RIMS Kôkyûroku Bessatsu
B56 (2016), 51–77

Orthonormal scaling functions generating fractional Hilbert transforms of an orthonormal wavelet

By

Takeshi MANDAI*

Abstract

The Hilbert transform is an important transform not only in Mathematics, but also in some applications. Since a wavelet function has zero average, the Hilbert transform of it is a good function in many cases. It is well-known that many wavelet functions, especially important ones, can be generated from scaling functions in the framework of multiresolution analysis (MRA). Hence, it is an important problem what is the scaling function from which the Hilbert transform of the wavelet function is generated. We consider two families of unitary operators. One is a family of extensions of the Hilbert transform called fractional Hilbert transforms. The other is a new family of operators which are a kind of modified translation operators. A fractional Hilbert transform of a given orthonormal wavelet (resp. scaling) function is also an orthonormal wavelet (resp. scaling) function, although a fractional Hilbert transform of a scaling function has bad localization in many cases. We show that a modified translation of a scaling function is also a scaling function, and it generates a fractional Hilbert transform of the corresponding wavelet function. Further, we show a good localization property of the modified translation operators. The modified translation operators act on the Meyer scaling functions as the ordinary translation operators. We give a class of scaling functions, on which the modified translation operators act as the ordinary translation operators.

This article is a survey, without proofs, of some results of collaborated research with Ryuichi ASHINO and Akira MORIMOTO at Osaka Kyoiku University, Japan.

Received September 30, 2014.
2010 Mathematics Subject Classification(s): Primary 42C40; Secondary 44A15
Key Words: *orthonormal wavelet*, *scaling function*, *MRA*, *Meyer wavelet*, *fractional Hilbert transform.*
This work was partially supported by JSPS KAKENHI Grant Numbers (C)25400202, (C)24540197, (C)23540135
*Osaka Electro-Communication University, 572-8530, Japan.
e-mail: `mandai@isc.osakc.ac.jp`

§ 1. Introduction

We denote the set of real numbers by $\mathbb{R}$, the space of square integrable functions on $\mathbb{R}$ by $L^2(\mathbb{R})$, the inner product of $f, g \in L^2(\mathbb{R})$ by $\langle f, g\rangle := \int_{\mathbb{R}} f(x)\overline{g(x)}\,dx$ and the norm of f by $\|f\| := \sqrt{\langle f, f\rangle}$. Let us define two unitary operators in $L^2(\mathbb{R})$:

$$\begin{aligned} T_b &: \text{ Translation operator, } b \in \mathbb{R},\ (T_b f)(x) := f(x-b), \\ D_a &: \text{ Dilation operator, } a \in \mathbb{R}_+,\ (D_a f)(x) := a^{-1/2} f(x/a), \end{aligned}$$

where $\mathbb{R}_\pm = \{x \in \mathbb{R} \mid \pm x > 0\}$. For $\psi \in L^2(\mathbb{R})$ and $(j,k) \in \mathbb{Z}^2$, where $\mathbb{Z}$ denotes the set of integers, we set

$$\psi_{j,k}(x) = (D_{2^{-j}} T_k \psi)(x) = 2^{j/2}\psi(2^j x - k). \tag{1.1}$$

A function $\psi \in L^2(\mathbb{R})$ is called an *orthonormal wavelet function*, if $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$. Then, $\psi_{j,k}$, $j, k \in \mathbb{Z}$ are called orthonormal wavelets. In order to construct an orthonormal wavelet function, a system of subspaces called a *multiresolution approximation* or a *multiresolution analysis* (MRA) ([7],[11]) is used, where an orthonormal *scaling function* ϕ generates an orthonormal wavelet function ψ in a fixed manner. Then, we say ψ is *associated with* ϕ. Scaling functions are important not only for the construction of wavelet functions, but also for stepwise decomposition and reconstruction of functions, based on the orthonormal basis $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$.

The Hilbert transform $\mathcal{H}$ ([9], [5] and so on) is a typical example of Calderón-Zygmund operator, which has a rich world full of mathematical results. Although it can be considered in various function spaces, we consider only $L^2(\mathbb{R})$ here. Let $\widehat{f}(\xi)$ be the Fourier transform of f:

$$\widehat{f}(\xi) = (f)^\wedge(\xi) = \mathcal{F}[f](\xi) := \int_{\mathbb{R}} f(x)e^{-i\xi x}\,dx,$$

where the operator $\mathcal{F}$: $f \mapsto \widehat{f}$ can be considered to be a bounded operator from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. The Hilbert transform $\mathcal{H}f$ of $f \in L^2(\mathbb{R})$ is defined by

$$(\mathcal{H}f)^\wedge(\xi) = -i(\operatorname{sgn}\xi)\widehat{f}(\xi), \tag{1.2}$$

where

$$\operatorname{sgn}\xi = \begin{cases} 1, \xi > 0, \\ -1, \xi < 0. \end{cases}$$

The operator $\mathcal{H}$ is a unitary operator of $L^2(\mathbb{R})$.

The Hilbert transform is also important in many applications. The most famous application would be the *analytic signal*. For a real signal (real-valued function) f, the complex signal $\mathcal{A}f = f + i\mathcal{H}f$, which is called the analytic signal of f, has an interesting properties illustrated later in Section 3.

In many applications of wavelets, Hilbert pairs $(\psi, \mathcal{H}\psi)$ of a wavelet function ψ plays an important role. Since $\mathcal{H}$ is a unitary operator which commutes with translations and dilations, if ψ is an orthonormal wavelet function, then $\mathcal{H}\psi$ is also an orthonormal wavelet function. The problem is what is the scaling function with which $\mathcal{H}\psi$ is associated. Let ϕ be a scaling function with which ψ is associated. Although $\mathcal{H}\phi$ is the orthonormal scaling function with which the wavelet function $\mathcal{H}\psi$ is associated, the scaling function $\mathcal{H}\phi$ is usually a very bad function as for the localization, while the wavelet function $\mathcal{H}\psi$ is not. When ψ is a so-called Meyer wavelet, which belongs to the Schwartz class $\mathscr{S}$ and has compactly supported Fourier transform, Toda and Zhang[14, 15] pointed out that $\mathcal{H}\psi$ is the orthonormal wavelet function associated with the scaling function $T_{1/2}\phi$, the half shift of ϕ, which has a good localization as ϕ. This was very unexpected and attractive for us. It is a natural question that what is the wavelet function generated from $T_c\phi$ for $c \neq 1/2$, and whether there is any scaling function with good localization which generates $\mathcal{H}\psi$ in the case of other wavelets than Meyer wavelets.

In this article, we consider two families of translation-invariant unitary operators $\mathcal{H}_c$ and $T_c^\dagger$ $(c \in \mathbb{R})$, where $\mathcal{H}_c$ is a *fractional Hilbert transform* ([10], [5]) with $\mathcal{H}_{1/2} = \mathcal{H}$, and $T_c^\dagger$ is a newly defined operator, a kind of modified translation operator. Let ϕ be an arbitrary orthonormal scaling function, and ψ be the wavelet function associated with ϕ. For every $c \in \mathbb{R}$, we prove that $T_c^\dagger\phi$ is also an orthonormal scaling function, and that $\mathcal{H}_c\psi$ is the wavelet function associated with the scaling function $T_c^\dagger\phi$. Further, we can easily show that $T_c^\dagger f = T_c f$ if $\operatorname{supp}\widehat{f} \subset [-2\pi, 2\pi]$. These clarify the remarkable situation explained above, since $\operatorname{supp}\widehat{\phi} \subset [-2\pi, 2\pi]$ for Meyer scaling functions. We also prove that $T_c^\dagger$ has a good localization property under vanishing moments condition.

In the next section, we give a definition and several examples of orthonormal wavelets. In Section 3, we explain the Hilbert transform, and the analytic signal as an important application. In Section 4, we give a short sketch of a theory of MRA. In Section 5, we present the main problems. In order to give our answers to the problems, we define two families of translation-invariant unitary operators $\mathcal{H}_c$ and $T_c^\dagger$ $(c \in \mathbb{R})$ in Section 6. In Section 7, our answers to the main problems are given. In Section 8, good properties of $T_c^\dagger$ are given. As an extension of Meyer scaling functions, a family of scaling functions ϕ satisfying $T_c^\dagger\phi = T_c\phi$ is given in the final section.

§ 2. Orthonormal Wavelets

If $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$ is an orthonormal basis of $L^2(\mathbb{R})$, then ψ is called an *orthonormal wavelet function* ([8], [16] and so on), which is referred to as a wavelet function for short in this article. As important examples, we give some well-known examples: the Haar wavelet, the Shannon wavelet, the Meyer wavelets, and the Daubechies wavelets.

Example 2.1. (1) The *Haar wavelet.* This would be the oldest orthonormal wavelet function (Figure 1).

$$\psi_{\mathtt{H}}(x) = \begin{cases} 1, & 0 < x < \dfrac{1}{2}, \\ -1, & \dfrac{1}{2} < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

This $\psi_{\mathtt{H}}$ has a compact support, but is discontinuous.

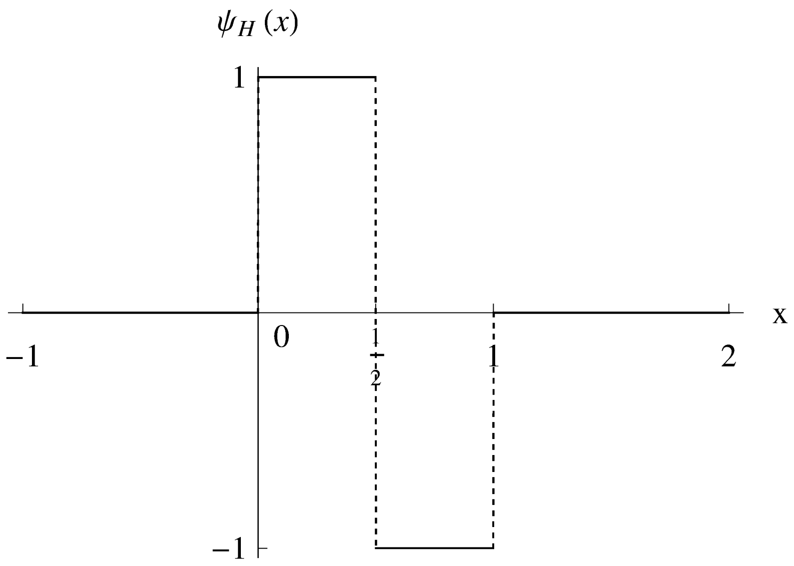

Figure 1. Haar wavelet.

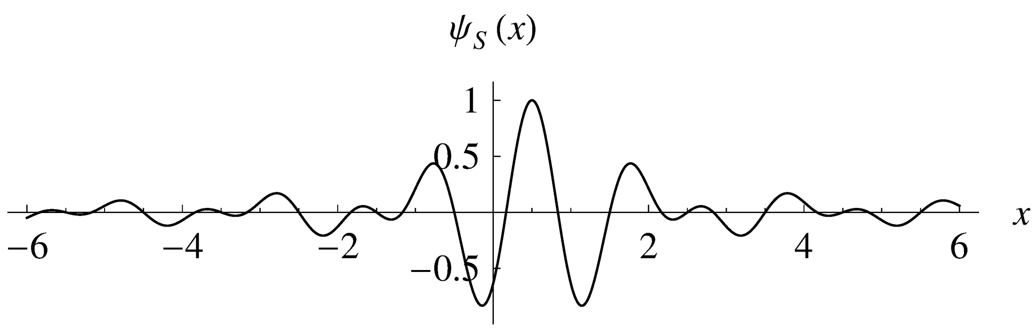

Figure 2. The Shannon wavelet

(2) The *Shannon wavelet.* The function (Figure 2)

$$\psi_{\mathtt{S}}(x) = 2\operatorname{sinc}(2x) - \operatorname{sinc}(x),$$

where $\operatorname{sinc}(x) := \dfrac{\sin \pi x}{\pi x}$ (Figure 9), is a wavelet function called the Shannon wavelet. In this case, $\psi_{\mathtt{S}}(x - 1/2)$ is also a wavelet function, and it is sometimes called the Shannon wavelet instead of $\psi_{\mathtt{S}}(x)$. The Fourier transform of $\psi_{\mathtt{S}}$ has a simple form (Figure 3).

$$\widehat{\psi_{\mathtt{S}}}(\xi) = \begin{cases} 1, & \pi < |\xi| < 2\pi, \\ 0, & \text{otherwise.} \end{cases}$$

This $\psi_{\mathtt{S}}(x)$ is an entire function, but has a bad localization. In fact, $\psi_{\mathtt{S}} \notin L^1(\mathbb{R})$, though

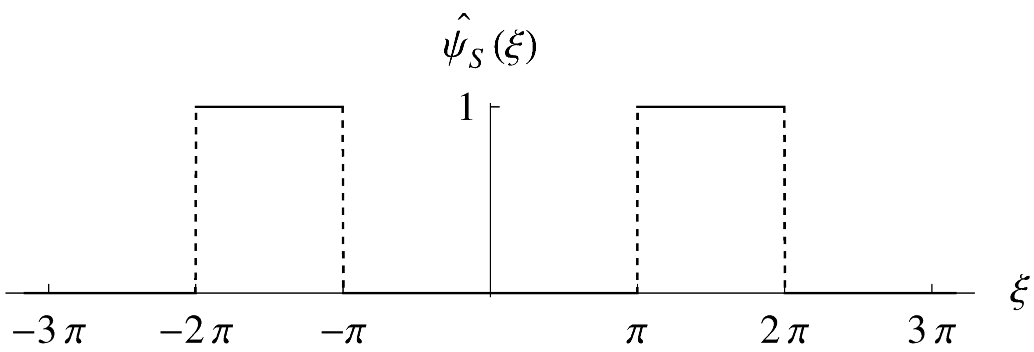

Figure 3. Fourier transform $\widehat{\psi_{\mathtt{S}}}$ of the Shannon wavelet.

