A Concise Lecture Note on Fourier Analysis

1 Review on Analysis/Advanced Calculus

1.1 Pointwise and Uniform Convergence（逐點收斂與均勻收斂）

Definition 1.1. Let $I \subseteq \mathbb{R}$ be an interval, and $f_k, f : I \to \mathbb{R}$ be functions for $k = 1, 2, \cdots$. The sequence of functions $\{f_k\}_{k=1}^\infty$ is said to ***converge pointwise*** if $\{f_k(a)\}_{k=1}^\infty$ converges for all $a \in I$. In other words, $\{f_k\}_{k=1}^\infty$ converges pointwise if there exists a function $f : I \to \mathbb{R}$ such that

$$\lim_{k\to\infty} |f_k(x) - f(x)| = 0 \qquad \forall\, x \in I.$$

In this case, $\{f_k\}_{k=1}^\infty$ is said to converge pointwise to f and is denoted by $f_k \to f$ p.w..

The sequence of functions $\{f_k\}_{k=1}^\infty$ is said to ***converge uniformly*** on I if there exists $f : I \to \mathbb{R}$ such that

$$\lim_{k\to\infty} \sup_{x\in I} |f_k(x) - f(x)| = 0\,.$$

In this case, $\{f_k\}_{k=1}^\infty$ is said to converge uniformly to f on I. In other words, $\{f_k\}_{k=1}^\infty$ converges uniformly to f on I if for every $\varepsilon > 0$, $\exists\, N > 0$ such that

$$|f_k(x) - f(x)| < \varepsilon \qquad \forall\, k \geqslant N \text{ and } x \in I\,.$$

Proposition 1.2. *Let $I \subseteq \mathbb{R}$ be an interval, and $f_k, f : I \to \mathbb{R}$ be functions for $k = 1, 2, \cdots$. If $\{f_k\}_{k=1}^\infty$ converges uniformly to f on I, then $\{f_k\}_{k=1}^\infty$ converges pointwise to I.*

Proposition 1.3 (Cauchy criterion for uniform convergence)**.** *Let $I \subseteq \mathbb{R}$ be an interval, and $f_k : I \to \mathbb{R}$ be a sequence of functions. Then $\{f_k\}_{k=1}^\infty$ converges uniformly (to some function f) on I if and only if for every $\varepsilon > 0$, $\exists\, N > 0$ such that*

$$|f_k(x) - f_\ell(x)| < \varepsilon \qquad \forall\, k, \ell \geqslant N \text{ and } x \in I\,.$$

Theorem 1.4. *Let $I \subseteq \mathbb{R}$ be an interval, and $f_k : I \to \mathbb{R}$ be a sequence of continuous functions converging to $f : I \to \mathbb{R}$ uniformly on I. Then f is continuous on I; that is,*

$$\lim_{x\to a} f(x) = \lim_{x\to a} \lim_{k\to\infty} f_k(x) = \lim_{k\to\infty} \lim_{x\to a} f_k(x) = f(a)\,.$$

Theorem 1.5. *Let $I \subseteq \mathbb{R}$ be a finite interval, $f_k : I \to \mathbb{R}$ be a sequence of differentiable functions, and $g : I \to \mathbb{R}$ be a function. Suppose that $\{f_k(a)\}_{k=1}^\infty$ converges for some $a \in I$, and $\{f_k'\}_{k=1}^\infty$ converges uniformly to g on I. Then*

1. *$\{f_k\}_{k=1}^\infty$ converges uniformly to some function f on I.*

2. *The limit function f is differentiable on I, and $f'(x) = g(x)$ for all $x \in I$; that is,*

$$\lim_{k\to\infty} f_k'(x) = \lim_{k\to\infty} \frac{d}{dx} f_k(x) = \frac{d}{dx} \lim_{k\to\infty} f_k(x) = f'(x)\,.$$

Theorem 1.6. *Let $f_k : [a,b] \to \mathbb{R}$ be a sequence of Riemann integrable functions which converges uniformly to f on $[a,b]$. Then f is Riemann integrable, and*

$$\lim_{k\to\infty}\int_a^b f_k(x)dx = \int_a^b \lim_{k\to\infty} f_k(x)dx = \int_a^b f(x)dx\,. \tag{1.1}$$

Definition 1.7. Let $I \subseteq \mathbb{R}$ be an interval. The collection of bounded continuous real-valued functions defined on I is denoted by $\mathscr{C}_b(I;\mathbb{R})$. The sup-norm of $\mathscr{C}_b(I;\mathbb{R})$, denoted by $\|\cdot\|_\infty$, is defined by

$$\|f\|_\infty = \sup_{x\in I}\big|f(x)\big| \qquad \forall\, f \in \mathscr{C}_b(I;\mathbb{R})\,.$$

If $I = [a,b] \subseteq \mathbb{R}$ is a closed interval (so that every continuous function on I is bounded), we simply use $\mathscr{C}([a,b];\mathbb{R})$ to denote $\mathscr{C}_b([a,b];\mathbb{R})$.

Having the definition above, we can rephrase Proposition 1.3 and Theorem 1.4 as follows.

Theorem 1.8. *Let $I \subseteq \mathbb{R}$ be an interval. Then $\big(\mathscr{C}_b(I;\mathbb{R}), \|\cdot\|_\infty\big)$ is a complete norm space; that is, every Cauchy sequence in $\big(\mathscr{C}_b(I;\mathbb{R}), \|\cdot\|_\infty\big)$ converges uniformly (to some limit) in $\mathscr{C}_b(I;\mathbb{R})$.*

1.2 Series of Functions and The Weierstrass M-Test

Definition 1.9. Let $I \subseteq \mathbb{R}$ be an interval, and $g_k : I \to \mathbb{R}$ (or $\mathbb{C}$) be a sequence of functions. We say that the series $\sum\limits_{k=1}^{\infty} g_k$ converges pointwise if the sequence of partitial sum $\{s_n\}_{n=1}^{\infty}$ given by

$$s_n = \sum_{k=1}^{n} g_k$$

converges pointwise. We say that $\sum\limits_{k=1}^{\infty} g_k$ converges uniformly on I if $\{s_n\}_{n=1}^{\infty}$ converges uniformly on I.

The following two corollaries are direct consequences of Proposition 1.3 and Theorem 1.4.

Corollary 1.10. *Let $I \subseteq \mathbb{R}$ be an interval, and $g_k : I \to \mathbb{R}$ be functions. Then $\sum\limits_{k=1}^{\infty} g_k$ converges uniformly on I if and only if*

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni \Big|\sum_{k=m+1}^{n} g_k(x)\Big| < \varepsilon \qquad \forall\, n > m \geqslant N \text{ and } x \in A\,.$$

Corollary 1.11. *Let $I \subseteq \mathbb{R}$ be an interval, and $g_k, g : I \to \mathbb{R}$ be functions. If $g_k : I \to \mathbb{R}$ are continuous and $\sum\limits_{k=1}^{\infty} g_k(x)$ converges to g uniformly on I, then g is continuous.*

Theorem 1.12 (Weierstrass M-test)**.** *Let $I \subseteq \mathbb{R}$ be an interval, and $g_k : I \to \mathbb{R}$ be a sequence of functions. Suppose that $\exists\, M_k > 0$ such that $\sup\limits_{x\in I}|g_k(x)| \leqslant M_k$ for all $k \in \mathbb{N}$ and $\sum\limits_{k=1}^{\infty} M_k$ converges. Then $\sum\limits_{k=1}^{\infty} g_k$ and $\sum\limits_{k=1}^{\infty} |g_k|$ both converge uniformly on I.*

Corollary 1.13. *Let $I \subseteq \mathbb{R}$ be an interval, and $g_k : I \to \mathbb{R}$ be a sequence of continuous functions. Suppose that $\exists\, M_k > 0$ such that $\sup\limits_{x\in I} |g_k(x)| \leqslant M_k$ for all $k \in \mathbb{N}$ and $\sum\limits_{k=1}^{\infty} M_k$ converges. Then $\sum\limits_{k=1}^{\infty} g_k$ is continuous on I.*

The following two theorems are direct consequences of Theorem 1.5 and 1.6.

Theorem 1.14. *Let $g_k : [a,b] \to \mathbb{R}$ be a sequence of Riemann integrable functions. If $\sum\limits_{k=1}^{\infty} g_k$ converges uniformly on $[a,b]$, then*

$$\int_a^b \sum_{k=1}^{\infty} g_k(x)dx = \sum_{k=1}^{\infty} \int_a^b g_k(x)dx\,.$$

Theorem 1.15. *Let $g_k : (a,b) \to \mathbb{R}$ be a sequence of differentiable functions. Suppose that $\sum\limits_{k=1}^{\infty} g_k(c)$ converges for some $c \in (a,b)$, and $\sum\limits_{k=1}^{\infty} g_k'$ converges uniformly on (a,b). Then*

$$\sum_{k=1}^{\infty} g_k'(x) = \frac{d}{dx}\sum_{k=1}^{\infty} g_k(x)\,.$$

1.3 Analytic functions and the Stone-Weierstrass theorem

Theorem 1.16. *Let $f : (a,b) \to \mathbb{R}$ be an infinitely differentiable functions; that is, $f^{(k)}(x)$ exists for all $k \in \mathbb{N}$ and $x \in (a,b)$. Let $c \in (a,b)$ and suppose that for some $0 < h < \infty$, $\big|f^{(k)}(x)\big| \leqslant M$ for all $x \in (c-h, c+h) \subseteq (a,b)$. Then*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k \qquad \forall\, x \in (c-h, c+h)\,.$$

Moreover, the convergence is uniform.

Proof. First, we claim that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + (-1)^n \int_c^x \frac{(y-x)^n}{n!} f^{(n+1)}(y)dy \qquad \forall\, x \in (a,b)\,. \tag{1.2}$$

By the fundamental theorem or Calculus it is clear that (1.2) holds for $n = 0$. Suppose that (1.2) holds for $n = m$. Then

$$\begin{aligned} f(x) &= \sum_{k=0}^{m} \frac{f^{(k)}(c)}{k!}(x-c)^k + (-1)^m \Big[\frac{(y-x)^{m+1}}{(m+1)!} f^{(m+1)}(y)\Big|_{y=c}^{y=x} - \int_c^x \frac{(y-x)^{m+1}}{(m+1)!} f^{(m+2)}(y)dy\Big] \\ &= \sum_{k=0}^{m+1} \frac{f^{(k)}(c)}{k!}(x-c)^k + (-1)^{m+1} \int_c^x \frac{(y-x)^{m+1}}{(m+1)!} f^{(m+2)}(y)dy \end{aligned}$$

which implies that (1.2) also holds for $n = m+1$. By induction (1.2) holds for all $n \in \mathbb{N}$.

Letting $s_n(x) = \sum\limits_{k=0}^{n} \dfrac{f^{(k)}(c)}{k!}(x-c)^k$, then if $x \in (c-h, c+h)$,

$$\big|s_n(x) - f(x)\big| \leqslant \Big|\int_c^x \frac{h^n}{n!} M dy\Big| \leqslant \frac{h^{n+1}}{n!} M\,.$$

Let $\varepsilon > 0$ be given. Since $\lim\limits_{n\to\infty} \dfrac{h^{n+1}}{n!} M = 0$, $\exists\, N > 0$ such that $\Big|\dfrac{h^{n+1}}{n!}\Big| M < \varepsilon$ if $n \geqslant N$. As a consequence, if $n \geqslant N$,

$$\big|s_n(x) - f(x)\big| < \varepsilon \quad \text{whenever } n \geqslant N\,.$$

□

Definition 1.17. Let $I \subseteq \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}$ is said to be ***real analytic*** at $a \in \text{int}(I)$ if $f(x) = \sum\limits_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$ in a neighborhood of a.

Theorem 1.18 (Weierstrass)**.** *For every given $f \in \mathscr{C}([0,1];\mathbb{R})$ there exists a sequence of polynomials $\{p_n\}_{n=1}^{\infty}$ such that $\{p_n\}_{n=1}^{\infty}$ converges uniformly to f on $[0,1]$. In other words, the collection of all polynomials is dense in the space $\big(\mathscr{C}([0,1];\mathbb{R}), \|\cdot\|_\infty\big)$.*

Proof. Let $r_k(x) = C_k^n x^k (1-x)^{n-k}$. By looking at the partial derivatives with respect to x of the identity $(x+y)^n = \sum\limits_{k=0}^{n} C_k^n x^k y^{n-k}$, we find that

1. $\sum\limits_{k=0}^{n} r_k(x) = 1$; 2. $\sum\limits_{k=0}^{n} k r_k(x) = nx$; 3. $\sum\limits_{k=0}^{n} k(k-1) r_k(x) = n(n-1)x^2$.

As a consequence,

$$\sum_{k=0}^{n}(k-nx)^2 r_k(x) = \sum_{k=0}^{n}\big[k(k-1) + (1-2nx)k + n^2x^2\big] r_k(x) = nx(1-x)\,.$$

Let $\varepsilon > 0$ be given. Since $f : [0,1] \to \mathbb{R}$ is continuous on a compact set $[0,1]$, f is uniformly continuous on $[0,1]$; thus

$$\exists\, \delta > 0 \ni \big|f(x) - f(y)\big| < \frac{\varepsilon}{2} \qquad \text{if } |x-y| < \delta,\ x, y \in [0,1]\,.$$

Consider the ***Bernstein polynomial*** $p_n(x) = \sum\limits_{k=0}^{\infty} f\big(\frac{k}{n}\big) r_k(x)$. Note that

$$\begin{aligned}
\big|f(x) - p_n(x)\big| &= \Big|\sum_{k=0}^{n}\Big(f(x) - f\big(\frac{k}{n}\big)\Big) r_k(x)\Big| \leqslant \sum_{k=0}^{n}\Big|f(x) - f\big(\frac{k}{n}\big)\Big| r_k(x) \\
&\leqslant \Big(\sum_{|k-nx|<\delta n} + \sum_{|k-nx|\geqslant\delta n}\Big)\Big|f(x) - f\big(\frac{k}{n}\big)\Big| r_k(x) \\
&< \frac{\varepsilon}{2} + 2\|f\|_\infty \sum_{|k-nx|\geqslant\delta n}\frac{(k-nx)^2}{(k-nx)^2} r_k(x) \\
&\leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n^2\delta^2}\sum_{k=0}^{n}(k-nx)^2 r_k(x) \leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n\delta^2}x(1-x) \leqslant \frac{\varepsilon}{2} + \frac{\|f\|_\infty}{2n\delta^2}\,.
\end{aligned}$$

Choose N large enough such that $\frac{\|f\|_\infty}{2N\delta^2} < \frac{\varepsilon}{2}$. Then for all $n \geqslant N$,

$$\|f - p_n\|_\infty = \sup_{x\in[0,1]}\big|f(x) - p_n(x)\big| < \varepsilon\,. \qquad \square$$

Remark 1.19. A polynomial of the form $p_n(x) = \sum\limits_{k=0}^{n} \beta_k r_k(x)$ is called a ***Bernstein polynomial of degree*** n, and the coefficients β_k are called Bernstein coefficients.

Corollary 1.20. *The collection of polynomials on $[a,b]$ is dense in $\big(\mathscr{C}([a,b];\mathbb{R}), \|\cdot\|_\infty\big)$; that is, for every $f \in \mathscr{C}([a,b];\mathbb{R})$ there exists a sequence of polynomials $\{p_k\}_{k=1}^{\infty}$ such that $\{p_k\}_{k=1}^{\infty}$ converges uniformly to f on $[a,b]$.*

Proof. We note that $g \in \mathscr{C}([a,b];\mathbb{R})$ if and only if $f(x) = g\big(x(b-a)+a\big) \in \mathscr{C}([0,1];\mathbb{R})$; thus

$$\big|f(x) - p(x)\big| < \varepsilon\ \forall\, x \in [0,1] \Leftrightarrow \Big|g(x) - p\big(\frac{x-a}{b-a}\big)\Big| < \varepsilon\ \forall\, x \in [a,b]\,. \qquad \square$$

1.4 Trigonometric polynomials and the space of 2π-periodic continuous functions

In this section, we focus on the approximations of a special class of functions, the collection of 2π-periodic continuous function. Let $\mathscr{C}(\mathbb{T})$ denote the collection of 2π-periodic continuous function (defined on $\mathbb{R}$):

$$\mathscr{C}(\mathbb{T}) = \left\{ f \in \mathscr{C}(\mathbb{R};\mathbb{R}) \,\middle|\, f(x+2\pi) = f(x) \quad \forall\, x \in \mathbb{R} \right\}.$$

The sup-norm on $\mathscr{C}(\mathbb{T})$ is denoted by $\|\cdot\|_{L^\infty(\mathbb{T})}$; that is, $\|f\|_{L^\infty(\mathbb{T})} \equiv \sup\limits_{x\in\mathbb{R}} |f(x)|$ if $f \in \mathscr{C}(\mathbb{T})$.

We note that $\mathscr{C}(\mathbb{T})$ can be treated as the collection of continuous functions defined on the unit circle $\mathbb{S}^1$ in the sense that every $f \in \mathscr{C}(\mathbb{T})$ corresponds to a unique $F \in \mathscr{C}(\mathbb{S}^1;\mathbb{R})$ such that

$$f(x) = F(\cos x, \sin x) \qquad \forall\, x \in \mathbb{R} \tag{1.3}$$

and vice versa.

Definition 1.21. A family of functions $\left\{\varphi_n \in \mathscr{C}(\mathbb{T}) \,\middle|\, n \in \mathbb{N}\right\}$ is said to be ***an approximation of the identity*** if

(1) $\varphi_n(x) \geqslant 0$;

(2) $\displaystyle\int_{\mathbb{T}} \varphi_n(x)\,dx = 1$ for every $n \in \mathbb{N}$, here we identify $\mathbb{T}$ with the interval $[-\pi, \pi]$;

(3) $\displaystyle\lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\,dx = 0$ for every $\delta > 0$.

Definition 1.22 (Convolutions on $\mathbb{T}$)**.** The convolution of two (continuous) function $f, g : \mathbb{T} \to \mathbb{R}$ is the function $f \star g : \mathbb{T} \to \mathbb{C}$ defined by the integral

$$(f \star g)(x) = \int_{\mathbb{T}} f(x-y) g(y)\,dy\,.$$

Theorem 1.23. *If $\{\varphi_n\}_{n=1}^\infty$ is an approximation of the identity and $f \in \mathscr{C}(\mathbb{T})$, then $\varphi_n \star f$ converges uniformly to f as $n \to \infty$.*

Proof. Without loss of generality, we may assume that $f \not\equiv 0$. By the definition of the convolution,

$$\begin{aligned}\big|(\varphi_n \star f)(x) - f(x)\big| &= \int_{\mathbb{T}} \varphi_n(x-y) f(y)\,dy - f(x) \\ &= \int_{\mathbb{T}} \varphi_n(x-y)\big(f(x) - f(y)\big)dy\,,\end{aligned}$$

where we use (2) of Definition 1.21 to obtain the last equality. Now given $\varepsilon > 0$. Since $f \in \mathscr{C}(\mathbb{T})$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \dfrac{\varepsilon}{2}$ whenever $|x-y| < \delta$. Therefore,

$$\begin{aligned}&|(\varphi_n \star f)(x) - f(x)| \\ &\quad\leqslant \int_{|x-y|<\delta} \varphi_n(x-y)\big|f(x) - f(y)\big|dy + \int_{\delta \leqslant |x-y|} \varphi_n(x-y)\big|f(x) - f(y)\big|dy \\ &\quad\leqslant \frac{\varepsilon}{2} \int_{\mathbb{T}} \varphi_n(x-y)\,dy + 2\max_{\mathbb{T}} |f| \int_{\delta \leqslant |z| \leqslant \pi} \varphi_n(z)\,dz\,.\end{aligned}$$

By (3) of Definition 1.21, there exists $N > 0$ such that if $n \geqslant N$,

$$\int_{\delta \leqslant |z| \leqslant \pi} \varphi_n(z)\, dx < \frac{\varepsilon}{4 \max_{\mathbb{T}} |f|}\,.$$

Therefore, for $n \geqslant N$, $\big|(\varphi_n \star f)(x) - f(x)\big| < \varepsilon$ for all $x \in \mathbb{T}$. □

Definition 1.24. A trigonometric polynomial $p(x)$ of degree n is a finite sum of the form

$$p(x) = \frac{c_0}{2} + \sum_{k=1}^{n} (c_k \cos kx + s_k \sin kx) \qquad x \in \mathbb{R}\,.$$

The collection of all trigonometric polynomial of degree n is denoted by $\mathscr{P}_n(\mathbb{T})$, and the collection of all trigonometric polynomials is denoted by $\mathscr{P}(\mathbb{T})$; that is, $\mathscr{P}(\mathbb{T}) = \bigcup\limits_{n=0}^{\infty} \mathscr{P}_n(\mathbb{T})$.

On account of the Euler identity $e^{i\theta} = \cos\theta + i \sin\theta$, a trigonometric polynomial of degree n can also be written as

$$p(x) = \sum_{k=-n}^{n} a_k e^{ikx} \quad \text{with} \quad a_k = \frac{c_{|k|} - i s_{|k|}}{2}\,,$$

where s_0 is taken to be 0. Therefore, every trigonometric polynomial of degree n is associated to a unique function of the form $\sum\limits_{k=-n}^{n} a_k e^{ikx}$ and vice versa.

Having defined trigonometric polynomials, we can show that every 2π-periodic function can be approximated by a sequence of trigonometric polynomials in the sense of uniform convergence.

Theorem 1.25. *The collection of all trigonometric polynomials $\mathscr{P}(\mathbb{T})$ is dense in $\mathscr{C}(\mathbb{T})$ with respect to the sup-norm; that is, for every $f \in \mathscr{C}(\mathbb{T})$ there exists a sequence $\{p_n\}_{n=1}^{\infty} \subseteq \mathscr{P}(\mathbb{T})$ such that $\{p_n\}_{n=1}^{\infty}$ converges uniformly to f on $\mathbb{T}$.*

Proof. Let $\varphi_n(x) = c_n(1+\cos x)^n$, where c_n is chosen so that $\int_{\mathbb{T}} \varphi_n(x)\, dx = 1$. By the residue theorem,

$$\int_{\mathbb{T}} (1+\cos x)^n dx = \oint_{\mathbb{S}^1} \Big(1 + \frac{z^2+1}{2z}\Big)^n \frac{dz}{iz} = \frac{1}{i2^n} \oint_{\mathbb{S}^1} \frac{(z+1)^{2n}}{z^{n+1}} dz = \frac{\pi}{2^{n-1}} \binom{2n}{n};$$

thus $c_n = \dfrac{2^{n-1}}{\pi} \dfrac{(n!)^2}{(2n)!}$.

Now $\{\varphi_n\}_{n=1}^{\infty}$ is clearly non-negative and satisfies (2) of Definition 1.21 for all $n \in \mathbb{N}$. Let $\delta > 0$ be given.

$$\int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\, dx \leqslant \int_{\delta \leqslant |x| \leqslant \pi} c_n (1+\cos\delta)^n dx \leqslant 2^{2n} \Big(\frac{1+\cos\delta}{2}\Big)^n \frac{(n!)^2}{(2n)!}\,.$$

By Stirling's formula $\lim\limits_{n\to\infty} \dfrac{n!}{\sqrt{2\pi n} n^n e^{-n}} = 1$,

$$\begin{aligned} \lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\, dx &\leqslant \lim_{n\to\infty} 2^{2n} \Big(\frac{1+\cos\delta}{2}\Big)^n \frac{\big(\sqrt{2\pi n} n^n e^{-n}\big)^2}{\sqrt{2\pi(2n)}(2n)^{2n} e^{-2n}} \\ &= \lim_{n\to\infty} \sqrt{\pi n} \Big(\frac{1+\cos\delta}{2}\Big)^n = 0\,. \end{aligned}$$

So $\{\varphi_n\}_{n=1}^{\infty}$ is an approximation of the identity. By Theorem 1.23, $\varphi_k \star f$ converges uniformly to f if $f \in \mathscr{C}(\mathbb{T})$, while $\varphi_n \star f$ is a trigonometric function. □

Remark 1.26. Theorem 1.25 can also be proved using the abstract version of the Stone-Weierstrass Theorem and the identification (1.3). See Theorem 7.32 in "Principles of Mathematics Analysis" by W. Rudin or Theorem 5.8.2 in Elementary Classical Analysis by J. Marsden and M. Hoffman for the Stone-Weierstrass Theorem.

2 Fourier Series

讓我們回顧一下之前已經有的一些結論。在 §1.3 中我們學到了 Stone-Weierstrass 定理，它告訴我們定義在 $[0,1]$ 上的連續函數 f 可以用多項式（例如 Bernstein 多項式）去逼近（在均勻收斂的意義下)，而我們也注意到 Bernstein 多項式，在取不同次數 n 的多項式做逼近時，每一個單項式 x^k 前面的係數都跟 n 和 k 有關。但是從定理 1.16 中我們又發現，對某些擁有很好性質的函數 f（叫做解析函數 Analytic functions)，即使取不同次數 n 的多項式做逼近時，每個單項式 x^k 前面的係數可以取成只跟函數 f 的 k 次導數有關（跟 n 無關)。這給了我們一個很粗略的概念，知道想用多項式去逼近連續函數時，多項式的係數有些時候會跟多項次的次數有關，有時則無關。

在這一章中，我們在前四節特別關注在週期為 2π 的連續函數。由定理 1.25 我們知道這樣的函數可用形如

$$p_n(x) = \frac{c_0^{(n)}}{2} + \sum_{k=1}^{n}(c_k^{(n)}\cos kx + s_k^{(n)}\sin kx)$$

的三角多項式 (trigonometric polynomials) 所逼近（在均勻收斂的意義下)，其中上標 (n) 代表的是係數可能與用來逼近的三角多項式的次數 n 有關係。跟前一段所述的經驗類似，在數學理論上我們想知道下面問題的答案：

1. 什麼樣的函數，可以用係數與逼近次數無關的三角多項式去逼近。對這樣的函數，三角多項式要怎麼挑？
2. 對於實在沒辦法用係數與逼近次數無關的三角多項式去逼近的連續週期函數，有什麼好的方法逼近？而上面所挑出來的那個係數跟逼近次數無關的三角多項式，在次數接近無窮大時出了什麼問題？

上述的問題解決之後，我們用變數變換，也可以得到對於週期為 $2L$ 的函數的相關理論。

另外，由於在進行的過程中，我們發現我們所關心用來逼近連續函數的三角多項式（叫富氏級數)，其係數的定法只要求函數可積分即可，因此，一個自然衍生的問題則是：對不連續（但可積分）的函數來說，有沒有什麼收斂理論可以說明？這個部份的研究則是第四、五節的主要重點。在第六節中，我們則提供了一個快速傅利葉變換 (FFT) 的演算法可供電腦去計算富氏級數（的係數)。

2.1 Basic properties of the Fourier series

Let $f \in \mathscr{C}(\mathbb{T})$ be given. We first assume that the trigonometric polynomials used to approximate f can be chosen in such a way that the coefficients does not depend on the degree of approximation; that is, $c_k^{(n)} = c_k$ and $s_k^{(n)} = s_k$. In this case, if $p_n \to f$ uniformy on $[-\pi,\pi]$, by Theorem 1.6 we must have

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} p_n(x)\cos kx\,dx = \int_{-\pi}^{\pi} f(x)\cos kx\,dx \qquad \forall\, k \in \{0,1,\cdots,n\}$$

and

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} p_n(x)\sin kx\,dx = \int_{-\pi}^{\pi} f(x)\sin kx\,dx \qquad \forall\, k \in \{1,\cdots,n\}\,.$$

Since

$$\int_{-\pi}^{\pi}\cos kx\cos \ell x\,dx = \int_{-\pi}^{\pi}\sin kx\sin \ell x\,dx = \pi\delta_{k\ell} \qquad \forall\, k,\ell \in \mathbb{N}$$

and

$$\int_{-\pi}^{\pi}\sin kx\cos \ell x\,dx = 0 \qquad \forall\, k \in \mathbb{N}, \ell \in \mathbb{N}\cup\{0\}\,,$$

we find that

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx \quad \text{and} \quad s_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx\,. \tag{2.1}$$

This induces the following

Definition 2.1. For a Riemann integrable function $f : [-\pi,\pi] \to \mathbb{R}$, the ***Fourier series representation*** of f, denoted by $s(f,\cdot)$, is given by

$$s(f,x) = \frac{c_0}{2} + \sum_{k=1}^{\infty}(c_k\cos kx + s_k\sin kx)$$

whenever the sum makes sense, where sequences $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ given by (2.1) are called the ***Fourier coefficients*** associated with f. The n-th partial sum of the Fourier series representation to f, denoted by $s_n(f,\cdot)$, is given by

$$s_n(f,x) = \frac{c_0}{2} + \sum_{k=1}^{n}(c_k\cos kx + s_k\sin kx)\,.$$

We note that for the Fourier series $s(f,x)$ to be defined, f is not necessary continuous. Our goal is to establish the convergence of Fourier series in various senses.

Remark 2.2. Because of the Euler identity $e^{i\theta} = \cos\theta + i\sin\theta$, we can write

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)(e^{iky} + e^{-iky})dy \quad \text{and} \quad s_k = \frac{1}{2\pi i}\int_{-\pi}^{\pi} f(y)(e^{iky} - e^{-iky})dy$$

thus

$$\begin{aligned}
s_n(f,x) &= \frac{c_0}{2} + \sum_{k=1}^{n}\Big(c_k\frac{e^{ikx}+e^{-ikx}}{2} + s_k\frac{e^{ikx}-e^{-ikx}}{2i}\Big) \\
&= \frac{1}{2}\Big[c_0 + \sum_{k=1}^{n}\big((c_k - is_k)e^{ikx} + (c_k + is_k)e^{-ikx}\big)\Big] \\
&= \frac{1}{2}\Big[c_0 + \sum_{k=1}^{n}\big((c_k - is_k)e^{ikx} + \sum_{k=-n}^{-1}(c_{-k} + is_{-k})e^{ikx}\Big] \\
&= \frac{1}{2}\Big[c_0 + \frac{1}{\pi}\sum_{k=1}^{n}\int_{-\pi}^{\pi} f(y)e^{-iky}dy e^{ikx} s + \frac{1}{\pi}\sum_{k=-n}^{-1}\int_{-\pi}^{\pi} f(y)e^{-iky}dy e^{ikx}\Big]\,.
\end{aligned}$$

Define $\widehat{f}_k = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(y)e^{-iky}dy$. Then

$$s_n(f,x) = \sum_{k=-n}^{n} \widehat{f}_k e^{ikx}\,.$$

The sequence $\{\widehat{f}_k\}_{k=-\infty}^{\infty}$ is also called the Fourier coefficients associated with f, and one can write the Foruier series representation of f as $\sum\limits_{k=-\infty}^{\infty} \widehat{f}_k e^{ikx}$.

Remark 2.3. Given a continuous function g with period $2L$, let $f(x) = g\big(\frac{Lx}{\pi}\big)$. Then f is a continuous function with period 2π, and the Fourier series representation of f is given by

$$s(f,x) = \frac{c_0}{2} + \sum_{k=1}^{n}(c_k \cos kx + s_k \sin kx)\,,$$

where c_k and s_k are given by (2.1). Now, define the Fourier series representation of g by $s(g,x) = s\big(f, \frac{\pi x}{L}\big)$. Then the Fourier series representation of g is given by

$$s(g,x) = \frac{c_0}{2} + \sum_{k=1}^{\infty}\Big(c_k \cos\frac{k\pi x}{L} + s_k \sin\frac{k\pi x}{L}\Big)\,,$$

where $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ is also called the Fourier coefficients associated with g and are given by

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx = \frac{1}{\pi}\int_{-\pi}^{\pi} g\Big(\frac{Lx}{\pi}\Big)\cos kx\,dx = \frac{1}{L}\int_{-L}^{L} g(x)\cos\frac{k\pi x}{L}\,dx$$

and similarly, $s_k = \dfrac{1}{L}\displaystyle\int_{-L}^{L} g(x)\sin\frac{k\pi x}{L}\,dx$.

Example 2.4. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

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

and $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$. To find the Fourier representation of f, we compute the Fourier coefficients by

$$s_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx = \frac{1}{\pi}\Big(\int_0^{\pi} x\sin kx\,dx - \int_{-\pi}^{0} x\sin kx\,dx\Big) = 0$$

and

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx = \frac{1}{\pi}\Big(\int_0^{\pi} x\cos kx\,dx - \int_{-\pi}^{0} x\cos kx\,dx\Big) = \frac{2}{\pi}\int_0^{\pi} x\cos kx\,dx\,.$$

If $k=0$, then $c_0 = \dfrac{2}{\pi}\displaystyle\int_0^{\pi} x\,dx = \pi$, while if $k \in \mathbb{N}$,

$$c_k = \frac{2}{\pi}\Big(\frac{x\sin kx}{k}\Big|_0^{\pi} - \int_0^{\pi}\frac{\sin kx}{k}\,dx\Big) = \frac{2}{\pi}\frac{\cos kx}{k^2}\Big|_0^{\pi} = \frac{2((-1)^k - 1)}{\pi k^2}\,.$$

Therefore, $c_{2k} = 0$ and $c_{2k-1} = -\dfrac{4}{\pi(2k-1)^2}$ for all $k \in \mathbb{N}$. Therefore, the Fourier series representation of f is given by

$$s(f,x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\cos(2k-1)x}{(2k-1)^2}\,.$$

Example 2.5. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } -\dfrac{\pi}{2} \leqslant x \leqslant \dfrac{\pi}{2}\,, \\ 0 & \text{if } -\pi \leqslant x < -\dfrac{\pi}{2} \text{ or } \dfrac{\pi}{2} < x \leqslant \pi\,, \end{cases}$$

and $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. We compute the Fourier coefficients of f and find that $s_k = 0$ for all $k \in \mathbb{N}$ and $c_0 = 1$, as well as

$$c_k = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos kx\, dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \cos kx\, dx = \frac{2\sin\frac{k\pi}{2}}{\pi k}\,.$$

Therefore, $c_{2k} = 0$ and $c_{2k-1} = \dfrac{2(-1)^{k+1}}{\pi(2k-1)}$ for all $k \in \mathbb{N}$; thus the Fourier series representation of f is given by

$$s(f, x) = \frac{1}{2} - \frac{2}{\pi}\sum_{k=1}^{\infty} \frac{(-1)^k}{2k-1}\cos(2k-1)x\,.$$

Example 2.6. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x \qquad \text{if} \quad -\pi < x \leqslant \pi$$

and $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. Then the Fourier coefficients of f are computed as follows: $c_k = 0$ for all $k \in \mathbb{N} \cup \{0\}$ since f is (more or less) an odd function, and

$$s_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin kx\, dx = \frac{2}{\pi}\int_0^{\pi} x\sin kx\, dx = \frac{2}{\pi}\Big(-\frac{x\cos kx}{k}\Big|_0^{\pi} + \int_0^{\pi}\frac{\cos kx}{k}\,dx\Big) = \frac{2(-1)^{k+1}}{k}\,.$$

Therefore, the Fourier series representation of f is given by

$$s(f, x) = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\sin kx\,.$$

2.2 Uniform Convergence of the Fourier Series

Before proceeding, we note that Remark 2.2 implies that

$$s_n(f, x) = \sum_{k=-n}^{n}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{ik(x-y)}\, dy = \int_{-\pi}^{\pi} f(y)\sum_{k=-n}^{n} e^{ik(x-y)}\, dy\,.$$

By defining

$$\begin{aligned} D_n(x) &= \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} = \frac{1}{2\pi}\frac{e^{-inx}\big[e^{i(2n+1)x} - 1\big]}{e^{ix} - 1} \\ &= \frac{1}{2\pi}\frac{e^{i(n+1/2)x} - e^{-i(n+1/2)x}}{e^{ix/2} - e^{-ix/2}} = \frac{\sin(n + \frac{1}{2})x}{2\pi\sin\frac{x}{2}}\,, \end{aligned}$$

we obtain that

$$s_n(f, x) = \int_{-\pi}^{\pi} f(y)D_n(x - y)\, dy \equiv (D_n \star f)(x)\,,$$

where we recall that the $\star$ operation is the convolution on the circle defined by

$$(f \star g)(x) = \int_{-\pi}^{\pi} f(y)g(x - y)\, dy\,.$$

Definition 2.7. The function

$$D_n(x) = \frac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}} \tag{2.2}$$

is called the ***Dirichlet kernel***.

In the following, we first consider an easier case $f \in \mathscr{C}^1(\mathbb{T})$; that is, f is 2π-periodic continuously differentiable on $\mathbb{R}$. We note that integration-by-parts formula provides that

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_{x=a}^{x=b} - \int_a^b f'(x)g(x)\,dx \qquad \forall\, f,g \in \mathscr{C}^1(\mathbb{T})\,.$$

The identity above allows us to prove the uniform convergence much more easily. We have the following

Theorem 2.8. *For any $f \in \mathscr{C}^1(\mathbb{T})$, $s_n(f,\cdot) = D_n \star f$ converges to f uniformly as $n \to \infty$.*

Proof. Since $\displaystyle\int_{\mathbb{T}} D_n(x-y)\,dy = 1$ for all $x \in \mathbb{T}$,

$$\begin{aligned} s_n(f,x) - f(x) = (D_n \star f - f)(x) &= \int_{\mathbb{T}} D_n(x-y)\big(f(y)-f(x)\big)dy \\ &= \int_{\mathbb{T}} D_n(y)\big(f(x-y)-f(x)\big)dy\,. \end{aligned}$$

We break the integral into two parts: one is the integral over $|y| \leqslant \delta$ and the other is the integral over $\delta < |y| \leqslant \pi$. Since $f \in \mathscr{C}^1(\mathbb{T})$,

$$|f(x-y)-f(x)| \leqslant \|f'\|_{L^\infty(\mathbb{T})}|y|\,;$$

thus using the fact that $\dfrac{x}{\sin x} \leqslant \dfrac{\pi}{2}$ for $0 < x < \dfrac{\pi}{2}$, we obtain that

$$\begin{aligned} &\Big|\int_{|y|\leqslant\delta} D_n(y)\big(f(x-y)-f(x)\big)dy\Big| \\ &\qquad \leqslant \int_{-\delta}^{\delta} \frac{|f(x-y)-f(x)|}{2\pi|\sin\frac{y}{2}|}dy \leqslant \frac{\|f'\|_{L^\infty(\mathbb{T})}}{2\pi}\int_{-\delta}^{\delta}\frac{y}{\sin\frac{y}{2}}dy \leqslant \|f'\|_{L^\infty(\mathbb{T})}\delta\,. \end{aligned} \tag{2.3}$$

Now we take care of the integral over $\delta < |y| \leqslant \pi$ by first looking at the integral over $\delta < y < \pi$. Integrating by parts,

$$\begin{aligned} &\int_\delta^\pi D_n(y)\big(f(x-y)-f(y)\big)\,dy = \frac{1}{2\pi}\int_\delta^\pi \sin\big(n+\frac{1}{2}\big)y\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}dy \\ &= -\frac{1}{2\pi}\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi} + \frac{1}{2\pi}\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}dy\,. \end{aligned}$$

For the first term on the right-hand side,

$$\Big|\frac{1}{2\pi}\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi}\Big| \leqslant \frac{2\|f\|_{L^\infty(\mathbb{T})}}{\pi n\sin\frac{\delta}{2}} \leqslant \frac{\|f\|_{L^\infty(\mathbb{T})}}{n\sin\frac{\delta}{2}} \qquad \forall\, x \in \mathbb{R}\,.$$

For the second term on the right-hand side,

$$
\begin{aligned}
&\Big|\frac{1}{2\pi}\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}dy\Big| \\
&\qquad \leqslant \frac{1}{2\pi}\Big[\Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f'(x-y)}{\sin\frac{y}{2}}dy\Big| + \Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{\cos\frac{y}{2}\big(f(x-y)-f(x)\big)}{\sin^2\frac{y}{2}}dy\Big|\Big] \\
&\qquad \leqslant \frac{1}{2\pi}\Big[\|f'\|_{L^\infty(\mathbb{T})}\frac{\pi-\delta}{\big(n+\frac{1}{2}\big)\sin\frac{\delta}{2}} + \|f\|_{L^\infty(\mathbb{T})}\frac{2(\pi-\delta)}{\big(n+\frac{1}{2}\big)\sin^2\frac{\delta}{2}}\Big] \leqslant \frac{\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}}\,.
\end{aligned}
$$

Similarly,

$$
\Big|\int_{-\pi}^{-\delta} D_n(y)\big(f(x-y)-f(x)\big)dy\Big| \leqslant \frac{\|f\|_{L^\infty(\mathbb{T})}}{n\sin\frac{\delta}{2}} + \frac{\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}}\,;
$$

thus for all $x \in \mathbb{R}$,

$$
\begin{aligned}
\big|s_n(f,x)-f(x)\big| &\leqslant \Big|\Big(\int_{-\delta}^{\delta} + \int_\delta^\pi + \int_{-\pi}^{-\delta}\Big)D_n(y)\big(f(x-y)-f(x)\big)\,dy\Big| \\
&\leqslant \|f'\|_{L^\infty(\mathbb{T})}\delta + \frac{2\|f\|_{L^\infty(\mathbb{T})}}{n\sin\frac{\delta}{2}} + \frac{2\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}} \leqslant \|f'\|_{L^\infty(\mathbb{T})}\delta + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}}
\end{aligned}
$$

Let $\varepsilon > 0$ be given. Choose a fixed $\delta > 0$ such that $\|f'\|_{L^\infty(\mathbb{T})}\delta < \dfrac{\varepsilon}{2}$. For this fixed δ, choose $N > 0$ such that

$$
\frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{N\sin^2\frac{\delta}{2}} < \frac{\varepsilon}{2}\,.
$$

Then if $n \geqslant N$ and $x \in \mathbb{R}$, we have

$$
\big|s_n(f,x)-f(x)\big| < \frac{\varepsilon}{2} + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}} \leqslant \frac{\varepsilon}{2} + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{N\sin^2\frac{\delta}{2}} < \varepsilon\,.
$$

This implies the uniform convergence of the sequence $\big\{s_n(f,\cdot)\big\}_{n=1}^\infty$ to f on $\mathbb{T}$. □

After showing the uniform convergence of the Fourier series representation of $\mathscr{C}^1$-functions, we next consider the convergence of the Fourier series representation of less regular functions. The functions of which we prove the convergence of the Fourier series representation belong to the so-called Hölder class continuous functions.

