UCLA

COMPUTATIONAL AND APPLIED MATHEMATICS

Multiresolution Representation of Data.

II. General Framework

Ami Harten

April 1994

CAM Report 94-10

Multiresolution Representation of Data.
II. General Framework

Ami Harten

School of Mathematical Sciences
Tel-Aviv University

and

Department of Mathematics
UCLA

Abstract

In this paper we present a general framework for multiresolution representation of data which is obtained by discretization of mappings. This framework , which can be viewed as a generalization of the theory of wavelets, includes discretizations corresponding to unstructured grids in several space dimensions, and thus is general enough to enable us to embed most numerical problems in a multiresolution setting. Furthermore, this framework allows for nonlinear (data-dependent) multiresolution representation schemes and thus enables us to design adaptive data-compression algorithms.

In this paper we also study the stability of linear schemes for multiresolution representation and derive sufficient conditions for existence of a multiresolution basis.

1. Introduction and Overview

The purpose of this paper is to integrate ideas from three different fields: (1)Theory of wavelets; (2)numerical solution of PDEs, and (3)subdivision schemes, in order to formulate a general framework for multiresolution of data.

Our starting point is high-order Godunov-type schemes for the numerical solution of hyperbolic conservation laws, from which we take the notion of reconstructible discretization: We consider a discretization $\mathcal{D}$ which assigns discrete values $v = \{v_i\}$ to a function $f \in \mathcal{F}$, and is reconstructible in the sense that it has a right inverse $\mathcal{R}$, i.e. $\mathcal{D}(\mathcal{R}v) = v$, where $\mathcal{R}v$ is an approximation to f for which $v = \mathcal{D}f$. In the context of Godunov-type schemes we use discretization by cell-averages and say that $\mathcal{R}v$ is a *conservative* reconstruction of f (see e.g. [HEOC]).

From multigrid methods (which typically use discretization by pointvalues and

1

reconstruction by interpolation) we get the idea that if we consider a sequence of grids with corresponding discretizations $\{\mathcal{D}_k\}$ and reconstructions $\{\mathcal{R}_k\}$, then the most natural way to go from the k-th grid to the coarser $(k-1)$-th grid is by the operator $D_k^{k-1} = \mathcal{D}_{k-1}\mathcal{R}_k$, and similarly to use the operator $P_{k-1}^k = \mathcal{D}_k\mathcal{R}_{k-1}$ in order to go from the $(k-1)$-th grid to the finer k-th grid.

Next we consider the notion of nested discretization in a more abstract setting and observe that the operators D_k^{k-1} and P_{k-1}^k can serve respectively as decimation and prediction in a pyramid scheme of the type that is used in signal processing . Using ideas from the theory of wavelets we remove the redundancy that is typical to frames which are obtained by pyramid schemes and get a multiresolution *representation* (tight frame). Furthermore, we use knowledge from the theory of wavelets in order to relate the discrete multiresolution representation to a multiresolution basis in the space of functions $\mathcal{F}$ (see e.g [Me], [Ma] and [Da]). We show that the construction of wavelets can be formulated in terms of discretization and reconstuction (corresponding to a nested dyadic sequence of uniform grids), and observe that it also includes discretizations which are different from the ones that are traditionally used in the numerical solution of PDEs.

From Finite Element Methods we borrow the notion of hierarchial bases (see e.g. [Y1] and [Y2]), and show that if the sequence of approximation is hierarchial then the discrete multiresolution representation (MR) scheme is stable and it corresponds to a multiresolution basis of $\mathcal{F}$.

From subdivision schemes we adopt the concept of "hierarchial form": We show that if the subdivision sequence $\{\Pi_k^L f\}_{L=k}^\infty$, where

$$\Pi_k^L = (\mathcal{R}_L\mathcal{D}_L)\cdots(\mathcal{R}_k\mathcal{D}_k)$$

is convergent, then the reconstruction $\mathcal{R}_k^H$ which is defined by

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

is hierarchial.

Our notion of stability is taken from the theory of finite-difference methods for initial value problems (see e.g. [RM]), and its role is to prevent unbounded growth of initial perturbations by repeated applications of an operator. We show that nestedness of the sequence of discretization implies stability of the direct MR transform with respect to repeated decimation, and that convergence of the subdivision scheme implies stability of the inverse MR transform with respect to repeated prediction. Extending the analysis of subdivision schemes to our general framework

2

we derive a sufficient condition for convergence of a subdivision sequence which also implies the stability of the corresponding discrete MR scheme and the existence of a multiresolution basis in the space $\mathcal{F}$.

What is basically new in this paper is the design of MR schemes from a nested sequence of discretization. Other results which are "new" in one field may be well-known in others. The usefulness of our general framework is that it enables us to transfer knowledge from one field to the others.

Its contribution to the theory of wavelets is in its extension to nonuniform grids, and in the treatment of boundaries which is adopted from the numerical solution of PDEs. We also show how to design adaptive (data-dependent) MR schemes - this is borrowed from the numerical solution of hyperbolic conservation laws where this need arises because of the presence of shock waves in the solution.

Its contribution to the numerical solution of PDEs is in providing capabilities for regularity analysis and data compression in unstructured meshes, which have not been fully explored. In [HY] we show that the multilevel matrix multiplication algorithm of [BL] which is designed from a multigrid point of view can be related to data-compression of the matrix, and thus can be fitted into the framework which was suggested in [BCR] from the point of view of wavelets. In [H5] and [H6] we show how to apply the discrete MR algorithm to hyperbolic systems of conservation laws in order to reduce the number of numerical flux computations. This approach is related to adaptive grids and may serve as a more convenient implementation of the same idea.

Its contribution to subdivision schemes is in providing a more general formulation of the "two scale difference equations" that are studied there, and by suggesting a sufficient condition for convergence of these more general subdivision sequences.

Many of the results of this paper are obtained by transfer of existing knowledge from one field to another. Because of the broad nature of this general framework we cannot do justice to the many works which constitute the infrastructure of this existing knowledge. What we do is give some references which may serve as a point of entry for the interested reader.

In the following we present a rather lengthy overview and some examples in order to motivate the abstract derivation of our framework, and in particular to show its relation to the theory of wavelets in [Da], [CDF] and [CDV].

This paper is a sequel to [H7], in which many specific examples were described.

1.A. Multiresolution Representation of Sequences

In this subsection we describe multiresolution schemes for representation of sequences of real numbers. Let $\hat{D}_k^{k-1}$ and $\hat{P}_{k-1}^k$ be a pair of operators which satisfy

(1.1a) $\hat{D}_k^{k-1} : S^k \to S^{k-1}, \quad \hat{D}_k^{k-1}$ a linear operator

(1.1b) $\hat{P}_{k-1}^k : S^{k-1} \to S^k$

(1.1c) $\hat{D}_k^{k-1} \hat{P}_{k-1}^k = \hat{I}_{k-1}, \quad \hat{I}_{k-1} =$ identity operator in S^{k-1},

and observe that (1.1c) implies that $\hat{D}_k^{k-1}$ maps S^k *onto* S^{k-1}, i.e.

$$(1.2) \qquad S^{k-1} = \hat{D}_k^{k-1}(S^k);$$

we refer to $\hat{D}_k^{k-1}$ as the decimation operator. Also observe that $\hat{P}_{k-1}^k$ is a right-inverse of $\hat{D}_k^{k-1}$ in S^{k-1}, and that unlike $\hat{D}_k^{k-1}$, it is not required to be a linear operator; we refer to $\hat{P}_{k-1}^k$ as the prediction operator.

In the following we describe both the finite-dimensional case and the infinite one with the same notation. To do so we use the notation agreement that in the finite-dimensional case

$$(1.3a) \qquad S^k = \{s | s = \{s_i\}_{i=1}^{J_k}, \ s_i \in \mathbf{R}\},$$

while in the infinite case

$$(1.3b) \qquad S^k = S = \{s | s = \{s_i\}_{i=-\infty}^{\infty}, \ s_i \in \mathbf{R}\}, \quad \text{for all} \ \ k.$$

Given a sequence s, finite or infinite, which we associate with the L-th level of resolution for some $L > 1$, say $s = s^L$, we generate $\{s^k\}_{k=0}^{L-1}$, $s^k \in S^k$, by successive decimation

$$(1.4) \qquad s^{k-1} = \hat{D}_k^{k-1} s^k, \qquad k = L, \ldots, 1.$$

For each level of resolution k, we use $\hat{P}_{k-1}^k$ to predict s^k from knowledge of s^{k-1} by

$$(1.5a) \qquad \tilde{s}^k = \hat{P}_{k-1}^k s^{k-1},$$

and observe that the prediction error e^k,

$$(1.5b) \qquad e^k = s^k - \tilde{s}^k,$$

satisfies

$$\hat{D}_k^{k-1} e^k = \hat{D}_k^{k-1} s^k - \hat{D}_k^{k-1} \hat{P}_{k-1}^k s^{k-1} = s^{k-1} - s^{k-1} = 0$$

4

i.e., it is in the null space of $\hat{D}_k^{k-1}$,

$$(1.5c) \qquad e^k \in \mathcal{N}(\hat{D}_k^{k-1}) = \{s | s \in S^k, \ \hat{D}_k^{k-1}s = 0\}.$$

We define $d^k = \{d_j^k\}$, the k-th scale coefficients, to be the coordinates of the prediction error e^k in some basis of $\mathcal{N}(\hat{D}_k^{k-1})$, and observe that since s^k can be recovered from knowledge of s^{k-1} and e^k , we also have

$$(1.6a) \qquad s^k \xleftrightarrow{1:1} \{s^{k-1}, d^k\},$$

which, when applied repeatedly, implies

$$(1.6b) \qquad s^L \xleftrightarrow{1:1} M(s^L) = \{s^0, d^1, \ldots, d^L\}.$$

We refer to $M(s^L)$ as the multiresolution representation of s^L.

Note that in the finite-dimensional case, (1.2) implies that

$$(1.7a) \qquad \dim \mathcal{N}(\hat{D}_k^{k-1}) = J_k - J_{k-1},$$

therefore the number of components in d^k is $(J_k - J_{k-1})$ and consequently the number of components in $M(s^L)$ is

$$(1.7b) \qquad J_0 + \sum_{k=1}^{L}(J_k - J_{k-1}) = J_L.$$

In section 2 we shall describe multiresolution representation (MR) schemes in the more general framework of linear spaces, and discuss the stability of such schemes.

1.B. Relation to Biorthogonal Wavelets

The framework (1.1)-(1.6) for the design of multiresolution representation (MR) schemes was developed in [H4] and [H7], and it generalizes the class of subband coding schemes which correspond to the bases of biorthogonal wavelets in [CDF]. We refer the reader to this paper for historical review of the development of this subject, as well as for description of recent related works.

To obtain the wavelet schemes of [CDF] from our framework we take in (1.1a)

$$(1.8) \qquad (\hat{D}_k^{k-1}s^k)_i = \sum_m \alpha_{m-2i}s_m^k = \sum_\ell \alpha_\ell s_{2i+\ell}^k,$$

and in (1.1b)

$$(1.9a) \qquad (\hat{P}_{k-1}^k s^{k-1})_i = \sum_m \beta_{i-2m} s_m^{k-1},$$

which can also be written as

$$(1.9b) \qquad \begin{cases} (\hat{P}_{k-1}^k s^{k-1})_{2i-1} = \sum_\ell \beta_{2\ell-1} s_{i-\ell}^{k-1}, \\ (\hat{P}_{k-1}^k s^{k-1})_{2i} = \sum_\ell \beta_{2\ell} s_{i-\ell}^{k-1}. \end{cases}$$

Here $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ are sequences of compact support which satisfy

$$(1.10a) \qquad \sum_\ell \alpha_\ell \beta_{\ell+2m} = \delta_{m,0},$$

$$(1.10b) \qquad \sum_\ell \alpha_\ell = 1, \quad \sum_\ell \beta_{2\ell-1} = \sum_\ell \beta_{2\ell} = 1.$$

Relation (1.10a), which is the biorthogonality condition in [CDF], is obtained from (1.1c), while (1.10b) are consistency relations which are obtained from the requirement that any constant sequence will remain unchanged under decimation and prediction. In [CDF] additional conditions on $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ are derived from requirements of regularity of the associated wavelet functions; we shall describe these conditions in subsections 1.G and 2.E.

1.C. Predictability and Scale-Decomposition

In the present paper, as in the previous ones [H4] and [H7], our main concern is the "quality" of the prediction $\hat{P}_{k-1}^k$: The notion of "k-th scale" is related to the information in s^k which cannot be predicted from knowledge of s^{k-1} (1.4) by *any* prediction scheme. When using a *particular* prediction scheme, the prediction error e^k (1.5b), and consequently the k-th scale coefficients d^k, includes, in addition to the "true" k-th scale, also a component of approximation error which is related to the "quality", or "accuracy", of the particular prediction scheme.

We recall that the MR scheme applies to any sequence $s = s^L$ of real numbers. But these numbers could have been generated by some stochastic process, or by some iterated function system (IFS), or by a numerical scheme for the solution of a PDE, or by who knows what? Thus the question of "quality", when posed in general, is not meaningful; it can be made meaningful only by limiting our interest to a subset of data, where we know something about the way it was generated.

1.D. From Prediction to Approximation

In this paper we consider subsets of data which are obtained from the following process of discretization: Let $\mathcal{F}$ be a linear space of mappings

$$(1.11a) \qquad\qquad \mathcal{F} \subset \{f | f : X \to Y\}$$

where X and Y are linear spaces, and let $\mathcal{D}_k$ be a linear operator which assigns to any $f \in \mathcal{F}$ a sequence $v^k = \mathcal{D}_k f$, $v^k = \{v_i^k\}$, $v_i^k \in Y$; this sequence corresponds to some k-th level discretization of X (see subsection 1.F and section 5). Let V^k denote the range of $\mathcal{D}_k$ and let $\{\eta_i^k\}$ denote any basis of this space,

$$(1.11b) \qquad\qquad \mathcal{D}_k(\mathcal{F}) = V^k = \text{span}\{\eta_i^k\};$$

we denote the coordinates of $v^k \in V^k$ in this basis by the sequence $\hat{v}^k = \{\hat{v}_i^k\}$, $\hat{v}^k \in S^k$, i.e.

$$(1.11c) \qquad\qquad v^k = \sum_i \hat{v}_i^k \eta_i^k.$$

We refer to such $\mathcal{D}_k$ as a discretization operator.

Our goal is to design an MR scheme (1.6) 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 (1.11).

Since $\mathcal{D}_k$ maps $\mathcal{F}$ *onto* V^k (1.11b), it follows that for any v^k in V^k there is at least one f in $\mathcal{F}$ such that $\mathcal{D}_k f = v^k$. We refer to such an assignment of $f \in \mathcal{F}$ to $v^k \in V^k$ as reconstruction, and denote the reconstruction operator by $\mathcal{R}_k$,

$$(1.12) \qquad\qquad \mathcal{R}_k : V^k \longrightarrow \mathcal{F}, \quad \mathcal{D}_k \mathcal{R}_k = I_k,$$

where I_k is the identity operator in V^k, i.e. $\mathcal{R}_k$ is a right-inverse of $\mathcal{D}_k$ in V^k.

Given a sequence of discretization $\{\mathcal{D}_k\}$ and any sequence of corresponding reconstruction operators $\{\mathcal{R}_k\}$, linear or not, we define the operators D_k^{k-1} and P_{k-1}^k by

$$(1.13a) \qquad\qquad D_k^{k-1} = \mathcal{D}_{k-1} \mathcal{R}_k : V^k \longrightarrow V^{k-1},$$

$$(1.13b) \qquad\qquad P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1} : V^{k-1} \longrightarrow V^k.$$

In section 3 we show that if $\{\mathcal{D}_k\}$ is *nested* (Definition 3.1) then

$$(1.13c) \qquad\qquad D_k^{k-1} P_{k-1}^k = I_{k-1},$$

7

and that the operators so constructed have the following properties:

(i) For any $f \in \mathcal{F}$

$$(1.14) \qquad\qquad D_k^{k-1}(\mathcal{D}_k f) = \mathcal{D}_{k-1} f;$$

(ii) For any $p \in \mathcal{F}$ for which the reconstruction $\mathcal{R}_{k-1}$ is exact,

$$(1.15a) \qquad\qquad \mathcal{R}_{k-1}(\mathcal{D}_{k-1} p) = p$$

we have likewise

$$(1.15b) \qquad\qquad P_{k-1}^k(\mathcal{D}_{k-1} p) = \mathcal{D}_k p,$$

i.e. the prediction P_{k-1}^k is also exact.

Let us consider now any $v^L \in V^L$, then there is $f \in \mathcal{F}$ such that

$$(1.16a) \qquad\qquad v^L = \mathcal{D}_L f,$$

and it follows from (1.14) that the process of successive decimation in (1.4)

$$(1.16b) \qquad\qquad v^{k-1} = D_k^{k-1} v^k, \quad k = L, \ldots, 1$$

yields for all k

$$(1.16c) \qquad\qquad v^k = \mathcal{D}_k f.$$

Thus the problem of prediction which is associated with the corresponding MR scheme, can now be stated as a problem of *approximation*:

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

The quality of this prediction can be judged by the class of $p \in \mathcal{F}$ for which it is exact; by (1.15) this includes the class for which the reconstruction is exact.

The above analysis shows that finding a suitable prediction for our MR scheme can be formulated as a typical problem in approximation theory, and if we solve it well then we have also accomplished our stated goal: Let $\hat{D}_{k-1}^k$ and $\hat{P}_{k-1}^k$ be the representation of D_k^{k-1} and P_{k-1}^k in the bases (1.11b) (see section 2), then (1.1c) follows from (1.13c) and the resulting MR scheme is applicable for all sequences $s^L \in S^L$; it is particularly adequate for data of the form $s^L = \hat{v}^L$, where $\hat{v}^L$ is the representation (1.11c) of $v^L = \mathcal{D}_L f$ with $f \in \mathcal{F}$.

8

1.E. Multiresolution Bases.

In section 2 we show that V^L has a multiresolution basis

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

and that any $v^L \in V^L$ can be written as

$$(1.17) \qquad v^L = \sum_i \hat{v}_i^0 \bar{\varphi}_i^{0,L} + \sum_{k=1}^L \sum_j d_j^k \bar{\psi}_j^{k,L},$$

where $\{d_j^k\}$ are the k-th scale coefficients of the associated multiresolution representation (1.6b), and $\{\hat{v}_i^0\}$ is defined by (1.11c) with $k = 0$. In section 3 we show that if $\{\mathcal{D}_k\}$ is a *nested* sequence of discretization, and $\{\mathcal{R}_k\}$ is any corresponding sequence of *linear* reconstruction operators, then taking in (1.17) $v^L = \mathcal{D}_L f$ and applying R_L to it we get

$$(1.18a) \qquad \mathcal{R}_L \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^{0,L} + \sum_{k=1}^L \sum_j d_j^k \psi_j^{k,L},$$

where

$$(1.18b) \qquad \varphi_i^{0,L} = R_L \bar{\varphi}_i^{0,L} \in \mathcal{F} \quad , \quad \psi_j^{k,L} = R_L \bar{\psi}_j^{k,L} \in \mathcal{F},$$

and

$$\mathcal{D}_0 f = \sum_i \hat{f}_i^0 \eta_i^0.$$

At this point we assume $\mathcal{F}$ to be a Banach space and formulate conditions for stability of the corresponding MR scheme.

In section 4 we present sufficient conditions which ensure that the limiting process $L \to \infty$ in (1.18) yields a multiresolution basis for $\mathcal{F}$. These conditions also imply the stability of the associated MR scheme.

In [Da] and [CDF] a similar limiting process is applied to the discrete MR scheme (1.17) with (1.8)-(1.10), except that $R_L v^L$ in (1.18a) is replaced by the piecewise-constant function

$$(1.19) \qquad \mathcal{R}_L v^L = \sum_i v_i^L \chi_{c_i^L}(x), \quad c_i^L = [x_{i-1}^L, x_i^L],$$

where $\chi_c(x)$ is the characteristic function of the set c (1.34). It is shown there that, under certain conditions on $\{\alpha_\ell\}$ and $\{\beta_\ell\}$, this limiting process results in a

9

multiresolution basis of wavelets. In this sense we referred in [H4] [H7] to the limit functions of (1.18b) in the general case as "generalized wavelets". This involves a certain degree of poetic freedom, since for unstructured meshes these limit functions are far from being translates and dilates of a single wavelet function.

We remark that in the case of discretization by pointvalue, the limiting process of taking $L \to \infty$ in (1.18b) is closely related to subdivision schemes (see [CD], [CDM], [DD], [DaL1], [DaL2] and [DGL]) which are used in Computer-Aided Design (CAD) to add points to a set of given ones, so that the added points lie on some relatively smooth surface that interpolates the given set of points.

1.F. Examples.

In this subsection we present some examples in order to demonstrate the various considerations in choosing discretization and reconstruction.

Example 1.1. Pointvalue discretization

Consider the case

$$(1.20) \qquad \mathcal{F} \subset \{f | f : X \subset \mathbf{R}^m \longrightarrow \mathbf{R}^n\},$$

take any sequence, finite or infinite,

$$(1.21a) \qquad X^k = \{x_i^k\}, \ x_i^k \in X,$$

and define $v^k = \mathcal{D}_k f$ by

$$(1.21b) \qquad v_i^k = (\mathcal{D}_k f)_i = f(x_i^k), \quad v^k = \{v_i^k\}, \quad v_i^k \in \mathbf{R}^n.$$

We refer to (1.21) as discretization by pointvalue. Each element v_i^k in (1.21b) is a vector in $\mathbf{R}^n$

$$(1.22a) \qquad v_i^k = (v_{i,1}^k, \dots, v_{i,n}^k).$$

We can represent $v^k = \{v_i^k\}$ as a sequence $\hat{v}^k$ in the space of sequences S^k by (1.11c), e.g.

$$(1.22b) \qquad \hat{v}^k = \{\cdots, \ \{v_{i-1,\ell}^k\}_{\ell=1}^n, \ \{v_{i,\ell}^k\}_{\ell=1}^n, \ \{v_{i+1,\ell}^k\}_{\ell=1}^n, \cdots\},$$

which corresponds to a particular choice of basis $\{\eta_i^k\}$.

The sequence of discretization $\{\mathcal{D}_k\}$ in (1.21) is nested if for all k

$$(1.23a) \qquad X^{k-1} \subset X^k;$$

10

decimation in this case amounts to removing from v^k components $v_i^k = f(x_i^k)$ for $x_i^k \notin X^{k-1}$. Note that the decimation operator D_k^{k-1} is defined directly from the sequence $\{\mathcal{D}_k\}$, and we do not have to use the formulation (1.13a); in section 3 we show that this is true for any *nested* sequence of discretization.