$\psi_{\mathtt{S}} \in L^2(\mathbb{R})$.

(3) The (*Lemarié-*)*Meyer wavelets.* These wavelet functions belong to the Schwartz class $\mathscr{S}$, that is, these are of C^∞ class and all the derivatives are rapidly decreasing. As a matter of fact, $\widehat{\psi_{\mathtt{M}}}$ has a compact support, and hence $\psi_{\mathtt{M}}$ is an entire function. It is known that there is no orthonormal wavelet function ψ with exponential decay such that $\psi \in C^\infty(\mathbb{R})$ and all the derivatives are bounded ([7] Corollary 5.5.3). Hence, the Meyer wavelets have a good balance between the smoothness and the localization as wavelet functions.

We explain the Meyer wavelets more precisely. Take a real-valued function $b(\xi)$ of C^∞ class as

$$\begin{gathered} b(\xi) \geq 0, \quad b(-\xi) = b(\xi), \\ \operatorname{supp} b \subset \Big[-\frac{8}{3}\pi, -\frac{2}{3}\pi\Big] \cup \Big[\frac{2}{3}\pi, \frac{8}{3}\pi\Big], \\ b(\pi + \xi) = b\big(2(\pi - \xi)\big) \quad \text{for } |\xi| \leq \frac{\pi}{3}, \\ b(\pi + \xi)^2 + b(\pi - \xi)^2 = 1 \quad \text{for } |\xi| \leq \frac{\pi}{3}, \end{gathered}$$

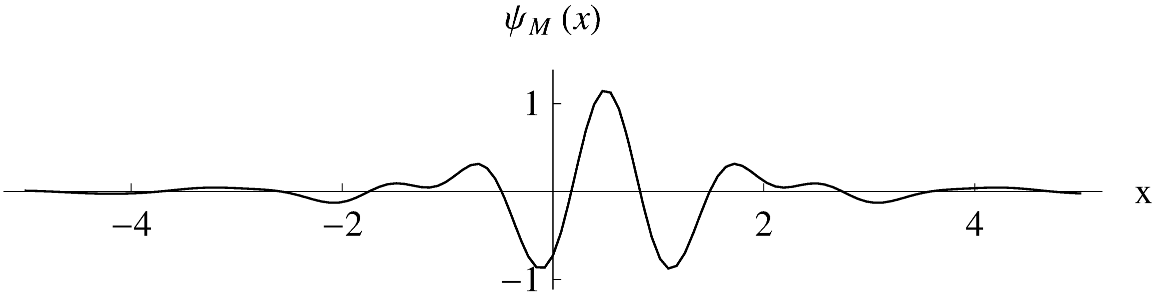

Figure 4. Meyer wavelet $\psi_{\mathtt{M}}$.

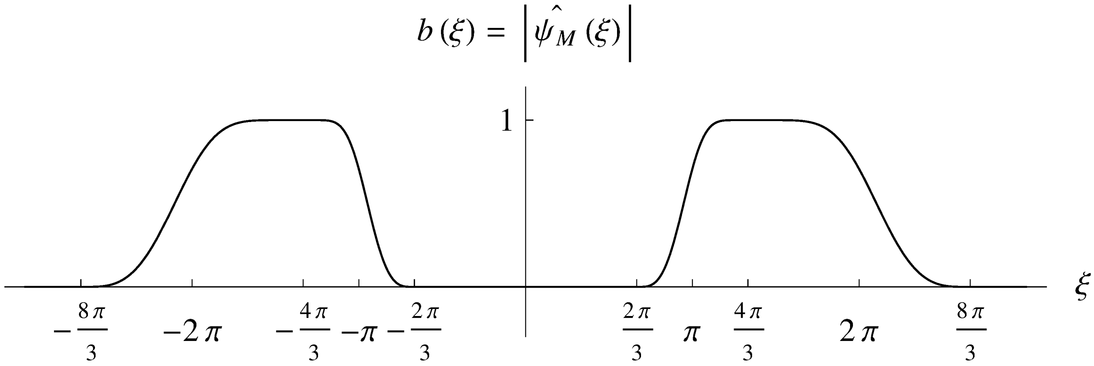

Figure 5. $b(\xi) = |\widehat{\psi_{\mathtt{M}}}(\xi)|$ for a Meyer wavelet.

and define $\psi_{\mathtt{M}}$ by $\widehat{\psi_{\mathtt{M}}}(\xi) := b(\xi)e^{-i\xi/2}$ (Figure 5). There are some freedom of the choice of $b(\xi)$. Sometimes, especially in applications, we take $b(\xi)$ not necessarily of C^∞ class, but only sufficiently smooth (for example [7], [11]), although $\psi \notin \mathscr{S}$ then.

(4) The *Daubechies wavelets.* These are a sequence of orthonormal wavelet functions ${}_N\psi$ with compact supports, $N \in \mathbb{N} = \{\text{positive integers}\}$, which have the following properties.

- $\operatorname{supp} {}_N\psi(x) = [-N+1, N]$.
- ${}_N\psi(x)$ has N vanishing moments, i.e. $\int_{\mathbb{R}} x^j\,{}_N\psi(x)\,dx = 0$ for $j \in \mathbb{Z}$, $0 \le j < N$.
- ${}_N\psi \in C^{\alpha(N)}(\mathbb{R})$ where $\alpha(N) \to \infty$ as $N \to \infty$.
- ${}_1\psi$ is the Haar wavelet function.

§ 3. Hilbert Transform

Although the Hilbert transform $\mathcal{H}$ is defined on many function spaces in several ways, it is simply defined on $L^2(\mathbb{R})$ by (1.2). If f is real-valued, then $\mathcal{H}f$ is also real-

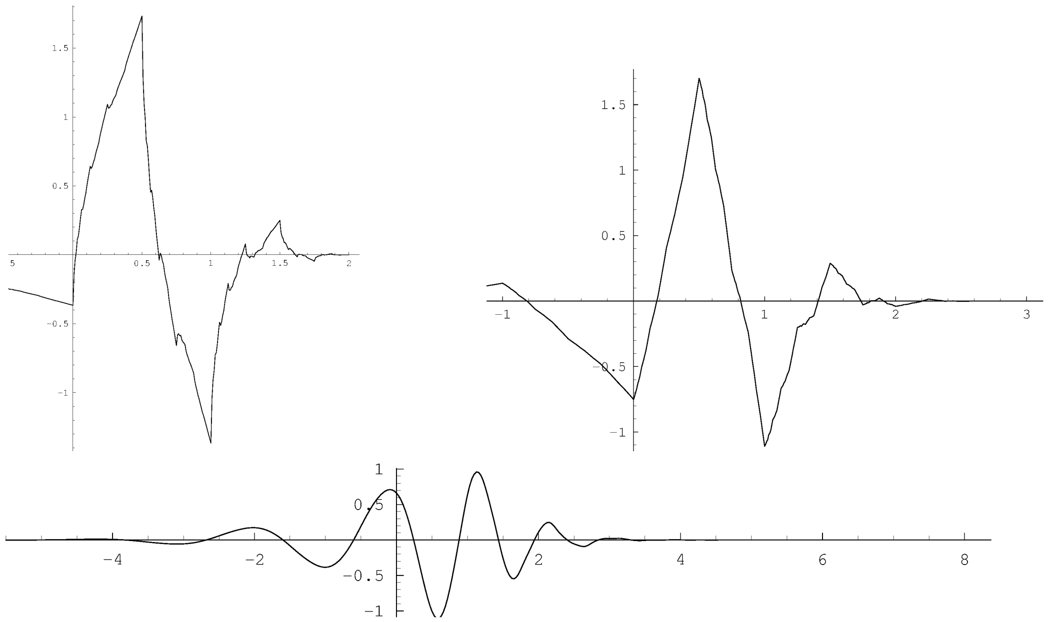

Figure 6. Daubechies wavelet functions: ${}_N\psi$, $N = 2, 3, 8$.

valued and $\mathcal{H}f$ is orthogonal to f: $\langle f, \mathcal{H}f\rangle = 0$. Moreover, $\mathcal{H}$ commutes with T_b for every $b \in \mathbb{R}$ and with D_a for every $a \in \mathbb{R}_+$. Hence, $(\mathcal{H}f)_{j,k} = \mathcal{H}(f_{j,k})$ for every $j, k \in \mathbb{Z}$.

The Hilbert transform is important not only in Mathematics, but also in many applications. A typical application is a so-called analytic signal. For a real signal (real-valued function) f, set $\mathcal{A}_\pm f = f \pm i\mathcal{H}f$. We have $\mathcal{A}_- f = \overline{\mathcal{A}_+ f}$. The complex signal $\mathcal{A}_+ f$ is called the *analytic signal*, or *analytic representation*, of f. Mathematically, $\mathcal{A}_\pm f$ are two times the limits as $\Im z \to \pm 0$ of the Cauchy extension

$$F_\pm(z) = \frac{1}{2\pi i}\int_{\mathbb{R}} \frac{f(x)}{z - x}\, dx, \qquad z \in \mathbb{C}, \pm\Im z > 0$$

on the upper (resp. lower) half plane. Namely,

$$(\mathcal{A}_\pm f)(x) = 2 \lim_{y \to +0} F_\pm(x \pm iy), \qquad x \in \mathbb{R}.$$

As for Fourier transform, $\mathcal{A}_\pm f$ has only positive (resp. negative) frequencies:

$$\widehat{\mathcal{A}_\pm f}(\xi) = 2\chi_{\mathbb{R}_\pm}(\xi)\widehat{f}(\xi), \qquad \xi \in \mathbb{R},$$

where χ_I denotes the characteristic function of I:

$$\chi_I(x) = \begin{cases} 1 & (x \in I) \\ 0 & (x \notin I) \end{cases}.$$

If f is a real signal, the Fourier transform $\widehat{f}$ satisfies $\widehat{f}(-\xi) = \overline{\widehat{f}(\xi)}$, and hence $\mathcal{A}_+ f$ does not lose any information of f. As a matter of fact, we have $f(x) = \Re(\mathcal{A}_\pm f)(x)$.

For many signals in the real world, the absolute value $|\mathcal{A}_+ f| = |\mathcal{A}_- f|$ of $\mathcal{A}_\pm f$ represents a rough variation of $f(x)$, and the graph of $|\mathcal{A}_\pm f(x)|$ can be considered to be an "envelope" of the graph of $f(x)$. Mathematically, it is still a wonder why such a phenomena occurs. There is no rigorous definition of "envelope" used here, and there are only many examples.

