S DTIC ELECTE DEC 27 1994 C **D**
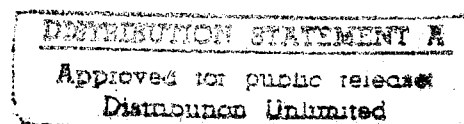
ICASE

MULTIRESOLUTION REPRESENTATION AND NUMERICAL ALGORITHMS: A BRIEF REVIEW

Ami Harten

19941221 043

U**SR**A   Operated by Universities Space Research Association

A M I R A M   H A R T E N
1 9 4 6 - 1 9 9 4

Amiram Harten, a professor of Mathematics at Tel-Aviv University and a consultant at ICASE, died of a massive heart attack on August 5, 1994. He was 47. He was associated with ICASE since 1976 and had a great impact on the ICASE program of research in numerical analysis and algorithm development. He was an active participant in ICASE activities and visited ICASE at least once a year. His last visit to ICASE was during the week of May 22, 1994, when he participated in the Parallel Numerical Algorithm Workshop. He planned to come back on August 29, 1994 to be at ICASE for a month. His untimely passing away is an irreparable loss to ICASE and the mathematics community. He will be greatly missed by his friends and colleagues even as his influence lives on.

This report is the written version of his lecture at the Parallel Numerical Algorithm Workshop, and it is sadly noted to be his last ICASE report.

Multiresolution Representation and Numerical Algorithms:
A Brief Review

Ami Harten[1]

School of Mathematical Sciences

Tel-Aviv University

Tel-Aviv, 69978 ISRAEL

and

Department of Mathematics

UCLA

ABSTRACT

In this paper we review recent developments in techniques to represent data in terms of its *local* scale components. These techniques enable us to obtain data compression by eliminating scale-coefficients which are sufficiently small. This capability for data compression can be used to reduce the cost of many numerical solution algorithms by either applying it to the numerical solution operator in order to get an approximate sparse representation, or by applying it to the numerical solution itself in order to reduce the number of quantities that need to be computed.

1. Introduction

Fourier analysis, which provides a way to represent square-integrable functions in terms of their sinusoidal scale-components, has contributed greatly to all fields of science. The main drawback of Fourier analysis is in its globality – a single irregularity in the function dominates the behavior of the scale-coefficients and prevents us from getting immediate information about the behavior of the function elsewhere.

The recent development of the theory of wavelets (see [Me] and [Ma]) was a great step towards local scale decomposition, and has already had great impact on several fields of science. In numerical analysis representation by compactly supported wavelets (see [Da] and [CDF]) is used to reduce the cost of many numerical solution algorithms by either applying it to the numerical solution operator to obtain an approximate sparse form (see [BCR]), or by applying it to the numerical solution itself to obtain an approximate reduced representation in order to solve for less quantities (see e.g. [LT], [MR] and [BMP]). The main drawback of the theory of wavelets is that it attempts to decompose any square integrable function into scale-components which are translates and dilates of a single function. Consequently there are conceptual difficulties in extending wavelets to bounded domains and general geometries. Furthermore, the uniformity of the underlying wavelet approximation makes it impossible to obtain an adaptive (data-dependent) multiresolution *representation* which fits the approximation to the local nature of the data. The only adaptivity which is possible within the theory of wavelets is through redundant "dictionaries."

In a series of works [H1-3] we have studied the question of how to represent discrete data which originates from unstructured grids in bounded domains in terms of scale decomposition. Combining ideas from multigrid methods, numerical solution of conservation laws, hierarchial bases of finite element spaces, subdivision schemes of Computer-Aided Design and of course – the theory of wavelets, we came up with the more general concept of "nested sequence of discretization." Given discrete data which can be associated with a nested sequence of discretization we show that it has a multiresolution representation, i.e., a one-to-one correspondence between the given data and its scale-decomposition. This framework is a generalization of the theory of wavelets in the sense that under conditions of uniformity its natural result is wavelets.

In this paper we review the work in [HY], [AC], [ACD], [H4-5] where the previously mentioned works on numerical solution algorithms with representation by wavelets are extended to the more general framework of nested discretization.

1

2. Nested Discretization

In this section we describe the class of discrete data for which we can obtain representation in terms of a scale-decomposition. We start with two examples.

Example 1.

Let us consider continuous functions f in the interval [0,1]

$$f \in \mathcal{F} = C^0[0,1],$$

and let $\{X^k\}_{k=0}^L$ be the following nested sequence of uniform grids

$$X^k = \{x_i^k\}_{i=0}^{J_k}, \quad x_i^k = ih_k, \quad h_k = 2^{-k}h_0, \quad J_k = 2^k J_0 \qquad (2.1a)$$

for some integer J_0 with $h_0 = 1/J_0$. Here $k = L$ is the finest grid and $k = 0$ is the coarsest. Observe that X^k is obtained by dyadic refinement of X^{k-1}, i.e.

$$x_{2i}^k = x_i^{k-1}, \quad i = 0, \ldots, J_{k-1} \qquad (2.1b)$$

$$x_{2i-1}^k = \frac{1}{2}(x_{i-1}^{k-1} + x_i^{k-1}), \quad i = 1, \ldots, J_{k-1}. \qquad (2.1c)$$

Let us consider now the discrete values $v^k = \{v_i^k\}_{i=0}^{J_k}$ which are obtained from point value discretization of $f \in \mathcal{F}$ in the k-th grid

$$v_i^k =: (\mathcal{D}_k f)_i = f(x_i^k), \quad i = 0, \ldots, J_k, \ 0 \leq k \leq L. \qquad (2.2a)$$

It follows from (2.1b) that v^{k-1} can be obtained from v^k by the decimation

$$v_i^{k-1} = v_{2i}^k, \quad i = 0, \ldots, J_{k-1}. \qquad (2.2b)$$

Let $I^{k-1}(x; v^{k-1})$ denote any continuous function in [0,1] which interpolates v^{k-1} at the grid points of X^{k-1}, i.e.

$$I^{k-1}(x; v^{k-1}) \in \mathcal{F} = C^0[0,1] \qquad (2.3a)$$

$$I^{k-1}(x_i^{k-1}; v^{k-1}) = v_i^{k-1} \quad \text{for} \quad i = 0, \ldots, J_{k-1} \qquad (2.3b)$$

e.g. we can take $I^{k-1}(x; v^{k-1})$ to be the piecewise-linear interpolation

$$I^{k-1}(x; v^{k-1}) = v_{i-1}^{k-1} + \frac{x - x_{i-1}^{k-1}}{h_{k-1}}(v_i^{k-1} - v_{i-1}^{k-1}), \quad \text{for} \quad x_{i-1}^{k-1} \leq x \leq x_i^{k-1}. \qquad (2.4)$$

Given v^{k-1} we can approximate v^k by

$$v_i^k \approx I^{k-1}(x_i^k; v^{k-1}), \quad i = 0, \ldots, J_k.$$

2

Let us denote

$$e_i^k = v_i^k - I^{k-1}(x_i^k; v^{k-1}), \quad i = 0, \ldots, J_k \qquad (2.5a)$$

and refer to $e^k = \{e_i^k\}_{i=0}^{J_k}$ as the prediction error. We observe that

$$e_{2i}^k = 0 \quad \text{for} \quad i = 0, \ldots, J_{k-1} \qquad (2.5b)$$

and denote the interpolation error at the odd grid points of X^k by $d^k = \{d_j^k\}_{j=1}^{J_{k-1}}$

$$d_j^k =: e_{2j-1}^k = v_{2j-1}^k - I^{k-1}(x_{2j-1}^k; v^{k-1}), \quad j = 1, \ldots, J_{k-1}. \qquad (2.5c)$$

Clearly there is a one-to-one correspondence between v^k and $\{d^k, v^{k-1}\}$: Knowing v^k we get v^{k-1} and d^k by (2.2b) and (2.5c), respectively. From $\{d^k, v^{k-1}\}$ we recover v^k by

$$\begin{cases} v_{2i}^k = v_i^{k-1}, \quad i = 0, \ldots, J_{k-1} \\[2mm] v_{2i-1}^k = I^{k-1}(x_{2i-1}^k; v^{k-1}) + d_i^k, \quad i = 1, \ldots, J_{k-1}. \end{cases} \qquad (2.6)$$

Since v^{k-1} can be likewise represented by $\{d^{k-1}, v^{k-2}\}$ we get that there is a one-to-one correspondence between the values of the finest level v^L and the sequence of $(J_L + 1)$ elements $\{d^L, \ldots, d^1, v^0\}$ which we denote by $\hat{v}_M$:

$$v^L \xleftrightarrow{1:1} \{d^L, \ldots, d^1, v^0\} =: \hat{v}_M. \qquad (2.7)$$

We refer to $\hat{v}_M$ as the multiresolution representation of v^L. It follows from (2.2b) and (2.5c) that the direct multiresolution representation (MR) transform $v^L \mapsto \hat{v}_M$ can be expressed by the following algorithm

$$\begin{cases} DO \ k = L, \ldots, 1 \\[2mm] v_i^{k-1} = v_{2i}^k, \quad i = 0, \ldots, J_{k-1} \\[2mm] d_j^k = v_{2j-1}^k - I^{k-1}(x_{2j-1}^k; v^{k-1}), \quad j = 1, \ldots, J_{k-1}. \end{cases} \qquad (2.8a)$$

We denote this transform by M, i.e.

