

MULTIRESOLUTION BASED ON WEIGHTED AVERAGES OF THE HAT FUNCTION I: LINEAR RECONSTRUCTION TECHNIQUES*

FRANCESC ARÀNDIGA[†], ROSA DONAT[‡], AND AMI HARTEN[§]

Abstract. In this paper we analyze a particular example of the general framework developed in [A. Harten, *SIAM J. Numer. Anal.*, 33 (1996) pp. 1205–1256], the case in which the discretization operator is obtained by taking local averages with respect to the hat function. We consider a class of reconstruction procedures which are appropriate for this multiresolution setting and describe the associated prediction operators that allow us to climb up the ladder from coarse to finer levels of resolution. In Part I we use data-independent (linear) reconstruction techniques as our approximation tool. We show how to obtain multiresolution transforms in bounded domains and analyze their stability with respect to perturbations.

Key words. multiscale decomposition, discretization, reconstruction

AMS subject classifications. 41A05, 41A15, 65O15

PII. S0036142996308770

1. Introduction. Multiresolution representations have become effective tools for analyzing the information contents of a given signal. In this respect, the recent development of the theory of wavelets (see, e.g., [8] and references therein) has been a giant leap toward local scale decompositions and has already had great impact on several fields of science.

Multiscale techniques do have an important role in numerical analysis. A wavelet type decomposition of a function is used to reduce the cost of many numerical algorithms either by applying it to the numerical solution operator to obtain an approximate sparse form [4, 14, 21, 2] or by applying it to the numerical solution itself to obtain an approximate reduced representation in order to solve for fewer quantities [19].

The building block of the wavelet theory is a square-integrable function whose dilates and translates form an orthonormal base of the space of square-integrable functions. Such uniformity leads to conceptual difficulties in extending wavelets to bounded domains and general geometries. Moreover, it is impossible to obtain adaptive (data-dependent) multiresolution representations which fit the approximation to the local nature of the data. Adaptivity is possible only by admitting “redundant” representations.

A combination of ideas from multigrid methods, numerical solution of conservation laws, hierarchical bases of finite element spaces, subdivision schemes of computer-aided design (CAD) and, of course, the theory of wavelets led A. Harten to the devel-

*Received by the editors September 3, 1996; accepted for publication (in revised form) January 16, 1998; published electronically December 2, 1998.

<http://www.siam.org/journals/sinum/36-1/30877.html>

[†]Departament de Matemàtica Aplicada, Universitat de València, València, Spain (arandiga@godella.matapl.uv.es). The research of this author was supported in part by DGICYT PB94-0987, in part by a University of Valencia grant, and in part by ONR-N00014-95-1-0272.

[‡]Departament de Matemàtica Aplicada, Universitat de València, València, Spain (donat@rocafort.matapl.uv.es). The research of this author was supported by in part by DGICYT PB94-0987, in part by a grant from the Generalitat Valenciana, and in part by ONR-N00014-95-1-0272.

[§]The author is deceased. Former address: School of Mathematical Sciences, Tel-Aviv University, Tel Aviv, Israel, and Department of Mathematics, UCLA, Los Angeles, CA 90024-1555.

opment of a “general framework” for multiresolution representation of discrete data.

Multiresolution representations à la Harten are constructed using two operators, decimation and prediction, which connect adjacent resolution levels. In turn, these operators are defined with two basic building blocks: the discretization and reconstruction operators. The former obtains discrete information from a given signal (belonging to a particular function space), and the latter produces an “approximation,” in the same function space, from the discrete information contents of the original signal.

Because of the essential role played by the discretization and reconstruction operators in Harten’s framework, building multiresolution schemes that are appropriate for a given application becomes a task which is very familiar to a numerical analyst. First one identifies a sense of discretization which is appropriate for the given application. Then one solves a problem in approximation theory.

The discretization operator specifies the process of generation of discrete data and, thus, it determines the nature of the discrete data to be analyzed. When reinterpreted within Harten’s framework, the discretization operator in the wavelet theory is obtained by taking local averages against the scaling function.

The strict requirements of the wavelet theory rule out many scaling functions that provide, nevertheless, appropriate discretization settings in many situations. For example, weighted averages against the δ -function lead to point-value discretizations, a natural discretization procedure for continuous functions which is widely used within the numerical analysis community. However, the δ -function is not square integrable; thus it is never considered as the basic building block of a wavelet-type multiresolution decomposition. The box function is a classical example in wavelet theory: It is square integrable and it leads to a wavelet basis, the Haar basis. From the point of view of numerical analysis, weighted averages against the box function provide the natural discretization setting for discontinuous functions. Following this line of reasoning, the hat function leads to a natural discretization setting for functions with a finite number of δ -type singularities. The space of such *generalized functions* is used in the design of vortex methods for the numerical solution of partial differential equations.

It is shown in [15, 16, 17] how to obtain stable multiscale decompositions using the δ -function and the box function in the discretization process. In this paper and its sequel (henceforth Part II) we study the case in which the discretization of a given function is carried through by taking weighted averages with respect to the hat function. We show how the theory in [17] can be applied to obtain appropriate decimation and prediction operators that allow for multiresolution representations of a set of data that can be considered as hat-averages of a given function.

In this paper (Part I) we shall limit ourselves to linear reconstruction techniques, which lead to linear prediction operators and to multiresolution transforms which can be reinterpreted as a change of basis functions. However, the relation between prediction and reconstruction gives Harten’s framework a degree of flexibility lacking in standard wavelet theory. In particular, in the linear case it is conceptually a simple matter to treat the bounded domain case (no-periodicity): It is enough to design the reconstruction operator in such a way that it solves an approximation problem that fits the boundary data.

However, the reconstruction, hence the prediction, operator does not need to be linear; adaptivity can thus be introduced into the multiresolution scheme by choosing nonlinear, adaptive reconstruction techniques. The nonlinear case is studied in Part II [1].

To apply these multiresolution schemes to real-life problems we need them to be *stable*. Our sequence of discrete data is assumed to be related not to an L^2 function but to a function in an appropriate normed space. The notion of stability we deal with (as in [17]) simply leads to a *well-posed algorithm* from the point of view of numerical analysis.

The present paper is organized as follows: In sections 2, 3, and 4, we describe those aspects of the general framework which are relevant to our discussion. Most of the material in these sections is taken from [16, 17] but we include it here to make this work almost self-contained.

Section 5 describes the general properties of multiresolution settings associated with the process of discretizing by integration against a compactly supported function that satisfies a dilation relation.

Interpolatory multiresolution settings are used in the construction of multiresolution schemes within the cell-average and hat-average frameworks. Therefore, they are studied with some detail in section 6. The stability of the multiresolution schemes in the interpolatory setting is related to the theory of refinement subdivision in CAD. We pay special attention to this connection in section 7, since it will be used later in the hat-average multiresolution setting.

Section 8 is the core of the paper. Here we describe the hat-average set-up and study its properties. Finally, some conclusions are drawn in section 9.

2. The general framework: A quick overview. Harten introduces its notion of multiresolution analysis in [15] and later generalizes it in [16, 17] where the theoretical foundation for multiresolution representation of data and operators is laid out.

This section and the next give a brief overview of the construction of a multiresolution scheme à la Harten. The material is taken mainly from [17]; we include it here to make this work almost self-contained, but also to emphasize what we think is essential in our later development. Proofs are given only when they are simple, short, and illustrative. For further details we refer the reader to [16, 17].

We start by recalling several useful definitions.

DEFINITION 2.1. A *multiresolution setting* is a sequence of linear spaces, $\{V^k\}$, which have denumerable basis, which we denote as $\{\eta_i^k\}$, together with a sequence of linear operators $\{D_k^{k-1}\}$ that map V^k onto V^{k-1} , i.e.,

$$D_k^{k-1} : V^k \rightarrow V^{k-1}, \quad V^{k-1} = D_k^{k-1}(V^k).$$

The operator D_k^{k-1} is called the *decimation operator*.

DEFINITION 2.2. We say that P_{k-1}^k is a *prediction operator* for the multiresolution setting $\{V^k\}, \{D_k^{k-1}\}$ if it is a right inverse of D_k^{k-1} in V^{k-1} , i.e.,

$$P_{k-1}^k : V^{k-1} \rightarrow V^k, \quad D_k^{k-1} P_{k-1}^k = I_{V^{k-1}}.$$

Note that P_{k-1}^k is not required to be a linear operator.

The spaces V^k represent the different levels of resolution (in our notation k increasing will imply more resolution). According to Definition 2.2, prediction operators should satisfy a basic consistency relation: Predicted values at the k th resolution level should contain the same discrete information as the original values, when restricted to the $(k-1)$ st level, that is, $D_k^{k-1} P_{k-1}^k = I_{V^{k-1}}$. On the other hand, the operator $P_{k-1}^k : V^{k-1} \rightarrow V^k$ produces approximations to each vector v^k in V^k from its

information contents at the level $k - 1$, i.e., $D_k^{k-1}v^k$. The prediction error

$$(1) \quad e^k := (I_{V^k} - P_{k-1}^k D_k^{k-1})v^k =: Q_k v^k$$

is a vector in V^k that belongs to the null space of the decimation operator

$$e^k \in \mathcal{N}(D_k^{k-1}) = \{v \mid v \in V^k, \quad D_k^{k-1}v = 0\},$$

since

$$D_k^{k-1}Q_k = D_k^{k-1}(I - P_{k-1}^k D_k^{k-1}) = D_k^{k-1} - (D_k^{k-1}P_{k-1}^k)D_k^{k-1} = 0.$$

Let us select a set of basis functions in $\mathcal{N}(D_k^{k-1})$

$$\mathcal{N}(D_k^{k-1}) = \text{span}\{\mu_j^k\}.$$

Let $G_k : \mathcal{N}(D_k^{k-1}) \rightarrow \mathcal{G}^k$ be the operator which assigns to any $e^k \in \mathcal{N}(D_k^{k-1})$ the sequence d^k of its coordinates in the basis $\{\mu_j^k\}$, and let E_k be the canonical injection $\mathcal{N}(D_k^{k-1}) \hookrightarrow V^k$. Obviously

$$G_k E_k = I_{\mathcal{G}^k}, \quad E_k G_k = I_{\mathcal{N}(D_k^{k-1})}.$$

It is then easy to prove that there is a one-to-one correspondence between v^k and $\{d^k, v^{k-1}\}$, where $d^k = G_k Q_k v^k$: Given v^k , we evaluate

$$(2) \quad \begin{cases} v^{k-1} &= D_k^{k-1}v^k, \\ d^k &= G_k(I_{V^k} - P_{k-1}^k D_k^{k-1})v^k. \end{cases}$$

Given v^{k-1} and d^k , we recover v^k by

$$(3) \quad \begin{aligned} P_{k-1}^k v^{k-1} + E_k d^k &= P_{k-1}^k D_k^{k-1}v^k + E_k G_k(I_{V^k} - P_{k-1}^k D_k^{k-1})v^k \\ &= P_{k-1}^k D_k^{k-1}v^k + (I_{V^k} - P_{k-1}^k D_k^{k-1})v^k \\ &= v^k. \end{aligned}$$

By repeating step (2) for each resolution level, we find that a multiresolution setting $(\{V_k\}_{k=0}^L, \{D_k^{k-1}\}_{k=1}^L)$ and a sequence of corresponding prediction operators $\{P_{k-1}^k\}_{k=1}^L$ (linear or nonlinear) define a *multiresolution transform*. The algorithms that compute this invertible transformation as well as its inverse are as follows:

$$(4) \quad v^L \rightarrow Mv^L \quad (\text{Encoding}) \quad \begin{cases} \text{Do} & k = L, \dots, 1, \\ v^{k-1} &= D_k^{k-1}v^k, \\ d^k &= G_k(v^k - P_{k-1}^k v^{k-1}) \end{cases}$$

$$Mv^L = \{v^0, d^1, \dots, d^L\},$$

$$(5) \quad Mv^L \rightarrow M^{-1}Mv^L \quad (\text{Decoding}) \quad \begin{cases} \text{Do} & k = 1, \dots, L, \\ v^k &= P_{k-1}^k v^{k-1} + E_k d^k. \end{cases}$$

We refer to $M(v^L)$ as the multiresolution representation of v^L and to algorithms (4) and (5) as the direct and inverse multiresolution transforms, respectively. Note that the simple algebraic relations (2) and (3) show that

$$v^L \xrightarrow{1:1} \{v^0, d^1, \dots, d^L\} = Mv^L.$$

REMARK 2.1. *In the finite dimensional case $\dim V^k = N_k$ we have $\dim \mathcal{N}(D_k^{k-1}) = N_k - N_{k-1}$. The number of components in d^k is thus $(N_k - N_{k-1})$ and, consequently, the number of components in $M(v^L)$ is*

$$N_0 + \sum_{k=1}^L (N_k - N_{k-1}) = N_L.$$

Hence, when v^L is a finite sequence, which is the case in most applications, $M(v^L)$ has exactly the same cardinality as v^L .

REMARK 2.2. *The operators D_k^{k-1} and P_{k-1}^k serve, respectively, as decimation and prediction in a pyramid scheme of the type that is used in signal processing. The redundancy which is typical of frames which are obtained by pyramid schemes is removed, much like in wavelet theory, by expressing the prediction errors in a basis of $\mathcal{N}(D_k^{k-1})$. This allows us to obtain a multiresolution representation (tight frame).*

Observe that algorithms (4) and (5) have the same structure as Mallat's decomposition and reconstruction algorithms. Decomposition is carried out by means of two filters; one of them might be nonlinear, if the prediction operator is too. In this context (and unlike wavelet algorithms), even in the linear case the filters need not be of convolution type.

In multiscale analysis, a “new scale” is loosely defined as the information on a given resolution level which cannot be inferred from information in lower resolution levels. We can interpret (3) as saying that d^k represents nonredundant information present in v^k and not predictable from v^{k-1} by the prediction scheme defined by P_{k-1}^k . Motivated by this interpretation, the components of d^k are referred to as the k th *scale coefficients* of the multiresolution representation. When using a *particular* prediction scheme, the errors e^k , and consequently the k th scale coefficients, include, in addition to the “true” k th scale, a component due to the approximation error which is related to the “quality” or “accuracy” of the particular prediction scheme. It is clear then that one of the main concerns should be the “quality” of the prediction P_{k-1}^k .

A given multiresolution scheme can be applied to *any* sequence of real numbers. These numbers could have been generated by some stochastic process, by an iterated function system (IFS) or a numerical scheme for the solution of a PDE, or by any other means. When posed in general, it is not clear how to give a precise meaning to the question of quality. It can, however, be made meaningful by restricting our attention to a subset of data for which we know something about the way it was generated.

The generation of discrete data is usually done through the application of a particular *discretization procedure*. In numerical analysis, different types of discretization operators are used to obtain discrete representations of “continuous” signals, according to their “nature.” For example, discretization by point-values is the natural way of associating a set of discrete values to a continuous function, while a “cell-average” discretization procedure is appropriate for piecewise continuous functions with jumps. Next we shall see how a sequence of decimation operators can be constructed from a *nested sequence of discretization operators*.

DEFINITION 2.3. Let \mathcal{D} be a linear operator on a linear space \mathcal{F} , and denote its range by V . If V has a denumerable basis, $\{\eta_i\}$, we say that \mathcal{D} is a discretization operator on \mathcal{F} and, for each $f \in \mathcal{F}$, we refer to $v = \mathcal{D}f$ as the discretization of f at the resolution level specified by V .

$$\mathcal{D} : \mathcal{F} \rightarrow V, \quad \text{where } V = \mathcal{D}(\mathcal{F}) = \text{span}\{\eta_i\}.$$

DEFINITION 2.4. Let $\{\mathcal{D}_k\}$ be a sequence of discretization operators on \mathcal{F}

$$\mathcal{D}_k : \mathcal{F} \rightarrow V^k, \quad \mathcal{D}_k(\mathcal{F}) = V^k = \text{span}\{\eta_i^k\}.$$

We say that the sequence $\{\mathcal{D}_k\}$ is nested if $\forall k$ and $\forall f \in \mathcal{F}$,

$$(6) \quad \mathcal{D}_k f = 0 \Rightarrow \mathcal{D}_{k-1} f = 0.$$

The nested property implies that the discrete information at a given resolution level is also included in the discrete information at all finer resolution levels.

A nested sequence of discretization defines a sequence of decimation operators and, thus, a multiresolution setting. This result follows from the following lemma.

LEMMA 2.1. If $\{\mathcal{D}_k\}$ is a nested sequence of discretization, then the mapping from V^k to V^{k-1} defined as

For $v \in V^k$ take any $f \in \mathcal{F}$ such that $v = \mathcal{D}_k f$ and assign to it $u := \mathcal{D}_{k-1} f$ is a well-defined mapping.

Each decimation operator is then defined as follows: For any $v^k \in V^k$, let $f \in \mathcal{F}$ be such that $\mathcal{D}_k f = v^k$; then $D_k^{k-1} v^k = \mathcal{D}_{k-1} f$. Lemma 2.1 implies that the definition is independent of f . Thus we have

$$(7) \quad D_k^{k-1} \mathcal{D}_k = \mathcal{D}_{k-1},$$

which readily shows that D_k^{k-1} is a linear operator. Moreover, for a nested sequence of discretization, (7) defines an operator that maps V^k onto V^{k-1} . To see this, let $u \in V^{k-1}$ and take $f \in \mathcal{F}$ such that $u = \mathcal{D}_{k-1} f$; then let $v = \mathcal{D}_k f$. Clearly $v \in V^k$ and (7) implies

$$D_k^{k-1} v = D_k^{k-1} (\mathcal{D}_k f) = \mathcal{D}_{k-1} f = u.$$

Given $v \in V^k$, any $f \in \mathcal{F}$ satisfying $v = \mathcal{D}_k f$ is called a “reconstruction” of v in \mathcal{F} .

A nested sequence of discretization operators defines a multiresolution setting, i.e., an appropriate framework to deal with discrete data coming from the discretization process at hand. To complete the design of a multiresolution scheme we still have two more *independent* choices to make:

1. a prediction operator P_{k-1}^k which is a right inverse of D_k^{k-1} ;
2. a basis $\{\mu_j^k\}$ of $\mathcal{N}(D_k^{k-1})$, which will be used to define the operators G_k and E_k .

Prediction operators are constructed using a sequence of appropriate “reconstruction operators.”

DEFINITION 2.5. We say that \mathcal{R} ,

$$\mathcal{R} : V \rightarrow \mathcal{F}, \quad V = \mathcal{D}(\mathcal{F}),$$

is a reconstruction operator for \mathcal{D} in V if it is a right inverse of \mathcal{D} , i.e., $\mathcal{D}\mathcal{R} = I_V$. Note that \mathcal{R} is not required to be a linear operator.

