Wavelets Based on Orthogonal Polynomials

Bernd Fischer
Institut für Mathematik
Medizinische Universität zu Lübeck
23560 Lübeck, Germany

Jürgen Prestin
Fachbereich Mathematik
Universität Rostock
18051 Rostock, Germany

Abstract

We present a unified approach for the construction of polynomial wavelets. Our main tool are orthogonal polynomials. With the help of their properties we devise schemes for the construction of time localized polynomial bases on bounded and unbounded subsets of the real line. Several examples illustrate the new approach.

1. Introduction

In this paper we introduce and discuss a new method for the construction of time localized bases for polynomial subspaces of an L_2-space with arbitrary weight. Our analysis is based upon the theory of orthogonal polynomials. Whereas the frequency localization will be predetermined by the choice of the polynomial spaces, the time localization will be realized by the choice of special basis functions. More precisely, such a basis function will be defined as the solution of a constrained approximation problem which is designed such that its solution is maximally localized around a specified point.

Starting with the paper of Chui and Mhaskar [2], discussing trigonometric polynomial multiresolution analysis, the theory has been adapted to the algebraic polynomial case, see, e.g. Kilgore, Prestin [6] and Tasche [10]. They investigated the special case of the Chebyshev weight of the first kind. Their analysis is based on the properties of ordinary Chebyshev polynomials and does not carry over to other weight functions. In contrast, our derivations make use of the general theory of kernel polynomials. This allows us to treat not only weight functions which are supported on a compact interval (e.g., Jacobi weights) but also weight functions which are supported on the real line (e.g., Hermite weight) or on the real half line (e.g., Laguerre weight).

The paper is organized as follows. In Section 2 we collect some basic properties of orthogonal polynomials. Beside more theoretical results we discuss in particular computational aspects of orthogonal polynomials. Then we define scaling functions and wavelets and investigate some of their properties. This includes questions concerning orthogonality, interpolatory properties, time localization, and the construction of dual functions. In Section 3 we discuss the algorithms for reconstruction and decomposition. In particular we devise a compact matrix formulation for these schemes. Section 4 is concerned with questions related to the Riesz stability and to the (generalized) translation invariance of the proposed basis functions. Finally in Section 5 we apply the new approach to two different Chebyshev weights.

2. Scaling Functions and Wavelets

After having collected some auxiliary results for orthogonal polynomials, we will define in this section scaling functions and wavelets with respect to arbitrary weight functions.

2.1 Orthogonal Polynomials

Let $d\sigma(t)$ be a nonnegative measure on the real line, with compact or infinite support $[a, b]$, $-\infty \le a < b \le \infty$, for which all moments

$$\nu_r := \int_a^b t^r \, d\sigma(t), \quad r = 0, 1, \ldots, \tag{2.1}$$

exist and are finite with $\nu_0 > 0$. With $d\sigma(t)$ there is associated an inner product and a norm

$$\langle p, q \rangle := \int_a^b p(t) q(t) d\sigma(t), \quad \|p\| := \sqrt{\langle p, p \rangle}, \tag{2.2}$$

on the vector space of all polynomials. It is well-known (see, for example Szegö [8, §2.2]) that there exists a unique system of polynomials that are orthonormal with respect to this inner product, *i.e.*, a set of polynomials $\{P_r\}$ such that

$$\langle P_k, P_l \rangle = \delta_{k,l}. \tag{2.3}$$

In general the system $\{P_r\}$ consists of infinitely many polynomials, but reduces to a finite number, if $\sigma(t)$ has only finitely many points of increase. Throughout this paper we assume that $\sigma(t)$ has at least $2n+1$ points of increase and consequently $\{P_r\}_{r=0}^{2n}$ forms a basis for V_{2n}, where

$$V_n := \operatorname{span}\{P_0, P_1, \ldots, P_n\}. \tag{2.4}$$

An important special case are distributions of the form $w(t)dt$. Here we assume that the weight function $w(t)$ is nonnegative with $\int_a^b w(t)dt > 0$.

The orthogonal polynomials P_k fulfill the following three-term recurrence relation

$$\begin{aligned} P_{-1}(t) :&= 0, \quad P_0(t) = \nu_0^{-1/2}, \\ b_{k+1} P_k(t) &= (t - a_k) P_{k-1}(t) - b_k P_{k-2}(t), \quad k \ge 1. \end{aligned} \tag{2.5}$$

Let us collect together the three-term recurrence coefficients of $\{P_r\}_{r=0}^{n}$ into an unreduced symmetric tridiagonal matrix $\boldsymbol{J}_n$, the so-called *Jacobi matrix*,

$$\boldsymbol{J}_n := \begin{pmatrix} a_1 & b_2 & 0 & \cdots & 0 \\ b_2 & a_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_n \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix}. \tag{2.6}$$

With the vector

$$\boldsymbol{v}_n(t) := (P_0(t), P_1(t), \ldots, P_n(t))^{\mathrm{T}} \tag{2.7}$$

we can rewrite the three-term recurrence relation of the orthonormal polynomials (2.5) in compact matrix notation as

$$t\boldsymbol{v}_{n-1}(t) = \boldsymbol{J}_n \boldsymbol{v}_{n-1}(t) + b_{n+1} P_n(t) \boldsymbol{e}_n \tag{2.8}$$

where $\boldsymbol{e}_n := (0, 0, \ldots, 0, 1)^{\mathrm{T}}$ denotes the nth unit vector.

The next lemma collects some properties of the zeros of orthogonal polynomials. A proof of part (a) and (b) may be found in Szegö [8, Theorem 3.3.1, 3.3.2], whereas (c) follows directly from (2.8).

Lemma 2.1. *Let $y_r^{(n)}$, $r = 0, 1, \ldots, n-1$, denote the zeros of P_n.*

(a) The zeros of P_n are all real, simple and are located in (a, b)

$$a < y_0^{(n)} < y_1^{(n)} < \cdots < y_{n-1}^{(n)} < b.$$

(b) The zeros of P_n and P_{n+1} separate each other

$$y_0^{(n+1)} < y_0^{(n)} < y_1^{(n+1)} < \cdots < y_{n-1}^{(n)} < y_n^{(n+1)}.$$

(c) Any zero $y_r^{(n)}$ of P_n is an eigenvalue of $\boldsymbol{J}_n$ with eigenvector $\boldsymbol{v}_{n-1}(y_r^{(n)})$.

For a given fixed number $\xi \in \mathbb{R}$ the polynomial

$$K_n(t;\xi) := \sum_{k=0}^{n} P_k(t) P_k(\xi) \tag{2.9}$$

is called the *kernel polynomial* with respect to $\langle \cdot, \cdot \rangle$ (and the parameter ξ). Note that

$$K_n(\xi;\xi) = \sum_{k=0}^{n} P_k(\xi)^2 > 0. \tag{2.10}$$

The name “kernel” is motivated by the following result, which is also known as the *reproducing property* of the kernel polynomials (see, *e.g.*, Davis [3, §10.1]),

$$\langle K_n(\cdot;\xi), p \rangle = \int_a^b K_n(t;\xi) p(t) d\sigma(t) = p(\xi), \quad \text{for all} \quad p \in V_n. \tag{2.11}$$

The nth kernel polynomial $K_n(t;\xi)$ is the unique solution of the following constrained approximation problem (cf. Szegö [8, Theorem 3.1.3])

$$\left\| \frac{K_n(\cdot;\xi)}{K_n(\xi;\xi)} \right\| = \min \big\{ \|p\| : \ p \in V_n, \ p(\xi) = 1 \big\}. \tag{2.12}$$

2.2 Scaling Functions

Equation (2.12) indicates that the kernel polynomials are localized around ξ. Motivated by this property we define *scaling functions* as kernel polynomials

$$\varphi_{n,r}(t) = \varphi_n(t; x_r^{(n+1)}) := K_n(t; x_r^{(n+1)}), \quad r = 0, 1, \ldots, n, \tag{2.13}$$

with respect to a suitable set of parameter

$$x_0^{(n+1)} < x_1^{(n+1)} < \cdots < x_n^{(n+1)}. \tag{2.14}$$

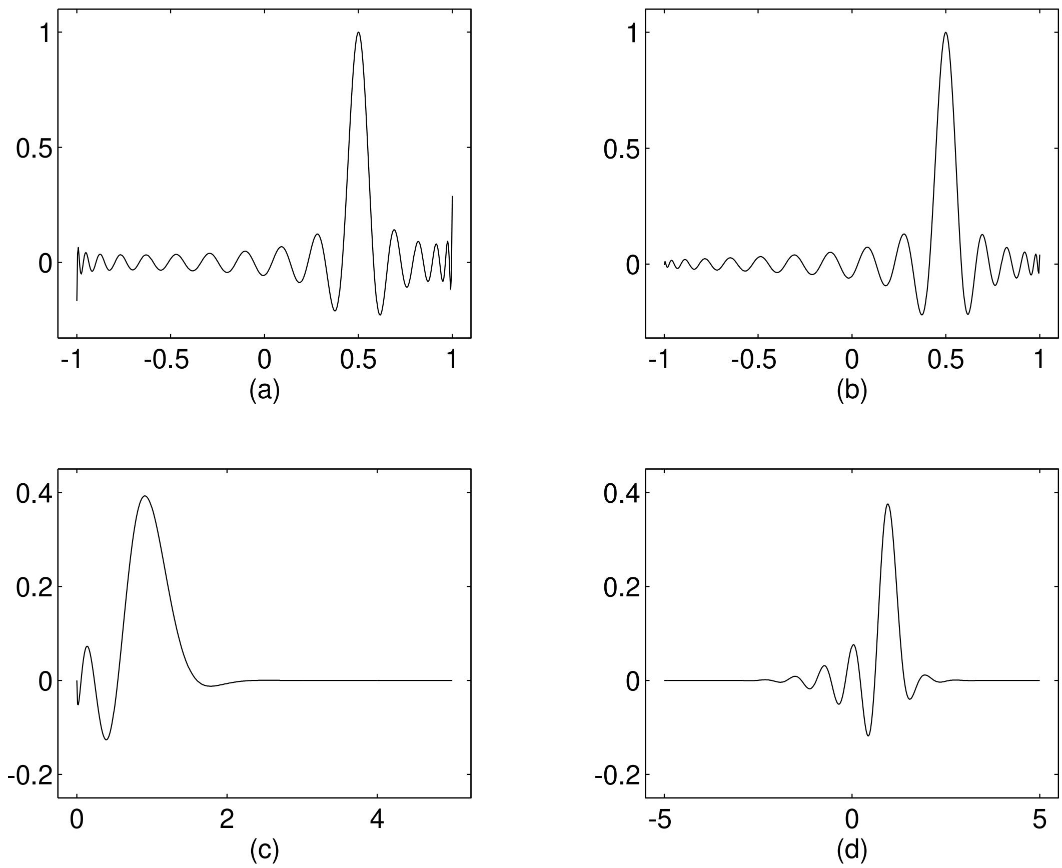

Figure 2.2. Various scaling functions of degree $n = 32$. (a) Legendre weight: $w(t) = 1$; $\varphi_{32}(t; 0.5)$; (b) Jacobi weight: $w(t) = (1-t)^{-0.5}(1+t)^{-0.8}$; $\varphi_{32}(t; 0.5)$; (c) Laguerre weight: $w(t) = \sqrt{t}\,\exp(-t)$; $\varphi_{32}(t; 1) \cdot w(t)$; (d) Hermite weight: $w(t) = \exp(-t^2)$; $\varphi_{32}(t; 1) \cdot w(t)$.

