Journal of Approximation Theory **96**, 258–280 (1999)
Article ID jath.1997.3234, available online at http://www.idealibrary.com on IDEAL®

Adaptive Multiresolution Analysis on the Dyadic Topological Group

Bl. Sendov*

Center of Informatics and Computer Technology, Bulgarian Academy of Sciences, "Acad. G. Bonchev" street, Block 25A, 1113 Sofia, Bulgaria
E-mail: sendov@amigo.acad.bg

Communicated by Charles K. Chui

Received March 31, 1997; accepted in revised form December 9, 1997

A type of multiresolution analysis on the space of continuous functions defined on the dyadic topological group is proposed, depending on free parameters. The appropriate choice of parameters is used to adapt this analysis to a given function. © 1999 Academic Press

1. INTRODUCTION

The classical multiresolution analysis in the space of the square-integrable functions $f \in L^2(R)$, where R is the set of real numbers, is based on a sequences of closed subspaces $\{V_j\}_{j \in Z}$ and a set of operators d, $\{t_{j,k}\}_{j,k \in Z}$, where Z is the set of integers, which satisfy the conditions

$$V_j \subset V_{j+1} \qquad \forall j \in Z, \tag{1.1}$$

$$\bigcup_{j=-\infty}^{+\infty} V_j \text{ is dense in } L^2(R) \text{ and } \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}, \tag{1.2}$$

$$f(x) \in V_j \Leftrightarrow d(f; x) = f(2x) \in V_{j+1} \qquad \forall j \in Z, \tag{1.3}$$

$$f(x) \in V_j \Rightarrow t_{j,k}(f; x) = f(x - 2^{-j}k) \in V_j \qquad \forall j, k \in Z. \tag{1.4}$$

The definitions of the operators d, $\{t_{j,k}\}_{j,k \in Z}$ explore transformations in the group R, dilate, and translate.

In this paper, we define a type of multiresolution analysis on $L^2(\mathbf{G})$, where $\mathbf{G}$ is the dyadic topological group [8], which is compact.

There are many general and profound studies of multiresolution analysis over different groups, see for example [1] and [5]. Our study differs in

* Supported by contract I-409/94 with the Bulgarian national research foundation.

0021-9045/99 $30.00

one important point from [5], where the fundamental theorem for the multiresolution analysis over dyadic group is proved. The study in [5] is oriented to "*locally compact* Abelian (LCA) groups, resembling the familiar constructions of Y. Meyer [7] or S. Mallat [6] on n-dimensional Euclidean space." We consider the *compact* dyadic group, in which the *diletes* do not work. The compact dyadic group is important for the applications as it is isomorphic to the unit interval $[0, 1]$ of real numbers, to the unit square $[0, 1]^2$, to the unite cube $[0, 1]^3$, and so on. The modifications of the conditions (1.1)–(1.4), which we use, replacing Z with Z^+, the natural numbers, and the operators in (1.3) and (1.4) are different as it is impossible to use dilate in the compact dyadic group.

To give an example of a multiresolution analysis over the compact dyadic group, we start by defining the *dyadic topological group* **G** itself.

G is the infinite direct product of the Abelian group with two elements 0 and 1, and group operation $\dot{+}$, addition modulo 2. Thus the dyadic group **G** is the set of all 0, 1 sequences $\mathbf{x} = (x_1, x_2, x_3, ...)$ in which the group operation is

$$\mathbf{x} \dot{+} \mathbf{y} = (x_1 \dot{+} y_1, x_2 \dot{+} y_2, x_3 \dot{+} y_3, ...).$$

The dyadic group is a compact metric space [8] with distance

$$\beta(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} 2^{-i} |x_i - y_i| = \sum_{i=1}^{\infty} 2^{-i}(x_i \dot{+} y_i). \tag{1.5}$$

Set the notation

$$\mathbf{2}^{-k} = (\underbrace{0, 0, 0, ..., 0,}_{k-1} 1, 0, 0, 0, ...),$$

and the correspondence l: $\mathbf{G} \to [0, 1]$,

$$l(\mathbf{x}) = l((x_1, x_2, x_3, ...)) = x = \sum_{i=1}^{\infty} 2^{-i} x_i \in [0, 1]. \tag{1.6}$$

If $f = f(\mathbf{x})$ is a function defined on **G**, then the function

$$\tilde{f}(x) = \tilde{f}(l(\mathbf{x})) = f(\mathbf{x}) \tag{1.7}$$

is a function defined on $[0, 1]$. We shall comment later on how to define $\tilde{f}$ when one real number $x \in [0, 1]$ corresponds to two different (binary rational) elements from **G**.

The example we have in mind, of a multiresolution analysis over the compact dyadic group $\mathbf{G}$, is defined by the following conditions [11] corresponding to (1.1)–(1.4)

$$V_j \subset V_{j+1} \qquad \text{for} \quad j=0, 1, 2, ..., \tag{1.8}$$

$$\bigcup_{j=0}^{+\infty} V_j \text{ is dense in } L^2(\mathbf{G}) \text{ and } V_0 \text{ is one dimensional,} \tag{1.9}$$

$$f(\mathbf{x}) \in V_j \Rightarrow f(\mathbf{x}) \in V_{j+1} \qquad \text{for} \quad j=0, 1, 2, ..., \tag{1.10}$$

$$f(\mathbf{x}) \in V_j \Rightarrow t_{j+1}(f; \mathbf{x}) = (-1)^{x_{j+1}} f(\mathbf{x} + \mathbf{2}^{-j-1}) \in V_{j+1} \tag{1.11}$$

for $j=0, 1, 2, ...$ In fact (1.10) is to ensure (1.8).

Let $\phi_0(\mathbf{x}) = 1$, $\mathbf{x} \in \mathbf{G}$ and V_0 be the span over ϕ_0. V_0 is one dimensional subspace of $L^2(\mathbf{G})$. If we apply (1.10) and (1.11), it is easy to see that V_j contains 2^j orthonormal functions $\{\phi_n\}_{n=0}^{2^j-1}$, such that $\{\tilde{\phi}_n\}_{n=0}^{2^j-1}$ are the first 2^j Walsh functions. From this it follows that (1.9) is satisfied.

These 2^j orthonormal functions are "translates" of the function ϕ_0,

$$\phi_1(\mathbf{x}) = t_1(\phi_0; \mathbf{x}), \qquad \phi_2(\mathbf{x}) = t_2(\phi_0; \mathbf{x}), \qquad \phi_3(\mathbf{x}) = t_2(\phi_1; \mathbf{x}) = t_2 t_1(\phi_0; \mathbf{x}),$$

$$\phi_4(\mathbf{x}) = t_3(\phi_0; \mathbf{x}), \qquad \phi_5(\mathbf{x}) = t_3(\phi_1; \mathbf{x}) = t_3 t_1(\phi_0; \mathbf{x}), ...,$$

$$\phi_{2^j-1}(\mathbf{x}) = t_j t_{j-1} t_{j-2} \cdots t_1(\phi_0; \mathbf{x}).$$

Following the analogy we may call the function ϕ_0 a wavelet. To distinguish between the different types of multiresolution analysis we have in hand, we call ϕ_0 *first function.*

Our goal further will be to find different first functions. A class of first functions is defined (Definition 2.5), called dyadic exponential functions. The reason for this name is that $\phi(\mathbf{x})$ is a first function if $\tilde{\phi}(x) = e^{\alpha x}$, where α is a real number. We do not know wether there exist first functions which are not dyadic exponential functions.

The dyadic exponential functions depend on free parameters. These parameters may be used to adapt the multiresolution analysis to an arbitrary given function.

Another difference of the considered multiresolution analysis is that its adaptation depends on continuous parameters. In the classical case of Malvar wavelets and wavelet packets [7], the adaptation depends on discrete choices.

In Section 2 some definitions and statements for the topological dyadic group $\mathbf{G}$ and for the real continuous functions $f \in \mathscr{C}$, defined on $\mathbf{G}$, are presented. The definitions of *dyadic exponential function* and *Rademacher transformation* of rank 1, 2, 3 are given, which are used to define the operators for multiresolution analysis.

