Chapter 1

An Overview

1.1 Brief Review on Fourier Series and Fourier Transform

1.1.1 Fourier series

Let $L^2(0, 2\pi)$ denote the collection of all measurable (complex-valued) functions f defined on the interval $(0, 2\pi)$ with

$$\int_0^{2\pi} \left|f(x)\right|^2 dx < \infty .$$

For the reader who is not familiar with the basic Lebesgue theory, the sacrifice is very minimal by assuming that f is a piecewise continuous function. It will always be assumed that functions in $L^2(0, 2\pi)$ are extended periodically to the real line $\mathbb{R} = (-\infty, \infty)$, namely: $f(x) - f(x - 2\pi)$ for all x. Hence, the collection $L^2(0, 2\pi)$ is often called the space of 2π-periodic square-integrable functions. That $L^2(0, 2\pi)$ is a vector space can be verified very easily. Any f in $L^2(0, 2\pi)$ has a Fourier series representation:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} , \tag{1.1.1}$$

where the constants c_n, called the Fourier coefficients of f, are defined by

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx . \tag{1.1.2}$$

The convergence of the series in (1.1.1) is in $L^2(0, 2\pi)$, meaning that

$$\lim_{N,M\to\infty} \int_0^{2\pi} \left| f(x) - \sum_{n=-M}^{N} c_n e^{inx} \right|^2 dx = 0 .$$

There are two distinct features in the Fourier series representation (1.1.1).

1. Any $f \in L^2(0, 2\pi)$ is decomposed into a sum of infinitely many mutually orthogonal components $g_n(x) \equiv c_n e^{inx}$, where orthogonality means that

$$\langle g_n, g_m \rangle_{L^2(0,2\pi)} = 0 \qquad \text{for all } m \neq n \tag{1.1.3}$$

with the "inner product" in (1.1.3) being defined by

$$\langle g_n, g_m \rangle_{L^2(0,2\pi)} = \frac{1}{2\pi} \int_0^{2\pi} g_n(x)\overline{g_m(x)}\, dx\,. \tag{1.1.4}$$

That (1.1.3) holds is a consequence of the important, yet simple fact that

$$\{e_n \in L^2(0,2\pi) \,|\, e_n(x) \equiv e^{inx}, n \in \mathbb{Z}\} \tag{1.1.5}$$

is an orthonormal (o.n.) "basis" of $L^2(0,2\pi)$.

2. The orthonormal basis $\{e_n\}_{n=-\infty}^{\infty}$ used in the Fourier series representation (1.1.1) is generated by "dilation" of a single function

$$e(x) = e^{ix}; \tag{1.1.6}$$

that is, $e_n(x) = e(nx)$ for all integers n. This will be called integral dilation. The theory of the Fourier series shows that "every 2π-periodic square-integrable function is generated by a "superposition" of integral dilations of the basic function e^{ix}.

Definition 1.1. An orthonormal set $\{e_n\}_{n=1}^{\infty}$ in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is said to be complete or is called an orthonormal basis if for every $v \in H$, we have

$$v = \sum_{n=1}^{\infty} \langle v, e_n \rangle e_n\,,$$

where, with $\|\cdot\|$ denote the norm induced by the inner product, the equality above is understood in the sense that

$$\lim_{n\to\infty} \Big\| v - \sum_{k=1}^{n} \langle v, e_k \rangle e_k \Big\| = 0\,.$$

Theorem 1.2. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space, and $\{e_n\}_{n=1}^{\infty}$ be an orthonormal set in H. The following three statements are equivalent:*

1. $\{e_n\}_{n=1}^{\infty}$ *is complete;*
2. $\langle v, e_n \rangle = 0$ *for all $n \in \mathbb{N}$ implies that $v = 0$;*
3. *for every $v \in H$, $\|v\|^2 \equiv \langle v, v \rangle = \sum\limits_{n=1}^{\infty} \big|\langle v, e_n \rangle\big|^2$.*
4. *for every $u, v \in H$, $\langle u, v \rangle = \sum\limits_{n=1}^{\infty} \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$.*

From the o.n. property of $\{e_n\}_{n=-\infty}^{\infty}$, the Fourier series representation (1.1.1) also satisfies the so-called Parseval identity:

$$\frac{1}{2\pi} \int_0^{2\pi} \big|f(x)\big|^2 dx = \sum_{n=-\infty}^{\infty} |c_n|^2\,. \tag{1.1.7}$$

Let ℓ^2 denote the space of all square-summable bi-infinite sequences; that is, $\{c_n\}_{n=-\infty}^{\infty} \in \ell^2$ if and only if

$$\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty .$$

Hence, if the square-root of the quantity on the left of (1.1.7) is used as the "norm" for the measurement of functions in $L^2(0, 2\pi)$, and similarly, the square-root of the quantity on the right of (1.1.7) is used as the norm for ℓ^2, then the function space $L^2(0, 2\pi)$ and the sequence space ℓ^2 are "isometric" to each other. Returning to the above mentioned observation on the Fourier series representation (1.1.1), we can also say that every 2π-periodic square-integrable function is an ℓ^2-linear combination of integral dilations of the basic function $e(x) = e^{lx}$.

We emphasize again that the basic function

$$e(x) = e^{ix} = \cos x + i \sin x$$

which is a "sinusoidal wave", is the only function required to generate all 2π-periodic square-summable functions. For any integer n with large absolute value, the wave $e_n(x) = e(nx)$ has high "frequency", and for n with small absolute value, the wave e_n has low frequency. So, every function in $L^2(0, 2\pi)$ is composed of waves with various frequencies.

In general, we can consider $L^2(0, L)$, the collection of square-integrable (complex-valued) functions defined on $(0, L)$, and extend $f \in L^2(0, L)$ periodically (with period L) to the real line satisfying $f(x) = f(x - L)$ for all $x \in \mathbb{R}$. In other words, $L^2(0, L)$ can be viewed as the collection of all L-periodic complex-valued functions defined on $\mathbb{R}$ satsifying

$$\int_0^L |f(x)|^2 \, dx < \infty .$$

The inner production of $L^2(0, L)$ is given by

$$\langle f, g \rangle_{L^2(0,L)} = \frac{1}{L} \int_0^L f(x) \overline{g(x)} \, dx$$

with induced norm $\| \cdot \|_{L^2(0,L)}$ given by

$$\|f\|_{L^2(0,L)} = \Big(\frac{1}{L} \int_0^L |f(x)|^2 \, dx \Big)^{\frac{1}{2}} .$$

Any $f \in L^2(0, L)$ has the Fourier series representation

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{\frac{2\pi i n x}{L}} , \tag{1.1.8}$$

where the Fourier coefficients $\{c_n\}_{n=-\infty}^{\infty}$ are given by $c_n = \dfrac{1}{L} \displaystyle\int_0^L f(x) e^{-\frac{2\pi i n x}{L}} \, dx$, and (1.1.8) is understood in the sense

$$\lim_{N,M \to \infty} \Big\| f - \sum_{n=-M}^{N} c_n d_{\frac{L}{2\pi}} e_n \Big\|_{L^2(0,L)} = 0 , \tag{1.1.9}$$

where for a constant $c > 0$, d_c is the dilation operator given by

$$(d_c f)(x) = f(x/c) .$$

1.1.2 Fourier transform

Next we consider functions defined on $\mathbb{R}$ without periodicity.

Definition 1.3. The space $L^p(\mathbb{R})$, where $1 \leqslant p < \infty$, is a normed space that consists of all complex-valued measurable functions satisfying

$$L^p(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, \int_{\mathbb{R}} |f(t)|^p \, dt < \infty \right\}$$

which is equipped with norm $\| \cdot \|_{L^p(\mathbb{R})}$ given by

$$\|f\|_{L^p(\mathbb{R})} = \left(\int_{\mathbb{R}} \big|f(t)\big|^p \, dt \right)^{\frac{1}{p}},$$

and the space $L^\infty(\mathbb{R})$ consists of all complex-valued (essentially) bounded measurable functions equipped with norm

$$\|f\|_{L^\infty(\mathbb{R})} = \operatorname*{ess\,sup}_{x \in \mathbb{R}} \big|f(x)\big|.$$

Note that $(L^p(\mathbb{R}), \| \cdot \|_{L^p(\mathbb{R})})$ is a Banach space; that is, a complete normed space.

Definition 1.4. For all $f \in L^1(\mathbb{R})$, the Fourier transform of f, denoted by $\mathscr{F} f$ or $\widehat{f}$, is a function defined by

$$(\mathscr{F} f)(\omega) = \widehat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-it\omega} dt \qquad \forall\, \omega \in \mathbb{R}. \tag{1.1.10}$$

Let $\mathscr{C}_0(\mathbb{R})$ denote the space of all bounded/continuous functions on $\mathbb{R}$ which decay at infinity; that is, $f \in \mathscr{C}_0(\mathbb{R})$ if and only if $f(t) \to 0$ as $t \to \infty$. $\mathscr{C}_0(\mathbb{R})$ is a normed space equipped the sup-norm defined by

$$\|f\| = \sup_{t \in \mathbb{R}} \big|f(t)\big|.$$

Theorem 1.5. *$\mathscr{F} : L^1(\mathbb{R}) \to \mathscr{C}_0(\mathbb{R})$ is a bounded linear map.*

Proof. To show this theorem, we need to establish the following properties:

1. $\mathscr{F}$ is linear;

2. there exists a constant C such that $\|\mathscr{F}[f]\|_{L^\infty(\mathbb{R})} \leqslant C \|f\|_{L^1(\mathbb{R})}$;

3. $\mathscr{F}[f]$ is continuous and $\mathscr{F}[f](\omega) \to 0$ as $|\omega| \to \infty$.

The linearity of $\mathscr{F}$ is trivial. Moreover, if $f \in L^1(\mathbb{R})$,

$$|\widehat{f}(\omega)| \leqslant \int_{\mathbb{R}} |f(t)| \, dt = \|f\|_{L^1(\mathbb{R})} < \infty$$

which shows that $\|\mathscr{F}[f]\|_{L^\infty(\mathbb{R})} \leqslant \|f\|_{L^1(\mathbb{R})}$. To show that $\widehat{f}$ is continuous, we in fact show that $\widehat{f}$ is uniformly continuous as follows. Let $\varepsilon > 0$ be given. Note that the Dominated Convergence Theorem implies that

$$\lim_{\Delta w \to 0} \int_{\mathbb{R}} |f(t)| \big|e^{-it\Delta w} - 1\big| \, dt = 0;$$

thus there exists $\delta > 0$ such that

$$\int_{\mathbb{R}} |f(t)| \left| e^{-it\Delta w} - 1 \right| dt < \varepsilon \quad \text{whenever} \quad |\Delta w| < \delta.$$

Therefore, if $\omega_1, \omega_2 \in \mathbb{R}$ satisfying $|\omega_1 - \omega_2| < \delta$, we have

$$\left| \widehat{f}(\omega_1) - \widehat{f}(\omega_2) \right| \leqslant \int_{\mathbb{R}} |f(t)| \left| e^{-it(\omega_1 - \omega_2)} - 1 \right| dt < \varepsilon$$

that shows that $\widehat{f}$ is uniformly continuous on $\mathbb{R}$. Finally, since $e^{-it\omega} = -e^{-i\omega(t+\frac{\pi}{\omega})}$,

$$\widehat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-it\omega} \, dt = -\int_{\mathbb{R}} f(t) e^{-i\omega(t+\frac{\pi}{\omega})} \, dt = -\int_{\mathbb{R}} f\Big(t - \frac{\pi}{\omega}\Big) e^{-i\omega t} \, dt.$$

Therefore,

$$\begin{aligned} \left| \widehat{f}(\omega) \right| &= \frac{1}{2} \left| \int_{\mathbb{R}} f(t) e^{-i\omega t} \, dt - \int_{\mathbb{R}} f\Big(t - \frac{\pi}{\omega}\Big) e^{-i\omega t} \, dt \right| \\ &= \frac{1}{2} \left| \int_{\mathbb{R}} \Big[f(t) - f\Big(t - \frac{\pi}{\omega}\Big) \Big] e^{-i\omega t} \, dt \right| \leqslant \frac{1}{2} \int_{\mathbb{R}} \left| f(t) - f\Big(t - \frac{\pi}{\omega}\Big) \right| dt \end{aligned}$$

which converges to 0 as $|\omega| \to \infty$. □

Remark 1.6. The result that $\lim_{|\omega| \to \infty} \left| \widehat{f}(\omega) \right| = 0$ for $f \in L^1(\mathbb{R})$ is often called the Riemann-Lebesgue Lemma.

Before proceeding, we define the following useful operators, called the translation, modulation and (scaled) dilation operators, respectively: for $f \in L^1(\mathbb{R})$,

$$(T_c f)(t) = f(t - c), \qquad (M_c f)(t) = e^{ict} f(t), \qquad (D_c f)(t) = \frac{1}{\sqrt{|c|}} f\Big(\frac{x}{c}\Big). \tag{1.1.11}$$

In particular, D_{-1} is called the parity/reflection operator.

Theorem 1.7. *Let $f \in L^1(\mathbb{R})$. Then*

1. **(Shifting)** $\mathscr{F}[T_c f](\omega) = e^{-ic\omega} \widehat{f}(\omega)$.
2. **(Scaling)** $\mathscr{F}[D_c f](\omega) = \big(D_{\frac{1}{c}} \widehat{f}\big)(\omega)$.
3. **(Conjugation)** $\mathscr{F}\big[\overline{D_{-1} f}\big](\omega) = \mathscr{F}\big[D_{-1} \overline{f}\,\big](\omega) = \overline{\widehat{f}}\,(\omega)$.
4. **(Modulation)** $\mathscr{F}[M_c f](\omega) = (T_c \widehat{f})(\omega)$.

Example 1.8. Let $\psi \in L^1(\mathbb{R})$, and $\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\Big(\frac{t-b}{a}\Big)$. Using the translation and dilation operators we have $\psi_{a,b}(t) = (T_b D_a \psi)(t)$; thus

$$\widehat{\psi_{a,b}}(\omega) = \mathscr{F}[T_b D_a \psi](\omega) = e^{-ib\omega} \mathscr{F}[D_a \psi](\omega) = e^{-ib\omega} D_{\frac{1}{a}} \widehat{\psi}(\omega) = \sqrt{|a|} e^{-ib\omega} \widehat{\psi}(a\omega).$$

Definition 1.9. Let f, g be complex-valued function defined on $\mathbb{R}$. The convolution of f and g, denoted by $f * g$, is the function defined by

$$(f * g)(t) = \int_{\mathbb{R}} f(\tau)g(t-\tau)\,d\tau$$

whenever the integral makes sense.

Note that by the change of variable formula, $f * g = g * f$ whenever the convolution makes sense.

Remark 1.10. The convolution $f * g$ makes sense if $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ with $\frac{1}{p} + \frac{1}{q} \geqslant 1$. In fact, for $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, then $f * g \in L^r(\mathbb{R})$ with $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ and one has Young's inequality

$$\|f * g\|_{L^r(\mathbb{R})} \leqslant \|f\|_{L^p(\mathbb{R})}\|g\|_{L^q(\mathbb{R})}.$$

Theorem 1.11. *If $f, g \in L^1(\mathbb{R})$, then $\mathscr{F}[f * g](\omega) = \widehat{f}(\omega)\widehat{g}(\omega)$.*

Theorem 1.12. *Let $g \in L^1(\mathbb{R})$ and define*

$$h(t) = \int_{\mathbb{R}} g(\omega)e^{it\omega}\,d\omega \left(= \widehat{g}(-t)\right).$$

If $h \in L^1(\mathbb{R})$, then for all $f \in L^1(\mathbb{R})$,

$$(f * h)(t) = \int_{\mathbb{R}} g(\omega)\widehat{f}(\omega)e^{i\omega t}\,d\omega.$$

Definition 1.13. A summability kernel on $\mathbb{R}$ is a family $\{K_\lambda\}_{\lambda>0}$ of continuous functions with the following properties:

(i) $\displaystyle\int_{\mathbb{R}} K_\lambda(x)\,dx = 1$ for all $\lambda > 0$;

(ii) there exists $M > 0$ such that

$$\int_{\mathbb{R}} \left|K_\lambda(x)\right|dx \leqslant M \qquad \forall\, \lambda > 0;$$

(iii) $\displaystyle\lim_{\lambda\to\infty} \int_{|x|>\delta} \left|K_\lambda(x)\right|dx = 0$ for all $\delta > 0$.

A simple construction of a summability on $\mathbb{R}$ is as follows. Suppose F is a continuous Lebesgue integrable function so that

$$\int_{\mathbb{R}} F(x)\,dx = 1.$$

Then, we set

$$K_\lambda(x) = \lambda F(\lambda x) \qquad \text{for all } \lambda > 0 \text{ and } x \in R. \tag{1.1.12}$$

Evidently, it follows that

$$\int_{\mathbb{R}} K_\lambda(x)\,dx = \int_{\mathbb{R}} \lambda F(\lambda x)\,dx = \int_{\mathbb{R}} F(x)\,dx = 1\,,$$
$$\int_{\mathbb{R}} \big|K_\lambda(x)\big|\,dx = \int_{\mathbb{R}} \lambda\big|F(\lambda x)\big|\,dx = \int_{\mathbb{R}} \big|F(x)\big|\,dx = \|F\|_{L^1(\mathbb{R})}$$

and for $\delta > 0$,

$$\int_{|x|>\delta} \big|K_\lambda(x)\big|\,dx = \int_{|x|>\delta} \lambda\big|F(\lambda x)\big|\,dx = \int_{|x|>\lambda\delta} \big|F(x)\big|\,dx \to 0 \quad \text{as } \lambda \to \infty.$$

Therefore, the family $\{K_\lambda\}_{\lambda>0}$ defined by (1.1.12) is a summability kernel on $\mathbb{R}$.

Example 1.14. 1. The family $\{K_\lambda\}_{\lambda>0}$ defined by $K_\lambda(x) = \lambda F(\lambda x)$, where

$$F(x) = \frac{1}{2\pi}\frac{\sin^2(x/2)}{(x/2)^2}$$

is called the Fejér kernel.

2. The family $\{K_\lambda\}_{\lambda>0}$ defined by $K_\lambda(x) = \lambda G(\lambda x)$, where

$$G(x) = \frac{1}{\sqrt{\pi}}\exp(-x^2)$$

is called the Gaussian kernel.

Both the Fejér kernel and the Guassian kernel are summability kernels.

Theorem 1.15. *Let $\{K_\lambda\}_{\lambda>0}$ be a summability kernel on $\mathbb{R}$. If $f \in L^1(\mathbb{R})$, then*

$$\lim_{\lambda\to\infty} \|f * K_\lambda - f\|_{L^1(\mathbb{R})} = 0\,.$$

Moreover, if in addition that f is essentially bounded and is continuous at c, then

$$\lim_{\lambda\to\infty} (f * K_\lambda)(c) = f(c)\,.$$

Proof. Let $f \in L^1(\mathbb{R})$ be given, and $M > 0$ be such that

$$\int_{\mathbb{R}} \big|K_\lambda(x)\big|\,dx \leqslant M\,.$$

Let $\varepsilon > 0$ be given. By the definition of summability kernels and the convolution,

$$\begin{aligned}(f * K_\lambda)(x) - f(x) &= \int_{\mathbb{R}} f(x-y)K_\lambda(y)\,dy - f(x)\int_{\mathbb{R}} K_\lambda(y)\,dy \\ &= \int_{\mathbb{R}} K_\lambda(y)\big[f(x-y) - f(x)\big]\,dy\,. \end{aligned} \tag{1.1.13}$$

1. Since $\lim\limits_{y\to 0} \displaystyle\int_{\mathbb{R}} \big|f(x) - f(x-y)\big|\,dx = 0$, there exists $\delta > 0$ such that if $|y| < \delta$,

$$\int_{\mathbb{R}} \big|f(x) - f(x-y)\big|\,dx < \frac{\varepsilon}{2M}\,.$$

By the Tonelli Theorem,

$$\begin{aligned}
\int_{\mathbb{R}} \big|(f * K_\lambda)(x) - f(x)\big| dx &\leqslant \int_{\mathbb{R}}\int_{\mathbb{R}} \big|K_\lambda(y)\big|\big|f(x-y) - f(x)\big| dydx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \big|K_\lambda(y)\big|\big|f(x-y) - f(x)\big| dxdy \\
&= \Big(\int_{|y|<\delta} + \int_{|y|\geqslant\delta}\Big)\big|K_\lambda(y)\big| \int_{\mathbb{R}} \big|f(x-y) - f(x)\big| dxdy \\
&\leqslant \frac{\varepsilon}{2M}\int_{|y|<\delta} \big|K_\lambda(y)\big| dy + 2\|f\|_{L^1(\mathbb{R})}\int_{|y|\geqslant\delta} \big|K_\lambda(y)\big| dy \\
&\leqslant \frac{\varepsilon}{2} + 2\|f\|_{L^1(\mathbb{R})}\int_{|y|\geqslant\delta} \big|K_\lambda(y)\big| dy .
\end{aligned}$$

Therefore, by the properties of summability kernels,

$$\limsup_{\lambda\to\infty} \int_{\mathbb{R}} \big|(f * K_\lambda)(x) - f(x)\big| dx \leqslant \frac{\varepsilon}{2} < \varepsilon .$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\limsup_{\lambda\to\infty} \int_{\mathbb{R}} \big|(f * K_\lambda)(x) - f(x)\big| dx = 0$$

which shows that $\lim\limits_{\lambda\to\infty} \|f * K_\lambda - f\|_{L^1(\mathbb{R})} = 0$.

2. Now suppose in addition that f is continuous at c. Then there exists $\delta > 0$ such that

$$\big|f(a-y) - f(a)\big| < \frac{\varepsilon}{2M} .$$

Therefore, (1.1.13) implies that

$$\begin{aligned}
\big|(f * K_\lambda)(c) - f(c)\big| &\leqslant \int_{\mathbb{R}} \big|K_\lambda(y)\big|\big|f(c-y) - f(c)\big| dy \\
&= \Big(\int_{|y|<\delta} + \int_{|y|\geqslant\delta}\Big)\big|K_\lambda(y)\big|\big|f(c-y) - f(c)\big| dy \\
&\leqslant \frac{\varepsilon}{2M}\int_{|y|<\delta} \big|K_\lambda(y)\big| dy + 2\|f\|_{L^\infty(\mathbb{R})}\int_{|y|\geqslant\delta} \big|K_\lambda(y)\big| dy .
\end{aligned}$$

The same argument as in Part 1 shows that $\lim\limits_{\lambda\to\infty} \big|(f * K_\lambda)(c) - f(c)\big| = 0$. □

Let Λ be the function defined by

$$\Lambda(x) = \begin{cases} \dfrac{1}{2\pi}(1 - |x|) & \text{if } |x| \leqslant 1, \\ 0 & \text{if } |x| > 1. \end{cases}$$

Then

$$\begin{aligned}
\widehat{\Lambda}(\omega) &= \frac{1}{2\pi}\int_{-1}^{1} (1 - |x|)e^{-ix\omega}\, dx = \frac{1}{\pi}\int_0^1 (1 - x)\cos(x\omega)\, dx \\
&= \frac{1}{\pi}\cdot\frac{1 - \cos\omega}{\omega^2} = \frac{1}{2\pi}\frac{\sin^2(\omega/2)}{(\omega/2)^2} = F(\omega);
\end{aligned}$$

thus the Fejér kernel is given by

$$K_\lambda(x) = \lambda F(\lambda x) = \sqrt{\lambda} D_{\frac{1}{\lambda}} \widehat{\Lambda}(x) = \mathscr{F}[\sqrt{\lambda} D_\lambda \Lambda](x).$$

By Theorem 1.12 (with h being the Fejér kernel), the fact that $\widehat{\Lambda}$ is even implies that

$$(f * K_\lambda)(t) = \int_{\mathbb{R}} \sqrt{\lambda}(D_\lambda \Lambda)(\omega)\widehat{f}(\omega)e^{i\omega t}\, d\omega = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\Big(1 - \frac{|\omega|}{\lambda}\Big)\widehat{f}(\omega)e^{i\omega t}\, d\omega.$$

Using the above identity, Theorem 1.15 and the Dominated Convergence Theorem show the following

Theorem 1.16 (Fourier Inversion Formula). *Let $f \in L^1(\mathbb{R})$. If $\widehat{f} \in L^1(\mathbb{R})$, then*

$$f(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\omega)e^{i\omega t}\, d\omega \qquad \text{for a.a. } t \in \mathbb{R}. \tag{1.1.14}$$

In particular, the identity above holds for t at which f is continuous.

Remark 1.17. Note that in order to show (1.1.14) holds for continuities of f, the boundedness of f is required. Nevertheless, since $\widehat{f} \in L^1(\mathbb{R})$, (1.1.14) shows that f is almost everywhere equal to a continuous function that decays at infinity; thus f is essentially bounded.

Remark 1.18. The integral operator

$$f \mapsto \frac{1}{2\pi}\int_{\mathbb{R}} M_\bullet(\omega)f(\omega)\, d\omega \qquad \Big(\text{or } f(t) = \frac{1}{2\pi}\int_{\mathbb{R}} M_t(\omega)f(\omega)\, d\omega\Big),$$

where $M_\bullet$ is the Modulation operator defined in (1.1.11), is called the inverse Fourier transform and is usually denoted by $\mathscr{F}^{-1}$ or $\check{\ }$. In other words,

$$\mathscr{F}^{-1}[f](t) = \check{f}(t) = \frac{1}{2\pi}\int_{\mathbb{R}} f(\omega)M_t(\omega)\, d\omega = \frac{1}{2\pi}\int_{\mathbb{R}} f(\omega)e^{i\omega t}\, d\omega.$$

Note that $\mathscr{F}^{-1} = \frac{1}{2\pi}D_{-1}\mathscr{F} = \frac{1}{2\pi}\mathscr{F}D_{-1}$. Similar to Theorem 1.5, $\mathscr{F}^{-1} : L^1(\mathbb{R}) \to \mathscr{C}_0(\mathbb{R})$ is a bounded linear map and $\lim\limits_{|t|\to\infty} \check{f}(t) = 0$.

Next we consider the Fourier transform of square-integrable functions. First we note that there exists square-integrable function which is not integrable. For example, the function

$$f(x) = \begin{cases} x^{-3/4} & \text{if } x > 1, \\ 0 & \text{otherwise} \end{cases}$$

belongs to $L^2(\mathbb{R})$ but not $L^1(\mathbb{R})$. It is not possible to find the Fourier transform for this f using Definition 1.4 since the integral $\int_{\mathbb{R}} f(t)e^{it\omega}\, dt$ does not exist. In other words, when we talk about Fourier transform of functions that are not integrable, we indeed try to extend the domain of the original Fourier transform $\mathscr{F}$.

Before proceeding, we introduce the inner product $\langle\cdot,\cdot\rangle_{L^2(\mathbb{R})}$ used in $L^2(\mathbb{R})$:

$$\langle f,g\rangle_{L^2(\mathbb{R})}=\int_{\mathbb{R}} f(x)\overline{g(x)}\,dx \qquad \forall\, f,g\in L^2(\mathbb{R}),$$

and the induced norm is indeed the L^2-norm; that is, $\|f\|_{L^2(\mathbb{R})}=\langle f,f\rangle_{L^2(\mathbb{R})}^{\frac{1}{2}}$. The space $(L^2(\mathbb{R}),\langle\cdot,\cdot\rangle_{L^2(\mathbb{R})})$ is a Hilbert space.

Since $L^1(\mathbb{R})\cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ (for example, if $f\in L^2(\mathbb{R})$, then $f_n\equiv f\mathbf{1}_{[-n,n]}$ is square-integrable and $\{f_n\}_{n=1}^{\infty}$ converges to f in $L^2(\mathbb{R})$), it is natural to define the Fourier transform of a square-integrable function as the limit of the Fourier transform of a sequence in $L^1(\mathbb{R})\cap L^2(\mathbb{R})$ whose L^2-limit is that function (if it exists). To talk about whether the limit of such a sequence exists, we need the following

Lemma 1.19. *Suppose that $f\in L^2(\mathbb{R})$ and f vanishes outside a bounded interval. Then $\widehat{f}\in L^2(\mathbb{R})$ and $\|\widehat{f}\|^2_{L^2(\mathbb{R})}=2\pi\|f\|^2_{L^2(\mathbb{R})}$.*

Proof. Note that since $f\in L^2(\mathbb{R})$ and f vanishes outside a bounded interval, Cauchy-Schwarz inequality shows that $f\in L^1(\mathbb{R})$ so $\widehat{f}\in\mathscr{C}_0(\mathbb{R})$ is well-defined.

Suppose f vanishes outside $[-R,R]$ for some $R>0$. Let $c=R/\pi$ and define

$$g(t)=(M_{-x}D_{\frac{1}{c}}f)(t)=\sqrt{c}e^{-ixt}f(ct),$$

where $x\in\mathbb{R}$ is arbitrarily given. Then g vanishes outside $[-\pi,\pi]$ and Theorem 1.7 shows that $\widehat{g}(\omega)=(T_{-x}D_c\widehat{f})(\omega)=(D_c\widehat{f})(\omega+x)$. On the other hand, the Parseval identity (1.1.7) implies that

$$\begin{aligned}\frac{1}{2\pi}\|g\|^2_{L^2(\mathbb{R})}&=\sum_{n=-\infty}^{\infty}\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}g(t)e^{-int}\,dt\right|^2=\sum_{n=-\infty}^{\infty}\left|\frac{1}{2\pi}\int_{\mathbb{R}}g(t)e^{-int}\,dx\right|^2\\&=\frac{1}{4\pi^2}\sum_{n=-\infty}^{\infty}\left|\widehat{g}(n)\right|^2=\frac{1}{4\pi^2}\sum_{n=-\infty}^{\infty}\left|(D_c\widehat{f})(n+x)\right|^2.\end{aligned}$$

Integrating the identity above in x from 0 to 1, we obtain that

$$\begin{aligned}\|g\|^2_{L^2(\mathbb{R})}&=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_0^1\left|(D_c\widehat{f})(n+x)\right|^2dx=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_n^{n+1}\left|(D_c\widehat{f})(x)\right|^2dx\\&=\frac{1}{2\pi}\int_{\mathbb{R}}\left|(D_c\widehat{f})(x)\right|^2dx=\frac{1}{2\pi}\|D_c\widehat{f}\|^2_{L^2(\mathbb{R})}=\frac{1}{2\pi}\|\widehat{f}\|^2_{L^2(\mathbb{R})}.\end{aligned}$$

The lemma is then concluded since $\|f\|_{L^2(\mathbb{R})}=\|g\|_{L^2(\mathbb{R})}$. □

The collection of square-integrable functions vanishing outside a bounded interval is denoted by $L^2_c(\mathbb{R})$; that is,

$$L^2_c(\mathbb{R})=\big\{f:\mathbb{R}\to\mathbb{C}\,\big|\,f\in L^2(\mathbb{R})\text{ and }f\text{ vanishes outside }[-R,R]\text{ for some }R>0\big\}.$$

Since the support of a function f ($\in L_c^2(\mathbb{R})$), denoted by $\mathrm{supp}(f)$, is defined as the closure of the collection of points at which f does not vanish; that is,

$$\mathrm{supp}(f) = \mathrm{cl}\big(\{x \in \mathbb{R} \,|\, f(x) \neq 0\}\big).$$

Functions in $L_c^2(\mathbb{R})$ are also called square-integrable functions with compact support.

Lemma 1.20. *Let $f \in L^2(\mathbb{R})$. There exists a (unique) function $F \in L^2(\mathbb{R})$ such that if $\{f_n\}_{n=1}^\infty \subseteq L_c^2(\mathbb{R})$ converges to f in $L^2(\mathbb{R})$, then $\{\widehat{f_n}\}_{n=1}^\infty$ converges to F in $L^2(\mathbb{R})$.*

Proof. Let $f \in L^2(\mathbb{R})$, and $\{f_n\}_{n=1}^\infty \subseteq L_c^2(\mathbb{R})$ that converges to f in $L^2(\mathbb{R})$. Then Lemma 1.19 shows that

$$\|\widehat{f_n} - \widehat{f_m}\|_{L^2(\mathbb{R})} = \sqrt{2\pi}\|f_n - f_m\|_{L^2(\mathbb{R})} \to 0 \qquad \text{as } n, m \to \infty.$$

In other words, $\{\widehat{f_n}\}_{n=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R})$; thus the completeness of $L^2(\mathbb{R})$ shows that $\{\widehat{f_n}\}_{n=1}^\infty$ converges to some function F in $L^2(\mathbb{R})$.

Next we show that such a function F is independent of the sequence $\{f_n\}_{n=1}^\infty$ that is used to approach f. Suppose that there is another sequence $\{g_n\}_{n=1}^\infty \subseteq L_c^2(\mathbb{R})$ that also converges to f in $L^2(\mathbb{R})$. Then the argument above shows that $\{\widehat{g_n}\}_{n=1}^\infty$ converges to some function G in $L^2(\mathbb{R})$. On the other hand, the sequence $\{h_n\}_{n=1}^\infty$ defined by

$$h_n = \begin{cases} f_n & \text{if } n \text{ is odd,} \\ g_n & \text{if } n \text{ is even} \end{cases}$$

also converges to f in $L^2(\mathbb{R})$; thus $\{\widehat{h_n}\}_{n=1}^\infty$ converges to some function $H \in L^2(\mathbb{R})$. Nevertheless, since the odd terms of $\{\widehat{h_n}\}_{n=1}^\infty$ is a subsequence of $\{\widehat{f_n}\}_{n=1}^\infty$ and the even terms of $\{\widehat{h_n}\}_{n=1}^\infty$ is a subsequence of $\{\widehat{g_n}\}_{n=1}^\infty$, we must have $F = H = G$. □

The lemma above induces the following

Definition 1.21. Let $f \in L^2(\mathbb{R})$. The Fourier transform of f, still denoted by $\mathscr{F}[F]$ and $\widehat{f}$, is the L^2-limit of the Fourier transform of (any) sequences $\{f_n\}_{n=1}^\infty$ in $L_c^2(\mathbb{R})$ that converges to f in $L^2(\mathbb{R})$. In other words,

$$\mathscr{F}[f] = \widehat{f} \equiv \lim_{n\to\infty} \widehat{f_n} \qquad \text{whenever } \{f_n\}_{n=1}^\infty \subseteq L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \text{ and } \lim_{n\to\infty} f_n = f,$$

where the two limits above are all in the L^2-sense.

Theorem 1.22 (Plancherel's identity)**.**

$$\langle f, g\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\langle \widehat{f}, \widehat{g}\rangle_{L^2(\mathbb{R})} \qquad \forall\, f, g \in L^2(\mathbb{R}). \tag{1.1.15}$$

Proof. First we note that

$$\|f\|^2_{L^2(\mathbb{R})} = \frac{1}{2\pi}\|\widehat{f}\|^2_{L^2(\mathbb{R})} \qquad \forall\, f \in L^2(\mathbb{R}) \tag{1.1.16}$$

since by choosing $\{f_n\}_{n=1}^\infty \subseteq L^2_c(\mathbb{R})$ with L^2-limit f, by Lemma 1.19 we have

$$\|f\|^2_{L^2(\mathbb{R})} = \lim_{n\to\infty} \|f_n\|^2_{L^2(\mathbb{R})} = \lim_{n\to\infty} \frac{1}{2\pi}\|\widehat{f_n}\|^2_{L^2(\mathbb{R})} = \frac{1}{2\pi}\|\widehat{f}\|^2_{L^2(\mathbb{R})}\,,$$

where the last identity follows from the fact that $\{\widehat{f_n}\}_{n=1}^\infty$ converges to $\widehat{f}$ in $L^2(\mathbb{R})$. Identity (1.1.15) then follows from the polarization identity

$$\langle f,g\rangle_{L^2(\mathbb{R})} = \frac{1}{4}\Big[\|f+g\|^2_{L^2(\mathbb{R})} - \|f-g\|^2_{L^2(\mathbb{R})} + i\|f+ig\|^2_{L^2(\mathbb{R})} - i\|f-ig\|^2_{L^2(\mathbb{R})}\Big]. \qquad \square$$

Theorem 1.23. *Let $f,g\in L^2(\mathbb{R})$.*

1. $\big\langle f,\widehat{\overline{g}}\big\rangle_{L^2(\mathbb{R})} = \big\langle \widehat{f},\overline{g}\big\rangle_{L^2(\mathbb{R})}$.
2. *If $g=\overline{\widehat{f}}$, then $f=\dfrac{1}{2\pi}\widehat{\overline{g}}$. In other words, the operator $\dfrac{1}{\sqrt{2\pi}}\,\overline{\widehat{\ }}$ is the inverse of itself.*

Remark 1.24. The Plancherel identity often refers to the following identity

$$\|f\|^2_{L^2(\mathbb{R})} = \frac{1}{2\pi}\|\widehat{f}\|^2_{L^2(\mathbb{R})} \qquad \forall\, f\in L^2(\mathbb{R})\,. \tag{1.1.17}$$

In this lecture the Plancherel identity of the form (1.1.16) will be used extensively. Nevertheless, the Plancherel identity of the form (1.1.17) can be applied to show that

$$\mathscr{F}\Big[\sum_{n=-\infty}^{\infty} c_n\phi_n\Big] = \sum_{n=-\infty}^{\infty} c_n\widehat{\phi_n} \quad \forall\,\{c_n\}_{n\in\mathbb{Z}}\in\ell^2 \text{ and } \{\phi_n\}_{n\in\mathbb{Z}} \text{ is orthonormal in } L^2(\mathbb{R})\,. \tag{1.1.18}$$

In other words, under the conditions stated above the Fourier transform $\mathscr{F}$ commutes with infinite sums. To see why (1.1.18) is true, we first note that the Plancherel identity (1.1.15) shows that $\Big\{\dfrac{\widehat{\phi_n}}{\sqrt{2\pi}}\Big\}_{n\in\mathbb{Z}}$ is an orthonormal set in $L^2(\mathbb{R})$. By setting

$$s_k = \sum_{n=-k}^{k} c_n\phi_n\,,$$

the Plancherel identity (1.1.17) implies that

$$\|\widehat{s_k}-\widehat{s_\ell}\|^2_{L^2(\mathbb{R})} = 2\pi\|s_k-s_\ell\|^2_{L^2(\mathbb{R})} = 2\pi \sum_{\min\{k,\ell\}<|n|\leqslant\max\{k,\ell\}} |c_n|^2 \to 0 \quad \text{as } k,\ell\to\infty\,. \tag{1.1.19}$$

This shows that $\{\widehat{s_k}\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbb{R})$; thus $\{\widehat{s_k}\}_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R})$. Nevertheless, $\widehat{s_k} = \sum\limits_{n=-k}^{k} c_n\widehat{\phi_n}$, so the fact that $\{c_n\}_{n\in\mathbb{Z}}\in\ell^2$ and that $\Big\{\dfrac{\widehat{\phi_n}}{\sqrt{2\pi}}\Big\}_{n\in\mathbb{Z}}$ is an orthonormal basis in $L^2(\mathbb{R})$ implies that

$$\lim_{k\to\infty}\widehat{s_k} = \sum_{n=-\infty}^{\infty} c_n\widehat{\phi_n} \qquad \text{and the convergence is in } L^2(\mathbb{R})\,.$$

On the other hand, part 1 of Theorem 1.23 for all functions $\phi\in L^2(\mathbb{R})$ we have

$$\begin{aligned}\Big\langle \sum_{n=-\infty}^{\infty} c_n\widehat{\phi_n},\phi\Big\rangle_{L^2(\mathbb{R})} &= \lim_{k\to\infty}\langle\widehat{s_k},\phi\rangle_{L^2(\mathbb{R})} = \lim_{k\to\infty}\Big\langle s_k,\widehat{\overline{\phi}}\Big\rangle_{L^2(\mathbb{R})} = \Big\langle \sum_{n=-\infty}^{\infty} c_n\phi_n,\widehat{\overline{\phi}}\Big\rangle_{L^2(\mathbb{R})}\\ &= \Big\langle \mathscr{F}\Big[\sum_{n=-\infty}^{\infty} c_n\phi_n\Big],\phi\Big\rangle_{L^2(\mathbb{R})}.\end{aligned}$$

Since this identity holds for all $\phi\in L^2(\mathbb{R})$, we conclude (1.1.18).

