dichotomy between randomness and structure in a number of ways, most strikingly in proving Roth's celebrated theorem (which we discuss in Chapter 10) that subsets of integers of positive upper density contain infinitely many progressions of length 3. (Progressions of higher length cannot be treated by linear Fourier techniques, requiring either higher order Fourier analysis or other approaches; see Chapter 11.)

Fourier analysis can be performed on any additive group Z (and even on non-abelian groups). However, we shall only need this transform on finite groups, where the theory is slightly simpler technically. Thus we shall restrict our attention exclusively to the finite case. The cases $Z = \mathbf{Z}$, $Z = \mathbf{R}/\mathbf{Z}$, and $Z = \mathbf{R}$ are also of importance to additive combinatorics (in particular leading to the *Hardy–Littlewood circle method* in analytic number theory), but it turns out that the finite Fourier theory forms an acceptable substitute for these infinite Fourier theories in our applications.

4.1 Basic theory

Let Z be a finite additive group (for instance, Z could be a cyclic group $Z = \mathbf{Z}_N$). In this section we recall the basic theory of the finite Fourier transform on such groups.

Fourier analysis relies on the duality between a group Z and its *Pontryagin dual* $\hat{Z}$, which can be defined as the space of homomorphisms from Z to the circle group $\mathbf{R}/\mathbf{Z}$. In the case of finite groups, it turns out that a group Z and its Pontryagin dual $\hat{Z}$ are always isomorphic, and so it shall be convenient to identify the two in order to simplify the theory slightly. This can be done by means of a non-degenerate bilinear form:

Definition 4.1 (Bilinear forms) A *bilinear form* on an additive group Z is a map $(\xi, x) \mapsto \xi \cdot x$ from $Z \times Z$ to $\mathbf{R}/\mathbf{Z}$, which is a homomorphism in each of the variables ξ, x separately. We say that the form is *non-degenerate* if for every non-zero ξ the map $x \mapsto \xi \cdot x$ is not identically zero, and similarly for every non-zero x the map $\xi \mapsto \xi \cdot x$ is not identically zero. We say the form is *symmetric* if $\xi \cdot x = x \cdot \xi$.

Examples 4.2 If Z is a cyclic group $\mathbf{Z}_N$ then the bilinear form $x \cdot \xi := x\xi/N$ is symmetric and non-degenerate. If Z is a standard vector space F^n over a finite field F, then the bilinear form $(x_1, \ldots, x_n) \cdot (\xi_1, \ldots, \xi_n) := \phi(x_1\xi_1 + \cdots + x_n\xi_n)$ is symmetric and non-degenerate whenever $\phi : F \to \mathbf{R}/\mathbf{Z}$ is any non-trivial homomorphism from F to $\mathbf{R}/\mathbf{Z}$ (e.g. if $F = \mathbf{Z}_p$ we can take $\phi(x) := x/p$). This particular choice has the useful additional property that $a\xi \cdot x = \xi \cdot ax$ for all $a \in F$ and $x, \xi \in Z$.

Lemma 4.3 (Existence of bilinear forms) *Every finite additive group Z has at least one non-degenerate symmetric bilinear form.*

Proof From Corollary 3.8 we know that every finite additive group is the direct sum of cyclic groups. We have already seen in Example 4.2 that each cyclic group has a symmetric non-degenerate bilinear form. Finally, observe that if Z_1 and Z_2 have symmetric non-degenerate bilinear forms, then the direct sum $Z_1 \oplus Z_2$ also has a symmetric non-degenerate bilinear form, defined by $(\xi_1, \xi_2) \cdot (x_1, x_2) := \xi_1 \cdot x_1 + \xi_2 \cdot x_2$. The claim follows. □

Remark 4.4 A given additive group Z generally has multiple bilinear forms (see Exercise 4.1.10), but from the point of view of Fourier analysis they are all equivalent[1]. The symmetry property has some minor aesthetic advantages but is not essential to the Fourier theory, as the physical space variable and the frequency space variable usually play completely different roles.

Henceforth we fix a finite additive group Z, equipped with a non-degenerate symmetric bilinear form $\xi \cdot x$; in practice we shall usually use one of the two examples from Example 4.2.

