Fourier Analysis

Topics in Finite Fields (Fall 2013)
Rutgers University
Swastik Kopparty
Last modified: Friday 18th October, 2013

1 Fourier basics

Let G be a finite abelian group.
A character of G is simply a homomorphism ψ from G to the multiplicative group of the complex numbers, $\mathbb{C}^*$: $\psi(a+b) = \psi(a)\psi(b)$, and $\psi(-a) = \frac{1}{\psi(a)}$. Since G is finite, we have that every element in the image of ψ is a root of unity, and thus $\frac{1}{\psi(a)} = \overline{\psi(a)}$.
Characters form a group under multiplication. We define the dual group of G, $\hat{G}$, to be the group of all characters of G. Let ψ_0 be the trivial character, which maps all of G to 1; this is the identity element of $\hat{G}$.

1.1 Examples

1. Let $G = \mathbb{Z}_p$. For $a \in \mathbb{Z}_p$, define $\psi_a : G \to \mathbb{C}$ by:

$$\psi_a(x) = e^{2\pi i a x/p} = \omega_p^{ax},$$

   (where ω_p is the primitive pth root of unity $e^{2\pi i/p}$). Then $\hat{G} = \{\psi_a \mid a \in \mathbb{Z}_p\}$. (Here p need not be prime).

2. Let $G = \mathbb{Z}_{2^n}$. For $a \in \mathbb{Z}_2^n$, define $\psi_a : G \to \mathbb{C}$ by:

$$\psi_a(x) = (-1)^{\langle a,x\rangle} = (-1)^{\sum_{j=1}^n a_i x_i}.$$

   Then $\hat{G} = \{\psi_a \mid a \in \mathbb{Z}_p\}$.

(In the above examples, it is clear that the given ψ_a are indeed characters. That these are *all* the characters can be verified using a minimal set of generators for G. Soon we will see a proof of the fact that in general, the size of the dual group is at most $|G|$).
For a general finite abelian group G, by the classification of finite abelian groups, we can write $G \equiv \bigoplus_{i=1}^k \mathbb{Z}_{d_i}$. This gives us, for each $a = (a_1, \ldots, a_k) \in \mathbb{Z}_{d_i}$, the character $\psi_a : G \to \mathbb{C}^*$ given by:

$$\psi_a(x_1, \ldots, x_k) = \prod_{i=1}^k \omega_{d_i}^{a_i x_i}.$$

Observe that these characters are distinct and form a group. Again, these are all the characters of G.
It so happens that in all these cases, $\hat{G}$ is isomorphic to G, but this is just a coincidence that occurs only in the finite case.

1.2 Characters form an orthonormal basis

We use $\mathbb{C}^G$ to denote the $\mathbb{C}$ vector space of all complex valued functions on G, equipped with the inner product $\langle f, g\rangle = \mathbb{E}_{x\in g}[f(x)\overline{g(x)}]$. We now show that the characters of G form an orthonormal basis for $\mathbb{C}^G$. While studying additive questions on G, it will be terribly useful to represent functions on G in this basis.

Lemma 1. *Let ψ be a character of G. Then:*

$$\mathbb{E}_{x\in G}[\psi(x)] = \begin{cases} 1 & \psi = \psi_0 \\ 0 & \psi \neq \psi_0 \end{cases}.$$

Proof. If $\psi = \psi_0$, then the lemma is obvious. Suppose $\psi \neq \psi_0$.
Let $S = \mathbb{E}_{x\in G}[\psi(x)]$. For a fixed $y \in G$, as x varies over all elements of the group, then so does xy. Then:

$$S = \mathbb{E}_{x\in G}[\psi(xy)] = \psi(y) \cdot S.$$

Take $y \in G$ to be such that $\psi(y) \neq 1$. Thus $S = 0$. □

This immediately gives us the orthonormality of characters:

Lemma 2. *Let ψ, ψ' be characters of G. Then:*

$$\langle \psi, \psi' \rangle = \mathbb{E}_{x\in G}[\psi(x)\overline{\psi'(x)}] = \begin{cases} 1 & \psi = \psi' \\ 0 & \psi \neq \psi' \end{cases}.$$

Proof. Apply the previous lemma to the character $\psi \cdot \overline{\psi'}$. □

Thus the number of characters of G is at most $|G|$. We already exhibited, for every finite abelian group G, a set of $|G|$ distinct characters of G. Thus those are all the characters, and the set of all characters forms a basis for $\mathbb{C}^G$.
Now that we have determined the dual group $\hat{G}$ of every finite abelian group, we change notation somewhat to make our future expressions cleaner. For $G \equiv \bigoplus_{i=1}^{k} \mathbb{Z}_{d_i}$, we will let $\hat{G} = \bigoplus_{i=1}^{k} \mathbb{Z}_{d_i}$, and for $a \in \hat{G}$, we usel ψ_a to denote the character of G corresponding to a as given in the previous section.

1.3 The Fourier transform

By the previous section, every function $f : G \to \mathbb{C}$ can be written as a linear combination of characters of G.