Remark 1.25. There are ways to define the Fourier transform for more general "functions" (termed tempered distributions). The key idea is to look at relationships between functions and their Fourier transforms (such as the Parseval identity) and define the Fourier transform in a way that preserves these relationships. For example, by Theorem 1.23, for all $f, g \in L^2(\mathbb{R})$ we have

$$\int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}} \widehat{f}(x)g(x)\,dx \tag{1.1.20}$$

Therefore, the Fourier transform of a general "function" f must satisfy (some variant of) equation (1.1.20) whenever g is a function whose Fourier transform $\widehat{g}$ is well-defined and the left-hand side of (1.1.20) makes sense. For instance, the function $f(x) = e^{itx}$ (with $t \in \mathbb{R}$) is neither integrable nor square-integrable, but the integral on the left-hand side:

$$\int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx \equiv \int_{\mathbb{R}} \widehat{g}(x)e^{itx}\,dx$$

is well-defined and equal (almost everywhere) to $2\pi g(t)$, provided that $g \in L^1(\mathbb{R})$ and $\widehat{g} \in L^1(\mathbb{R})$. In this case, the Fourier transform $\widehat{f}$ of f is defined as the "function" that satisfies:

$$2\pi g(t) = \int_{\mathbb{R}} \widehat{f}(x)g(x)\,dx$$

for any g satisfying certain conditions. However, it is important to note that no actual function satisfies this identity in the traditional sense. Eventually, the right-hand side integral must also be interpreted symbolically, no longer representing the usual Lebesgue integral. In other words, to define the Fourier transform of more general "functions," we:

1. Generalize the concept of the integral of the product of two functions by treating the integral

$$\int_{\mathbb{R}} f(x)g(x)\,dx$$

   as a bilinear form of the pair (f, g). This bilinear form is identical to the integral of fg whenever $fg \in L^1(\mathbb{R})$.

2. The Fourier transform of a general function f is defined as something that satisfies equation (1.1.20) whenever that identity makes sense.

This is essentially what the theory of tempered distributions is about, though we will not delve into it further here.

The bottom line is: everything we derived for the Fourier transform of f with $f \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$ is treated as identities that the Fourier transform of generalized functions must satisfy.

1.2 Poisson Summation Formula

Theorem 1.26. *If $f \in L^1(\mathbb{R})$, then the series*

$$\sum_{n=-\infty}^{\infty} f(t+2n\pi) \tag{1.2.1}$$

converges absolutely for almost all t in $(0, 2\pi)$ and its sum belongs to $L^1(0, 2\pi)$ and is 2π-periodic. If $\{a_n\}_{n\in\mathbb{Z}}$ denotes the Fourier coefficient of the series, then $a_n = \dfrac{1}{2\pi}\widehat{f}(n)$ for all $n \in \mathbb{Z}$.

Proof. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_0^{2\pi} \sum_{n=-\infty}^{\infty} \big|f(t+2n\pi)\big|\,dt &= \sum_{n=-\infty}^{\infty} \int_0^{2\pi} \big|f(t+2n\pi)\big|\,dt = \sum_{n=-\infty}^{\infty} \int_{2n\pi}^{2(n+1)\pi} \big|f(t)\big|\,dt \\ &= \int_{\mathbb{R}} \big|f(t)\big|\,dt = \|f\|_{L^1(\mathbb{R})} < \infty;\end{aligned}$$

thus the series $\sum\limits_{n=-\infty}^{\infty} |f(\cdot + 2n\pi)|$ belongs to $L^1(0, 2\pi)$ which shows that $\sum\limits_{n=-\infty}^{\infty} f(t+2n\pi)$ converges in $L^1(0, 2\pi)$ and also converges absolutely for a.a. $t \in (0, 2\pi)$.

By the definition of the Fourier coefficients,

$$\begin{aligned}a_m &= \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=-\infty}^{\infty} f(t+2n\pi)e^{-imt}\,dt = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_0^{2\pi} f(t+2n\pi)e^{-imt}\,dt \\ &= \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_{2n\pi}^{2(n+1)\pi} f(t)e^{-imt}\,dt = \frac{1}{2\pi}\int_{\mathbb{R}} f(t)e^{-imt}\,dt = \frac{1}{2\pi}\widehat{f}(m)\end{aligned}$$

which concludes the theorem. □

For $f \in L^1(\mathbb{R})$, define $F(t) = \sum\limits_{n=-\infty}^{\infty} f(t+2n\pi)$. Then Theorem 1.26 shows that $F \in L^1(0, 2\pi)$ and is 2π-periodic. If the Fourier series of F converges pointwise to F, then

$$\sum_{n=-\infty}^{\infty} f(t+2n\pi) = F(t) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi}\widehat{f}(m)e^{imt}. \tag{1.2.2}$$

In particular, letting $t = 0$ in (1.2.2) we obtain

$$\sum_{n=-\infty}^{\infty} f(2n\pi) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \widehat{f}(n).$$

This identity is called the Poisson summation formula. The following theorem provides a condition that (1.2.2) holds (if f is continuous).

Theorem 1.27. *If there exists $\delta > 0$ such that*

$$\big|f(x)\big| + \big|\widehat{f}(x)\big| \leqslant C(1+|x|)^{-1-\delta} \qquad \forall\, x \in \mathbb{R},$$

then

$$\sum_{n=-\infty}^{\infty} \check{\widehat{f}}(x+2n\pi) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx} \qquad \forall\, x \in [0, 2\pi]. \tag{1.2.3}$$

The proof of this theorem is left as an exercise.

1.3 Shannon Sampling Theorem

An analog signal is a piecewise continuous function of time defined on $\mathbb{R}$, with the exception of perhaps a countable number of jump discontinuities. Almost all analog signals of interest in engineering have finite energy. By this we mean that the signal is square-integrable. The norm of a signal $f : \mathbb{R} \to \mathbb{R}$ defined by

$$\|f\|_{L^2(\mathbb{R})} = \Big(\int_{\mathbb{R}} |f(t)|^2 \, dt \Big)^{\frac{1}{2}} \tag{1.3.1}$$

represents the square root of the total energy content of the signal f. The spectrum of a signal f is represented by its Fourier transform $\widehat{f}$, where the variable of $\widehat{f}$, usually denoted by ω, is called the frequency. The frequency is measured by $\nu = \dfrac{\omega}{2\pi}$ in terms of Hertz.

A signal f is called band-limited if its Fourier transform has a compact support; that is,

$$\widehat{f}(\omega) = 0 \qquad \text{whenever} \quad |\omega| > \omega_0 \tag{1.3.2}$$

for some $\omega_0 > 0$. If $\omega_0 > 0$ is the smallest value for which (1.3.2) holds, then it is called the bandwidth of the signal. Even if an analog signal f is not band-limited, we can reduce it to a band-limited signal by what is called an ideal low-pass filtering. To reduce f to a band-limited signal f_{ω_0} with bandwidth less than or equal to ω_0, we consider

$$\widehat{f_{\omega_0}}(\omega) = \begin{cases} \widehat{f}(\omega) & \text{if } |\omega| \leqslant \omega_0, \\ 0 & \text{if } |\omega| > \omega_0 \end{cases} \tag{1.3.3}$$

and we find the low-pass filter function f_{ω_0} by the inverse Fourier transform

$$f_{\omega_0}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{it\omega} \widehat{f_{\omega_0}}(\omega) \, d\omega = \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} e^{it\omega} \widehat{f}(\omega) \, d\omega .$$

Define the gate/window function

$$\Pi_{\omega_0}(\omega) = \mathbf{1}_{[-\omega_0,\omega_0]}(\omega) = \begin{cases} 1 & \text{if } |\omega| \leqslant \omega_0, \\ 0 & \text{if } |\omega| > \omega_0. \end{cases}$$

Then $\Pi_{\omega_0} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ whose inverse Fourier transform is given by

$$\widetilde{\Pi_{\omega_0}}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \Pi_{\omega_0}(\omega) e^{i\omega t} \, d\omega = \frac{\sin(\omega_0 t)}{\pi t} .$$

Therefore, by Theorem 1.12 we obtain that

$$f_{\omega_0}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \Pi_{\omega_0}(\omega) e^{it\omega} \widehat{f}(\omega) \, d\omega = (f * \widetilde{\Pi_{\omega_0}})(t) = \int_{\mathbb{R}} \frac{\sin(\omega_0(\tau - t))}{\pi(t-\tau)} f(\tau) \, d\tau .$$

This gives the sampling integral representation of a band-limited signal f_{ω_0}. Thus, $f_{\omega_0}(t)$ can be interpreted as the weighted average of f with the Fourier kernel $\dfrac{\sin \omega_0(t - \cdot)}{\pi(t - \cdot)}$ as weight.

In the field of digital signal processing, the ***sampling theorem*** is a fundamental bridge between continuous-time signals (often called "analog signals") and discrete-time signals (often called "digital signals"). It establishes a sufficient condition for a *sample rate* (取樣頻率) that permits a discrete sequence of samples to capture all the information from a continuous-time signal of finite bandwidth. To be more precise, Shannon's version of the theorem states that "if an analog signal contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $\dfrac{1}{2B}$ seconds apart."

Theorem 1.28. *Let $f:\mathbb{R}\to\mathbb{R}$ be a continuous integrable function. If* $\operatorname{supp}(\widehat{f})\subseteq[-\omega_0,\omega_0]$, *then f is fully determined by the sequence* $\Big\{f\Big(\dfrac{\pi k}{\omega_0}\Big)\Big\}_{k=-\infty}^{\infty}$, *and*

$$f(t)=\sum_{k=-\infty}^{\infty} f\Big(\frac{\pi k}{\omega_0}\Big)\operatorname{sinc}(\omega_0 t-k\pi)\qquad\forall\, t\in\mathbb{R},\tag{1.3.4}$$

where the sinc function is given by

$$\operatorname{sinc}(x)=\begin{cases}\dfrac{\sin x}{x} & \textit{if } x\neq 0,\\ 1 & \textit{if } x=0.\end{cases}\tag{1.3.5}$$

Proof. Let $f\in L^1(\mathbb{R})\cap\mathscr{C}(\mathbb{R};\mathbb{R})$ such that $\operatorname{supp}(\widehat{f})\subseteq[-B,B]$. Then $\widehat{f}\in L^1(\mathbb{R})$ (since it must be bounded by Theorem 1.5); thus the Fourier inversion formula implies that

$$f(t)=\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)e^{ix\omega}\,d\omega=\int_{-\omega_0}^{\omega_0}\widehat{f}(\omega)e^{ix\omega}\,d\omega\qquad\forall\, t\in\mathbb{R}.$$

In particular,

$$f\Big(\frac{k\pi}{\omega_0}\Big)=\frac{1}{2\pi}\int_{-\omega_0}^{\omega_0}\widehat{f}(\omega)e^{\frac{ik\pi\omega}{\omega_0}}\,d\omega.$$

Treating $\widehat{f}$ as a function defined on $[-\omega_0,\omega_0]$, the identity above shows that $\Big\{\dfrac{\pi}{\omega_0}f\Big(\dfrac{-k\pi}{\omega_0}\Big)\Big\}_{k=-\infty}^{\infty}$ is the Fourier coefficients of $\widehat{f}$.

Note that the boundedness of $\widehat{f}$ implies that $\widehat{f}\in L^2(-\omega_0,\omega_0)$. Therefore, if $g\in L^2(-\omega_0,\omega_0)$, the Parseval identity, together with the polarization identity, implies that

$$\frac{1}{2\omega_0}\int_{-\omega_0}^{\omega_0}\widehat{f}(\omega)\overline{g(\omega)}\,d\omega=\sum_{k=-\infty}^{\infty}\frac{\pi}{\omega_0}f\Big(\frac{-k\pi}{\omega_0}\Big)\,\overline{\widehat{g}_k},\tag{1.3.6}$$

where $\widehat{g}_k=\dfrac{1}{2\omega_0}\displaystyle\int_{-\omega_0}^{\omega_0}g(t)e^{-\frac{ik\pi x}{\omega_0}}\,dx$.

For each $t\in\mathbb{R}$, the Fourier coefficients of the function $g(\omega)=e^{-it\omega}$ is given by

$$\widehat{g}_k=\frac{1}{2\omega_0}\int_{-\omega_0}^{\omega_0}e^{-it\omega}e^{-\frac{ik\pi\omega}{\omega_0}}\,d\omega=\frac{1}{2\omega_0}\int_{-\omega_0}^{\omega_0}e^{-i\frac{(\omega_0 t+k\pi)}{\omega_0}\omega}\,d\omega=\operatorname{sinc}(\omega_0 t+k\pi);$$

thus the Fourier inversion formula and (1.3.6) imply that if $x\in\mathbb{R}$,

$$\begin{aligned}f(t)&=\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)e^{it\omega}\,d\omega=\frac{1}{2\pi}\int_{-\omega_0}^{\omega_0}\widehat{f}(\omega)\overline{g(\omega)}\,dx=\sum_{k=-\infty}^{\infty}f\Big(\frac{-k\pi}{\omega_0}\Big)\operatorname{sinc}(\omega_0 t+k\pi)\\&=\sum_{k=-\infty}^{\infty}f\Big(\frac{k\pi}{\omega_0}\Big)\operatorname{sinc}(\omega_0 t-k\pi).\end{aligned}$$

The identity above shows that f is fully determined by the sequence $\Big\{f\Big(\dfrac{k\pi}{\omega_0}\Big)\Big\}_{k=-\infty}^{\infty}$. ◻

Remark 1.29. Equation (1.3.4) is called the ***Whittaker–Shannon interpolation formula***.

Remark 1.30. The reconstruction formula (1.3.4) can be obtained formally as follows. In the proof of the sampling theorem, since we have obtained that the Fourier coefficient of $\widehat{f}$ is $\left\{\frac{\pi}{\omega_0} f\left(\frac{-k\pi}{\omega_0}\right)\right\}_{k=-\infty}^{\infty}$, so we have

$$\widehat{f}(\omega) = \sum_{k=-\infty}^{\infty} \frac{\pi}{\omega_0} f\left(\frac{-k\pi}{\omega_0}\right) e^{\frac{ik\pi\omega}{\omega_0}} = \sum_{k=-\infty}^{\infty} \frac{\pi}{\omega_0} f\left(\frac{k\pi}{\omega_0}\right) e^{-\frac{ik\pi\omega}{\omega_0}}.$$

Taking the inverse Fourier transform,

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} \frac{\pi}{\omega_0} f\left(\frac{k\pi}{\omega_0}\right) e^{-\frac{ik\pi\omega}{\omega_0}} e^{it\omega}\, d\omega = \frac{1}{2\omega_0} \int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\omega_0}\right) e^{-\frac{ik\pi\omega}{\omega_0}} e^{it\omega}\, d\omega;$$

thus if we can switch the order of the infinite sum and the integration, we then immediately obtain the reconstruction formula

$$f(t) = \frac{1}{2\omega_0} \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\omega_0}\right) \int_{\mathbb{R}} e^{-\frac{ik\pi\omega}{\omega_0}} e^{it\omega}\, d\omega = \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\omega_0}\right) \operatorname{sinc}(\omega_0 t - k\pi).$$

In fact, this kind of "switch of order of the infinite sum and the integral" is usually valid when the integral is due to the Fourier transform/inverse transform. In particular, later we will take the Fourier transform of series of the form

$$\sum_{n=-\infty}^{\infty} c_n \phi(x-n) = \sum_{n=-\infty}^{\infty} c_n (T_n \phi)(x)$$

where $\phi \in L^2(\mathbb{R})$ and $\{c_n\}_{n=-\infty}^{\infty} \in \ell^2$; that is, $\sum\limits_{n=-\infty}^{\infty} |c_n|^2 < \infty$, and directly switch the order of summation and the Fourier transform as

$$\mathscr{F}\left[\sum_{n=-\infty}^{\infty} c_n (T_n \phi)\right] = \sum_{n=-\infty}^{\infty} c_n \mathscr{F}[T_n \phi].$$

A rigorous proof of the identity above will be ignored.

1.4 Time-Frequency Analysis

In signal processing, time-frequency analysis comprises those techniques that study a signal in both the time and frequency domains simultaneously, using various time-frequency representations. The mathematical motivation for this study is that functions and their transform representation are tightly connected, and they can be understood better by studying them jointly, as a two-dimensional object, rather than separately. A simple example is that the 4-fold periodicity of the Fourier transform - and the fact that two-fold Fourier transform reverses direction - can be interpreted by considering the Fourier transform as a 90° rotation in the associated time-frequency plane: 4 such rotations yield the identity, and 2 such rotations simply reverse direction (reflection through the origin).

The practical motivation for time-frequency analysis is that classical Fourier analysis is quite inadequate for most applications. In the first place, the Fourier analysis assumes that signals are infinite in time or periodic, while many signals in practice are of short duration, and change substantially over their duration. For example, traditional musical instruments do not produce infinite duration sinusoids, but instead begin with an attack, then gradually decay. Moreover, to extract the spectral information $\widehat{f}(\omega)$ from the analog signal f, the Fourier transform takes an infinite amount of time, using both past and future information of the signal just to evaluate the spectrum at a single frequency ω. Besides, the formula

$$(\mathscr{F} f)(\omega) = \widehat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-it\omega}dt \qquad \forall\, \omega \in \mathbb{R} \tag{1.1.10}$$

does not even reflect frequencies that evolve with time. What is really needed is for one to be able to determine the time intervals that yield the spectral information on any desirable range of frequencies (or frequency band). In addition, since the frequency of a signal is directly proportional to the length of its cycle, it follows that for high-frequency spectral information, the time-interval should be relatively small to give better accuracy, and for low-frequency spectral information, the time-interval should be relatively wide to give complete information. In other words, it is important to have a flexible time-frequency window that automatically narrows at high "center-frequency" and widens at low "center-frequency". All these issues are poorly represented by traditional methods, which motivates time-frequency analysis.

In this lecture, concerning the time-frequency analysis we are going to study the following subjects:

1. **The Gabor Transform/Short-Time Fourier Transform (STFT)**: For a given $f \in L^2(\mathbb{R})$, we consider the following integral

   $$\mathcal{G}[f](t, \omega) = \int_{\mathbb{R}} f(\tau)\overline{g_{t,\omega}(\tau)}\, d\tau\,,$$

   where $g_{t,\omega} \in L^2(\mathbb{R})$ is a time-localization window function taking the form

   $$g_{t,\omega}(\tau) = \phi(\tau - t)e^{i\omega\tau} = (M_\omega T_t \phi)(\tau) \tag{1.4.1}$$

   for some $\phi \in L^2(\mathbb{R})$ and is used for extracting local information from a Fourier transform of a signal f.

   We are going to study the basic properties of the Gabor transform and derive the inversion formula

   $$f(t) = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i\omega t}\mathcal{G}[f](b, \omega)\phi(t - b)\, d\omega db = \int_{\mathbb{R}} \mathscr{F}^{-1}\big[\mathcal{G}[f](b, \cdot)\big](t)(T_b\phi)(x)\, db.$$

2. **The Wigner-Ville Distribution (WVD) and Transform (WVT)**: The cross Wigner-Ville Distribution of $f, g \in L^2(\mathbb{R})$ is the integral

$$W_{f,g}(t,\omega) = \int_{\mathbb{R}} f\Big(t + \frac{\tau}{2}\Big)\overline{g}\Big(t - \frac{\tau}{2}\Big)e^{-i\omega\tau}\, d\tau\,. \tag{1.4.2}$$

At the first glance it looks very similar to the STFT since with the substitution of variable $x = t + \frac{\tau}{2}$,

$$W_{f,g}(t,\omega) = 2\int_{\mathbb{R}} f(x)\overline{g}(x-2t)e^{-2i(x-t)\omega}\, dx = 2e^{2it\omega}\int_{\mathbb{R}} f(x)\overline{g}(x-2t)e^{-2ix\omega}\, dx$$

and the last integral is indeed a STFT of f with window function g (at the point $(2t, 2\omega)$). When the windows function is the signal itself, it is called the Wigner-Ville Transform of the signal. In other words, the Wigner-Ville Transform of $f \in L^2(\mathbb{R})$ is $W_{f,f}$ (which is also denoted by W_f in the textbook). We note that WVT is a nonlinear map of the input f.

The WVD is a great method to perform time-frequency analysis; however, the study of the Wigner-Ville transform is beyond the scope of this course, so we will only talk about this breifly.

3. **The Wavelet Transform**: A given function $\psi \in L^2(\mathbb{R})$ satisfying the "admissibility" condition

$$C_\psi = \int_{\mathbb{R}} \frac{|\widehat{\psi}(\omega)|^2}{|\omega|}\, d\omega < \infty\,,$$

is called a basic wavelet. The "integral wavelet transform" relative to the "basic wavelet" ψ, denoted by W_ψ (do not confuse with the Wigner-Ville transform), is an integral operator on $L^2(\mathbb{R})$ defined by

$$W_\psi[f](a,b) = |a|^{-\frac{1}{2}}\int_{\mathbb{R}} f(x)\overline{\psi\Big(\frac{t-b}{a}\Big)}\, dt = \int_{\mathbb{R}} f(t)\overline{(T_b D_a \psi)(t)}\, dt\,.$$

We are going to study the basic properties of the wavelet transform and derive the inversion formula

$$f(t) = \frac{1}{C_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} W_\psi[f](a,b)(T_b D_a\psi)(t)\,\frac{dbda}{a^2}\,. \tag{1.4.3}$$

1.5 The Wavelet Series and Frames

For the purpose of localization, the basis wavelet is usually chosen to have compact support or decay very fast, similar to the case of Fourier transform v.s. Fourier series, we might be able to reconstruct f using only values of the wavelet transform at discrete points. Throughout the lecture, any function $\psi \in L^2(\mathbb{R})$ is associated with a bi-infinite sequence of functions $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ defined by

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) \qquad \forall\, j,k \in \mathbb{Z}\,. \tag{1.5.1}$$

We note that the index j measures the width of the window/support, and the parameter k is used to represent translation of the window to cover the whole time domain.

Definition 1.31. A function $\psi \in L^2(\mathbb{R})$ is called ***an orthogonal wavelet*** (or o.n. wavelet), if the family $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$, as defined in (1.5.1), is an orthonormal basis of $L^2(\mathbb{R})$; that is,

$$\langle \psi_{j,k}, \psi_{\ell,m} \rangle_{L^2(\mathbb{R})} = \delta_{j\ell}\delta_{km} \qquad \forall\, j,k,\ell,m \in \mathbb{Z} \tag{1.5.2}$$

and every function $f \in L^2(\mathbb{R})$ can be written as

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(x), \tag{1.5.3}$$

where $c_{j,k} = \langle f, \psi_{j,k} \rangle_{L^2(\mathbb{R})}$ and the convergence of the series in (1.5.3) is in $L^2(\mathbb{R})$; that is,

$$\lim_{M_1,N_1,M_2,N_2\to\infty} \Big\| f - \sum_{j=-M_2}^{N_2} \sum_{k=-M_1}^{N_1} c_{j,k}\psi_{j,k} \Big\|_{L^2(\mathbb{R})} = 0\,.$$

Any arbitrarily given $\psi \in L^2(\mathbb{R})$ is most likely not an o.n. wavelet. For a given function $\psi \in L^2(\mathbb{R})$, the first few questions we would like to answer are

1. Is the linear span of $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ dense in $L^2(\mathbb{R})$?
2. If so, is there a effect way to express a function $f \in L^2(\mathbb{R})$ in terms of "linear combinations" of $\psi_{j,k}$'s?
3. Is the expression of a function $f \in L^2(\mathbb{R})$ as a "linear combinations" of $\psi_{j,k}$'s unique?

The abstraction of the aforementioned questions leads to the concept of frames and frame operators defined below.

Definition 1.32 (Frame)**.** A sequence $\{x_n\}$ in a separable Hilbert space $(H, \langle\cdot,\cdot\rangle)$ (not necessarily a basis of H) is called a ***frame*** if there exist two numbers A and B with $0 < A \leqslant B < \infty$ such that

$$A\|x\|^2 \leqslant \sum_n \big|\langle x, x_n \rangle\big|^2 \leqslant B\|x\|^2\,. \tag{1.5.4}$$

The numbers A and B are called the frame bounds.

Definition 1.33 (Frame Operator)**.** To each frame $\{x_n\}$ there corresponds an operator T, called the ***frame operator***, from H into itself defined by

$$Tx = \sum_n \langle x, x_n \rangle x_n \qquad \forall\, x \in H\,. \tag{1.5.5}$$

For a given frame $\{x_n\}$ of a Hilbert space, it can be shown that the frame operator is bounded, and eventually answer (partially) the questions above by the following

Theorem 1.34. *Suppose $\{x_n\}_{n=1}^{\infty}$ is a frame on a separable Hilbert space with frame bounds A and B, and T is the corresponding frame operator. Then,*

(a) *T is invertible and $B^{-1}\mathrm{I} \leqslant T^{-1} \leqslant A^{-1}\mathrm{I}$.*

(b) *$\{T^{-1}x_n\}_{n=1}^{\infty}$ is a frame, called the dual frame of $\{x_n\}_{n=1}^{\infty}$, with frame bounds B^{-1} and A^{-1}.*

(c) *Every $x \in H$ can be expressed in the form*

$$x = \sum_{n=1}^{\infty} \langle x, T^{-1}x_n \rangle x_n = \sum_{n=1}^{\infty} \langle x, x_n \rangle T^{-1}x_n. \tag{1.5.6}$$

Later we will prove the following theorem which provides a sufficient condition of a given ψ to generate a frame $\{\psi_{m,n}\}_{m,n}$ in $L^2(\mathbb{R})$.

Theorem 1.35. *Let $\phi \in L^2(\mathbb{R})$, and $a_0 > 1$. If*

(i) *there exist A, $B > 0$ such that $A \leqslant \sum_{m=-\infty}^{\infty} \left|\widehat{\phi}(a_0^m \omega)\right|^2 \leqslant B$ for all $1 \leqslant \omega \leqslant a_0$, and*

(ii) $\sup_{\omega \in \mathbb{R}} \sum_{m=-\infty}^{\infty} \left|\widehat{\phi}(a_0^m)\right| \left|\widehat{\phi}(a_0^m \omega + x)\right| \leqslant C(1+|x|)^{-(1+\delta)}$ *for some constants C and $\delta > 0$,*

then there exists $\widetilde{b} > 0$ such that the family $\{\phi_{m,n}\}_{m,n\in\mathbb{Z}}$ given by

$$\phi_{m,n} = a_0^{-m/2} \phi(a_0^{-m}x - nb_0) = (D_{a_0^m} T_{nb_0} \phi)(x)$$

forms a frame in $L^2(\mathbb{R})$ for any $b_0 \in (0, \widetilde{b})$.

Suppose that a function $\psi \in L^2(\mathbb{R})$ generates a frame $\mathbf{\Psi} = \{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ in $L^2(\mathbb{R})$. Even though in theory we have a representation formula (1.5.6) using the frame operator T associated with $\mathbf{\Psi}$, it is still not practical enough since it requires to compute the dual frame $\{T^{-1}\psi_{m,n}\}_{m,n\in\mathbb{Z}}$. We hope, just like how we obtain $\mathbf{\Psi}$, that the dual frame is a frame simply generated by a function $\widetilde{\psi} \in L^2(\mathbb{R})$; that is,

$$T^{-1}\psi_{m,n} = \widetilde{\psi}_{m,n} \qquad \forall\, m, n \in \mathbb{Z}.$$

However, this requires a subtle design of the function ψ.

Definition 1.36. A function $\psi \in L^2(\mathbb{R})$ is called an $\mathcal{R}$-function if $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$, as defined in (1.5.1), is a Riesz basis of $L^2(\mathbb{R})$, in the sense that the linear span of $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is dense in $L^2(\mathbb{R})$ and that there exist positive constants A and B, with $0 < A \leqslant B < \infty$, such that

$$A\left\|\{c_{j,k}\}_{j,k\in\mathbb{Z}}\right\|_{\ell^2}^2 \leqslant \left\| \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k} \right\|_{L^2(\mathbb{R})}^2 \leqslant B\left\|\{c_{j,k}\}_{j,k\in\mathbb{Z}}\right\|_{\ell^2}^2 \tag{1.5.7}$$

for all doubly bi-infinite square-summable sequences $\{c_{j,k}\}_{j,k\in\mathbb{Z}}$; that is,

$$\left\|\{c_{j,k}\}_{j,k\in\mathbb{Z}}\right\|_{\ell^2}^2 \equiv \sum_{j,k=-\infty}^{\infty} |c_{j,k}|^2 < \infty\,.$$

An $\mathcal{R}$-function $\psi \in L^2(\mathbb{R})$ is called an $\mathcal{R}$-wavelet (or wavelet), if there exists a function $\widetilde{\psi} \in L^2(\mathbb{R})$, such that $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ and $\{\widetilde{\psi}_{j,k}\}_{j,k\in\mathbb{Z}}$, as defined in (1.5.1), are dual bases of $L^2(\mathbb{R})$. If ψ is an $\mathcal{R}$-wavelet, then $\widetilde{\psi}$ is called a dual wavelet corresponding to ψ.

Remark 1.37. A collection of functions $\{\phi_n\}_{n\in\mathbb{Z}}$ in $L^2(\mathbb{R})$ is said to have Riesz bounds A and B, where $0 < A \leqslant B < \infty$, if

$$A\left\|\{c_n\}_{n\in\mathbb{Z}}\right\|_{\ell^2}^2 \leqslant \left\|\sum_{n=-\infty}^{\infty} c_n\phi_n\right\|_{L^2(\mathbb{R})}^2 \leqslant B\left\|\{c_n\}_{n\in\mathbb{Z}}\right\|_{\ell^2}^2 \qquad \forall\,\{c_n\}_{n\in\mathbb{Z}} \in \ell^2\,. \tag{1.5.8}$$

An orthonormal system in $L^2(\mathbb{R})$ indeed has Riesz bounds 1 and 1. Similar to (1.1.18), the argument used to establish (1.1.18) can be applied to show that

$$\mathscr{F}\left[\sum_{n=-\infty}^{\infty} c_n\phi_n\right] = \sum_{n=-\infty}^{\infty} c_n\widehat{\phi_n} \quad \forall\,\{c_n\}_{n\in\mathbb{Z}} \in \ell^2 \text{ and } \{\phi_n\}_{n\in\mathbb{Z}} \text{ satisfying (1.5.8)}. \tag{1.5.9}$$

The key difference in the argument is that one has to modify (1.1.19) using (1.5.8):

$$\|\widehat{s_k} - \widehat{s_\ell}\|_{L^2(\mathbb{R})}^2 = 2\pi\|s_k - s_\ell\|_{L^2(\mathbb{R})}^2 \leqslant 2\pi B \sum_{\min\{k,\ell\}<|n|\leqslant\max\{k,\ell\}} |c_n|^2 \to 0 \quad \text{as } k,\ell\to\infty\,.$$

We also remark that using (1.5.9), if $\{\phi_n\}_{n\in\mathbb{Z}} \subseteq L^2(\mathbb{R})$ has Riesz bounds A and B, then $\{\widehat{\phi_n}\}_{n\in\mathbb{Z}}$ has Riesz bounds $2\pi A$ and $2\pi B$ because of the Plancherel identity.

Now, the question is how do we construct such kind of wavelet?

1.6 Multi-resolution Analysis and Construction of O.N. Wavelets

A wavelet ψ in $L^2(\mathbb{R})$ is called a semi-orthogonal wavelet (or wavelet) if the family $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ it generates satisfies

$$\langle\psi_{j,k},\psi_{\ell,m}\rangle = 0 \qquad \text{if } j \neq \ell,\ j,k,\ell,m \in \mathbb{Z}. \tag{1.6.1}$$

Suppose that ψ is a semi-orthogonal wavelet and consider the family $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ it generates. For each $j \in \mathbb{Z}$, let W_j denote the closure of the linear span of $\{\psi_{j,k}\}_{k\in\mathbb{Z}}$, namely:

$$W_j \equiv \text{closure}_{\|\cdot\|_2}\big(\{\psi_{j,k}\,\big|\,k\in\mathbb{Z}\}\big)\,. \tag{1.6.2}$$

Then it is clear that $W_j \perp W_\ell$ if $j \neq \ell$, meaning that

$$\langle g_j, g_\ell\rangle = 0 \qquad \text{if } j \neq \ell,\ g_j \in W_j \text{ and } g_\ell \in W_\ell. \tag{1.6.3}$$

Moreover, by the fact that the linear span of $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is dense in $L^2(\mathbb{R})$, $L^2(\mathbb{R})$ can be decomposed as a direct sum of the spaces W_j:

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots, \tag{1.6.4}$$

in the sense that every function $f \in L^2(\mathbb{R})$ has a unique decomposition:

$$f(x) = \sum_{j=-\infty}^{\infty} g_j(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots, \tag{1.6.5}$$

where $g_j = \sum_{k=-\infty}^{\infty} \langle f, \widetilde{\psi}_{j,k}\rangle \psi_{j,k} \in W_j$ for all $j \in \mathbb{Z}$.

For each $j \in \mathbb{Z}$, let us consider the closed subspaces

$$V_j \equiv \bigoplus_{\ell=-\infty}^{j-1} W_\ell = \cdots \oplus W_{j-2} \oplus W_{j-1} \tag{1.6.6}$$

of $L^2(\mathbb{R})$. These subspaces clearly have the following properties:

(1°) $\cdots \subsetneq V_{-1} \subsetneq V_0 \subsetneq V_1 \subsetneq \cdots$ or $V_j \subsetneq V_{j+1}$ for all $j \in \mathbb{Z}$;

(2°) $\text{closure}_{\|\cdot\|_2}\Big(\bigcup_{j\in\mathbb{Z}} V_j\Big) = L^2(\mathbb{R})$;

(3°) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;

(4°) $V_{j+1} = V_j \oplus W_j$ for all $j \in \mathbb{Z}$; and

(5°) For all $j \in \mathbb{Z}$, $f \in V_j$ if and only if $d_{1/2}f \in V_{j+1}$, where for $\lambda > 0$ the dilation operator d_λ is given by $(d_\lambda f)(x) = f(\lambda^{-1}x)$.

Hence, in contrast to the subspaces W_j which satisfy

$$W_j \cap W_\ell = \{0\} \qquad \text{if } j \neq \ell,$$

the sequence of subspaces V_j is nested, as described by (1°), and has the property that every function f in $L^2(\mathbb{R})$ can be approximated as closely as desirable by its projections $P_j f$ in V_j, as described by (2°). But on the other hand, by decreasing j, the projections $P_j f$ could have arbitrarily small energy, as guaranteed by (3°). What is not described by (1°)-(3°) is the most important intrinsic property of these spaces which is that more and more "variations" of $P_j f$ are removed as $j \to -\infty$. In fact, these variations are peeled off, level by level in decreasing order of the "rate of variations" (better known as "frequency bands") and stored in the complementary subspaces W_j as in (4°). This process can be made very efficient by an application of the property (5°).

