$ and write $u = w^{k-j}$. Then the right side of (2.8) is a £nite geometric series in powers of u:

$$1 + u + u^2 + \cdots + u^{N-1} = \frac{1-u^N}{1-u}.$$

(Note that $u \neq 1$ because $0 < |j-k| < N$ and $w^p = 1$ only when p is an integer multiple of N.) But $u^N = w^{N(j-k)} = 1$, so we conclude that $\langle \mathbf{h}_j, \mathbf{h}_k \rangle = 0$ for $j \neq k$. These orthogonality relations can be written in matrix form as

$$F_N \big(F_N\big)^{\mathrm{H}} = N I_N\,, \tag{2.9}$$

where I_N is the $N \times N$ identity matrix. Since F_N is symmetric, we have $\big(F_N\big)^H = \overline{F}_N$. Hence the matrix $(1/N)F_N$ has inverse $\overline{F}_N$, as claimed. Equation (2.9) can be rewritten as

$$(1/\sqrt{N})F_N\,(1/\sqrt{N})\overline{F}_N = I_N$$

which shows that $(1/\sqrt{N})F_N$ is a unitary matrix. ∎

Corollary 2.1.3.

(a) *Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ be the standard basis for $\mathbb{C}^N$. Set $\mathbf{u}_j = (1/\sqrt{N})F_N\mathbf{e}_j$. Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ is an orthonormal basis for $\mathbb{C}^N$, called the* Fourier basis.

(b) *Let $\mathbf{y} \in \mathbb{C}^N$ and set $\mathbf{d} = (1/N)F_N\mathbf{y}$. Then $\frac{1}{N}||\mathbf{y}||^2 = ||\mathbf{d}||^2$.*

Proof. (a): Note that $\mathbf{u}_j$ is the jth column of the unitary matrix $(1/\sqrt{N})F_N$.

(b): Since $\sqrt{N}\mathbf{d} = (1/\sqrt{N})F_N\mathbf{y}$ and $(1/\sqrt{N})F_N$ is a unitary matrix, the vectors $\sqrt{N}\mathbf{d}$ and $\mathbf{y}$ have the same norm. ∎

Example 2.1.4. Suppose $N = 4$. The Fourier basis for $\mathbb{C}^4$ is

$$\mathbf{u}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -\mathrm{i} \\ -1 \\ \mathrm{i} \end{bmatrix}, \quad \mathbf{u}_3 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{u}_4 = \frac{1}{2}\begin{bmatrix} 1 \\ \mathrm{i} \\ -1 \\ -\mathrm{i} \end{bmatrix}.$$

If we think of the standard basis $\mathbf{e}_j$ as a sampled version of a signal, then the signal is *localized* in time, since only one component of $\mathbf{e}_j$ is nonzero. By contrast, all the entries in $\mathbf{u}_j$ are nonzero, so the Fourier matrix removes the time localization.

Let $\mathbf{y} = [\,1,\,2,\,-1,\,0\,]^T$. Then

$$\mathbf{d} = \frac{1}{4}F_4\mathbf{y} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ (1-\mathrm{i})/2 \\ -1/2 \\ (1+\mathrm{i})/2 \end{bmatrix}.$$

In this case $\frac{1}{4}||\mathbf{y}||^2 = \frac{1}{4}[1 + 2^2 + (-1)^2] = 2/3$ and

$$||\mathbf{d}||^2 = \frac{1}{4}[1^2 + (1-\mathrm{i})(1+\mathrm{i}) + (-1)^2 + (1+\mathrm{i})(1-\mathrm{i})] = 2/3,$$

as predicted by Corollary 2.1.3. ■

2.2 Discrete Periodic Signals and Convolution

Consider a £nite digital signal $\mathbf{y}$ with N values, say $\mathbf{y}[0], \mathbf{y}[1], \ldots, \mathbf{y}[N-1]$. In the previous section we viewed $\mathbf{y}$ as a column vector

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[1] \\ \vdots \\ \mathbf{y}[N-1] \end{bmatrix} \in \mathbb{C}^N. \tag{2.10}$$

We will also think of digital signals as functions. A basic operation in signal processing is to take a *moving average* of the signal. For example, we can replace each value $\mathbf{y}[j]$ by the average of the values $\mathbf{y}[j-1]$ and $\mathbf{y}[j+1]$. This gives a new signal $\mathbf{z}$ with

$$\mathbf{z}[j] = \frac{1}{2}(\mathbf{y}[j-1] + \mathbf{y}[j+1]). \tag{2.11}$$

There is a bug in formula (2.11), however. To calculate $\mathbf{z}[0]$ or $\mathbf{z}[N-1]$ we need the values $\mathbf{y}[-1]$ and $\mathbf{y}[N]$, which aren't available. We will solve this problem by using the *periodic extension* of $\mathbf{y}$:

$$\mathbf{y}[j+kN] = \mathbf{y}[j] \quad \text{for } j = 0, 1, \ldots, N-1 \text{ and all integers } k \tag{2.12}$$

Thus we set $\mathbf{y}[-1] = \mathbf{y}[N-1]$ and $\mathbf{y}[N] = \mathbf{y}[0]$, since $-1 = N - 1 + N$ and $N = 0 + N$. In terms of *modular arithmetic*, we have $\mathbf{y}[m] = \mathbf{y}[j]$ when $m \equiv j \pmod{N}$. Now formula (2.11) is well-de£ned. It can be written in a more cumbersome case-by-case way as

$$\mathbf{z}[0] = \frac{1}{2}(\mathbf{y}[N-1] + \mathbf{y}[1]), \quad \mathbf{z}[N-1] = \frac{1}{2}(\mathbf{y}[N-2] + \mathbf{y}[0]),$$

and

$$\mathbf{z}[j] = \frac{1}{2}(\mathbf{y}[j-1] + \mathbf{y}[j+1]) \quad \text{for } j = 1, \ldots, N-2.$$

For example, if $\mathbf{y} = [\,1,\ 2,\ -1,\ 0\,]^T$ as in Example 2.1.4, then

$$\mathbf{z}[0] = (0+2)/2, \quad \mathbf{z}[1] = (1-1)/2, \quad \mathbf{z}[2] = (2+0)/2, \quad \mathbf{z}[3] = (-1+1)/2.$$

Deﬁne the *shift operator* S on periodic signals $\mathbf{y}$ of period N by

$$S\mathbf{y}[j] = \mathbf{y}[j-1] \quad \text{for } j = 0, 1, \ldots, N-1.$$

Here $S\mathbf{y}[0] = \mathbf{y}[N-1]$, since $\mathbf{y}$ is periodic. It is clear from the deﬁnition that S is linear and invertible:

$$S^{-1}\mathbf{y}[j] = \mathbf{y}[j+1].$$

We can write formula (2.11) as

$$\mathbf{z} = \frac{1}{2}(S + S^{-1})\mathbf{y}. \tag{2.13}$$

It follows that formula (2.11) has satisﬁes the following:

(linearity) The output signal $\mathbf{z}$ depends linearly on the input signal $\mathbf{y}$.

(shift invariance) If the input signal $\mathbf{y}$ is replaced by $S\mathbf{y}$, then the output signal $\mathbf{z}$ is also replaced by $S\mathbf{z}$.

We now show that every shift-invariant linear transformation C can be expressed as a linear combination of powers of the shift operator S. We ﬁrst observe that the property of shift-invariance for C is the same as

$$CS = SC. \tag{Shift Invariance}$$

In particular, any linear combination of powers of S is shift invariant. To prove the converse, we identify the periodic signals of period N with $\mathbb{C}^N$ by (2.10). Then S becomes a linear transformation of $\mathbb{C}^N$. We calculate its matrix relative to the standard basis of $\mathbb{C}^N$ as follows: Suppose the signal $\mathbf{y}$ corresponds to the standard basis vector $\mathbf{e}_k$. Then $\mathbf{y}[j] = 1$ if $j+1 = k$, and otherwise $\mathbf{y}[j] = 0$ (note the index shift by one). Since $S\mathbf{y}[j] = \mathbf{y}[j-1]$, we see that $S\mathbf{y}[j] = 1$ if $j = k$ and $S\mathbf{y}[j] = 0$ if $j \neq k$. This shows that

$$S\mathbf{e}_k = \mathbf{e}_{k+1} \quad \text{for } k = 1, 2, \ldots, N$$

(for this formula to be valid we must label the basis vectors circularly modulo N: $\mathbf{e}_{N+1} = \mathbf{e}_1$, $\mathbf{e}_{N+2} = \mathbf{e}_2$ and so on). We see that S acts as a *circular permutation* of the standard basis vectors.

Example 2.2.1. Suppose $N = 3$. Then $S\mathbf{e}_1 = \mathbf{e}_2$, $S\mathbf{e}_2 = \mathbf{e}_3$, and $S\mathbf{e}_3 = \mathbf{e}_1$, so the matrix of the shift operator S relative to the standard basis for $\mathbb{C}^3$ is

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

Notice that $S^2\mathbf{e}_1 = \mathbf{e}_3$, $S^2\mathbf{e}_2 = \mathbf{e}_1$, and $S^2\mathbf{e}_3 = \mathbf{e}_2$. Also $S^3 = I$. Thus

$$S^{-1} = S^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = S^T.$$

We have $S^{-1} = S^T$ since $\{S\mathbf{e}_1, S\mathbf{e}_2, S\mathbf{e}_3\}$ is an orthonormal basis for $\mathbb{C}^3$. ■

The general features of Example 2.2.1 are valid for the shift operator for any value of N. Namely, $S^N = I_N$ and $S^{-1} = S^{N-1}$. The matrix of S relative to the standard basis for $\mathbb{C}^N$ is real and orthogonal, so in matrix form $S^{-1} = S^T$.

Theorem 2.2.2. *Let S be the shift operator, viewed as an $N \times N$ matrix relative to the standard basis for $\mathbb{C}^N$. Suppose C is any shift-invariant linear transformation of N-periodic signals. View C as an $N \times N$ matrix relative to the standard basis for $\mathbb{C}^N$ and let the £rst column of C be $[c_0, c_1, \ldots, c_{N-1}]^T$. Then*

$$C = c_0 I + c_1 S + c_2 S^2 + \cdots + c_{N-1} S^{N-1}, \tag{2.14}$$

where I denotes the $N \times N$ identity matrix.

Proof. The £rst column of C is the vector $C\mathbf{e}_1$, so this vector can be written in terms of the standard basis as

$$C\mathbf{e}_1 = c_0\mathbf{e}_1 + c_1\mathbf{e}_2 + \cdots + c_{N-1}\mathbf{e}_N\,. \tag{2.15}$$

Now we calculate the columns $C\mathbf{e}_k$ of C for $k = 2, \ldots, N$. Since C is shift-invariant we have $S^{k-1}C = CS^{k-1}$. Thus if we multiply both sides of (2.15) by S^{k-1} and use the property $S^{k-1}\mathbf{e}_1 = \mathbf{e}_k$, we obtain

$$\begin{aligned} C\mathbf{e}_k &= CS^{k-1}\mathbf{e}_1 \\ &= S^{k-1}C\mathbf{e}_1 \\ &= c_0 S^{k-1}\mathbf{e}_1 + c_1 S^{k-1}\mathbf{e}_2 + c_2 S^{k-1}\mathbf{e}_3 + \cdots + c_{N-1}S^{k-1}\mathbf{e}_N \\ &= c_0\mathbf{e}_k + c_1 S\mathbf{e}_k + c_2 S^2\mathbf{e}_k + \cdots + c_{N-1}S^{N-1}\mathbf{e}_k\,. \end{aligned}$$

This shows that the kth column of the matrix C is the same as the kth column of the matrix $c_0 I + c_1 S + c_2 S^2 + \cdots + c_{N-1}S^{N-1}$ for $k = 1, \ldots, N$. Hence the two matrices are equal. ∎

Example 2.2.3. Suppose $N = 3$ and $C = c_0 I + c_1 S + c_2 S^2$ is a 3×3 shift-invariant matrix. From Example 2.2.1 we have

$$C = c_0 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + c_1 \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} c_0 & c_2 & c_1 \\ c_1 & c_0 & c_2 \\ c_2 & c_1 & c_0 \end{bmatrix}.$$

Hence the successive columns of C are obtained by circular permutation of the £rst column. Matrices of this form are called *circulant matrices*. For example, when $N = 4$ the averaging operation from (2.11) is given by the circulant matrix

$$C = \frac{1}{2}(S + S^{-1}) = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

∎

We now obtain the connection between shift-invariant linear transformations and the Fourier matrix. Let $F_N = [\,\mathbf{h}_0\ \mathbf{h}_1\ \cdots\ \mathbf{h}_{N-1}\,]$ be the $N \times N$ Fourier matrix with columns

$$\mathbf{h}_j = \begin{bmatrix} 1 \\ w^{-j} \\ w^{-2j} \\ \vdots \\ w^{-(N-1)j} \end{bmatrix}, \quad \text{where } w = \mathrm{e}^{2\pi\,\mathrm{i}/N}.$$

Since S shifts the entries in $\mathbf{h}_j$ down one place, with the last entry moved to the top, we have

$$S\mathbf{h}_j = \begin{bmatrix} w^{-(N-1)j} \\ 1 \\ w^{-j} \\ \vdots \\ w^{-(N-2)j} \end{bmatrix} = w^j \begin{bmatrix} w^{-Nj} \\ w^{-j} \\ w^{-2j} \\ \vdots \\ w^{-(N-1)j} \end{bmatrix} = w^j \mathbf{h}_j\,.$$

De£ne a diagonal matrix with the Nth roots of 1 on the diagonal:

$$D_N = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & w & 0 & \cdots & 0 \\ 0 & 0 & w^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & w^{N-1} \end{bmatrix} \tag{2.16}$$

The calculation just made shows that

$$SF_N = \begin{bmatrix} \mathbf{h}_0 & w\mathbf{h}_1 & w^2\mathbf{h}_2 & \cdots & w^{N-1}\mathbf{h}_{N-1} \end{bmatrix} = F_N D_N. \tag{2.17}$$

By Theorem 2.1.2 the Fourier matrix is invertible. Hence multiplying (2.17) on the left by F_N^{-1}, we obtain

$$F_N^{-1} S F_N = D_N. \tag{2.18}$$

We can summarize these calculations as follows:

Theorem 2.2.4. *The $N \times N$ shift matrix S is diagonalized by the Fourier matrix F_N. The columns of F_N are eigenvectors of S, and the eigenvalues of S are the N complex numbers w^j for $j = 0, 1, \ldots, N-1$ (the Nth roots of unity).*

Combining the last two theorems gives us the main result of this section:

Theorem 2.2.5 (Diagonalization of Circulant Matrices)**.** *Suppose that C is a $N \times N$ shift-invariant (circulant) matrix. Write*

$$C = c_0 I + c_1 S + c_2 S^2 + \cdots + c_{N-1} S^{N-1}$$

and de£ne the polynomial $p(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_{N-1} z^{N-1}$. Then $\mathbf{h}_j$ is an eigenvector for C, with eigenvalue $p(w^j)$, for $j = 0, 1, \ldots, N-1$. Hence C is diagonalized by the Fourier matrix:

$$F_N^{-1} C F_N = p(D_N) = \begin{bmatrix} p(1) & 0 & 0 & \cdots & 0 \\ 0 & p(w) & 0 & \cdots & 0 \\ 0 & 0 & p(w^2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p(w^{N-1}) \end{bmatrix} \tag{2.19}$$

Proof. Since (2.18) implies that $F_N^{-1} S^k F_N = D_N^k$ for all integers k, the matrix C satis£es the corresponding equation:

$$F_N^{-1} C F_N = c_0 I + c_1 D_N + c_2 D_N^2 + \cdots + c_{N-1} D_N^{N-1}.$$

The right side of this equation is $p(D_N)$. ■

Example 2.2.6. Consider the 4×4 circulant matrix $C = \frac{1}{2}(S + S^{-1}) = \frac{1}{2}(S + S^3)$ from Example 2.2.3 (note that $S^{-1} = S^3$ since $S^4 = I$). Then $p(z) = \frac{1}{2} z + \frac{1}{2} z^3$. Since the fourth roots of 1 are $1, \mathrm{i}, -1, -\mathrm{i}$, the eigenvalues of C are

$$p(1) = 1, \quad p(\mathrm{i}) = (1/2)(\mathrm{i} + \mathrm{i}^3) = 0,$$

$$p(-1) = (1/2)(-1 + (-1)^3) = -1, \quad p(-\mathrm{i}) = (1/2)(-\mathrm{i} + (-\mathrm{i})^3) = 0.$$

■

Now we return to the digital signal point of view. Let C be a linear shift-invariant operator on signals periodic of period N. Then by Theorem 2.2.2 there are complex numbers $c_0, \ldots, c_{N-1}$ so that

$$C = c_0 I + c_1 S + c_2 S^2 + \cdots + c_{N-1} S^{N-1}.$$

If we apply C to a periodic signal $\mathbf{y}$, then we get the signal

$$C\mathbf{y}[j] = c_0 \mathbf{y}[j] + c_1 \mathbf{y}[j-1] + c_2 \mathbf{y}[j-2] + \cdots + c_{N-1} \mathbf{y}[j-N+1] \tag{2.20}$$

for $j = 0, 1, \ldots, N-1$. This shows that $C\mathbf{y}$ is a *moving average* of the original signal $\mathbf{y}$, generalizing the special case of (2.11). De£ne the function $\mathbf{f}[k] = c_k$ for $k = 0, 1, \ldots, N-1$. Then (2.20) can be written as

$$C\mathbf{y}[j] = \sum_{k=0}^{N-1} \mathbf{f}[k] \mathbf{y}[j-k]. \tag{2.21}$$

We call the function $C\mathbf{y}$ the *convolution* (folding) of $\mathbf{f}$ and $\mathbf{y}$ and we write $C\mathbf{y} = \mathbf{f} * \mathbf{y}$. An alternate statement of Theorem 2.2.2 is the following:

(Linear Shift-Invariant Filters) Every linear transformation of N-periodic signals $\mathbf{y}$ that is shift invariant is given by the convolution (moving average) operation $\mathbf{y} \mapsto \mathbf{f} * \mathbf{y}$ for some function $\mathbf{f}$ on the set $\{0, 1, \ldots, N-1\}$ (the *£lter*).

We can now obtain the linear £lter version of Theorem 2.2.5.

De£nition 2.2.7 (Discrete Fourier Transform)**.** If $\mathbf{y}$ is a periodic digital signal (of period N), then the *Fourier transform* of $\mathbf{y}$ is the function

$$\hat{\mathbf{y}}[k] = \sum_{j=0}^{N-1} \mathbf{y}[j] w^{-jk} \quad \text{for } k = 0, 1, \dots, N-1,$$

where $w = \mathrm{e}^{2\pi\,\mathrm{i}/N}$ (note that the function $\hat{\mathbf{y}}$ is also periodic of period N). Thus if $\mathbf{y}$ is viewed as a column vector in $\mathbb{C}^N$, then $\hat{\mathbf{y}}$ is the column vector $F_N \mathbf{y}$. ■

The £lter $\mathbf{f}$ corresponding to the circulant matrix C in (2.14) is de£ned by $\mathbf{f}[k] = c_k$ for $k = 0, 1, \dots, N-1$. The matrix-vector product $C\mathbf{y}$ becomes the convolution $\mathbf{f} * \mathbf{y}$ in the signal-processing picture. We can restate the result of Theorem 2.2.4 in terms of the Fourier transform and convolution as follows:

Theorem 2.2.8 (Diagonalization of Convolution Operators)**.** *Let* $C\mathbf{y} = \mathbf{f} * \mathbf{y}$ *be the convolution operator* (2.21) *on signals* $\mathbf{y}$ *that are periodic of period* N*. Then the Fourier transform of* $C\mathbf{y}$ *is the pointwise product:*

$$\widehat{C\mathbf{y}}[k] = \hat{\mathbf{f}}[k]\, \hat{\mathbf{y}}[k] \quad \textit{for } k = 0, 1, \dots, N-1. \tag{2.22}$$

Proof. The columns of the Fourier matrix are eigenvectors for the circulant matrix C, and the eigenvalues are the scalars $\hat{\mathbf{f}}[k]$. Thus when a vector is expressed in terms of the Fourier basis, C acts on the kth component by multiplying by the eigenvalue $\hat{\mathbf{f}}[k]$.

We can give a direct proof of this result, without using Theorem 2.2.4, as follows. By de£nition of the £nite Fourier transform, we have

$$\widehat{\mathbf{f} * \mathbf{y}}[k] = \sum_{j=0}^{N-1} (\mathbf{f} * \mathbf{y})[j] w^{-jk} = \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{f}[j-l]\, \mathbf{y}[l] w^{-jk}$$

for $k = 0, 1, \dots, N-1$. Making the substitution $m = j - l$ and using the periodicity of $\mathbf{f}$ and $\mathbf{y}$ so that the range of summation is $0 \le m < N$ and $0 \le l < N$, we obtain

$$\begin{aligned} \widehat{\mathbf{f} * \mathbf{y}}[k] &= \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{f}[m]\, \mathbf{y}[l] w^{-(m+l)k} = \sum_{m=0}^{N-1} \mathbf{f}[m] w^{-mk} \sum_{l=0}^{N-1} \mathbf{y}[l] w^{-lk} \\ &= \hat{\mathbf{f}}[k]\, \hat{\mathbf{y}}[k] \end{aligned}$$

This proves (2.22). ■

Example 2.2.9. Consider the averaging operator (2.11) from the beginning of this Section:

$$C\mathbf{y}[j] = \frac{1}{2}(\mathbf{y}[j-1] + \mathbf{y}[j+1]),$$

where $\mathbf{y}$ is a periodic signal of length N. We can write this as $C\mathbf{y} = \mathbf{f} * \mathbf{y}$, where

$$\mathbf{f}[1] = 1/2, \quad \mathbf{f}[N-1] = 1/2, \quad \text{and } \mathbf{f}[j] = 0 \text{ for } j \ne 1, N-1,$$

since $C = (1/2)(S + S^{-1})$ as in Example 2.2.6. In this case the polynomial $p(z) = (1/2)(z + z^{-1})$ and $\hat{\mathbf{f}}[k] = (1/2)(w^k + w^{-k})$. Thus

$$\widehat{C\mathbf{y}}[k] = \frac{1}{2}(w^k + w^{-k}) \hat{\mathbf{y}}[k] \quad \text{for } k = 0, 1, \dots, N-1.$$

■

2.3 Fast Fourier Transform

The effectiveness of the discrete Fourier transform (DFT) as a computational tool depends on a remarkable *fast algorithm* for calculating the matrix-vector product $F_n\mathbf{v}$ when $n = 2^k$ is a power of 2 (similar fast algorithms exist for every *highly composite* number n, such as $n = 2^k 3^m$). The Fast Fourier Transform (FFT) algorithm is based on the observation that the Fourier matrix F_{2n} can be written as product of a permutation matrix (which has no arithmetic computational cost) and a 2×2 *block matrix*, where the blocks are F_n or a diagonal matrix multiplying F_n.

Example 2.3.1. Consider $n = 2$. Recall that

$$F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix}.$$

Let $\mathbf{y} \in \mathbb{C}^4$. By the de£nition of matrix-vector multiplication we can write

$$F_4\mathbf{y} = \mathbf{y}[0]\,\mathbf{h}_0 + \mathbf{y}[1]\,\mathbf{h}[1] + \mathbf{y}[2]\,\mathbf{h}[2] + \mathbf{y}_3\,\mathbf{h}[3]$$

as a linear combination of the columns of the Fourier matrix. Rearrange this sum according to the even and odd indices:

$$\mathbf{y}[0]\,\mathbf{h}_0 + \mathbf{y}[2]\,\mathbf{h}[2] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[2] \end{bmatrix}, \qquad \mathbf{y}[1]\,\mathbf{h}[1] + \mathbf{y}_3\,\mathbf{h}[3] = \begin{bmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \\ -1 & -1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} \mathbf{y}[1] \\ \mathbf{y}[3] \end{bmatrix}. \tag{2.23}$$

De£ne

$$\mathbf{y}_{\text{even}} = \begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[2] \end{bmatrix}, \qquad \mathbf{y}_{\text{odd}} = \begin{bmatrix} \mathbf{y}[1] \\ \mathbf{y}[3] \end{bmatrix}, \qquad \widetilde{D}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\mathrm{i} \end{bmatrix}.$$

Then, using block multiplication of matrices, we can write the formulas (2.23) as

$$\mathbf{y}[0]\,\mathbf{h}_0 + \mathbf{y}[2]\,\mathbf{h}[2] = \begin{bmatrix} F_2 \\ F_2 \end{bmatrix} \mathbf{y}_{\text{even}}\,, \qquad \mathbf{y}[1]\,\mathbf{h}[1] + \mathbf{y}_3\,\mathbf{h}[3] = \begin{bmatrix} \widetilde{D}_2 F_2 \\ -\widetilde{D}_2 F_2 \end{bmatrix} \mathbf{y}_{\text{odd}}\,.$$

The splitting of $\mathbf{y}$ into even/odd vectors of half length can be accomplished by the permuation matrix

$$P_4 = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_3 & \mathbf{e}_2 & \mathbf{e}_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad P_4\mathbf{y} = \begin{bmatrix} \mathbf{y}_{\text{even}} \\ \mathbf{y}_{\text{odd}} \end{bmatrix}.$$

The calculations above show that

$$F_4\mathbf{y} = \begin{bmatrix} F_2\mathbf{y}_{\text{even}} + \widetilde{D}_2 F_2 \mathbf{y}_{\text{odd}} \\ F_2\mathbf{y}_{\text{even}} - \widetilde{D}_2 F_2 \mathbf{y}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} F_2 & \widetilde{D}_2 F_2 \\ F_2 & -\widetilde{D}_2 F_2 \end{bmatrix} P_4\mathbf{y}.$$

Hence the 4×4 Fourier matrix F_4 can be written in 2×2 block form:

$$F_4 = \begin{bmatrix} F_2 & \widetilde{D}_2 F_2 \\ F_2 & -\widetilde{D}_2 F_2 \end{bmatrix} P_4.$$

■

The same splitting into even and odd components works for the DFT of a signal

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}[0] & \mathbf{y}[1] & \ldots & \mathbf{y}[2n-2] & \mathbf{y}[2n-1] \end{bmatrix}^{\mathrm{T}}$$

of length $2n$. Let

$$\mathbf{y}_{\text{even}} = \begin{bmatrix} \mathbf{y}[0] & \mathbf{y}[2] & \ldots & \mathbf{y}[2n-2] \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{y}_{\text{odd}} = \begin{bmatrix} \mathbf{y}[1] & \mathbf{y}[3] & \ldots & \mathbf{y}[2n-1] \end{bmatrix}^{\mathrm{T}}.$$

Here we are using the terms *even* and *odd* because we view $\mathbf{y}$ as a function on $\{0, 1, \ldots, 2n-1\}$; the vector $\mathbf{y}_{\text{even}}$ contains components $1, 3, \ldots, 2n-1$ of the vector $\mathbf{y}$ when we use the MATLAB indexing convention. The splitting of $\mathbf{y}$ into $\mathbf{y}_{\text{even}}$ and $\mathbf{y}_{\text{odd}}$ of half length is called *downsampling*; it will play an important role in wavelet analysis in Chapters 3 and 4.

Write $w = \mathrm{e}^{2\pi\,\mathrm{i}/2n} = \mathrm{e}^{\pi\,\mathrm{i}/n}$ and $z = w^2 = \mathrm{e}^{2\pi\,\mathrm{i}/n}$. Then

$$\begin{aligned} F_{2n}\mathbf{y}[j] &= \sum_{k=0}^{2n-1} w^{-jk}\mathbf{y}[k] \\ \text{(split into even-odd)} \qquad &= \sum_{k=0}^{n-1} w^{-j(2k)}\mathbf{y}[2k] + \sum_{k=0}^{n-1} w^{-j(2k+1)}\mathbf{y}[2k+1] \\ &= \sum_{k=0}^{n-1} z^{-jk}\mathbf{y}_{\text{even}}[k] + w^{-j}\sum_{k=0}^{n-1} z^{-jk}\mathbf{y}_{\text{odd}}[k] \end{aligned}$$

for $j = 0, 1, 2, \ldots, 2n-1$. This shows that

$$F_{2n}\mathbf{y}[j] = F_n\mathbf{y}_{\text{even}}[j] + w^{-j}F_n\mathbf{y}_{\text{odd}}[j] \quad \text{for } j = 0, 1, \ldots, n-1.$$

Since $w^n = -1$ and $z^n = 1$, we have $w^{-(n+j)} = -w^{-j}$ and $z^{-(n+j)k} = z^{-jk}$. Furthermore, the functions $F_n\mathbf{y}_{\text{even}}$ and $F_n\mathbf{y}_{\text{odd}}$ are periodic of period n. Thus

$$F_{2n}\mathbf{y}[n+j] = F_n\mathbf{y}_{\text{even}}[j] - w^{-j}F_n\mathbf{y}_{\text{odd}}[j] \quad \text{for } j = 0, 1, \ldots, n-1.$$

We can write these formulas in block-matrix form, just as in the case $n = 2$. Let

$$\widetilde{D}_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & w^{-1} & 0 & \cdots & 0 \\ 0 & 0 & w^{-2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & w^{-(n-1)} \end{bmatrix} \qquad \text{(caution: } w^n = -1\text{).} \tag{2.24}$$

Note that the diagonal of $\widetilde{D}_n$ only contains half of the $2n$th roots of 1; it is not the same as the matrix D_N in equation (2.16) which has all Nth roots of 1. Let P_{2n} be the permutation matrix that splits $\mathbf{y}$ into its even and odd components:

$$P_{2n} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_3 & \dots & \mathbf{e}_{2n-1} & \mathbf{e}_2 & \mathbf{e}_4 & \dots & \mathbf{e}_{2n} \end{bmatrix}^{\mathrm{T}}, \quad P_{2n}\mathbf{y} = \begin{bmatrix} \mathbf{y}_{\text{even}} \\ \mathbf{y}_{\text{odd}} \end{bmatrix}.$$

Then, just as in the case $n = 2$, the equations for $F_{2n}\mathbf{y}$ can be written as

$$F_{2n}\mathbf{y} = \begin{bmatrix} F_n\mathbf{y}_{\text{even}} + \widetilde{D}_n F_n \mathbf{y}_{\text{odd}} \\ F_n\mathbf{y}_{\text{even}} - \widetilde{D}_n F_n \mathbf{y}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} F_n & \widetilde{D}_n F_n \\ F_n & -\widetilde{D}_n F_n \end{bmatrix} P_{2n}\mathbf{y} \tag{2.25}$$

Note: Equation (2.25) is incorrectly stated on page 285 of Leon, where the factor P_{2n} is on the wrong side of the equation–for the case $n = 2$ that we considered above it happens that $P_4 = P_4^{\mathrm{T}} = P_4^{-1}$, so the formula in Leon is correct in that case. For larger values of n the matrix P_{2n} is orthogonal but not symmetric, and hence $P_{2n}^{-1} = P_{2n}^{\mathrm{T}} \neq P_{2n}$. Also, the matrix D_m in Leon's formula should have jth diagonal entry ω_{2m}^{j-1} rather than ω_m^{j-1} (using Leon's notation).

