Math 259: Introduction to Analytic Number Theory

An application of Kloosterman sums

As promised, here is the analytic lemma from [Merel 1996]. The algebraic exponential sum that arises naturally here also arises in our investigation of the coefficients of modular forms.

Fix a prime p and a nonzero $c \bmod p$. [More generally we might ask the same question for any integer N and $c \in (\mathbf{Z}/N)^*$; see Exercise 1 below.] Let $I, J \subset (\mathbf{Z}/p)$ be intervals of size $A, B < p$. How many solutions $(x, y) \in I \times J$ are there to $xy \equiv c \bmod p$?

As usual we cannot reasonably hope for a meaningful exact answer, but on probabilistic grounds we expect the number to be roughly AB/p, and can analytically bound the difference between this and the actual number:

Lemma 1. *The number M of solutions $(x, y) \in I \times J$ of $xy \equiv c \bmod p$ is $AB/p + O(p^{1/2} \log^2 p)$.*

Proof: Let $\chi, \psi : (\mathbf{Z}/p) \to \mathbf{C}$ be the characteristic functions of I, J. Then the number of solutions to our equation is

$$M = \sum_{n \in (\mathbf{Z}/p)^*} \chi(n)\psi(cn^{-1}).$$

As it stands, this "formula" for M is just restating the problem. But we may expand χ, ψ in discrete Fourier series:

$$\chi(x) = \sum_{a \bmod p} \hat{\chi}(a) e_p(ax), \qquad \psi(x) = \sum_{b \bmod p} \hat{\psi}(b) e_p(bx).$$

[NB for $t \in (\mathbf{Z}/p)$ the notation $e_p(t)$ now means $e(t/p) = e^{2\pi i t/p}$, **not** $e(pt)$ as before.] So, we have

$$M = \sum_{x \bmod p} \sum\sum_{a,b \bmod p} \hat{\chi}(a)\hat{\psi}(b) e_p(ax + bcx^{-1}) = \sum\sum_{a,b \bmod p} \hat{\chi}(a)\hat{\psi}(b) K_p(a, bc), \quad (1)$$

where for $a, b \bmod p$ the *Kloosterman sum* $K_p(a, b)$ is defined by

$$K_p(a, b) := \sum_{n \in (\mathbf{Z}/p)^*} e_p(an + bn^{-1}).$$

Now clearly $K_p(0, 0) = p - 1$, and almost as clearly $K_p(0, b) = K_p(a, 0) = -1$ if $a, b \neq 0$. The interesting case is $a, b \in (\mathbf{Z}/p)^*$. We now encounter yet another algebraic result that we must cite without proof, again due to Weil [1948]: for a, b nonzero,

$$\boxed{|K_p(a, b)| < 2\sqrt{p}.} \quad (2)$$

[This comes from an interpretation of $K_p(a, b)$ as $\lambda + \bar{\lambda}$ where λ is an eigenvalue of Frobenius for the "Artin-Schreier curve" $Y^p - Y = aX + b/X$, though even

the connection with that curve is nontrivial — see [1948] again, which as usual generalizes to finite fields which need not have prime order.] Putting this into (1) we find

$$|M - \hat{\chi}(0)\hat{\psi}(0)(p-1)| < 2\sqrt{p} \sum_{a \bmod p} |\hat{\chi}(a)| \cdot \sum_{b \bmod p} |\hat{\psi}(b)|.$$

But $\hat{\chi}(0) = A/p$ and $\hat{\psi}(0) = B/p$. For nonzero $a, b \bmod p$ we obtain $\hat{\chi}(a), \hat{\psi}(b)$ as sums of geometric series and find (as in Polya-Vinogradov)

$$p\,|\hat{\chi}(a)| \ll \{a/p\}^{-1}, \qquad p\,|\hat{\psi}(b)| \ll \{b/p\}^{-1}.$$

Thus $\sum_a |\hat{\chi}(a)|, \sum_b |\hat{\psi}(b)| \ll \log p$, and Lemma 1 is proved. □

Corollary. ("Lemme 5" of [Merel 1996]) *If AB is a sufficiently high multiple of $p^{3/2}\log^2 p$ then there are $x \in I$, $y \in J$ such that $xy \equiv c \bmod p$.*

For instance it is enough for A, B to both be sufficiently high multiples of $p^{3/4}\log p$. Presumably $p^{1/2+\epsilon}$ suffices, but as far as I know even p^θ for any $\theta < 3/4$ is a difficult problem. We can, however, remove the log factors from the Corollary:

Lemma 2. *Suppose $I, J \subset (\mathbf{Z}/p)$ are intervals of sizes A, B with $AB \geq 8p^{5/2}/(p-1)$. Then there are $x \in I$, $y \in J$ such that $xy \equiv c \bmod p$.*

Proof: The idea is to replace χ, ψ by functions $f, g : (\mathbf{Z}/p) \to [0,1]$ supported on I, J whose discrete Fourier coefficients decay more rapidly than $p/\{\frac{a}{p}\}$, and sum to $O(1)$ instead of $O(\log p)$. This will yield an estimate on

$$M' := \sum_{n \in (\mathbf{Z}/p)^*} f(n) g(cn^{-1})$$

instead of M, but if there are *no* solutions $(x, y) \in I \times J$ of $xy \equiv c \bmod p$ then M' vanishes as well as M and a contradiction would arise just the same.

