A Multiresolution Approach to Fast Summation and Regularization of Singular Operators

by

Robert L. Cramer
B.Sc., Metropolitan State College, 1989
M.Sc., University of Colorado, 1995

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Department of Applied Mathematics
1996

This thesis for the Doctor of Philosophy degree by
Robert L. Cramer
has been approved for the
Department of
Applied Mathematics
by

Gregory Beylkin

Jerrold Bebernes

William Briggs

James Curry

Bengt Fornberg

Date

Cramer, Robert L. (Ph.D., Applied Mathematics)

A Multiresolution Approach to Fast Summation and Regularization of Singular Operators

Thesis directed by Professor Gregory Beylkin

A class of fast algorithms is introduced for the evaluation of discrete sums that utilizes projections on a multiresolution analysis. The discrete sums under consideration arise, for example, in the study of physical systems by means of particle simulations requiring long-range potentials. These include gravitational and electrostatic models, plasma physics, atmospheric physics, and vortex methods in fluid dynamics. These numerical models of particle interactions require the application of dense matrices which, done directly, requires $O(N^2)$ arithmetic operations. The algorithms we develop accomplish this task to within accuracy ϵ in $O(N)$ arithmetic operations.

There are two types of algorithms used today for the fast computation of discrete sums, namely, the Method of Local Corrections and the Fast Multipole Method. Our approach is related to both, but has its own unique features. We describe implementations in one and two dimensions, and present theoretical foundations for algorithms in higher dimensions.

In our approach to discrete summation problems, we construct explicit representations of singular operators on subspaces of the multiresolution analysis. These representations provide a definition for the regularization of such operators, as well as a practical algorithm for their computation. We present a new multiresolution approach to the regularization of singular operators, and show that our method coincides with the classical method, where the classical method is applicable.

Contents

Chapter 1

Introduction

1.1 Introductory Remarks

The study of physical systems by means of particle simulations is an important computational tool in many fields. Examples include plasma physics, atmospheric physics, N-body gravitational problems, and vortex methods in fluid dynamics. In most of these particle models the evaluation of discrete sums describing the pairwise interactions between particles occupies a central role, and is often the most expensive part of the computation. Discrete sums may also be encountered in the evaluation of integral equations obtained in the solution of boundary value problems. In any event, numerical models require the application of dense matrices which, done directly, requires an amount of work proportional to N^2 for an N particle system (or N point discretization). To overcome this computational hurdle there is a need for fast $O(N)$ algorithms.

In this thesis, we introduce a class of fast algorithms for the computation of discrete sums using projections on a multiresolution analysis. (A brief survey of existing algorithms is presented below in Section 1.3.) The algorithms are designed to compute

$$g(x_m) = \sum_{\substack{n=1 \\ y_n \neq x_m}}^{N} K(x_m - y_n) f(y_n) , \qquad 1 \leq m \leq M \tag{1.1}$$

for real or complex numbers f and g, in a number of operations proportional to $(M + N)$. In the context of particle simulations the numbers $f(y_n)$ and $g(x_m)$, respectively, may be interpreted as the charge carried by an individual particle positioned at y_n, and the value of the potential field generated by the entire ensemble of N particles, sampled at the point x_m. In case x_m is also the location of a particle we must exclude the self-interaction from the sum (which would in general be infinite), hence the requirement $x_m \neq y_n$.

As a special case of (1.1) we consider the sums

$$g(x_m) = \sum_{\substack{n=1 \\ n \neq m}}^{N} K(x_m - x_n) f(x_n)\,, \qquad 1 \leq m \leq N \tag{1.2}$$

where the N sampling locations coincide with the positions of the N particles. Indeed, there is no loss of generality in taking (1.2) as our model problem, and this we do from now on. The particle locations $\{x_n\}_{n=1}^N$ are points in $\mathbf{R}^d$, where $d = 1, 2, 3$. Numerical results for test problems in one and two dimensions are presented below.

We develop algorithms for kernels that might be described loosely as being of potential-type. We require that interaction between particles depends only on the distance separating them, which implies that the kernel must be convolutional, i.e. $K(x, y) = K(x - y)$. In addition, we allow the kernel to be singular on the diagonal $x = y$, but require it to be non-oscillatory and smooth away from the diagonal. Simple examples are $K(x - y) = 1/(x - y)$ and $K(x - y) = 1/|x - y|$.

We should point out that although the FFT diagonalizes a convolution, this method cannot be used directly for evaluation of sums when the kernel is singular. For such kernels, a prohibitively large number of points is needed for discretization in order to evaluate (1.2) on an equispaced grid, especially for non-uniform particle distributions.

In [4] algorithms have been developed for evaluating

$$g(x_m) = \sum_{n=1}^{N} K(x_m, x_n) f(x_n)\,, \qquad 1 \leq m \leq N \tag{1.3}$$

when this sum may be viewed as a discretization of an integral operator

$$g(x) = \int_{-\infty}^{\infty} K(x, y) f(y)\, dy\,. \tag{1.4}$$

Our task here is to start with a sum in (1.2) that may not involve a summable kernel, and such that interpretation as a discretization of a convergent integral is not valid. By transferring the problem to a multiresolution analysis we effectively regularize a singular kernel. The corresponding operator becomes nonsingular when restricted to the subspaces of the multiresolution analysis.

Our approach offers some advantages over existing algorithms. For example, we offer greater flexibility, since changing the algorithm to accommodate a new kernel requires only trivial modifications, and the ability to handle a wider class of kernels. In contrast to previous algorithms, our approach transfers the computation into a basis. For this reason, it is more easily combined with pseudo-spectral or wavelet-type PDE solvers.

In our approach to discrete summation problems, we construct explicit representations of singular operators on subspaces of the multiresolution analysis. These representations provide a definition for the regularization of such operators, as well as a practical algorithm for their computation. We present a new multiresolution approach to the regularization of singular operators, and show that our method coincides with the classical method, where the classical method is applicable.

1.2 Heuristic Considerations

By transferring the problem to a multiresolution analysis we are projecting onto a basis consisting of translations and dilations of a single function $\phi(x)$, called a "scaling function." This basis provides high order approximation for smooth kernels. Of considerable importance as well is the fact that the scaling function may be chosen to have compact support. This means that approximation errors due to singularities in the kernel are localized, and their affect is not felt by basis functions whose supports lie outside a small neighborhood of the singularity. This allows us to split the kernel into the sum of a low frequency, or smooth part, and a high frequency, or singular part. We note that all existing fast summation algorithms make use of this splitting, and differ only in the manner in which it is done.

We therefore express the original kernel K in the form

$$K = K_{LF} + K_{HF}\,. \tag{1.5}$$

The low frequency part K_{LF} accounts for long range, or "far-field" interactions. In our approach, this part of the kernel is obtained by projecting the kernel onto an appropriate subspace of the multiresolution analysis. Thus, we have

$$K_{LF} = P_j K P_j = T_j\,, \tag{1.6}$$

where P_j denotes the projector onto a subspace of the multiresolution analysis. The integer j is a scale parameter. The projection is a smoothed version of the original kernel, which can be applied to a vector efficiently. Current implementations of our algorithm use the FFT to accomplish this, but other methods are available.

The high fequency part K_{HF} is defined by the difference

$$K_{HF} = K - T_j\,. \tag{1.7}$$

It will be shown below that for any positive ϵ we have an error estimate of the form

$$|K(x-y) - T_j(x,y)| < \epsilon\,, \quad \text{if } |x-y| > 2^j\delta \tag{1.8}$$

for some positive δ. It follows from (1.7) that $|K_{HF}(x,y)|$ decays rapidly as $|x-y|$ increases. Therefore, this singular part of the kernel influences only the short range, or local interactions, and is represented by a banded matrix which can be applied to a vector in $O(N)$ operations.

1.3 A Brief Survey of Existing Algorithms

All fast algorithms for evaluation of (1.2) make use of the splitting (1.5) in one form or another. For example, in [21] we find the following statements, "Ewald summation separates the Green function for a cube into a high frequency localized part and a rapidly converging Fourier series", and "Our methods use Ewald summation to split the potential into a high frequency localized part and a low frequency part with separated variables." (Ewald summation [12] refers to the use of two different series expansions for the Green function of the Laplacian.) With the obvious modification, these statements apply equally well to the current work. We should point out, however, that Ewald summation is specific to the Laplacian operator, although a similar ansatz could conceivably be applied in other cases. Once the splitting has been achieved, the author employs a method appropriate to each part for evaluation. A fast Gauss transform is used to evaluate the high frequency part, while an early version of V. Rohklin's non-equidistant fast Fourier transform [10] is used to evaluate the low frequency part. It is interesting to note that both of these fast transforms are based on the Fast Multipole Method [16].

The most competitive algorithm in terms of CPU requirements that has yet appeared for fast summation problems is the Fast Multipole Method, or FMM. In this approach, a heirarchy of boxes is first constructed that refines the computational domain into smaller and smaller regions. The construction is adaptive in that at each level, only those boxes that contain a number of particles greater than a fixed parameter are subdivided. The potential due to the particles contained within each box is then represented by a single multipole expansion about the center of the box. The expansion for a box on a coarse scale is obtained in an efficient manner by merging the expansions for fine scale boxes contained within it (its "children"). Far field (low frequency) interaction between pairs of well-separated boxes on the same scale is computed as follows: the multipole expansion for a given box is shifted to the center of a second box, where it is added to the existing multipole expansion for that box. The contribution from the first box can then be distributed to each "child" of the second box, then to each of its children, and so on down the tree to the boxes on the finest scale. Local interactions (high frequency) are simply computed directly.

Implicit in the FMM is the simple splitting $K = K_{LF} + K_{HF}$, where

$$K_{LF} = \begin{cases} 0\,, & |x-y| \le \delta \\ K(x-y)\,, & |x-y| > \delta \end{cases} \quad \text{and} \quad K_{HF} = \begin{cases} K(x-y)\,, & |x-y| \le \delta \\ 0\,, & |x-y| > \delta \end{cases}$$

The low frequency part is applied by a clever use of multipole expansions in a divide and conquer strategy, taking advantage of the smoothness of the kernel on regions removed from the origin to truncate the expansions after a few terms. The high frequency part is applied directly. Despite its apparent simplicity, the algorithm is not so straightforward to implement. The explicit form of the multipole expansions and translation operators, which are responsible for shifting the expansions from one box center to another, must be worked out anew for each new kernel. However, when properly implemented, the FMM provides a very efficient $O(N)$ algorithm.

The Method of Local Corrections was introduced in [1] as a vortex method for solving problems in fluid mechanics, though the main ideas are certainly relevant in a more general context. This method is designed to approximate the velocity field due to a distribution of "vortex blobs" in a fluid, and to evaluate this field at the center of each blob. A vortex blob is a radially symmetric function, usually with compact support, that approximates a point vortex. In [1] it is observed that "...the difference between the velocity field due to a point vortex and a vortex blob located at the same point in space becomes very small as one moves away from the center of the vortices." (This statement should be compared to the error estimate (1.8).) The approximate velocity is first obtained on an equispaced grid via a fast Poisson solver (FFT), and then interpolated to the centers of the vortices. The approximation is then corrected locally, for all vortices that lie in close proximity to other vortices. The interpolation is accomplished by the use of a complex-valued interpolating polynomial, made possible by the fact that the $x-$ and $y-$ components of velocity are the real and imaginary parts of a harmonic function. This is a special feature of vortex methods, and is not likely to generalize to other applications. The correction step in this method is essentially equivalent to (1.7), and the splitting of the kernel is achieved by means analogous to that described in Section 1.2. However, in approximating the velocity field, a smoothed version of the original operator is not constructed, and a finite difference method is employed instead.

In [11], a class of algorithms for particle simulations involving Poisson's equation is developed. These algorithms are named PPPM by the authors, which stands for "particle-particle, particle-mesh." The name refers to the by now familiar splitting of the summation into low and high frequency contributions. The low frequency part is evaluated on an equispaced grid, or mesh. The values assigned to the mesh points are obtained from the charges on the particles by use of a "charge-assignment function." This is a piecewise constant, linear, or quadratic B-spline, centered at the mesh point. The value assigned to the mesh point is equal to the sum of the values of the assignment function at each particle location, for all particle locations that lie within the support of the assignment function. Once the mesh values have been obtained, the authors solve a finite difference approximation to the Poisson equation on the mesh using a fast Poisson solver, and then interpolate the result back to the particle locations using the charge assignment function. This phase of the algorithm is called "particle-mesh."

The result of the PM step must be corrected to account for the high frequency part of the kernel, and this is done in a manner similar to the Method of Local Corrections. It appears that the principle difference between algorithms described in [1] and those in [11] is that in the former, extensive use is made of the fact that the velocity field induced by a point vortex is harmonic away from the source, whereas this feature does not appear in the latter, where applications to vortex methods were not considered.

1.4 Organization of the Thesis

This thesis is organized as follows. In Chapter(2) we describe the mathematical structure of a multiresolution analysis. In an attempt to present our thesis in a reasonably self-contained format, we have included many well known results. However, we also present material that does not seem to have appeared previously.

Once the foundation has been established, we proceed in Chapter(3) to describe the one-dimensional version of our algorithm. The generalization to higher dimensions is fairly straightforward, the only additional tool utilized being singular value decomposition, and this is described in Chapter(4). In addition, numerical examples in one and two dimensions are presented in Chapters (3) and (4), respectively.

In Chapter(5) we describe a procedure for constructing representations of singular kernels. This representation provides a definition for the regularization of such kernels, as well as a practical algorithm for computation. It is a unique feature of our approach to discrete summation problems that we construct smooth and explicit representations for a wide class of singular kernels, and we describe this process in some detail.

Chapter 2

Multiresolution Analysis

2.1 Definition and Basic Properties

We first make some preliminary comments. Throughout this thesis, we use the notation $\hat{f}$ to refer to the Fourier transform of a function f,

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{ix\cdot\xi} f(x)\, dx\,, \qquad f(x) = \frac{1}{(2\pi)^d} \int_{-\infty}^{\infty} e^{-ix\cdot\xi} \hat{f}(\xi)\, d\xi\,.$$

for a function $f \in L^2(\mathbf{R}^d)$. We use $(\cdot\,,\cdot)$ to refer to the usual inner product on $L^2(\mathbf{R}^d)$,

$$(f, g) = \int_{-\infty}^{\infty} f(x)\overline{g}(x)\, dx\,.$$

A superscript and a subscript on a function will denote, respectively, a dilation and a translation,

$$f_k^j(x) = 2^{-dj/2} f(2^{-j}x - k)\,,$$

where $j \in \mathbf{Z}$ and $k \in \mathbf{Z}^d$. This will sometimes also be written as $f_{j,k}(x)$. The normalization factor $2^{-dj/2}$ insures that $\|f_k^j\| = \|f\|$ in the usual $L^2(\mathbf{R}^d)$ norm, $\|f\|^2 = (f, f)$.

We use the shorthand MRA to refer to a multiresolution analysis, or multiresolution approximation, of $L^2(\mathbf{R}^d)$. This consists of a nested sequence of subspaces (see Definition(2.1.1) below), ordered as follows,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots\,.$$

With this choice of notation, the sequence *increases* as the subscript j *decreases*. To further emphasize this point, we note that

$$\bigcap_{-\infty}^{\infty} V_j = \lim_{j\uparrow+\infty} V_j\,, \qquad \bigcup_{-\infty}^{\infty} V_j = \lim_{j\downarrow-\infty} V_j\,.$$

In what follows, we make repeated use of Poisson's summation formula,

$$\sum_{k=-\infty}^{\infty} f(k)e^{ik\xi} = \sum_{l=-\infty}^{\infty} \hat{f}(\xi + 2l\pi)\,. \tag{2.1}$$

There exist many proofs of this well-known result (see e.g. [18] or [22]).

The following definition is by now standard, and is borrowed from [20, p.21].

Definition 2.1.1 *A multiresolution approximation of $L^2(\mathbf{R}^d)$ is an increasing sequence $V_j\,, j \in \mathbb{Z}$ of closed linear subspaces of $L^2(\mathbf{R}^d)$ with the following properties:*

$$\bigcap_{-\infty}^{\infty} V_j = \{0\}\,, \qquad \bigcup_{-\infty}^{\infty} V_j \;\; \textit{is dense in } L^2(\mathbf{R}^d)\,; \tag{2.2}$$

for all $f \in L^2(\mathbf{R}^d)$ and all $j \in \mathbb{Z}$,

$$f(x) \in V_j \Longleftrightarrow f(2x) \in V_{j-1}\,; \tag{2.3}$$

for all $f \in L^2(\mathbf{R}^d)$ and all $k \in \mathbb{Z}^d$,

$$f(x) \in V_0 \Longleftrightarrow f(x-k) \in V_0\,; \tag{2.4}$$

there exists a function, $g(x) \in V_0$, such that the sequence

$$g(x-k)\,, \quad k \in \mathbb{Z}^d \tag{2.5}$$

is a Riesz basis for the space V_0.

The sequence (2.5) is a Riesz basis if it spans V_0, and there exist two constants, $c_2 \geq c_1 > 0$, such that for all sequences of scalars $\{\alpha_k\}$ we have

$$c_1 \sum_k |\alpha_k|^2 \leq \left\| \sum_k \alpha_k\, g(x-k) \right\|^2 \leq c_2 \sum_k |\alpha_k|^2\,.$$

The function $g(x)$ is normalized so that $\int g(x)\,dx = 1$. It follows that the sequence $g_k^j(x), k \in \mathbb{Z}^d$ is a Riesz basis for the subspace V_j. The function g is called a "scaling function". In this thesis, we consider only real-valued scaling functions. A given MRA may have several scaling functions.

The subspace V_j may be viewed as corresponding to an equispaced grid, after the manner of spaces of piecewise polynomial splines, where the stepsize is 2^j. The value assigned to the node $2^j k, k \in \mathbb{Z}^d$, is the coefficient corresponding to the basis function g_k^j. Note that the grid spacing becomes finer as the scale parameter j becomes more negative.

The following theorem shows that every MRA possesses an orthonormal basis which, furthermore, has the same structure as the Riesz basis (2.5). Equation (2.7) is a recipe for constructing the "canonical" orthonormal basis from a given Riesz basis. This construction will be used below (see Section 2.4.1) in connection with Riesz bases of B-splines. The theorem and its proof may be found in [20, pp.26-7].

Theorem 2.1.1 *Let $V_j\,,j \in \mathbb{Z}$ be a multiresolution approximation of $L^2(\mathbb{R}^d)$. Then there exist two constants, $c_2 \geq c_1 > 0$, such that for almost all $\xi \in \mathbb{R}^d$ we have*

$$c_1 \leq \left(\sum_{k \in \mathbb{Z}^d} |\hat{g}(\xi + 2k\pi|^2 \right)^{1/2} \leq c_2\,. \tag{2.6}$$

Further, if $\phi \in L^2(\mathbb{R}^d)$ is defined by

$$\hat{\phi}(\xi) = \hat{g}(\xi) \left(\sum_{k \in \mathbb{Z}^d} |\hat{g}(\xi + 2k\pi|^2 \right)^{-1/2}\,, \tag{2.7}$$

then $\phi(x-k)\,, k \in \mathbb{Z}^d$ is an orthonormal basis for V_0. Finally, let $f \in V_0$ be a function such that the sequence $f(x-k)\,, k \in \mathbb{Z}^d$ is orthonormal. Then the sequence is an orthonormal basis for V_0 and we have $\hat{f}(\xi) = \theta(\xi)\hat{\phi}(\xi)$, where $\theta(\xi) \in C^\infty(\mathbb{R}^d)$, $|\theta(\xi)| = 1$ almost everywhere, and $\theta(\xi + 2k\pi) = \theta(\xi)$, for each $k \in \mathbb{Z}^d$.

It follows that the sequence $\phi^j_k(x), k \in \mathbb{Z}^d$ is an orthonormal basis for the subspace V_j. We generally use the lower-case ϕ to refer to the orthonormal scaling function for the multiresolution analysis under discussion.

2.1.1 Tensor Products of One-Dimensional MRAs

The simplest method for constructing an MRA for $L^2(\mathbb{R}^d)$ is to form the tensor product of an MRA for $L^2(\mathbb{R})$. For example, if the sequence $\phi(x-k), k \in \mathbb{Z}$ is a basis for the subspace V_0 of an MRA for $L^2(\mathbb{R})$, then we may take as a basis for the subspace V_0 of an MRA for $L^2(\mathbb{R}^d)$ the sequence

$$\phi(x-k) = \phi(x_1 - k_1) \cdots \phi(x_d - k_d)\,, \qquad k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$$

where $x = (x_1, \ldots, x_d)$. As this construction is employed here, we restrict our attention to development of the one-dimensional theory for the remainder of this chapter.

However, the tensor product construction, though simple, does not address the issue of the speed of the resulting algorithms. It is possible that, in the future, other approaches will be found that lead to more efficient algorithms in higher dimensions.

At the present time, in order to achieve true separation of variables in two dimensions, we use a singular value decomposition of the coefficient matrix, but discussion of these details must be delayed until Chapter(4).

2.1.2 The Two-Scale Difference Equation

It is a consequence of the nesting of the subspaces of the multiresolution analysis, in particular $V_0 \subset V_{-1}$, that the scaling function $\phi(x)$ may be written as a linear combination of the basis functions in the next finer subspace. Thus, we have

$$\phi(x) = \sqrt{2} \sum_k h_k \, \phi(2x - k) \, . \tag{2.8}$$

Equation (2.8) is known as the "two-scale difference equation". In this thesis, we assume that the coefficients $\{h_k\}$ are real.

Taking the Fourier transform on both sides of (2.8), we obtain

$$\hat{\phi}(\xi) = m_0(\xi/2)\hat{\phi}(\xi/2) \, , \tag{2.9}$$

where

$$m_0(\xi) = \frac{1}{\sqrt{2}} \sum_k h_k e^{ik\xi} \, . \tag{2.10}$$

Equation (2.9) is the form taken by the two-scale difference equation in the frequency domain.

Orthonormality of the sequence $\phi(x-k), k \in \mathbb{Z}$ is equivalent to the following condition

$$|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2 = 1 \tag{2.11}$$

on the function $m_0(\xi)$ [7].

Since $m_0(0) = 1$ (as is implied by (2.9) since $\hat{\phi}(0) \neq 0$), it follows from (2.11) that $m_0(\pi) = 0$. Since m_0 has a zero at $\xi = \pi$, it is possible to write

$$m_0(\xi) = \left(\frac{1 + e^{i\xi}}{2} \right) F(\xi)$$

where F is a 2π-periodic function. We will see below that the number of zeroes that m_0 has at π is closely related to the order of approximation obtainable in a given MRA (see Section 2.2). In view of this fact it should come as no surprise that m_0 ordinarily has a zero of multiplicity greater than one at π and so, without loss of generality, we assume that m_0 has the following form,

$$m_0(\xi) = \left(\frac{1 + e^{i\xi}}{2} \right)^M F(\xi) \tag{2.12}$$

where F is a 2π-periodic function, $F \in L^2([0, 2\pi]$, and $F(\pi) \neq 0$.

2.1.3 The Autocorrelation of the Scaling Function

The autocorrelation of the orthonormal scaling function is an important function in its own right, particularly in regard to the approximation of kernels, a topic that we take up in Section 2.3. We denote this function by the upper-case Φ, where

$$\Phi(x) = \int_{-\infty}^{\infty} \phi(x+y)\phi(y)\, dy \,. \tag{2.13}$$

We first note that Φ is interpolating in the following sense,

$$\Phi(n) = \begin{cases} 1 & , n = 0 \\ 0 & , n \neq 0 \end{cases} \tag{2.14}$$

for $n \in \mathbb{Z}$.

The autocorrelation also satisfies a two-scale difference equation. This is most easily seen by examining the Fourier transform. Since

$$\hat{\Phi}(\xi) = |\hat{\phi}(\xi)|^2 \,, \tag{2.15}$$

it follows from (2.9) that

$$\hat{\Phi}(\xi) = M_0(\xi/2)\hat{\Phi}(\xi/2) \,, \tag{2.16}$$

where

$$M_0(\xi) = |m_0(\xi)|^2 \,. \tag{2.17}$$

It follows immediately from (2.11) that

$$M_0(\xi) + M_0(\xi+\pi) = 1 \,. \tag{2.18}$$

The following proposition (see e.g. [2]) establishes an important property of the function $\Phi(x)$, namely that it has a high number of vanishing moments. This property is useful in that it leads to a one-point quadrature formula for evaluating expressions of the form $\int K(x)\Phi(x-k)\, dx$, for functions K that are sufficiently smooth on the support of $\Phi(x-k)$ (cf. Proposition(5.3.2) below.)

Proposition 2.1.1 *In an orthonormal system, the autocorrelation of the scaling function has vanishing moments,*

$$\int_{-\infty}^{\infty} x^m \Phi(x)\, dx = \begin{cases} 1 & , m = 0 \\ 0 & , m > 0 \end{cases} \tag{2.19}$$

for $0 \leq m \leq 2M-1$, *where* M *is the multiplicity of the zero of* m_0 *at* π.

Proof: The moments of Φ are given in terms of the Fourier transform by the formula

$$\int_{-\infty}^{\infty} x^m \Phi(x)\, dx = (-i)^m \hat{\Phi}^{(m)}(0)$$

where $\hat{\Phi}^{(m)}$ denotes the mth derivative of $\hat{\Phi}$. From (2.15) it follows that $\hat{\Phi}(0) = 1$ since $\hat{\phi}(0) = 1$. By differentiating both sides of (2.16) we obtain

$$2^m \hat{\Phi}^{(m)}(2\xi) = \sum_{n=0}^{m} \binom{m}{n} \hat{\Phi}^{(m-n)}(\xi) M_0^{(n)}(\xi)\,.$$

When $\xi = 0$, we have

$$\hat{\Phi}^{(m)}(0) = \frac{1}{2^m - 1} \sum_{n=1}^{m} \binom{m}{n} \hat{\Phi}^{(m-n)}(0) M_0^{(n)}(0)\,. \tag{2.20}$$

Differentiating (2.18), we obtain

$$M_0^{(n)}(0) = -M_0^{(n)}(\pi)\,, \qquad n > 0\,.$$

The function M_0 has a zero of multiplicity $2M$ at $\xi = \pi$, due to the explicit form given by (2.17) and (2.12). Thus we have $M_0^{(n)}(0) = 0$ for $1 \le n \le 2M-1$, and using this fact in (2.20), we obtain (2.19). □

A well-known formula that relates the moments of the autocorrelation to the moments of the scaling function is

Lemma 2.1.1 *Let μ_m and $\mathcal{M}_m$ denote the mth moment of the scaling function and its autocorrelation, respectively. Then we have the formula*

$$\mathcal{M}_m = \sum_{n=0}^{m} \binom{m}{n} (-1)^n \mu_{m-n} \mu_n\,. \tag{2.21}$$

Proof:

$$\begin{aligned}
\mathcal{M}_m &= \int x^m \Phi(x)\, dx \\
&= \int x^m \int \phi(x+y)\phi(y)\, dy\, dx \\
&= \iint (x-y)^m \phi(x)\phi(y)\, dy\, dx \\
&= \sum_{n=0}^{m} \binom{m}{n} (-1)^n \int x^{m-n}\phi(x)\, dx \int y^n \phi(y)\, dy\,,
\end{aligned}$$

which proves the lemma. □

In order to compute the moments of the scaling function, it is not necessary to evaluate the defining integral,

$$\mu_m = \int_{-\infty}^{\infty} x^m \phi(x)\, dx\,.$$

Instead, we have the following recursive formula

$$\begin{aligned} \mu_0 &= 1\,, \\ \mu_m &= \frac{1}{2^m - 1} \sum_{n=1}^{m} \binom{m}{n} \nu_n \mu_{m-n}\,, \quad m \geq 1 \end{aligned} \tag{2.22}$$

where

$$\nu_n = \frac{1}{\sqrt{2}} \sum_l l^n h_l\,.$$

The numbers $\{\nu_n\}$ are the (normalized) moments of the sequence $\{h_l\}$, and are easily computed when $\{h_l\}$ is of finite length.

