

# Interpolating Wavelets and Difference Wavelets \*

I-Liang Chern <sup>†</sup>

Email: chern@math.ntu.edu.tw

## Abstract

This paper first shows that the Cohen-Daubechies-Feauveau (C-D-F) biorthogonal wavelet can be derived from the interpolating wavelet through a lifting process. Its high-pass filter measures the interpolation error of the averaged data. Next, we propose a new wavelet method, called the difference wavelet method, for efficient representation for functions on  $R$ . Its analysis part is simply averaging and finite differencing. The corresponding synthesis part involves a finite difference equation, which is solved by the cyclic reduction method to achieve fast transform. The associated wavelets constitute biorthogonal Riesz bases in  $L^2(R)$ . Their decay and regularity properties are investigated. A comparison study on the efficiency issues with C-D-F wavelet method is performed. The comparison includes (i) operation counts for performing wavelet transform, (ii) the approximate power (i.e. the coefficient in the approximation error estimate), and (iii) the compression ratio (by numerical experiments). The results show that the difference wavelet method is about twice more efficient than the C-D-F wavelet method. This efficiency is due to that the difference wavelets are more regular with just slightly bigger essential support than those of C-D-F wavelets.

## 1 Introduction

Wavelet expansion is an efficient way to represent functions and operators. In such an expansion, a function is decomposed into *fluctuations* at various resolutions plus an *averaged* information at the coarsest resolution. The expansion is done by performing a fast wavelet transform through so called analysis filter banks. The original function can be recovered from the wavelet coefficients through so called synthesis filter banks. The advantage of such an expansion is that only small amount of coefficients are needed to achieve accurate approximation, provided the function is piecewise smooth. This technique is widely used for data compression in digital signal processing [14], for designing fast algorithms for matrix-vector multiplications [1], for solving integral equations, etc. The efficiency of such a representation is due to that the wavelet coefficients indeed measure certain interpolation

---

\*This work was supported by the National Science Council of the Republic of China under Contract NSC86-2115-M002-008

<sup>†</sup>Department of Mathematics, National Taiwan University, Taipei, Taiwan, R.O.C.

errors of functions or operators at various scales. They are negligible, provided the given functions or operators are piecewise smooth.

In order to have such expansions in  $L^2(R)$ , the forementioned interpolation error should be taken to be the interpolation error of the *averaged* data instead of the *point-valued* data. Thus, there are two important parameters in a wavelet expansion method: the order of averaging  $r$ , and the order of interpolation  $\tilde{r}$ . In the orthogonal wavelet of Daubechies [5] and the semi-orthogonal wavelet of Chui-Wang [3], these two parameters are taken to be identical. In Cohen-Daubechies-Feauveau's (C-D-F) biorthogonal wavelet [4],  $\tilde{r}$  is allowed to be different from  $r$ , but with a constraint that  $r + \tilde{r}$  must be an even number in order to be bistable in  $L^2(R)$ . The C-D-F wavelet method is indeed more flexible to use. Numerical experiments also show that one can get better compression ratio for  $r < \tilde{r}$ .

In this article, we first show that the C-D-F wavelet can be derived from the interpolating wavelet through a lifting process. Its high-pass filter measures the interpolation error of the averaged data. We thus refer the C-D-F wavelets as the interpolating wavelets. Indeed, C-D-F wavelet method is the most ideal method in the class of finite (i.e. polynomial) filter banks that allows different  $r$  and  $\tilde{r}$ .

Next, we look for efficient wavelet expansion method in the class of *rational* filter banks that also allows different  $r$  and  $\tilde{r}$ . We would like to have such expansion method to be *bistable* in  $L^2(R)$ , more *efficient*, and with the corresponding expansion algorithm to be *fast*.

Notice that one way to derive the C-D-F filter banks is to the smallest degree polynomial that satisfies the no distortion condition:

$$Q(z) \left( \frac{1+z}{2} \right)^{2K} + Q(-z) \left( \frac{1-z}{2} \right)^{2K} = 1,$$

where  $r + \tilde{r} = 2K$ . After proper shift of power,  $Q(z)$  has the following explicit form

$$Q_K(z) = \sum_{n=0}^{K-1} \binom{K-1+n}{n} \left( \frac{2-z-z^{-1}}{4} \right)^n. \quad (1.1)$$

We notice that  $Q_K(-1)$  is a big number as  $K$  is big. This has the effect that the filter bank  $Q_K(z)$  amplifies the low-frequency modes during the analysis filtering and the high-frequency modes during synthesis filtering. In this paper, we search for efficient filter banks by finding rational functions of the form  $Q(z)/P(z^2)$  that has smallest magnitude for  $|z| \leq 1$  and satisfies

$$\left( \frac{1+z}{2} \right)^{2K} Q(z)/P(z^2) + \left( \frac{1-z}{2} \right)^{2K} Q(-z)/P(z^2) = 1.$$

It turns out the optimal choice is  $Q(z) \equiv 1$  and

$$P_K(z^2) = z^{-K} \left( \frac{1+z}{2} \right)^{2K} + (-z)^{-K} \left( \frac{1-z}{2} \right)^{2K}.$$

We call this wavelet method the difference wavelet because the corresponding high-pass filtering is simply finite differencing. In this case, the magnitude of  $|1/P_K(z^2)|$  for  $|z| \leq 1$  is much smaller than  $|Q_K(z)|$  in the C-D-F filter banks. As a result, the corresponding wavelet function is more regular, hence it has better approximate power and better compression ability. Although the filter bank  $1/P_K(z^2)$  is of infinite length, it can be performed by solving a finite difference equation by the cyclic reduction method. The operation count of such an algorithm is about the same as performing  $Q_K(z)$  in the C-D-F wavelet method. Also, the difference wavelet function is not of compact support, yet it decays exponentially fast at infinity. Its *essential support* is only slightly bigger than the corresponding C-D-F wavelet. Numerical experiments show that difference wavelets do have good “zoom-in” property.

