THE FRIEDLANDER–IWANIEC CHARACTER SUM

ÉTIENNE FOUVRY, EMMANUEL KOWALSKI, AND PHILIPPE MICHEL

In their paper on the exponent of distribution of the ternary divisor function, in which they break the large sieve barrier, Friedlander and Iwaniec [5, p. 329] encounter exponential sums which, for a prime modulus p, are of the form

$$T(\alpha,\beta,\gamma)=\sum_{\substack{x,y,z\in\mathbf{F}_p^\times\\ x\neq-\alpha}}\psi\Bigl(\frac{y}{x}+\frac{z}{x+\alpha}+\frac{\beta}{y}+\frac{\gamma}{z}\Bigr)$$

where $\psi(x)=e(x/p)$ and α, β, γ are integral parameters.

As observed by Bombieri and Birch [5, p. 347] in their Appendix to [5], summing over x first and making a change of variable gives the formula

$$T(\alpha,\beta,\gamma)=pS\Bigl(\frac{\beta}{\alpha},\frac{\gamma}{\alpha}\Bigr)$$

for $\alpha\neq 0$, where

$$S(\alpha,\beta)=\sum_{u\in\mathbf{F}_p^\times-\{-1\}}\mathrm{Kl}_2\Bigl(\frac{\alpha}{u}\Bigr)\,\mathrm{Kl}_2\Bigl(\frac{\beta}{1+u}\Bigr)$$

in terms of classical (normalized) Kloosterman sums

$$\mathrm{Kl}_2(a)=\frac{1}{\sqrt{p}}\sum_{x\in\mathbf{F}_p^\times}e\Bigl(\frac{ax+x^{-1}}{p}\Bigr).$$

Bombieri and Birch show ([1, Th. 1]; note the different normalization):

Theorem 1. *For any p and α, $\beta\in\mathbf{F}_p^\times$, we have*

$$|S(\alpha,\beta)|\ll p^{1/2}$$

where the implied constant is absolute.

As we observed in [3], one can also interpret the sums $S(\alpha;\beta)$ as special cases of the *correlation sums* arising in our work on algebraic twists of modular forms, and we also noted that our more general results and techniques had, as special case, this theorem of Birch and Bombieri.

Very recently, Y. Zhang [12, Lemma 12] made a crucial use of this estimate in his work on bounded gaps between primes. In view of this renewed interest, we spell out the argument of [3] for this special case, and make it more precise. This gives a conceptual proof of the result which is quite short, but which depends on two deep results of Deligne: the general form of the Riemann Hypothesis over finite fields [2], and the construction of Kloosterman sheaves (explained and studied by Katz in [8]).

Proposition 2. *For any prime p and any α, $\beta\in\mathbf{F}_p^\times$, we have*

$$|S(\alpha,\beta)|\leqslant 8p^{1/2}.$$

Date: June 14, 2013, 11:08.

Ph. M. was partially supported by the SNF (grant 200021-137488) and the ERC (Advanced Research Grant 228304). É. F. thanks ETH Zürich, EPF Lausanne and the Institut Universitaire de France for financial support.

We first set up some general notation from [3]. For a prime number p, and for a function $K : \mathbf{F}_p \to \mathbf{C}$ and $v \in \mathbf{Z}/p\mathbf{Z}$, we denote by

$$\hat{K}(v) = \frac{1}{p^{1/2}} \sum_{x \,(\mathrm{mod}\, p)} K(x) e\Big(\frac{vx}{p}\Big)$$

the (unitarily normalized) Fourier transform modulo p of K. Denoting by $\gamma \cdot z$ the action of PGL_2 (or GL_2) on $\mathbf{P}^1$ by homographies, the correlation sums $\mathcal{C}(K;\gamma)$ of K are defined for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbf{F}_p)$ by

$$\mathcal{C}(K;\gamma) = \sum_{\substack{z \in \mathbf{F}_p \\ z \neq -d/c}} \hat{K}(\gamma \cdot z) \overline{\hat{K}(z)}. \tag{1}$$

For $K(x) = e(\bar{x}/p)$ and $K(0) = 0$, we have

$$\hat{K}(z) = \frac{S(z,1;p)}{\sqrt{p}} = \overline{\hat{K}(z)},$$

and hence we deduce that

$$S(\alpha,\beta) = \ \mathcal{C}\Big(K; \begin{pmatrix} \beta & 0 \\ 1 & \alpha \end{pmatrix}\Big).$$

For a given prime p, we fix a prime $\ell \neq p$, and then we fix an isomorphism $\iota : \bar{\mathbf{Q}}_\ell \longrightarrow \mathbf{C}$, which we will use as an identification when considering ℓ-adic numbers as complex numbers. Most often, we will omit it from the notation. For instance, there is an additve character $\psi_\ell : \mathbf{F}_p \longrightarrow \bar{\mathbf{Q}}_\ell^\times$ such that $\iota(\psi_\ell(x)) = \psi(x) = e(x/p)$, and we will just denote it ψ.

