An Introduction to Kloosterman Sums

Franklyn Wang*

May 10, 2021

Abstract

In this expository note, we develop some of the theory regarding Kloosterman sums, and show how they can be applied to solve certain problems. First, we calculate estimates on the Kloosterman sums for prime p. Then, we show how this can be used to bound the number of solutions to $ab \equiv c \pmod p$ for $a, b \in (\mathbb{Z}/p\mathbb{Z})^*$.

Then, we provide two refinements of the original Kloosterman sums, one for nonprime moduli and one in the form of the Salié sum. We use these results to derive refinements and generalizations to the $ab \equiv c \pmod p$ problem.

1 Introduction

In analytic number theory, and in math writ large, there are often questions whose answers can be "guessed" in one way or another. Sometimes these guesses are right, and sometimes they are wrong. Often by using probabilistic heuristics, we can obtain guesses to the answers of certain questions. Yet proving that these guesses are correct can often be much, much harder – arguably the Riemann Hypothesis is a perfect example of this, in the sense that it states the prime counting function $\pi(x)$ *should be* very close to $\text{li}(x)$.

In this note, we will first describe the rich theory of Kloosterman sums, and derive some properties. Then, we will show how they can be used to solve a particularly interesting problem whose answer should be "guessable". To conclude, we will give a generalization of the theory of Kloosterman sums to Salié sums, which allow further work on the aforementioned guessable problem.

2 Kloosterman sums

The *Kloosterman sum* is defined as

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

where $e_N(x) = e^{2\pi i x/N}$, where n^{-1} is the inverse in $\mathbb{Z}/p\mathbb{Z}$.

When $a, b = 0$, we have that $K_p(a,b) = p-1$, and when exactly one of a, b is zero we have $K_p(a,b) = -1$. It remains to handle $K_p(a,b)$ for $a, b \in (\mathbb{Z}/p\mathbb{Z})^*$.

*Email: franklyn_wang@college.harvard.edu.

First, to slightly simplify the expression, note that $n \mapsto bn$ is a bijection on $(\mathbb{Z}/p\mathbb{Z})^*$. We thus substitute to obtain

$$\begin{aligned} K_p(a,b) &= \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} e_p(an+bn^{-1}) \\ &= \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} e_p(a(bn)+b(bn)^{-1}) \\ &= \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} e_p(abn+n^{-1}) \\ &= K_p(ab,1). \end{aligned}$$

Thus to evaluate all Kloosterman sums, it suffices to evaluate $K_p(c,1)$ for all c. Also, note that substituting $-n$ in for n (again noting that it's a bijection) gives us $K_p(a,b) = K_p(-a,-b)$. Now,

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

which implies $K_p(a,b) \in \mathbb{R}$ as desired.

Evaluating individual Kloosterman sums is a rather daunting task. Thus, we will try instead to calculate *aggregate properties* of Kloosterman sums, which will turn out to be much more tractable.

We write a formula for the sum of all nth powers of Kloosterman sums. Define $S_n = \sum_{b=1}^{p-1} K_p(1,b)^n$. Then, we first have

Theorem 2.1. *For each integer $n \geq 1$,*

$$S_n = \frac{p^2}{p-1} N_n(p) - (p-1)^{n-1} - 2(-1)^n$$

where $N_n(p)$ is the number of solutions to $\sum_{i=1}^n x_i = \sum_{i=1}^n 1/x_i = 0$ for $x_i \in (\mathbb{Z}/p\mathbb{Z})^*$.

Proof. We work in $\mathbb{F}_p$. Define H to be the hyperplane corresponding to $a_1 + a_2 + \ldots + a_n = 0$, and let H' be the hyperplane corresponding to $a_1^{-1} + a_2^{-1} + \ldots + a_n^{-1} = 0$. Furthermore, let $\ell_n = (k_1, k_2, \ldots k_n)$ represent n-tuples in $((\mathbb{Z}/p\mathbb{Z})^*)^n$.