In this proceeding paper, only the contents of theorems are stated here. Their proofs will be provided in a separated journal paper. This paper is organized as the follows. Section 2 is the notation and preliminary, where we review the notion of filter banks and biorthogonality. In Section 3, we review the C-D-F biorthogonal wavelet and show that they can be derived from interpolating filter by a lifting technique. We also explain that the corresponding wavelet coefficients measure the interpolation error of the averaged data. In Section 4, we illustrate the difference wavelet method, compute its operation counts, and show their biorthogonality property. We also investigate the regularity and decay properties of the corresponding scaling function. Section 5 is a comparison study between the difference wavelet method and the C-D-F wavelet method on efficiency issues, including the approximation power and compression ratio. We demonstrate that this better efficiency of the difference wavelet method is due its better regularity and its support being “essentially” finite.

## 2 Preliminary Notations

### 2.1 Filter banks and fast wavelet transform

A wavelet transform decomposes data (or functions, or operators) into fluctuations at various resolutions. It depends on four sets of coefficients:  $\{h_k\}_{k \in \mathbb{Z}}$ ,  $\{g_k\}_{k \in \mathbb{Z}}$ ,  $\{\tilde{h}_k\}_{k \in \mathbb{Z}}$  and  $\{\tilde{g}_k\}_{k \in \mathbb{Z}}$ . The first two are called the analysis filter banks, the latter two the synthesis filter banks. A data sequence  $u^\ell = \{u_j^\ell\}$  at resolution level  $\ell$  can be decomposed, through the analysis filter banks, into the following two sets of data sequences at level  $\ell - 1$ :

$$\begin{cases} \text{the low-pass data: } u_j^{\ell-1} = \sqrt{2} \sum_k h_k u_{2j-k}^\ell, \\ \text{the high-pass data: } w_j^{\ell-1} = \sqrt{2} \sum_k g_k u_{2j+1-k}^\ell. \end{cases} \quad (2.1)$$

Here,  $\sqrt{2}$  is a normalized scale factor [6]. Roughly speaking, the low-pass data are the averaged informations of  $u^\ell$  at coarse grid, while the high-pass data the fluctuation informations (or the differencing) of  $u^\ell$  away from  $u^{\ell-1}$ . By applying the above transform:  $u^\ell \mapsto (u^{\ell-1}, w^{\ell-1})$  recursively for  $\ell = L, L-1, \dots, 1$ , one can decompose a given

data  $u^L$  at the finest scale  $L$  into  $(u^0, w^0, w^1, \dots, w^{L-1})$ , the averaged information at coarsest resolution and the fluctuation informations at various resolutions. The mapping:  $u^L \mapsto (u^0, w^0, w^1, \dots, w^{L-1})$  is called a wavelet transform.

The data  $u^\ell$  can be recovered from  $(u^{\ell-1}, w^{\ell-1})$  through the synthesis filter banks  $\{\tilde{h}_k\}_{k \in Z}$  and  $\{\tilde{g}_k\}_{k \in Z}$ :

$$u_j^\ell = \sqrt{2} \sum_k \left[ \tilde{h}_{j-2k} u_k^{\ell-1} + \tilde{g}_{j-2k-1} w_k^{\ell-1} \right]. \quad (2.2)$$

By applying this inverse transform recursively for  $\ell = 1, \dots, L-1$ ,  $u^L$  can be recovered from  $(u^0, w^0, \dots, w^{L-1})$ .

Notice that the amount of information in  $(u^0, w^0, \dots, w^{L-1})$  is identical to the original  $u^L$ . The advantage of a wavelet representation is that many wavelet coefficients (i.e.  $w_j^\ell$ ,  $j \in Z$ ,  $\ell = 0, \dots, L-1$ ) become negligible when  $u^L$  is piecewise smooth. Thus, a data sequence can be represented more efficiently through a wavelet expansion.

Since the above filtering process is in convolution form, it is convenient to introduce the  $z$ -transform (or the generating function) of these filter banks. Let us define  $h(z) = \sum_k h_k z^k$ , etc. By taking the  $z$ -transforms of (2.1) and (2.2), we find that the filter banks satisfy

$$\begin{aligned} u^\ell(z) &= \sqrt{2} \tilde{h}(z) u^{\ell-1}(z^2) + \sqrt{2} \tilde{g}(z) z w^{\ell-1}(z^2) \\ &= \left( \tilde{h}(z) h(z) + \tilde{g}(z) g(z) \right) \cdot u^\ell(z) \\ &\quad + \left( \tilde{h}(z) h(-z) - \tilde{g}(z) g(-z) \right) \cdot u^\ell(-z). \end{aligned}$$

This leads to

$$h(z) \tilde{h}(z) + g(z) \tilde{g}(z) = 1 \quad (\text{no distortion}), \quad (2.3)$$

$$h(-z) \tilde{h}(z) - g(-z) \tilde{g}(z) = 0 \quad (\text{alias cancellation}). \quad (2.4)$$

