Current Development in Theory and Applications of Wavelets **1** (2007) 273–292

MULTILEVEL DECOMPOSITION IN HILBERT SPACE

SOMDEB MAJUMDAR, NHAN LEVAN, AND CARLOS S. KUBRUSLY

ABSTRACT. By decomposing an element of a sequence — of Hilbert space bounded linear operators — into the sum of a lower level element and several higher level elements, one obtains a Multilevel Decomposition (MLD) of the element. Moreover, as we shall show, such a decomposition can result in a Multilevel Approximation (MLA) of vectors of the space. In particular, for the function space $\mathcal{L}^2(\mathbb{R})$, the MLD of elements of a sequence of orthogonal projections P_m results in the well known Multiresolution Approximation (MRA) of Wavelet Theory. Also for the sequence of elements $P_m D^m$, where D is the $\mathcal{L}^2(\mathbb{R})$-dyadic-scale operator, MLD also yields an approximation for functions of the space $\mathcal{L}^2(\mathbb{R})$. An interesting feature of MLD is that it leads to a new interpretation of the Dilation-by-s operator as a "time-varying" shift operator.

1. INTRODUCTION

We study application of what we define as a natural multilevel decomposition, of elements of sequences of Hilbert space bounded linear operators, to approximation of vectors in the space.

Let $\mathbb{Z}$ stand for the set of all integers, and $\mathcal{H}$ for a (nonzero, complex, infinite-dimensional) separable Hilbert space. By an operator on $\mathcal{H}$ we mean a bounded linear transformation of $\mathcal{H}$ into itself. Let $\{A_m\}_{m\in\mathbb{Z}}$ be a family of operators on $\mathcal{H}$. Each A_m, to be referred to as the *level-m* element, can be naturally decomposed as

$$A_m = A_{m-1} + (A_m - A_{m-1}), \quad m \in \mathbb{Z}, \tag{1.1}$$

which is simply the sum of the lower level element A_{m-1} and the "difference" element $A_m - A_{m-1}$. Repeating the decomposition with A_j on the right hand side instead of A_{m-1}, for any $j, m \in \mathbb{Z}$ such that $j < m$, we obtain

$$A_m = A_j + A_m - A_j = A_j + \sum_{k=j+1}^{m} A_k - \sum_{k=j}^{m-1} A_k = A_j + \sum_{k=j}^{m-1} A_{k+1} - \sum_{k=j}^{m-1} A_k,$$

which yields a *Multilevel Decomposition* (MLD) of A_m,

$$A_m = A_j + \sum_{k=j}^{m-1} E_k, \quad j < m, \quad j, m \in \mathbb{Z}, \tag{1.2}$$

where A_j is a lower level element and E_k, called *difference element at level-k* (or , à la Wavelet Theory, *fluctuation* or *detail* element), is defined as

$$E_k := A_{k+1} - A_k, \quad k \in [j,\ m-1]. \tag{1.3}$$

Date: November 27, 2006.
2000 *Mathematics Subject Classification.* 42C40, 47A15.
Keywords. Multilevel Decomposition, Multilevel Approximation, Multiresolution Approximation, Scaling Operators and Time-Varying Shifts on $\mathcal{L}^2(\mathbb{R})$.

We therefore conclude that any element A_m of the family $\{A_m\}$ of operators can be decomposed into the sum of any lower level element and the difference elements for all levels not greater than $m-1$. Let $\mathcal{R}(A_m)$ denote the range of A_m and suppose

$$\mathcal{R}(A_m) = \mathcal{R}(A_j) \oplus \mathcal{R}(E_j) \oplus \cdots \oplus \mathcal{R}(E_{m-1}) = \mathcal{R}(A_j) \oplus \bigoplus_{k=j}^{m-1} \mathcal{R}(E_k) \tag{1.4}$$

for each $m \in \mathbb{Z}$. Moreover, if the orthogonal projection of any $h \in \mathcal{H}$ onto $\mathcal{R}(A_m)$ can be computed in terms of those of h onto the subspaces on the right hand side, then one obtains an approximation of h at level-m, in terms of the approximation at level-j and the difference-approximations at levels $k \in [j,\ m-1]$. These begin to sound like Multiresolution Approximation (or Analysis) (MRA) of Wavelet Theory.

This paper is a "follow-up" of [8] in which we show that Multiresolution Approximation involves both scale subspaces as well as time-shift subspaces. Here we show that approximation of vectors of a separable Hilbert space $\mathcal{H}$ can be obtained from a natural multilevel decomposition of "suitable" sequences of operators defined on the space — instead of multilevel-decomposing the vectors themselves. We begin in Section 2 by considering a sequence of orthogonal projections $\{P_m\}_{m\in\mathbb{Z}}$ on $\mathcal{H}$. This, for the function space $\mathcal{L}^2(\mathbb{R})$, results in the well known Wavelet Theory Multiresolution Approximation. We then turn to the case of a sequence of powers of an operator, in particular, powers of the *dyadic-scale* operator D on $\mathcal{L}^2(\mathbb{R})$, defined for every function $f \in \mathcal{L}^2(\mathbb{R})$ by

$$(Df)(t) = \sqrt{2}\, f\big(2(t)\big) \tag{1.5}$$

(for almost all t in $\mathbb{R}$ with respect to Lebesgue measure). We then combine those two cases to obtain a new algorithm for computing approximation at any level-m of a $\mathcal{L}^2(\mathbb{R})$ function. We close the paper with an original interpretation of the dyadic-scale operator D, via multilevel decomposition of powers of a "scaling" scalar.