Given a sequence of discretization operators $\{\mathcal{D}_k\}$ and any sequence of corresponding reconstruction operators $\{\mathcal{R}_k\}$, a right inverse of D_k^{k-1} can now easily be defined as follows:

$$(8) \quad P_{k-1}^k := \mathcal{D}_k \mathcal{R}_{k-1} : V^{k-1} \rightarrow V^k.$$

The prediction operator defined above is a right inverse of D_k^{k-1} , since

$$D_k^{k-1} P_{k-1}^k = D_k^{k-1} (\mathcal{D}_k \mathcal{R}_{k-1}) = (D_k^{k-1} \mathcal{D}_k) \mathcal{R}_{k-1} = \mathcal{D}_{k-1} \mathcal{R}_{k-1} = I_{V_{k-1}}.$$

REMARK 2.3. *An interesting by-product of the definition of discretization and reconstruction is that the decimation associated to the sequence of discretization $\{\mathcal{D}_k\}$ can also be characterized as*

$$(9) \quad D_k^{k-1} = \mathcal{D}_{k-1} \mathcal{R}_k,$$

since our definition in (7) called for “any $f \in \mathcal{F}$ such that $\mathcal{D}_k f = v^k$,” and $f = \mathcal{R}_k v^k$ satisfies $\mathcal{D}_k f = \mathcal{D}_k \mathcal{R}_k v^k = v^k$. Although there seems to be an explicit dependence of D_k^{k-1} on the reconstruction \mathcal{R}_k , because of Lemma 2.1, this dependence is totally fictitious. Nevertheless, expression (9) will turn out to be very useful.

Within this framework, it is clear that finding a suitable prediction for a multiresolution setting, and thus a suitable multiresolution scheme for a given application, can be formulated as a typical problem in approximation theory:

Knowing $\mathcal{D}_{k-1} f$, $f \in \mathcal{F}$, find a “good approximation” to $\mathcal{D}_k f$.

Moreover, an operative definition of the “quality” of the prediction can be stated as follows:

If $p \in \mathcal{F}$ is a function for which \mathcal{R}_{k-1} is exact, i.e.,

$$\mathcal{R}_{k-1}(\mathcal{D}_{k-1} p) = p,$$

we have likewise

$$P_{k-1}^k(\mathcal{D}_{k-1} p) = \mathcal{D}_k \mathcal{R}_{k-1} \mathcal{D}_{k-1} p = \mathcal{D}_k p,$$

i.e., the prediction P_{k-1}^k is also exact on the discrete values associated with the function p . The quality, or accuracy, of the prediction can thus be judged by the class of functions in \mathcal{F} for which the reconstruction used in its definition is exact.

A good solution to our approximation problem will bring us one step closer to our stated goal: The design of multiresolution schemes that apply to all sequences, but are particularly adequate for those sequences v^L which are obtained by the discretization process defined by the operators \mathcal{D}_k .

3. Stability analysis for linear reconstruction operators. We have seen in the previous section that there is a one-to-one correspondence between v^L , the discrete data at the finest resolution level, and its multiresolution representation, i.e., the set (v^0, d^1, \dots, d^L) . The *scale coefficients* are directly related to the prediction errors, thus, small scale coefficients on a given scale mean that v^k is properly represented by $P_{k-1}^k v^{k-1}$ on that particular scale. When they are “sufficiently small” (this should be appropriately quantified), scale coefficients can be truncated (or “quantized,” see, e.g., [15] or Part II), reducing the dimensionality of Mv^L with very little alteration of its (discrete) information contents. This is, in fact, the driving principle behind the design of compression algorithms from multiresolution schemes.

If a numerical problem can be embedded in a multiresolution setting, we can improve the efficiency of the numerical solution algorithm by applying data compression to the numerical solution [3, 19, 20] as well as to the multiresolution representation of the solution operator [4, 14, 2]. The numerical solution algorithm can be also reorganized as a multiscale computation, where the problem is solved directly on the coarsest level and then we advance from coarse to fine levels by prediction and correction [4, 21].

In order to apply these multiresolution schemes to real-life problems, we would like the scale coefficients

$$d^k = G_k(v^k - P_{k-1}^k D_k^{k-1} v^k)$$

to be a good approximation to the “true” k th scale. Although a crucial element in achieving this goal is the accuracy of the prediction, it is not the only consideration. We have to make sure that the direct multiresolution transform and its inverse are stable with respect to perturbations. The notion of stability we use is deeply rooted in the numerical analysis concept that a *useful* algorithm should not amplify noise, that is, the effect of an initial perturbation in the data of the algorithm should not destroy its ‘reasonable’ outcome.

This is a weaker concept than the usual stability requirements in wavelet theory, which involves the L^2 norm. The stability requirements we impose will imply that the multiresolution transform and its inverse are “usable” algorithms, and it requires only a normed linear space, not necessarily L^2 .

The decimation operators D_k^{k-1} are always linear. When the reconstruction operators, \mathcal{R}_k , are linear, the prediction operators P_{k-1}^k are linear too. In this case, the multiresolution transform becomes a linear operator describing a change of basis vectors in $\mathcal{D}_L(\mathcal{F})$, and the question of stability admits a relatively simple approach.

To see this, let us introduce the linear operator B_L^k of successive decimation

$$(10) \quad B_L^k = D_{k+1}^k \cdots D_L^{L-1} : V^L \rightarrow V^k$$

and observe that v^k in (4) can be written as

$$(11) \quad v^k = B_L^k v^L.$$

The direct multiresolution transform $v^L \mapsto M(v^L)$ can thus be expressed as

$$(12) \quad v^0 = B_L^0 v^L, \quad d^k = G_k Q_k B_L^k v^L, \quad 1 \leq k \leq L.$$

Likewise, let us introduce the operator A_k^L of successive prediction as

$$(13) \quad A_k^L = P_{L-1}^L \cdots P_k^{k+1} : V^k \rightarrow V^L.$$

When \mathcal{R}_k is linear $\forall k$ these operators are also linear, and this fact allows us to express the inverse multiresolution transform (5) directly in terms of $M(v^L)$ as follows:

$$(14) \quad v^L = A_0^L v^0 + \sum_{k=1}^L A_k^L E_k d^k.$$

3.1. Stability with respect to perturbations. Expressions (12) and (14) let us examine rather easily the influence of perturbations in the input data of the direct and inverse multiresolution transforms.

For purposes of analysis, if v^L is replaced by a perturbed v_ϵ^L , stability of the direct multiresolution transform should imply that the perturbation in the resulting scale coefficients and low-level approximation has to be bounded by the perturbation in the input. Under our linearity assumptions, we can write

$$(15) \quad \delta(v^0) = v_\epsilon^0 - v^0 = B_L^0(v_\epsilon^L - v^L), \quad \delta(d^k) = d_\epsilon^k - d^k = G_k Q_k B_L^k(v_\epsilon^L - v^L).$$

These relations show that the perturbation in the input is subject to successive decimation D_{m-1}^m for $m = L, \dots, k+1$, and in the case of the scale coefficients, projected into $\mathcal{N}(D_k^{k-1})$ and represented in some basis there. Clearly the “dangerous” process that needs to be controlled, from the point of view of error-amplification, is that of successive decimation; the choice of basis in $\mathcal{N}(D_k^{k-1})$ is not that important.

Similarly, for purposes of data compression if the scale coefficients $\{d^k\}$ are replaced by $\{d_\epsilon^k\}$ which are obtained either by quantification or by truncation, we want the perturbation in the output of the algorithm, the decompressed v_ϵ^L , to be bounded by the perturbation in the scale coefficients. Linearity of all operators involved leads now to

$$(16) \quad \delta(v^L) = v_\epsilon^L - v^L = A_0^L(v_\epsilon^0 - v^0) + \sum_{k=1}^L A_k^L E_k(d_\epsilon^k - d^k),$$

which shows that the perturbation in the scale coefficients is “translated” into a perturbation in the prediction error and then transmitted into higher levels of resolution by successive prediction P_{m-1}^m for $m = k+1, \dots, L$. The danger here is that the perturbation could be amplified by the process of successive prediction.

The role of stability is to prevent unbounded growth of initial perturbations by repeated applications of an operator. For a nested sequence of discretization, the successive decimation operator B_L^k controls the stability of the direct multiresolution transform. For linear prediction operators, the successive prediction operator A_k^L controls the stability of the inverse multiresolution transform.

The nested character of the sequence of discretization suffices to eliminate the possibility of amplification due to successive decimation. This is a consequence of the following lemma.

LEMMA 3.1. *If $\{\mathcal{D}_k\}$ is nested, then $\mathcal{D}_l(\mathcal{R}_m \mathcal{D}_m) = \mathcal{D}_l$ for $l \leq m$.*

Proof. For any $f \in \mathcal{F}$, let $g = \mathcal{R}_m \mathcal{D}_m f$; then

$$\mathcal{D}_m g = \mathcal{D}_m(\mathcal{R}_m \mathcal{D}_m) f = (\mathcal{D}_m \mathcal{R}_m) \mathcal{D}_m f = \mathcal{D}_m f \Rightarrow \mathcal{D}_m(f - g) = 0 \Rightarrow \mathcal{D}_l(f - g) = 0, \\ l \leq m. \quad \square$$

Lemma 3.1 implies that

$$(17) \quad B_L^k \mathcal{D}_L = D_{k+1}^k \cdots D_L^{L-1} \mathcal{D}_L = \mathcal{D}_k \mathcal{R}_{k+1} \mathcal{D}_{k+1} \cdots \mathcal{D}_{L-1} \mathcal{R}_L \mathcal{D}_L = \mathcal{D}_k.$$

The stability of the successive decimation step hinges on this purely algebraic relation. It essentially means that if we start at a given resolution level, L , and apply a number of decimation sweeps, say m , the discrete information we obtain is precisely what corresponds to the $L-m$ resolution level; in other words, the decimation operator does not introduce additional information or amplify noise.

To obtain explicit stability bounds we shall assume that \mathcal{F} is a Banach space. Let $\|\cdot\|$ denote the norm in \mathcal{F} ; we can define discrete norms in the spaces V^k and \mathcal{G}_k as follows (proving they are norms is an easy exercise; see [17]):

$$(18) \quad v^k \in V^k, \quad |v^k|_k := \|\mathcal{R}_k v^k\|,$$

$$(19) \quad d^k \in \mathcal{G}_k, \quad \langle d^k \rangle_k := |E_k d^k|_k.$$

These special norms are designed to accommodate the different dimensions of the various levels of resolution in the finite dimensional case.

We can then prove the following.

LEMMA 3.2. *Let us assume that $\mathcal{R}_k \mathcal{D}_k : \mathcal{F} \rightarrow \mathcal{F}$ is a bounded linear operator, i.e.,*

$$(20) \quad \|\mathcal{R}_k \mathcal{D}_k f\| \leq C^k \|f\|;$$

then $B_L^k : V^L \rightarrow V^k$ is a bounded linear operator, and

$$(21) \quad \|B_L^k\| \leq C^k \quad \text{for any } L.$$

Proof. Let $f := \mathcal{R}_L v^L$. Then relations (11) and (17) imply

$$v^k = B_L^k v^L = B_L^k \mathcal{D}_L f = \mathcal{D}_k f.$$

Therefore

$$|B_L^k v^L|_k = |B_L^k \mathcal{D}_L f|_k = |\mathcal{D}_k f|_k = \|\mathcal{R}_k \mathcal{D}_k f\| \leq C^k \|f\| = C^k \|\mathcal{R}_L v^L\| = C^k |v^L|_L,$$

i.e., $\|B_L^k\| \leq C^k$, which means that perturbation growth due to successive decimation is adequately controlled. \square

Stability of the inverse multiresolution transform is usually a more involved matter. There is one situation, however, where the analysis is particularly simple.

DEFINITION 3.1 (hierarchical sequence). *We say that the sequence $\{\mathcal{R}_k \mathcal{D}_k\}$ is hierarchical, if $\forall k$*

$$(22) \quad (\mathcal{R}_k \mathcal{D}_k) \mathcal{R}_{k-1} = \mathcal{R}_{k-1} \quad \equiv \quad \mathcal{R}_k P_{k-1}^k = \mathcal{R}_{k-1}.$$

Note that for a hierarchical sequence

$$(23) \quad \mathcal{R}_L A_k^L = \mathcal{R}_L \mathcal{D}_L \mathcal{R}_{L-1} \cdots \mathcal{D}_{k+1} \mathcal{R}_k = \mathcal{R}_k.$$

Relation (23) shows that a hierarchical structure in the sequence $\{\mathcal{R}_k \mathcal{D}_k\}$ prevents the amplification of perturbations due to successive prediction in the same way nestedness, i.e., $\mathcal{D}_{k-1}(\mathcal{R}_k \mathcal{D}_k) = \mathcal{D}_{k-1}$, prevents excessive perturbation growth in the successive decimation step.

The algebraic relation (23) is the equivalent to (17) for the successive prediction operator. It means that after a finite number of applications of the prediction operator the reconstruction from the discrete information obtained is the same as the reconstruction obtained with the discrete data we started with. This is enough to ensure that the successive prediction step does not introduce spurious information or amplify existing noise.

LEMMA 3.3. *If $\{\mathcal{R}_k \mathcal{D}_k\}$ is hierarchical, then*

$$(24) \quad \|A_k^L\| = 1.$$

Proof.

$$|A_k^L v^k|_L = |\mathcal{R}_L A_k^L v^k| = |\mathcal{R}_k v^k| = |v^k|_k \Rightarrow \|A_k^L\| = 1. \quad \square$$

The following two theorems give working stability bounds for the direct and inverse multiresolution transforms.

THEOREM 3.1. *If \mathcal{R}_k are linear operators and $\{\mathcal{R}_k \mathcal{D}_k\}$ are bounded operators, the direct multiresolution transform is stable, and*

$$\langle \delta(d^k) \rangle_k \leq C^k (1 + C^{k-1}) |\delta(v^L)|_L, \quad |\delta(v^0)|_0 \leq C^0 |\delta(v^L)|_L.$$

Proof. Let $f = \mathcal{R}_L v^L$, $f_\epsilon = \mathcal{R}_L v_\epsilon^L$. We have

$$\delta(v^0) = B_L^0 (v_\epsilon^L - v^L) \Rightarrow |\delta(v^0)|_0 \leq \|B_L^0\| \cdot |\delta(v^L)|_L \leq C^0 |\delta(v^L)|_L.$$

$$\delta(d^k) = G_k Q_k B_L^k \delta(v^L) \Rightarrow \langle \delta(d^k) \rangle_k = |E_k G_k Q_k B_L^k \delta(v^L)|_k = |Q_k B_L^k \delta(v^L)|_k.$$

Observe that

$$Q_k B_L^k \mathcal{D}_L = Q_k \mathcal{D}_k = (I - P_{k-1}^k D_k^{k-1}) \mathcal{D}_k = \mathcal{D}_k - (P_{k-1}^k)(D_k^{k-1} \mathcal{D}_k) = \mathcal{D}_k - \mathcal{D}_k \mathcal{R}_{k-1} \mathcal{D}_{k-1};$$

thus

$$\begin{aligned} \langle \delta(d^k) \rangle_k &= |\mathcal{D}_k (I - \mathcal{R}_{k-1} \mathcal{D}_{k-1}) \delta(f)|_k = |\mathcal{R}_k \mathcal{D}_k (I - \mathcal{R}_{k-1} \mathcal{D}_{k-1}) \delta(f)| \\ &\leq C^k (1 + C^{k-1}) \|\delta(f)\| = C^k (1 + C^{k-1}) \|\mathcal{R}_L \delta(v^L)\| \\ &= C^k (1 + C^{k-1}) |\delta(v^L)|_L. \quad \square \end{aligned}$$

THEOREM 3.2. *If \mathcal{R}_k are linear operators and $\{\mathcal{R}_k \mathcal{D}_k\}$ is a hierarchical sequence of bounded linear operators, the inverse multiresolution transform is stable and*

$$|\delta(v^L)|_L \leq |\delta(v^0)|_0 + \sum_{k=1}^L \langle \delta(d^k) \rangle_k.$$

Proof. Because of linearity and $\|A_k^L\| = 1$, we have

$$\begin{aligned} |\delta(v^L)|_L &\leq |A_0^L \delta(v^0)|_L + \sum_{k=1}^L |A_k^L E_k \delta(d^k)|_L \leq |\delta(v^0)|_0 + \sum_{k=1}^L |E_k \delta(d^k)|_k \\ &= |\delta(v^0)|_0 + \sum_{k=1}^L \langle \delta(d^k) \rangle_k. \quad \square \end{aligned}$$

3.2. Stability for nonhierarchical reconstructions: The hierarchical form.

Hierarchical reconstructions are guaranteed to be stable. However, many reconstruction techniques used in numerical analysis are not hierarchical. For example, piecewise interpolation, one of the most common procedures in numerical analysis, does not lead to hierarchical reconstruction procedures when the polynomial pieces are of degree strictly larger than one (see [16, 17] or next section). Checking stability might not be an easy task in this case, with the definitions we have covered so far. However, a sequence of approximation that is not hierarchical to begin with, has, in many cases, a hierarchical form which is obtained by considering a limiting process akin to

refinement in subdivision schemes [5, 13]. This hierarchical form leads to the same scale coefficients as the original one; hence the stability properties derived from the structure of the hierarchical reconstruction sequence are also inherited by the original (usually more manageable) one.

The main theoretical results are the following (we refer to [17] for proofs).

THEOREM 3.3. *Let $\{\mathcal{D}_k\}_{k=0}^\infty$ be a nested sequence of discretization operators and $\{\mathcal{R}_k\}_{k=0}^\infty$ a sequence of reconstruction operators satisfying $\mathcal{D}_k\mathcal{R}_k = I_{V_k}$. Assume that $\{\mathcal{R}_k\mathcal{D}_k\}_{k=0}^\infty$ is a sequence of bounded linear operators such that, for any $k \geq 0$ and any $f \in \mathcal{F}$, the following limit exists:*

$$(25) \quad \lim_{L \rightarrow \infty} \Pi_k^L f =: f_k^\infty \in \mathcal{F},$$

where $\Pi_k^L = (\mathcal{R}_L\mathcal{D}_L)(\mathcal{R}_{L-1}\mathcal{D}_{L-1}) \cdots \mathcal{R}_k\mathcal{D}_k$. Then

$$\mathcal{D}_l f_k^\infty = \mathcal{D}_l f \quad \text{for } l \leq k, \quad d^l(f_k^\infty) = 0 \quad \text{for } l \geq k+1.$$

Note that $\Pi_k^L f$ is described on a higher level of resolution (finer scale) than $\Pi_k^{L-1} f$, so in this respect f_k^∞ corresponds to “infinite resolution.” Nevertheless, Theorem 3.3 shows that f_k^∞ has exactly the same discrete information contents at the k th level as the initial data $\mathcal{R}_k\mathcal{D}_k f$. The limiting process (25) which assigns f_k^∞ to $\mathcal{R}_k\mathcal{D}_k f$ is called in [17] “cosmetic refinement,” in order to stress that unlike other refinement processes in numerical analysis, there is no addition of (discrete) information.