First, we can write

$$\begin{aligned} &\sum_{1\leq a\leq p-1} K_p(1,a)^n \\ &= \frac{1}{p-1} \sum_{1\leq a,b\leq p-1} K_p(a,b)^n \\ &= \frac{1}{p-1} \sum_{1\leq a,b\leq p-1} \sum_{k\in\ell_n} e_p(a(k_1+k_2+\ldots+k_n) + b(k_1^{-1}+k_2^{-1}+\ldots+k_n^{-1})) \\ &= \frac{1}{p-1} \sum_{k\in\ell_n} \sum_{1\leq a\leq p-1} e_p(a(k_1+k_2+\ldots+k_n)) \sum_{1\leq b\leq p-1} e_p(b(k_1^{-1}+k_2^{-1}+\ldots+k_n^{-1})) \\ &= \frac{1}{p-1} \sum_{k\in\ell_n} (p1_{k\in H}-1)(p1_{k\in H'}-1) \\ &= \frac{p^2}{p-1} |\{k\in\ell_n : k\in H\cap H'\}| - \frac{2p}{p-1} |\{k\in\ell_n : k\in H\}| + (p-1)^{n-1} \\ &= \frac{p^2}{p-1} N_n(p) - \frac{2}{p-1} \sum_{k\in\ell_n} p1_{k\in H} + (p-1)^{n-1} \end{aligned} \tag{1}$$

where we've used the fact that the number of solutions to $k_1 + k_2 + \ldots + k_n \equiv 0 \pmod{p}$ in ℓ_n is equal to the number of solutions to $k_1^{-1} + k_2^{-1} + \ldots + k_n^{-1} \equiv 0$ in ℓ_n (which can be seen by a straightforward bijection).

It remains to find $\sum_{k \in \ell_n} p 1_{k \in H}$. To find this, we use a similar trick.

$$\begin{aligned}
\sum_{k \in \ell_n} p 1_{k \in H} &= \sum_{k_i \in (\mathbb{Z}/p\mathbb{Z})^*} \sum_{\ell=0}^{p-1} e_p(\ell(k_1 + k_2 + \ldots + k_n)) \\
&= \sum_{\ell=0}^{p-1} \sum_{k \in \ell_n} e(\ell(k_1 + k_2 + \ldots + k_n)) \\
&= \sum_{\ell=0}^{p-1} \left(\sum_{k \in \ell_n} e_p(\ell(k_1)) \right)^n \\
&= (p-1)^n + (p-1)(-1)^n
\end{aligned}$$

and now substituting this into Equation (1) yields

$$S_n = \frac{p^2}{p-1} N_n(p) - (p-1)^{n-1} - 2(-1)^n$$

and we may conclude. □

Now, we calculate the values of $N_n(p)$ for various values of n.

Lemma 2.2. *$N_1(p) = 0$, $N_2(p) = p - 1$, $N_3(p) = (p-1)(1 + (-3/p))$, $N_4(p) = (p-1)(3p-6)$. Correspondingly, we have $S_1 = 1$, $S_2 = p^2 - p - 1$, $S_3 = p^2\chi_3(p) + 2p + 1$, and $S_4 = 2p^3 - 3p^2 - 3p - 1$.*

Proof. Note that $N_1(p) = 0$, $N_2(p) = p - 1$ by direct arguments.

Finding $N_3(p)$ To calculate $N_3(p)$, we note that this is equivalent to finding solutions to $x^{-1} + y^{-1} = (x+y)^{-1}$, which after clearing denominators is equivalent to $x^2 + xy + y^2 = 0$ which in turn is equivalent to $(2x+y)^2 + 3y^2 = 0$ in $(\mathbb{Z}/p\mathbb{Z})^2$, excluding solutions where $x = 0$, $y = 0$, or $x + y = 0$. Note that if $y \neq 0$, any solution to $x^2 + xy + y^2 = 0$ must satisfy $x \neq 0$ and $x + y \neq 0$, so it suffices to find the number of solutions when $y \neq 0$.

When $y \neq 0$, the number of solutions for x is equal to $1 + (-3/p)$. This gives us $N_3(p) = (p-1)(1 + (-3/p))$, and

$$(-3/p) = (-1/p)(3/p) = (p/3)(-1/p)^2 = (p/3) = \chi_3(p)$$

from quadratic reciprocity. Thus, we get

$$S_3 = p^2(1 + \chi_3(p)) - (p-1)^2 + 2 = p^2\chi_3(p) + 2p + 1.$$

Finding $N_4(p)$ To calculate $N_4(p)$, we note (by substituting $a_3 \mapsto -a_3$, $a_4 \mapsto -a_4$) that it's equivalent to finding the number of solutions to $a_1 + a_2 \equiv a_3 + a_4 \pmod{p}$ and $a_1^{-1} + a_2^{-1} \equiv a_3^{-1} + a_4^{-1} \pmod{p}$. Note that it suffices to calculate n_{xy}, the number of solutions to $a_1 + a_2 = x, a_1^{-1} + a_2^{-1} = y$, because then the answer will simply be $\sum_{x,y} n_{xy}^2$.