The *Fast Fourier Transform* algorithm calculates F_N when N is a power of 2 by iterating formula (2.25). For example, when $N = 256 = 2^8$ then (2.25) expresses $F_{256}\mathbf{y}$ in terms of F_{128} applied to signals of length 128. To calculate these Fourier transforms, we use (2.25) again to express F_{128} in terms of F_{64} applied to signals of length 64, and so on until we are down to F_2.

To determine the computational cost of the FFT algorithm, let $n = 2^k$, and de£ne $\phi(k)$ be the number of scalar multiplications needed to evaluate $F_n\mathbf{y}$ for a signal of length $n = 2^k$ using the FFT algorithm. When $k = 1$ then the entries in F_2 are ± 1, so no multiplications are needed (just sign changes). Hence $\phi(1) = 0$. If $\mathbf{y}$ is a signal of length $2n = 2^{k+1}$, then calculating $F_{2n}\mathbf{y}$ using (2.25) requires $2\phi(k)$ multiplications to obtain $F_n\mathbf{y}_{\text{even}}$ and $F_n\mathbf{y}_{\text{odd}}$, followed by 2^k multiplications to obtain $\widetilde{D}_n F_n \mathbf{y}_{\text{odd}}$. We are using the fact that $\widetilde{D}_n$ is a diagonal matrix, so it only requires n multiplications to calculate $\widetilde{D}_n\mathbf{b}$ for any vector $\mathbf{b}$. The matrix P_{2n} just sorts the entries of $\mathbf{y}$; no arithmetic is needed to calculate $P_{2n}\mathbf{y}$. Thus

$$\phi(k+1) = 2\phi(k) + 2^k \tag{2.26}$$

We can calculate $\phi(k)$ recursively from (2.26), starting with $\phi(1) = 0$:

$$\phi(2) = 2\phi(1) + 2 = 2, \quad \phi(3) = 2\phi(2) + 2^2 = 2 \cdot 2^2, \quad \phi(4) = 2\phi(3) + 2^3 = 3 \cdot 2^3$$

This suggests that

$$\phi(k) = (k-1)2^{k-1} \quad \text{for all positive integers } k = 1, 2, 3, \dots. \tag{2.27}$$

This formula, which we have just shown true for $k = 2$, 3, and 4, is easily veri£ed by induction: assuming it true for k and using (2.26), we get

$$\phi(k+1) = 2(k-1)2^{k-1} + 2^k = k2^k - 2^k + 2^k = k2^k,$$

so the formula is true for $k + 1$.

To appreciate the consequences of (2.27), note that direct evaluation of $F_n\mathbf{y}$ as a matrix-vector product requires $n^2 = 2^{2k}$ scalar multiplications (n for each of the n components of $\mathbf{y}$). Take

$k = 10$ and $n = 2^{10} = 1024$. Then direct evaluation of $F_n\mathbf{y}$ as a matrix-vector product requires $n^2 = 2^{20} = 1,048,576$ multiplications, whereas evaluation using the FFT only requires $9 \cdot 2^9 = 4608$ multiplications. This is a speedup by a factor of

$$\frac{2^{20}}{9 \cdot 2^9} = 228.$$

If we go to longer signals, such as $n = 2^{20} = 1,048,576$, then the speedup is by a factor of

$$\frac{2^{40}}{19 \cdot 2^{19}} = 110,376$$

(more than one hundred thousand times faster). The same sort of counting of the number of scalar addition operations needed in the FFT shows a similar dramatic improvement over calculations using the standard matrix-vector product. Without the FFT algorithm digital signal processing would be impractical.

2.4 Exercises

1. Let N be a positive integer and set $w = \mathrm{e}^{2\pi\,\mathrm{i}/N}$. View the columns of the Fourier matrix F_N as the functions $\mathbf{h}_0, \ldots, \mathbf{h}_{N-1}$ de£ned by $\mathbf{h}_k[j] = w^{-jk}$ (note that with this de£nition $\mathbf{h}_k$ is automatically periodic of period N.) Verify directly that each function $\mathbf{h}_k$ is an eigenfunction for the shift operator S, and determine the eigenvalue. Recall that S acts on a periodic function $\mathbf{f}$ by $S\,\mathbf{f}\,[j] = \mathbf{f}\,[j-1]$.

2. Let C be the linear shift-invariant transformation $C\mathbf{y}[j] = \mathbf{y}[j-1] - 2\mathbf{y}[j] + \mathbf{y}[j+1]$ for $\mathbf{y}$ a function periodic of period N.

   (a) Find the function $\mathbf{f}$ such that $C\mathbf{y} = \mathbf{f} * \mathbf{y}$. (HINT: Write C in terms of the shift operator.)

   (b) Find the Fourier transform of the function $\mathbf{f}$ in (a).

   (c) Suppose $N = 4$ and $\mathbf{y}$ corresponds to the vector $\mathbf{y} = [2\,3\,1\,5]^{\mathrm{T}} \in \mathbb{C}^4$. Calculate the vectors corresponding to $C\mathbf{y}$ and $\widehat{C\mathbf{y}}$.

3. Let $S = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ be the matrix for the shift operator relative to the standard basis for $\mathbb{C}^3$. Suppose the matrix $C = \begin{bmatrix} 4 & * & * \\ 7 & * & * \\ 5 & * & * \end{bmatrix}$ satis£es $CS = SC$.

   (a) Write C as a polynomial in S. Use this to £ll in the missing entries in C:

$$C = \begin{bmatrix} 4 & \_\_\_\_ & \_\_\_\_ \\ 7 & \_\_\_\_ & \_\_\_\_ \\ 5 & \_\_\_\_ & \_\_\_\_ \end{bmatrix}$$

(b) View vectors in $\mathbb{C}^3$ as periodic functions $\mathbf{y}$ on the integers: $\mathbf{y}[j] = \mathbf{y}[j+3]$ for all integers j. Let T be the linear transformation on such functions corrsponding to the matrix C above. Give explicit formulas (in terms of $\mathbf{y}[0]$, $\mathbf{y}[1]$, and $\mathbf{y}[2]$) for $T\mathbf{y}[j]$ for $j = 0, 1, 2$.

(c) Let F be the 3×3 Fourier matrix, and let $w = \mathrm{e}^{2\pi\,\mathrm{i}/3}$. Let C be the matrix in part **(a)**. Find complex numbers λ_0, λ_1, and λ_2 so that $F^{-1}CF = \begin{bmatrix} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}$.

Express your answer in terms of w and w^2 (no complex arithmetic is needed).

4. Let $n = 2^k$. De£ne $\psi(k)$ to be the number of scalar additions (or subtractions) needed to calculate $F_n\mathbf{c}$ by the Fast Fourier Transform (FFT) algorithm. Note that the product of a row vector and a column vector, each with n components, needs $n - 1$ additions.

   (a) Show that $\psi(1) = 2$ and that $\psi(k + 1) = 2\psi(k) + 2(2^k - 1)$.

   (b) Use the recursion in (a) to calculate $\psi(k)$ for $k = 2, 3, 4$.

   (c) Prove by induction that $\psi(k) \leq k2^k$ for all positive integers k.

   (d) Use the result in the notes and (c) to show that the total number of arithmetic operations (multiplications and additions) required for the FFT on vectors of size 2^k is less than $(3/2)k2^k$.

Chapter 3

Finite Wavelet Transforms

3.1 Prediction and Update Transforms

Example 3.1.1. Consider the following digital signal $\mathbf{s}_3[n]$ with 2^3 nonzero values (see Table 2.1 in Ripples):

n :	0	1	2	3	4	5	6	7
$\mathbf{s}_3[n]$:	56	40	8	24	48	48	40	16

There are two features of the signal that we want to analyze:

1. The overall *trend* of the signal. Notice that $\mathbf{s}_3[4] = \mathbf{s}_3[5]$, so there is no change in the signal when n goes from 4 to 5. Thus the value of $\mathbf{s}_3$ at $n = 4$ is a good predictor of the value at $n = 5$.

2. The *detail* in the signal. Notice the big change in the signal when n goes from 6 to 7. The ¤uctuations in the signal can be measured by the error when we predict the signal values for odd n based on the values for even n.

To carry out this analysis we *split* the signal into two half-length signals

$$(\mathbf{s}_3)_{\text{even}}[n] = \mathbf{s}_3[2n] \quad \text{and} \quad (\mathbf{s}_3)_{\text{odd}}[n] = \mathbf{s}_3[2n+1]$$

by taking every other value of $\mathbf{s}_3$ (this is also called *downsampling*). We then calculate the *level-two trend*:

$$\mathbf{s}_2[n] = \frac{1}{2}\Big\{(\mathbf{s}_3)_{\text{even}}[n] + (\mathbf{s}_3)_{\text{odd}}[n]\Big\}$$

(the average of the values of $\mathbf{s}_3$ at $2n$ and $2n+1$) and the *level-two detail*:

$$\mathbf{d}_2[n] = (\mathbf{s}_3)_{\text{even}}[n] - \mathbf{s}_2[n]$$

(the difference between the value of $\mathbf{s}_3$ at $2n$ and the average of $\mathbf{s}_3$ at $2n$ and $2n+1$). These signals of length 2^2 are given in the table below:

n :	0	1	2	3	n :	0	1	2	3
$(\mathbf{s}_3)_{\text{even}}[n]$:	56	8	48	40	$(\mathbf{s}_3)_{\text{odd}}[n]$:	40	24	48	16
$\mathbf{s}_2[n]$:	48	16	48	28	$\mathbf{d}_2[n]$:	8	-8	0	12

Notice the value $\mathbf{d}_2[2] = 0$ corresponding to $\mathrm{s}_3[4] = \mathrm{s}_3[5]$ and the large value $\mathbf{d}_2[3]$ corresponding to the big change between $\mathrm{s}_3[5]$ and $\mathrm{s}_3[7]$.

We can repeat this analysis on the trend vector s_2: De£ne the *level-one trend*:

$$\mathrm{s}_1[n] = \frac{1}{2}\Big\{(\mathrm{s}_2)_{\text{even}}[n] + (\mathrm{s}_2)_{\text{odd}}[n]\Big\}$$

and the *level-one detail*:

$$\mathbf{d}_1[n] = (\mathrm{s}_2)_{\text{even}}[n] - \mathrm{s}_1[n].$$

These signals of length 2^1 are given in the table below:

n :	0	1	n :	0	1
$(\mathrm{s}_2)_{\text{even}}[n]$:	48	48	$(\mathrm{s}_2)_{\text{odd}}[n]$:	16	28
$\mathrm{s}_1[n]$:	32	38	$\mathbf{d}_1[n]$:	16	10

Finally, we repeat this analysis on the trend vector s_1: De£ne the *level-zero trend*:

$$\mathrm{s}_0[n] = \frac{1}{2}\Big\{(\mathrm{s}_1)_{\text{even}}[n] + (\mathrm{s}_1)_{\text{odd}}[n]\Big\}$$

and the *level-zero detail*:

$$\mathbf{d}_0[n] = (\mathrm{s}_1)_{\text{even}}[n] - \mathrm{s}_0[n].$$

These signals of length 2^0 are given in the table below:

n :	0	n :	0
$(\mathrm{s}_1)_{\text{even}}[n]$:	32	$(\mathrm{s}_1)_{\text{odd}}[n]$:	38
$\mathrm{s}_0[n]$:	35	$\mathbf{d}_0[n]$:	-3

We can assemble this three-level decomposition into the following *multiresolution analysis* of the original signal:

s_0	$\mathbf{d}_0$	$\mathbf{d}_1$		$\mathbf{d}_2$			
35	-3	16	10	8	-8	0	12

(see Ripples, Table 2.1). We shall explain the utility of this decomposition in Section 3.2. ■

In the multiresolution analysis in Example 3.1.1 the £rst entry $\mathrm{s}_0[0]$ is the *average* of the original signal:

$$35 = (56 + 40 + 8 + 24 + 48 + 48 + 40 + 16)/8.$$

This is easy to check:

$$\mathrm{s}_0[0] = \frac{1}{2}\sum_{n=0}^{1}\mathrm{s}_1[n] = \frac{1}{4}\sum_{n=0}^{3}\mathrm{s}_2[n] = \frac{1}{8}\sum_{n=0}^{7}\mathrm{s}_3[n].$$

We now introduce some linear algebra to understand the rest of this decomposition and its purpose. Suppose $\mathbf{x}$ is a real-valued signal of length $N = 2k$ with values $\mathbf{x}[0], \mathbf{x}[1], \ldots, \mathbf{x}[N-1]$. We identify $\mathbf{x}$ with the $N \times 1$ column vector with these components (notice that we are indexing the components from 0 to $N-1$, whereas in MATLAB the indexing would go from 1 to N). Following the same pattern as for the fast Fourier transform, we split (downsample) $\mathbf{x}$ into even and odd signals of length $N/2$:

$$\mathbf{x}_{\text{even}}[n] = \mathbf{x}[2n] \quad \text{and} \quad \mathbf{x}_{\text{odd}}[n] = \mathbf{x}[2n+1] \quad \text{for } n = 0, 1, \ldots, N/2 - 1$$

These formulas de£ne a linear transformation that we denote by $\boxed{\mathsf{split}}$:

$$\boxed{\mathsf{split}}\,\mathbf{x} = \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix}$$

When we make $\mathbf{x}$ into a column vector the transformation $\boxed{\mathsf{split}}$ is given by the permutation matrix P_N in Section 2.3. We de£ne $\boxed{\mathsf{merge}}$ to be the inverse permutation matrix (the transpose of $\boxed{\mathsf{split}}$):

$$\boxed{\mathsf{merge}} \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \mathbf{x}$$

De£ne the *trend* vector $\mathbf{s}$ and *detail* vector $\mathbf{d}$ of $\mathbf{x}$ just as in Example 3.1.1 by

$$\mathbf{s} = \frac{1}{2}\Big(\mathbf{x}_{\text{even}} + \mathbf{x}_{\text{odd}}\Big) \quad \text{and} \quad \mathbf{d} = \mathbf{x}_{\text{even}} - \mathbf{s}$$

(notice that $\mathbf{s}$ and $\mathbf{d}$ are vectors with $N/2$ components). Let $\mathbf{T_a}$ be the linear transformation on $\mathbb{R}^N$ given by

$$\mathbf{T_a}\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix}$$

(the subscript $\mathbf{a}$ is for *analysis*). We can rewrite the formulas for $\mathbf{s}$ and $\mathbf{d}$ as

$$2\mathbf{d} = \mathbf{x}_{\text{even}} - \mathbf{x}_{\text{odd}} \quad \text{and} \quad \mathbf{s} = \mathbf{x}_{\text{even}} - \mathbf{d}$$

Hence we can factor $\mathbf{T_a}$ as a product of $\boxed{\mathsf{split}}$ and the following *elementary linear transformations* (where I denotes the $N/2 \times N/2$ identity matrix):

$$\begin{bmatrix} \mathbf{x}_{\text{even}} \\ -2\mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} - \mathbf{x}_{\text{even}} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} \qquad \text{(Prediction)}$$

$$\begin{bmatrix} \mathbf{s} \\ -2\mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{\text{even}} - \mathbf{d} \\ -2\mathbf{d} \end{bmatrix} = \begin{bmatrix} I & \frac{1}{2}I \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\text{even}} \\ -2\mathbf{d} \end{bmatrix} \qquad \text{(Update)}$$

$$\begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ \frac{1}{2}(2\mathbf{d}) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \frac{1}{2}I \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ 2\mathbf{d} \end{bmatrix} \qquad \text{(Normalization)}$$

We de£ne $N \times N$ matrices P, U, and D (in 2×2 block form) by

$$P = \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix}, \quad U = \begin{bmatrix} I & \frac{1}{2}I \\ 0 & I \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} I & 0 \\ 0 & -\frac{1}{2}I \end{bmatrix}.$$

Then

$$\mathbf{T_a}\mathbf{x} = DUP\,\boxed{\mathsf{split}}\,\mathbf{x}$$

(notice the order of multiplication of the factors). Thus we have factored $\mathbf{T}_a$ as the product of elementary matrices (the matrices for elementary row operations) to obtain a *lifting block* (see Fig. 3.1 in Ripples). The important point about elementary matrices is that they are invertible and the inverses are also elementary matrices:

$$P^{-1} = \begin{bmatrix} I & 0 \\ I & I \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} I & -\frac{1}{2}I \\ 0 & I \end{bmatrix}, \quad \text{and} \quad D^{-1} = \begin{bmatrix} I & 0 \\ 0 & -2I \end{bmatrix}.$$

Hence the factorization shows that $\mathbf{T_a}$ is invertible with inverse

$$\mathbf{T_a}^{-1} = \boxed{\text{merge}}\, P^{-1}U^{-1}D^{-1}.$$

(see Fig. 3.4 in *Ripples*). We call the inverse matrix the *synthesis matrix* and write $\mathbf{T_s} = \mathbf{T_a}^{-1}$. For any signal $\mathbf{x}$ of length N, we set $\mathbf{y} = \mathbf{T_a}\mathbf{x}$. Then we can reconstruct $\mathbf{x}$ from $\mathbf{y}$ by $\mathbf{x} = \mathbf{T_s}\mathbf{y}$.

Example 3.1.2. Using block multiplication of matrices, we calculate

$$\mathbf{T_a} = DUP\,\boxed{\text{split}} = \frac{1}{2}\begin{bmatrix} I & I \\ I & -I \end{bmatrix}\boxed{\text{split}}$$

and

$$\mathbf{T_s} = \boxed{\text{merge}}\, P^{-1}U^{-1}D^{-1} = \boxed{\text{merge}}\begin{bmatrix} I & I \\ I & -I \end{bmatrix}$$

For example, if $N = 4$, then

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

and

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

The calculation of $\mathbf{s}_1$ and $\mathbf{d}_1$ from $\mathbf{s}_2$ in Example 3.1.1 can then be obtained as the matrix-vector product

$$\begin{bmatrix} \mathbf{s}_1 \\ \mathbf{d}_1 \end{bmatrix} = \mathbf{T_a}\mathbf{s}_2 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} 48 \\ 16 \\ 48 \\ 28 \end{bmatrix} = \begin{bmatrix} 32 \\ 38 \\ 16 \\ 10 \end{bmatrix}$$

In the opposite direction, we can obtain $\mathbf{s}_2$ from $\mathbf{s}_1$ and $\mathbf{d}_1$ by

$$\mathbf{s}_2 = \mathbf{T_s}\begin{bmatrix} \mathbf{s}_1 \\ \mathbf{d}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 32 \\ 38 \\ 16 \\ 10 \end{bmatrix} = \begin{bmatrix} 48 \\ 16 \\ 48 \\ 28 \end{bmatrix}$$

The matrices $\mathbf{T_a}$ and $\mathbf{T_s}$ are called the one-scale Haar analysis and synthesis matrices (named after the Hungarian mathematician A. Haar). There is an important pattern to observe in these matrices: rows 2 and 4 of $\mathbf{T_a}$ are obtained from rows 1 and 3 by shifting to the right by two positions. Likewise, columns 2 and 4 of $\mathbf{T_s}$ are obtained from columns 1 and 3 by shifting down by two positions. We will study this pattern in more detail later in connection with other wavelet transforms. Notice also that $2\mathbf{T_a}^{\mathrm{T}} = \mathbf{T_s}$, and hence

$$(\sqrt{2}\mathbf{T_a})(\sqrt{2}\mathbf{T_a})^{\mathrm{T}} = \mathbf{T_s}\mathbf{T_a} = I.$$

Hence the *normalized Haar transform* $\sqrt{2}\mathbf{T_a}$ is an orthogonal matrix. ■

3.2 Multiple Scale Wavelet Transforms

In Example 3.1.1 we took a signal $\mathbf{s}_3$ of length 2^3 and transformed it into a *trend* $\mathbf{s}_2$ and a *detail* $\mathbf{d}_2$, each of half the length. We then repeated this operation on the trend portion. We can describe this algorithm in terms of the analysis matrices $\mathbf{T_a}$ introduced in Section 3.1, as follows:

Write $\mathbf{T_a^{(k)}}$ for the $2^k \times 2^k$ Haar analysis matrix:

$$\mathbf{T_a^{(k)}} = \frac{1}{2}\begin{bmatrix} I^{(k-1)} & I^{(k-1)} \\ I^{(k-1)} & -I^{(k-1)} \end{bmatrix} \boxed{\text{split}}.$$

Here we use the notation $I^{(p)}$ for the identity matrix of size $2^p \times 2^p$. The calculations in Example 3.1.1 can be written in block-matrix form as

$$\begin{bmatrix} \mathbf{s}_2 \\ \mathbf{d}_2 \end{bmatrix} = \mathbf{T_a^{(3)}}\mathbf{s}_3, \qquad \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{T_a^{(2)}} & 0 \\ 0 & I^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{s}_2 \\ \mathbf{d}_2 \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{s}_0 \\ \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{T_a^{(1)}} & 0 & 0 \\ 0 & I^{(1)} & 0 \\ 0 & 0 & I^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}.$$

In these formulas 0 denotes matrices of zeros of the appropriate sizes. We can describe the three-scale analysis transformation by the following diagram (see also Fig. 3.7 in Ripples):

$$\begin{array}{ccccccc}
 & & & & & & \longrightarrow \mathbf{s}_0 \\
 & & & & \xrightarrow{\mathbf{s}_1} & \boxed{\mathbf{T_a^{(1)}}} & \\
 & & \xrightarrow{\mathbf{s}_2} & \boxed{\mathbf{T_a^{(2)}}} & & & \longrightarrow \mathbf{d}_0 \\
\mathbf{x} \longrightarrow & \boxed{\mathbf{T_a^{(3)}}} & & & \to \mathbf{d}_1 & & \\
 & & \to \mathbf{d}_2 & & & &
\end{array}$$

We can combine the transformations to obtain the multiresolution analysis of $\mathbf{s}_3$ in matrix form:

$$\begin{bmatrix} \mathbf{s}_0 \\ \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \mathbf{W_a^{(3)}}\mathbf{s}_3, \quad \text{where} \quad \mathbf{W_a^{(3)}} = \begin{bmatrix} \mathbf{T_a^{(1)}} & 0 & 0 \\ 0 & I^{(1)} & 0 \\ 0 & 0 & I^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{T_a^{(2)}} & 0 \\ 0 & I^{(2)} \end{bmatrix} \mathbf{T_a^{(3)}} \tag{3.1}$$

(in equation (3.1) the matrix blocks are of different sizes, so the matrix multiplication cannot be simplified while still in block form). We call $\mathbf{W}_{\mathbf{a}}^{(3)}$ the *three-scale Haar wavelet analysis matrix* (see equation (5.2) in Ripples for the explicit form of this matrix).

The *three-scale Haar wavelet synthesis matrix* $\mathbf{W}_{\mathbf{s}}^{(3)}$ is the inverse of the three-scale analysis matrix. Since the one-scale synthesis matrices are the inverses of the one-scale analysis matrices, the factorization (3.1) implies that

$$\mathbf{W}_{\mathbf{s}}^{(3)} \begin{bmatrix} \mathbf{s}_0 \\ \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \mathbf{s}_3, \quad \text{where} \quad \mathbf{W}_{\mathbf{s}}^{(3)} = \mathbf{T}_{\mathbf{s}}^{(3)} \begin{bmatrix} \mathbf{T}_{\mathbf{s}}^{(2)} & 0 \\ 0 & I^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{\mathbf{s}}^{(1)} & 0 & 0 \\ 0 & I^{(1)} & 0 \\ 0 & 0 & I^{(2)} \end{bmatrix} \tag{3.2}$$

(see equation (5.1) in Ripples for the explicit form of this matrix). We can describe the three-scale synthesis transformation by the following diagram (see also Fig. 3.8 in Ripples):

$$\begin{array}{ccccccccc}
\mathbf{s}_0 \longrightarrow & & & & & & & & \\
 & \boxed{\mathbf{T}_{\mathbf{s}}^{(1)}} & \xrightarrow{\mathbf{s}_1} & & & & & & \\
\mathbf{d}_0 \longrightarrow & & & \boxed{\mathbf{T}_{\mathbf{s}}^{(2)}} & \xrightarrow{\mathbf{s}_2} & & & & \\
 & & \mathbf{d}_1 \rightarrow & & & \boxed{\mathbf{T}_{\mathbf{s}}^{(3)}} & \longrightarrow \mathbf{x} & & \\
 & & & & \mathbf{d}_2 \rightarrow & & & &
\end{array}$$

The matrix $\mathbf{W}_{\mathbf{s}}^{(3)} = [\mathbf{h}_0 \cdots \mathbf{h}_7]$ has the following special structure:

$$\mathbf{h}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{h}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad \mathbf{h}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{h}_4 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

and

$$\mathbf{h}_3 = S^4\mathbf{h}_1, \quad \mathbf{h}_5 = S^2\mathbf{h}_4, \quad \mathbf{h}_6 = S^4\mathbf{h}_4, \quad \mathbf{h}_7 = S^6\mathbf{h}_4,$$

where S is the 8×8 shift matrix (we are enumerating the columns by 0 to 7, just as we did for the Fourier matrix). The vector $\mathbf{h}_0$ describes a signal that is constantly 1 (the *DC* component). The vector $\mathbf{h}_1$ describes a signal that is 1 for four time units, then switches sign and is -1 for four time units (slow *AC* component). The vector $\mathbf{h}_2$ describes a signal that is 1 for two time units, switches sign and is -1 for two time units, and then is zero (faster *AC* component). The vector $\mathbf{h}_4$ describes a signal that is 1 for one time unit, switches sign and is -1 for one time unit, and then is zero (fastest *AC* component). The other vectors are shifts of these vectors by an even number of positions (see Fig. 5.1 in *Ripples*).

To obtain the *multiresolution representation* of the original signal, recall that the product of a matrix $\mathbf{A}$ and a vector $\mathbf{b}$ is the linear combination of the columns of $\mathbf{A}$ with coefficients the entries of $\mathbf{b}$. Thus

$$\mathbf{s}_3 = \mathbf{W}_{\mathbf{s}}^{(3)} \begin{bmatrix} \mathbf{s}_0 \\ \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} \begin{aligned} &= \mathbf{s}_0[0]\mathbf{h}_0 + \mathbf{d}_0[0]\mathbf{h}_1 + \{\mathbf{d}_1[0]\mathbf{h}_2 + \mathbf{d}_1[1]\mathbf{h}_3\} \\ &\quad + \{\mathbf{d}_2[0]\mathbf{h}_4 + \mathbf{d}_2[1]\mathbf{h}_5 + \mathbf{d}_2[2]\mathbf{h}_6 + \mathbf{d}_2[3]\mathbf{h}_7\}. \end{aligned}$$

In this formula we build up the signal by taking the overall average $\mathbf{s}_0[0]\mathbf{h}_0$, then adding the slow ¤uctuation term $\mathbf{d}_0[0]\mathbf{h}_1$, followed by the faster and shorter ¤uctuations $\{\mathbf{d}_1[0]\mathbf{h}_2 + \mathbf{d}_1[1]\mathbf{h}_3\}$, and £nally the fastest and shortest ¤uctuations $\{\mathbf{d}_2[0]\mathbf{h}_4 + \mathbf{d}_2[1]\mathbf{h}_5 + \mathbf{d}_2[2]\mathbf{h}_6 + \mathbf{d}_2[3]\mathbf{h}_7\}$. For the signal in Example 3.1.1 the multiresolution representation is thus

$$\mathbf{s}_3 = 35\mathbf{h}_0 - 3\mathbf{h}_1 + \{16\mathbf{h}_2 + 10\mathbf{h}_3\} + \{8\mathbf{h}_4 - 8\mathbf{h}_5 + 0\mathbf{h}_6 + 12\mathbf{h}_7\} \tag{3.3}$$

(see Table 2.1 in *Ripples*).

One of the main applications of the multiresolution representation is *compression*. We choose a *threshold* ϵ, and set to zero all coef£cients in the multiresolution representation whose absolute value is less than ϵ. For example, if we apply this procedure with $\epsilon = 4$ to (3.3), we get the modi£ed signal

$$\mathbf{y} = 35\mathbf{h}_0 + \{16\mathbf{h}_2 + 10\mathbf{h}_3\} + \{8\mathbf{h}_4 - 8\mathbf{h}_5 + 12\mathbf{h}_7\}$$

(see Fig. 2.1 and Table 2.2 of Ripples). Notice that the graphs of the modi£ed signal and the original signal are almost the same. To measure the relative difference between the two graphs, we use the ratios of the energies (the square of the norms):

$$\text{relative compression error} = \frac{||\mathbf{s}_3 - \mathbf{y}||^2}{||\mathbf{s}_3||^2}$$

For compression with $\epsilon = 4$ every entry in $\mathbf{y}$ happens to differ from the corresponding entry in $\mathbf{s}_3$ by ± 3, so the relative compression error is

$$\frac{3^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2}{56^2 + 40^2 + 8^2 + 24^2 + 48^2 + 48^2 + 40^2 + 16^2} = 0.0061 = 0.6\%$$

If we apply this procedure with $\epsilon = 9$ to (3.3), we get the we get the modi£ed signal

$$\mathbf{z} = 35\mathbf{h}_0 + 16\mathbf{h}_2 + 10\mathbf{h}_3 + 12\mathbf{h}_7.$$

(see Fig. 2.2 and Table 2.3 of Ripples). Now the graphs of the modi£ed signal and the original signal have the same general shape, but differ in details. The relative compression error is

$$\frac{5^2 + 11^2 + 11^2 + 5^2 + 3^2 + 3^2 + 3^2 + 3^2}{56^2 + 40^2 + 8^2 + 24^2 + 48^2 + 48^2 + 40^2 + 16^2} = 0.028 = 2.8\%.$$

This is about £ve times larger than the compression error with $\epsilon = 4$. On the other hand, the signal $\mathbf{z}$ only has four nonzero coef£cients in its multiresolution representation, so we have compressed the original signal by a ratio of 2 : 1 while still retaining its essential features.

3.3 Discrete Wavelet Transform via Lifting

The general approach to discrete wavelet transforms that we are using is to split a signal of even length N into even and odd parts, apply successive *prediction* and *update* linear transformations to the two parts of the signal, and £nally apply a normalization; this procedure is called *lifting* (see Ripples, Section 3.2).

CDF(2,2) Transform

In our £rst example of a discrete Wavelet transform of a signal $\mathbf{x}$ (the Haar Transform–Example 3.1.1), we predicted the value $\mathbf{x}[2n+1]$ by the previous value $\mathbf{x}[2n]$. Now we modify this scheme and predict $\mathbf{x}[2n+1]$ to be the *average* of the neighboring values $\mathbf{x}[2n]$ and $\mathbf{x}[2n+2]$. The deviation $\mathbf{d}$ from the predicted value is thus

$$\mathbf{d}[n] = \mathbf{x}[2n+1] - \frac{1}{2}\big(\mathbf{x}[2n] + \mathbf{x}[2n+2]\big) = \mathbf{x}_{\text{odd}}[n] - \frac{1}{2}\big(\mathbf{x}_{\text{even}}[n] + \mathbf{x}_{\text{even}}[n+1]\big)$$

(see Fig. 3.3 in Ripples). Notice that if $\mathbf{x}[n] = an + b$ is a linear function of n, then the deviation $\mathbf{d}$ is zero.