In Section 3 the theorem for multiresolution analysis in $\mathscr{C}$ is proved. In Section 4 the adaptation of a first function to a given function is considered. A criterion (see Definition 4.1) for adaptation of the considered multiresolution analysis, inspired by the criterion of R. Coifman and V. Wickerhauser [3], is formulated and a theorem, solving the problem for practical implementation of this criterion, is proved.

2. DYADIC FUNCTIONS

We start this section with some definitions and properties of the dyadic group $\mathbf{G}$ and the isomorphisms between $\mathbf{G}$ and the unit cube $[0, 1]^n$, $n = 1, 2, 3, \ldots$.

2.1. *Dyadic Topological Group*

In the Introduction the group $\mathbf{G}$ was introduced. Now set the additional notations

$$\mathbf{0} = (0, 0, 0, \ldots), \qquad \mathbf{1} = (1, 1, 1, \ldots),$$

and

$$\mathbf{G}_{nk}^{(j)} := \{\mathbf{x} : \mathbf{x} \in \mathbf{G},\ 2^{n-1}x_{n(k-1)+1} + 2^{n-2}x_{n(k-1)+2} + \cdots + x_{nk} = j\}.$$

An element $\mathbf{x}$ of $\mathbf{G}$ is called *left rational* if it has finite number of coordinates equal to 0 and *right rational* if it has finite number of coordinates equal to 1. The elements $\mathbf{0}$, $\mathbf{2}^{-k}$ are right rational, and $\mathbf{1}$ is left rational.

The order in $\mathbf{G}$ is lexicographic. For every element $\mathbf{x} \in \mathbf{G}$ we have $\mathbf{0} \leqslant \mathbf{x} \leqslant \mathbf{1}$. For $\mathbf{x} < \mathbf{y}$ denote by $[\mathbf{x}, \mathbf{y}]$ the set of all $\mathbf{z} \in \mathbf{G}$ such that $\mathbf{x} \leqslant \mathbf{z} \leqslant \mathbf{y}$.

Definition 2.1. Let

$$\mathbf{x} = (x_1, x_2, x_3, \ldots, x_k, 0, 0, 0, \ldots), \qquad \mathbf{y} = (x_1, x_2, x_3, \ldots, x_k, 1, 1, 1, \ldots).$$

The segment

$$[\mathbf{x}, \mathbf{y}] = [k; p], \qquad \text{where} \quad p = 2^{k-1}x_1 + 2^{k-2}x_2 + \cdots + 2x_{k-1} + x_k$$

is called a *pixel* of *rank* k.

Every pixel $[k; p]$ of rank k consists of two pixels $[k+1; 2p]$ and $[k+1; 2p+1]$ of rank $k+1$.

It is obvious that

$$[k; 0] \cup [k; 1] \cup [k; 2] \cup \cdots \cup [k; 2^k - 1] = [0; 0] = [\mathbf{0}, \mathbf{1}] = \mathbf{G},$$

where $[k; p] \cap [k; q] = \varnothing$ for $p \neq q$.

If $\mathbf{x} \in [k; p]$, then $x = l(\mathbf{x}) \in [p2^{-k}, (p+1)\, 2^{-k}]$.

The function l, defined by (1.6), does not have a single-valued inverse on the rational elements of $\mathbf{G}$. Consider the right rational in this case, with one exception. Let m: $[0, 1] \to \mathbf{G}$ be the inverse of the function l, with the agreement that if $x \in [0, 1]$ is a binary rational number, then $m(x)$ is the right rational element of $\mathbf{G}$ corresponding to x. The exception is for $1 \in [0, 1]$, namely, $m(1) = \mathbf{1} = (1, 1, 1, ...)$, which is left rational. Thus, for every $x \in [0, 1]$, $l(m(x)) = x$, and for every $\mathbf{x} \in \mathbf{G}$, $m(l(\mathbf{x})) = \mathbf{x}$. By the functions l and m, a correspondence between $\mathbf{G}$ and $[0, 1]$ is defined.

To define a correspondence between $\mathbf{G}$ and the unit square $[0, 1]^2$, we may proceed as follows. For every $\mathbf{x} = (x_1, x_2, x_3, ...) \in \mathbf{G}$, let

$$\xi = \sum_{i=1}^{\infty} 2^{-i} x_{2i} \in [0, 1], \qquad \eta = \sum_{i=1}^{\infty} 2^{-i} x_{2i+1} \in [0, 1],$$

and

$$l_2(\mathbf{x}) = (\xi, \eta) \in [0, 1]^2.$$

The inverse function m_2 to l_2 is defined obviously by m, the inverse function of l.

It is clear that there exist many different correspondences between $\mathbf{G}$ and $[0, 1]^n$ for $n = 3, 4, ...$.

2.2. *Continuous Dyadic Functions*

A function f, defined on $\mathbf{G}$ with real values $f(\mathbf{x}) \in R$, is called a *dyadic function*. The set of all these functions f: $\mathbf{G} \to R$ is denoted by $\mathcal{G}$.

Dyadic functions have been considered by many authors, see for example [4], in connection with the Walsh functions and other problems. Usually, the dyadic argument is used only for convenience and the topology of the group $\mathbf{G}$ is not explored for the properties of the considered functions. In this paper, the topology of $\mathbf{G}$ is used to define the continuity of a dyadic function [8].

The set of points $(\mathbf{x}, f(\mathbf{x})) \in \mathbf{G} \times R$ is the graph of the dyadic function f, which is impossible to draw on paper. Usually we consider the set of points $(l(\mathbf{x}), f(\mathbf{x})) \in [0, 1] \times R$, where l is defined by (1.6), as a "graphical representation" of the dyadic function f. When the values of f, for two rational elements $\mathbf{x}$, $\mathbf{x}' \in \mathbf{G}$ corresponding to one real number $x = l(\mathbf{x}) = l(\mathbf{x}')$, are different, we join the points $(x, f(\mathbf{x}))$, $(x, f(\mathbf{x}'))$ with a vertical segment.

The graphical representation of a dyadic function $f(\mathbf{x})$ is the *completed graph* [9, 10] of the multi-valued function $\tilde{f}(x) = f(\mathbf{x})$, where $x = l(\mathbf{x})$.

This procedure for drawing a graph of a function is widely accepted and used without comments. In [5], the graphs of the dyadic functions are depicted in the same manner.

The uniform norm of a bounded function $f \in \mathscr{G}$ is defined as

$$\|f\| = \sup\{|f(\mathbf{x})|;\ \mathbf{x} \in \mathbf{G}\}.$$

DEFINITION 2.2. A dyadic function $f \in \mathscr{G}$ is continuous in the point $\mathbf{x}_0 \in \mathbf{G}$ if for every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$, such that

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| < \varepsilon \qquad \text{for} \quad \beta(\mathbf{x}, \mathbf{x}_0) < \delta.$$

If a dyadic function f is continuous at every point $\mathbf{x} \in \mathbf{G}$, then it is uniformly continuous, as $\mathbf{G}$ is compact [8].

The set of all uniformly continuous dyadic functions is denoted by $\mathscr{C}$.

The module of continuity of a function $f \in \mathscr{G}$ is

$$\omega(f; \delta) = \sup\{|f(\mathbf{x}) - f(\mathbf{y})| : \beta(\mathbf{x}, \mathbf{y}) < \delta,\ \mathbf{x}, \mathbf{y} \in G\}.$$

A necessary and sufficient condition for a function $f \in \mathscr{G}$ to be uniformly continuous ($f \in \mathscr{C}$) is

$$\lim_{\delta \to +0} \omega(f; \delta) = 0. \tag{2.12}$$

DEFINITION 2.3. A function $f \in \mathscr{G}$ is called a *pixel function* of *rank* k if $f(\mathbf{x}) = c_p = \text{constant}$ for $\mathbf{x} \in [k; p]$; $p = 0, 1, 2, ..., 2^k - 1$. The set of all pixel functions of rank k is denoted by $\mathscr{P}_k$.

It is obvious that $\mathscr{P}_0 \subset \mathscr{P}_1 \subset \mathscr{P}_2 \subset \cdots \subset \mathscr{P}_s \subset \cdots$ and $\overline{\bigcup_{s=0}^{\infty} \mathscr{P}_s} = \mathscr{C}$.

