Chapter 4

Fourier Analysis

4.0.1 Intuition

This discussion is borrowed from [T. Tao, Fourier Transform].

Fourier transform/series can be viewed as a way to decompose a function from some given space V into a superposition of "symmetric" functions (whatever that means). To define the "symmetry", we will have a certain group G action of some maps $V \to V$ and a function f from V will be viewed as "symmetric" if action of any map from G on f preserves f up to a multiplicative factor (character) of modulus 1.

Some examples of this:

Examples 4.1. (i) (even–odd decomposition) V is the space of real-valued functions $f : \mathbb{R} \to \mathbb{R}$ with the group G of maps $f(x) \mapsto f(x)$ and $f(x) \mapsto f(-x)$ (isomorphic to $\mathbb{Z}_2$). "Symmetric" functions are then of two types: even or odd. The "Fourier" decomposition of a function $f : \mathbb{R} \to \mathbb{R}$ then is $f = f_e + f_o$, where $f_e = \frac{f(x)+f(-x)}{2}$ (even part), $f_o = \frac{f(x)-f(-x)}{2}$ (odd part).

(ii) (decomposition with respect to rotations by roots of unity) Fix $n \in \mathbb{N}$. Let V be the space of complex-valued functions $f : \mathbb{C} \to \mathbb{C}$ with the group G of maps $f(z) \mapsto f(e^{2\pi ij/n}z)$ for $j = 0, \dots, n-1$ (isomorphic to $\mathbb{Z}_n$). "Symmetric" functions are then of n types: each type is completely determined by action of the generating element of G: $f(e^{2\pi i/n}z) = e^{2\pi im/n}f(z)$ for some $m = 0, 1, \dots, n-1$ (action of the other elements are then uniquely determined). The "Fourier" decomposition of a function $f : \mathbb{C} \to \mathbb{C}$ then is

$$f = \sum_{m=0}^{n-1} f_m,$$

where

$$f_m(z) := \frac{1}{n}\sum_{k=0}^{n-1} f(e^{2\pi ik/n}z)e^{-2\pi imk/n}.$$

Note that if $n = 2$, we essentially get example (i) (but for complex functions).

(iii) (Fourier series) Now let V consists of (smooth) functions $f : \partial\mathbb{D} \to \mathbb{C}$ on the unit circle $\partial\mathbb{D} := \{z \in \mathbb{C} : |z| = 1\}$. Let G consists of (rotations) maps $f(z) \mapsto f(e^{i\phi}z)$ for some choice of $e^{i\phi} \in \partial\mathbb{D}$ (so that G is $\partial\mathbb{D}$ viewed as a group with the usual multiplication operation). Observe that a function $f_n(z) := c_n z^n$ (with some $n \in \mathbb{Z}$) satisfies $f(e^{i\phi}z) = e^{in\theta}f(z)$, and it can be shown that these are the only examples of "symmetric" functions. Then the "Fourier decomposition" takes any smooth f and decomposes it into

$$f(z) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)z^n,$$

where

$$\widehat{f}(n) := \int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\,\tfrac{d\theta}{2\pi}.$$

Note that this expression can be viewed as $n \to \infty$ limit of example (ii).

Note that we start with $G = \partial\mathbb{D}$ and end up with another group $\mathbb{Z}$ which indexes the "symmetry" types. This is known as the Pontryagin dual group of $\partial\mathbb{D}$.

This example can be generalized to Fourier series for functions $f : \partial\mathbb{D}^n \to \mathbb{C}$.

(iv) (Fourier transform) Finally, let us consider V to consist of (smooth) functions $f : \mathbb{R} \to \mathbb{C}$, and let the group G consist of (translations) maps $f(x) \mapsto f(x+t)$ for some choice of $t \in \mathbb{R}$ (so that G is $\mathbb{R}$ viewed as a group with the usual addition operation). Now we have more than countably many "symmetric" functions: indeed, note that for an arbitrary choice of $\xi \in \mathbb{R}$, function $f_\xi(x) = c_\xi e^{2\pi i x\xi}$ satisfies the "symmetry" $f(x+t) = e^{2\pi i t\xi} f(x)$, and it can be shown that these are the only examples of "symmetric" functions (see [F, 8.19]). Then the "Fourier decomposition" takes any (smooth) f and decomposes it into

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

where

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

Note that the "superposition" of "symmetric" components is interpreted now through an integral rather than through a sum.

Note that here $G = \mathbb{R}$, whose Pontryagin dual is again $\mathbb{R}$.

This example can be generalized to Fourier series for functions $f : \mathbb{R}^n \to \mathbb{C}$.

4.0.2 Preliminaries: some notation

Let us set some basic notation:

- Dot product and norm: for $x, y \in \mathbb{R}^n$, let $x \cdot y = \sum_{j=1}^n x_j y_j$ and $|x| = \sqrt{x \cdot x}$.
- Partial derivatives: $\partial_j^m f := \frac{\partial^m f}{\partial x_j^m}$.
- For a multi-index (or, an ordered n-tuple of nonnegative integers) $\alpha = (\alpha_1, \ldots, \alpha_n)$, we define

$$|\alpha| = \alpha_1 + \ldots + \alpha_n,$$
$$\partial^\alpha := \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n},$$
$$\alpha! := \prod_{j=1}^n \alpha_j!,$$
$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}$$

  ($|\cdot|$ is confusing but it should be clear from the context if the argument is an index or a vector).

- C^k denotes the space of all functions whose all derivatives of total order k are continuous. C^∞ denotes the space of all infinitely differentiable functions. C_c^∞ denotes C^∞ functions with compact support.

- Throughout this chapter, whenever we write $L^p(X)$ (for various choices of the space X), we always assume the Lebesgue measure on X. The Lebesgue integration will be denoted with dx (where x will typically be in $\mathbb{R}^n$).
- Let us denote $\widetilde{f}(x) := f(-x)$.
- Given $y \in \mathbb{R}^n$, let us denote $\tau_y(f)(x) := f(x-y)$.

4.0.3 Preliminaries: the n-torus $\mathbb{T}^n$ and functions on $\mathbb{T}^n$

Define the torus $\mathbb{T}$ as $[0,1]$ with endpoints identified (i.e., $\mathbb{R}/\mathbb{Z}$).

Let us regard functions on $\mathbb{T}$ as 1-*periodic* functions on the real line: $f : \mathbb{R} \to \mathbb{C}$ with $f(x) = f(x+1)$.

It is sometimes useful to identify $\mathbb{T}$ with the unit circle $\partial\mathbb{D}$ via the map from $[0,1]$ to $\{z \in \mathbb{C} : |z| = 1\}$ given by $x \mapsto e^{2\pi i x}$. This sets up the correspondence $g(x) = f(e^{2\pi i x})$ between functions $f(z) : \partial\mathbb{D} \to \mathbb{C}$ and functions $g(x) : \mathbb{T} \to \mathbb{C}$.

Similarly, we will view $\mathbb{T}^n$ (the "n-torus") as the cube $[0,1]^n$ with opposite sides identified, i.e., $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

Consequently, functions on $\mathbb{T}^n$ are functions $f : \mathbb{R}^n \to \mathbb{C}$ that are 1-periodic in each coordinate: $f(x+m) = f(x)$ for all $x \in \mathbb{R}^n$ and all $m \in \mathbb{Z}^n$.

Analogously, it is sometimes useful to identify $\mathbb{T}^n$ with $\partial\mathbb{D}^n \subset \mathbb{C}^n$ via the map from $[0,1]^n$ to $\partial\mathbb{D}^n$ given by $(x_1, \dots, x_n) \mapsto (e^{2\pi i x_1}, \dots, e^{2\pi i x_n})$.

4.0.4 Preliminaries: convolutions

Definition 4.2. *(i) For measurable functions $f, g : \mathbb{R}^n \to \mathbb{C}$ we define their* ***convolution***

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x-y) g(y)\, dy$$

for all x where this integral exists.

(ii) Convolution for functions $f, g : \mathbb{T}^n \to \mathbb{C}$ are defined analogously, but now the integral is over $\mathbb{T}^n$ and we treat $x - y$ as the subtraction in $\mathbb{R}^n/\mathbb{Z}^n$ (i.e., modulo 1 in each coordinate).

Remark 4.3. It is elementary to see that $f * g = g * f$.

Theorem 4.4 (Young's Inequality for convolutions)**.** *Suppose $1 \le p, q, r \le \infty$ and*

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

Suppose $f \in L^p$ and $g \in L^q$. Then $f * g \in L^r$, and

$$||f * g||_r \le ||f||_p \, ||g||_q. \tag{4.0.1}$$

Remarks 4.5. (a) This works both in $\mathbb{R}^n$ and in $\mathbb{T}^n$.

