Interpolating Wavelet Transforms

David L. Donoho
Department of Statistics
Stanford University

October, 1992

Abstract

We describe several "wavelet transforms" which characterize smoothness spaces and for which the coefficients are obtained by sampling rather than integration. We use them to re-interpret the empirical wavelet transform, i.e. the common practice of applying pyramid filters to samples of a function.

Key Words and Phrases. Interpolating Wavelet Transform. Interpolating Spline Wavelets. Interpolating Deslauriers-Dubuc Wavelets. Wavelet transforms of Sampled Data. Wavelet Interpolation of Sampled Data. Wavelets on the Interval.

Acknowledgements. This work was supported by grant NSF DMS 92-09130. These results were briefly described at the NATO-Advanced Study Institute "Wavelets and Applications", Il Ciocco, Italy, August 1992. The author is also with the Department of Statistics at the University of California, Berkeley.

Thanks to Albert Cohen of Université de Paris-Dauphine, for the suggestion to study Deslauriers-Dubuc interpolating wavelets rather than Spline interpolating wavelets. This has led to many fundamental improvements over an earlier version of this work.

Special thanks to Iain Johnstone for pointing out mistakes in an earlier version of this paper.

1 Introduction

1.1 Interpolating Transforms

Recently, several articles (e.g. (DeVore, Jawerth, Lucier, 1992), (Antonini, Barlaud, Daubechies, 1991)) have pointed out that compression of image data can be accomplished by quantization of wavelet coefficients. One calculates the wavelet coefficients, and quantizes them into discrete levels, which are represented by short bit strings (and perhaps further compressed by run-length schemes).

For compressing sequences of moving images, it is interesting to consider the possibility of a "wavelet compression machine" based on special-purpose massively-parallel hardware. In the abstract, such a machine would dedicate one processor per wavelet coefficient, and this processor would perform the calculation and subsequent quantization for that wavelet coefficient. The machine would operate in cycles where data would be acquired, all the wavelet coefficients computed and compressed in parallel, then the compressed data sent to a mass storage device. Ignoring the input/output issues, and focusing on the image processing issues, the processing cycle time of such a machine is determined by the longest time it takes to compute and compress any single wavelet coefficient.

For the usual wavelet transforms based on pyramid filtering, most coefficients take about $C \cdot \log_2(n)$ operations to compute, where n is the image extent. To compute a coefficient near the middle or top of the pyramid, it is necessary to compute coefficients at all finer scales, and there are as many as $\log_2(n)$ of those scales.

The "hardware" issues just raised lead to the following mathematical question: *What is the smallest amount of computation required in a wavelet transform to calculate one single wavelet coefficient?* We believe this question has practical interest, because it describes the parallel complexity of the wavelet transform, but we take this as a purely mathematical issue here.

In this paper we develop an *interpolating wavelet transform.* This is a non-orthogonal transform with formal resemblance to orthogonal wavelet transforms, in that it represents the object in terms of dilations and translations of a basic function – but for which the coefficients are obtained from linear combinations of samples rather than from integrals. The transform depends in a fundamental way on the interpolation scheme of Deslauriers-Dubuc.

The interpolating transform is optimal from the point of view of computing individual coefficients in parallel. The optimality is expressed by two properties. (For simplicity we focus on the 1-dimensional case in this paper.)

[IT1] *Computational Cost:* Each coefficient $\alpha_{j,k}$ of the transform can be calculated, independently of all other coefficients, in no more than $D+2$ multiply-adds.

[IT2] *Coefficient Decay.* The coefficients have decay properties for D-times differentiable functions which are comparable to the decay properties of coefficients of smooth orthogonal wavelet decompositions.

The interpolating transform has versions for the line and the interval, both of which we describe.

The "wavelet compression machine" described above would process wavelet coefficients entirely in parallel; that is, each wavelet coefficient would be compressed without waiting to inquire about the value of any other. This proposal seems reasonable because of results showing that simple thresholding of orthogonal wavelet coefficients is a kind of optimal compression algorithm (DeVore, Jawerth, and Popov, 1990). One may object to simplistic nature of the data compression model employed there, but such results provide some evidence that parallel thresholding of *orthogonal* coefficients is a reasonable procedure. In contrast, one may imagine that the non-orthogonality or some other non-standard feature of the interpolating transform spoils everything, and makes simple thresholding of *interpolating* coefficients a very bad idea. If so, we would need to use some other compression scheme, perhaps a computationally intensive one, and we would end up with a fast transform giving coefficients which are very time-consuming to compress. Fortunately, this does not seem to be the case.

[IT3] *Parallel Compression.* A simple thresholding rule, processing each interpolating coefficient independently of the value of any other coefficients, yields near-optimal compression, in the sense that the reconstruction accuracy for a fixed number of terms is near-optimal, in a minimax sense, over Besov classes, when error is measured in L^∞-norm.

Thresholded interpolating transforms are extremal with respect to machine cycles for parallel computation and compression, and yet may be quite acceptable with respect to reconstruction accuracy.

1.2 Empirical and Hybrid Transforms

Our emphasis in this article is mathematical. In fact, our interest in the interpolating transform actually arose, not from compression, but from trying to understand a sometimes confusing issue – the fact that the phrase "Wavelet Transform" is presently used in two distinctly different ways. (In this article we exclude discussion of the continuous wavelet transform).

In the mathematical literature it refers to a transform W from functions f on the line or the interval to coefficient sequences $(\alpha_{j,k})$. The coefficients are defined by integrals, and for coefficients defined in this way there are a variety of powerful results showing how the wavelet coefficients characterize the modulus of continuity of the transformed function.

In the signal processing and engineering literature, it refers to a transform W_n^n that takes n sampled data – a digitized sound or image – into n coefficients by a scheme of hierarchical pyramid filtering. The transform may be accomplished in order n time.

The link between these two uses of the phrase "wavelet transform" is, of course, that the mathematical wavelet transform has a representation in terms of filtering, in which one uses filters which could also be profitably applied in signal processing and engineering problems.

To be more precise we recall the pyramid filtering algorithm for obtaining theoretical wavelet coefficients of functions in $L^2[0,1]$, as described in [CDJV]. Given $n = 2^{j_1}$ integrals $\beta_{j_1,k} = \int_0^1 \varphi_{j_1,k}(t) f(t)dt$, $k = 0, \dots, 2^{j_1} - 1$, "sampling" f near $2^{-j_1}k$, one iteratively applies a sequence of decimating high pass and low pass operators $H_j, L_j : \mathbf{R}^{2^j} \to \mathbf{R}^{2^{j-1}}$ via

$$\begin{aligned} (\beta_{j-1,\cdot}) &= L_j \circ (\beta_{j,\cdot}) \\ (\alpha_{j-1,\cdot}) &= H_j \circ (\beta_{j,\cdot}) \end{aligned}$$

for $j = j_1, j_1 - 1, \dots, j_0 + 1$, producing a sequence of $n = 2^{j_1}$ coefficients

$$((\beta_{j_0,\cdot}), (\alpha_{j_0,\cdot}), (\alpha_{j_0+1,\cdot}), \dots, (\alpha_{j_1-1,\cdot})).$$

The transformation U_{j_0,j_1} mapping $(\beta_{j_1,\cdot})$ into this sequence is a real orthogonal transformation. The coefficients calculated by this filtering process are actually integrals:

$$\beta_{j_0,k} = \int_0^1 \varphi_{j_0,k}(t) f(t)dt, \qquad 0 \le k < 2^{j_0}$$

$$\alpha_{j,k} = \int_0^1 \psi_{j,k}(t) f(t) dt, \qquad 0 \le k < 2^j.$$

Hence, starting from fine-scale integrals $(\beta_{j_1,k})_k$, filtering delivers the integrals $(\beta_{j_0,k})_k$ and $(\alpha_{j,k})_k$ at coarser scales.

For empirical work, one does not have access to the fine scale integrals $(\beta_{j_1,k})$, and so one can not actually use this filtering to calculate the theoretical wavelet coefficients of a function $f(t)$. However, it is interesting to consider applying the same filtering algorithm to sampled data $(f(k/n))_k$. Observe that (for k away from the boundary) $\varphi_{j_1,k}$ has integral $2^{-j_1/2}$ and that it is supported near $k/2^{j_1}$. Hence the $\beta_{j_1,k}$ behave much like the *samples*

$$b_{j_1,k} = n^{-1/2} f(k/n) \qquad k = 0, \dots, n-1.$$

One adjusts for the fact that the $\varphi_{j_1,k}$ near the boundary do not have the same integral as at the interior by a *preconditioning transformation* $P_D b = (\tilde{\beta}_{j_1,\cdot})$, affecting only the $D+1$ values at each end of the segment $(b_{j_1,k})_{k=0}^{2^{j_1}-1}$. Then one applies the algorithm of [CDJV], to $(\tilde{\beta}_{j_1,\cdot})$ in place of $(\beta_{j_1,k})$ producing not theoretical wavelet coefficients but what we call *empirical wavelet* coefficients:

$$\tilde{\theta}^{(n)} = ((\tilde{\beta}_{j_0,\cdot}), (\tilde{\alpha}_{j_0,\cdot}), (\tilde{\alpha}_{j_0+1,\cdot}), \dots, (\tilde{\alpha}_{j_1-1,\cdot})).$$

What is the connection between these two "Wavelet Transforms?" For applications in Donoho and Johnstone (1992a,b,c) it would be extremely desirable for the first n coefficients in the empirical transform to be exactly the same as the first n coefficients in the theoretical transform. Then we could be sure that empirical coefficients of a smooth function have the same properties as theoretical coefficients.

This, however, is too much to hope for. Obviously in some sense the first n empirical wavelet coefficients of the samples of f are approximately the same as the first n theoretical wavelet coefficients of f. The more smooth and regular the function f, the better the agreement. However, we are not interested here in such approximation arguments.

Instead, we explore here the idea that the *empirical wavelet coefficients are* **precisely** *the first n theoretical coefficients of f in a slightly modified transform.* In what follows, we construct a "hybrid wavelet transform" W_n, depending on $n = 2^{j_1}$, which has the following character. One starts from a wavelet transform based on orthogonal wavelets $\bar{\varphi}$ and $\bar{\psi}$ of compact support where the moments of $\bar{\psi}$ through order $\bar{D}$ are zero, and the wavelets have $\bar{R}$ continuous derivatives.

[ET1] W_n maps continuous functions on $[0,1]$ into countable coefficient sequences

$$\tilde{\theta} = \left((\tilde{\beta}_{j_0,\cdot}), (\tilde{\alpha}_{j_0,\cdot}), (\tilde{\alpha}_{j_0+1,\cdot}), \dots, (\tilde{\alpha}_{j_1-1,\cdot}) \dots\right).$$

[ET2] There are smooth functions $(\tilde{\varphi}_{j_0,k})$, $(\tilde{\psi}_{j,k})$, $0 \le k < 2^j$, $j \ge j_0$ which are $\bar{R}$-times differentiable functions of support width $\le C2^{-j}$ so that every function $f \in C[0,1]$ has an expansion

$$f = \sum_k \tilde{\beta}_{j_0,k} \tilde{\varphi}_{j_0,k} + \sum_{j \ge j_0} \sum_k \tilde{\alpha}_{j,k} \tilde{\psi}_{j,k}.$$

with uniform convergence of partial sums (the convergence is unconditional if f has some regularity).

[ET3] The first n coefficients $\left((\tilde{\beta}_{j_0,k})_k, (\tilde{\alpha}_{j_0,k})_k, \ldots, (\tilde{\alpha}_{j_1-1,\cdot})\right)$ result from pre-conditioned pyramid filtering $U_{j_1,j_0} \circ P_D$ of the samples $(b_{j,k} = n^{-1/2} f(k/n))$.

[ET4] The first n basis functions are nearly orthogonal with respect to the sampling measure: $\langle f, g \rangle_n = n^{-1} \sum_{k=0}^{n-1} f(k/n) g(k/n)$. The constants of equivalence do not depend on $n = 2^{j_1}$ but may depend on j_0 and on the wavelet $\bar{\psi}$ being used.

[ET5] The coefficients $\theta = W_n f$ measure smoothness.

Each Besov space $B^\sigma_{p,q}[0,1]$ with $1/p < \sigma < \min(\bar{R}, \bar{D})$ and $0 < p, q \leq \infty$ is characterized by the coefficients in the sense that

$$\|\tilde{\theta}\|_{b^\sigma_{p,q}} \equiv \|(\tilde{\beta}_{j_0,k})_k\|_{\ell_p} + (\sum_{j \geq j_0} (2^{js} (\sum_k |\tilde{\alpha}_{j,k}|^p)^{1/p})^q)^{1/q},$$

is an equivalent norm to the norm of $B^\sigma_{p,q}[0,1]$; here $s \equiv \sigma + 1/2 - 1/p$, with constants of equivalency that do not depend on n, but which may depend on p, q, j_0 and the wavelet filters.

Each Triebel-Lizorkin space $F^\sigma_{p,q}[0,1]$ with $1/p < \sigma < \min(\bar{R}, \bar{D})$, $1 < p, q \leq \infty$ is characterized by the coefficients, in the sense that, with $\chi_{j,k}$ the indicator function of the interval $[k/2^j, (k+1)/2^j)$,

$$\|\tilde{\theta}\|_{f^\sigma_{p,q}} \equiv \|(\tilde{\beta}_{j_0,k})_k\|_{\ell_p} + \|(\sum_{j \geq j_0} 2^{jsq} \sum_k |\tilde{\alpha}_{j,k}|^q \chi_{j,k})^{1/q}\|_{L^p[0,1]},$$

is an equivalent norm to the norm of $F^\sigma_{p,q}[0,1]$; here $s \equiv \sigma + 1/2$.

In words, the empirical wavelet coefficients, which derive solely from finite filtering calculations, are actually the first n theoretical coefficients for a nicely behaved transform of continuous functions.

This *interpretation* gives us several insights immediately. It shows for example that empirical wavelet coefficients of a smooth function automatically obey the same type of decay estimates as theoretical orthogonal wavelet coefficients. It also shows that "shrinking" empirical wavelet coefficients towards zero (by any procedure whatever, linear or nonlinear) always acts as a "smoothing operator" in any of a wide range of smoothness measures. It also shows that sampling followed by appropriate interpolation of the sampled values is a "smoothing operator" in any of a wide range of smoothness classes. Finally it shows that the theoretical wavelet coefficients are close to the empirical wavelet coefficients in a precise sense. In Donoho and Johnstone (1992c) and Donoho (1992) these facts are crucial for the study of certain nonlinear methods for smoothing and de-noising noisy, sampled data.

We emphasize that our aim in developing this hybrid transform is *not* to change the practice of empirical wavelet transforms, but to allow us to interpret and understand the transforms as currently practiced.

1.3 Contents

In the paper which follows, we develop the interpolating and hybrid transforms in a leisurely series of steps, which we believe are interesting in their own right.

Section 2 develops the interpolating wavelet transform using a non-orthogonal analog of multiresolution analysis based on interpolation schemes. From an algorithmic point of view, a particularly interesting instance of the IWT is based on interpolating wavelets of compact support, which were previously defined for another purpose by Deslauriers-Dubuc. They involve two-scale relations with finitely many nonzero coefficients, and possess the property [IT1] above. This section formulates results showing that the wavelet coefficients for the interpolating transform have decay properties comparable to those of orthogonal wavelet coefficients in case the function is continuous. Formally, the coefficients characterize Besov and Triebel-Lizorkin spaces with smoothness parameter $\sigma > 1/p$. Such spaces are spaces of uniformly continuous functions. Special cases are Hölder(-Zygmund) spaces and Sobolev spaces.

Section 3 develops an interpolating wavelet transform for the interval. It is based on the interpolating wavelets of compact support for the line, and bears a striking similarity to the [CDJV] wavelet transform for the interval based on orthogonal wavelets of compact support. It has a structure which agrees with the interpolating transform for the line "at the heart" of the interval, while having boundary-corrected wavelets "at the edges". This transform also characterizes those Besov and Triebel-Lizorkin spaces on the interval which consist of continuous functions. This fact is used to study the quality of reconstructions based on simple thresholding of interpolating coefficients.

Section 4 develops a hybrid transform for functions on the line. It consists of wavelets which at fine resolutions are interpolating, and at coarse resolutions approach orthogonal wavelets. This transform also characterizes Besov and Triebel spaces with $\sigma > 1/p$.

Finally, Section 5 develops the hybrid transform on the interval, and establishes the properties [ET1]-[ET5] mentioned above. Because of what takes place in Sections 2-4, these properties are quite evident.

Section 6 sketches what happens in the critical case of analysing bounded, discontinuous functions; discusses the relation between this work and other work with some contact, for example work on Schauder bases, on other wavelet-like transforms, and on interpolating splines.

2 Interpolating Transforms on the Line

We now develop a transform for $C(\mathbf{R})$ which has the same formal dyadic structure as orthonormal wavelet transforms, but for which the coefficients are obtained by sampling rather than integration.

2.1 Interpolating Wavelets

Definition 2.1 *An* (R, D) **interpolating wavelet** *is a father function* φ *satisfying five conditions:*

[IW1] Interpolation. φ *interpolates the Kronecker sequence at the integers:*

$$\varphi(k) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} . \tag{2.1}$$

[IW2] Two-Scale Relation. φ *can be represented as a linear combination of dilates and translates of itself:*

$$\varphi(x) = \sum_k \varphi(k/2)\varphi(2x-k). \tag{2.2}$$

[IW3] Polynomial Span. For an integer $D \geq 0$*, the collection of formal sums* $\sum_k \beta_k \varphi(t-k)$ *contains all polynomials of degree* D*.*

[IW4] Regularity. For some real $R > 0$*,* φ *is Hölder continuous of order* R*.*

[IW5] Localization. φ *and all its derivatives through order* $\lfloor R \rfloor$ *decay rapidly:*

$$|\varphi^{(m)}(x)| \leq A_\ell \cdot (1+|x|)^{-\ell}, \qquad x \in \mathbf{R}, \quad \ell > 0, \quad 0 \leq m \leq \lfloor R \rfloor. \tag{2.3}$$

There are two well-known families of such functions. The first are the *interpolating spline wavelets.* Let D be an odd integer > 0, and L_D be the *fundamental polynomial spline of degree* D (Schoenberg, 1972), i.e. the piecewise polynomial with knots at the integers $k \in \mathbf{Z}$, continuity C^{D-1}, and satisfying the interpolation conditions (2.1). This function is a $(D-1, D)$ interpolating wavelet; it is regular of order $R = D-1$; its derivatives through order $D-1$ decay exponentially with distance from 0; it satisfies a two-scale relation; and it generates all polynomials of degree D through its translates.

The second family consists of the *Deslauriers-Dubuc Fundamental functions.* Dubuc (1986), Deslauriers and Dubuc (1987,1989). Let D be an odd integer > 0. These are functions F_D defined by interpolating the Kronecker sequence at the integers to a function on the binary rationals by repeated application of the following rule. If F_D has already been defined at all binary rationals with denominator 2^j, $j \geq 0$, extend it, by polynomial interpolation, to all binary rationals with denominator 2^{j+1}, i.e. all points halfway between previously defined points. Specifically, to define the function at $(k+1/2)/2^j$ when it is already defined at all $k/2^j$, fit a polynomial $\pi_{j,k}$ to the data $(k'/2^j, F_D(k'/2^j))$ for $k' \in \{(k-(D-1)/2)/2^j, \ldots, (k+(D+1)/2)/2^j\}$ – this polynomial is unique – and set

$$F_D((k+1/2)/2^j) \equiv \pi_{j,k}((k+1/2)/2^j).$$

It turns out that this scheme defines a function which is uniformly continuous at the rationals and hence has a unique continuous extension to the reals. This extension defines an (R, D) interpolating wavelet for an $R = R(D)$. Indeed, the first three properties [IW1]-[IW3] come automatically by construction; the last [IW5] by the compact support of F_D. The condition [IW4], regularity, is much less obvious. Deslauriers and Dubuc (1987, 1989) and Daubechies and Lagarias (1991) develop techniques to estimate the regularity in specific cases. Recently, Saitoh and Beylkin (1992) have shown that F_D *is the autocorrelation function of the Daubechies wavelet of degree* $D+1$; see also Daubechies (1992, Page 210). Hence, F_D is at least as smooth as the corresponding Daubechies wavelet, roughly twice as smooth. It follows from this and results for Daubechies wavelets that the regularity

$$R(D) \geq Const \cdot D, \qquad D = 3, 5, 7, \ldots$$

The connection between orthonormal wavelets and interpolating wavelets is valid generally. If $\bar{\varphi}$ is a nice orthonormal scaling function, then its autocorrelation $\varphi = \bar{\varphi} \star \bar{\varphi}(-\cdot)$ is an interpolating wavelet. Indeed, smoothness, localization, and the two-scale relation are inherited from ϕ, and the interpolation conditions on φ are identical to the orthonormality conditions on $\bar{\varphi}$. Thus, for example, the autocorrelation of the Haar scaling function is the interpolating Schauder wavelet, the autocorrelation of the Battle-Lemarié wavelets is an interpolating polynomial spline wavelet, and the autocorrelation of the Meyer wavelet is an interpolating wavelet which is C^∞ and of rapid decay at ∞. However, we are primarily interested here in Deslauriers-Dubuc wavelets.

We emphasize that although we use the term "wavelets" to describe interpolating wavelets, they are *not* orthonormal wavelets. At best, interpolating spline wavelets are cousins of the traditional spline wavelets and Deslauriers-Dubuc wavelets are cousins of Daubechies wavelets. Nevertheless, we shall operate with these functions in much the same way as we would with orthonormal wavelets.

2.2 Interpolating Multiresolutions & Transforms

Define $\psi(t) = \varphi(2(t - \frac{1}{2}))$, and, by analogy with the wavelet transform of $L^2(\mathbf{R})$, $\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k)$, $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$.

