A Lecture note on Wavelets

Kwan Pyo Ko

July 1, 2003

Chapter 1

Orthogonal Series

Orthogonal series play an important part in many areas of mathematics as well as in applications. They constitute an easy way of representing a function in terms of a series and may replace complicated operators on the function by simpler ones on the coefficients of the series. The most familiar orthogonal systems are the trigonometric and the various orthogonal polynomials.

1.1 General theory

We shall restrict ourselves to $L^2(a,b)$.

A nontrivial sequence $\{f_n\}_{n=0}^{\infty}$ of real (or complex) functions in $L^2(a,b)$ is said to be orthogonal if

$$< f_n, f_m >= \int_a^b f_n(x)\overline{f_m(x)}dx = 0, \quad n \neq m, \quad n, m = 0, 1, 2, \ldots$$

and orthonormal if in addition

$$< f_n, f_n >= 1, \quad n = 0, 1, 2, \ldots$$

For example, $f_n(x) = \sin(n+1)x$ is orthogonal on $(0, \pi)$. Another example is $f_n(x) = \chi_{[n,n+1)}(x)$ which is orthonormal on $[0, \infty)$.

The idea is to expand a given function $f(x) \in L^2(a,b)$ in an orthonormal series

$$f(x) = \sum_{n=0}^{\infty} c_n f_n \tag{1.1}$$

This is not always possible (e,g., take $f(x) = \chi_{[0.5,1)}(x)$ in the second example), but if it is, then the c_n's must have a special form. We shall use the usual notation for the L^2-norm, $\|f\| =< f, f >^{1/2}$.

Note: 1. Show that any orthogonal system $\{f_n\}$ is also linearly independent.

2. Let $\{f_n\}$ be a linearly independent, can you make an orthogonal system?

Proposition 1.1 *Let $\{c_n\}$ be a sequence such that the series in (1.1) converges in the sense of $L^2(a, b)$ to $f(x)$; then $c_n =< f, f_n >$.*

Proof We multiply both sides of (1.1) by $\overline{f}_m(x)$ and then integrate. Because of the orthogonality all terms in the series drop out except c_m. There is no problem with interchanging the integral and the summation because of the continuity of the inner product with respect to the norm. $\diamond$

Convergence in the sense of $L^2(a, b)$ is also known as mean square convergence, and the error

$$e_N = \|f - \sum_{n=0}^{N} c_n f_n\|^2$$

is called the mean square error. The coefficients appearing in Proposition 1.1 are called the Fourier coefficients of f with respect to $\{f_n\}$ and have another property that makes them useful.

Proposition 1.2 *Let $\{c_n\}$ be the Fourier coefficient of $f \in L^2(a, b)$ and $\{a_n\}$ any other sequence; then we have*

$$\|f - \sum_{n=0}^{N} c_n f_n\|^2 \leq \|f - \sum_{n=0}^{N} a_n f_n\|^2,$$

i.e., the mean square error is minimized for the series with Fourier coefficients.

Proof Exercise

The Fourier coefficients give the orthogonal projection of f onto the subspace V_N spanned by $(f_0, f_1, f_2, \ldots, f_N)$. We see that

$$< \sum_{n=0}^{N} c_n f_n, f - \sum_{n=0}^{N} c_n f_n >= 0$$

The best approximation to f in V_N is given by this sum, and the error is orthogonal to V_N.

Similar calculations lead to Bessel's inequality

$$\sum_{n=0}^{\infty} |c_n|^2 \leq \|f\|^2 \tag{1.2}$$

since

$$0 \leq \|f - \sum_{n=0}^{N} c_n f_n\|^2 = \|f\|^2 - \sum_{n=0}^{N} |c_n|^2,$$

and therefore $\{\sum_{n=0}^{N} |c_n|^2\}$ is a monotone sequence bounded by $\|f\|^2$. Thus the series of (1.2) converges and has the same bound. A simple consequence of Bessel's inequality is that $\{c_n\} \in l^2$ and $c_n \to 0$ as $n \to \infty$.

To round out our theory, we should like to have the series with Fourier coefficients $\sum c_n f_n$ converges to f. By Bessel's inequality the partial sums are a Cauchy sequence in $L^2(\mathbf{R})$, which because of the completeness of this space must converge in L^2 sense but not to f necessarily. To ensure this we need to add another condition, the completeness of $L^2(a, b)$ if no nontrivial $f \in L^2(a, b)$ is orthogonal to all the f_n's, i.e., if $< f, f_n >= 0, n = 0, 1, 2, \ldots$, for $f \in L^2(a, b)$ then $f = 0$, a.e.

Theorem 1.1 *Let $\{f_n\}$ be an orthonormal system in $L^2(a, b)$; let $f \in L^2(a, b)$ with Fourier coefficients $\{c_n\}$; then*

$$\|f - \sum_{n=0}^{N} c_n f_n\| \to 0 \quad \text{as } N \to \infty$$

if and only if $\{f_n\}$ is complete.

The conclusion of the theorem can restated as Parseval's equality

$$\|f\|^2 = \sum_{n=0}^{\infty} |c_n|^2 \tag{1.3}$$

1.2 Trigonometric system

The trigonometric system is a complete orthogonal system in $L^2(-\pi,\pi)$ given by

$$\begin{aligned} f_0(x) &= 1/2, f_1(x) = \sin x, f_2(x) = \cos x, \\ \cdots & f_{2n-1}(x) = \sin nx, f_{2n}(x) = \cos nx, \cdots \end{aligned}$$

It is usually not normalized, since $\|f_n\|^2 = \pi, n \neq 0$. The series is usually written in the form

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx. \tag{1.4}$$

If (1.4) is the Fourier series of a function $f \in L^2(-\pi,\pi)$, the coefficients are given by

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx \; dx, \quad n = 1, 2, \cdots \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx \; dx, \quad n = 1, 2, \cdots \end{aligned}$$

First showing that the Fourier series converges uniformly for certain functions. **It should be remarked that this is not true for all continuous functions**.

If the series (1.4) is to converges uniformly, then the limit function must be continuous and periodic of period 2π, which we assume f to be. We shall need an expression for the partial sums of the series,

$$\begin{aligned} S_n(x) &= \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{k=1}^{n} \cos kt \cos kx + \sin kt \sin kx\right\} dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{k=1}^{n} \cos k(x-t)\right\} dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\sin(n+1/2)(x-t)}{2\sin(x-t)/2} dt \\ &= \int_{-\pi}^{\pi} f(x-u)\frac{\sin(n+1/2)u}{2\pi \sin u/2} du \end{aligned}$$

The expression

$$D_n(x) = \frac{1}{\pi}\left[\frac{1}{2} + \sum_{k=1}^{n} \cos ku\right] = \frac{\sin(n+1/2)u}{2\pi \sin u/2} \tag{1.5}$$

is called the <u>Dirichlet kernel</u>, and it plays a central role in the study of point-wise convergence of Fourier series.

Note: Prove that

$$\frac{1}{2} + \sum_{k=1}^{n} \cos k(x-t) = \frac{\sin(n+1/2)(x-t)}{2\sin(x-t)/2}$$

Proposition 1.3 *Let f be a 2π periodic function in $C^2(\mathbf{R})$; then*

- $\sup|S_n(x) - f(x)| \to 0$
- $\|S_n - f\| \to 0$

Let $f \in L^2(0, 2\pi)$ and denote by $S_n f$ the nth partial sum of the Fourier series. Then the Nth Cesaro means of $\{S_n f\}$ is given by

$$\sigma_N f := \frac{S_0 f + \cdots + S_N f}{N+1}$$

Since $S_n f$ is the convolution of f with the Dirichlet kernel D_n, it follows that $\sigma_N f$ is the convolution of f with so-called <u>Fejer kernel</u>, defined by

$$K_N(x) := \frac{D_0(x) + \cdots + D_N(x)}{N+1} = \frac{1}{N+1}\frac{\sin^2(\frac{N+1}{2}x)}{2\sin^2(x/2)}$$

namely:

$$(\sigma_N f)(x) = \frac{1}{\pi}\int_0^{2\pi} f(x-t)K_N(t)dt.$$

Proposition 1.4 *Let $f \in L^2(0, 2\pi)$. Then*

$$\|\sigma_N f - f\|_2 \to 0$$

An alternative form for the trigonometric series (1.4) is the exponential form

$$S(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{1.6}$$

If (1.6) is a Fourier series, the coefficients are

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx. \tag{1.7}$$

Of course expression (1.6) is reducible to (1.4) by using $e^{\pm inx} = \cos nx \pm i\sin nx$.

The Fourier transform of the function $f \in L^1(-\infty,\infty)$ is the expression

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt, \quad \omega \in \mathbf{R}.$$

Note : Prove that

- $\hat{f} \in L^\infty(\mathbf{R})$ with $\|\hat{f}\|_\infty \leq \|f\|_1$
- $\hat{f}$ is uniformly continuous
- $\hat{f}'(\omega) = i\omega\hat{f}(\omega)$
- $\hat{f}(\omega) \to 0$

Note : Although $\hat{f}(\omega) \to 0$ for every $f \in L^1(R)$, it does not means that $\hat{f}$ is necessarily in $L^1(R)$. (example, $f(x) = e^{-x}u_0(x)$, where $u_0(x)$: Heaviside unit step function).

- $\int_{-\infty}^{\infty} e^{-i\omega t}e^{-at^2}dt = \sqrt{\pi/a}e^{-\omega^2/4a}$

The Fourier transform of the Gaussian function e^{-x^2} is $\sqrt{\pi}e^{-\omega^2/4}$

The image of the transform in this case is a continuous function on $\mathbf{R}$, which, if it is also in $L^1(\mathbf{R})$, leads to the inverse

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t}d\omega,$$

Versions of Parseval's equality also exist for Fourier transforms. They are for $f, g \in L^2$

$$\|f\|^2 = \frac{1}{2\pi}\|\hat{f}\|^2, \quad < f, g >= \frac{1}{2\pi} < \hat{f}, \hat{g} > .$$

Let f and g be functions in $L^1(R)$. Then the convolution of f and g is also an $L^1(R)$ function h defined by

$$h(x) = (f * g)(x) := \int_{-\infty}^{\infty} f(x-y)g(y)dy$$

It is clear that $h \in L^1(R)$, and in fact,

$$\|h\|_1 \leq \|f\|_1 \|g\|_1$$

Note :

- $f * g = g * f$
- $(f \hat{*} g)(\omega) = \hat{f}(\omega)\hat{g}(\omega)$

One can also go the other way and approximate (1.7) by a discrete sum. This gives us the discrete Fourier transform,

$$\gamma_k = \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{2\pi j}{N}\right) e^{-ijk2\pi/N}, \quad k = 0, \cdots, N-1$$

with the inverse given by

$$f\left(\frac{2\pi j}{N}\right) = \sum_{k=0}^{N-1} \gamma_k e^{ikj2\pi/M}, \quad j = 0, \cdots, N-1$$

This is the form that leads to the fast Fourier transform.

1.3 Poisson's summation formula

To periodize a function in $L^1(R)$, the simplest way to consider

$$\Phi_f(x) := \sum_{k=-\infty}^{\infty} f(x + 2\pi k)$$

Lemma 1.1 *Let $f \in L^1(R)$. Then*

$$\|\Phi_f\|_{L^1(0,2\pi)} \leq \frac{1}{2\pi} \|f\|_1$$

Proof

$$\begin{aligned}\int_0^{2\pi} |\Phi_f(x)|\ dx &\le \sum_{k=-\infty}^{\infty} \int_0^{2\pi} |f(x+2\pi k)|\ dx \\ &= \sum_{k=-\infty}^{\infty} \int_{2\pi k}^{2\pi(k+1)} |f(x)|\ dx = \int_{-\infty}^{\infty} |f(x)|\ dx < \infty\end{aligned}$$

In view of this lemma, we may consider the Fourier series of Φ_f, namely:

$$\Phi_f(x) = \sum_{k=-\infty}^{\infty} c_k(\Phi_f) e^{ikx},$$

where

$$\begin{aligned}c_k(\Phi_f) &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} \Phi_f(x)\ dx \\ &= \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \int_0^{2\pi} e^{-ikx} f(x+2\pi j)\ dx \\ &= \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \int_{2\pi j}^{2\pi(j+1)} e^{-ikx} f(x)\ dx = \frac{1}{2\pi} \hat{f}(k).\end{aligned}$$

Then the following Poisson Summation Formula holds:

$$\sum_{k=-\infty}^{\infty} f(x+2\pi k) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikx}, \quad x \in R$$

In particular,

$$\sum_{k=-\infty}^{\infty} f(2\pi k) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \hat{f}(k).$$

1.4 Haar system

1.4.1 Haar scaling function

Definition 1.1 *The Haar scaling function is defined as*

$$\phi(x) = \begin{cases} 1 & 0 \le x < 1 \\ 0 & o.w \end{cases}$$

Let V_0 be the space of all functions of the form

$$\sum_{k \in Z} a_k \phi(x-k)$$

V_0 consists of all piece-wise constant functions whose discontinuities are in the set of integers. We need blocks that are thinner to analyze signals of high frequency. The building block whose width is half that of the graph of ϕ is given by the graph of $\phi(2x)$. The function $\phi(2x-k) = \phi(2(x-k/2))$ is the same as the graph of the function of $\phi(2x)$ shifted to the right by $k/2$ units. Let V_1 be the space of functions of the form

$$\sum_{k \in Z} a_k \phi(2x-k)$$

We make the following more general definition.

Definition 1.2 *Suppose $j \in Z^+$. The space of step functions at level j, denoted by V_j, is defined to be the space spanned by the set*

$$\{\cdots, \phi(2^j x+1), \phi(2^j x), \phi(2^j x-1), \cdots\}$$

V_j is the space of piece-wise constant functions of finite support whose discontinuities are contained in the set

$$\{\cdots, -1/2^j, 0, 1/2^j, 2/2^j, 3/2^j, \cdots\}$$

We get the following containment.

$$V_0 \subset V_1 \subset \cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \cdots$$

V_j contains all relevant information up to a resolution scale of order 2^{-j}. As j gets larger, the resolution gets finer. The fact that $V_j \subset V_{j+1}$ means that no information is lost as the resolution gets finer.

The following theorem is an easy consequence of the definitions.

Theorem 1.2

- *A function $f(x) \in V_0$ if and only if $f(2^j x) \in V_j$.*
- *A function $f(x) \in V_j$ if and only if $f(2^{-j} x) \in V_0$*

It is desirable to have an efficient algorithm to decompose a signal into its V_j-components. One way to perform this decomposition efficiently is to construct an orthonormal basis for V_j. Lets start with V_0.

$$\|\phi(x-k)\|_{L^2}^2 = \int_{-\infty}^{\infty} \phi(x-k)^2 \ dx = \int_{k}^{k+1} 1 \ dx = 1$$

If j is different from k, then $\phi(x-k)$ and $\phi(x-j)$ have disjoint supports. Therefore,

$$< \phi(x-j), \phi(x-k) >= \int_{-\infty}^{\infty} \phi(x-j)\phi(x-k) \ \ dx = 0, \quad j \neq k$$

so that the set $\{\phi(x-k), k \in Z\}$ is an orhonormal basis for V_0. The same argument establishes the following more general result.

Theorem 1.3 *The set of functions* $\{2^{j/2}\phi(2^j x - k) : k \in Z\}$ *is orthonormal basis of* V_j.

1.4.2 Haar wavelet

In order to solve our noise-filtering problem, we need to have a way of isolating the "spikes" that belong to V_j but not members of V_{j-1}. The idea is to decompose V_j as on orthogonal sum of V_{j-1} and its complement. Let's start with $j = 1$ and identify the orthogonal complement of V_0 in V_1. Since V_0 is generated by ϕ and its translation, it is reasonable to expect that the orthogonal complement of V_0 is generated by the translates of some function ψ. Two key facts are needed to construct ψ:

- ψ is a member of V_1 and so ψ can be expressed as $\psi(x) = \sum_l a_l \phi(2x - l)$
- ψ is orthogonal to V_0

Definition 1.3 *The Haar wavelet is the function*

$$\psi(x) = \phi(2x) - \phi(2x-1)$$

You can show that any function

$$f_1 = \sum_k a_k \phi(2x - k) \in V_1$$

is orthogonal to V_0 if and only if

$$a_1 = -a_0, a_3 = -a_2, \cdots$$

In this case

$$f_1 = \sum_k a_{2k}(\phi(2x - 2k) - \phi(2x - 2k - 1)) = \sum_k a_{2k}\psi(x - k).$$

In other words, a function in V_1 is orthogonal to V_0 if and only if it is of the form $\sum_k a_k\psi(x - k)$.

Let W_0 be the space of all functions of the form

$$\sum_{k\in Z} a_k\psi(x - k), \quad x \in R$$

What we have just shown is that W_0 is the orthogonal complement of V_0 in V_1,

$$V_1 = V_0 \bigoplus W_0$$

Theorem 1.4 *Let W_j be the space of the functions of the form*

$$\sum_{k\in Z} a_k\psi(2^j x - k)$$

where we assume that only a finite number of a_k are nonzero. W_j is the orthogonal complement of V_j in V_{j+1} and

$$V_{j+1} = V_j \bigoplus W_j$$

By successively decomposing V_j, V_{j-1} and so on, we have

$$\begin{aligned} V_j &= W_{j-1} \bigoplus V_{j-1} \\ &= W_{j-1} \bigoplus W_{j-2} \bigoplus V_{j-2} \\ &\cdots \\ &= W_{j-1} \bigoplus W_{j-2} \bigoplus \cdots \bigoplus W_0 \bigoplus V_0 \end{aligned}$$

So each f in V_j can be decomposed uniquely as a sum

$$f = w_{j-1} + w_{j-2} + \cdots + w_0 + v_0$$

where each w_l belongs to W_l

Theorem 1.5 *The space $L^2(R)$ can be decomposed as an infinite orthogonal direct sum*

$$L^2(R) = V_0 \bigoplus W_0 \bigoplus W_1 \bigoplus \cdots$$

In particular, each $f \in L^2(R)$ can be written uniquely as

$$f = f_0 + \sum_{j=0}^{\infty} w_j,$$

where f_0 belongs to V_0 and w_j belong to W_j

1.4.3 Haar decomposition

Lemma 1.2 *The following relations hold for all $x \in R$:*

$$\begin{aligned}
\phi(2^j x) &= (\phi(2^{j-1}x) + \psi(2^{j-1}x))/2 \\
\phi(2^j x - 1) &= (\phi(2^{j-1}x) - \psi(2^{j-1}x))/2
\end{aligned}$$

First divide the sum $f_j(x) = \sum_k a_k \phi(2^j x - k)$ into even and odd terms:

$$f_j(x) = \sum_k a_{2k}\phi(2^j x - 2k) + \sum_k a_{2k+1}\phi(2^j x - 2k - 1).$$

By Lemma 1.2, we get

$$\begin{aligned}
\phi(2^j x - 2k) &= (\phi(2^{j-1}x - k) + \psi(2^{j-1}x - k))/2 \\
\phi(2^j x - 2k - 1) &= (\phi(2^{j-1}x - k) - \psi(2^{j-1}x - k))/2
\end{aligned}$$

Substituting these expression yields

$$\begin{aligned}
f_j(x) &= \sum_k a_{2k}(\phi(2^{j-1}x - k) + \psi(2^{j-1}x - k))/2 \\
&+ \sum_k a_{2k+1}(\phi(2^{j-1}x - k) - \psi(2^{j-1}x - k))/2 \\
&= \sum_k (\frac{a_{2k} - a_{2k+1}}{2})\psi(2^{j-1}x - k) + (\frac{a_{2k} + a_{2k+1}}{2})\phi(2^{j-1}x - k) \\
&= w_{j-1} + f_{j-1}
\end{aligned}$$

We summarize the preceding decomposition algorithm in the following theorem

Theorem 1.6 *Suppose*

$$f_j(x) = \sum_k a_k^j \phi(2^j x - k) \in V_j$$

Then f_j can be decomposed as

$$f_j = w_{j-1} + f_{j-1}$$

where

$$\begin{aligned} w_{j-1} &= \sum_k b_k^{j-1} \psi(2^{j-1}x - k) \in W_{j-1} \\ f_{j-1} &= \sum_k a_k^{j-1} \phi(2^{j-1}x - k) \in V_{j-1} \end{aligned}$$

with

$$b_k^{j-1} = \frac{a_{2k}^j - a_{2k+1}^j}{2}, \quad a_k^{j-1} = \frac{a_{2k}^j + a_{2k+1}^j}{2}$$

1.4.4 Haar reconstruction

Having decomposed f into its V_0 and $W_{j'}$-components for $0 \leq j' < j$, then what? The answer depends on goal.