To perform Fourier analysis, it will be convenient to adopt the following "ergodic" notation. Let $\mathbf{C}^Z$ denote the space of all complex-valued functions $f : Z \to \mathbf{C}$. If $f \in \mathbf{C}^Z$, we define the *mean* or *expectation* of f to be the quantity

$$\mathbf{E}_Z(f) = \mathbf{E}_{x \in Z} f(x) := \frac{1}{|Z|} \sum_{x \in Z} f(x).$$

Similarly, if $A \subseteq Z$, we define the *density* or *probability* of A as

$$\mathbf{P}_Z(A) = \mathbf{P}_{x \in Z}(x \in A) := \mathbf{E}_Z(1_A) = \frac{|A|}{|Z|}.$$

We can generalize this notation to other finite non-empty domains than Z, thus for instance $\mathbf{E}_{x \in A, y \in B} f(x, y) := \frac{1}{|A||B|} \sum_{x \in A, y \in B} f(x, y)$. This notation not only suggests the connections between Fourier analysis, ergodic theory, and probability, but is also useful in concealing from view a number of normalizing powers of $|Z|$ which would otherwise clutter the estimates. Generally, we shall use this ergodic notation for the physical variable, but use the discrete notation $\sum_{\xi \in Z} f(\xi)$ and $|A|$ (without the normalizing $|Z|$ factor) for the frequency variable. We shall also rely

[1] One way of viewing this is that the identification between $\hat{Z}$ and Z is non-canonical, and one should really be placing the frequency variable in $\hat{Z}$ instead of Z. This is ultimately the more correct viewpoint; however since we shall usually be working in very concrete situations such as cyclic groups $\mathbf{Z}_N$, where one does have a standard identification, we have chosen to rely on the bilinear form approach here rather than the abstract approach.

heavily on the exponential map $e : \mathbf{R}/\mathbf{Z} \to \mathbf{C}$, defined by

$$e(\theta) := e^{2\pi i\theta}. \tag{4.1}$$

The following two orthogonality properties form the foundation for Fourier analysis.

Lemma 4.5 (Orthogonality properties) *For any $\xi, \xi' \in Z$ we have*

$$\mathbf{E}_{x\in Z} e(\xi \cdot x)\overline{e(\xi' \cdot x)} = \mathbf{I}(\xi = \xi')$$

and for any $x, x' \in Z$ we have

$$\sum_{\xi\in Z} e(\xi \cdot x)\overline{e(\xi \cdot x')} = |Z|\mathbf{I}(x = x').$$

Proof We prove the first identity only, as the second is similar. Since $e(\xi \cdot x)\overline{e(\xi' \cdot x)} = e((\xi - \xi') \cdot x)$, it will suffice to show the claim in the $\xi' = 0$ case, i.e. it suffices to show

$$\mathbf{E}_{x\in Z} e(\xi \cdot x) = \mathbf{I}(\xi = 0).$$

This is clear in the case $\xi = 0$. If $\xi \neq 0$, then by non-degeneracy there exists $h \in Z$ such that $e(\xi \cdot h) \neq 1$. Shifting x by h we then have

$$\mathbf{E}_{x\in Z} e(\xi \cdot x) = \mathbf{E}_{x\in Z} e(\xi \cdot (x + h)) = e(\xi \cdot h)\mathbf{E}_{x\in Z} e(\xi \cdot x)$$

and hence $\mathbf{E}_{x\in Z} e(\xi \cdot x) = 0 = \mathbf{I}(\xi = 0)$ as desired. □

For every $\xi \in Z$, we can define the associated *character* $e_\xi \in \mathbf{C}^Z$ by $e_\xi(x) := e(\xi \cdot x)$. The above lemma then shows that the e_ξ are an orthonormal system in $\mathbf{C}^Z$, with respect to the complex Hilbert space structure

$$\langle f, g\rangle_{\mathbf{C}^Z} := \mathbf{E}_Z(f\overline{g}) = \mathbf{E}_{x\in Z} f(x)\overline{g(x)}.$$

Since the number $|Z|$ of characters equals the dimension $|Z|$ of the space, we see that this system is in fact a *complete* orthonormal system. This motivates

Definition 4.6 (Fourier transform) If $f \in \mathbf{C}^Z$, we define the *Fourier transform* $\hat{f} \in \mathbf{C}^Z$ by the formula

$$\hat{f}(\xi) := \langle f, e_\xi\rangle_{\mathbf{C}^Z} = \mathbf{E}_{x\in Z} f(x)\overline{e(\xi \cdot x)}.$$

We refer to $\hat{f}(\xi)$ as the *Fourier coefficient* of f at the frequency (or *mode*) ξ.

