FOURIER ANALYSIS–PART 2

XU WANG

This is the second part of the Fourier analysis notes for TMA4170.

CONTENTS

1. MORE FOURIER ANALYSIS

1.1. **Eigentheory of the Fourier transform.** The starting point is the following:

$$\widehat{f'} = i \cdot 2\pi x \hat{f}, \quad \widehat{2\pi x f} = i \cdot (\hat{f})'.$$

Consider the following operators A, B on the Schwartz space defined by

$$Af := f', \quad Bf := 2\pi x f.$$

Then

$$\widehat{Af} = iB\hat{f}, \quad \widehat{Bf} = iA\hat{f}.$$

Think of the Fourier transform as an operator

$$\mathcal{F}(f) := \hat{f}.$$

Then we have

$$\mathcal{F}((A+B)f) = i \cdot (A+B)\mathcal{F}(f),$$

and

$$\mathcal{F}((A-B)f) = -i \cdot (A-B)\mathcal{F}(f).$$

Now suppose that f is a fixed point of the Fourier transform, i.e. $\mathcal{F}(f) = f$, then

$$\mathcal{F}((A-B)f) = -i \cdot (A-B)\mathcal{F}(f) = -i \cdot (A-B)f.$$

Date: March 30, 2019.

Let us inductively define

$$(A-B)^k f := (A-B)((A-B)^{k-1}f), \ \ (A-B)^1 f := (A-B)f.$$

Then we have

$$\mathcal{F}((A-B)^k f) = (-i)^k (A-B)^k f, \ \ \mathcal{F}((A+B)^k f) = i^k (A+B)^k f.$$

Now suppose that $f = e^{-\pi x^2}$, a direct computation gives

$$(A+B)f = 0, \quad (A-B)f = -4\pi x f = -2Bf.$$

Definition 1. *We call*

$$h_k := (A-B)^k e^{-\pi x^2}, \ \ k = 0, 1, \cdots$$

the k-th Hermite function associated to the Fourier transform $\mathcal{F}$.

Remark 1. *We have*

$$h_0(x) = e^{-\pi x^2}, \ h_1(x) = -4\pi x e^{-\pi x^2}, \ \ h_2(x) = ((-4\pi x)^2 - 4\pi)e^{-\pi x^2},$$

In general, we can write

$$h_k(x) = P_k(x) \cdot e^{-\pi x^2},$$

where

$$P_k(x) = (-4\pi x)^k + \textit{lower order terms}.$$

We call the above degree k polynomial the k-th Hermite polynomial. From the definition, we have

$$(h_k, h_l) = \int_{\mathbb{R}} P_k(x) P_l(x) \, e^{-2\pi x^2} \, dx.$$

The right hand side is called the Segal–Bargmann–Fock inner product of P_k and P_l.

The main theorem in this section is the following:

Theorem 1. *$\{h_k\}_{k\geq 0}$ defines an orthogonal basis of the Schwartz space $\mathcal{S}$, moreover, we have*

$$\mathcal{F}(h_k) = (-i)^k h_k, \ \ k = 0, 1, 2, \cdots.$$

Since $A - B$ is an eigen-operator of $\mathcal{F}$, we have $\mathcal{F}(h_k) = (-i)^k h_k$. Now let us prove that $\{h_k\}_{k\geq 0}$ is an orthogonal system. It is enough to find an operator $\mathfrak{f}$ such that

$$\mathfrak{f}(h_k) = \lambda_k h_k$$

and $\lambda_k \neq \lambda_j$ if $j \neq k$. The main idea is to find a *fixed operator* of the Fourier transform. Since

$$\mathcal{F}((A-B)(A+B)f) = i(-i)(A-B)(A+B)\mathcal{F}(f) = (A-B)(A+B)\mathcal{F}(f),$$

we know that $(A-B)(A+B)$ is a fixed operator of $\mathcal{F}$. A direct computation gives

$$(A-B)(A+B) = 4\pi\mathfrak{f} + 2\pi = (A+B)(A-B) + 4\pi,$$

where

$$\mathfrak{f}(g) := \frac{g''}{4\pi} - \pi x^2 g.$$

We know that $\mathfrak{f}$ is also a fixed operator of $\mathcal{F}$. Now $(A+B)h_0 = 0$ gives