We start with a signal of the form

$$f(x) = f_0(x) + w_0(x) + \cdots + w_{j-1}(x), \quad w_l \in W_l,$$

where

$$f_0(x) = \sum_k a_k^0 \phi(x - k) \in V_0, \quad w_l = \sum_k b_k^l \psi(2^l x - k) \in W_l$$

Our goal is to rewrite f as $f(x) = \sum_l a_l^j \phi(2^j x - l)$ and find algorithm for the computation of the constants a_l^j. We use the equations

$$\begin{aligned} \phi(2^{j-1}x) &= \phi(2^j x) + \phi(2^j x - 1) \\ \psi(2^{j-1}x) &= \phi(2^j x) - \phi(2^j x - 1) \end{aligned}$$

Using equations with x replaced by $x - k$, we have

$$\begin{aligned} f_0(x) &= \sum_k a_k^0 \phi(x - k) \\ &= \sum_k a_k^0 \phi(2x - 2k) + a_k^0 \phi(2x - 2k - 1) \end{aligned}$$

So

$$f_0(x) = \sum_k \hat{a}_l^1 \phi(2x - l)$$

where

$$\hat{a}_l^1 = \begin{cases} a_k^0 & l = 2k \\ a_k^0 & l = 2k + 1 \end{cases}$$

Similarly, $w_0 = \sum_k b_k^0 \psi(x - k)$ can be written as

$$w_0(x) = \sum_k \hat{b}_l^1 \phi(2x - l),$$

where

$$\hat{b}_l^1 = \begin{cases} b_k^0 & l = 2k \\ -b_k^0 & l = 2k + 1 \end{cases}$$

Combining yields

$$f_0(x) + w_0(x) = \sum_l a_l^1 \phi(2x - l)$$

where

$$a_l^1 = \hat{a}_l^1 + \hat{b}_l^1 = \begin{cases} a_k^0 + b_k^0 & l = 2k \\ a_k^0 - b_k^0 & l = 2k + 1 \end{cases}$$

Next, $w_1 = \sum_k b_k^1 \psi(2x - k)$ is added to this sum in the same manner:

$$f_0(x) + w_0(x) + w_1(x) = \sum_l a_l^2 \phi(2^2 x - l)$$

where

$$a_l^2 = \begin{cases} a_k^1 + b_k^1 & l = 2k \\ a_k^1 - b_k^1 & l = 2k + 1 \end{cases}$$

Theorem 1.7 *Suppose*

$$f = f_0 + w_0 + w_1 + w_2 + \cdots + w_{j-1}$$

with

$$f_0(x) = \sum_k a_k^0 \phi(x - k) \in V_0, \quad w_{j'} = \sum_k b_k^{j'} \psi(2^{j'} x - k) \in W_{j'}$$

for $0 \leq j' < j$. *Then*

$$f(x) = \sum_l a_l^j \phi(2^j x - l) \in V_j$$

where

$$a_l^{j'} = \begin{cases} a_k^{j'-1} + b_k^{j'-1} & l = 2k \\ a_k^{j'-1} - b_k^{j'-1} & l = 2k+1 \end{cases}$$

The $\phi(x)$ is usually called the scaling function in wavelet terminology while $\psi(x)$ is the mother wavelet.

1.5 The Shannon system

It is the Fourier transform of the scaling function, taken to be

$$\hat{\phi}(\omega) = \begin{cases} 1 & -\pi \leq \omega < \pi \\ 0 & o.w \end{cases}$$

Its inverse Fourier transform is

$$\phi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega t} d\omega = \frac{\sin \pi t}{\pi t}$$

The orthogonality of $\phi(t)$ and $\phi(t-n)$ is based on the Parseval's equality, and the fact that $(\phi(t \hat{-} \alpha))(\omega) = \hat{\phi}(\omega) e^{-i\alpha\omega}$:

$$\begin{aligned} \int_{-\infty}^{\infty} \phi(t) \overline{\phi(t-n)} dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega) \overline{\hat{\phi}(\omega)} e^{i\omega n} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega n} d\omega = \frac{\sin \pi n}{\pi n} = 0, n \neq 0. \end{aligned}$$

Let $f(t)$ be a function that is square integrable and whose Fourier transform $\hat{f}(\omega)$ vanishes for $|\omega| > \pi$. It has a Fourier series given by

$$\hat{f}(\omega) = \sum_n c_n e^{i\omega n}, \quad |\omega| \leq \pi \tag{1.8}$$

where $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{f}(\omega)e^{-i\omega n}d\omega$. By the Fourier integral theorem this is just $f(-n)$.

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{f}(\omega)e^{i\omega t}d\omega \\ &= \sum_n f(-n)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega n}e^{i\omega t}d\omega \\ &= \sum_n f(-n)\frac{\sin \pi(t+n)}{\pi(t+n)} \end{aligned}$$

We denote by V_0 the set of all such functions. This is a linear space and is closed as well since limits of the sequences of functions in V_0 are also V_0. The formula is referred to as the Shannon sampling theorem. It enables one to recover a band-limited function in V_0 from its values on the integers. This is used by engineers to convert a digital to an analog signal.

Chapter 2

Scaling Functions and Wavelets

Any wavelet gives rise to some decomposition of the Hilbert space $L^2(\mathbf{R})$ into a direct sum of closed subspaces $W_j, j \in \mathbf{Z}$; in the sense that each subspace W_j is the closure in $L^2(\mathbf{R})$ of the linear span of the collection of functions

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \quad k \in \mathbf{Z}.$$

Hence, the corresponding subspaces

$$V_j := \cdots \dot{+} W_{j-2} \dot{+} W_{j-1}, \quad j \in \mathbf{Z}$$

form a nested sequence of subspaces of $L^2(\mathbf{R})$, whose union is dense in $L^2(\mathbf{R})$ and whose intersection is the null space $\{0\}$.

This observation motivates the following introduction of a very useful technique for constructing the wavelet ψ and its corresponding dual $\tilde{\psi}$, namely: the investigation of the existence, and a study the structure, of some scaling function ϕ that generates the spaces V_j, in the same manner as ψ generates the spaces W_j. In particular, the collection of functions

$$\phi(x - k), \quad k \in \mathbf{Z}$$

is to form a Riesz basis of V_0; and hence, ϕ generates a multiresolution analysis (MRA) $\{V_j\}$ of $L^2(\mathbf{R})$. Since $\phi \in V_0 \subset V_1$, there exists a unique sequence $\{p_n\} \in l^2$ that relates $\phi(x)$ with the functions $\phi(2x - k)$; and the structure of ϕ is governed by that of this *two-scale sequence* $\{p_n\}$. For instance, a finite two-scale sequence characterizes a scaling function ϕ with compact support.

We will see that there is quite a lot of freedom in choosing the corresponding wavelet ψ and its dual $\tilde{\psi}$, and another objective of this chapter is to investigate the structure of the complementary space W_j. Among those specification of special interest: decomposition of the space $L^2(\mathbf{R})$ as an orthogonal sum of the spaces W_j; an orthonormal basis of $L^2(\mathbf{R})$ generated by ψ; finite reconstruction and decomposition sequences as a result of compactly supported ψ and $\tilde{\psi}$; and symmetry or anti-symmetry of ψ and $\tilde{\psi}$. We will also discuss the relation between symmetric wavelets and linear-phase filtering.

2.1 Multiresolution analysis

If some wavelet $\psi \in L^2(\mathbf{R})$ has to be constructed, it is advisable to study the structure of the $L^2(\mathbf{R})$ decomposition it generates. Let $\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k)$ and

$$W_j := clos_{L^2(\mathbf{R})} < \psi_{j,k} : k \in \mathbf{Z} > . \tag{2.1}$$

Then every $f \in L^2(\mathbf{R})$ has a unique decomposition

$$f(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots , \tag{2.2}$$

where $g_j \in W_j$ for all $j \in \mathbf{Z}$, and we shall describe this by writing

$$L^2(\mathbf{R}) = \sum_{\mathbf{j}\in\mathbf{Z}} \mathbf{W_j} := \cdots \dot{+}\mathbf{W_{-1}}\dot{+}\mathbf{W_0}\dot{+}\mathbf{W_1}\dot{+}\cdots \tag{2.3}$$

Being in W_j, the component g_j of f has a unique wavelet series representation, where the coefficient sequence gives localized spectral information of f in the j-th octave (or frequency band) in terms of integral wavelet transform of f with the dual $\tilde{\psi}$ of ψ as the basic wavelet. Using the decomposition of $L^2(\mathbf{R})$ in (2.3), we also have a nested sequence of closed subspaces V_j defined by

$$V_j := \cdots \dot{+}W_{j-2}\dot{+}W_{j-1} \tag{2.4}$$

Lemma 2.1 *The subspaces V_j satisfy:*

1. $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$
2. $clos_{L^2(\mathbf{R})} \left(\cup_{j \in \mathbf{Z}} V_j \right) = L^2(\mathbf{R})$
3. $\cap_{j \in \mathbf{Z}} V_j = \{0\}$
4. $V_{j+1} = V_j \dot{+} W_j, \quad j \in \mathbf{Z}$
5. $f(x) \in V_j \leftrightarrow f(2x) \in V_{j+1}$
6. $f(x) \in V_j \leftrightarrow f(x + 1/2^j) \in V_j, \quad j \in \mathbf{Z}$

Definition 2.1 *A function $\phi \in L^2(\mathbf{R})$ is called a scaling function, if the subspaces V_j defined by*

$$V_j := clos_{L^2(\mathbf{R})} < \phi_{j,k} : k \in \mathbf{Z} >, \quad j \in \mathbf{Z}, \tag{2.5}$$

satisfy the properties 1,2,5, and 6 stated above in this section, and if $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is a Riesz basis of V_0. We say that the scaling function ϕ generates a multiresolution analysis $\{V_j\}$ of $L^2(\mathbf{R})$.

Remark: All properties 1-6 will be assumed in any MRA$\{V_j\}$ of $L^2(\mathbf{R})$.

If ϕ generates an MRA, then since $\phi \in V_0$ is also in V_1 and since $\{\phi_{1,k} : k \in \mathbf{Z}\}$ is a Riesz basis of V_1, there exists a unique l^2-sequence $\{p_k\}$ that describes the two-scale relation

$$\phi(x) = \sum_{k=-\infty}^{\infty} p_k \phi(2x - k) \tag{2.6}$$

of the scaling function ϕ. This sequence $\{p_k\}$ is called the two-scale sequence of ϕ. Let

$$P(z) := \frac{1}{2} \sum_{k=-\infty}^{\infty} p_k z^k. \tag{2.7}$$

The following Fourier transform formulation:

$$\hat{\phi}(\omega) = P(z) \hat{\phi}\left(\frac{\omega}{2}\right), \quad z = e^{-i\omega/2}, \tag{2.8}$$

of the identity (2.6). We will call P the two-scale symbol of the scaling function ϕ.

We will make the following assumption on ϕ and its two-scale sequence:

1. $\phi \in L^1(\mathbf{R})$
2. $\sum_{k=-\infty}^{\infty} \phi(x-k) = 1$
3. $\{p_k\} \in l^1$

The assumption in 2 is called the property of partition of unity of ϕ. Observe that every cardinal B-spline satisfies 2. Assumption 1 implies that $\hat{\phi}$ is a continuous function on $\mathbf{R}$. Hence, it follows from Poisson Summation formula that 2 is a consequence of the following condition on $\hat{\phi}$:

$$\hat{\phi}(0) = 1, \quad \hat{\phi}(2\pi k) = 0, \quad 0 \neq k \in \mathbf{Z} \tag{2.9}$$

Finally, the assumption in 3 guarantees that P is a continuous function on the unit circle $|z| = 1$.

From the continuity of P on $|z| = 1$ and the first condition in (2.9), and by applying (2.8), we have

$$P(1) = \frac{1}{2} \sum_k p_k = 1. \tag{2.10}$$

On the other hand, it follows from the assumption that $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is a Riesz basis of V_0 and second condition of $\hat{\phi}$ in (2.9) that $P(z)$ also satisfies

$$P(-1) = \frac{1}{2} \sum_k (-1)^k p_k = 0. \tag{2.11}$$

Of course, an equivalent statement of (2.10) and (2.11) is

$$\sum_k p_{2k} = \sum_k p_{2k+1} = 1. \tag{2.12}$$

As another consequence of the continuity of $\hat{\phi}$ and the condition $\hat{\phi}(0) = 1$, we observe, by repeated application of (2.8), that as $n \to \infty$,

$$\begin{aligned} \hat{\phi}(\omega) &= \left(\prod_{k=1}^{n} P(e^{-i\omega/2^k}) \right) \hat{\phi}\left(\frac{\omega}{2^n}\right) \\ &\to \prod_{k=1}^{\infty} P(e^{-i\omega/2^k}), \quad \omega \in \mathbf{R} \end{aligned}$$

pointwise, provided that the infinite product converges.

Example: For the m^{th} order cardinal B-spline N_m, we have

$$P(z) = \left(\frac{1+z}{2}\right)^m$$

so that

$$\begin{aligned}
\prod_{k=1}^{n} P(e^{-i\omega/2^k}) &= \prod_{k=1}^{n} \left(\frac{1+e^{-i\omega/2^k}}{2}\right)^m \\
&= \prod_{k=1}^{n} \left(\frac{1+e^{-i\omega/2^k}}{2}\frac{1-e^{-i\omega/2^k}}{1-e^{-i\omega/2^k}}\right)^m \\
&= \prod_{k=1}^{n} \left(\frac{1}{2}\frac{1-e^{-i\omega/2^{k-1}}}{1-e^{-i\omega/2^k}}\right)^m \\
&= \frac{1}{2^{mn}}\left(\frac{1-e^{i\omega}}{1-e^{-i\omega/2^n}}\right)^m \to \left(\frac{1-e^{-i\omega}}{i\omega}\right)^m \diamond
\end{aligned}$$

In view of the preceding spline example, we will restrict our attention to two-scale equation with governing sequence $\{p_k\}$ given by

$$P(z) = \frac{1}{2}\sum_k p_k z^k = \left(\frac{1+z}{2}\right)^N S(z), \tag{2.13}$$

where N is some positive integer, $S(1) = 1$, and $S(z)$ is sufficiently smooth on the unit circle $|z| = 1$.

Definition 2.2 *A Laurent series $P(z)$ of the form (2.13) is called an admissible two-scale symbol if S is a continuous on the unit circle satisfying*

- $S(1) = 1$
- *as a function of ω, the $L^\infty(0, 2\pi)$ modulus of continuity of $S(e^{-i\omega})$ is of order $O(\eta^\alpha)$ for some α, with $0 < \alpha \leq 1$, as $\eta \to 0^+$.*

For any admissible two-scale symbol P with factor S as in (2.13), let us consider the bounds $B_j = B_j(S)$ and $b_j = b_j(S)$ defined by

$$B_j = B_j(S) := \sup_\omega \left|\prod_{k=1}^{j} S(e^{-i\omega/2^k})\right|;$$

$$b_j = b_j(S) := \frac{1}{j} \log_2 B_j$$

We have the following convergence result.

Theorem 2.1 *Let P be an admissible two-scale symbol of the form (2.13). Then the infinite product*

$$g(\omega) := \prod_{k=1}^{\infty} P(e^{-i\omega/2^k}) \tag{2.14}$$

converges pointwise everywhere to some function g. Furthermore, for every positive integer n_0, there exists some positive constant C_{n_0}, such that the limit function g satisfies

$$|g(\omega)| \leq C_{n_0}(1 + |\omega|)^{-N+b_{n_0}} \tag{2.15}$$

where b_{n_0} is defined above. In particular, if there is some n_0 such that $b_{n_0} < N - 1/2$, then there exists a function $\phi \in L^2(\mathbf{R})$ such that $\hat{\phi} = g, \hat{\phi}(0) = 1$ and $\hat{\phi}$ satisfies the two-scale relation (2.8).

2.2 Finite two-scale relation

In this section, we restrict our attention to two-scale relation (2.6) described by finite sums. A very important consequence of this restriction is that the corresponding scaling functions necessarily have compact supports. We will also study the class of all scaling functions with finite two-scale relations that generate the same MRA, and investigate the ones with minimum supports. It will be clear that the smaller the supports of a scaling function ϕ and its corresponding wavelet ψ are, the shorter the reconstruction sequences used in the wavelet reconstruction algorithm become.

Let ϕ be a scaling function described by the two-scale relation

$$\phi(x) = \sum_{k=0}^{N} p_k \phi(2x - k), \quad p_0, p_N \neq 0. \tag{2.16}$$

We remark that by a change of index in p_k, any finite two-scale relation can be written as in (2.16). Let us first take care of the cases $N = 0, 1$ in (2.16).

(i) For $N = 0$ we have

$$\phi(x) = 2\phi(2x)$$

so that the two-scale symbol is $P(z) = 1$ and the infinite product in (2.14) is $g(\omega) = 1$ for all ω. So if $g = \hat{\phi}$, then ϕ must be the delta distribution.

(ii) For $N = 1$ we have

$$\phi(x) = \phi(2x) + \phi(2x-1)$$

which is the same as the two-scale relation of the first order cardinal B-spline N_1. So $\phi = N_1$.