Definition 2.9. A function $f \in \mathscr{C}(\mathbb{T})$ is said to be ***Hölder continuous with exponent*** $\alpha \in (0,1]$, denoted by $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, if $\displaystyle\sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|^\alpha} < \infty$. Let $\|\cdot\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}$ be defined by

$$
\|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})} = \sup_{x\in\mathbb{T}}\big|f(x)\big| + \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{\big|f(x)-f(y)\big|}{|x-y|^\alpha}\,.
$$

Then $\|\cdot\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}$ is a norm on $\mathscr{C}^{0,\alpha}(\mathbb{T})$, and

$$
\mathscr{C}^{0,\alpha}(\mathbb{T}) = \big\{f \in \mathscr{C}(\mathbb{T})\,\big|\,\|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})} < \infty\big\}\,.
$$

In particular, when $\alpha = 1$, a function in $\mathscr{C}^{0,1}(\mathbb{T})$ is said to be Lipschitz continuous on $\mathbb{T}$; thus $\mathscr{C}^{0,1}(\mathbb{T})$ consists of Lipschitz continuous functions on $\mathbb{T}$.

The uniform convergence of $s_n(f,\cdot)$ to f for $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha \in (0,1)$ requires a lot more work. The idea is to estimate $\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})}$ in terms of the quantity $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})}$. Since $s_n(f,\cdot) \in \mathscr{P}_n(\mathbb{T})$, it is obvious that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant \big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} .$$

The goal is to show the inverse inequality

$$\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant C_n \inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \tag{2.4}$$

for some constant C_n, and pick a suitable $p \in \mathscr{P}_n(\mathbb{T})$ which gives a good upper bound for $\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})}$. The inverse inequality is established via the following

Proposition 2.10. *The Dirichlet kernel D_n satisfies that for all $n \in \mathbb{N}$,*

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx \leqslant 2 + \log n . \tag{2.5}$$

Proof. The validity of (2.5) for the case $n = 1$ is left to the reader, and we provide the proof for the case $n \geqslant 2$ here. Recall that $D_n(x) = \sum\limits_{k=-n}^{n} \dfrac{e^{ikx}}{2\pi} = \dfrac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}}$. Therefore,

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx = 2\int_0^{\pi} \big|D_n(x)\big| dx = \int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx + \int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx .$$

Since $|D_n(x)| \leqslant \lim\limits_{t\to 0^+} |D_n(t)| = \dfrac{2n+1}{2\pi}$ for all $0 < x \leqslant \dfrac{1}{n}$, the first integral can be estimated by

$$\int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx \leqslant \frac{1}{\pi}\frac{2n+1}{n} . \tag{2.6}$$

Since $\dfrac{2x}{\pi} \leqslant \sin x$ for $0 \leqslant x \leqslant \dfrac{\pi}{2}$, the second integral can be estimated by

$$\int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx \leqslant \int_{\frac{1}{n}}^{\pi} \frac{1}{x}\, dx = \log \pi + \log n . \tag{2.7}$$

We then conclude (2.5) from (2.6) and (2.7) by noting that $\log \pi + \dfrac{2n+1}{n\pi} \leqslant 2$ for all $n \geqslant 2$. □

Remark 2.11. A more subtle estimate can be done to show that

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx \geqslant c_1 + c_2 \log n \qquad \forall\, n \in \mathbb{N}$$

for some positive constants c_1 and c_2. Therefore, the integral of $|D_n|$ over $[-\pi,\pi]$ blows up as $n \to \infty$.

With the help of Proposition 2.10, we are able to prove the inverse inequality (2.4). The following theorem is a direct consequence of Proposition 2.10.

Theorem 2.12. *Let $f \in \mathscr{C}(\mathbb{T})$; that is, f is a continuous function with period 2π. Then*

$$\big\|f - s_n(f,\cdot)\big\|_\infty \leqslant (3 + \log n) \inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty . \tag{2.8}$$

Proof. For $n \in \mathbb{N}$ and $x \in \mathbb{T}$,

$$\big|s_n(f,x)\big| \leqslant \int_{-\pi}^{\pi} \big|D_n(y)\big|\big|f(x-y)\big|dy \leqslant (2+\log n)\|f\|_\infty .$$

Given $\varepsilon > 0$, let $p \in \mathscr{P}_n(\mathbb{T})$ such that

$$\|f-p\|_\infty \leqslant \inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty + \varepsilon.$$

Then by the fact that $s_n(p,x) = p(x)$ if $p \in \mathscr{P}_n(\mathbb{T})$, we obtain that

$$\begin{aligned}
\big\|f - s_n(f,\cdot)\big\|_\infty &\leqslant \big\|f-p\big\|_\infty + \big\|p - s_n(f,\cdot)\big\|_\infty \leqslant \big\|f-p\big\|_\infty + \big\|s_n(f-p,\cdot)\big\|_\infty \\
&\leqslant \big\|f-p\big\|_\infty + (2+\log n)\|f-p\|_\infty \\
&\leqslant (3+\log n)\Big[\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty + \varepsilon\Big],
\end{aligned}$$

and (2.8) is obtained by passing $\varepsilon \to 0$. □

Having established Theorem 2.12, the study of the uniform convergence of $s_n(f,\cdot)$ to f then amounts to the study of the quantity $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty$. In Exercise Problem 3, the reader is asked to show that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty \leqslant \frac{1+2\log n}{2n}\|f\|_{\mathscr{C}^{0,1}(\mathbb{T})};$$

thus by Theorem 2.12, $s_n(f,\cdot)$ converges to f uniformly as $n \to \infty$ if $f \in \mathscr{C}^{0,1}(\mathbb{T})$.

The estimate of $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty$ for $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, where $\alpha \in (0,1)$, is more difficult, and requires a clever choice of p. We begin with the following

Lemma 2.13. *If f is a continuous function on $[a,b]$, then for all $\delta_1, \delta_2 > 0$,*

$$\sup_{|x-y|\leqslant\delta_1} \big|f(x)-f(y)\big| \leqslant \Big(1+\frac{\delta_1}{\delta_2}\Big) \sup_{|x-y|\leqslant\delta_2} \big|f(x)-f(y)\big| .$$

The proof of Lemma 2.13 is not very difficult, and is left to the readers.

Now we are in position to prove the theorem due to D. Jackson.

Theorem 2.14 (Jackson)**.** *There exists a constant $C > 0$ such that*

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant C \sup_{|x-y|\leqslant\frac{1}{n}} |f(x)-f(y)| \qquad \forall\, f \in \mathscr{C}(\mathbb{T}) .$$

Proof. Let $p(x) = 1 + c_1\cos x + \cdots + c_n \cos nx$ be a positive trigonometric function of degree n with coefficients $\{c_i\}_{i=1}^n$ determined later. Define an operator K on $\mathscr{C}(\mathbb{T})$ by

$$\mathrm{K}f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} p(y)f(x-y)\,dy .$$

Then $\mathrm{K}f \in \mathscr{P}_n(\mathbb{T})$. Lemma 2.13 then implies

$$\begin{aligned}
\big|\mathrm{K}f(x) - f(x)\big| &\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} p(y)\big|f(x-y)-f(x)\big|dy \\
&\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} p(y)\big(1+n|y|\big) \sup_{|x-y|\leqslant\frac{1}{n}} \big|f(x)-f(y)\big|dy \\
&= \Big[1+\frac{n}{2\pi}\int_{-\pi}^{\pi} |y|p(y)\,dy\Big] \sup_{|x-y|\leqslant\frac{1}{n}} \big|f(x)-f(y)\big| .
\end{aligned}$$

Since $y^2 \leqslant \dfrac{\pi^2}{2}(1-\cos y)$ for $y \in [-\pi,\pi]$, by Hölder's inequality we find that

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}|y|p(y)\,dy &\leqslant \Big[\frac{1}{2\pi}\int_{-\pi}^{\pi}y^2p(y)\,dy\Big]^{\frac12}\Big[\frac{1}{2\pi}\int_{-\pi}^{\pi}p(y)\,dy\Big]^{\frac12}\\ &\leqslant \Big[\frac{\pi}{4}\int_{-\pi}^{\pi}(1-\cos y)p(y)\,dy\Big]^{\frac12} = \frac{\pi}{2}\sqrt{2-c_1}\,.\end{aligned}$$

Therefore,

$$\|\mathrm{K}f-f\|_\infty \leqslant \Big(1+\frac{n\pi}{2}\sqrt{2-c_1}\Big)\sup_{|x-y|\leqslant\frac1n}\big|f(x)-f(y)\big|\,.$$

To conclude the theorem, we need to show that the number $n\sqrt{2-c_1}$ can be made bounded by choosing p properly. Nevertheless, let

$$\begin{aligned}p(x) &= c\Big|\sum_{k=0}^n \sin\frac{(k+1)\pi}{n+2}e^{ikx}\Big|^2 = c\sum_{k=0}^n\sum_{\ell=0}^n \sin\frac{(k+1)\pi}{n+2}\sin\frac{(\ell+1)\pi}{n+2}e^{i(k-\ell)x}\\ &= c\sum_{k=0}^n \sin^2\frac{(k+1)\pi}{n+2} + 2c\sum_{\substack{k,\ell=0\\k>\ell}}^n \sin\frac{(k+1)\pi}{n+2}\sin\frac{(\ell+1)\pi}{n+2}\cos(k-\ell)x\end{aligned}$$

and choose c so that $p(x) = 1 + c_1\cos x + \cdots + c_n\cos nx$. Then

$$\begin{aligned}c^{-1} &= \sum_{k=0}^n \sin^2\frac{(k+1)\pi}{n+2} = \frac12\sum_{k=0}^n\Big[1-\cos\frac{2(k+1)\pi}{n+2}\Big]\\ &= \frac{n+1}{2} - \frac{\sin\frac{(2n+3)\pi}{n+2}-\sin\frac{\pi}{n+2}}{4\sin\frac{\pi}{n+2}} = \frac{n+2}{2}\,,\end{aligned}$$

and

$$\begin{aligned}c_1 &= 2c\sum_{k=1}^n \sin\frac{(k+1)\pi}{n+2}\sin\frac{k\pi}{n+2} = c\sum_{k=1}^n\Big[\cos\frac{\pi}{n+2}-\cos\frac{(2k+1)\pi}{n+2}\Big]\\ &= c\Big[n\cos\frac{\pi}{n+2} - \frac{\sin\frac{(2n+2)\pi}{n+2}-\sin\frac{2\pi}{n+2}}{2\sin\frac{\pi}{n+2}}\Big]\\ &= c\Big[n\cos\frac{\pi}{n+2} + \frac{\sin\frac{2\pi}{n+2}}{\sin\frac{\pi}{n+2}}\Big]\\ &= c(n+2)\cos\frac{\pi}{n+2} = 2\cos\frac{\pi}{n+2}.\end{aligned}$$

As a consequence,

$$\begin{aligned}n\sqrt{2-c_1} &= n\Big(2-2\cos\frac{\pi}{n+2}\Big)^{\frac12} = 2n\sin\frac{\pi}{2(n+2)}\\ &= 2(n+2)\sin\frac{\pi}{2(n+2)} - 4\sin\frac{\pi}{2(n+2)}\\ &= \pi\frac{2(n+2)}{\pi}\sin\frac{\pi}{2(n+2)} - 4\sin\frac{\pi}{2(n+2)}\end{aligned}$$

which is bounded by π; thus

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})} \leqslant \|\mathrm{K}f-f\|_{L^\infty(\mathbb{T})} \leqslant \big(1+\frac{\pi^2}{2}\big)\sup_{|x-y|\leqslant\frac1n}\big|f(x)-f(y)\big|\,. \qquad \square$$

Finally, since $\lim_{n\to\infty} n^{-\alpha}\log n = 0$ for all $\alpha \in (0,1]$, we conclude the following

Theorem 2.15. *For all $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha \in (0,1]$, $s_n(f,\cdot) = D_n \star f$ converges to f uniformly as $n \to \infty$.*

Remark 2.16. The converse of Theorem 2.14 is the Bernstein theorem which states that if f is a 2π-periodic function such that for some constant C (independent of n) and $\alpha \in (0,1)$,

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant C n^{-\alpha} \tag{2.9}$$

for all $n \in \mathbb{N}$, then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$. In other words, (2.9) is an equivalent condition to the Hölder continuity with exponent α of 2π-periodic continuous functions. One way of proving the Bernstein theorem can be found in Exercise Problem 4.

2.3 Cesàro Mean of Fourier Series

While Theorem 1.25 shows that the collection of trigonometric polynomials

$$\left\{ \frac{c_0}{2} + \sum_{k=1}^{n} (c_k \cos kx + s_k \sin kx) \,\middle|\, \{c_k\}_{k=0}^{n}, \{s_k\}_{k=1}^{n} \subseteq \mathbb{R} \right\}$$

is dense in $\mathscr{C}(\mathbb{T})$, Theorem 2.15 only implies the uniform convergence of the Fourier series representation of Hölder continuous functions. Since the Fourier coefficients $\{c_k\}_{k=0}^{n}$ and $\{s_k\}_{k=1}^{n}$ are independent of the order of approximation n, as we discussed in the beginning of this chapter we do not expect that $s_n(f,\cdot)$ uniformly to f on $[-\pi,\pi]$ for general $f \in \mathscr{C}(\mathbb{T})$. To approximate continuous functions uniformly, the coefficients of the trigonometric polynomials should depend on the order of approximation.

The motivation of the discussion below is due to the following observation. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence. Define a new sequence $\{b_n\}_{n=1}^{\infty}$, called the ***Cesàro mean*** of the sequence $\{a_k\}_{k=1}^{\infty}$, by

$$b_n = \frac{a_1 + \cdots + a_n}{n} = \frac{1}{n}\sum_{k=1}^{n} a_k \,.$$

If $\{a_k\}_{k=1}^{\infty}$ converges to a, then $\{b_n\}_{n=1}^{\infty}$ converges to a as well. Even though the convergence of a sequence cannot be guaranteed by the convergence of its Cesàro mean, it is worthwhile investigating the convergence behavior of the Cesàro mean.

Let $\sigma_n(f,\cdot)$ denote the Cesàro mean of the Fourier series representation of f given by

$$\sigma_n(f,\cdot) \equiv \frac{1}{n+1}\sum_{k=0}^{n} s_k(f,\cdot) = \frac{1}{n+1}\sum_{k=0}^{n} (D_k \star f) = \Big(\frac{1}{n+1}\sum_{k=0}^{n} D_k\Big) \star f \,.$$

We note that the coefficients of the Cesàro mean $\sigma_n(f,\cdot)$ depend on the order of approximation n since

$$\sigma_n(f,x) = \frac{c_0}{2} + \sum_{k=1}^{n}\Big(\underbrace{\frac{n+1-k}{n+1} c_k}_{\equiv c_k^{(n)}} \cos kx + \underbrace{\frac{n+1-k}{n+1} s_k}_{\equiv s_k^{(n)}} \sin kx \Big) \,.$$

Recall that $D_k(x) = \dfrac{\sin(k+\frac{1}{2})x}{2\pi\sin\frac{x}{2}}$. By the product-to-sum（積化和差）formula, we find that

$$\begin{aligned}\sum_{k=0}^{n} D_k(x) &= \frac{1}{2\pi\sin\frac{x}{2}}\sum_{k=0}^{n}\sin(k+\frac{1}{2})x = \frac{1}{4\pi\sin^2\frac{x}{2}}\sum_{k=0}^{n} 2\sin\frac{x}{2}\sin(k+\frac{1}{2})x \\ &= \frac{1}{4\pi\sin^2\frac{x}{2}}\sum_{k=0}^{n}\big(\cos(k-1)x - \cos(k+1)x\big) \\ &= \frac{1}{4\pi\sin^2\frac{x}{2}}\big(1-\cos(n+1)x\big) = \frac{\sin^2\frac{n+1}{2}x}{2\pi\sin^2\frac{x}{2}}\,.\end{aligned}$$

This induces the following

Definition 2.17. The ***Fejér kernel*** is the Cesàro mean of the Dirichlet kernel given by

$$F_n(x) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(x) = \frac{1}{2\pi(n+1)}\frac{\sin^2\frac{(n+1)x}{2}}{\sin^2\frac{x}{2}}\,.$$

We note that $\sigma_n(f,\cdot) = F_n \star f$, where $F_n \geqslant 0$ and has the property that $\displaystyle\int_{-\pi}^{\pi} F_n(x)\,dx = 1$ (since the integral of the Dirichlet kernel is 1). Moreover, for any $\delta > 0$,

$$\limsup_{n\to\infty}\int_{\delta\leqslant|x|\leqslant\pi} F_n(x)\,dx \leqslant \limsup_{n\to\infty}\int_{\delta\leqslant|x|\leqslant\pi}\frac{1}{2\pi(n+1)\sin^2\frac{\delta}{2}}\,dx = 0\,. \tag{2.10}$$

Therefore, $\{F_n\}_{n=1}^{\infty}$ is an approximation of the identity, and Theorem 1.23 implies the following

Theorem 2.18. *For any $f \in \mathscr{C}(\mathbb{T})$, the Cesàro mean $\big\{\sigma_n(f,\cdot)\big\}_{n=1}^{\infty}$ of the Fourier series representation of f converges uniformly to f.*

2.4 Convergence of Fourier Series for Functions with Jump Discontinuity

In previous sections we discussed the convergence of the Fourier series representation of continuous functions. However, since the Fourier series can be defined for bounded Riemann integrable functions, it is natural to ask what happen if the function under consideration is not continuous. We note that in this case we cannot apply Theorem 1.25 at all so no uniform convergence is expected.

In this section, we focus on the convergence behavior of Fourier series representation of functions with only jump discontinuities.

Definition 2.19. A function $f : [-\pi,\pi] \to \mathbb{R}$ is said to have jump discontinuity at $a \in (-\pi,\pi)$ if

1. $\lim\limits_{x\to a^+} f(x)$ and $\lim\limits_{x\to a^-} f(x)$ both exist.
2. $\lim\limits_{x\to a^+} f(x) \neq \lim\limits_{x\to a^-} f(x)$.

Now suppose that $f : [-\pi,\pi] \to \mathbb{R}$ is piecewise Hölder continuous with exponent α and the discontinuities of f are all jump discontinuities. In other words, f has finitely many jump discontinuities $\{a_1,\cdots,a_m\}$ in $(-\pi,\pi)$, and $f \in \mathscr{C}^{0,\alpha}((a_j,a_{j+1}))$ for all $j \in \{0,\cdots,m\}$, where $a_0 = -\pi$ and $a_{m+1} = \pi$. Let $f(a_j^+) = \lim\limits_{x\to a_j^+} f(x)$, $f(a_j^-) = \lim\limits_{x\to a_j^-} f(x)$, and define $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(x) = \frac{1}{2\pi}(x-\pi) \qquad \forall\, x \in [0, 2\pi) \tag{2.11}$$

and $\phi(x+2\pi) = \phi(x)$ for all $x \in \mathbb{R}$. Since f has jump discontinuities at $\{a_1, \cdots, a_m\}$, with a_0^- denoting a_{m+1}^- the function $g : [-\pi, \pi) \to \mathbb{R}$ defined by

$$g(x) \equiv \begin{cases} f(x) + \displaystyle\sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big)\phi(x - a_j) & \text{if } x \neq a_k \text{ for all } k, \\ \dfrac{f(a_k^+) + f(a_k^-)}{2} + \displaystyle\sum_{\substack{0 \leqslant j \leqslant m \\ j \neq k}} \big(f(a_j^+) - f(a_j^-)\big)\phi(a_k - a_j) & \text{if } x = a_k \text{ for some } k, \end{cases} \tag{2.12}$$

is Hölder continuous with exponent α and $g(a_0^+) = g(a_0^-)$. Let G be the 2π-periodic extension of g; that is, $G = g$ on $[-\pi, \pi)$ and $G(x+2\pi) = G(x)$ for all $x \in \mathbb{R}$. Then $G \in \mathscr{C}^{0,\alpha}(\mathbb{T})$; thus Theorem 2.15 implies that $s_n(G, \cdot) \to G$ uniformly on $\mathbb{R}$. In particular, $s_n(g, \cdot) \to g$ uniformly on $[-\pi, \pi)$.

Using the identity

$$\int_{-\pi}^{\pi} \phi(x-a) e^{-ikx}\, dx = e^{-ika} \int_{-\pi}^{\pi} \phi(x) e^{-ikx}\, dx = \widehat{\phi}_k e^{-ika}\,,$$

we obtain that

$$s_n(\phi(\cdot - a), x) = \sum_{k=-n}^{n} \widehat{\phi}_k e^{ik(x-a)} = s_n(\phi, x-a)\,; \tag{2.13}$$

thus (2.12) implies that the Fourier series representation of f is given by

$$\begin{aligned} s_n(f, x) &= s_n(g, x) - \sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big) s_n(\phi(\cdot - a_j), x) \\ &= s_n(g, x) - \sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big) s_n(\phi, x - a_j)\,. \end{aligned} \tag{2.14}$$

Therefore, to understand the convergence of the Fourier series representation of f, without loss of generality it suffices to consider the convergence of $s_n(\phi, \cdot)$.

2.4.1 Uniform convergence on compact subsets

In this sub-section, we show that the Fourier series representation of a piecewise Hölder continuous function whose discontinuities are all jump discontinuities converges uniformly on each compact subset containing no jump discontinuities.

Based on the discussion above, we first study the convergence of $s_n(\phi, \cdot)$. Since ϕ is an odd function, for $k \in \mathbb{N}$,

$$\begin{aligned} s_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} \phi(x) \sin kx\, dx = \frac{1}{\pi^2} \int_0^{\pi} (x - \pi) \sin kx\, dx \\ &= \frac{1}{\pi^2} \Big[\frac{-(x-\pi)\cos kx}{k} \Big|_{x=0}^{x=\pi} + \int_0^{\pi} \frac{\cos kx}{k}\, dx \Big] = -\frac{1}{\pi k}\,. \end{aligned}$$

Therefore, the Fourier series representation of ϕ is given by

$$s_n(\phi, x) = -\frac{1}{\pi} \sum_{k=1}^{n} \frac{\sin kx}{k}\,. \tag{2.15}$$

Lemma 2.20. *The series* $\sum\limits_{k=1}^{\infty} \dfrac{\sin kx}{k}$ *converges uniformly on* $[-\pi,-\delta] \cup [\delta,\pi]$ *for all* $0<\delta<\pi$.

Proof. Let $S_n(x)$ denote the sum $\sum\limits_{k=1}^{n} \sin kx$. Using the identity

$$\sum_{k=1}^{n} \sin kx = \frac{\cos(n+\frac{1}{2})x - \cos\frac{x}{2}}{2\sin\frac{x}{2}} \qquad \forall\, x \in [-\pi,-\delta] \cup [\delta,\pi]\,,$$

we find that $|S_n| \leqslant M < \infty$ for some fixed constant M. For $m>n$,

$$\begin{aligned}\sum_{k=n+1}^{m} \frac{1}{k}\sin kx &= \frac{1}{m}(S_m - S_{m-1}) + \frac{1}{m-1}(S_{m-1}-S_{m-2}) + \cdots + \frac{1}{n+1}(S_{n+1}-S_n)\\ &= \frac{S_m}{m} - \frac{S_n}{n+1} + \frac{1}{m(m-1)}S_{m-1} + \frac{1}{(m-1)(m-2)}S_{m-2} + \cdots \frac{1}{(n+1)n}S_{n+1}\,;\end{aligned}$$

thus

$$\Big|\sum_{k=n+1}^{m} \frac{1}{k}\sin kx\Big| \leqslant M\Big(\frac{1}{m} + \frac{1}{n+1} + \sum_{k=n+1}^{m} \frac{1}{k(k-1)}\Big) \leqslant 2M\Big(\frac{1}{m}+\frac{1}{n}\Big)\,.$$

Since the right-hand side converges to 0 as $n,m\to\infty$, the Cauchy criteria implies that the series

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k}$$

converges uniformly on $[-\pi,-\delta]\cup[\delta,\pi]$. □

Lemma 2.20 provides the uniform convergence of $s_n(\phi,\cdot)$ in $[-\pi,-\delta]\cup[\delta,\pi]$. To see the limit is exactly ϕ, we consider an anti-derivative Φ of ϕ and establish that $\Phi' = s(\phi,\cdot)$.

Let $\Psi:\mathbb{R}\to\mathbb{R}$ be 2π-periodic and $\Psi(x) = \dfrac{x^2}{4\pi}$ for $x\in[-\pi,\pi]$. Then $\Psi \in \mathscr{C}^{0,1}(\mathbb{T})$ is an even function and the Fourier coefficients of Ψ is

$$\widehat{\Psi}_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{x^2}{4\pi}\,dx = \frac{\pi}{12}$$

and for $k\neq 0$,

$$\widehat{\Psi}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{x^2}{4\pi} e^{-ikx}\,dx = \frac{1}{8\pi^2}\int_{-\pi}^{\pi} x^2(\cos kx + i\sin kx)\,dx = \frac{(-1)^k}{2k^2\pi}\,.$$

Therefore, using (2.13) we find that the Fourier series representation of $\Phi \equiv \Psi(\cdot-\pi)$ is

$$s(\Phi,x) = s(\Psi,x-\pi) = \frac{\pi}{12} + \sum_{k\in\mathbb{Z},k\neq 0} \widehat{\Psi}_k e^{ik(x-\pi)} = \frac{\pi}{12} + \frac{1}{2\pi}\sum_{k\in\mathbb{Z},k\neq 0}\frac{e^{ikx}}{k^2} = \frac{\pi}{12} + \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\cos kx}{k^2}\,.$$

Since $\Phi\in\mathscr{C}^{0,1}(\mathbb{T})$, $s_n(\Phi,\cdot)$ converges uniformly to Φ on $\mathbb{R}$. Moreover, $s_n(\Phi,\cdot)' = s_n(\phi,\cdot)$ which converges uniformly on $[-\pi,-\delta]\cup[\delta,\pi]$. Therefore, Theorem 1.5 implies that $s(\phi,\cdot)$, the uniform limit of $s_n(\phi,\cdot)$, must equal Φ' on $[-\pi,-\delta]\cup[\delta,\pi]$. Finally, we note that $\phi=\Phi'$ on $[-\pi,-\delta]\cup[\delta,\pi]$, so we establish that $s_n(\phi,\cdot)\to\phi$ uniformly on $[-\pi,-\delta]\cup[\delta,\pi]$.

The uniform convergence of $s_n(\phi,\cdot)$ to ϕ on $[-\pi,-\delta]\cup[\delta,\pi]$ for all $0<\delta<\pi$ implies the following

Theorem 2.21. *Let $f : (-\pi, \pi) \to \mathbb{R}$ be piecewise Hölder continuous with exponent $\alpha \in (0, 1]$ such that its set of discontinuities consists of only jump discontinuities. If f is continuous on (a, b), then $s_n(f, \cdot) \to f$ uniformly on any compact subsets of (a, b).*

By Remark 2.3, we can also conclude the following

Corollary 2.22. *Let $f : (-L, L) \to \mathbb{R}$ be piecewise Hölder continuous with exponent $\alpha \in (0, 1]$ such that f has only jump discontinuities. If f is continuous on (a, b), then $s_n(f, \cdot) \to f$ uniformly on any compact subsets of (a, b) (where the Fourier series representation of f is given in Remark 2.3). In particular, $s_n(f, x_0) \to f(x_0)$ if f is continuous at x_0. In other words, the Fourier series representation of f converges pointwise to f except the jump discontinuities.*

2.4.2 Gibbs phenomenon

In this sub-section, we show that the Fourier series evaluated at the jump discontinuity converges to the average of the limits from the left and the right. Moreover, the convergence of the Fourier series is never uniform in the domain including these jump discontinuities due to the famous Gibbs phenomenon: near the jump discontinuity the maximum difference between the limit of the Fourier series and the function itself is at least 8% of the jump. To be more precise, we have the following

Theorem 2.23. *Let $f : \mathbb{R} \to \mathbb{R}$ be a $2L$-periodic piecewise Hölder continuous function with exponent $\alpha \in (0, 1]$ such that its set of discontinuities consists of only jump discontinuities. Suppose that at some point x_0 the limit from the left $f(x_0^-)$ and the limit from the right $f(x_0^+)$ of the function f exist and differ by a non-zero gap a :*

$$f(x_0^+) - f(x_0^-) = a \neq 0,$$

then there exists a constant $c > 0$, independent of f, x_0 and L (in fact, $c = \dfrac{1}{\pi}\displaystyle\int_0^\pi \frac{\sin x}{x}\,dx - \frac{1}{2} \approx 0.089490$), such that

$$\lim_{n\to\infty} s_n\big(f, x_0 + \frac{L}{n}\big) = f(x_0^+) + ca\,, \tag{2.16a}$$

$$\lim_{n\to\infty} s_n\big(f, x_0 - \frac{L}{n}\big) = f(x_0^-) - ca\,. \tag{2.16b}$$

Moreover,

$$\lim_{n\to\infty} s_n(f, x_0) = \frac{f(x_0^+) + f(x_0^-)}{2}\,. \tag{2.17}$$

Proof. By Remark 2.3, W.L.O.G. we can assume that $L = \pi$. Let $\{a_1, \cdots, a_m\} \subseteq (-\pi, \pi)$ be the collection of jump discontinuities of f, $a_0 = -\pi$, $a_{m+1} = \pi$, and define g by (2.12), where a_0^- is used to denote a_{m+1}^-. Then $g \in \mathscr{C}^{0,\alpha}(\mathbb{T})$. Since x_0 is a jump discontinuity of f, $x_0 = a_k$ for some $k \in \{0, 1, \cdots, m\}$. Therefore, by the fact that ϕ is continuous at $x_0 - a_j$ if $j \neq k$ and $s_n(\phi, 0) = 0$ for all $n \in \mathbb{N}$, Corollary 2.22 implies that

$$\begin{aligned}
&\sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big) \lim_{n\to\infty} s_n(\phi, x_0 - a_j) \\
&\quad = \sum_{\substack{0\leqslant j\leqslant m \\ j\neq k}} \big(f(a_j^+) - f(a_j^-)\big) \lim_{n\to\infty} s_n(\phi, x_0 - a_j) = \sum_{\substack{0\leqslant j\leqslant m \\ j\neq k}} \big(f(a_j^+) - f(a_j^-)\big)\phi(x_0 - a_j)\,.
\end{aligned}$$

On the other hand,

$$\lim_{n\to\infty} s_n(g,x_0) = g(x_0) = \frac{f(x_0^+)+f(x_0^-)}{2} + \sum_{\substack{0\leqslant j\leqslant m\\ j\neq k}} \big(f(a_j^+)-f(a_j^-)\big)\phi(x_0-a_j)\,.$$

Identity (2.17) is then concluded using (2.14).

Now we focus on (2.16a). Since $g\in\mathscr{C}^{0,\alpha}(\mathbb{T})$, $s_n(g,\cdot)\to g$ uniformly on $\mathbb{T}$; thus

$$\lim_{n\to\infty} s_n\big(g,x_0+\frac{\pi}{n}\big) = g(x_0)\,.$$

Similarly, since $s_n(\phi,\cdot)\to\phi$ uniformly on $[-\pi,-\delta]\cup[\delta,\pi]$ for all $\delta>0$, if $j\neq k$,

$$\lim_{n\to\infty} s_n\big(\phi,x_0+\frac{\pi}{n}-a_j\big) = \phi(x_0-a_j)\,.$$

On the other hand,

$$s_n\big(\phi,\frac{\pi}{n}\big) = -\sum_{k=1}^n \frac{1}{\pi k}\sin\frac{k\pi}{n} = -\frac{1}{\pi}\sum_{k=1}^n \frac{n}{k\pi}\sin\frac{k\pi}{n}\frac{\pi}{n} \to -\frac{1}{\pi}\int_0^\pi \frac{\sin x}{x}dx \equiv -\big(c+\frac{1}{2}\big)\,.$$

As a consequence,

$$\begin{aligned}\lim_{n\to\infty} s_n\big(f,x_0+\frac{\pi}{n}\big) &= \lim_{n\to\infty}\Big[s_n\big(g,x_0+\frac{\pi}{n}\big) - \sum_{j=0}^m \big(f(a_j^+)-f(a_j^-)\big)s_n\big(\phi,x_0+\frac{\pi}{n}-a_j\big)\Big]\\ &= g(x_0) - \sum_{\substack{0\leqslant j\leqslant m\\ j\neq k}} \big(f(a_j^+)-f(a_j^-)\big)\phi(x_0-a_j) + \big(c+\frac{1}{2}\big)\big(f(x_0^+)-f(x_0^-)\big)\\ &= f(x_0^+) + c\big(f(x_0^+)-f(x_0^-)\big)\,.\end{aligned}$$

Identity (2.16b) can be proved in the same fashion, and is left as an exercise. □

Remark 2.24. Let f be a function given in Theorem 2.23, x_0 be a jump discontinuity of f, and $I=(x_0,x_0+r)$ for some $r>0$. By the definition of the right limit, there exists $0<\delta<r$ such that

$$\big|f(x)-f(x_0^+)\big| < \frac{c|a|}{2} \qquad \forall\, x\in(x_0,x_0+\delta)\,.$$

Choose $N>0$ such that $\dfrac{L}{N}<\delta$. Then $x_0+\dfrac{L}{N}\in(x_0,x_0+\delta)$ for all $n\geqslant N$; thus if $n\geqslant N$,

$$\begin{aligned}\sup_{x\in I}\big|s_n(f,x)-f(x)\big| &\geqslant \big|s_n(f,x_0+\frac{L}{N}) - f(x_0+\frac{L}{N})\big|\\ &\geqslant \big|s_n(f,x_0+\frac{L}{N}) - f(x_0^+)\big| - \big|f(x_0+\frac{L}{N})-f(x_0^+)\big|\\ &\geqslant \big|s_n(f,x_0+\frac{L}{N}) - f(x_0^+)\big| - \frac{c|a|}{2}\end{aligned}$$

which implies that

$$\liminf_{n\to\infty}\big|s_n(f,x)-f(x)\big| \geqslant c|a| - \frac{c|a|}{2} = \frac{c|a|}{2}\,.$$

Therefore, $\big\{s_n(f,\cdot)\big\}_{n=1}^\infty$ does not converge uniformly (to f) on I, while Corollary 2.22 shows that $\big\{s_n(f,\cdot)\big\}_{n=1}^\infty$ converges pointwise to f. Similarly, if x_0 is a jump discontinuity of f and f is continuous on (x_0-r,x_0) for some $r>0$, then $\big\{s_n(f,\cdot)\big\}_{n=1}^\infty$ converge pointwise but not uniformly on (x_0-r,x_0).

For a function f given in Theorem 2.23, let $\widetilde{f}$ be defined by

$$\widetilde{f}(x)=\left\{\begin{array}{cl} f(x) & \text{if } f \text{ is continuous at } x\,,\\ \dfrac{f(x^+)+f(x^-)}{2} & \text{if } x \text{ is a jump discontinuity of } f\,. \end{array}\right.$$

Then $s_n(\widetilde{f},\cdot) = s_n(f,\cdot)$ for all $n \in \mathbb{N}$, and Corollary 2.22 and Theorem 2.23 together imply that $\{s_n(f,\cdot)\}_{n=1}^\infty$ converges pointwise to $\widetilde{f}$. However, the discussion above shows that $\{s_n(f,\cdot)\}_{n=1}^\infty$ cannot converge uniformly on I as long as I contains jump discontinuities of f.

2.5 The Inner-Product Point of View

除了逐點收斂或均勻收斂的觀點之外，還有一個更自然（就數學而言）的觀點可以用來看 Fourier series。我們可以把定義在 $[-\pi,\pi]$ 的所有 Riemann integrable 函數所形成的集合看成一個向量空間，然後在上面定義一個內積的結構。一個可積分函數（也可視為一個向量）的 Fourier series representation 可以看成這個向量在一組正交基底向量的線性組合。

Let $L^2(\mathbb{T})$ denote the collection of Riemann measurable, square integrable function over $[-\pi,\pi]$ modulo the relation that $f \sim g$ if $f-g=0$ except on a set of measure zero (or $f=g$ almost everywhere). In other words,

$$L^2(\mathbb{T})=\Big\{f:[-\pi,\pi)\to\mathbb{C}\,\Big|\int_{[-\pi,\pi)}|f(x)|^2\,dx<\infty\Big\}/\sim\ .$$

Here again we **abuse** the use of notation $L^2(\mathbb{T})$ for that it indeed denotes a more general space. We also note that the domain $[-\pi,\pi)$ can be replaced by any intervals with $-\pi$, π as end-points for we can easily modify functions defined on those domains to functions defined on $[-\pi,\pi)$ without changing the Riemann measurability and the square integrability.

