1

Fourier series: convergence and summability

1.1. The basics: partial sums and the Dirichlet kernel

1.1.1. Definitions

We begin with a basic object in analysis, namely the Fourier series associated with a function or a measure on the circle. To be specific, let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus (in other words, the circle). We will consider various function spaces on the torus $\mathbb{T}$, namely the space of continuous functions $C(\mathbb{T})$, the space of Hölder continuous functions $C^\alpha(\mathbb{T})$ where $0 < \alpha \le 1$, and the Lebesgue spaces $L^p(\mathbb{T})$ where $1 \le p \le \infty$. The space of complex Borel measures on $\mathbb{T}$ will be denoted by $\mathcal{M}(\mathbb{T})$. Any $\mu \in \mathcal{M}(\mathbb{T})$ has associated with it a *Fourier series*

$$\mu \sim \sum_{n=-\infty}^{\infty} \hat{\mu}(n) e(nx) \tag{1.1}$$

where $e(x) := e^{2\pi i x}$ and

$$\hat{\mu}(n) := \int_0^1 e(-nx)\,\mu(dx) = \int_{\mathbb{T}} e(-nx)\,\mu(dx).$$

The symbol $\sim$ in (1.1) is formal and simply means that the series on the right-hand side is associated with μ. If $\mu(dx) = f(x)\,dx$ where $f \in L^1(\mathbb{T})$, then we may write $\hat{f}(n)$ instead of $\hat{\mu}(n)$.

The central question which we wish to explore in this chapter is the following: when does μ equal the right-hand side in (1.1), that is, when does it represent f in a suitable sense? Note that if we start from a *trigonometric polynomial*

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e(nx)$$

where all but finitely many a_n are zero, then we see that

$$\hat{f}(n) = a_n \quad \forall n \in \mathbb{Z}. \tag{1.2}$$

In other words, we have the pointwise equality

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e_n(x),$$

with $e_n(x) := e(nx)$. Property (1.2) is equivalent to the basic *orthogonality relation*

$$\int_{\mathbb{T}} e_n(x) \overline{e_m(x)}\, dx = \delta_0(n-m), \tag{1.3}$$

where $\delta_0(j) = 1$ if $j = 0$ and $\delta_0(j) = 0$ otherwise.

It is therefore natural to explore the question of the *convergence* of the Fourier series for more general functions. Of course, precise meaning of the convergence of infinite series needs to be specified before this fundamental question can be answered. It is fair to say that much modern analysis (including functional analysis) arose out of the struggle with this question. For example, the notion of the Lebesgue integral was developed in order to overcome the deficiencies in the older Riemannian definition of the integral that had been revealed through the study of Fourier series. The reader will note the recurring theme of the convergence of Fourier series throughout both volumes of this book.

1.1.2. Dirichlet kernel

It is natural to start from the most basic notion of convergence, namely that of *pointwise convergence*, in the case where the measure μ is of the form $\mu(dx) = f(x)\,dx$ with $f(x)$ continuous or of even better regularity. The partial sums of $f \in L^1(\mathbb{T})$ are defined as

$$\begin{aligned} S_N f(x) &= \sum_{n=-N}^{N} \hat{f}(n) e(nx) = \sum_{n=-N}^{N} \int_{\mathbb{T}} e(-ny) f(y)\, dy\; e(nx) \\ &= \int_{\mathbb{T}} \sum_{n=-N}^{N} e(n(x-y)) f(y)\, dy = \int_{\mathbb{T}} D_N(x-y) f(y)\, dy, \end{aligned}$$

where $D_N(x) := \sum_{n=-N}^{N} e(nx)$ is the *Dirichlet kernel*. In other words, we have shown that the partial sum operator S_N is given by convolution with the *Dirichlet kernel* D_N:

$$S_N f(x) = (D_N * f)(x). \tag{1.4}$$

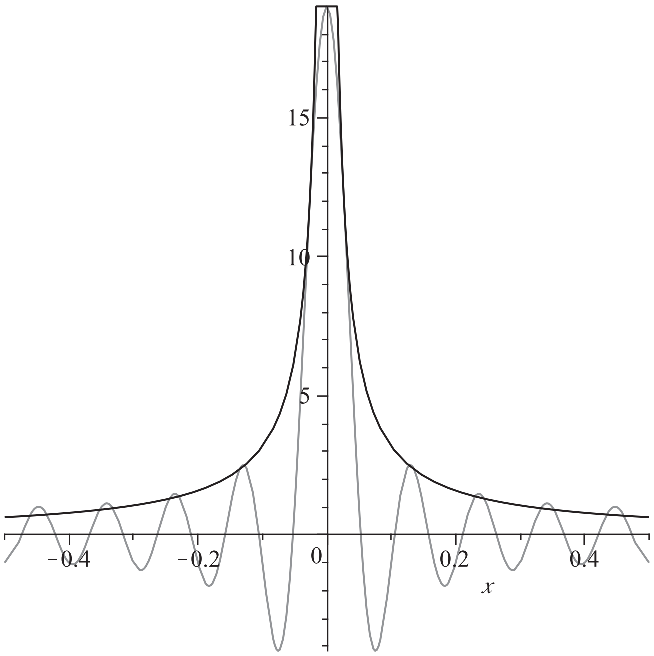

Figure 1.1. The Dirichlet kernel D_N and the upper envelope $\min(2N+1, |\pi x|^{-1})$ for $N = 9$. See Exercise 1.1.

In order to understand basic properties of this convolution, we first sum the geometric series defining $D_N(x)$ to obtain an explicit expression for the Dirichlet kernel.

Exercise 1.1 Verify that, for each integer $N \geq 0$,

$$D_N(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)} \tag{1.5}$$

and draw the graph of D_N for several different values of N, say $N = 2$ and $N = 5$; cf. Figure 1.1. Prove the bound

$$|D_N(x)| \leq C \min\left(N, \frac{1}{|x|}\right) \tag{1.6}$$

for all $N \geq 1$ and some absolute constant C. Finally, prove the bound

$$C^{-1} \log N \leq \|D_N\|_{L^1(\mathbb{T})} \leq C \log N \tag{1.7}$$

for all $N \geq 2$, where C is another absolute constant.

The growth of the bound in (1.7), as well as the oscillatory nature of D_N as given by (1.5), indicates that to understand the pointwise or almost everywhere convergence properties of $S_N f$ may be a very delicate matter. This will become clearer as we develop the theory.

1.1.3. Convolution

In order to study (1.4) we need to establish some basic properties of the convolution of two functions f, g on $\mathbb{T}$. If f and g are continuous, say, then define

$$(f * g)(x) := \int_{\mathbb{T}} f(x-y)g(y)\,dy = \int_{\mathbb{T}} g(x-y)f(y)\,dy. \tag{1.8}$$

It is helpful to think of $f * g$ as an average of translates of f by the measure $g(y)\,dy$ (or the same statement but with f and g interchanged). In particular, convolution commutes with the translation operator τ_z, which is defined for any $z \in \mathbb{T}$ by its action on functions, i.e., $(\tau_z f)(x) = f(x-z)$. Indeed, one may immediately verify that

$$\tau_z(f * g) = (\tau_z f) * g = f * (\tau_z g). \tag{1.9}$$

In passing, we mention the important relation between the Fourier transform and translations:

$$\widehat{(\tau_z \mu)}(n) = e(-zn)\hat{\mu}(n) \quad \forall n \in \mathbb{Z}.$$

In what follows, we will abbreviate *almost everywhere* or *almost every* by *a.e.*

Lemma 1.1 *The operation of convolution as defined in (1.8) satisfies the following properties.*

(i) *Let $f, g \in L^1(\mathbb{T})$. Then, for a.e. $x \in \mathbb{T}$, one has that $f(x-y)g(y)$ is L^1 in y. Thus, the integral in* (1.8) *is well defined for a.e. $x \in \mathbb{T}$ (but not necessarily for every x), and the bound $\|f * g\|_1 \le \|f\|_1\|g\|_1$ holds.*

(ii) *More generally, $\|f * g\|_r \le \|f\|_p\|g\|_q$ for all $1 \le r, p, q \le \infty$,*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \; f \in L^p, \; g \in L^q.$$

This is called Young's inequality.

