Introduction to harmonic analysis

Sasha Sodin and Matthias Täufer*

March 16, 2020

Lecture notes for an LTCC course, February–March, 2020. The sections marked with an asterisk were not covered in class. Please drop a note (at the email address below) if you spot any mistakes!

Contents

1 Introduction **2**

2 Fourier analysis on $\mathbb{Z}/N\mathbb{Z}$ **2**
- 2.1 Discrete Fourier transform . . . 2
- 2.2 Roth's theorem . . . 4
- 2.3 Fast Fourier transform* . . . 7

3 Fourier series **9**
- 3.1 Fourier coefficients and Plancherel theorem . . . 9
- 3.2 Convergence of Fourier series . . . 12
- 3.3 Féjér summation . . . 16

4 Fourier series: further topics and applications **18**
- 4.1 Polynomial approximation . . . 18
- 4.2 The heat equation . . . 23
- 4.3 Toeplitz matrices, and the Szegő theorem . . . 26
- 4.4 Central limit theorem . . . 32
- 4.5 Equidistribution modulo one . . . 35

5 Fourier transform **39**
- 5.1 Introduction . . . 39
- 5.2 Fourier transform in Schwartz space . . . 40
- 5.3 Poisson summation . . . 42
- 5.4 The uncertainty principle* . . . 44

Some notation **46**

References **47**

*School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom. E-mail: {a.sodin, m.taeufer}@qmul.ac.uk

1 Introduction

Harmonic analysis is used to study problems possessing translation invariance. For example, it can be used to diagonalise operators commuting with shifts.

Example 1.1. *For any $r_0, r_1, r_2 \in \mathbb{C}$, the matrix*

$$A = A[r_0, r_1, r_2] = \begin{pmatrix} r_0 & r_1 & r_2 \\ r_2 & r_0 & r_1 \\ r_1 & r_2 & r_0 \end{pmatrix}$$

is invariant under simultaneous cyclic shifts of the rows and the columns; in other words, it commutes with the shifts $A[0,1,0]$ and $A[0,0,1]$. Let $\omega = \exp(2\pi i/3)$, and

$$e_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ e_1 = \begin{bmatrix} 1 \\ \omega \\ \omega^2 \end{bmatrix}, \ e_2 = \begin{bmatrix} 1 \\ \omega^2 \\ \omega \end{bmatrix}.$$

Then

$$Ae_0 = (r_0 + r_1 + r_2)e_0 \ , \ Ae_1 = (r_0 + r_1\omega + r_2\omega^2)e_1 \ , \ Ae_2 = (r_0 + r_1\omega^2 + r_2\omega)e_2 \ .$$

Remarkably, the eigenvectors do not depend on r_0, r_1, r_2.[1]

There are many other operators thatr commute with shifts, e.g.

$$\frac{d}{dx} \ , \ \frac{d^2}{dx^2} \ , \ f \mapsto f * g \quad \text{et cet.}$$

and there are many problems with (explicit or hidden) translation invariance. This will be the scope of the applications that we shall consider in this minicourse.

Among the numerous books on Fourier analysis, we have been inspired by that of Dym and McKean [1972], which highlights the rôle of Fourier transform as a versatile tool in applications of various kinds. We have also borrowed some material from the monographs of Montgomery [2014], Katznelson [2004], Körner [1989], Stein and Shakarchi [2003], Lanczos [1966]; similarly to the last two of these, we do not rely on Lebesgue integration.

2 Fourier analysis on $\mathbb{Z}/N\mathbb{Z}$

2.1 Discrete Fourier transform

Equip the space of functions $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, i.e. of N-periodic functions from $\mathbb{Z}$ to $\mathbb{C}$, with the inner product

$$\langle f, g \rangle = \sum_{k=0}^{N-1} f(k)\overline{g(k)}$$

[1]Remarkably but not surprisingly: all the matrices A commute, and in fact form a commutative algebra.

This will be our inner product space, $\ell_2(\mathbb{Z}/N\mathbb{Z})$. Let

$$e_p(k) = \exp(2\pi i p k/N) \ , \quad p, k \in \mathbb{Z}/N\mathbb{Z} \ .$$

Claim 2.1.

1. $e_p : \mathbb{Z}/N\mathbb{Z}$ are (the) characters of $\mathbb{Z}/N\mathbb{Z}$, i.e. $e_p(k+\ell) = e_p(k)e_p(\ell)$.
2. $\frac{1}{\sqrt{N}} e_p$ form an orthonormal basis of $\ell_2(\mathbb{Z}/N\mathbb{Z})$.

Proof. Item 1. is obvious, item 2. is verified directly:

$$\frac{1}{N}\langle e_p, e_q\rangle = \frac{1}{N}\sum_{k=0}^{N-1} \exp(2\pi i k(p-q)/N) = \begin{cases} 1, & p \equiv q \mod N \\ 0 \end{cases} \qquad \square$$

Consider the (left) shift $S : \ell_2(\mathbb{Z}/N\mathbb{Z}) \to \ell_2(\mathbb{Z}/N\mathbb{Z})$, $(Sf)(k) = f(k+1)$. From item 1.

$$(Se_p)(k) = e_p(k+1) = e_p(1)e_p(k) \ , \quad \text{i.e.} \quad Se_p = e_p(1)e_p \ .$$

Theorem 2.2. *If $A : \ell_2(\mathbb{Z}/N\mathbb{Z}) \to \ell_2(\mathbb{Z}/N\mathbb{Z})$ is a linear map that commutes with S, i.e. $AS = SA$, then A is diagonal in the basis (e_p).*

Proof. We have: $SAe_p = ASe_p = e_p(1)Ae_p$, hence Ae_p is an eigenvector of S with eigenvalue $e_p(1)$, i.e. a multiple $\lambda_p e_p$ of e_p. $\square$

Exercise 2.3. Let $A : \ell_2(\mathbb{Z}/N\mathbb{Z}) \to \ell_2(\mathbb{Z}/N\mathbb{Z})$ be a linear map.

1. The map A commutes with S iff it is a convolution operator, i.e. there exists $r : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ such that

$$(Af)(k) = (r * f)(k) = \sum_{\ell=0}^{N-1} r(k-\ell)f(\ell) \ . \tag{2.1}$$

2. If A is of the form (2.1), the eigenvalue corresponding to e_p equals

$$\lambda_p = \sum_{k=0}^{N-1} r(k)\overline{e_p(k)} = \langle r, e_p\rangle = \hat{r}(p)$$

in the notation of

Definition 2.4. *The Fourier coefficients of $f \in \ell_2(\mathbb{Z}/N\mathbb{Z})$ are the numbers*

$$\hat{f}(p) = \langle f, e_p\rangle = \sum_{k=0}^{N-1} f(k)\exp(-2\pi i p k/N) \ ;$$

the map $\mathcal{F} : f \mapsto \hat{f}$ is called the discrete Fourier transform.

Remark 2.5. If we let $L_2(\mathbb{Z}/N\mathbb{Z})$ be the space of N-periodic functions equipped with the inner product

$$\langle\!\langle \phi, \psi \rangle\!\rangle = \frac{1}{N} \sum_{p=0}^{N-1} \phi(p)\overline{\psi}(p) \ ,$$

then $\mathcal{F} : \ell_2(\mathbb{Z}/N\mathbb{Z}) \to L_2(\mathbb{Z}/N\mathbb{Z})$ is an isometry. To keep notation consistent, we shall denote functions of k by Roman letters $f, g, \cdots$, and functions of p – by Greek letters $\phi, \psi, \cdots$.

Remark 2.6. By definition, $\hat{f}(p)$ measures the correlation between f and e_p. Informally, this coefficient is responsible for the correlation between f and patterns of period q, where $pq \equiv 1$ modulo N.

Exercise 2.7. The discrete Fourier transform boasts the following properties:

1. the inverse transform $\mathcal{F}^* : L_2(\mathbb{Z}/N\mathbb{Z}) \to \ell_2(\mathbb{Z}/N\mathbb{Z})$ is given by

$$\mathcal{F}^* : \phi \mapsto \check{\phi}, \ \check{\phi}(k) = \frac{1}{N} \sum_{k=0}^{N-1} \phi(p) e_k(p) \ .$$

2. Let $(f * g)(k) = \sum f(k-\ell)g(\ell)$, $(\phi \circledast \psi)(p) = \frac{1}{N}\sum \phi(p-q)\psi(q)$. Then $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$ and $\mathcal{F}(fg) = \mathcal{F}(f) \circledast \mathcal{F}(g)$.

3. For $f \equiv \frac{1}{N}$ one has $\hat{f}(p) = N\delta_{p0}$, and for $g(k) = \delta_{k0}$ one has $\hat{g}(p) \equiv 1$.

2.2 Roth's theorem

In this section we describe the following, somewhat sophisticated application.

Theorem 2.8 (Roth [1952]). *For any $\delta \in (0,1]$ there exists $N_0(\delta)$ such that for any $N \geq N_0(\delta)$ any $A \subset \{1, \cdots, N\}$ of cardinality $|A| \geq \delta N$ contains a three-term arithmetic progression.*

See Gowers [1998] and Soundararajan [2010] for a survey of the field including the more recent developments.

Proof. The property of being an arithmetic progression is translation invariant. Therefore it is not unexpected that Fourier analysis can be used. However, a few preliminary reductions have to be made. For convenience, denote the assertion of the theorem ("there exists $N(\delta)$ ... arithmetic progression") by Roth_δ.

First, Roth_δ is clearly true for $\delta > 2/3$ (in this case A, if large enough, contains three consecutive integers). Therefore it suffices to show that

$$\mathrm{Roth}_{\delta(1+\delta/25)} \Longrightarrow \mathrm{Roth}_\delta \ .$$

Second, it suffices to consider prime N. Formally, if, for a certain $\delta' > 0$ and a certain $N_0'(\delta')$, one has that for any prime $N' \geq N_0'(\delta)$ any set $A \subset \{1, \cdots, N'\}$ of cardinality $|A| \geq \delta' N$ contains a three-term arithmetic progression, then Roth_δ holds for all $\delta > \delta'$. Indeed, by Bertrand's postulate there is a prime

$$\lfloor (\delta - \delta') N / 2 \rfloor \leq N' \leq 2 \lfloor (\delta - \delta') N / 2 \rfloor \ .$$

Partitioning $\{1, \cdots, N' \lfloor N/N' \rfloor\}$ into $\lfloor N/N' \rfloor$ chunks C_j of size N', we find that for least one of the chunks $|A \cap C_j| \geq \delta' N'$, and hence $A \cap C_j$ contains a three-term arithmetic progression.

Third, to make use of the Fourier transform, we shall work modulo N. To rule out "wraparound" (i.e. to make sure that the three-term progression we will have found is not, say, $N-4, N-1, 2$), we consider several cases. If

$$|\{a \in A \, : \, a \leq N/3\}| > \delta(1 + \delta/20)|A|/3 \ , \tag{2.2}$$

the assumption $\mathrm{Roth}_{\delta(1+\delta/25)}$ ensures that there is an arithmetic progression in $A_1 = \{a \in A : \leq N_1 = \lfloor N/3 \rfloor\}$. We can argue similarly if

$$|\{a \in A \, : \, a > 2N/3\}| > \delta(1 + \delta/20)|A|/3 \ . \tag{2.3}$$

Therefore we assume for the rest of the proof that (2.2) and (2.3) fail, i.e. that

$$B = \{a \in A \, : \, N/3 < a \leq 2N/3\}$$

is of cardinality $|B| \geq \delta(1 - \delta/10)|A|/3$. Observe that if $a \in A$ and $b, c \in B$ are such that $a + c \equiv 2b \mod N$, then either (a, b, c) is a proper arithmetic progression (without wraparound) or $a = b = c$.

Now the Fourier-analytic part begins. The triples (a, b, c) with $a + c \equiv 2b \mod N$ can be counted as follows:

$$\begin{aligned} S &= \sum_{a \in A, b, c \in B} \mathbb{1}_{a+c=2b} = \sum_{a,b,c=1}^{N} \mathbb{1}_A(a) \mathbb{1}_B(b) \mathbb{1}_B(c) \mathbb{1}_{a+c=2b} && (2.4) \\ &= \frac{1}{N} \sum_{a,b,c=1}^{N} \sum_{p=0}^{N-1} \mathbb{1}_A(a) \mathbb{1}_B(b) \mathbb{1}_B(c) \exp(-2\pi i p (a - 2b + c)/N) && (2.5) \\ &= \frac{1}{N} \sum_{p=0}^{N-1} \hat{\mathbb{1}}_A(p) \hat{\mathbb{1}}_B(-2p) \hat{\mathbb{1}}_B(p) \ . && (2.6) \end{aligned}$$

Observing that $\hat{\mathbb{1}}_A(0) = |A|$ and similarly for B, C, and using the Cauchy–Schwarz inequality and that N is odd, we get:

$$S \geq \frac{1}{N}|A|\,|B|^2 - \max_{p\neq 0}\frac{1}{N}|\hat{\mathbb{1}}_A(p)| \times \sum_{p=1}^{N-1}|\hat{\mathbb{1}}_B(-2p)\hat{\mathbb{1}}_B(p)| \tag{2.7}$$

$$\geq \frac{1}{N}|A|\,|B|^2 - \max_{p\neq 0}\frac{1}{N}|\hat{\mathbb{1}}_A(p)| \times \sum_{p=1}^{N-1}|\hat{\mathbb{1}}_B(p)|^2 \tag{2.8}$$

$$= \frac{1}{N}|A|\,|B|^2 - \max_{p\neq 0}|\hat{\mathbb{1}}_A(p)|\,|B|\ . \tag{2.9}$$

This may be understood as follows. Among the $\sim N$ arithmetic progressions (a,b,c), $\sim |A|/N$ have $a \in A$, and similarly $\sim |B|/N$ have $b \in B$ or $c \in B$. If these events would be independent (which is the case if A is random), we would have $S \sim N^3 \times (|A|/N) \times (|B|/N)^2 = |A||B|^2/N$. If the non-zero Fourier coefficients of $\mathbb{1}_A$ are small, A has no distinguished patterns, and therefore we expect this to be a good approximation.

Proceeding with the argument, consider two cases. If $\max_{p\neq 0}|\hat{\mathbb{1}}_A(p)| \leq \delta|A|/6$, then

$$(2.9) \geq |A||B|(|B|/N - \delta/6) \geq |A||B|(\delta(1-\delta/10)/3 - \delta/6) \geq c(\delta)N^2\ ,$$

which is much greater than N, and we are done.

