The ℓ-adic Fourier Transform

Akshay Venkatesh∗

Mary 24, 2017

1 Kloosterman sums

Let \( \psi: \mathbb{F}_p \to \mathbb{C}^\times \) or \( \mathbb{Q}_p^\times \) be a character, e.g.
\[
    x \mapsto e^{2\pi i x/p}.
\]
An important property is that
\[
    \sum_{a \in \mathbb{F}_p} \psi(ay) = \begin{cases} 
        0 & y \neq 0, \\
        p & y = 0
    \end{cases} \quad (1.1)
\]

For \( a \in \mathbb{F}_p^\times \), the Kloosterman sum \( K(a) \) is
\[
    K(a) = \sum_{xy=a} \psi(x+y) = \sum_{x \in \mathbb{F}_p, x \neq 0} \psi(ax + x^{-1}). \quad (1.2)
\]

**Example 1.1.** \( K(a) \) is always a real number, because the sum is symmetric with respect to complex conjugation. For \( p = 7 \) and \( a = 1 \), it is
\[
    \zeta_7^2 + \zeta_7^{-1}s + \zeta_7 + \zeta_7^{-2} + \zeta_7 + \zeta_7^{-1}.
\]

**Remark 1.2.** The analogue of \( K(a) \) over \( \mathbb{R} \) would be something like
\[
    \int_{\mathbb{R}} e^{i(ax+1/x)} \, dx.
\]
This isn’t convergent, but
\[
    \int_{\mathbb{R}} e^{-ax+1/x} \, dx \sim \sqrt{a}K(\sqrt{a})
\]
where \( K \) is a Bessel function. There is a parallel between these special functions and the special sheaves that will arise later.

∗notes by Tony Feng
This was first studied by Kloosterman, in analyzing the Hardy-Littlewood circle method. Obviously \(|K(a)| < p\); Kloosterman wanted to show that \(|K(a)|\) is much less than \(p\). He showed that

\[ |K(a)| \leq p^{3/4}. \]

Weil later improved this to \(|K(a)| \leq 2\sqrt{p}\).

We’ll discuss Kloosterman’s proof. He studied the sum

\[ \sum_{a} |K(a)|^4 \]

and showed

\[ \sum_{a} |K(a)|^4 \leq cp^3. \]

We start off by writing

\[ \sum_{a} |K(a)|^4 = \sum_{a} \sum_{x,y,z,w,} \psi(a(x + y - z - w) + (x^{-1} + y^{-1} - z^{-1} - w^{-1}) \]

Using \(1.1\), this is

\[ = p \sum_{x+y=z+w} \psi(x^{-1} + y^{-1} - z^{-1} - w^{-1}) \]

\[ = p \left( \sum_{x+y=z+w, x^{-1}+y^{-1}=z^{-1}+w^{-1}, x^{-1}+y^{-1} \neq z^{-1}+w^{-1}} \psi(x^{-1} + y^{-1} - z^{-1} - w^{-1}) + \# \left\{ x^{-1} + y^{-1} = z^{-1} + w^{-1} \right\} \right) \]

Recall that we want to get a bound of \(p^3\), while the trivial bound is about \(p^4\). The second term \(\# \left\{ x^{-1} + y^{-1} = z^{-1} + w^{-1} \right\}\) has only about \(p^2\) terms, so that’s good. The first term is \(p\) times a sum over \(p^3\) things, so we need to do something intelligent there. Luckily, it has a scaling symmetry. The sum over each \(\mathbb{F}_p^*\)-orbits is \(-1\) by \(1.1\).

So, letting \(N = \# \{ x + y = z + w \}\) and \(A = \# \{ x^{-1} + y^{-1} = z^{-1} + w^{-1} \}\), we get

\[ \sum_{a} |K(a)|^4 = p \left( \frac{N - A}{p - 1}(-1) + A \right). \]

To conclude that, note that \(N\) has size about \(p^3\), \(A\) has size about \(p^2\).

**Remark 1.3.** You can evaluate \(N\) and \(A\) exactly. For odd \(p\), we think \(A = 3(p - 2)(p - 1)\). This gives

\[ \sum_{a} |K(a)|^4 = 2p^3 + \text{ (lower order terms)}. \]
Kloosterman’s argument is the earliest instance I know of the following principle: to bound a single value of a function, put that function in a family and raise it to a higher power. This idea was used again by Rankin in the context of modular forms, which Deligne said was an inspiration for his proof of the Weil conjectures.