In fact, if the reference subspace V_0, say, is generated by a single function $\phi \in L^2(\mathbb{R})$ in the sense that

$$V_0 \equiv \text{closure}_{\|\cdot\|_2}\big(\{\phi_{0,k} \,|\, k \in \mathbb{Z}\}\big), \tag{1.6.7}$$

where

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \tag{1.6.8}$$

then all the subspaces V_j are also generated by the same ϕ (just as the subspaces W_j are generated by as ψ in (1.6.2)), namely:

$$V_j \equiv \text{closure}_{\|\cdot\|_2}\left(\left\{\phi_{j,k} \,\middle|\, k \in \mathbb{Z}\right\}\right) \qquad \forall\, j \in \mathbb{Z}. \tag{1.6.9}$$

Hence, the "peeling-off" process from V_j to W_{j-1}, W_{j-2}, $\cdots$, $W_{j-\ell}$ can be accomplished efficiently.

Definition 1.38. An MRA consists of a sequence $\{V_m \,|\, m \in \mathbb{Z}\}$ of embedded closed subspaces of $L^2(\mathbb{R})$ that satisfy the following conditions:

(i) $\cdots \subseteq V_{-2} \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_m \subseteq V_{m+1} \subseteq \cdots$;

(ii) $\bigcup\limits_{m=-\infty}^{\infty} V_m$ is dense in $L^2(\mathbb{R})$; that is, $\text{closure}_{\|\cdot\|_2}\left(\bigcup\limits_{m=-\infty}^{\infty} V_m\right) = L^2(\mathbb{R})$.

(iii) $\bigcap\limits_{m=\infty}^{\infty} V_m = \{0\}$.

(iv) $f \in V_m$ if and only if $d_{1/2}f \in V_{m+1}$ for all $m \in Z$;

(v) there exists a function $\phi \in V_0$ such that $\left\{\phi_{0,n} = T_n\phi \,\middle|\, n \in \mathbb{Z}\right\}$ is an orthonormal basis for V_0; that is,

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{n=-\infty}^{\infty} \left|\langle f, \phi_n\rangle\right|^2 \qquad \forall\, f \in V_0\,.$$

The function ϕ is called the scaling function or father wavelet. If $\{V_m\}_{m\in\mathbb{Z}}$ is a multi-resolution of $L^2(\mathbb{R})$ and if V_0 is the closed subspace generated by the integer translates of a single function ϕ, then we say that ϕ generates the MRA.

Remark 1.39. To define scaling function, sometimes condition (v) is relaxed by assuming that $\left\{T_n\phi \,\middle|\, n \in \mathbb{Z}\right\}$ is a Riesz basis for V_0. Nevertheless, in most of the applications we look for ϕ such that $\{T_n\phi \,|\, n \in \mathbb{Z}\}$ is an orthonormal basis of V_0, so we simply use Definition 1.38 (which is the one used in the textbook) for the scaling functions.

Theorem 1.40 (Orthonormalization Process)**.** *If $\phi \in L^2(\mathbb{R})$ and if $\{T_n\phi \,|\, n \in \mathbb{Z}\}$ is a Riesz basis of V_0, then $\left\{T_n\widetilde{\phi} \,\middle|\, n \in \mathbb{Z}\right\}$ is an orthonormal basis of V_0 with*

$$\widehat{\widetilde{\phi}}(\omega) = \frac{\widehat{\phi}(\omega)}{\sqrt{\sum\limits_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + 2k\pi)\right|^2}}. \tag{1.6.10}$$

Theorem 1.41. *If $\{V_n\}_{n\in\mathbb{Z}}$ is an MRA with the scaling function ϕ, then there is an orthogonal wavelet ψ given by*

$$\psi(x) = \sum_{n=-\infty}^{\infty} (-1)^{-n-1}\overline{c_{-n-1}}(D_{1/2}T_n\phi)(x)\,, \tag{1.6.11}$$

where the coefficients c_n are given by

$$c_n = \langle \phi, \phi_{1,n} \rangle = \sqrt{2}\int_{\mathbb{R}} \phi(x)\overline{\phi(2x-n)}\,dx\,. \tag{1.6.12}$$

We will also give some examples of constructing orthogonal wavelet based on the theorems above.

Chapter 6

The Wavelet Transforms and Their Basic Properties

6.2 Continuous Wavelet Transforms and Examples

Definition 6.1 (Wavelet)**.** A wavelet is a function $\psi \in L^2(\mathbb{R})$ which satisfies the admissibility condition

$$C_\psi \equiv \int_{\mathbb{R}} \frac{\left|\widehat{\psi}(\omega)\right|^2}{|\omega|}\, d\omega < \infty, \tag{6.2.1}$$

where $\widehat{\psi}$ is the Fourier transform of ψ.

If $\psi \in L^2(\mathbb{R})$, then $\psi_{a,b} \in L^2(\mathbb{R})$ for all a and b since

$$\|\psi_{a,b}\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \frac{1}{|a|}\left|\psi\Big(\frac{t-b}{a}\Big)\right|^2 dt = \int_{\mathbb{R}} \left|\psi(x)\right|^2 dx = \|\psi\|^2_{L^2(\mathbb{R})} < \infty.$$

The Fourier transform of $\psi_{a,b}$ is given by

$$\widehat{\psi_{a,b}}(\omega) = \big(M_{-b}D_{\frac{1}{a}}\widehat{\psi}\big)(\omega) = \sqrt{|a|}e^{-ib\omega}\widehat{\psi}(a\omega). \tag{6.2.2}$$

Definition 6.2 (Continuous Wavelet Transform)**.** If $\psi \in L^2(\mathbb{R})$, and $\psi_{a,b}$ is given by

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\Big(\frac{t-b}{a}\Big) = (T_b D_a \psi)(t), \tag{6.2.3}$$

then the integral transformation W_ψ defined on $L^2(\mathbb{R})$ by

$$W_\psi[f](a,b) = \langle f, \psi_{a,b}\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} f(t)\overline{\psi_{a,b}(t)}\, dt \tag{6.2.4}$$

is called a continuous wavelet transform of f (relative to the wavelet ψ).

Using the Plancherel identity of the Fourier transform, it also follows from (6.2.4) that

$$\begin{aligned} W_\psi[f](a,b) = \langle f, \psi_{a,b}\rangle_{L^2(\mathbb{R})} &= \frac{1}{2\pi}\big\langle \widehat{f}, \widehat{\psi_{a,b}}\big\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_{\mathbb{R}} \sqrt{|a|}\widehat{f}(\omega)\overline{\widehat{\psi}}(a\omega)e^{ib\omega}\, d\omega \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} (\widehat{f}D_{\frac{1}{a}}\overline{\widehat{\psi}})(\omega)e^{ib\omega}\, d\omega. \end{aligned}$$

This means that for a fixed a $W_\psi[f](a,\cdot)$ is the inverse Fourier transform of the function $\widehat{f} D_{\frac{1}{a}} \overline{\widehat{\psi}}$ so the Fourier inversion formula shows that

$$\mathscr{F}\big[W_\psi[f](a,\cdot)\big](\omega) = \int_{\mathbb{R}} W_\psi[f](a,t)e^{-it\omega}\,dt = (\widehat{f} D_{\frac{1}{a}} \overline{\widehat{\psi}})(\omega) = \sqrt{|a|}\,\widehat{f}(\omega)\overline{\widehat{\psi}}(a\omega). \qquad (6.2.5)$$

Example 6.3 (The Haar Wavelet)**.** The Haar wavelet (Haar 1910) is one of the classic examples. It is defined by

$$\psi(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < \frac{1}{2}, \\ -1 & \text{if } \frac{1}{2} \leqslant t < 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (6.2.6)$$

The Haar wavelet has compact support. It is obvious that

$$\int_{\mathbb{R}} \psi(t)\,dt = 0, \qquad \int_{\mathbb{R}} \big|\psi(t)\big|^2\,dt = 1.$$

This wavelet is very well localized in the time domain, but it is not continuous. Its Fourier transform $\widehat{\psi}$ is calculated as follows: for $\omega \neq 0$ we have

$$\begin{aligned} \widehat{\psi}(\omega) &= \int_0^{\frac{1}{2}} e^{-i\omega t}\,dt - \int_{\frac{1}{2}}^{1} e^{-i\omega t}\,dt = \frac{e^{-i\omega t}}{-i\omega}\bigg|_{t=0}^{t=\frac{1}{2}} - \frac{e^{-i\omega t}}{-i\omega}\bigg|_{t=\frac{1}{2}}^{t=1} = \frac{i}{\omega}\Big(2e^{\frac{-i\omega}{2}} - 1 - e^{-i\omega}\Big) \\ &= \frac{-i}{\omega}\big(1 - e^{\frac{-i\omega}{2}}\big)^2 = \frac{-i}{\omega}\exp\Big(-\frac{i\omega}{2}\Big)\big(e^{\frac{i\omega}{4}} - e^{\frac{-i\omega}{4}}\big)^2 \\ &= i\exp\Big(-\frac{i\omega}{2}\Big)\frac{\sin^2(\omega/4)}{\omega/4} \end{aligned} \qquad (6.2.7)$$

and for $\omega = 0$ we have $\widehat{\psi}(0) = 0$; thus

$$\int_{\mathbb{R}} \frac{\big|\widehat{\psi}(\omega)\big|^2}{\omega}\,d\omega = 16\int_{\mathbb{R}} |\omega|^{-3}\sin^4\frac{\omega}{4}\,d\omega < \infty.$$

Both ψ and $\widehat{\psi}$ are plotted in Figure 6.1.

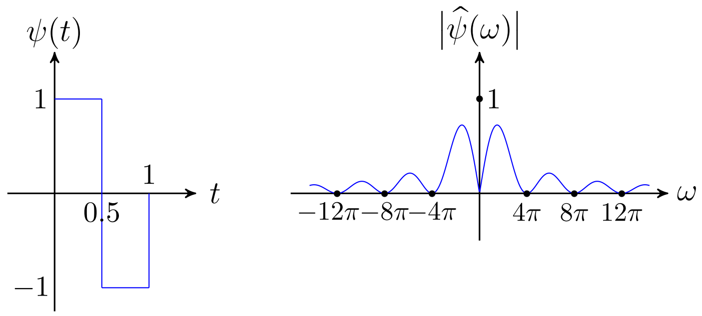

Figure 6.1: The Haar wavelet and its Fourier transform

Theorem 6.4. *If ψ is a wavelet and ϕ is a bounded integrable function, then the convolution function $\psi * \phi$ is a wavelet.*

Proof. By Young's inequality,

$$\|\psi * \phi\|_{L^2(\mathbb{R})} \leqslant \|\phi\|_{L^1(\mathbb{R})}\|\psi\|_{L^2(\mathbb{R})} < \infty,$$

so $\psi * \phi \in L^2(\mathbb{R})$. Moreover, by the fact that $\|\widehat{\phi}\|_{L^\infty(\mathbb{R})} \leqslant \|\phi\|_{L^1(\mathbb{R})}$,

$$\int_{\mathbb{R}} \frac{\left|\widehat{\psi * \phi}(\omega)\right|^2}{|\omega|}\, d\omega = \int_{\mathbb{R}} \frac{\left|\widehat{\psi}(\omega)\right|^2 \left|\widehat{\phi}(\omega)\right|^2}{|\omega|}\, d\omega \leqslant \|\phi\|^2_{L^1(\mathbb{R})} C_\psi < \infty.$$

Thus, the convolution function $\psi * \phi$ is a wavelet. □

Example 6.5. This example illustrates how to generate other wavelets by using Theorem 6.4. For example, if we take the Haar wavelet and convolute it with the following function

$$\phi(t) = \mathbf{1}_{[0,1]}(t) = \begin{cases} 1 & \text{if } 0 \leqslant t \leqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

we obtain a simple wavelet, as shown in Figure 6.2.

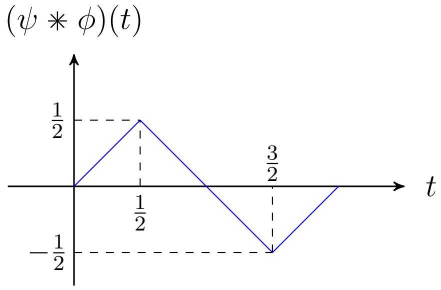

Figure 6.2: The wavelet $\psi * \phi$

Example 6.6. The convolution of the Haar wavelet with $\phi(t) = \exp(-t^2)$ generates a smooth wavelet, as shown in Figure 6.3.

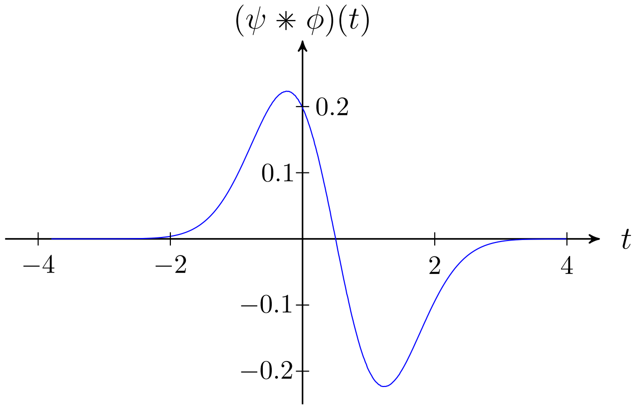

Figure 6.3: The wavelet $\psi * \phi$

Example 6.7 (The Mexican Hat Wavelet)**.** The Mexican hat wavelet is defined by the second derivative of a Gaussian function as

$$\begin{aligned} \psi(t) &= (1 - t^2)\exp\left(-\frac{t^2}{2}\right) = -\frac{d^2}{dt^2}\exp\left(-\frac{t^2}{2}\right) = \psi_{1,0}(t), \\ \widehat{\psi}(\omega) &= \widehat{\psi_{1,0}}(\omega) = \sqrt{2\pi}\,\omega^2 \exp\left(-\frac{\omega^2}{2}\right). \end{aligned} \tag{6.2.8}$$

The Mexican hat wavelet $\psi_{1,0}$ and its Fourier transform are shown in Figure 6.4(a)(b). This wavelet has excellent localization in time and frequency domains and clearly satisfies the admissibility condition.

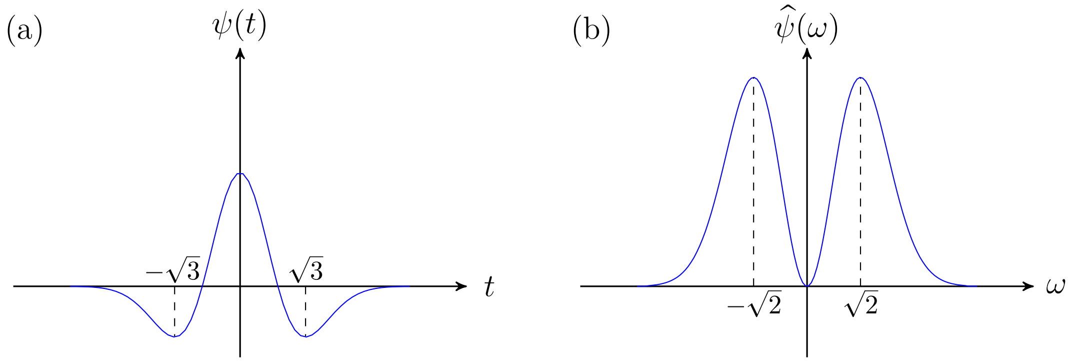

Figure 6.4: (a) The Mexican hat wavelet $\psi_{1,0}$ and (b) its Fourier transform $\widehat{\psi_{1,0}}$

Two other wavelets, $\psi_{\frac{3}{2},-2}$ and $\psi_{\frac{1}{4},\sqrt{2}}$, from the mother wavelet (6.2.8) can be obtained. These three wavelets, $\psi_{1,0}$, $\psi_{\frac{3}{2},-2}$, and $\psi_{\frac{1}{4},\sqrt{2}}$, are shown in Figure 6.5(i), (ii), and (iii), respectively.

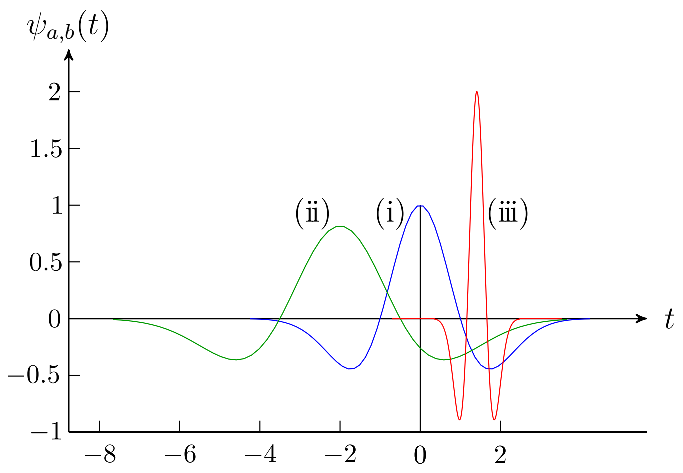

Figure 6.5: Three wavelets $\psi_{1,0}$, $\psi_{\frac{3}{2},-2}$, $\psi_{\frac{1}{4},\sqrt{2}}$

Remark 6.8. If in addition $\psi \in L^1(\mathbb{R})$ (which is usually the case if ψ has rapid decay), then its Fourier transform $\widehat{\psi}$ is bounded continuous. Since $\widehat{\psi}$ is continuous, C_ψ can be finite only if $\widehat{\psi}(0) = 0$ or, equivalently, $\int_{\mathbb{R}} \psi(t)\,dt = 0$. This means that ψ must be an oscillatory function with zero mean.

6.3 Basic Properties of Wavelet Transforms

The following theorem gives several properties of continuous wavelet transforms.

Theorem 6.9. *If ψ and ϕ are wavelets and f, g are functions which belong to $L^2(\mathbb{R})$, then*

(i) (Linearity) *For any scalars α and β,*

$$W_\psi[\alpha f + \beta g] = \alpha W_\psi[f] + \beta W_\psi[g].$$

(ii) (Translation) *With* T_c *denoting the translation operator defined by* $(T_c f)(t) = f(t-c)$,

$$W_\psi[T_c f](a,b) = W_\psi[f](a,b-c) \qquad \textit{and} \qquad W_{T_c\psi}[f](a,b) = W_\psi[f](a,b+ca).$$

(iii) (Dilation) *For* $c > 0$*, with* D_c *denoting the (scaled) dilation operator defined by* $(D_c f)(t) = \frac{1}{\sqrt{c}} f\Big(\frac{t}{c}\Big)$,

$$W_\psi[D_c f](a,b) = \frac{1}{\sqrt{c}} W_\psi[f]\Big(\frac{a}{c}, \frac{b}{c}\Big) \qquad \textit{and} \qquad W_{D_c\psi}[f](a,b) = \frac{1}{\sqrt{c}} W_\psi[f](ac,b).$$

(iv) (Symmetry) *For any* $a \neq 0$,

$$W_\psi[f](a,b) = W_f[\psi]\Big(1, -\frac{b}{a}\Big).$$

(v) (Parity) *With* P *denoting the parity operator defined by* $(Pf)(t) = f(-t)$ *(that is,* $P = D_{-1}$*),*

$$W_{P\psi}[Pf](a,b) = W_\psi[f](a,-b).$$

(vi) (Anti-linearity) *For any scalars* α, β,

$$W_{\alpha\psi+\beta\phi}[f] = \overline{\alpha} W_\psi[f] + \overline{\beta} W_\phi[g].$$

Proofs of the above properties are straightforward and are left as exercises.

Theorem 6.10 (Parseval's Formula for Wavelet Transforms)**.** *If* $\psi \in L^2(\mathbb{R})$ *is a wavelet and* $W_\psi[f]$ *is the wavelet transform of* f *(relative to* ψ*) defined by* (6.2.4)*, then, for any functions* f, $g \in L^2(\mathbb{R})$*, we obtain*

$$\int_{\mathbb{R}}\int_{\mathbb{R}} W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\,\frac{dbda}{a^2} = C_\psi \langle f, g\rangle_{L^2(\mathbb{R})}, \tag{6.3.1}$$

where C_ψ *is defined by* (6.2.1)*.*

Proof from the textbook. By Parseval's relation (3.4.37) for the Fourier transforms, we have

$$\begin{aligned} W_\psi[f](a,b) = \langle f, \psi_{a,b}\rangle_{L^2(\mathbb{R})} &= \frac{1}{2\pi}\big\langle \widehat{f}, \widehat{\psi_{a,b}}\big\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\big\langle \widehat{f}, \widehat{T_b D_a \psi}\big\rangle_{L^2(\mathbb{R})} \\ &= \frac{1}{2\pi}\langle \widehat{f}, M_{-b} D_{\frac{1}{a}} \widehat{\psi}\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\omega)|a|^{\frac{1}{2}} e^{ib\omega}\overline{\widehat{\psi}(a\omega)}\,d\omega, \end{aligned} \tag{6.3.2}$$

and substituting g for f in the identity above,

$$\overline{W_\psi[g](a,b)} = \frac{1}{2\pi}\int_{\mathbb{R}} \overline{\widehat{g}(\sigma)}|a|^{\frac{1}{2}} e^{-ib\sigma}\widehat{\psi}(a\sigma)\,d\sigma. \tag{6.3.3}$$

Substituting (6.3.2) and (6.3.3) in the left-hand side of (6.3.1) gives

$$\begin{aligned} &\int_{\mathbb{R}}\int_{\mathbb{R}} W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\,\frac{dbda}{a^2} \\ &= \frac{1}{(2\pi)^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{f}(\omega)|a|^{\frac{1}{2}}\overline{\widehat{\psi}(a\omega)}\,\overline{\widehat{g}(\sigma)}|a|^{\frac{1}{2}} e^{ib(\omega-\sigma)}\widehat{\psi}(a\sigma)\,d\sigma d\omega \frac{dbda}{a^2} \\ &= \frac{1}{(2\pi)^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{|a|}\widehat{f}(\omega)\overline{\widehat{\psi}(a\omega)}\,\overline{\widehat{g}(\sigma)}\widehat{\psi}(a\sigma)e^{ib(\omega-\sigma)}\,d\sigma d\omega dbda. \end{aligned} \tag{6.3.4}$$

Note that

$$f(t) = \mathscr{F}^{-1}[\widehat{f}](t) = \frac{1}{2\pi}\int_{\mathbb{R}}\Big(\int_{\mathbb{R}} f(s)e^{-is\omega}\Big)e^{i\omega t}\,dsd\omega = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} f(s)e^{i\omega(t-s)}\,dsd\omega;$$

thus interchanging the order of integration from $d\omega db$ to $dbd\omega$ in (6.3.4) we obtain that

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\,\frac{dbda}{a^2}\\
&\quad= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{|a|}\widehat{f}(\omega)\overline{\widehat{\psi}(a\omega)}\bigg(\frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\overline{\widehat{g}(\sigma)}\widehat{\psi}(a\sigma)e^{ib(\omega-\sigma)}\,d\omega db\bigg)d\sigma da\\
&\quad= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{|a|}\widehat{f}(\omega)\overline{\widehat{\psi}(a\omega)}\overline{\widehat{g}(\omega)}\widehat{\psi}(a\omega)d\sigma da = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\frac{|\widehat{\psi}(a\omega)|^2}{|a|}d\omega da\,.
\end{aligned}$$

A further interchanging the order of integration and putting $a\omega = x$ show that

$$\begin{aligned}
\int_{\mathbb{R}}\int_{\mathbb{R}} W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\,\frac{dbda}{a^2} &= \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\bigg(\int_{\mathbb{R}}\frac{|\widehat{\psi}(a\omega)|^2}{|a|}da\bigg)d\omega\\
= \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\bigg(\int_{\mathbb{R}}\frac{|\widehat{\psi}(x)|^2}{|x|}dx\bigg)d\omega &= \frac{C_\psi}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\,d\omega = \frac{C_\psi}{2\pi}\big\langle\widehat{f},\widehat{g}\big\rangle_{L^2(\mathbb{R})}\,,
\end{aligned}$$

and the Plancherel identity is used to conclude the theorem. □

Remark 6.11. The proof above is not rigorous since the first interchange of the order of integration (from $d\omega db$ to $dbd\omega$) cannot be valid since the integrand is indeed not integrable. The second interchange of the order of integration (from $d\omega da$ to $dad\omega$) is true due to the fact that $\widehat{f}$, $\widehat{g}$ and the admissibility condition.

Proof from another book. Define

$$F(\omega) = |a|^{-\frac{1}{2}}\mathscr{F}[W_\psi[f](a,\cdot)](\omega) \qquad \text{and} \qquad G(\omega) = |a|^{-\frac{1}{2}}\mathscr{F}[W_\psi[g](a,\cdot)](\omega)\,.$$

Using (6.2.5),

$$F(\omega) = \widehat{f}(\omega)\overline{\widehat{\psi}(a\omega)}\,, \qquad G(\omega) = \widehat{g}(\omega)\overline{\widehat{\psi}(a\omega)}\,. \tag{6.3.5}$$

Applying the Plancherel identity and using (6.2.2), we find that

$$\begin{aligned}
\int_{\mathbb{R}}\big[W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\big]db &= \big\langle W_\psi[f](a,\cdot), W_\psi[g](a,\cdot)\big\rangle_{L^2(\mathbb{R})}\\
&= \frac{1}{2\pi}\big\langle\mathscr{F}[W_\psi[f](a,\cdot)],\mathscr{F}[W_\psi[g](a,\cdot)\big\rangle = \frac{|a|}{2\pi}\langle F,G\rangle.
\end{aligned}$$

Hence, by substituting (6.3.5) into the above expression, integrating with respect to da/a^2 on $\mathbb{R}$, and recalling the definition of C_ψ from (6.2.1), we obtain

$$\begin{aligned}
&\int_{\mathbb{R}}\Big[\int_{\mathbb{R}}\big[W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\big]db\Big]\frac{da}{a^2} = \int_{\mathbb{R}}\Big[\frac{1}{2\pi|a|}\langle F,G\rangle\Big]da\\
&\quad= \frac{1}{2\pi}\int_{\mathbb{R}}\Big[\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\frac{|\widehat{\psi}(a\omega)|^2}{|a|}dx\Big]da = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\frac{|\widehat{\psi}(a\omega)|^2}{|a|}dadx\\
&\quad= \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\omega)\overline{\widehat{g}(\omega)}\Big(\int_{\mathbb{R}}\frac{|\widehat{\psi}(y)|^2}{|y|}dy\Big)dx = \frac{C_\psi}{2\pi}\langle\widehat{f},\widehat{g}\rangle_{L^2(\mathbb{R})} = C_\psi\langle f,g\rangle_{L^2(\mathbb{R})}\,.
\end{aligned}$$

Note that the switch of the order of the integration in the second line is due to the Fubini Theorem which requires that the integrand is integrable; nevertheless, the integrability of the integrand can be shown using the Tonelli Theorem as long as the admissibility condition (6.2.1) is satisfied. This completes the proof of the theorem. □

Remark 6.12. There is still one particular problem in the proof above: In order to apply the Plancherel identity it is required that $W_\psi[f](a,\cdot)$ and $W_\psi[g](a,\cdot) \in L^2(\mathbb{R})$ for $a \neq 0$. However, if $W_\psi[f](a,\cdot) \in L^2(\mathbb{R})$, then $\mathscr{F}[W_\psi[f](a,\cdot)] \in L^2(\mathbb{R})$ but if this might not be true since $\widehat{f}$, $\widehat{g}$, $\widehat{\psi} \in L^2(\mathbb{R})$ only guarantees that $\mathscr{F}[W_\psi[f](a,\cdot)]$, $\mathscr{F}[W_\psi[g](a,\cdot)] \in L^1(\mathbb{R})$.

Proof. We first prove that (6.3.1) holds for all $f,g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Let $f,g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and as in the previous proof we define

$$F(\omega) = |a|^{-\frac{1}{2}} \mathscr{F}[W_\psi[f](a,\cdot)](\omega) \qquad \text{and} \qquad G(\omega) = |a|^{-\frac{1}{2}} \mathscr{F}[W_\psi[g](a,\cdot)](\omega).$$

Then for $a \neq 0$, $F, G \in L^2(\mathbb{R})$; thus $W_\psi[f](a,\cdot)$, $W_\psi[g](a,\cdot) \in L^2(\mathbb{R})$. Therefore, the previous proof goes through and we obtain that (6.3.1) holds for $f,g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. In particular,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{1}{|a|} W_\psi[f](a,b) \right|^2 dbda = C_\psi \|f\|^2_{L^2(\mathbb{R})} \tag{6.3.6}$$

Now let $f \in L^2(\mathbb{R})$. Choose $\{f_n\}_{n=1}^\infty \subseteq L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\{f_n\}_{n=1}^\infty$ converges to f in $L^2(\mathbb{R})$. Then $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R})$ and the identity above, together with the linearity of the wavelet transform, shows that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{1}{|a|} W_\psi[f_n](a,b) - \frac{1}{|a|} W_\psi[f_m](a,b) \right|^2 dbda = C_\psi \|f_n - f_m\|^2_{L^2(\mathbb{R})} \to 0 \quad \text{as} \quad n,m \to \infty.$$

This shows that the sequence $\{h_n\}_{n=1}^\infty$ defined by

$$h_n(a,b) = \frac{1}{|a|} W_\psi[f_n](a,b) = \frac{1}{|a|} \langle f_n, \psi_{a,b} \rangle_{L^2(\mathbb{R})}$$

is a Cauchy sequence in $L^2(\mathbb{R}^2)$; thus $\{h_n\}_{n=1}^\infty$ converges to some function in $L^2(\mathbb{R}^2)$. On the other hand, the fact that $f_n \to f$ in $L^2(\mathbb{R})$ shows that $\{h_n\}_{n=1}^\infty$ converges a.e. to the function $h(a,b) = \frac{1}{|a|} W_\psi[f](a,b)$. Therefore, $\{h_n\}_{n=1}^\infty$ converges to g in $L^2(\mathbb{R}^2)$ and this shows that

$$\begin{aligned}\int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{1}{|a|} W_\psi[f](a,b) \right|^2 dbda = \|h\|^2_{L^2(\mathbb{R})} = \lim_{n\to\infty} \|h_n\|^2_{L^2(\mathbb{R})} = \lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{1}{|a|} W_\psi[f_n](a,b) \right|^2 dbda \\ = C_\psi \lim_{n\to\infty} \|f_n\|^2_{L^2(\mathbb{R})} = C_\psi \|f\|^2_{L^2(\mathbb{R})}.\end{aligned}$$

This establishes that (6.3.6) holds for all $f \in L^2(\mathbb{R})$, and we completes the proof of the theorem using the polarization identity. □

Theorem 6.13 (Inversion Formula). *If $f \in L^2(\mathbb{R})$, then f can be reconstructed by the formula*

$$f(t) = \frac{1}{C_\psi} \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f](a,b) \psi_{a,b}(t) \, \frac{dbda}{a^2}, \tag{6.3.7}$$

where the equality holds almost everywhere.

Proof from the textbook. For any $g \in L^2(\mathbb{R})$, we have, from Theorem 6.10, that

$$\begin{aligned} C_\psi \langle f, g\rangle_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f](a,b) \overline{W_\psi[g](a,b)} \, \frac{dbda}{a^2} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f](a,b) \overline{\Big(\int_{\mathbb{R}} g(t) \overline{\psi_{a,b}(t)} \, dt \Big)} \, \frac{dbda}{a^2} \\ &= \int_{\mathbb{R}} \bigg(\int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f](a,b) \psi_{a,b}(t) \, \frac{dbda}{a^2} \bigg) \overline{g(t)} \, dt . \end{aligned}$$

Since g is an arbitrary element of $L^2(\mathbb{R})$, the inversion formula (6.3.7) follows. □

Remark 6.14. The problem in this proof again is that the interchange of the order of integration cannot be guaranteed. Nevertheless, it is possible to show the validity of the inversion formula for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with $\widehat{f} \in L^1(\mathbb{R})$. However, even if this case is proved, we cannot prove the general result by the density argument since the convergence

$$\lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f_n](a,b) \psi_{a,b}(t) \frac{dbda}{a^2} = \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi[f](a,b) \psi_{a,b}(t) \frac{dbda}{a^2} .$$

cannot be guaranteed.

Proof. For fixed positive constants ε, A, B satisfying $\varepsilon < A$, we define an operator $S(\varepsilon, A, B)$ on $L^2(\mathbb{R})$ by

$$[S(\varepsilon, A, B) f](t) = \frac{1}{C_\psi} \int_{\varepsilon < |a| < A} \int_{|b| < B} W_\psi[f](a,b) \psi_{a,b}(t) \frac{dbda}{a^2} \qquad f \in L^2(\mathbb{R}) .$$

Our goal is to show that $S(\varepsilon, A, B) f \to f$ in $L^2(\mathbb{R})$ as $\varepsilon \to 0^+$ and $A, B \to \infty$.

To see the L^2 convergence, we note that

$$\big\| S(\varepsilon, A, B) f - f \big\|_{L^2(\mathbb{R})} = \sup_{\|g\|_{L^2(\mathbb{R})} = 1} \Big| \big\langle S(\varepsilon, A, B) f - f, g \big\rangle_{L^2(\mathbb{R})} \Big|$$

For $g \in L^2(\mathbb{R})$,

$$\big\langle S(\varepsilon, A, B) f, g \big\rangle_{L^2(\mathbb{R})} = \frac{1}{C_\psi} \int_{\mathbb{R}} \int_{\varepsilon < |a| < A} \int_{|b| < B} W_\psi[f](a,b) \psi_{a,b}(t) \overline{g(t)} \frac{dbda}{a^2} dt . \tag{6.3.8}$$

We first show that the integral above is absolutely convergent. By the Tonelli Theorem,

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\varepsilon < |a| < A} \int_{|b| < B} \big| W_\psi[f](a,b) \big| \big| \psi_{a,b}(t) \big| \big| g(t) \big| \, \frac{dbda}{a^2} dt \\ &= \int_{\varepsilon < |a| < A} \int_{|b| < B} \big| W_\psi[f](a,b) \big| \bigg(\int_{\mathbb{R}} \big| \psi_{a,b}(t) \big| \big| g(t) \big| \, dt \bigg) \frac{dbda}{a^2} \\ &= \bigg[\int_{\varepsilon < |a| < A} \int_{|b| < B} \big| W_\psi[f](a,b) \big| \bigg(\int_{\mathbb{R}} \big| \psi_{a,b}(t) \big|^2 \, dt \bigg)^{\frac{1}{2}} \frac{dbda}{a^2} \bigg] \|g\|_{L^2(\mathbb{R})} \\ &= \bigg(\int_{\varepsilon < |a| < A} \int_{|b| < B} \Big| \frac{1}{|a|} W_\psi[f](a,b) \Big|^2 dbda \bigg)^{\frac{1}{2}} \bigg(\int_{\varepsilon < |a| < A} \int_{|b| < B} \int_{\mathbb{R}} \big| \psi_{a,b}(t) \big|^2 dt \frac{dbda}{a^2} \bigg)^{\frac{1}{2}} \|g\|_{L^2(\mathbb{R})} \\ &= \sqrt{C_\psi} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \bigg(\int_{\varepsilon < |a| < A} \int_{|b| < B} \int_{\mathbb{R}} \big| \psi_{a,b}(t) \big|^2 dt \frac{dbda}{a^2} \bigg)^{\frac{1}{2}} \end{aligned}$$

and further computation shows that

$$\int_{\varepsilon<|a|<A}\int_{|b|<B}\int_{\mathbb{R}}\left|\psi_{a,b}(t)\right|^2 dt\frac{dbda}{a^2} = \|\psi\|^2_{L^2(\mathbb{R})}\int_{\varepsilon<|a|<A}\int_{|b|<B}\frac{1}{|a|^2}\,dbda < \infty.$$

Therefore, the integral on the RHS of (6.3.8) is absolutely convergent, so the Fubini Theorem implies that

$$\begin{aligned}
\big\langle S(\varepsilon,A,B)f,g\big\rangle_{L^2(\mathbb{R})} &= \frac{1}{C_\psi}\int_{\mathbb{R}}\int_{\varepsilon<|a|<A}\int_{|b|<B} W_\psi[f](a,b)\psi_{a,b}(t)\overline{g(t)}\frac{dbda}{a^2}dt\\
&= \frac{1}{C_\psi}\int_{\varepsilon<|a|<A}\int_{|b|<B} W_\psi[f](a,b)\int_{\mathbb{R}}\psi_{a,b}(t)\overline{g(t)}\,dt\frac{dbda}{a^2}\\
&= \frac{1}{C_\psi}\int_{\varepsilon<|a|<A}\int_{|b|<B} W_\psi[f](a,b)\overline{W_\psi[g](a,b)}\,\frac{dbda}{a^2};
\end{aligned}$$

thus (6.3.1) implies that

$$\begin{aligned}
\big\|S(\varepsilon,A,B)f-f\big\|_{L^2(\mathbb{R})} &= \sup_{\|g\|_{L^2(\mathbb{R})}=1}\left|\big\langle S(\varepsilon,A,B)f-f,g\big\rangle_{L^2(\mathbb{R})}\right|\\
&= \sup_{\|g\|_{L^2(\mathbb{R})}=1}\left|\frac{1}{C_\psi}\int_{\{(a,b)|\varepsilon<|a|<A,|b|<B\}^\complement} W_\psi[f](a,b)W_\psi[g](a,b)\,\frac{d(a,b)}{a^2}\right|\\
&\leqslant \frac{1}{\sqrt{C_\psi}}\int_{\{(a,b)|\varepsilon<|a|<A,|b|<B\}^\complement}\left|\frac{1}{|a|}W_\psi[f](a,b)\right|^2 d(a,b)\,.
\end{aligned}$$

The fact that the function $(a,b)\mapsto \dfrac{1}{|a|}W_\psi[f](a,b)$ belongs to $L^2(\mathbb{R}^2)$ shows that

$$\lim_{\varepsilon\to0^+,A,B\to\infty}\int_{\{(a,b)|\varepsilon<|a|<A,|b|<B\}^\complement}\left|\frac{1}{|a|}W_\psi[f](a,b)\right|^2 d(a,b) = 0\,.$$

This shows that $\lim\limits_{\varepsilon\to0^+,A,B\to\infty}\big\|S(\varepsilon,A,B)f-f\big\|_{L^2(\mathbb{R})} = 0$ and the proof is complete. □

Remark 6.15. The proof above indeed shows that the RHS integral of (6.3.7) is obtained by

$$\lim_{\varepsilon\to0^+,B\to\infty}\int_{\varepsilon<|a|}\int_{|b|<B} W_\psi[f](a,b)\psi_{a,b}(t)\frac{dbda}{a^2}\,.$$

On the other hand, for $f\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$ with $\widehat{f}\in L^1(\mathbb{R})$, the RHS integral of (6.3.7) means the usual Lebesgue integral.