Lemma 3. *Every $f : G \to \mathbb{C}$ has the following expression in terms of the characters of G:*

$$f = \sum_{a\in\hat{G}} \hat{f}(a) \cdot \psi_a,$$

where:

$$\hat{f}(a) = \langle f, \psi_a \rangle = \mathbb{E}_{x\in G}[f(x)\overline{\psi_a(x)}].$$

The function $\hat{f} : \hat{G} \to \mathbb{C}$ is called the Fourier transform of f.
There is also an inverse Fourier transform: given any function $h : \hat{G} \to \mathbb{C}$, there is a unique function $f : G \to \mathbb{C}$ such that $\hat{f} = h$.

1.4 Parseval and Plancherel

The orthonormality characters also implies that the Fourier transform is an isometry. This is given by the next two lemmas[1].

Lemma 4. *Let $f : G \to \mathbb{C}$. Then:*

$$\mathbb{E}_{x\in G}[|f(x)|^2] = \sum_{a\in\hat{G}} |\hat{f}(a)|^2.$$

[1]Somewhere I got the impression that the first of these lemmas is called the Parseval identity, and the second is called the Plancherel identity, but just now I wasn't able to confirm that from the internet.

Lemma 5. *Let $f, g : G \to \mathbb{C}$. Then:*

$$\mathbb{E}_{x\in G}[f(x)\overline{g(x)}] = \sum_{a\in\hat{G}} \hat{f}(a)\overline{\hat{g}(a)}.$$

The second lemma implies the first. To prove the second lemma, we expand and simplify:

$$\begin{aligned}
\mathbb{E}_{x\in G}[f(x)\overline{g(x)}] &= \mathbb{E}_{x\in G}[\left(\sum_{a_1\in\hat{G}} \hat{f}(a_1)\psi_{a_1}(x)\right)\left(\sum_{a_2\in\hat{G}} \overline{\hat{g}(a_2)\psi_{a_2}(x)}\right)] \\
&= \mathbb{E}_{x\in G}[\sum_{a_1,a_2\in\hat{G}} \hat{f}(a_1)\psi_{a_1}(x)\overline{\hat{g}(a_2)\psi_{a_2}(x)}] \\
&= \sum_{a_1,a_2\in\hat{G}} \mathbb{E}_{x\in G}[\hat{f}(a_1)\psi_{a_1}(x)\overline{\hat{g}(a_2)\psi_{a_2}(x)}] \\
&= \sum_{a_1,a_2\in\hat{G}} \hat{f}(a_1)\overline{\hat{g}(a_2)}\mathbb{E}_{x\in G}[\psi_{a_1}(x)\overline{\psi_{a_2}(x)}] \\
&= \sum_{a_1,a_2\in\hat{G}} \hat{f}(a_1)\overline{\hat{g}(a_2)}\mathbf{1}_{a_1=a_2} \\
&= \sum_{a\in\hat{G}} \hat{f}(a)\overline{\hat{g}(a)},
\end{aligned}$$

as desired.

1.5 Convolution

So far we have not seen why the Fourier transform has anything to do with the group structure on G. Now we will.
Given functions $f, g : G \to \mathbb{C}$, define their convolution $f * g : G \to \mathbb{C}$ as follows:

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

Convolution shows up naturally in many additive-combinatorial problems. For example, if $f = 1_S$,

$$f * f(x) = \frac{1}{q} \cdot (\text{ the number of } (y_1, y_2) \in S^2 \text{ s.t. } y_1 + y_2 = x).$$

Convolution is an associative operation (check this!), and thus we may write expressions such as $f * g * h$. We denote $f * f * \cdots * f$ (k times) by $f^{(*k)}$.
Convolution interacts nicely with the Fourier transform.

Lemma 6. *If $f, g : G \to \mathbb{C}$, and $h = f * g$.*
Then for every $a \in \hat{G}$,

$$\hat{h}(a) = \hat{f}(a) \cdot \hat{g}(a).$$

We also have a relation in the other direction.

Lemma 7. *If $f, g : G \to \mathbb{C}$, and $h = f \cdot g$.*
Then for every $a \in \hat{G}$,

$$\hat{h}(a) = \sum_{b\in\hat{G}} \hat{f}(b)\hat{g}(a-b).$$

1.6 Big and small

There is a lot of useful information in the magnitudes of Fourier coefficients. To get acclimatized, we should understand how big Fourier coefficents typically are, and how big can they possibly be.
There are two special scenarios we will be especially interested in.

1. S is a set, and $f = 1_S$. By Parseval, $\sum_{a\in\hat{G}} |\hat{f}(a)|^2 = \frac{|S|}{|G|}$, and thus at least one $|\hat{f}(a)|$ is at least $\frac{\sqrt{|S|}}{|G|}$. Furthermore, for all $a \in \hat{G}$, we have $|\hat{f}(a)| \leq \frac{|S|}{|G|}$.

2. $f : G \to \mathbb{C}$ is such that $|f(x)| = 1$ for all $x \in G$. By Parseval, $\sum_{a\in\hat{G}} |\hat{f}(a)|^2 = 1$, and thus at least one $|\hat{f}(a)| \geq \frac{1}{|\sqrt{G}|}$. Furthermore, for all $a \in \hat{G}$, we have $|\hat{f}(a)| \leq 1$.