$$\hat{v}_M = M \cdot v^L. \qquad (2.8b)$$

Similarly it follows from (2.6) that the inverse MR transform $\hat{v}_M \mapsto v^L$ can be expressed by

$$\begin{cases} DO \ k = 1, \ldots, L \\[2mm] v_{2i}^k = v_i^{k-1}, \quad i = 0, \ldots, J_{k-1} \\[2mm] v_{2j-1}^k = I^{k-1}(x_{2j-1}^k; v^{k-1}) + d_j^k, \quad j = 1, \ldots, J_{k-1}, \end{cases} \qquad (2.9a)$$

3

and we denote

$$v^L = M^{-1} \cdot \hat{v}_M. \tag{2.9b}$$

We refer to d^k as the scale-coefficients of the k-th level of resolution. For the piecewise-linear interpolation (2.4) we get from (2.5c) that

$$d_j^k = v_{2j-1}^k - \frac{1}{2}(v_{j-1}^{k-1} + v_j^{k-1}); \tag{2.10a}$$

in terms of $f \in \mathcal{F}$ for which $v^k = \mathcal{D}_k f$ this can be expressed by

$$d_j^k(f) = -\frac{1}{2}[f(x_{2j}^k) - 2f(x_{2j-1}^k) + f(x_{2j-2}^k)]. \tag{2.10b}$$

Hence if $f(x)$ is twice differentiable in $[x_{j-1}^{k-1}, x_j^{k-1}]$ we get that

$$d_j^k(f) = -\frac{1}{2}(h_k)^2 f''(\xi) \text{ for some } \xi \in (x_{j-1}^{k-1}, x_j^{k-1}) \tag{2.10c}$$

and consequently the scale coefficients in a region of smoothness of $f(x)$ tend to zero as $O((h_k)^2)$ for $k \to \infty$. However at a jump discontinuity $d_j^k(f)$ is proportional for the size of the jump and thus remains $0(1)$ as $k \to \infty$.

We can obtain data compression by setting to zero all scale coefficients which fall below a prescribed tolerance. Let us denote

$$\tilde{d}_j^k = \begin{cases} 0 & \text{if } |d_j^k| \leq \varepsilon_k \\ \\ d_j^k & \text{if } |d_j^k| > \varepsilon_k, \end{cases} \tag{2.11a}$$

and

$$\tilde{v}^L = M^{-1} \cdot \{\tilde{d}^L, \ldots, \tilde{d}^1, v^0\}. \tag{2.11b}$$

Based on the analysis in [H3] we get the following bound on the compression error in the case of piecewise-linear interpolation:

$$\max_{0 \leq i \leq J_L} |v_i^L - \tilde{v}_i^L| \leq \sum_{k=1}^{L} \varepsilon_k. \tag{2.12a}$$

Given any $\varepsilon > 0$ we can take

$$\varepsilon_k = 2^{k-L-1}\varepsilon \tag{2.12b}$$

and thus ensure by (2.12a) that

$$\max_{0 \leq i \leq J_L} |v_i^L - \tilde{v}_i^L| < \varepsilon. \tag{2.12c}$$

4

Example 2.

Let us consider absolutely integrable functions f in $[0,1]$

$$f \in \mathcal{F} = L^1[0,1],$$

and let

$$C^k = \{c_i^k\}_{i=1}^{J_k}, \quad c_i^k = (x_{i-1}^k, x_i^k), \tag{2.13a}$$

where $\{x_i^k\}$ are the gridpoints of X^k in (2.1); observe that

$$\overline{c_i^{k-1}} = \overline{c_{2i-1}^k \cup c_{2i}^k}. \tag{2.13b}$$

Let $v^k = \{v_i^k\}_{i=1}^{J_k}$ be discrete values which are obtained by taking the average over the cells in C^k of some function $f \in \mathcal{F} = L^1[0,1]$

$$v_i^k =: (\mathcal{D}_k f)_i = \frac{1}{|c_i^k|} \int_{c_i^k} f \, dx, \ |c_i^k| = h_k, \ i = 1, \dots, J_k, \ 1 \le k \le L. \tag{2.14a}$$

It follows from the additivity of the integral and (2.13b) that v^{k-1} can be obtained from v^k by the decimation

$$v_i^{k-1} = \frac{1}{2}(v_{2i-1}^k + v_{2i}^k), \ i = 1, \dots, J_{k-1}. \tag{2.14b}$$

Let $(\mathcal{R}_k v^k)(x)$ denote any function in $L^1[0,1]$ which satisfies

$$(\mathcal{D}_k \cdot \mathcal{R}_k v^k)_i = \frac{1}{|c_i^k|} \int_{c_i^k} (\mathcal{R}_k v^k)(x) dx = v_i^k, \quad i = 1, \dots, J_k \tag{2.15}$$

e.g. we can take $\mathcal{R}_k v^k$ to be the piecewise-constant function

$$(\mathcal{R}_k v^k)(x) = v_i^k \quad \text{for} \ x \in c_i^k. \tag{2.16}$$

In the context of ENO schemes for the numerical solution of conservation laws we refer to $\mathcal{R}_k$ as reconstruction from cell-averages and to (2.15) as "conservation" (see [HEOC]).

Given v^{k-1} we can approximate v^k by

$$v_i^k \approx (\mathcal{D}_k \cdot \mathcal{R}_{k-1} v^{k-1})_i = \frac{1}{|c_i^k|} \int_{c_i^k} (\mathcal{R}_{k-1} v^{k-1})(x) dx,$$

i.e. by first approximating f from v^{k-1} by $\mathcal{R}_{k-1} v^{k-1}$ and then taking cell-averages of this approximation over the finer level k. It follows from the conservation property (2.15) that the prediction error e^k,

$$e_i^k = v_i^k - (\mathcal{D}_k \cdot \mathcal{R}_{k-1} v^{k-1})_i, \quad i = 1, \dots, J_k \tag{2.17a}$$

satisfies the relation

$$e^k_{2i-1} + e^k_{2i} = 0 \quad \text{for} \quad i = 1, \ldots, J_{k-1}. \tag{2.17b}$$

Therefore, if we store

$$d^k_j = e^k_{2j-1} \quad \text{for} \quad j = 1, \ldots, J_{k-1} \tag{2.18a}$$

we can recover the prediction error e^k by

$$\begin{cases} e^k_{2i-1} = d^k_i \\[2mm] e^k_{2i} = -d^k_i \end{cases} \tag{2.18b}$$

and thus get a one-to-one correspondence between v^k and $\{d^k, v^{k-1}\}$. As in the previous example this leads to the multiresolution representation

$$v^L \xleftrightarrow{\ 1:1\ } \{d^L, \ldots, d^1, v^0\} =: \hat{v}_M$$

where the direct MR transform $\hat{v}_M = M \cdot v^L$ can be expressed by the algorithm

$$\begin{cases} DO\ k = L, \ldots, 1 \\[2mm] v^{k-1}_i = \frac{1}{2}(v^k_{2i-1} + v^k_{2i}), \quad i = 1, \ldots, J_{k-1} \\[2mm] d^k_j = v^k_{2j-1} - \frac{1}{|c^k_{2j-1}|} \int_{c^k_{2j-1}} (\mathcal{R}_{k-1} v^{k-1})(x) dx, \quad j = 1, \ldots, J_{k-1} \end{cases} \tag{2.19}$$

and the inverse MR transform $v^L = M^{-1} \cdot \hat{v}_M$ is given by

$$\begin{cases} DO\ k = 1, \ldots, L \\[2mm] DO\ i = 1, \ldots, J_{k-1} \\[2mm] v^k_{2i-1} = \frac{1}{|c^k_{2i-1}|} \int_{c^k_{2i-1}} (\mathcal{R}_{k-1} v^{k-1})(x) dx + d^k_i \\[2mm] v^k_{2i} = 2 v^{k-1}_i - v^k_{2i-1}. \end{cases} \tag{2.20}$$

Observe that the last statement of (2.20) is obtained from (2.14b).

In [HEOC] we showed that any interpolation method (2.3) gives rise to a reconstruction from cell-averages (2.15) by the following "reconstruction via primitive function" technique: Given cell-averages $v^k = \mathcal{D}_k f$ in (2.14a) we calculate

$$F^k_i = F(x^k_i), \quad F(x) = \int_0^x f(y) dy$$

by

$$F^k_0 = 0, \quad F^k_i = h_k \sum_{j=1}^i v^k_j, \quad 1 \le i \le J_k, \tag{2.21a}$$

and define

$$(\mathcal{R}_k v^k)(x) = \frac{d}{dx} I^k(x; F^k),$$
<div align="right">(2.21b)</div>

where $I^k(x; F^k)$ is any interpolation of the values $F^k = \{F_i^k\}_{i=0}^{J_k}$ at the grid points of X^k in (2.1a).

In the previous two examples we have shown how to design MR schemes for discrete data v^L which is obtained from discretization of functions by point values (Example 1) or by cell-averages (Example 2) in the nested sequence of uniform grids of $[0,1]$. In [H2] and [H3] we have presented a more general class of discrete data which can be represented by a scale decomposition. This class is characterized by the following notion of nested discretization.

Definition. We say that a sequence of linear operators $\{D_k\}_{k=0}^{\infty}$ is a nested sequence of discretization if

(i)

$$\mathcal{D}_k : \mathcal{F} \xrightarrow{\text{onto}} V^k, \quad \dim V^k = J_k,$$
<div align="right">(2.22a)</div>

(ii)

$$\mathcal{D}_k f = 0 \Rightarrow \mathcal{D}_{k-1} f = 0$$
<div align="right">(2.22b)</div>

Here $\mathcal{F}$ is a space of mappings and V^k is a linear space of dimension J_k.

In the next section we show how to obtain multiresolution representation of any discrete data $v^L = \mathcal{D}_L f$, where the scale-decomposition corresponds to the levels of resolution which are introduced in (2.22). This is a very general framework which allows for discretizations corresponding to unstructured grids in several space dimensions.