6.4 The Discrete Wavelet Transforms

It has been stated in the last section that the continuous wavelet transform (6.2.4) is a two-parameter representation of a function. In many applications, especially in signal processing, data are represented by a finite number of values, so it is important and often useful to consider discrete versions of the continuous wavelet transform (6.2.4). Our goal in this section is to answer the fundamental question whether we can reconstruct f from discrete values of its wavelet transform $W_\psi[f]$. In particular, we would like to reconstruct f using the discrete values of $W_\psi[f]$ at $a = a_0^m$ and $b = nb_0a_0^m$; that is,

$$(W_\psi[f])(a_0^m, nb_0a_0^m) = a_0^{-\frac{m}{2}} \int_{\mathbb{R}} f(t)\overline{\psi}(a_0^{-m}t - nb_0)dt,$$

where $a_0 \neq 0$, b_0 are some given and fixed constants, and m, n are integers. Define

$$\psi_{m,n}(x) = a_0^{-m/2}\psi(a_0^{-m}x - nb_0) = (D_{a_0^m}T_{nb_0}\psi)(x), \tag{6.4.1}$$

where we abuse the use of notation here and do not confuse with (6.2.3). Using (6.4.1), we have

$$(W_\psi[f])(a_0^m, nb_0a_0^m) = \langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}.$$

The discrete wavelet transform represents a function by a countable set of wavelet coefficients, which correspond to points on a two dimensional grid or lattice of discrete points in the scale-time domain indexed by m and n.

The answer to the fundamental question is positive if the wavelets form a complete system in $L^2(\mathbb{R})$. The problem is whether there exists another function $g \in L^2(\mathbb{R})$ such that

$$\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})} = \langle g, \psi_{m,n}\rangle_{L^2(\mathbb{R})}$$

for all m, $n \in \mathbb{Z}$ implies $f = g$. In practice, the evaluation/measurement of $\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}$ might not be very accurate, so the best we can hope is that f and g are "close" if the two sequences $\{\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\}_{m,n\in\mathbb{Z}}$ and $\{\langle g, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\}_{m,n\in\mathbb{Z}}$ are "close". This property can be guaranteed if there exists an $A > 0$ independent of f, such that

$$A\|f\|^2_{L^2(\mathbb{R})} \leqslant \sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2$$

since the inequality above implies that

$$A\|f - g\|^2_{L^2(\mathbb{R})} \leqslant \sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})} - \langle g, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2.$$

On the other hand, we also want two sequences $\{\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\}_{m,n\in\mathbb{Z}}$ to be "close" if f and g are "close". This will be guaranteed if there exists a $B > 0$ independent of f such that

$$\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 \leqslant B\|f\|^2_{L^2(\mathbb{R})}$$

since the inequality above implies that

$$\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})} - \langle g, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 \leqslant B\|f - g\|^2_{L^2(\mathbb{R})}.$$

These two requirements are best studied in terms of the so-called frames.

6.4.1 Frames and frame operators

In the following, when the inner product of a Hilbert space is specified, $\|\cdot\|$ is used to denote the induced norm of the inner product.

Definition 6.16 (Frames)**.** A collection of countably many vectors $\{x_n\}$ in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is called a frame if these exist constants A, $B > 0$ such that

$$A\|x\|^2 \leqslant \sum_n \left|\langle x, x_n\rangle\right|^2 \leqslant B\|x\|^2 \qquad \forall\, x \in H. \tag{6.4.2}$$

The constants A and B are called frame bounds, and a frame satisfying (6.4.2) is called a frame with frame bounds A and B. If $A = B$, then the frame is called tight. The frame is called exact if no proper subset of $\{x_n\}$ is also a frame.

The following example shows that tightness and exactness are not related.

Example 6.17. If $\{e_n\}$ is an orthonormal basis of H, then

(i) $\{e_1, e_1, e_2, e_2, e_3, e_3, \cdots\}$ is a tight frame with frame bounds $A = B = 2$, but it is not exact.

(ii) $\{\sqrt{2}e_1, e_2, e_3, \cdots\}$ is an exact frame but not tight since the frame bounds are easily seen as $A = 1$ and $B = 2$.

(iii) $\left\{e_1, \frac{e_2}{\sqrt{2}}, \frac{e_2}{\sqrt{2}}, \frac{e_3}{\sqrt{3}}, \frac{e_3}{\sqrt{3}}, \frac{e_3}{\sqrt{3}}, \cdots\right\}$ is a tight frame with the frame bound $A = B = 1$ but not an orthonormal basis.

(iv) $\left\{e_1, \frac{e_2}{2}, \frac{e_3}{3}, \cdots\right\}$ is a complete orthogonal sequence but is not a frame.

Theorem 6.18. *Let $\{x_n\}$ be a collection of countably many vectors in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$. Then the following two statements are equivalent.*

(a) *The operator $Tx = \sum\limits_n \langle x, x_n\rangle x_n$ is a bounded linear operator on H with $A\mathrm{I} \leqslant T \leqslant B\mathrm{I}$, where I is the identity operator on H.*

(b) *$\{x_n\}$ is a frame with frame bounds A and B.*

Proof. Before proceeding, we recall that the relation $A\mathrm{I} \leqslant T \leqslant B\mathrm{I}$ means

$$\langle A\mathrm{I}x, x\rangle \leqslant \langle Tx, x\rangle \leqslant \langle B\mathrm{I}x, x\rangle \qquad \forall\, x \in H, \tag{6.4.3}$$

and note that if the series $\sum\limits_n \langle x, x_n\rangle x_n$ converges (if it is an infinite sum) for some particular $x \in H$, then

$$\Big\langle \sum_n \langle x, x_n\rangle x_n, y\Big\rangle = \sum_n \Big\langle \langle x, x_n\rangle x_n, y\Big\rangle = \sum_n \langle x, x_n\rangle\langle x_n, y\rangle. \tag{6.4.4}$$

"(a) ⇒ (b)" Suppose that (a) holds. Since T is defined on H, the series $\sum_n \langle x, x_n\rangle x_n$ converges to Tx for all $x \in H$; thus (6.4.4) implies that

$$\langle Tx, x\rangle = \sum_n \langle x, x_n\rangle\langle x_n, x\rangle = \sum_n \left|\langle x, x_n\rangle\right|^2.$$

Using (6.4.3), we conclude that $\{x_n\}_n$ is a frame with frame bounds A and B. This shows that (a) implies (b).

"(b) ⇒ (a)" We next prove that (b) implies (a). Suppose (b) holds. First we claim that $Tx = \sum_n \langle x, x_n\rangle x_n$ converges for all $x \in H$. To see this, it suffices to show the case that $\{x_n\} = \{x_n\}_{n=1}^{\infty}$ is an infinite sequence. Recall that in any Hilbert space H the norm of any element $x \in H$ is given by

$$\|x\| = \sup_{\|y\|=1} \left|\langle x, y\rangle\right|.$$

For a fixed $x \in H$, we consider $T_N x = \sum_{n=1}^{N} \langle x, x_n\rangle x_n$. For $0 \leqslant M < N$, we have, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left\|T_N x - T_M x\right\|^2 &= \sup_{\|y\|=1} \left|\langle T_N x - T_M x, y\rangle\right|^2 = \sup_{\|y\|=1} \left|\sum_{M+1\leqslant n\leqslant N} \langle x, x_n\rangle\langle x_n, y\rangle\right|^2 \\
&\leqslant \sup_{\|y\|=1} \Big(\sum_{M+1\leqslant n\leqslant N} \left|\langle x, x_n\rangle\right|^2\Big)\Big(\sum_{M+1\leqslant n\leqslant N} \left|\langle x_n, y\rangle\right|^2\Big) \\
&\leqslant \sup_{\|y\|=1} \Big(\sum_{M+1\leqslant n\leqslant N} \left|\langle x, x_n\rangle\right|^2\Big) B\|y\|^2 \\
&= B\Big(\sum_{M+1\leqslant n\leqslant N} \left|\langle x, x_n\rangle\right|^2\Big) \to 0 \qquad \text{as } M, N \to \infty.
\end{aligned}$$

Thus, $\{T_N x\}_{N=1}^{\infty}$ is a Cauchy sequence in H and hence it is convergent as $N \to \infty$. Therefore, $Tx \equiv \sum_n \langle x, x_n\rangle x_n$ converges for all $x \in H$, and (6.4.4) implies that

$$\langle Tx, y\rangle = \sum_n \langle x, x_n\rangle\langle x_n, y\rangle.$$

Following the preceding argument we obtain that

$$\|Tx\|^2 = \sup_{\|y\|=1} \left|\langle Tx, y\rangle\right|^2 = \sup_{\|y\|=1} \left|\sum_n \langle x, x_n\rangle\langle x_n, y\rangle\right|^2 \leqslant B\Big(\sum_n \left|\langle x, x_n\rangle\right|^2\Big) \leqslant B^2\|x\|^2$$

which implies that $\|T\| \leqslant B$. This shows the boundedness of T. The relation $A\mathrm{I} \leqslant T \leqslant B\mathrm{I}$ follows from that $\langle Tx, x\rangle = \sum_n \left|\langle x, x_n\rangle\right|^2$ and the frame $\{x_n\}$ has frame bounds A and B. □

Definition 6.19 (Frame Operator)**.** To each frame $\{x_n\}$ in a Hilbert space there corresponds an operator T, called the ***frame operator***, from H into itself defined by

$$Tx = \sum_n \langle x, x_n\rangle x_n \qquad \forall\, x \in H. \tag{6.4.5}$$

We remark that the frame operator associated with a frame is self-adjoint because of (6.4.4):

$$\langle Tx, y\rangle = \Big\langle \sum_n \langle x, x_n\rangle x_n, y \Big\rangle = \sum_n \Big\langle \langle x, x_n\rangle x_n, y \Big\rangle = \sum_n \langle x, x_n\rangle\langle x_n, y\rangle$$
$$= \overline{\sum_n \overline{\langle x, x_n\rangle}\,\overline{\langle x_n, y\rangle}} = \overline{\sum_n \langle y, x_n\rangle\langle x_n, x\rangle} = \overline{\langle Ty, x\rangle} = \langle x, Ty\rangle.$$

Theorem 6.20. *Suppose $\{x_n\}$ is a frame on a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ with frame bounds A and B, and T is the corresponding frame operator. Then,*

(a) *T is invertible and $B^{-1}\mathrm{I} \leqslant T^{-1} \leqslant A^{-1}\mathrm{I}$. Furthermore, T^{-1} is a positive operator and hence it is self-adjoint.*

(b) *$\{T^{-1}x_n\}$ is a frame with frame bounds B^{-1} and A^{-1}.*

(c) *Every $x \in H$ can be expressed in the form*

$$x = \sum_n \langle x, T^{-1}x_n\rangle x_n = \sum_n \langle x, x_n\rangle T^{-1}x_n. \tag{6.4.6}$$

Proof. (a) Since the frame operator T satisfies the relation

$$A\mathrm{I} \leqslant T \leqslant B\mathrm{I};$$

it follows that

$$\mathrm{I} - B^{-1}T \leqslant \mathrm{I} - B^{-1}A\mathrm{I} = \Big(1 - \frac{A}{B}\Big)\mathrm{I}$$

and hence

$$\big\|\mathrm{I} - B^{-1}T\big\| \leqslant \Big(1 - \frac{A}{B}\Big)\|\mathrm{I}\| < 1.$$

Thus, $B^{-1}T$ is invertible and consequently so is T. In view of the fact that

$$\langle T^{-1}x, x\rangle = \langle T^{-1}x, TT^{-1}x\rangle \geqslant A\langle T^{-1}x, T^{-1}x\rangle = A\|T^{-1}x\|^2 \neq 0 \quad \text{whenever } x \neq 0,$$

we conclude that T^{-1} is a positive operator and hence it is self-adjoint. Finally, since T is a bounded positive operator, it is possible to define $T^{\frac{1}{2}}$ (using the spectral decomposition in functional analysis) and $T^{\frac{1}{2}}$ is also positive definite (hence self-adjoint) and invertible. Therefore, by $A\mathrm{I} \leqslant T \leqslant B\mathrm{I}$, the fact that

$$A^{-1}\mathrm{I} - T^{-1} = A^{-1}T^{-\frac{1}{2}}(T - A\mathrm{I})T^{-\frac{1}{2}} \quad \text{and} \quad T^{-1} - B^{-1}\mathrm{I} = B^{-1}T^{-\frac{1}{2}}(B\mathrm{I} - T)T^{-\frac{1}{2}}$$

shows that

$$B^{-1}\mathrm{I} \leqslant T^{-1} \leqslant A^{-1}\mathrm{I}.$$

(b) Since T^{-1} is self-adjoint, we have

$$\sum_n \langle x, T^{-1}x_n\rangle T^{-1}x_n = T^{-1}\Big(\sum_n \langle T^{-1}x, x_n\rangle x_n\Big) = T^{-1}\big(T(T^{-1}x)\big) = T^{-1}x. \tag{6.4.7}$$

This gives

$$\langle T^{-1}x, x\rangle = \left\langle \sum_n \langle x, T^{-1}x_n\rangle T^{-1}x_n, x \right\rangle = \sum_n \langle x, T^{-1}x_n\rangle\langle T^{-1}x_n, x\rangle.$$

Hence,

$$\langle T^{-1}x, x\rangle = \sum_n \langle x, T^{-1}x_n\rangle\overline{\langle x, T^{-1}x_n\rangle} = \sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2.$$

Using the result from (a); that is, $B^{-1}\mathrm{I} \leqslant T \leqslant A^{-1}\mathrm{I}$; it turns out that

$$B^{-1}\langle \mathrm{I}x, x\rangle \leqslant \langle T^{-1}x, x\rangle \leqslant A^{-1}\langle \mathrm{I}x, x\rangle$$

and hence

$$B^{-1}\|x\|^2 \leqslant \langle T^{-1}x, x\rangle \leqslant A^{-1}\|x\|^2.$$

By Theorem 6.18 this shows that $\{T^{-1}x_n\}$ is a frame with frame bounds B^{-1} and A^{-1}.

(c) We replace x by $T^{-1}x$ in (6.4.5) to derive

$$x = \sum_n \langle T^{-1}x, x_n\rangle x_n = \sum_n \langle x, T^{-1}x_n\rangle x_n.$$

Similarly, replacing x by Tx in (6.4.7) gives

$$x = \sum_n \langle Tx, T^{-1}x_n\rangle T^{-1}x_n = \sum_n \langle x, x_n\rangle T^{-1}x_n.$$

This completes the proof. □

Definition 6.21. Let H be a separable Hilbert space, $\{x_n\}$ be a frame in H, and T be the corresponding frame operator of frame $\{x_n\}$. The frame $\{T^{-1}x_n\}$ is called the ***dual frame*** of $\{x_n\}$.

By writing $T^{-1}x_n$ as $\widetilde{x}_n$, according to formula (6.4.6), the reconstruction formula for x has the form

$$x = \sum_n \langle x, \widetilde{x}_n\rangle x_n = \sum_n \langle x, x_n\rangle \widetilde{x}_n.$$

It is easy to verify that the dual frame of $\{\widetilde{x}_n\}$ is the original frame $\{x_n\}$.

Theorem 6.22. *Suppose $\{x_n\}$ is a frame in a separable Hilbert space H with frame bounds A and B. If there exists a sequence of scalars $\{c_n\}$ such that $x = \sum_n c_n x_n$, then*

$$\sum_n |c_n|^2 = \sum_n |a_n|^2 + \sum_n |a_n - c_n|^2,$$

where $a_n = \langle x, T^{-1}x_n\rangle$ so that $x = \sum_n a_n x_n$.

Proof. Note that $\langle x_n, T^{-1}x\rangle = \langle T^{-1}x_n, x\rangle = \overline{a_n}$. Substituting $x = \sum\limits_n a_n x_n$ in the first term in the inner product $\langle x, T^{-1}x\rangle$ gives

$$\langle x, T^{-1}x\rangle = \Big\langle \sum_n a_n x_n, T^{-1}x \Big\rangle = \sum_n a_n \langle x_n, T^{-1}x\rangle = \sum_n |a_n|^2 .$$

Similarly, substituting $x = \sum\limits_n c_n x_n$ into the first term in the inner product $\langle x, T^{-1}x\rangle$ yields

$$\langle x, T^{-1}x\rangle = \Big\langle \sum_n c_n x_n, T^{-1}x \Big\rangle = \sum_n c_n \langle x_n, T^{-1}x\rangle = \sum_n c_n \overline{a_n} .$$

Consequently,

$$\sum_n |a_n|^2 = \sum_n c_n \overline{a_n} = \sum_n \overline{c_n} a_n . \tag{6.4.8}$$

Finally, we obtain, by using (6.4.8),

$$\sum_n |a_n|^2 + \sum_n |a_n - c_n|^2 = \sum_n |a_n|^2 + \sum_n \big(|a_n|^2 - a_n\overline{c_n} - c_n\overline{a_n} + |c_n|^2\big) = \sum_n |c_n|^2 .$$

This completes the proof. □

Theorem 6.23. *Let $\{x_n\}$ be a frame in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$. If $\{x_n\}$ is exact, then $\{x_n\}$ and its dual frame $\{T^{-1}x_n\}$ is biorthonormal; that is,*

$$\langle x_m, T^{-1}x_n\rangle = \delta_{mn} \qquad \forall\, m, n \in \mathbb{Z} . \tag{6.4.9}$$

Proof. Suppose that $\{x_n\}$ is a frame with frame bounds A and B. Let $\mathcal{I}$ denote the index set, and $m \in \mathcal{I}$ be a fixed index. Since $\{x_n\}$ is an exact frame,

$$A \leqslant \sum_n \big|\langle x, x_n\rangle\big|^2 \leqslant B \qquad \forall\, x \in H \text{ with } \|x\| = 1$$

but the fact that $\{x_n\}_{n\in\mathcal{I}\setminus\{m\}}$ is not a frame implies that for each $k \in \mathbb{N}$, there exists $y_k \in H$ with $\|y_k\| = 1$ such that

$$1 > k \sum_{n\neq m} \big|\langle y_k, x_n\rangle\big|^2 .$$

Adding $\varepsilon\big|\langle y_k, x_m\rangle\big|^2$ on both sides of the inequality above, we find that

$$\begin{aligned} 1 + \varepsilon\big|\langle y_k, x_m\rangle\big|^2 &> (k-\varepsilon) \sum_{n\neq m} \big|\langle y_k, x_n\rangle\big|^2 + \varepsilon \sum_n \big|\langle y_k, x_n\rangle\big|^2 \\ &\geqslant (k-\varepsilon) \sum_{n\neq m} \big|\langle y_k, x_n\rangle\big|^2 + A\varepsilon . \end{aligned} \tag{6.4.10}$$

Letting $\varepsilon = 1/A$ in (6.4.10) and applying the Cauchy-Schwartz inequality we obtain that

$$(Ak-1) \sum_{n\neq m} \big|\langle y_k, x_n\rangle\big|^2 < \big|\langle y_k, x_m\rangle\big|^2 \leqslant \|x_m\|^2 ;$$

thus $\lim\limits_{k\to\infty} \sum\limits_{n\neq m} \big|\langle y_k, x_n\rangle\big|^2 = 0$. On the other hand, (6.4.6) implies that

$$\langle y_k, x_m\rangle = \sum_n \langle y_k, x_n\rangle \langle T^{-1}x_n, x_m\rangle$$

so that

$$\big(1 - \langle T^{-1}x_m, x_m\rangle\big)\langle y_k, x_m\rangle = \sum_{n\neq m}\langle y_k, x_n\rangle\langle T^{-1}x_n, x_m\rangle.$$

By Theorem 6.20 and the Cauchy-Schwartz inequality we have

$$\begin{aligned}\big|1 - \langle T^{-1}x_m, x_m\rangle\big|\big|\langle y_k, x_m\rangle\big| &\leqslant \Big(\sum_{n\neq m}\big|\langle y_k, x_n\rangle\big|^2\Big)^{\frac{1}{2}}\Big(\sum_{n\neq m}\big|\langle T^{-1}x_n, x_m\rangle\big|^2\Big)^{\frac{1}{2}}\\ &\leqslant \frac{1}{\sqrt{A}}\Big(\sum_{n\neq m}\big|\langle y_k, x_n\rangle\big|^2\Big)^{\frac{1}{2}}\|x_m\| \to 0 \quad\text{as}\quad k\to\infty.\end{aligned}$$

Letting $\varepsilon = k$ in (6.4.10) shows that $\big|\langle y_k, x_m\rangle\big|^2 \geqslant A - \frac{1}{k}$, so the inequality above shows that

$$\langle T^{-1}x_m, x_m\rangle = 1. \tag{6.4.11}$$

The rest of (6.4.9); that is, $\langle x_m, T^{-1}x_n\rangle = 0$ for $n \neq m$, follows from the identity

$$\langle x_m, T^{-1}x_m\rangle = \sum_n\langle x_m, T^{-1}x_n\rangle\langle x_n, T^{-1}x_m\rangle = \big|\langle x_m, T^{-1}x_m\rangle\big|^2 + \sum_{n\neq m}\big|\langle x_m, T^{-1}x_n\rangle\big|^2$$

and the identity (6.4.11). □

6.4.2 A sufficient condition for a function generating a frame

As pointed out above, we want the family of functions $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ to form a frame in $L^2(\mathbb{R})$. Obviously, the double indexing of the functions is irrelevant. The following theorem gives fairly general sufficient conditions for a sequence $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ to constitute a frame in $L^2(\mathbb{R})$.

Theorem 6.24. *Let $\psi \in L^2(\mathbb{R})$, and $a_0 > 1$. If*

(i) *there exist A, $B > 0$ such that $A \leqslant \sum_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m\omega)\big|^2 \leqslant B$ for all $1 \leqslant \omega \leqslant a_0$, and*

(ii) $\sup_{\omega\in\mathbb{R}}\sum_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m)\big|\big|\widehat{\psi}(a_0^m\omega + x)\big| \leqslant C(1+|x|)^{-(1+\delta)}$ *for some constants C and $\delta > 0$,*

then there exists $\widetilde{b} > 0$ such that the family $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ given by

$$\psi_{m,n} = a_0^{-m/2}\psi(a_0^{-m}x - nb_0) = (D_{a_0^m}T_{nb_0}\psi)(x)$$

forms a frame in $L^2(\mathbb{R})$ for any $b_0 \in (0, \widetilde{b})$.

Proof. Before proceeding, we remark that (i) is equivalent to that there exist A, $B > 0$ such that

$$A \leqslant \sum_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m\omega)\big|^2 \leqslant B \qquad \forall\,\omega \in \mathbb{R}\backslash\{0\} \tag{6.4.12}$$

due to the fact that $a_0 > 1$. Suppose $f \in L^2(\mathbb{R})$. By the Plancherel identity (1.1.15),

$$\begin{aligned}\sum_{m,n=-\infty}^{\infty}\big|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\big|^2 &= \sum_{m,n=-\infty}^{\infty}\big|\langle f, D_{a_0^m}T_{nb_0}\psi\rangle_{L^2(\mathbb{R})}\big|^2\\ &= \sum_{m,n=-\infty}^{\infty}\frac{1}{2\pi}\big|\langle \widehat{f}, D_{a_0^{-m}}M_{-nb_0}\widehat{\psi}\rangle_{L^2(\mathbb{R})}\big|^2 = \frac{1}{2\pi}\sum_{m,n=-\infty}^{\infty}\bigg|\int_{\mathbb{R}}\widehat{f}(\omega)a_0^{m/2}e^{ib_0a_0^m n\omega}\overline{\widehat{\psi}(a_0^m\omega)}\,d\omega\bigg|^2.\end{aligned}$$

Since, for any $s > 0$, the integral $\int_{\mathbb{R}} g(t)\,dt$ can be written as

$$\sum_{\ell=-\infty}^{\infty} \int_0^s g(t+\ell s)\,dt$$

provided that $g \in L^1(\mathbb{R})$, by taking $s = \dfrac{2\pi}{b_0 a_0^m}$, we obtain

$$\begin{aligned}
&\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 \\
&\quad = \frac{1}{2\pi} \sum_{m,n=-\infty}^{\infty} a_0^m \left| \sum_{\ell=-\infty}^{\infty} \int_0^s \widehat{f}(\omega+\ell s) e^{2\pi i n\omega/s} \overline{\widehat{\psi}(a_0^m(\omega+\ell s))}\,d\omega \right|^2 \\
&\quad = \frac{1}{2\pi} \sum_{m,n=-\infty}^{\infty} a_0^m \left| \int_0^s e^{2\pi i n\omega/s} \left(\sum_{\ell=-\infty}^{\infty} \widehat{f}(\omega+\ell s) \overline{\widehat{\psi}(a_0^m(\omega+\ell s))} \right) d\omega \right|^2 \\
&\quad = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} a_0^m s^2 \sum_{n=-\infty}^{\infty} \left| \frac{1}{s} \int_0^s e^{2\pi i n\omega/s} \left(\sum_{\ell=-\infty}^{\infty} \widehat{f}(\omega+\ell s) \overline{\widehat{\psi}(a_0^m(\omega+\ell s))} \right) d\omega \right|^2, \quad (6.4.13)
\end{aligned}$$

where we have used (a version of) Theorem 1.26 to conclude that the series

$$F_m(\omega) \equiv \sum_{\ell=-\infty}^{\infty} \widehat{f}(\omega+\ell s) \overline{\widehat{\psi}(a_0^m(\omega+\ell s))}$$

converges in $L^1(0,s)$ so we can switch the order of infinite sum and the integration. Next we show that $F_m \in L^2(0,s)$ by showing that the series

$$G_m(\omega) \equiv \sum_{\ell=-\infty}^{\infty} \left|\widehat{f}(\omega+\ell s)\right| \left|\widehat{\psi}(a_0^m(\omega+\ell s))\right|$$

converges in $L^2(0,s)$. To see this, we apply the Monotone Convergence Theorem and find

$$\begin{aligned}
\int_0^s \left|G_m(\omega)\right|^2 d\omega &= \int_0^s \left(\sum_{\ell=-\infty}^{\infty} \left|\widehat{f}(\omega+\ell s)\right| \left|\widehat{\psi}(a_0^m(\omega+\ell s))\right| \right) G_m(\omega)\,d\omega \\
&= \sum_{\ell=-\infty}^{\infty} \int_0^s \left|\widehat{f}(\omega+\ell s)\right| \left|\widehat{\psi}(a_0^m(\omega+\ell s))\right| G_m(\omega)\,d\omega \\
&= \sum_{\ell=-\infty}^{\infty} \int_{\ell s}^{(\ell+1)s} \left|\widehat{f}(\omega)\right| \left|\widehat{\psi}(a_0^m\omega)\right| G_m(\omega-\ell s)\,d\omega \\
&= \int_{\mathbb{R}} \left|\widehat{f}(\omega)\right| \left|\widehat{\psi}(a_0^m\omega)\right| G_m(\omega)\,d\omega\,,
\end{aligned}$$

where the s-periodicity of G_m is used to conclude the last equality. Summing over $m \in \mathbb{Z}$, we obtain that

$$\sum_{m=-\infty}^{\infty} \int_0^s \left|G_m(\omega)\right|^2 d\omega = \sum_{\ell=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \int_{\mathbb{R}} \left|\widehat{f}(\omega)\right| \left|\widehat{\psi}(a_0^m\omega)\right| \left|\widehat{f}(\omega+\ell s)\right| \left|\widehat{\psi}(a_0^m(\omega+\ell s))\right| d\omega\,.$$

By applying Cauchy-Schwarz inequality twice,

$$\begin{aligned}
&\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{f}(\omega+\ell s)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\\
&\quad\leqslant \sum_{m=-\infty}^{\infty}\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\bigg)^{\frac12}\times\\
&\qquad\times\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega+\ell s)\big|^2\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\bigg)^{\frac12}\\
&\quad\leqslant \bigg(\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\bigg)^{\frac12}\times\\
&\qquad\times\bigg(\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}}\big|\widehat{f}(\omega+\ell s)\big|^2\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\bigg)^{\frac12},
\end{aligned}$$

and the Monotone Convergence Theorem again allow us to switch the order of infinite sum and the integration so that

$$\begin{aligned}
&\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{f}(\omega+\ell s)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\\
&\quad\leqslant\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\sum_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\bigg)^{\frac12}\times\\
&\qquad\times\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\sum_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m(\omega-\ell s))\big|\,d\omega\bigg)^{\frac12}.
\end{aligned}$$

Define $\beta(\xi)=\sup\limits_{\omega\in\mathbb{R}}\sum\limits_{m=-\infty}^{\infty}\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{\psi}(a_0^m\omega+\xi)\big|$. The inequality above leads to that

$$\begin{aligned}
&\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|\big|\widehat{\psi}(a_0^m\omega)\big|\big|\widehat{f}(\omega+\ell s)\big|\big|\widehat{\psi}(a_0^m(\omega+\ell s))\big|\,d\omega\\
&\quad\leqslant\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\beta(a_0^m\ell s)\,d\omega\bigg)^{\frac12}\bigg(\int_{\mathbb{R}}\big|\widehat{f}(\omega)\big|^2\beta(-a_0^m\ell s)\big|\,d\omega\bigg)^{\frac12}\\
&\quad\leqslant\|f\|_{L^2(\mathbb{R})}^2\big[\beta(a_0^m\ell s)\beta(-a_0^m\ell s)\big]^{\frac12}=\|f\|_{L^2(\mathbb{R})}^2\Big[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\Big]^{\frac12}.
\end{aligned}\tag{6.4.14}$$

Since condition (ii) implies that $\sum\limits_{\ell=-\infty}^{\infty}\Big[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\Big]^{\frac12}<\infty$, we find that

$$\sum_{m=-\infty}^{\infty}\int_0^s\big|G_m(\omega)\big|^2\,d\omega\leqslant\|f\|_{L^2(\mathbb{R})}^2\sum_{\ell=-\infty}^{\infty}\Big[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\Big]^{\frac12}<\infty.$$

Therefore, $G_m\in L^2(0,s)$ for all $m\in\mathbb{Z}$; thus $F_m\in L^2(0,s)$ for all $m\in\mathbb{Z}$ as well.

We next repeat the steps above (but use the Domninated Convergence Theorem instead of the Monotone Convergence Theorem) to show that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is a frame. By the Parseval identity,

$$\sum_{n=-\infty}^{\infty}\bigg|\frac1s\int_0^s e^{2\pi in\omega/s}\bigg(\sum_{\ell=-\infty}^{\infty}\widehat{f}(\omega+\ell s)\overline{\widehat{\psi}(a_0^m(\omega+\ell s))}\bigg)\,d\omega\bigg|^2=\frac1s\int_0^s\big|F_m(\omega)\big|^2\,d\omega$$

so (6.4.13) and the Dominated Convergence Theorem implies that

$$\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} a_0^m s \int_0^s \left|F_m(\omega)\right|^2 d\omega = \frac{1}{b_0}\sum_{m=-\infty}^{\infty}\int_0^s \left|F_m(\omega)\right|^2 d\omega$$
$$= \frac{1}{b_0}\sum_{m,\ell=-\infty}^{\infty}\int_0^s \widehat{f}(\omega+\ell s)\widehat{\psi}(a_0^m(\omega+\ell s))\overline{F_m(\omega)}\,d\omega$$
$$= \frac{1}{b_0}\sum_{m,\ell=-\infty}^{\infty}\int_{\ell s}^{(\ell+1)s} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{F_m(\omega-\ell s)}\,d\omega\,.$$

Since F_m is s-periodic, the identity above implies that

$$\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2$$
$$= \frac{1}{b_0}\sum_{m,\ell=-\infty}^{\infty}\int_{\ell s}^{(\ell+1)s} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{F_m(\omega)}\,d\omega = \frac{1}{b_0}\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{F_m(\omega)}\,d\omega$$
$$= \frac{1}{b_0}\sum_{m,\ell=-\infty}^{\infty}\int_{\mathbb{R}} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{\widehat{f}(\omega+\ell s)\widehat{\psi}(a_0^m(\omega+\ell s))}\,d\omega$$
$$= \frac{1}{b_0}\sum_{m=-\infty}^{\infty}\int_{\mathbb{R}} \left|\widehat{f}(\omega)\right|^2\left|\widehat{\psi}(a_0^m\omega)\right|^2 d\omega + \frac{1}{b_0}\sum_{\substack{m,\ell\in\mathbb{Z}\\ \ell\neq 0}}\int_{\mathbb{R}} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{\widehat{f}(\omega+\ell s)\widehat{\psi}(a_0^m(\omega+\ell s))}\,d\omega$$
$$= \frac{1}{b_0}\int_{\mathbb{R}} \left|f(\omega)\right|^2 \sum_{m=-\infty}^{\infty}\left|\widehat{\psi}(a_0^m\omega)\right|^2 d\omega + \frac{1}{b_0}\sum_{\substack{m,\ell\in\mathbb{Z}\\ \ell\neq 0}}\int_{\mathbb{R}} \widehat{f}(\omega)\widehat{\psi}(a_0^m\omega)\overline{\widehat{f}(\omega+\ell s)\widehat{\psi}(a_0^m(\omega+\ell s))}\,d\omega\,.$$

Therefore, using (6.4.14) we obtain that

$$\left|\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 - \frac{1}{b_0}\int_{\mathbb{R}}\left|f(\omega)\right|^2 \sum_{m=-\infty}^{\infty}\left|\widehat{\psi}(a_0^m\omega)\right|^2 d\omega\right|$$
$$\leqslant \frac{1}{b_0}\sum_{\substack{m,\ell\in\mathbb{Z}\\ \ell\neq 0}}\int_{\mathbb{R}} \left|\widehat{f}(\omega)\right|\left|\widehat{\psi}(a_0^m\omega)\right|\left|\widehat{f}(\omega+\ell s)\right|\left|\widehat{\psi}(a_0^m(\omega+\ell s))\right| d\omega$$
$$\leqslant \frac{1}{b_0}\|f\|^2_{L^2(\mathbb{R})}\sum_{\substack{\ell\in\mathbb{Z}\\ \ell\neq 0}}\left[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\right]^{\frac{1}{2}} = \frac{2}{b_0}\|f\|^2_{L^2(\mathbb{R})}\sum_{\ell=1}^{\infty}\left[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\right]^{\frac{1}{2}}.$$

Define

$$A(b_0) = \frac{1}{b_0}\bigg(\inf_{\omega\in\mathbb{R}\setminus\{0\}}\sum_{m=-\infty}^{\infty}\left|\widehat{\psi}(a_0^m\omega)\right|^2 - 2\sum_{\ell=1}^{\infty}\left[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\right]^{\frac{1}{2}}\bigg),$$
$$B(b_0) = \frac{1}{b_0}\bigg(\sup_{\omega\in\mathbb{R}\setminus\{0\}}\sum_{m=-\infty}^{\infty}\left|\widehat{\psi}(a_0^m\omega)\right|^2 + 2\sum_{\ell=1}^{\infty}\left[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\right]^{\frac{1}{2}}\bigg).$$

Then

$$A(b_0)\|f\|^2_{L^2(\mathbb{R})} \leqslant \sum_{m,n=-\infty}^{\infty}\left|\langle f, \psi_{m,n}\rangle_{L^2(\mathbb{R})}\right|^2 \leqslant B(b_0)\|f\|^2_{L^2(\mathbb{R})},$$

so using (6.4.12) the theorem is concluded provided we show that

$$\lim_{b_0\to 0}\sum_{\ell=1}^{\infty}\Big[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\Big]^{\frac{1}{2}} = 0. \tag{6.4.15}$$

Nevertheless, condition (ii) implies that $\beta(\xi) \leqslant C(1+|\xi|)^{-1-\delta}$; thus

$$\sum_{\ell=1}^{\infty}\Big[\beta\Big(\frac{2\pi\ell}{b_0}\Big)\beta\Big(-\frac{2\pi\ell}{b_0}\Big)\Big]^{\frac{1}{2}} \leqslant C\sum_{\ell=1}^{\infty}\Big(1+\frac{2\pi\ell}{b_0}\Big)^{-1-\delta} \leqslant C\int_0^{\infty}\Big(1+\frac{2\pi x}{b_0}\Big)^{-1-\delta}dx = \frac{Cb_0}{2\pi\delta};$$

thus (6.4.15) is established. This completes the proof. □

Remark 6.25. Having established Theorem 6.24, we choose b_0 in the form a_0^{-N} for some $N \in \mathbb{N}$ such that $b_0 \in (0,\widetilde{b})$. Let $\phi = \psi_{N,0}$ or $\phi(x) = a_0^{-N/2}\psi(a_0^{-N}x)$. Then for this choice of b_0, we have

$$\begin{aligned}(D_{a_0^m}T_{nb_0}\psi)(x) &= a_0^{-\frac{m}{2}}\psi\big(a_0^{-m}x - na_0^{-N}\big) = a_0^{-\frac{m-N}{2}}a_0^{-\frac{N}{2}}\psi\big(a_0^{-N}(a_0^{-m+N}x-n)\big)\\ &= a_0^{-\frac{m-N}{2}}\phi(a_0^{-m+N}x-n) = \phi_{m-N,n}(x),\end{aligned}$$

where $\phi_{m,n}(x) \equiv a_0^{-m/2}\phi(a_0^{-m}x-n)$. By the fact that $\big\{D_{a_0^m}T_{nb_0}\psi\big\}_{m,n\in\mathbb{Z}}$ is a frame, there exist A, $B>0$ such that

$$A\|f\|^2_{L^2(\mathbb{R})} \leqslant \sum_{m,n\in\mathbb{Z}}\big|\langle f,\phi_{m-N,n}\rangle_{L^2(\mathbb{R})}\big|^2 \leqslant B\|f\|^2_{L^2(\mathbb{R})}.$$

Since $\sum_{m,n\in\mathbb{Z}}\big|\langle f,\phi_{m-N,n}\rangle_{L^2(\mathbb{R})}\big|^2 = \sum_{m,n\in\mathbb{Z}}\big|\langle f,\phi_{m,n}\rangle_{L^2(\mathbb{R})}\big|^2$, we conclude that $\{\phi_{m,n}\}_{m,n\in\mathbb{Z}}$ is also a frame.

