CS 880: Advanced Complexity Theory 1/28/2008

Lecture 3: Harmonic Analysis Over Finite Groups

Instructor: Dieter van Melkebeek Scribe: Adeel Pervez

In the last lecture we reviewed classical harmonic analysis over the reals and its generalization to locally compact abelian groups. We defined the notion of a character and exhibited some simple properties. Today we will first cover the basics of harmonic analysis over finite abelian groups. We then restrict ourselves to the special case of the Boolean cube, which will be the main arena for this course.

We begin by stating the key property of characters we will use. For the remainder of this lecture, let G denote a finite group under addition.

1 Orthonormality of Characters

To define orthonormality we first need to introduce the inner product of two functions over G.

Definition 1. *Let $f : G \to \mathbb{C}$ and $g : G \to \mathbb{C}$ be two functions. We define the inner product $\langle f, g\rangle$ of f and g as:*

$$\langle f, g\rangle = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}$$

This can also be thought of as an expected value over x where x is uniformly chosen from the group G:

$$\langle f, g\rangle = \mathrm{E}_x[f(x)\overline{g(x)}]$$

Proposition 1. *Characters are orthonormal with respect to $\langle,\rangle$. That is, given two characters χ and ψ over G, $\langle \chi, \psi\rangle = 0$ if $\chi \not\equiv \psi$ and $\langle \chi, \psi\rangle = 1$ if $\chi \equiv \psi$.*

Proof. Let χ and ψ be two distinct characters. Then $\phi(x) = \chi(x)\overline{\psi(x)}$ is a non-trivial character; i.e., ϕ is a character and $\phi \not\equiv 1$. Then by definition of inner product

$$\langle \chi, \psi\rangle = \mathrm{E}_x[\chi(x)\overline{\psi(x)}] = \mathrm{E}_x[\phi(x)].$$

Let $\phi(x) = \chi(x)\overline{\psi(x)}$. Then for each $a \in G$

$$\phi(a)\mathrm{E}_x[\phi(x)] = \frac{1}{|G|}\sum_{x\in \text{G}} \phi(a)\phi(x) = \frac{1}{|G|}\sum_{x\in \text{G}} \phi(a+x) = \frac{1}{|G|}\sum_{x\in \text{G}} \phi(x) = \mathrm{E}_x[\phi(x)],$$

so

$$(\phi(a) - 1)\mathrm{E}_x[\phi(x)] = 0.$$

Since ϕ is non-trivial

$$\mathrm{E}_x[\phi(x)] = 0.$$

This proves orthogonality. That $\langle \chi, \chi\rangle = 1$ follows from the fact that the complex conjugate of a character is equal to its inverse. □

As mentioned in the previous lecture, the set of characters is a group $\hat{G}$ under pointwise multiplication. Proposition 1 implies that the size of $\hat{G}$ is no greater than the size of G, since the dimension of the space of functions $f : G \to \mathbb{C}$ equals $|G|$.

The properties of characters mentioned previously hold for arbitrary finite groups. We now restrict ourselves to the case of finite abelian groups, where we will see that $\hat{G}$ is isomorphic to G, so in particular $|\hat{G}| = |G|$.

2 Harmonic Analysis Over Finite Abelian Groups

For finite abelian groups we state the following proposition without proof.

Proposition 2. *Let G be a finite abelian group. Then there exist $k, n_1, \ldots, n_k \in \mathbb{N}$ such that*

$$G \cong \mathbb{Z}_{n_1} \otimes \mathbb{Z}_{n_2} \cdots \otimes \mathbb{Z}_{n_k} \tag{1}$$

where $\otimes$ denotes a direct product of groups.

The above statement tells us that every finite abelian group can be viewed as a direct product of cyclic groups. For cyclic groups we can easily determine all characters. The proposition above then allows us to determine all characters for any finite abelian group.

Characters over $\mathbb{Z}_n$ can intuitively be viewed as rotations over a regular n-gon. Each character of $\mathbb{Z}_n$ is completely determined by the value $\chi(1)$. There are exactly n distinct characters corresponding to the n distinct values of $\chi(1)$ satisfying $(\chi(1))^n) = 1$, or equivalently unique rotations of the n-gon. We let $\omega = e^{\frac{2\pi i}{n}}$ correspond to the smallest non-trivial rotation of the n-gon. We can define the character χ_a for each $a \in \mathbb{N}$ in terms of ω as follows

$$\chi_a(x) = \omega^{ax}$$

χ_a thus corresponds to a rotation over a multiple of the smallest angle of rotation of the n-gon.

χ_a is easily seen to be a character. We also have the following simple proposition.