2. Multilevel Decomposition of Orthogonal Projections

Let $\{P_m\}_{m\in\mathbb{Z}}$ be a sequence of orthogonal projections on $\mathcal{H}$. According to (1.2)

$$P_m \;=\; P_j + \sum_{k=j}^{m-1} (P_{k+1} - P_k), \tag{2.1}$$

$$\;=\; P_j + \sum_{k=j}^{m-1} Q_k, \quad \text{for} \quad j < m, \quad m \in \mathbb{Z}, \tag{2.2}$$

where the difference element E_k of (1.2) is now re-denoted by

$$Q_k := P_{k+1} - P_k, \quad k \in [j,\ m-1].$$

It then follows that, for every $h \in \mathcal{H}$,

$$P_m h = P_j h + \sum_{k=j}^{m-1} (P_{k+1}h - P_k h), \quad j < m, \quad m \in \mathbb{Z}. \tag{2.3}$$

Thus the level-m approximation $P_m h$ can be "decomposed into", or "computed from", the sum of the approximation at level-j and the differences of the approximations at consecutive levels in the interval $[j,\ m-1]$. We note that $(P_{k+1} - P_k)h$

needs not be an approximation of h on the range space $\mathcal{R}(P_{k+1}-P_k)$ — unless $P_{k+1}-P_k$ is also an orthogonal projection. This important case is discussed below.

Now, as in the case of a $\mathcal{L}^2(\mathbb{R})$ MRA (see Definition 1 below), suppose

$$P_\infty := \lim_{m\to\infty} P_m = I, \tag{2.4}$$

$$P_{-\infty} := \lim_{m\to-\infty} P_m = O, \tag{2.5}$$

where I stands for the identity operator and O for the null operator, both on $\mathcal{H}$. It is worth noticing that the above limits are to be interpreted in the strong topology (and not in the uniform topology), which means that $\|(P_m - I)h\| \to 0$ and $\|(P_{-m} - O)h\| \to 0$ as $m \to \infty$ for every $h \in \mathcal{H}$. Thus, according to (2.2),

$$\sum_{k\in\mathbb{Z}}(P_{k+1}-P_k)h = \sum_{k\in\mathbb{Z}} Q_k h = \lim_{m\to\infty} P_m h - \lim_{j\to-\infty} P_j = Ih = h \tag{2.6}$$

for every $h \in \mathcal{H}$. The operators Q_k need not be orthogonal projections. However, it is easy to see that Q_k are orthogonal projections if and only if

$$\mathcal{V}_k \subseteq \mathcal{V}_{k+1}, \quad k \in \mathbb{Z}, \tag{2.7}$$

where $\mathcal{V}_k$ denotes the range $\mathcal{R}(P_k)$ of P_k, that is,

$$\mathcal{V}_k := \mathcal{R}(P_k), \quad k \in \mathbb{Z}.$$

It is worth noticing that each of the preceding equivalent assertions (viz. (i) Q_k are orthogonal projections and (ii) $\mathcal{V}_k \subseteq \mathcal{V}_{k+1}$, for every $k \in \mathbb{Z}$) are still equivalent to each of the following: (iii) $P_{k+1}P_k = P_k$, (iv) $P_k P_{k+1} = P_k$, or (v) $P_{k+1}P_kP_{k+1} = P_k$, for every $k \in \mathbb{Z}$ (cf. [6, Problem 2.9]). Moreover if (2.7) holds, then Q_k is actually the orthogonal projection onto the orthogonal complement of $\mathcal{V}_k$ in $\mathcal{V}_{k+1}$. That is, with $\ominus$ standing for orthogonal complement, if (2.7) holds, then Q_k is the orthogonal projection onto the subspace

$$\mathcal{W}_k := \mathcal{V}_{k+1} \ominus \mathcal{V}_k, \quad k \in \mathbb{Z} \tag{2.8}$$

(i.e., $\mathcal{R}(Q_k) = \mathcal{W}_k$ for each $k \in \mathbb{Z}$). We refer to $\mathcal{W}_k$ as the *difference subspace* at level-k. Observe that, according to (2.7) and (2.8), if $j, m \in \mathbb{Z}$ with $j < m$, then

$$\mathcal{V}_m = \mathcal{V}_j \oplus (\mathcal{V}_m \ominus \mathcal{V}_j) = \mathcal{V}_j \oplus \bigoplus_{k=j}^{m-1}(\mathcal{V}_{k+1} \ominus \mathcal{V}_k) = \mathcal{V}_j \oplus \bigoplus_{k=j}^{m-1} \mathcal{W}_k,$$

where $\oplus$ stands for orthogonal direct sum. More is true if (2.7) holds. Indeed, put

$$P_k' := I - P_k, \quad k \in \mathbb{Z}, \tag{2.9}$$

which is the orthogonal projection onto the subspace $\mathcal{V}_k^\perp = \mathcal{H} \ominus \mathcal{V}_k$ (i.e., P_k' is the complementary projection of P_k) so that

$$P_k' P_{k+1} = (I - P_k)\, P_{k+1} = P_{k+1} - P_k\, P_{k+1} \tag{2.10}$$

$$= P_{k+1} - P_k = Q_k \tag{2.11}$$

because of (2.7). Since $P_k h$ is called the level-k approximation of h, the vector

$$P_k' h = h - P_k h$$

can be regarded as the *error-vector* between h and its approximation at level-k. We call such vector a *k-error-vector*. Consequently, $\mathcal{V}_k^\perp$ is simply *the subspace of all k-error-vectors*. It then follows from (2.10) and (2.11) that $Q_k h$ is, on the one hand the difference between the two consecutive approximations of h at levels k and $k+1$ and, on the other hand, the approximation on the subspace $\mathcal{V}_k^\perp$ of the approximation $P_{k+1}h$. In MRA $Q_k h$ is simply referred to as the *detail* (*or fluctuation*) *at level-k* of h.