$$0 = (A-B)(A+B)h_0 = 4\pi\mathfrak{f}(h_0) + 2\pi h_0.$$

Thus

$$\mathfrak{f}(h_0) = -\frac{1}{2}h_0.$$

Assume that

$$\mathfrak{f}(h_k) = -\left(k + \frac{1}{2}\right) h_k.$$

Then $(A-B)(A+B)h_{k+1} = (A-B)(A+B)(A-B)h_k$ can be written as

$$(A-B)(4\pi\mathfrak{f} - 2\pi)h_k = 4\pi(-k-1)h_{k+1}.$$

Thus we have

$$(4\pi\mathfrak{f} + 2\pi)h_{k+1} = 4\pi(-k-1)h_{k+1},$$

which gives

$$\mathfrak{f}(h_{k+1}) = -\left(k + 1 + \frac{1}{2}\right) h_{k+1}.$$

Thus we have

Lemma 1. *Each h_k is an eigenfunction of $\mathfrak{f}$ with eigenvalue $-k-\frac{1}{2}$, in particular, it implies that $\{h_k\}$ is an orthogonal system.*

Now it suffices to show that $\{h_k\}$ is a basis of the Schwartz space $\mathcal{S}$ in order to prove our main theorem. The idea is consider the *generating function* of $\{h_k\}$ defined by

$$H(t,x) := \sum_{k=0}^{\infty} h_k(x)\frac{t^k}{k!}.$$

Lemma 2. *$H(t,x) = e^{-\pi(x+2t)^2} \cdot e^{2\pi t^2}$.*

Proof. Notice that

$$(A-B)f = e^{\pi x^2}\frac{d}{dx}(e^{-\pi x^2} f),$$

gives

$$(A-B)^k f = e^{\pi x^2}\left(\frac{d}{dx}\right)^k (e^{-\pi x^2} f).$$

Thus we have

$$H(t,x) = e^{\pi x^2}\sum_{k=0}^{\infty}\left(\frac{d}{dx}\right)^k e^{-2\pi x^2}\frac{t^k}{k!}.$$

Notice that

$$\left(\frac{d}{dx}\right)^k e^{-2\pi x^2} = \left(\left(\frac{d}{dt}\right)^k e^{-2\pi(x+t)^2}\right)\Big|_{t=0}.$$

Thus the taylor expansion of $e^{-2\pi(x+t)^2}$ gives

$$H(t,x) = e^{\pi x^2 - 2\pi(x+t)^2} = e^{-\pi(x+2t)^2} \cdot e^{2\pi t^2}.$$

□

Now we can prove the following lemma.

Lemma 3. *$\{h_k\}$ is a basis of the Schwartz space $\mathcal{S}$,*

Proof. Assume that $f \in \mathcal{S}$ and $(f, h_k) = 0$ for $k = 0, 1, \cdots$. Then we have $(f, H(t, \cdot)) = 0$ for every $t \in \mathbb{R}$, which implies that $f \star e^{-\pi x^2} \equiv 0$. Taking the Fourier transform, we get $\hat{f}(y)e^{-\pi y^2} \equiv 0$. Thus $\hat{f} = 0$ and we know that $f = \check{\hat{f}} = 0$. □

Remark 2 (**). *Up to a constant, $\mathfrak{f}$ is equal to the* Hermite operator *(see Week 8 exercise). Since*

$$\mathcal{F}(h_k) = (-i)^k h_k, \quad \mathfrak{f}(h_k) = -\left(k + \frac{1}{2}\right) h_k,$$

and

$$e^{\frac{i\pi}{2}\left(-k-\frac{1}{2}\right)} = (-i)^k e^{-\frac{i\pi}{4}}.$$

We get the following crucial identity

$$e^{\frac{i\pi}{2}\mathfrak{f}} = \mathcal{F} \cdot e^{-\frac{i\pi}{4}}.$$

Recall that we have

$$\mathfrak{f} = L + \Lambda, \quad L := -\pi x^2, \quad \Lambda := \left(\frac{d}{dx}\right)^2 / 4\pi.$$

Put $N := [L, \Lambda] = x\frac{d}{dx} + \frac{1}{2}$, then we have the following sl_2-identities

$$[N, L] = 2L, \quad [N, \Lambda] = -2\Lambda, \quad [L, \Lambda] = N.$$

Usually we call

$$e^{\frac{i\pi}{2}(L+\Lambda)} = e^{\frac{i\pi}{2}\mathfrak{f}}$$

the sl_2-star operator (up to a constant, the famous Hodge star operator is a special sl_2-star operator). Thus the Fourier transform is essentially the star operator in sl_2-representation.

1.2. **Central limit theorem.** From page 114–116 in [3]. The mathematical content of the "central limit theorem" of probability theory is the following

Theorem 2. *Let f be a non-negative function in $\mathcal{S}$ with $\int f = 1$, $\int xf = 0$ and $\int x^2 f = 1$, denote by f_n the n-fold convolution defined inductively by*

$$f_1 := f, \quad f_2 := f * f, \quad f_{n+1} := f * f_n.$$

Then $\sqrt{n} f_n(\sqrt{n}x)$ converges weakly to $(2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$.

