LECTURE NOTES 9 FOR 247B

TERENCE TAO

1. Fourier analysis on finite abelian groups

We have been using Fourier analysis on the Euclidean group $\mathbf{R}^d$ (and to a lesser extent, on the toral group $\mathbf{T}^d$) for some time now. It turns out that Fourier analysis can in fact be formalised on any locally compact Hausdorff abelian group. We will come to this generalisation later in this set of notes, but to keep things simple for now we will work with the special setting of *finite* abelian groups, in which qualitative issues such as integrability, measurability, continuity, etc. are irrelevant, and one can do things like induct on dimension without having to invoke things like Zorn's lemma.

Our viewpoint shall be representation-theoretic. Other approaches to the subject include the Gelfand-theoretic approach (viewing the convolution algebra as a B^*-algebra), spectral-theoretic (focusing on trying to diagonalise various translation-invariant operators), or algebraic (using explicit classifications of the group and its characters). Each of these approaches is worthwhile (and they are of course closely related to each other), but we select the representation-theoretic approach as it extends relatively easily to the non-abelian case, whereas the other approaches have more difficulty.

In contrast to previous notes, the material here will be largely algebraic in nature; what analysis there is here is largely of a qualitative nature (e.g. dealing with issues of continuity, integrability, etc.) rather than quantitative (estimates, etc.).

We now turn to the details. Fix a finite additive group $(G, +)$, where we use the addition symbol for the group operation to emphasise the abelian nature of the group; in particular we use 0 for the group identity and $-x$ for the group inverse of x. It will not be terribly relevant for this finitary analysis, but we give G the discrete topology and the discrete σ-algebra. Let $\#G$ denote the cardinality of G. We define the *normalised Haar measure* dx on this group to be the normalised counting measure

$$\int_G f(x)\ dx := \frac{1}{\#G} \sum_{x \in G} f(x).$$

The normalising factor $\frac{1}{\#G}$ will almost never be seen again, being concealed within the measure dx at all times.

We have the (finite-dimensional) Hilbert space $L^2(G)$ of functions $f : G \to \mathbf{C}$, with inner product

$$\langle f, g \rangle_{L^2(G)} := \int_G f(x)\overline{g(x)}\ dx.$$

Note that as G is finite, all norms are equivalent, thus for instance $L^2(G) = L^1(G) = L^\infty(G)$. Every group element $y \in G$ gives rise to a translation operator $\mathrm{Trans}_y : L^2(G) \to L^2(G)$, defined by

$$\mathrm{Trans}_y f(x) := f(x - y);$$

observe each such operator is a unitary operator on $L^2(G)$, and furthermore the map $y \mapsto \mathrm{Trans}_y$ is a homomorphism:

$$\mathrm{Trans}_y \mathrm{Trans}_z = \mathrm{Trans}_{y+z}. \tag{1}$$

In other words, the map $y \mapsto \mathrm{Trans}_y$ is a *unitary representation* of the group G acting on the Hilbert space $L^2(G)$; this particular representation is known as the *regular representation.* One consequence of (1) is that all the translations commute with each other.

Given two functions $f, g \in L^2(G)$, we can form their *convolution* $f * g \in L^2(G)$ by the formula

$$f * g(x) := \int_G f(y) g(x - y)\ dy.$$

One easily verifies that convolution is bilinear, associative, and commutative (the latter of course relies on the abelian nature of G). There is also an identity element $\delta \in L^2(G)$, defined by $\delta(x) := (\#G)1_{\{0\}}$, thus $f * \delta = \delta * f = 1$.

Given any $g \in L^2(G)$, we can define the associated convolution operator $T_g : L^2(G) \to L^2(G)$ by $T_g f := g * f$. Equivalently, we have

$$T_g = \int_G g(y) \mathrm{Trans}_y\ dy,$$

thus convolution operators are nothing more than linear combinations of translation operators. The convolution operators commute with each other, and are translation-invariant; conversely, it is not hard to show that every translation-invariant operator is a convolution operator. (One can view the map $g \mapsto T_g$ as a representation of the convolution algebra $L^2(G)$ acting on itself.)

Define a *translation-invariant subspace* V of $L^2(G)$ to be any subspace of V which is invariant under all of the translations Trans_y, thus $\mathrm{Trans}_y V = V$ for all y; one can also think of such a space as a component of the regular representation of G. Equivalently, V is preserved by all of the operators T_g. Besides the trivial examples of $\{0\}$ and $L^2(G)$, two other simple examples of such spaces are the space $\{c : c \in G\}$ of constant functions, and the space $\{f \in L^2(G) : \int_G f = 0\}$ of mean-zero functions. Another important pair of examples: given any convolution operator T_g, the kernel $\{f \in L^2(G) : T_g f = 0\}$ and the range $\{T_g f : f \in L^2(G)\}$ are also translation-invariant (this is basically because convolution operators commute with translations). The vector sum or intersection of two translation-invariant spaces is translation-invariant; since all the translation operators are unitary, we also see that the orthogonal complement of a translation-invariant space is translation-invariant.

Example 1.1. *In the case of the cyclic group $G = \mathbf{Z}/N\mathbf{Z}$, given any $\xi \in \mathbf{Z}/N\mathbf{Z}$ we can form the one-dimensional translation-invariant space V_ξ generated by the character $e_\xi : x \mapsto e^{2\pi i x\xi/N}$. If we secretly allow ourselves the use of the finite Fourier transform, it is not hard to see that a space is translation-invariant iff it is the direct sum of some of these V_ξ, or equivalently if it is of the form $\{f \in L^2(G) : \hat{f}|_E = 0\}$ for some fixed set of frequencies $E \subset \mathbf{Z}/N\mathbf{Z}$. Readers familiar with the finite Fourier transform may find it instructive to keep this example in mind for the rest of this section.*

Since $T_g\delta = g$ for all $g \in L^2(G)$, we see that any translation-invariant space which contains the convolution identity δ, must be all of $L^2(G)$. Equivalently, a translation-invariant space is proper if and only if it avoids δ.