What does it mean for a Fourier coefficient to be large? Take $G = \mathbb{Z}_2^n$, let $a \in \hat{G}$ be nonzero, let $S \subseteq G$, and let $f = 1_S$. For $|\hat{f}(a)|$ to be near its largest possible value $\frac{|S|}{|G|}$, we must have the following extreme scenario: Let $H_0 = \{x \in G \mid \langle x, a\rangle = 0\}$ and $H_1 = \{x \in G \mid \langle x, a\rangle = 1\}$ be the two hyperplanes defined by a that partition G. Then the manner in which S gets partitioned by H_0 and H_1 must be highly skewed. Conversely, the Fourier coefficent is very small ($o(\frac{|S|}{|G|})$) if S is partitioned almost equally into two parts by H_0 and H_1. Thus largeness of Fourier coefficients detects "nonuniform" distribution of the set S (with respect to the hyperplanes of the group G).
For $G = \mathbb{Z}_p$, the relationship is slightly less precise. For $S \subseteq G$ to have a large Fourier coefficient $a \in \hat{G}$, it must be the case that most of the elements of $\{\omega_p^{ax} \mid x \in S\}$ are "aligned" on the unit circle. Thus the scaled set aS should be strongly concentrated in an interval, or equivalently, S should be strongly concentrated in an arithmetic progression. Again, this is a form of nonuniform distribution.
What do we expect for typical f?

1. Let f be a uniformly random function from G to the unit circle of $\mathbb{C}$. Then for any fixed nonzero $a \in \hat{G}$, we have by the Chernoff/Hoeffding/Bernstein bounds:

$$\Pr[|\hat{f}(a)| > \frac{t}{|G|}] < e^{-t^2/2|G|}.$$

   Taking union bound over all nonzero $a \in \hat{G}$, we have that:

$$\Pr[\ \exists \text{ nonzero } a \in \hat{G} \text{ s.t. } |\hat{f}(a)| > \frac{t}{|G|}] < |G| \cdot e^{-t^2/2|G|}.$$

   Taking $t \gg \sqrt{|G| \cdot \log |G|}$, we have that with high probability, $|\hat{f}(a)| \leq O(\sqrt{\frac{\log |G|}{|G|}})$ for all $a \in G$.

2. Take S to be a random multiset of size k, selected by choosing $x_1, \ldots, x_k$ uniformly at random from G. Then for any fixed nonzero $a \in \hat{G}$, we have by the Chernoff/Hoeffding/Bernstein bounds:

$$\Pr[|\hat{f}(a)| > \frac{t}{|G|}] < e^{-t^2/2k}.$$

   Taking union bound over all nonzero $a \in \hat{G}$, we have that:

$$\Pr[\ \exists \text{ nonzero } a \in \hat{G} \text{ s.t. } |\hat{f}(a)| > \frac{t}{|G|}] < |G| \cdot e^{-t^2/2k}.$$

   Taking $t \gg \sqrt{k \cdot \log |G|}$, we have that with high probability, $|\hat{f}(a)| \leq O(\sqrt{k \log |G|}/|G|)$ for all $a \in G$.

Thus random functions/sets have really small Fourier coefficients. Thus we will informally say that having really small Fourier coefficients is a pseudorandom property.

1.6.1 A quick application

Let $S \subseteq G$. Let $E(S)$ denote the number of $(x, y, z, w) \in S^4$ such that $x+y = z+w$. This quantity measures the amount of additive structure in S, is called the additive energy of S. Let us find a formula for $E(S)$ in terms of Fourier coefficients.

$$\begin{aligned} E(S) &= |G|^3 \cdot \mathbb{E}_{r \in G} \left[\left(\mathbb{E}_{y \in G} [1_S(y) 1_S(r-y) \right)^2 \right] \\ &= |G|^3 \cdot \mathbb{E}_{r \in G} \left[(1_S * 1_S(r))^2 \right] \\ &= |G|^3 \cdot \left(\sum_{a \in \hat{G}} |\widehat{1_S * 1_S}(a)|^2 \right) \\ &= |G|^3 \cdot \left(\sum_{a \in \hat{G}} |\widehat{1_S}(a)|^4 \right). \end{aligned}$$

Note that all the terms of this expression are nonnegative. This can be used to get a lower bound on $E(S)$. Indeed, notice that:

$$\widehat{1_S}(0) = \langle 1_S, \psi_0 \rangle = \frac{|S|}{|G|}.$$

So $E(S) \geq \frac{|S|^4}{|G|}$. Furthermore, $E(S)$ is close to $\frac{|S|^4}{|G|}$ if and only if all the Fourier coefficients of 1_S are "small". We will find this fact very useful.

1.7 Examples

Compute the Fourier transforms in the following two very instructive cases. Note which are the "large" Fourier coefficients.

1. Let $G = \mathbb{Z}_2^n$. Let S be a subspace of G, and let $f = 1_S$.
2. Let $G = \mathbb{Z}_p$. Let S be an interval in G, and let $f = 1_S$.

2 Fourier transforms on finite fields

When we work with finite fields, there are two groups hanging around, and thus we get two kinds of Fourier transforms at our disposal. On top of that, it will often be interesting and useful to take additively-defined sets/functions, and take their multiplicative Fourier transform, or to take multiplicatively-defined sets/functions and take their additive Fourier transforms, or more generally, to take sets/functions defined by polynomials and to take their Fourier transform(s).
This makes the theory of Fourier transforms on finite fields much richer than what we see on just finite abelian groups.

