Fourier Analysis

MA746: Fourier Analysis, July–November, 2025

Rajesh Srivastava

Department of Mathematics, IIT Guwahati

Updated on March 24, 2026

Contents

Introduction

Fourier analysis is, at its core, the systematic study of how a function decomposes into elementary oscillations. The guiding principle is that *translation-invariant* phenomena are best understood in a basis of *characters* (complex exponentials), because translations act diagonally on those building blocks. This idea originates in classical boundary–value problems, where separation of variables leads naturally to eigenfunction expansions; it has since become a central viewpoint in modern analysis, linking partial differential equations, approximation theory, probability, signal processing, and geometry.

These lecture notes were prepared for the course MA746 (Fourier Analysis). They develop Fourier analysis in three complementary settings:

- the *periodic* setting, where a function is described by a discrete spectrum (Fourier series);
- the *Euclidean* setting, where a non-periodic function is analyzed via a continuous spectrum (Fourier transform);
- the *distributional* setting, which extends differentiation and Fourier methods beyond classical functions.

A recurring theme is the interplay between *physical space* and *frequency space*: regularity, decay, and localization of a function are reflected in corresponding properties of its Fourier coefficients or Fourier transform, and conversely. Convolution appears throughout as the natural operation encoding translation invariance, while summability kernels and approximate identities provide a robust mechanism for recovering functions from their spectral data.

Organization of the notes.

- **Chapter 1: Fourier Series.** We identify functions on the circle with 2π-periodic functions on $\mathbb{R}$ and develop the basic theory of Fourier series. After motivating examples from PDE, we study Fourier coefficients, uniqueness, and the Riemann–Lebesgue lemma. We then address the central question of *reconstruction*: in what sense does $\sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{in\theta}$ recover f? This leads naturally to summability methods (Abel/Poisson and Fejér/Cesàro means) and to Hilbert space techniques in $L^2(S^1)$, where orthogonality and completeness yield the cleanest theory.

- **Chapter 2: The Fourier Transform.** On $\mathbb{R}^n$, the translation group is non-compact, so the spectrum becomes continuous. We introduce the Fourier transform, establish its elementary properties (translations, modulations, scaling), and prove foundational results such as inversion and Plancherel's theorem. We emphasize the role of *good kernels* and approximate identities, which provide a unified framework for inversion and convergence, and we develop the basic L^p inequalities that govern convolution and Fourier decay.

- **Chapter 3: Distributions.** We introduce distributions as continuous linear functionals on spaces of test functions and develop the basic calculus: distributional derivatives, multiplication by smooth functions, convolution with test functions, and support. This framework clarifies how Fourier analysis interacts with weak notions of differentiation and enables one to formulate and solve problems where classical pointwise definitions are unavailable.

Prerequisites and conventions. The presentation assumes familiarity with Lebesgue integration and L^p spaces, basic analysis in $\mathbb{R}^n$, and the language of normed and Hilbert spaces. We adopt the standard identification $S^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$ and often regard a function on S^1 as a 2π-periodic function on $\mathbb{R}$. Fourier coefficients are normalized by

$$\widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-in\theta}\,d\theta,$$

This normalization makes the exponentials $\{e^{in\theta}\}_{n\in\mathbb{Z}}$ an orthonormal system in $L^2(S^1)$ with respect to the inner product

$$\langle f, g\rangle := \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\overline{g(\theta)}\,d\theta.$$

For the Fourier transform on $\mathbb{R}^n$ we adopt the *unitary* (probability/signal-processing) convention

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\,e^{-2\pi i\,x\cdot\xi}\,dx, \qquad f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi)\,e^{2\pi i\,x\cdot\xi}\,d\xi,$$

whenever the expressions make sense (e.g. $f \in \mathcal{S}(\mathbb{R}^n)$ or $f \in L^1 \cap L^2$), and in particular $\|\widehat{f}\|_2 = \|f\|_2$ (Plancherel).

Syllabus roadmap. The selection of topics and the order of presentation are meant to match a standard first graduate syllabus in Fourier analysis. For quick navigation, the main results appear in the following places:

- **Fourier series on S^1 (Chapter 1):** orthogonality and Fourier coefficients; Dirichlet and summability kernels (Fejér and Poisson); reconstruction and convergence in L^2 and at Lebesgue points; uniqueness; and classical applications (vibrating string, harmonic extension, and a Fourier-analytic proof of the isoperimetric inequality).
- **Fourier transform on $\mathbb{R}^n$ (Chapter 2):** definition and basic properties on L^1; the Riemann–Lebesgue lemma; approximate identities and "good kernels" on $\mathbb{R}^n$; Fourier inversion and Plancherel; convolution estimates (Young); interpolation (Riesz–Thorin) and Hausdorff–Young; Poisson summation; and selected applications (Gaussian integrals and the heat kernel).
- **Distributions (Chapter 3):** the space of test functions and its locally convex topology; distributions and their calculus (derivatives, multiplication, support); tempered distributions and the Fourier transform on $\mathcal{S}'(\mathbb{R}^n)$.

Each chapter ends with exercises designed both for practice and for extending the theory beyond the core statements proved in the text.

Chapter 1

Fourier Series

Fourier series provide a canonical way to represent periodic functions as superpositions of the basic characters *of the circle group, namely the complex exponentials* e^{inx}*,* $n \in \mathbb{Z}$*. Beyond their striking applications to boundary–value problems in physics, Fourier series form a central tool of analysis: they convert questions about a function into questions about its frequency spectrum* $\{\hat{f}(n)\}_{n\in\mathbb{Z}}$.

Chapter roadmap. This chapter moves from the spectral motivation coming from classical PDE to the analytic machinery of Fourier series on the circle. The logical progression is: define coefficients and partial sums; prove that coefficients determine the function; study reconstruction through summability kernels; then pass to the Hilbert-space framework, where orthogonality yields Parseval, mean-square convergence, and a representative geometric application.

1.1 Motivation: eigenfunction expansions in PDE

A guiding principle of Fourier analysis is that *translation-invariant* linear problems diagonalize in a basis of exponential functions. On $\mathbb{R}/2\pi\mathbb{Z}$ these exponentials are precisely e^{inx}, $n \in \mathbb{Z}$. One classical route to this conclusion is separation of variables in boundary–value problems.

The vibrating string on an interval

Consider the one-dimensional wave equation on the interval $(0, \pi)$ with Dirichlet boundary conditions,

$$u_{tt}(x,t) = u_{xx}(x,t), \qquad u(0,t) = u(\pi,t) = 0.$$

Seeking separable solutions $u(x,t) = X(x)T(t)$ and dividing by XT yields

$$\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} = -\lambda$$

for some constant $\lambda \in \mathbb{R}$, hence

$$X'' + \lambda X = 0, \qquad T'' + \lambda T = 0.$$

The boundary conditions force $X(0) = X(\pi) = 0$, so nontrivial solutions occur exactly for $\lambda = n^2$ with $n \in \mathbb{N}$, with eigenfunctions

$$X_n(x) = \sin(nx), \qquad n \in \mathbb{N}.$$

By linearity, one is led to expansions of the form

$$f(x) \ \sim \ \sum_{n=1}^{\infty} A_n \sin(nx),$$

where $f(x) = u(x,0)$ is the initial displacement and the coefficients are determined using orthogonality:

$$A_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx)\, dx.$$

A second initial condition $u_t(x,0) = g(x)$ determines the coefficients in front of $\sin(nx)\sin(nt)$ (or $\sin(nx)\cos(nt)$), and it already hints at a central theme: *regularity of the data controls decay of the coefficients.*

The Dirichlet problem on the disc

A second motivation comes from the Laplace equation on the unit disc

$$\Delta u = 0 \quad \text{on } D = \{(r,\theta) : 0 \le r < 1,\ 0 \le \theta < 2\pi\}, \qquad u(1,\theta) = f(\theta),$$

the classical Dirichlet problem. Writing $u(r,\theta) = F(r)G(\theta)$ leads to

$$G'' + \lambda G = 0, \qquad r^2F'' + rF' - \lambda F = 0.$$

Periodicity in θ forces $\lambda = n^2$ with $n \in \mathbb{Z}$, so $G(\theta) = e^{in\theta}$, and boundedness at $r = 0$ selects the radial solutions $F(r) = r^{|n|}$. Thus a bounded harmonic function admits an expansion

$$u(r,\theta) = \sum_{n\in\mathbb{Z}} a_n\, r^{|n|} e^{in\theta}, \qquad \text{so that} \qquad f(\theta) = u(1,\theta) \;\sim\; \sum_{n\in\mathbb{Z}} a_n e^{in\theta}.$$

Question 1.1 (Central question)**.** Given a function f on the circle (for instance $f \in L^1(S^1)$ or $f \in C(S^1)$), in what sense does the Fourier series

$$\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{in\theta}, \qquad \hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-in\theta}\, d\theta,$$

recover f?

We now formalize the identification of functions on the unit circle with 2π-periodic functions on $\mathbb{R}$, and then develop the basic tools needed to answer the question above.

1.2 Functions on the circle

Throughout these notes we identify the unit circle

$$S^1 := \{e^{it} : t \in \mathbb{R}\}$$

with the quotient group $\mathbb{R}/2\pi\mathbb{Z}$ via the map $t \mapsto e^{it}$. Under this identification, a function $f : S^1 \to \mathbb{C}$ may be viewed as a 2π-periodic function (again denoted by f) on $\mathbb{R}$.

The Lebesgue measure on S^1 corresponds to the usual Lebesgue measure dt on $[0, 2\pi)$ and is the (unique) translation-invariant probability measure up to scaling (Haar measure) on the circle. We use the normalization

$$\int_{S^1} f(t)\, dt := \int_0^{2\pi} f(t)\, dt, \qquad \text{so that} \qquad \int_{S^1} 1\, dt = 2\pi.$$

In particular, for every $t_0 \in S^1$ and every integrable f we have the translation invariance

$$\int_{S^1} f(t - t_0)\, dt = \int_{S^1} f(t)\, dt,$$

which follows immediately from the substitution $s = t - t_0$ and 2π-periodicity.

A *trigonometric polynomial* of degree at most N is an expression

$$P_N(t) = \sum_{k=-N}^{N} a_k e^{ikt},$$

and a *trigonometric series* is a formal sum $\sum_{k\in\mathbb{Z}} a_k e^{ikt}$.

Definition 1.2. Let $f \in L^1(S^1)$ and $n \in \mathbb{Z}$. The nth Fourier coefficient of f is defined by

$$\widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt.$$

Definition 1.3. The *Fourier series* of $f \in L^1(S^1)$ is the formal expansion

$$\sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{int}.$$

Its nth partial sum is

$$S_n(f)(t) := \sum_{k=-n}^{n} \widehat{f}(k)e^{ikt},$$

which is a trigonometric polynomial of degree at most n.

Example 1.4 (Model coefficients)**.** For $m \in \mathbb{Z}$, the character e^{imt} has a single nonzero Fourier coefficient:

$$\widehat{(e^{imt})}(n) = \delta_{mn}.$$

Consequently, for $m \geq 1$,

$$\widehat{\cos(mt)}(n) = \frac{1}{2}\Big(\delta_{n,m} + \delta_{n,-m}\Big), \qquad \widehat{\sin(mt)}(n) = \frac{1}{2i}\Big(\delta_{n,m} - \delta_{n,-m}\Big).$$

Thus the Fourier coefficients literally record the frequency content of the basic oscillations.

Remark 1.5. This elementary computation should be kept in mind throughout the chapter: Fourier coefficients are coordinates with respect to the exponential basis, and the later uniqueness and reconstruction theorems explain how much of the original function is encoded by these coordinates.

Lemma 1.6. *Let $f, g \in L^1(S^1)$. Then, for every $n \in \mathbb{Z}$,*

(i) $\widehat{f+g}(n) = \widehat{f}(n) + \widehat{g}(n)$,

(ii) $\widehat{\alpha f}(n) = \alpha\widehat{f}(n)$ *for every* $\alpha \in \mathbb{C}$,

(iii) $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$*,*

(iv) if $\tau_{t_0} f(t) := f(t - t_0)$ *with* $t_0 \in S^1$*, then*

$$\widehat{\tau_{t_0} f}(n) = e^{-int_0} \widehat{f}(n),$$

(v) $|\widehat{f}(n)| \le \frac{1}{2\pi} \int_0^{2\pi} |f(t)|\, dt = \|f\|_1$.

Corollary 1.7. *If* $f_j \in L^1(S^1)$ *and* $\|f_j - f\|_1 \to 0$*, then for each* $n \in \mathbb{Z}$ *we have* $\widehat{f_j}(n) \to \widehat{f}(n)$*. The convergence is uniform in* n *after multiplying by any bounded sequence of unimodular factors.*

Theorem 1.8. *Let* $f\colon [0, 2\pi] \to \mathbb{C}$*. Then* f *is absolutely continuous if and only if* f' *exists almost everywhere,* $f' \in L^1([0, 2\pi])$*, and*

$$f(x) = f(0) + \int_0^x f'(t)\, dt, \qquad 0 \le x \le 2\pi.$$

Remark 1.9. A proof may be found, for example, in Carothers, p. 374.

Theorem 1.10. *Let* $f \in L^1(S^1)$ *with* $\widehat{f}(0) = 0$*, and define*

$$F(t) := \int_0^t f(s)\, ds.$$

Then F *extends to a continuous* 2π*-periodic function on* $\mathbb{R}$*,* F *is absolutely continuous on* $[0, 2\pi]$*, and*

$$\widehat{F}(n) = \frac{\widehat{f}(n)}{in}, \qquad n \ne 0.$$

Proof. For $t, t_0 \in [0, 2\pi]$ we have

$$|F(t) - F(t_0)| \le \int_{\min\{t, t_0\}}^{\max\{t, t_0\}} |f(s)|\, ds,$$

and the right-hand side tends to 0 as $t \to t_0$. Hence F is continuous on $[0, 2\pi]$. Since

$$F(t + 2\pi) - F(t) = \int_t^{t+2\pi} f(s)\, ds = \int_0^{2\pi} f(s)\, ds = 2\pi \widehat{f}(0) = 0,$$

F is 2π-periodic and therefore extends continuously to S^1.

Moreover, for every partition $0 = t_0 < t_1 < \cdots < t_m = 2\pi$,

$$\sum_{j=1}^m |F(t_j) - F(t_{j-1})| \le \sum_{j=1}^m \int_{t_{j-1}}^{t_j} |f(s)|\, ds = \int_0^{2\pi} |f(s)|\, ds,$$

so F is absolutely continuous and $F' = f$ almost everywhere on $[0, 2\pi]$. For $n \neq 0$, integration by parts therefore gives

$$\widehat{F}(n) = \frac{1}{2\pi}\int_0^{2\pi} F(t)e^{-int}\,dt = \frac{1}{2\pi in}\int_0^{2\pi} F'(t)e^{-int}\,dt = \frac{\widehat{f}(n)}{in},$$

as claimed. □

Example 1.11. Let $f(\theta) = \theta$, $-\pi \le \theta < \pi$. Then

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \theta e^{-in\theta} d\theta = \frac{(-1)^{n+1}}{in}, \quad n \neq 0.$$

$\hat{f}(0) = 0$. Thus,

$$f(\theta) \sim \sum \frac{(-1)^{n+1}}{in} e^{in\theta} = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\,\sin(n\theta).$$

The series on the right converges for every θ, but identifying its sum with f requires convergence theory; we return to this issue after introducing summability methods and the Dirichlet/Fejér kernels.

Example 1.12. $f(\theta) = \frac{(\pi-\theta)^2}{4}$, $0 \le \theta \le 2\pi$

$$f(\theta) \sim \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{\cos n\theta}{n^2}$$

The Fourier Series is uniformly convergent, but it converges to $f(\theta)$ is not easy.

Theorem 1.13 (Convolution and Fourier coefficients)**.** *Let $f, g \in L^1(S^1)$ and define*

$$(f * g)(t) := \frac{1}{2\pi}\int_0^{2\pi} f(t-s)g(s)\,ds.$$

Then $f * g \in L^1(S^1)$ and

$$\|f * g\|_{L^1(S^1)} \le \|f\|_{L^1(S^1)}\,\|g\|_{L^1(S^1)}.$$

Moreover, for every $n \in \mathbb{Z}$,

$$\widehat{f * g}(n) = \widehat{f}(n)\,\widehat{g}(n).$$

Proof. Tonelli's theorem gives

$$\int_0^{2\pi} |(f * g)(t)|\,dt \le \frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi} |f(t-s)|\,|g(s)|\,ds\,dt.$$

Since translation preserves the L^1-norm on S^1,

$$\int_0^{2\pi} |f(t-s)|\,dt = \int_0^{2\pi} |f(t)|\,dt = 2\pi\|f\|_1$$

for every s, and therefore

$$\|f*g\|_1 \le \frac{1}{2\pi}\int_0^{2\pi}(2\pi\|f\|_1)|g(s)|\,ds = \|f\|_1\|g\|_1.$$

For the Fourier coefficients, Fubini's theorem applies because the preceding estimate shows the integrand is absolutely integrable. Hence

$$\begin{aligned}
\widehat{f*g}(n) &= \frac{1}{2\pi}\int_0^{2\pi}(f*g)(t)e^{-int}\,dt \\
&= \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} f(t-s)g(s)e^{-int}\,ds\,dt \\
&= \frac{1}{4\pi^2}\int_0^{2\pi} g(s)e^{-ins}\left(\int_0^{2\pi} f(t-s)e^{-in(t-s)}\,dt\right)ds \\
&= \frac{1}{4\pi^2}\int_0^{2\pi} g(s)e^{-ins}(2\pi\widehat{f}(n))\,ds = \widehat{f}(n)\widehat{g}(n).
\end{aligned}$$

□

Proposition 1.14 (Reconstruction operators are convolution operators)**.** *Let $f \in L^1(S^1)$.*

(i) If $P(t) = \sum_{m=-N}^{N} c_m e^{imt}$ is a trigonometric polynomial, then

$$(P*f)(t) = \sum_{m=-N}^{N} c_m\widehat{f}(m)e^{imt}.$$

In particular, $P*f$ is again a trigonometric polynomial.

(ii) If

$$D_N(t) := \sum_{m=-N}^{N} e^{imt}, \qquad S_N(f)(t) := \sum_{m=-N}^{N} \widehat{f}(m)e^{imt},$$

then

$$S_N(f) = D_N * f.$$

Moreover,

$$D_N(t) = \frac{\sin((N+\frac{1}{2})t)}{\sin(t/2)} \quad (t \ne 0), \qquad D_N(0) = 2N+1.$$

(iii) For $0 \le r < 1$, let

$$P_r(\theta) := \sum_{m\in\mathbb{Z}} r^{|m|} e^{im\theta} = \frac{1-r^2}{1-2r\cos\theta + r^2}.$$

Then the series converges absolutely and uniformly, $\widehat{P_r}(m) = r^{|m|}$, and the Abel mean of the Fourier series of f is

$$A_r f(\theta) := \sum_{m\in\mathbb{Z}} r^{|m|} \widehat{f}(m) e^{im\theta} = (P_r * f)(\theta).$$

Proof. For a single exponential $e_m(t) := e^{imt}$,

$$(e_m * f)(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{im(t-s)} f(s)\, ds = e^{imt}\widehat{f}(m).$$

By linearity this proves (i). Part (ii) is the special case $P = D_N$. The closed form for D_N follows from summing a geometric series. For (iii), absolute and uniform convergence are immediate from

$$\sum_{m\in\mathbb{Z}} r^{|m|} = 1 + 2\sum_{m=1}^{\infty} r^m < \infty,$$

so termwise integration is legitimate and yields $\widehat{P_r}(m) = r^{|m|}$. Applying part (i) to the uniformly convergent Fourier series of P_r gives the formula for $A_r f$. □

Remark 1.15 (Why kernels enter the reconstruction problem)*.* Part (ii) identifies the ordinary partial sums with convolution by the Dirichlet kernel, while part (iii) identifies Abel means with convolution by the Poisson kernel. The reconstruction problem for Fourier series is therefore the problem of deciding whether one can choose kernels K for which $f * K$ returns to f in a useful sense: pointwise, uniformly, or in L^1.

Definition 1.16. A sequence of functions $\{K_n\}_{n=1}^{\infty}$ is called a family of "good kernels" if

(i) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(t)\, dt = 1$ for all $n \ge 1$;

(ii) there exists $M > 0$ such that $\frac{1}{2\pi}\int_{-\pi}^{\pi} |K_n(t)|\, dt \le M$ for all $n \ge 1$;

(iii) for each $\delta > 0$,

$$\int_{\delta < |t| \le \pi} |K_n(t)|\, dt \longrightarrow 0 \qquad (n \to \infty).$$

The point is that such kernels are normalized, uniformly L^1-bounded, and asymptotically concentrated near the origin.

Theorem 1.17 (Approximate identities on S^1). *Let $\{K_n\}_{n\geq 1}$ be a family of good kernels on S^1, and let $f \in \mathcal{R}[-\pi, \pi]$. Then, at every continuity point x of f,*

$$(f * K_n)(x) \longrightarrow f(x).$$

If $f \in C(S^1)$, then in fact $f * K_n \to f$ uniformly on S^1.

Proof. Fix a continuity point x of f and write

$$(f * K_n)(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(y)\Big(f(x-y) - f(x)\Big)\, dy,$$

where we used property (i) of a good kernel. Let $\varepsilon > 0$. By continuity of f at x, choose $\delta \in (0, \pi)$ such that

$$|f(x-y) - f(x)| < \varepsilon \qquad (|y| < \delta).$$

Since f is Riemann integrable on a compact interval, it is bounded; write $|f| \leq B$. Splitting the integral into the regions $|y| < \delta$ and $\delta \leq |y| \leq \pi$, we obtain

$$\begin{aligned} |(f * K_n)(x) - f(x)| &\leq \frac{\varepsilon}{2\pi} \int_{|y|<\delta} |K_n(y)|\, dy + \frac{2B}{2\pi} \int_{\delta\leq|y|\leq\pi} |K_n(y)|\, dy \\ &\leq M\varepsilon + \frac{B}{\pi} \int_{\delta\leq|y|\leq\pi} |K_n(y)|\, dy, \end{aligned}$$

where M is the uniform L^1-bound from property (ii). Property (iii) sends the second term to 0, so

$$\limsup_{n\to\infty} |(f * K_n)(x) - f(x)| \leq M\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $(f * K_n)(x) \to f(x)$.

If f is continuous on S^1, then it is uniformly continuous, so one can choose the same δ for every x. The estimate above is therefore uniform in x, which yields uniform convergence. □

Corollary 1.18. *If $\{K_n\}_{n\geq 1}$ is a family of good kernels in $L^1(S^1)$ and $f \in L^1(S^1)$, then*

$$\|f * K_n - f\|_{L^1(S^1)} \longrightarrow 0 \qquad (n \to \infty).$$

Proof. Fix $\varepsilon > 0$. By density of $C(S^1)$ in $L^1(S^1)$, choose $g \in C(S^1)$ with

$$\|f - g\|_{L^1(S^1)} < \varepsilon. \tag{1.1}$$

By Young's inequality on S^1 and the uniform L^1-bound on K_n,

$$\|(f-g)*K_n\|_1 \le \|f-g\|_1\,\|K_n\|_1 \le M\varepsilon$$

for some constant M independent of n. Since g is continuous, Theorem 1.17 gives uniform convergence $g*K_n \to g$, hence also

$$\|g*K_n - g\|_1 \le 2\pi\,\|g*K_n - g\|_\infty \longrightarrow 0. \tag{1.2}$$

Therefore,

$$\|f*K_n - f\|_1 \le \|(f-g)*K_n\|_1 + \|g*K_n - g\|_1 + \|g-f\|_1 \le (M+1)\varepsilon + \|g*K_n - g\|_1.$$

Letting $n\to\infty$ and then $\varepsilon \downarrow 0$ proves the claim. □

Remark 1.19 (The Dirichlet kernel is not a good kernel). The Dirichlet kernel

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = \frac{\sin\!\big((n+\frac{1}{2})t\big)}{\sin(t/2)} \qquad (t\neq 0)$$

fails to satisfy the uniform L^1-bound (ii). Indeed, since $|\sin u| \le |u|$, we have

$$\begin{aligned}\int_{-\pi}^{\pi} |D_n(t)|\,dt &\ge \frac{2}{\pi}\int_0^{\pi} \Big|\sin\!\big((n+\tfrac{1}{2})t\big)\Big|\,\frac{dt}{t} = \frac{2}{\pi}\int_0^{(n+\frac{1}{2})\pi} \frac{|\sin s|}{s}\,ds \\ &\ge \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \xrightarrow[n\to\infty]{} \infty.\end{aligned}$$

In particular, $\sup_n \|D_n\|_{L^1(S^1)} = \infty$, although

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\,dt = 1.$$

This explains why ordinary partial sums are much harder to control than averaged ones.

Fejér kernels and Cesàro summation

Define the Fejér kernel by

$$F_n(t) := \frac{1}{n}\sum_{k=0}^{n-1} D_k(t), \qquad n\ge 1.$$

The associated Cesàro means are

$$\sigma_n(f)(x) := \frac{1}{n}\sum_{k=0}^{n-1} S_k(f)(x) = (f * F_n)(x).$$

In general, for a sequence $\{a_n\}_{n\geq 0}$ with partial sums $S_n = a_0 + \cdots + a_n$, the series $\sum_{n\geq 0} a_n$ is called *Cesàro summable* if the averages

$$\sigma_n := \frac{S_0 + S_1 + \cdots + S_{n-1}}{n}$$

converge.

Example 1.20. For the alternating series $\sum_{n=0}^{\infty}(-1)^n = 1 - 1 + 1 - 1 + \cdots$, the partial sums satisfy $S_n \in \{0, 1\}$, hence $\sigma_n \to \frac{1}{2}$.

Proposition 1.21 (Fejér kernel properties)**.** *For every $n \geq 1$ and $x \neq 0$,*

$$F_n(x) = \frac{1}{n}\frac{\sin^2(nx/2)}{\sin^2(x/2)}, \qquad F_n(0) = n.$$

In particular, $F_n \geq 0$ on S^1,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)\, dt = 1,$$

and $\{F_n\}_{n\geq 1}$ is a family of good kernels on S^1.