We summarize the above results in the next proposition.

Proposition 1. *If $\{P_m\}_{m\in\mathbb{Z}}$ is a family of orthogonal projections on a separable Hilbert space $\mathcal{H}$, then each P_m admits the multilevel decomposition*

$$P_m = P_j + \sum_{k=j}^{m-1}(P_{k+1} - P_k), \quad j < m, \quad m \in \mathbb{Z}.$$

Therefore, any $h \in \mathcal{H}$ admits the level-m approximation

$$P_m h = P_j h + \sum_{k=j}^{m-1}(P_{k+1}h - P_k h), \quad j < m, \quad m \in \mathbb{Z}.$$

If, in addition,

$$\lim_{m\to\infty} P_m = I \quad \text{and} \quad \lim_{m\to-\infty} P_m = O$$

strongly, then

$$\sum_{k\in\mathbb{Z}}(P_{k+1} - P_k) = I$$

strongly. Moreover, the difference operators

$$Q_k = P_{k+1} - P_k, \quad k \in \mathbb{Z},$$

are orthogonal projections if and only if the ranges $\mathcal{V}_k := \mathcal{R}(P_k)$ of P_k are nested,

$$\mathcal{V}_k \subseteq \mathcal{V}_{k+1}, \quad k \in \mathbb{Z},$$

and, in this case, Q_k is the orthogonal projection onto the difference subspace

$$\mathcal{W}_k = \mathcal{V}_{k+1} \ominus \mathcal{V}_k, \quad k \in \mathbb{Z},$$

and $\mathcal{V}_m$ admits the multilevel orthogonal decomposition

$$\begin{aligned}\mathcal{V}_m = \mathcal{R}(P_m) = P_m(\mathcal{H}) &= \Big(P_j + \sum_{k=j}^{m-1} Q_k\Big)\mathcal{H}, \\ &= \mathcal{V}_j \oplus \bigoplus_{k=j}^{m-1} \mathcal{W}_k, \quad j < m, \quad m \in \mathbb{Z}.\end{aligned}$$

This shows that the approximation at level-m of any vector $h \in \mathcal{H}$ can be computed from the approximation $P_j h$ and from the approximations $Q_k h$ on the difference subspaces $\mathcal{W}_k$:

$$P_m h = P_j h + \sum_{k=j}^{m-1} Q_k h, \quad h \in \mathcal{H} \tag{2.12}$$

which is a Multilevel Approximation (MLA) on $\mathcal{H}$. Note that computation of $P_m h$, in general, requires the knowledge of a basis of $\mathcal{V}_m$. However, the right hand sides of (2.12) offer an alternate effective method for computing $P_m h$.

To proceed, let us recall the Definition of an $\mathcal{L}^2(\mathbb{R})$ MRA [10, 11].

Definition 1. An $\mathcal{L}^2(\mathbb{R})$ MRA — with the *scaling function* $\phi \in \mathcal{L}^2(\mathbb{R})$ — is a sequence of closed subspaces $\{\mathcal{V}_m(\phi)\}_{m\in\mathbb{Z}}$, called *approximation subspaces*, satisfying the following properties:

(o) Generation of $\mathcal{V}_0(\phi)$: $\mathcal{V}_0(\phi)$ is spanned by the orthonormal set $\{\phi(\cdot\ - n)\}_{n\in\mathbb{Z}}$,

(i) Generation of $\mathcal{V}_m(\phi)$: $\quad \mathcal{V}_m(\phi) := D^m\mathcal{V}_0(\phi) = \bigvee_{n\in\mathbb{Z}} D^m\phi(\cdot\ - n)$ for $m \in \mathbb{Z}$,

(ii) $\mathcal{V}_m(\phi) \subset \mathcal{V}_{m+1}(\phi), \quad m \in \mathbb{Z}$,

(iii) $\overline{\bigcup}_{m\in\mathbb{Z}} \mathcal{V}_m(\phi) = \mathcal{L}^2(\mathbb{R})$,

(iv) $\bigcap_{m\in\mathbb{Z}} \mathcal{V}_m(\phi) = \{0\}$,

where D is the dyadic-scale operator defined in (1.5).

Remark 1. Consider the operator D on $\mathcal{L}^2(\mathbb{R})$ defined in (1.5), which is unitary (in fact, it is a bilateral shift of infinity multiplicity on $\mathcal{L}^2(\mathbb{R})$). Now consider another bilateral shift (thus unitary) on $\mathcal{L}^2(\mathbb{R})$, also of infinity multiplicity, which will be denoted by T and defined for every function $f \in \mathcal{L}^2(\mathbb{R})$ by

$$(Tf)(t) = f(t-1)$$

(for almost all t in $\mathbb{R}$ with respect to Lebesgue measure). Properties (o) and (i) of Definition 1 are written in terms of the bilateral shifts D and T as follows

$$\text{(o)} \qquad \{T^n\phi\}_{n\in\mathbb{Z}} \ \text{ is an orthonormal family and } \ \mathcal{V}_0(\phi) = \bigvee_{n\in\mathbb{Z}} T^n\phi,$$

$$\text{(i)} \qquad \mathcal{V}_m(\phi) = D^m\mathcal{V}_0(\phi) = \bigvee_{n\in\mathbb{Z}} D^mT^n\phi \ \text{ for each } \ m \in \mathbb{Z}.$$