3. General Multiresolution Representation Schemes

In this section we consider discrete data which is associated with a nested sequence of discretization $\{\mathcal{D}_k\}_{k=0}^{L}$ and show how to design schemes for its multiresolution representation.

First we show that a nested sequence of discretization comes equipped with a decimation operator D_k^{k-1} which is a linear mapping from $V^k = \mathcal{D}_k(\mathcal{F})$ onto $V^{k-1} = \mathcal{D}_{k-1}(\mathcal{F})$

$$D_k^{k-1} : V^k \xrightarrow{\text{onto}} V^{k-1}.$$
<div align="right">(3.1a)</div>

This decimation operator is defined as follows: For any v in V^k there is at least one $f \in \mathcal{F}$ such that $\mathcal{D}_k f = v$; the decimation of v is $\mathcal{D}_{k-1} f \in V^{k-1}$, i.e.,

$$v \in V^k, \quad v = \mathcal{D}_k f, \quad D_k^{k-1} v = \mathcal{D}_{k-1} f.$$
<div align="right">(3.1b)</div>

It follows from (2.22b) that D_k^{k-1} is well-defined by (3.1b), i.e. its definition is independent of the particular f. To see that let us take f_1 and f_2 in $\mathcal{F}$ such that

$$\mathcal{D}_k f_1 = v = \mathcal{D}_k f_2,$$

then by (2.22b)

$$0 = \mathcal{D}_k f_1 - \mathcal{D}_k f_2 = \mathcal{D}_k(f_1 - f_2) \Rightarrow 0 = \mathcal{D}_{k-1}(f_1 - f_2) = \mathcal{D}_{k-1} f_1 - \mathcal{D}_{k-1} f_2$$

which proves our claim.

Given $v^L \in V^L$ we can evaluate $\{v^k\}_{k=0}^{L-1}$ by repeated decimation

$$v^{k-1} = D_k^{k-1} v^k, \quad k = L, \ldots, 1. \tag{3.2}$$

Since (3.1b) implies that

$$D_k^{k-1}(\mathcal{D}_k f) = \mathcal{D}_{k-1} f \quad \text{for any } f \in \mathcal{F} \tag{3.3}$$

we get for any $f \in \mathcal{F}$ for which $v^L = \mathcal{D}_L f$, that $v^k = \mathcal{D}_k f$ for all k in (3.2). We would like to stress the point that, as in (2.2b) and (2.14b) in the previous examples, this decimation is done without explicit knowledge of f.

Since by (2.22a) $V^k = \mathcal{D}_k(\mathcal{F})$, it follows that $\mathcal{D}_k$ has a right-inverse (at least one) which we denote by $\mathcal{R}_k$:

$$\mathcal{R}_k : V^k \to \mathcal{F}, \quad \mathcal{D}_k \mathcal{R}_k = I_k, \tag{3.4}$$

where I_k denotes the identity operator in V^k. Since $(\mathcal{R}_k v^k) \in \mathcal{F}$ is an approximation to any $f \in \mathcal{F}$ for which $\mathcal{D}_k f = v^k$, we refer to $\mathcal{R}_k$ as a reconstruction of $\mathcal{D}_k$; see (2.3)-(2.4) and (2.15)-(2.16) in the previous examples.

Next we show that any sequence of corresponding reconstruction operators $\{\mathcal{R}_k\}_{k=0}^L$ defines a MR scheme for discrete data v^L in V^L. Starting from v^{k-1} in (3.2) we can get an approximation to v^k by

$$v^k \approx \mathcal{D}_k(\mathcal{R}_{k-1} v^{k-1}).$$

We denote

$$P_{k-1}^k =: \mathcal{D}_k \mathcal{R}_{k-1}, \quad P_{k-1}^k : V^{k-1} \to V^k \tag{3.5a}$$

and refer to it as prediction operator. It follows immediately from taking $f = \mathcal{R}_{k-1} v^{k-1}$ in (3.3) and using (3.4) that P_{k-1}^k is a right-inverse of the decimation D_k^{k-1}

$$D_k^{k-1} P_{k-1}^k = I_{k-1}. \tag{3.5b}$$

We observe that the prediction error e^k

$$e^k = v^k - P_{k-1}^k v^{k-1} = (I_k - P_{k-1}^k D_k^{k-1}) v^k \tag{3.6a}$$

8

satisfies the relation

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

and therefore it is in the null space of the decimation operator

$$e^k \in \mathcal{N}(D_k^{k-1}) = \{v| \quad v \in V^k, \ D_k^{k-1} v = 0\}. \tag{3.6b}$$

It follows from (3.1a) that

$$\dim \mathcal{N}(D_k^{k-1}) = J_k - J_{k-1} \tag{3.7a}$$

and hence

$$\mathcal{N}(D_k^{k-1}) = \operatorname{span}\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}, \tag{3.7b}$$

where $\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}$ is any basis of $\mathcal{N}(D_k^{k-1})$. Therefore the prediction error e^k, which is described in terms of J_k components in V^k, can be represented by its $(J_k - J_{k-1})$ coordinates d^k in (3.7b)

$$e^k = \sum_{j=1}^{J_k - J_{k-1}} d_j^k \mu_j^k =: E_k d^k, \quad d^k =: G_k e^k. \tag{3.8a}$$

Here G_k denotes the operator which assigns to $e^k \in \mathcal{N}(D_k^{k-1})$ its coordinates d^k in the basis $\{\mu_j^k\}_{j=1}^{J_k - J_{k-1}}$; observe that $E_k G_k$ is the identity operator in $\mathcal{N}(D_k^{k-1})$, i.e.

$$E_k G_k e^k = e^k \quad \text{for any } e^k \in \mathcal{N}(D_k^{k-1}). \tag{3.8b}$$

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

$$\begin{cases} v^{k-1} &= D_k^{k-1} v^k \\ d^k &= G_k(I_k - P_{k-1}^k D_k^{k-1}) v^k \quad ; \end{cases}$$

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

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

As in the previous examples this shows that

$$v^L \xleftrightarrow{1:1} \{d^L, \ldots, d^1, v^0\} =: \hat{v}_M, \tag{3.9}$$

9

where the direct MR transform $\hat{v}_M = M \cdot v^L$ is given by the algorithm

$$\begin{cases} DO \ k = L, \dots, 1 \\[2mm] v^{k-1} = D_k^{k-1} v^k \\[2mm] d^k = G_k (I_k - P_{k-1}^k D_k^{k-1}) v^k =: G_k^D v^k \end{cases} \tag{3.10}$$

and the inverse MR transform $v^L = M^{-1} \cdot \hat{v}_M$ can be calculated by

$$\begin{cases} DO \ k = 1, \dots, L \\[2mm] v^k = P_{k-1}^k v^{k-1} + E_k d^k. \end{cases} \tag{3.11}$$

We remark that in multigrid terminology D_k^{k-1} is "restriction" and P_{k-1}^k is "prolongation." In signal processing D_k^{k-1} plays the role of "low-pass filter" while G_k^D, which is defined in (3.10), plays the role of "high-pass filter."

Example 3. Biorthogonal Wavelets.

In this example we derive the MR schemes which correspond to the bases of biorthogonal wavelets in [CDF]. These MR schemes are obtained from nested discretization of functions in $L_{loc}^2(\mathbf{R})$ by taking weighted-averages on a nested sequence of uniform grids of $\mathbf{R}$, as follows:

$$(\mathcal{D}_k f)_i = \frac{1}{h_k} \int_{-\infty}^{\infty} f(x) w \left(\frac{x - x_i^k}{h_k} \right) dx, \quad -\infty < i < \infty, \tag{3.12a}$$

where $w \in L^2(\mathbf{R})$ is a weight-funcntion

$$\int_{-\infty}^{\infty} w(x) dx = 1 \tag{3.12b}$$

and

$$x_i^k = i h_k, \quad -\infty < i < \infty, \quad h_k = 2^{-k} h_0. \tag{3.12c}$$

In order to obtain a nested sequence of discretization we want to choose $w(x)$ so that

$$(\mathcal{D}_{k-1} f)_i = \sum_{\ell=0}^{N} \alpha_\ell (\mathcal{D}_k f)_{2i-\ell} \tag{3.12d}$$

where α_ℓ, $\ell = 0, \dots, N$ are real numbers. We observe that since

$$f(x) \equiv c = \text{constant} \Rightarrow (\mathcal{D}_k f)_i = c \quad \forall i, k$$

we have to limit the choice of $\{\alpha_\ell\}$ by

$$\sum_{\ell=0}^{N} \alpha_\ell = 1. \tag{3.13a}$$

10

From (3.12d) and (3.12a) we get that for any $f \in L^2_{\text{loc}}(\mathbf{R})$

$$0 = \int_{-\infty}^{\infty} \left[\frac{1}{h_{k-1}} w \left(\frac{x - x_i^{k-1}}{h_{k-1}} \right) - \frac{1}{h_k} \sum_{\ell=0}^{N} \alpha_\ell w \left(\frac{x - x_{2i-\ell}^k}{h_k} \right) \right] f(x) dx \quad ;$$

this shows that $w(x)$ has to satisfy the functional equation

$$w(x) = 2 \sum_{\ell=0}^{N} \alpha_\ell w(2x - \ell). \tag{3.13b}$$

This equation has been investigated in [Da] and [CDF] and got the name of "dilation equation" in [S]. It is shown there that subject to condition (3.13a) the dilation equation (3.13b) has a distribution solution which is determined up to a multiplicative constant and a shift, and that $w(x)$ has a support of size N. Furthermore, if

$$\sum_\ell \alpha_{2\ell} = \sum_\ell \alpha_{2\ell-1} \tag{3.13c}$$

then $w(x)$ is also square integrable.