(b) The most common special cases of Young's inequality are:

$$||f * g||_\infty \le ||f||_p \, ||g||_q \qquad \text{if } \frac{1}{p} + \frac{1}{q} = 1, \tag{4.0.2}$$

and

$$||f * g||_p \le ||f||_p \, ||g||_1 \qquad \text{if } 1 \le p \le \infty. \tag{4.0.3}$$

Proof. This can be proven via a careful application of generalized Hölder's inequality. Alternatively:

Firstly, for $p = q = 1$ (so that $r = 1$), the inequality follows from the triangle inequality and linear change of variables:

$$||f * g||_1 = \int \left| \int f(x-y)g(y)\, dy \right| dx \leq \int \int |f(x-y)|\, |g(y)|\, dy\, dx = ||f||_1\, ||g||_1. \tag{4.0.4}$$

Secondly, for $r = \infty$ and any conjugate p, q (i.e., $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p \leq \infty$), we have:

$$||f * g||_\infty = \left|\left| \int f(x-y)g(y)\, dy \right|\right| \leq ||f||_p\, ||g||_q \tag{4.0.5}$$

by the usual Hölder's inequality.

Now for any p, q, r with $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, let

Now let us interpolate these two results using two applications of the Riesz–Thorin theorem 3.35: first we interpolate between $L^1 \to L^1$ and $L^\infty \to L^\infty$ to obtain the $L^p \to L^p$ result, and then we interpolate between $L^1 \to L^p$ and $L^{p/(p-1)} \to L^\infty$ to get the general $L^q \to L^r$ result. Here are the details:

First, the convolution operator $T(f) := f * g$ is linear and satisfies $||T(f)||_1 \leq ||g||_1\, ||f||_1$ when viewed as in $L^1 \to L^1$; and $||T(f)||_\infty \leq ||g||_1\, ||f||_\infty$ when viewed as in $L^\infty \to L^\infty$. Then the Riesz–Thorin theorem says that $||T(f)||_s \leq ||g||_1\, ||f||_s$ for any $s \in [1, \infty]$ viewed as in $L^s \to L^s$.

Now, as we just proved, the convolution operator $S(g) := f * g$ is linear and satisfies $||T(g)||_p \leq ||f||_p\, ||g||_1$ when viewed as in $L^1 \to L^p$; also $||T(g)||_\infty \leq ||f||_p\, ||g||_{p/(p-1)}$ when viewed as in $L^{p/(p-1)} \to L^\infty$. Then the Riesz–Thorin theorem says that $||T(g)||_r \leq ||f||_p\, ||g||_s$ viewed as in $L^s \to L^r$, where $\frac{1}{s} = \frac{1-t}{1} + \frac{t}{p/(p-1)}$, $\frac{1}{r} = \frac{1-t}{p} + \frac{t}{\infty}$. Getting rid of t, we get $\frac{1}{s} + \frac{1}{p} = \frac{1}{r} + 1$, which is what we need. □

Proposition 4.6. *(i) If $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$ and $f \in L^p(\mathbb{T}^n)$, $g \in L^q(\mathbb{T}^n)$, then $f * g \in C(\mathbb{T}^n)$.*

(ii) If $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$ and $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, then $f * g \in C_0(\mathbb{R}^n)$.

Remark 4.7. If $p = 1$ or $p = \infty$ continuity in fact still holds [F241].

Proof. (i) and (ii): If $f, g \in C_c$, then the statement is clear. Now for any $f \in L^p$, $g \in L^q$, choose $f_n, g_n \in C_c$ such that $f_n \to f$ in L^p and $g_n \to g$ in L^q (see Theorem 3.24). By the Young's inequality,

$$||f * g - f_n * g_n||_\infty \leq ||f * (g - g_n)||_\infty + ||(f - f_n) * g_n||_\infty \leq ||f||_p\, ||g - g_n||_q + ||f - f_n||_p\, ||g_n||_q \to 0.$$

So $f * g$ is a uniform limit of continuous functions with compact support, so $f * g \in C(\mathbb{T}^n)$ or $f * g \in C_0(\mathbb{T}^n)$, respectively. □

4.0.5 Preliminaries: approximate identity

Definition 4.8. *An **approximate identity** on $\mathbb{T}^n$ is a sequence of functions $\{k_j(x)\}_{j=1}^\infty$ on $\mathbb{T}^n$ that satisfy*

(i) $k_j(x) \geq 0$;

(ii) $\int_{\mathbb{T}^n} k_j(x)\, dx = 1$ for all j;

(iii) For any $\varepsilon > 0$,

$$\lim_{j \to \infty} \int_{|x| > \varepsilon} |k_j(x)|\, dx = 0.$$

Remarks 4.9. (a) $|x| > \varepsilon$ is meant in the periodic sense, e.g., in $\mathbb{T}^1$ it means that $x \in (\varepsilon, 1-\varepsilon)$.

(b) One can relax the requirement $k_j \geq 0$ but then instead needs to require $||k_j||_1 \leq C$ for some choice of $C > 0$.

(c) Similarly, one defines an approximate identity on $\mathbb{R}^n$: just replace each $\mathbb{T}^n$ with $\mathbb{R}^n$.

(d) For an example of an approximate identity, we can set $k_j(x) = j^{-n}\phi(x/j)$ for any $\phi \in L^+ \cap L^1(\mathbb{T}^n)$ (or on $\mathbb{R}^n$).

Proposition 4.10. *Let $\{k_j\}_{j=1}^{\infty}$ be an approximate identity on $\mathbb{T}^n$ or on $\mathbb{R}^n$. Then:*

(i) If $f \in C(\mathbb{T}^n)$ then

$$||k_j * f - f||_\infty \to 0.$$

The same statement holds for $\mathbb{R}^n$ with the only change that in f needs to be bounded and uniformly continuous on $\mathbb{R}^n$.

(ii) If $f \in L^p(\mathbb{T}^n)$ for $1 \leq p < \infty$, then

$$||k_j * f - f||_p \to 0.$$

The same statement holds for $f \in L^p(\mathbb{R}^n)$.

Proof. Let us prove only the $\mathbb{T}^1$ case – cases $\mathbb{T}^n$ and $\mathbb{R}^n$ follow along the same lines.

(i) By periodicity and Definition 4.8(ii), we have

$$(k_j * f)(x) - f(x) = \int_0^1 k_j(y)(f(x-y) - f(x))\,dy.$$

Now divide the integral into regions $y \in (\varepsilon, 1-\varepsilon)$ and $|y| \leq \varepsilon$:

$$||k_j * f - f||_\infty \leq \sup_{|y|\leq\varepsilon, x\in[0,1]} |f(x-y) - f(x)| + 2||f||_\infty \int_{|y|>\varepsilon} |k_j(y)|\,dy,$$

so we get $\limsup ||k_j * f - f||_\infty \leq \sup_{|y|\leq\varepsilon, x\in[0,1]} |f(x-y) - f(x)|$. But the latter expression goes to 0 as $\varepsilon \to 0$ since any continuous function on $[0,1]$ is uniformly continuous.

(ii) The statement for f continuous with compact support follow from (i) since $||k_j * f - f||_p \leq ||k_j * f - f||_\infty$ on $\mathbb{T}^n$ (for $\mathbb{R}^n$ one should make use of compact support and property (iii) of Definition 4.8). Now for arbitrary $f \in L^p$, find $g \in C_c$ such that $||f-g||_p < \varepsilon$. Then using Young's inequality (4.0.3), we get:

$$\begin{aligned}\limsup_{j\to\infty} ||k_j * f - f||_p &\leq \limsup_{j\to\infty} \left(||k_j * (f-g)||_p + ||k_j * g - g||_p + ||f-g||_p\right) \\ &\leq \limsup_{j\to\infty} ||k_j||_1\, ||f-g||_p + 0 + ||f-g||_p < 2\varepsilon.\end{aligned}$$

Then we take $\varepsilon \to 0$. □

4.1 Fourier Analysis on $\mathbb{T}^n$

4.1.1 Fourier series on $L^2(\mathbb{T}^n)$: definition

Recall that if H is a Hilbert space and $\{\phi_j\}_{j=1}^{\infty}$ is an orthonormal basis, then for any $x \in H$, we have

$$x = \sum_{j=1}^{\infty} \langle x, \phi_j \rangle \phi_j,$$

where the infinite sum is meant as the converging in the norm of H (that is, $||x - \sum_{j=1}^{n} \langle x, \phi_j \rangle \phi_j|| \to 0$ as $n \to \infty$). Coefficients $\langle x, \phi_j \rangle$ above are called the (abstract) Fourier coefficients.

Now we consider $L^2(\mathbb{T}^n)$ with inner product $\langle f, g \rangle := \int_{\mathbb{T}^n} f(x)\overline{g(x)}\,dx$. Note that the sequence $\{e^{2\pi i m\cdot x}\}_{m\in\mathbb{Z}^n}$ is orthonormal (basic application of the Fubini theorem):

$$\int_{[0,1]^n} e^{2\pi i m\cdot x}\overline{e^{2\pi i k\cdot x}}\,dx = \begin{cases} 1, & \text{if } m = k, \\ 0, & \text{if } m \neq k. \end{cases}$$

In fact, $\{e^{2\pi i m\cdot x}\}_{m\in\mathbb{Z}^n}$ is an orthonormal basis of $L^2(\mathbb{T}^n)$ (it is proven below in Theorem 4.26).