THEOREM 3.4. *Let $\{\mathcal{R}_k\mathcal{D}_k\}$ be as in Theorem 3.3 and define*

$$\mathcal{R}_k^H : V^k \rightarrow \mathcal{F} \quad \mathcal{R}_k^H v^k = \lim_{L \rightarrow \infty} \Pi_{k+1}^L \mathcal{R}_k v^k = \lim_{L \rightarrow \infty} \mathcal{R}_L A_k^L v^k.$$

Then

1. \mathcal{R}_k^H is a reconstruction of \mathcal{D}_k in \mathcal{F} ;
2. $(P^H)_{k-1}^k := \mathcal{D}_k \mathcal{R}_{k-1}^H = \mathcal{D}_k \mathcal{R}_{k-1} = P_{k-1}^k$;
3. $\{\mathcal{R}_k^H \mathcal{D}_k\}$ is a hierarchical sequence, i.e., $(\mathcal{R}_{k+1}^H \mathcal{D}_{k+1}) \mathcal{R}_k^H = \mathcal{R}_k^H$.

This theorem states that the existence of the cosmetic refinement limit implies, in turn, the existence of a hierarchical reconstruction procedure that produces exactly the same prediction operator as the original one. Since hierarchical reconstructions lead naturally to stable multiresolution transforms, stability of the original scheme is a direct consequence of the *existence* of the cosmetic refinement limit.

For all practical purposes it is not important to know the explicit expression of the hierarchical form; however, knowledge of its *existence* is essential because it implies stability of the original multiresolution scheme.

Notice also that if $p \in \mathcal{F}$ is such that

$$\mathcal{R}_l \mathcal{D}_l p = p \quad \forall l \geq 0,$$

then $\Pi_k^L p = p$, $\forall k$ and $L \geq k$ and consequently

$$\mathcal{R}_k^H \mathcal{D}_k p = \lim_{L \rightarrow \infty} \Pi_k^L p = p,$$

which shows that the hierarchical reconstruction, \mathcal{R}_k^H , has the same “accuracy” as the original \mathcal{R}_k .

The existence of \mathcal{R}_k^H is directly related to the existence of the cosmetic refinement of the functions $\mathcal{R}_k \eta_i^k$. Let us consider the finite dimensional case. If $\varphi_i^k = \mathcal{R}_k^H \eta_i^k$ is well defined, then so is $\mathcal{R}_k^H v^k \forall v^k \in V^k$. Note that

$$v^k = \sum_i \hat{v}_i^k \eta_i^k \quad \Rightarrow \quad \mathcal{R}_k^H v^k = \sum_i \hat{v}_i^k \mathcal{R}_k^H \eta_i^k,$$

since the sum is finite and the reconstruction operators are linear. The existence of the limit functions

$$(26) \quad \varphi_i^k = \lim_{L \rightarrow \infty} \Pi_{k+1}^L \mathcal{R}_k \eta_i^k = \lim_{L \rightarrow \infty} \mathcal{R}_L A_k^L \eta_i^k$$

becomes a test for the stability of the multiresolution scheme derived from a particular sequence of discretization and reconstruction operators.

Orthogonal and biorthogonal wavelet algorithms can be seen as particular examples of this general framework [15, 16, 17]. The reconstruction operators used in these algorithms are hierarchical and, as a consequence, the associated compression algorithms are stable. As it turns out, they are the hierarchical form of other algorithms, which are more “natural” from the point of view of numerical analysis.

In the linear case, reconstruction sequences which are based on spectral expansions or splines are also hierarchical (see [11, 16, 17]); thus, the additional functional structure and stability properties just described also apply to the multiresolution schemes they define.

The general framework, however, allows for *any* type of reconstruction procedure, linear or nonlinear, as long as it is a right inverse of the discretization operator. Nonlinear reconstruction techniques can be used to optimize compression rates. In this case, stability must be ensured by a modified encoding-decoding procedure. We refer the reader to [15, 18] and to Part II [1] for descriptions of nonlinear reconstruction procedures in various contexts, as well as stability considerations in the nonlinear case.

4. Hierarchical sequences and their wavelet structure. Hierarchical sequences have an associated wavelet-like functional structure. Given $v^k = \mathcal{D}_k f$, one step of the inverse multiresolution transform (5) can be written as

$$\mathcal{D}_k f = \mathcal{D}_k \mathcal{R}_{k-1} \mathcal{D}_{k-1} f + \sum_j d_j^k \mu_j^k.$$

If the sequence $\{\mathcal{R}_k \mathcal{D}_k\}$ is hierarchical, applying \mathcal{R}_k leaves us with

$$(27) \quad \mathcal{R}_k \mathcal{D}_k f = \mathcal{R}_k \mathcal{D}_k \mathcal{R}_{k-1} \mathcal{D}_{k-1} f + \sum_j d_j^k \mathcal{R}_k \mu_j^k = \mathcal{R}_{k-1} \mathcal{D}_{k-1} f + \sum_j d_j^k \psi_j^k,$$

where $\psi_j^k := \mathcal{R}_k \mu_j^k$.

Relation (27) tells us that $\sum_j d_j^k \psi_j^k$ represents the difference in information between two functional approximations to the original signal f at consecutive resolution levels. The approximation $\mathcal{R}_k \mathcal{D}_k f$ is obtained by adding to a lower resolution approximation, $\mathcal{R}_{k-1} \mathcal{D}_{k-1} f$, the missing *details*.

The functional approximation $\mathcal{R}_k \mathcal{D}_k f$ belongs to the finite dimensional subspace $\mathcal{R}_k \mathcal{D}_k \mathcal{F}$. In fact,

$$\mathcal{R}_k \mathcal{D}_k \mathcal{F} = \text{span}\{\varphi_i^k\} =: \Phi^k,$$

where $\varphi_i^k := \mathcal{R}_k \eta_i^k$. This can easily be checked; since $\mathcal{D}_k f = \sum_i (\mathcal{D}_k f)_i \eta_i^k$ and \mathcal{R}_k is linear, we have

$$\mathcal{R}_k \mathcal{D}_k f = \mathcal{R}_k \sum_i (\mathcal{D}_k f)_i \eta_i^k = \sum_i (\mathcal{D}_k f)_i \varphi_i^k \quad \forall f \in \mathcal{F}.$$

Notice that

$$(\mathcal{R}_k \mathcal{D}_k)^2 = \mathcal{R}_k (\mathcal{D}_k \mathcal{R}_k) \mathcal{D}_k = \mathcal{R}_k \mathcal{D}_k;$$

therefore $\mathcal{R}_k \mathcal{D}_k$ is a projection onto Φ^k .

The spaces $\mathcal{R}_k \mathcal{D}_k \mathcal{F} = \Phi^k$ form a ladder of subspaces of \mathcal{F} . The following lemma readily implies that $\Phi^{k-1} \subset \Phi^k$.

LEMMA 4.1. *Let \hat{P}_{k-1}^k be the matrix representation of the operator P_{k-1}^k with respect to the basis $\{\eta_i^k\}$ and $\{\eta_i^{k-1}\}$. Then*

$$(28) \quad \varphi_i^{k-1} = \sum_l (\hat{P}_{k-1}^k)_{li} \varphi_l^k.$$

Proof. Let us define

$$(29) \quad \bar{\varphi}_i^{k,L} := A_k^L \eta_i^k, \quad \varphi_i^{k,L} := \mathcal{R}_L \bar{\varphi}_i^{k,L}.$$

Then we have

$$(30) \quad \bar{\varphi}_i^{k-1,L} = A_{k-1}^L \eta_i^{k-1} = A_k^L P_{k-1}^k \eta_i^{k-1} = A_k^L \sum_l (\hat{P}_{k-1}^k)_{li} \eta_l^k = \sum_l (\hat{P}_{k-1}^k)_{li} \bar{\varphi}_l^{k,L}.$$

When $\{\mathcal{R}_k \mathcal{D}_k\}$ is hierarchical, we have $\mathcal{R}_L A_k^L = \mathcal{R}_k$ and

$$\varphi_i^{k,L} = \mathcal{R}_L A_k^L \eta_i^k = \mathcal{R}_k \eta_i^k = \varphi_i^k \quad \forall L;$$

thus, applying \mathcal{R}_L , which is a linear operator, to (30), we get (28). \square

Let us define

$$(31) \quad \psi_j^{k,L} := \mathcal{R}_L A_k^L \mu_j^k$$

and denote by \hat{E}_k the matrix representation of the operator E_k with respect to the basis $\{\eta_i^k\}$ and $\{\mu_j^k\}$. As in the proof of Lemma 4.1, we have

$$\psi_j^{k,L} = \mathcal{R}_L A_k^L \sum_l (\hat{E}_k)_{lj} \eta_l^k = \sum_l (\hat{E}_k)_{lj} \varphi_l^{k,L}.$$

For a hierarchical reconstruction $\mathcal{R}_L A_k^L = \mathcal{R}_k$; thus we have $\psi_j^{k,L} = \mathcal{R}_k \mu_j^k = \psi_j^k \quad \forall L$. Hence

$$(32) \quad \psi_j^k = \sum_l (\hat{E}_k)_{lj} \varphi_l^k.$$

Let us define $\Psi^k := \text{span}\{\psi_j^k\}$. Relations (28) and (32) show that $\Phi^{k-1} \subset \Phi^k$ and $\Psi^k \subset \Phi^k$, and relation (27) proves that $\Phi^{k-1} + \Psi^k = \Phi^k$. It is not hard to prove that $\Phi^{k-1} \cap \Psi^k = \{0\}$; thus

$$\Phi^k = \Phi^{k-1} \oplus \Psi^k,$$

which also implies that

$$(33) \quad \Phi^L = \Psi^L \oplus \dots \oplus \Psi^1 \oplus \Phi^0$$

and

$$\mathcal{R}_L \mathcal{D}_L f = \mathcal{R}_0 \mathcal{D}_0 f + \sum_k \sum_j d_j^k \psi_j^k.$$

This decomposition directly implies that the set $\{\varphi_i^0\}, \{\{\psi_j^k\}\}_{k=1}^L$ is a basis of $\mathcal{R}_L \mathcal{D}_L \mathcal{F}$. For this reason, the functions ψ_j^k are called in [17] *generalized wavelets*.

5. Discretization by local averages. In numerical analysis, discretization processes are used as tools to go from a function in some functional space to a set of discrete values which give sensible information about the original function. As an example, the simplest discretization process when dealing with continuous functions is that of pointwise discretization; the values of a continuous function at a given set of points give relevant information about the function. This information can be effectively used to represent the function or to reconstruct it if necessary (via interpolation, for example).

On the other hand, pointwise discretization does not provide an appropriate discretization setting for functions that are only piecewise continuous, since information about the exact location of discontinuities that fall between grid-points is completely lost. A discretization process often used when this space of functions is relevant (e.g., numerical solution of hyperbolic conservation laws) is the so-called cell-average discretization which considers the mean values of the function in each interval.

Thus, different discretization procedures arise naturally in different contexts. The goal is to know whether or not they lead to multiresolution settings and to stable multiresolution transforms, since multiresolution representations can then be used to reduce the cost of a numerical algorithm or to compress the information in the discrete set for purposes of storage or transmission.

Many discretization procedures in numerical analysis, and in particular the two mentioned above, can be described as the process of obtaining local averages with respect to a weight function, which is imposed by the underlying context. The weight function, $\omega(x)$, is a function with compact support and it is usually required that

$$(34) \quad \int \omega(x) dx = 1.$$

Let us consider a grid X composed of a finite sequence of equally spaced points in $[0, 1]$:

$$X = \{x_i\}, \quad x_i \in [0, 1], \quad h = x_i - x_{i-1}.$$

For each function f in \mathcal{F} , a discretization operator, \mathcal{D} , associated to the resolution level defined by the grid X is defined as follows:

$$(35) \quad (\mathcal{D}f)_i \equiv \bar{f}_i := \frac{1}{h} \int f(x) \omega\left(\frac{x - x_i}{h}\right) dx =: \left\langle f, \frac{1}{h} \omega\left(\frac{x - x_i}{h}\right) \right\rangle, \quad x_i \in X.$$

The (linear) operator \mathcal{D} acts naturally on a space of functions \mathcal{F} for which the integral in (35) is well defined and produces discrete values which give local information on the behavior of the function f at the resolution level specified by the grid X .

Weighted averages against the “generalized” function $\omega(x) = \delta(x)$, *Dirac’s* delta-function, correspond precisely to the process of discretizing a continuous function by considering its values at the points of the grid X . A rather ‘natural’ function space \mathcal{F} for this discretization operator is $\mathcal{C}[0, 1]$. On the other hand, the cell-average discretization procedure corresponds to the local-average discretization (35) with $\omega(x)$ being the box function:

$$(36) \quad \omega(x) = \begin{cases} 1, & -1 \leq x < 0, \\ 0 & \text{otherwise.} \end{cases}$$

A natural function space in this context is $L^1[0, 1]$.

Do these discretization procedures lead to multiresolution settings? The answer is yes. A sequence of *nested* dyadic grids on $[0, 1]$ together with the *dilation relation* satisfied by each one of the aforementioned functions allow us to construct a nested sequence of discretization and thus a multiresolution setting.

Let us consider the set of nested dyadic grids (with which we associate the increasing levels of resolution) $\{X^k\}$, $k \geq 0$ of size $h_k = 2^{-k}h_0$,

$$(37) \quad X^k = \{x_j^k\} \quad x_j^k = j \cdot h_k, \quad j = 0, \dots, J_k, \quad J_k \cdot h_k = 1$$

(notice that $x_{2j}^k = x_j^{k-1}$) and the sequence of discretization operators

$$(38) \quad \mathcal{D}_k : \mathcal{F} \rightarrow S^k, \quad (\mathcal{D}_k f)_i = \bar{f}_i^k = \langle f, \omega_i^k \rangle, \quad \omega_i^k = \frac{1}{h_k} \omega \left(\frac{x - x_i^k}{h_k} \right),$$

where S^k is a space of sequences of dimension N_k related to J_k (e.g., for $\omega(x) = \delta(x)$, $N_k = J_k + 1$). In this context we shall always consider $\eta_i^k = \delta_i^k$, the canonical basis vectors in spaces of sequences.

When the weight function, $\omega(x)$, satisfies a dilation relation such as

$$(39) \quad \omega(y) = 2 \sum_l \alpha_l \omega(2y - l),$$

it is easy to see that the sequence of discretization (38) is nested. In fact, taking $y = (x - x_i^{k-1})/h_{k-1}$, we can rewrite (39) as

$$\begin{aligned} \omega \left(\frac{x - x_i^{k-1}}{h_{k-1}} \right) &= 2 \sum_l \alpha_l \omega \left(2 \frac{x - x_i^{k-1}}{h_{k-1}} - l \right) = 2 \sum_l \alpha_l \omega \left(\frac{x - x_{2i}^k}{h_k} - l \right) \\ &= 2 \sum_l \alpha_l \omega \left(\frac{x - x_{2i+l}^k}{h_k} \right), \end{aligned}$$

i.e.,

$$(40) \quad \omega_i^{k-1} = \sum_l \alpha_l \omega_{2i+l}^k = \sum_l \alpha_{l-2i} \omega_l^k.$$

The last relation can be expressed in terms of the discretization operators as

$$(41) \quad (\mathcal{D}_{k-1} f)_i = \sum_j \alpha_{j-2i} (\mathcal{D}_k f)_j,$$

which implies that the sequence $\{\mathcal{D}_k\}$ is nested. Therefore, discretizing by local averages with respect to a function that satisfies a dilation relation becomes a particular way of obtaining a nested sequence of discretization operators.

Formulas (40) and (41) also imply that the decimation operator, D_k^{k-1} , can be described by a matrix whose elements are

$$(D_k^{k-1})_{ij} = \alpha_{j-2i}.$$

Observe that D_k^{k-1} is independent of the level of resolution (except for its dimension).

The point-value and cell-average multiresolution settings have been extensively discussed in [15, 16, 17]. The main purpose of this paper is to analyze the multiresolution setting derived from discretizing by taking local averages with respect to the *hat* function:

$$(42) \quad \omega(x) = \begin{cases} 1+x, & -1 \leq x \leq 0, \\ 1-x, & 0 \leq x \leq 1, \end{cases}$$

which satisfies the following dilation relation:

$$(43) \quad \omega(x) = \frac{1}{2}[\omega(2x-1) + 2\omega(2x) + \omega(2x+1)] \Rightarrow \alpha_1 = \alpha_{-1} = \frac{1}{4}, \alpha_0 = \frac{1}{2}.$$

Since the hat function is continuous, integration against “functions” with δ -type singularities can be allowed. A “natural” function space, which is also of interest in vortex methods, is that of piecewise smooth functions in $[0, 1]$ with δ -type singularities in $(0, 1)$. This space is just the space of generalized derivatives of piecewise smooth functions with jumps, and it is a subspace of the Banach space $\mathcal{C}^*[0, 1]$, the space of continuous linear functionals on $\mathcal{C}[0, 1]$.

There is an interesting duality between the interpolatory and hat-average multiresolution settings. For the interpolatory set-up $\mathcal{F} = \mathcal{C}[0, 1]$, while for the hat-average set-up $\mathcal{F} \subset \mathcal{C}^*[0, 1]$. This duality will be exploited further in the following sections since it helps to study the stability (with respect to perturbations) properties of hat-average multiresolution transforms.

Once the weight function is fixed, the primary choice, that of the decimation operator, is already made. To construct an adequate multiresolution scheme, we still have two more *independent* choices to make:

1. a basis for the null space or, equivalently, an operative definition of the transfer operators G_k and E_k ;
2. a prediction operator P_{k-1}^k , which is a right inverse of D_k^{k-1} . This amounts to choosing appropriate reconstruction operators at each resolution level.

When the weight function satisfies a dilation relation, the null space of D_k^{k-1} is easily characterized:

$$(44) \quad \mathcal{N}(D_k^{k-1}) = \{s^k \in S^k \mid D_k^{k-1}s^k = 0\} = \left\{s^k \in S^k \mid \sum_l \alpha_{l-2m} s_l^k = 0\right\}.$$

In this case, the prediction errors (1) always satisfy the following system of equations:

$$(45) \quad \sum_l \alpha_l e_{2m+l}^k = 0, \quad m = 1, \dots, N_{k-1}.$$

These relations can be used to determine the scale coefficients. For the weight functions mentioned so far in this section it is possible to store the values e_i^k with odd indices, i.e.,

$$(46) \quad d_j^k = e_{2j-1}^k,$$

and use (45) in order to formulate a system of equations

$$(47) \quad \sum \alpha_{2l} e_{2j+2l}^k = - \sum \alpha_{2l-1} e_{2j+2l-1}^k$$

for the unknowns (e_{2j}^k) .

