UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Multiresolution Based on Weighted Averages of the Hat Function I: Linear Reconstruction Techniques

Francesc Arandiga
Rosa Donat
Ami Harten

August 1996

CAM Report 96-25

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

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

FRANCESC ARÀNDIGA*, ROSA DONAT† AND AMI HARTEN ‡

Abstract. We study the properties of the multiresolution analysis corresponding to discretization by 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. Only data-independent (i.e. linear) reconstruction operators are considered in Part I. Linear reconstruction techniques allow us, under certain circumstances, to construct a basis of generalized wavelets for the multiresolution representation of the original data. The stability of the associated multiresolution schemes is analyzed using the general framework developed by A. Harten in [18] and the connection with the theory of recursive subdivision.

Key Words. Multi-scale decomposition, discretization, reconstruction.

AMS(MOS) subject classifications. 41A05, 41A15, 65015

* Departament de Matemàtica Aplicada. Universitat de València (Spain). Research supported in part by by DGICYT PB94-0987, in part by a University of Valencia Grant and in part by ONR-N00014-95-1-0272 e-mail:arandiga@godella.matapl.uv.es

† Departament de Matemàtica Aplicada. Universitat de València (Spain). Research 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 e-mail:donat@godella.matapl.uv.es

‡ School of Mathematical Sciences Tel-Aviv University, Tel Aviv, Israel and Department of Methematics, UCLA, Los Angeles CA 90024-1555 until his death August 5, 1994

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. [9] and references therein) has been a giant leap towards local scale decompositions and has already had great impact on several fields of science.

In Numerical Analysis, a wavelet type decomposition of a function is used to reduce the cost of many numerical algorithms by either applying it to the numerical solution operator to obtain an approximate sparse form [6, 15, 21, 5] or by applying it to the numerical solution itself to obtain an approximate reduced representation in order to solve for less quantities [20].

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. Adaptativity is only possible 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 CAD (Computer-Aided Design) and, of course, the theory of wavelets, led A. Harten to the development of a "General Framework" for multiresolution representation of discrete data.

Harten's General Framework is built around two basic operators, decimation and prediction, which connect adjacent resolution levels. In turn, these operators are constructed from two basic building blocks: the discretization and reconstruction operators. The former obtains discrete information from a given (continuous) signal, and the latter produces an "approximation" to that signal, from the discrete values, in the same function space to which the original signal belongs.

A "new scale" is defined as the information on a given resolution level which cannot be predicted from discrete information at lower levels. If the discretization and reconstruction are local operators, the concept of "new scale" is also local. The scale coefficients are directly related to the prediction errors, and thus to the reconstruction procedure. If a scale coefficient is small at a certain location on a given scale, it means that the reconstruction procedure on that scale gives a proper approximation of the original signal at that particular location.

Under these premises, 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.

In a way, the discretization process specifies the nature of the discrete data to be analyzed, i.e. how it was generated. In [17] Harten introduces the concept of *nested sequence of discretization*, and shows that when we consider a nested sequence of discretization operators associated to increasing resolution levels, non-redundant multi-scale decompositions are always possible.

When reinterpreted within Harten's framework, the discretization operator in the wavelet theory is obtained by taking local averages against the scaling function. The dilation relation satisfied by the scaling function becomes a particular way of getting a nested structure in the discretization sequence.

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, well 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. It is shown in [16, 17, 18] how to obtain stable multi-scale decompositions using the δ-function in the discretization process.

The flexibility of Harten's general framework also allows for an easier way of handling boundaries. In [19], multiresolution settings obtained using the box function in the discretization process are considered. Under periodicity assumptions, the decimation and prediction operators are the same as those obtained in the biorthogonal framework, when the box function is chosen as the basic scaling function. However, within Harten's framework the bounded domain case (no periodicity) can be handled by reducing it again to an approximation problem, which is easy to solve with well known numerical techniques. We refer the reader to [18, 19] for a detailed account of the so-called "cell-average" framework.

It is our opinion that the ideas which form the basic core of the framework developed in [16, 17, 18] are general enough to enable an embedding of most numerical problems in a multiresolution setting. The general framework can be used to obtain multiresolution representations from discretizations corresponding to unstructured grids in several space dimensions [18, 19, 1]. Moreover, the notion of discretization is defined as a mapping from a "continuum" space $\mathcal{F}$ onto a "denumerable" V. Therefore $\mathcal{F}$ can be taken to be a family of operators, which would enable us to develop a multiresolution representation of operators, an approach that might be advantageous to the indirect derivation of [6] and [21] or [4].

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 [18] 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 consider only linear reconstruction techniques. We study the stability of the corresponding multiresolution decompositions and show how to obtain multiresolution representations of appropriate approximations to the original function. As in the cell-average formulation, periodicity assumptions lead to biorthogonal wavelets (with the hat functions as the scaling function used for discretization). We show how to obtain multiresolution decompositions in the case of a bounded domain and give numerical evidence of their stability.

In Part II [3], we consider nonlinear reconstruction techniques and thus introduce adaptativity into the hat-average multiresolution set-up.

The present paper is organized as follows: In Sections 2 and 3 we describe those aspects of the general framework which are relevant to our discussion. Most of the material in these sections is taken from [17, 18] but we include it here to make this work almost self-contained.

Section 4 describes the general properties of multiresolution settings associated to 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 4.1. The stability of the multiresolution schemes in the interpolatory setting is related to the theory of refinement subdivision in Computer Aided Design. We pay special attention to this connection, since it will be used later in the hat-average multiresolution setting.

Section 5 is the core of the paper. Here we describe the hat-average set-up and study its properties. Section 6 relates the uniform (i.e. periodic) case to the biorthogonal framework. Finally, some conclusions are drawn in Section 7.

2. The General Framework: A Quick Overview. Harten introduces its notion of multiresolution analysis in [16] and later generalizes it in [17, 18] where the theoretical foundation for multiresolution representation of data and operators is laid out. This general framework provides an appropriate set-up to study the stability of the corresponding multiresolution schemes as well as the functional structure associated to some of them.

This section is a brief overview of some of the results in [18]. Proofs are given only when they are simple, short and illustrative, for further details we refer the reader to [17, 18].

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.*

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

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

It follows directly from this definition that for any $v^{k-1} \in V^{k-1}$ there is at least one $u \in V^k$ such that $D_k^{k-1}u = v^{k-1}$.

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

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

The requirement in the definition above is nothing but a consistency relation: Predicted values at the k-th resolution level should contain the same discrete information as the original values, when restrited to the $k-1$st level. Note that $P_{k-1}^k$is not required to be a linear operator.

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:

$v^L \quad \to \quad Mv^L \qquad$ (Encoding)

$$(1) \qquad \begin{cases} \text{Do} \quad k = L, \ldots, 1 \\ v^{k-1} = D_k^{k-1} v^k \\ d^k = G_k(v^k - P_{k-1}^k v^{k-1}) \end{cases}$$

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

$Mv^L \quad \to \quad M^{-1}Mv^L \qquad$ (Decoding)

$$(2) \qquad \begin{cases} \text{Do} \quad k = 1, \ldots, L \\ v^k = P_{k-1}^k v^{k-1} + E_k d^k \end{cases}$$

The *scale coefficients*, d^k, are obtained from the prediction errors

$$e^k = v^k - P_{k-1}^k v^{k-1}$$

by removing the redundant information in them. Notice that

$$D_k^{k-1} e^k = D_k^{k-1}(v^k - P_{k-1}^k v^{k-1}) = 0,$$

in other words, e^k belongs to the null space of the decimation operator

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

If $\dim V^k = N_k$, then $\dim \mathcal{N}(D_k^{k-1}) = N_k - N_{k-1}$. Hence, if we select a set of basis functions in $\mathcal{N}(D_k^{k-1})$

$$\mathcal{N}(D_k^{k-1}) = \text{span}\{\mu_j^k\}_{j=1}^{N_k - N_{k-1}},$$

the prediction error e^k, which belongs to V^k and so is described in terms of N_k components, can be represented in terms of its $N_k - N_{k-1}$ coordinates in the base $\{\mu_j^k\}$:

At this point it is easy to prove that there is a one to one correspondence between v^k and $\{d^k, v^{k-1}\}$: Given v^k we evaluate

$$\text{(3)} \qquad \begin{cases} v^{k-1} &= D_k^{k-1} v^k \\ d^k &= G_k(I_k - P_{k-1}^k D_k^{k-1}) v^k. \end{cases}$$

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

$$\text{(4)} \qquad \begin{aligned} P_{k-1}^k v^{k-1} + E_k d^k &= P_{k-1}^k D_k^{k-1} v^k + E_k G_k (I_k - P_{k-1}^k D_k^{k-1}) v^k \\ &= P_{k-1}^k D_k^{k-1} v^k + (I_k - P_{k-1}^k D_k^{k-1}) v^k \\ &= v^k. \end{aligned}$$

This shows that

$$v^L \overset{1:1}{\leftrightarrow} \{v^0, d^1, \ldots, d^L\} = M v^L.$$

We refer to $M(v^L)$ as the multiresolution representation of v^L, and to algorithms (1) and (2) as the direct and inverse multiresolution transforms, respectively.

REMARK 2.1. *There is a one-to-one correspondence between an element of a linear vector space and its components in a given basis, i.e.* $v^L \overset{1:1}{\leftrightarrow} \bar{v}^L$. *Algorithms (1) and (2) can be used for* $\hat{M}(s^L)$, *the associated multiresolution representation of sequences* $s^L \in S^L$ *by defining*

$$\text{(5)} \qquad s^L \overset{1:1}{\leftrightarrow} \hat{M}(s^L) := M(v^L), \qquad v^L = \sum_i s_i^L \eta_i^L.$$

In the finite dimensional case, the number of components in d^k *is* $(N_k - N_{k-1})$ *and, consequently, the number of components in* $M(s^L)$ *is*

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

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

The scale coefficients are directly related to the prediction errors: if a scale coefficient is small at a certain location on a given scale, it means that s^k *is properly represented by* $\hat{P}_{k-1}^k s^{k-1}$ *on that particular scale at that location. For purposes of data compression, the small scale coefficients can then be quantized or truncated, reducing the dimensionality of the discrete representation without a significant alteration of the information contents of the original sequence.*

It is clear then that one of the main concerns should be the 'quality' of the prediction $\hat{P}^k_{k-1}$. The notion of 'k-th scale' is related to the information in s^k which cannot be predicted from knowledge of s^{k-1} by any prediction scheme. When using a particular one, 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.

A given multiresolution scheme applies to any *sequence $s = s^L$ of real numbers. These numbers could have been generated by some stochastic process, or by some 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 possible to give a precise meaning to the question of quality. It can only be made meaningful by restricting our attention to a subset of data for which we know something about the way it was generated.*

The main question becomes then that of the design of a stable multiresolution scheme that can be applied to any sequence in S^L, but that is particularly adequate for a given application.

REMARK 2.2. *Observe that algorithms (1) and (2) 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.*

The point of view in [16, 17, 18] is that the primary choice to be made in the design of a multiresolution scheme is that of the decimation operator. Following [18] we now introduce the concept of *nested discretization.* Here the 'discretization' specifies the nature of the data to be analyzed (i.e. how it was generated). The 'nestedness' concept introduces a sense of hierarchy among the levels of resolution that allows for the construction of a sequence of decimation 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*

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

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

$$\mathcal{D}_k : \mathcal{F} \to V^k, \qquad \mathcal{D}_k(\mathcal{F}) = V^k = span\{\eta_i^k\}.$$

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

$$\mathcal{D}_k f = 0 \Rightarrow \mathcal{D}_{k-1} f = 0 \tag{6}$$

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. *(Harten-Lax) If $\{\mathcal{D}_k\}$ is a nested sequence of discretization, then the mapping from V^k to V^{k-1} defined as follows:*

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

$$D_k^{k-1} \mathcal{D}_k = \mathcal{D}_{k-1} \tag{7}$$

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 decimation operators fix the multiresolution setting, once this selection is made we still have two more *independent* choices to make:

1. A prediction operator P_{k-1}^kwhich is a right inverse of D_k^{k-1}; this leads to the operator Q_k.
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.

The prediction operators can then be constructed using a sequence of appropriate 'reconstruction operators'.

DEFINITION 2.5. *We say that* $\mathcal{R}$

$$\mathcal{R} : V \to \mathcal{F}, \qquad V = \mathcal{D}(\mathcal{F})$$

is a reconstruction operator in V*, if it is a right inverse of* $\mathcal{D}$*, i.e.* $\mathcal{D}\mathcal{R} = I_V$ *where* I_V *is the identity operator in* 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

$$P_{k-1}^k := \mathcal{D}_k \mathcal{R}_{k-1} : V^{k-1} \to V^k. \tag{8}$$

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

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

The analysis above shows that finding a suitable prediction for a multiresolution setting, and thus a suitable multiresolution scheme for a given application, can now be formulated as a typical problem in approximation theory:

Knowing $\mathcal{D}_{k-1}f$*,* $f \in \mathcal{F}$*, find a 'good approximation' for* $\mathcal{D}_k f$*.*

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 to 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 it is exact.

A *good* solution to our approximation problem will bring us one step closer to our stated goal: The design a multiresolution scheme that applies to all sequences $s \in S^L$ but is particularly adequate for those sequences $\hat{v}^L \in S^L$ which are obtained by the discretization process defined by the operators $\mathcal{D}_k$.

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

$$d^k = G_k(v^k - P^k_{k-1}v^{k-1})$$

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.

For linear reconstruction operators, the question of stability admits a relatively simple approach in this general framework. We review this subject in the following section.

3. Stability Analysis for Linear Reconstruction Operators. The decimation operators D^{k-1}_k are always linear. When the reconstruction operators, $\mathcal{R}_k$, are linear functionals, the prediction operators P^k_{k-1} are linear too. In this case, the multiresolution transform becomes a linear operator describing a change of basis vectors in $\mathcal{D}_L(\mathcal{F})$.