We will only be concerned with scaling function which are continuous. Under this additional assumption, it is possible to show that for $N = 2$, the two-scale relation must be given by

$$\phi(x) = \frac{1}{2}\phi(2x) + \phi(2x-1) + \frac{1}{2}\phi(2x-2), \tag{2.17}$$

which is identical with the two-scale relation of the second order cardinal B-spline N_2, hence $\phi = N_2$. When $N = 3$, we have the quadratic cardinal B-spline N_3 whose two-scale relation is

$$N_3(x) = \frac{1}{4}N_3(2x) + \frac{3}{4}N_3(2x-1) + \frac{3}{4}N_3(2x-2) + \frac{1}{4}N_3(2x-3) \tag{2.18}$$

However, there is another alternative, namely: *Daubechies' scaling function* ϕ_3^D governed by

$$\begin{aligned} \phi_3^D(x) &= \frac{1+\sqrt{3}}{4}\phi_3^D(2x) + \frac{3+\sqrt{3}}{4}\phi_3^D(2x-1) \\ &+ \frac{3-\sqrt{3}}{4}\phi_3^D(2x-2) + \frac{1-\sqrt{3}}{4}\phi_3^D(2x-3) \end{aligned}$$

We have

$$\begin{aligned} P(z) &= \frac{1}{2}\sum_{k=0}^{3} p_k z^k \\ &= \frac{1}{2}\left\{\frac{1+\sqrt{3}}{4} + \frac{3+\sqrt{3}}{4}z + \frac{3-\sqrt{3}}{4}z^2 + \frac{1-\sqrt{3}}{4}z^3\right\} \\ &= \left(\frac{1+z}{2}\right)^2\left(\frac{(1+\sqrt{3}) + (1-\sqrt{3})z}{2}\right) \end{aligned}$$

which satisfies the admissibility condition, with $N = 2$ and $S(z)$ being a trigonometric polynomial such that $S(1) = 1$.

To understand a scaling function ϕ better, we consider the recursive scheme

$$\phi_n(x) = \sum_k p_k \phi_{n-1}(2x - k), \quad n = 1, 2, \ldots, \tag{2.19}$$

for some suitable initial function ϕ_0. By considering the Fourier transform of (2.19), we have

$$\begin{aligned} \hat{\phi}_n(\omega) &= P(e^{-i\omega/2})\hat{\phi}_{n-1}\left(\frac{\omega}{2}\right) = \cdots \\ &= \left\{\prod_{k=1}^{n} P(e^{-i\omega/2^k})\right\}\hat{\phi}_0\left(\frac{\omega}{2^n}\right) \end{aligned}$$

Hence, if P is an admissible two-scale symbol and the Fourier transform $\hat{\phi}_0$ of the initial choice ϕ_0 is continuous at $\omega = 0$ and satisfies $\hat{\phi}_0(0) = 1$, then by Theorem 2.1

$$\lim_{n\to\infty} \hat{\phi}_n(\omega) = g(\omega)$$

where g is the infinite product.

We see that for a two-scale sequence with at least three non-zero terms the second order cardinal B-spline N_2, being a continuous spline function with lowest order, provides a good choice as the initial function in (2.19) for producing the scaling function ϕ. That is, we recommend the following recursive scheme

$$\begin{cases} \phi(x) = \lim_{n\to\infty} \phi_n(x), & \textit{where} \\ \phi_n(x) = \sum_{k=0}^{N} p_k \phi_{n-1}(2x - k), & n = 1, 2, \ldots, \\ \phi_0(x) = N_2(x) \end{cases}$$

As a consequence of the process, we see that ϕ has compact support, and in fact, we can find its support exactly, provided that ϕ is continuous. It is interesting to observe that, supp ϕ_n increase monotonically with n. More precisely, we have

$$\begin{cases} \textit{supp}\ \phi_0 = [0, 2] \\ \textit{supp}\ \phi_1 = [0, \frac{1}{2}(2 + N)] = [0, \frac{2+N}{2}] \\ \textit{supp}\ \phi_2 = [0, \frac{1}{2}(\frac{2+N}{2} + N)] = [0, \frac{2+(2^2-1)N}{2^2}] \\ \cdots\cdots\cdots\cdots \\ \textit{supp}\ \phi_n = [0, \frac{2+(2^n-1)N}{2^n}] \end{cases}$$

and hence, since $N \geq 2$, we have

$$supp\ \phi_n \subseteq [0, N], \quad n = 1, 2, \ldots,$$

and it follows that

$$supp\ \phi = [0, N] \tag{2.20}$$

Knowing that the support of ϕ is $[0, N]$, is a tremendous help in computing $\phi(x)$, at least at all the dyadic points $x = k/2^j$, where $j, k \in \mathbf{Z}$. This is evident by referring to the two-scale relation (2.16). If fact, if the value of $\phi(1), \ldots, \phi(N-1)$ are known, then since $\phi(k) = 0$ for all $k \leq 0$ or $k \geq N$, the relations

$$\begin{cases} \phi\left(\frac{k}{2}\right) = \sum_l p_l \phi(k-l), \\ \phi\left(\frac{k}{2^2}\right) = \sum_l p_l \phi\left(\frac{k}{2} - l\right), \\ \cdots\cdots\cdots\cdots \end{cases}$$

uniquely determine all the values of $\phi(x)$ at $x = k/2^j, j, k \in \mathbf{Z}$.

To determine the values of $\phi(k), k \in \mathbf{Z}$, we again use the two-scale relation (2.16) with x being an integer. That is, in matrix notation, we have

$$m = Mm, \tag{2.21}$$

where m is the column vector

$$m := [\phi(1) \ldots \phi(N-1)]^T \tag{2.22}$$

and M the $(n-1) \times (n-1)$ matrix

$$M := [p_{2j-k}]_{1 \leq j,k \leq N-1}, \tag{2.23}$$

with j being the row index and k the column index. Recalling that ϕ generates a partition of unity, we can determine the values of $\phi(k), k \in \mathbf{Z}$, simply by finding the eigenvector m in (2.21) corresponding to the eigenvalue 1 and imposing the normalization condition

$$\phi(1) + \cdots \phi(N-1) = 1. \tag{2.24}$$

Example: Determine the value of $\phi_3^D(k), k \in \mathbf{Z}$, where the two-scale relation of ϕ_3^D.

Solution: We have $N = 3$ and the matrix M

$$M = \begin{bmatrix} p_1 & p_0 \\ p_3 & p_2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 3+\sqrt{3} & 1+\sqrt{3} \\ 1-\sqrt{3} & 3-\sqrt{3} \end{bmatrix}$$

It is easy to see that the solution space is

$$m = a[1+\sqrt{3} \;\; 1-\sqrt{3}]^T, \quad a \in \mathbf{R}$$

So, by the normalization condition (2.24), we have $a = 1/2$ and

$$\begin{cases} \phi_3^D(1) = \frac{1+\sqrt{3}}{2} \\ \phi_3^D(2) = \frac{1-\sqrt{3}}{2} \end{cases} \diamond$$

We now turn to a study of class Φ of all the scaling functions ϕ with finite two-scale relations that generate the same MRA $\{V_j\}$ of $L^2(\mathbf{R})$. Without loss of generality, we may assume that the two-scale relation of any $\phi \in \Phi$ takes on the form (2.16), and hence, the support of ϕ is precisely the interval $[0, N]$. So $\phi^* \in \Phi$ has minimum support if and only if

$$N_{\phi^*} \leq N_\phi, \quad \phi \in \Phi \tag{2.25}$$

Corresponding to any $\phi \in \Phi$, let us consider the autocorrelation function

$$F_\phi(x) := \int_\infty^\infty \phi(x+y)\overline{\phi(y)}dy \tag{2.26}$$

and the symbol of the sequence $\{F_\phi(k)\}$, namely:

$$E_\phi(z) := \sum_k F_\phi(k) z^k \tag{2.27}$$

Since F_ϕ clearly satisfies:

$$\begin{cases} F_\phi(-x) = \overline{F_\phi(x)}, \quad x \in \mathbf{R} \\ \quad supp\ F_\phi \subseteq [-N, N] \end{cases}$$

it follows that E_ϕ is a Laurent polynomial. Let k_ϕ denote the one-sided degree of E_ϕ; that is, k_ϕ is the largest integer for which $F_\phi(k_\phi) \neq 0$. Then

$$\Pi_\phi(z) := z^{k_\phi} E_\phi(z) \tag{2.28}$$

is an algebraic polynomial of degree $2k_\phi$, and the reciprocal polynomial of Π_ϕ is given by

$$\Pi^r_\phi(z) := z^{2k_\phi}\overline{\Pi_\phi\left(\frac{1}{\bar{z}}\right)} \tag{2.29}$$

It is clear that

$$\Pi^r_\phi(z) = \Pi_\phi(z), all\ z. \tag{2.30}$$

We call Π_ϕ the generalized Euler-Frobenius polynomial

and E_ϕ the generalized Euler-Frobenius Laurent polynomial relative ϕ.

Definition 2.3 *Let z_0 be a zero of an algebraic polynomial $p(z)$. Then we call z_0 a symmetric zero of $p(z)$ if $z_0 \neq 0$ and $p(-z_0) = p(z_0) = 0$.*

In the following theorem concerning scaling function ϕ with finite two-scale sequence, recall that $P_\phi(z)$ denote the two-scale symbol of ϕ.

Theorem 2.2 *let $\phi \in \Phi$ be any scaling function governed by (2.16). Then*

1. *both $E_\phi(z)$ and $\Pi_\phi(z)$ never vanish on $|z| = 1$*
2. *for all ω*
$$E_\phi(e^{-i\omega}) = \sum_{k=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi k)|^2$$
3. *for all ω*
$$|P_\phi(e^{-i\omega/2})|^2 E_\phi(e^{-i\omega/2}) + |P_\phi(-e^{-i\omega/2})|^2 E_\phi(-e^{-i\omega/2}) = E_\phi(e^{-i\omega})$$
4. *P_ϕ has no symmetric zeros that lie on $|z| = 1$.*

Theorem 2.3 *A scaling function $\phi^* \in \Phi$ has minimum support if and only if its two-scale symbol P_{ϕ^*} has no symmetric zeros*

Theorem 2.4 *For any ϕ_1, ϕ_2, the symbol $C(z)$ of the sequence $\{s_n\}$ relating ϕ_1 and ϕ_2, in the sense that*

$$\phi_2(x) = \sum_{n=-\infty}^{\infty} s_n \phi_1(x - n),$$

is a rational function which is both zero-free and pole-free on the unit circle $|z| = 1$. *In addition, if* $\phi_1 \in \Phi$ *has minimum support, then* $C(z)$ *is a polynomial: that is, every* $\phi_2 \in \Phi$ *is a finite linear combination of integer translates of the minimally supported* $\phi_1 \in \Phi$. *In particular, the minimally supported* $\phi_1 \in \Phi$ *is unique.*

2.3 Direct-sum decomposition of $L^2(\mathbf{R})$

Definition 2.4 *A Laurent series is said to belong to the* Wiener Class **W** *if its coefficient sequence is in* l^1.

Since the discrete convolution of two l^1-sequences is a sequence in l^1, it is clear that $\mathbf{W}$ is an algebra.

Theorem 2.5 *Let* $f \in \mathbf{W}$ *and suppose that* $f(z) \neq 0$ *for all* z *on the unit circle* $|z| = 1$. *Then* $\frac{1}{f} \in \mathbf{W}$ *also.*

Let ϕ be a scaling function whose two-scale symbol

$$P_\phi(z) = \frac{1}{2} \sum_{k=-\infty}^{\infty} p_k z^k \tag{2.31}$$

is in $\mathbf{W}$. Recall that P_ϕ governs the relation $V_0 \subset V_1$ in the sense that

$$\phi(x) = \sum_k p_k \phi(2x - k) \tag{2.32}$$

and ϕ generates V_0. Let us now consider any other l^1-sequence $\{q_k\}$ and its symbol

$$Q(z) = \frac{1}{2} \sum_{k=-\infty}^{\infty} q_k z^k \tag{2.33}$$

Then Q is also in $\mathbf{W}$ and define a function

$$\psi(x) := \sum_k q_k \phi(2x - k) \tag{2.34}$$

in V_1. This function ψ also generates a closed subspace W_0 in the same manner as ϕ generates V_0, namely:

$$W_0 := clos_{L^2(\mathbf{R})} < \psi(\cdot - k) : k \in \mathbf{Z} > . \tag{2.35}$$

Hence

$$P = P_\phi$$

does, the symbol Q governs the relation $W_0 \subset V_1$ in the sense (2.34) and (2.35) are satisfied.

Our main concern in the construction of wavelets is at least to ensure that V_0 and W_0 are complementary subspaces of V_1 in the sense that

$$V_0 \cap W_0 = \{0\} \quad V_1 = V_0 + W_0 \tag{2.36}$$

That is, V_1 is the direct sum of V_0 and W_0, and notation

$$V_1 = V_0 \dot{+} W_0 \tag{2.37}$$

In the following, we will see the matrix

$$M_{P,Q}(z) := \begin{bmatrix} P(z) & Q(z) \\ P(-z) & Q(-z) \end{bmatrix} \tag{2.38}$$

plays an essential role in characterizing (2.37). Hence, we must consider the determinant

$$\Delta_{P,Q}(z) := \det M_{P,Q}(z) \tag{2.39}$$

Since P and Q are in $\mathbf{W}$, we have $\Delta_{P,Q} \in \mathbf{W}$ also. In addition, if $\Delta_{P,Q}(z) \neq 0$ on $|z| = 1$, we have $1/\Delta_{P,Q} \in \mathbf{W}$. So, under the condition $\Delta_{P,Q} \neq 0$ on $|z| = 1$, the two functions

$$\begin{cases} G(z) := \frac{Q(-z)}{\Delta_{P,Q}(z)} \\ H(z) := \frac{-P(-z)}{\Delta_{P,Q}(z)} \end{cases} \tag{2.40}$$

are both in the Wiener class $\mathbf{W}$. The reason for considering the functions G and H is that the transpose $M^T_{G,H}$ of $M_{G,H}$ is the inverse of $M_{P,Q}$, namely:

$$\begin{cases} M_{P,Q}(z) M^T_{G,H}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ M^T_{G,H}(z) M_{P,Q}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad |z| = 1 \end{cases} \tag{2.41}$$

The first identity in (2.41) is equivalent to the pair of identities

$$\begin{cases} P(z)G(z) + Q(z)H(z) = 1 \\ P(z)G(-z) + Q(z)H(-z) = 0 \quad |z| = 1, \end{cases} \tag{2.42}$$

while the second identity in (2.41) is equivalent to the following set of four identities:

$$\begin{cases} P(z)G(z) + P(-z)G(-z) = 1; \\ P(z)H(z) + P(-z)H(-z) = 0; \\ G(z)Q(z) + G(-z)Q(-z) = 0; \\ Q(z)H(z) + Q(-z)H(-z) = 1; \quad |z| = 1, \end{cases} \tag{2.43}$$

For $L^2(\mathbf{R})$ decomposition, we do not need the identities in (2.43). However, this set of identities will be crucial to our discussion of *duality* in the next section.

Since $G, H \in \mathbf{W}$, we may write

$$\begin{cases} G(z) = \frac{1}{2}\sum_{n=-\infty}^{\infty} g_n z^n \\ H(z) = \frac{1}{2}\sum_{n=-\infty}^{\infty} h_n z^n \end{cases} \tag{2.44}$$

where $\{g_n\}, \{h_n\} \in l^1$, whenever $\Delta_{P,Q}(z) \neq 0$ on the unit circle. We now ready to formulate the following decomposition result.

Theorem 2.6 *A necessary and sufficient condition for the direct-sum decomposition (2.37) to hold is that the function $\Delta_{P,Q}$ never vanishes on the unit circle $|z| = 1$. Furthermore, if $\Delta_{P,Q}(z) \neq 0$ for all $|z| = 1$, then the family $\{\psi(\cdot - k) : k \in \mathbf{Z}\}$, governed by $Q(z)$ as in (2.34), is a Riesz basis of W_0, and the decomposition relation*

$$\phi(2x - l) = \frac{1}{2}\sum_{k=-\infty}^{\infty} \{g_{2k-l}\phi(x-k) + h_{2k-l}\psi(x-k)\}, l \in \mathbf{Z}, \tag{2.45}$$

holds for all $x \in \mathbf{R}$.

Proof: Assume that $\Delta_{P,Q}(z) \neq 0$ for all z satisfying $|z| = 1$. As a consequence, all sequence to be considered are in l^1 and there will be no danger in interchanging the orders of summation.

Observe that as an equivalent formulation of (2.42), we have

$$\begin{cases} P(z)(G(z)+G(-z))+Q(z)(H(z)+H(-z))=1; \\ P(z)(G(z)-G(-z))+Q(z)(H(z)-H(-z))=1, |z|=1 \end{cases} \tag{2.46}$$

which, in view of (2.44), may be written as

$$\begin{cases} P(z)\sum_k g_{2k}z^{2k}+Q(z)\sum_k h_{2k}z^{2k}=1; \\ P(z)\sum_k g_{2k-1}z^{2k-1}+Q(z)\sum_k h_{2k-1}z^{2k-1}=1 \end{cases} \tag{2.47}$$

Hence, by setting $z=e^{-i\omega/2}$ and multiplying the two identities in (2.47) by $\hat{\phi}\left(\frac{\omega}{2}\right)$ and $z\hat{\phi}\left(\frac{\omega}{2}\right)$, consequently, we have

$$\begin{cases} \hat{\phi}\left(\frac{\omega}{2}\right)=\sum_k \left(g_{2k}z^{2k}P(z)\hat{\phi}\left(\frac{\omega}{2}\right)+h_{2k}z^{2k}Q(z)\hat{\phi}\left(\frac{\omega}{2}\right)\right); \\ \hat{\phi}\left(\frac{\omega}{2}\right)e^{-i\omega/2}=\sum_k \left(g_{2k-1}z^{2k}P(z)\hat{\phi}\left(\frac{\omega}{2}\right)+h_{2k-1}z^{2k}Q(z)\hat{\phi}\left(\frac{\omega}{2}\right)\right) \end{cases}$$

which is equivalent to

$$\begin{cases} \hat{\phi}\left(\frac{\omega}{2}\right)=\sum_k (g_{2k}z^{2k}\hat{\phi}(\omega)+h_{2k}z^{2k}\hat{\psi}(\omega)); \\ \hat{\phi}\left(\frac{\omega}{2}\right)e^{-i\omega/2}=\sum_k (g_{2k-1}z^{2k}\hat{\phi}(\omega)+h_{2k-1}z^{2k}\hat{\psi}(\omega)) \end{cases} \tag{2.48}$$

where the Fourier transform formulations of (2.32) and (2.34) have been used. Consequently, by taking the inverse Fourier transform on both sides in (2.48), we obtain

$$\begin{cases} 2\phi(2x)=\sum_k (g_{2k}\phi(x-k)+h_{2k}\psi(x-k)); \\ 2\phi(2x-1)=\sum_k (g_{2k-1}\phi(x-k)+h_{2k-1}\psi(x-k)) \end{cases} \tag{2.49}$$

It is clear that (2.49) is equivalent to (2.45). As a consequence, since $\{g_k\}$ and $\{h_k\}$ are in l^1, and since

$$V_1 = clos_{L^2(\mathbf{R})} < \phi(2\cdot -k) : k \in \mathbf{Z} >,$$

we have now shown that $V_1 \subset V_0 + W_0$, so that

$$V_1 = V_0 + W_0$$

To prove that this is a direct sum, we consider

$$\sum_k a_k\phi(x-k)+\sum_k b_k\psi(x-k)=0, \tag{2.50}$$

where $\{a_k\}$ and $\{b_k\}$ are in l^1. Then by applying the two-scale relations in (2.32) and (2.34), we obtain

$$\sum_l \left(\sum_k a_k p_{l-2k} + \sum_k b_k q_{l-2k} \right) \phi(2x - l) = 0,$$

so that

$$\sum_k a_k p_{l-2k} + \sum_k b_k q_{l-2k} = 0, \quad l \in \mathbf{Z}, \tag{2.51}$$

by referring to the fact that $\{\phi(2 \cdot -l) : l \in \mathbf{Z}\}$ is a Riesz basis of V_1.