6.4.3 Riesz basis

A frame might have redundant vectors which makes the expression of a vector as the linear combination of vectors in a frame not unique. What we really want is a basis. The last part of this sub-section contributes to the Riesz basis.

Definition 6.26. Let H be a Hilbert space. A collection of vectors $\{x_n\}_{n\in\mathcal{I}}$ is called a Riesz basis of H if there exist a bounded linear bijection $T : H \to H$ and an orthonormal basis $\{e_n\}_{n\in\mathcal{I}}$ of H such that $Te_n = x_n$ for every $n \in \mathcal{I}$.

Theorem 6.27. *Let $\{x_n\}$ be a collection of countably many vectors in a separable Hilbert space $(H,\langle\cdot,\cdot\rangle)$. The following two statements are equivalent:*

1. *$\{x_n\}$ is a Riesz basis of H.*

2. *The linear span of $\{x_n\}$ is dense in H, and there exist A, $B>0$ such that*

$$A\sum_n|c_n|^2 \leqslant \Big\|\sum_n c_nx_n\Big\|^2 \leqslant B\sum_n|c_n|^2 \quad \textit{whenever} \quad \sum_n|c_n|^2<\infty. \tag{6.4.16}$$

Proof. W.L.O.G. we assume that H is infinite dimensional, and we write $\{x_n\}$ as $\{x_n\}_{n=1}^{\infty}$.

"$\Rightarrow$" Since $\{x_n\}_{n=1}^{\infty}$ is a Riesz basis of H, there exist a bounded invertible linear map T and an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ of H satisfying $Te_n = x_n$ for all $n \in \mathbb{N}$. By the boundedness and invertibility of T, there exist A, $B > 0$ such that

$$m\|x\| \leqslant \|Tx\| \leqslant M\|x\| \qquad \forall\, x \in H\,.$$

Note that the lower bound is due to the open mapping theorem. Since $x = \sum\limits_{n=1}^{\infty} \langle x, e_n\rangle e_n$, the Parseval identity implies that

$$m^2 \sum_{n=1}^{\infty} |c_n|^2 \leqslant \Big\| T\Big(\sum_{n=1}^{\infty} c_n e_n\Big)\Big\|^2 \leqslant M^2 \sum_{n=1}^{\infty} |c_n|^2 \qquad \forall\, \{c_n\}_{n=1}^{\infty} \in \ell^2.$$

By the boundedness of T and the convergence of $\sum\limits_{n=1}^{\infty} c_n e_n$ in H for any given $\{c_n\}_{n=1}^{\infty} \in \ell^2$, we conclude that

$$\Big\| T\Big(\sum_{n=1}^{\infty} c_n e_n\Big)\Big\| = \lim_{n\to\infty} \Big\| T\Big(\sum_{k=1}^{n} c_k e_k\Big)\Big\| = \lim_{n\to\infty} \Big\| \sum_{k=1}^{n} c_k x_k\Big\| = \Big\| \sum_{n=1}^{\infty} c_n x_n\Big\|.$$

Therefore, (6.4.16) holds for $A = m^2$ and $B = M^2$.

For the denseness of the linear span of $\{x_n\}$, we note that for each $x \in H$, $T^{-1}x \in H$ can be expressed as

$$T^{-1}x = \sum_{n=1}^{\infty} \big\langle T^{-1}x, e_n\big\rangle e_n\,;$$

thus the boundedness of T shows that

$$x = T\Big(\sum_{n=1}^{\infty} \big\langle T^{-1}x, e_n\big\rangle e_n\Big) = \sum_{n=1}^{\infty} \big\langle T^{-1}x, e_n\big\rangle Te_n = \sum_{n=1}^{\infty} \big\langle T^{-1}x, e_n\big\rangle x_n\,.$$

This shows that the linear span of $\{x_n\}$ is dense in H.

"$\Leftarrow$" Suppose that there exist A, $B > 0$ such that (6.4.16) holds. Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal basis of H, and define T on (a subset of) H by

$$Tx = \sum_{n=1}^{\infty} \langle x, e_n\rangle x_n \qquad \text{whenever the RHS makes sense.}$$

Using (6.4.16), we find that $\sum\limits_{n=1}^{\infty} \langle x, e_n\rangle x_n$ converges for all $x \in H$ so that $T : H \to H$ is well-defined and by (6.4.16) again we have

$$A \sum_{n=1}^{\infty} \big|\langle x, e_n\rangle\big|^2 \leqslant \Big\| \sum_{n=1}^{\infty} \langle x, e_n\rangle x_n\Big\|^2 \leqslant B \sum_{n=1}^{\infty} \big|\langle x, e_n\rangle\big|^2 \qquad \forall\, x \in H\,.$$

By the Parseval identity, the inequality above is equivalent to that

$$A\|x\|^2 \leqslant \|Tx\|^2 \leqslant B\|x\|^2 \qquad \forall\, x \in H\,. \tag{6.4.17}$$

Therefore, T is bounded and injective. Next we show that T is surjective by showing that the image of T is dense in H since the image of T must be closed due to (6.4.17).

Suppose that the image of T is not dense in H. Then there exists non-zero $z \in H$ such that $\langle z, Tx\rangle = 0$ for all $x \in H$. In particular,

$$\langle z, x_n\rangle = \langle z, Te_n\rangle = 0 \qquad \forall\, n \in \mathbb{N}.$$

However, since the linear span of $\{x_n\}$ is dense in H, the statement above implies that $z = 0$ which is a contradiction. Therefore, we establish the existence of a bounded surjective linear map $T : H \to H$ with the property that $Te_n = x_n$ for all $n \in \mathbb{N}$; thus $\{x_n\}_{n=1}^{\infty}$ is a Riesz basis of H. □

Theorem 6.28. *Let $\{x_n\}$ be a collection of countably many vectors in a Hilbert space $(H, \langle\cdot,\cdot\rangle)$. Then $\{x_n\}$ is an exact frame in H if and only if $\{x_n\}$ be a Riesz basis of H.*

Proof. Let $\{x_n\}$ be a collection of countably many vectors.

"$\Rightarrow$" Suppose that $\{x_n\}$ is an exact frame with frame bounds A and B, and T is the corresponding frame operator. Note that by Theorem 6.23 $\{x_n\}$ and $\{T^{-1}x_n\}$ are bi-orthonormal; that is, $\langle x_m, T^{-1}x_n\rangle = \delta_{mn}$.

We first show that for every $x \in H$ there exists a unique $\{c_n\}$ such that $x = \sum\limits_n c_n x_n$. The existence of $\{c_n\}$ is clear since $\{x_n\}$ is a frame so that Theorem 6.20 shows that

$$x = \sum_n \langle x, T^{-1}x_n\rangle x_n \qquad \forall\, x \in H.$$

For the uniqueness, we have to show that c_m has to agree with $\langle x, T^{-1}x_m\rangle$ for all $m \in \mathbb{N}$. To see this, we note that

$$\langle x, T^{-1}x_m\rangle = \Big\langle \sum_n c_n x_n, T^{-1}x_m \Big\rangle = \sum_n c_n \langle x_n, T^{-1}x_m\rangle = \sum_n c_n \delta_{nm} = c_m.$$

Therefore, $\{x_n\}$ is a basis.

Next we show that

$$\frac{A^2}{B}\sum_n |c_n|^2 \leqslant \Big\| \sum_n c_n x_n \Big\|^2 \leqslant \frac{B^2}{A}\sum_n |c_n|^2 \qquad \forall\, \{c_n\} \in \ell^2. \tag{6.4.18}$$

Nevertheless, the fact that $\{x_n\}$ is a frame with frame bounds A and B as well as that T is invertible implies that

$$A\|T^{-1}x\|^2 \leqslant \sum_n \big|\langle T^{-1}x, x_n\rangle\big|^2 \leqslant B\|T^{-1}x\|^2 \qquad \forall\, x \in H. \tag{6.4.19}$$

Since $B^{-1}\mathrm{I} \leqslant T^{-1} \leqslant A^{-1}\mathrm{I}$ and T^{-1} is self-adjoint,

$$\frac{1}{B^2}\|x\|^2 \leqslant \frac{1}{B}\langle x, T^{-1}x\rangle \leqslant \frac{1}{B}\|T^{-\frac{1}{2}}x\|^2 \leqslant \langle T^{-\frac{1}{2}}x, T^{-1}T^{-\frac{1}{2}}x\rangle = \|T^{-1}x\|^2$$

and similarly $\|T^{-1}x\|^2 \leqslant \dfrac{1}{A^2}\|x\|^2$. Therefore, (6.4.19) shows that

$$\frac{A}{B^2}\|x\|^2 \leqslant \sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2 \leqslant \frac{B}{A^2}\|x\|^2 \qquad \forall\, x \in H\,,$$

or equivalently,

$$\frac{A^2}{B}\sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2 \leqslant \|x\|^2 \leqslant \frac{B^2}{A}\sum_n \left|\langle x, T^{-1}x_n\rangle\right|^2 \qquad \forall\, x \in H\,.$$

Inequality (6.4.18) then follows from the fact that $x = \sum\limits_n c_n x_n$ if and only if $c_n = \langle x, T^{-1}x_n\rangle$ for all $n \in \mathbb{N}$.

"$\Leftarrow$" Suppose that $\{x_n\}$ is a Riesz basis of H. Then, there exists an orthonormal basis $\{e_n\}$ and a bounded linear bijection $T : H \to H$ such that $Te_n = x_n$ for all n. For $x \in H$, we have

$$\sum_n \left|\langle x, x_n\rangle\right|^2 = \sum_n \left|\langle x, Te_n\rangle\right|^2 = \sum_n \left|\langle T^*x, e_n\rangle\right|^2 = \|T^*x\|^2\,,$$

where T^* is the adjoint of T. On the other hand, the fact that T is a bounded linear bijection shows that $(T^*)^{-1}$ exists and is bounded. Moreover, we have

$$\left\|(T^*)^{-1}\right\|^{-1}\|x\| \leqslant \|T^*x\| \leqslant \|T^*\|\|x\|\,;$$

thus the collection $\{x_n\}$ is a frame (with frame bounds $\left\|(T^*)^{-1}\right\|^{-2}$ and $\|T^*\|^2$). The collection $\{x_n\}$ is obviously an exact frame because it ceases to be a basis whenever any element is deleted from the collection.

This completes the proof. □

6.5 Orthonormal Wavelets

Since the discovery of wavelets, orthonormal wavelets with good time-frequency localization are found to play an important role in wavelet theory and have a great variety of applications. In general, the theory of wavelets begins with a single function $\psi \in L^2(\mathbb{R})$, and a family of functions $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is generated from this single function ψ by the operation of binary dilations (that is, dilation by 2^m) and dyadic translation of $n2^{-m}$ so that

$$\psi_{m,n}(x) = 2^{m/2}\psi\Big(2^m\big(x - \frac{n}{2^m}\big)\Big) = 2^{m/2}\psi(2^m x - n)\,, \tag{6.5.1}$$

where the factor $2^{m/2}$ is introduced to ensure orthonormality so that $\|\psi_{m,n}\|_{L^2(\mathbb{R})} = \|\psi\|_{L^2(\mathbb{R})}$ for all $m, n \in \mathbb{Z}$.

A situation of interest in applications is to deal with an orthonormal family $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$; that is,

$$\langle \psi_{m,n}, \psi_{k,\ell}\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \psi_{m,n}(t)\overline{\psi_{k,\ell}(t)}\,dt = \delta_{mk}\delta_{n\ell}\,, \tag{6.5.2}$$

where m, n, k, $\ell \in \mathbb{Z}$.

To show how the inner products behave in this formalism, we prove the following lemma.

Lemma 6.29. *If ψ and $\phi \in L^2(\mathbb{R})$, then*

$$\langle \psi_{m,k}, \phi_{m,\ell} \rangle_{L^2(\mathbb{R})} = \langle \psi_{n,k}, \phi_{n,\ell} \rangle_{L^2(\mathbb{R})} \tag{6.5.3}$$

for all m, n, k, $\ell \in \mathbb{Z}$.

Proof. By the substitution of variable $x = 2^{n-m}t$, we have

$$\begin{aligned} \langle \psi_{m,k}, \psi_{m,\ell} \rangle_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} 2^m \psi(2^m x - k)\overline{\psi(2^m x - \ell)}\, dx = \int_{\mathbb{R}} 2^n \psi(2^n t - k)\overline{\psi(2^n t - \ell)}\, dt \\ &= \langle \psi_{n,k}, \psi_{n,\ell} \rangle_{L^2(\mathbb{R})}. \end{aligned}$$

□

Definition 6.30 (Orthonormal Wavelet)**.** A wavelet $\psi \in L^2(\mathbb{R})$ is called orthonormal if the family of functions $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ defined by (6.5.1) is an orthonormal basis in $L^2(\mathbb{R})$.

As in the classical Fourier series, the wavelet series for a function $f \in L^2(\mathbb{R})$ based on a given orthonormal wavelet ψ is given by

$$f(x) = \sum_{m,n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x), \tag{6.5.4}$$

where the wavelet coefficients $c_{m,n}$ are given by

$$c_{m,n} = \langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})} \tag{6.5.5}$$

and the double wavelet series (6.5.4) converges to the function f in the L^2-norm.

Example 6.31 (Discrete Haar Wavelet)**.** The simplest example of an orthonormal wavelet is the classic Haar wavelet (6.2.6). To prove this fact, we first show that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal set. With ψ defined by (6.2.6) and $\psi_{m,n}$ defined by (6.5.1), we have

$$\begin{aligned} \langle \psi_{m,n}, \psi_{k,\ell} \rangle_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} 2^{m/2} \psi(2^m x - n) \cdot 2^{k/2} \psi(2^k x - \ell)\, dx \\ &\overset{(2^m x - n = t)}{=} 2^{k/2} 2^{-m/2} \int_{\mathbb{R}} \psi(t) \psi(2^{k-m}(t+n) - \ell)\, dt \end{aligned} \tag{6.5.6}$$

1. For $m = k$, this result gives

   $$\langle \psi_{m,n}, \psi_{k,\ell} \rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \psi(t)\psi(t + n - \ell)\, dt = \delta_{0(n-\ell)} = \delta_{n\ell},$$

   where $\psi(t) \neq 0$ in $0 \leqslant t \leqslant 1$ and $\psi(t - (\ell - n)) \neq 0$ in $\ell - n \leqslant t < 1 + \ell - n$, and these intervals are disjoint from each other unless $n = \ell$.

2. We now consider the case $m \neq k$. In view of symmetry, it suffices to consider the case $m < k$. Putting $r = k - m > 0$ in (6.5.6), we can complete the proof by showing that, for $k \neq m$,

   $$\langle \psi_{m,n}, \psi_{k,\ell} \rangle_{L^2(\mathbb{R})} = 2^{r/2} \int_{\mathbb{R}} \psi(t)\psi(2^r t + s)\, dt,$$

   where $s = 2^r n - \ell \in \mathbb{Z}$. In view of the definition of the Haar wavelet ψ, we must prove that

   $$\int_0^{\frac{1}{2}} \psi(2^r t + s)\, dt - \int_{\frac{1}{2}}^1 \psi(2^r t + s)\, dt = 0.$$

Invoking a simple change of variables $2^r t + s = x$, with $a = s + 2^{r-1}$ and $b = s + 2^r$ we find that

$$\int_0^{\frac{1}{2}} \psi(2^r t + s)\, dt - \int_{\frac{1}{2}}^1 \psi(2^r t + s)\, dt = \int_s^a \psi(x)\, dx - \int_a^b \psi(x)\, dx = 0\,,$$

where we have used the fact that $|a - s| = |b - a| = 2^{r-1} \geqslant 1$ and the integral of ψ on an interval with length not less than 1 is zero to conclude the last equality.

This completes the proof that the Haar wavelet ψ is an orthonormal set.

Next we show that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is indeed a basis in $L^2(\mathbb{R})$. Using (6.5.1) the discrete Haar wavelet is defined by

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n) = \begin{cases} 2^{m/2} & \text{if } \dfrac{n}{2^m} \leqslant t < \dfrac{n+1/2}{2^m}, \\ -2^{m/2} & \text{if } \dfrac{n+1/2}{2^m} \leqslant t \leqslant \dfrac{n+1}{2^m}, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal set, any function $f \in L^2(\mathbb{R})$ can be expanded in the wavelet series in the form

$$f = \sum_{m,n=-\infty}^{\infty} \langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})} \psi_{m,n}\,, \tag{6.5.7}$$

as long as we can show that

$$\|f\|_{L^2(\mathbb{R})}^2 = \sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})}\right|^2 \qquad \forall\, f \in L^2(\mathbb{R})\,. \tag{6.5.8}$$

To prove this, it suffices to show that (6.5.7)/(6.5.8) holds for the function

$$f(t) = \mathbf{1}_{[0,1)}(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

since this will show that (6.5.7)/(6.5.8) also holds for a collection of characteristic functions $\left\{\mathbf{1}_{[\frac{n}{2^m}, \frac{n+1}{2^m})}\right\}_{m,n\in\mathbb{Z}}$ whose linear span is dense in $L^2(\mathbb{R})$. Evidently,

$$\langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})} = 0 \;\text{ for } m \geqslant 0 \text{ or } n \neq 0 \qquad \text{and} \qquad \langle f, \psi_{m,0} \rangle_{L^2(\mathbb{R})} = 2^{\frac{m}{2}} \;\text{ if }\; m < 0\,.$$

Consequently,

$$\sum_{m,n=-\infty}^{\infty} \left|\langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})}\right|^2 = \sum_{m=-\infty}^{-1} \left|\langle f, \psi_{m,0} \rangle_{L^2(\mathbb{R})}\right|^2 = \sum_{m=1}^{\infty} 2^{-m} = 1 = \|f\|_{L^2(\mathbb{R})}^2\,.$$

This verifies (6.5.8).

Example 6.32 (The Discrete Shannon Wavelet)**.** The Shannon function ψ whose Fourier transform satisfies

$$\widehat{\psi}(\omega) = \mathbf{1}_I(\omega)\,, \tag{6.5.9}$$

where $I = [-2\pi, -\pi) \cup [\pi, 2\pi)$ is called the Shannon wavelet. This wavelet ψ can directly be obtained from the inverse Fourier transform of $\widehat{\psi}$ so that

$$\begin{aligned}\psi(t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{i\omega t}\widehat{\psi}(\omega)\,d\omega = \frac{1}{2\pi}\left[\int_{-2\pi}^{-\pi} e^{i\omega}\,d\omega + \int_{\pi}^{2\pi} e^{i\omega}\,d\omega\right] \\ &= \frac{1}{\pi t}\big(\sin 2\pi t - \sin \pi t\big) = \frac{\sin(\pi t/2)}{\pi t/2}\cos\frac{3\pi t}{2} \end{aligned} \tag{6.5.10}$$

Both ψ and $\widehat{\psi}$ are shown in Figure 6.6.

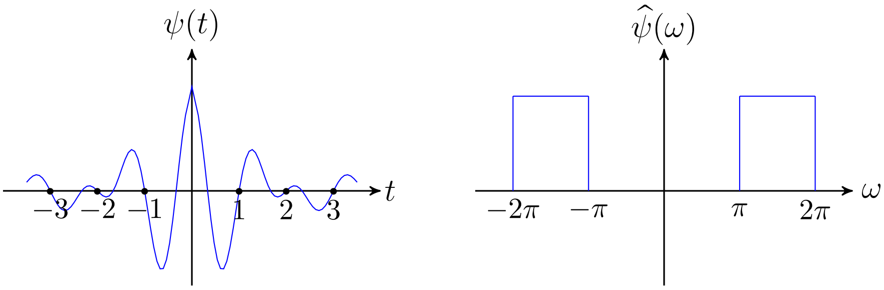

Figure 6.6: The Shannon wavelet and its Fourier transform

We define $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n) = (D_{2^m}T_n\psi)(t)$ whose Fourier transform is given by

$$\widehat{\psi_{m,n}}(\omega) = (D_{2^{-m}}M_{-n}\widehat{\psi})(\omega) = \begin{cases} 2^{\frac{m}{2}}\exp(-i\omega n 2^m) & \text{if } 2^m\omega \in I, \\ 0 & \text{otherwise.}\end{cases} \tag{6.5.11}$$

Evidently, $\operatorname{supp}(\widehat{\psi_{m,n}}) \cap \operatorname{supp}(\widehat{\psi_{k,\ell}}) = \varnothing$ for $m \neq k$. Hence, by the Plancherel identity (1.1.15), it turns out that, for $m \neq k$,

$$\langle \psi_{m,n}, \psi_{k,\ell}\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\langle \widehat{\psi_{m,n}}, \widehat{\psi_{k,\ell}}\rangle_{L^2(\mathbb{R})} = 0\,. \tag{6.5.12}$$

For $m = k$, again by the Plancherel identity (1.1.15) we have

$$\begin{aligned}\langle \psi_{m,n}, \psi_{k,\ell}\rangle_{L^2(\mathbb{R})} &= \frac{1}{2\pi}\langle \widehat{\psi_{m,n}}, \widehat{\psi_{m,\ell}}\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\psi_{m,n}}(\omega)\overline{\widehat{\psi_{m,\ell}}(\omega)}\,d\omega \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} (D_{2^m}M_{-n}\widehat{\psi})(\omega)\overline{(D_{2^m}M_{-\ell}\widehat{\psi})(\omega)}\,d\omega \\ &= \frac{1}{2\pi}\cdot 2^{-m}\int_{\mathbb{R}} \exp\big(-i\omega 2^{-m}(n-\ell)\big)\big|\widehat{\psi}(2^{-m}\omega)\big|^2\,d\omega \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \exp\big(-i(n-\ell)\sigma\big)\big|\widehat{\psi}(\sigma)\big|^2\,d\sigma \\ &= \frac{1}{2\pi}\left(\int_{-2\pi}^{2\pi} - \int_{-\pi}^{\pi}\right)\exp\big(-i(n-\ell)\sigma\big)\,d\omega = \delta_{n\ell}\,. \end{aligned} \tag{6.5.13}$$

This shows that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal set in $L^2(\mathbb{R})$.

Next we show that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal basis in $L^2(\mathbb{R})$. Let $I_m = \big\{\omega \,\big|\, 2^m\omega \in I\big\}$. Note that for a fixed $m \in \mathbb{Z}$, $\big\{\widehat{\psi_{m,n}}\big\}_{n\in\mathbb{Z}}$ is an orthogonal basis of $L^2(I_m)$. To see this, we first note that $L^2(-2^{-m}\pi, 2^{-m}\pi)$ has a Fourier basis $\big\{d_{2^{-m}}e_n\big\}_{n\in\mathbb{Z}}$, where $e_n(x) = e^{inx}$ as defined in

Section 1.1. Since every $f \in L^2(I_m)$ corresponds to a unique function $g \in L^2(-2^{-m}\pi, 2^{-m}\pi)$ satisfying

$$g(x) = \begin{cases} f(x+2^{-m}\pi) & \text{if } 0 \leqslant x < 2^{-m}\pi, \\ f(x-2^{-m}\pi) & \text{if } -2^{-m}\pi \leqslant x < 0, \end{cases}$$

and $g = \sum\limits_{n=-\infty}^{\infty} \big\langle g, d_{2^{-m}} e_n \big\rangle_{L^2(-2^{-m}\pi, 2^m\pi)} d_{2^{-m}} e_n$ in $L^2(-2^{-m}\pi, 2^{-m}\pi)$, the fact that

$$d_{2^{-m}} e_n(x \pm 2^{-m}\pi) = (-1)^n d_{2^{-m}} e_n(x)$$

shows that $f = \sum\limits_{n=-\infty}^{\infty} \big\langle f, d_{2^{-m}} e_n \big\rangle_{L^2(I_m)} d_{2^{-m}} e_n$ in $L^2(I_m)$, where

$$\big\langle f, d_{2^{-m}} e_n \big\rangle_{L^2(I_m)} = \frac{1}{|I_m|} \int_{I_m} f(x) \overline{d_{2^{-m}} e_n(x)}\, dx = \frac{2^{m-1}}{\pi} \int_{I_m} f(x) e^{-in2^m x}\, dx\,.$$

Since $d_{2^{-m}} e_n = 2^{-m/2} \widehat{\psi_{m,-n}}$ for all $m, n \in \mathbb{Z}$, we conclude that $\big\{\widehat{\psi_{m,n}}\big\}_{n\in\mathbb{Z}}$ is an orthogonal basis of $L^2(I_m)$ and we have

$$f = 2^{-m} \sum_{n=-\infty}^{\infty} \big\langle f, \widehat{\psi_{m,-n}} \big\rangle_{L^2(I_m)} \widehat{\psi_{m,-n}} = 2^{-m} \sum_{n=-\infty}^{\infty} \big\langle f, \widehat{\psi_{m,n}} \big\rangle_{L^2(I_m)} \widehat{\psi_{m,n}} \qquad \forall\, f \in L^2(I_m)\,.$$

Now, every $f \in L^2(\mathbb{R})$ can be expressed as

$$\begin{aligned} f &= \mathscr{F}^{-1}\big[\mathscr{F}[f]\big] = \mathscr{F}^{-1}\Big[\sum_{m=-\infty}^{\infty} \widehat{f}\mathbf{1}_{I_m}\Big] = \mathscr{F}^{-1}\Big[\sum_{m=-\infty}^{\infty} 2^{-m} \sum_{n=-\infty}^{\infty} \langle \widehat{f}\mathbf{1}_{I_m}, \widehat{\psi_{m,n}} \rangle_{L^2(I_m)} \widehat{\psi_{m,n}}\Big] \\ &= \mathscr{F}^{-1}\Big[\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \langle \widehat{f}, \widehat{\psi_{m,n}} \rangle_{L^2(\mathbb{R})} \widehat{\psi_{m,n}}\Big] = \mathscr{F}^{-1}\Big[\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})} \widehat{\psi_{m,n}}\Big] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{m,n} \rangle_{L^2(\mathbb{R})} \psi_{m,n}\,. \end{aligned}$$

This shows that the linear span of $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is dense in $L^2(\mathbb{R})$; thus $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal basis in $L^2(\mathbb{R})$.

Chapter 7

Multi-resolution Analysis and Construction of Wavelets

Throughout the chapter, for simplicity we use $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{L^2(\mathbb{R})}$.

7.2 Definition of MRA and Examples

Definition 7.1. An MRA consists of a sequence $\{V_m \mid m \in \mathbb{Z}\}$ of embedded closed subspaces of $L^2(\mathbb{R})$ that satisfy the following conditions:

(i) $\cdots \subseteq V_{-2} \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots$; that is, $V_m \subseteq V_{m+1}$ for all $m \in \mathbb{Z}$;

(ii) $\bigcup\limits_{m=-\infty}^{\infty} V_m$ is dense in $L^2(\mathbb{R})$; that is, $\text{closure}_{\|\cdot\|_2}\Big(\bigcup\limits_{m=-\infty}^{\infty} V_m\Big) = L^2(\mathbb{R})$.

(iii) $\bigcap\limits_{m=\infty}^{\infty} V_m = \{0\}$.

(iv) $f \in V_m$ if and only if $d_{1/2} f \in V_{m+1}$ for all $m \in Z$;

(v) there exists a function $\phi \in V_0$ such that $\big\{\phi_{0,n} = T_n\phi \,\big|\, n \in \mathbb{Z}\big\}$ is an orthonormal basis for V_0; that is,

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{n=-\infty}^{\infty} \big|\langle f, \phi_n\rangle\big|^2 \qquad \forall\, f \in V_0 .$$

The function ϕ is called the scaling function or father wavelet. If $\{V_m\}_{m\in\mathbb{Z}}$ is a multi-resolution of $L^2(\mathbb{R})$ and if V_0 is the closed subspace generated by the integer translates of a single function ϕ, then we say that ϕ generates the MRA.

Sometimes, condition (v) is relaxed by assuming that $\big\{T_n\phi \,\big|\, n \in \mathbb{Z}\big\}$ is a Riesz basis for V_0; that is, for every $f \in V_0$, there exists a unique sequence $\{c_n\}_{n=-\infty}^{\infty} \in \ell^2$ such that

$$f(x) = \sum_{n=-\infty}^{\infty} c_n (T_n\phi)(x) = \sum_{n=-\infty}^{\infty} c_n \phi(x-n) .$$

with convergence in $L^2(\mathbb{R})$ and there exist two positive constants A and B independent of $f \in V_0$ such that

$$A \sum_{n=-\infty}^{\infty} |c_n|^2 \leqslant \|f\|^2_{L^2(\mathbb{R})} \leqslant B \sum_{n=-\infty}^{\infty} |c_n|^2 ,$$

where $0 < A \leqslant B < \infty$. In this case, we have an MRA with a Riesz basis.

Note that condition (v) implies that $\{T_n\phi \,|\, n \in \mathbb{Z}\}$ is a Riesz basis for V_0 with frame bounds $A = B = 1$.

Since $\phi_{0,n} = T_n\phi \in V_0$ for all $n \in \mathbb{Z}$. Further, if $n \in \mathbb{Z}$, it follows from (iv) that

$$\phi_{m,n}(x) = (D_{2^{-m}}T_n\phi)(x) = 2^{\frac{m}{2}}\phi(2^m x - n), \qquad m \in \mathbb{Z} \tag{7.2.1}$$

is an orthonormal basis for V_m.

Consequences of Definition 7.1.

1. A repeated application of condition (iv) implies that $f \in V_m$ if and only if $d_{2^{-k}} \in V_{m+k}$ for all $m, k \in \mathbb{Z}$. In other words, $f \in V_m$ if and only if $d_{2^m} f \in V_0$ for all $m \in \mathbb{Z}$.

   This shows that functions in V_m are obtained from those in V_0 through a scaling 2^{-m}. If the scale $m = 0$ is associated with V_0, then the scale 2^{-m} is associated with V_m. Thus, subspaces V_m are just scaled versions of the central space V_0 which is invariant under translation by integers; that is, $T_n V_0 = V_0$ for all $n \in \mathbb{Z}$.

2. It follows from Definition 7.1 that an MRA is completely determined by the scaling function ϕ, but not conversely. For a given $\phi \in V_0$, we first define

$$V_0 = \left\{ f = \sum_{n=-\infty}^{\infty} c_n\phi_{0,n} = \sum_{n=-\infty}^{\infty} c_n T_n\phi \,\middle|\, \{c_n\}_{n=-\infty}^{\infty} \in \ell^2 \right\}.$$

   Condition (v) implies that V_0 has an orthonormal basis $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$. Then, V_0 consists of all functions $f = \sum\limits_{n=-\infty}^{\infty} c_n T_n\phi$ with finite energy $\|f\|^2_{L^2(\mathbb{R})} = \sum\limits_{n=-\infty}^{\infty} |c_n|^2 < \infty$. Similarly, the space V_m has the orthonormal basis $\{\phi_{m,n}\}_{n\in\mathbb{Z}}$ given by (7.2.1) so that f_m is given by

$$f_m(x) = \sum_{n=-\infty}^{\infty} c_{mn}\phi_{m,n}(x) \tag{7.2.2}$$

   with the finite energy

$$\|f_m\|^2_{L^2(\mathbb{R})} = \sum_{n=-\infty}^{\infty} |c_{mn}|^2 < \infty .$$

   Thus, f_m represents a typical function in the space V_m. It builds in self-invariance and scale invariance through the basis $\{\phi_{m,n}\}_{n\in\mathbb{Z}}$.

3. Conditions (ii) and (iii) can be expressed in terms of the orthogonal projections P_m onto V_m; that is, for all $f \in L^2(\mathbb{R})$,

$$\lim_{m\to-\infty} \mathrm{P}_m f = 0 \qquad \text{and} \qquad \lim_{m\to\infty} \mathrm{P}_m f = f . \tag{7.2.3}$$

   The projection $\mathrm{P}_m f$ can be considered as an approximation of f at the scale 2^{-m}. Therefore, the successive approximations of a given function f are defined as the orthogonal projections P_m onto the space V_m:

$$\mathrm{P}_m f = \sum_{n=-\infty}^{\infty} \langle f, \phi_{m,n} \rangle \phi_{m,n}, \tag{7.2.4}$$

where $\phi_{m,n}$ given by (7.2.1) is an orthonormal basis for V_m.

4. Since $V_0 \subseteq V_1$, the scaling function ϕ that leads to a basis for V_0 also belongs to V_1. Since $\phi \in V_1$ and $\{\phi_{1,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for V_1, ϕ can be expressed in the form

$$\phi(x) = \sum_{n=-\infty}^{\infty} c_n \phi_{1,n}(x) = \sqrt{2} \sum_{n=-\infty}^{\infty} c_n \phi(2x-n), \tag{7.2.5}$$

where

$$c_n = \langle \phi, \phi_{1,n} \rangle \qquad \text{and} \qquad \sum_{n=-\infty}^{\infty} |c_n|^2 = 1.$$

Equation (7.2.5) is called the dilation equation. It involves both x and $2x$ and is often referred to as the two-scale equation or refinement equation because it displays ϕ in the refined space V_1. The space V_1 has the finer scale 2^{-1} and it contains ϕ which has scale 1.

From the preceding facts MRA is described so that we can specify

(a) the subspaces V_m,

(b) the scaling function ϕ,

(c) the coefficient $\{c_n\}_{n=-\infty}^{\infty}$ in the dilation equation (7.2.5).

The real importance of an MRA lies in the simple fact that it enables us to construct an orthonormal wavelet for $L^2(\mathbb{R})$. In order to prove this statement, we first assume that $\{V_m\}_{m=-\infty}^{\infty}$ is an MRA. Since $V_m \subseteq V_{m+1}$, we define W_m as the orthogonal complement of V_m in V_{m+1} for every $m \in \mathbb{Z}$, so that we have

$$\begin{aligned} V_{m+1} &= V_m \oplus W_m = \left(V_{m-1} \oplus W_{m-1}\right) \oplus W_m = \cdots = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_m \\ &= V_0 \oplus \left(\bigoplus_{k=0}^{m} W_k \right) \end{aligned} \tag{7.2.6}$$

and $V_n \perp W_m$ for $n < m$.

