Centrum voor Wiskunde en Informatica

REPORT RAPPORT

The Fast Wavelet X-Ray Transform

R.A. Zuidwijk, P.M. de Zeeuw

Probability, Networks and Algorithms (PNA)

PNA-R9908 September 30, 1999

Report PNA-R9908
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

The Fast Wavelet X-Ray Transform

R.A. Zuidwijk
EUR
Rotterdam School of Management
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
R.Zuidwijk@fac.fbk.eur.nl

P.M. de Zeeuw
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Paul.de.Zeeuw@cwi.nl

ABSTRACT
The wavelet X-ray transform computes one-dimensional wavelet transforms along lines in Euclidian space in order to perform a directional time-scale analysis of functions in several variables. A fast algorithm is proposed which executes this transformation starting with values given on a cartesian grid that represent the underlying function. The algorithm involves a rotation step and wavelet analysis/synthesis steps. The number of computations required is of the same order as the number of data involved. The analysis/synthesis steps are executed by the pyramid algorithm which is known to have this computational advantage. The rotation step makes use of a wavelet interpolation scheme. The order of computations is limited here due to the localization of the wavelets. The rotation step is executed in an optimal way by means of quasi-interpolation methods using (bi-)orthogonal wavelets.

1991 Mathematics Subject Classification: 42C15, 44A12, 86A15
1991 ACM Computing Classification System: G.1.1, J.2
Keywords and Phrases: wavelet transform, X-ray transform, windowed X-Ray transform, local X-Ray transform, grid function, quasi-interpolation, (bi-)orthogonal wavelets.
Note: Work carried out under project PNA 4.2 "Wavelets: Analysis of Seismic Signals" and supported financially by the Technology Foundation (STW), project no. CWI44.3403.

1. Introduction

In this paper, we present results related to the fast wavelet X-ray transform, which performs a directional time-scale analysis of functions in several variables. Details concerning the wavelet transform and the X-ray transform, and combined use of these two transforms, is given in the second part of the introduction. First, we describe the results of this paper in general terms.

We start with a square integrable function $f \in L^2(\mathbb{R}^n)$ for which values are given on the standard grid $\mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n$ (here $e_1, \ldots, e_n$ is the standard orthonormal basis in $\mathbb{R}^n$). The wavelet X-ray coefficients

$$P_\psi f(\theta, x, b, a) = \int_{\mathbb{R}} f(x + t\theta) \frac{1}{\sqrt{a}} \overline{\psi\left(\frac{t-b}{a}\right)}\, dt$$

are computed by means of a uni-variate wavelet ψ. These coefficients are the result of one-dimensional wavelet transforms along lines $x + \mathbb{R}\theta$ which are parameterized by unit vectors θ and by vectors $x \perp \theta$. The dilation parameter $a > 0$ and the translation parameter $b \in \mathbb{R}$ arise from the wavelet transform. In order to compute these coefficients $P_\psi f(\theta, x, b, a)$ by means of a fast wavelet algorithm, we need grid values of the function f on these rotated lines $x + \mathbb{R}\theta$. We remark that the one-dimensional fast wavelet transform computes coefficients for translation-dilation pairs (b, a) from the countable set $\{(k2^{-j}, 2^{-j}) \mid j, k \in \mathbb{Z}\}$. In order to compute the fast wavelet transform along equi-spaced parallel lines, we actually need grid values of the function on a rotated and dilated grid.

Such a rotated and dilated grid can be defined using an orthonormal basis $\theta_1, \ldots, \theta_n$ and positive dilation factors $d_1, \ldots, d_n$. The rotated and dilated grid can be written as $d_1\mathbb{Z}\theta_1 + \cdots + d_n\mathbb{Z}\theta_n$. In Section 5, such a grid will be denoted by $\mathbb{G}_{\theta,d}$, where θ represents the orthonormal basis $\theta_1, \ldots, \theta_n$ instead of a single unit vector. In Section 5, some of the main results of this paper are proved. Indeed, Theorems 5.5 and 5.6 state that if $f \in H^{p+1}(\mathbb{R}^n)$ is an arbitrary function with partial derivatives up to order $p+1$ in $L^2(\mathbb{R}^n)$, and if g denotes the function arising from interpolation on the grid $\mathbb{G}_{\theta,d}$ (details are given in Section 5), then

$$\|f - g\|_{L^2(\mathbb{R}^n)} \leq C\,\beta^n \sum_{r=1}^{n} d_r^{p+1} \|\partial_r^{p+1} f\|_{L^2(\mathbb{R}^n)},$$

where C and β are positive constants which do not depend on f, d and p. In particular, the theorems state that the approximation of f by the interpolating function g improves as the mesh d of the grid becomes finer. This theorem requires some known approximation results as described in Section 3 and a result on bi-orthogonal Riesz systems given in Section 2. In Section 2, a number of new results are reported concerning the construction of bi-orthogonal systems of Riesz bases. These results are inspired by the papers [Uns, Uns2, UA].

The approximation of functions by interpolating functions on finer and finer grids is formalized in terms of multi-resolution analysis. The well-known theory on this matter is described in Section 4 and applied to rotated and dilated grids in Section 6. In the latter section, it is also shown that the interpolation method, called the *dual method*, only requires a number of computations proportional to the number of data involved.

We shall now present a concise review on wavelet and X-ray transforms. The wavelet transform is by now an established mathematical tool to perform a time-scale analysis of functions [Dau, Hol, Mey, LMR]. In this approach, a function ψ is fixed which plays the role of the *wavelet*. The wavelet should satisfy certain technical conditions which we will mention later on. Given $(b, a) \in \mathbb{H} = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$, a point in the open upper half plane, we consider a shifted and dilated version

$$\psi_{b,a}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$$

of the wavelet. Observe that the function $\psi_{b,a}$ is centered around $b \in \mathbb{R}$ and is dilated by the positive factor $a > 0$. The normalization constant is chosen in such a way that the energy norm is preserved, i.e.,

$$\|\psi_{b,a}\|_{L^2(\mathbb{R})} = \|\psi\|_{L^2(\mathbb{R})} = \sqrt{\int_{\mathbb{R}} |\psi(t)|^2\, dt},$$

where $L^2(\mathbb{R})$ is the Hilbert space consisting of functions for which the Lebesgue integral on the right hand side exists and is finite. For $(b, a) \in \mathbb{H}$, we consider the *wavelet coefficients*

$$W_\psi f(b, a) = \langle f, \psi_{b,a}\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} f(t)\overline{\psi_{b,a}(t)}\, dt.$$

Typically, the function ψ has compact support or is well-localized, e.g. has exponential decay outside a compact interval. In that case, the wavelet coefficient $W_\psi f(b, a)$ probes the function f around $b \in \mathbb{R}$ at scale $a > 0$. In this manner, a time-scale analysis is performed using the wavelet transform $f \mapsto W_\psi f$.

Such an analysis becomes particularly useful in the case when the original function can be reconstructed from its wavelet coefficients. The so-called *continuous wavelet transform* involves a reconstruction of the function from *all* wavelet coefficients $\{W_\psi f(b,a) \mid (b,a) \in \mathbb{H}\}$. In order to achieve this reconstruction, an admissibility condition is put upon the wavelet ψ which in quite a few cases boils down to the requirement that the wavelet has a vanishing mean, i.e., $\int_{\mathbb{R}} \psi(t)\,dt = 0$. For more details, see [Koo, LMR].

We could also try to reconstruct the function from only a countable set of its wavelet coefficients. It turns out that this is possible. One simply requires that a countable set of shifted and dilated versions of the wavelet constitutes an orthonormal basis in $L^2(\mathbb{R})$. The notion of orthonormal basis can be relaxed to the notion of Riesz (stable) basis, as explained in Section 2. However, an explicit construction of such wavelets is far from trivial, and these achievements (see e.g. [Chu, CDF, Dau2]) are keys to the present popularity of wavelet methods. In most cases, the aforementioned countable subset of the open upper half plane $\mathbb{H}$ is given by $K = \{(k2^j, 2^j) \mid j,k \in \mathbb{Z}\}$.

The *pyramid scheme* [Mal2] has made wavelet methods attractive for a wide range of applications such as seismology, diagnostic medicine and telecommunications. This algorithm, which is also known as the *fast wavelet transform*, computes wavelet coefficients of a function which is given by a finite dataset. The number of computations is of the same (asymptotic) order as the number of data involved. Since only a finite number of scales are used (e.g., the function is sampled at a finite rate), not only a wavelet ψ is required in the analysis, but also a *scaling function* φ. This scaling function typically satisfies $\int_{\mathbb{R}} \varphi(t)\,dt = 1$. The dataset $(a_k^{(0)})_k$ is usually interpreted as $f = \sum_k a_k^{(0)} \varphi(\cdot - k)$. It is assumed that the wavelet ψ and the scaling function φ satisfy *two-scale relations*

$$\varphi(x) = \sum_k h_k \sqrt{2}\varphi(2x-k), \quad \psi(x) = \sum_k g_k \sqrt{2}\varphi(2x-k), \quad \text{almost all } x \in \mathbb{R}. \tag{1.1}$$

One constructs φ in such a way that a countable set of its translates and dilates forms an orthononormal or Riesz basis, and ψ is derived from φ. In Section 4, some details are given in the context of multiresolution analysis.

To give an idea of how the algorithm works, we introduce the wavelet- and *scaling function coefficients* as ($j \geq 0$, $k \in \mathbb{Z}$)

$$d_k^{(j)} = \langle f, 2^{-j/2}\psi(2^{-j}\cdot - k)\rangle_{L^2(\mathbb{R})} = W_\psi f(k2^j, 2^j), \quad a_k^{(j)} = \langle f, 2^{-j/2}\varphi(2^{-j}\cdot - k)\rangle_{L^2(\mathbb{R})},$$

respectively. Observe that substituting $x \to 2^j x - l$ in (1.1) and taking the inner product of the result with $f \in L^2(\mathbb{R})$ provides

$$a_l^{(j)} = \sum_k h_k a_{2l+k}^{(j-1)}, \quad d_l^{(j)} = \sum_k g_k a_{2l+k}^{(j-1)}, \quad j \in \mathbb{Z}.$$

These formulas provide an efficient means of computing the wavelet coefficients of a function for a finite range of scales. Observe that the wavelet and scaling function coefficients at the coarse scale 2^j can be computed very efficiently by means of the scaling function coefficients corresponding to the finer scale 2^{j-1}. The computations involve the sequences $(g_k)_{k\in\mathbb{Z}}$ and $(h_k)_{k\in\mathbb{Z}}$. These coefficients are called *filter coefficients*, since the convolution with these coefficients allows for a fruitful interpretation as filter operations. In this terminology, $(g_k)_{k\in\mathbb{Z}}$ gives rise to a *high-pass filter* and $(h_k)_{k\in\mathbb{Z}}$ to a *low-pass filter*. We remark that for most practical applications, the sequences of filter coefficients are finite. Indeed, wavelets and scaling functions with compact support give rise to filters with finitely many non-vanishing coefficients. An introduction to wavelet filter bank theory is provided by [SN].

In this paper, we consider wavelet analysis of functions in several variables. However, all wavelet methods here are based on uni-variate wavelets. We consider the wavelet X-ray transform, which performs wavelet transforms along lines in $\mathbb{R}^n$. These lines are parameterized in the same fashion as for the X-ray transform. The *X-ray transform* [Nat]

$$Pf(\theta, x) = \int_{\mathbb{R}} f(x + t\theta)\,dt$$

integrates a function f on $\mathbb{R}^n$ along an affine line $x + \mathbb{R}\theta$, where $x \in \mathbb{R}^n$ is perpendicular to the direction θ. Observe that (θ, x), where θ is a unit vector and x a vector orthogonal to θ, parameterize all lines in $\mathbb{R}^n$. In particular, the distance of the line $x + \mathbb{R}\theta$ to the origin is given by $\|x\|$. This transformation (as the name suggests) has led to fruitful applications in diagnostic medicine [SSW].

Combined use of the X-Ray transform and the wavelet transform has also received attention. In particular, the *wavelet X-ray transform*

$$P_\psi f(\theta, x, b, a) = \int_{\mathbb{R}} f(x + t\theta)\overline{\psi_{b,a}(t)}\, dt$$

computes one-dimensional wavelet transforms along lines in $\mathbb{R}^n$ which are parameterized in the same fashion as for the X-ray transform. First properties of this transform were studied in [KS, Tak]. The transform was discretized using Fourier methods in [WD] and used to detect linear events in SAR images. The transform has been studied further in [ZZ, Zui, Zui2], where an alternative discretization was proposed. This discretization comes down to the computation of the fast wavelet transform along parallel lines in a rotated and dilated grid.

The motivation for the authors to study the fast wavelet X-ray transform comes from exploration seismology. In a seismic experiment, a dynamite explosion or a vibrating truck produces waves which propagate through the earth's sub-surface and reflect at geophysical interfaces. Wavefronts arriving at the surface of the earth are detected by geophones which are lined up in arrays. An array of vibration measurements in time gives rise to a two-dimensional data set, a *shot record*. This data set contains arrivals, denoted by *reflections*, corresponding to reflected waves which contain relevant information on the deep sub-surface. Unfortunately, the dataset is polluted with high-amplitude arrivals which correspond to diffused waves, known as *groundroll*, just below the surface. These arrivals do not contain relevant information and are simply disturbing. It is an aim of seismic processing to remove such unwanted components from the dataset. Numerical results are presented in [ZZ2].