An easy example is the case of

$$f(x) = \cos 10x + \cos(10x + ax) = \frac{1}{2}\{e^{i10x} + e^{i(10+a)x} + e^{-i10x} + e^{-i(10+a)x}\}$$

for $a \in (0, 10)$. (Though f does not belong to $L^2(\mathbb{R})$, we can explain the situation very well by considering this function itself, rather than modifying it so that it belongs to $L^2(\mathbb{R})$. It is easy to extend the definition of the Hilbert transform to a class of distributions including $f(x)$. Note that the support of $\widehat{f}$ does not include $\xi = 0$.) In this case, since $\widehat{f}(\xi) = \pi\{\delta(\xi - 10) + \delta(\xi - 10 - a) + \delta(\xi + 10) + \delta(\xi + 10 + a)\}$, we have $\mathcal{A}_+ f(x) = e^{i10x} + e^{i(10+a)x}$. Hence $|\mathcal{A}_+ f(x)| = |1 + e^{iax}| = 2|\cos(ax/2)|$. As is seen in Figure 7, if a is small, then $|\mathcal{A}_+ f(x)|$ seems to represent a rough variation of $f(x)$, and the graph of $|\mathcal{A}_+ f(x)|$ looks like a kind of "envelope" of the graph of $f(x)$. If a is large, the situation is very different, and the graph of $|\mathcal{A}_+ f(x)|$ cannot be seen as an "envelope" of the graph of $f(x)$. In many signals, the situation is near the case of small a, and the analytic signal is used to extract a rough variation of the signal. $|\mathcal{A}_+ f|$ is sometimes called the *amplitude envelope* or *instantaneous amplitude* of f.

The continuous wavelet transform

$$(W_\psi f)(b, a) = \langle f, T_b D_a \psi \rangle = \int_{\mathbb{R}} f(x) \overline{\psi\Big(\frac{x-b}{a}\Big)} a^{-1/2}\, dx, \qquad (b, a) \in \mathbb{R} \times \mathbb{R}_+,$$

with the wavelet function ψ, is very compatible with the analytic signal. In fact, we have the following ([1]):

$$(W_\psi \mathcal{A}_\pm f)(b, a) = (W_{\mathcal{A}_\pm \psi} f)(b, a) = \frac{1}{2}(W_{\mathcal{A}_\pm \psi} \mathcal{A}_\pm f)(b, a) = \mathcal{A}_\pm\Big((W_\psi f)(\cdot, a)\Big)(b) \tag{3.1}$$

for $(b, a) \in \mathbb{R} \times \mathbb{R}_+$. If we do not consider the inverse of W_ψ, then we need only $\psi, f \in L^2(\mathbb{R})$ for the definition of $W_\psi f$ and the validity of (3.1), although we imposed further assumptions on ψ in [1]. The property (3.1) is an advantage of the continuous wavelet transform compared to the windowed Fourier transform (or short-time Fourier transform). By this property, we can get the wavelet transforms of the analytic signals of various original signals, by computing once the analytic signal of ψ, without computing each analytic signal of the original signals.

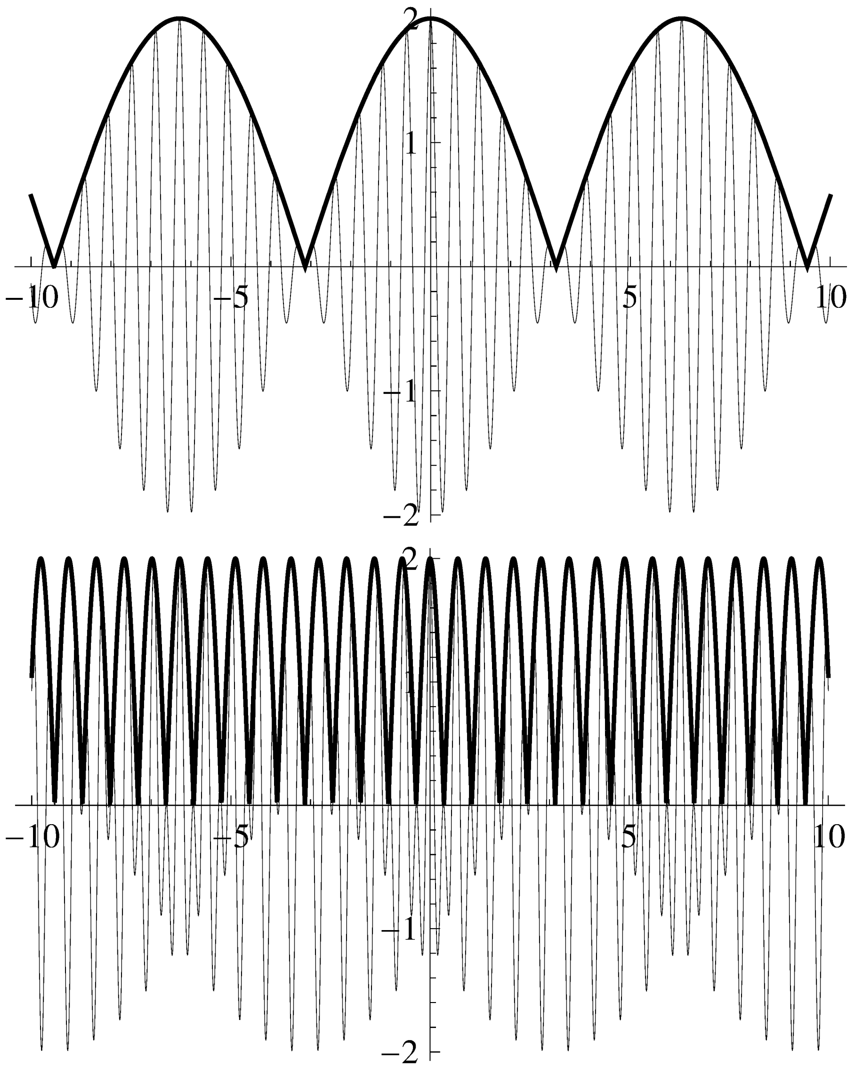

Figure 7. graphs of f(thin) and $|\mathcal{A}_{\pm}f|$(thick), $a = 1$ (above) and $a = 9$ (below).

A pair $(f, \mathcal{H}f)$ of a function (a signal) and its Hilbert transform are often useful ([13], [2] and so on). Chaudhury-Unser[6] investigated several properties of $\mathcal{H}\psi$ for a wavelet function ψ.

§ 4. MRA

In order to construct orthonormal wavelet functions systematically, a concept called multiresolution analysis (MRA) was developed.

In order to explain the idea of MRA, we begin with the situation where we have already a wavelet function ψ. Set

$$W_j = \overline{\mathrm{Span}}\{\psi_{j,k}\}_{k\in\mathbb{Z}}, \tag{4.1}$$

where $\overline{\mathrm{Span}}F$ denotes the closed subspace of $L^2(\mathbb{R})$ spanned by a set F. We can consider this space as the space representing the variation of the scale level j. The sequence $\{W_j\}_{j\in\mathbb{Z}}$ of closed subspaces satisfies the following properties.

- $f \in W_j \iff f(2\cdot) \in W_{j+1}$.
- $W_j \perp W_k$ $(j \neq k)$.
- $L^2(\mathbb{R}) = \overline{\bigoplus_{j\in\mathbb{Z}} W_j}$, where $\oplus$ denotes the orthogonal direct sum.

Next, we set

$$V_j = \overline{\bigoplus_{l<j} W_l}. \tag{4.2}$$

We can consider this space as the space representing the variation coarser than the level j. The sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces satisfies $V_{j+1} = V_j \oplus W_j$ and the following.

(i) $V_j \subset V_{j+1}$, $j \in \mathbb{Z}$.

(ii) $f \in V_j \iff f(2\cdot) \in V_{j+1}$.

(iii) $\cap_{j\in\mathbb{Z}} V_j = \{0\}$.

(iv) $\overline{\cup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$.

Moreover, V_0 is shift invariant: $f \in V_0 \implies T_k f \in V_0$, $k \in \mathbb{Z}$. Note that W_l for $l < 0$ is not necessarily shift-invariant, and the shift-invariance of V_0 does not straightly follows from (4.2). The shift-invariance of W_l for $l \geq 0$ is straightforward from the definition. Hence, $U_0 = \overline{\oplus_{l\geq 0} W_l}$ is also shift-invariant. Since $V_0 = U_0^{\perp}$, the orthogonal complement of U_0, and since the orthogonal complement of a shift-invariant subspace is also shift-invariant, we have the shift-invariance of V_0.

The idea of MRA is the reverse of this, that is, we construct ψ from a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$. Here is the definition of MRA.

Definition 4.1. If V_j, $j \in \mathbb{Z}$, are closed linear subspaces of $L^2(\mathbb{R})$ satisfying the conditions (i)–(v), where (i)–(iv) is given above and (v) is given below, then the sequence $\{V_j\}_{j\in\mathbb{Z}}$ is called a *multiresolution analysis* (*MRA*).

(v) there exists a function $\phi \in V_0$ such that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_0.

The condition (v) is stronger than the condition that V_0 is shift-invariant.

By the conditions (ii) and (v), we have $V_j = \overline{\mathrm{Span}}\{\phi_{j,k}\}_{k\in\mathbb{Z}}$. The function ϕ in the condition (v) is very important and called an orthonormal *scaling function*, which is referred to as a scaling function for short in this article. In this article, we do not assume any further conditions to ϕ, unless otherwise specified. In particular, it can be that $\phi \notin L^1(\mathbb{R})$, and the familiar condition $\int_{\mathbb{R}} \phi(x)\, dx = 1$ or $\widehat{\phi}(0) = 1$ is not assumed, may be even meaningless.

If ϕ is a scaling function, then by $V_0 \subset V_1$, there exists a unique sequence $\{h_k\}_{k\in\mathbb{Z}}$ such that

$$\phi(x) = \sum_{k\in\mathbb{Z}} h_k \phi_{1,k}(x) = \sum_{k\in\mathbb{Z}} \sqrt{2} h_k \phi(2x-k) \quad \text{a.e. on } \mathbb{R}. \tag{4.3}$$

If we set $m_0(\xi) = \sum_{k\in\mathbb{Z}} \frac{h_k}{\sqrt{2}} e^{-ik\xi}$, then we have

$$\widehat{\phi}(2\xi) = m_0(\xi)\widehat{\phi}(\xi) \quad \text{a.e. on } \mathbb{R}. \tag{4.4}$$

Each of these equations is called the *two scale equation*, and $m_0(\xi)$ is called the *low-pass filter* associated with ϕ.

It is well-known that from a scaling function ϕ we can construct a wavelet function ψ as follows. (See, for example, [8], [16].)

Theorem 4.2. *Let ϕ be a scaling function and m_0 be the low-pass filter. Let $\nu \in L^1_{loc}(\mathbb{R})$ be a 2π-periodic function such that $|\nu(\xi)| = 1$ a.e. We set*

$$m_1(\xi) = e^{-i\xi}\, \overline{m_0(\xi+\pi)}\, \nu(2\xi). \tag{4.5}$$

If we define ψ by

$$\widehat{\psi}(\xi) = m_1(\xi/2)\, \widehat{\phi}(\xi/2), \tag{4.6}$$

then ψ is a wavelet function.

If we expand m_1 as $m_1(\xi) = \sum_{k\in\mathbb{Z}} \frac{g_k}{\sqrt{2}} e^{-ik\xi}$, then we have

$$\psi(x) = \sum_{k\in\mathbb{Z}} g_k \phi_{1,k}(x) = \sum_{k\in\mathbb{Z}} \sqrt{2} g_k \phi(2x-k) \quad \text{a.e. on } \mathbb{R}. \tag{4.7}$$

This is called the *wavelet equation* and the 2π-periodic function m_1 is called the *high-pass filter*. These m_1 and ψ are said to be *associated with* ϕ. There are many choices of ν. If we take $\nu(\xi) = 1$, then we say that m_1 and ψ are *naturally* associated with ϕ:

$$\widehat{\psi}(\xi) = e^{-i\xi/2}\, \overline{m_0(\xi/2+\pi)}\, \widehat{\phi}(\xi/2). \tag{4.8}$$

Example 4.3. (1) (Haar) Let $\phi_{\mathtt{H}} = \chi_{[0,1)}$ (Figure 8). In this case, $V_j := \{ f \in L^2(\mathbb{R}) \mid \text{ constant on } \big[\frac{k}{2^j}, \frac{k+1}{2^j}\big)(k \in \mathbb{Z}) \}$. Since $\phi_{\mathtt{H}}(x) = \phi_{\mathtt{H}}(2x) + \phi_{\mathtt{H}}(2x-1)$, we have $m_0(\xi) = \dfrac{1+e^{-i\xi}}{2}$. If we take $\nu(\xi) = -1$, then $m_1(\xi) = \dfrac{1-e^{-i\xi}}{2}$, and we get the Haar wavelet function $\psi_{\mathtt{H}}$. If we take $\nu = 1$, then we have a wavelet function with the different sign, which is the naturally associated wavelet function.