Theorem 2.2 *Given (R,D)-interpolating wavelet φ, we may construct an* **interpolating wavelet transform**, *mapping continuous functions f into sequences $((\beta_{j_0,k}), (\alpha_{j_0,k}), (\alpha_{j_0+1,k}), \ldots)$ with each coefficient $\alpha_{j,k}$ depending only on samples of f at scale 2^{-j-1} and coarser, and that any f which is the sum of a polynomial of degree $\leq D$ and a function in $C_0(\mathbf{R})$ can be reconstructed from its coefficients,*

$$f = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j \geq j_0} \sum_k \alpha_{j,k}\psi_{j,k}, \tag{2.4}$$

with the infinite sum converging in sup norm when summed in the right order.

We prove this result by a leisurely discussion, occupying the remainder of this subsection. It is based on the $C_0(\mathbf{R})$ analog of $L^2(\mathbf{R})$ multiresolution analysis.

Let V_j denote the vector space of all sums $f(t) = \sum_k \beta_{j,k}\varphi_{j,k}(t)$, where the $\beta_{j,k}$ are of at-most polynomial growth in k. The spaces V_j have three important properties, which follow immediately from [IW1]-[IW3]:

Lemma 2.3 *Let V_j be generated by an (R,D)-interpolating wavelet. (1) For any $f \in V_j$ the coefficients $\beta_{j,k}$ in the sum $f = \sum_k \beta_{j,k}\varphi_{j,k}$ can be recovered by sampling:*

$$\beta_{j,k} = f(2^{-j}k)/2^{j/2}. \tag{2.5}$$

(2) We have the inclusion

$$V_j \subset V_{j+1}.$$

(3) If Π_D denotes all polynomials of degree $\leq D$, then

$$\Pi_D \subset V_j.$$

Because of (2.5), we formally define, for any continuous f, $P_j f$ as the interpolant $2^{-j/2} \sum f(2^{-j}k)\varphi_{j,k}(t)$. This linear operator acts as the identity on V_j, and so is a kind of non-orthogonal projection. It is actually well defined, by [IW5], for all continuous functions of at-most polynomial growth. For all continuous functions vanishing at ∞, $P_j f$ converges to f as j increases.

Lemma 2.4 *If $f \in C_0(\mathbf{R})$ then*

$$\|f - P_j f\|_\infty \to 0, \qquad j \to \infty, \tag{2.6}$$

Proof. Let $\omega(\delta; f)$ denote the modulus of continuity

$$\omega(\delta; f) = \sup_{|h|\leq\delta} \sup_x |f(x+h) - f(x)|.$$

For $f \in C_0(\mathbf{R})$, $\omega(\delta; f) \to 0$ as $\delta \to 0$. In the Appendix will prove the inequality

$$\omega(2^{-j}; P_j f) \leq C \cdot \omega(2^{-j}; f) \tag{2.7}$$

with C independent of f and j. From $(P_j f)(k/2^j) = f(k/2^j)$, we get that for $x \in [0,1]$

$$\begin{aligned}
|f((k+x)/2^j) - (P_j f)((k+x)/2^j)| &\leq |f((k+x)/2^j) - f(k/2^j)| \\
&\quad + |(P_j f)((k+x)/2^j) - (P_j f)(k/2^j)| \\
&\leq \omega(2^{-j}, f) + \omega(2^{-j}, P_j f) \\
&\leq C \cdot \omega(2^{-j}, f).
\end{aligned}$$

Hence

$$\|f - P_j f\|_\infty \leq C \cdot \omega(2^{-j}, f) \to 0, \qquad j \to \infty.$$

which proves (2.6). □

Let W_j be the vector space of all formal sums $f(t) = \sum_k \alpha_{j,k}\psi_{j,k}$, and note that, by the particular definition of $\psi_{j,k}$,

$$\psi_{j,k} = \sqrt{2} \cdot \varphi_{j+1,2k+1}.$$

Hence, $W_j \subset V_{j+1}$. Now suppose, given sequences $(\beta_{j,k})$ and $(\alpha_{j,k})$, we construct a function

$$f = \sum_k \beta_{j,k}\varphi_{j,k} + \sum_k \alpha_{j,k}\psi_{j,k}. \tag{2.8}$$

As $f \in V_{j+1}$, we also have a representation

$$f = \sum \beta_{j+1,k}\varphi_{j+1,k}. \tag{2.9}$$

What is the relation between these two representations? Because $\sum_k \beta_{j,k}\varphi_{j,k}$ and $\sum \beta_{j+1,k}\varphi_{j+1,k}$ agree on the coarse grid, we have

$$\beta_{j,k} = \sqrt{2} \cdot \beta_{j+1,2k}, \qquad k \in \mathbf{Z}.$$

Because $\sum_k \beta_{j,k}\varphi_{j,k} + \sum_k \alpha_{j,k}\psi_{j,k}$ and $\sum \beta_{j+1,k}\varphi_{j+1,k}$ agree on the fine grid, we have

$$2^{1/2}\beta_{j+1,2k+1} = \alpha_{j,k} + 2^{-j/2} \sum_{k'} \beta_{j,k'}\varphi_{j,k'}((k+1/2)/2^j).$$

More interestingly, we can go in the other direction, decomposing any sum of the form (2.9) as a sum (2.8), expressing f in terms of "gross structure" $\sum \beta_{j,k}\varphi_{j,k}$ and "detail corrections" $\sum \alpha_{j,k}\psi_{j,k}$.

Lemma 2.5 *Every $f \in V_{j+1}$ has a representation (2.8), with coefficients*

$$\alpha_{j,k} = 2^{-j/2} \cdot \left(f((k+\frac{1}{2})/2^j) - (P_j f)((k+\frac{1}{2})/2^j) \right) \tag{2.10}$$

This formula shows transparently that the wavelet coefficients $\alpha_{j,k}$ measure lack of approximation of f by $P_j f$.

Iterating this two-scale decomposition, we may express any $f \in V_{j_1}$ as a sum of a coarse-scale description in V_{j_0}, $j_0 < j_1$, and a series of detail corrections:

$$f = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j_0 \le j < j_1} \sum_k \alpha_{j,k}\psi_{j,k}.$$

For more general f, not in V_{j_1}, setting $f = P_{j_1} f + (f - P_{j_1} f)$, and letting j_1 tend to ∞, we get formally (assuming now that $P_j f \to f$ as $j \to \infty$ in some appropriate sense) the *inhomogeneous interpolating wavelet expansion* (2.4). In fact Lemma 2.4 shows that this formal expression makes sense under minimal regularity.

Theorem 2.6 *Consider the interpolating wavelet transform with respect to an (R, D) interpolating wavelet with $R \ge 0$, $D \ge 0$. Let $f \in C_0(\mathbf{R})$. Then the inhomogeneous interpolating expansion (2.4) holds, in the sense of uniform convergence:*

$$\|f - \sum_{|k| \le K} \beta_{j_0,k}\varphi_{j_0,k} - \sum_{j_0 \le j \le j_0+J} \sum_{|k| \le K} \alpha_{j,k}\psi_{j,k}\|_\infty \to 0 \tag{2.11}$$

as $J, K \to \infty$.

Proof. The partial sum operator $P_{J,K}$ implicit in (2.11) is uniformly bounded:

$$\|P_{J,K} f\|_\infty \le C\|f\|_\infty;$$

compare (7.9)-(7.10) in the appendix. The collection of continuous functions of compact support is dense in $C_0(\mathbf{R})$, and so for each $\epsilon > 0$ there is a compactly supported function f'

$$\|f - f'\|_\infty \le \epsilon.$$

Write

$$P_{J,K} f - f = (P_{J,K} f - P_{J,K} f') + (P_{J,K} f' - f') + (f' - f);$$

use the triangle inequality and boundedness of $P_{J,K}$ to get

$$\|P_{J,K} f - f\|_\infty \le C \cdot \epsilon + \|P_{J,K} f' - f'\|_\infty + \epsilon.$$

As f' is compactly supported, there exists K' so that

$$P_{J,K} f' = P_J f', \qquad K \ge K'.$$

By Lemma 2.4 $P_J f' \to f'$ in L^∞; hence

$$\limsup_{J,K \to \infty} \|P_{J,K} f - f\|_\infty \le C \cdot \epsilon.$$

As this is true for each $\epsilon > 0$, we have the desired convergence. □

More generally,if f is the sum of a function in $C_0(\mathbf{R})$ and a polynomial of degree $\le D$, the partial sums converge uniformly on compacts.

We thus have a wavelet decomposition which exhibits the $(\alpha_{j,k})$ explicitly as measures of error in approximation by V_j, and which reconstructs continuous functions.

2.3 Characterizing Smoothness

Despite the elementary nature of the interpolating transform, it yields reasonably strong characterization theorems for Hölder classes C^δ, $\delta > 0$, and for those Besov and Triebel classes embedding into some C^δ. We develop first a result for Besov classes. This implies immediately results also for Hölder (-Zygmund) and L^2-Sobolev classes.

To simplify our proofs in the appendix, it is convenient to assume an extra condition on our interpolating wavelet:

[IW6] *Piecewise Finite Dimensionality.* The collection $\mathcal{F}$ of restrictions $f|_{[0,1]}$ of sums $f = \sum_k \beta_{0,k}\varphi(t-k)$ is finite-dimensional.

Since we are really only interested in Deslauriers-Dubuc wavelets, this assumption is no restriction. Deslauriers-Dubuc wavelets satisfy [IW6] because they are of compact support. Spline interpolating wavelets also satisfy [IW6], because they are piecewise polynomial.

The significance of finite-dimensionality is the following. As the space $\mathcal{F}$ consists of uniformly continuous functions, we have norm equivalences between L^p and L^∞ norms, with constants

$$N_p(\mathcal{F}) = \sup_{f\in\mathcal{F}} \|f\|_{L^\infty[0,1]}/\|f\|_{L^p[0,1]} < \infty, \qquad p \in (0,\infty].$$

These constants, along with the localization constants in [IW5], appear implicitly in the proof of the following theorem.

Theorem 2.7 *Consider an interpolating wavelet transform with a wavelet satisfying [IW1]-[IW6]. Let $min(R,D) > \sigma > 1/p$, $p,q \in (0,\infty]$. Define a norm on the interpolating wavelet coefficients*

$$\theta = ((\beta_{j_0,\cdot}), (\alpha_{j_0,\cdot}), (\alpha_{j_0+1,\cdot}), \ldots)$$

by

$$\|\theta\|_{b^\sigma_{p,q}} = \|(\beta_{j_0,\cdot})\|_{\ell^p} + (\sum_{j\geq j_0}(2^{js}(\sum_k |\alpha_{j,k}|^p)^{1/p})^q)^{1/q}. \tag{2.12}$$

with the calibration $s \equiv \sigma + 1/2 - 1/p$. This is an equivalent norm for the Besov space $B^\sigma_{p,q}(\mathbf{R})$.

Important corollary: if $f \in B^\sigma_{p,q}$ with $\sigma > 1/p$, then the reconstruction of f from its wavelet coefficients converges unconditionally to f in the $B^\sigma_{p,q}$ norm – the order in which the individual terms are summed does not matter. Thus, as an example, if $f \in C^\alpha$, $\alpha \in (0,1)$, and we use interpolating wavelets of regularity ($R \geq 1$, $D \geq 1$), finite interpolating wavelet expansions converge unconditionally in C^α norm. The proof of this remark: simply that the norm of the difference between f and its approximation is equivalent to the sequence space norm of the sequence consisting of those coefficients in which were omitted in forming the approximation.

The limitation $\sigma > 1/p$ is essential rather than technical. Those Besov spaces with $\sigma < 1/p$ do not have point functionals $f(t)$ as bounded linear functionals. Hence a transform based on samples cannot be well-defined over such spaces. Those spaces with $\sigma = 1/p$ are critical cases, which need to be discussed separately (Section 6.1). Those spaces with $\sigma > 1/p$ embed into Hölder classes C^δ for $\delta = \sigma - 1/p$, and hence the interpolating wavelet transform makes sense and converges uniformly as in Theorem 2.2.

A complete proof of Theorem 2.7 is given in the appendix. If we were not interested in working hard to get an "optimal" theorem covering the full range $\sigma > 1/p$, some simple off-the-shelf results could take us part of the way.

It is known that using the wavelet coefficients $\bar\theta$ of a sufficiently nice orthonormal wavelet expansion, the norm $\|\bar\theta\|_{b^\sigma_{pq}}$ is an equivalent norm for Besov space (Meyer, 1990), Frazier, Jawerth, and Weiss (1991). A given function f with orthonormal coefficients $\bar\theta$ also has interpolating wavelet coefficients θ; and the coefficients in the two expansions are related, formally, by operators T and S:

$$\bar\theta = T\theta, \qquad \theta = S\bar\theta.$$

If the operators S and T are both bounded on $b^\sigma_{p,q}$ then the norm $\|\theta\|_{b^\sigma_{pq}}$ is equivalent to $\|\bar\theta\|_{b^\sigma_{pq}}$, and hence to the Besov space norm. Frazier, Jawerth and Weiss (1991, page 59), have defined a notion of almost diagonal operators between sequence spaces b^σ_{pq}. (Compare also the notion $Op(\mathcal{M}_\gamma)$ of Meyer (1990, Vol II) and the applications on pp.334-335 there, and the article of Jaffard (1990)). Almost-diagonal operators are bounded.

Simple calculations reveal that T and S are almost diagonal for the Besov sequence spaces b^σ_{pq}, $\sigma > 1$, $p \geq 1$. Let $\bar\psi$ denote the mother orthogonal wavelet; then $T((j,k),(j',k')) = \int \bar\psi_{j,k}\psi_{j',k'}$. For a sufficiently large $\ell > 0$, $\eta((j,k);(j',k')) = (1 + |2^{-j}k - 2^{-j'}k'|/2^{-\min(j,j')})^\ell$. Then calculations such as we give in the proof of Theorem 2.7 show

$$|T((j,k),(0,0))| = |\int \bar\psi_{j,k}\psi| \leq \begin{cases} C_\ell \cdot 2^{-j/2} 2^{-j\min(\bar D,R)} \eta((j,k);(0,0)) & j \geq 0 \\ C_\ell \cdot 2^{-j/2}\eta((j,k);(0,0)) & j \leq 0 \end{cases}.$$

The decay as $j \to +\infty$ is due to the regularity of ψ and the vanishing moments of $\bar\psi$. The lack of decay as $j \to -\infty$ is due to the lack of vanishing moments for ψ.

For S we have the expression that $S((j,k),(j',k'))$ = the (j,k) interpolating wavelet coefficient of $\bar\psi_{j',k'}$. Hence for some constants (c_k),

$$\begin{aligned} |S((j,k),(0,0))| &= 2^{-j/2} \cdot |\sum_k c_k \bar\psi(k/2^{j+1})| \\ &\leq \begin{cases} C_\ell \cdot 2^{-j/2} 2^{-j\min(D,\bar R)} \eta((j,k);(0,0)) & j \geq 0 \\ C_\ell \cdot 2^{-j/2}\eta((j,k);(0,0)) & j \leq 0 \end{cases}. \end{aligned}$$

Notice again the lack of decay as $j \to -\infty$.

These expressions, and obvious translations and dilations to other scales, give that for small enough $\epsilon > 0$

$$|S((j,k),(j',k'))| \leq C_{\sigma,\epsilon} \cdot 2^{(-1/2-\epsilon)|j'-j|} \cdot 2^{-j\sigma} \cdot \eta((j,k);(j',k'))$$

for any $\sigma \in (1+\epsilon, \min(D,\bar R) - \epsilon)$. From the definition of almost-diagonal operators (e.g. Frazier, Jawerth and Weiss (1991), Page 59), it follows that the operator S is almost diagonal for each $\sigma \in (1+\epsilon, \min(D,\bar R) - \epsilon)$. Similarly, the operator T is almost diagonal for each $\sigma \in (1+\epsilon, \min(\bar D, R) - \epsilon)$.

We stress that these "matrix arguments" are simply heuristic, and to get a full formal proof, covering the full range $\sigma > 1/p$, we use a different approach, based on approximation rates.

2.4 Triebel Smoothness Classes

We now state a result for Triebel classes $F^\sigma_{p,q}$. This implies results for L^p Sobolev classes W^m_p with $1 < p \le \infty$.

Theorem 2.8 *Consider an interpolating wavelet transform based on [IW1]-[IW6]. Let* $min(R, D) > \sigma > 1/p$, $p \in (1, \infty]$, $q \in [1, \infty]$. *Define a norm on the interpolating wavelet coefficients*

$$\theta = ((\beta_{j_0,\cdot}), (\alpha_{j_0,\cdot}), (\alpha_{j_0+1,\cdot}), \ldots)$$

by

$$\|\theta\|_{f^\sigma_{p,q}} \equiv \|\beta_{j_0,k}\|_{\ell_p} + \|(\sum_j 2^{jsq} \sum_k |\alpha_{j,k}|^q \chi_{j,k})^{1/q}\|_{L^p},$$

with the calibration $s \equiv \sigma + 1/2$, *and with* $\chi_{j,k}$ *the indicator function of the interval* $[k/2^j, (k+1)/2^j)$. *This is an equivalent norm to the norm of* $F^\sigma_{p,q}$.

The limitation $\sigma > 1/p$ applies, just as in the Besov case. Because the almost-diagonality concept applies to boundedness on Triebel spaces as well, (see Frazier, Jawerth, and Weiss, (1991) Page 59), the matrix argument given above can be used to get a quick proof for the case $\sigma > 1$, $p, q > 1$.

2.5 Sampling, Interpolation and Smoothing

Suppose that we take samples $(2^{-j_1} f(k/2^{j_1}))_{k \in \mathbf{Z}}$. We may, using just those samples, obtain the interpolating wavelet coefficients of f at all levels up to and including $j_1 - 1$. However, we cannot know the coefficients at any finer scales. Symbolize this truncation process by the operator $\mathcal{T}_{j_1}$; hence, with θ denoting the wavelet coefficients of f, we observe

$$\mathcal{T}_{j_1}\theta = ((\beta_{j_0,\cdot}), (\alpha_{j_0,\cdot}), (\alpha_{j_0+1,\cdot}), \ldots, (\alpha_{j_1-1,\cdot})).$$

Let $\mathcal{E}_{j_1}$ denote the operator that fills in all the unknown wavelet coefficients by **zero** ("zero-extension"), giving a complete array of wavelet coefficients,

$$\mathcal{E}_{j_1}\mathcal{T}_{j_1}\theta = ((\beta_{j_0,\cdot}), (\alpha_{j_0,\cdot}), (\alpha_{j_0+1,\cdot}), \ldots, (\alpha_{j_1-1,\cdot}), \mathbf{0}, \mathbf{0}, \ldots)$$

Because of the zero-extension process, we have

$$\|\mathcal{E}_{j_1}\mathcal{T}_{j_1}\theta\|_{b^s_{p,q}} \le \|\theta\|_{b^s_{p,q}}. \tag{2.13}$$

$$\|\mathcal{E}_{j_1}\mathcal{T}_{j_1}\theta\|_{f^s_{p,q}} \le \|\theta\|_{f^s_{p,q}}. \tag{2.14}$$

Consider now the inverse wavelet transform of $\mathcal{E}_{j_1}\mathcal{T}_{j_1}\theta$. By (2.13)-(2.14) this has smaller Besov and/or Triebel norms than f. This object has an alternate description. It is just $P_{j_1} f$. Hence, it is an interpolation operator: it takes the samples $(f(k/2^{j_1}))_{k \in \mathbf{Z}}$ and produces a function defined for all $t \in \mathbf{R}$. If the interpolating wavelet was a fundamental spline, this is a spline interpolation. If the interpolating wavelet was a Deslauriers-Dubuc fundamental function, this is a Deslauriers-Dubuc interpolation. With these re-interpretations, (2.13)-(2.14) say the following.

Corollary 2.9 *Consider an interpolation process, such as cardinal spline interpolation or Deslauriers-Dubuc interpolation, which can be interpreted as being of the form $P_j f$ for an (R, D) interpolating wavelet φ. There are equivalent norms for the function spaces $B^\sigma_{p,q}$ and $F^\sigma_{p,q}$ with $1/p < \sigma < \min(R, D)$ such that sampling followed by interpolation is norm-reducing on all those spaces.*

3 Interpolating Transform on the Interval

We now develop interpolating wavelet transforms for $C[0,1]$. These are based on adapting the inhomogeneous interpolating transforms for $\Pi_D + C_0(\mathbf{R})$ to "life on the interval".

We concentrate exclusively on Interpolating Wavelets of Compact Support. So in this section φ and ψ derive from the Deslauriers-Dubuc fundamental wavelet F_D. We will often write K_j as shorthand for $\{0 \leq k < 2^j\}$.