From the definition of the module of continuity it follows directly that the necessary and sufficient condition for a function $P \in \mathscr{G}$ to be a pixel function of rank k is

$$\omega(P; \delta) = 0 \qquad \text{for} \quad \delta \leqslant 2^{-k}. \tag{2.13}$$

The equality (2.13) shows that every pixel function is a continuous dyadic function. Let us point out that a dyadic pixel function f, defined on $\mathbf{G}$, is continuous, but the function $\tilde{f}$, defined on [0, 1] is a step function and it is not continuous.

2.3. *Dyadic Exponential Functions*

Let $p = p_0 + p_1 2 + p_2 2^2 + \cdots + p_s 2^s$ be the binary representation of the integer p, where $p_i = 0, 1$ are the binary digits of p. For a natural number n define

$$p_{n,i} = p_{ni} + p_{ni+1} 2 + \cdots + p_{ni+n-1} 2^{n-1}. \tag{2.14}$$

The addition $\dot{+}$ (addition modulo 2) of two integers m, n is defined as

$$m \dot{+} n = (m_0 \dot{+} n_0) + (m_1 \dot{+} n_1)\, 2 + (m_2 \dot{+} n_2)\, 2^2 + \cdots + (m_s \dot{+} n_s)\, 2^s.$$

For the natural number $p < 2^k$, the element $p\mathbf{2}^{-k} \in \mathbf{G}$ is defined as

$$\begin{aligned} p\mathbf{2}^{-k} &= (p_{k-1}, p_{k-2}, ..., p_0, 0, 0, 0, ...), \quad \text{or} \quad 1 \cdot \mathbf{2}^{-k} = \mathbf{2}^{-k}, \\ 2 \cdot \mathbf{2}^{-k} &= \mathbf{2}^{-k+1}, \quad 3 \cdot \mathbf{2}^{-k} = \mathbf{2}^{-k+1} + \mathbf{2}^{-k} \text{ and so on.} \end{aligned} \tag{2.15}$$

For $\mathbf{x} = (x_1, x_2, x_3, ...) \in \mathbf{G}$ define

$$x_{n,i} = x_{ni+1} 2^{n-1} + x_{ni+2} 2^{n-2} \cdots + x_{ni+n}. \tag{2.16}$$

Let $\Lambda = \{\lambda_i(u)\}_{i=1}^{\infty}$ be a sequence of real functions defined for $u = 0, 1, 2, ..., 2^n - 1$, $\lambda_i(0) = 1$ and

$$\bar{\lambda}_i = \max\{\lambda_i(u) : u = 0, 1, ..., 2^n - 1\},$$

$$\underline{\lambda}_i = \min\{\lambda_i(u) : u = 0, 1, ..., 2^n - 1\}.$$

Set the notations

$$\lambda_k = (\lambda_k(0), \lambda_k(1), \lambda_k(2), ..., \lambda_k(2^n - 1)),$$

the 2^n dimensional vector of the values of the function $\lambda_k(\cdot)$,

$$\langle \lambda_k, \mu_k \rangle = \sum_{j=0}^{2^n - 1} \lambda_k(j)\, \mu_k(j) \qquad \text{and} \qquad \|\lambda_k\| = \sqrt{\langle \lambda_k, \lambda_k \rangle}.$$

Definition 2.4. Call the sequence $\Lambda = \{\lambda_k\}_{k=1}^{\infty}$ *normal* of *rank* n if the products

$$\prod_{i=1}^{\infty} \bar{\lambda}_i \qquad \text{and} \qquad \prod_{i=1}^{\infty} \underline{\lambda}_i$$

are convergent.

From Definition 2.4 it follows that if Λ is a normal sequence of rank n, then exists a number $N(\Lambda)$ such that

$$1-2^{-n} \leqslant \underline{\lambda}_i^2 \leqslant \bar{\lambda}_i^2 \leqslant 1-2^{-n} \qquad \text{for} \quad i > N(\Lambda),$$

and for $j = 0, 1, 2, ..., 2^n - 1$

$$\lim_{k\to\infty} \prod_{i=k}^{\infty} \lambda_i(j) = \lim_{k\to\infty} \prod_{i=k}^{\infty} \bar{\lambda}_i = \lim_{k\to\infty} \prod_{i=k}^{\infty} \underline{\lambda}_i = 1.$$

A characteristic of a normal sequence Λ is

$$\varepsilon(\Lambda; 2^{-s}) = \max\left\{ \prod_{i=s+1}^{\infty} \bar{\lambda}_i^2 - 1, 1 - \prod_{i=s+1}^{\infty} \underline{\lambda}_i^2 \right\} \to 0 \qquad \text{for} \quad s\to\infty. \tag{2.17}$$

Definition 2.5. Let Λ be a normal sequence; then the function

$$\Lambda(\mathbf{x}) = c \prod_{i=0}^{\infty} \lambda_{i+1}(x_{n,i}); \qquad c = \Lambda(\mathbf{0})$$

is called a *dyadic exponential* function with sequence Λ of *rank n*.

For $i > N(\Lambda)$ the inequalities

$$\underline{\lambda}_i^2 \leqslant \frac{2^{n/2}\lambda_i(j)}{\|\lambda_i\|} \leqslant \bar{\lambda}_i^2; \qquad j = 0, 1, 2, ..., 2^n - 1$$

hold, and hence

$$\left| 1 - \prod_{i=s}^{\infty} \frac{2^{n/2}\lambda_{i+1}(x_{n,i})}{\|\lambda_{i+1}\|} \right| \leqslant \varepsilon(\Lambda; 2^{-s}). \tag{2.18}$$

The dyadic exponential function Λ is continuous as

$$\omega(\Lambda; 2^{-ns}) = c\left(\prod_{i=s+1}^{\infty} \bar{\lambda}_i - \prod_{i=s+1}^{\infty} \underline{\lambda}_i \right) = 2c\varepsilon(\Lambda, 2^{-s}) \to 0 \qquad \text{for} \quad s\to\infty.$$

The function

$$g(\mathbf{x}) = \prod_{i=1}^{\infty} e^{\alpha 2^{-i} x_i}, \qquad \tilde{g}(x) = e^{\alpha x}$$

is a dyadic exponential function of rank 1 with the sequence $\lambda_k = (1, \alpha e^{2^{-k}})$, $k = 1, 2, 3, ...$.

In Fig. 1 the completed graph of a dyadic exponential function Λ of rank 2 is depicted with the sequence shown in Table I.

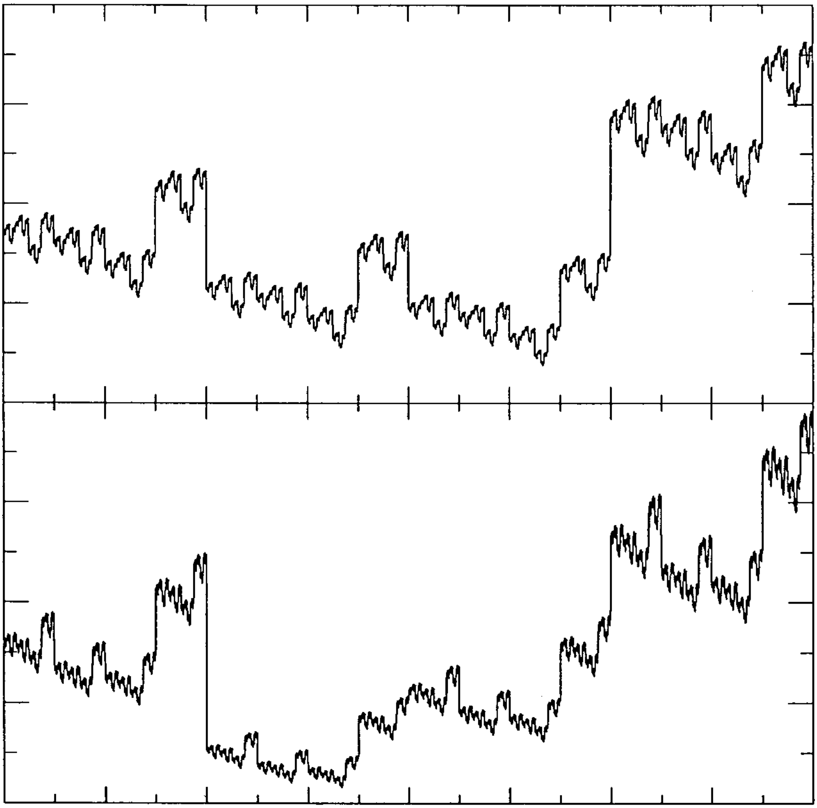

FIG. 1. Bottom: The complete graph of a dyadic exponential function Λ of rank 2 with sequence on Table I. Top: One quarter of the completed graph (supported by a pixel of rank 2) of Λ expanded (zoomed).