There is a problem with the formula for $\mathbf{d}$, however. The signal values are $\mathbf{x}[0]$, $\mathbf{x}[1]$, ..., $\mathbf{x}[N-1]$, and to calculate the last deviation $\mathbf{d}[N/2]$ we need the value $\mathbf{x}[N]$. We shall avoid this problem by extending the signal $\mathbf{x}$ to be *periodic* of period N:

$$\mathbf{x}[k+N] = \mathbf{x}[k] \quad \text{for all integers } k.$$

(There are other solutions to this problem that are discussed in Ch. 10 of Ripples). Let S be the $N/2 \times N/2$ shift operator. Then $\mathbf{x}_{\text{even}}[n+1] = \big(S^{-1}\mathbf{x}_{\text{even}}\big)[n]$. Hence the formula for $\mathbf{d}$ can be written in vector form as

$$\mathbf{d} = \mathbf{x}_{\text{odd}} - \frac{1}{2}\big(\mathbf{x}_{\text{even}} + S^{-1}\mathbf{x}_{\text{even}}\big) \tag{3.4}$$

Since $\mathbf{x}_{\text{even}}$ and $\mathbf{x}_{\text{odd}}$ are periodic of period $N/2$, so is $\mathbf{d}$:

$$\mathbf{d}[k+N/2] = \mathbf{d}[k] \quad \text{for all integers } k.$$

De£ne the *prediction* transformation by $P\begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{d} \end{bmatrix}$. From formula (3.4) we can write the matrix for P in block form as

$$P = \begin{bmatrix} I & 0 \\ -\frac{1}{2}(I + S^{-1}) & I \end{bmatrix}$$

where I is the $N/2 \times N/2$ identity matrix.

As the next step in the lifting procedure, we use the detail vector $\mathbf{d}$ to update $\mathbf{x}_{\text{even}}$ and obtain the *trend vector* $\mathbf{s}$:

$$\mathbf{s}[n] = \mathbf{x}_{\text{even}}[n] + \frac{1}{4}\big(\mathbf{d}[n] + \mathbf{d}[n-1]\big).$$

The choice of the constant $\frac{1}{4}$ makes $\mathbf{s}$ and $\mathbf{x}$ have the same average value:

$$\frac{2}{N}\sum_{n=0}^{N/2-1} \mathbf{s}[n] = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{x}[n]$$

(for a proof of this, see Ripples, p. 18). The formula for s can be written in vector form as

$$\mathbf{s} = \mathbf{x}_{\text{even}} + \frac{1}{4}(\mathbf{d} + S\mathbf{d}) \tag{3.5}$$

Since $\mathbf{x}_{\text{even}}$ and $\mathbf{d}$ are periodic of period $N/2$, so is $\mathbf{s}$:

$$\mathbf{s}[k + N/2] = \mathbf{s}[k] \quad \text{for all integers } k.$$

De£ne the *update* transformation by $U \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix}$. From formula (3.5) we can write the matrix for U in block form as

$$U = \begin{bmatrix} I & \frac{1}{4}(I + S) \\ 0 & I \end{bmatrix}$$

where I is the $N/2 \times N/2$ identity matrix.

The £nal step in the lifting process is a *normalization* $D \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \sqrt{2}\mathbf{s} \\ (1/\sqrt{2})\mathbf{d} \end{bmatrix}$. Thus D is given by the diagonal matrix

$$D = \begin{bmatrix} \sqrt{2}I & 0 \\ 0 & (1/\sqrt{2})I \end{bmatrix}.$$

The *one-scale* CDF(2,2) *analysis transform* is the product of these transformations:

$$\mathbf{T_a} = DUP\,\boxed{\text{split}}\,. \tag{3.6}$$

Example 3.3.1 (CDF(2,2) Analysis Transform)**.** Suppose $N = 4$. In this case $S = S^{-1}$ and $I + S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Hence

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus we £nd (with the aid of MATLAB) that

$$\begin{aligned} \mathbf{T_a} &= \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{bmatrix} \boxed{\text{split}} \\ &= \frac{1}{2\sqrt{2}} \begin{bmatrix} 3 & 1 & -1 & 1 \\ -1 & 1 & 3 & 1 \\ -1 & 2 & -1 & 0 \\ -1 & 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \mathbf{v}_0 \\ \mathbf{v}_1 \end{bmatrix}. \end{aligned}$$

The rows of $\mathbf{T_a}$ are related as follows:

1. The row $\mathbf{u}_1$ is obtained by shifting the previous row $\mathbf{u}_0$ to the right two positions (and using periodic wraparound).

2. The row $\mathbf{v}_1$ is obtained by shifting the previous row $\mathbf{v}_0$ to the right two positions (and using periodic wraparound).

We already saw this pattern in the Haar transform, and we will show in Section 3.4 that it holds for all one-scale wavelet analysis matrices. ■

As in the case of the Haar transform, it is easy to construct the inverse (synthesis) transform $\mathbf{T_s}$ by inverting the prediction, update, and normalization transforms:

$$P^{-1} = \begin{bmatrix} I & 0 \\ \frac{1}{2}(I + S^{-1}) & I \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} I & -\frac{1}{4}(I+S) \\ 0 & I \end{bmatrix}, \quad D^{-1} = \begin{bmatrix} (1/\sqrt{2})I & 0 \\ 0 & \sqrt{2}I \end{bmatrix},$$

Hence we obtain the *one-scale* CDF(2,2) *synthesis transform* as

$$\mathbf{T_s} = \boxed{\text{merge}}\, P^{-1}U^{-1}D^{-1}. \tag{3.7}$$

Example 3.3.2 (CDF(2,2) Synthesis Transform)**.** Suppose $N = 4$. Then we calculate (with the aid of MATLAB) that

$$\begin{aligned} \mathbf{T_s} &= \boxed{\text{merge}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 1 & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \\ &= \frac{\sqrt{2}}{4} \begin{bmatrix} 2 & 0 & -1 & -1 \\ 1 & 1 & 3 & -1 \\ 0 & 2 & -1 & -1 \\ 1 & 1 & -1 & 3 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{u}}_0 & \widetilde{\mathbf{u}}_1 & \widetilde{\mathbf{v}}_0 & \widetilde{\mathbf{v}}_1 \end{bmatrix} \end{aligned}$$

The columns of $\mathbf{T_s}$ are related as follows:

1. The column $\widetilde{\mathbf{u}}_1$ is obtained by shifting $\widetilde{\mathbf{u}}_0$ down two positions (and using periodic wraparound).

2. The column $\widetilde{\mathbf{v}}_1$ is obtained by shifting $\widetilde{\mathbf{v}}_0$ down two positions (and using periodic wraparound).

We already saw this pattern in the Haar transform; we will show in Section 3.4 that it holds for all one-scale wavelet synthesis matrices. ■

The CDF(2,2) discrete wavelet transform that we have just constructed is part of a family of transforms called the *Cohen-Daubechies-Feauveau* wavelets (in the Uvi-Wave implementation they are called *w-spline* wavelets); see Section 3.6 of Ripples for more examples of these transforms. The analysis matrices for these transforms are not orthogonal, unlike the case of the (normalized) Haar transform, so the columns of the synthesis matrix are quite different from the rows of the analysis matrix. However, this family of transforms is very easy to create (the matrices have simple rational coef£cents related to the binomial coef£cients), and these transforms have many desirable properties relating to smoothness and feature detection.

Daubechies 4 Transform

Our next example of a wavelet transform is orthogonal and is one of an in£nite family of orthogonal wavelet transforms. [1]

After splitting the signal $\mathbf{x}$ of even length $N \geq 4$ into $\mathbf{x}_{\text{even}}$ and $\mathbf{x}_{\text{odd}}$, we perform an *update* operation to de£ne the *£rst trend* $\mathbf{s}^{(1)}$ (see Ripples formula 3.18):

$$\mathbf{s}^{(1)}[n] = \mathbf{x}[2n] + \sqrt{3}\mathbf{x}[2n+1] = \mathbf{x}_{\text{even}}[n] + \sqrt{3}\mathbf{x}_{\text{odd}}[n]. \tag{3.8}$$

De£ne the *£rst update* transformation by $U_1 \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{x}_{\text{odd}} \end{bmatrix}$. From formula (3.8) we can write the matrix for U_1 in block form as

$$U_1 = \begin{bmatrix} I & \sqrt{3}I \\ 0 & I \end{bmatrix}$$

where I is the $N/2 \times N/2$ identity matrix. Next, we predict $\mathbf{x}_{\text{odd}}[n]$ using two adjacent values of $\mathbf{s}^{(1)}$:

$$\text{prediction of } \mathbf{x}_{\text{odd}}[n] = \frac{1}{4}\left\{\sqrt{3}\mathbf{s}^{(1)}[n] + (\sqrt{3}-2)\mathbf{s}^{(1)}[n-1]\right\},$$

where $\mathbf{s}^{(1)}$ is extended to be periodic of period $N/2$. The difference between the prediction and the actual value is the (unnormalized) detail $\mathbf{d}^{(1)}$ (see Ripples formula 3.19):

$$\mathbf{d}^{(1)} = \mathbf{x}_{\text{odd}} - \frac{1}{4}\left\{\sqrt{3}\mathbf{s}^{(1)} + (\sqrt{3}-2)S\mathbf{s}^{(1)}\right\}. \tag{3.9}$$

We have written this equation in vector form; S is the $N/2 \times N/2$ shift matrix. De£ne the *prediction* transformation by $P \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{d}^{(1)} \end{bmatrix}$. From formula (3.9) we can write the matrix for P in block form as

$$P = \begin{bmatrix} I & 0 \\ -\frac{\sqrt{3}}{4}I - \frac{\sqrt{3}-2}{4}S & I \end{bmatrix}.$$

As the next step in the lifting procedure, we use the detail vector $\mathbf{d}^{(1)}$ to update the £rst trend $\mathbf{s}^{(1)}$ and obtain the *second trend* $\mathbf{s}^{(2)}$ (see Ripples formula 3.20):

$$\mathbf{s}^{(2)}[n] = \mathbf{s}^{(1)}[n] - \mathbf{d}^{(1)}[n+1].$$

The formula for $\mathbf{s}^{(2)}$ can be written in vector form as

$$\mathbf{s}^{(2)} = \mathbf{s}^{(1)} - S^{-1}\mathbf{d}^{(1)}. \tag{3.10}$$

De£ne the *second update* transformation by $U_2 \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{d}^{(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(2)} \\ \mathbf{d}^{(1)} \end{bmatrix}$. From formula (3.10) we can write the matrix for U_2 in block form as

$$U_2 = \begin{bmatrix} I & -S^{-1} \\ 0 & I \end{bmatrix}.$$

[1]This family of transforms was created by Ingrid Daubechies, who was a mathematics professor at Rutgers in the early 1990's and is now a professor at Princeton.

The £nal step in the lifting process is a *normalization* $\begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = D \begin{bmatrix} \mathbf{s}^{(2)} \\ \mathbf{d}^{(1)} \end{bmatrix}$, where D is the diagonal matrix

$$D = \begin{bmatrix} \frac{\sqrt{3}-1}{\sqrt{2}} I & 0 \\ 0 & \frac{\sqrt{3}+1}{\sqrt{2}} I \end{bmatrix}.$$

Note that $\frac{\sqrt{3}-1}{\sqrt{2}} \cdot \frac{\sqrt{3}+1}{\sqrt{2}} = 1$, so $\det D = 1$. The *one-scale* Daub4 *analysis transform* is de£ned to be the product of these transformations:

$$\mathbf{T_a} = DU_2PU_1 \boxed{\text{split}}.$$

The complicated numerical coef£cients in the Daub4 transform are required to obtain the following result:

Theorem 3.3.3. *The* Daub4 *wavelet analysis matrix is orthogonal.*

Proof. We calculate (using $N/2 \times N/2$ blocks) that

$$\begin{aligned} DU_2PU_1 &= \begin{bmatrix} \frac{\sqrt{3}-1}{\sqrt{2}} I & 0 \\ 0 & \frac{\sqrt{3}+1}{\sqrt{2}} I \end{bmatrix} \begin{bmatrix} I & -S^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{\sqrt{3}}{4} I - \frac{\sqrt{3}-2}{4} S & I \end{bmatrix} \begin{bmatrix} I & \sqrt{3} I \\ 0 & I \end{bmatrix} \\ &= \frac{1}{4\sqrt{2}} \begin{bmatrix} (aI + cS^{-1}) & (bI + dS^{-1}) \\ -(bI + dS) & (aI + cS) \end{bmatrix}, \end{aligned} \tag{3.11}$$

where $a = 1 + \sqrt{3}$, $b = 3 + \sqrt{3}$, $c = 3 - \sqrt{3}$, and $d = 1 - \sqrt{3}$. Thus

$$\begin{aligned} \mathbf{T_a T_a^T} &= (DU_2PU_1) \boxed{\text{split}} \boxed{\text{split}}^{\mathrm{T}} (DU_2PU_1)^{\mathrm{T}} \\ &= \frac{1}{32} \begin{bmatrix} (aI + cS^{-1}) & (bI + dS^{-1}) \\ -(bI + dS) & (aI + cS) \end{bmatrix} \begin{bmatrix} (aI + cS^{-1}) & -(bI + dS) \\ (bI + dS^{-1}) & (aI + cS) \end{bmatrix}, \end{aligned}$$

since the permutation matrix $\boxed{\text{split}}$ is orthogonal. Carrying out the block multiplication, we £nd that

$$\mathbf{T_a T_a^T} = \frac{1}{32} \begin{bmatrix} \alpha I + \beta(S + S^{-1}) & 0 \\ 0 & \alpha I + \beta(S + S^{-1}) \end{bmatrix},$$

where $\alpha = a^2 + b^2 + c^2 + d^2$ and $\beta = ac + bd$. An easy calculation shows that $\alpha = 32$ and $\beta = 0$. Hence $\mathbf{T_a T_a^T}$ is the identity matrix. ∎

Example 3.3.4. When $N = 4$ then from (3.11) we £nd that the Daub4 analysis matrix is

$$\mathbf{T_a} = \frac{1}{4\sqrt{2}} \begin{bmatrix} a & c & b & d \\ c & a & d & b \\ -b & -d & a & c \\ -d & -b & c & a \end{bmatrix} \boxed{\text{split}} = \frac{1}{4\sqrt{2}} \begin{bmatrix} a & b & c & d \\ c & d & a & b \\ -b & a & -d & c \\ -d & c & -b & a \end{bmatrix}$$

(since right multiplication by $\boxed{\text{split}}$ interchanges columns 2 and 3). Note that the second row is obtained from the £rst row by shifting right two places (with wraparound). The same relation holds between the third and fourth rows. Also, the third row is obtained from the second row by reversing the entries and inserting alternating signs; this procedure on vectors of even length automatically produces a pair of orthogonal vectors.

Since the Daub4 analysis matrix is orthogonal, the synthesis matrix is the transpose:

$$\mathbf{T_s} = \frac{1}{4\sqrt{2}} \begin{bmatrix} a & c & -b & -d \\ b & d & a & c \\ c & a & -d & -b \\ d & b & c & a \end{bmatrix}$$

The columns of $\mathbf{T_s}$ have the same pattern as the rows of $\mathbf{T_a}$. ■

3.4 Wavelet Bases

A general one-scale periodic wavelet analysis transform matrix (implemented through the lifting procedure) is of the form

$$\mathbf{T_a} = D \cdot (\text{product of updates and predictions}) \cdot \boxed{\text{split}}. \tag{3.12}$$

If $\mathbf{T_a}$ is of size $N \times N$ with N even, then the update and prediction transformations are matrices in unit block-triangular form. The nonzero off-diagonal blocks are linear combinations of (positive and negative) powers of the $N/2 \times N/2$ shift matrix S (see the examples of CDF transforms on page 24 of Ripples). Recall from Section 2.2 that a polynomial $p(S)$ in the shift matrix S is called a *circulant matrix*. Hence the *update* matrices U and the *prediction* matrices P are in $N/2 \times N/2$ block form

$$U = \begin{bmatrix} I & C_1 \\ 0 & I \end{bmatrix}, \qquad P = \begin{bmatrix} I & 0 \\ C_2 & I \end{bmatrix},$$

where $C_1 = p_1(S)$ and $C_2 = p_2(S)$ are circulant matrices. The normalization matrix is $D = \begin{bmatrix} \alpha I & 0 \\ 0 & \beta I \end{bmatrix}$ with $\alpha\beta \neq 0$ (there can be several update and prediction matrices).

The factored form (3.12) shows that $\mathbf{T_a}$ is an invertible matrix. Products and sums of circulant matrices are still circulant matrices, so the product of the update, prediction, and normalization transforms has the block form

$$\mathbf{T_a} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \boxed{\text{split}}, \quad \text{where } C_{ij} = p_{ij}(S) \text{ is a } N/2 \times N/2 \text{ circulant matrix.} \tag{3.13}$$

Likewise, the inverse (synthesis) matrix has the block form

$$\mathbf{T_s} = \boxed{\text{merge}} \begin{bmatrix} \widetilde{C}_{11} & \widetilde{C}_{12} \\ \widetilde{C}_{21} & \widetilde{C}_{22} \end{bmatrix}, \quad \text{where } \widetilde{C}_{ij} = \widetilde{p}_{ij}(S) \text{ is a } N/2 \times N/2 \text{ circulant matrix.} \tag{3.14}$$

Example 3.4.1. For the Daub4 analysis transform, we found in (3.11) that

$$C_{11} = a + cS^{-1}, \quad C_{12} = b + dS^{-1}, \quad C_{21} = -b - dS, \quad C_{22} = a + cS,$$

where

$$a = \tfrac{1+\sqrt{3}}{4\sqrt{2}}, \quad b = \tfrac{3+\sqrt{3}}{4\sqrt{2}}, \quad c = \tfrac{3-\sqrt{3}}{4\sqrt{2}}, \quad \text{and} \quad d = \tfrac{1-\sqrt{3}}{4\sqrt{2}}.$$

Since this transform is orthogonal, the inverse matrix is the transpose. Thus $\widetilde{C}_{ij} = (C_{ji})^{\mathrm{T}}$ in this case.

For the CDF(2,2) analysis transform we have

$$\begin{aligned}\mathbf{T_a} &= \begin{bmatrix} \sqrt{2}I & 0 \\ 0 & (1/\sqrt{2})I \end{bmatrix} \cdot \begin{bmatrix} I & \frac{1}{4}(I+S) \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}(I+S^{-1}) & I \end{bmatrix} \boxed{\text{split}} \\ &= \frac{1}{4\sqrt{2}} \begin{bmatrix} -S^{-1}+6I-S & 2I+2S \\ -2I-2S^{-1} & 4I \end{bmatrix} \boxed{\text{split}}.\end{aligned}$$

For the CDF(2,2) synthesis transform we have

$$\begin{aligned}\mathbf{T_s} &= \boxed{\text{merge}} \begin{bmatrix} I & 0 \\ \frac{1}{2}(I+S^{-1}) & I \end{bmatrix} \begin{bmatrix} I & -\frac{1}{4}(I+S) \\ 0 & I \end{bmatrix} \begin{bmatrix} (1/\sqrt{2})I & 0 \\ 0 & \sqrt{2}I \end{bmatrix}. \\ &= \frac{1}{4\sqrt{2}} \boxed{\text{merge}} \begin{bmatrix} 4I & -2I-2S \\ 2I+2S^{-1} & -S^{-1}+6I-S \end{bmatrix}\end{aligned}$$

■

We return to a general wavelet transform of real N-periodic signals, where $N = 2m$ is even. Let $\mathbf{T_a}$ be a one-scale wavelet analysis matrix given by a formula of the type (3.12). Write $\mathbf{T_a}$ in terms of its rows as

$$\mathbf{T_a} = \begin{bmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{m-1} \\ \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{m-1} \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix},$$

where $\mathbf{u}_j$, $\mathbf{v}_j$ are $1 \times N$ row vectors and $\mathbf{U}$, $\mathbf{V}$ are $m \times N$ matrices. Let $\mathbf{T_s}$ be the inverse synthesis matrix. We write $\mathbf{T_s}$ in terms of its columns as

$$\mathbf{T_s} = \begin{bmatrix} \widetilde{\mathbf{u}}_0 & \dots & \widetilde{\mathbf{u}}_{m-1} & \widetilde{\mathbf{v}}_0 & \dots & \widetilde{\mathbf{v}}_{m-1} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{U}} & \widetilde{\mathbf{V}} \end{bmatrix},$$

where $\widetilde{\mathbf{u}}_j$, $\widetilde{\mathbf{v}}_j$ are $N \times 1$ column vectors and $\widetilde{\mathbf{U}}$, $\widetilde{\mathbf{V}}$ are $N \times m$ matrices. The matrix inverse property $\mathbf{T_a}\mathbf{T_s} = I$ can be expressed as the *biorthogonality relations*

$$\begin{aligned}\mathbf{u}_j\widetilde{\mathbf{u}}_k &= \delta[j-k], \qquad \mathbf{u}_j\widetilde{\mathbf{v}}_k = 0, \\ \mathbf{v}_j\widetilde{\mathbf{u}}_k &= 0, \qquad \mathbf{v}_j\widetilde{\mathbf{v}}_k = \delta[j-k],\end{aligned}$$

for $j, k = 0, 1, \dots, m-1$, where $\delta[n] = 1$ if $n = 0$ and $\delta[n] = 0$ if $n \neq 0$. These relations are the same as the matrix equations

$$\begin{aligned}\mathbf{U}\,\widetilde{\mathbf{U}} &= I, \qquad \mathbf{U}\,\widetilde{\mathbf{V}} = 0, \\ \mathbf{V}\,\widetilde{\mathbf{U}} &= 0, \qquad \mathbf{V}\,\widetilde{\mathbf{V}} = I,\end{aligned}$$

where I is the $m \times m$ identity matrix. When $\mathbf{T_a}$ is an orthogonal matrix (as for the Haar or Daub4 transforms) we have $\widetilde{\mathbf{U}} = \mathbf{U}^{\mathrm{T}}$ and $\widetilde{\mathbf{V}} = \mathbf{V}^{\mathrm{T}}$. In general, these matrices are all different.

We now show that the matrix $\mathbf{T_a}$ is completely determined by the row vectors $\mathbf{u}_0$ and $\mathbf{v}_0$. Likewise, the matrix $\mathbf{T_s}$ is completely determined by the column vectors $\widetilde{\mathbf{u}}_0$ and $\widetilde{\mathbf{v}}_0$. (We have already seen this pattern in the Haar, CDF(2,2) and Daub4 transforms.)

Theorem 3.4.2. *Let S_N be the $N \times N$ shift matrix ($N = 2m$). Then for $k = 1, \ldots, m-1$*

$$\begin{aligned} \mathbf{u}_k &= \mathbf{u}_0(S_N)^{-2k} && \textit{(shift components of } \mathbf{u}_0 \textit{ to right } 2k \textit{ positions with wraparound),} \\ \mathbf{v}_k &= \mathbf{v}_0(S_N)^{-2k} && \textit{(shift components of } \mathbf{v}_0 \textit{ to right } 2k \textit{ positions with wraparound).} \end{aligned}$$

Likewise,

$$\begin{aligned} \widetilde{\mathbf{u}}_k &= (S_N)^{2k}\widetilde{\mathbf{u}}_0 && \textit{(shift components of } \widetilde{\mathbf{u}}_0 \textit{ down } 2k \textit{ positions with wraparound),} \\ \widetilde{\mathbf{v}}_k &= (S_N)^{2k}\widetilde{\mathbf{v}}_0 && \textit{(shift components of } \widetilde{\mathbf{v}}_0 \textit{ down } 2k \textit{ positions with wraparound).} \end{aligned}$$

Proof. The key point is the relation between shifting and splitting (or merging). Let $\mathbf{x} \in \mathbb{R}^N$. Shifting the components of $\mathbf{x}$ down by *two positions* (with wraparound) is the same as shifting the even and odd parts of $\mathbf{x}$ down by *one position* (with wraparound). Thus

$$\boxed{\text{split}}\,(S_N)^2\mathbf{x} = \begin{bmatrix} S\mathbf{x}_{\text{even}} \\ S\mathbf{x}_{\text{odd}} \end{bmatrix},$$

where S is the $m \times m$ shift matrix. We can write this relation in matrix terms as

$$\boxed{\text{split}}\,(S_N)^2 = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \boxed{\text{split}}\,. \tag{3.15}$$

Now we write $\mathbf{T_a}$ as in (3.13) and use (3.15) to calculate

$$\begin{aligned} \mathbf{T_a}(S_N)^2 &= \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \boxed{\text{split}}\,(S_N)^2 = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \boxed{\text{split}}\,. \\ &= \begin{bmatrix} C_{11}S & C_{12}S \\ C_{21}S & C_{22}S \end{bmatrix} \boxed{\text{split}} = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \boxed{\text{split}}\,. \end{aligned}$$

Here we have used the fundamental property $C_{ij}S = SC_{ij}$ of a circulant matrix. Thus

$$\mathbf{T_a}(S_N)^2 = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \mathbf{T_a} \tag{3.16}$$

Now write $\mathbf{T_a}$ in terms of the matrices $\mathbf{U}$, $\mathbf{V}$ and apply the last equation:

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} (S_N)^2 = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix}$$

Carrying out the block multiplication on each side of this equation, we see that

$$\begin{bmatrix} \mathbf{U}(S_N)^2 \\ \mathbf{V}(S_N)^2 \end{bmatrix} = \begin{bmatrix} S\mathbf{U} \\ S\mathbf{V} \end{bmatrix}$$

Hence
$$S\mathbf{U} = \mathbf{U}(S_N)^2 \quad \text{and} \quad S\mathbf{V} = \mathbf{V}(S_N)^2.$$
(Remember that $\mathbf{U}$ and $\mathbf{V}$ are of size $m \times N$, so all the matrix products in these equations are de£ned.) Since left multiplication by S shifts the rows of $\mathbf{U}$ and $\mathbf{V}$ down one position, we obtain
$$\begin{bmatrix} \mathbf{u}_{m-1} \\ \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{m-2} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_0(S_N)^2 \\ \mathbf{u}_1(S_N)^2 \\ \vdots \\ \mathbf{u}_{m-1}(S_N)^2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{v}_{m-1} \\ \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{m-2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_0(S_N)^2 \\ \mathbf{v}_1(S_N)^2 \\ \vdots \\ \mathbf{v}_{m-1}(S_N)^2 \end{bmatrix}. \tag{3.17}$$
Matching up the rows in equations (3.17) and multiplying each row on the right by $(S_N)^{-2}$, we see that
$$\begin{aligned} \mathbf{u}_1 &= \mathbf{u}_0(S_N)^{-2}, \quad \mathbf{u}_2 = \mathbf{u}_1(S_N)^{-2} = \mathbf{u}_0(S_N)^{-4}, \ldots \\ \mathbf{v}_1 &= \mathbf{v}_0(S_N)^{-2}, \quad \mathbf{v}_2 = \mathbf{v}_1(S_N)^{-2} = \mathbf{v}_0(S_N)^{-4}, \ldots \end{aligned}$$
Multiplying a row vector $\mathbf{u}$ on the right by $(S_N)^{-2}$ shifts the components of $\mathbf{u}$ to the right two positions (with periodic wraparound). This proves that the rows of $\mathbf{T_a}$ follow the pattern asserted by Theorem 3.4.2.

The proof for the inverse matrix $\mathbf{T_s}$ follows the same pattern. Taking inverses in (3.16), we obtain the relation
$$(S_N)^{-2}\mathbf{T_s} = \mathbf{T_s} \begin{bmatrix} S^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix}.$$
Writing $\mathbf{T_s}$ in block form, we obtain
$$(S_N)^{-2} \begin{bmatrix} \widetilde{\mathbf{U}} & \widetilde{\mathbf{V}} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{U}} & \widetilde{\mathbf{V}} \end{bmatrix} \begin{bmatrix} S^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix}$$
Carrying out the block multiplication on each side of this equation, we see that
$$\begin{bmatrix} (S_N)^{-2}\widetilde{\mathbf{U}} & (S_N)^{-2}\widetilde{\mathbf{V}} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{U}}S^{-1} & \widetilde{\mathbf{V}}S^{-1} \end{bmatrix}$$
Hence
$$(S_N)^{-2}\widetilde{\mathbf{U}} = \widetilde{\mathbf{U}}S^{-1} \quad \text{and} \quad (S_N)^{-2}\widetilde{\mathbf{V}} = \widetilde{\mathbf{V}}S^{-1}$$
Since right multiplication by S^{-1} shifts the columns of $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$ to the right one position, we £nd that
$$\begin{aligned} \begin{bmatrix} (S_N)^{-2}\,\widetilde{\mathbf{u}}_0 & (S_N)^{-2}\,\widetilde{\mathbf{u}}_1 & \cdots & (S_N)^{-2}\,\widetilde{\mathbf{u}}_{m-1} \end{bmatrix} &= \begin{bmatrix} \widetilde{\mathbf{u}}_{m-1} & \widetilde{\mathbf{u}}_0 & \cdots & \widetilde{\mathbf{u}}_{m-2} \end{bmatrix} \\ \begin{bmatrix} (S_N)^{-2}\,\widetilde{\mathbf{v}}_0 & (S_N)^{-2}\,\widetilde{\mathbf{v}}_1 & \cdots & (S_N)^{-2}\,\widetilde{\mathbf{v}}_{m-1} \end{bmatrix} &= \begin{bmatrix} \widetilde{\mathbf{v}}_{m-1} & \widetilde{\mathbf{v}}_0 & \cdots & \widetilde{\mathbf{v}}_{m-2} \end{bmatrix}. \end{aligned}$$
Matching up columns in these equations and multiplying on the left by $(S_N)^2$, we obtain
$$\begin{aligned} \widetilde{\mathbf{u}}_1 &= (S_N)^2\,\widetilde{\mathbf{u}}_0, \quad \widetilde{\mathbf{u}}_2 = (S_N)^2\,\widetilde{\mathbf{u}_1} = (S_N)^4\,\widetilde{\mathbf{u}}_0, \ldots \\ \widetilde{\mathbf{v}}_1 &= (S_N)^2\,\widetilde{\mathbf{v}}_0, \quad \widetilde{\mathbf{v}}_2 = (S_N)^2\,\widetilde{\mathbf{v}}_1 = (S_N)^4\,\widetilde{\mathbf{v}}_0, \ldots \end{aligned}$$

This proves that the columns of $\mathbf{T_s}$ follow the pattern asserted by Theorem 3.4.2. ■

Given $\mathbf{x} \in \mathbb{R}^N$, we de£ne the one-scale *trend vector* $\mathbf{s}$ and *detail vector* $\mathbf{d}$ in $\mathbb{R}^m$ using the analysis matrix $\mathbf{T_a}$:

$$\begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = \mathbf{T_a}\mathbf{x} = \begin{bmatrix} \mathbf{U}\mathbf{x} \\ \mathbf{V}\mathbf{x} \end{bmatrix}$$

From Theorem 3.4.2 these vectors have components that are products of the even shifts of the row vectors $\mathbf{u}_0$ and $\mathbf{v}_0$ with the column vector $\mathbf{x}$:

$$\mathbf{s} = \begin{bmatrix} \mathbf{u}_0\mathbf{x} & \mathbf{u}_0 S^2\mathbf{x} & \cdots & \mathbf{u}_0 S^{2m-2}\mathbf{x} \end{bmatrix}^{\mathrm{T}} \quad \text{and} \quad \mathbf{d} = \begin{bmatrix} \mathbf{v}_0\mathbf{x} & \mathbf{v}_0 S^2\mathbf{x} & \cdots & \mathbf{v}_0 S^{2m-2}\mathbf{x} \end{bmatrix}^{\mathrm{T}}.$$

We can then reconstruct $\mathbf{x}$ from $\mathbf{s}$ and $\mathbf{d}$ by the synthesis matrix:

$$\mathbf{x} = \mathbf{T_s} \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{U}} & \widetilde{\mathbf{V}} \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix} = \widetilde{\mathbf{U}}\mathbf{s} + \widetilde{\mathbf{V}}\mathbf{d}. \tag{3.18}$$

De£ne the *trend subspace* to be the column space of $\widetilde{\mathbf{U}}$ and the *detail subspace* to be the column space of $\widetilde{\mathbf{V}}$. Then

$$\widetilde{\mathbf{U}}\mathbf{s} = (\mathbf{u}_0\mathbf{x})\,\widetilde{\mathbf{u}}_0 + (\mathbf{u}_0 S^2\mathbf{x})\,S^2\widetilde{\mathbf{u}}_0 + \cdots + (\mathbf{u}_0 S^{2m-2}\mathbf{x})\,S^{2m-2}\widetilde{\mathbf{u}}_0 \tag{3.19}$$

is in the trend subspace and

$$\widetilde{\mathbf{V}}\mathbf{d} = (\mathbf{v}_0\mathbf{x})\widetilde{\mathbf{v}}_0 + (\mathbf{v}_0 S^2\mathbf{x})\,S^2\widetilde{\mathbf{v}}_0 + \cdots + (\mathbf{v}_0 S^{2m-2}\mathbf{x})\,S^{2m-2}\widetilde{\mathbf{v}}_0 \tag{3.20}$$

is in the detail subspace. Decomposition (3.18) expresses $\mathbb{R}^N$ as the direct sum of the trend and detail subspaces.