Let $I^k(x; v^k)$ denote *any* interpolation of $\{v_i^k\}$ at the corresponding nodes $\{x_i^k\}$, i.e.

$$(1.23b) \qquad I^k(x_i^k; v^k) = v_i^k \qquad for \ \ all \ \ x_i^k \in X^k,$$

and observe that

$$(1.24a) \qquad \mathcal{D}_k I^k(\cdot; v^k) = v^k.$$

The above relation shows that reconstruction (1.12) in this case amounts to a selection of an interpolation technique in (1.23b). Given v^{k-1} we use (1.13b) to approximate v^k by $\tilde{v}^k$, i.e.

$$(1.24b) \qquad \tilde{v}_i^k = (P_{k-1}^k v^k)_i = I^{k-1}(x_i^k; v^{k-1}) \qquad for \ \ all \ \ x_i^k \in X^k;$$

using multigrid terminology this can be expressed by saying that we use injection of the values corresponding to x_i^k which are in X^{k-1}, and interpolation for those which are not in X^{k-1}. Observe that since the prediction error vanishes at X^{k-1}, we define the scale-coefficients $d^k = \{d_j^k\}$ by

$$(1.24c) \qquad d_j^k = e_{i_j}^k = v_{i_j}^k - \tilde{v}_{i_j}^k, \ \ for \ \ all \ \ x_{i_j}^k \notin X^{k-1};$$

in the finite case $j = 1, \ldots, (J_k - J_{k-1})$ and $\{x_{i_j}^k\}_{j=1}^{J_k - J_{k-1}} = X^k - X^{k-1}$.

In subsection 5.A we describe MR schemes for pointvalue discretization in a triangulated mesh in $\mathbf{R}^2$. In the following we take $m = n = 1$, $X = [0, 1]$ in (1.20) and restrict the prediction problem to finite sequences of real numbers which are obtained from scalar functions by pointvalue discretization at grid-points $X^k = \{x_i^k\}_{i=1}^{J_k}$ in $[0, 1]$

$$(1.25a) \qquad 0 \le x_1^k < x_2^k < \cdots < x_{J_k}^k \le 1.$$

Example 1.1.1 *Piecewise-Polynomial Interpolation*

We assume that the data is obtained from sampling of a continuous function and describe a general piecewise-polynomial interpolation of the pointvalue data. Let $\mathcal{S}_i^k$ denote a stencil of r consecutive points of X^k which includes x_i^k and x_{i+1}^k

11

and let $p_i^k(x; v^k)$ denote the unique polynomial of degree $(r-1)$ which interpolates v^k at the points of this stencil . We define the piecewise-polynomial interpolation $I^k(x; v^k)$ by

$$(1.25b) \qquad I^k(x; v^k) = p_i^k(x; v^k) \quad for \quad x_i^k \le x \le x_{i+1}^k.$$

Observe that by construction $I^k(x; v^k)$ is continuous and that

$$p_i^k(x; \mathcal{D}_k p) = p$$

for any polynomial p of degree less or equal $(r-1)$. Hence the interpolation is r-th order accurate, and the set of exactness in (1.15) is the linear span of

$$(1.26) \qquad p_\nu(x) = x^\nu, \quad 0 \le \nu \le r-1,$$

Up to now we have not specified the stencil $\mathcal{S}_i^k$ of r consecutive points of X^k that we assign to $[x_i^k, x_{i+1}^k]$. Clearly for $[x_1^k, x_2^k]$ at the left boundary we have no choice but to assign the one-sided stencil $\mathcal{S}_1^k = \{x_1^k, \ldots, x_r^k\}$; for the next interval $[x_2^k, x_3^k]$ we have two choices for $\mathcal{S}_2^k$: $\{x_1^k, \ldots, x_r^k\}$, $\{x_2^k, \ldots, x_{r+1}^k\}$; away from the boundaries we have $(r-1)$ choices for $\mathcal{S}_i^k$:

$$\{x_{i-r+2}^k, \ldots, x_{i+1}^k\}, \ldots, \{x_i^k, \ldots, x_{i+r-1}^k\}.$$

If we choose $\mathcal{S}_i^k$ independently of the data v^k then $I^k(\cdot; v^k)$ in (1.25b) is a linear functional of v^k. In this case the most accurate choice is that of a centered stencil (away from the boundaries) ,i.e. for $r = 2s$ we take

$$(1.27a) \qquad \mathcal{S}_i^k = \{x_{i-s+1}^k, \ldots, x_{i+s}^k\} \quad for \quad s \le i \le J_k - s,$$

and near the boundaries

$$(1.27a) \qquad \mathcal{S}_i^k = \{x_1^k, \ldots, x_r^k\} \quad for \quad 1 \le i \le s,$$

$$(1.27b) \qquad \mathcal{S}_i^k = \{x_{J_k-r+1}^k, \ldots, x_{J_k}^k\} \quad for \quad J_k - s \le i \le J_k.$$

Example 1.1.2 *Essentially Non-Oscillatory (ENO) Interpolation*

Let us consider discretized data which is obtained from sampling of piecewise-continuous data at the gridpoints X^k in (1.25a). It is well-known that any *data-independent* interpolation technique which is accurate for polynomials of degree larger or equal to 2 , has a Gibbs-like phenomenon of generating spurious oscillations near a discontinuity. In [H1] and [HEOC] we presented a data-dependent

12

piecewise-polynomial interpolation technique which avoids the Gibbs-phenomenon by an adaptive selection of stencil $\mathcal{S}_i^k$ in (1.25b) ; we refer to this technique as Essentially Non-Oscillatory (ENO) interpolation. The basic idea of ENO interpolation is to assign to $[x_i^k, x_{i+1}^k]$ which is in the smooth part of the sampled function, a stencil $\mathcal{S}_i^k = \{x_{i_0}^k, \ldots, x_{i_0+r-1}^k\}$ with $i_0 = i_0(i)$, which is likewise in the smooth part of the function (provided that this is possible ,i.e. that discontinuities are well separated and are far enough from the boundaries). This is done by choosing $\mathcal{S}_i^k$ to be the stencil for which the interpolation polynomial $p_i^k(x; v^k)$ in (1.25b) is the "smoothest" among all candidate-stencils, i.e. those of r consecutive points of X^k (starting with $x_{i_0}^k$) which contain both x_i^k and x_{i+1}^k ,e.g. by taking $i_0(i)$ to be the index for which

$$\min_{i_0} \left| \frac{d^{r-1}}{dx^{r-1}} p_i^k(x; v^k) \right|$$

is attained among all candidate-stencils.

Example 1.1.3 *Cubic-Spline Interpolation*

We assume now that the discretized data is generated by sampling of functions that are at least twice differentiable. In this case it makes sense to use cubic-spline interpolation (or higher-order ones). Take $x_1^k = 0$, $x_{J_k} = 1$ in (1.25a), and specify any two values q_0' and q_1'. We define the cubic-spline interpolation $I^k(x; v^k) = q(x)$ to be the unique piecewise-polynomial function $q(x)$ which is: (i) a cubic polynomial in each (x_{i-1}^k, x_i^k) (ii) twice differentiable in $[0, 1]$; (iii) interpolates v^k, and (iv) satisfies the boundary conditions $q'(0) = q_0'$, $q'(1) = q_1'$.

Example 1.1.4. *Trigonometric Interpolation*

We assume now that the discretized data is generated by pointvalue sampling of a smooth periodic signal, with say period 1, at the sampling points (1.25a) with $x_1^k = 0$, $x_{J_k}^k < 1$. In this case we can take the set of exactness in (1.15a) to be

$$(1.28a) \qquad p_\nu(x) = \cos \nu \pi x, \quad 0 \le \nu \le r - 1;$$

then the interpolation technique of choice is the Fourier-cosine collocation (trigonometric interpolation) which has $r = J_k$ (see [IK], [GO], [DH] and [H7]), i.e.

$$(1.28b) \qquad I^k(x; v^k) = \sum_{\nu=0}^{J_k-1} a_\nu(v^k) \cos \nu \pi x$$

where the coefficients $a_\nu(v^k)$ are the solution of the system of linear equations

$$(1.28c) \qquad I^k(x_i^k; v^k) = \sum_{\nu=0}^{J_k-1} a_\nu(v^k) \cos \nu \pi x_i^k = v_i^k, \quad 1 \le i \le J_k.$$

13

Example 1.2. Cell-average discretization.

Let $C^k = \{c_i^k\}$ be a covering of X by disjoint cells, i.e.

$$(1.29a) \qquad \overline{\cup_i c_i^k} = X, \qquad c_i^k \cap c_j^k = \emptyset \ \ for \ \ i \neq j,$$

and define the discretization

$$(1.29b) \qquad (\mathcal{D}_k f)_i = \frac{1}{|c_i^k|} \int_{c_i^k} f(x) dx, \qquad |c_i^k| = \int_{c_i^k} dx;$$

we refer to (1.29b) as discretization by cell-averages. Next let us consider a refinement sequence $\{C^k\}_{k=0}^L$, in which C^k is formed from C^{k-1} by dividing each cell c_i^{k-1} into, say q, disjoint cells $\{c_{i_\ell}^k\}_{\ell=1}^q$,

$$(1.30a) \qquad \overline{\cup_{\ell=1}^q c_{i_\ell}^k} = \overline{c_i^{k-1}}.$$

In this case the sequence of discretization $\{\mathcal{D}_k\}_{k=0}^L$ is nested and it follows from the additivity of the integral that

$$(1.30b) \qquad (\mathcal{D}_{k-1} f)_i = \frac{1}{|c_i^{k-1}|} \sum_{\ell=1}^q |c_{i_\ell}^k| (\mathcal{D}_k f)_{i_\ell} =: (D_k^{k-1} \mathcal{D}_k f)_i$$

which directly defines the decimation operator in (1.14). Let R_k denote any reconstruction from cell-averages, linear or not, and let e^k denote the prediction error in (1.5), then

$$(1.31) \qquad D_k^{k-1} e^k = 0 \Rightarrow \sum_{\ell=1}^q |c_{i_\ell}^k| e_{i_\ell}^k = 0.$$

This relation shows that we can define the scale coefficients d^k by taking $(q-1)$ properly chosen linear combinations of the q prediction errors $\{e_{i_\ell}^k\}_{\ell=1}^q$ in each cell c_i^{k-1}. These linear combinations should be chosen so that together with (1.31) they constitute an invertible system of q linear equations for the prediction errors $\{e_{i_\ell}^k\}_{\ell=1}^q$ in the cell c_i^{k-1}. (see e.g. [HY] for such combinations in representation of matrices).

Example 1.2.1. *Piecewise-Polynomial Reconstruction.*

Using notation similar to that of Example 1.1.1, let us denote by $\mathcal{S}_i^k$ a stencil of $s(r)$ *cells* in C^k which includes c_i^k, i.e.

$$(1.32a) \qquad \mathcal{S}_i^k = \{c_{i_m}^k\}_{m=1}^{s(r)}, \qquad c_i^k \in \mathcal{S}_i^k;$$

here $s(r)$ is the number of coefficients in a polynomial of degree $(r-1)$ in $\mathbf{R}^m$. Let $p_i^k(x; \mathcal{D}_k f)$ denote the unique polynomial of degree $(r-1)$ which attains the averages $(\mathcal{D}_k f)_{i_m}$ in $\mathcal{S}_i^k$, i.e. the one which satisfies the following system of $s(r)$ linear equations for its $s(r)$ coefficients:

$$(1.32b) \qquad \frac{1}{|c_{i_m}^k|} \int_{c_{i_m}^k} p_i^k(x; \mathcal{D}_k f) dx = (\mathcal{D}_k f)_{i_m}, \qquad m = 1, \ldots, s(r),$$

and define

$$(1.32c) \qquad (R_k \mathcal{D}_k f)(x) = p_i^k(x; \mathcal{D}_k f) \quad for \quad x \in c_i^k.$$

Clearly (1.32) defines a reconstruction of $\mathcal{D}_k f$ which is exact for polynomial functions of degree less or equal $(r-1)$, and thus is r-th order accurate. In subsection 5.B we describe techniques for selection of a "centered" stencil $\mathcal{S}_i^k$, as well as an "ENO stencil".

Note that for $r = 1$ in (1.32) we have $s(r) = 1$ and we get the piecewise-constant reconstruction

$$(1.33) \qquad (R_k \mathcal{D}_k f)(x) = \sum_i (\mathcal{D}_k f)_i \chi_{c_i^k}(x),$$

where $\chi_C(x)$ denotes the characteristic function of the set C,

$$(1.34) \qquad \chi_C(x) = \begin{cases} 1 & x \in C \\ 0 & \text{otherwise} \end{cases};$$

in subsection 5.B we show that this leads to a generalization of the Haar basis for unstructured grids.

Example 1.2.2. *Reconstruction via Primitive Function.*

We consider now the one-dimensional case with $X = [0, 1]$ and the grid

$$X^k = \{x_i^k\}_{i=0}^{J_k} \quad , \quad 0 = x_0^k < x_1^k < \cdots < x_{J_k}^k = 1.$$

We take $c_i^k = (x_{i-1}^k, x_i^k)$ and observe that

$$(\mathcal{D}_k f)_i = \frac{1}{|c_i^k|} \int_{x_{i-1}^k}^{x_i^k} f dx = \frac{1}{|c_i^k|} [F(x_i^k) - F(x_{i-1}^k)]$$

where $F(x)$ is the primitive function of $f(x)$, i.e.

$$F(x) = \int_0^y f(y) dy \quad , \quad \frac{d}{dx} F = f.$$

15

Given $v^k = \mathcal{D}_k f$ we set $F_0^k = 0$ and evaluate $F_i^k = F(x_i^k)$ for $1 \leq i \leq J_k$ by

$$(1.35a) \qquad F_i^k = \sum_{\ell=1}^{i} |c_\ell^k| v_\ell^k.$$

Let $I^k(x; F^k)$ denote *any* interpolation of the pointvalues of the primitive function at the gridpoints of X^k and define

$$(1.35b) \qquad \mathcal{R}_k v^k(x) = \frac{d}{dx} I^k(x; F^k);$$

using (1.35a) we get that

$$(\mathcal{D}_k \mathcal{R}_k v^k)_i = \frac{1}{|c_i^k|} \int_{x_{i-1}^k}^{x_i^k} \frac{d}{dx} I^k(x; F^k) dx = \frac{1}{|c_i^k|} [I^k(x_i^k; F^k) - I^k(x_{i-1}^k; F^k)]$$

$$= \frac{1}{|c_i^k|} (F_i^k - F_{i-1}^k) = v_i^k,$$

which shows that $\mathcal{R}_k$ in (1.35b) is indeed a reconstruction of the cell-average discretization.

We observe that if the interpolation in (1.35b) is taken to be the piecewise-polynomial interpolation in Examples 1.1.1 and 1.1.2, then the resulting reconstruction is likewise piecewise-polynomial, and because of uniqueness it is identical to the one which is defined in (1.32) for the one-dimensional case. Note however that the stencil in (1.32a) is of cells, and therefore the corresponding stencil for the pointvalues of the primitive function, which consists of the endpoints of the cells, has $s(r) + 1$ points of X^k. The technique of reconstruction of cell-average discretizaion via pointvalues of the primitive function can be extended to multidimensional cartesian grids (see [HC] and [HEOC]).

We turn now to consider discretization of *piecewise-continuous* functions. Unlike pointvalue discretization, where information about the *exact* location of a discontinuity is lost, cell-average discretization does retain this information. In [H2] we show that the location of a discontinuity in a piecewise-smooth function can be recovered from its cell-average data to any order of accuracy. Furthermore, we present a piecewise-polynomial reconstruction technique which is exact for piecewise-polynomial functions of the same polynomial degree as that of the reconstruction (provided that discontinuities in the function are well-separated and far enough from the boundaries); we refer to this technique as Essentially Nonoscillatory (ENO) reconstruction with subcell resolution.

1.G. Biorthogonal Wavelets Revisited.

We return now to the case of biorthogonal wavelets (1.8)-(1.10) of subsection 1.B and show that it can be cast into the framework of discretization and reconstruction, corresponding to $X = Y = \mathbf{R}$ in (1.11a) with $\mathcal{F} = L_2^{loc}(\mathbf{R})$. We assume now that $\{\alpha_\ell\}$ in (1.8) has its support in $[0, N]$ and consider the dilation equation for $w(x)$

$$(1.36a) \qquad w(x) = 2\sum_{\ell=0}^{N} \alpha_\ell w(2x - \ell), \quad \sum_{\ell=0}^{N} \alpha_\ell = 1.$$

It is shown in [Da] that $w(x)$ has likewise a compact support of size N, and is determined by (1.36a) up to a multiplicative constant and a shift, which we now fix by taking its support in $[-N/2, N/2]$ and normalizing it by

$$(1.36b) \qquad \int_{\mathbf{R}} w(x)dx = 1.$$

Furthermore, it is shown in [Da] that $w(x)$, the solution of (1.36a), is at least a distribution, and that by imposing additional conditions on $\{\alpha_\ell\}$, or equivalently on the Fourier symbol $m_0(\xi)$

$$(1.36c) \qquad m_0(\xi) = \sum_{\ell=0}^{N} \alpha_\ell e^{-i\ell\xi},$$

(and thus increasing the support N) it can be made regular to any desired degree; we assume now that $\{\alpha_\ell\}_{\ell=0}^{N}$ is so that $w(x)$ is at least square integrable.

Let $\{X^k\}_{k=0}^{L}$ be the following nested dyadic sequence of uniform grids in $\mathbf{R}$

$$(1.37) \qquad X^k = \{x_i^k\}_{i=-\infty}^{\infty}, \quad x_i^k = ih_k, \quad h_k = 2^{-k}h_0, \quad h_0 > 0,$$

denote

$$(1.38a) \qquad w_i^k(x) = \frac{1}{h_k}w\left(\frac{x - x_i^k}{h_k}\right),$$

and observe that

$$(1.38b) \qquad \int_{\mathbf{R}} w_i^k(x)dx = 1.$$

We now define the sequence of discretization operators $\{\mathcal{D}_k\}$

$$(1.39a) \qquad \mathcal{D}_k : L_2^{loc}(\mathbf{R}) \longrightarrow S,$$

17

where S is the space of infinite sequences (1.3b), by

$$(1.39b) \qquad v_i^k = (\mathcal{D}_k f)_i = \int_{\mathbf{R}} f(x) w_i^k(x) dx, \quad f \in L_2^{loc}(\mathbf{R}).$$

Since $w(x)$ is the solution of the dilation equation (1.36a), it follows from (1.38a) that

$$(1.40a) \qquad w_i^{k-1}(x) = \sum_{\ell=0}^{N} \alpha_\ell w_{2i+\ell}^k(x);$$

in subsection 5.B we show that (1.40a) implies that $\{\mathcal{D}_k\}$ in (1.39) is nested. Using (1.40a) in (1.39b) we get the following relation between $v^{k-1} = \mathcal{D}_{k-1} f$ and $v^k = \mathcal{D}_k f$, which is the decimation (1.8):

$$(1.40b) \qquad v_i^{k-1} = \sum_{\ell=0}^{N} \alpha_\ell v_{2i+\ell}^k =: (D_k^{k-1} v^k)_i;$$

note that in the scalar case $\hat{D}_k^{k-1} = D_k^{k-1}$.

In subsection 2.E we show that for each $k \geq 0$ the set $\{\mu_j^k\}$ which is defined by

$$(\mu_j^k)_i = (-1)^{i+1} \alpha_{2j-i-1}$$

is a basis for $\mathcal{N}(D_k^{k-1})$, the null space of the decimation operator. We define the scale coefficients $\{d_j^k\}$ in terms of the prediction error e^k by

$$e^k = \sum_j d_j^k \mu_j^k \quad,$$

and show that it follows from the biorthogonality relation (1.10a) that the scale coefficients can be expressed explicitly by

$$d_j^k = \sum_m (-1)^{m+1} \beta_{2j-m-1} e_m^k.$$

The prediction operator $P_{k-1}^k = \hat{P}_{k-1}^k$ in (1.9) is obtained by (1.13b) from the particular choice of reconstruction sequence $\{\mathcal{R}_k\}$ which is defined by

$$(1.41a) \qquad \mathcal{R}_k v^k = \sum_i v_i^k \varphi_i^k(x), \quad \varphi_i^k(x) = \varphi\left(\frac{x - x_i^k}{h_k}\right);$$

18

here $\varphi(x)$ is the solution of the dilation equation with the coefficients $\{\beta_\ell\}$ in (1.10)

$$(1.41b) \qquad \varphi(x) = \sum_\ell \beta_\ell \varphi(2x - \ell), \quad \sum_\ell \beta_{2\ell-1} = \sum_\ell \beta_{2\ell} = 1,$$

which is normalized by

$$(1.41c) \qquad \int_{\mathbf{R}} \varphi(x) w(x) dx = 1.$$

Since $\varphi(x)$ is a solution of the dilation equation (1.41b) we get that, as in (1.40a),

$$(1.41d) \qquad \varphi_i^{k-1}(x) = \sum_\ell \beta_\ell \varphi_{2i+\ell}^k(x).$$

It is shown in [CDF] that the biorthogonality condition (1.10a) implies that

$$(1.42a) \qquad \int_{\mathbf{R}} \varphi(x) w(x - \ell) dx = \delta_{0,\ell},$$

which in turn implies that $\{\varphi_i^k\}$ and $\{w_i^k\}$ are biorthogonal systems, i.e.

$$(1.42b) \qquad (\mathcal{D}_k \varphi_\ell^k)_m = \int_{\mathbf{R}} \varphi_\ell^k(x) w_m^k(x) dx = \delta_{\ell,m}.$$

In subsection 5.B we shall prove the biorthogonality (1.42b) as part of a more general result for unstructured grids. Choosing $\{\beta_\ell\}$ such that $\varphi \in L_2(\mathbf{R})$, we get that for any sequence v^k:

(i) $\quad \sum_i v_i^k \varphi_i^k(x) \in L_2^{loc}(\mathbf{R})$,

(ii) $\quad [\mathcal{D}_k(\sum_i v_i^k \varphi_i^k)]_\ell = \sum_i v_i^k (\mathcal{D}_k \varphi_i^k)_\ell = \sum_i v_i^k \delta_{i,\ell} = v_\ell^k$;

recall that since $\{\beta_\ell\}$ is a finite sequence, $\varphi(x)$ is of compact support. This shows that $\mathcal{R}_k v^k$ as defined by (1.41a) is indeed a proper reconstruction in the sense of (1.12). Finally we show that the prediction operator in (1.13b)

$$P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$$

is identical to (1.9a). It follows from (1.41d) and (1.42b) that

$$(\mathcal{D}_k \varphi_m^{k-1})_i = \sum_\ell \beta_\ell (\mathcal{D}_k \varphi_{2m+\ell}^k)_i = \sum_\ell \beta_\ell \delta_{2m+\ell,i} = \beta_{i-2m},$$

and therefore we get from (1.41a) that

$$(\mathcal{D}_k \mathcal{R}_{k-1} v^{k-1})_i = (\mathcal{D}_k \sum_m v_m^{k-1} \varphi_m^{k-1})_i = \sum_m v_m^{k-1} (\mathcal{D}_k \varphi_m^{k-1})_i$$

$$= \sum_m v_m^{k-1} \beta_{i-2m}.$$

We observe that due to the compact support of $\{\alpha_\ell\}$ and $\{\beta_\ell\}$, and the fact that polynomials are in $L_2^{loc}(\mathbf{R})$, we can take the polynomials in (1.26) as the set of exactness. Doing so we obtain in [H7] the following set of conditions

$$(1.43a) \qquad \sum_\ell \beta_{2\ell-1} = \sum_\ell \beta_{2\ell} = 1,$$

$$(1.43b) \qquad \sum_m m^\ell \sum_n \beta_{2n}\alpha_{2n+m} = 0 = \sum_m m^\ell \sum_n \beta_{2n+1}\alpha_{2n+m+1}$$
$$for \;\; 1 \le \ell \le r-1, \;\; r \ge 2;$$

and show that conditions (1.43) imply that

$$(1.44) \qquad \sum_m (-1)^m m^\ell \beta_m = 0 \;\; for \;\; 0 \le \ell \le r-1, \;\; r \ge 2.$$

The conditions on $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ in [CDF] are derived from considerations of regularity of $\varphi(x)$ and $w(x)$, and are expressed in terms of the Fourier symbols $\hat{\varphi}(\xi)$ and $\hat{w}(\xi)$. On the other hand, our conditions (1.43) are obtained from requirements of accuracy; however, if we assume duality (see subsection 2.E) then the two different formulations lead to the same choice of $\{\alpha_\ell\}$ and $\{\beta_\ell\}$.