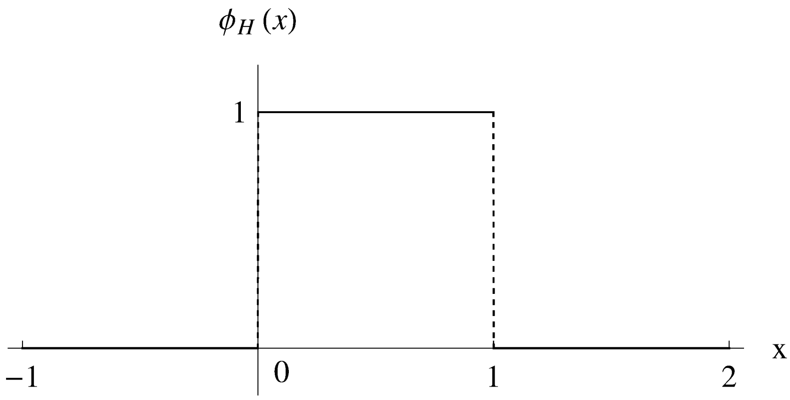

Figure 8. Haar scaling function.

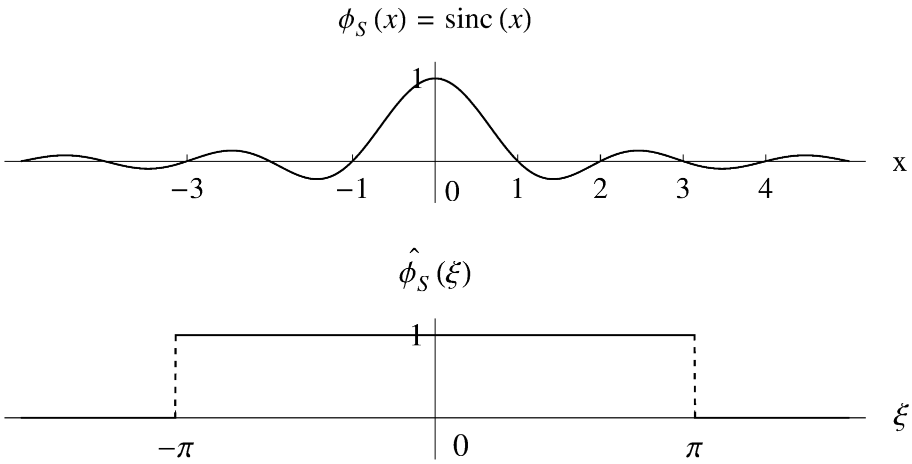

Figure 9. Shannon scaling function $\phi_{\mathtt{S}}(x) = \mathrm{sinc}(x)$ and its Fourier transform $\widehat{\phi_{\mathtt{S}}}(\xi)$.

(2) (Shannon) Let $\phi_{\mathtt{S}}(x) = \mathrm{sinc}(x)$, that is, $\widehat{\phi_{\mathtt{S}}}(\xi) = \chi_{[-\pi,\pi]}(\xi)$ (Figure 9). In this case,

$$V_j := \{ f \in L^2(\mathbb{R}) \mid \operatorname{supp} \widehat{f} \subset [-2^j\pi, 2^j\pi] \}.$$

We have $m_0(\xi) = \chi_{[-\pi/2,\pi/2]}(\xi)$ for $|\xi| \le \pi$, that is, $m_0(\xi) = \sum_{k\in\mathbb{Z}} \chi_{[-\pi/2,\pi/2]}(\xi + 2k\pi) = \chi_S(\xi)$, where $S = \bigcup_{k\in\mathbb{Z}}[-\pi/2 + 2k\pi, \pi/2 + 2k\pi]$ (Figure 10 Left). In this

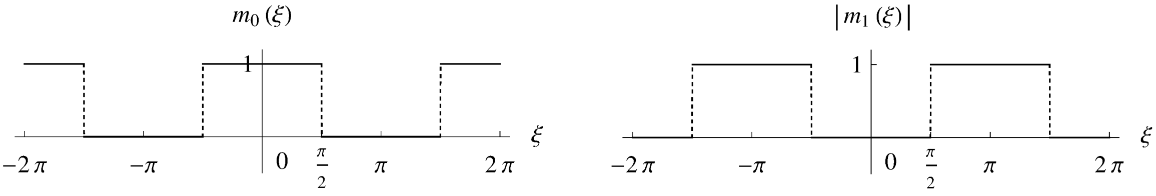

Figure 10. $m_0(\xi)$ and $|m_1(\xi)|$ for Shannon wavelet.

case, the naturally associated wavelet function is $\psi_{\mathtt{S}}(x-1/2)$ in Example 2.1 (2). In Shannon's case, by taking a suitable $\nu(\xi)$, we can omit the factor $e^{-i\xi}$ in the definition of $m_1(\xi)$, and can take $m_1(\xi) = m_0(\xi+\pi)$, which is real-valued. This leads to the Shannon wavelet function $\psi_{\mathtt{S}}(x)$ in Example 2.1 (2). It is a special property of the Shannon wavelet (and so-called MFS wavelets) that both $\psi_{\mathtt{S}}(x)$ and $\psi_{\mathtt{S}}(x-1/2)$ are wavelet functions. Usually, if ψ is a wavelet function, then $\psi(x-1/2)$ is not necessarily a wavelet function, while each shift $\psi(x-k)$ $(k \in \mathbb{Z})$ is a wavelet function.

(3) (Meyer) We can get scaling functions by smoothing $\widehat{\phi_{\mathtt{S}}}$, which leads to the Meyer scaling functions. Let $\phi_{\mathtt{M}}$ be a function satisfying the following conditions (Figure 11).

- $\widehat{\phi_{\mathtt{M}}} \in C^\infty(\mathbb{R})$, $\widehat{\phi_{\mathtt{M}}} \geq 0$, $\widehat{\phi_{\mathtt{M}}}$ is an even function.
- $\operatorname{supp}\widehat{\phi_{\mathtt{M}}} \subset \big[-\frac{4}{3}\pi, \frac{4}{3}\pi\big]$.
- $\widehat{\phi_{\mathtt{M}}}(\xi) = 1$ for $|\xi| \leq \frac{2}{3}\pi$.
- $|\widehat{\phi_{\mathtt{M}}}(\xi+\pi)|^2 + |\widehat{\phi_{\mathtt{M}}}(\xi-\pi)|^2 = 1$ for $|\xi| \leq \frac{\pi}{3}$.

Then, $\phi_{\mathtt{M}}$ is a scaling function and $m_0(\xi) = \widehat{\phi_{\mathtt{M}}}(2\xi)$ for $|\xi| \leq \pi$, that is, $m_0(\xi) = \sum_{k\in\mathbb{Z}} \widehat{\phi_{\mathtt{M}}}(2\xi+4k\pi)$ (Figure 12, Left).

Further, by taking $\nu(\xi) \equiv 1$, we get

$$m_1(\xi) = e^{-i\xi}\widehat{\phi_{\mathtt{M}}}(2\xi+2\pi) \quad \text{for } -2\pi \leq \xi \leq 0, \qquad \text{(Figure 12, Right)}$$
$$\widehat{\psi_{\mathtt{M}}}(\xi) = e^{-i\xi/2}\{\widehat{\phi_{\mathtt{M}}}(\xi+2\pi) + \widehat{\phi_{\mathtt{M}}}(\xi-2\pi)\}\widehat{\phi_{\mathtt{M}}}(\xi/2).$$

This $\psi_{\mathtt{M}}$ is a Meyer wavelet in Example 2.1 (3).

(4) (Daubechies) It is not easy to describe how to construct ϕ ([7]). We have a sequence ${}_N\phi$, $N \in \mathbb{N}$, of scaling functions with compact supports corresponding to ${}_N\psi$. We can construct ${}_N\phi$ satisfying the following.

- $\operatorname{supp}{}_N\phi(x) = [0, 2N-1]$.

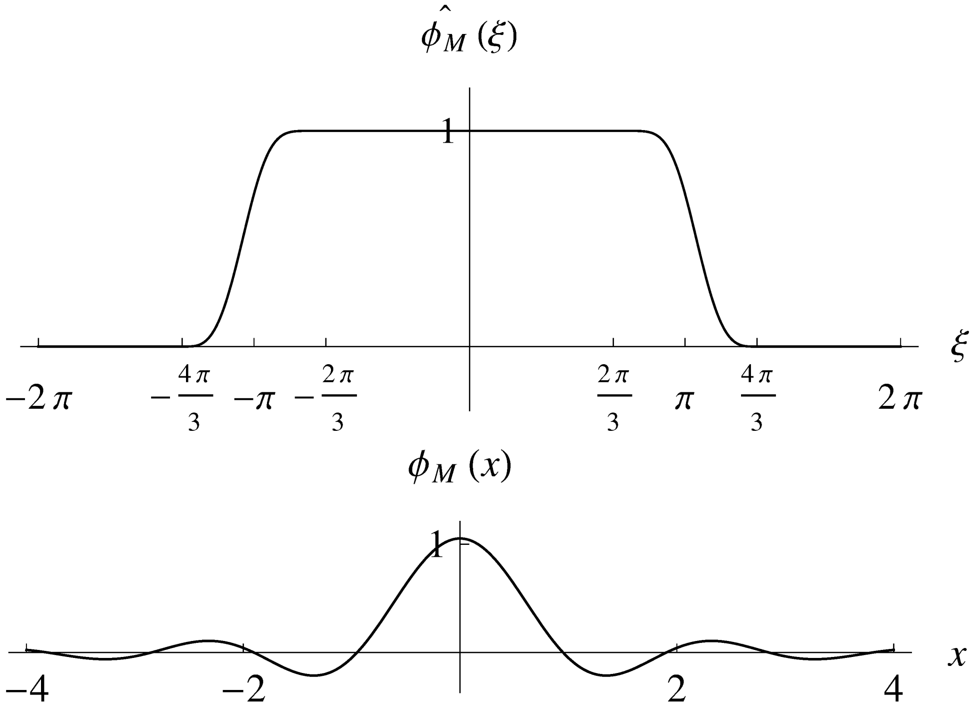

Figure 11. $\widehat{\phi_{\mathtt{M}}}(\xi)$ and $\phi_{\mathtt{M}}(x)$ for a Meyer wavelet.

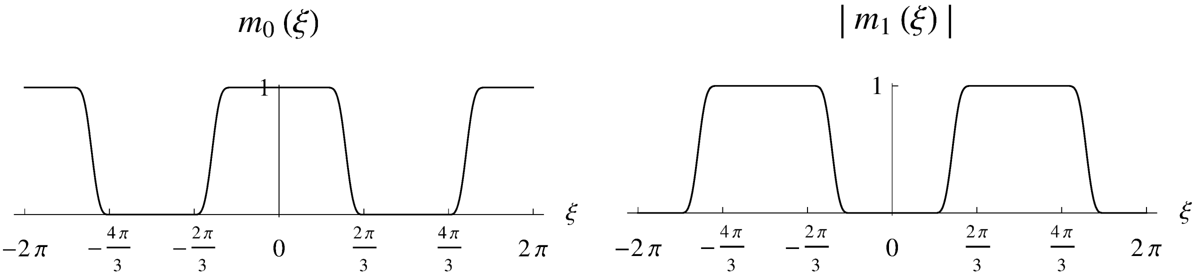

Figure 12. $m_0(\xi)$ and $|m_1(\xi)|$ for a Meyer wavelet.

- ${}_N\phi \in C^{\alpha(N)}(\mathbb{R})$ where $\alpha(N) \to \infty$ as $N \to \infty$.
- ${}_1\phi$ is the Haar scaling function.