The definition of scale coefficients given in (46) leads to a simple definition for the operator G_k :

$$(48) \quad (G_k)_{ij} = \delta_{2i-1,j}.$$

The operator E_k is then obtained from (47), and its columns provide a set of basis vectors for $\mathcal{N}(D_k^{k-1})$.

REMARK 5.1. *The wavelet theory provides many examples of weight functions that define a multiresolution setting by the local-average discretization procedure (35). In this context the space \mathcal{F} is forced to be $L^2(R)$; thus, in reward, we have all the geometric properties of a Hilbert space. This very natural requirement from the point of view of functional analysis leads to heavy restrictions on the dilation relation that the weight function has to satisfy. Many weight functions are automatically ruled out, in particular the δ -function which does not belong to $L^2(R)$. The box function leads to a stable multiresolution analysis: the Haar basis, which is rarely used because of its poor approximation properties. The hat function does not lead to an orthonormal wavelet basis, but it is used as the scaling function in the biorthogonal framework, an attempt to build stable multiresolution decompositions in $L^2(R)$ removing some of the heavy restrictions needed to obtain orthonormal wavelet basis.*

The prediction operator is constructed using an appropriate reconstruction technique which is very much linked to the space of functions on which the discretization operators are to be applied. We shall come back to this subject in each particular case.

In this paper we consider only linear reconstruction techniques. Thus, to ensure stability, we look for the existence of the hierarchical reconstruction obtained by “cosmetic refinement.” The existence of this operator guarantees the stability of the associated multiresolution transform.

6. Interpolatory MR analysis. We consider $\mathcal{F} = \mathcal{C}([0, 1])$ and

$$(49) \quad \mathcal{D}_k : \mathcal{C}[0, 1] \longrightarrow S^k, \quad \bar{f}_j^k = (\mathcal{D}_k f)_j = f(x_j^k), \quad 0 \leq j \leq J_k.$$

In this case, $N_k = \dim S^k = \dim X^k = J_k + 1$. Since the δ -function satisfies (in the sense of distributions) the dilation relation

$$(50) \quad \omega(x) = 2\omega(2x) \Rightarrow \alpha_0 = 1,$$

we have

$$(51) \quad \mathcal{N}(D_k^{k-1}) = \{s^k \in S^k \mid s_{2i}^k = 0\}; \quad (D_k^{k-1})_{ij} = \delta_{2i,j}.$$

The operators G_k (defined as in (48)) and E_k have the following expressions:

$$(52) \quad (G_k^{k-1})_{ij} = \delta_{2i-1,j}, \quad (E_k)_{ij} = \delta_{i,2j-1}.$$

REMARK 6.1. *It is interesting to notice that we can also define the set of basis functions for $\mathcal{N}(D_k^{k-1})$ “à la wavelet,” i.e.,*

$$(\mu_j^k)_l = (-1)^l \alpha_{2j-l+1} = (-1)^l \delta_{2j+1,l}$$

since $(D_k^{k-1} \mu_j^k)_i = \sum_{l=0}^{J_k-1} \delta_{2i,l} (-1)^l \delta_{2j+1,l} = 0$. The transfer matrix E_k obtained using this basis is (up to a sign change) the same as (52).

A reconstruction procedure for this discretization is given by any operator \mathcal{R}_k such that

$$(53) \quad \mathcal{R}_k : S^k \longrightarrow \mathcal{C}[0, 1]; \quad \mathcal{D}_k \mathcal{R}_k \bar{f}^k = \bar{f}^k,$$

which means

$$(54) \quad (\mathcal{R}_k \bar{f}^k)(x_j^k) = \bar{f}_j^k = f(x_j^k).$$

Therefore, $\mathcal{R}_k \bar{f}^k$ should be a continuous function that interpolates the data \bar{f}^k on X^k . From these considerations, it is clear that Dirac's delta function gives rise to interpolatory multiresolution settings, which should be appropriate for multiscale representations of continuous functions.

We shall follow a slightly different notation in this section, and we shall denote by \mathcal{I}_k any interpolatory reconstruction of the data \bar{f}^k , i.e.,

$$(\mathcal{R}_k \bar{f}^k)(x) = \mathcal{I}_k(x; \bar{f}^k).$$

The encoding and decoding algorithms (4) and (5) take the following simple form (see [15, 16]):

$$\mu(\bar{f}^L) = M \bar{f}^L \text{ (Encoding)}$$

$$(55) \quad \left\{ \begin{array}{ll} \text{Do } k = L, 1 & \\ \bar{f}_j^{k-1} = \bar{f}_{2j}^k, & 0 \leq j \leq J_{k-1}, \\ d_j^k = \bar{f}_{2j-1}^k - \mathcal{I}_{k-1}(x_{2j-1}^k; \bar{f}^k), & 1 \leq j \leq J_{k-1}. \end{array} \right.$$

$$\bar{f}^L = M^{-1} \mu(\bar{f}^L) \text{ (Decoding)}$$

$$(56) \quad \left\{ \begin{array}{ll} \text{Do } k = 1, L & \\ \bar{f}_{2j}^k = \bar{f}_j^{k-1}, & 0 \leq j \leq J_{k-1}, \\ \bar{f}_{2j-1}^k = \mathcal{I}_{k-1}(x_{2j-1}^k; \bar{f}^{k-1}) + d_j^k, & 1 \leq j \leq J_{k-1}. \end{array} \right.$$

Up to this point, the type of interpolatory procedure has yet to be specified. The framework allows the user to choose a particular type of procedure depending on the application at hand. Data-independent interpolatory techniques lead to linear reconstruction operators. We can then use the machinery developed in the previous sections to study the stability properties of the associated multiresolution schemes as well as the additional functional structure that comes with them.

In [15, 16] different types of data-independent interpolatory reconstructions are considered. Here we shall review the piecewise polynomial interpolation, because we shall use it to work out an example in the hat-average framework.

6.1. Piecewise polynomial interpolation. Let \mathcal{S} denote the stencil

$$\mathcal{S} = \mathcal{S}(r, s) = \{-s, -s+1, \dots, -s+r\}, \quad r \geq s > 0, \quad r \geq 1,$$

and let $\{L_m(y)\}_{m \in \mathcal{S}}$ denote the Lagrange interpolation polynomials for this stencil

$$L_m(y) = \prod_{l=-s, l \neq m}^{-s+r} \left(\frac{y-l}{m-l} \right), \quad L_m(i) = \delta_{i,m}, \quad i \in \mathcal{S}.$$

It is clear that

$$q_j^k(x; \bar{f}^k, r, s) = \sum_{m=-s}^{-s+r} \bar{f}_{j+m}^k L_m \left(\frac{x - x_j^k}{h_k} \right)$$

interpolates $f(x)$ at the points $\{x_{j-s}^k, \dots, x_{j-s+r}^k\}$. Thus,

$$(57) \quad \mathcal{I}_k(x, \bar{f}^k) := q_j^k(x; \bar{f}^k, r, s), \quad x \in [x_{j-1}^k, x_j^k], \quad 1 \leq j \leq J_k$$

is a piecewise polynomial function that interpolates $f(x)$ at the grid points $\{x_j^k\}$.

The situation $r = 2s - 1$ corresponds to an interpolatory stencil which is symmetric around the given interval. In this case, $q_j^k(x; \bar{f}^k, r, s)$ is the unique polynomial of degree r that interpolates $f(x)$ at the $r + 1 = 2s$ points $\{x_{j-s}^k, \dots, x_{j+s-1}^k\}$.

When the given function is periodic, i.e.,

$$\bar{f}_{-j}^k = \bar{f}_{J_k-j}^k, \quad \bar{f}_{J_k+j+1}^k = \bar{f}_{j+1}^k, \quad 0 \leq j \leq J_k,$$

the data to construct the polynomial function $\mathcal{I}_k(x, f^k)$ using centered stencils is always available. If the function is not periodic we choose one-sided stencils, of $r + 1 = 2s$ points, at intervals where the centered-stencil choice would require function values which are not available. The definition of $\mathcal{I}_k(x; \bar{f}^k)$ in the nonperiodic case is as follows:

$$(58) \quad \mathcal{I}_k(x; \bar{f}^k) = \begin{cases} q_j^k(x; \bar{f}^k; r, j), & 1 \leq j \leq s - 1, \\ q_j^k(x; \bar{f}^k; r, s), & s \leq j \leq J_k - s + 1, \\ q_j^k(x; \bar{f}^k; r, j - J_k + r), & J_k - s + 2 \leq j \leq J_k. \end{cases}$$

Periodic functions can be treated as nonperiodic ones, but it is customary not to do so because of the better approximation properties of centered reconstructions against noncentered ones.

Observe that if $f(x) = P(x)$, where $P(x)$ is a polynomial of degree less than or equal to r , then $q_j^k(x; f^k, r, s) = f(x)$ for $x \in [x_{j-1}^k, x_j^k]$, i.e., $\mathcal{I}_k(x, f^k) = f(x)$. Therefore, the space of polynomials of degree less than or equal to r satisfies $\mathcal{I}_k \mathcal{D}_k f = f$, and the order of the reconstruction procedure is $r + 1 = 2s$. In addition, for smooth functions,

$$\mathcal{I}_k(x, f^k) = f(x) + O(h_k)^{2s}.$$

The stability of the associated multiresolution scheme follows directly from the *existence* of the hierarchical form of the reconstruction procedure used to define the scheme, which in this context is the interpolatory reconstruction $\mathcal{I}_k(x; \bar{f}^k)$. Thus, convergence (as $L \rightarrow \infty$) of the functions $\varphi_i^{m,L}(x)$ obtained by repeated interpolation of $\eta_i^m = \delta_i^m$:

$$\varphi_i^{m,L}(x) = \mathcal{R}_L A_m^L \delta_i^m = \mathcal{I}_L(x; A_m^L \delta_i^m),$$

immediately implies the stability of the associated multiresolution scheme.

The set of values $A_m^L \delta_i^m$, used to construct the cosmetic refinement sequence $\varphi_i^{m,L}$, is obtained after repeatedly applying the prediction scheme to climb from the k th level to the L th level. Let us look a bit more closely at this refinement process for

data-independent interpolatory techniques (i.e., for linear reconstruction operators). For these, we can write

$$\mathcal{I}_k(x; \bar{f}^k) = \sum_i \bar{f}_i^k \mathcal{I}_k(x; \delta_i^k).$$

Then, since $P_{k-1}^k = \mathcal{D}_k \mathcal{I}_{k-1}$,

$$\tilde{f}_i^k = (P_{k-1}^k \bar{f}^{k-1})_i = \mathcal{I}_{k-1}(x_i^k; \bar{f}^{k-1}) = \sum_j \bar{f}_j^{k-1} \mathcal{I}_{k-1}(x_i^k, \delta_j^{k-1}).$$

Therefore (recall $x_{2i}^k = x_i^{k-1}$),

$$(59) \quad (P_{k-1}^k)_{2i,j} = \mathcal{I}_{k-1}(x_{2i}^k, \delta_j^{k-1}) = \delta_{i,j}, \quad (P_{k-1}^k)_{2i-1,j} = \mathcal{I}_{k-1}(x_{2i-1}^k, \delta_j^{k-1}),$$

and we can write

$$(60) \quad \begin{cases} \tilde{f}_{2i}^k &= (P_{k-1}^k \bar{f}^{k-1})_{2i} = \bar{f}_i^{k-1}, & 0 \leq i \leq J_{k-1}, \\ \tilde{f}_{2i-1}^k &= (P_{k-1}^k \bar{f}^{k-1})_{2i-1} = \sum_j \bar{f}_j^{k-1} \mathcal{I}_{k-1}(x_{2i-1}^k, \delta_j^{k-1}), & 1 \leq i \leq J_{k-1}. \end{cases}$$

The predicted values at even points of the level k are exactly the same as the values given on the $(k-1)$ st grid at those same points, while the odd points are computed using a *refinement rule* involving points in the $(k-1)$ st level.

These refinement rules define, in the limit as $k \rightarrow \infty$, an infinite set of points (corresponding to all dyadic rationals in $[0, 1]$). The question of whether or not these values admit a continuous extension to the whole interval $[0, 1]$ is directly related to the convergence properties of the sequences $\varphi_i^{m,L}(x)$ as $L \rightarrow \infty$. This is a consequence of the following lemma.

LEMMA 6.1. *If*

$$(61) \quad A_k^L \bar{f}^k \xrightarrow{L \rightarrow \infty} f(x) \in \mathcal{C}[0, 1] \quad \Rightarrow \quad I_L(x; A_k^L \bar{f}^k) \xrightarrow{L \rightarrow \infty} f(x) \text{ in } \mathcal{C}[0, 1].$$

The notion of convergence for the sequence $A_k^L \bar{f}^k$ to the function $f(x)$ is used in subdivision refinement. It means that we have

$$(62) \quad \lim_{L \rightarrow \infty} \|A_k^L \bar{f}^k - f(\cdot)\|_\infty = 0,$$

where $f(\cdot)$ denotes the sequence $\{f(x_j^L)\}_{j=0}^{J_L}$. The proof of the lemma uses only elementary arguments of piecewise polynomial interpolation theory and the fact that a continuous function on $[0, 1]$ is uniformly continuous.

Let us consider for a moment the periodic case. Since the symmetric interpolatory procedure can now be used at each subinterval, a straightforward algebraic manipulation shows that

$$(63) \quad \begin{cases} \tilde{f}_{2i}^k &= (P_{k-1}^k \bar{f}^{k-1})_{2i} = \bar{f}_i^{k-1}, \\ \tilde{f}_{2i-1}^k &= (P_{k-1}^k \bar{f}^{k-1})_{2i-1} = \sum_{m=-s}^{-s+r} L_m(-1/2) \bar{f}_{i+m}^{k-1}. \end{cases}$$

Observe that the periodicity assumption leads to refinement rules which are independent of the level of resolution. Moreover, we can consider (63) $\forall i \in \mathbb{Z}$. Formula (63) describes in fact the refinement rule of an *interpolatory subdivision scheme*, a special subclass of stationary subdivision schemes. The interesting connection with stationary subdivision allows us to prove many essential properties of the multiresolution scheme with very little effort.

7. Periodicity: The connection with uniform subdivision refinement.

In this section we summarize the main results in refinement subdivision theory which are of interest to us. These results have been extracted mainly from [5, 12] and [13]. We refer the interested reader to these works for proofs and more information on subdivision refinement as well as its relation to other branches of mathematics.

Subdivision schemes are efficient tools for computer-aided curve and surface design. A *uniform binary subdivision* scheme, S , is defined in terms of a *mask* consisting of a (finite) set of nonzero coefficients $\{\gamma_l : l \in Z\}$ (we restrict our attention to one dimension since it is the case of interest to us here). The scheme is defined by a refinement rule as follows:

$$p_i^k = (Sp^{k-1})_i = \sum_l \gamma_{i-2l} p_l^{k-1},$$

and the set $\{p_i^k\}$ is referred to as the set of *control points* at the k th level. The last relation can also be expressed as

$$(64) \quad \begin{aligned} p_{2i}^k &= \sum_l \gamma_{2l} p_{i-l}^{k-1}, \\ p_{2i-1}^k &= \sum_l \gamma_{2l-1} p_{i-l}^{k-1}, \end{aligned}$$

which implies that at every stage of the computation the values computed at previous stages are further “refined,” and new values at intermediate points are added. The “refinement rules” are the same for all stages of the computation and are given by the *mask* of the subdivision scheme.

Given an initial set of control points $\{p_i^0, i \in Z\}$, the binary subdivision refinement of this sequence defines, in the limit $k \rightarrow \infty$, an infinite set of points (corresponding to all rational numbers whose denominator is a power of 2 and integer translations of these) and the obvious question is then whether or not these values admit a continuous extension to the real axis. If this is so for any set of initial control points, the scheme is said to converge. Mathematically, this is expressed by saying that for every bounded sequence $\{p_i, i \in Z\}$, there exists a continuous function $f_p(x)$ such that

$$\lim_{k \rightarrow \infty} \left\| S^k p - f_p \left(\frac{\cdot}{2^k} \right) \right\|_{\infty} = 0,$$

where

$$\left\| S^k p - f_p \left(\frac{\cdot}{2^k} \right) \right\|_{\infty} = \max_{j \in Z} \left| (S^k p)_j - f_p \left(\frac{j}{2^k} \right) \right|.$$

A convergent subdivision scheme, S , uniquely determines a compactly supported function satisfying the functional equation

$$(65) \quad \varphi(x) = \sum_j \gamma_j \varphi(2x - j).$$

This function, also called the *fundamental* or *S-refinable* function, is obtained by recursive subdivision of the sequence $p_i^0 = \delta_{0,i}$, i.e., $\varphi(x) =: S^{\infty} \delta_0$.

Conversely, if (65) admits a continuous, compactly supported solution $\varphi(x)$ and if, in addition, the set $\{\varphi(x - j)\}, j \in Z$, is *linearly independent*, then the subdivision scheme defined by the mask $\{\gamma_l\}$ converges.

These facts play a pivotal role in the analysis of subdivision schemes. Some of the more immediate consequences are as follows.

1. Convergence needs to be checked only for the δ sequence, and the limiting curve obtained by recursive subdivision of p^0 is

$$(S^\infty p^0)(x) = \sum_i p_i^0 \varphi(x - i).$$

2. The regularity of the limit functions is the same as the regularity of $\varphi(x)$.

Let us go back to our interpolation-based prediction schemes. The periodicity assumption implies that we consider the same choice of stencil for all points and all resolution levels. As a consequence, the reconstruction operator is translation invariant and independent of the resolution level. Harten proves in [16] the following.

THEOREM 7.1. *Let \mathcal{R}_k be a linear operator which is translation invariant,*

$$(66) \quad (\mathcal{R}_k \delta_{j-q}^k)(x - qh_k) = (\mathcal{R}_k \delta_j^k)(x) \quad \forall q \in Z,$$

and independent of the level of resolution,

$$(67) \quad (\mathcal{R}_k \delta_0^k)(x) = (\mathcal{R}_{k-1} \delta_0^{k-1})(2x) \quad \forall k.$$

Then, the quantities $\langle \omega_l^k, \mathcal{R}_{k-1} \delta_0^{k-1} \rangle$ are independent of k , and if we define

$$(68) \quad \gamma_l := \langle \omega_l^k, \mathcal{R}_{k-1} \delta_0^{k-1} \rangle,$$

then

$$(69) \quad (P_{k-1}^k)_{i,j} = \gamma_{i-2j}.$$

Notice that $\gamma_l = 0$ when $\text{supp } \omega_l^k \cap \text{supp } \mathcal{R}_{k-1} \delta_0^{k-1} = \emptyset$; thus, for the piecewise polynomial reconstruction techniques of section 6, only a finite number of γ 's will be nonzero.