Note, that we plotted in (a) and (b) the "plain" polynomials and in (c) and (d) the polynomials times the underlying weight function. Actually one may view the scaling functions on the one hand as polynomial basis functions in a weighted L_2-space and on the other hand as weighted polynomial basis functions in an unweighted L_2-space.

Some properties of these polynomial scaling functions are summarized in the next theorem.

Theorem 2.3. *Let $\varphi_{n,r}(t) = \varphi_n(t; x_r^{(n+1)})$ denote the scaling functions with respect to a given set of parameter $x_0^{(n+1)} < x_1^{(n+1)} < \cdots < x_n^{(n+1)}$.*

(a) The inner product of scaling functions may be evaluated as follows

$$\langle \varphi_{n,r}, \varphi_{n,s} \rangle = \varphi_{n,r}(x_s^{(n+1)}), \quad r, s = 0, 1, \ldots, n.$$

(b) The scaling function $\varphi_{n,r}$ is localized around $x_r^{(n+1)}$. More precisely, we have

$$\left\| \frac{\varphi_{n,r}}{\varphi_{n,r}(x_r^{(n+1)})} \right\| = \min \left\{ \|p\| : \ p \in V_n, \ p(x_r^{(n+1)}) = 1 \right\}.$$

(c) The $\varphi_{n,r}$'s form a basis for V_n, i.e.,

$$V_n = \mathit{span}\{\varphi_{n,0}, \varphi_{n,1}, \ldots, \varphi_{n,n}\}.$$

(d) The scaling function $\varphi_{n,r}$ is orthogonal with respect to the "modified inner product" $\langle \cdot, \cdot \, (\cdot - x_r^{(n+1)}) \rangle$

$$\langle \varphi_{n,r}(\cdot), q(\cdot)(\cdot - x_r^{(n+1)}) \rangle = 0 \quad \textit{for all} \quad q \in V_{n-1}.$$

(e) The scaling function $\varphi_{n,r}$ satisfies the so-called Christoffel - Darboux identity

$$\varphi_{n,r}(t) = b_{n+2} \frac{P_{n+1}(t) P_n(x_r^{(n+1)}) - P_n(t) P_{n+1}(x_r^{(n+1)})}{t - x_r^{(n+1)}},$$

where b_{n+2} is a three-term recurrence coefficient of P_{n+1} (cf. (2.5)).

(f) Let $\{y_k^{(n)}\}_{k=0}^{n-1}$ and $\{y_k^{(n+1)}\}_{k=0}^{n}$ denote the zeros of P_n and P_{n+1}, respectively. Moreover, define $y_{-1}^{(n)} := -\infty$ and $y_n^{(n)} := \infty$.
If $x_r^{(n+1)} = y_j^{(n)}$ is a zero of P_n, then $\varphi_{n,r}$ has the $n-1$ zeros $y_k^{(n)}$, $k = 0, 1, \ldots, j-1, j+1, \ldots, n-1$.
If $x_r^{(n+1)} = y_j^{(n+1)}$ is a zero of P_{n+1}, then $\varphi_{n,r}$ has the n zeros $y_k^{(n+1)}$, $k = 0, 1, \ldots, j-1, j+1, \ldots, n$.
If $x_r^{(n+1)} \in (y_j^{(n+1)}, y_j^{(n)})$, then $\varphi_{n,r}$ has precisely one zero in each interval $(y_k^{(n+1)}, y_k^{(n)})$, $k = 0, 1, \ldots, j-1, j+1, \ldots, n$.
If $x_r^{(n+1)} \in (y_{j-1}^{(n)}, y_j^{(n+1)})$, then $\varphi_{n,r}$ has precisely one zero in each interval $(y_{k-1}^{(n)}, y_k^{(n+1)})$, $k = 0, 1, \ldots, j-1, j+1, \ldots, n$.

Proof. Part (a) and (b) follow immediately from (2.11) and (2.12), respectively.

To verify (c), assume that

$$\sum_{r=0}^{n} \tau_r \varphi_{n,r}(t) \equiv 0. \tag{2.15}$$

Furthermore, let $\{\ell_r\}_{r=0}^n$ denote the set of fundamental polynomials of Lagrange interpolation with respect to the knots (2.14), *i.e.*,

$$\ell_r \in V_n \quad \text{and} \quad \ell_r(x_s^{(n+1)}) = \delta_{r,s}, \quad r, s = 0, 1, \ldots, n. \tag{2.16}$$

In view of the assumption (2.15) and the reproducing property (2.11) we deduce

$$0 = \left\langle \sum_{r=0}^{n} \tau_r \varphi_{n,r}, \ell_s \right\rangle = \sum_{r=0}^{n} \tau_r \ell_s(x_r^{(n+1)}) = \tau_s,$$

for $s = 0, 1, \ldots, n$, which shows the linear independence of the $\varphi_{n,r}$'s.

(d) is nothing but the reproducing property (2.11) applied to the polynomial $p(t) = (t - x_r^{(n+1)})q(t)$.

(e) follows readily from the three-term recurrence relation (2.5) (compare Szegö [8, Theorem 3.2.2].

(f) is a direct consequence of the Christoffel - Darboux identity (e) (compare Fischer [4, Theorem 2.5.8]). ∎

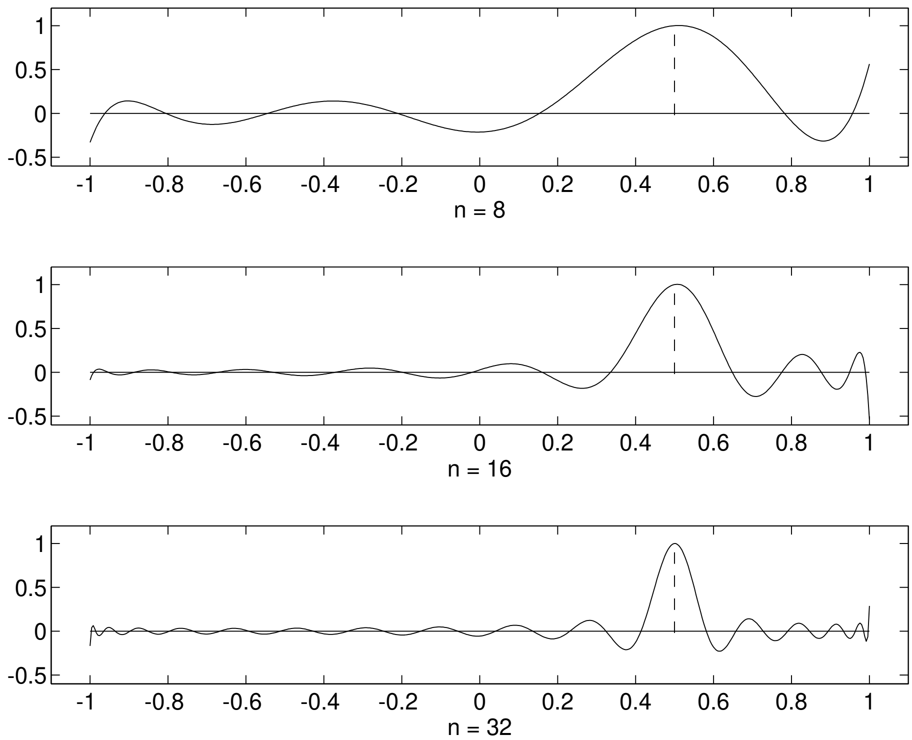

Figure 2.4. Scaling functions $\varphi_n(t; 0.5)$ of degree $n = 8, 16, 32$ with respect to the Legendre weight $w(t) = 1$ and the parameter $x_0^{(n+1)} = 0.5$.

Note that the interlacing property of the zeros of orthogonal polynomials (cf. Lemma 2.1(b)) together with part (f) implies that the "zero-free interval" around the constraint point shrinks with increasing degree, as is apparent from Figure 2.4.

Part (a) of the theorem above implies that the scaling functions (2.13) are orthogonal to each other if, and only if they fulfill the interpolatory property $\varphi_{n,r}(x_s^{(n+1)}) = d_r^{(n+1)}\delta_{r,s}$, $d_r^{(n+1)} \in \mathbb{R}$. This may be seen as a requirement for the parameter set (2.14). The next theorem characterizes the parameter sets which lead to orthogonal scaling functions.

Theorem 2.5. *Let $\varphi_{n,r}(t) = \varphi_n(t; x_r^{(n+1)})$ denote the scaling functions with respect to a given parameter set $x_0^{(n+1)} < x_1^{(n+1)} < \cdots < x_n^{(n+1)}$. Then the following conditions are equivalent to the orthogonality of the scaling functions.*

(a) *The scaling functions satisfy an interpolatory condition*

$$\varphi_{n,r}(x_s^{(n+1)}) = d_r^{(n+1)}\delta_{r,s}, \quad \text{for} \quad r,s = 0,1,\ldots,n,$$

where $d_r^{(n+1)} \in \mathbb{R}$.

(b) *The parameter $x_r^{(n+1)}$ define a quadrature rule which is exact for polynomials of degree $2n$, i.e.,*

$$\int_a^b p(t)d\sigma(t) = \sum_{r=0}^{n} \left(d_r^{(n+1)}\right)^{-1} p(x_r^{(n+1)}), \quad \text{for all} \quad p \in V_{2n},$$

where $d_r^{(n+1)} = \varphi_{n,r}(x_r^{(n+1)})$.

(c) *The polynomial $q_{n+1}(t) := \prod_{r=0}^{n}(t - x_r^{(n+1)})$ is quasi-orthogonal, i.e.,*

$$\langle q_{n+1}, t^k \rangle = 0, \quad \text{for} \quad k = 0,1,\ldots,n-1.$$

(d) *There exists a number τ_n with*

$$P_{n+1}(x_r^{(n+1)}) + \tau_n P_n(x_r^{(n+1)}) = 0, \quad \text{for} \quad r = 0,1,\ldots,n.$$

Proof. For convenience we drop the superscripts, *i.e.*, $x_r = x_r^{(n+1)}$ and $d_r = d_r^{(n+1)}$. We first show that (b) follows from (a). To this end we assume that

$$d_r^{-1}\varphi_{n,r}(x_s) = \sum_{k=0}^{n} d_r^{-1} P_k(x_r) P_k(x_s) = \delta_{r,s},$$

and conclude

$$\begin{aligned} P_l(x_s) &= \sum_{r=0}^{n} P_l(x_r)\delta_{r,s} \\ &= \sum_{r=0}^{n} P_l(x_r) \sum_{k=0}^{n} d_r^{-1} P_k(x_r) P_k(x_s) \\ &= \sum_{k=0}^{n} P_k(x_s) \sum_{r=0}^{n} d_r^{-1} P_l(x_r) P_k(x_r), \end{aligned}$$

for $s = 0, 1, \ldots, n$. Hence, the polynomial

$$P_l(t) - \sum_{k=0}^{n} P_k(t) \sum_{r=0}^{n} d_r^{-1} P_l(x_r) P_k(x_r)$$

has $n+1$ zeros x_s. For $l \leq n$ this is only possible if

$$\sum_{r=0}^{n} d_r^{-1} P_l(x_r) P_k(x_r) = \delta_{l,k}, \quad \text{for} \quad l, k = 0, 1, \ldots, n. \tag{2.17}$$

On the other hand, the orthonormality of the P_j's

$$\int_a^b P_l(t) P_k(t) d\sigma(t) = \delta_{l,k}, \quad l, k = 0, 1, \ldots, n,$$

implies that (2.17) constitutes a quadrature rule for polynomials of the form $P_l P_k$. Finally, observe that the product $P_l P_k$ has exact degree $l + k$ which clearly shows that

$$V_{2n} = \operatorname{span}\{P_l P_k : \ l, k = 0, 1, \ldots, n\}.$$

The proof for the statement that (a) follows from (b) is along the same lines and is therefore omitted here.