(iii) *If $f \in C(\mathbb{T})$, $\mu \in \mathcal{M}(\mathbb{T})$ then $f * \mu$ is well defined. For $1 \le p \le \infty$,*

$$\|f * \mu\|_p \le \|f\|_p\|\mu\|;$$

this allows one to extend $f * \mu$ to arbitrary $f \in L^p$.

(iv) *If $f \in L^p(\mathbb{T})$ and $g \in L^{p'}(\mathbb{T})$, where $1 \le p \le \infty$, and*

$$\frac{1}{p} + \frac{1}{p'} = 1$$

then $f * g$, originally defined only a.e., extends to a continuous function on $\mathbb{T}$, and

$$\|f * g\|_\infty \le \|f\|_p \|g\|_{p'}. \tag{1.10}$$

(v) *For $f, g \in L^1(\mathbb{T})$ one has, for all $n \in \mathbb{Z}$,*

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

Proof (i) is an immediate consequence of Fubini's theorem since $f(x-y)g(y)$ is jointly measurable on $\mathbb{T} \times \mathbb{T}$ and belongs to $L^1(\mathbb{T} \times \mathbb{T})$. For (ii), first let $q = 1$. Then (ii) can be obtained by interpolating between the case $p = 1$ covered in (i) and the easy bound for $p = \infty$. Alternatively, one can use Minkowski's inequality,

$$\|f * g\|_p \le \int_{\mathbb{T}} \|f(\cdot - y)\|_p |g(y)| \, dy \le \|f\|_p \|g\|_1,$$

which also implies (iii). The other extreme is $q = p'$, which is covered by (iv). The remaining choices of q follow by interpolation relative to g. The bound (1.10) is Hölder's inequality. Part (iv) follows from the fact that $C(\mathbb{T})$ is dense in $L^p(\mathbb{T})$ for $1 \le p < \infty$ and from the translation invariance (1.9). Indeed, one may verify from the uniform continuity of functions in $C(\mathbb{T})$ and (1.9) that $f * \mu \in C(\mathbb{T})$ for any $f \in C(\mathbb{T})$ and $\mu \in \mathcal{M}(\mathbb{T})$. Since uniform limits of continuous functions are continuous, (iv) now follows from the aforementioned denseness of $C(\mathbb{T})$ and (1.10).

Finally, (v) is a consequence of Fubini's theorem and the homomorphism property of the exponentials $e(n(x+y)) = e(nx)e(ny)$. □

The following exercise introduces the convolution as an operation acting on the Fourier coefficients of functions, rather than on the functions themselves. This is in the context of the largest class of functions where the respective Fourier series are absolutely convergent. This class of functions is necessarily a subalgebra of $C(\mathbb{T})$, called the *Wiener algebra*.

Exercise 1.2 Let $\mu \in \mathcal{M}(\mathbb{T})$ have the property that

$$\sum_{n \in \mathbb{Z}} |\hat{\mu}(n)| < \infty. \tag{1.11}$$

Show that $\mu(dx) = f(x)\,dx$, where $f \in C(\mathbb{T})$. Denote the space of all measures with this property by $\mathbb{A}(\mathbb{T})$ and identify these measure with their respective densities. Show that $\mathbb{A}(\mathbb{T})$ is an algebra under multiplication and that

$$\widehat{fg}(n) = \sum_{m \in \mathbb{Z}} \hat{f}(m)\hat{g}(n-m) \quad \forall n \in \mathbb{Z},$$

where the sum on the right-hand side is absolutely convergent for every $n \in \mathbb{Z}$ and indeed is absolutely convergent over all n. Moreover, show that $\|fg\|_{\mathbb{A}} \le \|f\|_{\mathbb{A}}\|g\|_{\mathbb{A}}$ where $\|f\|_{\mathbb{A}} := \|\hat{f}\|_{\ell^1}$. Finally, verify that if $f, g \in L^2(\mathbb{T})$ then $f * g \in \mathbb{A}(\mathbb{T})$.

Note that the Wiener algebra has a unit, namely the constant function 1. It is clear that if f has an inverse in $\mathbb{A}(\mathbb{T})$ then this is $1/f$, which in particular requires that $f \neq 0$ everywhere on $\mathbb{T}$. A remarkable theorem due to Norbert Wiener states that the converse holds, too; that is, if $f \in \mathbb{A}(\mathbb{T})$ does not vanish anywhere on $\mathbb{T}$ then $1/f \in \mathbb{A}(\mathbb{T})$. We present this result in Section 4.3 as an easy corollary to Gelfand's theory of commutative Banach algebras; see Corollary 4.27.