The wavelet functions in Example 2.1 (4) are constructed by taking $\nu(\xi) \equiv -1$.

We end this section by giving well-known conditions for a function to be a scaling function.

Theorem 4.4. *Let $\phi \in L^2(\mathbb{R})$. Then, ϕ is a scaling function if and only if the following three conditions hold* ([8] Chapter 7, Theorem 5.2).

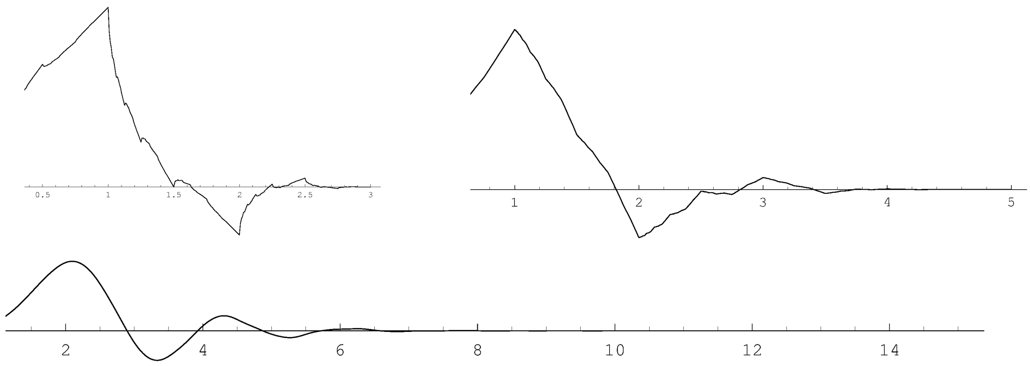

Figure 13. Daubechies scaling functions: ${}_N\phi$, $N = 2, 3, 8$.

(A1) *The equality*

$$\sum_{k\in\mathbb{Z}} |\widehat{\phi}(\xi + 2k\pi)|^2 = 1 \qquad \text{a.e. on } \mathbb{R} \tag{4.9}$$

is satisfied. This condition is equivalent to that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ *is an orthonormal system.*

(A2) *There exists a* 2π*-periodic function* $m_0(\xi)$ *such that* $\widehat{\phi}(2\xi) = m_0(\xi)\widehat{\phi}(\xi)$ *a.e. on* $\mathbb{R}$*. This condition is equivalent to that there exists a sequence* $\{h_k\}_{k\in\mathbb{Z}}$ *such that* (4.3) *holds.*

(A3) $\lim_{j\to\infty} |\widehat{\phi}(2^{-j}\xi)| = 1$ *a.e. on* $\mathbb{R}$.

§ 5. Main Problems

In the case of Meyer wavelets, Toda-Zhang[14, 15] pointed out the essential part of the following theorem, which shows that the Hilbert transform of ψ is associated with $T_{1/2}\phi$.

Theorem 5.1. *Let* ϕ *be a Meyer scaling function and* ψ *be the wavelet function naturally associated with* ϕ*. Fix arbitrary* $b \in \mathbb{R}$*, and set* $\phi_b := T_b\phi$*. Then we have the following.*

(1) ϕ_b *is also a scaling function.*
(2) *If* ψ_b *is the wavelet function naturally associated with* ϕ_b*, then the Hilbert transform* $\mathcal{H}\psi_b$ *is the wavelet function naturally associated with* $T_{1/2}\phi_b = \phi_{b+1/2}$.

The statement (1) had already been well-known in the field of wavelets. (2) was very unexpected and attractive to us. It is very natural to ask the following questions.

Main Questions:

[Q1] What happens for $T_c\phi_b$ with $c \neq 1/2$?

[Q2] From which characteristics of the Meyer scaling functions, do the properties described in the theorem come?

[Q3] What happens for other wavelets than the Meyer wavelets?

In order to give our answers, we define two families of unitary operators $\mathcal{H}_c$ and $T_c^\dagger$, $c \in \mathbb{R}$, in the next section.

§ 6. Unitary Operators $\mathcal{H}_c$ and $T_c^\dagger$

In this section, we define two families of unitary operators $\mathcal{H}_c$ and $T_c^\dagger$, $c \in \mathbb{R}$. The operators $\mathcal{H}_c$ are extensions of the Hilbert transform, called *fractional Hilbert transforms* ([10], [5] and so on).

Definition 6.1. We define unitary operators $\mathcal{H}_c$ on $L^2(\mathbb{R})$ by

$$\mathcal{H}_c = (\cos c\pi)I + (\sin c\pi)\mathcal{H}, \ c \in \mathbb{R}, \tag{6.1}$$

where I is the identity operator. In other words,

$$(\mathcal{H}_c f)^\wedge(\xi) = \{\cos c\pi - i(\sin c\pi)\operatorname{sgn}\xi\}\widehat{f}(\xi) = e^{-ic\pi(\operatorname{sgn}\xi)}\widehat{f}(\xi). \tag{6.2}$$

We have $\mathcal{H}_{1/2} = \mathcal{H}$, and $\mathcal{H}_c$ is called a *fractional Hilbert transform.* Here, we use a different parametrization from the definition in [5] for the compatibility with the other family of operators $T_c^\dagger$.

If f is real-valued, then $\mathcal{H}_c f$ is also real-valued. Further, we have

$$\langle f, \mathcal{H}_c f\rangle = (\cos c\pi)\|f\|^2, \tag{6.3}$$

which means that the "angle" between f and $\mathcal{H}_c f$ is $c\pi$.

The family $\{\mathcal{H}_c\}_{c\in\mathbb{R}}$ constitutes a one-parameter group of unitary operators: $\mathcal{H}_c\mathcal{H}_d = \mathcal{H}_{c+d}$, $\mathcal{H}_0 = I$. Further, we have $\mathcal{H}_{c+1} = -\mathcal{H}_c$, $\mathcal{H}_{c+2} = \mathcal{H}_c$, $\mathcal{H}_1 = -I$, $\mathcal{H}_c^* = \mathcal{H}_c^{-1} = \mathcal{H}_{-c}$, where U^* denotes the adjoint operator of U.

We also have the commutativity with translations and dilations:

$$\mathcal{H}_c T_b = T_b\mathcal{H}_c, \ \mathcal{H}_c D_a = D_a\mathcal{H}_c \ \text{ for } \ b, c \in \mathbb{R}, a \in \mathbb{R}_+. \tag{6.4}$$

In particular, $\mathcal{H}_c(f_{j,k}) = (\mathcal{H}_c f)_{j,k}$, $j, k \in \mathbb{Z}$.

The unitary operators $\mathcal{H}_c$ are natural operators in the sense of the following proposition. A limited version was given in [5, Theorem 3.1], where the domain of the operators consists of only real-valued functions.

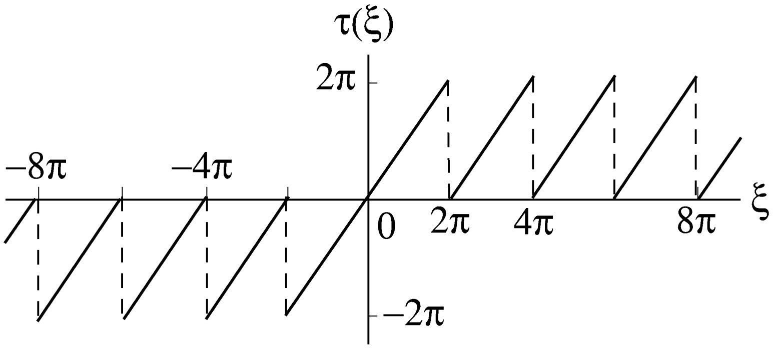

Figure 14. $\tau(\xi)$.

Proposition 6.2 ([3]). *Let U be a unitary operator which is commutative with T_b, D_a for every $b \in \mathbb{R}$, $a \in \mathbb{R}_+$. Then, we have the following.*
(1) *There exist constants $\theta, c \in \mathbb{R}$ such that $U = e^{i\theta}\mathcal{H}_c$.*
(2) *If further U maps real-valued functions to real-valued functions, then there exists $c \in \mathbb{R}$ such that $U = \mathcal{H}_c$.*
(3) *Moreover, if $\langle Uf, f\rangle = 0$ for every real-valued f, then $U = \pm\mathcal{H}_{1/2} = \pm\mathcal{H}$.*

Next, let us define the unitary operators $T_c^\dagger$, a kind of modified translation operators.

Definition 6.3. Set

$$\tau(\xi) = \left(|\xi| - 2\pi\left\lfloor\frac{|\xi|}{2\pi}\right\rfloor\right)\operatorname{sgn}\xi = \xi - 2\pi(\operatorname{sgn}\xi)\left\lfloor\frac{|\xi|}{2\pi}\right\rfloor \quad \text{(Figure 14)},$$

where $\lfloor x\rfloor = \max\{n \in \mathbb{Z} \mid n \le x\}$. We define unitary operators $T_c^\dagger$, $c \in \mathbb{R}$, by $(T_c^\dagger f)^\wedge(\xi) = e^{-ic\tau(\xi)}\widehat{f}(\xi)$.

If f is real-valued, then $T_c^\dagger f$ is also real-valued. The family $\{T_c^\dagger\}_{c\in\mathbb{R}}$ constitutes a one-parameter group of unitary operators: $T_c^\dagger T_d^\dagger = T_{c+d}^\dagger$, $T_0^\dagger = I$. Further, $T_c^\dagger$ are commutative with the translations (but not with the dilations): $T_b T_c^\dagger = T_c^\dagger T_b$, $b, c \in \mathbb{R}$.

Remark. (1) If $c = k$ is an integer, then $e^{-ik\tau(\xi)} = e^{-ik\xi}$, and hence $T_k^\dagger$ is just the translation: $T_k^\dagger = T_k$, $k \in \mathbb{Z}$.
(2) If $\operatorname{supp}\widehat{f} \subset [-2\pi, 2\pi]$, then $T_c^\dagger f = T_c f$, $c \in \mathbb{R}$. So, in a sense, $T_c^\dagger$ is the translation in a low frequency domain.
(3) In the signal processing community, filter design is important. Selesnick[12] designed a low-pass filter corresponding to $\mathcal{H}\psi$. This low-pass filter turned out to be the low-pass filter associated with the scaling function $T_{1/2}^\dagger\phi$.

At the end of the next section, we give several graphs of $T_{1/2}^\dagger\phi$ and related functions.

§ 7. Our answers to Main Problems

In this section, we consider general scaling functions. We assume the following.

AssumptionF ϕ is a scaling function, and ψ is the wavelet function naturally associated with ϕ.

By the commutativity (6.4), the unitary operator $\mathcal{H}_c$ preserves the MRA structure, and hence we have the following.

Proposition 7.1. *For every $c \in \mathbb{R}$, we have the followings.*
(1) *$\mathcal{H}_c\phi$ is a scaling function.*
(2) *$\mathcal{H}_c\psi$ is the wavelet function naturally associated with $\mathcal{H}_c\phi$.*

Unfortunately, $\mathcal{H}_c\phi$ has bad localization in general. In particular, if $\phi \in L^1(\mathbb{R})$ and $c \notin \mathbb{Z}$, then $\mathcal{H}_c\phi \notin L^1(\mathbb{R})$. In fact, $\widehat{\phi}$ is continuous and $\widehat{\phi}(0) \neq 0$, hence $\widehat{\mathcal{H}_c\phi}(\xi)$ has a jump at $\xi = 0$. In Figures 15, 16 and 17, the graphs of $\mathcal{H}_{1/2}\phi = \mathcal{H}\phi$ for several ϕ are illustrated.

The following is our answer to the main problems. Note that $T_c\phi$ ($c \notin \mathbb{Z}$) is not necessarily a scaling function.

Theorem 7.2 ([3]). *For every $c \in \mathbb{R}$, we have the following.*
(1) *$T_c^\dagger\phi$ is a scaling function.*
(2) *$\mathcal{H}_c\psi$ is the wavelet function naturally associated with $T_c^\dagger\phi$.*

By the property of $T_c^\dagger$ stated in Remark (2) after Definition 6.3, we have the following corollary.