Example 3.4.3. Consider the CDF(2,2) transform with $N = 4$. From Examples 3.3.1 and 3.3.2 we have

$$\mathbf{T_a} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 3 & 1 & -1 & 1 \\ -1 & 1 & 3 & 1 \\ -1 & 2 & -1 & 0 \\ -1 & 0 & -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{T_s} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 & 0 & -1 & -1 \\ 1 & 1 & 3 & -1 \\ 0 & 2 & -1 & -1 \\ 1 & 1 & -1 & 3 \end{bmatrix}$$

Then the trend subspace is the span of the vectors $\begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$; the detail subspace is the span of the vectors $\begin{bmatrix} -1 \\ 3 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix}$. Suppose $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$ (linearly increasing entries). Then $\mathbf{s} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 \\ 10 \end{bmatrix}$ and $\mathbf{d} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 0 \\ 4 \end{bmatrix}$. We can decompose $\mathbf{x} = \widetilde{U}\mathbf{s} + \widetilde{V}\mathbf{d}$ (sum of trend and detail), where

$$\widetilde{U}\mathbf{s} = \frac{2}{8} \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} + \frac{10}{8} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1.5 \\ 2.5 \\ 1.5 \end{bmatrix}, \quad \widetilde{V}\mathbf{d} = \frac{4}{8} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ 1.5 \end{bmatrix}.$$

Recall that we constructed the CDF(2,2) transform so that the trend would be an exact £t to linear signals. In this example we see the effect of periodic wraparound: the trend component $\widetilde{U}\mathbf{s}$ linearly increases (just like the signal $\mathbf{x}$) until the last entry, which then goes down because of periodicity. The detail component $\widetilde{V}\mathbf{d}$ is small but not zero. ■

3.5 Two-dimensional Wavelet Transforms

Images as Matrices

A two-dimensional black and white image can be digitized as a matrix $\mathbf{X}$ of size $M \times N$ by imposing a rectangular grid with M horizontal strips and N vertical strips on the image. Each rectangle in the grid is called a *pixel* (picture element) and is given a numerical value (*grayscale*) corresponding to the average darkness or brightness of the image in the pixel. With eight bit encoding the numbers range from 0 to 255 with 0 for black and 255 for white. The origin of coordinates is placed at the upper left-hand corner of the image and the vertical axis points down. Thus the entry $\mathbf{X}[i, j]$ in $\mathbf{X}$ encodes the average grayscale level of the pixel that is i units *down* and j units to the *right* of the upper left-hand corner of the image (this system of coordinates agrees with the usual labeling of matrix entries).

For example, we can digitize the image

$$\begin{array}{cccc} \blacksquare & \square & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \square & \square \\ \blacksquare & \square & \blacksquare & \square \\ \blacksquare & \square & \square & \blacksquare \end{array} \quad \text{by the matrix} \quad \mathbf{X} = \begin{bmatrix} 0 & 240 & 0 & 0 \\ 0 & 0 & 240 & 240 \\ 0 & 240 & 0 & 240 \\ 0 & 240 & 240 & 0 \end{bmatrix}.$$

(We are making the white squares slightly gray by encoding them using 240 instead of 255.) We shall only consider the case $M = N$ in the following.

One-scale 2D Wavelet Transform

Let $\mathbf{X}$ be an $N \times N$ matrix that encodes a grayscale image (assume N is even). Let $\mathbf{W_a}$ be the $N \times N$ one-scale analysis matrix for a wavelet transform (such as Haar, CDF(2,2), or Daub4). The transform of $\mathbf{X}$ is de£ned to be

$$\mathbf{Y} = \mathbf{W_a}\mathbf{X}\mathbf{W_a}^{\mathrm{T}}.$$

There are three ways to describe $\mathbf{Y}$:

1. Write $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_N \end{bmatrix}$ in terms of its columns and set

$$\mathbf{Z} = \mathbf{W_a}\mathbf{X} = \begin{bmatrix} \mathbf{W_a}\mathbf{x}_1 & \cdots & \mathbf{W_a}\mathbf{x}_N \end{bmatrix}.$$

   The columns of $\mathbf{Z}$ are the wavelet transforms of the columns of $\mathbf{X}$, and $\mathbf{Y} = \mathbf{Z}\mathbf{W_a}^{\mathrm{T}} = (\mathbf{W_a}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}$. Thus the *rows* of $\mathbf{Y}$ are the wavelet transforms of the rows of $\mathbf{Z}$.

2. Write $\mathbf{X}^{\mathrm{T}} = \begin{bmatrix} \widetilde{\mathbf{x}}_1 & \cdots & \widetilde{\mathbf{x}}_N \end{bmatrix}$ in terms of its columns (these are the rows of $\mathbf{X}$) and set

$$\widetilde{\mathbf{Z}} = \mathbf{X}\mathbf{W_a}^{\mathrm{T}} = (\mathbf{W_a}\mathbf{X}^{\mathrm{T}})^{\mathrm{T}} = \begin{bmatrix} \mathbf{W_a}\widetilde{\mathbf{x}}_1 & \cdots & \mathbf{W_a}\widetilde{\mathbf{x}}_N \end{bmatrix}^{\mathrm{T}}.$$

The rows of $\widetilde{\mathbf{Z}}$ are the wavelet transforms of the rows of $\mathbf{X}$, and $\mathbf{Y} = \mathbf{W_a}\widetilde{\mathbf{Z}}$. Thus the *columns* of $\mathbf{Y}$ are the wavelet transforms of the columns of $\widetilde{\mathbf{Z}}$.

3. Write $\mathbf{W_a} = \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix}$, where $\mathbf{U}$ consists of the $N/2$ *trend rows* and $\mathbf{V}$ consists of the $N/2$ *detail rows* (see Theorem 3.4.2). Then

$$\mathbf{Y} = \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} \mathbf{X} \begin{bmatrix} \mathbf{U}^{\mathrm{T}} & \mathbf{V}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \mathbf{UXU}^{\mathrm{T}} & \mathbf{UXV}^{\mathrm{T}} \\ \mathbf{VXU}^{\mathrm{T}} & \mathbf{VXV}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \mathbf{Y_{ss}} & \mathbf{Y_{sd}} \\ \mathbf{Y_{ds}} & \mathbf{Y_{dd}} \end{bmatrix}. \tag{3.21}$$

Example 3.5.1. Take $\mathbf{W_a} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ to be the (unnormalized) 2×2 Haar transform matrix and $\mathbf{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$\mathbf{Y} = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} (a+b+c+d) & (a-b+c-d) \\ (a-c+b-d) & (a-b-c+d) \end{bmatrix}$$

Thus we see that

$$\begin{aligned}
\mathbf{Y_{ss}} &= \frac{1}{4}(a+b+c+d) \quad \text{(overall average)} \\
\mathbf{Y_{sd}} &= \frac{1}{4}[(a-b)+(c-d)] \quad \text{(average of column-to-column differences)} \\
\mathbf{Y_{ds}} &= \frac{1}{4}[(a-c)+(b-d)] \quad \text{(average of row-to-row differences)} \\
\mathbf{Y_{dd}} &= \frac{1}{4}[(a-b)-(c-d)] \quad \text{(column difference of row differences)}
\end{aligned}$$

■

Let $\mathbf{W_s} = \mathbf{W_a}^{-1}$ be the one-scale synthesis matrix for the wavelet transform. Then the original matrix $\mathbf{X}$ can be reconstructed from the transform $\mathbf{Y}$:

$$\mathbf{X} = \mathbf{W_s}\mathbf{W_a}\mathbf{X}\mathbf{W_a^{\mathrm{T}}}\mathbf{W_s^{\mathrm{T}}} = \mathbf{W_s}\mathbf{Y}\mathbf{W_s^{\mathrm{T}}}$$

Write $\mathbf{W_s} = \begin{bmatrix} \widetilde{U} & \widetilde{V} \end{bmatrix}$ in terms of the trend and detail columns (as in Theorem 3.4.2). and use the block decomposition of $\mathbf{Y}$ in (3.21). Then

$$\begin{aligned}
\mathbf{X} &= \begin{bmatrix} \widetilde{U} & \widetilde{V} \end{bmatrix}\begin{bmatrix} \mathbf{Y_{ss}} & \mathbf{Y_{sd}} \\ \mathbf{Y_{ds}} & \mathbf{Y_{dd}} \end{bmatrix}\begin{bmatrix} \widetilde{U}^{\mathrm{T}} \\ \widetilde{V}^{\mathrm{T}} \end{bmatrix} \\
&= \widetilde{U}\mathbf{Y_{ss}}\widetilde{U}^{\mathrm{T}} + \widetilde{U}\mathbf{Y_{sd}}\widetilde{V}^{\mathrm{T}} + \widetilde{V}\mathbf{Y_{ds}}\widetilde{U}^{\mathrm{T}} + \widetilde{V}\mathbf{Y_{dd}}\widetilde{V}^{\mathrm{T}} \qquad (3.22)
\end{aligned}$$

We call (3.22) the *multiresolution decomposition* of $\mathbf{X}$. Descriptions 1. and 2. above of $\mathbf{Y}$ show that the four matrices in the decomposition (3.22) carry the following information about the image encoded by $\mathbf{X}$:

$$\begin{aligned}
\mathbf{X_{ss}} &= \widetilde{U}\mathbf{Y_{ss}}\widetilde{U}^{\mathrm{T}} \quad \text{column and row trend (overall features of image)} \\
\mathbf{X_{sd}} &= \widetilde{U}\mathbf{Y_{sd}}\widetilde{V}^{\mathrm{T}} \quad \text{column trend and row detail (vertical edges of image)} \\
\mathbf{X_{ds}} &= \widetilde{V}\mathbf{Y_{ds}}\widetilde{U}^{\mathrm{T}} \quad \text{column detail and row trend (horizontal edges of image)} \\
\mathbf{X_{dd}} &= \widetilde{V}\mathbf{Y_{dd}}\widetilde{V}^{\mathrm{T}} \quad \text{column and row detail (diagonal edges of image)}
\end{aligned}$$

(see Figures 6.3 and 6.4 in Ripples).

Example 3.5.2. Take $\mathbf{W_a}$ in Example 3.5.1. Then $\mathbf{W_s} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, and so $\widetilde{U} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\widetilde{V} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Suppose $\mathbf{X} = \begin{bmatrix} 14 & 2 \\ 4 & 0 \end{bmatrix}$. Then

$$\mathbf{Y} = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 14 & 2 \\ 4 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix}.$$

The multiresolution decomposition of $\mathbf{X}$ is

$$\begin{aligned} \begin{bmatrix} 14 & 2 \\ 4 & 0 \end{bmatrix} &= 5\widetilde{\mathbf{U}}\widetilde{\mathbf{U}}^{\mathrm{T}} + 4\widetilde{\mathbf{U}}\widetilde{\mathbf{V}}^{\mathrm{T}} + 3\widetilde{\mathbf{V}}\widetilde{\mathbf{U}}^{\mathrm{T}} + 2\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^{\mathrm{T}} \\ &= 5\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + 4\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} + 3\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + 2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \mathbf{X_{ss}} + \mathbf{X_{sd}} + \mathbf{X_{ds}} + \mathbf{X_{dd}}. \end{aligned}$$

If we represent a matrix entry 1 by a white box and a matrix entry -1 by a black box, then this last equation can be displayed as

$$\mathbf{X} = 5 \begin{matrix} \square & \square \\ \square & \square \end{matrix} + 4 \begin{matrix} \square & \blacksquare \\ \square & \blacksquare \end{matrix} + 3 \begin{matrix} \square & \square \\ \blacksquare & \blacksquare \end{matrix} + 2 \begin{matrix} \square & \blacksquare \\ \blacksquare & \square \end{matrix}$$

The coef£cient 5 is the average of the four entries in $\mathbf{X}$, and it multiplies the matrix with all entries white. The other three matrices detect the pattern of vertical, horizontal, and diagonal detail in $\mathbf{X}$. ∎

Example 3.5.3. For a slightly more complicated example, take $\mathbf{W_a}$ to be the (unnormalized) 4×4 Haar transform matrix and take $\mathbf{X}$ to be the matrix at the beginning of this section:

$$\mathbf{W_a} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 0 & 240 & 0 & 0 \\ 0 & 0 & 240 & 240 \\ 0 & 240 & 0 & 240 \\ 0 & 240 & 240 & 0 \end{bmatrix}.$$

Then we calculate (with the aid of MATLAB) that

$$\mathbf{Y} = \mathbf{W_a}\mathbf{X}\mathbf{W_a}^{\mathrm{T}} = \begin{bmatrix} 60 & 120 & -60 & 0 \\ 120 & 120 & -120 & 0 \\ 60 & -120 & -60 & 0 \\ 0 & 0 & 0 & -120 \end{bmatrix}.$$

Hence the four submatrices of $\mathbf{Y}$ are

$$\mathbf{Y_{ss}} = \begin{bmatrix} 60 & 120 \\ 120 & 120 \end{bmatrix}, \mathbf{Y_{sd}} = \begin{bmatrix} -60 & 0 \\ -120 & 0 \end{bmatrix}, \mathbf{Y_{ds}} = \begin{bmatrix} 60 & -120 \\ 0 & 0 \end{bmatrix}, \mathbf{Y_{dd}} = \begin{bmatrix} -60 & 0 \\ 0 & -120 \end{bmatrix}.$$

The synthesis matrix is

$$\mathbf{W_s} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{and so} \quad \widetilde{U} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \widetilde{V} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}.$$

The multiresolution decomposition (calculated with the aid of MATLAB) is

$$\begin{aligned} \mathbf{X} &= \widetilde{\mathbf{U}} \begin{bmatrix} 60 & 120 \\ 120 & 120 \end{bmatrix} \widetilde{\mathbf{U}}^{\mathrm{T}} + \widetilde{\mathbf{U}} \begin{bmatrix} -60 & 0 \\ -120 & 0 \end{bmatrix} \widetilde{\mathbf{V}}^{\mathrm{T}} \\ &\quad + \widetilde{\mathbf{V}} \begin{bmatrix} 60 & -120 \\ 0 & 0 \end{bmatrix} \widetilde{\mathbf{U}}^{\mathrm{T}} + \widetilde{\mathbf{V}} \begin{bmatrix} -60 & 0 \\ 0 & -120 \end{bmatrix} \widetilde{\mathbf{V}}^{\mathrm{T}} \\ &= \mathbf{X_{ss}} + \mathbf{X_{sd}} + \mathbf{X_{ds}} + \mathbf{X_{dd}}, \end{aligned}$$

where the components are

$$\mathbf{X_{ss}} = \begin{bmatrix} 60 & 60 & 120 & 120 \\ 60 & 60 & 120 & 120 \\ 120 & 120 & 120 & 120 \\ 120 & 120 & 120 & 120 \end{bmatrix}, \qquad \mathbf{X_{sd}} = \begin{bmatrix} -60 & -60 & 0 & 0 \\ -60 & -60 & 0 & 0 \\ -120 & -120 & 0 & 0 \\ -120 & -120 & 0 & 0 \end{bmatrix}$$

$$\mathbf{X_{ds}} = \begin{bmatrix} 60 & 60 & -120 & -120 \\ -60 & -60 & 120 & 120 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{X_{dd}} = \begin{bmatrix} -60 & 60 & 0 & 0 \\ 60 & -60 & 0 & 0 \\ 0 & 0 & -120 & 120 \\ 0 & 0 & 120 & -120 \end{bmatrix}$$

Comparing these four matrices with the original image, we see that $\mathbf{X_{ss}}$ gives the overall pattern (darker in the upper-left portion, lighter elsewhere), $\mathbf{X_{sd}}$ emphasizes the vertical edges, $\mathbf{X_{ds}}$ emphasizes the horizontal edges, and $\mathbf{X_{dd}}$ emphasizes the diagonal features. ■

Multi-scale 2D Wavelet Transform

The multiscale 2D wavelet tranform is obtained by the same pyramid algorithm used for one-dimensional signals: the three submatrices containing detail information are saved, and the pure trend submatrix is subjected to further wavelet transforms.

Let the image matrix $\mathbf{X}$ be of size $N \times N$, where now N is a multiple of 4, and let

$$\mathbf{Y}^{(1)} = \begin{bmatrix} \mathbf{Y}^{(1)}_{\mathbf{ss}} & \mathbf{Y}^{(1)}_{\mathbf{sd}} \\ \mathbf{Y}^{(1)}_{\mathbf{ds}} & \mathbf{Y}^{(1)}_{\mathbf{dd}} \end{bmatrix}$$

be the one-scale transform of $\mathbf{X}$. Let $\mathbf{W}^{(2)}_{\mathbf{a}}$ be the $N/2 \times N/2$ wavelet analysis matrix and write

$$\mathbf{W}^{(2)}_{\mathbf{a}} \mathbf{Y}^{(1)}_{\mathbf{ss}} (\mathbf{W}^{(2)}_{\mathbf{a}})^{\mathrm{T}} = \begin{bmatrix} \mathbf{Y}^{(2)}_{\mathbf{ss}} & \mathbf{Y}^{(2)}_{\mathbf{sd}} \\ \mathbf{Y}^{(2)}_{\mathbf{ds}} & \mathbf{Y}^{(2)}_{\mathbf{dd}} \end{bmatrix},$$

where each block matrix is of size $N/4 \times N/4$. The two-scale transform $\mathbf{Y}^{(2)}$ of $\mathbf{X}$ is obtained by replacing the block $\mathbf{Y}^{(1)}_{\mathbf{ss}}$ in $\mathbf{Y}^{(1)}$ by this block matrix form of $\mathbf{W}^{(2)}_{\mathbf{a}}\mathbf{Y}^{(1)}_{\mathbf{ss}}(\mathbf{W}^{(2)}_{\mathbf{a}})^{\mathrm{T}}$:

$$\mathbf{Y}^{(2)} = \begin{bmatrix} \begin{bmatrix} \mathbf{Y}^{(2)}_{\mathbf{ss}} & \mathbf{Y}^{(2)}_{\mathbf{sd}} \\ & \\ \mathbf{Y}^{(2)}_{\mathbf{ds}} & \mathbf{Y}^{(2)}_{\mathbf{dd}} \end{bmatrix} & \mathbf{Y}^{(1)}_{\mathbf{sd}} \\ & \\ \mathbf{Y}^{(1)}_{\mathbf{ds}} & \mathbf{Y}^{(1)}_{\mathbf{dd}} \end{bmatrix}$$

(see Fig. 6.5 of Ripples). The inverse transformation begins with

$$\mathbf{Y}^{(1)}_{\mathbf{ss}} = \mathbf{W}^{(2)}_{\mathbf{s}} \begin{bmatrix} \mathbf{Y}^{(2)}_{\mathbf{ss}} & \mathbf{Y}^{(2)}_{\mathbf{sd}} \\ & \\ \mathbf{Y}^{(2)}_{\mathbf{ds}} & \mathbf{Y}^{(2)}_{\mathbf{dd}} \end{bmatrix} (\mathbf{W}^{(2)}_{\mathbf{s}})^{\mathrm{T}},$$

where $\mathbf{W}^{(2)}_{\mathbf{s}}$ is the $N/2 \times N/2$ synthesis matrix. Then $\mathbf{X}$ is reconstructed from the one-scale transform $\mathbf{Y}^{(1)}$ using the $N \times N$ synthesis matrix $\mathbf{W}_{\mathbf{s}}$ as before. If $N/4$ is even, then this procedure can be continued by applying a 2D wavelet transform to $\mathbf{Y}^{(2)}_{\mathbf{ss}}$ (see Figures 6.6 and 6.7 of Ripples).

Example 3.5.4. Take $\mathbf{X}$ as in Example 3.5.3. Then the Haar transform of the one-scale trend submatrix $\mathbf{Y}^{(1)}_{\mathbf{ss}}$ is

$$\frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 60 & 120 \\ 120 & 120 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 105 & -15 \\ -15 & -15 \end{bmatrix}.$$

Hence the two-scale Haar transform of $\mathbf{X}$ is

$$\mathbf{Y}^{(2)} = \begin{bmatrix} 105 & -15 & -60 & 0 \\ -15 & -15 & -120 & 0 \\ 60 & -120 & -60 & 0 \\ 0 & 0 & 0 & -120 \end{bmatrix}.$$

■

3.6 Exercises

1. Let $\mathbf{x}$ be a real-valued function on $\{0, 1, 2, 3\}$. Extend $\mathbf{x}$ to be a periodic function on the integers of period 4. De£ne a *trend* function $\mathbf{s}$ and a *detail* function $\mathbf{d}$ by the following *lifting step* formulas for $n = 0, 1$:

$$\begin{aligned} \mathbf{d}[n] &= \mathbf{x}[2n+1] - \mathbf{x}[2n] - 2\mathbf{x}[2n+2] \\ \mathbf{s}[n] &= \mathbf{x}[2n] + \mathbf{d}[n] + 3\mathbf{d}[n-1] \end{aligned}$$

(a) Suppose $\mathbf{x} = [\,4,\ 7,\ 0,\ 3\,]$. Calculate $\mathbf{d}[0]$, $\mathbf{d}[1]$, $\mathbf{s}[0]$, and $\mathbf{s}[1]$.

(b) Let $\mathbf{x}_{\text{even}} = \begin{bmatrix} \mathbf{x}[0] \\ \mathbf{x}[2] \end{bmatrix}$ and $\mathbf{x}_{\text{odd}} = \begin{bmatrix} \mathbf{x}[1] \\ \mathbf{x}[3] \end{bmatrix}$. Identify $\mathbf{s}$ and $\mathbf{d}$ with column vectors in $\mathbb{R}^2$ as usual. Let P be the *prediction* linear transformation: $P \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{d} \end{bmatrix}$. Write down the matrix for P. First give the matrix in 2×2 block form (using the shift matrix S), and then give the 4×4 numerical matrix.

(c) Let U be the *update* linear transformation: $U \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ \mathbf{d} \end{bmatrix}$. Write down the matrix for U. First give the matrix in 2×2 block form (using the shift matrix S), and then give the 4×4 numerical matrix.

2. Show that the coef£cients in the Daub4 transform satisfy $a+b+c+d = 8$ and $a+c = b+d$ (see Example 3.3.4). Use this to calculate the transform $\mathbf{T_a x}$ when $\mathbf{x}$ is the constant signal $[1\,1\,1\,1]^{\mathrm{T}}$. Find the trend vector $\mathbf{s}$ and the detail vector $\mathbf{d}$. Check that $||\mathbf{x}||^2 = ||\mathbf{s}||^2 + ||\mathbf{d}||^2$.

3. Suppose A, B, C, D are $m \times m$ circulant matrices that satisfy $AD - BC = zI$, where $z \neq 0$ is a complex number.

   (a) Show that the $2m \times 2m$ matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is invertible with inverse $z^{-1} \begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$.

   (b) Use the result from **(a)** to obtain the formula for the CDF(2,2) synthesis transform from the analysis transform (see Example 3.4.1).

4. The CDF(3,1) wavelet transform of a vector $\mathbf{x}$ of length N (even) consists of the following lifting steps (in the order given) with a £nal normalization:

   First update U_1: $\quad \mathbf{s}^{(1)}[n] = \mathbf{x}_{\text{even}}[n] - \frac{1}{3}\mathbf{x}_{\text{odd}}[n-1]$

   Prediction P: $\quad \mathbf{d}^{(1)}[n] = \mathbf{x}_{\text{odd}}[n] - \frac{1}{8}(9\mathbf{s}^{(1)}[n] + 3\mathbf{s}^{(1)}[n+1])$

   Second update U_2: $\quad \mathbf{s}^{(2)}[n] = \mathbf{s}^{(1)}[n] + \frac{4}{9}\mathbf{d}^{(1)}[n]$

   (a) Draw a ¤ow-chart for this transform (as in Ripples).

   (b) Let U_1 be the £rst update transformation: $U_1 \begin{bmatrix} \mathbf{x}_{\text{even}} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{x}_{\text{odd}} \end{bmatrix}$. Write down the matrix for P in 2×2 block form using the $N/2 \times N/2$ identity matrix I and shift matrix S.

   (c) Let P be the prediction transformation: $P \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{x}_{\text{odd}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{d}^{(1)} \end{bmatrix}$. Write down the matrix for P in 2×2 block form using the $N/2 \times N/2$ identity matrix I and shift matrix S.

   (d) Let U_2 be the second update linear transformation: $U_2 \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{d}^{(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(2)} \\ \mathbf{d}^{(1)} \end{bmatrix}$. Write down the matrix for U_2 in 2×2 block form.

5. Consider the CDF(3,1) wavelet transform on $\mathbb{R}^6$.

(a) The one-step analysis matrix $\mathbf{T_a}$ is

$$\frac{1}{4\sqrt{2}}\begin{bmatrix} 6 & 6 & -1 & 0 & 0 & -2 \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ -3 & 3 & -1 & 0 & 0 & 1 \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \end{bmatrix}$$

Fill in the missing entries.

(b) The one-step synthesis matrix $\mathbf{T_s}$ is

$$\frac{1}{4\sqrt{2}}\begin{bmatrix} 3 & \_\_\_ & \_\_\_ & -6 & \_\_\_ & \_\_\_ \\ 3 & \_\_\_ & \_\_\_ & 6 & \_\_\_ & \_\_\_ \\ 1 & \_\_\_ & \_\_\_ & 2 & \_\_\_ & \_\_\_ \\ 0 & \_\_\_ & \_\_\_ & 0 & \_\_\_ & \_\_\_ \\ 0 & \_\_\_ & \_\_\_ & 0 & \_\_\_ & \_\_\_ \\ 1 & \_\_\_ & \_\_\_ & -2 & \_\_\_ & \_\_\_ \end{bmatrix}$$

Fill in the missing entries.

6. The unnormalized 2×2 Haar analysis matrix is $\mathbf{W_a} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

   (a) Calculate the one-scale Haar wavelet transform $\mathbf{Y}$ of the matrix $\mathbf{X} = \begin{bmatrix} 2 & 4 \\ 0 & 8 \end{bmatrix}$.

   (b) Calculate the *multiresolution representation* $\mathbf{X} = \mathbf{X_{ss}} + \mathbf{X_{sd}} + \mathbf{X_{ds}} + \mathbf{X_{dd}}$.

Chapter 4

Wavelet Transforms by Two-channel Filter Banks

4.1 Finite Signals and the z-Transform

Let $\mathbf{x}$ be a real-valued function on the integers $\mathbb{Z}$. Assume that $\mathbf{x}$ has *£nite support*: there are integers $p \le q$ so that $\mathbf{x}[n] = 0$ when $n < p$ or $n > q$. We call such a function a *signal*. If $\mathbf{x} \neq 0$ and we choose p, q so that $\mathbf{x}[p] \neq 0$ and $\mathbf{x}[q] \neq 0$, then we call the integer $q - p + 1$ the *length* of the signal (the zero signal has length 0).

De£ne δ_k (the *unit impulse* at k) by

$$\delta_k[n] = \begin{cases} 1 & \text{if } n = k, \\ 0 & \text{if } n \neq k. \end{cases}$$

Every signal $\mathbf{x}$ can be written uniquely as a linear combination of unit impulses:

$$\mathbf{x} = \sum_{k \in \mathbb{Z}} \mathbf{x}[k] \delta_k \,.$$

The set of all signals is a real vector space (in£nite-dimensional) with a basis given by the unit impulses.

Let $\mathbf{x}$ be a signal. We de£ne the *z-transform* of $\mathbf{x}$ to be

$$X(z) = \sum_{n \in \mathbb{Z}} \mathbf{x}[n]\, z^{-n} \quad \text{where } z \in \mathbb{C} \text{ and } z \neq 0.$$

Since $\mathbf{x}$ has £nite support, there are only a £nite number of nonzero terms in the sum, and $X(z)$ is a *Laurent polynomial* (£nite linear combination of positive and negative powers of z). Suppose $\mathbf{x}[n] = 0$ for $n < p$ and for $n > q$. Then we can write the z-transform of $\mathbf{x}$ as

$$X(z) = \mathbf{x}[p]\, z^{-p} + \mathbf{x}[p+1]\, z^{-p-1} + \cdots + \mathbf{x}[q]\, z^{-q}$$

(the sum has at most $q - p + 1$ nonzero terms).

When $z = \mathrm{e}^{\mathrm{i}\,\omega}$ has absolute value 1 (with ω real), then $z^{-n} = \mathrm{e}^{-\mathrm{i}\,n\omega}$ and

$$X(\mathrm{e}^{\mathrm{i}\,\omega}) = \sum_{n \in \mathbb{Z}} \mathbf{x}[n]\, \mathrm{e}^{-\mathrm{i}\,n\omega} \quad \text{(£nite Fourier series).}$$

In signal-processing language ω is the *frequency* variable, whereas n is the *discrete time* variable.

Example 4.1.1. Suppose the nonzero values of $\mathbf{x}$ are $\mathbf{x}[-1] = 2$, $\mathbf{x}[0] = 3$, and $\mathbf{x}[2] = 4$. Then

$$\mathbf{x} = 2\delta_{-1} + 3\delta_0 + 4\delta_2,$$

and $\mathbf{x}$ is a signal of length $2 - (-1) + 1 = 4$ (note that $\mathbf{x}[1] = 0$). The z-transform of $\mathbf{x}$ is $X(z) = 2z + 3 + 4z^{-2}$. When $z = \mathrm{e}^{\mathrm{i}\,\omega}$ then the z-transform becomes the £nite Fourier series:

$$X(\mathrm{e}^{\mathrm{i}\,\omega}) = 2\mathrm{e}^{\mathrm{i}\,\omega} + 3 + 4\mathrm{e}^{-2\,\mathrm{i}\,\omega}$$

■

The transformation from $\mathbf{x}$ to $X(z)$ is linear: adding signals $\mathbf{x}$ and $\mathbf{y}$ or multiplying $\mathbf{x}$ by a real number corresponds to the same operations on the z-transforms $X(z)$ and $Y(z)$. Every £nite Laurent polynomial with real coef£cients is the z transform of a unique signal.