The particular choice

$$(1.45) \qquad\qquad\qquad \beta_\ell = 2\alpha_\ell$$

corresponds to the case of orthonormal wavelets in [Da], in which case $\varphi(x) = w(x)$ and (1.42b) implies that $\{\varphi_i^k\}$ is an orthogonal sequence: To make it orthonormal we can redefine $\varphi_i^k(x) = (h_k)^{-\frac{1}{2}} \varphi\left(\frac{x-x_i^k}{h_k}\right)$ and modify the coefficients $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ accordingly. We remark that in the case of orthonormal wavelets (1.45), condition (1.44) which is used here for accuracy is the same one which is used in [Da] for regularity (see also [S] and [CDF]).

For $N = 0$ we get in (1.36) that $\alpha_0 = 1$ and the dilation equation becomes $w(x) = 2w(2x)$, which is solvable only in the sense of distributions. In this case $w(x) = \delta(x)$ (the Dirac distribution) and then $\{\mathcal{D}_k\}$ in (1.39b) is the pointvalue discretization (1.21) for the dyadic sequence of uniform grids (1.37)(see [S]). Since $\delta(x)$ is not in L_2 it is excluded from the class of wavelets; however the discretization (1.39b) is well-defined once we modify (1.39a) to read $\mathcal{D}_k : C^0(\mathbf{R}) \longrightarrow S$.

For $N = 1$ with the choice $\alpha_0 = \alpha_1 = \frac{1}{2}$ we get $w(x) = \chi_{[-1,0]}(x)$ in (1.36). In this case $\{\mathcal{D}_k\}$ in (1.39) is the discretization by cell-averages (1.29) in the dyadic uniform grids (1.37); note that we shifted the support of $w(x)$ from $[-1/2, 1/2]$ to $[-1, 0]$ in order to use the notation of Example 1.2.2.

The examples of subsection 1.E demonstrate that the framework of subsection 1.D. provides a broad and natural generalization of biorthogonal wavelets: It removes the need to discretize on the dyadic sequence of uniform grids of $\mathbf{R}$ in (1.37) and allows for discretizations in unstructured meshes in $\mathbf{R}^m$, finite or infinite, as long as $\{\mathcal{D}_k\}$ is nested. Furthermore, it allows for nonlinear reconstruction techniques which are needed for piecewise-smooth functions (Examples 1.1.2 and 1.2.2).

2. General Framework

In this section we generalize the multiresolution representation of sequences of subsection 1.A to obtain multiresolution representation of elements in a linear space which has a denumerable basis, and discuss the stability of the corresponding MR schemes.

Definition 2.1. Multiresolution Setting. .

Let $\{V^k\}$ be a sequence of linear spaces which have denumerable bases, and let $\{D_k^{k-1}\}$ be a sequence of *linear* operators which map V^k *onto* V^{k-1}, i.e.

$$(2.1) \qquad D_k^{k-1} : V^k \to V^{k-1}, \quad V^{k-1} = D_k^{k-1}(V^k).$$

We say that $(\{V^k\}, \{D_k^{k-1}\})$ is a multiresolution setting, and refer to D_k^{k-1} as the decimation operator.

2.A. Notation agreement:

Let V^k denote a linear space which has a basis $\{\eta_i^k\}$, finite or infinite, and denote the representation of $v^k \in V^k$ in this basis by the sequence $\hat{v}^k = \{\hat{v}_i^k\}$ of its coordinates,

$$(2.2) \qquad V^k = span\{\eta_i^k\}, \quad v^k \in V^k, \quad v^k = \sum_i \hat{v}_i^k \eta_i^k, \quad \hat{v}^k = \{\hat{v}_i^k\} \in S^k;$$

here we use for $\{\eta_i^k\}$ and $\{\hat{v}_i^k\}$ the same notation agreement as in (1.3); in the finite-dimensional case we denote $\dim V^k = J_k$.

We denote the identity operator in V^k by I_k.

Let $\mathcal{N}(D_k^{k-1})$ denote the null space of D_k^{k-1}

$$(2.3a) \qquad \mathcal{N}(D_k^{k-1}) = \{v^k | v^k \in V^k, \; D_k^{k-1} v^k = 0\},$$

and let $\{\mu_j^k\}$ denote any basis of $\mathcal{N}(D_k^{k-1})$. Observe that in the finite dimensional case, it follows from (2.1) that

$$(2.3b) \qquad \dim \mathcal{N}(D_k^{k-1}) = \dim V^k - \dim V^{k-1} = J_k - J_{k-1};$$

21

to distinguish between the different dimensions of the two bases $\{\eta_i^k\}$ and $\{\mu_j^k\}$ we reserve the index i to denote $i = 1, \ldots, J_k$ and j to denote $j = 1, \ldots, (J_k - J_{k-1})$. The index k is reserved for the level of resolution.

Let $d^k = \{d_j^k\}$ denote any sequence which has the dimension of $\mathcal{N}(D_k^{k-1})$, and denote the space of all such sequences by $\mathcal{G}^k$. We define the operator $E^k : \mathcal{G}^k \rightarrow \mathcal{N}(D_k^{k-1})$ by

$$(2.4a) \qquad E_k d^k = \sum_j d_j^k \mu_j^k.$$

Let $G^k : \mathcal{N}(D_k^{k-1}) \rightarrow \mathcal{G}^k$ be the operator which assigns to any $e^k \in \mathcal{N}(D_k^{k-1})$ the sequence d^k of its coordinates in the basis $\{\mu_j^k\}$; clearly

$$(2.4b) \qquad e^k \in \mathcal{N}(D_k^{k-1}) \;,\quad d^k = G_k e^k \Leftrightarrow e^k = \sum_j d_j^k \mu_j^k = E_k d^k.$$

It follows from (2.4b) that

$$(2.4c) \qquad d^k \in \mathcal{G}^k \Rightarrow G_k E_k d^k = d^k,$$

and

$$(2.4d) \qquad e^k \in \mathcal{N}(D_k^{k-1}) \Rightarrow E_k G_k e^k = e^k.$$

2.B. Multiresolution Representation.

It follows directly from (2.1) that for any $v^{k-1} \in V^{k-1}$ there is at least one $u \in V^k$ such that $D_k^{k-1} u = v^{k-1}$. We refer to such an assignment of u to v^{k-1} as prediction, and define:

Definition 2.2. Prediction operator.

We say that P_{k-1}^k is a prediction operator for the multiresolution setting (2.1), if it is a right-inverse of D_k^{k-1} in V^{k-1}, i.e.

$$(2.5) \qquad P_{k-1}^k : V^{k-1} \rightarrow V^k, \quad D_k^{k-1} P_{k-1}^k = I_{k-1}.$$

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

Lemma 2.1.

Let P_{k-1}^k be a prediction operaor, linear or not, and define (the possibly non-linear) operator Q_k by

$$(2.6a) \qquad Q_k = I_k - P_{k-1}^k D_k^{k-1}.$$

Then:

$$(2.6b) \qquad \text{(i)} \qquad e^k \in \mathcal{N}(D_k^{k-1}) \Rightarrow Q_k e^k = e^k,$$

$$(2.6c) \qquad \text{(ii)} \qquad Q_k : V^k \xrightarrow{\text{onto}} \mathcal{N}(D_k^{k-1}).$$

Proof: (i) $e^k \in \mathcal{N}(D_k^{k-1}) \Rightarrow D_k^{k-1} e^k = 0$ and therefore

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

(ii) For any $v^k \in V^k$ we get from the linearity of D_k^{k-1} and (2.5) that

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

and therefore $Q_k(V^k) \subseteq \mathcal{N}(D_k^{k-1})$. From part (i) we get that

$$\mathcal{N}(D_k^{k-1}) = Q_k(\mathcal{N}(D_k^{k-1})) \subseteq Q_k(V^k)$$

and we conclude that $Q_k(V^k) = \mathcal{N}(D_k^{k-1})$, which proves (2.6c).

<div align="right">⊓</div>

Theorem 2.1.

Let $(\{V^k\}_{k=0}^L, \{D_k^{k-1}\}_{k=1}^L)$ *be a multiresolution setting, and let* $\{P_{k-1}^k\}_{k=1}^L$ *be any sequence of corresponding prediction operators, linear or nonlinear. We define a transformation M,*

$$M : V^L \to S^L,$$

where S^L is the space of sequences (1.3), by the following algorithm: Given any $v \in V^L$

$$(2.7a) \qquad\qquad Set \ \ v^L = v,$$

$$(2.7b) \qquad\qquad \begin{cases} Do \ \ for \ \ k = L, \dots, 1 \\ v^{k-1} = D_k^{k-1} v^k \\ d^k = G_k(Q_k v^k), \end{cases}$$

<div align="center">23</div>

let $\hat{v}^0$ denote the coordinates of v^0 in the basis $\{\eta_i^0\}$ (i.e. (2.2) with $k = 0$), and define

(2.7c)
$$M(v) = \{\hat{v}^0, d^1, d^2, \ldots, d^L\}.$$

Then M is an invertible transformation,

(2.8)
$$v \in V^L \xrightarrow{\ 1:1\ } M(v) \in S^L,$$

and its inverse

$$M^{-1} : S^L \to V^L$$

is defined by the following algorithm: Given any sequence $\{\hat{v}^0, d^1, \ldots, d^L\} \in S^L$,

(2.9a)
$$Set \quad v^0 = \sum_i \hat{v}_i^0 \eta_i^0,$$

(2.9b)
$$\begin{cases} Do \quad for \quad k = 1, \ldots, L \\ v^k = P_{k-1}^k v^{k-1} + E_k d^k, \end{cases}$$

and define

(2.9c)
$$M^{-1}(\{\hat{v}^0, d^1, \ldots, d^L\}) = v^L.$$

Proof: (i) We want to show that for any $v \in V^L$, $M^{-1}(M(v)) = v$. This amounts to proving the claim that if

$$v^{k-1} = D_k^{k-1} v^k \quad , \quad d^k = G_k Q_k v^k,$$

then v^k is recovered by

$$v^k = P_{k-1}^k v^{k-1} + E_k d^k.$$

From Lemma 2.1 we get that $Q_k v^k \in \mathcal{N}(D_k^{k-1})$ and therefore it follows from (2.4d) that

$$E_k G_k Q_k v^k = Q_k v^k.$$

Using (2.6a) we get that

$$P_{k-1}^k v^{k-1} + E_k d^k = P_{k-1}^k (D_k^{k-1} v^k) + E_k (G_k Q_k v^k)$$
$$= (P_{k-1}^k D_k^{k-1}) v^k + Q_k v^k = v^k,$$

24

which proves our claim.

(ii)We want to show that for any sequence $\{\hat{v}^0, d^1, \ldots, d^L\} \in S^L$,

$$M(M^{-1}(\{\hat{v}^0, d^1, \ldots, d^L\})) = \{\hat{v}^0, d^1, \ldots, d^L\}.$$

This amounts to proving the claim that if

$$v^k = P_{k-1}^k v^{k-1} + E_k d^k,$$

then

$$v^{k-1} = D_k^{k-1} v^k \quad , \quad d^k = G_k Q_k v^k.$$

Since $E_k d^k \in \mathcal{N}(D_k^{k-1})$ it follows from (2.5) and the linearity of D_k^{k-1} that

$$D_k^{k-1} v^k = D_k^{k-1} P_{k-1}^k v^{k-1} + D_k^{k-1}(E_k d^k) = v^{k-1} + 0 = v^{k-1},$$

and therefore

$$\begin{aligned} Q_k v^k &= (I_k - P_{k-1}^k D_k^{k-1}) v^k = v^k - P_{k-1}^k v^{k-1} \\ &= (P_{k-1}^k v^{k-1} + E_k d^k) - P_{k-1}^k v^{k-1} = E_k d^k; \end{aligned}$$

using (2.4c) we get that

$$G_k Q_k v^k = G_k E_k d^k = d^k,$$

which proves our claim.

$$\sqcap$$

We refer to $M(v^L)$ as the multiresolution representation of v^L, and to algorithms (2.7), (2.9) as the direct and inverse multiresolution transforms, respectively.

Remark 2.1. Since by (2.2) $v^L \overset{1:1}{\leftrightarrow} \hat{v}^L$, where $\hat{v}^L$ is *any* sequence of numbers in S^L, it follows from Theorem 2.1 that the same algorithms (2.7), (2.9) can be used for $\hat{M}(s^L)$, the associated multiresolution representation of sequences s^L in S^L (1.3), by defining

$$(2.10) \qquad s^L \overset{1:1}{\leftrightarrow} \hat{M}(s^L) =: M(v^L), \quad v^L = \sum_i s_i^L \eta_i^L.$$

Remark 2.2. D_k^{k-1}, G_k and E_k are linear operators and therefore can be represented by matrices, finite or infinite, which we denote by $\hat{D}_k^{k-1}, \hat{G}_k$ and $\hat{E}_k$, respectively; in this case the sequences $\hat{v}^k$ and d^k are considered to be column-vectors; thus

$$v^{k-1} = D_k^{k-1} v^k \leftrightarrow \hat{v}^{k-1} = \hat{D}_k^{k-1} \hat{v}^k,$$
$$d^k = G_k e^k \leftrightarrow d^k = \hat{G}_k \hat{e}^k,$$
$$e^k = E_k d^k \leftrightarrow \hat{e}^k = \hat{E}_k d^k.$$

If P_{k-1}^k is a linear operator it can be likewise represented by a matrix, which we denote by $\hat{P}_{k-1}^k$.

2.C. Multiresolution Bases in V^L.

The multiresolution transform can be expressed directly in terms of v^L by introducing the linear operator B_L^k of successive decimation

$$(2.11a) \qquad B_L^k = D_{k+1}^k \cdots D_L^{L-1} : V^L \to V^k,$$

and observing that v^k in (2.7) are defined by

$$(2.11b) \qquad v^k = B_L^k v^L;$$

Thus $v^L \mapsto M(v^L)$ can be expressed by

$$(2.11c) \qquad v^0 = B_L^0 v^L, \quad d^k = G_k Q_k B_L^k v^L, \quad 1 \leq k \leq L.$$

If $\{P_{k-1}^k\}$ are all *linear* prediction operators, then the inverse multiresoluton transform (2.9) can be expressed directly in terms of $M(v^L)$ by introducing the linear operator A_k^L of successive prediction

$$(2.12a) \qquad A_k^L = P_{L-1}^L \cdots P_k^{k+1} : V^k \to V^L;$$

thus $M(v^L) \mapsto v^L$ can be expressed by

$$(2.12b) \qquad v^L = A_0^L v^0 + \sum_{k=1}^L A_k^L E_k d^k.$$

Denoting

$$(2.13a) \qquad \bar{\varphi}_i^{k,L} = A_k^L \eta_i^k, \quad 0 \leq k \leq L,$$
$$(2.13b) \qquad \bar{\psi}_j^{k,L} = A_k^L \mu_j^k, \quad 1 \leq k \leq L,$$

we get from (2.12b), (2.2) and (2.4a) that

$$(2.13c) \qquad v^L = \sum_i \hat{v}_i^0 \bar{\varphi}_i^{0,L} + \sum_{k=1}^L \sum_j d_j^k \bar{\psi}_j^{k,L}.$$

Thus we have shown:

26

Theorem 2.2.

Let $(\{V^k\}_{k=1}^L,\ \{D_k^{k-1}\}_{k=1}^L)$ be a multiresolution setting, and let $\{P_{k-1}^k\}_{k=1}^L$ be any sequence of linear prediction operators, then

$$(2.13d) \qquad \bar{\mathcal{B}}_M = \left(\{\bar{\varphi}_i^{0,L}\}_i, \left\{\{\bar{\psi}_j^{k,L}\}_j\right\}_{k=1}^L \right)$$

is a basis of V^L, and any $v^L \in V^L$ has the representation (2.13c) where the coordinates are given by the direct MR transform $M(v^L)$.

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

Remark 2.3. Let us denote the representation of $\mu_j^k \in \mathcal{N}(D_k^{k-1}) \subset V^k$ in the basis $\{\eta_i^k\}$ of V^k by

$$(2.14a) \qquad \mu_j^k = \sum_i (\widehat{\mu_j^k})_i \eta_i^k,$$

then $(\hat{E}_k)_{i,j} = (\widehat{\mu_j^k})_i$ (see Remark 2.2), and it follows from (2.13a) - (2.13b) that

$$(2.14b) \qquad \bar{\psi}_j^{k,L} = \sum_i (\hat{E}_k)_{i,j} \bar{\varphi}_i^{k,L};$$

thus $\bar{\psi}_j^{k,L}$ is a linear combination of $\{\bar{\varphi}_i^{k,L}\}_i$ which is *independent of L*.

Remark 2.4 Let $\hat{P}_{k-1}^k$ denote the matrix representation of the prediction P_{k-1}^k (see Remark 2.2). Since $P_{k-1}^k \eta_\ell^{k-1}$ is in V^k, this matrix representation is defined by

$$(2.15a) \qquad P_{k-1}^k \eta_\ell^{k-1} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \eta_i^k.$$

It follows from the definition of A_k^L in (2.12a) that

$$A_{k-1}^L = A_k^L P_{k-1}^k,$$

therefore we get from (2.13a) that

$$\bar{\varphi}_\ell^{k-1,L} = A_{k-1}^L \eta_\ell^{k-1} = A_k^L P_{k-1}^k \eta_\ell^{k-1} = A_k^L \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \eta_i^k;$$

thus

$$(2.15b) \qquad \bar{\varphi}_\ell^{k-1,L} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \bar{\varphi}_i^{k,L}.$$

This shows that $\bar{\varphi}_\ell^{k-1,L}$ is a linear combination of $\{\bar{\varphi}_i^{k,L}\}$, which is *independent of L*.

2.D. How to Design MR Schemes?

The point of view of this paper is that the primary choice to be made is that of the decimation operator D_k^{k-1}. In the next section we introduce the concept of nested discretization and show that it defines a decimation operator, and thus a multiresolution setting (2.1). Here the "discretization" specifies the nature of the data in $\hat{v}^L$ (i.e. how it was generated), and the "nestedness" induces a sense of hierarchical levels of resolution.

Once we have chosen a multiresolution setting, we have to make two more *independent* choices:

(1) A prediction operator P_{k-1}^k which is a right-inverse of D_k^{k-1};

(2) A basis $\{\mu_j^k\}$ of $\mathcal{N}(D_k^{k-1})$.

In the finite dimensional case we have shown that for any choice of prediction technique and any choice of basis in $\mathcal{N}(D_k^{k-1})$

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

has $(J_k - J_{k-1})$ components, and consequently the J_L coefficients of $\hat{v}^L$ are represented by the J_L coefficients of $M(v^L)$. We would like d^k to be a good approximation to the "true" k-th scale. Although the crucial element in achieving this goal is the accuracy of the prediction, it is not the only consideration.

In order to apply this multiresolution representation to real-life problems for purposes of analysis and data compression, we have to make sure that the direct MR transform and its inverse are stable with respect to perturbations

For purposes of analysis, if v^L is replaced by a perturbed v_ε^L, we want the perturbation in the resulting scale coefficients d_ε^k,

$$(2.16) \qquad d_\varepsilon^k - d^k = G_k Q_k B_L^k(v_\varepsilon^L - v^L),$$

to be "bounded" by the perturbation in the input. Relation (2.16) shows that the perturbation in the input is subject to successive decimation D_{m-1}^m for $m = L, \dots, k+1$, and then the result is projected into $\mathcal{N}(D_k^{k-1})$ and represented by some basis there. Clearly the "dangerous" process that we have to control 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 $\{d_\varepsilon^k\}$ which are obtained either by quantization (representation by less

bits according to some table of prescribed significance) or truncation (setting to zero of coefficients that fall below some prescribed tolerance), we want the perturbation in the decompressed v_ε^L

$$(2.17) \qquad v_\varepsilon^L - v^L = \sum_{k=1}^{L} A_k^L E_k (d_\varepsilon^k - d^k),$$

to be "bounded" by the perturbation in the scale coefficients. Examination of (2.17) shows that the perturbation in the k-th scale coefficients is "translated" into a perturbation in the prediction error, which is then transmitted into higher levels of resolution by successive prediciton P_{m-1}^m for $m = k+1, \ldots, L$. The danger here is that this perturbation could be amplified by this process.

In sections 3 and 4 we shall formulate conditions which ensure the stability of the direct MR transform and of its inverse.

2.E. Design of Biorthogonal Wavelets.

In this subsection we desscribe the derivation of biorthogonal wavelets in [CDF] and compare it to our approach. The MR scheme of [CDF] is taken to be of the following form: The direct multiresolution transform is

$$(2.18a) \qquad v^{k-1} = Dv^k, \quad d^k = G_P v^k, \quad k = L, \ldots, 1$$

and its inverse is

$$(2.18b) \qquad v^k = Pv^{k-1} + G_D^* d^k, \quad k = 1, \ldots, L.$$

The infinite matrices D and P in (1.8)-(1.9) are given by

$$(2.19a) \qquad D_{i,j} = \alpha_{j-2i}, \quad P_{i,j} = \beta_{i-2j};$$

G_D^* and G_P are expressed in terms of $\{\alpha_\ell\}$ and $\{\beta_\ell\}$, respectively, by

$$(2.19b) \qquad (G_D^*)_{i,j} = (-1)^{i+1} \alpha_{2j-i-1}, \quad (G_P)_{i,j} = (-1)^{j+1} \beta_{2i-j-1}.$$

The first condition on $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ is that of consistency: Using (2.18a) in (2.18b) we get the identity

$$(2.20) \qquad PD + G_D^* G_P = I.$$

It is easy to see that (2.20) is equivalent to the biorthogonality relation (1.10a). At this point we associate $w(x)$ and $\varphi(x)$ to $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ by defining them to be the

solution of the respective dilation equation (1.36) and (1.41b)-(1.41c); the functions $w(x)$ and $\varphi(x)$ are regarded as scaling functions in *dual* systems of wavelet bases. Additional conditions are then imposed on the Fourier symbols $\hat{w}(\xi)$ and $\hat{\varphi}(\xi)$ in order to ensure increasing regularity of the scaling functions $w(x)$ and $\varphi(x)$. Note that there are no direct requirements of accuracy in this formulation; it comes out later that increasing regularity implies higher accuracy.