2.1 Additive and multiplicative characters

Let $q = p^n$ with p prime.
As we already know, the additive group of $\mathbb{F}_q$ is isomorphic to $\mathbb{Z}_p^n$, and the multiplicative group of $\mathbb{F}_q$ is isomorphic to $\mathbb{Z}_{q-1}$. Thus we can get 'explicit" descriptions of the additive and multiplicative characters of $\mathbb{F}_q$.
For additive characters, the description from the previous section would have required us to choose a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. Instead we use the representation of linear functions from $\mathbb{F}_q$ to $\mathbb{F}_p$ in terms of the trace map from $\mathbb{F}_q$ to $\mathbb{F}_p$. For each $a \in \mathbb{F}_q$, define the character $\psi_a : \mathbb{F}_q \to \mathbb{C}$ by:

$$\psi_a(x) = \omega_p^{\mathsf{Tr}(ax)}.$$

Then $\{\psi_a \mid a \in \mathbb{F}_q\}$ is the set of all characters of the additive group $\mathbb{F}_q$.
For multiplicative characters, an explicit description can be given using a generator of $\mathbb{F}_q^*$. For now, we note that for every $d \mid (q-1)$, there are exactly d characters χ such that χ^d equals the trivial character χ_0. These characters interact with the set of perfect dth powers in $\mathbb{F}_q^*$ very nicely:

$$\sum_{\chi \mid \chi^d = \chi_0} \chi(x) = \begin{cases} d & x \text{ is a perfect } d\text{th power.} \\ 0 & \text{otherwise} \end{cases}.$$

2.2 Gauss sums

The Gauss sum is the interaction of additive and multiplicative Fourier analysis at its finest.
Let ψ be an additive character of $\mathbb{F}_p$, and let χ be a multiplicative character of $\mathbb{F}_p^*$. We extend χ to a function on $\mathbb{F}_p$ by defining $\chi(0) = 0$. Then the Gauss sum corresponding to this situation is their inner product:

$$G(\psi, \chi) = \sum_{x \in \mathbb{F}_q} \psi(x)\chi(x).$$

Theorem 8 (Gauss)**.** *The absolute value of my sum is given by:*

$$|G(\psi, \chi)| = \begin{cases} q-1 & \psi = \psi_0, \chi = \chi_0 \\ 1 & \psi \neq \psi_0, \chi = \chi_0 \\ 0 & \psi = \psi_0, \chi \neq \chi_0 \\ \sqrt{q} & \psi \neq \psi_0, \chi \neq \chi_0 \end{cases}$$

Proof. The first 3 cases are obvious.
Now suppose $\chi \neq \chi_0$.

Proof 1: In this proof we will compute $|G(\psi, \chi)|$ for all ψ simultaneously. Let us compute the additive Fourier transform of χ. Clearly,

$$\hat{\chi}(0) = 0.$$

For each nonzero $a \in \mathbb{F}_q$:

$$\hat{\chi}(a) = \langle \chi, \psi_a \rangle = \mathbb{E}_{x \in \mathbb{F}_q}[\overline{\chi(x)} \cdot \omega_p^{-\mathsf{Tr}(ax)}].$$

We now exploit the fact that χ is a multiplicative character.

$$\begin{aligned} \hat{\chi}(a) &= \mathbb{E}_{x \in \mathbb{F}_q}[\overline{\chi(x/a)} \cdot \omega_p^{-\mathsf{Tr}(x)}] \\ &= \overline{\chi(a)} \cdot \mathbb{E}_{x \in \mathbb{F}_q}[\omega_p^{\mathsf{Tr}(x)} \cdot \overline{\chi(x)}] \\ &= \overline{\chi(a)}\hat{\chi}(1). \end{aligned}$$

Thus all the nonzero Fourier coefficients of χ have equal absolute value!
By Parseval's identity, we have:

$$\sum_a |\hat{\chi}(a)|^2 = \mathbb{E}_{x \in \mathbb{F}_q}[|\chi(x)|^2] = \frac{q-1}{q}.$$

Thus $|\hat{\chi}(a)| = \frac{1}{\sqrt{q}}$ for each nonzero $a \in \mathbb{F}_q$, as desired.

Proof 2: This proof proceeds directly.

$$|G(\psi,\chi)|^2 = \left|\sum_{x\in\mathbb{F}_q}\psi(x)\chi(x)\right|^2 \tag{1}$$

$$= \left(\sum_{x\in\mathbb{F}_q}\psi(x)\chi(x)\right)\overline{\left(\sum_{y\in\mathbb{F}_q}\psi(y)\chi(y)\right)} \tag{2}$$

$$= \sum_{x\in\mathbb{F}_q,y\in\mathbb{F}_q^*}\psi(x)\psi(-y)\chi(x)\chi(1/y) \tag{3}$$

$$= \sum_{x\in\mathbb{F}_q,y\in\mathbb{F}_q^*}\psi(x-y)\chi(x/y) \tag{4}$$

$$= \sum_{x\in\mathbb{F}_q,y\in\mathbb{F}_q^*}\psi(y(\frac{x}{y}-1))\chi(x/y) \tag{5}$$

$$= \sum_{u\in\mathbb{F}_q,y\in\mathbb{F}_q^*}\psi(y(u-1))\chi(u) \tag{6}$$

$$= \sum_{u\in\mathbb{F}_q}\chi(u)\left(\sum_{y\in\mathbb{F}_q^*}\psi(y(u-1))\right) \tag{7}$$

$$= \chi(1)\cdot(q-1) + \sum_{u\in\mathbb{F}_q\setminus\{1\}}\chi(u)\cdot(-1) \tag{8}$$

$$= (q-1) - \sum_{u\in\mathbb{F}_q\setminus\{1\}}\chi(u) \tag{9}$$

$$= q. \tag{10}$$

□

Some remarks:

1. χ is extremely pseudorandom from the point of view of the additive Fourier transform (all Fourier coefficients are $O(\frac{1}{\sqrt{q}})$).
2. ψ is extremely random from the point of view of the multiplicative Fourier transform (all Fourier coefficients are $O(\frac{1}{\sqrt{q}})$).
3. A random function from $\mathbb{F}_q$ to the unit circle will, with high probability, have its largest Fourier coefficient equal to $\Theta(\sqrt{\frac{\log q}{q}})$. Thus the characters are even more pseudorandom than random functions!

Next, we will see some applications of Gauss sums.

3 Correlation between additive and multiplicative objects

We will be dealing with some additive and multiplicative objects in $\mathbb{F}_q$.
A *subspace* is an additive subgroup (equivalently, an $\mathbb{F}_p$-linear subspace of $\mathbb{F}_q$, where p is the characteristic of q). An arithmetic progression (AP) is a set of the form $\{b, b+c, b+2c, \ldots, b+kc\}$. Both of these sets have nice additive Fourier transforms.
A multiplicative subgroup is a subgroup of the multiplicative group $\mathbb{F}_q^*$. A geometric progression (GP) is a set of the form $\{b, bc, bc^2, \ldots, bc^k\}$. Note that by cyclicity of $\mathbb{F}_q^*$, every multiplicative subgroup is a geometric progression. Geometric progressions have nice multiplicative Fourier transforms.

Theorem 9 (Polya-Vinogradov)**.** *Suppose S is either:*

- *a subspace in $\mathbb{F}_q$ with q not prime, or*
- *an AP in $\mathbb{F}_q$ with q prime.*

Let χ be a nontrivial multiplicative character of $\mathbb{F}_q$. Then:

$$\left|\sum_{x\in S}\chi(x)\right| \leq C\cdot\sqrt{q},$$

where $C = 1$ if S is a subspace, and $C = O(\log q)$ if S is an AP.

Proof. The crux of the proof is to realize that we want to compute the inner product of χ with the indicator function of S, and to then apply the Parseval identity to reduce it to a question about the Fourier coefficients of 1_S.

$$\begin{aligned}
\sum_{x\in S}\chi(x) &= \sum_{x\in\mathbb{F}_q}\chi(x)1_S(x)\\
&= q\cdot\langle\chi, 1_S\rangle\\
&= q\cdot\sum_{a\in\mathbb{F}_q}\hat{\chi}(a)\overline{\hat{1_S}(a)}\\
&\leq q\cdot\sum_{a\in\mathbb{F}_q}|\hat{\chi}(a)|\cdot|\hat{1_S}(a)|\\
&\leq q\cdot\sum_{a\in\mathbb{F}_q}\frac{1}{\sqrt{q}}\cdot|\hat{1_S}(a)|\\
&\leq \sqrt{q}\cdot\sum_{a\in\mathbb{F}_q}|\hat{1_S}(a)|.
\end{aligned}$$

It remains to bound $\sum_{a\in\mathbb{F}_q}|\hat{1_S}(a)|$ for S as in the theorem statement.
For S a subspace, it is a simple linear algebra exercise to show that:

$$\hat{1}_S(a) = \begin{cases}\frac{|S|}{q} & a\in S^\perp\\ 0 & \text{otherwise}\end{cases},$$

where[2]

$$S^\perp = \{a\in\mathbb{F}_q \mid \mathsf{Tr}(ax) = 0 \text{ for all } x\in S\}.$$

Note that $S^\perp$ has size $\frac{q}{|S|}$. Thus $\sum_{a\in\mathbb{F}_q}|\hat{1_S}(a)| = 1$, as desired.

[2]**Quick linear algebra recap + warning.** If S is a d dimensional $\mathbb{F}_p$-subspace of $\mathbb{F}_p^n$, then $S^\perp = \{a\in\mathbb{F}_p^n \mid \langle a, x\rangle = 0\}$ is a $\mathbb{F}_p$-linear space of dimension $n-d$. BUT, $S^\perp$ need not be disjoint from S, and it need not be true that $S + S^\perp = \mathbb{F}_p^n$. There can be nonzero vectors x such that $\langle x, x\rangle = 0$. If you are seeing these facts for the first time, it is a great exercise to go over the proofs of the basic facts of "usual" linear algebra, see what exactly is different between $\mathbb{F}_p$ and the $\mathbb{R}$ cases, and take a moment to reflect some of the elementary properties of $\mathbb{R}$ that we take for granted.