To see this, let us introduce the linear operator B^k_L of successive decimation

$$B^k_L = D^k_{k+1} \cdots D^{L-1}_L : V^L \to V^k \tag{9}$$

and observe that v^k in (1) can be written as

$$v^k = B^k_L v^L. \tag{10}$$

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

$$v^0 = B^0_L v^L, \qquad d^k = G_k Q_k B^k_L v^L, \qquad 1 \le k \le L. \tag{11}$$

Likewise, let us introduce the operator A^L_k of successive prediction as

$$A^L_k = P^L_{L-1} \cdots P^{k+1}_k : V^k \to V^L. \tag{12}$$

When $\mathcal{R}_k$ is linear $\forall k$, these operators are also linear. This fact allows us to express the inverse multiresolution transform (2) directly in terms of $M(v^L)$ as follows:

$$v^L = A^L_0 v^0 + \sum_{k=1}^{L} A^L_k E_k d^k. \tag{13}$$

With the following definitions:

$$\bar{\varphi}^{k,L}_i = A^L_k \eta^k_i, \quad 0 \le k \le L \qquad \bar{\psi}^{k,L}_j = A^L_k \mu^k_j, \quad 1 \le k \le L \tag{14}$$

and using the relations $v^k = \sum_i \hat{v}^k \eta^k_i$ and $E_k d^k = \sum_j d^k_j \mu^k_j$, (13) can be expressed as

$$v^L = \sum_i \hat{v}^0_i \bar{\varphi}^{0,L}_i + \sum_{k=1}^{L} \sum_j d^k_j \bar{\psi}^{k,L}_j. \tag{15}$$

Thus, we have the following

THEOREM 3.1. *Let* $(\{V^k\}^L_{k=0}, \{D^{k-1}_k\}^L_{k=1})$ *be a multiresolution setting and let* $\{P^k_{k-1}\}$ *be any sequence of linear prediction operators, then*

$$\bar{\mathcal{B}}_M = \left(\{\bar{\varphi}^{0,L}_i\}_i, \{\{\bar{\psi}^{k,L}_j\}_j\}^L_{k=1}\right) \tag{16}$$

is a basis of V^L and any $v^L \in V^L$ has the representation (15) where the coordinates are given by the direct multiresolution transform $M(v^L)$.

We refer to $\bar{\mathcal{B}}_M$ as a multiresolution basis of V^L.

Within this framework, it is easy to see that, for linear prediction operators, the successive decimation and prediction operators B_L^k and A_k^L do control the stability of the direct and inverse multiresolution transforms with respect to perturbations.

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

$$\delta(d^k) = d_\epsilon^k - d^k = G_k Q_k B_L^k (v_\epsilon^L - v^L). \tag{17}$$

This relation shows that the perturbation in the input is subject to successive decimation D_{m-1}^m for $m = L, \ldots, k+1$, projected into $\mathcal{N}(D_k^{k-1})$ and represented in some basis there. Clearly the 'dangerous' process that has to be controlled is that of successive decimation; the choice of basis in $\mathcal{N}(D_k^{k-1})$ is not that important: the basis need not be 'orthogonal' but it should not be 'too distorted' either.

Similarly, for purposes of data compression if the scale coefficients $\{d^k\}$ are replaced by $\{d_\epsilon^k\}$ which are obtained either by quantization or 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

$$\delta(v^L) = v_\epsilon^L - v^L = \sum_{k=1}^{L} A_k^L E_k (d_\epsilon^k - d^k), \tag{18}$$

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 $m = k+1, \ldots L$. The danger here is that the perturbation could be amplified by the process of successive prediction.

In the general framework we are considering, the decimation and prediction operators are obtained from a sequence of discretization $\{\mathcal{D}_k\}$ and corresponding reconstruction $\{\mathcal{R}_k\}$. The properties of the sequence $\{\mathcal{R}_k\mathcal{D}_k\}$ play a fundamental role in the stability of the direct and inverse multiresolution transforms.

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

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

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

□

Lemma 3.1 implies that

$$B_L^k \mathcal{D}_L = D_{k+1}^k \cdots D_L^{L-1} \mathcal{D}_L = \mathcal{D}_k \mathcal{R}_{k+1} \mathcal{D}_{k+1} \cdots \mathcal{D}_{L-1} \mathcal{R}_L \mathcal{D}_L = \mathcal{D}_k. \tag{19}$$

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.

Stability of the inverse multiresolution transform is usually more involved. 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 for all k*

$$(\mathcal{R}_k\mathcal{D}_k)\mathcal{R}_{k-1} = \mathcal{R}_{k-1} \qquad \equiv \qquad \mathcal{R}_k P_{k-1}^k = \mathcal{R}_{k-1} \tag{20}$$

Note that for a hierarchical sequence

$$\mathcal{R}_L A_k^L = \mathcal{R}_L \mathcal{D}_L \mathcal{R}_{L-1} \cdots \mathcal{D}_{k+1}\mathcal{R}_k = \mathcal{R}_k. \tag{21}$$

The hierarchical structure in the sequence of reconstruction operators 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 (21) is the equivalent to (19) 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. Therefore, the successive prediction step does not introduce spurious information or amplify existing noise, and consequently, the corresponding inverse multiresolution transform is a stable algorithm.

When $\mathcal{F}$ is a Banach space, the norm in $\mathcal{F}$ can be used to obtain working stability bounds for the direct and inverse multiresolution transforms, but the essential relations that lead to those stability bounds are (19) and (21). We refer the reader to [18] for further details.

Hierarchical sequences have an associated wavelet-like functional structure. First, notice that we can obtain a relation analogous to (15) in the space of functions $\mathcal{F}$. For this, let us consider the following definitions

$$\varphi_i^{k,L} = \mathcal{R}_L \bar{\varphi}_i^{k,L} = \mathcal{R}_L A_k^L \eta_i^k \tag{22}$$

$$\psi_j^{k,L} = \mathcal{R}_L \bar{\psi}_j^{k,L} = \mathcal{R}_L A_k^L \mu_j^k \tag{23}$$

Now, applying $\mathcal{R}_L$ to (15) we have

$$\mathcal{R}_L\mathcal{D}_L f = \sum_i (\mathcal{D}_0 f)_i \varphi_i^{0,L} + \sum_{k=1}^{L}\sum_j d_j^k \psi_j^{k,L}. \tag{24}$$

The functions $\{\varphi_i^{0,L}\}, \{\{\psi_j^{k,L}\}\}_{k=1}^L$ constitute, in fact, a basis for the space $\mathcal{R}_L\mathcal{D}_L\mathcal{F}$. These spaces can be thought of as subspaces of $\mathcal{F}$ with increasing approximation power (see [18]).Notice that the coefficients in expression (24) (which should be regarded as a formal derivation in the infinite-dimensional case) are computed by the direct multiresolution transform (1) of $\mathcal{D}_L f$.

For a hierarchical sequence of discretization

$$\varphi_i^{k,L} = \mathcal{R}_L A_k^L \eta_i^k = \mathcal{R}_k \eta_i^k =: \varphi_i^k, \qquad \psi_j^{k,L} = \mathcal{R}_L A_k^L \mu_j^k = \mathcal{R}_k \mu_j^k =: \psi_j^k \qquad \forall L.$$

Thus, the same functions, $\{\varphi_i^0\}, \{\{\psi_j^k\}\}$ can be used in each of the approximating spaces $\mathcal{R}_L\mathcal{D}_L\mathcal{F}$.

The functions $\{\{\varphi_i^k\}_i\}_k$ are linked by a two-level relationship. To see this, 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

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

Thus,

$$\varphi_i^{k-1} = \sum_l (\hat{P}_{k-1}^k)_{li} \varphi_l^k. \tag{25}$$

Likewise, let $\hat{E}_k$ be the matrix representation of the operator E_k in the basis $\{\eta_i^k\}$ and $\{\mu_j^k\}$ then

$$\psi_j^k = \sum_l (\hat{E}_k)_{lj} \varphi_l^k. \tag{26}$$

Let us define

$$\Phi^k := \text{span}\{\varphi_i^k, \}, \qquad \Psi^k := \text{span}\{\psi_j^k\},$$

then (25) implies that $\Phi^{k-1} \subset \Phi^k$, while (26) implies that $\Psi^k \subset \Phi^k$. Since

$$\begin{aligned} \mathcal{R}_k \mathcal{D}_k f &= \mathcal{R}_k \sum_i (\mathcal{D}_k f)_i \eta_i^k = \sum_i (\mathcal{D}_k f)_i \varphi_i^k \in \Phi^k, \quad \forall f \in \mathcal{F} \\ (\mathcal{R}_k \mathcal{D}_k)^2 &= \mathcal{R}_k (\mathcal{D}_k \mathcal{R}_k) \mathcal{D}_k = \mathcal{R}_k \mathcal{D}_k, \end{aligned}$$

$\mathcal{R}_k \mathcal{D}_k$ is a projection onto Φ^k. Moreover, for $v^k = \mathcal{D}_k f$, one step of the inverse multiresolution transform (2) can be written as

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

Thus applying $\mathcal{R}_k$ leaves us with

$$\mathcal{R}_k \mathcal{D}_k f = \mathcal{R}_k \mathcal{D}_k \mathcal{R}_{k-1} \mathcal{D}_{k-1} f + \sum_j d_j^k \mathcal{R}_k \mu_j^k = \mathcal{R}_{k-1} \mathcal{D}_{k-1} f + \sum_j d_j^k \psi_j^k. \tag{27}$$

It is not hard to prove that $\Phi^{k-1} \cap \Psi^k = \emptyset$, thus we can write

$$\Phi^k = \Phi^{k-1} \oplus \Psi^k, \tag{28}$$

which implies

$$\Phi^L = \Psi^L \oplus \cdots \Psi^1 \oplus \Phi^0. \tag{29}$$

This direct sum decomposition is not, in general, an orthogonal decomposition.

Relation (27) tells us that the d_j^k represent indeed the difference in information between two functional approximations to the original signal f at consecutive resolution levels. How well these scale coefficients represent a 'true' new scale depends, to a fairly good degree, on the 'accuracy' of the underlying reconstruction operator.

Orthogonal and biorthogonal wavelet algorithms can be seen as particular examples of this general framework. The reconstruction operators used in these algorithms are hierarchical; as a consequence, the associated compression algorithms are stable.

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.

In the linear case, reconstruction sequences which are based on spectral expansions or splines are also hierarchical (see [13, 17, 18]), thus, the additional functional structure and stability properties just described also apply to the multiresolution schemes they define.

Nonlinear reconstruction techniques can be used to optimize compression rates. In this case, stability is ensured by a modified encoding-decoding procedure. We refer the reader to [16, 19] and to Part II [3] for descriptions of non-linear reconstruction procedures in various contexts, as well as stability considerations.

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 [17, 18] or next section). Checking stability is not an easy task in this case, with the definitions we have covered so far. However, in many cases a sequence of approximation that is not hierarchical to begin with, has a hierarchical form which is obtained by considering a limiting process akin to refinement in subdivision schemes [7, 14]. The hierarchical form has the same scale coefficients as the original one, thus the stability properties derived from the hierarchical structure of the new reconstruction sequence are also inherited by the original (usually more manageable) one.

The main theoretical results are the following (we refer to [18] for proofs):

THEOREM 3.2. *Let* $\{\mathcal{D}_k\}_{k=0}^{\infty}$ *a nested sequence of discretization operators and* $\{\mathcal{R}_k\}_{k=0}^{\infty}$ *be a sequence of reconstruction operators satisfying* $\mathcal{D}_k\mathcal{R}_k = I_{V_k}$ *and such that for any* $k \geq 0$ *and any* $f \in \mathcal{F}$ *the following limit exists*

$$\lim_{L\to\infty} \Pi_k^L f =: f_k^\infty \in \mathcal{F}. \tag{30}$$

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

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

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 3.2 shows that f_k^∞ has exactly the same discrete information contents as the initial data $\mathcal{R}_k\mathcal{D}_k f$. The limiting process (30) which assigns f_k^∞ to $\mathcal{R}_k\mathcal{D}_k f$ is called in [18] "cosmetic refinement", in order to stress that unlike other refinement processes in numerical analysis, there is no addition of (discrete) information.

THEOREM 3.3. *Let* $\{\mathcal{R}_k\mathcal{D}_k\}$ *be as in Theorem 3.2 and define*

$$\mathcal{R}_k^H : V^k \to \mathcal{F} \qquad\qquad \mathcal{R}_k^H v^k = \lim_{L\to\infty} \Pi_{k+1}^L \mathcal{R}_k v^k.$$

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.* $(\mathcal{R}_{k+1}^H\mathcal{D}_{k+1})\mathcal{R}_k^H = \mathcal{R}_k^H$.

As a consequence of the above theorem, the multiresolution scheme associated to the hierarchical form is *the same* as that of the original sequence. Since the hierarchical form leads naturally to a stable multiresolution transform, stability of the original scheme is a consequence of the *existence* of the hierarchical reconstruction. For all practical purposes it is not important to know the explicit expression of the hierarchical form, however knowledge of its *existence* is important because it implies stability of the original multiresolution scheme.

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

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

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

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

which shows that the hierarchical form has the same 'accuracy' as the original sequence.

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 $\mathcal{R}_k^H v^k$ for all $v^k \in V^k$. Note that

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

since the sum is finite and the reconstruction operators are linear. Hence,

$$\exists \lim_{L\to\infty} \varphi_i^{k,L} = \varphi_i^k \in \mathcal{F} \qquad \equiv \qquad \exists \mathcal{R}_k^H v^k \quad \forall v^k \in V^k.$$

Thus, the existence of the limit functions becomes a test for the stability of the multiresolution scheme derived from a particular sequence of discretization.

Notice that applying $\mathcal{R}_L^H$ to (15) we obtain the following relation (instead of (24))

$$\mathcal{R}_L^H \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^0 + \sum_{k=1}^{L} \sum_j d_j^k(f) \psi_j^k$$

where

$$\varphi_i^k = \mathcal{R}_k^H \eta_i^k = \lim_{L\to\infty} \Pi_{k+1}^L \mathcal{R}_k \eta_i^k, \qquad \psi_j^k = \mathcal{R}_k^H \mu_j^k = \lim_{L\to\infty} \Pi_{k+1}^L \mathcal{R}_k \mu_j^k,$$

and $d_j^k(f)$ are the original scale coefficients $d_j^k(f) = G_k \mathcal{D}_k (I - \mathcal{R}_{k-1}\mathcal{D}_{k-1}) f$.

The existence of the limiting process does not guarantee, by itself, existence of a multiresolution set of basis functions in $\mathcal{F}$. A sufficient condition for this to happen is that $\{\mathcal{R}_k^H \mathcal{D}_k\}$ be a sequence of approximation. We refer the reader to [18] for details, as well as sufficient conditions to be imposed on the original sequence $\{\mathcal{R}_k, \mathcal{D}_k\}$ to guarantee the existence of a multiresolution basis on the space $\mathcal{F}$.

4. Discretization by local averages. Discretizing by local averages against an *appropriate* weight function is one of the most usual processes in numerical analysis. The weighted averages carry local information about the given function and very often the weight function is imposed by the underlying context. In this section we examine the multiresolution settings that correspond to these discretization processes (see also [16, 17, 18]).

Let us consider a sequence, finite or infinite, of equally spaced points on the real line:

$$X = \{x_i\}, \qquad x_i \in R \qquad h = x_i - x_{i-1}.$$

We take $\omega(x)$, the weight function, to be a function with compact support satisfying

$$\int \omega(x) dx = 1 \tag{31}$$

and define the discretization operator, $\mathcal{D}$, associated to the resolution level defined by the grid X as follows: for each f in an appropriate function space, $\mathcal{F}$,

$$(\mathcal{D}f)_i = \bar{f}_i = < f, \frac{1}{h}\omega(\frac{x - x_i}{h}) > = \frac{1}{h}\int f(x)\omega(\frac{x - x_i}{h})dx, \quad x_i \in X. \tag{32}$$

The operator $\mathcal{D}$ acts on a space of functions $\mathcal{F}$ for which the integral in (32) is well defined which, in a sense, specifies the nature of the data to be analyzed.