Theorem 3.1 *Let j_0 be a non-negative integer satisfying $2^{j_0} > 2D + 2$ (non-interacting boundaries). There exists a collection of functions $\varphi^{[]}_{j,k}$ and $\psi^{[]}_{j,k}$ such that every $f \in C[0,1]$ has a representation*

$$f = \sum_{k=0}^{2^{j_0}-1} \beta_{j,k} \varphi^{[]}_{j_0,k} + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \alpha_{j,k} \psi^{[]}_{j,k},$$

with uniform convergence of partial sums $j \leq j_1$ as $j_1 \to \infty$. Here the wavelet coefficients are given by

$$\beta_{j,k} = 2^{-j/2} f(k/2^j), \qquad k \in K_j,$$

and, with $P^{[0,1]}_j f = \sum_{k=0}^{2^j-1} \beta_{j,k} \varphi^{[]}_{j,k}$,

$$\alpha_{j,k} = 2^{-j/2} \left(f((k+1/2)/2^j) - (P^{[0,1]}_j f)((k+1/2)/2^j) \right).$$

The functions $\varphi^{[]}_{j,k}$ and $\psi^{[]}_{j,k}$ are interpolating:

$$\varphi^{[]}_{j,k}(k' 2^{-j}) = 2^{j/2} \delta_{k,k'}, \qquad k' \in K_j$$

$$\psi^{[]}_{j,k}(k' 2^{-j-1}) = 2^{j/2} \delta_{2k+1,k'}, \qquad k' \in K_{j+1}$$

The $\varphi^{[]}_{j,k}$ derive from the interpolating wavelets $\varphi_{j,k}$ for the line as follows. At the "heart" of the interval, they are the same:

$$\varphi^{[]}_{j,k} = \varphi_{j,k}|_{[0,1]}, \qquad D < k < 2^j - D - 1$$

and at the "edges" of the interval, they are dilations of certain special boundary-adjusted wavelets:

$$\varphi^{[]}_{j,k}(x) = 2^{j/2} \varphi^{\#}_k (2^j x - k), \qquad 0 \leq k \leq D,$$

$$\varphi^{[]}_{j,2^j-k-1}(x) = 2^{j/2}\varphi^{\flat}_k(2^j x - 2^j - k - 1), \qquad 0 \le k \le D.$$

The boundary-adjusted wavelets $\varphi^{\#}_k$ *and* $\varphi^{\flat}_k$ *are compactly supported and exhibit the same degree of regularity as the corresponding wavelets on the line. The functions* $\psi^{[]}_{j,k}$ *derive from the wavelets on the line* $\psi_{j,k}$ *as follows. At the "heart" of the interval, they are the same:*

$$\psi^{[]}_{j,k} = \psi_{j,k}|_{[0,1]}, \qquad \lfloor D/2 \rfloor \le k < 2^j - \lfloor D/2 \rfloor$$

and at the "edges" of the interval, they are dilations of certain special boundary wavelets:

$$\psi^{[]}_{j,k} = 2^{j/2}\psi^{\#}_k(2^j x - k), \qquad 0 \le k < \lfloor D/2 \rfloor,$$

$$\varphi^{[]}_{j,2^j-k-1} = 2^{j/2}\psi^{\flat}_k(2^j x - 2^j - k - 1), \qquad 0 \le k < \lfloor D/2 \rfloor.$$

The boundary-adjusted wavelets $\psi^{\#}_k$ *and* $\psi^{\flat}_k$ *are compactly supported and exhibit the same degree of regularity as the corresponding wavelets on the line.*

It is interesting to compare this interpolating wavelet transform with the orthogonal transforms of Meyer (1991) and [CDJV]. It has a similar formal structure – all the basis functions are dilations and translations of functions in a finite list; functions in that list are all built by boundary correcting the wavelets for the line.

3.1 Construction of the Basis

We now prove Theorem 3.1 by an extended discussion. Suppose we are given normalized samples $(\beta_{j,k} = 2^{-j/2} f(k/2^j) : 0 \le k < 2^j)$ only for lattice sites in the unit interval $[0,1)$, and that $2^j > 2D+2$ ("non-interacting boundaries"). The *Extension operator* $\mathcal{E}_{j,D}$ of degree D generates an extrapolation of these samples to a bilateral infinite sequence $(\tilde\beta_{j,k})_{k\in\mathbf{Z}}$ which agrees with $\beta_{j,k}$ for $0 \le k < 2^j$. The extrapolated samples $(\tilde\beta_{j,k} : k < 0)$ are filled in by fitting a polynomial $\pi^{\#}_j$ of degree D to the leftmost $D+1$ samples:

$$\pi^{\#}_j(k2^{-j}) = f(k/2^j), \quad 0 \le k \le D$$

and using this to extrapolate to $k < 0$:

$$\tilde\beta_{j,k} = 2^{-j/2}\pi^{\#}_j(k2^{-j}), \qquad k < 0.$$

Similarly, the missing samples $(\tilde\beta_{j,k} : k \ge 2^j)$ are obtained by fitting a separate polynomial $\pi^{\flat}_j$ of degree D to the rightmost $D+1$ samples:

$$\pi^{\flat}_j(k2^{-j}) = f(k/2^j), \quad 2^j - 1 - D \le k \le 2^j - 1;$$

and using this to extrapolate to $k \ge 2^j$:

$$\tilde\beta_{j,k} = 2^{-j/2}\pi^{\flat}_j(k2^{-j}), \qquad k \ge 2^j.$$

One then has data at every lattice site and can form the interpolate

$$\tilde f = \sum_{k=\infty}^{\infty} \tilde\beta_{j,k}\varphi_{j,k}. \tag{3.1}$$

Definition 3.2 *$V_j[0,1]$ is the collection of (restrictions to [0,1] of) functions $\tilde{f}$ defined by (3.1), where $\tilde{\beta} = \mathcal{E}_{j,D} \circ (\beta_{j,k})_{k=0}^{2^j-1}$.*

This gives us a 2^j-dimensional vector space of functions with two key properties. First, the coefficients arise by sampling:

$$\beta_{j,k} = \tilde{f}(k2^{-j})/2^{j/2}, \qquad 0 \le k < 2^j.$$

This property is simply inherited from the interpolating transform on $\mathbf{R}$. Second:

$$\Pi_D \subset V_j[0,1]$$

Indeed, if π is a polynomial of degree $\le D$, then $\pi_j^{\#} = \pi_j^{\flat} = \pi$. Hence the extrapolated samples satisfy

$$\tilde{\beta}_{j,k} = 2^{-j/2}\pi(k2^{-j}), \qquad k \in \mathbf{Z},$$

and so we may use the reproducing formula for the transform on $\mathbf{R}$,

$$\pi = 2^{-j/2} \sum_{k=-\infty}^{\infty} \pi(k/2^j)\varphi_{j,k} = \sum_{k=-\infty}^{\infty} \tilde{\beta}_{j,k}\varphi_{j,k}$$

which exhibits π as an element of $V_j[0,1]$.

We now seek a representation of the finite-dimensional space $V_j[0,1]$ involving finite sums. Each extrapolated coefficient $\tilde{\beta}_{j,k}$ with $k < 0$ is a linear functional of the extrapolating polynomial and hence of the left-end samples ($\beta_{j,k} : 0 \le k \le D$). It follows that we may define *extrapolation weights* $e^{\#}_{k,k'}$ such that

$$\tilde{\beta}_{j,k} = \sum_{k'=0}^{D} e^{\#}_{k,k'}\beta_{j,k'}, \qquad k < 0.$$

Similarly, for $k \ge 2^j$ we have weights $e^{\flat}_{k,k'}$

$$\tilde{\beta}_{j,k} = \sum_{k'=2^j-D-1}^{2^j-1} e^{\flat}_{k,k'}\beta_{j,k'}, \qquad k \ge 2^j.$$

Now define *boundary-wavelets*

$$\varphi^{\#}_{j,k'} = \varphi_{j,k'} + (\sum_{k<0} e^{\#}_{k,k'}\varphi_{j,k}), \qquad 0 \le k' \le D$$

and

$$\varphi^{\flat}_{j,k'} = \varphi_{j,k'} + (\sum_{k \ge 2^j} e^{\flat}_{k,k'}\varphi_{j,k}), \qquad 2^j - D - 1 \le k' < 2^j.$$

These sums are actually finite because of the compact support of the fundamental wavelets. Moreover, the new boundary wavelets are interpolating. To see this, note that the non-boundary corrected $\varphi_{j,k}$ are interpolating, and the correction terms to $\varphi_{j,k}$ are sums of $\varphi_{j,k'}$, $k' \notin \{0, \ldots, 2^j - 1\}$ which vanish identically on the dyadic grid $\{0, 2^{-j}, \ldots, (2^j-1)/2^j\}$.

We therefore have, by rearranging the sum (3.1)

Lemma 3.3 *Every $f \in V_j[0,1]$ has the representation*

$$\tilde{f} = \sum_{k=0}^{D} \beta_{j,k} \varphi^{\#}_{j,k} \quad + \sum_{k=D+1}^{2^j-D-2} \beta_{j,k}\varphi_{j,k} \quad + \sum_{k=2^j-D-1}^{2^j-1} \beta_{j,k}\varphi^{\flat}_{j,k};$$

where again, the coefficients

$$\beta_{j,k} = 2^{-j/2} f(k/2^j).$$

Two crucial but trivial comments. First, the boundary wavelets are dilation-homogeneous:

$$\begin{aligned} \varphi^{\#}_{j+\ell,k}(t) &= 2^{\ell/2}\varphi^{\#}_{j,k}(2^\ell t), \qquad 0 \le k \le D \\ \varphi^{\#}_{j+\ell,2^{j+\ell}-k} &= 2^{\ell/2}\varphi^{\#}_{j,2^j-k}(2^\ell t), \qquad 0 \le k \le D. \end{aligned}$$

This follows immediately from the observation that the edge weights $e^{\#}_{k,k'}$ and $e^{\flat}_{k,k'}$ do not depend on j.

Secondly, the boundary wavelets $\varphi^{\#}_{j,k}$ are compactly supported, with support width $\le C\,2^{-j}$; indeed, the corresponding uncorrected $\varphi_{j,k}$ have this property; the sum of corrections has a bounded number of terms each of which has this property.

To finish up our analysis of $V_j[0,1]$ we record

$$\varphi^{[]}_{j,k} = \begin{cases} \varphi^{\#}_{j,k} & 0 \le k \le D \\ \varphi_{j,k} & D < k < 2^j - D - 1 \\ \varphi^{\flat}_{j,k} & 2^j - D - 1 \le k < 2^j \end{cases}$$

Our discussion so far proves all the properties of $\varphi^{[]}_{j,k}$ claimed in the theorem.

We now turn to the analysis of the detail spaces $W_j[0,1]$. To define these we first need a key observation:

Lemma 3.4

$$V_j[0,1] \subset V_{j+1}[0,1]$$

Proof. The result amounts to the assertion that

$$P^{[0,1]}_{j+1} f = f, \qquad f \in V_j[0,1].$$

The key point is that for an $f \in V_j[0,1]$, the edge polynomials at the next finer scale agree:

$$\pi^{\#}_{j+1} = \pi^{\#}_j, \qquad \pi^{\flat}_{j+1} = \pi^{\flat}_j. \tag{3.2}$$

To see this, note that the edge polynomial $\pi^{\#}_j$ is defined by the property

$$f(k/2^j) = \pi^{\#}_j(k/2^j), \qquad 0 \le k \le D. \tag{3.3}$$

However, when we use it to extend to samples at negative k, and then apply the Deslauriers-Dubuc interpolation scheme to the extended samples to get halfway points, we get agreement at halfway points as well:

$$f((k+1/2)/2^j) = \pi_j^\#((k+1/2)/2^j), \qquad 0 \le k \le (D-1)/2. \tag{3.4}$$

For example, $f((0+1/2)/2^j)$ is produced by fitting a polynomial to samples $(f(k/2^j) : (-D+1)/2 \le k \le (D+1)/2)$. But $\pi_j^\#$ passes through all D of these points, and is the only polynomial that does so. Hence, when fitting a polynomial $\pi_{j,0}$ to these points we get $\pi_{j,0} = \pi_j^\#$, and the Deslauriers-Dubuc interpolation rule $f((0+1/2)/2^j) = \pi_{j,0}((0+1/2)/2^j)$ gives (3.4). Combining (3.3) and (3.4) we conclude

$$f(k/2^{j+1}) = \pi_j^\#(k/2^{j+1}), \qquad 0 \le k \le D. \tag{3.5}$$

But $\pi_{j+1}^\#$ is defined by the property that it interpolates the points $f(k/2^{j+1})$, $0 \le k \le D$. And only one polynomial can solve this interpolation. Hence (3.5) implies (3.2). $\square$

Now define $Q_j^{[0,1]} = P_{j+1}^{[0,1]} - P_j^{[0,1]}$ and set $W_j[0,1] = Range(Q_j^{[0,1]})$. Because of the lemma we have just proved, this is a 2^j-dimensional function space.

We now seek an accessible representation of elements of $W_j[0,1]$. Define $\tilde{\alpha}_{j,k}$ by, if $0 \le k < 2^j$,

$$\tilde{\alpha}_{j,k} = 2^{-j/2} \cdot (f((k+\frac{1}{2})/2^j) - (P_j^{[0,1]}f)((k+\frac{1}{2})/2^j)),$$

and 0 for other values of k. The sum $\sum_{k=-\infty}^{\infty} \tilde{\beta}_{j,k}\varphi_{j,k} + \sum_{k=-\infty}^{\infty} \tilde{\alpha}_{j,k}\psi_{j,k}$ belongs to $V_{j+1}(-\infty,\infty)$, and so there is a sequence $(\bar{\beta}_{j+1,k})_k$ for which

$$\sum_{k=-\infty}^{\infty} \tilde{\beta}_{j,k}\varphi_{j,k} + \sum_{k=-\infty}^{\infty} \tilde{\alpha}_{j,k}\psi_{j,k} = \sum_{k=-\infty}^{\infty} \bar{\beta}_{j+1,k}\varphi_{j+1,k}.$$

We calculate that

$$\bar{\beta}_{j+1,k} = \begin{cases} 2^{-(j+1)/2} f(k2^{-j-1}) & k \in K_{j+1} \\ 2^{-(j+1)/2} \pi_j^\#(k2^{-j-1}) & k < 0 \\ 2^{-(j+1)/2} \pi_j^\flat(k2^{-j-1}) & k \ge 2^{j+1} \end{cases}.$$

Indeed, for each $k < 0$ the $D+1$ lattice sites $k'/2^j$ nearest to $(k+1/2)/2^j$ have values $\tilde{\beta}_{j,k'}$ which are reproduced exactly by $\pi_j^\#$. Therefore, fitting a polynomial $\pi_{j,k}$ to those values only reproduces the polynomial $\pi_j^\#$.

Let, as before, $\tilde{\beta}_{j+1,k}$ denote the polynomial extension of the values with $k \in K_{j+1}$. Then for $k < 0$,

$$\tilde{\beta}_{j+1,k} - \bar{\beta}_{j+1,k} = \pi_{j+1}^\#(k2^{-j-1}) - \pi_j^\#(k2^{-j-1}),$$

and similarly for $k \ge 2^j$. Hence we may write

$$\sum_{k=-\infty}^{\infty} \tilde{\beta}_{j,k}\varphi_{j,k} + \sum_{k=0}^{2^j-1} \tilde{\alpha}_{j,k}\psi_{j,k} + A^\#(t) + A^\flat(t) = \sum_{k=-\infty}^{\infty} \tilde{\beta}_{j+1,k}\varphi_{j+1,k}$$

where, for example,

$$A^\#(t) = \sum_{k=-\infty}^{-1} (\pi_{j+1}^\#(k2^{-j-1}) - \pi_j^\#(k2^{-j-1}))\varphi_{j+1,k}(t).$$

We analyze the $A^{\#}$ term closely. The polynomials $\pi_j^{\#}(t)$ and $\pi_{j+1}^{\#}(t)$ necessarily agree for values $t = k/2^j$, $0 \leq k \leq (D+1)/2$. Hence the space of difference polynomials $\pi_{j+1}^{\#}(t) - \pi_j^{\#}(t)$ is $\lfloor D/2 \rfloor$ dimensional. A basis for this space is $\{p_{j,k}^{\#}\}$, where $p_{j,k}^{\#}(t)$ is a polynomial satisfying the interpolation conditions

$$p_{j,k}^{\#}(t) = \begin{cases} 1 & t = (2k+1)/2^{j+1}, \\ 0 & t = k'/2^{j+1}, 0 \leq k' \leq D, k' \neq (2k+1) \end{cases}.$$

Then

$$A^{\#}(t) = \sum_{k=0}^{\lfloor D/2 \rfloor} \alpha_{j,k}^{\#} p_{j,k}^{\#}(t)$$

where

$$\alpha_{j,k}^{\#} = f((k+1/2)/2^j) - \pi_j^{\#}((k+1/2)/2^j).$$

We now define boundary wavelets $\psi_{j,k}^{\#}(t)$, for $k = 0, \ldots, \lfloor D/2 \rfloor - 1$ (if $D = 1$ just skip what follows).

$$\psi_{j,k}^{\#}(t) = \psi_{j,k}(t) + \sum_{k'<0} p_{j,k}^{\#}(k'/2^{j+1}) \varphi_{j+1,k'}(t).$$

These sums are actually finite because the wavelets ψ are compactly supported. These new functions vanish at all points of the dyadic grid $\{k'/2^{j+1}\}$ where $k' \in K_{j+1}$ except at $k' = 2k+1$. Indeed, the uncorrected $\psi_{j,k}$ has this property, and the correction terms vanish at all points of the dyadic grid $k' \in K_j$. These new functions have support width $\leq C2^{-j}$ – because they are finite sums of functions with such properties.

Because (pleasant surprise!) $\alpha_{j,k}^{\#} = \tilde{\alpha}_{j,k}$, and similarly for $\alpha_{j,k}^{\flat}$, we may group together all terms containing $\alpha_{j,k}^{\#}$ and the corresponding $\tilde{\alpha}_{j,k}$ and conclude that

$$\begin{aligned} \sum_{k=0}^{2^j-1} \tilde{\alpha}_{j,k}\psi_{j,k} + A^{\#}(t) + A^{\flat}(t) &= \sum_{k=0}^{\lfloor D/2 \rfloor - 1} \tilde{\alpha}_{j,k}\psi_{j,k}^{\#}(t) \\ &+ \sum_{k=\lfloor D/2 \rfloor}^{2^j - \lfloor D/2 \rfloor - 1} \tilde{\alpha}_{j,k}\psi_{j,k}(t) + \sum_{k=2^j-\lfloor D/2 \rfloor}^{2^j-1} \tilde{\alpha}_{j,k}\psi_{j,k}^{\flat}(t) \end{aligned}$$

If we now define

$$\psi_{j,k}^{[]} = \begin{cases} \psi_{j,k}^{\#} & 0 \leq k < \lfloor D/2 \rfloor \\ \psi_{j,k} & \lfloor D/2 \rfloor \leq k < 2^j - \lfloor D/2 \rfloor \\ \psi_{j,k}^{\flat} & 2^j - \lfloor D/2 \rfloor \leq k < 2^j \end{cases},$$

the above discussion proves

Lemma 3.5 *Every $f \in V_{j+1}[0,1]$ can be written*

$$f = \sum_{k=0}^{2^j-1} \beta_{j,k}\varphi_{j,k}^{[]} + \sum_{k=0}^{2^j-1} \alpha_{j,k}\psi_{j,k}^{[]},$$

where

$$\beta_{j,k} = 2^{-j/2} f(k/2^j), \qquad k \in K_j$$

$$\alpha_{j,k} = 2^{-j/2} \cdot \left(f((k+1/2)/2^j) - (P_j^{[0,1]} f)((k+1/2)/2^j) \right), \qquad k \in K_j.$$

As in the case of the transform on the line, we can iterate this decomposition, writing every element of $V_{j_1}[0,1]$ as a sum of a term in $V_{j_0}[0,1]$, $j_0 < j_1$, and intervening terms from $W_j[0,1]$, $j_0 \le j < j_1$:

$$f = \sum_k \beta_{j_0,k}\varphi^{[]}_{j_0,k} + \sum_{j_0 \le j < j_1} \sum_k \alpha_{j,k}\psi^{[]}_{j,k}$$

More generally, if f is not in V_{j_1}, we have that f almost obeys this decomposition, with an error $f - P_{j_1} f$. Now if $||f - P^{[0,1]}_j f||_\infty \to 0$ as $j \to \infty$, we can rigorously justify an infinite sum decomposition. But $||f - P^{[0,1]}_j f||_\infty \to 0$ if f is uniformly continuous. Hence this decomposition holds for every $f \in C[0,1]$.

3.2 Smoothness Classes on the Interval

The interpolating wavelet transform on the interval gives a characterization of Besov classes on the interval.

Theorem 3.6 *Let θ denote the wavelet coefficients for the interpolating wavelet transform on the interval. Then*

$$\|\theta\|_{b^\sigma_{p,q}} \equiv \|(\beta_{j_0,\cdot})\|_{\ell^p} + (\sum_{j \ge j_0} (2^{js}(\sum_{k=0}^{2^j-1} |\alpha_{j,k}|^p)^{1/p})^q)^{1/q},$$

where $s = \sigma + 1/2 - 1/p$, gives an equivalent norm for the Besov spaces $B^\sigma_{p,q}[0,1]$ where $1/p < \sigma < \min(R,D)$, $p,q \in (0,\infty]$.

The proof for the case of the line has been constructed so that it will continue to work here, with only slight modifications. See the appendix.

Triebel-Lizorkin equivalence is also proved in the appendix.

Theorem 3.7 *Let θ denote the wavelet coefficients for an interpolating wavelet transform on the interval built from the Deslauriers-Dubuc fundamental wavelet F_D. Then*

$$\|\theta\|_{f^\sigma_{p,q}} \equiv \|(\beta_{j_0,\cdot})\|_{\ell^p} + \|(\sum_{j \ge j_0} 2^{jsq} \sum_{k=0}^{2^j-1} |\alpha_{j,k}|^q \chi_{j,k})^{1/q}\|_p$$

where $s = \sigma + 1/2$, gives an equivalent norm for the Triebel spaces $F^\sigma_{p,q}[0,1]$ where $1/p < \sigma < \min(R,D)$ and $p,q \in (1,\infty]$.

3.3 Fast Compression of Interpolating Expansions

This interpolating wavelet transform for the interval has a very simple algorithmic structure. The calculation of $(P^{[0,1]}_j f)((k+1/2)/2^j)$ requires only $D+1$ multiplications and additions. The calculation of $\alpha_{j,k}$ requires only $D+2$. Moreover the multiplications involve rational coefficients with small denominators.

The interpolating transform for the interval gives a sparse representation of certain functions – particularly so for piecewise polynomials: *If* f *is a piecewise polynomial on* $[0,1]$*, with* $\le P$ *pieces, each of degree* D*, sampled on a dyadic grid of* $n = 2^{j_1}$ *points, then from those samples, we can calculate the wavelet coefficients at scale* $j_1 - 1$ *and each coarser scale, and there are only* $C_0 + P \cdot (D+1) \cdot \log_2(n)$ *nonzero wavelet coefficients among them.* Since a piecewise polynomial has in general $(D+1)\cdot P$ parameters, this behavior is, in a natural sense, within a logarithmic factor of optimal.

Of course, general functions have many nonzero wavelet coefficients. But often these can be well approximated by a sparse sequence. Suppose that for we have the collection of all interpolating wavelet coefficients θ and that we "sparsify" these as follows.

Let $\epsilon > 0$ denote a thresholding control parameter. Define gross-structure coefficients $\beta^{(\epsilon)}_{j_0,k} = \beta_{j_0,k}$, $k \in K_{j_0}$, and coefficients

$$\alpha^{(\epsilon)}_{j,k} = \alpha_{j,k} \cdot 1_{\{|\alpha_{j,k}| > \epsilon \cdot 2^{-j/2}\}}, \qquad j \ge j_0,\ k \in K_j.$$

Hence, at each level $j \ge j_0$, we set to zero coefficients which are smaller in amplitude than $\epsilon \cdot 2^{-j/2}$.