Next we would like to compare (2.18)-(2.19) to our MR scheme (2.7)-(2.9) for the same case. The direct MR transform is

$$(2.21a) \qquad v^{k-1} = Dv^k \quad , \quad d^k = GQv^k \quad k = l, \ldots, 1$$

and its inverse is

$$(2.21b) \qquad v^k = Pv^{k-1} + Ed^k \quad , \quad k = 1, \ldots, L.$$

It is easy to see that as in the orthogonal case of [Da]

$$(2.22a) \qquad DG_D^* = 0 \quad ,$$

$$(2.22b) \qquad G_P P = 0 \quad ,$$

and that the biorthogonality relation (1.10a) implies

$$(2.22c) \qquad G_P G_D^* = I.$$

It follows from (2.22a) that we can take the basis $\{\mu_j\}$ of $\mathcal{N}(\mathcal{D})$ in (2.4) to be the columns of G_D^*, i.e.

$$(2.23a) \qquad (\mu_j)_i = (-1)^{i+1} \alpha_{2j-i-1}.$$

For this choice of basis we get

$$(2.23b) \qquad e^k = \sum_j d_j^k \mu_j = G_D^* d^k \Rightarrow E = G_D^*;$$

multiplying the above equality from the left by G_P and using (2.22c), we get

$$(2.23c) \qquad G_P e^k = G_P G_D^* d^k = d^k \Rightarrow G = G_P.$$

It follows then from (2.22b) that

$$(2.23d) \qquad GQ = G_P(I - PD) = G_P,$$

30

and we conclude that the two algorithms are identical.

We remark that the assumption of duality, i.e. that replacing $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ by $\{\tilde{\alpha}_\ell\}$ and $\{\tilde{\beta}_\ell\}$,

$$(2.24a) \qquad\qquad \tilde{\alpha}_\ell = \beta_\ell/2 \quad , \quad \tilde{\beta}_\ell = 2\alpha_\ell$$

results in a consistent MR scheme, amounts to adding the condition (see (1.10b))

$$(2.24b) \qquad \sum_\ell \tilde{\beta}_{2\ell} = \sum_\ell \tilde{\beta}_{2\ell+1} = 1 \Rightarrow \sum_\ell \alpha_{2\ell} = \sum_\ell \alpha_{2\ell+1} = 1/2;$$

this condition implies that $w(x)$ is square integrable. Under this assumption of duality we can exchange the roles of the sequences $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ in (1.43) and obtain in (1.44) the additional condition

$$(2.25) \qquad \sum_m (-1)^m m^\ell \alpha_m = 0 \qquad for \;\; 0 \le \ell \le \tilde{r} - 1, \;\; \tilde{r} \ge 1$$

for some $\tilde{r} \ge 1$. The system of algebraic equations for $\{\alpha_\ell\}$ and $\{\beta_\ell\}$ which is used in [CDF] is (2.24b), (1.43a), (1.10a), (1.44) and (2.25); hence under assumption of duality the two formulations are identical.

However, unlike [CDF] , we do not attempt to explicitly solve this system of algebraic equations. Our approach is to first choose $w(x)$ (which is equivalent to choosing $\{\alpha_\ell\}$ in (1.36)) and then solve the *approximation* problem of finding $\mathcal{R}_k$ in (1.12) which is an r-th order accurate reconstruction of the discretization $\mathcal{D}_k$ in (1.39). Typically, as in Example 1.1.1, $\mathcal{R}_k v^k$ is a piecewise-polynomial function. In [H7] we show that if we use the same reconstruction technique at all the points and all the levels of the dyadic sequence of uniform grids (1.37), then the prediction operator (1.13b) $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ has the Töplitz -like representation (2.19a) ,(1.9) , and the coefficients $\{\beta_\ell\}$ *automatically* satisfy the set of equations (1.43).

We remark that the piecewise-polynomial reconstruction $\mathcal{R}_k$ is different from the wavelet reconstruction in (1.41) ; the latter can be obtained from the piecewise-polynomial $\mathcal{R}_k$ by the limiting process $L \to \infty$ in (1.18) (see subsection 5.B).

The main advantage of our approach is that it applies directly to bounded domains – all we have to do is modify the reconstruction near the boundaries (as in Example 1.1.1). Thus "biorthogonal wavelets in the interval" are obtained by using one-sided stencils near the boundaries (see [H7]).

3. Design of MR Schemes from Nested Discretization

In this section we show how to construct a multiresolution setting (Definition 2.1) from a nested sequence of discretization and discuss the stability of MR schemes.

Definition 3.1 Discretization.

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, say $\{\eta_i\}$, we say that $\mathcal{D}$ is a discretization operator and refer to $v = \mathcal{D}f$ as the discretization of f

$$(3.1) \qquad \mathcal{D} : \mathcal{F} \longrightarrow V, \qquad V = \mathcal{D}(\mathcal{F}) = span\{\eta_i\}.$$

Definition 3.2. Nested discretization.

Let $\{\mathcal{D}_k\}$ be a sequence of discretization operators

$$(3.2a) \qquad \mathcal{D}_k : \mathcal{F} \longrightarrow V^k, \qquad V^k = \mathcal{D}_k(\mathcal{F}) = span\{\eta_i^k\}.$$

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

$$(3.2b) \qquad \mathcal{D}_k f = 0 \Rightarrow \mathcal{D}_{k-1} f = 0.$$

3.A. Multiresolution Setting.

Let $\{\mathcal{D}_k\}$ be a sequence of discretization as in (3.2a), and consider the following mapping from V^k to V^{k-1}: For $v \in V^k$ take any f in $\mathcal{F}$ such that $v = \mathcal{D}_k f$ and assign to it $u = \mathcal{D}_{k-1} f$, i.e.

$$(3.3) \qquad v \mapsto u = \mathcal{D}_{k-1} f, \quad where \quad v = \mathcal{D}_k f.$$

Lemma 3.1. (Harten-Lax)

If $\{\mathcal{D}_k\}$ is nested, then (3.3) is a well-defined mapping.

Proof: Since $V^k = \mathcal{D}_k(\mathcal{F})$, any $v \in V^k$ has at least one $f \in \mathcal{F}$ such that $v = \mathcal{D}_k f$. We want to prove that the assignment $u = \mathcal{D}_{k-1} f$ to v is independent of the particular choice of such f. To do so we take any $g \in \mathcal{F}$ such that

$$\mathcal{D}_k f = v = \mathcal{D}_k g$$

and show that

$$\mathcal{D}_{k-1} f = \mathcal{D}_{k-1} g.$$

Since $\mathcal{D}_k$ and $\mathcal{D}_{k-1}$ are linear operators

$$0 = \mathcal{D}_k f - \mathcal{D}_k g = \mathcal{D}_k(f - g)$$

we get from (3.2b) that

$$0 = \mathcal{D}_{k-1}(f - g) = \mathcal{D}_{k-1}f - \mathcal{D}_{k-1}g.$$

⊓

We denote the mapping (3.3) by $u = D_{k-1}^k v$, and observe that for any $f \in \mathcal{F}$

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

Theorem 3.1.

Let $\{\mathcal{D}_k\}$ be a nested sequence of discretization (3.2) and let $\{D_k^{k-1}\}$ be the corresponding mappings (3.3), then

$$(\{V^k\}, \{D_k^{k-1}\})$$

is a multiresolution setting.

Proof: The linearity of the mapping D_k^{k-1} follows immediately from the linearity of $\mathcal{D}_k$, $\mathcal{D}_{k-1}$ and the space $\mathcal{F}$. To show that D_k^{k-1} maps V^k *onto* V^{k-1}, for any u in V^{k-1} take a $f \in \mathcal{F}$ such that $u = \mathcal{D}_{k-1}f$ and let $v = \mathcal{D}_k f$. Clearly v is in V^k and by (3.4)

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

⊓

Let $\mathcal{D}$ be the discretization (3.1), then to any v in V we can assign a $f \in \mathcal{F}$ such that $\mathcal{D}f = v$. We refer to such an assignment of $f \in \mathcal{F}$ to $v \in V$, as a reconstruction of f from $v = \mathcal{D}f$.

Definition 3.3. Reconstruction.

We say that $\mathcal{R}$

(3.5a) $$\mathcal{R} : V \longrightarrow \mathcal{F}, \qquad V = \mathcal{D}(\mathcal{F})$$

is a reconstruction operator in V, if it is a right-inverse of $\mathcal{D}$, i.e.

(3.5b) $$\mathcal{D}\mathcal{R} = I,$$

where I is the identity operator in V.

Note that $\mathcal{R}$ is not required to be a linear operator.

Lemma 3.2.

If $\{\mathcal{D}_k\}$ is nested, then

$$(3.6) \qquad \mathcal{D}_\ell(\mathcal{R}_m\mathcal{D}_m) = \mathcal{D}_\ell \quad \text{for} \quad \ell \le m \ .$$

Proof: For any $f \in \mathcal{F}$ denote

$$g = \mathcal{R}_m\mathcal{D}_m f \ ;$$

it follows from Definitions 3.2 and 3.3 that

$$\mathcal{D}_m g = (\mathcal{D}_m\mathcal{R}_m)\mathcal{D}_m f = \mathcal{D}_m f \Rightarrow \mathcal{D}_\ell g = \mathcal{D}_\ell f \quad \text{for} \quad \ell \le m$$
$$\Rightarrow \mathcal{D}_\ell(\mathcal{R}_m\mathcal{D}_m)f = \mathcal{D}_\ell f \ .$$

$\sqcap$

Theorem 3.2.

Let $\{\mathcal{D}_k\}$ be a nested sequence of discretization and let $\{\mathcal{R}_k\}$ be a corresponding sequence of reconstruction (3.5), linear or not. Then

$$(3.7\text{a}) \qquad \text{(i)} \qquad D_k^{k-1} = \mathcal{D}_{k-1}\mathcal{R}_k$$

is an expression for the decimation operator (3.3)–(3.4);

$$(3.7\text{b}) \qquad \text{(ii)} \qquad P_{k-1}^k = \mathcal{D}_k\mathcal{R}_{k-1}$$

is a corresponding prediction operator (Definition 2.2).

Proof: It follows from Lemma 3.2 that

$$\mathcal{D}_{k-1}(\mathcal{R}_k\mathcal{D}_k) = \mathcal{D}_{k-1} \ ;$$

therefore D_k^{k-1} in (3.7a) satisfies (3.4), i.e.

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

(ii) Multiplying the above from the right by $\mathcal{R}_{k-1}$ we get

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

$\sqcap$

Remark 3.1. The operator $(\mathcal{D}_{k-1}\mathcal{R}_k)$ in the RHS of (3.7a) is defined for any two discretizations, nested or not. However, it is easy to see that it is a decimation operator (Definition 2.1) *only if* they are nested. In this case the decimation operator D_k^{k-1} is defined directly by (3.3), and then (3.7a) becomes a conceptual description which is independent of the particular $\mathcal{R}_k$, linear or nonlinear. Likewise, the operator $(\mathcal{D}_k\mathcal{R}_{k-1})$ in the RHS of (3.7b) is defined for any two discretizations, nested or not. However, it is a prediction operator, i.e. a right inverse of $(\mathcal{D}_{k-1}\mathcal{R}_k)$, *only if* the two discretizations are nested.

Thus the converse to Theorem 3.2 is also true.

3.B. Sequence of Discrete Approximation.

Let v^L be any vector in V^L and denote

(3.8a)
$$f = \mathcal{R}_L v^L \in \mathcal{F} \,,$$

then it follows from (3.4) that the sequence $\{v^k\}_{k=0}^L$ in the direct MR transform (2.7) can be expressed by

(3.8b)
$$v^k = \mathcal{D}_k f, \quad k = L, \ldots, 0 \,.$$

The prediction error e^k,

(3.8c)
$$e^k = v^k - P_{k-1}^k v^{k-1} \,,$$

can be expressed as $e^k = e^k(f)$ by

(3.8d)
$$e^k(f) = \mathcal{D}_k f - (\mathcal{D}_k \mathcal{R}_{k-1})\mathcal{D}_{k-1} f = \mathcal{D}_k(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})f \,.$$

Let us assume that $\{\mathcal{R}_k\}$ are linear operators and denote

(3.9)
$$\Pi_k^L = (\mathcal{R}_L \mathcal{D}_L) \cdots (\mathcal{R}_k \mathcal{D}_k) : \mathcal{F} \to \mathcal{F}$$

and observe that

(3.10a) $\quad \mathcal{R}_L \mathcal{D}_L = \Pi_L^L = \Pi_0^L + \sum_{k=1}^{L} \left(\Pi_k^L - \Pi_{k-1}^L \right) = \Pi_0^L + \sum_{k=1}^{L} \Pi_k^L (I - \mathcal{R}_{k-1}\mathcal{D}_{k-1}) \,.$

Applying (3.10a) to f in (3.8) we get

(3.10b)
$$\mathcal{R}_L v^L = \mathcal{R}_L \mathcal{D}_L f = \Pi_1^L \cdot \mathcal{R}_0 v^0 + \sum_{k=1}^{L} \Pi_{k+1}^L \cdot \mathcal{R}_k e^k \,.$$

Observe that $\Pi_{k+1}^L \mathcal{R}_k = \mathcal{R}_L A_k^L$ and therefore (3.10b) can also be obtained by applying $\mathcal{R}_L$ to the inverse MR transform (2.12).

The analysis above shows that the MR scheme in this case can be expressed in terms of the sequence $\{(\mathcal{R}_k \mathcal{D}_k)\}$.

From now on we take $\mathcal{F}$ to be a Banach space with a norm $\| \ \|$.

35

Definition 3.4. Sequence of (discrete) approximation.

We say that $\{(\mathcal{R}_k \mathcal{D}_k)\}$,

$$(3.11a) \qquad\qquad (\mathcal{R}_k \mathcal{D}_k) : \mathcal{F} \to \mathcal{F} \ ,$$

is a sequence of (discrete) approximation in a Banach space $\mathcal{F}$ if for any $f \in \mathcal{F}$

$$(3.11b) \quad \text{(i)} \qquad \|\mathcal{R}_k \mathcal{D}_k f\| \le C_A^k \|f\|$$

$$(3.11c) \quad \text{(ii)} \qquad \|\mathcal{R}_k \mathcal{D}_k f - f\| \to 0 \quad \text{as} \quad k \to \infty \ .$$

Lemma 3.3.

If $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is a sequence of approximation, then it is uniformly bounded, i.e. there exists a constant C_A such that for any $f \in \mathcal{F}$

$$(3.12) \qquad\qquad \|\mathcal{R}_k \mathcal{D}_k f\| \le C_A \|f\| \ .$$

Proof: From (3.11c) we get that

$$\|\mathcal{R}_k \mathcal{D}_k f\| \to \|f\| \quad \text{as} \quad k \to \infty$$

and therefore $\{\|\mathcal{R}_k \mathcal{D}_k f\|\}$ is bounded, i.e. there exists a bound $B(f)$ which does not depend on k such that for any $f \in \mathcal{F}$

$$\|\mathcal{R}_k \mathcal{D}_k f\| \le B(f) \ .$$

This bound, together with our assumption (3.11b), implies (3.12) by the principle of uniform boundedness (see e.g. [RM, pp. 34-36]).

$\square$

Next we define the discrete norms $| \ |_k$ for elements in V^k, and $\langle \ \rangle_k$ for the scale coefficients d^k in $\mathcal{G}^k$; these special norms are designed to accommodate the different dimensions of the various levels of resolution in the finite-dimensional case.

Lemma 3.4.

Let $\{(\mathcal{R}_k \mathcal{D}_k)\}$ be a sequence of approximation and denote for any $v^k \in V^k$

$$(3.13a) \qquad\qquad |v^k|_k = \|\mathcal{R}_k v^k\| \ ;$$

then

(i) $|\ |_k$ *is a norm in* V^k,

(ii) *for any* $f \in \mathcal{F}$ *such that* $v^k = \mathcal{D}_k f$

$$(3.13b) \qquad |v^k|_k = |\mathcal{D}_k f|_k \leq C_A \|f\| \ .$$

Proof: (i) Clearly $|v^k| = \|\mathcal{R}_k v^k\| \geq 0$, and

$$|v^k| = 0 \Leftrightarrow \|\mathcal{R}_k v^k\| = 0 \Leftrightarrow \mathcal{R}_k v^k = 0 \Leftrightarrow v^k = \mathcal{D}_k \mathcal{R}_k v^k = \mathcal{D}_k \cdot 0 = 0 \ .$$

The triangle inequality for $|\ |_k$ follows immediately from that of $\|\ \|$

$$|\alpha_1 v_1^k + \alpha_2 v_2^k|_k = \|\mathcal{R}_k(\alpha_1 v_1^k + \alpha_2 v_2^k)\| = \|\alpha_1 \mathcal{R}_k v_1^k + \alpha_2 \mathcal{R}_k v_2^k\|$$
$$\leq |\alpha_1| \, \|\mathcal{R}_k v_1^k\| + |\alpha_2| \, \|\mathcal{R}_k v_2^k\| = |\alpha_1| \, |v_1^k|_k + |\alpha_2| \, |v_2^k|_k \ .$$

(ii)
$$|v^k|_k = |\mathcal{D}_k f|_k = \|\mathcal{R}_k \mathcal{D}_k f\|$$

and (3.13b) follows from(3.12).

$$\sqcap$$

Lemma 3.5.

Let $\{(\mathcal{R}_k \mathcal{D}_k)\}$ *be a sequence of approximation and denote for any element* $d^k \in \mathcal{G}^k$ *in* (2.4)

$$(3.14a) \qquad \langle d^k \rangle_k = |E_k d^k|_k \ ;$$

then

(i) $\langle \ \rangle_k$ *is a norm in* $\mathcal{G}^k$;

(ii) *for any* $f \in \mathcal{F}$, $d^k(f) = G_k e^k(f)$ *satisfies*

$$(3.14b) \qquad \langle d^k(f) \rangle_k = |e^k(f)|_k \leq C_A \|(I - \mathcal{R}_{k-1} \mathcal{D}_{k-1})f\| \ .$$

37

Proof: (i) $\langle d^k \rangle_k = |E_k d^k|_k \geq 0$. It follows from Lemma 3.4 and (2.4c) that

$$\langle d^k \rangle_k = 0 \Leftrightarrow |E_k d^k|_k = 0 \Leftrightarrow E^k d^k = 0 \Leftrightarrow d^k = G^k(E^k d^k) = G^k \cdot 0 = 0 .$$

The triangle inequality for $\langle \ \rangle_k$ follows immediately from that of $|\ |_k$.
(ii) It follows from Lemma 3.3 and (3.8d) that

$$\langle d^k(f) \rangle_k = |E_k d^k(f)|_k = |e^k(f)|_k = \|\mathcal{R}_k e^k(f)\| = \|\mathcal{R}_k \mathcal{D}_k(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})f\|$$
$$\leq C_A \|(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})f\| .$$

$\sqcap$

Corollary 3.1.

If $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is a sequence of approximation then for any $f \in \mathcal{F}$

(3.14c)
$$\langle d^k(f) \rangle_k = |e^k(f)|_k \to 0 \quad \text{as} \quad k \to \infty .$$

and

(3.14d)
$$\langle d^k(f) \rangle_k \leq C_A(1 + C_A)\|f\|$$

3.C. Stability and Data Compression.

In this subsection we discuss the question of stability of the MR scheme that we raised in subsection 2.D.

First let us consider the direct MR transform (2.7) and denote by $\delta(v^L)$ the perturbation in its input, and by $\delta(d^k)$ and $\delta(v^0)$ the perturbation in its output. Because of the linearity of the operators we get from (3.8) for $f = \mathcal{R}_L \delta(v^L)$ that

$$\mathcal{R}_k E_k \delta(d^k) = \mathcal{R}_k \delta(e^k) = \mathcal{R}_k \mathcal{D}_k(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})\mathcal{R}_L \delta(v^L)$$

and

$$\mathcal{R}_0 \delta(v^0) = \mathcal{R}_0 \mathcal{D}_0 \cdot \mathcal{R}_L \delta(v^L) .$$

Using the discrete norms (3.13a), (3.14a) and Lemma 3.3 we conclude that

Theorem 3.3.

If $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is a sequence of approximation then the direct MR transform (2.7) is stable and

(3.15a) $\qquad \langle\delta(d^k)\rangle_k = |\delta(e^k)|_k \leq C_A(1+C_A)|\delta(v^L)|_L$

(3.15b) $\qquad |\delta(v^0)|_0 \leq C_A|\delta(v^L)|_L$.

Observe that the assumption of nestedness implies through Lemma 3.2 that $B_L^k\mathcal{D}_L = \mathcal{D}_k$; this eliminates the possibility of amplification of the perturbation $\delta(v^L)$ by the successive decimation in (2.16).

Definition 3.5. (Uniformly) π-Stable sequence of approximation.

We say that the sequence of approximation $\{(\mathcal{R}_\ell\mathcal{D}_\ell)\}$ is π-stable if for all $k \geq 0$ there exists a constant $\tilde{C}_k$ such that for any $f \in \mathcal{F}$

(3.16a) $\qquad \|\Pi_k^L \cdot f\| \leq \tilde{C}_k\|f\|, \quad L \geq k$.

We say that the sequence is *uniformly* π-stable if there exists a constant $\tilde{C}_\pi$,such that for all $k \geq 0$ in (3.16a)

(3.16b) $\qquad \tilde{C}_k \leq \tilde{C}_\pi < \infty.$

Lemma 3.6.

The sequence of approximation $\{(\mathcal{R}_\ell\mathcal{D}_\ell)\}$ is π-stable if and only if for any $k \geq 0$ there exists a constant C_k such that for any $v \in V^k$

(3.17a) $\qquad \|\Pi_{k+1}^L \cdot \mathcal{R}_kv\| \leq C_k|v|_k, \quad L \geq k+1$.

The sequence is uniformly π-stable if for all $k \geq 0$ in (3.17a)

(3.17b) $\qquad C_k \leq C_\pi < \infty.$

Proof: If the sequence is π-stable, for any $v \in V^k$ take $f = \mathcal{R}_kv$ in (3.16a); then

$$\Pi_k^L f = \Pi_k^L \mathcal{R}_kv = \Pi_{k+1}^L(\mathcal{R}_k\mathcal{D}_k)\mathcal{R}_kv = \Pi_{k+1}^L \mathcal{R}_kv$$

and we get that

$$\|\Pi_{k+1}^L \mathcal{R}_k v\| = \|\Pi_k^L f\| \leq \tilde{C}_k \|f\| = \tilde{C}_k \|\mathcal{R}_k v\|$$

and (3.17a) follows from the definition (3.13a) of $|v|_k$. Clearly, if the sequence is uniformly π-stable, then (3.17b) is also true.

Conversely, if (3.17a) is true then for any $f \in \mathcal{F}$ we take $v = \mathcal{D}_k f$; then

$$\Pi_{k+1}^L \mathcal{R}_k v = \Pi_{k+1}^L (\mathcal{R}_k \mathcal{D}_k) f = \Pi_k^L f$$

and we get from (3.17a) and (3.13b) that

$$\|\Pi_k^L f\| = \|\Pi_{k+1}^L \mathcal{R}_k v\| \leq C_k |\mathcal{D}_k f|_k \leq C_k C_A \|f\|.$$

If (3.17b) is true, then

$$\|\Pi_k^L f\| \leq C_\pi C_A \|f\|$$

and the sequence is uniformly π-stable.

$$\sqcap$$

We turn now to consider the stability of the inverse MR transform and prove that

Theorem 3.4.