From these facts it is clear that (a) every proper translation-invariant subspace is contained in a maximal translation-invariant subspace; and (b) every translation-invariant subspace can be expressed (possibly non-uniquely) as the direct sum of *irreducible* translation-invariant subspaces, i.e. non-zero subspaces which cannot be nontrivially split as the sum of two smaller translation-invariant subspaces. It is not hard to show that a translation-invariant subspace is irreducible iff its orthogonal complement is maximal.

Remark 1.2. The orthogonal projection to a translation-invariant space is a translation-invariant operator, and thus given by convolution with some function μ, in particular $\mu * \mu = \mu$. Such functions are known as *idempotent measures* and were influential in the classical development of harmonic analysis. We have however structured our presentation here so that these measures are not needed.

We now come to the first really non-trivial result about these spaces. There are two equivalent forms of this result. The first is phrased in terms of maximal translation-invariant subspaces:

Proposition 1.3 (Gelfand-Mazur theorem, special case)**.** *All maximal translation-invariant subspaces are hyperplanes (i.e. their codimension is one).*

Proof Suppose for contradiction that we have a maximal translation-invariant subspace V of codimension two or greater. Observe that V is an ideal of the Banach algebra $L^2(G)$, and hence the quotient $L^2(G)/V$ is also a Banach algebra, with dimension at least two. In particular, it contains an element g which is not a multiple of the identity. From Liouville's theorem we know that $(g - z)^{-1}$ cannot exist for every complex z, thus there is a $z \in \mathbf{C}$ for which $g - z$ is non-invertible. Thus the kernel of $g - z$ is a proper non-zero ideal in $L^2(G)/V$, which induces a proper translation-invariant subspace of $L^2(G)$ which strictly contains V, contradicting maximality. ■

The second is phrased in terms of the orthogonal complement, irreducible translation-invariant subspaces:

Proposition 1.4. *All irreducible translation-invariant subspaces have dimension one.*

Proof Let V be an irreducible translation-invariant subspace, and suppose for contradiction that V had dimension at least two. Then V will contain a non-zero function f which vanishes at at least one point. Then there exists $y \in G$ such that $\mathrm{Trans}_y f$ is not a constant multiple of f (simply choose y so that Trans_y shifts a zero of f to a non-zero of f). In particular, Trans_y is not a constant multiple of the identity on V. On the other hand, it *is* a unitary transformation on V. Thus by the spectral theorem, it has an eigenvalue, and hence a non-trivial eigenspace W in V. Since all the translations commute with each other, we see that W is also translation-invariant, as is the orthogonal complement of W in V. But this contradicts the irreducibility of V. ■

Of course, the above two propositions are equivalent to each other; interestingly, they both secretly use complex analysis in the proof (the latter via the spectral theorem for unitary operators). In the special case of finite abelian groups, another way to establish these propositions is to use the classification of finite abelian groups and then explicitly construct a Fourier transform with which to analyse these translation-invariant spaces. This is the most direct proof, but relies heavily on the classification, which is not easily available for more general abelian groups.

Now let V be an irreducible translation-invariant subspace, thus one-dimensional. This means that each of the translation operators Trans_y acts by multiplication by a complex constant $\chi_V(y)$ on V; since Trans_y is unitary, we must have $|\chi_V(y)| = 1$. Also, from (1) we see that χ_V is a homomorphism: $\chi_V(y+z) = \chi_V(y)\chi_V(z)$. In other words, $\chi_V : G \to S^1$ is a *multiplicative character* of G; conversely, given any multiplicative character $\chi : G \to S^1$, the one-dimensional space generated by χ is clearly an irreducible translation-invariant space. Thus the irreducible translation-invariant spaces are in one-to-one correspondence with multiplicative characters. Introducing the exponential function $e : \mathbf{R}/\mathbf{Z} \to S^1$ by $e(x) := e^{2\pi i x}$, we can write any multiplicative character χ as $\chi = e(\xi)$, where $\xi : G \to \mathbf{R}/\mathbf{Z}$ is an *additive character*, i.e. an additive homomorphism from G to $\mathbf{R}/\mathbf{Z}$. Thus the irreducible translation-invariant spaces are also in one-to-one correspondence with additive characters.

Remark 1.5. As a corollary, one also concludes that every maximal translation-invariant subspace is the orthogonal complement of a multiplicative character χ.

Define the *Pontryagin dual* $\hat{G}$ of G to be the space of all additive characters of G; this is clearly an additive group. Given $x \in G$ and $\xi \in \hat{G}$, we shall also write $\xi \cdot x$ for $\xi(x) \in \mathbf{R}/\mathbf{Z}$. We refer to elements ξ of $\hat{G}$ as *frequencies*. Given a frequency ξ, the associated irreducible translation-invariant space V_ξ is described as the linear span of the multiplicative character $e_\xi : x \mapsto e(\xi \cdot x)$.