One of the most basic as well as oldest results on the pointwise convergence of Fourier series is the following theorem. We shall see later that it *fails* for functions that are merely continuous.

Theorem 1.2 *If* $f \in C^\alpha(\mathbb{T})$ *with* $0 < \alpha \le 1$ *then* $\|S_N f - f\|_\infty \to 0$ *as* $N \to \infty$.

Proof One has, with $\delta \in (0, \frac{1}{2})$ to be determined,

$$\begin{aligned} S_N f(x) - f(x) &= \int_0^1 (f(x-y) - f(x)) D_N(y)\, dy \\ &= \int_{|y|\le\delta} (f(x-y) - f(x)) D_N(y)\, dy \\ &\quad + \int_{1/2>|y|>\delta} (f(x-y) - f(x)) D_N(y)\, dy. \end{aligned} \tag{1.12}$$

Here we have exploited the fact that

$$\int_{\mathbb{T}} D_N(y)\, dy = 1.$$

We now use the bound from (1.6), i.e.,

$$|D_N(y)| \le C \min\left(N, \frac{1}{|y|}\right).$$

Here and in what follows, C will denote a numerical constant that can change from line to line. The first integral in (1.12) can be estimated as follows:

$$\int_{|y|\le\delta} |f(x) - f(x-y)| \frac{1}{|y|}\, dy \le [f]_\alpha \int_{|y|\le\delta} |y|^{\alpha-1}\, dy \le C[f]_\alpha \delta^\alpha, \tag{1.13}$$

with the usual C^α semi-norm

$$[f]_\alpha = \sup_{x,y} \frac{|f(x) - f(x-y)|}{|y|^\alpha}.$$

To bound the second term in (1.12) one needs to invoke the oscillation of $D_N(y)$. In fact, we have

$$
\begin{aligned}
B &:= \int_{1/2>|y|>\delta} (f(x-y)-f(x))D_N(y)\,dy \\
&= \int_{1/2>|y|>\delta} \frac{f(x-y)-f(x)}{\sin(\pi y)} \sin((2N+1)\pi y)\,dy \\
&= -\int_{1/2>|y|>\delta} h_x(y) \sin\left((2N+1)\pi\left(y+\frac{1}{2N+1}\right)\right)dy
\end{aligned}
$$

where

$$h_x(y) := \frac{f(x-y)-f(x)}{\sin(\pi y)}.$$

Therefore, with all integrals understood to be in the interval $(-\frac{1}{2}, \frac{1}{2})$,

$$
\begin{aligned}
2B &= \int_{|y|>\delta} h_x(y)\sin((2N+1)\pi y)\,dy \\
&\quad - \int_{|y-1/(2N+1)|>\delta} h_x\left(y-\frac{1}{2N+1}\right)\sin((2N+1)\pi y)\,dy \\
&= \int_{|y|>\delta} \left(h_x(y)-h_x\left(y-\frac{1}{2N+1}\right)\right)\sin((2N+1)\pi y)\,dy \\
&\quad - \int_{[-\delta,-\delta+1/(2N+1)]} h_x\left(y-\frac{1}{2N+1}\right)\sin((2N+1)\pi y)\,dy \\
&\quad + \int_{[\delta,\delta+1/(2N+1)]} h_x\left(y-\frac{1}{2N+1}\right)\sin((2N+1)\pi y)\,dy.
\end{aligned}
$$

These integrals are estimated by putting absolute values inside. To do so we use the bounds

$$
\begin{aligned}
|h_x(y)| &< C\frac{\|f\|_\infty}{\delta}, \\
|h_x(y)-h_x(y+\tau)| &< C\left(\frac{|\tau|^\alpha[f]_\alpha}{\delta}+\frac{\|f\|_\infty}{\delta^2}|\tau|\right),
\end{aligned}
$$

if $|y|>\delta>2\tau$.