Let $\delta(d^k)$ and $\delta(v^0)$ denote the perturbation in the input of the inverse MR transform (2.9) and let $\delta(v^L)$ denote the resulting perturbation in its output. If $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is a π-stable sequence of approximation then the inverse MR transform (2.9) is stable and

$$(3.18a); \qquad |\delta(v^L)|_L \leq C_0 |\delta(v^0)|_0 + \sum_{k=1}^L C_k \langle \delta(d^k) \rangle_k \ .$$

If the sequence is uniformly π-stable , then

$$(3.18b); \qquad |\delta(v^L)|_L \leq C_\pi \big[|\delta(v^0)|_0 + \sum_{k=1}^L \langle \delta(d^k) \rangle_k \ \big].$$

Proof: It follows from (3.10b) and $\delta(e^k) = E_k \delta(d^k)$ that

$$(3.19) \qquad \mathcal{R}_L \delta(v^L) = \Pi_1^L \cdot \mathcal{R}_0 \delta(v^0) + \sum_{k=1} \Pi_{k+1}^L \cdot \mathcal{R}_k E_k \delta(d^k) \ .$$

40

Using Lemma 3.6 we get

$$\|\mathcal{R}_L\delta(v^L)\| \leq C_0\|\mathcal{R}_L\delta(v^0)\| + \sum_{k=1}^{L} C_k\|E_k\delta(d^k)\|,$$

and (3.18a) follows from the definition of the various discrete norms.

If the sequence is uniformly π-stable, then (3.18b) follows from (3.17b).

$$\sqcap$$

Next let us consider the application of the MR scheme to data compression. In this case we replace the scale coefficients $\{d^k\}$ by $\{d_\varepsilon^k\}$ which are obtained either by quantization or by truncation. Assuming that $\delta(v^0) = 0$ and that the data compression introduces purturbations $\delta(d^k)$ such that

$$(3.20a) \qquad\qquad C_k\langle\delta(d^k)\rangle_k \leq \varepsilon_k$$

we get from (3.18) that

$$(3.20b) \qquad\qquad |\delta(v^L)|_L \leq \sum_{k=1}^{L} \varepsilon_k \ .$$

The "art" of data compression is to find a sequence ε_k, $k = 1,\ldots,L$ which sums up to a given ε and maximizes the rate of compression.

4. Multiresolution Bases of $\mathcal{F}$

In this section we assume that $\mathcal{F}$ is a Banach space with norm $\| \ \|$, and derive multiresolution bases for it. To simplify our presentation we continue to use the notation agreement of subsection 2.A. However, in the infinite case the following should be regarded as a formal derivation, which can be made rigorous later by adding conditions to ensure convergence of the infinite sums and to allow for term-by-term operations when needed.

Let $\{(\mathcal{R}_k\mathcal{D}_k)\}$ be a sequence of approximation and denote

$$(4.1a) \qquad\qquad \varphi_i^{k,L} = \Pi_{k+1}^{L} \cdot \mathcal{R}_k\eta_i^k$$

$$(4.1b) \qquad\qquad \psi_j^{k,L} = \Pi_{k+1}^{L} \cdot \mathcal{R}_k\mu_j^k \ .$$

Using

$$e^k = \sum_i \hat{e}_i^k \eta_i^k = \sum_j d_j^k \mu_j^k$$

41

we get from (3.10b) that

$$(4.2a) \qquad \mathcal{R}_L \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^{0,L} + \sum_{k=1}^{L} \sum_i \hat{e}_i^k \varphi_i^{k,L}$$

or

$$(4.2b) \qquad \mathcal{R}_L \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^{0,L} + \sum_{k=1}^{L} \sum_j d_j^k \psi_j^{k,L} \; ,$$

where the coefficients in (4.2b) are computed by the direct MR transform (2.7) of $\mathcal{D}_L f$, i.e.

$$(4.2c) \qquad \{\hat{f}^0, d^1, d^2, \ldots, d^L\} = M(\mathcal{D}_L f).$$

Observe that as in (2.14)-(2.15)

$$(4.2d) \qquad \psi_j^{k,L} = \sum_i (\hat{E}_k)_{i,j} \varphi_i^{k,L}$$

$$(4.2e) \qquad \varphi_\ell^{k-1,L} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \varphi_i^{k,L}.$$

4.A. Hierarchial Sequence of Approximation.

In this subsection we describe multiresolution bases which correspond to hierarchial bases of $\mathcal{F}$; the latter are a generalization of the ones that are used within the context of finite-element methods (see e.g. [Y1] and [Y2]).

Definition 4.1. Hierarchial sequence. We say that the sequence $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is hierarchial, if for all k

$$(4.3a) \qquad (\mathcal{R}_k \mathcal{D}_k) \mathcal{R}_{k-1} = \mathcal{R}_{k-1}.$$

Observe that since $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ (3.8), another way to express (4.3a) is by

$$(4.3b) \qquad \mathcal{R}_k P_{k-1}^k = \mathcal{R}_{k-1}.$$

Lemma 4.1.

If $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is an hierarchial sequence then for all k and $L \geq k$

$$(4.4) \qquad \Pi_k^L = \mathcal{R}_k \mathcal{D}_k \; .$$

Proof: Using (4.3a) we get

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

$\sqcap$

Corollary 4.1.

If $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is an hierarchial sequence of approximation, then:

(i) it is uniformly π-stable ;

(ii) perturbations in the inverse MR transform are bounded by

(4.5a) $$|\delta(v^L)|_L \leq |\delta(v^0)|_0 + \sum_{k=1}^{L} \langle \delta(d^k) \rangle_k .$$

Proof: (i) It follows from (4.4) and Lemma 3.3 that

$$\|\Pi_k^L f\| = \|\mathcal{R}_k \mathcal{D}_k f\| \leq C_A \|f\|.$$

(ii) It follows from (4.4) and (4.3a) that

(4.5b) $$\Pi_{k+1}^{\ell} \mathcal{R}_k = \mathcal{R}_{k+1} \mathcal{D}_{k+1} \mathcal{R}_k = \mathcal{R}_k;$$

therefore in (3.19)

$$\mathcal{R}_L \delta(v^L) = \mathcal{R}_0 \delta(v^0) + \sum_{k=1}^{L} \mathcal{R}_k E_k \delta(d^k) ,$$

and (4.5a) follows from the various definitions of the discrete norms.

$$\sqcap$$

Observe that the hierarchial structure (4.3a) eliminates the possibility of amplification by successive prediction of perturbations $\delta(d^k)$ in the input of the inverse MR transform in (2.17); this is analogous to what the nestedness $\mathcal{D}_{k-1}(\mathcal{R}_k\mathcal{D}_k) = \mathcal{D}_{k-1}$ does for the stability of the direct MR transform (Theorem 3.3).

Theorem 4.1.

If $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is an hierarchial sequence of approximation then

(4.6a) $$\mathcal{B}_M = \left(\{\varphi_i^0\}, \left\{ \{\psi_j^k\}_j \right\}_{k=1}^{\infty} \right) ,$$

with

(4.6b) $$\varphi_i^k = \mathcal{R}_k \eta_i^k ,$$

(4.6c) $$\psi_j^k = \mathcal{R}_k \mu_j^k ,$$

43

is a multiresolution basis of $\mathcal{F}$, and for any $f \in \mathcal{F}$

$$(4.6d) \qquad f = \sum_i \hat{f}_i^0 \varphi_i^0 + \sum_{k=1}^{\infty} \sum_j d_j^k \psi_j^k \ ,$$

where the coefficients $d_j^k = d_j^k(f)$ are the scale coefficients (4.2c), i.e.

$$(4.6e) \qquad d^k(f) = G_k e^k(f) = G_k(\mathcal{D}_k f - P_{k-1}^k \mathcal{D}_{k-1} f).$$

Proof: It follows from (4.5b) that in (4.1a)-(4.1b)

$$\varphi_i^{k,L} = \mathcal{R}_k \eta_i^k = \varphi_i^k,$$
$$\psi_j^{k,L} = \mathcal{R}_k \mu_j^k = \psi_j^k,$$

and therefore (4.2b) becomes

$$(4.7) \qquad \mathcal{R}_L \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^0 + \sum_{k=1}^{L} \sum_j d_j^k \psi_j^k,$$

where the coefficients in this expansion are given by (4.2c). Since by (3.11c), the LHS of (4.7) converges to f in the norm $\| \ \|$ we get (4.6d) and thus $\mathcal{B}_M$ (4.6a) is a basis of $\mathcal{F}$.

$$\sqcap$$

Remark 4.1. Since for hierarchial sequences $\varphi_i^{k,L} = \varphi_i^k$ and $\psi_j^{k,L} = \psi_j^k$ it follows from (4.2d)-(4.2e) that

$$(4.8a) \qquad \varphi_\ell^{k-1} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \varphi_i^k,$$

$$(4.8b) \qquad \psi_j^k = \sum_i (\hat{E}_k)_{i,j} \varphi_i^k.$$

Denoting the linear span of $\{\varphi_i^m\}$ by Φ^m, we get from (4.8a) that for all k

$$(4.9a) \qquad \Phi^{k-1} \subset \Phi^k.$$

Furthermore,

$$(4.9b) \qquad (\mathcal{R}_k \mathcal{D}_k) f = \sum_i \hat{f}_i^k \varphi_i^k \implies (\mathcal{R}_k \mathcal{D}_k) f \in \Phi^k$$

and

$(4.9c)$ $$(\mathcal{R}_k \mathcal{D}_k)^2 = \mathcal{R}_k(\mathcal{D}_k \mathcal{R}_k)\mathcal{D}_k = \mathcal{R}_k \mathcal{D}_k;$$

Thus $(\mathcal{R}_k \mathcal{D}_k)$ is a projection of $\mathcal{F}$ onto Φ^k. Observe that

$(4.10a)$ $$\mathcal{R}_k \mathcal{D}_k f = \mathcal{R}_{k-1}\mathcal{D}_{k-1}f + \sum_j d_j^k \psi_j^k = \sum_i \hat{f}_i^{k-1}\varphi_i^{k-1} + \sum_j d_j^k \psi_j^k$$

corresponds to the direct sum decomposition

$(4.10b)$ $$\Phi^k = \Phi^{k-1} \oplus \Psi^k,$$

where Ψ^k, the linear span of $\{\psi_j^k\}$, is the complement of Φ^{k-1} in Φ^k.

Remark 4.2. The sequence $\{(\mathcal{R}_k \mathcal{D}_k)\}$ of the biorthogonal wavelets in (1.40)-(1.41) is hierarchial. To show that we use (1.9a) and (1.41d) to prove (4.3b), as follows: For any $v^{k-1} \in V^{k-1}$

$$\mathcal{R}_k P_{k-1}^k v^{k-1} = \sum_m (P_{k-1}^k v^{k-1})_m \varphi_m^k = \sum_m (\sum_i \beta_{m-2i}v_i^{k-1})\varphi_m^k$$
$$= \sum_i v_i^{k-1}(\sum_\ell \beta_\ell \varphi_{\ell+2i}^k) = \sum_i v_i^{k-1}\varphi_i^{k-1} = \mathcal{R}_{k-1}v^{k-1}.$$

It is easy to see that the sequences of discrete approximation in Examples 1.1.3 and 1.1.4 are also hierarchial; this is true in general for reconstruction sequences $\{\mathcal{R}_k\}$ which are based on splines or spectral expansion (see [DH], [H4] and [H7]). The standard piecewise-polynomial interpolation in Example 1.1.1 which uses a sliding stencil is not hierarchial. This method is the most commonly used interpolation technique in numerical analysis, e.g. finite-difference schemes for the numerical solution of PDEs. In subsection 4.D we show that in many cases, a sequence of approximation which is not hierarchial to begin with, has an hierarchial form which is obtained by taking $L \to \infty$ in (4.1)-(4.2); this hierarchial form has the same scale coefficients as the original one.

4.B. Cosmetic Refinement.

In this subsection we prepare the framework for taking $L \to \infty$ in (4.1)–(4.2).

Definition 4.2. π-Convergent sequence of approximation.

We say that the sequence of approximation $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}_{\ell=0}^\infty$ is π-convergent if for any $k \geq 0$ and any $f \in \mathcal{F}$ the sequence $\{\Pi_k^L f\}_{L=k}^\infty$ is a Cauchy sequence in $\| \ \|$.

Theorem 4.2.

If $\{(\mathcal{R}_\ell\mathcal{D}_\ell)\}_{\ell=0}^\infty$ is a π-convergent sequence of approximation then it is also π-stable.

Proof: Since $\{\Pi_k^L f\}_{L=k}^\infty$ is a Cauchy sequence for any $f \in \mathcal{F}$ we get that $\{\|\Pi_k^L f\|\}$ is convergent and therefore bounded, i.e. for any $f \in \mathcal{F}$

$$(4.11a) \qquad \|\Pi_k^L f\| \leq \tilde{B}_k(f) \ .$$

From Lemma 3.3 we get that for $L \geq k$

$$(4.11b) \qquad \|\Pi_k^L f\| \leq C_A \|\Pi_k^{L-1} f\| \leq \cdots \leq (C_A)^{L-k} \|f\| \ ;$$

since $\mathcal{F}$ is assumed to be a Banach space, it follows from (4.11a), (4.11b) and the principle of uniform boundedness that for any $f \in \mathcal{F}$ and any $k \geq 0$ there exists a constant $\tilde{C}_k$ (independent of f) such that for all $L \geq k$

$$(4.11c) \qquad \|\Pi_k^L f\| \leq \tilde{C}_k \|f\|.$$

$\sqcap$

Theorem 4.3.

Let $\{(\mathcal{R}_k\mathcal{D}_k)\}_{k=0}^\infty$ be a π-convergent sequence of approximation and denote

$$(4.12a) \qquad \lim_{L \to \infty} \Pi_k^L f = f_k^\infty \in \mathcal{F} \ ;$$

then

$$(4.12b) \qquad \mathcal{D}_\ell f_k^\infty = \mathcal{D}_\ell f \quad \text{for} \quad \ell \leq k$$
$$(4.12c) \qquad d^\ell(f_k^\infty) = 0 \quad \text{for} \quad \ell \geq k+1 \ .$$

Proof: It follows from (3.9) and Lemma 3.2 that

$$(4.13a) \qquad \mathcal{D}_\ell\Pi_k^L = \mathcal{D}_\ell\Pi_k^{\ell-1} \quad \text{for} \quad k+1 \leq \ell \leq L$$

and

$$(4.13b) \qquad \mathcal{D}_\ell\Pi_k^L = \mathcal{D}_\ell \quad \text{for} \quad 0 \leq \ell \leq k \ .$$

46

Therefore using (4.13b) for $0 \le \ell \le k$ we get from (3.13b) that

$$0 \le |\mathcal{D}_\ell f_k^\infty - \mathcal{D}_\ell f|_\ell = |\mathcal{D}_\ell f_k^\infty - \mathcal{D}_\ell \Pi_k^L f|_\ell \le C_A \|f_k^\infty - \Pi_k^L f\| \underset{L \to \infty}{\longrightarrow} 0$$

which proves (4.12b). Using (3.8d) we get

$$e^\ell(f_k^\infty) = \mathcal{D}_\ell(I - \mathcal{R}_{\ell-1}\mathcal{D}_{\ell-1})f_k^\infty = \mathcal{D}_\ell(I - \mathcal{R}_{\ell-1}\mathcal{D}_{\ell-1})\Pi_k^L f$$
$$+ \mathcal{D}_\ell(I - \mathcal{R}_{\ell-1}\mathcal{D}_{\ell-1})(f_k^\infty - \Pi_k^L f) \ ;$$

it is easy to see that (4.13a) implies for $L \ge \ell \ge k+1$

$$\mathcal{D}_\ell \Pi_k^L = \mathcal{D}_\ell \Pi_k^{\ell-1} = \mathcal{D}_\ell \mathcal{R}_{\ell-1} \mathcal{D}_{\ell-1} \Pi_k^L,$$

therefore

$$\mathcal{D}_\ell(I - \mathcal{R}_{\ell-1}\mathcal{D}_{\ell-1})\Pi_k^L = 0$$

and thus using (3.13b) and (3.12) we get

$$0 \le |e^\ell(f_k^\infty)|_\ell = |\mathcal{D}_\ell(I - \mathcal{R}_{\ell-1}\mathcal{D}_{\ell-1})(f_k^\infty - \Pi_k^L f)|_\ell$$
$$\le C_A(1 + C_A)\|f_k^\infty - \Pi_k^L f\| \underset{L \to \infty}{\longrightarrow} 0 \ ;$$

this shows that $e^\ell(f_k^\infty) = 0$ for $\ell \ge k+1$, which implies (4.12c).

$$\sqcap$$

$\Pi_k^L f$ is described on a higher level of resolution than $\Pi_k^{L-1} f$, and in this respect f_k^∞ corresponds to "infinite resolution". Nevertheless, Theorem 4.3 shows that f_k^∞ has exactly the same (discrete) information contents as the initial data $\mathcal{R}_k \mathcal{D}_k f$. Therefore we refer to the limiting process (4.12a) which assigns f_k^∞ to $\mathcal{R}_k \mathcal{D}_k f$ as a cosmetic refinement scheme; the qualification "cosmetic" is used in order to stress that unlike other refinement processes in numerical analysis, there is no addition of information in (4.12a).

We remark that the linear operator which assigns f_k^∞ (4.12a) to each $f \in \mathcal{F}$ is bounded, since by (4.11c)

$$(4.13c) \qquad \qquad \|f_k^\infty\| = \lim_{L \to \infty} \|\Pi_k^L f\| \le \tilde{C}_k \|f\| \ .$$

4.C. Sufficient Conditions for π-Convergence

In this subsection we consider a sequence of approximation $\{(\mathcal{R}_k \mathcal{D}_k)\}$ and derive sufficient conditions for its π-convergence (Definition 4.3). Let h_k be a parameter which measures the "coarseness" of the discretization $\mathcal{D}_k$. Since $\{\mathcal{D}_k\}$ is nested, h_k is decreasing with k and we assume that for all k

$$(4.14) \qquad \qquad h_{k+1} \le \tilde{q} \cdot h_k, \quad 0 < \tilde{q} < 1 \ .$$

Let $\tau_k(f)$ denote $\|(I - \mathcal{R}_k\mathcal{D}_k)f\|$, or any error bound of the approximation $(\mathcal{R}_k\mathcal{D}_k)$, i.e.

$$(4.15a) \qquad \|(I - \mathcal{R}_k\mathcal{D}_k)f\| \le \tau_k(f) \ ,$$

$$(4.15b) \qquad \tau_k(f) \to 0 \quad \text{as} \quad k \to \infty \ ;$$

in the latter case we assume that $\tau_k(\cdot)$ is a semi-norm in $\mathcal{F}$, and that there exists a constant C_0 such that for all $k \ge 0$ and any $f \in \mathcal{F}$

$$(4.15c) \qquad \tau_k(f) \le C_0\|f\| \ .$$

Note that these requirements are satisfied for the choice $\tau_k(f) = \|(I - \mathcal{R}_k\mathcal{D}_k)f\|$ in (4.15a).

Let us suppose now that the sequence is π-convergent and that the limit function f_k^∞,

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

is of such "regularity" that for large ℓ

$$\tau_\ell(f_k^\infty) = C_\tau(f_k^\infty)(h_\ell)^\alpha + o\big((h_\ell)^\alpha\big), \quad \alpha > 0 \ .$$

Our assumptions on $\tau_\ell(\cdot)$ imply that it is continuous in $\mathcal{F}$. If also

$$\tau_\ell(\Pi_k^\ell f) = \tau_\ell(f_k^\infty) + o\big((h_\ell)^\alpha\big) = C_\tau(f_k^\infty)(h_\ell)^\alpha + o\big((h_\ell)^\alpha\big),$$

then for large ℓ we expect

$$\tau_{\ell+1}(\Pi_k^{\ell+1} f) \le (\widetilde{q})^\alpha \tau_\ell(\Pi_k^\ell f) \ ,$$

where $\widetilde{q}$ is defined in (4.14) and $0 < (\widetilde{q})^\alpha < 1$.

Let us denote $\tau_{\ell+1}(\mathcal{R}_\ell v^\ell)$ by $\sigma_\ell(v^\ell)$ and observe that

$$(4.15d) \qquad \Pi_k^\ell = \mathcal{R}_\ell A_k^\ell \mathcal{D}_k,$$

where A_k^ℓ is the operator defined by (2.12a). Using this notation we can rewrite the asymptotic relation above for $v^k = \mathcal{D}_k f$ as

$$\sigma_\ell(A_k^\ell v^k) \le q\sigma_{\ell-1}(A_k^{\ell-1} v^k) \quad , \quad q = (\widetilde{q})^\alpha.$$

Since $\tau_k(\cdot)$ is a semi-norm in $\mathcal{F}$ we get from (4.15c) that

$$(4.15e) \quad \sigma_k(\mathcal{D}_k f) = \tau_{k+1}(\mathcal{R}_k\mathcal{D}_k f) \le \tau_{k+1}(f) + \tau_{k+1}(\mathcal{R}_k\mathcal{D}_k f - f)$$

$$\le \tau_{k+1}(f) + C_0\|\mathcal{R}_k\mathcal{D}_k f - f\| \le \tau_{k+1}(f) + C_0\tau_k(f) \ .$$

48

and therefore by (4.15b) $\sigma_k(\mathcal{D}_k f) \to 0$ as $k \to \infty$.

Next we present a sufficient condition for π-convergence. Prior to that let us generalize $\sigma_\ell(v^\ell)$ to denote any upper bound of $\tau_{\ell+1}(\mathcal{R}_\ell v^\ell)$, i.e.

$$(4.16a) \qquad \sigma_\ell(v^\ell) \geq \tau_{\ell+1}(\mathcal{R}_\ell v^\ell) \geq \|(I - \mathcal{R}_{\ell+1}\mathcal{D}_{\ell+1})\mathcal{R}_\ell v^\ell\| \quad , \quad v^\ell \in V^\ell$$

such that for any $f \in \mathcal{F}$

$$(4.16b) \qquad \sigma_\ell(\mathcal{D}_\ell f) \to 0 \quad as \quad \ell \to \infty;$$

we assume that $\sigma_\ell(v^\ell)$ is a semi-norm in V^ℓ which satisfies

$$(4.16c) \qquad \sigma_\ell(v^\ell) \leq C_0 |v^\ell|_\ell$$

for all $\ell \geq 0$ and any $v^\ell \in V^\ell$. Note that these requirements are satisfied for the choice $\sigma_\ell(v^\ell) = \tau_{\ell+1}(\mathcal{R}_\ell v^\ell)$ in (4.16a).

Motivated by the above analysis we now formulate the following condition on the sequence $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$:

Definition 4.3. (Uniformly) σ-Contractive Sequence.