Lemma 1.6 (Orthogonality). *If ξ, η are two distinct frequencies, then the associated translation-invariant subspaces V_ξ and V_η are orthogonal to each other.*

Proof It suffices to show that e_ξ and e_η are orthogonal, or in other words that the expression

$$I := \int_G e(\xi \cdot x) e(-\eta \cdot x)\ dx$$

is zero. Shifting x by y we see that

$$I = e(\xi \cdot y)e(-\eta \cdot y)I$$

for all $y \in G$. But since ξ, η are distinct, then there exists y such that $\xi \cdot y \neq \eta \cdot y$, and so $I = 0$ as claimed. ∎

Splitting the regular representation into irreducible (and now also orthogonal) components, we conclude

Corollary 1.7 (Peter-Weyl theorem, finite abelian case)**.** $L^2(G) = \bigoplus_{\xi \in \hat{G}} V_\xi$. *In particular (by dimension count)* $\#G = \#\hat{G}$. *Equivalently, the space* $\{e_\xi : \xi \in \hat{G}\}$ *of multiplicative characters is an orthonormal basis for* $L^2(G)$.

Observe that if $f \in L^2(G)$ and $\xi \in \hat{G}$, the orthogonal projection of f to V_ξ is given by $\hat{f}(\xi)e_\xi$, where $\hat{f}(\xi) = \langle f, e_\xi \rangle_{L^2(G)} = \int_G f(x)e(-\xi \cdot x)\, dx$. We thus conclude

Corollary 1.8 (Fourier inversion formula)**.** *For any* $f \in L^2(G)$, *we have* $f = \sum_{\xi \in \hat{G}} \hat{f}(\xi)e_\xi$.

Other quick consequences are the *Plancherel identity*

$$\|f\|_{L^2(G)} = \|\hat{f}\|_{l^2(\hat{G})}$$

and more generally *Parseval identity*

$$\langle f, g \rangle_{L^2(G)} = \langle \hat{f}, \hat{g} \rangle_{l^2(\hat{G})}.$$

It is also easy to check the convolution identity

$$\widehat{f * g} = \hat{f}\hat{g}$$

and dually that

$$\widehat{fg} = \hat{f} * \hat{g}$$

where the $*$ on the right now refers to discrete convolution (using counting measure on $\hat{G}$ rather than normalised counting measure on G).

Observe that every $x \in G$ can be viewed as a character $x \mapsto \xi \cdot x$ on $\hat{G}$, thus providing a canonical map from G to $\hat{\hat{G}}$. This map is injective. To see this, suppose that x was in the kernel of this map, then $\xi \cdot x = 0$ for all $\xi \in \hat{G}$, or equivalently Trans_x fixes each of the characters e_ξ. By the Fourier inversion formula this implies that Trans_x fixes all functions, which is observe. From the Peter-Weyl theorem we also know that G and $\hat{\hat{G}}$ have the same cardinality. Thus the map is bijective. In other words, the Pontryagin dual of $\hat{G}$ is canonically identifiable with G itself.

Example 1.9. *When* $G = \mathbf{Z}/N\mathbf{Z}$, *each* $\xi \in \mathbf{Z}/N\mathbf{Z}$ *generates a character, defined by* $\xi \cdot x := \xi x/N$. *These are* N *distinct characters, and so by the Peter-Weyl theorem there are no further characters. Thus the abstract Fourier transform given here corresponds to the usual finite Fourier transform on* $\mathbf{Z}/N\mathbf{Z}$.

We have the usual duality relationship between translation and modulation: for every $f \in L^2(G)$ and $y \in G$, $\xi \in \hat{G}$ we have

$$\widehat{\mathrm{Trans}_y f} = \mathrm{Mod}_{-y} \hat{f}; \quad \widehat{\mathrm{Mod}_\xi f} = \mathrm{Trans}_\xi \hat{f}$$

where $\mathrm{Mod}_{-y} F(\xi) := e(-\xi \cdot y) F(\xi)$ and $\mathrm{Mod}_\xi f(x) := e(\xi \cdot x) f(x)$.

2. Locally compact abelian groups

Many (though not all) of the theory developed above continues to apply in the more general category of *locally compact abelian (LCA) groups* - topological groups which are both abelian and locally compact. (A topological group is a group with a topology under which the group operations are continuous. To avoid minor technicalities, we also assume that the topology is Hausdorff.) We will omit some details, and refer the reader to Rudin's " Fourier analysis on groups" for a full treatment.

Fix a LCA group G. The first task is to determine the analogue of the measure dx used earlier. This is *Haar measure* - a non-negative translation-invariant Radon measure which is not identically zero. The fact that such a measure exists at all is non-trivial - akin to the construction of Lebesgue measure, which is Haar measure for $\mathbf{R}^d$. Roughly speaking, the way the measure is constructed is to use covering by translates of a small neighbourhood of the origin to create some sort of outer measure, and then send that neighbourhood to zero (normalising the measure so that some fixed compact set has measure 1, say). We will not give the details here, as it is a little tricky. However, it is relatively easy to show that Haar measure is unique up to constants. For if dx and $\tilde{dx}$ are two Haar measures, then for any $f, g \in C_c(G)$ one easily verifies using Fubini's theorem that

$$\int_G \int_G g(y) f(x+y)\ \tilde{dx} dy = (\int_G f(x)\ \tilde{dx})(\int_G g(y) dy) = (\int_G f(x)\ dx)(\int_G g(-y)\ \tilde{dy})$$

and so by choosing g so that the two g-integrals are non-zero, we see that dx and $\tilde{dx}$ differ by a constant multiple. From this we also conclude that Haar measure is always reflection-symmetric, since the reflection of Haar measure is also translation-invariant and assigns the same weight to symmetric sets as the original measure.