We make the choice of $w(x)$ unique by imposing (3.12b) and fixing its support in, say, $[-N, 0]$. With this choice of a weight-function, the sequence of discretization in (3.12) is nested and its decimation operator is given by

$$(D_k^{k-1} v)_i =: \sum_\ell \alpha_\ell v_{2i-\ell} = \sum_m \alpha_{2i-m} v_m. \tag{3.14a}$$

In [Da] and [H3] it is shown that $\{\mu_j^k\}_{j=-\infty}^{\infty}$,

$$(\mu_j^k)_i = (-1)^{i+1} \alpha_{2j-i-1}, \quad -\infty < i < \infty \tag{3.14b}$$

is a basis of $\mathcal{N}(D_k^{k-1})$ in (3.6b).

We reconstruct the discretization $\mathcal{D}_k$ in (3.12a) as follows: We take a sequence $\{\beta_\ell\}$ of compact support which satisfies

$$\begin{cases} \sum_\ell \beta_{2\ell} = \sum_\ell \beta_{2\ell-1} = 1 \\ \\ \sum_\ell \alpha_\ell \beta_{\ell+2m} = \delta_{m,0} \end{cases} \tag{3.15}$$

and define

$$(\mathcal{R}_k v)(x) = \sum_i v_i \varphi \left(\frac{x - x_i^k}{h_k} \right), \tag{3.16a}$$

where $\varphi(x)$ is a solution of the dilation equation

$$\varphi(x) = \sum_\ell \beta_\ell \varphi(2x - \ell) \tag{3.16b}$$

11

which is normalized by

$$\int \varphi(x)w(x)dx = 1.$$
(3.16c)

It is easy to see that the corresponding prediction operator $P_{k-1}^k = \mathcal{D}_k \mathcal{R}_{k-1}$ is given by

$$(P_{k-1}^k v)_i =: \sum_m \beta_{i-2m} v_m$$
(3.17)

Daubechies' orthonormal wavelets are obtained by imposing the additional condition

$$\beta_\ell = 2\alpha_\ell$$

(see [H2] and [H3] for more details).

We remark that the "fundamental solution" for the dilation equation (3.13b) with

$$\alpha_\ell = \delta_{\ell,0}$$
(3.18a)

is

$$w(x) = \delta(x),$$
(3.18b)

where $\delta(x)$ is the Dirac distribution (see [S]). In this case (3.12a) becomes the point value discretization (2.2a) of Example 1. However, since $\delta(x)$ is not square integrable, point value discretization is excluded from the theory of wavelets. For

$$\alpha_\ell = \frac{1}{2}(\delta_{\ell,0} + \delta_{\ell,1})$$
(3.19a)

we get

$$w(x) = \begin{cases} 1 & -1 < x < 0 \\ 0 & \text{otherwise} \end{cases}$$
(3.19b)

which is square integrable, and the discretization (3.12a) becomes the cell-average discretization (2.14a) of Example 2. Observe, however, that the theory of wavelets is for the infinite domain R, while our formulation is suitable for both the finite (Example 2) and the infinite case (Example 3). Furthermore, unlike the theory of wavelets which uses translates and dilates of a single function for both discretization and reconstruction and consequently is restricted to uniform grids, our framework of nested discretization allows for general geometries. In [H3] and [AH1] we extend the multiresolution representation of cell-averaged data in Example 2 to unstructured meshes in bounded domains of $R^m, m > 1$, by using agglomeration of cells to generate a nested sequence of discretization.

In [ADH] we consider the case where $w(x)$ in (3.12) is the "hat-function" and show how to improve data compression by using adaptive prediction techniques near discontinuities and distributions.

12

4. Multiresolution Representation of a 2-Dimensional Array

In this section we consider functions f

$$f : [0,1] \times [0,1] \longrightarrow R \qquad (4.1a)$$

which are discretized on the tensor-product grid

$$X^k = \{(x_i^k, x_j^k)\}_{i,j=0}^{N_k} \qquad (4.1b)$$

by

$$A_{i,j}^k = \frac{1}{(h_k)^2} \int_0^1 \int_0^1 f(x_1, x_2) w\left(\frac{x_1 - x_i^k}{h_k}\right) w\left(\frac{x_2 - x_i^k}{h_k}\right) dx_1 dx_2 \qquad (4.1c)$$

where $\{x_i^k\}$ are the one-dimensional gridpoints in (2.1) and $w(x)$ is a weight-function as in Example 3; by (3.18) and (3.19) this includes point value and cell-average discretizations. Although this case is covered by the general framework in (3.9)-(3.11), it is convenient here to represent the two-dimensional array in (4.1) as the $N_k \times N_k$ matrix A^k, and to use tensor-product extension of the corresponding one-dimensional operators to get a MR scheme for the input A^L.

Let us denote the matrix representation of the various one-dimensional operators by

$$D_k^{k-1} \longrightarrow (D)_{N_{k-1} \times N_k}$$

$$P_{k-1}^k \longrightarrow (P)_{N_k \times N_{k-1}}$$

$$G_k^D = G_k(I_k - P_{k-1}^k D_k^{k-1}) \longrightarrow (G^D)_{N_{k-1} \times N_k} = G(I - PD) \qquad (4.2)$$

$$E_k \longrightarrow (E)_{N_k \times N_{k-1}}.$$

These matrix representations are obtained by taking v^k and d^k in (3.10)-(3.11) to be column-vectors, e.g.

$$v_i^{k-1} = (D_k^{k-1} v^k)_i =: \sum_{j=1}^{N_k} D_{ij} v_j^k, \qquad 1 \le i \le N_{k-1} \qquad (4.3a)$$

where by (3.14a)

$$D_{ij} = \alpha_{2i-j} ; \qquad (4.3b)$$

for simplicity we drop the index k.

Starting with A^L we decimate to get

$$A^{k-1} = DA^k D^*, \qquad k = L, \ldots, 1; \qquad (4.4)$$

here $(\cdot)^*$ denotes the transpose. Given A^{k-1} we get an approximation to A^k by

$$A^k \approx PA^{k-1}P^*$$

13

and observe that the prediciton error *matrix* e^k

$$e^k = A^k - PA^{k-1}P^* \qquad (4.5a)$$

satisfies

$$\cdot \, De^k D^* = 0. \qquad (4.5b)$$

The dimension of the null space of the decimation operators here is

$$J_k - J_{k-1} = (2N_{k-1})^2 - (N_{k-1})^2 = 3(N_{k-1})^2$$

and we store the scale-coefficients d^k in three $N_{k-1} \times N_{k-1}$ matrices $\Delta_1^k, \Delta_2^k, \Delta_3^k$.

Using the matrix identity

$$I = PD + EG^D, \qquad G^D = G(I - PD), \qquad (4.6)$$

we show that if we take A^k and A^{k-1} from the sequence (4.4) and define

$$\Delta_1^k = G^D A^k (G^D)^*, \qquad \Delta_2^k = DA^k (G^D)^*, \qquad \Delta_3^k = G^D A^k D^* \qquad (4.7)$$

then A^k can be recovered from A^{k-1} and the above by

$$A^k = PA^{k-1}P^* + E\Delta_1^k E^* + P\Delta_2^k E^* + E\Delta_3^k P^*. \qquad (4.8)$$

This follows immediately from the identity

$$\begin{aligned} A^k &= (PD + EG^D)A^k(PD + EG^D)^* = P(DA^k D^*)P^* \\ &+ E[G^D A^k(G^D)^*]E^* + P[DA^k(G^D)^*]E^* + E(G^D A^k D^*)P^*. \end{aligned}$$

We conclude from (4.7)-(4.8) that A^L has a multiresolution representation $\hat{A}_M$,

$$A^L \overset{1:1}{\leftrightarrow} \hat{A}_M =: \left(\left\{ \Delta_m^L \right\}_{m=1}^3, \ldots, \left\{ \Delta_m^1 \right\}_{m=1}^3, A^0 \right), \qquad (4.9)$$

where the direct MR transform is given by

$$\left\{ \begin{aligned} &DO \quad k = L, \ldots, 1 \\ &A^{k-1} = DA^k D^* \\ &\Delta_1^k = G^D A^k (G^D)^*, \quad \Delta_2^k = DA^k (G^D)^*, \quad \Delta_3^k = G^D A^k D^*, \end{aligned} \right. \qquad (4.10)$$

and the inverse MR transform is

$$\left\{ \begin{aligned} &DO \quad k = 1, \ldots, L \\ &A^k = PA^{k-1}P^* + E\Delta_1^k E^* + P\Delta_2^k E^* + E\Delta_3^k P^*. \end{aligned} \right. \qquad (4.11)$$

14

In Figure 1, which is taken from [HY], we show the result of applying thresholding with tolerance $\varepsilon = 10^{-7}$ to $\hat{A}_M$ in (4.9) where

$$A_{ij} = \begin{cases} \frac{1}{i-j} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{4.12}$$

Here we take $N_L = 512$ and use point value discretization which is reconstructed by a sixth-order accurate piecewise-polynomial interpolation with a centered stencil; near the boundaries we use one-sided stencils. The rate of compression, i.e. the ratio of $(N_L)^2$ to the number of entries in $\hat{A}^M$ which are above the tolerance $\varepsilon = 10^{-7}$, is 8.57. This example is taken from [BCR] where it is done with MR schemes which use Daubechies' wavelets; the corresponding rate of compression for wavelets with six vanishing moments is reported to be 7.33 .