Since $\bigcup_{m=-\infty}^{\infty} V_m$ is dense in $L^2(\mathbb{R})$, we may take the limit as $m \to \infty$ to obtain

$$V_0 \oplus \left(\bigoplus_{m=0}^{\infty} W_m \right) = L^2(\mathbb{R}).$$

Similarly, we may go in the other direction to write

$$V_0 = V_{-1} \oplus W_{-1} = (V_{-2} \oplus W_{-2}) \oplus W_{-1} = \cdots = V_{-m} \oplus W_{-m} \oplus \cdots \oplus W_{-1}.$$

We may again take the limit as $m \to \infty$. Since $\bigcap_{m=-\infty}^{\infty} V_m = \{0\}$, it follows that $\lim_{m\to\infty} V_{-m} = \{0\}$. Consequently, it turns out that

$$\bigoplus_{m=-\infty}^{\infty} W_m = L^2(\mathbb{R}). \tag{7.2.7}$$

Finally, the difference between the two successive approximations $\mathrm{P}_m f$ and $\mathrm{P}_{m+1} f$ is given by the orthogonal projection $\mathrm{Q}_m f$ of f onto the orthogonal complement W_m of V_m in V_{m+1} so that

$$\mathrm{Q}_m f = \mathrm{P}_{m+1} f - \mathrm{P}_m f .$$

It follows from conditions (i)-(v) in Definition 7.1 that the spaces W_m are also scaled versions of W_0 and, for $f \in L^2(\mathbb{R})$,

$$f \in W_m \quad \text{if and only if} \quad d_{2^m} f \in W_0 \qquad \forall\, m \in \mathbb{Z} \tag{7.2.8}$$

since

$$f \in W_m \;\Leftrightarrow\; f \in V_{m+1} \text{ and } f \perp V_m \;\Leftrightarrow\; d_{2^m} f \in V_1 \text{ and } d_{2^m} f \perp V_0 \;\Leftrightarrow\; d_{2^m} f \in W_0 .$$

Moreover, W_m's are mutually orthogonal spaces generating all $L^2(\mathbb{R})$; that is,

$$W_m \perp W_k \text{ if } m \neq k \quad \text{and} \quad \bigoplus_{m\in\mathbb{Z}} W_m = L^2(\mathbb{R}), \tag{7.2.9}$$

and are translation-invariant for the discrete translations $n \in \mathbb{Z}$; that is,

$$f \in W_0 \quad \text{if and only if} \quad T_n f \in W_0 ,$$

where the translation-invariant is due to the following equivalence:

$$\begin{aligned} f \in W_0 \Leftrightarrow f \in V_1 \text{ and } f \perp V_0 \Leftrightarrow f \overset{L^2}{=} \sum_{k=-\infty}^{\infty} c_k \phi_{1,k} \text{ for some } \{c_k\}_{k=-\infty}^{\infty} \in \ell^2 \text{ and } f \perp V_0 \\ \Leftrightarrow T_n f \overset{L^2}{=} \sum_{k=-\infty}^{\infty} c_{k-2n} \phi_{1,k} \text{ for some } \{c_k\}_{k=-\infty}^{\infty} \in \ell^2 \text{ and } T_n f \perp V_0 \\ \Leftrightarrow T_n f \in V_1 \text{ and } T_n f \perp V_0 \Leftrightarrow T_n f \in W_0 . \end{aligned}$$

Moreover, it can be shown that there exists a function $\psi \in W_0$ such that $\psi_{0,n} = T_n \psi$ constitutes an orthonormal basis for W_0. It follows from (7.2.8) that

$$\psi_{m,n}(x) = 2^{m/2} \psi(2^m x - n) \qquad \text{for } n \in \mathbb{Z} \tag{7.2.10}$$

constitute an orthonormal basis for W_m. Thus, the family $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ represents an orthonormal basis of wavelets for $L^2(\mathbb{R})$.

Example 7.2 (Characteristic Function and Piecewise Constant Function)**.** We assume that $\phi = \mathbf{1}_{[0,1]}$ is the characteristic function of the interval $[0, 1]$. Define spaces V_m by

$$V_m = \left\{ \sum_{k=-\infty}^{\infty} c_k \phi_{m,k} \,\middle|\, \{c_k\}_{k=-\infty}^{\infty} \in \ell^2 \right\},$$

where

$$\phi_{m,n}(x) = 2^{m/2}\phi(2^m x - n).$$

The spaces V_m satisfy all the conditions of Definition 7.1, and so, $\{V_m\}_{m\in\mathbb{Z}}$ is an MRA.

The space V_m consists of functions in $L^2(\mathbb{R})$ which are constant on intervals $\left[\frac{n}{2^m}, \frac{n+1}{2^m}\right]$, where $n \in \mathbb{Z}$. Obviously, $V_m \subseteq V_{m+1}$ because any function that is constant on intervals of length 2^{-m} is automatically constant on intervals of half that length. The space V_0 contains all functions f in $L^2(\mathbb{R})$ that are constant on $n \leqslant x < n+1$. The function $d_{1/2}f$ in V_1 is then constant on $\frac{n}{2} \leqslant x < \frac{n+1}{2}$. Intervals of length 2^{-m} are usually referred to as dyadic intervals. A sample function in spaces V_m is shown in Figure 7.1.

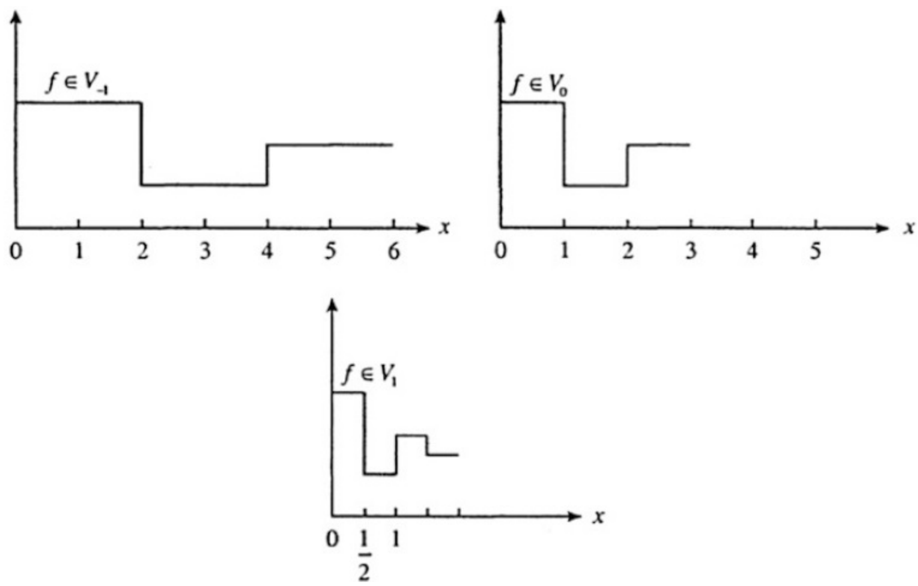

Figure 7.1: Piecewise constant functions in V_{-1}; V_0 and V_1

As we shall see later, this MRA is related to the classic Haar wavelet.

7.3 Properties of Scaling Functions and Orthonormal Wavelet Bases

Lemma 7.3. *Let $f, g \in L^2(\mathbb{R})$. Then the function*

$$\sum_{k=-\infty}^{\infty} [T_{2k\pi}(\widehat{f}\,\overline{\widehat{g}})](\cdot) \equiv \sum_{k=-\infty}^{\infty} \widehat{f}(\cdot + 2k\pi)\overline{\widehat{g}}(\cdot + 2k\pi)$$

belongs to $L^1(0, 2\pi)$, and for all $n \in \mathbb{Z}$,

$$\langle f, T_n g\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{in\omega}\left[\sum_{k=-\infty}^{\infty} \widehat{f}(\omega + 2k\pi)\overline{\widehat{g}}(\omega + 2k\pi)\right] d\omega. \tag{7.3.1}$$

In other words, $\{\langle f, T_{-n}g\rangle\}_{n\in\mathbb{Z}}$ is the Fourier coefficient of $\sum_{k=-\infty}^{\infty} [T_{2k\pi}(\widehat{f}\,\overline{\widehat{g}})]$.

Proof. Let $f, g \in L^2(\mathbb{R})$ and $n \in \mathbb{Z}$ be given. Then $\widehat{f}, \widehat{g} \in L^2(\mathbb{R})$ so that $\widehat{f}\,\overline{\widehat{g}} \in L^1(\mathbb{R})$; thus Theorem 1.26 shows that the series $\sum\limits_{k=-\infty}^{\infty} [T_{2k\pi}(\widehat{f}\,\overline{\widehat{g}})]$ converges in $L^1(0, 2\pi)$. Therefore, the fact that $T_n g \in L^2(\mathbb{R})$ and the Plancherel identity (1.1.15) show that

$$\begin{aligned}
\langle f, T_n g\rangle &= \frac{1}{2\pi}\langle \widehat{f}, \widehat{T_n g}\rangle = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\omega)e^{in\omega}\overline{\widehat{g}}(\omega)\,d\omega = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\int_{2k\pi}^{2(k+1)\pi} \widehat{f}(\omega)\overline{\widehat{g}}(\omega)e^{in\omega}\,d\omega \\
&= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\int_0^{2\pi} \widehat{f}(\omega+2k\pi)\overline{\widehat{g}}(\omega+2k\pi)e^{in\omega}\,d\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{in\omega}\Big[\sum_{k=-\infty}^{\infty} \widehat{f}(\omega+2k\pi)\overline{\widehat{g}}(\omega+2k\pi)\Big]\,d\omega\,.
\end{aligned}$$

□

Corollary 7.4. *Let $f, \phi \in L^2(\mathbb{R})$, and $\{\phi_{0,n} \equiv T_n\phi \,|\, n \in \mathbb{Z}\}$ be an orthonormal sysmtem. Then $\langle f, \phi_{0,n}\rangle = 0$ for all $n \in \mathbb{Z}$ (this can be expressed as $f \perp V_0$, where V_0 is the closure of the linear span of the orthonormal system) if and only if*

$$\sum_{k=-\infty}^{\infty} \widehat{f}(\omega+2k\pi)\overline{\widehat{\phi}}(\omega+2k\pi) = 0 \quad \textit{for a.a. } \omega \in \mathbb{R}. \tag{7.3.2}$$

Proof. Let ϕ be the role of g in Lemma 7.3. Since $\phi_{0,n} = T_n\phi$, (7.3.1) shows that

$$\langle f, \phi_{0,n}\rangle = \frac{1}{2\pi}\int_0^{2\pi}\Big[\sum_{k=-\infty}^{\infty} \widehat{f}(\omega+2k\pi)\overline{\widehat{\phi}}(\omega+2k\pi)\Big]e^{in\omega}\,d\omega\,.$$

Consequently, it follows from the completeness of $\{e^{in\omega} \,|\, n \in \mathbb{Z}\}$ (which holds for functions in $L^1(0, 2\pi)$ as well since the Cesàro mean of the Fourier series of f converges to f in $L^1(0, 2\pi)$ if $f \in L^1(0, 2\pi)$) that $\langle f, \phi_{0,n}\rangle = 0$ for all $n \in \mathbb{Z}$ if and only if (7.3.2) holds. □

Theorem 7.5. *For any function $\phi \in L^2(\mathbb{R})$, the following conditions are equivalent.*

(a) *The system $\{\phi_{0,n} \equiv T_n\phi \,|\, n \in \mathbb{Z}\}$ is orthonormal.*

(b) $\sum\limits_{k=-\infty}^{\infty} \big|\widehat{\phi}(\omega+2k\pi)\big|^2 = 1$ *for a.a.* $\omega \in \mathbb{R}$.

Proof. Letting $f = g = \phi$ in Lemma 7.3, we find that

$$\langle \phi_{0,n}, \phi_{0,m}\rangle = \langle \phi_{0,0}, \phi_{0,m-n}\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)\omega}\sum_{k=-\infty}^{\infty}\big|\widehat{\phi}(\omega+2k\pi)\big|^2\,d\omega\,.$$

By the fact that $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} e^{i(m-n)\omega}\,d\omega = \delta_{nm}$, the identity above implies that

$$\int_0^{2\pi} e^{i(m-n)\omega}\Big[\sum_{k=-\infty}^{\infty}\big|\widehat{\phi}(\omega+2k\pi)\big|^2 - 1\Big]\,d\omega = \langle \phi_{0,n}, \phi_{0,m}\rangle - \delta_{mn} \qquad \forall\, n, m \in \mathbb{Z}.$$

Thus, $\langle \phi_{0,n}, \phi_{0,m}\rangle = \delta_{nm}$ if and only if $\sum\limits_{k=-\infty}^{\infty}\big|\widehat{\phi}(\omega+2k\pi)\big|^2 = 1$ almost everywhere. □

Theorem 7.6. *For any two functions $\phi, \psi \in L^2(\mathbb{R})$, the sets of functions $\{\phi_{0,n} \equiv T_n\phi \,|\, n \in \mathbb{Z}\}$ and $\{\psi_{0,m} \equiv T_m\psi \,|\, m \in \mathbb{Z}\}$ are bi-orthogonal; that is,*

$$\langle \phi_{0,n}, \psi_{0,m} \rangle = 0 \qquad \forall\, n, m \in \mathbb{Z}$$

if and only if

$$\sum_{k=-\infty}^{\infty} \widehat{\phi}(\omega + 2k\pi)\overline{\widehat{\psi}}(\omega + 2k\pi) = 0 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

Proof. Letting $f = \phi$ and $g = \psi$ in Lemma 7.3, we find that

$$\langle \phi_{0,n}, \psi_{0,m} \rangle = \langle \phi, \psi_{0,m-n} \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)\omega} \Big[\sum_{k=-\infty}^{\infty} \widehat{\phi}(\omega + 2k\pi)\overline{\widehat{\psi}}(\omega + 2k\pi) \Big] d\omega .$$

Thus, the same reason for proving Theorem 7.5, we conclude that

$$\langle \phi_{0,n}, \psi_{0,m} \rangle = 0 \qquad \forall\, n, m \in \mathbb{Z}$$

if and only if

$$\sum_{k=-\infty}^{\infty} \widehat{\phi}(\omega + 2k\pi)\overline{\widehat{\psi}}(\omega + 2k\pi) = 0 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$ □

A somewhat weaker property than the property of orthonormality in the previous theorem is the "Riesz (or unconditional) condition", which we study in the following.

Theorem 7.7. *For any function $\phi \in L^2(\mathbb{R})$ and constants $0 < A \leqslant B < \infty$, the following two statements are equivalent:*

(i) $\{T_k\phi = \phi(\cdot - k)\}_{k\in\mathbb{Z}}$ *satisfies the Riesz condition with Riesz bounds A and B; that is,*

$$A\|\{c_k\}_{k\in\mathbb{Z}}\|_{\ell^2}^2 \leqslant \Big\| \sum_{k=-\infty}^{\infty} c_k \phi(\cdot - k) \Big\|_{L^2(\mathbb{R})}^2 \leqslant B\|\{c_k\}_{k\in\mathbb{Z}}\|_{\ell^2}^2 \qquad \forall\, \{c_k\}_{k\in\mathbb{Z}} \in \ell^2 . \tag{7.3.3}$$

(ii) *The Fourier transform $\widehat{\phi}$ of ϕ satisfies*

$$A \leqslant \sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(x + 2k\pi)\big|^2 \leqslant B \qquad \text{for a.a. } x \in \mathbb{R}. \tag{7.3.4}$$

Proof. Let $\phi \in L^2(\mathbb{R})$ and $0 < A \leqslant B < \infty$ be given. For each $n \in \mathbb{Z}$, let $e_n(x) = e^{-inx}$. Since each sequence in ℓ^2 corresponds to a unique function in $L^2(0, 2\pi)$ and vice versa, (7.3.3) is equivalent to that

$$A\|C\|_{L^2(0,2\pi)}^2 \leqslant \Big\| \sum_{k=-\infty}^{\infty} \langle C, e_k \rangle_{L^2(0,2\pi)} T_k\phi \Big\|_{L^2(\mathbb{R})}^2 \leqslant B\|C\|_{L^2(0,2\pi)}^2 \qquad \forall\, C \in L^2(0, 2\pi), \tag{7.3.5}$$

here we recall that $\langle f, g \rangle_{L^2(0,2\pi)} = \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} f(x)\overline{g(x)}\, dx$. Moreover, for every function $C \in$

$L^2(0,2\pi)$, $C(\omega) = \sum\limits_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} e^{-ik\omega}$ for almost all $\omega \in \mathbb{R}$; thus the fact that $\widehat{T_k\phi}(\omega) = e^{-ik\omega}\widehat{\phi}(\omega)$ shows that

$$C(\omega)\widehat{\phi}(\omega) = \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}(\omega) \qquad \text{for a.a. } \omega \in \mathbb{R}$$

so that the monotone convergence theorem further implies that

$$\begin{aligned}\int_{\mathbb{R}} \Big| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}(\omega)\Big|^2 d\omega &= \int_{\mathbb{R}} \big|C(\omega)\widehat{\phi}(\omega)\big|^2 d\omega = \sum_{k=-\infty}^{\infty} \int_{2k\pi}^{2(k+1)\pi} \big|C(\omega)\widehat{\phi}(\omega)\big|^2 d\omega \\ &= \sum_{k=-\infty}^{\infty} \int_0^{2\pi} \big|C(t)\widehat{\phi}(t+2k\pi)\big|^2 dt \\ &= \int_0^{2\pi} \big|C(t)\big|^2 \sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(t+2k\pi)\big|^2 dt. \qquad (7.3.6)\end{aligned}$$

"(i) ⇒ (ii)" Let $C \in L^2(0,2\pi)$ be given. Note that (7.3.5) implies that $\sum\limits_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} T_k\phi$ converges in $L^2(\mathbb{R})$. By the Plancherel identity (1.1.16),

$$\begin{aligned}&\Big\| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} T_k\phi\Big\|^2_{L^2(\mathbb{R})} = \lim_{n\to\infty} \Big\| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} T_k\phi\Big\|^2_{L^2(\mathbb{R})} \\ &\quad = \lim_{n\to\infty} \frac{1}{2\pi} \Big\| \sum_{k=-n}^{n} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}\Big\|^2_{L^2(\mathbb{R})} = \frac{1}{2\pi} \Big\| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}\Big\|^2_{L^2(\mathbb{R})}; \quad (7.3.7)\end{aligned}$$

and (7.3.6) further shows that

$$\Big\| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} T_k\phi\Big\|^2_{L^2(\mathbb{R})} = \frac{1}{2\pi} \int_0^{2\pi} \big|C(t)\big|^2 \sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(t+2k\pi)\big|^2 dt.$$

Therefore, condition (7.3.5) and the Parseval identity show that

$$A\|C\|^2_{L^2(0,2\pi)} \leqslant \frac{1}{2\pi} \int_0^{2\pi} \big|C(t)\big|^2 \sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(t+2k\pi)\big|^2 dt \leqslant B\|C\|^2_{L^2(0,2\pi)} \quad \forall\, C \in L^2(0,2\pi).$$

Let $\{g_n\}_{n=1}^{\infty}$ be an approximation of the identity. Replacing $\frac{1}{2\pi}|C|^2$ by $T_x g_n$ in the inequality above and passing to the limit as $n \to \infty$, we conclude (7.3.4).

"(ii) ⇒ (i)" Let $C \in L^2(0,2\pi)$. Then (7.3.4) and (7.3.6) imply that

$$A\|C\|^2_{L^2(0,2\pi)} \leqslant \frac{1}{2\pi} \int_{\mathbb{R}} \Big| \sum_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}(\omega)\Big|^2 d\omega \leqslant B\|C\|^2_{L^2(0,2\pi)}.$$

This shows the series $\sum\limits_{k=-\infty}^{\infty} \langle C, e_k\rangle_{L^2(0,2\pi)} \widehat{T_k\phi}$ converges in $L^2(\mathbb{R})$. The desired inequality (7.3.5) then follows from (7.3.7). □

We next proceed to the construction of a mother wavelet by introducing a generating function in $L^2(0, 2\pi)$. Before proceeding, we first establish the following

Lemma 7.8. *For every $f \in V_1$, there exists $\widehat{m}_f \in L^2(0, 2\pi)$ such that*

$$\widehat{f}(\omega) = \widehat{m}_f\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big). \tag{7.3.8}$$

Indeed, $\widehat{m}_f$ is given by

$$\widehat{m}_f(\omega) = \frac{1}{\sqrt{2}} \sum_{n=-\infty}^{\infty} \langle f, \phi_{1,n} \rangle e^{-in\omega}. \tag{7.3.9}$$

Proof. Let $c_n = \langle f, \phi_{1,n} \rangle$. Since $f \in V_1$,

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \phi_{1,n}(x) = \sum_{n=-\infty}^{\infty} c_n (D_{1/2} T_n \phi)(x),$$

where $\sum_{n=-\infty}^{\infty} |c_n|^2 = \|f\|^2_{L^2(\mathbb{R})} < \infty$. Using (1.1.18), the Fourier transform of the identity above gives

$$\widehat{f}(\omega) = \sum_{n=-\infty}^{\infty} c_n \big(D_2 M_{-n} \widehat{\phi}\big)(\omega) = \frac{1}{\sqrt{2}} \sum_{n=-\infty}^{\infty} c_n e^{-\frac{in\omega}{2}} \widehat{\phi}\Big(\frac{\omega}{2}\Big). \qquad \square$$

The mother wavelet ψ can be generated by the generating function $\widehat{m} \in L^2(0, 2\pi)$ in the following lemma.

Lemma 7.9. *The Fourier transform of the scaling function ϕ satisfies the following conditions:*

$$\sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(\omega + 2k\pi)\big|^2 = 1 \qquad \text{for a.a. } \omega \in \mathbb{R} \tag{7.3.10}$$

$$\widehat{\phi}(\omega) = \widehat{m}\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big) \tag{7.3.11}$$

where

$$\widehat{m}(\omega) \equiv \widehat{m}_\phi(\omega) = \frac{1}{\sqrt{2}} \sum_{n=-\infty}^{\infty} \langle \phi, \phi_{1,n} \rangle e^{-in\omega} \tag{7.3.12}$$

is a 2π-periodic function and satisfies the so-called the orthogonality condition

$$\big|\widehat{m}(\omega)\big|^2 + \big|\widehat{m}(\omega + \pi)\big|^2 = 1 \quad \text{for a.a. } \omega \in \mathbb{R}. \tag{7.3.13}$$

Remark 7.10. The Fourier transform $\widehat{\phi}$ of the scaling function ϕ satisfies the functional equation (7.3.11). The function $\widehat{m}$ is called the generating function of the MRA. This function is often called the discrete Fourier transform of the sequence $\{c_n\} \equiv \{\langle \phi, \phi_{1,n} \rangle\}$. In signal processing, $\widehat{m}$ is called the transfer function of a discrete filter with impulse response $\{c_n\}$ or the low-pass filter associated with the scaling function ϕ.

Proof. Condition (7.3.10) follows from Theorem 7.5, and (7.3.11) follows from Lemma 7.8 (with f being the scaling function ϕ in the lemma).

To verify the orthogonality condition (7.3.13), we substitute (7.3.11) in (7.3.10) so that condition (7.3.10) becomes

$$1 = \sum_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + 2k\pi)\right|^2 = \sum_{k=-\infty}^{\infty} \left|\widehat{m}\Big(\frac{\omega}{2} + k\pi\Big)\right|^2 \left|\widehat{\phi}\Big(\frac{\omega}{2} + k\pi\Big)\right|^2 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

This is true for almost all $\omega \in \mathbb{R}$ and hence, replacing ω by 2ω gives

$$1 = \sum_{k=-\infty}^{\infty} \left|\widehat{m}(\omega + k\pi)\right|^2 \left|\widehat{\phi}(\omega + k\pi)\right|^2 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

We now split the above infinite sum over k into even and odd integers and use the 2π-periodic property of the function $\widehat{m}$ to obtain that for almost all $\omega \in \mathbb{R}$,

$$\begin{aligned}
1 &= \sum_{k=-\infty}^{\infty} \left|\widehat{m}(\omega + 2k\pi)\right|^2 \left|\widehat{\phi}(\omega + 2k\pi)\right|^2 + \sum_{k=-\infty}^{\infty} \left|\widehat{m}(\omega + (2k+1)\pi)\right|^2 \left|\widehat{\phi}(\omega + (2k+1)\pi)\right|^2 \\
&= \sum_{k=-\infty}^{\infty} \left|\widehat{m}(\omega)\right|^2 \left|\widehat{\phi}(\omega + 2k\pi)\right|^2 + \sum_{k=-\infty}^{\infty} \left|\widehat{m}(\omega + \pi)\right|^2 \left|\widehat{\phi}(\omega + (2k+1)\pi)\right|^2 \\
&= \left|\widehat{m}(\omega)\right|^2 \sum_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + 2k\pi)\right|^2 + \left|\widehat{m}(\omega + \pi)\right|^2 \sum_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + \pi + 2k\pi)\right|^2. \qquad (7.3.14)
\end{aligned}$$

Using (7.3.10),

$$\sum_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + 2k\pi)\right|^2 = \sum_{k=-\infty}^{\infty} \left|\widehat{\phi}(\omega + \pi + 2k\pi)\right|^2 = 1 \qquad \text{for a.a. } \omega \in \mathbb{R};$$

thus (7.3.14) leads to the desired condition (7.3.13). □

The following lemma is useful for reducing the computation in the follow up theorems.

Lemma 7.11. *Let $\{V_m\}_{m\in\mathbb{Z}}$ be an MRA with the scaling function ϕ. Then for all $f, g \in V_1$,*

$$\sum_{k=-\infty}^{\infty} \widehat{f}(\omega + 2k\pi)\overline{\widehat{g}(\omega + 2k\pi)} = \big(\widehat{m}_f\overline{\widehat{m}_g}\big)\Big(\frac{\omega}{2}\Big) + \big(\widehat{m}_f\overline{\widehat{m}_g}\big)\Big(\frac{\omega}{2} + \pi\Big) \quad \textit{for a.a. } \omega \in \mathbb{R}, \ (7.3.15)$$

where $\widehat{m}_f$ and $\widehat{m}_g$ are functions satisfying

$$\widehat{f}(\omega) = \widehat{m}_f\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big), \qquad \widehat{g}(\omega) = \widehat{m}_g\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big).$$

Proof. Using (7.3.8) and (7.3.11) (as well as the case with g replacing f),

$$\begin{aligned}
&\sum_{k=-\infty}^{\infty} \widehat{f}(\omega + 2k\pi)\overline{\widehat{g}(\omega + 2k\pi)} \\
&\qquad = \sum_{k=-\infty}^{\infty} \widehat{m}_f\Big(\frac{\omega}{2} + k\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2} + k\pi\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2} + k\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2} + k\pi\Big)} \\
&\qquad = \sum_{k=-\infty}^{\infty} \widehat{m}_f\Big(\frac{\omega}{2} + k\pi\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2} + k\pi\Big)}\left|\widehat{\phi}\Big(\frac{\omega}{2} + k\pi\Big)\right|^2.
\end{aligned}$$

Splitting the sum into even and odd integers k, by the 2π-periodicity of $\widehat{m}_f$ and $\widehat{m}$ we obtain the for almost all $\omega \in \mathbb{R}$,

$$\begin{aligned}
&\sum_{k=-\infty}^{\infty} \widehat{f}(\omega + 2k\pi)\overline{\widehat{g}(\omega + 2k\pi)} \\
&\quad = \sum_{k=-\infty}^{\infty} \widehat{m}_f\Big(\frac{\omega}{2} + 2k\pi\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2} + 2k\pi\Big)}\Big|\widehat{\phi}\Big(\frac{\omega}{2} + 2k\pi\Big)\Big|^2 \\
&\qquad + \sum_{k=-\infty}^{\infty} \widehat{m}_f\Big(\frac{\omega}{2} + (2k+1)\pi\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2} + (2k+1)\pi\Big)}\Big|\widehat{\phi}\Big(\frac{\omega}{2} + (2k+1)\pi\Big)\Big|^2 \\
&\quad = \widehat{m}_f\Big(\frac{\omega}{2}\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2}\Big)} \sum_{k=-\infty}^{\infty} \Big|\widehat{\phi}\Big(\frac{\omega}{2} + 2k\pi\Big)\Big|^2 \\
&\qquad + \widehat{m}_f\Big(\frac{\omega}{2} + \pi\Big)\overline{\widehat{m}_g\Big(\frac{\omega}{2} + \pi\Big)} \sum_{k=-\infty}^{\infty} \Big|\widehat{\phi}\Big(\frac{\omega}{2} + (2k+1)\pi\Big)\Big|^2 .
\end{aligned}$$

The desired identity (7.3.15) then follows from Theorem 7.5. □

We next prove the following major technical lemma.

Lemma 7.12. *Let $\{V_n\}_{n\in\mathbb{Z}}$ be an MRA with the scaling function ϕ, and $\widehat{m}$ be the associated generating function given by* (7.3.12)*. Then the Fourier transform of any function $f \in W_0$ can be expressed in the form*

$$\widehat{f}(\omega) = \widehat{\nu}(\omega) d_2\Big[M_1\big(\widehat{\phi} T_{-\pi}\overline{\widehat{m}}\big)\Big](\omega) = \widehat{\nu}(\omega) \exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2} + \pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big), \tag{7.3.16}$$

where $\widehat{\nu}$ is a 2π-periodic function satisfying

$$\frac{1}{2\pi}\int_0^{2\pi} \big|\widehat{\nu}(\omega)\big|^2 d\omega = \|f\|^2_{L^2(\mathbb{R})}, \tag{7.3.17}$$

and the factor $d_2\Big[M_1\big(\widehat{\phi} T_{-\pi}\overline{\widehat{m}}\big)\Big]$ is independent of f.

Proof. Since $f \in W_0$, it follows from $V_1 = V_0 \oplus W_0$ that $f \in V_1$ and is orthogonal to V_0. By Lemma 7.8,

$$\widehat{f}(\omega) = \widehat{m}_f\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big). \tag{7.3.18}$$

where, with c_n denoting $\langle f, \phi_{1,n}\rangle$, the function $\widehat{m}_f$ is given by

$$\widehat{m}_f(\omega) = \frac{1}{\sqrt{2}}\sum_{n=-\infty}^{\infty} c_n e^{-in\omega}.$$

Evidently, $\widehat{m}_f$ is a 2π-periodic function which belongs to $L^2(0, 2\pi)$. Moreover, since $f \perp V_0$, by Corollary 7.4 and Lemma 7.11, we have

$$\widehat{m}_f\Big(\frac{\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}\Big) + \widehat{m}_f\Big(\frac{\omega}{2} + \pi\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2} + \pi\Big) = 0 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

Replacing ω by 2ω in the identity above gives

$$0 = \widehat{m}_f(\omega)\overline{\widehat{m}}(\omega) + \widehat{m}_f(\omega+\pi)\overline{\widehat{m}}(\omega+\pi) \qquad \text{for a.a. } \omega \in \mathbb{R},$$

or, equivalently,

$$\begin{vmatrix} \widehat{m}_f(\omega) & \overline{\widehat{m}}(\omega+\pi) \\ -\widehat{m}_f(\omega+\pi) & \overline{\widehat{m}}(\omega) \end{vmatrix} = 0 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

This can be interpreted as the linear dependence of two vectors

$$\big(\widehat{m}_f(\omega), -\widehat{m}_f(\omega+\pi)\big) \qquad \text{and} \qquad \big(\overline{\widehat{m}}(\omega+\pi), \overline{\widehat{m}}(\omega)\big)$$

for almost all $\omega \in \mathbb{R}$. Since (7.3.13) implies that the vector $\big(\overline{\widehat{m}}(\omega+\pi), \overline{\widehat{m}}(\omega)\big)$ is not a zero vector for all $\omega \in \mathbb{R}$, there exists a function $\widehat{\lambda}$, depending on f, such that

$$\big(\widehat{m}_f(\omega), -\widehat{m}_f(\omega+\pi)\big) = \widehat{\lambda}(\omega)\big(\overline{\widehat{m}}(\omega+\pi), \overline{\widehat{m}}(\omega)\big) \qquad \text{for a.a. } \omega \in \mathbb{R}. \tag{7.3.19}$$

Using (7.3.19), the 2π-periodicity of $\widehat{m}$ and $\widehat{m}_f$ implies that $\widehat{\lambda}$ is also 2π-periodic. Furthermore,

$$\widehat{\lambda}(\omega) + \widehat{\lambda}(\omega+\pi) = 0 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

Thus, the function $M_{-1}\widehat{\lambda}$ (or the function $y = e^{-i\omega}\widehat{\lambda}(\omega)$) is π-periodic, so there exists a 2π-periodic function $\widehat{\nu}$ defined by

$$\widehat{\lambda}(\omega) = e^{i\omega}\widehat{\nu}(2\omega). \tag{7.3.20}$$

A simple combination of (7.3.18), (7.3.19), and (7.3.20) gives the desired representation (7.3.16).

Finally, by (7.3.13) the π-periodicity of $|\widehat{\lambda}|$ implies that

$$\begin{aligned}
\int_0^{2\pi} |\widehat{\nu}(\omega)|^2 \, d\omega &= 2\int_\pi^{2\pi} |\widehat{\lambda}(\omega)|^2 \, d\omega = 2\int_\pi^{2\pi} |\widehat{\lambda}(\omega)|^2 \Big(|\widehat{m}(\omega)|^2 + |\widehat{m}(\omega+\pi)|^2\Big) \, d\omega \\
&= 2\bigg[\int_0^{\pi} |\widehat{\lambda}(\omega+\pi)|^2 |\widehat{m}(\omega+\pi)|^2 \, d\omega + \int_\pi^{2\pi} |\widehat{\lambda}(\omega)|^2 |\widehat{m}(\omega+\pi)|^2 \, d\omega\bigg] \\
&= 2\int_0^{2\pi} |\widehat{\lambda}(\omega)|^2 |\widehat{m}(\omega+\pi)|^2 \, d\omega .
\end{aligned}$$

Using (7.3.19) and the Parseval identity,

$$\int_0^{2\pi} |\widehat{\nu}(\omega)|^2 \, d\omega = 2\int_0^{2\pi} |\widehat{m}_f(\omega)|^2 \, d\omega = 2\pi \sum_{n=-\infty}^{\infty} |c_n|^2 = 2\pi \|f\|^2_{L^2(\mathbb{R})} < \infty .$$

This completes the proof of Lemma 7.12. □

Now, we return to the main problem of constructing a mother wavelet $\psi(x)$. Suppose that there is a function ψ such that $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is a basis for the space W_0. Then, every function $f \in W_0$ has a series representation

$$f(x) = \sum_{n=-\infty}^{\infty} h_n \psi_{0,n}(x) = \sum_{n=-\infty}^{\infty} h_n \psi(x-n), \tag{7.3.21}$$

where $h_n = \langle f, \psi_{0,n} \rangle$ satisfies

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{n=-\infty}^{\infty} |h_n|^2 < \infty .$$

Using (1.1.18), the application of the Fourier transform to (7.3.21) gives

$$\hat{f}(\omega) = \Big(\sum_{n=-\infty}^{\infty} h_n e^{-in\omega} \Big) \hat{\psi}(\omega) = \hat{h}(\omega)\hat{\psi}(\omega), \tag{7.3.22}$$

where the function $\hat{h}$ is

$$\hat{h}(\omega) = \sum_{n=-\infty}^{\infty} h_n e^{-in\omega},$$

and it is a square integrable and 2π-periodic function in $[0, 2\pi]$, and the Parseval identity implies that

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{n=-\infty}^{\infty} |h_n|^2 = \frac{1}{2\pi} \int_0^{2\pi} \left|\hat{h}(\omega)\right|^2 d\omega \tag{7.3.23}$$

When (7.3.22) and (7.3.23) are compared with (7.3.16) and (7.3.17), by picking up the terms independent of f we see that one possible choice of $\hat{\psi}$ should be

$$\hat{\psi}(\omega) = \exp\Big(\frac{i\omega}{2}\Big) \overline{\hat{m}}\Big(\frac{\omega}{2} + \pi\Big) \hat{\phi}\Big(\frac{\omega}{2}\Big) \equiv \hat{m}_{\psi}\Big(\frac{\omega}{2}\Big)\hat{\phi}\Big(\frac{\omega}{2}\Big), \tag{7.3.24}$$

where the function $\hat{m}_{\psi}$ is given by

$$\hat{m}_{\psi}(\omega) = \overline{\hat{m}}(\omega + \pi) e^{i\omega} = \overline{\hat{m}_{\phi}}(\omega + \pi) e^{i\omega} .$$

In other words, we choose ψ so that for every $f \in W_0$ we have $\hat{h} = \hat{\nu}$. The function $\hat{m}_{\psi}$ is called the filter conjugate to $\hat{m}$ and hence, $\hat{m}$ and $\hat{m}_{\psi}$ are called conjugate quadratic filters in signal processing.

In the following, we show that the function ψ whose Fourier transform is given by (7.3.24) is indeed an orthonormal wavelet. We start with the following

Lemma 7.13. *Let $\{V_m\}_{m\in\mathbb{Z}}$ be an MRA with the scaling function ϕ and its associated generating function*

$$\hat{m}(\omega) = \frac{1}{\sqrt{2}} \sum_{n=-\infty}^{\infty} c_n e^{-in\omega}, \qquad c_n = \langle \phi, \phi_{1,n} \rangle .$$

Then the Fourier transform of a function ψ is given by

$$\hat{\psi}(\omega) = \exp\Big(\frac{i\omega}{2}\Big) \overline{\hat{m}}\Big(\frac{\omega}{2} + \pi\Big) \hat{\phi}\Big(\frac{\omega}{2}\Big) \tag{7.3.25}$$

if and only if $\psi \in V_1$ takes the form

$$\psi(x) = \sum_{n=-\infty}^{\infty} (-1)^{-n-1} \overline{c_{-n-1}} \phi_{1,n}(x) . \tag{7.3.26}$$

Proof. By the fact that $\{c_n\}_{n\in\mathbb{Z}} \in \ell^2$, (1.1.18) implies that

$$\mathscr{F}\Big[\sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\phi_{1,n}\Big] = \sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\widehat{\phi_{1,n}}.$$

Therefore, by the injectivity of the Fourier transform and the fact that $\{\phi_{1,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for V_1, it suffices to show that

$$\sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\widehat{\phi_{1,n}}(\omega) = \exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big).$$

Nevertheless, the fact that $\phi_{1,n} = D_{1/2}T_n\phi$ implies that

$$\begin{aligned}
\sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\widehat{\phi_{1,n}}(\omega) &= \sum_{n=-\infty}^{\infty}\overline{c_{-n-1}}(-1)^{-n-1}(D_2M_{-n}\widehat{\phi})(\omega)\\
&= \frac{1}{\sqrt{2}}\sum_{n=-\infty}^{\infty}\overline{c_{-n-1}}e^{-i(n+1)\pi}e^{-in\frac{\omega}{2}}\widehat{\phi}\Big(\frac{\omega}{2}\Big) = \frac{1}{\sqrt{2}}\sum_{n=-\infty}^{\infty}\overline{c_n}e^{in\pi+i(n+1)\frac{\omega}{2}}\widehat{\phi}\Big(\frac{\omega}{2}\Big)\\
&= \exp\Big(\frac{i\omega}{2}\Big)\overline{\frac{1}{\sqrt{2}}\sum_{n=-\infty}^{\infty}c_ne^{-in(\frac{\omega}{2}+\pi)}}\widehat{\phi}\Big(\frac{\omega}{2}\Big) = \exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big),
\end{aligned}$$

so the lemma is concluded. □

Lemma 7.14. *Let $\{V_m\}_{m\in\mathbb{Z}}$ be an MRA with the scaling function ϕ and its associated generating function $\widehat{m}$, and ψ be the function whose Fourier transform is given by*

$$\widehat{\psi}(\omega) = \exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big). \tag{7.3.25}$$

Then the system $\{\psi_{0,n} \,|\, n \in \mathbb{Z}\}$ is an orthonormal system in W_0.