Given signals $\mathbf{x}$ and $\mathbf{y}$, we de£ne their *inner product* to be

$$\langle \mathbf{x},\, \mathbf{y} \rangle = \sum_{n \in \mathbb{Z}} \mathbf{x}[n]\, \mathbf{y}[n]$$

(by the £nite support condition, there are only £nitely many nonzero terms in the sum).

De£ne the *energy* $||\mathbf{x}||^2$ of $\mathbf{x}$ by

$$||\mathbf{x}||^2 = \langle \mathbf{x},\, \mathbf{x} \rangle = \sum_{n \in \mathbb{Z}} \mathbf{x}[n]^2.$$

The energy in $\mathbf{x}$ is expressed in terms of the z-transform by *Parseval's relation:*

$$||\mathbf{x}||^2 = \frac{1}{2\pi} \int_0^{2\pi} |X(\mathrm{e}^{\mathrm{i}\,\omega})|^2 \, d\omega \tag{4.1}$$

This follows from equation (1.8) because the function $f(\omega) = X(\mathrm{e}^{\mathrm{i}\,\omega})$ has Fourier coef£cients $\mathbf{x}[-n]$:

$$\frac{1}{2\pi} \int_0^{2\pi} f(\omega) \mathrm{e}^{-\,\mathrm{i}\,n\omega} \, d\omega = \sum_k \mathbf{x}[k] \left\{ \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{-\,\mathrm{i}\,(k+n)\omega} \, d\omega \right\} = \mathbf{x}[-n]$$

(the integral in brackets is zero except when $k + n = 0$, and then it is 1).

Discrete Fourier Transform and z-Transform

Let $\mathbf{x}$ be a £nite signal. Assume that $p \leq q$ are integers so that all the nonzero values of $\mathbf{x}[n]$ occur for $p \leq n \leq q$, and that $\mathbf{x}[p] \neq 0$, $\mathbf{x}[q] \neq 0$. Thus $\mathbf{x}$ has length $q - p + 1$. Fix any integer $N \geq q - p + 1$. Then we de£ne the N*-periodic extension* of $\mathbf{x}$ as follows: Every integer n can be expressed uniquely as $n = k + mN$ where $p \leq k < p + N$ and $m \in \mathbb{Z}$. We de£ne

$$\mathbf{x}_{\mathrm{per},N}[k + mN] = \begin{cases} \mathbf{x}[k] & \text{if } p \leq k \leq q, \\ 0 & \text{if } q < k < p + N. \end{cases}$$

This de£nition makes sense because $q < p + N$ (since $q - p < N$). It is clear that $\mathbf{x}_{\mathrm{per},N}[n + N] = \mathbf{x}_{\mathrm{per},N}[n]$ for all $n \in \mathbb{Z}$.

We identify $\mathbf{x}_{\mathrm{per},N}$ with the N-component column vector

$$\begin{bmatrix} \mathbf{x}_{\mathrm{per},N}[0] \\ \mathbf{x}_{\mathrm{per},N}[1] \\ \vdots \\ \mathbf{x}_{\mathrm{per},N}[N-1] \end{bmatrix}$$

Notice that we have inserted zeros in the vector $\mathbf{x}_{\mathrm{per},N}$ as needed to obtain a vector with N components (this is called *zero padding*). The N-periodic signal has the same energy as the original signal:

$$||\mathbf{x}_{\mathrm{per},N}||^2 = ||\mathbf{x}||^2$$

since both vectors have the same nonzero entries.

Theorem 4.1.2. *The discrete Fourier transform of $\mathbf{x}_{\mathrm{per},N}$ is obtained by sampling the z-transform $X(z)$ at the Nth roots of unity (going* counterclockwise *around the unit circle):*

$$\widehat{\mathbf{x}}_{\mathrm{per},N}[k] = X(w^k) \quad \textit{for } k = 0, 1, \ldots, N-1, \textit{ where } w = \mathrm{e}^{2\pi\,\mathrm{i}\,k/N}.$$

Proof. The discrete Fourier transform of an N-periodic function can be calculated by summing over any set of representatives of the integers modulo N. If we use the set $p \le n \le p + N$ and the de£nition of $\mathbf{x}_{\mathrm{per},N}$, we obtain

$$\begin{aligned} \widehat{\mathbf{x}}_{\mathrm{per},N}[k] &= \sum_{p \le n < p+N} \mathbf{x}_{\mathrm{per},N}[n]\, w^{-nk} \\ &= \sum_{p \le n \le q} \mathbf{x}[n]\, w^{-nk} = X(w^k). \end{aligned}$$

Notice that in the second sum we only need the values $p \le n \le q$ since $q < p + N$ and $\mathbf{x}[n] = 0$ for $q < n < p + N$. ∎

Example 4.1.3. Suppose $\mathbf{x}$ is the function in Example 4.1.1. Then $p = -1$ and $q = 2$, so we can make an N-periodic extension of $\mathbf{x}$ for any integer $N \ge 4$. The 4-periodic extension has values

$$\mathbf{x}_{\mathrm{per},4}[0] = 3, \quad \mathbf{x}_{\mathrm{per},4}[1] = 0, \quad \mathbf{x}_{\mathrm{per},4}[2] = 4, \quad \mathbf{x}_{\mathrm{per},4}[3] = -2$$

(the value at 3 is $\mathbf{x}[-1]$ since $3 = -1 + 4$). Since $\mathrm{e}^{2\pi\,\mathrm{i}/4} = \mathrm{i}$, we have

$$\widehat{\mathbf{x}}_{\mathrm{per},4}[k] = X(\mathrm{i}^k) \quad \text{for } k = 0, 1, 2, 3$$

in this case. For this example $X(z) = 2z + 3 + 4z^{-2}$; thus Theorem 4.1.2 shows that $\widehat{\mathbf{x}}_{\mathrm{per},4}$ corresponds to the column vector

$$\begin{bmatrix} X(1) \\ X(\mathrm{i}) \\ X(-1) \\ X(-\mathrm{i}) \end{bmatrix} = \begin{bmatrix} 9 \\ 2\,\mathrm{i} + 3 - 4 \\ -2 + 3 + 4 \\ -2\,\mathrm{i} + 3 - 4 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 + 2\,\mathrm{i} \\ 5 \\ -1 - 2\,\mathrm{i} \end{bmatrix}$$

We can also de£ne a 5-periodic extension of $\mathbf{x}$. In this case

$$\mathbf{x}_{\mathrm{per},5}[0] = 3, \quad \mathbf{x}_{\mathrm{per},5}[1] = 0, \quad \mathbf{x}_{\mathrm{per},5}[2] = 4, \quad \mathbf{x}_{\mathrm{per},5}[3] = 0, \quad \mathbf{x}_{\mathrm{per},5}[4] = -2,$$

since $3 \equiv -2 \pmod 5$ and $4 \equiv -1 \pmod 5$. Note that $\mathbf{x}_{\mathrm{per},5}[3] \ne \mathbf{x}_{\mathrm{per},4}[3]$. ∎

Convolution

If $\mathbf{x}$ and $\mathbf{y}$ are £nite signals, we de£ne their *convolution* $\mathbf{x} * \mathbf{y}$ as the function

$$(\mathbf{x} * \mathbf{y})[n] = \sum_{j+k=n} \mathbf{x}[j]\mathbf{y}[k].$$

The summation over the pairs (j, k) in this equation can be written in two forms:

$$(\mathbf{x} * \mathbf{y})[n] = \sum_{k \in \mathbb{Z}} \mathbf{x}[n-k]\mathbf{y}[k] = \sum_{j \in \mathbb{Z}} \mathbf{x}[j]\mathbf{y}[n-j].$$

The convolution product can be described in words as follows: take the function $\mathbf{x}[j]\mathbf{y}[k]$ of the two (discrete) variables (j, k) and add the values of this function at all integer points along the diagonal line $j + k = n$. If $|n|$ is suf£ciently large, this diagonal line does not intersect the set of points where $\mathbf{x}[j]\mathbf{y}[k] \neq 0$, and hence $\mathbf{x} * \mathbf{y}[n] = 0$ in this case. Thus $\mathbf{x} * \mathbf{y}$ is a £nite signal. From the de£nition we see that $\mathbf{x} * \mathbf{y} = \mathbf{y} * \mathbf{x}$.

If $\mathbf{x} = \delta_p$ and $\mathbf{y} = \delta_q$ are unit impulses, then the function $\mathbf{x}[j]\mathbf{y}[k]$ is zero except when $j = p$ and $k = q$. Hence we get

$$\delta_p * \delta_q = \delta_{p+q} \tag{4.2}$$

in this case. In general, if $\mathbf{x}[j] = 0$ for $j < p$ or $j > q$ then

$$(\mathbf{x} * \mathbf{y})[n] = \sum_{p \le j \le q} \mathbf{x}[j]\mathbf{y}[n-j].$$

This shows that the value of $\mathbf{x} * \mathbf{y}$ at n depends on the values of $\mathbf{y}$ between $n - p$ and $n - q$. There is no wraparound in this formula, however, and it is not the same as the circular convolution of periodic extensions of $\mathbf{x}$ and $\mathbf{y}$ de£ned in Section 2.2.

Example 4.1.4. Take $\mathbf{x}[n] = 1/3$ for $n = -1, 0, 1$ and zero otherwise. Then

$$(\mathbf{x} * \mathbf{y})[n] = \frac{1}{3}(\mathbf{y}[n-1] + \mathbf{y}[n] + \mathbf{y}[n+1])$$

is a *moving average*. If $\mathbf{y}[n] = 0$ for $n < 0$ and $n > 7$, for example, then $(\mathbf{x} * \mathbf{y})[n] = 0$ for $n < -1$ and $n > 8$. However $(\mathbf{x} * \mathbf{y})[-1] \neq 0$ and $(\mathbf{x} * \mathbf{y})[8] \neq 0$, in general, even though $\mathbf{y}[-1] = 0$ and $\mathbf{y}[8] = 0$. Thus convolution with $\mathbf{x}$ spreads out the support of $\mathbf{y}$. ■

Here is one of the most important properties of convolution.

Theorem 4.1.5. *Let $\mathbf{x}$ and $\mathbf{y}$ be functions on $\mathbb{Z}$ with £nite support. Then the z-transform of $\mathbf{x} * \mathbf{y}$ is the pointwise product $X(z)Y(z)$ of the z-transforms.*

Proof. The de£nition of convolution can be written in terms of unit impulses as

$$\mathbf{x} * \mathbf{y} = \sum_{p,q} \mathbf{x}[p]\mathbf{y}[q]\delta_{p+q}\,.$$

Since the z-transform of the unit impulse δ_{p+q} is $z^{-p-q} = z^{-p}z^{-q}$, we see that the z-transform of $\mathbf{x} * \mathbf{y}$ is

$$\sum_{p,q} \mathbf{x}[p]\mathbf{y}[q]z^{-p-q} = X(z)Y(z).$$

■

Shift Operator

We de£ne the shift operator S applied to a function $\mathbf{x}$ on $\mathbb{Z}$ by

$$(S\mathbf{x})[n] = \mathbf{x}[n-1] \quad \text{for } n \in \mathbb{Z}$$

(the minus sign appears in the formula so that the graph of $S\mathbf{x}$ is obtained by shifting the graph of $\mathbf{x}$ to the *right* one unit). From the de£nition of convolution we see that $S\mathbf{x} = \delta_1 * \mathbf{x}$. More generally, for any integer k we have

$$S^k\mathbf{x} = \delta_k * \mathbf{x}.$$

Hence the shift-invariant linear transformation $\sum_k a_k S^k$ acts by convolution with the function $\sum_k a_k \delta_k$. This has the following important consequence:

Theorem 4.1.6. *Let $\mathbf{x}$ be a signal. Then the z-transform of $S\mathbf{x}$ is $z^{-1}X(z)$. More generally, the z-transform of a linear combination $\sum_k a_k S^k \mathbf{x}$ of shifts of $\mathbf{x}$ is $\left(\sum_k a_k z^{-k}\right) X(z)$.*

Proof. The unit impulse at k obviously has z-transform z^{-k}. Since

$$\sum_k a_k S^k \mathbf{x} = \sum_k a_k \delta_k * \mathbf{x},$$

the theorem follows from Theorem 4.1.5. ■

Periodic Shift Operator

Now that we have used the symbol S to denote the shift operator on nonperiodic signals, we will write S_N to denote the shift operator on periodic functions of period N, to avoid confusion.

Suppose $\mathbf{x}$ is a signal such that $\mathbf{x}[n] = 0$ when $n < p$ or $n > q$. Fix $N > q - p$. Then it is easy to check that

$$S_N \mathbf{x}_{\text{per},N} = (S\mathbf{x})_{\text{per},N}. \tag{4.3}$$

This follows from a basic property of modular arithmetic: if $a \equiv b \pmod N$, then $a + 1 \equiv b + 1 \pmod N$. For example, the signal $\mathbf{x} = 3\delta_3 + 4\delta_4 + 5\delta_5 + 6\delta_6$ has length 4. If we take $N = 4$, then $\mathbf{x}_{\text{per},4}$ corresponds to the column vector $\mathbf{u} = \begin{bmatrix} 4 & 5 & 6 & 3 \end{bmatrix}^{\mathrm{T}}$ (the positions of the components of $\mathbf{u}$ are determined by reading the subscripts on the unit impulse functions modulo 4). The shifted (nonperiodic) signal

$$S\mathbf{x} = 3\delta_4 + 4\delta_5 + 5\delta_6 + 6\delta_7$$

also has length 4 and $(S\mathbf{x})_{\text{per},4}$ corresponds to the column vector $\begin{bmatrix} 3 & 4 & 5 & 6 \end{bmatrix}^{\mathrm{T}} = S_4\mathbf{u}$ (note the wraparound).

Theorem 4.1.7. *The discrete Fourier transform of $S_N \mathbf{x}_{\text{per},N}$ is obtained by sampling $z^{-1}X(z)$ at the Nth roots of unity.*

Proof. Set $w = \mathrm{e}^{2\pi\,\mathrm{i}/N}$ and $\mathbf{y} = S\mathbf{x}$. From Theorem 4.1.2 we have

$$\widehat{\mathbf{y}_{\text{per},N}}[k] = Y(w^k) = w^{-k}X(w^k) = w^{-k}\widehat{\mathbf{x}}_{\text{per},N}[k].$$

Now apply (4.3). ■

4.2 Wavelet Transforms and Polyphase Matrices

We now construct some linear transformations on the vector space of signals that are fundamental for digital signal processing and wavelet theory.

Downsampling and Upsampling

Let $\mathbf{x}$ be a signal. The *downsampling* by 2 of $\mathbf{x}$ is the signal

$$\mathbf{x}_{2\downarrow}[n] = \mathbf{x}[2n] \quad \text{for } n \in \mathbb{Z}.$$

This de£nes a linear transformation: $\boxed{2\downarrow}\,\mathbf{x} = \mathbf{x}_{2\downarrow}$. We denote the z-transform of $\mathbf{x}_{2\downarrow}$ by $X_{2\downarrow}(z)$. It is given by the formula

$$X_{2\downarrow}(z) = \sum_n \mathbf{x}[2n]\, z^{-n} \tag{4.4}$$

The *upsampling* by 2 of $\mathbf{x}$ is the signal

$$\mathbf{x}_{2\uparrow}[n] = \begin{cases} \mathbf{x}[m] & \text{if } n = 2m \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

This de£nes a linear transformation: $\boxed{2\uparrow}\,\mathbf{x} = \mathbf{x}_{2\uparrow}$. We denote the z-transform of $\mathbf{x}_{2\uparrow}$ by $X_{2\uparrow}(z)$. It is given by the formulas

$$X_{2\uparrow}(z) = \sum_n \mathbf{x}[n]\, z^{-2n} = X(z^2). \tag{4.5}$$

The signal $\mathbf{x}_{2\uparrow}$ is a stretched version of $\mathbf{x}$ (with zeros interlaced); we recover $\mathbf{x}$ from $\mathbf{x}_{2\uparrow}$ by downsampling:

$$\boxed{2\downarrow}\ \boxed{2\uparrow}\,\mathbf{x} = \mathbf{x}.$$

Thus the transformation $\boxed{2\downarrow}$ is a *left inverse* to the transformation $\boxed{2\uparrow}$. In the opposite order, we have

$$\left(\boxed{2\uparrow}\ \boxed{2\downarrow}\,\mathbf{x}\right)[n] = \begin{cases} \mathbf{x}[n] & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd.} \end{cases} \tag{4.6}$$

Hence $\boxed{2\uparrow}\ \boxed{2\downarrow}\,\mathbf{x}$ is the projection of $\mathbf{x}$ onto the signals that are zero at all odd integers. If $\mathbf{x}[n] = 0$ when n is even, then $\boxed{2\uparrow}\ \boxed{2\downarrow}\,\mathbf{x} = 0$. Thus $\boxed{2\downarrow}$ is not a right inverse to $\boxed{2\uparrow}$ and neither $\boxed{2\uparrow}$ nor $\boxed{2\downarrow}$ are invertible linear transformations on the vector space of signals. (This is a signi£cant change from the £nite-dimensional case of N-periodic signals).

Example 4.2.1. Suppose the nonzero values of $\mathbf{x}$ are $\mathbf{x}[-1] = 2$, $\mathbf{x}[0] = 3$, $\mathbf{x}[1] = 4$, and $\mathbf{x}[2] = 1$. Then the nonzero values of $\mathbf{x}_{2\downarrow}$ are $\mathbf{x}_{2\downarrow}[0] = 3$ and $\mathbf{x}_{2\downarrow}[1] = 1$. The z-transforms are

$$X(z) = 2z + 3 + 4z^{-1} + z^{-2} \quad \text{and} \quad X_{2\downarrow}(z) = 3 + z^{-1}.$$

If we set $\mathbf{y} = \mathbf{x}_{2\downarrow}$, then the nonzero values of $\mathbf{y}_{2\uparrow}$ are $\mathbf{y}_{2\uparrow}[0] = 3$ and $\mathbf{y}_{2\uparrow}[2] = 1$. Thus

$$Y_{2\uparrow}(z) = Y(z^2) = 3 + z^{-2}$$

(the even terms in $X(z)$). ■

Even-Odd Splitting

In performing a wavelet decomposition of a periodic function $\mathbf{x}$ on $\mathbb{Z}$ of even period N, the £rst operation was to split $\mathbf{x}$ into $\mathbf{x}_{\text{even}}$ and $\mathbf{x}_{\text{odd}}$. We can do a similar splitting for a nonperiodic signal $\mathbf{x}$. De£ne

$$\mathbf{x}_0 = \boxed{2\downarrow}\mathbf{x} \quad \text{and} \quad \mathbf{x}_1 = \boxed{2\downarrow}S^{-1}\mathbf{x},$$

(the operator S^{-1} shifts $\mathbf{x}$ *left* by one unit). This splitting corresponds to the ¤ow chart

$$\begin{array}{ccccccc}
 & & & \boxed{2\downarrow} & \longrightarrow & & \mathbf{x}_0 \\
 & \nearrow & & & & & \\
\mathbf{x} & & & & & & \\
 & \searrow & & & & & \\
 & & \boxed{S^{-1}} & \rightarrow & \boxed{2\downarrow} & \rightarrow & \mathbf{x}_1
\end{array}$$

We can reconstruct $\mathbf{x}$ from $\mathbf{x}_0$ and $\mathbf{x}_1$:

$$\mathbf{x} = \boxed{2\uparrow}\mathbf{x}_0 + S\boxed{2\uparrow}\mathbf{x}_1. \tag{4.7}$$

This relation corresponds to the ¤ow chart

$$\begin{array}{cccccccc}
\mathbf{x}_0 & \longrightarrow & \boxed{2\uparrow} & & & & & \\
 & & & & \searrow & & & \\
 & & & & & \boxed{+} & \longrightarrow & \mathbf{x} \\
 & & & & \nearrow & & & \\
\mathbf{x}_1 & \rightarrow & \boxed{2\uparrow} & \rightarrow & \boxed{S} & & &
\end{array}$$

To prove (4.7), note from (4.6) that

$$\boxed{2\uparrow}\mathbf{x}_0[n] = \begin{cases} \mathbf{x}[n] & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Likewise,

$$\left(S\boxed{2\uparrow}\mathbf{x}_1\right)[n] = \left(\boxed{2\uparrow}\mathbf{x}_1\right)[n-1] = \left(\boxed{2\uparrow}\,\boxed{2\downarrow}S^{-1}\mathbf{x}\right)[n-1] = \begin{cases} 0 & \text{if } n \text{ is even,} \\ \mathbf{x}[n] & \text{if } n \text{ is odd,} \end{cases}$$

since $(S^{-1}\mathbf{x})[n-1] = \mathbf{x}[n]$. Thus the two vectors on the right side of (4.7) £t together like a zipper to give the values of $\mathbf{x}[n]$ for all n. From (4.7) we have the *perfect reconstruction formula*

$$\mathbf{x} = \boxed{2\uparrow}\,\boxed{2\downarrow}\mathbf{x} + S\boxed{2\uparrow}\,\boxed{2\downarrow}S^{-1}\mathbf{x} \tag{4.8}$$

This formula corresponds to the *lazy wavelet transform*:

$$\begin{array}{cccccccccccc}
 & & & \boxed{2\downarrow} & \longrightarrow & \mathbf{x}_0 & \longrightarrow & \boxed{2\uparrow} & & & & \\
 & \nearrow & & & & & & & \searrow & & & \\
\mathbf{x} & & & & & & & & & \boxed{+} & \longrightarrow & \mathbf{x} \\
 & \searrow & & & & & & & \nearrow & & & \\
 & & \boxed{S^{-1}} \rightarrow & \boxed{2\downarrow} & \rightarrow & \mathbf{x}_1 & \rightarrow & \boxed{2\uparrow} & \rightarrow \boxed{S} & & &
\end{array}$$

The signal is split into even/odd parts that are immediately recombined without further transformation.

The z-transforms of $\mathbf{x}_0$ and $\mathbf{x}_1$ are given by

$$\begin{aligned} X_0(z) &= \sum_n \mathbf{x}[2n]z^{-n} = X_{2\downarrow}(z) \\ X_1(z) &= \sum_n \mathbf{x}[2n+1]z^{-n} = (zX(z))_{2\downarrow} \end{aligned}$$

(see Fig. 7.1 in Ripples). Notice in the second formula that we multiply by z *before* downsampling.

We can express $X_0(z)$ and $X_1(z)$ directly in terms of $X(z)$ by introducing the variable $z^{1/2}$:

$$X_0(z) = \frac{1}{2}\left\{X(z^{1/2}) + X(-z^{1/2})\right\} \quad \text{and} \quad X_1(z) = \frac{z^{1/2}}{2}\left\{X(z^{1/2}) - X(-z^{1/2})\right\}. \tag{4.9}$$

The notation in (4.9) means that we substitute $z^{1/2}$ in z-transform of $\mathbf{x}$ and simplify by the usual algebraic rules for exponents:

$$(z^{1/2})^n = z^{n/2} \quad \text{and} \quad (-z^{1/2})^n = (-1)^n z^{n/2}.$$

To verify the correctness of (4.9), observe that the terms with n odd cancel in the formula for $X_0(z)$, whereas the terms with $n = 2m$ even contribute $\mathbf{x}[2m]z^{-m}$. Likewise, in the formula for $X_1(z)$ the terms with n even cancel, whereas the terms with $n = 2m+1$ odd contribute

$$\mathbf{x}[2m+1]z^{1/2}z^{-m-1/2} = \mathbf{x}[2m+1]z^{-m}.$$

Thus (4.9) follows from these observations.

The z-transform version of equation (4.7) is

$$X(z) = X_0(z)_{2\uparrow} + z^{-1}\left(X_1(z)_{2\uparrow}\right) = X_0(z^2) + z^{-1}X_1(z^2) \tag{4.10}$$

(see Fig. 7.2 in Ripples). Notice in this equation that we multiply by z^{-1} *after* upsampling.

Example 4.2.2. Suppose the nonzero values of $\mathbf{x}$ are $\mathbf{x}[-3] = 2$, $\mathbf{x}[0] = 3$, $\mathbf{x}[1] = 4$, and $\mathbf{x}[2] = 1$. Then $X(z) = 2z^3 + 3 + 4z^{-1} + z^{-2}$ and we have

$$\begin{aligned} X_0(z) &= X(z)_{2\downarrow} = 3 + z^{-1}, \\ X_1(z) &= (zX(z))_{2\downarrow} = \left(2z^4 + 3z + 4 + z^{-1}\right)_{2\downarrow} = 2z^2 + 4 \end{aligned}$$

(notice the shift of the exponents). We recover $X(z)$ by (4.10):

$$X_0(z^2) + z^{-1}X_1(z^2) = (3 + z^{-2}) + z^{-1}(2z^4 + 4) = X(z).$$

■

Lifting and the Polyphase Matrix

We return to the one-scale analysis and synthesis wavelet transformations. The formulas for analysis and synthesis transforms (such as CDF(2,2) and Daub4) are all expressed in terms of the shift operator S (but without the assumption of periodicity)—see Ripples, §3.6. The splitting of a periodic signal of period N into even and odd signals of period $N/2$ is now replaced by the downsampling of the signal $\mathbf{x}$ into the pair of signals $\mathbf{x}_0$ and $\mathbf{x}_1$ (this is called the *polyphase decomposition* of the signal). Since S becomes the operator of multiplication by z^{-1} when we use z-transforms, it is easy to calculate the prediction and update steps in terms of the z-transforms of $X_0(z)$ and $X_1(z)$.

Example 4.2.3 (CDF(2,2))**.** This transform consists of a prediction step, followed by an update and a normalization:

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} = \begin{bmatrix} \sqrt{2}I & 0 \\ 0 & (1/\sqrt{2})I \end{bmatrix} \begin{bmatrix} I & \frac{1}{4}(I+S) \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}(I+S^{-1}) & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{bmatrix}$$

Thus the CDF(2,2) analysis transform becomes a matrix multiplication

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \mathbf{H}_p(z) \begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix},$$

on the vector of z-transforms of the downsampled signal. Here the *analysis polyphase matrix* $\mathbf{H}_p(z)$ is de£ned by

$$\begin{aligned} \mathbf{H}_p(z) &= \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{4}(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1+z) & 1 \end{bmatrix} \\ &= \frac{\sqrt{2}}{8} \begin{bmatrix} (-z+6-z^{-1}) & (2+2z^{-1}) \\ -(2+2z) & 4 \end{bmatrix}. \end{aligned}$$

The inverse transform is obtained by inverting the individual lifting steps:

$$\begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ \frac{1}{2}(I+S^{-1}) & I \end{bmatrix} \begin{bmatrix} I & -\frac{1}{4}(I+S) \\ 0 & I \end{bmatrix} \begin{bmatrix} (1/\sqrt{2})I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix}$$

In terms of z-transforms, this transformation becomes

$$\begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} = \mathbf{G}_p(z) \begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix},$$

where the *synthesis polyphase matrix* $\mathbf{G}_p(z)$ is de£ned by

$$\begin{aligned} \mathbf{G}_p(z) &= \begin{bmatrix} 1 & 0 \\ \frac{1}{2}(1+z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{4}(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \\ &= \frac{\sqrt{2}}{8} \begin{bmatrix} 4 & -(2+2z^{-1}) \\ (2+2z) & (-z+6-z^{-1}) \end{bmatrix} \end{aligned}$$

Since $\mathbf{G}_p(z)$ is the inverse matrix to $\mathbf{H}_p(z)$, the formula for $\mathbf{G}_p(z)$ can be obtained directly from the formula for $\mathbf{H}_p(z)$ using *Cramer's Rule:*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{where } \delta = ad - bc. \tag{4.11}$$

(the lifting-step factorization of $\mathbf{H}_p(z)$ shows that it has determinant one). We will use this formula for the inverse of a 2×2 matrix several times. ■

The general one-scale wavelet analysis and synthesis transforms are of the same form as the CDF(2,2) transform (see the discussion on p. 67 of Ripples). The *analysis polyphase matrix* is a product

$$\mathbf{H}_p(z) = \begin{bmatrix} \kappa & 0 \\ 0 & 1/\kappa \end{bmatrix}\begin{bmatrix} 1 & F_1(z) \\ 0 & 1 \end{bmatrix}\cdots\begin{bmatrix} 1 & 0 \\ F_k(z) & 1 \end{bmatrix} = \begin{bmatrix} H_{00}(z) & H_{01}(z) \\ H_{10}(z) & H_{11}(z) \end{bmatrix}, \tag{4.12}$$

where $F_1(z), \ldots, F_k(z)$ and $H_{ij}(z)$ are Laurent polynomials and $\kappa \neq 0$ (there may be several prediction and update factors). The signal $\mathbf{x}$ is downsampled into $\mathbf{x}_0$ and $\mathbf{x}_1$ and this pair of signals is transformed into the pair $\mathbf{y}_0$ and $\mathbf{y}_1$, whose z-transforms are

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \mathbf{H}_p(z)\begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix}.$$

This transformation is described by the flow chart

$$\begin{tikzcd} & & \boxed{2\downarrow} \arrow[r] & \mathbf{X}_0(z) \arrow[dr] & & \mathbf{Y}_0(z) \\ \mathbf{X}(z) \arrow[urr] \arrow[dr] & & & & \boxed{\mathbf{H}_p(z)} \arrow[ur] \arrow[dr] & \\ & \boxed{z} \arrow[r] & \boxed{2\downarrow} \arrow[r] & \mathbf{X}_1(z) \arrow[ur] & & \mathbf{Y}_1(z) \end{tikzcd}$$

Notice that the polyphase matrix acts on the z-transform of the signal *after* the signal has been split, and the splitting involves a time shift in one channel.

The inverse transform is obtained using the *synthesis polyphase matrix*:

$$\mathbf{G}_p(z) = \mathbf{H}_p(z)^{-1} = \begin{bmatrix} 1 & 0 \\ -F_k(z) & 1 \end{bmatrix}\cdots\begin{bmatrix} 1 & -F_1(z) \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1/\kappa & 0 \\ 0 & \kappa \end{bmatrix}.$$

The flow chart for the inverse transform is

$$\begin{tikzcd} \mathbf{Y}_0(z) \arrow[dr] & & \mathbf{X}_0(z) \arrow[r] & \boxed{2\uparrow} \arrow[drr] & & \\ & \boxed{\mathbf{G}_p(z)} \arrow[ur] \arrow[dr] & & & & \boxed{+} \arrow[r] & \mathbf{X}(z) \\ \mathbf{Y}_1(z) \arrow[ur] & & \mathbf{X}_1(z) \arrow[r] & \boxed{2\uparrow} \arrow[r] & \boxed{z^{-1}} \arrow[ur] & \end{tikzcd}$$

Notice that the inverse polyphase matrix acts on the two z-transforms *before* they are upsampled, shifted, and combined.

We can calculate $\mathbf{G}_p(z)$ directly from the four Laurent polynomials in the matrix $\mathbf{H}_p(z)$ using Cramer's rule:

$$\mathbf{G}_p(z) = \begin{bmatrix} H_{11}(z) & -H_{01}(z) \\ -H_{10}(z) & H_{00}(z) \end{bmatrix} \tag{4.13}$$

(note that the normalization matrix has been chosen to ensure $\det \mathbf{H}_p(z) = 1$). The *perfect reconstruction* (PR) property holds:

$$\begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} = \mathbf{G}_p(z) \begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix}.$$

4.3 Filter Banks and Modulation Matrices

The de£nition of wavelet transforms using the lifting method and the polyphase matrix assures that the PR property always holds, since each step of the lifting process (prediction, update, normalization) uses an invertible elementary matrix. However, this approach doesn't explain how the lifting steps are chosen to obtain desirable properties in the wavelet transform, such as separation of trend and detail. To understand this aspect, we need an alternate description of wavelet transforms using ideas from signal processing (low-pass and high-pass £lters).