Since $\{T^n\phi\}_{n\in\mathbb{Z}}$ is an orthonormal family that spans $\mathcal{V}_0(\phi)$, and since D is unitary, $\{D^mT^n\phi\}_{n\in\mathbb{Z}}$ is also an orthonormal family for each $m \in \mathbb{Z}$ (spanning each subspace $\mathcal{V}_m$), and so each $\mathcal{V}_m(\phi)$ in (i) can be identified with the orthogonal direct sum

$$\mathcal{V}_m(\phi) = \bigoplus_{n\in\mathbb{Z}} D^mT^n\phi.$$

However, $D^mT^n\phi$ is not orthogonal to $D^kT^n\phi$ if $k \neq m$, since $\mathcal{V}_m(\phi) \subset \mathcal{V}_{m+1}(\phi)$ as in Definition 1(ii), and therefore

$$\{D^mT^n\phi\}_{(m,n)\in\mathbb{Z}\times\mathbb{Z}} \ \text{ is a family of unitary vectors but not an orthogonal family,}$$

although, by using all items of Definition 1, we can infer that it spans $\mathcal{L}^2(\mathbb{R})$,

$$\bigvee_{m,n\in\mathbb{Z}\times\mathbb{Z}} D^mT^n\phi = \bigvee_{m\in\mathbb{Z}}\bigvee_{n\in\mathbb{Z}} D^mT^n\phi = \mathcal{L}^2(\mathbb{R}).$$

Now recall that a *wavelet* on $\mathcal{L}^2(\mathbb{R})$ (with respect to the bilateral shifts D and T) is precisely a function $\psi \in \mathcal{L}^2(\mathbb{R})$ such that $\{D^mT^n\psi\}_{(m,n)\in\mathbb{Z}\times\mathbb{Z}}$ is an orthonormal basis for $\mathcal{L}^2(\mathbb{R})$, and so it also spans $\mathcal{L}^2(\mathbb{R})$ but, unlike $\{D^mT^n\phi\}_{(m,n)\in\mathbb{Z}\times\mathbb{Z}}$,

$$\{D^mT^n\psi\}_{(m,n)\in\mathbb{Z}\times\mathbb{Z}} \ \text{ is an orthonormal family.}$$

However, a wavelet ψ can be constructed from a scaling function ϕ. Indeed, for each $m \in \mathbb{Z}$, consider the orthogonal differences (equivalently, the orthogonal complements) $\mathcal{V}_{m+1}(\phi) \ominus \mathcal{V}_m(\phi)$ — cf. Definition 1(ii). It happens that there exists a wavelet ψ such that, for each $m \in \mathbb{Z}$,

$$D^m\bigvee_{n\in\mathbb{Z}} T^n\psi = \bigvee_{n\in\mathbb{Z}} D^mT^n\psi = \mathcal{V}_{m+1}(\phi) \ominus \mathcal{V}_m(\phi) = D\mathcal{V}_m(\phi) \ominus \mathcal{V}_m(\phi);$$

in particular, for each scaling function ϕ there exists a wavelet ψ such that

$$\bigvee_{n\in\mathbb{Z}} T^n\psi = D\bigvee_{n\in\mathbb{Z}} T^n\phi \ominus \bigvee_{n\in\mathbb{Z}} T^n\phi.$$

In other words, given a scaling function ϕ as in Definition 1, there exists a (not necessarily unique) function ψ in $\mathcal{L}^2(\mathbb{R})$ that satisfies the above expressions and is a wavelet — i.e., such that $\{D^mT^n\psi\}_{(m,n)\in\mathbb{Z}\times\mathbb{Z}}$ is an orthonormal basis for $\mathcal{L}^2(\mathbb{R})$) — see e.g., [4, Proposition 2.13, p.57].

Let ψ be a wavelet constructed from the scaling function ϕ and define the *detail* orthogonal subspaces

$$\mathcal{W}_k(\psi) := \bigvee_{n\in\mathbb{Z}} D^k\psi\big(\,\cdot - n\big), \quad k\in\mathbb{Z}, \tag{2.13}$$

so that

$$\mathcal{V}_{m+1}(\phi) = \mathcal{V}_m(\phi)\oplus\mathcal{W}_m(\psi), \quad m\in\mathbb{Z}. \tag{2.14}$$

This means that the *finer* approximation subspace $\mathcal{V}_{m+1}(\phi)$ is obtained from the *coarser* approximation subspace $\mathcal{V}_m(\phi)$ by adding the detail subspace $\mathcal{W}_m(\psi)$ to it. The orthogonal projection P_m onto $\mathcal{V}_m(\phi)$ (cf. Proposition 1) now describes the Discrete Wavelet Transform (DWT) [5]

$$P_m\big(\mathcal{L}^2(\mathbb{R})\big) = \mathcal{V}_m(\phi) = \mathcal{V}_j(\phi)\oplus\bigoplus_{k=j}^{m-1}\mathcal{W}_k(\psi), \quad j<m, \quad m\in\mathbb{Z}. \tag{2.15}$$

The following result on families of orthogonal projections on $\mathcal{L}^2(\mathbb{R})$ follows from Proposition 1 and from the above discussion.