Under these premises the predicted values, $\tilde{f}^k = P_{k-1}^k \bar{f}^{k-1}$, satisfy

$$(70) \quad \tilde{f}_i^k = \sum_j \gamma_{i-2j} \bar{f}_j^{k-1} \quad \equiv \quad \begin{aligned} \tilde{f}_{2i}^k &= \sum_l \gamma_{2l} \bar{f}_{i-l}^{k-1}, \\ \tilde{f}_{2i+1}^k &= \sum_l \gamma_{2l+1} \bar{f}_{i-l}^{k-1}. \end{aligned}$$

Thus, for a reconstruction sequence satisfying the hypothesis of Theorem 7.1, the prediction process (70) takes the form of a *uniform binary subdivision scheme*. The mask of this scheme is given in terms of the prediction operator which, thanks to the periodicity assumption, is independent of k and can be considered as an infinite matrix. We have, in fact,

$$\gamma_j = (P \delta_0)_j,$$

where P is the infinite matrix representation of the k -independent prediction operator.

The particular case described by (63) corresponds to a subdivision scheme with the property that at each stage of the iteration the previous control points are left unchanged. It is then obvious that, if the limit curve exists, it must interpolate the original control points. Stationary subdivision schemes with these properties are referred to as *interpolatory subdivision schemes*.

It is easy to see that for interpolatory subdivision schemes, the existence of a compactly supported $\varphi(x)$ satisfying (65) guarantees the convergence of the subdivision

scheme, since the interpolatory property readily implies that the set $\{\varphi(x-j), j \in Z\}$ is linearly independent.

Later in this article, we shall also use a characterization of the regularity of $\varphi(x)$ (and thus of the limit functions obtained by the recursive subdivision) in terms of the *divided-difference schemes* (see [13]). A necessary condition for a subdivision scheme to converge is that

$$(71) \quad \sum_l \gamma_{2l} = 1 = \sum_l \gamma_{2l+1}$$

(this condition simply guarantees that the subdivision scheme reproduces constants). This condition also guarantees the existence of the so-called *divided difference scheme* which we shall denote as S_1 . When (71) holds, there exists a subdivision scheme, S_1 , that satisfies

$$dp^{k+1} = S_1 dp^k,$$

where $(dp^k)_i = (p_{i+1}^k - p_i^k)/h_k$ is the vector whose components are the “divided differences” of p^k .

Divided difference schemes can be used to study the smoothness properties of the limit functions. For our purpose, it is enough to cite the following theorem, which can be found, e.g., in [12, Theorem 5.3].

THEOREM 7.2. *Let S be an interpolatory subdivision scheme which reproduces polynomials up to degree ν . Then the following two conditions are equivalent:*

1. S converges to C^ν -limit functions;
2. S_n converges to $C^{\nu-n}$ -limit functions for $n = 1, 2, \dots, \nu$.

Moreover, it is not hard to relate the limiting functions of S and S_1 . Let us apply S to a sequence of initial control points p^0 and let us call $f(x) = S^\infty p^0$. Then it is easy to see (see [13]) that for $q^0 = dp^0$ we have $S_1^\infty q^0 = f'(x)$. If we denote $g(x) = S_1^\infty p^0$, we must have

$$f'(x) = g(x+1) - g(x) \quad \equiv \quad f(x) = \int_x^{x+1} g(s) ds.$$

The argument can be repeated to obtain the relation between the limiting functions of S and S_2 and so on.

7.1. Interpolatory multiresolution: The existence of the hierarchical form.

The periodic case. Relation (63) shows that the prediction process in the periodic case is just an interpolatory subdivision scheme with odd mask coefficients $\gamma_{2l-1} = L_{-l}(-1/2)$, $-s \leq l \leq -s + r$. Thus, Lemma 6.1 plus the connection with stationary subdivision imply that the existence of the cosmetic refinement limit in $\mathcal{F} = \mathcal{C}[0, 1]$ for any discrete set of initial data is a direct consequence of the existence of such a limit for $\bar{f}^0 = \delta_0^0$.

Here, we are only concerned with symmetric interpolatory procedures, $r = 2s - 1$. For $r = 1$, \mathcal{I}_k is the piecewise-linear interpolation which is hierarchical and, as a consequence, the associated multiresolution scheme is stable. Deslauriers and Dubuc were the first to study interpolatory subdivision schemes based on symmetric interpolation. In [10], they proved convergence of the recursive refinement process (63) for $r = 3$ and $r = 5$. The considerations above lead us to conclude that the hierarchical

reconstruction (75) exists for $r = 3$ and $r = 5$; hence the associated multiresolution schemes are stable.

The connection with subdivision refinement leads to further consequences. If the cosmetic refinement limit of the sequence $\bar{f}^0 = \delta_0^0$ exists, then it satisfies a functional relation given by the *mask* of the subdivision scheme, i.e., the coefficients of the prediction matrix. In our case, relations (59) and (65) imply that the fundamental function $\varphi(x)$ must satisfy the dilation relation

$$(72) \quad \varphi(x) = \varphi(2x) + \sum_{m=-s}^{-s+r} L_m(-1/2)\varphi(2x+2m+1)$$

and $\text{supp}(\varphi) = [-(2s+1), 2s+1]$.

Thus, under the periodicity assumption, if

$$(73) \quad \exists \quad \mathcal{I}_k^\infty(x; \delta_0^0) := \lim_{k \rightarrow \infty} \mathcal{I}_k(x; \bar{A}_0^k \delta_0^0) = \varphi_0(x),$$

then $\varphi_0(x) = \varphi(\frac{x}{h_0})$ and $\forall i$ and k

$$(74) \quad \exists \quad \mathcal{I}_k^\infty(x; \delta_i^k) := \lim_{L \rightarrow \infty} \mathcal{I}_k(x; A_k^L \delta_i^k) = \varphi_i^k(x) = \varphi\left(\frac{x}{h_k} - i\right).$$

These facts were also proven in [16], without explicitly using the subdivision refinement theory.

The bottom figure in Figure 1 shows (a scaled and shifted version of) the fundamental function of the interpolatory subdivision scheme defined by (63) for $r = 3$. In this case $\text{supp}\varphi(x) = [-3, 3]$. The function displayed is a translated version of $\varphi_0(x) = \varphi(x/h_0)$, with $h_0 = 1/8$, thus it has $\text{supp}\varphi_0(x) = [-3h_0, 3h_0]$ and has been obtained computing $M^{-1}(\delta_4, 0, \dots, 0)$ with $J_0 = 8$ and $L = 7$. Results for larger L are indistinguishable from these.

It is interesting to notice that the hierarchical form of \mathcal{I}_k , which we refer to as \mathcal{I}_k^∞ since it is obtained via the cosmetic refinement limiting process, has the form

$$(75) \quad \mathcal{I}_k^\infty(x; \bar{f}^k) = \sum_i \bar{f}_i^k \varphi_i^k$$

and that the two-level relation between the limiting functions φ_i^k in (28) can also be obtained directly, using the dilation relation for $\varphi(x)$.

On the other hand, (32) becomes in this context

$$\psi_j^k = \sum_l (E_k)_{lj} \varphi_l^k = \sum_l \delta_{l, 2j-1} \varphi_l^k = \varphi_{2j-1}^k.$$

Thus

$$\psi_j^k = \psi\left(\frac{x}{2h_k} - j\right), \quad \text{where} \quad \psi(x) = \varphi(2x+1).$$

Moreover, using the hierarchical interpolation (75), we can write (33) as

$$(76) \quad I_L^\infty(x; \mathcal{D}_L f) = I_0^\infty(x; \mathcal{D}_0 f) + \sum_{k=1}^L \sum_{j=1}^{J_{k-1}} d_j^k(f) \varphi_{2j-1}^k(x)$$

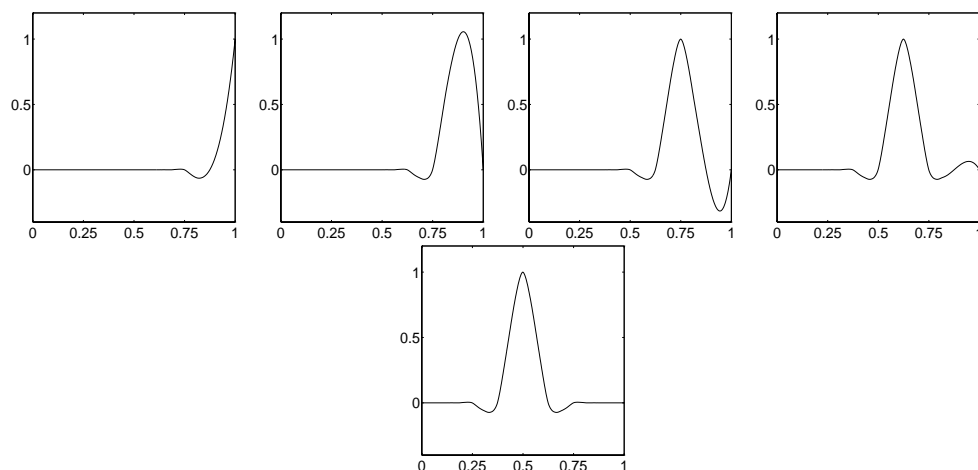


FIG. 1. Limiting functions for interpolatory multiresolution. $J_0 = 8$, $r = 3$. Top: “special” boundary functions. Bottom: periodic case.

with $d_j^k(f) = f(x_{2j-1}^k) - \mathcal{I}_{k-1}(x_{2j-1}^k; \mathcal{D}_{k-1}f)$.

In the finite element context, the set

$$\left(\left\{ \left\{ \varphi_{2i-1}^k \right\}_{i=1}^{J_{k-1}} \right\}_{k=1}^L, \left\{ \varphi_i^0 \right\}_{i=0}^{J_0} \right)$$

is referred to as a hierarchical basis (see [22]).

The nonperiodic case. Because of Lemma 6.1, we study the convergence properties of the sequence $A_k^L \delta_i^k$. It is a simple matter to check, numerically, the convergence properties of this sequence. For this, we simply apply the inverse multiresolution transform (5) to the sequence $(u^0, 0, \dots, 0)$ (for L large enough) with $u^0 = \delta_i^k$, i.e., taking the starting grid as the k th grid and all the scale coefficients as zero.

For illustration purposes we display in Figure 1 the $r = 3$ case. The displayed results correspond to $A_0^7 \delta_i^0$, with $J_0 = 8$, $L = 7$, and $i = 4, \dots, 8$. These are basically indistinguishable from $A_0^L \delta_i^0$ for L larger than 7. The remaining functions (not shown), $\varphi_i^{0,7}$, $i = 0, \dots, 3$, are specular images (with respect to the left boundary) of $\varphi_i^{0,7}$, $i = 8, \dots, 5$. The regular behavior of the displayed functions seems to indicate that these limiting functions are at least continuous.

In Figure 2 we display the limiting functions for $J_0 = 16$ and $i = 12, \dots, 16$. The limiting functions displayed are the “boundary” limiting functions. For $r = 3$ there are four special boundary functions at each resolution level (starting from $J_0 = 8$). All ‘interior’ limiting functions appear to be translates of the same function which in turn is a scaled version of the bottom figure in Figure 1. This is a direct consequence of the fact that these ‘interior’ functions have been obtained using only the centered interpolation procedure. It is worth noticing that all boundary limit functions corresponding to higher levels of resolution are also scaled versions of those for lower resolution levels as long as only one of the boundaries makes its influence felt (compare the top row of Figure 1 and Figure 2).

An analytic proof of the existence of these boundary functions seems to be feasible [6]. This would justify the stability of the nonperiodic multiresolution transforms.

8. Hat-weighted multiresolution. The discretization process defined in (35) requires that the function under consideration be integrated against scaled translates

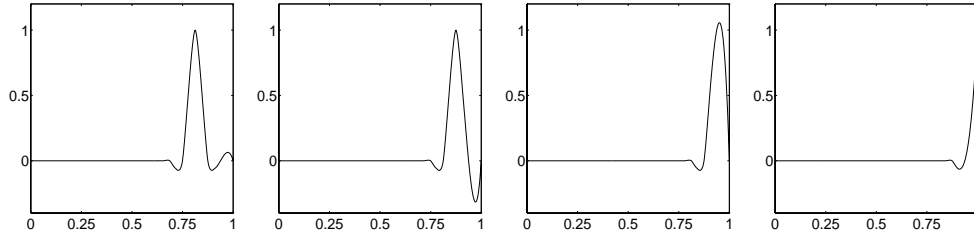


FIG. 2. Limiting functions for interpolatory multiresolution. $J_0 = 16$, $r = 3$. “Special” boundary functions.

of the weight function $\omega(x)$. When the weight function is the hat function, (42), these are continuous functions; thus δ -type singularities can be allowed.

Let us notice first that it is sufficient to consider weighted averages \bar{f}_i^k for $1 \leq i \leq N_k = J_k - 1$, since these averages contain information on f over the whole interval $[0, 1]$. Therefore,

$$\mathcal{D}_k : \mathcal{F} \longrightarrow S^k, \quad \bar{f}_i^k = (\mathcal{D}_k f)_i = \langle f, \omega_i^k \rangle, \quad 1 \leq i \leq N_k = J_k - 1.$$

The weighted averages \bar{f}_i^k provide a representation of the information contents at the k th level of resolution of any piecewise smooth function defined on the unit interval with a finite number of δ -type singularities in the open interval. Thus, we consider \mathcal{F} to be the space of piecewise smooth functions in $[0, 1]$ with, at most, a finite number of δ -type singularities in $(0, 1)$, and S^k is the space of finite sequences of $N_k = J_k - 1$ components.

The dilation relation for the hat function (43) leads to

$$(77) \quad \omega_i^{k-1} = \frac{1}{4}\omega_{2i-1}^k + \frac{1}{2}\omega_{2i}^k + \frac{1}{4}\omega_{2i+1}^k$$

or, in terms of the weighted averages,

$$(78) \quad \bar{f}_i^{k-1} = \frac{1}{4}\bar{f}_{2i-1}^k + \frac{1}{2}\bar{f}_{2i}^k + \frac{1}{4}\bar{f}_{2i+1}^k.$$

The relation above can be used to compute the N_{k-1} hat averages at the $(k-1)$ st level from the N_k hat averages on the k th level. The decimation matrix is then an $N_{k-1} \times N_k$ matrix given explicitly by the following expression:

$$(79) \quad (D_k^{k-1})_{ij} = \frac{1}{4}\delta_{2i-1,j} + \frac{1}{2}\delta_{2i,j} + \frac{1}{4}\delta_{2i+1,j}.$$

Any reconstruction procedure \mathcal{R}_k for this multiresolution setting must satisfy

$$(80) \quad \mathcal{R}_k : S_k \longrightarrow \mathcal{F}, \quad (\mathcal{D}_k \mathcal{R}_k \bar{f}^k)_i = \langle \mathcal{R}_k \bar{f}^k, \omega_i^k \rangle = \bar{f}_i^k.$$

In what follows we shall describe a reconstruction procedure for the hat-averaged framework which is a generalization of the “reconstruction via primitive function” developed in the context of cell averages (see [16, 18]). We refer to this procedure as “reconstruction via second primitive.”

Let f be a piecewise smooth function in $[0, 1]$ with a finite number of δ -jumps in $(0, 1)$. Then, f can be expressed as

$$(81) \quad f = f_p + \sum_l h_l \delta(x - a_l), \quad 0 < a_l < 1,$$

where f_p represents the piecewise smooth part of f and there are only a finite number of terms in the sum at the right-hand side of the expression.

We define its “second primitive” as follows:

$$(82) \quad H(x) := \int_0^x \int_0^y f_p(z) dz dy + \sum h_l (x - a_l)_+,$$

where

$$(x - a)_+ = \begin{cases} 0, & x \leq a, \\ x - a, & x \geq a. \end{cases}$$

Then $H(x)$ is a continuous piecewise smooth function which satisfies the following relation:

$$(83) \quad \bar{f}_i^k = \langle f, \omega_i^k \rangle = \frac{1}{h_k^2} (H_{i+1}^k - 2H_i^k + H_{i-1}^k), \quad 1 \leq i \leq J_k - 1,$$

where $H_i^k = H(x_i^k)$, $0 \leq i \leq J_k$. When $f(x)$ is an integrable function, (83) is easily proven by integrating by parts. If $f(x) = \delta(x - a)$, $a \in (0, 1)$, then its second primitive (82) is $H(x) = (x - a)_+$ and it is also a straightforward matter to prove that

$$\bar{f}_i^k = \langle \delta(x - a), \omega_i^k \rangle = \omega_i^k(a) = \frac{(x_{i+1}^k - a)_+ - 2(x_i^k - a)_+ + (x_{i-1}^k - a)_+}{h_k^2}.$$

The definition of $H(x)$ in (82) leads to $H_{J_k}^k = H(1) = \int_0^1 \int_0^y f_p(z) dz dy + \sum h_l = q$ and $H_0^k = H(0) = 0 \forall k$. Changing the lower limits in (82) is equivalent to modifying the basic definition (82) by a first-order polynomial, and it amounts to computing different values for $H(0)$ and $H(1)$. Once these have been specified, (83) establishes a one-to-one correspondence between the sets $\{\bar{f}_i^k\}_{i=1}^{J_k-1}$ and $\{H_i^k\}_{i=1}^{J_k-1}$. $H(x)$ in (82) is one of the second primitives of $f(x)$ (i.e., a function satisfying (83)); however, another second primitive, which we also label $H(x)$,

$$(84) \quad H(x) = \int_0^x \int_0^y f_p(z) dz dy + \sum h_l (x - a_l)_+ - qx, \quad q = \int_0^1 \int_0^y f_p(z) dz dy + \sum h_l,$$

is theoretically more appealing because, for this particular one $H_0^k = H(0) = 0$ and $H_{J_k}^k = H(1) = 0$. In this case we have

$$(85) \quad h_k^2 \bar{f}^k = M H^k, \quad M_{i,j} = \begin{cases} -2, & i = j, \\ 1, & |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, knowledge of the hat averages of a given function $f \in \mathcal{F}$ is equivalent to knowledge of the point values of its second primitive (84). We can then interpolate the point-values of the “second primitive” by any interpolation procedure $\mathcal{I}_k(x; H^k)$ and define

$$(86) \quad (\mathcal{R}_k \bar{f}^k)(x) := \frac{d^2}{dx^2} \mathcal{I}_k(x; H^k).$$

In most cases (for all elementary interpolation techniques), $\mathcal{I}_k(x; H^k)$ is a continuous, piecewise smooth function. Its first derivative will also be a piecewise smooth function possibly with discontinuities at the grid points of the k th level. Thus its second

derivative must be considered in the distribution sense. $R_k \bar{f}^k(x)$ may have a finite number of δ -type singularities, which will be located, by construction, at those (interior) grid points where the first derivative of $\mathcal{I}_k(x; H^k)$ has a jump discontinuity. This fact is a consequence of the following lemma, whose proof is a straight application of the definition of a distributional derivative and shall be omitted.