Proof. Write

$$\sum_{k=0}^{n-1} e^{ikx} = \frac{1 - e^{inx}}{1 - e^{ix}} \qquad (x \notin 2\pi\mathbb{Z}).$$

Taking absolute values and using

$$|1 - e^{i\theta}| = 2\left|\sin\frac{\theta}{2}\right|$$

gives

$$\left|\sum_{k=0}^{n-1} e^{ikx}\right|^2 = \frac{\sin^2(nx/2)}{\sin^2(x/2)}.$$

On the other hand,

$$\left|\sum_{k=0}^{n-1} e^{ikx}\right|^2 = \sum_{j,k=0}^{n-1} e^{i(j-k)x} = \sum_{|m|\leq n-1}(n - |m|)e^{imx} = nF_n(x),$$

which proves the closed form and also shows $F_n \geq 0$. The value at $x = 0$ follows either by continuity or directly from the definition.

Since $\widehat{D_k}(0) = 1$ for every k, we have $\widehat{F_n}(0) = 1$, equivalently

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(t)\, dt = 1.$$

Fix $\delta > 0$. Because $\sin^2(x/2)$ is bounded below by a positive constant on $\{x : \delta \leq |x| \leq \pi\}$, there exists $c_\delta > 0$ such that

$$0 \leq F_n(x) \leq \frac{1}{nc_\delta} \qquad (\delta \leq |x| \leq \pi).$$

Hence

$$\int_{\delta \leq |x| \leq \pi} F_n(x)\, dx \leq \frac{2(\pi - \delta)}{nc_\delta} \xrightarrow[n\to\infty]{} 0.$$

Because $F_n \geq 0$ and its normalized integral is 1, we also have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |F_n(t)|\, dt = 1.$$

Thus $\{F_n\}$ satisfies all three good-kernel conditions. □

Corollary 1.22 (Fejér theorem)**.** *Let $f \in L^1(S^1)$. Then*

$$\sigma_n(f) = f * F_n \longrightarrow f \quad \text{in } L^1(S^1).$$

If $f \in C(S^1)$, then $\sigma_n(f) \to f$ uniformly on S^1; and if x is a point of continuity of f, then $\sigma_n(f)(x) \to f(x)$.

Proof. The pointwise and uniform assertions follow from Theorem 1.17 together with Proposition 1.21. The L^1 statement follows from Corollary 1.18. □

Remark 1.23 (From uniqueness to reconstruction)*.* Fejér's theorem is more than a convergence result: it is the structural bridge from uniqueness to reconstruction. If all Fourier coefficients of f vanish, then every Fejér mean vanishes identically; but the Fejér means also converge back to f. Thus the same averaged kernels that reconstruct the function also prove that the coefficient data determine it uniquely.

1.3 Uniqueness Theorem

We can now make the uniqueness principle conceptually clean: once Fejér kernels are known to reconstruct L^1-functions, uniqueness becomes an immediate corollary.

Theorem 1.24 (Uniqueness of Fourier coefficients)**.** *If $f \in L^1(S^1)$ satisfies $\widehat{f}(m) = 0$ for every $m \in \mathbb{Z}$, then $f = 0$ almost everywhere on S^1.*

Proof. For each $n \geq 1$, Proposition 1.21 shows that F_n is a trigonometric polynomial of degree at most $n - 1$ with Fourier coefficients

$$\widehat{F_n}(m) = \begin{cases} 1 - \dfrac{|m|}{n}, & |m| < n, \\ 0, & |m| \geq n. \end{cases}$$

Hence, by the convolution theorem,

$$\widehat{f * F_n}(m) = \widehat{f}(m)\,\widehat{F_n}(m) = 0 \qquad \text{for every } m \in \mathbb{Z}.$$

But $f * F_n$ is itself a trigonometric polynomial of degree at most $n - 1$. A trigonometric polynomial is determined by its Fourier coefficients, so all of its coefficients being zero implies

$$f * F_n \equiv 0 \qquad (n \geq 1).$$

Now apply Corollary 1.22: since $f * F_n \to f$ in $L^1(S^1)$, we get

$$\|f\|_1 = \lim_{n\to\infty} \|f - f * F_n\|_1 = 0.$$

Therefore $f = 0$ almost everywhere. □

Remark 1.25 (Density of trigonometric polynomials). The same argument also clarifies why trigonometric polynomials are dense. If $f \in C(S^1)$, then Corollary 1.22 gives uniform approximation by the trigonometric polynomials $\sigma_n(f) = f * F_n$. If $f \in L^1(S^1)$, the same corollary gives approximation in L^1. Thus the Fejér means provide an explicit reconstruction scheme, not merely an abstract density statement.

1.4 Riemann-Lebesgue Lemma

The next basic principle is a qualitative decay statement: integrability in physical space already implies that high frequencies become negligible.

Lemma 1.26. *If* $f \in L^1(S^1)$, *then* $\lim_{|n|\to\infty} \hat{f}(n) = 0$.

Proof. For $\epsilon > 0$, there exists a trigonometric polynomial P such that $\|f - P\|_1 < \epsilon$ (where $P = f * F_n$ etc.). Let $|n| > \deg P$. Then

$$|\hat{f}(n)| = |\hat{f}(n) - \hat{P}(n)| \leq \|f - P\|_1 < \epsilon, \quad \text{if } |n| > \deg P.$$

That is, $|\hat{f}(n)| < \epsilon$ for large n. Hence, $\lim_{|n|\to\infty} \hat{f}(n) = 0$. □

Remark 1.27 (Decay versus regularity). The Riemann–Lebesgue lemma is qualitative rather than quantitative. Extra smoothness produces faster decay: one derivative typically yields $O(1/|n|)$, k derivatives yield $O(|n|^{-k})$, and analyticity leads to exponential decay. Much of Fourier analysis can be read as a precise dictionary between smoothness in physical space and decay in frequency space.

1.5 Abel Means Summability

A series $\sum_{n=0}^{\infty} a_n$ is said to be **Abel summable** to s if the series

$$A(r) = \sum_{n=0}^{\infty} a_n r^n$$

is convergent for each $0 \leq r < 1$, and $\lim_{r\to 1} A(r) = s$.

Example 1.28. Every convergent series is Abel summable. Consider

$$1 - 2 + 3 - 4 + 5 - \cdots = \sum_{n=0}^{\infty} (-1)^n (n+1).$$

Then

$$A(r) = \sum_{n=0}^{\infty} (-1)^n (n+1) r^n = \frac{1}{(1+r)^2} \to \frac{1}{4}$$

Show that the above series is **not Cesaro summable**.

Now, consider the Fourier series of $f \in R[-\pi, \pi]$ as

$$f(t) \sim \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int}$$

Let

$$A_r f(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} \hat{f}(n) e^{in\theta}$$

then

$$A_r f(\theta) = (f * P_r)(\theta)$$

where

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

Lemma 1.29 (Poisson kernels form an approximate identity)**.** *For $0 \le r < 1$, the Poisson kernel*

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}$$

is nonnegative and satisfies

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta)\, d\theta = 1.$$

Moreover, for every $\delta > 0$,

$$\int_{\delta \le |\theta| \le \pi} P_r(\theta)\, d\theta \longrightarrow 0 \qquad (r \to 1^-).$$

Consequently, $\{P_r\}_{0 \le r < 1}$ is a good-kernel family indexed by $r \to 1^-$.

Proof. The closed form shows immediately that $P_r(\theta) \ge 0$. Since $\widehat{P_r}(0) = 1$, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta)\, d\theta = 1.$$

Because $P_r \ge 0$, this also implies $\|P_r\|_{L^1(S^1)} = 1$ for every r.

Now fix $\delta > 0$. For $|\theta| \ge \delta$ and $r \in [1/2, 1)$,

$$1 - 2r\cos\theta + r^2 = (1 - r)^2 + 2r(1 - \cos\theta) \ge 1 - \cos\delta =: c_\delta > 0.$$

Hence

$$0 \le P_r(\theta) \le \frac{1 - r^2}{c_\delta} \qquad (\delta \le |\theta| \le \pi),$$

and therefore

$$\int_{\delta \le |\theta| \le \pi} P_r(\theta)\, d\theta \le \frac{2(\pi - \delta)(1 - r^2)}{c_\delta} \xrightarrow[r \to 1^-]{} 0.$$

This is exactly the concentration property required of a good kernel. ◻

Theorem 1.30 (Abel summation via the Poisson kernel)**.** *Let $f \in L^1(S^1)$ and define the Abel means*

$$A_r f(\theta) := (P_r * f)(\theta) = \sum_{n \in \mathbb{Z}} r^{|n|}\, \hat{f}(n)\, e^{in\theta}, \qquad 0 \le r < 1.$$

Then $A_r f \to f$ *in* $L^1(S^1)$ *as* $r \to 1^-$. *If* $f \in C(S^1)$, *then* $A_r f \to f$ *uniformly on* S^1. *Moreover, if* θ *is a point of continuity of* f, *then* $A_r f(\theta) \to f(\theta)$.

Proof. Assume first that $f \in C(S^1)$. Fix $\varepsilon > 0$. By uniform continuity of f, choose $\delta \in (0, \pi)$ such that

$$|f(\theta - t) - f(\theta)| < \varepsilon \qquad (|t| < \delta,\ \theta \in S^1).$$

Using the normalization of P_r and splitting the convolution integral,

$$\begin{aligned} |A_r f(\theta) - f(\theta)| &\le \frac{1}{2\pi} \int_{|t|<\delta} P_r(t) |f(\theta - t) - f(\theta)| \, dt \\ &\quad + \frac{1}{2\pi} \int_{\delta \le |t| \le \pi} P_r(t) |f(\theta - t) - f(\theta)| \, dt \\ &\le \varepsilon \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(t) \, dt + 2\|f\|_\infty \frac{1}{2\pi} \int_{\delta \le |t| \le \pi} P_r(t) \, dt. \end{aligned}$$

The first term equals ε, while the second tends to 0 by Lemma 1.29. Because the estimate is uniform in θ, we obtain $A_r f \to f$ uniformly.

If $f \in L^1(S^1)$, fix $\varepsilon > 0$ and choose $g \in C(S^1)$ with $\|f - g\|_1 < \varepsilon$. Since $\|P_r\|_1 = 1$, Young's inequality yields

$$\|(f - g) * P_r\|_1 \le \|f - g\|_1 \|P_r\|_1 < \varepsilon.$$

Also, the continuous case shows $g * P_r \to g$ uniformly, hence in L^1. Therefore

$$\|A_r f - f\|_1 \le \|(f - g) * P_r\|_1 + \|g * P_r - g\|_1 + \|g - f\|_1,$$

and the right-hand side is eventually at most $2\varepsilon + o(1)$. This proves $A_r f \to f$ in L^1.

Finally, if θ is merely a continuity point of f, the same ε-δ decomposition as above, now with δ chosen only for that point, gives $A_r f(\theta) \to f(\theta)$. □

Corollary 1.31 (Abel summability)**.** *For every* $f \in L^1(S^1)$, *the Fourier series of* f *is Abel summable to* f *in* $L^1(S^1)$. *If* $f \in C(S^1)$, *it is Abel summable to* f *uniformly.*

Theorem 1.32 (Poisson integral and the Dirichlet problem)**.** *Let* $f \in C(S^1)$ *and define*

$$U(r, \theta) := (P_r * f)(\theta), \qquad 0 \le r < 1,\ \theta \in [-\pi, \pi].$$

Then $U \in C^\infty(D)$ *is harmonic on the unit disc* D *and extends continuously to* $\overline{D}$ *with boundary values* $\lim_{r \to 1^-} U(r, \theta) = f(\theta)$ *uniformly in* θ. *Moreover, if* v *is a bounded* *harmonic function on D such that $v(r,\theta) \to f(\theta)$ as $r \to 1^-$ uniformly in θ, then $v \equiv U$ on D.*

Proof. Since

$$U(r,\theta) = \sum_{n\in\mathbb{Z}} r^{|n|}\hat{f}(n)e^{in\theta},$$

the series converges absolutely and uniformly on $\{(r,\theta) : 0 \le r \le r_0\}$ for each $r_0 < 1$, and the same holds for all ∂_r and ∂_θ derivatives. Termwise differentiation therefore yields $\Delta U = 0$ on D.

The boundary convergence statements are exactly those of Theorem 1.30.

For uniqueness, let v be as in the statement and define the Fourier coefficients

$$a_n(r) := \frac{1}{2\pi}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\,d\theta, \qquad n \in \mathbb{Z},\ 0 < r < 1.$$

Differentiating under the integral sign gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\partial^2 v}{\partial\theta^2}(r,\theta)\,e^{-in\theta}\,d\theta = -n^2 a_n(r).$$

Using $\Delta v = v_{rr} + \frac{1}{r}v_r + \frac{1}{r^2}v_{\theta\theta} = 0$ we obtain the ODE

$$a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r) = 0, \qquad 0 < r < 1.$$

For $n \neq 0$ the solutions are of the form $a_n(r) = A_n r^{|n|} + B_n r^{-|n|}$; boundedness of v on D forces $B_n = 0$, so $a_n(r) = A_n r^{|n|}$. For $n = 0$, the ODE reduces to $a_0'' + \frac{1}{r}a_0' = 0$, hence a_0 is constant.

The hypothesis $v(r,\theta) \to f(\theta)$ uniformly implies $a_n(r) \to \hat{f}(n)$ as $r \to 1^-$ for each fixed n, and therefore $A_n = \hat{f}(n)$. Consequently

$$v(r,\theta) = \sum_{n\in\mathbb{Z}} a_n(r)e^{in\theta} = \sum_{n\in\mathbb{Z}} \hat{f}(n)r^{|n|}e^{in\theta} = U(r,\theta),$$

as claimed. □

Remark 1.33 (Regularity gain by integration). The condition $\widehat{f}(0) = 0$ is exactly the compatibility condition needed to integrate a periodic function and remain periodic. On the Fourier side, passing from f to its primitive F divides the nonzero coefficients by n; this is the first concrete manifestation of the principle that integration improves spectral decay.

Exercise 1.34. If $\{J_n\}_{n=1}^{\infty}$ and $\{K_n\}_{n=1}^{\infty}$ are two families of good kernels for $L^1(S^1)$, then $\{J_n * K_n\}_{n=1}^{\infty}$ is a good kernel for $L^1(S^1)$.

(i)

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} J_n * k_n(t)dt &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\int_{-\pi}^{\pi} J_n(t-s)k_n(s)dsdt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} J_n(t-s)dt\right)k_n(s)ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1\cdot k_n(s)ds \quad (\text{ since } L^1(S^1) \text{ is translation invariant}) \\
&= 1
\end{aligned}$$

(ii)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|J_n * k_n(t)|dt \le \frac{1}{2\pi}\int_{-\pi}^{\pi} M|k_n(s)|ds \le MN < \infty$$

(iii) Let $\delta > 0$, then

$$\int_{\delta<|t|\le\pi}|K_n * J_n(t)|dt \le \int_{s=-\pi}^{\pi}\left(\int_{\delta<|t|\le\pi}|K_n(t-s)|dt\right)|J_n(s)|ds$$

Let $|s| < \delta/2$, then $r = t-s \in (-\delta/2, \delta/2)$. Now

$$(**)\int_{|s|<\delta/2}\left(\int_{\delta/2<|r|<\pi}|K_n(r)|dr\right)|J_n(s)|ds \to 0 \text{ as } n\to\infty,$$

since $\int_{\delta/2<|s-t|<\pi}|K_n(t-s)|dt \to 0$ as $n\to\infty$. (Exercise)

Since $|s| < \delta/2$, (use the fact that $\tau_x f \to f$ is continuous on $L^1(S^1)$). That is, if

$$\int_{\delta<|t|\le\pi}|K_n(t)|dt \to 0 \quad \text{for all } \delta > 0,$$

then

$$\left|\int_{\delta<|t|\le\pi}(\tau_s K_n(t) - K_n(t))dt\right| < \int_{\delta<|t|\le\pi}|(\tau_s K_n(t) - K_n(t))|\,dt \le \epsilon$$

For $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$, such that $\int_{|t|>\delta}|K_n(t)|dt < \epsilon$ for all $n \ge n_0$ and for small $|s| < \delta^1$. However,

$$\int_{|s|>\delta/2}\int_{|t|>\delta}|K_n(t-s)||J_n(s)|ds\ dt \le \int_{|s|>\delta/2} M|J_n(s)|ds \to 0 \text{ as } n\to\infty.$$

Lemma 1.35. *Let* $f : [-\pi, \pi] \to \mathbb{C}$ *be such that*

$$|f(x) - f(y)| \leq M|x - y| \quad \textit{for all } x, y \in [-\pi, \pi]$$

for some $M > 0$*. Then* $S_n(f) \to f$ *uniformly. Note that* $|x-y| = \min\{|x-y|, |x-y\pm 2\pi|\}$*, that is, the distance between* x *and* y *modulo* 2π*.*

Proof. Calculate

$$S_n(f)(x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}(f(x-t) - f(x))D_n(t)dt.$$

Since

$$D_n(t) = \frac{\sin((n+\frac{1}{2})t)}{\sin(t/2)}, \quad t \neq 0,$$

$$\begin{aligned}|S_n(f)(x) - f(x)| \leq \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}(f(x-t) - f(x))\frac{\cos t/2}{\sin t/2}\sin nt\, dt\right| \\ +\frac{1}{2\pi}\left|\int_{-\pi}^{\pi}(f(x-t) - f(x))\cos nt\, dt\right|.\end{aligned}$$

Let

$$g(t) = \frac{f(x+t) - f(x)}{t/2}\cos\frac{t}{2}, \quad t \neq 0.$$

Then $|g(t)| \leq 2M|\frac{t/2}{\sin t/2}|$, if $t \neq 0$.

Since $\lim_{t\to 0}\frac{t/2}{\sin(t/2)} = 1$, it follows that g is a bounded function on $[-\pi, \pi]$ and continuous on $[-\pi, \pi] \setminus \{0\}$. Hence, $g \in R[-\pi, \pi]$.

Let $h(t) = f(x-t) - f(x)$. Then

$$\begin{aligned}|S_n(f)(x) - f(x)| &\leq \frac{1}{2\pi}|\int_{-\pi}^{\pi} g(t)\sin(nt)dt| + \frac{1}{2\pi}|\int_{-\pi}^{\pi} h(t)\cos(nt)dt| \\ &= \frac{1}{2}|\hat{g}(n) - \hat{g}(-n)| + \frac{1}{2}|\hat{h}(n) + \hat{h}(-n)| \to 0 \text{ (by R-L Lemma)}\end{aligned}$$

whenever $x \in [-\pi, \pi]$. □

Corollary 1.36. *If* $f \in R[-\pi, \pi]$ *and* f *is differentiable at* x_0*, then* $S_n(f)(x_0) \to f(x_0)$*.*

Define $g(t) = \begin{cases} \frac{f(x_0-t)-f(x_0)}{t}, t \neq 0; \\ -f'(x_0), \text{ otherwise} \end{cases}$

Corollary 1.37. *If* $f \in C'[-\pi, \pi]$*, then* $S_n(f) \to f$ *uniformly. (Hint: Use MVT.)*

Notice that if f is piecewise C^1-function, then $S_n(f) \to f$ uniformly too.

Question 1.38. Does every continuous function f on S^1 have a Fourier series which converges to f at each point of S^1?

To discuss this, we need the following lemma.

Lemma 1.39. *Let $f \in R[-\pi, \pi]$ and f is bounded on $[-\pi, \pi]$ by M. Then there exists a sequence f_n of continuous functions on $[-\pi, \pi]$ such that*

(i) $|f_n(x)| \leq M$ for all $n \in \mathbb{N}$, $x \in [-\pi, \pi]$.

(ii) $\int_{-\pi}^{\pi} |f_n(x) - f(x)| dx \to 0$ as $n \to \infty$.

Proof. First consider f as a real-valued function. For $\epsilon > 0$, there exists a partition P of $[-\pi, \pi]$ such that

$$U(P, f) - L(P, f) < \epsilon,$$

where

$$P = \{-\pi = x_0 < x_1 < \cdots < x_i < x_{i+1} < \cdots < x_N = \pi\}$$

For $x \in [x_{i-1}, x_i]$, define $g(x) = \sup\{f(y) : x_{i-1} \leq y \leq x_i\}$. Then g is bounded by M.

$$\int_{-\pi}^{\pi} |g(x) - f(x)| dx = \int_{-\pi}^{\pi} (g(x) - f(x)) dx < \epsilon \quad \text{(by (1))}$$

Let $\delta > 0$ and $x \in (x_i - \delta, x_i + \delta)$, define $\tilde{g}(x)$ be the linear function joining $g(x - \delta)$ and $g(x + \delta)$, and $\tilde{g} = 0$ near $-\pi$ and π. Then $\tilde{g}$ is a continuous periodic function which differs with g on N many intervals, each of length less than 2δ surrounding the partitioning points. Hence,

$$\int_{-\pi}^{\pi} |g(x) - \tilde{g}(x)| \, dx \leq (2M)N(2\delta).$$

For δ sufficiently small,

$$\int_{-\pi}^{\pi} |g(x) - \tilde{g}(x)| \, dx < \epsilon.$$

$$\implies \int_{-\pi}^{\pi} |f(x) - \tilde{g}(x)| \, dx < 2\epsilon.$$

For $2\epsilon = \frac{1}{n}$, take $\tilde{g} = f_n$. Thus

$$\int_{-\pi}^{\pi} |f(x) - f_n(x)| dx \to 0 \quad \text{as} \quad n \to \infty.$$

□

Remark 1.40. If $f \in R[-\pi,\pi]$ has only finitely many discontinuities, then $\tilde{g}_n(x) \to f(x)$ pointwise.

Now let $X = C(S^1)$ and define $\Lambda_n : X \to \mathbb{C}$ by

$$\Lambda_n(f) := S_n(f)(0).$$

Then each Λ_n is a bounded linear functional and

$$|\Lambda_n(f)| \le \|D_n\|_1 \|f\|_\infty, \qquad \text{hence} \qquad \|\Lambda_n\| \le \|D_n\|_1.$$

We claim that in fact $\|\Lambda_n\| = \|D_n\|_1 = \int_{-\pi}^{\pi} |D_n(t)|\,dt$.

To see this, let $g(t) = \operatorname{sign}(D_n(-t))$. For fixed n, the function g has only finitely many discontinuities. By the previous lemma there exists a sequence $g_m \in C[-\pi,\pi]$ with $|g_m(t)| \le 1$ and $g_m(t) \to g(t)$ pointwise. Hence, by dominated convergence,

$$\begin{aligned}
\lim_{m\to\infty} \Lambda_n(g_m) &= \lim_{m\to\infty} \int_{-\pi}^{\pi} g_m(-t) D_n(t)\,dt \\
&= \int_{-\pi}^{\pi} g(-t) D_n(t)\,dt \\
&= \int_{-\pi}^{\pi} |D_n(t)|\,dt = \|D_n\|_1.
\end{aligned}$$

Therefore $\|\Lambda_n\| = \|D_n\|_1$. Since $\|D_n\|_1 \to \infty$, the sequence $\{\Lambda_n\}$ is not uniformly bounded. By the Uniform Boundedness Principle, there exists $f \in C([-\pi,\pi])$ such that

$$\sup_n |S_n(f)(0)| = \infty.$$

In particular, the Fourier series of f does not converge at 0.

By translation, the same conclusion holds at every prescribed point $x \in [-\pi,\pi]$: for each such x there exists $f \in C[-\pi,\pi]$ whose Fourier series fails to converge to $f(x)$ at x. In fact, for each x one can construct a dense subset $E_x \subset C[-\pi,\pi]$ such that $S_n(f)(x) \to \infty$ for every $f \in E_x$; see Rudin, *Real and Complex Analysis*.

1.6 Hilbert space methods and L^2-theory

The space $L^2(S^1)$ is a Hilbert space with inner product

$$\langle f, g\rangle := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\,\overline{g(\theta)}\,d\theta, \qquad \|f\|_2 := \langle f, f\rangle^{1/2}.$$

For each $n \in \mathbb{Z}$ set $e_n(\theta) := e^{in\theta}$. A direct computation shows that $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal set in $L^2(S^1)$, and the Fourier coefficients may be written as $\hat{f}(n) = \langle f, e_n \rangle$.

For $N \geq 0$ we denote by

$$\mathcal{T}_N := \operatorname{span}\{e_n : \ |n| \leq N\}$$

the space of trigonometric polynomials of degree at most N, and we recall the Nth partial sum

$$S_N(f)(\theta) = \sum_{|n|\leq N} \hat{f}(n)\, e^{in\theta}.$$

Lemma 1.41 (Best approximation / orthogonal projection)**.** *For each $N \geq 0$, the map $S_N : L^2(S^1) \to \mathcal{T}_N$ is the orthogonal projection onto $\mathcal{T}_N$. Equivalently, for every $P \in \mathcal{T}_N$ one has*

$$\|f - S_N(f)\|_2 \leq \|f - P\|_2,$$

with equality if and only if $P = S_N(f)$.

Proof. Write $P(\theta) = \sum_{|n|\leq N} c_n e^{in\theta}$. Then

$$f - P = \Big(f - S_N(f)\Big) + \sum_{|n|\leq N} \Big(\hat{f}(n) - c_n\Big) e^{in\theta}.$$

The second term on the right lies in $\mathcal{T}_N$, while $f - S_N(f)$ is orthogonal to every e_n with $|n| \leq N$. Hence the two terms are orthogonal in L^2, and Pythagoras' theorem gives

$$\|f - P\|_2^2 = \|f - S_N(f)\|_2^2 + \left\| \sum_{|n|\leq N} \Big(\hat{f}(n) - c_n\Big) e^{in\theta} \right\|_2^2,$$

which immediately implies the claimed inequality and the equality condition. □

Corollary 1.42 (Bessel inequality)**.** *For every $f \in L^2(S^1)$ and every $N \geq 0$,*

$$\sum_{|n|\leq N} |\hat{f}(n)|^2 \leq \|f\|_2^2.$$

Theorem 1.43 (Parseval and mean-square convergence)**.** *For every $f \in L^2(S^1)$ one has*

$$\|f\|_2^2 = \|f - S_N(f)\|_2^2 + \sum_{|n|\leq N} |\hat{f}(n)|^2, \qquad N \geq 0. \tag{1.3}$$

In particular,

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2, \tag{1.4}$$

and $S_N(f) \to f$ *in* $L^2(S^1)$ *as* $N \to \infty$.