To show that (c) follows from (b) observe that

$$\langle q_{n+1}, t^k \rangle = \int_a^b q_{n+1}(t) t^k d\sigma(t) = \sum_{r=0}^{n} d_r^{-1} q_{n+1}(x_r^{(n+1)})(x_r^{(n+1)})^k = 0,$$

for $t^k q_{n+1} \in V_{2n}$, *i.e.*, for $k \leq n-1$.

Conversely, let $p_{2n} \in V_{2n}$ be given. Then there exist polynomials $p_{n-1} \in V_{n-1}$ and $p_n \in V_n$ with

$$p_{2n}(t) = p_{n-1}(t) q_{n+1}(t) + p_n(t).$$

Now we make use of the quasi-orthogonality of q_{n+1}, and the fact that any polynomial of degree n can be integrated by an interpolatory quadrature rule based on $n+1$ given knots, to obtain

$$\begin{aligned}
\int_a^b p_{2n}(t) d\sigma(t) &= \int_a^b p_{n-1}(t) q_{n+1}(t) d\sigma(t) + \int_a^b p_n(t) d\sigma(t) \\
&= \int_a^b p_n(t) d\sigma(t) \\
&= \sum_{r=0}^{n} e_r^{-1} p_n(x_r^{(n+1)}) \\
&= \sum_{r=0}^{n} e_r^{-1} p_{2n}(x_r^{(n+1)}).
\end{aligned} \tag{2.18}$$

It remains to show that $e_r^{-1} = d_r^{-1}$, $r = 0, 1, \ldots, n$. This, however, follows from the implication (b) $\Rightarrow$ (a).

For the rest of the proof we refer to Chihara [1, Ch. II, Theorem 5.1, 5.3]. $\square$

In particular part (d) of the theorem above is quite useful for actually computing parameters $x_r^{(n+1)}$ which correspond to orthogonal scaling functions. Note, that the interlacing property (cf. Lemma 2.1(b)) immediately implies that the polynomial

$$P_{n+1}(t) + \tau_n P_n(t)$$

has $n+1$ real and simple zeros, where at most one of these zeros lies outside the "orthogonality interval" $[a, b]$ (compare Chihara [1, Ch. I, Theorem 5.2]).

Probably the most important special case is provided by the choice $\tau_n = 0$.

Corollary 2.6. *Let $y_r^{(n+1)}$, $r = 0, 1, \ldots, n$, denote the zeros of P_{n+1} and let $\varphi_{n,r}(t) = \varphi_n(t; y_r^{(n+1)})$ denote the associated scaling functions (2.13). Then*

$$\langle \varphi_{n,r}, \varphi_{n,s} \rangle = \varphi_{n,r}(y_s^{(n+1)}) = c_r^{(n+1)} \ \delta_{r,s}, \quad r, s = 0, 1, \ldots, n,$$

where the $c_r^{(n+1)}$'s are given by the weights in the classical Gaussian quadrature rule

$$\int_a^b p(t) d\sigma(t) = \sum_{r=0}^{n} \left(c_r^{(n+1)} \right)^{-1} p(y_r^{(n+1)}), \quad \textit{for all} \quad p \in V_{2n+1}.$$

We remark that the $\varphi_{n,r}$ may be viewed as fundamental polynomials of Lagrange interpolation with respect to the knots $y_r^{(n+1)}$.

2.3 Wavelets

In this section we define our wavelets and discuss some of their properties. To this end let

$$W_n := V_{2n} \ominus V_n = \text{span}\{P_{n+1}, P_{n+2}, \ldots, P_{2n}\}. \tag{2.19}$$

Note that

$$\dim W_n = n. \tag{2.20}$$

The goal is to identify functions, our wavelets, which define a localized basis for W_n.

In accordance with the definition of the scaling function (2.13) we define the wavelets, for $r = 0, 1, \ldots, n-1$, in terms of kernel functions

$$\begin{aligned} \psi_{n,r}(t) = \psi_n(t; z_r^{(n)}) &:= K_{2n}(t; z_r^{(n)}) - K_n(t; z_r^{(n)}) \\ &= \sum_{k=n+1}^{2n} P_k(z_r^{(n)}) P_k(t), \end{aligned} \tag{2.21}$$

for a suitable set of parameter

$$z_0^{(n)} < z_1^{(n)} < \cdots < z_{n-1}^{(n)}. \tag{2.22}$$

Note that the interlacing property implies $\psi_n(z_r^{(n)}) > 0$, for $n > 1$.

The next figure shows some typical wavelets. For a plot of the corresponding scaling functions we refer to Figure 2.2.

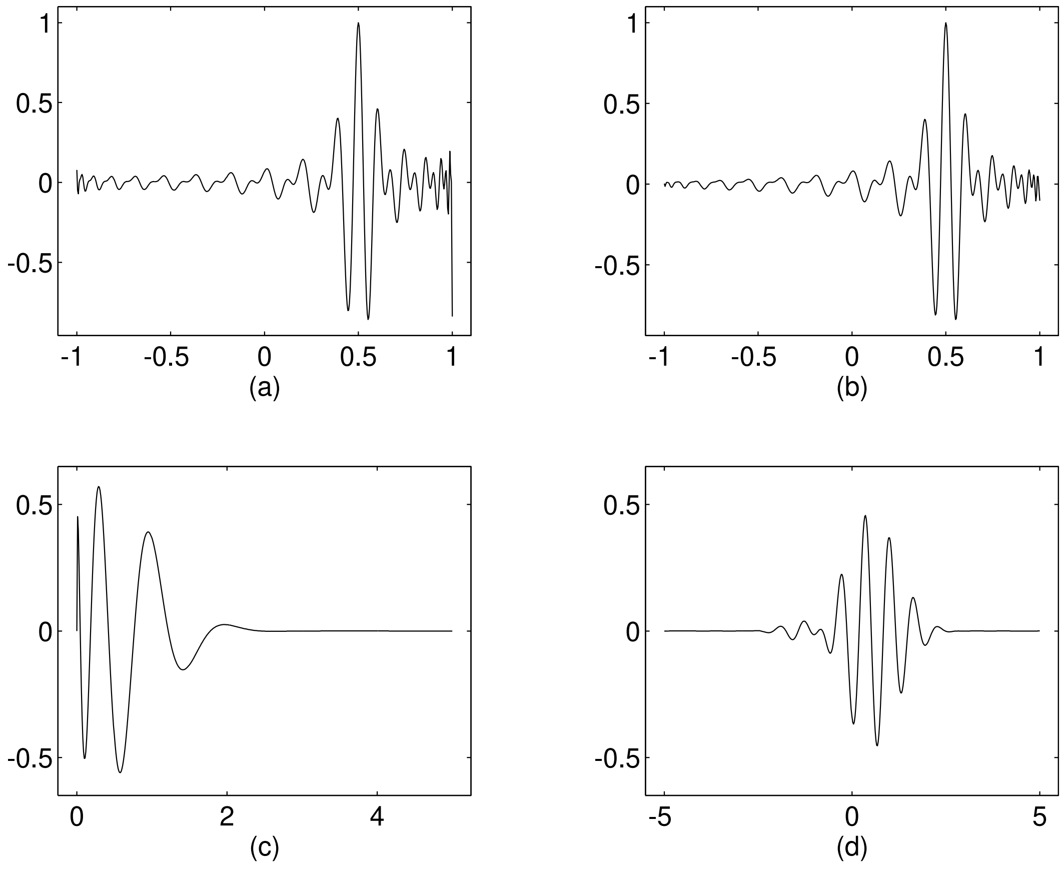

Figure 2.7. Various wavelets of degree $n = 32$. (a) Legendre weight: $w(t) = 1$; $\psi_{32}(t; 0.5)$; (b) Jacobi weight: $w(t) = (1-t)^{-0.5}(1+t)^{-0.8}$; $\psi_{32}(t; 0.5)$; (c) Laguerre weight: $w(t) = \sqrt{t}\,\exp(-t)$; $\psi_{32}(t; 1)\cdot w(t)$; (d) Hermite weight: $w(t) = \exp(-t^2)$; $\psi_{32}(t; 1)\cdot w(t)$.

The next theorem collects some properties of the wavelets $\psi_{n,r}$. Note, that part (a) and (b) are similar to the one for the associated scaling functions (cf. Theorem 2.3). We stress that these properties do not depend on the particular choice of the parameter set $\{z_r^{(n)}\}_{r=0}^{n-1}$.

Theorem 2.8. *Let $\psi_{n,r}(t) := \psi_n(t; z_r^{(n)})$ denote the wavelets with respect to the parameter $z_0^{(n)} < z_1^{(n)} < \cdots < z_{n-1}^{(n)}$.*

(a) The inner product of wavelets may be evaluated as follows

$$\langle \psi_{n,r}, \psi_{n,s} \rangle = \psi_{n,r}(z_s^{(n)}), \quad r, s = 0, 1, \ldots, n-1.$$

(b) The wavelet $\psi_{n,r}$ is localized around $z_r^{(n)}$

$$\left\| \frac{\psi_{n,r}}{\psi_{n,r}(z_r^{(n)})} \right\| = \min \left\{ \|p\| : \ p \in W_n, \ p(z_r^{(n)}) = 1 \right\}.$$

(c) Let $\varphi_{n,r}(t) := \varphi_n(t; x_r^{(n+1)})$ (cf. (2.13)) denote the scaling functions with respect to the parameter $x_0^{(n+1)} < x_1^{(n+1)} < \cdots < x_n^{(n+1)}$. The wavelets and the scaling functions are orthogonal to each other

$$\langle \psi_{n,r}, \varphi_{n,s} \rangle = 0, \quad r, s = 0, 1, \ldots, n-1.$$

Proof. For convenience we drop the superscript $z_r := z_r^{(n)}$.

To verify (a), we show that the wavelets fulfill a reproducing property with respect to W_n. In fact, for $p \in W_n$ we have by (2.21) and (2.11)

$$\begin{aligned} \langle \psi_{n,r}, p \rangle &= \langle K_{2n}(\cdot; z_r) - K_n(\cdot; z_r), p \rangle \\ &= \langle K_{2n}(\cdot; z_r), p \rangle - \langle K_n(\cdot; z_r), p \rangle \\ &= p(z_r). \end{aligned} \tag{2.23}$$

(b) The proof is along the lines of the proof for the standard case of kernel polynomials (cf. Chihara [1, Ch.I, Theorem 7.3]). Let $p \in W_n$ with $p(z_r) = 1$, *i.e.*,

$$p(t) = \sum_{k=n+1}^{2n} d_k P_k(t), \quad p(z_r) = \sum_{k=n+1}^{2n} d_k P_k(z_r) = 1. \tag{2.24}$$

The orthonormality of the P_j's implies

$$\langle p, p \rangle = \sum_{k=n+1}^{2n} d_k^2.$$

This identity together with (2.24) and the Cauchy-Schwarz inequality (applied to the Euclidian inner product) yields

$$\begin{aligned} 1 = p^2(z_r) &= ((d_{n+1}, d_{n+2}, \ldots, d_{2n})(P_{n+1}(z_r), P_{n+2}(z_r), \ldots, P_{2n}(z_r))^{\mathrm{T}})^2 \\ &\leq \|p\|^2 \sum_{k=n+1}^{2n} P_k^2(z_r). \end{aligned}$$

On the other hand we have

$$\left\| \frac{\psi_{n,r}}{\psi_{n,r}(z_r)} \right\|^2 = \frac{\langle \psi_{n,r}, \psi_{n,r} \rangle}{\left(\sum_{k=n+1}^{2n} P_k^2(z_r) \right)^2} = \frac{1}{\sum_{k=n+1}^{2n} P_k^2(z_r)},$$

which concludes the proof of statement (b).