A sufficient condition for (2.4) is

$$\begin{aligned} g(z) &= \tilde{h}(-z) P(z^2), \\ \tilde{g}(z) &= h(-z) / P(z^2). \end{aligned}$$

A general restriction on  $P(z^2)$  is

$$P(z^2) = \sum_k p_{2k} z^{2k} \text{ with } \sum |p_{2k}| < \infty.$$

(see Chui [2]). Thus,  $h$  and  $\tilde{h}$  satisfy

$$h(z) \tilde{h}(z) + h(-z) \tilde{h}(-z) = 1. \quad (2.5)$$

We shall also normalize  $h$  and  $\tilde{h}$  by

$$h(1) = \tilde{h}(1) = 1. \quad (2.6)$$

Thus, we shall design filter banks  $h(z)$  and  $\tilde{h}(z)$  to satisfy (2.5) (2.6) and choose a proper  $P(z^2)$  to obtain the filter banks  $g(z)$  and  $\tilde{g}(z)$ .

A general design principle of filter banks in applications is to have the corresponding wavelet transform to be *fast, bistable and efficient* in the following sense. The term “fast” means that the wavelet transform and its inverse is of linear computational complexity. The term “bistable” means that the transform is bistable in  $\ell^2$ , independent of the resolution  $L$ . The term “efficient” means that a small amount of nonzero wavelet coefficients  $w_j^\ell$  suffices to approximate the original data accurately.

## 2.2 Bistability

The bistability of the wavelet transform is studied through the helps of the scaling functions and wavelets. Such bistability is equivalent to the biorthogonality of a pair of scaling function  $\phi$  and  $\tilde{\phi}$  in  $L^2(R)$ , or equivalent to the biorthogonality of the wavelets  $\psi$  and  $\tilde{\psi}$  [4].

### 2.2.1 Scaling functions and wavelets

Associated with the analysis filter banks, one can define the scaling functions  $\phi$  and wavelet function  $\psi$  by

$$\phi(x) = 2 \sum_k h_{-k} \phi(2x - k), \quad (2.7)$$

$$\psi(x) = 2 \sum_k g_{1-k} \phi(2x - k). \quad (2.8)$$

The scaling function  $\phi$  exists as a distribution if and only if  $h(1) = 1$ . Once  $\phi$  is found,  $\psi$  can be obtained from (2.8). The connection between the scaling function, the wavelet and the forementioned analysis filtering is as the follows. First, define  $\phi_j^\ell(\cdot) := 2^{\frac{\ell}{2}} \phi(2^\ell \cdot - j)$ ,  $\psi_j^\ell(\cdot) := 2^{\frac{\ell}{2}} \psi(2^\ell \cdot - j)$ . Here, the factor  $2^{\ell/2}$  is for normalization in  $L^2(R)$ . Next, given a function  $u \in L^2(R)$ , define  $u_j^\ell := (u, \phi_j^\ell)$  and  $w_j^\ell := (u, \psi_j^\ell)$ . Then it is easy to see that the mapping  $(u_j^\ell)_{j \in \mathbb{Z}} \mapsto (u_j^{\ell-1}, w_j^{\ell-1})_{j \in \mathbb{Z}}$  is the analysis filtering (2.1) if and only if  $\phi$  and  $\psi$  satisfy (2.7) and (2.8).

The synthesis filter banks correspond to a dual pair of functions  $\tilde{\phi}$  and  $\tilde{\psi}$  defined by

$$\tilde{\phi}(x) = 2 \sum_k \tilde{h}_k \tilde{\phi}(2x - k), \quad (2.9)$$

$$\tilde{\psi}(x) = 2 \sum_k \tilde{g}_{k-1} \tilde{\phi}(2x - k). \quad (2.10)$$

Let us also define  $\tilde{\phi}_j^\ell(\cdot) := 2^{\frac{\ell}{2}} \tilde{\phi}(2^\ell \cdot - j)$  and  $\tilde{\psi}_j^\ell(\cdot) := 2^{\frac{\ell}{2}} \tilde{\psi}(2^\ell \cdot - j)$ . Then the condition (2.2) in the synthesis filtering is equivalent to

$$\sum_j (u, \phi_j^\ell) \tilde{\phi}_j^\ell = \sum_j (u, \phi_j^{\ell-1}) \tilde{\phi}_j^{\ell-1} + \sum_j (u, \psi_j^{\ell-1}) \tilde{\psi}_j^{\ell-1}. \quad (2.11)$$

Let  $\tilde{V}^\ell$  denote the space  $\text{span}\{\tilde{\phi}_k^\ell \mid k \in Z\}$ . Given a function  $u \in L^2(R)$ , the data  $\{u_j^L \equiv (u, \phi_j^L)\}_{j \in Z}$  is equivalent to the projection of  $u$  in  $V^L$ :

$$u^L(x) = \sum_j (u, \phi_j^L) \tilde{\phi}_j^L(x).$$

The wavelet expansion  $(u^0, w_0, \dots, w^{L-1})$  of  $\{u_j^L\}_{j \in Z}$  is equivalent to the expansion of  $u^L(x)$  with the basis  $\{\tilde{\phi}_j^0, \tilde{\psi}_j^0, \dots, \tilde{\psi}_j^{L-1}\}_{j \in Z}$ :

$$u^L(x) = \sum_j (u, \phi_j^L) \tilde{\phi}_j^L(x) \quad (2.12)$$

$$= \sum_j (u, \phi_j^0) \tilde{\phi}_j^0(x) + \sum_{\ell=0}^{L-1} \sum_j (u, \psi_j^\ell) \tilde{\psi}_j^\ell(x). \quad (2.13)$$