Corollary 7.3. *If $\operatorname{supp}\widehat{\phi} \subset [-2\pi, 2\pi]$, then $T_c\phi$ is a scaling function. Further, $\mathcal{H}_c\psi$ is the wavelet function naturally associated with $T_c\phi$.*

The scaling function $T_c^\dagger\phi$ does not have so bad localization in many cases. In particular, if ϕ is a Meyer scaling function, then $T_c^\dagger\phi = T_c\phi \in \mathscr{S}$. We give more properties of $T_c^\dagger$ in Section 8.

This theorem gives answers to the main questions in Section 5.

[Ans1] In the case of Meyer wavelets, $\mathcal{H}_c\psi_b$ is naturally associated with $T_c\phi_b = T_{c+b}\phi$, $c, b \in \mathbb{R}$. (Now we have $\psi_b = \mathcal{H}_b\psi$ and hence $\mathcal{H}_c\psi_b = \mathcal{H}_{b+c}\psi$.)

[Ans2] $\operatorname{supp}\widehat{\phi} \subset [-2\pi, 2\pi]$ implies that $T_c\phi$ is a scaling function, and $\mathcal{H}_c\psi$ is associated with $T_c\phi$. (Corollary 7.3.)

[Ans3] In general, $\mathcal{H}_c\psi$ is naturally associated with $T_c^\dagger\phi$. (Theorem 7.2.)

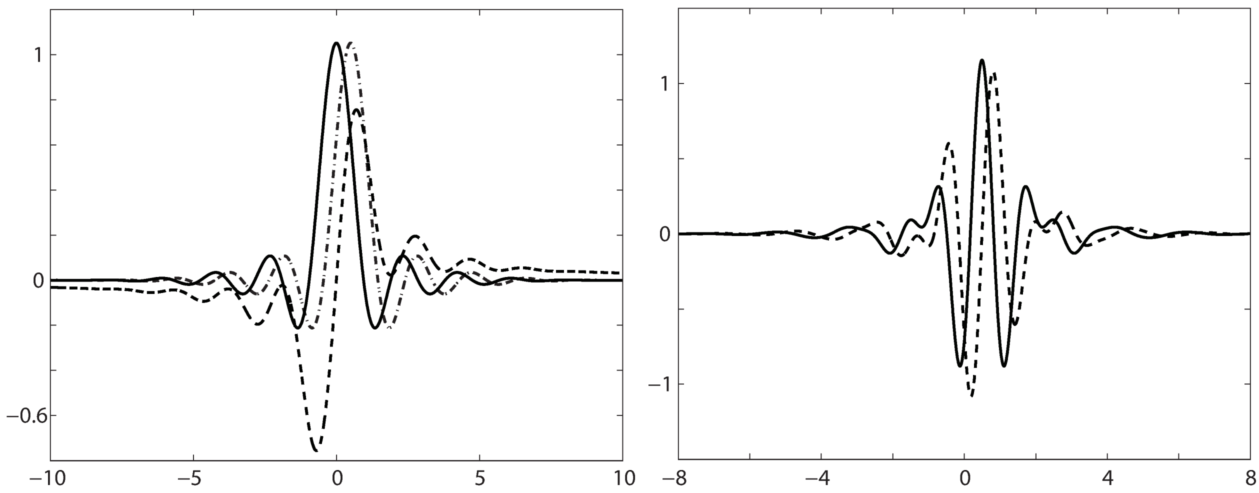

Figure 15. Case of Meyer wavelets. Left: ϕ (solid), $\mathcal{H}\phi$ (broken), $T^{\dagger}_{1/2}\phi$ (dash-dot). Right: ψ (solid), $\mathcal{H}\psi$ (broken)

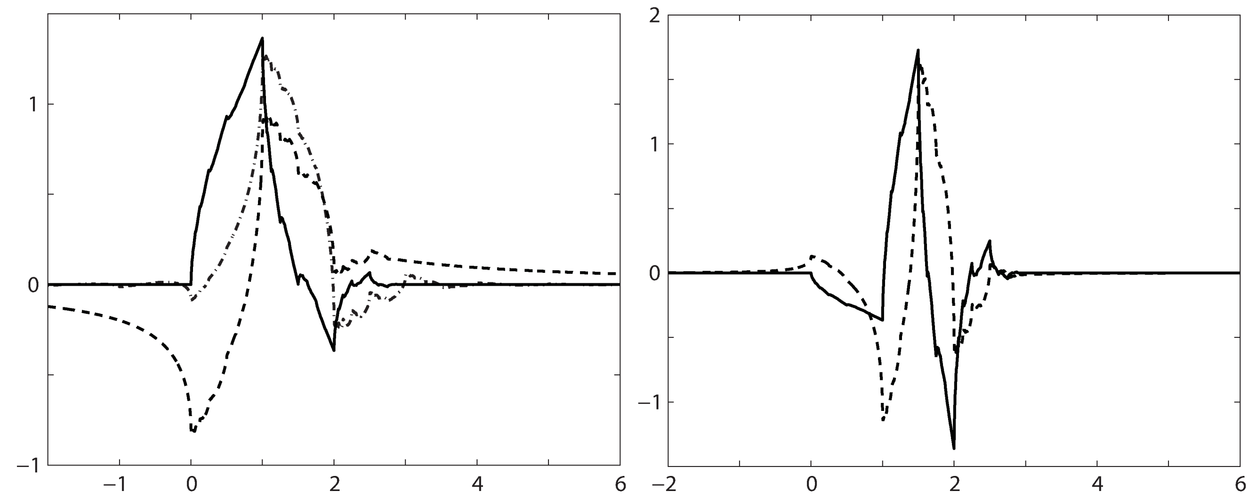

Figure 16. Case of Daubechies wavelets $N = 2$. Left: ${}_2\phi$ (solid), $\mathcal{H}_2\phi$ (broken), $T^{\dagger}_{1/2}\, {}_2\phi$ (dash-dot). Right: ${}_2\psi$ (solid), $\mathcal{H}_2\psi$ (broken)

In Figures 15–17, we show the graphs of ϕ, $\mathcal{H}\phi = \mathcal{H}_{1/2}\phi$, $T^{\dagger}_{1/2}\phi$, ψ, and $\mathcal{H}\psi = \mathcal{H}_{1/2}\psi$ for the case of the Meyer wavelets and the Daubechies wavelets. In the case of Meyer wavelets, we have $T^{\dagger}_{1/2}\phi = T_{1/2}\phi$. In the case of Daubechies wavelets, $T^{\dagger}_{1/2}\,{}_N\phi$ approaches $T_{1/2}\,{}_N\phi$ as $N \to \infty$, since $\widehat{{}_N\phi}$ concentrate in $[-2\pi, 2\pi]$ as seen in Figure 18. In Figure 17, $T^{\dagger}_{1/2}\,{}_8\phi$ looks as if $T_{1/2}\,{}_8\phi$. In both cases, the scaling functions $\mathcal{H}\phi \notin L^1(\mathbb{R})$ have bad localization. $T^{\dagger}_{1/2}\phi$ and $\mathcal{H}\psi$ have far better localization than $\mathcal{H}\phi$. We give a rigorous result about localization in Section 8.

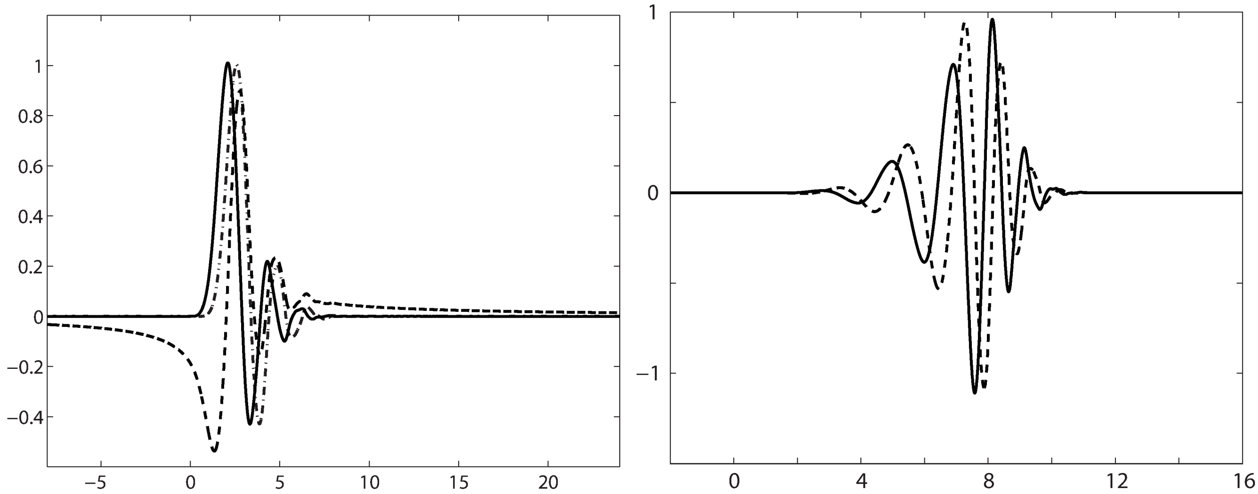

Figure 17. Case of Daubechies wavelets $N = 8$. Left: ${}_8\phi$ (solid), $\mathcal{H}_8\phi$ (broken), $T^\dagger_{1/2}\,{}_8\phi$ (dash-dot). Right: ${}_8\psi$ (solid), $\mathcal{H}_8\psi$ (broken)

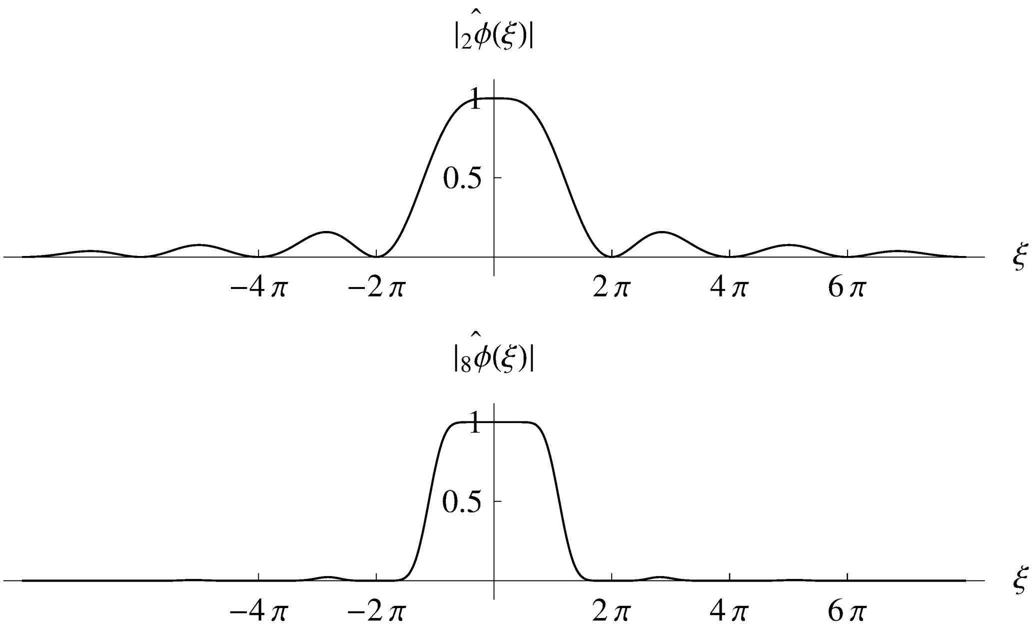

Figure 18. Fourier transforms of Daubechies scaling functions: $|\widehat{{}_2\phi}|, |\widehat{{}_8\phi}|$

§ 8. Properties of $T^\dagger_c$

In this section, we give several properties of $T^\dagger_c$, especially localization property. Let $\mathscr{S}'$ be the space of tempered distributions on $\mathbb{R}$. The operator $(1+|D|^2)^{s/2}$, $s \in \mathbb{R}$, is defined as $\left\{(1+|D|^2)^{s/2}f\right\}^\wedge(\xi) = (1+|\xi|^2)^{s/2}\widehat{f}(\xi)$ for $f \in \mathscr{S}'$. From now on, the derivatives are considered in the distribution sense.

As for smoothness, $T^\dagger_c f$ and $\mathcal{H}_c f$ have the same smoothness as f in the following sense. For $s \in \mathbb{R}$, set $H^s = \{ f \in \mathscr{S}' \mid (1+|D|^2)^{s/2} f \in L^2(\mathbb{R}) \}$, which is the Sobolev space of order s. The following is almost trivial by the boundedness of $e^{-ic\tau(\xi)}$ and $\operatorname{sgn}\xi$.