Now, taking the symbols of both sides of (2.51), we have

$$A(z^2)P(z) + B(z^2)Q(z) = 0, \tag{2.52}$$

where A and B denote the symbols of $\{a_k\}$ and $\{b_k\}$, respectively. So if z is also replaced by $-z$, then (2.52) gives rise to the linear equations:

$$\begin{cases} P(z)A(z^2) + Q(z)B(z^2) = 0; \\ P(-z)A(z^2) + Q(-z)B(z^2) = 0, \end{cases}$$

with two unknowns $A(z^2)$ and $B(z^2)$, where the coefficient matrix is $M_{P,Q}(z)$, which is nonvanishing for all z on $|z| = 1$. Hence $A(z^2)$ and $B(z^2)$ must be zero, and the l^2-sequences $\{a_k\}$ and $\{b_k\}$ in (2.50) are trivial. This proves that $V_0 \cap W_0 = \{0\}$.

To prove that the family $\{\psi(\cdot - k) : k \in \mathbf{Z}\}$ is a Riesz basis of W_0. In particular, since $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is a Riesz basis of V_0, we have

$$0 < A \leq \sum_k |\hat{\phi}(\omega + 2\pi k)|^2 \leq B < \infty, \quad \omega \in \mathbf{R} \tag{2.53}$$

Also, it follows from the Fourier transform of (2.34) that

$$\begin{aligned} \sum_k |\hat{\psi}(\omega + 2\pi k)|^2 &= \sum_k \left| Q(e^{-i(\omega/2+\pi k)}) \right|^2 \left| \hat{\phi}(\frac{\omega}{2} + \pi k) \right|^2 \\ &= |Q(z)|^2 \sum_k \left| \hat{\phi}(\frac{\omega}{2} + 2\pi k) \right|^2 \\ &+ |Q(-z)|^2 \sum_k \left| \hat{\phi}(\frac{\omega}{2} + \pi + 2\pi k) \right|^2 \end{aligned}$$

where $z = e^{-i\omega/2}$, so that an application of (2.53) yields

$$\begin{aligned} A\{|Q(z)|^2 + |Q(-z)|^2\} &\leq \sum_k |\hat{\psi}(\omega + 2\pi k)|^2 \\ &\leq B\{|Q(z)|^2 + |Q(-z)|^2\} \end{aligned}$$

Since $Q \in \mathbf{W}$, it is continuous on $|z| = 1$, and we have

$$B' := 2 \max_{|z|=1} |Q(z)| < \infty \tag{2.54}$$

On the other hand, in view of

$$\Delta_{P,Q}(z) = \det \begin{bmatrix} P(z) & Q(z) \\ P(-z) & Q(-z) \end{bmatrix} \neq 0, \quad |z| = 1$$

we see that not both $Q(z)$ and $Q(-z)$ can vanish at the same z on the unit circle, and so, by the continuity of Q on $|z| = 1$, we have

$$A' := \min_{|z|=1} (|Q(z)|^2 + |Q(-z)|^2) > 0 \tag{2.55}$$

hence, it follows from (2.54) and (2.55) that

$$AA' \leq \sum_k |\hat{\psi}(\omega + 2\pi k)|^2 \leq BB' \tag{2.56}$$

or $\{\psi(\cdot - k) : k \in \mathbf{Z}\}$ is a Riesz basis of W_0 $\diamond$

Remark: (i) From (2.33), it follows that

$$\begin{aligned} \int_{-\infty}^{\infty} \psi(x) dx &= \sum_{k=-\infty}^{\infty} q_k \int_{-\infty}^{\infty} \phi(2x - k) dx \\ &= \frac{1}{2} \sum_{k=-\infty}^{\infty} q_k \hat{\phi}(0) = Q(1) \end{aligned}$$

As usual, let

$$\psi_{j,k}(x) := 2^{j/2} \psi(2^j x - k).$$

Theorem 2.6 says that for each $j \in \mathbf{Z}$, the family $\{\psi_{j,k} : k \in \mathbf{Z}\}$ is a Riesz basis of W_j. However, the whole family $\{\psi_{j,k} : j, k \in \mathbf{Z}\}$ is not necessarily Riesz basis of $L^2(\mathbf{R})$. Indeed, for a function ψ to generate a Riesz basis of $L^2(\mathbf{R})$ such that $\hat{\psi}$ is continuous, its integral over must be zero, and so, a necessary condition is that

$$Q(1) = 0 \tag{2.57}$$

(ii) Even if ψ would generate a Riesz basis of $L^2(\mathbf{R})$, ψ may bot be a wavelet, since the existence of the dual $\tilde{\psi}$ of ψ still has to be investigated. Recall that in any series representation

$$f(x) = \sum_{j,k} c_{j,k}\psi_{j,k}(x), \quad f \in L^2(\mathbf{R}),$$

it requires a dual $\tilde{\psi}$ of ψ to extract any time-frequency information of f from the coefficients $c_{j,k}$.

2.4 Wavelets and their duals

Wavelet decomposition requires the function ψ, governed by the Laurent series $Q \in \mathbf{W}$ according to (2.34), to be a wavelet with some dual wavelet $\tilde{\psi}$. In particular, Q must satisfy (2.57). Recall that the two-scale symbol $P \in \mathbf{W}$ must also satisfy the conditions in (2.10) and (2.11). Hence, P and Q necessarily satisfy the conditions

$$\begin{cases} P(1) = 1 \quad P(-1) = 0; \\ Q(1) = 0 \end{cases} \tag{2.58}$$

Let G and H be the Laurent series defined in previous section. Then we have $G, H \in \mathbf{W}$, and the four Laurent series P, Q, G, H satisfy the identities in (2.42). It follows from this set of identities and (2.58) that G must also satisfy the conditions

$$G^*(1) = 1 \quad G^*(-1) = 0, \tag{2.59}$$

where the notation

$$G^*(z) := \overline{G(z)} = \overline{G}\left(\frac{1}{z}\right), \quad |z| = 1 \tag{2.60}$$

The similarity between P and G^*, as described by (2.58) and (2.59), suggests that

$$G^*(z) = \frac{1}{2} \sum_{n=-\infty}^{\infty} \overline{g}_{-n} z^n, \quad |z| = 1 \tag{2.61}$$

should also be chosen as the two-scale symbol of scaling function that generates a possibly different MRA of $L^2(\mathbf{R})$.

This motivates the following strategy for constructing wavelets and their duals. We will start from the two admissible two-scale symbols $P = P_\phi$ and $G^* = G^*_{\tilde{\phi}}$ such that

$$\begin{cases} \hat{\phi}(\omega) = \prod_{k=1}^{\infty} P(e^{-i\omega/2^k}) \\ \hat{\tilde{\phi}}(\omega) = \prod_{k=1}^{\infty} G^*(e^{-i\omega/2^k}) \end{cases} \tag{2.62}$$

are in $L^2(\mathbf{R})$. Moreover, we assume that ϕ generates as MRA $\{V_j\}$ and $\tilde{\phi}$ generates an MRA $\{\tilde{V}_j\}$ of $L^2(\mathbf{R})$. Then according to Theorem 2.6, selecting any two arbitrary Laurent series Q and H that satisfy

$$\Delta_{P,Q}(z) \neq 0 \quad \Delta_{G,H}(z) \neq 0, \quad |z| = 1 \tag{2.63}$$

We will make use of the first identity in (2.43) to make a connection between these two decomposition.

Definition 2.5 *The two-scale symbols P and G^* are said to be* <u>*duals*</u> *of each other if they satisfy the identity*

$$P(z)G(z) + P(-z)G(-z) = 1, \quad |z| = 1 \tag{2.64}$$

Hence, if the two Laurent series Q and H are so chosen that the two nonvanishing matrices $M_{P,Q}(z)$ and $M^T_{G,H}(z)$ are inverse of each other on $|z| = 1$, that is,

$$M_{P,Q}(z)M^T_{G,H}(z) = M^T_{G,H}(z)M_{P,Q}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, |z| = 1 \tag{2.65}$$

then by (2.43) and the equivalence between this identity, we have

$$\begin{cases} P(z)H(z) + P(-z)H(-z) = 0; \\ G(z)Q(z) + G(-z)Q(-z) = 0; \\ Q(z)H(z) + Q(-z)H(-z) = 1, \quad |z| = 1 \end{cases} \tag{2.66}$$

Of course (2.65) is also equivalent to

$$\begin{cases} P(z)G(z) + Q(z)H(z) = 1; \\ P(-z)G(z) + Q(-z)H(z) = 0;, \quad |z| = 1 \end{cases} \tag{2.67}$$

Theorem 2.7 *Let P and G^* be dual two-scale symbols. Then the Laurent series Q and H in $\mathbf{W}$ satisfy (2.65) if and only if they are chosen from the class:*

$$Q(z) = z^{-1}G(-z)K(z^2) \quad H(z) = zP(-z)K^{-1}(z^2) \tag{2.68}$$

where $K \in \mathbf{W}$, with $K(z) \neq 0$ on $|z| = 1$.

Proof: It is easy to verify that every pair of Q and H from (2.68) satisfies (2.65). To derive the converse, we rely on the equivalence between (2.65) and (2.67). So, by applying Cramer's rule, we may express G and H in terms of P and Q, namely:

$$G(z) = \frac{Q(-z)}{\Delta_{P,Q}(z)} \quad H(z) = \frac{-P(-z)}{\Delta_{P,Q}(z)}, \quad |z| = 1 \tag{2.69}$$

where $\Delta_{P,Q}(z) = P(z)Q(-z) - P(-z)Q(z) \neq 0$ on $|z| = 1$. Since $\Delta_{P,Q}(-z) = -\Delta_{P,Q}(z)$, we may define

$$K(z^2) := z\Delta_{P,Q}(-z), \quad |z| = 1 \tag{2.70}$$

so that $K \in \mathbf{W}$ by Theorem 2.5, and $K(z) \neq 0$ for $|z| = 1$. Now (2.68) follows from (2.69) and (2.70). $\diamond$

We remark that by (2.59) and the first identity in (2.66), the pair (G^*, H^*) satisfies the condition

$$\begin{cases} G^*(1) = 1 \quad G^*(-1) = 0 \\ H^*(1) = 0 \end{cases} \tag{2.71}$$

which is the same set of conditions as in (2.58) for the pair (P, Q). In addition, in our strategy for constructing wavelets and dual wavelets through Q and H of the class described by (2.68), the two-scale symbols P and G^* play the same leading roles. Hence, the two pairs (P, Q) and (G^*, H^*) are interchangeable. This is called the duality principle

It is important to study the two admissible two-scale symbols P and G^* in more detail. According to Definition 2.2, we can write

$$\begin{cases} P(z) = \left(\frac{1+z}{2}\right)^N S(z), \\ G^*(z) = \left(\frac{1+z}{2}\right)^{\tilde{N}} \tilde{S}(z), \quad |z| = 1 \end{cases} \tag{2.72}$$

where N and $\tilde{N}$ are positive integers, $S(1) = \tilde{S}(1) = 1$, and the $L^\infty(0, 2\pi)$ moduli of continuity of both $S(e^{-i\omega})$ and $\tilde{S}(e^{-i\omega})$ are of order $O(\eta^\alpha)$ and $O(\eta^{\tilde{\alpha}})$. We will require the factors S and $\tilde{S}$ to satisfy,

$$\begin{cases} B := \max_{|z|=1} |S(z)| < 2^{N-1/2}; \\ \tilde{B} := \max_{|z|=1} |\tilde{S}(z)| < 2^{\tilde{N}-1/2} \end{cases} \tag{2.73}$$

Recall that the two admissible two-scale symbols P and G^* give rise to two scaling functions ϕ and $\tilde{\phi}$, as in (2.62). Although ϕ and $\tilde{\phi}$ might generate two different MRA's of $L^2(\mathbf{R})$, they could still be related in the following sense.

Definition 2.6 *Two scaling functions ϕ and $\tilde{\phi}$, generating possibly different MRA's $\{V_j\}$ and $\{\tilde{V}_j\}$, are said to be dual scaling functions, if they satisfy the condition*

$$< \phi(\cdot - j), \tilde{\phi}(\cdot - k) >= \int_{-\infty}^{\infty} \phi(x-j)\overline{\tilde{\phi}(x-k)}dx = \delta_{j,k}, j, k \in \mathbf{Z} \tag{2.74}$$

In the following, we shall give the connection between dual scaling functions and admissible two-scale symbols that are dual each other.

Theorem 2.8 *Let P and G^* be two admissible two-scale symbols as defined in (2.72). Also, let ϕ and $\tilde{\phi}$ be the corresponding scaling functions whose Fourier transforms are given by (2.62). If ϕ and $\tilde{\phi}$ are dual scaling functions, then P and G^* are dual to each other in the sense of (2.64). Conversely, if P and G^* are dual to each other and satisfy (2.73), then ϕ and $\tilde{\phi}$ are dual scaling functions.*

Proof: Let ϕ and $\tilde{\phi}$ be dual scaling functions. For each $n \in \mathbf{Z}$, we have

$$\delta_{n,0} \quad = \quad < \phi, \tilde{\phi}(\cdot - n) >= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\phi}(\omega)\overline{\hat{\tilde{\phi}}(\omega)} e^{in\omega} d\omega$$

$$= \sum_{k=-\infty}^{\infty} \frac{1}{2\pi} \int_{2\pi k}^{2\pi(k+1)} \hat{\phi}(\omega)\overline{\hat{\tilde{\phi}}(\omega)} e^{in\omega} d\omega$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=-\infty}^{\infty} \hat{\phi}(\omega + 2\pi k)\overline{\hat{\tilde{\phi}}(\omega + 2\pi k)} \right) e^{in\omega} d\omega$$

so that

$$\sum_{k=-\infty}^{\infty} \hat{\phi}(\omega + 2\pi k)\overline{\hat{\tilde{\phi}}(\omega + 2\pi k)} = 1, a.e. \tag{2.75}$$

Hence, setting $z = e^{-i\omega/2}$ and applying (2.75), we obtain

$$\begin{aligned}
\delta_{n,0} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} P(z)\overline{G^*(z)}\hat{\phi}\left(\frac{\omega}{2}\right) \overline{\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right)} e^{in\omega} d\omega \\
&= \frac{1}{2\pi} \int_0^{2\pi} \sum_k [P(z)G(z)\hat{\phi}\left(\frac{\omega}{2} + 2\pi k\right) \overline{\hat{\tilde{\phi}}\left(\frac{\omega}{2} + 2\pi k\right)} \\
&+ P(-z)G(-z)\hat{\phi}\left(\frac{\omega}{2} + \pi + 2\pi k\right) \overline{\hat{\tilde{\phi}}\left(\frac{\omega}{2} + \pi + 2\pi k\right)}] e^{in\omega} d\omega \\
&= \frac{1}{2\pi} \int_0^{2\pi} [P(z)G(z) + P(-z)G(-z)] e^{in\omega} d\omega,
\end{aligned}$$

so that, by the continuity of P and G on $|z| = 1$, we have

$$P(z)G(z) + P(-z)G(-z) = 1, \quad |z| = 1$$

That is, P and G^* are dual to each other.

To prove the converse, we fix a $j \in \mathbf{Z}$ and consider, for any positive integer n,

$$I_n := \frac{1}{2\pi} \int_{-2^n\pi}^{2^n\pi} \left(\prod_{k=1}^{n} P(e^{-i\omega/2^k})G(e^{-i\omega/2^k}) \right) e^{ij\omega} d\omega \tag{2.76}$$

Then by a change of variable $x = 2^{-n}\omega$, we have

$$\begin{aligned}
I_n &= 2^n \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\prod_{k=1}^{n} P(e^{-i2^{n-k}x})G(e^{-i2^{n-k}x}) \right) e^{ij2^n x} dx \\
&= 2^n \frac{1}{2\pi} \int_0^{\pi} \left(\prod_{k=1}^{n-1} P(e^{-i2^{n-k}x})G(e^{-i2^{n-k}x}) \right) \\
&\times [P(e^{-ix})G(e^{-ix}) + P(-e^{-ix})G(-e^{-ix})] e^{ij2^n x} dx
\end{aligned}$$

Now, by invoking the duality between P and G^* and making another change of variable $y = 2x$, it follows that

$$\begin{aligned} I_n &= 2^{n-1}\frac{1}{2\pi}\int_0^{2\pi}\left(\prod_{k=1}^{n-1} P(e^{-i2^{n-k-1}y})G(e^{-i2^{n-k-1}y})\right)e^{ij2^{n-1}y}dy \\ &= 2^{n-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\prod_{k=1}^{n-1} P(e^{-i2^{(n-1)-k}y})G(e^{-i2^{(n-1)-k}x})\right)e^{ij2^{n-1}y}dy \end{aligned}$$

Hence, comparing two equations above, we have $I_n = I_{n-1}$. Since this conclusion is valid for any positive integer n, we obtain

$$I_n = I_{n-1} = \cdots = I_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ij\omega}d\omega = \delta_{j,0} \tag{2.77}$$

Finally, the result (2.77) yields

$$\begin{aligned} <\phi, \tilde{\phi}(\cdot - j)> &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{\phi}(\omega)\overline{\hat{\tilde{\phi}}(\omega)}e^{ij\omega}d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\prod_{k=1}^{\infty} P(e^{-i\omega/2^k})G(e^{-i\omega/2^k})\right)e^{ij\omega}d\omega \\ &= \lim_{n\to\infty} I_n = I_0 = \delta_{j,0} \end{aligned}$$

This completes the proof of the theorem. ◇

Let us now select any Q and H from the class of functions in (2.69). By Theorem 2.7, the matrices $M_{P,Q}$ and M_{G^*,H^*} are invertible on $|z| = 1$, and so Theorem 2.6 applies. By considering the functions

$$\begin{cases} \psi(x) := \sum_k q_k\phi(2x-k); \\ \tilde{\psi}(x) := \sum_k \overline{h}_{-k}\tilde{\phi}(2x-k) \end{cases} \tag{2.78}$$

where

$$\begin{cases} Q(z) := \frac{1}{2}\sum_k q_k z^k; \\ H^*(z) := \frac{1}{2}\sum_k \overline{h}_{-k} z^k \end{cases} \tag{2.79}$$

and setting

$$\begin{cases} \psi_{j,k} := 2^{j/2}\psi(2^j\cdot -k); \\ \tilde{\psi}_{j,k} := 2^{j/2}\tilde{\psi}(2^j\cdot -k) \end{cases} \tag{2.80}$$

as well as

$$\begin{cases} W_j := clos_{L^2(\mathbf{R})} < \psi_{j,k} : k \in \mathbf{Z} >: \\ \tilde{W}_j := clos_{L^2(\mathbf{R})} < \tilde{\psi}_{j,k} : k \in \mathbf{Z} > \end{cases} \tag{2.81}$$

we have

$$\begin{cases} V_{j+1} = V_j \dot{+} W_j; \\ \tilde{V}_{j+1} = \tilde{V}_j \dot{+} \tilde{W}_j \end{cases} \tag{2.82}$$

Here, as usual, we set

$$\begin{cases} V_j := clos_{L^2(\mathbf{R})} < \phi_{j,k} : k \in \mathbf{Z} >: \\ \tilde{V}_j := clos_{L^2(\mathbf{R})} < \tilde{\phi}_{j,k} : k \in \mathbf{Z} > \end{cases} \tag{2.83}$$

where

$$\begin{cases} \phi_{j,k} := 2^{j/2}\phi(2^j \cdot -k); \\ \tilde{\phi}_{j,k} := 2^{j/2}\tilde{\phi}(2^j \cdot -k) \end{cases} \tag{2.84}$$

with ϕ and $\tilde{\phi}$ being the scaling functions whose two-scale symbols are P and G^*, respectively.