Suppose that $f \in B^{\sigma}_{p,q}[0,1]$ for some $\sigma > 1/p$. Then $\theta^{(\epsilon)}$ has finitely many nonzero terms. The series obtained by summing the wavelet series corresponding to $\theta^{(\epsilon)}$ produces a reconstruction $f^{(\epsilon)}$; $f^{(\epsilon)} \to f$ in $B^{\sigma}_{p,q}[0,1]$ as $\epsilon \to 0$.

The sparse representation $f^{(\epsilon)}$ has advantages over another finite series, namely $P^{[0,1]}_j f$. The slogan is that $f^{(\epsilon)}$ contains the terms which are important, while $P^{[0,1]}_j f$ contains all terms which might possibly be important. This can be seen in the case of piecewise polynomials, where there are only order $J \cdot P$ nonzero wavelet coefficients up to and including level J of the wavelet expansion, but they are scattered around, with some present at each resolution level. In order to perfectly reconstruct the piecewise polynomial on a grid of size $n = 2^J$ points by a non-adaptive levelwise scheme $P^{[0,1]}_j$ $j_0 \le j \le J$ we have to capture all the levels where nonzero coefficients occur (i.e. all the levels) and so we would use the full representation $P^{[0,1]}_J f$ which requires order 2^J storage. In contrast $f^{(\epsilon)}$ uses only order $J \cdot P$ storage, and this is within a logarithmic factor of optimal. For more general classes of functions, we have the following result.

Theorem 3.8 *Suppose* $f \in B^{\sigma}_{p,q}[0,1]$ *with* $D > \sigma > 1/p$*. Then we have*

$$\|f - f^{(\epsilon)}\|_\infty \le \epsilon \cdot \left(c_0(\sigma,p) + c_1(\sigma,p) \log_2(\|\theta\|_{b^{\sigma}_{p,q}}/\epsilon) \right) \qquad \epsilon > 0 \tag{3.6}$$

while the number of nonzero terms

$$N(\epsilon) = \#\{(j,k) : \alpha^{(\epsilon)}_{j,k} \ne 0\}$$

satisfies

$$N(\epsilon) \le c_2(\sigma,p) \cdot \epsilon^{-p} \cdot \|\theta\|^p_{b^{\sigma}_{p,q}}. \tag{3.7}$$

In words, $f^{(\epsilon)}$ is roughly ϵ away from f; and has roughly ϵ^{-p} nonzero coefficients.

For better insight, we rewrite the result in terms of n rather than ϵ. If $\sigma > 1/p$ then, for each $\eta > 0$ we can pick $\tilde{p} < p$ so that $\sigma - \eta < 1/\tilde{p} < \sigma$. Using $\|\theta\|_{b^\sigma_{\tilde{p},\infty}} \le \|\theta\|_{b^\sigma_{p,q}}$, and applying Theorem 3.8, we can define a sequence ϵ_n so that $N(\epsilon_n) \le n$ and so that

$$\|f - f^{(\epsilon_n)}\|_\infty \le C(\sigma, p, \eta, \|f\|_{B^\sigma_{p,q}[0,1]}) \cdot n^{-(\sigma-\eta)}, \qquad n \to \infty. \tag{3.8}$$

This is near optimal in a certain *minimax compression model.* Let $\mathcal{F}$ be a library of functions, any of which we would like to repesent efficiently. Let $\mathcal{S}_n$ be the set of all interpolating wavelet sums with at most n nonzero coefficients (all possible positions of the n nonzero coefficients being allowed). The best we can hope to do, taking worst-case error over $\mathcal{F}$ as a criterion, is the *minimax error*:

$$\sup_{f \in \mathcal{F}} \inf_{f_n \in \mathcal{S}_n} \|f - f_n\|_\infty .$$

Suppose that our library of functions to be compressed is a Besov ball. In the appendix we prove the following lower bound on the minimax error:

Lemma 3.9 *Let $\mathcal{F}$ be the Besov ball of all functions f with $B^\sigma_{p,q}[0,1]$ norm bounded by B, $\sigma > 1/p$*

$$\sup_{f \in \mathcal{F}} \inf_{f_n \in \mathcal{S}_n} \|f - f_n\|_\infty \ge C(\sigma, p, q) \cdot B \cdot n^{-\sigma} \tag{3.9}$$

Hence no essentially better rate of approximation than $n^{-\sigma}$ is possible, by any method of sparsifying the representation of f to n terms, no matter how nonlinear or computationally intensive the scheme might be. The same type of lower bound can be developed for sparsification of other wavelet expansions, including orthogonal wavelets. Comparing (3.8) to (3.9), we see that simple thresholding of interpolating coefficients is nearly as good, in a minimax sense, as some computationally more extravagant, perhaps yet to be invented, sparsification of orthogonal coefficients.

How does $f^{(\epsilon_n)}$ compare with $P_J f$ as an approximation scheme? For functions in $B^\sigma_{p,q}[0,1]$ we have

$$\begin{aligned}
\|f - P_J f\|_\infty &\le \sum_{j \ge J} \|Q_j f\|_\infty \\
&\le \sum_{j \ge J} \|(\alpha_{j,k})_k\|_{\ell^\infty} 2^{j/2} \cdot C \\
&\le \sum_{j \ge J} 2^{-j(\sigma+1/2-1/p)} \|\theta\|_{b^\sigma_{p,q}} 2^{j/2} \cdot C \\
&\le 2^{-J(\sigma-1/p)} \cdot \|\theta\|_{b^\sigma_{p,q}} \cdot C
\end{aligned}$$

Moreover there is a lower bound indicating that this upper bound is sharp, to within constant factors. Now P_J has $n = 2^J$ nonzero terms, and so we have

$$\|f - P_J f\|_\infty \le C \cdot \|f\|_{B^\sigma_{p,q}[0,1]} \cdot n^{-(\sigma-1/p)}$$

Except in the case $p = \infty$ this rate is slower than the rate $n^{-\sigma}$ nearly attained in (3.8). Hence P_J has a slower minimax rate of approximation than $f^{(\epsilon_n)}$ in general; it gives worse reconstructions for a given number of nonzero terms. This is a generalization of what we saw already for approximations to piecewise polynomials. For an example of a function where the two methods differ, we need an $f \in B^\sigma_{p,q}[0,1]$ but $f \notin B^\sigma_{\infty,\infty}[0,1]$. An example is the cusp function $\exp(-|x-1/2|)$ which is piecewise C^∞ but which does not belong to $C^{1+\delta}$. Except for a constant number of terms at each resolution level which "feel" the cusp, and which decay like $2^{-j(3/2)}$, the wavelet coefficients decay like $2^{-j(1/2+D)}$. The function f is in $B^\sigma_{p,q}[0,1]$ whenever $\sigma - 1/p < 1$, and so, the rate of approximation by $f^{(\epsilon_n)}$ is at least $n^{-D+\eta}$, $\eta > 0$. In contrast, f is in $B^\sigma_{\infty,\infty}[0,1]$ only if $\sigma \leq 1$, and the rate of approximation by $P_J f$ is only n^{-1}. Roughly speaking, the difference between the schemes is that $f^{(\epsilon_n)}$ adaptively spends its budget of n terms predominantly near the cusp, while $P_J f$ uses its n terms in a non-adaptive way.

Optimal results for minimax compression were obtained by DeVore, Jawerth, and Lucier (1992), who developed quasi-interpolating wavelet expansions with sparsification schemes giving functions f_n with n nonzero components and

$$\|f - f_n\|_\infty \leq C(\sigma, p) \cdot \|f\|_{B^\sigma_{p,p}[0,1]} \cdot n^{-\sigma}.$$

The method of DeVore, Jawerth, and Lucier is more complicated than simple thresholding; in particular, it is not based on applying thresholds in a completely parallel fashion. Whether a certain coefficient is set to zero by their method depends on what happens to as many as $O(\log(n))$ other coefficients. On the other hand, the reconstructions they give are very nice.

Remark: although the minimax compression model provides theoretical motivation for the use of simple thresholding schemes, one can easily question the choice of Besov balls $\mathcal{F}$ as models for empirical data. These balls model certain properties of interest, but actual images contain spatial correlations and inter-level correlations which are nowhere captured by the Besov formalism. Compression schemes which exploit correlations between resolution levels and spatial locations might turn out to give dramatically better reconstructions than simple thresholding. Our aim here is only to point out that interpolating wavelet transforms are equally as amenable to simple thresholding as orthogonal wavelet transforms.

Another interpretation of compression is uniform *quantization*. We define a quantum $q > 0$ and obtain q-quantized coefficients $\theta^{\{q\}}$ via

$$\alpha^{\{q\}}_{j,k} = q_j \cdot Round(\alpha_{j,k}/q_j)$$

where $q_j = q \cdot 2^{-j/2}$ and so on. The arguments used to study thresholding will also show that, if f has regularity, the bulk of these coefficients are zero, and the reconstruction $f^{\{q\}}$ is near-optimal in terms of minimax closeness as a function of number of bits used.

Such a quantization procedure, while having near-optimal L^∞ reconstruction errors, is completely parallelizable, meaning that each wavelet coefficient and its quantization may calculated independently of all others. The whole procedure may be calculated in $D+3$ parallel arithmetic operations. Moreover, the arithmetic required is all fixed-point integer arithmetic.

4 A Hybrid Transform

We now construct a hybrid transform, half-way between an orthonormal wavelets basis and an interpolating basis.

4.1 Hybrid Wavelets

Let $\bar{\varphi}_{j,k}$ and $\bar{\psi}_{j,k}$ denote the wavelets of compact support associated with a pyramidal filtering scheme based on quadrature mirror filters of finite length. Suppose that these wavelets have $C^{\bar{R}}$ regularity and that the $\bar{\psi}$ are orthogonal to all polynomials of degree $\bar{D}$.

Definition 4.1 *We say that the interpolating wavelet φ is* **linked** *to the orthogonal wavelet $\bar{\varphi}$ if φ is the autocorrelation $\varphi = \bar{\varphi}(-\cdot) \star \bar{\varphi}(\cdot)$ of $\bar{\varphi}$.*

Below we always assume that our interpolating wavelets are linked to the orthogonal wavelets. As our orthogonal wavelets are Daubechies wavelets we are therefore restricting ourselves to Deslauriers-Dubuc interpolating wavelets. And also to $D \geq \bar{D}$, $R \geq \bar{R}$.

We now construct *hybrid* wavelets "in between" these two particular systems. Let j_1 be a fixed integer, and consider the normalized samples

$$\beta_{j_1,k} = 2^{-j_1/2} f(k/2^{j_1}).$$

Now consider applying the pyramidal algorithm U_{j_0,j_1} to $(\beta_{j_1,k})$ with quadrature mirror filters of compact support corresponding to the wavelets $\bar{\varphi}$. This results in a finite collection of infinite sequences

$$\tilde{\theta} = \Big((\tilde{\beta}_{j_0,\cdot}), (\tilde{\alpha}_{j_0,\cdot}), \ldots, (\tilde{\alpha}_{j_0+1,\cdot})\Big)$$

obtained as follows. There are orthogonal transformations $H, L : \ell^2(\mathbf{Z}) \to \ell^2(2\mathbf{Z})$ which consist of convolution by filters of finite length composed with a factor of two downsampling. Applying these gives

$$(\tilde{\alpha}_{j,\cdot}) = H \circ L^{j-j_1-1} \circ (\beta_{j_1,\cdot}), \qquad j_0 \leq j < j_1,$$

$$(\tilde{\beta}_{j_0,\cdot}) = L^{j_0-j_1} \circ (\beta_{j_1,\cdot}).$$

These coefficients, which simply arise from filtering normalized samples, are the coefficients of f in an expansion with respect to modified functions $\tilde{\varphi}_{j_0,k}$, $\tilde{\psi}_{j,k}$ $j_0 \leq j < j_1$, defined as follows. Let $\theta^{(j_0,k')}$ be the sequence

$$((\tilde{\beta}_{j_0,\cdot}), (\tilde{\alpha}_{j_0,\cdot}), \ldots)$$

with all entries zero, except $\tilde{\beta}_{j_0,k'} = 1$. Then define the coefficient sequences $(u_k^{(j_0,k')})$ by

$$(U_{j_0,j_1})^T \cdot \theta^{(j_0,k')} = (u_k^{(j_0,k')})_{k \in \mathbf{Z}}.$$

Because our filters are of compact support, $u^{(j_0,k')}$ has only finitely many nonzero terms. Define

$$\tilde{\varphi}_{j_0,k'} = \sum_{k \in \mathbf{Z}} u_k^{(j_0,k')} \varphi_{j_1,k}. \tag{4.1}$$

These wavelets are finite linear combinations of interpolating wavelets at finer scales. Similarly, if $j > j_1$, we can construct vectors $(u_k^{(j,k')})_{k\in\mathbf{Z}}$ and define

$$\tilde{\psi}_{j,k'} = \sum_{k\in\mathbf{Z}} u_k^{(j,k')} \varphi_{j_1,k}. \tag{4.2}$$

Because of the orthogonality of U_{j_0,j_1}, and the interpolating property of the $\varphi_{j_1,k}$, these new functions are orthonormal with respect to the sampling inner product $\langle f, g\rangle_{j_1} = \sum_{k\in\mathbf{Z}} f(k/2^{j_1})g(k/2^{j_1})$.

Complete this collection of functions to $j \geq j_1$ using the interpolating wavelets:

$$\begin{aligned} \tilde{\psi}_{j,k} &= \psi_{j,k}, \qquad j \geq j_1, \\ \tilde{\alpha}_{j,k} &= \alpha_{j,k}, \qquad j \geq j_1. \end{aligned}$$

Theorem 4.2 *Let $f \in C_0(\mathbf{R})$. Then*

$$f = \sum_{k=-\infty}^{\infty} \tilde{\beta}_{j_0,k}\tilde{\varphi}_{j_0,k} + \sum_{j\geq j_0}\sum_k \tilde{\alpha}_{j,k}\tilde{\psi}_{j,k}$$

in the sense of uniform convergence of finite partial sums, $j \leq J$, $|k| \leq K$, $J, K \to \infty$.

Proof. This is just a rearrangement of the identity in Theorem 2.6,

$$f = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j\geq j_0}\sum_k \alpha_{j,k}\psi_{j,k}, \qquad f \in C_0(\mathbf{R}),$$

by finite linear combinations. □

We make some simple remarks. First, let $\tilde{V}_{j_0}$ denote the collection of sums $\sum_{k=-\infty}^{\infty} \tilde{\beta}_{j_0,k}\tilde{\varphi}_{j_0,k}$ where the sequence $\tilde{\beta}_{j_0,k}$ is of at-most polynomial growth. Then

$$\Pi_{\min(D,\tilde{D})} \subset \tilde{V}_{j_0}.$$

Second, the functions $\tilde{\varphi}_{j,k}$ and $\tilde{\psi}_{j,k}$ are *not* dilations and translations of a single pair of functions; instead the functions differ from one resolution level j to the next. To make this precise, let $\sigma_{j,k} f = 2^{-j/2} f(2^{-j}(x+k))$ be the standardization operator, so that, for example $\sigma_{j,k}\varphi_{j,k} = \varphi$ for each j and k. Then $\sigma_{j,k}\tilde{\psi}_{j,k}$ is a function which, in general, depends on j but not on k. In fact it is a function $\tilde{\psi}^{[\ell]}$, $\ell = j - j_1$. Similarly $\sigma_{j_0,k}\tilde{\varphi}_{j_0,k}$ is a function $\tilde{\varphi}^{[\ell]}$. At one limit, $\ell = 0$, $\tilde{\psi}^{[\ell]}$ and $\tilde{\varphi}^{[\ell]}$ are the interpolating wavelets. At the other limit, $\ell \to \infty$ they are orthogonal wavelets:

$$\|\tilde{\psi}^{[\ell]} - \bar{\psi}\|_1 \to 0, \qquad \ell \to \infty, \tag{4.3}$$

and

$$\|\tilde{\varphi}^{[\ell]} - \bar{\varphi}\|_1 \to 0, \qquad \ell \to \infty. \tag{4.4}$$

Indeed, Daubechies (1988, section 3.B) discusses a *Cascade algorithm* in which one iteratively applies a two-scale filtering operator starting from a "tent" function and proves that the result converges in L^1 to a limit, which with the coefficient set-up assumed here must be the compactly supported orthogonal scaling function. The tent function is of course the Schauder interpolating wavelet.

Our functions $\tilde{\varphi}^{[\ell]}$ and $\tilde{\psi}^{[\ell]}$ are the result of running the Cascade algorithm ℓ times starting from regular interpolating wavelets. A modification of Daubechies' proof of convergence of the cascade algorithm would show that (4.3)-(4.4) hold. However, our linking assumption, and Lemma 4.3 below furnish immediately, without any honest work, the convergence (4.3)-(4.4) as well as convergence in many stronger norms.

Third, as the functions $\tilde{\varphi}_{j_0,k}$ and $\tilde{\psi}_{j,k}$ are finite linear combinations of the interpolating wavelets, they are all R-regular. However due to the asymptotic agreement with the orthogonal wavelets, at best, we can hope for results saying that

$$\sigma_{j,k}\tilde{\varphi}_{j,k} \in C^{\min(R,\bar{R})}$$

uniformly in j,k. Establishing such bounds in a general situation, i.e. without linking, would be quite tedious since one must study the regularity, uniformly in $\ell > 0$, of the functions $\tilde{\varphi}^{[\ell]}$, and $\tilde{\psi}^{[\ell]}$. The customary way is perhaps to adapt the methods of Daubechies (1988) and Daubechies and Lagarias(1991), to study the decay of such functions in the Fourier domain, where they have product representations:

$$\hat{\varphi}^{[\ell]}(\xi) = \Pi_{j=1}^{\ell} \bar{m}_0(2^{-j}\xi)\, \Pi_{j=\ell+1}^{\infty} m_0(2^{-j}\xi). \tag{4.5}$$

Here $\bar{m}_0$ and m_0 are the 'two-scale' symbols of the orthogonal wavelet system and of the interpolating wavelet system, respectively. However, to establish quantitative bounds, uniform in ℓ, for general choices of such two-scale symbols, seems challenging.

The special choice of the interpolating wavelets we have made above, linking them to the orthogonal wavelets of compact support, gives an easy proof of the properties of the hybrid wavelets $\tilde{\varphi}^{[\ell]}$, uniformly in ℓ. Our assumption links the different two-scale symbols by

$$|\bar{m}_0|^2 = m_0; \tag{4.6}$$

so, comparing with (4.5), we have (with z^* the complex conjugate of z)

$$\begin{aligned} \hat{\tilde{\varphi}}^{[\ell]}(\xi) &= \Pi_{j=1}^{\ell} \bar{m}_0(2^{-j}\xi) \Pi_{j=\ell+1}^{\infty} |\bar{m}_0(2^{-j}\xi)|^2 \\ &= \Pi_{j=1}^{\infty} \bar{m}_0(2^{-j}\xi) \Pi_{j=\ell+1}^{\infty} \bar{m}_0^*(2^{-j}\xi) \\ &= \hat{\bar{\varphi}}(\xi) \cdot \hat{\bar{\varphi}}^*(2^{-\ell}\xi). \end{aligned}$$

This proves

Lemma 4.3 *Let the interpolating wavelet φ be linked to the orthogonal wavelet $\bar{\varphi}$. Then the hybrid wavelet $\tilde{\varphi}^{[\ell]}$ is the convolution of the orthogonal wavelet with an approximate identity manufactured from the orthogonal wavelet:*

$$\tilde{\varphi}^{[\ell]} = K_\ell \star \bar{\varphi},$$

where

$$(K_\ell)(t) = 2^\ell \cdot \bar{\varphi}(-2^\ell t).$$

Similarly,

$$\tilde{\psi}^{[\ell]} = K_{\ell+1} \star \bar{\psi}.$$

Uniform estimates of regularity and localization now follow by inspection:

Corollary 4.4 *With the interpolating wavelet linked to the orthogonal wavelet of compact support, the hybrid wavelets satisfy the uniform support bounds*

$$supp(\tilde{\varphi}^{[\ell]}) \subset [-B, B], \qquad supp(\tilde{\psi}^{[\ell]}) \subset [-B, B], \qquad \ell > 0$$

for a constant B independent of $\ell > 0$. The hybrid wavelets satisfy the uniform smoothness bounds

$$\|\frac{d^m}{dt^m}\tilde{\varphi}^{[\ell]}\|_\infty \leq A, \qquad 0 \leq m \leq \bar{R}, \quad \ell > 0,$$

$$\|\frac{d^m}{dt^m}\tilde{\psi}^{[\ell]}\|_\infty \leq A, \qquad 0 \leq m \leq \bar{R}, \quad \ell > 0.$$

4.2 Smoothness Characterization

Our assumptions linking the orthogonal wavelet and the interpolating wavelet lead easily to characterization theorems.

Theorem 4.5 *Build the hybrid transform for $C_0(\mathbf{R})$ from an interpolating wavelet φ linked to an orthogonal wavelet $\bar{\varphi}$ of regularity $\bar{R}$ and polynomial span $\bar{D}$. Let $f \in B^\sigma_{p,q}(\mathbf{R})$, $\min(\bar{D}, \bar{R}) > \sigma > 1/p$, $p, q \in (0, \infty]$. Then*

$$\|f\|_{B^\sigma_{p,q}} \asymp \|\tilde{\theta}\|_{b^\sigma_{p,q}},$$

with constants independent of $j_1 - j_0 > 0$.

As before, this result has the consequence that the convergence of the wavelet reconstruction is unconditional for functions f belonging to $B^\sigma_{p,q}$ with $\sigma > 1/p$ – it does not depend on the order in which the terms are summed.

Theorem 4.6 *Build the hybrid transform for $C_0(\mathbf{R})$ from compactly supported $(\bar{R}, \bar{D})$-orthogonal wavelets, using their autocorrelations as the interpolating wavelets. Let $\min(\bar{D}, \bar{R}) > \sigma > 1/p$, $p, q \in (1, \infty]$. Then*

$$\|f\|_{F^\sigma_{p,q}} \asymp \|\tilde{\theta}\|_{f^\sigma_{p,q}},$$

with constants independent of $j_1 - j_0 > 0$.