Let χ_0 be the characteristic function of an interval of size $A' = \lceil A/2 \rceil$, and let f_0 be the convolution $\chi_0 * \chi_0$. Then f_0 is a function from $(\mathbf{Z}/p)$ to $[0, A']$, supported on an interval of size $\leq A$ centered at the origin, and with nonnegative discrete Fourier coefficients $\hat{f}_0(a)$. Thus

$$\sum_{a \bmod p} |\hat{f}_0(a)| = \sum_{a \bmod p} \hat{f}_0(a) = f(0) = A'.$$

Moreover $\sum_{x \bmod p} f_0(x) = A'^2$. Let f be a translate of f_0/A' supported on I. Then $\sum_{a \bmod p} |\hat{f}(a)| = 1$ and $\sum_{x \bmod p} f_0(x) = A'$. Define ψ_0, g_0, g similarly. Arguing as before, we find that

$$|M' - \frac{p-1}{p^2} A'B'| < 2\sqrt{p}.$$

Thus if $M' = 0$ then $A'B' < 2p^{5/2}/(p-1)$ and $AB \leq 4A'B' < 8p^{5/2}/(p-1)$, Q.E.D. □

Exercises

1. Show that (unless a, b both vanish mod p) the Kloosterman sum $K_p(a,b)$ depends only on ab mod p. In particular $K_p(a,b) \in \mathbf{R}$.

2. Calculate $\sum_{b=1}^{p-1} |K_p(a,b)|^2$ (for any prime p and integer a not divisible by p). In particular, show that (2) cannot be improved to $|K_p(a,b)| = o(\sqrt{p})$.

It is known that in fact for large p there exist a, b such that $p^{-1/2}K_p(a,b)$ is arbitrarily close to θ for each $\theta \in [-2,2]$. This is a consequence of the result that, for each t_1, t_2 with $-2 \le t_1 \le t_2 \le 2$, as $p \to \infty$ the proportion of (a,b) such that $p^{-1/2}K_p(a,b) \in [t_1,t_2]$ approaches $(2/\pi)\int_{t_1}^{t_2} \sqrt{1-\theta^2}\, d\theta$ ("Sato-Tate distribution for Kloosterman sums", see [Katz 1988] and [Adolphson 1989]).

3. For any integer N and any $a, b \in \mathbf{Z}/N\mathbf{Z}$, define

$$K_N(a,b) := \sum_{n \in (\mathbf{Z}/N)^*} e_N(an + bn^{-1}).$$

Prove that if N is squarefree then $K_N(a,b) = \prod_{p|N} K_p(a,b)$. Deduce results analogous to our Lemmas 1,2 for composite moduli. What can you say about $K_{p^r}(a,b)$ for $r > 1$, and $K_N(a,b)$ for general N?

4. Show using only the "Riemann hypothesis" for elliptic curves over $\mathbf{F}_p$ that $K_p(a,b) \ll p^{3/4}$. [Expand $|K_p(a,b)|^2$ and collect like terms. The point is that while the bound is worse than (2), it is still effectively $o(p)$, which suffices for many purposes (including Merel's), while the proof is more elementary in that RH for elliptic curves is easier to prove (and was already done by Hasse in 1936) and the resulting bound on $K_p(a,b)$ is obtained more directly than the one in [Weil 1948].]

5. Let p be an odd prime, and $\chi = (\cdot/p)$ the nontrivial real character mod p. Evaluate the *Salié sum*

$$S_p(a,b) := \sum_{n=1}^{p-1} \chi(n) e_p(an + bn^{-1})$$

in closed form.

As with Gauss sums, there is an analogy between Kloosterman sums and certain definite integrals, in this case the integral $\int_0^\infty \exp(-ax - b/x)\, dx/x$ which gives twice the Bessel function $K_0(2\sqrt{ab})$. The Salié sum is analogous to $\int_0^\infty \exp(-ax - b/x)\, dx/\sqrt{x}$, which involves a Bessel function $K_{1/2}$ of half-integer order and so (unlike the K_ν for $\nu \in \mathbf{Z}$) known in closed form. See for instance [GR 1980, 3.471 9. and 8.468] for the relevant formulas.

References

[Adolphson 1989] Adolphson, A.: On the distribution of angles of Kloosterman sums. *J. reine angew. Math.* **395** (1989), 214–220.

[GR 1980] Gradshteyn, I.S., Ryzhik, I.M.: *Table of Integrals, Series, and Products.* New York: Academic Press 1980. [D 9.80.1 / BASEMENT REFERENCE QA55.G6613]

[Katz 1988] Katz, N.: *Gauss Sums, Kloosterman Sums, and Monodromy Groups.* Princeton, NJ 1988 (#116 in *The Annals of Math. Studies*). [QA246.8.G38 K37]

[Merel 1996] Merel, L.: Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Invent. Math.* **124** (1996), 437–449.

[Weil 1948] Weil, A.: On some exponential sums. Item 1948c (pages 386–389) in his *Collected Papers I.* [O 9.79.1 (I) / QA3.W43].