The formula (2.22) is well-known and is derived as follows. Using the two-scale difference equation (2.8), we have

$$\begin{aligned} \mu_m &= \int x^m \phi(x)\, dx \\ &= \sqrt{2} \sum_l h_l \int x^m \phi(2x - l)\, dx \\ &= \frac{1}{\sqrt{2}} \sum_l h_l \int \left(\frac{x + l}{2}\right)^m \phi(x)\, dx \\ &= \frac{1}{2^m} \sum_{n=0}^{m} \binom{m}{n} \frac{1}{\sqrt{2}} \sum_l l^n h_l \int x^{m-n} \phi(x)\, dx \\ &= \frac{1}{2^m} \sum_{n=0}^{m} \binom{m}{n} \nu_n \mu_{m-n}\,, \end{aligned}$$

from which (2.22) easily follows.

We now consider the coefficients of the two-scale difference equation for Φ. Using (2.10) we have

$$\begin{aligned} M_0(\xi) &= m_0(\xi)\, \overline{m_0(\xi)} \\ &= \frac{1}{\sqrt{2}} \sum_k h_k e^{ik\xi} \frac{1}{\sqrt{2}} \sum_l h_l e^{-il\xi} \\ &= \frac{1}{4} \sum_m a_m e^{im\xi}\,, \end{aligned} \tag{2.23}$$

where

$$a_m = 2 \sum_l h_l h_{l+m}\,. \tag{2.24}$$

Thus, the two-scale difference equation for Φ may be expressed in terms of the variable x as

$$\Phi(x) = \frac{1}{2}\sum_m a_m \Phi(2x - m)\,. \tag{2.25}$$

Setting $x = n$, $n \in \mathbb{Z}$ we have

$$\Phi(n) = \frac{1}{2}\sum_m a_m \Phi(2n - m)\,.$$

The interpolating property (2.14) implies that $a_{2n} = 2\delta_{n,0}$, where δ denotes the Kronecker delta. Using (2.24) it is easy to show that $a_{-m} = a_m$, and we use these observations to write

$$\Phi(x) = \Phi(2x) + \frac{1}{2}\sum_{m\geq 1} a_{2m-1}\left[\Phi(2x - 2m + 1) + \Phi(2x - 1 + 2m)\right]\,. \tag{2.26}$$

If Φ is compactly supported, as it must be if ϕ has compact support, then we understand that only finitely many of the coefficients in (2.26) are non-zero.

2.1.4 Examples

Perhaps the simplest example of an MRA is one whose elements are piecewise constant on dyadic intervals. The scaling function for this MRA is the characteristic function of the interval $[0, 1)$,

$$\phi(x) = \begin{cases} 1\,, & 0 \leq x < 1 \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{2.27}$$

For historical reasons this is known as the Haar system. The Fourier transform is

$$\begin{aligned} \hat{\phi}(\xi) &= \int_0^1 e^{ix\xi}\,dx = \frac{e^{i\xi} - 1}{i\xi} \\ &= \left(\frac{e^{i\xi/2} - 1}{i\xi/2}\right)\left(\frac{e^{i\xi/2} + 1}{2}\right) \\ &= \hat{\phi}(\xi/2) m_0(\xi/2)\,. \end{aligned}$$

From this expression, we can read off the trigonometric polynomial m_0, i.e.

$$m_0(\xi) = \left(\frac{1 + e^{i\xi}}{2}\right) = \frac{1}{\sqrt{2}}\sum_l h_l e^{il\xi}\,,$$

which implies that $h_0 = h_1 = 1/\sqrt{2}$, and $h_l = 0$ otherwise. Furthermore, from the explicit form of $m_0(\xi)$, we see that this function has a single zero at $\xi = \pi$.

The two-scale difference equation satisfied by (2.27) is

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

The autocorrelation of (2.27) is

$$\Phi(x) = \begin{cases} 1+x\,, & -1 \le x < 0 \\ 1-x\,, & 0 \le x < 1 \\ 0\,, & \text{otherwise.} \end{cases}$$

It is easily verified that

$$\int_{-\infty}^{\infty} \Phi(x)\,dx = 1\,, \qquad \int_{-\infty}^{\infty} x\Phi(x)\,dx = 0\,.$$

Example(1) is the lowest order member of both families of scaling functions mentioned in this thesis, namely the central B-splines and the orthonormal scaling functions with compact support constructed by Daubechies (see [7] or [8]). Spline spaces will be described in more detail in Section 2.4. Daubechies scaling functions satisfy the two-scale difference equation

$$\phi(x) = \sqrt{2}\sum_{l=0}^{2M-1} h_l \phi(2x-l)\,, \qquad \text{where} \quad \sum_{l=0}^{2M-1} h_l^2 = 1\,.$$

This scaling function provides an orthonormal basis for the subspaces of an MRA with M vanishing moments, for $M = 1, 2, \dots$, where M is the multiplicity of the zero at $\xi = \pi$ in (2.12). The support of ϕ is the interval $[0, 2M-1]$. The Haar function described above corresponds to $M = 1$. The autocorrelation Φ satisfies the two-scale difference equation

$$\Phi(x) = \Phi(2x) + \sum_{m=1}^{M} a_{2m-1}\left[\Phi(2x-2m+1) + \Phi(2x-1+2m)\right]\,, \tag{2.28}$$

and is supported on the interval $[1-2M, 2M-1]$. The coefficients $a_{2m-1}, 1 \le m \le M$ that appear in (2.28) are rational, and we have the following formula from [2],

$$(2m-1)a_{2m-1} = \frac{(-1)^{m-1}}{(M-m)!(M+m-1)!}\left[\frac{(2M-1)!}{(M-1)!4^{M-1}}\right]^2\,. \tag{2.29}$$

For example, when $M = 2$, we have

$$\begin{aligned}\Phi(x) = \Phi(2x) \quad &+ \quad \frac{9}{16}\left[\Phi(2x-1) + \Phi(2x+1)\right] \\ &- \quad \frac{1}{16}\left[\Phi(2x-3) + \Phi(2x+3)\right]\,.\end{aligned} \tag{2.30}$$

These scaling functions possess good approximation properties, but are difficult to evaluate pointwise. One method that is available for this purpose is a recursion based on the two-scale difference equation. Since $\phi(2^j k) = \sqrt{2}\sum_{l=0}^{2M-1} h_l \phi(2^{j+1}k - l)$, it follows that the values at dyadic rationals $2^j k, j < 0$, may be calculated if the values at all rationals of the form $2^{j+1}k$ are known. For $j = 0$, the values $\phi(1), \ldots, \phi(2M-2)$ may be obtained by solving a small eigenvalue problem. Since ϕ is continuous if $M > 1$, we have $\phi(0) = \phi(2M-1) = 0$.

After tabulating $\phi(x)$ on a sufficiently fine grid, values at intermediate points may be approximated by linear interpolation. The degree of differentiability of these scaling functions increases linearly, and slowly, with M. For values of M less than about ten or twelve, $\phi(x)$ is no more than twice differentiable. For this reason, it is not advisable to use a higher order interpolating polynomial.

2.2 Multiresolution Approximation of Functions

In this section we consider the representation of functions in an MRA. It will be shown that for smooth functions, arbitrarily high order of approximation can be attained with error estimate that depends only on the high order derivatives.

The multiplicity of the zero at $\xi = \pi$ of the function $m_0(\xi)$ (that appears in the two-scale difference equation (2.9)) is the relevant parameter, and determines the order of approximation of a given MRA. It is customary to refer to this parameter as the number of vanishing moments, and this we do from now on. The reason for this is that when the multiplicity of the zero is M, the associated wavelet will have exactly M vanishing moments. Every MRA must have at least one vanishing moment.

We assume without loss of generality that the scaling function ϕ either has compact support, or has exponential decay at infinity. In the second case $\phi(x)$ satisfies an inequality of the form

$$|\phi(x)| \leq Ae^{-\alpha|x|}, \quad x \in \mathbf{R} \tag{2.31}$$

for positive constants A and α.

An important result concerning approximation is contained in the following proposition.

Proposition 2.2.1 *Let ϕ be a scaling function in an MRA with M vanishing moments. Let μ_m denote the mth moment of the scaling function, $\mu_m = \int_{-\infty}^{\infty} x^m \phi(x)\,dx$. Then the following identities are satisfied,*

$$\sum_{-\infty}^{\infty} (x-k)^m \phi(x-k) = \mu_m \tag{2.32}$$

for $0 \leq m \leq M-1$.

Due to our assumption of compact support or exponential decay at infinity, the numbers μ_m are well-defined for every integer $m \geq 0$. Before giving the proof of the proposition, we state and prove a series of lemmas.

Lemma 2.2.1 *If $\phi(x)$ is compactly supported or satisfies (2.31), then for each non-negative integer m there exists a constant C_m such that*

$$\sum_{k=-\infty}^{\infty} |x-k|^m |\phi(x-k)| \leq C_m \tag{2.33}$$

for every real x, and the sum converges uniformly.

Proof: First assume that $\phi(x)$ is compactly supported. Now

$$\sum_{k=-\infty}^{\infty} |x-k|^m |\phi(x-k)| = \lim_{N\to\infty} \sum_{k=-N}^{N} |x-k|^m |\phi(x-k)|\,,$$

provided that this limit exists. If the limit exists, then the limit function must be 1-periodic, so it is sufficient to consider $0 \leq x < 1$. Let $[a,b]$ be the smallest closed interval that contains the support of $\phi(x)$. Then $\phi(x-k) = 0$ if $x-k > b$ or $x-k < a$, so it follows that

$$\lim_{N\to\infty} \sum_{k=-N}^{N} |x-k|^m |\phi(x-k)| = \sum_{k=k_0}^{k_1} |x-k|^m |\phi(x-k)|\,,$$

where $k_0 = -\lfloor b \rfloor$, $k_1 = -\lfloor a \rfloor$, and $\lfloor \cdot \rfloor$ denotes the greatest integer less than or equal to $(\cdot)$. Thus the sequence of partial sums converges for each x, and as the sum involves only a finite number of terms it converges uniformly. For $0 \leq x < 1$, we have $|k| \leq |x-k| \leq |k|+1$, so that

$$\sum_{k=k_0}^{k_1} |x-k|^m |\phi(x-k)| \leq \|\phi\|_\infty \sum_{k=k_0}^{k_1} (1+|k|)^m\,, \tag{2.34}$$

where $\|\phi\|_\infty = \sup|\phi(x)|, a \leq x \leq b$. It follows that the sum is uniformly bounded by the constant on the right-hand side of (2.34).

Next assume that $\phi(x)$ has exponential decay at infinity, i.e. $|\phi(x)| \leq Ae^{-\alpha|x|}$, for some positive constants A and α. Now

$$\sum_{k=-\infty}^{\infty} |x-k|^m |\phi(x-k)| = \lim_{N\to\infty} \sum_{k=-N}^{N} |x-k|^m |\phi(x-k)|\,,$$

provided that this limit exists, and as before it is sufficient to consider $0 \le x < 1$. Then $|k| \le |x-k| \le |k|+1$ implies that

$$|x-k|^m |\phi(x-k)| \le A|x-k|^m e^{-\alpha|x-k|} \le A(1+|k|)^m e^{-\alpha|k|} .$$

Thus,

$$\sum_{k=-N}^{N} |x-k|^m |\phi(x-k)| \le A \sum_{k=-N}^{N} (1+|k|)^m e^{-\alpha|k|} , \tag{2.35}$$

and as $N \to \infty$ the sequence of partial sums on the right converges for any non-negative integer m. It follows that the sequence of partial sums on the left-hand side of (2.35) converges uniformly for each x, and the limit function is uniformly bounded by the constant $A\sum_{-\infty}^{\infty}(1+|k|)^m e^{-\alpha|k|}$. □

Lemma 2.2.2 *If $\phi(x)$ is compactly supported or satisfies (2.31), then $x^m\phi(x)$ is in $L^1(\mathbf{R})$ for each non-negative integer m, and we have the estimate*

$$\int_{-\infty}^{\infty} |x|^m |\phi(x)|\, dx \le C_m , \tag{2.36}$$

where the constants in (2.33) and (2.36) are identical.

Proof: Using (2.33), and the uniform convergence established by Lemma(2.2.1), we have

$$\begin{aligned}
\int_{-\infty}^{\infty} |x|^m |\phi(x)|\, dx &= \sum_{k=-\infty}^{\infty} \int_{k}^{k+1} |x|^m |\phi(x)|\, dx \\
&= \int_{0}^{1} \sum_{k=-\infty}^{\infty} |x-k|^m |\phi(x-k)|\, dx \\
&\le C_m \int_0^1 dx ,
\end{aligned}$$

which proves the lemma. □

Lemma 2.2.3 *If $\phi(x)$ is compactly supported or satisfies (2.31), then for each non-negative integer m we have the identity*

$$\sum_{k=-\infty}^{\infty} (x-k)^m \phi(x-k) = (-i)^m \sum_{l=-\infty}^{\infty} \hat{\phi}^{(m)}(2l\pi) e^{-2\pi i l x} . \tag{2.37}$$

Proof: Let m be a non-negative integer. By Lemma(2.2.2), $x^m\phi(x)$ is in $L^1(\mathbf{R})$, hence we are allowed to differentiate the Fourier transform $\hat{\phi}(\xi)$ m times. Thus we have

$$(-i)^m \hat{\phi}^{(m)}(\xi) = \int_{-\infty}^{\infty} e^{ix\xi} x^m \phi(x)\, dx ,$$

where $\hat{\phi}^{(m)}(\xi) = (d/d\xi)^m \phi(\xi)$. Now

$$\begin{aligned}
(-i)^m \hat{\phi}^{(m)}(2l\pi) &= \sum_{k=-\infty}^{\infty} \int_k^{k+1} e^{2\pi i l x} x^m \phi(x)\, dx \\
&= \sum_{k=-\infty}^{\infty} \int_0^1 e^{2\pi i l x} (x-k)^m \phi(x-k)\, dx\,.
\end{aligned}$$

As the series $\sum_{-\infty}^{\infty} |x-k|^m |\phi(x-k)|$ converges uniformly by Lemma(2.2.1), so therefore does the series $\sum_{-\infty}^{\infty} (x-k)^m \phi(x-k)$. Accordingly, we may reverse the order of integration and summation to obtain

$$(-i)^m \hat{\phi}^{(m)}(2l\pi) = \int_0^1 e^{2\pi i l x} \sum_{k=-\infty}^{\infty} (x-k)^m \phi(x-k)\, dx\,. \tag{2.38}$$

The inequality (2.33) shows that the series $\sum_{-\infty}^{\infty} (x-k)^m \phi(x-k)$ is in $L^1([0,1])$, and the coefficients of its Fourier series are given by (2.38). The identity (2.37) now follows. □

Lemma 2.2.4 *Let ϕ be a scaling function in an MRA with M vanishing moments. Then for $0 \le m \le M-1$ we have*

$$\hat{\phi}^{(m)}(2l\pi) = 0 \tag{2.39}$$

for all non-zero integers l.

Proof: The proof is by induction. First take $m = 0$. Let l be a non-zero integer. Then l can be expressed in the form $l = 2^j k$, where $j \ge 0$ and k is an odd integer. Using the two-scale difference equation (2.9) we have

$$\begin{aligned}
\hat{\phi}(2l\pi) &= \hat{\phi}(l\pi) m_0(l\pi) \\
&= \hat{\phi}(2^j k\pi) m_0(2^j k\pi)\,.
\end{aligned}$$

We recall that m_0 is 2π-periodic, $m_0(0) = 1$, and $m_0(\pi) = 0$. Now if $j > 0$, then since $m_0(2^j k\pi) = 1$, we have

$$\begin{aligned}
\hat{\phi}(2l\pi) &= \hat{\phi}(2^j k\pi) \\
&= \hat{\phi}(2^{j-1} k\pi) m_0(2^{j-1} k\pi)
\end{aligned}$$

and so on. Iterating this equation j times we obtain

$$\hat{\phi}(2l\pi) = \hat{\phi}(k\pi) m_0(k\pi) = 0\,,$$

since k is odd. This verifies (2.33) for $m = 0$.

Now assume that $\hat{\phi}^{(n)}(2l\pi) = 0$ for all non-zero integers l and for $0 \le n \le m-1$. Differentiate (2.9) m times to obtain

$$\hat{\phi}^{(m)}(\xi) = \frac{1}{2^m} \sum_{n=0}^{m} \binom{m}{n} \hat{\phi}^{(m-n)}(\xi/2) m_0^{(n)}(\xi/2) \,.$$

Setting $\xi = 2l\pi$ we have

$$\begin{aligned} \hat{\phi}^{(m)}(2l\pi) &= \frac{1}{2^m} \sum_{n=0}^{m} \binom{m}{n} \hat{\phi}^{(m-n)}(2^j k\pi) m_0^{(n)}(2^j k\pi) \\ &= \frac{1}{2^m} \hat{\phi}^{(m)}(2^j k\pi) \,, \quad \text{etc.} \end{aligned}$$

and upon iterating this equation j times we obtain

$$\hat{\phi}^{(m)}(2l\pi) = \frac{1}{2^{mj}} \sum_{n=0}^{m} \binom{m}{n} \hat{\phi}^{(m-n)}(k\pi) m_0^{(n)}(k\pi) = 0$$

since k is odd and $m_0^{(n)}(\pi) = 0$ for $0 \le n \le m$. As this argument holds for any m up to $(M-1)$ the lemma is proved. $\square$

Proof of Proposition(2.2.1): Since $\phi(x)$ either has compact support, or exponential decay at infinity, the identity (2.37) holds for all integers $m \ge 0$. In particular, we have

$$\sum_{k=-\infty}^{\infty} (x-k)^m \phi(x-k) = (-i)^m \sum_{l=-\infty}^{\infty} \hat{\phi}^{(m)}(2l\pi) e^{-2\pi i l x}$$

for $0 \le m \le M-1$. Since $\phi(x)$ belongs to an MRA with M vanishing moments, it follows by Lemma(2.2.4) that for $0 \le m \le M-1$, we have $\hat{\phi}^{(m)}(2l\pi) = 0$ if $l \ne 0$. Thus,

$$\sum_{k=-\infty}^{\infty} (x-k)^m \phi(x-k) = (-i)^m \hat{\phi}^{(m)}(0)$$

for $0 \le m \le M-1$. To complete the proof, note that $(-i)^m \hat{\phi}^{(m)}(0) = \int_{-\infty}^{\infty} x^m \phi(x)\, dx$. $\square$

Using (2.32), we obtain the following identities,

$$\begin{aligned} \sum \phi(x-k) &= 1 \\ \sum k\phi(x-k) &= x - \mu_1 \,, \quad \text{if } M \ge 2 \\ \sum k^2 \phi(x-k) &= x^2 - 2x\mu_1 + \mu_2 \,, \quad \text{if } M \ge 3 \\ &\vdots \end{aligned}$$

and in general we have

$$\sum_{-\infty}^{\infty} k^m \phi(x-k) = \sum_{n=0}^{m} \binom{m}{n} x^{m-n}(-1)^n \mu_n\,, \qquad 0 \le m \le M-1\,. \tag{2.40}$$

Let $p_m(x)$ denote the polynomial on the right-hand side of (2.40). It is interesting to note that

$$p'_m(x) = m p_{m-1}(x)\,, \quad 1 \le m \le M-1$$

or equivalently

$$p_m(x) = \int_0^x p_{m-1}(t)\,dt + (-1)^m \mu_m\,, \quad m \ge 1\,.$$

Setting $p_0(x) = 1$, we can use this iteration to construct a sequence of polynomials $p_m(x), m \ge 0$ which is, in a sense, associated to the given MRA.

Since $p_m(x)$ has degree m, it follows that the polynomials $p_0(x), \ldots, p_{M-1}(x)$ are linearly independent. Since $p_m(x) = \sum k^m \phi(x-k)$, and the series converges uniformly, it follows that any polynomial of degree less than or equal to $(M-1)$ may be expressed as a combination of the functions $\phi(x-k)$, $k \in \mathbb{Z}$.

2.2.1 The Projection onto V_j

The projection of a function $f \in L^2(\mathbb{R})$ onto the subspace V_j is given by

$$(P_j f)(x) = \sum_{-\infty}^{\infty} (f, \phi_k^j)\, \phi_k^j(x)\,. \tag{2.41}$$

where

$$(f, \phi_k^j) = 2^{-j/2} \int_{-\infty}^{\infty} f(x)\phi(2^{-j}x - k)\,dx\,.$$

The projection operator may be written explicitly as an integral operator,

$$(P_j f)(x) = \int_{-\infty}^{\infty} P_j(x,y) f(y)\,dy\,,$$

where $P_j(x,y) = 2^{-j} P(2^{-j}x, 2^{-j}y)$, and

$$P(x,y) = \sum_{-\infty}^{\infty} \phi(x-k)\phi(y-k)\,. \tag{2.42}$$

We use the notation s_k^j to denote the coefficient of the basis function ϕ_k^j. Thus,

$$s_k^j = (f, \phi_k^j)\,. \tag{2.43}$$

If $f \in L^2(\mathbf{R})$, then $\sum |s^j_k|^2 < \infty$. Since the basis $\phi^j_k, k \in \mathbf{Z}$ is orthonormal, it follows that

$$\|P_j f\|^2 = \sum_{-\infty}^{\infty} |s^j_k|^2 .$$

We may also allow f to be a generalized function [3]. In this context, we take $\phi^j_k, k \in \mathbf{Z}$ to be the test functions, and V_j to be the space of test functions. Then f is a continuous linear functional that assigns a unique real number (f, ϕ^j_k) to each $\phi^j_k \in V_j$. If f is locally summable in every bounded interval on the line, then we have

$$(f, \phi^j_k) = 2^{-j/2} \int_{-\infty}^{\infty} f(x)\phi(2^{-j}x - k)\, dx$$

as before. If $\sum (f, \phi^j_k)^2 < \infty$, then we define the projection of f onto V_j to be

$$(P_j f)(x) = \sum_{-\infty}^{\infty} (f, \phi^j_k)\, \phi^j_k(x) .$$

As an example, consider the δ-function, $\delta_{x_0} = \delta(x - x_0)$. For this generalized function the projection is

$$(P_j \delta_{x_0})(x) = \sum_{-\infty}^{\infty} \phi^j_k(x_0)\phi^j_k(x) .$$

Now let us consider the difference between a smooth function and its projection onto a subspace of the multiresolution analysis.

Proposition 2.2.2 *Let ϕ be a compactly supported scaling function in an MRA with M vanishing moments. Let $P_j f$ denote the projection of a function f onto the subspace V_j, given by*

$$(P_j f)(x) = \sum_{k=-\infty}^{\infty} s^j_k \phi^j_k(x) \tag{2.44}$$

where the coefficients are given by (2.43). For a given point $x \in \mathbf{R}$, let $I_j(x)$ be the interval formed by the union of the supports of all basis functions which are non-zero at x. Thus

$$I_j(x) = \bigcup_{k \in \mathcal{K}} \mathrm{supp}\,(\phi^j_k) , \qquad \mathcal{K} = \{k \in \mathbf{Z} | \phi^j_k(x) \neq 0\} .$$

Suppose that f is at least M times continuously differentiable on $I_j(x)$. Then we have

$$(P_j f)(x) = f(x) + E_j(x) , \tag{2.45}$$

where

$$|E_j(x)| \leq 2^{Mj} C \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!} . \tag{2.46}$$

The constant C depends on ϕ but not on f.

We point out that in the inequality (2.46), in all practical cases the scale parameter j satisfies $j \leq 0$. Before proving the proposition we state and prove a combinatorial lemma.

Lemma 2.2.5 *Let $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ be arbitrary sequences of length M, then we have*

$$\sum_{n=0}^{M-1} a_n \sum_{m=0}^{M-1-n} \binom{n+m}{n} b_m c_{n+m} = \sum_{m=0}^{M-1} c_m \sum_{n=0}^{m} \binom{m}{n} a_n b_{m-n} \,. \tag{2.47}$$

Proof: For convenience denote the left-hand side of (2.47) by l.h.s. Expanding the summation over n we have

$$\text{l.h.s.} = a_0 \sum_{m=0}^{M-1} \binom{0+m}{0} b_m c_{0+m} + \cdots + a_{M-1} \sum_{m=0}^{0} \binom{M-1+m}{M-1} b_m c_{M-1+m} \,.$$

Now regroup these terms, factoring out in turn c_0, c_1, etc. Since c_0 appears only in the first term, c_1 appears only in the first and second terms, etc., we obtain

$$\text{l.h.s.} = c_0 \left\{ \binom{0}{0} a_0 b_0 \right\} + c_1 \left\{ \binom{1}{0} a_0 b_1 + \binom{1}{1} a_1 b_0 \right\}$$

$$+ \cdots + c_{M-1} \left\{ \binom{M-1}{0} a_0 b_{M-1} + \binom{M-1}{1} a_1 b_{M-2} + \cdots + \binom{M-1}{M-1} a_{M-1} b_0 \right\} \,.$$

This can evidently be written in compact notation as

$$\text{l.h.s.} = \sum_{m=0}^{M-1} c_m \sum_{n=0}^{m} \binom{m}{n} a_m b_{m-n} \,,$$

which verifies (2.47). $\square$

Proof of the Proposition: We first obtain an expression for the coefficient s_k^j in terms of the derivatives of f. Expanding f in a Taylor series about $(2^j k)$ we have

$$f(x) = \sum_{m=0}^{M-1} \frac{f^{(m)}(2^j k)}{m!} (x - 2^j k)^m + \frac{f^{(M)}(\xi_k)}{M!} (x - 2^j k)^M$$

where ξ_k lies between x and $(2^j k)$. Using this expression we have

$$\begin{aligned} s_k^j &= 2^{-j/2} \int f(x) \phi(2^{-j} x - k)\, dx \\ &= 2^{-j/2} \sum_{m=0}^{M-1} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} \int (2^{-j} x - k)^m \phi(2^{-j} x - k)\, dx + 2^{j/2} \varepsilon_k^j \\ &= 2^{j/2} \sum_{m=0}^{M-1} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} \mu_m + 2^{j/2} \varepsilon_k^j \end{aligned}$$

where

$$\varepsilon_k^j = 2^{Mj} \int \frac{f^{(M)}(\xi_k)}{M!}(2^{-j}x - k)^M \phi(2^{-j}x - k)\, dx\,. \tag{2.48}$$

Now substitute this into (2.44) to obtain

$$(P_j f)(x) = \sum_{m=0}^{M-1} 2^{mj} \mu_m \sum_k \frac{f^{(m)}(2^j k)}{m!} \phi(2^{-j}x - k) + E_j^1(x) \tag{2.49}$$

where we have set

$$E_j^1(x) = \sum_k \varepsilon_k^j \phi(2^{-j}x - k)\,.$$

Using the Taylor expansion for the mth derivative we can write

$$f^{(m)}(2^j k) = \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x)}{n!}(2^j k - x)^n + \frac{f^{(M)}(\xi_{m,k})}{M!}(2^j k - x)^M\,,$$

where $\xi_{m,k}$ lies between x and $(2^j k)$ for each m and k. Substituting this expression into (2.49) we obtain

$$\begin{aligned}(P_j f)(x) \;&=\; \sum_{m=0}^{M-1} 2^{mj} \mu_m \sum_k \phi(2^{-j}x - k) \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x)}{m!n!}(2^j k - x)^n \\ &\quad + E_j(x)\end{aligned} \tag{2.50}$$

where we have set $E_j = E_j^1 + E_j^2$, and

$$E_j^2(x) = \sum_{m=0}^{M-1} \frac{\mu_m}{m!} 2^{mj} \sum_k \frac{f^{(M)}(\xi_{m,k})}{M!}(2^j k - x)^M \phi(2^{-j}x - k)\,.$$

We can rearrange (2.50) to obtain

$$(P_j f)(x) - E_j(x) = \sum_{m=0}^{M-1} 2^{mj} \mu_m \sum_{n=0}^{M-1-m} \binom{n+m}{m} \frac{f^{(n+m)}(x)}{(n+m)!}(-1)^n 2^{nj} \mu_n \tag{2.51}$$

having also used Proposition(2.2.1). Now use Lemma(2.2.5) to transform the right-hand side of (2.51), thus obtaining

$$(P_j f)(x) - E_j(x) = \sum_{m=0}^{M-1} \frac{f^{(m)}(x)}{m!} 2^{mj} \sum_{n=0}^{m} \binom{m}{n} (-1)^n \mu_{m-n} \mu_n\,. \tag{2.52}$$

Finally, using Lemma(2.1.1) together with Proposition(2.1.1), we have

$$\sum_{n=0}^{m} \binom{m}{n} (-1)^n \mu_{m-n} \mu_n = \delta_{m,0}\,,$$

so that (2.52) reduces to

$$(P_j f)(x) - E_j(x) = f(x) .$$

This proves (2.45).