Note that $n_{0\ell} = p - 1$ if $\ell = 0$ and $n_{0\ell} = 0$ otherwise. Now, we find n_{xy} for $x \neq 0$. Note that $n_{x0} = 0$, so it suffices to find n_{xy} for $x, y \neq 0$. Now, we can see that $a_1^{-1} + a_2^{-1} = (a_1 a_2)^{-1}(a_1 + a_2)$,

so it suffices to determine the number of values attained by $a_1 a_2 = a_1(x - a_1)$. It will takes on the value of $x^2/4$ exactly once (when $a_1 = x/2$) and takes on all other values exactly twice (at most twice because a quadratic has at most two roots, and at least twice because plugging in a_1 and $x - a_1$ yield the same result.

Thus, for each x, $n_{xy} = 2$ for precisely $(p-3)/2$ values of y, $n_{xy} = 1$ for one value of y, and $n_{xy} = 0$ for all other values of y.

The above discussion yields

$$\sum_{x,y} n_{xy}^2 = (p-1)^2 + (p-1)(1^2 + (p-3)/2 \cdot 2^2) = (p-1)^2 + (p-1)(2p-5) = (p-1)(3p-6).$$

Finally, using Theorem 2.1 we have

$$S_4 = (3p-6)p^2 - (p-1)^3 - 2 = 2p^3 - 3p^2 + 3p - 1,$$

and we're done.

□

It turns out that finding this value for higher powers like $n = 5, 6 \ldots$ and so on is much more difficult and requires nonelementary functions. Nonetheless, note that the $n = 4$ case alone is enough to conclude that $|K_p(a,b)| \leq 2^{1/4} p^{3/4}$, which is already enough for many applications. It turns out that a far stronger bound is possible – namely

Theorem 2.3 ([Wei48]). *$|K_p(a,b)| \leq 2\sqrt{p}$ for all p.*

– but the proof is beyond the scope of the paper, and we will not cover it here.

Our method for pinning down the $K_{a,b}$ relied on using the moments of the distribution. It's a well known that the moments of a distribution generally tend to yield information on the distribution itself. An analagous idea is used in the celebrated *method of moments* from economics and statistics. This begs the question – can we calculate the asymptotic distribution of $\{K_p(1,b)\}_{b=1}^{p-1}$? The answer somewhat miraculously turns out to be yes, as shown in [Kat88].

Theorem 2.4 (Sato-Tate distribution). *For each t_1, t_2 with $-2 \leq t_1 \leq t_2 \leq 2$, as $p \to \infty$ the proportion of (a,b) such that $p^{-1/2} K_p(a,b) \in [t_1, t_2]$ approaches $(2\pi)^{-1} \int_{t_1}^{t_2} \sqrt{4 - \theta^2} \mathrm{d}\theta$.*

While we will not be able to give a complete proof of Theorem 2.4, we can calculate the moments of the Sato-Tate distribution, and show that they are consistent with the values implied by Lemma 2.2.

First, we give the moments of the Sato-Tate distribution.

Lemma 2.5 (Sato-Tate Moments). *Let Z be a random variable distributed as the Sato-Tate distribution, so the probability that $Z \in [t_1, t_2]$ is $(2\pi)^{-1} \int_{t_1}^{t_2} \sqrt{4 - \theta^2} \mathrm{d}\theta$. Then, for $n \geq 0$*

$$\mathbb{E}[Z^n] = \begin{cases} 0 & \text{if } n = 2k+1, k \in \mathbb{Z} \\ C_k & \text{if } n = 2k, k \in \mathbb{Z} \end{cases}$$

where $C_k = \frac{1}{k+1} \binom{2k}{k}$ is the kth Catalan number.