Proof. Identity (1.3) is the Pythagorean identity coming from the orthogonal decomposition $f = S_N(f) + (f - S_N(f))$, together with $\|S_N(f)\|_2^2 = \sum_{|n|\le N} |\hat{f}(n)|^2$.

To prove L^2-convergence, fix $\varepsilon > 0$. Since trigonometric polynomials are dense in $L^2(S^1)$, there exists a trigonometric polynomial $P \in \mathcal{T}_M$ such that $\|f - P\|_2 < \varepsilon$. For every $N \ge M$ we have $S_N(P) = P$, and therefore

$$\|f - S_N(f)\|_2 \le \|f - P\|_2 + \|S_N(f - P)\|_2 \le 2\|f - P\|_2 < 2\varepsilon,$$

since S_N is an orthogonal projection and hence a contraction on L^2. Thus $\|f - S_N(f)\|_2 \to 0$. Taking the limit $N \to \infty$ in (1.3) yields (1.4). □

1.7 Isoperimetric problem

Theorem 1.44 (Isoperimetric inequality)**.** *Let* γ *be a simple closed curve in* $\mathbb{R}^2$ *of length* l *enclosing an area* A. *Then*

$$A \le \frac{l^2}{4\pi},$$

with equality if and only if γ *is a circle.*

Proof. By scaling we may assume $l = 2\pi$; then the claim becomes $A \le \pi$. Parametrize γ by arclength: $\gamma(t) = (x(t), y(t))$ for $t \in [0, 2\pi]$ with

$$x'(t)^2 + y'(t)^2 = 1 \quad \text{a.e. on } [0, 2\pi]. \tag{1.5}$$

Since γ is closed, x and y are 2π-periodic. Write their Fourier series in L^2,

$$x(t) \sim \sum_{n\in\mathbb{Z}} a_n e^{int}, \qquad y(t) \sim \sum_{n\in\mathbb{Z}} b_n e^{int},$$

so that $x'(t) \sim \sum_{n\in\mathbb{Z}} (in) a_n e^{int}$ and similarly for y'. Applying Parseval's identity (1.4) to x' and y' and using (1.5) gives

$$\sum_{n\in\mathbb{Z}} n^2 \Big(|a_n|^2 + |b_n|^2\Big) = 1. \tag{1.6}$$

By Green's theorem, the signed area enclosed by γ satisfies

$$A = \frac{1}{2}\left|\int_0^{2\pi}\Big(x(t)y'(t) - y(t)x'(t)\Big)\,dt\right|.$$

Using the bilinear form of Parseval's identity yields

$$A = \pi\left|\sum_{n\in\mathbb{Z}} n\Big(a_n\overline{b_n} - b_n\overline{a_n}\Big)\right|. \tag{1.7}$$

For each n we have

$$\Big|a_n\overline{b_n} - b_n\overline{a_n}\Big| \le 2|a_n||b_n| \le |a_n|^2 + |b_n|^2,$$

and therefore, from (1.7),

$$A \le \pi\sum_{n\in\mathbb{Z}}|n|\Big(|a_n|^2 + |b_n|^2\Big) \le \pi\sum_{n\in\mathbb{Z}} n^2\Big(|a_n|^2 + |b_n|^2\Big) = \pi,$$

where the last equality uses (1.6). This proves $A \le \pi$ for curves of length 2π.

If equality holds, then equality must hold in each of the preceding inequalities. In particular, $|n| = n^2$ whenever $|a_n|^2 + |b_n|^2 \neq 0$, hence $a_n = b_n = 0$ for all $|n| \ge 2$. Thus

$$x(t) = a_0 + a_1e^{it} + \overline{a_1}e^{-it}, \qquad y(t) = b_0 + b_1e^{it} + \overline{b_1}e^{-it}.$$

Moreover, equality in $2|a_1||b_1| \le |a_1|^2 + |b_1|^2$ forces $|a_1| = |b_1|$, and equality in $\Big|a_1\overline{b_1} - b_1\overline{a_1}\Big| \le 2|a_1||b_1|$ forces a phase difference of $\pm\pi/2$ between a_1 and b_1. Consequently, after a translation and a rotation of the plane, one may write

$$x(t) = a_0 + \cos(t+\alpha), \qquad y(t) = b_0 + \sin(t+\alpha),$$

which parametrizes a circle of radius 1.

Rescaling back to a curve of length l completes the proof. □

1.8 Problem Sets (Chapter 1: Fourier Series)

Conventions. Throughout these problem sets, $S^1 = \mathbb{R}/2\pi\mathbb{Z}$. Unless otherwise specified, Fourier coefficients are

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt,$$

and convolution on S^1 is normalized so that $\widehat{f * g}(n) = \widehat{f}(n)\widehat{g}(n)$.

> **Problem-set architecture.** The problems are arranged in four tiers. *Tier I* builds fluency with coefficients and basic estimates; *Tier II* develops kernels, convolution, and operator bounds; *Tier III* centers on convergence and summability; and *Tier IV* collects synthesis problems and applications. A first pass through the chapter should prioritize Tiers I–III in order; Tier IV is ideal for take-home assignments, oral-exam preparation, or project work.

Tier I. Fourier coefficients: identities, estimates, and examples

1. Determine whether each statement is **TRUE** or **FALSE**, and justify your answer rigorously.

   (a) Let D_n be the Dirichlet kernel on S^1. Must the identity $D_n * D_n = D_n$ hold (with the normalized convolution used in these notes)?

   (b) Does there exist $f \in L^1(S^1)$ such that $\sum_{n\in\mathbb{Z}} |n\,\widehat{f}(n)|^2 = \infty$?

2. Suppose $f \in C^1(S^1)$. Prove that

$$\widehat{f'}(n) = in\,\widehat{f}(n) \qquad (n \in \mathbb{Z}),$$

   and deduce that $|\widehat{f}(n)| \le C/|n|$ for all $n \neq 0$. Does the same conclusion remain valid if f is merely absolutely continuous?

3. Let f be of bounded variation on $[-\pi, \pi]$. Prove that

$$|\widehat{f}(n)| \le \frac{\mathrm{Var}(f)}{2\pi|n|} \qquad (n \in \mathbb{Z} \setminus \{0\}).$$

4. For $f \in L^1(S^1)$, establish the identity

$$\widehat{f}(n) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\Big[f(x) - f\Big(x + \frac{\pi}{n}\Big)\Big]e^{-inx}\,dx \qquad (n \neq 0),$$

and use it to prove the Riemann–Lebesgue lemma on S^1.

5. Let $f \in L^1(S^1)$ satisfy the Hölder condition

$$|f(x+h) - f(x)| \le M|h|^\alpha, \qquad 0 < \alpha < 1.$$

Show that $\widehat{f}(n) = O(|n|^{-\alpha})$.

6. Compute the Fourier series (and identify the pointwise limit) for each of the following 2π-periodic functions:

   (a) $f(t) = t$ on $(-\pi, \pi)$ (sawtooth wave).

   (b) $f(t) = \operatorname{sgn}(t)$ on $(-\pi, \pi)$.

   (c) $f(t) = |t|$ on $(-\pi, \pi)$.

   Describe the Gibbs phenomenon at a jump discontinuity (no quantitative estimate required).

7. (Parseval and classical sums.) Assume $f \in L^2(S^1)$. Prove Parseval's identity and use it to evaluate:

$$\sum_{n=1}^{\infty} \frac{1}{n^2}, \qquad \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}, \qquad \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2}.$$

Tier II. Kernels and convolution on the circle

1. (Fejér kernel as a Fourier multiplier.) Show that Fejér's kernel F_n satisfies

$$F_n(t) = \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) e^{ijt},$$

and deduce that $\widehat{F_n}(j) = \max\{1 - |j|/(n+1), 0\}$.

2. Prove that $F_n(t) \ge 0$ for all t, $\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)\,dt = 1$, and that $\{F_n\}$ is an approximate identity on S^1.

3. Given $f \in L^1(S^1)$ and $m \in \mathbb{N}$, define $f_m(t) = f(mt)$. Prove that

$$\widehat{f_m}(n) = \begin{cases} \widehat{f}(n/m), & m \mid n, \\ 0, & m \nmid n. \end{cases}$$

4. For $1 \le p < \infty$, let $\tau_x f(y) = f(y-x)$. Prove that $x \mapsto \tau_x f$ is continuous from S^1 into $L^p(S^1)$, i.e. $\|\tau_x f - f\|_p \to 0$ as $x \to 0$. Does the same conclusion hold for $p = \infty$?

5. Let $f \in L^1(S^1)$ and define $T_f(g) = f * g$ on $L^1(S^1)$. Prove that T_f is bounded and that $\|T_f\| = \|f\|_1$.

6. (Bernstein inequality.) Let P be a trigonometric polynomial of degree at most n. Prove that
$$\|P'\|_\infty \le C\, n\, \|P\|_\infty$$
for an absolute constant C (identify the best constant if you wish).

7. (Dirichlet kernel growth.) Show that $\|D_n\|_{L^1(S^1)} \to \infty$ and prove the estimate
$$\|D_n\|_1 \simeq \log(n+1)$$
(up to absolute multiplicative constants).

Tier III. Summability, convergence, and qualitative phenomena

1. Let $f \in L^1(S^1)$ and $g \in L^\infty(S^1)$. Prove that
$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, g(nt)\, dt = \widehat{f}(0)\, \widehat{g}(0).$$

2. Let f be bounded and monotone on $[-\pi, \pi]$ (extended periodically). Show that $\widehat{f}(n) = O(1/|n|)$.

3. Let f be Riemann integrable on $[-\pi, \pi]$. Prove that $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 < \infty$. Conclude that $\widehat{f}(n) = o(1)$.

4. Prove that if $\sum_{n\ge 0} a_n$ converges, then it is both Cesàro and Abel summable to the same sum.

5. Prove that if $\sum_{n\ge 0} a_n$ is Cesàro summable to σ, then it is Abel summable to σ. Give a counterexample showing that Abel summability does *not* imply Cesàro summability.

6. Let P_r be the Poisson kernel on the unit disk and define $A_r(f) = f * P_r$.
   (a) If f has a jump discontinuity at θ, prove that $\lim_{r\to 1^-} A_r(f)(\theta) = \frac{f(\theta^-)+f(\theta^+)}{2}$.
   (b) Prove the analogous statement for Fejér means $\sigma_n(f) = f * F_n$.

7. Show that there exists $f \in L^1(S^1)$ for which the partial sums $S_n(f) = D_n * f$ do *not* converge to f in L^1-norm.

8. Let $f \in L^1(S^1)$. Prove that $\|S_n(f)/n\|_1 \to 0$ as $n \to \infty$.

9. Suppose f is Riemann integrable and differentiable at t_0. Prove that $S_n(f; t_0) \to f(t_0)$.

10. Suppose $f \in L^\infty(S^1)$ satisfies $|\widehat{f}(n)| \le k/|n|$ for all $n \ne 0$. Prove that

$$|S_n(f)(t)| \le \|f\|_\infty + 2k \qquad \text{for all } t \in S^1.$$

Tier IV. Applications and synthesis problems

1. (Heat equation on the circle.) Let $u_t = u_{\theta\theta}$ on S^1 with initial data $u(0,\theta) = f(\theta) \in L^2(S^1)$. Use Fourier series to construct a solution $u(t,\theta)$ and prove that $u(t,\cdot) \to f$ in $L^2(S^1)$ as $t \downarrow 0$. What additional regularity does $u(t,\cdot)$ gain for $t > 0$?

2. (A rigidity problem.) Suppose $f \in C^1(S^1)$ satisfies

$$(f * (1+f))(t) = f'(t) \qquad (t \in S^1).$$

   Prove that f is a trigonometric polynomial.

3. (Optional/advanced.) Let $f \in L^2(S^1)$ with $f' \in L^2(S^1)$ in the distributional sense. Show that

$$\|f\|_\infty \le |\widehat{f}(0)| + 2\sqrt{\sum_{n=1}^{\infty} \frac{1}{n^2}}\, \|f'\|_2.$$

Chapter 2

The Fourier Transform

On $\mathbb{R}^n$, translations form a non-compact abelian group, so the Fourier expansion of a non-periodic function is no longer discrete. The Fourier transform replaces the Fourier coefficients $\{\hat{f}(n)\}_{n\in\mathbb{Z}}$ by a continuous *frequency variable $\xi \in \mathbb{R}^n$. It linearizes convolution, converts differentiation into multiplication, and provides the natural L^2 isometry (Plancherel).*

Chapter roadmap. The chapter begins by identifying the characters of $\mathbb{R}^n$ and using them to motivate the definition of the Fourier transform. We then develop its basic functorial rules, decay properties, and interaction with convolution. Good kernels provide the bridge to inversion; Plancherel and interpolation place the transform in the L^p framework; and the final sections gather model computations, Poisson summation, and derivative criteria that tie the theory together.

Fourier analysis can be viewed as the study of functions through the symmetries of the ambient space. For Fourier series, periodicity reduces the problem to a single fundamental interval, and the function is encoded by a discrete sequence of Fourier coefficients. On $\mathbb{R}^n$, however, translations form a non-compact group, so one expects a continuous rather than a discrete frequency parameter. The relevant building blocks are again the eigenfunctions of translations.

Suppose that a nonzero function f satisfies

$$f(x+y) = \varphi(x)f(y), \qquad |\varphi(x)| = 1.$$

Setting $y = 0$ gives

$$f(x) = \varphi(x)f(0),$$

so f is completely determined by the scalar factor φ. Substituting this back into the relation yields

$$\varphi(x+y)f(0) = f(x+y) = \varphi(x)f(y) = \varphi(x)\varphi(y)f(0).$$

Since $f \not\equiv 0$, we obtain the multiplicative identity

$$\varphi(x+y) = \varphi(x)\varphi(y).$$

Thus the problem reduces to describing measurable unimodular characters of $\mathbb{R}^n$.

Theorem 2.1 (Characters of $\mathbb{R}^n$)**.** *Let $\varphi : \mathbb{R}^n \to \mathbb{C}$ be measurable and satisfy*

$$\varphi(x+y) = \varphi(x)\varphi(y), \qquad |\varphi(x)| = 1, \qquad x, y \in \mathbb{R}^n.$$

Then there exists $\xi \in \mathbb{R}^n$ such that

$$\varphi(x) = e^{2\pi i\, x\cdot\xi}, \qquad x \in \mathbb{R}^n.$$

Proof strategy. The substantive point is that measurability already forces regularity. Averaging φ over a short interval recovers φ from an absolutely continuous primitive, so the functional equation upgrades from a purely algebraic identity to a first-order differential equation. The n-dimensional statement is then obtained by restricting to the coordinate axes.

Proof. We begin with the case $n = 1$. Since $|\varphi| = 1$, the function φ belongs to $L^1_{\mathrm{loc}}(\mathbb{R})$. There must exist $a \in \mathbb{R}$ such that

$$\int_0^a \varphi(t)\, dt \neq 0.$$

Indeed, if this integral vanished for every a, then the primitive $A(x) := \int_0^x \varphi(t)\, dt$ would be identically zero, which would force $\varphi = 0$ a.e., contradicting $|\varphi| = 1$. Set

$$c^{-1} := \int_0^a \varphi(t)\, dt.$$

Using the multiplicative law $\varphi(x+t) = \varphi(x)\varphi(t)$, we obtain

$$\varphi(x) = c\int_0^a \varphi(x+t)\, dt = c\int_x^{x+a} \varphi(u)\, du.$$

The right-hand side is a difference of values of an absolutely continuous primitive of φ, so φ is continuous. Differentiating the displayed identity gives

$$\varphi'(x) = c\Big(\varphi(x+a) - \varphi(x)\Big).$$

Applying the functional equation once more,

$$\varphi(x+a) = \varphi(x)\varphi(a),$$

and hence

$$\varphi'(x) = c\Big(\varphi(a) - 1\Big)\varphi(x) =: B\,\varphi(x).$$

Thus φ satisfies the linear ODE $\varphi' = B\varphi$, so

$$\varphi(x) = \varphi(0)e^{Bx} = e^{Bx},$$

since $\varphi(0) = \varphi(0)^2$ and $|\varphi(0)| = 1$ imply $\varphi(0) = 1$. Finally, $|\varphi(x)| = 1$ for all x, so $\Re B = 0$; therefore $B = 2\pi i\xi$ for some $\xi \in \mathbb{R}$, and

$$\varphi(x) = e^{2\pi i\xi x}.$$

For general n, let $e_1, \ldots, e_n$ denote the standard basis of $\mathbb{R}^n$ and define

$$\varphi_j(t) := \varphi(te_j), \qquad t \in \mathbb{R}.$$

Each φ_j satisfies the one-dimensional hypotheses, so there exists $\xi_j \in \mathbb{R}$ such that

$$\varphi_j(t) = e^{2\pi i\xi_j t}.$$

If $x = \sum_{j=1}^n x_j e_j$, repeated use of the functional equation yields

$$\varphi(x) = \varphi\left(\sum_{j=1}^n x_j e_j\right) = \prod_{j=1}^n \varphi(x_j e_j) = \prod_{j=1}^n e^{2\pi i\xi_j x_j} = e^{2\pi i x\cdot\xi},$$

where $\xi = (\xi_1, \ldots, \xi_n)$. This proves the theorem. $\square$

Corollary 2.2. *If $\varphi : \mathbb{T} \to \mathbb{C}$ is measurable, satisfies $\varphi(x+y) = \varphi(x)\varphi(y)$, and has $|\varphi(x)| = 1$, then*

$$\varphi(x) = e^{2\pi i n x}$$

for some $n \in \mathbb{Z}$.

Proof. By the theorem, $\varphi(x) = e^{2\pi i \xi x}$ for some $\xi \in \mathbb{R}$. Since φ is defined on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, it must be 1-periodic. Hence

$$e^{2\pi i \xi} = \varphi(1) = \varphi(0) = 1,$$

so $\xi \in \mathbb{Z}$. □

Exercise 2.3. If $\varphi : \mathbb{T}^n \to \mathbb{C}$ is measurable, $|\varphi(x)| = 1$, and

$$\varphi(s+t) = \varphi(s)\varphi(t),$$

show that

$$\varphi(t) = e^{2\pi i\, t\cdot\alpha} \qquad \text{for some } \alpha \in \mathbb{Z}^n.$$

Therefore the relevant model eigenfunctions for translations are precisely the exponentials

$$x \longmapsto e^{2\pi i x\cdot\xi}, \qquad \xi \in \mathbb{R}^n$$

(and, on the torus, $\xi \in \mathbb{Z}^n$). These are the basic building blocks that will appear in the Fourier transform.

2.1 Definition of the Fourier transform

Definition 2.4. Let $f \in L^1(\mathbb{R}^n)$. We define its Fourier transform by

$$\widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i\, x\cdot\xi}\, dx.$$

Example 2.5 (A first transform)**.** For $f = \chi_{[0,1]}$ on $\mathbb{R}$, a direct computation gives

$$\widehat{f}(\xi) = \int_0^1 e^{-2\pi i x\xi}\, dx = e^{-\pi i\xi}\frac{\sin(\pi\xi)}{\pi\xi},$$

with the value at $\xi = 0$ understood by continuity. This example already displays two recurring themes: compact support in physical space creates oscillatory decay in frequency space, while the jump discontinuities of f prevent rapid decay.

Lemma 2.6. *Let $f \in L^1(\mathbb{R}^n)$. Then*

(i) $(\tau_y f)^\wedge(\xi) = e^{-2\pi i\xi\cdot y}\widehat{f}(\xi)$, *where* $\tau_y f(x) = f(x-y)$.

(ii) *If* $g(x) = e^{2\pi i\alpha\cdot x} f(x)$, *then* $\widehat{g}(\xi) = \widehat{f}(\xi - \alpha)$.

(iii) *If* $g(x) = \overline{f(-x)}$, *then* $\widehat{g}(\xi) = \overline{\widehat{f}(\xi)}$.

(iv) *If* $g(x) = f(\frac{x}{\lambda})$ *with* $\lambda \neq 0$, *then* $\widehat{g}(\xi) = |\lambda|^n \widehat{f}(\lambda\xi)$.

(v) $|\widehat{f}(\xi)| \leq \|f\|_1$ *(uniformly bounded).*

(vi) *If* $f, g \in L^1(\mathbb{R}^n)$, *then* $(f * g)^\wedge(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$.
(Hint: use Fubini's theorem and change of variable.)

Lemma 2.7. *Let* $f \in L^1(\mathbb{R}^n)$. *Then* $\widehat{f}$ *is uniformly continuous on* $\mathbb{R}^n$.

Proof. For $h \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^n$ we have

$$\widehat{f}(\xi + h) - \widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi} \left(e^{-2\pi i x\cdot h} - 1\right) dx,$$

so

$$|\widehat{f}(\xi + h) - \widehat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)|\, |e^{-2\pi i x\cdot h} - 1|\, dx.$$

Fix $\varepsilon > 0$. Choose $R > 0$ so that

$$\int_{|x|>R} |f(x)|\, dx < \frac{\varepsilon}{4}.$$

On the compact ball $\{x : |x| \leq R\}$, the function $x \mapsto e^{-2\pi i x\cdot h}$ converges uniformly to 1 as $h \to 0$. Hence there exists $\delta > 0$ such that

$$|h| < \delta \quad \Longrightarrow \quad \sup_{|x|\leq R} |e^{-2\pi i x\cdot h} - 1| < \frac{\varepsilon}{2\|f\|_1 + 1}.$$

For such h and every $\xi \in \mathbb{R}^n$,

$$\begin{aligned} |\widehat{f}(\xi + h) - \widehat{f}(\xi)| &\leq \int_{|x|\leq R} |f(x)|\, |e^{-2\pi i x\cdot h} - 1|\, dx + \int_{|x|>R} |f(x)|\, |e^{-2\pi i x\cdot h} - 1|\, dx \\ &\leq \|f\|_1 \frac{\varepsilon}{2\|f\|_1 + 1} + 2\int_{|x|>R} |f(x)|\, dx < \varepsilon. \end{aligned}$$

The estimate is uniform in ξ, so $\widehat{f}$ is uniformly continuous. □

Lemma 2.8. *Let* $f \in L^1(\mathbb{R})$ *and suppose that* f *is uniformly continuous on* $\mathbb{R}$. *Then*

$$\lim_{|x|\to\infty} f(x) = 0.$$

Proof. Argue by contradiction. If $f(x) \not\to 0$ as $|x| \to \infty$, then there exist $\varepsilon_0 > 0$ and a sequence $(x_k)_{k\geq 1}$ with $|x_k| \to \infty$ such that $|f(x_k)| \geq \varepsilon_0$ for all k. Uniform continuity yields $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon_0}{2}.$$

In particular, for every k and every $x \in (x_k - \delta, x_k + \delta)$ we have $|f(x)| \geq \varepsilon_0/2$. Passing to a subsequence, we may assume the intervals $(x_k - \delta, x_k + \delta)$ are pairwise disjoint. Therefore,

$$\|f\|_1 \;\geq\; \sum_{k=1}^{\infty} \int_{x_k-\delta}^{x_k+\delta} |f(x)|\, dx \;\geq\; \sum_{k=1}^{\infty} (2\delta)\, \frac{\varepsilon_0}{2} \;=\; \infty,$$

contradicting $f \in L^1(\mathbb{R})$. Hence $f(x) \to 0$ as $|x| \to \infty$. □

We use this fact to prove the following result.

Theorem 2.9. *Let $f \in L^1(\mathbb{R})$ and assume that $x\, f(x) \in L^1(\mathbb{R})$. Then $\widehat{f}$ is differentiable and*

$$\frac{d}{d\xi}\widehat{f}(\xi) = -2\pi i\, (\widehat{xf})(\xi).$$

Proof strategy. Differentiate the oscillatory factor rather than the function f itself. The hypothesis $xf \in L^1$ supplies exactly the integrable majorant needed to pass the limit through the integral.