We now introduce a sense of hierarchy by constructing a nested sequence of discretization operators of the type defined by (32). To this end, we consider the set of nested dyadic grids (to which we associate the increasing levels of resolution) $\{X^k\}$, $k \geq 0$ of size $h_k = 2^{-k}h_0$:

$$X^k = \{x_j^k\} \qquad x_j^k = j \cdot h_k. \tag{33}$$

Notice that $x_{2j}^k = x_j^{k-1}$.

The sequence of discretization is defined as

$$\mathcal{D}_k : \mathcal{F} \to S^k, \qquad (\mathcal{D}_k f)_i = \bar{f}_i^k = < f, \frac{1}{h_k}\omega(\frac{x - x_i^k}{h_k}) > \equiv < f, \omega_i^k >, \tag{34}$$

where the space S^k is an appropriate space of sequences, and ω_i^k are scaled translates of $\omega(x)$,

$$\omega_i^k = \frac{1}{h_k}\omega(\frac{x - x_i^k}{h_k}).$$

The weighted averages (34) give information on the behavior of the function f at different resolution levels. When the weight function $w(x)$ satisfies a dilation relation such as

$$\omega(y) = 2\sum_l \alpha_l \omega(2y - l), \tag{35}$$

knowledge of the weighted averages at a certain level of resolution determines, without further reference to the function itself, the weighted averages of that function at all coarser levels of resolution. This is the basic trend in all multiresolution representations.

In fact, taking $y = (x - x_i^{k-1})/h_{k-1}$, we can rewrite (35) as

$$\omega_i^{k-1} = \sum_l \alpha_l \omega_{2i+l}^k = \sum_l \alpha_{l-2i}\omega_l^k. \tag{36}$$

In terms of the discretization operators, (36) can be expressed as follows

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

which immediately implies that the sequence $\{\mathcal{D}_k\}$ is nested. Thus, 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.

Formula (36) implies that the decimation operator, D_k^{k-1}, can be described by a matrix whose elements are

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

Observe that $D_k^{k-1}$is independent of the level of resolution. In what follows, we shall always assume that only a finite number of α's are non zero.

The theory of wavelets provides examples of weight functions that satisfy a dilation relation (called scaling functions in the wavelet world). If this dilation relation satisfies special properties then the associated multiresolution analysis has a special functional structure and one obtains orthonormal basis of $L^2(R)$ that come from considering dilates and translates of a single function: the wavelet.

Many of the functions $\omega(x)$ that are used in numerical analysis automatically satisfy a dilation equation. Some of the easiest ones are the following:

1. $\omega(x) = \delta(x)$, where δ is the Dirac distribution, satisfies

$$\omega(x) = 2\omega(2x) \Rightarrow \alpha_0 = 1; \tag{37}$$

2.

$$\text{The box function} \quad \omega(x) = \begin{cases} 1 & -1 \leq x < 0 \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

satisfies

$$\omega(x) = \omega(2x) + \omega(2x+1) \Rightarrow \alpha_0 = \alpha_{-1} = \frac{1}{2}; \tag{39}$$

3.

$$\text{The hat function} \quad \omega(x) = \begin{cases} 1+x & -1 \leq x \leq 0 \\ 1-x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

satisfies

$$\omega(x) = \frac{1}{2}[\omega(2x-1) + 2\omega(2x) + \omega(2x+1)] \Rightarrow \alpha_1 = \alpha_{-1} = \frac{1}{4}, \alpha_0 = \frac{1}{2}. \tag{41}$$

All these functions $\omega(x)$ form a hierarchy of functions $\omega^m(x)$ which is obtained by repeated convolution with a characteristic function (see [17, 18])

$$\omega^{m+1} = \omega^m * \chi_{[-1+s_m, s_m]}, \quad s_m = \frac{1}{2}[1-(-1)^m], \quad \omega^0 = \delta(x). \tag{42}$$

From these, only the box function leads to an orthonormal wavelet basis in $L^2(R)$, the Haar basis, the most elementary of the wavelet basis. However, they are the natural weight functions in many contexts and, as we shall see later, they also give rise to stable multiresolution algorithms, which can be used in the same manner as those derived from wavelets.

Once the weight function is fixed, the primary choice, that of the decimation operator, is already made. With respect to this, we shall see later (see also [17, 18]) that local averages with respect to Dirac's Delta function are appropriate to obtain multiscale decompositions of continuous functions, while the box and hat function lead to appropriate multiresolution representations of piecewise continuous functions with jump discontinuities and with δ-discontinuities respectively.

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.

In the dilation relation case, the null space of $D_k^{k-1}$is easily characterized, in fact

$$\mathcal{N}(D_k^{k-1}) = \{s^k \in S^k \mid D_k^{k-1}s^k = 0\} = \{s^k \in S^k \mid \sum_l \alpha_{l-2m} s_l^k = 0\}. \tag{43}$$

A choice of basis in the null space determines the final value of the k-th scale coefficients d^k. In the infinite dimensional case, the vectors μ_j^k defined as

$$(\mu_j^k)_l = (-1)^l \alpha_{2j-l+1}, \tag{44}$$

belong to $\mathcal{N}(D_k^{k-1})$ since

$$(D_k^{k-1}\mu_j^k)_i = \sum_l (D_k^{k-1})_{il}(\mu_j^k)_l = \sum_l \alpha_{l-2i} \cdot (-1)^l \alpha_{2j-l+1} = 0.$$

If we define E_k as the matrix whose columns are the vectors μ_j^k, then

$$(E_k^* E_k)_{ij} = (\mu_i^k)^* \mu_j^k = \sum_l \alpha_l \alpha_{2(j-i)+l}. \tag{45}$$

The right hand side of the expression above is directly related to the properties of the dilation relation. In fact, it is proven in [9] that when the dilation relation leads to an associated wavelet basis, the coefficients α_l must necessarily satisfy the relation

$$\sum_l \alpha_l \alpha_{2m+l} = \frac{1}{2}\delta_{0,m},$$

i.e. $(E_k^* E_k)$ is a multiple of the identity matrix. In most cases (and specifically in all examples we shall treat in this paper) $(E_k^* E_k)$ is an invertible matrix and, thus, the μ_j^k in (44) are a set of basis functions for $\mathcal{N}(D_k^{k-1})$. If

$$e^k \in \mathcal{N}(D_k^{k-1}) \qquad e^k = \sum_j d_j^k \mu_j^k = E_k d^k$$

then, multiplying the expression above by E_k^*

$$E_k^* e^k = E_k^* E_k d^k$$

we obtain

$$G_k = (E_k^* E_k)^{-1} E_k^*. \tag{46}$$

Other realizations of the operators G_k and E_k are also possible. Notice that (43) implies that the prediction errors always satisfy the following system of equations:

$$\sum_l \alpha_l e_{2m+l}^k = 0. \tag{47}$$

In the finite dimensional case, if $\dim S^k = N_k = 2N_{k-1}$, then $\dim \mathcal{N}(D_k^{k-1}) = N_k - N_{k-1} = N_{k-1}$, and (47) implies that only half of the prediction errors are independent. It is then possible to store the values e_i^k with odd indices, i.e.

$$d_j^k = e_{2j-1}^k, \quad 1 \le j \le N_{k-1} \tag{48}$$

and use relation (47) in order to formulate a system of equations

$$\sum \alpha_{2l} e_{2j+2l}^k = -\sum \alpha_{2l-1} e_{2j+2l-1}^k \tag{49}$$

for the unknowns $(e_2^k, e_4^k, \ldots, e_{N_k}^k)$. For the weight functions in the chain (42), system (49) has an invertible, well conditioned coefficient matrix, and the procedure is well defined.

The definition of scale coefficients given in (48) leads to a simple definition for the operator G_k:

$$(G_k)_{ij} = \delta_{2i-1,j}. \tag{50}$$

The operator E_k is then obtained from (49). The columns of E_k provide a set of basis vectors for $\mathcal{N}(D_k^{k-1})$.

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 describe the possible reconstruction techniques in each particular case.

Unlike the wavelet framework, the reconstruction operators we shall consider will not be hierarchical in general. Thus, checking the stability of the algorithms shall be one of our major concerns.

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 mere existence of this operator guarantees stability of the associated compression scheme.

It is worth taking a look at what the general relations (28) and (29) mean when the sequence of discretization is given by (34). Since $V^k = S^k$ are spaces of sequences, we shall always consider $\eta_i^k = \delta_i^k$, the canonical basis functions.

Let $\{\mathcal{R}_k^H\}$ be a sequence of reconstruction operators for the sequence $\{\mathcal{D}_k\}$ given by (34), so that $\{\mathcal{R}_k^H \mathcal{D}_k\}$ is hierarchical. Then

$$\varphi_i^k := \mathcal{R}_k^H \delta_i^k \quad \Rightarrow \quad \mathcal{D}_k \varphi_i^k = \mathcal{D}_k \mathcal{R}_k^H \delta_i^k = \delta_i^k \qquad \equiv \qquad <\omega_l^k, \varphi_i^k> = \delta_{il} \quad \forall i, \forall l \tag{51}$$

i.e. the sets $\{\omega_i^k\}$ and $\{\varphi_i^k\}$ are biorthogonal. Moreover, if $\psi_j^k := \mathcal{R}_k^H \mu_j^k$

$$\mathcal{D}_{k-1}\psi_j^k = D_k^{k-1}\mathcal{D}_k\psi_j^k = D_k^{k-1}\mathcal{D}_k\mathcal{R}_k^H\mu_j^k = D_k^{k-1}\mu_j^k = 0 \quad \equiv \quad <\omega_i^{k-1}, \psi_j^k> = 0 \quad \forall i, \forall j \tag{52}$$

The dilation relation for $\omega(x)$ implies then that $<\omega_i^m, \psi_j^k> = 0, \quad \forall m < k$ and $\forall i$. Defining

$$\Omega_k = \text{span}\{\omega_i^k\}$$

(52) implies that

$$\Omega_{k-1} \perp \Psi^k.$$

In general the direct sum decomposition $\Phi^k = \Phi^{k-1} \oplus \Psi^k$ is not an orthogonal decomposition. If $\Omega_k = \Phi^k$, as is the case when the decomposition and reconstruction filters are the same to start with (i.e. $\omega_i^k = \varphi_i^k$), then (51) implies that $\{\varphi_i^k\}$ is an orthonormal basis in Φ^k, and (52) implies that the direct sum decompositions in (28) and (29) are orthogonal. Notice also that (26) leads to

$$\begin{aligned} <\psi_j^k, \psi_l^k> &= <\sum_m (\hat{E}_k)_{mj}\varphi_m^k, \sum_p (\hat{E}_k)_{pl}\varphi_p^k> = \sum_m\sum_p (\hat{E}_k)_{mj}(\hat{E}_k)_{pl} <\varphi_m^k, \varphi_p^k> \\ &= \sum_m\sum_p (\hat{E}_k)_{mj}(\hat{E}_k)_{pl}\delta_{pm} = \sum_m (\hat{E}_k)_{mj}(\hat{E}_k)_{ml} = (\hat{E}_k^* \cdot \hat{E}_k)_{jl}. \end{aligned}$$

Thus, if $(\hat{E}_k^* \cdot \hat{E}_k) = I$, as is the case in the orthogonal wavelet case, the set $\{\psi_j^k\}$ is also orthonormal. It is interesting to notice that these orthogonality relations are a consequence, rather than the essence of the derivation.

Orthogonal and biorthogonal wavelets can be cast into the framework of discretization and reconstruction with $\mathcal{F} = L_2^{loc}(R)$ (see [17, 18]). However, there is a difference in emphasis between Harten's general framework and the classical wavelet theory: The starting point, and also the main tool, is not functional analysis but approximation theory. The functional structure comes as a consequence.

The greatest advantage of the general framework developed in [16, 17, 18] lies in its flexibility. Even in these simple settings, it allows for an easy, and natural, handling of boundaries. It makes also possible to consider data-dependent reconstruction techniques, which are needed to obtain near-to-optimal data-compression rates (see [16, 19] and Part II).

In the remainder of this section we re-examine the multiresolution schemes derived from the first member of the hierarchy (42), the interpolatory multiresolution seting, because of its intrinsic relation to the hat-weighted multiresolution .

4.1. Interpolatory MR analysis. In order for the integral in (32) to make sense when $\omega(x) = \delta(x)$ we need to consider functions which are continuous. We shall thus consider $f \in \mathcal{F} = \mathcal{C}([0,1])$ and $X^k = \{x_i^k\}_{i=0}^{J_k}$, $x_i^k = ih_k$, $h_k = 2^{-k}h_0$, $J_k = 2^k J_0$. In this case, $\dim S^k = \dim X^k = J_k + 1$.

Here we have

$$\mathcal{D}_k : \mathcal{C}[0,1] \longrightarrow S^k \qquad \bar{f}_j^k = (\mathcal{D}_k f)_j = f(x_j^k), \quad 0 \le j \le J_k \tag{53}$$

$$(D_k^{k-1})_{ij} = \delta_{2i,j} \qquad \mathcal{N}(D_k^{k-1}) = \{s^k \in S^k \quad | \quad s_{2i}^k = 0\} \tag{54}$$

To define the operators G_k and E_k we can use any of the two alternatives described in the beginning of the section, however both alternatives lead to the same operators in this case, (see also [17]). Thus, G_k is defined as in (50) and E_k has the following expression

$$(E_k)_{ij} = \delta_{2j-1,i}. \tag{55}$$

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

$$\mathcal{R}_k : S^k \longrightarrow \mathcal{C}[0,1]; \qquad \mathcal{D}_k\mathcal{R}_k\bar{f}^k = \bar{f}^k, \tag{56}$$

which means

$$(\mathcal{R}_k\bar{f}^k)(x_j^k) = \bar{f}_j^k = f(x_j^k). \tag{57}$$

Therefore, $\mathcal{R}_k$ 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.

If we denote by I_k any interpolatory reconstruction of the data $\bar{f}^k$, i.e.

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

the encoding and decoding algorithms (1) and (2) take the following simple form (see [16, 17]):

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

$$\left\{\begin{array}{lll} \text{Do} & k = L, 1 & \\ & \bar{f}_j^{k-1} = \bar{f}_{2j}^k, & 0 \le j \le J_{k-1} \\ & d_j^k = \bar{f}_{2j-1}^k - I_{k-1}(x_{2j-1}^k;\bar{f}^k), & 1 \le j \le J_{k-1} \end{array}\right. \tag{58}$$

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

$$(59) \qquad \begin{cases} \text{Do} \quad k = 1, L & \\ \quad \bar{f}^k_{2j} = \bar{f}^{k-1}_j, & 0 \le j \le J_{k-1} \\ \quad \bar{f}^k_{2j-1} = I_{k-1}(x^k_{2j-1}; \bar{f}^{k-1}) + d^k_j, & 1 \le j \le J_{k-1} \end{cases}$$

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 [16, 17] different types of interpolatory reconstructions are considered. Here we shall review just one of them, the piecewise polynomial interpolation.