4.3 Hybrid Transform and Special Interpolation

The hybrid transform gives a new interpretation of the operation of pyramid filtering applied to samples $(2^{-j_1/2} f(k/2^{j_1}))$. The blind hope would be, since the filtering is the filtering associated with orthogonal wavelets of compact support, that the partial reconstruction

$$\bar{f} = \sum_k \tilde{\beta}_{j_0,k} \bar{\varphi}_{j_0,k} + \sum_{j_0 \leq j < j_1} \tilde{\alpha}_{j,k} \bar{\psi}_{j,k}$$

in some sense reconstructs f. However, if we sample $\bar{f}$ and calculate its coefficients, we don't get back the coefficients we started with: the method is not self-consistent. The alternative reconstruction

$$\tilde{f} = \sum_k \tilde{\beta}_{j_0,k}\tilde{\varphi}_{j_0,k} + \sum_{j_0 \le j < j_1} \tilde{\alpha}_{j,k}\tilde{\psi}_{j,k}$$

has the interpolation property

$$\tilde{f}(k/2^{j_1}) = f(k/2^{j_1}), \qquad k \in \mathbf{Z},$$

so it is self-consistent.

Because of the special assumption linking our interpolating wavelets to orthogonal ones, there is a close connection between the naive reconstruction and the improved reconstruction. Indeed, by Lemma 4.3

$$\tilde{f} = K_{j_1} \star \bar{f}.$$

Simply smoothing $\bar{f}$ on the same scale as the sampling results in an overall procedure which interpolates f at the sampling grid.

5 Hybrid Transform on the Interval

Let $n = 2^{j_1}$ and $(S_n f)$ be the normalized samples of f at k/n $0 \le k < n$, (normed by $n^{-1/2}$). Let $\theta^{(n)}$ be the result of filtering these samples via the finite, boundary corrected algorithm of [CDJV]. This transforms the n samples to a vector of n empirical wavelet coefficients

$$\tilde{\theta}^{(n)} = \Big((\tilde{\beta}_{j_0,\cdot}), (\tilde{\alpha}_{j_0,\cdot}), (\tilde{\alpha}_{j_0+1,\cdot}), (\tilde{\alpha}_{j_1-1,\cdot}) \Big).$$

We may interpret these as the first n coefficients of f in a hybrid expansion. In a fashion similar to (4.1), we obtain coefficients $u_k^{(j_0,k')}$ via

$$(u_k^{(j_0,k')})_{k=0}^{2^{j_1}-1} = P_D^{-1} \circ (U_{j_0,j_1})^T \circ \theta^{(j_0,k)}$$

where P_D is the pre-conditioning operator and U_{j_0,j_1} is the boundary-adjusted pyramid operator defined in [CDJV]. This lets us define

$$\tilde{\varphi}^{[]}_{j_0,k'} = \sum_{k=0}^{2^j-1} u_k^{(j_0,k')} \varphi^{[]}_{j_1,k}$$

Similarly for $j_0 \le j < j_1$ and $0 \le k' < 2^j$

$$\tilde{\psi}^{[]}_{j,k'} = \sum_{k=0}^{2^j-1} u_k^{(j,k')} \varphi^{[]}_{j_1,k}.$$

Define now $\tilde{\alpha}_{j,k}$ and $\tilde{\psi}_{j,k}$ for $j \ge j_1$ by

$$\begin{aligned} \tilde{\alpha}^{[]}_{j,k} &= \alpha_{j,k} \qquad j \ge j_1,\ k \in K_j, \\ \tilde{\psi}^{[]}_{j,k} &= \psi^{[]}_{j,k} \qquad j \ge j_1,\ k \in K_j. \end{aligned}$$

Arguing exactly as in earlier sections gives the following result. We omit the proof.

Theorem 5.1 *Let $f \in C[0,1]$. Then*

$$f = \sum_{k=0}^{2^j-1} \tilde{\beta}_{j_0,k}\tilde{\varphi}^{[]}_{j_0,k} + \sum_{j \geq j_0} \sum_{k \in K_j} \tilde{\alpha}_{j,k}\tilde{\psi}^{[]}_{j,k}$$

with uniform convergence of finite partial sums:

$$\|f - \sum_{k \in K_{j_0}} \tilde{\beta}_{j_0,k}\tilde{\varphi}^{[]}_{j_0,k} + \sum_{j_0 \leq j \leq J} \sum_{k \in K_j} \tilde{\alpha}_{j,k}\tilde{\psi}^{[]}_{j,k}\|_\infty \to 0, \quad J \to \infty.$$

The coefficients $\tilde{\beta}_{j_0,k}$ and $\tilde{\alpha}_{j,k}$, $j < j_1$, can be evaluated by applying pre-conditioned pyramid filtering to the samples $(f(k2^{-j_1}))_{k=0}^{2^{j_1}-1}$.

To complete our results we have smoothness characterizations. The proof is discussed in the appendix.

Theorem 5.2 *Let $\frac{1}{p} < \sigma < \min(\bar{R}, \bar{D})$ and $p, q, \in (0, \infty]$. Define*

$$\tilde{\theta} = ((\tilde{\beta}_{j_0,k})_{k \in K_{j_0}}, (\tilde{\alpha}_{j_0,k})_{k \in K_{j_0}}, \dots, (\tilde{\alpha}_{j_1,\cdot})_{k \in K_{j_1}}, \dots).$$

The norm

$$\|\tilde{\theta}\|_{b^\sigma_{p,q}} \equiv \|(\tilde{\beta}_{j_0,\cdot})\|_{\ell^p} + (\sum_{j \geq j_0} (2^{js}(\sum_{k=0}^{2^j-1} |\tilde{\alpha}_{j,k}|^p)^{1/p})^q)^{1/q}$$

is an equivalent norm for $B^\sigma_{p,q}[0,1]$, with constants of equivalency independent of $j_1 > j_0$.

Theorem 5.3 *Let $\frac{1}{p} < \sigma < \min(\bar{R}, \bar{D})$ and $p, q, \in (1, \infty]$. Then*

$$\|f\|_{F^\sigma_{p,q}[0,1]} \asymp \|\theta\|_{f^\sigma_{p,q}}$$

with constants of equivalency independent of $j_1 > j_0$.

Together theorems 5.2 and 5.3 furnish [ET1]-[ET5] of the introduction and complete the proof of all main results.

6 Discussion

6.1 The Critical Case: $\sigma = 1/p$

Interpolating wavelet transforms are not suited for dealing with Besov and Triebel spaces if $\sigma = 1/p$. First, members of $B^\sigma_{p,q}[0,1]$ and $F^\sigma_{p,q}[0,1]$ in the critical case are guaranteed to be bounded only if $q \leq 1$. For example, $\log(x) \in B^{1/p}_{p,\infty}[0,1]$ for all $p > 0$. Hence, sampling of functions in the critical spaces with $q > 1$ doesn't really make sense. In particular, no general inequality of the form

$$\|\theta\|_{b^\sigma_{p,q}} \leq C \cdot \|f\|_{B^\sigma_{p,q}} \tag{6.1}$$

with C independent of f. Indeed, the presence of an infinite value of $f(t)$ at a single binary rational makes the sequence space norm infinite while the function space norm is finite.

Second, even in the case $q \leq 1$ where samples "exist", no general inequality (6.1) is possible. The homogeneous space $\dot{B}^{1/p}_{p,q}$ is invariant under dilation, so the right side of (6.1) is bounded under the rescaling $f(a(t-t_0))$, $a > 1$. Now let f be a smooth function supported in $[0,1]$ with $f(0) = 1$. Set $f^{[j_1]} = f(2^{j_1}t)$. Then $\|f^{[j_1]}\|_{B^\sigma_{p,q}[0,1]}$ is bounded as $j_1 \to \infty$. On the other hand, we get $f^{[j_1]}(2^{-j-1}) = 0$ for $j_0 \leq j < j_1$. Also, there is a unique polynomial P_D interpolating the sequence $(1,0,\dots,0)$ at $(0,1,2,\dots,D)$, and $\pi_{j,0}(\cdot) = P_D(2^{-j}\cdot)$ for $j_0 < j < j_1$. Hence $\alpha_{j,0} = -P_D(1/2)\cdot 2^{-j/2}$ for $j_0 < j < j_1$. We conclude that $\|\theta^{[j_1]}\|_{b^\sigma_{p,q}} \geq |P_D(1/2)|(j_1 - j_0 - 1)^{1/q}$, which is unbounded as $j_1 \to \infty$.

For discussion of the critical case $\sigma = 1/p$, $1 \leq p < \infty$, it is therefore natural to abandon the Besov and Triebel scales. Consider instead the ℓ^p variation spaces V_p of Peetre (1976). These have seminorm

$$\|f\|_{V_p} = \sup_{t_i < t_{i+1}} \|(f(t_{i+1}) - f(t_i))_i\|_p$$

where the sup is over all partitions of the line. Of course, the case $p = 1$ is just the bounded variation seminorm, and the case $p = 2$ is Wiener's quadratic variation. These are all spaces of bounded functions where functions have left-hand and right-hand limits. Hence sampling and interpolation still make sense (if only just barely). Peetre points out that

$$\dot{B}^{1/p}_{p,1} \subset V_p \subset \dot{B}^{1/p}_{p,\infty}, \qquad p \in [1,\infty),$$

where dots denote homogeneous Besov spaces. Hence V_p belongs to the critical case $\sigma = 1/p$.

Now for nice orthogonal wavelet bases we have

$$c \cdot \|\bar{\theta}\|_{\dot{b}^{1/p}_{p,\infty}} \leq \|f\|_{V_p} \leq C\|\bar{\theta}\|_{\dot{b}^{1/p}_{p,1}}.$$

For interpolating wavelets satisfying [IW1]-[IW6], we have, in an ℓ^p variant of Lemma 2.4, that $\|(\alpha_{j,k})_k\|_{\ell^p} \leq C \cdot 2^{-j/2}\|f\|_{V_p}$. Hence, the interpolating wavelet coefficients satisfy

$$\|\theta\|_{\dot{b}^{1/p}_{p,\infty}} \leq C \cdot \|f\|_{V_p}.$$

Similarly, we get

$$\|\theta\|_{\dot{b}^{1/p}_{p,1}} \geq c \cdot \|f\|_{V_p}.$$

These inequalities parallel the orthogonal case, and are essentially the best we can expect.

6.2 Compression in the Critical Case

Real digital imagery generally behaves, along individual scan lines, as samples of a bounded function with discontinuities. Because our discussion to date emphasizes continuous functions, considerations of credibility demand some discussion of what happens when discontinuities are present. The spaces V_p include bounded functions with discontinuities and minimax compression over such spaces may be taken as a model for problems of compressing real images.

We suppose that a vector of samples $S_{j_1} f = (f(k/2^{j_1}))_{k \in K_{j_1}}$ is available, and that the underlying function belongs to V_p. We apply the thresholding scheme of section 3.3 to the wavelet coefficients of these samples, i.e. to the wavelet coefficients at scales $j_0 \le j < j_1$. We get the reconstruction

$$f_{j_1}^{(\epsilon)} = \sum_k \beta_{j_0,k} \varphi_{j_0,k}^{[]} + \sum_{j=j_0}^{j_1-1} \alpha_{j,k}^{(\epsilon)} \psi_{j,k}^{[]}$$

We evaluate error in compressing the vector $S_{j_1} f$ by the ℓ^∞ norm of the difference between original samples and the samples of the reconstruction $(f_{j_1}^{(\epsilon)}(k/2^{j_1}))$.

Slight adaptations of the reasoning behind Theorem 3.8 give

Lemma 6.1 *Suppose $f \in V_p$ and that the wavelet transform satifies $D \ge 1 \ge 1/p$. Then with constants c_1, c_2 depending only on the wavelet chosen,*

$$\|S_{j_1} f - f_{j_1}^{(\epsilon)}\|_{\ell^\infty} \le \epsilon \cdot j_1 \cdot c_1, \qquad \epsilon > 0 \tag{6.2}$$

while the number of nonzero terms $N(\epsilon) = \#\{(j,k) : \alpha_{j,k}^{(\epsilon)} \ne 0\}$ satisfies

$$N(\epsilon) \le c_2 \cdot j_1 \cdot \epsilon^{-p} \cdot \|f\|_{V_p}^p. \tag{6.3}$$

As a particular example, we let $p = 1$, and we define a sequence $\epsilon_n = c_2 \cdot j_1 \cdot \|f\|_{V_1} \cdot \frac{1}{n}$ so that $N(\epsilon_n) \le n$ and we get

$$\|S_{j_1} f - f_{j_1}^{(\epsilon_n)}\|_{\ell^\infty} \le n^{-1} \cdot c_2 \cdot c_1 \cdot j_1^2 \cdot \|f\|_{V_1}, \qquad n \to \infty.$$

This behavior is optimal to within logarithmic terms. Indeed since we have the estimate

$$\|(\alpha_{j,k})_k\|_1 \le C \cdot 2^{-j/2} \cdot \|f\|_{V_1}$$

the argument of Lemma 3.9 shows that the best reconstruction f_n based on summing n wavelet terms satisfies

$$\sup_{f : \|f\|_{V_1} \le 1} \|S_{j_1} f - f_n\|_{\ell^\infty} \ge c \cdot \|f\|_{V_1}/n, \qquad n \ge n_0$$

Compare this with the discussion of Kahane's Theorem in (DeVore, 1989).

6.3 Alternate Interpolating Bases

Our construction of an interpolating basis for the interval is built to resemble, as nearly as possible, the construction of orthogonal wavelet bases for the interval as in Meyer (1991) and [CDJV]. It is built from translates and dilates of a finite collection of functions, at the "heart" of which are functions from the transform on the real line. However, this is just one way to construct interpolants. There is a dyadic construction of an interpolating spline Schauder basis of $C[0,1]$ by Domsta (1976) for which function space characterizations have been proven; but Domsta's basis does not have dilation and translation homogeneity. Compare also the work of Schonefeld (1972) and Subbotin (1972).

6.4 Related Wavelet Transforms

The concept of interpolating transform is related to a great deal of existing wavelet work, which we attempt to mention here in alphabetical order.

Aldroubi-Unser. Aldroubi and Unser (1992) develop non-orthogonal interpolating scaling functions, and mention the associated wavelets. Their approach is based on a signal processing/filtering point of view, emphasizing interpolating splines (which are not compactly supported). They claim, in passing, that the wavelet (as well as the scaling function) is also interpolating. It is not clear, from their filtering representation, whether their wavelet satisfies $\psi(x) = \varphi(2(x-1/2))$, so I am unable to say that they are actually performing the interpolating transform as defined here. They give examples of "cardinal spline wavelet" decompositions of images, whose captions suggest they may be instances of interpolating transforms.

Saito-Beylkin. The interpolating wavelet transform is closely related to the Saito-Beylkin (1992) transform, which uses wavelets derived from Deslauriers-Dubuc interpolants, but which does not decimate by factors of 2 after filtering. The aim of Saitoh and Beylkin is not to construct an interpolating transform, but to avoid phase shifts in edge detection. Also, the two transforms differ in their choice of ψ. Saito and Beylkin use the autocorrelation of an orthogonal wavelet $\bar{\psi}$, while we do not.

Cohen-Daubechies-Feauveau. Interpolating wavelet transforms represent a degenerate instance of the biorthogonal wavelet transforms of Cohen, Daubechies, Feauveau (1990), in which one wavelet of the biorthogonal pair is no longer a function, but a linear combination of Dirac masses.

DeVore-Popov. The interpolating wavelet transform is also closely related to the DeVore-Popov (1988) transform, a nonlinear transform based on spline quasi-interpolants which is useful for characterizing Besov spaces on the interval. A number of useful technical tools such as those underlying e.g. Lemma 7.2 may be found in that work.

DeVore-Jawerth-Lucier. DeVore, Jawerth, and Lucier (1990) construct a 2-dimensional wavelet expansion based on quasi-interpolating box-spline wavelets. Although this expansion is not an interpolating transform, it has a similar domain of applicability (i.e. to $\sigma > d/p$ in dimension d). The article gives very nice examples where such a transform may be used to compress smooth nonparametric surfaces. The compression results of our section 3.3 are similar in spirit to those of DeVore, Jawerth, and Lucier, who use, as we have said, something besides simple thresholding in order to compress expansions optimally.

Lemarié-Malgouyres. P.G. Lemarié and G. Malgouyres (1989) construct orthogonal scaling functions which are also interpolating. Such scaling functions are necessarily of non-compact support (Daubechies,1992, Page 211). However, even if we had used such wavelets, the interpolating transform defined above would not reduce to the orthogonal transform. This is because in our definition, $\psi = \varphi(2(x-1/2))$ is not orthogonal to V_{-1} etc.

Sickel. A particular scaling function which is both interpolating and orthogonal to translates is the cardinal sine, $\sin(\pi x)/(\pi x)$. (It is not of rapid decay). Work characterizing Besov and Triebel-Lizorkin spaces on the line, $\sigma > 1/p$, in terms of samples and using properties of the cardinal series was given by Sickel (1992). Although Sickel's paper did not use the language of wavelets it might, in retrospect, be interpreted as saying that function space characterizations may be carried out even using wavelets without good decay.

6.5 Variations

Normalization. Much of the "accounting" work with L^2-normalized interpolating wavlelets $2^{j/2}\psi(2^j x - k)$ is counterintuitive. We have used it in this article only to emphasize the formal similarity with the orthogonal wavelet transform. It makes perfectly good sense to use instead L^∞-normalization $\psi(2^j x - k)$.

Periodization. The periodized wavelet for $[0, 2\pi]$

$$\varphi^{\circ}_{j,k}(\omega) = \sum_{k'} \varphi_{j,k}(\omega/2\pi + 2^j k')$$

is well-defined (assuming rapid decay of φ) and is interpolating. The corresponding transform characterizes Besov and Triebel spaces on the circle in the range $\sigma > 1/p$.

Higher Dimensions. When refining a dyadic grid in dimension $d > 1$, each scale halving produces $2^d - 1$ new mesh points per old mesh point. Hence there need to be $2^d - 1$ ψ functions, which could be simply

$$\psi^{[\epsilon]} = 2^{jd/2}\phi(2^j x - (k + \epsilon/2))$$

for ϵ a nonzero binary d-vector, and φ an interpolating wavelet. Then as long as $\sigma > d/p$, we get that the interpolating coefficients give equivalent norms to Besov and Triebel spaces.

Interpolating Spline Wavelets for $[0, 1]$. We have not so far described the interpolating spline wavelets for $[0, 1]$. We define $V_j[0, 1]$ as the set of piecewise polynomials of degree D with knots at $k/2^j$ for $k = \lceil D/2 \rceil, ..., 2^j - \lceil D/2 \rceil - 1$. Then $V_j[0, 1]$ contains all polynomials of degree D, and satisfies the nesting condition $V_j[0, 1] \subset V_{j+1}[0, 1]$. A basis for $V_j[0, 1]$ is given by those splines $\varphi^{[]}_{j,k}$ satisfying Kronecker interpolation conditions. Unfortunately, this basis does not consist of functions related to each other by dilation and translation. So it is not really a wavelet basis.

There is, however, a connection between interpolating splines on the interval and interpolating spline wavelets on the line. It is based on the same principle of extension and folding that we exploited in extending the wavelet transforms for Deslauriers-Dubuc wavelets on the line to the interval.

Lemma 6.2 *Let φ be a fundamental spline of degree D, D an odd integer greater than 1. Given $(\beta_{j,k})_{k=0}^{2^j-1}$, there is a unique extension by a polynomial of degree D,*

$$\tilde{\beta}_{j,k} = 2^{-j/2} \cdot \pi^{[]}(k/2^j), \qquad k \notin K_j$$

to a sequence $(\tilde{\beta}_{j,k})_{k \in \mathbf{Z}}$ which ensures that the infinite sum

$$\tilde{f} = \sum_{k \in \mathbf{Z}} \tilde{\beta}_{j,k}\varphi_{j,k}$$

satisfies

$$\tilde{f}|_{[0,1]} \in V_j[0, 1]$$

This lemma establishes a correspondence between the interpolating spline spaces on the interval and the interpolating spline spaces on the line. It follows from this lemma that the interpolating splines $\varphi^{[]}_{j,k}$ are corrected versions of the interpolating splines $\varphi_{j,k}$ for the line:

$$\varphi^{[]}_{j,k} = \varphi_{j,k} + \sum_{k' \notin K_j} \pi^{[]}_{j,k}(k'/2^j)\varphi_{j,k'}, \qquad k \in K_j.$$

Here the correction terms vanish on the dyadic grid $K_j/2^j$, and are exponentially decaying with distance from 0 and 1.

7 Appendix: Proofs

7.1 Proofs for Section 2.2

To prove (2.7), it is enough to do so at scale $j = 0$, since it follows for other j by a dilation. Let $x, h \in [0, 1]$. We have, using summation by parts,

$$(P_0 f)(x+h) - (P_0 f)(x) = \sum_k (\beta_{0,k+1} - \beta_{0,k})\Phi^{(h)}(x-k),$$

where

$$\Phi^{(h)}(x) = \sum_{\ell=-\infty}^{-1} (\varphi(x+h+\ell) - \varphi(x+\ell)),$$

and the summation by parts is justified by the rapid decay of φ. Hence

$$|(P_0 f)(x+h) - (P_0 f)(x)| \le \|(\beta_{0,k+1} - \beta_{0,k})\|_{\ell^\infty} \cdot \sum_k |\Phi^{(h)}(x-k)|.$$

Now by the rapid decay of φ, there is a finite constant C_Φ with

$$\sum_k |\Phi^{(h)}(x-k)| \le C_\Phi, \qquad x, h \in [0, 1].$$

Hence

$$\begin{aligned} \omega(1, P_0 f) &\le C_\Phi \cdot \|(\beta_{0,k+1} - \beta_{0,k})_k\|_\infty \\ &= C_\Phi \cdot \|(f(k+1) - f(k))_k\|_\infty \\ &\le C_\Phi \cdot \omega(1, f). \end{aligned}$$

This is (2.7) at scale $j = 0$.