We can now define convolution operators as before, but on $L^1(G)$ rather than $L^2(G)$. This turns $L^1(G)$ into a B^*-algebra (a Banach algebra with a conjugation, namely $\tilde{f}(x) := \overline{f(-x)}$). The analogue of "maximal translation-invariant subspace" is now "maximal convolution ideal" of $L^1(G)$. The Gelfand-Mazur theorem argument works in this setting and shows that such maximal convolution ideals have codimension one, and are therefore identifiable with non-trivial B^*-algebra homomorphisms λ from $L^1(G)$ to $\mathbf{C}$. By the duality of $L^1(G)$ and $L^\infty(G)$, such homomorphisms take the form $\lambda : f \mapsto \int_G f(x)\overline{\chi(x)}\ dx$ for some $\chi \in L^\infty(G)$. Using the homomorphism property and approximations to the identity we can show that

$$\lambda \circ \mathrm{Trans}_y = \overline{\chi(y)} \lambda;$$

since translations are continuous in $L^1(G)$ we conclude that χ is continuous. Also, the identity also reveals that χ is multiplicative: $\chi(x+y) = \chi(x)\chi(y)$. Since χ is also in $L^\infty(G)$ and non-trivial, we conclude that $|\chi| = 1$. Thus χ is a continuous multiplicative map from G to S^1, otherwise known as a *multiplicative character*. Again, we can write $\chi(x) = e(\xi \cdot x)$, where $\xi : G \to \mathbf{R}/\mathbf{Z}$ is a continuous homomorphism from G to $\mathbf{R}/\mathbf{Z}$, otherwise known as a *additive character* or *frequency.*

Let $\hat{G}$ be the collection of additive characters; this is an abelian group. We can define the Fourier transform $\hat{f} : \hat{G} \to \mathbf{C}$ for any $f \in L^1(G)$ by

$$\hat{f}(\xi) := \int_G f(x) e(-\xi \cdot x)\ dx.$$

Thus each frequency ξ generates a linear functional on $L^1(G)$, i.e. we have an embedding of $\hat{G}$ in $L^1(G)^*$. We give $\hat{G}$ the topology induced by the weak-* topology on $L^1(G)$; this is easily verified to make $\hat{G}$ a locally compact abelian group. By the preceding discussion, $\hat{f}$ is essentially the *Gelfand transform* of f in the B^*-algebra $L^1(G)$.

We let $A(\hat{G}) := \{\hat{f} : f \in L^1(G)\}$ (this is known as the *Wiener algebra* of $\hat{G}$.) One easily verifies that $\widehat{f * g} = \hat{f}\hat{g}$, and so $A(\hat{G})$ becomes a B^*-algebra with pointwise product (rather than convolution) as the multiplication operation. It is also translation-invariant. There is an analogue of the Riemann-Lebesgue lemma, which asserts that $A(\hat{G}) \subset C_0(\hat{G})$, where $C_0(\hat{G})$ is the space of continuous functions which go to zero at infinity. As $A(\hat{G})$ is also closed under conjugation and separates points, the Stone-Weierstrass theorem then shows that $A(\hat{G})$ is in fact dense in $C_0(\hat{G})$ in the uniform topology. This makes $A(\hat{G})$ a useful class of "test functions" to verify a variety of Fourier-analytic identities.

Now we move towards the all-important Fourier inversion formula. Let $M(\hat{G})$ be the space of finite Radon measures on $\hat{G}$. Given such a measure ν, one can define its inverse Fourier transform $\mathcal{F}^*\nu \in C(G)$ by

$$\mathcal{F}^*\nu(x) := \int_{\hat{G}} e(\xi \cdot x)\ d\nu(\xi).$$

We let $B(G)$ denote the space of all such inverse Fourier transforms. One easily verifies that this is a translation-invariant sub-algebra of $C(G)$. From Fubini's theorem we have the duality relationship

$$\int_G f \overline{\mathcal{F}^*\nu(x)}\ dx = \int_{\hat{G}} \hat{f}(\xi) \overline{d\nu(\xi)}$$

for all $f \in L^1(G)$ and $\nu \in M(\hat{G})$.

So far we have not shown that $\hat{G}$ is large in any sense. One convenient way to obtain this largeness is to show that $B(G)$ is large. A key tool for achieving this is

Theorem 2.1 (Bochner's theorem)**.** *Let $f \in C(G)$ be* positive semi-definite*, in the sense that we have $\langle f * \mu, \mu \rangle \geq 0$ for any finitely supported complex measure μ. Then $f \in B(G)$, and furthermore we have $f = \mathcal{F}^*\nu$ for a* non-negative *measure ν.*

We remark that the converse is easily established: if ν is a non-negative finite Radon measure on $\hat{G}$, then $\mathcal{F}^* f$ is positive semi-definite.

Proof Let f be positive semi-definite. It is easy to see that $f(x) = \overline{f(-x)}$ and $|f(x)| \leq f(0)$ for all x; we can then normalise $f(0) = 1$.

It is not hard to use the hypothesis and limiting arguments to show that $\langle f*g, g\rangle \geq 0$ for all simple functions g of finite measure support, and thence also for all $g \in L^1(G)$. This implies that the inner product $\langle g, h\rangle_f := \langle f * g, h\rangle$ is positive semi-definite, and thus by the Cauchy-Schwarz inequality

$$|\langle g, h\rangle_f| \leq |\langle g, g\rangle_f|^{1/2}|\langle h, h\rangle_f|^{1/2}.$$