Define a bilinear function $\langle\cdot,\cdot\rangle$ on $L^2(\mathbb{T})\times L^2(\mathbb{T})$ by

$$\langle f,g\rangle=\frac{1}{2\pi}\int_{[-\pi,\pi)} f(x)\overline{g(x)}\,dx\,.$$

Then $\langle\cdot,\cdot\rangle$ is an inner product on $L^2(\mathbb{T})$. Indeed, if f,g belong to $L^2(\mathbb{T})$, then the product $f\bar{g}$ is also Riemann measurable, and the Cauchy-Schwartz inequality as well as the monotone convergence theorem imply that

$$\begin{aligned}\big|\langle f,g\rangle\big| &\leqslant \lim_{k\to\infty}\frac{1}{2\pi}\int_{[-\pi,\pi)}\big|(f\wedge k)(x)\big|\big|(g\wedge k)(x)\big|\,dx\\ &\leqslant \lim_{k\to\infty}\frac{1}{2\pi}\Big(\int_{[-\pi,\pi)}\big|(f\wedge k)(x)\big|^2\,dx\Big)^{\frac12}\Big(\int_{[-\pi,\pi)}\big|(g\wedge k)(x)\big|^2\,dx\Big)^{\frac12}\\ &=\frac{1}{2\pi}\Big(\int_{[-\pi,\pi)}\big|f(x)\big|^2\,dx\Big)^{\frac12}\Big(\int_{[-\pi,\pi)}\big|g(x)\big|^2\,dx\Big)^{\frac12}<\infty\,;\end{aligned}$$

thus the definition of the inner product $\langle\cdot,\cdot\rangle$ given above is well-defined. The norm induced by the inner product above is denoted by $\|\cdot\|_{L^2(\mathbb{T})}$ or $\|\cdot\|_{L^2(-\pi,\pi)}$.

For $k\in\mathbb{Z}$, define $\mathbf{e}_k:[-\pi,\pi]\to\mathbb{C}$ by $\mathbf{e}_k(x)=e^{ikx}$. Then $\{\mathbf{e}_k\}_{k=-\infty}^\infty$ is an orthonormal set in $L^2(\mathbb{T})$ since

$$\langle\mathbf{e}_k,\mathbf{e}_\ell\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{ikx}e^{-i\ell x}\,dx=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i(k-\ell)x}\,dx=\left\{\begin{array}{ll}1 & \text{if } k=\ell\,,\\ 0 & \text{if } k\neq\ell\,.\end{array}\right.$$

Let $\mathcal{V}_n = \operatorname{span}(\mathbf{e}_{-n}, \mathbf{e}_{-n+1}, \cdots, \mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n) = \Big\{ \sum\limits_{k=-n}^{n} a_k \mathbf{e}_k \,\Big|\, \{a_k\}_{k=-n}^{n} \subseteq \mathbb{C} \Big\}$. For each vector $f \in L^2(\mathbb{T})$, the orthogonal projection of f onto $\mathcal{V}_n$ is, conceptually, given by

$$\sum_{k=-n}^{n} \langle f, \mathbf{e}_k \rangle \mathbf{e}_k = \sum_{k=-n}^{n} \Big(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx}\, dx \Big) \mathbf{e}_k = \sum_{k=-n}^{n} \widehat{f}_k \mathbf{e}_k \,.$$

By the definition of $\mathbf{e}_k$, we obtain that the projection of f on $\mathcal{V}_n$ is exactly $s_n(f, \cdot)$. We also note that $\mathcal{V}_n = \mathscr{P}_n(\mathbb{T})$.

Now we prove that $s_n(f, \cdot)$ is exactly the orthogonal projection of f onto $\mathcal{V}_n = \mathscr{P}_n(\mathbb{T})$.

Proposition 2.25. *Let $f \in L^2(\mathbb{T})$. Then*

$$\langle f - s_n(f, \cdot), p \rangle = 0 \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}) \,.$$

Proof. Let $p \in \mathscr{P}_n(\mathbb{T})$. Then $p = s_n(p, \cdot)$; thus

$$\begin{aligned} \langle f - s_n(f, \cdot), p \rangle &= \langle f, p \rangle - \langle s_n(f, \cdot), p \rangle = \Big\langle f, \sum_{k=-n}^{n} \widehat{p}_k \mathbf{e}_k \Big\rangle - \Big\langle \sum_{k=-n}^{n} \widehat{f}_k \mathbf{e}_k, p \Big\rangle \\ &= \sum_{k=-n}^{n} \overline{\widehat{p}_k} \langle f, \mathbf{e}_k \rangle - \sum_{k=-n}^{n} \widehat{f}_k \overline{\langle p, \mathbf{e}_k \rangle} = \sum_{k=-n}^{n} \overline{\widehat{p}_k} \widehat{f}_k - \sum_{k=-n}^{n} \widehat{f}_k \overline{\widehat{p}_k} = 0 \,. \end{aligned}$$

□

Theorem 2.26. *Let $f \in L^2(\mathbb{T})$. Then*

$$\|f - p\|_{L^2(\mathbb{T})}^2 = \|f - s_n(f, \cdot)\|_{L^2(\mathbb{T})}^2 + \|s_n(f, \cdot) - p\|_{L^2(\mathbb{T})}^2 \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}) \,. \tag{2.18}$$

Proof. By Proposition 2.25, if $p \in \mathscr{P}_n(\mathbb{T})$,

$$\begin{aligned} \int_{-\pi}^{\pi} \big| f(x) - p(x) \big|^2 dx &= \int_{-\pi}^{\pi} \big| f(x) - s_n(f, x) + s_n(f, x) - p(x) \big|^2 dx \\ &= \int_{-\pi}^{\pi} \big| f(x) - s_n(f, x) \big|^2 dx + \int_{-\pi}^{\pi} \big| s_n(f - p, x) \big|^2 dx \end{aligned}$$

which concludes the proposition. □

We note that (2.18) implies that

$$\|f - s_n(f, \cdot)\|_{L^2(\mathbb{T})} \leqslant \|f - p\|_{L^2(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}) \,. \tag{2.19}$$

Since $s_n(f, \cdot) \in \mathscr{P}_n(\mathbb{T})$, we conclude that

$$\|f - s_n(f, \cdot)\|_{L^2(\mathbb{T})} \leqslant \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^2(\mathbb{T})} \,.$$

Moreover, letting $p = 0$ in (2.18) we establish the famous Bessel's inequality.

Corollary 2.27. *Let $f \in L^2(\mathbb{T})$. Then for all $n \in \mathbb{N}$,*

$$\|s_n(f, \cdot)\|_{L^2(\mathbb{T})} \leqslant \|f\|_{L^2(\mathbb{T})} \,. \tag{2.20}$$

In particular,

$$\frac{1}{2}\Big[\frac{c_0^2}{2} + \sum_{k=1}^{\infty} (c_k^2 + s_k^2) \Big] = \sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} \big| f(x) \big|^2 dx \,. \qquad \textit{(Bessel's inequality)}$$

Next, we prove that the Bessel inequality is in fact an equality, called the Parseval identity. It is actually equivalent to that $\{s_n(f,\cdot)\}_{n=1}^{\infty}$ converges to f in the sense of L^2-norm; that is,

$$\lim_{n\to\infty} \big\|s_n(f,\cdot) - f\big\|_{L^2(\mathbb{T})} = 0 \qquad \forall\, f \in L^2(\mathbb{T})\,.$$

Before proceeding, we first prove that every $f \in L^2(\mathbb{T})$ can be approximated by a sequence $\{g_n\}_{n=1}^{\infty} \subseteq \mathscr{C}(\mathbb{T})$ in the sense of L^2-norm.

Proposition 2.28. *Let $f \in L^2(\mathbb{T})$. Then for all $\varepsilon > 0$ there exists $g \in \mathscr{C}(\mathbb{T})$ such that*

$$\|f - g\|_{L^2(\mathbb{T})} < \varepsilon\,.$$

In other words, $\mathscr{C}(\mathbb{T})$ is dense in $\big(L^2(\mathbb{T}), \|\cdot\|_{L^2(\mathbb{T})}\big)$.

Proof. W.L.O.G., we can assume that f is real-valued and non-zero. Let $\varepsilon > 0$ be given. Since $f \in L^2(\mathbb{T})$, the monotone convergence theorem implies that

$$\lim_{k\to\infty} \|f - (-k)\vee(f\wedge k)\|^2_{L^2(\mathbb{T})} = \lim_{k\to\infty}\int_{[-\pi,\pi)} \mathbf{1}_{\{|f(x)|>k\}}(x)|f(x)|^2\,dx = 0\,;$$

thus there exists $K > 0$ such that

$$\|f - (-k)\vee(f\wedge k)\|_{L^2(\mathbb{T})} < \frac{\varepsilon}{2} \qquad \forall\, k \geqslant K\,.$$

Let $h = (-K)\vee(f\wedge K)$. Then h is bounded and Riemann measurable; thus h is Riemann integrable over $[-\pi,\pi)$. Therefore, there exists a partition $\mathcal{P} = \{-\pi = x_0 < x_1 < \cdots < x_n = \pi\}$ of $[-\pi,\pi)$ such that $U(h,\mathcal{P}) - L(h,\mathcal{P}) < \dfrac{\varepsilon}{2}$. Define

$$S(x) = \sum_{k=0}^{n-1} \sup_{\xi\in[x_k,x_{k+1}]} f(\xi)\mathbf{1}_{[x_k,x_{k+1})}(x) \qquad \text{and} \qquad s(x) = \sum_{k=0}^{n-1} \inf_{\xi\in[x_k,x_{k+1}]} f(\xi)\mathbf{1}_{[x_k,x_{k+1})}(x)\,,$$

where $\mathbf{1}_A$ denotes the characteristic/indicator function of set A. Then

1. $-K \leqslant s \leqslant h \leqslant S \leqslant K$ on $[-\pi,\pi)\backslash\{x_1,\cdots,x_{n-1}\}$;
2. $\displaystyle\int_{-\pi}^{\pi} S(x)\,dx = U(h,\mathcal{P})$; 3. $\displaystyle\int_{-\pi}^{\pi} s(x)\,dx = L(h,\mathcal{P})$.

The properties above show that

$$\int_{-\pi}^{\pi} \big|h(x) - s(x)\big|\,dx = \int_{-\pi}^{\pi} h(x) - s(x)\,dx \leqslant \int_{-\pi}^{\pi} \big(S(x) - s(x)\big)\,dx \leqslant U(h,\mathcal{P}) - L(h,\mathcal{P}) < \frac{\varepsilon^2}{16K}\,.$$

Now, for the step function s defined on $[-\pi,\pi)$, we can always find a continuous function $g \in \mathscr{C}(\mathbb{T})$ (for example, g can be a trapezoidal function) such that

1. $\|g\|_{L^\infty(\mathbb{T})} \leqslant K$. 2. $\displaystyle\int_{-\pi}^{\pi} \big|s(x) - g(x)\big|\,dx < \frac{\varepsilon^2}{16K}$.

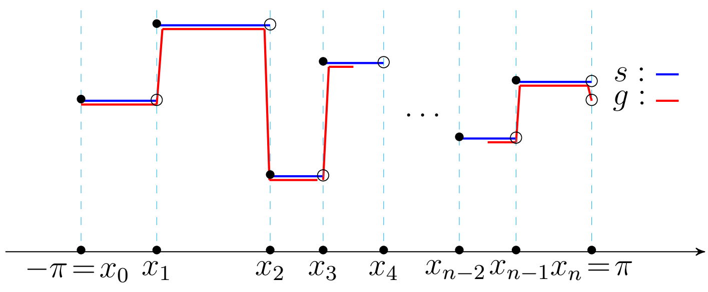

Figure 1: One way of constructing $g \in \mathscr{C}(\mathbb{T})$ given step function s

Then that $\|g\|_{L^\infty(\mathbb{T})} \leqslant K$ follows from that $\|s\|_{L^\infty(\mathbb{T})} \leqslant K$, and

$$\int_{[-\pi,\pi)} |h(x) - g(x)|\, dx \leqslant \int_{[-\pi,\pi)} |h(x) - s(x)|\, dx + \int_{[-\pi,\pi)} |s(x) - g(x)|\, dx < \frac{\varepsilon^2}{8K}.$$

Therefore,

$$\int_{-\pi}^{\pi} |h(x) - g(x)|^2\, dx \leqslant 2K \int_{[-\pi,\pi)} |h(x) - g(x)|\, dx < \frac{\varepsilon^2}{4}$$

which implies that $\|h - g\|_{L^2(\mathbb{T})} < \frac{\varepsilon}{2}$. The proposition is then concluded by the triangle inequality. □

Theorem 2.29. *Let $f : [-\pi, \pi] \to \mathbb{R}$ be bounded Riemann integrable. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |s_n(f, x) - f(x)|^2\, dx = 0 \tag{2.21}$$

and

$$\int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 = \pi\Big[\frac{c_0^2}{2} + \sum_{k=1}^{\infty}(c_k^2 + s_k^2)\Big]. \quad (\textit{Parseval's identity}) \tag{2.22}$$

Proof. Let $\varepsilon > 0$ be given. By Proposition 2.28 there exists $g \in \mathscr{C}(\mathbb{T})$ such that

$$\int_{-\pi}^{\pi} |f(x) - g(x)|^2\, dx < \frac{\varepsilon^2}{9}.$$

By the denseness of the trigonometric polynomials in $\mathscr{C}(\mathbb{T})$, there exists $h \in \mathscr{P}(\mathbb{T})$ such that $\sup_{x\in\mathbb{R}} |g(x) - h(x)| < \frac{\varepsilon}{\sqrt{18\pi}}$. Suppose that $h \in \mathscr{P}_N(\mathbb{T})$. Using (2.19),

$$\int_{-\pi}^{\pi} |g(x) - s_N(g, x)|^2\, dx \leqslant \int_{-\pi}^{\pi} |g(x) - h(x)|^2\, dx \leqslant 2\pi \cdot \frac{\varepsilon^2}{18\pi} = \frac{\varepsilon^2}{9}.$$

Since $s_N(g, \cdot) \in \mathscr{P}_n(\mathbb{T})$ if $n \geqslant N$, we must have

$$\int_{-\pi}^{\pi} |g(x) - s_n(g, x)|^2\, dx \leqslant \int_{-\pi}^{\pi} |g(x) - s_N(g, x)|^2\, dx \leqslant \frac{\varepsilon^2}{9} \qquad \forall\, n \geqslant N.$$

Therefore, for $n \geqslant N$, inequality (2.20) and the triangle inequality yield that

$$\begin{aligned}\|f - s_n(f, \cdot)\|_{L^2(\mathbb{T})} &\leqslant \|f - g\|_{L^2(\mathbb{T})} + \|g - s_n(g, \cdot)\|_{L^2(\mathbb{T})} + \|s_n(g - f, \cdot)\|_{L^2(\mathbb{T})} \\ &\leqslant 2\|f - g\|_{L^2(\mathbb{T})} + \|g - s_n(g, \cdot)\|_{L^2(\mathbb{T})} < \varepsilon;\end{aligned}$$

thus (2.21) is concluded. Finally, using (2.18) with $p = 0$ we obtain that

$$\int_{-\pi}^{\pi} \left|f(x)\right|^2 dx = \int_{-\pi}^{\pi} \left|s_n(f,x)\right|^2 dx + \int_{-\pi}^{\pi} \left|f(x) - s_n(f,x)\right|^2 dx\,.$$

Using the fact that $\displaystyle\int_{-\pi}^{\pi} \left|s_n(f,x)\right|^2 dx = \pi\Big[\frac{c_0^2}{2} + \sum_{k=1}^{n}(c_k^2 + s_k^2)\Big]$ and passing to the limit as $n \to \infty$, we conclude (2.22). ◻

Example 2.30. Example 2.6 provides that $\displaystyle\int_{-\pi}^{\pi} x^2\,dx = \pi \sum_{k=1}^{\infty} \frac{4}{k^2}$; thus $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

2.6 The Discrete Fourier "Transform" and the Fast Fourier "Transform"

Let $f : \mathbb{R} \to \mathbb{R}$ be $2L$-periodic such that f is bounded Riemann integrable over $[-L, L)$. Similar to Remark 2.2, the Fourier series representation of f, defined in Remark 2.3, can be written as

$$s(f,x) = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{\frac{i\pi kx}{L}}\,,$$

where $\widehat{f}_k = \displaystyle\frac{1}{2L}\int_{-L}^{L} f(y)e^{\frac{-i\pi ky}{L}}\,dy$. Due to the periodicity, $\widehat{f}_k$ can also be computed via the formula $\widehat{f}_k = \displaystyle\frac{1}{2L}\int_{0}^{2L} f(y)e^{\frac{-i\pi ky}{L}}\,dy$; thus $\widehat{f}_k$ can be approximated by the Riemann sum

$$\frac{1}{2L}\sum_{\ell=0}^{N-1} f\Big(\frac{2L\ell}{N}\Big)e^{\frac{-2\pi ik\ell}{N}}\frac{2L}{N} = \frac{1}{N}\sum_{\ell=0}^{N-1} f\Big(\frac{2L\ell}{N}\Big)e^{\frac{-2\pi ik\ell}{N}}\,.$$

In other words, the values of f at N points can be used to determine an approximation of the Fourier coefficients of f. This induces the following

Definition 2.31. The ***discrete Fourier transform***, symbolized by DFT, of a sequence of N complex numbers $\{x_0, x_1, \cdots, x_{N-1}\}$ is a sequence $\{X_k\}_{k\in\mathbb{Z}}$ defined by

$$X_k = \sum_{\ell=0}^{N-1} x_\ell e^{\frac{-2\pi ik\ell}{N}} \qquad \forall\, k \in \mathbb{Z}\,.$$

We note that the sequence $\{X_k\}_{k\in\mathbb{Z}}$ is N-periodic; that is, $X_{k+N} = X_k$ for all $k \in \mathbb{Z}$. Therefore, often time we only focus on one of the following N consecutive terms of the DFT:

1. $\{X_0, X_1, \cdots, X_{N-1}\}$;
2. $\big\{X_{-\frac{N}{2}}, \cdots, X_{\frac{N}{2}-1}\big\}$ if N is even;
3. $\big\{X_{-\frac{N-1}{2}}, \cdots, X_{\frac{N-1}{2}}\big\}$ if N is odd.

Example 2.32. The DFT of the sequence $\{x_0, x_1\}$ is $\{x_0 + x_1, x_0 - x_1\}$.

Remark 2.33. Let $x = [x_0, x_1, \cdots, x_{N-1}]$ be a sequence of numbers. The matlab® command **fft(x)** outputs the sequence

$$\begin{cases} [X_0, X_1, \cdots, X_{\frac{N}{2}-1}, X_{-\frac{N}{2}}, \cdots, X_{-2}, X_{-1}] & \text{if } N \text{ is even}\,, \\ [X_0, X_1, \cdots, X_{\frac{N-1}{2}}, X_{-\frac{N-1}{2}}, \cdots, X_{-2}, X_{-1}] & \text{if } N \text{ is odd}\,, \end{cases}$$

where $\big\{X_{-\frac{N}{2}}, \cdots, X_{\frac{N}{2}-1}\big\}$ and $\big\{X_{-\frac{N-1}{2}}, \cdots, X_{\frac{N-1}{2}}\big\}$ are the DFT of $\{x_0, x_1, \cdots, x_{N-1}\}$ when N is even and odd, respectively.

2.6.1 The inversion formula

Let $\{X_k\}_{k=0}^{N-1}$ be the discrete Fourier transform of the sequence $\{x_\ell\}_{\ell=0}^{N-1}$. Then $\{x_\ell\}_{\ell=0}^{N-1}$ can be recovered given $\{X_k\}_{k=0}^{N-1}$ by the inversion formula

$$x_\ell = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi i k\ell}{N}} . \tag{2.23}$$

To see this, we compute $\sum\limits_{k=0}^{N-1}\Big(\sum\limits_{j=0}^{N-1} x_j e^{\frac{-2\pi i k j}{N}}\Big)e^{\frac{2\pi i k\ell}{N}}$ and obtain that

$$\begin{aligned}\sum_{k=0}^{N-1}\Big(\sum_{j=0}^{N-1} x_j e^{\frac{-2\pi i k j}{N}}\Big)e^{\frac{2\pi i k\ell}{N}} &= \sum_{j=0}^{N-1}\Big(x_j\sum_{k=0}^{N-1} e^{\frac{2\pi i k(\ell-j)}{N}}\Big) = Nx_\ell + \sum_{\substack{j=0\\ j\neq \ell}}^{N-1}\Big(x_j\sum_{k=0}^{N-1} e^{\frac{2\pi i k(\ell-j)}{N}}\Big)\\ &= Nx_\ell + \sum_{\substack{j=0\\ j\neq \ell}}^{N-1}\Big(x_j\frac{e^{2\pi i(\ell-j)}-1}{e^{\frac{2\pi i(\ell-j)}{N}}-1}\Big) = Nx_\ell .\end{aligned}$$

The map from $\{X_k\}_{k=0}^{N-1}$ to $\{x_\ell\}_{\ell=0}^{N-1}$ is called the ***discrete inverse Fourier transform***.

We note that the inversion formula (2.23) is an analogy of

$$f(x) = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{ikx}$$

for all piecewise constant function f and $x \in \mathbb{R}$ at which f is continuous.

2.6.2 The fast Fourier transform

Let $M = [m_{k\ell}]$ be an $N\times N$ matrix with entry $m_{k\ell}$ defined by

$$m_{k\ell} = e^{\frac{-2\pi i k\ell}{N}} \qquad 0 \leqslant k,\ell \leqslant N-1 ,$$

and write $\boldsymbol{x} = (x_0, x_1, \cdots, x_{N-1})^{\rm T}$ and $\boldsymbol{X} = (X_0,\cdots,X_{N-1})^{\rm T}$. Then $\boldsymbol{X} = M\boldsymbol{x}$ and it requires N^2 multiplications to compute $\boldsymbol{X}$. The ***fast Fourier transform***, symbolized by FFT, is a much faster way to compute $\boldsymbol{X}$. In the following, we show that when $N = 2^\gamma$ for some $\gamma \in \mathbb{N}$, then there is a way to compute the DFT with at most $N\log_2 N$ multiplications.

With $N = 2^\gamma$, suppose that $(x_0,\cdots,x_{N-1})$ is a given sequence, and $\{X_k\}_{k=0}^{N-1}$ is the DFT of $\{x_k\}_{k=0}^{N-1}$. Let $\omega = e^{-\frac{2\pi i}{N}}$, and

$$\boldsymbol{x}_{\rm even} = \begin{bmatrix} x_0 & x_2 & x_4 & \cdots & x_{N-2}\end{bmatrix} \quad\text{and}\quad \boldsymbol{x}_{\rm odd} = \begin{bmatrix} x_1 & x_3 & x_5 & \cdots & x_{N-1}\end{bmatrix}$$

Then

$$\begin{aligned} X_j = \sum_{\ell=0}^{N-1} x_\ell\omega^{j\ell} &= \sum_{\substack{0\leqslant \ell\leqslant N-1\\ \ell \text{ is even}}} x_\ell\omega^{j\ell} + \omega^j \sum_{\substack{0\leqslant \ell\leqslant N-1\\ \ell \text{ is odd}}} x_\ell\omega^{j(\ell-1)}\\ &= \boldsymbol{x}_{\rm even}\cdot\begin{bmatrix}\omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)}\end{bmatrix} + \omega^j\boldsymbol{x}_{\rm odd}\cdot\begin{bmatrix}\omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)}\end{bmatrix} .\end{aligned}$$

In particular, for $0 \leqslant j \leqslant \dfrac{N}{2} - 1$,

$$\begin{aligned} X_{\frac{N}{2}+j} &= \boldsymbol{x}_{\rm even} \cdot \left[\omega^0 \ \omega^{2(\frac{N}{2}+j)} \ \omega^{4(\frac{N}{2}+j)} \ \cdots \ \omega^{(\frac{N}{2}+j)(N-2)}\right] \\ &\quad + \omega^{\frac{N}{2}+j} \boldsymbol{x}_{\rm odd} \cdot \left[\omega^0 \ \omega^{2(\frac{N}{2}+j)} \ \omega^{4(\frac{N}{2}+j)} \ \cdots \ \omega^{(\frac{N}{2}+j)(N-2)}\right] \\ &= \boldsymbol{x}_{\rm even} \cdot \left[\omega^0 \ \omega^{2j} \ \omega^{4j} \ \cdots \ \omega^{j(N-2)}\right] - \omega^j \boldsymbol{x}_{\rm odd} \cdot \left[\omega^0 \ \omega^{2j} \ \omega^{4j} \ \cdots \ \omega^{j(N-2)}\right], \end{aligned}$$

where we have used the fact that $\omega^{\frac{N}{2}} = -1$. We note that

$$\left\{\boldsymbol{x}_{\rm even} \cdot \left[\omega^0 \ \omega^{2j} \ \omega^{4j} \ \cdots \ \omega^{j(N-2)}\right]\right\}_{j=0}^{N/2}$$

is exactly the DFT of the sequence $\{x_0, x_2, \cdots, x_{N-2}\}$ and

$$\left\{\boldsymbol{x}_{\rm odd} \cdot \left[\omega^0 \ \omega^{2j} \ \omega^{4j} \ \cdots \ \omega^{j(N-1)}\right]\right\}_{j=0}^{N/2}$$

is exactly the DFT of the sequence $\{x_1, x_3, \cdots, x_{N-1}\}$. In other words, to compute the DFT of $\{x_0, \cdots, x_{N-1}\}$, where $N = 2^\gamma$, it suffices to compute the DFTs of the sequence $\{x_0, x_2, \cdots, x_{N-2}\}$ and $\{x_1, x_3, \cdots, x_{N-1}\}$. As long as the DFTs of the sequences $\{x_0, x_2, \cdots, x_{N-2}\}$ and $\{x_1, x_3, \cdots, x_{N-1}\}$ are known, it requires another $\dfrac{N}{2}$ multiplications to compute the DFT of $\{x_0, x_1, \cdots, x_{N-1}\}$.

Now we compute the total multiplications it requires to compute the DFT of the sequence $\{x_k\}_{k=0}^{2^\gamma-1}$ using the procedure above. Suppose that to compute the DFT of $\{x_k\}_{k=0}^{2^\gamma-1}$ requires $f(\gamma)$ multiplications. Then

$$f(\gamma) = 2f(\gamma - 1) + 2^{\gamma-1}.$$

It is easy to see that it requires no multiplication to compute the DFT of $\{x_0, x_1\}$ since it is simply $\{x_0+x_1, x_0-x_1\}$; thus $f(1) = 0$. Solving the iteration relation above, we obtain that $f(\gamma) = 2^{\gamma-1}(\gamma-1)$ which implies the total multiplications requires to compute the DFT of $\{x_k\}_{k=0}^{N-1}$, where $N = 2^\gamma$, is $\dfrac{N}{2}(\log_2 N - 1)$.

Example 2.34. To compute the DFT of $\{x_0, x_1, \cdots, x_7\}$, we compute the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$ first, and it requires another 4 multiplications (to compute the multiplication of ω^j and the j-th term of the DFT of $\{x_1, x_3, x_5, x_7\}$ for $0 \leqslant j \leqslant 3$). Nevertheless, instead of computing the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$ directly using matrix multiplication $\boldsymbol{X} = M\boldsymbol{x}$, we again divide the sequence of length 4 into further shorter sequence $\{x_0, x_4\}$, $\{x_2, x_6\}$, $\{x_1, x_5\}$ and $\{x_3, x_7\}$. Once the DFT of those sequence of length 2 are computed, it requires another $2 \times 2 = 4$ multiplications to compute the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$. By Example 2.32, it does not require any multiplications to compute the DFT of sequences of length 2; thus the total multiplications required to compute the DFT of $\{x_0, x_1, \cdots, x_7\}$ is $4 + 4 = 8$.

2.7 Exercise

Problem 1. Let $f \in \mathscr{C}(\mathbb{T})$ and $\{\widehat{f}_k\}_{k=-\infty}^{\infty}$ be the Fourier coefficients (given in Remark 2.2). Show that if $\sum\limits_{k=-\infty}^{\infty} |\widehat{f}_k| < \infty$, then $s_n(f, \cdot) \to f$ uniformly on $\mathbb{T}$, where $s_n(f, x) = \sum\limits_{k=-n}^{n} \widehat{f}_k e^{ikx}$.

Problem 2. This problem contributes to another proof of showing that the n-th partial sum of the Fourier series representation $s_n(f,\cdot)$ converges uniformly to f on $\mathbb{T}$ if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ for $\frac{1}{2} < \alpha \leqslant 1$. Complete the following.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be 2π-periodic such that f is bounded, Riemann integrable over $[-\pi, \pi]$. Show that
$$\widehat{f}_k = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f\big(x + \frac{\pi}{k}\big)e^{-ikx}\,dx$$
and hence
$$\widehat{f}_k = \frac{1}{4\pi}\int_{-\pi}^{\pi} \big[f(x) - f\big(x + \frac{\pi}{k}\big)\big]e^{-ikx}\,dx\,.$$
Therefore, if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, the Fourier coefficients $\widehat{f}_k$ satisfies $|\widehat{f}_k| \leqslant \dfrac{\pi^\alpha \|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}}{2k^\alpha}$.

2. Let $f : \mathbb{R} \to \mathbb{R}$ be 2π-periodic such that f is bounded, Riemann integrable over $[-\pi, \pi]$. Show that
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x+h) - f(x-h)\big|^2\,dx = \sum_{k=-\infty}^{\infty} 4\sin^2(kh)|\widehat{f}_k|^2\,.$$
Therefore, if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, the Fourier coefficients $\widehat{f}_k$ satisfies
$$\sum_{k=-\infty}^{\infty} \sin^2(kh)|\widehat{f}_k|^2 \leqslant \|f\|^2_{\mathscr{C}^{0,\alpha}(\mathbb{T})}2^{2(\alpha-1)}|h|^{2\alpha} \tag{2.24}$$

3. Let $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, and $p \in \mathbb{N}$. Show that
$$\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}_k|^2 \leqslant \frac{\|f\|^2_{\mathscr{C}^{0,\alpha}(\mathbb{T})}\pi^{2\alpha}}{2^{2\alpha p+1}}\,.$$
Hint: Let $h = \dfrac{\pi}{2^{p+1}}$ in (2.24).

4. Show that if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ for some $\frac{1}{2} < \alpha \leqslant 1$, then $\sum\limits_{k=-\infty}^{\infty} |\widehat{f}_k| < \infty$; thus Problem 1 implies that $s_n(f,\cdot) \to f$ uniformly on $\mathbb{T}$.

Problem 3. Let f be a 2π-periodic Lipchitz function. Show that for $n \geqslant 2$,
$$\|f - F_{n+1} \star f\|_{L^\infty(\mathbb{T})} \leqslant \frac{1 + 2\log n}{2n}\|f\|_{\mathscr{C}^{0,1}(\mathbb{T})} \tag{2.25}$$
and
$$\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant \frac{2\pi(1+\log n)^2}{n}\|f\|_{\mathscr{C}^{0,1}(\mathbb{T})}\,. \tag{2.26}$$
Hint: For (2.25), apply the estimate
$$F_n(x) \leqslant \min\Big\{\frac{n+1}{2\pi}, \frac{\pi}{2(n+1)x^2}\Big\}$$
in the following inequality:
$$\big|f(x) - F_{n+1} \star f(x)\big| \leqslant \Big[\int_{-\delta}^{\delta} + \int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\Big]\big|f(x+y) - f(x)\big|F_{n+1}(y)\,dy$$

with $\delta = \dfrac{\pi}{n+1}$. For (2.26), use (2.8) and note that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant \|f - F_n \star f\|_{L^\infty(\mathbb{T})}\,.$$

Problem 4. In this problem, we are concerned with the following

Theorem 2.35 (Bernstein)**.** *Suppose that f is a 2π-periodic function such that for some constant C and $\alpha \in (0,1)$,*

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant Cn^{-\alpha}$$

for all $n\in\mathbb{N}$. Then $f\in\mathscr{C}^{0,\alpha}(\mathbb{T})$.

Complete the following to prove the theorem.

1. Suppose that there is $p\in\mathscr{P}_n(\mathbb{T})$ such that

   $$\|p'\|_{L^\infty(\mathbb{T})} > n\,,\quad \|p\|_{L^\infty(\mathbb{T})} < 1\,,\quad \text{and}\quad p'(0) = \|p'\|_{L^\infty(\mathbb{T})}.$$

   Choose $\gamma \in \big[-\frac{\pi}{n}, \frac{\pi}{n}\big]$ such that $\sin(n\gamma) = -p(0)$ and $\cos(n\gamma) > 0$, and define $\alpha_k = \gamma + \frac{\pi}{n}\big(k+\frac{1}{2}\big)$ for $-n \leqslant k \leqslant n$. Show that the function $r(x) = \sin n(x-\gamma) - p(x)$ has at least one zeros in each interval (α_k, α_{k+1}).

2. Let $s\in\mathbb{N}$ be such that such that $0\in(\alpha_s,\alpha_{s+1})$. Show that r has at least 3 distinct zeros in (α_s,α_{s+1}) by noting that $r'(0) < 0$ and $r(0) = 0$.

3. Combining 1 and 2, show that

   $$\|p'\|_{L^\infty(\mathbb{T})} \leqslant n\|p\|_{L^\infty(\mathbb{T})} \qquad \forall\, p\in\mathscr{P}_n(\mathbb{T}). \tag{2.27}$$

4. Choose $p_n\in\mathscr{P}_n(\mathbb{T})$ such that $\|f-p_n\| \leqslant 2Cn^{-\alpha}$ for $n\in\mathbb{N}$. Define $q_0 = p_1$, and $q_n = p_{2^n} - p_{2^{n-1}}$ for $n\in\mathbb{N}$. Show that $\sum\limits_{n=0}^{\infty} q_n = f$ and the convergence is uniform.

5. Show that $\|q_n\|_{L^\infty(\mathbb{T})} \leqslant 6C2^{-n\alpha}$. As a consequence, show that

   $$\big|q_n(x)-q_n(y)\big| \leqslant 6Cn2^{n(1-\alpha)}|x-y| \ \text{ and } \ \big|q_n(x)-q_n(y)\big| \leqslant 12C2^{-n\alpha}.$$

6. For any $x,y\in\mathbb{T}$ with $|x-y|\leqslant 1$, choose $m\in\mathbb{N}$ such that $2^{-m}\leqslant|x-y|\leqslant 2^{1-m}$. Then use the inequality

   $$\big|f(x)-f(y)\big| \leqslant \sum_{n=0}^{m-1}\big|q_n(x)-q_n(y)\big| + \sum_{n=m}^{\infty}\big|q_n(x)-q_n(y)\big|$$

   to show that $\big|f(x)-f(y)\big| \leqslant B|x-y|^\alpha$ for some constant $B>0$.

3 Fourier Transforms

Before introducing the Fourier transform, let us "motivate" the idea a little bit. In Section 2.5 we show that $\{\mathbf{e}_k\}_{k=-\infty}^{\infty}$, where $\mathbf{e}_k(x) = e^{ikx}$, is an orthonormal basis in $L^2(\mathbb{T})$. Similarly, with $L^2([-K,K])$ denoting the inner-product space

$$L^2([-K,K]) = \big\{f : [-K,K] \to \mathbb{C} \,\big|\, f \text{ is Riemann measurable and square integrable}\big\}/\sim$$

equipped with the inner product

$$\langle f, g\rangle = \frac{1}{2K}\int_{-K}^{K} f(x)\overline{g(x)}\,dx\,,$$

the set $\Big\{\exp\big(\frac{ik\pi x}{K}\big)\Big\}_{k=-\infty}^{\infty}$ is an orthonormal basis of $L^2([-K,K])$; that is, any functions $f \in L^2([-K,K])$ can be expressed as

$$f(x) = \sum_{k=-\infty}^{\infty} \widehat{f}(k)e^{\frac{ik\pi x}{K}}\,, \quad \text{where } \widehat{f}(k) = \frac{1}{2K}\int_{-K}^{K} f(y)e^{-\frac{ik\pi y}{K}}\,dy\,. \tag{3.1}$$

Moreover, $\sum\limits_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 = \dfrac{1}{2K}\displaystyle\int_{-K}^{K} |f(x)|^2\,dx$. In other words, there is a one-to-one correspondence between $f \in L^2([-K,K])$ and $\widehat{f} \in \ell_2$. We look for a space X so that there is also a one-to-one correspondence between the square integrable functions on $\mathbb{R}$ and X. Intuitively, we can check what "might" happen by letting $K \to \infty$ in (3.1).

Making use of the Riemann sum to approximate the integral (by partition $[-K,K]$ into $2K^2$ intervals), we find that

$$\begin{aligned}
f(x) &= \frac{1}{2K}\sum_{k=-\infty}^{\infty}\int_{-K}^{K} f(y)e^{\frac{ik\pi(x-y)}{K}}dy \approx \frac{1}{2K}\sum_{k=-K^2}^{K^2}\int_{-K}^{K} f(y)e^{\frac{ik\pi(x-y)}{K}}dy \\
&\approx \frac{1}{2K}\sum_{k=-K^2}^{K^2}\sum_{\ell=1}^{2K^2} f\big(-K+\frac{\ell}{K}\big)\exp\Big(\frac{ik\pi(x+K-\frac{\ell}{K})}{K}\Big)\frac{1}{K} \\
&\approx \frac{1}{2K}\sum_{k=-K^2}^{K^2}\sum_{\ell=-K^2}^{K^2} f\big(\frac{\ell}{K}\big)\exp\Big(\frac{ik\pi(x-\frac{\ell}{K})}{K}\Big)\frac{1}{K} \qquad \big(y_\ell = \frac{\ell}{K}\,, \Delta y = \frac{1}{K}\big) \\
&= \frac{1}{2\pi}\sum_{\ell=-K^2}^{K^2}\sum_{k=-K^2}^{K^2} f\big(\frac{\ell}{K}\big)\exp\Big(i\frac{k\pi}{K}\big(x-\frac{\ell}{K}\big)\Big)\frac{\pi}{K}\frac{1}{K} \\
&\approx \frac{1}{2\pi}\sum_{\ell=-K^2}^{K^2}\int_{-K\pi}^{K\pi} f\big(\frac{\ell}{K}\big)\exp\Big(i\xi\big(x-\frac{\ell}{K}\big)\Big)\,d\xi\frac{1}{K} \qquad \big(\xi_k = \frac{k\pi}{K}\,, \Delta\xi = \frac{\pi}{K}\big) \\
&\approx \frac{1}{2\pi}\int_{-K}^{K}\int_{-K\pi}^{K\pi} f(y)e^{i\xi(x-y)}d\xi dy = \frac{1}{2\pi}\int_{-K}^{K}\int_{-K\pi}^{K\pi} f(y)e^{i\xi(x-y)}dyd\xi \\
&\approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Big[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(y)e^{-i\xi y}dy\Big]e^{i\xi x}d\xi\,.
\end{aligned}$$

Therefore, if we define $\widehat{f}(\xi) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} f(y) e^{-i\xi y} dy$, then the formal computation above suggests that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{i\xi x} d\xi \,. \tag{3.2}$$

In the rest of this section, we are going to verify the identity above rigorously (for functions f with certain properties).

3.1 The definition and basic properties of the Fourier transform

For notational convenience, we **abuse** the following notion from real analysis.

Definition 3.1. The space $L^1(\mathbb{R}^n)$ consists of all functions that are integrable over $\mathbb{R}^n$ and whose integrals are absolute convergent. In other words,

$$L^1(\mathbb{R}^n) = \Big\{ f : \mathbb{R}^n \to \mathbb{C} \,\Big|\, \int_{\mathbb{R}^n} |f(x)|\, dx < \infty \Big\} \,;$$

that is, $f \in L^1(\mathbb{R}^n)$ if the limit $\displaystyle\lim_{R\to\infty} \int_{B(0,R)} \big|f(x)\big|\, dx$ exists.

Remark 3.2. Even though we have not defined the integral for complex-valued function, the definition of $L^1(\mathbb{R}^n)$ should be clear: when f is complex-valued function, the absolute integrability of f is equivalent to that the real part and the imaginary part of f are both absolutely integrable, and

$$\begin{aligned} \int_{\mathbb{R}^n} f(x)\, dx &= \int_{\mathbb{R}^n} \mathrm{Re}(f)(x)\, dx + i \int_{\mathbb{R}^n} \mathrm{Im}(f)(x)\, dx \\ &= \int_{\mathbb{R}^n} \frac{f(x) + \overline{f(x)}}{2}\, dx + \int_{\mathbb{R}^n} \frac{f(x) - \overline{f(x)}}{2}\, dx \,, \end{aligned}$$

where $\overline{f(x)}$ is the complex conjugate of $f(x)$.