Definition 4.11. *For $f \in L^2(\mathbb{T}^n)$:*

*(i) Define its **Fourier coefficients** to be*

$$\widehat{f}(m) := \int_{\mathbb{T}^n} f(x)e^{-2\pi i m\cdot x}\,dx, \qquad m \in \mathbb{Z}^n. \tag{4.1.1}$$

*(ii) Define its **Fourier series** to be*

$$\sum_{m\in\mathbb{Z}^m} \widehat{f}(m)e^{2\pi i m\cdot x}. \tag{4.1.2}$$

As a corollary of the general Hilbert space theory, we immediately get:

Corollary 4.12. *For any $f, g \in L^2(\mathbb{T}^n)$:*

(i) $f(t)$ is the $L^2(\mathbb{T}^n)$-limit of the sequence of its partial Fourier sums $S_N(f) := \sum_{|m|\le N} \widehat{f}(m)e^{2\pi i m\cdot t}$:

$$||S_N(f) - f||_2 \to 0.$$

(ii) (Plancherel's identity)

$$||f||_2^2 = \sum_{m\in\mathbb{Z}^n} |\widehat{f}(m)|^2$$

(iii) (Parseval's relation)

$$\int_{\mathbb{T}^n} f(t)\overline{g(t)}\,dt = \sum_{m\in\mathbb{Z}^n} \widehat{f}(m)\overline{\widehat{g}(m)}.$$

(iv) The map $f \mapsto \{\widehat{f}(m)\}_{m\in\mathbb{Z}^n}$ is an isometry from $L^2(\mathbb{T}^n)$ onto $\ell^2(\mathbb{Z}^n)$.

Proof. (i) and (ii) are standard facts about orthonormal bases. (iii) follows from (ii) and the polarization identity

$$\langle f, g \rangle = \tfrac{1}{4}\left(||x+y||^2 - ||x-y||^2 + i||x-iy||^2 - i||x+iy||^2\right).$$

(iv) is then almost proved except for surjectivity: given $\{a_m\}_{m\in\mathbb{Z}^n}$ we can form partial Fourier sums $f_M(t) = \sum_{|m|\le M} a_m e^{2\pi i m\cdot t}$ and show that it's Cauchy in $L^2(\mathbb{T}^n)$, therefore converging to some $f \in L^2(\mathbb{T}^n)$. Then Fourier's coefficients of f are a_m's and we are done. □

4.1.2 Fourier series on $L^1(\mathbb{T}^n)$: definition and basic properties

Note that the definition (4.1.1) and (4.1.2) work under the mere assumption that $f \in L^1(\mathbb{T}^n)$. However the convergence of (4.1.2) is then no longer clear. Let us state some of the properties of the Fourier series on L^1.

Recall the notation $\widetilde{f}(x) := f(-x)$ and $\tau_y(f)(x) := f(x-y)$.

Proposition 4.13. *Let f, g be in $L^1(\mathbb{T}^n)$. Then for all $m, k \in \mathbb{Z}^n$, $\lambda \in \mathbb{C}$, $y \in \mathbb{T}^n$:*

(i) $\widehat{f+g}(m) = \widehat{f}(m) + \widehat{g}(m)$;

(ii) $\widehat{\lambda f}(m) = \lambda \widehat{f}(m)$;

(iii) $\widehat{\overline{f}}(m) = \overline{\widehat{f}(-m)}$;

(iv) $\widehat{\widetilde{f}}(m) = \widehat{f}(-m)$;

(v) $\widehat{\tau_y(f)}(m) = \widehat{f}(m) e^{-2\pi i m \cdot y}$;

(vi) $\left(e^{2\pi i k \cdot x} f(x)\right)^{\widehat{}}(m) = \widehat{f}(m-k)$;

(vii) $\sup_{m \in \mathbb{Z}^n} |\widehat{f}(m)| \leq ||f||_1$, *in particular, if $f \in L^1(\mathbb{T}^n)$ then $\widehat{f} \in \ell^\infty(\mathbb{Z}^n)$;*

(viii) $\widehat{f * g}(m) = \widehat{f}(m)\widehat{g}(m)$;

(ix) $\widehat{\partial^\alpha f}(m) = (2\pi i m)^\alpha \widehat{f}(m)$ *for $f \in C^{|\alpha|}$.*

Proof. To prove (viii), we use Tonelli+Fubini theorems to interchange the integrals:

$$\widehat{f * g}(m) = \int_{\mathbb{T}^n} \int_{\mathbb{T}^n} f(x-y) g(y) e^{-2\pi i m \cdot (x-y)} e^{-2\pi i m \cdot y} \, dy \, dx = \widehat{f}(m)\widehat{g}(m).$$

□

4.1.3 Fourier series on $L^1(\mathbb{T}^n)$: various types of pointwise convergence

Intuition. Recall that an infinite series $\sum_{j=1}^\infty a_j$ is called *summable* if its partial sums $s_n := \sum_{j=1}^n a_j$ has a finite limit, and *Cesàro summable* if the arithmetic mean of its partial sum has a finite limit: $\frac{1}{n}\sum_{j=1}^n s_j$. It is elementary to show that if a series is summable then it is Cesàro summable (with the same limit), but not necessarily vice versa.

We want to discuss the issue of pointwise convergence of Fourier series, and Cesàro summation is natural here.

Definition 4.14. *Let $f \in L^1(\mathbb{T}^n)$ and $N \in \mathbb{N}$.*

(i) The expression

$$S_N(f)(x) := \sum_{\substack{m \in \mathbb{Z}^n \\ |m_j| \leq N}} \hat{f}(m) e^{2\pi i m \cdot x} \tag{4.1.3}$$

is called the ***square partial sum of the Fourier series of*** *f.*

(ii) Let $S_{k_1,\dots,k_n}(f)(x) := \sum_{\substack{m\in\mathbb{Z}^n \\ |m_j|\le k_j}} \hat{f}(m)e^{2\pi i m\cdot x}$ *(the "rectangular partial sum of Fourier series"). The expression*

$$C_N(f)(x) := \frac{1}{(N+1)^n} \sum_{\substack{m\in\mathbb{Z}^n \\ 0\le m_j\le N}} S_{m_1,\dots,m_n}(f)(x)$$

is called the ***square Cesàro mean of the Fourier series of*** f.

Remarks 4.15. (i) One can also form circular (spherical) partial sums and circular (spherical) Cesàro means: now one would sum over $m \in \mathbb{Z}^n, \sum |m_j|^2 \le M$ instead of $m \in \mathbb{Z}^n, |m_j| \le M$.

(ii) Yet another common way to sum the Fourier series is via the Abel–Poisson summation (convolution with a Poisson kernel). There're also notions of Riesz and Bochner–Riesz summations.

4.1.4 Fourier series on $L^1(\mathbb{T}^n)$: Dirichlet and Fejér kernels

Intuition. One of the themes in the discussion that follows will be that partial Fourier sums $S_M(f)$ do not behave as well as their Cesàro means $C_M(f)$ do. The main explanation for this is that the Fejér kernels form an approximate identity which we explore in this section.

Definition 4.16. *Let* $N \in \mathbb{N}$.

(i) On $\mathbb{T}^1$*: the* ***Dirichlet kernel*** *on* $\mathbb{T}^1$ *is defined to be*

$$D_N(x) = \sum_{m=-N}^{N} e^{2\pi i m x}.$$

(ii) On $\mathbb{T}^1$*: the* ***Fejér kernel*** *on* $\mathbb{T}^1$ *is defined to be*

$$F_N(x) = \frac{1}{N+1}\sum_{j=0}^{N} D_j(x)$$

(iii) On $\mathbb{T}^n$*: the* ***square Dirichlet kernel*** *on* $\mathbb{T}^n$ *is defined to be*

$$D(n,N)(x) = \sum_{\substack{m\in\mathbb{Z}^n \\ |m_j|\le N}} e^{2\pi i m\cdot x} = D_N(x_1)\dots D_N(x_n).$$

(iv) On $\mathbb{T}^n$*: the* ***square Fejér kernel*** *on* $\mathbb{T}^n$ *is defined to be*

$$F(n,N)(x) = \frac{1}{(N+1)^n}\sum_{k_1=0}^{N}\dots\sum_{k_n=0}^{N} D_{k_1}(x_1)\cdots D_{k_n}(x_n) = F_N(x_1)\cdots F_N(x_n)$$

Proposition 4.17. *(i) On* $\mathbb{T}^1$*:*

$$D_N(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)}$$

for $x \in [0,1]$.

(ii) On $\mathbb{T}^1$:

$$F_N(x) = \sum_{j=-N}^{N} \left(1 - \frac{|j|}{N+1}\right) e^{2\pi ijx} = \frac{1}{N+1} \left(\frac{\sin(\pi(N+1)x)}{\sin(\pi x)}\right)^2.$$

Remark 4.18. Combining (ii) with Definition 4.16(iv), we get

$$F(n,N)(x) = \sum_{\substack{m\in\mathbb{Z}^n \\ |m_j|\le N}} \left(1 - \frac{|m_1|}{N+1}\right) \cdots \left(1 - \frac{|m_n|}{N+1}\right) e^{2\pi i m\cdot x}. \tag{4.1.4}$$

Proof. (i) follows from from the formula for the geometric series.