In view of the preceding discussion, one obtains

$$|B| \le C\left(\frac{N^{-\alpha}[f]_\alpha}{\delta}+\frac{N^{-1}\|f\|_\infty}{\delta^2}\right), \tag{1.14}$$

provided that $\delta > 1/N$. Choosing $\delta = N^{-\alpha/3}$ one concludes from (1.12), (1.13), and (1.14) that

$$|(S_N f)(x) - f(x)| \le C\left(N^{-\alpha^2/3}+N^{-2\alpha/3}+N^{-1+2\alpha/3}\right) \tag{1.15}$$

for any function with $\|f\|_\infty + [f]_\alpha \le 1$, which proves the theorem. □

Cambridge University Press
978-0-521-88245-3 - Classical and Multilinear Harmonic Analysis: Volume I
Camil Muscalu and Wilhelm Schlag
Excerpt
More information

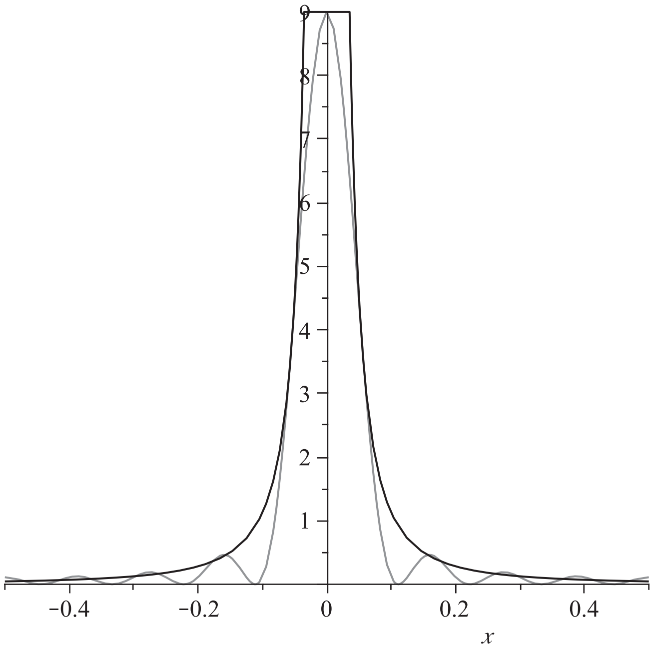

Figure 1.2. The Fejér kernel K_N and the upper envelope $\min(N, (N(\pi x)^2)^{-1})$ for $N = 9$.

The reader is invited to optimize the rate of decay derived in (1.15). In other words, the challenge is to obtain the largest $\beta > 0$ in terms of α such that the bound in (1.15) becomes $CN^{-\beta}$ for any f with $\|f\|_\infty + [f]_\alpha \le 1$.

1.2. Approximate identities, Fejér kernel

1.2.1. Cesáro means of partial sums

The difficulties with the Dirichlet kernel (see Figure 1.1), such as its slow $1/|x|$ decay, can be regarded as a result of the "discontinuity" in $\widehat{D_N} = \chi_{[-N,N]}$: this indicator function on the lattice $\mathbb{Z}$ jumps at $\pm N$. Therefore we may hope to obtain a kernel that is easier to analyze – in a sense that will be made precise below by means of the notion of *approximate identity* – by substituting for D_N a suitable average whose Fourier transform does not exhibit such jumps.

An elementary way of carrying this out is given by the Cesàro mean, i.e.,

$$\sigma_N f := \frac{1}{N}\sum_{n=0}^{N-1} S_n f.$$

Setting

$$K_N := \frac{1}{N} \sum_{n=0}^{N-1} D_n,$$

where K_N is called the *Fejér kernel*, one therefore has $\sigma_N f = K_N * f$.

Exercise 1.3 Let K_N be a Fejér kernel with N a positive integer.

(a) Verify that $\widehat{K}_N$ looks like a triangle (see Figure 1.3), i.e., for all $n \in \mathbb{Z}$,

$$\widehat{K}_N(n) = \left(1 - \frac{|n|}{N}\right)^+. \tag{1.16}$$

(b) Show that

$$K_N(x) = \frac{1}{N} \left(\frac{\sin(N\pi x)}{\sin(\pi x)}\right)^2. \tag{1.17}$$

(c) Conclude that

$$0 \le K_N(x) \le C\, N^{-1} \min(N^2, x^{-2}). \tag{1.18}$$

We remark that the square and thus the positivity in (1.17) are not entirely surprising, since the triangle in (1.16) can be written as the convolution of two rectangles (the convolution is now at the level of the Fourier coefficients on the lattice $\mathbb{Z}$). Therefore we should expect K_N to have the form of the square of a version of D_M, where M is about half the size of N, suitably normalized.