Now assume that there exists $p \neq 0$ such that $|\hat{\mathbb{1}}_A(p)| > \delta|A|/6$. (Recall that this means that A is correlated with some pattern of period p.) In this case, we shall decompose $\{1,\cdots,N\}$ into several arithmetic progressions, so that the density of A in at least one of them is at least $\delta(1+\delta/50)$. Formally, let $m = \lceil\frac{100}{\delta}\rceil$, and let

$$P_j = \{0 \leq x \leq N-1 \,:\, \frac{jN}{m} \leq px < \frac{(j+1)N}{m}\}\ , \quad 0 \leq j < m\ .$$

These are arithmetic progressions modulo N: indeed, if r be such that $pr \equiv 1 \mod N$, then

$$P_j = \{kr \,:\, \frac{jN}{m} \leq k < \frac{(j+1)N}{m}\}.$$

Then

$$|\hat{\mathbb{1}}_A(p)| = |\sum_j \sum_{x\in P_j}\mathbb{1}_A(x)\exp(-2\pi i p x/N)| \tag{2.10}$$

$$\leq |\sum_j \sum_{x\in P_j}\mathbb{1}_A(x)\exp(-2\pi i j/m)| + \frac{2\pi}{m}\sum_j\sum_{x\in P_j}\mathbb{1}_A(x) \tag{2.11}$$

$$= \left|\sum_j |A\cap P_j|\exp(-2\pi i j/m)\right| + \frac{2\pi|A|}{m}\ . \tag{2.12}$$

Hence

$$\left|\sum_j |A \cap P_j| \exp(-2\pi i j/m)\right| \geq |\hat{\mathbb{1}}_A(p)| - \frac{2\pi|A|}{m} \geq \frac{\delta|A|}{12} \ . \tag{2.13}$$

If all the sets $A \cap P_j$ were of the same size, the sum would vanish. Quantitatively, we have:

Exercise 2.9. Prove: if $f : \mathbb{Z}/m\mathbb{Z} \to \mathbb{R}$ and $\bar{f} = \frac{1}{m}\sum_{j=0}^{m-1} f(j)$, then

$$\max_{j=0}^{m-1} f(j) \geq \bar{f} + \frac{1}{2m}\left|\sum_{j=0}^{m-1} f(j) \exp(-2\pi i j/m)\right| \ . \tag{2.14}$$

Applying (2.14) to $f(j) = |A \cap P_j|$, we obtain that there exists j such that

$$|A \cap P_j| \geq \frac{|A|}{m} + \frac{\delta|A|}{24m} \geq \delta(1 + \delta/25)|P_j| \ .$$

Now we further decompose this P_j into proper arithmetic progressions. Among the $\lfloor\sqrt{N}\rfloor$ numbers $r, 2r, \cdots, \lfloor\sqrt{N}\rfloor r$, there are two which are not far from one another: $r^* = k_2 r - k_1 r \mod N \in \{1, \cdots, \lfloor\sqrt{N}\rfloor\}$. Let $k^* = k_2 - k_1$; note that $k^* \leq \sqrt{N}$. For each $1 \leq k \leq k^*$, split P_j into k^* arithmetic progressions of step r^* (still possibly with wraparound); and then split each of these into proper arithmetic progressions without wraparound. All but at most $C\sqrt{N}$ of these progressions have length $\geq c\sqrt{N}$, hence at least one of them, $P'_j \subset P_j$, satisfies $|P'_j| \geq c\sqrt{N}$, $|A \cap P'_j| \geq \delta(1+\delta/25)|P'_j|$. By the induction assumption $A \cap P'_j$ contains a three-term arithmetic progression. □

Exercise 2.10. Show that

1. $N_0(\delta) \leq \exp\exp(C/\delta)$;
2. the assumption $|A| \geq \delta N$ can be replaced with $|A| \geq C_1 N/\log\log N$, for a sufficiently large constant C_1.

2.3 Fast Fourier transform*

The naïve implementation of the discrete Fourier transform on $\mathbb{Z}/N\mathbb{Z}$ requires $\sim \mathrm{const}\, N^2$ arithmetic operations. Fast Fourier transform allows to do the job with $O(N \log N)$ operations. See Montgomery [2014] for a historical discussion.

Denote the smallest number of arithmetic operations required for the computation of the Fourier transform by $\mathfrak{n}(N)$.

Proposition 2.11. $\mathfrak{n}(2N) \leq 2\mathfrak{n}(N) + 3N$.

Proof. Let $f : \mathbb{Z}/2N\mathbb{Z} \to \mathbb{C}$. Construct two auxiliary functions $g, h : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$,

$$g(n) = f(2n) \ , \ h(n) = f(2n+1) \ .$$

Then

$$\begin{aligned}
\hat{f}(p) &= \sum_{n=0}^{2N-1} f(n) \exp(-2\pi i np/(2N)) \\
&= \sum_{n=0}^{N-1} f(2n) \exp(-2\pi i \, 2np/(2N)) + \sum_{n=0}^{N-1} f(2n+1) \exp(-2\pi i \, (2n+1)p/(2N)) \\
&= \hat{g}(p) + \exp(-\pi i p/N) \hat{h}(p) \ ,
\end{aligned}$$

which means that for $0 \le p \le N-1$

$$\hat{f}(p) = \hat{g}(p) + \exp(-\pi i p/N)\hat{h}(p) \ , \quad \hat{f}(p+N) = \hat{g}(p) - \exp(-\pi i p/N)\hat{h}(p) \ . \tag{2.15}$$

The computation of $\hat{f}$ from $\hat{g}$ and $\hat{h}$ using (2.15) requires $3N$ arithmetic operations (if we assume that the values of $\exp(-\pi i p/N)$ can be precomputed). □

Corollary 2.12. $\mathfrak{n}(2^n) \le (3/2)n2^n$ *for* $n \ge 1$.

Proof. By induction. □

This corollary implies that $\mathfrak{n}(N) \le CN \log N$ when N is a power of 2. Similar reasoning applies to N which do not have large prime factors. The general case requires a separate argument for large prime N, which we do not reproduce here.

Applications From item 2 of Exercise 2.7 we obtain that the convolution $f * g = (\hat{f}\hat{g})^\vee$ of two functions $f, g : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ can be computed in $O(\mathfrak{n}(N))$ steps; for $N = 2^n$, this is $O(N \log N)$.

Next, $p(x) = \sum_{k=0}^d a_k x^k$ and $q(x) = \sum_{k=0}^d b_k x^k$ be two polynomials. Their product is equal to $(pq)(x) = \sum_{k=0}^{2d} (\sum_{l=0}^k a_l b_{k-l}) x^k$. This looks quite similar to a convolution, except that there is no wraparound.

Exercise 2.13. The product of two polynomials of degree d can be computed in $O(d \log d)$ arithmetic operations.

Exercise 2.14. Long multiplication of two d-digit numbers takes $O(d^2)$ operations (where an operation means addition or multiplication of single-digit numbers). Propose an algorithm which would only require $O(d \log^{10} d)$ operations.

3 Fourier series

3.1 Fourier coefficients and Plancherel theorem

From the finite group $\mathbb{Z}/N\mathbb{Z}$ we pass to the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (defined in the natural way, e.g., $[1/3] + [3/4] = [1/12]$). We identify functions on the torus with 1-periodic functions of a real variable.

A *character* of $\mathbb{T}$ is a continuous function $\chi : \mathbb{T} \to \mathbb{C}^\times$ such that

$$\chi(x+y) = \chi(x)\chi(y) \ . \tag{3.1}$$

Exercise 3.1. Any character of $\mathbb{T}$ is of the form $e_p(x) = \exp(2\pi i p x)$ for some $p \in \mathbb{Z}$.

The goal is to expand f in a series in e_p. The questions of convergence are delicate, as we shall see. We start with the discussion of mean-square convergence, for which the theory is simplest. We think of $\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}dx$ as an inner product, however, we should keep in mind that, if we work with Riemann integrals, the space of square-integrable functions is not complete. We shall focus on piecewise-continuous functions, i.e. $[0,1]$ is a finite union of closed intervals such that the function is uniformly continuous in the interior of each of them, and try to extract the most of the pre-Hilbert point of view. (Ultimately, a better approach would be to work with Lebesgue integrals and enjoy the benefits of a proper Hilbert space.)

The following exercise shows that e_p are orthonormal:

Exercise 3.2. $\int_0^1 e_p(x)\overline{e_q(x)}dx = \begin{cases} 1 \ , & p = q \\ 0 \ . \end{cases}$

This leads us to

Definition 3.3. *Let $f : \mathbb{T} \to \mathbb{C}$ be piecewise continuous. The numbers $\hat{f}(p) = \langle f, e_p\rangle$ are called the Fourier coefficients of f, and the formal*[2] *series $\sum_{p=-\infty}^{\infty} \hat{f}(p)e_p$ is called the Fourier series representing f.*

The discussion of convergence starts with

Theorem 3.4 (Plancherel). *Let $f : \mathbb{T} \to \mathbb{C}$ be piecewise continuous. Then*

$$\sum_{p=-\infty}^{\infty} |\hat{f}(p)|^2 = \int_0^1 |f(x)|^2 dx \ .$$

Proof. One direction holds for any orthonormal system.

[2]here *formal* refers to manipulations with formulæ, disregarding questions of convergence.

Claim 3.5. Let H be an inner product space, and let $(u_\alpha)_{\alpha\in A}$ be an orthonormal system of vectors. Then for any $f \in H$

$$\sum_{\alpha\in A} |\langle f, u_\alpha\rangle|^2 \leq \|f\|^2 = \langle f, f\rangle \ .$$

Proof of Claim. Observe that for any finite $A' \subset A$

$$0 \leq \|f - \sum_{\alpha\in A'} \langle f, u_\alpha\rangle u_\alpha\|^2 = \|f\|^2 - \sum_{\alpha\in A} |\langle f, u_\alpha\rangle|^2 \leq \|f\|^2 \ . \qquad \square$$

The inequality $\leq$ of the theorem follows by applying the claim to the space of piecewise continuous functions. The other direction, $\geq$, requires rolling up the sleeves. Let f be piecewise continuous, and let $M = \max|f|$. Fix $\epsilon > 0$, and choose $f_1 \in C(\mathbb{T})$ such that $\max|f_1| = M$ and $f = f_1$ outside a union of intervals of total length $\leq \epsilon$. Now choose[3] a trigonometric polynomial $f_2 = \sum_{p=-N}^{N} c_p e_p$ such that $\max|f_1 - f_2| \leq \epsilon$. Then (still using the notation $\|f\|^2 = \langle f, f\rangle$)

$$\|f\| \leq \|f_2\| + \|f_1 - f_2\| + \|f - f_1\| \leq \|f_2\| + \epsilon + 2M\sqrt{\epsilon}$$

whereas (similarly)

$$\sqrt{\sum_{p=-N}^{N} |\hat{f}(p)|^2} \geq \sqrt{\sum_{p=-N}^{N} |\hat{f}_2(p)|^2} - \epsilon - 2M\sqrt{\epsilon} = \|f_2\| - \epsilon - 2M\sqrt{\epsilon} \ ,$$

i.e.

$$\sqrt{\sum_{p=-\infty}^{\infty} |\hat{f}(p)|^2} \geq \|f\| - 2(\epsilon + 2M\sqrt{\epsilon})$$

for any $\epsilon > 0$. $\square$

The Plancherel theorem implies mean-square convergence of Fourier series:

Exercise 3.6. If $f : \mathbb{T} \to \mathbb{C}$ is piecewise continuous, then

$$\lim_{N\to\infty} \int_0^1 |f(x) - \sum_{p=-N}^{N} \hat{f}(p) e_p(x)|^2 dx = 0 \ . \qquad (3.2)$$

We remind that (3.2) does not imply pointwise convergence, and in fact does not even imply the existence of one point at which the series converges. We shall return to the question of pointwise convergence later.

Here is another corollary of the Plancherel theorem:

Corollary 3.7. *If $f, g : \mathbb{T} \to \mathbb{C}$ are piecewise continuous and $\hat{f} \equiv \hat{g}$, then $f(x) = g(x)$ at all the continuity points of f.*

[3] We shall prove the Weierstrass approximation theorem in Section 4.1, without circular reasoning.

Exercise 3.8. If $f, g : \mathbb{T} \to \mathbb{C}$ are piecewise continuous, then

$$\sum_{p=-\infty}^{\infty} \hat{f}(p)\overline{\hat{g}(p)} = \langle f, g \rangle \ .$$

Let us see a couple of applications before we move on. More applications of Fourier series will be discussed in the next section.

The value of $\zeta(2)$. Let $f(x) = x$ (for $x \in (0,1)$), so that $\|f\|^2 = \frac{1}{3}$. The Fourier coefficients of f can be computed as follows:

$$\begin{aligned} \hat{f}(p) &= \int_0^1 x \exp(-2\pi i p x) dx = -\frac{1}{2\pi i p} \int_0^1 x \frac{d}{dx} \exp(-2\pi i p x) dx \\ &= -\frac{1}{2\pi i p} x \exp(-2\pi i p x)\Big|_0^1 + \frac{1}{2\pi i p} \int_0^1 \exp(-2\pi i p x) dx = -\frac{1}{2\pi i p} \end{aligned} \tag{3.3}$$

for $p \neq 0$, and $\hat{f}(0) = \frac{1}{2}$. Hence

$$\frac{1}{4} + 2 \sum_{p=1}^{\infty} \frac{1}{4\pi^2 p^2} = \frac{1}{3} \ , \quad \zeta(2) = \sum_{p=1}^{\infty} \frac{1}{p^2} = \frac{\pi^2}{6} \ .$$

Exercise 3.9. Compute $\zeta(4) = \sum_{p=1}^{\infty} p^{-4}$.

Poincaré inequality Here is another application:

Exercise 3.10. Prove that for any reasonable (differentiable with piecewise continuous derivative) $f : \mathbb{T} \to \mathbb{C}$,

$$\int_0^1 |f(x) - \bar{f}|^2 dx \leq \frac{1}{4\pi^2} \int_0^1 |f'(x)|^2 dx \ , \quad \text{where} \quad \bar{f} = \int_0^1 f(x) dx \ . \tag{3.4}$$

The right-hand side of (3.4) measures the global magnitude of fluctuations (it is the variance of $f(X)$, where $X \in \mathbb{T}$ is chosen uniformly at random), whereas the right-hand side measures the local fluctuations. To appreciate the meaning of such estimates, it may be useful to consider the following d-dimensional generalisation of (3.4): for any reasonable $f : \mathbb{T}^d \to \mathbb{C}$,

$$\int_{\mathbb{T}^d} |f(x) - \bar{f}|^2 dx \leq \frac{1}{4\pi^2} \int_{\mathbb{T}^d} \|\nabla f(x)\|^2 dx \ , \quad \text{where} \quad \bar{f} = \int_{\mathbb{T}^d} f(x) dx \tag{3.5}$$

(and $\|\nabla f\|^2 = \sum_{j=1}^d |\frac{\partial f}{\partial x_j}|^2$). In particular, if f satisfies the Lipschitz estimate

$$|f(x) - f(y)| \leq \operatorname{dist}(x, y) \ ,$$

where dist is the usual Euclidean distance on the d-dimensional torus, then the variance of $f(X)$ (where $X \in \mathbb{T}^d$ is chosen uniformly at random) is bounded by $1/(4\pi^2)$. This estimate is much better than the naïve bound

$$\int_{\mathbb{T}^d} |f(x) - \bar{f}|^2 dx \leq \operatorname{diam}^2 \mathbb{T}^d = \frac{d}{4} ,$$

which grows with the dimension (it is instructive to compare the two bounds for the simple special case $f(x) = \sum_{j=1}^d f_j(x_j)$). This is an instance of the *concentration phenomenon* in high dimension, put forth in the 1970s by V. Milman. See Giannopoulos and Milman [2000], Ledoux [2001] and references therein.

3.2 Convergence of Fourier series

We return to the main narrative. Let $f : \mathbb{T} \to \mathbb{C}$ be piecewise continuous. We are interested in the convergence of partial sums

$$\sum_{p=-N}^{N} \hat{f}(p) e_p(x) \tag{3.6}$$

to $f(x)$. We first evaluate:

$$\begin{aligned}(3.6) &= \sum_{p=-N}^{N} \int f(y)\bar{e}_p(y)dy \; e_p(x) \\ &= \int f(y) \left(\sum_{p=-N}^{N} e_p(x-y) \right) dy = (D_N * f)(y) ,\end{aligned}$$

where

$$D_N(x) = \sum_{p=-N}^{N} e_p(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)} \tag{3.7}$$

is the Dirichlet kernel (see Figure 4.1).

Lemma 3.11 (Riemann–Lebesgue)**.** *Let $f : \mathbb{T} \to \mathbb{C}$ be piecewise continuous. Then $\hat{f}(p) \to 0$ as $p \to \pm\infty$.*

The intuitive explanation may go as follows. If a function is constant, $\hat{f}(p) = 0$ for all $p \neq 0$. A typical function, e.g. $f(x) = \operatorname{dist}(x, \mathbb{Z})$ is not constant, however, for large p it is almost constant on scale $1/p$, which means that the p-th Fourier coefficient should be small. The following argument formalises this idea.