Proposition 2. *Let $\{P_m\}_{m\in\mathbb{Z}}$ be a sequence of $\mathcal{L}^2(\mathbb{R})$ orthogonal projections satisfying equations* (2.4) *and* (2.5)*. Moreover suppose that the difference operators*

$$Q_m = P_{m+1} - P_m, \quad m\in\mathbb{Z},$$

are also orthogonal projections over $\mathcal{L}^2(\mathbb{R})$. Then the subspaces

$$\mathcal{V}_m = \mathcal{R}(P_m), \quad m\in\mathbb{Z},$$

are nested and $\{\mathcal{V}_m\}_{m\in\mathbb{Z}}$ is an $\mathcal{L}^2(\mathbb{R})$ MRA if and only if

$$\mathcal{V}_0 = \bigvee_{n\in\mathbb{Z}}\phi(\cdot - n) = \mathcal{V}_0(\phi),$$

and

$$\mathcal{V}_m = D^m\mathcal{V}_0(\phi) = \mathcal{V}_m(\phi), \quad m\in\mathbb{Z},$$

for some scaling function $\phi\in\mathcal{L}^2(\mathbb{R})$ which is such that $\{\phi(\cdot - n)\}_{n\in\mathbb{Z}}$ is an orthonormal family. Consequently,

$$\mathcal{R}(Q_0) = \mathcal{W}_0(\psi) = \bigvee_{n\in\mathbb{Z}}\psi(\cdot - n),$$

and

$$\mathcal{R}(Q_m) = \mathcal{W}_m(\psi) = D^m\mathcal{W}_0(\psi), \quad m\in\mathbb{Z},$$

where ψ is a wavelet constructed from the scaling function ϕ.

Note that Proposition 2 explains why a MLA needs not be a MRA (cf. [1]).

3. Multilevel Decomposition of Powers of an Operator

We now turn to a class of sequences of powers of a (bounded linear) operator. Let A be an operator on $\mathcal{H}$. For any integers $0 \le j < m$, we get from (1.2) that

$$A^m = A^j + \sum_{k=j}^{m-1} (A^{k+1} - A^k), \tag{3.1}$$

which holds for every $j, m \in \mathbb{Z}$ such that $j < m$, whenever A is invertible.

Remark 2. Equation (3.1) leads to the following generalizations. For any pair $\{A, B\}$ of operators on $\mathcal{H}$, and any pair of integers $0 \le j < m$,

$$\begin{aligned} A^m - A^j B^{m-j} &= A^m - A^j B^{m-j} + \sum_{k=j+1}^{m-1} A^k B^{m-k} - \sum_{k=j+1}^{m-1} A^k B^{m-k} \\ &= \sum_{k=j+1}^{m} A^k B^{m-k} - \sum_{k=j}^{m-1} A^k B^{m-k} \\ &= A\Big(\sum_{k=j+1}^{m} A^{k-1} B^{m-k} \Big) - \Big(\sum_{k=j}^{m-1} A^k B^{m-k-1} \Big) B \\ &= A\Big(\sum_{k=j}^{m-1} A^k B^{m-k-1} \Big) - \Big(\sum_{k=j}^{m-1} A^k B^{m-k-1} \Big) B. \end{aligned}$$

Particular cases: If $B = I$, then we get (3.1):

$$A^m - A^j = (A - I) \sum_{k=j}^{m-1} A^k. \tag{3.2}$$

If $j = 0$, then

$$A^m - B^m = A\Big(\sum_{k=0}^{m-1} A^k B^{m-k-1} \Big) - \Big(\sum_{k=0}^{m-1} A^k B^{m-k-1} \Big) B.$$

From now on suppose A and B commute. In this case,

$$A^m - A^j B^{m-j} = (A - B) \sum_{k=j}^{m-1} A^k B^{m-k-1}$$

and, for $j = 0$,

$$A^m - B^m = (A - B) \sum_{k=0}^{m-1} A^k B^{m-k-1}$$

so that

$$\begin{aligned} (A^m - A^j) - (B^m - B^j) &= (A^m - B^m) - (A^j - B^j) \\ &= (A - B)\Big(\sum_{k=0}^{m-1} A^k B^{m-k-1} - \sum_{k=0}^{j-1} A^k B^{j-k-1} \Big) \\ &= (A - B)\Big(\sum_{k=0}^{j-1} A^k B^{j-k-1} (B^{m-j} - I) + \sum_{k=j}^{m-1} A^k B^{m-k-1} \Big). \end{aligned}$$

If, in addition, B is invertible, then

$$(3.3)\quad (A^m - A^j) - (B^m - B^j) = (A - B)\sum_{k=0}^{j-1} A^k (B^m - B^j) B^{-(k+1)} + (A^m - A^j B^{m-j}).$$

Observe that if A and B are invertible, then all the above expressions hold for every $j, m \in \mathbb{Z}$ such that $j < m$.

Now put $A = D$, the dyadic-scale operator on $\mathcal{L}^2(\mathbb{R})$, which is invertible (since it is unitary). In this case we get from (3.1) that

$$(3.4)\qquad D^m = D^j + \sum_{k=j}^{m-1} (D^{k+1} - D^k)$$

$$(3.5)\qquad = D^j + E_0 \sum_{k=j}^{m-1} D^k, \quad j < m, \quad m \in \mathbb{Z},$$

where E_0 is the difference element at level-0:

$$E_0 := D - D^0 = D - I.$$

Note that conditions (2.4) and (2.5) certainly do not apply to the sequence $\{D^m\}_{m\in\mathbb{Z}}$ since D is unitary.