Now let us examine the error terms. Clearly

$$|E_j^1(x)| \leq C_0 \sup_k |\varepsilon_k^j| ,$$

where we have used Lemma(2.2.1). Now consider

$$\begin{aligned} |\varepsilon_k^j| &\leq 2^{Mj} \int \frac{|f^{(M)}(\xi_k)|}{M!} |2^{-j}x - k|^M |\phi(2^{-j}x - k)| \, dx \\ &\leq 2^{(M+1)j} \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!} \int |x|^M |\phi(x)| \, dx \\ &\leq 2^{(M+1)j} C_M \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!} \end{aligned}$$

where we have used Lemma(2.2.2). Thus we have

$$|E_j^1(x)| \leq 2^{(M+1)j} C_0 C_M \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!} \tag{2.53}$$

Similarly we have

$$|E_j^2(x)| \leq 2^{Mj} C' C_M \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!}$$

where $C' = \sum_{m=0}^{M-1} (|\mu_m|/m!) 2^{mj}$. This proves (2.46). □

2.3 Multiresolution Approximation of Kernels

2.3.1 Kernels on V_j

In general, a kernel on V_j is an expression of the form

$$T_j(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t_{m,n}^j \phi_m^j(x) \phi_n^j(y) ,$$

where the function $\phi(x) \in V_0$ is the scaling function. In this thesis, we approximate kernels of the form $K(x, y) = K(x - y)$, and in this case the coefficients satisfy

$$t_{m,n}^j = t_{m-n}^j . \tag{2.54}$$

Therefore, for our purposes it is sufficient to view a kernel as an expression of the form

$$T_j(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m-n} \phi^j_m(x) \phi^j_n(y)\,,$$

where the sequence $\{t^j_n\}$ belongs to $l^2(\mathbb{Z})$. Thus, to build an approximation to a given kernel $K(x-y)$ on the subspace V_j, it is necessary only to compute the appropriate coefficients. Conversely, the kernel T_j is completely determined once the coefficients $\{t^j_n\}$ are known.

We note that (2.54) does not imply that $T_j(x,y) = T_j(x-y)$, i.e. T_j is not a convolutional kernel. However, we do have the identity

$$T_j(x + 2^j k, y + 2^j k) = T_j(x,y)\,,$$

which shows that T_j is a "block convolution". To state this another way, T_j is a periodic function of period 2^j along any diagonal line $x - y = constant$.

2.3.2 Projection of a Kernel

Denote by $(Kf)(x)$ the integral operator

$$(Kf)(x) = \int_{-\infty}^{\infty} K(x-y) f(y)\, dy\,.$$

Here we allow K to be a singular operator, and may exist only as a principle value, or only for a certain class of functions f. Assume $\phi^j_k(x), k \in \mathbb{Z}$ is an orthonormal basis for the subspace V_j. The projection of a function $f \in L^2(\mathbf{R})$ onto the subspace V_j (see Section 2.2.1) is given by

$$(P_j f)(x) = \sum_{n=-\infty}^{\infty} (f, \phi^j_n) \phi^j_n(x) = \sum_{n=-\infty}^{\infty} s^j_n \phi^j_n(x)\,.$$

Apply the operator K to this projection to obtain

$$\begin{aligned}
(KP_j f)(x) &= \int_{-\infty}^{\infty} K(x-y)(P_j f)(y)\, dy \\
&= \sum_{n=-\infty}^{\infty} s^j_n \int_{-\infty}^{\infty} K(x-y) \phi^j_n(y)\, dy \\
&= \sum_{n=-\infty}^{\infty} s^j_n (K\phi^j_n)(x)\,.
\end{aligned}$$

Now project this result onto the subspace V_j to obtain

$$\begin{aligned}
(P_j K P_j f)(x) &= \sum_{m=-\infty}^{\infty} (KP_j f, \phi^j_m) \phi^j_m(x) \\
&= \sum_{m=-\infty}^{\infty} \phi^j_m(x) \sum_{n=-\infty}^{\infty} s^j_n (K\phi^j_n, \phi^j_m)\,.
\end{aligned}$$

However, we also have

$$\begin{aligned}(P_jKP_jf)(x) &= \int_{-\infty}^{\infty} T_j(x,y)f(y)\,dy \\ &= \int \sum_m \sum_n t^j_{m-n}\phi^j_m(x)\phi^j_n(y)f(y)\,dy \\ &= \sum_{m=-\infty}^{\infty} \phi^j_m(x) \sum_{n=-\infty}^{\infty} t^j_{m-n} \int \phi^j_n(y)f(y)\,dy \\ &= \sum_{m=-\infty}^{\infty} \phi^j_m(x) \sum_{n=-\infty}^{\infty} s^j_n t^j_{m-n}\,.\end{aligned}$$

Equating these two expressions, we obtain

$$t^j_{m-n} = (K\phi^j_n, \phi^j_m) = \iint K(x-y)\phi^j_m(x)\phi^j_n(y)\,dy\,dx\,. \tag{2.55}$$

Using a change of variables and reversing the order of integration we can rewrite (2.55) as

$$t^j_n = \int_{-\infty}^{\infty} K(x)\Phi(2^{-j}x-n)\,dx\,, \tag{2.56}$$

where Φ is the autocorrelation introduced in Section 2.1.3. Note also the operator identity

$$T_j = P_jKP_j\,.$$

The following proposition establishes a bound on the difference between a kernel K and its multiresolution approximation T_j.

Proposition 2.3.1 *Let ϕ be a compactly supported scaling function in an MRA with M vanishing moments. Let T_j denote the projection of a kernel K onto the subspace V_j, given by*

$$T_j(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} t^j_{k-l}\phi^j_k(x)\phi^j_l(y) \tag{2.57}$$

where the coefficients are given by (2.56). For a given point $(x,y) \in \mathbb{R}^2$, let $R_j(x,y)$ be the rectangle formed by the union of the supports of all basis functions which are non-zero at (x,y). Thus, if

$$I_j(x) = \bigcup_{k\in\mathcal{K}} \mathrm{supp}\,(\phi^j_k)\,, \qquad \mathcal{K} = \{k \in \mathbb{Z} | \phi^j_k(x) \neq 0\}$$

and

$$I_j(y) = \bigcup_{l\in\mathcal{L}} \mathrm{supp}\,(\phi^j_l)\,, \qquad \mathcal{L} = \{l \in \mathbb{Z} | \phi^j_l(y) \neq 0\}$$

then

$$R_j(x,y) = I_j(x) \times I_j(y)\,.$$

Suppose that K is at least M times continuously differentiable on $R_j(x,y)$. Then we have

$$T_j(x,y) = K(x-y) + E_j(x,y)\,, \tag{2.58}$$

where

$$|E_j(x,y)| \le 2^{Mj} C \sup_{(\xi,\eta)\in R_j(x,y)} \frac{|K^{(M)}(\xi-\eta)|}{M!}\,. \tag{2.59}$$

The constant C depends on ϕ but not on K.

Proof: Using Proposition(5.3.2), we can write

$$t_n^j = 2^j K(2^j n) + \varepsilon_n^j\,,$$

and we substitute this into (2.57) to obtain

$$\begin{aligned} T_j(x,y) &= \sum_k \sum_l K\left(2^j(k-l)\right) \phi(2^{-j}x-k)\phi(2^{-j}y-l) \\ &\quad + \sum_k \sum_l \varepsilon_{k-l}^j \phi_k^j(x)\phi_l^j(y)\,. \end{aligned} \tag{2.60}$$

Expanding K in a Taylor series about (x,y), we have

$$\begin{aligned} K\left(2^j(k-l)\right) &= \sum_{m=0}^{M-1} \frac{K^{(m)}(x-y)}{m!} \left[2^j(k-l)-(x-y)\right]^m \\ &+ \frac{K^{(M)}(\xi_k-\eta_l)}{M!} \left[2^j(k-l)-(x-y)\right]^M\,, \end{aligned} \tag{2.61}$$

where ξ_k lies between x and $2^j k$, and η_l lies between y and $2^j l$. Substituting (2.61) into (2.60), we have

$$\begin{aligned} T_j(x,y) &= \sum_{m=0}^{M-1} \frac{K^{(m)}(x-y)}{m!} (-1)^m 2^{mj} \sum_{n=0}^{m} \binom{m}{n} (-1)^n \\ &\quad \times \left(\sum_k (2^{-j}x-k)^{m-n} \phi(2^{-j}x-k) \right) \left(\sum_k (2^{-j}y-l)^n \phi(2^{-j}y-l) \right) \\ &\quad + E_j(x,y)\,, \end{aligned} \tag{2.62}$$

where we have put

$$\begin{aligned} E_j(x,y) &= \sum_k \sum_l \varepsilon_{k-l}^j \phi_k^j(x) \phi_l^j(y) \\ &+ \sum_k \sum_l \frac{K^{(M)}(\xi_k-\eta_l)}{M!} \left[2^j(k-l)-(x-y)\right]^M \phi_k^j(x)\phi_l^j(y)\,. \end{aligned}$$

Now use Proposition(2.2.1) to rewrite (2.62) as

$$\begin{aligned} T_j(x,y) - E_j(x,y) &= \sum_{m=0}^{M-1} \frac{K^{(m)}(x-y)}{m!}(-1)^m 2^{mj} \sum_{n=0}^{m} \binom{m}{n} (-1)^n \mu_{m-n}\mu_n \\ &= K(x-y) \end{aligned}$$

having also used Lemma(2.1.1), and Proposition(2.1.1). This proves (2.58).

Now consider the error terms. Taking $m = M$ in (5.39), we have

$$\sup_{(k,l)\in\mathcal{K}\times\mathcal{L}} |\varepsilon^j_{k-l}| \leq 2^{(M+1)j} C_M \sup_{(\xi,\eta)\in R_j(x,y)} \frac{|K^{(M)}(\xi-\eta)|}{M!}.$$

Using this inequality, we have

$$\begin{aligned} & |E_j(x,y)| \\ \leq \; & \sup_{(k,l)\in\mathcal{K}\times\mathcal{L}} |\varepsilon^j_{k-l}| \left(\sum_k |\phi(2^{-j}x-k)|\right) \left(\sum_l |\phi(2^{-j}y-l)|\right) \\ & + 2^{Mj} \sup_{(\xi,\eta)\in R_j(x,y)} \frac{|K^{(M)}(\xi-\eta)|}{M!} \sum_{n=0}^{M} \binom{M}{n} \\ & \times \left(\sum_k |2^{-j}x-k|^{M-n} |\phi(2^{-j}x-k)|\right) \left(\sum_l |2^{-j}y-l|^n |\phi(2^{-j}y-l)|\right) \\ \leq \; & 2^{Mj} \sup_{(\xi,\eta)\in R_j(x,y)} \frac{|K^{(M)}(\xi-\eta)|}{M!} \left(C_1^2 + \sum_{n=0}^{M} \binom{M}{n} C_{M-n} C_n\right), \end{aligned}$$

which proves (2.59). The constants are those provided by Lemma(2.2.1). □

2.3.3 Trigonometric Expansion of a Kernel

The following theorem provides an efficient method for computing the value of $T_j(x,y)$ at a given point (x,y). As stated above (Section 2.3.1), the kernel $T_j(x,y)$ is not a function of the difference $(x-y)$ alone, a fact which makes it difficult to tabulate. However, it turns out that T_j may be represented by a sum of functions that depend only on $(x-y)$, and being functions of a single variable, are easily tabulated. The proof of this exploits the fact that T_j is periodic on a fixed diagonal $x - y = \textit{constant}$.

The resulting series expansion (2.64) converges rapidly, and in our implementation we retain terms only for $|n| \leq 3$. This high rate of convergence follows from the rapid decay of the Fourier transform of the scaling function.

This result does not require the scaling function ϕ to belong to an orthonormal system.

Theorem 2.3.1 *Let $\phi(x)$ be a scaling function for the subspace V_0, which is continuous and has a piecewise continuous derivative. Let $T_j(x,y)$ be a kernel on V_j, which has the form*

$$T_j(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m-n} \phi^j_m(x) \phi^j_n(y)\,, \tag{2.63}$$

where $\sum |t^j_k|^2 < \infty$. Then we have the identity,

$$2^j T_j(2^j x, 2^j y) = \sum_{n=-\infty}^{\infty} e^{in\pi(x+y)} I_n(x-y)\,, \tag{2.64}$$

where

$$I_n(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iz\xi} \hat{t}^j(\xi - n\pi) \hat{\Phi}_n(\xi)\, d\xi\,, \tag{2.65}$$

$$\hat{t}^j(\xi) = \sum_{-\infty}^{\infty} t^j_k e^{ik\xi}\,, \tag{2.66}$$

and

$$\hat{\Phi}_n(\xi) = \hat{\phi}(\xi - n\pi)\overline{\hat{\phi}}(\xi + n\pi)\,. \tag{2.67}$$

Furthermore, for each $(x,y) \in \mathbf{R}^2$, the right-hand-side of (2.64) converges uniformly to $2^j T_j(2^j x, 2^j y)$.

Example: In order to illustrate the decay of the functions $\hat{\Phi}_n(\xi)$, assume that $\phi(x)$ is the central B-spline of degree $(2m-1)$. Then we have

$$\hat{\Phi}_n(\xi) = \left[\frac{\sin\left(\frac{\xi}{2} + \frac{n\pi}{2}\right)}{\frac{\xi}{2} + \frac{n\pi}{2}}\right]^{2m} \left[\frac{\sin\left(\frac{\xi}{2} - \frac{n\pi}{2}\right)}{\frac{\xi}{2} - \frac{n\pi}{2}}\right]^{2m}$$

which obviously decays rapidly as $|n| \to \infty$.

Proof of the Theorem: We begin by rewriting (2.63) as

$$2^j T_j(2^j x, 2^j y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m-n} \phi(x-m)\phi(y-n)\,. \tag{2.68}$$

Let us restrict the point (x,y) to lie on the line $x - y = z$, the value of z being fixed but arbitrary. On this line, the right-hand-side of (2.68) may be viewed as a function of a single variable, given by

$$\begin{aligned} F(x) &= \sum_m \sum_n t^j_{m-n} \phi(x-m)\phi(x-z-n) \\ &= \sum_m \phi(x-z+m) \sum_k t^j_k \phi(x-k+m)\,. \end{aligned}$$

The function $F(x)$ is 1-periodic, and in addition is continuous and has a piecewise continuous derivative, due to the conditions placed on $\phi(x)$. Hence, by Theorem(2) on p.81 of [23], the Fourier series of $F(x)$ converges uniformly to the value $F(x \bmod 1)$ for each $x \in \mathbf{R}$. Thus, we have

$$F(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x}. \tag{2.69}$$

Now let us compute the coefficients, which must satisfy

$$\begin{aligned} c_n &= \int_0^1 e^{-2\pi i n x} F(x)\, dx \\ &= \int_0^1 e^{-2\pi i n x} \sum_m \phi(x - z + m) \sum_k t_k^j \phi(x - k + m)\, dx. \end{aligned}$$

Making the change of variables $u = x + m$, we have

$$\begin{aligned} c_n &= \sum_m \int_m^{m+1} e^{-2\pi i n(u-m)} \phi(u - z) \sum_k t_k^j \phi(u - k)\, du \\ &= \int_{-\infty}^{\infty} e^{-2\pi i n u} \phi(u - z) \sum_k t_k^j \phi(u - k)\, du. \end{aligned} \tag{2.70}$$

Since, by assumption, $\sum |t_k^j|^2 < \infty$, the series in (2.70) may be expressed in terms of the Fourier transform as

$$\sum_k t_k^j \phi(u - k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iu\xi}\, \hat{t}^j(\xi) \hat{\phi}(\xi)\, d\xi. \tag{2.71}$$

Substituting (2.71) into (2.70) and reversing the order of integration, we obtain

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{t}^j(\xi) \hat{\phi}(\xi) \int_{-\infty}^{\infty} e^{-i(u+z)(\xi + 2n\pi)} \phi(u)\, du\, d\xi \\ &= e^{-2\pi i n z} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iz\xi}\, \hat{t}^j(\xi) \hat{\phi}(\xi) \overline{\hat{\phi}}(\xi + 2n\pi)\, d\xi \\ &= e^{-i\pi n z} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iz\xi}\, \hat{t}^j(\xi - n\pi) \hat{\phi}(\xi - n\pi) \overline{\hat{\phi}}(\xi + n\pi)\, d\xi. \end{aligned}$$

Substituting this result back into (2.69) we have

$$F(x) = \sum_n e^{in\pi(2x - z)} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iz\xi}\, \hat{t}^j(\xi - n\pi) \hat{\Phi}_n(\xi)\, d\xi. \tag{2.72}$$

Using the fact that z is arbitrary, we can equate the right-hand-side of (2.72) with the left-hand-side of (2.68), which yields the result (2.64). $\square$

Alternatively, we can write

$$2^j T_j(2^j x, 2^j y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi(x-y)}\hat{t}^j(\xi)\left|\hat{\phi}(\xi)\right|^2 d\xi$$

$$+\sum_{n=1}^{\infty}\frac{1}{\pi}\int_{-\infty}^{\infty} e^{-i\xi(x-y)}t^j(\xi-n\pi)\mathrm{Re}\left\{e^{in\pi(x+y)}\hat{\Phi}_n(\xi)\right\} d\xi\,. \tag{2.73}$$

When the functions $\hat{\Phi}_n(\xi)$ are real, this reduces to

$$2^j T_j(2^j x, 2^j y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi(x-y)}\hat{t}^j(\xi)\left|\hat{\phi}(\xi)\right|^2 d\xi$$

$$+\sum_{n=1}^{\infty}\cos[n\pi(x+y)]\cdot\frac{1}{\pi}\int_{-\infty}^{\infty} e^{-i\xi(x-y)}t^j(\xi-n\pi)\hat{\Phi}_n(\xi)\, d\xi\,. \tag{2.74}$$

2.4 Spline MRAs

An important class of multiresolution analyses use splines as a starting point (see e.g. [5] or [20]). In this thesis, we restrict our attention to the central B-splines of odd degree, though splines of even degree could also be used. Integer translates of these functions form a Riesz basis.

We denote by $\beta^{(M-1)}(x)$ the spline that is piecewise polynomial of degree $(M-1)$, for $M = 2, 4, \ldots$. By way of example, the lowest order case, corresponding to $M = 2$, is the well-known "hat function",

$$\beta^{(1)}(x) = \begin{cases} 1-|x|, & |x| < 1 \\ 0, & |x| \geq 1\,. \end{cases} \tag{2.75}$$

These functions are compactly supported, and satisfy $\beta^{(M-1)}(x) = 0$ for $|x| \geq M/2$. In addition, they are even, and $\beta^{(M-1)}$ has $(M-2)$ continuous derivatives. Higher order derivatives can also be defined, provided they are taken in the sense of generalized functions.

The spline MRA is obtained by defining the subspaces

$$V_j = \text{closure of span}\,\{\beta_{j,k}^{(M-1)} \,|\, k \in \mathbb{Z}\}$$

for $j \in \mathbb{Z}$.

The following recursive formula provides the simplest definition for the central B-splines,

$$\beta^{(m)}(x) = \int_{-\infty}^{\infty} \beta^{(m-1)}(x-t)\beta^{(0)}(t)\, dt\,, \quad m \geq 1$$

where $\beta^{(0)}(x)$ is the characteristic function of the interval $[-1/2, 1/2)$. (Here we allow m to be any non-negative integer.) We can also write

$$\beta^{(m)}(x) = \int_{-1/2}^{1/2} \beta^{(m-1)}(x-t)\,dt\,, \quad m \geq 1\,.$$

Using the fact that $\beta^{(m)}$ is an $(m+1)$-fold convolution of the characteristic function $\chi_{[-1/2,1/2)}$, it is a simple exercise to compute the Fourier transform,

$$\hat{\beta}^{(m)}(\xi) = \left(\hat{\beta}^{(0)}(\xi)\right)^{m+1} = \left(\frac{\sin \xi/2}{\xi/2}\right)^{m+1}. \tag{2.76}$$

Let us derive the two-scale difference equation (2.9) for the central B-splines of odd degree. Using a trigonometric identity in (2.76), we have

$$\begin{aligned} \hat{\beta}^{(M-1)}(\xi) &= \left(\frac{\sin \xi/4 \cos \xi/4}{\xi/4}\right)^{M} \\ &= \hat{\beta}^{(M-1)}(\xi/2) m_0(\xi/2) \end{aligned}$$

where

$$m_0(\xi) = (\cos \xi/2)^M\,. \tag{2.77}$$

We can also express m_0 in the following form,

$$\begin{aligned} m_0(\xi) &= \left(\frac{e^{i\xi/2} + e^{-i\xi/2}}{2}\right)^{M} \\ &= e^{-iM\xi/2}\left(\frac{1+e^{i\xi}}{2}\right)^{M} \\ &= \frac{1}{2^M} \sum_{k=-M/2}^{M/2} \binom{M}{M/2+k} e^{ik\xi}\,. \end{aligned} \tag{2.78}$$

It follows that (cf. equation (2.10))

$$\beta^{(M-1)}(x) = \frac{1}{2^{M-1}} \sum_{k=-M/2}^{M/2} \binom{M}{M/2+k} \beta^{(M-1)}(2x-k)\,. \tag{2.79}$$

Note that the function $m_0(\xi)$ associated to $\beta^{(M-1)}$ has M zeroes at $\xi = \pi$, as can be seen from (2.78). Thus the MRA generated by $\beta^{(M-1)}$ has M vanishing moments, and in accordance with Proposition(2.2.1) we have

$$\sum_{-\infty}^{\infty} (x-k)^m \beta^{(M-1)}(x-k) = \mu_m\,, \quad 0 \leq m \leq M-1 \tag{2.80}$$

where μ_m denotes the mth moment of $\beta^{(M-1)}(x)$.

2.4.1 The Battle-Lemarié Scaling Function

The scaling function for the orthonormal system in a spline MRA is known as the Battle-Lemarié scaling function (see e.g. [8, pp.146-152]). This scaling function is the result of applying the orthogonalization procedure (2.7) to the B-spline. The Battle-Lemarié function is thus defined in the Fourier domain by the equation

$$\hat{\phi}^{(M-1)}(\xi) = \frac{\beta^{(M-1)}(\xi)}{\sqrt{a^{(M-1)}(\xi)}}\,, \tag{2.81}$$

where

$$a^{(M-1)}(\xi) = \sum_{-\infty}^{\infty} |\hat{\beta}^{(M-1)}(\xi + 2l\pi)|^2\,. \tag{2.82}$$

Using (2.76) we can write

$$a^{(M-1)}(\xi) = \sum_{-\infty}^{\infty} \hat{\beta}^{(2M-1)}(\xi + 2l\pi)\,. \tag{2.83}$$

Using Poisson's summation formula we have

$$\sum_{-\infty}^{\infty} \hat{\beta}^{(2M-1)}(\xi + 2l\pi) = \sum_{-\infty}^{\infty} \beta^{(2M-1)}(k)e^{ik\xi}\,, \tag{2.84}$$

and since $\beta^{(2M-1)}(k)$ is nonzero for $k \in \mathbb{Z}$ if and only if $|k| \leq M-1$, we can rewrite (2.83) as

$$\begin{aligned} a^{(M-1)}(\xi) &= \sum_{1-M}^{M-1} \beta^{(2M-1)}(k)e^{ik\xi} &(2.85)\\ &= \beta^{(2M-1)}(0) + 2\sum_{k=1}^{M-1} \beta^{(2M-1)}(k)\cos k\xi\,. &(2.86) \end{aligned}$$

For example, associated to the hat function (2.75) we have

$$\begin{aligned} a^{(1)}(\xi) &= \sum_{-1}^{1} \beta^{(3)}(k)e^{ik\xi} \\ &= \frac{2}{3} + \frac{1}{6}\left(e^{i\xi} + e^{-i\xi}\right) \\ &= \frac{2}{3} + \frac{1}{3}\cos\xi\,. \end{aligned} \tag{2.87}$$

It is easily seen that $1/3 \leq a^{(1)}(\xi) \leq 1$.

2.4.2 The Dual of the B-spline

In spite of its importance as the scaling function for the orthonormal system, the Battle-Lemarié scaling function is difficult to work with directly. This is due both to the fact that it is not compactly supported in the space variable x, and to the presence of the square root in the Fourier transform (2.81). In practice we prefer to work always with the compactly supported B-spline.

The B-spline is obviously not orthogonal to its translates. However, it is possible to construct another scaling function γ with the following property,

$$\begin{aligned}(\beta^j_m, \gamma^j_n) &= 2^{-j} \int_{-\infty}^{\infty} \beta(2^{-j}x - m)\gamma(2^{-j}x - n)\, dx \\ &= \int_{-\infty}^{\infty} \beta(x - m)\gamma(x - n)\, dx \\ &= \delta_{m,n}\,. \end{aligned} \tag{2.88}$$

The two functions β and γ form a basis known as a "biorthogonal system", and the two scaling functions are known as "duals" of each other. For the sake of completeness one may consider an orthonormal scaling function to be its own dual. It is important here that the dual may be chosen so that translates of β and γ span the same subspace. This is not true of all biorthogonal systems. (For a study of these systems, we refer the reader to [6]).

Thus, for example, the coefficients of the projection of a function f onto the subspace V_0,

$$(P_0 f)(x) = \sum_{k=-\infty}^{\infty} s^0_k \beta(x - k)$$

are given by

$$s^0_k = \int_{-\infty}^{\infty} f(x)\gamma(x - k)\, dx\,.$$

Let us derive an expression for the Fourier transform of γ. First, we want $\gamma \in V_0$. This means that

$$\gamma(x) = \sum_{k=-\infty}^{\infty} c_k \beta(x - k)$$

for some coefficients $\{c_k\}$. Taking the Fourier transform on both sides of this expression we obtain

$$\hat{\gamma}(\xi) = \hat{c}(\xi)\hat{\beta}(\xi)\,, \tag{2.89}$$

where $\hat{c}$ denotes the formal trigonometric series

$$\hat{c}(\xi) = \sum_{k=-\infty}^{\infty} c_k e^{ik\xi}\,.$$

Using Poisson's summation formula, the relationship (2.88) can be expressed in terms of the Fourier transform as

$$\sum_{l=-\infty}^{\infty} \hat{\gamma}(\xi + 2l\pi)\overline{\hat{\beta}}(\xi + 2l\pi) = 1\,, \quad \text{a.e.} \tag{2.90}$$

Substituting (2.89) into (2.90), and using the fact that the B-splines form a Riesz basis (which implies that $0 < c_1 \leq \sum |\hat{\beta}(\xi + 2l\pi)|^2 \leq c_2$), we have

$$\sum_{l=-\infty}^{\infty} [\hat{\beta}(\xi + 2l\pi)\hat{c}(\xi)]\overline{\hat{\beta}}(\xi + 2l\pi) = 1\,,$$

which is satisfied if we take

$$\hat{c}(\xi) = \frac{1}{\sum |\hat{\beta}(\xi + 2l\pi)|^2}\,. \tag{2.91}$$

Using (2.82), we write $\hat{c}(\xi) = 1/a(\xi)$. Thus,

$$\hat{\gamma}^{(M-1)}(\xi) = \frac{\hat{\beta}^{(M-1)}(\xi)}{a^{(M-1)}(\xi)}\,. \tag{2.92}$$

2.4.3 Spline Approximation of Functions and Kernels

In this section we present a practical method for the computation of coefficients to represent functions and kernels using the compactly supported B-splines. Our goal is to compute coefficients for representations of the form

$$(P_j f)(x) = \sum_{k=-\infty}^{\infty} s_k^j \beta_k^j(x)\,, \tag{2.93}$$

and

$$T_j(x,y) = P_j K P_j = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} t_{k-l}^j \beta_k^j(x) \beta_l^j(y)\,. \tag{2.94}$$

For this it is necessary to compute the integrals

$$s_k^j = \int_{-\infty}^{\infty} f(x)\gamma_k^j(x)\,dx\,, \tag{2.95}$$

$$t_n^j = \int_{-\infty}^{\infty} K(x)\Gamma(2^{-j}x - n)\,dx\,. \tag{2.96}$$

The function Γ is the autocorrelation of the dual scaling function γ.