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

$$\mathcal{S} = \mathcal{S}(r, s) = \{-s, -s+1 \cdots, -s+r\}, \qquad r \ge s > 0, \quad r \ge 1$$

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

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

It is clear that

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

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

$$(60) \qquad I_k(x, \bar{f}^k) = q^k_j(x; \bar{f}^k, r, s) \qquad x \in [x^k_{j-1}, x^k_j], \quad 1 \le j \le J_k$$

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

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

When the given function is periodic, i.e.

$$\bar{f}^k_{-j} = \bar{f}^k_{J_k - j}, \qquad \bar{f}^k_{J_k+j+1} = \bar{f}^k_{j+1}, \quad 0 \le j < J_k$$

the data to construct the polynomial function $I_k(x, f^k)$ 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 $I_k(x; \bar{f}^k)$ in the non-periodic case is, thus, as follows:

$$(61) \qquad I_k(x; \bar{f}^k) = \begin{cases} q^k_j(x; \bar{f}^k; r, j) & 1 \le j \le s-1 \\ q^k_j(x; \bar{f}^k; r, s) & s \le j \le J_k - s + 1 \\ q^k_j(x; \bar{f}^k; r, j - J_k + r) & J_k - s + 2 \le j \le J_k \end{cases}$$

Observe that if $f(x) = P(x)$, where $P(x)$ is a polynomial of degree less than or equal to r, then $q^k_j(x; f^k, r, s) = f(x)$ for $x \in [x^k_{j-1}, x^k_j]$, i.e. $I_k(x, f^k) = f(x)$. Thus, the space of polynomials of degree less than or equal to r satisfy $I_k \mathcal{D}_k f = f$. In addition, for smooth functions

$$I_k(x, f^k) = f(x) + O(h_k)^{2s}.$$

Therefore, the order of the reconstruction procedure is $r + 1 = 2s$.

As usual, we consider $\eta_i^k = \delta_i^k$ with $(\delta_i^k)_l = \delta_{il}$ as the basis functions for the spaces S^k. The results of section 3 imply that 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 $I_k(x; \bar{f}^k)$. Thus, convergence (as $L \to \infty$) of the functions $\varphi_i^{m,L}$ obtained by repeated interpolation of δ_i^m:

$$\bar{\varphi}_i^{m,L} = (\Pi_{k=m+1}^{L} P_{k-1}^k)\delta_i^m = A_m^L \delta_i^m; \qquad \varphi_i^{m,L} = I_L(x; \bar{\varphi}_i^{m,L}),$$

immediately implies the stability of the associated multiresolution scheme.

Here, we are only concerned with symmetric interpolatory procedures, $r = 2s - 1$. For $r = 1$, I_k is the piecewise-linear interpolation which is hierarchical and, as a consequence, the associated multiresolution scheme is stable. For $r = 3, 5$, we can use the connection between the interpolatory multiresolution framework and the theory of recursive subdivision (at least under periodicity assumptions) to prove the stability of the multiresolution schemes.

Since we are considering only data-independent interpolatory techniques, the reconstruction operator is linear and we can write

$$I_k(x; \bar{f}^k) = \sum_i \bar{f}_i^k u_i^k, \qquad u_i^k(x) = I_k(x; \delta_i^k).$$

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

$$\bar{f}_i^k = (P_{k-1}^k \bar{f}^{k-1})_i = I_{k-1}(x_i^k; \bar{f}^{k-1}) = \sum_j \bar{f}_j^{k-1} u_j^{k-1}(x_i^k).$$

Therefore

$$(P_{k-1}^k)_{2i,j} = u_j^{k-1}(x_{2i}^k) = \delta_{i,j}, \qquad (P_{k-1}^k)_{2i-1,j} = u_j^{k-1}(x_{2i-1}^k). \tag{62}$$

Let us consider first the periodic case. A straightforward algebraic manipulation shows that

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

Because of periodicity, we can consider (63) for all $i \in Z$. Observe that the coefficients of $\bar{f}_{i+m}^{k-1}$ depend only on m and they vanish if $m < -s$ or $m > -s + r$. Formula (63) describes in fact the refinement rule of an *interpolatory subdivision scheme*, a special subclass of stationary subdivision or refinement schemes.

Stationary subdivision schemes are appear in computer-aided geometric design as a method for the definition and generation of curves and surfaces. The general form of a stationary subdivision scheme in one dimension (univariate) is as follows:

$$p_i^k = \sum_l \gamma_{i-2l} p_l^{k-1}.$$

The coefficients $\{\gamma_l\}$ are the *mask* of the scheme (in the terminology of the refinement literature) and it is assumed that only finitely many γ's are non-zero. The last relation can also be expressed as

$$\begin{aligned} p^k_{2i} &= \sum_l \gamma_{2l} p^{k-1}_{i-l} \\ p^k_{2i-1} &= \sum_l \gamma_{2l-1} p^{k-1}_{i-l} \end{aligned}$$

Thus, 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 $\{p^0_i, i \in Z\}$, the binary subdivision refinement of this sequence defines, in the limit $k \to \infty$, an infinite set of points (corresponding to all rational numbers whose denominator is a power of 2 and integer translations of these). The main question is then whether or not these values admit a continuous extension to the real axis, i.e. whether or not the subdivision scheme converges uniformly for any set of initial control points.

Interpolatory subdivision schemes have the property that the limit curve interpolates the original control points. This requirement is met if at each stage of the iteration the previous control points are left unchanged. Thus, for an interpolatory subdivision scheme, $\gamma_{2l} = \delta_{l,0}$. The odd mask coefficients of the interpolatory subdivision defined by (63) are $\gamma_{2l-1} = L_{-l}(-1/2)$, $-s \leq l \leq -s + r$.

Under the assumption of periodicity, the sequence $A^L_k \bar{f}^k$ can then be re-interpreted as the sequence of control points obtained after L applications of the interpolatory subdivision scheme (63) on the set of control points $\{\bar{f}^k\}$. Using the results of [14, 7], it is proven in [18] that the convergence of the interpolatory subdivision scheme given by (63) implies convergence of the cosmetic refinement sequence. In fact,

(64) $$A^L_k \bar{f}^k \to f(x) \quad \in \mathcal{C}[0,1] \qquad \Rightarrow \qquad I_L A^L_k \bar{f}^k \to f(x) \text{ in } \mathcal{C}[0,1] \quad \text{as } L \to \infty$$

Thus, the convergence of the cosmetic refinement limiting process in $\mathcal{F} = \mathcal{C}[0,1]$ is intrinsically connected to the convergence of the subdivision scheme (63), which is now a well documented issue ([7] is an excellent review on stationary subdivision and [22] on interpolatory subdivision and its relation to wavelets; see also [14, 10, 11] and references therein).

Deslauries and Dubuc [12] proved convergence of the recursive refinement process (63) for $r = 3$ and 5. The considerations above lead us to conclude that the the hierarchical reconstruction (69) exists for $r = 3$ and 5; hence the associated multiresolution schemes are stable.

There is a close connection between the convergence of a subdivision scheme and the existence of a compactly supported function satisfying the functional equation

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

A convergent subdivision scheme determines uniquely $\varphi(x)$, which is obtained by applying the subdivision scheme to the sequence $\bar{f}^0_i = \delta_{0,i}$. In addition, for any $\bar{f}^0$ the limiting curve obtained by recursive subdivision is

$$(S^\infty \bar{f}^0)(x) = \sum_i \bar{f}^0_i \varphi(x - i).$$

Moreover, if an interpolatory subdivision process converges for the δ sequence, then it converges for all bounded sequences.

The connection with stationary subdivision implies that the existence of the cosmetic refinement limit in $\mathcal{F} = \mathcal{C}[0,1]$ for any discrete set of initial data depends only on the existence of such a limit for $\bar{f}^0 = \delta_0^0$. If this limit exists, then it satisfies a functional relation given by the *mask* of the subdivision scheme, i.e. the coefficients of the prediction matrix.

Thus, under the periodicity assumption, if

$$\exists \lim_{k\to\infty} I_k(x; \bar{\varphi}_0^{0,k}) = \varphi_0(x) \tag{66}$$

then for all i and m

$$\exists \lim_{k\to\infty} I_k(x; \bar{\varphi}_i^{m,m+k}) = \varphi_i^m(x) = \varphi(\frac{x}{h_m} - i) \tag{67}$$

where

$$\varphi(x) = \varphi_0(xh_0), \quad \text{support}(\varphi) = [2s-2r-1, 2s-1]. \tag{68}$$

Moreover, it follows from (62) and (65) that $\varphi(x)$ satisfies the dilation relation

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

These facts were also proven in [17], without explicitly using the subdivision refinement theory.

In addition, Deslauries and Dubuc [12] proved that for $r = 3$ the limit function corresponding to the set of data $\bar{\varphi}_0^{0,0} = \delta_0^0$ is 'almost' $\mathcal{C}^2$ (it is $\mathcal{C}^1$ and its first derivative is differentiable almost everywhere). For $r = 5$, the limit function is $\mathcal{C}^2$.

We recall that the hierarchical form of I_k, which we refer to as I_k^∞ since it is obtained via the cosmetic refinement limiting process, has the form

$$I_k^\infty(x; \bar{f}^k) = \sum_i \bar{f}_i^k \varphi_i^k. \tag{69}$$

The two level relation between the limiting functions φ_i^k in (25) can also be obtained directly, using the dilation relation for $\varphi(x)$.

Notice that (26) 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

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

Moreover, using the hierarchical interpolation (69), we can write (27) as

$$I_L^\infty(x; \mathcal{D}_L f) = I_0^\infty(x; \mathcal{D}_0 f) + \sum_{k=1}^{L}\sum_{j=1}^{J_{k-1}} d_j^k(f)\varphi_{2j-1}^k(x) \tag{70}$$

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

When I_k is hierarchical to begin with (e.g. $r = 1$ or piecewise interpolation techniques using splines) $I_k^\infty = I_k$, $\varphi_i^k(x) = I_k(x; \delta_i^k)$. In the finite element context, the set

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

is referred to as a hierarchical basis (see [23]).

In order to deduce the stability of the multiresolution scheme associated to the non-periodic case, we consider again the cosmetic refinement process. We start by setting $\bar{\varphi}_i^{m,m} = \delta_i^m$ at the points of the m-grid and apply repeatedly the prediction operator. Convergence of (66) implies also convergence of (67) for those indexes i such that the support of the limit function in the periodic case does not intersect any of the intervals affected by the one-sided interpolatory procedure. Since in the symmetric case the support of $\varphi(x)$ in (68) is $[-(2s-1),(2s-1)]$, there are $2s$ 'special' boundary functions (at each boundary) at each resolution level.

At this moment, there are no convergence results available to us in order to prove, analytically, the existence of these boundary functions. In Figure 1 we show numerical evidence that the limit functions do indeed exist and have the desired interpolation properties. This accounts for the evident stability of the associated compression algorithms.

Figure 1 shows the limiting functions when $J_0 = 8$ and $r = 3$. The support of $\varphi(x)$ in (68) is $[-3,3]$, so we get 4 'special' limiting functions due to boundary effects. The numerical results are obtained by applying the inverse multiresolution transform (for L large enough) to compute $M^{-1}(u^0, 0, \ldots, 0)$, where u^0 is the unit sequence $u_i^0 = \delta_{i,0}$. The convergence of the limiting process seems to be very fast; we display $\varphi_i^{0,7}, i = 4, \ldots, 8$. These are basically indistinguishable from $\varphi_i^{0,L}$ for L larger than 7.

It is worth noticing that the boundary limit functions corresponding to higher levels of resolution are scaled versions of those for lower resolution levels, as long as only one of the boundaries makes its influence felt (see Figure 2).

The functions $\psi_j^k = \varphi_{2j-1}^k$, which would be obtained by computing $M^{-1}(0, \delta_j^1, 0, \ldots, 0)$, are just scaled versions of the odd-indexed φ's and are not shown.

5. Hat-weighted Multiresolution. Let us describe now multiresolution schemes corresponding to the third member of the hierarchy (42), the hat function. The discretization process defined in (32) requires that the function under consideration be integrated against scaled translates of the hat function. Since these are continuous functions, δ-type singularities can be allowed.

Let us consider, as in the interpolatory and cell-average frameworks, the unit interval $[0,1]$ and the sequence of nested dyadic grids $X^k = \{x_i^k\}_{i=0}^{J_k}$, $x_i^k = ih_k$, $h_k = 2^{-k}h_0$, $J_k = 2^kJ_0$.

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

$$\mathcal{D}_k : \mathcal{F} \longrightarrow S^k \qquad \bar{f}_j^k = (\mathcal{D}_k f)_i =< f, \omega_i^k >, \quad 1 \le i \le N_k = J_k - 1$$

The weighted averages $\bar{f}_j^k$ 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 a finite number of 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 (41), leads to

$$\omega_i^{k-1} = \frac{1}{4}\omega_{2i-1}^k + \frac{1}{2}\omega_{2i}^k + \frac{1}{4}\omega_{2i+1}^k \tag{71}$$

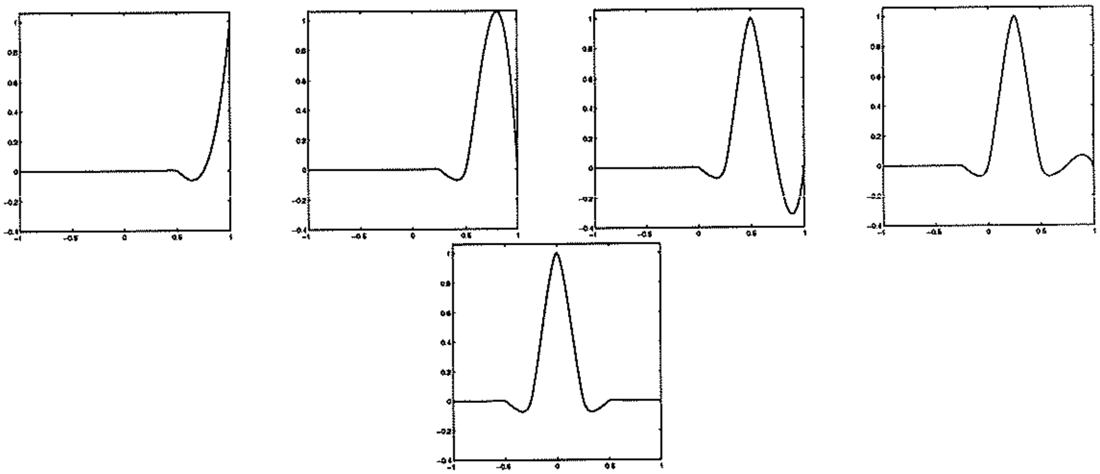

FIG. 1. *Limiting functions for interpolatory multiresolution.* $J_0 = 8$ $r = 3$. *Top: 'special' boundary functions. Bottom: periodic case.*

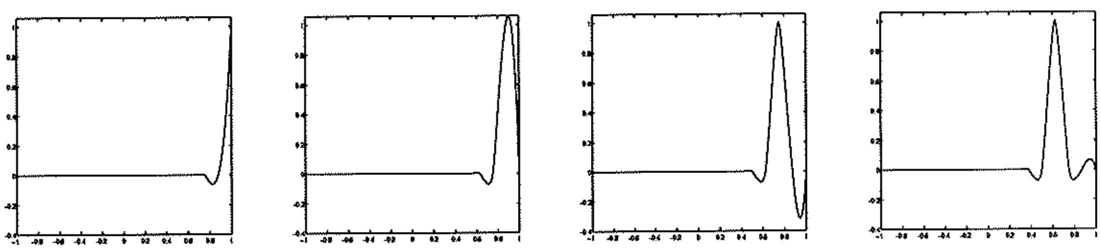

FIG. 2. *Limiting functions for interpolatory multiresolution.* $J_0 = 16$ $r = 3$. *'special' boundary functions.*

or, in terms of the weighted averages

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

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:

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

The definition of the decimation operator determines the multiresolution setting. To obtain a multiresolution scheme we need an appropriate prediction operator, which in our context is equivalent to finding an appropriate reconstruction operator for $\mathcal{D}_k$.