Definition 3.3. For all $f \in L^1(\mathbb{R}^n)$, the Fourier transform of f, denoted by $\mathscr{F}f$ or $\widehat{f}$, is defined by

$$(\mathscr{F}f)(\xi) = \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}^{\,n}} \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} dx \qquad \forall\, \xi \in \mathbb{R}^n \,,$$

where $x \cdot \xi = x_1\xi_1 + x_2\xi_2 + \cdots + x_n\xi_n$.

3.2 Some Properties of the Fourier Transform

Proposition 3.4. *$\mathscr{F} : L^1(\mathbb{R}^n) \to \mathscr{C}_b(\mathbb{R}^n; \mathbb{C})$, and*

$$\|\mathscr{F}f\|_\infty \equiv \sup_{\xi\in\mathbb{R}^n} \big|(\mathscr{F}f)(\xi)\big| \leqslant \|f\|_{L^1(\mathbb{R}^n)} \,. \tag{3.3}$$

Proof. First we show that $\mathscr{F}f$ is continuous if $f \in L^1(\mathbb{R}^n)$. Let $\xi \in \mathbb{R}^n$ and $\varepsilon > 0$ be given. Since $f \in L^1(\mathbb{R}^n)$, there exists $R > 0$ such that

$$\int_{B(0,r)^{\complement}} \big|f(x)\big|\, dx < \frac{\varepsilon}{3} \qquad \forall\, r \geqslant R \,.$$

Moreover, there exists $M > 0$ such that

$$\int_{\mathbb{R}^n} |f(x)|\, dx \leqslant M < \infty\,.$$

Since $\phi(x, y) = e^{-ix\cdot y}$ is uniformly continuous on $A \equiv B(0, R) \times B(\xi, 1)$, there exists $0 < \delta < 1$ such that

$$\big|\phi(x_1, y_1) - \phi(x_2, y_2)\big| < \frac{\varepsilon}{3M} \quad \text{whenever} \quad \big|(x_1, y_1) - (x_2, y_2)\big| < \delta \text{ and } (x_1, y_1), (x_2, y_2) \in A\,.$$

In particular, for all $x \in B(0, R)$ and $\eta \in B(\xi, \delta)$,

$$\big|e^{-ix\cdot\xi} - e^{-ix\cdot\eta}\big| < \frac{\varepsilon}{3M}\,.$$

Therefore, for $\eta \in B(\xi, \delta)$,

$$\begin{aligned}
\big|\widehat{f}(\eta) - \widehat{f}(\xi)\big| &\leqslant \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} |f(x)|\big|e^{-ix\cdot\eta} - e^{-ix\cdot\xi}\big|\, dx \\
&\leqslant \frac{2}{\sqrt{2\pi}^n} \int_{B(0,R)^\complement} |f(x)|\, dx + \frac{1}{\sqrt{2\pi}^n} \int_{B(0,R)} |f(x)|\big|e^{-ix\cdot\eta} - e^{-ix\cdot\xi}\big|\, dx \\
&\leqslant \frac{1}{\sqrt{2\pi}^n} \Big[\frac{2\varepsilon}{3} + \frac{\varepsilon}{3M} \int_{B(0,R)} |f(x)|\, dx\Big] < \varepsilon\,;
\end{aligned}$$

thus $\mathscr{F} f$ is continuous. The validity of (3.3) should be clear, and is left as an exercise. □

Definition 3.5. A function f on $\mathbb{R}^n$ is said to have rapid decrease/decay if for all integers $N \geqslant 0$, there exists a_N such that

$$|x|^N |f(x)| \leqslant a_N, \qquad \text{as } x \to \infty.$$

Definition 3.6. The *Schwartz space* $\mathscr{S}(\mathbb{R}^n)$ is the collection of all (complex-valued) smooth functions f on $\mathbb{R}^n$ such that f and all of its derivatives have rapid decrease. In other words,

$$\mathscr{S}(\mathbb{R}^n) = \big\{u \in \mathscr{C}^\infty(\mathbb{R}^n) \,\big|\, |\cdot|^N D^k u \text{ is bounded for all } k, N \in \mathbb{N} \cup \{0\}\big\}.$$

Elements in $\mathscr{S}(\mathbb{R}^n)$ are called Schwartz functions.

The prototype element of $\mathscr{S}(\mathbb{R}^n)$ is $e^{-|x|^2}$ which is not compactly supported, but has rapidly decreasing derivatives.

The reader is encouraged to verify the following basic properties of $\mathscr{S}(\mathbb{R}^n)$:

1. $\mathscr{S}(\mathbb{R}^n)$ is a vector space.

2. $\mathscr{S}(\mathbb{R}^n)$ is an algebra under the pointwise product of functions.

3. $\mathcal{P}u \in \mathscr{S}(\mathbb{R}^n)$ for all $u \in \mathscr{S}(\mathbb{R}^n)$ and all polynomial functions $\mathcal{P}$.

4. $\mathscr{S}(\mathbb{R}^n)$ is closed under differentiation.

5. $\mathscr{S}(\mathbb{R}^n)$ is closed under translations and multiplication by complex exponentials $e^{ix\cdot\xi}$.

Remark 3.7. Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and $\mathscr{C}_c^\infty(\Omega)$ denote the collection of all smooth functions with compact support in Ω; that is,

$$\mathscr{C}_c^\infty(\Omega) \equiv \big\{ u \in \mathscr{C}^\infty(\Omega) \,\big|\, \{x \in \Omega \,|\, f(x) \neq 0\} \subset\subset \Omega \big\}\,,$$

then $\mathscr{C}^\infty(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n)$. The set $\mathrm{cl}\big(\{x \in \Omega \,\big|\, f(x) \neq 0\}\big)$ is called the ***support*** of f.

The following lemma allows us to take the Fourier transform of Schwartz functions.

Lemma 3.8. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $f \in L^1(\mathbb{R}^n)$.*

Proof. If $f \in \mathscr{S}(\mathbb{R}^n)$, then $(1 + |x|)^{n+1}|f(x)| \leqslant C$ for some $C > 0$. Therefore, with ω_{n-1} denoting the the surface area of the $(n-1)$-dimensional unit sphere,

$$\begin{aligned}\int_{\mathbb{R}^n} \big|f(x)\big|\, dx &\leqslant \int_{\mathbb{R}^n} \frac{C}{(1+|x|)^{n+1}} dx = \int_{\mathbb{S}^{n-1}} \int_0^\infty \frac{C}{(1+r)^{n+1}} r^{n-1} dr dS \\ &\leqslant C\omega_{n-1} \int_0^\infty (1+r)^{-2} dr = C\omega_n\end{aligned}$$

which is a finite number. □

Now we check if $\widehat{f}$ is differentiable if $f \in \mathscr{S}(\mathbb{R}^n)$. Note that if $f \in \mathscr{S}(\mathbb{R}^n)$, then the function $y_j = x_j f(x)$ belongs to $\mathscr{S}(\mathbb{R}^n)$ for all $1 \leqslant j \leqslant n$.

Lemma 3.9. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $\widehat{f}$ is differentiable, and for each $j \in \{1, \cdots, n\}$, $\dfrac{\partial \widehat{f}}{\partial \xi_j}$ exists is given by*

$$\frac{\partial \widehat{f}}{\partial \xi_j}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} (-ix_j) f(x) e^{-ix\cdot\xi} dx = \Big[\frac{1}{i} x_j f(x)\Big]^\wedge(\xi)\,. \tag{3.4}$$

Proof. Let g_j be defined by $g_j(x) = -ix_j f(x)$. Since f and g_j are both Schwartz functions,

$$\lim_{k\to\infty} \int_{B(0,k)^\complement} \big|f(x)\big|\, dx = 0 \quad \text{and} \quad \lim_{k\to\infty} \int_{B(0,k)^\complement} \big|g_j(x)\big|\, dx = 0\,.$$

Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be a smooth decreasing function such that

$$\chi(r) = \begin{cases} 1 & \text{if } 0 \leqslant r \leqslant 1\,, \\ 0 & \text{if } r > 2\,. \end{cases}$$

Define $f_k(x) = \chi\big(\frac{|x|}{k}\big) f(x)$. We first show that

$$\frac{\partial \widehat{f}_k}{\partial \xi_j}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \chi\big(\frac{|x|}{k}\big) g_j(x) e^{-ix\cdot\xi}\, dx\,. \tag{3.5}$$

Too see this, we note that

$$\begin{aligned}&\frac{\widehat{f}_k(\xi + he_j) - \widehat{f}_k(\xi)}{h} - \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \chi\big(\frac{|x|}{k}\big) g_j(x) e^{-ix\cdot\xi}\, dx \\ &\quad = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \chi\big(\frac{|x|}{k}\big) f(x) e^{-ix\cdot\xi} \Big[\frac{e^{-ihx_j} - 1}{h} + ix_j\Big]\, dx \\ &\quad = \frac{1}{\sqrt{2\pi}^n} \int_{B(0,2k)} \chi\big(\frac{|x|}{k}\big) f(x) e^{-ix\cdot\xi} \Big[\frac{e^{-ihx_j} - 1}{h} + ix_j\Big]\, dx\,;\end{aligned}$$

thus by the fact that $\dfrac{e^{-ihx_j}-1}{h}+ix_j \to 0$ uniformly on $B(0,2k)$ as $h \to 0$, Theorem 1.6 implies that

$$\lim_{h\to 0} \frac{\widehat{f}_k(\xi+he_j)-\widehat{f}_k(\xi)}{h} - \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \chi\big(\frac{|x|}{k}\big) g_j(x) e^{-ix\cdot\xi}\,dx = 0\,;$$

hence (3.5) is established. Therefore, for each $k \in \mathbb{N}$,

$$\sup_{\xi\in\mathbb{R}^n}\Big|\frac{\partial \widehat{f}_k}{\partial \xi_j}(\xi) - \widehat{g}_j(\xi)\Big| \leqslant \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\big|1-\chi\big(\frac{|x|}{k}\big)\big|\big|g_j(x)\big|\,dx \leqslant \frac{1}{\sqrt{2\pi}^n}\int_{B(0,k)^{\complement}}\big|g_j(x)\big|\,dx$$

which converges to zero as $k \to \infty$. In other words, $\dfrac{\partial \widehat{f}_k}{\partial \xi_j} \to \widehat{g}_j$ uniformly on $\mathbb{R}^n$ as $k \to \infty$.

Similarly,

$$\sup_{\xi\in\mathbb{R}^n}\Big|\widehat{f}_k(\xi) - \widehat{f}(\xi)\Big| \leqslant \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\big|1-\chi\big(\frac{|x|}{k}\big)\big|\big|f(x)\big|\,dx \leqslant \frac{1}{\sqrt{2\pi}^n}\int_{B(0,k)^{\complement}}\big|f(x)\big|\,dx$$

which converges to zero as $k \to \infty$. Therefore, $\widehat{f}_k \to \widehat{f}$ uniformly on $\mathbb{R}^n$. By Theorem 1.5, $\dfrac{\partial \widehat{f}}{\partial \xi_j} = \widehat{g}_j$ so the lemma is concluded. □

Corollary 3.10. *For $f \in \mathscr{S}(\mathbb{R}^n)$, $\widehat{f} \in \mathscr{C}^\infty(\mathbb{R}^n)$ and*

$$D_\xi^\alpha \widehat{f}(\xi) = \frac{1}{i^{|\alpha|}}\Big[x_1^{\alpha_1}\cdots x_n^{\alpha_n} f(x)\Big]^\wedge(\xi)\,,$$

where for a ***multi-index*** *$\alpha = (\alpha_1,\cdots,\alpha_n)$, $|\alpha| \equiv \alpha_1+\cdots+\alpha_n$ and $D_\xi^\alpha \equiv \dfrac{\partial^{\alpha_1}}{\partial \xi_1^{\alpha_1}}\cdots\dfrac{\partial^{\alpha_n}}{\partial \xi_n^{\alpha_n}} = \dfrac{\partial^{|\alpha|}}{\partial \xi_1^{\alpha_1}\cdots\partial \xi_n^{\alpha_n}}$.*

Lemma 3.11. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $\mathscr{F}\Big(\dfrac{1}{i}\dfrac{\partial f}{\partial x_k}\Big)(\xi) = \xi_k \widehat{f}(\xi)$.*

Proof. Since $f \in \mathscr{S}(\mathbb{R}^n)$, $f(x) \to 0$ as $|x| \to \infty$; thus $\lim\limits_{R\to\infty} f(x)e^{-ix\cdot\xi}\big|_{x_k=-R}^{x_k=R} = 0$. Therefore, integrating by parts formula we find that

$$\begin{aligned}
\mathscr{F}\Big(\frac{1}{i}\frac{\partial f}{\partial x_k}\Big)(\xi) &= \frac{1}{i}\frac{1}{\sqrt{2\pi}^n}\lim_{R\to\infty}\int_{[-R,R]^n}\frac{\partial f}{\partial x_k}(x)e^{-ix\cdot\xi}dx \\
&= \frac{1}{i}\frac{1}{\sqrt{2\pi}^n}\lim_{R\to\infty}\Big[f(x)e^{-ix\cdot\xi}\Big|_{x_k=-R}^{x_k=R} + i\xi_k\int_{[-R,R]^n} f(x)e^{-ix\cdot\xi}dx\Big] \\
&= \xi_k\frac{1}{\sqrt{2\pi}^n}\lim_{R\to\infty}\int_{[-R,R]^n} f(x)e^{-ix\cdot\xi}dx = \xi_k\widehat{f}(\xi)\,.
\end{aligned}$$ □

Corollary 3.12. $\mathcal{P}(\xi_1,\cdots,\xi_n)\widehat{f}(\xi) = \mathscr{F}\Big[\mathcal{P}\Big(\dfrac{1}{i}\dfrac{\partial}{\partial x_1},\cdots,\dfrac{1}{i}\dfrac{\partial}{\partial x_n}\Big)f\Big](\xi)$ *for all $f \in \mathscr{S}(\mathbb{R}^n)$ and polynomial $\mathcal{P}$.*

Corollary 3.13. $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$.

Proof. Let $\mathcal{P}$ be a polynomial and α be a multi-index. By Corollary 3.10 and 3.12,

$$\begin{aligned}\mathcal{P}(\xi)D^\alpha \widehat{f}(\xi) &\equiv \mathcal{P}(\xi_1,\cdots,\xi_n)\frac{\partial^{|\alpha|}\widehat{f}}{\partial\xi_1^{\alpha_1}\cdots\partial\xi_n^{\alpha_n}}(\xi)\\ &= \frac{1}{i^{|\alpha|}}\Big[\mathcal{P}\big(\frac{1}{i}\frac{\partial}{\partial x_i},\cdots,\frac{1}{i}\frac{\partial}{\partial x_n}\big)\big[x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}f(x)\big]\Big]^{\wedge}(\xi)\,;\end{aligned}$$

thus $\mathcal{P}D^\alpha \widehat{f}$ is the Fourier transform of a Schwartz function. By Proposition 3.4 and Lemma 3.8, $\mathcal{P}D^\alpha \widehat{f}$ is bounded. □

Remark 3.14. There exists a duality under $\wedge$ between differentiability and rapid decrease: the more differentiability f possesses, the more rapid decrease $\widehat{f}$ has and vice versa.

Definition 3.15. For all $f \in L^1(\mathbb{R}^n)$, we define operator $\mathscr{F}^*$ by

$$(\mathscr{F}^* f)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}d\xi\,.$$

The function $\mathscr{F}^* f$ sometimes is also denoted by $\check{f}$.

The operator $\mathscr{F}^*$, indicated implicitly by the way it is written, is the formal adjoint of $\mathscr{F}$. To be more precise, we have the following

Lemma 3.16. $(\mathscr{F}u, v)_{L^2(\mathbb{R}^n)} = (u, \mathscr{F}^* v)_{L^2(\mathbb{R}^n)}$ *for all* $u, v \in \mathscr{S}(\mathbb{R}^n)$*, where* $(\cdot,\cdot)_{L^2(\mathbb{R}^n)}$ *is an inner product on* $\mathscr{S}(\mathbb{R}^n)$ *given by*

$$(u, v)_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} u(x)\overline{v(x)}\,dx\,.$$

Proof. Since $u, v \in \mathscr{S}(\mathbb{R}^n)$, by Fubini's Theorem,

$$\begin{aligned}(\mathscr{F}u, v)_{L^2(\mathbb{R}^n)} &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} u(x)e^{-ix\cdot\xi}dx\Big)\overline{v(\xi)}\,d\xi\\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x)\overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx\\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u(x)\int_{\mathbb{R}^n}\overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx = (u, \mathscr{F}^* v)_{L^2(\mathbb{R}^n)}\,.\end{aligned}$$

□

3.3 The Fourier Inversion Formula

We remind the readers that our goal is to prove (3.2), while having introduced operators $\mathscr{F}$ and $\mathscr{F}^*$, it is the same as showing that $\mathscr{F}$ and $\mathscr{F}^*$ are inverse to each other; that is, we want to show that

$$\mathscr{F}\mathscr{F}^* = \mathscr{F}^*\mathscr{F} = \mathrm{Id} \qquad \text{on} \quad \mathscr{S}(\mathbb{R}^n)\,.$$

For $t > 0$ and $x \in \mathbb{R}$, let $P_t(x) = \frac{1}{\sqrt{t}}e^{-\frac{x^2}{2t}}$. Note that $P_t \in \mathscr{S}(\mathbb{R})$ and P_t is normalized so that

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} P_t(x)\,dx = 1\,.$$

Now we compute the Fourier transform of P_t. By Lemma 3.9, we find that

$$\frac{d\widehat{P}_t}{d\xi}(\xi) = \frac{-i}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x)e^{-ix\xi}\,dx = \frac{-i}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x)\cos(\xi x)\,dx - \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x)\sin(\xi x)\,dx\,.$$

Since the functions $y = xP_t(x)$ is absolutely integrable over $\mathbb{R}$ for each fixed $t > 0$, the integral $\int_{\mathbb{R}} xP_t(x)\cos(\xi x)\,dx$ converges absolutely; thus by the fact that $x\cos(\xi x)$ are odd functions in x, we have

$$\int_{\mathbb{R}} xP_t(x)\cos(\xi x)\,dx = \lim_{R\to\infty} \int_{-R}^{R} xP_t(x)\cos(\xi x)\,dx = 0\,.$$

As a consequence,

$$\frac{d\widehat{P}_t}{d\xi}(\xi) = -\frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xe^{-\frac{x^2}{2t}}\sin(x\xi)dx\,.$$

Similarly, $\widehat{P}_t(\xi) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-\frac{x^2}{2t}}\cos(x\xi)dx$, and the integration by parts formula implies that

$$\begin{aligned}
\frac{d\widehat{P}_t}{d\xi}(\xi) &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \frac{\partial}{\partial\xi}\Big(e^{-\frac{x^2}{2t}}\cos(x\xi)\Big)dx = -\frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xe^{-\frac{x^2}{2t}}\sin(x\xi)dx \\
&= -\frac{1}{\sqrt{2\pi t}} \lim_{R\to\infty} \Big[-te^{-\frac{x^2}{2t}}\sin(x\xi)\Big|_{x=-R}^{x=R} + \int_{-R}^{R} \xi te^{-\frac{x^2}{2t}}\cos(x\xi)\,dx\Big] \\
&= -\frac{\xi t}{\sqrt{2\pi t}} \lim_{R\to\infty} \int_{-R}^{R} e^{-\frac{x^2}{2t}}\cos(x\xi)\,dx = -\frac{\xi t}{\sqrt{2\pi t}} \lim_{R\to\infty} \int_{-R}^{R} e^{-\frac{x^2}{2t}}\big[\cos(x\xi) - i\sin(x\xi)\big]\,dx \\
&= -\frac{\xi t}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-\frac{x^2}{2t}}e^{-ix\xi}dx = -\xi t\widehat{P}_t(\xi)\,;
\end{aligned}$$

thus $\widehat{P}_t(\xi) = Ce^{-\frac{t\xi^2}{2}}$. By the fact that $\widehat{P}_t(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} P_t(x)dx = 1$, we must have

$$\widehat{P}_t(\xi) = e^{-\frac{1}{2}t\xi^2}\,. \tag{3.6}$$

For $x \in \mathbb{R}^n$, if we define $\mathrm{P}_t(x) = \prod_{k=1}^{n} P_t(x_k) = \Big(\frac{1}{\sqrt{t}}\Big)^n e^{-\frac{|x|^2}{2t}}$, then (3.6) implies that $\widehat{\mathrm{P}}_t(\xi) = e^{-\frac{1}{2}t|\xi|^2}$. Therefore,

$$\widehat{\mathrm{P}}_t(\xi) = \Big(\frac{1}{\sqrt{t}}\Big)^n \mathrm{P}_{\frac{1}{t}}(\xi)$$

which, together with the fact that $\check{f}(x) = \widehat{f}(-x)$, further shows that

$$\check{\widehat{\mathrm{P}}}_t(x) = \Big(\frac{1}{\sqrt{t}}\Big)^n \widehat{\mathrm{P}_{\frac{1}{t}}}(-x) = \Big(\frac{1}{\sqrt{t}}\Big)^n e^{-\frac{1}{2}t^{-1}|x|^2} = \mathrm{P}_t(x)\,.$$

Similarly, $\widehat{\check{\mathrm{P}}}_t(\xi) = \mathrm{P}_t(\xi)$, so we establish that

$$\mathscr{F}^*\mathscr{F}(\mathrm{P}_t) = \mathscr{F}\mathscr{F}^*(\mathrm{P}_t) = \mathrm{P}_t\,. \tag{3.7}$$

The proof of the following lemma is similar to that of Theorem 2.18.

Lemma 3.17. *If $g \in \mathscr{S}(\mathbb{R}^n)$, then $\mathrm{P}_t * g \to g$ uniformly on $\mathbb{R}^n$ as $t \to 0^+$, where the convolution operator $*$ is given by*

$$(\mathrm{P}_t * g)(x) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \mathrm{P}_t(x-y)g(y)\,dy = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \mathrm{P}_t(y)g(x-y)\,dy\,. \tag{3.8}$$

Proof. Let $\varepsilon > 0$ be given. Since $g \in \mathscr{S}(\mathbb{R}^n)$, g is uniformly continuous; thus there exists $\delta > 0$ such that

$$\big|g(x) - g(y)\big| < \frac{\varepsilon}{2} \quad \forall\, |x-y| < \delta\,.$$

Since $\dfrac{1}{\sqrt{2\pi}^n}\displaystyle\int_{\mathbb{R}^n} \mathrm{P}_t(x)\,dx = 1$, for all $x \in \mathbb{R}^n$ we have

$$\begin{aligned}
\big|(\mathrm{P}_t * g)(x) - g(x)\big| &= \frac{1}{\sqrt{2\pi}^n}\Big|\int_{\mathbb{R}^n} g(x-y)\mathrm{P}_t(y)\,dy - \int_{\mathbb{R}^n} g(x)\mathrm{P}_t(y)\,dy\Big| \\
&= \frac{1}{\sqrt{2\pi}^n}\Big|\int_{\mathbb{R}^n} \big((g(x-y) - g(x)\big)\mathrm{P}_t(y)\,dy\Big| \\
&\leqslant \frac{\varepsilon}{2}\frac{1}{\sqrt{2\pi}^n}\int_{|y|<\delta} \mathrm{P}_t(y)\,dy + \frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy\,,
\end{aligned}$$

so we obtain that

$$\big\|(\mathrm{P}_t * g) - g\big\|_\infty \leqslant \frac{\varepsilon}{2} + \frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy\,.$$

Note that as $t \to 0^+$,

$$\int_{|y|>\delta} \mathrm{P}_t(y)\,dy = \frac{1}{\sqrt{t}^n}\int_{|y|>\delta} e^{-\frac{|y|^2}{2t}}\,dy = \int_{|z|>\frac{\delta}{\sqrt{t}}} e^{-\frac{|z|^2}{2}}\,dz \to 0\,;$$

thus there exists $h > 0$ such that if $0 < |t| < h$,

$$\frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy < \frac{\varepsilon}{2}\,.$$

Therefore, we conclude that

$$\big\|(P_t * g) - g\big\|_\infty < \varepsilon \qquad \forall\, 0 < t < h$$

which shows that $\mathrm{P}_t * g \to g$ uniformly as $t \to 0^+$. □

Before proceeding, we establish a special case of the Fubini theorem for improper integrals which will be used in the following discussion.

Proposition 3.18 (Fubini theorem - special case)**.** *Let $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be absolutely integrable, and $g, h \in L^1(\mathbb{R}^n)$. If $|f(x,y)| \leqslant |g(x)||h(y)|$ for all $x, y \in \mathbb{R}^n$, then*

$$\begin{aligned}
\int_{\mathbb{R}^{2n}} f(x,y)d(x,y) &\equiv \lim_{R\to\infty}\int_{[-R,R]^{2n}} f(x,y)d(x,y) \\
&= \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx = \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x,y)\,dx\Big)dy\,.
\end{aligned}$$

Proof. Let $\varepsilon > 0$ be given. Since $g, h \in L^1(\mathbb{R}^n)$, there exists $R_0 > 0$ such that

$$\int_{([-R,R]^n)^\complement} \big[|g(x)| + |h(x)|\big]\,dx < \frac{\varepsilon}{1 + \|g\|_{L^1(\mathbb{R}^n)} + \|h\|_{L^1(\mathbb{R}^n)}} \qquad \text{whenever} \quad R > R_0.$$

Therefore, the Fubini theorem for Riemann integral implies that

$$\begin{aligned}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx &= \int_{[-R,R]^n}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx + \int_{([-R,R]^n)^\complement}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx \\ &= \int_{[-R,R]^n}\Big[\Big(\int_{[-R,R]^n} + \int_{([-R,R]^n)^\complement}\Big) f(x,y)\,dy\Big]dx + \int_{([-R,R]^n)^\complement}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx \\ &= \int_{[-R,R]^{2n}} f(x,y)d(x,y) + \int_{[-R,R]^n}\Big(\int_{([-R,R]^n)^\complement} f(x,y)\,dy\Big)dx + \int_{([-R,R]^n)^\complement}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx\,;\end{aligned}$$

thus by the fact that $|f(x,y)| \leqslant |g(x)||h(y)|$,

$$\begin{aligned}&\Big|\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x,y)\,dy\Big)dx - \int_{[-R,R]^{2n}} f(x,y)d(x,y)\Big| \\ &\qquad \leqslant \int_{[-R,R]^n}\Big(\int_{([-R,R]^n)^\complement} |g(x)||h(y)|\,dy\Big)dx + \int_{([-R,R]^n)^\complement}\Big(\int_{\mathbb{R}^n} |g(x)||h(y)|\,dy\Big)dx \\ &\qquad \leqslant \|g\|_{L^1(\mathbb{R}^n)}\int_{([-R,R]^n)^\complement} |h(y)|\,dy + \|h\|_{L^1(\mathbb{R}^n)}\int_{([-R,R]^n)^\complement} |g(x)|\,dx \\ &\qquad < \frac{\big(\|g\|_{L^1(\mathbb{R}^n)} + \|h\|_{L^1(\mathbb{R}^n)}\big)\varepsilon}{1 + \|g\|_{L^1(\mathbb{R}^n)} + \|h\|_{L^1(\mathbb{R}^n)}} < \varepsilon\end{aligned}$$

whenever $R > R_0$. ◻

Lemma 3.19. *If f and $g \in \mathscr{S}(\mathbb{R}^n)$, then*

$$(\check{f} * g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi\,.$$

Proof. By definition of $\check{f}$ and convolution,

$$(\check{f} * g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\check{f}(x-y)g(y)\,dy = \Big(\frac{1}{2\pi}\Big)^n\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(\xi)e^{-i(x-y)\cdot\xi}g(y)\,d\xi\Big)dy\,.$$

Since the function $F(\xi,y) \equiv f(\xi)g(y)e^{-i(x-y)\cdot\xi}$ has the property that $|F(\xi,y)| \leqslant |f(\xi)||g(y)|$ for some $f, g \in \mathscr{S}(\mathbb{R}^n)$, the Fubini theorem (Proposition 3.18) implies that

$$\begin{aligned}(\check{f} * g)(x) &= \Big(\frac{1}{2\pi}\Big)^n\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(\xi)e^{-ix\cdot\xi}e^{iy\cdot\xi}g(y)\,dy\Big)d\xi \\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{-ix\cdot\xi}\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{iy\cdot\xi}g(y)\,dy\Big)d\xi = \Big(\frac{1}{2\pi}\Big)^n\int_{\mathbb{R}^n} f(\xi)e^{-ix\cdot\xi}\widehat{g}(\xi)\,d\xi\,.\end{aligned}$$ ◻

Theorem 3.20 (Fourier Inversion Formula)**.** *If $g \in \mathscr{S}(\mathbb{R}^n)$, then $\check{\widehat{g}} = \widehat{\check{g}}(\xi) = g(\xi)$. In other words, $\mathscr{F}\mathscr{F}^* = \mathscr{F}^*\mathscr{F} = \mathrm{Id}$.*

Proof. Apply Lemma 3.19 with $f(\xi) = \widehat{\mathrm{P}}_t(\xi) = e^{-\frac{1}{2}t|\xi|^2}$, using (3.7) we find that

$$(\mathrm{P}_t * g)(x) = (\check{f} * g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-\frac{1}{2}t|\xi|^2}e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi\,.$$

Letting $t \to 0^+$, by Lemma 3.17 it suffices to show that

$$\lim_{t\to 0^+}\int_{\mathbb{R}^n} e^{-\frac{1}{2}t|\xi|^2}e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi = \int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{g}\,d\xi\,.$$

To see this, let $\varepsilon > 0$ be given. Since $\widehat{g} \in \mathscr{S}(\mathbb{R}^n)$, there exists $R > 0$ such that

$$\int_{B(0,R)^\complement} \big|\widehat{g}(\xi)\big|\, d\xi < \frac{\varepsilon}{2}\,.$$

For this particular R, there exists $\delta > 0$ such that if $0 < t < \delta$,

$$\frac{tR^2}{2}\|\widehat{g}\|_{L^1(\mathbb{R}^n)} < \frac{\varepsilon}{2}\,.$$

Therefore, if $0 < t < \delta$, using the fact that $1 - e^{-x} \leqslant x$ for $x > 0$,

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^n} e^{-\frac{1}{2}t|\xi|^2} e^{ix\cdot\xi}\widehat{g}(\xi)\, d\xi - \int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{g}\, d\xi\Big| \\
&\qquad \leqslant \Big(\int_{B(0,R)} + \int_{B(0,R)^\complement}\Big)\big|e^{-\frac{1}{2}t|\xi|^2} - 1\big|\big|\widehat{g}(\xi)\big|\, d\xi \\
&\qquad \leqslant \frac{1}{2}tR^2 \int_{B(0,R)} \big|\widehat{g}(\xi)\big|\, d\xi + \int_{B(0,R)^\complement} \big|\widehat{g}(\xi)\big|\, d\xi < \varepsilon\,.
\end{aligned}$$

Therefore,

$$g(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \widehat{g}(\xi)e^{ix\cdot\xi}d\xi = \check{\widehat{g}}(x)\,.$$

Let $\widetilde{\ }$ denote the reflection operator given by $\widetilde{f}(x) = f(-x)$. Then the change of variable formula implies that

$$\begin{aligned}
\check{g}(\xi) &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{ix\cdot\xi}dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{-i(-x)\cdot\xi}dx \\
&= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(-x)e^{-ix\cdot\xi}dx = \widehat{\widetilde{g}}(\xi)\,.
\end{aligned}$$

On the other hand,

$$\check{g}(\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{-ix\cdot(-\xi)}dx = \widehat{g}(-\xi) = \widetilde{\widehat{g}}(\xi)\,;$$

thus $\widehat{\check{g}}(\xi) = \widehat{\widetilde{\widehat{g}}}(\xi) = \check{\widehat{g}}(\xi) = g(\xi)$. □

Corollary 3.21. *$\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is a bijection.*

Remark 3.22. In view of the Fourier Inversion Formula (Theorem 3.20), $\mathscr{F}^*$ sometimes is written as $\mathscr{F}^{-1}$, and is called the ***inverse Fourier transform***.

We have established the Fourier inversion formula for Schwartz class functions. Our goal next is to show that the Fourier inversion formula holds (in certain sense) for absolutely integrable function whose Fourier transform is also absolutely integrable. Motivated by the Fourier inversion formula, we would like to show, if possible, that

$$\widehat{\check{f}} = \check{\widehat{f}} = f \qquad \forall\, f \in L^1(\mathbb{R}^n) \text{ such that } \widehat{f} \in L^1(\mathbb{R}^n).$$

The above assertion cannot be true since $\widehat{\check{f}}$ and $\check{\widehat{f}}$ are both continuous (by Proposition 3.3) while $f \in L^1(\mathbb{R}^n)$ which is not necessary continuous. However, we will prove that the identity above holds for points x at which f is continuous.

Before proceeding, let us discuss some properties concerning the Fourier transform the product and the convolution of two Schwartz class functions.

Theorem 3.23. *If $f, g \in \mathscr{S}(\mathbb{R}^n)$, then $\mathscr{F}(f*g) = \widehat{f}\,\widehat{g}$. In particular, $f*g \in \mathscr{S}(\mathbb{R}^n)$ if $f, g \in \mathscr{S}(\mathbb{R}^n)$.*

Proof. By the definition of the Fourier transform and the convolution,

$$\begin{aligned}
\widehat{f*g}(\xi) &= \frac{1}{\sqrt{2\pi}^n}\mathscr{F}\Big(\int_{\mathbb{R}^n} f(\cdot - y)g(y)\,dy\Big)(\xi) \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\Big[\int_{\mathbb{R}^n} f(x-y)g(y)\,dy\Big]e^{-ix\cdot\xi}dx \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(x)\Big(\int_{\mathbb{R}^n} g(y)e^{-i(x+y)\cdot\xi}dx\Big)dy \\
&= \Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}dx\Big)\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(y)e^{-iy\cdot\xi}dy\Big)
\end{aligned}$$

which concludes the theorem. □

Corollary 3.24. *$\mathscr{F}^*(f*g) = \check{f}\,\check{g}$, $\widehat{fg} = \widehat{f}*\widehat{g}$ and $\check{\widetilde{fg}} = \check{f}*\check{g}$ for all $f, g \in \mathscr{S}(\mathbb{R}^n)$.*

Exercise: Show that $\mathscr{F}(\widehat{f}\cdot g) = \widetilde{f}*\widehat{g}$ for all $f,\ g \in \mathscr{S}(\mathbb{R}^n)$.

Theorem 3.25 (Plancherel formula for $\mathscr{S}(\mathbb{R}^n)$)**.** *If $f,\ g \in \mathscr{S}(\mathbb{R}^n)$, then*

$$(f,g)_{L^2(\mathbb{R}^n)} = (\widehat{f},\widehat{g})_{L^2(\mathbb{R}^n)}.$$

Proof. Recall that $(f,g)_{L^2(\mathbb{R}^n)} = \displaystyle\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx$. By the Fubini theorem (Proposition (3.18)),

$$\begin{aligned}
(\check{f},g)_{L^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n}\check{f}(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n}\Big[\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}d\xi\Big]\overline{g(x)}\,dx \\
&= \int_{\mathbb{R}^n} f(\xi)\Big[\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\overline{g(x)e^{-ix\cdot\xi}}dx\Big]d\xi = (f,\widehat{g})_{L^2(\mathbb{R}^n)}\,.
\end{aligned}$$

Therefore, $(f,g)_{L^2(\mathbb{R}^n)} = (\check{\widehat{f}},g)_{L^2(\mathbb{R}^n)} = (\widehat{f},\widehat{g})_{L^2(\mathbb{R}^n)}$. □

Lemma 3.26. *Let $f \in L^1(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$. Then $\langle\widehat{f},g\rangle = \langle f,\widehat{g}\rangle$ and $\langle\check{f},g\rangle = \langle f,\check{g}\rangle$, where $\langle f,g\rangle = \displaystyle\int_{\mathbb{R}^n} f(x)g(x)\,dx$.*

Proof. We only prove $\langle\widehat{f},g\rangle = \langle f,\widehat{g}\rangle$ if $f \in L^1(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$. By Proposition 3.4, $\widehat{f}$ is bounded and continuous on $\mathbb{R}^n$; thus $\widehat{f}g$ is an absolutely integrable continuous function. By the Fubini Theorem (Proposition 3.18),

$$\begin{aligned}
\langle\widehat{f},g\rangle &= \int_{\mathbb{R}^n}\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}dx\Big)g(\xi)d\xi = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x)g(\xi)e^{-ix\cdot\xi}dx\Big)d\xi \\
&= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x)g(\xi)e^{-ix\cdot\xi}d\xi\Big)dx = \int_{\mathbb{R}^n} f(x)\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(\xi)e^{-ix\cdot\xi}d\xi\Big)dx
\end{aligned}$$

which is exactly $\langle f,\widehat{g}\rangle$. □

Remark 3.27. Even though in general an square integrable function might not be integrable, using the Plancherel formula the Fourier transform of L^2-functions can still be defined. Note that the Plancherel formula provides that

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2(\mathbb{R}^n)} \qquad \forall\, f \in \mathscr{S}(\mathbb{R}^n)\,. \tag{3.9}$$

If $f \in L^2(\mathbb{R}^n)$; that is, $|f|$ is square integrable, by the fact that $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, there exists a sequence $\{f_k\}_{k=1}^\infty \subseteq \mathscr{S}(\mathbb{R}^n)$ such that $\lim\limits_{k\to\infty} \|f_k - f\|_{L^2(\mathbb{R}^n)} = 0$. Then $\{f_k\}_{k=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$; thus (3.9) implies that $\{\widehat{f}_k\}_{k=1}^\infty$ is also a Cauchy sequence in $L^2(\mathbb{R}^n)$. By the completeness of $L^2(\mathbb{R}^n)$, there exists $g \in L^2(\mathbb{R}^n)$ such that

$$\lim_{k\to\infty} \|\widehat{f}_k - g\|_{L^2(\mathbb{R}^n)} = 0\,.$$

We note that such a limit g is independent of the choice of sequence $\{f_k\}_{k=1}^\infty$ used to approximate f; thus we can denote this limit g as $\widehat{f}$. In other words, $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Moreover, by that $f_k \to f$ and $\widehat{f}_k \to \widehat{f}$ in $L^2(\mathbb{R}^n)$ as $k \to \infty$, we find that

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2(\mathbb{R}^n)} \qquad \forall\, f \in L^2(\mathbb{R}^n)\,,$$

and the parallelogram law further implies that $(f, g)_{L^2(\mathbb{R}^n)} = (\widehat{f}, \widehat{g})_{L^2(\mathbb{R}^n)}$ for all $f, g \in L^2(\mathbb{R}^n)$. Similar argument applies to the case of inverse transform of L^2-functions; thus we conclude that

$$(f, g)_{L^2(\mathbb{R}^n)} = (\widehat{f}, \widehat{g})_{L^2(\mathbb{R}^n)} = (\check{f}, \check{g})_{L^2(\mathbb{R}^n)} \qquad \forall\, f, g \in L^2(\mathbb{R}^n)\,. \tag{3.10}$$

Next, we shall establish some useful tools in analysis that can be applied in a wide range of applications. Those tools are fundamental in real analysis; however, we assume only knowledge of elementary analysis again to derive those results. We first define the class of locally integrable functions.