2 The sheaf-function correspondence

We now want to implement this idea with sheaves. Let $X$ be a variety over $\mathbb{F}_p$. Suppose you have a Weil sheaf $\mathcal{F}$ on $X$, meaning a sheaf on $X_{\mathbb{F}_p}$ with a Frobenius endomorphism. Then we get a function $f$ on $|X|$ or $X(\mathbb{F}_p^n)$, given by $f(x) =$ trace of geometric Frobenius at $x$.

Remark 2.1. If for example $\mathcal{F}$ is lisse and semisimple, then the function $f$ determines the sheaf, because the Frobenii are dense in the monodromy group.

2.1 Translation of sheaf-theoretic operations

Operations on sheaves can be translated into operations on functions.

- The tensor product of sheaves translates into product of functions.
  $$\mathcal{F} \otimes \mathcal{G} \mapsto f_\mathcal{F} \cdot f_\mathcal{G}$$

- The pullback of sheaves translates into pullback of functions.
  $$\pi^* \mathcal{F} \mapsto f_\mathcal{F} \circ \pi$$

- If $\mathcal{F}$ is pure of weight $w \in \mathbb{Z}$, then $\mathcal{F}^\vee$ corresonds to the function $\overline{f} p^{-w} \deg$. For instance, the Kloosterman sums being real-valued corresponds to the Kloosterman sheaves being self-dual up to Tate twist.

- If $\mathcal{F}$ is in the derived category of Weil sheaves\footnote{Although we have glided over this point in this seminar, the construction of the “derived category of $\ell$-adic sheaves” (or Weil sheaves) is actually quite subtle. It is not obtained by the naïve construction taking the derived category of a category of $\ell$-adic sheaves, although this is often what one pretends for practical purposes. Suffice it to say that working rigorously with the “derived category of $\ell$-adic sheaves” requires a good deal more care than one might think; “arguments” which treat this category as a genuine derived category are merely reasoning by analogy.} then
  $$R\pi_! \mathcal{F} \mapsto \sum_i (-1)^i f_{H^i, \mathcal{F}}$$

With these conventions, the Lefschetz trace formula translates into the statement that the derived pushforward of sheaf corresponds to the pushforward of $f$ as defined by

$$y \in Y(\mathbb{F}_p^n) \mapsto \sum_{x \in X(\mathbb{F}_p^n), \pi(x) = y} f(x).$$
For example, given a map \( \pi: X \to \mathbb{A}^1 \), then the function
\[
y \in \mathbb{F}_p^m \mapsto \#X_y(\mathbb{F}_p^m)
\]
comes from \( R\pi_! \mathbb{Q}_\ell \).

**Example 2.2.** We’re going to make a sheaf corresponding to the function \( \psi \).

We start out with the Artin-Schreier cover
\[
y^p - y = x \subset \mathbb{A}^2.
\]
This maps via the \( x \)-coordinate to \( \mathbb{A}^1 \), which is an étale cover. The Galois group is canonically \( \mathbb{Z}/p\mathbb{Z} \), generated by \( y \mapsto y + 1 \). In other words, this cover induces a map
\[
\pi_1(\mathbb{A}^1_{\mathbb{F}_p}) \to \mathbb{Z}/p\mathbb{Z} \xrightarrow{\psi} \mathbb{Q}_\ell^*.
\]
Let \( \mathcal{L}_\psi \) be the corresponding rank 1 lisse sheaf on \( \mathbb{A}^1 \). We compute the associated function. We need to figure out where Frobenius goes. The Frobenius at \( x \in \mathbb{A}^1(\mathbb{F}_p^m) \) takes \( (x, y) \mapsto (x^{p^m}, y^{p^m}) \). Of course \( x^{p^m} = x \). We have
\[
\begin{align*}
y^p &= y + x \\
y^{p^2} &= y^p + x^p = y + x + x^p \\
&
\vdots \\
y^{p^m} &= y + x + x^p + \ldots + x^{p^{m-1}}
\end{align*}
\]
Therefore, the Frobenius at \( x \) takes
\[
(x, y) \mapsto (x, y + x + x^p + x^{p^2} + \ldots + x^{p^{m-1}}).
\]
The conclusion is that Frobenius acts on the stalk \( \mathcal{L}_\psi \) as multiplication by \( \psi(\text{Tr}_{\mathbb{F}_p^{p^m}/\mathbb{F}_p}(x)) \).
Therefore geometric Frobenius acts as multiplication by \( \psi(- \text{Tr}_{\mathbb{F}_p^{p^m}/\mathbb{F}_p}(x)) = \overline{\psi}(\text{Tr}_{\mathbb{F}_p^{p^m}/\mathbb{F}_p}(x)) \).