Proof. The odd moments of Z are clearly 0. To find the even moments, note that if we let (X, Y) be distributed as a random point chosen inside of a disk centered at the origin with radius 2, we have $Z \sim X$. With polar coordinates, we can write X as $R\cos(\theta)$, so $R \sim 2\text{Beta}(2,1) \sim 2U^{1/2}$, and $\theta \sim \text{Unif}[0, 2\pi]$. Then, since X is distributed according to the semicircle distribution, note that

$$\begin{aligned}
\mathbb{E}[X^{2n}] &= \mathbb{E}[R^{2n}\cos(\theta)^{2n}] \\
&= 2^{2n}\mathbb{E}[U^n]\mathbb{E}[\cos(\theta)^{2n}] \\
&= \frac{2^{2n}}{n+1}\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^{2n}\mathrm{d}\theta \\
&= \frac{1}{n+1}\frac{1}{2\pi}\int_0^{2\pi}\sum_{k=0}^{2n}\binom{2n}{k}e^{(2k-2n)i\theta}\mathrm{d}\theta \\
&= \frac{1}{n+1}\frac{1}{2\pi}\sum_{k=0}^{2n}\int_0^{2\pi}\binom{2n}{k}e^{(2k-2n)i\theta}\mathrm{d}\theta \\
&= \frac{1}{n+1}\binom{2n}{n}
\end{aligned} \tag{2}$$

which is the nth Catalan number, as each term in the sum (except $k = n$) vanishes. Another method of evaluating Equation (2) is by expressing it as a Beta Integral. □

Now we check that Lemma 2.2 is consistent with Theorem 2.4.

Formally, one can state Theorem 2.4 as saying

$$\frac{1}{\sqrt{p}}\text{DUnif}\{K_p(1,b)|1 \le b \le p-1\} \xrightarrow{\mathcal{L}} S,$$

where S follows the Sato-Tate semicircle distribution.

Note that if this is true, the moments of the LHS and RHS must be equal as $p \to \infty$. And now we can check this for the lower moments. Observe that the moments of the LHS are

$$M_{n,p} := \frac{1}{p^{n/2+1}}\sum_{b=1}^{p-1}K_p(1,b)^n = \frac{S_n}{p^{n/2+1}}.$$

For $n = 1$, we have $M_{n,p} \to 0$. For $n = 2$, we have $M_{n,p} \to 1$. For $n = 3$, we have $M_{n,p} \to 0$. And for $n = 4$, we have $M_{n,p} \to 2$. These are consistent with the moments obtained from Lemma 2.5.

3 A "guessable" problem

We will now consider a problem whose answer appears to be "guessable".

Letting I, J be intervals in $\mathbb{R}/p\mathbb{Z}$, find the number of a, b so that $ab = c \pmod p$ and $a \in I, b \in J$.

Of course the answer to this question is "obvious". It should just be $|I||J|/p$. Is this right? This relationship superficially resembles certain edge concentration results in expander graphs:

Theorem 3.1 ([Vad12], Lemma 4.15)**.** *Let G be a D-regular, N-vertex undirected graph with spectral expansion $1-\lambda$. Then for all sets of vertices I, J of densities $\alpha = |I|/N$ and $\beta = |J|/N$, we have*

$$\frac{1}{N}\left|\frac{e(I,J)}{D} - \frac{|I||J|}{N}\right| = \left|\frac{e(I,J)}{N\cdot D} - \frac{|I||J|}{N^2}\right| = \left|\frac{e(I,J)}{N\cdot D} - \alpha\beta\right| \le \lambda\sqrt{\alpha\beta}.$$

Proof. The key idea in the proof is to write χ_S and χ_T to be the vectors corresponding to characteristic functions of the sets S and T respectively.

Now it suffices to look at (where A is the adjacency matrix of G)

$$\begin{aligned}\left|\frac{e(S,T)}{N\cdot D}-\alpha\beta\right| &= \left|\frac{\chi_S^\top A\chi_T}{N\cdot D}-\alpha\beta\right| \\ &= \left|\frac{(\alpha 1+\chi_S^\perp)^\top A(\beta 1+\chi_T^\perp)}{N\cdot D}-\alpha\beta\right| \\ &= \left|\frac{(\chi_S^\perp A\chi_T^\perp)}{N\cdot D}\right| \\ &\le \lambda\sqrt{\alpha\beta}\end{aligned}$$

as $|\chi_S^\perp| = \sqrt{N\alpha(1-\alpha)}$ and $|\chi_T^\perp| = \sqrt{N\beta(1-\beta)}$. and we may conclude. □

The key idea in this proof was looking at χ_S and χ_T in an orthogonal eigenbasis, and using the fact that all the eigenvalues of A are quite small.