In Figure 1 we display the results by writing $\hat{A}_M$ in (4.9) as the matrix

and marking the entries that are larger in absolute value than 10^{-7} by a black dot.

5. Multiresolution Algorithm for Matrix-Vector Multiplication.

In this section we describe an algorithm to reduce the cost of evaluating the matrix-vector multiplication c,

$$c = Ab, \tag{5.1}$$

where A is a $N_L \times N_L$ matrix which can be thought of as the discretization (4.1) of some piecewise-smooth function, and b is *any* vector of N_L components; note that we do not make

15

any smoothness assumptions on b. As an example let us consider the case where (5.1) is obtained from discretization of the integral transform

$$c(x) = \int_0^1 a(x,y)b(y)dy; \qquad (5.2a)$$

here

$$c_i \approx c(x_i^L), \quad A_{ij} \approx h_L \cdot a(x_i^L, x_j^L), \quad b_j \approx b(x_j^L), \qquad (5.2b)$$

and we want to evaluate (5.1) to a specified tolerance which is taken to be of the order of the discretization error in (5.2).

The basic idea is that the product $c = Ab$, which is being set up as $c^L = A^L b^L$, has a meaningful analog

$$c^k = A^k b^k, \qquad k = L-1,\ldots,0 \qquad (5.3a)$$

corresponding to the k-th grid, and that

$$c^k = Pc^{k-1} + \text{correction}, \qquad (5.3b)$$

where P is the matrix representation (4.2) of the one-dimensional prediction operator, and the "correction" comes from locations in $[0,1] \times [0,1]$ where $a(x,y)$ is still not "sufficiently" resolved on the $(k-1)$-th grid. To obtain (5.3) let us multiply (4.8) by b^k to get

$$A^k b^k = PA^{k-1}(P^* b^k) + E\Delta_1^k(E^* b^k) + P\Delta_2^k(E^* b^k) + E\Delta_3^k(P^* b^k),$$

from which we see that if we define

$$b^{k-1} = P^* b^k, \qquad k = L,\ldots,1 \qquad (5.4a)$$

and denote

$$s^{k-1} = E^* b^k, \qquad k = 1,\ldots,L \qquad (5.4b)$$

then we can rewrite the above identity as

$$c^k = Pc^{k-1} + E(\Delta_1^k s^{k-1} + \Delta_3^k b^{k-1}) + P(\Delta_2^k s^{k-1}). \qquad (5.4c)$$

Using this observation we get the following algorithm for the approximate (up to a specified error) evaluation of the matrix-vector multiplication (5.1):

(A) **Preparation:**

(i) Given A, set $A^L = A$ and use the direct MR transform (4.10) to obtain the MR representation $\hat{A}_M$ (4.9).

(ii) Apply thresholding to $\hat{A}_M$ in order to obtain a sparse representation.

(B) **Multiplication:**

Given any vector b

(i) Set $b^L = b$

$$\begin{cases} DO \quad k = L, \ldots, 1 \\[2mm] b^{k-1} = P^* b^k \\[2mm] s^{k-1} = E^* b^k \end{cases} \tag{5.5}$$

(ii) Calculate directly

$$c^0 = A^0 b^0 \tag{5.6a}$$

(this is done on the coarsest grid).

(iii)

$$\begin{cases} DO \ for \quad k = 1, \ldots, L \\[2mm] c^k = P(c^{k-1} + \Delta_2^k s^{k-1}) + E(\Delta_1^k s^{k-1} + \Delta_3^k b^{k-1}) \end{cases} \tag{5.6b}$$

Set $c = c^L$.

In the case of pointvalue discretization this algorithm turns out to be identical to the multilevel matrix multiplication of Brandt and Lubrecht in [BL]. The algorithm which corresponds to Daubechies' orthonormal wavelets is identical to the "non-standard form" of Beylkin, Coifman and Rokhlin in [BCR]. We remark that (5.5)-(5.6) is a slight generalization of the algorithm which was presented in [HY].

Based on the analysis in [BCR] we show in [HY] that if the kernel $a(x, y)$ in (5.2a) satisfies

$$|\partial_\ell a(x, y)| \le \frac{C_a}{|x - y|^{\ell+1}}, \quad \ell = 0, \ldots, r - 1 \tag{5.7}$$

then outside a diagonal band of width B, the entries of the matrices $\{\Delta_m^k\}_{m=1}^3$ are not larger than $O(B^{-r})$, where r is the order of accuracy of the reconstruction technique. Since the matrices P and E are banded, we get that the complexity of the algorithm in (5.5)-(5.6) is $O(N_L)$.

In [HY] we used the above matrix-vector multiplication algorithm to apply the matrix in (4.12) to a vector b with "randomly generated" components. Using single precision we obtained for the case in Figure 1 a relative residual error $\|Ab - c\|/\|b\|$ which was 7.52×10^{-6} in the ℓ_1 norm, and 4.41×10^{-5} in the ℓ_∞ norm.

17

6. Multiresolution Form of Numerical Schemes.

In this section we consider the numerical algorithm

$$v^{n+1} = Av^n + g \tag{6.1a}$$

which describes either an iterative procedure or a numerical scheme for the solution of an initial value problem. Taking the MR transform of (6.1a) we get

$$\hat{v}_M^{n+1} = Mv^{n+1} = (MAM^{-1})(Mv^n) + Mg =: \hat{A}_S \hat{v}_M^n + \hat{g}_M; \tag{6.1b}$$

observe that $\hat{A}_S = MAM^{-1}$, the multiresolution form of the matrix (= operator) A, is different from $\hat{A}_M$, the multiresolution representation (4.9), where A is treated as a two-dimensional array and not as an operator.

From (3.10) we get that Mv^L can be expressed by

$$\begin{cases} d^k = G_k^D(D_{k+1}^k \cdots D_L^{L-1}) \cdot v^L =: B_L^k v^L, \quad k = 1, \dots, L \\ v^0 = (D_1^0 \cdots D_L^{L-1}) \cdot v^L =: B_L^0 v^L. \end{cases} \tag{6.2}$$

From (3.11) we get that $M^{-1}\hat{v}_M$ can be expressed by

$$M^{-1}\hat{v}_M = \sum_{k=1}^L C_k^L d^k + C_0^L v^0, \tag{6.3a}$$

where

$$\begin{cases} C_k^L = (P_{L-1}^L \cdots P_k^{k+1})E_k, \qquad k = 1, \dots, L \\ C_0^L = P_{L-1}^L \cdots P_0^1. \end{cases} \tag{6.3b}$$

Using (6.2)-(6.3) to epxress (6.1b) with $g = 0$, we get

$$\begin{cases} d^{k,n+1} = \sum_{\ell=1}^L (B_L^k A C_\ell^L) \cdot d^{\ell,n} + (B_L^k A C_0^L) \cdot v^{0,n}, \quad k = 1, \dots, L \\ v^{0,n+1} = \sum_{\ell=1}^L (B_L^0 A C_\ell^L) \cdot d^{\ell,n} + (B_L^0 A C_0^L) \cdot v^{0,n} \end{cases} \tag{6.4a}$$

which shows that

$$\hat{A}_S = \qquad \begin{array}{c} B_L^L AC_1^L \\ \downarrow \end{array} \qquad \longleftarrow B_L^L AC_0^L$$

$$B_L^L AC_L^L \qquad \cdots$$

$$\longleftarrow B_1^L AC_1^L$$

$$B_L^1 AC_L^L$$

$$B_L^0 AC_L^L \qquad \cdots \qquad \longleftarrow B_L^0 AC_0^L$$

$$\begin{array}{c} \uparrow \\ B_L^0 AC_1^L \end{array}$$

(6.4b)

Observe that the block $B_L^k AC_\ell^L$, $1 \le k, \ell \le L$, is of size $N_{k-1} \times N_{\ell-1}$.

In the Appendix we show that $C_\ell^L = (\widetilde{B}_L^\ell)^*$ where $\widetilde{B}_L^\ell$ is given by (6.2) for the *dual* MR scheme; hence

$$(6.5) \qquad\qquad B_L^k AC_\ell^L = B_L^k A(\widetilde{B}_L^\ell)^*$$

and each column of $(B_L^k A)$ is d^k, the scale coefficients (6.2) of the k-th column of A, while each row of $A(\widetilde{B}_L^\ell)^*$ is $(\tilde{d}^\ell)^*$, where $\tilde{d}^\ell$ is the column-vector of scale coefficients of the data in the ℓ-th row of A which is obtained by the dual MR scheme.

We remark that when M in (6.1b) corresponds to Daubechies' orthonormal wavelets, $\hat{A}_S$ is identical to the standard form which was introduced in [BCR]. In this case the MR scheme is dual to itself, i.e. $\widetilde{B}_L^k = B_L^k$ and therefore $M^{-1} = M^*$. It was shown in [BCR] that if A is a discretization of a Calderon-Zygmund operator, then data compression of each block in (6.5) results in a "diagonal" band, the width of which is independent of k and ℓ. Consequently applying data compression to $\hat{A}_S$ results in a finger-like structure of non-zero entries, and their total number is $O(N_L \log_2 N_L)$.

In [AC] and [ACD] the standard form of [BCR] was extended by (6.1b) to point value and cell-average discretizations. It is shown there that, inspite of the lack of symmetry, the rate of compression compares favorably to that of orthonormal wavelets: It is about the same in the periodic case, but it is significantly better in presence of boundaries.

In Figure 2, which is taken from [AC], we show the nonzero entries of $\hat{A}_S$, the multiresolution form of the matrix in (4.12). As in Figure 1 we use point value discretization which is reconstructed by piecewise-polynomial interpolation, however here $N_L = 64$.