We say that $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$ is σ-contractive if there exists $0 < q < 1$ such that for all $k \geq 0$ and any $v^k \in V^k$

$$(4.16d) \qquad \sigma_\ell(A_k^\ell v^k) \leq C_1^k q^{\ell-k} \sigma_k(v^k) \qquad for \quad all \quad \ell \geq k+1;$$

here C_1^k is a constant which may depend on k but not on v^k. If in (4.16d)

$$(4.16e) \qquad \sup_{k \geq 0}\{C_1^k\} \leq C_\sigma < \infty,$$

then we say that the sequence is *uniformly* σ-contractive

Theorem 4.4.

If the sequence of approximation $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$ *is σ-contractive then it is also π-convergent.*

Proof: Using the identity

$$\Pi_k^{L+n} - \Pi_k^L = \sum_{\ell=L}^{L+n-1} (\Pi_k^{\ell+1} - \Pi_k^\ell) = \sum_{\ell=L}^{L+n-1} (\mathcal{R}_{\ell+1}\mathcal{D}_{\ell+1} - I)\Pi_k^\ell$$

we get from (4.15a), (4.15d) and (4.16) with $v^k = \mathcal{D}_k f$ that

$$\|(\Pi_k^{L+n} - \Pi_k^L)f\| \leq \sum_{\ell=L}^{L+n-1} \|(\mathcal{R}_{\ell+1}\mathcal{D}_{\ell+1} - I)\Pi_k^\ell f\|$$

$$\leq \sum_{\ell=L}^{L+n-1} \tau_{\ell+1}(\Pi_k^\ell f) \leq \sum_{\ell=L}^{L+n-1} \sigma_\ell(A_k^\ell \mathcal{D}_k f)$$

$$\leq C_1^k \sigma_k(\mathcal{D}_k f) \sum_{\ell=L}^{L+n-1} q^{\ell-k} .$$

Since $0 < q < 1$ we get for any $n \geq 0$

$$(4.17a) \qquad \|(\Pi_k^{L+n} - \Pi_k^L)f\| \leq C_1^k \sigma_k(\mathcal{D}_k f)\frac{q^{L-k}}{1-q} \underset{L\to\infty}{\longrightarrow} 0 ,$$

which shows that $\{\Pi_k^L f\}_{L=k}^\infty$ is a Cauchy sequence for any $f \in \mathcal{F}$.

$\sqcap$

Corollary 4.2.

If $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is a (uniformly) σ-contractive sequence of approximation then it is also (uniformly) π-stable.

Proof: If the sequence of approximation is σ-contractive, then by Theorem 4.4 it is π-convergent and therefore by Theorem 4.2 it is also π-stable. We can get an expression for the constant $\tilde{C}_k$ in the definition 3.5 by taking $L = k$ in (4.17a), and then use (4.16c) and (3.13b) to obtain

$$\|(\Pi_k^{k+n} - \mathcal{R}_k\mathcal{D}_k)f\| \leq \frac{C_1^k}{1-q}\sigma_k(\mathcal{D}_k f) \leq \frac{C_1^k}{1-q}C_0 C_A\|f\|;$$

which implies that for all $n \geq 0$

$$(4.17b) \qquad \|\Pi_k^{k+n}f\| \leq \|\mathcal{R}_k\mathcal{D}_k f\| + \|(\Pi_k^{k+n} - \mathcal{R}_k\mathcal{D}_k)f\|$$

$$\leq C_A\left(1 + \frac{C_0 C_1^{\prime k}}{1-q}\right)\|f\| =: \tilde{C}_k\|f\|.$$

If the sequence is uniformly σ-contractive, then $C_1^k \leq C_\sigma$ and we get in (4.17b)

$$(4.17c) \qquad \tilde{C}_k \leq C_A\left(1 + \frac{C_0 C_\sigma}{1-q}\right) =: \tilde{C}_\pi,$$

which shows that the sequence is uniformly π-stable.

$\sqcap$

50

Lemma 4.2.

If there exist $0 < q < 1$, a constant C_Δ and a convergent series $\sum\limits_{\ell=1}^{\infty} \Delta_\ell < \infty$ of positive numbers such that for any $v^k \in V^k$ and all $\ell \geq k+1$

$$(4.18a) \qquad \sigma_\ell(A_k^\ell v^k) \leq (1 + C_\Delta \Delta_\ell) q \sigma_{\ell-1}(A_k^{\ell-1} v^k) ,$$

then condition (4.16d) is satisfied with

$$(4.18b) \qquad C_1^k = \exp(C_\Delta \sum_{m=k+1}^{\infty} \Delta_m) \leq \exp(C_\Delta \sum_{m=1}^{\infty} \Delta_m) =: C_\sigma .$$

and the sequence $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$ is uniformly σ-contractive.

Proof: Since $1 + C_\Delta \Delta_\ell \leq \exp(C_\Delta \Delta_\ell)$ we get in (4.18a)

$$
\begin{aligned}
\sigma_\ell(A_k^\ell v^k) &\leq \exp(C_\Delta \Delta_\ell) q \sigma_{\ell-1}(A_k^{\ell-1} v^k) \\
&\leq \exp\big[C_\Delta(\Delta_\ell + \Delta_{\ell-1})\big] q^2 \sigma_{\ell-2}(A_k^{\ell-2} v^k) \leq \ldots \\
&\leq \exp(C_\Delta \sum_{m=k+1}^{\ell} \Delta_m) q^{\ell-k} \sigma_k(v^k) \\
&\leq C_1^k q^{\ell-k} \sigma_k(v^k) .
\end{aligned}
$$

$\sqcap$

Using Lemma 4.3 we now formulate the following sufficient condition (4.19), which is easier to verify.

Corollary 4.3.

If there exist $0 < q < 1$, a constant C_Δ and a convergent series $\sum\limits_{\ell=1}^{\infty} \Delta_\ell < \infty$ of positive numbers such that for all $\ell \geq k+1$ and any $v^{\ell-1} \in V^{\ell-1}$

$$(4.19) \qquad \sigma_\ell(P_{\ell-1}^\ell v^{\ell-1}) \leq (1 + C_\Delta \Delta_\ell) q \sigma_{\ell-1}(v^{\ell-1}) ,$$

where $P_{\ell-1}^\ell$ is the prediction operator (3.7b), then the sequence of approximation $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$ is uniformly σ-contractive.

Proof: For any $v^k \in V^k$ take in (4.19)

$$v^{\ell-1} = A_k^{\ell-1} v^k ;$$

51

then

$$P_{\ell-1}^{\ell} A_k^{\ell-1} v^k = A_k^{\ell} v^k$$

and (4.19) becomes (4.18a). Hence by Lemma 4.3 we get that condition (4.16) is satisfied with $C_1^k \leq C_\sigma$ which are defined in (4.18b).

$$\sqcap$$

Example 4.1.

Let us assume that for any $f \in \mathcal{F}$ and all $k \geq 0$

$$(4.20a) \qquad \|(I - \mathcal{R}_k \mathcal{D}_k)f\| \leq \tau_k(f) = C_\tau(f)(h_k)^\alpha, \quad \text{for some} \quad \alpha > 0 ,$$

and consider the linear operators T_k

$$T_k = (h_k)^{-\alpha}(\mathcal{R}_k \mathcal{D}_k - I) : \mathcal{F} \longrightarrow \mathcal{F} .$$

It follows from (4.20a) that

$$\|T_k f\| \leq C_\tau(f) ,$$

and from (3.12) that for all $f \in \mathcal{F}$

$$\|T_k f\| \leq (h_k)^{-\alpha}(1 + C_A)\|f\| .$$

Hence by the principle of uniform boundedness there exists a constant C_T (independent of k and f) such that

$$\|T_k f\| \leq C_T \|f\| ,$$

i.e.

$$\|(\mathcal{R}_k \mathcal{D}_k - I)f\| \leq C_T (h_k)^\alpha \|f\| .$$

This shows that if $\tau_k(f)$ in (4.20a) is an error bound so is

$$(4.20b) \qquad \tilde{\tau}_k(f) = C_T (h_k)^\alpha \|f\| ,$$

and that for any $f \in \mathcal{F}$

$$(4.20c) \qquad \|\mathcal{R}_k \mathcal{D}_k f\| \leq \|f\| + \|(\mathcal{R}_k \mathcal{D}_k - I)f\| \leq \left[1 + C_T (h_k)^\alpha\right]\|f\| .$$

For

$$(4.21a) \qquad \sigma_\ell(v^l) = \tilde{\tau}_{\ell+1}(\mathcal{R}_\ell v^\ell) = C_T (h_{\ell+1})^\alpha \|\mathcal{R}_\ell v^\ell\|$$

we get from (4.20c) and (4.14) that

$$(4.21b) \quad \sigma_\ell(P_{\ell-1}^\ell v^{\ell-1}) = C_T(h_{\ell+1})^\alpha \|\mathcal{R}_\ell \mathcal{D}_\ell \mathcal{R}_{\ell-1} v^{\ell-1}\|$$

$$\leq C_T(h_{\ell+1})^\alpha [1 + C_T(h_\ell)^\alpha] \|\mathcal{R}_{\ell-1} v^{\ell-1}\|$$

$$= [1 + C_T(h_\ell)^\alpha] \cdot \left(\frac{h_{\ell+1}}{h_\ell}\right)^\alpha \cdot C_T(h_\ell)^\alpha \|\mathcal{R}_{\ell-1} v^{\ell-1}\|$$

$$\leq [1 + C_T(h_\ell)^\alpha] \cdot (\widetilde{q})^\alpha \cdot \sigma_{\ell-1}(v^{\ell-1}) .$$

The above inequality shows that condition (4.19) is satisfied with $C_\Delta = C_T$,

$$(4.21c) \qquad 0 < q = (\widetilde{q})^\alpha < 1 ,$$

and

$$(4.21d) \qquad \Delta_\ell = (h_\ell)^\alpha, \quad \sum_{\ell=1}^\infty \Delta_\ell \leq (\sum_{\ell=1}^\infty q^\ell)(h_0)^\alpha = \frac{q}{1-q}(h_0)^\alpha ;$$

thus, by Corollary 4.3, $\{(\mathcal{R}_\ell \mathcal{D}_\ell)\}$ is uniformly σ-contractive and consequently by Theorem 4.4 it is a π-convergent sequence.

Example 4.2

Let us consider the pointvalue discretization of Example 1.1.1 in the Banach space $\mathcal{F} = C^0[0,1]$ with

$$\|f\| = \max_{0 \leq x \leq 1} |f(x)| =: \|f\|_\infty,$$

and define

$$h_k = \max_{1 \leq i \leq J_k - 1} |x_{i+1}^k - x_i^k|.$$

We assume that $\{\mathcal{D}_k\}$ is nested and that there exists $0 < \tilde{q} < 1$ such that for all $k \geq 1$

$$h_k \leq \tilde{q} h_{k-1}.$$

For any $f \in \mathcal{F}$ let us consider the sequence of cosmetic refinement (4.12) which is defined by

$$(4.22a) \qquad f_k^k(x) = I^k(x; \mathcal{D}_k f),$$

$$(4.22b) \qquad f_k^{\ell+1}(x) = I^{\ell+1}(x; \mathcal{D}_{\ell+1} f_k^\ell) \quad , \quad \ell = k, \ldots, \infty.$$

In the following we take $I^k(x; .)$ to be the piecewise-polynomial interpolation (1.25b) and derive a sufficient condition for convergence in the maximum norm of the sequence above. The piecewise-polynomial interpolation (1.25b) has an error bound of the form

$$(4.23a) \qquad \|I^k(\cdot; f) - f\|_\infty \leq \text{const} \cdot \omega(f; h_k) =: \tau_k(f),$$

where $\omega(f; \delta)$ is the modulus of continuity of f

$$(4.23b) \qquad \omega(f; \delta) = \max_{|x-y| \leq \delta} |f(x) - f(y)|.$$

It is easy to see (e.g. by using the Newton form of $p_i^k(x; v^k)$ in (1.25b)) that

$$\omega\big(I^\ell(\cdot; v^\ell); h_{\ell+1}\big) = \max_{|x-y| \leq h_{\ell+1}} |I^\ell(x; v^\ell) - I^\ell(y; v^\ell)| \leq \text{const} \cdot \max_{1 \leq i \leq J_\ell - 1} |v_{i+1}^\ell - v_i^\ell| \quad ;$$

therefore we can take in (4.16) $\sigma_\ell(v^\ell) = \text{const} \cdot \hat{\sigma}(v^\ell)$ where

$$(4.24a) \qquad \hat{\sigma}(v^\ell) = \max_{1 \leq i \leq J_\ell - 1} |v_{i+1}^\ell - v_i^\ell|,$$

and conclude by Theorem 4.4 that if there exists $0 < q < 1$ such that for all $k \geq 0$ ans any $v^k \in V^k$

$$(4.24b) \qquad \hat{\sigma}(A_k^\ell v^k) \leq C_1^k q^{\ell-k} \hat{\sigma}(v^k) \qquad \text{for all } \ell \geq k,$$

then the sequence f_k^ℓ converges in the maximum norm as $\ell \to \infty$ to a function in $\mathcal{F}$, which we denote by f_k^∞. Hence f_k^∞ is continuous and by Theorem 4.3

$$(4.25a) \qquad \mathcal{D}_k f_k^\infty = \mathcal{D}_k f = v^k,$$

which implies that

$$(4.25b) \qquad I^{k,\infty}(x; v^k) =: f_k^\infty(x)$$

is an interpolation of v^k. In the next subsection 4.D we prove that $I^{k,\infty}(x; v^k)$ is an hierarchial interpolation.

Let us consider now the case where we use the piecewise-polynomial interpolation (1.25) with the same choice of stencil for all intervals $[x_i^k, x_{i+1}^k]$ of the dyadic sequence of uniform grids (1.37). In this case $\tilde{q} = \frac{1}{2}$ in (4.14) and the sequence of cosmetic refinement (4.22) can be described in terms of the pointvalues $(\bar{f}_k^\ell)_i = f_k^\ell(x_i^\ell)$ by

$$(4.26) \qquad \begin{cases} (\bar{f}_k^{\ell+1})_{2i} = (\bar{f}_k^\ell)_i, \\ (\bar{f}_k^{\ell+1})_{2i-1} = \sum_m \beta_{2m-1} (\bar{f}_k^\ell)_{i-m} \quad , \quad \sum_m \beta_{2m-1} = 1. \end{cases}$$

Applying the results of [DaL1] and [DaL2] to this case we can conclude that if the sequence $\bar{f}_k^\ell = A_k^\ell \bar{f}_k^k$ satisfies condition (4.24), then $f_k^\infty(x)$ is uniformly Hölder-continuous

$$(4.27a) \qquad |f_k^\infty(x) - f_k^\infty(y)| \leq C_2 |x - y|^\alpha$$

where

(4.27b)
$$\alpha = \min(1, -\log_2 q).$$

and C_2 is a constant which depends on the coefficients $\{\beta_{2m-1}\}$ and the initial data $\bar{f}_k^k$; observe that the exponent α in (4.27b) is computed from $q = (\tilde{q})^\alpha = (\frac{1}{2})^\alpha$.

Applying the results of [CDM] and [DL] to (4.26) we can conclude that condition (4.24) is *necessary*, as well as sufficient, for convergence in the maximum norm.

Remark 4.3. Taking equality in (4.15a) and (4.16a) we get

$$\sigma_\ell(v^\ell) = \|(I - \mathcal{R}_{\ell+1}\mathcal{D}_{\ell+1})\mathcal{R}_\ell v^\ell\|.$$

The analysis at the beginning of this subsection suggests that if we restrict ourselves to the case where $f_k^\infty \in \mathcal{F}$ is slightly "more regular" than what is required of elements of $\mathcal{F}$, then condition (4.16d) is *necessary*, as well as sufficient, for convergence in the norm of $\mathcal{F}$. Because of the abstract nature of our framework it is difficult to express what "regularity" is. For this reason we formulated our condition in (4.16a) via the error bound $\tau_k(f)$, as it enables us to relate the rate of convergence q to the "regularity α" (as was done in (4.27)).

4.D. Hierarchial Form.

In this subsection we show that if the sequence of approximation $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is not hierarchial to begin with, then if the sequence is π-convergent it has an hierarchial form $\{(\mathcal{R}_k^H\mathcal{D}_k)\}$ which is obtained by cosmetic refinement. We also show that if $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is *uniformly* σ-contractive, then it generates an MR basis in $\mathcal{F}$.

Theorem 4.5.

If $\{(\mathcal{R}_k\mathcal{D}_k)\}$ is a π-convergent sequence of approximation, define

(4.28a)
$$\mathcal{R}_k^H : V^k \to \mathcal{F}$$

by

(4.28b)
$$\mathcal{R}_k^H \cdot v^k = \lim_{L \to \infty} \Pi_{k+1}^L \cdot \mathcal{R}_k v^k ;$$

then:

(i) $\mathcal{R}_k^H$ *is a reconstruction of $\mathcal{D}_k$ in $\mathcal{F}$;*

(ii) *the prediction operator* $(P^H)_k^{k+1}$ *of the sequence* $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ *is identical to that of the original sequence ,i.e.*

(4.29a)
$$(P^H)_k^{k+1} = \mathcal{D}_{k+1}\mathcal{R}_k^H = \mathcal{D}_{k+1}\mathcal{R}_k = P_k^{k+1} \; ;$$

(iii) *the sequence* $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ *is hierarchial :*

(4.29b)
$$(\mathcal{R}_{k+1}^H \mathcal{D}_{k+1})\mathcal{R}_k^H = \mathcal{R}_k^H \; .$$

Proof: It follows from Theorem 4.3 that

(4.29c)
$$\mathcal{D}_\ell \mathcal{R}_k^H v^k = \mathcal{D}_\ell \lim_{L \to \infty} \Pi_{k+1}^L \cdot \mathcal{R}_k v^k = \mathcal{D}_\ell \mathcal{R}_k v^k \quad \text{for} \quad \ell \le k+1 \; .$$

(i) Using (4.29c) with $\ell = k$ we get

$$\mathcal{D}_k \mathcal{R}_k^H = \mathcal{D}_k \mathcal{R}_k = I_k \; .$$

(ii) Using (4.29c) with $\ell = k+1$ we get

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

(iii) Using (4.29a) and the definition (4.28) we get for any $v^k \in V^k$

$$\mathcal{R}_{k+1}^H(\mathcal{D}_{k+1}\mathcal{R}_k^H v^k) = \mathcal{R}_{k+1}^H(\mathcal{D}_{k+1}\mathcal{R}_k v^k)$$
$$= \lim_{L \to \infty} \Pi_{k+2}^L \cdot \mathcal{R}_{k+1}(\mathcal{D}_{k+1}\mathcal{R}_k v^k) = \lim_{L \to \infty} \Pi_{k+1}^L \cdot \mathcal{R}_k v^k = \mathcal{R}_k^H \cdot v^k \; .$$

$$\sqcap$$

We refer to $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ as the hierarchial form of $\{(\mathcal{R}_k \mathcal{D}_k)\}$.

Corollary 4.4.

Let $\{(\mathcal{R}_k \mathcal{D}_k)\}$ be a π-convergent sequence, then:
(i) *The MR scheme of the hierarchial form is the same as the original one, and*

(4.30a)
$$\mathcal{R}_L^H v^L = \mathcal{R}_L^H \mathcal{D}_L f = \mathcal{R}_0^H \mathcal{D}_0 f + \sum_{k=1}^{L} \mathcal{R}_k^H e^k(f) \; ,$$

(4.30b)
$$e^k(f) = \mathcal{D}_k(I - \mathcal{R}_{k-1}\mathcal{D}_{k-1})f \; .$$

(ii) *The MR scheme is stable and*

$$(4.31a) \qquad |\delta(v^L)|_{H,L} \leq |\delta(v^0)|_{H,0} + \sum_{k=1}^{L} \langle \delta(d^k) \rangle_{H,k}$$

where $\delta(\cdot)$ denotes the perturbation in $(\cdot)$, and

$$(4.31b) \qquad |v^k|_{H,k} = \|\mathcal{R}_k^H v^k\|, \quad \langle d^k \rangle_{H,k} = |E^k d^k|_{H,k} \ .$$

Proof: (i) It follows from (4.29a) that

$$\mathcal{D}_k(I - \mathcal{R}_{k-1}^H \mathcal{D}_{k-1}) = \mathcal{D}_k(I - \mathcal{R}_{k-1} \mathcal{D}_{k-1})$$

and therefore the prediction error (3.8d) for the hierarchial form $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ is the same as that of $\{(\mathcal{R}_k \mathcal{D}_k)\}$ in (4.30b). Using (4.29b) we get, as in Lemma 4.1, that

$$(\mathcal{R}_L^H \mathcal{D}_L) \cdots (\mathcal{R}_{k+1}^H \mathcal{D}_{k+1}) \mathcal{R}_k^H = \mathcal{R}_k^H \ ,$$

and therefore (4.30a) follows from (3.10b).

(ii) Since $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is π-convergent it follows from Theorem 4.2 that it is π-stable, and therefore by Theorem 3.4 the associated MR scheme is likewise stable. The bound (4.31) follows from Corollary 4.1.

$\sqcap$

Remark 4.4. It follows from (4.13c) that for any $v^k \in V^k$

$$|v^k|_{H,k} = \|\mathcal{R}_k^H v^k\| \leq C_k |v^k|_k, \quad |v^k|_k = \|\mathcal{R}_k v^k\| \ ,$$

and that for any $f \in \mathcal{F}$

$$\|\mathcal{R}_k^H \mathcal{D}_k f\| \leq C_k C_A \|f\| \ ,$$

where C_k is defined in (3.17a) and C_A is defined in (3.13b).

Remark 4.5. If $p \in \mathcal{F}$ is such that for all $\ell \geq 0$

$$\mathcal{R}_\ell \mathcal{D}_\ell p = p \ ,$$

then for all k and $L \geq k$

$$\Pi_k^L p = p$$

and consequently

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

57

This shows that the hierarchial form $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ is of the same "accuracy" as the original sequence $\{(\mathcal{R}_k \mathcal{D}_k)\}$.

At this point we have shown that if $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is a π-convergent sequence of approximation, then it has an hierarchial form $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$. The MR scheme (2.7) – (2.9) of the hierarchial form is identical to that of the original one, and it is likewise stable. The sequence $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ has the same accuracy as $\{(\mathcal{R}_k \mathcal{D}_k)\}$ (Remark 4.5), and by Corollary 4.3 we get that

$$(4.32a) \qquad \mathcal{R}_L^H \mathcal{D}_L f = \sum_i \hat{f}_i^0 \varphi_i^0 + \sum_{k=1}^{L} \sum_j d_j^k(f) \psi_j^k$$

where

$$(4.32b) \qquad \varphi_i^k = \mathcal{R}_k^H \eta_i^k = \lim_{L \to \infty} \Pi_{k+1}^L \cdot \mathcal{R}_k \eta_i^k \ ,$$

$$(4.32c) \qquad \psi_j^k = \mathcal{R}_k^H \mu_j^k = \lim_{L \to \infty} \Pi_{k+1}^L \cdot \mathcal{R}_k \mu_j^k \ ,$$

and $d^k(f)$ are the original scale-coefficients

$$(4.32d) \qquad d^k(f) = G_k e^k(f) = G_k \mathcal{D}_k (I - \mathcal{R}_{k-1} \mathcal{D}_{k-1}) f \ .$$

From Theorem 4.4 we conclude that if $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is σ-contractive, then it is also π-convergent, and consequently it has an hierarchial form $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ which satisfies (4.31)-(4.32). In order to be able to conclude by Theorem 4.1 that

$$(4.32e) \qquad \mathcal{B}_M = \left(\{\varphi_i^0\}, \ \{\{\psi_j^k\}_j\}_{k=1}^{\infty} \right) \ ,$$

is a multiresolution basis for $\mathcal{F}$, we have to show that the hierarchial form is also a sequence of approximation (Definition 3.4). To do this we assume further that the sequence is *uniformly* σ-contractive (in which case it is also uniformly π-stable) and prove:

Theorem 4.6.