We shall next show that if the admissible two-scale symbols P and G^* are dual to each other in the sense that the identity

$$P(z)G(z) + P(-z)G(-z) = 1, \quad |z| = 1$$

is satisfied, then not only are $\{\psi_{j,k}\}$ and $\{\tilde{\psi}_{j,k}\}$ dual to each other, but additional orthogonality properties are achieved as well.

Theorem 2.9 *Let P and G^* be two admissible two-scale symbols which satisfy (2.73) and are dual to each other. Then for any $Q, H \in \mathbf{W}$ chosen from the class (2.68), the functions $\phi, \tilde{\phi}, \psi$, and $\tilde{\psi}$ defined as in (2.62) and (2.78) satisfy*

$$< \psi_{j,k}, \tilde{\psi}_{l,m} >= \delta_{j,l}\delta_{k,m}, \quad j,k,l,m \in \mathbf{Z}; \tag{2.85}$$

and

$$\begin{cases} < \phi_{j,k}, \tilde{\psi}_{j,l} >= 0; \\ < \tilde{\phi}_{j,k}, \psi_{j,l} >= 0, \quad j,k,l \in \mathbf{Z} \end{cases} \tag{2.86}$$

That is, $V_j \perp \tilde{W}_j$ and $\tilde{V}_j \perp W_j$ for all $j \in \mathbf{Z}$.

Proof: Let us first consider the case $j = l$ in (2.85). In this case, by the third identity in (2.66) and (2.75) we have, using the notation $z = e^{-i\omega/2}$,

$$\begin{aligned}
< \psi_{j,k}, \tilde{\psi}_{j,m} > &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}(\omega)\overline{\hat{\tilde{\psi}}(\omega)}e^{-i(k-m)\omega}d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} Q(z)\overline{H^*(z)}\hat{\phi}\left(\frac{\omega}{2}\right)\overline{\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right)}e^{-i(k-m)\omega}d\omega \\
&= \frac{1}{2\pi}\sum_l \int_0^{2\pi}[Q(e^{-i(\omega/2+\pi l)})H(e^{i(\omega/2+\pi l)}) \\
&\times \hat{\phi}\left(\frac{\omega}{2}+\pi l\right)\overline{\hat{\tilde{\phi}}}\left(\frac{\omega}{2}+\pi l\right)]e^{-i(k-m)\omega}d\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi}\sum_l[Q(z)H(z)\hat{\phi}\left(\frac{\omega}{2}+2\pi l\right)\overline{\hat{\tilde{\phi}}}\left(\frac{\omega}{2}+2\pi l\right) \\
&+ Q(-z)H(-z)\hat{\phi}\left(\frac{\omega}{2}+\pi+2\pi l\right)\overline{\hat{\tilde{\phi}}}\left(\frac{\omega}{2}+\pi+2\pi l\right)]e^{-i(k-m)\omega}d\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi}[Q(z)H(z)+Q(-z)H(-z)]e^{-i(k-m)\omega}d\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{-i(k-m)\omega}d\omega = \delta_{k,m}
\end{aligned}$$

Proceeding to the general case, we observe, by applying the first two identities in (2.66) instead, that the same derivation given above also yields (2.86), so that

$$V_j \perp \tilde{W}_j \quad \tilde{V}_j \perp W_j, \quad j \in \mathbf{Z} \tag{2.87}$$

Hence, if $j < l$, then

$$\psi_{j,k} \in W_j \subset V_{j+1} \subset V_l$$

and by the first assertion in (2.87), we have

$$< \psi_{j,k}, \tilde{\psi}_{l,m} >= 0 \quad k, m \in \mathbf{Z}$$

For $j > l$, the same conclusion can be drawn by applying the second assertion in (2.87). $\diamond$.

As a consequence of the biorthogonality property in (2.85), both families $\{\psi_{j,k}\}$ and $\{\tilde{\psi}_{j,k}\}$ are l^2-linearly independent. Therefore, since

$$\begin{aligned} L^2(\mathbf{R}) &= \cdots \dot{+} W_{-1} \dot{+} W_0 \dot{+} W_1 \dot{+} \cdots \\ &= \cdots \dot{+} \tilde{W}_{-1} \dot{+} \tilde{W}_0 \dot{+} \tilde{W}_1 \dot{+} \cdots \end{aligned}$$

both $\psi_{j,k}$ and $\tilde{\psi}_{j,k}$ are basis of $L^2(\mathbf{R})$. In fact, under the hypotheses of Theorem 2.9, it follows that both $\psi_{j,k}$ and $\tilde{\psi}_{j,k}$ are frames of $L^2(\mathbf{R})$. We may now conclude that $\psi_{j,k}$ and $\tilde{\psi}_{j,k}$ are actually Riesz basis of $L^2(\mathbf{R})$.

Theorem 2.10 *Under the hypotheses of Theorem 2.9, the two functions $\psi \in W_0$ and $\tilde{\psi} \in \tilde{W}_0$ are wavelets which are dual to each other.*

Consequently, every function $f \in L^2(\mathbf{R})$ has two (unique) wavelet series representation:

$$\begin{cases} f(x) = \sum_{j,k} < f, \tilde{\psi}_{j,k} > \psi_{j,k}(x); \\ f(x) = \sum_{j,k} < f, \psi_{j,k} > \tilde{\psi}_{j,k}(x), \end{cases} \tag{2.88}$$

where the coefficients are the values of the IWT of f, relative to the basic wavelets $\tilde{\psi}$ and ψ respectively, evaluated at the time-scale positions

$$(b, a) = \left(\frac{k}{2^j}, \frac{1}{2^j} \right)$$

It is very important to derive efficient algorithms for finding these IWT values from f and for reconstructing f from these IWT values. It turns out that the two-scale sequences $\{\bar{g}_{-n}\}$ and $\{\bar{h}_{-n}\}$-whose two-scale symbols are G^* and H^*-can be used for obtaining the IWT values $< f, \tilde{\psi}_{j,k} >$. This computational scheme, called decomposition algorithm is a consequence of the decomposition relation (2.45).

On the other hand, the two-scale sequences $\{p_n\}$ and $\{q_n\}$ can be used for reconstructing f from its IWT values. This computational scheme, called the reconstruction algorithm is a consequence of the two-scale relations (2.32) and (2.34).

If we wish to use ψ, instead of $\tilde{\psi}$, as the basic wavelet, then the two-scale sequences $\{p_n\}$ and $\{q_n\}$ are used in the decomposition algorithm, while the two-scale sequences $\{\bar{g}_{-n}\}$ and $\{\bar{h}_{-n}\}$ are used in the reconstruction algorithm.

In other words, the roles of the pairs

$$(\{\bar{g}_{-n}\}, \{\bar{h}_{-n}\}) \quad (\{p_n\}, \{q_n\})$$

for decomposition and reconstruction purposes are interchanged, if the IWT information

$$\left\{(W_{\tilde{\psi}} f)\left(\frac{k}{2^j}, \frac{1}{2^j}\right) : j, k \in \mathbf{Z}\right\}$$

is replaced by the IWT information

$$\left\{(W_{\psi} f)\left(\frac{k}{2^j}, \frac{1}{2^j}\right) : j, k \in \mathbf{Z}\right\}$$

This is called the duality principle in wavelet decomposition-reconstruction.

For any $f \in L^2(\mathbf{R})$, let f_N be some approximant of f from V_N for a fixed $N \in \mathbf{Z}$. Note that this approximation does not have to be the $L^2(\mathbf{R})$ orthogonal projection. We may consider V_N as the sample space and f_N the date of f on V_N. Since

$$\begin{aligned} V_N &= W_{N-1} \dot{+} V_{N-1} \\ &= \cdots = W_{N-1} \dot{+} \cdots \dot{+} W_{N-M} \dot{+} V_{N-M} \end{aligned}$$

for any positive integer M, f_N has a unique decomposition:

$$f_N(x) = g_{N-1}(x) + g_{N-2}(x) + \cdots + g_{N-M}(x) + f_{N-M}(x), \qquad (2.89)$$

where

$$\begin{cases} g_j(x) \in W_j, \quad j = N-M, \ldots, N-1; \\ f_{N-M}(x) \in V_{N-M} \end{cases} \qquad (2.90)$$

Let us write

$$\begin{cases} f_j(x) = \sum_k c_k^j \phi(2^j x - k) \in V_j, \\ c^j := \{c_k^j\}, k \in \mathbf{Z} \end{cases} \qquad (2.91)$$

and

$$\begin{cases} g_j(x) = \sum_k d_k^j \psi(2^j x - k), \\ d^j := \{d_k^j\}, k \in \mathbf{Z} \end{cases} \qquad (2.92)$$

then the decomposition in (2.89) is uniquely determined by the sequences c^j and d^j in (2.91) and (2.92). It is important to note that

$$d_k^j = (W_{\tilde{\psi}} f_N)\left(\frac{k}{2^j}, \frac{1}{2^j}\right), j, k \in \mathbf{Z} \tag{2.93}$$

are the values of the IWT of f_N, using $\tilde{\psi}$ as the basic wavelet.

In the wavelet decomposition and reconstruction schemes to be discussed below, we will use the *digital representations* c^j, d^j of $f_j(x)$ and $g_j(x)$. To facilitate our discussion, we introduce the notation

$$\begin{cases} a_n := \frac{1}{2} g_{-n}; \\ b_n := \frac{1}{2} h_{-n}, \end{cases} \tag{2.94}$$

where $\{\bar{g}_{-n}\}$ and $\{\bar{h}_{-n}\}$ are two-scale sequences corresponding to the two-scale symbols G^* and H^*. Hence, the decomposition relation (2.45) becomes

$$\phi(2x - l) = \sum_{k=-\infty}^{\infty} \{a_{l-2k}\phi(x-k) + b_{l-2k}\psi(x-k)\}, l \in \mathbf{Z} \tag{2.95}$$

Decomposition algorithm

$$\begin{cases} c_k^{j-1} = \sum_l a_{l-2k} c_l^j; \\ d_k^{j-1} = \sum_l b_{l-2k} c_l^j \end{cases} \tag{2.96}$$

Proof: By applying the decomposition relation (2.93), we have

$$\begin{aligned} f_j(x) &= \sum_l c_l^j \phi(2^j x - l) \\ &= \sum_l c_j^l \left[\sum_k \{a_{l-2k}\phi(2^{j-1}x - k) + b_{l-2k}\psi(2^{j-1}x - k)\}\right] \\ &= \sum_k \left\{\sum_l a_{l-2k} c_l^j\right\} \phi(2^{j-1}x - k) \\ &+ \sum_k \left\{\sum_l b_{l-2k} c_l^j\right\} \psi(2^{j-1}x - k) \end{aligned}$$

hence, from the decomposition $f_j(x) = f_{j-1}(x)+g_{j-1}(x)$, where $f_{j-1}(x)$ and $g_{j-1}(x)$ are given as in (2.91) and (2.92) with j replaced by $j-1$, it follows that

$$\begin{aligned} &\sum_k \left\{ \sum_l a_{l-2k} c_l^j - c_k^{j-1} \right\} \phi(2^{j-1}x - k) \\ + &\sum_k \left\{ \sum_l b_{l-2k} c_l^j - d_k^{j-1} \right\} \psi(2^{j-1}x - k) = 0 \end{aligned}$$

so that (2.96) follows by invoking the l^2-linear independent of $\{\phi_{j-1,k} : k \in \mathbf{Z}\}$ and $\{\psi_{j-1,k} : k \in \mathbf{Z}\}$ and the fact that $V_{j-1} \cap W_{j-1} = \{0\}$. $\diamond$

Reconstruction algorithm

$$c_k^j = \sum_l [p_{k-2l} c_l^{j-1} + q_{k-2l} d_l^{j-1}] \tag{2.97}$$

Proof: By applying the two-scale relations (2.32) and (2.34), we have

$$\begin{aligned} f_{j-1}(x) + g_{j-1}(x) &= \sum_l [c_l^{j-1}\phi(2^{j-1}x - l) + d_l^{j-1}\psi(2^{j-1}x - l)] \\ &= \sum_l [c_l^{j-1} \sum_k p_k \phi(2^j x - 2l - k) \\ &+ d_l^{j-1} \sum_k q_k \phi(2^j x - 2l - k)] \\ &= \sum_l \sum_k (c_l^{j-1} p_{k-2l} + d_l^{j-1} q_{k-2l}) \phi(2^j x - k) \\ &= \sum_k \left\{ \sum_l [p_{k-2l} c_l^{j-1} + q_{k-2l} d_l^{j-1}] \right\} \phi(2^j x - k) \end{aligned}$$

Since $f_{j-1}(x)+g_{j-1}(x) = f_j(x)$, we obtain (2.97) by referring to the representation formula (2.91) of $f_j(x)$ and l^2-linear independent of $\{\phi_{j,k} : k \in \mathbf{Z}\}$. $\diamond$

2.5 Linear-phase filtering

Scaling functions and wavelets can be considered as filter functions. If the space $L^2(\mathbf{R})$ represents the space of all analog signals with finite energy, and $\{V_j\}$ is an MRA of $L^2(\mathbf{R})$, then sampling an analog signal $f \in L^2(\mathbf{R})$ is accomplished by approximation from some sample space V_N.

Definition 2.7 *Let $f \in L^2(\mathbf{R})$. Then f is said to have* <u>*linear phase*</u> *if its Fourier transform satisfies*

$$\hat{f}(\omega) = \pm|\hat{f}(\omega)|e^{-ia\omega}, \quad a.e., \tag{2.98}$$

where a is some real constant and the $+$ and $-$ sign is independent of ω. Also, f is said to have <u>*generalized linear phase*</u> *if*

$$\hat{f}(\omega) = F(\omega)e^{-i(a\omega+b)}, \quad a.e., \tag{2.99}$$

where $F(\omega)$ is a real-valued function and a, b are real constants. The constant a in both (2.98) and (2.99) is called the <u>*phase*</u> *of $\hat{f}$.*

Example: The Fourier transform of the m-th order cardinal B-spline N_m is given by

$$\hat{N}_m(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^m e^{-im\omega/2},$$

and hence, N_m has linear phase, and the phase of $\hat{N}_m$ is $m/2$.

Definition 2.8 *Let $\{a_n\} \in l^1$ and $A(e^{-i\omega})$ be its discrete Fourier transform (or Fourier series). Then $\{a_n\}$ is said to have* <u>*linear phase*</u> *if*

$$A(e^{-i\omega}) = \pm|A(e^{-i\omega})|e^{-in_0\omega}, \quad \omega \in \mathbf{R}, \tag{2.100}$$

where $n_0 \in \frac{1}{2}\mathbf{Z}$ and the $+$ and $-$ sign is independent of ω. Also, $\{a_n\}$ is said to have <u>*generalized linear phase*</u> *if*

$$A(e^{-i\omega}) = F(\omega)e^{-i(n_0\omega+b)}, \quad \omega \in \mathbf{R}, \tag{2.101}$$

for some real-valued function $F(\omega), n_0 \in \frac{1}{2}\mathbf{Z}$ and $b \in \mathbf{R}$. The value n_0 in both (2.100) and (2.101) is called the <u>*phase*</u> *of the symbols of $\{a_n\}$.*

Let us first give a characterization of both functions and sequences with generalized linear phases.

Lemma 2.2 *(1) A function $f \in L^2(\mathbf{R})$ has generalized linear phase in the sense of (2.99), where $a, b \in \mathbf{R}$, if and only if $e^{ib}f(x)$ is skew-symmetric with respect to a in the sense that*

$$e^{ib}f(a+x) = \overline{e^{ib}f(a-x)}, \quad x \in \mathbf{R} \tag{2.102}$$

(2) A sequence $\{a_n\} \in l^1$ has generalized linear phase in the sense of (2.101), where $n_0 \in \frac{1}{2}\mathbf{Z}$ and $b \in \mathbf{R}$, if and only if $\{e^{ib}a_n\}$ is skew-symmetric with respect to n_0 in the sense that

$$e^{ib}a_n = \overline{e^{ib}a_{2n_0-n}}, \quad n \in \mathbf{Z} \tag{2.103}$$

Proof: (1) Suppose that $f \in L^2(\mathbf{R})$ satisfies (2.99). Then

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{-i(a\omega+b)}e^{ix\omega}d\omega,$$

or equivalently,

$$e^{ib}f(a-x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{-ix\omega}d\omega. \tag{2.104}$$

Since $F(\omega)$ is real, assertion (2.102) follows by equating the complex conjugate of the expression in (2.104) with itself.

Conversely, if (2.102) is satisfied, then taking the Fourier transform of both sides of (2.102) yields

$$\begin{aligned} e^{ib}\hat{f}(\omega)e^{ia\omega} &= e^{-ib}\int_{-\infty}^{\infty} \overline{f(a-x)}e^{-i\omega x}dx \\ &= e^{-ib}\int_{-\infty}^{\infty} \overline{f(a-x)e^{i\omega x}}dx = \overline{e^{ib}\hat{f}(\omega)e^{ia\omega}} \end{aligned}$$

Hence this quantity is real, and (2.99) follows by setting this real-valued function to be $F(\omega)$.