(ii) The first expression is immediate from the definition. The second uses (i) and a bit more of trigonometry. □

Proposition 4.19. *Let $f \in L^1(\mathbb{T}^n)$ and $M \in \mathbb{N}$.*

(i) Square partial sum of the Fourier series of f is a convolution with the square Dirichlet kernel $D(n,N)$:

$$S_N(f)(x) = (f * D(n,N))(x).$$

(ii) Square Cesàro mean of the Fourier series of f is a convolution with the square Fejér kernel $F(n,N)$:

$$C_N(f)(x) = (f * F(n,N))(x) = \sum_{\substack{m\in\mathbb{Z}^n \\ |m_j|\le N}} \hat{f}(m) \left(1 - \frac{|m_1|}{N+1}\right) \cdots \left(1 - \frac{|m_n|}{N+1}\right) e^{2\pi i m\cdot x}. \tag{4.1.5}$$

Proof. (i) follows directly from the definitions of convolution and of D_N. Note that by Proposition 4.13(viii) and Definition 4.16 we can immediately see that the Fourier series of left-hand side and right-hand side coincide, but we do not know the uniqueness result yet (see Proposition 4.27 below).

(ii) follows from in the exact same way by using also (4.1.4). □

Proposition 4.20. *Fejér kernel $\{F(n,N)\}_{N=1}^{\infty}$ forms an approximate identity on $\mathbb{T}^n$.*

Remark 4.21. Dirichlet kernel, on the other hand, has $||D_N||_1 \sim \log n \to \infty$.

Proof. For $\mathbb{T}^1$ this is an easy check using Proposition 4.17(ii). For $\mathbb{T}^n$ this follows if one uses the fact that $F(n,N)(x)$ is the product of $F(1,N)(x_j)$. □

4.1.5 Fourier series on $L^1(\mathbb{T}^n)$: uniform and L^p convergence of $C_N(f)$

Theorem 4.22. *(i) (Fejér's Theorem) For any $f \in C(\mathbb{T}^n)$, $C_N(f) \to f$ in the $||\cdot||_\infty$ norm (i.e., uniformly on $\mathbb{T}^n$).*

(ii) For any $f \in L^p(\mathbb{T}^n)$ $(1 \le p < \infty)$, $C_N(f) \to f$ in the L^p norm.

Remarks 4.23. (a) Note that continuity is necessary for (i) to hold as uniform limit of continuous functions is continuous.

(b) Part (i) for the usual partial sums $S_N(f)$ fails: there exist continuous functions whose Fourier series do not converge at infinitely many points (however, see Theorem 4.39 below). So for the uniform convergence of $S_N(f)$ one needs some stronger condition than mere continuity – e.g., differentiability or Hölder continuity suffices for $\mathbb{T}^1$, see Theorem 4.33.

(c) Part (ii) for the usual partial sums $S_N(f)$ works only for $1 < p < \infty$, see Theorem 4.35.

Proof. Both (i) and (ii) follow immediately from the fact that Fejér kernel forms an approximate identity and combining it with Proposition 4.10. □

4.1.6 Fourier series on $L^1(\mathbb{T}^n)$: some corollaries of Fejér

Theorem 4.24 (Weierstrass approximation theorem on the circle/torus)**.** *The set of trigonometric polynomials (i.e., functions $\mathbb{T}^n \to \mathbb{C}$ of the form*

$$\sum_{\substack{m\in\mathbb{Z}^n \\ |m|\le N}} c_m e^{2\pi i m\cdot x}$$

with $N \in \mathbb{N}$ and $c_j \in \mathbb{C}$) is $||\cdot||_\infty$-dense in $C(\mathbb{T}^n)$.

Remarks 4.25. (a) If we chose different scaling, then this theorem in dimension 1 says that polynomials in $\pm e^{i\theta}$ (sometimes called Laurent polynomials) are $||\cdot||_\infty$-dense in $C(\partial\mathbb{D})$.

(b) If one takes real parts, then one obtains yet another restatement of the same theorem: function of the type $\sum_{j=1}^N a_j\cos(j\theta) + b_j\sin(j\theta)$ are $||\cdot||_\infty$-dense in the space of 2π-periodic functions on $\mathbb{R}$.

(c) There is also the Weierstrass approximation theorem on the real line which states that polynomials in x are $||\cdot||_\infty$-dense in $C([a,b])$ for any bounded interval $[a,b]\subset\mathbb{R}$.

(d) All of these Weierstrass theorems are in fact special cases of the much more general Stone–Weierstrass theorem which states that if X is compact Hausdorff, then any sub-algebra of $C_{\mathbb{R}}(X)$ (index $\mathbb{R}$ here means real-valued) that contains the function $f(x)\equiv 1$ and separates points is necessarily $||\cdot||_\infty$-dense in $C_R(X)$.

Proof. Given any $f \in C(\mathbb{T}^n)$, we know (Theorem 4.22(i)) that sequence of Cesàro means converges to f uniformly. But each Cesàro mean is a trigonometric polynomial, so we are done. □

Now we are able to prove the following theorem which we stated without a proof before:

Theorem 4.26. *(i) For any $1 \le p < \infty$, the set of trigonometric polynomials is $||\cdot||_p$-dense in $L^p(\mathbb{T}^n)$.*

(ii) Moreover, $\{e^{2\pi i m\cdot x}\}_{m\in\mathbb{Z}^n}$ is an orthonormal basis of $L^2(\mathbb{T}^n)$.

Proof. For (i), just use Theorem 4.22(ii) as in the previous proof.

(ii) then follows as orthonormality (follows from direct computation along with Tonelli+Fubini theorems) and density is enough. □

Proposition 4.27 (Uniqueness of Fourier coefficients)**.** *Let $f, g \in L^1(\mathbb{T}^n)$ satisfy $\widehat{f}(m) = \widehat{g}(m)$ for all $m \in \mathbb{Z}^n$. Then $f = g$ a.e.*

Remark 4.28. For $f, g \in L^2(\mathbb{T}^n)$ this follows from the previous theorem. However for L^1 we need a proof.

Proof. By linearity we can assume $g = 0$, so that $\hat{f}(m) = \hat{g}(m) = 0$ for all $m \in \mathbb{Z}^n$. From the definition of the Cesàro means (Definition 4.14(ii)), $C_N(f)(x) = 0$ for all N. But we know that $||f - C_N(f)||_1 \to 0$ by Theorem 4.22, which implies $||f||_1 = 0$, i.e., $f = 0$ a.e. □

Theorem 4.29 (Fourier inversion)**.** *Suppose $f \in L^1(\mathbb{T}^n)$ and $\hat{f} \in \ell^1(\mathbb{Z}^n)$ (that is, $\sum|\widehat{f}(m)| < \infty$). Then $S_N(f)$ converges uniformly (and absolutely) to a continuous function that is equal to f a.e.*

Remarks 4.30. (a) While the conclusion of the theorem is very nice, it is very unsatisfying as there are conditions on both f and $\widehat{f}$.

(b) Compare this Theorem with Theorem 4.58 for Fourier inversion on $\mathbb{R}^n$.

Proof. Uniform convergence of $S_N(f)$ is clear from the ℓ^1 condition and elementary convergence test. The limit function g must be continuous. It is easy to see that both g and f have the same Fourier coefficients. By Proposition 4.27, we get $f = g$ a.e. □

Theorem 4.31 (Riemann–Lebesgue lemma)**.** *Let $f \in L^1(\mathbb{T}^n)$. Then $|\widehat{f}(m)| \to 0$ as $|m| \to \infty$.*

Remarks 4.32. (a) The converse to this is not true: not every sequence satisfying $|a_m| \to 0$ is the Fourier sequence of some $f \in L^1$.

(b) On the other hand, the following can be proved (see [G, Thm 3.2.2]): given any sequence $|d_m| \to 0$, there exists $f \in L^1(\mathbb{T}^n)$ whose Fourier coefficients have slower rate of decay than $|d_m|$, i.e., $|\widehat{f}(m)| \geq |d_m|$ for all $m \in \mathbb{Z}^n$.

Proof. Given any $\varepsilon > 0$, find a trigonometric polynomial P such that $||f - P||_1 < \varepsilon$ (use Theorem 4.26(i)). Then for any large enough $m \in \mathbb{Z}^n$ we have $\widehat{P}(m) = 0$, which implies

$$|\widehat{f}(m)| = |\widehat{f}(m) - \widehat{P}(m)| = \left| \int_{\mathbb{T}^n} (f(x) - P(x)) e^{-2\pi i m \cdot x}\, dx \right| \leq ||f - P||_1 \leq \varepsilon.$$

This proves that $|\widehat{f}(m)| \to 0$ as $|m| \to \infty$. □

4.1.7 Fourier series on $L^1(\mathbb{T})$: uniform and L^p convergence of $S_N(f)$

Theorem 4.33. *Suppose $f : \mathbb{T}^1 \to \mathbb{C}$ is Hölder continuous of order $\alpha \in (0, 1]$, which means that there is some $C > 0$ such that for all $x, y \in \mathbb{T}^1$*

$$|f(x) - f(y)| \leq C|x - y|^\alpha.$$

Then $S_N(f) \to f$ in the $|| \cdot ||_\infty$ norm (i.e., uniformly on $\mathbb{T}^1$).