Letting h be an approximation to the identity, and taking limits, using the fact that f is continuous and equals 1 near the origin, one soon concludes that

$$|f * g(0)| \leq |f * g * \tilde{g}(0)|^{1/2}$$

where $\tilde{g}(x) := \overline{g(-x)}$. We can iterate this repeatedly and obtain

$$\begin{aligned} |f * g(0)| &\leq |(T_g T_g^*)^{2^n} f(0)|^{1/2^{n+1}} \\ &\leq \|(T_g T_g^*)^{2^n}\|_{L^\infty(G)\to L^\infty(G)}^{1/2^{n+1}}, \\ &= \|(T_g T_g^*)^{2^n}\|_{L^1(G)\to L^1(G)}^{1/2^{n+1}}, \end{aligned}$$

where T_g is the operation of convolution by g. Taking limits as $n \to \infty$ we can thus bound $|f * g(0)|$ by the spectral radius $\rho(T_g)$. But Gelfand theory tells us that

$$\rho(T_g) \leq \|\hat{g}\|_{L^\infty(\hat{G})}.$$

(Quick sketch of proof: observe that $g-z$ is invertible (hence holomorphic) in $L^1(G)$ if $|z| > \|\hat{g}\|_{L^\infty(\hat{G})}$ (otherwise the range of $g - z$ is a proper ideal, hence contained in a maximal ideal, contradiction) and Cauchy's integral formula outside of a circle then quickly shows that

$$\|(ff^*)^n\|_{L^1(G)}/\|\hat{g}\|_{L^\infty(\hat{G})}^{2n} \to 0 \text{ as } n \to \infty$$

and the claim follows. The reverse inequality is also true.)

To summarise so far, we have shown that

$$|f * g(0)| \leq \|\hat{g}\|_{L^\infty(\hat{G})}$$

for all $g \in L^1(G)$. Thus the map $\hat{g} \mapsto f * g(0)$ is a bounded linear functional with norm at most 1 on $A(\hat{G})$, and hence extends by density to $C(\hat{G})$. Applying the Riesz representation theorem, we obtain a finite Radon measure ν on $\hat{G}$ with total mass one

$$f * g(0) = \int_{\hat{G}} \hat{g}\ d\nu$$

for all $g \in L^1(G)$. The right-hand side is also equal to $\langle g, \mathcal{F}^*\overline{\nu}\rangle$, and so by duality we see that f is (up to some harmless conjugations) the inverse Fourier transform of a measure of total mass 1. Since $f(0) = 1$, we also see that this measure has to be non-negative as required. ∎

One easily verifies that for any $f \in L^2(G)$, the function $f * \tilde{f}$ is positive definite, indeed $\langle f * \tilde{f} * \mu, \mu\rangle = \|f * \mu\|_{L^2(G)}^2 \geq 0$. By Bochner's theorem we conclude that $f * \tilde{f} \in B(G)$. From Young's inequality we thus also see that if $f \in L^1(G) \cap L^2(G)$

then $f * \tilde{f} \in L^1(G) \cap B(G)$. From this and translation invariance, we obtain a large class of functions in $L^1(G) \cap B(G)$. The significance of this is that we can establish a preliminary inversion formula in this class:

Proposition 2.2 (Preliminary inversion theorem)**.** *There exists a Haar measure $d\xi$ on $\hat{\Gamma}$ such that whenever $f \in L^1(G) \cap B(G)$, then $\hat{f} \in L^1(\hat{G})$ and*

$$f(x) = \int_{\hat{G}} \hat{f}(\xi) e(\xi \cdot x)\ d\xi.$$

Proof Since $f \in B(G)$, there exists $\nu_f \in M(\hat{G})$ such that $f = \mathcal{F}^* \nu_f$. Observe that if $f, g \in L^1(G) \cap B(G)$ and $h \in L^1(G)$, then

$$\int_{\hat{G}} \hat{f}\hat{h}\ d\nu_g = \int_{\hat{G}} \widehat{f * h}\ d\nu_g = f * h * \mathcal{F}\nu_g(0) = f * g * h(0).$$

By symmetry of convolution we thus have

$$\int_{\hat{G}} \hat{f}\hat{h}\ d\nu_g = \int_{\hat{G}} \hat{g}\hat{h}\ d\nu_f$$

for all $\hat{h} \in A(\hat{G})$. Since $A(\hat{G})$ is dense in $C(\hat{G})$, we conclude from the Riesz representation theorem that

$$\hat{f} d\nu_g = \hat{g} d\nu_f.$$

To put this another way, the quantity $d\nu_f / \hat{f}$ is independent of f on any region on which the quotient is well-defined (i.e. $\hat{f}$ is non-zero). Given any $\xi \in \hat{G}$, one can engineer an f in $L^1(G) \cap B(G)$ whose Fourier transform does not vanish near ξ by taking an expression of the form $f = h * \tilde{h}$ for some approximation to the identity h, by the previous discussion; we can then glue all the $d\nu_f / \hat{f}$ measures together to create a Radon measure $d\xi$ such that $d\nu_f = \hat{f}\ d\xi$ for all $f \in L^1(G) \cap B(G)$. By translating ν_f (which modulates f, which translates $\hat{f}$) we see that $d\xi$ is translation-invariant and is thus a Haar measure. Since $f = \mathcal{F}^* \nu_f$, the claim now follows. ■

One consequence of the inversion theorem is that

$$f(0) = \int_{\hat{G}} \hat{f}(\xi)\ d\xi$$

whenever $f \in L^1(G) \cap B(G)$, which implies also that

$$f * g(0) = \int_{\hat{G}} \hat{f}(\xi) \hat{g}(\xi)\ d\xi$$

for $f, g \in L^1(G) \cap B(G)$. This easily implies the Plancherel and Parseval identities in this class. Using approximations to the identity, one can show that $L^1(G) \cap B(G)$ is dense in $L^2(G)$, and so we obtain Plancherel's theorem, namely that there is a unique unitary extension of the Fourier transform to $L^2(G)$.