(2) Suppose that $\{a_n\} \in l^1$ satisfies (2.101). Then we have

$$e^{i(n_0\omega+b)}A(e^{-i\omega}) = F(\omega) = \overline{F(\omega)} = e^{-i(n_0\omega+b)}\overline{A(e^{-i\omega})},$$

or equivalently,

$$e^{i2n_0\omega}e^{ib}A(e^{-i\omega}) = \overline{e^{ib}A(e^{-i\omega})}. \tag{2.105}$$

Hence, assertion (2.103) follows by comparing the coefficients of $e^{in\omega}$ in (2.105).

Conversely, if (2.103) holds, then we have (2.105), and consequently

$$e^{i(n_0\omega+b)}A(e^{-i\omega}) = \overline{e^{i(n_0\omega+b)}A(e^{-i\omega})}$$

and we define this real-valued expression to be $F(\omega)$. This yields (2.101). $\diamond$

Remark: The notion of skew-symmetry in (2.102) and (2.103) is not satisfying because of the necessary complex conjugation. When $f(x)$ is real-valued, however, it is clear that for (2.102) to hold, e^{i2b} must also be real, or $b = \frac{1}{2}\pi k$. That is, (2.102) becomes

1. $f(a+x) = f(a-x)$, (symmetry)
2. $f(a+x) = -f(a-x)$, (antisymmetry)

Theorem 2.11 *(1) A real-valued function $f \in L^2(\mathbf{R})$ has generalized linear phase if and only if it is either symmetric or antisymmetric.*

(2) A real-valued sequence $\{a_n\} \in l^1$ has generalized linear phase if and only if it is either symmetric or antisymmetric.

Since the phase property of a two-scale sequence directly influences that of the corresponding scaling function, we give the following characterization of linear-phase sequences.

Lemma 2.3 *A real-valued l^1-sequence $\{a_n\}$, with symbol $A(e^{-i\omega})$, has linear phase if and only if there is some $n_0 \in \frac{1}{2}\mathbf{Z}$, such that $A(e^{-i\omega})e^{in_0\omega}$ is real-valued, even, and has no sign changes.*

Lemma 2.4 *A real-valued finite sequence $\{a_n\}$ with support $[0, N]$ has linear phase if and only if the following statements hold:*

1. $a_{N-n} = a_n, \quad n \in \mathbf{Z}$; *and*
2. *the symbol*

$$A(z) = \sum_{n=0}^{N} a_n z^n$$

has only zeros of even order on the unit circle.

Proof: By Lemma 2.3, the real-valued finite sequence $\{a_n\}, n = 0, \ldots, N$ has linear phase if and only if there exists some $n_0 \in \frac{1}{2}\mathbf{Z}$, such that the function

$$F(\omega) := A(e^{-i\omega})e^{in_0\omega}$$

is real-valued, even. and has no sign changes. On the other hand, it is clear that $F(\omega) = F(-\omega)$ is equivalent to

$$\begin{aligned}\sum_{n=0}^{N} a_n e^{in\omega} &= A(e^{i\omega}) = e^{i2n_0\omega}A(e^{-i\omega}) \\ &= \sum_{n=0}^{N} a_n e^{i(2n_0-n)\omega} = \sum_{n=2n_0-N}^{2n_0} a_{2n_0-n}e^{in\omega},\end{aligned}$$

which , in turn, is also equivalent to $n_0 = \frac{1}{2}N$ and $a_{N-n} = a_n$ for all $n \in \mathbf{Z}$. Of course, the real-valued function $F(\omega)$ has no sign changes if and only if its real zeros (if any) of even order, and this, in turn, is equivalent to the statement that $A(z)$ has only zeros of even order on the unit circle $\diamond$.

Theorem 2.12 *Let ϕ be a scaling function with two-scale sequence $\{p_n\} \in l^1$. Also, let P denote the two-scale symbol of ϕ. then*

1. *ϕ has generalized linear phase if and only if*

$$\overline{P(z)} = z^{2n_0}P(z), \quad |z| = 1, \tag{2.106}$$

   for some $n_0 \in \frac{1}{2}\mathbf{Z}$, and

2. *ϕ has linear phase if and only if*

$$P(e^{-i\omega}) = |P(e^{-i\omega})|e^{-in_0\omega}, \tag{2.107}$$

   where $n_0 \in \frac{1}{2}\mathbf{Z}$.

Proof: If ϕ has generalized linear phase, then by Definition 2.7, we have

$$\hat{\phi}(\omega) = F(\omega)e^{-i(a\omega+b)}, a.e.$$

for some real-valued function $F(\omega)$ and some $a, b \in \mathbf{R}$. Hence, $\overline{\hat{\phi}(\omega)} = F(\omega)e^{i(a\omega+b)}$ and therefore

$$\begin{aligned} \overline{P(e^{-i\omega/2})} &= \frac{\overline{\hat{\phi}(\omega)}}{\overline{\hat{\phi}(\omega/2)}} = e^{ia\omega/2}\frac{F(\omega)}{F(\omega/2)} \\ &= e^{ia\omega}\frac{\hat{\phi}(\omega)}{\hat{\phi}(\omega/2)} = e^{ia\omega}P(e^{-i\omega/2}), \end{aligned}$$

for almost all $\omega \in \mathbf{R}$. This implies that

$$n_0 := a \in \frac{1}{2}\mathbf{Z},$$

and (2.106) holds. If ϕ has linear phase, then by Definition 2.7, we have $b = 0$ and $F(\omega)$ has no sign changes. Consequently,

$$\begin{aligned} P(e^{-i\omega/2}) &= \frac{\hat{\phi}(\omega)}{\hat{\phi}(\omega/2)} = e^{-ia\omega/2}\frac{F(\omega)}{F(\omega/2)} \\ &= e^{-ia\omega}\left|\frac{F(\omega)}{F(\omega/2)}\right| = e^{-in_0\omega/2}\left|P(e^{-i\omega/2})\right|, \end{aligned}$$

which agrees with (2.107).

Conversely, if (2.106) holds, then we have

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} P(e^{-i\omega/2^k}) = \left\{\prod_{k=1}^{\infty} \overline{P(e^{-i\omega/2^k})}\right\} e^{-i2n_0\omega} \tag{2.108}$$

$$= \overline{\hat{\phi}(\omega)}e^{-i2n_0\omega} \tag{2.109}$$

Hence, the function

$$F(\omega) := e^{in_0\omega}\hat{\phi}(\omega)$$

is real-valued, and since

$$\hat{\phi}(\omega) = F(\omega)e^{-in_0\omega},$$

ϕ has generalized linear phase. If the hypothesis (2.107) is assumed, then we have

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} P(e^{-i\omega/2^k}) = \left|\prod_{k=1}^{\infty} P(e^{-i\omega/2^k})\right| e^{-in_0\omega} \tag{2.110}$$

$$= |\hat{\phi}(\omega)|e^{-in_0\omega} \tag{2.111}$$

so that ϕ has linear phase. $\diamond$

Remark: For a scaling function ϕ to have generalized linear phase, it is necessary and sufficient that ϕ is skew-symmetric with respect to some $n_0 = \frac{1}{2}\mathbf{Z}$, in the sense that

$$\phi(n_0 + x) = \overline{\phi(n_0 - x)}, a.e. \tag{2.112}$$

Indeed, for ϕ to have generalized linear phase, we have (2.106) and consequently (2.108), so that $\hat{\phi}(\omega) = F(\omega)e^{-in_0\omega}$. Hence, (2.112) follows from Lemma 2.2. The converse is trivial.

If the two-scale sequence $\{p_k\}$ is real-valued and finite, then by applying Theorem 2.12 and Lemma 2.4, we can say a little more, as in the following.

Theorem 2.13 *Let ϕ be a real-valued scaling function whose two-scale sequence $\{p_n\}$ is a finite real sequence with support $[0, N]$. Then*

1. *ϕ has generalized linear phase if and only if $p_{N-n} = p_n$ for all $n \in \mathbf{Z}$;*

2. *ϕ has linear phase if and only if $p_{N-n} = p_n$ for all n and all zeros of two-scale symbol P that lie on the circle, if any, have even multiplicities.*

To investigate the phase properties of a wavelet, one has to have some knowledge of its two-scale symbol Q. For instance, if the scaling function ϕ has generalized linear phase, then by the two-scale relation $\hat{\psi}(\omega) = Q(e^{-i\omega/2})\hat{\phi}(\omega)$ and Definition 2.7 and 2.8, it follows that ψ also has generalized linear phase provided the sequence $\{q_n\}$ has generalized linear phase, and an analogous conclusion can be made concerning the property of linear phase.

Example: Consider the first order cardinal B-spline N_1, and its corresponding Haar wavelet $\psi_1(x) := N_1(2x) - N_1(2x-1)$. We see that N_1 has linear phase. Since the two-scale symbol Q for ψ_1 is

$$Q(z) = \frac{1}{2}(1 - z) = \left(\sin\frac{\omega}{4}\right)e^{-i(1/4\omega - \pi/2)}, \tag{2.113}$$

where $z = e^{-i\omega/2}$, we also see that

$$\hat{\psi}_1(\omega) = Q(z)\hat{N}_1\left(\frac{\omega}{2}\right) = \frac{(\sin\omega/2)^2}{\omega/4}e^{-i(\omega/2-\pi/2)} \tag{2.114}$$

has generalized linear phase, but does not have linear phase. $\diamond$

Theorem 2.14 *Let ϕ be a scaling function governed by a finite two-scale relation*

$$\phi(x) = \sum_{k=0}^{N} p_k \phi(2x-k), \quad p_0, p_N \neq 0. \tag{2.115}$$

Suppose that $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is an orthonormal family that constitutes a partition of unity, and ϕ is skew-symmetric in the sense that

$$\phi(a+x) = \overline{\phi(a-x)}, \quad x \in \mathbf{R}, \tag{2.116}$$

for some $a \in \mathbf{R}$. Then ϕ must be the first order cardinal B-spline.

Proof: Let

$$P(z) = \frac{1}{2}\sum_{k=0}^{N} p_k z^k$$

and $z = e^{-i\omega/2}$. Then (2.115) is equivalent to

$$\hat{\phi}(\omega) = P(z)\hat{\phi}\left(\frac{\omega}{2}\right), \quad \omega \in \mathbf{R}. \tag{2.117}$$

On the other hand, assertion (2.116) is equivalent to $\hat{\phi}(\omega)e^{ia\omega} = \overline{\hat{\phi}(\omega)e^{ia\omega}}, \omega \in \mathbf{R}$. Hence, from (2.117), we have

$$P(z) = \frac{\hat{\phi}(\omega)}{\hat{\phi}(\frac{\omega}{2})} = e^{-ia\omega}\frac{\overline{\hat{\phi}(\omega)}}{\overline{\hat{\phi}(\frac{\omega}{2})}} = z^{2a}\overline{P(z)}$$

for $|z| = 1$. Since $P(z)$ is a polynomial of degree N with nonzero leading coefficient and nonzero constant term, it follows that $2a = N$, so that

$$\overline{P(z)} = z^{-N}P(z), \quad |z| = 1. \tag{2.118}$$

Let us now consider the hypothesis that $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is an orthonormal family. This is equivalent to

$$\sum_{k=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi k)|^2 = 1, \tag{2.119}$$

so that an application of (2.117) yields

$$|P(z)|^2 + |P(-z)|^2 = 1, \quad |z| = 1. \tag{2.120}$$

So, by substituting (2.118) into (2.120), we have

$$(P(z))^2 + (-1)^N (P(-z))^2 = z^N, \quad |z| = 1. \tag{2.121}$$

Hence, from (2.121) and the hypothesis $p_N \neq 0$, we see that N must be an odd integer. Now by setting

$$\begin{cases} P_e(z) = \frac{1}{2}\sum_k p_{2k} z^k; \\ P_o(z) = \frac{1}{2}\sum_k p_{2k+1} z^k, \end{cases} \tag{2.122}$$

we can write

$$P(z) = P_e(z^2) + zP_o(z^2), \tag{2.123}$$

which the identity (2.120) yields

$$|P_e(z)|^2 + |P_o(z)|^2 = \frac{1}{2}, \quad |z| = 1. \tag{2.124}$$

Hence, by applying (2.118) and (2.123), we obtain

$$\begin{aligned} P_e(z^2) + zP_o(z^2) &= P(z) = z^N \overline{P(z)} \\ &= z^N \overline{[P_e(z^2) + zP_o(z^2)]} \\ &= z^N \overline{P_e(z^2)} + z^{N-1} \overline{P_o(z^2)}, \quad |z| = 1 \end{aligned}$$

Since N is odd, equating the odd and even parts gives

$$\begin{cases} P_e(z^2) = z^{N-1} \overline{P_o(z^2)} : \\ P_o(z^2) = z^{N-1} \overline{P_e(z^2)}, \quad |z| = 1 \end{cases}$$

so that

$$|P_e(z)|^2 = |P_o(z)|^2, \quad |z| = 1.$$

Applying this identity to (2.124) gives rise to

$$|P_e(z)|^2 = |P_o(z)|^2 = \frac{1}{4}, \quad |z| = 1 \tag{2.125}$$

This is not possible unless P_e and P_o are monomials, or

$$P(z) = \frac{1}{2}(p_0 + p_N), \quad N; odd$$

Since $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ is a partition of unity, we have $P(1) = 1$ and $P(-1) = 0$ so that

$$P(z) = \frac{1 + z^N}{2}$$

and hence,

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} \left(\frac{1 + e^{-i\omega N/2^k}}{2} \right) = \frac{1 - e^{-i\omega N}}{i\omega N}$$

Consequently, we obtain

$$\sum_k |\hat{\phi}(\omega + 2\pi k)|^2 = \frac{\sin^2(\omega N/2)}{N^2 \sin^2(\omega/2)} \tag{2.126}$$

It now follows from (2.119) that $N = 1$. That is, $P(z) = (1 + z)/2$, or ϕ is the first order cardinal B-spline N_1. $\diamond$.

2.6 Compactly supported wavelets

The objective of this section is to investigate the structure of wavelets with compact supports. We are particularly interested in skew-symmetric wavelets. We will consider a pair of admissible two-scale symbols P and G^* which are dual to each other in the sense that

$$P(z)G(z) + P(-z)G(-z) = 1, \quad |z| = 1 \tag{2.127}$$

where $G^*(z) = \overline{G(z)}, |z| = 1$. Then the wavelet ψ and its dual $\tilde{\psi}$ have two-scale symbols in the sense that

$$\begin{cases} \hat{\psi}(\omega) = Q(e^{-i\omega/2})\hat{\phi}\left(\frac{\omega}{2}\right); \\ \hat{\tilde{\psi}}(\omega) = H^*(e^{-i\omega/2})\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right), \end{cases} \tag{2.128}$$

where $H^*(z) = \overline{H(z)}, |z| = 1$, and Q and H are arbitrarily, but necessarily, selected from the class in (2.68), namely:

$$\begin{cases} Q(z) = z^{-1}G(-z)K(z^2); \\ H(z) = zP(-z)K^{-1}(z^2), \quad |z| = 1 \end{cases} \tag{2.129}$$

where K is in Wiener's class W with $K(z) \neq 0$ for $|z| = 1$. Also recall that the space $\{V_j\}, \{W_j\}, \{\tilde{V_j}\}$ and $\{\tilde{W_j}\}$ generated by $\phi, \psi, \tilde{\phi}$, and $\tilde{\psi}$, respectively, satisfy

$$\begin{cases} \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \\ V_{j+1} = V_j \dot{+} W_j, \quad j \in \mathbf{Z} \end{cases} \tag{2.130}$$

$$\begin{cases} \cdots \subset \tilde{V_{-1}} \subset \tilde{V_0} \subset \tilde{V_1} \subset \cdots \\ \tilde{V_{j+1}} = \tilde{V_j} \dot{+} \tilde{W}_j, \quad j \in \mathbf{Z} \end{cases} \tag{2.131}$$

and

$$\begin{cases} V_j \perp \tilde{W}_j, \quad j \in \mathbf{Z}; \\ \tilde{V_j} \perp W_j, \quad j \in \mathbf{Z}. \end{cases} \tag{2.132}$$

In addition, the pairs $(\phi, \tilde{\phi})$ and $(\psi, \tilde{\psi})$ are dual pairs, in the sense that

$$\begin{cases} < \phi(\cdot - k), \tilde{\phi}(\cdot - m) >= \delta_{k,m}, \quad k, m \in \mathbf{Z}; \\ < \psi_{j,k}, \tilde{\psi_{l,m}} >= \delta_{j,l}\delta_{k,m}, \quad j, k, l, m \in \mathbf{Z}. \end{cases} \tag{2.133}$$

Let us first study the structure of semi-orthogonal (s.o.) wavelets, and particularly orthogonal (o.n.) wavelets. It is clear that the spaces W_j and $\tilde{W}_j, j \in \mathbf{Z}$, generated by any s.o. wavelet ψ and its dual $\tilde{\psi}$ are identical, namely: $W_j = \tilde{W}_j$ for all $j \in \mathbf{Z}$. Hence, from (2.4), we have $V_j = \tilde{V_j}$ for all $j \in \mathbf{Z}$, so that the scaling function ϕ and its dual $\tilde{\phi}$ generate the same MRA. In fact, we observe the (unique) dual $\tilde{\phi}$ of ϕ is given by

$$\hat{\tilde{\phi}}(\omega) = \frac{\hat{\phi}(\omega)}{\sum_{k=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi k)|^2} \tag{2.134}$$

Now let us restrict our attention to scaling function ϕ with finite two-scale sequences $\{p_n\}$, namely:

$$\phi(x) = \sum_{n=1}^{N} p_n \phi(2x - n), \quad p_0, p_N \neq 0,$$

Recall from (2.26),(2.27) and Theorem 2.2, that the generalized Euler-Frobenius Laurent polynomial

$$\begin{aligned} E(z) &= \sum_{k=-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \phi(k+y)\overline{\phi(y)}dy \right\} z^k \\ &= \sum_{k=-\infty}^{\infty} \left| \hat{\phi}(\omega/2 + 2\pi k) \right|^2 \end{aligned}$$

relative to ϕ, where $z = e^{-i\omega/2}$, is zero-free and pole-free on $|z| = 1$. Hence, it follows from (2.134) that

$$\hat{\tilde{\phi}}(\omega) = \frac{1}{E(z^2)} \hat{\phi}(\omega), \tag{2.135}$$

and consequently the dual $\tilde{\phi}$ of the compactly supported scaling function ϕ does not have compact support unless $E(z)$ is a positive constant, although it is of exponential decay. First, we must find the two-scale symbol G^* of $\tilde{\phi}$. This is easily done by applying (2.135) and the two-scale relation of ϕ as follows:

$$\begin{aligned} \hat{\tilde{\phi}}(\omega) &= \frac{1}{E(z^2)} \hat{\phi}(\omega) = \frac{1}{E(z^2)} P(z) \hat{\phi}\left(\frac{\omega}{2}\right) \\ &= \frac{E(z)}{E(z^2)} P(z) \hat{\tilde{\phi}}\left(\frac{\omega}{2}\right), \end{aligned}$$

so that

$$G^*(z) = \frac{E(z)}{E(z^2)} P(z), \quad z = e^{-i\omega/2}. \tag{2.136}$$

It is easy to verify that for this G^*, the dual relation (2.127) is equivalent to the identity Theorem 2.2 for generalized Euler-Frobeniuos Laurent polynomials. Now by Theorem 2.7 and (2.136), the two-scale symbol Q for any wavelet ψ relative to the scaling function ϕ is given by

$$\begin{aligned} Q(z) &= z^{-1} G(-z) K(z^2) \\ &= z^{-1} \frac{E(-z)\overline{P(-z)}}{E(z^2)} K(z^2) \end{aligned}$$

where $K \in \mathbf{W}$ with $K(z) \neq 0$ for $|z| = 1$. Thus, we have some freedom in the choice of ψ. In particular, the wavelet ψ with minimum support is obtained by selecting the admissible $K \in \mathbf{W}$ such that $Q(z)$ is a polynomial with lowest degree.