19

The present work extends the scope of application of these algorithms to the general class of MR schemes in (3.9)-(3.11).

7. Multiresolution Application of Operators.

In this section we describe multiresolution algorithms for the solution of integral equations and for the numerical solution of initial-value problems. We shall say that a matrix is *T-Sparse* if it becomes sparse under thresholding, and refer to the number of non-zero elements as its *T-sparseness*. In this language we can express the main result of the previous section by saying that the multiresolution form $\hat{A}_S$ of Calderon-Zygmund operators is *T*-sparse and that the *T*-sparseness of $\hat{A}_S$ is $O(N_L \log_2 N_L)$.

7.A. Integral Transforms and Equations.

In section 5 we described a multilevel algorithm for matrix-vector multiplication (5.1) and showed that if it corresponds to a discretization of an integral transform of the form (5.2a) with a kernel which satisfies (5.7), then it can be performed in $O(N_L)$ operations. We can also evaluate $c = Ab$ from its multiresolution form by

$$\begin{cases} \hat{b}_M = Mb, \\ \hat{c}_M = \hat{A}_S \hat{b}_M, \\ c = M^{-1} \hat{c}_M. \end{cases} \tag{7.1}$$

Since the *T*-sparseness of $\hat{A}_S$ is $O(N_L \log_2 N_L)$ this is also the cost of the product $\hat{A}_S \hat{b}_M$; in addition we also have to apply the direct MR transform to b and its inverse to $\hat{c}_M$. Hence, unless we are in a special situation in which $\hat{c}_M$ and/or $\hat{b}_M$ are also *T*-sparse, it is more efficient to calculate integral transforms by the multilevel algorithm (5.5)-(5.6).

A situation of this type occurs in a matrix-matrix multiplication $C = AB$ where both $\hat{A}_S$ and $\hat{B}_S$ are *T*-sparse

$$\hat{C}_S = MCM^{-1} = (MAM^{-1})(MBM^{-1}) = \hat{A}_S \hat{B}_S. \tag{7.2a}$$

Of particular interest is the case where the *T*-sparseness of

$$M(A)^n M^{-1} = (\hat{A}_S)^n \tag{7.2b}$$

is uniform in n. In this case $(\hat{A}_S)^n$ for $n = 2^m$ can be computed in m steps of squaring and thresholding

$$(\hat{A}_S)^{2^k} = \left[(\hat{A}_S)^{2^{k-1}} \right]^2, \quad k = 1, \ldots, m, \tag{7.3}$$

where each product above is between matrices with *T*-sparseness of $O(N_L \log_2 N_L)$.

20

The numerical algorithm (6.1a) can be written as

$$v^n = (A)^n v^0 + \sum_{j=0}^{n-1} (A)^{j-1} g \qquad (7.4a)$$

or in its multiresolution form (6.1b) as

$$\hat{v}_M^n = (\hat{A}_S)^n \hat{v}_M^0 + \sum_{j=0}^{n-1} (\hat{A}_S)^{j-1} \hat{g}_M. \qquad (7.4b)$$

Following [BCR], [EOZ] and [ACD] we use the following algorithm for the fast evaluation of (7.4a) for $n = 2^m$:

(i) Set

$$B = MAM^{-1}$$

$$C = I$$

(ii)

$$\left\{ \begin{array}{l} \text{DO} \ m \ \text{times} \\[2mm] C = tr(C + BC; \varepsilon) \\[2mm] B = tr(BB; \varepsilon) \end{array} \right. \qquad (7.5)$$

Calculate

(iii)

$$v^n = M^{-1}(BMv^0 + CMg);$$

here $tr(A; \varepsilon)$ denotes the truncation operation

$$[tr(A; \varepsilon)]_{ij} = \left\{ \begin{array}{ll} A_{ij} & \text{if } |A_{i,j}| \geq \varepsilon \\[2mm] 0 & \text{if } |A_{ij}| < \varepsilon \end{array} \right. . \qquad (7.6)$$

Fredholm integral equations of the second kind are usually solved by iterative procedures of the form (6.1a). In this case the gain in efficiency which is offered by algorithm (7.5) is based both on the compression of the operator A by

$$A \longrightarrow tr(\hat{A}_s; \varepsilon) \qquad (7.7)$$

and using the fact that the T-sparseness of $(\hat{A}_S)^n$ remains constant as n increases.

7.B. Initial Value Problems.

Consider the evolution equation

$$\partial_t u + \mathcal{L}(x, \partial_x)u = f(x) \quad , x \in \Omega \quad t > 0$$
$$u(x, 0) = u_0(x) \tag{7.8}$$

with boundary conditions, where $\mathcal{L}$ is a differential operator. An explicit discretization typically takes the form (6.1a)

$$v^{n+1} = Av^n + g \tag{7.9a}$$

where

$$v_j^n \approx u(x_j, t_n), \quad g_j \approx \Delta t f(x_j), \tag{7.9b}$$

$t_n = n\Delta t$ and $\{x_j\}$ is a grid in Ω.

It is shown in [EOZ] that using the multiresolution algorithm (7.5) to calculate large-time solutions of one-dimensional hyperbolic problems to a fixed predetermined accuracy can be reduced from the standard $O((N_L)^2)$ to $O(N_L(\log_2 N_L)^3)$. For parabolic equations, a standard explicit calculation with complexity $O((N_L)^3)$ can be likewise reduced by (7.5) to $O(N_L(\log_2 N_L)^3)$. The multiresolution algorithm (7.5) in [EOZ] is based on Daubechies' orthonormal wavelets. In [ACD] this algorithm is extended to point value and cell-average discretizations.

As an example let us consider the simplest hyperbolic problem

$$u_t + u_x = 0 \tag{7.10a}$$

and its solution by the Lax-Wendroff scheme

$$v_i^{n+1} = v_i^n - \frac{\lambda}{2}(v_{i+1}^n - v_{i-1}^n) + \frac{\lambda^2}{2}(v_{i+1}^n - 2v_i^n + v_{i-1}^n) =: (Av^n)_i \tag{7.10b}$$

where $\lambda = \Delta t / h_L$.

The matrix A is a tridiagonal matrix and thus has $3N_L$ non-zero entries. On the other hand $\hat{A}_S$, the multiresolution form of the scheme, has a finger-like structure with $O(N_L \log_2 N_L)$ non-zero entries. In Figure 3, which is taken from [AC], we show the multiresolution form of the Lax-Webdroff scheme for $N_L = 64$. This shows that unlike the application to iterative solution of integral equations where (7.7) results in a "compressed" representation of the operator, the gain in efficiency in large-time computation of hyperbolic problems is only due to the uniform T-sparseness of powers of $\hat{A}_S$, i.e. while $(A)^n$ fills up linearly in n, the T-sparseness of $(\hat{A}_S)^n$ remains $O(N_L \log_2 N_L)$ for all n.

In the following we describe another application of the multiresolution form (6.1b) which uses data-compression of the numerical solution v^n and the finite-speed of propagation in

hyperbolic problems in order to produce a multiresolution analog to adaptive grids. Let

$$\Gamma_\varepsilon(v^n) = \{(j,k)| \quad |d_j^k(v^n)| \ge \varepsilon_k\} \tag{7.11a}$$

denote the domain in the $(j-k)$ plane which contains all the significant scale-coefficients of v^n, and let $\tilde{\Gamma}_\varepsilon^n$ denote the domain which is obtained by enlarging $\Gamma_\varepsilon(v^n)$ by adding side-neighbors of the cells in $\Gamma_\varepsilon(v^n)$ and allowing for a growth of one scale per time-step where needed. Due to the finite speed of propagation in hyperbolic problems

$$\Gamma_\varepsilon(v^{n+1}) \subset \tilde{\Gamma}_\varepsilon^n. \tag{7.11b}$$

Therefore we can set the components of $\hat{v}_M^{n+1}$ which are not listed in $\tilde{\Gamma}_\varepsilon^n$ to zero, and evaluate the product $\text{row}(\hat{A}_S)$ times $\hat{v}_M^n$ only for those rows which are listed in $\tilde{\Gamma}_\varepsilon^n$. The computational work can be further reduced by taking into account the T-sparseness of $\hat{v}_M^n$.

This technique can be extended to nonlinear problems. In [LT], [MR] and [BMP] it is shown how to derive a multiresolution scheme for the numerical solution of Burgers' equation

$$u_t + u u_x = \nu u_{xx}, \qquad \nu > 0 \tag{7.12}$$

in which the time-evolution is restricted to the significant scale-coefficients of the numerical solution. This numerical scheme is obtained by a Galerkin approach in which the PDE is projected on a basis of wavelets. We remark that this Galerkin-type scheme is not suitable for the "inviscid" Burgers' equation ($\nu = 0$ in (7.12)) in the sense that it generates spurious oscillations at shocks, and may even become unstable in some cases – thus some form of artificial viscosity is needed.

In [H4] and [H5] we consider a hyperbolic system of conservation laws

$$u_t + f(u)_x = 0 \tag{7.13a}$$

and its numerical solution by *any* scheme in conservation form

$$v_j^{n+1} = v_j^n - \lambda(\bar{f}_j - \bar{f}_{j-1}) \tag{7.13b}$$

where

$$\bar{f}_j = \bar{f}(v_{j-K+1}^n, \ldots, v_{j+K}^n) \quad \text{for some} \quad K \ge 1, \tag{7.13c}$$

and $\bar{f}$ is the numerical flux function. Observe that the computational task here is the evaluation of the numerical flux function (7.13c) at all the gridpoints. Using the multiresolution form of the numerical scheme (7.13b) with respect to cell-average discretization we derive an algorithm for the time-evolution of the scale-coefficients, and show that data compression of

23

the numerical solution can be translated into reduction of the number of flux calculations in (7.13c).