In [FKV], the X-ray (or Radon) transform was used in a cascade with wavelet transforms in order to remove groundroll from shot records. A cascaded use of the wavelet and Radon transform was also proposed in [MPO] for the detection of linear features in aerial images.

We conclude the introduction with some remarks on notational conventions. The inner product and norm on a Hilbert space H is denoted by $\langle \cdot, \cdot \rangle_H$ and $\| \cdot \|_H$, respectively. In case when $H = \mathbb{R}^n$, the subscript is omitted. The norm of a bounded operator A between Hilbert spaces is denoted by $\|A\|$. The distinction between the norm of a bounded operator and the norm of a vector in $\mathbb{R}^n$ should be clear from the context. The range and kernel of a bounded operator $A : H \to K$ between Hilbert spaces are denoted by $\operatorname{ran} A = \{Ax \mid x \in H\}$ and by $\ker A = \{x \in H \mid Ax = 0\}$. If $\mathbb{G}$ is a countable set, then the Kronecker symbol $\delta_{p,q}$ for $p, q \in \mathbb{G}$ is defined as follows: $\delta_{p,q} = 1$ whenever $p = q$, and $\delta_{p,q} = 0$ whenever $p \neq q$. Further, we define the sequence space $\ell^2(\mathbb{G}) = \{(a_p)_{p\in\mathbb{G}} \mid \sum_{p\in\mathbb{G}} |a_p|^2 < \infty\}$, and we introduce the shorthand notation $\ell^2 = \ell^2(\mathbb{Z})$. The Hilbert space of square integrable functions on $\mathbb{R}^n$ is denoted by $L^2(\mathbb{R}^n)$. The set of measurable functions which are square integrable on compacta is denoted by $L^2_{\text{loc}}(\mathbb{R}^n)$. By abuse of notation, we denote the Fourier transform on ℓ^2 and on $L^2(\mathbb{R})$ by one and the same superscript: Indeed, if $d \in \ell^2$, then $\widehat{d}(\omega) = \sum_{k\in\mathbb{Z}} d_k e^{-i\omega k}$, which defines a 2π-periodic function in $L^2_{\text{loc}}(\mathbb{R})$. On the other hand, if $f \in L^2(\mathbb{R})$, then $\widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t) e^{-i\omega t}\, dt$ defines a unitary transformation $f \mapsto \widehat{f}$ on $L^2(\mathbb{R})$.

2. Bi-orthogonal Riesz systems

We introduce Riesz systems in Hilbert space and focus on geometric issues related to bi-orthogonal pairs of Riesz systems. For general reading on Riesz systems, we refer to [You], although wavelet textbooks like [Chu, CR, Dau, Hol] contain material on the subject relevant to wavelets as well. Recall that a *Riesz system* in a Hilbert space H is a sequence of vectors $(x_k)_{k=1}^\infty$ with two positive constants $0 < A \leq B$ such that

$$A\sum_{k=1}^{N}|a_k|^2 \leq \left\|\sum_{k=1}^{N}a_kx_k\right\|_H^2 \leq B\sum_{k=1}^{N}|a_k|^2$$

for all finite sequences $a_1, \ldots, a_N$. If the constants A, B are chosen optimally for the Riesz system under consideration, then A, B are called the *Riesz bounds* of the Riesz system. Let $(x_k)_{k=1}^{\infty}$ be a Riesz system in the Hilbert space H with closed linear span $V = \operatorname{span}(x_k)_{k=1}^{\infty}$. We first remark that there exists a *unique* sequence of vectors $(\widetilde{x}_k)_{k=1}^{\infty}$ in V such that the *bi-orthogonality condition*

$$\langle x_k, \widetilde{x}_l\rangle_H = \delta_{k,l}, \quad k, l \in \mathbb{Z}^+ \tag{2.1}$$

is satisfied. It turns out that $(\widetilde{x}_k)_{k=1}^{\infty}$ is a Riesz system and we get $\operatorname{span}(\widetilde{x}_k)_{k=1}^{\infty} = V$. Finally, if the Riesz bounds of $(x_k)_{k=1}^{\infty}$ are given by A, B, then the Riesz bounds of $(\widetilde{x}_k)_{k=1}^{\infty}$ are given by $\widetilde{A} = B^{-1}, \widetilde{B} = A^{-1}$.

A Riesz system $(\widetilde{x}_k)_{k=1}^{\infty}$ which satisfies (2.1) will be called a *dual Riesz system* with respect to $(x_k)_{k=1}^{\infty}$. Observe that a Riesz system is its own dual if and only if it is orthonormal. This particular situation corresponds to the case when $A = B = 1$. In the case when the Riesz bounds are the same, this procedure leads to the normalization $A = B = 1$.

If we allow the sequence $(\widetilde{x}_k)_{k=1}^{\infty}$ to span a subspace $\widetilde{V} \subseteq H$ which is different from the subspace V, then the bi-orthogonality condition (2.1) does not imply that $(\widetilde{x}_k)_{k=1}^{\infty}$ is a Riesz system. However, we can state the following.

Theorem 2.1 *Let $(x_k)_{k=1}^{\infty}$ be a Riesz system with Riesz bounds A, B and closed linear span V, and $(\widetilde{x}_k)_{k=1}^{\infty}$ a Riesz system with Riesz bounds $\widetilde{A}, \widetilde{B}$ and closed linear span $\widetilde{V}$. Assume that the bi-orthogonality condition (2.1) is satisfied. In that case, there exists a bounded projection P onto V along $\widetilde{V}^{\perp}$. Moreover, the Riesz bounds of the systems are subject to*

$$\alpha^2 \leq \widetilde{A}B \leq \beta^2, \quad \alpha^2 \leq A\widetilde{B} \leq \beta^2,$$

where

$$\alpha = \inf_{0\neq x\in\widetilde{V}}\frac{\|Px\|_H}{\|x\|_H} = \inf_{0\neq y\in V}\frac{\|P^*y\|_H}{\|y\|_H}, \quad \beta = \sup_{0\neq x\in\widetilde{V}}\frac{\|Px\|_H}{\|x\|_H} = \sup_{0\neq y\in V}\frac{\|P^*y\|_H}{\|y\|_H}.$$

.

Proof Define the operators $L : \ell^2 \to H$ and $\widetilde{L} : \ell^2 \to H$ by

$$L(c_k)_{k=1}^{\infty} = \sum_{k=1}^{\infty}c_kx_k, \quad \widetilde{L}(c_k)_{k=1}^{\infty} = \sum_{k=1}^{\infty}c_k\widetilde{x}_k.$$

Observe that $A \leq L^*L \leq B$ and $\widetilde{A} \leq \widetilde{L}^*\widetilde{L} \leq \widetilde{B}$ (in the sense of positive self-adjoint operators) and that $\widetilde{L}^*L = L^*\widetilde{L} = I$. Consequently, $P = L\widetilde{L}^*$ defines a bounded projection with $\operatorname{ran} P = V$ and $\ker P = \widetilde{V}^{\perp}$.

First, we prove that α and β are well-defined. We observe that $P : \widetilde{V} \to V$ and $P^* : V \to \widetilde{V}$ are bijections. Indeed, If $x \in \widetilde{V}$ and $y = Px \in V$, then there exist $c, d \in \ell^2$ such that $x = \widetilde{L}c$ and $y = Ld$. This gives $Ld = L\widetilde{L}^*\widetilde{L}c$ and hence $c = (\widetilde{L}^*\widetilde{L})^{-1}d$ or $x = \widetilde{L}(\widetilde{L}^*\widetilde{L})^{-1}\widetilde{L}^*y = \widetilde{\Pi}y$, where $\widetilde{\Pi}$ denotes the orthoprojector onto $\widetilde{V}$. Consequently,

$$\alpha = \inf_{0\neq x\in\widetilde{V}}\frac{\|Px\|_H}{\|x\|_H} = \left(\sup_{0\neq y\in V}\frac{\|\widetilde{\Pi}y\|_H}{\|y\|_H}\right)^{-1} = \left(\sup_{0\neq x\in\widetilde{V}, 0\neq y\in V}\frac{\langle\widetilde{\Pi}y, x\rangle_H}{\|y\|_H\|x\|_H}\right)^{-1} =$$

$$\left(\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle y,x\rangle_H}{\|y\|_H\|x\|_H}\right)^{-1}=\inf_{0\neq y\in V}\frac{\|P^*y\|_H}{\|y\|_H},$$

and

$$\beta=\sup_{0\neq x\in\widetilde{V}}\frac{\|Px\|_H}{\|x\|_H}=\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle Px,y\rangle_H}{\|x\|_H\|y\|_H}=\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle x,P^*y\rangle_H}{\|x\|_H\|y\|_H}=\sup_{0\neq y\in V}\frac{\|P^*y\|_H}{\|y\|_H}.$$

For later use, we observe that

$$\beta=\left(\inf_{0\neq y\in V}\frac{\|\widetilde{\Pi}y\|_H}{\|y\|_H}\right)^{-1}.$$

Next, with $c=(c_k)_{k=1}^{\infty}\in\ell^2$, we get

$$\langle\widetilde{L}^*\widetilde{L}c,c\rangle_{\ell^2}\leq\|\widetilde{L}\|^2\langle c,c\rangle_{\ell^2}=\|\widetilde{L}^*\|^2\langle c,c\rangle_{\ell^2}=\sup_{0\neq x\in\widetilde{V}}\frac{\langle\widetilde{L}\widetilde{L}^*x,x\rangle_H}{\langle x,x\rangle_H}\langle c,c\rangle_{\ell^2}. \tag{2.2}$$

Note that for $0\neq x\in\widetilde{V}$,

$$\frac{\langle P^*Px,x\rangle_H}{\langle x,x\rangle_H}=\frac{\langle\widetilde{L}L^*L\widetilde{L}^*x,x\rangle_H}{\langle x,x\rangle_H}\geq A\,\frac{\langle\widetilde{L}\widetilde{L}^*x,x\rangle_H}{\langle x,x\rangle_H},$$

so

$$\frac{\langle\widetilde{L}\widetilde{L}^*x,x\rangle_H}{\langle x,x\rangle_H}\leq\frac{1}{A}\frac{\|Px\|_H^2}{\|x\|_H^2}\leq\frac{\beta^2}{A}. \tag{2.3}$$

Formulas (2.2) and (2.3) imply $A\widetilde{B}\leq\beta^2$. In the same fashion, one proves that $\widetilde{A}B\leq\beta^2$.

In order to prove the other inequalities, observe that

$$\widetilde{A}=\inf_{0\neq c\in\ell^2}\frac{\langle\widetilde{L}c,\widetilde{L}c\rangle_H}{\langle c,c\rangle_{\ell^2}}=\inf_{0\neq x\in\widetilde{V}}\frac{\langle x,x\rangle_H}{\langle L^*x,L^*x\rangle_{\ell^2}}=\left(\sup_{0\neq x\in\widetilde{V}}\frac{\langle L^*x,L^*x\rangle_{\ell^2}}{\langle x,x\rangle_H}\right)^{-1}.$$

We continue as follows:

$$\widetilde{A}^{-1/2}=\sup_{0\neq x\in\widetilde{V}}\frac{\|L^*x\|_{\ell^2}}{\|x\|_H}=\sup_{0\neq x\in\widetilde{V},0\neq c\in\ell^2}\frac{\langle x,Lc\rangle_H}{\|x\|_H\|c\|_{\ell^2}}=\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle x,y\rangle_H}{\|x\|_H\|\widetilde{L}^*y\|_{\ell^2}}\leq$$

$$\|L\|\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle x,y\rangle_H}{\|x\|_H\|y\|_H}=B^{1/2}\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle x,y\rangle_H}{\|x\|_H\|y\|_H}.$$

This provides

$$\widetilde{A}B\geq\left(\sup_{0\neq x\in\widetilde{V},0\neq y\in V}\frac{\langle x,y\rangle_H}{\|x\|_H\|y\|_H}\right)^{-2}=\alpha^2,$$

by the argument given to prove that α was well-defined. Similarly, one proves $A\widetilde{B}\geq\alpha^2$. □

Observe that in the case when $V=\widetilde{V}$, the theorem reproduces the fact that $\widetilde{A}=B^{-1}$ and $\widetilde{B}=A^{-1}$. Moreover, in that case, the projection P is orthogonal, i.e., $P=\Pi$. The bounds in the theorem are not sharp, but can be attained in specific cases. We give a finite-dimensional example to illustrate this.

Example 2.2 Let $H = \mathbb{R}^4$ and consider the bi-orthogonal Riesz systems ($q \in \mathbb{R}$)

$$x_1 = (1,0,0,0)^T, \quad x_2 = (1,1,0,0)^T$$

and

$$\widetilde{x}_1 = (1,-1,1,0)^T, \quad \widetilde{x}_2 = (0,1,0,q).$$

One can calculate that

$$A = \frac{3}{2} - \frac{1}{2}\sqrt{5}, \quad B = \frac{3}{2} + \frac{1}{2}\sqrt{5}$$

and

$$\widetilde{A} = 2 + \frac{1}{2}q^2 - \frac{1}{2}\sqrt{q^4 - 4q^2 + 8}, \quad \widetilde{B} = 2 + \frac{1}{2}q^2 + \frac{1}{2}\sqrt{q^4 - 4q^2 + 8}.$$

The projection P is given by

$$\begin{pmatrix} 1 & 0 & 1 & q \\ 0 & 1 & 0 & q \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and results in the values

$$\alpha^2 = \frac{3}{2} + q^2 + \frac{1}{2}\sqrt{1+4q^4}, \quad \beta^2 = \frac{3}{2} + q^2 - \frac{1}{2}\sqrt{1+4q^4}.$$

For $q \in \mathbb{R}$, as in the theorem, $\widetilde{A}B \leq \beta^2$ and $A\widetilde{B} \geq \alpha^2$, but only for the values $q = \pm 1$, both inequalities are actually equalities. □

The following result deals with the case when one of the systems is orthonormal.