From Definition 2.5 and (2.15) it follows that

$$f(\mathbf{x}+j\mathbf{2}^{-nk}) = \lambda_k(j)\, f(\mathbf{x}) \qquad \text{for} \quad \mathbf{x} \in \mathbf{G}_{nk}^{(0)}, \tag{2.19}$$

$j = 1, 2, 3, ..., 2^n - 1$, $k = 1, 2, 3, ...$

The relation (2.19) shows a type of self-similarity of the dyadic exponential functions, which is seen in Fig. 1.

If for the sequence of the dyadic exponential function f, $\lambda_k(j) = 1$ for $j = 0, 1, 2, ..., 2^n - 1$ and $k = s+1, s+2, s+3, ...$, then f is a pixel function of rank ns.

TABLE I

The sequence of Λ

i	$\lambda_1(i)$	$\lambda_2(i)$	$\lambda_3(i)$	$\lambda_4(i)$	$\lambda_5(i)$	$\lambda_6(i)$	$\lambda_6(i)$	$\cdots$
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	$\cdots$
1	0.600	0.900	0.980	1.015	1.010	1.002	1.000	$\cdots$
2	0.825	0.866	0.938	0.864	0.985	0.995	1.000	$\cdots$
3	1.400	1.200	1.077	1.020	1.005	1.003	1.000	$\cdots$

Let f, g be dyadic exponential functions with sequences $\{\lambda_k\}_{k=1}^{\infty}$, $\{\mu_k\}_{k=1}^{\infty}$; then from Definition 2.5 and (2.19) it follows that

$$\begin{aligned}\langle f, g\rangle &= \int_{[0;0]} f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x} = \sum_{j=0}^{2^n-1} \int_{[n;j]} f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x} \\ &= \sum_{j=0}^{2^n-1} \lambda_{1,j}\mu_{1,j} \int_{[n;0]} f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x} = \frac{\langle \lambda_1, \mu_1\rangle}{2^n}\, 2^n \int_{[n;0]} f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x} \\ &= \cdots = f(\mathbf{0})\, g(\mathbf{0}) \prod_{i=1}^{\infty} \frac{\langle \lambda_i, \mu_i\rangle}{2^n},\end{aligned} \tag{2.20}$$

and

$$\|\Lambda(\cdot)\|_2 = |c| \prod_{i=1}^{\infty} 2^{-n/2}\, \|\lambda_i\|. \tag{2.21}$$

3. MULTIRESOLUTION ANALYSIS IN $\mathscr{C}$

The type of multiresolution analysis considered in this paper differs from the classical one mainly by the type of operators used.

3.1. *Rademacher Set of Operators*

Let r_j; $j = 0, 1, 2, \ldots, k-1$ be transformations in the set of k dimensional vectors $a = (a_0, a_1, a_2, \ldots, a_{k-1})$ of the form

$$r_j(a) = (\varepsilon_{j,0} a_j, \varepsilon_{j,1} a_{j \dotplus 1}, \varepsilon_{j,2} a_{j \dotplus 2}, \ldots, \varepsilon_{j,k-1} a_{j \dotplus (k-1)}), \tag{3.22}$$

where $\varepsilon_{j,l} = 1, -1$; $j, l = 0, 1, 2, \ldots, k-1$ and $\varepsilon_{0,l} = \varepsilon_{l,0} = 1$; $l = 0, 1, 2, \ldots, k-1$.

Call the transformations (3.22) Rademacher transformations and the matrix $\{\varepsilon_{j,l}\}$ a Rademacher ε-matrix if

$$\langle r_j(a), r_l(a)\rangle = \sum_{i=0}^{k-1} \varepsilon_{j,i} a_{j \dotplus i} \varepsilon_{l,i} a_{l \dotplus i} = 0 \qquad \text{for} \quad j \neq l, \tag{3.23}$$

$j, l = 0, 1, 2, ..., k-1$, for an arbitrary vector a.

There are Rademacher transformations only for $k = 2, 2^2, 2^3$, corresponding to the complex numbers, the quaternions, and the octets of Kelly, respectively. Three Rademacher ε-matrices are

$$\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{vmatrix}, \tag{3.24}$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{vmatrix}. \tag{3.25}$$

Definition 3.1. For a fixed natural number s, a set of 2^n operators $r_{s,j}\colon \mathscr{C} \to \mathscr{C}$ is called Rademacher set of operators of rank n if

$$r_{s,j}(f; \mathbf{x}) = \varepsilon_{j,l} f(\mathbf{x} + j\mathbf{2}^{-ns}), \qquad \text{for} \quad \mathbf{x} \in \mathbf{G}_{ns}^{(l)}; \qquad j, l = 0, 1, 2, ..., 2^n - 1,$$

where $\{\varepsilon_{j,l}\}$ is a Rademacher ε-matrix

The operator $r_{s,0}$ is the identity for $s = 1, 2, 3, \ldots$.

It is easy to see that every Rademacher operator preserves the scalar product, or for every two functions $f, g \in \mathscr{C}$,

$$\langle r_{s,j}(f; \cdot), r_{s,j}(g; \cdot)\rangle = \langle f, g\rangle. \tag{3.26}$$

Lemma 3.1. *Let $f \in \mathscr{C}$, s be a natural number and $r_{s,j}$ be a Rademacher operator of rank $n = 1, 2, 3$. Then*

$$\langle r_{s,j}(f; \cdot), f\rangle = 0 \qquad \textit{for} \quad j = 1, 2, ..., 2^n - 1.$$

The proof follows directly from the definition of a Rademacher transformation.

3.2. *Main Theorem*

Let $\Lambda(\mathbf{x})$ be a dyadic exponential function of rank n with the sequence $\{\lambda_i\}_{i=1}^{\infty}$ and let

$$w_0(\mathbf{x}) = \|\Lambda(\cdot)\|_2^{-1}\,\Lambda(\mathbf{x}); \qquad \Lambda(\mathbf{x}) = \prod_{i=0}^{\infty} \lambda_{i+1}(x_{n,i}), \tag{3.27}$$

where

$$\|\Lambda(\cdot)\|_2 = \prod_{i=1}^{\infty} 2^{-n/2}\,\|\lambda_i\| \tag{3.28}$$

is the L_2 norm of the dyadic exponential function Λ; see (2.21).

THEOREM 3.1. *Let $\{V_s\}_{s=0}^{\infty}$ be a sequence of subspaces of the functional space $\mathscr{C}$, where V_0 is the span over the dyadic exponential function Λ, described above, and*

$$f(\mathbf{x}) \in V_s \Rightarrow r_{s+1,j}(f;\mathbf{x}) \in V_{s+1}; \qquad s=0,1,..., \quad j=0,1,...,2^n-1, \tag{3.29}$$

where, for every natural s, the operator $r_{s,j}$, $j=0,1,2,...,2^n-1$ form in Rademacher set of operators of rank n (Definition 3.1).

Then

$$V_s \subset V_{s+1}; \qquad j=0,1,2,...,$$

$$\bigcup_{s=0}^{+\infty} V_s \qquad \text{is dense in } \mathscr{C} \qquad (V_0 \text{ is one dimensional}).$$

Define the sequence $\{w_i\}_{i=0}^{\infty}$ by

$$w_{j2^{ns-n}+p}(\mathbf{x}) = r_{s,j}(w_p;\mathbf{x}); \qquad j=1,2,...,2^n-1, \tag{3.30}$$

$p=0,1,...,2^{ns-n}-1$, $s=1,2,3,...$ and w_0 is defined by (3.27).