Given that Haar measure is defined up to a multiplicative constant, the only remaining question is to determine what this constant is. But by testing against a single explicit non-trivial function, one can recover this constant. Some standard examples:

Example 2.3. *Let $G = \mathbf{Z}$. Then it is easy to see that the only additive characters $x \mapsto \xi \cdot x$ from $\mathbf{Z}$ to $\mathbf{R}/\mathbf{Z}$ can be parameterised by a frequency $\xi \in \mathbf{R}/\mathbf{Z}$, given by the formula $\xi \cdot x = x\xi$. Thus $\hat{G} \equiv \mathbf{R}/\mathbf{Z}$. Testing against $f(x) := 1_{x=0}$ (so $\hat{f}(\xi) = 1$ we see that the Fourier measure $d\xi$ on $\hat{G}$ is just Lebesgue measure.*

Example 2.4. *Let $G = \mathbf{Z}/N\mathbf{Z}$. Then the additive characters $x \mapsto \xi \cdot x$ from $\mathbf{Z}/N\mathbf{Z}$ to $\mathbf{R}/\mathbf{Z}$ lift to one from $\mathbf{Z}$ to $\mathbf{R}/\mathbf{Z}$ so are parameterised by $\xi \in \mathbf{R}/\mathbf{Z}$ as before, but also we must have $N\xi = 0$ in order to be able to descend back to $\mathbf{Z}/N\mathbf{Z}$. Thus $\hat{G}$ is equivalent to the N^{th} roots of $\mathbf{R}/\mathbf{Z}$, which can also be identified with $\mathbf{Z}/N\mathbf{Z}$ (if we change the pairing $(x, \xi) \mapsto \xi \cdot x$ to $\xi \cdot x = \xi x/N$). Thus $\hat{G} \equiv \mathbf{Z}/N\mathbf{Z}$. Testing against $f(x) := 1_{x=0}$ (so $\hat{f}(\xi) = 1$ we see that the Fourier measure $d\xi$ on $\hat{G}$ is just counting measure.*

Example 2.5. *Let $G = \mathbf{R}$. Then it is easy to see that the only additive characters $x \mapsto \xi \cdot x$ from $\mathbf{R}$ to $\mathbf{R}/\mathbf{Z}$ can be parameterised by a frequency $\xi \in \mathbf{R}$, and given by the formula $\xi \cdot x = \xi x \mod 1$; this can be seen by first using the continuity to observe that for x close enough to zero, $\xi \cdot x$ lies in (say) a 0.1 neighbourhood of the origin in $\mathbf{R}/\mathbf{Z}$, and then one can easily show using the homomorphism property and continuity that $\xi \cdot x$ is linear near the origin, then one can extend to all of $\mathbf{R}$ using the homomorphism property again. Thus $\hat{G} \equiv \mathbf{R}$. Testing against the function $f(x) := e^{-\pi x^2}$ (say) one sees that the required Fourier measure $d\xi$ on $\hat{G}$ is just Lebesgue measure.*

Example 2.6. *Let $G = \mathbf{R}/\mathbf{Z}$. The additive characters $x \mapsto \xi \cdot x$ from $\mathbf{R}/\mathbf{Z}$ to $\mathbf{R}/\mathbf{Z}$ descend from characters from $\mathbf{R}$ to $\mathbf{R}/\mathbf{Z}$, but in order for the descent to work properly the frequency ξ must be an integer. Thus $\hat{G} = \mathbf{Z}$. Testing against $f(x) := 1$ (so $\hat{f}(\xi) = 1_{\xi=0}$) we see that $d\xi$ is just counting measure.*

Observe that each $x \in G$ induces an additive character on $\hat{G}$ by the map $x : \xi \mapsto \xi \cdot x$. This allows one to canonically embed G in $\hat{\hat{G}}$; this map is easily seen to be an injective homomorphism and homeomorphism. We have two Plancherel theorems on $\hat{G}$, one with frequencies in $\hat{\hat{G}}$ and one with frequencies in G. The only way these can be compatible is if G has full measure in $\hat{\hat{G}}$ and the Fourier measure of $\hat{\hat{G}}$ is dx. In fact one can show that G is equal to all of $\hat{\hat{G}}$, by observing that (by the full measure property) G is dense, while also being locally compact. This equivalence $G \equiv \hat{\hat{G}}$ is sometimes referred to as *Pontryagin duality.*

3. The Walsh ring

For most applications in real-variable harmonic analysis, the general theory of Fourier analysis on LCA groups can be specialised to the classical cases of the real line, torus, integers, cyclic groups, and products thereof. There are however a number of other interesting groups (or rings) on which this theory is useful. One is Fourier analysis on the ring $A_{\mathbf{Q}}$ of adeles, which is of significant importance in algebraic number theory but which will not be discussed here. Another is to *multiplicative* abelian groups such as $\mathbf{R}^+$ (which leads to the Mellin transform) or

$(\mathbf{Z}/N\mathbf{Z})^*$ (which leads to the theory of characters in number theory). The other is the *Walsh ring* or *Cantor-Walsh ring*[1] W_p, which is a dyadic variant of the real number ring $\mathbf{R}$ and which serves as a useful model for harmonic analysis on the real line (among other things, it manages to avoid the perennial presence of Schwartz tails which plague the analysis on $\mathbf{R}$.