Proof. For $h \neq 0$,

$$\frac{\widehat{f}(\xi + h) - \widehat{f}(\xi)}{h} = \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi}\, \frac{e^{-2\pi i x h} - 1}{h}\, dx.$$

By the mean value theorem applied to $t \mapsto e^{-2\pi i x t}$,

$$\left| \frac{e^{-2\pi i x h} - 1}{h} \right| \leq 2\pi |x|,$$

and, for each fixed x,

$$\frac{e^{-2\pi i x h} - 1}{h} \to -2\pi i x \qquad (h \to 0).$$

Since $2\pi |x|\, |f(x)| \in L^1(\mathbb{R})$, dominated convergence yields

$$\frac{d}{d\xi}\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} (-2\pi i x)\, dx = -2\pi i\, (\widehat{xf})(\xi).$$

□

Theorem 2.10. *Let $f \in L^1(\mathbb{R})$ and define*

$$F(x) := \int_{-\infty}^{x} f(y)\,dy.$$

If $F \in L^1(\mathbb{R})$, then for every $\xi \neq 0$,

$$\widehat{F}(\xi) = \frac{1}{2\pi i\,\xi}\,\widehat{f}(\xi).$$

Equivalently, if $f, f' \in L^1(\mathbb{R})$, then

$$\widehat{f'}(\xi) = 2\pi i\,\xi\,\widehat{f}(\xi).$$

Proof. Since

$$F(x) - F(y) = \int_{y}^{x} f(t)\,dt,$$

the function F is absolutely continuous on bounded intervals and uniformly continuous on $\mathbb{R}$. Because $F \in L^1(\mathbb{R})$, Lemma 2.8 implies

$$F(x) \to 0 \qquad (|x| \to \infty).$$

Now fix $\xi \neq 0$. Integration by parts gives

$$\widehat{F}(\xi) = \int_{\mathbb{R}} F(x) e^{-2\pi i x\xi}\,dx = \left[\frac{F(x)e^{-2\pi i x\xi}}{-2\pi i\xi}\right]_{-\infty}^{\infty} + \frac{1}{2\pi i\xi}\int_{\mathbb{R}} F'(x) e^{-2\pi i x\xi}\,dx.$$

The boundary term vanishes because $F(x) \to 0$ at both ends, and $F' = f$ almost everywhere by the fundamental theorem of calculus. Therefore

$$\widehat{F}(\xi) = \frac{1}{2\pi i\xi}\widehat{f}(\xi), \qquad \xi \neq 0.$$

Applying this identity with $F = f$ whenever $f, f' \in L^1(\mathbb{R})$ yields

$$\widehat{f'}(\xi) = 2\pi i\,\xi\,\widehat{f}(\xi).$$

□

Lemma 2.11. *The space $C_c^\infty(\mathbb{R})$ is dense in $L^1(\mathbb{R})$.*

Proof. Fix $f \in L^1(\mathbb{R})$ and $\varepsilon > 0$. Since $C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$, choose $g \in C_c(\mathbb{R})$ such that

$$\|f-g\|_1 < \varepsilon.$$

Let $\varphi \in C_c^\infty(\mathbb{R})$ satisfy $\int_{\mathbb{R}} \varphi = 1$, and for $t > 0$ set

$$\varphi_t(x) := t^{-1}\varphi(x/t).$$

Then $g * \varphi_t \in C_c^\infty(\mathbb{R})$ and

$$(g*\varphi_t)(x) - g(x) = \int_{\mathbb{R}} \Big(g(x-tz) - g(x)\Big)\varphi(z)\,dz. \tag{2.1}$$

By Minkowski's integral inequality,

$$\|g*\varphi_t - g\|_1 \le \int_{\mathbb{R}} \|\tau_{tz}g - g\|_1\,|\varphi(z)|\,dz.$$

Since translations are continuous in $L^1(\mathbb{R})$ and $g \in C_c(\mathbb{R})$, the integrand tends to 0 for each fixed z as $t \to 0$, and it is bounded by $2\|g\|_1|\varphi(z)|$, which is integrable. Dominated convergence therefore gives

$$\|g*\varphi_t - g\|_1 \to 0 \qquad (t\to 0).$$

For sufficiently small $t > 0$ we obtain

$$\|g*\varphi_t - f\|_1 \le \|g*\varphi_t - g\|_1 + \|g-f\|_1 < 2\varepsilon.$$

Thus $C_c^\infty(\mathbb{R})$ is dense in $L^1(\mathbb{R})$. □

Exercise 2.12. For $1 \le p < \infty$, show that

$$\overline{C_c^\infty(\mathbb{R})} = L^p(\mathbb{R}), \qquad \overline{C_c^\infty(\mathbb{R})} = C_0(\mathbb{R}).$$

(Hint: use Minkowski integral inequality in (2.1)*.)*

2.2 Riemann-Lebesgue Lemma

Theorem 2.13. *If $f \in L^1(\mathbb{R})$, then*

$$\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0.$$

Proof strategy. Approximate f in L^1 by a smooth compactly supported function. For the smooth approximation, one integration by parts produces explicit $|\xi|^{-1}$ decay; the L^1 approximation error controls the difference of the Fourier transforms uniformly in ξ.

Proof. Fix $\varepsilon > 0$. By the preceding lemma, choose $g \in C_c^\infty(\mathbb{R})$ such that

$$\|f - g\|_1 < \varepsilon.$$

Since $g' \in L^1(\mathbb{R})$, Theorem 2.10 gives

$$2\pi i \, \xi \, \widehat{g}(\xi) = \widehat{g'}(\xi),$$

and therefore

$$|\widehat{g}(\xi)| \le \frac{\|g'\|_1}{2\pi|\xi|} \to 0 \qquad (|\xi| \to \infty).$$

Also,

$$|\widehat{f}(\xi) - \widehat{g}(\xi)| \le \|f - g\|_1 < \varepsilon \qquad \text{for all } \xi \in \mathbb{R}.$$

Hence

$$\limsup_{|\xi|\to\infty} |\widehat{f}(\xi)| \le \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$. □

Remark 2.14 (The range of $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is proper). Every Fourier transform of an L^1-function belongs to $C_0(\mathbb{R})$ by the preceding theorem, but the inclusion is strict. Indeed, suppose $g = \widehat{f}$ for some odd function g and some $f \in L^1(\mathbb{R})$. Then

$$g(x) = -i \int_{\mathbb{R}} f(t) \sin(2\pi t x) \, dt.$$

Consequently, for $1 < A < B < \infty$,

$$\left| \int_A^B \frac{g(x)}{x} \, dx \right| \le \int_{\mathbb{R}} |f(t)| \left| \int_A^B \frac{\sin(2\pi t x)}{x} \, dx \right| dt.$$

After the change of variables $u = 2\pi t x$, the inner integral becomes

$$\int_{2\pi t A}^{2\pi t B} \frac{\sin u}{u} \, du,$$

which is bounded uniformly in t, A, B because the sine integral is bounded on $\mathbb{R}$. Hence

$$\sup_{1<A<B<\infty} \left| \int_A^B \frac{g(x)}{x}\, dx \right| < \infty.$$

Now consider

$$g_0(x) := \frac{x}{(1+|x|)\log(e+|x|)}.$$

Then $g_0 \in C_0(\mathbb{R})$ and g_0 is odd, but

$$\int_2^B \frac{g_0(x)}{x}\, dx \asymp \int_2^B \frac{dx}{x \log x} \to \infty \qquad (B \to \infty).$$

Therefore g_0 cannot be the Fourier transform of any L^1-function.

Example 2.15 (The Gaussian is a fixed point of $\mathcal{F}$)**.** Let $f(x) = e^{-\pi x^2}$ and set $F(\xi) := \widehat{f}(\xi)$. Then $F(\xi) = e^{-\pi\xi^2}$ for all $\xi \in \mathbb{R}$.

Indeed, $f, xf \in L^1(\mathbb{R})$ and differentiation under the integral sign gives

$$F'(\xi) = \int_{\mathbb{R}} (-2\pi i x)\, f(x)\, e^{-2\pi i x\xi}\, dx = -2\pi i\, \widehat{(xf)}(\xi).$$

Since $f'(x) = -2\pi x f(x)$, we have $xf = -(2\pi)^{-1} f'$, hence

$$F'(\xi) = i\, \widehat{f'}(\xi).$$

Using the differentiation rule $\widehat{f'}(\xi) = 2\pi i\, \xi\, \widehat{f}(\xi)$ we obtain the ODE

$$F'(\xi) = -2\pi\xi\, F(\xi).$$

Finally $F(0) = \int_{\mathbb{R}} e^{-\pi x^2} dx = 1$, so $F(\xi) = e^{-\pi\xi^2}$.

Remark 2.16. For $\delta > 0$, let $f_\delta(x) = \delta^{1/2} e^{-\pi x^2/\delta}$. Then $\widehat{f_\delta}(x) = e^{-\pi\delta x^2} \to 0$ as $\delta \to 0$, however, $f_\delta(x) \to 1$ as $\delta \to 0$. Hence, we cannot see both f_δ & $\hat{f}_\delta$ exist together. That is, f_δ and $\hat{f}_\delta$ cannot be localized together. (This is known as the Heisenberg uncertainty principle; we elaborate later.)

Example 2.17. If $f(x) = e^{-\pi x^2}$ then show that $|f(x)| \leq \frac{M}{1+x^2}$

Lemma 2.18. *Let $f, h, H \in L^1(\mathbb{R})$ and assume that*

$$h(x) = \int_{\mathbb{R}} H(\xi) e^{2\pi i x\xi}\, d\xi \qquad (x \in \mathbb{R}).$$

Then

$$(h * f)(x) = \int_{\mathbb{R}} H(\xi)\widehat{f}(\xi)e^{2\pi i x\xi}\, d\xi.$$

Proof. Since $H \in L^1(\mathbb{R})$ and $|\widehat{f}(\xi)| \le \|f\|_1$, the right-hand side is absolutely integrable. Using the assumed representation of h and Fubini's theorem, we compute

$$\begin{aligned}(h * f)(x) &= \int_{\mathbb{R}} h(x-y)f(y)\, dy \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} H(\xi)e^{2\pi i(x-y)\xi} f(y)\, d\xi\, dy \\ &= \int_{\mathbb{R}} H(\xi)e^{2\pi i x\xi}\left(\int_{\mathbb{R}} f(y)e^{-2\pi i y\xi}\, dy\right) d\xi \\ &= \int_{\mathbb{R}} H(\xi)\widehat{f}(\xi)e^{2\pi i x\xi}\, d\xi.\end{aligned}$$

□

2.3 The Schwartz space and rapid decay

For many structural results (Fourier inversion, Poisson summation, and later the passage to tempered distributions) it is convenient to work first with a class of functions that is simultaneously: *smooth*, *closed under differentiation*, and *rapidly decaying* together with all derivatives. The standard choice is the Schwartz space.

Notation 2.19 (Multi-index notation)*.* For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ we write $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$.

Definition 2.20 (Schwartz space)**.** The *Schwartz space* $\mathcal{S}(\mathbb{R}^n)$ consists of all C^∞ functions $f : \mathbb{R}^n \to \mathbb{C}$ such that for every pair of multi-indices $\alpha, \beta \in \mathbb{N}_0^n$,

$$\sup_{x\in\mathbb{R}^n} |x^\alpha\, \partial^\beta f(x)| < \infty.$$

Equivalently, $\mathcal{S}(\mathbb{R}^n)$ is the Fréchet space whose topology is generated by the seminorms $p_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^n} |x^\alpha\, \partial^\beta f(x)|$.

Lemma 2.21. *If $f \in \mathcal{S}(\mathbb{R}^n)$, then $f \in L^p(\mathbb{R}^n)$ for every $1 \le p \le \infty$. Moreover, $\partial^\beta f \in L^p(\mathbb{R}^n)$ for every multi-index β and every $1 \le p \le \infty$.*

Proof. Fix β and choose $N > n/p$ (with the convention $n/\infty = 0$). Since $f \in \mathcal{S}(\mathbb{R}^n)$, we have $|\,\partial^\beta f(x)\,| \lesssim (1 + |x|)^{-N}$, and $(1 + |x|)^{-N} \in L^p(\mathbb{R}^n)$ for such N. The L^∞ case is immediate from the defining seminorms. □

Proposition 2.22 (Fourier transform preserves $\mathcal{S}$)**.** *If $f \in \mathcal{S}(\mathbb{R}^n)$, then $\widehat{f} \in \mathcal{S}(\mathbb{R}^n)$. More precisely, for all multi-indices α, β,*

$$\widehat{\partial^\beta f}(\xi) = (2\pi i\,\xi)^\beta\,\widehat{f}(\xi), \qquad \widehat{x^\alpha f}(\xi) = \left(\frac{i}{2\pi}\right)^{|\alpha|} \partial^\alpha \widehat{f}(\xi). \tag{2.2}$$

Proof. For $f \in \mathcal{S}$, all integrals below are absolutely convergent and differentiation under the integral sign is justified by Lemma 2.21 and dominated convergence. The first identity in (2.2) follows by integrating by parts in each variable. For the second identity, differentiate $\widehat{f}(\xi) = \int f(x)e^{-2\pi i x\cdot\xi}\,dx$ with respect to ξ and use that $\partial_{\xi_j} e^{-2\pi i x\cdot\xi} = -2\pi i x_j e^{-2\pi i x\cdot\xi}$.

To see that $\widehat{f}$ is Schwartz, fix α, β. Using (2.2) and the uniform bound $|\widehat{g}(\xi)| \le \|g\|_1$ valid for $g \in L^1$, we obtain

$$\sup_{\xi\in\mathbb{R}^n} |\xi^\alpha\,\partial^\beta \widehat{f}(\xi)| \;\lesssim\; \sum_{|\gamma|\le|\alpha|} \left\| x^\gamma\,\partial^\beta f \right\|_1 \;<\; \infty,$$

since $x^\gamma \partial^\beta f \in \mathcal{S} \subset L^1$ by Lemma 2.21. □

Remark 2.23. We have $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$. In particular, whenever a statement is first proved for Schwartz functions, it can often be extended to larger function spaces by density arguments (for example, C_c^∞ is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$).

2.4 Good Kernels on $\mathbb{R}$

In the non-periodic setting, the role of averaging kernels is played by *approximate identities* (also called *summability kernels*).

Definition 2.24 (Good kernels / approximate identities)**.** A family $\{K_\lambda\}_{\lambda>0} \subset L^1(\mathbb{R})$ is called a family of *good kernels* if

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

(ii) $\sup_{\lambda>0} \|K_\lambda\|_{L^1(\mathbb{R})} < \infty$;

(iii) for every $\delta > 0$, $\int_{|x|>\delta} |K_\lambda(x)|\,dx \to 0$ as $\lambda \to \infty$.

Example 2.25 (Standard construction)**.** Let $K \in L^1(\mathbb{R})$ satisfy $\int_{\mathbb{R}} K(x)\,dx = 1$, and define

$$K_\lambda(x) := \lambda\,K(\lambda x), \qquad \lambda > 0.$$

Then $\{K_\lambda\}_{\lambda>0}$ is a family of good kernels.

Theorem 2.26 (Approximation by good kernels)**.** *Let* $\{K_\lambda\}$ *be a family of good kernels and let* $1 \le p < \infty$*. Then for every* $f \in L^p(\mathbb{R})$*,*

$$\|f * K_\lambda - f\|_{L^p(\mathbb{R})} \longrightarrow 0 \qquad (\lambda \to \infty).$$

If, in addition, f *is bounded and uniformly continuous, then* $f * K_\lambda(x) \to f(x)$ *uniformly in* $x \in \mathbb{R}$*.*

Proof. We prove the L^p-convergence. Fix $\varepsilon > 0$. By the continuity of translations in $L^p(\mathbb{R})$, there exists $\delta > 0$ such that

$$\|\tau_y f - f\|_{L^p(\mathbb{R})} < \varepsilon \qquad \text{whenever } |y| < \delta, \tag{2.3}$$

where $\tau_y f(x) := f(x - y)$. Using Minkowski's integral inequality and (2.3),

$$\begin{aligned} \|f * K_\lambda - f\|_p &= \left\| \int_{\mathbb{R}} K_\lambda(y) \left(\tau_y f - f\right) dy \right\|_p \\ &\le \int_{|y|<\delta} |K_\lambda(y)| \, \|\tau_y f - f\|_p \, dy + \int_{|y|\ge\delta} |K_\lambda(y)| \, \|\tau_y f - f\|_p \, dy \\ &\le \varepsilon \|K_\lambda\|_1 + 2\|f\|_p \int_{|y|\ge\delta} |K_\lambda(y)| \, dy. \end{aligned}$$

By Definition 2.24(ii)–(iii), the first term is $\le \varepsilon \sup_\lambda \|K_\lambda\|_1$, while the second term tends to 0 as $\lambda \to \infty$. This proves the claim. □

2.5 The Fejér Kernel on $\mathbb{R}$

A particularly important approximate identity is obtained by taking a compactly supported multiplier in the frequency domain. For $\lambda > 0$, define the triangular cutoff

$$G_\lambda(\xi) := \chi_{[-\lambda,\lambda]}(\xi) \left(1 - \frac{|\xi|}{\lambda}\right),$$

and set

$$K_\lambda(x) := \int_{\mathbb{R}} G_\lambda(\xi) \, e^{2\pi i x \xi} \, d\xi. \tag{2.4}$$

Then $K_\lambda \in L^1(\mathbb{R})$, $\int_{\mathbb{R}} K_\lambda(x) \, dx = 1$, and $\{K_\lambda\}$ is a family of good kernels. Moreover, $\widehat{K_\lambda} = G_\lambda$.

In fact, writing $G(\xi) = (1 - |\xi|)_+$, one has the explicit formula

$$K(x) = \left(\frac{\sin(\pi x)}{\pi x}\right)^2, \qquad K_\lambda(x) = \lambda\, K(\lambda x).$$

2.6 Fourier uniqueness theorem

Theorem 2.27 (Uniqueness)**.** *If $f \in L^1(\mathbb{R})$ and $\widehat{f}(\xi) = 0$ for all $\xi \in \mathbb{R}$, then $f = 0$ almost everywhere.*

Proof. For each $\lambda > 0$, we have $\widehat{f * K_\lambda} = \widehat{f}\,\widehat{K_\lambda} = \widehat{f}\,G_\lambda \equiv 0$. Hence $f * K_\lambda \equiv 0$. Since $\{K_\lambda\}$ is a family of good kernels, Theorem 2.26 gives $f * K_\lambda \to f$ in $L^1(\mathbb{R})$, and therefore $f = 0$ a.e. □

2.7 Fourier Inversion

Proof strategy. The kernels K_λ simultaneously localize frequency and approximate the identity in physical space. The inversion formula emerges by computing the same quantity $f * K_\lambda$ in these two complementary ways and then letting $\lambda \to \infty$.

Theorem 2.28 (Inversion)**.** *Let $f \in L^1(\mathbb{R})$ and assume $\widehat{f} \in L^1(\mathbb{R})$. Then for almost every $x \in \mathbb{R}$,*

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

Proof. Using (2.4) and the convolution theorem,

$$f * K_\lambda(x) = \int_{\mathbb{R}} \widehat{f}(\xi)\, G_\lambda(\xi)\, e^{2\pi i x\xi}\, d\xi.$$

Since $0 \le G_\lambda \le 1$ and $G_\lambda(\xi) \to 1$ pointwise as $\lambda \to \infty$, dominated convergence (with $\widehat{f} \in L^1$) yields, for every fixed x,

$$\lim_{\lambda\to\infty} f * K_\lambda(x) = \int_{\mathbb{R}} \widehat{f}(\xi)\, e^{2\pi i x\xi}\, d\xi.$$

On the other hand, Theorem 2.26 implies $f * K_\lambda \to f$ in $L^1(\mathbb{R})$, hence along a subsequence also pointwise a.e. Therefore,

$$f(x) = \lim_{\lambda\to\infty} \int_{\mathbb{R}} \widehat{f}(\xi)\, G_\lambda(\xi)\, e^{2\pi i x\xi}\, d\xi$$

for almost every x, and the right-hand side equals $\int_{\mathbb{R}} \widehat{f}(\xi)e^{2\pi ix\xi}\,d\xi$ by dominated convergence.

□

Remark 2.29. The conclusion is stated almost everywhere because an L^1 function is only defined up to null sets, and approximate identities recover such a function at Lebesgue points in general. If f is continuous and $\widehat{f} \in L^1$, then the inversion formula holds pointwise for every x.

2.8 Plancherel theorem

The Fourier transform maps $L^1(\mathbb{R})$ into the space of bounded uniformly continuous functions. On $L^2(\mathbb{R})$, however, it admits a much stronger extension: it becomes an isometry, and after the inverse transform is identified on a dense class one sees that it is in fact unitary. More precisely, for $f \in L^2(\mathbb{R})$ one has

$$\|\widehat{f}\|_2 = \|f\|_2.$$

We first establish this on the dense subspace $L^1 \cap L^2$, and then extend by continuity. Later, using the Riesz–Thorin interpolation theorem, we shall obtain the Hausdorff–Young inequality for $1 \leq p \leq 2$ and see how the Fourier transform extends further in the distributional setting.

Theorem 2.30. *There exists a unique operator* $\mathscr{F}\colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$ *having the following properties:*

$$\mathscr{F}f = \widehat{f} \text{ for } f \in L^1 \cap L^2(\mathbb{R}),$$

$$\|\mathscr{F}f\|_2 = \|f\|_2.$$

Proof strategy. On the dense class $L^1 \cap L^2$, the key identity is that the transform of $f * \widetilde{f}$ equals $|\widehat{f}|^2$. Evaluating this relation at the origin converts convolution into the L^2 inner product and yields the isometry from which the extension to all of L^2 follows.

Proof. For $f \in L^1 \cap L^2(\mathbb{R})$, set

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi ix\xi}\,dx.$$

Let $\widetilde{f}(x) := \overline{f(-x)}$ and define $g := f * \widetilde{f}$. Then $g \in L^1(\mathbb{R})$ and

$$\widehat{g}(\xi) = \widehat{f}(\xi)\,\overline{\widehat{f}(\xi)} = |\widehat{f}(\xi)|^2.$$

Moreover,

$$g(0) = \int_{\mathbb{R}} f(y)\overline{f(y)}\,dy = \|f\|_2^2.$$

Let K_λ denote the Fejér kernel on $\mathbb{R}$, with Fourier transform

$$G_\lambda(\xi) = \left(1 - \frac{|\xi|}{\lambda}\right)_+.$$

Since $g \in L^1(\mathbb{R})$ and $\widehat{g} = |\widehat{f}|^2 \geq 0$, the inversion formula applied to $g * K_\lambda$ at the origin gives

$$(g * K_\lambda)(0) = \int_{\mathbb{R}} G_\lambda(\xi)\,|\widehat{f}(\xi)|^2\,d\xi.$$

Because $g \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ and $\{K_\lambda\}$ is a family of good kernels,

$$(g * K_\lambda)(0) \to g(0) = \|f\|_2^2.$$

On the other hand, $0 \leq G_\lambda \uparrow 1$ pointwise, so the monotone convergence theorem yields

$$\int_{\mathbb{R}} |\widehat{f}(\xi)|^2\,d\xi = \|f\|_2^2.$$

Thus the Fourier transform is an isometry on $L^1 \cap L^2(\mathbb{R})$.

Now let $f \in L^2(\mathbb{R})$. Choose a sequence $f_n \in L^1 \cap L^2(\mathbb{R})$ such that $f_n \to f$ in L^2. The isometry just proved implies

$$\|\widehat{f}_n - \widehat{f}_m\|_2 = \|f_n - f_m\|_2,$$

so $(\widehat{f}_n)$ is Cauchy in $L^2(\mathbb{R})$. Define

$$\mathscr{F} f := \lim_{n\to\infty} \widehat{f}_n \qquad \text{in } L^2(\mathbb{R}).$$

This definition is independent of the approximating sequence, extends the classical Fourier transform on $L^1 \cap L^2$, and satisfies

$$\|\mathscr{F} f\|_2 = \|f\|_2.$$

Uniqueness follows from the density of $L^1 \cap L^2$ in L^2. □

Remark 2.31. Once the inverse transform is identified on a dense class (for example on Schwartz functions), the same construction applied to that inverse shows that $\mathscr{F}$ is surjective as well; equivalently, $\mathscr{F}$ is unitary on $L^2(\mathbb{R})$.

Remark 2.32. If $f \in L^2(\mathbb{R})$, then $\chi_{[-n,n]} f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Writing

$$\varphi_n := \chi_{[-n,n]} f,$$

we obtain

$$\|\widehat{\varphi_n} - \mathscr{F} f\|_2 = \|\varphi_n - f\|_2 \to 0.$$

Thus the L^2-Fourier transform may be recovered as an L^2-limit of the classical truncated Fourier integrals.

Example 2.33. For $H(x) = e^{-|x|}$ one computes directly that

$$\widehat{H}(\xi) = \int_{\mathbb{R}} e^{-|t|} e^{-2\pi i t \xi}\, dt = \frac{2}{1 + 4\pi^2 \xi^2}.$$

In particular, the Plancherel identity extends the familiar polarization formula:

$$\int_{\mathbb{R}} f(x)\overline{g(x)}\, dx = \int_{\mathbb{R}} \mathscr{F} f(\xi)\, \overline{\mathscr{F} g(\xi)}\, d\xi, \qquad f, g \in L^2(\mathbb{R}).$$

2.9 A model computation: the Gaussian and the heat kernel

A single explicit computation already illustrates several recurring themes: rapid decay in physical space, rapid decay in frequency space, and the way Fourier analysis diagonalizes constant-coefficient PDE.

Proposition 2.34 (Fourier transform of a Gaussian)**.** *For $a > 0$ define $g_a(x) = e^{-\pi a |x|^2}$ on $\mathbb{R}^n$. Then $g_a \in \mathcal{S}(\mathbb{R}^n)$ and*

$$\widehat{g_a}(\xi) = a^{-n/2}\, e^{-\pi |\xi|^2 / a} \qquad (\xi \in \mathbb{R}^n).$$

In particular, g_1 is an eigenfunction of the Fourier transform: $\widehat{g_1} = g_1$.

Proof. By Fubini and the product structure, it suffices to treat the one-dimensional integral

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

Complete the square:

$$-\pi a x^2 - 2\pi i x\xi = -\pi a\Big(x + \frac{i\xi}{a}\Big)^2 - \pi\frac{\xi^2}{a}.$$

Shifting the contour in the complex plane (justified since the integrand is entire and rapidly decaying in horizontal strips), or equivalently differentiating I_a with respect to ξ and solving the resulting ODE, gives $I_a(\xi) = a^{-1/2}e^{-\pi\xi^2/a}$. Taking products over the n coordinates yields the stated formula. □

Corollary 2.35 (Heat kernel on $\mathbb{R}^n$)**.** *Let $u_0 \in \mathcal{S}(\mathbb{R}^n)$ and consider the Cauchy problem*

$$\partial_t u = \Delta u, \qquad u(\,\cdot\,,0) = u_0.$$

Then the unique solution in $\mathcal{S}(\mathbb{R}^n)$ is

$$u(\,\cdot\,,t) = p_t * u_0, \qquad p_t(x) = (4\pi t)^{-n/2}\, e^{-|x|^2/(4t)} \qquad (t>0).$$

Equivalently, in frequency space

$$\widehat{u}(\xi,t) = e^{-4\pi^2 t|\xi|^2}\,\widehat{u_0}(\xi).$$

Proof. Taking Fourier transforms in x and using $\widehat{\partial_j^2 u}(\xi,t) = -(2\pi\xi_j)^2\widehat{u}(\xi,t)$ gives the ODE

$$\partial_t\widehat{u}(\xi,t) = -4\pi^2|\xi|^2\,\widehat{u}(\xi,t),$$

whose solution is $\widehat{u}(\xi,t) = e^{-4\pi^2 t|\xi|^2}\widehat{u_0}(\xi)$. By Proposition 2.34, the inverse transform of $e^{-4\pi^2 t|\xi|^2}$ is precisely p_t, and the inversion formula yields $u(\cdot,t) = p_t * u_0$. □

Remark 2.36. The kernel $(p_t)_{t>0}$ forms an approximate identity as $t \downarrow 0$ and satisfies $\int_{\mathbb{R}^n} p_t = 1$. In particular, $u(\cdot,t) \to u_0$ in $L^p(\mathbb{R}^n)$ for $1 \le p \le \infty$ whenever $u_0 \in L^p$ and the convolution is defined.