The reconstruction procedure $\mathcal{R}_k$ must satisfy

(74) $$\mathcal{R}_k : S_k \longrightarrow \mathcal{F} \qquad (\mathcal{D}_k \mathcal{R}_k \bar{f}^k)_j = \bar{f}_j^k.$$

In what follows we describe a procedure to compute $\mathcal{R}_k$ for the hat-averaged framework which is a generalization of the 'reconstruction via primitive function' developed in the context of cell averages (see [17, 19]). 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 $(0, 1)$. Define its "second primitive" as

(75) $$H(x) = \int_0^x \int_0^y f(z) dz dy.$$

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

$$\bar{f}_i^k = \langle f, \omega_i^k \rangle = \frac{1}{h_k^2}(H_{i+1}^k - 2H_i^k + H_{i-1}^k), \quad 1 \leq i \leq J_k - 1, \tag{76}$$

$$\text{where} \quad H_l^k = H(x_l^k) \quad 0 \leq l \leq J_k$$

When $f(x)$ is an integrable function, (76) is easily proven integrating by parts. If $f(x) = \delta(x-a)$, $a \in (0,1)$, then

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

and it is also a straightforward matter to prove that

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

For a general $f(x) \in \mathcal{F}$, linearity of the integral implies (76).

The definition of $H(x)$ in (75) leads to $H_{J_k}^k = H(1) = \int_0^1 \int_0^y f(z)dzdy$ and $H_0^k = H(0) = 0$, $\forall k$. $H(x)$ in (75) is one of the second primitives of $f(x)$, (i.e. a function satisfying (76)); however another second primitive, which we also label $H(x)$, is sometimes more convenient

$$H(x) = \int_0^x \int_0^y f(z)dzdy - \alpha x \qquad \alpha = \int_0^1 \int_0^y f(z)dzdy \tag{77}$$

because, for this particular one $H_0^k = H(0) = 0$, and $H_{J_k}^k = H(1) = 0$. Choosing the lower limits to be zero, or modifying the basic definition of $H(x)$ given by (75) by a first order polynomial, as in (77), amounts to computing different values for $H(0)$ and $H(1)$. Once these have been specified, (76) establishes a one-to-one correspondence between the sets $\{\bar{f}_i^k\}_{i=1}^{J_k-1}$ and $\{H_i^k\}_{i=1}^{J_k-1}$. In fact,

$$h_k^2 \bar{f}^k = MH^k, \qquad M_{i,j} = \begin{cases} -2 & i = j \\ 1 & |i-j| = 1 \\ 0 & \text{else} \end{cases} \tag{78}$$

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 (77). We can then interpolate the point-values of the "second primitive" by any interpolation procedure $I_k(x; H^k)$ and define

$$(\mathcal{R}_k \bar{f}^k)(x) = \frac{d^2}{dx^2} I_k(x; H^k). \tag{79}$$

In general, $I_k(x; H^k)$ 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. $R_k \bar{f}^k$ may have a finite number of δ-type singularities, which will be located at those (interior) grid points where the first derivative of $I_k(x; H^k)$ 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 5.1. *Let $I(x)$ be a piecewise smooth function of the form*

$$I(x) = \begin{cases} I_L(x) & \text{if } x < 0 \\ I_R(x) & \text{if } x > 0. \end{cases} \tag{80}$$

Then, its derivative in the distribution sense is

$$\frac{d}{dx}I(x) = \tilde{I}(x) + (I_R(0) - I_L(0))\delta(x) \tag{81}$$

where

$$\tilde{I}(x) = \begin{cases} \frac{d}{dx}I_L(x) & \text{if } x < 0 \\ \frac{d}{dx}I_R(x) & \text{if } x > 0. \end{cases} \tag{82}$$

Lemma 5.1 implies that if the interpolatory function is defined as

$$I_k(x; H^k) = I_{k,j}(x; H^k) \qquad \text{for } x \in [x_{j-1}^k, x_j^k] \tag{83}$$

in the distribution sense, we have

$$(\mathcal{R}_k \bar{f}^k)(x) = \tilde{I}_k(x) + \sum_{j=1}^{J_k - 1} s_j^k \delta(x - x_j^k) \tag{84}$$

where $\tilde{I}_k$ is defined as

$$\tilde{I}_k(x) = \frac{d^2}{dx^2} I_{k,j}(x; H^k) \qquad \text{for } x \in [x_{j-1}^k, x_j^k], \tag{85}$$

and

$$s_j^k = \left[\frac{d}{dx}I_{k,j+1}(x; H^k) - \frac{d}{dx}I_{k,j}(x; H^k)\right]_{x = x_j^k} = I_k{}'(x_j^k + 0; H^k) - I_k{}'(x_j^k - 0; H^k). \tag{86}$$

Obviously, $(\mathcal{R}_k \bar{f}^k)(x) \in \mathcal{F}$. To prove that $\mathcal{R}_k$ in (79) is a right inverse of $\mathcal{D}_k$ we need the following lemma:

LEMMA 5.2. *Let $f(x)$ be an integrable function and let $H(x)$ be a continuous function in $[0,1]$ twice differentiable in $[x_{l-1}^k, x_l^k]$, $l = j, j+1$ such that $H''(x) = f(x)$ almost everywhere in $[x_{l-1}^k, x_l^k]$ for $l = j, j+1$. Then*

$$< f, \omega_j^k > = \bar{f}_j^k = \frac{1}{h_k}\left[H'(x_j^k - 0) - H'(x_j^k + 0)\right] + \frac{1}{h_k^2}(H_{j+1}^k - 2H_j^k + H_{j-1}^k) \tag{87}$$

Proof. For an integrable function $f(x)$,

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

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

Notice that (85) implies that we can apply lemma 5.2 to $\tilde{I}_k$ and I_k. We thus obtain,

$$(\mathcal{D}_k \mathcal{R}_k \bar{f}^k)_j = < \mathcal{R}_k \bar{f}^k, \omega_j^k > = < \tilde{I}_k(x) + \sum_{l=1}^{J_k - 1} \delta(x - x_l^k) s_l^k, \omega_j^k >$$

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

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

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

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

Lemma 5.2 implies

$$
\begin{aligned}
< \tilde{I}_{k-1}, \omega_j^k > &= \frac{1}{h_k}(I'_{k-1}(x_j^k - 0; H^{k-1}) - I'_{k-1}(x_j^k + 0; H^{k-1})) \\
&+ \frac{1}{h_k^2}\left(I_{k-1}(x_{j-1}^k; H^{k-1}) - 2I_{k-1}(x_j^k; H^{k-1}) + I_{k-1}(x_{j+1}^k; H^{k-1})\right),
\end{aligned}
$$

so that for $j = 2m$ we have

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

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

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

$$
I'_{k-1}(x_j^k - 0) - I'_{k-1}(x_j^k + 0) = 0, \qquad < \delta(x - x_l^{k-1}), \omega_j^k >= 0.
$$

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

$$
(P_{k-1}^k \bar{f}^{k-1})_j = \frac{1}{h_k^2}\left(I_{k-1}(x_{j-1}^k; H^{k-1}) - 2I_{k-1}(x_j^k; H^{k-1}) + I_{k-1}(x_{j+1}^k; H^{k-1})\right). \tag{88}
$$

This expression is particularly useful because, once the interpolation procedure is specified, it will allow us to write the predicted values in terms of $\{\bar{f}^{k-1}\}$. Thus, as in the cell-averaged framework, there is no need to compute explicitly the point values of the second primitive (or the values s_j^k).

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 (47) takes the form

$$
e_{2i}^k = -\frac{1}{2} e_{2i-1}^k - \frac{1}{2} e_{2i+1}^k. \tag{89}
$$

Thus, the natural choice of transfer operators described in Section 4 is

$$
d^k = G_k e^k \qquad d_j^k = e_{2j-1}^k, \quad 1 \le j \le J_{k-1} \tag{90}
$$

and

$$(91) \qquad e^k = E_k d^k \qquad \begin{cases} e^k_{2j-1} &= d^k_j \\ e^k_{2j} &= -\frac{1}{2}(d^k_j + d^k_{j+1}). \end{cases}$$

Therefore

$$(G_k)_{ij} = \delta_{2i-1,j} \qquad (E_k)_{ij} = \begin{cases} \delta_{i,2j-1} - \frac{1}{2}(\delta_{i,2j} + \delta_{i,2(j-1)}) & 1 < j < J_{k-1} \\ \delta_{i,2j-1} - \frac{1}{2}\delta_{i,2j} & j = 1 \\ \delta_{i,2j-1} - \frac{1}{2}\delta_{i,2(j-1)} & j = J_{k-1} \end{cases}$$

On the other hand, we can also define the transfer matrices in a wavelet-like manner, i.e., for each j, $1 \leq j \leq J_{k-1}$

$$(92) \qquad (\bar{\mu}^k_j)_l = (-1)^l \alpha_{2j-l-1}, \qquad 1 \leq l \leq J_k - 1.$$

It is easy to see that $\bar{E}_k = [\bar{\mu}^k_1, \ldots, \bar{\mu}^k_{J_{k-1}}] = -\frac{1}{2}E_k$. Thus the choice

$$\bar{G}_k = (\bar{E}^*_k \bar{E}_k)^{-1}\bar{E}^*_k = -2(E^*_k E_k)^{-1}E^*_k$$

basically amounts to computing the scale coefficients using *another* left inverse of E_k in S^k, namely $B_k = (E^*_k E_k)^{-1}E^*_k$. Since

$$\bar{\mu}^k_i \cdot \bar{\mu}^k_j = \begin{cases} \frac{1}{16} & |i-j| = 1 \\ \frac{5}{16} & i = j \neq 1, J_{k-1} \\ \frac{5}{16} & i = j = 1, J_{k-1} \\ 0 & \text{otherwise,} \end{cases}$$

the computation of the scale coefficients using $\bar{G}_k$ requires solving a tridiagonal system. Although this is an $O(n)$ task, the extra amount of calculation is not necessary since both left inverses give identical values when acting on $\mathcal{N}(D^{k-1}_k)$. To see this, let μ^k_j be the j-th column of E_k. Since these form a basis of $\mathcal{N}(D^{k-1}_k)$, each $e^k \in \mathcal{N}(D^{k-1}_k)$ can be uniquely expressed as $e^k = \sum d^k_j \mu^k_j$. Then

$$B_k E_k = I \Rightarrow B_k \mu^k_j = \delta_j \qquad \text{and} \qquad G_k E_k = I \Rightarrow G_k \mu^k_j = \delta_j.$$

Thus

$$(B_k - G_k)e^k = \sum d^k_j (B_k - G_k)\mu^k_j = 0 \Rightarrow d^k = G_k e^k = B_k e^k.$$

Once all the necessary ingredients of the multiresolution scheme have been specified, we can give an explicit description of the hat-based multiresolution transform and its inverse,

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

$$(93) \qquad \begin{cases} \text{Do} \quad k = L, 1 \\ \quad \bar{f}^{k-1}_i = \frac{1}{4}(\bar{f}^k_{2i-1} + 2\bar{f}^k_{2i} + \bar{f}^k_{2i+1}), & 1 \leq i \leq J_{k-1} - 1 \\ \quad d^k_i = \bar{f}^k_{2i-1} - (P^k_{k-1}\bar{f}^{k-1})_{2i-1} & 1 \leq i \leq J_{k-1} \end{cases}$$

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

$$
(94) \qquad \begin{cases} \text{Do} \quad k=1,L \\ \bar{f}^k_{2i-1} = (P^k_{k-1}f^{k-1})_{2i-1} + d^k_i & 1 \le i \le J_{k-1} \\ \bar{f}^k_{2i} = 2\bar{f}^{k-1}_i - \frac{1}{2}(\bar{f}^k_{2i-1} + \bar{f}^k_{2i+1}), & 1 \le i \le J_{k-1}-1 \end{cases}
$$

Using (88), in (93) and (94), the algorithms for encoding and decoding have the general form of a fast discrete wavelet-type algorithm.

Notice that defining

$$
\bar{f}^k_{2i} = (P^k_{k-1}\bar{f}^{k-1})_{2i} + (E_k d^k)_{2i}
$$

is equivalent to

$$
\bar{f}^k_{2i} = 2\bar{f}^{k-1}_i - \frac{1}{2}(\bar{f}^k_{2i-1} + \bar{f}^k_{2i+1}),
$$

because, from the definition of the odd averages in (94) and (88), we obtain

$$
\begin{aligned}
\bar{f}^k_{2i} + \frac{1}{2}(\bar{f}^k_{2i-1} + \bar{f}^k_{2i+1}) &= (P^k_{k-1}\bar{f}^{k-1})_{2i} - \frac{1}{2}(d^k_i + d^k_{i+1}) \\
&+ \frac{1}{2}[(P^k_{k-1}\bar{f}^{k-1})_{2i-1} + d^k_i + (P^k_{k-1}\bar{f}^{k-1})_{2i+1} + d^k_{i+1}] \\
&= (P^k_{k-1}\bar{f}^{k-1})_{2i} + \frac{1}{2}[(P^k_{k-1}\bar{f}^{k-1})_{2i-1} + (P^k_{k-1}\bar{f}^{k-1})_{2i+1}] \\
&= \frac{1}{2h_k^2}(H^{k-1}_{i-1} - 2H^{k-1}_i + H^{k-1}_{i+1}) = 2\bar{f}^{k-1}_i
\end{aligned}
$$

As pointed out in Section 3, stability of the direct multiresolution transform is a consequence of the nested character of the sequence of discretization. Stability of the inverse multiresolution transform follows from the *existence* of a hierarchical form for the reconstruction operator. The hierarchical reconstruction is constructed via the *cosmetic refinement* limiting process (see Theorems 3.2 and 3.3).