Thus, the bistability of the discrete wavelet transform is equivalent to that the formula (2.13) is valid in  $L^2(R)$  sense. It turns out that this is equivalent to that

$$u(x) = \lim_{L \rightarrow \infty} \sum_j (u, \phi_j^L) \tilde{\phi}_j^L(x) \quad (2.14)$$

$$= \sum_{\ell=-\infty}^{\infty} \sum_j (u, \psi_j^\ell) \tilde{\psi}_j^\ell(x). \quad (2.15)$$

is valid in  $L^2(R)$  sense [4].

### 2.2.2 Biorthogonality

In order to have the expansion (2.12) valid in  $L^2(R)$ , we need the following definition.

**Definition 1** We say that a function  $\phi \in L^2(R)$  has the Riesz basis property if  $\{\phi(\cdot - k)\}_{k \in Z}$  constitute a Riesz basis in the space  $V^0 := \text{span}\{\phi(\cdot - k) \mid k \in Z\}$ , i.e. there exist constants  $C_1 > 0, C_2 < \infty$ , such that for all these finite sums,

$$C_1 \sum |c_k|^2 \leq \left\| \sum c_k \phi(\cdot - k) \right\|_{L^2}^2 \leq C_2 \sum |c_k|^2. \quad (2.16)$$

In order to have the limit (2.14) valid in  $L^2(R)$ , Mallat [10] introduced the concept of multiresolution analysis.

**Definition 2** A sequence of function spaces  $\{V^\ell\}_{\ell \in Z}$  is said to constitute a multiresolution analysis in  $L^2(R)$  if

1.  $V^{\ell-1} \subset V^\ell$ ,  $\overline{\bigcup_{\ell \in Z} V^\ell} = L^2(R)$ , and  $\bigcap_{\ell \in Z} V^\ell = \{0\}$ ,
2. translation invariance: if  $u \in V^0$ , then  $u(\cdot - j) \in V^0$  for all  $j \in Z$ ,

3. dilation invariance:  $u \in V^0$  if and only if  $u(2^\ell \cdot) \in V^\ell$ ,
4. Riesz basis property: there exists a function  $\phi \in L^2$  such that  $\phi$  has the Riesz basis property and  $V^0 = \text{Span} \{\phi(\cdot - j)\}_{j \in \mathbb{Z}}$ .

That the expansion (2.14) in  $L^2(R)$  sense leads to the following definition of biorthogonality.

**Definition 3** Two scaling functions  $\phi$  and  $\tilde{\phi}$  are called biorthogonal in  $L^2(R)$  if

1.  $\phi$  and  $\tilde{\phi}$  have the Riesz basis property,
2. let  $V^\ell = \text{Span}\{\phi_j^\ell\}_{j \in \mathbb{Z}}$  and  $\tilde{V}^\ell = \text{Span}\{\tilde{\phi}_j^\ell\}_{j \in \mathbb{Z}}$ , then both  $\{V^\ell\}_{\ell \in \mathbb{Z}}$  and  $\{\tilde{V}^\ell\}_{\ell \in \mathbb{Z}}$  constitute multiresolution analysis in  $L^2(R)$ ,
3. for each  $\ell$ ,  $\{\phi_j^\ell\}_{j \in \mathbb{Z}}$  and  $\{\tilde{\phi}_k^\ell\}_{k \in \mathbb{Z}}$  are dual to each other, i.e.  $(\phi_j^\ell, \tilde{\phi}_k^\ell) = \delta_{jk}$ .

**Definition 4** Two wavelets  $\psi$  and  $\tilde{\psi}$  are said to be biorthogonal in  $L^2(R)$  if

1. both  $\{\psi_j^\ell\}_{j, \ell \in \mathbb{Z}}$  and  $\{\tilde{\psi}_j^\ell\}_{j, \ell \in \mathbb{Z}}$  constitute Riesz bases in  $L^2(R)$ ,
2.  $\{\psi_j^\ell\}_{j, \ell \in \mathbb{Z}}$  and  $\{\tilde{\psi}_j^\ell\}_{j, \ell \in \mathbb{Z}}$  are dual to each other:

$$(\psi_j^\ell, \tilde{\psi}_k^m) = \delta_{jk} \delta_{\ell m}. \quad (2.17)$$

The main theorems that characterize the biorthogonality are the following.

**Theorem 2.1 (Cohen-Daubechies-Feauveau)** The following two statements are equivalent.

1. The scaling functions  $\phi$  and  $\tilde{\phi}$  have the Riesz basis property and the corresponding filter banks  $h$  and  $\tilde{h}$  satisfy (2.5), (2.6).
2.  $\phi$  and  $\tilde{\phi}$  are biorthogonal in  $L^2(R)$ .

When  $g(z) = \tilde{h}(-z)$  and  $\tilde{g}(z) = h(-z)$ , Cohen-Daubechies-Feauveau also showed that the biorthogonality of  $\phi$  and  $\tilde{\phi}$  in  $L^2(R)$  is equivalent to the biorthogonality of  $\psi$  and  $\tilde{\psi}$  in  $L^2(R)$ . A more general theorem due to Chui [2] is stated below.