Let $\{(\mathcal{R}_k \mathcal{D}_k)\}$ *be a sequence of approximation which is uniformly σ-contractive, then:*

(i) *The hierarchial form* $\{(\mathcal{R}_k^H \mathcal{D}_k)\}$ *is a sequence of approximation;*

(ii) $\mathcal{B}_M$ *in (4.32e) is a multiresolution basis of $\mathcal{F}$.*

Proof: (i) From (4.13c) and (4.17c) we get that

$$\|\mathcal{R}_k^H \mathcal{D}_k f\| \le \tilde{C}_\pi \|f\|,$$

which proves (3.11b). To prove (3.11c) we take $L = k$ in (4.17a) and use the uniformity to obtain

$$\|(\Pi_{k+1}^{k+n}\mathcal{R}_k - \mathcal{R}_k)\mathcal{D}_k f\| \le \frac{C_\sigma}{1-q}\sigma_k(\mathcal{D}_k f);$$

letting $n \to \infty$ in the above inequality we get that

$$\|(\mathcal{R}_k^H - \mathcal{R}_k)\mathcal{D}_k f\| \le \frac{C_\sigma}{1-q}\sigma_k(\mathcal{D}_k f) ,$$

and therefore using (4.15b) and (4.16b) we get

$$\|(\mathcal{R}_k^H \mathcal{D}_k - I)f\| \le \|(\mathcal{R}_k^H - \mathcal{R}_k)\mathcal{D}_k f\| + \|(\mathcal{R}_k \mathcal{D}_k - I)f\|$$

$$\le \frac{C_\sigma}{1-q}\sigma_k(\mathcal{D}_k f) + \tau_k(f) \to 0 \quad as \quad k \to \infty .$$

(ii) is now a direct consequence of Theorems 4.5 and 4.1.

$$\sqcap$$

Remark 4.6. Let us take $\sigma_k(\mathcal{D}_k f) = \tau_{k+1}(\mathcal{R}_k \mathcal{D}_k f)$ in which case we obtained in (4.15e) that

$$\sigma_k(\mathcal{D}_k f) \le \tau_{k+1}(f) + C_0 \tau_k(f) .$$

Using the above inequality in Theorem 4.6 we get the following error bound for the hierarchial form

(4.33) $$\|(\mathcal{R}_k^H \mathcal{D}_k - I)f\| \le \frac{C_\sigma}{1-q}\big[\tau_{k+1}(f) + C_0 \tau_k(f)\big] + \tau_k(f) .$$

This shows, as in Remark 4.5, that the approximations $(\mathcal{R}_k^H \mathcal{D}_k)$ and $(\mathcal{R}_k \mathcal{D}_k)$ are of comparable accuracy.

Remark 4.7 We observe that (4.18) and (4.19) imply *uniform* σ-contraction which we assumed in Theorem 4.6. From (4.33) we see that if we replace C_σ by C_1^k and give up the assumption of uniform σ-contraction, then we have to add the requirement

$$C_1^k \tau_k(f) \to 0 \quad as \quad k \to \infty$$

instead. Note that uniformity is not needed for π-convergence, and neither for the stability of the associated MR scheme.

59

Remark 4.8. Letting $L \to \infty$ in (4.2d) and (4.2e) we get for $\{\varphi_i^k\}$ and $\{\psi_j^k\}$ in (4.32) that

$$(4.34a) \qquad \psi_j^k = \sum_i (\hat{E}_k)_{i,j} \varphi_i^k,$$

$$(4.34b) \qquad \varphi_\ell^{k-1} = \sum_i (\hat{P}_{k-1}^k)_{i,\ell} \varphi_i^k;$$

here $\hat{E}_k$ is the matrix representation of E^k and $\hat{P}_{k-1}^k$ is the matrix representation of the *original* prediction operator (4.29a). Note that when $\hat{P}_{k-1}^k$ is the Töplitz-like matrix in (1.9) , then (4.34b) is identical to (1.41d), in which case all $\varphi_i^k(x)$ are generated from a single function $\varphi(x)$ which is the solution of the dilation equation (1.41b). Hence (4.34b) is a generalization of the relation (1.41d) (for wavelets) to unstructured grids.

5. Nested Discretization of Functions

In this section we take $\mathcal{F}$ to be a space of functions

$$(5.1) \qquad \mathcal{F} = \{f | f : X \subset \mathbf{R}^m \to \mathbf{R}^n\}$$

and describe several cases of nested discretization. We do so in an increasing order of "domain of dependence": In subsection 5.A we examine nestedness of pointvalue discretization; in subsection 5.B we describe discretization by local (weighted) averages, and finally in subsection 5.C we consider global (spectral) discretization.

To simplify our presentation we take in (5.1) $n = 1$, and unless otherwise stated, we also take $m = 2$ and assume that X is a compact set, i.e. we consider scalar functions which are defined in a compact set of $\mathbf{R}^2$.

5.A. Discretization by Pointvalue

The most commonly used discretization in numerical analysis is that by pointvalues (Example 1.1), where we take $\mathcal{F} = C^0(X)$ and define the discretization $\mathcal{D}_k$ on the grid

$$(5.2a) \qquad X^k = \{x_i^k\}, \qquad x_i^k \in X,$$

by

$$(5.2b) \qquad \mathcal{D}_k f = \{f(x_i^k)\}, \qquad x_i^k \in X^k.$$

Clearly, if we form X^k from X^{k-1} by adding more points to it, then the sequence of discretization is nested (Definition 3.2).

Next let us consider the sequence $\{I^k(x; \mathcal{D}_k f)\}_{k=0}^{\infty}$ in the Banach space which consists of continuous functions in X with the maximum norm

$$(5.3) \qquad \|f\| = \|f\|_\infty =: \max_{x \in X} |f(x)|.$$

The conditions (3.11b) and (3.11c), which are required of a sequence of approximation in this space, state that for any f which is continuous in X

$$(5.4a) \qquad \max_{x \in X} |I^k(x; \mathcal{D}_k f)| \le C_A^k \cdot \max_{x \in X} |f(x)|,$$

and

$$(5.4b) \qquad \lim_{k \to \infty} \max_{x \in X} |I^k(x; \mathcal{D}_k f) - f(x)| = 0.$$

Since it is difficult to introduce a sense of refinement for completely unstructured grids, it is common practice in numerical analysis to work with triangulated meshes. Let us denote the triangles in such a mesh by $\{t_i^k\} =: T^k \subseteq X$ and denote by X^k the set of vertices in these triangles; we refer to T^k as a triangulation of X. We consider now a *dyadic* refinement sequence where each triangle t_i^{k-1} of T^{k-1} is divided into four triangles of T^k by connecting the midpoints of its sides; these midpoints become vertices in the triangulation T^k, and therefore they are added to X^{k-1} to form X^k; consequently the discretization (5.2) is nested. Observe that the 3 midpoints of the sides of each triangle are shared with its 3 side-neighbors (except at boundaries); thus the number of gridpoints is approximately doubled at each refinement.

Let $h_{i,k}$ denote the infimum over the radii of all the circles which contain the triangle t_i^k, and let h_k be the supremum of $h_{i,k}$ over all triangles in T^k. Clearly

$$(5.5) \qquad h_k = \frac{1}{2} h_{k-1}$$

and therefore $\tilde{q} = \frac{1}{2}$ in (4.14).

In the following we describe several techniques of interpolation for values at the vertices X^k of the triangulation T^k.

Example 5.1. Piecewise-Polynomial Interpolation.

Let $\mathcal{S}_i^k$ be a stencil of $s = r(r + 1)/2$ points of X^k which is assigned to the triangle $t_i^k \in T^k$ and includes its vertices for $r \ge 2$. Let $p_i^k(x; \mathcal{D}_k f)$ denote the

unique polynomial of degree $(r-1)$ which interpolates $f(x)$ at the points of the stencil $\mathcal{S}_i^k$, and define

$$(5.6) \qquad I^k(x; \mathcal{D}_k f) = p_i^k(x; \mathcal{D}_k f) \quad for \quad x \in t_i^k$$

Clearly this technique is exact for data of polynomial functions of degree less or equal $r-1$.

For $r = 2$ we get that the 3 points of $\mathcal{S}_i^k$ are necessarily the vertices of t_i^k, and $p_i^k(x; \mathcal{D}_k f)$ is the piecewise-linear function which interpolates f at these three points. Observe that in this case $\{I^k(x; \mathcal{D}_k f)\}$ is an hierarchial sequence of approximation.

Example 5.1.1. *Hierarchial Piecewise-quadratic interpolation.*

Take $r = 3$ and for each $t_i^k \in T^k$ identify the triangle $t_{i_*}^{k-1}$ to which it "belongs" by being a part of its division. Let $\mathcal{S}_i^k$ be the stencil of 6 points which is composed of the vertices of $t_{i_*}^{k-1}$ and the midpoints of its sides, and denote the quadratic polynomial which interpolates $f(x)$ at these 6 points by $p_{i_*}^{k-1}(x; \mathcal{D}_k f)$. The piecewise-polynomial interpolation (5.6) can be expressed in this case by

$$(5.7) \qquad I^k(x; \mathcal{D}_k f) = p_{i_*}^{k-1}(x; \mathcal{D}_k f) \quad for \quad x \in t_{i_*}^{k-1};$$

clearly this sequence of interpolation is hierarchial. Observe that this technique is rather efficient because there is no need to search for a stencil, and due to the fact that the same quadratic polynomial (5.7) serves 4 triangles of T^k.

Similar techniques of refinement that automatically provide a convenient stencil for higher-order hierarchial interpolation in $\mathbf{R}^2$ and $\mathbf{R}^3$ were developed within the context of Finite-Elemet Methods.

Example 5.1.2. *Stencil of closest neighbors.*

For any $r \geq 2$ let $\delta(\mathcal{S}_i^k, t_i^k)$ denote the largest Euclidian distance between points in $\mathcal{S}_i^k$ and the triangle t_i^k. The stencil of closest neighbors is the one which minimizes $\delta(\mathcal{S}_i^k, t_i^k)$ among all possible choices of $\mathcal{S}_i^k$. For $r = 2$, the 3 closest neighbors are the vertices of t_i^k and the resulting interpolation is the same as in Example 5.1. However for $r = 3$, the stencil of closest neighbors is different from that of Example 5.1.1. The interpolation of the closest neighbors is of better "quality" than that of Example 5.1.1, but at a greater computational cost: One has to find this stencil and to compute a different polynomial for each triangle of T^k.

Clearly the resulting sequence of interpolation is not hierarchial, but we expect it to have an hierarchial form in many cases of practical importance.

Example 5.1.3. *Adaptive choice of stencil.*

In this example we describe the Essentially Non-Oscillatory (ENO) interpolation method of [HC] and [Ab], which is designed to avoid the Gibbs phenomenon. This is accomplished, as in Example 1.1.2, by an adaptive (data-dependent) choice of stencil in (5.6), where we assign to t_i^k the stencil of $r(r+1)/2$ points (including the vertices of t_i^k) in which $f(x)$ is "smoothest" in some sense, e.g. one in which the following minimum

$$(5.8) \qquad \min_{S_i^k} \left| \frac{d^{r-1}}{dx^{r-1}} p_i^k(x; \mathcal{D}_k f) \right|,$$

over an appropriate set of candidate-stencils S_i^k, is obtained. Note that since $p_i^k(x; \mathcal{D}_k f)$, the interpolating polynomial of $f(x)$ in S_i^k, is of polynomial degree $(r-1)$, the terms to be minimized in (5.8) are constants (independent of x).

The ENO interpolation is a highly nonlinear operator which leads to compression algorithms that are highly effective for discontinuous data. Due to the strong nonlinearity of the operator, one should not use the encoding algorithm (2.7), but rather a "nonlinear" version of it which is described in [H4]. This encoding algorithm, which we refer to as an "error control" algorithm, allows one to specify the maximal error in the decompression (which is done at the cost of losing direct control over the rate of compression).

5.B. Discretization by Local Averages

Let $\{w_i^k(x)\}$,

$$(5.9a) \qquad w_i^k : X \subset \mathbf{R}^m \longrightarrow \mathbf{R}$$

be square integrable functions of compact support which satisfy

$$(5.9b) \qquad \int w_i^k(x) dx = 1,$$

and define the discretization $\mathcal{D}_k$ for $f \in \mathcal{F} = L_2(X)$ by the inner product

$$(5.9c) \qquad (\mathcal{D}_k f)_i = (f, w_i^k) =: \int f(x) w_i^k(x) dx.$$

We refer to $w_i^k(x)$ as a weight-function and to the discretization (5.9) as local (weighted) averages.

63

Definition 5.1. Nested Sequence of Functions.

Let Ω^k denote the linear span of $\{w_i^k\}$. We say that the sequence of functions $\{\{w_i^k\}\}_{k \geq 0}$ is nested if $\Omega^{k-1} \subset \Omega^k$ for all $k \geq 1$.

The following Lemma 5.1 can be viewed as a generalization to unstructured grids of Mallat's notion of multiresolution analysis ([Ma]).

Lemma 5.1.

If the sequence of weight-functions is nested, then the corresponding sequence of discretization (5.9) is also nested.

Proof:

$$\Omega^{k-1} \subset \Omega^k \Rightarrow w_i^{k-1} = \sum_\ell \alpha_{i,\ell}^k w_\ell^k \Rightarrow$$

$$(\mathcal{D}_{k-1} f)_i = \int f w_i^{k-1} dx = \sum_\ell \alpha_{i,\ell}^k \int f w_\ell^k dx$$

$$= \sum_\ell \alpha_{i,\ell}^k (\mathcal{D}_k f)_\ell =: (D_k^{k-1} \cdot \mathcal{D}_k f)_i.$$

$\sqcap$

Let us assume that the sequence $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is hierarchial and examine the relation between the weight-functions $\{w_i^k\}$, and the functions $\{\varphi_i^k\}$ and $\{\psi_j^k\}$ in (4.6b)-(4.6c).

Lemma 5.2.

Let $\{\mathcal{R}_k \mathcal{D}_k)\}$ be an hierarchial sequence, and and let φ_i^k and ψ_j^k be defined by (4.6b)-(4.6c), then:

(5.10a) $$(\varphi_i^k, w_\ell^k) = (\eta_i^k)_\ell,$$

and

(5.10b) $$(\psi_j^k, w_\ell^{k-1}) = 0.$$

Proof: It follows immediately from (4.6b) with the definition (5.9c) that

$$(\varphi_i^k, w_\ell^k) = (\mathcal{D}_k \varphi_i^k)_\ell = (\mathcal{D}_k \mathcal{R}_k \eta_i^k)_\ell = (\eta_i^k)_\ell$$

Similarly, since μ_j^k is in the null space of D_k^{k-1}, it follows from (3.7a) and (4.6c) that

$$(\psi_j^k, w_\ell^{k-1}) = (\mathcal{D}_{k-1} \psi_j^k)_\ell = (\mathcal{D}_{k-1} \mathcal{R}_k \mu_j^k)_\ell = (D_k^{k-1} \mu_j^k)_\ell = 0.$$

$\sqcap$

64

Corollary 5.1. *Let $\{\eta_i^k\}$ be the cardinal basis of S^k (1.3) i.e.*

(5.10c)
$$(\eta_i^k)_\ell = \delta_{i,\ell},$$

then $\{w_i^k\}$ and $\{\varphi_i^k\}$ are biorthonormal, i.e.

(5.10d)
$$(\varphi_i^k, w_\ell^k) = \delta_{i,\iota},$$

Observe that this result shows that biorthogonality is a consequence of (2.5), the nestedness of the weight-functions and the hierarchial form; having a dilation equation for $w(x)$ and $\varphi(x)$ is a particular way of getting nestedness and an hierarchial structure.

Next we show that, as in the case of biorthogonal wavelets in [CDF], the roles of the two systems $\{w_i^k\}$ and $\{\varphi_i^k\}$ are interchangeable; this introduces the notion of a *dual* MR scheme.

It follows from (4.9a) that $\{\varphi_i^k\}$ is a nested sequence; therefore, normalizing it to have an integral of 1 by

(5.11a)
$$\tilde{w}_i^k(x) = \varphi_i^k(x) / \int \varphi_i^k dx$$

we get that $\{\tilde{w}_i^k\}$ is a nested sequence of weight-functions. Hence by Lemma 5.1, $\{\tilde{\mathcal{D}}_k\}$,

(5.11b)
$$(\tilde{\mathcal{D}}_k f)_i = (f, \tilde{w}_i^k) = \int f(x)\tilde{w}_i^k(x)dx,$$

is a nested sequence of discretization.

Theorem 5.1. *Let $\tilde{R}_k$,*
$$\tilde{R}_k : S^k \to \mathcal{F}$$

be defined by

(5.12a)
$$\tilde{R}_k s^k = \sum_i s_i^k \tilde{\varphi}_i^k(x), \quad \tilde{\varphi}_i^k(x) = w_i^k(x) \cdot \int \varphi_i^k dx$$

then
(i) $\tilde{R}_k$ is a reconstruction of $\tilde{\mathcal{D}}_k$, i.e.

(5.12b)
$$\tilde{\mathcal{D}}_k \tilde{R}_k = I_k,$$

65

(ii) $(\mathcal{R}_k \mathcal{D}_k)$ is the adjoint of $(\tilde{\mathcal{R}}_k \tilde{\mathcal{D}}_k)$, i.e. for any $f, g \in \mathcal{F}$

$$(5.13) \qquad (\tilde{\mathcal{R}}_k \tilde{\mathcal{D}}_k f,\ g) = (f, \mathcal{R}_k \mathcal{D}_k g).$$

Proof: It follows from relation (5.10a) and the definitions (5.11) and (5.12) that

(i)
$$(\tilde{\mathcal{D}}_k \tilde{\mathcal{R}}_k s^k)_\ell = \sum_i s_i^k (\tilde{\varphi}_i^k, \tilde{w}_\ell^k) = \sum_i s_i^k (\varphi_i^k, w_\ell^k)$$
$$= \sum_i s_i^k (\eta_i^k)_\ell = s_\ell^k.$$

(ii)
$$(\tilde{\mathcal{R}}_k \tilde{\mathcal{D}}_k f, g) = \sum_i (\tilde{\mathcal{D}}_k f)_i (\tilde{\varphi}_i^k, g) = \sum_i (\tilde{w}_i^k, f)(\tilde{\varphi}_i^k, g)$$
$$= \sum_i (\varphi_i^k, f)(w_i^k, g) = (f, \sum_i (\mathcal{D}_k g)_i \varphi_i^k)$$
$$= (f, \mathcal{R}_k \mathcal{D}_k g).$$

$\sqcap$

Remark 5.1. The results of Lemma 5.2 and Theorem 5.1 are a generalization of the results of [CDF] for the case of wavelets. The duality between the two MR schemes is a direct consequence of (1.1): Taking the transpose of the matrices in (1.1) we get that

$$(5.14a) \qquad \hat{D}_k^{k-1} \hat{P}_{k-1}^k = I_{k-1} \Rightarrow (\hat{P}_{k-1}^k)^* (\hat{D}_k^{k-1})^* = I_{k-1},$$

and since

$$(5.14b) \qquad (\hat{P}_{k-1}^k)^* : S^k \longrightarrow S^{k-1},$$
$$(5.14c) \qquad (\hat{D}_k^{k-1})^* : S^{k-1} \longrightarrow S^k$$

we can define a dual MR scheme by taking $(\hat{P}_{k-1}^k)^*$ to be the decimation operator, and $(\hat{D}_k^{k-1})^*$ to be the prediction.

Example 5.2. Discretization by Cell-Averages.

As in Example 1.2 let $C'^k = \{c_i^k\}$ be a covering of X by disjoint cells,

$$(5.15a) \qquad \overline{\cup_i c_i^k} = X, \qquad c_i^k \cap c_j^k = \emptyset \quad for \quad i \ne j,$$

and define the weight functions

$$(5.15b) \qquad w_i^k(x) = \frac{1}{|c_i^k|} \chi_{c_i^k}(x), \qquad |c_i^k| = \int_{c_i^k} dx,$$

where $\chi_c(x)$ is the characteristic function of the cell c (1.34). In this case (5.9) becomes the cell-average discretization (1.29b). If, as in Example 2.1, we consider a refinement sequence $\{C^k\}_{k=0}^L$, in which C^k is formed from C^{k-1} by dividing each cell c_i^{k-1} into q disjoint cells $\{c_{i_\ell}^k\}_{i=1}^q$, then

$$(5.15c) \qquad w_i^{k-1}(x) = \frac{1}{|c_i^{k-1}|} \sum_{\ell=1}^q |c_{i_\ell}^k| \cdot w_{i_\ell}^k(x).$$

and the sequence of discretization is nested by Theorem 5.1.

The "natural" function space for cell-average discretization is $\mathcal{F} = L_1(X)$, and there (3.11b)-(3.11c) take the following form:

$$(5.16a) \qquad \int_X |(R_k \mathcal{D}_k f)(x)| dx \le C_A^k \int_X |f(x)| dx,$$

$$(5.16b) \qquad \lim_{k \to \infty} \int_X |(R_k \mathcal{D}_k f)(x) - f(x)| dx = 0.$$

In Example 1.2.1 we have described a general piecewise-polynomial reconstruction of the cell-average discretization for unstructured grids. Let us consider now the simplest case which is the piecewise-constant reconstruction (1.34). Clearly $\{(\mathcal{R}_k \mathcal{D}_k)\}$ is an hierarchial sequence of approximation, and therefore it generates a multiresolution basis (4.6) of $L_1(X)$,where

$$(5.17) \qquad \varphi_i^k(x) = \chi_{c_i^k}(x) \quad ;$$

the expression for $\psi_j^k(x)$ depends on the particular way we choose to project into the null space of the decimation operator in (1.31), and can be made to look like a "generalized Haar basis".

Let $h_{i,k}$ denote the infimum over the radii of all balls in $\mathbf{R}^m$ that contain the cell c_i^k, and let h_k be the supremum of $h_{i,k}$ over all the cells in C^k. We observe that the analysis of Example 4.1 indicates that the refinement sequence should satisfy not only $h_k \longrightarrow 0$ as $k \longrightarrow \infty$, but more than that, namely that $\sum_k (h_k)^\alpha < \infty$ for some (possibly small) $\alpha > 0$.

Example 5.2.1. *Triangulated Meshes.*

The refinement technique in Example 5.1 of connecting the midpoints of the sides of each triangle t_i^{k-1}, provides a convenient way to assign a stencil of 6 points to each of the 4 smaller triangles of T^k in t_i^{k-1}. This stencil is not of "closest neighbors", but the points are reasonably close, and this arrangement is computationally

efficient – there is no need to search for a stencil, and a single reconstruction serves 4 triangles.