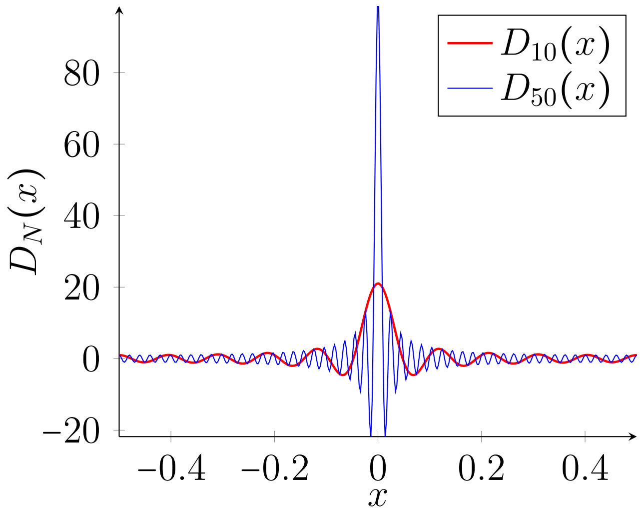

Figure 3.1: The Dirichlet kernel for $N = 10, 50$

Proof. We perform a change of variables $x \to x + \frac{1}{2p}$:

$$\begin{aligned}
\hat{f}(p) &= \int f(x)\bar{e}_p(x)dx \\
&= \int f(x + \frac{1}{2p})\bar{e}_p(x + \frac{1}{2p})dx \\
&= \int f(x)\bar{e}_p(x + \frac{1}{2p})dx + \int (f(x) - f(x + \frac{1}{2p}))\bar{e}_p(x + \frac{1}{2p})dx \ .
\end{aligned}$$

The first term on the right-hand side is equal to the negative of the left-hand side. The second term is bounded by

$$\left|\int (f(x) - f(x + \frac{1}{2p}))\bar{e}_p(x + \frac{1}{2p})dx\right| \le \int |f(x) - f(x + \frac{1}{2p})|dx \ ,$$

therefore

$$|\hat{f}(p)| \le \frac{1}{2}\int |f(x) - f(x + \frac{1}{2p})|dx$$

and this tends to zero as $p \to \pm\infty$ (why?). □

Now we can prove:

Proposition 3.12. *Let $f : \mathbb{T} \to \mathbb{C}$ be piecewise continuous, and suppose $x_0 \in \mathbb{T}$ is such that there exist $a \in (0, 1]$ and $C > 0$ for which*

$$|f(x) - f(x_0)| \le C|x - x_0|^a \ . \tag{3.8}$$

Then $(D_N * f)(x_0) \to f(x_0)$ as $N \to \infty$.

Proof. Without loss of generality f is real-valued, and $x_0 = 0$, $f(x_0) = 0$. Decompose

$$\int f(x)D_N(x)dx = \int_{|x|\le\delta} + \int_{\delta\le x\le\frac{1}{2}} .$$

Recalling that $|\sin(\pi x)| \ge 2|x|$ for $|x| \le \frac{1}{2}$, the first integral is bounded by

$$\Big|\int_{|x|\le\delta}\Big| \le \int_{|x|\le\delta} C|x|^a \frac{dx}{2|x|} \le C'\delta^a$$

The second integral is equal to

$$\int_{\delta\le x\le\frac{1}{2}} = \int_{-\frac{1}{2}}^{\frac{1}{2}} [f_1(x)\sin(2N\pi x) + f_2(x)\cos(2N\pi x)]\, dx , \tag{3.9}$$

where $f_1(x) = f(x)\cot(\pi x)\mathbb{1}_{|x|\ge\delta}$ and $f_2(x) = f(x)\mathbb{1}_{|x|\ge\delta}$. These two functions are piecewise continuous (for any $\delta > 0$), hence the Riemann–Lebesgue lemma implies that (3.9) $\to 0$ as $N \to \infty$. It remains to let $\delta \to +0$. $\square$


(Pisa, Italy)

Exercise 3.13. Amplify Proposition 3.12 as follows: if $f : \mathbb{T} \to \mathbb{C}$ is piecewise continuous and there exists $A \in \mathbb{C}$ such that

$$\int_0^{\frac{1}{2}} \frac{|f(x_0 + t) + f(x_0 - t) - 2A|}{|t|} dt < \infty , \tag{3.10}$$

then $(D_N * f)(x_0) \to A$ as $N \to \infty$. If f satisfies (3.8), then (3.10) holds with $A = f(x_0)$. What happens if f has a jump continuity at x_0, e.g. $f = \mathbb{1}_{[0,1/2]}$? What is the connection to the picture on the left?

Another important corollary of the Riemann–Lebesgue lemma is the Riemann localisation principle, which asserts that the convergence of the Fourier series at a point only depends on the behaviour of the function in the vicinity of this point.

Exercise 3.14. If $f, g : \mathbb{T} \to \mathbb{C}$ are piecewise continuous functions that coincide in a neighbourhood of x_0, then $(D_N * f)(x_0) \to A$ iff $(D_N * g)(x_0) \to A$.

Remark 3.15. The Riemann–Lebesgue lemma does not provide any quantitative rate of decay. This can not be improved without introducing extra assumptions:

Exercise 3.16. Let $(\epsilon(p))_{p\ge1}$ be a sequence of positive numbers tending to zero as $p \to \infty$. Show that one can choose $1 \le p_1 < p_2 < \cdots$ such that

$$f(x) = \sum_{p=1}^{\infty} \epsilon(p_j) e_{p_j}(x)$$

is a continuous function. Clearly, $\epsilon(p)^{-1}|\hat{f}(p)| \not\to 0$ as $p \to \infty$.

Divergence If f is only assumed to be continuous, the Fourier series may diverge. That is, f is not the pointwise sum of its Fourier series.

Proposition 3.17 (Du Bois-Reymond)**.** *There exists $f \in C(\mathbb{T})$ such that $(D_N * f)(0)$ diverges as $N \to \infty$.*

The proof is based on the estimate

$$\begin{aligned}\int |D_N(x)|dx &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{|\sin((2N+1)\pi x)|}{|\sin \pi x|}dx \\ &> 2\sum_{k=1}^{N} \int_{\frac{k-1}{2N+1}}^{\frac{k}{2N+1}} \frac{|\sin((2N+1)\pi x)|}{|\sin \pi x|}dx \\ &\geq 2\sum_{k=1}^{N} \int_{\frac{k-1}{2N+1}}^{\frac{k}{2N+1}} \frac{|\sin((2N+1)\pi x)|}{\pi k/(2N+1)}dx \geq c\log N \to \infty \ .\end{aligned} \tag{3.11}$$

In particular, one can find a continuous function f_N with $\sup|f_N| \leq 1$ and $\int D_N(x)f_N(x)dx \to \infty$, as $N \to \infty$.[4]

The idea is to construct f_N which will approximate $\mathrm{sign}(D_N)$ along a subsequence of N, and such that the Fourier coefficients of f_N are small when p is far from N, and then make a convergent series of of f_{N_k} (when N_k are sparse enough to avoid interference between different F_{N_k}). This is implemented as follows.

Proof. Let

$$f_N(x) = \sin((2N+1)\pi x) \ , \quad 0 \leq x < 1 \ ,$$

and continue it periodically.

Exercise 3.18. Prove that $\int D_N(x)f_N(x)dx \geq c_1 \log N$.

For M not very close to M, $\int D_N(x)f_M(x)dx$ is not very large, as made quantitative by

Exercise 3.19. $|\int D_N(x)f_M(x)dx| \leq \begin{cases} CM/N \ , & M \leq N/2 \\ CN/M \ , & M \geq 2N \end{cases}.$

Now set

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k^2} f_{N_k}(x) \ ,$$

where we shall take $N_k = 3^{k^3}$. According to Exercises 3.18 and 3.19,

$$(D_{N_k} * f)(0) \geq \frac{c_1 \log N_k}{k^2} - \sum_{j=1}^{k-1} \frac{1}{j^2}\frac{CN_j}{N_k} - \sum_{j=k+1}^{\infty} \frac{1}{j^2}\frac{CN_k}{N_j} \geq c_2 k \ . \qquad \square$$

[4] One can prove the theorem using this but property of D_N, if one relies on the Banach–Steinhaus theorem from functional analysis. The advantage of the more classical proof reproduced below is that it gives an explicit estimate on the rate of divergence.

3.3 Féjér summation

The negative result of the previous paragraph is somewhat surprising, given that by Weierstrass theorem any continuous function on $\mathbb{T}$ can be uniformly approximated by trigonometric polynomials. Of course, $D_N * f$ is the orthogonal projection of f onto the space of trigonometric polynomials of degree $\le N$, i.e. it is the trigonometric polynomial P_N of degree N which minimises $\int |P_N - f|^2 dx$. This means that the best approximation in the norm $\|\cdot\| = \|\cdot\|_2$ may not be a very good approximations in the maximum norm. However, good uniform approximations can be constructed from $D_N * f$ using the following simple procedure.

Theorem 3.20 (Féjér). *If $f \in C(\mathbb{T})$, then*

$$\frac{1}{N}\sum_{n=0}^{N-1}(D_n * f) = \sum_{p=-N+1}^{N-1}(1 - |p|/N)\hat{f}(p)e_p \rightrightarrows f \quad \text{as } N \to \infty.$$

The arithmetic mean appearing in the left-hand side is known as Cesàro summation. If $(a_n)_{n\ge 0}$ is a sequence of numbers, one constructs the Cesàro means $A_N = \frac{1}{N}\sum_{n=0}^{N-1} a_n$.

Exercise 3.21. If (a_n) converges to a limit L, then also $A_N \to L$.

This procedure is useful since there are divergent sequences, such as $a_n = (-1)^n$, for which the Cesàro means converge; the limit may be considered as a generalised limit of (a_n). It is remarkable that such a simple resummation procedure works simultaneously for all continuous functions.

Proof of Theorem 3.20. We first observe that $\frac{1}{N}\sum_{n=0}^{N}(D_n * f) = S_N * f$, where

$$S_N(x) = \frac{1}{N}\sum_{n=0}^{N-1} D_n(x) = \frac{1}{N}\sum_{n=0}^{N-1}\frac{\sin((2N+1)\pi x)}{\sin(\pi x)} = \frac{1}{N}\frac{\sin^2(N\pi x)}{\sin^2(\pi x)} .$$

The Féjér kernel S_N boasts the following properties:

1. $S_N \ge 0$;
2. $\int_0^1 S_N(x)dx = 1$;
3. For any $\delta > 0$, $\int_{\delta \le |x| \le \frac{1}{2}} S_N(x)dx \to 0$ as $N \to \infty$.

From these properties,

$$\begin{aligned} |f(x) - (S_N * f)(x)| &\underset{2.}{=} \left|\int (f(x) - f(x-y))S_N(y)dy\right| \\ &\underset{1.}{\le} \int |f(x) - f(x-y)|S_N(y)dy . \end{aligned}$$

Fix $\epsilon > 0$, and choose δ so that $|f(x) - f(x')| \leq \frac{\epsilon}{2}$ for $|x - x'| \leq \delta$, and then use 3. to choose N_0 so that for $N \geq N_0$

$$\int_{\delta \leq |y| \leq \frac{1}{2}} S_N(y) dy \leq \frac{\epsilon}{2 \max |f|} .$$

Then for $N \geq N_0$

$$\int |f(x) - f(x-y)| S_N(y) dy = \int_{|y| \leq \delta} + \int_{\delta \leq |y| \leq \frac{1}{2}} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon .$$

□

Remark 3.22. A kernel $S_N(x)$ satisfying 1.–3. [5] is called a non-negative summability kernel. A kernel is called a (general) summability kernel if

1*. $\sup_N \int |S_N(x)| dx < \infty$;

2*. $\int_0^1 S_N(x) dx = 1$;

3*. For any $\delta > 0$, $\int_{\delta \leq |x| \leq \frac{1}{2}} |S_N(x)| dx \to 0$ as $N \to \infty$.

Note that any non-negative summability kernel is a summability kernel.

Exercise 3.23. Show that the conclusion $S_N * f \rightrightarrows f$ of Féjér's theorem holds for any summability kernel S_N.

Here is one additional example of a summability kernel. We shall see another one in Section 4.1 and yet another one in Section 4.2.

Exercise 3.24. Prove the following.

1. The Poisson kernel

$$P_r(x) = \Re \left[\frac{1 + r \exp(2\pi i x)}{1 - r \exp(2\pi i x)} \right] = \frac{1 - r^2}{1 - 2r \cos(2\pi x) + r^2} , \quad r \in [0, 1)$$

satisfies that $(P_{1-\eta})_{\eta \to +0}$ is a non-negative summability kernel.

2. If $f \in C(T)$, then

$$u(re^{2\pi i x}) = (P_r * f)(x)$$

is continuous in the closed unit disk in $\mathbb{R}^2$ and satisfies $\Delta u = 0$ in the interior of the disk (i.e. it is harmonic).

[5] or more precisely, a family of kernels depending on an additional large parameter $N \to \infty$. To emphasise this, we may sometimes write $(S_N)_{N \to \infty}$. We shall also consider kernels $(S_\eta)_{\eta \to +0}$ depending on a small parameter $\eta \to +0$; the definitions are then adjusted in an obvious way

4 Fourier series: further topics and applications

4.1 Polynomial approximation

As we noted, Féjér's theorem implies the Weierstrass approximation theorem: for any $f \in C(\mathbb{T})$,

$$\lim_{N\to\infty} E_N(f) = 0 \ , \quad \text{where} \quad E_N(f) = \inf_{c_{-N},\cdots,c_N\in\mathbb{C}} \|f - \sum_{p=-N}^{N} c_p e_p\|_\infty \ . \tag{4.1}$$

Moreover, similar arguments yield good quantitative bounds. The following propositions are based on the work of Dunham Jackson, and are known as Jackson's theorems.

Theorem 4.1. *For any* $f \in C(\mathbb{T})$*,* $E_N(f) \le C\omega_f(\frac{1}{N})$*, where*

$$\omega_f(\delta) = \sup_{|x-x'|\le\delta} |f(x) - f(x')| \tag{4.2}$$

is the modulus of continuity.

Proof. Let $J_N(x)$ be a non-negative trigonometric polynomial of degree, say, $2N$ such that $\int J_N(x)dx = 1$. Then

$$\begin{aligned}|(J_N * f)(x) - f(x)| &\le \int_{-\frac{1}{2}}^{\frac{1}{2}} |f(x) - f(x-y)|J_N(y)dy \\ &= \sum_{j=1}^{N} \int_{\frac{j-1}{2N}\le|y|\le\frac{j}{2N}} |f(x) - f(x-y)|J_N(y)dy \ .\end{aligned}$$

The j-th integral is bounded by

$$\omega_f(j/(2N)) \int_{\frac{j-1}{2N}\le|y|\le\frac{j}{2N}} \Delta_N(y)dy \le j\omega_f(1/(2N)) \int_{\frac{j-1}{2N}\le|y|\le\frac{j}{2N}} \Delta_N(y)dy \ ,$$

hence

$$E_{2N}(f) \le \omega_f(1/(2N)) \sum_{j=1}^{N} j \int_{\frac{j-1}{2N}\le|y|\le\frac{j}{2N}} J_N(y)dy \ .$$

If we find J_N such that

$$\sup_N \sum_{j=1}^{N} j \int_{\frac{j-1}{2N}\le|y|\le\frac{j}{2N}} J_N(y)dy < \infty \ , \tag{4.3}$$

the sum to be bounded uniformly N. This condition is strictly stronger than property 3. of non-negative summability kernels: (4.3) implies 3. (why?), but not vice versa, since it does not hold for the Féjér kernel S_N (why?). Therefore we take the Jackson kernel

$$J_N(x) = \frac{3}{2N^3+N}\frac{\sin^4(\pi N x)}{\sin^4(\pi x)} = \frac{3N}{2N^2+1} S_N(x)^2 \ .$$

The faster decay ensures that (4.3) holds (why?) The second representation ensures that indeed this is a trigonometric polynomial of degree $2N$. The normalisation constant is chosen to ensure that the integral is one:

$$\begin{aligned}\int_{-\frac12}^{\frac12} S_N(x)^2 dx &= \sum_{p=-N+1}^{N-1} (1-|p|/N)^2 \\ &= 1 + \frac{2}{N^2}\sum_{p=1}^{N-1} p^2 \\ &= 1 + \frac{(N-1)(2N-1)}{3N} = \frac{1+2N^2}{3N} \ .\end{aligned}$$

Thus $E_{2N}(f) \le C\omega_f(1/(2N))$. □

Exercise 4.2. Prove: if $f \in C^k(\mathbb{T})$ for some $k \ge 1$, then $E_N(f) \le C_k N^{-k}\omega_{f^{(k)}}(\frac{1}{N})$.