FIR Filters

Let $\mathbf{h}$ be a £xed signal. The linear transformation T de£ned by

$$T\mathbf{x} = \mathbf{h} * \mathbf{x}, \quad \text{for all signals } \mathbf{x},$$

is called a *£nite impulse response £lter* (FIR £lter). If we take $\mathbf{x} = \delta_0$ (the unit impulse at 0), then

$$T\delta_0[n] = \sum_k \mathbf{h}[n-k]\delta_0[k] = \mathbf{h}[n]$$

by de£nition of δ_0. Thus $\mathbf{h}$ is uniquely determined by T and is called the *impulse response function* of the £lter. Since $\mathbf{h}$ has £nite support, it follows that $T\mathbf{x}$ also has £nite support for every signal $\mathbf{x}$. Write $H(z)$ for the z-transform of $\mathbf{h}$. Then the z-transform of $T\mathbf{x}$ is $H(z)X(z)$ by Theorem 4.1.5. Thus the action of the £lter on the z-transform is to multiply by the function $H(z)$. Parseval's relation (4.1) gives

$$||T\mathbf{x}||^2 = \frac{1}{2\pi}\int_0^{2\pi} |H(\mathrm{e}^{\mathrm{i}\omega})|^2 |X(\mathrm{e}^{\mathrm{i}\omega})|^2 \, d\omega. \tag{4.14}$$

This shows, for example, that if $|H(\mathrm{e}^{\mathrm{i}\omega})| = 1$ for $0 \le \omega \le 2\pi$, then $||T\mathbf{x}||^2 = ||\mathbf{x}||^2$ (thus T is energy-preserving in this case).

Example 4.3.1. The transformation S^p, where S is the shift operator and p is an integer, is a FIR £lter, with $\mathbf{h} = \delta_p$:

$$(\delta_p * \mathbf{x})[n] = \sum_k \delta_p[k]\mathbf{x}[n-k] = \mathbf{x}[n-p] = (S^p\mathbf{x})[n] \tag{4.15}$$

The z-transform of δ_p is z^{-p}. Since $|\mathrm{e}^{\,\mathrm{i}p\omega}| = 1$ for $0 \le \omega \le 2\pi$, we have $||\delta_p * \mathbf{x}||^2 = ||\mathbf{x}||^2$ by (4.14); this is obvious (without z-transforms) in this case since

$$||\delta_p * \mathbf{x}||^2 = \sum_n |\mathbf{x}[n-p]|^2 = \sum_n |\mathbf{x}[n]|^2 = ||\mathbf{x}||^2.$$

If T is any FIR £lter with impulse response function $\mathbf{h}$, then T is a linear combination of powers of the shift operator:

$$T = \sum_p \mathbf{h}[p] S^p$$

(this follows from (4.15) because $\mathbf{h} = \sum_p \mathbf{h}[p]\delta_p$). Thus FIR £lters are the linear transformations of nonperiodic signals that are analogous to $N \times N$ circulant matrices acting on N-periodic signals (see Theorem 2.2.2). ■

Two-channel Filter Banks

A *two-channel analysis £lter* uses two FIR £lters with impulse responses $\mathbf{h}_0$ (*low pass*) and $\mathbf{h}_1$ (*high pass*) to transform the input signal $\mathbf{x}$ into

$$\mathbf{T_a}\mathbf{x} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix}, \quad \text{where } \mathbf{y}_0 = \boxed{2\downarrow}\,\mathbf{h}_0 * \mathbf{x} \quad \text{and} \quad \mathbf{y}_1 = \boxed{2\downarrow}\,\mathbf{h}_1 * \mathbf{x}. \tag{4.16}$$

Since convolution and downsampling are linear processes, this gives a linear transformation $\mathbf{T_a}$ whose output is the pair of signals $\mathbf{y}_0, \mathbf{y}_1$:

$$\begin{array}{lllllll} & & & \boxed{\mathbf{h}_0 *} & \longrightarrow & \boxed{2\downarrow} & \longrightarrow \ \mathbf{y}_0 \\ & & \nearrow & & & & \\ \mathbf{T_a}: & \mathbf{x} & & & & & \\ & & \searrow & & & & \\ & & & \boxed{\mathbf{h}_1 *} & \longrightarrow & \boxed{2\downarrow} & \longrightarrow \ \mathbf{y}_1 \end{array}$$

Thus the entire signal £rst passes through each £lter separately and then the two £ltered signals are downsampled (see Fig. 7.4 in Ripples). By contrast, in the lifting procedure the operations are in the opposite order and there is a shift on one branch.

Since the z-transform of $\mathbf{h}_0 * \mathbf{x}$ is $H_0(z)X(z)$ and the z-transform of $\mathbf{h}_1 * \mathbf{x}$ is $H_1(z)X(z)$, we can use (4.9) to express the z-transforms of $\mathbf{y}_0$ and $\mathbf{y}_1$ as

$$\begin{aligned} Y_0(z) &= \frac{1}{2}\left\{H_0(z^{1/2})X(z^{1/2}) + H_0(-z^{1/2})X(-z^{1/2})\right\}, \\ Y_1(z) &= \frac{1}{2}\left\{H_1(z^{1/2})X(z^{1/2}) + H_1(-z^{1/2})X(-z^{1/2})\right\}. \end{aligned}$$

De£ne the *analysis modulation matrix*

$$\mathbf{H}_m(z) = \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \tag{4.17}$$

(the entries in this matrix are Laurent polynomials). Then the formulas for $Y_0(z)$ and $Y_1(z)$ can be combined into a single vector-matrix equation

$$2\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \mathbf{H}_m(z^{1/2})\begin{bmatrix} X(z^{1/2}) \\ X(-z^{1/2}) \end{bmatrix}. \tag{4.18}$$

The term *modulation* is used to describe the matrix $\mathbf{H}_m(z)$ because replacing z by $-z$ corresponds to a half band *frequency modulation* (shift in frequency) $\omega \to \omega + \pi$ when $z = \mathrm{e}^{\,\mathrm{i}\,\omega}$. The entries $H_0(-z)$ and $H_1(-z)$ in the second column of $\mathbf{H}_m(z)$ are frequency modulations by π of the entries in the £rst column.

A *two-channel synthesis £lter* uses two FIR £lters with impulse responses $\mathbf{g}_0$ (*low pass*) and $\mathbf{g}_1$ (*high pass*) to transform a pair of input signals $\mathbf{y}_0$, $\mathbf{y}_1$ into

$$\widetilde{\mathbf{x}} = \mathbf{T_s}\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} = \mathbf{g}_0 * (\boxed{2\uparrow}\mathbf{y}_0) + \mathbf{g}_1 * (\boxed{2\uparrow}\mathbf{y}_1) \tag{4.19}$$

Since upsampling and convolution are linear processes, this gives a linear transformation $\mathbf{T_s}$ whose output we have denoted as $\widetilde{\mathbf{x}}$:

$$\mathbf{T_s}:\quad \begin{tikzcd} \mathbf{y}_0 \arrow[r] & \boxed{2\uparrow} \arrow[r] & \boxed{\mathbf{g}_0 *} \arrow[dr] & & \\ & & & \boxed{+} \arrow[r] & \widetilde{\mathbf{x}} \\ \mathbf{y}_1 \arrow[r] & \boxed{2\uparrow} \arrow[r] & \boxed{\mathbf{g}_1 *} \arrow[ur] & & \end{tikzcd}$$

Thus each signal is £rst upsampled and then £ltered (see Fig. 7.4 in Ripples). By contrast, for the polyphase synthesis transform the operations are in the opposite order and there is a shift in one branch.

Since the z-transform of $\boxed{2\uparrow}\mathbf{y}_0$ is $Y_0(z^2)$ and the z-transform of $\boxed{2\uparrow}\mathbf{y}_1$ is $Y_1(z^2)$, it follows that the z-transform of $\widetilde{\mathbf{x}}$ is

$$\widetilde{X}(z) = G_0(z)Y_0(z^2) + G_1(z)Y_1(z^2).$$

If we apply the analysis transform $\mathbf{T_a}$ to a signal $\mathbf{x}$ and then apply the synthesis transform $\mathbf{T_s}$ to $\mathbf{T_a}\mathbf{x}$, we obtain a signal $\widetilde{\mathbf{x}}$. We want to express the z-transform $\widetilde{X}(z)$ of the output in terms of the z-transform $X(z)$ of the input. We have just calculated that

$$\begin{aligned} \widetilde{X}(z) &= G_0(z)Y_0(z^2) + G_1(z), \\ \widetilde{X}(-z) &= G_0(-z)Y_0(z^2) + G_1(-z)Y_1(z^2) \end{aligned}$$

(since $(-z)^2 = z^2$). We can write this pair of equations in matrix form as

$$\begin{bmatrix} \widetilde{X}(z) \\ \widetilde{X}(-z) \end{bmatrix} = \mathbf{G}_m(z)\begin{bmatrix} Y_0(z^2) \\ Y_1(z^2) \end{bmatrix},$$

where

$$\mathbf{G}_m(z) = \begin{bmatrix} G_0(z) & G_1(z) \\ G_0(-z) & G_1(-z) \end{bmatrix} \tag{4.20}$$

is called the *synthesis modulation matrix*. Notice that the entries $G_0(-z)$ and $G_1(-z)$ in the second row of $\mathbf{G}_m(z)$ are frequency modulations by π of the entries in the £rst row.

From (4.18) we know that

$$2\begin{bmatrix} Y_0(z^2) \\ Y_1(z^2) \end{bmatrix} = \mathbf{H}_m(z)\begin{bmatrix} X(z) \\ X(-z) \end{bmatrix}.$$

Hence the z-transforms of the input $\mathbf{x}$ and output $\widetilde{\mathbf{x}}$ of a two-channel analysis-synthesis £lter bank are related by

$$2\begin{bmatrix} \widetilde{X}(z) \\ \widetilde{X}(-z) \end{bmatrix} = \mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} X(z) \\ X(-z) \end{bmatrix}. \tag{4.21}$$

We say that the £lter bank has the *perfect reconstruction* (PR) property if $\mathbf{x} = \widetilde{\mathbf{x}}$ for all signals $\mathbf{x}$.

Theorem 4.3.2. *The perfect reconstruction property holds if and only if the modulation matrices satisfy $\mathbf{G}_m(z)\mathbf{H}_m(z) = 2I$. In this case $\mathbf{H}_m(z)\mathbf{G}_m(z) = 2I$.*

Proof. Suppose $\mathbf{G}_m(z)\mathbf{H}_m(z) = 2I$. Then it follows from (4.21) that PR holds. Also, since $\mathbf{G}_m(z) = 2\mathbf{H}_m(z)^{-1}$ (as a 2×2 matrix), we also have $\mathbf{H}_m(z)\mathbf{G}_m(z) = 2I$.

Conversely, if PR holds, take $\mathbf{x} = \delta_0$. Then $X(z) = 1$ and (4.21) implies that

$$2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{4.22}$$

Now take $\mathbf{x} = \delta_1$. Then $X(z) = z^{-1}$ and (4.21) implies that

$$2\begin{bmatrix} z^{-1} \\ -z^{-1} \end{bmatrix} = \mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} z^{-1} \\ -z^{-1} \end{bmatrix}.$$

Multiply this equation by z to obtain

$$2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{4.23}$$

Adding and subtracting equations (4.22) and (4.23), we £nd that

$$\mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{G}_m(z)\mathbf{H}_m(z)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Hence $\mathbf{G}_m(z)\mathbf{H}_m(z) = 2I$. This shows that $\mathbf{H}_m(z)$ is an invertible matrix with inverse $\mathbf{G}_m(z)$. Thus we also have $\mathbf{H}_m(z)\mathbf{G}_m(z) = 2I$. ■

To obtain the main result on PR £lter banks, we need the following algebraic lemma:

Lemma 4.3.3. *Suppose $g(z)$ and $h(z)$ are Laurent polynomials such that $g(z)h(z) = c$ (a nonzero complex number). Then $g(z)$ and $h(z)$ are monomials.*

Proof. We can write $g(z) = c_m z^m + \cdots + c_n z^n$, where $m \le n$, $c_m \ne 0$, and $c_n \ne 0$. Likewise $h(z) = d_p z^p + \cdots + d_q z^q$ where $p \le q$, $d_p \ne 0$, and $d_q \ne 0$. The product is

$$g(z)h(z) = c_m d_p z^{m+p} + \cdots + c_n d_q z^{n+q}.$$

By assumption the right side of this equation is a constant c. Hence $m + p = n + q = 0$, $m = n$, and $p = q$. Thus $g(z) = c_n z^n$ and $h(z) = d_{-n} z^{-n}$, where $c_n d_{-n} = c$. ■

Theorem 4.3.4. *Suppose* $\mathbf{h}_0$ *and* $\mathbf{h}_1$ *are FIR £lters. These £lters are the analysis part of a two-channel FIR £lter bank with perfect reconstruction if and only if the corresponding modulation matrix* $\mathbf{H}_m(z)$ *satis£es*

$$\det \mathbf{H}_m(z) = cz^{2k+1} \quad (c \neq 0 \text{ a constant and } k \text{ an integer}). \tag{4.24}$$

When (4.24) *is satis£ed then the synthesis £lters* $\mathbf{g}_0$ *and* $\mathbf{g}_1$ *in the £lter bank are uniquely determined by the analysis £lters:*

$$G_0(z) = \frac{2}{d(z)} H_1(-z) \quad \text{and} \quad G_1(z) = -\frac{2}{d(z)} H_0(-z), \tag{4.25}$$

where $d(z) = \det \mathbf{H}_m(z) = H_0(z)H_1(-z) - H_0(-z)H_1(z)$.

Proof. Suppose that there exist FIR £lters $\mathbf{g}_0$ and $\mathbf{g}_1$ so that the synthesis £lter bank with these £lters and the analysis £lter bank with £lters $\mathbf{h}_0$ and $\mathbf{h}_1$ give perfect reconstruction. Then the modulation matrices satisfy $\mathbf{G}_m(z)\mathbf{H}_m(z) = 2I$ by Theorem 4.3.2. Hence

$$\det \mathbf{G}_m(z) \det \mathbf{H}_m(z) = \det(2I) = 4.$$

Lemma 4.3.3 implies that $d(z) = \det \mathbf{H}_m(z)$ is a nonzero monomial. Since $d(-z) = -d(z)$, it must be a monomial of odd degree. Furthermore, $\mathbf{G}_m(z) = 2\mathbf{H}_m(z)^{-1}$, so by Cramer's rule

$$\begin{bmatrix} G_0(z) & G_1(z) \\ G_0(-z) & G_1(-z) \end{bmatrix} = \frac{2}{d(z)} \begin{bmatrix} H_1(-z) & -H_0(-z) \\ -H_1(z) & H_0(z) \end{bmatrix} \tag{4.26}$$

Comparing entries in these matrices yields equations (4.25).

Conversely, if $\det \mathbf{H}_m(z)$ is a monomial, then we can de£ne Laurent polynomials $G_0(z)$ and $G_1(z)$ by (4.25). The synthesis modulation matrix is then given by (4.26) and the PR condition is satis£ed. ■

Equations (4.25) show that the low pass synthesis £lter $\mathbf{g}_0$ is obtained from the high pass analysis £lter $\mathbf{h}_1$ by the following operations:

- *half-band frequency shift*: when $z = \mathrm{e}^{\mathrm{i}\,\omega}$ then $-z = \mathrm{e}^{\mathrm{i}\,(\omega+\pi)}$ (recall that for discrete signals the frequency range is $0 \leq \omega \leq 2\pi$);
- *time shift*: multiplication of z-transforms by $z^{-(2k+1)}$ corresponds in the time domain to applying the operator $S^{-(2k+1)}$;
- *rescaling*: multiplication by a constant.

The high pass synthesis £lter $\mathbf{g}_1$ is obtained from the low pass analysis £lter $\mathbf{h}_0$ in the same way.

Construction of PR Filter Banks

The terminology *low pass* and *high pass* for the £lters in a £lter bank describes their frequency response. An *ideal* pair of £lters $\mathbf{h}_0$ (low pass) and $\mathbf{h}_1$ (high pass) would have z-transforms that satisfy

$$H_0(\mathrm{e}^{\mathrm{i}\,\omega}) = \begin{cases} 1 & \text{when } 0 \leq \omega < L, \\ 0 & \text{when } L \leq \omega \leq \pi, \end{cases} \quad \text{and} \quad H_1(\mathrm{e}^{\mathrm{i}\,\omega}) = \begin{cases} 0 & \text{when } 0 \leq \omega < L, \\ 1 & \text{when } L \leq \omega \leq \pi. \end{cases} \tag{4.27}$$

where L is the *crossover frequency* ($0 < L < \pi$) between the two £lters. (Note that the real-valued £lter $\mathbf{h}$ has a Fourier transform satisfying $H(\mathrm{e}^{-\mathrm{i}\omega}) = \overline{H(\mathrm{e}^{\mathrm{i}\omega})}$; since the Fourier transform is periodic of period 2π we only need to specify it in the range $0 \leq \omega \leq \pi$.) For such £lters, the £ltered signal $\mathbf{h}_0 * \mathbf{x}$ only has low frequencies ($|\omega| < L$), since its Fourier transform $H_0(\mathrm{e}^{\mathrm{i}\omega})X(\mathrm{e}^{\mathrm{i}\omega})$ is zero when $L \leq |\omega| < \pi$. Likewise, the £ltered signal $\mathbf{h}_1 * \mathbf{x}$ only has high frequencies ($L \leq |\omega| \leq \pi$).

However, the £lters described by (4.27) have an in£nite number of nonzero coef£cients, since

$$\mathbf{h}_0[n] = \frac{1}{2\pi}\int_{-L}^{L} \mathrm{e}^{\mathrm{i}n\omega}\, d\omega = \frac{\sin(nL)}{n\pi}$$

and $\sin(nL) \neq 0$ for in£nitely many integers n (recall that $0 < L < \pi$).

To obtain FIR £lters we must allow some overlap between the low and high frequency bands; the partial separation into high and low frequencies is made by requiring that

$$H_0(z) = (1+z)^p\varphi(z) \quad \text{and} \quad H_1(z) = (1-z)^q\psi(z) \tag{4.28}$$

for some positive integers p and q, where $\varphi(z)$ and $\psi(z)$ are Laurent polynomials with $\varphi(-1) \neq 0$ and $\psi(1) \neq 0$. This ensures that $H_0(z)$ vanishes at $z = -1$ ($\omega = \pi$) to order p and $H_1(z)$ vanishes at $z = 1$ ($\omega = 0$) to order q. A large value of p means that $H_0(\mathrm{e}^{\mathrm{i}\omega})$ stays very close to zero when ω is near π, while a large value of q means that $H_1(\mathrm{e}^{\mathrm{i}\omega})$ stays very close to zero when ω is near 0. Thus $\mathbf{h}_0$ will be a low pass £lter, and $\mathbf{h}_1$ will be a high pass £lter. We can describe this low frequency/high frequency separation property in terms of the modulation matrix:

$$\mathbf{H}_m(1) \text{ is a diagonal matrix} \iff H_0(-1) = 0 \text{ and } H_1(1) = 0$$

To obtain a two-channel PR £lter bank, the Laurent polynomials $\varphi(z)$ and $\psi(z)$ must be chosen so that (4.24) holds. Write

$$Q(z) = H_0(z)H_1(-z) = (1+z)^{p+q}\varphi(z)\psi(-z).$$

Then $Q(z) - Q(-z)$ is twice the sum of the odd-degree terms in $Q(z)$; thus the PR condition (4.24) is the same as

$$Q(z) \text{ contains exactly } \textit{one} \text{ term of odd degree.} \tag{4.29}$$

Example 4.3.5. Take $\varphi(z) = \psi(z) = 1$ and $p = q = 1$. Then $H_0(z) = 1 + z$ and $H_1(z) = 1 - z$; thus $Q(z) = (1+z)^2 = 1 + 2z + z^2$ in this case. Hence condition (4.29) is satis£ed, and $\mathbf{h}_0$, $\mathbf{h}_1$ are the analysis £lters for a PR £lter bank. Up to a normalizing factor, this is the Haar transform (see Example 4.5.2). Graphs of $|H_0(\mathrm{e}^{\mathrm{i}\omega})|$ and $|H_1(\mathrm{e}^{\mathrm{i}\omega})|$ are shown in Ripples, Fig. 7.5. ■

Example 4.3.6. Take $\varphi(z) = \psi(z) = 1$ and $p + q > 2$. In this case

$$Q(z) = (1+z)^{p+q} = 1 + (p+q)z + \cdots + (p+q)z^{p+q-1} + z^{p+q}.$$

If $p + q$ is even, then $(p+q)z^{p+q-1}$ has odd degree, whereas if $p + q$ is odd, then z^{p+q} has odd degree. So when $p + q > 2$ the polynomial $Q(z)$ has two or more terms of odd degree. Hence condition (4.29) is *not* satis£ed, and $\mathbf{h}_0$, $\mathbf{h}_1$ cannot be the analysis £lters for a PR £lter bank. ■

Example 4.3.7. We modify Example 4.3.6 by taking $\varphi(z) = 1 + bz + cz^2$, $\psi(z) = 1$, and $p = 2$, $q = 2$, where b and c are real parameters to be determined. Then $H_0(z) = (1 + z)^2(1 + bz + cz^2)$ and $H_1(z) = (1 - z)^2$; thus

$$Q(z) = (1 + z)^4(1 + bz + cz^2) = (1 + 4z + 6z^2 + 4z^3 + z^4)(1 + bz + cz^2).$$

The terms of odd degree in $Q(z)$ are

$$(4c + b)z^5 + (4 + 6b + 4c)z^3 + (4 + b)z.$$

So if we take $b = -4$ and $c = 1$, then condition (4.29) is satis£ed, and $\mathbf{h}_0$, $\mathbf{h}_1$ are the analysis £lters for a PR £lter bank. We have

$$\begin{aligned} H_0(z) &= (1 + z)^2(1 - 4z + z^2) = 1 - 2z - 6z^2 - 2z^3 + z^4, \\ H_1(z) &= (1 - z)^2 = 1 - 2z + z^2. \end{aligned}$$

After applying normalizing factors and multiplying $H_0(z)$ by z^{-2} (a time shift that preserves the PR property), we obtain the analysis £lters of the CDF(2,2) transform (see Example 4.5.3). Graphs of $|H_0(\mathrm{e}^{\mathrm{i}\omega})|$ and $|H_1(\mathrm{e}^{\mathrm{i}\omega})|$ are shown in Ripples, Fig. 7.7. ■

4.4 Constructing PR Filter Banks

To construct a two-channel FIR £lter bank with the PR property, it suf£ces to specify two of the four £lters. In Section 4.3 we used the low-pass and high-pass analysis £lters $\mathbf{h}_0$, $\mathbf{h}_1$, and expressed the PR condition as (4.29). We can also express the PR condition in terms of the low-pass analysis and synthesis £lters, as follows:

Theorem 4.4.1. *Suppose $\mathbf{h}_0$ and $\mathbf{g}_0$ are FIR £lters. These £lters are the low-pass part of a two-channel FIR £lter bank with perfect reconstruction if and only if $H_0(-1) = 0$, $G_0(-1) = 0$, and*

$$H_0(z)G_0(z) + H_0(-z)G_0(-z) = 2. \tag{4.30}$$

Conversely, if condition (4.30) *is satis£ed, de£ne FIR £lters $\mathbf{h}_1$ and $\mathbf{g}_1$ by*

$$H_1(z) = zG_0(-z) \quad \text{and} \quad G_1(z) = z^{-1}H_0(-z). \tag{4.31}$$

Then $\mathbf{h}_1$ and $\mathbf{g}_1$ are high-pass £lters, and the set of £lters $\mathbf{h}_0, \mathbf{h}_1$ (analysis) and $\mathbf{g}_0, \mathbf{g}_1$ (synthesis) give a PR £lter bank.

Proof. The left side of (4.30) is the upper left entry in the matrix $\mathbf{H}_m(z)\mathbf{G}_m(z)$. If $\mathbf{h}_0$ and $\mathbf{g}_0$ are low-pass £lters for a PR £lter bank, then $\mathbf{H}_m(z)\mathbf{G}_m(z) = 2I$, and hence (4.30) holds. Conversely, assume condition (4.30) holds and de£ne $H_1(z)$ and $G_1(z)$ by (4.31). Then $H_1(1) = G_0(-1) = 0$ and $G_1(1) = H_0(-1) = 0$, so $\mathbf{h}_1$ and $\mathbf{g}_1$ are high-pass FIR £lters. We calculate

$$H_0(z)H_1(-z) - H_0(-z)H_1(z) = -z\{H_0(z)G_0(z) + H_0(-z)G_0(-z)\} = -2z.$$

Hence the £lter bank has the PR property by Theorem 4.3.4. ■

Suppose that $\mathbf{h}_0$ and $\mathbf{g}_0$ are FIR low-pass £lters. Since $H_0(-1) = 0$ and $G_0(-1) = 0$, we can write $H_0(z) = (1+z)^p\varphi(z)$ and $G_0(z) = (1+z)^q\psi(z)$, where p, q are positive integers and $\varphi(z), \psi(z)$ are Laurent polynomials that do not vanish at $z = -1$. The PR condition (4.30) can be expressed as

$$(1+z)^n f(z) + (1-z)^n f(-z) = 2, \tag{4.32}$$

where $n = p + q$ and $f(z) = \varphi(z)\psi(z)$ does not vanish at $z = -1$.

It will be useful to make a quadratic change of variable

$$y = \frac{1}{4}\big(-z + 2 - z^{-1}\big) = (-4z)^{-1}(1-z)^2 \tag{4.33}$$

(notice that y is unchanged when z is replaced by z^{-1}). To understand the choice of this transformation, we observe that when $z = \mathrm{e}^{\,\mathrm{i}\,\omega}$ then

$$y = \frac{1}{4}\big(-\mathrm{e}^{\,\mathrm{i}\,\omega} + 2 - \mathrm{e}^{-\,\mathrm{i}\,\omega}\big) = \Big(\frac{\mathrm{e}^{\,\mathrm{i}\,\omega/2} - \mathrm{e}^{-\,\mathrm{i}\,\omega/2}}{2i}\Big)^2 = \sin^2\frac{\omega}{2}$$

Thus $1 - y = \cos^2\frac{\omega}{2}$. The values $z = 1$ ($\omega = 0$) and $z = -1$ ($\omega = \pi$) correspond to $y = 0$ and $y = 1$. Furthermore,

$$1 - y = \frac{1}{4}\big(z + 2 + z^{-1}\big) = (4z)^{-1}(1+z)^2. \tag{4.34}$$

Thus the replacement of z by $-z$ (frequency modulation by π) corresponds to replacing y by $1-y$. We now prove the following key algebraic result:

Proposition 4.4.2 (Bezout's Theorem)**.** *For every integer $n \geq 1$ there is a unique polynomial $B_n(y)$ of degree $n-1$ that satis£es*

$$y^n B_n(1-y) + (1-y)^n B_n(y) = 1. \tag{4.35}$$

It is given by

$$B_n(y) = 1 + ny + \frac{n(n+1)}{1\cdot 2}y^2 + \cdots + \binom{n+k-1}{k}y^k + \cdots + \binom{2n-1}{n-1}y^{n-1}. \tag{4.36}$$

Furthermore, $B_n(y) > 0$ for all $y \geq 0$.

Proof. The uniqueness is easy: If $B(y)$ and $\widetilde{B}(y)$ are polynomials of degree $n-1$ that satisfy (4.35), then $C(y) = B(y) - \widetilde{B}(y)$ satis£es

$$y^n C(1-y) + (1-y)^n C(y) = 0.$$

Hence $C(y) = -y^n\left\{(1-y)^{-n}C(1-y)\right\}$ vanishes to order n at $y = 0$. Since $C(y)$ is a polynomial of degree $n-1$, this order of vanishing is only possible if $C(y) = 0$.

The formula for $B_n(y)$ (assuming existence) uses a similar argument: Multiply equation (4.35) by $(1-y)^{-n}$ and regroup the terms:

$$(1-y)^{-n} - B_n(y) = y^n D(y)$$

where $D(y) = (1-y)^{-n} B_n(1-y)$. This equation shows that

$$\left(\frac{d}{dy}\right)^k B_n(y)\bigg|_{y=0} = \left(\frac{d}{dy}\right)^k (1-y)^{-n}\bigg|_{y=0} = n(n+1)\cdots(n+k-1)$$

for $k = 0, 1, \ldots, n-1$. Hence the right side of (4.36) is the Taylor polynomial of $B_n(y)$ of degree $n-1$ around $y = 0$. But $B_n(y)$, having degree $n-1$, is equal to this Taylor polynomial. Since all the binomial coeff£cients in this polynomial are positive and the constant term is 1, we see from (4.36) that $B_n(y) > 0$ for all $y \geq 0$. We shall prove the existence of $B_n(y)$ at the end of this section. ■

We now apply Bezout's Theorem to construct the CDF(p,q) family of wavelet transforms, where p and q are positive integers and $p + q = 2n$ is even. Using (4.33) and (4.34), we write the Bezout equation (4.35) in terms of z as

$$(1+z)^{2n}(4z)^{-n} B_n\left(\frac{-z+2-z^{-1}}{4}\right) + (1-z)^{2n}(-4z)^{-n} B_n\left(\frac{z+2+z^{-1}}{4}\right) = 1.$$

Thus if we de£ne

$$H_0(z) = \frac{\sqrt{2}}{2^p}(1+z)^p \tag{4.37}$$

$$G_0(z) = \frac{\sqrt{2}}{2^q}(1+z)^q z^{-(p+q)/2} B_{(p+q)/2}\left(\frac{-z+2-z^{-1}}{4}\right) \tag{4.38}$$

then the low-pass £lters $\mathbf{h}_0$ and $\mathbf{g}_0$ satisfy the conditions of Theorem 4.4.1 (the factors of $\sqrt{2}$ are needed to change the right side of the Bezout equation from 1 to 2). Taking the high-pass £lters to have z transforms $H_1(z) = zG_0(-z)$ and $G_1(z) = z^{-1}H_0(-z)$, as in (4.31), we obtain the £lters for the CDF(p,q) transform (see Example 4.5.3 for the CDF(2,2) £lters). The parameters (p, q) give the orders of vanishing of $H_0(z)$ and $G_0(z)$ at $z = -1$.

Remark. Both factors of $\sqrt{2}$ can be put on one of the CDF low-pass £lters; when this is done all the £lter coeff£cients become rational numbers with denominators that are powers of 2 (bit shifts), since the polynomial $B_n(y)$ has integer coeff£cients.