In Figure 4 we show the results of [H5] for the multiresolution form of the Lax-Wendroff scheme which is applied to the periodic initial value problem for the "inviscid" Burgers' equation.

$$u_t + uu_x = 0 \quad, -1 \leq x \leq 1 \quad, t > 0$$

$$u(x,0) = 2 + sin(\pi x)$$

(7.14)

with $J_L = 256$, $CFL = 0.8$, and tolerance $\varepsilon = 10^{-3}$. Figure 4 consists of 3 snapshots corresponding to $n = 25, 150, 400$ time-steps. In the upper part of each snapshot we compare the solution of the MR scheme (circles) to the solution of Lax-Wendroff scheme on the finest grid (continuous line), which is computed independently. In the lower part of each snapshot we display $\Gamma_\varepsilon(v^n)$ (circles) and its corresponding estimate $\tilde{\Gamma}_\varepsilon^n$ (dots). This is done by drawing the symbol at (x_{2j-1}^{k-1}, k) in the $x - k$ plane for each (j, k) in the set; note that due to a different notation in [H5] $k = 0$ is the finest grid, and $k = L = 5$ is the coarsest. In the Table we list the efficiency (i.e. the ratio between the fine grid calculation of 256 fluxes over the number that we actually had to compute) and the difference $E_m, m = 1, 2, \infty$ in the corresponding norm between the solution of the MR scheme and the independent finest-grid calculation.

In [BH1] we extend this technique to the numerical solution of

$$u_t + div \ f(u) = 0$$

(7.15)

on Cartesian grids, and in [AH2] we generalize it further to unstructured meshes where the coarser levels of resolution are obtained by agglomeration of cells.

8. Conclusions.

In this paper we reviewed recent developments in techniques to represent data in terms of its *local* scale components. These techniques enabled us to obtain data compression by eliminating scale-coefficients which are sufficiently small. This capability for data compression can be used to reduce the cost of many numerical solution algorithms by either applying it to the numerical solution operator in order to get an approximate sparse representation, or by applying it to the numerical solution itself in order to reduce the number of quantities that need to be computed.

24

Appendix: The Dual MR Scheme.

In this appendix we describe the MR scheme which is dual to the one in (3.10)-(3.11). In order to better see the duality we rewrite (3.10)-(3.11) as follows: we express the direct MR transform

$$\hat{v}_M = M \cdot v^L \qquad (A.1a)$$

by

$$\begin{cases} \text{DO} \quad k = L, \dots, 1 \\ v^{k-1} = D_k^{k-1} v^k \\ d^k = G_k^D v^k \end{cases} \qquad (A.1b)$$

where

$$G_k^D = G_k(I_k - P_{k-1}^k D_k^{k-1}). \qquad (A.1c)$$

We observe that

$$E_k d^k \in \mathcal{N}(D_k^{k-1}) \Rightarrow (I_k - P_{k-1}^k D_k^{k-1}) E_k d^k = E_k d^k$$

and therefore we can rewrite the inverse MR transform

$$v^L = M^{-1} \hat{v}_M \qquad (A.2a)$$

as

$$\begin{cases} \text{DO} \quad k = 1, \dots, L \\ v^k = P_{k-1}^k v^{k-1} + E_k^P d^k \end{cases} \qquad (A.2b)$$

where

$$E_k^P = (I_k - P_{k-1}^k D_k^{k-1}) E_k. \qquad (A.2c)$$

To simplify our presentation we shall use the matrix representation of the various operators and define

$$\widetilde{D}_k^{k-1} =: (P_{k-1}^k)^*, \quad \widetilde{P}_{k-1}^k =: (D_k^{k-1})^*, \quad \widetilde{G}_k^{\widetilde{D}} =: (E_k^P)^*, \quad \widetilde{E}_k^{\widetilde{P}} =: (G_k^D)^*. \qquad (A.3)$$

Observe that $\widetilde{D}_k^{k-1}$ is a $J_{k-1} \times J_k$ matrix, $\widetilde{P}_{k-1}^k$ is a $J_k \times J_{k-1}$ matrix, $\widetilde{G}_k^{\widetilde{D}}$ is a $(J_k - J_{k-1}) \times J_k$ matrix, and $\widetilde{E}_k^{\widetilde{P}}$ is a $J_k \times (J_k - J_{k-1})$. With these definitions we obtain the dual MR scheme from (A.1)-(A.2) by applying $\widetilde{(\cdot)}$ to all the operators, i.e. the direct MR transform of the dual scheme

$$\hat{v}_{\widetilde{M}} = \widetilde{M} \cdot v^L \qquad (A.4a)$$

is given by

$$\begin{cases} \text{DO} \quad k = L, \dots, 1 \\ v^{k-1} = \widetilde{D}_k^{k-1} v^k \\ d^k = \widetilde{G}_k^{\widetilde{D}} v^k \end{cases} \qquad (A.4b)$$

25

where $\widetilde{G}_k^{\widetilde{D}}$ is defined in (A.3). The inverse MR transform of the dual scheme

$$v^L = \widetilde{M}^{-1} \cdot \hat{v}_{\widetilde{M}} \qquad (A.5a)$$

is given by

$$\begin{cases} \text{DO} \quad k = 1, \ldots, L \\[2mm] v^k = \widetilde{P}_{k-1}^k v^{k-1} + \widetilde{E}_k^{\widetilde{P}} d^k \end{cases} \qquad (A.5b)$$

where $\widetilde{E}_k^{\widetilde{P}}$ is defined in (A.3). Observe that the dual of the dual is the original scheme.

It follows from the above definitions that for $1 \le k \le L$

$$C_k^L = (P_{L-1}^L \cdots P_k^{k+1})E_k^P = [(E_k^P)^*(P_k^{k+1})^* \cdots (P_{L-1}^L)^*]^* = [\widetilde{G}_k^{\widetilde{D}}(\widetilde{D}_{k+1}^k \cdots \widetilde{D}_L^{L-1})]^* = (\widetilde{B}_L^k)^*$$

and

$$C_0^L = P_{L-1}^L \cdots P_0^1 = [(P_0^1)^* \cdots (P_{L-1}^L)^*]^* = [\widetilde{D}_1^0 \cdots \widetilde{D}_L^{L-1}]^* = (\widetilde{B}_L^0)^* \ ;$$

thus

$$M^{-1} = (\widetilde{M})^* \ .$$

In [H3] we also show that $\widetilde{M}$ is associated with discretization $\widetilde{\mathcal{D}}_k$ and reconstruction $\widetilde{\mathcal{R}}_k$ such that $(\widetilde{\mathcal{R}}_k\widetilde{\mathcal{D}}_k) : \mathcal{F} \to \mathcal{F}$ is the adjoint of $(\mathcal{R}_k\mathcal{D}_k)$.

Acknowledgements.

References

[AC] F. Arandiga and V. Candela, "Multiresolution standard form of a matrix", UCLA CAM Report 92-37, August 1992.

[ACD] F. Arandiga, V. Candela and R. Donat, " Fast multiresolution algorithms for solving linear equations: A comparative study", UCLA CAM Report 92-52, December 1992.

[ADH] F. Arandiga, R. Donat and A. Harten, "Multiresolution based on weighted averages of the hat function", UCLA CAM Report 93-34, September 1993.

[AH1] R. Abgrall and A. Harten, "Multiresolution analysis on unstructured meshes: preliminary report", CAM report 94-26, July 1994.

[AH2] R. Abgrall and A. Harten, "Multiresolution schemes for hyperbolic conservation laws on unstructured meshes", To be published as an INRIA report.

[BCR] G. Beylkin, R. Coifman and V. Rokhlin, "Fast wavelet tranform and numerical algorithms. I," *Comm. Pure Appl. Math.* **44**, pp. 141-183, 1991.

[BH1] B. Bihari and A. Harten, "Application of generalized wavelets: An adaptive multiresolution scheme", Preprint, October 1993; to appear in Jour. Appl. Num. Math.

[BL] A. Brandt and A.A. Lubrecht, "Multilevel matrix multiplication and fast solution of integral equations," *Jour. Comput. Phys.* **90**, pp. 348-370, 1990.

[BMP] E. Bacry, S. Mallat and G. Papanicolau, "A wavelet based space-time adaptive numerical methods for partial differential equations," *Mathematical Modelling and Numerical Analysis* **26**, pp. 703-834, 1992.

[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.

[Da] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.* **41**, pp. 909-996, 1988.

[EOZ] B. Engquist, S. Osher and S. Zhong, "Fast wavelet algorithms for linear evolution equations," ICASE Report 92-14, 1992.

[H1] A. Harten, "Discrete multiresolution analysis and generalized wavelets," *J. Appl. Num. Math.* **12**, pp. 153-193, 1993; also UCLA CAM Report 92-08, February 1992.

[H2] A. Harten, "Multiresolution representation of data. I. Preliminary Report," UCLA CAM Report No. 93-13, June 1993.

[H3] A. Harten, "Multiresolution representation of data. II. General Framework," UCLA CAM Report No. 94-10, April 1994; also Technical Report 3-94, Dept. of Applied Mathematics, Tel Aviv University, April 1994.

[H4] A. Harten, "Multiresolution algorithms for the numerical solution of hyperbolic conservation laws," UCLA CAM Report 93-03, March 1993; to appear in *Comm. Pure Appl. Math.*

[H5] A. Harten, "Adaptive multiresolution schemes for shock computations," UCLA CAM Report 93-06, April 1993; to appear in *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.*

[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.* **315**, pp. 69-87, 1989.