**Theorem 2.2 (Chui)** If  $\phi$  and  $\tilde{\phi}$  are biorthogonal in  $L^2(R)$ , then  $\psi$  and  $\tilde{\psi}$  are biorthogonal in  $L^2(R)$  if and only if  $g(z) = \tilde{h}(-z)P(z^2)$  and  $\tilde{g}(z) = h(-z)/P(z^2)$  with  $P(z) = \sum_k p_k z^k$  satisfying  $\sum_k |p_k| < \infty$ .

### 3 The Interpolating Wavelet Method

#### 3.1 The pointwise interpolating wavelet

The interpolating filter has been introduced by many researchers [9, 7, 8]. It depends on a parameter  $K > 0$ . The analysis filtering is defined by

$$\begin{aligned} u_j^{\ell-1} &= u_{2j}^\ell \\ w_j^{\ell-1} &= u_{2j+1}^\ell - R_K(x_{2j+1}^\ell, u^{\ell-1}), \end{aligned} \quad (3.1)$$

where  $R_K(x, u^{\ell-1})$  is the interpolating polynomial that interpolates  $u_{j-k}^{\ell-1}$  at  $x_{j-k}^{\ell-1}$  for  $k = -K+1, \dots, K$ , and  $x_j^\ell \equiv j2^{-\ell}$ . In other words,  $w_j^{\ell-1}$  measures the interpolation error at  $x_{2j+1}^\ell$ . The synthesis filter is given by

$$\begin{aligned} u_{2j}^\ell &= u_j^{\ell-1} \\ u_{2j+1}^\ell &= w_j^{\ell-1} + R_K(x_{2j+1}^\ell, u^{\ell-1}). \end{aligned}$$

Thus, the corresponding filter banks are

$$h(z) = \tilde{g}(-z) \equiv 1 \quad (3.2)$$

$$\tilde{h}(z) = g(-z) = 1 + \sum_{k=1}^K \left( \prod_{\substack{m \neq k \\ m = -K+1}}^K \frac{\frac{1}{2} - m}{k - m} \right) (z^{2k-1} + z^{-2k+1}). \quad (3.3)$$

Since  $h(z) = 1$ , the corresponding scaling function is a delta function. The feature of the interpolating filter is that smooth functions can be approximated by  $\tilde{\phi}$  in some proper function space through interpolation [8]:

$$u(x) = \lim_{L \rightarrow \infty} \sum_j u(x_j^L) \tilde{\phi}_j^L(x).$$

#### 3.2 The Cohen-Daubechies-Feauveau biorthogonal wavelets

Notice that the forementioned pointwise version of interpolating wavelet is not biorthogonal in  $L^2(R)$  because the corresponding  $\phi$  is a delta function. Cohen-Daubechies-Feauveau introduced biorthogonal wavelets. It turns out that their biorthogonal wavelet can be derived from the interpolating wavelets through a lifting process. Let us first give a brief review of the C-D-F wavelet method [4]. Their filter banks are defined by

$$h(z) = \tilde{g}(-z) = z^{-[r/2]} \left( \frac{1+z}{2} \right)^r, \quad (3.4)$$

$$\tilde{h}(z) = g(-z) = z^{-[\tilde{r}/2]-1} \left( \frac{1+z}{2} \right)^{\tilde{r}} Q(z). \quad (3.5)$$



Here,  $r + \tilde{r}$  is required to be an even number  $2K$ , and  $Q(z)$  is the minimal degree polynomial satisfying

$$z^{-K} \left( \frac{1+z}{2} \right)^{2K} Q(z) + (-z)^{-K} \left( \frac{1-z}{2} \right)^{2K} Q(-z) = 1. \quad (3.6)$$

It has the following explicit expression

$$Q(z) \equiv Q_K(z) = \sum_{n=0}^{K-1} \binom{K-1+n}{n} \left( \frac{2-z-z^{-1}}{4} \right)^n. \quad (3.7)$$

The C-D-F wavelets have the following advantages.

1. The expansion algorithm is fast.
2. The corresponding wavelets  $\psi$  and  $\tilde{\psi}$  are biorthogonal.
3. The scaling functions and wavelets are of compact support and symmetric, thus have good zoom-in property and non-bias approximation ability.
4. They are flexible, namely, allowing different orders of averaging and interpolation.
5. It has the shortest length of filter banks that have the above properties.

### 3.3 The C-D-F wavelet as a lifted interpolating wavelet

The C-D-F filter can be viewed as a lifted interpolating filter. The lifting process is a process to derive the filter banks of  $(r+1, \tilde{r}-1)$  from the filter banks of  $(r, \tilde{r})$ . Let  $\phi^r$  and  $\tilde{\phi}^{\tilde{r}}$  (resp.  $h^r(z)$  and  $\tilde{h}^{\tilde{r}}(z)$ ) be the scaling functions (resp. the filter banks) associated with the C-D-F filter banks with averaging order  $r$  and interpolation order  $\tilde{r}$ .