LEMMA 8.1. *Let $P(x)$ be a piecewise smooth function of the form*

$$(87) \quad P(x) = \begin{cases} P_L(x) & \text{if } x < 0, \\ P_R(x) & \text{if } x > 0. \end{cases}$$

Then, its derivative in the distribution sense is

$$(88) \quad \frac{d}{dx} P(x) = \tilde{P}(x) + (P_R(0) - P_L(0))\delta(x),$$

where

$$(89) \quad \tilde{P}(x) = \begin{cases} \frac{d}{dx} P_L(x) & \text{if } x < 0, \\ \frac{d}{dx} P_R(x) & \text{if } x > 0. \end{cases}$$

Lemma 8.1 implies that if the interpolatory function is defined as

$$(90) \quad \mathcal{I}_k(x; H^k) = q_j^k(x; H^k) \quad \text{for } x \in [x_{j-1}^k, x_j^k],$$

with $q_j^k(x; H^k)$ smooth, we have in the distribution sense

$$(91) \quad (\mathcal{R}_k \bar{f}^k)(x) = \tilde{\mathcal{I}}_k(x) + \sum_{j=1}^{J_k-1} s_j^k \delta(x - x_j^k),$$

where $\tilde{\mathcal{I}}_k$ is a piecewise smooth function defined as

$$(92) \quad \tilde{\mathcal{I}}_k(x) = \frac{d^2}{dx^2} q_j^k(x; H^k) \quad \text{for } x \in [x_{j-1}^k, x_j^k],$$

and

$$(93) \quad s_j^k = \left(\frac{d}{dx} q_{j+1}^k(x; H^k) - \frac{d}{dx} q_j^k(x; H^k) \right)_{x=x_j^k} = \mathcal{I}_k'(x_j^k+0; H^k) - \mathcal{I}_k'(x_j^k-0; H^k).$$

Obviously, $(\mathcal{R}_k \bar{f}^k)(x) \in \mathcal{F}$. To prove that \mathcal{R}_k in (86) is a right inverse of \mathcal{D}_k we need the following lemma.

LEMMA 8.2. *Let $g(x)$ be a locally integrable function and let $G(x)$ be a continuous function in $[0, 1]$ twice differentiable in $[x_{l-1}^k, x_l^k]$, $l = j, j+1$, such that $G''(x) = g(x)$ almost everywhere in $[x_{l-1}^k, x_l^k]$ for $l = j, j+1$. Then*

$$(94) \quad \langle g, \omega_j^k \rangle = \bar{g}_j^k = \frac{1}{h_k} [G'(x_j^k - 0) - G'(x_j^k + 0)] + \frac{1}{h_k^2} (G_{j+1}^k - 2G_j^k + G_{j-1}^k).$$

Proof. For a locally integrable function $g(x)$,

$$\begin{aligned}\bar{g}_j^k &= \frac{1}{h_k} \int_{x_{j-1}^k}^{x_j^k} g(x) \left(1 + \frac{x - x_j^k}{h_k}\right) dx + \frac{1}{h_k} \int_{x_j^k}^{x_{j+1}^k} g(x) \left(1 - \frac{x - x_j^k}{h_k}\right) dx \\ &= \frac{1}{h_k} \int_{x_{j-1}^k}^{x_j^k} G''(x) \left(1 + \frac{x - x_j^k}{h_k}\right) dx + \frac{1}{h_k} \int_{x_j^k}^{x_{j+1}^k} G''(x) \left(1 - \frac{x - x_j^k}{h_k}\right) dx.\end{aligned}$$

The result follows from integration by parts in the expression above. \square

Notice that (92) implies that we can apply Lemma 8.2 to $g = \tilde{\mathcal{I}}_k$ and $G = \mathcal{I}_k$. We thus obtain

$$\begin{aligned}(\mathcal{D}_k \mathcal{R}_k \bar{f}^k)_j &= \langle \mathcal{R}_k \bar{f}^k, \omega_j^k \rangle = \langle \tilde{\mathcal{I}}_k(x) + \sum_{l=1}^{J_k-1} \delta(x - x_l^k) s_l^k, \omega_j^k \rangle \\ &= \langle \tilde{\mathcal{I}}_k(x), \omega_j^k \rangle + \sum_{l=1}^{J_k-1} s_l^k \langle \delta(x - x_l^k), \omega_j^k \rangle \\ &= \frac{1}{h_k^2} (\mathcal{I}_k(x_{j+1}^k; H^k) - 2\mathcal{I}_k(x_j^k; H^k) + \mathcal{I}_k(x_{j-1}^k; H^k)) \\ &\quad + \frac{1}{h_k} (\mathcal{I}_k'(x_j^k - 0; H^k) - \mathcal{I}_k'(x_j^k + 0; H^k) + s_j^k) \\ &= \frac{1}{h_k^2} [H_{j+1}^k - 2H_j^k + H_{j-1}^k] = \bar{f}_j^k, \quad 1 \leq j \leq J_k - 1.\end{aligned}$$

Hence $\mathcal{D}_k \mathcal{R}_k = I_{S^k}$.

REMARK 8.1. Our function space \mathcal{F} is a subspace of $\mathcal{C}^*[0, 1]$, the space of continuous linear functionals on $\mathcal{C}[0, 1]$. This is a Banach space with the norm

$$(95) \quad f \in \mathcal{C}^*[0, 1] \quad \rightarrow \quad \|f\| = \max\{|\langle f, \alpha \rangle|; \quad \alpha(x) \in \mathcal{C}[0, 1], \|\alpha\|_\infty = 1.\}$$

The Riesz representation theorem identifies any $f \in \mathcal{C}^*[0, 1]$ with a function σ_f of bounded variation ($BV[0, 1]$ from now on) so that for each $\alpha(x) \in \mathcal{C}[0, 1]$,

$$(96) \quad \langle f, \alpha \rangle = \int_0^1 \alpha(x) d\sigma_f$$

(the integral above is a Riemann-Stieltjes integral). Our construction shows that $\sigma_{\mathcal{R}_k \bar{f}^k}$ can be chosen as $d/dx(\mathcal{I}_k(x; H^k))$, i.e.,

$$(97) \quad \langle \mathcal{R}_k \bar{f}^k, \alpha \rangle = \int_0^1 \alpha(x) d\left(\frac{d}{dx} \mathcal{I}_k(x; H^k)\right).$$

We could choose $\mathcal{F} = \mathcal{C}^*[0, 1]$, but we would run into some difficulties in proving the fundamental relation (83). It turns out that our development can also be carried out when \mathcal{F} is the subspace of $\mathcal{C}^*[0, 1]$ corresponding to BV functions that are the sum of an absolutely continuous function and a jump function (i.e., BV functions without singular part). However, our description of \mathcal{F} seems much more natural from the numerical analysis point of view, and it is also the natural function space in some potential applications (e.g., vortex methods).

The prediction operator is now computed from \mathcal{R}_k using (8):

$$\begin{aligned} (P_{k-1}^k \bar{f}^{k-1})_j &= (\mathcal{D}_k \mathcal{R}_{k-1} \bar{f}^{k-1})_j = \langle \mathcal{R}_{k-1} \bar{f}^{k-1}, \omega_j^k \rangle \\ &= \langle \tilde{\mathcal{I}}_{k-1}, \omega_j^k \rangle + \sum_{l=1}^{J_{k-1}-1} s_l^{k-1} \langle \delta(x - x_l^{k-1}), \omega_j^k \rangle, \quad 1 \leq j \leq J_k - 1. \end{aligned}$$

Observe that Lemma 8.2 implies

$$\begin{aligned} \langle \tilde{\mathcal{I}}_{k-1}, \omega_j^k \rangle &= \frac{1}{h_k} (\mathcal{I}'_{k-1}(x_j^k - 0; H^{k-1}) - \mathcal{I}'_{k-1}(x_j^k + 0; H^{k-1})) \\ &+ \frac{1}{h_k^2} (\mathcal{I}_{k-1}(x_{j-1}^k; H^{k-1}) - 2\mathcal{I}_{k-1}(x_j^k; H^{k-1}) + \mathcal{I}_{k-1}(x_{j+1}^k; H^{k-1})), \end{aligned}$$

so that for $j = 2m$ we have

$$\mathcal{I}'_{k-1}(x_j^k - 0) - \mathcal{I}'_{k-1}(x_j^k + 0) = \mathcal{I}'_{k-1}(x_m^{k-1} - 0) - \mathcal{I}'_{k-1}(x_m^{k-1} + 0) = -s_m^{k-1},$$

$$\langle \delta(x - x_l^{k-1}), \omega_j^k \rangle = \delta_{j,m},$$

while for $j = 2m + 1$ we obtain

$$\mathcal{I}'_{k-1}(x_j^k - 0) - \mathcal{I}'_{k-1}(x_j^k + 0) = 0, \quad \langle \delta(x - x_l^{k-1}), \omega_j^k \rangle = 0.$$

Thus, for each j , $1 \leq j \leq J_k - 1$

$$(98) \quad (P_{k-1}^k \bar{f}^{k-1})_j = \frac{1}{h_k^2} (\mathcal{I}_{k-1}(x_{j-1}^k; H^{k-1}) - 2\mathcal{I}_{k-1}(x_j^k; H^{k-1}) + \mathcal{I}_{k-1}(x_{j+1}^k; H^{k-1})).$$

REMARK 8.2. Our description of the reconstruction operator allows us to exploit an interesting interplay between the hat and interpolatory multiresolution settings. Let $\tilde{\mathcal{D}}_k$ denote the discretization by point-values operator. Then, if $\{\{\bar{f}_j^k\}_{j=1}^{N_k}\}_{k=0}^L$ is a hat-multiresolution decomposition of $\{\mathcal{D}_L f\}$, the sequence $\{\{H_j^k\}_{j=0}^{J_k}\}_{k=0}^L$ is an interpolatory-multiresolution decomposition of $\{\tilde{\mathcal{D}}_L H\}$ (see section 6). The interpolatory procedure \mathcal{I}_k used to define \mathcal{R}_k in (86) serves as reconstruction operator in interpolatory multiresolution settings. Let us denote the prediction operator for the latter as \tilde{P}_{k-1}^k , i.e., $\tilde{P}_{k-1}^k = \tilde{\mathcal{D}}_k \mathcal{I}_{k-1}$. Then (98) can also be expressed as follows:

$$(99) \quad (P_{k-1}^k \bar{f}^{k-1})_j = \frac{1}{h_k^2} \left((\tilde{P}_{k-1}^k H^{k-1})_{j-1} - 2(\tilde{P}_{k-1}^k H^{k-1})_j + (\tilde{P}_{k-1}^k H^{k-1})_{j+1} \right).$$

I.e., the predicted values in the hat-framework are simply the second divided differences of the corresponding set of predicted values in the interpolatory framework.

REMARK 8.3. The function space \mathcal{F} can also be chosen as $L^1[0, 1]$ or even $L^2[0, 1]$ (since the hat function belongs to $L^\infty[0, 1]$ and $L^2[0, 1]$). As suggested by one of the referees of this paper, an alternative description of the reconstruction operator consists in defining a polynomial of degree p that satisfies

$$\int_{x_i^k}^{x_{i+1}^k} R(x) \omega_j^k(x) = h_k \bar{f}_j^k$$

and which is built from a stencil that contains the i th cell. If the nodes are two-by-two distinct and the stencil contains $p+1$ elements the resulting linear system is solvable.

However, this construction obscures the relation between the hat and interpolatory multiresolution settings. Since much of our development is based on exploiting this relation, we shall stick to our original derivation, which also provides us with an explicit formula for $\mathcal{R}_k \bar{f}^k$ as an element of $\mathcal{C}^*[0, 1]$.

To complete the multiresolution scheme we need to give an explicit description of the transfer operators E_k and G_k . The value of the scale coefficients d^k will be directly related to the definition of these operators. Notice that $\dim \mathcal{N}(D_k^{k-1}) = N_k - N_{k-1} = J_k - J_{k-1} = J_{k-1}$.

Because of the dilation relation satisfied by the hat function, system (45) takes the form

$$(100) \quad e_{2i}^k = -\frac{1}{2}e_{2i-1}^k - \frac{1}{2}e_{2i+1}^k.$$

The natural choice of transfer operators described in section 5 is

$$(101) \quad d^k = G_k e^k, \quad d_j^k = e_{2j-1}^k, \quad 1 \leq j \leq J_{k-1},$$

and

$$(102) \quad e^k = E_k d^k \quad \begin{cases} e_{2j-1}^k &= d_j^k, & 1 \leq j \leq J_{k-1}, \\ e_{2j}^k &= -\frac{1}{2}(d_j^k + d_{j+1}^k), & 1 \leq j \leq J_{k-1} - 1. \end{cases}$$

With all the necessary ingredients of the multiresolution scheme appropriately specified, the explicit description of the hat-based multiresolution transform and its inverse is as follows:

$$\mu(\bar{f}^L) = M \bar{f}^L \text{ (Encoding)}$$

$$(103) \quad \begin{cases} \text{Do } k = L, \dots, 1, \\ \bar{f}_i^{k-1} = \frac{1}{4}(\bar{f}_{2i-1}^k + 2\bar{f}_{2i}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1} - 1, \\ d_i^k = \bar{f}_{2i-1}^k - (P_{k-1}^k \bar{f}^{k-1})_{2i-1}, & 1 \leq i \leq J_{k-1}, \end{cases}$$

$$\bar{f}^L = M^{-1} \mu(\bar{f}^L) \text{ (Decoding)}$$

$$(104) \quad \begin{cases} \text{Do } k = 1, \dots, L, \\ \bar{f}_{2i-1}^k = (P_{k-1}^k \bar{f}^{k-1})_{2i-1} + d_i^k, & 1 \leq i \leq J_{k-1}, \\ \bar{f}_{2i}^k = 2\bar{f}_i^{k-1} - \frac{1}{2}(\bar{f}_{2i-1}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1} - 1. \end{cases}$$

Notice that

$$\bar{f}_{2i}^k = (P_{k-1}^k \bar{f}^{k-1})_{2i} + (E_k d^k)_{2i} \quad \equiv \quad \bar{f}_{2i}^k = 2\bar{f}_i^{k-1} - \frac{1}{2}(\bar{f}_{2i-1}^k + \bar{f}_{2i+1}^k),$$

because, from the definition of the odd averages in (104) and (98), we obtain

$$\bar{f}_{2i}^k + \frac{1}{2}(\bar{f}_{2i-1}^k + \bar{f}_{2i+1}^k) = (P_{k-1}^k \bar{f}^{k-1})_{2i} - \frac{1}{2}(d_i^k + d_{i+1}^k)$$

$$\begin{aligned}
& + \frac{1}{2}[(P_{k-1}^k \bar{f}^{k-1})_{2i-1} + d_i^k + (P_{k-1}^k \bar{f}^{k-1})_{2i+1} + d_{i+1}^k] \\
& = (P_{k-1}^k \bar{f}^{k-1})_{2i} + \frac{1}{2}[(P_{k-1}^k \bar{f}^{k-1})_{2i-1} + (P_{k-1}^k \bar{f}^{k-1})_{2i+1}] \\
& = \frac{1}{2h_k^2}(H_{i-1}^{k-1} - 2H_i^{k-1} + H_{i+1}^{k-1}) = 2\bar{f}_i^{k-1}.
\end{aligned}$$

REMARK 8.4. *In algorithms (103) and (104), only the predicted values at the odd-indexed grid points are computed. With the definition given in (90), relation (98) becomes*

$$(P_{k-1}^k \bar{f}^{k-1})_{2i-1} = \frac{1}{h_k^2} [q_i^{k-1}(x_{2i-2}^k, H^{k-1}) - 2q_i^{k-1}(x_{2i-1}^k, H^{k-1}) + q_i^{k-1}(x_{2i}^k, H^{k-1})].$$

Using the Newton form of the polynomial pieces $q_l^{k-1}(x, H^{k-1})$, it is easy to write $(P_{k-1}^k \bar{f}^{k-1})_{2i-1}$ in terms of $\{\bar{f}_l^{k-1}\}$ and their divided differences. Hence, the role of $H(x)$ is that of a design tool, and it never needs to be computed explicitly.

8.1. Stability with respect to perturbations. As pointed out in section 3, Theorem 3.1, when $\mathcal{R}_k \mathcal{D}_k$ are bounded linear operators stability of the direct multiresolution transform is a consequence of the nested character of the sequence of discretization. Boundedness of the operator $\mathcal{R}_k \mathcal{D}_k$ is a consequence of the fact that \mathcal{R}_k is a linear operator onto a finite dimensional subspace and therefore is bounded. In fact, we can write

$$\mathcal{R}_k \mathcal{D}_k f = \mathcal{R}_k \sum_{i=1}^{N_k} \bar{f}_i^k \delta_i^k \quad \Rightarrow \quad \|\mathcal{R}_k \mathcal{D}_k f\| \leq \sum_{i=1}^{N_k} |\bar{f}_i^k| \|\mathcal{R}_k \delta_i^k\|.$$

Since $f \in C^*[0, 1]$, and $\|\omega_j^k(x)\|_\infty = 1/h_k$, we have

$$|\bar{f}_j^k| = |\langle f, \omega_j^k \rangle| \leq \|\omega_j^k\|_\infty \|f\| = \frac{\|f\|}{h_k};$$

then

$$(105) \quad \|\mathcal{R}_k \mathcal{D}_k f\| \leq \|f\| \sum_{i=1}^{N_k} \frac{1}{h_k} \|\mathcal{R}_k \delta_i^k\| = C^k \|f\|$$

with

$$C^k = \sum_{i=1}^{N_k} \frac{1}{h_k} \|\mathcal{R}_k \delta_i^k\|.$$

This proves that the direct multiresolution transform is a stable algorithm.

Stability of the inverse multiresolution transform follows from the *existence* of a hierarchical form of the reconstruction sequence. We use again the link between the hat and interpolatory frameworks to prove the following.