Our design of multiresolution schemes for the hat-average framework is based on the underlying interpolatory multiresolution schemes. We shall prove next that multiresolution schemes for hat-average multiresolution settings are stable (that is, the reconstruction operator has a hierarchical form) provided the underlying interpolatory reconstruction used in its design has a hierarchical form.

Let us denote by $\tilde{\mathcal{D}}_k$ and I_k the discretization and reconstruction operators corresponding to the interpolatory framework of Section 4.1, while $\mathcal{D}_k$ and $\mathcal{R}_k$ are the discretization and reconstruction in the hat-average framework.

It is easy to see that, if the interpolation is itself hierarchical, then so is the reconstruction in the hat-average context,

LEMMA 5.3. *If the interpolatory reconstructions I_k are hierarchical, then so are the corresponding reconstructions obtained by (79).*

Proof. The hierarchical property of the interpolatory reconstructions means that

$$
I_k \tilde{\mathcal{D}}_k I_{k-1} H^{k-1} = I_{k-1} H^{k-1},
$$

where $(\tilde{\mathcal{D}}_k H)_i = H(x^k_i)$ is the discretization by point values operator which defines interpolatory multiresolution settings. The above relation can also be expressed as follows:

$$
(95) \qquad \text{If} \quad \tilde{H}^k_j = I_{k-1}(x^k_j; H^{k-1}) \quad 0 \le j \le J_k, \quad \text{then} \quad I_k(x;\tilde{H}^k) = I_{k-1}(x, H^{k-1}).
$$

To prove that

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

is also hierarchical, observe that (88) can be expressed as

$$\begin{aligned} (\mathcal{D}_k \mathcal{R}_{k-1} \bar{f}^{k-1})_j &= \frac{1}{h_k^2}\left\{ I_{k-1}(x_{j-1}^k; H^{k-1}) - 2I_{k-1}(x_j^k; H^{k-1}) + I_{k-1}(x_{j+1}^k; H^{k-1}) \right\} \\ &= \frac{1}{h_k^2}(\tilde{H}_{j-1}^k - 2\tilde{H}_j^k + \tilde{H}_{j+1}^k). \end{aligned}$$

Therefore, the set $\{\tilde{H}_j^k\}$ in (95) are the point values of a "second primitive" of the function $(\mathcal{R}_{k-1}\bar{f}^{k-1})(x)$ on the k-th grid (notice that, by construction $\tilde{H}_0^k = H_0^k; \tilde{H}_{J_k}^k = H_{J_k}^k$). Hence

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

which completes the proof. □

Thus, interpolatory techniques that lead to hierarchical reconstructions in the interpolatory multiresolution set-up lead directly to hierarchical reconstructions in the hat-average framework and, thus, to stable multiresolution schemes.

Let us assume now that that the reconstruction I_k of the underlying interpolatory framework admits a hierarchical form I_k^∞. Notice that $I_k^\infty(x; \tilde{\mathcal{D}}_k H)$ is a continuous function which interpolates $H(x)$ on the k-th grid. Lemma 5.3 implies that

$$\mathcal{R}_k^\infty(x; \mathcal{D}_k f) = \frac{d^2}{dx^2} I_k^\infty(x; \tilde{\mathcal{D}}_k H) \tag{96}$$

(where $H(x)$ is a second primitive of $f(x)$) is a hierarchical reconstruction in the hat-average framework. We then have

THEOREM 5.4. *$\mathcal{R}_k^\infty$ in (96) is the hierarchical form of the reconstruction $\mathcal{R}_k$; i.e.*

$$\lim_{L\to\infty} \mathcal{R}_L A_k^L = \mathcal{R}_k^\infty \tag{97}$$

Proof. Notice that $\mathcal{F}$ is a space of distributions, and therefore (97) is a limit in the distribution sense.

I_k^∞ is the hierarchical form of I_k, thus

$$I_k^\infty(x; \tilde{\mathcal{D}}_k H) = \lim_{L\to\infty} I_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H$$

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

Let us consider $\phi(x) \in C_0^\infty[0,1]$, an infinitely differentiable function with compact support in $[0,1]$. Then for any $f \in \mathcal{F}$

$$< \mathcal{R}_L A_k^L \mathcal{D}_k f, \phi > = < \frac{d^2}{dx^2} I_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \phi > = < I_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \frac{d^2}{dx^2}\phi > . \tag{98}$$

Since

$$\lim_{L\to\infty} < I_L \tilde{A}_k^L \tilde{\mathcal{D}}_k H, \frac{d^2}{dx^2}\phi > = < I_k^\infty(\cdot; \tilde{\mathcal{D}}_k H), \frac{d^2}{dx^2}\phi > = < \frac{d^2}{dx^2} I_k^\infty(\cdot; \tilde{\mathcal{D}}_k H), \phi > . \tag{99}$$

the proof is complete. □

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

$$I_{k-1}(x_{2i}^k; H^{k-1}) = I_{k-1}(x_i^{k-1}; H^{k-1}) = H_i^{k-1} = H_{2i}^k.$$

Thus

$$(100)\quad \begin{aligned} e_{2j-1}^k(f) &= \bar{f}_{2j-1}^k - (P_{k-1}^k \bar{f}^{k-1})_{2i-1} \\ &= \frac{1}{h_k^2}[H_{2j}^k - 2H_{2j-1}^k + H_{2j-2}^k] - \\ &\quad \frac{1}{h_k^2}\left(I_j^{k-1}(x_{2j-2}^k; H^{k-1}) - 2I_j^{k-1}(x_{2j-1}^k; H^{k-1}) + I_j^{k-1}(x_{2j}^k; H^{k-1})\right) \\ &= -\frac{2}{h_k^2}(H_{2j-1}^k - I_j^{k-1}(x_{2j-1}^k; H^{k-1})) = -\frac{2}{h_k^2} e_{2j-1}^k(H). \end{aligned}$$

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.

Let us see now how to obtain multi-scale decompositions in the hat-average multiresolution framework from multi-scale decompositions in the underlying interpolatory framework.

Consider $\{\{\bar{f}_i^k\}_{i=1}^{J_k-1}\}_{k=0}^L$ a hat-average multiresolution analysis of a function $f \in \mathcal{F}$, and let $\{\{H_i^k\}_{i=0}^{J_k}\}_{k=0}^L$ be the corresponding interpolatory multiresolution of its second primitive. Any linear interpolatory reconstruction satisfies

$$(101)\qquad I_k(x; H^k) = \sum_{j=0}^{J_k} H_j^k \tilde{\varphi}_j^k(x); \quad \tilde{\varphi}_j^k(x) = I_k(x; \delta_j^k).$$

Because of (76), it is easy to see that

$$H_{l+1}^k - H_l^k = H_1^k + h_k^2(\bar{f}_1^k + \bar{f}_2^k + \cdots + \bar{f}_l^k).$$

Thus, differentiating (101), we get

$$(102)\quad \begin{aligned} \mathcal{R}_k \bar{f}^k(x) &= \sum_{j=0}^{J_k} H_j^k \frac{d^2}{dx^2} \tilde{\varphi}_j^k(x); \\ &= H_1^k \frac{d^2}{dx^2}\tilde{\varphi}_1^k(x) + \sum_{j=2}^{J_k}\left(H_1 + \sum_{l=1}^{j-1}(H_{l+1}^k - H_l^k)\right)\frac{d^2}{dx^2}\tilde{\varphi}_j^k(x) \\ &= H_1^k \sum_{j=1}^{J_k} j \frac{d^2}{dx^2}\tilde{\varphi}_j^k(x) + \sum_{j=2}^{J_k}\left(\sum_{i=1}^{j-1}(j-i)\bar{f}_i^k\right) h_k^2 \frac{d^2}{dx^2}\tilde{\varphi}_j^k(x) \\ &= H_1^k \sum_{j=0}^{J_k} j \frac{d^2}{dx^2}\tilde{\varphi}_j^k(x) + \sum_{j=1}^{J_k-1} \bar{f}_j^k \left(\sum_{i=j+1}^{J_k}(i-j)h_k^2 \frac{d^2}{dx^2}\tilde{\varphi}_i^k(x)\right) \\ &= \sum_{i=j}^{J_k-1} \bar{f}_j^k \varphi_j^k(x), \end{aligned}$$

where

$$(103)\quad \begin{aligned} \varphi_i^k(x) &= \sum_{l=i+1}^{J_k}(l-i)h_k^2 \frac{d^2}{dx^2}\tilde{\varphi}_l^k(x) = h_k \frac{d^2}{dx^2} I_k(x; \sum_{l=i+1}^{J_k} h_k(l-i)\delta_l^k) \\ &= h_k \frac{d^2}{dx^2} I_k(x; \tilde{\mathcal{D}}_k(x - x_i^k)_+). \end{aligned}$$

To deduce (102) we have assumed that I_k is an interpolatory reconstruction that reproduces first order polynomials exactly. Then

$$\sum_{j=0}^{J_k} j \frac{d^2}{dx^2}\tilde{\varphi}_j^k(x) = \frac{1}{h_k}\frac{d^2}{dx^2} I_k(x; \sum_{j=0}^{J_k} j h_k \delta_j^{J_k}) = \frac{d^2}{dx^2} I_k(x; \tilde{\mathcal{D}}_k x) = 0$$

Notice that in order to get a consistent reconstruction in the hat-average case, i.e. $\mathcal{R}_k 1 = 1$ we need to consider interpolatory techniques that are at least third order accurate, i.e. that reproduce exactly polynomials of at least second degree.

Relation (102) implies that $\mathcal{R}_k S^k = \text{span}\{\varphi_1^k, \dots, \varphi_{N_k}^k\}$, with φ_i^k as defined in (103). If $\{I_k \tilde{\mathcal{D}}_k\}$ is hierarchical, then (26) and (27) lead, in the interpolatory context, to the following expression

$$I_k(x; H^k) - I_{k-1}(x; H^{k-1}) = \sum_{j=1}^{J_{k-1}} d_j^k(H) \tilde{\varphi}_{2j-1}^k(x). \tag{104}$$

Lemma 5.3 implies that the reconstruction for the hat-average set up which is obtained from I_k is also hierarchical. Then differentiating (104) and taking into account that

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

we obtain

$$(\mathcal{R}_k \bar{f}^k)(x) - (\mathcal{R}_{k-1}\bar{f}^{k-1})(x) = -\frac{h_k^2}{2} \sum_{j=1}^{J_{k-1}} d_j^k(f) \frac{d^2}{dx^2}\tilde{\varphi}_{2j-1}^k(x). \tag{105}$$

Hence

$$(\mathcal{R}_k \bar{f}^k)(x) - (\mathcal{R}_{k-1}\bar{f}^{k-1})(x) = \sum_{j=1}^{J_k - 1} d_j^k(f) \psi_j^k(x), \tag{106}$$

and so

$$(\mathcal{R}_L \bar{f}^L)(x) = \sum_{i=1}^{J_0 - 1} \bar{f}_i^0 \varphi_i^0 + \sum_{k=1}^{L} \sum_{j=1}^{J_k - 1} d_j^k(f) \psi_j^k(x). \tag{107}$$

with φ_i^k as in (103) and

$$\psi_j^k(x) = -\frac{h_k^2}{2}\frac{d^2}{dx^2}\tilde{\varphi}_{2j-1}^k(x), \quad 1 \le j \le J_{k-1}. \tag{108}$$

The functional structure in the hat-average framework can, thus, be deduced directly from the interpolatory framework.

Up to this point, we have not specified the kind of interpolatory procedure to be used in each multiresolution set-up. In the next section we describe the reconstruction operator obtained from the piecewise interpolation of Section 4.1.1 (see [2] for reconstructions derived from interpolation by splines).

5.1. Piecewise polynomial interpolation. Following the notation in Section 4.1.1, let $q_j^k(x; H^k, r, s)$ be the (unique) polynomial of degree r that interpolates $H(x)$ at the points As in Section 4.1.1, 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; thus we define I_k as follows

$$I_k(x, H^k) = q_j^k(x; H^k, r, s), \qquad x \in [x_{j-1}^k, x_j^k], \quad 1 \leq j \leq J_k.$$

In the non periodic case, I_k is defined using one sided stencils for some of its polynomial pieces. The definition of I_k in this case is given by (61).

Notice that if $f(x) = Q(x)$, a polynomial function of degree q, then $H(x) = P(x)$ where $P(x)$ is also polynomial of degree $q + 2$ that satisfies $P''(x) = Q(x)$. Then

$$I_k(x, H^k) = H(x)$$

for $q + 2 \leq r$, which implies (because of (79)) that

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

Hence, polynomials of degree up to $r - 2$ are reconstructed exactly and, therefore, the reconstruction is formally of order $p = r - 1 = r - 1 = 2s - 2$.

In what follows we give an explicit description of multiresolution schemes for the hat-average framework which are based on the piecewise interpolatory techniques of Section 4.1.1.

As in Section 4.1.1, we start by considering the periodic case. A periodic function can be treated as a non periodic one. However, when possible, centered reconstructions are preferred because their approximation errors are usually smaller than those of their non-centered counterparts.