Remarks 4.34. (a) If $\alpha = 1$, such functions are also called Lipschitz continuous.

(b) E.g., functions in $C^1(\mathbb{T})$ are Hölder continuous with $\alpha = 1$.

(c) Note that if we assume Hölder continuity at one point x_0 only, the the proof shows that $S_N(f)(x_0)$ goes $f(x_0)$.

(d) Instead of the Hölder continuity, one could instead impose a weaker "Dini condition" $\int_{\mathbb{T}} \frac{|f(x) - f(x_0)|}{|x - x_0|} dx < \infty$ for the convergence at a point x_0, or a uniform Dini-type condition for the uniform convergence. The proof goes in the exact same manner as below.

(e) There is an analogue of this in $\mathbb{T}^n$ due to Tonelli – see [G, Sect. 3.3.3].

Proof. Let us prove pointwise convergence only, and leave uniform estimates as an exercise. So let us fix $x \in \mathbb{T}$.

Note that the Dirichlet kernel (see Propositions 4.19(i) and 4.17(i)), satisfies $\int_{\mathbb{T}} D_N(y)\, dy = 1$. So we can write

$$|S_N(f)(x) - f(x)| = \left| \int_{\mathbb{T}} D_N(y)(f(x - y) - f(x))\, dy \right| = \left| \int_{\mathbb{T}} \sin((2N + 1)\pi y) h_x(y)\, dy \right|, \qquad (4.1.6)$$

where $|h_x(y)| := \left| \frac{f(x-y) - f(x)}{\sin(\pi y)} \right| \leq \frac{C|y|^\alpha}{\sin(\pi y)} \leq C|y|^{-1+\alpha}$. Note that $h_x \in L^1(\mathbb{T})$. Using $\sin((2N + 1)\pi y) = \frac{1}{2i}(e^{(2N+1)\pi i y} + e^{-(2N+1)\pi i y})$, we can rewrite (4.1.6) as

$$\left| \tfrac{1}{2i} \left((e^{\pi i y} h_x(y))\widehat{\ }(-N) - (e^{\pi i y} h_x(y))\widehat{\ }(N) \right) \right|$$

which goes to 0 by the Riemann–Lebesgue lemma (Theorem 4.31). □

Theorem 4.35. *If $1 < p < \infty$ and $f \in L^p(\mathbb{T}^n)$, then $S_N(f) \to f$ in the $||\cdot||_p$-norm.*

Remarks 4.36. (a) This fails for circular partial Fourier sums if $n \geq 2$ and $p \neq 2$.
(b) This fails for $p = 1$ even in the one-dimensional case $n = 1$!

Proof. For $p = 2$ we already discussed it.
For $p \neq 2$, see, e.g., [G, Sect. 3.5.1] (requires a lot more machinery than we currently have). □

4.1.8 Fourier series on $L^1(\mathbb{T}^1)$: pointwise convergence of $S_N(f)$ and $C_N(f)$

We saw that $C_N(f) \to f$ if $f \in C(\mathbb{T})$ and $S_N(f) \to f$ if $f \in C(\mathbb{T})$ and f is Hölder continuous. We also saw that Hölder or Dini conditions at a point x_0 (both of which are stronger than continuity at x_0) suffice for the convergence of $C_N(f)(x_0)$ to $f(x_0)$. In this section we show that a jump discontinuity (under some extra mild conditions) usually do not spoil the picture – but the limit of $S_N(x)$ or $C_N(x)$ will no longer be $f(x)$.

Theorem 4.37. *(i) Suppose $f \in L^1(\mathbb{T})$ has one sided limits $f(x_0-)$ and $f(x_0+)$. Then*

$$C_N(f)(x_0) \to \frac{1}{2}(f(x_0+) + f(x_0-))$$

as $N \to \infty$.

(ii) Suppose $f : \mathbb{T} \to \mathbb{C}$ is a function of bounded variation (=difference of two bounded monotonically non-decreasing functions), and f has one sided limits $f(x_0-)$ and $f(x_0+)$. Then

$$S_N(f)(x_0) \to \frac{1}{2}(f(x_0+) + f(x_0-))$$

as $N \to \infty$.

Proof. (i) See [G, Sect. 3.3.1] for all the details. By translation, we may assume $x_0 = 0$. We write

$$\begin{aligned} C_N(f)(0) - \frac{1}{2}(f(0+) + f(0-)) &= \int_{-1/2}^{1/2} F_N(0-x)f(x)\,dx - \frac{1}{2}(f(0+) + f(0-)) \\ &= \int_0^{1/2} (f(x) - f(0+))F_N(x)\,dx + \int_{-1/2}^{0} (f(x) - f(0-))F_N(x)\,dx. \end{aligned}$$

We divide $\int_0^{1/2}$ into $\int_0^\delta + \int_\delta^{1/2}$, where $\delta > 0$ is small enough to make $|f(\delta) - f(0+)| < \varepsilon$. Then $\int_0^\delta$ can be bounded by $\varepsilon \int_0^\delta F_N < \varepsilon$, while the second integral can be bounded by $\int_\delta^{1/2}(|f(x)| + |f(0+)|)\,dx \times \sup_{[\delta,1/2]} F_N(x)$ and then showing that $\sup_{[\delta,1/2]} F_N(x) \to 0$ for any $\delta > 0$.

(ii) See [F, Thm 8.43] for all the details. We just sketch some ideas similar to (i). First we may assume that $x_0 = 0$ and f is real-valued, non-decreasing, and right-continuous. Given $\varepsilon > 0$, we write

$$S_N(f)(0) - \frac{1}{2}(f(0+) + f(0-)) = \int_0^{1/2} (f(x) - f(0+))D_N(x)\,dx + \int_{-1/2}^{0} (f(x) - f(0-))D_N(x)\,dx.$$

The difference from the proof of Theorem 4.33 is that at $t = 0$ we no longer have a bound on $|f(x) - f(0)| \leq Cx^\alpha$. However we still have $f(x) - f(0+) \to 0$. The difference with (i) is that D_N doesn't behave as well as F_N. Just like in (i), we split $\int_0^{1/2}$ into $\int_0^\delta + \int_\delta^{1/2}$, where $\delta > 0$ is small enough to make $f(\delta) - f(0+) < \varepsilon$. Then $\int_0^\delta$ can be shown to be comparable with $\varepsilon \int_\eta^\delta D_N(x)\,dx$ for some $\eta \in [0, \delta]$ (important that f is monotone here! otherwise oscillations in D_N could add up and one would get a contribution of order $\int |D_N|\,dx$ which is large); and the second integral is small by the Riemann–Lebesgue lemma as in the proof of Theorem 4.33. □

Note that alongside we proved the following

Corollary 4.38. *Suppose $f_n \to f$ in $||\cdot||_p$-norm (for any $1 \leq p \leq \infty$). Then there exists an subsequence $n_k \subseteq \mathbb{N}$ such that $f_{n_k}(x) \to f(x)$ for almost all x.*

4.1.9 Fourier series on $L^1(\mathbb{T}^1)$: a.e. pointwise convergence of $S_N(f)$ and $C_N(f)$

If we only require convergence a.e., then the conditions can be significantly relaxed:

Theorem 4.39. *(i) If $f \in L^1(\mathbb{T})$ then $C_N(f)(x) \to f(x)$ for a.e. $x \in \mathbb{T}$.*

(ii) (Carleson's theorem) If $f \in L^p(\mathbb{T})$ (for some $1 < p < \infty$), then $S_N(f)(x) \to f(x)$ for a.e. $x \in \mathbb{T}$.

Remarks 4.40. (a) Both parts work for $\mathbb{T}^n$ without a change.
(b) Part (ii) fails for the case $p = 1$ (so, again, C_N behaves better than S_N!).
(c) There are examples of (continuous!) functions where $S_N(f) \to f$ fails on an arbitrary set of Lebesgue measure 0.

Proof. We will omit the proofs. (i) is proved in [G, Thm 3.3.3], (ii) in [G, Thm 3.6.14]. □

4.1.10 Fourier series on $L^p(\mathbb{T}^n)$: summability of coefficients and Hausdorff–Young

Recall that we showed:

- $f \in L^2(\mathbb{T}^n)$ iff $\widehat{f} \in \ell^2(\mathbb{Z}^n)$;
- $f \in L^1(\mathbb{T}^n)$ implies $\widehat{f} \in \ell^\infty(\mathbb{Z}^n)$ (in fact, more is true: $\widehat{f}(m) \to 0$ by the Riemann–Lebesgue lemma).

Theorem 4.41 (Hausdorff–Young Inequality)**.** *Suppose that $1 \leq p \leq 2$, and let $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mathbb{T}^n)$, then $\widehat{f} \in \ell^q(\mathbb{Z}^n)$. Moreover, $||\widehat{f}||_q \leq ||f||_p$.*

Remark 4.42. The optimal constants are in fact $||\widehat{f}||_q \leq p^{1/2p} q^{-1/2q} ||f||_p$.