The functions γ and Γ are not compactly supported in the space variable x, and hence these integrals are difficult to evalute directly. However, they can be approximated as follows.

Let M be an even, positive integer and let $(M-1)$ be the degree of the B-spline β. Construct two sets of numbers $\{q_m\}_{m=0}^{M-1}$ and $\{Q_m\}_{m=0}^{2M-1}$ according to the formulae

$$\sum_{n=0}^{m} \binom{m}{n} (-1)^n q_{m-n} \mu_n = \begin{cases} 1, & m = 0 \\ 0, & m > 0 \end{cases} \qquad \text{for } 0 \leq m \leq M-1 \tag{2.97}$$

and

$$\sum_{n=0}^{m} \binom{m}{n} (-1)^n Q_{m-n} \mathcal{M}_n = \begin{cases} 1, & m = 0 \\ 0, & m > 0 \end{cases} \qquad \text{for } 0 \leq m \leq 2M-1 \tag{2.98}$$

where μ_n denotes the nth moment of the B-spline of degree $(M-1)$ and $\mathcal{M}_n$ denotes the nth moment of its autocorrelation, i.e. the B-spline of degree $(2M-1)$.

Define the coefficients

$$s_k^j = 2^{j/2} \sum_{m=0}^{M-1} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} q_m \tag{2.99}$$

and

$$t_n^j = 2^j \sum_{m=0}^{2M-1} \frac{K^{(m)}(2^j n)}{m!} 2^{mj} Q_m \tag{2.100}$$

which are to be used in the representations (2.93) and (2.94), respectively. If derivatives are not available, then finite difference formulas may be used instead.

We prove the following two propositions.

Proposition 2.4.1 *Let β be the central B-spline of degree $(M-1)$ where M is an even positive integer. Let $P_j f$ denote the projection of a function f onto the subspace V_j,*

$$(P_j f)(x) = \sum_{k=-\infty}^{\infty} s_k^j \beta_k^j(x) \tag{2.101}$$

where the coefficients are given by (2.99). For a given point $x \in \mathbf{R}$, let $I_j(x)$ be the interval formed by the union of the supports of all basis functions which are non-zero at x. Thus

$$I_j(x) = \bigcup_{k \in \mathcal{K}} \operatorname{supp}(\beta_k^j), \qquad \mathcal{K} = \{k \in \mathbf{Z} | \beta_k^j(x) \neq 0\}.$$

Suppose that f is at least M times continuously differentiable on $I_j(x)$. Then we have

$$(P_j f)(x) = f(x) + E_j(x), \tag{2.102}$$

where

$$|E_j(x)| \leq 2^{Mj} C \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!} . \tag{2.103}$$

The constant C depends on β but not on f.

Proof: Let x_0 be an arbitrary real number. Expand f in a Taylor series about x_0 to obtain the expression for the mth derivative,

$$f^{(m)}(x) = \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(M)}(\xi_m)}{M!}(x-x_0)^M ,$$

where $0 \leq m \leq M-1$, and ξ_m lies between x and x_0. Set $x = 2^j k$ to obtain

$$f^{(m)}(2^j k) = \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x_0)}{n!}(2^j k - x_0)^n + \frac{f^{(M)}(\xi_{m,k})}{M!}(2^j k - x_0)^M ,$$

where $\xi_{m,k}$ lies between x_0 and $(2^j k)$ for each m. Now use this expression together with (2.99) to write

$$\begin{aligned}
(P_j f)(x_0) &= 2^{-j/2} \sum_k s_k^j \beta(2^{-j} x_0 - k) \\
&= \sum_k \beta(2^{-j}x_0 - k) \sum_{m=0}^{M-1} \frac{f^{(m)}(2^j k)}{m!} \, 2^{mj} q_m \\
&= \sum_k \beta(2^{-j}x_0 - k) \sum_{m=0}^{M-1} \frac{q_m}{m!} \left\{ \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x_0)}{n!}(2^j k - x_0)^n \right. \\
&\quad \left. + \frac{f^{(M)}(\xi_{m,k})}{M!}(2^j k - x_0)^M \right\} \\
&= \sum_{m=0}^{M-1} q_m \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x_0)}{m!n!} 2^{(m+n)j}(-1)^n \\
&\quad \times \sum_k (2^{-j}x_0 - k)^n \beta(2^{-j}x_0 - k) + E_j(x_0) ,
\end{aligned} \tag{2.104}$$

where we have put

$$E_j(x) = \sum_{m=0}^{M-1} \frac{q_m}{m!} \sum_k \frac{f^{(M)}(\xi_{m,k})}{M!}(2^j k - x)^M \beta(2^{-j}x - k) . \tag{2.105}$$

Continuing from (2.104), and using the identity (2.80), we have

$$\begin{aligned}
(P_j f)(x_0) - E_j(x_0) &= \sum_{m=0}^{M-1} q_m \sum_{n=0}^{M-1-m} \frac{f^{(n+m)}(x_0)}{m!n!} 2^{(m+n)j}(-1)^n \mu_n \\
&= \sum_{m=0}^{M-1} q_m \sum_{n=0}^{M-1-m} \binom{n+m}{m} \frac{f^{(n+m)}(x_0)}{(n+m)!} 2^{(m+n)j}(-1)^n \mu_n .
\end{aligned}$$

Now use (2.47) to transform the right-hand side of this equation, thus obtaining

$$\begin{aligned}(P_j f)(x_0) - E_j(x_0) &= \sum_{m=0}^{M-1} \frac{f^{(m)}(x_0)}{m!} 2^{mj} \sum_{n=0}^{m} \binom{m}{n} (-1)^n q_{m-n}\mu_n \\ &= f(x_0)\,,\end{aligned}$$

where we have used (2.97). This proves (2.102).

Finally, we observe that, since β is compactly supported, there is a constant C_M such that

$$\sum_{-\infty}^{\infty} |x-k|^M |\beta(x-k)| \le C_M\,,$$

by Lemma(2.2.1). Thus, from (2.105) we have

$$\begin{aligned}|E_j(x)| &\le \left(\sum_{m=0}^{M-1} \frac{|q_m|}{m!} 2^{mj}\right) 2^{Mj} \sum_k |2^{-j}x - k|^M |\beta(2^{-j}x-k)| \frac{|f^{(M)}(\xi_{m,k})|}{M!} \\ &\le 2^{Mj} C' C_M \sup_{\xi \in I_j(x)} \frac{|f^{(M)}(\xi)|}{M!}\,,\end{aligned}$$

where $C' = \sum_{m=0}^{M-1} 2^{mj}(|q_m|/m!)$. This verifies (2.103). □

Proposition 2.4.2 *Let β be the central B-spline of degree $(M-1)$, where M is an even positive integer. Let $T_j = P_j K P_j$ denote the projection of a kernel K onto the subspace V_j, given by*

$$T_j(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} t^j_{k-l} \beta^j_k(x) \beta^j_l(y) \tag{2.106}$$

where the coefficients are given by (2.100). For a given point $(x,y) \in \mathbb{R}^2$, let $R_j(x,y)$ be the rectangle formed by the union of the supports of all basis functions which are non-zero at (x,y). Thus, if

$$I_j(x) = \bigcup_{k\in\mathcal{K}} \operatorname{supp}(\beta^j_k)\,, \qquad \mathcal{K} = \{k \in \mathbb{Z} | \beta^j_k(x) \neq 0\}$$

and

$$I_j(y) = \bigcup_{l\in\mathcal{L}} \operatorname{supp}(\beta^j_l)\,, \qquad \mathcal{L} = \{l \in \mathbb{Z} | \beta^j_l(y) \neq 0\}$$

then

$$R_j(x,y) = I_j(x) \times I_j(y)\,.$$

Suppose that K is at least M times continuously differentiable on $R_j(x,y)$. Then we have

$$T_j(x,y) = K(x-y) + E_j(x,y)\,, \tag{2.107}$$

where

$$|E_j(x,y)| \le 2^{Mj} C \sup_{(\xi,\eta)\in R_j(x,y)} \frac{|K^{(M)}(\xi-\eta)|}{(M)!}\,. \tag{2.108}$$

The constant C depends on β but not on K.

The proof of this proposition is similar to those for Propositions (2.4.1) and (2.3.1), and we omit the details.

The expansions (2.97) and (2.98) may be viewed as asymptotic expansions for the integrals in (2.95) and (2.96), respectively, obtained by straightforward Taylor expansion. For example, formally expanding a function f about $(2^j k)$ we obtain

$$f(x) = \sum_{m=0}^{\infty} \frac{f^{(m)}(2^j k)}{m!}(x-2^j k)^m\,.$$

Substituting this into the integral (2.95) we have

$$\begin{aligned}
s_k^j &= 2^{-j/2}\int_{-\infty}^{\infty} \gamma(2^{-j}x-k) \sum_{m=0}^{\infty} \frac{f^{(m)}(2^j k)}{m!}(x-2^j k)^m\,dx \\
&= 2^{-j/2} \sum_{m=0}^{\infty} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} \int_{-\infty}^{\infty} (2^{-j}x-k)^m \gamma(2^{-j}x-k)\,dx \\
&= 2^{j/2} \sum_{m=0}^{\infty} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} \int_{-\infty}^{\infty} x^m \gamma(x)\,dx \\
&= 2^{j/2} \sum_{m=0}^{\infty} \frac{f^{(m)}(2^j k)}{m!} 2^{mj} q_m
\end{aligned}$$

where we define q_m to be the mth moment of the scaling function γ,

$$q_m = \int_{-\infty}^{\infty} x^m \gamma(x)\,dx\,.$$

It is not practical to compute the moments $\{q_m\}$ directly from this expression, because γ is not compactly supported. However, we can determine the first M moments as follows.

Let us define the function A to be the correlation of the B-spline and its dual,

$$A(x) = \int \beta(x+y)\gamma(y)\,dy\,.$$

The Fourier transform of A is then given by

$$\begin{aligned}
\hat{A}(\xi) &= \hat{\beta}(\xi)\overline{\hat{\gamma}}(\xi) \\
&= \frac{|\hat{\beta}(\xi)|^2}{a(\xi)}
\end{aligned}$$

which follows from (2.92). Comparison of this expression with (2.81) shows that

$$\hat{A}(\xi) = |\hat{\phi}(\xi)|^2$$

where ϕ is the Battle-Lemarié scaling function. Thus the correlation function A is equal to the autocorrelation of the Battle-Lemarié scaling function almost everywhere, which by Proposition(2.1.1) must have vanishing moments. Let $\mathcal{A}_m$ denote the mth moment of A. Then we have

$$\mathcal{A}_m = \begin{cases} 1\,, & m = 0 \\ 0\,, & m > 0 \end{cases}$$

for $0 \leq m \leq 2M - 1$, and as in Lemma(2.1.1) we have

$$\mathcal{A}_m = \sum_{n=0}^{m} \binom{m}{n} (-1)^n q_{m-n} \mu_n \,.$$

Since the moments of the B-spline $\{\mu_n\}$ are easy to compute, (use the recursive formula in equation 2.22) we can use the above two equations to obtain $\{q_m\}_{m=0}^{2M-1}$.

Chapter 3

The Fast Summation Algorithm in One Dimension

In this chapter we describe our approach for problems in one dimension. Our approach in higher dimensions is similar. Indeed, to develop an algorithm in two dimensions, the only additional machinery used is singular value decomposition of the coefficient matrix, but this will be discussed in Chapter(4).

3.1 General Description

1. Our goal is to compute the numbers $\{g_m\}$, where

$$g_m = \sum_{\substack{n=1 \\ n \neq m}}^{N} K(x_m - x_n) f_n \,, \qquad 1 \leq m \leq N \,. \tag{3.1}$$

We assume that the kernel K is singular on the line $x - y = 0$, i.e.

$$|K(x-y)| \to \infty \quad \text{as} \quad (x-y) \to 0 \,,$$

but is at least M times continuously differentiable on any region that does not contain the line $x = y$. Moreover, we assume that there is a "bandwidth" B such that the Mth derivative is uniformly bounded outside the band $|x - y| < B$, i.e.

$$|K^{(M)}(x-y)| \leq C \quad \text{for } |x-y| \geq B. \tag{3.2}$$

2. We choose a level of refinement ($j \leq 0$), a multiresolution analysis with M vanishing moments, and construct the projection of the kernel onto the subspace V_j, given by

$$T_j(x,y) = \sum_k \sum_l t^j_{k-l} \phi^j_k(x) \phi^j_l(y) \,. \tag{3.3}$$

As discussed in Chapter(2) this construction requires only the computation of the coefficients $\{t_n^j\}$, given by

$$t_n^j = \int_{-\infty}^{\infty} K(x)\Phi(2^{-j}x - n)\, dx\,, \quad n \in \mathbb{Z}$$

where Φ is the autocorrelation of the scaling function ϕ. This can be done ahead of time and the coefficients $\{t_n^j\}$ are then stored in memory. This portion of the computation is part of the initialization.

3. Using Proposition(2.3.1) together with (3.2) we have the estimate

$$|K(x-y) - T_j(x,y)| \leq \epsilon \quad \text{for} \quad |x-y| \geq 2^j B\,, \tag{3.4}$$

for any $\epsilon > 0$, provided that M and $j \leq 0$ have been chosen so that $C(2^{Mj}/M!) < \epsilon$. (There is some abuse of notation here, since the constants C and B are not necessarily identical to the constants that appear in (3.2), but this detail is not important.)

4. The first step, which we refer to as the **"low frequency approximation"**, is the application of the kernel T_j to the particle configuration

$$f(x) = \sum_{n=1}^{N} f_n \delta(x - x_n)\,. \tag{3.5}$$

Thus, we compute the approximations

$$g_m^j = \sum_{\substack{n=1\\ n\neq m}}^{N} T_j(x_m, x_n) f_n \tag{3.6}$$

for $m = 1, \ldots, N$. The most expensive part of this computation is the application of a dense Toeplitz matrix to a vector, and is done using the FFT. Substituting (3.3) into (3.6) we have

$$\begin{aligned} g_m^j &= \sum_{\substack{n=1\\ n\neq m}}^{N} f_n \sum_k \sum_l t_{k-l}^j \phi_k^j(x_m)\phi_l^j(x_n) \\ &= \sum_k \phi_k^j(x_m) \sum_l t_{k-l}^j \sum_{n=1}^{N} f_n \phi_l^j(x_n) - T_j(x_m, x_m) f_m \\ &= \sum_k \hat{s}_k^j \phi_k^j(x_m) - T_j(x_m, x_m) f_m\,, \end{aligned} \tag{3.7}$$

where

$$\hat{s}_k^j = \sum_l t_{k-l}^j s_l^j \tag{3.8}$$

and

$$s_l^j = \sum_{n=1}^{N} f_n \phi_l^j(x_n)\,. \tag{3.9}$$

The term $T_j(x_m, x_m) f_m$ is the self-interaction and must be subtracted after the coefficients (3.8) of the product have been computed and the sum in (3.7) has been evaluated. Note that $T_j(x_m, x_m)$ is finite, as T_j is a regularized version of K.

5. Let us compare the approximations $\{g_m^j\}$ to the exact values $\{g_m\}$. For a given index m, let us split the indices $1, \ldots, N$ into two distinct sets, depending on the distance between the point x_m and the points $\{x_n\}$. All indices corresponding to particles that are sufficiently removed from x_m we place in

$$S_{j,m}^{far} = \{n : |x_n - x_m| \geq 2^j B\}\,.$$

All indices corresponding to particles that are near to x_m we place in

$$S_{j,m}^{near} = \{n : |x_n - x_m| < 2^j B\}\,.$$

Now the approximation (3.6) can be written

$$g_m^j = \left(\sum_{n \in S_{j,m}^{far}} + \sum_{n \in S_{j,m}^{near}} \right) T_j(x_m, x_n) f_n\,. \tag{3.10}$$

Comparing the above expression to the actual values (3.1) we have

$$\begin{aligned} |g_m^j - g_m| &\leq \sum_{n \in S_{j,m}^{far}} |f_n| |T_j(x_m, x_n) - K(x_m - x_n)| \\ &+ \sum_{n \in S_{j,m}^{near}} |f_n| |T_j(x_m, x_n) - K(x_m - x_n)|\,. \end{aligned} \tag{3.11}$$

For particle indices contained in the index set $S_{j,m}^{far}$ we can use the estimate (3.4). Thus

$$\begin{aligned} \sum_{n \in S_{j,m}^{far}} |f_n| |T_j(x_m, x_n) - K(x_m - x_n)| &\leq \mathcal{F} \max_{n \in S_{j,m}^{far}} |T_j(x_m, x_n) - K(x_m - x_n)| \\ &\leq \mathcal{F} \epsilon\,, \end{aligned} \tag{3.12}$$

where we have put

$$\mathcal{F} = \sum_{n=1}^{N} |f_n|\,. \tag{3.13}$$

6. We cannot, however, give a similar bound for the error in the second sum. This is because the second sum involves particle separations which are small, and hence points (x_m, x_n) which lie near the singularity of the kernel K. The projection T_j cannot be expected to provide an acceptable approximation to K in this region. The contribution to the approximation (3.10) from particles near x_m must be corrected, and we refer to this second step of our algorithm as the **"high frequency correction"**. A second operator must now be applied that preserves the accuracy of the initial approximation for particles that are far apart but corrects the errors due to particles that are close together.

7. The correction operator is defined by

$$C_j(x,y) = \begin{cases} K(x-y) - T_j(x,y)\,, & 0 < |x-y| < 2^j B \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3.14}$$

We update the approximation (3.10) according to the formula

$$\begin{aligned} g_m^j &\leftarrow g_m^j + \sum_{n=1}^{N} C_j(x_m, x_n) f_n \\ &= \sum_{n \in S_{j,m}^{far}} T_j(x_m, x_n) f_n + \sum_{n \in S_{j,m}^{near}} \left[C_j(x_m, x_n) + T_j(x_m, x_n) \right] f_n\,, \end{aligned}$$

for $m = 1, \ldots, N$. It now follows from (3.12) and (3.14) that

$$|g_m^j - g_m| < \mathcal{F}\epsilon\,,$$

for $m = 1, \ldots, N$.

3.2 Additional Details

Current implementations of the algorithm use B-splines as the basis functions. This is primarily due to the simplicity of obtaining values of the B-splines at a given point. This is not a straightforward affair for many other families of scaling functions, since explicit expressions are not available and one must exploit the two-scale difference equation or perhaps numerical integration of the Fourier transform. This property of the B-splines greatly enhances their utility in this context.

3.2.1 Low Frequency Approximation

1. We assume that the particle locations $\{x_n\}_{n=1}^N$ lie in the unit interval $[0,1]$ and are ordered. Thus, we have

$$0 < x_1 < x_2 < \cdots < x_N < 1\,.$$

2. The number of vanishing moments M corresponds to a B-spline of degree $(2m-1)$. More precisely, we have $M-1=2m-1$, so that M must be an even, positive integer. Once this parameter and the level of refinement $j \le 0$ have been chosen, the projection of the kernel K is

$$T_j(x,y) = \sum_{k=0}^{J-1} \sum_{l=0}^{J-1} t^j_{k-l} \beta^j_k(x) \beta^j_l(y)\,, \tag{3.15}$$

where $J = 2^{-j}$, and β denotes the B-spline of degree $(2m-1)$. The coefficients $\{t^j_n\}$ for $|n| \le (J-1)$ are computed using the formula (2.100).

3. Note that, in anticipation of using the FFT, we have restricted the indices in (3.15) to lie in the range $0 \le k, l \le 2^{-j}-1$. This is to ensure that the dimension of the resulting matrix is a power of two, which is convenient when using the FFT. (This is also the reason for choosing the interval $[0,1]$.) However, this means that we must restrict the particle locations so that no particles lie in the supports of the basis functions β^j_l for $l<0$ or $l \ge 2^{-j}$. (If the kernel K is periodic then this restriction is not necessary.) Since the support of $\beta^{(2m-1)}$ is the interval $[-m,m]$, the restriction is

$$2^j(m-1) \le x_1 < x_2 < \cdots < x_N \le 1 - 2^j m\,. \tag{3.16}$$

4. The coefficients of the projection of the particle configuration (3.5) are given by (3.9), which here takes the form

$$s^j_l = \sum_{n=1}^{N} f_n \beta^j_l(x_n)\,, \quad 0 \le l \le J-1\,. \tag{3.17}$$

The values of the B-spline may be computed using the following recursion formula,

$$\beta^{(m)}(x) = \frac{(m+1)/2 + x}{m} \beta^{(m-1)}(x+1/2) + \frac{(m+1)/2 - x}{m} \beta^{(m-1)}(x-1/2)$$

for $m \ge 1$, where $\beta^{(0)}(x)$ is the characteristic function of the interval $[-1/2, 1/2)$. This formula may be found in, for example, [3], and a derivation of the formula in a slightly different form may be found in [5]. However, for practical implementation a significant savings in speed may be realized by setting up an interpolation table for the B-spline.

5. Next we compute the matrix-vector product

$$\hat{s}^j_k = \sum_{l=0}^{J-1} t^j_{k-l} s^j_l\,, \quad 0 \le k \le J-1 \tag{3.18}$$

using the FFT. If the kernel K is not periodic we use an FFT of length $(2J)$.

6. Having obtained the coefficients (3.18) we can now evaluate the expansion

$$g^j(x) = \sum_{k=0}^{J-1} \hat{s}^j_k \beta^j_k(x)$$

at the points $x_1, \ldots, x_N$ to obtain the approximations

$$g^j_m = g^j(x_m) = \sum_{k=0}^{J-1} \hat{s}^j_k \beta^j_k(x_m)\,. \qquad (3.19)$$

3.2.2 High Frequency Correction

The first step is to determine, for a given index m, all those particle locations $\{x_n\}$ such that $|x_m - x_n| < 2^j B$. However, this poses no difficulty when the particle locations are ordered as in (3.16). Thus, for each index m, we determine the set of indices $S^{near}_{j,m} = \{n : |x_m - x_n| < 2^j B\}$. The band-width B is determined numerically.

2. Next we compute the correction, and use it to update the approximation g^j_m,

$$g^j_m \leftarrow g^j_m + \sum_{n \in S^{near}_{j,m}} C_j(x_m, x_n) f_n\,, \quad 1 \le m \le N\,. \qquad (3.20)$$

In order to evaluate the kernel $T_j(x, y)$, we use the series expansion established by Theorem(2.3.1). Since we use B-splines as basis functions, the B-spline having a real Fourier transform allows us to write the expansion in the form (2.74). This series converges very quickly, and we typically retain only the first four terms. Hence, we use the truncated series

$$\begin{aligned} 2^j T_j(2^j x, 2^j y) \approx I^j_0(z) &+ 2\cos(\pi w) I^j_1(z) \\ &+ 2\cos(2\pi w) I^j_2(z) + 2\cos(3\pi w) I^j_3(z) \end{aligned} \qquad (3.21)$$

where

$$z = x - y\,, \qquad w = x + y$$

and where

$$I^j_n(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iz\xi} \hat{t}^j(\xi + p_n) \hat{\Phi}_n(\xi)\, d\xi\,, \qquad (3.22)$$

for $n \ge 0$, with $p_n = \pi$ if n is odd and $p_n = 0$ if n is even. In order to avoid computing cosines, and to optimize the evaluation of (3.21), we use a summation technique found in [14]. For example, to compute $C = \sum_{k=0}^{3} I_k \cos(k\pi w)$, we put $r = \cos(\pi w)$, and then perform the following steps,

$$\begin{aligned} u_3 &= I_3\,, \\ u_2 &= I_2 + 2ru_3\,, \\ u_1 &= I_1 + 2ru_2 - u_3\,, \\ C &= I_0 + ru_1 - u_2\,. \end{aligned}$$

3. The functions $\hat{\Phi}_n$ are given in this instance by

$$\hat{\Phi}_n(\xi) = \hat{\beta}(\xi + n\pi)\hat{\beta}(\xi - n\pi)\,. \tag{3.23}$$

Note that for $n = 0$, (3.23) is the autocorrelation of the B-spline. For the B-splines, explicit expressions in terms of the space variable z can be derived for the functions

$$\Phi_n(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iz\xi}\hat{\Phi}_n(\xi)\,d\xi\,.$$

In the Appendix B, it is shown that

$$\Phi_n(z) = \sum_{k=0}^{2m-1} \frac{(-i)^{2m-k} a_{2m,k}}{(2n\pi)^{2m+k}} \left[e^{in\pi x} + (-1)^k e^{-in\pi x}\right] \left(\frac{d}{dz}\right)^{2m+k} \beta^{(4m-1)}(z)\,,$$

when the basis function in (3.23) is the B-spline of degree $(2m-1)$, where

$$a_{m,k} = \frac{(m-1+k)!}{(m-1)!\,k!}\,.$$

We then rewrite (3.22) as

$$I_n^j(z) = \sum_k t_k^j \Phi_n(z-k)\,. \tag{3.24}$$

The functions $I_n^j(z)$, so defined, are easily approximated by polynomials. Thus we build an interpolation table for each of $I_0^j(z), I_1^j(z), I_2^j(z)$, and $I_3^j(z)$.

3.2.3 Complexity Analysis

Suppose a system of N particles, located at the points $x_1, \ldots, x_N$, carrying "charges" $f_1, \ldots, f_N$, respectively. The level of refinement, or projection scale, is $j \leq 0$, and $J = 2^{-j}$. The basis function is the central B-spline of degree $(2m_d - 1)$, where $m_d = 1, 2, \ldots$.

Let $\lfloor x \rfloor$ denote the greatest integer that is less than or equal to x.

We make the following estimates on the number of operations required by each step of our algorithm.

Step	Procedure	Complexity
1.	Project particles onto the basis: for each integer $n = 1, \ldots, N$ we compute the values $\beta_l^j(x_n)$ for $2^{-j}\lfloor x_n \rfloor - m_d < l < 2^{-j}\lfloor x_n \rfloor + m_d$.	$O(2m_d \cdot N)$
2.	Use the result of Step 1 to obtain the coefficients of the projection, $s_l^j = \sum_{n=1}^{N} f_n \beta_l^j(x_n)$, for $0 \leq l \leq J-1$.	$O(2m_d \cdot N)$
3.	Apply the projection of the kernel via the FFT to obtain the coefficients of the product, $\hat{s}_k^j = \sum_{l=0}^{J-1} t_{k-l}^j s_l^j$, for $0 \leq k \leq J-1$. (If the kernel K is periodic, then the matrix of coefficients $\{t_{k-l}^j\}$ is a circulant, and the resulting estimate is $O(J \log_2 J)$.)	$O(2J \log_2 2J)$
4.	Use the result of Step 3 to obtain the low frequency approximations, $g_n^j = \sum_{k=0}^{J-1} \hat{s}_k^j \beta_k^j(x_n)$, for $1 \leq n \leq N$.	$O(2m_d \cdot N)$
5.	Determine the set of indices $S_{j,m}^{near} = \{n : \lvert x_m - x_n \rvert < 2^j B\}$ for each $m = 1, \ldots, N$. We assume that $S_{j,m}^{near} \leq s_j$ for each m.	$O(s_j N)$
6.	Compute the high frequency corrections and update the approximations obtained in Step 4, $g_m^j \leftarrow g_m^j + \sum_{n \in S_{j,m}^{near}} C_j(x_m, x_n) f_n$, for $1 \leq m \leq N$. There are four terms to be evaluated in the series (3.21) for each point (x_m, x_n), and each term is approximated by an interpolating polynomial of degree $ideg$.	$O(s_j N \cdot 4 \cdot ideg)$

Steps 1 through 4 constitue the low frequency approximation, while Steps 5 and 6 constitute the high frequency correction.