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.

$$\mathcal{D}_k f = (\bar{f}_i^k)_{i=1}^{J_k}, \qquad \bar{f}_i^k = < f, \omega_i^k > .$$

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

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

$$(109) \qquad \begin{cases} \text{Do} \quad k = L, 1 \\ \quad \bar{f}_i^{k-1} = \frac{1}{4}(\bar{f}_{2i-1}^k + 2\bar{f}_{2i}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1} \\ \quad d_i^k = \bar{f}_{2i-1}^k - \sum_{l=1}^{s-1} \beta_l(\bar{f}_{i+l-1}^{k-1} + \bar{f}_{i-l}^{k-1}), & 1 \leq i \leq J_{k-1} \end{cases}$$

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

$$(110) \qquad \begin{cases} \text{Do} \quad k = 1, L \\ \quad \bar{f}_{2i-1}^k = d_i^k + \sum_{l=1}^{s-1} \beta_l(\bar{f}_{i+l-1}^{k-1} + \bar{f}_{i-l}^{k-1}), & 1 \leq i \leq J_{k-1} \\ \quad \bar{f}_{2i}^k = 2\bar{f}_i^{k-1} - \frac{1}{2}(\bar{f}_{2i-1}^k + \bar{f}_{2i+1}^k), & 1 \leq i \leq J_{k-1} \end{cases}$$

where p is the order of $\mathcal{R}_k$ and

$$
(111) \qquad \begin{cases} p=2 & \Rightarrow \quad \beta_1 = \frac{1}{2} \\ p=4 & \Rightarrow \quad \beta_1 = \frac{19}{32}, \beta_2 = -\frac{3}{32} \\ p=6 & \Rightarrow \quad \beta_1 = \frac{162}{256}, \beta_2 = -\frac{39}{256}, \beta_3 = \frac{5}{256} \end{cases}
$$

In the non periodic case, the algorithm above needs to be modified to account for the boundaries. The first i-loop in (109) and the second i-loop in (110) run only from $i = 1$ to $N_k = J_k - 1$. 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 follows:

$$
p=2 \quad \left\{ \begin{array}{lcl} d_1^k & = & \bar{f}_1^k - (\frac{3}{2}\bar{f}_1^{k-1} - \frac{1}{2}\bar{f}_2^{k-1}) \end{array} \right.
$$

$$
p=4 \quad \left\{ \begin{array}{lcl} d_1^k & = & \bar{f}_1^k - (\frac{65}{32}\bar{f}_1^{k-1} - \frac{57}{32}\bar{f}_2^{k-1} + \frac{31}{32}\bar{f}_3^{k-1} - \frac{7}{32}\bar{f}_4^{k-1}) \\ d_2^k & = & \bar{f}_3^k - (\frac{7}{32}\bar{f}_1^{k-1} + \frac{37}{32}\bar{f}_2^{k-1} - \frac{15}{32}\bar{f}_3^{k-1} + \frac{3}{32}\bar{f}_4^{k-1}) \end{array} \right.
$$

$$
p=6 \quad \left\{ \begin{array}{lcl} d_1^k & = & \bar{f}_1^k - (\frac{595}{256}\bar{f}_1^{k-1} - \frac{789}{256}\bar{f}_2^{k-1} + \frac{830}{256}\bar{f}_3^{k-1} - \frac{554}{256}\bar{f}_4^{k-1} + \frac{207}{256}\bar{f}_5^{k-1} - \frac{33}{256}\bar{f}_6^{k-1}) \\ d_2^k & = & \bar{f}_3^k - (\frac{33}{256}\bar{f}_1^{k-1} + \frac{397}{256}\bar{f}_2^{k-1} - \frac{294}{256}\bar{f}_3^{k-1} + \frac{170}{256}\bar{f}_4^{k-1} - \frac{59}{256}\bar{f}_5^{k-1} + \frac{9}{256}\bar{f}_6^{k-1}) \\ d_3^k & = & \bar{f}_5^k - (\frac{-9}{256}\bar{f}_1^{k-1} + \frac{87}{256}\bar{f}_2^{k-1} + \frac{262}{256}\bar{f}_3^{k-1} - \frac{114}{256}\bar{f}_4^{k-1} + \frac{35}{256}\bar{f}_5^{k-1} - \frac{5}{256}\bar{f}_6^{k-1}) \end{array} \right.
$$

Modifications at the right boundary can be obtained by symmetry.

The stability of these algorithms follows from the stability of their associated interpolatory schemes and Theorem 5.4.

The case $p = 2$ corresponds to $r = 3$, i.e. I_k uses third degree polynomial pieces. Figure 3 shows $M^{-1}(u^0, 0, \cdots, 0)$ with $J_0 = 8$, $L = 7$ and $u^0 = \delta_i, i = 1, 2, 3$. The limiting functions for $u^0 = \delta_i, i = 4, 5$ are translated versions of the limiting function for $i = 3$. The limiting functions for $i = 6, 7$ are the symmetric reflections of the $i = 2, 1$ limits with respect to the right boundary. The limiting functions exist but they are not continuous.

The case $p = 4$ corresponds to $r = 5$. Figure 4 displays $M^{-1}(u^0, 0, \cdots, 0)$ with $J_0 = 8$, $L = 7$ and $u^0 = \delta_i, i = 1, 2, 3, 4$. All limiting functions are affected by boundary effects. When $J_0 \geq 16$, we have 4 limiting functions at each boundary altered by boundary effects; the rest are translates of the same function (in fact they are scaled translations of the recursive subdivision limit corresponding to the periodic case). Numerical results for $J_0 = 16$, $L = 7$ and $p = 4$ are displayed in Figure 5.

Under periodicity assumptions, all φ_i^k are translated versions of the same function φ_0^k whose support is $[-ph_0, ph_0]$. 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. These facts can be explained using recursive subdivision theory, and we shall do so in the next section.

In the non-periodic case, we retain the periodic limit whenever the new non-zero refined values at each resolution level can be computed without using one-sided reconstructions. For a reconstruction of order p, there are p limiting functions affected by boundary effects. These functions have compact support. In fact, support $\varphi_i^k = [0, (p+i)h_k]$ for $i = 1, \ldots, p$. Moreover, when increasing the resolution level, these limiting functions are also scaled versions of the same functions in lower levels (see Figures 4 and 5).

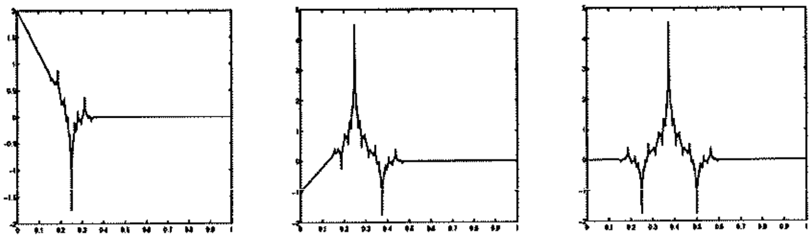

FIG. 3. *Hat-average limiting functions.* $J_0 = 8$, $p = 2$.

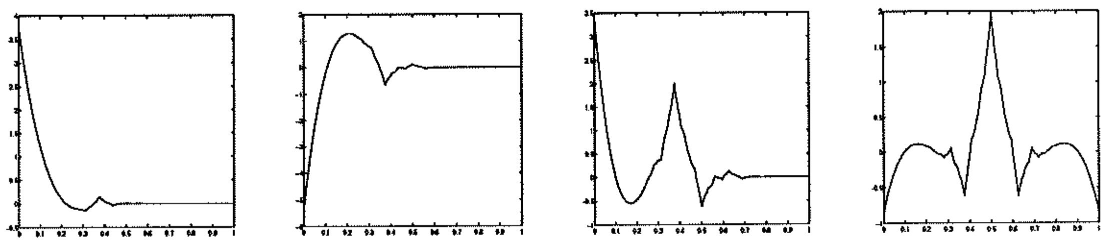

FIG. 4. *Hat-average limiting functions.* $J_0 = 8$, $p = 4$.

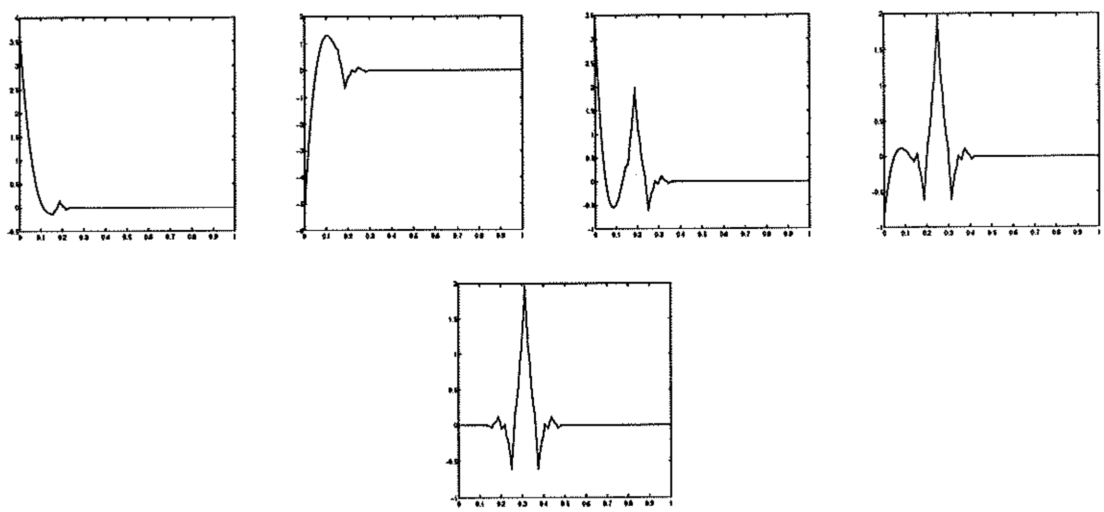

FIG. 5. *Hat-average limiting functions.* $J_0 = 16$, $p = 4$.

Figure 6 displays the relevant limiting functions for $J_0 = 16$, $L = 7$, $p = 6$. All limiting functions have been obtained applying the inverse multiresolution transform to the initial sequence $(\delta_i^0, 0, \ldots, 0)$. We show the limiting functions at the left boundary for $i = 1, \ldots, p$ and the first function, $i = p + 1$, which is not affected by one sided reconstructions at the boundary. Their right boundary counterparts are their mirror images with respect to the right boundary.

In Figures 7, 8, and 9 we show $M^{-1}(0, \delta_i, 0, \ldots, 0)$ for $L = 7$ and J_0 as specified. These are the 'generalized wavelets'. There appear to be $p/2 + 1$ limiting functions affected by each boundary. Outside the influence of the boundary, all limiting functions are translated versions of the same function. This basic 'wavelet' is precisely the one obtained in the biorthogonal framework with the hat function as one of the scaling functions (see next section).

6. Periodicity: The connection with Biorthogonal Wavelets. When dealing with piecewise-polynomial reconstructions based on polynomial interpolation, the periodicity assumption implies that we consider the same choice of stencil for all points and all resolution

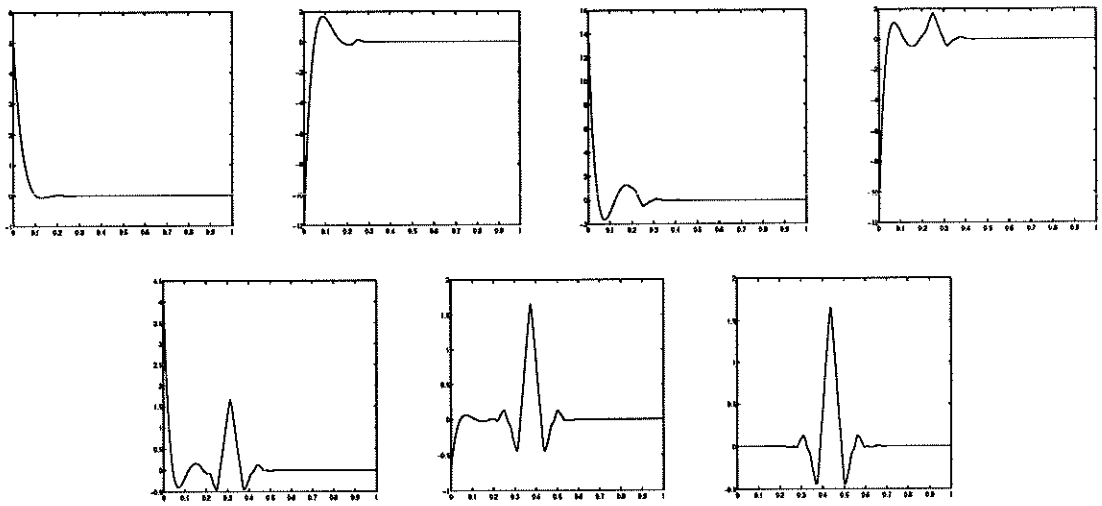

FIG. 6. *Hat-average limiting functions.* $J_0 = 16$, $p = 6$.

levels. This implies that the reconstruction operator is translation invariant and independent of the resolution level. Then we have

THEOREM 6.1. *Let $\mathcal{R}_k$ be a linear operator which is translation invariant*

$$(\mathcal{R}_k \delta^k_{j-q})(x - qh_k) = (\mathcal{R}_k \delta^k_j)(x) \qquad \forall q \in Z \tag{113}$$

and independent of the level of resolution

$$(\mathcal{R}_k \delta^k_0)(x) = (\mathcal{R}_{k-1}\delta^{k-1}_0)(2x) \quad \forall k. \tag{114}$$

Then, the quantities $< \omega^k_l, \mathcal{R}_{k-1}\delta^{k-1}_0 >$ are independent of k, and if we define

$$\gamma_l =< \omega^k_l, \mathcal{R}_{k-1}\delta^{k-1}_0 >, \tag{115}$$

then

$$(P^k_{k-1})_{i,j} = \gamma_{i-2j}. \tag{116}$$

The proof is an easy exercise (see [17]).

Under these premises the predicted values, $\tilde{f}^k = P^k_{k-1}\bar{f}^{k-1}$, can be computed as follows:

$$\begin{array}{rcl} \tilde{f}^k_{2i} & = & \sum_l \gamma_{2l} \bar{f}^{k-1}_{i-l} \\ \tilde{f}^k_{2i+1} & = & \sum_l \gamma_{2l+1} \bar{f}^{k-1}_{i-l}. \end{array} \tag{117}$$

Therefore, the prediction process takes the form of a uniform binary subdivision scheme. The mask of this scheme is given in terms of the prediction operator, which is now independent of k and can, thanks to the periodicity assumption, 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.