However, it's clear that such a universal approach cannot work for our problem, because there *are* large I, J for which no such $ab \equiv c \pmod p$ – for example, take I to have size $p/2$, and let J contain every number *not of the form* c/i for $i \in I$. Thus, such solutions will not easily generalize to this setting.

However, the first step is still similar, as we can decompose the characteristic functions of I and J in a Fourier basis. The key difference, and what allows us to refine our bounds, is that by virtue of being intervals, we can calculate the Fourier coefficients explicitly – and they will be small. Of course, this means that any characteristic functions with small Fourier coefficients will also work, a fact we take to its logical conclusion in Theorem 3.3.

Let

$$U(N) = \max_{(a,b)\neq(0,0)} |K_N(a,b)|.$$

Then the results from the previous part give us that $U(p) < 2\sqrt{p}$ (Theorem 2.3) and $U(p) < 2^{1/4}p^{3/4}$ (Lemma 2.2). We state our results in terms of U, in order to show how bounds on U naturally lead to bounds in the Kloosterman sums. This is done to make it precise how bounds on U directly lead to better bounds on various problems we seek to solve.

Theorem 3.2. *The number M of solutions $(x,y) \in I \times J$ to $xy \equiv c \bmod p$ is $AB/p + O(U(p)\log^2 p)$.*

Proof. Letting χ, ψ be the characteristic functions of I, J, the number of solutions to our equation is given by

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

We can now decompose χ, ψ into the Fourier basis, so

$$\chi(x) = \sum_{a \bmod p} \hat{\chi}(a)e_p(ax) \text{ and } \psi(x) = \sum_{b \bmod p} \hat{\psi}(b)e_p(bx)$$

We now have that

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

Now, note that we can bound

$$|M - \hat{\chi}(0)\hat{\psi}(0)(p-1)| \le \max_{(a,b)\neq(0,0)} |K_p(a,b)| \sum_{a,b \pmod p} |\hat{\chi}(a)||\hat{\psi}(b)| = U(p) \sum_{a,b \pmod p} |\hat{\chi}(a)||\hat{\psi}(b)|$$

We now bound each element of $\hat{\chi}(a)$. For $a = 0$, we have that $|\hat{\chi}(a)| \le 1$. For $a \neq 0$, note that we have

$$|\hat{\chi}(a)| = \left|\frac{1}{p}\sum_{x\in I} e^{2\pi a x i/p}\right| = \left|\frac{1}{p}\frac{e^{2\pi i a(I_{\max}+1)/p} - e^{2\pi i a I_{\min}/p}}{e^{2\pi i a/p}-1}\right| \le \frac{2}{p}\left|\frac{1}{e^{2\pi i a/p}-1}\right| = \frac{1}{p\sin(\pi a/p)} \ll \frac{1}{p\|a/p\|}$$

where $\|x\|$ is the absolute difference between x and the nearest integer.

We can now bound $\sum_{a,b \pmod p} |\hat{\chi}(a)||\hat{\chi}(b)| \ll \log^2 p$ by a simple calculation and we may conclude.

□

Now a direct consequence of this lemma is

Corollary 3.2.1. *If AB is a sufficiently high multiple of $pU(p)\log^2 p$ then there are $x \in I, y \in J$ such that $xy = c \pmod p$.*

With slightly more work, we can remove the log-factors.

Theorem 3.3. *Suppose $I, J \subseteq \mathbb{Z}/p\mathbb{Z}$ are intervals of sizes A, B with $AB \ge 4p^2U(p)/(p-1)$. Then there are $x \in I, y \in J$ such that $xy \equiv c \pmod p$.*

Proof. The key idea is to replace χ, ψ by functions f, g supported on I, J whose Fourier coefficients will be altogether smaller. This yields an estimate on $M' = \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} f(n)g(cn^{-1})$ instead of M, but this is okay as if there are *no* solutions $(x,y) \in I \times J$ of $xy \equiv c \bmod p$ then M' vanishes.

Let χ_0 be the characteristic function for any interval of width $A' = \lceil A/2 \rceil$, and let f_0 be the convolution $(\chi_0 * \chi_0)/A'$. This function is supported on $[-A', A']$. Then because f_0 is a convolution of a function with itself, all of its Fourier coefficients are positive (as self-convolution in the time basis is squaring in the frequency basis). Finally, define f to be a translate of f_0 which is supported on I. Define g similarly for the interval B.