2.10 More on Convolution

Theorem 2.37. *Let $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$, and suppose $\frac{1}{p} + \frac{1}{q} = 1$. Then $f * g$ is a bounded, uniformly continuous function on $\mathbb{R}$ and*

$$\|f * g\|_\infty \le \|f\|_p\|g\|_q.$$

In particular, if $1 < p < \infty$*, then* $f * g \in C_0(\mathbb{R})$*.*

Proof. Hölder's inequality gives, for every $x \in \mathbb{R}$,

$$|f * g(x)| \le \int_{\mathbb{R}} |f(x-y)|\,|g(y)|\,dy \le \|\tau_x f\|_p \, \|g\|_q = \|f\|_p \, \|g\|_q,$$

so $f * g$ is bounded. Moreover,

$$\begin{aligned} |(f*g)(y+x) - (f*g)(y)| &= \left| \int_{\mathbb{R}} \Big(f(y+x-\xi) - f(y-\xi) \Big) g(\xi)\, d\xi \right| \\ &\le \|\tau_x f - f\|_p \, \|g\|_q. \end{aligned}$$

Taking the supremum in y yields

$$\|\tau_x(f*g) - (f*g)\|_\infty \le \|\tau_x f - f\|_p \, \|g\|_q.$$

Since translations are continuous on $L^p(\mathbb{R})$, the right-hand side tends to 0 as $x \to 0$; hence $f * g$ is uniformly continuous.

Assume now that $1 < p < \infty$, so also $1 < q < \infty$. Choose sequences $f_n \in C_c^\infty(\mathbb{R})$ and $g_n \in C_c^\infty(\mathbb{R})$ such that

$$\|f_n - f\|_p \to 0, \qquad \|g_n - g\|_q \to 0.$$

Each $f_n * g_n$ belongs to $C_c^\infty(\mathbb{R}) \subset C_0(\mathbb{R})$. Furthermore,

$$\begin{aligned} \|f_n * g_n - f * g\|_\infty &\le \|(f_n - f) * g_n\|_\infty + \|f * (g_n - g)\|_\infty \\ &\le \|f_n - f\|_p \, \|g_n\|_q + \|f\|_p \, \|g_n - g\|_q. \end{aligned}$$

Since (g_n) is bounded in L^q, the right-hand side tends to 0. Thus $f_n * g_n \to f * g$ uniformly. Because $C_0(\mathbb{R})$ is closed in $L^\infty(\mathbb{R})$, it follows that $f * g \in C_0(\mathbb{R})$. □

2.11 Riesz-Thorin Interpolation Theorem

Theorem 2.38. *Let* (X, S, μ) *and* (Y, T, ν) *be two* σ*-finite measure spaces. Let* $p_i, q_i \in [1, \infty]$*,* $i = 0, 1$ *and define*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \quad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}$$

where $0 \leq t \leq 1$*. If* T *is a linear map from*

$$L^{p_0}(\mu) + L^{p_1}(\mu) \to L^{q_0}(\nu) + L^{q_1}(\nu)$$

such that

$$\|Tf\|_{q_i} \leq M_i \|f\|_{p_i}, \quad i = 0, 1,$$

then

$$\|Tf\|_{q_t} \leq M_0^{1-t} M_1^t \|f\|_{p_t}$$

(For a proof, see Real Analysis *by G.B. Folland.)*

Using R-T theorem we see that the Fourier transform of a function $f \in L^p(\mathbb{R}),\ 1 \leq p \leq 2$, exists as a function in L^q, $\frac{1}{p} + \frac{1}{q} = 1$.

2.12 Hausdorff-Young Inequality

Theorem 2.39. *Let* $1 \leq p \leq 2$*. Then for* $f \in L^p(\mathbb{R})$*,* $\hat{f} \in L^q(\mathbb{R})$*, with* $\|\hat{f}\|_q \leq \|f\|_p$*, where* $\frac{1}{p} + \frac{1}{q} = 1$*.*

Note that if $1 \leq p < 2$*, then* $q \in [2, \infty]$*.*

Similarly, if $f \in L^p(S^1)$*,* $1 \leq p \leq 2$*, then* $\hat{f} \in l^q(\mathbb{Z})$*, with* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $\|\hat{f}\|_q \leq \|f\|_p$*.*

Proof. We know that $\mathcal{F} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ satisfies

$$\|\mathcal{F}(f)\|_\infty \leq \|f\|_1$$

and $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ with $\|\mathcal{F}(f)\|_2 = \|f\|_2$.

Let

$$\frac{1}{p_t} = \frac{1-t}{1} + \frac{t}{2}, \quad \frac{1}{q_t} = \frac{1-t}{\infty} + \frac{t}{2}$$

Note that

$$\frac{1}{p_t} + \frac{1}{q_t} = 1, \frac{1}{p} + \frac{1}{q} = 1.$$

so we can choose $t \in (0, 1)$ such that $\frac{1}{q} = \frac{t}{2}$ and $\frac{1}{p} = \frac{1-t}{1} + \frac{t}{2}$. Hence by R-T inequality, we get

$$\|\mathcal{F}(f)\|_q \leq \|f\|_p$$

Thus, the Fourier transform is a bounded linear function from L^p to L^q. □

2.13 Young's Inequality

Theorem 2.40. *Let* $1 \le p, q, r \le \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$*. If* $f \in L^p$ *and* $g \in L^q$*, then* $f * g \in L^r$ *and*

$$\|f * g\|_r \le \|f\|_p \|g\|_q$$

Proof. We prove the endpoint estimates first.

Case 1: $p = 1$ *and* $r = q$*.* By Minkowski's integral inequality,

$$\|f * g\|_q \le \|f\|_1 \, \|g\|_q.$$

Case 2: $r = \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$*.* This is exactly the previous boundedness theorem:

$$\|f * g\|_\infty \le \|f\|_p \, \|g\|_q.$$

For the general case, fix $g \in L^q$ and consider the linear operator

$$T_g(f) := f * g.$$

By Case 1 and Case 2,

$$T_g : L^1 \to L^q \quad \text{with norm at most } \|g\|_q,$$

and, writing q' for the conjugate exponent of q,

$$T_g : L^{q'} \to L^\infty \quad \text{with norm at most } \|g\|_q.$$

Apply the Riesz–Thorin interpolation theorem with

$$p_0 = 1, \quad q_0 = q, \qquad p_1 = q', \quad q_1 = \infty.$$

Then for $0 \le t \le 1$,

$$\frac{1}{p_t} = \frac{1-t}{1} + \frac{t}{q'}, \qquad \frac{1}{r_t} = \frac{1-t}{q}.$$

Eliminating t gives precisely

$$\frac{1}{p_t} + \frac{1}{q} = 1 + \frac{1}{r_t}.$$

Therefore, for every triple (p, q, r) satisfying

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

we obtain

$$\|f * g\|_r \leq \|f\|_p \, \|g\|_q.$$

This is Young's inequality. □

Notice that, by the Hausdorff-Young inequality, if $1 \leq p \leq 2$, then for $f \in L^p(\mathbb{R})$, $\hat{f} \in L^q(\mathbb{R})$ where $\frac{1}{p} + \frac{1}{q} = 1$. Hence by continuity we can define

$$\widehat{f}(\xi) \stackrel{L^2}{:=} \lim_{n\to\infty} \int_{-n}^{n} e^{-2\pi i x \xi} f(x)\, dx.$$

However, if $1 < p < 2$, we do not know how the $\hat{f}$ looks like. For example, if $f \in L^1(\mathbb{R})$, then

$$\lim_{\lambda\to\infty} \|f * K_\lambda - f\|_1 = 0$$

and

$$f(x) = \lim_{\lambda\to\infty} \int_{\mathbb{R}} G_\lambda(\xi) \widehat{f}(\xi) e^{2\pi i x \xi} d\xi$$

holds in $L^1(\mathbb{R})$.

For $1 < p < 2$, we can generalize (*). For this, we need to verify the following: If $f \in L^1(\mathbb{R})$ and $g \in L^p(\mathbb{R}), 1 < p < 2$, then $f * g \in L^p$ and $(f * g)^\wedge = \hat{f}\hat{g}$. Since $C_0^\infty(\mathbb{R})$ is dense in $L^p(\mathbb{R})$, for $\epsilon > 0$, there exists $g_n \in C_0^\infty(\mathbb{R})$ so that $\|g - g_n\|_{L^p} < \epsilon$.

Note that $\widehat{g_n} \in L^1(\mathbb{R})$ (since second derivative of g satisfies $\hat{g^2}_n(x) = (ix)^2 \hat{g}_n(x)$) and

$$\mathcal{F}(g_n * f) = \mathcal{F}(g_n)\mathcal{F}(f).$$

As $\mathcal{F} : L^p \to L^q$, is a continuous linear map, from (**) it follows that

$$\mathcal{F}(g * f) = \mathcal{F}(g)\mathcal{F}(f).$$

Now, consider $f = K_\lambda$ (Fejér kernel on $\mathbb{R}$), then

$$(K_\lambda * g)^\wedge = \hat{K}_\lambda \hat{g} = G_\lambda \hat{g},$$

where

$$G_\lambda(\xi) = (1 - |\xi|/\lambda)\chi_{[-\lambda,\lambda]}(\xi)$$

Since $\hat{g} \in L^q(\mathbb{R})$, $q > 2$, it is easy to see that $G_\lambda \hat{g} \in L^2(\mathbb{R})$. By inversion formula,

$$K_\lambda * g(x) = \int G_\lambda(\xi)\, \widehat{g}(\xi)\, e^{2\pi i x\xi}\, d\xi,$$

and $K_\lambda * g \in L^2(\mathbb{R})$. Since K_λ is a good kernel and $K_\lambda * g \to g$ in $L^p(\mathbb{R})$, we can write the following result:

Theorem 2.41. *Let $1 \leq p \leq 2$ and $g \in L^p(\mathbb{R})$. Then*

$$g(x) = \lim_{\lambda\to\infty} \int_{\mathbb{R}} G_\lambda(\xi)\widehat{g}(\xi)e^{2\pi i x\xi} d\xi$$

in $L^p(\mathbb{R})$.

Corollary 2.42. $\{f \in L^p, 1 \leq p \leq 2 \ \operatorname{supp} \hat{f}$ *is compact* $\}$*, is dense in $L^p(\mathbb{R})$.*

Notice that, if $f, g \in L^1(\mathbb{R})$, then $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$ where $\mathcal{F}$ is the Fourier transform.

Question 2.43. Does F is unique that satisfies $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$?

Note that if we write

$$\mathcal{F}(f) = \int f(x)e^{-2\pi i t_0 x} dx = \widehat{f}(t_0),$$

then $\mathcal{F}$ is a continuous linear functional on $L^1(\mathbb{R})$. We then shall see that such any continuous linear functional is only the Fourier transform

2.14 Riesz Theorem

Theorem 2.44. *Let $1 \leq p < \infty$ and (X, S, μ) be a σ-finite measure space. Then for every continuous linear functional T on $L^p(\mu)$, there exists a unique $g \in L^q(X)$, where $1/p + 1/q = 1$, such that*

$$Tf = \int fg$$

Fourier transform is unique. Suppose φ is a continuous linear functional on $L^1(\mathbb{R})$ with $\|\varphi\| \leq 1$ and

$$\varphi(f * g) = \varphi(f)\varphi(g), \qquad f, g \in L^1(\mathbb{R}).$$

By the Riesz representation theorem, there exists $\beta \in L^\infty(\mathbb{R})$ such that

$$\varphi(f) = \int_{\mathbb{R}} f(x)\beta(x)\, dx.$$

Hence, using Fubini's theorem,

$$\varphi(f * g) = \int_{\mathbb{R}} g(y)\,\varphi(f_y)\,dy, \qquad f_y(x) := f(x-y).$$

On the other hand,

$$\varphi(f * g) = \varphi(f)\varphi(g) = \varphi(f)\int_{\mathbb{R}} g(y)\beta(y)\,dy.$$

Therefore,

$$\int_{\mathbb{R}} \Big(\varphi(f_y) - \varphi(f)\beta(y)\Big) g(y)\,dy = 0 \qquad \text{for all } g \in L^1(\mathbb{R}),$$

so uniqueness in the Riesz representation theorem implies

$$\varphi(f_y) = \varphi(f)\beta(y) \qquad \text{for a.e. } y.$$

Since $y \mapsto f_y$ is continuous from $\mathbb{R}$ into $L^1(\mathbb{R})$ and φ is continuous, the function $y \mapsto \varphi(f_y)$ is continuous. Replacing β by this continuous representative, we may therefore assume that β is continuous.

Now replace y by $x + y$ to obtain

$$\varphi(f)\beta(x+y) = \varphi(f_{x+y}) = \varphi((f_x)_y) = \varphi(f_x)\beta(y) = \varphi(f)\beta(x)\beta(y).$$

Because φ is nonzero, there exists $f \in L^1(\mathbb{R})$ with $\varphi(f) \neq 0$, and hence

$$\beta(x+y) = \beta(x)\beta(y).$$

By Theorem 2.1, there exists $t_0 \in \mathbb{R}$ such that $\beta(x) = e^{-2\pi i t_0 x}$. Hence

$$\varphi(f) = \int_{\mathbb{R}} f(x) e^{-2\pi i t_0 x}\,dx = \widehat{f}(t_0).$$

□

Notice that for every φ (except $\varphi = 0$), there exists unique $t \in \mathbb{R}$ such that $\varphi(f) = \widehat{f}(t)$, because if $s \neq t$, then there exists $f \in L^1(\mathbb{R})$ such that $\widehat{f}(t) \neq \widehat{f}(s)$.

2.15 Poisson Summation Formula

One of the most useful bridges between Fourier series and the Fourier transform is the *Poisson summation formula.* With our Fourier transform convention

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx,$$

the cleanest statement is obtained for rapidly decaying functions.

Theorem 2.45 (Poisson summation, period 1)**.** *Let $f \in \mathcal{S}(\mathbb{R})$ and define its* periodization

$$(\mathcal{P}f)(t) := \sum_{k\in\mathbb{Z}} f(t+k), \qquad t \in \mathbb{R}.$$

Then $\mathcal{P}f$ is 1-periodic and admits the Fourier series expansion

$$\sum_{k\in\mathbb{Z}} f(t+k) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, e^{2\pi i n t} \qquad (t \in \mathbb{R}). \tag{2.5}$$

In particular, at $t = 0$ one has $\sum\limits_{k\in\mathbb{Z}} f(k) = \sum\limits_{n\in\mathbb{Z}} \widehat{f}(n)$.

Proof. Since $f \in \mathcal{S}(\mathbb{R})$, the series defining $\mathcal{P}f$ converges absolutely and uniformly on compact sets, so $\mathcal{P}f \in C^\infty(\mathbb{R})$ and termwise integration is justified. For $n \in \mathbb{Z}$, the n-th Fourier coefficient of $\mathcal{P}f$ is

$$\widehat{(\mathcal{P}f)}(n) = \int_0^1 \Big(\sum_{k\in\mathbb{Z}} f(t+k)\Big) e^{-2\pi i n t}\, dt = \sum_{k\in\mathbb{Z}} \int_0^1 f(t+k)\, e^{-2\pi i n t}\, dt.$$

With the change of variables $x = t + k$ we obtain

$$\widehat{(\mathcal{P}f)}(n) = \int_{\mathbb{R}} f(x)\, e^{-2\pi i n x}\, dx = \widehat{f}(n).$$

The Fourier series of $\mathcal{P}f$ is therefore (2.5), and uniqueness completes the proof. □

Remark 2.46 (A 2π-periodic version). If one prefers 2π-periodic functions, set $\Phi(t) := 2\pi \sum_{k\in\mathbb{Z}} f(t + 2\pi k)$. Then Φ is 2π-periodic and

$$\Phi(t) = \sum_{n\in\mathbb{Z}} \widehat{f}\Big(\frac{n}{2\pi}\Big) e^{int}, \qquad t \in \mathbb{R},$$

which is exactly (2.5) after the change of variables $t \mapsto t/(2\pi)$.

Example 2.47. Using Poisson summation one can derive the classical identity

$$\sum_{n\in\mathbb{Z}} \frac{1}{(n+x)^2} = \frac{\pi^2}{\sin^2(\pi x)}, \qquad x \notin \mathbb{Z}.$$

Hint: apply Theorem 2.45 to the "tent" function $g(x) = (1-|x|)_+$, whose Fourier transform is $\widehat{g}(\xi) = \left(\frac{\sin(\pi\xi)}{\pi\xi}\right)^2$, and differentiate an appropriate Fourier series identity.

2.16 L^p-Derivative of a Function on $\mathbb{R}$

For $h \in \mathbb{R}$ and f a function on $\mathbb{R}$, define

$$D_h f(x) = \frac{f(x+h) - f(x)}{h}.$$

Definition 2.48. A function $f \in L^p(\mathbb{R})$ is said to be differentiable in L^p sense if there exists $g \in L^p(\mathbb{R})$ such that

$$\lim_{h\to 0} \|D_h f - g\|_p = 0.$$

Lemma 2.49. *Let $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Suppose $f \in L^p$, has derivatives f' in L^p sense , then $(f * g)'$ exists in the ordinary sense when $g \in L^q$ and*

$$(f * g)' = f' * g.$$

Proof. We know that $f * g$ is continuous and $f' \in L^p$, therefore $f' * g$ is also continuous. Thus

$$|D_h(f * g)(x) - f' * g(x)| = |(D_h f - f') * g(x)| \le \|D_h f - f'\|_p \|g\|_q \to 0 \text{ as } |h| \to 0$$

Hence

$$(f * g)' = f' * g$$

□

Theorem 2.50. *Let $1 \le p < \infty$ and $f \in L^p(\mathbb{R})$. Then f has an L^p-derivative if and only if, after modifying f on a set of measure zero, the resulting representative is absolutely continuous on every bounded interval $[a, b]$ and its classical derivative belongs to $L^p(\mathbb{R})$.*

Proof of Theorem 2.50. Assume first that $D_h f \to g$ in $L^p(\mathbb{R})$ as $h \to 0$. Let (φ_ε) be a standard mollifier and set

$$f_\varepsilon := f * \varphi_\varepsilon.$$

Then $f_\varepsilon \in C^\infty(\mathbb{R})$ and, by the previous lemma,

$$f_\varepsilon' = g * \varphi_\varepsilon.$$

Fix a bounded interval $I = [a, b]$. Since mollifiers approximate the identity in L^p, we have

$$f_\varepsilon \to f \quad \text{and} \quad f_\varepsilon' \to g \qquad \text{in } L^p(I),$$

hence also in $L^1(I)$ because I has finite measure. Passing to a sequence $\varepsilon_k \downarrow 0$, we may assume that $f_{\varepsilon_k}(x) \to f(x)$ for a.e. $x \in I$. Choose a Lebesgue point $x_0 \in I$ of f; then $f_{\varepsilon_k}(x_0) \to f(x_0)$ as well. For every $x \in I$,

$$f_{\varepsilon_k}(x) = f_{\varepsilon_k}(x_0) + \int_{x_0}^{x} f_{\varepsilon_k}'(t)\, dt.$$

Letting $k \to \infty$ and using the $L^1(I)$-convergence of f_{ε_k}' to g, we obtain for a.e. $x \in I$,

$$f(x) = f(x_0) + \int_{x_0}^{x} g(t)\, dt.$$

Therefore f agrees almost everywhere on I with the absolutely continuous function

$$F_I(x) := f(x_0) + \int_{x_0}^{x} g(t)\, dt,$$

whose classical derivative is g a.e. on I. Since I was arbitrary, f has an absolutely continuous representative on every bounded interval and that representative has derivative $g \in L^p(\mathbb{R})$.

Conversely, suppose that f is absolutely continuous on every bounded interval and that its classical derivative f' belongs to $L^p(\mathbb{R})$. Then for almost every x and every $h \neq 0$,

$$D_h f(x) - f'(x) = \frac{1}{h} \int_0^h \Big(f'(x+t) - f'(x) \Big)\, dt.$$

Applying Minkowski's integral inequality yields

$$\|D_h f - f'\|_p \le \frac{1}{|h|} \int_0^{|h|} \|\tau_{\mathrm{sgn}(h)t} f' - f'\|_p\, dt.$$

Translations are continuous in $L^p(\mathbb{R})$, so the integrand tends to 0 uniformly for $0 \le t \le |h|$ as $h \to 0$. Hence

$$\|D_h f - f'\|_p \to 0,$$

which shows that f' is the L^p-derivative of f. $\square$

2.17 C^∞ form of Urysohn lemma

Lemma 2.51. *Let K be a compact set that is contained in an open set $\mathcal{O} \subset \mathbb{R}$. Then there exists $f \in C_c^\infty(\mathbb{R})$ such that $0 \le f \le 1, f|_K = 1$ and $\mathrm{supp} f \subset \mathcal{O}$.*

Proof. Let $\delta = d(K, \mathcal{O}^c)$. Then $\delta > 0$, and let

$$V = \{x : d(x, K) < \delta/3\}.$$

Suppose $\varphi \in C_c^\infty(\mathbb{R})$ such that $\int \varphi = 1, \varphi(x) = 0$ if $|x| > \delta/3$. Write $f = \chi_V * \varphi$. Then $f|_K = 1$, $0 \le f \le 1$, and $\mathrm{supp}(f) \subset \{x : d(x, K) < 2\delta/3\} \subset \mathcal{O}$, and $f \in C_c^\infty(\mathbb{R})$. Note that φ can be constructed by choosing

$$\varphi(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & |x| < 1 \\ 0 & |x| \ge 1 \end{cases}$$

$\square$

2.18 Problem Sets (Chapter 2: The Fourier Transform)

Conventions. The Fourier transform is

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i\, x\cdot\xi}\, dx,$$

with inverse (when valid) given by

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi)\, e^{2\pi i\, x\cdot\xi}\, d\xi.$$

Unless stated otherwise, functions are complex-valued.

Problem-set architecture. These problems are organized from local technique to global structure. *Tier I* treats basic identities and decay; *Tier II* focuses on convolution and approximation; *Tier III* develops inversion, density, and spectral localization; *Tier IV* contains model computations and uniqueness arguments; and *Tier V* gathers synthesis problems involving periodization, Poisson summation, and L^p differentiability. For a clean progression, work through the tiers in order.

Tier I. Basic properties and quantitative decay

1. **True/false with justification.**

   (a) Let $f \in C_c^\infty(\mathbb{R})$ be nonzero and let P be a polynomial of degree $n \geq 1$. Is $P(\xi)\widehat{f}(\xi)$ necessarily bounded on $\mathbb{R}$?

   (b) Is the subspace $\{f \in L^2(\mathbb{R}) : \operatorname{supp} \widehat{f} \text{ is compact}\}$ dense in $L^2(\mathbb{R})$?

2. Suppose f is continuously differentiable on $[-R, R]$. Prove that there exists $C > 0$ such that

   $$|\widehat{f}(\xi)| \leq \frac{C}{|\xi|} \qquad (\xi \neq 0).$$

   What changes if f is absolutely continuous on $[-R, R]$ with $f' \in L^1([-R, R])$?

3. Let $f \in L^1(\mathbb{R})$ satisfy $f(x) > 0$ for all x. Prove that there exists $\delta > 0$ such that

   $$|\widehat{f}(\xi)| < \widehat{f}(0) \qquad \text{for all } |\xi| > \delta.$$

4. Let $f \in L^1(\mathbb{R})$ with $f \geq 0$. Show that

$$\|\widehat{f}\|_\infty = \widehat{f}(0) = \|f\|_1.$$

5. Let $f \in L^1(\mathbb{R})$ be continuous at 0 and assume $\widehat{f}(\xi) \geq 0$ for all ξ. Prove that $\widehat{f} \in L^1(\mathbb{R})$ and that

$$f(0) = \int_{\mathbb{R}} \widehat{f}(\xi)\, d\xi.$$

6. (Compact support $\Rightarrow$ analyticity.) If $f \in L^1(\mathbb{R})$ has compact support, prove that $\widehat{f}$ is real-analytic on $\mathbb{R}$. Is it true that $\widehat{f} \in L^1(\mathbb{R})$ necessarily? What additional decay can you prove if $f \in C_c^2(\mathbb{R})$?

Tier II. Convolution, kernels, and approximation

1. Let $f, g \in L^2(\mathbb{R})$. Show that $f * g$ admits a bounded continuous representative and that

$$\lim_{|x|\to\infty} (f * g)(x) = 0.$$

2. For $n \in \mathbb{N}$, define $F_n = \chi_{[-1,1]} * \chi_{[-n,n]}$. Verify that $F_n \in C_c(\mathbb{R})$ and $\|F_n\|_\infty = 2$. Does $F_n \to 2$ uniformly on $\mathbb{R}$? Does $F_n \to 2$ pointwise?

3. For $1 \leq p < \infty$, let $f \in L^p(\mathbb{R})$ and define $F(x) = \int_x^{x+1} f(t)\, dt$. Show that $F \in C_0(\mathbb{R})$. Does this remain valid for $f \in L^\infty(\mathbb{R})$?

4. For $f \in L^1(\mathbb{R})$, prove the identity

$$2\widehat{f}(\xi) = \int_{\mathbb{R}} \left[f(x) - f\left(x - \frac{\pi}{\xi}\right)\right] e^{-2\pi i \xi x}\, dx \qquad (\xi \neq 0),$$

and deduce the Riemann–Lebesgue lemma on $\mathbb{R}$.

5. (Good kernels.) Let $\{k_\lambda\} \subset L^1(\mathbb{R})$ be a family of good kernels. If $f \in L^\infty(\mathbb{R}) \cap C(\mathbb{R})$, prove that $f * k_\lambda \to f$ uniformly on every compact subset of $\mathbb{R}$.

6. (Fejér kernel on $\mathbb{R}$.) Let f be of moderate decrease and define

$$f * K_\lambda(x) = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \widehat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

Show that $f * K_\lambda \to f$ uniformly as $\lambda \to \infty$.

Tier III. Inversion, density, and spectral localization

1. Let $f, g \in L^1(\mathbb{R})$. Prove the duality identity

$$\int_{\mathbb{R}} f(y)\widehat{g}(y)\, dy = \int_{\mathbb{R}} \widehat{f}(\xi) g(\xi)\, d\xi.$$

   If $\widehat{f} \in L^1(\mathbb{R})$, deduce Fourier inversion for f.

2. For $1 \leq p \leq 2$, prove that

$$\{f \in L^p(\mathbb{R}) : \operatorname{supp} \widehat{f} \text{ compact}\}$$

   is dense in $L^p(\mathbb{R})$.