Since E is a Laurent polynomial which is zero-free and pole-free on $|z| = 1$, we may choose $K(z) = -zE(z)$, so that the two-scale symbols Q and H^* for a compactly supported s.o. wavelet ψ and its dual $\tilde{\psi}$, respectively, are given by

$$\begin{cases} Q(z) = -zE(-z)\overline{P(-z)}; \\ H(z) = -z^{-1}\frac{P(-z)}{E(z^2)} \end{cases} \tag{2.137}$$

The reason for this choice of K, instead of simply $K = E$, is that this normalization is consistent with the formulation of the Haar function. Observe that if E is not a constant, then the dual wavelet $\tilde{\psi}$ does not have compact support, although it decays exponentially. One advantage of the choice of Q in (2.137) is that it is very easy to determine whether or not the wavelet ψ has generalized linear phase. Indeed, since the finite coefficient sequence of the Laurent polynomial $E(z)$ is skew-symmetric, it is clear that the coefficient sequence of the polynomial $Q(z)$in (2.137) is also skew-symmetric provided that the two-scale sequence $\{p_n\}$ has this property.

Theorem 2.15 *Let $\{p_n\}$ be a finite,symmetric,real-valued two-scale sequence of a scaling function of ϕ. Also, let $\psi, \tilde{\phi}$, and $\tilde{\psi}$ be the s.o. wavelet, dual scaling function and dual wavelet with two-scale symbols Q, G^* and H^*, respectively, as given by (2.137) and (2.136). Then ϕ and its corresponding s.o. wavelet ψ have compact supports, $\tilde{\phi}$ and $\tilde{\psi}$ are of exponential decay, and all of $\phi, \psi, \tilde{\phi}, \tilde{\psi}$ have generalized linear phases.*

Next, let us consider orthogonal (o.n.) wavelet ψ relative to compactly supported scaling function ϕ. The general approach is to construct ϕ which is orthonormal (o.n.), in the sense that

$$\{\phi(\cdot - k) : k \in \mathbf{z}\}$$

is an o.n. family. For such a ϕ, it follows that the generalized Euler-Frobenius polynomial $E(z)$ is constant 1. Hence, by (2.137), we have

$$Q(z) = -z\overline{P(-z)}, \quad |z| = 1$$

Consequently, if the two-scale relation of ϕ is given by

$$\phi(x) = \sum_{n=0}^{N} p_n \phi(2x - n), \quad p_0, p_N \neq 0,$$

then the two-scale relation of the o.n. wavelet ψ is given by

$$\psi(x) = \sum_{n=-N+1}^{1} (-1)^n \bar{p}_{1-n} \phi(2x - n). \tag{2.138}$$

Observe that since the coefficient sequences $\{p_n\}$ and $\{(-1)^n \bar{p}_{1-n}\}$ for ϕ and ψ have similar phase properties, one expects, in view of Theorem 2.14, that the compactly supported o.n. wavelets also fail to have generalized phases. Let us reformulate Theorem 2.14 as the following.

Theorem 2.16 *Let ψ be a compactly supported o.n. wavelet as given by (2.138) whose corresponding o.n. scaling function ϕ generates a partition of unity. Suppose that ϕ is skew-symmetric in the sense of (2.116). Then ψ must be the Haar function ψ_H.*

Hence, for a compactly supported continuous wavelet function ψ to have generalized linear phase, there seem to be only two alternatives. First, we may settle for semi-orthogonality. This certainly works, provided that we are willing to accept duals which are only of exponential decay. In the following, we will totally give up orthogonality and look for compactly supported ψ and $\tilde{\psi}$ with generalized linear phases.

Recall from Theorem 2.12 that ϕ has generalized linear phase if and only if its two-scale symbol P satisfies

$$P(z) = z^m \overline{P(z)}, \quad |z| = 1, \tag{2.139}$$

for some integer m. In order to be able to construct a compactly supported dual $\tilde{\phi}$ of ϕ that also generalized linear phase, we have to come up with an admissible two-scale polynomial symbol of $\tilde{\phi}$ that also satisfies (2.139) for some integer power of z. In this direction, we have the following result.

Theorem 2.17 *Let P and G^* be admissible two-scale Laurent polynomial symbols that are dual to each other, such that P satisfies (2.139). Then*

$$G_1(z) := \frac{1}{2}\{G(z) + z^{-m}G^*(z)\} \tag{2.140}$$

satisfies the same duality relation

$$P(z)G_1(z) + P(-z)G_1(-z) = 1, \quad |z| = 1 \tag{2.141}$$

as $G(z)$, and moreover,

$$G_1(z) = z^{-m}\overline{G_1(z)}, \quad |z| = 1 \tag{2.142}$$

or equivalently,

$$G_1^*(z) = z^m\overline{G_1^*(z)}, \quad |z| = 1 \tag{2.143}$$

Proof: It is clear that G_1 satisfies (2.142). Indeed, for $|z| = 1$, we have

$$\begin{aligned} G_1^*(z) &= \overline{G_1(z)} = \frac{1}{2}\{\overline{G(z)} + z^m\overline{G^*(z)}\} \\ &= \frac{1}{2}\{G^*(z) + z^m G(z)\} = \frac{1}{2}\{G(z) + z^{-m}G^*(z)\}z^m \\ &= z^m G_1(z) = z^m\overline{G_1^*(z)}. \end{aligned}$$

To verify (2.141), we simply apply (2.139) and (2.127), and obtain

$$\begin{aligned} & P(z)G_1(z) + P(-z)G_1(-z) \\ = \; & \frac{1}{2}\{P(z)[G(z) + z^{-m}G^*(z)] + P(-z)[G(-z) + (-z)^{-m}G^*(-z)]\} \\ = \; & \frac{1}{2}\{[P(z)G(z) + P(-z)G(-z)] + [\overline{P(z)}G^*(z) + \overline{P(-z)}G^*(-z)]\} \\ = \; & \frac{1}{2}\{[P(z)G(z) + P(-z)G(-z)] + [\overline{P(z)G(z) + P(-z)G(-z)}]\} \\ = \; & \frac{1}{2}(1+1) = 1, \quad |z| = 1 \end{aligned}$$

This completes the proof of the theorem. ◇

In the following, we will only consider finite two-scale sequences that are real-valued. For such sequences the generalized linear phase property (2.139) becomes

$$P(e^{-i\omega}) = e^{-im\omega}P(e^{i\omega}), \quad \omega \in \mathbf{R} \tag{2.144}$$

Lemma 2.5 *Let P be a Laurent polynomial with real coefficients that satisfy (2.144) for some $m \in \mathbf{Z}$. Then there is another polynomial P_1 with real coefficients such that*

$$P(e^{-i\omega}) = \begin{cases} e^{-im\omega/2} P_1(\cos\omega), & for\ even\ m; \\ e^{-im\omega/2}(\cos\omega/2) P_1(\cos\omega), & for\ odd\ m. \end{cases} \tag{2.145}$$

Proof: From the assumption (2.144), we see that $e^{im\omega/2}P(e^{-i\omega})$ is an even function of ω. Thus, if m is an even integer, then $e^{im\omega/2}P(e^{-i\omega})$ is a 2π-periodic function and is therefore a real polynomial in $\cos\omega$. This gives (2.145) for even m. On the other hand, if m is an odd integer, then by selecting $\omega = \pi$ in (), we have $P(-1) = -P(-1)$, so that $P(-1) = 0$. Therefore, we can write

$$P(z) = \left(\frac{1+z}{2}\right) P_0(z), \tag{2.146}$$

for some polynomial P_0 with real coefficients. Substituting (2.146) into (2.144) yields

$$P_0(e^{-i\omega}) = e^{-i(m-1)\omega} P_0(e^{i\omega})$$

Now, since $m - 1$ is even, we have

$$P_0(e^{-i\omega}) = e^{-i(m-1)\omega/2} P_1(\cos\omega), \tag{2.147}$$

for some polynomial P_1 with real coefficients. Hence, by putting (2.147) into (2.146), we obtain (2.145) for odd m. $\diamond$

In addition to the result in Lemma (), we recall that, as a two-scale symbol, P can be written as

$$P(z) = \left(\frac{1+z}{2}\right)^l P_2(z), \tag{2.148}$$

where P_2 is a Laurent polynomial with real coefficients satisfying

$$P_2(1) = 1 \quad P_2(-1) \neq 0, \tag{2.149}$$

and l is some positive integer.

Lemma 2.6 *Let P be a Laurent polynomial with real coefficients that satisfies (2.144) and (2.148)-(2.149). Then $(m-l)$ must be an even integer, and*

$$P(e^{-i\omega}) = e^{-im\omega/2}(\cos\omega/2)^l S(\cos\omega), \tag{2.150}$$

where S is polynomial with real coefficients that satisfies

$$S(1) = 1 \quad S(-1) \neq 0 \tag{2.151}$$

Proof: From (2.148), we have

$$P(e^{-i\omega}) = e^{-il\omega/2}(\cos\omega/2)^l P_2(e^{-i\omega}) \tag{2.152}$$

Hence, by (2.144), we obtain

$$P_2(e^{-i\omega}) = e^{-i(m-l)\omega} P_2(e^{i\omega}) \tag{2.153}$$

We will first show that $(m-l)$ is an even integer. Suppose, on the contrary, that $(m-l)$ is odd. Then as before, we see, by applying (2.153) with $\omega = \pi$, that $P_2(-1) = 0$. This is a contradiction to (2.149). Now, since $(m-l)$ is even, we may apply Lemma 2.5 to write

$$P_2(e^{-i\omega}) = e^{-i(m-l)\omega/2} S(\cos\omega) \tag{2.154}$$

for some polynomial S with real coefficients. Hence, assertion (2.150) is established by substituting (2.154) into (2.152). In addition, by (2.149), it is clear that by polynomial S in (2.154) satisfies (2.151). $\diamond$.

Chapter 3

Quadrature Formulae

Almost every application using wavelets at some point involves the calculation of the multiresolution coefficients of a function f. Remember that these coefficients are defined as inner products with dual functions. However, we can hardly calculate these exactly. Therefore, we need to construct schemes to approximate them numerically. A classic method is so-called quadrature formula, which approximate an integral of a function, possibly multiplied with a weight function, with a linear combination of evaluation of that function at particular points called abscissae.

We discuss the construction and use of quadrature formulae in connection with multiresolution analysis. Since the construction involves the moments of the scaling function we also derive a recursion formula to calculate them. We show that the construction using monomial moments is ill-conditioned and build a well-conditioned construction using Chebyshev polynomials.

3.1 General Idea

Consider a MRA where the finest level is n and assume that the coarsest level is 0. This implies that we need to calculate the coefficients $\lambda_{j,l}$ and $\gamma_{j,l}$ for $0 \leq j < n$. The coefficients $\lambda_{j,l}$ and $\gamma_{j,l}$ with $j < n$ can be calculated from the $\lambda_{n,l}$ using the fast wavelet transform. We therefore use the quadrature formula on the finest level to approximate the $\lambda_{n,l}$.

Remember that these coefficients are defined as

$$\lambda_{n,l} = \sqrt{2^n} \int_{-\infty}^{\infty} f(x) \overline{\tilde{\phi}(2^n x - l)} dx$$

We only considered real-valued functions, so the complex conjugation is superfluous. Because of the translation and dilation properties we can focus on the case $n = l = 0$,

$$\lambda_{0,0} = \int_{-\infty}^{\infty} f(x) \tilde{\phi}(x) dx.$$

Here we consider the dual scaling function as a weight function. For notational simplicity, we from now on omit the tilde in the notation of the dual scaling function. We keep in mind that the coefficients are given by inner products with the dual functions and remember that, basis functions and dual functions are always interchangeable.

The idea of a quadrature formula now is to find weights w_k and abscissae x_k so that

$$\int_{-\infty}^{\infty} f(x)\phi(x)dx \approx Q[f(x)] = \sum_{k=0}^{r} w_k f(x_k). \tag{3.1}$$

Once the weights and abscissae are known, the coefficients on the finest level can be approximated by

$$\lambda_{n,l} \approx \frac{1}{\sqrt{2^n}} \sum_{k=0}^{r} w_k f\left(\frac{x_k + l}{2^n}\right). \tag{3.2}$$

Evidently, the quadrature formula can also be used to calculate the coefficients $\lambda_{j,l}$ with $j < n$. Why it is better to use the quadrature formula only on the finest level and the fast wavelet transform on the coarsest level will become apparent later.

We need to address two important issues: how to choose the abscissae and how to find the weights. The weights are determined by the fact that the quadrature formula is, in some sense, a good approximation for the integral.

The abscissae have to be chosen equidistant for two reasons. First of all, in many application such as signal processing, image processing and time series, the function f is only known through its evaluation at equidistant points. The second reason follows from the translation properties of the MRA. One typically needs to calculate the coefficients $\lambda_{n,l}$ for a wide range of the parameter l. If the abscissae are not equidistant, each coefficients needs r function evaluations. In case the abscissae are equidistant, quadrature formulae for neighboring coefficients can share function evaluations. More precisely, if the distance between two abscissae is 2^{-n}, each extra coefficient only needs one extra function evaluation.

The fact that they are equidistant does not pin down their location exactly. We still can allow a shift τ and a different spacing. Therefore, we let the abscissae be of the form $k2^{-s} + \tau$ with $k \in \mathbf{Z}$. The shift τ is an extra unknown and is determined together with the weights by the fact that quadrature formula needs to approximate the integral.

Let us consider the problem of finding the unknowns. A popular technique to solve problems involving functions in numerical analysis is to design an approximate solution scheme that is exact for polynomials. When numerically approximating integrals, this leads to the following definition.

Definition 3.1 *The degree of accuracy of a quadrature formula is q if it yields the exact result for every polynomial of degree less than or equal to q.*

We can write this as

$$Q[x^i] = M_i \quad 0 \le i \le q \tag{3.3}$$

The importance of the degree of accuracy can be understood as follows: if the function f is smooth, it locally resembles a polynomial and quadrature formula gives an accurate result. Note that we do not impose any regularity condition on ϕ. Equation (3.2) now lead to algebraic system in the unknowns w_k and τ. The smaller the number of abscissae r, the more efficient the quadrature formula since the number of function evaluations and algebraic operations for one coefficient is proportional to r. From simply counting the number of unknowns and equations we can hope for $q = r$. However, this is not guaranteed as the algebraic system is nonlinear.

3.2 Special Case

3.2.1 Trapezoidal Rule

A simple quadrature formula is the trapezoidal rule, where

$$Q[f(x)] = \sum_k \phi(k) f(k). \tag{3.4}$$

In general, the use of this is limited because it only has a degree of accuracy equal to one. In connection with MRA, however, the following lemma holds:

Lemma 3.1 *If the scaling function satisfies the Strang-Fix condition,*

$$\hat{\phi}^{(p)}(k2\pi) = 0 \quad 0 \leq p < N \quad k \neq 0$$

the degree of accuracy of the trapezoidal rule (3.4) is equal to $N - 1$.

In other words, if the scaling function and its integer translates can be reproduce polynomials up to degree $N-1$, the trapezoidal rule has a degree of accuracy equal to N. This simple result is remarkable. It was already known that the trapezoidal rule is more accurate than expected in special cases such as periodic integrands. This lemma adds another case, namely when the weight function satisfies the Strang-Fix condition up to some order. Note that no regularity of the weight function ϕ is required. The function only needs to be continuous at the integers. Typical functions that satisfy the Strang-Fix condition are the functions of finite element methods such as B-splines. In case ϕ satisfies a refinement relation, the value $\phi(k)$ can easily be calculated as solutions of an eigenvalue problem.

A disadvantage is that the trapezoidal rule is not very efficient. In case ϕ is not compactly supported, the sum (3.4) has to be broken off, which usually leads to a large number of abscissae. But also when ϕ is compactly supported, the trapezoidal rule is not really efficient: the Daubechies orthogonal scaling functions has a support length of $2N-1$, and hence $r = 2N-2 = 2q$. Remember that we are hoping for $r = q$. Only in the case of cardinal B-sples is the trapezoidal rule useful. The B-spline of order m has support width of m and can reproduce polynomials of degree less than or equal to $m-1$. Consequently, $q = r = m-1$.

3.2.2 One-point Formula

Before we consider the general construction, let us take a look at the case where the number of abscissae is one. Since the integral of the scaling function is one, we can write a one-point formula as $Q[f(x)] = f(x_1)$. This means that we take the function value as an approximation for the coefficient. Evidently, if $x_1 = M_1$, the degree of accuracy is equal to one. In the case of orthogonal wavelets, the following theorem holds:

Theorem 3.1 *If $\phi(x)$ is an orthogonal scaling function with $N > 1$, then*

$$M_2 = M_1^2$$

Proof: Define

$$\kappa_m =< x, \phi(x)\phi(x-m) > .$$

Because of the orthogonality it holds that

$$\kappa_{-m} =< x-m, \phi(x-m)\phi(x) >= \kappa_m$$

Consequently,

$$0 = \sum_m m\kappa_m =< x, \phi(x) \sum_m m\phi(x-m) > .$$

Since $N > 1$, we have that

$$\sum_m m\phi(x-m) = x - M_1.$$

Combining the last two equations yields $M_2 - M_1^2 = 0$. ◇

This implies that the degree of accuracy of a one-point quadrature formula is two, which is one more than expected.

3.2.3 Coiflets

The idea of a one-point quadrature is attractive because of its simplicity. Its degree of accuracy is, however, limited. Therefore, Ingrid Daubechies constructed orthogonal scaling functions with compact support for which the one-point quadrature has a higher degree of accuracy.

These scaling functions again can reproduce polynomials up to degree $N-1$, but moreover have $N-1$ vanishing moments,

$$M_p = 0 \quad 1 \leq p < N. \tag{3.5}$$

The corresponding wavelets were called coiflets after Ronald Coifman, who asked for their construction, because he and his collaborators wanted to use them in numerical analysis applications such as the solution of integral equations.

We see from that

$$\sum_l (x-l)^p \phi(x-l) = M_p \quad 0leqp < N$$

they also satisfy

$$\sum_k k^p \phi(k) = \delta_p \quad 0 \leq p < N. \tag{3.6}$$

In this case the one-point quadrature formula with $x_1 = 0$ immediately has a degree of accuracy of $N-1$.