**Theorem 3.1** *The C-D-F filter banks with  $r = 0, \tilde{r} = 2K$  is identical to the order  $2K$  interpolating filter banks (3.2) (3.3)*

**Theorem 3.2** *The C-D-F filter bank with  $(r+1, \tilde{r}-1)$  can be derived from the case  $(r, \tilde{r})$  as the follows.*

$$\begin{aligned} h^{r+1}(z) &= \left( \frac{1+z}{2} \right) h^r(z) \\ \tilde{h}^{\tilde{r}}(z) &= \left( \frac{1-z}{2} \right) \tilde{h}^{\tilde{r}-1}(z) \\ \phi^{r+1} &= 1_{[0,1)} * \phi^r \\ \tilde{\phi}^{\tilde{r}} &= 1_{[0,1)} * \tilde{\phi}^{\tilde{r}-1} \end{aligned}$$

**Remarks.**

1. The lifting process on filter banks corresponds to an averaging process on the data (i.e.  $h(z) = \left(\frac{1+z}{2}\right)^r$ ). Thus, the C-D-F high-pass filtering can be viewed as an “interpolation error” of the “averaged data.”
2. Harten’s cell average version of interpolating wavelet [9] is exactly identical to the C-D-F wavelet with  $r = 1$  and  $\tilde{r} = 2K - 1$ .
3. The price to pay for the filtering  $Q_K(z)$  is the following. Since

$$Q_K(-1) = \sum_{n=0}^{K-1} \binom{K+n-1}{n}.$$

Hence, low frequency modes are amplified by  $Q_K$  during high-pass filtering, and high frequency modes are also amplified during low-pass synthesis filtering. In the next section, we replace  $Q_K(z)$  by  $1/P_K(z^2)$ . Figure 1 demonstrates that the magnitude of  $|1/P(z^2)|$  is much smaller than that of  $|Q(z)|$ . As a result, the corresponding scaling function is more regular, hence has better approximation ability.

## 4 Difference Wavelets

### 4.1 The filter banks of the difference wavelet method

In order to have better approximation ability, we search for filter banks in the class of rational functions. Given  $K > 0$ , we find rational function of the form  $Q(z)/P(z^2)$  that satisfies

$$\left(\frac{1+z}{2}\right)^{2K} Q(z)/P(z^2) + \left(\frac{1-z}{2}\right)^{2K} Q(-z)/P(z^2) = 1,$$

and has minimal magnitude in absolute value for  $|z| \leq 1$ . It turns out that the optimal choice is, after proper shift of power, that

$$Q(z) \equiv 1 \tag{4.1}$$

$$P_K(z) = z^{-K} \left(\frac{1+z}{2}\right)^{2K} + (-z)^{-K} \left(\frac{1-z}{2}\right)^{2K}. \tag{4.2}$$

We then define the filter banks to be

$$\begin{aligned} h(z) &= z^{-[\frac{r}{2}]} \left(\frac{1+z}{2}\right)^r \\ g(z) &= z^{-[\frac{\tilde{r}}{2}]-1} \left(\frac{1-z}{2}\right)^{\tilde{r}} \\ \tilde{h}(z) &= z^{-[\frac{\tilde{r}}{2}]-1} \left(\frac{1+z}{2}\right)^{\tilde{r}} / P_K(z^2) \\ \tilde{g}(z) &= z^{-[\frac{r}{2}]} \left(\frac{1-z}{2}\right)^r / P_K(z^2). \end{aligned}$$

Thus, the only difference from C-D-F filter is to replace the  $Q(z)$  filter by  $1/P_K(z^2)$ . We call this method the difference wavelet method because its high-pass filtering is simply a finite difference operator. To justify this method is valuable we need to show that

1. the operation count to perform  $1/P_K(z^2)$  is almost the same as that of  $Q(z)$ ,
2. it is bistable,
3. it is more efficient in the sense that it has better approximation ability and better compression ratio.

We shall devote to these issues below.

## 4.2 Fast algorithm for difference wavelet transform

The filtering  $1/P_K(z^2)$  is performed by the following two steps.

Step 1. Factorize  $P_K(z^2)$  into

$$P_K(z^2) = \prod_{k=1}^{[K/2]} \frac{1}{1 + 2\alpha_k} \left( \alpha_k z^{-2} + 1 + \alpha_k z^2 \right) \equiv \prod_{k=1}^{[K/2]} P^k(z^2),$$

where

$$\begin{aligned} \alpha_k &= 1 / \left( \tan^2 \theta_k + 1 / \tan^2 \theta_k \right) < 1/2, \\ \theta_k &= \begin{cases} \frac{(2k-1)\pi}{4K} & \text{if } K \text{ is even} \\ \frac{k\pi}{2K} & \text{if } K \text{ is odd} \end{cases} \end{aligned}$$

Step 2. For each  $k = 1, \dots, [K/2]$ , solve  $1/P^k(z^2)$  by the cyclic reduction method.

The operation count for performing  $1/P_K(z^2)$  is less than  $(4A + 3M)[K/2]$  per each datum, whereas the operation count for performing  $Q_K(z)$  is  $(2A + 1M)K$  per each datum. Here  $A$  is the addition operation and  $M$  is the multiplication. Thus, the operation counts for these two methods are about the same.

## 4.3 Properties of difference wavelets

Consider the difference filter banks with averaging order  $r$  and differencing order  $\tilde{r}$ , and  $r + \tilde{r} = 2K > 0$ . Then the scaling function  $\phi$  is the spline:  $\phi = 1_{[0,1]}^*$ . The dual scaling function  $\tilde{\phi}$  has the following properties.