Exercise 4.3. Let $f \in C(\mathbb{T})$. Then

$$\sup|f - D_N * f| \le C \log N \, \omega_f(1/N) \ .$$

In particular, we recover the Dini–Lipschitz criterion: $D_N * f \rightrightarrows f$ whenever $\omega_f(t)\log\frac{1}{t} \to 0$ as $t \to +0$.

Exercise 4.4. State and prove a version of Theorem 4.1 for the approximation of a function in $C[-1,1]$ by algebraic polynomials of degree $\le N$. You are welcome to rely on Theorem 4.1, if needed.

Constructive function theory In the early XXth century, Serge Bernstein put forth the idea that the analytic properties of a function, such as continuity and smoothness, are related to the rate of approximation by polynomials (or trigonometric polynomials). This circle of questions and results known as *constructive function theory*. For example, $f : \mathbb{T} \to \mathbb{C}$ is continuous if and only if $E_N(f) \to 0$. As a more sophisticated illustration, we prove the following converse to Jackson's theorem (Theorem 4.1):

Theorem 4.5 (Bernstein)**.** *Let $a \in (0,1)$. For any $f : \mathbb{T} \to \mathbb{C}$,*

$$\sup_{\delta \in (0,\frac12]} \omega_f(\delta)\delta^{-a} \le C_a \sup_{N\ge1} E_N(f) N^a \ , \quad C_a = 2\pi\frac{2^{1-a}+1}{2^{1-a}-1} \ , \tag{4.4}$$

where E_N are as in (4.1). In particular, the left-hand side is finite whenever the right-hand side is finite.

Remark 4.6. Note that we assume $a < 1$. For $a = 1$ the statement does not hold, and in fact $E_N(f) = O(1/N)$ iff

$$\sup_{x \in \mathbb{T},\, h \in (0, \frac{1}{2})} \frac{|f(x+h) + f(x-h) - 2f(x)|}{h} < \infty \ .$$

The proof relies on the following very useful inequality, also due to Bernstein.

Theorem 4.7 (Bernstein)**.** *Let $P(x) = \sum_{p=-N}^{N} c_p e_p(x)$ be a trigonometric polynomial of degree N. Then $\|P'\|_\infty \le 2\pi N \|P\|_\infty$.*

Proof of Theorem 4.7. (cf. Montgomery [2014]) We look for $2N$ coefficients $a_1, \cdots, a_{2N}$ (independent of P but dependent on N) such that

$$P'(0) = \sum_{k=1}^{2N} a_k P(\frac{2k-1}{4N}) \tag{4.5}$$

so that we can bound $|P'(0)| \le \sum_{k=1}^{2N} |a_k| \|P\|_\infty$. It suffices to have the identity (4.5) for $P = e_p$, $-N \le p \le N$:

$$2\pi i p = e_p'(0) = \sum_{k=1}^{2N} a_k e_p(\frac{2k-1}{4N}) = \sum_{k=1}^{2N} a_k e_p(\frac{k}{2N}) e_p(-\frac{1}{4N}) \ ,$$

which we further rewrite as

$$\sum_{k=1}^{2N} a_k e_p(\frac{k}{2N}) = 2\pi i p \, e_p(\frac{1}{4N}) \ , \quad -N \le p \le N \tag{4.6}$$

Note that the first and the last equations coincide, so we have $2N$ equations in $2N$ variables. Inverting the discrete Fourier transform in $\mathbb{Z}/(2N\mathbb{Z})$, we obtain:

$$a_k = -\frac{\pi i}{N} \sum_{p=-N}^{N} p e_p(\frac{2k-1}{4N}) + \pi(-1)^k \ ,$$

where the extra term appears since we insist on summing over $2N+1$ (rather than $2N$) values of p.

Exercise 4.8. Show that

$$a_k = \frac{(-1)^{k-1}\pi}{2N \sin^2(\frac{(2k-1)\pi}{4N})} \ .$$

This implies that

$$\sum_{k=1}^{2N} |a_k| = \sum_{k=1}^{2N} (-1)^{k-1} a_k = 2\pi N \ ,$$

where the last equality follows from (4.6) with $p = N$. □

Exercise 4.9. For any trigonometric polynomial $P \in \operatorname{span}(e_{-N}, \cdots, e_N)$,

$$\int_0^1 |P'(x)| dx \le 2\pi N \int_0^1 |P(x)| dx \ , \quad \int_0^1 |P'(x)|^2 dx \le 4\pi N^2 \int_0^1 |P(x)|^2 dx \ .$$

Now we can prove Theorem 4.5.

Proof of Theorem 4.5. Let P_{2^k} be a trigonometric polynomial of degree $\le 2^k$ such that $\|f - P_{2^k}\|_\infty \le A 2^{-ak}$. Then

$$|f(x) - f(y)| \le \frac{2A}{2^{ak}} + |P_{2^k}(x) - P_{2^k}(y)| \le \frac{2A}{2^{ak}} + \sum_{j=1}^{k} |(P_{2^j} - P_{2^{j-1}})(x) - (P_{2^j} - P_{2^{j-1}})(y)|$$

(Please convince yourself that the dyadic decomposition is needed, i.e. if we directly apply the Bernstein inequality to estimate $|P_{2^k}(x) - P_{2^k}(y)|$, the conclusion we get is much weaker than the claimed (4.4).)

The j-th term in the sum is bounded using Bernstein's inequality:

$$\begin{aligned} |(P_{2^j} - P_{2^{j-1}})(x) - (P_{2^j} - P_{2^{j-1}})(y)| &\le |x - y| \|(P_{2^j} - P_{2^{j-1}})'\|_\infty \\ &\le |x - y| \times 2\pi 2^j \times 2A\, 2^{-(j-1)a} \end{aligned}$$

whence

$$|f(x) - f(y)| \le CA \left[2^{-ka} + |x - y|\, 2^{(1-a)k} \right] \ , \quad C = \frac{2\pi}{2^{1-a} - 1} \ .$$

Choosing k so that $2^{-k} \le |x - y| \le 2^{-k+1}$, we obtain:

$$|f(x) - f(y)| \le C' |x - y|^a \ , \quad C' = C(2^{1-a} + 1) \ . \qquad \square$$

Remark 4.10. As we see, the relation between the continuity properties of f and the decay of E_N is rather tight. This can be compared with the much looser connection between the continuity properties of f and the decay of its Fourier coefficients. If f is continuous, the Fourier coefficients tend to zero by the Riemann–Lebesgue lemma (Lemma 3.11); no quantitative estimate can be made, in general (Exercise 3.16). On the other hand, the Fourier coefficients of a discontinuous function can decay as fast as $1/p$ (see (3.3)), though not much faster:

Exercise 4.11. Let $(\phi_p)_{p \in \mathbb{Z}}$ be a sequence of complex numbers.

1. If $\sum_p |\phi_p| < \infty$, then
$$f(x) = \sum_{p \in \mathbb{Z}} \phi_p e_p(x)$$
is a continuous function.

2. If $\sup_p |\phi_p| |p|^{1+a} < \infty$ for some $a \in (0, 1]$, then $f(x)$ is a-Hölder, i.e. $\omega_f(\delta) \le C\delta^a$.

Fourier series of analytic functions* Here is another constructive characterisation. In this case there is no major difference between the rates of decay of E_N and of the Fourier coefficients.

Proposition 4.12. *Let $f : \mathbb{T} \to \mathbb{C}$, and let $\epsilon > 0$. Then the following are equivalent:*

(1) f admits an analytic extension to $|\Im z| < \epsilon$;

(2) $\limsup_{p\to\pm\infty} \frac{1}{|p|} \log|\hat{f}(p)| \le -2\pi\epsilon$;

(3) $\limsup_{N\to\infty} \frac{1}{N} \log E_N(f) \le -2\pi\epsilon$.

Proof. (1) $\Longrightarrow$ (2): Choose $\epsilon' \in (0,\epsilon)$. Then

$$\hat{f}(p) = \int_0^1 f(x - i\epsilon') e_p(-(x - i\epsilon'))dx$$

(why?), and this expression is bounded in absolute value by $C(\epsilon')\exp(-2\pi p\epsilon')$. Similarly $|\hat{f}(p)| \le C(\epsilon')\exp(+2\pi p\epsilon')$.

(2) $\Longrightarrow$ (1): Let $\epsilon' \in (0,\epsilon)$. Choose $\epsilon'' \in (\epsilon',\epsilon)$, then the Fourier coefficients of f admit the bound

$$|\hat{f}(p)| \le C(\epsilon'')\exp(-2\pi|p|\epsilon'') ,$$

whence the series

$$\sum_{p=-\infty}^{\infty} \hat{f}(p)e_p(z)$$

of analytic functions converges uniformly in the strip $|\Im z| < \epsilon'$, and hence defines an analytic function there. This is true for any $\epsilon' \in (0,\epsilon)$, whence the claim.

(2) $\Longrightarrow$ (3): For $\epsilon' \in (0,\epsilon)$

$$|\hat{f}(p)| \le C(\epsilon')\exp(-2\pi|p|\epsilon') ,$$

hence

$$\|f - \sum_{p=-N}^{N} \hat{f}(p)e_p\|_\infty \le \sum_{|p|\ge N} C(\epsilon')\exp(-2\pi|p|\epsilon') \le C'(\epsilon')\exp(-2\pi N\epsilon') .$$

(3) $\Longrightarrow$ (2): Let $\epsilon' \in (0,\epsilon)$. If for any N one can find P_N for which

$$\|f - P_N\|_\infty \le C(\epsilon')\exp(-2\pi N\epsilon') ,$$

then definitely

$$|\hat{f}(p)| = |\hat{f}(p) - \hat{P}_{|p|-1}(p)| \le \|f - P_{|p|-1}\|_\infty \le 2\pi C(\epsilon')\exp(-2\pi N\epsilon') . \qquad \square$$

Exercise 4.13. Let $f : \mathbb{T} \to \mathbb{C}$ be a function satisfying the equivalent conditions of Proposition 4.12 with some $\epsilon > 0$. Then for any $\epsilon' \in (0, \epsilon)$

$$|\frac{1}{N}\sum_{k=0}^{N-1} f(k/N) - \int_0^1 f(x)dx| \le C(\epsilon')\exp(-2\pi N\epsilon') ,$$

i.e. the Riemann sums converge exponentially fast.

4.2 The heat equation

This was historically the first application of Fourier series (the name of the treatise of Fourier is "Théorie analytique de la chaleur"). The partial differential equation

$$\frac{\partial u(t,x)}{\partial t} = \frac{\kappa}{2}\frac{\partial^2 u(t,x)}{\partial x^2} , \tag{4.7}$$

where $\kappa > 0$ is a parameter, describes the propagation of heat in a one-dimensional medium. If $u(t,x)$ is the temperature at time t at position x of a long rod, the equation says that the rate of increase of the temperature is proportional to the difference between the average temperature in an infinitesimal neighbourhood of x and the temperature at x (in this case κ is called the thermal diffusivity). More formally, (4.7) is a limit of the difference equation:

$$u(t+\Delta t, x) - u(t,x) = \kappa\left[\frac{u(t, x+\Delta x) + u(t, x-\Delta x)}{2} - u(t,x)\right] \tag{4.8}$$

as $\Delta t, \Delta x \to +0$, $\Delta t = \Delta x^2$. The left-hand side of (4.8) is the increment in the temperature at x, while the right-hand side is proportional to the difference between the average temperature in the neighbourhood of x and the temperature at x. That is, the temperature at x evolves in the direction of the local average.

For (4.7), as well as for (4.8), one needs to impose an initial condition:

$$u(0,x) = u_0(x) \tag{4.9}$$

We shall assume that the rod is circular, i.e. we impose periodic boundary conditions $u(t, x+1) = u(t,x)$, and we set $\kappa = 1$ (since the case of general κ can be recovered by scaling). Expand $u(t,x)$ in a Fourier series:

$$u(t,x) \sim \sum_{p=-\infty}^{\infty} \hat{u}(t,p) e_p(x) .$$

The sign "~" is there to remind us that we do not discuss convergence yet.

Exercise 4.14. If $f : \mathbb{T} \to \mathbb{C}$ is differentiable and f' is piecewise continuous, then $\hat{f'}(p) = 2\pi i p \hat{f}(p)$.

Using this formula, we rewrite (4.7) as an equation on $\hat{u}$:

$$\frac{\partial \hat{u}(t,p)}{\partial t} = \frac{1}{2}(-4\pi^2 p^2)\hat{u}(t,p) \ . \tag{4.10}$$

Note that the p-th equation of (4.10) only involves $\hat{u}(\cdot,p)$, i.e. we have traded a partial differential equation for an infinite list of (uncoupled) ordinary differential equations. These are solved exactly:

$$\hat{u}(t,p) = \hat{u}(0,p)\exp(-2\pi^2 p^2 t) \ ,$$

and we finally have:

$$\begin{aligned} u(t,x) &\sim \sum_{p=-\infty}^{\infty} \hat{u}(0,p)\exp(-2\pi^2 p^2 t + 2\pi i p x) \\ &\sim \int_0^1 u(0,y) \sum_{p=-\infty}^{\infty} \exp(-2\pi^2 p^2 t + 2\pi i p (x-y)) dy \\ &= [P_t * u(0,\cdot)](x) \ , \end{aligned} \tag{4.11}$$

where

$$P_t(x) = \sum_{p=-\infty}^{\infty} \exp(-2\pi^2 p^2 t + 2\pi i p x) \tag{4.12}$$

is called the heat kernel.

The discussion so far has been purely formal. However, having the answer (4.11) at hand, we can justify it:

Proposition 4.15. *Let $u_0 \in C(\mathbb{T})$. Then*

$$u(t,x) = \begin{cases} u_0(x) \ , & t = 0 \\ (P_t * u_0)(x) \ , & t > 0 \end{cases}$$

defines the unique function in $C([0,\infty)\times\mathbb{T}) \cap C^\infty((0,\infty)\times\mathbb{T})$ which satisfies the heat equation (4.7) for $t > 0$ and $x \in \mathbb{T}$, with the initial condition (4.9).

Exercise 4.16. Prove the uniqueness part of Proposition 4.15.