Now, we can write

$$\sum_{a \bmod p} |\hat{f}(a)| = \sum_{a \bmod p} |\hat{f}_0(a)| = \sum_{a \bmod p} \hat{f}_0(a) = f_0(0) = 1$$

Thus, by the same arguments as on the previous part we have

$$|M' - \hat{f}(0)\hat{g}(0)(p-1)| \le U(p) \sum_{a,b \bmod p} |\hat{f}(a)||\hat{g}(b)| = U(p)$$

so that we have (since $\hat{f}(0) = A'/p$ and $\hat{g}(0) = B'/p$)

$$\left|M' - \frac{p-1}{p^2}A'B'\right| < U(p)$$

and thus if $M' = 0$ then $A'B' < p^2U(p)/(p-1)$ and $AB \le 4A'B' < 4p^2U(p)/(p-1)$ as desired. □

Note that even using the strongest bound we have for $U(p)$ ($2\sqrt{p}$) only gives a $p^{3/2}$ bound in Theorem 3.3. It is conjectured that we can strengthen $3/2$ to $1+\epsilon$, but it is not known definitively whether $3/2$ can even be improved to $3/2 - \epsilon$. Compare this to the much easier lemma that holds when instead of needing $ab \equiv c \pmod p$, we need $ab^{-1} \equiv c \pmod p$.

Lemma 3.4 ([Thu02])**.** *For any prime p and integer $1 \le a \le p$, there exists $|x| < \lceil\sqrt{p}\rceil$, $0 < y < \lceil\sqrt{p}\rceil$ so that $ay \equiv x \pmod p$.*

4 Extensions

In this section, we consider two different extensions: one of Kloosterman sums to nonprime N, and one to the Salié sums, which appear to be more complicated than Kloosterman sums but in actuality are much simpler.

4.1 Generalized Sums

4.1.1 Nonprime bases

Prime Powers Kloosterman sums can sometimes be evaluated for prime powers $N = p^k$, but the analysis can be quite complicated, and we do not go into more details here.

General Composite Numbers In many multiplicative number theory problems, we can solve the problem on general composite numbers by breaking it up into many prime powers, and then by applying the Chinese Remainder theorem. This is precisely what we will use here.

To evaluate

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

where $N = p_1^{e_1} p_2^{e_2} \dots p_z^{e_z}$, note first that it's far from clear how to even enumerate all members of $(\mathbb{Z}/N\mathbb{Z})^*$! One way to organize the sum would be to find a function $g(c_1, c_2, \dots c_z)$ (where $c_i \in (\mathbb{Z}/p_i^{e_i}\mathbb{Z})$), which gives the number $v \in (\mathbb{Z}/N\mathbb{Z})^*$ so that $v \equiv c_i \pmod{p_i}$. Then, in theory we would just have to sum

$$\sum_{c_i \in \left(\mathbb{Z}/p_i^{e_i}\mathbb{Z}\right)^*} e_N(ag(c_1, c_2, \dots, c_z) + bg(c_1^{-1}, c_2^{-1}, \dots, c_z^{-1})).$$

However, in its current form we are no closer to a solution. It will turn out, however, that g is a linear function! Note that we can write

$$g(c_1, c_2, \dots, c_z) = \sum_i c_i \frac{N d_i}{p_i^{e_i}},$$

where d_i is chosen so $d_i(N/p_i^{e_i}) \equiv 1 \pmod{p_i}$. Now we can simply write

$$\begin{aligned}
K_N(a,b) &= \sum_{n \in (\mathbb{Z}/N\mathbb{Z})^*} e_N(an + bn^{-1}) \\
&= \sum_{c_i \in (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*} e_N\left(\sum_i a c_i \frac{N d_i}{p_i^{e_i}} + \sum_i b c_i^{-1} \frac{N d_i}{p_i^{e_i}}\right) \\
&= \prod_{i=1}^{z} \left(\sum_{c_i \in (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*} e_N\left(a c_i \frac{N d_i}{p_i^{e_i}} + b c_i^{-1} \frac{N d_i}{p_i^{e_i}}\right)\right) \\
&= \prod_{i=1}^{z} \left(\sum_{c_i \in (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^*} e_{p_i^{e_i}}\left(a c_i d_i + b c_i^{-1} d_i\right)\right) \\
&= \prod_{i=1}^{z} K_{p_i^{e_i}}(a d_i, b d_i)
\end{aligned}$$

4.1.2 Salié Sum

The Salié sum is defined as $S'_p(a,b) = \sum_{n=1}^{p-1} \chi(n) e_p(an + bn^{-1})$ where χ is the nontrivial real character modulo p. While the Salié sum appears to be more difficult to evaluate than the Kloosterman sum, quite the opposite is true!