Proof. Use Riesz–Thorin Theorem 3.35 to interpolate between the above two $L^2 \to \ell^2$ and $L^1 \to \ell^\infty$ results. □

4.1.11 Fourier series on $L^1(\mathbb{T}^n)$: decay of the Fourier coefficients

Intuition. By the Riemann–Lebesgue lemma, $\widehat{f}(m) \to 0$ as $|m| \to \infty$ for any $f \in L^1(\mathbb{T}^n)$. In this section we discuss the speed of decay of $\widehat{f}$. This is closely connected with the smoothness properties of f: the smoother f is, the faster is the decay of $\widehat{f}$ at infinity.
Note that the speed of decay of $\widehat{f}$ is also related to the summability properties ($\widehat{f} \in \ell^q$ as in the previous section).

Theorem 4.43. *Let $s \in \mathbb{N}$.*

(i) Suppose $\partial^\alpha f$ exist and $\partial^\alpha f \in L^1$ for $|\alpha| \leq s$. Then

$$|\widehat{f}(m)| \leq \left(\frac{\sqrt{n}}{2\pi}\right)^s \frac{\sup_{|\alpha|=s} \left|\widehat{\partial^\alpha f}(m)\right|}{|m|^s}, \tag{4.1.7}$$

in particular $|\widehat{f}(m)|$ goes to 0 as $|m| \to \infty$ faster than $\frac{1}{|m|^s}$.

(ii) Suppose that

$$|\widehat{f}(m)| \leq C \frac{1}{1+|m|^{s+n}} \tag{4.1.8}$$

for some $C > 0$ and all $m \in \mathbb{Z}^n$. Then $\partial^\alpha f$ exist for all $|\alpha| \leq s-1$.

Proof. (i) Given $m \in \mathbb{Z}^n$, choose j such that $|m_j| = \max_{1\leq k\leq n} |m_k|$. Note that $|m_j| \geq \frac{1}{\sqrt{n}}|m|$ (here $|m| = \sqrt{m \cdot m}$). Then we use integration by parts with respect to x_j (see property (ix) in Proposition 4.13) s times to get:

$$\widehat{f}(m) = \int f(x) e^{-2\pi i x\cdot m}\, dx = \frac{(-1)^s}{(2\pi i m_j)^s} \int (\partial_j^s f)(x) e^{-2\pi i x \cdot m}\, dx. \tag{4.1.9}$$

This proves (4.1.7). The statement about the limit follows from $\partial^\alpha f \in L^1$ and the Riemann–Lebesgue lemma.

(ii) Since $\int_{\mathbb{R}^n} \frac{1}{1+|x|^{n+1}}\, dx < \infty$, we have that (4.1.8) implies $\widehat{f} \in \ell^1$, so Theorem 4.29 tells us that $\sum \widehat{f}(m) e^{2\pi i x\cdot m} = f$ uniformly. We can apply ∂^α (with $|\alpha| \leq s-1$) term-wise since the resulting equality

$$\partial^\alpha f = \sum_{m\in\mathbb{Z}^n} \widehat{f}(m)(2\pi i m)^\alpha e^{2\pi i x\cdot m}$$

is still uniformly convergent by (4.1.8). Note that with $|\alpha| = s$ the uniform convergence of the last series is no longer clear. □

4.2 Fourier Analysis on $\mathbb{R}^n$

4.2.1 Schwartz space

Intuition. Instead of defining the Fourier transform for $L^1(\mathbb{R}^n)$ functions, let us introduce the class of Schwartz functions (smooth functions that decay fast an infinity) and work with Fourier transform for only those. Because of the decay, technicalities will be very mild, and then afterwards we will be able to extend the Fourier transform and its properties to the general setting.

Definition 4.44. *(i) A function $f : \mathbb{R}^n \to \mathbb{C}$ is called a **Schwartz function** (we write $f \in \mathcal{S}(\mathbb{R}^n)$) if $f \in C^\infty(\mathbb{R}^n)$ and for any pair of multi-indices α and β,*

$$||f||_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty. \tag{4.2.1}$$

*(ii) Let $f_n, f \in \mathcal{S}(\mathbb{R}^n)$. We say that the **sequence** f_n **converges to** f **in** $\mathcal{S}(\mathbb{R}^n)$ iff for all α, β:*

$$||f_n - f||_{\alpha,\beta} \to 0$$

as $n \to \infty$.

Remarks 4.45. (a) $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$.

(b) Any function of the form $x^\kappa e^{-|x|^2}$ (where κ is some multi-index) is in $\mathcal{S}(\mathbb{R}^n)$.

(c) We defined a Schwartz function via $\sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty$ for all α, β. Equivalently, one can use $\sup_{x\in\mathbb{R}^n} |\partial^\beta (x^\alpha f(x))| < \infty$ for all α, β. Yet another way to characterize Schwartz functions is by saying that for all $N \in \mathbb{N}$ and all α there is $C_{\alpha,N} > 0$ such that

$$|(\partial^\alpha f)(x)| \le C_{\alpha,N} \frac{1}{(1+|x|)^N}.$$

Definition 4.46. *Let X be a Hausdorff topological vector space. Let $\{||\cdot||_j\}_{j=1}^\infty$ be a family of semi-norms on X which generate the topology on X. We say that X is a* ***Fréchet space*** *if X is complete with respect to this family of semi-norms: that is, if $x_j \in X$ is a Cauchy sequence (for every j: $||x_k - x_m||_j \to 0$ as $k, m \to \infty$), then x_j converges to some element x in X (that is, for every j: $||x_n - x||_j \to 0$ as $n \to \infty$).*

Remark 4.47. One can define a complete metric

$$d(x,y) = \sum_{k=0}^\infty 2^{-j} \frac{||x-y||_j}{1 + ||x-y||_j}$$

that induces the same topology on X.

Proposition 4.48. $\mathcal{S}(\mathbb{R}^n)$ *with the family of semi-norms $\{||\cdot||_{\alpha,\beta}\}$ forms a Fréchet space.*

Proof. To prove completeness, let $||f_j - f_k||_{\alpha,\beta} \to 0$ as $j, k \to \infty$. Then for any β, $\{\partial^\beta f_k\}_{k=1}^\infty$ is Cauchy in $||\cdot||_\infty$, so $\partial^\beta f_k \to g_\beta$ as $k \to \infty$ uniformly, for some continuous g_β. Let e_j be the j-th standard vector in $\mathbb{R}^n$. Then by the Fundamental Theorem of Calculus, we have

$$f_k(x + te_j) - f_k(x) = \int_0^t \partial_j f_k(x + se_j)\, ds.$$

Taking $k \to \infty$ (using uniform convergence), we get

$$g_0(x + te_j) - g_0(x) = \int_0^t g_{e_j}(x + se_j)\, ds.$$

Differentiating this, we get $\partial_j g_0 = g_{e_j}$. By a similar argument and an induction, we get $\partial^\beta g_0 = g_\beta$ for all β. Now note that $\{x^\alpha f_k\}_{k=1}^\infty$ is also uniformly convergent, and it is easy to see that the limit therefore must be $x^\alpha g_0$. Similarly for $\{x^\alpha \partial^\beta f_k\}_{k=1}^\infty$, which shows that $g_0 \in \mathcal{S}(\mathbb{R}^n)$ and $f_k \to g_0$ in $\mathcal{S}(\mathbb{R}^n)$. ☐

Proposition 4.49. *(i) If $f, g \in \mathcal{S}(\mathbb{R}^n)$ then fg and $f * g$ are in $\mathcal{S}(\mathbb{R}^n)$. Moreover,*

$$\partial^\alpha (f * g) = (\partial^\alpha f) * g = g * (\partial^\alpha g)$$

for any multi-index α.

(ii) For any $1 \le p < \infty$, $\mathcal{S}(\mathbb{R}^n)$ is $||\cdot||_p$-dense in $L^p(\mathbb{R}^n)$.

Proof. Left as an exercise. ☐

4.2.2 Fourier transform on $\mathcal{S}(\mathbb{R}^n)$: definition and basic properties

Definition 4.50. *Let $f \in \mathcal{S}(\mathbb{R}^n)$. Define the* ***Fourier transform of*** f *to be*

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi}\, dx. \tag{4.2.2}$$

Recall the notation $\widetilde{f}(x) := f(-x)$ and $\tau_y(f)(x) := f(x-y)$.

Proposition 4.51. *Let f, g be in $\mathcal{S}(\mathbb{R}^n)$. Then for all $y \in \mathbb{R}^n$, $\lambda \in \mathbb{C}$:*

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

(ii) $\widehat{\lambda f} = \lambda \widehat{f}$;

(iii) $\widehat{\overline{f}}(m) = \overline{\widetilde{\widehat{f}}}$;

(iv) $\widehat{\widetilde{f}} = \widetilde{\widehat{f}}$;

(v) $\widehat{\tau_y(f)}(\xi) = \widehat{f}(\xi) e^{-2\pi i \xi\cdot y}$;

(vi) $\left(e^{2\pi i y\cdot x} f(x)\right)\widehat{\ }(\xi) = \tau_y(\widehat{f})(\xi)$;

(vii) $||\widehat{f}||_\infty \leq ||f||_1$;

(viii) $\widehat{f * g} = \widehat{f}\widehat{g}$;

(ix) $\widehat{\partial^\alpha f}(\xi) = (2\pi i \xi)^\alpha \widehat{f}(\xi)$;

(x) $(\partial^\alpha \widehat{f})(\xi) = ((-2\pi i x)^\alpha f(x))\widehat{\ }(\xi)$;

(xi) $f \in \mathcal{S}(\mathbb{R}^n)$ *implies* $\widehat{f} \in \mathcal{S}(\mathbb{R}^n)$.