Now let us relate the parameters J and s_j to the number of particles N. Suppose that the particles are evenly spaced, with a stepsize 2^{j_0}, for some integer $j_0 < 0$. Let the particle locations be given by $x_n = 2^{j_0} n$, for $0 \leq n \leq N-1$, where $N = 2^{-j_0}$. The projection scale is $j < 0$, and $J = 2^{-j}$, as above. The inequality $\lvert x_m - x_n \rvert < 2^j B$ (that appears in, e.g. (3.14)) becomes

$$\lvert x_m - x_n \rvert = 2^{j_0} \lvert m - n \rvert < 2^j B \,,$$

or equivalently

$$\lvert m - n \rvert < \left(\frac{N}{J} \right) B \,.$$

Thus, for a given index m, the number of indices n that fall into the indicated range is approximately $2B(N/J)$, and we estimate $s_j \approx 2B(N/J)$, where s_j is defined in Step 5 above. The operation count for Step 6 is now proportional to the quantity

$$2B\left(\frac{N^2}{J}\right)\cdot 4\cdot ideg\,.$$

In order to obtain an operation count proportional to N for this step it is necessary that $J = O(N)$. Let us assume that the ratio (J/N) is a power of two. We refer to this ratio as the "oversampling factor", and assume that $(J/N) = 2^s$, or $J = 2^s N$, where s is an integer. The right-hand column of the above table can now be rewritten to provide the following estimates.

Step	**Complexity**
1.	$O\left(2m_d \cdot N\right)$
2.	$O\left(2m_d \cdot N\right)$
3.	$O\left(2^{s+1}N\log_2 2^{s+1}N\right)$
4.	$O\left(2m_d \cdot N\right)$
5.	$O\left(2^{1-s}B \cdot N\right)$
6.	$O\left(2^{3-s}B \cdot ideg \cdot N\right)$

It is clear from the above table that in computing the low frequency approximation it is desirable to have the parameter s as small as possible, while in computing the high frequency correction it is desirable to have this parameter as large as possible. The optimal value of s varies slightly with different computer systems and implementations, but in practice we find that $s = 0, 1$, or 2 works best.

3.3 Examples

3.3.1 The Kernel $(1/x)$

As our model problem in one dimension we choose the kernel

$$K(x-y) = 1/(x-y)\,, \qquad x \neq y\,. \tag{3.25}$$

This example is closely related to the Hilbert transform,

$$(Hf)(x) = -\frac{1}{\pi}\text{p.v.}\int_{-\infty}^{\infty}\frac{f(y)}{x-y}\,dy \tag{3.26}$$

which is a bounded operator on $L^2(\mathbb{R})$ (see e.g. [22]). In the context of particle simulations, we can think of applying the integral operator with kernel function (3.25) to a sum of delta functions

$$f(x) = \sum_{n=1}^{N} f_n\delta(x-x_n)\,, \tag{3.27}$$

and then evaluating the result at the points $x_1, \ldots, x_N$. Thus, the sums to be computed are

$$g_m = \sum_{\substack{n=1 \\ n \neq m}}^{N} \frac{f_n}{x_m - x_n}, \tag{3.28}$$

for $m = 1, \ldots, N$.

The kernel (3.25) is homogeneous, satisfying $aK(ax) = K(x), a \neq 0$. It follows that the coefficients, given by

$$t_n^j = \text{p.v.} \int_{-\infty}^{\infty} \frac{\Phi(2^{-j}x - n)}{x} \, dx$$

satisfy $t_n^j = t_n^0$, and hence the projection of K onto the subspace V_j is a simple rescaling of the projection onto V_0. That is,

$$2^j T_j(2^j x, 2^j y) = T_0(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} t_{k-l} \phi(x - k) \phi(y - l) \,,$$

where

$$t_n = \text{p.v.} \int_{-\infty}^{\infty} \frac{\Phi(x - n)}{x} \, dx \,.$$

In the tables below we present the result of timing our algorithm against direct evaluation using the formula (3.28). In the tables, the parameter N is the number of particles, "T_{LF}" is the time required to compute the low frequency approximations, "T_{HF}" is the time required to compute the high frequency corrections, and "T_{tot}" is the total time for the algorithm. The errors are computed according to the following formulae. Suppose that $\mathbf{x}$ is the "exact" N-length vector, and $\tilde{\mathbf{x}}$ is the approximation. Then we compute the relative errors

$$E = \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \,,$$

where

$$\|\mathbf{x}\|_2^2 = \frac{1}{N} \sum_{i=1}^{N} |x_i|^2 \,, \qquad \|\mathbf{x}\|_\infty = \max_{1 \leq i \leq N} |x_i| \,.$$

The timings were done on 11/20/96 using a Sun Sparc-20 workstation. The "break-even point" is estimated by linear interpolation of the data in the tables. The particle locations $\{x_n\}_{n=1}^{N}$ are chosen randomly in $(0, 1)$, the distribution having a fair degree of non-uniformity.

As an evaluation of performance, we remark that these results are comparable to performance results for the same problem reported in [9].

N	T_{LF}	T_{HF}	T_{tot}	T_{dir}	E_2	E_∞
64	0.0036	0.0037	0.0073	0.0077	0.46860E-05	0.35852E-05
128	0.0066	0.0079	0.0145	0.0278	0.41383E-05	0.47034E-05
256	0.0138	0.0155	0.0293	0.1044	0.38780E-05	0.42643E-05
512	0.0264	0.0300	0.0564	0.4042	0.36356E-05	0.43642E-05
1024	0.0552	0.0591	0.1143	1.5946	0.33201E-05	0.33570E-05
2048	0.1089	0.1175	0.2264	6.3151	0.31335E-05	0.33471E-05
4096	0.2379	0.2351	0.4731	25.2104	0.33210E-05	0.31553E-05

Table 3.1: Implementation in one dimension using B-splines of degree 3, break-even at about 64 particles.

N	T_{LF}	T_{HF}	T_{tot}	T_{dir}	E_2	E_∞
64	0.0085	0.0090	0.0175	0.0083	0.24752E-09	0.14817E-09
128	0.0162	0.0160	0.0322	0.0284	0.13568E-09	0.74756E-10
256	0.0291	0.0280	0.0571	0.1049	0.57723E-10	0.60517E-10
512	0.0582	0.0526	0.1108	0.4046	0.27059E-10	0.52330E-10
1024	0.1188	0.1016	0.2204	1.6041	0.14141E-10	0.46223E-10
2048	0.2515	0.2002	0.4517	6.3056	0.81066E-11	0.41690E-10
4096	0.5187	0.4057	0.9244	25.1213	0.50117E-11	0.38148E-10

Table 3.2: Implementation in one dimension using B-splines of degree 7, break-even at about 150 particles.

3.3.2 FFT for Unequally Spaced Data

In this section we indicate how our algorithm could be used for fast evaluation of the sums

$$\hat{f}_j = \sum_{n=0}^{N-1} f(x_n)\, e^{2\pi i j x_n}\,, \qquad 0 \le j \le N-1\,. \tag{3.29}$$

The points $\{x_n\}$ satisfy $0 \le x_0 < x_1 < \cdots < x_{N-1} < 1$, but are otherwise arbitrary. Our main idea is to construct an interpolating trigonometric polynomial of the form

$$g(\theta) = \sum_{m=0}^{N-1} g_m e^{2\pi i m\theta} \tag{3.30}$$

which takes the values $\hat{f}_j$ at the equally spaced points $\theta_j = j/N$, $0 \le j \le N-1$. The sums

$$g(\theta_j) = \hat{f}_j = \sum_{m=0}^{N-1} g_m e^{2\pi i m j/N} \tag{3.31}$$

N	T_{LF}	T_{HF}	T_{tot}	T_{dir}	E_2	E_∞
64	0.0138	0.0130	0.0268	0.0086	0.50929E-14	0.26622E-14
128	0.0261	0.0260	0.0521	0.0318	0.23669E-14	0.16785E-14
256	0.0525	0.0477	0.1002	0.1244	0.43058E-14	0.18982E-14
512	0.1111	0.0925	0.2036	0.4910	0.85681E-14	0.43438E-14
1024	0.2344	0.1770	0.4114	2.0113	0.78821E-14	0.25543E-14
2048	0.5288	0.3593	0.8881	8.0242	0.91355E-14	0.27125E-14
4096	1.0738	0.7384	1.8122	32.1545	0.11375E-13	0.66191E-14

Table 3.3: Implementation in one dimension using B-splines of degree 11, break-even at about 220 particles.

can then be computed with the standard FFT. In order to determine the coefficients $\{g_m\}$, we invert (3.31) to get

$$g_m = \frac{1}{N}\sum_{j=0}^{N-1} \hat{f}_j e^{-2\pi i m j/N} .$$

Next substitute expression (3.29) for $\hat{f}_j$ to get

$$\begin{aligned} g_m &= \frac{1}{N}\sum_{j=0}^{N-1} e^{-2\pi i m j/N} \sum_{n=0}^{N-1} f_n\, e^{2\pi i j x_n} \\ &= \sum_{n=0}^{N-1} f_n \frac{1}{N}\sum_{j=0}^{N-1} e^{-2\pi i j(m/N - x_n)} , \end{aligned} \tag{3.32}$$

where $f_n = f(x_n)$. Now we have

$$\begin{aligned} \sum_{j=0}^{N-1} e^{-2\pi i j(m/N-x_n)} &= \frac{1-e^{-2\pi i(m-Nx_n)}}{1-e^{-2\pi i(m/N-x_n)}} \\ &= -2e^{i\pi N x_n}\sin(\pi N x_n) e^{i\pi(m/N-x_n)} \frac{\sin\left[\pi(\frac{m}{N}-x_n)\right]}{1-\cos\left[2\pi(\frac{m}{N}-x_n)\right]} \\ &= -e^{i\pi N x_n}\sin(\pi N x_n)\left\{\cot\left[\pi\left(\frac{m}{N}-x_n\right)\right]+i\right\} . \end{aligned}$$

Substituting this result into (3.32), we have

$$g_m = \sum_{n=0}^{N-1} \tilde{f}_n\, K(\theta_m - x_n) \,+\, i\mathcal{F} \tag{3.33}$$

for $0 \le m \le N-1$, where

$$\tilde{f}_n = -\frac{1}{N} e^{-i\pi N x_n}\sin(\pi N x_n) f(x_n) , \tag{3.34}$$

$$\mathcal{F} = \sum_{n=0}^{N-1} \tilde{f}_n \,, \tag{3.35}$$

and

$$K(x-y) = \cot\left[\pi(x-y)\right] \,, \qquad 0 < |x-y| < \frac{1}{2} \,. \tag{3.36}$$

This kernel is obviously 1-periodic. More important is the fact that this kernel can be handled by our algorithm. Thus, our algorithm is used to evaluate the sums (3.33), and once the numbers $\{g_0, \ldots, g_{N-1}\}$ are known, we use a standard FFT to compute $\{\hat{f}_0, \ldots, \hat{f}_{N-1}\}$.

We remark that the derivation given above has previously been published in [10], but of course the authors there use a different algorithm (Fast Multipole) to evaluate the sums in (3.33).

3.4 Splitting over Multiple Scales

In Section 1.2 we introduced the following splitting of the kernel K,

$$K = T_j + (K - T_j) \,. \tag{3.37}$$

Here $T_j = P_j K P_j$ is the projection of K onto the subspace V_j (cf. equations (1.5), (1.6), and (1.7); see also Sections 2.3.1 and 2.3.2). However, in principle there is no reason why we could not split the kernel over several scales, thus obtaining

$$\begin{aligned} K &= (T_j - T_{j+1}) + (T_{j+1} - T_{j+2}) + \cdots \\ &\quad + (T_{j+n-1} - T_{j+n}) + T_{j+n} + (K - T_j) \\ &= T_{j+n} + \sum_{k=0}^{n-1} (T_{j+k} - T_{j+k+1}) + (K - T_j) \,, \end{aligned} \tag{3.38}$$

where $n > 0$. Note that if $n = 0$, then (3.38) reduces to (3.37). Such a splitting may be useful when the particle distribution is strongly clustered.

In (3.38), the operator T_{j+n} corresponds to the low-frequency approximation, and $(K - T_j)$ corresponds to the high-frequency correction. These operators are applied in the same manner as described above. (Note that $(j+n)$ is the coarsest scale and j is the finest scale.) We now describe the remaining operators in (3.38).

Let us define

$$\widetilde{T}_j = T_j - T_{j+1} \,. \tag{3.39}$$

Since $T_j : V_j \to V_j$, and since $V_{j+1} \subset V_j$, it follows that $\widetilde{T}_j : V_j \to V_j$. Thus, we rewrite (3.38) as

$$K = T_{j+n} + \sum_{k=0}^{n-1} \widetilde{T}_{j+k} + (K - T_j) \,. \tag{3.40}$$

Now, in order to find an explicit representation for $\widetilde{T}_j$, we proceed as follows. Assume that we have constructed T_j and T_{j+1}. Then we have

$$T_j(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m-n} \phi^j_m(x) \phi^j_n(y)\,, \tag{3.41}$$

and

$$T_{j+1}(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^{j+1}_{m-n} \phi^{j+1}_m(x) \phi^{j+1}_n(y)\,. \tag{3.42}$$

In order to express T_{j+1} in terms of the basis functions $\phi^j_n\,, n \in \mathbb{Z}$ in V_j, we use the two-scale difference equation (2.8) satisfied by the scaling function $\phi(x)$, which takes the general form

$$\phi^{j+1}_k(x) = \sum_l h_{l-2k} \phi^j_l(x)\,. \tag{3.43}$$

Using (3.43) in (3.42), we have

$$\begin{aligned} T_{j+1}(x,y) &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^{j+1}_{m-n} \sum_k h_{k-2m} \phi^j_k(x) \sum_l h_{l-2n} \phi^j_l(y) \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi^j_k(x) \phi^j_l(y) \sum_m \sum_n t^{j+1}_{m-n} h_{k-2m} h_{l-2n}\,. \end{aligned}$$

Subtracting this expression from (3.41), we have

$$\begin{aligned} \widetilde{T}_j(x,y) &= T_j(x,y) - T_{j+1}(x,y) \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \tilde{t}^j_{k,l} \phi^j_k(x) \phi^j_l(y)\,, \end{aligned} \tag{3.44}$$

where we have set

$$\tilde{t}^j_{k,l} = t^j_{k-l} - \sum_m \sum_n t^{j+1}_{m-n} h_{k-2m} h_{l-2n}\,. \tag{3.45}$$

Note that the coefficient $\tilde{t}^j_{k,l}$ does not depend only on the difference of indices. However, it is not difficult to show that the coefficient matrix $\{\tilde{t}^j_{k,l}\}$ is a Toeplitz in (2×2) blocks. In addition, this matrix is banded, since the difference on the right-hand side of (3.45) decays quickly as the distance $|k-l|$ from the main diagonal increases. Due to their banded structure, it seems best to apply the matrices $\widetilde{T}_j$ directly.

When the multiple scale splitting (3.40) is used, it is also necessary to project the particle distribution $\sum f_n \delta(x - x_n)$ onto each of the subspaces

$V_j, \ldots, V_{j+n}$. This is done as follows. Recall that (see equation (3.9)) the coefficients $\{s_l^j\}$ that represent the particle distribution on V_j are given by

$$s_l^j = \sum_{n=1}^{N} f_n \phi_l^j(x_n)\,. \tag{3.46}$$

Now use (3.43) to write

$$\begin{aligned} s_k^{j+1} &= \sum_{n=1}^{N} f_n \phi_k^{j+1}(x_n) \\ &= \sum_{n=1}^{N} f_n \sum_l h_{l-2k} \phi_l^j(x_n) \\ &= \sum_l h_{l-2k} \sum_{n=1}^{N} f_n \phi_l^j(x_n) \\ &= \sum_l h_{l-2k} s_l^j\,. \end{aligned} \tag{3.47}$$

Hence, the coefficients $\{s_k^{j+1}\}$ may be computed directly from the coefficients $\{s_l^j\}$ on the next finer scale. We thus use (3.46) to obtain the projection onto the finest subspace V_j, then use the decomposition formula (3.47) to obtain the projection onto the coarser subspaces $V_{j+1}, \ldots, V_{j+n}$.

3.5 Algorithm for Nonsingular Kernels

We wish to point out that if the kernel $K(x-y)$ is globally bounded, then the low-frequency approximation alone provides an efficient algorithm for its application. As examples, consider

$$K(x-y) = \sin[\pi(x-y)]\,, \qquad K(x-y) = \frac{1}{1+(x-y)^2}\,.$$

For such kernels, the estimate $|K(x-y) - T_j(x,y)| < \epsilon$ holds for all (x,y), and there is no need to apply the high-frequency step. In this case, the approach outlined in Section 3.2.1 provides an algorithm of complexity $O(J \log_2 J)$, where J is the number of grid points. Here J depends only on the desired accuracy ϵ, and is independent of the number of particles N.

Chapter 4

The Fast Summation Algorithm in Higher Dimensions

In this chapter, we explain how to extend our one-dimensional scheme to higher dimensions. The algorithm has been implemented in two dimensions, and we give details of this and also present some numerical results. As mentioned above, there is little here beyond the use of vector notation that is not completely analogous to the one-dimensional case, except for singular value decomposition of the coefficient matrix (this matrix is defined by equations (4.2) and (4.3) below).

4.1 General Description

1. Our goal is to compute the numbers $\{g_m\}$, where

$$g_m = \sum_{\substack{n=1 \\ n \neq m}}^{N} K(x_m - x_n) f_n \,, \qquad 1 \leq m \leq N \,.$$

Here we assume that the particle locations $\{x_n\}$ are points in $\mathbf{R}^d, d = 1, 2, \ldots$. We assume that the line $x - y = 0$ is a singularity of the kernel $K(x-y)$, i.e.

$$|K(x-y)| \to \infty \quad \text{as} \quad |x-y| \to 0 \,,$$

but that the partial derivatives $\partial^\alpha K$, for $|\alpha| \leq M$, where $\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} + \cdots + \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d}$ and $|\alpha| = \alpha_1 + \cdots + \alpha_d$, exist and are continuous on any open set not containing this line.

We assume that the kernel function K assigns a single real number to each point $x \in \mathbf{R}^d, x \neq 0$. For vector-valued kernels, the algorithm must be applied separately to each component.

We assume moreover the existence of a "bandwidth" B, such that all partial derivatives $\partial^\alpha K$, for $|\alpha| = M$, are uniformly bounded outside the band $|x-y| < B$, i.e.

$$|\partial^\alpha K(x-y)| \le C \quad \text{for} \quad |x-y| \ge B\,,$$

for all multi-indices α such that $|\alpha| = M$.

2. We choose a level of refinement ($j \le 0$), an MRA with M vanishing moments, and construct the projection of the kernel onto the subspace V_j, given by

$$T_j(x,y) = \sum_{k\in\mathbf{Z}^d}\sum_{l\in\mathbf{Z}^d} t^j_{k-l}\phi^j_k(x)\phi^j_l(y)\,.$$

As discussed in Section 2.1.1, the scaling function here is a tensor product of a one-dimensional scaling function. For example,

$$\phi^j_k(x) = 2^{-jd/2}\phi(2^{-j}x_1 - k_1)\cdots\phi(2^{-j}x_d - k_d)\,, \qquad k = (k_1,\ldots,k_d) \in \mathbf{Z}^d$$

where $x = (x_1,\ldots,x_d) \in \mathbf{R}^d$. Construction of T_j requires only the computation of the coefficients $\{t^j_n\}$, given by

$$t^j_n = \int_{\mathbf{R}^d} K(x_1,\ldots,x_d)\Phi(2^{-j}x_1 - n_1)\cdots\Phi(2^{-j}x_d - n_d)\,dx\,, \quad n \in \mathbf{Z}^d$$

where $dx = dx_1\cdots dx_d$, and $\Phi(x)$ is the autocorrelation of the one-dimensional scaling function $\phi(x)$.

3. As in the one-dimensional case, we have an estimate of the form

$$|K(x-y) - T_j(x,y)| \le \epsilon \quad \text{for } |x-y| \ge 2^j B\,,$$

for any $\epsilon > 0$, for all scales such that the scale parameter $j \le 0$ is sufficiently small.

4. The low-frequency approximation is analogous to the one-dimensional case, requiring the application of a $2d$-dimensional "matrix" to a d-dimensional "vector". The matrix is Toeplitz in each index, and the multiplication can be accomplished using the FFT.

5. The high-frequency correction step is also analogous to the one-dimensional case. The only difference is that we use a singular value decomposition of the coefficient matrix $\{t^j_n, n \in \mathbf{Z}^d\}$, in order to express the kernel $T_j(x,y)$ as a product of one-dimensional kernels, which greatly facilitates the pointwise evaluation. When $d = 2$, this is straightforward, and will be discussed in greater detail below. For $d \ge 3$, it is still an open question whether or not such a decomposition is feasible.

4.2 The Fast Summation Algorithm in Two Dimensions

1. Our goal is to compute

$$g(x_m, y_m) = \sum_{\substack{n=1 \\ n \neq m}}^{N} K(x_m - x_n, y_m - y_n) f(x_n, y_n)\,, \qquad 1 \leq m \leq N\,. \tag{4.1}$$

In order to approximate the kernel $K(x - x', y - y')$, we construct the kernel

$$T_j(x, y, x', y') = \sum_{k,k'} \sum_{l,l'} t^j_{k-k',l-l'} \phi^j_k(x) \phi^j_l(y) \phi^j_{k'}(x') \phi^j_{l'}(y')\,. \tag{4.2}$$

The algorithm is currently implemented using central B-splines as the scaling function. For this choice, in two dimensions, the coefficients will be computed according to the formula

$$t^j_{k,l} = \sum_{m=0}^{2M-1} \frac{2^{mj}}{m!} \sum_{n=0}^{m} \binom{m}{n} \partial_x^{m-n} \partial_y^n K(2^j k, 2^j l) Q_{m-n} Q_n\,. \tag{4.3}$$

Formula (4.3) is the two-dimensional analogue of (2.100).

2. The first step is to compute the low-frequency approximations

$$g^j_m = \sum_{n=1}^{N} T_j(x_m, y_m, x_n, y_n) f_n\,, \qquad 1 \leq m \leq N \tag{4.4}$$

for $m = 1, \ldots, N$, and where $f_n = f(x_n, y_n)$. This can be written as

$$g^j_m = \sum_{k,l} \hat{s}^j_{k,l} \phi^j_k(x_m) \phi^j_l(y_m)\,, \tag{4.5}$$

where

$$\hat{s}^j_{k,l} = \sum_{k',l'} t^j_{k-k',l-l'}\, s^j_{k',l'} \tag{4.6}$$

and

$$s^j_{k',l'} = \sum_{n=1}^{N} f_n \phi^j_{k'}(x_n) \phi^j_{l'}(y_n)\,. \tag{4.7}$$

Equation (4.6) indicates the operation of applying a four-dimensional matrix to a two-dimensional vector, but as the matrix is a Toeplitz this can be accomplished using the (two-dimensional) fast Fourier transform.

3. The high-frequency correction step requires that we evaluate the kernel $T_j(x, y, x', y')$ at specific locations (x, y, x', y'), wherever $|(x, y) - (x', y')| < B$. To do this, we use the trigonometric expansion of the kernel obtained Section

2.3.3. Since we are using central B-splines as basis functions, the functions $\hat{\Phi}_n(\xi)$ are real. In this case, we can write

$$\begin{aligned}
& 4^j\, T_j(2^j x, 2^j y, 2^j x', 2^j y') \\
= \;& I_{0,0}(z,z') + 2\sum_{n=1}^{\infty} \cos(n\pi w') I_{0,n}(z,z') + 2\sum_{m=1}^{\infty} \cos(m\pi w) I_{m,0}(z,z') \\
+ \;& 4\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \cos(m\pi w)\cos(n\pi w') I_{m,n}(z,z') \qquad (4.8)
\end{aligned}$$

where

$$I_{m,n}(z,z') = \frac{1}{4\pi^2}\iint e^{-iz\xi - iz'\eta}\hat{t}^j(\xi - m\pi, \eta - n\pi)\hat{\Phi}_m(\xi)\hat{\Phi}_n(\eta)\, d\xi\, d\eta \qquad (4.9)$$

and

$$\begin{aligned}
z &= x - x', \qquad z' = y - y' \\
w &= x + x', \qquad w' = y + y'.
\end{aligned}$$

4. In two dimensions, the trigonometric series $\hat{t}^j$ is given by

$$\hat{t}^j(\xi,\eta) = \sum_{k,l} t^j_{k,l} e^{i(k\xi + l\eta)}. \qquad (4.10)$$

Now suppose that the singular value decomposition of the matrix $\{t^j_{k,l}\}$ is known, such that

$$t^j_{k,l} = \sum_{r=1}^{R} \sigma_r u^{(r)}_k v^{(r)}_l \qquad (4.11)$$

where $u^{(r)}_k$ and $v^{(r)}_l$ denote elements of the left and right singular vectors, respectively. The parameter R is the numerical rank. This is determined by truncating the SVD expansion when we have achieved the desired accuracy (as is implied by this remark, the two sides of equation (4.11) may agree only to within the specified accuracy ϵ).

Using (4.11) and (4.10) in (4.9), we have

$$\begin{aligned}
I_{m,n}(z,z') &= \frac{1}{4\pi^2}\sum_{k,l} t^j_{k,l}\iint e^{-i[\xi(z-k) + \eta(z'-l)]}\hat{\Phi}_m(\xi)\hat{\Phi}_n(\eta)\, d\xi\, d\eta \\
&= \sum_{k,l} t^j_{k,l}\Phi_m(z-k)\Phi_n(z'-l) \\
&= \sum_{r=1}^{R}\sigma_r \sum_k u^{(r)}_k \Phi_m(z-k)\sum_l v^{(r)}_l \Phi_n(z'-l) \qquad (4.12)
\end{aligned}$$

Defining

$$\begin{aligned} U_m^{(r)}(z) &= \sum_k u_k^{(r)} \Phi_m(z-k) \\ V_n^{(r)}(z') &= \sum_l v_l^{(r)} \Phi_n(z'-l)\,, \end{aligned}$$

we can express $I_{m,n}$ in the form

$$I_{m,n}(z,z') = \sum_{r=1}^{R} \sigma_r U_m^{(r)}(z) V_n^{(r)}(z')\,. \tag{4.13}$$

5. Substituting (4.13) into (4.8), we obtain our final expression for the trigonometric expansion of the kernel,

$$\begin{aligned} 4^j\, T_j(2^j x, 2^j y, 2^j x', 2^j y') &= \sum_{r=1}^{R} \sigma_r \left\{ U_0^{(r)}(z) + 2\sum_{m=1}^{\infty} \cos(m\pi w) U_m^{(r)}(z) \right\} \\ &\times \left\{ V_0^{(r)}(z') + 2\sum_{n=1}^{\infty} \cos(n\pi w') V_n^{(r)}(z') \right\}. \end{aligned} \tag{4.14}$$

Once the functions $U_m^{(r)}(z)$ and $V_n^{(r)}(z')$ have been tabulated, the cost of evaluating the two-dimensional kernel (4.14) is $(2R)$ times the cost of evaluating a one-dimensional kernel. In practice, the summations over m and n are also truncated, and we retain terms only for $m \leq 3$ and $n \leq 3$.