If $\{P_m\}_{m\in\mathbb{Z}}$ is the $\mathcal{L}^2(\mathbb{R})$ family of orthogonal projections of a MRA, then we have the following important relationship between P_m and D^m:

$$(3.6)\qquad \mathcal{R}(P_m) = P_m\big(\mathcal{L}^2(\mathbb{R})\big) = \mathcal{V}_m(\phi) = D^m\big(\mathcal{V}_0(\phi)\big),$$

that is, the range space of P_m is the image of the level-0 approximation subspace $\mathcal{V}_0(\phi)$ under D^m. We have seen from (2.15) that $P_m\big(\mathcal{L}^2(\mathbb{R})\big)$ describes a DWT. We now show that this can also be obtained directly from (3.5). To see this, we begin with

$$(3.7)\qquad E_0\mathcal{V}_0(\phi) = (D - I)\mathcal{V}_0(\phi),$$

and, by Definition 1(i),

$$(3.8)\qquad D\mathcal{V}_0(\phi) = \mathcal{V}_1(\phi).$$

Moreover,

$$(3.9)\qquad \mathcal{V}_0(\phi) \perp \mathcal{W}_0(\psi) \quad \text{and} \quad \mathcal{V}_1(\phi) = \mathcal{V}_0(\phi) \oplus \mathcal{W}_0(\psi).$$

Thus (3.7) can be rewritten as

$$(3.10)\qquad E_0\mathcal{V}_0(\phi) = D\mathcal{V}_0(\phi) \ominus \mathcal{V}_0(\phi) = \mathcal{W}_0(\psi).$$

Next let us act both sides of (3.5) on the scaling subspace $\mathcal{V}_0(\phi)$ to get

$$\begin{aligned} D^m\big(\mathcal{V}_0(\phi)\big) &= \Big(D^j + \sum_{k=j}^{m-1} D^k E_0\Big)\mathcal{V}_0(\phi), \\ &= D^j\mathcal{V}_0(\phi) \oplus \bigoplus_{k=j}^{m-1} D^k \mathcal{W}_0(\psi), \end{aligned}$$

where we have made use of (3.10) and of the fact that

$$D^j\big(\mathcal{V}_0(\phi)\big) = \mathcal{V}_j(\phi) \quad \perp \quad D^k\mathcal{W}_0(\psi) = \mathcal{W}_k(\psi), \quad j \le k.$$

Therefore,

$$D^m\big(\mathcal{V}_0(\phi)\big) = \mathcal{V}_m(\phi) = \mathcal{V}_j(\phi) \oplus \bigoplus_{k=j}^{m-1} \mathcal{W}_k(\psi),$$

which indeed describes a DWT.

We summarize the above results as follows.

Proposition 3. *A DWT is described by the range space of a multiresolution decomposition of the orthogonal projection $P_m\colon \mathcal{L}^2(\mathbb{R}) \to \mathcal{L}^2(\mathbb{R})$ such that $\mathcal{R}(P_m) = \mathcal{V}_m(\phi)$ —* i.e., *of the unique orthogonal projection onto $\mathcal{V}_m(\phi)$ — as well as the image of the scaling subspace $\mathcal{V}_0(\phi)$ under a multilevel decomposition of the operator D^m,*

$$P_m\big(\mathcal{L}^2(\mathbb{R})\big) = \mathcal{V}_m(\phi) = \mathcal{V}_j(\phi) \oplus \bigoplus_{k=j}^{m-1} \mathcal{W}_k(\psi) = D^m\big(\mathcal{V}_0(\phi)\big). \tag{3.11}$$

To proceed, we define $D^m f$ as the *level-m dyadic-scale-transform* of f, then

$$D^m f = D^j f + \sum_{k=j}^{m-1} D^k E_0 f, \quad j \in [0, m-1]. \tag{3.12}$$

This leads to the following further result.

Proposition 4. *Let $f \in \mathcal{L}^2(\mathbb{R})$ then its level-m dyadic-scale-transform $D^m f$ can be computed from any lower level transform $D^j f$, and from the transforms — at all levels in $[j,\ m-1]$ — of the difference vector $E_0 f$.*

Setting $j = 0$ in (3.5) we obtain

$$D^m = I + \sum_{k=0}^{m-1} D^k E_0 \tag{3.13}$$

so that the identity I on $\mathcal{L}^2(\mathbb{R})$ admits the multilevel decomposition

$$I = D^m + \sum_{k=0}^{m-1} D^k \widehat{E}_0, \tag{3.14}$$

where

$$\widehat{E}_0 := -E_0 = I - D. \tag{3.15}$$

Consequently, any $f \in \mathcal{L}^2(\mathbb{R})$ admits the multilevel dyadic-scale-transform representation

$$f = D^m f + \sum_{k=0}^{m-1} D^k \widehat{E}_0 f, \quad f \in \mathcal{L}^2(\mathbb{R}). \tag{3.16}$$

4. Multilevel Decomposition of $\{P_m D^m\}_{m\in\mathbb{Z}}$

Let $\{P_m\}_{m\in\mathbb{Z}}$ be a sequence of orthogonal projections of an $\mathcal{L}^2(\mathbb{R})$ MRA and consider the sequence $\{P_m D^m\}_{m\in\mathbb{Z}}$. We have, as in equation (1.2), the multilevel decomposition

$$P_m D^m = P_\ell D^\ell + \sum_{k=\ell}^{m-1} (P_{k+1}D^{k+1} - P_k D^k), \quad m \in \mathbb{Z}. \tag{4.1}$$