As mentioned in Section 4.1.1, there is always a compactly supported continuous function $\varphi(x)$ associated to a convergent subdivision scheme S which is obtained by refining the

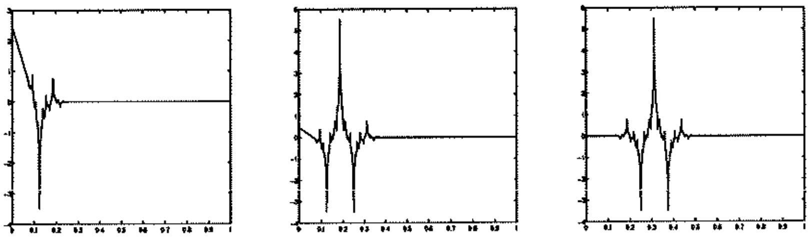

FIG. 7. *Hat-average limiting functions.* $J_0 = 8$, $p = 2$. ψ *functions*

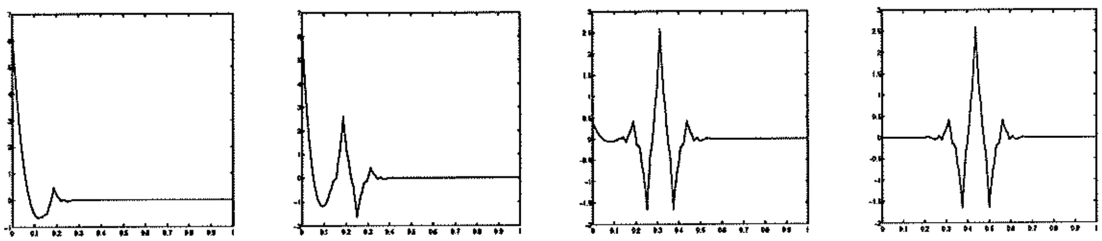

FIG. 8. *Hat-average limiting functions.* $J_0 = 8$, $p = 4$. ψ *functions*

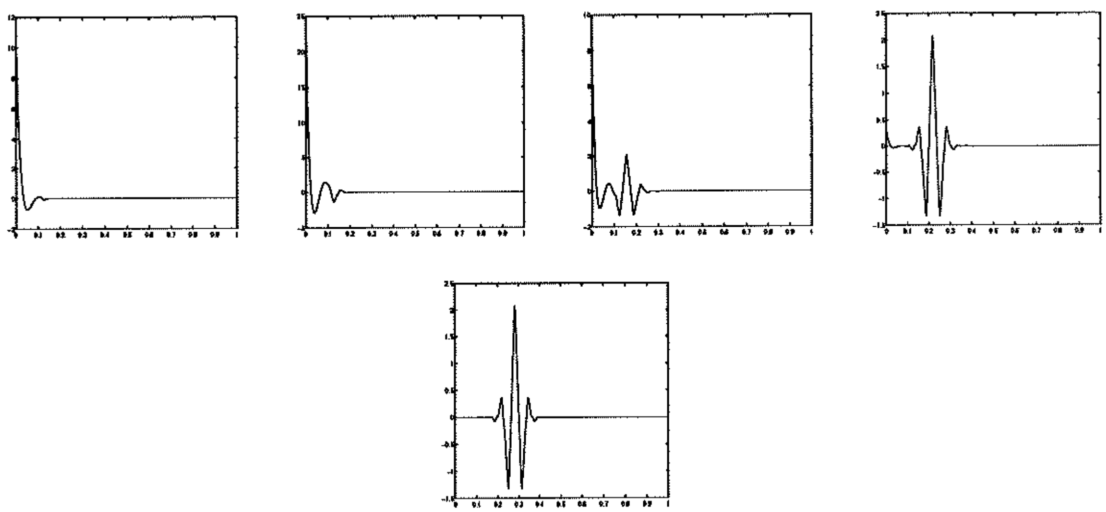

FIG. 9. *Hat-average limiting functions.* $J_0 = 16$, $p = 6$ ψ *functions*

sequence δ_0. The sometimes called 'fundamental' (or S-refinable) function $\varphi(x)$ satisfies the dilation relation

$$\varphi(x) = \sum_j \gamma_j \varphi(2x - j) = \sum_j (P\delta_0)_j \varphi(2x - j). \tag{118}$$

The relation between the limiting processes in the interpolatory and hat-average frameworks stems, in fact, from the relation between the subdivision schemes defined by the prediction operators in each of these frameworks.

Following the notation of last section, let $\tilde{P}^k_{k-1} = \tilde{\mathcal{D}}_k I_{k-1}$ be the prediction operator in the interpolatory framework, while we reserve $P^k_{k-1}$to denote the prediction operator in the hat-average case. Notice that relation (88) can be expressed as follows:

$$(P^k_{k-1}\bar{f}^{k-1})_i = \frac{(\tilde{P}^k_{k-1}H^{k-1})_{i+1} - 2(\tilde{P}^k_{k-1}H^{k-1})_i + (\tilde{P}^k_{k-1}H^{k-1})_{i-1}}{h_k^2},$$

which implies that the subdivision scheme defined by $P_{k-1}^k$is the 'second divided difference' scheme of the interpolatory subdivision scheme defined by $\tilde{P}_{k-1}^k$. In the refinement subdivision terminology, if S is the scheme defined by $\tilde{P}_{k-1}^k$, then the scheme defined by $P_{k-1}^k$is D_2S.

This fact has important consequences. First, S converges uniformly to $\mathcal{C}^\nu$ limit functions if and only if D_2S converges uniformly to $\mathcal{C}^{\nu-2}$ limit functions, provided S is reproduces polynomials of degree less than or equal to ν, or, in our terminology, provided the interpolation is exact for polynomials of degree up to ν (see e.g. [7] or [14] for details and generalizations). Moreover, if $\nu \geq 2$, the limit function for D_2S is the second derivative of the limit function for S.

In our notation these facts can be expressed as follows:

$$(119) \qquad H(x) = \lim_{L\to\infty} \tilde{A}_0^L H^0 \in \mathcal{C}^2, \qquad \Rightarrow \qquad f(x) = \lim_{L\to\infty} A_0^L \bar{f}^0 \in \mathcal{C}$$

$$(120) \qquad f(x) = H''(x),$$

where $\bar{f}^0$ is the set of second divided differences of H^0.

A convergent stationary subdivision scheme is completely determined by its fundamental function. It is shown in [14] that the relation between the fundamental function associated to D_2S, which we denote by $\varphi(x)$, and the fundamental function associated to S, which we denote by $\tilde{\varphi}(x)$, is as follows:

$$\tilde{\varphi}(x) = \int_{x-1}^{x}\int_{y}^{y+1} \varphi(z)dzdy \qquad \text{hence,} \qquad \tilde{\varphi}''(x) = \varphi(x+1) - 2\varphi(x) + \varphi(x-1)$$

The fundamental function of the scheme satisfies the dilation relation (118). Using (109) it is easy to deduce that the fundamental function associated to the hat average prediction operators defined in Section 5.1 (under periodicity assumptions) satisfies

$$(121) \qquad \begin{aligned} \varphi(x) &= (2-\beta_1)\varphi(2x) - \sum_{l=1}^{s-1} \frac{\beta_l + \beta_{l+1}}{2} [\varphi(2x-2l) + \varphi(2x+2l)] \\ &+ \sum_{l=1}^{s-1} \beta_l [\varphi(2x-2l+1) + \varphi(2x+2l-1)] \end{aligned}$$

The coefficients β_l are those defined in (111).

For $p = 2$ $(r = 3)$, $\tilde{\varphi}(x)$ is differentiable and $\tilde{\varphi}'(x)$ is Hölder continuous with exponent $1 - \epsilon$, with ϵ arbitrarily small. $\tilde{\varphi}''(x)$ does not exist at any dyadic rational (see [11] or [12]), thus the limit of the binary subdivision scheme D_2S cannot be continuous. In fact the solution of the dilation relation (121) in this case is an L^2 function (see [8]) and D_2S converges weakly in L_2 (see [7]). Observe that $\mathcal{R}_k^\infty$ is nevertheless well defined, which accounts for the stability of the associated multiresolution scheme.

For $p = 4$ $(r = 5)$, $\tilde{\varphi}(x)$ is a $\mathcal{C}^2$ function, thus D_2S converges uniformly and $\mathcal{R}_k^\infty(x, \bar{f}^k)$ is a continuous function. For the case $p = 6$ $(r = 7)$, our numerical results indicate that the second divided difference scheme converges to a continuous function (which seems to be smoother than in the $p = 4$ case), and thus the original interpolatory subdivision scheme must converge uniformly to a function which is at least $\mathcal{C}^2$. Convergence can be proven analytically using the results of [11].

Convergence of the interpolatory subdivision scheme corresponding to $\hat{P}_{k-1}^k$ implies existence of the hierarchical interpolation $I_k^\infty(x; H^k)$. In turn, theorem 5.4 implies the existence of $\mathcal{R}_k^\infty$ and the stability of the hat-average multiresolution scheme corresponding to P_{k-1}^k. We do remark that stability follows solely from the existence of $\mathcal{R}_k^\infty$; however, the connection with stationary subdivision provides also the *form* of the hierarchical reconstruction.

We saw in Section 3 that

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

with $\varphi_i^k := \mathcal{R}_k^\infty(x;\delta_i^k)$. Using the fundamental function of the scheme we get

$$\varphi_i^k := \mathcal{R}_k^\infty(x;\delta_i^k) = \varphi(\frac{x}{h_k} - i)$$

i.e.

$$\mathcal{R}_k^\infty(x;\bar{f}^k) = \sum_i \bar{f}_i^k \varphi(\frac{x}{h_k} - i). \tag{122}$$

The multiresolution schemes which correspond to the bases of biorthogonal wavelets in [8] can be cast as particular examples of Harten's framework (see [18, 19]). In [8] the decomposition and reconstruction filters are constructed using *two* scaling functions satisfying two different dilation relations (which have to satisfy a biorthogonality condition). In terms of Harten's framework, the discretization process used to define the scheme is as in (32), $\omega(x)$ being one of the scaling functions, while the reconstruction procedure is defined by the right hand side of (122), $\varphi(x)$ being the other scaling function. If

$$\omega(x) = 2\sum_l \alpha_l \omega(2x-l), \qquad \varphi(x) = \sum_l \gamma_l \varphi(2x-l)$$

the biorthogonality condition is equivalent to

$$\sum_l \alpha_l \gamma_{l+2m} = \delta_{m,0} \tag{123}$$

The prediction operator P_{k-1}^k is given by

$$(P_{k-1}^k v^k)_i = \sum_m \gamma_{i-2m} v_m^k.$$

The hat function and the functions satisfying (121) for $p = 2,4,6$ are the two scaling functions in one of the families of biorthogonal wavelets (the *spline* examples, see [8] or [9]).

The observations above imply that the biorthogonal algorithms derived from the pairs of scaling functions ω and φ are in fact the hierarchical form of the centered-stencil piecewise-polynomial reconstruction in Harten's framework. The prediction operators in these two schemes are the same (see also Theorem 3.3), thus both schemes are equivalent. Since periodicity is an essential ingredient of the formulation, it is clear that these biorthogonal schemes will have poor approximation properties for non-periodic data.

We proved in Section 4 that the sets $\{\omega_i^k\}$ and $\{\varphi_i^k\}$ are biorthogonal, i.e.

$$< \omega_i^k, \varphi_l^k >= \delta_{il}, \qquad \forall i, \forall l.$$

In terms of the coefficients of the decimation and prediction matrices, the last relation can be written as (123).

Notice that the biorthogonality conditions are never *imposed* in Harten's framework. They come as an automatic consequence of the relation $\mathcal{D}_k\mathcal{R}_k = I$.

Notice also that (26), which in this case reads

$$\psi_j^k = \sum_i (E_k)_{ij} \varphi_i^k = \varphi_{2j-1}^k - \frac{1}{2}(\varphi_{2j}^k - \varphi_{2j-2}^k),$$

leads to

$$\psi_j^k = \psi(\frac{x}{2h_k} - j),$$

where $\psi(x)$ satisfies

$$\psi(x) = -\frac{1}{2}\varphi(2x) + \varphi(2x+1) - \frac{1}{2}\varphi(2x+2).$$

7. 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 multi-scale 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 [17, 18] 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 [8].

Harten's framework allows also to consider nonlinear reconstruction operators. We consider the multiresolution schemes derived from these in Part II [3].

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

REFERENCES

[1] R. Abgrall, A. Harten *Multiresolution Representation in Structured Meshes. I Preliminary Report*, UCLA CAM Report 94-20 (1994)

[2] F.Aràndiga, R. Donat and A. Harten, *Multiresolution Based on Weighted Averages of the Hat Function*, UCLA CAM Report 93-34 (1993)

[3] F.Aràndiga, R. Donat and A. Harten, *Multiresolution Based on Weighted Averages of the Hat Function II: Nonlinear Reconstruction Operators*, UCLA CAM Report 96-26

[4] F. Aràndiga and V. Candela, *Multiresolution Standard Form of a Matrix*, SIAM J. Numer. Anal., **33** (1996), pp. 417-434; also UCLA CAM Report 92-37 (1992)

[5] 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; also UCLA CAM Report 92-52 (1992)

[6] G. Beylkin, R. Coifman and V. Rokhlin, *Fast Wavelet Transform and Numerical Algorithms I*, Comm. Pure Appl. Math., **XLIV** (1991), pp. 141-183.

[7] A. Cavaretta, W. Dahmen and C. Micchelli *Stationary Subdivision*, AMS Memoirs, **453** (1991)

[8] A. Cohen, I. Daubechies and J.C. Feauveau, *Biorthogonal Bases of Compactly Supported Wavelets*, Comm. Pure Applied Math., **45** (1992), pp. 485-560.

[9] I. Daubechies *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics (1992).

[10] I. Daubechies and J. Lagarias *Two Scale Difference Equations I*, SIAM J. Math. Anal., **22** (1991), pp. 1388-1410.

[11] I. Daubechies and J. Lagarias *Two Scale Difference Equations II*, SIAM J. Math. Anal., **23** (1992), pp. 1031-1079.

[12] G. Deslauriers and S. Duboc, *Symmetric Iterative Interpolation Scheme*, Constr. Approx., **5** (1989), pp. 49-68.

[13] R. Donat and A. Harten, *Data Compression Algorithms for Locally Oscilatory Data*, UCLA CAM Report 93-26 (1993)

[14] N. Dyn, J.A. Gregory and D. Levin *Analysis of Linear Binary Subdivision Schemes for Curve Design*, Constr. Approx., **7** (1991), pp. 127-147.

[15] B. Engquist, S. Osher and S. Zhong, *Fast Wavelet Algorithms for Linear Evolution Equations*, SIAM J. Sci. Comput., **15** (1994), pp. 755-775; also ICASE Report 92-14 (1992)

[16] A. Harten, *Discrete Multiresolution Analysis and Generalized Wavelets*, J. Appl. Numer. Math., **12** (1993), pp. 153-192; also UCLA CAM Report 92-08 (1992)

[17] A. Harten, *Multiresolution Representation of Data*, UCLA CAM Report 93-13 (1993)

[18] A. Harten, *Multiresolution Representation of Data II: General Framework*, UCLA CAM Report 94-10 (1994); SIAM J. Numer. Anal. **33** #3 pp 1205-1256 (1996)

[19] A. Harten, *Multiresolution Representation of Cell-Averaged Data*, UCLA CAM Report 94-21 (1994)

[20] A. Harten, *Multiresolution Algorithms for the Numerical Solution of Hyperbolic Conservation Laws*, UCLA CAM Report 93-03 (1993); Comm. Pure Appl. Math.**48** #12 pp 1305-1342 (1995)

[21] A. Harten and I. Yad-Shalom *Fast Multiresolution Algorithms for Matrix-Vector Multiplication*, SIAM J. Numer. Anal., **31** (1994), pp. 1191-1218; also UCLA CAM Report 92-31 (1992)

[22] C. A. Micchelli, *Interpolatory Subdivision Schemes and Wavelets* J. Approximation Theory **86**, pp. 41-71 (1996)

[23] H. Yserentant, *On the multi-level splitting of finite element spaces* , Numer. Math., **49** (1986), pp. 379-412.