It is also clear that $u(t,\cdot) = P_t * u_0$ satisfies the heat equation for $t > 0$. What remains is to prove that $u(t,\cdot) \rightrightarrows u_0$ as $t \to +0$. This is straightforward for, say, C^∞ initial conditions u_0 (why?). To treat general continuous initial conditions, we need to show that P_t is a non-negative summability kernel as in the proof of Theorem 3.20 (but with the $t \to +0$ limit in place of $N \to \infty$). This is easy to check using the following dual representation:

Lemma 4.17 (Jacobi identity)**.** *For any $t > 0$, $x \in \mathbb{T}$*

$$P_t(x) = \frac{1}{\sqrt{2\pi t}} \sum_{n=-\infty}^{\infty} \exp(-(x-n)^2/(2t)) \ . \tag{4.13}$$

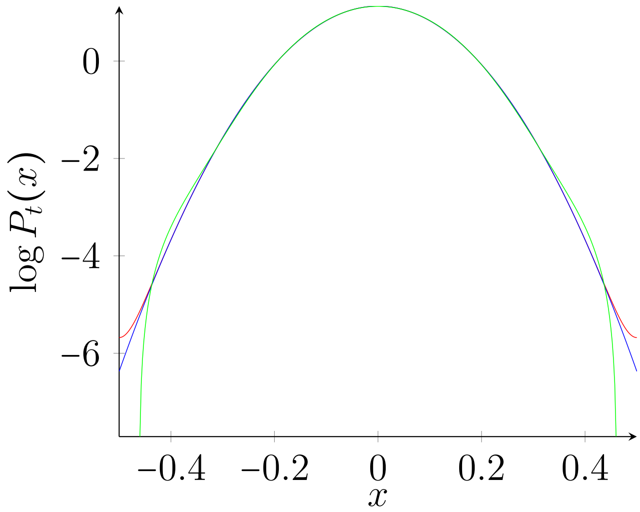

Figure 4.1: The logarithm of: the heat kernel for $t = 1/15$ (red); the zeroth term of (4.13) (blue); $\sum_{-3}^{3}$ of (4.12) (green).

Note that for small t the series (4.13) converges much faster than (4.12).

Proof. Introduce the C^∞ periodic function

$$f(x) = \sum_{n=-\infty}^{\infty} \exp(-(x-n)^2/(2t)) \ ;$$

its Fourier coefficients are given by

$$\begin{aligned}\hat{f}(p) &= \sum_{n=-\infty}^{\infty} \exp(-(x-n)^2/(2t) - 2\pi i p x) dx \\ &= \int_{-\infty}^{\infty} \exp(-x^2/(2t) - 2\pi i p x) dx = \sqrt{2\pi t} \exp(-2\pi^2 p^2 t) \ ,\end{aligned}$$

where the last step follows from the Gaussian integral

$$\int_{-\infty}^{\infty} \exp(-x^2/(2t) + ax) dx = \sqrt{2\pi t} \exp(a^2 t/2) \ , \quad a \in \mathbb{C} \ , \ t > 0 \ . \tag{4.14}$$

Exercise 4.18. Prove (4.14). It may be convenient to start with $a = 0$.

Thus

$$f(x) = \sqrt{2\pi t} \sum_{p=-\infty}^{\infty} \exp(-2\pi^2 p^2 t) e_p(x) \ ,$$

as claimed. □

Exercise 4.19. Verify that P_t is a non-negative summability kernel as in the proof of Theorem 3.20, and compelte the proof of Proposition 4.15.

It is instructive to inspect the solutiuon

$$u(t,x) = \sum_{p=-\infty}^{\infty} \hat{u}(0,p) \exp(-2\pi^2 p^2 t) e_p(x) \ .$$

The high Fourier coefficients are damped by the super-exponentially decaying term $\exp(-2\pi^2 p^2 t)$, i.e. by time t all the oscillations at frequencies $|p| \gg t^{-1/2}$ are smoothened out. In particular, $u(t,\bullet)$ is analytic for any $t > 0$ (why?). In the large t limit, u_t converges to its average $\bar{u} = \int_0^1 u_0(x)dx$. If u is smooth, this can be quantified as follows:

Exercise 4.20. Let $u_0 \in C(\mathbb{T})$ have a piecewise continuous derivative. Then

$$\|u_t - \bar{u}\|_2 \le \exp(-2\pi^2 t)\|u_0 - \bar{u}\|_2$$

(where $\bar{u}$ is understood as the constant function).

4.3 Toeplitz matrices, and the Szegő theorem

Convolution operators We start with a subject close to the initial point of these lecture notes (Example 1.1 and Exercise 2.3). If $r = (r_p)_{p\in\mathbb{Z}}$ is a two-sided sequence of complex numbers, we construct the infinite matrix

$$A[r] = (r_{p-q})_{p,q=-\infty}^{\infty} = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ \ddots & r_0 & r_{-1} & r_{-2} & & \\ & r_1 & r_0 & r_{-1} & r_{-2} & \\ & & r_1 & r_0 & r_{-1} & r_{-2} \\ & & & r_1 & r_0 & r_{-1} & \ddots \end{pmatrix}$$

representing convolution with the sequence $(r_p)_{p\in\mathbb{Z}}$. This matrix commutes with shifts, therefore Fourier transform should help us diagonalise it. And indeed, if we apply it to the vector $e_\bullet(x) = (e_p(x))_{p\in\mathbb{Z}}$, we get

$$(A[r]e_\bullet(x))_p = \sum_{q\in\mathbb{Z}} r_{p-q} e_q(x) = \sum_{q\in\mathbb{Z}} r_{p-q} e_{q-p}(x) e_p(x) = \left[\sum_{q\in\mathbb{Z}} r_q e_q(-x)\right] e_\bullet(x)_p \ .$$

That is, for each $x \in \mathbb{T}$, $e_\bullet(x)$ is a kind of eigenvector, and the expression $\check{r}(-x)$, where

$$\check{r}(x) = \sum_{q\in\mathbb{Z}} r_q e_q(x) \ ,$$

is the corresponding eigenvalue.

Naturally, this should not be taken too literally. To speak of eigenvectors and eigenvalues, one first needs to construct an operator acting on some space. We focus on the Hilbert space $\ell_2(\mathbb{Z})$ of two-sided sequences $\phi = (\phi_p)_{p\in\mathbb{Z}}$ with

$$\|\phi\|_2^2 = \sum_p |\phi_p|^2 < \infty \ .$$

Then $A[r]$ acts on this space in a natural way:

$$(A[r]\phi)_p = \sum_{q=\infty}^{\infty} r_{p-q}\phi_q = \sum_{q=-\infty}^{\infty} r_q \phi_{p-q} = (r * \phi)_p \ .$$

To ensure that $A[r]$ maps $\ell_2(\mathbb{Z})$ to itself, we assume that r is summable:

$$\|r\|_1 = \sum_{p=-\infty}^{\infty} |r_p| < \infty \ , \tag{4.15}$$

so that by Cauchy–Schwarz

$$|\langle r * \phi, \psi\rangle| = |\sum_p \sum_q r_q \phi_{p-q}\overline{\psi_p}| \le \sum_q |r_q| \sum_p |\phi_{p-q}\psi_p| \le \|r\|_1 \|\phi\|_2 \|\psi\|_2 \ , \tag{4.16}$$

whence $\|A[r]\| = \|A[r]\|_{2\to 2} \le \|r\|_1$. This is great, but unfortunately our wannabe-eigenvectors $e_\bullet(x)$ do not lie in the space.

Still, we would like to be able to compute the matrix elements of functions of $A[r]$, e.g. $(A[r]^5)_{00}$. We now try to make rigorous the idea that $A[r]$ is diagonalised by the Fourier transform.

First observe that $A[r]_{00} = r_0 = \int \check{r}(x)dx$. Next, $A[r]^2 = A[r * r]$, therefore $(A[r]^2)_{00} = \int \check{r}(x)^2 dx$. Similarly, $(A[r]^2)_{pq} = \int \check{r}(x)^2 e_{q-p}(x)dx$. Finally we obtain:

Proposition 4.21. *Let r be a summable sequence of complex numbers. Then, for any polynomial $u(\lambda)$ and any $p, q \in \mathbb{Z}$,*

$$u(A[r])_{pq} = \int_0^1 u(\check{r}(x)) e_{q-p}(x) dx = ((u \circ \check{r})^\wedge)_{p-q} \ . \tag{4.17}$$

This fact can be concisely written as

$$(\mathcal{F}^* u(A[r])\mathcal{F} f)(x) = u(\check{r}(-x)) f(x) \ . \tag{4.18}$$

Observe that the right-hand side of (4.17) (or of (4.18)) makes sense for any $u \in C(\mathbb{R})$. Therefore we can view (4.17) as a definition of the operator $u(A[r])$.

Exercise 4.22. The map $u \mapsto u(A[r]) = A[(u \circ \check{r})^\wedge]$ from $C(\mathbb{R})$ to the space of bounded operators on $\ell_2(\mathbb{Z})$ equipped with norm topology is continuous.

Now we discuss an application for which polynomials suffice.

Exercise 4.23. Let T be the transition matrix of the random walk on $\mathbb{Z}$, i.e. $T_{ij} = \frac{1}{2}\mathbb{1}_{|i-j|=1}$. Prove that:

1. $\sum_{k=0}^{\infty} [T^k]_{00} = \sum_{k=0}^{\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos^k(2\pi x) dx = \infty$, i.e. the number of returns of the random walk to its starting point has infinite expectation;

2. with probability one a random walk on $\mathbb{Z}$ returns to its starting point infinitely many times. (The same is true on $\mathbb{Z}^2$ but not on $\mathbb{Z}^3$ – why??)

Finally we comment on the condition $\|r\|_1 < \infty$, which is a special case of

Exercise 4.24 (Schur's test). A matrix $A = (a_{pq})_{p,q=-\infty}^{\infty}$ such that

$$\alpha = \sup_p \sum_q |a_{pq}| < \infty \ , \quad \beta = \sup_q \sum_p |a_{pq}| < \infty$$

defines a bounded operator on $\ell_2(\mathbb{Z})$, and moreover $\|A\| \leq \sqrt{\alpha\beta}$.

Exercise 4.25. Does there exist $r = (r_p)_{p\in\mathbb{Z}}$ such that $\|r\|_1 = \infty$ and yet $\|A[r]\| < \infty$?

Toeplitz matrices Our next goal is to study the finite restrictions of $A[r]$, namely, the $N \times N$ matrices

$$(A_N[r])_{p,q} = r_{p-q} \ , \quad 1 \leq p, q \leq N \ . \tag{4.19}$$

A matrix of this form is called a Toeplitz matrix. Our goal is to find the asymptotic behaviour of $\det A_N[r]$. The answer turns out to be

$$\lim_{N\to\infty} [\det A_N[r]]^{\frac{1}{N}} = \exp\left[\int_0^1 \log \check{r}(x)dx\right] \ , \tag{4.20}$$

where the right-hand side is the geometric mean of $\check{r}$ (at least, if $\check{r} \geq 0$). This answer is natural for the following reason:

$$\det A_N[r] = \exp \operatorname{tr} \log A_N[r] = \exp \sum_{p=1}^{N} (\log A_N[r])_{pp} \ ; \tag{4.21}$$

if we would be able to replace $A_N[r]$ with $A[r]$ in the last equality, we would get by (4.17)

$$(4.21) \overset{??}{\approx} \exp \sum_{p=1}^{N} (\log A[r])_{pp} = \exp\left[N \int_0^1 \log \check{r}(x)dx\right] \ , \tag{4.22}$$

as claimed. However, the justification of the approximation $(\log A_N[r])_{pp} \approx (\log A[r])_{pp}$ is a delicate matter, having to do with the influence of the boundary, and we shall proceed along different lines. We start with a digression.

Szegő theorem For an integrable $w : \mathbb{T} \to \mathbb{R}_+$, denote by

$$\mathfrak{G}(w) = \exp\left[\int_0^1 \log w(x)dx\right]$$

the geometric mean of w, understood to be zero if the integral diverges to $-\infty$.

Theorem 4.26 (Szegő)**.** *If $w : \mathbb{T} \to \mathbb{R}_+$ is piecewise continuous, then*

$$\lim_{N\to\infty} \inf_{c_0,\cdots,c_{N-1}\in\mathbb{C}} \int_0^1 |e_{-1}(x) - \sum_{p=0}^{N-1} c_p e_p(x)|^2 w(x) dx = \mathfrak{G}(w) \ .$$

Example 4.27. *If $w \equiv 1$, the left-hand side is obviously equal to one (the infimum is attained when all the coefficients are zero), and so is the right-hand side. On the other hand, if w vanishes on an open interval, the right-hand side equals zero (in fact, in this case the convergence to zero is exponentially fast). You are welcome to prove the following: for any open interval $I \subset \mathbb{T}$, any $f \in C(\mathbb{T})$ can be approximated by linear combinations of $(e_p)_{p\geq 0}$ uniformly on the complement of I.*

The conclusion of the Szegő theorem can be amplified, as follows.

Exercise 4.28. If $w : \mathbb{T} \to \mathbb{R}_+$ is piecewise continuous, then $\mathfrak{G}(w) = 0$ if and only if for *any* piecewise continuous $\phi : \mathbb{T} \to \mathbb{C}$

$$\lim_{N\to\infty} \inf_{c_0,\cdots,c_{N-1}\in\mathbb{C}} \int_0^1 |\phi(x) - \sum_{p=0}^{N-1} c_p e_p(x)|^2 w(x) dx = 0 \ .$$

(You are welcome to use the theorem, if needed)

We also note the following corollary/restatement (cf. (4.20)):

Exercise 4.29. In the setting of the Szegő theorem, we have for $A_N[\hat{w}]$ from (4.19):

$$\lim_{N\to\infty} [\det A_N[\hat{w}]]^{\frac{1}{N}} = \mathfrak{G}(w) \ .$$

Now we prove the theorem. We need a few basic properties of subharmonic functions (see any textbook in complex analysis, e.g. Ahlfors [1978]).

Definition 4.30. *Let $D \subset \mathbb{R}^2$ be a domain. An upper semicontinuous function $u : D \to \mathbb{R} \cup \{-\infty\}$ is called subharmonic if for any $(x_0, y_0) \in D$ and any $0 < r <$* $\mathrm{dist}((x_0, y_0), \partial D)$,

$$u(x_0, y_0) \leq \int_0^1 u(x_0 + r\cos(2\pi x), y_0 + r\sin(2\pi x)) dx \ . \tag{4.23}$$

Remark 4.31.

1. It is sufficient to impose the condition (4.23) for $0 < r < r_0(x_0, y_0)$. Another equivalent property is that $\Delta u \geq 0$ in distribution sense, i.e.

$$\int_D u(x, y) \Delta\phi(x, y) dx dy \geq 0 \tag{4.24}$$

   for any smooth non-negative test function ϕ which is compactly supported in D.

2. If $F(z)$ is analytic in D, then $u(x,y) = \log|F(x+iy)|$ is subharmonic. Further, if F does not have zeros in D, then u is harmonic, meaning that (4.23) (or (4.24)) is always an equality. Both of these facts follow from the identity

$$\int_D \log|F(x+iy)|\,(\Delta\phi)(x,y)dxdy = 2\pi \sum_{F(x+iy)=0} \phi(x,y) \ ,$$

valid for any ϕ which is smooth and compactly supported in D (with multiplicity taken into account on the right-hand side).