We now recall from [7] that

$$P_m D^\mu = D^\mu P_{m-\mu}, \tag{4.2}$$

$$D^\nu P_m = P_{m+\nu} D^\nu. \tag{4.3}$$

The above results connect a level-m projection with a lower-level or a higher-level projections. What is interesting is the fact that the connection is achieved by means of a dyadic-scale transform and its adjoint transform. It follows from (4.2) that (4.1) becomes

$$P_m D^m = D^m P_0 \tag{4.4}$$

$$= \Big[D^\ell + \sum_{k=\ell}^{m-1}(D^{k+1} - D^k)\Big] P_0. \tag{4.5}$$

Equation (4.4) shows that the level-m approximation of $D^m f$ is the level-m scale transform of the level-0 approximation $P_0 f$, while equation (4.5) implies that $D^m P_0$ can be computed from a MLD.

If $f \in \mathcal{L}^2(\mathbb{R})$, then, by (4.4), its level-m approximation can be expressed as

$$P_m f = D^m P_0 D^{*m} f. \tag{4.6}$$

This suggests the following three steps for computing $P_m f$.

(i) Calculate $D^{*m} f$ from a MLD

$$D^{*m} f(\cdot) = D^{*\ell} f(\cdot) + \sum_{k=\ell}^{m-1} E_k^* f(\cdot).$$

(ii) Calculate the level-0 approximation of $D^{*m} f$ from Definition 1(o)

$$P_0 D^{*m} f(\cdot) = \sum_{n\in\mathbb{Z}} \langle D^{*m} f(\cdot)\,;\phi(\cdot - n)\rangle\, \phi(\cdot - n).$$

(iii) Calculate $D^m P_0 D^{*m} f$ directly.

These have been applied to Audio Signal Analysis [9].

5. Scaling-by-s Operator and Time-Varying Shift

We close the paper with a characterization of the scaling-by-s operator D_s on $\mathcal{L}^2(\mathbb{R})$, defined for every function $f \in \mathcal{L}^2(\mathbb{R})$ as

$$(D_s f)(t) = \sqrt{s}\, f\big(s(t)\big) \tag{5.1}$$

(for almost all t in $\mathbb{R}$ with respect to Lebesgue measure), where the real scalar $s \geq 2$ is called a *scale.* We note that s^m for each $m > 0$ is called the *level-m scale* while $\frac{1}{s^m}$ for each $m > 0$ is the *level-m resolution.* Also the dyadic-scale operator D is now the operator D_2.

If $f \in \mathcal{L}^2(\mathbb{R})$, then, for each $t \in \mathbb{R}$, $f(t-\tau)$ for every $\tau < t$ in $\mathbb{R}$ is simply $f(t)$ "delayed" by a *time-shift* of τ units, or simply a *time-shift* τ. This can be expressed in terms of the unitary group $\{T(\tau),\, \tau \in \mathbb{R}\}$ of bilateral right-shift operators over the space $\mathcal{L}^2(\mathbb{R})$ as

$$T(\tau)f(t) = f(t-\tau), \tag{5.2}$$

and the adjoint group $\{T(\tau)^*,\, \tau \in \mathbb{R}\}$ is the group of bilateral left-shift operators

$$T(\tau)^* f(t) = f(t+\tau) \tag{5.3}$$

(for almost all t in $\mathbb{R}$ with respect to Lebesgue measure). Setting $\tau = 1$ we obtain

$$T(1)f(t) = f(t-1), \tag{5.4}$$

which is the *translation-by-*1 bilateral shift operator T defined in Remark 1. The operator T and the dyadic-scale operator D play a central role in Wavelet Theory. What interesting is the fact that both T and D are bilateral shifts of infinite multiplicity, and the families $\{T^n\}_{n\in\mathbb{Z}}$ and $\{D^m\}_{m\in\mathbb{Z}}$ are discrete groups of bilateral shifts over $\mathcal{L}^2(\mathbb{R})$. However, T is time-invariant while D is time-varying, yet they are unitarily equivalent — since they have the same multiplicity [3, 12].

To proceed we represent s^1 by the level-0 decomposition

$$s^1 = s^0 + (s^1 - s^0)s^0 = s^0 + \alpha_0\, s^0, \tag{5.5}$$

where α_0, defined by

$$\alpha_0 = s^1 - s^0, \tag{5.6}$$

is the "difference" scale at level-0, that is, it is the scalar counterpart of the difference element E_0. It then follows that (with $t \in \mathbb{R}$)

$$\begin{aligned} D_s f(t) &= \sqrt{s}\, f(s\,t) \\ &= \sqrt{1+\alpha_0}\; f\big(t + [\alpha_0 s^0]\, t\big). \end{aligned} \tag{5.7}$$

This "looks" like (5.3) except that the time-shift τ now depends on $\alpha_0 s^0$ and t. We therefore define the *time-scale-shift function* τ_0 as

$$\tau_0(t) = [\alpha_0 s^0]\, t, \quad t \in \mathbb{R}. \tag{5.8}$$

Hence (5.7) becomes

$$D_s f(t) = \sqrt{1+\alpha_0}\, f\big(t + \tau_0(t)\big). \tag{5.9}$$

The operator D_s can therefore be considered as a *left-time-varying shift* — with the time-shift function τ_0. We note that time-varying amplitude and time-varying phase have been defined for analytic signals, see [2] and the references therein.