7.2 Proof of Besov Equivalence

Here is the idea. Consider the orthogonal wavelet transform based on Daubechies wavelets $\bar\psi_{j,k}$ of compact support, orthogonal to polynomials of degree $\bar D$, and of regularity $\bar R$. Meyer (1990) has already proved that the orthogonal wavelet transform based on these wavelets of characterizes Besov space $B^\sigma_{p,q}$, $\sigma < \min(\bar R, \bar D)$, $p, q \in [1, \infty]$.

One way of stating this characterization is as follows. Let $\bar{\varphi}_{j,k}$ denote the orthogonal scaling function for the Daubechies wavelets, and let $\bar{P}_j$ denote the orthogonal projection operator $\bar{P}_j f = \sum \bar{\beta}_{j,k}\bar{\varphi}_{j,k}$ where $\bar{\beta}_{j,k} = \int f(t)\bar{\varphi}_{j,k}(t)dt$. $\bar{P}_j f$ is orthogonal projection of f on $\bar{V}_j$, say. Membership in Besov space can be measured by the rate at which $\bar{P}_j f \to f$ in L^p. We will show that, if $\sigma > 1/p$, the non-orthogonal interpolating wavelet transform derives from a linear approximation scheme $(P_j f)$ which converges to $f \in B^\sigma_{p,q}$ at the same rate as the linear approximation scheme $(\bar{P}_j f)$ generated from the corresponding wavelet analysis. This equivalence of approximation power implies the result.

Now the formal proof. We assume to begin with that $p, q \in [1, \infty]$. After we give the formal argument we will describe modifications to cover $p, q \in (0, \infty]$.

It is known that Besov space $B^\sigma_{p,q}(\mathbf{R})$ is normed as follows (Meyer, 1990, Vol. I, Chapter 2, pages 50-51). Define $(\bar{e}_j)_{j=j_0}^\infty$ by

$$\|f - \bar{P}_j f\|_{L^p} = \bar{e}_j 2^{-j\sigma}, \quad j \geq j_0$$

and set

$$\|f\|_{\bar{B}^\sigma_{p,q}} = \|\bar{P}_{j_0} f\|_{L^p} + \|(\bar{e}_j)\|_{\ell^q}.$$

Then $\|f\|_{\bar{B}^\sigma_{p,q}}$ is an equivalent norm for $B^\sigma_{p,q}(\mathbf{R})$. We will first show that this norm is equivalent, if $\sigma > 1/p$, to a similar norm based on the (non-orthogonal) interpolating approximation P_j. Define $(e_j)_{j \geq j_0}$ by

$$\|f - P_j f\|_p = e_j 2^{-j\sigma}, \quad j \geq j_0,$$

and set

$$\|f\|_{B^\sigma_{p,q}} = \|P_{j_0} f\|_p + \|(e_j)\|_{\ell^q}.$$

The equivalence of these norms depends on properties of the projection norms

$$\|P_j\|_{\bar{V}_{j'}} \equiv \sup\{\|P_j f\|_p : \ f \in \bar{V}_{j'}, \ \|f\|_p = 1\},$$

and

$$\|I - P_j\|_{\bar{V}_{j'}} \equiv \sup\{\|f - P_j f\|_p : \ f \in \bar{V}_{j'}, \ \|f\|_p = 1\},$$

and their counterparts with $\bar{P}_j$ and V_j.

The first property: although P_j is not a nicely bounded operator on general L^p spaces, P_j is bounded on $\bar{V}_j'$ even if $j' \gg j$.

Lemma 7.1

$$\|P_j\|_{\bar{V}_{j+\ell}} \leq C \cdot 2^{\ell/p}, \qquad \ell > 0 \tag{7.4}$$

$$\|\bar{P}_j\|_p \leq C, \forall j \tag{7.5}$$

This Lemma is proved below.

The second main result on projections P_j, $\bar{P}_j$ concerns the fact that, if $j' \ll j$, then P_j is a near-identity on $\bar{V}_j'$. The error in this approximation depends on the degree of smoothness of elements of $\bar{V}_j$ and on the degree of polynomial span of V_j.

Lemma 7.2

$$\|I - P_j\|_{\bar{V}_{j-\ell}} \leq C \cdot 2^{-\ell \min(D,\bar{R})}, \qquad \ell > 0 \tag{7.6}$$

$$\|I - \bar{P}_j\|_{V_{j-\ell}} \leq C \cdot 2^{-\ell \min(R,\bar{D})}, \qquad \ell > 0 \tag{7.7}$$

The lemma will also be proved below.

We assemble these facts to bound the behavior of (e_j) in terms of $(\bar{e}_j)$. Develop the telescoping sum,

$$(I - P_j)f = (I - P_j)(\bar{P}_{j_0} f + \sum_{j' \geq j_0} \bar{Q}_{j'} f)$$

giving

$$\|(I - P_j)f\|_p \leq \|(I - P_j)\bar{P}_{j_0} f\|_p + \sum_{j' \geq j_0} \|(I - P_j)\bar{Q}_{j'} f\|_p$$

Now

$$\|(I - P_j)\bar{P}_{j_0} f\|_p \leq \|I - P_j\|_{\bar{V}_{j_0}} \|\bar{P}_{j_0} f\|_p.$$

Applying Lemma 7.2, we have $\|I - P_j\|_{\bar{V}_{j_0}} \leq C \cdot 2^{-(j-j_0)\min(D,\bar{R})}$. Hence, defining $\tilde{e}_j = C \cdot 2^{-(j-j_0)\min(D,\bar{R})} \|f\|_{\bar{B}^\sigma_{pq}}$

$$\|(I - P_j)\bar{P}_{j_0} f\|_p \leq \tilde{e}_j, \qquad j \geq j_0.$$

Because $\bar{W}_{j'} \subset \bar{V}_{j'+1}$

$$\|(I - P_j)\bar{Q}_{j'} f\|_p \leq \|(I - P_j)\|_{\bar{V}_{j'+1}} \|\bar{Q}_{j'} f\|_p$$

and because $\bar{Q}_{j'} f = (f - \bar{P}_{j+1} f) - (f - \bar{P}_{j+1} f)$, the triangle inequality gives

$$\|\bar{Q}_{j'} f\|_p \leq 2^{-j'\sigma}(\bar{e}_{j'} + \bar{e}_{j'+1}).$$

Therefore we can write

$$\sum_{j' \geq j_0} \|(I - P_j)\bar{Q}_{j'} f\|_p \leq \sum_{j' \geq j_0} H(j, j')(\bar{e}_{j'} + \bar{e}_{j'+1}), \qquad j \geq j_0$$

where

$$H(j, j') = C \cdot 2^{-j'\sigma} \|I - P_j\|_{\bar{V}_{j'}}.$$

Combining these observations, we have

$$e_j \leq \tilde{e}_j + \sum_{j'} H(j, j')(\bar{e}_{j'} + \bar{e}_{j'+1}), \qquad j \geq j_0. \tag{7.8}$$

By Lemmas 7.1-7.2,

$$H(j, j') \leq C \cdot 2^{-|j-j'| \cdot \delta},$$

where $\delta = \min(\min(D, \bar{R}) - \sigma, \sigma - 1/p)$. The assumption $1/p < \sigma < \min(D, \bar{R})$ makes $\delta > 0$ so that H decays exponentially fast away from the diagonal and furnishes a bounded linear transformation from ℓ^q to ℓ^q for every $q \in [1, \infty]$. Also as $\sigma < \min(D, \bar{R})$,

$$\|\tilde{e}\|_{\ell^q} \leq C \cdot \|f\|_{\bar{B}^\sigma_{pq}} \cdot \|(2^{-(j-j_0)(\min(D,\bar{R})-\sigma)})_{j \geq j_0}\|_{\ell^q} = C \cdot \|f\|_{\bar{B}^\sigma_{pq}}$$

and so,

$$\|e\|_{\ell^q} \leq C \cdot \|f\|_{\bar{B}^\sigma_{pq}} + 2 \cdot \|H\|_q \|f\|_{\bar{B}^\sigma_{pq}}.$$

A similar argument gives

$$\begin{aligned}\|P_{j_0} f\|_p &\leq C \cdot \|P_{j_0}\|_{\bar{V}_{j_0}} \cdot \|\bar{P}_{j_0} f\|_p + C \sum_{j' \geq j_0} \|P_{j_0}\|_{\bar{V}_{j'}} \cdot (\bar{e}_{j'} + \bar{e}_{j'+1}) \\ &\leq C \cdot \|P_{j_0} f\|_p + C \sum_{j' \geq j_0} 2^{-(j'-j_0)\sigma} (\bar{e}_{j'} + \bar{e}_{j'+1}) \leq C \cdot \|f\|_{B^\sigma_{p,q}}\end{aligned}$$

and so, by (7.8),

$$\|f\|_{B^\sigma_{pq}} \leq C \cdot \|f\|_{\bar{B}^\sigma_{pq}}.$$

To get an inequality in the reverse direction, we systematically reverse the roles of barred and unbarred quantities in the above argument, and get equivalence of the two norms.

Now we show that

$$\|\theta\|_{b^\sigma_{pq}} = \|(\beta_{j_0,\cdot})\|_{\ell^p} + (\sum_{j \geq j_0} (2^{js} (\sum |\alpha_{j,k}|^p)^{1/p})^q)^{1/q}$$

is an equivalent norm to $\|f\|_{B^\sigma_{p,q}}$.

Define $Q_j f = P_{j+1} f - P_j f$, the non-orthogonal projection on W_j. By the triangle inequality, $\|Q_j f\|_{L^p} \leq (e_{j+1} + e_j) \cdot 2^{-j\sigma}$.

We need the fact that the wavelet coefficients of an $f \in V_j$ have an ℓ^p-norm comparable to the L^p norm of f. This is, of course, standard in the study of orthonormal wavelets (compare Meyer (1990, page 31, Lemme 8)) and in spline analysis (DeVore and Popov, 1988). It is a corollary of Lemmas 7.6-7.7 below.

Lemma 7.3 *For (R, D)-interpolating wavelets we have the inequalities*

$$\begin{aligned}\|\sum_k \alpha_{j,k} \psi_{j,k}\|_{L^p} &\asymp \|(\alpha_{j,\cdot})\|_{\ell^p} \cdot 2^{j(1/2-1/p)} \\ \|\sum_k \beta_{j_0,k} \varphi_{j_0,k}\|_{L^p} &\asymp \|(\beta_{j,\cdot})\|_{\ell^p} \cdot 2^{j_0(1/2-1/p)}.\end{aligned}$$

with implied constants of equivalence independent of j_0, j.

Using Lemma 7.3,

$$\begin{aligned}\|P_{j_0} f\|_{L^p} &\asymp \|(\beta_{j_0,\cdot})\|_{\ell^p} \cdot 2^{j_0(1/2-1/p)} \\ \|Q_j f\|_{L^p} &\asymp \|(\alpha_{j,\cdot})\|_{\ell^p} \cdot 2^{j(1/2-1/p)}\end{aligned}$$

and so from $s = \sigma + 1/2 - 1/p$ we see that $C \cdot (\|P_{j_0} f\|_p + \|e_j\|_{\ell^q}) \geq \|\theta\|_{b^\sigma_{p,q}}$.

In the other direction, telescoping sums give

$$\begin{aligned}e_j &= 2^{j\sigma} \cdot \|\sum_{\ell \geq 0} Q_{j+\ell} f\|_{L^p} \leq 2^{j\sigma} \cdot \sum_{\ell \geq 0} \|Q_{j+\ell} f\|_{L^p} \\ &\leq C \cdot 2^{j\sigma} \cdot \sum_{\ell \geq 0} \|(\alpha_{j+\ell,\cdot})\|_{\ell^p} \cdot 2^{-(j+\ell)(1/p-1/2)}.\end{aligned}$$

Hence, defining d_j by $2^{-js} d_j = \|(\alpha_{j,\cdot})\|_{\ell^p}$, and noting that $s = \sigma + 1/2 - 1/p$,

$$e_j \leq C \cdot \sum_{\ell \geq 0} d_{j+\ell} 2^{-\ell\sigma}.$$

Hence formally we have the convolution $\sum_{\ell \geq 0} d_{j+\ell} 2^{-\ell\sigma} = (h * d)_j$. Here the kernel h is in ℓ^1 if $\sigma > 0$. Hence,

$$\|(e_j)\|_{\ell^q} \leq \|h\|_{\ell^1} \cdot \|(d_j)\|_{\ell^q}$$

and so

$$\|P_{j_0} f\| + \|(e_j)\|_{\ell^q} \leq C \cdot \|\theta\|_{b^\sigma_{p,q}}.$$

This finishes the proof in the case $p, q \geq 1$. The general case is not essentially different, but we need some prefatory comments. There are two definitions of Besov Space if $p < 1$: one used by Peetre (1974), and by Frazier and Jawerth (1985); the other used by DeVore and Popov (1988). The two definitions look different, and for certain ranges of parameters they are different. However, in the range $\sigma > 1/p$ they give the same spaces. Orthogonal wavelets of compact support give an unconditional basis of these spaces. And these spaces are again characterized by rate of approximation through smooth orthogonal wavelets of compact support.

The proof given above is general, the Lemmas are proved for all $p > 0$, so the proof works for general p, with one proviso. The L^p norm does not obey the triangle inequality, but instead the p-triangle inequality

$$\|f + g\|_p^p \leq \|f\|_p^p + \|g\|_p^p, \qquad 0 < p < 1;$$

compare DeVore and Popov (1988). Therefore, every place the triangle inequality would be used, one should substitute the p-triangle inequality, and make the modifications this would require. The changes are all entirely superficial. The real work in the extension to $p \in (0, 1)$ is in the proofs of the Lemmas.

Also, one should note that a matrix decaying exponentially fast away from the diagonal and a convolution operator with kernel decaying exponentially fast away from the origin will both be bounded on ℓ^q for any $q > 0$. □

7.3 Lemmas Supporting Besov Equivalence

Here we prove the lemmas just referred to

7.3.1 Lemma 7.1

The assertion (7.5) concerning the orthogonal projection operator is well-known; it can be easily had from a simple analysis of the cases $p = 1, \infty$, and interpolation.

The bound (7.4) may also be had by analysis of extreme cases $p = 1, \infty$; the cases $p \in (0, 1]$ all follow the same pattern of argument; and $p \in (1, \infty)$ follow by interpolation of extreme cases.

Note that P_j is bounded as an operator between L^∞ spaces. Indeed, using Lemma 7.3,

$$\|P_j f\|_\infty \leq C \|(\beta_{j,\cdot})\|_{\ell^\infty} \cdot 2^{j/2} \tag{7.9}$$

and

$$|\beta_{j,k}| \leq \|f\|_\infty \cdot 2^{-j/2} \tag{7.10}$$

so $\|P_j f\|_\infty \leq C\|f\|_\infty$.

At the other extreme, $p = 1$, things are more complicated. We need the fact that samples of $f \in \bar{V}_j$ at the natural rate have an ℓ^p norm bounded the L^p norm of f.

Lemma 7.4 *Let S_j denote the sampling operator*

$$S_j f = (f(k/2^j))_{k \in \mathbf{Z}}$$

Let $\bar{V}_j$ denote the collection of sums $\sum_k \bar{\beta}_{j,k} \bar{\varphi}_{j,k}$ generated by a continuous wavelet $\bar{\varphi}$ of compact support. For all $0 < p \leq \infty$, we have the inequality

$$\|S_j f\|_{\ell^p} \leq C(p) \cdot 2^{j/p} \|f\|_{L^p} \qquad f \in \bar{V}_j,$$

where $C(p)$ does not depend on j or f.

This lemma will be proved after Lemma 7.3 below.

Combining Lemmas 7.3-7.4, if $f \in \bar{V}_{j'}$ then

$$\|(\beta_{j',\cdot})\|_1 \leq C \cdot 2^{j'/2} \cdot \|f\|_1$$

Now by definition

$$\beta_{j+h, 2^h k} = 2^{-h/2} \beta_{j,k}$$

and so

$$\|(\beta_{j,\cdot})\|_1 \leq 2^{(j'-j)/2} \|(\beta_{j',\cdot})\|_1$$

which yields

$$\|(\beta_{j,\cdot})\|_1 \leq C \cdot 2^{j'-j/2} \cdot \|f\|_1.$$

Using again Lemma 7.3, we have $\|P_j f\|_1 \leq C \cdot 2^{-j/2} \|(\beta_{j,\cdot})\|_1$. Combining these yields

$$\|P_j f\|_1 \leq C \cdot 2^{j'-j} \cdot \|f\|_1.$$

Interpolating between the cases $p = 1, \infty$ gives the full result (7.4).

7.3.2 Lemma 7.2

We begin by studying the operator norms for scale $j' = 0$, $j' < j$, and for $p \leq 1$ and $p = \infty$. If $p \leq 1$, and $f \in \bar{V}_0$ then

$$\|f - P_j f\|_p \leq (\sum_k |\bar{\beta}_{0,k}|^p \ \ \|(I - P_j)\bar{\varphi}_{0,k}\|_p^p)^{1/p} \leq \|(\bar{\beta}_{0,k})_k\|_{\ell^p} \cdot \sup_k \|(I - P_j)\bar{\varphi}_{0,k}\|_p.$$

But $\|(\bar{\beta}_{0,k})_k\|_{\ell^p} \leq C(p) \cdot \|f\|_p$ by Lemma 7.3, so

$$\|I - P_j\|_{\bar{V}_0} \leq C(p) \cdot \sup_k \|(I - P_j)\bar{\varphi}_{0,k}\|_p, \qquad p \leq 1.$$

For $p=\infty$ we have

$$\begin{aligned}|(f-P_jf)(x)| &\leq \sum_k |\bar\beta_{0,k}|\;|(I-P_j)\bar\varphi_{0,k}(x)| \\ &\leq \|(\bar\beta_{0,k})_k\|_{\ell^\infty}\cdot\sum_k |(I-P_j)\bar\varphi_{0,k}(x)|.\end{aligned}$$

Again by Lemma 7.3 we have

$$\|I-P_j\|_{\bar V_0} \leq C(\infty)\cdot\sum_k |(I-P_j)\bar\varphi_{0,k}(x)|, \qquad p=\infty.$$

Now Lemma 7.5 below gives immediately that for an (R,D)-interpolating wavelet transform and a $(\bar R,\bar D)$ orthogonal wavelet transform, we have

$$|(I-P_j)\bar\varphi(t)| \leq A_\ell\cdot(1+|t|)^{-\ell}\cdot 2^{-j\min(D,\bar R)}, \qquad j\geq 0, \ell>0, t\in\mathbf{R} \tag{7.11}$$

Hence

$$\sup_k \|(I-P_j)\bar\varphi_{0,k}\|_p \leq C(p)\cdot 2^{-j\min(D,\bar R)}, \qquad p\leq 1,$$

and

$$\sup_x \sum_k |(I-P_j)\bar\varphi_{0,k}(x)| \leq C(\infty)\cdot 2^{-j\min(D,\bar R)}.$$

For any $p\in(1,\infty)$ the operator norm is not larger than the maximum at the two endpoints $p=1,\infty$, so

$$\|I-P_j\|_{\bar V_0} \leq C(p)\cdot 2^{-j\min(D,\bar R)}, \qquad p\in[1,\infty].$$

The inequality extends by dilation to other scales.

The argument for the norm of the operator $\bar P_j$ is exactly the same.

7.4 General Technical Lemmas

For use in later sections, we state a general result. We recall that $\sigma_{j,k}$ denotes the standardization operator $(\sigma_{j,k}f)(t) = 2^{-j/2}f(2^{-j}(t+k))$, so that for wavelets on the line, $\sigma_{j,k}\psi_{j,k}=\psi$ and $\sigma_{j,k}\varphi_{j,k}=\varphi$.

Lemma 7.5 *Let $(f_{j,k})$ be a system of functions, not necessarily dilations and translations of a single function, indexed by a subset of $\mathbf{Z}^2$ and satisfying the uniform smoothness condition*

$$(\sigma_{j,k}f_{j,k})_{j,k} \text{ occupies a bounded subset of } C^{\bar R}$$

and the localization condition

$$\frac{d^m}{dt^m}(\sigma_{j,k}f_{j,k})(t) \leq A_\ell\cdot(1+|t|)^{-\ell}, \qquad t\in\mathbf{R},\ \ell>0,\ 0\leq m\leq\lfloor\bar R\rfloor.$$

Suppose that $(v_{j,k})$ is a localized family of functions

$$(\sigma_{j,k}v_{j,k})(t) \leq A_\ell\cdot(1+|t|)^{-\ell}, \qquad t\in\mathbf{R},\ \ell>0,$$

generating a family of operators

$$P_j f = \sum_k f(k/2^j) v_{j,k}$$

which preserves polynomials of degree D:

$$P_j \pi = \pi, \qquad \pi \in \Pi_D.$$

Then, uniformly in $j' \leq j$ and k

$$|\sigma_{j',k}(I - P_j) f_{j',k}|(t) \leq 2^{(j'-j)\min(\bar{R}, D)} \cdot C_\ell \cdot (1 + |t|)^{-\ell} \qquad t \in \mathbf{R},\ \ell > 0. \tag{7.12}$$

Proof. Let $\alpha = min(D, \bar{R})$, $m = \lfloor \alpha \rfloor$, and $\delta = \alpha - m$. Let $\phi = \sigma_{j,k} f_{j,k}$. Then $\phi \in C^m$.