4.2.1 An Example in Two Dimensions

As our model problem in two dimensions we choose the kernel

$$K(x,y) = \frac{1}{x^2+y^2}\,, \quad x^2+y^2 \neq 0\,.$$

The results are shown in the table below, where N is the number of particles, which were distributed randomly in the unit square $[0,1]\times[0,1]$. The columns marked T_{LF} and T_{HF} give the times for the low-frequency approximation and high-frequency correction, respectively, the column marked T_{tot} gives the total time for the algorithm, and the column marked T_{dir} gives the time required to compute the exact expressions

$$g_m = \sum_{\substack{n=1 \\ n\neq m}}^{N} \frac{f_n}{(x_m-x_n)^2+(y_m-y_n)^2}\,,$$

for $m = 1, \ldots, N$.

The errors are computed according to the following formulae. Suppose that $\mathbf{x}$ is the "exact" N-length vector, and $\tilde{\mathbf{x}}$ is the approximation. Then we compute the relative errors

$$E = \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|},$$

where

$$\|\mathbf{x}\|_2^2 = \frac{1}{N}\sum_{i=1}^{N} |x_i|^2, \qquad \|\mathbf{x}\|_\infty = \max_{1 \le i \le N} |x_i|.$$

The timings were done on 10/30/96 using a Sun Sparc-20 workstation.

N	T_{LF}	T_{HF}	T_{tot}	T_{dir}	E_2	E_∞
64	0.0920	0.0435	0.1355	0.0124	0.27021E-04	0.11911E-04
128	0.1047	0.1281	0.2328	0.0423	0.29654E-04	0.15733E-04
256	0.3559	0.1442	0.5001	0.1621	0.31086E-04	0.19491E-04
512	0.3769	0.4730	0.8499	0.6057	0.71012E-05	0.16866E-05
1024	1.4692	0.5130	1.9822	2.4040	0.99095E-06	0.14316E-06
2048	1.5518	1.8517	3.4034	9.4216	0.36296E-07	0.44216E-08
4096	5.4488	2.1030	7.5518	38.0664	0.25482E-05	0.33993E-06
8192	6.3597	7.7037	14.0633	150.1953	0.20096E-05	0.27010E-06

Table 4.1: Implementation in two dimensions using B-splines of degree 3, break-even at about 750 particles.

Chapter 5

Regularization of Singular Operators

One often encounters discrete sums that do not have a direct analogue as an integral operator. For example, the expression

$$(Kf)(x) = \int_{-\infty}^{\infty} \frac{f(y)}{|x-y|^2}\, dy \tag{5.1}$$

requires special interpretation as an operator. By contrast, the discrete sum

$$g(x_m) = \sum_{\substack{n=1 \\ n \neq m}}^{N} \frac{f(x_n)}{|x_m - x_n|^2}\,, \quad 1 \leq m \leq N \tag{5.2}$$

presents no difficulties. In our approach to the computation of such sums, we construct kernels $T_j : V_j \to V_j\,, j \in \mathbb{Z}$ which are approximations of the kernel functions that appear in the summation problem. However, it is clear that the same approximation T_j may be used to provide a definition of the corresponding integral operator.

Thus, for example, if we have constructed a kernel T_j such that the sum

$$g^j(x_m) = \sum_{\substack{n=1 \\ n \neq m}}^{N} T_j(x_m, x_n) f(x_n)\,, \quad 1 \leq m \leq N \tag{5.3}$$

is an approximation to (5.2), then the same kernel in the expression

$$(T_j f)(x) = \int_{-\infty}^{\infty} T_j(x,y) f(y)\, dy \tag{5.4}$$

provides a meaning to the integral operator (5.1), along with a practical computational algorithm.

We may therefore view the expression (5.4) as a regularization of the operator (5.1). In this chapter, we investigate the construction of regularizations for a certain class of linear operators.

5.1 Preliminary Considerations

In order to construct a kernel T_j that represents a linear operator K on the subspace V_j of a multiresolution analysis, it is necessary only to compute the coefficients $\{t^j_{m,n}\}$ that appear in the expansion

$$T_j(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m,n} \phi^j_m(x) \phi^j_n(y) \, . \tag{5.5}$$

In general, these coefficients are computed according to the formula

$$t^j_{m,n} = (\phi^j_m \, , K\phi^j_n) = \int_{-\infty}^{\infty} \phi^j_m(x)(K\phi^j_n)(x)\, dx \, . \tag{5.6}$$

Let us assume that

$$(\phi^j_m \, , K\phi^j_n) = (\phi^j_{m-n} \, , K\phi^j_0) \, , \tag{5.7}$$

which implies that the coefficient depends only on the difference of the two indices,

$$t^j_{m,n} = t^j_{m-n} \, . \tag{5.8}$$

Then the kernel (5.5) has the form

$$T_j(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t^j_{m-n} \phi^j_m(x) \phi^j_n(y) \, , \tag{5.9}$$

where

$$t^j_n = \int_{-\infty}^{\infty} \phi^j_n(x)(K\phi^j_0)(x)\, dx \, . \tag{5.10}$$

So long as (5.10) is well-defined for all $n \in \mathbb{Z}$, there is no difficulty in constructing the kernel (5.9). However, for many operators K that we wish to consider, coefficients t^j_n as defined by (5.10) do not exist for all integers n.

This motivates the following approach. To construct the kernel (5.9), we must produce the appropriate sequence $t^j_n \, , n \in \mathbb{Z}$. Let t^j_n be given by (5.10), whenever this expression is well-defined. For those integers n for which (5.10) fails to exist, we will propose an alternate method for assigning a value to the functional $(\phi^j_n \, , K\phi^j_0) = t^j_n$.

We now give a precise definition to the class of linear operators that will be considered in this chapter. We assume that K commutes with the operation of translation, that is, if τ_a is the translation operator defined by $(\tau_a f)(x) = f(x-a)$, then K satisfies $\tau_a K = K \tau_a$. This property implies (5.7). We consider operators that are homogeneous of some degree, and we have

$$t^j_n = 2^{-\alpha j}(\phi^0_n \, , K\phi^0_0) = 2^{-\alpha j}\left(\phi(x-n) \, , (K\phi)(x)\right) = 2^{-\alpha j} t^0_n \tag{5.11}$$

when the operator K is homogeneous of degree α.

For operators satisfying these conditions, the projections of the corresponding kernels $T_j : V_j \to V_j$ have the form

$$T_j(x,y) = 2^{-j(1+\alpha)} T(2^{-j}x, 2^{-j}y)\,, \tag{5.12}$$

for any $j \in \mathbb{Z}$, where

$$T(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t_{m-n} \phi(x-m)\phi(y-n)\,, \tag{5.13}$$

and the coefficients are given by the functional

$$t_n = (\phi_n^0\,, K\phi_0^0)\,, \tag{5.14}$$

where

$$(\phi_n^0\,, K\phi_0^0) = \int_{-\infty}^{\infty} \phi(x-n)(K\phi)(x)\,dx \tag{5.15}$$

whenever the right-hand side of (5.15) exists.

As examples, we consider below the following integral operators with algebraic singularities

$$(Kf)(x) = \int_{-\infty}^{\infty} \frac{f(y)}{(x-y)^{1+\alpha}}\,dy\,, \quad \alpha \geq -1 \tag{5.16}$$

and

$$(Kf)(x) = \int_{-\infty}^{\infty} \frac{f(y)}{|x-y|^{1+\alpha}}\,dy\,, \quad \alpha \geq -1 \tag{5.17}$$

For $\alpha = 0$, the expression (5.15) defines a bounded operator on $L^2(\mathbf{R})$, provided that we consider the principle value at $x = y$. When suitably scaled, this example is known as the Hilbert transform. However, for most values of α, the operators (5.16) and (5.17) are not bounded on $L^2(\mathbf{R})$. It has been shown in [2] that the derivative operators $(d/dx)^\alpha$, for $\alpha = 1, 2, \ldots$ also satisfy the conditions stated above, and have representations of the form (5.12).

5.2 Classical Regularization of Divergent Integrals

Our task is to give a meaning to the functional (5.14) when the integral (5.15) is divergent. To this end, we first recall how such problems have been addressed up to now.

As an example, consider the expression

$$\int_{-\infty}^{\infty} \frac{\phi(x)}{x}\,dx\,, \tag{5.18}$$

where ϕ is an infinitely differentiable function with compact support, $\phi \in C_0^\infty(\mathbf{R})$. This integral converges for all test functions ϕ that vanish in a neighborhood of the origin. However, the integral diverges if $\phi(0) \neq 0$. We now ask whether it is possible to define a functional (x^{-1}, ϕ) such that, for all test functions ϕ that vanish in a neighborhood of zero, the functional has the value given by (5.18). The functional (x^{-1}, ϕ) is then called a regularization of the divergent integral (5.18) (see e.g. [13, pp.10-12]).

The functional

$$(x^{-1}, \phi) = \int_{-\infty}^{-\epsilon} \frac{\phi(x)}{x}\, dx + \int_{-\epsilon}^{\epsilon} \frac{\phi(x) - \phi(0)}{x}\, dx + \int_{\epsilon}^{\infty} \frac{\phi(x)}{x}\, dx\,, \tag{5.19}$$

where ϵ is any positive real number, obviously reduces to (5.18) if $\phi(0) = 0$. Moreover, since $\phi(x) - \phi(0) \sim x\phi'(0)$ as $x \to 0$, it follows that (5.19) is well-defined for all $\phi \in C_0^\infty(\mathbf{R})$, and we may therefore consider this functional to be a regularization of (5.18).

More generally, consider the following integrals,

$$\int_{-\infty}^{\infty} \frac{\phi(x)}{x^{2m}}\, dx = \int_0^\infty \frac{\phi(x) + \phi(-x)}{x^{2m}}\, dx\,, \tag{5.20}$$

and

$$\int_{-\infty}^{\infty} \frac{\phi(x)}{x^{2m+1}}\, dx = \int_0^\infty \frac{\phi(x) - \phi(-x)}{x^{2m+1}}\, dx\,, \tag{5.21}$$

where $m = 0, 1, 2, \ldots$. In order to assign a meaning to the right-hand sides of (5.20) and (5.21) for test functions ϕ that do not vanish at $x = 0$, the main idea (see e.g. [13, pp.45-82]) is to replace the numerator with a function which has enough zeroes at the origin to insure convergence of the integral. This is accomplished by subtracting enough terms of the Taylor series expansion of the numerator about $x = 0$ to leave a remainder of order greater than or equal to that of the demoninator. (However, in so doing, we must also take care not to destroy the convergence at ∞.) Hence, we replace (5.20) with the functional

$$(x^{-2m}, \phi) = \int_0^\infty \frac{dx}{x^{2m}} \left[\phi(x) + \phi(-x) - 2\sum_{k=0}^{m-1} \frac{\phi^{(2k)}(0)}{(2k)!} x^{2k} \right]\,, \tag{5.22}$$

and we replace (5.21) with the functional

$$(x^{-2m-1}, \phi) = \int_0^\infty \frac{dx}{x^{2m+1}} \left[\phi(x) - \phi(-x) - 2\sum_{k=0}^{m-1} \frac{\phi^{(2k+1)}(0)}{(2k+1)!} x^{2k+1} \right]\,. \tag{5.23}$$

Although the approach outlined above is a standard method for dealing with the regularization of divergent integrals with algebraic singularities, it does not address the issue of numerical computation of such integrals.

In the following section, we present a method for regularization of the integral (5.15), thus providing a meaning to the functional (5.14), which utilizes the multiresolution approach. For integral operators with algebraic singularities, our approach produces the same results as the classical regularization method, with the additional benefit of a practical scheme for computation. Furthermore, it appears that our approach may also be applied to classes of operators with more general types of singularities.

5.3 A Multiresolution Approach to Regularization

5.3.1 Choice of the Scaling Function

Throughout this chapter, we limit our consideration to a specific family of MRAs, namely, those belonging to the compactly supported scaling functions constructed by I. Daubechies [7]. These scaling functions and their autocorrelations were introduced in Section 2.1.4.

We recall that for the MRA with M vanishing moments, the autocorrelation $\Phi(x)$ is supported on the interval $[1-2M, 2M-1]$, and also that the coefficients $\{a_m\}$ in the two-scale difference equation (2.28) for $\Phi(x)$ are given explicitly in (2.29).

5.3.2 Two-Scale Difference Equation for the Coefficients

In this section we derive the following necessary condition on the functional (5.14),

$$2^{-\alpha} t_n = t_{2n} + \frac{1}{2} \sum_{m=1}^{M} a_{2m-1} \left[t_{2n-2m+1} + t_{2n-1+2m} \right] . \qquad (5.24)$$

We refer to (5.24) as the two-scale difference equation. This equation is the tool that we will exploit for the purpose of defining the coefficient t_n when the integral (5.15) is divergent. If a sequence $t_n\,, n \in \mathbb{Z}$ can be found that satisfies (5.24), and agrees with the expression (5.15) whenever it exists, then the corresponding kernel (5.12) is our regularization of the operator K on the subspace $V_j\,, j \in \mathbb{Z}$.

Let K be a linear, homogeneous operator that commutes with translation. Consider the expression for the coefficient

$$t_n = (\phi(x-n), (K\phi)(x))\,, \qquad (5.25)$$

obtained by setting $j = 0$ in (5.11). The scaling function $\phi(x)$ satisfies the two-scale difference equation

$$\phi(x) = \sqrt{2} \sum_{k=0}^{2M-1} h_k \phi(2x - k)\,,$$

which we rewrite as

$$\phi_n^{(0)}(x) = \sum_{k=0}^{2M-1} h_k \phi_{2n+k}^{(-1)}(x)\,.$$

(Note that the superscript indicates the scale in the MRA.) Substituting this expression into (5.25) and using the linearity of K, we have

$$t_n = \left(\sum_{k=0}^{2M-1} h_k \phi_{2n+k}^{(-1)}\,, \sum_{l=0}^{2M-1} h_l (K\phi_l^{(-1)}) \right)\,.$$

Using (5.8) this becomes

$$t_n = \sum_{k=0}^{2M-1} \sum_{l=0}^{2M-1} h_k h_l \left(\phi_{2n+k-l}^{(-1)}\,, K\phi_0^{(-1)} \right)\,.$$

Finally, using (5.11) we obtain

$$\begin{aligned} t_n &= 2^\alpha \sum_{k=0}^{2M-1} \sum_{l=0}^{2M-1} h_k h_l \left(\phi_{2n+k-l}^{(0)}\,, K\phi_0^{(0)} \right) \\ &= 2^\alpha \sum_{k=0}^{2M-1} \sum_{l=0}^{2M-1} h_k h_l \, t_{2n+k-l} \\ &= 2^\alpha \sum_{1-2M}^{2M-1} \frac{1}{2} a_m \, t_{2n+m}\,, \end{aligned}$$

since

$$a_m = 2 \sum_{l=0}^{2M-1-m} h_l h_{m+l}\,.$$

Noting that $a_{2m} = 0$ for $m \neq 0$, $a_0 = 2$, and $a_{1-2m} = a_{2m-1}$, we arrive at equation (5.24). The coefficients $\{a_{2m-1}\}$ are given by equation (2.29).

5.3.3 Multiresolution Definition of Regularized Operators

In this section we describe an algorithm which utilizes (5.24) to compute the coefficients $t_n, n \in \mathbb{Z}$ for a linear, homogeneous operator K that commutes with translation. This algorithm may be considered as a multiresolution definition for the regularization of such operators. We show its consistency with the classical definition (see Section 5.2) in Section 5.3.4.

Step 1. Assume that we have established the asymptotic condition

$$t_n = F(n) + O\left(\frac{1}{n^{2M}}\right), \qquad \text{as } |n| \to \infty \tag{5.26}$$

with some function F. (We have shown in Section 5.3.5 how to find F when K is an integral operator.) Using (5.26), we determine all coefficients t_n for large $|n|$ as follows. Given $\epsilon > 0$, we choose a positive integer n_0, sufficiently large to insure that $|t_n - F(n)| < \epsilon$ whenever $|n|$ is greater than n_0. Then we set

$$t_n = F(n), \quad \text{for } |n| > n_0 . \tag{5.27}$$

Step 2. Next we use (5.24) to compute $\{t_n\}$ for $2M \le |n| \le n_0$. This is done in reverse order, beginning with $n = \pm n_0$ and ending with $n = \pm 2M$. This is made possible by the fact that for $|n| \ge 2M$, the right-hand side of (5.24) does not contain the coefficient t_n. For $|n| \ge 2M$ equation (5.24) is an expression for t_n in terms of $t_{2n+1-2M}, \ldots, t_{2n-1}, t_{2n}, t_{2n+1}, \ldots, t_{2n-1+2M}$.

Step 3. Finally, we solve the linear system defined by (5.24) to obtain the coefficients for $|n| \le 2M - 1$. Coefficients in this range appear on both sides of the two-scale difference equation (5.24). This system can be written in matrix notation as

$$\lambda\tau = A\tau + \mathbf{b}, \tag{5.28}$$

where we have set $\lambda = 2^{-\alpha}$, and $\tau = \{t_{1-2M}, \ldots, t_{2M-1}\}$. In (5.28), A is a square matrix of dimension $(4M - 1)$, and $\mathbf{b}$ is a vector containing the information obtained from the asymptotic condition (5.26). The entries of $\mathbf{b}$ are combinations of the known coefficients t_n in the range $|n| \ge 2M$.

The solution of (5.28) may be written as

$$\tau = (\lambda I - A)^{-1}\mathbf{b}, \tag{5.29}$$

provided that $\lambda = 2^{-\alpha}$ is not an eigenvalue of the matrix A.

If λ is an eigenvalue of A, we proceed as follows. Let V_λ be the invariant subspace belonging to the eigenvalue λ, and $V_\lambda^\perp$ the orthogonal complement. Suppose that the dimension of V_λ is r. (Recall that A is a square matrix of dimension $4M - 1$.) Let S be a matrix with $(4M - 1 - r)$ columns that are an orthonormal basis for $V_\lambda^\perp$. If the multiplicity of λ is also r, then the system (5.28) has the solution

$$\tau = S(\lambda I - S^T AS)^{-1} S^T \mathbf{b}, \tag{5.30}$$

if and only if $\mathbf{b}$ is orthogonal to every eigenvector of A^T belonging to λ, and moreover the vector τ in (5.30) is the unique solution to (5.28) that satisfies $SS^T\tau = \tau$, i.e. $\tau \in V_\lambda^\perp$. This result is proved in the Appendix A. $\square$

We are now in a position to give a constructive definition of the multiresolution regularization of the operator K.

Definition 5.3.1 *If the linear system (5.28) has a solution, then we obtain the sequence* $t_n\,, n \in \mathbb{Z}$ *by combining the solution* τ *with the coefficients* $t_n\,, |n| \geq 2M$ *computed in Steps 1 and 2. We define the kernel (5.12) corresponding to this sequence to be the regularization of the operator* K *on the subspace* V_j*.*

Example 1: For $M = 1$, let us write out (5.28) explicitly.

$$\lambda \begin{bmatrix} t_{-1} \\ t_0 \\ t_1 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 0 \\ 1/2 & 1 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} t_{-1} \\ t_0 \\ t_1 \end{bmatrix} + \begin{bmatrix} (1/2)t_{-3} + t_{-2} \\ 0 \\ t_2 + (1/2)t_3 \end{bmatrix},$$

where t_{-3}, t_{-2}, t_2, t_3 can be determined with any desired accuracy ϵ from the asymptotic condition (5.26). The eigenvalues for this matrix are $\{1, 1/2, 1/2\}$, corresponding to $\alpha = 0$ or 1.

Example 2: For $M = 2$, let us write out (5.28) explicitly.

$$\lambda \begin{bmatrix} t_{-3} \\ t_{-2} \\ t_{-1} \\ t_0 \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} -1/16 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9/16 & 0 & -1/16 & 0 & 0 & 0 & 0 \\ 9/16 & 1 & 9/16 & 0 & -1/16 & 0 & 0 \\ -1/16 & 0 & 9/16 & 1 & 9/16 & 0 & -1/16 \\ 0 & 0 & -1/16 & 0 & 9/16 & 1 & 9/16 \\ 0 & 0 & 0 & 0 & -1/16 & 0 & 9/16 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1/16 \end{bmatrix} \begin{bmatrix} t_{-3} \\ t_{-2} \\ t_{-1} \\ t_0 \\ t_1 \\ t_2 \\ t_3 \end{bmatrix}$$
$$+ \begin{bmatrix} t_{-6} + \frac{9}{16}(t_{-7} + t_{-5}) - \frac{1}{16}t_{-9} \\ t_{-4} + \frac{9}{16}t_{-5} - \frac{1}{16}t_{-7} \\ -\frac{1}{16}t_{-5} \\ 0 \\ -\frac{1}{16}t_5 \\ t_4 + \frac{9}{16}t_5 - \frac{1}{16}t_7 \\ t_6 + \frac{9}{16}(t_7 + t_5) - \frac{1}{16}t_9 \end{bmatrix},$$

where as before the coefficients t_n for $n = \pm 4, \pm 5, \pm 6, \pm 7, \pm 9$ can be determined from the asymptotic condition (5.26). The eigenvalues for this matrix are $\{1, 1/2, 1/4, 1/4, 1/8, -1/16, -1/16\}$, the first four of which correspond to $\alpha = 0, 1, 2,$ or 3.

Example 3: For the system of Example(1), consider $\lambda = 1$. For the matrix

$$A = \begin{bmatrix} 1/2 & 0 & 0 \\ 1/2 & 1 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix},$$

it is readly verified that an eigenvector is $[0,1,0]^T$, and that the projector onto the subspace $V_\lambda^\perp$ is the matrix

$$SS^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Computing the solution (5.30) we have

$$\begin{bmatrix} t_{-1} \\ t_0 \\ t_1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} (1/2)t_{-3} + t_{-2} \\ 0 \\ t_2 + (1/2)t_3 \end{bmatrix} = \begin{bmatrix} t_{-3} + 2t_{-2} \\ 0 \\ 2t_2 + t_3 \end{bmatrix}.$$

Example 4: For the system of Example(1), consider $\lambda = 1/2$. This eigenvalue has multiplicity two, and the projector onto the subspace $V_\lambda^\perp$ is given by

$$SS^T = \left(\frac{1}{3}\right) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}.$$

Computing the solution (5.30) for this eigenvalue we have

$$\begin{bmatrix} t_{-1} \\ t_0 \\ t_1 \end{bmatrix} = \left(-\frac{2}{3}\right) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} (1/2)t_{-3} + t_{-2} \\ 0 \\ t_2 + (1/2)t_3 \end{bmatrix},$$

which is an equation for determining t_{-1}, t_0, t_1.

Example 5: For the system of Example(2), consider $\lambda = 1$. Analogous to Example(3), an eigenvector is $[0,0,0,1,0,0,0]^T$, and we have

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

The matrix $S(I - S^TAS)^{-1}S^T$ which appears in equation (5.30) may be computed explicitly, and we have $S(I - S^TAS)^{-1}S^T =$

$$\begin{bmatrix} 16/17 & 0 & 0 & 0 & 0 & 0 & 0 \\ 47/119 & 55/63 & -8/63 & 0 & 1/63 & 1/63 & 2/119 \\ 256/119 & 128/63 & 128/63 & 0 & -16/63 & -16/63 & -32/119 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -32/119 & -16/63 & -16/63 & 0 & 128/63 & 128/63 & 256/119 \\ 2/119 & 1/63 & 1/63 & 0 & -8/63 & 55/63 & 47/119 \\ 0 & 0 & 0 & 0 & 0 & 0 & 16/17 \end{bmatrix}.$$

5.3.4 Classical vs Multiresolution Regularization

Consider the integral operator

$$(Kf)(x) = \int_{-\infty}^{\infty} \frac{f(y)}{(x-y)^{1+\alpha}}\, dy\,,$$

where $\alpha = 0, 1, 2, \ldots$ as an example. Projection of this operator onto the subspace V_0 of the MRA with M vanishing moments (see Section 5.3.1) requires evaluation of the integral

$$t_n = \int_{-\infty}^{\infty} \phi(x-n)(K\phi)(x)\, dx\,,$$

which can also be written as

$$t_n = \int_{-\infty}^{\infty} \frac{\Phi(x-n)}{x^{1+\alpha}}\, dx\,, \tag{5.31}$$

where $\Phi(x)$ is the autocorrelation of $\phi(x)$ (see equation (2.56)). The support of $\Phi(x)$ is the closed interval $[1-2M, 2M-1]$. Hence the integral (5.31) is well defined whenever $|n| \geq 2M$, since for these values of n, the support of $\Phi(x-n)$ does not contain the origin.

We consider a regularization of the integral (5.31) for the coefficients $t_n, |n| < 2M$ that parallels the classical method as outlined in Section 5.2.

If $\alpha = 1, 3, \ldots$ so that $(1+\alpha)$ is an even, positive integer, then (5.31) may be written as

$$t_n = \int_{0}^{\infty} \frac{\Phi(x-n)+\Phi(x+n)}{x^{1+\alpha}}\, dx\,, \tag{5.32}$$

since $\Phi(-x) = \Phi(x)$. In accordance with the classical method, we subtract enough terms of the Taylor expansion about zero to leave a remainder of order $(1+\alpha)$ in the numerator. Thus, we replace (5.32) with the functional

$$\begin{aligned} t_n &= \left(x^{-1-\alpha}, \Phi(x-n)\right) \\ &= \int_0^{\infty} \frac{dx}{x^{1+\alpha}} \left[\Phi(x-n)+\Phi(x+n) - 2\sum_{k=0}^{(\alpha-1)/2} \frac{\Phi^{(2k)}(n)}{(2k)!} x^{2k}\right] \end{aligned} \tag{5.33}$$

The functional (5.33) is well-defined for all n, provided that $\Phi(x)$ has at least $(1+\alpha)$ derivatives.

A similar situation holds when $(1+\alpha)$ is odd. In this case (5.31) may be written as

$$t_n = \int_{0}^{\infty} \frac{\Phi(x-n)-\Phi(x+n)}{x^{1+\alpha}}\, dx\,, \tag{5.34}$$

and the regularization of this integral is the functional

$$
\begin{aligned}
t_n &= \left(x^{-1-\alpha}, \Phi(x-n)\right) \\
&= \int_0^\infty \frac{dx}{x^{1+\alpha}} \left[\Phi(x-n) - \Phi(x+n) - 2\sum_{k=0}^{(\alpha-2)/2} \frac{\Phi^{(2k+1)}(n)}{(2k+1)!} x^{2k+1}\right] \quad (5.35)
\end{aligned}
$$

This functional is well-defined for all n, if $\Phi(x)$ has at least $(1+\alpha)$ derivatives.

Lemma 5.3.1 *The classical regularization of (5.31), given by the functionals (5.33) or (5.35), satisfies the two-scale difference equation (5.24).*

Proof: From the identity

$$
\Phi(x) = \frac{1}{2}\sum_{1-2M}^{2M-1} a_m \Phi(2x-m)\,,
$$

(see equations 2.25 and 2.28) it follows that

$$
\begin{aligned}
\Phi(x \pm n) &= \frac{1}{2}\sum_{1-2M}^{2M-1} a_m \Phi(2x \pm 2n - m) \\
\Phi^{(k)}(n) &= \frac{1}{2}\sum_{1-2M}^{2M-1} a_m 2^k \Phi^{(k)}(2n-m)
\end{aligned}
$$

provided that the kth derivative exists.