In general we have, for $m > 0$,

$$s^m = s^0 + \sum_{k=0}^{m-1} \alpha_0\, s^k. \tag{5.10}$$

This can be considered as an "unfolding" of the level-m scale s^m into scales at levels not greater than $m-1$, beginning with the "initial" scale s^0. It is easy to see that

$$D_s^m f(t) = \sqrt{1+\alpha_0}^{\,m} f\big(t + \tau_{m-1}(t)\big), \tag{5.11}$$

where $\tau_{m-1}(t)$ is defined by

$$\tau_{m-1}(t) = \sum_{k=0}^{m-1} [\alpha_0\, s^k]\, t, \quad m \geq 1, \tag{5.12}$$

Similarly, $\frac{1}{s^m}$ for each $m \geq 1$ admits the multilevel decomposition

$$\frac{1}{s^m} = \frac{1}{s^0} - \sum_{k=0}^{m-1} \beta_0\, \frac{1}{s^k}, \quad m \geq 1, \tag{5.13}$$

where

$$\beta_0 = \frac{1}{s^0} - \frac{1}{s^1} = \frac{s^1 - s^0}{s^1} = \frac{\alpha_0}{1+\alpha_0} > 0. \tag{5.14}$$

It is easy to see that the adjoint operator D_s^* is given, for each $t \in \mathbb{R}$, by

$$D_s^*\, f(t) = \frac{1}{\sqrt{s}}\, f\Big(\frac{1}{s}\, t\Big) \tag{5.15}$$

$$= \frac{1}{\sqrt{1+\alpha_0}}\, f\Big(\frac{1}{1+\alpha_0}\, t\Big). \tag{5.16}$$

But, from (5.14)

$$\frac{1}{1+\alpha_0} = 1 - \beta_0.$$

Therefore

$$D_s^* f(t) = \sqrt{1-\beta_0}\, f\Big(\big[1 - \beta_0 \frac{1}{s^0}\big] t\Big), \tag{5.17}$$

or

$$D_s^* f(t) = \sqrt{1-\beta_0}\, f\big(t - \tau_{*,0}(t)\big). \tag{5.18}$$

Consequently, the operator D_s^* is a right-time-varying shift, with the *time-resolution-shift function* given, for each $t \in \mathbb{R}$, by

$$\tau_{*,0}(t) = \beta_0\, \frac{1}{s^0}\, t. \tag{5.19}$$

Finally,

$$D_s^{*\,m} f(t) = \sqrt{1-\beta_0}^{\,m}\, f\big(t - \tau_{*,m-1}(t)\big), \tag{5.20}$$

where, for each $t \in \mathbb{R}$,

$$\tau_{*,m-1} = \sum_{k=0}^{m-1} \Big(\beta_0\, \frac{1}{s^k}\Big)\, t, \quad m \geq 1. \tag{5.21}$$

The above establish another connection between the two bilateral shifts D and T. Application of this in Wavelet Theory will be reported elsewhere.

We have in this paper shown that the “natural” Multilevel Decomposition of a sequence of bounded linear Hilbert space operators can, under suitable conditions, lead to Multilevel Approximation of vectors in the space. This is indeed the case of the Multiresolution Approximation of Wavelet Theory, as well as the case of the sequence of the operators $P_m D^m$. A new interpretation of a Dilation-by-s operator as a time-varying shift operator was also obtained from a MLD of the scale s.

References

1. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM, Philadelphia, 1992.
2. P. Flandrin, *Time-Frequency / Time-Scale Analysis*, Academic Press, San Diego, 1999.
3. P.R. Halmos, *Shifts on Hilbert spaces*, J. Reine Angew. Math. **208** (1961), 102–112.
4. E. Hernandez and G. Weiss, *A First Course on Wavelets*, CRC Press, Boca Raton, 1996.
5. F. Keinert, *Wavelets and Multiwavelets*, CRC Press, Boca Raton, 2004.
6. C.S. Kubrusly, *Hilbert Space Operators*, Birkhäuser, Boston, 2003.
7. C.S. Kubrusly and N. Levan, *Abstract wavelets generated by Hilbert space shift operators*, Adv. Math. Sci. Appl. **16** (2006), 643–660.
8. N. Levan and C.S. Kubrusly, *Multiresolution approximation scale and time-shift subspaces*, Multidimens. Syst. Signal Process. **17** (2006), 343–354.
9. S. Majumdar, *New Approach to Operator-Theoretic Frameworks for Wavelet Multiresolution Approximation and Analysis*, Ph.D Dissertation, Department of Electrical Engineering, UCLA, Los Angeles, June 2006.
10. S.G. Mallat, *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbb{R})$, Trans. Amer. Math. Soc. **315** (1989), 69–87.
11. Y. Meyer, *Ondelettes et Functions Spline*, Séminaire EDP, École Polytechnique, Paris, 1986.
12. B. Sz.-Nagy and C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space,* North-Holland, Amsterdam, 1970.

Qualcomm Inc., 5775 Morehouse Dr., San Diego, CA 92121
E-mail address: smajumda@qualcomm.com

Department of Electrical Engineering, University of California in Los Angeles, Los Angeles, CA 90024-1594, USA
E-mail address: levan@ee.ucla.edu

Catholic University of Rio de Janeiro, 22453-900, Rio de Janeiro, RJ, Brazil
E-mail address: carlos@ele.puc-rio.br