Part (c) follows directly from the definition of the participating functions. □

It is worth noticing that in accordance with the properties of scaling functions (cf. Theorem 2.3(a)) the wavelets are orthogonal if, and only if they satisfy an interpolatory condition $\psi_{n,r}(z_s^{(n)}) = 0$, for $r \neq s$. We will present in Section 5 an example of orthogonal wavelets. In general, however, it is not clear whether there exist orthogonal wavelets for a given inner product.

Moreover, not any set $\{z_r^{(n)}\}_{r=0}^{n-1}$ leads to linear independent wavelet functions. For example, let the $z_r^{(n)}$ be zeros of P_s, *i.e.*,

$$P_s(z_r^{(n)}) = 0, \quad r = 0, 1, \ldots, n-1, \; n+1 \leq s \leq 2n.$$

Then the wavelets

$$\psi_n(t; z_r^{(n)}) = \sum_{k=n+1, k\neq s}^{2n} P_k(z_r^{(n)}) P_k(t), \quad r = 0, 1, \ldots, n-1$$

can span at best a space of dimension $n-1$. However, we have the following theorem.

Theorem 2.9. *Let* $z_r^{(n)} = y_r^{(n)}$, $r = 0, 1, \ldots, n-1$, *denote the zeros of* P_n *and let* $\psi_{n,r}(t) = \psi_n(t; y_r^{(n)})$ *denote the associated wavelets. Then*

$$W_n = \mathit{span}\{\psi_{n,0}, \psi_{n,1}, \ldots, \psi_{n,n-1}\}.$$

Proof. We show that the $\{\psi_{n,r}\}_{r=0}^{n-1}$ are linearly independent. To this end, assume that

$$\sum_{r=0}^{n-1} \sigma_r \psi_{n,r}(t) \equiv 0.$$

Since the P_j's are orthogonal we have

$$\langle P_n, P_i P_j \rangle \begin{cases} \neq 0 & \text{for } i+j = n, \\ = 0 & \text{for } i+j < n. \end{cases}$$

This together with the reproducing property (2.11) implies, for $i = 1, 2, \dots, n$,

$$\begin{aligned}
0 &= \left\langle \sum_{r=0}^{n-1} \sigma_r \psi_{n,r}, P_n P_i \right\rangle \\
&= \sum_{r=0}^{n-1} \sigma_r \langle K_{2n}(\cdot; y_r^{(n)}) - K_n(\cdot; y_r^{(n)}), P_n P_i \rangle \\
&= -\sum_{r=0}^{n-1} \sigma_r \langle K_n(\cdot; y_r^{(n)}), P_n P_i \rangle \\
&= -\sum_{r=0}^{n-1} \sigma_r \sum_{j=0}^{n-1} P_j(y_r^{(n)}) \langle P_n, P_i P_j \rangle \\
&= -\sum_{j=0}^{n-1} \langle P_n, P_i P_j \rangle \sum_{r=0}^{n-1} \sigma_r P_j(y_r^{(n)}) \\
&= -\sum_{j=n-i}^{n-1} \langle P_n, P_i P_j \rangle \sum_{r=0}^{n-1} \sigma_r P_j(y_r^{(n)}).
\end{aligned}$$

In other words, we end up with a triangular homogeneous linear system in the unknown $\sum_{r=0}^{n-1} \sigma_r P_j(y_r^{(n)})$. Since the entries on the main diagonal do not vanish, it has the unique solution

$$\sum_{r=0}^{n-1} \sigma_r P_j(y_r^{(n)}) = 0, \quad j = 0, 1, \dots, n.$$

This, however, is only possible for $\sigma_0 = \sigma_1 = \dots = \sigma_{n-1} = 0$, because the vectors

$$\boldsymbol{v}_{n-1}(y_r^{(n)}) = (P_0(y_r^{(n)}), P_1(y_r^{(n-1)}), \dots, P_n(y_r^{(n)}))^{\mathrm{T}}, \quad r = 0, 1, \dots, n-1,$$

are linearly independent as eigenvectors of $\boldsymbol{J}_{n-1}$ (cf. Lemma 2.1(d)). □

2.4 Dual functions

For practical purposes it is important to get a hand on the dual functions $\tilde{\varphi}_{n,r} \in V_n$ and $\tilde{\psi}_{n,r} \in W_n$. They are uniquely determined by the following biorthogonality relations

$$\begin{aligned}
\langle \varphi_{n,s}, \tilde{\varphi}_{n,r} \rangle &= \delta_{r,s}, \quad r, s = 0, 1, \dots, n, \\
\langle \psi_{n,s}, \tilde{\psi}_{n,r} \rangle &= \delta_{r,s}, \quad r, s = 0, 1, \dots, n-1.
\end{aligned} \tag{2.25}$$

Of course, here we have to assume that the wavelets $\psi_{n,r}$ constitute a basis for W_n. The next theorem shows that the dual functions are easy to identify. The proof follows directly from (2.25) and the reproducing properties (2.11) and (2.23), respectively.

Theorem 2.10. *Let V_n and W_n be defined as in (2.4) and (2.7), respectively.*

(a) *The dual scaling functions $\tilde{\varphi}_{n,r} = \ell_r$ (cf. (2.16)) are the fundamental polynomials of Lagrange interpolation with respect to the given parameter set $x_0^{(n+1)}, x_1^{(n+1)}, \ldots, x_n^{(n+1)}$, i.e.,*

$$\tilde{\varphi}_{n,r} \in V_n \quad \text{and} \quad \tilde{\varphi}_{n,r}(x_s^{(n+1)}) = \delta_{r,s} \quad r,s = 0,1,\ldots,n.$$

(b) *Let $\{\psi_{n,r}\}_{r=0}^{n-1}$ be a basis for W_n. Then the dual wavelet functions $\tilde{\psi}_{n,r} \in W_n$ are the fundamental polynomials of Lagrange interpolation with respect to the given parameter set $z_0^{(n)}, z_1^{(n)}, \ldots, z_{n-1}^{(n)}$, i.e.,*

$$\tilde{\psi}_{n,r} \in W_n \quad \text{and} \quad \tilde{\psi}_{n,r}(z_s^{(n)}) = \delta_{r,s} \quad r,s = 0,1,\ldots,n-1.$$

For the actual computation of the dual functions we refer to the next section. We would like to point out that the dual functions as well satisfy a localization property with respect to a discrete measure. More precisely, it holds (compare Theorem 2.3(b))

$$\|\tilde{\varphi}_{n,r}\|_{n+1} = \min\left\{\|p\|_{n+1} : \ p \in V_n, \ p(x_r^{(n+1)}) = 1\right\},$$

where

$$\|p\|_{n+1} := \left(\sum_{s=0}^{n} |p(x_s^{(n+1)})|^2\right)^{1/2}.$$

Analogously, we have for the wavelet space (compare Theorem 2.8(b))

$$\left\|\tilde{\psi}_{n,r}\right\|_n = \min\left\{\|p\|_n : \ p \in W_n, \ p(z_r^{(n)}) = 1\right\},$$

where

$$\|p\|_n := \left(\sum_{s=0}^{n-1} |p(z_s^{(n)})|^2\right)^{1/2}.$$

Finally, let us mention that the dual functions are in general no kernel functions with respect to the set of orthonormal polynomials P_k as in (2.13) and (2.21). On the other hand, however, they do have a representation in terms of kernel polynomials with respect to the orthonormal polynomials defined by the corresponding discrete inner product.

3. Two-Scale Relations and Decomposition

The purpose of this section is to describe reconstruction and decomposition algorithms of given functions. The schemes are based on the space representation $V_{2n} = V_n \oplus W_n$. Clearly, a repeated application of this step would result in a multiresolution of a weighted L_2-space.

3.1 Matrix Notation

We start by noting that in view of (2.13) any function $f_n \in V_n$, represented by the vector $\boldsymbol{a}^{(n)} := \left(a_0^{(n)}, \ldots, a_n^{(n)}\right)^T$,

$$(3.1) \qquad f_n(t) = \sum_{r=0}^{n} a_r^{(n)} \varphi_n(t; x_r^{(n+1)}) = (P_0(t), \ldots, P_n(t))\, \boldsymbol{A}_n\ \boldsymbol{a}^{(n)},$$

may be written in terms of the matrix

$$(3.2) \qquad \begin{aligned} \boldsymbol{A}_n := \left(P_k(x_r^{(n+1)})\right)_{k,r=0,1,\ldots,n} &= \begin{pmatrix} P_0(x_0^{(n+1)}) & \cdots & P_0(x_n^{(n+1)}) \\ \vdots & \ddots & \vdots \\ P_n(x_0^{(n+1)}) & \cdots & P_n(x_n^{(n+1)}) \end{pmatrix} \\ &= \left(\boldsymbol{v}_n(x_0^{(n+1)}), \ldots, \boldsymbol{v}_n(x_n^{(n+1)})\right). \end{aligned}$$

Analogously, we obtain for $g_n \in W_n$, with $\boldsymbol{b}^{(n)} := \left(b_0^{(n)}, \ldots, b_{n-1}^{(n)}\right)^T$, the representation

$$(3.3) \qquad g_n(t) = \sum_{r=0}^{n-1} b_r^{(n)} \psi_n(t; z_r^{(n)}) = (P_{n+1}(t), \ldots, P_{2n}(t))\, \boldsymbol{B}_n\ \boldsymbol{b}^{(n)},$$

where

$$(3.4) \quad \boldsymbol{B}_n := \left(P_{k+n+1}(z_r^{(n)})\right)_{k,r=0,1,\ldots,n-1} = \begin{pmatrix} P_{n+1}(z_0^{(n)}) & \cdots & P_{n+1}(z_{n-1}^{(n)}) \\ \vdots & \ddots & \vdots \\ P_{2n}(z_0^{(n)}) & \cdots & P_{2n}(z_{n-1}^{(n)}) \end{pmatrix}.$$

It is the purpose of this section to study the matrices $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$, respectively, in more detail.

Recall that by Theorem 2.3(c) the scaling functions are linearly independent, *i.e.*,

$$\sum_{r=0}^{n} \sigma_r \varphi_n(t; x_r^{(n+1)}) = \sum_{r=0}^{n} \sigma_r \sum_{k=0}^{n} P_k(x_r^{(n+1)}) P_k(t) = \sum_{k=0}^{n} P_k(t) \sum_{r=0}^{n} \sigma_r P_k(x_r^{(n+1)}) = 0$$

implies $\sigma_0 = \sigma_1 = \ldots = \sigma_n = 0$. We learn from the above equation that also the vectors $\boldsymbol{v}_n(x_r^{(n+1)})$ are linearly independent, which are just the columns of $\boldsymbol{A}_n$. In fact, $\boldsymbol{A}_n$ has to be regular as coefficient matrix for the interpolation problem at the knots $x_r^{(n+1)}$ with respect to the space spanned by the P_j's.