Definition 3.28. The space $L^1_{\rm loc}(\mathbb{R}^n)$ consists of all functions (defined on $\mathbb{R}^n$) that are absolutely integrable over all bounded open subsets of $\mathbb{R}^n$ and whose integrals are absolute convergent. In other words,

$$L^1_{\rm loc}(\mathbb{R}^n) = \Big\{ f : \mathbb{R}^n \to \mathbb{C} \,\Big|\, \int_{\mathcal{U}} f(x)\, dx \text{ is absolutely convergent for all bounded open } \mathcal{U} \subseteq \mathbb{R}^n \Big\}\,.$$

Again, we emphasize that we ***abuse*** the notation $L^1_{\rm loc}(\mathbb{R}^n)$ which in fact stands for a larger class of functions. We also note that $L^1(\mathbb{R}^n) \subseteq L^1_{\rm loc}(\mathbb{R}^n)$.

Lemma 3.29. *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with compact support, and $f \in L^1_{loc}(\mathbb{R}^n)$. Then $\displaystyle\int_{-\infty}^{\infty} \phi(x-y) f(y)\, dy$ is smooth.*

Proof. It suffices to show that

$$\frac{\partial}{\partial x_j} \int_{-\infty}^{\infty} \phi(x-y) f(y)\, dy = \int_{-\infty}^{\infty} \phi_{x_j}(x-y) f(y)\, dy\,.$$

Since ϕ has compact support, ϕ_{x_j} is uniformly continuous on $\mathbb{R}$; thus there exists $\delta > 0$ such that

$$\big|\phi_{x_j}(z_1) - \phi_{x_j}(z_2)\big| < \frac{\varepsilon}{1 + \displaystyle\int_{-\infty}^{\infty} |f(y)|\, dy} \qquad \forall\, |z_1 - z_2| < \delta\,.$$

Define $g(x) = \displaystyle\int_{-\infty}^{\infty} \phi(x-y) f(y)\, dy$. Then for some function $\vartheta : \mathbb{R} \to (0,1)$,

$$\frac{\phi(x + he_j - y) - \phi(x-y)}{h} = \phi_{x_j}(x - y + \vartheta(h) h e_j)\,;$$

thus if $|h| < \delta$,

$$\begin{aligned}\Big|\frac{g(x+he_j)-g(x)}{h} &- \int_{-\infty}^{\infty} \phi_{x_j}(x-y)f(y)\,dy\Big| \\ &= \int_{-\infty}^{\infty} \Big|\frac{\phi(x+he_j-y)-\phi(x-y)}{h} - \phi_{x_j}(x-y)\Big|\big|f(y)\big|\,dy \\ &= \int_{-\infty}^{\infty} \big|\phi_{x_j}(x-y+\vartheta(h)he_j) - \phi_{x_j}(x-y)\big|\big|f(y)\big|\,dy < \varepsilon\,.\end{aligned}$$

This implies that $g_{x_j}(x) = \displaystyle\int_{-\infty}^{\infty} \phi_{x_j}(x-y)f(y)\,dy$. □

A special class of functions will be used as the role of ϕ in Lemma 3.29. Let $\zeta : \mathbb{R} \to \mathbb{R}$ be a smooth function defined by

$$\zeta(x) = \begin{cases} \exp\big(\frac{1}{x^2-1}\big) & \text{if } |x| < 1\,, \\ 0 & \text{if } |x| \geqslant 1\,. \end{cases}$$

For $x \in \mathbb{R}^n$, define $\eta_1(x) = C\zeta(|x|)$, where C is chosen so that $\displaystyle\int_{\mathbb{R}^n} \eta_1(x)\,d = 1$. The change of variables formula then implies that $\eta_\varepsilon(x) \equiv \varepsilon^{-n}\eta_1(x/\varepsilon)$ has integral 1.

Definition 3.30. The sequence $\{\eta_\varepsilon\}_{\varepsilon>0}$ is called the ***standard mollifiers***.

Example 3.31. Let $f = \mathbf{1}_{[a,b]}$, the characteristic/indicator function of the closed interval $[a,b]$. Then for $\varepsilon \ll 1$, the function $\eta_\varepsilon * f = \sqrt{2\pi}\eta_\varepsilon \ast f$ is smooth and has the property that

$$(\eta_\varepsilon * f)(x) = \begin{cases} 1 & \text{if } x \in [a+\varepsilon, b-\varepsilon]\,, \\ 0 & \text{if } x \in [a-\varepsilon, b+\varepsilon]^\complement, \end{cases}$$

and $0 \leqslant f \leqslant 1$. Therefore, $\eta_\varepsilon * f$ converges pointwise to f on $\mathbb{R}\backslash\{a,b\}$.

Since η_ε is supported in the closure of $B(0,\varepsilon)$, Lemma 3.29 implies that for any $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, $\eta_\varepsilon * f$ is smooth function. The following lemma shows that $\eta_\varepsilon * f$ converges to f at points of continuity of f.

Lemma 3.32. *Let $f \in L^1(\mathbb{R}^n)$ and x_0 be a continuity of f. Then*

$$(\eta_\varepsilon * f)(x_0) = \sqrt{2\pi}^{\,n}(\eta_\varepsilon \ast f)(x_0) \to f(x_0) \quad \textit{as} \quad \varepsilon \to 0\,.$$

Proof. Let $\epsilon > 0$ be given. Since f is continuous at x_0, there exists $\delta > 0$ such that

$$\big|f(y) - f(x_0)\big| < \frac{\epsilon}{2} \qquad \forall\, |y - x_0| < \delta\,.$$

Therefore, by the fact that $\displaystyle\int_{\mathbb{R}^n} \eta_\varepsilon(x_0-y)\,dy = 1$, if $0 < \varepsilon < \delta$,

$$\begin{aligned}\big|(\eta_\varepsilon * f)(x_0) - f(x_0)\big| &= \Big|\int_{\mathbb{R}^n} \eta_\varepsilon(x_0-y)f(y)\,dy - \int_{\mathbb{R}^n} \eta_\varepsilon(x_0-y)f(x_0)\,dy\Big| \\ &\leqslant \int_{B(x_0,\varepsilon)} \eta_\varepsilon(x_0-y)\big|f(y)-f(x_0)\big|\,dy \\ &\leqslant \frac{\epsilon}{2}\int_{B(x_0,\varepsilon)} \eta_\varepsilon(x_0-y)\,dy < \epsilon\end{aligned}$$

which implies $(\eta_\varepsilon * f)(x_0) \to f(x_0)$ as $\varepsilon \to 0$. □

Lemma 3.33. *Let $f \in L^1_{\rm loc}(\mathbb{R}^n)$. If $\langle f, g\rangle = 0$ for all $g \in \mathscr{S}(\mathbb{R}^n)$, then $f(x) = 0$ for all $x \in \{y \in \mathbb{R}^n \,|\, f$ is continuous at $y\}$.*

Proof. W.L.O.G. we can assume that f is real-valued. Let $\{\eta_\varepsilon\}_{\varepsilon>0}$ be the standard mollifiers, x_0 be a point of continuity of f, and $f_\varepsilon \equiv \eta_\varepsilon * f = \sqrt{2\pi}^n(\eta_\varepsilon \ast f)$. Then Lemma 3.29 shows that f_ε are smooth for all $\varepsilon > 0$.

Define $g(x) \equiv \eta_1(x - x_0)f_\varepsilon(x)$. Then $g \in \mathscr{S}(\mathbb{R}^n)$ since f_ε, η_1 are smooth and $\eta_1(\cdot - x_0)$ vanishes outside $D(x_0, 1)$. Since $\eta_\varepsilon, g \in \mathscr{S}(\mathbb{R}^n)$, Theorem 3.23 implies that $\eta_\varepsilon * g \equiv \sqrt{2\pi}^n(\eta_\varepsilon \ast g) \in \mathscr{S}(\mathbb{R}^n)$; thus

$$\langle f, \eta_\varepsilon * g\rangle = 0 \qquad \forall\, \varepsilon > 0\,.$$

Since $f \in L^1(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$, the function $F(x, y) = f(x)g(y)$ is absolutely integrable over $\mathbb{R}^n \times \mathbb{R}^n$. Moreover, by the boundedness and continuity of η_ε, the comparison test implies that the function $G(x, y) = F(x, y)\eta_\varepsilon(x - y)$ is also absolutely integrable over $\mathbb{R}^n \times \mathbb{R}^n$. Since $\big|G(x, y)\big| \leqslant C\big|f(x)\big|\big|g(y)\big|$, the Fubini theorem (Proposition 3.18) implies that

$$\langle f, \eta_\varepsilon * g\rangle = \int_{\mathbb{R}^n} f(x)\Big(\int_{\mathbb{R}^n} \eta_\varepsilon(x - y)g(y)\,dy\Big)dx = \int_{\mathbb{R}^n} g(y)\Big(\int_{\mathbb{R}^n} \eta_\varepsilon(x - y)f(x)\,dx\Big)dy\,;$$

thus by the fact that $\eta_\varepsilon(x-y) = \eta_\varepsilon(y-x)$ we conclude that $\langle f, \eta_\varepsilon * g\rangle = \langle \eta_\varepsilon * f, g\rangle$. As a consequence,

$$0 = \langle f, \eta_\varepsilon * g\rangle = \langle \eta_\varepsilon * f, \eta_1(\cdot - x_0)(\eta_\varepsilon * f)\rangle = \int_{\mathbb{R}^n} \eta_1(x - x_0)\big|(\eta_\varepsilon * f)(x)\big|^2\,dx$$

which implies that $\eta_\varepsilon * f = 0$ on $B(x_0, 1)$. We then conclude from Lemma 3.32 that $(\eta_\varepsilon * f)(x_0) \to f(x_0)$ as $\varepsilon \to 0$. □

Now we state the Fourier inversion formula for functions of more general class.

Theorem 3.34 (Fourier Inversion Formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be an absolutely integrable function such that $\widehat{f}$ is also absolutely integrable. Then*

$$\check{\widehat{f}}(x) = f(x) \qquad \forall\, x \in \big\{y \in \mathbb{R}^n \,\big|\, f \textit{ is continuous at } y\big\}\,.$$

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ be such that $f, \widehat{f} \in L^1(\mathbb{R}^n)$. By Lemma 3.26 and the Fourier inversion formula for Schwartz class functions (Theorem 3.20),

$$\langle \check{\widehat{f}}, g\rangle = \langle \widehat{f}, \check{g}\rangle = \langle f, \widehat{\check{g}}\rangle = \langle f, g\rangle \qquad \forall\, g \in \mathscr{S}(\mathbb{R}^n)\,.$$

In other words, if $f, \widehat{f} \in L^1(\mathbb{R}^n)$,

$$\langle \check{\widehat{f}} - f, g\rangle = 0 \qquad \forall\, g \in \mathscr{S}(\mathbb{R}^n)\,.$$

Noting that Proposition 3.4 implies that $\check{\widehat{f}} \in L^1_{\rm loc}(\mathbb{R}^n)$, Lemma 3.33 shows that $\check{\widehat{f}} - f = 0$ on the set $\big\{y \in \mathbb{R}^n \,\big|\, \check{\widehat{f}} - f$ is continuous at $y\big\}$. The theorem is then concluded since $\check{\widehat{f}}$ is continuous, again by Proposition 3.4. □

Remark 3.35. Since an integrable function $f : \mathbb{R}^n \to \mathbb{R}$ must be continuous **almost everywhere** on $\mathbb{R}^n$, Theorem 3.34 implies that if $f : \mathbb{R}^n \to \mathbb{R}$ is a function such that $f, \widehat{f} \in L^1(\mathbb{R}^n)$, then $\check{\widehat{f}} = f$ almost everywhere.

Remark 3.36. In some occasions (especially in engineering applications), the Fourier transform and inverse Fourier transform of a (Schwartz) function f are defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-i2\pi x\cdot\xi}dx \quad \text{and} \quad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi)e^{i2\pi x\cdot\xi}d\xi\,. \tag{3.11}$$

Using this definition, we still have

1. $\check{\widehat{f}} = \widehat{\check{f}} = f$ for all $f \in \mathscr{S}(\mathbb{R}^n)$;
2. if $f \in L^1(\mathbb{R}^n)$ and $\widehat{f} \in L^1(\mathbb{R}^n)$, then $\check{\widehat{f}}(x) = f(x)$ for all x at which f is continuous.

3.4 The Fourier Transform of Generalized Functions

It is often required to consider the Fourier transform of functions which do not belong to $L^1(\mathbb{R}^n)$. For example, the **normalized sinc function** $\text{sinc} : \mathbb{R} \to \mathbb{R}$ defined by

$$\text{sinc}(x) = \begin{cases} \dfrac{\sin(\pi x)}{\pi x} & \text{if } x \neq 0\,, \\ 1 & \text{if } x = 0\,, \end{cases} \tag{3.12}$$

does not belong to $L^1(\mathbb{R})$ but it is a very important function in the study of signal processing.

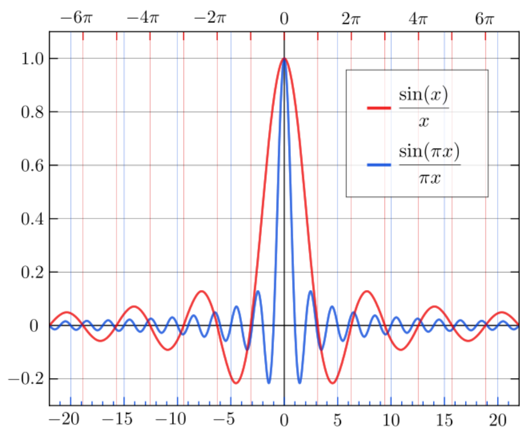

Figure 2: The graphs of unnormalized and normalized sinc functions (from wiki)

Moreover, there are "functions" that are not even functions in the traditional sense. For example, in physics and engineering applications the Dirac delta "function" δ is defined as the "function" which validates the relation

$$\int_{\mathbb{R}^n} \delta(x)\phi(x)\,dx = \phi(0) \qquad \forall\, \phi \in \mathscr{C}(\mathbb{R}^n)$$

In fact, there is no function (in the traditional sense) satisfying the property given above. Can we take the Fourier transform of those "functions" as well? To understand this topic better, it is required to study the theory of distributions.

The fundamental idea of the theory of distributions (generalized functions) is to identify a function v defined on $\mathbb{R}^n$ with the family of its integral averages

$$v \approx \int_{\mathbb{R}^n} v(x)\phi(x)\,dx \qquad \forall\, \phi \in \mathscr{C}_c^\infty(\mathbb{R}^n)\,,$$

where $\mathscr{C}_c^\infty(\mathbb{R}^n)$ denotes the collection of $\mathscr{C}^\infty$-functions with compact support, and is often denoted by $\mathcal{D}(\mathbb{R}^n)$ in the theory of distributions. Note that this makes sense for any locally integrable function v, and $\mathcal{D}(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n)$.

To understand the meaning of distributions, let us turn to a situation in physics: measuring the temperature. To measure the temperature T at a point a, instead of outputting the exact value of $T(a)$ the thermometer instead outputs the **overall value** of the temperature near a point. In other words, the reading of the temperature is determined by a pairing of the temperature distribution with the thermometer. The role of the test function ϕ is like the thermometer used to measure the temperature.

The Fourier transform can be defined on the space of tempered distributions, a smaller class of generalized functions. A tempered distribution on $\mathbb{R}^n$ is a continuous linear functional on $\mathscr{S}(\mathbb{R}^n)$. In other words, T is a tempered distribution if

$$T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C},\ T(c\phi + \psi) = cT(\phi) + T(\psi) \text{ for all } c \in \mathbb{C} \text{ and } \phi, \psi \in \mathscr{S}(\mathbb{R}^n),$$
$$\text{and } \lim_{j\to\infty} T(\phi_j) = T(\phi) \text{ if } \{\phi_j\}_{j=1}^\infty \subseteq \mathscr{S}(\mathbb{R}^n) \text{ and } \phi_j \to \phi \text{ in } \mathscr{S}(\mathbb{R}^n).$$

The convergence in $\mathscr{S}(\mathbb{R}^n)$ is described by semi-norms, and is given in the following

Definition 3.37 (Convergence in $\mathscr{S}(\mathbb{R}^n)$)**.** For each $k \in \mathbb{N}$, define the semi-norm

$$p_k(u) = \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x\rangle^k |D^\alpha u(x)|\,,$$

where $\langle x\rangle = (1+|x|^2)^{\frac{1}{2}}$. A sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{S}(\mathbb{R}^n)$ is said to converge to u in $\mathscr{S}(\mathbb{R}^n)$ if $p_k(u_j - u) \to 0$ as $j \to \infty$ for all $k \in \mathbb{N}$.

We note that $p_k(u) \leqslant p_{k+1}(u)$, so $\{u_j\}_{j=1}^\infty \subseteq \mathscr{S}(\mathbb{R}^n)$ converges to u in $\mathscr{S}(\mathbb{R}^n)$ if $p_k(u_j - u) \to 0$ as $j \to \infty$ for $k \gg 1$. We also note that if $\{u_j\}_{j=1}^\infty$ converge to u in $\mathscr{S}(\mathbb{R}^n)$, then $\{u_j\}_{j=1}^\infty$ converges uniformly to u on $\mathbb{R}^n$.

Definition 3.38 (Tempered Distributions)**.** A linear map $T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is continuous if for each $k \in \mathbb{N}$, there exists some constant C_k such that

$$\big|\langle T, u\rangle\big| \leqslant C_k p_k(u) \quad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,,$$

where $\langle T, u\rangle \equiv T(u)$ is the usual notation for the value of T at u. The collection of continuous linear functionals on $\mathscr{S}(\mathbb{R}^n)$ is denoted by $\mathscr{S}(\mathbb{R}^n)'$. Elements of $\mathscr{S}(\mathbb{R}^n)'$ are called ***tempered distributions***.

Remark 3.39. Every L^p-function $f : \mathbb{R}^n \to \mathbb{C}$ can be viewed as a tempered distribution for all $p \in [1, \infty]$. In fact, the tempered distribution T_f associated with f is defined by

$$T_f(\phi) = \int_{\mathbb{R}^n} f(x)\phi(x)\, dx \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,. \tag{3.13}$$

Since we have use $\langle \cdot, \cdot \rangle$ for the integral of product of functions, the value of the tempered distribution of f at ϕ is exactly $\langle f, \phi \rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. This should explain the use of the notation $\langle T, \phi \rangle$.

Now we show that T_f given by (3.13) is indeed a tempered distribution. Let $\phi \in \mathscr{S}(\mathbb{R}^n)$ be given. Then $\|\phi\|_{L^\infty(\mathbb{R}^n)} \leqslant p_k(\phi)$ for all $k \in \mathbb{N}$, while for $1 \leqslant q < \infty$ and $k > \dfrac{n}{q}$,

$$\begin{aligned}\|\phi\|_{L^q(\mathbb{R}^n)} &= \Big(\int_{\mathbb{R}^n} \langle x \rangle^{-kq} \big[\langle x \rangle^k |\phi(x)| \big]^q\, dx \Big)^{\frac{1}{q}} \leqslant \Big(\int_{\mathbb{R}^n} \langle x \rangle^{-kq}\, dx \Big)^{\frac{1}{q}} p_k(\phi) \\ &\leqslant \Big(\omega_{n-1} \int_0^\infty (1+r^2)^{-\frac{kq}{2}} r^{n-1} dr \Big)^{\frac{1}{q}} p_k(\phi)\,.\end{aligned}$$

Note that $\displaystyle\int_0^\infty (1+r^2)^{-\frac{kq}{2}} r^{n-1} dr < \infty$ if $k > \dfrac{n}{q}$; thus for all $q \in [1, \infty]$, there exists $C_{k,q,n} > 0$ such that

$$\|\phi\|_{L^q(\mathbb{R}^n)} \leqslant C_{k,q,n} p_k(\phi) \qquad \forall\, k > \frac{n}{q}\,. \tag{3.14}$$

Therefore, if $f \in L^p(\mathbb{R}^n)$,

$$\big| \langle f, \phi \rangle \big| \leqslant \|f\|_{L^p(\mathbb{R}^n)} \|\phi\|_{L^{p'}(\mathbb{R}^n)} \leqslant C_{k,p',n} \|f\|_{L^p(\mathbb{R}^n)} p_k(\phi) \qquad \forall\, k \gg 1$$

which shows that $T_f \in \mathscr{S}(\mathbb{R}^n)'$ if $f \in L^p(\mathbb{R}^n)$.

Example 3.40. Let $B > 0$ be given. For each sequence $\{c_k\}_{k \in \mathbb{Z}} \in \ell^p$ $\Big($that is, $\sum\limits_{k=-\infty}^{\infty} |c_k|^p < \infty\Big)$, we are interested in the function

$$f(x) = \sum_{k=-\infty}^{\infty} c_k \text{sinc}(2Bx - k)\,.$$

We note that if $1 \leqslant p < \infty$, the fact that

$$\big|\text{sinc}(2Bx - k)\big| \leqslant \min\Big\{1, \frac{1}{\pi|2Bx - k|}\Big\}$$

implies that the series $\sum\limits_{k=-\infty}^{\infty} c_k \text{sinc}(2Bx - k)$ converges pointwise on $\mathbb{R}$. Therefore, $f(x)$ is a well-defined function. We would like to know if f defines a tempered distribution.

Define $T : \mathscr{S}(\mathbb{R}) \to \mathbb{R}$ (or $\mathbb{C}$) by

$$\langle T, \phi \rangle = \frac{1}{2B} \sum_{k=-\infty}^{\infty} c_k \big\langle \text{sinc}, d_{2B} \tau_{\frac{-k}{2B}} \phi \big\rangle\,. \tag{3.15}$$

We would like to show that $T \in \mathscr{S}(\mathbb{R})'$. To see this, we note that if $k \neq 0$, using the Plancherel formula and integrating by parts,

$$\begin{aligned}\big\langle \text{sinc}, d_{2B} \tau_{\frac{-k}{2B}} \phi \big\rangle &= \big\langle \widecheck{\text{sinc}}, \widehat{d_{2B} \tau_{\frac{-k}{2B}} \phi} \big\rangle = 2B \big\langle \Pi, d_{\frac{1}{2B}} \widehat{\tau_{\frac{-k}{2B}} \phi} \big\rangle = \big\langle d_{2B} \Pi, \widehat{\phi}(\xi) e^{\frac{\pi i k \xi}{B}} \big\rangle \\ &= \int_{-B}^{B} \widehat{\phi}(\xi) e^{\frac{\pi i k \xi}{B}}\, d\xi = \frac{B}{\pi i k} \Big[\widehat{\phi}(\xi) e^{\frac{\pi i k \xi}{B}} \Big|_{\xi=-B}^{\xi=B} - \int_{-B}^{B} \widehat{\phi}'(\xi) e^{\frac{\pi i k \xi}{B}}\, d\xi \Big] \\ &= \frac{B}{\pi i k} \Big[(-1)^k \big(\widehat{\phi}(B) - \widehat{\phi}(-B) \big) - \int_{-B}^{B} \widehat{\phi}'(\xi) e^{\frac{\pi i k \xi}{B}}\, d\xi \Big]\,.\end{aligned}$$

By the Cauchy-Schwarz inequality, the Plancherel formula and Corollary 3.10, we find that for $\ell \in \mathbb{N}$,

$$
\begin{aligned}
\Big|\int_{-B}^{B} \widehat{\phi}'(\xi) e^{\frac{\pi i k \xi}{B}}\, d\xi\Big| &\leqslant \sqrt{2B}\|\widehat{\phi}'\|_{L^2(\mathbb{R})} = \frac{\sqrt{2B}}{\pi}\Big(\int_{\mathbb{R}} \big|\widehat{\pi i x \phi(x)}\big|^2(\xi)\, d\xi\Big)^{\frac{1}{2}} \\
&\leqslant \sqrt{2B}\Big(\int_{\mathbb{R}} x^2 \phi(x)^2\, dx\Big)^{\frac{1}{2}} \leqslant \sqrt{2B}\Big(\int_{\mathbb{R}} \langle x\rangle^{-2}\langle x\rangle^{2\ell+2}\phi(x)^2\, dx\Big)^{\frac{1}{2}} \\
&\leqslant \sqrt{2B}\Big(p_{\ell+1}(\phi)^2 \int_{\mathbb{R}} \langle x\rangle^{-2}\, dx\Big)^{\frac{1}{2}} \leqslant C_\ell p_{\ell+1}(\phi) \qquad \forall\, \ell \in \mathbb{N}\,.
\end{aligned}
$$

Therefore, by the fact that $\|\widehat{\phi}\|_{L^\infty(\mathbb{R})} \leqslant \|\phi\|_{L^1(\mathbb{R})}$, using (3.14) we find that if $k \neq 0$,

$$
\big|\big\langle \text{sinc}, d_{2B}\tau_{\frac{-k}{2B}}\phi\big\rangle\big| \leqslant \frac{B}{\pi|k|}\Big[2\|\phi\|_{L^1(\mathbb{R})} + \Big|\int_{-B}^{B} \widehat{\phi}'(\xi) e^{\frac{\pi i k \xi}{B}}\, d\xi\Big|\Big] \leqslant \frac{1}{|k|} C_\ell p_\ell(\phi) \qquad \forall\, \ell \gg 1\,.
$$

On the other hand, using (3.14) again we have

$$
\big|\big\langle \text{sinc}, d_{2B}\tau_{\frac{0}{2B}}\phi\big\rangle\big| \leqslant 2B\|\phi\|_{L^1(\mathbb{R})} \leqslant C_\ell p_\ell(\phi) \qquad \forall\, \ell \gg 1\,.
$$

Therefore, if $1 \leqslant p < \infty$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$, the Hölder inequality shows that

$$
\big|\langle T, \phi\rangle\big| \leqslant \frac{1}{2B}\Big(\sum_{k=-\infty}^{\infty} |c_k|^p\Big)^{\frac{1}{p}}\Big(\sum_{k=-\infty}^{\infty} \big|\big\langle \text{sinc}, d_{2B}\tau_{\frac{-k}{2B}}\phi\big\rangle\big|^q\Big)^{\frac{1}{q}} \leqslant C_\ell p_\ell(\phi) \qquad \forall\, \ell \gg 1\,.
$$

From now on, we identify f with the tempered distribution T_f if $f \in L^p(\mathbb{R}^n)$. For example, if $T \in \mathscr{S}(\mathbb{R}^n)'$ and $f : \mathbb{R}^n \to \mathbb{C}$ is bounded or integrable, we say that $T = f$ in $\mathscr{S}(\mathbb{R}^n)'$ if $T = T_f$, where T_f is the tempered distribution associated with the function f.

Remark 3.41. Let $f(x) = e^{x^4} \in L^1_{\text{loc}}(\mathbb{R}^n)$. Then $\langle T_f, e^{-x^2}\rangle = \infty$. Therefore, being in $L^1_{\text{loc}}(\mathbb{R}^n)$ is not good enough to generate elements in $\mathscr{S}(\mathbb{R}^n)'$, and it requires that $|f(x)| \leqslant C(1 + |x|^N)$ for any N. In such a case, $T_f \in \mathscr{S}(\mathbb{R}^n)'$ is well-defined.

Example 3.42 (Dirac delta function)**.** Consider the map $\delta : \mathscr{C}(\mathbb{R}^n) \to \mathbb{R}$ defined by $\delta(\phi) = \phi(0)$. Since the convergence in $\mathscr{S}(\mathbb{R}^n)$ implies pointwise convergence, we immediately conclude that $\delta \in \mathscr{S}(\mathbb{R}^n)'$.

As shown in the example above, a tempered distribution might not be defined in the pointwise sense. Therefore, how to define usual operations such as translation, dilation, and reflection on generalized functions should be answered prior to define the Fourier transform of tempered distributions. For completeness, let us start from providing the definitions of translation, dilation and reflection operators.

Definition 3.43 (Translation, dilation, and reflection)**.** Let $f : \mathbb{R}^n \to \mathbb{C}$ be a function.

1. For $h \in \mathbb{R}^n$, the translation operator τ_h maps f to $\tau_h f$ given by $(\tau_h f)(x) = f(x - h)$.

2. For $\lambda > 0$, the dilation operator $d_\lambda : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ maps f to $d_\lambda f$ given by $(d_\lambda f)(x) = f(\lambda^{-1}x)$.

3. The Reflection operator $\widetilde{\ }$ maps f to $\widetilde{f}$ given by $\widetilde{f}(x) = f(-x)$.

Now suppose that $T \in \mathscr{S}(\mathbb{R}^n)'$. We expect that $\tau_h T$, $d_\lambda T$ and $\widetilde{T}$ are also tempered distributions, so we need to provide the values of $\langle \tau_h T, \phi \rangle$, $\langle d_\lambda T, \phi \rangle$ and $\langle \widetilde{T}, \phi \rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. If $T = T_f$ is the tempered distribution associated with $f \in L^1(\mathbb{R}^n)$, then for $\phi \in \mathscr{S}(\mathbb{R}^n)$, the change of variable formula implies that

$$
\begin{aligned}
\langle \tau_h f, g \rangle &= \int_{\mathbb{R}^n} f(x-h) g(x)\, dx = \int_{\mathbb{R}^n} f(x) g(x+h)\, dx = \langle f, \tau_{-h} g \rangle, \\
\langle d_\lambda f, g \rangle &= \int_{\mathbb{R}^n} f(\lambda^{-1} x) g(x)\, dx = \int_{\mathbb{R}^n} f(x) g(\lambda x) \lambda^n\, dx = \langle f, \lambda^n d_{\lambda^{-1}} g \rangle, \\
\langle \widetilde{f}, g \rangle &= \int_{\mathbb{R}^n} f(-x) g(x)\, dx = \int_{\mathbb{R}^n} f(x) g(-x)\, dx = \langle f, \widetilde{g} \rangle.
\end{aligned}
$$

The computations above motivate the following

Definition 3.44. Let $h \in \mathbb{R}^n$, $\lambda > 0$, and τ_h and d_λ be the translation and dilation operator given in Definition 3.43. For $T \in \mathscr{S}(\mathbb{R}^n)'$, $\tau_h T$, $d_\lambda T$ and $\widetilde{T}$ are the tempered distributions defined by

$$
\langle \tau_h T, \phi \rangle = \langle T, \tau_{-h} \phi \rangle, \quad \langle d_\lambda T, \phi \rangle = \langle T, \lambda^n d_{\lambda^{-1}} \phi \rangle \quad \text{and} \quad \langle \widetilde{T}, \phi \rangle = \langle T, \widetilde{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).
$$

From the experience of defining the translation, dilation and reflection of tempered distribution, now we can talk about how to defined Fourier transform of tempered distributions. Recall that in Lemma 3.26 we have established that

$$
\langle \widehat{f}, g \rangle = \langle f, \widehat{g} \rangle \quad \text{and} \quad \langle \check{f}, g \rangle = \langle f, \check{g} \rangle \qquad \forall\, f \in L^1(\mathbb{R}^n), g \in \mathscr{S}(\mathbb{R}^n).
$$

Since the identities above hold for all L^1-functions f (and L^1-functions corresponds to tempered distributions T_f through (3.13)), we expect that the Fourier transform of tempered distributions has to satisfy the identities above as well. Let $T \in \mathscr{S}(\mathbb{R}^n)'$ be given, and define $\widehat{T} : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ by

$$
\widehat{T}(\phi) = \langle \widehat{T}, \phi \rangle \equiv \langle T, \widehat{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n). \tag{3.16}
$$

Since for $k \in \mathbb{N}$,

$$
\begin{aligned}
p_k(\widehat{\phi}) &= \sup_{\xi \in \mathbb{R}^n, |\alpha| \leqslant k} \langle \xi \rangle^k \big| D^\alpha \widehat{\phi}(\xi) \big| = \sup_{\xi \in \mathbb{R}^n, |\alpha| \leqslant k} \langle \xi \rangle^k \big| \widehat{x^\alpha \phi(x)} \big| (\xi) \\
&\leqslant C \sum_{|\beta| = k} \sup_{\xi \in \mathbb{R}^n, |\alpha| \leqslant k} \Big| \mathscr{F}\big[D^\beta \big(x^\alpha \phi(x) \big) \big](\xi) \Big| \leqslant C \sum_{|\alpha| \leqslant k, |\beta| = k} \big\| D^\beta \big(x^\alpha \phi(x) \big) \big\|_{L^1(\mathbb{R}^n)} \\
&\leqslant C \sum_{|\alpha| \leqslant k, |\beta| = k} \big\| \langle x \rangle^{-\mathrm{n}-1} \big\|_{L^1(\mathbb{R}^n)} \big\| \langle x \rangle^{\mathrm{n}+1} D^\beta \big(x^\alpha \phi(x) \big) \big\|_{L^\infty(\mathbb{R}^n)} \leqslant C p_{\mathrm{n}+k+1}(\phi)
\end{aligned}
$$

for some constant $C > 0$, by the fact that for each $k \in \mathbb{N}$ there exists $C_k > 0$ such that $\big| \langle T, \phi \rangle \big| \leqslant C_k p_k(\phi)$ for all $k \in \mathbb{N}$, we find that

$$
\big| \langle \widehat{T}, \phi \rangle \big| = \big| \langle T, \widehat{\phi} \rangle \big| \leqslant C_k p_k(\widehat{\phi}) \leqslant \widetilde{C}_k p_{k+n+1}(\phi) \qquad \forall\, k \in \mathbb{N}
$$

for some constant $\widetilde{C}_k > 0$. Therefore, $\widehat{T}$ defined by (3.16) is a tempered distribution. Similarly, $\check{T} : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ defined by $\langle \check{T}, \phi \rangle = \langle T, \check{\phi} \rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$ is also a tempered distribution. The discussion above leads to the following

Definition 3.45. Let $T \in \mathscr{S}(\mathbb{R}^n)'$. The Fourier transform of T and the inverse Fourier transform of T, denoted by $\widehat{T}$ and $\check{T}$ respectively, are tempered distributions satisfying

$$\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle \quad \text{and} \quad \langle \check{T}, \phi \rangle = \langle T, \check{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

In other words, if $T \in \mathscr{S}(\mathbb{R}^n)'$, then $\widehat{T}, \check{T} \in \mathscr{S}(\mathbb{R}^n)'$ as well and the actions of $\widehat{T}$, $\check{T}$ on $\phi \in \mathscr{S}(\mathbb{R}^n)$ are given in the relations above.

Example 3.46 (The Fourier transform of the Dirac delta function)**.** Consider the Dirac delta function $\delta : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ defined in Example 3.42. Then for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle \delta, \widehat{\phi} \rangle = \widehat{\phi}(0) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot 0}\, dx = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x)\, dx = \langle \frac{1}{\sqrt{2\pi}^n}, \phi \rangle\,;$$

thus the Fourier transform of the Dirac delta function is a constant function and $\widehat{\delta}(\xi) = \dfrac{1}{\sqrt{2\pi}^n}$. Similarly, $\check{\delta}(\xi) = \dfrac{1}{\sqrt{2\pi}^n}$, so $\widehat{\delta} = \check{\delta}$.

One can also consider the Dirac delta function at point ω, denoted by δ_ω or $\delta(\cdot - \omega)(\equiv \tau_\omega \delta)$, given by

$$\delta_\omega(\phi) = \phi(\omega) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)$$

which is often written as $\displaystyle\int_{\mathbb{R}^n} \delta(x-\omega)\phi(x)\, dx = \langle \delta_\omega, \phi \rangle$. Then for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle \delta_\omega, \widehat{\phi} \rangle = \widehat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot\omega}\, dx = \langle \frac{e^{-ix\cdot\omega}}{\sqrt{2\pi}^n}, \phi \rangle \equiv \langle \widehat{\delta_\omega}, \phi \rangle\,;$$

thus the Fourier transform of the Dirac delta function at point ω is the function $\widehat{\delta_\omega}(\xi) = \dfrac{e^{-i\xi\cdot\omega}}{\sqrt{2\pi}^n}$. The inverse Fourier transform of δ_ω can be computed in the same fashion and we have $\check{\delta_\omega}(\xi) = \dfrac{e^{i\xi\cdot\omega}}{\sqrt{2\pi}^n}$. We note that $\check{\delta_\omega} = \widehat{\widehat{\widehat{\delta_\omega}}} = \widehat{\widehat{\delta}}_\omega$.

Symbolically, "assuming" that $\delta_\omega(\phi) = \phi(\omega)$ for all continuous function ϕ,

$$\widehat{\delta_\omega}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \delta_\omega(x) e^{-ix\cdot\xi}\, dx = \frac{1}{\sqrt{2\pi}^n} e^{-ix\cdot\xi}\Big|_{x=\omega} = \frac{e^{-i\xi\cdot\omega}}{\sqrt{2\pi}^n}$$

and

$$\check{\delta_\omega}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \delta_\omega(x) e^{ix\cdot\xi}\, dx = \frac{1}{\sqrt{2\pi}^n} e^{ix\cdot\xi}\Big|_{x=\omega} = \frac{e^{i\xi\cdot\omega}}{\sqrt{2\pi}^n}$$

Example 3.47 (The Fourier transform of $e^{ix\cdot\omega}$)**.** By "definition", for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle e^{ix\cdot\omega}, \widehat{\phi} \rangle = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}} e^{i\xi\cdot\omega} \Big(\int_{\mathbb{R}} \phi(x) e^{-ix\cdot\xi}\, dx \Big) d\xi = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} \phi(x) e^{-i(x-\omega)\cdot\xi}\, dx \Big) d\xi\,.$$

Noting that the Fourier inversion formula implies that

$$\phi(\omega) = \check{\widehat{\phi}}(\omega) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} \phi(x) e^{-i(x-\omega)\cdot\xi}\, dx \Big) d\xi\,;$$

thus

$$\langle e^{ix\cdot\omega}, \widehat{\phi} \rangle = \sqrt{2\pi}^n \phi(\omega) = \langle \sqrt{2\pi}^n \delta_\omega, \phi \rangle\,.$$

Therefore, the Fourier transform of the function $s(x) = e^{ix\cdot\omega}$ is $\sqrt{2\pi}^n\delta_\omega$. where δ_ω is the Dirac delta function at point ω. We note that this result also implies that

$$\check{\hat{\delta}}_\omega = \delta_\omega \qquad \forall\,\omega \in \mathbb{R}^n\,.$$

Similarly, $\hat{\check{\delta}}_\omega = \delta_\omega$ for all $\omega \in \mathbb{R}^n$; thus the Fourier inversion formula is also valid for the Dirac δ function.

Example 3.48 (The Fourier Transform of the Sine function)**.** Let $s(x) = \sin\omega x$, where ω denotes the frequency of this sine wave. Since $\sin\omega x = \dfrac{e^{i\omega x} - e^{-i\omega x}}{2i}$, we conclude that the Fourier transform of $s(x) = \sin\omega x$ is

$$\frac{\sqrt{2\pi}}{2i}\big(\delta_\omega - \delta_{-\omega}\big)\,.$$

Theorem 3.49. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$. Then $\check{\hat{T}} = \hat{\check{T}} = T$.*

Proof. To see that $\check{\hat{T}}$ and T are the same tempered distribution, we need to show that $\langle \check{\hat{T}}, \phi\rangle = \langle T, \phi\rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. Nevertheless, by the defintion of the Fourier transform and the inverse Fourier transform of tempered distributions,

$$\langle \check{\hat{T}}, \phi\rangle = \langle \hat{T}, \check{\phi}\rangle = \langle T, \hat{\check{\phi}}\rangle = \langle T, \phi\rangle \qquad \forall\,\phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

That $\hat{\check{T}} = T$ can be proved in the same fashion. □

Theorem 3.50. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$. Then*

$$\langle \widehat{\tau_h T}, \phi\rangle = \langle \hat{T}, \phi(\xi)e^{-i\xi\cdot h}\rangle\,, \quad \langle \widehat{d_\lambda T}, \phi\rangle = \langle \hat{T}, d_\lambda\phi\rangle \quad \text{and} \quad \langle \hat{\hat{T}}, \phi\rangle = \langle \check{T}, \phi\rangle \qquad \forall\,\phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

A short-hand notation for identities above are $\widehat{\tau_h T}(\xi) = \hat{T}(\xi)e^{-i\xi\cdot h}$, $\widehat{d_\lambda T}(\xi) = \lambda^n\hat{T}(\lambda\xi)$, and $\hat{\hat{T}}(\xi) = \check{T}(\xi)$.