Lemma 3 (Deligne)**.** *With notation as above, there exists a lisse geometrically irreducible ℓ-adic sheaf $\mathcal{K}\ell_2$ of rank 2 on the multiplicative group over $\mathbf{F}_p$ such that for $a \in \mathbf{F}_p^\times$, the trace of the geometric Frobenius of $\mathbf{F}_p$ acting on the stalk over a is equal to $-\mathrm{Kl}_2(a)$. The pullback of $\mathcal{K}\ell_2$ to any open dense subset of the multiplicative group is still geometrically irreducible.*

Furthermore, this sheaf is self-dual, tamely ramified at 0 and totally wildly ramified at ∞ with only break $1/2$, hence Swan conductor 1. It is pointwise ι-pure of weight 0.

See [8] for this result, and much more concerning Kloosterman sheaves (including ones parameterizing hyper-Kloosterman sums) in particular [8, Th. 4.1.1].

As a corollary, given α and β in $\mathbf{F}_p^\times$ and an element $\gamma = \begin{pmatrix} \beta & 0 \\ 1 & \alpha \end{pmatrix} \in \mathrm{PGL}_2(\mathbf{F}_p)$, the sheaf

$$\mathcal{F} = \mathcal{K}\ell_2 \otimes \gamma^* \mathcal{K}\ell_2$$

is lisse and pointwise of weight 0 on $U = \mathbf{P}^1 - \{0, -\alpha, \infty\}$, and has trace of Frobenius equal to

$$\mathrm{Kl}_2(x)\,\mathrm{Kl}_2(\gamma \cdot x)$$

for all $x \in U(\mathbf{F}_p)$. Hence $S(\alpha,\beta)$ is the sum of the local traces of of this sheaf.

This sheaf is of rank 4. It is tamely ramified at 0 (because $\mathcal{K}\ell_2$ and $\gamma^*\mathcal{K}\ell_2$ are tame at 0) and totally wildly ramified at $-\alpha$ and ∞, with Swan conductor 2 at each of these singularities (it has unique break $1/2$ with multiplicity 4 at these points).

Because the sheaf is lisse on U, we have

$$H_c^0(U \times \bar{\mathbf{F}}_p, \mathcal{F}) = 0$$

(see,e.g., [2, (1.4.1)b]) and because the two tensor factors are geometrically irreducible on U and have different singularities, so that they are not geometrically isomorphic, we have also

$$H_c^2(U \times \bar{\mathbf{F}}_p, \mathcal{F}) = 0$$

(see [2, (1.4.1)b]).

Because of these two properties of cohomological vanshing, the Grothendieck-Lefschetz trace formula (see, e.g., [8, 2.3.2]) gives

$$S(\alpha,\beta) = -\operatorname{tr}(\mathrm{Fr} \mid H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}))$$

and Deligne's general form of the Riemann Hypothesis [2, Th. 3.3.1] implies that all eigenvalues of the Frobenius acting on $H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F})$ have modulus at most $p^{1/2}$, so that

$$|S(\alpha,\beta)| \leqslant (\dim H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}))p^{1/2}.$$

We conclude by applying:

Lemma 4. *We have*

$$\dim H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) = 8$$

for $\alpha,\ \beta\in\mathbf{F}_p^\times$.