Corollary 3.11. *Let $x_0^{(n+1)} < x_1^{(n+1)} < \cdots < x_n^{(n+1)}$ be given.*

(a) *The matrix $\boldsymbol{A}_n$ is regular.*

(b) *The scaling functions $\varphi_n(t; x_r^{(n+1)})$, $r = 0, 1, \ldots, n$, are orthogonal and interpolatory (cf. Theorem 2.3(a)) if, and only if $\boldsymbol{A}_n^{\mathrm{T}} \boldsymbol{A}_n$ is a diagonal matrix.*

In the light of Corollary 2.6 it should come as no surprise that the matrix $\boldsymbol{A}_n$ based on the zeros of P_{n+1} is special.

Corollary 3.12. *Let $x_r^{(n+1)} = y_r^{(n+1)}$, $r = 0, 1, \ldots, n$, denote the zeros of P_{n+1} and let $c_r^{(n+1)}$ denote the weights of the Gaussian quadrature rule (cf. Corollary 2.6). Then the columns of $\boldsymbol{A}_n$ are the eigenvectors of $\boldsymbol{J}_n$. Moreover,*

$$\begin{aligned}\boldsymbol{A}_n^{\mathrm{T}} \boldsymbol{A}_n &= \left(\boldsymbol{v}_n(y_k^{(n+1)})^{\mathrm{T}} \boldsymbol{v}_n(y_r^{(n+1)}) \right)_{k,r=0}^{n} \\ &= \operatorname{diag}\left((c_0^{(n+1)})^{-1}, \ldots, (c_n^{(n+1)})^{-1} \right) =: \boldsymbol{D}_n,\end{aligned}$$

and

$$\boldsymbol{A}_n^{-1} = \boldsymbol{D}_n^{-1} \boldsymbol{A}_n^{\mathrm{T}}.$$

To discuss properties of $\boldsymbol{B}_n$ note that

$$\sum_{r=0}^{n-1} \sigma_r \psi_n(t; z_r^{(n)}) = \sum_{r=0}^{n-1} \sigma_r \sum_{k=n+1}^{2n} P_k(z_r^{(n)}) P_k(t) = \sum_{k=n+1}^{2n} P_k(t) \sum_{r=0}^{n-1} \sigma_r P_k(z_r^{(n)}).$$

Hence, the wavelets are linear independent if, and only if the matrix $\boldsymbol{B}_n$ is regular. The next corollary follows from Section 2.3 and in particular from Theorem 2.9.

Corollary 3.13. *Let $z_0^{(n)} < z_1^{(n)} < \cdots < z_{n-1}^{(n)}$ be given.*

(a) *The matrix $\boldsymbol{B}_n$ is not necessarily regular.*

(b) *The wavelets $\psi_n(t; z_r^{(n)})$, $r = 0, 1, \ldots, n-1$, are orthogonal and interpolatory (cf. Theorem 2.8(a)) if, and only if $\boldsymbol{B}_n^{\mathrm{T}} \boldsymbol{B}_n$ is a diagonal matrix.*

In Theorem 2.9 we identified a set of parameter which leads to linear independent wavelets or, equivalently, to a regular $\boldsymbol{B}_n$.

Corollary 3.14. *Let $z_r^{(n)} = y_r^{(n)}$, $r = 0, 1, \ldots, n-1$, denote the zeros of P_n. Then $\boldsymbol{B}_n$ is regular.*

Proof. For later reference we offer a proof which is different from the one of Theorem 2.9. It provides a convenient expression for $\boldsymbol{B}_n^{-1}$. Namely, a straightforward computation shows that

$$\begin{aligned}\boldsymbol{B}_n \boldsymbol{A}_{n-1}^{-1} &= \boldsymbol{B}_n \boldsymbol{D}_{n-1}^{-1} \boldsymbol{A}_{n-1}^{\mathrm{T}} \\ &= \left(\sum_{r=0}^{n-1} c_r^{(n)} P_{k+n+1}(y_r^{(n)}) P_l(y_r^{(n)}) \right)_{k,l=0,1,\ldots,n-1},\end{aligned} \tag{3.5}$$

where the matrix $\boldsymbol{A}_{n-1}$ is based on the parameter set $y_r^{(n)}$. It turns out that this matrix is triangular with nonvanishing anti-main diagonal entries. To justify this statement observe that by Gaussian quadrature

$$0 = \int_a^b P_{k+n+1}(t) P_l(t) d\sigma(t) = \sum_{r=0}^{n-1} c_r^{(n)} P_{k+n+1}(y_r^{(n)}) P_l(y_r^{(n)}),$$

for $k+n+1+l \leq 2n-1$. It follows that $\det(\boldsymbol{B}_n \boldsymbol{D}_{n-1}^{-1} \boldsymbol{A}_{n-1}^{\mathrm{T}}) \neq 0$ and consequently $\det \boldsymbol{B}_n \neq 0$. ∎

Finally, let us summarize the relationships between the various introduced bases for V_n and W_n, respectively.

Corollary 3.15. *Let $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ be defined by (3.2) and by (3.4), respectively.*

(a) *For a given arbitrary parameter set $x_r^{(n+1)}$, $r = 0, 1, \ldots, n$, we have*

$$\begin{pmatrix} \varphi_{n,0} \\ \vdots \\ \varphi_{n,n} \end{pmatrix} = \boldsymbol{A}_n^T \begin{pmatrix} P_0 \\ \vdots \\ P_n \end{pmatrix},$$

$$\begin{pmatrix} \tilde{\varphi}_{n,0} \\ \vdots \\ \tilde{\varphi}_{n,n} \end{pmatrix} = \boldsymbol{A}_n^{-1} \begin{pmatrix} P_0 \\ \vdots \\ P_n \end{pmatrix} = \left(\boldsymbol{A}_n^T \boldsymbol{A}_n \right)^{-1} \begin{pmatrix} \varphi_{n,0} \\ \vdots \\ \varphi_{n,n} \end{pmatrix}.$$

(b) *For a given parameter set $z_r^{(n)}$, $r = 0, 1, \ldots, n-1$, such that $\boldsymbol{B}_n$ is regular, we have*

$$\begin{pmatrix} \psi_{n,0} \\ \vdots \\ \psi_{n,n-1} \end{pmatrix} = \boldsymbol{B}_n^T \begin{pmatrix} P_{n+1} \\ \vdots \\ P_{2n} \end{pmatrix},$$

$$\begin{pmatrix} \tilde{\psi}_{n,0} \\ \vdots \\ \tilde{\psi}_{n,n-1} \end{pmatrix} = \boldsymbol{B}_n^{-1} \begin{pmatrix} P_{n+1} \\ \vdots \\ P_{2n} \end{pmatrix} = \left(\boldsymbol{B}_n^T \boldsymbol{B}_n \right)^{-1} \begin{pmatrix} \psi_{n,0} \\ \vdots \\ \psi_{n,n-1} \end{pmatrix}.$$

Recall that $\boldsymbol{A}_n^T \boldsymbol{A}_n$ and $\boldsymbol{B}_n^T \boldsymbol{B}_n$ are the Gram matrices for our scaling functions and wavelets, respectively.

3.2 Two-Scale Relations and Decomposition

In this section we work out the relationship between the coefficient vectors $\boldsymbol{a}^{(2n)}, \boldsymbol{a}^{(n)}$, and $\boldsymbol{b}^{(n)}$ in the so-called two-scale relation

$$
\begin{aligned}
f_{2n}(t) &= \sum_{r=0}^{2n} a_r^{(2n)} \varphi_{2n}(t; x_r^{(2n+1)}) \\
&= \sum_{r=0}^{n} a_r^{(n)} \varphi_n(t; x_r^{(n+1)}) + \sum_{r=0}^{n-1} b_r^{(n)} \psi_n(t; z_r^{(n)}) \qquad (3.6)\\
&= f_n(t) + g_n(t).
\end{aligned}
$$

In view of (3.1) and (3.3) the above equation may rewritten as follows

$$
\begin{aligned}
&(P_0(t), \ldots, P_{2n}(t))\, \boldsymbol{A}_{2n}\ \boldsymbol{a}^{(2n)} \\
&\qquad = (P_0(t), \ldots, P_n(t))\, \boldsymbol{A}_n\ \boldsymbol{a}^{(n)} + (P_{n+1}(t), \ldots, P_{2n}(t))\, \boldsymbol{B}_n\ \boldsymbol{b}^{(n)},
\end{aligned}
$$

which then implies

$$
\boldsymbol{A}_{2n}\ \boldsymbol{a}^{(2n)} = \begin{pmatrix} \boldsymbol{A}_n & 0 \\ 0 & \boldsymbol{B}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{a}^{(n)} \\ \boldsymbol{b}^{(n)} \end{pmatrix}. \qquad (3.7)
$$

The next theorem shows how to decompose a function from V_{2n} into wavelets from W_n and scaling functions from V_n and states how to reverse this process. The proof follows directly from (3.7).

Theorem 3.16. *Let the scaling functions $\varphi_n(t; x_r^{(n+1)})$, $\varphi_{2n}(t; x_r^{(2n+1)})$, the wavelets $\psi_n(t; z_r^{(n)})$ and the corresponding matrices $\boldsymbol{A}_n$, $\boldsymbol{A}_{2n}$, $\boldsymbol{B}_n$ are based on arbitrary parameter sets.*

(a) (Reconstruction) Let the coefficient vectors $\boldsymbol{a}^{(n)}$ and $\boldsymbol{b}^{(n)}$ in (3.6) be given. Then

$$
\boldsymbol{a}^{(2n)} = \boldsymbol{A}_{2n}^{-1} \begin{pmatrix} \boldsymbol{A}_n & 0 \\ 0 & \boldsymbol{B}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{a}^{(n)} \\ \boldsymbol{b}^{(n)} \end{pmatrix}.
$$

(b) (Decomposition) Let the coefficient vector $\boldsymbol{a}^{(2n)}$ in (3.6) be given. If $\boldsymbol{B}_n$ is regular, then

$$
\begin{pmatrix} \boldsymbol{a}^{(n)} \\ \boldsymbol{b}^{(n)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_n^{-1} & 0 \\ 0 & \boldsymbol{B}_n^{-1} \end{pmatrix} \boldsymbol{A}_{2n} \boldsymbol{a}^{(2n)}.
$$

As it is not surprising, the above formulae simplify in the orthogonal case. In particular, the inversion of matrices can be avoided. Note, however, that the orthogonality of the wavelets is only known for special cases (see Section 5).