3. Show that $X = \{\widehat{f} : f \in L^1(\mathbb{R})\}$ is dense in $C_0(\mathbb{R})$ (with the uniform norm).

4. Let $f \in C_c^2(\mathbb{R})$. Prove that there exists $g \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ such that $\widehat{g} = f$.

5. (Cyclic translations in L^2.) For $f \in L^2(\mathbb{R})$, let $\tau_x f(y) = f(y - x)$. Show that $\{\tau_x f : x \in \mathbb{R}\}$ is dense in $L^2(\mathbb{R})$ if and only if $\widehat{f}(\xi) \neq 0$ for a.e. ξ.

6. (A sharp obstruction in L^1.) For $n \in \mathbb{N}$, let $g_n = \chi_{[-1,1]} * \chi_{[-n,n]}$. Show that g_n is the Fourier transform of

$$f_n(x) = \frac{\sin x \, \sin(nx)}{\pi^2 x^2} \in L^1(\mathbb{R}),$$

   and that $\|f_n\|_1 \to \infty$. Conclude that the Fourier transform maps $L^1(\mathbb{R})$ into a proper subspace of $C_0(\mathbb{R})$.

Tier IV. Gaussians, uniqueness, and model computations

1. For $n \in \mathbb{N}$, define $f(x) = \frac{x^n}{\sqrt{2\pi}} e^{-x^2/2}$. Show that $\widehat{f}(\xi) = P_n(\xi) e^{-\xi^2/2}$, where P_n is a polynomial of degree n. (Identify P_n in terms of Hermite polynomials if you wish.)

2. A continuous function $f : \mathbb{R} \to \mathbb{C}$ is of *moderate decrease* if $|f(x)| \leq A/(1 + x^2)$. Suppose f is of moderate decrease and satisfies

$$\int_{\mathbb{R}} f(y)\, e^{-y^2} e^{2xy}\, dy = 0 \qquad \forall x \in \mathbb{R}.$$

   Prove that $f \equiv 0$.

3. (Optional/advanced: uncertainty principle.) Prove the Heisenberg inequality

$$\left(\int_{\mathbb{R}} x^2|f(x)|^2\,dx\right)\left(\int_{\mathbb{R}} \xi^2|\widehat{f}(\xi)|^2\,d\xi\right) \geq \frac{1}{16\pi^2}\|f\|_2^4$$

for $f \in \mathcal{S}(\mathbb{R})$, and identify (up to the Fourier normalization) the extremizers.

Tier V. Periodization, Poisson summation, and differentiation in L^p

1. For $f \in L^1(\mathbb{R})$, define $f_\lambda(x) = \lambda f(\lambda x)$ and

$$\varphi_\lambda(t) = 2\pi \sum_{j\in\mathbb{Z}} f_\lambda(t + 2\pi j).$$

Show that $\varphi_\lambda \in L^1(S^1)$ and

$$\lim_{\lambda\to\infty} \|\varphi_\lambda\|_{L^1(S^1)} = \|f\|_{L^1(\mathbb{R})}.$$

2. For $f \in L^1(\mathbb{R})$, define $g(t) = 2\pi\sum_{n\in\mathbb{Z}} f(t + 2\pi n)$. Show that g is 2π-periodic and $\|g\|_{L^1(S^1)} \leq \|f\|_{L^1(\mathbb{R})}$. Derive a Poisson summation identity for sufficiently nice f from this periodization.

3. (Difference quotients in L^p.) For $1 \leq p < \infty$, let $f \in L^p(\mathbb{R})$ and define

$$\Delta_h f(x) = \frac{f(x+h) - f(x)}{h}.$$

Show that there exists $g \in L^p(\mathbb{R})$ such that $\|\Delta_h f - g\|_p \to 0$ as $h \to 0$ if and only if f has an absolutely continuous representative on bounded intervals (modulo null sets) with $f' \in L^p(\mathbb{R})$. Discuss what fails for $p = \infty$.

4. (A clean L^∞ substitute.) Suppose $f \in L^\infty(\mathbb{R})$ admits a representative (still denoted f) for which

$$\sup_{h\neq 0} \|\Delta_h f\|_\infty < \infty.$$

Show that f has a Lipschitz representative and that its a.e. derivative belongs to $L^\infty(\mathbb{R})$. Conversely, show that any Lipschitz function satisfies the displayed bound.

5. Give an example of $f \in L^\infty(0,\infty)$ such that f' exists pointwise on $(0,\infty)$ but $f' \notin L^\infty(0,\infty)$.

6. For $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$ with $1 < p < 2$, prove that $f * g \in L^p(\mathbb{R}^n)$, and deduce (by density) that $\widehat{f * g} = \widehat{f}\,\widehat{g}$ whenever both sides make sense.

Chapter 3

Distributions

Many operations in analysis — differentiation, convolution, Fourier transformation — extend well beyond smooth functions. The language of distributions (generalized functions) provides a precise framework for these extensions while remaining compatible with classical calculus whenever the latter makes sense. In this chapter we introduce test function spaces, distributions, and their basic operations, with an eye toward applications in Fourier analysis.

Chapter roadmap. We first build the locally convex topology on spaces of test functions, since continuity is part of the definition of a distribution. With this topological background in place, we introduce distributions, derivatives, multiplication, support, and convergence. The chapter culminates in the tempered setting, where the Fourier transform extends naturally by duality and the formal identities of harmonic analysis become rigorous statements.

In the previous section we saw that some L^p-functions possess derivatives in the L^p sense: there exists $g \in L^p$ such that

$$\|D_h f - g\|_p \to 0 \qquad \text{as } |h| \to 0.$$

However, many important functions are neither classically differentiable nor differentiable in this L^p sense. The theory of distributions provides a broader notion of derivative, defined through its action on a distinguished class of smooth compactly supported functions, called *test functions*.

The guiding identity is the integration-by-parts formula. If f is differentiable and g is a compactly supported differentiable function on $\mathbb{R}$, then

$$\int_{-\infty}^{\infty} f'g = -fg\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} fg' = -\int_{-\infty}^{\infty} fg',$$

because g vanishes outside a compact set. This suggests a definition that continues to make sense even when f' does not exist classically. For $f \in L^1_{\mathrm{loc}}(\mathbb{R})$, define

$$\Lambda_f(g) := \int_{\mathbb{R}} fg, \qquad g \in C_c^\infty(\mathbb{R}).$$

We then define the derivative of Λ_f by

$$\Lambda_f'(g) := -\int_{\mathbb{R}} fg'.$$

More generally,

$$D^k\Lambda_f(g) = (-1)^k \int_{\mathbb{R}} f\, D^k g,$$

where $D = \dfrac{d}{dx}$. In this way differentiation is transferred from the possibly rough function f to the smooth test function g.

To develop distributions systematically, we must endow $C_c^\infty(\mathbb{R}^n)$ with a suitable complete locally convex topology. The supremum norm alone is not adequate for this purpose, because it neither controls derivatives nor records support conditions. The correct topology is built from a family of seminorms on compact subsets.

Example 3.1 (Why distributions are natural)**.** Let $H = \chi_{(0,\infty)}$. Classically, H' does not exist at the origin; distributionally, however, one has $H' = \delta_0$. Thus differentiation of a jump function produces a singular object concentrated at the jump. This example is the prototype for many later constructions in PDE and harmonic analysis.

3.1 Locally Convex Topology

Let $\{p_i : i \in I\}$ be a family of seminorms on a vector space X. For a finite subset $F \subset I$, define

$$U_{F,\varepsilon} := \bigcap_{i\in F} \{x \in X : p_i(x) < \varepsilon\}.$$

Each set $U_{F,\varepsilon}$ is convex and balanced. Set

$$\mathcal{B} := \{U_{F,\varepsilon} : \varepsilon > 0,\ F \subset I,\ \#(F) < \infty\}.$$

We then define

$$\mathcal{T} := \Big\{O \subset X : \text{for every } x \in O \text{ there exists } U \in \mathcal{B} \text{ such that } x + U \subset O\Big\}.$$

This is a topology on X.

Indeed, $\mathcal{T}$ contains $\emptyset$ and X and is closed under arbitrary unions. To verify stability under finite intersections, let $O = \bigcap_{j=1}^{k} O_j$ with each $O_j \in \mathcal{T}$, and fix $x \in O$. For each j there exists $U_{F_j,\varepsilon_j} \in \mathcal{B}$ such that $x + U_{F_j,\varepsilon_j} \subset O_j$. If we set

$$\varepsilon := \min_{1\le j\le k} \varepsilon_j, \qquad F := \bigcup_{j=1}^{k} F_j,$$

then $\varepsilon > 0$, the set F is finite, and

$$x + U_{F,\varepsilon} \subset \bigcap_{j=1}^{k} (x + U_{F_j,\varepsilon_j}) \subset O.$$

The resulting space $(X, \mathcal{T})$ is called a **locally convex topological vector space**.

Example 3.2. Show that a locally convex topological vector space X is Hausdorff if and only if $\{p_i : i \in I\}$ separates points in X *i.e.*, given $x \in X, x \neq 0$, there exists $i \in I$ such that $p_i(x) \neq 0$.

Example 3.3. Let X be a locally convex Hausdorff space whose topology is induced by $\{p_i : i \in I\}$. Define

$$d(x, y) = \sum 2^{-n} \frac{p_n(x-y)}{1 + p_n(x-y)}$$

Show that topology τ_d coincides with $\mathcal{T}$.

Note that, in general settings, $U_{F,\epsilon}$ plays the role of $B_\epsilon(0)$ in $\mathbb{R}^n$ as $B_\epsilon(0), \epsilon > 0$ forms a local base at 0. Therefore,

$$\mathcal{B} = \{U_{F,\epsilon} : \epsilon > 0, F \subset I, \#(F) < \infty\}$$

is a local base at $0 \in X$.

Definition 3.4. (i) A sequence $(x_j)_{j=1}^{\infty} \subset X$ *converges* to $x \in X$ if for every $U \in \mathcal{B}$ there exists $N \in \mathbb{N}$ such that $x_j - x \in U$ for all $j \geq N$.

(ii) A sequence $(x_j)_{j=1}^{\infty} \subset X$ is a *Cauchy sequence* if for every $U \in \mathcal{B}$ there exists $N \in \mathbb{N}$ such that $x_k - x_\ell \in U$ for all $k, \ell \geq N$.

(iii) The space X is *sequentially complete* if every Cauchy sequence in X converges to a point of X.

Lemma 3.5. *A sequence $(x_j)_{j=1}^{\infty} \subset X$ converges to $x \in X$ if and only if*

$$p_i(x_j - x) \longrightarrow 0 \qquad \text{for every } i \in I.$$

Proof. If $x_j \to x$ in the topology generated by $\mathcal{B}$, then for every $i \in I$ and every $\varepsilon > 0$ the neighborhood $U_{\{i\},\varepsilon} = \{y \in X : p_i(y) < \varepsilon\}$ eventually contains $x_j - x$. Hence $p_i(x_j - x) \to 0$. The converse follows immediately from the definition of the basic neighborhoods $U_{F,\varepsilon}$. □

Theorem 3.6. *Let $\{p_i\}_{i \in I}$ be a separating family of seminorms on a vector space X, and define*

$$V_{p,n} := \{x \in X : p(x) < 1/n\}.$$

Then

$$J := \{V_{p_i,n} : i \in I,\ n \in \mathbb{N}\}$$

forms a convex balanced local base for a topology $\mathscr{T}$ on X. With this topology, X becomes a locally convex space such that

(i) each seminorm p_i is continuous;

(ii) a set $E \subset X$ is bounded if and only if $p_i(E)$ is bounded for every $i \in I$.

Proof. Because the family $\{p_i\}$ separates points, for each nonzero $x \in X$ there exists $i \in I$ with $p_i(x) > 0$. Hence $x \notin V_{p_i,n}$ for all sufficiently large n, so $\{0\}$ is closed. By translation invariance, every singleton is closed.

Addition is continuous. Let U be a neighborhood of 0. By definition, there exist indices $i_1, \ldots, i_m$ and integers $n_1, \ldots, n_m$ such that

$$\bigcap_{j=1}^{m} V_{p_{i_j},n_j} \subset U.$$

Set

$$V := \bigcap_{j=1}^{m} V_{p_{i_j},2n_j}.$$

Then $V+V \subset U$. Consequently, if U is an open neighborhood of x_1+x_2, then $U-(x_1+x_2)$ is a neighborhood of 0, so there exists a neighborhood V of 0 such that

$$V + V \subset U - (x_1 + x_2).$$

Hence

$$(x_1 + V) + (x_2 + V) \subset U,$$

which proves continuity of addition.

Scalar multiplication is continuous. Let $x \in X$, $\alpha \in \mathbb{C}$, and let U be a neighborhood of 0. Choose a balanced neighborhood V of 0 such that $V + V \subset U$. Since V absorbs X, there exists $s > 0$ with $x \in sV$. Set $t := \dfrac{s}{1+|\alpha|s}$. If $y \in x + tV$ and $|\beta - \alpha| < 1/s$, then

$$\beta y - \alpha x = \beta(y - x) + (\beta - \alpha)x \in |\beta|tV + |\beta - \alpha|sV \subset V + V \subset U.$$

Because V is balanced and $|\beta|t \leq (|\alpha| + 1/s)t = 1$, scalar multiplication is continuous.

For (ii), first suppose that $E \subset X$ is bounded. Since each $V_{p_i,1}$ is a neighborhood of 0, there exists $k_i > 0$ such that

$$E \subset k_i V_{p_i,1} = \{x \in X : p_i(x) < k_i\}.$$

Thus $p_i(E)$ is bounded for every i. Conversely, assume that each $p_i(E)$ is bounded. Let U be a neighborhood of 0. Then U contains a basic neighborhood of the form

$$\bigcap_{j=1}^{m} V_{p_{i_j},n_j}.$$

Choose $M > 0$ so large that $p_{i_j}(x) < M/n_j$ for every $x \in E$ and every $j = 1, \ldots, m$. Then $E \subset MU$, so E is bounded. □

3.2 Topology of the spaces $C^\infty(\Omega)$ and $\mathcal{D}_K$

We now define the standard topology on $C^\infty(\Omega)$, under which it becomes a Fréchet space with the Heine–Borel property and for which

$$\mathscr{D}_K = \{\varphi \in C^\infty(\Omega) : \operatorname{supp}(\varphi) \subset K\}$$

is a closed subspace whenever $K \subset \Omega$ is compact.

Choose an increasing sequence of compact sets (K_i) such that $K_i \subset K_{i+1}$ and

$$K_i := \Big\{x \in \Omega : d\Big(x, \mathbb{R}^n \setminus \Omega\Big) \geq 1/i\Big\} \cap \overline{B_i},$$

where $B_i := \{x \in \mathbb{R}^n : |x| < i\}$.

For $f \in C^\infty(\Omega)$, define

$$p_N(f) := \sup\Big\{|D^\alpha f(x)| : x \in K_N,\ |\alpha| \leq N\Big\}.$$

Then $\{p_N\}_{N=1}^\infty$ is a separating family of seminorms, and the resulting topology makes $C^\infty(\Omega)$ a metrizable locally convex topological vector space.

For each $x \in \Omega$, define $\delta_x(f) := f(x)$. Then δ_x is continuous for this topology, because

$$p_N(f_j) \to 0 \quad \Longrightarrow \quad |f_j(x)| \leq p_N(f_j) \to 0$$

for all sufficiently large N with $x \in K_N$. Moreover,

$$\mathscr{D}_K = \bigcap_{x \in \Omega \setminus K} \ker \delta_x,$$

so $\mathscr{D}_K$ is a closed subspace of $C^\infty(\Omega)$.

The sets

$$V_N := \{f \in C^\infty(\Omega) : p_N(f) < 1/N\}, \qquad N = 1, 2, \dots,$$

form a convex balanced local base at $0 \in C^\infty(\Omega)$.

Suppose $\{f_j\}$ is a Cauchy sequence in $C^\infty(\Omega)$. Then for each V_N there exists $l_N \in \mathbb{N}$ such that

$$\begin{aligned} & f_i - f_j \in V_N \qquad \text{for all } i, j > l_N, \\ & \Longrightarrow\ p_N(f_i - f_j) < 1/N, \\ & \Longrightarrow\ |D^\alpha f_i(x) - D^\alpha f_j(x)| < 1/N, \qquad x \in K_N. \end{aligned}$$

Thus, for each multi-index α, the sequence $D^\alpha f_i$ converges uniformly on every K_N to a limit g_α. In particular, $f_i \to g_0$ pointwise, $g_0 \in C^\infty(\Omega)$, and $g_\alpha = D^\alpha g_0$. Therefore $f_i \to g_0$ in the topology of $C^\infty(\Omega)$. Hence $C^\infty(\Omega)$ is a Fréchet space, and the same is true for $\mathscr{D}_K$.

Suppose $E \subset C^\infty(\Omega)$ is closed and bounded. By the preceding theorem, for each N there exists $0 < M_N < \infty$ such that $p_N(f) < M_N$ for all $f \in E$.

Thus $|D^\alpha f| < M_N$ on K_N whenever $|\alpha| \le N$. Hence

$$\{D^\beta f : f \in E\}$$

is an equicontinuous family on K_{N-1} whenever $|\beta| \le N-1$. Indeed, by the mean value theorem,

$$|f(x) - f(y)| \le N\|D^1 f\|_\infty |x-y|,$$

and replacing f by $D^\beta f$ yields

$$|D^\beta f(x) - D^\beta f(y)| \le \|D^{\beta+1} f\|_\infty \, |x-y| \le \|f\|_N \, |x-y|.$$

By the Arzelà–Ascoli theorem, every sequence (f_n) in E has a convergent subsequence. Hence E is compact in $C^\infty(\Omega)$, so $C^\infty(\Omega)$ has the Heine–Borel property. Since

$$d(f,0) \le \sum 2^{-n} \frac{p_N(f)}{1+p_N(f)} < 2,$$

the topology on $C^\infty(\Omega)$ is not normable.

For each fixed compact $K \subset \Omega$, the space $\mathscr{D}_K$ is therefore a Fréchet space, and

$$\mathscr{D}(\Omega) = C_c^\infty(\Omega) = \bigcup_{K \subset \Omega} \mathscr{D}_K.$$

This is the space of test functions.

For $\varphi \in \mathscr{D}(\Omega)$, define

$$\|\varphi\|_N := \sup\Big\{|D^\alpha \varphi(x)| : x \in \Omega,\ |\alpha| \le N\Big\}, \qquad N = 0,1,2,\dots.$$

The restriction of these norms to $\mathscr{D}_K$ induces the same topology as the seminorms $\{p_N\}_{N=1}^\infty$. Indeed, if $K \subset \Omega$ is compact, then there exists $N_0 \in \mathbb{N}$ such that $K \subset K_N$ for all $N \ge N_0$, and for such N one has

$$\|\varphi\|_N = p_N(\varphi), \qquad \forall\, \varphi \in \mathscr{D}_K.$$

Since both sequences are increasing in N, they generate the same topology on $\mathscr{D}_K$. Thus

$$V_N := \left\{\varphi \in \mathscr{D}_K : \|\varphi\|_N < \frac{1}{N}\right\}$$

forms a local base for $\mathscr{D}_K$.

Notice that the seminorm family $\{\|\cdot\|_N\}_{N=0}^\infty$ defines a locally convex metrizable topology on $\mathscr{D}(\Omega)$, but this topology is not complete. For example, if $\varphi \in \mathscr{D}(\Omega)$ satisfies $\operatorname{supp}\varphi \subset [0,1]$ and $\varphi > 0$ on $(0,1)$, then

$$\varphi_m(x) = \varphi(x-1) + \frac{1}{2}\varphi(x-2) + \frac{1}{m}\varphi(x-m)$$

is a Cauchy sequence for this topology, but the sequence is not eventually supported in any fixed compact subset of Ω. This shows that the seminorms $\{\|\cdot\|_N\}_{N=0}^{\infty}$ do not prevent Cauchy sequences from drifting to infinity.

We therefore define a finer topology τ on $\mathscr{D}(\Omega)$, under which Cauchy sequences do converge, although τ is not metrizable:

(i) let $\mathcal{B}$ be the collection of all subsets $W \subset \mathscr{D}(\Omega)$ that are convex and balanced and satisfy $\mathscr{D}_K \cap W \in \tau_K$ for every compact $K \subset \Omega$;

(ii) let Σ be the collection of all unions of sets of the form $\varphi + W$, where $\varphi \in \mathscr{D}(\Omega)$ and $W \in \mathcal{B}$.

Note that τ is strictly finer than the topology generated by the seminorms p_N, because it incorporates additional seminorms. For example, if $\{x_i\} \subset \Omega$ has no limit point and $C_i > 0$, then

$$p(\varphi) := \sup_i C_i|\varphi(x_i)| < \infty$$

defines a seminorm on $\mathscr{D}(\Omega)$, since each test function has compact support and therefore meets only finitely many points x_i. Its restriction to every $\mathscr{D}_K$ is continuous. Consequently,

$$W := \{\varphi \in \mathscr{D}(\Omega) : p(\varphi) < C\}$$

is a convex balanced τ-neighborhood of 0. This indicates that every τ-bounded set, and hence every Cauchy sequence, must be concentrated in a common compact subset of Ω.

Theorem 3.7. *(a) τ is a topology on $\mathscr{D}(\Omega)$, and $\mathcal{B}$ is a local base for τ.*

(b) Σ makes $\mathscr{D}(\Omega)$ into a locally convex topological vector space.

Proof. To prove (a), let $V_1, V_2 \in \tau$ and $\varphi \in V_1 \cap V_2$. By definition, there exist $\varphi_i \in \mathscr{D}(\Omega)$ and $W_i \in \mathcal{B}$ such that

$$\varphi \in \varphi_i + W_i \subset V_i, \qquad i = 1, 2.$$

Choose a compact set $K \subset \Omega$ such that $\varphi, \varphi_1, \varphi_2 \in \mathscr{D}_K$. Since $\mathscr{D}_K \cap W_i$ is open in $\mathscr{D}_K$ and

$\varphi - \varphi_i \in \mathscr{D}_K \cap W_i$, there exists $0 < \delta_i < 1$ such that

$$\varphi - \varphi_i + \delta_i(\mathscr{D}_K \cap W_i) \subset \mathscr{D}_K \cap W_i.$$

Because W_i is convex and balanced, this implies

$$\varphi + \delta_i W_i \subset \varphi_i + W_i \subset V_i, \qquad i = 1, 2.$$

Now set $W := \delta_1 W_1 \cap \delta_2 W_2$. Then $W \in \mathcal{B}$ and

$$\varphi + W \subset V_1 \cap V_2,$$

which proves (a).

For (b), translation invariance is immediate from the definition of Σ. Moreover, if $W \in \mathcal{B}$ and $\psi_1, \psi_2 \in \mathscr{D}(\Omega)$, then

$$(\psi_1 + \tfrac{1}{2}W) + (\psi_2 + \tfrac{1}{2}W) = (\psi_1 + \psi_2) + W,$$

so addition is continuous.

To verify continuity of scalar multiplication, fix $\alpha_0 \in \mathbb{C}$, $\varphi_0 \in \mathscr{D}(\Omega)$, and $W \in \mathcal{B}$. Choose $s > 0$ such that $\varphi_0 \in sW$, and set

$$t := \frac{s}{2(1 + s|\alpha_0|)}.$$

If $|\alpha - \alpha_0| < 1/s$ and $\varphi \in \varphi_0 + tW$, then $\varphi - \varphi_0 \in tW$ and therefore

$$\begin{aligned} \alpha\varphi - \alpha_0\varphi_0 &= \alpha(\varphi - \varphi_0) + (\alpha - \alpha_0)\varphi_0 \\ &\in |\alpha| tW + \frac{1}{s}(sW). \end{aligned}$$

Since

$$|\alpha| t \le \Big(|\alpha_0| + 1/s\Big) t = \frac{1}{2},$$

we obtain

$$\alpha\varphi - \alpha_0\varphi_0 \in \tfrac{1}{2}W + \tfrac{1}{2}W = W.$$

Hence scalar multiplication is continuous. From now on, by $D(\Omega)$ we mean $(D(\Omega), \tau)$. □

Theorem 3.8. *(a) A convex balanced subset $V \in \mathscr{D}(\Omega)$ is open if and only if $V \in \mathcal{B}$.*

(b) The topology τ_K of $\mathscr{D}_K \subset \mathscr{D}(\Omega)$ coincides with the topology on $\mathscr{D}_K$ that is inherited from $\mathscr{D}(\Omega)$.

(c) If E is a bounded subset of $\mathscr{D}(\Omega)$, then $E \subset \mathscr{D}_K$ for some compact $K \subset \Omega$ and there exists $0 \leq M_N < \infty$ such that

$$\|\varphi\|_N \leq M_N, \forall \varphi \in E, \qquad N = 0, 1, 2, \dots$$

(d) $D(\Omega)$ has the Heine-Borel property.

(e) $\{\varphi_i\}$ is a Cauchy sequence in $\mathscr{D}(\Omega)$, then $\{\varphi_i\} \in \mathscr{D}_K$ for some $K \subset \Omega$, K compact.

(f) If $\varphi_i \to 0$ in $\mathscr{D}(\Omega)$, then there exists compact set $K \subset \Omega$ such that $\operatorname{supp}\varphi_i \subset K$ for all i, and $D^\alpha \varphi_i \to 0$ uniformly for all α.

(g) In $\mathscr{D}(\Omega)$, every Cauchy sequence is convergent.

Proof. (a) Suppose $V \in \tau$. Claim $V \in \mathcal{B}$. Consider $\varphi \in \mathscr{D}_K \cap V$. By previous theorem, there exists $W \in \mathcal{B}$ such that $\varphi + W \subset V$.

$$\Rightarrow \varphi + (\mathscr{D}_K \cap W) \subset \mathscr{D}_K \cap V$$

Since $\mathscr{D}_K \cap W$ is open in $\mathscr{D}_K$, it implies $\mathscr{D}_K \cap V$ is open in $\mathscr{D}_K$ for each $V \in \tau$.

Conversely, if $V \in \mathcal{B}$, then $V \in \tau$, since $\mathcal{B} \subset \tau$.

(b) Let $V \in \tau$, then $\mathscr{D}_K \cap V \in \tau_K$ (by (a)). That is, $\tau \cap \mathscr{D}_K \in \tau_K$ for all $K \subset \Omega$.