To define the Walsh ring, we need an prime p; in many applications we set $p = 2$ (giving the *dyadic* Walsh ring) or $p = 3$ (giving the *ternary* Walsh ring). We let F_p be the field of p elements (identified with $\mathbf{Z}/p\mathbf{Z}$ and with $\{0, \ldots, p-1\}$ in the usual manner), and let $F_p(t)$ be the ring of Laurent polynomials in one indeterminate t, thus a typical non-zero element of $F_p(t)$ is of the form

$$x = \sum_{n=A}^{B} a_n t^n$$

for some finite integers $A \leq B$ and coefficients $a_n \in F_p$ with $a_B \neq 0$. We can also represent this element more suggestively in "base p notation" as

$$x = a_B a_{B-1} \ldots a_0 . a_{-1} \ldots a_A.$$

Indeed, if we identify an element $\sum_{n=A}^{B} a_n t^n$ of $F_p(t)$ with the base p terminating decimal $\sum_{n=A}^{B} a_n p^{-n}$ (recall that we identify F_p with $\{0, \ldots, p-1\}$), we obtain a ring which resembles the ring of base p terminating decimals in the reals, but where addition and multiplication are performed without the "carry" operation (or in more modern jargon, the cocycle is trivial).

Just as the reals $\mathbf{R}$ are the metric completion of the rationals $\mathbf{Q}$, the Walsh ring W_p are the metric completion of the Laurent ring $F_p(t)$, using the metric $d(x, y) := \|x - y\|_p := p^{-B}$, where B is the degree of $x - y$ (i.e. the largest power of t which appears; we use the convention that the degree of 0 is $-\infty$). One easily observes that the ring operations are locally uniformly continuous on $F_p(t)$ and thus extend in a unique continuous manner to W_p. The completion can also be identified with the one-sided infinite series

$$x = \sum_{n=-\infty}^{B} a_n t^n$$

over F_p; one can also *almost* identify this ring with the half-line $\mathbf{R}^+$ by identifying x with $\sum_{n=-\infty}^{B} a_n p^{-n}$, but of course the infamous $0.999\ldots = 1$ problem shows that there is a problem with doing so at every element of $F_p(t)$. Strictly speaking, one thus has to adjoin an additional element infinitesimally to the left of every non-zero terminating base p decimal in $\mathbf{R}^+$ before one has a fully accurate representation of the Walsh ring W_p, but we will gloss over this annoying technicality; in particular, we shall abuse notation and describe elements of W_p using the formula $\sum_{n=-\infty}^{B} a_n p^{-n}$ rather than the more accurate $\sum_{n=-\infty}^{B} a_n t^n$. (An alternate approach is to identify x with $\sum_{n=-\infty}^{B} a_n q^{-n}$ for some $q > p$, which identifies W_p bijectively with a Cantor set, but then the analogy between W_p and $\mathbf{R}^+$ is diminished.) We can pull back Lebesgue measure on $\mathbf{R}^+$ to create a measure dx on W_p, which is easily seen to be a Haar measure.

[1]This notation is not standard.

The metric d extends of course to the completion W_p; it is somewhat related to the usual Euclidean metric on $\mathbf{R}^+$ by observing that $d(x,y) \geq \frac{1}{2}|x-y|$, but it is possible for $d(x,y)$ to much larger than $|x-y|$ (this occurs for instance if x and y are close to but on opposite sides of a terminating decimal). The metric d is in fact better than a metric, it is an *ultrametric* (or *non-archimedean metric*), thus

$$d(x,y) \leq \max(d(x,z), d(z,y)).$$

This implies that the metric balls are nested. Indeed, the metric balls here are nothing more than the p-dyadic intervals $[\frac{j}{p^n}, \frac{j+1}{p^n})$ for $n \in \mathbf{Z}$, $j \geq 0$ (note that we include the element $\frac{j}{p^n} + \sum_{m=-\infty}^{n-1} \frac{p-1}{p^n}$, which is infinitesimally to the left of $\frac{j+1}{p^n}$, in this ball).

The space W_p is homeomorphic to an unbounded Cantor set, as the identification $x \leftrightarrow \sum_{n=-\infty}^{B} a_n q^{-n}$ shows; in particular, W_p is locally compact as well as abelian. Now let's compute the Pontryagin dual of W_p. Let $\xi : x \mapsto \xi \cdot x$ be an additive character. Observe that for each integer n, $\xi \cdot p^n$ must be a p^{th} root of unity in $\mathbf{R}/\mathbf{Z}$, since adding p^n to itself p times in W_p gives zero (note this is *not* the same as $p \times p^n$ in W_p!). Thus we can write $\xi \cdot p^n = a_{-n-1}/p$ for some $a_{-n-1} \in F_p$. By continuity, we see that $\xi \cdot p^n \to 0$ as $n \to +\infty$, and hence $a_n = 0$ for all sufficiently large n, say $n > N$. If we then write

$$\tilde{\xi} := \sum_{n=-\infty}^{N} a_n p^{-n} \in W_p$$

and define the dot product

$$(\sum_{n=-\infty}^{N} a_n p^{-n}, \sum_{m=-\infty}^{M} b_m p^{-m}) := \sum_{n=-M}^{N} a_n b_{-1-m}/p \in \mathbf{R}/\mathbf{Z}$$

then we see that

$$\xi \cdot x = \tilde{\xi} \cdot x$$

for all $x \in F_p(t)$, and hence by continuity for all $x \in W_p$ also. Thus $\hat{W}_p$ is identifiable with W_p, using the dot product $\cdot : W_p \times W_p \to \mathbf{R}/\mathbf{Z}$ described above.