Substituting these expressions into (5.33), we have

$$
\begin{aligned}
t_n &= \int_0^\infty \frac{dx}{x^{1+\alpha}} \frac{1}{2}\sum_m a_m[\Phi(2x-2n-m) - \Phi(2x+2n-m) \\
&\quad - 2\sum_{k=0}^{m-1} \frac{\Phi^{(2k)}(2n-m)}{(2k)!} 2^{2k} x^{2k}] \\
&= 2^\alpha \int_0^\infty \frac{dx}{x^{1+\alpha}} \frac{1}{2}\sum_m a_m[\Phi(x-2n-m) - \Phi(x+2n-m) \\
&\quad - 2\sum_{k=0}^{(\alpha-1)/2} \frac{\Phi^{(2k)}(2n-m)}{(2k)!} 2^{2k} \left(\frac{x}{2}\right)^{2k}] \\
2^{-\alpha} t_n &= \int_0^\infty \frac{dx}{x^{1+\alpha}} \frac{1}{2}\sum_m a_m[\Phi(x-2n+m) - \Phi(x+2n-m) \\
&\quad - 2\sum_{k=0}^{(\alpha-1)/2} \frac{\Phi^{(2k)}(2n-m)}{(2k)!} x^{2k}] \\
&\quad - \int_0^\infty \frac{dx}{x^{1+\alpha}} \frac{1}{2}\sum_m a_m \left\{\Phi(x-2n-m) - \Phi(x-2n+m)\right\}\,.
\end{aligned}
$$

Now, since $a_{-m} = a_m$, it follows that

$$\int_0^\infty \frac{dx}{x^{1+\alpha}} \frac{1}{2} \sum_m a_m \left\{ \Phi(x - 2n - m) - \Phi(x - 2n + m) \right\} = 0\,.$$

Therefore, we have

$$\begin{aligned} 2^{-\alpha} t_n &= \frac{1}{2} \sum_m a_m \int_0^\infty \frac{dx}{x^{1+\alpha}} [\Phi(x - 2n + m) - \Phi(x + 2n - m) \\ &\quad - 2 \sum_{k=0}^{(\alpha-1)/2} \frac{\Phi^{(2k)}(2n - m)}{(2k)!} x^{2k}] \\ &= \frac{1}{2} \sum_m a_m t_{2n-m} = \frac{1}{2} \sum_m a_m t_{2n+m}\,. \end{aligned}$$

As this result is valid for all $n \in \mathbb{Z}$, the two-scale difference equation (5.24) is satisfied. The procedure for (5.35) is identical. $\Box$

Proposition 5.3.1 *The classical regularization and the multiresolution regularization of (5.31) produce identical results.*

Proof: The coefficients t_n for $|n| \geq 2M$ computed in Steps 1 and 2 of the algorithm of Section 5.3 agree with the integral expression (5.31) to within the chosen accuracy ϵ. In addition, the coefficients produced by (5.33) or (5.35) satisfy the two-scale difference equation (5.24), as is shown by Lemma(5.3.1). Therefore, the two methods produce sequences $t_n, n \in \mathbb{Z}$ that are identical (to within ϵ) for $|n| \geq 2M$. However, for $|n| < 2M$ the solution of the two-scale difference equation, once the coefficients for $|n| \geq 2M$ have been specified, is unique only up to addition of an eigenvector of the matrix A in (5.28) belonging to the eigenvalue $2^{-\alpha}$. Thus we now consider the coefficients in this range.

Note that the vector $\mathbf{b}$ which appears in (5.28) is known only approximately, since the coefficients $t_n, |n| \geq 2M$ on which it depends are computed only to precision ϵ. Thus, $\mathbf{b} = \mathbf{b}(\epsilon)$, and because of this (5.28) is a perturbed system (the entries of A are known exactly). However, the two-scale difference equation (5.24), viewed as an infinite matrix, has a finite condition number (which is in fact low) that is independent of ϵ. Hence, as $\epsilon \to 0, \mathbf{b}(\epsilon) \to \mathbf{b}$ where $\mathbf{b}$ is the exact vector. It follows that the solution of the perturbed system $\tau(\epsilon) \to \tau$, where τ is the solution of the unperturbed system.

In the Appendix A, it is shown that the solution $\tau = \{t_n, |n| < 2M\}$ to (5.28) is orthogonal to the eigenvector of A, and that it is the unique solution which has this property. Thus, to show that the coefficients $t_n\,, |n| < 2M$ produced by the two methods are identical, it is sufficient to show that the vector $\tau = \{t_{1-2M}, \ldots, t_{2M-1}\}$ produced by (5.33) or (5.35) is also orthogonal to the eigenvector of A.

It is shown in [2] that eigenvectors of A belonging to the eigenvalues $2^{-\alpha}$ are composed of the values of the αth derivative of $\Phi(x)$, taken at the integers. That is, if

$$v = \{\Phi^{(\alpha)}(1-2M), \ldots, \Phi^{(\alpha)}(2M-1)\},$$

then $Av = 2^{-\alpha}v$. Furthermore, the αth derivative $\Phi^{(\alpha)}(x)$ exists and is continuous if and only if $2^{-\alpha}$ is an eigenvalue of A, and is nondegenerate.

Hence, to prove the proposition, it is necessary and sufficient to show that

$$\sum_{1-2M}^{2M-1} t_n \, \Phi^{(\alpha)}(n) = 0, \tag{5.36}$$

where the coefficients t_n are given by (5.33) or (5.35).

Now consider the following. Since $\Phi(-x) = \Phi(x)$, it follows that if α is even, then $\Phi^{(\alpha)}(-x) = \Phi^{(\alpha)}(x)$, and if α is odd, then $\Phi^{(\alpha)}(-x) = -\Phi^{(\alpha)}(x)$. Now, from (5.31) it follows that if α is even, then $t_{-n} = -t_n$, and if α is odd, then $t_{-n} = t_n$. Thus (5.36) is always satisfied. □

Remark: The fact that the classical regularization and the multiresolution regularization of (5.31) produce identical results proves that the limit of the regularized kernel T_j, as $j \to -\infty$, is independent of the basis chosen, at least in those cases in which both methods are applicable.

Example: Consider $\alpha = 1$ in (5.33). The regularization is

$$t_n = \int_0^\infty \frac{dx}{x^2} \{\Phi(x-n) + \Phi(x+n) - 2\Phi(n)\}.$$

Since $t_{-n} = t_n$, it is sufficient to consider only $n \geq 0$. Also, since $\Phi(0) = 1$, and $\Phi(n) = 0$ if $n \neq 0$, we have

$$\begin{aligned} t_0 &= 2\int_0^\infty \frac{\Phi(x)-1}{x^2}\, dx \\ t_n &= \int_0^\infty \frac{\Phi(x-n)+\Phi(x+n)}{x^2}\, dx, \quad n \geq 1. \end{aligned} \tag{5.37}$$

We note that, in (5.37), we must choose $\Phi(x)$ that belongs to an MRA with $M \geq 3$, since for $M = 1$ or 2 the function $\Phi(x)$ does not have a second derivative, and the integrals in (5.37) do not converge.

To convince the reader that the classical approach (i.e. equations (5.37)) produces the same results as the algorithm of Section 5.3.3, we have computed the coefficients t_n, for $0 \leq n \leq 5$ for the MRA with $M = 3$. We list the values in the table below, rounded to three digits. The column marked "Classical" is obtained by evaluating (5.37) using quadrature formulae, and has only two digits of accuracy. By contrast, the column marked "MRA" is correct to the accuracy ϵ chosen in Step 1 of the algorithm.

n	**Classical**	**MRA**
0	-5.471	-5.508
1	2.359	2.375
2	-0.058	-0.059
3	0.150	0.152
4	0.064	0.064
5	0.040	0.041

5.3.5 Asymptotic Condition for Integral Operators

For integral operators, the coefficients $t_n, n \in \mathbb{Z}$ are computed according to the formula

$$t_n^j = \int_{-\infty}^{\infty} K(x)\Phi(2^{-j}x - n)\, dx\,,$$

where $\Phi(x)$ is the autocorrelation of the scaling function. The following proposition enables us to determine the coefficients $\{t_n^j\}$ for all n such that $(2^j n)$ is sufficiently removed from a singularity of the kernel K.

Proposition 5.3.2 *Let Φ be the autocorrelation of a compactly supported scaling function in an MRA with M vanishing moments. Let I_n^j denote the support interval of $\Phi(2^{-j}x - n)$. Then*

$$t_n^j = 2^j K(2^j n) + \varepsilon_n^j\,, \tag{5.38}$$

where

$$|\varepsilon_n^j| \le 2^{(m+1)j}\mathcal{C}_m \sup_{\xi \in I_n^j} \frac{|K^{(m)}(\xi)|}{m!}\,, \tag{5.39}$$

for $1 \le m \le 2M$. The constant $\mathcal{C}_m$ depends on Φ but not on K.

Proof: Expanding $K(x)$ in a Taylor series about the point $(2^j n)$ we have

$$K(x) = \sum_{l=0}^{m-1} \frac{K^{(l)}(2^j n)}{l!}(x - 2^j n)^l + \frac{K^{(m)}(\xi)}{m!}(x - 2^j n)^m\,,$$

where $1 \le m \le 2M$. Substituting this expansion into (2.56), we have

$$t_n^j = \sum_{l=0}^{m-1} \frac{K^{(l)}(2^j n)}{l!} 2^{lj} \int_{-\infty}^{\infty} (2^{-j}x - n)^l \Phi(2^{-j}x - n)\, dx + \varepsilon_n^j\,,$$

where

$$\varepsilon_n^j = 2^{mj} \int_{-\infty}^{\infty} \frac{K^{(m)}(\xi_n)}{m!}(2^{-j}x - n)^m \Phi(2^{-j}x - n)\, dx\,. \tag{5.40}$$

With the change of variable $y = 2^{-j}x - n$ this reduces to

$$\begin{aligned} t_n^j &= \sum_{l=0}^{m-1} \frac{K^{(l)}(2^j n)}{l!} 2^{(l+1)j} \int_{-\infty}^{\infty} y^l \Phi(y)\, dy + \varepsilon_n^j \\ &= 2^j K(2^j n) + \varepsilon_n^j\,, \end{aligned}$$

which follows from Proposition(2.1.1), since $m - 1 \leq 2M - 1$.

Now let us consider the error term (5.40). Since the integrand vanishes outside the support interval I_n^j, we have

$$\begin{aligned} |\varepsilon_n^j| &\leq 2^{mj} \int_{-\infty}^{\infty} \frac{|K^{(m)}(\xi_n)|}{m!} |2^{-j}x - n|^m |\Phi(2^{-j}x - n)|\, dx \\ &\leq 2^{mj} \sup_{\xi \in I_n^j} \frac{|K^{(m)}(\xi)|}{m!} \int |2^{-j}x - n|^m |\Phi(2^{-j}x - n)|\, dx \\ &\leq 2^{(m+1)j} \sup_{\xi \in I_n^j} \frac{|K^{(m)}(\xi)|}{m!} \int |y|^m |\Phi(y)|\, dy\,. \end{aligned}$$

To complete the proof, set

$$\begin{aligned} \mathcal{C}_m &= \int |x|^m |\Phi(x)|\, dx \\ &\leq \int\!\!\int |x|^m |\phi(x+y)||\phi(y)|\, dy\, dx \\ &\leq \int\!\!\int (|x| + |y|)^m |\phi(x)||\phi(y)|\, dy\, dx \\ &\leq \sum_{n=0}^{m} \binom{m}{n} \int |x|^{m-n} |\phi(x)|\, dx \int |y|^n |\phi(y)|\, dy \\ &\leq \sum_{n=0}^{m} \binom{m}{n} C_{m-n} C_n\,, \end{aligned}$$

where we have used Lemma(2.2.2). □

5.4 Two-Scale Difference Equation in the Fourier Domain

In this section we introduce an operator on the space $L^2([0, 2\pi])$ of 2π-periodic functions that is an anologue of the two-scale difference equation (5.24).

Define the formal trigonometric series

$$\hat{t}(\xi) = \sum_{-\infty}^{\infty} t_n e^{in\xi}\,, \tag{5.41}$$

where $t_n, n \in \mathbb{Z}$ are the coefficients that satisfy the two-scale difference equation (5.24). Recall the definition of the trigonometric series $M_0(\xi)$ from equation (2.17),

$$M_0(\xi) = \frac{1}{4}\sum_{-\infty}^{\infty} a_m e^{im\xi},$$

which for the MRAs under consideration (see Section 5.3.1) has the form

$$\begin{aligned} M_0(\xi) &= \frac{1}{2} + \frac{1}{4}\sum_{m=1}^{M} a_{2m-1}[e^{i\xi(2m-1)} + e^{i\xi(1-2m)}] \\ &= \frac{1}{2} + \frac{1}{2}\sum_{m=1}^{M} a_{2m-1}\cos[(2m-1)\xi]\,. \end{aligned} \tag{5.42}$$

where the coefficients a_{2m-1} are given in (2.29). We show that the two-scale difference equation (5.24) can be expressed as an operator equation. We state this as a proposition.

Proposition 5.4.1 *Define the operator M_0 on $L^2([0,2\pi])$ by*

$$(M_0 f)(\xi) = M_0(\xi/2)f(\xi/2) + M_0(\xi/2+\pi)f(\xi/2+\pi)\,. \tag{5.43}$$

Then

$$(M_0\hat{t})(\xi) = 2^{-\alpha}\,\hat{t}(\xi)\,, \tag{5.44}$$

where $\hat{t}(\xi)$ is the trigonometric series (5.41).

Proof: Formally multiply the two series $M_0(\xi)$ and $\hat{t}(\xi)$ to obtain

$$M_0(\xi)\hat{t}(\xi) = \frac{1}{4}\sum_n e^{in\xi}\sum_m a_m t_{n-m}\,.$$

Now form the sum

$$\begin{aligned} (M_0\hat{t})(\xi) &= M_0(\xi/2)\hat{t}(\xi/2) + M_0(\xi/2+\pi)\hat{t}(\xi/2+\pi) \\ &= \frac{1}{4}\sum_n e^{in\xi/2}\sum_m a_m t_{n-m} + \frac{1}{4}\sum_n e^{in(\xi/2+\pi)}\sum_m a_m t_{n-m} \\ &= \frac{1}{4}\sum_n e^{in\xi/2}\,(1+(-1)^n)\sum_m a_m t_{n-m} \\ &= \frac{1}{2}\sum_n e^{in\xi}\sum_m a_m t_{2n-m}\,. \end{aligned}$$

Using the fact that $a_{2m} = 0$ if $m \neq 0$, $a_0 = 2$, and $a_{1-2m} = a_{2m-1}$, we have

$$(M_0\hat{t})(\xi) = \sum_n e^{in\xi}\left\{t_{2n} + \frac{1}{2}\sum_{m=1}^{M} a_{2m-1}\left[t_{2n-2m+1} + t_{2n-1+2m}\right]\right\}. \tag{5.45}$$

Comparing (5.45) to (5.24), we have

$$(M_0\hat{t})(\xi) = \sum_n e^{in\xi} 2^{-\alpha} t_n = 2^{-\alpha}\hat{t}(\xi)\,,$$

which proves the proposition. □

Some relevant properties of this operator are given by the following proposition.

Proposition 5.4.2 *The adjoint of M_0 is the operator M_0^*, defined by*

$$(M_0^* g)(\xi) = 2M_0(\xi)g(2\xi)\,, \tag{5.46}$$

and we also have

$$\|M_0\| = 2\,. \tag{5.47}$$

Proof: The adjoint is defined by the equation

$$(M_0 f, g) = (f, M_0^* g)$$

where f and g are 2π-periodic functions in $L^2([0, 2\pi])$. Using (5.43), we compute

$$\begin{aligned}
(M_0 f, g) &= \int_0^{2\pi} (M_0 f)(\xi)\overline{g}(\xi)\,d\xi \\
&= \int_0^{2\pi} [M_0(\xi/2)f(\xi/2) + M_0(\xi/2+\pi)f(\xi/2+\pi)]\overline{g}(\xi)\,d\xi \\
&= 2\int_0^{\pi} M_0(\xi)f(\xi)\overline{g}(2\xi)\,d\xi + 2\int_\pi^{2\pi} M_0(\xi)f(\xi)\overline{g}(2\xi - 2\pi)\,d\xi\,.
\end{aligned}$$

But since g is 2π-periodic, we have

$$\begin{aligned}
(M_0 f, g) &= 2\int_0^{2\pi} M_0(\xi)f(\xi)\overline{g}(2\xi)\,d\xi \\
&= \int_0^{2\pi} f(\xi)(\overline{M_0^* g})(\xi)\,d\xi\,,
\end{aligned}$$

which implies (5.46) (it can be seen from the explicit formula (5.42) that $M_0(\xi)$ is real).

To prove (5.47), we make use of the well-known fact that $\|M_0\| = \|M_0^*\|$, and compute the norm of the adjoint, which is easier. The following proof makes use of standard arguments, which can be found e.g. in [15, pp.55-56].

First, consider

$$\begin{aligned}
\|M_0^* g\|^2 &= \int_0^{2\pi} |(M_0^* g)(\xi)|^2\,d\xi \\
&= \int_0^{2\pi} |2M_0(\xi)g(2\xi)|^2\,d\xi \\
&\le 4\sup|M_0(\xi)|^2 \int_0^{2\pi} |g(2\xi)|^2\,d\xi\,.
\end{aligned}$$

But $\sup |M_0(\xi)|^2 = 1$, so we have

$$\|M_0^* g\|^2 \leq 4\|g\|^2 ,$$

which implies that $\|M_0^*\| \leq 2$.

Next, we make use of the fact that $M_0(0) = 1$. Define the sequence of functions $g_n(\xi) = \sqrt{n}$ if $|\xi| < 1/n, 0$ otherwise. Then we have

$$\begin{aligned} \|M_0^* g_n\|^2 &= \int_{-\pi}^{\pi} |(M_0^* g_n)(\xi)|^2 \, d\xi \\ &= n \int_{-1/2n}^{1/2n} |2M_0(\xi)|^2 \, d\xi , \end{aligned}$$

and by the continuity of M_0 we have

$$\|M_0^* g_n\|^2 \to |2M_0(0)|^2 = 4 , \quad \text{as } n \to \infty .$$

This implies that $\|M_0^*\| \geq 2$, and hence $\|M_0^*\| = 2$. □

Corollary 5.4.1 *If λ is an eigenvalue of M_0, then $|\lambda| \leq 2$.*

Proof: This follows immediately from (5.47), since $|\lambda| \leq \|M_0\|$. □

Remark: Corollary(5.4.1) shows that, if λ is an eigenvalue of M_0, and $\lambda = 2^{-\alpha}$, then $\alpha \geq -1$.

The operator M_0 has been considered in [6] and [19], for the eigenvalue $2^{-\alpha} = 1$, in connection with the construction of orthonormal wavelet bases. More generally, the operator equation (5.41) has been considered in [2] for eigenvalues $2^{-\alpha} = 1/2, 1/4, \ldots$, corresponding to $\alpha = 1, 2, \ldots$, in connection with differential operators.

We now show that $\lambda = 1$ is always an eigenvalue of the operator (5.43). We recall equation (2.18),

$$M_0(\xi) + M_0(\xi + \pi) = 1 .$$

Comparing this to (5.43), it is easily verified that $\hat{t}(\xi) = 1$ satisfies (5.44) with $2^{-\alpha} = 1$. This trigonometric polynomial corresponds to the coefficient sequence

$$t_n = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

and the corresponding kernel is

$$T(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} t_{m-n} \phi(x - m) \phi(y - n) = \sum_{-\infty}^{\infty} \phi(x - n) \phi(y - n) ,$$

which is the projector onto the subspace V_0 (see Section 2.2.1 and equation (2.42)).

Our interest in the operator M_0 is summed up in the following propositions, the first of which is essentially a restatement of Proposition(5.4.1).

Proposition 5.4.3 *Let K be a linear, homogeneous operator that commutes with translation. If the quantities*

$$t_n = \int_{-\infty}^{\infty} \phi(x-n)(K\phi)(x)\, dx$$

are well-defined for every integer n, then the series $\hat{t}(\xi) = \sum t_n e^{in\xi}$ satisfies the eigenvalue equation

$$(M_0 \hat{t})(\xi) = 2^{-\alpha}\, \hat{t}(\xi)\,.$$

Thus, the sequence $t_n\,, n \in \mathbb{Z}$ may be identified with a trigonometric series $\hat{t}(\xi)$, which is an eigenvector of the operator M_0. The question therefore arises whether or not the regularization produced by the algorithm of Section 5.3.3 is an eigenvector of M_0.

Proposition 5.4.4 *Let K be a linear operator that is homogeneous of degree α and commutes with translation. Let $t_n, n \in \mathbb{Z}$ be the regularization computed according to the algorithm of Section 5.3.3. Let $\hat{t}(\xi)$ be the formal trigonometric series $\sum t_n e^{in\xi}$. If $2^{-\alpha}$ is not an eigenvalue of the matrix A, then $\hat{t}$ satisfies the eigenvalue equation*

$$(M_0 \hat{t})(\xi) = 2^{-\alpha}\, \hat{t}(\xi)\,.$$

If $2^{-\alpha}$ is an eigenvalue of A, then $\hat{t}$ satisfies the equation

$$(M_0 \hat{t})(\xi) = 2^{-\alpha}\, \hat{t}(\xi) + r(\xi)\,, \tag{5.48}$$

where $r(\xi)$ is a trigonometric polynomial, and $(M_0 r)(\xi) = 2^{-\alpha} r(\xi)$. Furthermore, we have

$$r(\xi) = \sum_{1-2M}^{2M-1} r_n e^{in\xi}\,,$$

where the vector $\mathbf{r} = \{r_{1-2M}\,, \ldots\,, r_{2M-1}\}$ is given by

$$\mathbf{b} - (\lambda I - A)\tau = \mathbf{r}\,.$$

Thus, if the system (5.28) has a solution, then $\mathbf{r} = 0$, and in place of (5.48) we have

$$(M_0 \hat{t})(\xi) = 2^{-\alpha}\, \hat{t}(\xi)\,.$$

Proof: Let us define the trigonometric series σ and trigonometric polynomial τ by,

$$\sigma(\xi) = \sum_{-\infty}^{\infty} \sigma_n e^{in\xi}\,, \quad \text{where } \sigma_n = \begin{cases} t_n\,, & |n| \geq 2M \\ 0\,, & |n| < 2M \end{cases}$$

and

$$\tau(\xi) = \sum_{1-2M}^{2M-1} t_n e^{in\xi}\,.$$

Then $\hat{t}(\xi) = \sigma(\xi) + \tau(\xi)$. We also define the trigonometric polynomial

$$b(\xi) = \sum_{1-2M}^{2M-1} b_n e^{in\xi} \,,$$

where $\mathbf{b} = \{b_{1-2M}, \ldots, b_{2M-1}\}$ is the vector that appears in equation (5.28). Put $\lambda = 2^{-\alpha}$. Then

$$M_0 \sigma = \lambda\sigma + b \,,$$

and also

$$M_0 \tau = A\tau \,,$$

where A is the matrix that appears in (5.28). If λ is not an eigenvalue of A, then $b - (\lambda I - A)\tau = 0$, so that $A\tau = \lambda\tau - b$. If λ is an eigenvalue of A, then $b - (\lambda I - A)\tau = r$, where $Ar = \lambda r$ (see the Appendix A), so that $A\tau = r + \lambda\tau - b$. In the first case, we have

$$\begin{aligned} M_0 \hat{t} &= M_0\sigma + M_0\tau \\ &= (\lambda\sigma + b) + (\lambda\tau - b) \\ &= \lambda(\sigma + \tau) \\ &= \lambda\hat{t} \,. \end{aligned}$$

In the second case, we have

$$\begin{aligned} M_0 \hat{t} &= M_0\sigma + M_0\tau \\ &= (\lambda\sigma + b) + (r + \lambda\tau - b) \\ &= \lambda(\sigma + \tau) + r \\ &= \lambda\hat{t} + r \,. \end{aligned}$$

which proves the proposition. □

Bibliography

[1] C. Anderson. A method of local corrections for computing the velocity field due to a distribution of vortex blobs. *Journal of Computational Physics*, 62, 1986.

[2] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM Journal of Numerical Analysis*, 6(6):1716–1740, 1992.

[3] G. Beylkin. On the fast fourier transform of functions with singularities. *Applied and Computational Harmonic Analysis*, 2:363–381, 1995.

[4] G. Beylkin, R. Coifman, and V. Rohklin. Fast wavelet transforms and numerical algorithms I. In *Communications on Pure and Applied Mathematics*, volume XLIV. John Wiley and Sons, Inc., 1991.

[5] C. Chui. *An Introduction to Wavelets.* Academic Press, Inc., 1992.

[6] A. Cohen, I. Daubechies, and J. Feauveau. Biorthogonal bases of compactly supported wavelets. In *Communications on Pure and Applied Mathematics*, volume XLV, pages 485–560. John Wiley and Sons, Inc., 1992.

[7] I. Daubechies. Orthonormal bases of compactly supported wavelets. In *Communications on Pure and Applied Mathematics*, volume XLI, pages 909–996. John Wiley and Sons, Inc., 1988.

[8] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, 1992.

[9] A. Dutt, M. Gu, and V. Rohklin. Fast algorithms for polynomial interpolation, integration and differentiation. Technical Report YALEU/DCS/RR-977, Yale University, July 1993.

[10] A. Dutt and V. Rohklin. Fast fourier transforms for nonequispaced data II. *Applied and Computational Harmonic Analysis*, 2(1):85–100, 1995.

[11] J. Eastwood and R. Hockney. *Computer Simulation Using Particles.* IOP Publishing Ltd, 1988.

[12] P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 64:253–287, 1921.

[13] I.M. Gel'fand and G.E. Shilov. *Generalized Functions*, volume 1. Academic Press, 1964.

[14] G. Goertzel. An algorithm for the evaluation of finite trigonometric series. *American Mathematical Monthly*, 65:34–35, 1958.

[15] I. Gohberg and S. Goldberg. *Basic Operator Theory.* Birkhäuser, 1981.

[16] L. Greengard and V. Rohklin. A fast algorithm for particle simulations. *Journal of Computational Physics*, 73, 1987.

[17] A.S. Householder. *The Theory of Matrices in Numerical Analysis.* Dover Publications, Inc., 1964.

[18] Y. Katznelson. *An Introduction to Harmonic Analysis.* Dover Publications, Inc., 1976.

[19] W. Lawton. Necessary and sufficient conditions for constructing orthonormal wavelet bases. *Journal of Mathematical Physics*, 1(32):57–61, 1991.

[20] Y. Meyer. *Wavelets and Operators*, volume 37 of *Cambridge studies in advanced mathematics.* Cambridge University Press, 1992.

[21] J. Strain. Fast potential theory II, layer potentials and discrete sums. *Journal of Computational Physics*, 99(2), 1992.

[22] E. Titchmarsh. *Introduction to the Theory of Fourier Integrals.* Chelsea Publishing Company, third edition, 1986.

[23] G. Tolstov. *Fourier Series.* Dover Publications, Inc., 1962.

Appendix A

Solution of the System $\lambda x = Ax + b$

In this appendix we consider solution of a linear system

$$\lambda x = Ax + b\,, \tag{A.1}$$

where A is a square, diagonalizable matrix, but is otherwise arbitrary. We assume that λ is an eigenvalue of A (otherwise there is no difficulty). When A is symmetric, the situation is well understood.

In the symmetric case, invariant subspaces belonging to distinct eigenvalues are orthogonal. If a symmetric matrix A has m distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, where $m \leq n$, and if P_i is the projector onto the invariant subspace belonging to λ_i, then we have the operator identity

$$I = P_1 + \cdots + P_m\,,$$

called a partition of unity. Moreover, the system

$$\lambda_i x = Ax + b \tag{A.2}$$

has a solution if and only if $P_i b = 0$, and the solution is unique up to addition of an eigenvector of A belonging to λ_i.

It will be shown below that, even if A is not symmetric, a sequence of matrices D_i can be constructed that satisfies

$$I = D_1 + \cdots + D_m\,,$$

the system (A.2) has a solution if and only if $D_i b = 0$, and the solution is unique up to addition of an eigenvector of A belonging to λ_i.