**Theorem 4.1** 1.  $\tilde{\phi}$  is symmetric about 0 and decays at  $\infty$  exponentially fast:

$$\begin{aligned} \tilde{\phi}(x) &= O(e^{-\sigma|x|}), \\ \sigma &= 2 \ln \tan \left( \frac{K+1}{4K} \right) \pi > 0. \end{aligned}$$

2. The Fourier transform of  $\tilde{\phi}$  satisfies

$$|\hat{\tilde{\phi}}(\xi)| = O\left(|\xi|^{-(\tilde{r}-r)/2-1}\right). \quad (4.3)$$

**Theorem 4.2** *The difference wavelets with  $\tilde{r} \geq r \geq 1$  are biorthogonal in  $L^2(R)$ .*

We compare the regularity of the scaling functions corresponding the difference wavelet method and the C-D-F wavelet method. The Fourier transform of  $\tilde{\phi}$  of both method are given by

$$\begin{aligned} \hat{\tilde{\phi}}(\xi) &= \left(\frac{\sin \frac{\xi}{2}}{\frac{\xi}{2}}\right)^{\tilde{r}} \cdot \prod_{\ell=1}^{\infty} Q_K(e^{-i\xi/2^\ell}) \quad (\text{C-D-F}) \\ \hat{\tilde{\phi}}(\xi) &= \left(\frac{\sin \frac{\xi}{2}}{\frac{\xi}{2}}\right)^{\tilde{r}} \cdot \prod_{\ell=1}^{\infty} \frac{1}{P_K(e^{-2i\xi/2^\ell})} \quad (\text{difference}). \end{aligned}$$

Table 1 is the growth power of  $\prod_{\ell=1}^{\infty} Q_K(e^{-i\xi/2^\ell})$  and  $\prod_{\ell=1}^{\infty} 1/P_K(e^{-2i\xi/2^\ell})$ . It shows that the scaling function  $\tilde{\phi}$  of the difference wavelet is smoother than that of the C-D-F wavelet.

K	2	3	4	5	6	7	8	9	10	20
$\alpha$	2	3.17	4.45	5.81	7.22	8.68	10.17	11.68	13.20	28.61
$\alpha_0$	1	2	3	4	5	6	7	8	9	19

Table 1:  $K = (r + \tilde{r})/2$ ,  $\prod Q_K(e^{-i\xi/2^\ell}) \sim |\xi|^\alpha$  and  $\prod 1/P_K(e^{-2i\xi/2^\ell}) \sim |\xi|^{\alpha_0}$

## 5 Efficiency Issues

In this section, we will compare the C-D-F wavelet expansion and difference wavelet expansion on the following efficiency issues: approximation power and compression ratio.

### 5.1 Approximation power

We approximate  $u \in L^2(R)$  by the projection

$$u^L(x) = \sum_j (u, \phi_j^L) \tilde{\phi}_j^L(x).$$

The approximation power of this expansion was known as the Strang-Fix theorem. A refine result on the estimate of the constant was found first by Sweldens and Piessen [12], then by Unser [13] with sharpest estimation. Unser's theorem is stated as the follows.

**Theorem 5.3 (Unser)** *If  $\tilde{\phi}$  has Riesz basis property and the corresponding  $\tilde{h}(z)$  has  $\tilde{r}$  multiply zeros at  $z = -1$ , then*

$$\|u^L - u\|_{L^2} = C_{\tilde{\phi}} 2^{-\tilde{r}L} \|u^{(\tilde{r})}\|_{L^2} + O(2^{-(\tilde{r}+1)L}) \text{ as } L \rightarrow \infty$$

where

$$C_{\tilde{\phi}} = \frac{1}{\tilde{r}!} \left( \sum_{k \neq 0} |\hat{\tilde{\phi}}^{(\tilde{r})}(2k\pi)|^2 \right)^{1/2}.$$

We demonstrate by Table 2 to show that the difference wavelet has better approximation ability than the C-D-F wavelet.

$r$	$\tilde{r}$	$A$ (C-D-F)	$A$ (Difference)	$r$	$\tilde{r}$	$A$ (C-D-F)	$A$ (Difference)
1	1	0.35	0.35	3	3	2.93	0.21
1	3	0.24	0.08	3	5	8.57	0.23
1	5	0.93	0.09	3	7	76.07	0.60
1	7	8.43	0.24	3	9	1241.35	2.69
1	9	136.57	1.09	4	4	27.84	0.30
2	2	0.68	0.21	4	8	888.39	1.88
2	4	1.23	0.12	5	5	612.81	0.58
2	6	7.66	0.22	5	7	1220.81	1.48
2	8	95.61	0.76	5	9	12302.14	6.63

Table 2: Comparison of the approximation power between C-D-F wavelet and difference wavelet.  $A = C_{\tilde{\phi}} \tilde{r}!$  denotes for the coefficient appeared in the approximation error estimate.