Proof of Theorem 4.26. We shall prove the following identity (equivalent to the claimed one):

$$\lim_{N\to\infty} \inf_{c_1,\cdots,c_N\in\mathbb{C}} \int_0^1 |1 - \sum_{p=1}^N c_p e_p(x)|^2 w(x)dx = \mathfrak{G}(w) \ .$$

We start with the inequality $\boxed{\geq}$. By Jensen's inequality,

$$\begin{aligned}\log \int_0^1 |1 - \sum_{p=1}^N c_p e_p(x)|^2 w(x)dx &\geq \int_0^1 \log\left[|1 - \sum_{p=1}^N c_p e_p(x)|^2 w(x)\right]dx \\ &= \int_0^1 \log|1 - \sum_{p=1}^N c_p e_p(x)|^2 dx + \int_0^1 \log w(x)dx \ .\end{aligned} \tag{4.25}$$

From (4.23) applied to $u(x,y) = \log|1 - \sum_{p=1}^N c_p (x+iy)^p|$, the first integral is

$$= 2\int_0^1 u(\cos(2\pi x), \sin(2\pi x))dx \geq u(0,0) = 0 \ , \tag{4.26}$$

whence

$$\int_0^1 |1 - \sum_{p=1}^N c_p e_p(x)|^2 w(x)dx \geq \mathfrak{G}(w) \ .$$

Now we prove $\boxed{\leq}$. To saturate both inequalities (4.25) and (4.26), we need to find $P(z) = 1 - \sum_{p=1}^N c_p z^p$ such that $|P(e^{2\pi i x})|^2 w(x)$ is approximately constant and also P has no zeros in the interior of the unit disc. We can assume without loss of generality that w is bounded away from zero. Indeed, let $w_\delta = \max(w,\delta)$; then

$$\lim_{\delta\to+0} \int \log w_\delta(x)dx \to \int \log w(x)dx$$

(why?), whereas

$$\int |P(x)|^2 w_\delta(x)dx \geq \int |P(x)|^2 w(x)dx \ .$$

Therefore we assume that there exists $M \geq \delta > 0$ such that $0 < \delta \leq w \leq M$. Let Q_η be a trigonometric polynomial such that $|Q_\eta - 1/w| \leq \eta$ except for a union of intervals of total length $\leq \eta$, and such that $M^{-1} - \eta \leq Q_\eta \leq \delta^{-1} + \eta$. Then

$$|\int Q_\eta(x)w(x)dx - 1| \leq M(1 + \delta^{-1} + \eta)\eta \to 0 \quad \text{as} \quad \eta \to +0 \ .$$

Thus we have as $\eta \to +0$

$$\log \int Q_\eta(x)w(x)dx \leq \int Q_\eta(x)w(x)dx - 1 \to 0 \ ,$$
$$\int \log(Q_\eta(x)w(x))dx \geq \log(1 - M\eta) - \eta \log(M(\delta^{-1} + \eta)) \to 0 \ ,$$

hence $Q_\eta w$ almost saturates Jensen's inequality:

$$\log \int Q_\eta(x)w(x)dx - \int \log(Q_\eta(x)w(x))dx \to 0 \quad \text{as} \quad \eta \to +0 \ . \tag{4.27}$$

Now we need

Lemma 4.32 (Riesz–Féjér)**.** *Let $Q(x)$ be a trigonometric polynomial which is non-negative on $\mathbb{T}$. Then there exists a polynomial $P(z)$ with no zeros in $\mathbb{D}$ such that $Q(x) = |P(e^{2\pi i x})|^2$ on $\mathbb{T}$.*

Applying the lemma to Q_η, we get a polynomial P_η all the zeros of which lie outside $\overline{\mathbb{D}}$. Set $\tilde{P}_\eta(z) = P_\eta(z)/P_\eta(0)$. Then (4.27) implies that

$$\begin{aligned}\log \int |\tilde{P}_\eta(e^{2\pi i x})|^2 w(x)dx - \mathfrak{G}(w) \\ = \log \int |\tilde{P}_\eta(e^{2\pi i x})|^2 w(x)dx - \int \log(|\tilde{P}_\eta(e^{2\pi i x})|^2 w(x))dx \to 0 \quad \text{as} \quad \eta \to +0 \ ,\end{aligned}$$

as claimed. Now we proceed to

Proof of Lemma 4.32.

Let $Q(x) = \sum_{p=-N}^{N} c_p e_p(x)$, so that $c_{-N} \neq 0$. Then $R(z) = \sum_{p=-N}^{N} c_p z^{N+p}$ is an algebraic polynomial that satisfies $R(0) \neq 0$, and $Q(x) = |R(e^{2\pi i x})|$. Due to the relation $R(1/z) = z^{-2N}\overline{R(\bar{z})}$, the zeros of R on the unit circle are of even multiplicity, whereas the roots not on the unit circle come in pairs $(z, 1/\bar{z})$. Therefore

$$R(z) = c \prod_j (z - z_j)(z - 1/\bar{z}_j) \ , \quad \text{where} \quad c > 0 \ , |z_j| \geq 1 \ ,$$

and we can set

$$P(z) = \sqrt{c} \prod_j (z - z_j) \ . \qquad \square$$

Predictability of Gaussian processes Here is an application. Let $(X_j)_{j\in\mathbb{Z}}$ be a stationary Gaussian process with $\mathbb{E}X_j = 0$. It is called predictable if

$$\lim_{N\to\infty} \inf_{c_0,\cdots,c_{N-1}} |X_1 - \sum_{p=0}^{N-1} c_p X_{-p}|^2 = 0 \ .$$

The Szegő theorem allows to derive a necessary and sufficient condition for predictability, we shall discuss a special case.

The covariances

$$r_{p-q} = \mathbb{E}X_p \bar{X}_q$$

always form a positive semidefinite matrix $(r_{p-q})_{p,q\in\mathbb{Z}} \geq 0$, i.e.

$$\sum_{p,q} c_p r_{p-q} \bar{c}_q = \mathbb{E}|\sum c_p X_p|^2 \geq 0 \tag{4.28}$$

for any complex coefficients c_p. Vice versa, any sequence $(r_p)_{p\in\mathbb{Z}}$ for which $(r_{p-q})_{p,q} \geq 0$ corresponds to a stationary Gaussian process. Here is a way to construct such sequences:

$$r_p = \hat{w}(p) = \int_0^1 e^{-2\pi i x} w(x) dx \ , \quad w \geq 0 \ . \tag{4.29}$$

Exercise 4.33. Any sequence r_p of the form (4.29) satisfies $(r_{p-q})_{p,q} \geq 0$.

The function $w(x)$ is called the spectral function of the process $(X_p)_{p\in\mathbb{Z}}$.

Exercise 4.34. Let $w : \mathbb{T} \to \mathbb{R}_+$ be piecewise continuous. The stationary Gaussian process (X_j) defined by

$$\mathbb{E}X_p \bar{X}_q = \hat{w}(p-q)$$

is predictable if and only if $\mathfrak{G}(w) = 0$.

Remark 4.35. The construction of stationary Gaussian processes described above is almost the most general one: if we replace the piecewise continuous function $w(x)$ with a non-negative measure on the circle, we get the general form of a sequence (r_p) satisfying $(r_{p-q}) \geq 0$ (see Exercise 4.57 below), leading to the general form of a stationary Gaussian process.

See Grenander and Szegő [1984] or Dym and McKean [1972] for the general version of the Szegő theorem and the general version of Exercise 4.34.

4.4 Central limit theorem

Let $S_N = S_1 + \cdots + S_N$ be independent, identically distributed random variables with $\mathbb{E}X_1 = a$ and $\mathbb{E}(X_1 - a)^2 = \sigma^2$. The Levy–Lindeberg Central Limit Theorem asserts that

$$\frac{S_N - aN}{\sigma\sqrt{N}} \to N(0,1)$$

in distribution, i.e.

$$\forall \alpha \le \beta \qquad \lim_{N\to\infty} \mathbb{P}\left\{\frac{S_N - aN}{\sigma\sqrt{N}} \in [\alpha,\beta]\right\} = \int_\alpha^\beta \frac{1}{\sqrt{2\pi}} e^{-s^2/2} ds \ . \tag{4.30}$$

The simplest case (studied by de Moivre and Laplace) is when X_j are Bernoulli(1/2) distributed,

$$X_j = \begin{cases} -1 & \text{with probability } 1/2 \\ +1 & \text{with probability } 1/2. \end{cases}$$

In this case

$$\mathbb{P}(S_N = k) = \begin{cases} 2^{-N}\binom{N}{(N-k)/2} \ , & -N \le k \le N, \ N-k \in 2\mathbb{Z} \\ 0 \end{cases}$$

and Stirling's approximation $n! = (1+o(1))\sqrt{2\pi e}(n/e)^n$ implies that for any $\delta < 1/4$ one has

$$\mathbb{P}(S_N = k) = (1+o(1))\,\frac{2}{\sqrt{2\pi N}} e^{-k^2/(2N)} \ , \quad |k| \le N^{1/2+\delta} \ , \ N-k \in 2\mathbb{Z} \tag{4.31}$$

with a uniform $o(1)$ term.

Exercise 4.36. Check that (4.31) implies (4.30).

The combinatorial approach does not easily extend to other random variables. However, the following result is quite general.

Theorem 4.37 (Gnedenko)**.** *Let X_j be independent copies of an integer-valued random variable X such that $\mathbb{E}X = a$ and $\mathbb{E}(X-a)^2 = \sigma^2$. If* $\operatorname{supp} X$ *does not lie in any sublattice of $\mathbb{Z}$, in other words, if the differences $\{k - k' \mid k, k' \in \operatorname{supp} X\}$ generate $\mathbb{Z}$ as a group, then*

$$\lim_{N\to\infty} \sup_{k\in\mathbb{Z}} \sqrt{N}\left|\mathbb{P}(S_N = k) - \frac{1}{\sqrt{2\pi\sigma^2 N}} \exp\left[-\frac{(k-a)^2}{2\sigma^2 N}\right]\right| = 0 \ . \tag{4.32}$$

Exercise 4.38. Explain the inconsistency between (4.31) and (4.32).

Exercise 4.39. Check that (4.32) implies (4.30).

The proof uses characteristic functions. Let $\pi(k) = \mathbb{P}(X = k)$ and $\pi_N(k) = \mathbb{P}(S_N = k)$. Then

$$\pi_N(k) = \sum_{k_1+\dots+k_N = k} \pi(k_1)\cdots\pi(k_N) = \underbrace{(\pi * \cdots * \pi)}_{N \text{ times}}(k) \ ,$$

i.e. π_N is the N-fold convolution power of π. Set

$$\check{\pi}(x) = \sum_{k\in\mathbb{Z}} \pi(k) e_k(x) \ , \quad \check{\pi}_N(x) = \sum_{k\in\mathbb{Z}} \pi_N(k) e_k(x) \ ,$$

so that $(\check{\pi})^\wedge = \pi$ and $(\check{\pi}_N)^\wedge = \pi_N$. Then $\check{\pi}_N = (\check{\pi})^N$.

Exercise 4.40. Let π be a probability distribution on the integers.

1. $\check{\pi} \in C(\mathbb{T})$;
2. if, for some $m \in \mathbb{N}$, the m-th absolute moment $\sum_k \pi(k)|k|^m < \infty$, then $\check{\pi} \in C^m(\mathbb{T})$.

In our case, $\check{\pi} \in C^2(\mathbb{T})$, and

$$\check{\pi}(0) = 1 \ , \quad \check{\pi}'(0) = 2\pi i a \ , \quad \check{\pi}''(0) = -4\pi^2(a^2 + \sigma^2) \ , \tag{4.33}$$

i.e.

$$\check{\pi}(x) \approx 1 + 2\pi i a x - 2\pi^2(a^2 + \sigma^2)x^2 \approx \exp(2\pi i a x - 2\pi^2\sigma^2 x^2) \ , \tag{4.34}$$

and one can hope that

$$\pi_N(k) \approx \int_{-1/2}^{1/2} \exp(2\pi i(aN - k)x - 2\pi^2 N\sigma^2 x^2) dx \ . \tag{4.35}$$

For large N, the integrand is very small outside $(-1/2, 1/2)$, hence we can further hope that

$$(4.35) \approx \int_{-\infty}^{\infty} \exp(2\pi i(aN-k)x - 2\pi^2 N\sigma^2 x^2)dx = \frac{1}{\sqrt{2\pi}} \exp(-(k-aN)^2/(2N\sigma^2)) \ .$$

Here the last equality follows from (4.14). To justify this argument, we first need to handle x which are not very close to 0. For such x (4.34) and (4.35) are not valid, however, we can hope that $\check{\pi}_N$ is small.

Example 4.41. *For Bernoulli(1/2) random variables,* $\check{\pi}(x) = \cos(2\pi x)$*, hence* $|\check{\pi}_N(1/2)| = 1$ *whereas the right-hand side of (4.35) is exponentially small.*

Luckily, Bernoulli(1/2) random variables do not satisfy the assumptions of the theorem (see Exercise 4.38 above!) More generally, we have:

Lemma 4.42. *Let* π *be a probability distribution on the integers such that the differences*

$$\left\{ k - k' \,\middle|\, k, k' \in \operatorname{supp} \pi \right\}$$

generate $\mathbb{Z}$ *(as a group). Then* $\max_{x \in \mathbb{T}} |\check{\pi}(x)|$ *is uniquely attained at* $x = 0$.

Proof. Suppose $|\check{\pi}(x)| = 1$ for some $x \neq 0$. Then

$$\left| \sum_{k \in \mathbb{Z}} \pi(k) e_k(x) \right| = \sum_{k \in \mathbb{Z}} \pi(k) |e_k(x)| \ ,$$

i.e. $e_k(x)$ is constant on $\operatorname{supp} \pi$, whence $e_{k-k'}(x) = 1$ for any $k, k' \in \operatorname{supp} \pi$. This implies that x is rational, $x = a/b$, and all the differences $k - k'$ lie in $b\mathbb{Z}$. $\square$

Proof of Theorem 4.37. According to (4.33) and the Taylor expansion with Peano remainder,

$$|\check{\pi}(x) - \exp(2\pi i a x - 2\pi^2\sigma^2 x^2)| \leq x^2\omega(x) \ , \quad |x| \leq \frac{1}{2} \ ,$$

where $\omega : [0, 1/2] \to \mathbb{R}_+$ is non-decreasing and tends to 0 as the argument approaches zero. This implies

$$|\log \check{\pi}(x) - (2\pi i a x - 2\pi^2\sigma^2 x^2)| \leq 2x^2\omega(x) \ , \quad |x| \leq c_1 \ ,$$

where c_1 is small enough to ensure that $\Re\check{\pi} > 0$ in this region, whence:

$$|\log \check{\pi}_N(x) - N(2\pi i a x - 2\pi^2\sigma^2 x^2)| \leq 2Nx^2\omega(x) \ , \quad |x| \leq c_1 \ .$$

Now we write

$$\begin{aligned}&\left|\pi_N(k) - \frac{1}{\sqrt{2\pi\sigma^2 N}} e^{-(k-a)^2/(2\sigma^2 N)}\right| \\ &\quad = \left|\int_{-1/2}^{1/2} \check{\pi}_N(x) e_k(x) dx - \int_{-\infty}^{\infty} e^{-N(2\pi i a x - 2\pi^2\sigma^2 x^2)} e_k(x) dx\right| \leq A' + A'' + A''' \ ,\end{aligned}$$

where

$$\begin{aligned} A' &= \int_{-R/\sqrt{N}}^{R/\sqrt{N}} \left|\check{\pi}_N(x) - e^{-N(2\pi i a x - 2\pi^2\sigma^2 x^2)}\right| dx \ , \\ A'' &= \int_{\frac{R}{\sqrt{N}} \leq |x| \leq \frac{1}{2}} |\check{\pi}_N(x)| dx \ , \\ A''' &= \int_{\frac{R}{\sqrt{N}} \leq |x|} |e^{-N(2\pi i a x - 2\pi^2\sigma^2 x^2)}| dx = \int_{\frac{R}{\sqrt{N}} \leq |x|} e^{-2N\pi^2\sigma^2 x^2} dx \end{aligned}$$

Exercise 4.43. Show that $A' \leq CR^2\omega(R/\sqrt{N})/\sqrt{N}$, $A'' + A''' \leq \frac{C}{R\sqrt{N}} \exp(-2\pi^2\sigma^2 R^2)$ and complete the proof of the theorem. □

4.5 Equidistribution modulo one

The following is based on the work of Weyl. We mostly follow Montgomery [2014]. A sequence $(a_n)_{n\geq 1}$ of real numbers is said to be *equidistributed modulo one* if for any arc $I \subset \mathbb{T}$ one has:

$$\lim_{N\to\infty} \frac{1}{N} \# \{1 \leq n \leq N : \ a_n + \mathbb{Z} \in I\} = |I| \ , \tag{4.36}$$

i.e. the number of points in each arc is asymptotically proportional to the length of the arc.

Exercise 4.44. Let $X_1, X_2, \cdots$ be a sequence of independent random variables uniformly distributed in $[0,1]$. Then almost surely (X_n) is equidistributed modulo one.