1. Note that if π_x is the Taylor polynomial of degree m at x, then for $t \in [x - h, x + h]$,

$$|\phi(t) - \pi_x(t)| \leq C_\alpha \cdot h^\alpha \cdot |\phi_x^{(m)}(t)|_{C^\delta[x-h,x+h]}, \tag{7.13}$$

where C_α is an absolute constant depending only on α. (For $m > 0$ this can be seen from the identity, for $t > x$,

$$\phi(t) - \pi_x(t) = 1/m! \int_x^t (u - x)^{m-1} (\phi^{(m)}(u) - \phi^{(m)}(x)) du.)$$

2. From the inequality

$$\sum_k (1 + a|k|)^\alpha (1 + |t - k|)^{-\ell} \leq B(\ell, \alpha) \cdot (1 + a|t|)^\alpha, \qquad \ell \geq \alpha + 2, t \in \mathbf{R},$$

and the rapid decay of $v_{j,k}$, we get

$$\sum_k (1 + a|k|)^\alpha v_{0,k}(t) \leq C \cdot (1 + a|t|)^\alpha, \qquad t \in \mathbf{R},$$

and similarly for other scales j. As a result, the inequality

$$|r(t)| \leq \epsilon \cdot C \cdot (1 + a|t - x|)^\alpha, \qquad t \in \mathbf{R},$$

for $x \in \mathbf{R}$, implies

$$|(P_j r)(t)| \leq \epsilon \cdot C' \cdot (1 + a|t - x|)^\alpha, \qquad t \in \mathbf{R}.$$

3. Combining items 1 and 2, we have that with $x = 2^{-j}(k + 1/2)$, $h = 2^{-j-1}$, $\epsilon = h^\alpha = 2^{-(j+1)\alpha}$, and $a = h^{-1}$,

$$\begin{aligned} |\phi(t) - \pi_x(t)| &\leq C \cdot h^\alpha \cdot |\phi^{(m)}|_{C^\delta[x-h,x+h]} \\ &\leq \epsilon \cdot (1 + |t - x|/h)^\alpha \cdot C \cdot |\phi^{(m)}|_{C^\delta[x-h,x+h]} \\ &\leq \epsilon \cdot (1 + |t - x|/h)^\alpha \cdot C \cdot A'_{\ell,\alpha} \cdot (1 + |x|)^{-\ell}. \end{aligned}$$

Hence, putting $r(t) \equiv \phi(t) - \pi_x(t)$ we have

$$|(P_j r)(t)| \leq \epsilon \cdot (1 + |t - x|/h)^\alpha \cdot C \cdot (1 + |x|)^{-\ell}$$

where C does not depend on x or j.

4. By assumption, for any polynomial of degree $m \le D$, we have $P_j\pi = \pi$ and so

$$\phi - (P_j\phi) = (\phi - \pi) - P_j(\phi - \pi).$$

In particular,

$$|\phi(t) - (P_j\phi)(t)| \le |r(t)| + |(P_jr)(t)|.$$

For $t \in [x-h, x+h]$, $(1 + |t-x|/h)^\alpha \le 2^\alpha$, so

$$|\phi(t) - (P_j\phi)(t)| \le \epsilon \cdot C \cdot (1+|t|)^{-\ell}, \qquad t \in [x-h, x+h].$$

As this is true, with the same constant C for each k, the same inequality holds for all $t \in \mathbf{R}$. Hence (7.11) follows.

Lemma 7.6 *Let $(w_{j,k})_k$ be any collection of continuous functions (not necessarily dilates and translates of a single function) satisfying the rapid decay estimates*

$$|\sigma_{j,k}w_{j,k}| \le A_\ell \cdot (1+|t|)^{-\ell} \qquad t \in \mathbf{R},\ \ell > 0,\ j,k \in \mathbf{Z}.$$

Then, for each $p \in (0, \infty]$

$$\|\sum_k c_{j,k}w_{j,k}\|_{L^p} \le C(p) \cdot 2^{j(1/2-1/p)} \cdot \|(c_{j,k})_k\|_{\ell^p} \tag{7.14}$$

where $C(p)$ is independent of j and depends on the system of functions $(w_{j,k})$ only through the constants (A_ℓ).

First, consider $p \le 1$:

$$\|f\|_p \le (\sum_k |c_{j,k}|^p \, \|w_{j,k}\|_p^p)^{1/p} \le \left(\sup_k \|w_{j,k}\|_p\right) \cdot \|(c_{j,k})_k\|_{\ell^p}.$$

The rapid decay estimates give that for $\ell = p^{-1} + 1$,

$$\|w_{j,k}\|_p \le 2^{j(1/2-1/p)} \cdot A_\ell \cdot (\int (1+|t|)^{-\ell p} dt)^{1/p} \equiv C(p) \cdot 2^{j(1/2-1/p)}.$$

Second, consider $p = \infty$

$$\|f(x)\| \le \sum_k |c_{j,k}| \, |w_{j,k}(x)| \le \|(c_{j,k})_k\|_{\ell^\infty} \cdot \sum_k |w_{j,k}(x)|$$

The rapid decay estimates give that for $\ell > 1$,

$$\sup_x \sum_k |w_{j,k}(x)| \le 2^{j/2} \cdot A_\ell \cdot \sup_x \sum_k (1+|x-k|)^{-\ell} \equiv C(\infty) 2^{j/2}$$

and so

$$\|f\|_\infty \le 2^{j/2} \cdot C(\infty) \cdot \|(c_{j,k})_k\|_{\ell^\infty}.$$

By interpolation we get for all $p \in (1, \infty)$

$$\|f\|_p \le C(1)^{1/p} C(\infty)^{1-1/p} \cdot 2^{j(1/2-1/p)} \cdot \|(c_{j,k})_k\|_{\ell^p}. \Box$$

Lemma 7.7 *Let $(w_{j,k})_k$ be a collection of continuous functions, not necessarily translates and dilates of a single function. Suppose that each space $\mathcal{F}_{j,k}$ consisting of restrictions to $[0,1]$ of functions $\sigma_{j,k}f$ where $f = \sum_k c_{j,k}w_{j,k}$, is finite dimensional. Suppose that the $\mathcal{F}_{j,k}$, although not necessarily the same, satisfy the uniform estimate*

$$N_p(\mathcal{F}_{j,k}) \leq N(p) < \infty, \qquad k \in \mathbf{Z}.$$

Finally suppose that the $w_{j,k}$ are interpolating, so that $w_{j,k}(k'/2^j) = 0$, unless $k' = k$ where $w_{j,k}(k/2^j) = 2^{j/2}$. Then, for each $p \in (0,\infty]$,

$$N(p)\|\sum_k c_{j,k}w_{j,k}\|_{L^p} \geq 2^{j(1/2-1/p)} \cdot \|(c_{j,k})_k\|_{\ell^p}. \tag{7.15}$$

Proof. Define

$$C_p(\mathcal{F}) = \sup\{|f(0)| : f \in \mathcal{F}, \int_0^1 |f(t)|^p dt \leq 1\}.$$

We get by dilation and the interpolation condition that

$$|c_{j,k}| \leq 2^{-j/2} \cdot C_p(\mathcal{F}_{j,k}) \cdot (2^{-j} \int_{k/2^j}^{(k+1)/2^j} |f(t)|^p dt)^{1/p}.$$

Summing across k,

$$\begin{aligned} \sum_k |c_{j,k}|^p &\leq 2^{-jp(1/2-1/p)} \cdot \max_k C_p(\mathcal{F}_{j,k})^p \cdot \sum_k \int_{2^{-j}k}^{2^{-j}(k+1)} |f|^p \\ &= \max_k C_p(\mathcal{F}_{j,k})^p \cdot 2^{-jp(1/2-1/p)} \cdot \|f\|_p^p. \end{aligned}$$

Now let $f \in \mathcal{F}_{j,k}$. Then

$$|f(t)| \leq \|f\|_{L^\infty[0,1]} \leq N_p(\mathcal{F}_{j,k})\|f\|_{L^p[0,1]}$$

Hence $C_p(\mathcal{F}_{j,k}) \leq N_p(\mathcal{F}_{j,k}) \leq N(p)$.

Lemma 7.8 *Let $(w_{j,k})_k$ be a collection of continuous functions, not necessarily translates and dilates of a single function. Suppose that each space $\mathcal{F}_{j,k}$ consisting of restrictions to $[0,1]$ of functions $\sigma_{j,k}f$ where $f = \sum_k c_{j,k}w_{j,k}$, is finite dimensional. Suppose that the $\mathcal{F}_{j,k}$, although not necessarily the same, satisfy the uniform estimate*

$$N_p(\mathcal{F}_{j,k}) \leq N(p) < \infty, \qquad k \in \mathbf{Z}.$$

Let S_j denote the sampling operator $(S_jf)_k = f(k/2^j)$. Then

$$\|S_jf\|_p \leq N(p) \cdot 2^{j/p} \cdot \|f\|_p \tag{7.16}$$

Proof. It is enough to prove the result at scale $j = 0$

$$|f(0)| \leq \|f\|_{L^\infty[0,1]} \leq N(p) \cdot \|f\|_{L^p[0,1]}.$$

Hence

$$\sum_k |f(k)|^p \leq N(p)^p \int |f(t)|^p dt$$

and we are done. □

7.5 Proofs for Triebel Equivalence

As usual for Triebel-Lizorkin spaces, the proof is a vector-valued version of the Besov case. Compare Frazier and Jawerth (1990). Define $d_j(t) = 2^{j\sigma}(Q_j f)(t)$ and $h_j(t) = 2^{j\sigma}\sum \alpha_{j,k}\chi_{j,k}(t)$, where $\chi_{j,k}$ is the L^2-normalized characteristic function $2^{j/2}1_{I_{j,k}}$, and the coefficients $\alpha_{j,k}$ are the interpolating wavelet coefficients of f. Define similar quantities for based on a smooth orthonormal wavelet expansion: $\bar{d}_j = 2^{j\sigma}(\bar{Q}_j f)(t)$ and $\bar{h}_j(t) = 2^{j\sigma}\sum \bar{\alpha}_{j,k}\chi_{j,k}(t)$, where now the $\bar{\alpha}_{j,k}$ are orthonormal wavelet coefficients of f.

It is known (Frazier and Jawerth, 1990) that for compactly supported orthogonal wavelets $\bar{\psi}_{j,k}$ of regularity $\bar{R}$ and moments through order $\bar{D}$ vanishing, if $\min(\bar{R},\bar{D}) > \sigma$ the norm $\|\bar{\theta}\|_{f^\sigma_{pq}}$ of the orthogonal wavelet coefficients is an equivalent norm for Triebel-Lizorkin space. For the case $p = 2$ see also Meyer (1990).

We will argue below that this norm is equivalent to a measure of approximation speed by orthogonal projection on $\bar{V}_j$, and establish that, if $\sigma > 1/p$, the interpolation projection on V_j has comparable speed of convergence.

Let $(f_j(t))_{j\ge j_0}$ be a sequence of functions in $L^p(\mathbf{R})$. Define the norm

$$\|(f_j(t))\|_{L^p(\ell^q)} \equiv \left(\int (\sum_{j\ge j_0} |f_j(t)|^q)^{p/q} dt\right)^{1/p}.$$

Note that

$$\|\bar{\theta}\|_{f^\sigma_{pq}} = \|(\bar{h}_j(t))\|_{L^p(\ell^q)} + \|(\beta_{j_0,k})\|_{L^p}.$$

The proof goes in two stages. The first shows the equivalence of $\|\bar{\theta}\|_{f^\sigma_{pq}}$ with

$$\|f\|_{\bar{F}^\sigma_{pq}} \equiv \|(\bar{d}_j(t))\|_{L^p(\ell^q)} + \|\bar{P}_{j_0} f\|_{L^p}$$

and that the corresponding unbarred quantity

$$\|f\|_{F^\sigma_{pq}} \equiv \|(d_j(t))\|_{L^p(\ell^q)} + \|P_{j_0} f\|_{L^p}$$

is equivalent to $\|\theta\|_{f^\sigma_{pq}}$.

In the second stage, we will argue that

$$\|(d_j(t))\|_{L^p(\ell^q)} \le C \cdot \|f\|_{\bar{F}^\sigma_{pq}} \tag{7.17}$$

and

$$\|(\bar{d}_j(t))\|_{L^p(\ell^q)} \le C \cdot \|f\|_{F^\sigma_{pq}}. \tag{7.18}$$

These inequalities imply the equivalences

$$\|\theta\|_{f^\sigma_{pq}} \asymp \|f\|_{F^\sigma_{pq}} \asymp \|f\|_{\bar{F}^\sigma_{pq}} \asymp \|\bar{\theta}\|_{f^\sigma_{pq}}, \tag{7.19}$$

and hence the theorem.

For the first stage we need two technical facts. Following Frazier and Jawerth (1990) we use the Hardy-Littlewood maximal function. For $r \in (0,\infty)$ define the Hardy-Littlewood Maximal Function

$$M_r(f)(t) = \left(\sup_{t\in[a,b]} (b-a)^{-1}\int_a^b |f|^r(u)du\right)^{1/r}.$$

This is typically larger than f; but the Fefferman-Stein Vector-Valued Maximal Inequality (Frazier and Jawerth, 1990, Appendix A) says that is not much larger if $r < min(p,q)$:

$$\|(M_r(f_j)(t))\|_{L^p(\ell^q)} \leq C(r,p,q)\|(f_j(t))\|_{L^p(\ell^q)}. \tag{7.20}$$

The second fact derives from Lemma 7.9 below: with a constant C independent of j or f or t

$$|d_j|(t) \leq C \cdot M_1(h_j)(t), \qquad t \in \mathbf{R}$$
$$|h_j|(t) \leq C \cdot M_1(d_j)(t), \qquad t \in \mathbf{R}.$$

Armed with these inequalities, we have, in one direction

$$\|(d_j(t))\|_{L^p(\ell^q)} \leq C \cdot \|(M_1(h_j)(t))\|_{L^p(\ell^q)} \leq C \cdot \|((h_j)(t))\|_{L^p(\ell^q)}.$$

In the other,

$$\|(h_j(t))\|_{L^p(\ell^q)} \leq C \cdot \|(M_1(d_j)(t))\|_{L^p(\ell^q)} \leq C \cdot \|((d_j)(t))\|_{L^p(\ell^q)}.$$

Combining these gives the equivalence

$$\|\theta\|_{f^\sigma_{pq}} \asymp \|f\|_{F^\sigma_{pq}}.$$

The equivalence of corresponding barred quantities is precisely analogous. This completes the first stage.

For the second stage, develop the telescoping sum

$$2^{-j\sigma} d_j = (I - P_j)f = (I - P_j)(\bar{P}_{j_0} f + \sum_{j' \geq j_0} \bar{Q}_{j'} f)$$

which gives

$$d_j = 2^{j\sigma} \cdot (I - P_j)\bar{P}_{j_0} f + \sum_{j' \geq j_0} 2^{(j-j')\sigma}(I - P_j)\bar{d}_{j'}.$$

Now by Lemma 7.5

$$|(I - P_j)\varphi_{j_0,0}|(t) \leq 2^{-(j-j_0)\min(D,\bar{R})} \cdot w_{j,k;\ell}(t) \cdot A_\ell$$

Hence, with $\bar{P}_{j_0} f = \sum_k \bar{\beta}_{j_0,k}\bar{\varphi}_{j_0,k}$,

$$|(I - P_j)\bar{P}_{j_0} f|(t) \leq 2^{-(j-j_0)\min(D,\bar{R})} \cdot A_\ell \cdot \sum_k |\bar{\beta}_{j_0,k}| \cdot w_{j,k;\ell}(t).$$

We let $r < p$ be a constant to determine later, and set $w_{j,k;\ell}(t) = (1 + |2^j t - k|)^{-\ell} 2^{j/2}$, with $\ell = 2 + 1/r$. Applying Lemma 7.9 ,

$$\sum_k |\bar{\beta}_{j_0,k}| \cdot w_{j,k;\ell}(t) \leq C \cdot M_r(\sum_k \bar{\beta}_{j_0,k}\bar{\varphi}_{j_0,k})(t) \qquad t \in \mathbf{R},$$

and so

$$(I - P_j)(\bar{P}_{j_0} f)2^{j\sigma} \leq 2^{j_0\sigma} 2^{-(j-j_0)(\min(D,\bar{R})-\sigma)} \cdot C \cdot M_r(\bar{P}_{j_0} f)(t) = \tilde{d}_j(t), \qquad \text{say.}$$

Similarly, for $j_0 \le j' \le j$,

$$|(I-P_j)\bar{\psi}_{j',0}|(t) \le 2^{-(j-j_0)\min(D,\bar{R})} \cdot w_{j,k;\ell}(t) \cdot A_\ell$$

whence

$$(I-P_j)(\bar{d}_{j'}) \le 2^{-(j'-j)(\min(D,\bar{R})-\sigma)} \cdot C \cdot M_r(\bar{d}_{j'})(t).$$

Finally, we need Lemma 7.10. For $j' \ge j$, it gives

$$|(I-P_j)\bar{d}_{j'}|(t) \le 2^{(j'-j)/r} \cdot M_r(\bar{d}'_j)(t). \tag{7.21}$$

Combining these estimates, we get

$$|d_j(t)| \le |\tilde{d}_j(t)| + C \cdot \sum_{j' \ge j_0} M_r(\bar{d}_{j'})(t) \cdot 2^{-|j'-j| \cdot \delta}$$

where $\delta = \min(\min(D,\bar{R}) - \sigma, \sigma - 1/r)$. If $\sigma > 1/p$ we can choose $r < p$ so that $\sigma - 1/r > 0$. Then, putting $\bar{m}_{j'}(t) = M_r(\bar{d}_{j'})(t)$ and $H(j,j') = C \cdot 2^{-|j-j'| \cdot \delta}$, we have

$$\|H\bar{m}\|_{\ell^q}(t) \le C \cdot \|\bar{m}\|_{\ell^q}(t)$$

and so

$$\|(d_j(t))\|_{L^p(\ell^q)} \le \|(\tilde{d}_j(t))\|_{L^p(\ell^q)} + C \cdot \|(\bar{m}_j(t))\|_{L^p(\ell^q)}.$$

We conclude by Fefferman-Stein that

$$\begin{aligned} \|(d_j(t))\|_{L^p(\ell^q)} &\le \|(\tilde{d}_j(t))\|_{L^p(\ell^q)} + C \cdot \|(\bar{d}_j(t))\|_{L^p(\ell^q)} \\ &\le C(\|\bar{P}_{j_0} f\|_p + \|(\bar{d}_j(t))\|_{L^p(\ell^q)}) = C\|f\|_{\dot{F}^\sigma_{pq}} \end{aligned}$$

Similar arguments yield the companion inequality (7.18).

7.5.1 Lemmas Supporting Triebel Equivalence

Lemma 7.9 *Let* $w_{j,k;\ell} = 2^{j/2}(1 + |2^j t - k|)^{-\ell}$. *Suppose that the sums* $\sum_k \alpha_{j,k} w_{j,k}$ *are piecewise finite-dimensional. Let* $\mathcal{F}_{j,k}$ *denote the space of renormalized restrictions* $\sigma_{j,k} f$ *where* $f = \sum_k \alpha_{j,k} w_{j,k}$. *Suppose that these spaces obey the uniform estimate*

$$N_r(\mathcal{F}_{j,k}) \le N(r).$$

Suppose in addition that the $w_{j,k}$ *are interpolating, so that*

$$\alpha_{j,k} = f(k/2^j) 2^{-j/2}.$$

Then for $\ell > 1 + 1/r$, *and* $r \le 1$

$$\sum_k |\alpha_{j,k}| w_{j,k;\ell}(t) \le C \cdot M_r(\sum_k \alpha_{j,k} w_{j,k})(t)$$

where the constant is an absolute multiple of $N(r)$.

Proof.

$$|\alpha_{0,k}| \le \|f\|_{L^\infty[k,k+1]} \le N(r) \cdot \|f\|_{L^r[k,k+1]}.$$

and

$$\sum_k |\alpha_{0,k}| w_{j,k;\ell}(0) \le N(r) \cdot A_\ell \cdot \sum_k \|f\|_{L^r[k,k+1]} (1+|k|)^{-\ell}.$$

The result now follows from

$$\sum_k \|f\|_{L^r[k,k+1]} (1+|k|)^{-\ell} \le C \cdot M_r(f)(0)$$

which is an easy exercise. (Compare Frazier and Jawerth (1990) Appendix A).

Lemma 7.10 *Suppose that the vector space $\bar{W}_j$ of sums $\sum_k \bar{\alpha}_{j,k} \bar{\psi}_{j,k}$ is piecewise finite-dimensional, with uniform bound $N(r)$, and that the terms are uniformly of rapid decay. For $\ell = 2 + 1/r$, and $r \in (0, \infty)$*

$$|P_j(\sum_k \bar{\alpha}_{j',k} \bar{\psi}_{j',k})(t) \le C \cdot 2^{(j'-j)/r} M_r(\sum_k \bar{\alpha}_{j',k} \bar{\psi}_{j',k})(t), \qquad j' > j.$$

where the constant C may be chosen independently of j' and j as an absolute multiple of $N(r)$.

Proof. Set $f_{j'}(t) = \sum_k \bar{\alpha}_{j',k} \bar{\psi}_{j',k}$. By rapid decay

$$|P_j f_{j'}|(t) \le A_\ell \sum_k |f_{j'}(k/2^j)| w_{j,k;\ell}(t)$$

Now suppose that $k/2^{j'} < t < (k+1)/2^{j'} \le k'/2^{j'} \le u \le (k'+1)/2^{j'}$. Set $\tilde{t} = k/2^{j'}$, $\tilde{u} = (k'+1)/2^{j'}$ then by finite-dimensionality,

$$\begin{aligned}
|f_{j'}|(u) &\le N(r) \cdot (2^{j'} \int_{I_{j',k}} |f_{j'}|^r(t) dt)^{1/r} \\
&\le N(r) \cdot 2^{j'/r} \cdot (\tilde{u} - \tilde{t})^{1/r} \cdot ((\tilde{u} - \tilde{t})^{-1} \int_{\tilde{t}}^{\tilde{u}} |f_{j'}|^r(t) dt)^{1/r} \\
&\le N(r) \cdot 2^{j'/r} \cdot (\tilde{u} - \tilde{t})^{1/r} \cdot M_r(f_{j'})(t) \\
&\le N(r) \cdot (k' - k + 1)^{1/r} \cdot 2^{(j'-j)/r} \cdot M_r(f_{j'})(t).
\end{aligned}$$

Hence,

$$\sum_k |f_{j'}(k/2^j)| w_{j,k;\ell}(t) \le C \cdot 2^{(j'-j)/r} \cdot M_r(f_{j'})(t) \cdot \sum_k (|k|+1)^{1/r} (1+|k|)^{-\ell},$$

and the Lemma follows. □

7.6 Proofs for Section 3.2

Besov Equivalence: Our proof of Theorem 2.7 is actually quite general and adapts without essential change to give a proof of Theorem 3.6. One adapts the proof to life on the interval by substituting $L^p[0,1]$ for L^p, $P_j^{[0,1]}$ for P_j and $V_j[0,1]$ for V_j, and pursues the same proof line-by line. In the end, the proof rests on the establishment of the four conclusions (7.12), (7.14), (7.15), (7.16) offered by Lemmas 7.5-7.8. Those lemmas apply equally well in the interval case; one only has to establish the needed uniform estimates

As we have seen, the wavelets $\psi_{j,k}^{[]}$ and $\varphi_{j,k}^{[]}$ are of uniformly compact support. Let $\sigma_{j,k}$ denote the standardization operator which maps f into $2^{-j/2}f(2^{-j}(t-k))$. Then we have, for constants A,B,

$$|\sigma_{j,k}\psi_{j,k}^{[]}| \le A \cdot 1_{[-B,B]}, \qquad j \ge j_0, k \in K_j.$$

$$|\sigma_{j,k}\varphi_{j_0,k}^{[]}| \le A \cdot 1_{[-B,B]}, \qquad j \ge j_0, k \in K_j.$$

The same localization bounds hold for derivatives through order $\lfloor R \rfloor$. Hence the needed uniformities for Lemmas 7.5-7.6 hold, and the key conclusions (7.12)-(7.14) follow.