Theorem 4.1 (Salié Sum)**.** *If a, b are relatively prime to p, then the Salié sum*

$$S'_p(a,b) = g(a;p) \sum_{y^2 \equiv 4ab} e_p(an)$$

where $g(a;p) = \sum_{n=0}^{p-1} \chi(n) e_p(an)$ is a Gauss sum.

Proof. Note that $S'_p(a,b) = 0$ when ab is a QNR (Quadratic Non-Residue). To see this, note that if ab is a QNR, so is b/a. Then, we have (via the bijection $n \mapsto b/an$)

$$S'_p(a,b) = \sum_{n=1}^{p-1} \chi(n) e_p(an+bn^{-1}) = \sum_{n=1}^{p-1} \chi(b/an) e_p(a(b/an) + b(b/an)^{-1}) = -\sum_{n=1}^{p-1} \chi(n) e_p(an+bn^{-1})$$

implying that when ab is a QNR, the Salié sum is zero, as the equation would imply.

Now suppose that ab is a quadratic residue, so $ab = u^2$. We can write

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

We reduce this to the case where $a = b = u$. Note that we can calculate (via the bijection $n \mapsto un/a$)

$$S'_p(a,b) = \sum_{n=1}^{p-1} \chi(n) e_p(an + bn^{-1}) = \sum_{n=1}^{p-1} \chi(un/a) e_p(un + un^{-1}) = \chi(u/a) S'_p(u,u)$$

Now, we find $S'_p(u,u)$.

Lemma 4.2.

$$S'_p(u,u) = \sum_{n=1}^{p-1} \chi(n) e_p(un+un^{-1}) = g(u;p)(e_p(2u)+e_p(-2u)) = \left(\sum_{n=1}^{p-1} \chi(n) e_p(un) \right) (e_p(2u)+e_p(-2u))$$

Proof. To complete this proof, we compare the coefficients of $e_p(au)$ for each $a \in 0, 1, \ldots p-1$ (in other words, we prove that the Fourier coefficients are equal). On the RHS, we obtain a coefficient of $\chi(a-2) + \chi(a+2)$. On the LHS, we get

$$\sum_{k, k+k^{-1}=a} \chi(k)$$

It now suffices to show that these are always equal, which we do via casework on $\chi(a+2) + \chi(a-2)$.

Claim.

$$\chi(a-2) + \chi(a+2) = \sum_{k, k+k^{-1}=a} \chi(k).$$

Proof. If a^2-4 is not a quadratic residue, then the RHS is zero by the quadratic formula. Likewise, since $\chi(a-2)\chi(a+2)=-1$, the LHS will be zero.

If a^2-4 is a quadratic residue, then $a-2=bu^2, a+2=bv^2$, note that the solutions for k are $(-a\pm\sqrt{a^2-4})/2 = b((u-v)/2)^2, b((u+v)/2)^2$. Then, all we need to show is

$$\chi(a-2)+\chi(a+2)=\chi(bu^2)+\chi(bv^2)=2\chi(b)=\chi(b((u-v)/2)^2)+\chi(b((u+v)/2)^2)=\sum_{k+k^{-1}=a}\chi(k).$$

and we may conclude.

□

Now, note that

$$S'_p(a,b)=\chi(u/a)S'_p(u,u)=\chi(u/a)g(u;p)(e_p(2u)+e_p(-2u))=g(a;p)(e_p(2u)+e_p(-2u))$$

and we may conclude. □

Since Gauss sums have absolute value $\sqrt{p}$, and the $a=0, b\neq 0$ and $a\neq 0, b=0$ reduce to Gauss sums, we have.

Corollary 4.2.1. *Any Salié sum $K'_p(a,b)$ when $(a,b)\neq(0,0)$ has absolute value at most $2\sqrt{p}$.*

4.2 Generalizations of Section 3

In this section, we generalize the results of section 3 to nonprime bases.