Then $\{w_i\}_{i=0}^{2^{ns}-1}$ is an orthonormal basis in V_s.

Proof. The relation $V_s \subset V_{s+1}$ follows from (3.29), as the operator $r_{s,0}$ is the identity for every natural s. The orthonormality of (3.30) directly follows from Lemma 3.1 and (3.26).

To prove that $\bigcup_{s=0}^{\infty} V_s$ is uniformly dense in $\mathscr{C}$, consider the projection $\Lambda_s\colon \mathscr{C} \to V_s$:

$$\Lambda_s(f;\mathbf{x}) = \sum_{k=0}^{2^{ns}-1} c_k(f)\,w_k(\mathbf{x}), \qquad \text{where} \quad c_k(f) = \int_{\mathbf{G}} f(\mathbf{x})\,w_k(\mathbf{x})\,d\mathbf{x}.$$

The operator Λ_s is represented in the form

$$\Lambda_s(f; \mathbf{x}) = \int_{\mathbf{G}} f(\mathbf{t})\, K_s(\mathbf{x}, \mathbf{t})\, d\mathbf{t}; \qquad K_s(\mathbf{x}, \mathbf{t}) = \sum_{k=0}^{2^{ns}-1} w_k(\mathbf{x})\, w_k(\mathbf{t}),$$

where

$$K_s(\mathbf{x}, \mathbf{t}) = \begin{cases} 2^{ns} \prod_{i=s}^{\infty} 2^n \, \|\lambda_{i+1}\|_2^{-2}\, \lambda_{i+1}(x_{n,i})\, \lambda_{i+1}(t_{n,i}); & \beta(\mathbf{x}, \mathbf{t}) \leqslant 2^{-ns}, \\ 0; & \beta(\mathbf{x}, \mathbf{t}) > 2^{-ns}. \end{cases} \tag{3.31}$$

To prove (3.31), use the representation of w_k, which follows from (3.30):

$$w_k(\mathbf{x}) = \|\Lambda(\cdot)\|_2^{-1} \prod_{i=0}^{\infty} \varepsilon_{k_{n,i}, x_{n,i}} \lambda_{i+1}(x_{n,i} \dotplus k_{n,i}).$$

Then

$$\begin{aligned} &\|\Lambda(\cdot)\|_2^2\, K_s(\mathbf{x}, \mathbf{t}) \\ &\quad = \sum_{k=0}^{2^{ns}-1} \prod_{i=0}^{\infty} \varepsilon_{k_{n,i}, x_{n,i}} \lambda_{i+1}(x_{n,i} \dotplus k_{n,i})\, \varepsilon_{k_{n,i}, t_{n,i}} \lambda_{i+1}(t_{n,i} \dotplus k_{n,i}) \\ &\quad = \prod_{i=s}^{\infty} \lambda_{i+1}(x_{n,i})\, \lambda_{i+1}(t_{n,i}) \sum_{k=0}^{2^{ns}-1} \prod_{i=0}^{s-1} \varepsilon_{k_{n,i}, x_{n,i}} \lambda_{i+1}(x_{n,i} \dotplus k_{n,i}) \\ &\qquad \times \varepsilon_{k_{n,i}, t_{n,i}} \lambda_{i+1}(t_{n,i} \dotplus k_{n,i}). \end{aligned} \tag{3.32}$$

On the other side

$$\begin{aligned} M_s &= \sum_{k=0}^{2^{ns}-1} \prod_{i=0}^{s-1} \varepsilon_{k_{n,i}, x_{n,i}} \lambda_{i+1}(x_{n,i} \dotplus k_{n,i})\, \varepsilon_{k_{n,i}, t_{n,i}} \lambda_{i+1}(t_{n,i} \dotplus k_{n,i}) \\ &= \sum_{k=0}^{2^{ns-n}} \prod_{i=0}^{s-2} \varepsilon_{k_{n,i}, x_{n,i}} \lambda_{i+1}(x_{n,i} \dotplus k_{n,i})\, \varepsilon_{k_{n,i}, t_{n,i}} \lambda_{i+1}(t_{n,i} \dotplus k_{n,i}) \\ &\quad \times \sum_{j=0}^{2^n-1} \varepsilon_{j, x_{n,s-1}} \lambda_s(x_{n,s-1} \dotplus j)\, \varepsilon_{j, t_{n,s-1}} \lambda_s(t_{n,s-1} \dotplus j) \\ &= M_{s-1} \sum_{j=0}^{2^n-1} \varepsilon_{j, x_{n,s-1}} \lambda_s(x_{n,s-1} \dotplus j)\, \varepsilon_{j, t_{n,s-1}} \lambda_s(t_{n,s-1} \dotplus j) \end{aligned}$$

or

$$M_s = \begin{cases} \prod_{i=1}^{s} \|\lambda_i\|^2 & \text{for} \quad \beta(\mathbf{x}, \mathbf{t}) \leqslant 2^{-ns}, \\ 0 & \text{for} \quad \beta(\mathbf{x}, \mathbf{t}) > 2^{-ns} \end{cases} \tag{3.33}$$

as

$$\sum_{j=0}^{2^n-1} \varepsilon_{j,\, x_{n,i}} \lambda_{i+1}(x_{n,i} \dot{+} j)\, \varepsilon_{j,\, t_{n,i}} \lambda_{i+1}(t_{n,i} \dot{+} j) = \begin{cases} \|\lambda_i\|^2 & \text{for} \quad x_{n,i} = t_{n,i}, \\ 0 & \text{for} \quad x_{n,i} \neq t_{n,i}. \end{cases}$$

Then (3.31) follows from (3.32), (3.28), and (3.33).

From (3.31) calculate

$$\int_{\mathbf{G}} K_s(\mathbf{x}, \mathbf{t})\, d\mathbf{t} = \prod_{i=s+1}^{\infty} \frac{2^{n/2} \lambda_i(x_{n,i-1})}{\|\lambda_i\|_2};$$

hence, according to (2.18),

$$\left\| 1 - \int_{\mathbf{G}} K(\mathbf{x}, \mathbf{t})\, d\mathbf{t} \right\| \leqslant \epsilon(\Lambda; 2^{-s}).$$

Finally for every continuous function f, for $s > N(\Lambda)$,

$$\begin{aligned} |f(\mathbf{x}) - \Lambda_s(f; \mathbf{x})| &\leqslant \left| f(\mathbf{x}) - f(\mathbf{x}) \int_{\mathbf{G}} K_s(\mathbf{x}, \mathbf{t})\, d\mathbf{t} \right| + \int_{\mathbf{G}} |f(\mathbf{x}) - f(\mathbf{t})|\, K_s(\mathbf{x}, \mathbf{t})\, d\mathbf{t} \\ &\leqslant \|f\| \left| 1 - \int_{\mathbf{G}} K_s(\mathbf{x}, \mathbf{t})\, d\mathbf{t} \right| + \omega(f; 2^{-s})(1 + \epsilon(f; 2^{-s})), \end{aligned}$$

or

$$|f(\mathbf{x}) - \Lambda_s(f; \mathbf{x})| \leqslant \omega(f; 2^{-s}) + (\|f\| + \omega(f; 2^{-s}))\, \epsilon(f; 2^{-s}). \tag{3.34}$$

This completes the proof, according to (2.12) and (2.17).

For the function $\Lambda(\mathbf{x}) = 1$, according to (2.17), the estimate (3.34) has the form

$$|f(\mathbf{x}) - \Lambda_s(f; \mathbf{x})| \leqslant \omega(f; 2^{-s}). \tag{3.35}$$

The comparison of the inequalities (3.34) and (3.35) shows that for the entire class $\mathscr{C}$ of the continuous dyadic functions, the choice of the constant 1 as a first function is the best one. It is useful to use another first function if this first function is especially adapted to the function f, which has to be approximated.