Example 4.52. If $f = e^{-\pi a|x|^2}$ with $a > 0$ (note that $f \in \mathcal{S}(\mathbb{R}^n)$), then its Fourier transform is $a^{-n/2} e^{-\pi|\xi|^2/a}$.

To prove this for $n = 1$, we use properties (x) and then (ix) (see above) to see that $\frac{d}{d\xi}\widehat{f}(\xi) = (-2\pi i x f(x))\widehat{\ }(\xi) = (-2\pi i x e^{-\pi a x^2})\widehat{\ }(xi) = \frac{i}{a}(f')\widehat{\ }(\xi) = -\frac{2\pi}{a}\xi \widehat{f}(\xi)$. This is an ODE which integrates to $e^{\pi\xi^2/a}\widehat{f}(\xi) = c$. Taking $\xi = 0$ gives $c = \widehat{f}(0) = \int e^{-\pi a x^2}\, dx = \frac{1}{\sqrt{a}}$.

The case for general n follows from Fubini.

4.2.3 Fourier transform on $\mathcal{S}(\mathbb{R}^n)$: the inverse Fourier transform and Fourier inversion

Definition 4.53. *For a Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$, define its* ***inverse Fourier transform*** *to be*

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

Remark 4.54. It is clear that all the properties of Fourier transform from the previous section, hold for the inverse transform with the straightforward modifications.

Theorem 4.55. *Let $f, g \in \mathcal{S}(\mathbb{R}^n)$.*

(i) $\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}^n} \widehat{f}(x)g(x)\,dx$.

(ii) *(Fourier Inversion)* $(\widehat{f})^\vee = f = (f^\vee)^\wedge$;

(iii) *(Parseval's relation)* $\langle f, g\rangle = \left\langle \widehat{f}, \widehat{g}\right\rangle$;

(iv) *(Plancherel's identity)* $||f||_2 = ||\widehat{f}||_2 = ||f^\vee||_2$;

Proof. (i) follows from Fubini theorem: $\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)e^{-2\pi i y\cdot x}\,dy\,dx = \widehat{f}(y)g(y)\,dy$.

(ii) Let $\varepsilon > 0$, $y \in \mathbb{R}^n$, and $g(x) = e^{2\pi i y\cdot x}e^{-\pi\varepsilon^2|x|^2}$. By combining Example 4.52 and Proposition 4.51(vi), we have

$$\widehat{g}(\xi) = \frac{1}{\varepsilon^n}e^{-\pi|\xi-y|^2/\varepsilon^2}.$$

The important observation is that $\{\widehat{g}(\xi)\}_{\varepsilon\to 0}$ is an approximate identity. By (i) we get

$$\int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi|x-y|^2/\varepsilon^2}\,dx = \int_{\mathbb{R}^n}\widehat{f}(x)e^{2\pi i y\cdot x}e^{-\pi\varepsilon^2|x|^2}\,dx. \tag{4.2.3}$$

Now take $\varepsilon \downarrow 0$. The left-hand side of (4.2.3) goes to $f(y)$ since f is continuous and $\{\widehat{g}(\xi)\}_{\varepsilon\to 0}$ is an approximate identity (see Proposition 4.10). The right-hand side of (4.2.3) converges to $(\widehat{f})^\vee(y)$ by the Lebesgue Dominated Convergence, so we are done.

(iii) is obtained by plugging in $g = (\bar{h})^\vee$ into (i).

(iv) follows by choosing $f = g$ in (iii). □

4.2.4 Fourier transform on $L^1(\mathbb{R}^n)$: definition and basic properties

Note that Definition 4.50 of the Fourier transform works for any $f \in L^1(\mathbb{R}^n)$.

Proposition 4.56. *Let $f, g \in L^1(\mathbb{R}^n)$. Then properties (i)–(viii) of Proposition 4.51 hold.*

Theorem 4.57 (Riemann–Lebesgue lemma)**.** *Let $f \in L^1(\mathbb{R}^n)$. Then $\widehat{f} \in C_0(\mathbb{R}^n)$.*

Proof. Continuity (in fact, uniform continuity) follows from

$$\lim_{h\to 0}\widehat{f}(\xi+h) - \widehat{f}(\xi) = \lim_{h\to 0}\int_{\mathbb{R}^n} f(x)e^{-2\pi i x\cdot\xi}(e^{-2\pi i x\cdot h} - 1)\,dx = 0$$

where we use the Dominated Convergence Theorem – the integrand is bounded by $2|f(x)| \in L^1$.

Convergence to 0 can be proved by first proving this for simple functions (Fourier transform of $\chi_{[a,b]}(x)$ is $\frac{1}{2\pi i\xi}(e^{-2\pi i\xi a} - e^{-2\pi i\xi b})$ which goes to 0), and then approximating $f \in L^1$ by a simple function ϕ with $||\phi - f||_1 \le \varepsilon$, and then $|\widehat{f}(\xi) - \widehat{\phi}(\xi)| \le ||f - \phi||_1 \le \varepsilon$. □

4.2.5 Fourier transform on $L^1(\mathbb{R}^n)$: the Fourier Inversion formula

Similarly, the inverse Fourier transform (Definition 4.53) work for any $f \in L^1(\mathbb{R}^n)$. However the Fourier Inversion formula does not necessarily holds for all functions in L^1, compare with Theorem 4.29.

Theorem 4.58. *Suppose $f \in L^1(\mathbb{T}^n)$ and $\widehat{f} \in L^1(\mathbb{T}^n)$. Then $(\widehat{f})^\vee = (f^\vee)^\wedge$ is a continuous function which is equal to f a.e.*

Proof. The proof follows exactly same lines as the proof of Theorem 4.55(i)–(ii). The main difference is that $f \in L^1$, so that we do not get pointwise convergence of the left-hand side of (4.2.3) to f but only L^1-convergence. The right-hand side of (4.2.3) still converges pointwise to $(\widehat{f})^\vee$, so we are done. □

4.2.6 Fourier transform on $L^2(\mathbb{R}^n)$

Unlike the case of n-torus, $L^2(\mathbb{R}^n)$ is not a subspace of $L^1(\mathbb{R}^n)$, and the definition of the Fourier transform does *not* work for any $f \in L^2(\mathbb{R}^n)$.

However the Fourier transform is well-defined on $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ which is dense in $L^2(\mathbb{R}^n)$ (e.g., because it contains the class of simple functions which we know are $||\cdot||_2$-dense in L^2). Moreover, on $L^1 \cap L^2$ the Fourier transform is an isometry (see Theorem 4.55(iv) – these properties can be extended to $f \in L^1 \cap L^2$).

Definition 4.59. *For $f \in L^2(\mathbb{R}^n)$, choose any $f_N \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ such that $||f_N - f||_2 \to 0$. Then the $L^2(\mathbb{R}^n)$-limit of $\widehat{f_N}$ is defined to be the **Fourier transform of** f, denoted by $\mathcal{F}(f)$.*

Remarks 4.60. (a) If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ then $\widehat{f} \in C_0(\mathbb{R}^n)$ and $\mathcal{F}(f) \in L^2(\mathbb{R}^n)$. It can be shown that these two functions are equal a.e.. Therefore it is not much abuse of notation to use again $\widehat{f}$ for $\mathcal{F}(f)$.

(b) In particular, we can take $f_N(x) := f(x)\chi_{|x|\le N} \in L^1 \cap L^2$ in the definition to get that $\mathcal{F}(f)(\xi)$ is the L^2-limit of $(f(x)\chi_{|x|\le N})^\wedge$.

$$\mathcal{F}(f)(\xi) = L^2\text{-}\lim_{N\to\infty} \int_{|x|\le N} f(x) e^{-2\pi i x\cdot\xi}\, dx.$$

The existence of L^2-limit implies that we also have a pointwise (a.e.) convergence along a subsequence.

Definition 4.61. *For $f \in L^2(\mathbb{R}^n)$ we define the **inverse Fourier transform of** f to be the $L^2(\mathbb{R}^n)$ function given by*

$$\mathcal{F}^{-1}(f)(x) := \mathcal{F}(f)(-x).$$

Remarks 4.62. (a) Again, if $f \in L^1 \cap L^2$, it can be shown that $\mathcal{F}^{-1}(f) = f^\vee$ a.e., so we will use either notation interchangeably.

(b) Because we have (the Fourier inversion) $\mathcal{F} \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ \mathcal{F}$ on a dense subset $\mathcal{S}(\mathbb{R}^n)$ of L^2, we get that it holds for all $f \in L^2(\mathbb{R}^n)$.