Example 4.4.3 (CDF(3,1) Transform)**.** Take $p = 3$ and $q = 1$ in (4.37). Then $(p+q)/2 = 2$ and $B_2(y) = 1 + 2y$. Since $-z + 2 - z^{-1} = -z^{-1}(z-1)^2$, we have

$$B_2\left(\frac{-z+2-z^{-1}}{4}\right) = 1 - \frac{1}{2}z^{-1}(z-1)^2 = \frac{1}{2}\left\{-z+4-z^{-1}\right\}.$$

Thus the low-pass £lters have z transforms

$$H_0(z) = \frac{\sqrt{2}}{8}(z+1)^3 = \frac{\sqrt{2}}{8}\left\{z^3 + 3z^2 + 3z + 1\right\}$$

$$G_0(z) = \frac{\sqrt{2}}{4}(z+1)z^{-2}\left\{-z+4-z^{-1}\right\} = \frac{\sqrt{2}}{4}\left\{-1 + 3z^{-1} + 3z^{-2} - z^{-3}\right\}.$$

The high-pass £lters have z transforms

$$\begin{aligned} H_1(z) &= zG_0(-z) = \frac{\sqrt{2}}{4}\{-z-3+3z^{-1}+z^{-2}\} \\ G_1(z) &= z^{-1}H_0(-z) = \frac{\sqrt{2}}{8}\{-z^2+3z-3+z^{-1}\}. \end{aligned}$$

Thus the £lters all have length four and are the following linear combinations of unit impulses:

$$\mathbf{h}_0 = \frac{\sqrt{2}}{8}\big(\delta_{-3}+3\delta_{-2}+3\delta_{-1}+\delta_0\big) \qquad \mathbf{h}_1 = \frac{\sqrt{2}}{4}\big(-\delta_{-1}-3\delta_0+3\delta_1+\delta_2\big)$$

$$\mathbf{g}_0 = \frac{\sqrt{2}}{4}\big(-\delta_0+3\delta_1+3\delta_2-\delta_3\big) \qquad \mathbf{g}_1 = \frac{\sqrt{2}}{8}\big(-\delta_{-2}+3\delta_{-1}-3\delta_0+\delta_1\big)$$

■

Completion of proof of Bezout's Theorem. To prove the existence of $B_n(y)$ we only have to £nd polynomials $c(y)$ and $d(y)$ of degree at most $n-1$ that satisfy

$$c(y)y^n + d(y)(1-y)^n = 1. \tag{4.39}$$

Once we have such polynomials, then $B(y) = \frac{1}{2}(c(1-y)+d(y))$ has degree at most $n-1$ and will satisfy (4.35).

The existence of $c(y)$ and $d(y)$ follows from the Euclidean division algorithm for polynomials. It is convenient to express this algorithm in terms of upper triangular and lower triangular matrices applied to the column vector $\begin{bmatrix} y^n \\ (1-y)^n \end{bmatrix}$.

We illustrate the method with the case $n=2$. First we choose a unit lower trangular matrix (prediction step) to reduce the degree of the second component of the vector:

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} y^2 \\ (1-y)^2 \end{bmatrix} = \begin{bmatrix} y^2 \\ 1-2y \end{bmatrix}.$$

Next, we use a unit upper-triangular matrix (update step), chosen to reduce the degree of the £rst component:

$$\begin{bmatrix} 1 & \frac{1}{2}y \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y^2 \\ 1-2y \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y \\ 1-2y \end{bmatrix}.$$

We repeat with a unit lower-triangular matrix (prediction step) chosen to lower the degree of the second component:

$$\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2}y \\ 1-2y \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y \\ 1 \end{bmatrix}.$$

We stop at this point, since the second component of the vector is now a constant. Combining these transformations, we have

$$\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{2}y \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} y^2 \\ (1-y)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y \\ 1 \end{bmatrix}.$$

Write the product of the prediction and update matrices as

$$\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2}y \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} a(y) & b(y) \\ c(y) & d(y) \end{bmatrix},$$

where the entries $a(y), b(y), c(y), d(y)$ are polynomials of degree at most 1. Since

$$\begin{bmatrix} a(y) & b(y) \\ c(y) & d(y) \end{bmatrix} \begin{bmatrix} y^2 \\ (1-y)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y \\ 1 \end{bmatrix},$$

it follows that $c(y)y^2 + d(y)(1-y)^2 = 1$, as needed. Notice that we only need the existence of the polynomials $c(y)$ and $d(y)$ and the fact that they are of degree $n-1 = 1$, but not any explicit formula for them.

The general case follows the same pattern. Let $f(y)$ and $g(y)$ be polynomials of degrees m and n. If $m \geq n$, there is a polynomial $h(y)$ of degree $m-n$ so that $f(y) - g(y)h(y)$ has degree less than $m-n$ (this is the *Euclidean division algorithm*). We can express this in matrix form as

$$\begin{bmatrix} 1 & 0 \\ -h(y) & 1 \end{bmatrix} \begin{bmatrix} g(y) \\ f(y) \end{bmatrix} = \begin{bmatrix} g(y) \\ f(y) - g(y)h(y) \end{bmatrix}.$$

If $m \leq n$, we divide $g(y)$ by $f(y)$ and use a unit upper triangular matrix to replace the component $g(y)$ by $g(y) - f(y)h(y)$. Starting with $g(y) = y^n$ and $f(y) = (1-y)^n$, we repeat this process until we have transformed one component of the vector into a constant. Multiplying all the unit upper/lower triangular matrices, we obtain a 2×2 matrix with determinant 1 and entries that are polynomials of degree at most n, such that

$$\begin{bmatrix} a(y) & b(y) \\ c(y) & d(y) \end{bmatrix} \begin{bmatrix} y^n \\ (1-y)^n \end{bmatrix} = \begin{bmatrix} \varphi(y) \\ \alpha \end{bmatrix},$$

where α is a constant and $\varphi(y)$ is a polynomial of degree at least one. (We multiply by the rotation matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ if necessary to ensure that α is the second component.) By Cramer's rule,

$$\begin{bmatrix} d(y) & -b(y) \\ -c(y) & a(y) \end{bmatrix} \begin{bmatrix} \varphi(y) \\ \alpha \end{bmatrix} = \begin{bmatrix} y^n \\ (1-y)^n \end{bmatrix}.$$

Hence the polynomial $d(y)$ has degree at most $n-1$, since $d(y)\varphi(y) - \alpha b(y) = y^n$ and $b(y)$ has degree at most n. Furthermore, we have the relation $c(y)y^n + d(y)(1-y)^n = \alpha$, and hence $c(y)$ also has degree at most $n-1$. The constant α cannot be zero: as we showed in the proof of uniqueness, that would force $d(y) = 0$ (since the degree of $d(y)$ is at most $n-1$), and hence $c(y) = 0$. But then the matrix

$$\begin{bmatrix} a(y) & b(y) \\ c(y) & d(y) \end{bmatrix}$$

would have determinant zero, which is a contradiction. Thus we can divide $c(y)$ and $d(y)$ by α to obtain relation (4.39). ■

4.5 Comparison of Lifting and Filter Banks

A one-scale wavelet analysis transform can be implemented in two ways:

Lifting: The signal $\mathbf{x}$ is split by downsampling into $\mathbf{x}_0 = \boxed{2\downarrow}\,\mathbf{x}$ and $\mathbf{x}_1 = \boxed{2\downarrow}\,S^{-1}\mathbf{x}$. Then lifting steps (predictions, updates, and a normalization) are applied to $\begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{bmatrix}$ to give an output $\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix}$. The z-transform of the output is

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \mathbf{H}_p(z) \begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} = \frac{1}{2}\mathbf{H}_p(z) \begin{bmatrix} X(z^{1/2}) + X(-z^{1/2}) \\ z^{1/2}X(z^{1/2}) - z^{1/2}X(-z^{1/2}) \end{bmatrix},$$

where $\mathbf{H}_p(z) = \begin{bmatrix} H_{00}(z) & H_{01}(z) \\ H_{10}(z) & H_{11}(z) \end{bmatrix}$ is the polyphase analysis matrix. Replacing z by z^2 in these equations, we get

$$\begin{bmatrix} Y_0(z^2) \\ Y_1(z^2) \end{bmatrix} = \frac{1}{2}\mathbf{H}_p(z^2) \begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix} \begin{bmatrix} X(z) \\ X(-z) \end{bmatrix}. \tag{4.40}$$

Two-channel Filter Bank: First the signal $\mathbf{x}$ is £ltered by $\mathbf{h}_0$ and by $\mathbf{h}_1$. Then the two £ltered signals are downsampled to give the output $\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} = \begin{bmatrix} \boxed{2\downarrow}(\mathbf{h}_0 * \mathbf{x}) \\ \boxed{2\downarrow}(\mathbf{h}_1 * \mathbf{x}) \end{bmatrix}$. The z-transform of the output is

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} H_0(z^{1/2})X(z^{1/2}) + H_0(-z^{1/2})X(-z^{1/2}) \\ H_1(z^{1/2})X(z^{1/2}) + H_1(-z^{1/2})X(-z^{1/2}) \end{bmatrix}.$$

Replacing z by z^2 in this equation, we get the relation

$$\begin{bmatrix} Y_0(z^2) \\ Y_1(z^2) \end{bmatrix} = \frac{1}{2}\mathbf{H}_m(z) \begin{bmatrix} X(z) \\ X(-z) \end{bmatrix}, \tag{4.41}$$

where $\mathbf{H}_m(z)$ is the modulation analysis matrix.

Theorem 4.5.1. *Let $\mathbf{H}_p(z)$ be the polyphase matrix for a one-scale wavelet analysis transform obtained by the lifting procedure. De£ne*

$$\mathbf{H}_m(z) = \mathbf{H}_p(z^2) \begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix}. \tag{4.42}$$

Then $\mathbf{H}_m(z)$ is the analysis modulation matrix for a two-channel £lter bank with perfect reconstruction. The analysis £lters for this £lter bank have z-transforms

$$H_0(z) = H_{00}(z^2) + zH_{01}(z^2) \quad \textit{and} \quad H_1(z) = H_{10}(z^2) + zH_{11}(z^2). \tag{4.43}$$

The synthesis £lters have z-transforms

$$G_0(z) = -(cz)^{-1}\left\{H_{10}(z^2) - zH_{11}(z^2)\right\} \quad \textit{and} \quad G_1(z) = (cz)^{-1}\left\{H_{00}(z^2) - zH_{01}(z^2)\right\}, \tag{4.44}$$

where $c = \det \mathbf{H}_p(z^2)$ is a nonzero real constant.

Proof. By de£nition

$$\mathbf{H}_m(z) = \begin{bmatrix} H_{00}(z^2) + zH_{01}(z^2) & H_{00}(z^2) - zH_{01}(z^2) \\ H_{10}(z^2) + zH_{11}(z^2) & H_{10}(z^2) - zH_{11}(z^2) \end{bmatrix}.$$

This is the analysis modulation matrix for the £lters de£ned by equations (4.43). Its determinant is $-2z \det \mathbf{H}_p(z)$. But a polyphase matrix obtained by the lifting procedure is the product of upper and lower triangular matrices of determinant 1 and a diagonal normalization matrix whose determinant is a nonzero constant c. Now apply Theorem 4.3.4. ■

Example 4.5.2 (Haar Transform)**.** The (unnormalized) polyphase matrix of the Haar transform is $\mathbf{H}_p(z) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Hence the modulation matrix is

$$\mathbf{H}_m(z) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1+z & 1-z \\ 1-z & 1+z \end{bmatrix}.$$

Thus $H_0(z) = \frac{1}{2}(1+z)$ and $H_1(z) = \frac{1}{2}(1-z)$. These are the z-transforms of the £lters $\mathbf{h}_0 = \frac{1}{2}(\delta_0 + \delta_{-1})$ and $\mathbf{h}_1 = \frac{1}{2}(\delta_0 - \delta_{-1})$ that take averages and differences of adjacent signal values. Since $\det \mathbf{H}_p(z) = -1/2$, equations (4.25) give

$$G_0(z) = 2z^{-1}H_1(-z) = 1 + z^{-1} \quad \text{and} \quad G_1(z) = -2z^{-1}H_0(-z) = 1 - z^{-1}.$$

These are the z-transforms of the £lters $\mathbf{g}_0 = \delta_0 + \delta_1$ and $\mathbf{g}_1 = \delta_0 - \delta_1$. ■

Example 4.5.3 (CDF(2,2) Transform)**.** The polyphase matrix of the CDF(2,2) transform is

$$\mathbf{H}_p(z) = \frac{\sqrt{2}}{8}\begin{bmatrix} (-z + 6 - z^{-1}) & (2 + 2z^{-1}) \\ -(2 + 2z) & 4 \end{bmatrix}.$$

(see Example 4.2.3). Hence the modulation matrix is

$$\begin{aligned} \mathbf{H}_m(z) &= \frac{\sqrt{2}}{8}\begin{bmatrix} (-z^2 + 6 - z^{-2}) & (2 + 2z^{-2}) \\ -(2 + 2z^2) & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix} \\ &= \frac{\sqrt{2}}{8}\begin{bmatrix} (-z^2 + 2z + 6 + 2z^{-1} - z^{-2}) & (-z^2 - 2z + 6 - 2z^{-1} - z^{-2}) \\ (-2z^2 + 4z - 2) & (-2z^2 - 4z - 2) \end{bmatrix}. \end{aligned}$$

Thus the analysis £lters have z-transforms $H_0(z) = \frac{\sqrt{2}}{8}(-z^2 + 2z + 6 + 2z^{-1} - z^{-2})$ and $H_1(z) = \frac{\sqrt{2}}{8}(-2z^2 + 4z - 2)$. These are the z-transforms of the £lters

$$\mathbf{h}_0 = \frac{\sqrt{2}}{8}(-\delta_{-2} + 2\delta_{-1} + 6\delta_0 + 2\delta_1 - \delta_2) \quad \text{and} \quad \mathbf{h}_1 = \frac{\sqrt{2}}{8}(-2\delta_{-2} + 4\delta_{-1} - 2\delta_0).$$

From the factored form of $\mathbf{H}_p(z)$ in Example 4.2.3 we see that $\det \mathbf{H}_p(z) = 1$. Hence equations (4.25) give

$$\begin{aligned} G_0(z) &= -z^{-1}H_1(-z) = \frac{\sqrt{2}}{8}(2z + 4 + 2z^{-1}) \quad \text{and} \\ G_1(z) &= z^{-1}H_0(-z) = \frac{\sqrt{2}}{8}(-z - 2 + 6z^{-1} - 2z^{-2} - z^{-3}). \end{aligned}$$

These are the z-transforms of the £lters

$$\mathbf{g}_0 = \frac{\sqrt{2}}{8}(2\delta_{-1} + 4\delta_0 + 2\delta_1) \quad \text{and} \quad \mathbf{g}_1 = \frac{\sqrt{2}}{8}(-\delta_{-1} - 2\delta_0 + 6\delta_1 - 2\delta_2 - \delta_3).$$

■

The converse to Theorem 4.5.1 is also true: Given a two-channel £lter bank with perfect reconstruction, we de£ne the analysis polyphase matrix by (4.42). The PR condition that $d(z) = \det \mathbf{H}_m(z)$ be a nonzero monomial allows us to reduce to the case of a matrix with determinant 1. Then $\mathbf{H}_p(z)$ can be factored into a product of a diagonal matrix $\operatorname{diag}[c\, c^{-1}]$ (for some constant $c \neq 0$) and upper-triangular or lower-triangular matrices with 1 in the diagonal positions and zero or a Laurent polynomial in the off-diagonal positions. Each of the factors corresponds to a lifting step (prediction, update, or normalization).

Factoring the polyphase matrix is carried out by the familiar elementary row operation steps (Gaussian elimination), but using arithmetic with Laurent polynomials instead of complex numbers, as we already did in Section 4.4 for the proof of Proposition 4.4.2 (Bezout's Theorem). There is a signi£cant complication: in carrying out row reduction of a matrix with real (or complex) entries, we can divide a row by any nonzero element in the matrix. However, when the matrix has entries that are Laurent polynomials, we can only divide by matrix elements that are pure monomials cz^m with $c \neq 0$ (by Lemma 4.3.3 these are the only invertible Laurent polynomials). In other words, if we de£ne the *degree* of a nonzero Laurent polynomial

$$f(z) = a_p z^p + \cdots + a_q z^q \quad \text{(where } p \le q,\ a_p \neq 0, \text{ and } a_q \neq 0)$$

to be $\deg(f) = q - p$, then $1/f(z)$ is a Laurent polynomial if and only if $\deg(f) = 0$. The way around this complication is to use the *Euclidean division algorithm* with remainder for Laurent polynomials: when $\deg(f) \ge \deg(g) > 0$, then we can write $f(z) = g(z)h(z) + r(z)$, where $h(z)$ and $r(z)$ are Laurent polynomials and $\deg(r) < \deg(g)$. Repeated application of this division algorithm allows us to reduce the degree of the remainder to zero, and then division by the remainder is possible.

We illustrate the process with the following example (see Chapter 12 of Ripples for more details and examples).

Example 4.5.4 (Factoring a Polyphase Matrix)**.** Consider the £lter bank with £lters

$$H_0(z) = \frac{1}{2}z^2 + z + \frac{1}{2} = \frac{1}{2}(z+1)^2 \quad \text{and} \quad H_1(z) = -\frac{3}{4} + \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2}.$$

These £lters satisfy the lowpass/highpass conditions $H_0(-1) = 0$ and $H_1(1) = 0$. Furthermore,

$$H_0(z)H_1(-z) = -\frac{3}{8}z^2 - z - \frac{3}{4} - \frac{1}{8}z^{-2}$$

has exactly one term of odd degree, so the PR condition is satis£ed. By (4.43) we calculate that polyphase matrix for this £lter bank is

$$\mathbf{H}_p(z) = \begin{bmatrix} (\frac{1}{2}z + \frac{1}{2}) & 1 \\ (-\frac{3}{4} + \frac{1}{4}z^{-1}) & \frac{1}{2}z^{-1} \end{bmatrix} = \begin{bmatrix} H_{00}(z) & H_{01}(z) \\ H_{10}(z) & H_{11}(z) \end{bmatrix}.$$

The modulation matrix has determinant $H_0(z)H_1(-z) - H_0(-z)H_1(z) = -2z$, and the polyphase matrix has determinant 1.

To factor this polyphase matrix, we use the row-reduction method (Gaussian elimination), as implemented by elementary matrices.

We begin by dividing $H_{10}(z)$ by $H_{00}(z)$ (with remainder). This is done by a unit lower triangular matrix multiplication:

$$\begin{bmatrix} 1 & 0 \\ \frac{3}{2}z^{-1} & 1 \end{bmatrix} \begin{bmatrix} (\frac{1}{2}z + \frac{1}{2}) & 1 \\ (-\frac{3}{4} + \frac{1}{4}z^{-1}) & \frac{1}{2}z^{-1} \end{bmatrix} = \begin{bmatrix} (\frac{1}{2}z + \frac{1}{2}) & 1 \\ z^{-1} & 2z^{-1} \end{bmatrix}.$$

In the new matrix the element in the lower left position is now invertible (as a Laurent polynomial), so we divide it into $H_{00}(z)$. This is done by an unit upper triangular matrix multiplication:

$$\begin{bmatrix} 1 & -\frac{1}{2}z^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} (\frac{1}{2}z + \frac{1}{2}) & 1 \\ z^{-1} & 2z^{-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & (-z+1) \\ z^{-1} & 2z^{-1} \end{bmatrix}.$$

Now both entries in the £rst column of the resulting matrix are invertible (as Laurent polynomials), so we can make the lower left entry zero by another unit lower triangular matrix multiplication, just as we would for a matrix of real numbers:

$$\begin{bmatrix} 1 & 0 \\ -2z^{-1} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & (-z+1) \\ z^{-1} & 2z^{-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & (-z+1) \\ 0 & 2 \end{bmatrix}.$$

Finally, we factor this upper triangular matrix as a diagonal matrix times a unit upper triangular matrix:

$$\begin{bmatrix} \frac{1}{2} & (-z+1) \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & (-2z+2) \\ 0 & 1 \end{bmatrix}.$$

Thus the polyphase matrix factors as

$$\mathbf{H}_p(z) = \begin{bmatrix} 1 & 0 \\ -\frac{3}{2}z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2}z^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2z^{-1} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & (-2z+2) \\ 0 & 1 \end{bmatrix} \tag{4.45}$$

Here we have moved the unit upper/lower triangular matrices to the right side of (4.45) using the relations

$$\begin{bmatrix} 1 & f(z) \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -f(z) \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ f(z) & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -f(z) & 1 \end{bmatrix}$$

(where $f(z)$ is any Laurent polynomial). Finally, we can move the diagonal matrix $D = \mathrm{diag}[\frac{1}{2},\ 2]$ in the factorization (4.45) to the left using the relations

$$D^{-1} \begin{bmatrix} 1 & f(z) \\ 0 & 1 \end{bmatrix} D = \begin{bmatrix} 1 & 4f(z) \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad D^{-1} \begin{bmatrix} 1 & 0 \\ f(z) & 1 \end{bmatrix} D = \begin{bmatrix} 1 & 0 \\ \frac{1}{4}f(z) & 1 \end{bmatrix}.$$

This gives the £nal factorization of the polyphase matrix:

$$\begin{aligned} \mathbf{H}_p(z) &= \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{3}{8}z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 2z^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2}z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & (-2z+2) \\ 0 & 1 \end{bmatrix} \\ &= DP_2U_2P_1U_1. \end{aligned}$$

This factorization means that the £lter bank can be implemented as follows. The signal $\mathbf{x}$ is split into $\mathbf{x}_{\text{even}}[n] = \mathbf{x}[2n]$ and $\mathbf{x}_{\text{odd}}[n] = \mathbf{x}[2n+1]$. Then the pair of signals $\mathbf{x}_{\text{even}}$ and $\mathbf{x}_{\text{odd}}$ are transformed into the trend $\mathbf{s}$ and detail $\mathbf{d}$ by the following lifting steps:

$$\begin{aligned}
&\text{(First Update)} \quad U_1: && \mathbf{s}^{(1)}[n] = \mathbf{x}_{\text{even}}[n] + 2\mathbf{x}_{\text{odd}}[n] - 2\mathbf{x}_{\text{odd}}[n+1] \\
&\text{(First Prediction)} \quad P_1: && \mathbf{d}^{(1)}[n] = \frac{1}{2}\mathbf{s}^{(1)}[n-1] + \mathbf{x}_{\text{odd}}[n] \\
&\text{(Second Update)} \quad U_2: && \mathbf{s}^{(2)}[n] = \mathbf{s}^{(1)}[n] + 2\mathbf{d}^{(1)}[n+2] \\
&\text{(Second Prediction)} \quad P_2: && \mathbf{d}^{(2)}[n] = -\frac{3}{8}\mathbf{s}^{(2)}[n-1] + \mathbf{d}^{(1)}[n] \\
&\text{(Normalization)} \quad D: && \mathbf{s}[n] = \frac{1}{2}\mathbf{s}^{(2)}[n], \qquad \mathbf{d}[n] = 2\mathbf{d}^{(2)}[n]
\end{aligned}$$

(recall that the shift operator $(S\mathbf{y})[n] = \mathbf{y}[n-1]$ corresponds to multiplication by z^{-1}). ■

4.6 Trend-Detail Decomposition for PR Filter Banks

Assume we have a two-channel FIR £lter bank with perfect reconstruction (PR). Let $\mathbf{h}_0$ (lowpass) and $\mathbf{h}_1$ (highpass) be the analysis £lters and let $\mathbf{g}_0$ and $\mathbf{g}_1$ be the corresponding synthesis £lters. (Recall that the PR property implies that the synthesis £lters are uniquely determined by the analysis £lters.)

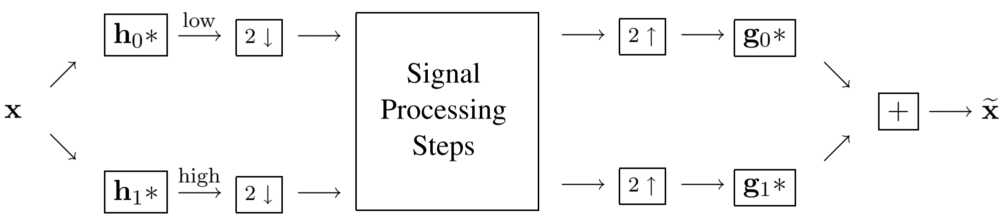
The analysis part of the £lter bank takes an input signal $\mathbf{x}$ and passes it through the £lters $\mathbf{h}_0$ and $\mathbf{h}_1$. Then the two £ltered signals are downsampled to give the output

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \end{bmatrix} = \begin{bmatrix} \boxed{2\downarrow}(\mathbf{h}_0 * \mathbf{x}) \\ \boxed{2\downarrow}(\mathbf{h}_1 * \mathbf{x}) \end{bmatrix}.$$

The synthesis part of the £lter bank takes the pair of signals $\mathbf{y}_0$ and $\mathbf{y}_1$, upsamples each of them, passes $\boxed{2\uparrow}\mathbf{y}_0$ and $\boxed{2\uparrow}\mathbf{y}_1$ through the synthesis £lters, and then adds the result to give the output

$$\widetilde{\mathbf{x}} = \mathbf{g}_0 * \left(\boxed{2\uparrow}\mathbf{y}_0\right) + \mathbf{g}_1 * \left(\boxed{2\uparrow}\mathbf{y}_1\right)$$

The *signal processing*, which is usually a nonlinear operation (such as setting small values to zero), occurs between the analysis and synthesis stages, and an input $\mathbf{x}$ produces an output $\tilde{\mathbf{x}}$:



When the signal processing is absent, then $\widetilde{\mathbf{x}} = \mathbf{x}$ and the PR property can be stated in the time domain as

$$\mathbf{x} = \underbrace{\mathbf{g}_0 * \left(\boxed{2\uparrow}\,\boxed{2\downarrow}(\mathbf{h}_0 * \mathbf{x})\right)}_{\text{trend}} + \underbrace{\mathbf{g}_1 * \left(\boxed{2\uparrow}\,\boxed{2\downarrow}(\mathbf{h}_1 * \mathbf{x})\right)}_{\text{detail}} = \mathbf{x}_s + \mathbf{x}_d \tag{4.46}$$

for all signals $\mathbf{x}$. Here the *trend* part $\mathbf{x}_s$ contains the low-frequency information in the signal, whereas the *detail* part $\mathbf{x}_d$ carries the high-frequency information. The analogous formula for periodic signals is (3.18).

Since convolution can be viewed as forming a moving average, formula (4.46) expresses $\mathbf{x}$ as a linear combination of shifts of the synthesis £lters $\mathbf{g}_0$ and $\mathbf{g}_1$. To make this precise, we need a formula for the linear transformations $\mathbf{x} \mapsto \mathbf{x}_s$ and $\mathbf{x} \mapsto \mathbf{x}_d$.

If $\mathbf{h}$ is a FIR £lter we write $\check{\mathbf{h}}$ for the time-reversed £lter:

$$\check{\mathbf{h}}[k] = \mathbf{h}[-k] \quad \text{for } k \in \mathbb{Z}.$$

Then the z-transform of $\check{\mathbf{h}}$ is $H(z^{-1})$. Note that if $z = \mathrm{e}^{\mathrm{i}\omega}$ then $z^{-1} = \mathrm{e}^{-\mathrm{i}\omega}$, so time reversal corresponds to frequency reversal. Recall that the inner product of signals $\mathbf{x}$ and $\mathbf{y}$ is

$$\langle \mathbf{x},\, \mathbf{y} \rangle = \sum_k \mathbf{x}[k]\, \mathbf{y}[k]$$

Lemma 4.6.1. *Suppose* $\mathbf{h}$ *and* $\mathbf{g}$ *are FIR £lters. Then for every signal* $\mathbf{x}$

$$\mathbf{g} * \left(\boxed{2\uparrow}\, \boxed{2\downarrow} (\mathbf{h} * \mathbf{x}) \right) = \sum_m \langle S^{2m} \check{\mathbf{h}},\, \mathbf{x} \rangle\, S^{2m} \mathbf{g} \tag{4.47}$$

Here S *is the right-shift operator and the coef£cients* $\langle S^{2m}\check{\mathbf{h}},\, \mathbf{x}\rangle$ *are zero for* $|m|$ *suf£ciently large.*

Proof. Recall that the linear transformation $\boxed{2\uparrow}\, \boxed{2\downarrow}$ (downsampling followed by upsampling) projects a signal $\mathbf{y}$ onto its even part:

$$\boxed{2\uparrow}\, \boxed{2\downarrow} \mathbf{y}[k] = \begin{cases} \mathbf{y}[k] & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Hence

$$\begin{aligned} \mathbf{g} * \left(\boxed{2\uparrow}\, \boxed{2\downarrow} (\mathbf{h} * \mathbf{x}) \right)[n] &= \sum_m \mathbf{g}(n - 2m)(\mathbf{h} * \mathbf{x})[2m] \\ &= \sum_m \Big\{ \sum_k \mathbf{h}[2m - k]\mathbf{x}[k] \Big\} (S^{2m}\mathbf{g})[n] \\ &= \sum_m \langle S^{2m}\check{\mathbf{h}},\, \mathbf{x} \rangle\, (S^{2m}\mathbf{g})[n]\,, \end{aligned}$$

which proves (4.47). Since $\mathbf{x}$ is a signal and $\mathbf{h}$ is a FIR £lter, there is an integer N so that $\mathbf{x}[k] = 0$ and $\mathbf{h}[k] = 0$ for $|k| > N$. Hence

$$\langle S^{2m}\check{\mathbf{h}},\, \mathbf{x} \rangle = \sum_{|k| \le N} \mathbf{h}[2m - k]\mathbf{x}[k].$$

Now take $|m| > N$ and $|k| \le N$. Then $|2m - k| \ge 2|m| - |k| > 2N - N = N$, and so $\mathbf{h}[2m - k] = 0$. Thus $\mathbf{h}[2m - k]\mathbf{x}[k] = 0$ and hence $\langle S^{2m}\check{\mathbf{h}},\, \mathbf{x} \rangle = 0$. ■

Applying Lemma 4.6.1 to equation (4.46) we obtain the generalization to nonperiodic signals of the trend/detail decomposition for periodic signals (equations (3.19) and (3.20)) that was proved in Section 3.4:

Theorem 4.6.2. *For the one-scale PR wavelet transform de£ned by the FIR analysis £lters* $\mathbf{h}_0$*,* $\mathbf{h}_1$ *and synthesis £lters* $\mathbf{g}_0$*,* $\mathbf{g}_1$*, the* trend *(low-frequency) component of a signal* $\mathbf{x}$ *is*

$$\mathbf{x}_s = \sum_m \langle S^{2m}\check{\mathbf{h}}_0,\, \mathbf{x}\rangle\, S^{2m}\mathbf{g}_0 \tag{4.48}$$

and the detail *(high-frequency) component of the signal* $\mathbf{x}$ *is*

$$\mathbf{x}_d = \sum_m \langle S^{2m}\check{\mathbf{h}}_1,\, \mathbf{x}\rangle\, S^{2m}\mathbf{g}_1. \tag{4.49}$$

Every signal $\mathbf{x}$ *has a decomposition* $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_d$*.*

Example 4.6.3 (CDF(2,2) Transform)**.** For the CDF(2,2) transform, the £lters are

$$\mathbf{h}_0 = \frac{\sqrt{2}}{8}\big(-\delta_{-2} + 2\delta_{-1} + 6\delta_0 + 2\delta_1 - \delta_2\big) \qquad \mathbf{h}_1 = \frac{\sqrt{2}}{8}\big(-2\delta_{-2} + 4\delta_{-1} - 2\delta_0\big)$$
$$\mathbf{g}_0 = \frac{\sqrt{2}}{8}\big(2\delta_{-1} + 4\delta_0 + 2\delta_1\big) \qquad \mathbf{g}_1 = \frac{\sqrt{2}}{8}\big(-\delta_{-1} - 2\delta_0 + 6\delta_1 - 2\delta_2 - \delta_3\big)$$

(see Example 4.5.3). The normalization factors $\sqrt{2}/8$ can be combined to give a single normalization of $1/32$ (binary shift) in the analysis £lters, for example. Then all arithmetic on a rational signal becomes rational.