Proposition 8.1. *Let $s \geq 0$. If $f \in H^s$, then $T_c^\dagger f \in H^s$ and $\mathcal{H}_c f \in H^s$.*

Next, we measure the order of localization of $f(x)$ by the index $r \in \mathbb{N} \cup \{0\}$ such that $(1+|\cdot|)^r f \in L^2(\mathbb{R})$, which is equivalent to $\widehat{f}^{(j)} \in L^2(\mathbb{R})$, $0 \leq j \leq r$.

For $r \in \mathbb{N} \cup \{0\}$ and $s \in \mathbb{R}$, we set

$$
\begin{aligned}
H_r^s :&= \{ f \in \mathscr{S}' \mid (1+|\cdot|)^r (1+|D|^2)^{s/2} f \in L^2(\mathbb{R}) \} \\
&= \{ f \in \mathscr{S}' \mid (\cdot)^j (1+|D|^2)^{s/2} f \in L^2(\mathbb{R}) \ \text{ for } \ 0 \leq j \leq r \} \\
&= \{ f \in \mathscr{S}' \mid \partial_\xi^j \{ (1+|\xi|^2)^{s/2} \widehat{f}(\xi) \} \in L^2(\mathbb{R}) \ \text{ for } \ 0 \leq j \leq r \} \\
&= \{ f \in \mathscr{S}' \mid (1+|\cdot|^2)^{s/2} \widehat{f}^{(j)} \in L^2(\mathbb{R}) \ \text{ for } \ 0 \leq j \leq r \}.
\end{aligned}
\tag{8.1}
$$

The condition $f \in H_r^s$ implies an estimate of decreasing order of f as follows.

Lemma 8.2. *If $f \in H_r^s$, then there exists a constant C such that*

$$|(1+|D|^2)^{s/2-1/4-\epsilon} f(x)| \leq \frac{C}{(1+|x|)^r}, \ x \in \mathbb{R}.$$

In particular, if $s > 1/2$, then $f \in H_r^s$ implies

$$|f(x)| \leq \frac{C}{(1+|x|)^r}, \ x \in \mathbb{R}.$$

Proof. Note that $(1+|\cdot|^2)^{-1/4-\epsilon} \in L^2(\mathbb{R})$ for every $\epsilon > 0$, and hence $f \in H_r^s$ implies that $(1+|\cdot|^2)^{s/2-1/4-\epsilon} \widehat{f}^{(j)} \in L^1(\mathbb{R})$ for $0 \leq j \leq r$. This is equivalent to $\partial_\xi^j \{ (1+|\xi|^2)^{s/2-1/4-\epsilon} \widehat{f}(\xi) \} \in L^1(\mathbb{R})$ for $0 \leq j \leq r$, which implies $(\cdot)^j (1+|D|^2)^{s/2-1/4-\epsilon} f \in L^\infty(\mathbb{R})$ for $0 \leq j \leq r$. Thus, there exists a constant C such that

$$|(1+|D|^2)^{s/2-1/4-\epsilon} f(x)| \leq \frac{C}{(1+|x|)^r}, \ x \in \mathbb{R}.$$

□

We found that the vanishing moments property of ψ is closely relevant to the localization not only of $\mathcal{H}_c \psi$, but also of $T_c^\dagger \phi$. For $r \in \mathbb{N}$, we say that a wavelet function ψ has r *vanishing moments* if $(1+|x|)^{r-1} \psi(x) \in L^1(\mathbb{R})$ and

$$\int_{\mathbb{R}} x^j \psi(x)\, dx = 0, \quad 0 \leq j < r.$$

The Haar wavelet has one vanishing moment, the Meyer wavelets have ∞ vanishing moments, and the Daubechies wavelet ${}_N\psi$ has N vanishing moments. As for the Shannon wavelet, we have $\psi_{\mathrm{S}} \notin L^1(\mathbb{R})$, and hence we cannot consider the integral in the usual sense.

Note that if $r \in \mathbb{N}$ and $f \in H^0_r$, then $(1+|\cdot|)^{r-1} f \in L^1(\mathbb{R})$ and hence $\widehat{f} \in C^{r-1}(\mathbb{R})$, which allows us to talk about $\int_{\mathbb{R}} x^j f(x)\,dx$ and $\widehat{f}^{(j)}(0)$ for $0 \le j < r$. The following is a variant of a well-known result, and it can be proved in the same way as in [4], though the assumptions are a little different.

Theorem 8.3. *Assume that $r \in \mathbb{N}$ and $\phi, \psi \in H^0_r$. Then, $\widehat{\phi}$ and $\widehat{\psi}$ are of C^{r-1} class. Also assume that*

$$\text{(8.2)} \qquad \textit{there exists } l_0 \in \mathbb{Z} \textit{ such that } \ \widehat{\phi}(\pi + 2l_0\pi) \neq 0.$$

Then, m_0 is also of C^{r-1} class. Further, ψ has r vanishing moments if and only if each of the following conditions is satisfied.

(1) $\widehat{\psi}^{(j)}(0) = 0,\ 0 \le j < r.$
(2) $m_0^{(j)}(\pi) = 0,\ 0 \le j < r.$
(3) $\widehat{\phi}^{(j)}(2k\pi) = 0,\ 0 \le j < r,\ k \in \mathbb{Z} \setminus \{0\}.$

Thus, we can consider the condition (3) as a moment condition.

Remark. The assumption (8.2) is a technical condition, and it is satisfied in most cases where $\widehat{\phi} \in C^0(\mathbb{R})$. It is well-known that if ϕ is a scaling function, then we have

$$\text{(8.3)} \qquad \sum_{k\in\mathbb{Z}} |\widehat{\phi}(\xi + 2k\pi)|^2 = 1 \quad \text{a.e. on } \mathbb{R}.$$

But this holds only a.e. in ξ, and it does not necessarily imply (8.2), even if $\widehat{\phi} \in C^0(\mathbb{R})$. If we further impose some conditions which imply the uniform convergence of the series in (8.3), then we can show that (8.3) holds for every ξ, and hence (8.2) holds. For example, the condition that there exists a constant $\epsilon > 0$ such that $\phi \in H^{1/2+\epsilon}_{1/2+\epsilon}$ implies the uniform convergence. In fact, if $\phi \in H^{1/2+\epsilon}_{1/2+\epsilon}$, then we have $(1+|D|^2)^{1/4+\epsilon/2}\phi = (1+|x|)^{1/2+\epsilon}(1+|D|^2)^{1/4+\epsilon/2}\phi \times (1+|x|)^{-1/2-\epsilon} \in L^2 \times L^2 \subset L^1$. Hence, $(1+|\xi|^2)^{1/4+\epsilon/2}\widehat{\phi}(\xi) \in L^\infty$, which implies that there exists a constant C such that $(1+|\xi|^2)^{1/4+\epsilon/2}|\widehat{\phi}(\xi)| \le C$. We have only to show the uniform convergence of (8.3) on $[-\pi, \pi]$, which follows from $|\widehat{\phi}(\xi + 2k\pi)|^2 \le \dfrac{C}{(1+|\xi+2k\pi|)^{1+2\epsilon}} \le \dfrac{C(1+|\xi|)^{1+2\epsilon}}{(1+2|k|\pi)^{1+2\epsilon}}$.

Now, we fix $r \in \mathbb{N}$ and $s \in \mathbb{R}$ with $s \ge 0$. We show that the localization condition $\phi \in H^s_r$ together with the moment condition (3) in Theorem 8.3 are preserved by $T^\dagger_c$. We also give a similar result about $\mathcal{H}_c$, whose proof is similar and simpler. As for $\mathcal{H} = \mathcal{H}_{1/2}$, a similar result on localization was obtained in [6].

Theorem 8.4 ([3]). *Let $r \in \mathbb{N}$ and $s \in \mathbb{R}$, $s \geq 0$.*
(1) *If $f \in H_r^s$ and if $\widehat{f}^{(j)}(2k\pi) = 0$ for $0 \leq j < r$, $k \in \mathbb{Z}\backslash\{0\}$, then $T_c^\dagger f$ also satisfies the same conditions, that is, $T_c^\dagger f \in H_r^s$ and $(\widehat{T_c^\dagger f})^{(j)}(2k\pi) = 0$ for $0 \leq j < r$, $k \in \mathbb{Z} \setminus \{0\}$.*
(2) *If $f \in H_r^s$ and if $\widehat{f}^{(j)}(0) = 0$ for $0 \leq j < r$, then $\mathcal{H}_c f$ also satisfies the same conditions, that is, $\mathcal{H}_c f \in H_r^s$ and $(\widehat{\mathcal{H}_c f})^{(j)}(0) = 0$ for $0 \leq j < r$.*

Remark. (1) We can also show the following by similar (and easier) proofs.

(i) If $\widehat{f} \in C^{r-1}(\mathbb{R})$ and if $\widehat{f}^{(j)}(2k\pi) = 0$ for $0 \leq j < r$, $k \in \mathbb{Z}\backslash\{0\}$, then $\widehat{T_c^\dagger f} \in C^{r-1}(\mathbb{R})$ and $(\widehat{T_c^\dagger f})^{(j)}(2k\pi) = 0$ for $0 \leq j < r$, $k \in \mathbb{Z} \setminus \{0\}$.

(ii) If $\widehat{f} \in C^{r-1}(\mathbb{R})$ and if $\widehat{f}^{(j)}(0) = 0$ for $0 \leq j < r$, then $\widehat{\mathcal{H}_c f} \in C^{r-1}(\mathbb{R})$ and $(\widehat{\mathcal{H}_c f})^{(j)}(0) = 0$ for $0 \leq j < r$.

We first proved these, but we are not satisfied with these, because the condition $\widehat{f} \in C^{r-1}(\mathbb{R})$ is not a good condition as a localization condition of f.

(2) We restricted ourselves to the case $s \geq 0$ since we defined the operators $T_c^\dagger$ and $\mathcal{H}_c$ only on $L^2(\mathbb{R})$. We can extend the results to the case $s < 0$ by extending the operators $T_c^\dagger$ and $\mathcal{H}_c$ on H^s.

Example 8.5. (1) In the case of Meyer wavelets, we can apply our theorem for all $r, s \in \mathbb{N}$, and hence we have $T_c^\dagger\phi, \mathcal{H}_c\psi \in \mathscr{S}$ by Lemma 8.2. Although $T_c^\dagger\phi, \mathcal{H}_c\psi \in \mathscr{S}$ is almost trivial by the definition, this shows that Theorem 8.4 has enough power to derive this strong property.
(2) If $\phi = {}_N\phi$ and $\psi = {}_N\psi$ are the Daubechies scaling function and wavelet function, then we can apply our theorem for $r = N$ and $s = 0$. In particular, $\mathcal{H}_c\, {}_N\psi$ has also N vanishing moments.

If $N \geq 3$, then we can apply our theorem for $r = N$ and $s = 1$, since it is known that ${}_N\phi, {}_N\psi \in C^1(\mathbb{R})$ for $N \geq 3$. In particular, there exists a constant C such that

$$|(T_c^\dagger\, {}_N\phi)(x)| \leq \frac{C}{(1+|x|)^N}, \quad |(\mathcal{H}_c\, {}_N\psi)(x)| \leq \frac{C}{(1+|x|)^N},$$

by Lemma 8.2.

For $N = 2$, we can show that ${}_2\phi, {}_2\psi \in H_2^{1/2+\epsilon}$ (see [3] for details). Thus, we can use our results for $r = 2$ and $s = 1/2 + \epsilon$. This implies that there exists a constant C such that

$$|(T_c^\dagger\, {}_2\phi)(x)| \leq \frac{C}{(1+|x|)^2}, \quad |(\mathcal{H}_c\, {}_2\psi)(x)| \leq \frac{C}{(1+|x|)^2},$$

by Lemma 8.2.