Since the e_ξ are a complete orthonormal basis, we have the *Parseval identity*

$$(\mathbf{E}_Z|f|^2)^{1/2} = \left(\sum_{\xi\in Z} |\hat{f}(\xi)|^2\right)^{1/2} \tag{4.2}$$

the *Plancherel theorem*

$$\langle f, g\rangle_{\mathbf{C}^Z} = \sum_{\xi\in Z} \hat{f}(\xi)\overline{\hat{g}(\xi)} \tag{4.3}$$

and the *Fourier inversion formula*

$$f = \sum_{\xi\in Z} \hat{f}(\xi) e_\xi. \tag{4.4}$$

In particular we see that two functions are equal if and only if their Fourier coefficients match at every frequency. In other words, the Fourier transform is a bijection from $\mathbf{C}^Z$ to $\mathbf{C}^Z$ (in fact it is a unitary isometry, thanks to (4.2), (4.3)).

From Lemma 4.5 we see that the Fourier coefficients of a character e_ξ are just a Kronecker delta function:

$$\widehat{e_\xi}(\xi') = \mathbf{I}(\xi = \xi').$$

In particular $\hat{1}(\xi) = \mathbf{I}(\xi = 0)$.

A special role in the additive theory of the Fourier transform is played by the *zero frequency* $\xi = 0$. This is because the zero Fourier coefficient is same concept as expectation:

$$\hat{f}(0) = \langle f, 1\rangle_{\mathbf{C}^Z} = \mathbf{E}_Z(f). \tag{4.5}$$

If S is any subset of Z, define the *orthogonal complement* $S^\perp \subseteq Z$ of S to be the set

$$S^\perp := \{\xi \in Z : \xi \cdot x = 0 \text{ for all } x \in S\}.$$

One can easily verify that $S^\perp$ is a subgroup of Z. Also one has the pleasant identity

$$\widehat{1_G} = \mathbf{P}_Z(G) 1_{G^\perp} \tag{4.6}$$

whenever G is a subgroup; see Exercise 4.1.6. Applying (4.2) we see in particular that

$$|G||G^\perp| = |Z|. \tag{4.7}$$

We now introduce the fundamental notion of *convolution*, which links the Fourier transform to the theory of sum sets.

Definition 4.7 (Convolution) If $f, g \in L^2(Z)$ are random variables, we define their *convolution* $f * g$ to be the random variable

$$f * g(x) = \mathbf{E}_{y\in Z} f(x-y)g(y) = \mathbf{E}_{y\in Z} f(y)g(x-y).$$

We also define the *support* $\mathrm{supp}(f)$ of f to be the set $\mathrm{supp}(f) = \{f \neq 0\} = \{x \in Z : f(x) \neq 0\}$.

The significance of convolution to sum sets lies in the obvious inclusion

$$\mathrm{supp}(f * g) \subseteq \mathrm{supp}(f) + \mathrm{supp}(g)$$

and particularly in the identity

$$A + B = \mathrm{supp}(1_A * 1_B).$$

Indeed we have the more precise statement

$$1_A * 1_B(x) := \mathbf{P}_Z(A \cap (x - B)). \tag{4.8}$$

The relevance of the Fourier transform to convolution lies in the easily verified identity

$$\widehat{f * g} = \hat{f} \cdot \hat{g} \tag{4.9}$$

Applying (4.9) at the zero frequency we have the basic formula

$$\mathbf{E}_Z(f * g) = (\mathbf{E}_Z f) \cdot (\mathbf{E}_Z g). \tag{4.10}$$

In particular, if f or g has mean zero, then so does $f * g$.

As one consequence of (4.9) we see that convolution is bilinear, symmetric, and associative. We also have a dual version of (4.9), namely the formula

$$\widehat{fg}(\xi) = \sum_{\eta \in Z} \hat{f}(\eta)\hat{g}(\xi - \eta) \tag{4.11}$$

which converts pointwise product back to convolution; we leave the verification of these identities as an exercise.

In the exercises below, Z is a fixed finite additive group, with a fixed symmetric non-degenerate bilinear form $\cdot$.

Exercises

4.1.1 Let $\hat{Z}$ be the additive group consisting of all the homomorphisms from Z to $\mathbf{R}/\mathbf{Z}$. Show that the identification of a frequency $\xi \in Z$ with the homomorphism $x \mapsto \xi \cdot x$ gives an isomorphism from Z to $\hat{Z}$.