### 2.2 The method of families

Suppose we have an open subset \( U \subset \mathbb{A}^1 \), and \( \mathcal{G} \) is a Weil sheaf on \( U \), associated to a function \( g \). Assume \( g \geq 0 \). (This can be arranged by taking the sum of \( \mathcal{G} \) with its conjugate.)

We can then bound a single value of \( g \) by a sum:
\[
g(x) \leq \sum_{x \in \mathbb{F}_p^m} g(x).
\]
Of course this isn’t sharp, but if you apply this to large powers of \( g \) then it will be sharp.
To simplify things, assume \( H^0_c(G) = 0 \). Then
\[
\sum_{x \in U(F_p^m)} g(x) = \sum_{\beta = \text{eig. of } F \text{ on } H^2_c} \beta^m - \sum_{\alpha = \text{eig. of } F \text{ on } H^1_c} \alpha^m.
\]

We can easily analyze the \( \beta \)'s, by using Poincaré duality to relate \( H^2_c \) to the coinvariants of geometric \( \pi_1 \). But the point is that \( \max |\alpha| \leq \max |\beta| \), which allows you to ignore \( \alpha \). Why? This is because the expression is positive.

**Remark 2.3.** This same observation appears elsewhere. For instance, the Weil bound is not optimal. For a curve it gives \( p + 1 + 2g\sqrt{p} \), but this can’t be attained because it would give a negative number of points over \( F_p^2 \).

This argument (applied to a high power of \( g \)) is the engine that provides the bounds in the proof.

The central part of the proof will be the following statement: if \( F \) is pure of weight \( w \) on \( U \subset \mathbb{A}^1 \), then the weights of \( H^1_c(U, F) \) are \( \leq w + 1 \). In terms of the functions \( f_F \) associated to \( F \), this statement translates to the bound
\[
\sum_{x \in U(F_p^m)} f_F(x) \leq p^{m(w+1)/2}.
\]

If \( H^0_c = 0 \) and \( H^2_c = 0 \) (it is easy to reduce to this case), then
\[
\text{Tr}(\text{Frob}^m, H^1_c) = - \sum_{x \in U(F_p^m)} f_F(x)
\]
so this becomes a question of bounding the eigenvalues of Frobenius on cohomology.

To obtain this estimate, we try to embed it into a family. We want to find a function \( g \) on \( \mathbb{A}^1 \) (associated to a sheaf) such that \( g(0 \in F_p^m) = \sum f_F(x) \). Then we’ll use the method of families. The punchline is that we take \( g \) to be the Fourier transform of \( f \). So next we’ll make a sheaf associated to the function \( g = \text{FT}(f) \), defined by
\[
\text{FT}(f_F)(y) = \sum_{x \in F_p^m} f_F(x) \psi \circ \text{Tr}(yx)
\]
for \( y \in F_p^m \).

### 3 Fourier transform

Let \( F \) be a sheaf on \( \mathbb{A}^1 \). We make a new sheaf \( \text{FT}_\psi F \) such that
\[
f_{\text{FT}_\psi}(y) = \sum_{x \in F_p^m} f_F(x) \psi \circ \text{Tr}(yx).
\]

We just replicate the Fourier transform step-by-step.
We start with $F$, pull it back to $A_2$ via $(x, y) \mapsto x$. Then we tensor with $m^*L_\psi$, where $m(x, y) = xy$. Finally, to sum over the first variable we push forward via $(x, y) \mapsto y$. The last step is to shift by degree 1, basically to preserve the property of being a sheaf (but it still might not quite).

Denote this functor by $F_\psi$.

**Theorem 3.1.** We have

$$F_\psi \circ F_\psi = \text{Id (up to Tate twist)}.$$ 

This mirrors the usual calculation

$$\sum_y \psi(-yz) \sum_x f(x) \psi(yx) = \sum_{x,y} f(x) \psi(y(x - z)) = p^m f(z).$$

The proof replicates this calculation at the level of sheaves. The only step that wasn’t formal was the calculation

$$\sum_{a \in \mathbb{F}_p} \psi(ay) = \begin{cases} 
0 & y \neq 0 \\
p & 
\end{cases}$$

so we need a sheaf-theoretic analogue of it, which is

$$H^*_c(\mathbb{A}_{\mathbb{F}_p}^1, L_\psi) = 0.$$ 

To prove this, recall that the sheaf $L_\psi$ came from the covering

$$y^p - y = x$$

by taking the $\psi$-component of the pushforward of the constant sheaf. Then $H^*_c$ is the $\psi^{\pm 1}$-component of $H^*_c(C, \overline{\mathbb{Q}}_C)$, which is 0 (since $C = \mathbb{A}^1$).