Proof. Since $\hat{\mathcal{S}} = \mathcal{S}$, it is enough to check that

$$\lim_{n\to\infty} \int_{\mathbb{R}} \sqrt{n} f_n(\sqrt{n}x)\hat{g}(x)\, dx = \int_{\mathbb{R}} (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} \hat{g}(x)\, dx, \qquad \forall\, g \in \mathcal{S}.$$

Notice that

$$\widehat{\sqrt{n} f_n(\sqrt{n}x)} = \left(\hat{f}(\gamma/\sqrt{n})\right)^n$$

gives

$$\int_{\mathbb{R}} \sqrt{n} f_n(\sqrt{n}x)\hat{g}(x)\, dx = \int_{\mathbb{R}} \left(\hat{f}(\gamma/\sqrt{n})\right)^n g(\gamma)\, d\gamma.$$

By definition, we have

$$\hat{f}(\gamma/\sqrt{n}) = \int_{\mathbb{R}} e^{-2\pi i \gamma x/\sqrt{n}} f(x)\, dx.$$

Since $f \in \mathcal{S}$, we know $c_l := \sup_{x\in\mathbb{R}} |f(x)x^{2l+3}| < \infty$ and

$$\Big|\int_{|x|\geq n^{1/l}} e^{-2\pi i \gamma x/\sqrt{n}} f(x)\, dx\Big| \leq c_l \int_{|x|\geq n^{1/l}} x^{-2l-1}\, dx = c_l/(ln^2).$$

When $|x| \leq n^{1/l}$ and $l > 6$, then $d := 3(1/2 - 1/l) > 1$ and Taylor expansion of e^x gives

$$\left|e^{-2\pi i \gamma x/\sqrt{n}} - 1 + 2\pi i \gamma x/\sqrt{n} + 2\pi^2\gamma^2 x^2/n\right| \leq \sum_{k\geq 3} (2\pi\gamma)^k/(n^{(k/2-k/l)}k!) \leq n^{-d} e^{2\pi|\gamma|}.$$

Thus

$$\left|\hat{f}(\gamma/\sqrt{n}) - \int_{|x|\leq n^{1/l}} (1 - 2\pi i\gamma x/\sqrt{n} - 2\pi^2\gamma^2x^2/n) f(x)\, dx\right| \leq c_l/(ln^2) + n^{-d}e^{2\pi|\gamma|}.$$

Moreover, we have

$$\Big|\int_{|x|\geq n^{1/l}} (1 - 2\pi i\gamma x/\sqrt{n} - 2\pi^2\gamma^2x^2/n) f(x)\, dx\Big| \leq c_l(1 + 2\pi\gamma/\sqrt{n} + 2\pi^2\gamma^2/n)/(ln^2).$$

Now we have

$$\left|\hat{f}(\gamma/\sqrt{n}) - \int_{\mathbb{R}} (1 - 2\pi i\gamma x/\sqrt{n} - 2\pi^2\gamma^2x^2/n) f(x)\, dx\right| \leq c(\gamma)n^{-d},\ d > 1.$$

But our assumption gives

$$\int_{\mathbb{R}} (1 - 2\pi i\gamma x/\sqrt{n} - 2\pi^2\gamma^2x^2/n) f(x)\, dx = 1 - 2\pi^2\gamma^2/n.$$

Thus we get

$$\left|\hat{f}(\gamma/\sqrt{n}) - (1 - 2\pi^2\gamma^2/n)\right| \leq c(\gamma)n^{-d},$$

which gives

$$\lim_{n\to\infty} \left(\hat{f}(\gamma/\sqrt{n})\right)^n = \lim_{n\to\infty} (1 - 2\pi^2\gamma^2/n)^n = e^{-2\pi^2\gamma^2}.$$

Notice that $|\hat{f}(\gamma/\sqrt{n})| \leq \int f = 1$, the above formula gives (by the dominated convergence theorem, see wikipedia page)

$$\lim_{n\to\infty} \int_{\mathbb{R}} \left(\hat{f}(\gamma/\sqrt{n})\right)^n g(\gamma)\, d\gamma = \int_{\mathbb{R}} e^{-2\pi^2\gamma^2} g(\gamma)\, d\gamma.$$

Now it is enough to show

$$\int_{\mathbb{R}} (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} \hat{g}(x)\, dx = \int_{\mathbb{R}} e^{-2\pi^2\gamma^2} g(\gamma)\, d\gamma,$$

which follows from

$$\widehat{(2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}} = e^{-2\pi^2\gamma^2}.$$

The proof is complete. □

Remark 3. *By approximating the indicator function of $[a, b]$ above and below by functions in $\mathcal{D}$, the above theorem implies that*

$$\lim_{n\to\infty} \int_a^b \sqrt{n} f_n(\sqrt{n}x)\, dx = \int_{a\sqrt{n}}^{b\sqrt{n}} f_n(x)\, dx = \int_a^b \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{x^2}{2}}\, dx$$

for any $a < b$.