4.1.2 Define a *character* to be any map $\chi : Z \to \mathbf{C}$ with $\chi(0) = 1$ and $\chi(x + y) = \chi(x)\chi(y)$ for all $x, y \in Z$. Show that the set of all characters is precisely $\{e_\xi : \xi \in Z\}$.

4.1.3 Show that for any $\xi \in Z$, e_ξ takes values in the $|Z|$th roots of unity.

4.1.4 Define a *linear phase function* to be any map $\phi : Z \to \mathbf{R}/\mathbf{Z}$ with the property that

$$\phi(x + h_1 + h_2) - \phi(x + h_1) - \phi(x + h_2) + \phi(x) = 0 \text{ for all } x, h_1, h_2 \in Z.$$

Show that $\phi : Z \to \mathbf{R}/\mathbf{Z}$ is a linear phase function if and only if there exists $\xi \in Z$ and $c \in \mathbf{R}/\mathbf{Z}$ such that $\phi(x) = \xi \cdot x + c$ for all c. (The map ϕ is also a *Freiman homomorphism of order* 2; see Definition 5.21.)

4.1.5 Let x be an element of Z chosen uniformly at random. Show that the random variables $\{e_\xi(x) : \xi \in Z\}$ are pairwise independent, and have variance 1 and mean zero for $\xi \neq 0$, and variance 0 and mean 1 for $\xi = 0$. Use this and (1.9), (4.4) to give an alternative proof of (4.2).

4.1.6 Prove (4.6).

4.1.7 Let $f : Z \to \mathbf{C}$. If H is a subgroup of Z, and $g := f 1_H$, show that

$$\hat{g}(\xi) = \mathbf{E}_{\eta \in H^\perp} \hat{f}(\xi + \eta) \text{ for all } \xi \in Z$$

and conclude in particular the *Poisson summation formula*

$$\mathbf{E}_{x \in H} f(x) = \mathbf{E}_{\xi \in H^\perp} \hat{f}(\xi).$$

In the converse direction, if $h = f * \frac{1}{\mathbf{P}_Z(H)} 1_H$ is the average of f on cosets of H, i.e.

$$h(x) := \mathbf{E}_{y \in H} f(x + y),$$

show that $\hat{h} = \hat{f} \cdot 1_{H^\perp}$.

4.1.8 If $\phi : Z \to Z$ is a group isomorphism of Z, then there exists a unique group isomorphism $\phi^\dagger : Z \to Z$, called the *adjoint* of ϕ, such that $\xi \cdot \phi(x) = \phi^\dagger(\xi) \cdot x$ for all $x, \xi \in Z$. Furthermore if $g(x) = f(\phi(x))$ for all $x \in Z$ then $\hat{g}(x) = f((\phi^\dagger)^{-1}(x))$ for all $x \in Z$.

4.1.9 If $\phi : Z \to Z$ and $\psi : Z \to Z$ are group isomorphisms, show that $(\phi \circ \psi)^\dagger = \psi^\dagger \circ \phi^\dagger$.

4.1.10 Let $\bullet : Z \times Z \to \mathbf{R}/\mathbf{Z}$ and $\tilde{\bullet} : Z \times Z \to C$ be two non-degenerate symmetric bilinear forms on a finite additive group Z. Show that there exists a self-adjoint group isomorphism $\phi : Z \to Z$ such that $\xi \tilde{\bullet} x = \xi \bullet \phi(x) = \phi^\dagger(\xi) \bullet x$ for all $x, \xi \in Z$. This shows that all Fourier transforms are equivalent up to isomorphisms of either the x or ξ variable.

4.1.11 Prove (4.9) and (4.11).

4.1.12 Let x be an element of Z chosen uniformly at random, and let $\xi_1, \ldots, \xi_n \in Z$. Show that the random variables $e_{\xi_1}(x), \ldots, e_{\xi_n}(x)$ are jointly independent if and only if the group $\langle \xi_1, \ldots, \xi_n \rangle$ generated by $\xi_1, \ldots, \xi_n$ has order $\mathrm{ord}(\xi_1) \ldots \mathrm{ord}(\xi_n)$.

4.1.13 Let G, H be two subgroups of Z. Show that $(G + H)^\perp = G^\perp \cap H^\perp$, $(G \cap H)^\perp = G^\perp + H^\perp$, and $d(G^\perp, H^\perp) = d(G, H)$, where d is the Ruzsa distance defined in Definition 2.5. This may help explain the symmetric nature of $G + H$ and $G \cap H$ in the estimates in Exercise 2.3.11.