Conversely, suppose $E \in \tau_K$ for some $K \subset \Omega$.

Claim. $E = \mathscr{D}_K \cap V$ for some $V \in \tau$. Let $\varphi \in E$, then there exists N and $\delta > 0$ such that

$$\{\psi \in \mathscr{D}_K : \|\psi - \varphi\|_N < \delta\} \subset E$$

or

$$\{\psi \in \mathscr{D}_K : \|\psi\|_N < \delta\} \subset E - \varphi$$

Let $W_\varphi = \{\psi \in \mathscr{D}_K : \|\psi\|_N < \delta\}$, then $W_\varphi \cap \mathscr{D}_K \in \tau_K$ (an open ball in $\mathscr{D}_K$). Hence $W_\varphi \in \mathcal{B}$, and

$$\mathscr{D}_K \cap (\varphi + W_\varphi) = \varphi + W_\varphi \cap \mathscr{D}_K \subset \varphi + E - \varphi = E$$

Let $V = \bigcup_{\varphi \in E}(\varphi + W_\varphi)$, then

$$\begin{aligned} E &= \bigcup_{\varphi \in E} (\varphi + W_\varphi) \cap \mathscr{D}_K \\ &= \text{ union of all balls around } \varphi \in E \\ &= V \cap \mathscr{D}_K. \end{aligned}$$

(c) Let E be a bounded set in $\mathscr{D}(\Omega)$. Suppose $E \notin \mathscr{D}_K$ for any K. Then there exists $\varphi_m \in E$ and a sequence $\{x_m\} \in \Omega$ having no limit point such that $\varphi_m(x_m) \neq 0$, $m = 1, 2, \dots$

Let
$$W = \left\{\varphi \in \mathscr{D}(\Omega) : |\varphi(x_m)| < \tfrac{1}{m}\varphi_m(x_m), m = 1, 2, \dots\right\}$$

Since each K contains only finitely many x_m,

$$W \cap \mathscr{D}_K = \left\{\varphi \in \mathscr{D}_K : |\varphi(x_m)| < \tfrac{1}{m}\varphi_m(x_m)\right\}$$

is open in $\mathscr{D}_K$. For this, let $\varphi \in W \cap \mathscr{D}_K$. Then $|\varphi(x_m)| < \frac{1}{m}|\varphi_m(x_m)|, m = 1, 2, \dots l$
Let
$$p(\varphi) = \sup_{1 \le m \le l} |\varphi(x_m)| < C_l, \quad \text{where } C_l = \max_{1 \le m \le l} \tfrac{1}{m}|\varphi_m(x_m)|$$

Since p is continuous, it follows that $W \cap \mathscr{D}_K$ is open in $\mathscr{D}_K$. Thus $W \in \mathcal{B}$. Since $\varphi_m \notin mW$ for any m, it follows that E is not bounded.

Thus every bounded set $E \subset \mathscr{D}(\Omega)$ must lie in some $\mathscr{D}_K$. By (b), E is bounded in $\mathscr{D}_K$. This implies

$$\sup\{\|\psi\|_N : \psi \in E\} \le M_N < \infty, \quad N = 0, 1, 2, \dots$$

(d) It follows from (c), since $\mathscr{D}_K$ has the Heine-Borel property. If E is a closed and bounded set in $\mathscr{D}(\Omega)$, then E is closed and bounded in $\mathscr{D}_K$, hence compact. Thus, E is compact in $\mathscr{D}(\Omega)$.

(e) If $\{\varphi_i\}$ is a Cauchy Sequence in $\mathscr{D}(\Omega)$, then it is bounded and hence $\varphi_i \in \mathscr{D}_K$ for some K. By (b), $\{\varphi_i\}$ is Cauchy Sequence relative to $\mathscr{D}_K$.

(f) It is just restatement of (e).

Finally, (g) follows from (b), (e) and completeness of $\mathscr{D}_K$ (i.e., D_K is a Fréchet space).

□

Theorem 3.9. *Let* Λ *be a linear map from* $\mathscr{D}(\Omega)$ *to a locally convex space* Y*. Then the following are equivalent:*

(i) Λ *is continuous.*

(ii) Λ *is bounded.*

(iii) If $\varphi_i \to 0$ *in* $\mathscr{D}(\Omega)$*, then* $\Lambda\varphi_i \to 0$ *in* Y*.*

(iv) For all $K \subset \Omega$*, the restriction* $\Lambda : \mathscr{D}_K \to Y$ *is continuous.*

Proof. $(i) \implies (ii)$: Known.

$(ii) \implies (iii)$: Suppose Λ is bounded and $\varphi_i \to 0$ in $\mathscr{D}(\Omega)$. Then $\varphi_i \to 0$ in some $\mathscr{D}_K$, and hence $\Lambda/\mathscr{D}_K$ is bounded. Therefore, $\Lambda : \mathscr{D}_K \to Y$ is continuous, and thus $\Lambda\varphi_i \to 0$ in Y.

$(iii) \implies (iv)$: Suppose $\{\varphi_i\} \subset \mathscr{D}_K$ and $\varphi_i \to 0$ in $\mathscr{D}_K$. Then by (b) of the previous theorem, $\varphi_i \to 0$ in $\mathscr{D}(\Omega)$. By (iii), $\Lambda\varphi_i \to 0$ in Y.

$(iv) \implies (i)$: Let U be a convex balanced neighborhood of 0 in Y, and write $V = \Lambda^{-1}(U)$. Then V is a convex, also balanced set in $\mathscr{D}(\Omega)$. By (a) of the previous theorem, $V \in \tau$ if and only if $\mathscr{D}_K \cap V \subset \tau_K$ for each $K \subset \Omega$. By (iv), $\mathscr{D}_K \cap V \in \tau_K$, hence $V \in \tau$. Hence Λ is continuous.

□

Definition 3.10. A linear functional Λ on $\mathscr{D}(\Omega)$ that is continuous with respect to the topology τ of $\mathscr{D}(\Omega)$ is called a **distribution**.

The space of all distributions on Ω is denoted by $\mathscr{D}'(\Omega)$.

Theorem 3.11. *Let* Λ *be a linear functional on* $(\mathscr{D}(\Omega), \tau)$*. Then the following are equivalent:*

(i) $\Lambda \in \mathscr{D}'(\Omega)$*.*

(ii) For every compact set $K \subset \Omega$*, there exist* $N \in \mathbb{N}$ *and* $C > 0$ *such that*

$$|\Lambda\psi| \leq C\|\psi\|_N \qquad \text{for all } \psi \in \mathscr{D}_K.$$

This is precisely the equivalence of (i) and (iv) in the previous theorem.

If the integer N above can be chosen independently of the compact set K, then the least such N is called the **order** of the distribution Λ. If no such finite N exists, we say that Λ is of infinite order.

Remark 3.12. Each $\mathscr{D}_K$ is closed in $\mathscr{D}(\Omega)$ and has empty interior. If (K_j) is an increasing exhaustion of Ω by compact sets,

$$\Omega = \bigcup_{j=1}^{\infty} K_j, \qquad K_j \subset K_{j+1},$$

then

$$\mathscr{D}(\Omega) = \bigcup_{j=1}^{\infty} \mathscr{D}_{K_j}.$$

Since $\mathscr{D}(\Omega)$ is sequentially complete, the Baire category theorem implies that $\mathscr{D}(\Omega)$ cannot be metrizable.

Example 3.13. Let $f \in L^{loc}(\mathbb{R}^n)$, then

$$\Lambda_f(\varphi) = \int f\varphi, \quad \varphi \in \mathcal{D}(\mathbb{R}^n)$$

defines a distribution on $\mathcal{D}(\mathbb{R}^n)$. However, every distribution cannot be generated by a function in this way.

For example, Dirac distribution δ_0 cannot be produced by any $f \in L_{loc}(\mathbb{R}^n)$.

On contrary, suppose, there exists $f(\neq 0) \in L_{loc}(\mathbb{R}^n)$ such that $\delta_0(\varphi) = \int f\varphi$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Consider $\varphi_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$ such that support of $\varphi_\varepsilon \subseteq B_\varepsilon(0)$, $0 \leq \varphi_\varepsilon \leq 1$, $\varphi_\varepsilon = 1$ on $B_{\varepsilon/2}(0)$. Then

$$\delta_0(\varphi_\varepsilon) = \int f\varphi_\varepsilon$$

$$\implies 1 = \varphi_\epsilon(0) = \int_{B_\varepsilon(0)} f\varphi_\varepsilon \leq \int_{B_\varepsilon(0)} |f| \to 0 \text{ as } \varepsilon \to 0.$$

However, every distribution is weakly assigned to some derivative of a continuous function. We see it later. Notice that

$$|\delta_0(\varphi)| = |\varphi(0)| \leq \|\varphi\|_\infty = \|\varphi\|_0, \ \forall \varphi \in \mathcal{D}(\mathbb{R}^n)$$

Hence, δ_0 is a distribution of order 0.

Example: Let μ be a Radon measure on Ω. Then

$$\Lambda(\varphi) = \int \varphi(x)\, d\mu(x)$$

defines a distribution and

$$|\Lambda(\varphi)| \le \|\varphi\|_\infty \mu(K),\ \varphi \in \mathcal{D}_K,\ \text{and for every choice of } K,\ \text{compact in } \Omega.$$

Hence, $\Lambda = \Lambda_\mu$ is a distribution of order 0. Later, we see that every distribution of order zero is given by a Radon measure.

3.3 Local Equality of Distribution

Let $\Lambda_i \in \mathcal{D}'(\Omega),\ i = 1, 2,$ and let $O \subset \Omega$ be open. Then we say $\Lambda_1 = \Lambda_2$ in O if

$$\Lambda_1\varphi = \Lambda_2\varphi,\ \forall\varphi \in \mathcal{D}(O).$$

For example, if $f \in L^{loc}(\mathbb{R})$ and $\varphi \in \mathcal{D}(O)$, then $\Lambda_f = 0$ if and only if $f = 0$ almost everywhere on O.

Similarly, if μ is a Radon measure, then $\Lambda_\mu = 0$ if $\mu(B) = 0$, for all $B \in \mathcal{B}(O)$, the Borel σ-algebra on O.

Therefore, distribution can be discussed locally, and that leads to ways to describe distributions globally, if its behavior is known locally.

For this, we need to describe "partition of unity".

Theorem 3.14 (Partition of unity)**.** *Let $\mathcal{A} = \{O_i : i \in I\}$ be an open cover of Ω. Then there exists a sequence $\{\psi_j\}_{j\ge1} \subset \mathcal{D}(\Omega)$ with $\psi_j \ge 0$ such that*

(i) $\operatorname{supp}\psi_j \subset O_{i(j)}$ *for some* $O_{i(j)} \in \mathcal{A}$*;*

(ii) $\sum_{j=1}^{\infty}\psi_j(x) = 1$ *for every* $x \in \Omega$*;*

(iii) for every compact $K \subset \Omega$ there exist $N \in \mathbb{N}$ and an open neighborhood V of K such that

$$\psi_1 + \cdots + \psi_N = 1 \qquad \text{on } V.$$

In particular, the family $\{\psi_j\}$ is locally finite.

Proof. Since $\Omega \subset \mathbb{R}^n$ is paracompact and second countable, the cover $\mathcal{A}$ admits a countable locally finite refinement $\{V_j\}_{j\ge1}$ with $\overline{V_j} \Subset O_{i(j)}$ for suitable $O_{i(j)} \in \mathcal{A}$. Choose $\eta_j \in C_c^\infty(\Omega)$ such that

$$0 \le \eta_j \le 1, \qquad \eta_j > 0 \text{ on } V_j, \qquad \operatorname{supp}\eta_j \subset O_{i(j)}.$$

Because the family $\{V_j\}$ is locally finite, the sum

$$s(x) := \sum_{j=1}^{\infty} \eta_j(x)$$

is locally finite, hence defines a smooth positive function on Ω. Set

$$\psi_j(x) := \frac{\eta_j(x)}{s(x)}, \qquad x \in \Omega.$$

Then each $\psi_j \in \mathcal{D}(\Omega)$, $\psi_j \geq 0$, and $\operatorname{supp} \psi_j \subset \operatorname{supp} \eta_j \subset O_{i(j)}$, proving (i). Also

$$\sum_{j=1}^{\infty} \psi_j(x) = \frac{\sum_j \eta_j(x)}{s(x)} = 1,$$

so (ii) holds.

Finally, let $K \subset \Omega$ be compact. Local finiteness implies that only finitely many $\operatorname{supp} \eta_j$ meet K. Hence there exists an open neighborhood V of K on which at most finitely many η_j are nonzero, say $\eta_1, \ldots, \eta_N$. On V we then have

$$\psi_1 + \cdots + \psi_N = \sum_{j=1}^{\infty} \psi_j = 1,$$

which proves (iii). □

Proposition 3.15 (Local equality is a local property)**.** *Let $\Lambda_1, \Lambda_2 \in \mathcal{D}'(\Omega)$. Suppose that for every $x \in \Omega$ there exists an open neighborhood $O_x \subset \Omega$ such that $\Lambda_1 = \Lambda_2$ in O_x. Then $\Lambda_1 = \Lambda_2$ in Ω.*

Proof. Choose a partition of unity $\{\psi_j\}_{j \geq 1} \subset \mathcal{D}(\Omega)$ subordinate to the cover $\{O_x : x \in \Omega\}$. If $\varphi \in \mathcal{D}(\Omega)$, then only finitely many products $\psi_j \varphi$ are nonzero, and each such product belongs to $\mathcal{D}(O_{x_j})$ for a suitable x_j. Therefore

$$\Lambda_1(\varphi) = \sum_j \Lambda_1(\psi_j \varphi) = \sum_j \Lambda_2(\psi_j \varphi) = \Lambda_2(\varphi).$$

Hence $\Lambda_1 = \Lambda_2$ on Ω. □

Theorem 3.16 (Gluing compatible local distributions)**.** *Let $\mathcal{A}$ be an open cover of Ω, and for each $O \in \mathcal{A}$ let $\Lambda_O \in \mathcal{D}'(O)$. Assume that whenever $O, O' \in \mathcal{A}$, the restrictions of Λ_O and $\Lambda_{O'}$ agree on $O \cap O'$. Then there exists a unique distribution $\widetilde{\Lambda} \in \mathcal{D}'(\Omega)$ whose restriction to each $O \in \mathcal{A}$ is Λ_O.*

Proof. Choose a partition of unity $\{\psi_i\}_{i\geq 1} \subset \mathcal{D}(\Omega)$ subordinate to $\mathcal{A}$, and for each i choose $O_i \in \mathcal{A}$ such that $\operatorname{supp}\psi_i \subset O_i$. Define, for $\varphi \in \mathcal{D}(\Omega)$,

$$\widetilde{\Lambda}(\varphi) := \sum_{i=1}^{\infty} \Lambda_{O_i}(\psi_i\varphi).$$

The sum is finite because $\operatorname{supp}\varphi$ is compact and the partition of unity is locally finite. Hence $\widetilde{\Lambda}$ is well defined and linear.

To prove continuity, fix a compact set $K \subset \Omega$. Only finitely many $\operatorname{supp}\psi_i$ meet K; call them $\psi_1, \dots, \psi_N$. For $\varphi \in \mathcal{D}_K$ we then have

$$\widetilde{\Lambda}(\varphi) = \sum_{i=1}^{N} \Lambda_{O_i}(\psi_i\varphi).$$

For each i, multiplication by the fixed smooth function ψ_i is a continuous map $\mathcal{D}_K \to \mathcal{D}(\Omega)$, and its image is contained in $\mathcal{D}(\operatorname{supp}\psi_i) \subset \mathcal{D}(O_i)$. Since each Λ_{O_i} is continuous on $\mathcal{D}(O_i)$, the map $\varphi \mapsto \Lambda_{O_i}(\psi_i\varphi)$ is continuous on $\mathcal{D}_K$. A finite sum of continuous maps is continuous, so $\widetilde{\Lambda} \in \mathcal{D}'(\Omega)$.

Now fix $O \in \mathcal{A}$ and $\varphi \in \mathcal{D}(O)$. For every i with $\psi_i\varphi \neq 0$ we have $\psi_i\varphi \in \mathcal{D}(O_i \cap O)$, and by compatibility

$$\Lambda_{O_i}(\psi_i\varphi) = \Lambda_O(\psi_i\varphi).$$

Therefore

$$\widetilde{\Lambda}(\varphi) = \sum_i \Lambda_O(\psi_i\varphi) = \Lambda_O\Big(\sum_i \psi_i\varphi\Big) = \Lambda_O(\varphi),$$

so $\widetilde{\Lambda}|_O = \Lambda_O$.

Uniqueness follows from the preceding proposition on local equality: if another distribution has the same restriction to every $O \in \mathcal{A}$, then the two distributions agree in a neighborhood of each point of Ω, hence agree on all of Ω. □

Theorem 3.17 (Order-zero distributions are measures)**.** *A distribution $\Lambda \in \mathcal{D}'(\Omega)$ is of order 0 if and only if there exists a (possibly complex-valued) Radon measure μ on Ω such that*

$$\Lambda(\varphi) = \int_\Omega \varphi\, d\mu, \qquad \varphi \in \mathcal{D}(\Omega).$$

Proof. First suppose that μ is a Radon measure. If $\varphi \in \mathcal{D}_K$ for a compact set $K \subset \Omega$, then

$$|\Lambda_\mu(\varphi)| = \left|\int_K \varphi\, d\mu\right| \leq \int_K |\varphi|\, d|\mu| \leq |\mu|(K)\, \|\varphi\|_0.$$

Hence Λ_μ has order 0.

Conversely, assume that Λ has order 0. Choose a locally finite open cover $\{U_i\}_{i\geq 1}$ of Ω by relatively compact sets. For each i, the closure $\overline{U_i}$ is compact, so by the order-zero hypothesis there exists $C_i > 0$ such that

$$|\Lambda(\varphi)| \leq C_i \|\varphi\|_\infty, \qquad \varphi \in \mathcal{D}(U_i).$$

Since $\mathcal{D}(U_i)$ is dense in $C_0(U_i)$ for the supremum norm, $\Lambda|_{\mathcal{D}(U_i)}$ extends uniquely to a continuous linear functional on $C_0(U_i)$. By the Riesz–Markov theorem, there exists a finite complex Radon measure μ_i on U_i such that

$$\Lambda(\varphi) = \int_{U_i} \varphi \, d\mu_i, \qquad \varphi \in \mathcal{D}(U_i).$$

Thus the local distributions Λ_{μ_i} agree with Λ on U_i, and therefore agree with one another on overlaps.

Let $\{\psi_i\}_{i\geq 1} \subset \mathcal{D}(\Omega)$ be a partition of unity subordinate to $\{U_i\}$. Because the family is locally finite, the expression

$$d\mu := \sum_{i=1}^{\infty} \psi_i \, d\mu_i$$

defines a complex Radon measure on Ω: on each compact set only finitely many terms contribute. Finally, if $\varphi \in \mathcal{D}(\Omega)$, then

$$\int_\Omega \varphi \, d\mu = \sum_{i=1}^{\infty} \int_{U_i} \psi_i \varphi \, d\mu_i = \sum_{i=1}^{\infty} \Lambda(\psi_i \varphi) = \Lambda\Big(\sum_{i=1}^{\infty} \psi_i \varphi \Big) = \Lambda(\varphi).$$

Hence $\Lambda = \Lambda_\mu$, as required. □

3.4 Derivative of a distribution

Let $\Omega \subset \mathbb{R}^n$ be open and let $\mathcal{D}(\Omega) = C_c^\infty(\Omega)$ denote the space of test functions. For $\Lambda \in \mathcal{D}'(\Omega)$ and a multi-index $\alpha \in \mathbb{N}_0^n$ we define the *distributional derivative* $\partial^\alpha \Lambda \in \mathcal{D}'(\Omega)$ by

$$(\partial^\alpha \Lambda)(\varphi) := (-1)^{|\alpha|} \Lambda(\partial^\alpha \varphi), \qquad \varphi \in \mathcal{D}(\Omega). \tag{3.1}$$

Linearity is immediate. To see continuity, fix a compact $K \Subset \Omega$. Since Λ is continuous on $\mathcal{D}_K$, there exist $C > 0$ and $N \in \mathbb{N}$ such that

$$|\Lambda(\psi)| \leq C \|\psi\|_N, \qquad \psi \in \mathcal{D}_K.$$

Applying this to $\psi = \partial^\alpha \varphi$ yields

$$|(\partial^\alpha \Lambda)(\varphi)| = |\Lambda(\partial^\alpha \varphi)| \le C\|\partial^\alpha \varphi\|_N \le C'\|\varphi\|_{N+|\alpha|}, \qquad \varphi \in \mathcal{D}_K,$$

hence $\partial^\alpha \Lambda \in \mathcal{D}'(\Omega)$.

The definition (3.1) immediately implies $\partial^\alpha \partial^\beta \Lambda = \partial^{\alpha+\beta}\Lambda = \partial^\beta \partial^\alpha \Lambda$.

Example 3.18 (Distributions induced by functions)**.** If $f \in L^1_{\mathrm{loc}}(\Omega)$, then

$$\Lambda_f(\varphi) := \int_\Omega f(x)\varphi(x)\, dx, \qquad \varphi \in \mathcal{D}(\Omega),$$

defines a distribution. For every multi-index α,

$$(\partial^\alpha \Lambda_f)(\varphi) = (-1)^{|\alpha|} \int_\Omega f(x)\, \partial^\alpha \varphi(x)\, dx.$$

If, in addition, f admits a weak derivative $\partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$ (for instance if $f \in C^{|\alpha|}(\Omega)$), then the usual integration-by-parts identity on compact supports shows $\partial^\alpha \Lambda_f = \Lambda_{\partial^\alpha f}$.

Example 3.19 (Cantor function)**.** Let $\Omega = (-2, 2)$ and let F be the Cantor function on $[0, 1]$, extended by constants outside $[0, 1]$. Then $F \in L^1_{\mathrm{loc}}(\Omega)$ and $F' = 0$ almost everywhere in the classical sense. Nevertheless, the distributional derivative F' is *not* the zero distribution: it is the (singular) Cantor measure μ_C, characterized by

$$\Lambda'_F(\varphi) = -(\Lambda_F)(\varphi') = \int_\Omega \varphi\, d\mu_C, \qquad \varphi \in \mathcal{D}(\Omega).$$

Example 3.20 (Absolutely continuous functions in one dimension)**.** If $\Omega \subset \mathbb{R}$ and f is absolutely continuous on every compact interval $[a, b] \Subset \Omega$, then $f' \in L^1_{\mathrm{loc}}(\Omega)$ and $\Lambda'_f = \Lambda_{f'}$, i.e.

$$\int_\Omega f'(x)\varphi(x)\, dx = -\int_\Omega f(x)\varphi'(x)\, dx, \qquad \varphi \in \mathcal{D}(\Omega).$$

3.5 Multiplication by a smooth function

Let $\Lambda \in \mathcal{D}'(\Omega)$ and let $f \in C^\infty(\Omega)$. Define the distribution $f\Lambda \in \mathcal{D}'(\Omega)$ by

$$(f\Lambda)(\varphi) := \Lambda(f\varphi), \qquad \varphi \in \mathcal{D}(\Omega).$$

To verify continuity, fix a compact $K \Subset \Omega$. By Leibniz' rule, for every multi-index α,

$$\partial^\alpha(f\varphi) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} (\partial^{\alpha-\beta} f)(\partial^\beta \varphi).$$

Thus there exists $C_{K,N}(f) > 0$ such that $\|f\varphi\|_N \le C_{K,N}(f)\|\varphi\|_N$ for all $\varphi \in \mathcal{D}_K$, and consequently $|(f\Lambda)(\varphi)| \le C'\|\varphi\|_N$ on $\mathcal{D}_K$.

3.6 Sequences of distributions

A sequence $\{\Lambda_j\}_{j\ge1} \subset \mathcal{D}'(\Omega)$ is said to converge to $\Lambda \in \mathcal{D}'(\Omega)$, written $\Lambda_j \to \Lambda$ in $\mathcal{D}'(\Omega)$, if

$$\Lambda_j(\varphi) \longrightarrow \Lambda(\varphi) \qquad \text{for every } \varphi \in \mathcal{D}(\Omega).$$

In particular, if $f_j \in L^1_{\text{loc}}(\Omega)$ and Λ_{f_j} denotes the associated distribution, then $\Lambda_{f_j} \to \Lambda_f$ in $\mathcal{D}'(\Omega)$ precisely when

$$\int_\Omega f_j(x)\varphi(x)\,dx \longrightarrow \int_\Omega f(x)\varphi(x)\,dx \qquad \text{for all } \varphi \in \mathcal{D}(\Omega).$$

Theorem 3.21 (Pointwise limits)**.** *Let $\{\Lambda_j\}_{j\ge1} \subset \mathcal{D}'(\Omega)$ and suppose that for each $\varphi \in \mathcal{D}(\Omega)$ the limit $\Lambda(\varphi) := \lim_{j\to\infty} \Lambda_j(\varphi)$ exists. Then $\Lambda \in \mathcal{D}'(\Omega)$ and $\Lambda_j \to \Lambda$ in $\mathcal{D}'(\Omega)$. Moreover, for every multi-index α one has $\partial^\alpha\Lambda_j \to \partial^\alpha\Lambda$ in $\mathcal{D}'(\Omega)$.*

Proof. The mapping $\Lambda : \mathcal{D}(\Omega) \to \mathbb{C}$ is linear by construction. Continuity follows from the Banach–Steinhaus theorem (uniform boundedness) applied to the barreled locally convex space $\mathcal{D}(\Omega)$: pointwise boundedness of $\{\Lambda_j\}$ on $\mathcal{D}(\Omega)$ implies equicontinuity on each $\mathcal{D}_K$, hence Λ is continuous on each $\mathcal{D}_K$ and therefore $\Lambda \in \mathcal{D}'(\Omega)$.