LEMMA 8.3. *If the piecewise polynomial interpolatory reconstructions \mathcal{I}_k (90) are hierarchical, then so are the corresponding reconstructions obtained by (86).*

Proof. The hierarchical property of the interpolatory reconstructions means that (we follow the notation of Remark 8.2)

$$\mathcal{I}_k \tilde{\mathcal{D}}_k \mathcal{I}_{k-1} H^{k-1} = \mathcal{I}_{k-1} H^{k-1} \quad \equiv \quad \mathcal{I}_k \tilde{P}_{k-1}^k H^{k-1} = \mathcal{I}_{k-1} H^{k-1}.$$

The above relation can also be expressed as follows:

$$(106) \quad \text{If } \tilde{H}^k = \tilde{P}_{k-1}^k H^{k-1}, \quad \text{then} \quad \mathcal{I}_k(x; \tilde{H}^k) = \mathcal{I}_{k-1}(x, H^{k-1}).$$

To prove that

$$\mathcal{R}_k \bar{f}^k(x) = \frac{d^2}{dx^2} \mathcal{I}_k(x; H^k)$$

is also hierarchical, observe that (99) implies

$$(107) \quad (\mathcal{D}_k \mathcal{R}_{k-1} \bar{f}^{k-1})_j = (P_{k-1}^k \bar{f}^{k-1})_j = \frac{1}{h_k^2} (\tilde{H}_{j-1}^k - 2\tilde{H}_j^k + \tilde{H}_{j+1}^k).$$

Therefore, the set $\tilde{P}_{k-1}^k H^{k-1} = \{\tilde{H}_j^k\}$ in (106) corresponds to the point values of a “second primitive” of the function $(\mathcal{R}_{k-1} \bar{f}^{k-1})(x)$ on the k th grid. Hence,

$$\mathcal{R}_k \mathcal{D}_k \mathcal{R}_{k-1} \bar{f}^{k-1}(x) = \frac{d^2}{dx^2} \mathcal{I}_k(x; \tilde{H}^k) = \frac{d^2}{dx^2} \mathcal{I}_{k-1}(x; H^{k-1}) = \mathcal{R}_{k-1} \bar{f}^{k-1}(x),$$

which completes the proof. \square

As a consequence, interpolatory techniques that lead to hierarchical reconstructions in the interpolatory multiresolution set-up (e.g. splines) lead directly to hierarchical reconstructions in the hat-average framework and, thus, to stable multiresolution schemes.

Let us assume now that the reconstruction \mathcal{I}_k of the underlying interpolatory framework is not hierarchical itself but admits a hierarchical form \mathcal{I}_k^∞ . Notice that $\mathcal{I}_k^\infty(x; \tilde{\mathcal{D}}_k H)$ must be a continuous function which interpolates $H(x)$ on the k th grid but it might not be a piecewise smooth function (derivatives might fail to exist at an infinite set of points, for example). In any case,

$$(108) \quad \mathcal{R}_k^\infty(x; \mathcal{D}_k f) := \frac{d^2}{dx^2} \mathcal{I}_k^\infty(x; \tilde{\mathcal{D}}_k H)$$

(where $H(x)$ is a second primitive of $f(x)$) is always well defined in the distribution sense. Moreover, we have the following theorem.

THEOREM 8.4. *In the distribution sense,*

$$(109) \quad \lim_{L \rightarrow \infty} \mathcal{R}_L A_k^L \bar{f}^k = \mathcal{R}_k^\infty \bar{f}^k.$$

Proof. Since \mathcal{I}_k^∞ is the hierarchical form of \mathcal{I}_k ,

$$\mathcal{I}_k^\infty(x; \tilde{\mathcal{D}}_k H) = \lim_{L \rightarrow \infty} \mathcal{I}_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H$$

and the convergence takes place in $\mathcal{C}[0, 1]$.

Let us consider $\phi(x) \in \mathcal{C}_0^\infty[0, 1]$, an infinitely differentiable function with compact support in $[0, 1]$. Let $f \in \mathcal{F}$ and $H(x)$ be its second primitive. Then (see Remark 8.2)

$$(110) \quad \langle \mathcal{R}_L A_k^L \mathcal{D}_k f, \phi \rangle = \left\langle \frac{d^2}{dx^2} \mathcal{I}_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \phi \right\rangle = \left\langle \mathcal{I}_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \frac{d^2}{dx^2} \phi \right\rangle.$$

Now, since $\mathcal{I}_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H \rightarrow \mathcal{I}_k^\infty(x; \tilde{\mathcal{D}}_k H)$ in $\mathcal{C}[0, 1]$,

$$(111) \lim_{L \rightarrow \infty} \left\langle \mathcal{I}_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \frac{d^2}{dx^2} \phi \right\rangle = \left\langle \mathcal{I}_k^\infty(\cdot; \tilde{\mathcal{D}}_k H), \frac{d^2}{dx^2} \phi \right\rangle = \left\langle \frac{d^2}{dx^2} \mathcal{I}_k^\infty(\cdot; \tilde{\mathcal{D}}_k H), \phi \right\rangle,$$

and the proof is complete. \square

The space of distributions is not a normed space, thus Theorem 8.4 does not guarantee the stability of the inverse multiresolution transform. To guarantee that \mathcal{R}_k^∞ gives the cosmetic refinement limiting functions we need convergence in $\mathcal{C}^*[0, 1]$. In certain cases, convergence in $\mathcal{C}^*[0, 1]$ can also be proven, thus the hat-multiresolution schemes derived from these interpolatory procedures will be stable.

THEOREM 8.5. *Let us assume that $\mathcal{I}_k^\infty(x; H^k) \in \mathcal{C}^2[0, 1]$ for any sequence $\{H^k\}$. Then, with the usual notation,*

$$\mathcal{R}_L A_k^L \bar{f}^k \xrightarrow{L \rightarrow \infty} \mathcal{R}_k^\infty \bar{f}^k = \frac{d^2}{dx^2} \mathcal{I}_k^\infty(x; H^k) \quad \text{in } \mathcal{C}^*[0, 1].$$

Proof. Notice that $\mathcal{R}_k^\infty \bar{f}^k(x) \in \mathcal{C}[0, 1]$ (and also to $\mathcal{C}^*[0, 1]$).

To prove the theorem we need to estimate the operator norm

$$\|\mathcal{R}_L A_k^L \bar{f}^k - \mathcal{R}_k^\infty \bar{f}^k\| = \max_{\|\alpha\|_\infty=1} |\langle \mathcal{R}_L A_k^L \bar{f}^k - \mathcal{R}_k^\infty \bar{f}^k, \alpha \rangle|.$$

Let us consider $\alpha(x) \in \mathcal{C}[0, 1]$ and evaluate

$$\Lambda_\alpha = \int_0^1 \alpha(x) [\mathcal{R}_L A_k^L \bar{f}^k(x) - \mathcal{R}_k^\infty \bar{f}^k(x)] dx.$$

We recall that

$$\mathcal{R}_L A_k^L \bar{f}^k = \tilde{\mathcal{I}}_L(x) + \sum_{j=1}^{N_L} s_j^L \delta(x - x_j^L)$$

with

$$\begin{aligned} \tilde{\mathcal{I}}_L(x) &= \frac{d^2}{dx^2} q_j^L(x; \tilde{A}_k^L H^k), \quad x \in [x_{j-1}^L, x_j^L], \\ s_j^L &= \mathcal{I}'_L(x_j + 0; \tilde{A}_k^L H^k) - \mathcal{I}'_L(x_j - 0; \tilde{A}_k^L H^k). \end{aligned}$$

As usual, \tilde{A}_k^L stands for the successive prediction operator in the interpolatory multiresolution framework.

Then

$$\Lambda_\alpha = \sum_{j=1}^{N_L} \int_{x_{j-1}^L}^{x_j^L} \alpha(x) \frac{d^2}{dx^2} \left[q_j^L(x; \tilde{A}_k^L H^k) - \mathcal{I}_k^\infty(x; H^k) \right] dx + \sum_{j=1}^{N_L} \alpha(x_j^L) s_j^L.$$

Defining

$$g_{L,j}(x) := \frac{d^2}{dx^2} \left[q_j^L(x; \tilde{A}_k^L H^k) - \mathcal{I}_k^\infty(x; H^k) \right]$$

we get

$$|\Lambda_\alpha| \leq \sum_{j=1}^{N_L} \|\alpha(x)\|_\infty \int_{x_{j-1}^L}^{x_j^L} |g_{L,j}(x)| dx + \sum_{j=1}^{N_L} \|\alpha(x)\|_\infty |s_j^L|.$$

Theorem 3.3 implies that $\mathcal{I}_k^\infty(x_j^L; H^k) = (\tilde{A}_k^L H^k)_j$, hence $q_j^L(x; \tilde{A}_k^L H^k)$ is a polynomial that interpolates the function $\mathcal{I}_k^\infty(x; H^k)$ at the points of the stencil used for its construction (which necessarily contains the two points x_{j-1}^L and x_j^L).

It is not hard to prove that if $\mathcal{I}_k^\infty(x; H^k)$ is a function in $\mathcal{C}^2[0, 1]$,

$$|g_{L,j}(x)| = o(1), \quad x \in [x_{j-1}^L, x_j^L], \quad \text{and} \quad |s_j^L| = o(h_L).$$

Then

$$|\Lambda_\alpha| \leq \sum_{j=1}^{N_L} o(h_L) = o(1),$$

which proves the theorem. \square

8.2. Hierarchical reconstructions and their wavelet structure. The close relationship between the interpolatory multiresolution and the hat-based multiresolution analysis can be further exploited. In fact, the latter directly inherits much of the algebraic structure of the former. To see how this works, let us notice first that

$$\mathcal{I}_{k-1}(x_{2i}^k; H^{k-1}) = \mathcal{I}_{k-1}(x_i^{k-1}; H^{k-1}) = H_i^{k-1} = H_{2i}^k.$$

Thus

$$\begin{aligned} (112) \quad e_{2j-1}^k(f) &= \bar{f}_{2j-1}^k - (P_{k-1}^k \bar{f}^{k-1})_{2i-1} \\ &= \frac{1}{h_k^2} [H_{2j}^k - 2H_{2j-1}^k + H_{2j-2}^k] \\ &\quad - \frac{1}{h_k^2} (I_j^{k-1}(x_{2j-2}^k; H^{k-1}) - 2I_j^{k-1}(x_{2j-1}^k; H^{k-1}) + I_j^{k-1}(x_{2j}^k; H^{k-1})) \\ &= -\frac{2}{h_k^2} (H_{2j-1}^k - I_j^{k-1}(x_{2j-1}^k; H^{k-1})) = -\frac{2}{h_k^2} e_{2j-1}^k(H). \end{aligned}$$

In the above relation, $e_j^k(f)$ are the prediction errors in the multiresolution scheme associated to the hat averages, while $e_j^k(H)$ are the prediction errors in the related interpolatory multiresolution scheme.

If $\{\mathcal{I}_k \mathcal{D}_k\}$ is hierarchical, then (32) and (27) lead, in the interpolatory context, to the following expression:

$$(113) \quad \mathcal{I}_k(x; H^k) - \mathcal{I}_{k-1}(x; H^{k-1}) = \sum_{j=1}^{J_{k-1}} d_j^k(H) \tilde{\varphi}_{2j-1}^k(x)$$

with $\tilde{\varphi}_j^k = \mathcal{I}_k(x; \delta_j^k)$. Lemma 8.3 implies that the reconstruction for the hat-average set-up which is obtained from \mathcal{I}_k is also hierarchical (provided it is a well-defined reconstruction in the hat-average framework). Then differentiating (113) and taking into account that

$$d_j^k(f) = -\frac{2}{h_k^2} d_j^k(H), \quad 1 \leq j \leq J_{k-1},$$

we obtain

$$(114) \quad (\mathcal{R}_k \bar{f}^k)(x) - (\mathcal{R}_{k-1} \bar{f}^{k-1})(x) = -\frac{h_k^2}{2} \sum_{j=1}^{J_k-1} d_j^k(f) \frac{d^2}{dx^2} \tilde{\varphi}_{2j-1}^k(x) = \sum_{j=1}^{J_k-1} d_j^k(f) \psi_j^k(x),$$

with

$$(115) \quad \psi_j^k(x) = -\frac{h_k^2}{2} \frac{d^2}{dx^2} \tilde{\varphi}_{2j-1}^k(x), \quad 1 \leq j \leq J_k-1.$$

Hence

$$(116) \quad (\mathcal{R}_L \bar{f}^L)(x) = \sum_{i=1}^{J_0-1} \bar{f}_i^0 \varphi_i^0 + \sum_{k=1}^L \sum_{j=1}^{J_k-1} d_j^k(f) \psi_j^k(x)$$

with $\varphi_i^k = \mathcal{R}_k \delta_i^k$. The functional structure in the hat-average framework can, thus, be deduced directly from the interpolatory framework.

8.3. Piecewise polynomial interpolation. In this section we shall study in detail one particular type of multiresolution scheme in the hat framework. These schemes are constructed using the piecewise interpolatory techniques of section 6.1.

Let $f \in \mathcal{F}$ and $H(x)$ its second primitive (84). Following the notation of section 6.1, let $q_j^k(x; H^k, r, s)$ be the (unique) polynomial of degree r that interpolates $H(x)$ at the points $\mathcal{S}_j = \{x_{j-s}^k, \dots, x_{j+r-s}^k\}$. Then

$$\mathcal{I}_k(x, H^k) := q_j^k(x; H^k, r, s), \quad x \in [x_{j-1}^k, x_j^k], \quad 0 \leq j \leq J_k$$

is a piecewise polynomial function that interpolates $H(x)$ at the grid points $\{x_j^k\}$. As in section 6.1, we shall use centered stencils, i.e., $r = 2s - 1$, except at those intervals where the centered choice would require function values which are not available. When the given function is periodic, centered stencils can always be chosen. In the nonperiodic case, \mathcal{I}_k is defined using one-sided stencils for some of its polynomial pieces (see (58)). We recall that a periodic function can always be treated as a nonperiodic one. However, when possible, centered reconstructions are preferred because their approximation errors are usually smaller than those of their noncentered counterparts.

Notice that if $f(x) = Q(x)$, a polynomial function of degree q , then $H(x) = P(x)$ where $P(x)$ is also a polynomial of degree $q + 2$ satisfying $P''(x) = Q(x)$. If $q + 2 \leq r$ we shall have

$$\mathcal{I}_k(x, H^k) = H(x)$$

and the definition (86) of the reconstruction operator readily implies that

$$\mathcal{R}_k(x; \bar{f}^k) = f(x).$$

Hence, the space of polynomials of degree up to $r - 2$ satisfies $\mathcal{R}_k \mathcal{D}_k f = f$. Moreover, if $f(x)$ is a smooth function, so is $H(x)$, thus

$$\mathcal{I}_k(x, H^k) = H(x) + O(h_k^{r+1}) \quad \text{and} \quad \mathcal{R}_k \bar{f}^k(x) = f(x) + O(h_k^{r-1}).$$

The *formal order* of the reconstruction is thus $p = r - 1 = 2s - 2$.

To study the stability properties of these schemes, we consider the cosmetic refinement limiting process and analyze the convergence properties of the functions $\varphi_i^{m,L}$:

$$\varphi_i^{m,L} = \mathcal{R}_L A_m^L \delta_i^m.$$

As in section 6.1, we start by considering the periodic case first.

The periodic case. To treat periodic functions, it is simpler to include also the hat average of f at one of the endpoints of the interval (the hat average at the other endpoint is the same due to periodicity) in the multiresolution scheme. Then in the periodic case we consider the range of \mathcal{D}_k to be the space of sequences with J_k components, i.e.,

$$\mathcal{D}_k f = (\bar{f}_i^k)_{i=1}^{J_k}, \quad \bar{f}_i^k = \langle f, \omega_i^k \rangle.$$

A straightforward, but rather lengthy, algebraic computation leads to the following multiresolution algorithm:

$$\mu(\bar{f}^L) = M \bar{f}^L \text{ (Encoding)}$$

$$(117) \quad \left\{ \begin{array}{ll} \text{Do } k = L, 1 \\ \bar{f}_i^{k-1} = \frac{1}{4}(\bar{f}_{2i-1}^k + 2\bar{f}_{2i}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1}, \\ d_i^k = \bar{f}_{2i-1}^k - \sum_{l=1}^{s-1} \beta_l (\bar{f}_{i+l-1}^{k-1} + \bar{f}_{i-l}^{k-1}), & 1 \leq i \leq J_{k-1}. \end{array} \right.$$

$$\bar{f}^L = M^{-1} \mu(\bar{f}^L) \text{ (Decoding)}$$

$$(118) \quad \left\{ \begin{array}{ll} \text{Do } k = 1, L \\ \bar{f}_{2i-1}^k = d_i^k + \sum_{l=1}^{s-1} \beta_l (\bar{f}_{i+l-1}^{k-1} + \bar{f}_{i-l}^{k-1}), & 1 \leq i \leq J_{k-1}, \\ \bar{f}_{2i}^k = 2\bar{f}_i^{k-1} - \frac{1}{2}(\bar{f}_{2i-1}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1}. \end{array} \right.$$

The coefficients β_l for reconstruction procedures of order $p = 3, 5, 7$ are as follows:

$$(119) \quad \left\{ \begin{array}{ll} p = 2 & \Rightarrow \beta_1 = \frac{1}{2}, \\ p = 4 & \Rightarrow \beta_1 = \frac{19}{32}, \quad \beta_2 = -\frac{3}{32}, \\ p = 6 & \Rightarrow \beta_1 = \frac{162}{256}, \quad \beta_2 = -\frac{39}{256}, \quad \beta_3 = \frac{5}{256}. \end{array} \right.$$

Because of periodicity, the reconstruction sequence $\{\mathcal{R}_k\}$ satisfies the hypothesis of Theorem 7.1 and as a consequence the prediction process takes the form of a uniform subdivision scheme. In our case it is not hard to see that the prediction process can be written as

$$(120) \quad \begin{aligned} p_{2i}^k &= \sum_l \gamma_{2l} p_{i-l}^{k-1}, \\ p_{2i+1}^k &= \sum_l \gamma_{2l+1} p_{i-l}^{k-1} \end{aligned}$$

with $\gamma_l = (P\delta_0)_l$ given as (assume $\beta_l = 0, l \geq s$)

$$(121) \quad \gamma_{2l} = \begin{cases} 2 - \beta_1, & l = 0, \\ -\frac{\beta_l + \beta_{l+1}}{2}, & 0 < l \leq s-1, \\ -\frac{\beta_{-l} + \beta_{-l+1}}{2}, & -(s-1) \leq l < 0, \end{cases} \quad \gamma_{2l-1} = \begin{cases} \beta_l, & 1 \leq l \leq (s-1), \\ \beta_{-l+1}, & -s+2 \leq l \leq 0. \end{cases}$$

Thus, the sequence $A_k^L \bar{f}_i^k$ can be interpreted as the sequence of control points obtained after L applications of the subdivision scheme on the initial set of control points $\{\bar{f}^k\}$. The convergence of this sequence to a continuous function plays a very important role, as in the interpolatory case, in proving the existence of the cosmetic refinement limit, and thus the stability of the associated multiresolution scheme.

To see this, let us follow the notation of [12, 13] and denote by S the subdivision scheme associated to a centered interpolation procedure (degree r) of the type described in section 6.1. Relation (99) shows that the predicted values in the hat framework are precisely the second divided differences of the predicted values in the interpolatory framework, hence the subdivision scheme (120) given by the hat prediction process is nothing but S_2 , its *second divided difference* scheme, i.e., the subdivision scheme for the second divided differences of the control points generated by S .