Corollary 3.17. *Let the scaling functions $\varphi_{n,r}(t) = \varphi_n(t; x_r^{(n+1)})$, $\varphi_{2n,r}(t) = \varphi_{2n}(t; x_r^{(2n+1)})$, and the wavelets $\psi_{n,r}(t) = \psi_n(t; z_r^{(n)})$ are given.*

(a) *(Reconstruction) Let the coefficient vectors $\boldsymbol{a}^{(n)}$ and $\boldsymbol{b}^{(n)}$ in (3.6) be given. If the $\varphi_{2n,r}$, $r = 0, 1, \ldots, 2n$, are orthogonal, then*

$$a_r^{(2n)} = \frac{1}{\varphi_{2n,r}(x_r^{(2n+1)})} \left(\sum_{s=0}^{n} a_s^{(n)} \varphi_{n,s}(x_r^{(2n+1)}) + \sum_{s=0}^{n-1} b_s^{(n)} \psi_{n,s}(x_r^{(2n+1)}) \right).$$

(b) *(Decomposition) Let the coefficient vector $\boldsymbol{a}^{(2n)}$ in (3.6) be given. If the $\varphi_{n,r}$, $r = 0, 1, \ldots, n$, and the $\psi_{n,r}$, $r = 0, 1, \ldots, n-1$, are orthogonal, then*

$$a_r^{(n)} = \frac{1}{\varphi_{n,r}(x_r^{(n+1)})} \sum_{s=0}^{2n} a_s^{(2n)} \varphi_{n,r}(x_s^{(2n+1)}),$$

$$b_r^{(n)} = \frac{1}{\psi_{n,r}(z_r^{(n+1)})} \sum_{s=0}^{2n} a_s^{(2n)} \psi_{n,r}(x_s^{(2n+1)}).$$

Proof. (a) The orthogonality and (3.6) imply

$$\begin{aligned} a_r^{(2n)} &= \frac{\langle f_n + g_n, \varphi_{2n,r} \rangle}{\langle \varphi_{2n,r}, \varphi_{2n,r} \rangle} \\ &= \frac{1}{\langle \varphi_{2n,r}, \varphi_{2n,r} \rangle} \left(\sum_{s=0}^{n} a_s^{(n)} \langle \varphi_{n,s}, \varphi_{2n,r} \rangle + \sum_{s=0}^{n-1} b_s^{(n)} \langle \psi_{n,s}, \varphi_{2n,r} \rangle \right). \end{aligned}$$

The remaining part follows from the reproducing property of $\varphi_{2n,r}$ (cf. (2.11)). Part (b) is along the same lines. Here we have

$$a_r^{(n)} = \frac{\langle f_{2n}, \varphi_{n,r} \rangle}{\langle \varphi_{n,r}, \varphi_{n,r} \rangle} \quad \text{and} \quad b_r^{(n)} = \frac{\langle f_{2n}, \psi_{n,r} \rangle}{\langle \psi_{n,r}, \psi_{n,r} \rangle}.$$

□

To decompose a given function f one first has to approximate f by a suitable function f_{2n} in V_{2n}. Let us assume that the scaling functions $\varphi_{2n,r}(t) = \varphi_{2n}(t; y_r^{(2n+1)})$ are based on the zeros $y_r^{(2n+1)}$ of P_{2n+1}, *i.e.*, they are orthogonal

$$\langle \varphi_{2n,r}, \varphi_{2n,s} \rangle = \varphi_{2n,r}(y_s^{(2n+1)}) = c_r^{(2n+1)} \delta_{r,s}.$$

Then the approximation is typically done by an orthogonal projection

$$f(t) \approx \sum_{r=0}^{2n} \langle f, \varphi_{2n,r} \rangle \, \frac{\varphi_{2n,r}(t)}{\langle \varphi_{2n,r}, \varphi_{2n,r} \rangle},$$

or by an interpolatory process

$$f(t) \approx \sum_{r=0}^{2n} f(y_r^{(2n+1)}) \, \frac{\varphi_{2n,r}(t)}{\langle \varphi_{2n,r}, \varphi_{2n,r} \rangle}.$$

Actually, if on computes $\langle f, \varphi_{2n,r}\rangle$ by the Gaussian quadrature both approaches provide the same approximation. The proof of the next lemma follows directly from Corollary 2.6.

Lemma 3.18. *Let $y_r^{(2n+1)}$, $r = 0, 1, \ldots, 2n$, denote the zeros of P_{2n+1} and let $\varphi_{2n,r}$ denote the associated scaling functions (2.13). Furthermore, let f denote a given smooth function. Then the Gaussian quadrature of $f\varphi_{2n,r}$ simplifies*

$$\langle f, \varphi_{2n,r}\rangle \approx \sum_{s=0}^{2n} \left(c_s^{(2n+1)}\right)^{-1} f(y_s^{(2n+1)})\varphi_{2n,r}(y_s^{(2n+1)}) = f(y_r^{(2n+1)}).$$

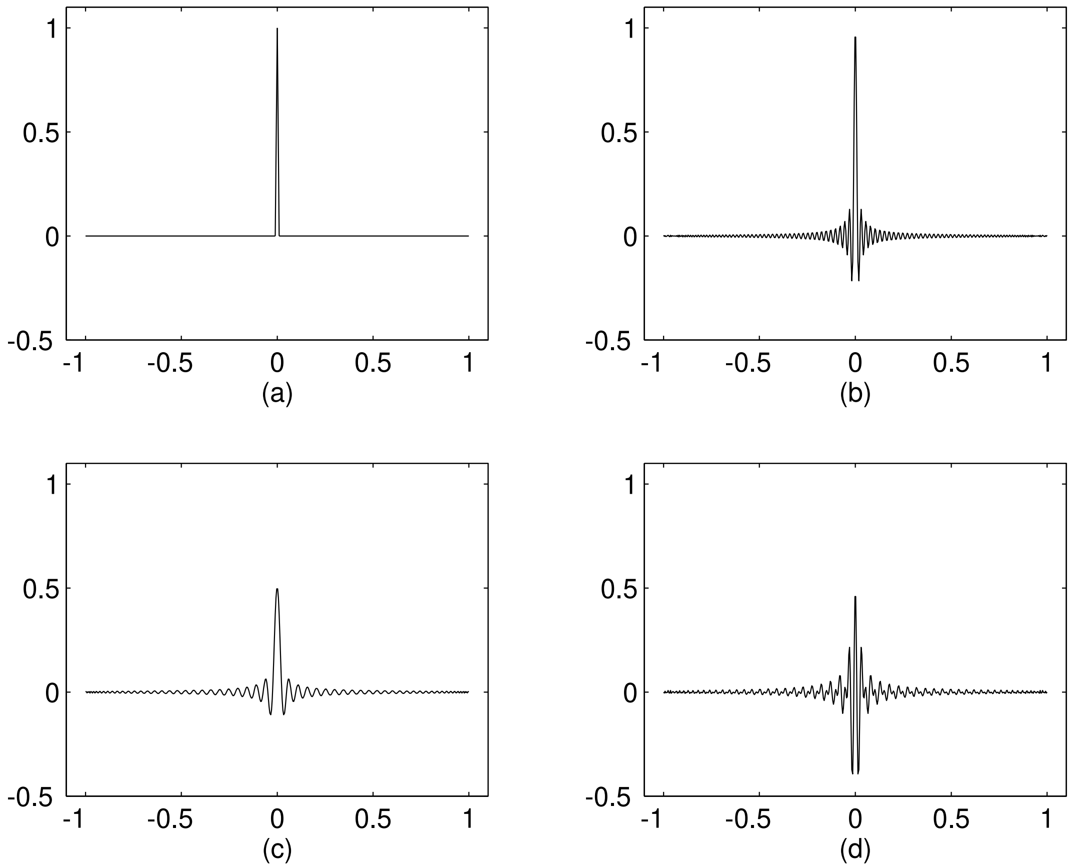

Figure 3.19. Decomposition with respect to $w(t) = (1 - t^2)^{-1/2}$ for $n = 128$. (a) Given function: f; (b) projection on V_{2n}: f_{2n}; decomposition: (c) $f_n \in V_n$; (d) $g_n \in W_n$.

Let us finish this section with an example. Here we decompose a piecewise linear “hat function” f which is zero on $[-1, 1] \setminus (-0.01, 0.01)$ and one at the origin (compare Figure 3.19(a)). The scaling function spaces V_{2n} and V_n were defined by the zeros of P_{2n+1} and P_{n+1}, respectively. The wavelet space W_n was defined by the zeros of P_n, which ensures that the $\psi_{n,r}$'s constitute a basis. The approximation f_{2n} of f in V_{2n} was computed by the above described interpolatory process.

It is important to note that the underlying numerical computations make use of the properties of orthogonal polynomials. In particular, we computed the corresponding parameter sets as eigenvalues of the associated Jacobi matrix (cf. Lemma 2.1(c)) and the resulting polynomials were evaluated by means of their three-term recurrence relations (2.5).

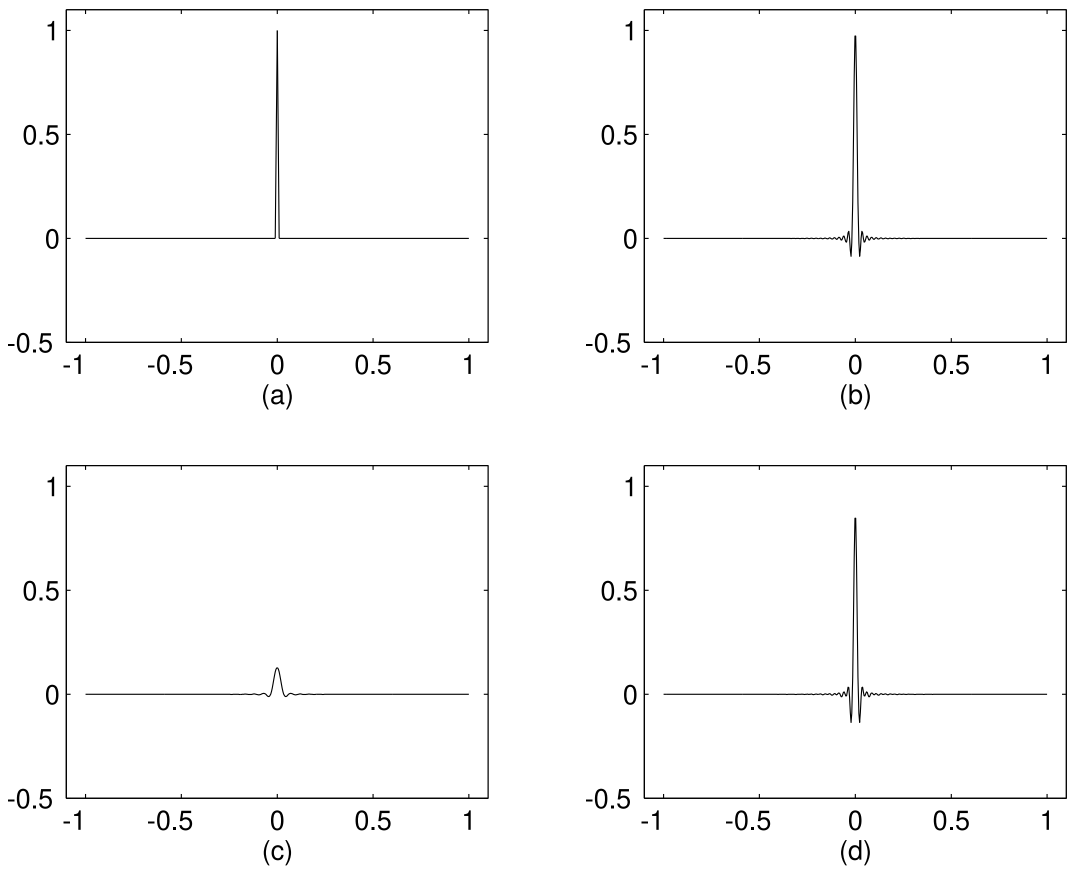

Figure 3.20. Decomposition with respect to $w(t) = t^2(1-t^2)^{-1/2}$ for $n = 128$. (a) Given function: f; (b) projection on V_{2n}: f_{2n}; decomposition: (c) $f_n \in V_n$; (d) $g_n \in W_n$.