Proof. Let $\phi \in \mathscr{S}(\mathbb{R}^n)$. For $h \in \mathbb{R}^n$, define $\phi_h(x) = \phi(x)e^{-ix\cdot h}$. Then

$$(\tau_{-h}\hat{\phi})(\xi) = \hat{\phi}(\xi + h) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)e^{-ix\cdot(\xi+h)}\,dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)e^{-ix\cdot h}e^{-ix\cdot\xi}\,dx = \widehat{\phi_h}(\xi)\,.$$

By the definition of the Fourier transform of tempered distribution and the translation operator,

$$\langle \widehat{\tau_h T}, \phi\rangle = \langle T, \tau_{-h}\hat{\phi}\rangle = \langle T, \widehat{\phi_h}\rangle = \langle \hat{T}(x), \phi(x)e^{-ix\cdot h}\rangle = \langle \hat{T}(\xi), \phi(\xi)e^{-i\xi\cdot h}\rangle\,.$$

On the other hand, for $\lambda > 0$,

$$(d_{\lambda^{-1}}\hat{\phi})(\xi) = \hat{\phi}(\lambda\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)e^{-ix\cdot(\lambda\xi)}\,dx = \lambda^{-n}\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi\big(\frac{x}{\lambda}\big)e^{-ix\cdot\xi}\,dx = \lambda^{-n}\widehat{d_\lambda\phi}(\xi)\,.$$

Therefore,

$$\langle \widehat{d_\lambda T}, \phi\rangle = \langle T, \lambda^n d_{\lambda^{-1}}\hat{\phi}\rangle = \langle T, \widehat{d_\lambda\phi}\rangle = \langle \hat{T}, d_\lambda\phi\rangle = \langle \lambda^n d_{\lambda^{-1}}\hat{T}, \phi\rangle\,.$$

The identity $\langle \hat{\hat{T}}, \phi\rangle = \langle \check{T}, \phi\rangle$ follows from that $\tilde{\hat{\phi}} = \check{\phi}$, and the detail proof is left to the readers. □

Remark 3.51. One can check (using the change of variable formula) that $\widehat{\tau_h f}(\xi) = \widehat{f}(\xi)e^{-i\xi\cdot h}$ and $\widehat{d_\lambda f}(\xi) = \lambda^n \widehat{f}(\lambda\xi)$ if $f \in L^1(\mathbb{R}^n)$.

Example 3.52 (The Fourier Transform of the sinc function)**.** The rect/rectangle function, also called the gate function or windows function, is a function $\Pi : \mathbb{R} \to \mathbb{R}$ defined by

$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < 1\,, \\ 0 & \text{if } |x| \geqslant 1\,. \end{cases}$$

Since $\Pi \in L^1(\mathbb{R})$, we can compute its (inverse) Fourier transform in the usual way, and we have

$$\widehat{\Pi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Pi(x) e^{-ix\xi}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-ix\xi}\, dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-ix\xi}}{-i\xi}\Big|_{x=-1}^{x=1} = \sqrt{\frac{2}{\pi}} \frac{\sin\xi}{\xi} \qquad \forall\, \xi \neq 0$$

and $\widehat{\Pi}(0) = \sqrt{\dfrac{2}{\pi}}$. Define the ***unnormalized sinc function*** $\operatorname{sinc}(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0\,. \end{cases}$ Then $\widehat{\Pi}(\xi) = \sqrt{\dfrac{2}{\pi}} \operatorname{sinc}(\xi)$. Similar computation shows that $\check{\Pi}(\xi) = \widehat{\Pi}(\xi) = \sqrt{\dfrac{2}{\pi}} \operatorname{sinc}(\xi)$.

Even though the sinc function is not integrable, we can apply Theorem 3.49 and see that

$$\widehat{\operatorname{sinc}}(\xi) = \widecheck{\operatorname{sinc}}(\xi) = \sqrt{\frac{2}{\pi}}\, \Pi(\xi) \qquad \forall\, \xi \in \mathbb{R}\,.$$

Next we define the convolution of a tempered distribution and a Schwartz function. Before proceeding, we note that if $f, g \in \mathscr{S}(\mathbb{R}^n)$, then

$$\begin{aligned} \langle f * g, \phi\rangle &= \int_{\mathbb{R}^n} (f*g)(x)\phi(x)\, dx = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(y) g(x-y)\, dy \Big) \phi(x)\, dx \\ &= \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} g(x-y)\phi(x)\, dx \Big) f(y)\, dy \\ &= \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \widetilde{g}(y-x)\phi(x)\, dx \Big) f(y)\, dy = \langle f, \widetilde{g} * \phi\rangle\,. \end{aligned}$$

The change of variable formula implies that

$$\begin{aligned} (\widetilde{g} * \phi)(y) &= \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \widetilde{g}(x)\phi(y-x)\, dx \Big) f(y)\, dy \\ &= \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \widetilde{g}(-x)\phi(y+x)\, dx \Big) f(y)\, dy \\ &= \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} g(x)\widetilde{\phi}(-y-x)\, dx \Big) f(y)\, dy = (g * \widetilde{\phi})(-y) = \widetilde{g * \widetilde{\phi}}(y)\,; \end{aligned}$$

thus

$$\langle f * g, \phi\rangle = \langle f, \widetilde{g} * \phi\rangle = \langle f, \widetilde{g * \widetilde{\phi}}\rangle = \langle \widetilde{f}, g * \widetilde{\phi}\rangle\,.$$

The identity above serves as the origin of the convolution of a tempered distribution and a Schwartz function.

Definition 3.53 (Convolution)**.** Let $T \in \mathscr{S}(\mathbb{R}^n)'$ and $f \in \mathscr{S}(\mathbb{R}^n)$. The convolution of T and f, denoted by $T * f$, is the tempered distribution given by

$$\langle T * f, \phi\rangle = \langle T, \widetilde{g} * \phi\rangle = \langle \widetilde{T}, f * \widetilde{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

where $\widetilde{T}$ is the tempered distribution given in Definition 3.44.

Remark 3.54. 1. If $S \in \mathscr{S}(\mathbb{R}^n)'$ satisfies that $S * \phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$, we can also define the convolution of T and S by

$$\langle T * S, \phi \rangle = \langle \widetilde{T}, S * \widetilde{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

In other words, it is possible to define the convolution of two tempered distributions.

2. Suppose that $S \in \mathscr{S}(\mathbb{R}^n)'$ satisfies that $S*\phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$ so that $T*S \in \mathscr{S}(\mathbb{R}^n)$ is well-defined for all $T \in \mathscr{S}(\mathbb{R}^n)'$. Then

$$\langle \widehat{T * S}, \phi \rangle = \langle T * S, \widehat{\phi} \rangle = \langle \widetilde{T}, S * \widetilde{\widehat{\phi}} \rangle = \langle \check{\widetilde{T}}, \widehat{S * \widetilde{\phi}} \rangle = \langle \widehat{T}, \widehat{S}\widehat{\widetilde{\phi}} \rangle$$

Similar to Theorem 3.23 and Corollary 3.24, the product and the convolutions of functions are related under Fourier transform.

Theorem 3.55. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$ and $f \in \mathscr{S}(\mathbb{R}^n)$. Then*

$$\langle \widehat{T * f}, \phi \rangle = \langle \widehat{T}, \widehat{f}\phi \rangle \quad \textit{and} \quad \langle \widecheck{T * f}, \phi \rangle = \langle \check{T}, \check{f}\phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

and

$$\langle \widehat{fT}, \phi \rangle = \langle \widehat{T} * \widehat{f}, \phi \rangle \quad \textit{and} \quad \langle \widecheck{fT}, \phi \rangle = \langle \check{T} * \check{f}, \phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

where $fT \in \mathscr{S}(\mathbb{R}^n)'$ is defined by $\langle fT, \phi \rangle = \langle T, f\phi \rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. A short-hand notation for the identities above are $\widehat{T * f} = \widehat{f}\widehat{T}$, $\widecheck{T * f} = \check{f}\check{T}$, $\widehat{fT} = \widehat{T} * \widehat{f}$ and $\widecheck{fT} = \check{T} * \check{f}$ in $\mathscr{S}(\mathbb{R}^n)'$.

Proof. By Theorem 3.23,

$$\langle \widehat{T * f}, \phi \rangle = \langle T * f, \widehat{\phi} \rangle = \langle \widetilde{T}, f * \widetilde{\widehat{\phi}} \rangle = \langle \widetilde{T}, f * \check{\widehat{\phi}} \rangle = \langle \check{\widetilde{T}}, \mathscr{F}(f * \check{\phi}) \rangle = \langle \widehat{T}, \widehat{f}\phi \rangle$$

and by the definition of the convolution of tempered distributions and Schwartz functions,

$$\langle \widehat{fT}, \phi \rangle = \langle T, f\widehat{\phi} \rangle = \langle \widehat{T}, \mathscr{F}^*(f\widehat{\phi}) \rangle = \langle \widehat{T}, \check{f} * \phi \rangle = \langle \widehat{T}, \widetilde{\widehat{f}} * \phi \rangle = \langle \widehat{T} * \widehat{f}, \phi \rangle\,.$$

The counterpart for the inverse Fourier transform can be proved similarly. □

4 Application on Signal Processing

In the study of signal processing, the Fourier transform and the inverse Fourier transform are often defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i2\pi x\cdot\xi} dx \quad \text{and} \quad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi) e^{i2\pi x\cdot\xi} d\xi \qquad \forall\, f \in L^1(\mathbb{R}^n)\,. \tag{3.11}$$

Then for $T \in \mathscr{S}(\mathbb{R}^n)'$, the Fourier transform of T is defined again by

$$\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

We also note that the definitions of the translation, dilation, and reflection of tempered distributions are independent of the Fourier transform, and are still defined by

$$\langle \tau_h T, \phi \rangle = \langle T, \tau_{-h}\phi \rangle\,, \quad \langle d_\lambda T, \phi \rangle = \langle T, \lambda^n d_{\lambda^{-1}}\phi \rangle \quad \text{and} \quad \langle \widetilde{T}, \phi \rangle = \langle T, \widetilde{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

Concerning the convolution, when the Fourier transform is given by (3.11), we usually consider the $*$ convolution operator

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy \qquad \forall\, f, g \in L^1(\mathbb{R}^n)\,.$$

instead of $\ast$ convolution operators. The convolution of T and $f \in \mathscr{S}(\mathbb{R}^n)$ is defined by

$$\langle T * f, \phi \rangle = \langle T, \widetilde{f} * \phi \rangle = \langle \widetilde{T}, f * \widetilde{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

Then similar to Theorem 3.49, 3.50, and 3.55, we have

1. $\check{\widehat{T}} = \widehat{\check{T}} = T$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.
2. $\widehat{\tau_h T}(\xi) = \widehat{T}(\xi)e^{-2\pi i \xi\cdot h}$, $\widehat{d_\lambda T}(\xi) = \lambda^n \widehat{T}(\lambda\xi)$, and $\widehat{\widehat{T}}(\xi) = \check{T}(\xi)$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.
3. $\widehat{T * f} = \widehat{T}\widehat{f}$ and $\widehat{fT} = \widehat{f} * \widehat{T}$ for all $f \in \mathscr{S}(\mathbb{R}^n)$ and $T \in \mathscr{S}(\mathbb{R}^n)'$. Moreover, if $S \in \mathscr{S}(\mathbb{R}^n)'$ has the property that $S * \phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\phi \in \mathbb{R}^n$, then $\widehat{T * S} = \widehat{T}\widehat{S}$ in $\mathscr{S}(\mathbb{R}^n)'$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.

Moreover,

1. $\widehat{\delta} = \check{\delta} = 1$ in $\mathscr{S}(\mathbb{R}^n)'$, and $\widehat{\delta_h}(\xi) = \widehat{\tau_h \delta}(\xi) = \widecheck{\delta_{-h}} = \widecheck{\tau_{-h}\delta} = e^{-2\pi i h\cdot\xi}$ in $\mathscr{S}(\mathbb{R}^n)'$ for all $h \in \mathbb{R}^n$.
2. By Euler's identity, $\widehat{\cos(2\pi\omega x)}(\xi) = \dfrac{1}{2}(\delta_\omega + \delta_{-\omega})$ and $\widehat{\sin(2\pi\omega x)}(\xi) = \dfrac{1}{2i}(\delta_\omega - \delta_{-\omega})$.
3. $\delta * \delta = \delta$, and $\delta_a * \delta_b = \delta_{a+b}$ for all $a, b \in \mathbb{R}^n$.
4. $\delta * \phi = \phi$ and $(\delta_a * \phi)(x) = \phi(x-a)$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$.
5. Re-define the rect function $\Pi : \mathbb{R} \to \mathbb{R}$ by
$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < \dfrac{1}{2}\,, \\ 0 & \text{if } |x| \geqslant \dfrac{1}{2}\,. \end{cases} \tag{4.1}$$
Then $\widehat{\Pi}(\xi) = \check{\Pi}(\xi) = \operatorname{sinc}(\xi)$, where sinc is the normalized sinc function given by (3.12).
6. Let $\Lambda : \mathbb{R} \to \mathbb{R}$ be the triangle function define by
$$\Lambda(x) = \begin{cases} 1 - |x| & \text{if } |x| < 1\,, \\ 0 & \text{if } |x| \geqslant 1\,. \end{cases}$$
Then by the fact that Λ is an even function, if $\xi \neq 0$,
$$\begin{aligned} \widehat{\Lambda}(\xi) &= 2\int_0^1 (1-x)\cos(2\pi x\xi)\,dx = 2\Big[(1-x)\frac{\sin(2\pi x\xi)}{2\pi\xi}\Big|_{x=0}^{x=1} + \int_0^1 \frac{\sin(2\pi x\xi)}{2\pi\xi}\,dx\Big] \\ &= \frac{1-\cos(2\pi\xi)}{2\pi^2\xi^2} = \frac{\sin^2 \pi\xi}{\pi^2\xi^2}\,, \end{aligned}$$
while $\widehat{\Lambda}(0) = 1$. Therefore, $\widehat{\Lambda}(\xi) = \operatorname{sinc}^2(\xi)$. Using the property of convolution, we have $\Pi * \Pi = \Lambda$.

4.1 The Sampling Theorem and the Nyquist Rate

When a continuous function, $x(t)$, is sampled at a constant rate f_s samples per second（以每秒 f_s 次取樣）, there is always an unlimited number of other continuous functions that fit the same set of samples; however, only one of them is bandlimited to $\frac{1}{2}f_s$ cycles per second (hertz), which means that its Fourier transform, $X(f)$, is 0 for all $|f| \geqslant \frac{1}{2}f_s$.

Definition 4.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. f is said to be a ***bandlimited*** function if $\text{supp}(\widehat{f})$ is bounded. The ***bandwidth*** of a bandlimited function f is the number $\sup \text{supp}(\widehat{f})$. f is said to be ***timelimited*** if $\text{supp}(f)$ is bounded.

Definition 4.2. In signal processing, the ***Nyquist rate*** is twice the bandwidth of a bandlimited function or a bandlimited channel.

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 a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $\frac{1}{2B}$ seconds apart."

Let us start from the following famous Poisson summation formula to demonstrate why countable sampling is possible to reconstruct the full signal.

Lemma 4.3 (Poisson summation formula)**.** *Let the Fourier transform and the inverse Fourier transform be defined by* (3.11)*. Then*

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{k=-\infty}^{\infty} \widehat{f}(k) e^{2\pi i k x} \qquad \forall\, f \in \mathscr{S}(\mathbb{R})\,. \tag{4.2}$$

The convergences on both sides are uniform.

Proof. Let $f \in \mathscr{S}(\mathbb{R})$ be given. Then there exists $C > 0$ such that

$$\big|f(x)\big| + \big|f'(x)\big| \leqslant \frac{C}{1+|x|^2} \qquad \forall\, x \in \mathbb{R}.$$

Define $F(x) = \sum\limits_{n=-\infty}^{\infty} f(x+n)$. Then for $x \in [0,1]$,

$$\big|f(x+n)\big| + \big|f'(x+n)\big| \leqslant \frac{C}{1+n^2}\ \forall\, n \geqslant 0 \quad \text{and} \quad \big|f(x+n)\big| + \big|f'(x+n)\big| \leqslant \frac{C}{1+(n+1)^2}\ \forall\, n < 0\,.$$

By the fact that

$$\sum_{n=0}^{\infty} \frac{C}{1+n^2} < \infty \quad \text{and} \quad \sum_{n=-\infty}^{-1} \frac{C}{1+|1+n|^2} < \infty\,,$$

the Weierstrass M-test implies that the series $\sum\limits_{n=-\infty}^{\infty} f(x+n)$ and $\sum\limits_{n=-\infty}^{\infty} f'(x+n)$ both converge uniformly on $[0,1]$. Therefore, $F:[0,1]$ is differentiable. Noting that $F(x) = F(x+1)$, so F has period 1.

Since $F \in \mathscr{C}^1([0,1])$ and is periodic with period 1, Theorem 2.15 implies that

$$F(x) = \sum_{k=-\infty}^{\infty} \widehat{F}_k e^{2\pi ikx} \qquad \forall\, x \in \mathbb{R}\,, \tag{4.3}$$

where $\{\widehat{F}_k\}_{k=-\infty}^{\infty}$ are the Fourier coefficients of F defined by $\widehat{F}_k = \int_0^1 F(x)e^{-2\pi ikx}\,dx$. By the uniform convergence of $\sum\limits_{n=-\infty}^{\infty} f(x+n)$ in $[0,1]$, we find that

$$\widehat{F}_k = \sum_{n=-\infty}^{\infty} \int_0^1 f(x+n)e^{-2\pi ikx}\,dx = \sum_{n=-\infty}^{\infty} \int_n^{n+1} f(x)e^{-2\pi ik(x-n)}\,dx = \int_{\mathbb{R}} f(x)e^{-2\pi ikx}\,dx = \widehat{f}(k)\,.$$

The Poisson summation formula (4.2) then follows from (4.3) and the identity above. □

Remark 4.4. Using Definition 3.3 of the Fourier transform, for $f \in \mathscr{S}(\mathbb{R})$ one has

$$\sum_{n=-\infty}^{\infty} f(x+2n\pi) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}\,.$$

Corollary 4.5. *Let the Fourier transform and the inverse Fourier transform be defined by* (3.11)*. Then*

$$\sum_{k=-\infty}^{\infty} \widehat{f}(\xi - \frac{k}{T}) = T \sum_{n=-\infty}^{\infty} f(nT)e^{-i2\pi nT\xi} \qquad \forall\, f \in \mathscr{S}(\mathbb{R})\,. \tag{4.4}$$

Proof. For a given $g \in \mathscr{S}(\mathbb{R})$, let $h = d_\lambda g$, where d_λ is a dilation operator. Then h is also a Schwartz function, and

$$\widehat{h}(\xi) = (\lambda d_{\lambda^{-1}}\widehat{g})(\xi) = \lambda \widehat{g}(\lambda\xi)\,;$$

thus the Poisson summation formula (4.2) (with $x = 0$ and $f = g$) implies that

$$\sum_{n=-\infty}^{\infty} h(n\lambda) = \sum_{n=-\infty}^{\infty} g(n) = \sum_{k=-\infty}^{\infty} \widehat{g}(k) = \frac{1}{\lambda} \sum_{k=-\infty}^{\infty} \widehat{h}\big(\frac{k}{\lambda}\big)\,. \tag{4.5}$$

Now let $s = \tau_t h$ for some $t \in \mathbb{R}$, where τ_t is a translation operator. Then $s \in \mathscr{S}(\mathbb{R})$, and

$$\widehat{s}(\xi) = \widehat{h}(\xi)e^{-2\pi it\xi}\,.$$

Therefore, (4.5) implies that

$$\sum_{n=-\infty}^{\infty} s(t+n\lambda) = \frac{1}{\lambda} \sum_{k=-\infty}^{\infty} \widehat{s}\big(\frac{k}{\lambda}\big)e^{\frac{2\pi ikt}{\lambda}}\,. \tag{4.6}$$

Finally, for $f \in \mathscr{S}(\mathbb{R})$, let $s = \check{f}$. Then using $\check{f} = \widetilde{\widehat{f}}$, $\lambda = \dfrac{1}{T}$ and $t = -\xi$ in the identity above, we obtain that

$$\sum_{k=-\infty}^{\infty} \widehat{f}(\xi - \frac{k}{T}) = \sum_{k=-\infty}^{\infty} \widehat{f}(\xi + \frac{k}{T}) = \sum_{k=-\infty}^{\infty} \check{f}(-\xi - \frac{k}{T}) = T \sum_{k=-\infty}^{\infty} f(kT)e^{-2\pi ikT\xi}$$

which shows (4.4). □

Remark 4.6. Identity (4.4) can be shown to hold for all continuous function f satisfying

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

for some C, $\delta > 0$. Therefore, if $\widehat{f}$ has compact support, as long as the decay rate of f is bigger than 1, (4.4) is a valid identity.

A direct consequence of the corollary above is the following sampling theorem. Suppose that $f \in \mathscr{S}(\mathbb{R})$ and $\operatorname{supp}(\widehat{f}) \subseteq \big[0, \frac{1}{T}\big]$. Then (4.4) implies that

$$\widehat{f}(\xi) = \sum_{n=-\infty}^{\infty} f(nT) e^{-i2\pi nT\xi} \qquad \forall\, \xi \in \Big[0, \frac{1}{T}\Big] .$$

This shows that if $\widehat{f}$ has compact support in $\big[0, \frac{1}{T}\big]$, f can be reconstructed based on partial knowledge of f, namely $f(nT)$.

Recall that the Fourier transform of the sine wave with frequency ω is "supported" in a symmetric domain $\{\omega, -\omega\}$. Therefore, in reality it is better to assume that the Fourier transform of a bandlimited signal is supported in a symmetric domain $[-B, B]$. In such a case we need $\big|\xi - \frac{k}{T}\big| \geqslant B$ for all $k \in \mathbb{Z}\backslash\{0\}$ and $\xi \in [-B, B]$, where T is the sampling frequency, in order to make use of (4.4) to gain all the information of the Fourier transform of the signal. Therefore, the sampling frequency T has to obey $\frac{1}{T} \geqslant 2B$ or $T \leqslant \frac{1}{2B}$ in order to gain the Fourier transform of the bandlimited signal.

Theorem 4.7 (Sampling theorem). *If a (Schwartz) function f contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $\frac{1}{2B}$ seconds apart.*

Alternative proof of Theorem 4.7. By the Fourier inversion formula,

$$f(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i x\xi}\, d\xi\,, \qquad \text{where} \quad \widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x\xi}\, dt\,.$$

By assumption, $\operatorname{supp}(\widehat{f}) \subseteq [-B, B]$; thus $f(x) = \int_{-B}^{B} \widehat{f}(\xi) e^{2\pi i x\xi}\, d\xi$ which implies that

$$f\Big(\frac{k}{2B}\Big) = \int_{-B}^{B} \widehat{f}(\xi) e^{\frac{i\pi k\xi}{B}}\, d\xi\,.$$

Treating $\widehat{f}$ as a function defined on $[-B, B]$, the identity above implies that $\Big\{\frac{1}{2B} f\Big(\frac{-k}{2B}\Big)\Big\}_{k=-\infty}^{\infty}$ is the Fourier coefficients of $\widehat{f}$ and

$$\widehat{f}(\xi) = \sum_{k=-\infty}^{\infty} \frac{1}{2B} f\Big(\frac{-k}{2B}\Big) e^{\frac{i\pi k\xi}{B}} = \sum_{k=-\infty}^{\infty} \frac{1}{2B} f\Big(\frac{k}{2B}\Big) e^{-\frac{i\pi k\xi}{B}} \qquad \forall\, \xi \in [-B, B] \tag{4.7}$$

which, together with the fact that $\widehat{f} = 0$ outside $[-B, B]$, allows us to reconstruct f using the Fourier inversion formula. □

Taking the Fourier inverse transform of $\widehat{f}(\xi)$ obtained by (4.7), we find that

$$\begin{aligned} f(x) &= \sum_{k=-\infty}^{\infty} \frac{1}{2B} f\Big(\frac{k}{2B}\Big) \int_{-B}^{B} e^{2\pi i x\xi - \frac{i\pi k\xi}{B}}\, d\xi = \sum_{k=-\infty}^{\infty} \frac{1}{2B} f\Big(\frac{k}{2B}\Big) \int_{-B}^{B} \cos\Big(\frac{2\pi Bx - \pi k}{B}\xi\Big)\, d\xi \\ &= \sum_{k=-\infty}^{\infty} f\Big(\frac{k}{2B}\Big) \frac{\sin \pi(2Bx-k)}{\pi(2Bx-k)}\,. \end{aligned}$$

Using the normalized sinc function defined by (3.12), we recover the so-called ***Whittaker–Shannon interpolation formula***:

$$f(x) = \sum_{k=-\infty}^{\infty} f\Big(\frac{k}{2B}\Big)\mathrm{sinc}(2Bx-k) \qquad \forall\, f \in \mathscr{S}(\mathbb{R}) \text{ with } \mathrm{supp}(\widehat{f}) \subseteq [-B,B]\,. \tag{4.8}$$

In the following, we examine the Whittaker - Shannon interpolation formula (4.8) for the case that $f \notin \mathscr{S}(\mathbb{R})$. In fact, since

$$\int_{\mathbb{R}} \mathrm{sinc}(2Bx-k)\phi(x)\, dx = \int_{\mathbb{R}} \big(d_{\frac{1}{2B}}\mathrm{sinc}\big)\Big(\frac{k}{2B} - x\Big)\phi(x)\, dx = \big[\big(d_{\frac{1}{2B}}\mathrm{sinc}\big) * \phi\big]\Big(\frac{k}{2B}\Big)\,,$$

instead of (4.8) we show that

$$\langle f, \phi\rangle = \sum_{k=-\infty}^{\infty} f\Big(\frac{k}{2B}\Big)\big[\big(d_{\frac{1}{2B}}\mathrm{sinc}\big) * \phi\big]\Big(\frac{k}{2B}\Big) = \sum_{k=-\infty}^{\infty} \big\langle f\tau_{\frac{k}{2B}}\delta, \big(d_{\frac{1}{2B}}\mathrm{sinc}\big) * \phi\big\rangle\,. \tag{4.9}$$

Suppose that $1 \leqslant p \leqslant \infty$ and $g : \mathbb{R} \to \mathbb{R}$ is an L^p-function $\Big($that is, $\displaystyle\int_{\mathbb{R}} |g(x)|^p\, dx < \infty$ if $1 \leqslant p < \infty$ or g is bounded if $p = \infty\Big)$ supported in an open interval of length $2B$ (later we will let g be the Fourier transform of a bandlimited signal f; so it is reasonable to assume that g is compactly supported). Define

$$G(x) = \sum_{n=-\infty}^{\infty} g(x+2Bn) = \sum_{n=-\infty}^{\infty} (\tau_{-2Bn}g)(x)\,. \tag{4.10}$$

Let q satisfies $\dfrac{1}{p} + \dfrac{1}{q} = 1$ and $\phi \in \mathscr{S}(\mathbb{R})$. The monotone convergence theorem shows that

$$\begin{aligned} \int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \big|(\tau_{-2Bn}g)(x)\phi(x)\big|\, dx &= \sum_{n=-\infty}^{\infty} \int_{-(2n+1)B}^{-(2n-1)B} \big|(\tau_{-2Bn}g)(x)\big|\langle x\rangle^{-2}\langle x\rangle^{2}\big|\phi(x)\big|\, dx \\ &\leqslant \sum_{n=-\infty}^{\infty} \int_{-(2n+1)B}^{-(2n-1)B} \big|g(x+2Bn)\big|\langle x\rangle^{-2} p_2(\phi)\, dx = p_2(\phi) \sum_{n=-\infty}^{\infty} \int_{-B}^{B} |g(x)|\langle x - 2Bn\rangle^{-2}\, dx\,. \end{aligned}$$

If $1 < p < \infty$, Hölder's inequality implies that

$$\begin{aligned} \int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \big|(\tau_{-2Bn}g)(x)\langle x\rangle^{-2}\big|\, dx &\leqslant \sum_{n=-\infty}^{\infty} \int_{-B}^{B} |g(x)|\langle x - 2Bn\rangle^{-2}\, dx \\ &\leqslant \sum_{n=-\infty}^{\infty} \Big(\int_{-B}^{B} |g(x)|^p\, dx\Big)^{\frac{1}{p}} \Big(\int_{-B}^{B} \frac{dx}{(1+|x-2Bn|^2)^q}\Big)^{\frac{1}{q}} \\ &\leqslant \|g\|_{L^p(\mathbb{R})} \sum_{n=-\infty}^{\infty} \frac{(2B)^{\frac{1}{q}}}{1+(2|n|-1)^2B^2} < \infty \end{aligned}$$

while if $p = 1$,

$$\sum_{n=-\infty}^{\infty} \int_{-B}^{B} |g(x)| \langle x - 2Bn \rangle^{-2} \, dx \leqslant \|g\|_{L^1(\mathbb{R})} \sum_{n=-\infty}^{\infty} \frac{1}{1 + (2|n| - 1)^2 B^2} < \infty \, ;$$

thus if $1 \leqslant p < \infty$,

$$\int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \big| (\tau_{-2Bn} g)(x) \phi(x) \big| \, dx \leqslant C_\ell \|g\|_{L^p(\mathbb{R})} p_\ell(\phi) \qquad \forall \, \ell \gg 1 \tag{4.11}$$

for some constant $C_\ell > 0$. On the other hand, if $p = \infty$,

$$\int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \big| (\tau_{-2Bn} g)(x) \phi(x) \big| \, dx \leqslant \|g\|_\infty \|\phi\|_{L^1(\mathbb{R})} \leqslant C_\ell \|g\|_\infty p_\ell(\phi) \qquad \forall \, \ell \gg 1 \, . \tag{4.12}$$

Therefore, $G \in \mathscr{S}(\mathbb{R})'$ since

$$\big| \langle G, \phi \rangle \big| \leqslant \int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \big| (\tau_{-2Bn} g)(x) \phi(x) \big| \, dx \leqslant C_\ell p_\ell(\phi) \qquad \forall \, \phi \in \mathscr{S}(\mathbb{R}) \text{ and } \ell \gg 1 \, .$$

Moreover, it follows from (4.11) and (4.12) that $\sum\limits_{n=-\infty}^{\infty} \big| (\tau_{-2Bn} g) \phi \big| \in L^1(\mathbb{R})$. By the fact that

$$\Big| \sum_{n=-k}^{k} (\tau_{-2Bn} g)(x) \phi(x) \Big| \leqslant \sum_{n=-\infty}^{\infty} \big| (\tau_{-2Bn} g)(x) \phi(x) \big| \qquad \forall \, x \in \mathbb{R} \, ,$$

the dominated convergence theorem implies that

$$\begin{aligned} \langle G, \phi \rangle &= \int_{\mathbb{R}} \lim_{k \to \infty} \sum_{n=-k}^{k} (\tau_{-2Bn} g)(x) \phi(x) \, dx = \lim_{k \to \infty} \int_{\mathbb{R}} \sum_{n=-k}^{k} (\tau_{-2Bn} g)(x) \phi(x) \, dx \\ &= \sum_{n=-\infty}^{\infty} \langle \tau_{-2Bn} g, \phi \rangle \qquad \forall \, \phi \in \mathscr{S}(\mathbb{R}) \, . \end{aligned}$$

Suppose that $\text{supp}(g) \subseteq (a - B, a + B)$. Then $G = g$ on $(a - B, a + B)$. In addition, if $x \in [a + (k-1)B, a + (k+1)B]$, then $G(x) = g(x - kB)$; thus $G(x + 2B) = G(x)$ for all $x \in \mathbb{R}$. In other words, G can be viewed as the $2B$-periodic extension of non-vanishing part of g.

Let $\phi \in \mathscr{S}(\mathbb{R})$. By the definition of the inverse Fourier transform of tempered distributions,

$$\langle \check{G}, \phi \rangle = \langle G, \check{\phi} \rangle = \sum_{n=-\infty}^{\infty} \langle \tau_{-2Bn} g, \check{\phi} \rangle = \sum_{n=-\infty}^{\infty} \langle g, \tau_{2Bn} \check{\phi} \rangle \, .$$

By the Poisson summation formula,

$$\begin{aligned} \sum_{n=-\infty}^{\infty} (\tau_{2Bn} \check{\phi})(x) &= \sum_{n=-\infty}^{\infty} \check{\phi}(x - 2Bn) = \sum_{n=-\infty}^{\infty} \check{\phi}(x + 2Bn) = \sum_{n=-\infty}^{\infty} (d_{\frac{1}{2B}} \check{\phi})\big(\frac{x}{2B} + n\big) \\ &= \sum_{k=-\infty}^{\infty} \widehat{d_{\frac{1}{2B}} \check{\phi}}(k) e^{\frac{\pi i k x}{B}} = \sum_{k=-\infty}^{\infty} \frac{1}{2B} \phi\big(\frac{k}{2B}\big) e^{\frac{\pi i k x}{B}} \end{aligned}$$

and the convergence is uniform. Therefore,

$$\begin{aligned}\langle \check{G}, \phi\rangle &= \sum_{n=-\infty}^{\infty} \langle g, \tau_{2Bn}\check{\phi}\rangle = \sum_{k=-\infty}^{\infty} \frac{1}{2B}\phi\big(\frac{k}{2B}\big)\int_{\text{supp}(g)} g(x)e^{\frac{\pi i k x}{B}}\,dx = \sum_{k=-\infty}^{\infty} \frac{1}{2B}\phi\big(\frac{k}{2B}\big)\int_{\mathbb{R}} g(x)e^{\frac{\pi i k x}{B}}\,dx \\ &= \sum_{k=-\infty}^{\infty} \frac{1}{2B}\phi\big(\frac{k}{2B}\big)\check{g}\big(\frac{k}{2B}\big) = \frac{1}{2B}\sum_{k=-\infty}^{\infty} \check{g}\big(\frac{k}{2B}\big)\langle \tau_{\frac{k}{2B}}\delta, \phi\rangle .\end{aligned}$$

Similarly, $\langle \widehat{G}, \phi\rangle = \dfrac{1}{2B}\sum\limits_{k=-\infty}^{\infty} \widehat{g}\big(\frac{k}{2B}\big)\langle \tau_{\frac{k}{2B}}\delta, \phi\rangle$ or one can use the formula that $\widehat{G} = \tilde{\check{G}}$ to deduce that

$$\langle \widehat{G}, \phi\rangle = \langle \check{G}, \widetilde{\phi}\rangle = \sum_{k=-\infty}^{\infty} \frac{1}{2B}\widetilde{\phi}\big(\frac{k}{2B}\big)\check{g}\big(\frac{k}{2B}\big) = \sum_{k=-\infty}^{\infty} \frac{1}{2B}\phi\big(\frac{k}{2B}\big)\widehat{g}\big(\frac{k}{2B}\big) = \frac{1}{2B}\sum_{k=-\infty}^{\infty} \widehat{g}\big(\frac{k}{2B}\big)\langle \tau_{\frac{k}{2B}}\delta, \phi\rangle .$$

Symbolically, we can write $\widehat{G} = \dfrac{1}{2B}\sum\limits_{k=-\infty}^{\infty} \widehat{g}\big(\frac{k}{2B}\big)\tau_{\frac{k}{2B}}\delta$ and $\check{G} = \dfrac{1}{2B}\sum\limits_{k=-\infty}^{\infty} \check{g}\big(\frac{k}{2B}\big)\tau_{\frac{k}{2B}}\delta$ in $\mathscr{S}(\mathbb{R})'$.

Remark 4.8. Let Ш denote the tempered distribution

$$\langle Ш, \phi\rangle = \sum_{n=-\infty}^{\infty} \phi(n) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

We note that the sum above makes sense if $\phi \in \mathscr{S}(\mathbb{R})$, and

$$\sum_{n=-\infty}^{\infty} \phi(n) = \sum_{n=-\infty}^{\infty} \langle n\rangle^{-k}\langle n\rangle^{k}\phi(n) \leqslant \Big(\sum_{n=-\infty}^{\infty} \langle n\rangle^{-k}\Big)p_k(\phi) = C_k p_k(\phi) \qquad \forall\, k \geqslant 2\,.$$

Therefore, Ш is indeed a tempered distribution. Since $\phi(n) = \langle \tau_n\delta, \phi\rangle$, symbolically we also write $Ш = \sum\limits_{n=-\infty}^{\infty} \tau_n\delta$.

By the definition of the Fourier transform of tempered distributions,

$$\langle \widehat{Ш}, \phi\rangle = \langle Ш, \widehat{\phi}\rangle = \sum_{n=-\infty}^{\infty} \widehat{\phi}(n) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,,$$

and the Poisson summation formula implies that

$$\langle \widehat{Ш}, \phi\rangle = \sum_{k=-\infty}^{\infty} \phi(k) = \langle Ш, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

Therefore, Theorem 3.49 implies that $\widehat{Ш} = \check{Ш} = Ш$ in $\mathscr{S}(\mathbb{R})'$. Define $Ш_p = \dfrac{1}{p}d_p Ш$, where d_p is a dilation operator. Then

$$\langle Ш_p, \phi\rangle = \langle Ш, d_{p^{-1}}\phi\rangle = \sum_{n=-\infty}^{\infty} (d_{p^{-1}}\phi)(n) = \sum_{n=-\infty}^{\infty} \phi(pn) = \sum_{n=-\infty}^{\infty} \langle \tau_{pn}\delta, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,. \tag{4.13}$$

Symbolically, $Ш_p = \sum\limits_{n=-\infty}^{\infty} \tau_{pn}\delta$. Moreover, $\widehat{Ш_p} = \widecheck{Ш_p} = d_{p^{-1}}Ш = \dfrac{1}{p}Ш_{\frac{1}{p}}$ which is the same as saying that

$$\langle \widecheck{Ш_p}, \phi\rangle = \langle d_{p^{-1}}Ш, \varphi\rangle = p^{-1}\langle Ш, d_p\varphi\rangle = \frac{1}{p}\sum_{n=-\infty}^{\infty} \langle \tau_{\frac{n}{p}}\delta, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

Symbolically, $\widecheck{\text{Ш}}_p = \dfrac{1}{p} \sum\limits_{n=-\infty}^{\infty} \tau_{\frac{n}{p}} \delta$.

Formally speaking, G given by (4.10) can be expressed as $G = \text{Ш}_{2B} * g$. Using this representation, $\check{G} = \widecheck{\text{Ш}_{2B}} \check{g} = \dfrac{1}{2B} \sum\limits_{n=-\infty}^{\infty} \check{g} \tau_{\frac{n}{2B}} \delta$. Therefore, by the fact that $g = (d_{2B}\Pi)G$ in $\mathscr{S}(\mathbb{R})'$, we find that

$$
\begin{aligned}
\check{g}(x) &= (\widecheck{d_{2B}\Pi} * \check{G})(x) = \big[(2Bd_{\frac{1}{2B}}\check{\Pi}) * \check{G}\big](x) = 2B \int_{\mathbb{R}} \check{G}(y)\check{\Pi}(2B(x-y))\, dy \\
&= \sum_{n=-\infty}^{\infty} \int_{\mathbb{R}} \check{g}(y)\tau_{\frac{n}{2B}}(y)\text{sinc}(2B(x-y))\, dy = \sum_{n=-\infty}^{\infty} \check{g}\big(\frac{n}{2B}\big)\text{sinc}(2Bx-n)\,.
\end{aligned}
$$

The Whittaker–Shannon interpolation formula (4.8) then follows from letting $g = \widehat{f}$ in the identity above.