As mentioned in section 7, S converges to \mathcal{C}^2 functions if and only if S_2 converges to continuous functions. If this is the case, $\mathcal{I}_k^\infty(x; H^k)$ is a \mathcal{C}^2 function, hence, by Theorem 8.5, $\mathcal{R}_k^\infty(x; \bar{f}^k) = \frac{d^2}{dx^2} \mathcal{I}_k^\infty(x; H^k)$ is the cosmetic refinement limit of the sequence \bar{f}^k . The existence of the cosmetic refinement limit implies the stability of the associated multiresolution scheme.

Other interesting consequences are obtained directly from the relation with stationary subdivision. Convergence needs to be checked only for the initial sequence $p_i^0 = \delta_{0,i}$. The limiting function corresponding to this sequence is the *fundamental* function of the scheme, $\varphi(x)$. This function satisfies the following dilation relation:

$$(122) \quad \begin{aligned} \varphi(x) = (2 - \beta_1)\varphi(2x) - \sum_{l=1}^{s-1} \frac{\beta_l + \beta_{l+1}}{2} [\varphi(2x - 2l) + \varphi(2x + 2l)] \\ + \sum_{l=1}^{s-1} \beta_l [\varphi(2x - 2l + 1) + \varphi(2x + 2l - 1)]. \end{aligned}$$

It is easy to see that $\text{supp } \varphi(x) = [-p, p]$.

Let us denote by $\tilde{\varphi}(x)$ the fundamental function corresponding to S . The relation between $\varphi(x)$ and $\tilde{\varphi}(x)$ is as follows (see section 7):

$$\tilde{\varphi}(x) = \int_{x-1}^x \int_y^{y+1} \varphi(z) dz dy, \quad \text{hence,} \quad \tilde{\varphi}''(x) = \varphi(x+1) - 2\varphi(x) + \varphi(x-1).$$

Figure 3 displays a numerical computation of scaled translates of the fundamental functions for $r = 3, 5, 7$. They have been obtained numerically by computing $M^{-1}(\delta_i^0, 0, \dots, 0)$ with $J_0 = 8$ and $L = 7$ (as in the interpolatory case, results for larger L are indistinguishable from these). Thanks to the connection with stationary subdivision, and in particular with interpolatory subdivision, we know what are the smoothness properties of these functions.

Deslauriers and Dubuc [10] realize that the fundamental function of the interpolatory scheme for $r = 3$ is not \mathcal{C}^2 . In fact, Daubechies and Lagarias [9], with more sophisticated techniques, prove that $\tilde{\varphi}''(x)$ fails to exist at all dyadic rationals. Since

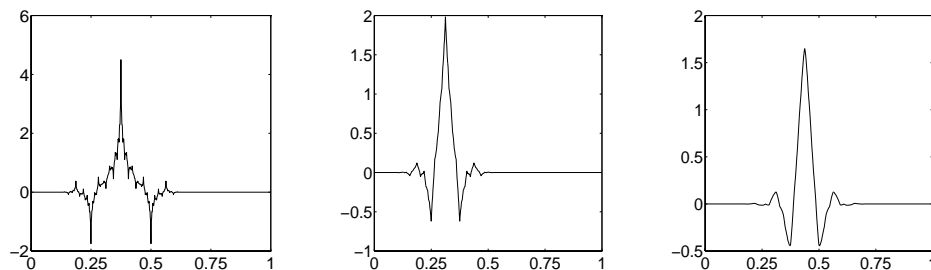


FIG. 3. Limiting functions for the periodic case: left $r = 3$, center $r = 5$, right $r = 7$.

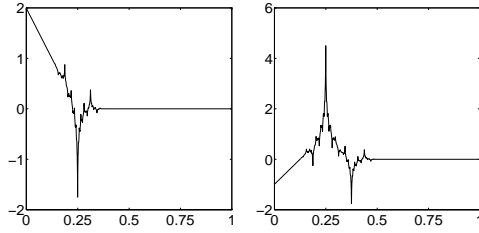
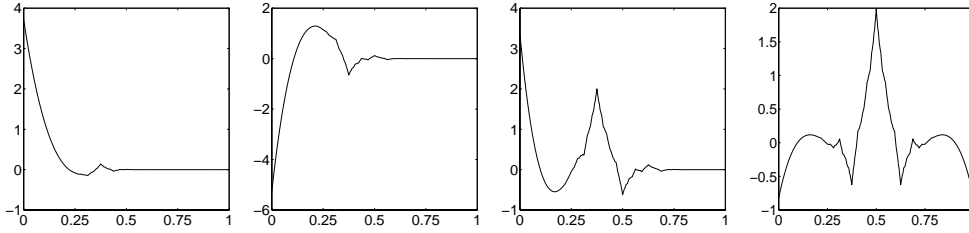
the fundamental function of the interpolatory scheme is not \mathcal{C}^2 , its second-divided-difference scheme, i.e., the subdivision scheme corresponding to the hat prediction process, cannot converge uniformly. This explains the ‘rugged’ behavior observed in the numerical display.

The situation for $r = 5$ is more favorable. Deslauriers and Dubuc [10] prove that $\tilde{\varphi}(x)$ is \mathcal{C}^2 , thus the subdivision scheme corresponding to the hat prediction process converges uniformly, i.e., $\varphi(x)$ must be continuous. As a consequence of Theorem 8.5, the hat multiresolution transform and its inverse are stable with respect to perturbations. The same holds for $r = 7$. The numerical limits appear to be continuous functions in both cases (see Figure 3).

The limiting functions displayed in Figure 3 can, in fact, be found in [8, section 8.3]. In the biorthogonal framework, one uses two scaling functions to construct the multiresolution scheme. If one of the scaling functions is the hat function, the restrictions imposed on the biorthogonal framework lead to, precisely, the dilation relations (122). There is one free parameter left which is used to decide the order of the scheme, which can be measured by the space of polynomials which are reproduced exactly.

Harten proves in [17] that the biorthogonal framework can be seen as the hierarchical form of the corresponding piecewise-polynomial reconstructions, under periodicity assumptions. From this point of view, our construction could be seen as an independent derivation of these multiresolution schemes, in which the basic building blocks are not related to functional analysis, but to approximation theory. This point of view makes it conceptually simpler to adjust to the boundaries.

The nonperiodic case. In the nonperiodic case, algorithms (117) and (118) above need to be modified to account for the boundaries. The first i -loop in (117) and the second i -loop in (118) run only from $i = 1$ to $N_k = J_k - 1$. A moment’s reflection shows that, for $p = 2s - 2$, there are s intervals at each boundary that require one-sided reconstruction procedures. This implies that, at each resolution level, the scale coefficients d_j^k , $1 \leq j \leq s - 1$, $J_{k-1} - s + 2 \leq j \leq J_{k-1}$ are to be computed in a special manner. The necessary modifications for $p = 2, 4$, and 6 at the left boundary are as

FIG. 4. *Hat-average limiting functions.* $J_0 = 8$, $p = 2$.FIG. 5. *Hat-average limiting functions.* $J_0 = 8$, $p = 4$.

follows:

$$\begin{aligned}
 p = 2 \quad & \left\{ \begin{aligned} d_1^k &= \bar{f}_1^k - \left(\frac{3}{2} \bar{f}_1^{k-1} - \frac{1}{2} \bar{f}_2^{k-1} \right), \end{aligned} \right. \\
 p = 4 \quad & \left\{ \begin{aligned} d_1^k &= \bar{f}_1^k - \left(\frac{65}{32} \bar{f}_1^{k-1} - \frac{57}{32} \bar{f}_2^{k-1} + \frac{31}{32} \bar{f}_3^{k-1} - \frac{7}{32} \bar{f}_4^{k-1} \right), \\ d_2^k &= \bar{f}_3^k - \left(\frac{7}{32} \bar{f}_1^{k-1} + \frac{37}{32} \bar{f}_2^{k-1} - \frac{15}{32} \bar{f}_3^{k-1} + \frac{3}{32} \bar{f}_4^{k-1} \right), \end{aligned} \right. \\
 p = 6 \quad & \left\{ \begin{aligned} d_1^k &= \bar{f}_1^k - \left(\frac{595}{256} \bar{f}_1^{k-1} - \frac{789}{256} \bar{f}_2^{k-1} + \frac{830}{256} \bar{f}_3^{k-1} - \frac{554}{256} \bar{f}_4^{k-1} + \frac{207}{256} \bar{f}_5^{k-1} - \frac{33}{256} \bar{f}_6^{k-1} \right), \\ d_2^k &= \bar{f}_3^k - \left(\frac{33}{256} \bar{f}_1^{k-1} + \frac{397}{256} \bar{f}_2^{k-1} - \frac{294}{256} \bar{f}_3^{k-1} + \frac{170}{256} \bar{f}_4^{k-1} - \frac{59}{256} \bar{f}_5^{k-1} + \frac{9}{256} \bar{f}_6^{k-1} \right), \\ d_3^k &= \bar{f}_5^k - \left(\frac{-9}{256} \bar{f}_1^{k-1} + \frac{87}{256} \bar{f}_2^{k-1} + \frac{262}{256} \bar{f}_3^{k-1} - \frac{114}{256} \bar{f}_4^{k-1} + \frac{35}{256} \bar{f}_5^{k-1} - \frac{5}{256} \bar{f}_6^{k-1} \right), \end{aligned} \right.
 \end{aligned}$$

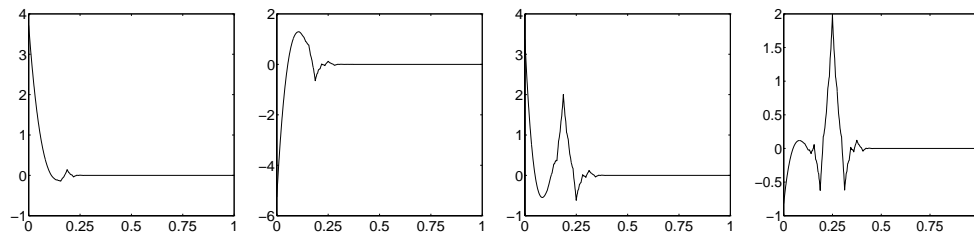
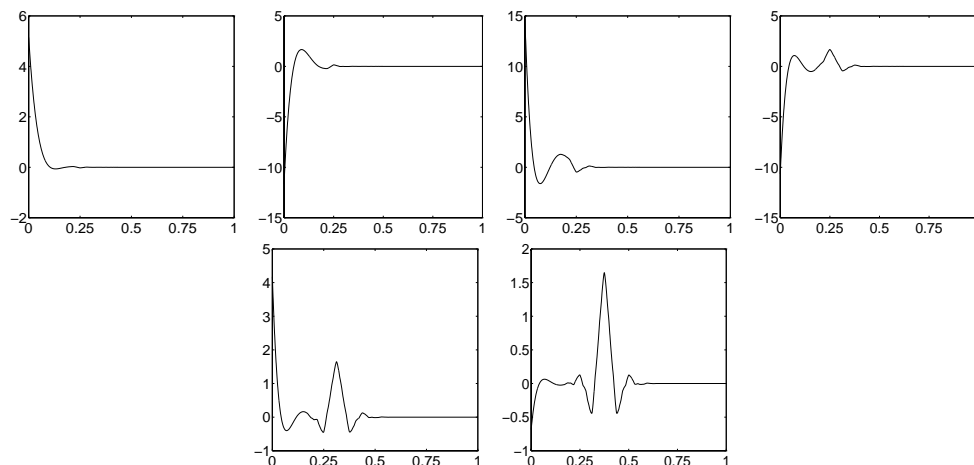
Modifications at the right boundary can be obtained by symmetry.

It is simple to study numerically the convergence properties of the sequence $\{A_k^L \delta_i^k\}_{L=k}^\infty$.

The case $p = 2$ corresponds to $r = 3$, i.e., \mathcal{I}_k uses third-degree polynomial pieces. Figure 4 shows $M^{-1}(\delta_i^0, 0, \dots, 0)$, $i = 1, 2$, with $J_0 = 8$, $L = 7$. Since only the periodic interpolation is involved in their computation, the limiting functions for $u^0 = \delta_i^0$, $i = 3, 4, 5$ are scaled translations of the recursive subdivision limit corresponding to the periodic case (they are not shown). The limiting functions for $i = 6, 7$ are the symmetric reflections of the $i = 2, 1$ limits with respect to the right boundary. The numerical results indicate that the boundary limiting functions exist and appear to have the same regularity properties as the periodic limit.

The case $p = 4$ corresponds to $r = 5$. Figure 5 displays $M^{-1}(\delta_i^0, 0, \dots, 0)$, $i = 1, 2, 3, 4$ with $J_0 = 8$, $L = 7$. All limiting functions are affected by boundary effects. When $J_0 \geq 16$, we have four limiting functions at each boundary altered by boundary effects; the remaining limiting functions are, of course, translates of the periodic limit. Numerical results for $J_0 = 16$, $L = 7$, and $p = 4$ are displayed in Figure 6.

The limiting functions obtained starting from higher resolution levels are scaled versions of the ones obtained starting the limiting process at a lower resolution level

FIG. 6. *Hat-average limiting functions.* $J_0 = 16$, $p = 4$.FIG. 7. *Hat-average limiting functions.* $J_0 = 16$, $p = 6$.

(see Figures 5 and 6). It can be easily observed that for a reconstruction of order p , there are p limiting functions affected by boundary effects. These functions have compact support. In fact, $\text{support } \varphi_i^k = [0, (p+i)h_k]$ for $i = 1, \dots, p$.

The case $p = 6$ ($r = 7$) is displayed in Figure 7, where $M^{-1}(\delta_i^0, 0, \dots, 0)$, $i = 1, 2, \dots, 6$, are shown.

As in the interpolatory case, the limiting function for the periodic case seems to be an important ingredient of these boundary limits, and their degree of smoothness seems to be directly related to that of the periodic limit. At the present moment we cannot prove analytically their existence. We think that an argument of the type used in the interpolatory setting might lead to an analytic proof, which would in turn prove analytically the stability of the multiresolution schemes.

9. Conclusions. We consider the multiresolution setting derived from discretizing by local averages with respect to the hat function and define a reconstruction technique which enables us to construct multiresolution schemes that are adequate for this multiresolution setting. It is observed that these multiresolution schemes are appropriate to obtain multiscale decompositions of piecewise smooth functions with a finite number of δ -type singularities.

In this paper we consider only linear reconstruction techniques. Even in this simple setting, the development of multiresolution schemes based on discretization and reconstruction, the two basic building blocks of Harten's framework, leads to a set-up in which multiscale decompositions are easily understood in terms of approximation

theory. Troublesome questions in wavelet theory, like boundary handling, admit also a relatively simple approach, once rephrased as approximation problems.

We obtain multiresolution schemes for functions defined in a bounded interval. Their stability properties are analyzed using the general framework developed in [16, 17] as well as the connection between the hat-weighted and interpolatory multiresolution settings. The link with the theory for stationary subdivision is exploited to show that, under periodicity assumptions, we recover several well-known multiresolution schemes within the biorthogonal framework of [7].

Harten's framework also allows us to consider nonlinear reconstruction operators. We consider the multiresolution schemes derived from these in Part II [1].

Acknowledgments. The first two authors would like to thank S. Osher and E. Tadmor for their comments and their support. We especially thank V. Candela and the referees of the paper for their constructive criticism on the first version of this paper.

REFERENCES

- [1] F. ARÀNDIGA, R. DONAT, AND A. HARTEN, *Multiresolution based on weighted averages of the hat function II: Nonlinear reconstruction operators*, SIAM J. Sci. Comput., to appear.
- [2] F. ARÀNDIGA, V. CANDELA, AND R. DONAT, *Fast multiresolution algorithms for solving linear equations: A comparative study*, SIAM J. Sci. Comput., 16 (1995), pp. 581–600.
- [3] E. BACRY, S. MALLAT, AND G. PAPANICOLAOU, *A wavelet based space-time adaptive numerical method for partial differential equations*, Math. Model. Numer. Anal., 26 (1992), pp. 703–834.
- [4] G. BEYLKIN, R. COIFMAN, AND V. ROKHLIN, *Fast wavelet transform and numerical algorithms I*, Comm. Pure Appl. Math., 44 (1991), pp. 141–183.
- [5] A. CAVARETTA, W. DAHMEN, AND C. MICCHELLI, *Stationary subdivision*, AMS Memoirs 453, AMS, Providence, RI, 1991.
- [6] A. COHEN, *private communication*.
- [7] A. COHEN, I. DAUBECHIES, AND J.C. FEAUVEAU, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.
- [8] I. DAUBECHIES, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in App. Math. 61, SIAM, Philadelphia, 1992.
- [9] I. DAUBECHIES AND J. LAGARIAS, *Two scale difference equations II*, SIAM J. Math. Anal., 23 (1992), pp. 1031–1079.
- [10] G. DESLAURIERS AND S. DUBUC, *Symmetric iterative interpolation scheme*, Constr. Approx., 5 (1989), pp. 49–68.
- [11] R. DONAT AND A. HARTEN, *Data Compression Algorithms for Locally Oscillatory Data*, CAM Report 93-26, UCLA, 1993.
- [12] N. DYN, *Subdivision schemes in computer-aided geometric design*, in Advances in Numerical Analysis II, Wavelets, Subdivision Algorithms and Radial Basis Functions, W. A. Light, ed., Clarendon Press, Oxford, 1992, pp. 36–104.
- [13] N. DYN, J.A. GREGORY, AND D. LEVIN, *Analysis of linear binary subdivision schemes for curve design*, Constr. Approx., 7 (1991), pp. 127–147.
- [14] B. ENGQUIST, S. OSHER, AND S. ZHONG, *Fast wavelet algorithms for linear evolution equations*, SIAM J. Sci. Comput., 15 (1994), pp. 755–775.
- [15] A. HARTEN, *Discrete multiresolution analysis and generalized wavelets*, J. Appl. Numer. Math., 12 (1993), pp. 153–192.
- [16] A. HARTEN, *Multiresolution Representation of Data*, CAM Report 93-13, UCLA, 1993.
- [17] A. HARTEN, *Multiresolution representation of data II: General framework*, SIAM J. Numer. Anal., 33 (1996), pp. 1205–1256.
- [18] A. HARTEN, *Multiresolution Representation of Cell-Averaged Data*, CAM Report 94-21, UCLA, 1994.
- [19] A. HARTEN, *Multiresolution algorithms for the numerical solution of hyperbolic conservation laws*, Comm. Pure Appl. Math., 48 (1995), pp. 1305–1342.
- [20] A. HARTEN, *Adaptive multiresolution schemes for shock computations*, J. Comput. Phys., 115 (1994), pp. 319–338.

- [21] A. HARTEN AND I. YAD-SHALOM, *Fast multiresolution algorithms for matrix-vector multiplication*, SIAM J. Numer. Anal., 31 (1994), pp. 1191–1218.
- [22] H. YSERENTANT, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.