Figure 3.19 shows the decomposition $f_{2n} = f_n + g_n$ with respect to the Chebyshev weight function of the first kind $w(t) = (1-t^2)^{-1/2}$. Whereas Figure 3.20 shows the same decomposition but with respect to the modified weight function $w(t) = t^2(1-t^2)^{-1/2}$ (explicit expressions for the corresponding orthogonal polynomials may be found in Chihara [1, pp. 155]). Here, the time localization is considerably improved.

Some comments are in order. The purpose of the example is to show that the choice of the weight function may have quite some affect on the decomposition. Here, we designed the given function f such that in both cases the approximation f_{2n} consists of only one scaling function, that is, f_{2n} is "maximally localized" with respect to the chosen weight function (cf. Theorem 2.3(b)). It is interesting to note that for both weights g_n has a full expansion into wavelets, *i.e.*, $b_r^{(n)} \neq 0,\ r = 0, 1, \ldots, n-1$.

4. Stability and Translation Invariance

The above introduced scaling functions and wavelets do not provide a multiresolution in the classical sense. However, there are some relationships which will be pointed out in this section.

4.1 Riesz Stability

We start with a result on Riesz stability. Here we establish a two-sided estimate between the weighted L_2-norm $\|f_n\|$ ($\|g_n\|$) (cf. (2.2)) of an arbitrary function $f_n \in V_n$ ($g_n \in W_n$) and the Euclidian norm of the coefficients of f_n (g_n) with respect to the basis of scaling functions (wavelets). The Euclidian norm of a vector $\boldsymbol{a}^{(n)} \in \mathbb{R}^{n+1}$ is defined as usual by $\|\boldsymbol{a}^{(n)}\|_2 = \left(\sum_{r=0}^{n} a_r^2\right)^{1/2}$ with corresponding spectral norm $\|\boldsymbol{A}\|_2$.

Theorem 4.21. *Let $\boldsymbol{A}_n$ (cf. (3.2)) and $\boldsymbol{B}_n$ (cf. (3.4)) denote the matrices associated with the parameter sets $x_r^{(n+1)}$ and $z_r^{(n)}$, respectively. Furthermore, let $\varphi_{n,r}$ and $\psi_{n,r}$ denote the corresponding scaling functions and wavelets.*

(a) For $f_n = \sum_{r=0}^{n} a_r^{(n)} \varphi_{n,r}$, we have

$$\frac{1}{\|\boldsymbol{A}_n^{-1}\|_2}\, \|\boldsymbol{a}^{(n)}\|_2 \;\leq\; \|f_n\| \;\leq\; \|\boldsymbol{A}_n\|_2\, \|\boldsymbol{a}^{(n)}\|_2.$$

(b) For $g_n = \sum_{r=0}^{n-1} b_r^{(n)} \psi_{n,r}$, we have

$$\|g_n\| \;\leq\; \|\boldsymbol{B}_n\|_2\, \|\boldsymbol{b}^{(n)}\|_2,$$

and if in addition $\boldsymbol{B}_n$ is regular, then also

$$\frac{1}{\|\boldsymbol{B}_n^{-1}\|_2}\, \|\boldsymbol{b}^{(n)}\|_2 \;\leq\; \|g_n\|.$$

Proof. By Parseval's equation we obtain for $f_n \in V_n$

$$\|f_n\|^2 = \left\| \sum_{r=0}^{n} a_r^{(n)} \sum_{k=0}^{n} P_k(x_r^{(n+1)}) P_k \right\|^2 = \sum_{k=0}^{n} \left| \sum_{r=0}^{n} a_r^{(n)} P_k(x_r^{(n+1)}) \right|^2 = \|\boldsymbol{A}_n \boldsymbol{a}^{(n)}\|_2^2.$$

Now (a) and analogously (b) follow by standard arguments. □

Hence, as it is not surprising, the Riesz stability can be measured by the spectral condition number of $\boldsymbol{A}_n$

$$\|\boldsymbol{A}_n^{-1}\|_2 \|\boldsymbol{A}_n\|_2 = \sqrt{\frac{\lambda_{\max}(\boldsymbol{A}_n^{\mathrm{T}}\boldsymbol{A}_n)}{\lambda_{\min}(\boldsymbol{A}_n^{\mathrm{T}}\boldsymbol{A}_n)}} \tag{4.1}$$

and by the spectral condition number of $\boldsymbol{B}_n$

$$\|\boldsymbol{B}_n^{-1}\|_2 \|\boldsymbol{B}_n\|_2 = \sqrt{\frac{\lambda_{\max}(\boldsymbol{B}_n^{\mathrm{T}}\boldsymbol{B}_n)}{\lambda_{\min}(\boldsymbol{B}_n^{\mathrm{T}}\boldsymbol{B}_n)}}\ , \tag{4.2}$$

respectively. Here $\lambda_{\max}$ and $\lambda_{\min}$ denote the extreme eigenvalues of the corresponding matrices.

4.2 Generalized Translation

Usually, in a multiresolution analysis time localization is realized by taking shifts of one given function. Also, Euler's functional equation is used to advantage. Namely, a shift in the time space is equivalent to a multiplication by an exponential in the Fourier space. In this section, we briefly outline how to generalize this concept to the present polynomial approach. Here, we restrict ourselves to the Jacobi polynomials $P_k^{(\alpha,\beta)}$. These polynomials are orthogonal with respect to the weight $w(t) = (1-t)^\alpha(1+t)^\beta$, $-1 < t < 1$.

For a given f in this weighted L_2-space with Fourier-Jacobi-coefficients

$$f\,\hat{}(k) = \int_{-1}^{1} f(t) P_k^{(\alpha,\beta)}(t)(1-t)^\alpha(1+t)^\beta dt$$

we consider the operator $S_\lambda : L_2(w) \to L_2(w)$, $-1 \le \lambda \le 1$, defined by a multiplication in the frequency domain

$$(S_\lambda f)\,\hat{}(k) := \frac{P_k^{(\alpha,\beta)}(\lambda)}{P_k^{(\alpha,\beta)}(1)} f\,\hat{}(k)\ .$$

For $-1 < \beta \le \alpha,\ -1 \le \alpha+\beta$, the operator S_λ has the properties (see Gasper [5])

$$\|S_\lambda f\| \le C\|f\|, \quad \text{for all} \quad \lambda \in (-1,1),$$

and

$$\lim_{\lambda\to 1-} \|S_\lambda f - f\| = 0.$$

Hence S_λ may be seen as a generalized translation operator.

In this context, it is possible to recover our scaling functions and wavelets, respectively, as generalized translations of a given function. More precisely, with

$$f_V(t) := \varphi_n(t;1) = \sum_{k=0}^{n} P_k^{(\alpha,\beta)}(1) P_k^{(\alpha,\beta)}(t)$$

it is straightforward to verify that

$$\varphi_n(\cdot;\lambda) = S_\lambda f_V .$$

Here, we used that

(4.3) $$f_V\,\check{}(k) = \begin{cases} P_k^{(\alpha,\beta)}(1) & \text{for } 0 \le k \le n , \\ 0 & \text{for } k > n . \end{cases}$$

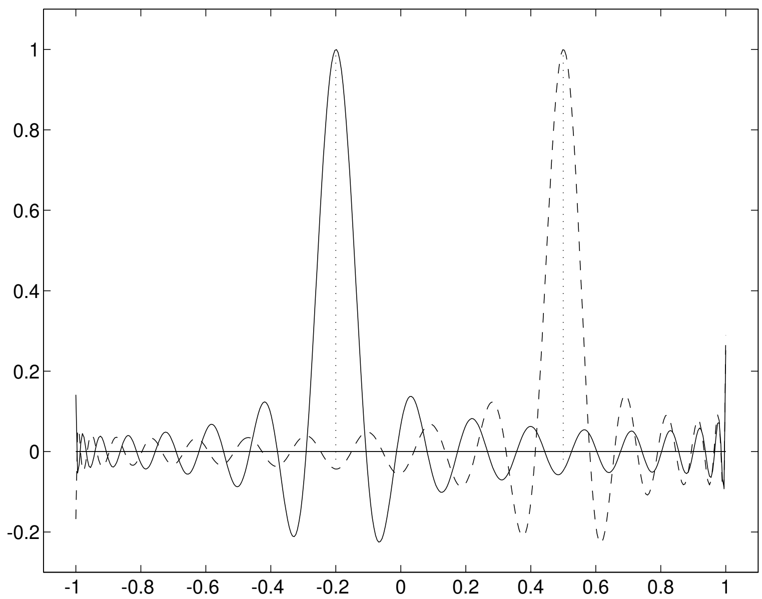

Figure 4.22. Scaling functions of degree $n = 32$ with respect to the Legendre weight $w(t) = 1$ and the parameter $x_0^{(n+1)} = -0.2$ (solid line) and $x_1^{(n+1)} = 0.5$ (dashed line).

Analogously, we have

$$\psi_n(\cdot;\lambda) = S_\lambda f_W$$

with

(4.4) $$f_W(t) := \psi_n(t;1) = \sum_{k=n+1}^{2n} P_k^{(\alpha,\beta)}(1) P_k^{(\alpha,\beta)}(t).$$

We would like to mention that modifications of (4.3) and of (4.4) which at least remain supp$f_V\,\hat{}=\{0,\ldots,n\}$ and supp$f_W\,\hat{}=\{n+1,\ldots,2n\}$, respectively, do affect the algorithms of Section 3 only by the multiplication of $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ by certain regular diagonal matrices.

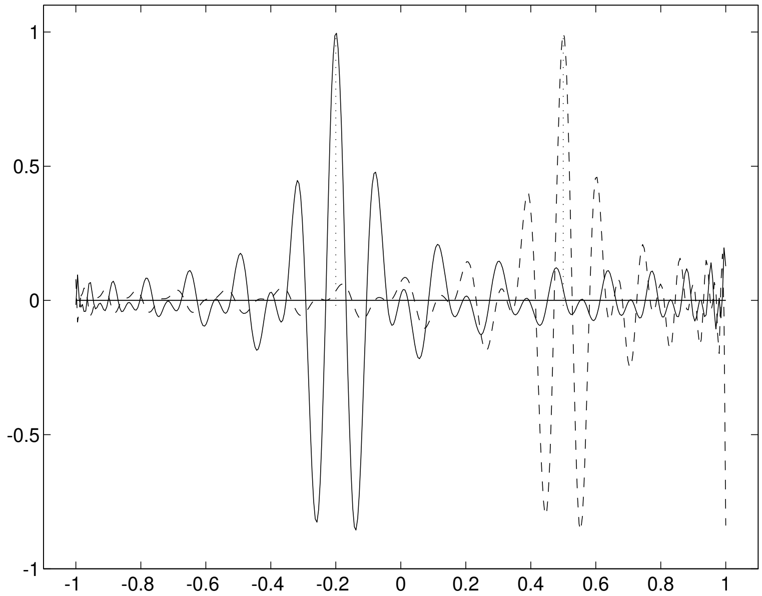

Figure 4.23. Wavelets of degree $n=32$ with respect to the Legendre weight $w(t)=1$ and the parameter $z_0^{(n+1)}=-0.2$ (solid line) and $z_1^{(n+1)}=0.5$ (dashed line).

The two figures illustrate, that scaling functions (wavelets) with respect to different parameters look almost like a shift of each other.