Proposition 2.3 *Let $(x_k)_{k=1}^\infty$ be an orthonormal system with closed linear span V, and $(\widetilde{x}_k)_{k=1}^\infty$ a Riesz system with closed linear span $\widetilde{V}$. Assume that the bi-orthogonality condition (2.1) is satisfied. In that case, there exists a bounded projection P onto V along $\widetilde{V}^\perp$. Moreover, the Riesz bounds $\widetilde{A}, \widetilde{B}$ of $(\widetilde{x}_k)_{k=1}^\infty$ are given by $\widetilde{A} = \alpha^2$ and $\widetilde{B} = \beta^2$, where*

$$\alpha = \inf_{0 \neq x \in \widetilde{V}} \frac{\|Px\|_H}{\|x\|_H}, \quad \beta = \sup_{0 \neq x \in \widetilde{V}} \frac{\|Px\|_H}{\|x\|_H}.$$

Proof In this case, the orthogonal projection onto V is given by $\Pi = LL^*$, hence, with the substitution $x = \widetilde{L}c$,

$$\widetilde{A} = \inf_{0 \neq c \in \ell^2} \frac{\langle \widetilde{L}c, \widetilde{L}c \rangle_H}{\langle c, c \rangle_{\ell^2}} = \inf_{0 \neq x \in \widetilde{V}} \frac{\langle x, x \rangle_H}{\langle L^*x, L^*x \rangle_{\ell^2}} = \left(\sup_{0 \neq x \in \widetilde{V}} \frac{\|\Pi x\|_H}{\|x\|_H} \right)^{-2} = \alpha^2.$$

Here we use

$$\sup_{0 \neq x \in \widetilde{V}} \frac{\|\Pi x\|_H}{\|x\|_H} = \sup_{0 \neq x \in \widetilde{V}, 0 \neq y \in V} \frac{\langle \Pi x, y \rangle_H}{\|x\|_H \|y\|_H} = \sup_{0 \neq x \in \widetilde{V}, 0 \neq y \in V} \frac{\langle x, y \rangle_H}{\|x\|_H \|y\|_H} = \alpha^{-1}.$$

In the same fashion, using the substitution $y = Lc$ (and $c = L^*y$),

$$\widetilde{B} = \sup_{0 \neq c \in \ell^2} \frac{\langle \widetilde{L}c, \widetilde{L}c \rangle_H}{\langle c, c \rangle_{\ell^2}} = \sup_{0 \neq y \in V} \frac{\langle P^*y, P^*y \rangle_H}{\langle L^*y, L^*y \rangle_{\ell^2}} = \left(\sup_{0 \neq y \in V} \frac{\|P^*y\|_H}{\|y\|_H} \right)^2 = \beta^2.$$ □

The following proposition shows that the existence of a bi-orthogonal system in two subspaces $V, \widetilde{V} \subseteq H$ is equivalent to the geometric condition $\widetilde{V} \oplus V^\perp = H$.

Proposition 2.4 *Let $V, \widetilde{V} \subseteq H$ be subspaces in Hilbert space, and assume that V contains a Riesz basis $(x_k)_{k=1}^\infty$ with Riesz bounds A, B. Then $\widetilde{V}$ contains a Riesz basis $(\widetilde{x}_k)_{k=1}^\infty$ bi-orthogonal to $(x_k)_{k=1}^\infty$ if and only if $\widetilde{V} \oplus V^\perp = H$, i.e., if and only if there exists a bounded projection onto $\widetilde{V}$ along $V^\perp$. In that case, the Riesz basis $(\widetilde{x}_k)_{k=1}^\infty$ in $\widetilde{V}$ is uniquely determined. The Riesz bounds of this system are subject to the same estimates as in Theorem 2.1.*

Proof Assume that $\widetilde{V}$ contains a Riesz basis $(\widetilde{x}_k)_{k=1}^\infty$ which is bi-orthogonal to $(x_k)_{k=1}^\infty$. If $x \in \widetilde{V}$, then $x = \sum_{k=1}^\infty \alpha_k \widetilde{x}_k$ with $\alpha_k = \langle x, x_k \rangle_H$ for all $k \in \mathbb{Z}^+$. However, if $x \in V^\perp$, then $\langle x, x_k \rangle_H = 0$ for all $k \in \mathbb{Z}^+$. Therefore, $\widetilde{V} \cap V^\perp = (0)$. Given $y \in H$, write $y = \sum_{k=1}^\infty \langle y, x_k \rangle_H \widetilde{x}_k + z$ with $z \in H$. It is immediate that $z \in V^\perp$. This implies $\widetilde{V} + V^\perp = H$.

To prove the converse, assume that $\widetilde{V} \oplus V^\perp = H$. For each $k \in \mathbb{Z}^+$, there exists a unique $y_k \in V$ such that $\langle x_l, y_k \rangle_H = \delta_{k,l}$ for all $l \in \mathbb{Z}^+$. In fact, the system $(y_k)_{k=1}^\infty$ is a Riesz system in V. Indeed, if $L : \ell^2 \to H$ is given by $L(c_k)_{k=1}^\infty = \sum_{k=1}^\infty c_k x_k$, then $x_k = Le_k$ and $y_k = Ke_k = L(L^*L)^{-1}e_k$, where $e_k \in \ell^2$ is the k-th standard basis vector. Observe that $K^*K = (L^*L)^{-1}L^*L(L^*L)^{-1} = (L^*L)^{-1}$, so

$$B^{-1} \le K^*K \le A^{-1}$$

in the sense of positive self-adjoint operators. It follows that the Riesz bounds of $(y_k)_{k=1}^\infty$ are given by B^{-1}, A^{-1}.

By assumption, given $y_k \in V$, there exist unique $\widetilde{x}_k \in \widetilde{V}$ and $z_k \in V^\perp$, such that $y_k = \widetilde{x}_k + z_k$. Obviously,

$$\langle x_l, \widetilde{x}_k \rangle_H = \langle x_l, y_k \rangle_H = \delta_{k,l}, \quad k, l \in \mathbb{Z}^+.$$

If P denotes the projection onto V along $\widetilde{V}^\perp$, then $\widetilde{x}_k = P^* y_k$. Therefore, using the same notation as in Theorem 2.1,

$$\left\| \sum_{k=1}^N a_k \widetilde{x}_k \right\|_H^2 = \left\| P^* \left(\sum_{k=1}^N a_k y_k \right) \right\|_H^2 \le \beta^2 \left\| \sum_{k=1}^N a_k y_k \right\|_H^2 \le \frac{\beta^2}{A} \sum_{k=1}^N |a_k|^2,$$

$$\left\| \sum_{k=1}^N a_k \widetilde{x}_k \right\|_H^2 = \left\| P^* \left(\sum_{k=1}^N a_k y_k \right) \right\|_H^2 \ge \alpha^2 \left\| \sum_{k=1}^N a_k y_k \right\|_H^2 \ge \frac{\alpha^2}{B} \sum_{k=1}^N |a_k|^2.$$

This implies that $(\widetilde{x}_k)_{k=1}^\infty$ is a Riesz system in $\widetilde{V}$ with Riesz bounds $\widetilde{A}, \widetilde{B}$, which satisfy $\alpha^2 \le \widetilde{A}B \le \beta^2$ and $\alpha^2 \le A\widetilde{B} \le \beta^2$ as in Theorem 2.1. In order to prove that $(\widetilde{x}_k)_{k=1}^\infty$ is a Riesz basis in $\widetilde{V}$, observe that if $z \in H$, then

$$z = \sum_{k=1}^\infty \langle z, x_k \rangle_H y_k + u,$$

where $u \in V^\perp$. If P^* is the projection along $V^\perp$ onto $\widetilde{V}$, then

$$P^* z = \sum_{k=1}^\infty \langle z, x_k \rangle_H \widetilde{x}_k,$$

so $\widetilde{V} = \operatorname{ran} P^* \subseteq \operatorname{span}(\widetilde{x}_k)_{k=1}^\infty$.

Finally, assume there exists $u_k \in \widetilde{V}$ which satisfies $\langle x_l, u_k \rangle_H = \delta_{kl}$ for all $l \in \mathbb{Z}^+$. It follows that $\widetilde{x}_k - u_k \in \widetilde{V} \cap V^\perp = (0)$. This proves uniqueness of the system in $\widetilde{V}$ bi-orthogonal to $(x_k)_{k=1}^\infty$. □

We shall now specialize to Riesz systems which consist of integer translates of a single function $\varphi \in L^2(\mathbb{R})$. More explicitly, we will assume that the sequence $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ is a Riesz system in $L^2(\mathbb{R})$. If V is the closed linear span of this Riesz system, then there exists a unique Riesz system in V bi-orthogonal to $(\varphi(\cdot-k))_{k\in\mathbb{Z}}$ which is a basis in V. In particular, for the given function φ, one may construct a function $\widetilde{\varphi} \in V$, such that the dual Riesz system of $(\varphi(\cdot-k))_{k\in\mathbb{Z}}$ is given by $(\widetilde{\varphi}(\cdot-k))_{k\in\mathbb{Z}}$.

In order to construct such a function, we quote from [Chu, CR, Dau] that for $\varphi \in L^2(\mathbb{R})$, the function

$$S = 2\pi \sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\cdot + 2n\pi)|^2$$

defines a 2π-periodic function in $L^1_{\text{loc}}(\mathbb{R})$. Put $V = \text{span}\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$. The system $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis in V with Riesz bounds $0 < A \leq B$ if and only if $\inf_{\mathbb{R}} S = A$ and $\sup_{\mathbb{R}} S = B$. In other words, the function S should be bounded and bounded away from zero in order to correspond with a Riesz basis. In particular, the system $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis in V if and only if $S \equiv 1$.

In the case when $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis, then $\{\varphi^o(\cdot - k)\}_{k\in\mathbb{Z}}$, given by $\widehat{\varphi}^o = S^{-1/2}\widehat{\varphi}$, is an orthonormal basis in V. The dual Riesz basis $\{\widetilde{\varphi}(\cdot - k)\}_{k\in\mathbb{Z}}$ of $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ in V is constructed by means of $\widehat{\widetilde{\varphi}} = S^{-1}\widehat{\varphi}$. If in addition, φ has compact support and is bounded, then φ^o and $\widetilde{\varphi}$ have exponential decay at infinity; see [CR].

We now specify Theorem 2.1 to this situation. The periodic function in this proposition can also be found in [UA].

Proposition 2.5 *Let $\varphi, \widetilde{\varphi} \in L^2(\mathbb{R})$ such that $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ and $(\widetilde{\varphi}(\cdot - k))_{k\in\mathbb{Z}}$ are Riesz systems in $L^2(\mathbb{R})$ with closed linear spans V and $\widetilde{V}$, respectively. The function $Z \in L^1_{\text{loc}}(\mathbb{R})$ given by*

$$Z(\omega) = \frac{\sum_{n\in\mathbb{Z}} \widehat{\varphi}(\omega + 2n\pi)\overline{\widehat{\widetilde{\varphi}}(\omega + 2n\pi)}}{\sqrt{\sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\omega + 2n\pi)|^2 \cdot \sum_{n\in\mathbb{Z}} |\widehat{\widetilde{\varphi}}(\omega + 2n\pi)|^2}}$$

is 2π-periodic. Define the numbers

$$\alpha = \left(\text{ess}\sup_{0\leq\omega\leq 2\pi} |Z(\omega)|\right)^{-1}, \quad \beta = \left(\text{ess}\inf_{0\leq\omega\leq 2\pi} |Z(\omega)|\right)^{-1}.$$

If A, B and $\widetilde{A}, \widetilde{B}$ are the Riesz bounds of $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ and $(\widetilde{\varphi}(\cdot - k))_{k\in\mathbb{Z}}$, respectively, then

$$\alpha^2 \leq \widetilde{A}B \leq \beta^2, \quad \alpha^2 \leq A\widetilde{B} \leq \beta^2.$$

Proof Using Theorem 2.1, it suffices to prove that

$$\inf_{0\neq f\in V} \frac{\|\widetilde{\Pi} f\|_{L^2(\mathbb{R})}}{\|f\|_{L^2(\mathbb{R})}} = \text{ess}\inf_{0\leq\omega\leq 2\pi} |Z(\omega)|, \tag{2.4}$$

$$\sup_{0\neq f\in V} \frac{\|\widetilde{\Pi} f\|_{L^2(\mathbb{R})}}{\|f\|_{L^2(\mathbb{R})}} = \text{ess}\sup_{0\leq\omega\leq 2\pi} |Z(\omega)|. \tag{2.5}$$

Here $\widetilde{\Pi}$ is the orthogonal projection onto $\widetilde{V}$. Let $\varphi^o \in V$ and $\widetilde{\varphi}^o \in \widetilde{V}$ be given so that $(\varphi^o(\cdot - k))_{k\in\mathbb{Z}}$ is an orthonormal basis in V and $(\widetilde{\varphi}^o(\cdot - k))_{k\in\mathbb{Z}}$ is an orthonormal basis in $\widetilde{V}$. The construction of such functions has been indicated above.