4.2.1 Nonprime Bases

Straightforwardly applying the results of section 3, we obtain the following generalizations of Theorem 3.2 and Theorem 3.3. Recall that $U(N)$ is the bound on all Kloosterman sums of the form $K_N(a,b)$ where $(a,b)\neq(0,0)$.

Theorem 4.3. *The number M of solutions $(x,y)\in I\times J$ to $xy\equiv c \bmod p$ is $AB/N+U(N)\log^2(N)$*

Theorem 4.4. *Suppose $I,J\subseteq\mathbb{R}/N\mathbb{Z}$ are intervals of sizes A,B with $AB\geq\frac{4N^2U(N)}{N-1}$. Then there are $x\in I, y\in J$ such that $xy\equiv c \pmod N$.*

4.3 Salié Sums

Now we show how Salié sums can be used to refine Theorem 3.2 and Theorem 3.3.

Theorem 4.5. *The number of solutions (a,b) so that $ab\equiv c \bmod p$ and $\chi(a)=1$ minus the number of solutions (a,b) so that $ab\equiv c \bmod p$ and $\chi(a)=-1$ has absolute value at most $O(\sqrt{p}\log^2(p))$.*

Proof. We bound the expression given by (where f_I and f_J are the characteristic functions of I,J respectively)

$$\begin{aligned}
M &= \#\{a,b \text{ s.t. } \chi(a)=1, ab\equiv c \bmod p\}-\#\{a,b \text{ s.t. } \chi(a)=-1, ab\equiv c \bmod p\}\\
&= \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*}\chi(n)f_I(n)f_J(cn^{-1})\\
&= \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*}\chi(n)\sum_{a \bmod p}\hat{f}_I(a)e_p(an)\sum_{b \bmod p}\hat{f}_J(b)e_p(bcn^{-1})\\
&= \sum_{a,b \bmod p}\hat{f}_I(a)\hat{f}_J(b)\sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*}\chi(n)e_p(an+bcn^{-1})
\end{aligned}$$

Then by Corollary 4.2.1,

$$|M - \hat{f}_I(0)\hat{f}_J(0)(0)| \leq \max_{(a,b)\neq(0,0)} |K'_p(a,b)| \sum_{a,b \bmod p} |\hat{f}_I(a)||\hat{g}_J(a)|$$
$$\leq O(\sqrt{p}\log^2(p)).$$

□

Theorem 4.6. *Suppose $I, J \subseteq \mathbb{R}/p\mathbb{Z}$ are intervals of sizes A, B with*

$$AB \geq \frac{4p^2U(p)}{p-1} + \frac{8p^2\sqrt{p}}{p-1}.$$

Then there are $x \in I, y \in J$ such that $xy \equiv c \pmod p$ and $\chi(x) = 1$ (respectively $\chi(x) = -1$).

Proof. The idea is to look at the expression $M' = \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*}(1+\chi(n))f(n)g(cn^{-1})$. If we can bound M' away from zero, this implies the result. We break this expression into

$$S_1 + S_2, \text{ where } S_1 := \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} f(n)g(cn^{-1}) \text{ and } S_2 := \sum_{n\in(\mathbb{Z}/p\mathbb{Z})^*} \chi(n)f(n)g(cn^{-1})$$

Note that for the choice of functions f, g used in Theorem 3.3, we have

$$S_1 \geq \frac{p-1A'B'}{p^2} - U(p)$$

and

$$S_2 \geq -2\sqrt{p}$$

so $S_1 + S_2 \geq \frac{p-1A'B'}{p^2} - U(p) - 2\sqrt{p} > 0$ and we may conclude. □

References

[Kat88] Nicholas M Katz. *Gauss sums, Kloosterman sums, and monodromy groups*. Number 116. Princeton university press, 1988.

[Thu02] Axel Thue. Et par autydninger til en taltheoretisk methode. *Kra. Vidensk. Selsk. Forh.*, 7:57–75, 1902.

[Vad12] Salil P. Vadhan. Pseudorandomness. *Found. Trends Theor. Comput. Sci.*, 7(1-3):1–336, 2012.

[Wei48] André Weil. On some exponential sums. *Proceedings of the National Academy of Sciences of the United States of America*, 34(5):204, 1948.