4. ADAPTATION OF A MULTIRESOLUTION ANALYSIS

In this section we consider finite dimensional subspaces of $\mathscr{C}$. Let s be a natural number, $n = 1, 2, 3$, and $\mathscr{P}_{ns}$ be the set of the pixel functions of rank ns. Theorem 3.1 provides the possibility to construct orthonormal bases in $\mathscr{P}_{ns}$, which depend on $(2^n - 1)\, s$ free parameters. These parameters may be used to adapt a basis to a given function $f \in \mathscr{C}$. Note that the number of the free parameters has the order of $\ln_2 N$, where N is the number of the pixels.

4.1. *Entropy Criterion*

Let $\{\phi_i\}_{i=0}^{2^{ns}-1}$ be an orthonormal basis in $\mathscr{P}_{ns}$, and

$$c_i(\phi; f) = \int_{\mathbf{G}} f(\mathbf{x})\, \phi_i(\mathbf{x})\, d\mathbf{x}, \qquad i = 0, 1, 2, ..., 2^{ns} - 1$$

be the Fourier coefficients of the function $f \in \mathscr{P}_{ns}$.

The entropy criterion, for the adaptation of an orthonormal basis $\{\phi_i\}$ to a given function f with $\|f\|_2 = 1$, is to minimize the value of the entropy

$$\epsilon^2(f, \{\phi_i\}) = -\sum_{i=0}^{2^{ns}-1} |c_i(\phi; f)|^2 \ln |c_i(\phi; f)|^2. \tag{4.36}$$

This criterion is the basis for adaptation of Malvar wavelets and wavelet packets [7].

To calculate the parameters of the first function ϕ_0, for a given function f ($\|f\|_2 = 1$), in such a way that the entropy (4.36) is minimal, the function may be solved for small values of s. We do not know practical methods to solve this problem for values of s, which are interesting for the applications.

4.2. *First Coefficient* (*FC*) *Criterion*

A criterion for adaptation may be to maximize the first coefficient $|c_0(\phi; f)|$, that is, to maximize the energy taken by the first Fourier coefficient from the expansion of f in this basis.

For simplicity, we consider this adaptation according to FC criterion only for first functions, which are dyadic exponential functions of rank 1 ($n = 1$). In the definition of the dyadic exponential function of rank 1 let $\lambda_k(1) = \xi_k$. The optimization problem we have in hand is to find the maximum of the function

$$F(\xi_1, \xi_2, ..., \xi_s) = \left(\prod_{i=1}^{s} \frac{1 + \xi_i^2}{2} \right)^{-1/2} \int_{\mathbf{G}} f(\mathbf{x}) \prod_{i=1}^{s} \xi_i^{x_i}\, d\mathbf{x}. \tag{4.37}$$

The necessary conditions for an extremum of (4.37),

$$\frac{\partial F}{\partial \xi_1} = \frac{\partial F}{\partial \xi_2} = \cdots = \frac{\partial F}{\partial \xi_s} = 0,$$

are equivalent to the system

$$\xi_k = \int_{\mathbf{G}_k^{(1)}} f(\mathbf{x})\, \xi_k^{-1} \prod_{i=1}^{s} \xi_i^{x_i}\, d\mathbf{x} \Big/ \int_{\mathbf{G}_k^{(0)}} f(\mathbf{x}) \prod_{i=1}^{s} \xi_i^{x_i}\, d\,\mathbf{x}, \qquad k = 1, 2, 3, ..., s. \tag{4.38}$$

For $s = 3$, the equations (4) are

$$\xi_1 = \frac{f_4 + f_5\xi_3 + f_6\xi_2 + f_7\xi_2\xi_3}{f_0 + f_1\xi_3 + f_2\xi_2 + f_3\xi_2\xi_3},$$

$$\xi_2 = \frac{f_2 + f_3\xi_3 + f_6\xi_1 + f_7\xi_1\xi_3}{f_0 + f_1\xi_3 + f_4\xi_1 + f_5\xi_2\xi_3},$$

$$\xi_3 = \frac{f_1 + f_3\xi_2 + f_5\xi_1 + f_7\xi_1\xi_2}{f_0 + f_2\xi_2 + f_4\xi_1 + f_3\xi_1\xi_2},$$

where

$$f_k = \int_{[3;\,k]} f(\mathbf{x})\, d\mathbf{x}, \qquad k = 0, 1, 2, ..., 7.$$

The equations (4.38) may be solved directly by iteration, as in the right hand side of every equation from (4.38), the unknown of the left does not appear. Of course, the iteration, starting with $\xi_k^{(0)} = 1$, $k = 1, 2, 3, ..., s$, is not always convergent and we do not know any specific conditions for this convergence. The Newton method for solving (4.38) works very well.

EXAMPLE 1. To illustrate the adaptation by FC criterion, we consider a simple Iterated Function System (IFS) [2]:

$$\tilde{g}(x) = \begin{cases} 0.7\tilde{g}(2x) + 50 & \text{for} \quad x \in [0, 1/2) \\ 0.7\tilde{g}(2x-1) + 1200 & \text{for} \quad x \in [1/2, 1]. \end{cases} \tag{4.39}$$

The completed graph of the function $\tilde{g}$ is depicted in the bottom of Fig. 2.

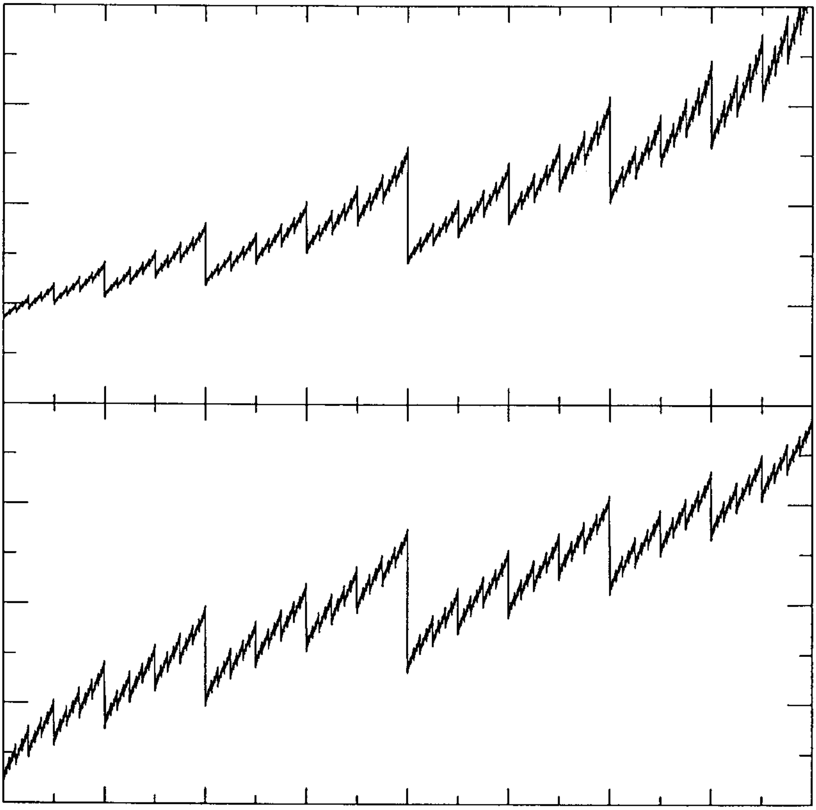

FIG. 2. Bottom: The completed graphs of the function $\tilde{g}$ defined by (4.39). Top: The dyadic exponential function of rank 1, which maximize the first coefficient of $\tilde{g}$.

The completed graph of the dyadic exponential function Λ of rank 1, which maximizes the first coefficient of the function $\tilde{g}$, is depicted in Fig. 2 above the function $\tilde{g}$. The sequence of Λ is shown in Table II.

Comparing the graphs of two functions in Fig. 2, we see the similarity between a fractal function and a dyadic exponential function. Both functions have self-similarity.

TABLE II

The sequence of Λ

i	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	1.70	1.42	1.27	1.18	1.12	1.09	1.06	1.04	1.03	1.02	1.01	1.01

It is interesting to compare the entropies of the pixel representation of $\tilde{g}$ (2^{12} pixels), and the representations of $\tilde{g}$ by the same number of Fourier coefficients for three orthonormal systems: Haar, Walsh, and the orthonormal system generated by the first function, which maximizes the first coefficient. The values of these entropies are shown here:

Method	Entropy
Pixels	11.657
Haar	0.937
Walsh	0.811
Adapted	0.066

The influence of the adaptation, according to FC criterion, on the minimization of the entropy (4.36) is obvious.