If $f \in V$, then $f = \sum_{k\in\mathbb{Z}} c_k \varphi^o(\cdot - k)$ with $c = (c_k)_{k\in\mathbb{Z}} \in \ell^2$. Moreover,

$$\widetilde{\Pi} f = \sum_{k,l\in\mathbb{Z}} c_k \langle \varphi^o(\cdot - k), \widetilde{\varphi}^o(\cdot - l)\rangle_{L^2(\mathbb{R})} \widetilde{\varphi}^o(\cdot - l).$$

If we write $d_l = \sum_{k\in\mathbb{Z}} c_k \langle \varphi^0(\cdot + l - k), \widetilde{\varphi}^o\rangle_{L^2(\mathbb{R})}$, then $d = (d_l)_{l\in\mathbb{Z}}$ is the discrete convolution of c with $\rho = (\rho_k)_{k\in\mathbb{Z}}$, where $\rho_k = \langle \varphi^0(\cdot + k), \widetilde{\varphi}^o\rangle_{L^2(\mathbb{R})}$. We write $d = c * \rho$. Observe that $d = c*\rho$ implies $\widehat{d} = \widehat{c}\cdot\widehat{\rho}$. Therefore,

$$\|\widetilde{\Pi} f\|^2_{L^2(\mathbb{R})} = \sum_{l\in\mathbb{Z}} |d_l|^2 = \frac{1}{2\pi}\int_0^{2\pi} |\widehat{c}(\omega)|^2 |\widehat{\rho}(\omega)|^2|\, d\omega,$$

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{k\in\mathbb{Z}} |c_k|^2 = \frac{1}{2\pi}\int_0^{2\pi} |\widehat{c}(\omega)|^2\, d\omega.$$

Note that $(k\in\mathbb{Z})$

$$\rho_k = \langle \varphi^o(\cdot + k), \widetilde{\varphi}^o\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \widehat{\varphi}^o(\omega)\overline{\widehat{\widetilde{\varphi}^o}(\omega)} e^{ik\omega}\, d\omega = \frac{1}{2\pi}\int_0^{2\pi} Z(\omega) e^{ik\omega}\, d\omega.$$

Consequently, $\widehat{\rho} = Z$ almost everywhere, and equations (2.5) and (2.5) follow. □

In the papers [Uns, Uns2, UA], the constant β as in this section arises as the reciprocal cosine of the angle between V and $\widetilde{V}$ and is used there to obtain the following result.

Proposition 2.6 *Let $V, \widetilde{V} \subseteq H$ be subspaces in Hilbert space such that $V \oplus \widetilde{V}^\perp = H$. Let P be the projection onto V along $\widetilde{V}^\perp$ and let Π be the orthoprojector onto V. For all $g\in H$, we get*

$$\|g - \Pi g\|_H \le \|g - Pg\|_H \le \beta \|g - \Pi g\|_H,$$

where

$$\beta = \sup_{0\neq k\in V} \frac{\|P^* k\|_H}{\|k\|_H}.$$

Proof Observe that

$$\beta^2 = \sup_{0\neq k\in V} \frac{\|P^*k\|^2_{L^2(\mathbb{R})}}{\|k\|^2_{L^2(\mathbb{R})}} = \sup_{\substack{0\neq k\in V,\, g\in V\\ f\in V^\perp,\, g+f\neq 0}} \frac{\langle P^*k, g+f\rangle^2_{L^2(\mathbb{R})}}{\|k\|^2_{L^2(\mathbb{R})}\left(\|g\|^2_{L^2(\mathbb{R})} + \|f\|^2_{L^2(\mathbb{R})}\right)} =$$

$$\sup_{\substack{0\neq k\in V,\, g\in V\\ f\in V^\perp,\, g+f\neq 0}} \frac{\left(\langle k, g\rangle_{L^2(\mathbb{R})} + \langle k, Pf\rangle_{L^2(\mathbb{R})}\right)^2}{\|k\|^2_{L^2(\mathbb{R})}\left(\|g\|^2_{L^2(\mathbb{R})} + \|f\|^2_{L^2(\mathbb{R})}\right)} \ge$$

(here we restrict ourselves to $g = k = tPf$ for some $t\in\mathbb{R}$)

$$\sup_{0\neq f\in V^\perp} \frac{\left(t^2\|Pf\|^2_{L^2(\mathbb{R})} + t\|Pf\|^2_{L^2(\mathbb{R})}\right)^2}{t^2\|Pf\|^2_{L^2(\mathbb{R})}\left(t^2\|Pf\|^2_{L^2(\mathbb{R})} + \|f\|^2_{L^2(\mathbb{R})}\right)} =$$

$$\sup_{0\neq f\in V^\perp} \frac{(t+1)^2\|Pf\|^2_{L^2(\mathbb{R})}}{t^2\|Pf\|^2_{L^2(\mathbb{R})} + \|f\|^2_{L^2(\mathbb{R})}} = \sup_{0\neq f\in V^\perp} \frac{\|f - Pf\|^2_{L^2(\mathbb{R})}}{\|f\|^2_{L^2(\mathbb{R})}},$$

if we choose

$$t = \frac{\|f\|^2_{L^2(\mathbb{R})}}{\|Pf\|^2_{L^2(\mathbb{R})}}.$$

Observe that the case when $Pf = 0$ is not relevant for the estimate. □

In the next section, we will elaborate on estimates from above of least squares distances of the form $\|g - \Pi g\|_H$.

3. Estimates on the least squares distance

Let $h > 0$ be a (small) positive number. We consider estimates from above of the least squares distance $\|f - \Pi_h f\|_{L^2(\mathbb{R}^n)}$, where $f \in L^2(\mathbb{R})$ and Π_h is the orthogonal projection onto V_h, the closed linear subspace spanned by the Riesz system $(h^{-1/2}\varphi(h^{-1}\cdot - k))_{k=1}^{\infty}$. It turns out that if $f \in L^2(\mathbb{R})$ is *smooth* in a sense to be explained below, and if φ satisfies the *Strang-Fix conditions* and has suitable decay, then useful estimates can be given.

For ease of notation, we deal with uni-variate functions. These results allow for immediate application to our grid functions. We remark, however, that corresponding results have been proved for multi-variate functions and, in addition, there are even results describing the case when the single function φ is replaced by a finite family of functions; see [HL, JL, SF].

The Sobolev space $H^q(\mathbb{R})$ consists of functions $f \in L^2(\mathbb{R})$ which have all derivatives up to order q in $L^2(\mathbb{R})$. Here q is a nonnegative integer. The space $H^q(\mathbb{R})$ becomes a Banach space with the norm $\|f\|_{H^q(\mathbb{R})} = \sum_{k=0}^{q} \|f^{(k)}\|_{L^2(\mathbb{R})}$.

Theorem 3.1 *Fix $\varepsilon > 0$. Assume that $\varphi \in L^2(\mathbb{R})$ induces a Riesz system and that it has sufficient decay in Fourier domain:*

$$\operatorname{ess\,sup}_{\omega\in\mathbb{R}} \widehat{\varphi}(\omega)(1 + |\omega|)^{1+\varepsilon} < \infty.$$

Moreover, assume there exists $c > 0$ such that $R(\omega) = \sum_{n\in\mathbb{Z}} \widehat{\varphi}(\omega + 2\pi n)$ satisfies $|R(\omega)| \geq c$ for almost all $\omega \in \mathbb{R}$. In that case, the function φ_{QI}, defined through $\widehat{\varphi}_{QI} = R^{-1}\widehat{\varphi}$, is interpolatory*, i.e., φ_{QI} is continuous and $\varphi_{QI}(k) = \delta_{k,0}$ for $k \in \mathbb{Z}$. If, in addition, φ_{QI} has polynomial decay*

$$\sup_{x\in\mathbb{R}} |\varphi_{QI}(x)|(1 + |x|)^{p+1+\varepsilon} < \infty,$$

then the following statements are equivalent:

(1) For each $f \in H^{p+1}(\mathbb{R})$ and each $h > 0$,

$$\left\| f - \sum_{k\in\mathbb{Z}} f(hk)\varphi_{QI}\left(\frac{\cdot}{h} - k\right) \right\|_{L^2(\mathbb{R})} \leq Ch^{p+1}\|f^{(p+1)}\|_{L^2(\mathbb{R})}.$$

The constant C does not depend on f and h.

(2) The function $\varphi_{QI} \in V$ is an quasi-interpolant *of order p, i.e., all polynomials Q of degree $\leq p$ satisfy*

$$Q(x) = \sum_{k\in\mathbb{Z}} Q(k)\varphi_{QI}(x - k).$$

The series converges for each $x \in \mathbb{R}$.

(3) The function $\varphi_{QI} \in V$ satisfies the Strang-Fix conditions of order p:

$$\widehat{\varphi}_{QI}(0) \neq 0, \quad \widehat{\varphi}_{QI}^{(k)}(2\pi n) = 0, \quad k = 0, \dots, p, \quad n \neq 0.$$

Under the condition that $\varphi \in L^2(\mathbb{R})$ has compact support, an equivalence comparable to (1) $\Leftrightarrow$ (3) was proved in [SF]. For a proof of Theorem 3.1, we refer to [Uns, UD]. Theorem 3.1 can be applied as follows to estimate $\|f - \Pi_h f\|_{L^2(\mathbb{R})}$. First of all, the operator $J_h : H^q(\mathbb{R}) \to L^2(\mathbb{R})$ given by

$$J_h f = \sum_{k\in\mathbb{Z}} f(hk)\varphi_{QI}\left(\frac{\cdot}{h} - k\right)$$

maps $H^q(\mathbb{R})$ into V_h. Therefore,

$$\|f - \Pi_h f\|_{L^2(\mathbb{R})} \leq \|f - J_h f\|_{L^2(\mathbb{R})} \leq Ch^{p+1}\|f^{(p+1)}\|_{L^2(\mathbb{R})}.$$

4. Multi-resolution Analysis

In this section, we define the notions of scaling function and wavelet in the context of multi-resolution analysis. Multi-resolution analysis forms one of the key issues in wavelet theory; see [CR, Dau, Mal]. A *multi-resolution analysis* (MRA) in $L^2(\mathbb{R})$ is a sequence of subspaces $(V_j)_{j\in\mathbb{Z}}$ in $L^2(\mathbb{R})$ with the following properties: For each $j \in \mathbb{Z}$, we have

(i) $V_j \subset V_{j+1}$,

(ii) $f \in V_j \ \Leftrightarrow\ f(2\cdot) \in V_{j+1}$,

and

(iii) $\bigcap_{j\in\mathbb{Z}} V_j = (0)$,

(iv) $\bigcup_{j\in\mathbb{Z}} V_j \subseteq L^2(\mathbb{R})$ dense,

(v) there exists $\varphi \in V_0$ such that $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ is a Riesz basis in V_0.

It is by now well-known that certain functions $\varphi \in L^2(\mathbb{R})$ give rise to a multi-resolution analysis. The construction, if possible, goes along the following lines: Conditions which are put on $\varphi \in L^2(\mathbb{R})$ will come along the way. Of course, condition (v) is satisfied only if $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ is a Riesz system in $L^2(\mathbb{R})$. The closed linear span of this Riesz system is -following notation above- denoted by V_0. It is not difficult to see that for each $j \in \mathbb{Z}$, the sequence $(2^{j/2}\varphi(2^j \cdot - k))_{k\in\mathbb{Z}}$ is a Riesz system in $L^2(\mathbb{R})$ with the same Riesz bounds as the original Riesz system. For each $j \in \mathbb{Z}$, we denote the closed linear span of the Riesz system $(2^{j/2}\varphi(2^j \cdot - k))_{k\in\mathbb{Z}}$ by V_j. Observe that condition (i) is satisfied if and only if there exists $h = (h_k)_{k\in\mathbb{Z}} \in \ell^2$, such that

$$\varphi(x) = \sum_{k\in\mathbb{Z}} h_k \sqrt{2}\varphi(2x - k), \quad \text{almost all } x \in \mathbb{R},$$

or, after Fourier transformation,

$$\sqrt{2}\,\widehat{\varphi}(2\omega) = \widehat{h}(\omega)\widehat{\varphi}(\omega), \quad \text{almost all } \omega \in [0, 2\pi].$$

In the wavelet literature, one often writes $m_0(\omega) = \widehat{h}(\omega) = \sum_{k\in\mathbb{Z}} h_k e^{-i\omega k}$, and we shall adopt this convention. Condition (iii) follows from (v). Moreover, if we assume that $\widehat{\varphi}$ is uniformly bounded on $\mathbb{R}$ and continuous at zero, and if $\widehat{\varphi}(0) \neq 0$, then (v) also implies (iv); see [Dau]. A function $\varphi \in L^2(\mathbb{R})$ will be called a *scaling function* if it induces a multi-resolution analysis as indicated above. The following lemma shows under which circumstances the dual wavelet is also a scaling function. Many elements from the considerations below are taken from [CDF].