Proof. By Lemma 7.13, $\psi \in V_1$, so it suffices to show that $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system and $\psi \perp V_0$. Since $\widehat{\psi}$ is given by (7.3.25), by setting

$$\widehat{m}_\psi(\omega) = \overline{\widehat{m}}(\omega+\pi)e^{i\omega}$$

we have $\widehat{\psi}(\omega) = \widehat{m}_\psi\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big)$. Letting $f = g = \psi$ in (7.3.15), we obtain

$$\sum_{k=-\infty}^{\infty}\big|\widehat{\psi}(\omega+2k\pi)\big|^2 = \Big|\widehat{m}_\psi\Big(\frac{\omega}{2}\Big)\Big|^2 + \Big|\widehat{m}_\psi\Big(\frac{\omega}{2}+\pi\Big)\Big|^2 = \Big|\widehat{m}\Big(\frac{\omega}{2}\Big)\Big|^2 + \Big|\widehat{m}\Big(\frac{\omega}{2}+\pi\Big)\Big|^2,$$

and the orthogonality condition (7.3.13) further shows that

$$\sum_{k=-\infty}^{\infty}\big|\widehat{\psi}(\omega+2k\pi)\big|^2 = 1 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

By Theorem 7.5, $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system. Moreover, letting $f = \psi$ and $g = \phi$ in (7.3.15), we obtain

$$\begin{aligned}
\sum_{k=-\infty}^{\infty}\widehat{\psi}(\omega+2k\pi)\overline{\widehat{\phi}(\omega+2k\pi)} &= \big(\widehat{m}_\psi\overline{\widehat{m}}\big)\Big(\frac{\omega}{2}\Big) + \big(\widehat{m}_\psi\overline{\widehat{m}}\big)\Big(\frac{\omega}{2}+\pi\Big)\\
&= \overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)e^{i\frac{\omega}{2}}\cdot\overline{\widehat{m}}\Big(\frac{\omega}{2}\Big) + \overline{\widehat{m}}\Big(\frac{\omega}{2}+2\pi\Big)e^{i(\frac{\omega}{2}+\pi)}\cdot\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big),
\end{aligned}$$

and the 2π-periodicity of $\widehat{m}$ shows that

$$\sum_{k=-\infty}^{\infty} \widehat{\psi}(\omega + 2k\pi)\overline{\widehat{\phi}(\omega + 2k\pi)} = \Big[e^{i\frac{\omega}{2}} + e^{i(\frac{\omega}{2}+\pi)}\Big]\overline{\widehat{m}}\Big(\frac{\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2} + \pi\Big) = 0.$$

By Corollary 7.4, $\psi \perp V_0$ for all $n \in \mathbb{Z}$. The translation invariant property then shows that $\{\psi_{0,n}\}_{n\in\mathbb{Z}} \perp V_0$; thus $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system in W_0. □

Lemma 7.15. *Let $\vartheta \in L^2(\mathbb{R})$ be such that $\{\vartheta_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system in $L^2(\mathbb{R})$. If $f \in L^2(\mathbb{R})$ and $\widehat{f} = \widehat{\nu}\widehat{\vartheta}$ for some 2π-periodic function $\widehat{\nu} \in L^2(0, 2\pi)$, then*

$$f \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\vartheta_{0,n}\}_{n\in\mathbb{Z}})\big);$$

that is, there exists $\{h_n\}_{n=-\infty}^{\infty} \in \ell^2$ such that $f = \sum\limits_{n=-\infty}^{\infty} h_n\vartheta_{0,n}$.

In particular, if ψ is the function given in Lemma 7.14, then $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of W_0.

Proof. Suppose that $f \in L^2(\mathbb{R})$ satisfies $\widehat{f} = \widehat{\nu}\widehat{\vartheta}$ for some 2π-periodic function $\widehat{\nu} \in L^2(0, 2\pi)$. Since $\widehat{\nu} \in L^2(0, 2\pi)$, there exists $\{h_n\}_{n\in\mathbb{Z}} \in \ell^2$ such that

$$\widehat{\nu}(\omega) = \sum_{n=-\infty}^{\infty} h_n e^{-in\omega}$$

and the convergence is in $L^2(0, 2\pi)$. Therefore,

$$\widehat{f}(\omega) = \sum_{n=-\infty}^{\infty} h_n e^{-in\omega}\widehat{\vartheta}(\omega) = \sum_{n=-\infty}^{\infty} h_n(M_{-n}\widehat{\vartheta})(\omega) = \sum_{n=-\infty}^{\infty} h_n\widehat{T_n\vartheta}(\omega) = \sum_{n=-\infty}^{\infty} h_n\widehat{\vartheta_{0,n}}(\omega).$$

Using (1.1.18) (with $\mathscr{F}^{-1}$ and $\check{}$ replacing $\mathscr{F}$ and $\hat{}$, respectively), the fact that $\left\{\frac{\widehat{\vartheta_{0,n}}}{\sqrt{2\pi}}\right\}_{n\in\mathbb{Z}}$ is an orthonormal system (which is a direct consequence of the Plancherel identity (1.1.15)) and $\{h_n\}_{n\in\mathbb{Z}} \in \ell^2$ imply that

$$f = \sum_{n=-\infty}^{\infty} h_n\vartheta_{0,n}.$$

This shows that $f \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\vartheta_{0,n}\}_{n\in\mathbb{Z}})\big)$.

Next we establish that $\{\psi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of W_0. By Lemma 7.14, it suffices to show that $W_0 \subseteq \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\vartheta_{0,n}\}_{n\in\mathbb{Z}})\big)$. Nevertheless, by Lemma 7.12 every $f \in W_0$ corresponds to a 2π-periodic function $\widehat{\nu} \in L^2(0, 2\pi)$ such that $\widehat{f} = \widehat{\nu}\widehat{\psi}$; thus the argument above then shows that $W_0 \subseteq \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\psi_{0,n}\}_{n\in\mathbb{Z}})\big)$. □

Remark 7.16. The key element to establish Lemma 7.15 is (1.1.18). Due to its similar version (1.5.9), one can relaxed the condition that "$\{\vartheta_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system in $L^2(\mathbb{R})$" to that "$\{\vartheta_{0,n}\}_{n\in\mathbb{Z}}$ has Riesz bounds A and B for some positive A and B".

The combination of Lemma 7.13, 7.14 and 7.15 leads to the main theorem of this section.

Theorem 7.17. *If $\{V_n\}_{n\in\mathbb{Z}}$ is an MRA with the scaling function ϕ, then there is a mother wavelet ψ given by*

$$\psi(x) = \sqrt{2}\sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\phi(2x-n) = \sum_{n=-\infty}^{\infty}(-1)^{-n-1}\overline{c_{-n-1}}\phi_{1,n}(x)\,, \tag{7.3.26}$$

where the coefficients c_n are given by

$$c_n = \langle \phi, \phi_{1,n}\rangle = \sqrt{2}\int_{\mathbb{R}} \phi(x)\overline{\phi(2x-n)}\,dx\,.$$

That is, the system $\{\psi_{m,n}\,|\,n\in\mathbb{Z}\}$ is an orthonormal basis for W_m.

Example 7.18 (The Shannon Wavelet)**.** We consider the Fourier transform $\widehat{\phi}$ of a scaling function ϕ defined by $\widehat{\phi}(\omega) = \mathbf{1}_{[-\pi,\pi]}(\omega)$ so that

$$\phi(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ix\omega}\,d\omega = \operatorname{sinc}(\pi x) = \begin{cases} \dfrac{\sin \pi x}{\pi x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

This is also known as the Shannon sampling function. Clearly, the Shannon scaling function does not have finite support. However, its Fourier transform has a finite support (band-limited) in the frequency domain and has good frequency localization. Evidently, the system

$$\phi_{0,k}(x) = \phi(x-k) = \operatorname{sinc}(\pi(x-k)) \qquad k\in\mathbb{Z}$$

is orthonormal because

$$\begin{aligned}\big\langle \phi_{0,k}, \phi_{0,\ell}\big\rangle_{L^2(\mathbb{R})} &= \big\langle \phi_{0,k-\ell}, \phi_{0,0}\big\rangle_{L^2(\mathbb{R})} = \big\langle T_{k-\ell}\phi, \phi\big\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\big\langle \widehat{T_{k-\ell}\phi}, \widehat{\phi}\big\rangle_{L^2(\mathbb{R})} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(k-\ell)\omega}\,d\omega = \delta_{k\ell}\,.\end{aligned}$$

In general, we define, for $m=0$,

$$V_0 = \left\{ \sum_{k=-\infty}^{\infty} c_k \operatorname{sinc}(\pi(x-k)) \,\middle|\, \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty \right\}$$

and, for other $m\neq 0$, $m\in\mathbb{Z}$,

$$V_m = \left\{ \sum_{k=-\infty}^{\infty} c_k 2^{m/2} \operatorname{sinc}(\pi(2^m x-k)) \,\middle|\, \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty \right\}.$$

It is easy to check that all conditions of Definition 7.1 are satisfied. We next find out the coefficients c_k defined by

$$\begin{aligned} c_k &= \langle \phi, \phi_{1,k}\rangle_{L^2(\mathbb{R})} = \big\langle \phi, D_{\frac{1}{2}}T_k\phi\big\rangle_{L^2(\mathbb{R})} = \frac{1}{2\pi}\big\langle \widehat{\phi}, \widehat{D_{\frac{1}{2}}T_k\phi}\big\rangle_{L^2(\mathbb{R})} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{(D_2M_{-k}\widehat{\phi})(\omega)}\,d\omega = \frac{1}{2\sqrt{2}\pi}\int_{-\pi}^{\pi} e^{-\frac{ik\omega}{2}}\widehat{\phi}\Big(\frac{\omega}{2}\Big)\,d\omega = \frac{1}{2\sqrt{2}\pi}\int_{-\pi}^{\pi} e^{-\frac{ik\omega}{2}}\,d\omega \\ &= \begin{cases} \dfrac{1}{\sqrt{2}} & \text{if } k=0, \\ \dfrac{\sqrt{2}}{k\pi}\sin\dfrac{k\pi}{2} & \text{if } k\neq 0. \end{cases} = \begin{cases} \dfrac{1}{\sqrt{2}} & \text{if } k=0, \\ 0 & \text{if } k \text{ is even and } k\neq 0, \\ \dfrac{\sqrt{2}}{k\pi}(-1)^{\frac{k-1}{2}} & \text{if } k \text{ is odd.} \end{cases}\end{aligned}$$

Consequently, we can use the formula (7.3.26) to find the Shannon mother wavelet

$$\begin{aligned}\psi(x) &= \sqrt{2}\sum_{n=-\infty}^{\infty}(-1)^{n-1}c_{-n-1}\phi(2x-n)\\ &= \sqrt{2}\Big[\frac{1}{\sqrt{2}}\phi(2x+1)+\sum_{n\in\mathbb{Z},n\neq-1}(-1)^{n-1}c_{-n-1}\phi(2x-n)\Big]\\ &= \sqrt{2}\Big[\frac{1}{\sqrt{2}}\phi(2x+1)+\sum_{\ell=-\infty}^{\infty}(-1)^{2\ell-1}c_{-2\ell-1}\phi(2x-2\ell)\Big]\\ &= \mathrm{sinc}\big(\pi(2x+1)\big)-\frac{1}{\pi}\sum_{\ell=-\infty}^{\infty}\frac{2(-1)^{\ell}}{(2\ell+1)}\mathrm{sinc}(2\pi(x-\ell)).\end{aligned}$$

By Theorem 7.17, the system $\{\psi_{m,n}\,|\,m,n\in\mathbb{Z}\}$ is an orthonormal basis in $L^2(\mathbb{R})$. It is known as the Shannon system.

Theorem 7.19. *Let ϕ be a scaling function for an MRA, and ψ be the mother wavelet given by Theorem 7.17. Then a function $\vartheta\in W_0$ is an orthonormal wavelet for $L^2(\mathbb{R})$ if and only if*

$$\widehat{\vartheta}(\omega)=\widehat{\nu}(\omega)\widehat{\psi}(\omega) \tag{7.3.27}$$

for some 2π-periodic function $\widehat{\nu}$ such that $\big|\widehat{\nu}(\omega)\big|=1$.

Proof. First we recall Lemma 7.13 that the mother wavelet ψ satisfies

$$\widehat{\psi}(\omega)=\exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big). \tag{7.3.25}$$

Suppose that $g\in W_0$ satisfying $\widehat{g}=\widehat{\nu}\widehat{\psi}$ for some 2π-periodic $\widehat{\nu}\in L^2(0,2\pi)$. By setting

$$\widehat{m}_g(\omega)=\widehat{\nu}(2\omega)\overline{\widehat{m}}(\omega+\pi)e^{i\omega},$$

we have $\widehat{g}(\omega)=\widehat{m}_g\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big)$. By Lemma 7.11, the 2π-periodicity of $\widehat{m}$ and $\widehat{\nu}$ implies that for almost all $\omega\in\mathbb{R}$,

$$\begin{aligned}\sum_{k=-\infty}^{\infty}\big|\widehat{g}(\omega+2k\pi)\big|^2 &= (\widehat{m}_g\overline{\widehat{m}_g})\Big(\frac{\omega}{2}\Big)+(\widehat{m}_g\overline{\widehat{m}_g})\Big(\frac{\omega}{2}+\pi\Big)=\Big|\widehat{m}_g\Big(\frac{\omega}{2}\Big)\Big|^2+\Big|\widehat{m}_g\Big(\frac{\omega}{2}+\pi\Big)\Big|^2\\ &= \big|\widehat{\nu}(\omega)\big|^2\Big|\widehat{m}\Big(\frac{\omega}{2}+\pi\Big)\Big|^2+\big|\widehat{\nu}(\omega+2\pi)\big|^2\Big|\widehat{m}\Big(\frac{\omega}{2}+2\pi\Big)\Big|^2\\ &= \big|\widehat{\nu}(\omega)\big|^2\Big[\Big|\widehat{m}\Big(\frac{\omega}{2}+\pi\Big)\Big|^2+\Big|\widehat{m}\Big(\frac{\omega}{2}\Big)\Big|^2\Big],\end{aligned}$$

and the orthogonality condition (7.3.13) further implies that

$$\sum_{k=-\infty}^{\infty}\big|\widehat{g}(\omega+2k\pi)\big|^2=\big|\widehat{\nu}(\omega)\big|^2 \qquad \text{for a.a. } \omega\in\mathbb{R}. \tag{7.3.28}$$

"$\Rightarrow$" Suppose that $\vartheta\in W_0$ is an orthonormal wavelet. By Lemma 7.12, there must be a 2π-periodic function $\widehat{\nu}\in L^2(0,2\pi)$ such that (7.3.27) holds. Letting $g=\vartheta$ in (7.3.28), Theorem 7.5 implies that

$$1=\sum_{k=-\infty}^{\infty}\big|\widehat{\vartheta}(\omega+2k\pi)\big|^2=\big|\widehat{\nu}(\omega)\big|^2 \qquad \text{for a.a. } \omega\in\mathbb{R}.$$

"$\Leftarrow$" Suppose that for some $\widehat{\nu}$ satisfying $|\widehat{\nu}| = 1$ a.e. the function ϑ satisfies (7.3.27). Lemma 7.15 shows that $\vartheta \in W_0 \subseteq V_1$; thus letting $g = \vartheta$ in (7.3.28) shows that

$$\sum_{k=-\infty}^{\infty} \big|\widehat{\vartheta}(\omega + 2k\pi)\big|^2 = \big|\widehat{\nu}(\omega)\big|^2 = 1 \qquad \text{for a.a. } \omega \in \mathbb{R}.$$

Theorem 7.5 then implies that $\{\vartheta_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system in W_0.

Next we show that $\{\vartheta_{0,n}\}_{n\in\mathbb{Z}}$ is indeed an orthonormal basis of W_0. Let $f \in W_0$ be given. By Lemma 7.12 there exists a 2π-periodic function $\widehat{\mu} \in L^2(0, 2\pi)$ satisfying

$$\widehat{f}(\omega) = \widehat{\mu}(\omega)\widehat{\psi}(\omega) = \frac{\widehat{\mu}(\omega)}{\widehat{\nu}(\omega)}\widehat{\vartheta}(\omega).$$

Since $\dfrac{\widehat{\mu}}{\widehat{\nu}}$ is 2π periodic and belongs to $L^2(0, 2\pi)$, Lemma 7.15 implies that

$$f \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\vartheta_{0,n}\}_{n\in\mathbb{Z}})\big);$$

thus $W \subseteq \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\vartheta_{0,n}\}_{n\in\mathbb{Z}})\big)$. □

If the scaling function ϕ of an MRA is not an orthonormal basis of V_0 but rather is a Riesz basis, we can use the following orthonormalization process to generate an orthonormal basis.

Theorem 7.20 (Orthonormalization Process)**.** *Let $\phi \in L^2(\mathbb{R})$ be such that $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$ is a Riesz basis of V_0; that is, the linear span of $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$ is dense in V_0 and (by Theorem 7.7) there exists two constants A, $B > 0$ such that*

$$0 < A \leqslant \sum_{k=-\infty}^{\infty} \big|\widehat{\phi}(\omega + 2k\pi)\big|^2 \leqslant B < \infty. \tag{7.3.29}$$

Then $\big\{\widetilde{\phi}_{0,n}\big\}_{n\in\mathbb{Z}}$ is an orthonormal basis of V_0 with

$$\widehat{\widetilde{\phi}}(\omega) = \frac{\widehat{\phi}(\omega)}{\sqrt{\sum\limits_{k=-\infty}^{\infty} \big|\widehat{\phi}(\omega + 2k\pi)\big|^2}}. \tag{7.3.30}$$

Proof. It follows from (7.3.30) that

$$\sum_{k=-\infty}^{\infty} \Big|\widehat{\widetilde{\phi}}(\omega + 2k\pi)\Big|^2 = 1.$$

Thus, the function $\widetilde{\phi}$ satisfies condition (b) of Theorem 7.5, and this shows that $\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal set in $L^2(\mathbb{R})$.

Next we show that $\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}} \subseteq V_0$. We consider a 2π-periodic function $\widehat{\nu}$ defined by

$$\widehat{\nu}(\omega) = \frac{1}{\sqrt{\sum\limits_{k=-\infty}^{\infty} \big|\widehat{\phi}(\omega + 2k\pi)\big|^2}}.$$

Using (7.3.29), $\widehat{\nu} \in L^2(0,2\pi)$. Since $\widehat{\widetilde{\phi}} = \widehat{\nu}\widehat{\phi}$, Lemma 7.15 and Remark 7.16 together imply that

$$\widetilde{\phi} \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\phi_{0,n}\}_{n\in\mathbb{Z}})\big).$$

The fact that $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$ is a Riesz basis of V_0 shows that $\widetilde{\phi} \in V_0$; thus $\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}} \subseteq V_0$.

On the other hand, we also have $\widehat{\phi} = \dfrac{1}{\widehat{\nu}}\widehat{\widetilde{\phi}}$. Since $\dfrac{1}{\widehat{\nu}}$ is 2π-periodic and belongs to $L^2(0,2\pi)$ (due to (7.3.29)), Lemma 7.15 shows that

$$\phi \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}})\big).$$

This further implies that

$$\{\phi_{0,n}\}_{n\in\mathbb{Z}} \subseteq \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}})\big). \tag{7.3.31}$$

Since $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$ is a Riesz basis, we have

$$L^2(\mathbb{R}) = \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\phi_{0,n}\}_{n\in\mathbb{Z}})\big);$$

thus (7.3.31) shows that $L^2(\mathbb{R}) = \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}(\{\widetilde{\phi}_{0,n}\}_{n\in\mathbb{Z}})\big)$. □

Finally, we provide a sufficient condition for a function being a qualified scaling function.

Theorem 7.21. *Let ϕ be a bounded function with compact support, $\widehat{\phi}(0) = 1$, and $\{\phi_{0,n}\}_{n\in\mathbb{Z}}$ is an orthonormal system in $L^2(\mathbb{R})$. If it holds the two-scale equation*

$$\phi(x) = \sum_{n=-\infty}^{\infty} \langle \phi, \phi_{1,n}\rangle \phi_{1,n}(x), \tag{7.2.5}$$

then V_m defined $V_m = \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\big(\text{span}\big(\{\phi_{m,n}\}_{n\in\mathbb{Z}}\big)\big)$ forms an MRA $\{V_m\}_{m\in\mathbb{Z}}$.

Proof. W.L.O.G, we assume that $\text{supp}(\phi) \subseteq [0,L]$ for some $L \in \mathbb{N}$. First we note that $\phi \in L^2(\mathbb{R})$ since it is continuous with compact support. Due to the two-scale equation (7.2.5), we have

$$V_m \subseteq V_{m+1} \quad \forall\, m \in \mathbb{Z} \qquad \text{and} \qquad f \in V_m \Leftrightarrow d_{1/2} f \in V_{m+1}.$$

Therefore, it suffices to show (iii) and (iv) in the Definition of MRA; that is,

$$\text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\Big(\bigcup_{m=-\infty}^{\infty} V_m\Big) = L^2(\mathbb{R}) \qquad \text{and} \qquad \bigcap_{m=-\infty}^{\infty} V_m = \{0\}.$$

We first focus on showing that $\text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\Big(\bigcup_{m=-\infty}^{\infty} V_m\Big) = L^2(\mathbb{R})$. Let $\mathrm{P}_m : L^2(\mathbb{R}) \to V_m$ be the orthogonal projection defined by

$$\mathrm{P}_m f = \sum_{\ell=-\infty}^{\infty} \langle f, \phi_{m,\ell}\rangle \phi_{m,\ell}.$$

By the nature of the orthogonality,

$$\|f\|^2_{L^2(\mathbb{R})} = \|f - \mathrm{P}_m f\|^2_{L^2(\mathbb{R})} + \|\mathrm{P}_m f\|^2_{L^2(\mathbb{R})}.$$

The identity above implies that to show that $\lim\limits_{m\to\infty}\|\mathrm{P}_m f - f\|_{L^2(\mathbb{R})} = 0$, we only need to that $\lim\limits_{m\to\infty}\|\mathrm{P}_m f\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}$.

Claim: For all finite sequence $\{h_k\}_{k=1}^n \subseteq \mathbb{C}$ and $-\infty < a_1 < b_1 \leqslant a_2 < b_2 \leqslant \cdots \leqslant a_k < b_k \leqslant \cdots \leqslant a_n < b_n < \infty$,

$$\sum_{k=1}^n h_k \mathbf{1}_{[a_k,b_k)} \in \mathrm{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\Big(\bigcup_{m=-\infty}^{\infty} V_m\Big).$$

Proof of Claim: We first prove the case that $n=1$ and $h_1 = 1$. Let $a = a_1$ and $b = b_1$. For each $m \in \mathbb{Z}$,

$$\mathrm{P}_m \mathbf{1}_{[a,b)} = \sum_{\ell=-\infty}^{\infty}\Big(\int_a^b \overline{\phi}_{m,\ell}(x)\,dx\Big)\phi_{m,\ell}.$$

Therefore, the orthonormality of $\{\phi_{m,\ell}\}_{\ell\in\mathbb{Z}}$ shows that for $m \gg 1$,

$$\begin{aligned}\|\mathrm{P}_m \mathbf{1}_{[a,b)}\|^2_{L^2(\mathbb{R})} &= \sum_{\ell=-\infty}^{\infty}\Big|\int_a^b \phi_{m,\ell}(x)\,dx\Big|^2 = 2^{-m}\sum_{\ell=-\infty}^{\infty}\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2 \\ &= 2^{-m}\Big(\sum_{\ell=-\infty}^{[2^m a]-L} + \sum_{\ell=[2^m a]-L+1}^{[2^m a]} + \sum_{\ell=[2^m a]+1}^{[2^m b]-L} + \sum_{\ell=[2^m b]-L+1}^{[2^m b]} + \sum_{\ell=[2^m b]+1}^{\infty}\Big)\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2.\end{aligned}$$

1. For $\ell \leqslant [2^m a] - L$ or $\ell \geqslant [2^m b] + 1$, we have

   $$\big[2^m a - \ell, 2^m b - \ell\big] \cap [0, L] = \varnothing.$$

   Therefore,

   $$\Big(\sum_{\ell=-\infty}^{[2^m a]-L} + \sum_{\ell=[2^m b]+1}^{\infty}\Big)\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2 = 0.$$

2. Since $\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big| \leqslant \|\phi\|_{L^1(\mathbb{R})}$, we have

   $$\Big(\sum_{\ell=[2^m a]-L+1}^{[2^m a]} + \sum_{\ell=[2^m b]-L+1}^{[2^m b]}\Big)\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2 \leqslant 2L\|\phi\|^2_{L^1(\mathbb{R})};$$

   thus

   $$\lim_{m\to\infty} 2^{-m}\Big(\sum_{\ell=[2^m a]-L}^{[2^m a]-1} + \sum_{\ell=[2^m b]-L+1}^{[2^m b]}\Big)\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2 = 0.$$

3. For $[2^m a] + 1 \leqslant \ell \leqslant [2^m b] - L$, we have

   $$2^m a - \ell \leqslant 2^m a - [2^m a] - 1 < 0 \qquad \text{and} \qquad 2^m b - \ell \geqslant 2^m b - [2^m b] + L \geqslant L;$$

   thus the fact that $\mathrm{supp}(\phi) \subseteq [0, L]$ implies that

   $$\begin{aligned}\sum_{\ell=[2^m a]+1}^{[2^m b]-L}\Big|\int_{2^m a-\ell}^{2^m b-\ell}\phi(x)\,dx\Big|^2 &= \sum_{\ell=[2^m a]}^{[2^m b]-L}\Big|\int_{\mathbb{R}}\phi(x)\,dx\Big|^2 = \sum_{\ell=[2^m a]}^{[2^m b]-L}\big|\widehat{\phi}(0)\big|^2 \\ &= [2^m b] - [2^m a] - L.\end{aligned}$$

Therefore,

$$\lim_{m\to\infty} 2^{-m} \sum_{\ell=[2^m a]}^{[2^m b]-L} \left| \int_{2^m a-\ell}^{2^m b-\ell} \phi(x)\,dx \right|^2 = \lim_{m\to\infty} 2^{-m}\big([2^m b]-[2^m a]-L\big)$$
$$= \lim_{m\to\infty} 2^{-m}\Big[2^m b - 2^m a + ([2^m b]-2^m b) + (2^m a - [2^m a]) - L\Big] = b-a\,.$$

The discussion above shows that $\lim\limits_{m\to\infty} \big\|\mathrm{P}_m \mathbf{1}_{[a,b)}\big\|_{L^2(\mathbb{R})} = b-a = \|\mathbf{1}_{[a,b)}\|^2_{L^2(\mathbb{R})}$.

Now if $s = \sum\limits_{k=1}^{n} h_k \mathbf{1}_{[a_k,b_k)}$, where $\{h_k\}_{k=1}^n \subseteq \mathbb{C}$ and $-\infty < a_1 < b_1 \leqslant a_2 < b_2 \leqslant \cdots \leqslant a_n < b_n < \infty$, the linearity of P_m implies that

$$\big\|\mathrm{P}_m s - s\big\|_{L^2(\mathbb{R})} = \bigg\| \sum_{k=1}^{n} h_k \Big(\mathrm{P}_m \mathbf{1}_{[a_k,b_k)} - \mathbf{1}_{[a_k,b_k)}\Big)\bigg\|_{L^2(\mathbb{R})} \leqslant \sum_{k=1}^{n} |h_k| \Big\|\mathrm{P}_m \mathbf{1}_{[a_k,b_k)} - \mathbf{1}_{[a_k,b_k)}\Big\|_{L^2(\mathbb{R})}$$

which converges to 0 as $m \to \infty$ (because it is a finite sum). This concludes the claim.

Let $f \in L^2(\mathbb{R})$ be given, and $\varepsilon > 0$ be given. There exists finite sequence $\{h_k\}_{k=1}^n \subseteq \mathbb{C}$ and $2n$ real numbers satisfying $-\infty < a_1 < b_1 \leqslant a_2 < b_2 \leqslant \cdots \leqslant a_n < b_n < \infty$ such that the step function $s = \sum\limits_{k=-\infty}^{\infty} h_k \mathbf{1}_{[a_k,b_k)}$ satisfies

$$\|f - s\|_{L^2(\mathbb{R})} < \frac{\varepsilon}{3}\,.$$

The orthogonality of P_m further shows that

$$\big\|\mathrm{P}_m f - \mathrm{P}_m s\big\|_{L^2(\mathbb{R})} \leqslant \|f - s\|_{L^2(\mathbb{R})} < \frac{\varepsilon}{3}\,.$$

By Claim 1, there exists $N > 0$ such that

$$\big\|\mathrm{P}_m s - s\big\|_{L^2(\mathbb{R})} < \frac{\varepsilon}{3} \qquad \text{whenever} \quad m \geqslant N\,.$$

Therefore, if $m \geqslant N$,

$$\|f - \mathrm{P}_m f\|_{L^2(\mathbb{R})} \leqslant \|f - s\|_{L^2(\mathbb{R})} + \big\|\mathrm{P}_m s - s\big\|_{L^2(\mathbb{R})} + \big\|\mathrm{P}_m f - \mathrm{P}_m s\big\|_{L^2(\mathbb{R})} < \varepsilon\,,$$

and this shows that $f \in \text{closure}_{\|\cdot\|_{L^2(\mathbb{R})}}\Big(\bigcup\limits_{m=-\infty}^{\infty} V_m\Big)$.

Next we show that $\bigcap\limits_{m=-\infty}^{\infty} V_m = \{0\}$ or equivalently,

$$\lim_{m\to-\infty} \|\mathrm{P}_m f\|_{L^2(\mathbb{R})} = 0 \qquad \forall\, f \in L^2(\mathbb{R})\,.$$

Let $f \in L^2(\mathbb{R})$ and $\varepsilon > 0$ be given. Choose $g \in \mathscr{C}_c(\mathbb{R})$ be such that $\|f - g\|_{L^2(\mathbb{R})} < \frac{\varepsilon}{2}$. Suppose that $\text{supp}(g) \subseteq [-R, R]$ and $N \in \mathbb{N}$ satisfies $2^N \geqslant R$. Then the fact that $\text{supp}(\phi_{m,n}) \subseteq [2^{-m}n, 2^{-m}(n+L)]$ implies that if $m \leqslant -N$,

$$\langle g, \phi_{m,n}\rangle = \int_{-R}^{R} g(x)\overline{\phi_{m,n}}(x)\,dx = 0 \qquad \text{whenever} \quad n \leqslant -L-1 \text{ or } n \geqslant 1\,.$$

The discussion above shows that if $m \leqslant -N$,

$$\mathrm{P}_m g = \sum_{n=-\infty}^{\infty} \langle g, \phi_{m,n} \rangle \phi_{m,n} = \sum_{n=-L}^{0} \langle g, \phi_{m,n} \rangle \phi_{m,n} \,.$$

Therefore, using the estimate

$$\left| \langle g, \phi_{m,n} \rangle \right| = \left| \int_{\mathbb{R}} g(x) 2^{\frac{m}{2}} \overline{\phi}(2^m x - n)\, dx \right| \leqslant 2^{\frac{m}{2}} \|\phi\|_{L^\infty(\mathbb{R})} \|g\|_{L^1(\mathbb{R})} \,,$$

we have for $m \leqslant -N$,

$$\|\mathrm{P}_m g\|^2_{L^2(\mathbb{R})} = \sum_{n=-L}^{0} \left| \langle g, \phi_{m,n} \rangle \right|^2 \leqslant 2^m (L+1) \|\phi\|^2_{L^\infty(\mathbb{R})} \|g\|^2_{L^1(\mathbb{R})} \,.$$

Therefore, by choosing N even larger, we have for $m \leqslant -N$,

$$\|\mathrm{P}_m f\|_{L^2(\mathbb{R})} \leqslant \|\mathrm{P}_m (f-g)\|_{L^2(\mathbb{R})} + \|\mathrm{P}_m g\|_{L^2(\mathbb{R})} \leqslant \|f-g\|_{L^2(\mathbb{R})} + \|\mathrm{P}_m g\|_{L^2(\mathbb{R})} < \varepsilon$$

which concludes that $\lim\limits_{m \to 0} \|\mathrm{P}_m f\|_{L^2(\mathbb{R})} = 0$ for all $f \in L^2(\mathbb{R})$. □

7.4 Construction of Orthonormal Wavelets

We now use the properties of scaling functions and filters for constructing orthonormal wavelets.

Example 7.22 (The Haar Wavelet)**.** Example 7.2 shows that spaces of piecewise constant functions constitute an MRA with the scaling function $\phi = \mathbf{1}_{[0,1)}$. Moreover, ϕ satisfies the dilation equation

$$\phi(x) = \sqrt{2} \sum_{n=-\infty}^{\infty} c_n \phi(2x - n) \,, \tag{7.4.1}$$

where the coefficients c_n are given by

$$c_n = \langle \phi, \phi_{1,n} \rangle_{L^2(\mathbb{R})} = \sqrt{2} \int_{\mathbb{R}} \phi(x) \phi(2x-n)\, dx \,. \tag{7.4.2}$$

Evaluating this integral with $\phi = \mathbf{1}_{[0,1)}$ gives c_n as follows:

$$c_0 = c_1 = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad c_n = 0 \ \text{ for } n \neq 0,\, 1.$$

Consequently, the dilation equation becomes

$$\phi(x) = \phi(2x) + \phi(2x-1) \,. \tag{7.4.3}$$

This means that ϕ is a linear combination of the even and odd translates of $d_{1/2}\phi$ and satisfies a very simple two-scale relation (7.4.3), as shown in Figure 7.2.

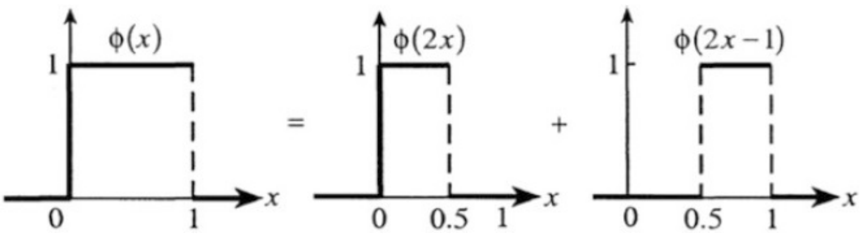

Figure 7.2: Two-scale relation of $\phi(x) = \phi(2x) + \phi(2x-1)$

Thus, the Haar mother wavelet is obtained from (7.3.26) as a simple two-scale relation

$$\psi(x) = \phi_{1,-1}(x) - \phi_{1,-2}(x) = \phi(2x+1) - \phi(2x+2) = \mathbf{1}_{[-\frac{1}{2},0)}(x) - \mathbf{1}_{[-1,-\frac{1}{2})}(x)$$
$$= \begin{cases} 1 & \text{if } x \in \big[-\frac{1}{2}, 0\big), \\ -1 & \text{if } x \in \big[-1, -\frac{1}{2}\big), \\ 0 & \text{otherwise.} \end{cases} \tag{7.4.4}$$

This two-scale relation (7.4.4) of ψ is represented in Figure 7.3.