The matrices D_i imitate the projection matrices P_i in that they are idempotent ($D_i^2 = D_i$), and satisfy $D_i x = x \Leftrightarrow x = P_i x$. However, they are not symmetric unless A is symmetric, in which case $D_i = P_i$.
In this appendix, we prove the following proposition.

Proposition A.0.5 *Let A be a diagonalizable $(n \times n)$ matrix. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of A, where $m \leq n$. Let V_i be the invariant subspace of A belonging to λ_i. Let U_i be the invariant subspace of A^T belonging to λ_i. Let D_i be the unique $(n \times n)$ matrix that satisfies*

$$v \in V_i \Leftrightarrow D_i v = v\,, \tag{A.3}$$

and

$$u \perp U_i \Leftrightarrow D_i u = 0\,. \tag{A.4}$$

Then

$$D_1 + \cdots + D_m = I\,, \tag{A.5}$$

and there exists a unique vector x that satisfies $x \perp V_i$ and

$$\lambda_i x = Ax + b \tag{A.6}$$

if and only if $D_i b = 0$. When this is the case, x is given by

$$x = S_i(\lambda_i I - S_i^T A S_i)^{-1} S_i^T b\,, \tag{A.7}$$

where S_i is a matrix whose columns form an orthonormal basis for the orthogonal complement of V_i.

The proof of the proposition will be given below after we have established several lemmas.

Let us choose a fixed index i from the list, $i \in \{1, \ldots, m\}$. In what follows, our notation will be simplified if we suppress the subscript, yet keeping in mind all the while that we are discussing a fixed eigenvalue of A, together with its invariant subspace, etc. Thus, we understand that $\lambda = \lambda_i$, $V = V_i$, and $U = U_i$. We also take $V^\perp$ and $U^\perp$ to stand for the orthogonal complements of U_i and V_i, respectively, and let $S = S_i$ be a matrix whose columns are an orthonormal basis for $V^\perp$. Then $S^T S = I$, and SS^T is the projector onto $V^\perp$. In the following lemma, we keep in mind that $D = D_i$.

Lemma A.0.1 *There exists a unique matrix D that satisfies $v \in V \Leftrightarrow Dv = v$, and $u \perp U \Leftrightarrow Du = 0$.*

Proof: We show how to construct D. Let R be a matrix with linearly independent columns that are a basis for V, and let L be a matrix with linearly independent columns that are a basis for U. By definition, we have

$$AR = \lambda R\,, \qquad L^T A = \lambda L^T\,. \tag{A.8}$$

We also have the following expressions for the projectors P_R and P_L onto the column spaces of R and L, i.e. onto V and U, respectively (see e.g. [17, p.8]),

$$P_R = R(R^T R)^{-1} R^T\,, \qquad P_L = L(L^T L)^{-1} L^T\,. \tag{A.9}$$

Now set

$$D = R(L^T R)^{-1} L^T\,, \tag{A.10}$$

where the linear independence of the columns of R and L implies the existence of the inverse. We now show that the matrix in (A.10) has the required properties.

First note that $u \perp U$ if and only if $L^T u = 0$, which is a simple consequence of the definition of L, but this implies that $u \perp U \Leftrightarrow Du = 0$.

To show that $v \in V \Leftrightarrow Dv = v$, we have

$$\begin{aligned} P_R D &= [R(R^T R)^{-1} R^T][R(L^T R)^{-1} L^T] \\ &= R(R^T R)^{-1}(R^T R)(L^T R)^{-1} L^T \\ &= R(L^T R)^{-1} L^T = D\,, \end{aligned}$$

and also

$$\begin{aligned} DP_R &= [R(L^T R)^{-1} L^T][R(R^T R)^{-1} R^T] \\ &= R(L^T R)^{-1}(L^T R)(R^T R)^{-1} R^T \\ &= R(R^T R)^{-1} R^T = P_R\,. \end{aligned}$$

Now, if $P_R v = v$, then $P_R = DP_R$ implies that $v = P_R v = DP_R v = Dv$. Conversely, if $Dv = v$, then $D = P_R D$ implies that $v = Dv = P_R Dv = P_R v$. Hence, $Dv = v \Leftrightarrow P_R v = v \Leftrightarrow v \in V$.

To show that D is unique, we note that the dimension of U is equal to the dimension of V, whence it follows that

$$\dim(V) + \dim(U^\perp) = n\,.$$

Since $v \in V \Rightarrow Dv = v$ and $u \in U^\perp \Rightarrow Du = 0$, it follows that the null space of D is $U^\perp$ and the range of D is V. As we have specified its nullspace, and specified the value it assigns to each element of its range, the matrix D is uniquely determined. □

Lemma A.0.2 *The matrix D satisfies $D^2 = D$, and*

$$AD = DA = \lambda D\,. \tag{A.11}$$

Proof: We verify directly that D is idempotent,

$$\begin{aligned} D^2 &= [R(L^T R)^{-1} L^T][R(L^T R)^{-1} L^T] \\ &= R(L^T R)^{-1}(L^T R)(L^T R)^{-1} L^T \\ &= R(L^T R)^{-1} L^T = D\,. \end{aligned}$$

Next, use (A.8) to compute

$$\begin{aligned} AD &= A[R(L^TR)^{-1}L^T] = (AR)(L^TR)^{-1}L^T = \lambda R(L^TR)^{-1}L^T = \lambda D\,, \\ DA &= [R(L^TR)^{-1}L^T]A = R(L^TR)^{-1}(L^TA) = R(L^TR)^{-1}\lambda L^T = \lambda D\,, \end{aligned}$$

which proves (A.11). □

Lemma A.0.3 *Let v be an eigenvector of A. If v belongs to λ, i.e. $Av = \lambda v$, then $Dv = v$. If v belongs to any eigenvalue of A distinct from λ, then $Dv = 0$.*

Proof: If $Av = \lambda v$, then by definition of the projector P_R we have $P_R v = v$. Hence by Lemma(A.0.1) we have $Dv = v$. Now suppose that $Av = \lambda_0 v$, where $\lambda_0 \neq \lambda$. Then we have

$$D(Av) = D(\lambda_0 v) = \lambda_0(Dv)\,,$$

and also using (A.11) we have

$$D(Av) = (DA)v = \lambda(Dv)\,.$$

Subtracting these two equations, we obtain

$$0 = (\lambda - \lambda_0)Dv\,,$$

which implies that $Dv = 0$, since $(\lambda - \lambda_0) \neq 0$. □

Lemma A.0.4 *Let $B = S^TAS$. Then λ is not an eigenvalue of B.*

Proof: Let y be any vector with dimension equal to the rank of S. Then

$$(\lambda I - B)y = (\lambda I - S^TAS)y = S^T(\lambda I - A)Sy\,,$$

since $S^TS = I$. Let $x = Sy$. Then x has length n, and we have

$$(\lambda I - B)y = S^T(\lambda I - A)x\,. \tag{A.12}$$

Now x is a vector in the space spanned by the columns of S, which is to say that $x \in V^\perp$, and if $y \neq 0$ then $x \neq 0$. But then $(\lambda I - A)x$ does not vanish, for if so then $x \in V$. Thus, the right-hand side of (A.12) vanishes only when the vector $v = (\lambda I - A)x$ satisfies $S^Tv = 0$. But this means that $v \perp V^\perp$, equivalently $v \in V$, which in turn implies that x satisfies

$$(\lambda I - A)^2 x = 0\,.$$

However, this equation has a non-zero solution if and only if λ is degenerate, but as we have assumed that A is diagonalizable this is not the case. Thus the right-hand side of (A.12) is nonzero for $y \neq 0$, and therefore λ is not an eigenvalue of B. □

Lemma A.0.5 *The following equation holds,*

$$D = I - (\lambda I - A)S(\lambda I - S^T AS)^{-1}S^T . \tag{A.13}$$

Proof: Let $D_1 = I - (\lambda I - A)S(\lambda I - S^T AS)^{-1}S^T$. We show that the nullspace and the range of D and D_1 are identical, and furthermore that for all x in the range of D, we have $Dx = D_1 x$, whence it must be the case that $D = D_1$.

It has been shown in the proof of Lemma(A.0.1) that the nullspace of D is $U^\perp$, the range of D is V, and $Dx = x$ for all $x \in V$.

Now consider D_1. If $x \in V$, then by definition we have $S^T x = 0$. Hence

$$D_1 x = x - (\lambda I - A)S(\lambda I - S^T AS)^{-1}S^T x = x .$$

Since $\mathbb{C}^n$ is equal to the direct sum of the nullspace of $(\lambda I - A^T)$ and the range of $(\lambda I - A)$, it follows that the range of $(\lambda I - A)$ is $U^\perp$. Thus for $u \in U^\perp$, $u \neq 0$, there exists a unique $x \in V^\perp$ such that $u = (\lambda I - A)x$. Since $SS^T x = x$, we have

$$u = (\lambda I - A)SS^T x ,$$

$$S^T u = [S^T(\lambda I - A)S]S^T x = (\lambda I - S^T AS)S^T x ,$$

and using Lemma(A.0.4) we can invert $(\lambda I - S^T AS)$ to obtain

$$S^T x = (\lambda I - S^T AS)^{-1}S^T u ,$$

$$SS^T x = x = S(\lambda I - S^T AS)^{-1}S^T u .$$

Now we have

$$\begin{aligned} D_1 u &= u - (\lambda I - A)S(\lambda I - S^T AS)^{-1}S^T u \\ &= u - (\lambda I - A)x = 0 . \end{aligned}$$

Since $u \in U^\perp$ was arbitrary, we conclude that the nullspace of D_1 is $U^\perp$, that the range of D_1 is V, and furthermore that $D_1 x = Dx$ for all $x \in V$. □

Proof of the Proposition: Equations (A.3) and (A.4) have been proved by Lemma(A.0.1).

To prove (A.5) we note that, since A is diagonalizable, there exists a basis of eigenvectors. To simplify the notation, we assume without loss of generality that $m = n$. Then for each vector x, we have

$$x = c_1 v_1 + \cdots + c_n v_n , \tag{A.14}$$

for some sequence of scalars $c_1, \ldots, c_n$, where $Av_i = \lambda_i v_i$. Multiplying (A.14) from the left by D_i, we have

$$\begin{aligned} D_i x &= D_i(c_1 v_1 + \cdots + c_n v_n) \\ &= c_1(D_i v_1) + \cdots + c_n(D_i v_n) \\ &= c_i v_i , \end{aligned}$$

since by Lemma(A.0.3) $D_i v_j = 0$ if $j \neq i$ and $D_i v_i = v_i$. It follows that (A.14) may be written as

$$x = D_1 x + \cdots + D_n x\,,$$

which implies the identity (A.5).

We now investigate the solution of (A.6). First note that Lemma(A.0.4) implies the existence of the inverse in (A.7). Write

$$r = b - (\lambda_i I - A)x\,. \tag{A.15}$$

Now, x is a solution to (A.6) if and only if $r = 0$ in (A.15). Substituting the expression (A.7) for x, equation (A.15) becomes

$$\begin{aligned} r &= b - (\lambda_i I - A)x \\ &= b - (\lambda_i I - A)S_i(\lambda_i I - S_i^T A S_i)^{-1} S_i^T b \\ &= [I - (\lambda_i I - A)S_i(\lambda_i I - S_i^T A S_i)^{-1} S_i^T]b \\ &= D_i b\,, \end{aligned}$$

where we have used Lemma(A.0.5). Thus $r = 0 \Leftrightarrow D_i b = 0$.

To show that the solution $x \perp V_i$, we recall that $S_i^T S_i = I$ and compute

$$\begin{aligned} S_i S_i^T x &= (S_i S_i^T) S_i(\lambda_i I - S_i^T A S_i)^{-1} S_i^T b \\ &= S_i(S_i^T S_i)(\lambda_i I - S_i^T A S_i)^{-1} S_i^T b \\ &= S_i(\lambda_i I - S_i^T A S_i)^{-1} S_i^T b = x\,. \end{aligned}$$

Thus, $x \perp V_i$ since $S_i S_i^T$ is the projector onto the orthogonal complement of V_i.

Finally, if ξ is any vector that satisfies $\xi \perp V_i$ and $\lambda_i \xi = A\xi + b$, then we have

$$\begin{aligned} (\lambda_i I - A)(x - \xi) &= 0 \\ S_i^T(\lambda_i I - A)(S_i S_i^T)(x - \xi) &= 0 \\ (\lambda_i I - S_i^T A S_i) S_i^T (x - \xi) &= 0 \end{aligned}$$

which implies that $S_i^T(x - \xi) = 0$ since $(\lambda_i I - S_i^T A S_i)$ is nonsingular. This means that $x - \xi \in V_i$, but we also have $x - \xi \perp V_i$ since both x and ξ are orthogonal to V_i. Therefore $x - \xi = 0$, and x is unique. □

Appendix B

Explicit Expression for $B_{m,n}(x)$

In this appendix we derive explicit expressions for the functions

$$B_{m,n}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\xi} \hat{\beta}^{(m)}(\xi - n\pi) \hat{\beta}^{(m)}(\xi + n\pi)\, d\xi \,. \tag{B.1}$$

These are (Fourier transforms of) the functions $\hat{\Phi}_n(\xi)$ introduced in Section 2.3.3, equation (2.67), in connection with the trigonometric expansion of a kernel. See also equation (3.23). Equation (B.1) corresponds to Φ_n when the basis function is the central B-spline $\beta^{(m)}(x)$.

It will be more convenient for the derivation to use the B-spline $N_m(x)$, in place of $\beta^{(m)}(x)$. This notation is used by some authors (e.g. [5]) to denote the spline which is piecewise polynomial of degree $(m-1)$, and is supported on the interval $[0, m]$. The relationship between $N_m(x)$ and $\beta^{(m)}(x)$ is

$$\beta^{(m-1)}(x) = N_m(x + m/2)\,.$$

It follows from this relationship that

$$\hat{N}_m(\xi - n\pi) \overline{\hat{N}}_m(\xi + n\pi) = \hat{\beta}^{(m-1)}(\xi - n\pi) \hat{\beta}^{(m-1)}(\xi + n\pi)\,.$$

Thus, it is sufficient to prove the following proposition.

Proposition B.0.6 *For $m \geq 1$, let $N_m(x)$ denote the spline function which is piecewise polynomial of degree $(m-1)$, and which vanishes outside the interval $[0, m]$. Define*

$$B_{m-1,n}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\xi} \hat{N}_m(\xi - n\pi) \overline{\hat{N}}_m(\xi + n\pi)\, d\xi \,. \tag{B.2}$$

Then, for $n \geq 1$, we have

$$B_{m-1,n}(x) = \sum_{k=0}^{m-1} \frac{(-i)^{m-k} a_{m,k}}{(2n\pi)^{m+k}} \left[e^{in\pi x} + (-1)^{m-k} e^{-in\pi x} \right] \left(\frac{d}{dx} \right)^{m+k} N_{2m}(x + m)\,, \tag{B.3}$$

where

$$a_{m,k} = \frac{(m-1+k)!}{(m-1)!\,k!}\,. \tag{B.4}$$

Note that the spline $N_{2m}(x+m)$ is equal to the central B-spline of degree $(2m-1)$.

Example: Take $m = 4$. The spline $N_4(x+2)$ is equal to the central B-spline $\beta^{(3)}(x)$. When $n = 0$, it is obvious that

$$B_{4,0}(x) = N_8(x+4) = \beta^{(7)}(x)\,,$$

which is the autocorrelation of $N_4(x)$. If $n \geq 1$, then using the formula (B.3), we have $B_{3,n}(x) =$

$$\left\{\frac{\cos(n\pi x)}{8(n\pi)^4}\frac{d^4}{dx^4} - \frac{\sin(n\pi x)}{4(n\pi)^5}\frac{d^5}{dx^5} - \frac{5\cos(n\pi x)}{16(n\pi)^6}\frac{d^6}{dx^6} + \frac{5\sin(n\pi x)}{16(n\pi)^7}\frac{d^7}{dx^7}\right\}\beta^{(7)}(x)\,.$$

Explicit expressions for the various derivatives of the central B-spline of degree seven are listed in the following table. We only consider $x \geq 0$ since the central B-splines are even functions. Note also that the last column is piecewise constant, thus the seventh derivative does not exist at integer points, but since it is multiplied by the term $\sin(n\pi x)$, which is zero at integer points, this does not cause any discontinuity in the function $B_{4,n}(x)$.

Interval	4^{th}	5^{th}	6^{th}	7^{th}
$0 \leq x \leq 1$	$\frac{35}{6}x^3 - 10x^2 + \frac{8}{3}$	$\frac{35}{2}x^2 - 20x$	$35x - 20$	35
$1 \leq x \leq 2$	$-\frac{21}{6}x^3 + 18x^2 - 28x + 12$	$-\frac{21}{2}x^2 + 36x - 28$	$-21x + 36$	-21
$2 \leq x \leq 3$	$\frac{7}{6}x^3 - 10x^2 + 28x - \frac{152}{6}$	$\frac{7}{2}x^2 - 20x + 28$	$7x - 20$	7
$3 \leq x \leq 4$	$-\frac{1}{6}x^3 + 2x^2 - 8x + \frac{32}{3}$	$-\frac{1}{2}x^2 + 4x - 8$	$-x + 4$	-1

We now state and prove three lemmas which will be used in the derivation of (B.3).

Lemma B.0.6 *Let $\triangle$ denote the backward difference operator. Then we can write*

$$(e^{-i\xi} - 2 + e^{i\xi})^m e^{-ix\xi} = \triangle^{2m} e^{-i\xi(x+m)}\,. \tag{B.5}$$

Proof: The operator $\triangle^m$ is defined recursively by

$$\begin{aligned}\triangle f(x) &= f(x) - f(x-1)\,,\\ \triangle^m f(x) &= \triangle\left(\triangle^{m-1} f(x)\right)\,.\end{aligned}$$

It follows from this that

$$\triangle^m f(x) = \sum_{n=0}^{m} \binom{m}{n} (-1)^n f(x-n)\,.$$

To prove the lemma, we use induction. First, consider

$$\begin{aligned}(e^{-i\xi} - 2 + e^{i\xi})e^{-ix\xi} &= e^{-i\xi(x+1)} - 2e^{-i\xi x} + e^{-i\xi(x-1)} \\ &= \triangle^2 e^{-i\xi(x+1)}\,.\end{aligned}$$

(In what follows, the algebra is somewhat tedious but entirely straightforward, and for the most part has been omitted.) Now, assume that (B.5) holds, and compute

$$\begin{aligned}&\left(e^{-i\xi} - 2 + e^{i\xi}\right)^{m+1} e^{-ix\xi} \\ =\;& \left(e^{-i\xi} - 2 + e^{i\xi}\right) \triangle^{2m} e^{-i\xi(x+m)} \\ =\;& \left(e^{-i\xi} - 2 + e^{i\xi}\right) \sum_{n=0}^{2m} \binom{2m}{n} (-1)^n e^{-i\xi(x+m-n)} \\ =\;& \sum_{n=0}^{2m} \binom{2m}{n} (-1)^n \left[e^{-i\xi(x+m+1-n)} - 2e^{-i\xi(x+m-n)} + e^{-i\xi(x+m-1-n)}\right] \\ =\;& \sum_{n=0}^{2m+2} \binom{2m+2}{n} (-1)^n e^{-i\xi(x+m+1-n)} \\ =\;& \triangle^{2m+2} e^{-i\xi(x+m+1)}\,,\end{aligned}$$

which establishes (B.5) for arbitrary m. □

Lemma B.0.7 *Define $R_m(z) = 1/z^m(z+1)^m$, for $m \geq 1$. This function has the partial fraction decomposition*

$$R_m(z) = (-1)^m \sum_{k=0}^{m-1} a_{m,k} \left\{ \frac{1}{(z+1)^{m-k}} + \frac{(-1)^{m-k}}{z^{m-k}} \right\}, \tag{B.6}$$

where

$$a_{m,k} = \frac{(m+k-1)!}{(m-1)!\,k!}\,. \tag{B.7}$$

Proof: Defining $f(z) = 1/z^m$, we compute

$$f^{(k)}(z) = \frac{(m+k-1)!}{(m-1)!} \frac{(-1)^k}{z^{m+k}}\,.$$

Expanding $R_m(z)$ about $z = 0$, we have

$$\begin{aligned}R_m(z) &= \frac{f(z+1)}{z^m} \\ &= \frac{1}{z^m} \sum_{k=0}^{\infty} \frac{f^{(k)}(1)}{k!} z^k \\ &= \sum_{k=0}^{\infty} a_{m,k} (-1)^k z^{k-m}\,.\end{aligned}$$

Similarly, expanding $R_m(z)$ about $z = -1$, we have

$$\begin{aligned} R_m(z) &= \frac{f(z)}{(z+1)^m} \\ &= \frac{1}{(z+1)^m} \sum_{k=0}^{\infty} \frac{f^{(k)}(-1)}{k!} z^k \\ &= (-1)^m \sum_{k=0}^{\infty} a_{m,k} (z+1)^{k-m} . \end{aligned}$$

The partial fraction decomposition is the sum of the singular parts of these two expansions, which consists of the first $(m-1)$ terms of each expansion. Combining these two finite sums gives (B.6). □

In (B.8), we consider the "function" $F_m(x)$ as a formal expression only, which will be useful later on. We do not claim that the integral converges for all positive integers m.

Lemma B.0.8 *Define*

$$F_m(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ix\xi}}{\xi^m}\, d\xi\,, \qquad m \geq 1\,. \tag{B.8}$$

These functions are related to the B-splines through the equation

$$\triangle^m F_m(x) = (-i)^m N_m(x)\,, \quad \text{a.e.} \tag{B.9}$$

Proof: We first note that for $m = 1$, the integral is well-known, and we have

$$F_1(x) = \frac{-i}{\pi} \int_0^{\infty} \frac{\sin(x\xi)}{\xi}\, d\xi = \frac{-i}{2} \operatorname{sgn}(x)\,. \tag{B.10}$$

An integration by parts yields the recurrence formula,

$$F_m(x) = \frac{-ix}{m-1} F_{m-1}(x)\,, \qquad m \geq 2\,. \tag{B.11}$$

We first verify that $\triangle F_1(x) = -iN_1(x)$, where $N_1(x) = \chi_{[0,1)}$. To do this, we simply note that

$$\frac{1}{2} \triangle \operatorname{sgn}(x) = \frac{1}{2} [\operatorname{sgn}(x) - \operatorname{sgn}(x-1)] = \begin{cases} 1 & , 0 < x < 1 \\ 1/2 & , x = 0 \text{ or } 1 \\ 0 & , \text{otherwise} \end{cases}$$

Thus, it follows that $\triangle F_1(x) = -iN_1(x)$, for all x except $x = 0$ or $x = 1$. In order to proceed with our induction proof, we make use of the following recurrence formula satisfied by the B-splines (see e.g. [5, p.86]),

$$N_m(x) = \frac{x}{m-1} N_{m-1}(x) + \frac{m-x}{m-1} N_{m-1}(x-1)\,, \qquad m \geq 2\,. \tag{B.12}$$

With this in hand, assume that (B.9) holds for index $m-1$, and use (B.11) to compute

$$\begin{aligned}
\triangle^m F_m(x) &= \triangle^m \frac{-ix}{m-1} F_{m-1}(x) \\
&= \frac{-i}{m-1} \sum_{n=0}^{m} \binom{m}{n} (\triangle^n x)\, \triangle^{m-n} F_{m-1}(x-n) \\
&= \frac{-i}{m-1} \left\{ \binom{m}{0} x \triangle^m F_{m-1}(x) + \binom{m}{1} \triangle^{m-1} F_{m-1}(x-1) \right\}.
\end{aligned}$$

Here we have used the product formula for the backward difference operator, as well as the fact that $\triangle^n x = 0$, for $n \geq 2$. Now we have

$$\begin{aligned}
\triangle^m F_m(x) &= \frac{-i}{m-1} \left\{ x \triangle^m F_{m-1}(x) + m \triangle^{m-1} F_{m-1}(x-1) \right\} \\
&= \frac{-ix}{m-1} \left[\triangle^{m-1} F_{m-1}(x) - \triangle^{m-1} F_{m-1}(x-1) \right] \\
&\quad - \frac{im}{m-1} \triangle^{m-1} F_{m-1}(x-1) \\
&= \frac{-i}{m-1} \left\{ x \triangle^{m-1} F_{m-1}(x) + (m-x) \triangle^{m-1} F_{m-1}(x-1) \right\} \\
&= (-i)^m \left\{ \frac{x}{m-1} N_{m-1}(x) + \frac{m-x}{m-1} N_{m-1}(x-1) \right\} \\
&= (-i)^m N_m(x),
\end{aligned}$$

where we have used (B.12) and our induction hypothesis. □

Proof of Proposition: Making the change of variable $\xi \leftarrow \xi + n\pi$, and using the formula

$$\hat{N}_m(\xi) = \left(\hat{N}_1(\xi)\right)^m = \left(\frac{e^{i\xi} - 1}{i\xi}\right)^m,$$

we can rewrite (B.2) as

$$\begin{aligned}
e^{in\pi x} B_{m-1,n}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\xi} \left(\frac{e^{i\xi} - 1}{i\xi}\right)^m \left(\frac{e^{-i\xi} - 1}{-i(\xi + 2n\pi)}\right)^m d\xi \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\xi} \frac{\left(e^{-i\xi} - 2 + e^{i\xi}\right)^m}{[-\xi(\xi + 2n\pi)]^m} d\xi.
\end{aligned}$$

Using Lemma(B.0.6), this equality can be written as

$$\begin{aligned}
e^{in\pi x} B_{m-1,n}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\triangle^{2m} e^{-i\xi(x+m)}}{[-\xi(\xi + 2n\pi)]^m} d\xi \\
&= \triangle^{2m} G_{m,n}(x+m),
\end{aligned}$$

where we have put

$$G_{m,n}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ix\xi}}{[-\xi(\xi+2n\pi)]^m}\,d\xi\,.$$

Now, using Lemma(B.0.7), we can express the denominator of the integrand in the form

$$\begin{aligned}
\frac{1}{[-\xi(\xi+2n\pi)]^m} &= \frac{(-1)^m}{(2n\pi)^m} R_m\left(\frac{\xi}{2n\pi}\right) \\
&= \sum_{k=0}^{m-1} \frac{a_{m,k}}{(2n\pi)^{m+k}}\left\{\frac{1}{(\xi+2n\pi)^{m-k}} + \frac{(-1)^{m-k}}{\xi^{m-k}}\right\}.
\end{aligned}$$

Using this equality, we can write

$$G_{m,n}(x) = \sum_{k=0}^{m-1} \frac{a_{m,k}}{(2n\pi)^{m+k}}\left[e^{2\pi inx} + (-1)^{m-k}\right]\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ix\xi}}{\xi^{m-k}}\,d\xi\,.$$

For convenience, let us introduce the functions

$$F_k(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ix\xi}}{\xi^k}\,d\xi\,, \qquad k \geq 1\,.$$

Then, using the easily verified identity

$$\triangle e^{2\pi inx} f(x) = e^{2\pi inx}\triangle f(x)\,,$$

we can write

$$e^{in\pi x} B_{m-1,n}(x) = \sum_{k=0}^{m-1} \frac{a_{m,k}}{(2n\pi)^{m+k}}\left[e^{2\pi inx} + (-1)^{m-k}\right]\triangle^{2m} F_{m-k}(x+m)\,. \quad \text{(B.13)}$$

Using Lemma(B.0.8), we have

$$\triangle^{2m} F_{m-k}(x) = \triangle^{m+k}\triangle^{m-k} F_{m-k} = (-i)^{m-k}\triangle^{m+k} N_{m-k}(x)\,.$$

The derivative formula $N_m'(x) = \triangle N_{m-1}(x)$ is well-known, and can be applied recursively to obtain

$$\triangle^{m+k} N_{m-k}(x) = \left(\frac{d}{dx}\right)^{m+k} N_{2m}(x)\,.$$

Finally, substitute the last two equations into (B.13) to obtain (B.3). □