Proposition 3. *$\chi_a \equiv \chi_b$ if and only if $a = b$ (mod n).*

This implies that we have n characters for $\mathbb{Z}_n$, which is the total number of characters since there cannot be more than $|\mathbb{Z}_n| = n$. Using (1) a general character for the finite abelian group G can now be written as follows:

$$\chi_a(x) = \prod_{j=1}^{k} \omega_j^{a_j x_j}, \tag{2}$$

where $\omega_j = e^{\frac{2\pi i}{n_j}}$ and $a_j \in \mathbb{Z}_{n_j}$ for each $j = 1, \ldots, k$.

This relation gives us a total of $\prod_{j=1}^{k} n_j = |G|$ distinct characters. The relation, therefore, accounts for all characters of G. Moreover $G \cong \hat{G}$ under the isomorphism $a \to \chi_a$, since $\chi_{a+b} = \chi_a \cdot \chi_b$.

The group $\hat{G} \cong G$ is an orthonormal basis for the functions $f : G \to \mathbb{C}$. This implies that f can be written as

$$f(x) = \sum_{a \in G} \hat{f}(a)\chi_a(x). \tag{3}$$

The function $\hat{f} : G \to \mathbb{C}$ is the *Fourier transform* of the function f over the group G. Due to orthonormality of characters the Fourier coefficients $\hat{f}(a)$ for $a \in G$ can be determined by taking the inner product with χ_a on both sides of (3), resulting in

$$\hat{f}(a) = \langle f, \chi_a \rangle = \mathrm{E}_x[f(x)\overline{\chi_a(x)}] = \frac{1}{|G|} \sum_{x \in G} f(x)\overline{\chi_a(x)}.$$

2.1 Properties of the Fourier Transform

Now we mention some properties of the Fourier transform that carry over from the transform over the reals to the transform over finite abelian groups.

Linearity

For functions f and g and a scalar α

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

and

$$\widehat{\alpha f} = \alpha \hat{f}.$$

These properties follow from (2) and the linearity of the inner product.

(Almost) Involution

Since $\overline{\chi_a(x)} = \chi_a(-x) = \chi_{-x}(a)$,

$$\hat{f}(a) = \frac{1}{|G|} \sum_{x \in G} f(x)\overline{\chi_a(x)} = \frac{1}{|G|} \sum_{x \in G} f(-x)\chi_x(a).$$

So,

$$\hat{\hat{f}}(x) = \frac{1}{|G|} f(-x).$$

Thus applying the Fourier transform twice results in the original function upto constant factors and replacing x by $-x$ (hence the ‘almost’ qualifier). If we wanted to, we could tweak the $1/|G|$ factor in the definition of the Fourier transform so that it does not appear in the last equation, but the sign change would still remain.

(Almost) Isometry

Fourier transforms preserve angles and distances upto constant factors. Let f and g be functions then

$$\langle f, g \rangle = \langle \sum_{a \in G} \hat{f}(a)\chi_a, \sum_{b \in G} \hat{g}(b)\chi_b \rangle.$$

By linearity of $\langle , \rangle$

$$\langle f, g \rangle = \sum_{a, b \in G} \hat{f}(a)\overline{\hat{g}(b)} \langle \chi_a, \chi_b \rangle.$$

By orthonormality of characters

$$\langle f, g \rangle = \sum_{a \in G} \hat{f}(a)\overline{\hat{g}(a)},$$

which is *Plancherel's equality*. A special case is when f and g are the same function. In this case

$$\langle f, f\rangle = \sum_{a\in G} |\hat{f}(a)|^2 = |G| \cdot \langle \hat{f}, \hat{f}\rangle = |G| \cdot \langle \hat{f}, \hat{f}\rangle,$$

which is *Parseval's equality*.

The Fourier transform is thus isometric upto a $\frac{1}{|G|}$ factor. By rescaling the inner product one can redistribute the factor and turn the Fourier transform into a perfect isometry, but we won't do that.

Convolution

Convolution of functions translates to multiplication in the Fourier domain. This is seen as follows:

$$\begin{aligned}(f * g)(x) &= \mathrm{E}_y[f(y)g(x-y)]\\ &= \mathrm{E}_y[(\sum_{a\in G} \hat{f}(a)\chi_a(y))(\sum_{b\in G} \hat{g}(b)\chi_b(x-y)]\\ &= \mathrm{E}_y[(\sum_{a\in G} \hat{f}(a)\chi_a(y))(\sum_{b\in G} \hat{g}(b)\chi_b(x)\overline{\chi_b(y)})]\\ &= \sum_{a,\, b\in G} \hat{f}(a)\hat{g}(b)\chi_b(x)E_y[\chi_a(y)\overline{\chi_b(y)}]\\ &= \sum_{a\in G} \hat{f}(a)\hat{g}(a)\chi_a(x)\end{aligned}$$