The probabilistic content of the central limit theorem is as follows: Think of an infinite number of statistically independent copies $u_1, u_2, \cdots, u_n, \cdots$ of a statistical quantity u distributed according to the rule

$$P(a \leq u < b) = \int_a^b f(x)\, dx,$$

in which the letter P stands for the probability of the indicated event. The adjective "independent" means that probabilities multiply:

$$P(a_j \leq u_j < b_j, j = 1, 2, \cdots) = \int_{a_1}^{b_1} f(x)\, dx \cdot \int_{a_2}^{b_2} f(x)\, dx \cdots$$

and you infer that the sum $s_n = u_1 + \cdot + u_n$ is a distribution according to the rule

$$P(a \leq s_n < b) = \int_{a \leq x_1 + \cdots + x_n < b} f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n = \int_a^b f_n(x)\, dx.$$

The content of the central limit theorem is now seen to be that the scale sum $s_n/\sqrt{n}$ is nearly Gaussian distributed for large n:

$$P(a \leq s_n/\sqrt{n} < b) = \sqrt{n} \int_a^b f_n(\sqrt{n}x)\, dx$$

is approximately

$$\int_a^b (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}\, dx.$$

The fact goes back to Moivre and Laplace in the 18th century; for additional information, see [4] (Vol. 1, pp. 174–195, Vol. 2, pp, 252–259).

2. Wavelet analysis

This part is a list of results from Chapter 4-7 in [1].

2.1. Haar wavelets.

Definition 2 (Haar scaling function). $\phi(x) = 1$ *if* $0 \le x < 1$*;* $\phi(x) = 0$ *otherwise.*

Definition 3. *The space of step functions at level* j*,* $j = 0, 1, \cdots$*, is defined to the* $\mathbb{R}$*-linear space spanned by* $\{\phi(2^j x + k)\}_{k\in\mathbb{Z}}$.

Fact 1: $V_0 \subset V_1 \subset \cdots V_j \subset V_{j+1} \subset \cdots$;
Fact 2: $f(x) \in V_0$ iff $f(2^j x) \in V_j$, $f(x) \in V_j$ iff $f(2^{-j}x) \in V_0$;
Fact 3: $\{2^{j/2}\phi(2^j x + k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_j.

The so called Haar wavelet is a function that can be used to express the *orthogonal complement* of V_j in V_{j+1}.

Definition 4. *The Haar wavelet is the function* $\psi(x) := \phi(2x) - \phi(2x - 1)$.

Fact 4: $V_{j+1} = V_j \oplus W_j$, where W_j is spanned by $\{\psi(2^j x + k)\}_{k\in\mathbb{Z}}$;
Fact 5: $L^2(\mathbb{R}) = \overline{V_0 \oplus W_0 \oplus W_1 \oplus \cdots}$.

Remark: The so called Haar decomposition and reconstruction algorithms are essentially based on the following orthogonal decomposition

$$V_j = V_0 \oplus W_1 \oplus \cdots W_{j-1}.$$

Step 1: Sample Approximated the original signal by a step function of the following form

$$f_j(x) = \sum a_k \phi(2^j x - k) \in V_j;$$

Step 2: Decompose Use $\phi(2x) = (\phi(x) + \psi(x))/2$, $\phi(2x - 1) = (\phi(x) - \psi(x))/2$ with x replaced by $2^{j-1}x - k$ to write

$$f_j = f_{j-1} + w_{j-1}, \;\; f_{j-1} \in V_{j-1}, \; w_{j-1} \in W_{j-1};$$

Repeating the process gives

$$f_j = f_0 + w_0 + \cdots + w_{j-1}, \;\; f_0 \in V_0, \; w_k \in W_k;$$

Step 3: Surgery Remove unwanted w_k (filter out noise or data compression);
Step 4: Recover: Use $\phi(x) = \phi(2x) + \phi(2x - 1)$, $\psi(x) = \phi(2x) - \phi(2x - 1)$ with x replaced by $2^{j-1}x - k$ to inductively write $f_0 + w_0 + \cdots + w_{j-1}$ as $\sum a_k \phi(2^k x - k)$.