The price to pay for these extra moment conditions is a lager support of the basis functions. The support width of the coiflets and the associated wavelets is $3N-1$, as opposed to $2N-1$ for the original Daubechies orthogonal scaling functions and wavelets. This implies that all fast wavelet transform filters $\{h_k\}$ and $\{g_k\}$ will be approximately 50% longer. This can be very high price, especially in real time applications or situations where one needs to calculate the fast wavelet transform many times. In these cases one might prefer to use a more complicated quadrature formula on the finest level, which has a sufficiently high degree of accuracy but does not imply the use of longer filters.

3.2.4 Practical Aspects

In applications such as signal and image processing, usually the data is given in the form of discrete sample $\{d_l\}$. One then has several possibilities regarding how to relate these samples to a continuous function. What is often done is to take the sample values as coefficients of the basis functions. It was proposed to construct a function $v_n \in V_n$ as

$$v_n(x) = \sqrt{h} \sum_l d_l \phi_{n,l}(x), \quad h = 2^{-n}$$

to start the MRA. We see that the continuous function v_n will in a way "follow" the discrete sample d_l. The quadrature formula function can help us to find a relationship between the function v_n and the discrete sample d_l. Indeed, using the notation with tilde again,

$$\sqrt{h}d_l =< v_n, \tilde{\phi}_{n,l} >$$

and

$$< v_n, \tilde{\phi}_{n,l} >= \sqrt{h}[v_n(h(M_1 + l)) + O(h^t)] \quad h = 2^{-n},$$

so

$$d_l = v_n(h(M_1 + l)) + O(h^t).$$

This means that v_n satisfies a quasi-interpolating property. Here $t = 2$ in general, $t = 3$ for orthogonal wavelets and $t = N$ for coiflets.

Consider the case where the samples d_l can be seen as function evaluations of a smooth function $f, d_l = f(hl)$. Then the following theorem is important:

Theorem 3.2 *If $f \in C^N$ with $f^{(i)}$ bounded for $i \leq N$, then*

$$\sum_l f(hl)\phi(2^n x - l) = \sum_l \phi(l) f(x - hl) + O(h^N) \quad h = 2^{-n}$$

Proof:

$$\begin{aligned}
\sum_l f(hl)\phi(2^j x - l) &= \sum_l \sum_{i=0}^{N-1} \frac{(hl - x)^i}{i!} f^{(i)}(x)\phi(2^j x - l) + O(h^N) \\
&= \sum_{i=0}^{N-1} f^{(i)}(x) \frac{(-h)^i}{i!} \sum_l (2^j x - l)^i \phi(2^j x - l) + O(h^N) \\
&= \sum_{i=0}^{N-1} f^{(i)}(x) \frac{(-h)^i}{i!} \sum_l l^i \phi(l) + O(h^N) \\
&= \sum_l \phi(l) \sum_{i=0}^{N-1} f^{(i)}(x) \frac{(-hl)^i}{i!} + O(h^N) \\
&= \sum_l \phi(l) f(x - hl) + O(h^N) \quad \diamond
\end{aligned}$$

This theorem states that taking function evaluations as coefficients results in approximating a different function $\tilde{f}_n(x) = \sum_l \phi(l) f(x - hl)$

with an error of $O(h^N)$. This function can be seen as a "blurred" version of $f(x)$, as $\{\phi(l)\}$ is a low pass filter. Now, $\tilde{f}_n(x)$ will converge to $f(x)$ for $n \to \infty$ since $\sum_k \phi(k) = 1$. However, in general this convergence is only $O(h)$. In this case of the coiflets we see from (3.6) that $\tilde{f}_n(x) = f(x) + O(h^N)$.

3.3 General Case

3.3.1 Construction Scheme

Since the degree of accuracy of a one-point formula in general is limited, we construct multiple-point quadrature formulae. We assume that ϕ has compact support $[0, L]$ and satisfies a refinement equation with $L+1$ non-zero coefficients h_k. We focus on scaling functions with compact support, since in this case w have the extra limitation that the abscissae should fall inside the integration interval. We construct an r point quadrature formula with $x_k = a_k - \tau, a_k = (k-1)2^s$ and $(r-1)2^s - L \leq \tau \leq 0$. The range of the shift τ is determined by the requirement that no abscissae should fall outside the integration interval. in order to have a non-zero range for the shift τ, the parameters r and s should be chosen so that $(r-1)2^s < L$.

Since there are $(r+1)$ unknowns $\{\tau, w_1, \ldots, w_r\}$, one can try to achieve a degree of accuracy $q = r$. This results in the following system, which is nonlinear in the unknown τ,

$$\sum_{k=1}^{r} w_k [a_k - \tau]^i = M_i \quad 0 \leq i \leq r \tag{3.7}$$

The value of the shift τ can be determined using the product polynomial $Pi(x)$. This polynomial is defined as

$$\Pi(x) = \prod_{k=1}^{r} (x - x_k) = \prod_{k=1}^{r} (x + r - a_k) = \sum_{i=0}^{r} p_i(\tau) x^i$$

where $p_i(\tau)$ is a polynomial of degree $r - i$. Since the degree of accuracy is r, the quadrature formula gives the exact result for the product polynomial $\Pi(x)$ so,

$$\begin{aligned} 0 &= Q_r[\Pi(x)] = \int_0^L \phi(x)\Pi(x)dx = \int_0^L \phi(x)\sum_{i=0}^{r} p_i(\tau)x^i dx \\ &= \sum_{i=0}^{r} p_i(\tau)M_i = \Gamma(\tau) \end{aligned}$$

Here $\Gamma(\tau)$ is a polynomial of degree r in τ. For the quadrature formula to exist, $\Gamma(\tau)$ must have a root in the interval $[(r-1)2^s-L, 0]$. However, the existence of such a toot is not theoretically guaranteed. If there is no root in this interval, an arbitrary value for τ must be choisen and one degree of accuracy is lost. Once τ is determined, the weights are the solution of the linear system formed by r equation (). In order to construct $\Gamma(\tau)$ we write

$$p_i(\tau) = \sum_{j=0}^{r-i} p_{i,j}\tau^j$$

and

$$\Gamma(\tau) = \sum_{i=0}^{r}\sum_{j=0}^{r-i} p_{i,j}\tau^j M_i = \sum_{j=0}^{r}\left(\sum_{i=0}^{r-j} M_i p_{i,j}\right)\tau^j$$

The coefficients $p_{i,j}$ are symmetric, since the product polynomial is symmetric in τ and x, and can be found as $p_{i,j}^{(r)}$ where

$$\Pi^{(m)}(x) = \prod_{k=1}^{m}(x + r - a_k) = \sum_{i=0}^{m}\sum_{j=0}^{m-i} p_{i,j}^{(m)}\tau^j x^i$$

An algorithm to calculate the $p_{i,j}$ can be derived by writing

$$\pi^{(m)}(x) = (x + r - a_m)\Pi^{(m-1)}(x),$$

and identifying the coefficients of the powers of x and r.

A disadvantage of this construction is that the system of equations () is ill-conditioned if r is large. We see that the condition becomes very poor in case r is large and consequently the numerical results for the weights can no longer be trusted. In case $N = 8$, the system is even singular within working precision (16 decimal digits).

3.3.2 Calculation of the Moments

in order to construct the quadrature formula, we first need an algorithm to calculate the moments of a scaling function numerically. Simple algebraic manipulations lead to

$$\begin{aligned} M_p &= \int_{-\infty}^{\infty} x^p \phi(x) dx \\ &= 2 \sum_k h_k \int_{-\infty}^{\infty} x^p \phi(2x-k) dx \\ &= 2^{-p} \sum_k k_k \int_{-\infty}^{\infty} (x+k)^p \phi(x) dx \\ &= 2^{-p} \sum_{i=0}^{p} \left(\begin{array}{c} p \\ i \end{array} \right) \sum_k h_k k^i \int_{-\infty}^{\infty} x^{p-i} \phi(x) dx \\ &= 2^{-p} \sum_{i=0}^{p} \left(\begin{array}{c} p \\ i \end{array} \right) m_i M_{p-i} \end{aligned}$$

where the m_i are the discrete moments of the sequence $\{h_k\}$,

$$m_i = \sum_k k^i h_k$$

Consequently the moments can be calculated with a p-terms recursion relation,

$$M_p = \frac{1}{2^p - 1} \sum_{i=1}^{p} \left(\begin{array}{c} p \\ i \end{array} \right) m_i M_{p-i}$$

3.3.3 Modified Construction

The ill-conditioning problem of the construction using monomial moments can be overcome if we use the basis of Chebyshev polynomials.

The Chebyshev polynomial $T_k(x)$ of degree k is defined by

$$T_k(\cos\theta) = \cos k\theta.$$

The Chebyshev polynomials are orthogonal with respect to q weight function,

$$\frac{1}{\pi} \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_k(x) T_l(x) dx = \begin{cases} 1/2 \; if \; k = l = 0 \\ 1 \; if \; k = l \neq 0 \\ 0 \; otherwise, \end{cases}$$

and have equioscillationg properties in the interval $[-1, 1]$.

Since the interesting properties of these polynomials only hold in the interval $[-1, 1]$, we transform the scaling function $\phi(x)$ to this interval giving a function $\phi^*(x)$ and use the notations y for an independent variable that varies between -1 and 1,

$$2\phi^*(y) = L\phi(x) \quad 2x = L(y+1)$$

The refinement equation becomes

$$\phi^*(y) = 2\sum_k h_k \phi^*(2y - 2k/L + 1).$$

We construct a quadrature formula,

$$\begin{aligned} \nu_{0,0} &= \int_0^L \phi(x) f(x) dx = \int_{-1}^1 \phi^*(y) f\left(\frac{L(y+1)}{2}\right) dy \\ &= \int_{-1}^1 \phi^*(y) f^*(y) dy \approx \sum_{k=1}^r w_k^* f^*(y_k) \\ &= \sum_{k=1}^r w_k f(x_k) = Q_r[f(x)], \end{aligned}$$

with $w_k = W_k^*, y_k = a_k^* - \tau^*, a_k^8 = 2a_k/L - 1$ and $\tau^* = 2\tau/L$. Let M_p^t denote the moments of the transformed scaling function

$$M_p^t = \int_{-1}^1 y^p \phi^*(y) dy,$$

and M_p^* denote the modified moments,

$$M_p^* = \int_{-1}^1 T_p(y) \phi^*(y) dy.$$

The new system can be written as

$$\sum_{k=1}^r w_k T_i(a_k^* - \tau^*) = M_i^* \quad 0 \le i \le r \tag{3.8}$$

The solution procedure is similar to the one in the previous section. We construct a polynomial $\Gamma^*(\tau^*)$, written as a linear combination of

Chebyshev polynomials, and try to find a root in the appropriate interval. In order to construct $\Gamma^*(\tau^*)$ we write

$$\Pi(y) = 2^{-(r-1)} \sum_{i=0}^{r} \sum_{j=0}^{r-i} q_{i,j} T_j(\tau^*) T_i(y)$$

and

$$\Gamma^*(\tau^*) = 2^{-(r-1)} \sum_{j=0}^{r} \sum_{i=0}^{r-j} q_{i,j} M_i^* T_j(\tau^*)$$

Now let

$$\begin{aligned} 2^{(m-1)} \Pi^{(m)}(y) &= 2^{(m-1)} \prod_{k=1}^{m} (y + \tau^* - a_k^*) \\ &= \sum_{i=0}^{m} \sum_{j=0}^{m-i} q_{i,j}^{(m)} T_j(\tau^*) T_i(y) \end{aligned}$$

3.4 Calculation of the modified moments

It is possible to calculate the modified moment as a linear combination of the monomial moments using the confinements $t_i^{(p)}$ of the Chebyshev polynomials,

$$T_p(y) = \sum_{i=0}^{p} t_i^{(p)} y^i \quad M_p^8 = \sum_{i=0}^{p} t_i^{(p)} M_i^t \tag{3.9}$$

However, a considerable loss of significant digits will occur since these coefficients tend to be large and different in sign, The condition would be as bad as in the construction using the monomial moments since calculating the modified moments like this essentially does not change the problem. We therefore need a formula to calculate the modified moments directly. We know that

$$\begin{aligned} M_p^* &= \int_{-1}^{1} T_p(y) \phi^*(y) dy \\ &= 2 \sum_k h_k \int_{k/L-1}^{k/L} T_p(y) \phi^*(2y + 1 - 2k/L) dy \\ &= \sum_k h_k \int_{-1}^{1} T_p \left(\frac{u - 1 + 2k?l}{2} \right) \phi^*(u) du. \end{aligned}$$

In order to find a recursion formula, we write this last shorted and dilated Chebyshev polynomial as a sum of Chebyshev polynomials of degree less than or equal to p,

$$T_p\left(\frac{y-1+2k/L}{2}\right) = 2^{-p}\sum_{i=0}^{p} w_i^{(p)}(k)T_i(y)$$

Hence

$$M_p^* = \frac{1}{2^p-1}\sum_{i=0}^{p-1}\left(\sum_{k=0}^{L} h_k w_i^{(p)}(k)\right)M_i^* \tag{3.10}$$

the $w_i^{(p)}(k)$ can be calculated recursively. We will use the notation $w_i^{(p)} = w_i^{(p)}(k)$ and $\lambda = 2k/L - 1$ for simplicity here. Now

$$T_{p+1}\left(\frac{y+\lambda}{2}\right) = 2^{-(p+1)}\sum_{i=0}^{p+1} w_i^{(p+1)}T_i(y) \tag{3.11}$$

3.5 Error analysis of the quadrature formula

Remember that the criterion of the degree of accuracy was chosen t have accurate results in case of smooth functions. In this section we make this statement more precise and also show why we need the fast wavelet transform to calculate the coefficients on the coarser levels.

Let $F_r[f(x)]$ denote the error of the quadrature formula,

$$F_r[f(x)] = \int_{-\infty}^{\infty} f(y)\phi(y)dy - Q_r[f(x)].$$

Suppose that $f \in C^{q+1}(supp\ \phi)$. One then immediately understands that the first $q+1$ terms of the Taylor formula of f around a point of $supp\ \phi$ are integrated exactly and that the error somehow depends on the $(q+1)$-th derivative of f.

To get a more precise formulation in the case $\tau = q$, we can reason along the following lines. Let x_0 be an arbitrary point of the interval $(0, L)$ not equal to one of the abscissae and let P_r be the polynomial of degree r that interpolates the function f in $x_0, \dots, x_r$. If f belongs to $C^{q+1}[0, L]$ and has bounded derivatives, then

$$\forall x \in supp\ \phi : \exists \xi(x) \in supp\ \phi : f(x) = P_r(x) + e_r(x)$$

with

$$e_r(x) = \frac{\Pi(x)(x - x_0)}{(r+1)!} f^{(r+1)}(\xi(x))$$

Then

$$\begin{aligned} F_r[f(x)] &= f_r[P_r(x) + e_r(x)] = F_r[P_r(x)] + F_r[e_r(x)] = F_r[e_r(x)] \\ &= < \phi, e_r > -Q_r[e_r(x)] \\ &= \frac{1}{(r+1)!} \int_r \phi(x)\Pi(x)(x - x_0) f^{(r+1)}(\xi(x)) dx \end{aligned}$$

with $\xi(x) \in supp\ \phi$. This error estimate, however, is not very useful in practice. One usually does not know the $(r+1)$-th derivative of the function. Moreover, on e has no control over the function $\xi(x)$. Deriving an upper bound usually leads to very pessimistic estimates.

However, our purpose is not so much to get accurate estimates on the error as to understand its asymptotic behavior. For the remainder of this section we switch back to the biorthogonal notation, i.e., the notation with tilde for the dual functions. Let $\lambda'_{n,l}$ be the computed approximation of the coefficients, where $\tilde{\phi}$ is taken as weight function,

$$\lambda'_{n,l} = 2^{-n/2} Q[f(2^{-n}(x + l)]$$

It then follows that

$$\frac{\lambda_{n,l} - \lambda'_{n,l}}{\lambda_{n,l}} = O(2^{-n(q+1)}). \qquad (3.12)$$

If one now uses the fast wavelet transform to calculate the coefficients on the coarser levels, they will all have a relative error of order $2^{-n(q+1)}$. On the other hand, the use of the quadrature formula directly on the coarser levels $j (j < n)$ will give an error of order $2^{-j(q+1)}$. This is precisely why we only use the quadrature formula on the finest level.

This analysis also helps us to choose the degree of accuracy. Remember that the multiresolution approximation of a function converges as

$$\|P_n f - f\| = O(h^N) \quad h = 2^{-n}$$

This immediately shows that the degree of accuracy should at least be $N-1$, otherwise the use of quadrature formula ruins the convergence rate of the multiresolution approximation. In other words, let

$$P_n^{'} f(x) = \sum_l \lambda_{n,l}^{'} \phi_{n,l}(x),$$

then

$$\|P_n^{'} f - f\} = O(h^N)$$

If $q \geq N-1$. At this moment $q = N-1$ seams the most natural choice.

Remember that the trapezoidal rile for $\tilde{\phi}$ has a degree of accuracy equal to $\tilde{N}-1$. This means that it does not ruin the convergence rate of the multiresolution approximation in case $\tilde{N} \geq N$¿ This is true in the orthogonal and semiorthogonal case.

Suppose we are only interested in one coefficient on a fixed level, let us say $\lambda_{0,0}$. One can then use a quadrature formula on a finer level n and use the fast wavelet transform to calculate $\lambda_{0,0}$. This scheme converges with an error of order $h^{q=1}$, with $h = 2^{-n}$. Using a generalization of Bernoulli polynomials it is possible to derive an asymptotic error expansion in powers of h for this scheme.

3.6 Fitting the formulae in the MRA

3.6.1 Using a quadrature formula at the finest level

Suppose we have to calculate T coefficients $\lambda_{n,l}$ yielding a function v_n,

$$v_n(x) = \sum_{l=0}^{T-1} \lambda_{n,l} \phi_{n,l}(x) \quad \lambda_{n,L} =< f.\tilde{\phi}_{n,l} > .$$

The quadrature formula and error estimation yield

$$\lambda_{n,l} = \sqrt{h} \left[\sum_{k=1}^{r} w_k f(h((k-1)2^s - \tau + l)) + O(h^{r+1}) \right] \tag{3.13}$$

where $h = 2^{-n}$. One usually wants to avoid evaluating the function f at abscissae with spacing smaller than 2^{-n}. This means that $s \geq 0$. The total number of evaluation for the calculation of T inner products is equal to $T + (r-1)2^s$. Note that this number is dominated by the first term and is only slightly dependent on r. As a result the one-point quadrature formula, which needs T evaluation , can be replaced with a quadrature formula with a lighter degree of accuracy, which requires in total almost the same number of evaluations. The workload is equal to $T \cdot r$ multiplication and $T(r-1)$ additions.

For a certain r, one usually wants to choose the maximal s to have abscissae spread put over the whole integration interval. The maximal s within the requirement that $(r-1)2^s < L$, however, corresponds to the smallest admittance interval for τ. If for a formula with $s > 0$ no τ can be found, one can always try to find a formula with spacing 2^{s-1}.