5. Examples

In this section we want to discuss two examples in more detail. Both belong to the class of Chebyshev weights, i.e. Jacobi weights with $|\alpha|=|\beta|=\frac{1}{2}$. These weights are of particular interest, because here one can handle the computations with the help of fast algorithms based on the Discrete Cosine Transform (see, for example Tasche et. al. [9], [10], [7]).

Let us start with the Chebyshev weight of the first kind

$$w(t)=\frac{1}{\sqrt{1-t^2}}, \quad t\in(-1,1).$$

The corresponding orthonormal polynomials, the Chebyshev polynomials of the first kind, can conveniently be written

$$P_n(t)=\sqrt{\frac{1}{\pi}}\cdot\begin{cases}\sqrt{2}\cos n\theta & \text{if } n>0\,,\\ 1 & \text{if } n=0\,,\end{cases}$$

in terms of $t = \cos\theta$, $0 \le \theta \le \pi$. If we take as parameter set for the scaling functions the zeros of P_{n+1} (cf. Corollary 2.6)

$$y_r^{(n+1)} = \cos\frac{(2r+1)\pi}{2n+2}, \quad r = 0, \ldots, n,$$

then

$$\varphi_{n,r}(t) = \frac{1}{\pi} + \frac{2}{\pi}\sum_{k=0}^{n} \cos\frac{k(2r+1)\pi}{2n+2}\cos k\theta$$

and

$$\boldsymbol{A}_n = \sqrt{\frac{2}{\pi}}\,\mathrm{diag}\left(\frac{\sqrt{2}}{2}, 1, 1, \ldots, 1\right) \cdot \left(\cos\frac{k(2r+1)\pi}{2n+2}\right)_{k,r=0,\ldots,n}.$$

With the help of trigonometric identities it is easy to see that

$$\boldsymbol{A}_n^{-1} = \frac{\pi}{n+1}\boldsymbol{A}_n^{\mathrm{T}}. \tag{5.1}$$

Following Corollary 3.12 we notice that this is the only situation where Gaussian quadrature coincides with Chebyshev quadrature (i.e., all weights are equal $\pi/(n+1)$).

Analogously we choose for the wavelets the zeros $y_r^{(n)}$ of P_n as the set of parameter. Then

$$\psi_{n,r}(t) = \frac{2}{\pi}\sum_{k=n+1}^{2n} \cos\frac{k(2r+1)\pi}{2n}\cos k\theta, \quad r = 0, \ldots, n-1,$$

and

$$\begin{aligned} \boldsymbol{B}_n &= \sqrt{\frac{2}{\pi}}\left(\cos\frac{(n+1+k)(2r+1)\pi}{2n}\right)_{k,r=0,\ldots,n-1} \\ &= \sqrt{\frac{2}{\pi}}\left((-1)^{r+1}\sin\frac{(k+1)(2r+1)\pi}{2n}\right)_{k,r=0,\ldots,n-1}. \end{aligned}$$

In this case we know from Corollary 3.14 that the wavelets are linear independent. However, the following lemma shows that they are not orthogonal to each other (compare Corollary 3.13(b)).

Lemma 5.24. *The inverse of $\boldsymbol{B}_n$ is given by*

$$\boldsymbol{B}_n^{-1} = \frac{\pi}{n}\boldsymbol{B}_n^{\mathrm{T}} \cdot \mathrm{diag}\left(1, 1, \ldots, 1, \frac{1}{2}\right). \tag{5.2}$$

with

$$\boldsymbol{B}_n^{\mathrm{T}}\boldsymbol{B}_n = \left(\frac{n}{\pi}\delta_{k,r} + \frac{1}{\pi}\right)_{k,r=0,\ldots,n-1}.$$

Proof. The first assertion is equivalent to

$$\delta_{r,s} = \frac{2}{n} \sum_{k=0}^{n-1}{}' (-1)^{r+s} \sin \frac{(k+1)(2r+1)\pi}{2n} \sin \frac{(k+1)(2s+1)\pi}{2n},$$

where the prime indicates that the last term in the sum has to be divided by 2. Having performed an index shift, we obtain for the right-hand side by the addition formula

$$\frac{1}{n}(-1)^{r+s} \sum_{k=1}^{n}{}' \cos \frac{k(r-s)\pi}{n} - \cos \frac{k(r+s+1)\pi}{n}.$$

The rest of the proof is just an application of the well-known summation formula for Dirichlet kernels

$$\frac{1}{2} + \sum_{k=1}^{n}{}' \cos \frac{k\ell\pi}{n} = \begin{cases} 0 & \text{if } |\ell| = 1, \dots, 2n-1, \\ n & \text{if } \ell = 0. \end{cases} \tag{5.3}$$

The second statement follows directly from the first one. ∎

We conclude this example with the computation of the Riesz stability constants (cf. (4.1), (4.2)). From (5.1) we obtain the best possible bounds for the scaling functions

$$\|\boldsymbol{A}_n^{-1}\|_2 \|\boldsymbol{A}_n\|_2 = 1,$$

whereas for the wavelets we deduce from (5.2) that

$$\|\boldsymbol{B}_n^{-1}\|_2 \|\boldsymbol{B}_n\|_2 = \sqrt{2}.$$

Let us now consider as a second example the Chebyshev weight of the second kind, i.e.,

$$w(t) = \sqrt{1-t^2}, \quad t \in (-1,1),$$

and the corresponding orthonormal polynomials

$$P_n(t) = \sqrt{\frac{2}{\pi}} \frac{\sin(n+1)\theta}{\sin\theta}.$$

Again we take as parameter set for the scaling functions the zeros of P_{n+1}, i.e.,

$$y_r^{(n+1)} = \cos \frac{(r+1)\pi}{n+2}, \quad r = 0, \dots, n.$$

This choice leads to

$$\varphi_{n,r}(t) = \frac{2}{\pi} \sum_{k=0}^{n} \frac{\sin \frac{(k+1)(r+1)\pi}{n+2}}{\sin \frac{(r+1)\pi}{n+2}} \frac{\sin(k+1)\theta}{\sin\theta}$$

and

$$\boldsymbol{A}_n = \sqrt{\frac{2}{\pi}} \left(\frac{\sin \frac{(k+1)(r+1)\pi}{n+2}}{\sin \frac{(r+1)\pi}{n+2}} \right)_{k,r=0,\ldots,n} .$$

For completeness we mention the result on the Gaussian quadrature

$$\boldsymbol{A}_n^{\mathrm{T}} \boldsymbol{A}_n = \operatorname{diag} \left(\frac{n+2}{\pi \sin^2 \frac{(r+1)\pi}{n+2}} \right)_{r=0,\ldots,n} . \tag{5.4}$$

Choosing the zeros $y_r^{(n)}$ of P_n as the set of parameter for the wavelets we obtain

$$\psi_{n,r}(t) = \frac{2}{\pi} \sum_{k=n+1}^{2n} \frac{\sin \frac{(k+1)(r+1)\pi}{n+1}}{\sin \frac{(r+1)\pi}{n+1}} \frac{\sin(k+1)\theta}{\sin \theta}, \qquad r = 0, \ldots, n-1,$$

and

$$\boldsymbol{B}_n = \sqrt{\frac{2}{\pi}} \left(\frac{\sin \frac{(k+n+2)(r+1)\pi}{n+1}}{\sin \frac{(r+1)\pi}{n+1}} \right)_{k,r=0,\ldots,n-1} .$$

The next lemma shows that this time we have an orthogonal set of wavelets.

Lemma 5.25. *For the above defined matrix $\boldsymbol{B}_n$ we have*

$$\boldsymbol{B}_n^{\mathrm{T}} \boldsymbol{B}_n = \mathit{diag} \left(\frac{n+1}{\pi \sin^2 \frac{(r+1)\pi}{n+1}} \right)_{r=0,\ldots,n-1} . \tag{5.5}$$

Proof. The (r,s) element of $\boldsymbol{B}_n^{\mathrm{T}} \boldsymbol{B}_n$ looks like

$$\frac{2}{\pi} \sum_{k=0}^{n-1} \frac{\sin \frac{(k+n+2)(r+1)\pi}{n+1}}{\sin \frac{(r+1)\pi}{n+1}} \cdot \frac{\sin \frac{(k+n+2)(s+1)\pi}{n+1}}{\sin \frac{(s+1)\pi}{n+1}} .$$

It may be simplified to

$$\frac{(-1)^{r+s}}{\pi \sin \frac{(r+1)\pi}{n+1} \sin \frac{(s+1)\pi}{n+1}} \sum_{k=1}^{n} \left(\cos \frac{k(r-s)\pi}{n+1} - \cos \frac{k(r+s+2)\pi}{n+1} \right) .$$

Now the statement follows from (5.3). ∎

Again, we finish by computing the quotient for the Riesz bounds. Here, we obtain from (5.4) that

$$\|\boldsymbol{A}_n^{-1}\|_2 \|\boldsymbol{A}_n\|_2 = \begin{cases} \sin^{-1} \frac{\pi}{n+2} & \text{for even } n, \\ \frac{1}{2} \sin^{-1} \frac{\pi}{2n+4} & \text{for odd } n, \end{cases}$$

and from (5.5) that

$$\|\boldsymbol{B}_n^{-1}\|_2\|\boldsymbol{B}_n\|_2 = \begin{cases} \sin^{-1}\frac{\pi}{n+1} & \text{for odd } n, \\ \frac{1}{2}\sin^{-1}\frac{\pi}{2n+2} & \text{for even } n. \end{cases}$$

Let us summarize these two examples. We have constructed for the Chebyshev weight of the first kind orthogonal scaling functions and nonorthogonal wavelets where the quotient of the associated Riesz bounds are uniformly bounded. On the other hand, for the Chebyshev weight of the second kind we constructed both orthogonal scaling functions and orthogonal wavelets where the quotient of the Riesz constants are growing linearly in n.

Acknowledgement

Part of this work was done while the second author was visiting the Institute of Mathematics at the Medical University of Lübeck. He is grateful for the support and the warm hospitality at this institute.

References

[1] T. S. Chihara. *An introduction to orthogonal polynomials.* Gordon and Breach, New York, London, Paris, 1978.

[2] C.K. Chui and H.N. Maskar. On trigonometric wavelets. *Constr. Approx.*, 9:167–190, 1993.

[3] P. J. Davis. *Interpolation & Approximation.* Blaisdell, Waltham, Massachusetts, 1963.

[4] B. Fischer. *Polynomial Based Iteration Methods for Symmetric Linear Systems.* Teubner, Stuttgart, 1996.

[5] G. Gasper. Banach algebras for Jacobi series and positivity of a kernel. *Ann. Math.*, 95:261–280, 1972.

[6] T. Kilgore and J. Prestin. Polynomial wavelets on the interval. Preprint 94/3, FB Mathematik, Universität Rostock, 1994. To appear in Constr. Approx.

[7] G. Plonka, K. Selig, and M. Tasche. On the construction of wavelets on an bounded interval. Preprint 94/22, FB Mathematik, Universität Rostock, 1994. To appear in Adv. Comp. Math.

[8] G. Szegö. *Orthogonal polynomials.* AMS Colloquium Publications XXIII, American Mathematical Society, New York, revised edition, 1959.

[9] M. Tasche. Fast algorithms for discrete Chebyshev - Vandermonde transforms and applications. *Numer. Alg.*, 5:453–464, 1993.

[10] M. Tasche. Polynomial wavelets on $[-1, 1]$. In S. P. Singh, editor, *Approximation Theory, Wavelets and Applications*, pages 497–512, Dordrecht, 1995. Kluwer Academic Publ.