2.2. **Multiresolution analysis.** The Haar wavelet theory is based on the Haar scaling function ϕ. One drawback is that ϕ is not continuous. The wavelet theory based on a general scaling function (can be continuous) is called a multiresolution analysis.

Definition 5 (Multiresolution analysis). *A sequence of linear spaces $\{V_j\}_{j\in\mathbb{Z}}$ in $L^2(\mathbb{R})$ is called a multiresolution analysis with scaling function ϕ if*

1) $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ *is an orthonormal basis of* V_0*;*
2) $V_j = \{f : f(2^{-j}x) \in V_0\}$, $j\in\mathbb{Z}$*;*
3) $V_j \subset V_{j+1}$*;*
4) $\cap_{j\in\mathbb{Z}} V_j = \{0\}$*;*
4) $\overline{\cup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$.

It is clear that the Haar scaling function defines a multiresolution analysis. Another interesting scaling function is the Shannon scaling function $\phi(x) = \mathrm{sinc}(x)$, where

$$\mathrm{sinc}(x) := \frac{\sin \pi x}{\pi x}$$

on $\mathbb{R}$ (with the understanding that $\mathrm{sinc}(0) = 1$). Notice that

$$\mathrm{sinc} = \widehat{1_{[-1/2,1/2]}}$$

Thus the Shannon scaling function is equal to the Fourier transform of the Haar scaling function up to a translation by $1/2$.

By a linear change of variable, assumption 1) in the definition of multiresolution analysis gives

Theorem 3. *For every $j\in\mathbb{Z}$, $\{\phi_k^j := 2^{j/2}\phi(2^jx-k)\}_{k\in\mathbb{Z}}$ defines an orthonormal basis for V_j.*

The orthogonal decomposition gives the following

$$\phi_l^{j-1} = \sum_{k\in\mathbb{Z}} (\phi_l^{j-1}, \phi_k^j)\phi_k^j.$$

Notice that (replace $2^{j-1}x$ by x)

$$(\phi_l^{j-1}, \phi_k^j) = 2^{-1/2}2^j \int_{\mathbb{R}} \phi(2^{j-1}x-l)\overline{\phi(2^jx-k)}\, dx = 2^{1/2}\int_{\mathbb{R}} \phi(x-l)\overline{\phi(2x-k)}\, dx.$$

Replace $x-l$ by x, we get

$$(\phi_l^{j-1}, \phi_k^j) = 2^{1/2}\int_{\mathbb{R}} \phi(x)\overline{\phi(2x+2l-k)}\, dx = (\phi, \phi_{k-2l}^1) := p_{k-2l},$$

which gives the following theorem:

Theorem 4 (Scaling relation theorem).

$$\phi_l^{j-1} = \sum_{k\in\mathbb{Z}} p_{k-2l}\, \phi_k^j.$$

Example: For the Haar scaling function ϕ, we have

$$p_0 = p_1 = 2^{-1/2}, \ \ p_k = 0, \ \forall\, k \neq 0, 1.$$

The following results contains identities for the p_k which will be important later.

Theorem 5. *We have the following formula*

$$\sum_{k\in\mathbb{Z}} |p_k|^2 = 1$$

and

$$\sum_{k\in\mathbb{Z}} p_{k-2l}\overline{p_k} = 0, \quad \forall\, l \neq 0.$$

Assume further that ϕ has compact support and $\int_{\mathbb{R}} \phi(x)\,dx \neq 0$. Then

$$\sum_{k\in\mathbb{Z}} p_{2k} = \sum_{k\in\mathbb{Z}} p_{2k+1} = \frac{1}{2}\sum_{k\in\mathbb{Z}} p_k = 2^{-1/2}.$$

Proof. The first two formulas follows from the orthogonal decomposition. For the remaining ones, notice that if ϕ has compact support then only a finite number of p_k is not zero. Thus we have integrate $\phi = \sum p_k \phi_k^1$ termwise, which gives

$$\int_{\mathbb{R}} \phi(x)\,dx = 2^{1/2}\sum_{k\in\mathbb{Z}} p_k \int_{\mathbb{R}} \phi(2x-k)\,dx = 2^{-1/2}\sum_{k\in\mathbb{Z}} p_k \int_{\mathbb{R}} \phi(x)\,dx,$$

which gives (here we use $\int_{\mathbb{R}} \phi(x)\,dx \neq 0$) $\sum_{k\in\mathbb{Z}} p_k = 2^{1/2}$. Also notice that

$$\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} p_{k-2l}\overline{p_k} = 1.$$