4.3. *Backward Minimization* (*BWM*) *Criterion*

Let $W_0 = V_0$ and W_i; $i = 1, 2, 3, \ldots, s$ be the orthogonal compliment of V_{i-1} to V_i, or $V_i = V_{i-1} \oplus W_i$. Then $\mathscr{P}_{ns}$ is an orthogonal direct sum of W_i,

$$\mathscr{P}_{ns} = \oplus \sum_{i=0}^{s} W_i. \tag{4.40}$$

Let $f \in \mathscr{P}_{ns}$, $\|f\|_2 = 1$, f_i be the projection of f on W_i and $V_0 = W_0$ be the span over the dyadic exponential function $\Lambda(\mathbf{x})$ with a sequence $\{\lambda_i(u)\}_{i=0}^{\infty}$, where $\lambda_i(u) = 1$ for $u = 0, 1, 2, \ldots, 2^n - 1$, $i = s+1, s+2, \ldots$. Then $V_s = \mathscr{P}_{ns} = \oplus \sum_{i=0}^{s} W_i$ and

$$\|f\|_2^2 = \|f_0\|_2^2 + \|f_1\|_2^2 + \|f_2\|_2^2 + \cdots + \|f_s\|_2^2. \tag{4.41}$$

R. Coifman and V. Wickerhauser [3] introduce the entropy

$$\epsilon^2(f, \{W_i\}) = -\sum_{i=0}^{s} \|f_i\|_2^2 \ln \|f_i\|_2^2 = F(f; \lambda_1, \lambda_2, \ldots, \lambda_s)$$

as a measure of distance between f and the orthogonal decomposition (4.40).

It is possible, for small values of s, to find the values of these parameters for which $\epsilon^2(f, \{W_i\})$ is minimum. This explicitly solves the problem for adaptation of the decomposition (4.40) to the function f with respect to the measure of R. Coifman and V. Wickerhauser [3]. This problem is numerically difficult to solve for large values of s, which are practically interesting.

We shall modify the criterion for optimization and adaptation of the decomposition (4.40) to a given function f as follows.

DEFINITION 4.1. First make $\|f_s\|_2^2$ as small as possible. Second make $\|f_{s-1}\|_2^2$ as small as possible. Continue so until $\|f_1\|_2^2$ is made as small as possible.

The motivation of the BWM criterion is that $\|f_s\|_2^2$ is equal to the sum of the squares of not less than half of the Fourier coefficients of the function f. Then $\|f_{s-1}\|_2^2$ comes, which is equal to the sum of the squares of not less than one half of the remaining Fourier coefficients of the function f, and so on. It is natural to make first $\|f_s\|_2^2$ as small as possible as it contributes with the biggest number of coefficients to the entropy of the Fourier coefficients of f. Having done this, the remaining member of the sum with most of the Fourier coefficients of f is $\|f_{s-1}\|_2^2$, and so on.

The practical implementation of the BWM criterion is based on the following statement.

THEOREM 4.1. *Let f be a fixed function in $\mathscr{C}$, $\Lambda(\mathbf{x}) = \prod_{i=0}^{\infty} \lambda_{i+1}(x_{n,i})$ be a dyadic exponential function of rank n, and f_k be the projection of f on W_k (see Theorem 3.1). Then $\|f_k\|_2^2$ is minimal if $\lambda_k = (\lambda_k(0), ..., \lambda_k(2^n-1))$ is the eigenvector of a symmetric quadratic form*

$$\sum_{l=0}^{2^n-1} \sum_{m=0}^{2^n-1} a_{l,m} \xi_l \xi_m,$$

corresponding to the largest eigenvalue of this quadratic form, where $a_{l,m}$ depends only on f and on λ_i for $i > k$.

Proof. Let

$$\hat{K}_i(\mathbf{x}, \mathbf{t}) = K_i(\mathbf{x}, \mathbf{t}) - K_{i-1}(\mathbf{x}, \mathbf{t}), \qquad \hat{K}_0(\mathbf{x}, \mathbf{t}) = K_0(\mathbf{x}, \mathbf{t}).$$

From (3.31) we have

$$\hat{K}_j(\mathbf{x}, \mathbf{t}) = \begin{cases} \left(1 - \dfrac{\lambda_j(x_{n,j-1})\, \lambda_j(t_{n,j-1})}{\|\lambda_j\|_2^2}\right) \alpha_j(\mathbf{x})\, \alpha_j(\mathbf{t}) \\ \qquad \text{for} \quad \beta(\mathbf{x}, \mathbf{t}) \leqslant 2^{-nj}, \\ -\dfrac{\lambda_j(x_{n,j-1})\, \lambda_j(t_{n,j-1})}{\|\lambda_j\|_2^2}\, \alpha_j(\mathbf{x})\, \alpha_j(\mathbf{t}) \\ \qquad \text{for} \quad 2^{-nj} < \beta(\mathbf{x}, \mathbf{t}) \leqslant 2^{-nj+n}, \\ 0 \qquad \text{for} \quad \beta(\mathbf{x}, \mathbf{t}) > 2^{-nj+n}, \end{cases} \tag{4.42}$$

where

$$\alpha_j(\mathbf{x}) = 2^{nj/2} \prod_{i=j+1}^{s} 2^{n/2} \, \|\lambda_i\|_2^{-1} \, \lambda_i(x_{n,\, i-1}),$$

$$\int_{[nj;\,0]} \alpha_j^2(\mathbf{x}) \, d\mathbf{x} = 1, \quad \alpha_s(\mathbf{x}) = 1, \tag{4.43}$$

and

$$f_j(\mathbf{x}) = \int_{\mathbf{G}} f(\mathbf{t}) \, \hat{K}_j(\mathbf{x}, \mathbf{t}) \, d\mathbf{t}.$$

From (4.42) we see that $\hat{K}_j(\mathbf{x}, \mathbf{t})$, and hence $\|f_j\|_2$ depends only on λ_j, $\lambda_{j+1}, \lambda_{j+2}, ..., \lambda_s$.

From (4.42) for $\mathbf{x} \in [nj; 2^n p + l]$; $p = 0, 1, 2, ..., 2^{nj-n} - 1$, $l = 0, 1, 2, ..., 2^n - 1$ there follows

$$\begin{aligned} f_j(\mathbf{x}) \sum_{m=0}^{2^n-1} \int_{[nj;\, 2^n p + m]} & f(\mathbf{t}) \, \hat{K}_j(\mathbf{x}, \mathbf{t}) \, d\mathbf{t} \\ &= \alpha_j(\mathbf{x}) \left(f_{p,l}^{n,j} - \|\lambda_j\|^{-2} \lambda_j(l) \sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \right), \end{aligned} \tag{4.44}$$

where

$$f_{p,l}^{n,j} = \int_{[nj;\, 2^n p + l]} f(\mathbf{t}) \, \alpha_j(\mathbf{t}) \, d\mathbf{t}. \tag{4.45}$$