Exercise 4.45. A sequence (a_n) is equidistributed modulo one if and only if for any $f \in C(\mathbb{T})$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(a_n) = \int_0^1 f(x)dx \ .$$

Corollary 4.46 (Weyl's criterion). *A sequence (a_n) is equidistributed modulo one if and only if for any $p \geq 1$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_p(a_n) = 0 \ . \tag{4.37}$$

The convenient feature of Weyl's criterion (4.37) is that it is a countable list of conditions which are often not hard to check. Informally, the p-th condition is responsible for equidistribution on scale $\approx 1/p$.

Exercise 4.47. Solve Exercise 4.44 using Weyl's criterion.

Here is a more interesting example.

Claim 4.48. If $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$, the sequence $(a_n = \alpha n)_{n\geq 1}$ is equidistributed modulo one.

The conclusion clearly fails if α is rational (why?)

Proof. For each $p \neq 0$,

$$\sum_{n=1}^{N} e_p(a_n) = \frac{e_p(\alpha)(1 - e_p(N\alpha))}{1 - e_p(\alpha)} \tag{4.38}$$

is bounded (as N varies), since the absolute value of the numerator is at most two, and the denominator does not vanish. Therefore the left-hand side of (4.37) is $O(1/N)$. □

Remark 4.49. The better α is approximated by rationals, the slower is the convergence in (4.36). A quantitative version of Weyl's criterion was found by Erdős and Turán; see Montgomery [2014].

Exercise 4.50. Let $f \in C(\mathbb{T})$. Then for any $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$

$$\frac{1}{N} \sum_{n=1}^{N} f(x + n\alpha) \rightrightarrows \int_0^1 f(x)dx \ , \quad N \to \infty \ .$$

Theorem 4.51 (Weyl). *Let $P \in \mathbb{R}[x]$ be a polynomial with at least one irrational coefficient (next to a power ≥ 1). Then $(P(n))_{n\geq 1}$ is equidistributed modulo one.*

Exercise 4.52. It is sufficient to prove the theorem for polynomials with irrational leading coefficient.

Proof. We argue by induction on $\deg P$. Claim 4.48 takes care of the induction base. The induction step follows from the next theorem. □

Theorem 4.53 (van der Corput). *Let $(a_n)_{n\geq 1}$ be a sequence of real numbers. If, for any $h \geq 1$, $(a_{n+h} - a_n)_{n\geq 1}$ is equidistributed modulo one, then $(a_n)_{n\geq 1}$ is equidistributed modulo one.*

Exercise 4.54. Is the converse true?

The claim is plausible since

$$\left|\frac{1}{N}\sum_{n=1}^{N} e_p(a_n)\right|^2 = \frac{1}{N}\sum_{h=1}^{N}\left(\frac{1}{N}\sum_{n=1}^{N} e_p(a_n^{(N)} - a_{n+h}^{(N)})\right) , \tag{4.39}$$

where $a_n^N = a_n$ for $1 \leq n \leq N$ and then continues periodically. Now,

$$\left|\sum_{n=1}^{N} e_p(a_n^{(N)} - a_{n+h}^{(N)}) - \sum_{n=1}^{N} e_p(a_n - a_{n+h})\right| \leq 2h ,$$

therefore for each fixed h the expression in the large parentheses of (4.39) tends to zero. The problem is that we need to sum over all $1 \leq h \leq N$, whereas the assumptions of the theorem imply no uniformity in h. The argument of van der Corput presented below allows to restrict the summation to a finite range of h.

Lemma 4.55 (van der Corput). *For any $z_1, \cdots, z_N \in \mathbb{C}$ and any $H \geq 1$,*

$$H^2\left|\sum_{n=1}^{N} z_n\right|^2 \leq H(N+H-1)\sum_{n=1}^{N}|z_n|^2 + 2(N+H-1)\sum_{h=1}^{H-1}(H-h)\left|\sum_{n=1}^{N-h} z_{n+h}\bar{z}_n\right| .$$

Proof. Set $z_n = 0$ for $n \geq N+1$ and for $n \leq 0$. Then

$$H\sum_{n=1}^{N} z_n = \sum_{r=0}^{H-1}\sum_{n=1}^{N+H-1} z_{n-r} = \sum_{n=1}^{N+H-1}\sum_{r=0}^{H-1} z_{n-r} ,$$

whence

$$H^2\left|\sum_{n=1}^{N} z_n\right|^2 \leq (N+H-1)\sum_{n=1}^{N+H-1}\left|\sum_{r=0}^{H-1} z_{n-r}\right|^2 \qquad \text{(Jensen)} \tag{4.40}$$

$$= (N+H-1)\sum_{n=1}^{N+H-1}\sum_{r,s=0}^{H-1} z_{n-r}\bar{z}_{n-s} \tag{4.41}$$

$$\leq (N+H-1)H\sum_{n=1}^{N}|z_n|^2 \tag{4.42}$$

$$+ 2(N+H-1)\sum_{h=1}^{H-1}(H-h)\left|\sum_{n=1}^{N-h} z_{n+h}\bar{z}_n\right| .$$

□

Proof of Theorem 4.53. Let

$$A_{N,p} = \frac{1}{N} \sum_{n=1}^{N} e_p(a_n) \ , \quad A_{N,p,h} = \frac{1}{N} \sum_{n=1}^{N} e_p(a_{n+h} - a_n) \ .$$

Fix $p \geq 1$; we know that $A_{N,p,h} \to 0$ for any $h \geq 1$, and we need to show that $A_{N,p} \to 0$. Applying van der Corput's lemma to $z_n = e_p(a_n)$, we obtain:

$$H^2 N^2 |A_{N,p}|^2 \leq H(N+H-1)N + 2(N+H-1) \sum_{h=1}^{H-1} (H-h)(N-h)|A_{N-h,p,h}| \ ,$$

whence for large N

$$|A_{N,p}|^2 \leq \frac{2}{H} + \frac{3}{H} \sum_{h=1}^{H-1} |A_{N-h,p,h}| \ .$$

It remains to let $N \to \infty$ and then $H \to \infty$. □

Exercise 4.56. The sequence $(\sqrt{n})_{n\geq 1}$ is equidistributed modulo 1.

Fourier series of functionals* Weyl's criterion is not a statement about Fourier series in the generality of our discussion so far. One way to embed it into the harmonic analysis framework is as follows. Let $\mu : C(\mathbb{T}) \to \mathbb{C}$ be a continuous functional. One may think e.g. of the functionals

$$\nu_g : f \mapsto \int_0^1 fg dx, \quad \delta_{x_0} : f \mapsto f(x_0) \ , \quad \cdots \tag{4.43}$$

Define the Fourier coefficients $\hat{\mu}(p) = \mu(e_{-p})$. Then

$$\mu_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{a_n} \to \nu_1 \tag{4.44}$$

if and only if $\widehat{\mu_N}(p) \to \widehat{\nu_1}(p) = 0$ for any $p \in \mathbb{N}$ (what about $p = 0$ and negative p?) The convergence (4.43) of functionals on $C(\mathbb{T})$ is equivalent (in this case) to the equidistribution of (a_n), cf. Exercise 4.45.

The functionals which we considered in this section had a special positivity property, which we highlight below.

Exercise 4.57. Let $\mu : C(\mathbb{T}) \to \mathbb{C}$ be a continuous functional. The following are equivalent:

1. μ sends (pointwise) non-negative functions to non-negative numbers;

2. μ sends (pointwise) non-negative trigonometric polynomials to non-negative numbers;

3. the matrix $(\hat{\mu}_{p-q})_{p,q\in\mathbb{Z}}$ is non-negative semi-definite (i.e. any principal sub-matrix of finite size is non-negative semi-definite).

Exercise 4.58. Let $(r_p)_{p\in\mathbb{Z}}$ be a sequence such that $(r_{p-q})_{p,q\in\mathbb{Z}}$ is non-negative semi-definite. Then there exists μ satisfying the equivalent conditions of Exercise 4.57 such that $r = \hat{\mu}$ (cf. Remark 4.35).

One can go further and consider functionals on smaller functional spaces containing e_p, e.g. the space $C^\infty(\mathbb{T})$.

Exercise 4.59. Compute the Fourier coefficients of the functional $f \mapsto f''(x_0)$.

5 Fourier transform

5.1 Introduction

Now we consider the group $\mathbb{R}$. Consider the characters

$$e_p(x) = \exp(2\pi i p x) \ , \quad p \in \mathbb{R} \ .$$

Note that we insist that $p \in \mathbb{R}$ (these are singled out, for example, by the requirement that a character be a closed map, or by insisting that the image lie in the unit circle). For a nice (piecewise continuous and absolutely integrable) function $f(x)$, set

$$(\mathcal{F}f)(p) = \hat{f}(p) = \int_{-\infty}^{\infty} f(x)\overline{e_p(x)}dx \ .$$

Similarly, for a nice $\phi(p)$ set

$$(\mathcal{F}^*\phi)(x) = \check{\phi}(x) = \int_{-\infty}^{\infty} \phi(p)e_p(x)dp \ .$$

Similarly to Section 3.2,

$$\int_{-R}^{R} \hat{f}(p)e_p(x)dp = \int_{-\infty}^{\infty} D_R(x-y)f(y)dy = (D_R * f)(x) \ ,$$

where the Dirichlet kernel is now given by

$$D_R(y) = \int_{-R}^{R} e_p(y)dp = \frac{\sin(2\pi R y)}{\pi y} \ . \tag{5.1}$$

Exercise 5.1. Prove the Riemann–Lebesgue lemma: if f is piecewise continuous and absolutely integrable, then $\hat{f}(p) \to 0$ as $p \to \pm\infty$.

Exercise 5.2. Let $a \in (0, 1]$. If f is piecewise continuous and absolutely integrable, and if

$$|f(x) - f(x_0)| \le C|x - x_0|^a$$

for all x in a neighbourhood of a certain $x_0 \in \mathbb{R}$, then $(D_R * f)(x_0) \to f(x_0)$, and, moreover,

$$\lim_{R_+, R_- \to +\infty} \int_{-R_-}^{R_+} \hat{f}(p) e_p(x_0) dp = f(x_0) \ .$$

Proposition 5.3. *Let $f : \mathbb{R} \to \mathbb{C}$ be bounded, absolutely integrable and uniformly continuous. Then*

$$\frac{1}{R} \int_0^R dr \int_{-r}^r \hat{f}(p) e_p(x) dp \rightrightarrows f(x) \ .$$

Proof. The proof follows the lines of Section 3.3. The starting point is the identity

$$\frac{1}{R} \int_0^R dr \int_{-r}^r \hat{f}(p) e_p(x) dp = (S_R * f)(x) \ , \quad S_R(x) = \frac{1}{R} \frac{\sin^2(\pi R x)}{(\pi x)^2} \ .$$

□

5.2 Fourier transform in Schwartz space

The Schwartz space $S(\mathbb{R})$ is defined as

$$S(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}) \,\middle|\, \forall k, \ell \ge 0 \ |||f|||_{k,l} = \sup_{x \in \mathbb{R}} |x|^k |f^{(\ell)}(x)| < \infty \right\} \ .$$

Observe that every $f \in S(\mathbb{R})$ is absolutely integrable, hence $\hat{f}$ is well-defined. We also note that $S(\mathbb{R})$ is closed under differentiation and multiplication by x (why?)

Proposition 5.4. *The map $\mathcal{F} : f \mapsto \hat{f}$ is a bijection $S(\mathbb{R}) \to S(\mathbb{R})$, and $\mathcal{F}^*$ is its inverse.*

Proof. In view of Exercise 5.2, it suffices to check that $\mathcal{F}(S(\mathbb{R})) \subset S(\mathbb{R})$. To this end, we compute

$$\begin{aligned}
|||\hat{f}|||_{k,\ell} &= \sup_{p \in \mathbb{R}} |p|^k |\hat{f}^{(\ell)}(p)| \\
&= \sup_{p \in \mathbb{R}} |p|^k |((-2\pi i p)^\ell f)\hat{}\,(p)| \\
&= (2\pi)^{\ell - k} \sup_{p \in \mathbb{R}} |(\frac{d^k}{dx^k} x^\ell f)\hat{}\,(p)| \\
&\le (2\pi)^{\ell - k} ||| \frac{d^k}{dx^k}(x^\ell f) |||_{0,0} \ .
\end{aligned}$$

□

Exercise 5.5. (cf. Exercise 4.18)

$$f(x) = \exp(-\frac{x^2}{2t} + ax) \Longrightarrow \hat{f}(p) = \sqrt{2\pi t}\exp((2\pi i p - a)^2 t/2) \ .$$

Plancherel indentity

Proposition 5.6 (Plancherel)**.** *Let $f \in S(\mathbb{R})$. Then*

$$\int |\hat{f}(p)|^2 dp = \int |f(x)|^2 dx \ .$$

We shall use the following converse to Exercise 3.21.

Exercise 5.7. Let $u : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-decreasing function. If

$$\lim_{R\to\infty} \frac{1}{R}\int_0^R u(r)dr = L \in [0, +\infty] \ ,$$

then also

$$\lim_{r\to\infty} u(r) = L \ .$$

Proof of Proposition 5.6. The integral $\int_{-r}^r |\hat{f}(p)|^2 dp$ is non-decreasing in r, therefore it suffices to show that

$$\lim_{R\to\infty} \frac{1}{R}\int_0^R \int_{-r}^r |\hat{f}(p)|^2 dp = \int |f(x)|^2 dx \ . \tag{5.2}$$

Interchanging the limits (by what rights?) in the identity

$$\int_{-r}^r |\hat{f}(p)|^2 dp = \int_{-r}^r \left[\int f(x)\overline{e_p(x)}dx \int \overline{f(y)}e_p(y)dy\right] dr \ ,$$

we obtain:

$$\int_{-r}^r |\hat{f}(p)|^2 dp = \int dy\overline{f(y)} \int dx f(x) D_r(y-x) \ ,$$

whence

$$\text{LHS of (5.2)} = \int dy\overline{f(y)} \int dx f(x) S_r(y-x) \ .$$

By Proposition 5.3,

$$\int dx f(x) S_r(y-x) \rightrightarrows f(y) \ , \quad r \to +\infty \ ,$$

hence

$$\text{LHS of (5.2)} \to \int |f(x)|^2 dx \ .$$

□

Exercise 5.8. Show that for any $f, g \in S(\mathbb{R})$,

$$\int \hat{f}(p)\overline{\hat{g}(p)}dp = \int f(x)\overline{g(x)}dx \ .$$

Remark 5.9. The Plancherel identity allows to extend the Fourier transform to an isometry of the space of (Lebesgue-) square-summable functions.

5.3 Poisson summation

Proposition 5.10. *Let $f \in S(\mathbb{R})$. Then*

$$\sum_{p=-\infty}^{\infty} \hat{f}(p) = \sum_{n=-\infty}^{\infty} f(n) \ .$$

Proof. Let $g(x) = \sum_{n\in\mathbb{Z}} f(x+n)$. Then $g \in C^\infty(\mathbb{T})$, and

$$\hat{g}(p) = \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(x)\overline{e_p(x)}dx = \int_{-\infty}^{\infty} f(x)\overline{e_p(x)}dx = \hat{f}(p), \quad p \in \mathbb{Z} \ ,$$

where the hat on the left-hand side stands for the Fourier coefficients, and on the right-hand side — for the Fourier transform. Thus

$$\sum_{n\in\mathbb{Z}} f(n) = g(0) = \sum_{p\in\mathbb{Z}} \hat{g}(p) = \sum_{p\in\mathbb{Z}} \hat{f}(p) \ . \qquad \square$$

One application is the Jacobi identity, which we have already discussed in Section 4.2 (using a similar argument). Here we employ the following notation: the Jacobi theta function is defined as

$$\theta(z) = \sum_{n=-\infty}^{\infty} \exp(\pi i n^2 z) \ , \quad \Im z > 0 \ .$$

Exercise 5.11. $\theta(-1/z) = \sqrt{-iz}\,\theta(z)$, where the principal branch of the square root is taken.