The properties established in Exercise 1.3 ensure that the K_N form what is called an *approximate identity*.

Definition 1.3 The family $\{\Phi_N\}_{N=1}^{\infty} \subset L^\infty(\mathbb{T})$ forms an approximate identity provided that

(A1) $\int_0^1 \Phi_N(x)\, dx = 1$ for all N,
(A2) $\sup_N \int_0^1 |\Phi_N(x)|\, dx < \infty$,
(A3) for all $\delta > 0$ one has $\int_{|x|>\delta} |\Phi_N(x)|\, dx \to 0$ as $N \to \infty$.

The term "approximate identity" derives from the fact that $\Phi_N * f \to f$ as $N \to \infty$ in any reasonable sense; see Proposition 1.5. In other words, $\Phi_N \rightharpoonup \delta_0$ in the weak-$*$ sense of measures. Clearly, the so-called *box kernels*

$$\Phi_N(x) = N\chi_{[|x|N<1/2]}, \qquad N \ge 1,$$

satisfy (A1)–(A3) and as a family consistitute the most basic example of an approximate identity. Note that the set $\{D_N\}_{N\ge 1}$ does *not* form an approximate identity. Finally, we remark that Definition 1.3 applies not just to the torus $\mathbb{T}$ but equally well to the line $\mathbb{R}$, the tori $\mathbb{T}^d$, or the Euclidean spaces $\mathbb{R}^d$.

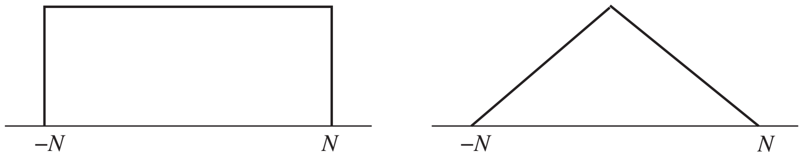

Figure 1.3. The Dirichlet and Fejér kernels as Fourier multipliers.

Next, we verify that the K_N belong to this class.

Lemma 1.4 *The Fejér kernels $\{K_N\}_{N=1}^{\infty}$ form an approximate identity.*

Proof We clearly have $\int_0^1 K_N(x)\,dx = 1$ and $K_N(x) \geq 0$, so that (A1) and (A2) hold. Condition (A3) follows from the bound (1.18). □

1.2.2. Convergence properties of approximate identities

Now we establish the basic convergence property of families that form approximate identities.

Proposition 1.5 *For any approximate identity $\{\Phi_N\}_{N=1}^{\infty}$ one has the following.*

(i) *If $f \in C(\mathbb{T})$ then $\|\Phi_N * f - f\|_\infty \to 0$ as $N \to \infty$.*
(ii) *If $f \in L^p(\mathbb{T})$, where $1 \leq p < \infty$, then $\|\Phi_N * f - f\|_p \to 0$ as $N \to \infty$.*
(iii) *For any measure $\mu \in \mathcal{M}(\mathbb{T})$, one has*

$$\Phi_N * \mu \rightharpoonup \mu, \qquad N \to \infty,$$

in the weak-$*$ sense.

Proof We begin with the uniform convergence. Since $\mathbb{T}$ is compact, f is uniformly continuous. Given $\varepsilon > 0$, let $\delta > 0$ be such that

$$\sup_x \sup_{|y|<\delta} |f(x-y) - f(x)| < \varepsilon.$$

Then, by Definition 1.3,

$$\begin{aligned}
|(\Phi_N * f)(x) - f(x)| &= \left| \int_{\mathbb{T}} (f(x-y) - f(x))\Phi_N(y)\,dy \right| \\
&\leq \sup_{x\in\mathbb{T}} \sup_{|y|<\delta} |f(x-y) - f(x)| \int_{\mathbb{T}} |\Phi_N(t)|\,dt + \int_{|y|\geq\delta} |\Phi_N(y)| 2\|f\|_\infty\,dy \\
&< C\varepsilon,
\end{aligned}$$

provided that N is large.