For S an AP and q prime, we can proceed as follows. Suppose $S = \{b, b+c, \ldots, b+kc\}$. Then

$$\begin{aligned}
|\hat{1}_S(a)| &= \left| \frac{1}{q} \sum_{j=0}^{k} \psi_a(b+jc) \right| \\
&= \left| \frac{1}{q} \cdot \psi_a(b) \cdot \left(\sum_{j=0}^{k} \psi_a(jc) \right) \right| \\
&= \frac{1}{q} \cdot \left| \left(\sum_{j=0}^{k} \omega_q^{jac} \right) \right|
\end{aligned}$$

We may bound this expression in two ways. The first bound is just the trivial:

$$|\hat{1}_S(a)| \leq \frac{k+1}{q} = \frac{|S|}{q}.$$

Alternately, we can say:

$$\begin{aligned}
|\hat{1}_S(a)| &= \frac{1}{q} \left| \frac{1-\omega_q^{(k+1)ac}}{1-\omega_q^{ac}} \right| \\
&\leq \frac{1}{q} \frac{2}{\left|1-e^{2\pi iac/q}\right|} \\
&\leq \frac{1}{q} \frac{2}{|\sin(\pi iac/q)|} \\
&\leq \frac{1}{q} \frac{2}{2\|ac/q\|} \\
&\leq \frac{1}{q} \frac{1}{\|ac/q\|},
\end{aligned}$$

where, for a real number h, $\|h\|$ denotes the distance of h from the nearest integer.

Thus

$$\begin{aligned}
\sum_{a\in\mathbb{F}_q}|\hat{1}_S(a)| &\le \sum_{0\le a\le q-1}\min\left(\frac{|S|}{q},\frac{1}{q}\frac{1}{\|ac/q\|}\right)\\
&\le \sum_{a:\|ac/q\|<\frac{1}{|S|}}\frac{|S|}{q}+\sum_{a:\|ac/q\|\ge\frac{1}{|S|}}\frac{1}{q}\frac{1}{\|ac/q\|}\\
&= \left|\{0\le a'<q-1:\|a'/q\|<\frac{1}{|S|}\}\right|\cdot\frac{|S|}{q}+\frac{1}{q}\cdot\sum_{0\le a'<q-1:\|a'/q\|\ge\frac{1}{|S|}}\frac{1}{\|a'/q\|}\\
&\qquad \text{where } a'\in\{0,1,\ldots,q-1\} \text{ is such that } a'=ac \mod q\\
&\qquad \text{note that } \|a'/q\|=\|ac/q\|.\\
&\le \frac{q}{|S|}\cdot\frac{|S|}{q}+2\cdot\frac{1}{q}\cdot\sum_{q/|S|\le a'\le q/2}\frac{q}{a'}\\
&\le 1+2\cdot\sum_{q/|S|\le a'\le q/2}\frac{1}{a'}\\
&\le 1+2\cdot\sum_{q/|S|\le a'\le q/2}\frac{1}{a'}\\
&\le O(\log|S|)\\
&\le O(\log q).
\end{aligned}$$

□

Lemma 10. *Let $H\subseteq\mathbb{F}_q^*$ be a geometric progression. Let ψ be a nontrivial additive character of $\mathbb{F}_q$. Then:*

$$\left|\sum_{x\in H}\psi(x)\right|\le C\cdot\sqrt{q},$$

where $C=1$ if H is a multiplicative subgroup, and $C=O(\log q)$ if H is a geometric progression that need not be a subgroup.

Proof. The proof is very similar to the previous proof.
Now we use the *multiplicative* Fourier transform. We are interested in bounding $\left|\sum_{x\in H}\psi(x)\right|$, which equals $|(q-1)\cdot\langle\psi,1_H\rangle|$. By Parseval, this equals $\left|(q-1)\cdot\sum_\chi\hat{\psi}(\chi)\overline{\hat{1_H}(\chi)}\right|$. Using the Gauss sum and the triangle inequality, this is at most:

$$\sqrt{q}\cdot\sum_{\chi\neq\chi_0}|\hat{1_H}(\chi)|.$$

Thus we need to bound $\sum_{\chi\neq\chi_0}|\hat{1_H}(\chi)|$. If H is a geometric progression, then via the isomorphism $\mathbb{F}_q^*\cong\mathbb{Z}_{q-1}$, we need to bound the sum of absolute values of Fourier coefficients of an AP in $\mathbb{Z}_{q-1}$, and we already showed in the previous section that it is at most $O(\log q)$. If H is in fact a subgroup of Z_{q-1}, then we can write the Fourier expansion of 1_H directly:

$$1_H=\sum_{\chi:\chi^d=\chi_0}\frac{1}{d}\cdot\chi,$$

and thus $\sum_{\chi\neq\chi_0}|\hat{1_H}(\chi)|=1$. This completes the proof. □

Lemma 11. *Let S be a subspace or an AP in $\mathbb{F}_q$, and let H be a multiplicative subgroup of $\mathbb{F}_q^*$. Then:*

$$|S\cap H|=\frac{|S|\cdot|H|}{q}+O(\sqrt{q}\log q).$$

The proof is similar to the earlier proofs. We write 1_H as a combination of multiplicative characters, 1_S as a combination of additive characters, and estimate their inner product using the Gauss sum. The term coming from χ_0 and ψ_0 contributes the main term $\frac{|S|\cdot|H|}{q}$, and the remaining terms sum up to a small quantity because of what we know about the Fourier coefficients of 1_H and 1_S. You should work out the details.
This is again a pseudorandom phenomenon. If S, H were random sets of given sizes, we would expect their intersection to have size around $\frac{|S|\cdot|H|}{q}$.