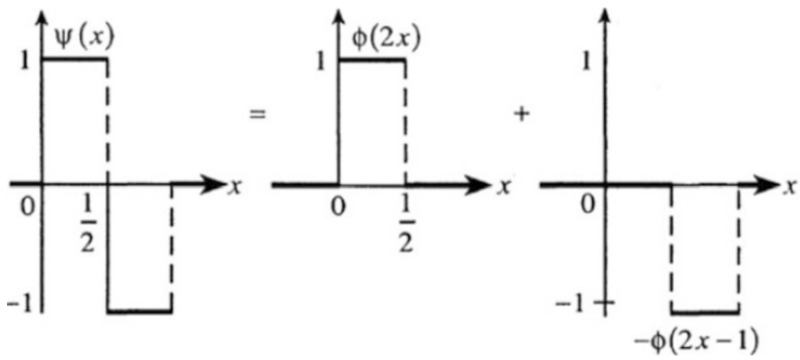

Figure 7.3: Two-scale relation of $\psi(x) = \phi(2x) + \phi(2x-1)$

Alternatively, the Haar wavelet can be obtained from the Fourier transform of the scaling function $\phi = \mathbf{1}_{[0,1)}$ so that

$$\begin{aligned} \widehat{\phi}(\omega) &= \widehat{\mathbf{1}_{[0,1)}}(\omega) = \exp\Big(-\frac{i\omega}{2}\Big)\text{sinc}\Big(\frac{\omega}{2}\Big) \\ &= \exp\Big(-\frac{i\omega}{4}\Big)\cos\Big(\frac{\omega}{4}\Big)\exp\Big(-\frac{i\omega}{4}\Big)\text{sinc}\Big(\frac{\omega}{4}\Big) \\ &= \widehat{m}\Big(\frac{\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big), \end{aligned} \tag{7.4.5}$$

where the associated filter $\widehat{m}$ and its complex conjugate are given by

$$\widehat{m}(\omega) = \exp\Big(-\frac{i\omega}{2}\Big)\cos\Big(\frac{\omega}{2}\Big) = \frac{1}{2}(1+e^{-i\omega}), \tag{7.4.6a}$$

$$\overline{\widehat{m}}(\omega) = \exp\Big(\frac{i\omega}{2}\Big)\cos\Big(\frac{\omega}{2}\Big) = \frac{1}{2}(1+e^{i\omega}). \tag{7.4.6b}$$

Thus, the Haar wavelet can be obtained form (7.3.24) or (7.3.27) and is given by

$$\begin{aligned} \widehat{\psi}(\omega) &= \widehat{\nu}(\omega)\exp\Big(\frac{i\omega}{2}\Big)\overline{\widehat{m}}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big) \\ &= \widehat{\nu}(\omega)\cdot\exp\Big(\frac{i\omega}{2}\Big)\cdot\frac{1}{2}\big(1-e^{\frac{i\omega}{2}}\big)\cdot\widehat{\phi}\Big(\frac{\omega}{2}\Big), \end{aligned}$$

where $\widehat{\nu}(\omega) = -\exp(-i\omega)$ is chosen to find the exact result (7.4.4) since using this $\widehat{\nu}$, we obtain

$$\begin{aligned}\widehat{\psi}(\omega) &= -\exp\Big(-\frac{i\omega}{2}\Big)\cdot\frac{1}{2}\big(1-e^{\frac{i\omega}{2}}\big)\cdot\widehat{\phi}\Big(\frac{\omega}{2}\Big) = \frac{1}{2}\widehat{\phi}\Big(\frac{\omega}{2}\Big) - \frac{1}{2}\exp\Big(-\frac{i\omega}{2}\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big)\\ &= \frac{1}{\sqrt{2}}\big[(D_2\widehat{\phi})(\omega) - (M_{-\frac{1}{2}}D_2\widehat{\phi})(\omega)\big] = \frac{1}{\sqrt{2}}\big[\widehat{D_{\frac{1}{2}}\phi(\omega) - T_{\frac{1}{2}}D_{\frac{1}{2}}\phi(\omega)}\big]\end{aligned}$$

so that the inverse Fourier transform gives the exact result (7.4.4) as

$$\psi(x) = \frac{1}{\sqrt{2}}\big[(D_{\frac{1}{2}}\phi)(x) - (T_{\frac{1}{2}}D_{\frac{1}{2}}\phi)(x)\big] = \phi(2x) - \phi(2x-1).$$

On the other hand, using (7.3.24) also gives the Haar wavelet as

$$\begin{aligned}\widehat{\psi}(\omega) &= \exp\Big(\frac{i\omega}{2}\Big)\overline{m}\Big(\frac{\omega}{2}+\pi\Big)\widehat{\phi}\Big(\frac{\omega}{2}\Big) = \exp\Big(\frac{i\omega}{2}\Big)\cdot\frac{1}{2}\big(1-e^{\frac{i\omega}{2}}\big)\cdot\exp\Big(-\frac{i\omega}{4}\Big)\mathrm{sinc}\Big(\frac{\omega}{4}\Big)\\ &= \exp\Big(\frac{i\omega}{2}\Big)\cdot\frac{1}{2}\big(e^{\frac{-i\omega}{4}} - e^{\frac{i\omega}{4}}\big)\cdot\mathrm{sinc}\Big(\frac{\omega}{4}\Big) = -\exp\Big(\frac{i\omega}{2}\Big)\cdot\sin\frac{\omega}{4}\cdot\mathrm{sinc}\Big(\frac{\omega}{4}\Big)\\ &= \left[i\exp\Big(-\frac{i\omega}{2}\Big)\frac{\sin^2(\omega/4)}{\omega/4}\right]\big(-\exp(i\omega)\big). \qquad (7.4.7)\end{aligned}$$

This corresponds to the same Fourier transform (6.2.7) of the Haar wavelet (7.4.4) except for the factor $-\exp(i\omega)$. This means that this factor induces a translation of the Haar wavelet to the left by one unit. Thus, we have chosen $\widehat{\nu}(\omega) = -\exp(-i\omega)$ in (7.3.27) to find the same value (7.4.4) for the classic Haar wavelet.

Example 7.23 (Cardinal B-splines and Spline Wavelets)**.** When we talk about "cardinal splines", we mean "polynomial spline functions with equally spaced simple knots". We first consider the set $\mathbb{Z}$ of all integers as the "knot sequence". Let π_n denotes the collection of all algebraic polynomials of degree at most n, and $\mathscr{C}^n(\mathbb{R})$ denote the collection of all functions f such that f, f', $\cdots$, $f^{(n)}$ are continuous everywhere, with the understanding that $\mathscr{C}^{-1}(\mathbb{R})$ is the space of piecewise continuous functions.

For each positive integer m, the space S_m of cardinal splines of order m and with knot sequence $\mathbb{Z}$ is the collection of all functions $f \in \mathscr{C}^{m-2}(\mathbb{R})$ such that the restrictions of f to any interval $[k, k+1)$, $k \in \mathbb{Z}$, are in π_{m-1}; that is,

$$f\restriction_{[k,k+1)} \in \pi_{m-1}, \qquad k \in \mathbb{Z}.$$

The space S_1 of piecewise constant functions is easy to understand. The most convenient basis to use is $\{B_1(x-k)\,|\,k \in \mathbb{Z}\}$, where $B_1 = \mathbf{1}_{[0,1)}$ is the scaling function given in the previous example. To give a basis of S_m, $m \geqslant 2$, let us first consider the space $S_{m,N}$ consisting of the restrictions of functions $f \in S_m$ to the interval $[-N, N]$, where N is a positive integer. In other words, we may consider $S_{m,N}$ as the subspace of functions $f \in S_m$ such that the restrictions

$$f\restriction_{(-\infty,-N+1)} \qquad \text{and} \qquad f\restriction_{[N-1,\infty)}$$

of f are polynomials in π_{m-1}. This subspace is easy to characterize. Indeed, for an arbitrary function f in $S_{m,N}$, by setting $p_{m,j} = f\restriction_{[j,j+1)} \in \pi_{m-1}$, $j = -N, \cdots, N-1$, we have, in view of the fact that $f \in \mathscr{C}^{m-2}$,

$$\left(p_{m,j}^{(\ell)} - p_{m,j-1}^{(\ell)}\right)(j) = 0, \qquad \ell = 0, 1, \cdots, m-2;\ m \geqslant 2.$$

That is, by considering the "jumps" of $f^{(m-1)}$ at the knot sequence $\mathbb{Z}$, namely:

$$\begin{aligned} c_j &= p_{m,j}^{(m-1)}(j^+) - p_{m,j-1}^{(m-1)}(j^-) \equiv \lim_{x \to j^+} p_{m,j}^{(m-1)}(x) - \lim_{x \to j^0} p_{m,j-1}^{(m-1)}(x) \\ &= \lim_{\varepsilon \to 0^+} \left[f^{(m-1)}(j+\varepsilon) - f^{(m-1)}(j-\varepsilon)\right], \end{aligned} \tag{7.4.8}$$

the adjacent polynomial pieces of f are related by the identity

$$p_{m,j}(x) = p_{m,j-1}(x) + \frac{c_j}{(m-1)!}(x-j)^{m-1}. \tag{7.4.9}$$

Hence, by introducing the notation

$$\begin{cases} x_+ = \max\{0, x\}, \\ x_+^{m-1} = (x_+)^{m-1},\ m \geqslant 2, \end{cases} \tag{7.4.10}$$

it follows from (7.4.9), that for all $x \in [-N, N]$,

$$f(x) = f\restriction_{[-N,-N+1)}(x) + \sum_{j=-N+1}^{N-1} \frac{c_j}{(m-1)!}(x-j)_+^{m-1}. \tag{7.4.11}$$

This holds for every $f \in S_{m,N}$, with the constants c_j given by (7.4.8). Consequently, the collection

$$\left\{1, \cdots, x^{m-1}, (x+N-1)_+^{m-1}, \cdots, (x-N+1)_+^{m-1}\right\} \tag{7.4.12}$$

of $m + 2N - 1$ functions is a basis of $S_{m,N}$. This collection consists of both monomials and "truncated powers". Since we restrict our attention to the interval $[-N, N]$, it is also possible to replace the monomials $1, \cdots, x^{m-1}$ in (7.4.12) by the truncated powers:

$$(x+N+m-1)_+^{m-1}, \cdots, (x+N)_+^{m-1}. \tag{7.4.13}$$

That is, the following set of truncated powers, which are generated by using integer translates of a single function x_+^{m-1}, is also a basis of $S_{m,N}$:

$$\left\{(x-k)_+^{m-1} \,\middle|\, k = -N-m+1, \cdots, N-1\right\}. \tag{7.4.14}$$

This basis is more attractive than the basis in (7.4.12) for the following reasons. Firstly, each function $(x-j)_+^{m-1}$ vanishes to the left of j; secondly all the basis functions in (7.4.14) are generated by a single function x_+^{m-1} which is independent of N. Moreover, since

$$S_m = \bigcup_{N=1}^{\infty} S_{m,N},$$

it follows that the basis in (7.4.14) can also be extended to be a "basis" $\mathcal{T}$ of the infinite dimensional space S_m, simply by taking the union of the bases in (7.4.14); that is, we have

$$\mathcal{T} = \left\{ (x-k)_+^{m-1} \,\middle|\, k \in \mathbb{Z} \right\}. \tag{7.4.15}$$

Since we are mainly concerned with the Hilbert space $L^2(\mathbb{R})$, we are especially interested in cardinal splines that are in $L^2(\mathbb{R})$. Unfortunately, there is not a single function in $\mathcal{T}$ that qualifies to be a function in $L^2(\mathbb{R})$, and in fact, each $(x-k)_+^{m-1}$ grows to infinity fairly rapdily as $x \to +\infty$. To create $L^2(\mathbb{R})$ functions from $\mathcal{T}_N$, we must tame the polynomial growth of $(x-k)_+^{m-1}$.Define the difference operator Δ recursively by

$$\begin{cases} (\Delta f)(x) = f(x) - f(x-1), \\ (\Delta^n f)(x) = \big(\Delta^{n-1}(\Delta f)\big)(x), \qquad n = 2, 3, \cdots. \end{cases} \tag{7.4.16}$$

Observe that just like the m^{th} order differential operator, the m^{th} order difference of any polynomial of degree $m-1$ or less is identically zero, that is,

$$\Delta^m f = 0, \qquad f \in \pi_{m-1}. \tag{7.4.17}$$

This motivates the definition of the sequence $\{M_m\}_{m=1}^{\infty}$: the function $M_1 \equiv \mathbf{1}_{[0,1)}$, and

$$M_m(x) \equiv \frac{1}{(m-1)!} \Delta^m x_+^{m-1} \qquad \forall\, m \geqslant 2. \tag{7.4.18}$$

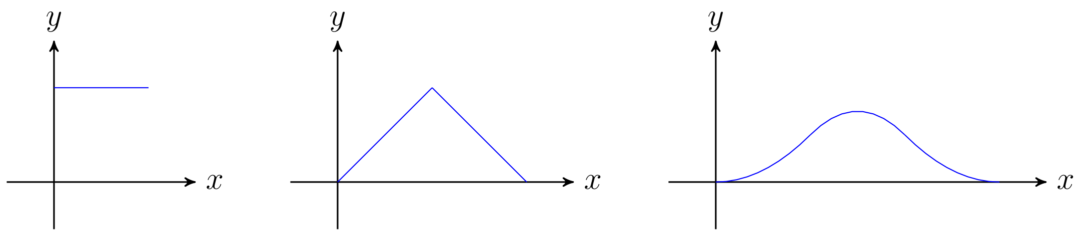

Figure 7.4: The graph of functions M_m, $m = 1, 2, 3$, from left to right.

It is clear from the definition that M_m is a linear combination of the basis functions in (7.4.15). In fact, it is easy to verify by induction that if $\ell \in \mathbb{N}$, then for all $m \in \mathbb{N}$,

$$\Delta^m x_+^{\ell} = \sum_{j=0}^{m} (-1)^j \binom{m}{j} (x-j)_+^{\ell},$$

so we indeed have

$$M_m(x) = \frac{1}{(m-1)!} \sum_{k=0}^{m} (-1)^k \binom{m}{k} (x-k)_+^{m-1}. \tag{7.4.19}$$

From (7.4.17), it follows that $M_m(x) = 0$ for all $x \geqslant m$. Since $M_m(x)$ clearly vanishes for $x < 0$, we have $\operatorname{supp} M_m \subseteq [0, m]$. By working a little harder, we can even conclude that

$$\operatorname{supp}(M_m) = [0, m]. \tag{7.4.20}$$

So, M_m is certainly in $L^2(\mathbb{R})$. However, is the collection

$$\mathcal{B} \equiv \{T_k M_m\}_{k\in\mathbb{Z}} \tag{7.4.21}$$

of integer-translates of M_m a "basis" of S_m? Let us again return to $S_{m,N}$ which, according to (7.4.12) or (7.4.14), has dimension $m + 2N - 1$. Now, by using the support property (7.4.20), each function in the collection

$$\{T_k M_m\}_{k=-N-m+1}^{N-1} \tag{7.4.22}$$

is non-trivial (at least one function in this collection is non-zero) on the interval $[-N, N]$ and $M_m(x-k)$ vanishes identically on $[-N, N]$ for $k < -N - m + 1$ or $k > N - 1$. Since it can be shown that (7.4.22) is a linearly independent set, we have obtained another basis of $S_{m,N}$. So, analogous to (7.4.15), if we take the union of the bases in (7.4.22), $N = 1, 2, \cdots$, we arrive at $\mathcal{B}$ in (7.4.21). One advantage of $\mathcal{B}$ over $\mathcal{T}$ in (7.4.15) is that we can now talk about a spline series

$$f(x) = \sum_{k=-\infty}^{\infty} c_k M_m(x-k) \tag{7.4.23}$$

without worrying too much about convergence. Indeed, for each fixed $x \in \mathbb{R}$, since M_m has compact support, all except for a finite number of terms in the infinite series (7.4.23) are zero.

As mentioned earlier, we are mainly interested in those cardinal splines that belong to $L^2(\mathbb{R})$, namely: $S_m \cap L^2(\mathbb{R})$. Let V_0^m denote its $L^2(\mathbb{R})$-closure. That is, V_0^m is the smallest closed subspace of $L^2(\mathbb{R})$ that contains $S_m \cap L^2(\mathbb{R})$. Since M_m has compact support, we see that $\mathcal{B} \subseteq V_0^m$. We can even show that $\mathcal{B}$ is a Riesz basis of V_0^m, so a scaling function ϕ for V_0^m can be constructed using the orthonormalization process described by Theorem .

So far, we have only considered cardinal splines with knot sequence $\mathbb{Z}$. More generally, we will also consider the spaces S_m^j of cardinal splines with knot sequences $2^{-j}\mathbb{Z}$, $j \in \mathbb{Z}$. Since a spline function with knot sequence $2^{-j_1}\mathbb{Z}$ is also a spline function with knot sequence $2^{-j_2}\mathbb{Z}$ whenever $j_1 < j_2$, we have a (doubly-infinite) nested sequence

$$\cdots \subseteq S_m^{-1} \subseteq S_m^0 \subseteq S_m^1 \subseteq \cdots$$

of cardinal spline spaces, with $S_m^0 = S_m$. Analogous to the definition of V_0^m, we will let V_j^m denote the $L^2(\mathbb{R})$-closure of $S_m^j \cap L^2(\mathbb{R})$. Hence, we have a nested sequence

$$\cdots \subseteq V_{-1}^m \subseteq V_0^m \subseteq V_1^m \subseteq \cdots \tag{7.4.24}$$

of closed cardinal spline subspaces of $L^2(\mathbb{R})$. It will be clear that this nested sequence of subspaces satisfies:

$$\begin{cases} \text{closure}_{\|\cdot\|_2}\left(\bigcup_{j\in\mathbb{Z}} V_j^m\right) = L^2(\mathbb{R}), \\ \bigcap_{j\in\mathbb{Z}} V_j^m = \{0\}. \end{cases} \tag{7.4.25}$$

Furthermore, it is clear that once we have shown that $\mathcal{B}$ is a Riesz basis of V_0^m, then for any $j \in \mathbb{Z}$, the collection

$$\left\{2^{j/2} M_m(2^j x - k) \,\middle|\, k \in \mathbb{Z}\right\} \tag{7.4.26}$$

is also a Riesz basis of V_j^m with the same Riesz bounds as those of $\mathcal{B}$.

The cardinal B-splines (basis splines) consist of functions in $\mathscr{C}^{m-1}(\mathbb{R})$ with equally spaced integer knots that coincide with polynomials of degree n on the intervals $\left[2^{-m}k, 2^{-m}(k+1)\right]$. These B-splines of order n with compact support generate a linear space V_0 in $L^2(\mathbb{R})$. This leads to an MRA $\{V_m \,|\, m \in \mathbb{Z}\}$ by defining $f \in V_m$ if and only if $d_{1/2} f \in V_{m+1}$.

The cardinal B-splines B_m of order m are defined by the following convolution product

$$B_1 = \mathbf{1}_{[0,1]} \qquad \text{and} \qquad B_m = \underbrace{B_1 * B_1 * \cdots * B_1}_{\text{there are } m \ B_1\text{'s}} = B_1 * B_{m-1} \quad n \geqslant 2\,, \tag{7.4.27}$$

where m factors are involved in the convolution product. Obviously,

$$B_m(x) = \int_{\mathbb{R}} B_{n-1}(x-t) B_1(t)\,dt = \int_0^1 B_{m-1}(x-t)\,dt = \int_{x-1}^{x} B_{m-1}(t)\,dt\,. \tag{7.4.28}$$

Using the formula (7.4.28), we can obtain the explicit representation of splines B_2, B_3, and B_4 as follows:

$$B_2(x) = \int_{x-1}^{x} B_1(t)\,dt = \int_{x-1}^{x} \mathbf{1}_{[0,1]}(t)\,dt\,.$$

Evidently, it turns out that

$$\begin{aligned}
B_2(x) &= 0 && \text{if } x \leqslant 0 \text{ or } x \geqslant 2,\\
B_2(x) &= \int_0^x dt = x && \text{if } 0 \leqslant x \leqslant 1,\\
B_2(x) &= \int_{x-1}^{1} dt = 2 - x && \text{if } 1 \leqslant x \leqslant 2.
\end{aligned}$$

Or, equivalently,

$$B_2(x) = x\mathbf{1}_{[0,1]}(x) + (2-x)\mathbf{1}_{[1,2]}(x)\,. \tag{7.4.29}$$

Similarly, using

$$B_3(x) = \int_{x-1}^{x} B_2(t)\,dt\,,$$

we find the explicit expression of B_3:

$$\begin{aligned}
B_3(x) &= 0 && \text{if } x \leqslant 0 \text{ or } x \geqslant 3,\\
B_3(x) &= \int_0^x B_2(t)\,dt = \int_0^x t\,dt = \frac{x^2}{2} && \text{if } 0 \leqslant x \leqslant 1,\\
B_3(x) &= \left(\int_{x-1}^{1} + \int_1^x\right) B_2(t)\,dt = \int_{x-1}^{1} t\,dt + \int_1^x (2-t)\,dt && \text{if } 1 \leqslant x \leqslant 2,\\
&= \frac{1}{2}\left[1-(x-1)^2\right] + 2x - 2 - \frac{1}{2}(x^2-1) = \frac{1}{2}(6x - 2x^2 - 3) &&\\
B_3(x) &= \int_{x-1}^{2} B_2(t)\,dt = \int_{x-1}^{2} (2-t)\,dt = \frac{-1}{2}(2-t)^2\Big|_{t=x-1}^{t=2} = \frac{1}{2}(3-x)^2 && \text{if } 2 \leqslant x \leqslant 3.
\end{aligned}$$

Or, equivalently,

$$B_3(x) = \frac{x^2}{2}\mathbf{1}_{[0,1]}(x) + \frac{1}{2}(6x - 2x^2 - 3)\mathbf{1}_{[1,2]}(x) + \frac{1}{2}(3-x)^2\mathbf{1}_{[2,3]}(x). \tag{7.4.30}$$

Finally, $B_4(x) = \int_{x-1}^{x} B_3(t)\,dt$ we have

$$\begin{aligned}
B_4(x) &= 0 && \text{if } x \leqslant 0 \text{ or } x \geqslant 4,\\
B_4(x) &= \int_0^x B_3(t)\,dt = \frac{1}{6}x^3 && \text{if } 0 \leqslant x \leqslant 1,\\
B_4(x) &= \int_{x-1}^x B_3(t)\,dt = \frac{1}{6}(-3x^3 + 12x^2 - 12x + 4) && \text{if } 1 \leqslant x \leqslant 2,\\
B_4(x) &= \int_{x-1}^x B_3(t)\,dt = \frac{1}{6}(3x^3 - 24x^2 + 60x - 44) && \text{if } 2 \leqslant x \leqslant 3,\\
B_4(x) &= \int_{x-1}^3 B_3(t)\,dt = \frac{1}{6}(4-x)^3 && \text{if } 3 \leqslant x \leqslant 4.
\end{aligned}$$

Or, equivalently,

$$\begin{aligned}
B_4(x) = &\frac{1}{6}x^3\mathbf{1}_{[0,1)}(x) + \frac{1}{6}(-3x^3 + 12x^2 - 12x + 4)\mathbf{1}_{[1,2)}(x)\\
&+ \frac{1}{6}(3x^3 - 24x^2 + 60x - 44)\mathbf{1}_{[2,3)}(x) + \frac{1}{6}(4-x)^3\mathbf{1}_{[3,4]}(x).
\end{aligned} \tag{7.4.31}$$

In general, we have the following

Theorem 7.24. *The m^{th} order cardinal B-spline B_m satisfies the following properties:*

(i) *For every $f \in \mathscr{C}(\mathbb{R})$,*

$$\int_{\mathbb{R}} f(x)B_m(x)\,dx = \int_0^1 \cdots \int_0^1 f(x_1 + x_2 + \cdots + x_m)\,dx_1 \cdots dx_m. \tag{7.4.32}$$

(ii) *For every $g \in \mathscr{C}^m(\mathbb{R})$,*

$$\int_{\mathbb{R}} g^{(m)}(x)B_m(x)\,dx = \sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}g(k). \tag{7.4.33}$$

(iii) *$B_m(x) = M_m(x)$ for all $x \in \mathbb{R}$; that is,*

$$B_m(x) = \frac{1}{(m-1)!}\sum_{k=0}^{m}(-1)^k\binom{m}{k}(x-k)_+^{m-1}.$$

(iv) $\operatorname{supp}(B_m) = [0, m]$.

(v) *$B_m(x) > 0$ for all $x \in (0, m)$.*

(vi) *$\sum_{k=-\infty}^{\infty} B_m(x-k) = 1$ for all $x \in \mathbb{R}$.*

(vii) $B_m'(x) = (\Delta B_{m-1})(x) \equiv B_{m-1}(x) - B_{m-1}(x-1)$ *for all* $x \in \mathbb{R}$.

(viii) *The cardinal B-splines* B_m *and* B_{m-1} *are related by the identity:*

$$B_m(x) = \frac{x}{m-1} B_{m-1}(x) + \frac{m-x}{m-1} B_{m-1}(x-1) \qquad \forall\, x \in \mathbb{R}. \tag{7.4.34}$$

(ix) B_m *is symmetric with respect to the center of its support, namely:*

$$B_m\Big(\frac{m}{2} + x\Big) = B_m\Big(\frac{m}{2} - x\Big) \qquad \forall\, x \in \mathbb{R}.$$

Proof. (i) Assertion (7.4.32) certainly holds for $m = 1$. Suppose it also holds for $m - 1$ (for some $m \geqslant 2$), then by the definition of B_m in (7.4.28) and this induction hypothesis, we have

$$\begin{aligned}
\int_{\mathbb{R}} f(t) B_m(t)\, dt &= \int_{\mathbb{R}} f(t) \int_0^1 B_{m-1}(t - x_m)\, dx_m dt = \int_0^1 \int_{\mathbb{R}} f(t) B_{m-1}(t - x_m)\, dt dx_m \\
&= \int_0^1 \int_{\mathbb{R}} f(t + x_m) B_{m-1}(t)\, dt\, dx_m \\
&= \int_0^1 \Big(\int_0^1 \cdots \int_0^1 f(x_1 + \cdots x_{m-1} + x_m)\, dx_1 \cdots dx_{m-1} \Big)\, dx_m .
\end{aligned}$$

It follows from induction that (7.4.32) holds for all m.

(ii) Assertion (7.4.33) holds for $m = 1$. Suppose that it also holds for $m - 1$ (for some $m \geqslant 2$). Then by part (i) we obtain that

$$\begin{aligned}
\int_{\mathbb{R}} g^{(m)}(t) B_m(t)\, dt &= \int_0^1 \cdots \int_0^1 g^{(m)}(x_1 + \cdots + x_m)\, dx_m \cdots dx_1 \\
&= \int_0^1 \cdots \int_0^1 g^{(m-1)}(x_1 + \cdots + x_{m-1} + 1)\, dx_{m-1} \cdots dx_1 \\
&\quad - \int_0^1 \cdots \int_0^1 g^{(m-1)}(x_1 + \cdots + x_{m-1})\, dx_{m-1} \cdots dx_1 \\
&= \int_{\mathbb{R}} g^{(m-1)}(t+1) B_{m-1}(t)\, dt - \int_{\mathbb{R}} g^{(m-1)}(t) B_{m-1}(t)\, dt \\
&= \sum_{k=0}^{m-1} (-1)^{m-1-k} \binom{m-1}{k} g(k+1) - \sum_{k=0}^{m-1} (-1)^{m-1-k} \binom{m-1}{k} g(k).
\end{aligned}$$

Since

$$\sum_{k=0}^{m-1} (-1)^{m-1-k} \binom{m-1}{k} g(k+1) = \sum_{k=1}^{m} (-1)^{m-k} \binom{m-1}{k-1} g(k),$$

we find that

$$\begin{aligned}
&\sum_{k=0}^{m-1}(-1)^{m-1-k}\binom{m-1}{k}g(k+1)-\sum_{k=0}^{m-1}(-1)^{m-1-k}\binom{m-1}{k}g(k)\\
&\quad=\sum_{k=1}^{m}(-1)^{m-k}\binom{m-1}{k-1}g(k)-\sum_{k=0}^{m-1}(-1)^{m-1-k}\binom{m-1}{k}g(k)\\
&\quad=g(m)+\sum_{k=1}^{m-1}(-1)^{m-k}\binom{m-1}{k-1}g(k)+(-1)^m g(0)+\sum_{k=1}^{m-1}(-1)^{m-k}\binom{m-1}{k}g(k)\\
&\quad=g(m)+(-1)^m g(0)+\sum_{k=1}^{m-1}(-1)^{m-k}\left[\binom{m-1}{k-1}+\binom{m-1}{k}\right]g(k)\\
&\quad=g(m)+(-1)^m g(0)+\sum_{k=1}^{m-1}(-1)^{m-k}\binom{m}{k}g(k)=\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}g(k).
\end{aligned}$$

It follows from induction that (7.4.33) holds for all $m \in \mathbb{N}$.

(iii) Clearly $B_1(x) = M_1(x) = \mathbf{1}_{[0,1)}(x)$ for all $x \in \mathbb{R}$. Assume that $B_{m-1}(x) = M_{m-1}(x)$ for all $x \in \mathbb{R}$ (for some $m \geqslant 2$). Then

$$\begin{aligned}
B_m(x) &= \int_0^1 B_{m-1}(x-t)\,dt = \int_0^1 M_{m-1}(x-t)\,dt\\
&= \frac{1}{(m-1)!}\sum_{k=0}^{m-1}(-1)^k\binom{m-1}{k}\int_0^1 (x-t-k)_+^{m-1}\,dt.
\end{aligned}$$

Since

$$\int_0^1 (x-t-k)_+^{m-1}\,dt = \begin{cases} \displaystyle\int_0^1 (x-t-k)^{m-1}\,dt & \text{if } x-k \geqslant 1,\\ \displaystyle\int_0^{x-k} (x-t-k)^{m-1}\,dt & \text{if } 0 \leqslant x-k < 1,\\ 0 & \text{if } x-k < 0,\end{cases}$$

we have

$$\int_0^1 (x-t-k)_+^{m-1}\,dt = \frac{1}{m}\big[(x-k)_+^m - (x-k-1)_+^m\big].$$

Therefore,

$$\begin{aligned}
B_m(x) &= \frac{1}{(m-1)!}\sum_{k=0}^{m-1}(-1)^k\binom{m-1}{k}\int_0^1 (x-t-k)_+^{m-1}\,dt\\
&= \frac{1}{(m-1)!}\sum_{k=0}^{m-1}(-1)^k\binom{m-1}{k}\frac{1}{m}\big[(x-k)_+^m - (x-k-1)_+^m\big]\\
&= \frac{1}{m!}\left[\sum_{k=0}^{m-1}(-1)^k\binom{m-1}{k}(x-k)_+^m + \sum_{k=1}^{m}(-1)^k\binom{m-1}{k-1}(x-k)_+^m\right]\\
&= \frac{1}{m!}\sum_{k=1}^{m}(-1)^k\binom{m}{k}(x-k)_+^m = M_m(x).
\end{aligned}$$

By induction, $B_m(x) = M_m(x)$ for all $x \in \mathbb{R}$ and $m \in \mathbb{N}$.

(vii) Using (7.4.28) again, the Fundamental Theorem of Calculus shows that

$$B'_m(x) = \frac{d}{dx}\int_{x-1}^{x} B_{m-1}(t)\,dt = -B_{m-1}(x-1) + B_{m-1}(x) = (\Delta B_{m-1})(x).$$

(viii) To verify the identity in (viii), we use the definition of M_m in (7.4.18) instead. Of course, we have already shown in (iii) that $B_m = M_m$. The idea is to represent as the product of a monomial and a truncated power, namely:

$$x_+^{m-1} = x \cdot x_+^{m-2}$$

and then apply the following "Leibniz Rule" for differences:

$$\big[\Delta^n(fg)\big](x) = \sum_{k=0}^{n}\binom{n}{k}(\Delta^k f)(x)(\Delta^{n-k}g)(x-k). \tag{7.4.35}$$

This identity for differences can be easily established by induction. It is almost exactly the same as the Leibniz Rule for derivatives. Now, if we set $f(x) = x$ and $g(x) = x_+^{m-2}$ in (7.4.35) and recall that $\Delta^k f = 0$ for $k \geqslant 2$ from (7.4.17), we then have

$$\begin{aligned}
B_m(x) = M_m(x) &= \frac{1}{(m-1)!}\Delta^m x_+^{m-1} = \frac{1}{(m-1)!}\Big[x\Delta^m x_+^{m-2} + m\Delta^{m-1}(x-1)_+^{m-2}\Big] \\
&= \frac{1}{(m-1)!}\Big[x\big(\Delta^{m-1}x_+^{m-2} - \Delta^{m-1}(x-1)_+^{m-2}\big) + m\Delta^{m-1}(x-1)_+^{m-2}\Big] \\
&= \frac{x}{m-1}B_{m-1}(x) + \frac{m-x}{m-1}B_{m-1}(x-1)
\end{aligned}$$

Assertions (iv), (v), (vi) and (ix) can also be easily derived by induction, using the definition of B_m in (7.4.28). This completes the proof of Theorem 7.24. □

In order to obtain the two-scale relation for the B-splines of order n, we apply the Fourier transform of (7.4.27) so that

$$\widehat{B_1}(\omega) = \exp\Big(-\frac{i\omega}{2}\Big)\operatorname{sinc}\Big(\frac{\omega}{2}\Big). \tag{7.4.36}$$

Using (7.4.5) and (7.4.6a) we can also express (7.4.36) in terms of $z = \exp\big(-\frac{i\omega}{2}\big)$ as

$$\widehat{B_1}(\omega) = \frac{1}{2}(1+z)\widehat{B_1}\Big(\frac{\omega}{2}\Big). \tag{7.4.37}$$

Application of the convolution theorem of the Fourier transform to (7.4.27) gives

$$\begin{aligned}
\widehat{B_n}(\omega) &= \big[\widehat{B_1}(\omega)\big]^n = \widehat{B_1}(\omega)\widehat{B_{n-1}}(\omega) = \Big(\frac{1+z}{2}\Big)^n\Big[\widehat{B_1}\Big(\frac{\omega}{2}\Big)\Big]^n = \Big(\frac{1+z}{2}\Big)^n\widehat{B_n}\Big(\frac{\omega}{2}\Big) \\
&= \widehat{M_n}\Big(\frac{\omega}{2}\Big)\widehat{B_n}\Big(\frac{\omega}{2}\Big),
\end{aligned}$$

where the associated filter $\widehat{M_n}$ is given by

$$\begin{aligned}
\widehat{M_n}\Big(\frac{\omega}{2}\Big) &= \Big(\frac{1+z}{2}\Big)^n = \frac{1}{2^n}\big(1+e^{-\frac{i\omega}{2}}\big)^n = \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\exp\Big(-\frac{ik\omega}{2}\Big) \\
&\equiv \frac{1}{\sqrt{2}}\sum_{k=-\infty}^{\infty} c_{n,k}\exp\Big(-\frac{ik\omega}{2}\Big),
\end{aligned} \tag{7.4.38}$$

where the coefficients $c_{n,k}$ are given by

$$c_{n,k} = \begin{cases} \dfrac{\sqrt{2}}{2^n}\dbinom{n}{k} & \text{if } 0 \leqslant k \leqslant n, \\ 0 & \text{otherwise.} \end{cases} \tag{7.4.39}$$

Therefore, the spline function in the time domain is

$$B_n(x) = \sqrt{2}\sum_{k=-\infty}^{\infty} c_{n,k} B_n(2x-k) = \sum_{k=0}^{n} 2^{1-n}\binom{n}{k} B_n(2x-k). \tag{7.4.40}$$

This may be referred to as the two-scale relation for the B-splines of order n.

We next show that the cardinal B-spline basis

$$\mathcal{B} = \{T_k B_m \,|\, k \in \mathbb{Z}\} \tag{7.4.41}$$

is a Riesz (or unconditional) basis of V_0^m in the sense of Definition 6.26 (or Theorem 6.27). By Theorem 7.7, this is equivalent to investigating the existence of lower and upper bounds A, B in (7.3.4). From (7.4.27), we see that so that $\widehat{B_m} = \widehat{B_1}^m$, so that

$$\left|\widehat{B_m}(\omega)\right|^2 = \left|\frac{1-e^{-i\omega}}{i\omega}\right|^{2m} = \left|\frac{\sin(\omega/2)}{\omega/2}\right|^{2m}.$$

Hence, replacing ω by 2ω, we have

$$\sum_{k=-\infty}^{\infty} \left|\widehat{B_m}(2\omega+2k\pi)\right|^2 = \sum_{k=-\infty}^{\infty} \frac{\sin^{2m}(\omega+k\pi)}{(\omega+k\pi)^{2m}} = \sin^{2m}\omega \sum_{k=-\infty}^{\infty} \frac{1}{(\omega+k\pi)^{2m}}. \tag{7.4.42}$$

By the residue theorem, for $\omega \neq k\pi$ for all $k \in \mathbb{Z}$,

$$\lim_{N\to\infty} \oint_{C_N} \frac{\pi\cot(\pi z)}{(\omega+z\pi)^2}\,dz = 2\pi i\left[\sum_{k=-\infty}^{\infty} \frac{1}{(\omega+k\pi)^2} + \frac{d}{dz}\Big|_{z=-\pi/\omega} \frac{\cot(\pi z)}{\pi}\right],$$

where C_N is the square contour with four corners $\big(\pm(N+0.5), \pm(N+0.5)\big)$. The choice of C_N leads to that

$$\lim_{N\to\infty} \oint_{C_N} \frac{\pi\cot(\pi z)}{(\omega+z\pi)^2}\,dz = 0,$$

so we conclude that

$$\sum_{k=-\infty}^{\infty} \frac{1}{(\omega+k\pi)^2} = -\frac{d}{dz}\Big|_{z=-\pi/\omega} \frac{\cot(\pi z)}{\pi} = \csc^2\omega \qquad \text{whenever} \quad \omega \notin \{k\pi \,|\, k \in \mathbb{Z}\}.$$

Taking another $2m-2$ derivative w.r.t. ω yields that

$$\sum_{k=-\infty}^{\infty} \frac{1}{(\omega+k\pi)^{2m}} = \frac{1}{(2m-1)!}\frac{d^{2m-2}}{dx^{2m-2}}\csc^2\omega. \tag{7.4.43}$$

Therefore, substituting (7.4.43) into (7.4.42), we obtain

$$\sum_{k=-\infty}^{\infty} \left|\widehat{B_m}(2\omega+2k\pi)\right|^2 = \frac{\sin^{2m}\omega}{(2m-1)!}\frac{d^{2m-2}}{dx^{2m-2}}\csc^2\omega. \tag{7.4.44}$$

This shows that the first order B-spline B_1 defined by (7.4.27) is a scaling function that generates the classic Haar wavelet.