Divide the sum over k into even and odd terms gives

$$1 = |a|^2 + |b|^2, \quad a := \sum_{k\in\mathbb{Z}} p_{2k}, \; b := \sum_{k\in\mathbb{Z}} p_{2k+1}.$$

Since $a+b=\sqrt{2}$, we have

$$\sqrt{2} = a + b = \sqrt{2(|a|^2+|b|^2)},$$

which gives $a = b = \sqrt{2}/2$ by the Schwarz inequality. □

The so called *wavelet function* in multiresolution analysis is the following single function

$$\psi := \sum (-1)^k \overline{p_{1-k}}\phi_k^1.$$

that can be used to decode the orthogonal complement of V_j in V_{j+1}. Notice that $\phi = \sum p_k \phi_k^1$ gives

$$(\phi, \psi) = \sum (-1)^k p_k p_{1-k} = \cdots - p_{-1}p_2 + p_0 p_1 - p_1 p_0 + p_2 p_{-1} + \cdots = 0.$$

Put

$$\psi_k^j := 2^{j/2}\psi(2^j x - k)$$

and defines

$$W_j := \mathrm{Span}\{\psi_k^j\}_{k\in\mathbb{Z}},$$

then similar to the Haar wavelet case, we have

Theorem 6. *Each W_j is the orthogonal complement of V_j in V_{j+1} and*

$$L^2(\mathbb{R}) = \oplus_{j\in\mathbb{Z}} W_j$$

has an orthonormal basis $\{\psi_k^j\}_{j,k\in\mathbb{Z}}$.

The above theorem gives the decomposition and reconstruction algorithms in multiresolution analysis (see section 5.1 and 5.2 in [1] for details). The following theorem (see Appendix A 2.1 in [1] for the proof) provides an equivalent formulation of multiresolution analysis in terms of ϕ.

Theorem 7. *Let ϕ be a continuous function with compact support such that $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is an orthogonal system. Put*

$$V_j := \mathrm{Span}\{\phi(2^j x - k)\}_{k\in\mathbb{Z}}.$$

Then $\cap_{j\in\mathbb{Z}} V_j = \{0\}$. Assume further that

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

and ϕ is a finite combination of $\phi(2x-k)$, then $L^2(\mathbb{R}) = \overline{\cup_{j\in\mathbb{Z}} V_j}$.

How is a scaling function ϕ constructed ? The idea is to first assume a sufficiently regular ϕ exists and see what properties must be satisfied, then we reverse the process. The main idea is to use *Fourier tranform*. Recall that

$$\hat{\phi}(\xi) := \int_{\mathbb{R}} \phi(x) e^{-2\pi i x\xi}\, dx.$$

Thus $\int \phi = 1$ iff $\hat{\phi}(0) = 1$, moreover notice that

$$\widehat{\phi(x-k)}(\xi) = \hat{\phi}(\xi) e^{-2\pi i k\xi}.$$

Thus by the Plancherel identity, $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is an orthogonal system iff $\{\hat{\phi}(\xi)e^{-2\pi i k\xi}\}_{k\in\mathbb{Z}}$ is an orthogonal system, i.e.

$$||\hat{\phi}|| = 1 \text{ and } (\hat{\phi}, \hat{\phi} e^{-2\pi i k\xi}) = 0, \qquad \forall\, k \neq 0.$$

The above identities are equivalent to that

$$\widehat{|\hat{\phi}|^2}(0) = 1 \text{ and } \widehat{|\hat{\phi}|^2}(k) = 1, \qquad \forall\, k \neq 0.$$

Thus the Poisson summation formula gives

$$\sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi+k)|^2 = \sum_{k\in\mathbb{Z}} \widehat{|\hat{\phi}|^2}(k) e^{2\pi i k\xi} \equiv 1.$$

Now let us look at the final assumption in the above theorem: ϕ is a finite combination of $\phi(2x-k)$. As before we write $\phi_k^j := 2^{j/2}\phi(2^j x - k)$ and assume that $\phi = \sum p_k \phi_k^1$. Taking the Fourier transform, we get

$$\hat{\phi}(\xi) = \sum p_k 2^{-1/2} \hat{\phi}(\xi/2) e^{-\pi i k\xi}.$$

Thus

$$\hat{\phi}(2\xi) = \hat{\phi}(\xi) P(z), \quad P(z) := 2^{-1/2} \sum p_k z^k, \quad z = e^{-2\pi i\xi}.$$

Now

$$1=\sum_{k\in\mathbb{Z}}|\hat{\phi}(2\xi+k)|^2=\sum_{k\in\mathbb{Z}}|\hat{\phi}(\xi+k/2)P((-1)^k z)|^2=|P(z)|^2+|P(-z)|^2.$$

Also notice that $\hat{\phi}(0)=1$ implies that $P(1)=1$. The following theorem (Theorem 5.23 in [1]) says that one may get ϕ from P with one extra condition: $P(z)\neq 0$ if $\mathrm{Re}\, z\geq 0,\ |z|=1$.