The idea of the proof of the Weil conjectures is to bound $F$-eigenvalues on $H^1_c(U \subset \mathbb{A}^1, \mathcal{G})$, which is the fiber at 0 of $FT_\psi(\mathcal{G})$.

## 4 Kloosterman sheaves

Recall that we defined the Kloosterman function

$$K(a) = \sum_{xy = a} \psi(x + y) = \sum_{x \in \mathbb{F}_p, x \neq 0} \psi(ax + x^{-1}).$$

We’re going to make a sheaf $Kl$ on $\mathbb{G}_m$ such that

$$f_{Kl}(a \in \mathbb{F}_p^m) = \sum_{x \in \mathbb{F}_p^m} \psi \circ \text{Tr}(ax + x^{-1}).$$
We start with $L_\psi$ on $G_m$ to get $\psi(x)$, apply inversion to get $\psi(x^{-1})$, and apply $\text{FT}_\psi$.

This gives a lisse sheaf $K_l$ on $G_m$, pure of weight 1. Since $\text{rank}(K_l) = 2$, this corresponds to a representation $\pi_1(G_m) \to \text{GL}_2(\overline{Q}_l)$ whose Zariski closure is $\text{SL}_2$ (the real-ness suggests the sheaf is self-dual).

Suppose we want to understand $K_{l,a}$ for $a \in \mathbb{F}_p$. Take $a = 1$. It is the $\psi$-component of $H^1_c(y^p - y = x + x^{-1})$. We will show that $\dim H^1_c(y^p - y = x + x^{-1}) = 2(p - 1) + 1$. This strongly suggests that, because there are $p - 1$ characters $\psi$, each piece has dimension 2.

If you actually want to compute, you have to understand the behavior of the sheaf at $\infty$. Consider

$$y^p - y = x + x^{-1} \to G_m$$

and compactify it to $X \to \mathbb{P}^1$, of degree $p$.

By Riemann-Hurwitz,

$$2g_X - 2 = p(-2) + \deg(\text{ram. divisor}).$$

The ramification is supported at 0, $\infty$. Since the equation $y^p - y = x + x^{-1}$ is symmetric, the answer will be the same at both points, so we just to the calculation 0. Localizing at 0, we need to consider the field extension $L/K$ where $K = \mathbb{F}_p((x))$ and $L = K(y)$ with $y^p - y = x + x^{-1}$. Since $v(x) = 1$, we have $v(y) = -1/p$, $\tau = y^{-1}$ is a uniformizer. The discriminant is the field extension

$$\prod_{i \neq j} (\tau_i - \tau_j)$$

Since the conjugates just add, a typical term is

$$\frac{1}{y+1} - \frac{1}{y} = \frac{-1}{y(y+1)}$$

with valuation $2/p$. So the discriminant has valuation $p(p - 1)(2/p) = 2(p - 1)$.

(This is double what one would expect in characteristic 0.) So the conclusion is that

$$2g_X - 2 = p(-2) + 4(p - 1) \implies \dim H^1(X) = 2(p - 1).$$

To get $C$ from the compactified guy, you delete 2 points so

$$\dim H^1(C) = 2(p - 1) + 1.$$

So we’ve verified that $\text{rank } KL = 2$.

**Remark 4.1.** This is related to the fact that the $K$-Bessel function from Remark 1.2 satisfies a second-order differential equation.
Let’s return to the estimate:

$$\sum |K(a)|^4 \sim 2p^3.$$ 

This sum can be interpreted as

$$\sum \text{Tr}(F|H_c^i(Kl \otimes Kl^\vee \otimes Kl \otimes Kl^\vee)(-1)^i).$$

By Deligne, $H_c^1$ contributes a second-order term, so the leading term comes from

$$H_c^2 = (V \otimes V^\vee \otimes V \otimes V^\vee)_{\pi_{\text{geom}}^1}(-1) = (V \otimes V^\vee \otimes V \otimes V^\vee)_{\pi_{\text{geom}}^1}.$$ 

We can interpret $V \otimes V^\vee \otimes V \otimes V^\vee = \text{End}(V \otimes V^\vee)$. So we want to compute

$$\dim \text{End}(V \otimes V^\vee)_{\pi_{\text{geom}}^1}.$$ 

We have an irreducible decomposition of $V \otimes V^\vee$ into the direct sum of a 3-dimensional representation and a 1-dimensional representation, so there are indeed two independent $\pi_{\text{geom}}^1$-equivariant endomorphisms.