Lemma 4.1 *Let $\varphi \in L^2(\mathbb{R})$ be a scaling function and assume that $\widetilde{\varphi} \in L^2(\mathbb{R})$ induces a Riesz system bi-orthogonal to φ. If there exists $\widetilde{h} = (\widetilde{h}_k)_{k\in\mathbb{Z}} \in \ell^2$ such that*

$$\widetilde{\varphi}(x) = \sum_{k\in\mathbb{Z}} \widetilde{h}_k \sqrt{2}\widetilde{\varphi}(2x-k), \quad \text{almost all } x \in \mathbb{R},$$

then $\widetilde{\varphi}$ is a scaling function too.

Proof We need to show that $\widetilde{\varphi}$ induces a multi-resolution analysis. Define $\widetilde{V}_j$ to be the closed linear span of $(2^{j/2}\widetilde{\varphi}(2^j\cdot -k))_{k\in\mathbb{Z}}$ for $j \in \mathbb{Z}$. We prove that the sequence $(\widetilde{V}_j)_{j\in\mathbb{Z}}$ satisfies properties (i)-(v) of a multi-resolution analysis. It is obvious that (i) and (ii) are satisfied. Proposition 2.4 implies that $V_j^{\perp} \oplus \widetilde{V}_j = L^2(\mathbb{R})$ for all $j \in \mathbb{Z}$, from which we can derive (iii) and (iv). Condition (v) is satisfied by assumption. □

We shall now construct wavelets starting with the bi-orthogonal scaling functions φ and $\widetilde{\varphi}$. Define (for almost all $\omega \in [0, 2\pi]$)

$$m_0(\omega) = \sum_k h_k e^{i\omega k}, \quad \widetilde{m}_0(\omega) = \sum_k \widetilde{h}_k e^{i\omega k},$$

and

$$m_1(\omega) = e^{-i\omega}\overline{\widetilde{m}_0(\omega+\pi)}, \quad \widetilde{m}_1(\omega) = e^{-i\omega}\overline{m_0(\omega+\pi)}.$$

The wavelets $\psi, \widetilde{\psi}$ are defined through

$$\widehat{\psi}(\omega) = \frac{1}{2}\sqrt{2}m_1(\omega/2)\widehat{\varphi}(\omega/2), \quad \widehat{\widetilde{\psi}}(\omega) = \frac{1}{2}\sqrt{2}\widetilde{m}_1(\omega/2)\widehat{\widetilde{\varphi}}(\omega/2).$$

Observe that this comes down to

$$\psi(x) = \sum_k g_k \sqrt{2}\varphi(2x-k), \quad \widetilde{\psi}(x) = \sum_k \widetilde{g}_k \sqrt{2}\varphi(2x-k),$$

where $g_k = (-1)^k \widetilde{h}_{1-k}$ and $\widetilde{g}_k = (-1)^k h_{1-k}$.

Lemma 4.2 *The wavelets $\psi, \widetilde{\psi}$ induce Riesz bases in $W_0, \widetilde{W_0}$, respectively.*

Proof Observe that

$$\sum_{n\in\mathbb{Z}} |\widehat{\psi}(2\omega+2n\pi)|^2 = \frac{1}{2}\sum_{n\in\mathbb{Z}} |m_1(\omega+n\pi)|^2|\widehat{\varphi}(\omega+n\pi)|^2 =$$

$$\frac{1}{2}\sum_{n\in\mathbb{Z}} \left\{|m_1(\omega)|^2|\widehat{\varphi}(\omega+2n\pi)|^2 + |m_1(\omega+\pi)|^2|\widehat{\varphi}(\omega+\pi+2n\pi)|^2\right\} \le \frac{B}{2}\left\{|\widetilde{m}_0(\omega)|^2 + |\widetilde{m}_0(\omega+\pi)|^2\right\}.$$

Further,

$$\widetilde{B} \ge \sum_{n\in\mathbb{Z}} |\widetilde{\varphi}(2\omega+2n\pi)|^2 = \frac{1}{2}\sum_{n\in\mathbb{Z}} \left\{|\widetilde{m}_0(\omega)|^2|\widehat{\widetilde{\varphi}}(\omega+2n\pi)|^2 + |\widetilde{m}_0(\omega+\pi)|^2|\widehat{\widetilde{\varphi}}(\omega+\pi+2n\pi)|^2\right\} \ge$$

$$\frac{\widetilde{A}}{2}\left\{|\widetilde{m}_0(\omega)|^2 + |\widetilde{m}_0(\omega+\pi)|^2\right\},$$

which implies that

$$\sum_{n\in\mathbb{Z}} |\widehat{\psi}(2\omega + 2n\pi)|^2 \le \widetilde{B}B\widetilde{A}^{-1}.$$

In the same fashion, one proves that

$$\sum_{n\in\mathbb{Z}} |\widehat{\psi}(2\omega + 2n\pi)|^2 \ge \widetilde{A}A\widetilde{B}^{-1}.$$

The proof that $\widetilde{\psi}$ induces a Riesz basis goes along the same lines. □

In addition to the result of the preceding lemma, note that if

$$W_j = \operatorname{span}(2^{j/2}\psi(2^j\cdot - k))_{k\in\mathbb{Z}}, \quad \widetilde{W_j} = \operatorname{span}(2^{j/2}\widetilde{\psi}(2^j\cdot - k))_{k\in\mathbb{Z}},$$

then we have the (not necessarily orthogonal) decompositions

$$W_j \oplus V_j = V_{j+1}, \quad \widetilde{W_j} \oplus \widetilde{V}_j = \widetilde{V}_{j+1}.$$

Further,

$$V_j \perp \widetilde{W_j}, \quad \widetilde{V}_j \perp W_j,$$

so we may conclude that $(2^{j/2}\psi(2^j\cdot - k))_{j,k\in\mathbb{Z}}$ and $(2^{j/2}\widetilde{\psi}(2^j\cdot - k))_{j,k\in\mathbb{Z}}$ are bi-orthogonal Riesz *bases* in $L^2(\mathbb{R})$.

Let $0 \ne \varphi \in L^2(\mathbb{R})$ be a real-valued function with compact support $[R_1, R_2]$, say, and assume that $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$ is a Riesz system in $L^2(\mathbb{R})$ with closed linear span V. There exists a unique $\widetilde{\varphi} \in V$ such that $(\widetilde{\varphi}(\cdot - k))_{k\in\mathbb{Z}}$ is a Riesz basis in V bi-orthogonal to $(\varphi(\cdot - k))_{k\in\mathbb{Z}}$. The construction of $\widetilde{\varphi}$, which has been mentioned before, will be made more explicit here. The construction goes along the lines as in [Chu] for the spline case. One needs to determine the sequence $(c_k)_{k\in\mathbb{Z}} \in \ell^2$ such that $\widetilde{\varphi} = \sum_{k\in\mathbb{Z}} c_k\varphi(\cdot - k)$. The bi-orthogonality condition implies

$$\delta_{k,0} = \sum_{m\in\mathbb{Z}} c_m\langle\varphi(\cdot + m - k), \varphi\rangle_{L^2(\mathbb{R})}, \quad k \in \mathbb{Z}.$$

Define $a_k = \langle\varphi(\cdot + k), \varphi\rangle_{L^2(\mathbb{R})}$. Observe that $a_k = a_{-k}$ and that $a_k = 0$ whenever $|k| > R_2 - R_1$. If we apply the Fourier transform, we get $1 = a(e^{i\theta})c(e^{i\theta})$, where $c(e^{i\theta}) = \sum_{k\in\mathbb{Z}} c_k e^{ik\theta}$ and, for some positive integer $m \le R_2 - R_1$,

$$a(e^{i\theta}) = a_0 + \sum_{k=1}^{m} a_k(e^{ik\theta} + e^{-ik\theta}), \quad a_m \ne 0.$$

Observe that $a(\cdot)$ does not vanish on the unit circle. Therefore, $p(z) = z^m a(z)$ is a polynomial of degree $2m$ with roots inside and outside the unit circle. In fact, $p(0) = a_m \ne 0$ and if $\lambda \ne 0$ satisfies $p(\lambda) = 0$, then $p(1/\lambda) = 0$. Therefore, we may write

$$a(z) = a_m z^{-m}\prod_{j=1}^{m}(z - \lambda_j)(z - \frac{1}{\lambda_j}), \quad z \ne 0,$$

where $1 > |\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_m|$. We state the following simple lemma.

Lemma 4.3 *Let a, b be complex numbers such that $|a|, |b| < 1$. Then*

$$\sum_{k=0}^{\infty} a^k z^k \cdot \sum_{k=0}^{\infty} b^k z^{-k} = \sum_{k\in\mathbb{Z}} A_k z^k, \quad |b| < |z| < |a|^{-1},$$

where

$$A_k = \begin{cases} \frac{a^k}{1-ab}, & k \geq 0 \\ \frac{b^k}{1-ab}, & k \leq 0 \end{cases}.$$

The lemma provides

$$c(z) = \frac{1}{a_m} \prod_{j=1}^{m} \frac{1}{1-\lambda_j/z} \cdot \frac{-\lambda_j}{1-\lambda_j z} = \frac{1}{a_m} \prod_{j=1}^{m} \left(\frac{-\lambda_j}{1-\lambda_j^2} \cdot \sum_{k\in\mathbb{Z}} \lambda_j^{|k|} z^k \right)$$

for $|\lambda_1| < |z| < |\lambda_1|^{-1}$, and we arrive at ($k \in \mathbb{Z}$)

$$c_k = \frac{1}{a_m} \sum_{k_1+\cdots+k_m=k} \prod_{j=1}^{m} \frac{-\lambda_j^{|k_j|+1}}{1-\lambda_j^2}.$$

5. Wavelet interpolation method

Throughout this section, we will assume that φ induces a multi-resolution analysis as described in the previous section, i.e., φ is a scaling function. We will also fix a dual scaling function $\widetilde{\varphi} \in L^2(\mathbb{R})$.

Given an orthonormal basis $\theta_1, \ldots, \theta_n$ in $\mathbb{R}^n$, we consider the grid

$$\mathbb{G}_{\theta,d} = \{\sum_{r=1}^{n} p_r d_r \theta_r \mid (p_1, \ldots, p_n) \in \mathbb{Z}^n\},$$

where $d_1, \ldots, d_n$ are positive real numbers. The standard basis in $\mathbb{R}^n$ will be denoted by $e_1, \ldots, e_n$ and the standard grid by $\mathbb{G}_e$ accordingly. Observe that $\mathbb{G}_{\theta,d}$ is the image of $\mathbb{G}_e$ under reflection, rotation and dilation. In fact, $\mathbb{G}_{\theta,d} = DRS\mathbb{G}_e$, where S is an $n \times n$ reflection matrix, R is an $n \times n$ rotation matrix and D is an $n \times n$ diagonal matrix with diagonal $d = (d_1, \ldots, d_n)^T$.

In order to deal with functions on $\mathbb{R}^n$ for $n \geq 2$, we shall construct multi-variate functions by means of products of uni-variate ones. Indeed, given an orthonormal basis $\theta_1, \ldots, \theta_n$ in $\mathbb{R}^n$, we shall write

$$\Phi_{\theta,d}(y) = \prod_{r=1}^{n} d_r^{-1/2} \varphi(d_r^{-1} \langle y, \theta_r \rangle), \quad \text{almost all } y \in \mathbb{R}^n,$$

where φ is a scaling function. In the same fashion, one defines $\widetilde{\Phi}_{\theta,d}$ using the dual scaling function $\widetilde{\varphi}$. The next two lemmas justify that $\Phi_{\theta,d}$ will be called a *multi-variate scaling function*.

Lemma 5.1 *The system $(\Phi_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ is a Riesz system in $L^2(\mathbb{R}^n)$. The dual Riesz system is given by $(\widetilde{\Phi}_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$. In particular,*

$$\langle \Phi_{\theta,d}(\cdot - p), \widetilde{\Phi}_{\theta,d}(\cdot - q) \rangle_{L^2(\mathbb{R}^n)} = \delta_{p,q}.$$

The Riesz bounds of the system are given by A^n, B^n.