Example 4.9. Let $f : \mathbb{R} \to \mathbb{R}$ be a function supported in $[0, T]$ (thus one can view f as a signal recorded in the time interval $[0, T]$). Define $F(x) = \sum\limits_{n=-\infty}^{\infty} f(x + nT)$. Then

$$
\widehat{F}(\xi) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \widehat{f}\big(\frac{k}{T}\big)(\tau_{\frac{k}{T}}\delta)(\xi)\,.
$$

On the other hand, the Fourier series of $\sum\limits_{n=-\infty}^{\infty} f(x+nT)$, the T-periodic extension of $f\mathbf{1}_{[0,T]}$, is

$$
s(f, x) = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{\frac{2\pi ikx}{T}}\,, \quad \text{where } \widehat{f}_k = \frac{1}{T}\int_0^T f(x)e^{-\frac{2\pi ikx}{T}}\, dx = \frac{1}{T}\widehat{f}\big(\frac{k}{T}\big)\,.
$$

Therefore, $\widehat{F} = \sum\limits_{k=-\infty}^{\infty} \widehat{f}_k \tau_{\frac{k}{T}} \delta$, and accordingly, $F(x) = \sum\limits_{k=-\infty}^{\infty} \widehat{f}_k e^{\frac{2\pi ikx}{T}}$ in $\mathscr{S}(\mathbb{R})'$.

Now suppose that the signal is sampled with sampling rate F_s (times per second). Then in total there are $N = TF_s$ samples of the signal. Write these samples as $\{x_0, x_1, \cdots, x_{N-1}\}$. Then $x_\ell = f\big(\dfrac{\ell}{F_s}\big)$. We remark that the set $\{x_0, x_1, \cdots, x_{N-1}\}$ resembles a digitalized version of the signal and is usually called a digital signal. The DFT of the digital signal is given by

$$
X_k = \sum_{\ell=0}^{N-1} x_\ell e^{\frac{-2\pi ik\ell}{N}} = \sum_{\ell=0}^{N-1} f\big(\frac{\ell}{F_s}\big) e^{\frac{-2\pi ik}{T}\cdot\frac{\ell}{F_s}} \qquad \forall\, k \in \mathbb{Z}
$$

and the inverse DFT of $\{X_k\}_{k\in\mathbb{Z}}$ is given by

$$
x_\ell = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi ik\ell}{N}} = \frac{1}{N}\sum_{\ell=0}^{N-1} X_k e^{\frac{2\pi ik}{T}\cdot\frac{\ell}{F_s}} \qquad \forall\, \ell \in \mathbb{Z}\,.
$$

4.1.1 The inner-product point of view

Let $e_k(x) = \text{sinc}(x-k) = (\tau_k\text{sinc})(x)$. Then $e_k \in L^2(\mathbb{R})$ since

$$
\int_{\mathbb{R}} \text{sinc}^2(x-k)\, dx = \int_{\mathbb{R}} \text{sinc}^2 x dx = \int_{\mathbb{R}} \frac{\sin^2 \pi x}{\pi^2 x^2} dx = \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2 x}{x^2} dx < \infty\,.
$$

By the Plancherel formula (3.10),

$$(e_k, e_\ell)_{L^2(\mathbb{R})}s = \big(\widetilde{\tau_k \text{sinc}}, \widetilde{\tau_\ell \text{sinc}}\big)_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \Pi(\xi) e^{2\pi i k\xi} \overline{\Pi(\xi) e^{2\pi i \ell\xi}}\, d\xi = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i (k-\ell)\xi}\, d\xi$$

which is 0 if $k \neq \ell$ and is 1 is $k = \ell$. Therefore, we find that $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal set in $L^2(\mathbb{R})$.

Now suppose that $f \in L^2(\mathbb{R})$ such that $\text{supp}(\widehat{f}) \subseteq \big[-\frac{1}{2}, \frac{1}{2}\big]$. Then

$$\begin{aligned}(f, e_k)_{L^2(\mathbb{R})} &= \big(\widehat{f}, \widehat{\tau_k \sin c}\big)_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \widehat{f}(\xi) \overline{\Pi(-\xi) e^{-2\pi i k\xi}}\, d\xi = \int_{-\frac{1}{2}}^{\frac{1}{2}} \widehat{f}(\xi) e^{2\pi i k\xi}\, d\xi \\ &= \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i k\xi}\, d\xi = \check{\widehat{f}}(k) = f(k)\end{aligned}$$

if f is continuous at k. In other words, if $f \in L^2(\mathbb{R}) \cap \mathscr{C}(\mathbb{R})$ such that $\text{supp}(\widehat{f}) \subseteq \big[-\frac{1}{2}, \frac{1}{2}\big]$, then

$$\sum_{k=-\infty}^{\infty} f(k)\text{sinc}(x-k) = \sum_{k=-\infty}^{\infty} (f, e_k)_{L^2(\mathbb{R})} e_k(x)$$

which, by the Whittaker–Shannon interpolation formula (4.8), further shows that

$$f = \sum_{k=-\infty}^{\infty} (f, e_k)_{L^2(\mathbb{R})} e_k \qquad \text{in} \quad \mathscr{S}(\mathbb{R})'.$$

In other words, one can treat $\{e_k\}_{k\in\mathbb{Z}}$ as an "orthonormal basis" in the space

$$\Big\{ f \in L^2(\mathbb{R}) \big(\cap \mathscr{C}(\mathbb{R})\big) \,\Big|\, \text{supp}(\widehat{f}) \subseteq \big[-\tfrac{1}{2}, \tfrac{1}{2}\big] \Big\}.$$

4.1.2 Sampling periodic functions

A bandlimited signal cannot be timelimited; thus when applying the sampling theorem, it always requires infinitely many sampling to construct the signal perfectly. On the other hand, it is possible that to construct a bandlimited signal perfectly using finitely many sampling provided that the bandlimited signal is periodic. In the following, we discuss why this is true.

Suppose that f is a q-periodic bandlimited signals such that $\text{supp}(f) \subseteq [-B, B]$. Then the Whittaker–Shannon interpolation formula implies that

$$f(x) = \sum_{k=-\infty}^{\infty} f\big(\frac{k}{p}\big)\text{sinc}(px-k)$$

as long as $p > 2B$. If $pq \in \mathbb{N}$, then by the fact that $f\big(\frac{k}{p}\big) = f\big(\frac{k+mpq}{p}\big)$ for all $m \in \mathbb{Z}$, the sum above can be re-grouped as

$$\begin{aligned}f(x) &= \sum_{\ell=0}^{pq-1} \sum_{m=-\infty}^{\infty} f\big(\frac{\ell+mpq}{p}\big)\text{sinc}(px-\ell-mpq) \\ &= \sum_{\ell=0}^{pq-1} f\big(\frac{\ell}{p}\big) \sum_{m=-\infty}^{\infty} \text{sinc}(px-\ell-mpq)\,;\end{aligned} \tag{4.14}$$

thus if we can find the sum $\sum\limits_{m=-\infty}^{\infty} \text{sinc}(px-\ell-mpq)$, f can be rewritten as finite sum.

Lemma 4.10. *If $p, q > 0$ and pq is an odd number, then*

$$\sum_{m=-\infty}^{\infty} \operatorname{sinc}(px - mpq) = \frac{\operatorname{sinc}(px)}{\operatorname{sinc}\frac{x}{q}} \qquad \forall\, x \in \mathbb{R}\,.$$

Proof. Note that

$$\sum_{m=-\infty}^{\infty} \operatorname{sinc}(px - mpq) = \sum_{m=-\infty}^{\infty} (d_{\frac{1}{p}}\operatorname{sinc})(x - mq) = (\text{Ш}_q * d_{\frac{1}{p}}\operatorname{sinc})(x)\,,$$

where $\text{Ш}_q = \dfrac{1}{q} d_q \text{Ш}$ is defined in Remark 4.8. Taking the Fourier transform, we find that

$$\mathscr{F}\big[(\text{Ш}_q * d_{\frac{1}{p}}\operatorname{sinc})\big](\xi) = \big(\widehat{\text{Ш}_q}\, \widehat{d_{\frac{1}{p}}\operatorname{sinc}}\big)(\xi) = \frac{1}{pq}\big(\text{Ш}_{\frac{1}{q}} d_p \Pi\big)(\xi) = \frac{1}{pq}(d_p\Pi)(\xi) \sum_{k=-\infty}^{\infty} \big(\tau_{\frac{k}{q}}\delta\big)(\xi)\,.$$

If pq is odd, then $\dfrac{p}{2} \neq \dfrac{k}{q}$ for all $k \in \mathbb{N}$; thus by the fact that $\operatorname{supp}(d_p\Pi) \subseteq \big[-\frac{p}{2}, \frac{p}{2}\big]$, we have

$$(d_p\Pi)(\xi) \sum_{k=-\infty}^{\infty} \big(\tau_{\frac{k}{q}}\delta\big)(\xi) = \sum_{-\frac{p}{2}<\frac{k}{q}<\frac{p}{2}} \big(\tau_{\frac{k}{q}}\delta\big)(\xi) = \sum_{k=-\frac{pq-1}{2}}^{\frac{pq-1}{2}} \big(\tau_{\frac{k}{q}}\delta\big)(\xi)\,.$$

Therefore,

$$\mathscr{F}\big[(\text{Ш}_q * d_{\frac{1}{p}}\operatorname{sinc})\big](\xi) = \frac{1}{pq} \sum_{k=-\frac{pq-1}{2}}^{\frac{pq-1}{2}} \big(\tau_{\frac{k}{q}}\delta\big)(\xi)\,.$$

Taking the inverse Fourier transform,

$$\sum_{m=-\infty}^{\infty} \operatorname{sinc}(px - mpq) = \mathscr{F}^*\mathscr{F}\big[(\text{Ш}_q * d_{\frac{1}{p}}\operatorname{sinc})\big](x) = \frac{1}{pq} \sum_{k=-\frac{pq-1}{2}}^{\frac{pq-1}{2}} e^{\frac{2\pi i k x}{q}} = \frac{1}{pq}\frac{\sin \pi p x}{\sin \frac{\pi x}{q}} = \frac{\operatorname{sinc}(px)}{\operatorname{sinc}\frac{x}{q}}\,. \qquad \square$$

By Lemma 4.10, (4.14) implies that

$$\sum_{m=-\infty}^{\infty} \operatorname{sinc}(px - \ell - mpq) = \frac{\operatorname{sinc}(p(x - \frac{\ell}{p}))}{\operatorname{sinc}\frac{x-\frac{\ell}{p}}{q}} = \frac{\operatorname{sinc}(px - \ell)}{\operatorname{sinc}\frac{px-\ell}{pq}}\,;$$

thus we obtain that

$$f(x) = \sum_{\ell=0}^{pq-1} f\Big(\frac{\ell}{p}\Big) \frac{\operatorname{sinc}(px - \ell)}{\operatorname{sinc}\frac{px-\ell}{pq}}\,. \tag{4.15}$$

Example 4.11. Let $f(x) = \cos(2\pi x)$. Then f is 1-periodic and $\operatorname{supp}(\widehat{f}) \subseteq [-1.2, 1.2]$. Letting $p = 3$ in (4.15), we find that

$$\cos(2\pi x) = \frac{\operatorname{sinc}(3x)}{\operatorname{sinc} x} + \cos\frac{2\pi}{3}\frac{\operatorname{sinc}(3x-1)}{\operatorname{sinc}\frac{3x-1}{3}} + \cos\frac{4\pi}{3}\frac{\operatorname{sinc}(3x-2)}{\operatorname{sinc}\frac{3x-2}{3}}\,.$$

4.2 Necessary Conditions for Sampling of Entire Functions

The sampling theorem provides a way of reconstructing signals based on sampled signals with sampling frequency larger than twice of the bandwidth of bandlimited signals. It is natural to ask whether we can reduce the sampling frequency for perfect reconstruction of bandlimited signals or not. Moreover, it is also possible that the support of the Fourier transform of a signal (usually called spectrum of the signal) is contained in a "small" portion of the interval $[-B, B]$, and in this case we hope to reduce the sampling frequency for the reconstruction of the signal.
Question: Is there a lower bound of the sampling frequency for perfect reconstruction of bandlimited signals?

Generally speaking, the way of sampling does not have to be uniform as long as the samples from a signal are enough to reconstruct the signal. A good choice of sampled set should obey

1. the signal is uniquely determined by the set of sampled signals - the ***uniqueness*** of the reconstruction of signals;

2. each set of sampled signals should come from a possible bandlimited signal - the ***existence*** of the reconstruction of signals.

These two requirements for sets on which the signals are sampled, together with the idea that the sampled set is not necessary uniform, induce the following

Definition 4.12. Let $S \subseteq \mathbb{R}^n$ be a measurable set in $\mathbb{R}^n$, and $\mathcal{B}(S)$ denote the subspace of $L^2(\mathbb{R}^n)$ consisting of those functions whose Fourier transform (given by (3.11)) is supported on S; that is,

$$\mathcal{B}(S) \equiv \left\{ f \in L^2(\mathbb{R}^n) \,\middle|\, \operatorname{supp}(\widehat{f}) \subseteq S \right\}.$$

A subset Λ of $\mathbb{R}^n$ is said to be ***uniformly discrete*** if the distance between any two distinct points of Λ exceeds some positive quantity; that is, there exists $\lambda_0 > 0$ such that $\|x - y\|_{\mathbb{R}^n} \geqslant \lambda_0$ for all $x, y \in \Lambda$ and $x \neq y$. Such a λ_0 is called a ***separation number***. A uniformly discrete set Λ is said to be

1. a ***set of sampling*** for $\mathcal{B}(S)$ if there exists a constant K such that

$$\|f\|^2_{L^2(\mathbb{R}^n)} \leqslant K \sum_{\lambda \in \Lambda} \big|f(\lambda)\big|^2 \qquad \forall\, f \in \mathcal{B}(S)\,;$$

2. a ***set of interpolation*** for $\mathcal{B}(S)$ if for each square-summable collection of complex numbers $\{a_\lambda\}_{\lambda \in \Lambda}$ there exists $f \in \mathcal{B}(S)$ with $f(\lambda) = a_\lambda$ for all $\lambda \in \Lambda$.

Example 4.13. Let I be an interval of length 1. Then $\mathbb{Z}$ is a set of sampling and interpolation for $\mathcal{B}(I)$:

1. If $f \in \mathcal{B}(I)$, then $\widehat{f}$ has the following Fourier series representation

$$\widehat{f}(x) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi i k x} \qquad \text{for almost every } x \in I\,,$$

where $c_k = \int_I \widehat{f}(x)e^{-2\pi ikx}\,dx$. By the fact that $\text{supp}(\widehat{f}) \subseteq I$, the Fourier inversion formula implies that

$$c_k = \int_{\mathbb{R}} \widehat{f}(x)e^{-2\pi ikx}\,dx = f(-k)\,;$$

thus $\{f(-k)\}_{k=-\infty}^{\infty}$ is the Fourier coefficients of $\widehat{f}$. The Plancherel identity and the Parseval identity then imply that

$$\|f\|_{L^2(\mathbb{R})}^2 = \|\widehat{f}\|_{L^2(\mathbb{R})}^2 = \|\widehat{f}\|_{L^2(I)}^2 = \sum_{k=-\infty}^{\infty} |c_k|^2 = \sum_{k=-\infty}^{\infty} \|f(k)|^2\,;$$

thus $\mathbb{Z}$ is a set of sampling for $\mathcal{B}(I)$.

2. Let $\{c_k\}_{k=-\infty}^{\infty}$ be a square-summable sequence. Define

$$g(x) = \mathbf{1}_I(x) \sum_{k=-\infty}^{\infty} c_k e^{-2\pi ikx}\,.$$

Then the Parseval identity implies that

$$\|g\|_{L^2(\mathbb{R})}^2 = \|g\|_{L^2(I)}^2 = \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty\,;$$

thus $g \in L^2(\mathbb{R})$. Let $f = \check{g}$. Then $f \in L^2(\mathbb{R})$ and the Fourier inversion formula implies that for all $k \in \mathbb{Z}$,

$$f(k) = \int_{\mathbb{R}} \widehat{f}(\xi)e^{2\pi ik\xi}\,d\xi = \int_{\mathbb{R}} g(\xi)e^{2\pi ik\xi}\,d\xi = \int_I g(\xi)e^{2\pi ik\xi}\,d\xi = \widehat{g}_{-k} = c_k\,;$$

thus $\mathbb{Z}$ is a set of interpolation for $\mathcal{B}(I)$.

Remark 4.14. Suppose that f is an L^2-signal satisfying $\text{supp}(\widehat{f}) \subseteq (B-1, B)$ for some $B \gg 1$. Then certainly $\text{supp}(\widehat{f}) \subseteq (-B, B)$ and the sample theorem implies that to perfectly reconstruct the signal one can consider sampling f every $\frac{1}{2B}$ seconds. On the other hand, Example 4.13 shows that one can reconstruct the signal by sampling the signal once per second. This is a huge amount of reduction of sampling if $B \gg 1$. Therefore, the sampling rate provided by the sampling theorem is only a sufficient condition for perfect reconstruction of bandlimited signals, but possibly can be reduced for specific cases.

For $n = 1$, Landau in his paper "Necessary density conditions for sampling and interpolation of certain entire functions" shows the following

Theorem 4.15. *Let S be the union of a finite number of intervals of total measure $|S|$.*

1. *If Λ is a set of sampling for $\mathcal{B}(S)$, then there exist generic constants A, B such that*

$$n^-(r) \equiv \inf_{y\in\mathbb{R}} \#\big(\Lambda \cap [y, y+r]\big) \geqslant |S|r - A\log^+ r - B \qquad \forall\, r > 0\,. \tag{4.16}$$

2. *If Λ is a set of interpolation for $\mathcal{B}(S)$, then there exist generic constants A, B such that*

$$n^+(r) \equiv \sup_{y\in\mathbb{R}} \#\big(\Lambda \cap [y, y+r]\big) \leqslant |S|r + A\log^+ r + B \qquad \forall\, r > 0\,.$$

In the following, we only focus on the proof of the first case in Theorem 4.15.

Before proceeding, we need to introduce some terminologies. Let $Q, S \subseteq \mathbb{R}^n$, and $\mathcal{D}(Q)$ be the subspace of $L^2(\mathbb{R}^n)$ consisting of functions supported on Q. Let D_Q and B_S denote the orthogonal projection of $L^2(\mathbb{R}^n)$ onto $\mathcal{D}(S)$ and $\mathcal{B}(S)$, respectively. Then

$$B_S = \mathscr{F}^* \chi_S \mathscr{F} \qquad \text{and} \qquad D_Q = \chi_Q\,, \tag{4.17}$$

where χ_A denotes the operator defined by multiplying by the characteristic function of A.

Proposition 4.16. *Let $k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be square integrable, $K(x,y) = \overline{K(y,x)}$ for all $x, y \in \mathbb{R}^n$, and $K : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ be an operator defined by*

$$(Kf)(x) = \int_{\mathbb{R}^n} k(x,y) f(y)\, dy\,.$$

Then

1. *$k(x,y) = \sum\limits_{k=1}^{\infty} \mu_k \varphi_k(x)\overline{\varphi_k(y)}$, where $\{\varphi_k\}_{k=1}^{\infty}$ denotes the orthonormal sequence of eigenfunctions, and $\{\mu_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ denotes the sequence of corresponding eigenvalues of K;*

2. *$\sum\limits_{k=1}^{\infty} \mu_k = \int_{\mathbb{R}^n} k(x,x)\, dx$;* 3. *$\sum\limits_{k=1}^{\infty} \mu_k^2 = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \big|k(x,y)\big|^2\, dxdy$.*

Theorem 4.17. *Let $Q, S \subseteq \mathbb{R}^n$ be bounded measurable sets, and D_Q, B_S be the projection operators of $L^2(\mathbb{R}^n)$ defined in (4.17). Denoting the eigenvalues of $B_S D_Q B_S$, arranged in non-increasing order, by $\lambda_k(S,Q)$, where $k \in \mathbb{N} \cup \{0\}$. Then*

(i) *$\lambda_k(S,Q) = \lambda_k(Q,S)$.*

(ii) *$\lambda_k(S,Q) = \lambda_k(S+\sigma, Q+\tau) = \lambda_k(\alpha S, \alpha^{-1}Q)$ for all $\sigma, \tau \in \mathbb{R}^n$ and $\alpha > 0$.*

(iii) *$\sum\limits_{k=0}^{\infty} \lambda_k(S,Q) = |S||Q|$.*

(iv) *$\sum\limits_{k=0}^{\infty} \lambda_k^2(S,Q) \geqslant \sum\limits_{k=0}^{\infty} \lambda_k^2(S,Q) + \sum\limits_{k=0}^{\infty} \lambda_k^2(S,Q)$ if $Q = Q_1 \cup Q_2$ and $Q_1 \cap Q_2 = \varnothing$.*

(v) *$\sum\limits_{k=0}^{\infty} \lambda_k^2(S,Q) \geqslant \big(sq - \frac{2}{\pi^2}\log^+(sq) - \frac{6}{\pi^2}\big)^n$, where S and Q are cubes with edges parallel to the coordinate axes with $|S| = s^n$, $|Q| = q^n$, and $\log^+ x = \max\{0, \log x\}$.*

(vi) *For any k-dimensional subspace C_k of $L^2(\mathbb{R}^n)$,*

$$\lambda_k(S,Q) \leqslant \sup_{\substack{f\in\mathcal{B}(S)\\ f\perp C_k, f\neq 0}} \frac{\|D_Q f\|^2_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} \qquad \textit{and} \qquad \lambda_{k-1}(S,Q) \geqslant \inf_{\substack{f\in\mathcal{B}(S)\cap C_k\\ f\neq 0}} \frac{\|D_Q f\|^2_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}}\,.$$

Proof. For two (completely continuous) operators A and B, we write $A \sim B$ if A and B has the same nonzero eigenvalues, including multiplicities. Suppose that $\lambda \neq 0$ is an eigenvalue of $B_S D_Q B_S$. Then $B_S D_Q B_S \varphi = \lambda\varphi$ for some $\varphi \neq 0$. By the fact that B_S is a projection, we have

$$\lambda B_S \varphi = B_S B_S D_Q B_S \varphi = B_S D_Q B_S \varphi = \lambda\varphi$$

which implies that $B_S\varphi = \varphi$. Moreover, $D_Q B_S \varphi \neq 0$. Applying D_Q to the equation above, we find that

$$D_Q B_S D_Q D_Q B_S \varphi = D_Q B_S D_Q B_S \varphi = \lambda D_Q B_S \varphi$$

which, by the fact that $D_Q B_S \varphi \neq 0$, implies that λ is also a eigenvalue of $D_Q B_S D_Q$. As a consequence,

$$B_S D_Q B_S \sim D_Q B_S D_Q \,. \tag{4.18}$$

Therefore, to study the nonzero eigenvalues of the operator $B_S D_Q B_S$, it suffices to study the operator $D_Q B_S D_Q$.

Let C denoted the complex conjugate operator; that is, $Cf = \bar{f}$. Then $C\mathscr{F}C = \mathscr{F}^{-1}$ and $C\mathscr{F}^{-1}C = \mathscr{F}$. By the fact that $\mathscr{F}$ is unitary and $D_Q B_S D_Q$ is symmetric (so the eigenvalues are real),

$$\begin{aligned} D_Q B_S D_Q &\sim C D_Q B_S D_Q C = \chi_Q C\mathscr{F}^{-1} C \chi_S C \mathscr{F} C \chi_Q = \chi_Q \mathscr{F} \chi_S \mathscr{F}^{-1}\chi_Q \\ &\sim \mathscr{F}^{-1}\chi_Q \mathscr{F}\chi_S \mathscr{F}^{-1}\chi_Q \mathscr{F} = B_Q D_S B_Q \,. \end{aligned}$$

This proves (ii). Since S and Q are bounded, the Fubini theorem implies that

$$\begin{aligned} (D_Q B_S D_Q f)(x) &= \chi_Q(x)\Big[\int_{\mathbb{R}^n} \chi_S(\xi)\Big(\int_{\mathbb{R}^n} (\chi_Q f)(y)^{-2\pi i y\cdot\xi}\, dy\Big) e^{2\pi i x\cdot\xi}\, d\xi\Big] \\ &= \Big[\int_{\mathbb{R}^n} \chi_Q(x)\chi_Q(y) f(y)\Big(\int_{\mathbb{R}^n} \chi_S(\xi) e^{-2\pi i (y-x)\cdot\xi}\, d\xi\Big) dy\Big] \\ &= \int_{\mathbb{R}^n} \chi_Q(x)\chi_Q(y)\widehat{\chi_S}(y-x) f(y)\, dy \,. \end{aligned}$$

Using (4.18), the change of variables formula together with (i) shows (ii).

Let $k(x,y) = \chi_Q(x)\chi_Q(y)\widehat{\chi_S}(y-x)$ and K be the operator defined by $(Kf)(x) = \displaystyle\int_{\mathbb{R}^n} k(x,y) f(y)\, dy$. Then $k(x,y) = \overline{k(y,x)}$; thus Proposition 4.16 implies that

$$\sum_{k=0}^{\infty} \lambda_k(S,Q) = \int_{\mathbb{R}^n} k(x,x)\, dx = \int_Q \widehat{\chi_S}(0)\, d\xi = |S||Q|$$

which establishes (iii).

To prove (iv), we make use of Proposition 4.16 and find that

$$\sum_{k=0}^{\infty} \lambda_k(S,Q)^2 = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \big|k(x,y)\big|^2\, dx\Big) dy = \int_{Q\times Q} \big|\widehat{\chi_S}(y-x)\big|^2\, d(x,y) \,.$$

Since $Q\times Q \subseteq (Q_1\times Q_1) \cup (Q_2\times Q_2)$ and $(Q_1\times Q_1)\cap(Q_2\times Q_2) = \varnothing$, by the identity above we conclude that

$$\begin{aligned} \sum_{k=0}^{\infty} \lambda_k(S,Q)^2 &\geqslant \int_{Q_1\times Q_1} \big|\widehat{\chi_S}(y-x)\big|^2\, d(x,y) + \int_{Q_2\times Q_2} \big|\widehat{\chi_S}(y-x)\big|^2\, d(x,y) \\ &= \lambda_k(S,Q_1) + \lambda_k(S,Q_2) \,. \end{aligned}$$

Let S and Q be cubes with volume s^n and q^n. Using (ii) we can assume that S and Q are centered at the origin; that is, $S = \left[-\frac{s}{2}, \frac{s}{2}\right]^n$ and $Q = \left[-\frac{q}{2}, \frac{q}{2}\right]^n$. Then

$$\widehat{\chi_S}(y-x) = \int_{[-\frac{s}{2},\frac{s}{2}]^n} e^{2\pi i(x-y)\cdot\xi}\, d\xi = \prod_{i=1}^{n} \frac{\sin\pi(x_i-y_i)s}{\pi(x_i-y_i)}\,;$$

thus Proposition 4.16 provides that

$$\begin{aligned}\sum_{k=0}^{\infty}\lambda_k(S,Q)^2 &= \int_{[-\frac{q}{2},\frac{q}{2}]^n}\Big(\int_{[-\frac{q}{2},\frac{q}{2}]^n}\prod_{i=1}^{n}\frac{\sin^2(\pi|x_i-y_i|s)}{\pi^2|x_i-y_i|^2}\,dx\Big)dy\\ &= \Big(\int_{-\frac{q}{2}}^{\frac{q}{2}}\int_{-\frac{q}{2}}^{\frac{q}{2}}\frac{\sin^2(\pi|x-y|s)}{\pi^2|x-y|^2}\,dxdy\Big)^n.\end{aligned}$$

By the fact that $\displaystyle\int_{-\infty}^{\infty}\frac{\sin^2 t}{t^2}\,dt = \pi$,

$$\begin{aligned}\int_{-\frac{q}{2}}^{\frac{q}{2}}\int_{-\frac{q}{2}}^{\frac{q}{2}}\frac{\sin^2(\pi|x-y|s)}{\pi^2|x-y|^2}\,dxdy &= \frac{s}{\pi}\int_{-\frac{q}{2}}^{\frac{q}{2}}\Big(\int_{\pi(-\frac{q}{2}-y)s}^{\pi(\frac{q}{2}-y)s}\frac{\sin^2 t}{t^2}\,dt\Big)dy\\ &= \frac{s}{\pi}\int_{-\frac{q}{2}}^{\frac{q}{2}}\Big(\int_{-\infty}^{\infty}\frac{\sin^2 t}{t^2}\,dt - \int_{\pi(\frac{q}{2}-y)s}^{\infty}\frac{\sin^2 t}{t^2}\,dt - \int_{-\infty}^{\pi(-\frac{q}{2}-y)s}\frac{\sin^2 t}{t^2}\,dt\Big)dy\\ &= sq - \frac{2s}{\pi}\int_{-\frac{q}{2}}^{\frac{q}{2}}\Big(\int_{\pi(\frac{q}{2}-y)s}^{\infty}\frac{\sin^2 t}{t^2}\,dt\Big)dy\\ &= sq - \frac{q}{\pi}\int_{-1}^{1}\Big(\int_{\frac{q\pi}{2}(1-y)}^{\infty}\frac{\sin^2(st)}{t^2}\,dt\Big)dy\,.\end{aligned}$$

Note that

$$\begin{aligned}\frac{q}{\pi}\int_{-1}^{1}\Big(\int_{\frac{q\pi}{2}(1-y)}^{\infty}\frac{\sin^2(st)}{t^2}\,dt\Big)dy &= \frac{q}{\pi}\int_{q\pi}^{\infty}\Big(\int_{-1}^{1}\frac{\sin^2(st)}{t^2}\,dy\Big)dt + \frac{q}{\pi}\int_{0}^{q\pi}\Big(\int_{1-\frac{2t}{q\pi}}^{1}\frac{\sin^2(st)}{t^2}\,dy\Big)dt\\ &= \frac{2}{\pi}\int_{\pi}^{\infty}\frac{\sin^2(sqt)}{t^2}\,dt + \frac{2}{\pi^2}\int_{0}^{q\pi}\frac{\sin^2(st)}{t}\,dt\\ &\leqslant \frac{2}{\pi}\int_{\pi}^{\infty}\frac{1}{t^2}\,dt + \frac{2}{\pi^2}\int_{0}^{sq}\frac{\sin^2(\pi t)}{t}\,dt\\ &\leqslant \frac{2}{\pi^2}\Big[1+\int_{0}^{1}\frac{\sin^2(\pi t)}{t}\,dt + \int_{1}^{sq}\frac{\sin^2(\pi t)}{t}\,dt\Big] \leqslant \frac{2}{\pi^2}\Big[3+\log^+(sq)\Big]\,,\end{aligned}$$

so (v) is established.

For a given k-dimensional subspace C_k, the subspace $B_S C_k$ has dimension $d \leqslant k$. Moreover, $f \perp B_S C_k$ if and only if $B_S f \perp C_k$. By the fact that $\|B_S f\|_{L^2(\mathbb{R}^n)} \leqslant \|f\|_{L^2(\mathbb{R}^n)}$ and

$$\lambda_k(S,Q) \leqslant \sup_{f\perp C_k, f\neq 0}\frac{(B_S D_Q B_S f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}}$$

for any k-dimensional subspace C_k of $L^2(\mathbb{R}^n)$, we conclude that

$$\begin{aligned}\lambda_k(S,Q) \leqslant \lambda_d(S,Q) &\leqslant \sup_{\substack{f\perp B_S C_k\\ f\neq 0, B_S f\neq 0}}\frac{(B_S D_Q B_S f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} \leqslant \sup_{\substack{B_S f\perp C_k\\ B_S f\neq 0}}\frac{(D_Q B_S f, B_S f)_{L^2(\mathbb{R}^n)}}{\|B_S f\|^2_{L^2(\mathbb{R}^n)}}\\ &\leqslant \sup_{\substack{f\in\mathcal{B}(S)\\ f\perp C_k, f\neq 0}}\frac{(D_Q f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} = \sup_{\substack{f\in\mathcal{B}(S)\\ f\perp C_k, f\neq 0}}\frac{(D_Q f, D_Q f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} = \sup_{\substack{f\in\mathcal{B}(S)\\ f\perp C_k, f\neq 0}}\frac{\|D_Q f\|^2_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}}\,.\end{aligned}$$

On the other hand, by the fact that

$$\lambda_{k-1}(S,Q) \geqslant \inf_{f\in C_k, f\neq 0} \frac{(B_S D_Q B_S f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}}$$

for any k-dimensional subspace of $L^2(\mathbb{R}^n)$, choosing $C_k \subseteq \mathcal{B}(S)$ we obtain that

$$\begin{aligned}\lambda_{k-1}(S,Q) &\geqslant \inf_{f\in C_k, f\neq 0} \frac{(B_S D_Q B_S f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} = \inf_{\substack{f\in\mathcal{B}(S)\cap C_k\\ f\neq 0}} \frac{(D_Q B_S f, B_S f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} \\ &= \inf_{\substack{f\in\mathcal{B}(S)\cap C_k\\ f\neq 0}} \frac{(D_Q f, f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} = \inf_{\substack{f\in\mathcal{B}(S)\cap C_k\\ f\neq 0}} \frac{(D_Q f, D_Q f)_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}} = \inf_{\substack{f\in\mathcal{B}(S)\cap C_k\\ f\neq 0}} \frac{\|D_Q f\|^2_{L^2(\mathbb{R}^n)}}{\|f\|^2_{L^2(\mathbb{R}^n)}}\,;\end{aligned}$$

thus (vi) is established. □

Lemma 4.18. *For any bounded measurable set $S \subseteq \mathbb{R}^n$ and $d > 0$, there exists a Schwartz function $h : \mathbb{R}^n \to \mathbb{C}$ such that $\operatorname{supp}(h) \subseteq B(0,d)$ and $|\widehat{h}(\xi)| \geqslant 1$ for all $\xi \in S$.*

Proof. Since S is bounded, $S \subseteq B(0,R)$ for some $R > 0$. Let $f \in \mathscr{S}(\mathbb{R}^n)$ be such that $f > 2$ on $B(0,R)$. Since $\check{f} \in \mathscr{S}(\mathbb{R}^n)$, there exists $g \in \mathscr{C}^\infty_c(\mathbb{R}^n)$ such that $\|\check{f} - g\|_{L^1(\mathbb{R}^n)} < 1$. Choose $r > d$ such that $\operatorname{supp}(g) \subseteq B(0,r)$, and defined the function h by

$$h(x) \equiv \frac{r^n}{d^n} g\Big(\frac{rx}{d}\Big)\,.$$

Then h is supported in $B(0,d)$. Moreover,

$$\widehat{h}(\xi) = \int_{\mathbb{R}^n} \frac{r^n}{d^n} g\Big(\frac{rx}{d}\Big) e^{2\pi i x\cdot\xi}\, dx = \widehat{g}\Big(\frac{d\xi}{r}\Big) \qquad \forall\, \xi \in \mathbb{R}^n\,,$$

and the Fourier inversion formula implies that

$$\sup_{\xi\in\mathbb{R}^n} \Big|f\Big(\frac{d\xi}{r}\Big) - \widehat{h}(\xi)\Big| = \sup_{\xi\in\mathbb{R}^n} \Big|f\Big(\frac{d\xi}{r}\Big) - \widehat{g}\Big(\frac{d\xi}{r}\Big)\Big| = \|f - \widehat{g}\|_{L^\infty(\mathbb{R}^n)} \leqslant \|\check{f} - g\|_{L^1(\mathbb{R}^n)} < 1\,.$$

Therefore, if $|\xi| < R$, we must have $\dfrac{d|\xi|}{r} < R$; hence

$$\big|\widehat{h}(\xi)\big| \geqslant \Big|f\Big(\frac{d\xi}{r}\Big)\Big| - 1 \geqslant 1 \qquad \forall\, |\xi| \leqslant R\,.$$

Since $S \subseteq B(0,R)$, $\big|\widehat{h}\big| \geqslant 1$ on S. □

Lemma 4.19. *Let $S \subseteq \mathbb{R}$ be a bounded set and Λ be a uniformly discrete set of sampling for $\mathcal{B}(S)$ with separation number d and counting function n. For a compact set I, I^+ denotes the set of points whose distance to I is less than $\frac{d}{2}$. Then*

$$\lambda_{n(I^+)}(S,I) \leqslant \gamma < 1 \tag{4.19}$$

for some γ depending on S, Λ but not on I.

Proof. By Lemma 4.18, there exists a Schwartz function h such that h vanishes outside $B(0,\frac{d}{2})$ and $|\widehat{h}| \geqslant 1$ on S. Let C be the subspace of $L^2(\mathbb{R})$ spanned by the functions $\overline{h(\lambda-\cdot)}$ for $\lambda \in \Lambda \cap I^+$. Since

$$\big(\overline{h(\lambda_i-\cdot)},\overline{h(\lambda_j-\cdot)}\big)_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \overline{h(\lambda_i-x)}h(\lambda_j-x)\,dx = 0 \qquad \text{if } \lambda_i \neq \lambda_j\,,$$

the dimension of C is $n(I^+)$.

For a given $f \in \mathcal{B}(S)$ be given, we define $g = f * h$; that is,

$$g(x) = \int_{\mathbb{R}} f(y)h(x-y)\,dy = \int_{|y-x|<\frac{d}{2}} f(y)h(x-y)\,dy\,.$$

Then $\widehat{g} = \widehat{f}\,\widehat{h}$ which further implies that $g \in \mathcal{B}(S)$. Therefore, by the fact that Λ is a set of sampling for $\mathcal{B}(S)$,

$$\|g\|^2_{L^2(\mathbb{R})} \leqslant K \sum_{\lambda\in\Lambda} \big|g(\lambda)\big|^2\,.$$

Moreover, the Plancherel identity shows that

$$\|g\|_{L^2(\mathbb{R})} = \|\widehat{g}\|_{L^2(\mathbb{R})} = \|\widehat{f}\|_{L^2(\mathbb{R})}\|\widehat{h}\|_{L^2(\mathbb{R})} \geqslant \|\widehat{f}\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})} \tag{4.20}$$

and the Cauchy-Schwarz inequality shows that

$$|g(x)|^2 \leqslant \|h\|^2_{L^2(\mathbb{R})} \int_{|y-x|<\frac{d}{2}} f(y)^2\,dy\,.$$

Therefore, if $f \in \mathcal{B}(S)$ and $f \perp C$, we have

$$\begin{aligned}\|f\|^2_{L^2(\mathbb{R})} \leqslant \|g\|^2_{L^2(\mathbb{R})} \leqslant K\sum_{\lambda\in\Lambda}\big|g(\lambda)\big|^2 &= K\sum_{\lambda\in\Lambda,\lambda\notin I^+}\big|g(\lambda)\big|^2\\ &\leqslant K\|h\|^2_{L^2(\mathbb{R})}\sum_{\lambda\in\Lambda,\lambda\notin I^+}\int_{|y-\lambda|<\frac{d}{2}}\big|f(y)\big|^2\,dy\\ &\leqslant K\|h\|^2_{L^2(\mathbb{R})}\int_{I^\complement}\big|f(y)\big|^2\,dy = K\|h\|^2_{L^2(\mathbb{R})}\Big[\|f\|^2_{L^2(\mathbb{R})} - \int_{\mathbb{R}}\big|D_If(y)\big|^2\,dy\Big]\,.\end{aligned}$$

As a consequence, letting $\gamma \equiv 1 - \dfrac{1}{K\|h\|^2_{L^2(\mathbb{R})}}$, we have

$$\frac{\|D_If\|^2_{L^2(\mathbb{R})}}{\|f\|^2_{L^2(\mathbb{R})}} \leqslant 1 - \frac{1}{K\|h\|^2_{L^2(\mathbb{R})}} = \gamma < 1\,.$$

Inequality (4.19) then follows from (vi) of Theorem 4.17. □

Lemma 4.20. *Let $S \subseteq \mathbb{R}$ be a bounded set and Λ be a uniformly discrete set of interpolation for $\mathcal{B}(S)$ with separation number d and counting function n. For a compact set I, I^- denotes the set of points whose distance to $I^\complement$ exceeds $\frac{d}{2}$. Then*

$$\lambda_{n(I^-)-1}(S,I) \geqslant \delta > 0$$

for some δ depending on S and Λ but not on I.