Proof. We have

$$\dim H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) = -\chi_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}),$$

where

$$\chi_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) = \dim H^0_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) - \dim H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) + \dim H^1_c(U\times\bar{\mathbf{F}}_p,\mathcal{F})$$

since the H^0_c and H^2_c cohomology groups vanish. By the Euler-Poincaré characteristic formula (see, e.g, [8, 2.3.1]), we have

$$\chi_c(U\times\bar{\mathbf{F}}_p,\mathcal{F}) = \operatorname{rank}(\mathcal{F})\chi_c(U\times\bar{\mathbf{F}}_p) - \mathrm{Swan}_0(\mathcal{F}) - \mathrm{Swan}_{-\alpha}(\mathcal{F}) - \mathrm{Swan}_\infty(\mathcal{F}).$$

Since $U=\mathbf{P}^1-\{0,-\alpha,\infty\}$, its Euler-Poincaré characteristic us -1, and since we have

$$\mathrm{Swan}_0(\mathcal{F}) = 0, \qquad \mathrm{Swan}_{-\alpha}(\mathcal{F}) = \mathrm{Swan}_\infty(\mathcal{F}) = 2,$$

by the above, we get the result. ☐

Remark 5. In fact, one can show with some more knowledge of Kloosterman sheaves that an estimate of the form $|\mathcal{C}(K;\gamma)| \leqslant Cp^{1/2}$, with C absolute, holds (in that case) for all $\gamma\neq 1$ in $\mathrm{PGL}_2(\mathbf{F}_p)$.

We add some further comments:

(1) The Friedlander-Iwaniec sum is also used by Heath-Brown [6] in his improvement of the exponent of distribution for d_3. As in our own further improvement [4] (for prime moduli) the sum appears more naturally by "dimension reduction" from another exponential sum (see [6, (3.10), p. 36]):

Proposition 6. *For p prime and $a\in\mathbf{F}_p^\times$, let*

$$\mathrm{Kl}_3(a) = \frac{1}{p}\sum_{xyz=a} e\Big(\frac{x+y+z}{p}\Big)$$

be the hyper-Kloosterman sum in two variables. We have

$$\sum_{a\in\mathbf{F}_p^\times} \mathrm{Kl}_3(a)\overline{\mathrm{Kl}_3(\alpha a)}e\Big(\frac{\beta a}{p}\Big) = \sum_{t\neq 0,-\beta} \mathrm{Kl}_2\Big(\frac{1}{t}\Big)\mathrm{Kl}_2\Big(\frac{\alpha}{t+\beta}\Big) - \frac{1}{p^2} = \mathcal{C}\Big(K;\begin{pmatrix}\alpha & 0\\ \beta & 1\end{pmatrix}\Big) - \frac{1}{p^2}.$$

Proof. Expanding the Kloosterman sums and exchanging the sums, the left-hand side is

$$\frac{1}{p^2}\sum_{x,y,u,v\in\mathbf{F}_p^\times}\psi(x+y+u+v)\Big\{\delta\Big(\frac{1}{xy}-\frac{\alpha}{uv}+\beta\Big)-1\Big\}$$

by orthogonality of characters. Introducing a variable $t=(xy)^{-1}=\alpha(uv)^{-1}-\beta$ which is in $\mathbf{F}_p^\times-\{-\beta\}$, and summing over t first, we get

$$\frac{1}{p}\sum_{t\neq 0,-\beta}\sum_{xy=t^{-1}}\psi(x+y)\sum_{uv=\alpha/(t+\beta)}\psi(u+v)-\frac{1}{p^2}$$

which is equal to

$$\sum_{t\neq 0,-\beta}\mathrm{Kl}_2\Big(\frac{1}{t}\Big)\,\mathrm{Kl}_2\Big(\frac{\alpha}{t+\beta}\Big)-\frac{1}{p^2}.$$

□

The left-hand side, when spelled out explicitly, is a five-variable character sum, and the Bombieri-Birch estimate gives square-root cancellation in these terms. In terms of correlation sums, the left-hand side is a correlation sum corresponding to the Fourier transform of $a\mapsto\mathrm{Kl}_3(a)$ and to an upper-triangular matrix. We note that we can also prove the Bombieri-Birch estimate directly using this different form.

(2) The correlation sums $\mathcal{C}(K;\gamma)$ also occur (often for other matrices γ than the ones above) in other papers: one of Pitt [11, Th. 3] and one of Munshi [10, §5.2, p. 8, line -6] (and they are special cases of the sums appearing in [3]).

(3) One can investigate some properties of the Friedlander-Iwaniec sums numerically. For $p=541$, 1151 and 1451, we found for instance that the second moment M_2 is very close to 1, where

$$M_k=\frac{1}{(p-1)^2}\sum_{\alpha,\beta\in\mathbf{F}_p^\times}\Big(\frac{S(\alpha,\beta)}{\sqrt{p}}\Big)^k,$$

while the fourth moment M_4 is close to 3. Because of the Larsen alternative (see [9]; note that $S(\alpha,\beta)$ is real), this suggests that $S(\alpha,\beta)$ is itself the local trace of a lisse sheaf of rank 8 on $\mathbf{G}\times\mathbf{G}$ (which can indeed be proved, at least on a dense subset, using higher-direct images) which is geometrically irreducible and has geometric monodromy group either (special?) orthogonal or symplectic. Thus one may expect an equidistribution statement for these sums according to the traces of random Haar-distributed matrices of elements in USp_8, SO_8 or O_8. We hope to come back to this question.