4 The Waring problem in finite fields

Many centuries ago, Waring made a conjecture. He conjectured that for every integer $d > 0$, there is an integer k such that every nonnegative integer can be written as the sum of at most k perfect d'th powers. This generalized the four squares theorem, which was well known at the time.
The Waring problem was first solved by Hilbert in the early 1900s using some delicate identities (I don't know much about this proof other than it was very clever). The definitive approach to the Waring problem came a few years later, by Hardy and Littlewood in the 1920s, who used (essentially) Fourier analysis.
We will study the analogue of the Waring question in finite fields: can every element of $\mathbb{F}_q$ be written as the sum of perfect dth powers. The approach will be through Fourier analysis, and the details will be significantly cleaner than what is required in the integer case.
We will do even more: we will try to *count* the number of ways of expressing an element of $\mathbb{F}_q$ as a sum of perfect dth powers. First note that without loss of generality, we may assume that $d \mid q-1$, because the set of perfect dth powers equals the set of perfect d'th powers, where $d' = GCD(d, q-1)$. We will also focus on *nonzero* perfect dth powers; later we will discuss the impact of this nonzeroness condition.

4.1 The basic Fourier argument

Let $H \subseteq \mathbb{F}_q^*$ be the set of nonzero perfect dth powers. Thus $|H| = (q-1)/d$. The number of ways x can be written as a sum of k nonzero perfect dth powers can be expressed in terms of 1_H as follows:

$$\begin{aligned}
|\{(y_1,\ldots,y_k) \in H^k \mid \sum y_i = x\}| &= \sum_{y_1,\ldots,y_{k-1}\in\mathbb{F}_q} 1_H(y_1)\cdot 1_H(y_2)\cdots 1_H(y_{k-1}) 1_H(x - \sum_{j=1}^{k-1} y_j) \\
&= q^{k-1}\cdot \mathbb{E}_{y_1,\ldots,y_{k-1}\in\mathbb{F}_q}\left[1_H(y_1)\cdot 1_H(y_2)\cdots 1_H(y_{k-1}) 1_H(x - \sum_{j=1}^{k-1} y_j)\right] \\
&= q^{k-1} 1_H * 1_H * \ldots * 1_H(x) \\
&= q^{k-1} 1_H^{(*k)}(x).
\end{aligned}$$

Thus we want to compute (or at least estimate) $1_H^{(*k)}(x)$. We can do this via the (additive) Fourier expansion:

$$\begin{aligned}
1_H^{(*k)}(x) &= \sum_{a\in\mathbb{F}_q} \widehat{1_H^{(*k)}}(a)\psi_a(x) \\
&= \sum_{a\in\mathbb{F}_q} (\hat{1_H}(a))^k \psi_a(x).
\end{aligned}$$

By the result from the previous section, for every nonzero a we have $|\hat{1_H}(a)| \le \frac{1}{\sqrt{q}}$. We also have $\hat{1_H}(0) = \frac{|H|}{q}$.

Thus:

$$\begin{aligned} 1_H^{(*k)}(x) &= \left(\frac{|H|}{q}\right)^k + \sum_{a\neq 0} (\hat{1_H}(a))^k \psi_a(x) \\ &= \left(\frac{|H|}{q}\right)^k \pm \sum_{a\neq 0} q^{-k/2} \\ &= \left(\frac{|H|}{q}\right)^k \pm q^{-(k-2)/2} \end{aligned}$$

Thus we have proved:

Theorem 12. *For d, k, q, H as above, for every $x \in \mathbb{F}_q$, we have:*

$$\left|\left\{(y_1, \ldots, y_k) \in H^k \mid \sum y_i = x\right\}\right| = \frac{|H|^k}{q} \pm q^{k/2}$$

ways.

Note the strong pseudorandomness undercurrents here: the main term is precisely what we would expect if H was a random set of a given size.
Two important corollaries:

- for every d, there exists q_0, such that for all $q > q_0$, every $x \in \mathbb{F}_q$ can be written as the sum of exactly **THREE** nonzero d'th powers in $\mathbb{F}_q$
- for every $\alpha < 1/2$, for every $k > \frac{1}{\frac{1}{2}-\alpha}$, for every q and $d < q^{1-\alpha}$, every $x \in \mathbb{F}_q$ can be written as the sum of exactly k nonzero d'th powers in $\mathbb{F}_q$.

Note that the above theorem does not say anything interesting when $|H| < \sqrt{q}$. THIS IS FOR A GOOD REASON. The homework will talk further about this.
The argument above is also very general. It used nothing more than the smallness of the Fourier coefficients of 1_H.