## 5.2 Compression ratio

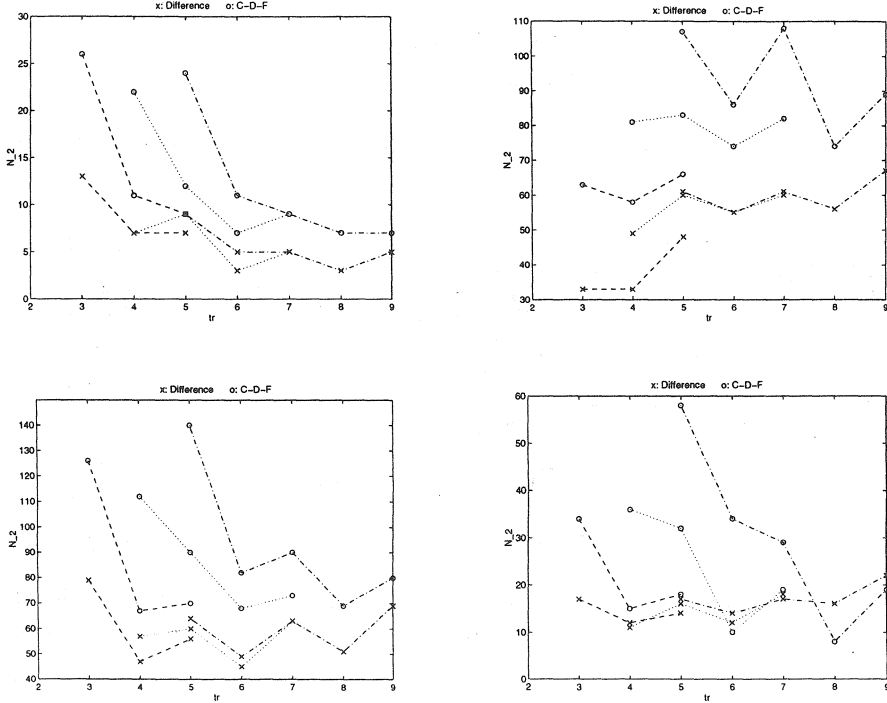
The compression ration is measured by the following quantity. Given a sequence of data  $u_{j,j}^L$ , we first transform expansion  $u^L$  to  $(u^0, w^0, \dots, w^{L-1})$ . Next, we truncate  $(u^0, w^0, \dots, w^{L-1})$  by a threshold  $\delta$  to yield  $(\bar{u}^0, \bar{w}^0, \dots, \bar{w}^{L-1})$ .  $\delta$  is chosen so that the backward transform, say  $\bar{u}^L$ , of the truncated wavelet expansion is with  $\epsilon$  neighborhood of  $u^L$  in  $\ell^2$ , i.e.

$$\|u^L - \bar{u}^L\|_{\ell^2} \leq \epsilon.$$

Let  $N_2(\epsilon, L)$  be the number of nonzero elements in  $(\bar{u}^0, \bar{w}^0, \dots, \bar{w}^{L-1})$ . We perform four tests to compare the compression ratio of the difference wavelet method and the C-D-F method. The four test functions are

$$\begin{aligned} u(x) &= \sin x \\ u(x) &= \chi_{[0,1]} \end{aligned}$$

$$u(x) = \begin{cases} 1/|x| & \text{for } |x| < 1/2 \\ 0 & \text{otherwise.} \end{cases}$$



In Figure 1, (a)-(d), each marked point specifies an  $N_2(\epsilon, L)$  for a run of a given  $(r, \tilde{r})$ . The tolerant error  $\epsilon$  is chosen to be  $10^{-14}$ , and  $L = 10$ . The circle "o" is for the C-D-F method, and the cross "x" is for the difference wavelet method. Marked points of the same method with the same  $K = (r + \tilde{r})/2$  are connected. In all tests, we observe that the difference wavelet method is about twice more efficient than the C-D-F method. The test 4 case (the truncated parabola) shows that the difference wavelet does have good zoom-in property, even its support is infinite.

The efficiency of the difference wavelet method is due to its better regularity and with only slightly bigger "essential support," as compared with the C-D-F method.

### Acknowledgments

This work was completed when the author visited the Department of Mechanical Engineering of the University of California at Berkeley. He would like to express his thanks for their warm hospitality. The author also thanks Mr. Chien-Chang Yen for preparing the figures.

## References

- [1] G. Beylkin, R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms, *Comm. Pure Appl. Math.* **44**, 141-183 (1991).
- [2] C. K. Chui, An Introduction of Wavelets, Academic Press, San Diego, CA (1992).
- [3] C.K Chui and J. Z. Wang, A cardinal spline approach to wavelets, *Proc. Amer. Math. Soc.* **113** 785-793 (1991).
- [4] A. Cohen, I. Daubechies and J. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **45** 485-560 (1992).
- [5] Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **41** 909-996 (1988).
- [6] Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Series in Appl. Math. Vol. 61, SIAM, Philadelphia, PA (1992).
- [7] G. Deslauriers and S. Dubuc, Symmetric iterative interpolation process, *Constr. Approx.*, **5** (1), 49-68 (1989).
- [8] D.L. Donoho, Interpolating wavelet transforms, *Stanford preprint* (1992).
- [9] A. Harten, Discrete multiresolution analysis and generalized wavelets, *J. of Appl. Numer. Math.* **12** 153-193 (1993).
- [10] S. Mallat, Multiresolution approximations and wavelet orthonormal bases of  $L^2(R)$ , *Trans. Amer. Math. Soc.* **315** 69-87 (1989).
- [11] W. Sweldens, The lifting scheme: a custom-design construction of biorthogonal wavelets, *preprint* (1994).
- [12] W. Sweldens and R. Piessens, Quadrature formulae and asymptotic error expansions for wavelet approximations of smooth functions, I, II, *SIAM J. Numer. Anal.* **31** 1240-1264 (1994), *Numer. Math.* **68** 377-401 (1994).
- [13] M. Unser, Approximation power of biorthogonal wavelet expansions, *IEEE Trans. SP.* to appear.
- [14] M. Vetterli and C. Herley, Wavelets and filter banks: theory and design, *IEEE Trans. ASSP* **40** 2207-2232 (1992).