Theorem 8. *Let $P(z):=2^{-1/2}\sum p_k z^k$ be a polynomial satisfies*

$$P(1)=1,\quad |P(z)|^2+|P(-z)|^2=1,\qquad \forall\, |z|=1,$$

and $P(z)\neq 0$ if $\mathrm{Re}\, z\geq 0,\ |z|=1$. Let ϕ_0 be the Haar scaling function, define ϕ_n inductively by

$$\phi_n(x)=\sum p_k\cdot 2^{1/2}\phi_{n-1}(2x-k),\qquad n\geq 1.$$

Then $\phi:=\lim_{n\to\infty}\phi_n$ satisfies the assumption in Theorem 7.

Remark: Put

$$P(z):=\frac{1}{8}\left((1+\sqrt{3})+(3+\sqrt{3})z+(3-\sqrt{3})z^2+(1-\sqrt{3})z^3\right).$$

One may check that $P(z)$ satisfies the assumptions above. The associated wavelet is a special Daubechies wavelet. For another approach of constructing ϕ and ψ, see Theorem 42.5.3 and Theorem 42.5.8 in [5].

2.3. **Wavelet transform.** Let $\psi\in L^2(\mathbb{R})\cap L^1(\mathbb{R})$. Assume that

$$\int_{\mathbb{R}}\psi(x)\,dx=0.$$

Then we know that $\hat{\psi}$ is an L^2 continuous function such that $\hat{\psi}(0)=0$. We need an extra condition to define the wavelet transform

$$C_\psi:=\int_{\mathbb{R}}\frac{|\hat{\psi}(\xi)|^2}{|\xi|}\,d\xi<\infty.$$

Notice that the above inequality implies $\hat{\psi}(0)=0$, which is equivalent to $\int\psi=0$. In all practical cases $\int\psi=0$ also implies $C_\psi<\infty$ (in fact assume further that $x\psi\in L^1(\mathbb{R})$ then $\hat{\psi}\in C^1(\mathbb{R})$ and $C_\psi<\infty$). The wavelet transform is defined by

$$(Tf)(a,b):=(f,\psi^{a,b}),$$

where

$$\psi^{a,b}(x):=|a|^{-1/2}\psi\left(\frac{x-b}{a}\right),\qquad a,b\in\mathbb{R},\ \ a\neq 0.$$

Since $||\psi^{a,b}||=||\psi||$, the Cauchy–Schwarz inequality gives $|(Tf)(a,b)|\leq ||f||\cdot||\psi||<\infty$, thus Tf is bounded. A function f can also be recovered from Tf (see page 24 in [2]):

Theorem 9. *Assume that* $\psi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ *and* $C_\psi = 1$. *Then for all* $f, g \in L^2(\mathbb{R})$,

$$(Tf, Tg) = (f, g),$$

where $(f,g) := \int f\bar{g}$ *and* $(Tf, Tg) = \int Tf \cdot \overline{Tg}\, \frac{dadb}{a^2}$.

Proof. The main idea of the proof is to use the Plancherel identity for Fourier transform:

$$Tf = (f, \psi^{a,b}) = (\hat{f}, \widehat{\psi^{a,b}}).$$

Notice that

$$\widehat{\psi^{a,b}}(\xi) = \int_{\mathbb{R}} |a|^{-1/2}\psi\left(\frac{x-b}{a}\right) e^{-2\pi i x\cdot\xi}\, dx,$$

change of variable $x - b = ay$ gives

$$\widehat{\psi^{a,b}}(\xi) = |a|^{1/2}\hat{\psi}(a\xi)e^{-2\pi i b\xi}.$$

Thus

$$Tf = (\hat{f}, \widehat{\psi^{a,b}}) = \int_{\mathbb{R}} \hat{f}(\xi)|a|^{1/2}\overline{\hat{\psi}(a\xi)}e^{2\pi i b\xi}\, d\xi = \mathcal{F}^{-1}\left(\hat{f}(\xi)|a|^{1/2}\overline{\hat{\psi}(a\xi)}\right),$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Apply the Plancherel identity again, we get

$$\int_{\mathbb{R}} Tf\overline{Tg}\, db = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} \cdot |a| \cdot |\hat{\psi}(a\xi)|^2\, d\xi,$$

which implies

$$(Tf, Tg) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \frac{|a| \cdot |\hat{\psi}(a\xi)|^2}{a^2}\, da\right) \hat{f}(\xi)\overline{\hat{g}(\xi)}\, d\xi.$$