[Me] Y. Meyer, "Ondelettes et Opérateurs," Hermann, 1990.

[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.

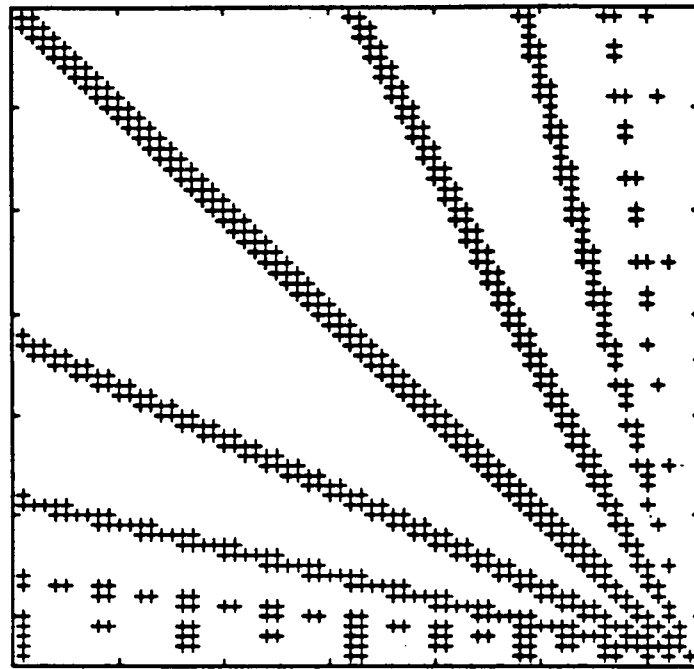[S] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Review* **31**, pp. 614-627, 1989.
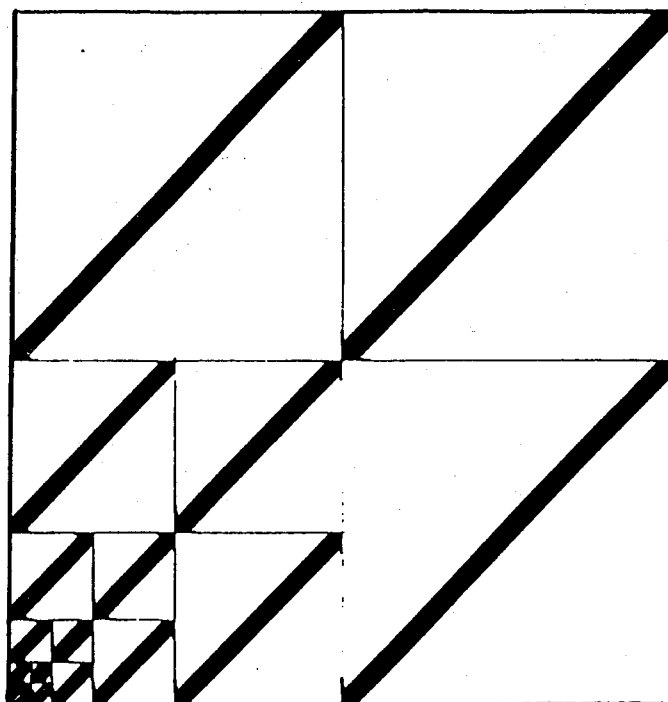
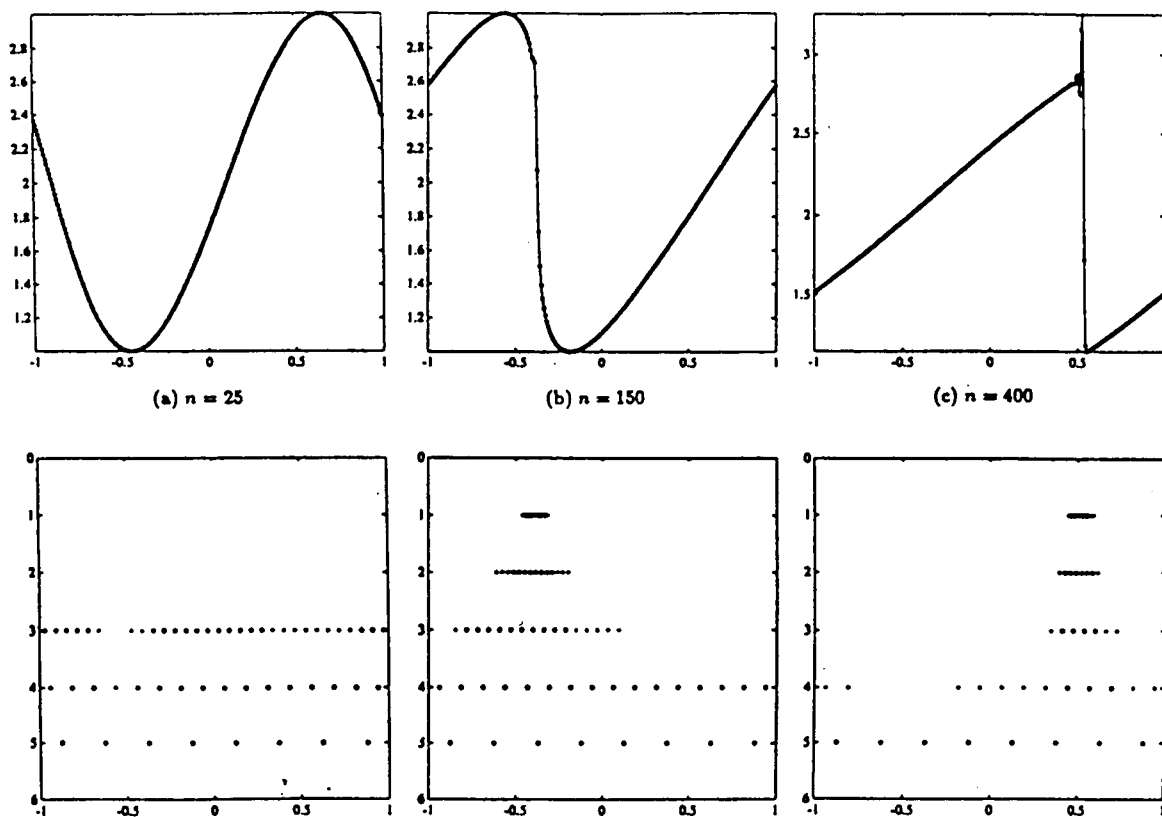**Figure 1.** Multiresolution representation of the 2-D array .
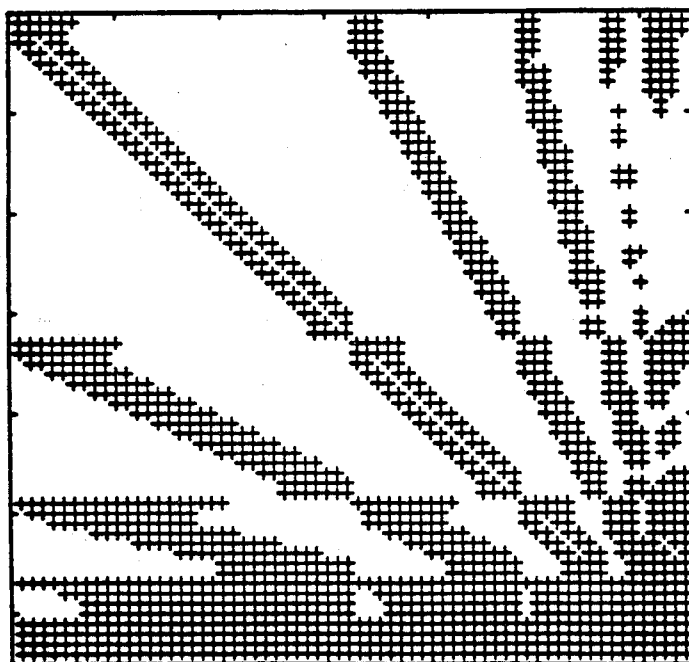
Figure 2. Multiresolution form of the operator (matrix).

29

Figure 3. Multiresolution form of the Lax-Wendroff scheme. (linear advection)

(a) n = 25

(b) n = 150

(c) n = 400

Figure 4. Multiresolution application of the Lax-Wendroff scheme. (Inviscid Burgers' equation).

30

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1994	Contractor Report

4. TITLE AND SUBTITLE
MULTIRESOLUTION REPRESENTATION AND NUMERICAL ALGORITHMS: A BRIEF REVIEW

5. FUNDING NUMBERS
C NAS1-18605, NAS1-19480
WU 505-90-52-01

6. AUTHOR(S)
Ami Harten

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Computer Applications in Science
and Engineering
Mail Stop 132C, NASA Langley Research Center
Hampton, VA 23681-0001

8. PERFORMING ORGANIZATION REPORT NUMBER
ICASE Report No. 94-59

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA CR-194949
ICASE Report No. 94-59

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Michael F. Card
Final Report
Proceedings from the ICASE/LaRC Workshop on Parallel Numerical Algorithms

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified–Unlimited
Subject Category 64

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
In this paper we review recent developments in techniques to represent data in terms of its *local* scale components. These techniques enable us to obtain data compression by eliminating scale-coefficients which are sufficiently small. This capability for data compression can be used to reduce the cost of many numerical solution algorithms by either applying it to the numerical solution operator in order to get an approximate sparse representation, or by applying it to the numerical solution itself in order to reduce the number of quantities that need to be computed.

14. SUBJECT TERMS
multiresolution representation, fast multiresolution transforms, generalized wavelets

15. NUMBER OF PAGES
34

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		