(4) Here is a slightly different viewpoint on the sum which might clarify the argument we use for readers with analytic number theory background: we can view the coefficients

$$a_1(x)=-\mathrm{Kl}_2(x),\qquad a_2(x)=-\mathrm{Kl}_2(\gamma\cdot x)$$

for $x\in\mathbf{F}_p$ as "Fourier coefficients" (or better Hecke eigenvalues) of a cusp form on a subgroup of $\mathrm{GL}_2(\mathbf{F}_p(T))$, associated to the primes $T-x$ of the ring $\mathbf{F}_p[T]$ (this follows from Drinfeld's work on the Langlands correspondance for GL_2 over function fields; these automorphic forms are cuspidal because the Kloosterman sheaves are geometrically irreducible). Then an estimate of the type

$$\sum_x a_1(x)a_2(x)\ll\sqrt{p}$$

is an incarnation of the Riemann Hypothesis for the Rankin-Selberg L-function associated to these two cusp forms, together with the absence of pole of this L-function, a property which holds because the cusp forms are distinct – essentially for the trivial reason that their conductor (in the sense of set of ramified primes) are not the same... This is therefore in exact analogy with the Prime

Number Theorems for classical cusp forms over $\mathbf{Q}$ (see, e.g.,[7, Th. 5.15, Prop. 5.22] for this type of conditional results); the crucial uniformity in p can then be seen as coming from a uniform conductor estimate.

However, although one could quote results to prove the Birch–Bombieri estimate in this manner, this would be a very perverse approach, as the Riemann Hypothesis for cusp forms (and Rankin-Selberg convolutions) over function fields, which was proved by Lafforgue, proceeds by reducing to Deligne's geometric results anyway...

References

[1] B.J. Birch and E. Bombieri: *On some exponential sums*, Appendix to [5].
[2] P. Deligne: *La conjecture de Weil, II*, Publ. Math. IHÉS 52 (1980), 137–252.
[3] É. Fouvry, E. Kowalski and Ph. Michel: *Algebraic twists of modular forms and Hecke orbits*, preprint (2012).
[4] É. Fouvry, E. Kowalski and Ph. Michel: *On the exponent of distribution of the ternary divisor function*, preprint (2013).
[5] J. Friedlander and H. Iwaniec: *Incomplete Kloosterman sums and a divisor problem*, Annals of Math. 121 (1985), 319–350.
[6] D.R. Heath–Brown: *The divisor function $d_3(n)$ in arithmetic progressions*, Acta Arith. 47 (1986), 29–56.
[7] H. Iwaniec and E. Kowalski: *Analytic number theory*, A.M.S. Coll. Publ. 53 (2004).
[8] N.M. Katz: *Gauss sums, Kloosterman sums and monodromy groups*, Annals of Math. Studies 116, Princeton Univ. Press (1988).
[9] N.M. Katz: *Larsen's alternative, moments and the monodromy of Lefschetz pencils*, in "Contributions to automorphic forms, geometry, and number theory (collection in honor of J. Shalika's 60th birthday)", J. Hopkins Univ. Press (2004), 521–560.
[10] R. Munshi: *Shifted convolution sums for $GL(3)\times GL(2)$*, preprint arXiv:1202.1157v1
[11] N. Pitt: *On shifted convolutions of $\zeta^3(s)$ with automorphic L-functions*, Duke Math. J. 77 (1995), 383–406.
[12] Y. Zhang: *Bounded gaps between primes*, to appear in Annals of Math.

Université Paris Sud, Laboratoire de Mathématique, Campus d'Orsay, 91405 Orsay Cedex, France
E-mail address: etienne.fouvry@math.u-psud.fr

ETH Zürich – D-MATH, Rämistrasse 101, CH-8092 Zürich, Switzerland
E-mail address: kowalski@math.ethz.ch

EPFL/SB/IMB/TAN, Station 8, CH-1015 Lausanne, Switzerland
E-mail address: philippe.michel@epfl.ch