For $N = 1$ (Haar), we can have only that $(1+|x|)T_c^\dagger\, {}_1\phi$, $(1+|x|)\mathcal{H}_c\, {}_1\psi \in L^2(\mathbb{R})$, which implies $T_c^\dagger\, {}_1\phi, \mathcal{H}_c\, {}_1\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

(3) As for the Shannon wavelet, since $\phi_{\mathsf{S}}, \psi_{\mathsf{S}} \notin H_1^0$, we cannot use our theorem. Just like $\phi_{\mathsf{S}}, \psi_{\mathsf{S}}$ themselves, we have $T_c^\dagger \phi_{\mathsf{S}}, \mathcal{H}_c \psi_{\mathsf{S}} \notin L^1(\mathbb{R})$ for every $c \in \mathbb{R}$.

§ 9. A generalization of the Meyer scaling functions

If $\operatorname{supp} \widehat{\phi} \subset [-2\pi, 2\pi]$, then we have $T_c^\dagger \phi = T_c \phi$. In this last section, we give a class of scaling functions with this property, which generalizes the Meyer scaling functions.

Definition 9.1 ([3]). A scaling function $\phi \in L^2(\mathbb{R})$ is called a *generalized Meyer* scaling function if $\operatorname{supp} \widehat{\phi} \subset [-a_1, a_2]$, $0 < a_1 < 2\pi$, $0 < a_2 < 2\pi$, $a_1/2 + a_2 \leq 2\pi$, $a_1 + a_2/2 \leq 2\pi$. A wavelet function associated with a generalized Meyer scaling function is also called a generalized Meyer wavelet function. Note that the condition (A1) in Theorem 4.4 implies $a_1 + a_2 \geq 2\pi$, and the equality holds only if $|\widehat{\phi}| = \chi_{[-a_1, a_2]}$. The region of possible (a_1, a_2) is illustrated as the gray region in Figure 19.

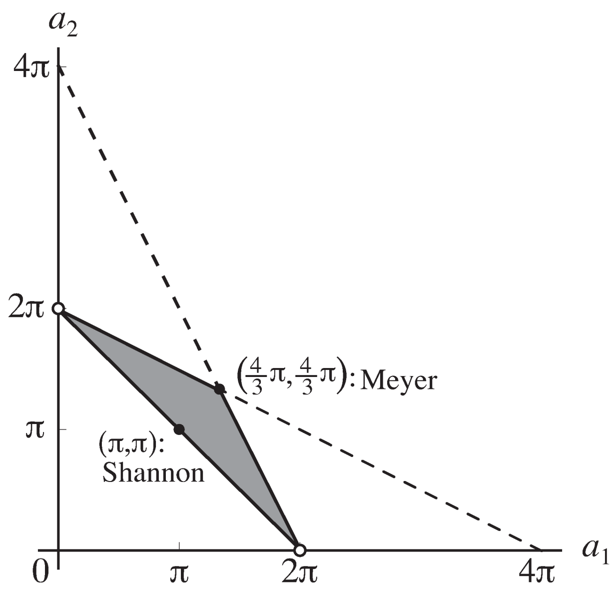

Figure 19. The region of (a_1, a_2). The boundary is included except $(2\pi, 0), (0, 2\pi)$.

Note that the Shannon wavelet is a generalized Meyer wavelet by our definition. The Meyer scaling functions are the case when $a_1 = a_2 = (4/3)\pi$, and the Shannon scaling function is the case when $a_1 = a_2 = \pi$.

In the definition above, it is assumed that ϕ is a scaling function. The following gives the conditions for a function to be a generalized Meyer scaling function.

Proposition 9.2 ([3]). *A function $\phi \in L^2(\mathbb{R})$ is a generalized Meyer scaling function if and only if the following three conditions hold* (Figure 20).

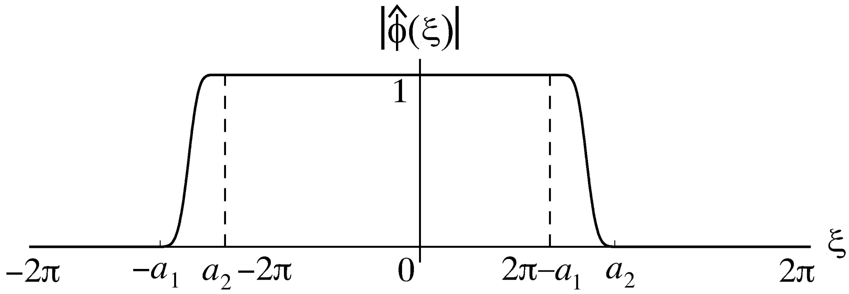

Figure 20. Graph of $|\widehat{\phi}(\xi)|$ for a generalized Meyer scaling function.

(gM1) $\operatorname{supp}\widehat{\phi} \subset [-a_1, a_2]$, $0 < a_1 < 2\pi$, $0 < a_2 < 2\pi$, $a_1/2 + a_2 \le 2\pi$, $a_1 + a_2/2 \le 2\pi$, $a_1 + a_2 \ge 2\pi$.

(gM2) $|\widehat{\phi}(\xi)| = 1$ *a.e. on* $[a_2 - 2\pi, 2\pi - a_1]$.

(gM3) $|\widehat{\phi}(\xi)|^2 + |\widehat{\phi}(\xi - 2\pi)|^2 = 1$ *a.e. on* $[2\pi - a_1, a_2]$.

Note that (gM1) implies $-2\pi < -a_1 \le a_2 - 2\pi < 2\pi - a_1 \le a_2 < 2\pi$, and the width of the support of $\widehat{\phi}$ is not greater than $a_1 + a_2 \le (8/3)\pi$. Also note that the conditions depend only on the absolute value of $\widehat{\phi}$, and hence if ϕ is a generalized Meyer scaling function and if $|\alpha(\xi)| = 1$, then $\alpha(D)\phi$ is also a generalized Meyer scaling function. In particular, if ϕ is a generalized Meyer scaling function, then $T_c\phi$ is also a generalized Meyer scaling function, which also follows from that $T_c\phi = T_c^\dagger\phi$.

Let ϕ be a generalized Meyer scaling function, and ψ be the wavelet function naturally associated with ϕ. If $\phi \in \mathscr{S}$, then the three functions $T_c^\dagger\phi = T_c\phi$, ψ, and $\mathcal{H}_c\psi$ also belong to $\mathscr{S}$, while $\mathcal{H}_c\phi \notin L^1(\mathbb{R})$ unless $c \in \mathbb{Z}$.

The generalized Meyer wavelet functions have the following properties.

Proposition 9.3. *If ϕ is a generalized Meyer scaling function, then any associated wavelet function ψ has the following properties* (Figure 21).

(gMw1) $\operatorname{supp}\widehat{\psi} \subset [-2a_1, a_2 - 2\pi] \cup [2\pi - a_1, 2a_2]$.

(gMw2) $|\widehat{\psi}(\xi)| = 1$ *a.e. on* $[2a_2 - 4\pi, -a_1] \cup [a_2, 4\pi - 2a_1]$,

(gMw3) $|\widehat{\psi}(2\xi + 4\pi)| = |\widehat{\psi}(\xi)|$ *a.e. on* $[-a_1, a_2 - 2\pi]$, $|\widehat{\psi}(2\xi - 4\pi)| = |\widehat{\psi}(\xi)|$ *a.e. on* $[2\pi - a_1, a_2]$, $|\widehat{\psi}(\xi)|^2 + |\widehat{\psi}(\xi - 2\pi)|^2 = 1$ *a.e. on* $[2\pi - a_1, a_2]$.

This proposition easily follows from (4.5) and (4.6).

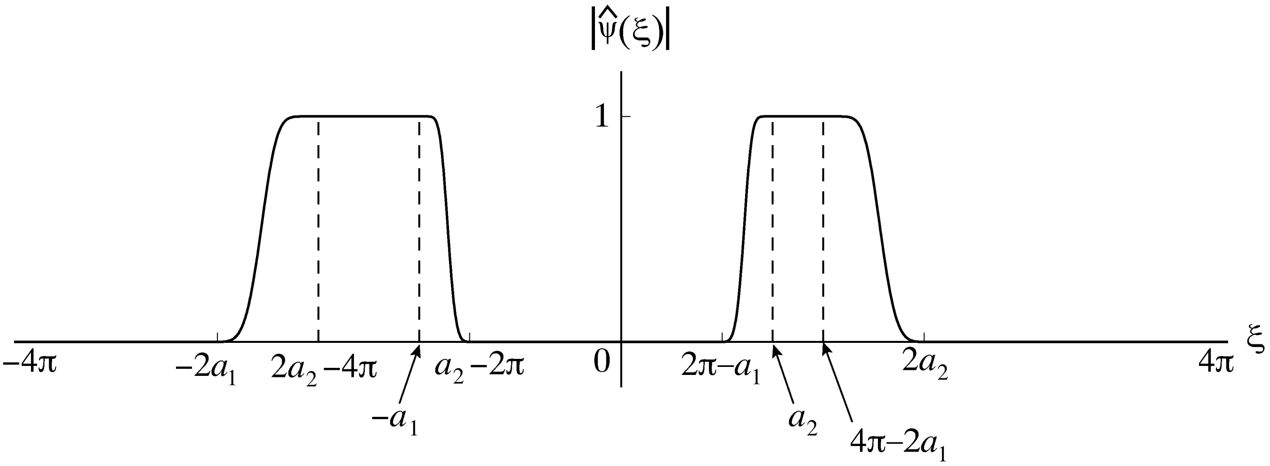

Figure 21. Graph of $|\widehat{\psi}(\xi)|$ for a generalized Meyer wavelet function.

References

[1] Ashino, R., Mandai, T., Morimoto, A. and Sasaki, F., Blind source separation of spatio-temporal mixed signals using time-frequency analysis, *Appl. Anal.*, **88** (2009), 425–456.

[2] Ashino, R., Mandai, T. and Morimoto, A., Blind source separation of spatio-temporal mixed signals using phase information of analytic wavelet transform, *Int. J. Wavelets Multiresolut. Inf. Process.*, **8** (2010), 575–594.

[3] Ashino, R., Mandai, T. and Morimoto, A., Scaling functions generating fractional Hilbert transforms of a wavelet function, *J. Math. Soc. Japan*, to appear.

[4] Butzer, P.L., Fischer, A. and Rückforth, K., Scaling functions and wavelets with vanishing moments, *Comput. Math. Appl.*, **27** (1994), 33–39.

[5] Chaudhury, K.N. and Unser, M., On the Shiftability of Dual-Tree Complex Wavelet Transforms, *IEEE Trans. Signal Process.*, **58** (2010), 221-232.

[6] Chaudhury, K.N. and Unser, M., On the Hilbert Transform of Wavelets, *IEEE Trans. Signal Process.*, **59** (2011), 1890–1894.

[7] Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Math., SIAM, Philadelphia, PA, 1992.

[8] Hernández, E. and Weiss, G., *A First Course on Wavelets*, CRC Press, Boca Raton, FL, 1996.

[9] King, F.W., *Hilbert Transforms, Volume 1*, Encyclopedia of Mathematics and its Applications 124, Cambridge University Press, Cambridge, UK, 2009.

[10] King, F.W., *Hilbert Transforms, Volume 2*, Encyclopedia of Mathematics and its Applications 125, Cambridge University Press, Cambridge, UK, 2009.

[11] Mallat, S., *A Wavelet Tour of Signal Processing — The Sparse Way*, Third Edition. Elsevier/Academic Press, Amsterdam, 2009.

[12] Selesnick, I.W., Hilbert transform pairs of wavelet bases, *IEEE Signal Processing Letters*, **8** (2001), 170–173.

[13] Selesnick, I.W., Baraniuk, R.G. and Kingsbury, N.G., The dual-tree complex wavelet transform, *IEEE Signal Processing Magazine*, **22** (2005), 123–151.

[14] Toda, H. and Zhang, Z., Perfect translation invariance with a wide range of shapes of Hilbert transform pairs of complex wavelet, In: *Proceedings of the International Confer-*

ence on Wavelet Analysis and Pattern Recognition 2007(ICWAPR 2007), Beijing, IEEE, 2007, pp.1565–1570.

[15] Toda, H. and Zhang, Z., Perfect translation invariance with a wide range of shapes of Hilbert transform pairs of wavelet bases, *Int. J. Wavelets Multiresolut. Inf. Process.*, **8** (2010), 501–520.

[16] Wojtaszczyk, P., *A Mathematical Introduction to Wavelets*, London Mathematical Society Student Texts 37, Cambridge University Press, Cambridge, 1997.