Here, the first line is the definition of convolution; the functions have been expanded in their character bases in the second line; the fourth line follows from linearity of expectation; and the last equation follows from the orthonormality of characters. From the last equation, the Fourier coefficients of $(f*g)(x)$ can be seen to be $\hat{f}(a)\hat{g}(a)$. Taking its Fourier transform, therefore, results in:

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

3 Harmonic Analysis Over the Boolean Cube

In this course we will be looking at functions on the Boolean cube, and mostly at Boolean ones. For this case the underlying group will be the group of n-bit strings $G = (\mathbb{Z}_2^n, +)$ and the functions will be $f : G \to \{0, 1\}$. For the Boolean cube our previously defined $n_j = 2$ so $\omega_j = -1$ for each $j = 1, 2, \ldots, n$, and the characters can be written as

$$\chi_a(x) = (-1)^{a\cdot x},$$

where $a \cdot x$ denotes the dot product $\sum_{j=1}^n a_j \cdot x_j$ (either over $\mathbb{Z}$ or over $\mathbb{Z}_2$).

The characters for the Boolean case can be viewed as parity functions over subsets of the bit positions where -1 denotes odd parity and 1 even parity. In particular, χ_a represents parity over the subset of bit positions j for which $a_j = 1$.

3.1 $(0,1) \to (1,-1)$ Transformation

The harmonic analysis of Boolean functions becomes simpler if the consider the transformation $0 \to 1$ and $1 \to -1$. That is, instead of taking 0 and 1 to mean false and true we use 1 and -1 respectively. With this transformation our functions are now from $\{1,-1\}^n$ to $\{1,-1\}$. An immediate consequence of using this transformation on the domain is that the characters become products of variables:

$$\chi_a(x) = \prod_{j \in S} x_j$$

where $S = \{j \in [n] | a_j = 1\}$. It is convenient to represent a in the above equation by the set S just defined. The above equation can then be rewritten as

$$\chi_S(x) = \prod_{j \in S} x_j,$$

and the Fourier expansion of a function $f : G \to \mathbb{C}$ becomes

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x).$$

A consequence of this transformation on the range is that Fourier coefficients can be interpreted as measures of correlation between the given function and a parity function using the correspondence between characters and parity given above. This can be seen as follows

$$\begin{aligned} \hat{f}(S) &= \mathrm{E}_x[f(x)\chi_S(x)] \\ &= \Pr_x[f(x) = \chi_S(x)] - \Pr_x[f(x) \neq \chi_S(x)] \\ &= 2\Pr_x[f(x) = \chi_S(x)] - 1 \\ &= 1 - 2\Pr_x[f(x) \neq \chi_S(x)]. \end{aligned} \tag{4}$$

Another consequence of this transformation on the range follows from Parseval's equality:

$$\langle f, f \rangle = \sum_{a \in G} |\hat{f}(a)|^2 = \sum_{a \in G} (\hat{f}(a))^2 = 1.$$

Therefore, Fourier coefficients squared can be interpreted as a probability distribution on $G = \{-1,1\}^n$.

3.2 Linear Functions

Definition 2. *Let f be a function from $G = (\mathbb{Z}_2^n, +) \to \{0,1\}$. Then f is linear if $f(x+y) = f(x) + f(y)$.*

Using the $0 \to 1$, $1 \to -1$ transformation described above, a function $f : (\{-1,1\}^n, \cdot) \to \{-1,1\}$, where $\cdot$ denotes component-wise multiplication, is linear if $f(x \cdot y) = f(x) \cdot f(y)$.

Notice that the second definition of a linear function is just that of a character. Thus we can measure the 'linearness' of a Boolean function by comparing it with a character. By (4), the largest Fourier coefficient corresponds to the linear function the given function is closest to. In a previous lecture we defined the distance $D(f,g)$ between two functions f and g as the relative Hamming distance between them, and the distance $D(f,\mathcal{F})$ of f to a class of functions $\mathcal{F}$ as $\min_{g\in\mathcal{F}} D(f,g)$. Then using (4) we can write the distance between a function f and the set $LINEAR$ of linear function as:

$$D(f, LINEAR) = \frac{1}{2} - \frac{1}{2}\max_{S\subseteq[n]} \hat{f}(S).$$

For affine functions we have:

$$D(f, AFFINE) = \frac{1}{2} - \frac{1}{2}\max_{S\subseteq[n]} |\hat{f}(S)|.$$

The latter follows from the former because $AFFINE = LINEAR \ \cup \ -LINEAR$.

4 Next Time

Next time we will introduce an efficient randomized test to distinguish between linear functions and functions that are far from linear, and will prove that the test succeeds with high probability.