Then

$$\begin{aligned} \|f_j\|_2^2 &= \int_{[0;\,0]} f_j(\mathbf{x})^2 \, d\mathbf{x} = \sum_{p=0}^{2^{nj-n}} \sum_{l=0}^{2^n-1} \int_{[nj;\, 2^n p + l]} f_j(\mathbf{x})^2 \, d\mathbf{x} \\ &= \sum_{p=0}^{2^{nj-n}} \sum_{l=0}^{2^n-1} \left(f_{p,l}^{n,j} - \|\lambda_j\|_2^{-2} \lambda_j(l) \sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \right)^2 \\ &= \|\lambda_j\|^{-4} \sum_{p=0}^{2^{nj-n}} \sum_{l=0}^{2^n-1} \Bigg(\|\lambda_j\|_2^4 \, (f_{p,l}^{n,j})^2 - 2 \, \|\lambda_j\|_2^2 \, \lambda_j(l) \, f_{p,l}^{n,j} \sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \\ &\qquad + \lambda_j(l)^2 \left(\sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \right)^2 \Bigg) \\ &= \|\lambda_j\|^{-2} \sum_{p=0}^{2^{nj-n}} \left(\|\lambda_j\|^2 \sum_{l=0}^{2^n-1} (f_{p,l}^{n,j})^2 - \left(\sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \right)^2 \right) \\ &= \sum_{p=0}^{2^{nj-n}} \sum_{l=0}^{2^n-1} (f_{p,l}^{n,j})^2 - \|\lambda_j\|^{-2} \sum_{p=0}^{2^{nj-n}-1} \left(\sum_{m=0}^{2^n-1} \lambda_j(m) \, f_{p,m}^{n,j} \right)^2, \end{aligned}$$

or to minimize $\|f_j\|_2^2$ it is necessary to maximize the symmetric quadratic form

$$\sum_{l,\, m=0}^{2^n-1} a_{l,\, m} \xi_l \xi_m \qquad \text{on the unit sphere } \sum_{l=0}^{2^n-1} \xi_l^2 = 1, \tag{4.46}$$

where

$$\xi_l = \lambda_j(l)/\|\lambda_j\|_2, \qquad a_{l,\, m} = a_{m,\, l} = \sum_{p=0}^{2^{nj-n}-1} f_{p,\, l}^{n,\, j} f_{p,\, m}^{n,\, j}. \tag{4.47}$$

It is well known that the maximum of the symmetric quadratic form (4.46) on the unit sphere is obtained if $\xi = \xi^* = (\xi_0^*, \xi_1^*, ..., \xi_{2^n-1}^*)$ is the eigenvector of the quadratic form (4.46) corresponding to the largest eigenvalue of this form.

Finaly, $\|f_j\|_2^2$ will be minimal if $\lambda_j = \lambda_j^* = (1, \xi_1^*/\xi_0^*, \xi_2^*/\xi_0^*, ..., \xi_{2^n-1}^*/\xi_0^*)$. This completes the proof as from (4.43), (4.45), and (4.47) is seen that $a_{l,\, m}$ depends only on f and on λ_i for $i > k$.

To adapt a dyadic exponential function Λ, to a given function f, according to the BWM criterion, we calculate first the values $\lambda_s(j)$, $j = 1, 2, 3, ..., 2^n - 1$, according to Theorem 4.1. For this purpose we find the eigenvector, corresponding to the biggest eigenvalue of a quadratic form with coefficients depending only on f. Then we calculate, according to Theorem 4.1, the values $\lambda_{s-1}(j)$, $j = 1, 2, 3, ..., 2^n - 1$. (For this we find the eigenvector, corresponding to the biggest eigenvalue of a quadratic form with coefficients depending only on f and the already calculated values $\lambda_s(j)$, $j = 1, 2, 3, ..., 2^n - 1$.) Next, in the same way we calculate the values $\lambda_k(j)$, $j = 1, 2, 3, ..., 2^n - 1$ for $k = s-2, s-3, ..., 1$.

Example 2. To illustrate the BWM method, we consider the function h in the bottom of Fig. 3, which represents the Solar flux (in 2800 MHz) for a period of 4096 days, represented with 2^{12} pixels. In the top of Fig. 3 is the dyadic exponential function Λ of rank 3, adapted according to the BWM method. The sequence of Λ is shown here:

i	$\lambda_1(i)$	$\lambda_2(i)$	$\lambda_3(i)$	$\lambda_4(i)$
0	1.000	1.000	1.000	1.000
1	0.711	0.950	1.009	0.999
2	0.711	0.978	0.977	0.998
3	1.471	1.007	0.985	0.999
4	2.246	1.023	1.011	1.001
5	2.172	1.040	1.014	1.000
6	1.531	1.027	0.987	0.999
7	0.960	1.021	0.955	1.003

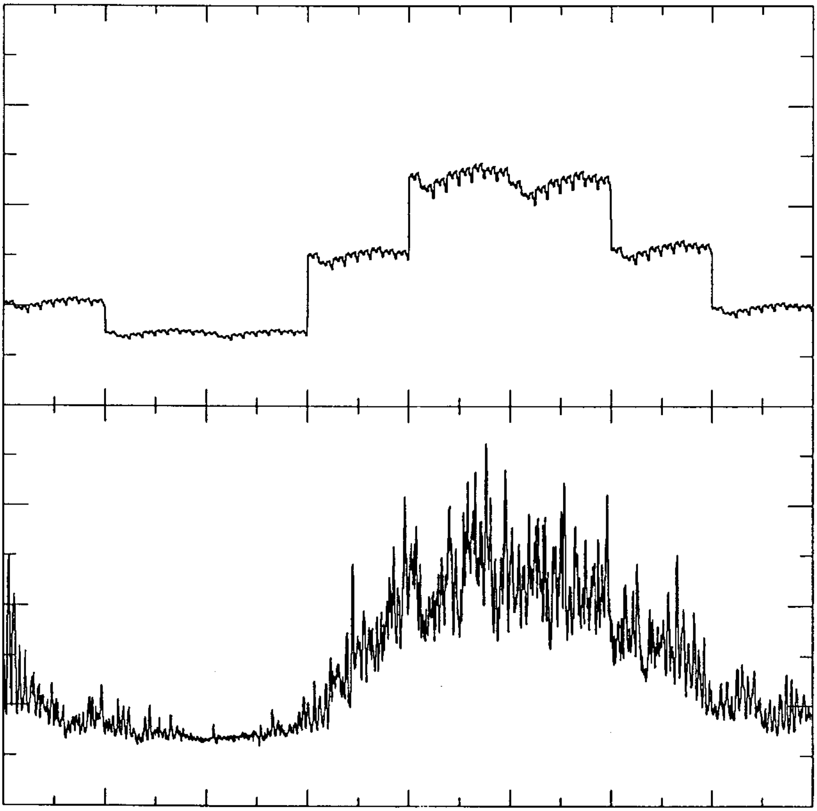

FIG. 3. Bottom: Data from Solar flux. Top: Dyadic exponential function of rank 3, optimized for these data according to the BWM criterion.

The respective entropies, considered in the Example 1, are shown here:

Method	Entropy
Pixels	11.465
Haar	1.349
Walsh	1.398
Adapted	0.497

Remarks. 1. There exist algorithms of complexity $O(N\,ln_2\,N)$, for N pixels, to adapt the dyadic exponential function to a given function f in respect to the BWM criterion.

2. The functions λ_i, forming the dyadic exponential function in Theorem 4.1, may be of different rank. Then the Rademacher set of operators from V_i to V_{i+1} has to be chosen accordingly.

3. The orthonormal system (3.30) is of Walsh type (functions with full support). It is possible to construct orthonormal system of Haar type to replace (3.30) in Theorem 3.1.

REFERENCES

1. L. A. Bagget, A. Carey, W. Moran, and Ohring, General existence theorem for orthonormal wavelets: An abstract approach, *Publ. Res. Inst. Math. Sci.* **31** (1995), 95–112.
2. M. F. Barnsley and L. P. Hurd, “Fractal Image Compression,” Peters, Wellesley, MA, 1993.
3. R. R. Coifman, Adapted multiresolution analysis, computation, signal processing and operator theory, *in* “Proc. of the Int. Congr. of Math.,” pp. 879–887, Kyoto, Japan, 1990.
4. N. J. Fine, On the Walsh functions, *Trans. Amer. Math. Soc.* **65** (1949), 373–414.
5. W. C. Lang, Orthogonal wavelets on the Cantor dyadic group, *SIAM J. Math. Anal.* **27** (1996), 305–312.
6. S. G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, *Trans. Amer. Math. Soc.* **315** (1989), 69–87.
7. Y. Meyer, “Ondelettes et Opérateurs, I,” Herman, Paris, 1990.
8. L. S. Pontrjagin, “Topologische Gruppen,” Teubner Verlag, Leipzig, 1957.
9. Bl. Sendov, “Hausdorff Approximations,” Kluwer Academic, Dordrecht/Boston/London, 1990.
10. Bl. Sendov, Mathematical modeling of real-world images, *Constr. Approx.* **12** (1996), 31–65.
11. Bl. Sendov, Multiresolution analysis of functions defined on the dyadic topological group, *East J. Approx.* **3** (1997), 225–239.