(c) Part (b) says that if $f \in L^2$, then the inverse Fourier transform $(\widehat{f})^\vee$ of $\widehat{f}$ converges to f in the L^2-sense. Just like in the Remark 4.60(b), we can take an approximating sequence to be $\widehat{f}(x)\chi_{\max|x_j|\le N}$. In other words, if

$$s_N(f) := \left(\widehat{f}(x)\chi_{\max|x_j|\le N}\right)^\vee = \int_{\max|x_j|\le N} \widehat{f}(x) e^{2\pi i x\cdot\xi}\, dx. \tag{4.2.4}$$

(this is the analogue of the square partial Fourier sum $S_N(f)$ that we had on the n-torus), then (b) says that

$$||f - s_N||_2 \to 0$$

(compare with Corollary 4.12(i)). Of course there is no significance in the square $\max|x_j| \le N$ here and circles or any other expanding sequence of domains would also work (just like on the n-torus for L^2-convergence).

4.2.7 Fourier transform on $L^1(\mathbb{R}^n)$: various types of convergence

Intuition. Recall (Sections 4.1.3—4.1.9) that on $\mathbb{T}^n$ we had different ways to sum the Fourier series (pointwise, L^p, Cesàro, Abel (see HW), etc) which allowed us to recover f (at least a.e.) from the Fourier series.

Similar theory can be developed for the Fourier transform. As a prototype, let us discuss the analogue of Theorem 4.22 for the uniform and L^p convergence of the Cesàro means of Fourier series.

Given $f \in L^1(\mathbb{R}^n)$ and its Fourier transform $\widehat{f}$, define

$$s_N(f) := \left(\widehat{f}(x)\chi_{\max|x_j|\le N}\right)^\vee = \int_{\max|x_j|\le N} \widehat{f}(x)e^{2\pi i x\cdot\xi}\,dx, \tag{4.2.5}$$

and

$$\sigma_N(f) := \left(\widehat{f}(x)\prod_{j=1}^n\left(1-\frac{|x_j|}{N}\right)\chi_{\max|x_j|\le N}\right)^\vee = \int_{\max|x_j|\le N}\prod_{j=1}^n\left(1-\frac{|x_j|}{N}\right)\widehat{f}(x)e^{2\pi i x\cdot\xi}\,dx. \tag{4.2.6}$$

These should be viewed as the analogues of $S_N(f)$ and $C_N(f)$ from the n-torus: cf. (4.1.3) and (4.1.5), respectively.

To make the analogy even more transparent, let us note the following:

Lemma 4.63. *Let $f \in L^1(\mathbb{R}^n)$. Then*

(i) $s_N(f) = f * d_N$, where

$$d_N(x) := \prod_{j=1}^n \frac{\sin(2\pi N x_j)}{\pi x_j} = \left(\chi_{\max|x_j|\le N}(x)\right)^\vee.$$

(ii) $\sigma_N(f) = f * k_N$, where

$$k_N(x) := \prod_{j=1}^n \frac{\sin^2(\pi N x_j)}{\pi^2 N x_j^2} = \left(\prod_{j=1}^n\left(1-\frac{|x_j|}{N}\right)\chi_{\max|x_j|\le N}(x)\right)^\vee.$$

Remarks 4.64. (a) More generally, both (i) and (ii) are consequences of the following fact: assume that $f, h \in L^1(\mathbb{R}^n)$ and that $\widehat{h} \in L^1(\mathbb{R}^n)$, then we have that

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

(b) d_N and k_N are the real-line analogues of the Dirichlet and Fejér kernels D_N and K_N that we had on the n-torus.

(c) Compare with Proposition 4.19 as well as Proposition 4.17.

Proof. Let us prove Remark (a). By the Property (viii) of Proposition 4.51 we have $\widehat{f*h} = \widehat{f}\,\widehat{h} = \widehat{f}H$. Since $H \in L^1$ and $\widehat{f} \in C_0$, we have that $\widehat{f}H \in L^1$, so we can apply the Inversion Formula (Theorem 4.58) to get $f*h = (\widehat{f}H)^\vee$ which is what we wanted to show. □

Proposition 4.65. *$\{k_N(x)\}_{N=1}^\infty$ forms an approximate identity on $\mathbb{R}^n$.*

Remark 4.66. Just like before, the Dirichlet kernels $\{d_N\}_{N=1}^\infty$ do not form an approximate identity: they satisfy $\int d_N\,dx = 1$ for all N, but fail $d_N \ge 0$ or $\int |d_N|\,dx < C$.

Proof. Just a calculation. □

Corollary 4.67. *(i) For any $f \in L^p(\mathbb{R}^n)$ $(1 \le p < \infty)$, $\sigma_N(f) \to f$ in the $||\cdot||_p$-norm.*

(ii) For any $f \in L^1(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, $\sigma_N(f) \to f$ uniformly on compacts of $\mathbb{R}^n$.

Remark 4.68. This is the analogue of Theorem 4.22.

Proof. Immediate from Propositions 4.65 and 4.10. ☐

4.2.8 Fourier transform on $L^p(\mathbb{R}^n)$: summability of $\widehat{f}$ and Hausdorff–Young

Intuition. This is the exact analogue of Section 4.1.10

Recall that we showed:

- $f \in L^2(\mathbb{R}^n)$ iff $\widehat{f} \in L^2(\mathbb{R}^n)$;
- $f \in L^1(\mathbb{T}^n)$ implies $\widehat{f} \in C_0(\mathbb{R}^n)$, in particular $\widehat{f} \in L^\infty(\mathbb{R}^n)$.

Theorem 4.69 (Hausdorff–Young Inequality)**.** *Suppose that $1 \le p \le 2$, and let $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mathbb{R}^n)$, then $\widehat{f} \in L^q(\mathbb{R}^n)$. Moreover, $||\widehat{f}||_q \le ||f||_p$.*

Proof. Use Riesz–Thorin Theorem 3.35 to interpolate between the above two $L^2 \to L^2$ and $L^1 \to L^\infty$ results. ☐

4.2.9 Fourier series on $L^1(\mathbb{R}^n)$: decay of $\widehat{f}$ at ∞

Intuition. The analogue of the results from 4.1.11 can also be proved in the same way: the smoother f is, the faster is the decay of $\widehat{f}$ at infinity. We omit the details.

4.2.10 The Poisson Summation Formula

Intuition. Given a function $f \in L^1(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, let us form its "periodization"

$$F(x) = \sum_{m \in \mathbb{Z}^n} f(x+m). \tag{4.2.7}$$

F can now be viewed as a periodic $\mathbb{T}^n \to \mathbb{C}$ function (in fact, $F \in L^1(\mathbb{T}^n)$ since $f \in L^1(\mathbb{R}^n)$). It is not hard to see that the Fourier coefficients $\widehat{F}(m)$ of F coincide with the Fourier transform $\widehat{f}(\xi)$ evaluated at $\xi = m$. Assuming the Fourier series of F converge to F, we therefore should expect $F(x) = \sum \widehat{f}(m) e^{2\pi i m \cdot x}$, in particular $F(0) = \sum \widehat{f}(m)$ which is the Poisson Summation Formula, see (4.2.10).

Theorem 4.70 (Poisson Summation Formula)**.** *Suppose $f \in L^1(\mathbb{R}^n) \cap C(\mathbb{R}^n)$ and $\widehat{f} \in L^1(\mathbb{R}^n)$ satisfy*

$$|f(x)| + |\widehat{f}(x)| \le C(1+|x|)^{-n-\delta} \tag{4.2.8}$$

for some $C > 0$, $\delta > 0$. Then for all $x \in \mathbb{R}^n$ we have

$$\sum_{m \in \mathbb{Z}^n} f(x+m) = \sum_{m \in \mathbb{Z}^n} \widehat{f}(m) e^{2\pi i m \cdot x}, \tag{4.2.9}$$

and in particular

$$\sum_{m\in\mathbb{Z}^n} f(m) = \sum_{m\in\mathbb{Z}^n} \widehat{f}(m). \tag{4.2.10}$$

Remark 4.71. Note that (4.2.8) automatically implies $\widehat{f} \in L^1(\mathbb{R}^n)$

Proof. Define F as in (4.2.7). Applying Theorem 2.26 (together with $f \in L^1(\mathbb{R}^n)$) we obtain that (4.2.7) converges a.e. and $F \in L^1(\mathbb{T}^n)$. In fact, it converges uniformly on $\mathbb{T}^n$ from the condition $|f(x)| \le C(1+|x|)^{-n-\delta}$; in particular F is continuous. Applying Theorem 2.26 again we get

$$\widehat{F}(m) = \widehat{f}(m)$$

for each $m \in \mathbb{Z}^n$. Since

$$\sum_{m\in\mathbb{Z}^n} |\widehat{F}(m)| = \sum_{m\in\mathbb{Z}^n} |\widehat{f}(m)| \le C \sum_{m\in\mathbb{Z}^n} \frac{1}{(1+|m|)^{n+\delta}} < \infty,$$

we get (see Theorem 4.29) that the Fourier series of F converge uniformly to F. This is precisely (4.2.9). ☐