In the case of cell-average discretization we take $c_i^k = t_i^k$ and look for refinement methods that generate convenient assignment of stencils of cells for the purpose of reconstruction. These are typically of the "dual mesh" type: e.g. one can start with a covering by hexagons, each containing 6 triangles, and then divide each triangle into 4 by connecting the midpoints of its sides as before. This results in a new system of smaller hexagons (4 times as many). To each triangle of T^k we assign the 6 triangles of the hexagon to which it belongs – this is a convenient assignment of stencil for 3-rd order accurate reconstruction from cell-averages.

Example 5.2.2. *Centered Stencils.*

In [HC] we present an hierarchial algorithm for the selection of a "centered" stencil, which is applicable even to completely unstructured coverings C^k in $\mathbf{R}^m$. In this context the "centered" stencil is defined as the one which minimizes the reconstruction error for the one-higher degree polynomials (i.e. degree r). This algorithm is of "crystal growth" type: starting with the cell c_i^k we begin to add successively, one cell at a time, to the cluster of cells that we have at the beginning of each step. The cell which is being added is selected from the set of all side-neighbors of the existing cluster by the requirment that it will minimize the reconstruction error of suitably chosen monomials.

This choice of "centered" stencil depends on the geometry of the covering C^k, but not on the data $\mathcal{D}_k f$, and therefore the reconstruction (1.32) is a linear operator. For $r \geq 2$ the sequence of approximation is not hierarchial, but we expect it to have an hierarchial form in many cases of practical importance.

Example 5.2.3. *Adaptive selection of stencils (ENO reconstruction).*

In [HC] we also present a modification of the "crystal growth" algorithm of the previous example 5.2.2, which is designed to assign a stencil $\mathcal{S}_i^k$ from the smooth part of $f(x)$, if available, to all cells c_i^k which are themselves in the smooth part of $f(x)$. This way a Gibbs-like phenomenon is avoided, and the resulting approximation is r-th order accurate everywhere, except at cells which contain a discontinuity. This is accomplished by selecting the cell from the set of side-neighbors which minimizes the derivatives of the so-defined reconstruction. We refer to this technique as Essentially Non-Oscillatory (ENO) reconstruction; observe that this technique is highly-nonlinear, and therefore (as in Example 5.1.3) it should be applied to data compression in a proper "Error Control" encoding (see [H4]).

We refer the reader to [Ab] for details of special ENO reconstruction techniques for triangulated meshes.

Remark 5.2. The problem of reconstruction from cell-averages comes up in the numerical solution of hyperbolic systems of conservation laws, where it is needed in order to ensure conservation on the discrete level – see e.g. [H2-3], [HEOC], [HC] and [Ab]. On the other hand, the new capability of representing the cell-averaged solution in a multiresolution basis, has been used recently to improve the efficiency of the numerical solution of the initial value-problem by evolving it in its multiresolution form – see e.g. [H3,5-6] and [EOZ]. Other techniques that use projection of the PDE on wavelet bases are described e.g. in [BMP], [LT] and [MR].

Example 5.3. Biorthogonal Wavelets

In this example we start with (1.37), the dyadic sequence of uniform grids in **R**, and show that under assumptions of "uniformity", the natural outcome of our framework is the biorthogonal wavelets of [CDF].

Relation (1.40a) of subsection 1.F. shows that

$$(5.18a) \qquad w_i^k(x) = \frac{1}{h_k} w \left(\frac{x - x_i^k}{h_k} \right),$$

where $w(x)$ is the solution of the dilation equation (1.36), is a nested sequence of weight functions; by Lemma 5.1, this leads to a nested sequence of discretization in (5.9c). Of particular interest are the weight-functions $w^{(s)}(x)$ which are generated by repeated convolution with $\chi_{[-\frac{1}{2},\frac{1}{2}]}$, i.e.

$$(5.18b) \qquad w_{(x)}^{(0)} = \delta(x)$$

$$(5.18c) \qquad w^{(s)} = w^{(s-1)} * \chi_{[-\frac{1}{2},\frac{1}{2}]}, \qquad s = 1,2,\dots \;\; ;$$

see e.g. [S],[CDF] and [H4]. For $s = 0$ we get discretization by pointvalue, which is reconstructed by interpolation. For $s = 1$ we get discretization by cell-averages; in Example 1.2.2 we showed that knowledge of cell-averages equivalent to knowledge of the pointvalues of the primitive function $F(x_j^k)$,

$$(5.19a) \qquad F(x) = \int_0^x f(y)dy,$$

and that the cell-average discretization can be reconstructed by

$$(5.19b) \qquad (R_k \mathcal{D}_k f)(x) = \frac{d}{dx} I^k(x; F^k),$$

69

where $I^k(x; F^k)$ is any interpolation of the pointvalues of the primitive function of $f(x)$. For $s = 2$ we get discretization by hat-function averages; in [ADH] we show that knowledge of the hat-averages is equivalent to knowledge of the pointvalues $H(x_j^k)$ of the "second primitive" function $H(x)$,

$$(5.20a) \qquad\qquad H(x) = \int_0^x \int_0^y f(z) dz dy,$$

and that the hat-average discretization can be reconstructed by

$$(5.20b) \qquad\qquad (R_k \mathcal{D}_k f)(x) = \frac{d^2}{dx^2} I^k(x; H^k),$$

where $I^k(x; H^k)$ is any interpolation of the pointvalues of the second-primitive function of $f(x)$.

In [H7] and [ADH] we showed that if the interpolation $I^k(x; f^k)$ of the pointvalues has an hierarchial form, so do the reconstruction from cell-averages in (5.19) and the reconstruction from hat-averages (5.20). Using the results of [DD] and [DGL] for piecewise-polynomial interpolation with the centered-stencil (1.27a) we could then conclude that the corresponding sequence of approximation in (5.19) and (5.20) also has an hierarchial form. Since the reconstruction is the same at all points and all levels of the dyadic sequence of uniform grids (1.37), the prediction operator has the representation (1.9) and consequently in (4.34b) we get (see [H7]) that

$$(5.21a) \qquad\qquad \varphi_\ell^{k-1}(x) = \sum_m \beta_{\ell - 2m} \varphi_m^k(x)$$

where

$$(5.21b) \qquad\qquad \varphi_i^k(x) = \varphi\left(\frac{x - x_i^k}{h_k} \right) \ .$$

This shows that $\varphi(x)$ satisfies the dilation equation

$$(5.21c) \qquad\qquad \varphi(x) = \sum_\ell \beta_\ell \varphi(2x - \ell),$$

and thus the biorthogonal wavelets of [CDF] in (1.41)-(1.42) are the hierarchial form of the corresponding piecewise-polynomial reconstruction.

When X is the interval $[0, 1]$, we use the centered stencil (1.27a) in the interior, and suitable one-sided stencils (1.27b) and (1.27c) near the respective boundaries $x = 0$ and $x = 1$. The hierarchial form of the so-modified piecewise-polynomial

reconstruction in (5.19) and (5.20) generates a multiresolution basis that consists of the same wavelet functions as in (5.21) for the interior of the interval, and special "boundary" functions for the vicinity of its endpoints; these functions are generated by cosmetic refinement with the nonuniform reconstruction which is introduced by (1.27a) and (1.27b). We refer the reader to [CDV] for modification of Daubechies' orthonormal wavelets for the interval.

Remark 5.3. In our framework we start with the discrete approximation $\{(\mathcal{R}_k \mathcal{D}_k)\}$ and obtain the basis in $\mathcal{F}$ by taking $L \to \infty$ in (4.1). However, in applications we use the piecewise-polynomial approximation $(R_L \mathcal{D}_L f)(x)$, and not the basis in $\mathcal{F}$; recall that both expansions have exactly the same scale-coefficients. The computational difficulty in using the basis (4.2) is a result of the fact that $\varphi(x)$ is defined by the limit (4.32b), and in general does not have a closed-form expression; furthermore, for order of accuracy r, $\varphi(x)$ is not in C^r, and one cannot use standard numerical quadrature rules. In [BCR], a special class of wavelets (which is known as "Coiflets") that satisfy an r-th order accurate one-point quadrature rule, were developed to circumvent this difficulty; unfortunately this is accomplished at the high computational cost of doubling the support of the wavelets to $(4r)$.

For all practical purposes it is not important to know the explicit expression of the hierarchial form, however knowledge of its *existence* is important, because it implies stability of the original MR scheme (Corollary 4.4).

Remark 5.4. In [H4] we raise the question whether there is any advantage of using a particular weight function rather than other for the discretization (5.14). We conclude that there is no real difference at parts where $f(x)$ is smooth; however there is an important difference in the information which is retained by the discretization about singularities of $f(x)$. In [H2]and [H4] (see also Example 1.2) we show that one can recover the location of jump-discontinuities in $f(x)$ to any polynomial accuracy, from the discrete data of its cell-averages. In [H4] and [ADH] we show that the location x^* of distributions $\delta(x - x^*)$ in $f(x)$, as well as location of jump-discontinuities, can be recovered to any polynomial accuracy from the discrete data of the hat-averages of $f(x)$.

Remark 5.5 When the weight functions $w_i^k(x)$ are defined by (5.18a) from a "mother" function $w(x)$, the discretization (5.9c) can be viewed as taking point-values of $\bar{f}^k(x)$, the "sliding-average" function of $f(x)$

$$(5.22a) \quad \bar{f}^k(x) = \frac{1}{h_k} \int f(y) w(\frac{y-x}{h_k}) dy = f * \bar{w}^k \;, \qquad \bar{w}^k(y) = \frac{1}{h_k} w\left(\frac{-y}{h_k}\right);$$

71

at the gridpoints of (1.37), i.e.

$$(5.22b) \qquad (\mathcal{D}_k f)_i = \bar{f}^k(x_i^k).$$

Hence the reconstruction $(\mathcal{R}_k \mathcal{D}_k f)(x) \approx f(x)$ is an approximate deconvolution of $f *$ $\bar{w}^k$), and the prediction operator P_{k-1}^k produces an approximation to the pointvalues of $f * \bar{w}^k$ at X^k from knowledge of $f * \bar{w}^{k-1}$ at the points of X^{k-1}.

Using the above interpretation we can derive multiresolution representation for data which corresponds to discretization on nested dyadic sequences of uniform cartesian grids in $\mathbf{R}^m$ by a "tensor-product" extension of the one-dimensional operators. We refer the reader to [BCR] and [HY] where the case $m = 2$, which corresponds to representation of matrices, is described in detail.

5.C. Global (Spectral) Discretization.

Let $\mathcal{F}$ be a Hilbert space with inner-product $(\ ,\ )$, and assume that $\{\varphi_\nu\}_{j=0}^\infty$ is an orthonormal basis in $\mathcal{F}$, i.e. for any $f \in \mathcal{F}$

$$(5.34a) \qquad f = \sum_{\nu=0}^\infty (f, \varphi_\nu)\varphi_\nu;$$

we refer to (f, φ_ν) as the (generalized) Fourier coefficients. Let us define the spectral discretization $\mathcal{D}_k$ by

$$(5.34b) \qquad (\mathcal{D}_k f)_i = (f, \varphi_i), \qquad 0 \le i \le J_k,$$

with monotone increasing $\{J_k\}$. Clearly $\{\mathcal{D}_k\}$ is a nested sequence of discretization which can be *trivially* reconstructed by

$$(5.34c) \qquad R_k v^k = \sum_{i=0}^{J_k} v_i^k \varphi_i^k.$$

It is easy to see that the scale-coefficients of the corresponding MR-scheme in this trivial case are

$$(5.35a) \qquad d^k = \{(f, \varphi_j)\}_{j=J_{k-1}+1}^{J_k},$$

and that the MR basis amounts to rewriting (5.34a) as

$$(5.35b) \qquad f = \sum_{i=0}^{J_0}(f, \varphi_i)\varphi_i + \sum_{k=1}^\infty \sum_{j=J_{k-1}+1}^{J_k}(f, \varphi_j)\varphi_j.$$

72

It is well known (see [GO]) that due to the global nature of the spectral discretization (5.34b), if $f(x)$ is discontinuous then

$$|(f, \varphi_\nu)| = O\left(\frac{1}{\nu}\right)$$

and $(R_k \mathcal{D}_k f)(x)$ in (5.34c) has a Gibbs phenomenon. This observation was the main motivation for developing the wavelet theory, which provides a tool for *local* scale-decomposition. We would like to point out that due to recent development of *local* nonlinear reconstruction techniques which recover accurate pointwise information from the global Fourier coefficients, spectral discretization *can* be made useful for purposes of data compression.

The question whether one can recover accurate pointwise information from global quantities, came up in the context of numerical solution of hyperbolic PDEs with discontinuous initial data. In this case schemes with formal order of accuracy $r \geq 2$ have a Gibbs phenomenon, and therefore they are not r-th order accurate in a pointwise sense. Mock (Sever) and Lax showed that, with some pre-processing of the initial data, one can ensure that the *first r-moments* of the numerical solution will approximate those of the exact solution to r-th order accuracy, and that *pointwise information can be recovered from these moments to the same order of accuracy, arbitrarily close to the discontinuity* – see [ML] and also [MMO]. This observation gave rise to development of techniques to obtain accurate pointwise information from Fourier coefficients in general, and also to nonlinear reconstruction techniques that identify the location of discontinuities and use polynomial approximation in-between – see e.g. [CGS] and [GT]. We can use such special piecewise-polynomial reconstruction techniques (instead of (5.34c)) to predict the given Fourier coefficients from those of the k-th level. Once we are satisfied with the accuracy of the prediction, we consider $R_k \mathcal{D}_k f$ to be an adequate approximation to $f(x)$; the description of $f(x)$ can then be compressed by storing the parameters that define this piecewise-polynomial reconstruction, instead of the given data of the Fourier coefficients.

Concluding Remarks.

(1) In this paper we show how to design an MR scheme which is suitable for a given application, and derive sufficient conditions that ensure its stability; it seems to us that these sufficient conditions are not far from being necessary. It is interesting to note that, unlike other fields of numerical analysis, the notion of π-convergence seems to be more manageable than that of π-stability – this is so because the question of stability here is represented by a product of matrices

73

which are not of the same size. Note also that under conditions of uniformity, π-convergence is implied by that of the single sequence $\lim_{L\to\infty} \Pi_1^L \mathcal{R}_0 u^0$, where the initial data u^0 is the unit-sequence $u_i^0 = \delta_{i,0}$. This can be tested numerically by applying the inverse MR transform (2.9) to compute $M^{-1}(\{u^0, 0, 0, \ldots, 0\})$ for large enough L; similar tests can be made for boundary terms.

(2) Our framework allows for adaptive MR schemes, which are needed for applications that involve "mixed data", such as image compression (where one has piecewise-smooth sections, as well as regular patterns and various kinds of "texture"). The stability theory of the present paper applies only to linear (data-independent) MR schemes. In [H4] and [ADH] we have presented special nonlinear encoding techniques that ensure "error control" in some particular cases. More work is needed in order to establish stability of adaptive MR schemes in more general cases.

(3) The notion of discretization in this paper is very general: It is defined as a mapping from the "continuum" $\mathcal{F}$ onto the "denumerable" V. Therefore $\mathcal{F}$ can be taken to be a family of operators, which enables us to develop a theory of multiresolution representation of operators. This direct approach may prove to be advantageous over the indirect derivation of [BCR] and [HY], where such a multiresolution representation of operators is obtained by tranforming both input and output into a MR basis.

(4) In section 5 we have examined the case where the data originates from discretization of functions which are basically piecewise-smooth. In this case our tool of prediction is basically interpolation (with possible adaptive treatment of neighborhoods of irregularities). We need a different approach for compression of data which is globally nonsmooth (e.g. fractal data). If this nonsmooth data is generated by some "rule" and we have a way to find an approximation to the "generator" from analysis of the scale coefficients, then we could possibly achieve data compression by representing this nonsmooth but predictable data by the parameters that are needed to define its approximate "generator".

Acknowledgements.

74

References

[Ab] R. Abgrall, "Design of an essentially nonoscillatory reconstruction procedure on finite element type meshes", ICASE report 91-84, (Dec. 1991), in revised form INRIA report No. 1592 (February 1992), submitted to Math. of Comp.

[ADH] F. Arandiga, R. Donat and A. Harten, "Multiresolution based on weighted averages of the hat function", UCLA CAM Report 93-34, September 1993.

[BCR] G. Beylkin, R. Coifman and V. Rokhlin, "Fast wavelet tranform and numerical algorithms. I", Comm. Pure Appl. Math., Vol. 44, pp. 141-183, 1991.

[BL] A. Brandt and A.A. Lubrecht, "Multilevel matrix multiplication and fast solution of integral equations", Jour. Comput. Phys., Vol. 90, pp. 348-370, 1990.

[BMP] E. Bacry, S. Mallet and G. Papanicolau, "A wavelet based space-time adaptive numerical methods for partial differential equations", Mathematical Modelling and Numerical Analysis, Vol. 26, pp. 703-834, 1992.

[CD] A. Cohen and N. Dyn, "Nonstationary subdivision schemes and multiresolution analysis", Preprint, December 1993.

[CDF] A. Cohen, I. Daubechies and J.-C. Feauveau, "Biorthogonal bases of compactly suported wavelets", Comm. Pure Appl. Math., Vol. 45, pp. 485-560, 1992.

[CDM] A. Caveretta, W. Dahmen and C. Miccheli, "Stationary subdivision", Mem. Amer. Math. Soc., Vol. 93, pp. 1-186, 1991.

[CDV] A. Cohen, I. Daubechies and P. Vial, "Wavelets on the interval and fast wavelets transforms", CEREMADE Report No. 9307, March 1993.

[CGS] W. Cai, D. Gottlieb and C.-W. Shu, "On one-sided filters for spectral Fourier approximations of discontinuous functions", SIAM J. Numer. Anal., Vol. 29, pp. 905-916, 1992.

[Da] I. Daubechies, "Orthonormal bases of compactly supported wavelets", Comm. Pure Appl. Math., Vol. 41, pp. 909-996, 1988.

[DaL1] I. Daubechies and J. Lagarias, "Two scale difference equations. I", SIAM J. Math. Anal., Vol. 22, pp. 1388-1410, 1991.

[DaL2] I. Daubechies and J. Lagarias", Two scale difference equations. II", SIAM J.

Math. Anal., Vol. 23, pp. 1031-1079, 1992.

[DD] G. Deslauriers and S. Duboc, "Symmetric iterative interpolation scheme", Constructive Approximation, Vol. 5, pp. 49-68, 1989.

[DGL] N. Dyn, J. A. Gregory and D. Levin, "Analysis of linear binary subdivision schemes for curve design", Constructive Approximation, Vol. 7, pp. 127-147, 1991.

[DH] R. Donat and A. Harten, "Data compression algorithms for locally oscillatory data: A preliminary report", UCLA Preprint, October 1992.

[DL] N. Dyn and D. Levin, "Interpolating Subdivision Schemes for the Generation of Curves and Surfaces ", in Multivariate Interpolation and Approximation, W. Haussman and K. Jetter, Eds., Birkhauser Verlag, Basel, 1990, pp. 91-106.

[EOZ] B. Engquist, S. Osher and S. Zhong, "Fast wavelet algorithms for linear evolution equations", ICASE Report 92-14, 1992.

[GO] D. Gottlieb and S. Orszag, "Numerical analysis of spectral methods: Theory and applications", CBMS-NSF, Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Vol. 26, Philadelphia, PA, 1977.

[GT] D. Gottlieb and E. Tadmor, "Recovering pointwise values of discontinuous data within spectral accuracy", in Progress and Supercomputing in Computational Fluid Dynamics, E.M. Murman and S.S. Abarbanel, eds., Birkhauser, Boston, MA 1985, pp. 357-375.

[H1] A. Harten, "On high order accurate interpolation for non-oscillatory shock capturing schemes", IMA Program on Continuum Physics and Its Applications, pp. 72-105, 1984-85.

[H2] A. Harten, "ENO schemes with subcell resolution", J. Comp. Phys., Vol. 83, pp. 148-184, 1989; also ICASE Report No. 87-56.

[H3] A. Harten, "Multi-resolution analysis for ENO schemes", ICASE Report 91-77, September 1991; also in Proceedings in ICASE/LaRC workshop on algorithmic trends in CFD for the 90's.

[H4] A. Harten, "Discrete multiresolution analysis and generalized wavelets", J. Appl. Num. Math., Vol. 12, pp. 153-193, 1993; also UCLA CAM Report 92-08, February 1992.

[H5] A. Harten, "Multiresolution algorithms for the numerical solution of hyper-

bolic conservation laws", UCLA CAM Report 93-03, March 1993; to appear in Comm. Pure Appl. Math.

[H6] A. Harten, "Adaptive multiresolution schemes for shock computations", UCLA CAM Report 93-06, April 1993; to appear in Jour. Comput. Phys.

[H7] A. Harten, "Multiresolution representation of data. I. Preliminary Report", UCLA CAM Report No. 93-13, June 1993.

[HC] A. Harten and S. R. Chakravarthy, "Multi-dimensional ENO schemes for general geometries", ICASE Report 91-76, September 1991; also submitted to Jour. Comput. Phys.

[HEOC] A. Harten, B. Engquist, S. Osher, and S. Chakravarthy, "Uniformly high order accurate essentially non-oscillatory schemes, III", J. Comput. Phys., **71**, pp. 231-303, 1987.

[HY] A. Harten and I. Yad-Shalom, "Fast multiresolution algorithms for matrix-vector multiplication", ICASE Report 92-55, October 1992; to appear in SIAM Jour. Num. Anal.

[IK] E. Isaacson and H.B. Keller,"Analysis of Numerical Methods", John Wiley & Sons, 1966.

[LT] J. Liandrat and Ph. Tchamitchian, "Resolution of the 1D regularized Burgers' equation using a spatial wavelet approximation", ICASE Report 90-83, December 1990.

[Ma] S. Mallat, "Multiresolution approximation and wavelets orthonormal bases of $L^2(\mathbf{R})$", Trans. Amer. Math. Soc., Vol. 315, pp. 69-87, 1989.

[Me] Y. Meyer, "Ondelettes et Opérateurs", Hermann, 1990.

[ML] M. Mock and P.D. Lax, "The computation of discontinuous solutions of linear hyperbolic equations", Comm. Pure Appl. Math., **31**, pp. 423-430, 1978.

[MMO] A. Majda, J. McDonough, and S. Osher, "The Fourier method for nonsmooth initial data", Math. Comput., **32**, pp. 1041-1081, 1978.

[MR] Y. Maday and J.C. Ravel, "Adaptivité par Ondelettes: condition aux limites et dimensions supérieures", Tech. Report, Université Pierre et Marie Curie, Lab. D'Analyse Numérique, January 1992.

[RM] R.D. Richtmyer and K.W. Morton, "Difference methods for initial value problems", Second Edition, Interscience Publishers, 1967.

[S] G. Strang, "Wavelets and dilation equations: A brief introduction", SIAM Review, Vol. 31, pp. 614-627, 1989.

[Y1] H. Yserentant, "On the multi-level splitting of finite element spaces", Numer. Math., Vol. 49, pp. 379-412, 1986.

[Y2] H. Yserentant, "Old and new convergence proofs for multigrid methods", Acta Numerica, pp. 285-326, 1993.