Moreover, the spaces $\mathcal{V}_{j,k}$ given by the restriction of renormalized functions

$$(\sigma_{j,k}f)|_{[0,1]}, \qquad f \in V_j[0,1]$$

are all finite-dimensional, because of the uniform support bounds. Similarly, the spaces $\mathcal{W}_{j,k}$ given by

$$(\sigma_{j,k}f)|_{[0,1]}, \qquad f \in W_j[0,1]$$

are finite dimensional. Even looking over all scales $j \ge j_0$ there are only a finite number of different such spaces, because they are made of combinations of restrictions $\varphi(\cdot - k')|_{[0,1]}$, $\varphi_k^{\#}(\cdot - k')|_{[0,1]}$, $\varphi_k^{\flat}(\cdot - k')|_{[0,1]}$, $\psi(\cdot - k')|_{[0,1]}$, $\psi_k^{\#}(\cdot - k')|_{[0,1]}$, $\psi_k^{\flat}(\cdot - k')|_{[0,1]}$. Hence, recalling the definition of $N_p(\mathcal{F})$ in section 2.2, we have

$$N_p^{[]} = \sup_{j \ge j_0} max(N_p(\mathcal{V}_{j,k}), N_p(\mathcal{W}_{j,k})) < \infty.$$

Consequently, the needed uniform finite-dimensionality holds, and we may apply Lemmas 7.7 and 7.8 to get the conclusions (7.15), (7.16).

With these technical underpinnings established, and the relevant notational substitutions made, the proof of Theorem 2.7 becomes also a proof of Theorem 3.6.

Triebel Equivalence: The argument is similar; one adapts the proof of Theorem 2.8 to the interval case by making the obvious substitutions in notation. The needed uniformities for the technical lemmas are established by the arguments we just gave in the Besov case.

7.7 Proofs for Section 3.3

Proof of Theorem 3.8 We first note that there is $J = J(\sigma, p, \|\theta\|_{b_{p,q}^\sigma})$ so that

$$\alpha_{j,k}^{(\epsilon)} = 0, \qquad j \ge J, k \in K_j.$$

Indeed,

$$\|(\alpha_{j,k})_k\|_{\ell^\infty} \le \|(\alpha_{j,k})_k\|_{\ell^p} \le 2^{-j(\sigma+1/2-1/p)}\|\theta\|_{b_{p,q}^\sigma}$$

so defining J as the unique real root of

$$2^{-j(\sigma+1/2-1/p)}\|\theta\|_{b^\sigma_{p,q}} = \epsilon 2^{-j/2},$$

does the trick. Evidently,

$$J(\sigma, p, \|\theta\|_{b^\sigma_{p,q}}) = (\sigma - 1/p)^{-1} \cdot \log_2(\|\theta\|_{b^\sigma_{p,q}}/\epsilon).$$

Hence,

$$\|f - f^{(\epsilon)}\|_\infty \le \sum_{j=j_0}^{\lfloor J \rfloor} \|\sum_k (\alpha_{j,k} - \alpha^{(\epsilon)}_{j,k})\psi^{[]}_{j,k}\|_\infty + \sum_{j \ge J} \|\sum_k \alpha_{j,k}\psi^{[]}_{j,k}\|_\infty$$

Let M_1 be the maximum number of wavelets at any level j "touching" any single point $x \in [0,1]$. Let M_2 be the maximum L^∞ norm of any standardized wavelet (ψ or φ, boundary-corrected or not). Then for any set of coefficients $(c_{j,k})_k$,

$$\|\sum_k c_{j,k}\psi^{[]}_{j,k}\|_\infty \le M_1 \cdot M_2 \cdot \|(c_{j,k})_k\|_\infty \cdot 2^{j/2}$$

Applying this at levels $j \le J$ with $c_{j,k} = \alpha_{j,k} - \alpha^{(\epsilon)}_{j,k}$, and noting that $|\alpha_{j,k} - \alpha^{(\epsilon)}_{j,k}| \le \epsilon \cdot 2^{-j/2}$, gives

$$\begin{aligned}\sum_{j=j_0}^{\lfloor J \rfloor} \|\sum_k (\alpha_{j,k} - \alpha^{(\epsilon)}_{j,k})\psi^{[]}_{j,k}\|_\infty &\le \sum_{j=j_0}^{\lfloor J \rfloor} M_1 \cdot M_2 \cdot \epsilon \cdot 2^{-j/2} \cdot 2^{j/2} \\ &\le M_1 \cdot M_2 \cdot \epsilon \cdot J\end{aligned}$$

Applying this at levels $j \ge J$ gives

$$\sum_{j \ge J} \|\sum_k \alpha_{j,k}\psi^{[]}_{j,k}\|_\infty \le \sum_{j \ge J} M_1 \cdot M_2 \cdot \|(\alpha_{j,k})\|_{\ell^\infty} \cdot 2^{j/2}.$$

Using

$$\|(\alpha_{j,k})\|_{\ell^\infty} \cdot 2^{j/2} \le \|\theta\|_{b^\sigma_{p,q}} \cdot 2^{-j(\sigma-1/p)}$$

gives

$$\begin{aligned}\sum_{j \ge J} \|\sum_k \alpha_{j,k}\psi^{[]}_{j,k}\|_\infty &\le M_1 \cdot M_2 \cdot \|\theta\|_{b^\sigma_{p,q}} \cdot 2^{-J(\sigma-1/p)}/(1 - 2^{-(\sigma-1/p)}) \\ &= M_1 \cdot M_2 \cdot \epsilon/(1 - 2^{-(\sigma-1/p)}).\end{aligned}$$

Combining results:

$$\|f - f^{(\epsilon)}\|_\infty \le M_1 \cdot M_2 \cdot \epsilon \cdot ((1 - 2^{-(\sigma-1/p)})^{-1} + J)$$

which is (3.6).

We now count the number of nonzero wavelet coefficients. Note that

$$\delta^p \#\{k : |a_{j,k}| \ge \delta\} \le \|(a_{j,k})_k\|_p^p$$

Hence, putting $\delta = \epsilon \cdot 2^{-j/2}$,

$$\begin{aligned} N(\epsilon) &\leq \sum_{j=j_0}^{J} \|(a_{j,k})_k\|_p^p / (\epsilon 2^{-j/2})^p \\ &\leq \|\theta\|_{b^\sigma_{p,q}} \cdot \epsilon^{-p} \cdot \sum_{j=j_0}^{J} 2^{j(\sigma-1/p)p} = \|\theta\|_{b^\sigma_{p,q}} \cdot \epsilon^{-p} \cdot c_2(\sigma, p) \end{aligned}$$

which is (3.7).

Proof of Lemma 3.9.

The best $\ell^\infty(n+m)$ approximation to a vector v by a vector v_n with n nonzero entries is obtained by setting to zero the m entries which are smallest in absolute value. The resulting approximation error is the size of the largest of those entries which were set to zero. Considering the example where all elements of the vector have the same absolute value, we see that the error in this approximation can be as large as that of the ℓ^∞ norm of the vector v itself. Hence the minimax error over a hypercube $\{v : \|v\|_{\ell^\infty(n+m)} \leq \rho\}$ is

$$\sup_{\|v\|_{\ell^\infty(n+m)} \leq \rho} \inf_{v_n : \#\{i : (v_n)_i \neq 0\} \leq n} \|v - v_n\|_\infty = \rho.$$

The idea is to embed a 2^{j+1}-dimensional hypercube into $B^\sigma_{p,q}[0,1]$, $2^{j+1} > n$ and use the above observation to bound the L^∞ error on its approximation by an n-term sum.

Consider the ℓ^p ball

$$\Theta_{j,p}(r) = \{\theta : \|(\alpha_{j,k})_k\|_{\ell^p} \leq r \qquad \alpha_{j',k'} = 0, j' \neq j;\ \beta_{j_0,k} = 0 \forall k\}.$$

Now evidently

$$\|(\alpha_{j,k})_k\|_{\ell^p} \leq 2^{-j(\sigma+1/2-1/p)} \|\theta\|_{b^\sigma_{p,q}}.$$

This tells us that the class of objects with $B^\sigma_{p,q}[0,1]$ norm $\leq B$ contains an ℓ^p-ball $\Theta_{j,p}(r_j)$of radius $r_j = C \cdot B \cdot 2^{-j(\sigma+1/2-1/p)}$. As $\|(\alpha_{j,k})_k\|_p \leq 2^{j/p} \cdot \|(\alpha_{j,k})_k\|_\infty$, the class also contains the hypercube $\Theta_{j,\infty}(\rho_j)$ with radius $\rho_j = r_j \cdot 2^{-j/p} = C \cdot B \cdot 2^{-j(\sigma+1/2)}$.

Fix now j so that $2^j \leq n < 2^{j+1}$. Let $\mathcal{F}_j$ be the collection of functions with wavelet coefficients in the hypercube $\Theta_{j,\infty}(\rho_j)$ with ρ_j as above. The problem of approximating this collection furnishes a lower bound on the problem of approximating the (larger) ball, because of

$$\sup_{f \in \mathcal{F}} \inf_{f_n \in \mathcal{S}_n} \|f - f_n\|_\infty \geq \sup_{f \in \mathcal{F}_j} \inf_{f_n \in \mathcal{S}_n} \|f - f_n\|_\infty,$$

which expresses the setwise monotonicity of the minimax error.

Now let f_n be any approximant with at most n nonvanishing wavelet coefficients, and write it as

$$f_n = P_j f_n + Q_j f_n + (I - P_{j+1}) f_n = f_n^< + f_n^= + f_n^>,$$

say. Now put for short $\|g\|_j = \|(g(k/2^{j+1}))_{k \in K_{j+1}}\|_{\ell^\infty}$. Then $\|f_n^>\|_j = 0$, $\|f\|_{j-1} = \|f_n^=\|_{j-1} = 0$ and

$$\|f_n - f\|_j = \|(f_n^< + f_n^=) - f\|_j \geq \max(\|f_n^<\|_{j-1}, \|f_n^= - f\|_j - \|f_n^<\|_j)$$

As $f_n^< \in V_j[0,1]$, $\|f_n^<\|_{j-1} \geq c\|f_n^<\|_{L^\infty(\mathbf{R})} \geq c\|f_n^<\|_j$ and $\max(a, b - a\cdot c) \geq b/(1+c)$, so

$$\|f_n - f\|_j \geq \max(\|f_n^<\|_{j-1}, \|f_n^= - f\|_j - c\|f_n^<\|_{j-1}) \geq \|f_n^= - f\|_j/(1+c).$$

Finally, $f_n^=$ is a sum of $\leq n$ terms $\sum_k \alpha_{j,k}^{[n]}\psi_{j,k}^{[]}$. Hence

$$\|f_n^= - f\|_j \geq c \cdot 2^{-j/2} \cdot \max_{k\in K_j} |(\alpha_{j,k} - \alpha_{j,k}^{[n]})|.$$

Now by the above argument

$$\sup_{\mathcal{F}_j} \inf_{f_n \in \mathcal{S}_n} \max_{k\in K_j} |(\alpha_{j,k} - \alpha_{j,k}^{[n]})| = \rho_j$$

and so

$$\sup_{f\in\mathcal{F}_j} \inf_{f_n\in n} \|f - f_n\|_\infty \geq c\rho_j \cdot 2^{j/2} \geq cn^{-\sigma}. \qquad \square$$

7.8 Proof for Section 4.2

Besov Equivalence: The proof of Theorem 2.7 also furnishes a proof of Theorem 4.5, once we make the substitutions of $\tilde{V}_j$ for V_j and so on. In the end, the proof reduces the result to the establishment of the four conclusions (7.12), (7.14), (7.15), (7.16) offered by Lemmas 7.5-7.8. Those lemmas apply equally well in the hybrid case; one only has to establish the needed uniform estimates.

Because of the uniform bounds of Lemma 4.2, all the needed localizations and smoothness bounds for Lemmas 7.5-7.6 follow easily. Also, Lemma 7.8 applies immediately, because it refers to sampling of the standard orthogonal wavelet bases and was already proved.

It only remains to establish a conclusion like (7.15). Now the argument is slightly different than before.

Lemma 7.11 *Let the hybrid wavelets be defined by linked interpolating and orthogonal wavelets. Then, for a constant C and for each $p \in (0,\infty]$,*

$$C \cdot N(p)\|\sum_k \alpha_{j,k}\tilde{\psi}_{j,k}\|_{L^p} \geq 2^{j(1/2-1/p)} \cdot \|(\alpha_{j,k})_k\|_{\ell^p}. \tag{7.22}$$

and similarly for sums involving φ. The constant C depends on p but not on j_0 or j_1 or f.

Proof. Let

$$f = \sum_k \alpha_{0,k}\tilde{\psi}_{0,k}^{[\ell]}.$$

Then, for appropriate constants $u_{k'}^{(\ell,k)}$,

$$\alpha_{0,k} = 2^{-\ell/2}\sum_{k'} u_{k'}^{(\ell,k)} f(k'/2^\ell).$$

Concerning the constants $u_{k'}^{(\ell,k)}$ we make two observations. First, by the compact support of the wavelet filters,

$$\#\{k' : u_{k'}^{(\ell,k)} \neq 0\} \leq C_1 2^\ell, \qquad \forall k,\ell, \tag{7.23}$$

and

$$\#\{k : u_{k'}^{(\ell,k)} \neq 0\} \leq C_2, \qquad \forall k', \ell. \tag{7.24}$$

Second, by orthonormality of the pyramid filtering operator

$$\sum_{k'} (u_{k'}^{(\ell,k)})^2 = 1, \qquad \forall k, \ell; \tag{7.25}$$

and

$$\max_{k'} |u_{k'}^{(\ell,k)}| \leq C_3 2^{-\ell/2}, \qquad \forall k, \ell. \tag{7.26}$$

Suppose $p = 1$. Then (7.24)-(7.26) give

$$\begin{aligned}
\sum_k |\alpha_{0,k}| &\leq \sum_{k'} \left(|f(k'/2^\ell)| \cdot 2^{-\ell/2} \sum_k |u_{k'}^{(\ell,k)}| \right) \\
&= \left(2^{-\ell/2} \sup_{k'} \sum_k |u_{k'}^{(\ell,k)}| \right) \cdot \left(\sum_k |f(k'/2^\ell)| \right) \\
&\leq \left(C_2 \cdot C_3 \cdot 2^{-\ell} \right) \cdot \left(N(1) \int |f(t)| dt \cdot 2^\ell \right) \\
&= C_2 \cdot C_3 \cdot N(1) \cdot \int |f(t)| dt
\end{aligned}$$

If $0 < p < 1$ arguing similarly gives

$$\sum_k |\alpha_{0,k}|^p \leq C_2 \cdot C_3 \cdot N(p)^p \cdot \int |f(t)|^p dt.$$

Suppose $p = \infty$. Then (7.24), (7.25) and Cauchy-Schwartz give $\sum_{k'} |u_{k'}^{(\ell,k)}| \leq C_1^{1/2} \cdot 2^{\ell/2}$ and

$$\begin{aligned}
\sup_k |\alpha_{0,k}| &\leq \sup_k \sum_{k'} |f(k'/2^\ell)| \cdot 2^{-\ell/2} |u_{k'}^{(\ell,k)}| \\
&\leq \left(2^{-\ell/2} \sup_k \sum_{k'} |u_{k'}^{(\ell,k)}| \right) \cdot \left(\sup_k |f(k'/2^\ell)| \right) \\
&\leq C_1^{1/2} \|f\|_\infty.
\end{aligned}$$

Hence, (7.22) holds with $C = \max(C_2 \cdot C_3, C_1)$.

Triebel Equivalence: The proof of Theorem 2.8 equally well gives a proof of 4.6. The needed uniform estimates necessary to make the proof go through were already established in the Besov case.

7.9 Proof for Section 5.2

The theorems follow from combining the reasoning for sections 3.2 and 4.2. The hybrid wavelets for the interval are smooth and uniformly localized by the observation that they are finite linear combinations, with bounded numbers of terms, of neighboring hybrid wavelets for the line, together with Lemma 4.4. They obey a uniform bound like (7.22) by the same reasoning. Equipped with these technical tools, the proofs for the Besov and Triebel cases given earlier go through line-by-line.

References

[1] Antonini, M., Barlaud, M., Mathieu, P. and Daubechies, I. (1991) Image coding using wavelet transforms, *IEEE Proc. Acoustics, Speech, Signal Processing*, to appear.

[2] Aldroubi, A. and Unser, M. (1992) Polynomial Splines and Wavelets – a signal processing perspective. in *Wavelets: a tutorial in theory and applications* ed C.K. Chui. Academic Press, San Diego.

[3] Cohen, A., Daubechies, I., and Feauveau, J.C. (1989) Biorthogonal bases of compactly supported wavelets. *Comm. Pure and Apl. Math.*, to appear 1991.

[4] Cohen, A., Daubechies, I., Jawerth, B., and Vial, P. (1992). Multiresolution analysis, wavelets, and fast algorithms on an interval. To appear, *Comptes Rendus Acad. Sci. Paris* (A).

[5] Daubechies, I. (1988) Orthonormal bases of compactly supported wavelets. *Comm. Pure and Apl. Math.*, **41**, 909-996.

[6] Daubechies, I and Lagarias, J. (1991) Two-scale difference equations, I. Global regularity of solutions; II. Local Regularity, infinite products of matrices, and fractals. *SIAM J. Math. Anal* **22** 1388-1410.

[7] Daubechies, I. (1992) *Ten Lectures on Wavelets.* Philadelphia: SIAM.

[8] Deslauriers, G. and Dubuc, S. (1987) Interpolation dyadique. in *Fractals, Dimensions non-entières et applications.* Paris: Masson.

[9] Deslauriers, G. and Dubuc, S. (1989) Symmetric iterative interpolation processes. *Constructive Approximation*, **5**, 49-68.

[10] Dubuc, S. (1986) Interpolation through an iterative scheme. *J. Math. Anal. and Appl.* **114** 185-204.

[11] DeVore, R.A. (1989) Degree of nonlinear approximation. *Approximation Theory VI*, Chui and Schumaker, eds. Academic.

[12] DeVore, R.A., Jawerth, B. and Lucier, B.J. (1990) Surface Compression. *Computer-Aided Geometric Design.* To Appear.

[13] DeVore, R.A., Jawerth, B., and Lucier, B.J. (1992) Image compression through wavelet transform coding. *IEEE Trans. Info Theory.* **38**,2,719-746.

[14] DeVore, R.A., Jawerth, B. and Popov, V. (1990) Compression of Wavelet decompositions. Preprint.

[15] DeVore, R.A., and Popov, V. (1988). Interpolation of Besov spaces. *Trans. Amer. Math. Soc. 305*, **1**, 397-414.

[16] Domsta, J. (1976). Approximation by spline interpolating bases. *Studia. Math.* **58** 223-237.

[17] Donoho, D.L. & Johnstone, I.M. (1992a). Ideal spatial adaptation via wavelet shrinkage.

[18] Donoho, D.L. & Johnstone, I.M. (1992b). New minimax theorems, thresholding, and adaptation.

[19] Donoho, D.L. & Johnstone, I.M. (1992d). Adapting to unknown smoothness by wavelet shrinkage.

[20] Jaffard, S. (1990) Propriétés de matrices "bien localisées" près de leur diagonale et quelques applications. *Ann. Inst. Henir Poincaré* **7** 461-476.

[21] Lemarié, P.G. and Malgouyres, G. (1989) Ondelettes sans peine. Prépublication Université de Paris-Sud.

[22] Meyer, Y. (1990a). *Ondelettes et opérateurs I: Ondelettes.* Hermann, Paris.

[23] Meyer, Y. (1990b). *Ondelettes et opérateurs II: Opérateurs de Calderón-Zygmund.* Hermann, Paris.

[24] Meyer, Y. (1991) Ondelettes sur l'intervalle. *Revista Mat. Ibero-Americana.*

[25] Peetre, J. (1976) *New thoughts on Besov Spaces.* Duke Univ. Mathematics Series 1.

[26] Saito, N. and Beylkin, G. (1992) Multiresolution representations using the autocorrelation functions of compactly supported wavelets.

[27] Schoenberg, I.J. (1972). Cardinal interpolation and spline functions: II. Interpolation of data of power growth. *Journ. Approx. Theory* **6** 404-420.

[28] Schonefeld, S. (1972). Schauder Bases in the Banach spaces $C^k(T^q)$. *Trans. Amer. Math. Soc.* **165** 309-318.

[29] Sickel, Winfried (1992) Characterization of Besov-Triebel-Lizorkin spaces via Approximation by Whittaker Cardinal Series. *Constructive Approximation* **8** (3).

[30] Subbotin, Y.N. (1972). Spline approximants as smooth bases in $C(0, 2\pi)$. *Mat. Zametki* **12** 45-51 (In Russian).