Proof First, we show that $(\Phi_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ is a Riesz system with Riesz bounds A^n, B^n. We give the proof for the upper bound for $n = 2$. Observe that by orthogonality of θ_1, θ_2, and with the substitution $y = y_1 d_1 \theta_1 + y_2 d_2 \theta_2$, we get

$$\int_{\mathbb{R}^2} \left| \sum_{p_1,p_2} a_{p_1,p_2} \Phi_{\theta,d}(y - p_1 d_1 \theta_1 - p_2 d_2 \theta_2) \right|^2 dy = \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \sum_{p_1,p_2} a_{p_1,p_2} \varphi(y_1 - p_1) \varphi(y_2 - p_2) \right|^2 dy_1\, dy_2$$

$$\leq B \sum_{p_1} \int_{\mathbb{R}} \left| \sum_{p_2} a_{p_1,p_2} \varphi(y_2 - p_2) \right|^2 dy_2 \leq B^2 \sum_{p_1,p_2} |a_{p_1,p_2}|^2.$$

The bound is optimal: Let $\varepsilon > 0$ and let $(a_{p_1})_{p_1 \in \mathbb{Z}}$ and $(b_{p_2})_{p_2 \in \mathbb{Z}}$ be sequences in ℓ^2 such that

$$\left\| \sum_{p_1 \in \mathbb{Z}} a_{p_1} \varphi(\cdot - p_1) \right\|^2_{L^2(\mathbb{R})} \geq (B - \varepsilon) \sum_{p_1 \in \mathbb{Z}} |a_{p_1}|^2, \quad \left\| \sum_{p_2 \in \mathbb{Z}} b_{p_2} \varphi(\cdot - p_2) \right\|^2_{L^2(\mathbb{R})} \geq (B - \varepsilon) \sum_{p_2 \in \mathbb{Z}} |b_{p_2}|^2.$$

This implies

$$\int_{\mathbb{R}^2} \left| \sum_{p_1,p_2} a_{p_1} b_{p_2} \Phi_{\theta,d}(y - p_1 d_1 \theta_1 - p_2 d_2 \theta_2) \right|^2 dy =$$

$$\int_{\mathbb{R}} \left| \sum_{p_1 \in \mathbb{Z}} a_{p_1} \varphi(y_1 - p_1) \right|^2 dy_1 \cdot \int_{\mathbb{R}} \left| \sum_{p_2 \in \mathbb{Z}} b_{p_2} \varphi(y_2 - p_2) \right|^2 dy_2 \geq (B - \varepsilon)^2 \sum_{p_1,p_2} |a_{p_1} b_{p_2}|^2.$$

In the same fashion, lower and upper bounds for general $n \geq 2$ are obtained, also for the dual Riesz system. We will check the bi-orthogonality condition for $n = 2$:

$$\langle \Phi_{\theta,d}(\cdot - p), \widetilde{\Phi}_{\theta,d}(\cdot - q) \rangle_{L^2(\mathbb{R}^2)} =$$

$$d_1^{-1} d_2^{-1} \int_{\mathbb{R}^2} \varphi(d_1^{-1} \langle y, \theta_1 \rangle - p_1) \varphi(d_2^{-1} \langle y, \theta_2 \rangle - p_2) \times$$

$$\times \overline{\widetilde{\varphi}(d_1^{-1} \langle y, \theta_1 \rangle - q_1)}\, \overline{\widetilde{\varphi}(d_2^{-1} \langle y, \theta_2 \rangle - q_2)}\, dy =$$

$$\int_{\mathbb{R}} \varphi(y_1 - p_1) \overline{\widetilde{\varphi}(y_1 - q_1)}\, dy_1 \cdot \int_{\mathbb{R}} \varphi(y_2 - p_2) \overline{\widetilde{\varphi}(y_2 - q_2)}\, dy_2 = \delta_{p_1,q_1} \cdot \delta_{p_2,q_2} = \delta_{p,q}.$$

□

A *multi-resolution analysis in $L^2(\mathbb{R}^n)$ associated with a grid $\mathbb{G}_{\theta,d}$* is a sequence of subspaces $(V_{\theta,d,j})_{j \in \mathbb{Z}}$ in $L^2(\mathbb{R}^n)$ with the properties: For each $j \in \mathbb{Z}$, we have

(i) $V_{\theta,d,j} \subset V_{\theta,d,j+1}$,

(ii) $f \in V_{\theta,d,j} \;\Leftrightarrow\; f(2\cdot) \in V_{\theta,d,j+1}$,

and

(iii) $\bigcap_{j \in \mathbb{Z}} V_{\theta,d,j} = (0)$,

(iv) $\bigcup_{j \in \mathbb{Z}} V_{\theta,d,j} \subseteq L^2(\mathbb{R}^n)$ dense,

(v) there exists $\Phi \in V_{\theta,d,0}$ such that $(\Phi(\cdot - p))_{p \in \mathbb{G}_{\theta,d}}$ is a Riesz basis in $V_{\theta,d,0}$.

Lemma 5.2 *If φ induces a multi-resolution analysis in $L^2(\mathbb{R})$, then $\Phi_{\theta,d}$ induces a multi-resolution analysis in $L^2(\mathbb{R}^n)$ associated with the grid $\mathbb{G}_{\theta,d}$.*

Proof Define $V_{\theta,d,j}$ as the closed linear span of $(2^{nj/2}\Phi_{\theta,d}(2^j\cdot -p))_{p\in\mathbb{G}_{\theta,d}}$. If φ satisfies $\varphi(x) = \sum_{k\in\mathbb{Z}} h_k\sqrt{2}\varphi(2x-k)$ for almost all $x\in\mathbb{R}$, then

$$\Phi_{\theta,d}(y) = \prod_{r=1}^{n}\sqrt{\frac{1}{d_r}}\varphi\left(d_r^{-1}\langle y,\theta_r\rangle\right) =$$

$$\prod_{r=1}^{n}\left(\sqrt{\frac{2}{d_r}}\sum_{p_r\in\mathbb{Z}}\varphi(2d_r^{-1}\langle y,\theta_r\rangle - p_r)\right) = \sum_{p\in\mathbb{G}_{\theta,d}} h_p 2^{n/2}\Phi_{\theta,d}(2y-p),$$

where $p=\sum_{r=1}^{n} d_r p_r\theta_r$ and $h_p = h_{p_1}\cdots h_{p_n}$. In this manner, it follows that properties (i) and (ii) hold. Fubini's theorem implies that, for $f\in L^2(\mathbb{R}^n)$, $f\in V_{\theta,d,j}$ if and only if $f(x+\cdot\theta_r)\in V_j$ for almost all $x\in\theta_r^{\perp}$ and $r=1,\dots,n$. Properties (iii) and (iv) now follow easily. Property (v) is obtained by using Lemma 5.1. □

We shall introduce a multi-variate interpolation operator which incorporates the Riesz system induced by the multi-variate scaling function. Define $L_{\theta,d}:\ell^2(\mathbb{G}_{\theta,d})\to L^2(\mathbb{R}^n)$ by

$$L_{\theta,d}(a_p)_{p\in\mathbb{G}_{\theta,d}}(y) = \sum_{p\in\mathbb{G}_{\theta,d}} a_p\Phi_{\theta,d}(y-p)$$

for almost all $y\in\mathbb{R}^n$. The operator $\widetilde{L}_{\theta,d}$ is defined in the same way using the dual scaling function. Observe that by Lemma 5.1, we get

$$A^n\|(a_p)_{p\in\mathbb{G}_{\theta,d}}\|^2_{\ell^2(\mathbb{G}_{\theta,d})} \le \|L_{\theta,d}(a_p)_{p\in\mathbb{G}_{\theta,d}}\|^2_{L^2(\mathbb{R}^n)} \le B^n\|(a_p)_{p\in\mathbb{G}_{\theta,d}}\|^2_{\ell^2(\mathbb{G}_{\theta,d})}.$$

This implies that $L_{\theta,d}:\ell^2(\mathbb{G}_{\theta,d})\to L^2(\mathbb{R}^n)$ is a bounded injective operator with closed range. It follows that $L_{\theta,d}$ has a bounded left-inverse. We shall discuss these matters in more detail below. The adjoint operator $L^*_{\theta,d}:L^2(\mathbb{R}^n)\to\ell^2(\mathbb{G}_{\theta,d})$ of $L_{\theta,d}$ is given by

$$(L^*_{\theta,d}g) = \left(\langle g,\Phi_{\theta,d}(\cdot-p)\rangle_{L^2(\mathbb{R}^n)}\right)_{p\in\mathbb{G}_{\theta,d}},\quad g\in L^2(\mathbb{R}^n).$$

We now identify a left inverse of $L_{\theta,d}$.

Proposition 5.3 *The operator $L_{\theta,d}:\ell^2(\mathbb{G}_{\theta,d})\to L^2(\mathbb{R}^n)$ has $\widetilde{L}^*_{\theta,d}:L^2(\mathbb{R}^n)\to\ell^2(\mathbb{G}_{\theta,d})$ as a left inverse.*

Proof Note that Lemma 5.1 and the continuity of the inner product implies

$$\langle\widetilde{L}^*_{\theta,d}L_{\theta,d}(a_p)_{p\in\mathbb{G}_{\theta,d}},(b_q)_{q\in\mathbb{G}_{\theta,d}}\rangle_{\ell^2(\mathbb{G}_{\theta,d})} = \langle L_{\theta,d}(a_p)_{p\in\mathbb{G}_{\theta,d}},\widetilde{L}_{\theta,d}(b_q)_{q\in\mathbb{G}_{\theta,d}}\rangle_{L^2(\mathbb{R}^n)} =$$

$$\langle\sum_{p\in\mathbb{G}_{\theta,d}} a_p\Phi_{\theta,d}(\cdot-p),\sum_{q\in\mathbb{G}_{\theta,d}} b_q\widetilde{\Phi}_{\theta,d}(\cdot-q)\rangle_{L^2(\mathbb{R}^n)} = \sum_{p\in\mathbb{G}_{\theta,d}} a_p\overline{b_p}.$$

□

In the particular case when $\Phi_{\theta,d}=\widetilde{\Phi}_{\theta,d}$, i.e., in the orthonormal case, we obviously get $L_{\theta,d}=\widetilde{L}_{\theta,d}$ and by Proposition 5.3, we see that in this case $L^*_{\theta,d}L_{\theta,d} = I_{\ell^2(\mathbb{G}_{\theta,d})}$. In the general situation, the self-adjoint operator $L^*_{\theta,d}L_{\theta,d}$ will not be the identity, although the following holds true.

Lemma 5.4 *The operator $L^*_{\theta,d}L_{\theta,d}:\ell^2(\mathbb{G}_{\theta,d})\to\ell^2(\mathbb{G}_{\theta,d})$ is a strictly positive, hence boundedly invertible, operator which satisfies the estimates (in the sense of self-adjoint operators)*

$$A^n\,I \le L^*_{\theta,d}L_{\theta,d} \le B^n\,I.$$

The proof of this lemma is straightforward and omitted. Observe that the expression

$$(f_p)_{p\in\mathbb{G}_{\theta,d}} = (L^*_{\theta,d}L_{\theta,d})^{-1}L^*_{\theta,d}g \tag{5.1}$$

makes sense and provides the least-squares solution to the equation

$$L_{\theta,d}(f_p)_{p\in\mathbb{G}_{\theta,d}} = g \tag{5.2}$$

where $g \in L^2(\mathbb{R}^n)$ is a given function. Another approximate solution to (5.2) is given by

$$(f_p)_{p\in\mathbb{G}_{\theta,d}} = \widetilde{L}^*_{\theta,d}g. \tag{5.3}$$

We will now look at equation (5.2) and its approximate solutions (5.1) and (5.3) more closely. If we apply the operator $L_{\theta,d}$ to the right hand side of (5.1), we get $L_{\theta,d}(L^*_{\theta,d}L_{\theta,d})^{-1}L^*_{\theta,d}g$. The operator $\Pi_{\theta,d} = L_{\theta,d}(L^*_{\theta,d}L_{\theta,d})^{-1}L^*_{\theta,d}$ is the orthogonal projection onto ran $L_{\theta,d}$. The *least squares solution* (5.1) produces an error $\|g - \Pi_{\theta,d}g\|_{L^2(\mathbb{R}^n)}$ which equals zero whenever $g \in$ ran $L_{\theta,d}$. In general, this error is minimal among all attainable errors, since $\|g - \Pi_{\theta,d}g\|_{L^2(\mathbb{R}^n)}$ equals the distance between g and ran $L_{\theta,d}$.

On the other hand, if we apply $L_{\theta,d}$ to the right hand side of (5.3), we get $L_{\theta,d}\widetilde{L}^*_{\theta,d}g$. The operator $P_{\theta,d} = L_{\theta,d}\widetilde{L}^*_{\theta,d}$ is a not necessarily orthogonal projection onto ran $L_{\theta,d}$. This *dual solution* produces an error $\|g - P_{\theta,d}g\|_{L^2(\mathbb{R}^n)}$. Since $g - \Pi_{\theta,d}g \perp \Pi_{\theta,d}g - P_{\theta,d}g$, we get

$$\|g - P_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)} = \|P_{\theta,d}g - \Pi_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)} + \|g - \Pi_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)}.$$

It is immediate that the error caused by the least squares solution is majorized by the error caused by the dual solution. Moreover, the difference between the two errors can be measured by $\|P_{\theta,d}g - \Pi_{\theta,d}g\|_{L^2(\mathbb{R}^n)}$. First of all, we recall from Section 2 that $P_{\theta,d} = \Pi_{\theta,d}$ whenever $\varphi, \widetilde{\varphi} \in V$. Secondly, we will use the following result to pursue these matters somewhat further. We prove the analogue of Proposition 2.6 for multi-variate scaling functions.