If one lets $f := 1_{[0,1)}$ in W_p, one easily checks that $\hat{f} = f$, which by Plancherel's theorem tells us that the Fourier measure $d\xi$ is the same as the original measure dx. Note that in sharp contrast to the Fourier theory on $\mathbf{R}$, in W_p it is perfectly possible to have a function of compact support whose Fourier transform is also of compact support. This leads to a very pleasant harmonic analysis in W_p, which can be regarded as a simplified model of harmonic analysis in $\mathbf{R}$: many of the foundational theorems of the subject are somewhat simpler in this setting while still capturing the main essence of the real-variable theory (a good example arises by comparing the ordinary Hardy-Littlewood maximal function with its Walsh variant, namely the dyadic (or p-adic) maximal function). We will however not pursue this theme in detail here.

4. Exercises

- Q1. (Poisson summation formula for finite abelian groups) Let G be a finite abelian group, and let H be a subgroup of G. Define the *orthogonal complement* $H^\perp \subset \hat G$ of H as $H^\perp := \{\xi \in \hat G : \xi \cdot x = 0 \text{ for all } x \in H\}$. Show that $H^\perp$ is a subgroup of $\hat G$, and that $\hat G / H^\perp$ is canonically identifiable with $\hat H$. In particular we see that
$$\#G = (\#H)(\#H^\perp).$$
Show that
$$\hat 1_H(\xi) = \frac{\#H}{\#G} 1_{H^\perp}(\xi)$$
and conclude the *Poisson summation formula*
$$\frac{1}{\#H} \sum_{x \in H} f(x) = \sum_{\xi \in H^\perp} \hat f(\xi)$$
for all $f \in L^2(G)$. More generally, show that
$$\frac{\#G}{\#H} \widehat{f 1_H}(\xi) = \sum_{\eta \in H^\perp} \hat f(\xi + \eta)$$
and
$$\frac{\#G}{\#H} \widehat{f * 1_H}(\xi) = 1_{H^\perp}(\xi) \hat f(\xi)$$
for all $f \in L^2(G)$ and $\xi \in \hat G$ (thus "the Fourier transform of restriction is projection", and vice versa).
- Q2. (Uncertainty principle for finite abelian groups) Let G be a finite abelian group. Show that for any non-zero $f \in L^2(G)$ we have
$$\#(\mathrm{supp}(f)) \#(\mathrm{supp}(\hat f)) \geq \#G.$$
(Hint: use the Plancherel and Hölder inequalities, together with the trivial bound $\|\hat f\|_{l^\infty(\hat G)} \leq \|f\|_{L^1(G)}$.
- Q3. (Entropy uncertainty principle for finite abelian groups) Let G be a finite abelian group. Establish the Hausdorff-Young inequality
$$\|\hat f\|_{l^{p'}(\hat G)} \leq \|f\|_{L^p(G)}$$
for all $1 \leq p \leq 2$ and $f \in L^p(G)$; differentiate this at $p = 2$ to conclude the entropy uncertainty principle
$$\int_{x \in G} |f(x)|^2 \log \frac{1}{|f(x)|^2} + \sum_{\xi \in \hat G} |\hat f(\xi)|^2 \log \frac{1}{|\hat f(\xi)|^2} \geq 0$$
for all $f \in L^2(G)$ with $\|f\|_{L^2(G)} = 1$, using the convention $0 \log \frac{1}{0} = 0$. Using Jensen's inequality, give a second proof of the uncertainty principle from Q2.
- Q4. (Adjoints in finite abelian groups) Let G be a finite abelian group, and let $\phi : G \to G$ be an automorphism. Define the adjoint operator $\phi^* : \hat G \to \hat G$ by
$$\phi^*(\xi) \cdot x := \xi \cdot \phi(x).$$

Show that ϕ^* is also an automorphism, and that for any $f \in L^2(G)$ we have

$$\widehat{f \circ \phi} = \hat{f} \circ (\phi^*)^{-1}.$$

- Q5. Show that the Pontryagin dual of a compact abelian group is a discrete abelian group, and vice versa.
- Q6. Let W_p be a Walsh group. Define the operators $\mathrm{Dil}^q_{p^n}$ for $1 \leq q \leq \infty$ and $n \in \mathbf{Z}$ and $f \in L^q(W_p)$ by

$$\mathrm{Dil}^q_{p^n} f(x) := p^{n/q} f(p^{-n}x).$$

(Note that Walsh multiplication of x by p^{-n} corresponds with classical multiplication.) Show that $\mathrm{Dil}^q_{p^n}$ is an isometry on $L^q(W_p)$, and that

$$\widehat{\mathrm{Dil}^q_{p^n} f} = \mathrm{Dil}^{q'}_{p^{-n}} \hat{f}.$$

- Q7 (Walsh uncertainty principle). Let W_p be a Walsh group, $f \in L^2(W_p)$, and $n \in \mathbf{Z}$. Show that $\hat{f}$ is supported in the interval $[0, p^n)$ if and only if f is constant on every ball of radius p^{-n}. We remark that the Walsh-Fourier projection to the interval $[0, p^n)$ is highly analogous to the Littlewood-Paley projection $\psi_{\leq n}$, and enjoys essentially the same theory (in fact, the Walsh theory is cleaner and simpler, and is also closely related to martingale theory).
- Q8 (Tiles). Define a *tile* to be a set of the form $P = I \times w \subset W_p \times W_p$, where I, w are balls with $|I||w| = 1$. Show that the space $V_P := \{f \in L^2(I) : \hat{f} \in L^2(w)\}$ is one-dimensional, and that V_P and V_Q are orthogonal whenever P, Q are disjoint tiles. If a set $\Omega \subset W_p \times W_p$ can be partitioned $\Omega = P_1 \cup \ldots \cup P_n$ into finitely many disjoint tiles, show that the vector space $V_\Omega := V_{P_1} + \ldots + V_{P_n}$ is independent of the choice of partition.

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: tao@math.ucla.edu