As an example of the expansion in Theorem 4.6.2, take $\mathbf{x} = \delta_0$. Then

$$\langle S^{2m}\check{\mathbf{h}}_0, \delta_0\rangle = S^{2m}\check{\mathbf{h}}_0[0] = \check{\mathbf{h}}[-2m] = \mathbf{h}[2m].$$

Thus the trend component of δ_0 is

$$\begin{aligned}\sum_m \langle S^{2m}\check{\mathbf{h}}_0, \delta_0\rangle\, S^{2m}\mathbf{g}_0 &= \sum_m \mathbf{h}_0[2m]\, S^{2m}\mathbf{g}_0 = \frac{1}{4\sqrt{2}}\Big\{-S^{-2}\mathbf{g}_0 + 6\mathbf{g}_0 - S^2\mathbf{g}_0\Big\} \\ &= \frac{1}{16}\big(-\delta_{-3} - 2\delta_{-2} + 5\delta_{-1} + 12\delta_0 + 5\delta 1 - 2\delta_2 - \delta_3\big)\end{aligned}$$

Likewise, the detail component of δ_0 is

$$\begin{aligned}\sum_m \langle S^{2m}\check{\mathbf{h}}_1, \delta_0\rangle\, S^{2m}\mathbf{g}_1 &= \sum_m \mathbf{h}_1[2m]\, S^{2m}\mathbf{g}_1 = \frac{1}{4\sqrt{2}}\Big\{-2S^{-2}\mathbf{g}_1 - 2\mathbf{g}_1\Big\} \\ &= \frac{1}{16}\big(\delta_{-3} + 2\delta_{-2} - 5\delta_{-1} + 4\delta_0 - 5\delta 1 + 2\delta_2 + \delta_3\big)\end{aligned}$$

It is clear that these two components add to δ_0. The sum of the trend entries is 1, while the detail component oscillates and the sum of its entries is 0 (see Figure 4.1, where we have drawn the piecewise linear graphs that interpolate the values of the trend and detail at integer times). ■

We now describe the trend-detail decomposition in Theorem 4.6.2 from the point of view of linear algebra, generalizing the case of periodic signals treated in Section 3.4.

Theorem 4.6.4. *For a PR £lter bank, the set of all even-shifted £lters* $\{S^{2m}\mathbf{g}_0,\, S^{2n}\mathbf{g}_1 \,:\, m, n \in \mathbb{Z}\}$ *is linearly independent. Hence the decomposition of a signal* $\mathbf{x}$ *into a trend* $\mathbf{x}_s$ *in equation* (4.48) *and a detail* $\mathbf{x}_d$ *in equation* (4.49) *is unique.*

Figure 4.1: Trend and Detail for CDF(2,2) Decomposition of δ_0

Proof. Suppose some £nite linear combination of the even-shifted £lters adds up to zero:

$$\sum_m c_m S^{2m}\mathbf{g}_0 + \sum_n d_n S^{2n}\mathbf{g}_1 = 0.$$

Taking the z-transform of the left side of this equation, we obtain the relation

$$\varphi(z)G_0(z) + \psi(z)G_1(z) = 0,$$

where $\varphi(z) = \sum_m c_m z^{-2m}$ and $\psi(z) = \sum_n d_n z^{-2n}$. Since the Laurent polynomials φ and ψ only have even powers of z, they satisfy $\varphi(z) = \varphi(-z)$ and $\psi(z) = \psi(-z)$. Hence we get another linear relation

$$\varphi(z)G_0(-z) + \psi(z)G_1(-z) = 0.$$

These two relations can be written in matrix-vector form as

$$\mathbf{G}_m(z)\begin{bmatrix}\varphi(z)\\ \psi(z)\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}, \tag{4.50}$$

where $\mathbf{G}_m(z)$ is the synthesis modulation matrix for the £lter bank. From the PR property we know that $\mathbf{G}_m(z)$ is an invertible matrix. Hence the only solution to equation (4.50) is $\varphi(z) = 0$, $\psi(z) = 0$. This means that all the coef£cents $c_m = 0$ and $d_n = 0$, which proves linear independence. ■

Corollary 4.6.5. *The shifted synthesis £lters satisfy the* biorthogonality relations

$$\langle S^{2m}\check{\mathbf{h}}_0,\, S^{2n}\mathbf{g}_0\rangle = \delta_{m,n} \qquad \langle S^{2m}\check{\mathbf{h}}_1,\, S^{2n}\mathbf{g}_0\rangle = 0 \tag{4.51}$$

$$\langle S^{2m}\check{\mathbf{h}}_0,\, S^{2n}\mathbf{g}_1\rangle = 0 \qquad \langle S^{2m}\check{\mathbf{h}}_1,\, S^{2n}\mathbf{g}_1\rangle = \delta_{m,n} \tag{4.52}$$

for all integers m, n (where $\delta_{n,n} = 1$ and $\delta_{m,n} = 0$ for $m \neq n$).

Proof. By Theorem 4.6.2 the signal $\mathbf{x} = S^{2n}\mathbf{g}_0$ has a trend-detail decomposition

$$S^{2n}\mathbf{g}_0 = \sum_m \langle S^{2m}\check{\mathbf{h}}_0,\, S^{2n}\mathbf{g}_0\rangle\, S^{2m}\mathbf{g}_0 + \sum_m \langle S^{2m}\check{\mathbf{h}}_1,\, S^{2n}\mathbf{g}_0\rangle\, S^{2m}\mathbf{g}_1.$$

But by Theorem 4.6.4 we know that this decomposition is unique. Hence all the coef£cents on the right side of this equation must be zero, except for the coef£cient of $S^{2n}\mathbf{g}_0$, which must be one. This gives the biorthogonality relations (4.51). Now take the signal $\mathbf{x} = S^{2n}\mathbf{g}_1$ and apply the same argument to get the the biorthogonality relations (4.52). ■

Example 4.6.6 (CDF(2,2) Transform)**.** For the CDF(2,2) transform in Example 4.6.3,

$$\check{\mathbf{h}}_0 = \frac{\sqrt{2}}{8}\big(-\delta_{-2} + 2\delta_{-1} + 6\delta_0 + 2\delta_1 - \delta_2\big), \quad \check{\mathbf{h}}_1 = \frac{\sqrt{2}}{8}\big(-2\delta_0 + 4\delta_1 - 2\delta_2\big),$$
$$\mathbf{g}_0 = \frac{\sqrt{2}}{8}\big(2\delta_{-1} + 4\delta_0 + 2\delta_1\big), \qquad \mathbf{g}_1 = \frac{\sqrt{2}}{8}\big(-\delta_{-1} - 2\delta_0 + 6\delta_1 - 2\delta_2 - \delta_3\big)$$

The biorthogonality relations (4.51) and (4.51) can be checked directly (with some tedious calculations) using the orthogonality of the unit impulses. For example, $\langle S^{2m}\check{\mathbf{h}}_0,\, \mathbf{g}_0\rangle = 0$ if $|m| \geq 2$ since the supports of $S^{2m}\check{\mathbf{h}}_0$ and $\mathbf{g}_0$ are disjoint in this case. When there is overlapping of supports, then cancellation produces biorthogonality:

$$\begin{aligned}
\langle \check{\mathbf{h}}_0, \mathbf{g}_0\rangle &= \frac{1}{32}(2\cdot 2 + 4\cdot 6 + 2\cdot 2) = 1, \quad \langle S^{\pm 2}\check{\mathbf{h}}_0, \mathbf{g}_0\rangle = \frac{1}{32}((-1)\cdot 4 + 2\cdot 2) = 0,\\
\langle \check{\mathbf{h}}_1, \mathbf{g}_1\rangle &= \frac{1}{32}((-2)\cdot(-2) + 4\cdot 6 + (-2)\cdot(-2)) = 1,\\
\langle \check{\mathbf{h}}_0, \mathbf{g}_1\rangle &= \frac{1}{32}((-2)\cdot(-1) + 2\cdot 6 + 6\cdot(-2) + 2\cdot(-1)) = 0,\\
\langle \check{\mathbf{h}}_1, \mathbf{g}_0\rangle &= \frac{1}{32}((-2)\cdot 4 + 4\cdot 2) = 0.
\end{aligned}$$

4.7 Orthogonal Filter Banks

Assume we have a two-channel FIR £lter bank with perfect reconstruction. Let $\mathbf{h}_0$ and $\mathbf{h}_1$ be analysis £lters and let $\mathbf{g}_0$ and $\mathbf{g}_1$ be the synthesis £lters.

De£nition 4.7.1. *The £lter bank is* orthogonal *if* $\mathbf{g}_0 = \check{\mathbf{h}}_0$ *and* $\mathbf{g}_1 = \check{\mathbf{h}}_1$ *(the synthesis £lters are the* time-reversed *analysis £lters).*

The term *orthogonal* is justi£ed by Corollary 4.6.5, since the biorthogonality relations now become orthogonality relations:

$$\langle S^{2m}\mathbf{g}_0,\, S^{2n}\mathbf{g}_0\rangle = \delta_{m,n} \qquad \langle S^{2m}\mathbf{g}_1,\, S^{2n}\mathbf{g}_0\rangle = 0 \tag{4.53}$$
$$\langle S^{2m}\mathbf{g}_0,\, S^{2n}\mathbf{g}_1\rangle = 0 \qquad \langle S^{2m}\mathbf{g}_1,\, S^{2n}\mathbf{g}_1\rangle = \delta_{m,n} \tag{4.54}$$

In this case the one-scale wavelet decomposition is $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_d$ with trend $\mathbf{x}_s$ and detail $\mathbf{x}_d$ given by

$$\mathbf{x}_s = \sum_{m\in\mathbb{Z}} \langle S^{2m}\mathbf{g}_0,\, \mathbf{x}\rangle S^{2m}\mathbf{g}_0 \quad \text{and} \quad \mathbf{x}_d = \sum_{m\in\mathbb{Z}} \langle S^{2m}\mathbf{g}_1,\, \mathbf{x}\rangle S^{2m}\mathbf{g}_1.$$

Thus the set of vectors $\{S^{2m}\mathbf{g}_0,\, S^{2n}\mathbf{g}_1 \,:\, m, n \in \mathbb{Z}\}$ is an orthonormal basis for the vector space of signals. The trend component of the signal is orthogonal to the detail component, and the wavelet transform is energy-preserving:

$$||\mathbf{x}||^2 = ||\mathbf{x}_s||^2 + ||\mathbf{x}_d||^2 = \sum_{m\in\mathbb{Z}} |\langle S^{2m}\mathbf{g}_0,\, \mathbf{x}\rangle|^2 + \sum_{m\in\mathbb{Z}} |\langle S^{2m}\mathbf{g}_1,\, \mathbf{x}\rangle|^2$$

by Parseval's relation for an orthonormal basis.

Theorem 4.7.2. *A two-channel £lter bank is orthogonal if and only if the analysis modulation matrix* $\mathbf{H}_m(z)$ *satis£es*

$$\mathbf{H}_m(z)\mathbf{H}_m(z^{-1})^{\mathrm{T}} = 2I. \tag{4.55}$$

Proof. The de£nition (4.7.1) of orthogonality can be stated in terms of z transforms as

$$G_0(z) = H_0(z^{-1}) \quad \text{and} \quad G_1(z) = H_1(z^{-1}).$$

This means that the synthesis and analysis modulation matrices satisfy $\mathbf{G}_m(z) = \mathbf{H}_m(z^{-1})^{\mathrm{T}}$. By Theorem 4.3.2 the PR property is equivalent to $\mathbf{H}_m(z)\mathbf{G}_m(z) = 2I$. Hence the £lter bank is orthogonal if and only if (4.55) holds. ■

When $z = \mathrm{e}^{\mathrm{i}\omega}$ with ω real, then $z^{-1} = \mathrm{e}^{-\mathrm{i}\omega} = \overline{z}$ (complex conjugate). Since the Laurent polynomials $H_0(z)$ and $H_1(z)$ have real coef£cients, it follows that

$$H_0(\mathrm{e}^{-\mathrm{i}\omega}) = \overline{H_0(\mathrm{e}^{\mathrm{i}\omega})} \quad \text{and} \quad H_1(\mathrm{e}^{-\mathrm{i}\omega}) = \overline{H_1(\mathrm{e}^{\mathrm{i}\omega})}.$$

Thus $\mathbf{H}_m(z^{-1})^{\mathrm{T}} = \overline{\mathbf{H}_m(z)}^{\mathrm{T}}$ when $z = \mathrm{e}^{\mathrm{i}\omega}$. So condition (4.55) for an orthogonal £lter bank implies that the matrix $(1/\sqrt{2})\mathbf{H}_m(\mathrm{e}^{\mathrm{i}\omega})$ is *unitary* for all real ω. (The converse is also true and easy to prove.)

Example 4.7.3 (Haar)**.** For the Haar wavelet transform, the normalized analysis modulation matrix is

$$\mathbf{H}_m(z) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1+z & 1-z \\ 1-z & 1+z \end{bmatrix}.$$

(see Example 4.5.2). In this case

$$\mathbf{H}_m(z^{-1})^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1+z^{-1} & 1-z^{-1} \\ 1-z^{-1} & 1+z^{-1} \end{bmatrix},$$

and we calculate that $\mathbf{H}_m(z)\mathbf{H}_m(z^{-1})^{\mathrm{T}} = 2I$. Thus the normalized Haar transform is orthogonal. ■

We now can prove that a two-channel FIR orthogonal £lter bank is determined by the low pass analysis £lter (which must satisfy a single quadratic relation) and the time shift between the low pass and high pass channel:

Theorem 4.7.4. *Let* $\mathbf{h}_0$ *be a FIR £lter with* $H_0(-1) = 0$.

(1) *If* $\mathbf{h}_0$ *is the low pass analysis £lter for an orthogonal PR £lter bank then its z-transform satis£es*

$$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2. \tag{4.56}$$

(2) *Conversely, if condition* (4.56) *is satis£ed, let* $L = 2k+1$ *be an odd integer and let* $\mathbf{h}_1$ *be the FIR £lter with z-transform*

$$H_1(z) = z^L H_0(-z^{-1}). \tag{4.57}$$

De£ne $\mathbf{g}_0 = \check{\mathbf{h}}_0$ *and* $\mathbf{g}_1 = \check{\mathbf{h}}_1$. *Then the two-channel FIR £lter bank with analysis £lters* $\mathbf{h}_0$, $\mathbf{h}_1$ *and synthesis £lters* $\mathbf{g}_0$, $\mathbf{g}_1$ *has the perfect reconstruction property and is orthogonal.*

Proof. For any £lter bank with the PR property, the synthesis modulation matrix satis£es $\mathbf{G}_m(z) = 2\mathbf{H}_m(z)^{-1}$. If the £lter bank is orthogonal, then (4.55) implies that

$$\mathbf{G}_m(z) = \mathbf{H}_m(z^{-1})^{\mathrm{T}}.$$

Comparing matrix entries on each side of this equation, we see that $G_0(z) = H_0(z^{-1})$. Thus equation (4.30 in Theorem 4.4.1 becomes equation (4.56), proving part (1).

For the converse, we take $G_0(z) = H_0(z^{-1})$ and $d(z) = -2z^L$ in Theorem 4.4.1; this gives

$$H_1(z) = z^L G_0(-z) = z^L H_0(-z^{-1}) \quad \text{and} \quad G_1(z) = z^{-L} H_0(-z) = H_1(z^{-1})$$

The synthesis modulation matrix is thus $\mathbf{G}_m(z) = \mathbf{H}_m(z^{-1})^{\mathrm{T}}$, so the £lter bank is orthogonal, by Theorem 4.7.2. ■

Construction of Orthogonal Filter Banks

Recall that if $\mathbf{h}$ is a nonzero signal, then the *length* of $\mathbf{h}$ is $n - m + 1$, where m is the smallest integer such that $\mathbf{h}[m] \neq 0$, and $\mathbf{h}[n]$ is the largest integer such that $\mathbf{h}[n] \neq 0$.

Lemma 4.7.5. *Let* $\mathbf{h}_0$ *be a FIR lowpass £lter (*$H_0(-1) = 0$*). If* $H_0(z)$ *satis£es* (4.56) *then* $\mathbf{h}_0$ *has even length.*

Proof. Since $H_0(-1) = 0$, the length of $\mathbf{h}_0$ must be greater than one. Write $H_0(z) = a_m z^{-m} + \cdots + a_n z^{-n}$, where $a_m \neq 0$, $a_n \neq 0$, and $m < n$. Then

$$\begin{aligned} H_0(z)H_0(z^{-1}) &= a_m a_n z^{m-n} + \cdots + a_m a_n z^{n-m}, \\ (-1)^{-m+n} H_0(-z)H_0(-z^{-1}) &= a_m a_n z^{m-n} + \cdots + a_m a_n z^{n-m}, \end{aligned}$$

where the omitted terms on the right are linear combinations of z^p with $m - n < p < n - m$. If $m - n$ were even, then $(-1)^{m-n} = 1$ and adding these equations could not produce a constant. This would violate equation (4.56). Hence $m - n$ must be odd, and so $m - n + 1$, which is the length of $\mathbf{h}_0$, must be even. ■

Suppose $\mathbf{h}_0$ is a low pass £lter such that $H_0(z)$ satis£es (4.56) and is not a monomial. Then $\mathbf{h}_0$ has even length $2K \geq 2$. If $H_0(z)$ satis£es (4.56) then so does $z^q H_0(z)$ and $z^q H_0(z^{-1})$ for any integer q (multiplying by z^q gives a time shift). Both of these £lters vanish at $z = -1$ and hence are low pass. So we may assume that

$$H_0(z) = \mathbf{h}_0[0] + \mathbf{h}_0[1]\, z^{-1} + \cdots + \mathbf{h}_0[2K-1]\, z^{-2K+1} \tag{4.58}$$

with $\mathbf{h}_0[0] \neq 0$ and $\mathbf{h}_0[2K-1] \neq 0$. Set $P(z) = H_0(z)H_0(z^{-1})$ (the *power spectral response function* for the £lter). Since the £lter $\mathbf{h}_0$ is real, its z-transform satis£es $H_0(\bar{z}) = \overline{H_0(z)}$. Hence

$$P(\mathrm{e}^{\mathrm{i}\omega}) = H_0(\mathrm{e}^{\mathrm{i}\omega})H_0(\mathrm{e}^{-\mathrm{i}\omega}) = |H_0(\mathrm{e}^{\mathrm{i}\omega})|^2.$$

Thus $P(z)$ satis£es the following conditions:

symmetry: $P(z) = P(z^{-1})$ and $P(z) = cz^{-2K+1} + \cdots + cz^{2K-1}$ with $c \neq 0$.

positivity: $P(\mathrm{e}^{\mathrm{i}\omega}) \geq 0$ for $0 \leq \omega \leq \pi$.

half-band: $P(z) + P(-z) = 2$ (the only even term in $P(z)$ is the constant term 2).

low pass: $P(-1) = 0$.

We already know how to construct a Laurent polynomial that satis£es these four conditions: Let $B_K(y)$ be the Bezout polynomial (4.36) of degree $K - 1$. Since the binomial coef£cients in $B_K(y)$ are all positive, we have $B_K(y) > 0$ for all $y \geq 0$. Now set

$$y = \frac{1}{4}(-z + 2 - z^{-1})$$

(see equation (4.33)). When $z = \mathrm{e}^{\mathrm{i}\omega}$ then $y = \sin^2 \frac{\omega}{2} \geq 0$. Thus the Laurent polynomial

$$P(z) = 2(1-y)^K B_K(y) = 2(1+z)^{2K}(4z)^{-K} B_K\left(\frac{-z+2-z^{-1}}{4}\right) \tag{4.59}$$

(which we used to construct the CDF(p,q) £lters in Section 4.4) satis£es the symmetry condition, the positivity condition, and the low pass condition. The half-band condition (4.56) follows from the Bezout equation (4.39); note that we have multiplied $B_K(y)$ by 2.

The second step is to factor $P(z) = H_0(z)H_0(z^{-1})$. The symmetry and positivity conditions ensure that this can always be done, theoretically (the positivity condition implies that the roots z with $|z| = 1$ have even multiplicity, whereas by the symmetry condition every root z with $|z| \neq 1$ is paired with the root z^{-1} outside the unit circle). However obtaining the factorization involves £nding all complex roots of $P(z)$; this is a dif£cult numerical calculation when K is large since the roots can occur in clusters.

After we have found $H_0(z)$, then we obtain the high pass £lter by (4.57):

$$H_1(z) = z^L H_0(-z^{-1}),$$

where L can be any odd integer. Suppose $H_0(z)$ is given by (4.58). If we choose $L = -2K + 1$, then

$$H_1(z) = -\mathbf{h}_0[2K-1] + \mathbf{h}_0[2K-2]\,z^{-1} - \cdots - \mathbf{h}_0[1]z^{-2K+2} + \mathbf{h}_0[0]\,z^{-2K+1}. \tag{4.60}$$

Thus $\mathbf{h}_1$ is obtained from $\mathbf{h}_0$ by reversing the nonzero coef£cients and inserting alternating $\pm$. For example, if $\mathbf{h}_0 = a\delta_0 + b\delta_1 + c\delta_2 + d\delta_3$ has length four, then

$$\mathbf{h}_1 = -d\delta_0 + c\delta_1 - b\delta_2 + a\delta_3 \tag{4.61}$$

Example 4.7.6 (Haar)**.** Take $K = 1$ in (4.59). Since $B_1(y) = 1$, we get

$$P(z) = \frac{1}{2}(1+z)^2 z^{-1} = \frac{1}{2}(1+z)(1+z^{-1})$$

Thus $P(z) = H_0(z)H_0(z^{-1})$, where

$$H_0(z) = \frac{1}{\sqrt{2}}(1+z)$$

is the low pass £lter for the normalized Haar transform. ■

Example 4.7.7 (Daub4)**.** Take $K = 2$ in (4.59). Then

$$B_2\left(\frac{-z+2-z^{-1}}{4}\right) = 1 - \frac{1}{2}z^{-1}(z-1)^2 = \frac{1}{2}\left\{-z+4-z^{-1}\right\},$$

as in Example 4.4.3. Thus

$$P(z) = \frac{1}{16}(z+2+z^{-1})^2(-z+4-z^{-1}).$$

The second step is to factor $P(z)$. Let

$$H_0(z) = (z+2+z^{-1})(\alpha z^2+\beta z) = \alpha z^3 + (2\alpha+\beta)z^2 + (\alpha+2\beta)z + \beta.$$

Then H_0 satisfies the low-pass condition $H_0(-1) = 0$. If we choose the coefficients α and β so that

$$16(\alpha z^2+\beta z)(\beta z^{-1}+\alpha z^{-2}) = -z+4-z^{-1}$$

then we will have the factorization $H_0(z)H_0(z^{-1}) = P(z)$. Since

$$(\alpha z^2+\beta z)(\beta z^{-1}+\alpha z^{-2}) = \alpha\beta z + \alpha^2 + \beta^2 + \alpha\beta z^{-1},$$

we obtain the pair of quadratic equations $\alpha^2+\beta^2 = 1/4$ and $\alpha\beta = -1/16$ for α and β (the four points of intersection of a circle with the two branches of a hyperbola). It is easy to check that the pair

$$\alpha = \frac{1-\sqrt{3}}{4\sqrt{2}}, \quad \beta = \frac{1+\sqrt{3}}{4\sqrt{2}}$$

satisfies both equations.[1] Then $2\alpha+\beta = (3-\sqrt{3})/(4\sqrt{2})$ and $\alpha+2\beta = (3+\sqrt{3})/(4\sqrt{2})$. Hence

$$4\sqrt{2}H_0(z) = dz^3 + cz^2 + bz + a,$$

where $a = 1+\sqrt{3}$, $b = 3+\sqrt{3}$, $c = 3-\sqrt{3}$, and $d = 1-\sqrt{3}$. Now use equation (3.11) for the polyphase matrix of the Daub4 transform:

$$\mathbf{H}_p(z) = \frac{1}{4\sqrt{2}}\begin{bmatrix} (a+cz) & (b+dz) \\ -(b+dz^{-1}) & (a+cz^{-1}) \end{bmatrix}$$

(recall that the shift operator S becomes multiplication by z^{-1} on z-transforms). Hence by Theorem 4.5.1 the modulation matrix for the *Daub4* analysis transform is

$$\begin{aligned}\mathbf{H}_m(z) &= \frac{1}{4\sqrt{2}}\begin{bmatrix} (a+cz^2) & (b+dz^2) \\ -(b+dz^{-2}) & (a+cz^{-2}) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix} \\ &= \frac{1}{4\sqrt{2}}\begin{bmatrix} (dz^3+cz^2+bz+a) & * \\ (az-b+cz^{-1}-dz^{-2}) & * \end{bmatrix}\end{aligned}$$

[1] By symmetry the other points of intersection are $(-\alpha,-\beta)$, (β,α), and $(-\beta,-\alpha)$.

(where the second column is the modulation of the £rst column). Thus $H_0(z)$ is the z-transform of the Daub4 low pass £lter:

$$\mathbf{h}_0 = \frac{1}{4\sqrt{2}}(d\delta_{-3} + c\delta_{-2} + b\delta_{-1} + a\delta_0)$$

of length 4 (Ripples, equation 7.76). The high pass £lter is obtained by reversing the coef£cients, alternating the signs, and shifting, as in (4.61):

$$\mathbf{h}_1 = \frac{1}{4\sqrt{2}}(a\delta_{-1} - b\delta_0 + c\delta_1 - d\delta_2)$$

■

4.8 Exercises

1. Let $\mathbf{x}$ and $\mathbf{y}$ be the signals that are the following linear combinations of unit impulses:

$$\mathbf{x} = 3\delta_{-1} + 2\delta_0 - 5\delta_1 + 4\delta_2, \qquad \mathbf{y} = 7\delta_0 + 6\delta_1$$

   (a) Express $\mathbf{x}_{2\downarrow}$ and $(\mathbf{x}_{2\downarrow})_{2\uparrow}$ as linear combinations of unit impulses.

   (b) Calculate the z-transforms $\mathbf{X}(z)$ and $\mathbf{Y}(z)$.

   (c) Use the result of (b) to calculate the z-transform of the signal $\mathbf{w} = \mathbf{x} * \mathbf{y}$.

   (d) Let $\mathbf{y}_{\mathrm{per},4}$ be the periodic extension of $\mathbf{y}$ of period 4. Use your calculation in (b) to evaluate the discrete Fourier transform (DFT) $\widehat{\mathbf{y}}_{\mathrm{per},4}[k]$ for $k = 0, 1, 2, 3$.

2. Consider the lazy wavelet transform in Section 4.2.

   (a) Find the analysis and synthesis £lters, the modulation matrix, and the polyphase matrix for this transform.

   (b) Do the £lters for this transform satisfy the low pass and high pass conditions?

3. Let the FIR £lters $\mathbf{h}_0$ and $\mathbf{h}_1$ have z-transforms

$$H_0(z) = (1 + z)(1 + az) \quad \text{and} \quad H_1(z) = (1 - z)(1 + bz),$$

   where a and b are constants. (Notice that $a = 0$ and $b = 0$ give the £lters for the Haar transform).

   (a) Find all values of a and b so that $\det \mathbf{H}_m(z)$ is a nonzero monomial, where $\mathbf{H}_m(z)$ is the modulation matrix for these £lters.

   (b) With $H_1(z)$ determined as in part (a), £nd the FIR synthesis £lters $\mathbf{g}_0$ and $\mathbf{g}_1$ that go with $\mathbf{h}_0$ and $\mathbf{h}_1$ to give a two-channel PR £lter bank.

4. Let the FIR £lters $\mathbf{h}_0$ and $\mathbf{h}_1$ have z-transforms

$$H_0(z) = (1+z)^3 \quad \text{and} \quad H_1(z) = (1-z)(1+bz+cz^2).$$

(a) Find the values of the constants b and c so that $\det \mathbf{H}_m(z)$ is a nonzero monomial, where $\mathbf{H}_m(z)$ is the modulation matrix for these £lters.

(b) With $H_1(z)$ determined as in part (a), £nd the synthesis £lters $\mathbf{g}_0$ and $\mathbf{g}_1$ that go with $\mathbf{h}_0$ and $\mathbf{h}_1$ to give a two-channel PR £lter bank.

5. Factor $\begin{bmatrix} 1 & -3z \\ 2z^{-1} & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ f(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & g(z) \\ 0 & 1 \end{bmatrix}$ with $f(z)$, $g(z)$ Laurent polynomials.

6. Consider a two-channel £lter bank having FIR analysis £lters $\mathbf{h}_0$ (lowpass) and $\mathbf{h}_1$ (highpass) with z-transforms $H_0(z)$ and $H_1(z)$. Suppose the *polyphase matrix* $\mathbf{H}_p(z)$ for the analysis £lter bank is $\mathbf{H}_p(z) = \begin{bmatrix} 1+z & 2 \\ 1-3z & 2 \end{bmatrix}$.

(a) Find $H_0(z)$ and $H_1(z)$ and show that the lowpass/highpass conditions $H_0(-1) = 0$ and $H_1(1) = 0$ are satis£ed.

(b) Show that the condition for PR satis£ed and £nd the synthesis £lters $\mathbf{g}_0$ and $\mathbf{g}_1$.

7. Consider a two-channel £lter bank having FIR analysis £lters $\mathbf{h}_0$ (lowpass) and $\mathbf{h}_1$ (highpass) with z-transforms $H_0(z)$ and $H_1(z)$. Suppose the *polyphase matrix* $\mathbf{H}_p(z)$ for the analysis £lter bank is $\mathbf{H}_p(z) = \begin{bmatrix} 1 & 1-z \\ 1+z & 2-z^2 \end{bmatrix}$.

(a) Find $H_0(z)$ and $H_1(z)$ and determine whether the lowpass/highpass conditions $H_0(-1) = 0$ and $H_1(1) = 0$ are satis£ed.

(b) Is the condition for PR satis£ed by these £lters?

8. (a) Find the z-transforms of the low-pass £lters for the CDF(3,1) transform (see Section 4.4).

(b) Find the polyphase matrix for the CDF(3,1) transform.

(c) Show that the formulas on page 24 of Ripples give a factorization of the polyphase matrix in part (a) into lifting steps.

9. Let $\mathbf{x} = \delta_1$. Follow the method of Example 4.6.6 to £nd the decomposition of $\mathbf{x}$ into a trend $\mathbf{s}$ and detail $\mathbf{d}$ for the CDF(2,2) transform. (Notice that $\mathbf{s}$ and $\mathbf{d}$ are *not* obtained by shifting the trend/detail vectors for δ_0.)

10. Let $\mathbf{x} = \delta_0$. Follow the method of Example 4.6.6 to £nd the decomposition of $\mathbf{x}$ into a trend $\mathbf{s}$ and detail $\mathbf{d}$ for the CDF(3,1) transform (see Example 4.4.3 for the £lters).