Finally,

$$(\partial^\alpha\Lambda_j)(\varphi) = (-1)^{|\alpha|}\Lambda_j(\partial^\alpha\varphi) \to (-1)^{|\alpha|}\Lambda(\partial^\alpha\varphi) = (\partial^\alpha\Lambda)(\varphi)$$

for each $\varphi \in \mathcal{D}(\Omega)$, which is exactly $\partial^\alpha\Lambda_j \to \partial^\alpha\Lambda$ in $\mathcal{D}'(\Omega)$. □

3.7 Support of a Distribution

Let $O \subset \Omega$ be open and let $\Lambda \in \mathcal{D}'(\Omega)$. We say that Λ vanishes on O if

$$\Lambda(\varphi) = 0, \qquad \forall\, \varphi \in \mathcal{D}(O).$$

Let

$$W := \bigcup \{O \subset \Omega : \Lambda|_O = 0\}.$$

Then $\Lambda|_W = 0$, and the complement of W is called the *support* of Λ.

Indeed, the family of open sets on which Λ vanishes forms an open cover of W. Choose a partition of unity $\{\psi_i\}$ subordinate to this cover. Then, for every $\varphi \in \mathcal{D}(W)$,

$$\varphi = \sum_{i=1}^{\infty} \psi_i \varphi,$$

and each summand $\psi_i \varphi$ is supported in an open set on which Λ vanishes. Hence

$$\Lambda(\varphi) = \sum_{i=1}^{\infty} \Lambda(\psi_i \varphi) = 0,$$

so $\Lambda|_W = 0$ as claimed.

Theorem 3.22. *Let $\Lambda \in \mathcal{D}'(\Omega)$ and set $S_\Lambda := \operatorname{supp} \Lambda$.*

(a) If $\operatorname{supp} \varphi \cap S_\Lambda = \varnothing$ for some $\varphi \in \mathcal{D}(\Omega)$, then $\Lambda(\varphi) = 0$.

(b) If $S_\Lambda = \varnothing$, then $\Lambda = 0$.

(c) If $\psi \in C^\infty(\Omega)$ and $\psi = 1$ on an open set $V \supset S_\Lambda$, then $\psi\Lambda = \Lambda$.

(d) If S_Λ is compact, then Λ is of finite order. More precisely, there exist $C > 0$ and $N \in \mathbb{N} \cup \{0\}$ such that

$$|\Lambda \varphi| \le C \|\varphi\|_N, \qquad \forall\, \varphi \in \mathcal{D}(\Omega).$$

Moreover, Λ extends uniquely to a continuous linear functional on $C^\infty(\Omega)$.

Proof. Parts (a) and (b) follow directly from the definition of the support.

(c) If $\psi = 1$ on $V \supset S_\Lambda$, then

$$\operatorname{supp}(\varphi - \psi\varphi) \cap S_\Lambda = \varnothing, \qquad \forall\, \varphi \in \mathcal{D}(\Omega).$$

Hence part (a) gives $\Lambda(\varphi - \psi\varphi) = 0$, so

$$\Lambda(\varphi) = \Lambda(\psi\varphi), \qquad \forall\, \varphi \in \mathcal{D}(\Omega),$$

which is exactly the identity $\psi\Lambda = \Lambda$.

(d) Assume that S_Λ is compact. Choose $\psi \in C_c^\infty(\Omega)$ such that $\psi = 1$ on an open neighborhood of S_Λ, and set $K := \operatorname{supp}\psi$. By part (c),

$$\Lambda(\varphi) = \Lambda(\psi\varphi), \qquad \forall\, \varphi \in \mathcal{D}(\Omega).$$

Since $\Lambda \in \mathscr{D}'(\Omega)$, there exist $C_1 > 0$ and $N \in \mathbb{N} \cup \{0\}$ such that

$$|\Lambda(\eta)| \le C_1 \|\eta\|_N, \qquad \forall\, \eta \in \mathscr{D}_K.$$

By Leibniz' rule there exists $C_2 > 0$ such that

$$\|\psi\varphi\|_N \le C_2 \|\varphi\|_N, \qquad \forall\, \varphi \in \mathcal{D}(\Omega).$$

Therefore,

$$|\Lambda(\varphi)| = |\Lambda(\psi\varphi)| \le C_1 \|\psi\varphi\|_N \le C_1 C_2 \|\varphi\|_N,$$

which proves that Λ has finite order.

Now define, for $f \in C^\infty(\Omega)$,

$$\widetilde{\Lambda}(f) := \Lambda(\psi f).$$

If $f_j \to 0$ in $C^\infty(\Omega)$, then all derivatives $D^\alpha f_j \to 0$ uniformly on compact subsets of Ω. Since ψ has compact support, Leibniz' rule implies that $\psi f_j \to 0$ in $\mathcal{D}(\Omega)$, and hence $\widetilde{\Lambda}(f_j) = \Lambda(\psi f_j) \to 0$. Thus $\widetilde{\Lambda}$ is continuous on $C^\infty(\Omega)$.

Finally, if $f \in C^\infty(\Omega)$ and $K_0 \subset \Omega$ is compact, then by Urysohn's lemma there exists $\chi \in \mathcal{D}(\Omega)$ such that $\chi = 1$ on K_0. Therefore $f\chi \in \mathcal{D}(\Omega)$ and agrees with f on K_0, which shows that $\mathcal{D}(\Omega)$ is dense in $C^\infty(\Omega)$. Hence the extension is unique. □

3.8 Schwartz space and tempered distributions

The test-function space $\mathcal{D}(\mathbb{R}^n) = C_c^\infty(\mathbb{R}^n)$ is tailored to local questions, but Fourier analysis on $\mathbb{R}^n$ interacts more naturally with the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ introduced in Section 2.3. Rapid decay at infinity is exactly what allows differentiation, multiplication by polynomials, and the Fourier transform to remain inside a single function class.

Definition 3.23 (Tempered distributions)**.** A *tempered distribution* on $\mathbb{R}^n$ is a continuous linear functional on the Fréchet space $\mathcal{S}(\mathbb{R}^n)$. The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$.

Remark 3.24*.* The inclusion $\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathcal{S}(\mathbb{R}^n)$ is continuous. Consequently every tempered distribution restricts canonically to a distribution on $\mathbb{R}^n$. The converse is false: distributions whose growth at infinity is faster than polynomial need not act continuously on $\mathcal{S}(\mathbb{R}^n)$.

Proposition 3.25 (Compact support implies tempered)**.** *Every compactly supported distribution on $\mathbb{R}^n$ belongs to $\mathcal{S}'(\mathbb{R}^n)$.*

Proof. Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and assume $\operatorname{supp} T \subset K$ for some compact set K. Then there exist $C > 0$ and $N \in \mathbb{N}$ such that

$$|T(\phi)| \le C \sum_{|\alpha| \le N} \sup_{x \in K} |\partial^\alpha \phi(x)|, \qquad \phi \in \mathcal{D}(\mathbb{R}^n).$$

Choose $\chi \in C_c^\infty(\mathbb{R}^n)$ such that $\chi = 1$ on a neighborhood of K. For $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we have $T(\varphi) = T(\chi\varphi)$, because $(1-\chi)\varphi$ vanishes on a neighborhood of K. Therefore

$$|T(\varphi)| = |T(\chi\varphi)| \le C \sum_{|\alpha| \le N} \sup_{x \in K} |\partial^\alpha (\chi\varphi)(x)|.$$

By Leibniz' rule, the right-hand side is bounded by a finite linear combination of Schwartz seminorms of φ. Hence T is continuous on $\mathcal{S}(\mathbb{R}^n)$, i.e. $T \in \mathcal{S}'(\mathbb{R}^n)$. □

Proposition 3.26 (Polynomial growth defines a tempered distribution)**.** *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. If*

$$|f(x)| \le C(1 + |x|)^m \qquad (x \in \mathbb{R}^n)$$

for some constants $C, m \ge 0$, then

$$T_f(\varphi) := \int_{\mathbb{R}^n} f(x)\varphi(x)\, dx, \qquad \varphi \in \mathcal{S}(\mathbb{R}^n),$$

defines a tempered distribution.

Proof. Choose an integer $M > m + n$. Then for $\varphi \in \mathcal{S}(\mathbb{R}^n)$,

$$|T_f(\varphi)| \leq C \int_{\mathbb{R}^n} (1+|x|)^m |\varphi(x)|\, dx \leq C \Big(\sup_{x \in \mathbb{R}^n} (1+|x|)^M |\varphi(x)| \Big) \int_{\mathbb{R}^n} (1+|x|)^{m-M}\, dx.$$

The last integral is finite because $M - m > n$. Thus $|T_f(\varphi)|$ is bounded by a Schwartz seminorm of φ, so $T_f \in \mathcal{S}'(\mathbb{R}^n)$. □

Example 3.27 (Dirac masses and their derivatives)**.** For $x_0 \in \mathbb{R}^n$, the Dirac mass δ_{x_0} is tempered and

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0), \qquad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Since δ_{x_0} has compact support, Proposition 3.25 applies. The same proposition also yields that every derivative $\partial^\alpha \delta_{x_0}$ is tempered.

3.9 Fourier transform on tempered distributions

The Fourier transform on $\mathcal{S}(\mathbb{R}^n)$ (Proposition 2.22) is a continuous automorphism of the Schwartz space. Duality therefore extends it automatically to $\mathcal{S}'(\mathbb{R}^n)$.

Definition 3.28 (Fourier transform of a tempered distribution)**.** Let $T \in \mathcal{S}'(\mathbb{R}^n)$. Its *Fourier transform* $\widehat{T} \in \mathcal{S}'(\mathbb{R}^n)$ is defined by

$$\langle \widehat{T}, \varphi \rangle := \langle T, \widehat{\varphi} \rangle, \qquad \varphi \in \mathcal{S}(\mathbb{R}^n),$$

where $\widehat{\varphi}$ is the Fourier transform in the convention of Chapter 2.

Proposition 3.29 (Well-definedness and inversion)**.** *The assignment $T \mapsto \widehat{T}$ defines a linear automorphism of $\mathcal{S}'(\mathbb{R}^n)$. Its inverse is given by*

$$\langle \mathcal{F}^{-1} T, \varphi \rangle := \langle T, \mathcal{F}^{-1} \varphi \rangle, \qquad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Proof. Since $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is continuous, the map

$$\varphi \longmapsto \langle T, \widehat{\varphi} \rangle$$

is a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$ for every $T \in \mathcal{S}'(\mathbb{R}^n)$. Thus $\widehat{T}$ is well defined and tempered. Linearity is immediate.

The same argument applies to $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. For $\varphi \in \mathcal{S}(\mathbb{R}^n)$, Fourier inversion on $\mathcal{S}(\mathbb{R}^n)$ gives

$$\langle \mathcal{F}^{-1}(\widehat{T}), \varphi \rangle = \langle \widehat{T}, \mathcal{F}^{-1}\varphi \rangle = \langle T, \widehat{(\mathcal{F}^{-1}\varphi)} \rangle = \langle T, \varphi \rangle.$$

Hence $\mathcal{F}^{-1}(\widehat{T}) = T$. Similarly, $\widehat{(\mathcal{F}^{-1}T)} = T$. Therefore $T \mapsto \widehat{T}$ is a linear automorphism of $\mathcal{S}'(\mathbb{R}^n)$. □

Proposition 3.30 (Basic identities)**.** *Let $T \in \mathcal{S}'(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$. Then, in $\mathcal{S}'(\mathbb{R}^n)$,*

$$\widehat{\partial^\alpha T} = (2\pi i\, \xi)^\alpha\, \widehat{T}, \qquad \widehat{x^\alpha T} = \left(\frac{i}{2\pi}\right)^{|\alpha|} \partial^\alpha \widehat{T}.$$

If translations are defined by $(\tau_y T)(\varphi) := T(\tau_{-y}\varphi)$, then

$$\widehat{\tau_y T} = e^{-2\pi i\, \xi \cdot y}\, \widehat{T}.$$

Proof. We use the Schwartz identities from (2.2).

For the derivative formula, let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then

$$\langle \widehat{\partial^\alpha T}, \varphi \rangle = \langle \partial^\alpha T, \widehat{\varphi} \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \widehat{\varphi} \rangle.$$

By Proposition 2.22,

$$\partial^\alpha \widehat{\varphi} = \widehat{\left(\frac{i}{2\pi}\right)^{|\alpha|} x^\alpha \varphi}.$$

Hence

$$\langle \widehat{\partial^\alpha T}, \varphi \rangle = \left(\frac{-i}{2\pi}\right)^{|\alpha|} \langle T, \widehat{x^\alpha \varphi} \rangle = \left(\frac{-i}{2\pi}\right)^{|\alpha|} \langle \widehat{T}, x^\alpha \varphi \rangle = \langle (2\pi i\, \xi)^\alpha \widehat{T}, \varphi \rangle,$$

which proves the first identity.

For multiplication by x^α,

$$\langle \widehat{x^\alpha T}, \varphi \rangle = \langle x^\alpha T, \widehat{\varphi} \rangle = \langle T, x^\alpha \widehat{\varphi} \rangle.$$

Again by (2.2),

$$x^\alpha \widehat{\varphi} = \widehat{\left(\frac{i}{2\pi}\right)^{|\alpha|} \partial^\alpha \varphi}.$$

Therefore

$$\langle \widehat{x^\alpha T}, \varphi \rangle = \left(\frac{i}{2\pi}\right)^{|\alpha|} \langle T, \widehat{\partial^\alpha \varphi} \rangle = \left(\frac{i}{2\pi}\right)^{|\alpha|} \langle \widehat{T}, \partial^\alpha \varphi \rangle = \left(\frac{i}{2\pi}\right)^{|\alpha|} \langle \partial^\alpha \widehat{T}, \varphi \rangle.$$

For translations, note that for every Schwartz function φ,

$$\widehat{e^{-2\pi i\,\xi\cdot y}\varphi}(\xi)\,(x) = \widehat{\varphi}(x+y) = \tau_{-y}\widehat{\varphi}(x).$$

Hence

$$\langle \widehat{\tau_y T}, \varphi\rangle = \langle \tau_y T, \widehat{\varphi}\rangle = \langle T, \tau_{-y}\widehat{\varphi}\rangle = \left\langle T,\, \widehat{e^{-2\pi i\,\xi\cdot y}\varphi}(\xi)\right\rangle = \langle \widehat{T}, e^{-2\pi i\,\xi\cdot y}\varphi\rangle,$$

which is exactly the claimed identity. □

Example 3.31 (The transforms of 1 and δ_0)**.** With our convention,

$$\widehat{\delta_0} = 1 \qquad \text{in } \mathcal{S}'(\mathbb{R}^n),$$

because for $\varphi \in \mathcal{S}(\mathbb{R}^n)$,

$$\langle \widehat{\delta_0}, \varphi\rangle = \langle \delta_0, \widehat{\varphi}\rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}^n} \varphi(x)\,dx = \langle 1, \varphi\rangle.$$

Applying Proposition 3.29 to the inverse transform yields $\widehat{1} = \delta_0$.

Remark 3.32. The tempered-distribution viewpoint turns many formal Fourier computations into honest identities. For example, if $u \in \mathcal{S}'(\mathbb{R}^n)$, then

$$\widehat{\Delta u}(\xi) = -4\pi^2|\xi|^2\widehat{u}(\xi)$$

in $\mathcal{S}'(\mathbb{R}^n)$. This is one of the basic reasons that tempered distributions are indispensable in PDE and harmonic analysis.

3.10 Problem Sets (Chapter 3: Distributions and Tempered Distributions)

Conventions. We write $\mathcal{D}(\Omega) = C_c^\infty(\Omega)$ for test functions, $\mathcal{D}'(\Omega)$ for distributions, $\mathcal{S}(\mathbb{R}^n)$ for Schwartz functions, and $\mathcal{S}'(\mathbb{R}^n)$ for tempered distributions. Pairings are denoted by $\langle \Lambda, \varphi\rangle$.

> **Problem-set architecture.** The five tiers mirror the logical development of the chapter. *Tier I* treats foundational structure questions; *Tier II* emphasizes localization and compact support; *Tier III* studies convergence phenomena and

the failure of nonlinear operations; *Tier IV* develops singular distributions and multiplication rules; and *Tier V* moves to the tempered setting and Fourier-analytic applications. Tiers I–III form the natural core sequence for a first reading.

Tier I. Order, support, and basic structure theorems

1. **True/false with justification.**
   (a) If Λ' is a compactly supported distribution on $\mathbb{R}$, must Λ be compactly supported?
   (b) Is every compactly supported distribution necessarily of finite order?
   (c) Must the Fourier transform of every compactly supported function in $L^1(\mathbb{R})$ be real analytic?
   (d) Determine the distributional support of $\chi_{\mathbb{Q}}$.
   (e) For $n \in \mathbb{N}$, let δ_n be the Dirac delta at n. Does $\delta_n \to 0$ in the weak-$*$ topology of $C_0(\mathbb{R})$?
   (f) Determine the order of the distribution $\Lambda \in \mathcal{D}'(\mathbb{R})$ defined by
   $$\Lambda(\varphi) = \int_{|x|>1} \log(x)\, \varphi(x)\, dx.$$
2. Suppose f is continuous on $\mathbb{R}^n$ and $\int_{\mathbb{R}^n} f(x)\varphi(x)\, dx = 0$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Show that $f \equiv 0$.
3. Let $\Lambda = \Lambda_f$ where f is continuous on $\mathbb{R}^n$. Show that $\operatorname{supp} \Lambda_f = \operatorname{supp} f$. Does the same statement remain valid for merely locally integrable f?
4. (Solvability of $\psi^{(k)} = \varphi$.) Show that there exists $\psi \in \mathcal{D}(\mathbb{R})$ such that $\varphi = \psi^{(k)}$ if and only if
   $$\int_{\mathbb{R}} p(x)\varphi(x)\, dx = 0$$
   for every polynomial p of degree $\leq k-1$.
5. If $\Lambda \in \mathcal{D}'(\mathbb{R})$ satisfies $\Lambda' = 0$, prove that $\Lambda = \Lambda_c$ for some constant c.
6. Show that every $\varphi \in \mathcal{D}(\mathbb{R})$ can be written as
   $$\varphi = \psi' + c\,\varphi_0,$$

where $\varphi_0 \in \mathcal{D}(\mathbb{R})$ is fixed with $\int_{\mathbb{R}} \varphi_0 \neq 0$.

7. Show that every $\varphi \in \mathcal{D}(\mathbb{R})$ can be written as

$$\varphi = x\psi + c\,\varphi_0,$$

where $\varphi_0 \in \mathcal{D}(\mathbb{R})$ is fixed with $\varphi_0(0) \neq 0$. Deduce that if $\Lambda \in \mathcal{D}'(\mathbb{R})$ and $x\Lambda = 0$, then $\Lambda = c\,\delta_0$.

8. Determine all $f \in C^\infty(\mathbb{R})$ such that $f\,\delta_0' = 0$.

Tier II. Compact support, finite order, and localization

1. Show that if $\Lambda \in \mathcal{D}'(\mathbb{R})$ is compactly supported, then Λ' is also compactly supported.

2. Verify that

$$\langle \Lambda, \varphi \rangle = \sum_{n=1}^{\infty} \varphi^{(n)}(n)$$

defines a distribution on $\mathbb{R}$. Is Λ compactly supported?

3. (Local-to-global uniqueness.) Let $\Lambda_1, \Lambda_2 \in \mathcal{D}'(\mathbb{R})$ be such that

$$\langle \Lambda_1, \varphi \rangle = 0 \iff \langle \Lambda_2, \varphi \rangle = 0 \qquad \forall \varphi \in \mathcal{D}(\mathbb{R}).$$

Show that $\Lambda_1 = c\,\Lambda_2$ for some constant c.

4. If $\Lambda \in \mathcal{D}'(\mathbb{R})$ has order N, show that $\Lambda = f^{(N+2)}$ in $\mathcal{D}'(\mathbb{R})$ for some continuous function f. If $\Lambda = \delta_0$, describe all possible choices of such f.

5. (Infinite order and non-extendability.) Let $\Omega = (0, \infty)$ and define $\Lambda \in \mathcal{D}'(\Omega)$ by

$$\langle \Lambda, \varphi \rangle = \sum_{n=1}^{\infty} \varphi^{(n)}\left(\tfrac{1}{n}\right), \qquad \varphi \in \mathcal{D}(\Omega).$$

Show that Λ has infinite order on Ω, and prove that Λ cannot be extended to a distribution on $\mathbb{R}$.

6. Let Λ be a distribution on $\mathbb{R}$ such that $x^2\Lambda = 0$. Show that $\Lambda = c\delta_0 + d\delta_0'$ for some constants c, d.

Tier III. Sequences of distributions and nonlinearity pitfalls

1. Let $H = \chi_{(-\infty,0)}$ and let h_n be differentiable functions with $h_n \to H$ in $\mathcal{D}'(\mathbb{R})$. Show that $h_n' \to \delta_0$ in $\mathcal{D}'(\mathbb{R})$. Does the conclusion remain valid if $H = \chi_{(-\infty,0]}$?
2. Let $\Lambda_n \in \mathcal{D}'(\mathbb{R})$ be defined by
$$\langle \Lambda_n, \varphi \rangle = n\Big(\varphi\big(\tfrac{1}{n}\big) - \varphi\big(-\tfrac{1}{n}\big)\Big).$$
Determine $\lim_{n\to\infty} \Lambda_n$ in $\mathcal{D}'(\mathbb{R})$.
3. For $k \in \mathbb{N}$, define $f_k = k\,\chi_{(1/k,\,2/k)}$. Show that $f_k \to \delta_0$ in $\mathcal{D}'(\mathbb{R})$. Moreover, show that although $f_k^2(x) \to 0$ pointwise, the sequence $\{f_k^2\}$ does *not* converge in $\mathcal{D}'(\mathbb{R})$.
4. Let $\{x_k\} \subset \mathbb{R}$ with $|x_k| \to \infty$. Show that $\delta_{x_k} \to 0$ in $\mathcal{D}'(\mathbb{R})$.
5. For $n \in \mathbb{N}$, let $f_n = \chi_{[0,n)}$. Compute f_n' in $\mathcal{D}'(\mathbb{R})$ and determine $\lim_{n\to\infty} f_n'$ in $\mathcal{D}'(\mathbb{R})$.

Tier IV. Singular distributions, principal values, and multiplication rules

1. For $a > 0$, define
$$\langle \Lambda_a, \varphi \rangle = \left(\int_{-\infty}^{-a} + \int_a^{\infty}\right) \frac{\varphi(x)}{|x|}\,dx + \int_{-a}^{a} \frac{\varphi(x) - \varphi(0)}{|x|}\,dx.$$
Show that Λ_a defines a distribution on $\mathcal{D}(\mathbb{R})$. Find $\lim_{a\to 0} \Lambda_a$ in $\mathcal{D}'(\mathbb{R})$ and compute its distributional derivative.
2. Define
$$f(x) = \begin{cases} x^2, & x < 1, \\ x^2 + 2x, & 1 \le x \le 2, \\ 2x, & x \ge 2. \end{cases}$$
Compute the distributional derivative f'.
3. Determine all $f, g \in C^\infty(\mathbb{R})$ such that $f\,\delta_0 + g\,\delta_0' = 0$.
4. If $\Lambda \in \mathcal{D}'(\mathbb{R})$ satisfies $\Lambda^k = 0$ (multiplication in the sense of distributions, whenever defined), prove that Λ must be a polynomial of degree at most $k-1$. (Explain carefully what notion of product you are using and where it is defined.)

Tier V. Tempered distributions and Fourier transform: identities and PDE

1. For $\Lambda \in \mathcal{D}'(\mathbb{R})$, define $G \in \mathcal{D}'(\mathbb{R}^2)$ by

$$\langle G, \varphi \rangle = \int_{\mathbb{R}} \langle \Lambda, \varphi_y \rangle \, dy, \qquad \varphi_y(x) = \varphi(x, y).$$

   Show that G is a distribution on $\mathbb{R}^2$.

2. Define

$$f(t) = \begin{cases} e^{-t}, & t > 0, \\ -e^{t}, & t < 0. \end{cases}$$

   Show that $f'' = 2\delta_0' + f$ in $\mathcal{D}'(\mathbb{R})$. Deduce that

$$\widehat{f}(\xi) = -\frac{2i\xi}{1+\xi^2}$$

   (with the Fourier normalization used in these notes; track constants if you use a different convention).

3. Let $H = \chi_{(-\infty,0)}$. Prove the convolution identities (in $\mathcal{D}'(\mathbb{R})$):

$$\begin{aligned}
&(a) \quad H * \varphi(x) = \int_{-\infty}^{x} \varphi(t)\, dt, \\
&(b) \quad \delta_0' * H = \delta_0, \\
&(c) \quad 1 * \delta_0' = 0, \\
&(d) \quad 1 * (\delta_0' * H) = 1 * \delta_0 = 1, \\
&(e) \quad (1 * \delta_0') * H = 0.
\end{aligned}$$

4. Define $f(x) = e^{x^2}\chi_{[0,1]}(x)$. Compute f' as a distribution and identify the singular part at the endpoints.

5. Define

$$f(x) = \begin{cases} e^{-x}, & x \geq 0, \\ 1, & x < 0. \end{cases}$$

   Show that $(1 - i\xi)\widehat{f} = \widehat{H}$ in $\mathcal{S}'(\mathbb{R})$, where $H = \chi_{(-\infty,0)}$.

6. (Uniqueness from Gaussian testing.) Suppose $f \in L^\infty(\mathbb{R})$ satisfies

$$\int_{\mathbb{R}} f(y) e^{-y^2} e^{2xy}\, dy = 0 \qquad \forall x \in \mathbb{R}.$$

   Prove that $f \equiv 0$.

7. (Classification.) Classify all continuous functions f on $\mathbb{R}$ which define tempered distributions (i.e. $\Lambda_f \in \mathcal{S}'(\mathbb{R})$).

Bibliography

[1] I. H. Dym and H. P. McKean, *Fourier Series and Integrals*, Academic Press, 1985.

[2] G. B. Folland, *Fourier Analysis and Applications*, Brooks/Cole Mathematics Series, 1972.

[3] Y. Katznelson, *An Introduction to Harmonic Analysis*, Dover, New York, 1976.

[4] T. W. Körner, *Fourier Analysis*, Cambridge University Press, 1989.

[5] L. Grafakos, *Classical Fourier Analysis*, 3rd ed., Springer, 2014.

[6] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, 2nd ed., Springer, 1990.

[7] E. M. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.

[8] R. S. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, CRC Press, 1994.

[9] W. Rudin, *Functional Analysis*, 2nd ed., McGraw-Hill, 1991.