Digression: Riemann ζ-fiunction We now use the Jacobi identity to derive the functional equation for the Riemann ζ-function. For $\Re s > 1$, denote

$$\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-s/2} \times \left[\int_0^\infty u^{\frac{s}{2}} e^{-u}\frac{du}{u}\right] \times \left[\sum_{n=1}^{\infty} n^{-s}\right] \ .$$

Theorem 5.12. *$\xi(s)$ admits a meromorphic continuation to $\mathbb{C}$, with simple poles at $s = 0, 1$ and no other poles. It satisfies the functional equation $\xi(1-s) = \xi(s)$.*

We follow Green [2016].

Proof. We first observe that for $s > 1$

$$\xi(s) = \int_0^\infty \left(\frac{\theta(iu) - 1}{2}\right) u^{\frac{s}{2}} \frac{du}{u} \ . \tag{5.3}$$

Indeed,

$$\int_0^\infty \exp(-\pi n^2 u)\, u^{\frac{s}{2}} \frac{du}{u} = \pi^{-\frac{s}{2}}\Gamma(-s/2) n^{-s} \ .$$

Now we split the integral (5.3) in two parts. The integral from 1 to ∞ is already an entire function, therefore we keep it as it is. The integral from 0 to 1 can be transformed using Exercise 5.11:

$$\begin{aligned}
\int_0^1 \left(\frac{\theta(iu)-1}{2}\right) u^{\frac{s}{2}} \frac{du}{u} &= \frac{1}{2}\int_0^1 \theta(iu) u^{\frac{s}{2}} \frac{du}{u} - \frac{1}{s} \\
&= \frac{1}{2}\int_1^\infty \theta(-(iu)^{-1}) u^{-\frac{s}{2}} \frac{du}{u} - \frac{1}{s} = \frac{1}{2}\int_1^\infty \sqrt{u}\theta(iu) u^{-\frac{s}{2}} \frac{du}{u} - \frac{1}{s} \\
&= \frac{1}{2}\int_1^\infty \theta(iu) u^{\frac{1-s}{2}} \frac{du}{u} - \frac{1}{s} = \int_1^\infty \left(\frac{\theta(iu)-1}{2}\right) u^{\frac{1-s}{2}} \frac{du}{u} - \frac{1}{s} - \frac{1}{1-s} \ .
\end{aligned}$$

Thus

$$\xi(s) = \int_1^\infty \left(\frac{\theta(iu)-1}{2}\right) \left[u^{\frac{s}{2}} + u^{\frac{1-s}{2}}\right] \frac{du}{u} - \frac{1}{s} - \frac{1}{1-s}$$

which enjoys the claimed properties. □

Now we can show that ζ has no zeros on the boundary of the critical strip.

Theorem 5.13. *The ζ-function has no zeros on the line $\Re s = 1$.*

This theorem is the key ingredient in the proof of the prime number theorem,

$$(\text{the number of primes} \le x) = (1 + o(1))\frac{x}{\log x} \ , \quad x \to +\infty \ .$$

See further Green [2016] or Dym and McKean [1972].

Proof. Following Gorin [2011/12], let $\sigma > 1$, and consider the function $f_\sigma(t) = \log|\zeta(\sigma + it)|$. By the Euler product formula

$$\zeta(s) = \prod_{\text{prime } p} (1 - p^{-s})^{-1}$$

we have:

$$\log \zeta(s = \sigma + it) = -\sum_p \log(1 - p^{-s}) = \sum_p \sum_{n \ge 1} \frac{1}{np^{sn}} = \sum_p \sum_{n \ge 1} \frac{1}{np^{\sigma n}} e_{n \log p}(t) \ .$$

Therefore

$$f_\sigma(t) = \sum_p \sum_{n \ge 1} \frac{1}{np^{\sigma n}} [e_{n \log p}(t) + e_{-n \log p}(t)] \ .$$

The key feature of this formula is that all the coefficients are positive, i.e.

$$\sum_{j,k=1}^N f_\sigma(t_j - t_k) c_j \overline{c_k} \ge 0$$

for any $c_j \in \mathbb{C}$, $t_j \in \mathbb{R}$ (does this remind you of (4.28)?). Take $t_1 = t > 0$, $t_2 = 0$, $t_3 = -t$, and $c_1 = c_2 = c_3 = 1$. We obtain:

$$3f_\sigma(0) + 4f_\sigma(t) + 2f_\sigma(2t) \geq 0 \ .$$

As $\sigma \to 1+0$, $f_\sigma(t) \sim -m(t)|\log(\sigma - 1)|$, where $m(t) = +m$ if $1 + it$ is a zero of multiplicity m and $m(t) = -m$ if t is a pole of multiplicity m. By Theorem 5.12, $m(0) = -1$ and $m(t) \geq 0$ for any other t, therefore $m(t) \leq 3/4$, whence it is zero. □

5.4 The uncertainty principle*

A digression to quantum mechanics In the Schrödinger picture, a state of a system is described by a unit vector $\psi \in H$, where H is a Hilbert space. An observable is described by a self-adjoint operator acting on H. The average of an observable A in a state ψ is given by $\mathbb{E}_\psi A = \langle A\psi, \psi\rangle$. The variance of A is

$$\mathrm{Var}_\psi A = \langle A^2\psi, \psi\rangle - \langle A\psi, \psi\rangle^2 = \mathbb{E}_\psi (A - (\mathbb{E}_\psi A)\mathbb{1})^2 \ .$$

One can further construct a random variable $\mathfrak{a}$ such that for any decent function u one has $\langle u(A)\psi, \psi\rangle = \mathbb{E}u(\mathfrak{a})$. The probability distribution of $\mathfrak{a}$ is called the distribution of A in the state ψ. If ψ is an eigenvector of A with eigenvalue λ, then $\langle u(A)\psi, \psi\rangle = u(\lambda)$, hence $\mathfrak{a} \equiv \lambda$. If ψ is not an eigenvalue of A, then the distribution of A is non-trivial.

In classical mechanics, one can measure the position x and the momentum p of a (say, one-dimensional) particle; these are examples of classical observables. In quantum mechanics, these are reblaced by quantum observables, X and P. The fundamental relation is $XP - PX = i\hbar\mathbb{1}$, where $\hbar \approx 10^{-34}\,\mathrm{J}\cdot\mathrm{sec}$ is a structural constant of quantum mechanics (the reduced Planck constant). It implies that X and P can not have joint eigenvectors, i.e. for each ψ at least one of them has non-zero variance.

The description of quantum mechanics is unitary-invariant, and does not depend on the choice of the underlying Hilbert space H. However, it is often useful to have in mind a realisation; a particularly convenient one is $H = L_2(\mathbb{R})$ (the space of Lebesgue square-integrable functions), X is the operator of multiplication by x, and p is a multiple of the differentiation operator:

$$(X\psi)(x) = x\psi(x) \ , \quad (P\psi)(x) = -i\hbar\psi'(x) \ .$$

Then indeed $(XP - PX)\psi = i\hbar\psi$ (at least, this relation formally holds for polynomials $\psi(x)$). For the sake of concreteness, we confine ourselves to this realisation.

Exercise 5.14. For any $f, g \in S(\mathbb{R})$, $(XP - PX)f = -i\hbar f$.

Theorem 5.15 (Heisenberg uncertainty principle). *For any $\psi \in S(\mathbb{R})$ such that $\|\psi\| = 1$,*

$$\mathrm{Var}_\psi X \times \mathrm{Var}_\psi P \geq \frac{\hbar^2}{4} \ .$$

Proof. By the commutation relation,

$$\begin{aligned} i\hbar = i\hbar\langle \psi, \psi\rangle &= \langle (XP - PX)\psi, \psi\rangle \\ &= \langle ((X - (\mathbb{E}_\psi X)\,\mathbb{1})(P - (\mathbb{E}_\psi P)\,\mathbb{1}) - (P - (\mathbb{E}P)\,\mathbb{1})(X - (\mathbb{E}_\psi X)\,\mathbb{1}))\,\psi, \psi\rangle \ . \end{aligned}$$

Now we estimate

$$\begin{aligned} |\langle (X - (\mathbb{E}_\psi X)\,\mathbb{1})(P - (\mathbb{E}_\psi P)\,\mathbb{1})\psi, \psi\rangle| &= |\langle (P - (\mathbb{E}_\psi P)\,\mathbb{1})\psi, (X - (\mathbb{E}_\psi X)\mathbb{1})\psi\rangle| \\ &\leq \|(P - (\mathbb{E}_\psi P)\,\mathbb{1})\psi\|\|(X - (\mathbb{E}_\psi X)\mathbb{1})\psi\| = \sqrt{\mathrm{Var}_\psi X \times \mathrm{Var}_\psi P} \ , \end{aligned}$$

and we are done. □

Back to harmonic analysis We now return to the mathematical framework of harmonic analysis, and from now on we set $\hbar = \frac{1}{2\pi}$. With this choice,

$$\mathcal{F}^* P \mathcal{F} = X \ ,$$

and we obtain:

Corollary 5.16. *For any $f \in S(\mathbb{R})$,*

$$\inf_{x_0, p_0 \in \mathbb{R}} \left[\int |x - x_0|^2 |f(x)|^2 dx \times \int |p - p_0|^2 |\hat{f}(p)|^2 dp \right] \geq \frac{\|f\|_2^4}{16\pi^2} \ . \qquad (5.4)$$

With some more work, one can show that this inequality holds whenever the quantities that appear in it are finite.

Fourier transform of compactly supported functions The inequality (5.4) is one of many results asserting that f and $\hat{f}$ can not be simultaneously localised. Here is another one:

Exercise 5.17. Suppose $f \in S(\mathbb{R})$ is compactly supported, and also $\hat{f}$ is compactly supported. Then $f \equiv 0$.

We now prove a much stronger statement.

Theorem 5.18 (Paley–Wiener). *Let $R > 0$. A function $f : \mathbb{R} \to \mathbb{C}$ can be analytically continued to an entire function satisfying the estimates*

$$\forall k \geq 0 \quad \sup_{z \in \mathbb{C}} |f(z)|(1 + |z|^k) \exp(-2\pi R |\Im z|) < \infty \qquad (5.5)$$

if and only if $f \in S(\mathbb{R})$ and $\operatorname{supp} \hat{f} \subset [-R, R]$.

Proof. Suppose $f \in S(\mathbb{R})$ and $\operatorname{supp} \hat{f} \subset [-R, R]$. Then we can define an analytic extension

$$f(x+iy) = \int_{-R}^{R} \hat{f}(p) \exp(2\pi i p(x+iy)) dp \ ,$$

and

$$|f(x+iy)| \le \int_{-R}^{R} |\hat{f}(p)| \exp(2\pi |p||y|) dp \le \int_{-R}^{R} |\hat{f}(p)| dp \, \exp(2\pi R|y|) = C_0 \exp(2\pi R|y|) \ .$$

This proves the case $k = 0$ of (5.5). Applying this to the functions $f_k(x) = f(x)x^k$, we obtain:

$$|f(x+iy)||x+iy|^k \le C_k \exp(2\pi R|y|) \ .$$

Vice versa, if f satisfies (5.5), its restriction to $\mathbb{R}$ definitely lies in the Schwartz space;

$$\hat{f}(p) = \int f(x) \exp(-2\pi i x p) dx \ .$$

For $p > 0$, we can deform the contour of integration to the line $\Im z = -r < 0$. We get:

$$|\hat{f}(p)| \le \int_{-\infty}^{\infty} |f(x-ir)| \exp(-2\pi r p) dx \le \int_{-\infty}^{\infty} \frac{C_2 \exp(+2\pi r R)}{1+x^2} \exp(-2\pi r p) dx$$

and for $p > R$ this tends to zero as $r \to \infty$. Similarly, $\hat{f}(p)$ vanishes for $p < -R$. □

See e.g. Dym and McKean [1972] for the more delicate L_2 version of this result.

In fact, functions with compactly supported Fourier transform share many properties with (trigonometric) polynomials. Here is a Bernstein-type inequality:

Exercise 5.19. If $f \in S(\mathbb{R})$ and $\operatorname{supp} \hat{f} \subset [-R, R]$, then $\|f'\|_2 \le 2\pi R \|f\|_2$.

A similar inequality is true for the sup-norm (cf. Theorem 4.5).

Exercise 5.20. Let $f \in S(\mathbb{R})$, $\operatorname{supp} \hat{f} \subset [-R, R]$. Prove that

$$f(x) = \sum_{n=-\infty}^{\infty} f(n/R) \frac{\sin(\pi(xR-n))}{\pi(xR-n)} \ .$$

Consequently, if $f, g \in S(\mathbb{R})$, $\operatorname{supp} \hat{f}, \hat{g} \subset [-R, R]$, then

$$\int f(x)\overline{g(x)} dx = \sum_{n=-\infty}^{\infty} f(n/R)\overline{g(n/R)} \ .$$

Some notation

$e_p(x)$	$\exp(2\pi ipx/N)$ (Sections 1, 2); $\exp(2\pi ipx)$ (Sections 3–5)
$\mathbb{T}$	The circle $\mathbb{R}/\mathbb{Z}$
$\hat{f}$	Fourier coefficient of f (Sections 1–4); Fourier transform of f (Section 5)
$\mathcal{F}$	The map $f \mapsto \hat{f}$
$\|f\| = \|f\|_2$	$[\int \lvert f(x)\rvert^2 dx]^{1/2}$
$\|f\|_\infty$	$\sup \lvert f\rvert$
$D_N; D_R$	Dirichlet kernel, (3.7), (5.1)
$\omega_f(\delta)$	Modulus of continuity; see (4.2)
$E_N(f)$	the measure of approximation of f by (trigonometric) polynomials; see (4.1)
$S(\mathbb{R})$	Schwartz space

References

Lars V. Ahlfors. *Complex analysis.* McGraw-Hill Book Co., New York, third edition, 1978. ISBN 0-07-000657-1. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

H. Dym and H. P. McKean. *Fourier series and integrals.* Academic Press, New York-London, 1972. Probability and Mathematical Statistics, No. 14.

A. A. Giannopoulos and V. D. Milman. Concentration property on probability spaces. *Adv. Math.*, 156(1):77–106, 2000. ISSN 0001-8708. doi: 10.1006/aima.2000.1949.

E. A. Gorin. Positive-definite functions as a tool of mathematical analysis. *Fundam. Prikl. Mat.*, 17(7):67–95, 2011/12. ISSN 1560-5159. doi: 10.1007/s10958-014-1730-5.

W. T. Gowers. Fourier analysis and Szemerédi's theorem. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, number Extra Vol. I, pages 617–629, 1998.

Ben Green. Analytic number theory. 2016. URL `http://people.maths.ox.ac.uk/greenbj/papers/primenumbers.pdf`.

Ulf Grenander and Gábor Szegő. *Toeplitz forms and their applications.* Chelsea Publishing Co., New York, second edition, 1984. ISBN 0-8284-0321-X.

Yitzhak Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004. ISBN 0-521-83829-0; 0-521-54359-2. doi: 10.1017/CBO9781139165372.

T. W. Körner. *Fourier analysis*. Cambridge University Press, Cambridge, second edition, 1989. ISBN 0-521-38991-7.

Cornelius Lanczos. *Discourse on Fourier series*. Hafner Publishing Co., New York, 1966.

Michel Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001. ISBN 0-8218-2864-9.

Hugh L. Montgomery. *Early Fourier analysis*, volume 22 of *Pure and Applied Undergraduate Texts*. American Mathematical Society, Providence, RI, 2014. ISBN 978-1-4704-1560-0.

Klaus Roth. Sur quelques ensembles d'entiers. *C. R. Acad. Sci. Paris*, 234: 388–390, 1952. ISSN 0001-4036.

K. Soundararajan. Additive combinatorics: Winter 2007. `http://math.stanford.edu/~ksound/Notes.pdf`, 2010.

Elias M. Stein and Rami Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. ISBN 0-691-11384-X. An introduction.