Proof. Again by Lemma 4.18, there exists a Schwartz function h such that h vanishes outside $B(0, \frac{d}{2})$ and $|\widehat{h}| \geqslant 1$ on S.

Define a bounded linear operator A on $\mathcal{B}(S)$ by $Ag = \big\{g(\lambda)\big\}_{\lambda\in\Lambda}$ if $g \in \mathcal{B}(S)$. To see the boundedness of A, let $g \in \mathcal{B}(S)$ be given, and let $f \in \mathcal{B}(S)$ be such that $\widehat{g} = \widehat{f}\,\widehat{h}$; that is,

$$f(x) = \int_{\mathbb{R}} \frac{\widehat{g}(\xi)}{\widehat{h}(\xi)} e^{2\pi i x\xi}\, d\xi\,.$$

The same as (4.20), we have $\|f\|_{L^2(\mathbb{R})} \leqslant \|g\|_{L^2(\mathbb{R})}$, and the Cauchy-Schwarz inequality implies that

$$\big|g(x)\big|^2 \leqslant \|h\|^2_{L^2(\mathbb{R})} \int_{|y-x|<\frac{d}{2}} |f(y)|^2\, dy\,.$$

Since Λ is uniformly discrete with separation number d, by the fact that $g = f * h$, we have

$$\sum_{\lambda\in\Lambda} \big|g(\lambda)\big|^2 \leqslant \|h\|^2_{L^2(\mathbb{R})} \sum_{\lambda\in\Lambda} \int_{|y-\lambda|<\frac{d}{2}} \big|f(y)\big|^2\, dy \leqslant \|h\|^2_{L^2(\mathbb{R})}\|f\|^2_{L^2(\mathbb{R})} \leqslant \|h\|^2_{L^2(\mathbb{R})}\|g\|^2_{L^2(\mathbb{R})}\,. \tag{4.21}$$

Therefore, $A : \mathcal{B}(S) \to \ell^2$ is bounded.

Define $\mathcal{E}(S) \equiv \big\{f \in \mathcal{B}(S)\,\big|\, f(\lambda) = 0 \text{ for all } \lambda \in \Lambda\big\}$. For $f \in \mathcal{B}(S)$, the Cauchy-Schwarz inequality and the Plancherel identity imply that

$$|f(x)|^2 \leqslant \Big(\int_S \big|\widehat{f}(y)\big|\, dy\Big)^2 \leqslant |S|\|\widehat{f}\|^2_{L^2(\mathbb{R})} = |S|\|f\|^2_{L^2(\mathbb{R})}\,,$$

so if $\{f_k\}_{k=1}^\infty \subseteq \mathcal{B}(S)$ converges to f in $L^2(\mathbb{R})$ (that means $\|f_k - f\|_{L^2(\mathbb{R})} \to 0$ as $k \to \infty$), $\{f_k\}_{k=1}^\infty$ also converges to f uniformly on S. In particular, if $\{f_k\}_{k=1}^\infty \subseteq \mathcal{E}(S)$ converges to f in L^2 sense, then for $\lambda \in \Lambda$,

$$\big|f(\lambda)\big| = \lim_{k\to\infty} \big|f(\lambda) - f_k(\lambda)\big| \leqslant \lim_{k\to\infty} \sqrt{|S|}\|f_k - f\|_{L^2(\mathbb{R})} = 0$$

which implies that $f \in \mathcal{E}(S)$. In other words, $\mathcal{E}(S)$ is a closed subspace.

Let $\mathcal{E}^\perp(S)$ denote the orthogonal complement of $\mathcal{E}(S)$, and $\{a_\lambda\}_{\lambda\in\Lambda} \in \ell^2$ be given. Since Λ is a set of interpolation for $\mathcal{B}(S)$, there exists $f \in \mathcal{B}(S)$ such that

$$f(\lambda) = a_\lambda \qquad \forall\, \lambda \in \Lambda\,.$$

By the fact that $\mathcal{B}(S) = \mathcal{E}(S) \oplus \mathcal{E}^\perp(S)$, there exist (unique) $f_1 \in \mathcal{E}(S)$ and $f_2 \in \mathcal{E}^\perp(S)$ such that $f = f_1 + f_2$. Therefore, since $f_1(\lambda) = 0$ for all $\lambda \in \Lambda$, we have

$$f_2(\lambda) = f_1(\lambda) + f_2(\lambda) = f(\lambda) = a_\lambda \qquad \forall\, \lambda \in \Lambda\,.$$

Therefore, Λ is a set of interpolation for $\mathcal{E}^\perp(S)$. This also implies that $A : \mathcal{E}^\perp(S) \to \ell^2$ is surjective.

Moreover, noting that $A : \mathcal{E}^\perp(S) \to \ell^2$ is one-to-one, we find that $A : \mathcal{E}^\perp(S) \to \ell^2$ is a bounded linear bijective operator. Therefore, the bounded inverse theorem (from functional analysis) implies that $A^{-1} : \ell^2 \to \mathcal{E}^\perp(S)$ is also bounded linear; thus there exists $K > 0$ such that

$$\|g\|^2_{L^2(\mathbb{R})} \leqslant K \sum_{\lambda\in\Lambda} \big|g(\lambda)\big|^2 \qquad \forall\, g \in \mathcal{E}^\perp(S)\,. \tag{4.22}$$

In other words, Λ is a set of sampling for $\mathcal{E}^\perp(S)$ as well.

For each $\lambda \in \Lambda$, let $\varphi_\lambda \in \mathcal{E}^\perp(S)$ be the function whose value is 1 at λ and 0 at other point of Λ. We remark that such a φ_λ exists since Λ is a set of interpolation for $\mathcal{E}^\perp(S)$. Clearly $\{\varphi_\lambda\}_{\lambda\in\Lambda}$ is a set of linear independent functions. Let $\psi_\lambda \in \mathcal{B}(S)$ be such that $\widehat{\varphi_\lambda} = \widehat{\psi_\lambda}\widehat{h}$; that is,

$$\psi_\lambda(x) = \int_S \frac{\widehat{\varphi_\lambda}(\xi)}{\widehat{h}(\xi)} e^{2\pi i x\cdot\xi}\, d\xi\,.$$

Then $\{\psi_\lambda\}_{\lambda\in\Lambda}$ is also a set of linear independent functions. Let C be the subspace of $\mathcal{B}(S)$ spanned by $\{\psi_\lambda\}_{\lambda\in\Lambda\cap I^-}$. Then $\dim(C) = n(I^-) = \#(\Lambda\cap I^-)$. For a given function $f \in C$, $f = \sum\limits_{\lambda\in\Lambda\cap I^-} c_\lambda \psi_\lambda$ for some $\{c_\lambda\}_{\lambda\in\Lambda\cap I^-}$; thus

$$\widehat{f * h} = \widehat{f}\,\widehat{h} = \sum_{\lambda\in\Lambda\cap I^-} c_\lambda \widehat{\psi_\lambda}\widehat{h} = \sum_{\lambda\in\Lambda\cap I^-} c_\lambda \widehat{\varphi_\lambda}$$

which shows that $f*h$ is a linear combination of $\{\varphi_\lambda\}_{\lambda\in\Lambda\cap I^-}$. This further implies that

$$f*h \in \mathcal{E}^\perp(S) \quad \text{and} \quad (f*h)(\lambda) = 0 \quad \forall\, \lambda \notin \Lambda\cap I^- \qquad \text{whenever} \quad f \in C\,.$$

As a consequence, using (4.20) and (4.22), we obtain that if $f \in C$,

$$\begin{aligned} K^{-1}\|f\|^2_{L^2(\mathbb{R})} &\leqslant K^{-1}\|(f*h)\|^2_{L^2(\mathbb{R})} \leqslant \sum_{\lambda\in\Lambda} \big|(f*h)(\lambda)\big|^2 = \sum_{\lambda\in\Lambda\cap I^-} \big|(f*h)(\lambda)\big|^2 \\ &\leqslant \|h\|^2_{L^2(\mathbb{R})} \sum_{\lambda\in\Lambda\cap I^-} \int_{|y-\lambda|<\frac{d}{2}} \big|f(y)\big|^2\, dy \leqslant \|h\|^2_{L^2(\mathbb{R})}\|f\|^2_{L^2(I)} = \|h\|^2_{L^2(\mathbb{R})}\|D_I f\|^2_{L^2(\mathbb{R})}\,; \end{aligned}$$

thus for $f \in C$,

$$\frac{\|D_I f\|^2_{L^2(\mathbb{R})}}{\|f\|^2_{L^2(\mathbb{R})}} \geqslant \frac{1}{K\|h\|^2_{L^2(\mathbb{R})}} = \delta > 0\,,$$

where we note that δ depends only on S (due to the dependence on h) and Λ but not on I. The lemma is then concluded by (vii) of Theorem 4.17. $\square$

Proof of (4.16). Let d be a separation number of Λ, $I = [-\frac{1}{2}, \frac{1}{2}]$ be a unit interval, and J be an interval of length r such that $n^-(r) = n(J) = \#(\Lambda\cap J)$. Since J is a single interval, then J^+, the set of points whose distance to J is less than $\frac{d}{2}$, satisfies $n(J^+) \leqslant n(J)+2$; thus (ii) of Theorem 4.17 and Lemma 4.19 imply that

$$\lambda_{n(J)+2}(S, rI) \leqslant \lambda_{n(J^+)}(S, J) \leqslant \gamma < 1 \tag{4.23}$$

for some γ independent of r.

Suppose that S consists of p disjoint intervals $J_1, \cdots, J_p$. By Example 4.13, the set of integers $\mathbb{Z}$ is a uniformly discrete set of sampling and interpolation of $\mathcal{B}(I)$ with separation number 1. The set $(rS)^-$, the collection of points whose distance to $(rS)^\complement$ exceeds $\frac{1}{2}$, consists of at most p disjoint intervals, so

$$\#((rS)^- \cap \mathbb{Z}) \geqslant |(rS)^-| - p = r|S| - 2p\,.$$

By (i) and (ii) of Theorem 4.17 and Lemma 4.20, we find that

$$\lambda_{r|S|-2p-1}(S, rI) = \lambda_{r|S|-2p-1}(I, rS) \geqslant \lambda_{\#((rS)^-\cap\mathbb{Z})-1}(I, rS) \geqslant \delta > 0 \tag{4.24}$$

for some δ independent of r.

Let $\mu(S, rI) = \sum_{k=0}^{\infty} \lambda_k(S, rI)\big(1 - \lambda_k(S, rI)\big)$. By (iii)-(v) of Theorem 4.17,

$$\begin{aligned}\mu(S, rI) &= r|S| - \sum_{k=0}^{\infty} \lambda_k^2(S, rI) \leqslant r|S| - \sum_{j=1}^{p} \sum_{k=0}^{\infty} \lambda_k^2(J_j, rI) \\ &\leqslant r|S| - \sum_{j=1}^{p} \Big(r|J_j| - \frac{2}{\pi^2} \log^+(r|J_j|) - \frac{6}{\pi^2}\Big) \\ &\leqslant \frac{2}{\pi^2} \sum_{j=1}^{p} \log^+ r|J_j| + \frac{6p}{\pi^2} \leqslant A \log^+ r + B\end{aligned}$$

for some constants A, B depending only on S.

Now suppose that $n(J) + 2 \leqslant r|S| - 2p - 1$, then (4.23) and (4.24) imply that

$$0 < \delta \leqslant \lambda_k(S, rI) \leqslant \gamma < 1 \qquad \forall\, k \in [n(J) + 2, r|S| - 2p - 1]\,.$$

Therefore,

$$\big(r|S| - 2p - 1 - n(J) - 2 + 1\big) \min\big\{\delta(1-\delta), \gamma(1-\gamma)\big\} \leqslant \mu(S, rI) \leqslant A \log^+ r + B$$

which shows that

$$n(J) \geqslant r|S| - A \log^+ r - B \tag{4.25}$$

for some constants A, B depending on S and Λ but not r. On the other hand, if $n(J)+2 > r|S|-2p-1$, (4.25) holds automatically (for proper choices of A and B); thus (4.16) is established. □

We can measure the density of a uniformly discrete set Λ in terms of function $n^{\pm}(r)$.

Definition 4.21. The ***Beurling upper and lower uniform densities*** of a uniformly discrete set Λ, denoted by $D^+(\Lambda)$ and $D^-(\Lambda)$, respectively, are the numbers defined by

$$D^{\pm}(\Lambda) = \lim_{r \to \infty} \frac{n^{\pm}(r)}{r}\,.$$

The Beurling density reduces to the usual concept of average sampling rate for uniform and periodic non-uniform sampling.

Corollary 4.22. *Let $S \subseteq \mathbb{R}$ be a bounded set with measure $|S|$ and Λ be a uniformly discrete set.*

1. *If Λ is a set of sampling for $\mathcal{B}(S)$, then $D^-(\Lambda) \geqslant |S|$.*

2. *If Λ is a set of interpolation for $\mathcal{B}(S)$, then $D^+(\Lambda) \leqslant |S|$.*

5 Applications on Partial Differential Equations

5.1 Heat Conduction in a Rod

Consider the heat distribution on a rod of length L: Parameterize the rod by $[0, L]$, and let t be the time variable. Let $\rho(x)$, $s(x)$, $\kappa(x)$ denote the density, specific heat, and the thermal conductivity of the rod at position $x \in (0, L)$, respectively, and $u(x, t)$ denote the temperature at position x and time t. For $0 < x < L$, and $\Delta x, \Delta t \ll 1$,

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[u(y, t+\Delta t) - u(y, t)\big]\, dy = \int_t^{t+\Delta t} \big[-\kappa(x)u_x(x, t') + \kappa(x+\Delta x)u_x(x+\Delta x, t')\big]\, dt' \,,$$

where the left-hand side denotes the change of the total heat in the small section $(x, x+\Delta x)$, and the right-hand side denotes the heat flows from outside. Divide both sides by $\Delta x \Delta t$ and letting Δx and Δt approach zero, if all the functions appearing in the equation above are smooth enough, we find that

$$\rho(x)s(x)u_t(x, t) = \big[\kappa(x)u_x(x, t)\big]_x \qquad 0 < x < L \,, \quad t > 0 \,. \tag{5.1}$$

Assuming uniform rod; that is, ρ, s, κ are constant, then (5.1) reduces to that

$$u_t(x, t) = \alpha^2 u_{xx}(x, t) \,, \qquad 0 < x < L \,, \quad t > 0 \,, \tag{5.2a}$$

where $\alpha^2 = \dfrac{\kappa}{\rho s}$ is called the ***thermal diffusivity***. By re-scaling the length scale, we can assume that $\alpha = 1$.

To determine the state of the temperature, we need to impose that initial condition

$$u(x, 0) = u_0(x) \qquad 0 < x < L \tag{5.2b}$$

for some given function $u_0 : [0, L] \to \mathbb{R}$ and a boundary condition. Usually one of the following four types of boundary conditions is imposed:

1. ***Dirichlet boundary condition***: The Dirichlet boundary condition is used to describe the phenomena that the temperature at the end points of the rod is known/controable. Mathematically, it is expressed by

   $$u(0, t) = a(t) \quad \text{and} \quad u(L, t) = b(t) \qquad \forall\, t > 0$$

   for some given functions $a(t)$ and $b(t)$.

2. ***Neumann boundary condition***: The Neumann boundary condition is used to describe the phenomena of insulation; that is, there is no heat flow at the end points. Mathematically, it is expressed by

   $$u_x(0, t) = u_x(L, t) = 0 \qquad \forall\, t > 0 \,.$$

   In general, we can consider the boundary condition

   $$u_x(0, t) = a(t) \quad \text{and} \quad u_x(L, t) = b(t) \qquad \forall\, t > 0$$

   for some given functions $a(t)$ and $b(t)$.

3. ***Mixed type boundary condition***: We can also consider the case that at one end point the temperature is known while there is no heat flow on the other end point. In general, this is expressed by

   $$u(0, t) = a(t) \quad \text{and} \quad u_x(L, t) = b(t) \qquad \forall\, t > 0$$

or

$$u_x(0,t) = a(t) \quad \text{and} \quad u(L,t) = b(t) \qquad \forall\, t > 0$$

for some given functions $a(t)$ and $b(t)$.

4. ***Periodic boundary condition***: Suppose that instead of rods we consider modelling the temperature distribution in a (big) ring (with perimeter L). Choosing a point on the ring as the "left-end" point and parameterizing the point of the ring by arc-length, we then have the "boundary" condition

$$u(0,t) = u(L,t) \qquad \forall\, t > 0\,.$$

This is called the periodic boundary condition.

5.1.1 The Dirichlet problem

In this sub-section we consider the heat equation with Dirichlet boundary condition:

$$u_t - u_{xx} = 0 \qquad \text{in} \quad (0,L) \times (0,\infty)\,, \tag{5.3a}$$

$$u = u_0 \qquad \text{on} \quad (0,L) \times \{t = 0\}\,, \tag{5.3b}$$

$$u(0,t) = a\,, \quad u(L,t) = b \qquad \text{for all } t > 0\,, \tag{5.3c}$$

where a and b are given constants. Let $v(x,t) = u(x,t) - \dfrac{b-a}{L}x - a$. Then v satisfies

$$v_t - v_{xx} = 0 \qquad \text{in} \quad (0,L) \times (0,\infty)\,, \tag{5.4a}$$

$$v = v_0 \qquad \text{on} \quad (0,L) \times \{t = 0\}\,, \tag{5.4b}$$

$$v(0,t) = v(L,t) = 0 \qquad \text{for all } t > 0\,, \tag{5.4c}$$

where $v_0 : [0,L] \to \mathbb{R}$ is given by $v_0(x) = u_0(x) - \dfrac{b-a}{L}x - a$. As long as the solution v to (5.4) is found, the solution u to (5.3) can be constructed using $u(x,t) = v(x,t) + \dfrac{b-a}{L}x + a$. Therefore, we focus on solving (5.4) (using the Fourier series method).

The idea of using the Fourier series to solve (5.4) is that for each fixed $t > 0$ we express v in terms of its Fourier series representation (using proper "basis"). Recall that for a function $f : [0,L] \to \mathbb{R}$, we have the Fourier representation

$$f(x) \text{ ``$=$'' } \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos\frac{2\pi kx}{L} + s_k \sin\frac{2\pi kx}{L} \qquad x \in [0,L]\,,$$

where $c_k = \dfrac{2}{L}\displaystyle\int_0^L f(x)\cos\frac{2\pi kx}{L}\,dx$ and $s_k = \dfrac{2}{L}\displaystyle\int_0^L f(x)\sin\frac{2\pi kx}{L}\,dx$, so

$$v(x,t) \text{ ``$=$'' } \frac{c_0(t)}{2} + \sum_{k=1}^{\infty} c_k(t) \cos\frac{2\pi kx}{L} + s_k(t) \sin\frac{2\pi kx}{L} \qquad x \in [0,L]\,,$$

for some sequence of functions $\{c_k(t)\}_{k=0}^{\infty}$ and $\{s_k(t)\}_{k=1}^{\infty}$. However, this particular Fourier series representation of v is not a good choice of solving (5.4) since it is difficult to validate the boundary condition (5.4c).

Note that for $f : [0, L] \to \mathbb{R}$ instead of the Fourier series representation above we can also consider the "cosine" series or "sine" series that are obtained by treating f as the restriction of an even or an odd function defined on $[-L, L]$ to $[0, L]$. In other words, define $f_e, f_o : [-L, L] \to \mathbb{R}$, called the ***even*** and ***odd extension*** of f respectively, by

$$f_e(x) = \begin{cases} f(x) & \text{if } x \in [0, L]\,, \\ f(-x) & \text{if } x \in [-L, 0)\,, \end{cases} \qquad \text{and} \qquad f_o(x) = \begin{cases} f(x) & \text{if } x \in [0, L]\,, \\ -f(-x) & \text{if } x \in [-L, 0)\,, \end{cases}$$

then $f = f_o = f_e$ on $[0, L]$. Since

$$f_e(x) \text{ ``$=$'' } \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos\frac{\pi k x}{L} \qquad \text{and} \qquad f_o(x) \text{ ``$=$'' } \sum_{k=1}^{\infty} s_k \sin\frac{\pi k x}{L} \qquad x \in [-L, L]\,,$$

where $c_k = \dfrac{2}{L}\displaystyle\int_0^L f(x)\cos\frac{\pi k x}{L}\,dx$ and $s_k = \dfrac{2}{L}\displaystyle\int_0^L f(x)\sin\frac{\pi k x}{L}\,dx$, we have

$$f(x) \text{ ``$=$'' } \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos\frac{\pi k x}{L} \qquad \text{and} \qquad f(x) \text{ ``$=$'' } \sum_{k=1}^{\infty} s_k \sin\frac{\pi k x}{L} \qquad x \in [0, L]\,.$$

Using the sine series, for each $t > 0$ $v(x, t)$ can be expressed as

$$v(x, t) \text{ ``$=$'' } \sum_{k=1}^{\infty} d_k(t) \sin\frac{\pi k x}{L} \qquad x \in [0, L]\,.$$

for some sequence of function $\big\{d_k(t)\big\}_{k=1}^{\infty}$ to be determined. We note that using this particular representation of v the boundary condition (5.4c) automatically holds. Therefore, it suffices to find $\big\{d_k(t)\big\}_{k=1}^{\infty}$ such that (5.4a,b) hold.

Assume that the differentiation of the series can be obtained by term-by-term differentiation; that is,

$$\frac{\partial}{\partial t}\sum_{k=1}^{\infty} d_k(t)\sin\frac{\pi k x}{L} = \sum_{k=1}^{\infty}\frac{\partial}{\partial t}\Big(d_k(t)\sin\frac{\pi k x}{L}\Big) = \sum_{k=1}^{\infty} d_k'(t)\sin\frac{\pi k x}{L}$$

and

$$\frac{\partial^2}{\partial x^2}\sum_{k=1}^{\infty} d_k(t)\sin\frac{\pi k x}{L} = \sum_{k=1}^{\infty}\frac{\partial^2}{\partial x^2}\Big(d_k(t)\sin\frac{\pi k x}{L}\Big) = -\sum_{k=1}^{\infty}\frac{k^2\pi^2}{L^2} d_k(t)\sin\frac{\pi k x}{L}\,.$$

Then (5.4a) implies that

$$\sum_{k=1}^{\infty}\Big[d_k'(t) + \frac{k^2\pi^2}{L^2} d_k(t)\Big]\sin\frac{\pi k x}{L} = 0 \qquad x \in [0, L]\,.$$

As a consequence,

$$d_k'(t) + \frac{k^2\pi^2}{L^2} d_k(t) = 0 \qquad \forall\, k \in \mathbb{N}\,. \tag{5.5a}$$

To determine d_k uniquely, an initial condition for d_k has to be imposed. Noting that (5.4b) implies that

$$v_0(x) \text{ ``$=$'' } \sum_{k=1}^{\infty} d_k(0)\sin\frac{\pi k x}{L} \qquad x \in [0, L]\,;$$

thus

$$d_k(0) = \widehat{v}_{0k} \equiv \frac{2}{L}\int_0^L v_0(x)\sin\frac{\pi k x}{L}\,dx\,. \tag{5.5b}$$

Solving the initial value problem (5.5), we find that

$$d_k(t) = \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \qquad \forall\, k \in \mathbb{N}\,;$$

thus the solution to (5.4) can be written as

$$v(x,t) = \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t}\sin\frac{k\pi x}{L}\,.$$

Therefore, the solution to (5.3) can be written as

$$u(x,t) = \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t}\sin\frac{k\pi x}{L} + \frac{b-a}{L}x + a\,. \tag{5.6}$$

• the long time behavior: Suppose that the temperature at the left-end and right-end points are fixed as a and b (as described in the boundary condition (5.3c)). Then we expect that no matter what the temperature distribution is given initially, the temperature distribution approaches a linear distribution; that is, we expect that $u(x,t) \to \dfrac{b-a}{L}x + a$ as $t \to \infty$ for all $x \in [0,L]$. This expectation is in fact true, and we try to prove this here.

Using (5.6), we obtain that

$$\Big|u(x,t) - \frac{b-a}{L}x - a\Big| \leqslant \sum_{k=1}^{\infty} \big|\widehat{v}_{0k}\big| e^{-\frac{k^2\pi^2}{L^2}t}\,.$$

By the fact that

$$\big|\widehat{v}_{0k}\big| \leqslant \frac{2}{L}\int_0^L \big|v_0(x)\big|\,dx = \frac{2}{L}\|v_0\|_{L^1(0,L)}\,,$$

we find that

$$\begin{aligned}\Big|u(x,t) - \frac{b-a}{L}x - a\Big| &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}t} \leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}(t-1)} e^{-\frac{k^2\pi^2}{L^2}} \\ &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} e^{-\frac{\pi^2}{L^2}(t-1)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}\,;\end{aligned}$$

thus with C denoting the constant $\dfrac{2}{L}\|v_0\|_{L^1(0,L)} e^{\frac{\pi^2}{L^2}}\sum\limits_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}$, we have

$$\sup_{x\in[0,L]}\Big|u(x,t) - \frac{b-a}{L}x - a\Big| \leqslant Ce^{-\frac{\pi^2}{L^2}t}\,. \tag{5.7}$$

Since $C < \infty$, we conclude that the function $u(\cdot,t)$ converges to the function $\dfrac{b-a}{L}x + a$ uniformly on $[0,L]$ as $t \to \infty$.

5.1.2 The Neumann problem

In this sub-section we consider the heat equation with Neumann boundary condition:

$$u_t - u_{xx} = 0 \qquad \text{in} \quad (0, L) \times (0, \infty)\,, \tag{5.8a}$$

$$u = u_0 \qquad \text{on} \quad (0, L) \times \{t = 0\}\,, \tag{5.8b}$$

$$u_x(0, t) = a\,, \quad u_x(L, t) = b \qquad \text{for all } t > 0\,, \tag{5.8c}$$

where a and b are given constants. Let $v(x, t) = u(x, t) - \dfrac{b-a}{2L}(x^2 + 2t) - ax$. Then v satisfies

$$v_t - v_{xx} = 0 \qquad \text{in} \quad (0, L) \times (0, \infty)\,, \tag{5.9a}$$

$$v = v_0 \qquad \text{on} \quad (0, L) \times \{t = 0\}\,, \tag{5.9b}$$

$$v_x(0, t) = v_x(L, t) = 0 \qquad \text{for all } t > 0\,, \tag{5.9c}$$

where $v_0 : [0, L] \to \mathbb{R}$ is given by $v_0(x) = u_0(x) - \dfrac{b-a}{2L}x^2 - ax$. As long as the solution v to (5.9) is found, the solution u to (5.8) can be constructed using $u(x, t) = v(x, t) + \dfrac{b-a}{2L}x^2 + ax$. Therefore, we focus on solving (5.9) (using the Fourier series method). We look for $\{d_k(t)\}_{k=0}^{\infty}$ such that

$$v(x, t) = \frac{d_0(t)}{2} + \sum_{k=1}^{\infty} d_k(t) \cos \frac{\pi k x}{L}$$

validates (5.9a,b).

Assume that the differentiation of the series can be obtained by term-by-term differentiation; that is,

$$\frac{\partial}{\partial t} \sum_{k=1}^{\infty} d_k(t) \cos \frac{\pi k x}{L} = \sum_{k=1}^{\infty} \frac{\partial}{\partial t}\Big(d_k(t) \cos \frac{\pi k x}{L}\Big) = \sum_{k=1}^{\infty} d_k'(t) \cos \frac{\pi k x}{L}\,,$$

$$\frac{\partial}{\partial x} \sum_{k=1}^{\infty} d_k(t) \cos \frac{\pi k x}{L} = \sum_{k=1}^{\infty} \frac{\partial}{\partial x}\Big(d_k(t) \cos \frac{\pi k x}{L}\Big) = -\sum_{k=1}^{\infty} \frac{k\pi}{L} d_k(t) \sin \frac{\pi k x}{L}\,,$$

and

$$\frac{\partial^2}{\partial x^2} \sum_{k=1}^{\infty} d_k(t) \cos \frac{\pi k x}{L} = \sum_{k=1}^{\infty} \frac{\partial^2}{\partial x^2}\Big(d_k(t) \cos \frac{\pi k x}{L}\Big) = -\sum_{k=1}^{\infty} \frac{k^2\pi^2}{L^2} d_k(t) \cos \frac{\pi k x}{L}\,.$$

Then (5.9c) holds automatically, and (5.9a) implies that

$$\frac{d_0'(t)}{2} + \sum_{k=1}^{\infty} \Big[d_k'(t) + \frac{k^2\pi^2}{L^2} d_k(t)\Big] \cos \frac{\pi k x}{L} = 0\,.$$

Therefore, d_0 is a constant and d_k satisfies (5.5) as well. Moreover, expressing v_0 in terms of cosine series, (5.9b) implies that

$$d_k(0) = \widehat{v}_{0k} \equiv \frac{2}{L} \int_0^L v_0(x) \cos \frac{\pi k x}{L}\, dx \qquad \forall\, k \in \mathbb{N} \cup \{0\}\,.$$

Solving (5.5) with the initial condition above, we obtain that

$$d_k(t) = \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \qquad \forall\, k \in \mathbb{N}\,;$$

thus the solution to (5.9) can be written as

$$v(x,t) = \frac{1}{L}\int_0^L v_0(x)\,dx + \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \cos\frac{k\pi x}{L}\,.$$

Therefore, the solution to (5.3) can be written as

$$u(x,t) = \frac{1}{L}\int_0^L v_0(x)\,dx + \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \cos\frac{k\pi x}{L} + \frac{b-a}{2L}(x^2+2t) + ax\,.$$

• the long time behavior: Suppose that the rod is insulated at the end-points; that is, the temperature u satisfies $u_x(0,t) = u_x(L,t) = 0$ for all $t > 0$. Then $v_0 = u_0$ and we expect that no matter what the temperature distribution is given initially, the temperature distribution approaches the average temperature; that is, we expect that $u(x,t) \to \frac{1}{L}\int_0^L u_0(x)\,dx$ as $t \to \infty$ for all $x \in [0,L]$. Similar to the derivation of (5.7),

$$\begin{aligned}\Big|u(x,t) - \frac{1}{L}\int_0^L u_0(x)\,dx\Big| &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}t} \leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}(t-1)} e^{-\frac{k^2\pi^2}{L^2}} \\ &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} e^{-\frac{\pi^2}{L^2}(t-1)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}\,;\end{aligned}$$

thus with C denoting the constant $\frac{2}{L}\|v_0\|_{L^1(0,L)} e^{\frac{\pi^2}{L^2}} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}$, we have

$$\sup_{x\in[0,L]} \Big|u(x,t) - \frac{1}{L}\int_0^L u_0(x)\,dx\Big| \leqslant Ce^{-\frac{\pi^2}{L^2}t}\,. \tag{5.10}$$

Since $C < \infty$, we conclude that the function $u(\cdot,t)$ converges to the function $\frac{1}{L}\int_0^L u_0(x)\,dx$ uniformly on $[0,L]$ as $t \to \infty$.

5.2 Heat Conduction on $\mathbb{R}^n$

Consider the heat equation on $\mathbb{R}^n$

$$u_t - \Delta u = 0 \qquad \text{in} \quad \mathbb{R}^n \times (0,\infty)\,, \tag{5.11a}$$

$$u = u_0 \qquad \text{on} \quad \mathbb{R}^n \times \{t=0\}\,, \tag{5.11b}$$

where Δ is the Laplace operator, called Laplacian, defined by

$$\Delta u = \sum_{k=1}^{n} \frac{\partial^2 u}{\partial x_k^2}\ (= \mathrm{div}\nabla u)\,.$$

For a function f of x (and probably also t), let $\mathscr{F}(f) = \widehat{f}$ denote the Fourier transform of f in x; that is,

$$\mathscr{F}(f)(\xi,t) = \widehat{f}(\xi,t) = \frac{1}{\sqrt{2\pi}^{\,n}}\int_{\mathbb{R}^n} f(x,t)e^{-ix\cdot\xi}\,dx\,.$$

Then by assuming that $\widehat{u}(\cdot,t)$ exists for all $t>0$, Lemma 3.11 implies that

$$\begin{aligned}\mathscr{F}(\Delta u)(\xi,t) &= \mathscr{F}\Big(\sum_{k=1}^n \frac{\partial}{\partial x_k}\frac{\partial u}{\partial x_k}\Big)(\xi,t) = \sum_{k=1}^n \mathscr{F}\Big(\frac{\partial}{\partial x_k}\frac{\partial u}{\partial x_k}\Big)(\xi,t)\\ &= \sum_{k=1}^n i\xi_k \mathscr{F}\Big(\frac{\partial u}{\partial x_k}\Big)(\xi,t) = \sum_{k=1}^n (i\xi_k)^2\widehat{u}(\xi,t) = -|\xi|^2\widehat{u}(\xi,t)\,.\end{aligned}$$

Assume further that

$$\begin{aligned}\frac{\partial}{\partial t}\widehat{u}(\xi,t) &= \frac{\partial}{\partial t}\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u(x,t)e^{-ix\cdot\xi}\,dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\frac{\partial}{\partial t}u(x,t)e^{-ix\cdot\xi}\,dx\\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u_t(x,t)e^{-ix\cdot\xi}\,dx = \widehat{u_t}(\xi,t)\,.\end{aligned}$$

Taking the Fourier transform of (5.11a), we find that

$$\frac{\partial}{\partial t}\widehat{u}(\xi,t) + |\xi|^2\widehat{u}(\xi,t) = \mathscr{F}(u_t - \Delta u)(\xi,t) = 0\,. \tag{5.12}$$

Since $\widehat{u}(\xi,0) = \mathscr{F}(u(\cdot,0))(\xi) = \widehat{u_0}(\xi)$, solving the ODE (5.12) with this initial condition we obtain that

$$\widehat{u}(\xi,t) = \widehat{u_0}(\xi)e^{-|\xi|^2 t} = \widehat{u_0}(\xi)\widehat{P_{2t}}(\xi)\,,$$

where $P_t(x) = \dfrac{1}{\sqrt{t}^n}e^{-\frac{|x|^2}{2t}}$. By Theorem 3.23, we conclude that

$$u(x,t) = (u_0 * P_{2t})(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\frac{1}{\sqrt{2t}^n}e^{-\frac{|x-y|^2}{4t}}u_0(y)\,dy = \frac{1}{\sqrt{4\pi t}^n}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}u_0(y)\,dy\,. \tag{5.13}$$

This induces the following

Definition 5.1. The function $\mathcal{H}(x,t) = \dfrac{1}{\sqrt{4\pi t}^n}e^{-\frac{|x|^2}{4t}}$ is called the ***heat kernel***.

Having introduced the heat kernel, the solution to (5.11), given by (5.13) can be expressed by

$$u(x,t) = \big(\mathcal{H}(\cdot,t) * u_0\big)(x)\,.$$

• Non-uniqueness of solutions: The Fourier transform method only picks up solutions whose Fourier transform is defined, and it is possible that there are other solutions to (5.11). Consider the function

$$u(x,t) = \sum_{k=0}^\infty \frac{g^{(k)}(t)}{(2k)!}x^{2k}\,, \tag{5.14}$$

where g is given by

$$g(t) = \begin{cases} \exp(-t^{-2}) & \text{if } t>0\,,\\ 0 & \text{if } t=0\,.\end{cases}$$

Then there exists $\theta>0$ such that

$$\big|g^{(k)}(t)\big| \leqslant \frac{k!}{(\theta t)^k}\exp\big(-\frac{1}{2}t^{-2}\big) \qquad \forall\, t>0\,. \tag{5.15}$$

In fact, using the Cauchy integral formula,

$$g^{(k)}(t) = \frac{k!}{2\pi i}\oint_{|z-t|=\theta t}\frac{g(z)}{(z-t)^{k+1}}\,dz$$

where $\theta \in (0,1)$ is chosen so small such that $|g(z)| \leqslant \exp\left(-\frac{1}{2}t^{-2}\right)$ on $|z-t| = \theta t$. The choice of such a θ is possible since by writing $z = x + iy$,

$$|g(x+iy)| = \left|\exp\left(\frac{y^2 - x^2 + 2ixy}{(x^2+y^2)^2}\right)\right| = \exp\left(\frac{y^2-x^2}{(x^2+y^2)^2}\right) \leqslant \exp\left(\frac{\theta^2 - (1-\theta)^2}{((1+\theta)^2+\theta^2)^2}t^{-2}\right) \quad \text{if } \theta < \frac{1}{2}\,.$$

By the fact that $\dfrac{k!}{(2k)!} \leqslant \dfrac{1}{k!}$, we find that

$$\sum_{k=0}^{\infty} \left|\frac{g^{(k)}(t)}{(2k)!}x^{2k}\right| \leqslant \sum_{k=0}^{\infty} \frac{x^{2k}}{k!(\theta t)^k} \exp\left(-\frac{1}{2}t^{-2}\right) = \exp\left[\frac{1}{t}\left(\frac{x^2}{\theta} - \frac{1}{2}t^{-2}\right)\right] \qquad \forall\, t > 0, x \in \mathbb{R}\,. \tag{5.16}$$

Therefore, by comparison the series in (5.14) converges for all $t > 0$, and trivially also converges for $t = 0$. Moreover, (5.16) shows that for each $t \in \mathbb{R}$, (5.14) is a convergent power series; thus

$$u_{xx}(x,t) = \frac{\partial^2}{\partial x^2}u(x,t) = \sum_{k=0}^{\infty} \frac{g^{(k)}(t)}{(2k)!}\frac{\partial^2 x^{2k}}{\partial x^2} = \sum_{k=1}^{\infty} \frac{g^{(k)}(t)}{(2k-2)!}x^{2k-2} \qquad \forall\, t > 0, x \in \mathbb{R}\,.$$

On the other hand, by the fact that $\dfrac{(k+1)!}{(2k)!} \leqslant \dfrac{1}{(k-1)!}$ if $k \geqslant 1$, using (5.15) we find that for all $t > 0$ and $x \in \mathbb{R}$,

$$\begin{aligned}\sum_{k=0}^{\infty} \left|\frac{g^{(k+1)}(t)}{(2k)!}x^{2k}\right| &\leqslant |g'(t)| + \sum_{k=1}^{\infty} \frac{x^{2k}}{(k-1)!(\theta t)^{k+1}} \exp\left(-\frac{1}{2}t^{-2}\right) \\ &= |g'(t)| + \frac{x^2}{(\theta t)^2} \exp\left[\frac{1}{t}\left(\frac{x^2}{\theta} - \frac{1}{2}t^{-2}\right)\right].\end{aligned}$$

Therefore, the series $\sum\limits_{k=0}^{\infty} \dfrac{g^{(k+1)}(t)}{(2k)!}x^{2k}$ converges uniformly on any bounded set of $\mathbb{R}$; thus

$$u_t(x,t) = \sum_{k=0}^{\infty} \frac{g^{(k+1)}(t)}{(2k)!}x^{2k} = \sum_{k=1}^{\infty} \frac{g^{(k)}(t)}{(2k-2)!}x^{2k-2} = u_{xx}(x,t) \qquad \forall\, t > 0, x \in \mathbb{R}\,.$$

This implies that u satisfies the heat equation

$$\begin{aligned} u_t - u_{xx} &= 0 \qquad \text{in} \quad \mathbb{R} \times (0,\infty)\,, \\ u &= 0 \qquad \text{on} \quad \mathbb{R} \times \{t = 0\}\,. \end{aligned}$$

Note that using the Fourier transform method to solve the PDE above we obtain trivial solution. The reason for not seeing the solution given by (5.14) using the Fourier transform method is that the Fourier transform of the function u given by (5.14) does not exists.