Notice that the change of variable $a\xi = \eta$ gives

$$\int_{\mathbb{R}} \frac{|a| \cdot |\hat{\psi}(a\xi)|^2}{a^2}\, da = \int_{\mathbb{R}} \frac{|\hat{\psi}(\eta)|^2}{|\eta|}\, d\eta = 1.$$

Thus we get

$$(Tf, Tg) = (\hat{f}, \hat{g})$$

and Theorem 9 follows from the Plancherel identity (third time). □

Remark: Think of T as an operator, formally we have

$$(Tf, Tg) = (T^*Tf, g),$$

where T^* denotes the adjoint T. Thus the above theorem implies the following inversion formula:

$$f = T^*Tf.$$

Now let us find a formula for T^*:

$$(F, Tg) = \int_{\mathbb{R}^2} F(a,b)\overline{(g, \psi^{a,b})}\, \frac{dadb}{a^2} = \int_{\mathbb{R}^3} F(a,b)\psi^{a,b}(x)\overline{g(x)}\, \frac{dadbdx}{a^2},$$

which gives

$$(T^*F)(x) = \int_{\mathbb{R}^2} F(a,b)\psi^{a,b}(x)\, \frac{dadb}{a^2}, \qquad \forall\, F \in T(L^2(\mathbb{R})).$$

Of course the above formula is well defined only if ψ is sufficiently regular, but if we think of T^*F which maps $g \in L^2(\mathbb{R})$ to

$$\int_{\mathbb{R}^2} F(a,b)\left(\int_{\mathbb{R}} \psi^{a,b}(x)\overline{g(x)}\,dx\right)\frac{dadb}{a^2},$$

then T^*f is well defined and satisfies $(T^*F, g) = (F, Tg)$.

Compare to windowed Fourier transform (or short time Fourier transform): The main idea of the windowed Fourier transform is the following: given $w(t)$ such that

$$\int_{\mathbb{R}} w(t)\,dt = 1,$$

we call w a window function. Notice that a change of variable gives $\int w(t-b)\,db = 1$ for every $t \in \mathbb{R}$ thus

$$f(t) = f(t)\cdot\int_{\mathbb{R}} w(t-b)\,db = \int_{\mathbb{R}} f(t)w(t-b)\,db, \ \ \forall\, b \in \mathbb{R}.$$

Now we can write the Fourier transform $\hat{f}(\lambda) = \int f(t)e^{-2\pi it\lambda}\,dt$ as

$$\hat{f}(\lambda) = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(t)w(t-b)\,db\right)e^{-2\pi it\lambda}\,dt = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(t)w(t-b)e^{-2\pi it\lambda}\,dt\right)db.$$

We call

$$(Wf)(\lambda,b) := \int_{\mathbb{R}} f(t)w(t-b)e^{-2\pi it\lambda}\,dt$$

the windowed Fourier transform, which gives another expression for the Fourier tarnsform:

$$\hat{f}(\lambda) = \int_{\mathbb{R}} (Wf)(\lambda,b)\,db.$$

The Fourier inversion formula gives

$$f(t) = \int_{\mathbb{R}} \hat{f}(\lambda)e^{2\pi it\lambda}\,d\lambda = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} (Wf)(\lambda,b)\,db\right)e^{2\pi it\lambda}\,d\lambda.$$

For more on windowed Fourier transform (usually we consider the case that $w = |\tilde{w}|^2$, $||\tilde{w}|| = 1$ and define the windowed Fourier transform using $\tilde{w}$), see section 2.7–2.8 in [2].

REFERENCES

[1] A. Boggess and F. Narcowich, *A first course in Wavelets with Fourier Analysis*, Wiley, 2nd Edition, 2009.
[2] I. Daubechies, *Ten lectures on wavelets*, Philadelphia, PA, 1992.
[3] H. Dym and H. P. McKean, *Fourier Series and Integrals*, Academic Press, 1972.
[4] W. Feller, *An Introduction to Probability Theory and its Applications*, Volume I, II around 1968.
[5] C. Gasquest and P. Witomski, *Fourier analysis and applications*, TAM 30.
[6] E. Stein and R. Shakarchi, *Fourier analysis*, Princeton lectures in analysis.

DEPARTMENT OF MATHEMATICAL SCIENCES, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY, NO-7491 TRONDHEIM, NORWAY

E-mail address: xu.wang@ntnu.no