4.2 An improved argument

We will now see that for $d \ll q^{1/4}$, every **nonzero** element of $\mathbb{F}_q$ can be written as the sum of **TWO** nonzero d'th powers. In fact, we will show that each nonzero element of $\mathbb{F}_q$ can be written as a sum of two elements from H in about $\frac{|H|^2}{q}$ ways, as we would expect from a random set.
There is something fundamentally different about what happens when we sum k elements of H if $k = 2$ or $k > 2$. This difference has to do with 0. If -1 happens to not be a perfect dth power in $\mathbb{F}_q$, then 0 cannot be written as a sum of nonzero d'th powers at all. If -1 happens to be a perfect d'th power in $\mathbb{F}_q$, then 0 can be written as a sum of nonzero d'th powers in $|H|$ ways. In either case, the number of ways we can write 0 as a sum of nonzero d'th powers is very different from the $\frac{|H|^2}{q}$ we would expect if H was a random set. This is the major difference between the $k = 2$ case and the $k > 2$ case; the argument of the previous section based only on the smallness of $|\hat{1_H}(a)|$ could not distinguish between 0 and nonzero elements, and thus cannot be extended to give something interesting when $k = 2$.
Instead, we use some properties strongly exploiting the fact that H is a subgroup. We do this via the relations between the different Gauss sums, going beyond just the simple bound on their absolute value.

Theorem 13. *For d, q, H as above, for every $x \in \mathbb{F}_q^*$, we have:*

$$\left|\{(y_1, y_2) \in H^2 \mid y_1 + y_2 = x\}\right| = \frac{|H|^2}{q} \pm \sqrt{q}.$$

Proof. In the proof of Theorem 8, we showed that the different Gauss sums are related to one another in a nice way. Let us write that down explicitly

$$\begin{aligned} G(\psi_a, \chi) &= \sum_{x \in \mathbb{F}_q} \psi_a(x)\chi(x) = \sum_x \psi_1(ax)\chi(x) \\ &= \sum_x \psi_1(x)\chi(x/a) = \overline{\chi(a)} \cdot \left(\sum_x \psi_1(x)\chi(x) \right) \\ &= \overline{\chi(a)} \cdot G(\psi_1, \chi). \end{aligned}$$

Now the quantity of interest can be written as follows:

$$\begin{aligned} \left|\{(y_1, y_2) \in H^2 \mid y_1 + y_2 = x\}\right| &= \sum_{y_1, y_2 \in \mathbb{F}_q} 1_H(y_1) 1_H(y_2) 1_{y_1 + y_2 = x} \\ &= \frac{1}{q} \sum_{y_1, y_2 \in \mathbb{F}_q} \sum_{a \in \mathbb{F}_q} 1_H(y_1) 1_H(y_2) \psi_a(y_1 + y_2 - x) \\ &= \frac{1}{q} \sum_{y_1, y_2 \in \mathbb{F}_q} \sum_{a \in \mathbb{F}_q} 1_H(y_1) 1_H(y_2) \psi_a(y_1)\psi_a(y_2)\psi_a(-x) \\ &= \frac{1}{qd^2} \sum_{y_1, y_2 \in \mathbb{F}_q} \sum_{a \in \mathbb{F}_q} \left(\sum_{\chi : \chi^d = \chi_0} \chi(y_1) \right) \left(\sum_{\chi' : (\chi')^d = \chi_0} \chi'(y_2) \right) \psi_a(y_1)\psi_a(y_2)\psi_a(-x) \\ &= \frac{1}{qd^2} \sum_{a \in \mathbb{F}_q} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} \sum_{y_1, y_2 \in \mathbb{F}_q} \chi(y_1)\chi'(y_2)\psi_a(y_1)\psi_a(y_2)\psi_a(-x) \\ &= \frac{1}{qd^2} \sum_{a \in \mathbb{F}_q} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_a, \chi) G(\psi_a, \chi')\psi_a(-x) \\ &= \frac{(q-1)^2}{qd^2} + \sum_{a \in \mathbb{F}_q^*} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_a, \chi) G(\psi_a, \chi')\psi_a(-x). \end{aligned}$$

Now we use the relations between the Gauss sums:

$$\begin{aligned} \left|\{(y_1, y_2) \in H^2 \mid y_1 + y_2 = x\}\right| &= \frac{|H|^2}{q} + \frac{1}{qd^2} \sum_{a \in \mathbb{F}_q^*} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_1, \chi) G(\psi_1, \chi') \overline{\chi(a)\chi'(a)} \psi_a(-x) \\ &= \frac{|H|^2}{q} + \frac{1}{qd^2} \sum_{a \in \mathbb{F}_q^*} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_1, \chi) G(\psi_1, \chi') \overline{\chi \cdot \chi'}(a) \psi_a(-x) \\ &= \frac{|H|^2}{q} + \frac{1}{qd^2} \sum_{a \in \mathbb{F}_q^*} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_1, \chi) G(\psi_1, \chi') \overline{\chi \cdot \chi'}(a) \psi_{-x}(a) \\ &= \frac{|H|^2}{q} + \frac{1}{qd^2} \sum_{\chi : \chi^d = \chi_0} \sum_{\chi' : (\chi')^d = \chi_0} G(\psi_1, \chi) G(\psi_1, \chi') G(\psi_{-x}, \overline{\chi \cdot \chi'}). \end{aligned}$$

This final sum we can estimate using our bounds for the absolute value of the Gauss sum, and noting that ψ_1, ψ_{-x} are nontrivial characters (this is where the nonzeroness of x gets used!):

$$\left|\{(y_1, y_2) \in H^2 \mid y_1 + y_2 = x\}\right| = \frac{|H|^2}{q} \pm \sqrt{q}.$$

□