Theorem 5.5 *Let $\Phi_{\theta,d}$ and $\widetilde{\Phi}_{\theta,d}$ be multi-variate scaling functions as defined before which satisfy the bi-orthogonality condition*

$$\langle \Phi_{\theta,d}(\cdot - p), \widetilde{\Phi}_{\theta,d}(\cdot - q) = \delta_{p,q}, \quad p, q \in \mathbb{G}_{\theta,d},$$

and let β be defined as in Proposition 2.6. The closed linear span of $(\Phi_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ is denoted by $V_{\theta,d}$ and the closed linear span of $(\widetilde{\Phi}_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ is denoted by $\widetilde{V}_{\theta,d}$. The orthoprojector onto $V_{\theta,d}$ is given by $\Pi_{\theta,d}$, and the projection onto $V_{\theta,d}$ along $\widetilde{V}^{\perp}_{\theta,d}$ reads $P_{\theta,d}$. We get for $g \in L^2(\mathbb{R}^n)$,

$$\|g - \Pi_{\theta,d}g\|_{L^2(\mathbb{R}^n)} \le \|g - P_{\theta,d}g\|_{L^2(\mathbb{R}^n)} \le \beta^n\|g - \Pi_{\theta,d}g\|_{L^2(\mathbb{R}^n)}.$$

Before we prove the theorem, we state the following consequence of its result: Observe that

$$\|P_{\theta,d}g - \Pi_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)} = \|g - P_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)} - \|g - \Pi_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)} \le$$

$$(\beta^{2n} - 1)\|g - \Pi_{\theta,d}g\|^2_{L^2(\mathbb{R}^n)}.$$

Moreover, $\beta \ge 1$, and $\beta = 1$ if and only if $V = \widetilde{V}$. In this manner, the coefficient $\beta^{2n} - 1$ measures the relative distance between the dual solution and the least squares solution error.

Proof of Theorem 5.5 Let φ^o and $\widetilde{\varphi}^o$ induce orthonormal bases in V and $\widetilde{V}$, respectively. Define

$$\Phi^o_{\theta,d} = \prod_{r=1}^n d_r^{-1/2}\varphi^o(d_r^{-1}\langle\cdot,\theta_r\rangle), \quad \widetilde{\Phi}^o_{\theta,d} = \prod_{r=1}^n d_r^{-1/2}\widetilde{\varphi}^o(d_r^{-1}\langle\cdot,\theta_r\rangle).$$

Observe that $(\Phi^o_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ and $(\widetilde{\Phi}^o_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ are orthonormal bases in $V_{\theta,d}$ and $\widetilde{V}_{\theta,d}$, respectively. Each $f \in V_{\theta,d}$ admits an orthonormal expansion $f = \sum_{p\in\mathbb{G}_{\theta,d}} c_p\Phi^o_{\theta,d}(\cdot - p)$, with $c_p = \langle f, \Phi^o_{\theta,d}(\cdot - p)\rangle_{L^2(\mathbb{R}^n)}$. As a consequence, if $f \in V_{\theta,d}$, we get

$$\widetilde{\Pi}_{\theta,d}f = \sum_{p,q\in\mathbb{G}_{\theta,d}} c_p\langle\Phi^o_{\theta,d}(\cdot - p), \widetilde{\Phi}^o_{\theta,d}(\cdot - q)\rangle_{L^2(\mathbb{R}^n)}\widetilde{\Phi}^o_{\theta,d}(\cdot - q).$$

Here $\widetilde{\Pi}_{\theta,d}$ denotes the orthoprojector onto $\widetilde{V}_{\theta,d}$. The sequence $(d_q)_{q\in\mathbb{G}_{\theta,d}}$, which is defined by $d_q = \sum_{p\in\mathbb{G}_{\theta,d}} c_p\langle\Phi^o_{\theta,d}(\cdot - p), \widetilde{\Phi}^o_{\theta,d}(\cdot - q)\rangle_{L^2(\mathbb{R}^n)}$, is a discrete convolution $d = \rho * c$, where $\rho = (\rho_q)_{q\in\mathbb{G}_{\theta,d}}$ is given by $\rho_q = \langle\Phi^o_{\theta,d}(\cdot + q), \widetilde{\Phi}^o_{\theta,d}\rangle_{L^2(\mathbb{R}^n)}$ and $c = (c_p)_{p\in\mathbb{G}_{\theta,d}}$.

Recall that if $a = (a_p)_{p\in\mathbb{G}_{\theta,d}} \in \ell^2(\mathbb{G}_{\theta,d})$, then $\widehat{a}$, defined through $\widehat{a}(\eta) = \sum_{p\in\mathbb{G}_{\theta,d}} a_p e^{-i\langle\eta,p\rangle}$ for $\eta \in \mathbb{R}^n$, satisfies $\widehat{a} \in L^2([0, 2\pi]^n)$. Using this notation, we get $\widehat{d} = \widehat{\rho}\cdot\widehat{c}$. Therefore, if $f \in V_{\theta,d}$, we get

$$\|\widetilde{\Pi}_{\theta,d}f\|^2_{L^2(\mathbb{R}^n)} = \sum_{q\in\mathbb{G}_{\theta,d}} |d_q|^2 = (2\pi)^{-n}\int_{[0,2\pi]^n} |\widehat{d}(\eta)|^2\,d\eta = (2\pi)^{-n}\int_{[0,2\pi]^n} |\widehat{\rho}(\eta)|^2|\widehat{c}(\eta)|^2\,d\eta,$$

while

$$\|f\|^2_{L^2(\mathbb{R}^n)} = (2\pi)^{-n}\int_{[0,2\pi]^n} |\widehat{c}(\eta)|^2\,d\eta.$$

Write

$$\eta = \sum_{r=1}^n \eta_r d_r^{-1}\theta_r, \quad q = \sum_{r=1}^n q_r d_r\theta_r,$$

then $\langle\eta, q\rangle = \sum_{r=1}^n \eta_r q_r$. Moreover,

$$\rho_q = \langle\Phi^o_{\theta,d}(\cdot + q), \widetilde{\Phi}^o_{\theta,d}\rangle_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n}\prod_{r=1}^n d_r^{-1}\varphi^o(d_r^{-1}\langle y,\theta_r\rangle + q_r)\overline{\widetilde{\varphi}^o(d_r^{-1}\langle y,\theta_r\rangle)}\,dy =$$

(write $y = \sum_{r=1}^n y_r d_r\theta_r$)

$$\prod_{r=1}^n\int_{\mathbb{R}}\varphi^o(y_r + q_r)\overline{\widetilde{\varphi}^o(y_r)}\,dy_r = \prod_{r=1}^n \rho_{q_r}.$$

This implies

$$\widehat{\rho}(\eta) = \prod_{r=1}^n\left(\sum_{q_r\in\mathbb{Z}}\rho_{q_r}e^{-\eta_r q_r}\right) = \prod_{r=1}^n\widehat{\rho}_r(\eta_r).$$

We conclude that

$$\beta_{\theta,d} = \left(\inf_{0\neq f\in V_{\theta,d}}\frac{\|\widetilde{\Pi}_{\theta,d}f\|_{L^2(\mathbb{R}^n)}}{\|f\|_{L^2(\mathbb{R}^n)}}\right)^{-1} \leq \left(\operatorname{ess\,inf}_{\eta\in[0,2\pi]^n}|\widehat{\rho}(\eta)|\right)^{-1} = \prod_{r=1}^n\left(\operatorname{ess\,inf}_{\eta_r\in[0,2\pi]}|\widehat{\rho}_r(\eta_r)|\right)^{-1}$$

$$= \beta^n,$$

by the proof of Proposition 2.5. We may now apply Proposition 2.6. □

In proposition 2.5, the function Z was introduced. If we set ($\eta \in \mathbb{R}^n$)

$$Z_{\theta,d}(\eta) = \frac{\sum_{p\in\mathbb{G}_{\theta,d}} \widehat{\Phi}_{\theta,d}(\eta+2\pi p)\overline{\widehat{\widetilde{\Phi}}_{\theta,d}(\eta+2\pi p)}}{\sqrt{\sum_{p\in\mathbb{G}_{\theta,d}} |\widehat{\Phi}_{\theta,d}(\eta+2\pi p)|^2}\sqrt{\sum_{p\in\mathbb{G}_{\theta,d}} |\widehat{\widetilde{\Phi}}_{\theta,d}(\eta+2\pi p)|^2}},$$

then, with $\eta = \sum_{r=1}^n \eta_r d_r^{-1}\theta_r$ and $p = \sum_{r=1}^n \eta_r d_r^{-1}\theta_r$,

$$Z_{\theta,d}(\eta) = \prod_{r=1}^n \frac{\sum_{p_r\in\mathbb{Z}} \widehat{\varphi}(\eta_r+2\pi p_r)\overline{\widehat{\widetilde{\varphi}}(\eta_r+2\pi p_r\rangle)}}{\sqrt{\sum_{p_r\in\mathbb{Z}} |\widehat{\varphi}(\eta_r+2\pi p_r)|^2}\sqrt{\sum_{p_r\in\mathbb{Z}} |\widehat{\widetilde{\varphi}}(\eta_r+2\pi p_r)|^2}} = \prod_{r=1}^n Z(\eta_r).$$

Theorem 5.6 *Let $\Pi_{\theta,d}$ be defined the orthoprojector onto $V_{\theta,d}$ and assume that $f \in H^{p+1}(\mathbb{R}^n)$, i.e., all partial derivatives of f up to order $p+1$ are in $L^2(\mathbb{R}^n)$. Then*

$$\|f - \Pi_{\theta,d} f\|_{L^2(\mathbb{R}^n)} \le C \sum_{r=1}^n d_r^{p+1} \|\partial_r^{p+1} f\|_{L^2(\mathbb{R}^n)},$$

where ∂_r denotes the r-th partial derivative operator.

Proof An orthonormal basis $(\Phi^o_{\theta,d}(\cdot - p))_{p\in\mathbb{G}_{\theta,d}}$ of $V_{\theta,d}$ is constructed from an orthonormal basis $(\varphi^o(\cdot - k))_{k\in\mathbb{Z}}$ as in the proof of Theorem 5.5. Define for $r = 1,\dots,n$ the operator $\Pi^{(r)}_{\theta,d}$ by

$$\Pi^{(r)}_{\theta,d} f(y) = \sum_{k\in\mathbb{Z}} d_r^{-1} \int_{\mathbb{R}} f(E_{\theta_r} y + t\theta_r)\varphi^o(d_r^{-1}t - k)\,dt \cdot \varphi^o(d_r^{-1}\langle y,\theta_r\rangle - k), \quad y \in \mathbb{R}^n.$$

It is not difficult to see that $\Pi^{(r)}_{\theta,d}$ are mutually commuting orthogonal projections which satisfy

$$\prod_{r=1}^n \Pi^{(r)}_{\theta,d} = \Pi_{\theta,d}.$$

Moreover, we get for each $1 \le r \le n$,

$$\|f - \Pi^{(r)}_{\theta,d} f\|^2_{L^2(\mathbb{R})} = \int_{\theta_r^\perp} \|f(x + \cdot\theta_r) - \Pi_{d_r} f(x + \cdot\theta_r)\|^2_{L^2(\mathbb{R})}\,dx \le$$

$$C^2 d_r^{2(p+1)} \|\partial_r^{p+1} f\|^2_{L^2(\mathbb{R}^n)},$$

by Theorem 3.1. These considerations imply that

$$\|f - \Pi_{\theta,d} f\|_{L^2(\mathbb{R}^n)} \le$$

$$\|f - \Pi^{(1)}_{\theta,d} f\|_{L^2(\mathbb{R}^n)} + \|\Pi^{(1)}_{\theta,d}(f - \Pi^{(2)}_{\theta,d} f)\|_{L^2(\mathbb{R}^n)} + \cdots + \|\prod_{r=1}^{n-1} \Pi^{(r)}_{\theta,d}(f - \Pi^{(n)}_{\theta,d} f)\|_{L^2(\mathbb{R}^n)} \le$$

$$\sum_{r=1}^n C d_r^{p+1} \|\partial_r^{p+1} f\|_{L^2(\mathbb{R}^n)}.$$

□

6. Interpolation between rotated grids

We shall now apply the results from the preceding sections to interpolation between dilated, reflected and rotated Cartesian grids. The situation is as follows. In Section 5, we mentioned that the standard grid

$$\mathbb{G}_e = \{\sum_{r=1}^{n} q_r e_r \mid (q_1, \dots, q_n) \in \mathbb{Z}^n\} \subseteq \mathbb{R}^n$$

is defined using the standard orthonormal basis $e_1, \dots, e_n$, and that for any orthonormal basis $\theta_1, \dots, \theta_n$ and vector with strictly positive entries $d = (d_1, \dots, d_n)^T$, we may define the grid

$$\mathbb{G}_{\theta,d} = \{\sum_{r=1}^{n} p_r d_r e_r \mid (p_1, \dots, p_n) \in \mathbb{Z}^n\}.$$

Given is a collection of data (complex numbers) $(a_q)_{q\in\mathbb{G}_e}$, and a function $g \in L^2(\mathbb{R}^n)$ is attached to this collection by means of interpolation on the standard grid. Indeed, we assume that

$$g(y) = \sum_{q\in\mathbb{G}_e} a_q \Phi_e(y-q), \quad \text{almost all } y \in \mathbb{R}^n.$$

Here $\Phi_e(y) = \prod_{r=1}^{n} \varphi(\langle y, \theta \rangle)$, for almost all $y \in \mathbb{R}^n$, as in Section 5. The dual function $\widetilde{\Phi}_e$ is defined accordingly. Note that we have constructed a function g such that

$$a_q = \langle g, \widetilde{\Phi}_e(\cdot - q)\rangle_{L^2(\mathbb{R}^n)}, \quad q \in \mathbb{G}_e.$$

In words, the function g reproduces the original data when inner products are taken with translates along the standard grid $\mathbb{G}_e$ of the dual multi-variate function $\widetilde{\Phi}_e$. We now search for a collection of data $(f_p)_{p\in\mathbb{G}_{\theta,d}}$ associated with the dilated, reflected and rotated grid $\mathbb{G}_{\theta,d}$. This data should reproduce the function g as an interpolant of $(f_p)_{p\in\mathbb{G}_{\theta,d}}$. In general, this is not possible. Therefore, we shall write

$$f(y) = \sum_{p\in\mathbb{G}_{\theta,d}} f_p \Phi_{\theta,d}(y-p), \quad \text{almost all } y \in \mathbb{R}^n,$$

and we shall try to minimize $\|f-g\|_{L^2(\mathbb{R}^n)}$.

In Section 5, we defined $\Phi_{\theta,d}(y) = \prod_{r=1}^{n} d_r^{-1/2} \varphi(d_r^{-1}\langle y, \theta_r\rangle)$, which implies that if $y = \sum_{r=1}^{n} y_r \theta_r$, then $\Phi_{\theta,d}(y) = \prod_{r=1}^{n} d_r^{-1/2} \varphi(d_r^{-1} y_r)$. In terms of the interpolation operators (details are given in Section 5)

$$L_e(a_q)_{q\in\mathbb{G}_e} = \sum_{q\in\mathbb{G}_e} a_q \Phi_e(y-q), \quad L_{\theta,d}(f_p)_{p\in\mathbb{G}_{\theta,d}} = \sum_{p\in\mathbb{G}_{\theta,d}} f_p \Phi_{\theta,d}(y-p),$$

we have the following situation. Given the data $(a_q)_{q\in\mathbb{G}_e} \in \ell^2(\mathbb{G}_e)$, we construct $g = L_e(a_q)_{q\in\mathbb{G}_e}$. The least squares solution to

$$L_{\theta,d}(f_p)_{p\in\mathbb{G}_{\theta,d}} = g = L_e(a_q)_{q\in\mathbb{G}_e}$$

provides us with data $(f_p)_{p\in\mathbb{G}_{\theta,d}}$ on the rotated and dilated grid, such that $f = L_{\theta,d}(f_p)_{p\in\mathbb{G}_{\theta,d}}$ minimizes $\|f-g\|_{L^2(\mathbb{R}^n)}$ among all possible solutions.

We propose the dual method

$$(f_p)_{p\in\mathbb{G}_{\theta,d}} = \widetilde{L}^*_{\theta,d} g = \widetilde{L}^*_{\theta,d} L_e(a_q)_{q\in\mathbb{G}_e},$$

whic requires a number of computations proportional to the number of data involved, as will be shown at the end of this section. Moreover, the dual solution can be taken equal to or arbitrarily close to the least squares solution, depending on the choice of the dual wavelet. This fact is expressed by the consequence mentioned after Theorem 5.5.

We shall describe the dual method by means of the operator

$$\widetilde{L}^*_{\theta,d} L_e : \ell^2(\mathbb{G}_e) \to \ell^2(\mathbb{G}_{\theta,d}).$$

Observe that for $p \in \mathbb{G}_{\theta,d}$,

$$\left(L^*_{\theta,d} L_e (a_q)_{q\in\mathbb{G}_e}\right)_p = \int_{\mathbb{R}^n} L_e(a_q)_{q\in\mathbb{G}_e}(y)\overline{\widetilde{\Phi}_{\theta,d}(y-p)}\,dy =$$

$$\int_{\mathbb{R}^n} \sum_{q\in\mathbb{G}_e} a_q \Phi_e(y-q)\overline{\widetilde{\Phi}_{\theta,d}(y-p)}\,dy = \sum_{q\in\mathbb{G}_e} a_q \langle \Phi_e(\cdot - q), \widetilde{\Phi}_{\theta,d}(\cdot - p)\rangle_{L^2(\mathbb{R}^n)}.$$

If we define the infinite matrix $M = (M_{p,q})_{p\in\mathbb{G}_{\theta,d}, q\in\mathbb{G}_e}$, with matrix elements given by the inner products $M_{p,q} = \langle \Phi_e(\cdot - q), \widetilde{\Phi}_{\theta,d}(\cdot - p)\rangle_{L^2(\mathbb{R}^n)}$, then the solution to the dual method is given by

$$(f_p)_{p\in\mathbb{G}_{\theta,d}} = M\,(a_q)_{q\in\mathbb{G}_e}.$$

Note that $M_{p,q}$ does only depend on the difference between the indices. Indeed, if $\delta = \sum_{r=1}^n d_r\delta_r\theta_r \in \mathbb{G}_{\theta,d}$, then

$$M_{p+\delta,q+\delta} = \int_{\mathbb{R}^n} \Phi_e(y-q-\delta)\overline{\widetilde{\Phi}_{\theta,d}(y-p-\delta)}\,dy \int_{\mathbb{R}^n} \Phi_e(y-q)\overline{\widetilde{\Phi}_{\theta,d}(y-p)}\,dy = M_{p,q}.$$

In order to describe the number of calculations required to perform the dual method, we consider the case when the data set $(a_q)_{q\in\mathbb{G}_e}$ is finite. In particular, we denote a compact set in $\mathbb{R}^n$ by K and assume that $a_q = 0$ for $q \notin K$. Further, we will assume that the underlying scaling functions φ and $\widetilde{\varphi}$ have compact support. Indeed, assume that $\rho, \widetilde{\rho} > 0$ are chosen such that supp $\varphi \subseteq [-\rho, \rho]$ and supp $\widetilde{\varphi} \subseteq [-\widetilde{\rho}, \widetilde{\rho}]$. Let $(p_1, \dots, p_n) \in \mathbb{Z}^n$ and $(q_1, \dots, q_n) \in \mathbb{Z}^n$, then $q = \sum_{r=1}^n q_r e_r \in \mathbb{G}_e$ and $p = \sum_{r=1}^n p_r\theta_r d_r \in \mathbb{G}_{\theta,d}$. The matrix element

$$M_{p,q} = \int_{\mathbb{R}^n} \prod_{r=1}^n \varphi(\langle y, e_r\rangle - q_r) d_r^{-1/2}\overline{\widetilde{\varphi}(d_r^{-1}\langle y, \theta_r\rangle - p_r)}\,dy$$

is nonzero, only if

$$|\langle y, e_r\rangle - q_r| \le \rho, \quad |d_r^{-1}\langle y, \theta_r\rangle - p_r| \le \widetilde{\rho}, \quad r = 1, \dots, n.$$

We get

$$\|y-q\|^2 = \sum_{r=1}^n |\langle y, e_r\rangle - q_r|^2 \le \rho^2 n,$$

$$\|y-p\|^2 = \sum_{r=1}^n |\langle y, \theta_r\rangle - d_r p_r|^2 = \sum_{r=1}^n d_r^2 |d_r^{-1}\langle y, \theta_r\rangle - p_r|^2 \le \widetilde{\rho}^2 \|d\|^2,$$

where $\|d\|$ is the Euclidian norm of $d = (d_1, \dots, d_n)^T$. This provides

$$0 = \|(y-q) + (q-p) + (p-y)\| \ge \|p-q\| - \|y-p\| - \|y-q\| \ge \|p-q\| - (\rho\sqrt{n} + \widetilde{\rho}\|d\|).$$

As a result, we see that

$$\|p-q\| \leq (\rho\sqrt{n} + \widetilde{\rho}\|d\|).$$

We may conclude that if $\|p-q\| > (\rho\sqrt{n} + \widetilde{\rho}\|d\|)$, then $M_{p,q} = 0$. This means that the matrix M has a *band structure* with *band width* majorized by $(\rho\sqrt{n} + \widetilde{\rho}\|d\|)$. The number of data N, i.e., the number of grid points of $\mathbb{G}_{\theta,d}$ in K, is proportional to $|K|\Delta^{-1}$, where $\Delta = d_1 \ldots d_n$ and $|K|$ is the volume of K. The number of computations required to calculate $(f_p)_{p\in\mathbb{G}_{\theta,d}}$ is majorized by the number $N(\rho\sqrt{n} + \widetilde{\rho}\|d\|)$ which is of the same order as the number of data involved.

References

[Ba] G. Battle, A block spin construction of ondelettes, Part I: Lemarié functions, *Comm. Math. Phys.* 110: 601-615 (1987).

[Chu] C.K. Chui, An introduction to wavelets, Academic Press, Boston (1992).

[CDF] A. Cohen, I. Daubechies, J.-C. Feauveau, Bi-orthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* 45: 485-560 (1992).

[CR] A. Cohen, R.D. Ryan, Wavelets and multiscale signal processing, Chapmann & Hall, London (1995).

[Dau] I. Daubechies, Ten lectures on wavelets, SIAM-CBMS Series 61, Philadelphia, PA (1992).

[Dau2] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* 41 (7): 909-996 (1988).

[FKV] L. Faqi, M.M.N. Kabir, D.J. Verschuur, Cascaded application of the linear Radon and wavelet transform in preprocessing, *Annual Meeting Society Exploration Geophysicists, Expanded Abstracts*: 1373-1376 (1995).

[GL] G.H. Golub, C.F. van Loan, Matrix Computations, John Hopkins University Press (1996).

[HL] E.J. Halton, W.A. Light, On local and controlled approximation order, *J. Approx. Th.* 72: 268-277 (1993).

[Hol] M. Holschneider, Wavelets: An analysis tool, Clarendon Press, Oxford (1995).

[JL] R.-Q. Jia, J. Lei, Approximation by multi-integer translates of functions having global support, *J. Approx. Th.* 72: 2-23 (1993).

[KS] G. Kaiser, R.F. Streater, Windowed Radon transforms, analytic signals, and the wave equation, Wavelets: A tutorial in theory and applications, C.K. Chui (ed.), pp. 399-441, Academic Press (1992).

[Koo] T.H. Koornwinder, The continuous wavelet transform, Wavelets: An elementary treatment of theory and applications, T.H. Koornwinder (ed.), pp. 27-48, World Scientific (1993).

[Le] P.G. Lemarié, Ondelettes à localisation exponentielles, *J. Math. Pure et Appl.* 67: 227-236 (1988).

[LMR] A.K. Louis, P. Maass, A. Rieder, Wavelets: Theory and Applications, John Wiley & Sons, Chicester (1997).

[Mal] S.G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$, *Trans. A.M.S.* 315(1): 69-87 (1989).

[Mal2] S.G. Mallat, A theory for multiresolution signal decomposition, *IEEE Tr. Pattern Anal. Machine Intelligence* 11: 674-693 (1989).

[Mey] Y. Meyer, R.D. Ryan, Wavelets: Algorithms and applications, SIAM, Philadelphia (1993).

[MPO] E. Magli, L. Lo Presti, G. Olmo, A joint compression and target detection algorithm for aerial images, unpublished (1996).

[Nat] F. Natterer, The mathematics of computerized tomography, John Wiley & Sons, Chicester (1986).

[RV] O. Rioul, M. Vetterli, Wavelets and signal processing, *IEEE Signal Processing Magazine*, October: 14-38 (1991).

[Ser] E.P. Serrano, Some remarks about compactly supported spline wavelets, *Appl. Comp. Harm. Anal.* 3: 57-64 (1996).

[SSW] K.T. Smith, D.C. Solmon, S.L. Wagner, Practical and mathematical aspects of the problem of reconstructing objects from radiographs *Bull. A.M.S.* 82(6): 1227-1270 (1977).

[SF] G. Strang, G. Fix, A Fourier analysis of the finite element variational method, in: Constructive aspects of functional analysis, G. Geymonat, ed., pp. 793-840, CIME (1973).

[SN] G. Strang, T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press (1996).

[Tak] T. Takiguchi, On invertibility of the windowed Radon transform, *J. Math. Sci. Univ. Tokyo* 2: 621-636 (1995).

[Uns] M. Unser, Quasi-orthogonality and quasi-projections, *Appl. Comp. Harm. Anal.* 3: 201-214 (1996).

[Uns2] M. Unser, Approximation power of bi-orthogonal wavelet expansions, *IEEE Tr. Signal Proc.* 44(3): 519-527 (1996).

[UA] M. Unser, A. Aldroubi, A general sampling theory for nonideal acquisition devices, *IEEE Tr. Signal Proc.* 42 (11): 2915-2925 (1994).

[UD] M. Unser, I. Daubechies, On the approximation power of convolution-based least squares versus interpolation, *IEEE Tr. Signal Proc.* 45(7): 1697-1711 (1997).

[You] R.M. Young, An introduction to nonharmonic Fourier series, Academic Press, New York (1980).

[WD] A.L. Warrick, P.A. Delaney, A wavelet localized Radon transform, SPIE Proceedings 2569, Wavelet Applications in Signal and Image Processing III: 632-643 (1995).

[ZZ] P.M. de Zeeuw, R.A. Zuidwijk, Fast algorithm for directional time-scale analysis using wavelets, SPIE proceedings 3458, Wavelet Applications in Signal and Image Processing VI (1998), pp. 222-231.

[ZZ2] P.M. de Zeeuw, R.A. Zuidwijk, Numerical methods for decomposition of 2D signals by rotation and wavelet techniques, CWI report PNA-R99xx (1999).

[Zui] R.A. Zuidwijk, The wavelet X-ray transform, CWI-report PNA-R9703 (1997).

[Zui2] R.A. Zuidwijk, The discrete and continuous wavelet X-ray transform, SPIE proceedings 3169, Wavelet Applications in Signal and Image Processing V (1997), pp. 357-366.