ON THE DISTRIBUTION OF KLOOSTERMAN SUMS

IGOR E. SHPARLINSKI

(Communicated by Wen-Ching Winnie Li)

Abstract. For a prime \( p \), we consider Kloosterman sums
\[
K_p(a) = \sum_{x \in \mathbb{F}_p^*} \exp(2\pi i (x + ax^{-1})/p), \quad a \in \mathbb{F}_p^*,
\]
over a finite field of \( p \) elements. It is well known that due to results of Deligne, Katz and Sarnak, the distribution of the sums \( K_p(a) \) when \( a \) runs through \( \mathbb{F}_p^* \) is in accordance with the Sato–Tate conjecture. Here we show that the same holds where \( a \) runs through the sums \( a = u + v \) for \( u \in \mathcal{U}, \ v \in \mathcal{V} \) for any two sufficiently large sets \( \mathcal{U}, \mathcal{V} \subseteq \mathbb{F}_p^* \).

We also improve a recent bound on the nonlinearity of a Boolean function associated with the sequence of signs of Kloosterman sums.

1. Introduction

For a prime \( p \) we use \( \mathbb{F}_p \) to denote the finite field of \( p \) elements. For \( a \in \mathbb{F}_p^* \) we consider the Kloosterman sum
\[
K_p(a) = \sum_{x \in \mathbb{F}_p^*} e_p(x + ax^{-1}),
\]
where
\[
e_p(z) = \exp(2\pi iz/p)
\]
(we identify \( \mathbb{F}_p \) with the set \( \{0, 1, \ldots, p-1\} \)). Since for the complex conjugated sum we have
\[
\overline{K_p(a)} = \sum_{x \in \mathbb{F}_p^*} e_p(-x - ax^{-1}) = K_p(a),
\]
the values of \( K_p(a) \) are real.

According to the Weil bound, see [15],
\[
|K_p(a)| \leq 2\sqrt{p}, \quad a \in \mathbb{F}_p^*.
\]
Therefore, we can define the angles \( \psi_p(a) \) by the relations
\[
K_p(a) = 2\sqrt{p} \cos \psi_p(a) \quad \text{and} \quad 0 \leq \psi_p(a) \leq \pi.
\]
The famous Sato–Tate conjecture asserts that for any fixed nonzero integer \(a\), when \(p\) varies, the angles \(\psi_p(a)\) are distributed according to the Sato–Tate density
\[
\mu_{ST}(\alpha, \beta) = \frac{2}{\pi} \int_0^\beta \sin^2 \gamma \, d\gamma;
\]
see [1] [6] [7] [8] [9] [11] [12] [14] [15] [17] [18] for various modifications and generalisations of this conjecture and further references.

It is also known that when a sufficiently large prime \(p\) is fixed and \(a\) runs through \(F_p^*\), then, as has been shown by Katz [11] Chapter 13, the work of Deligne on the Weil conjecture implies that the distribution of the sums \(K_p(a)\) is in accordance with the Sato–Tate density; see also [12]. Furthermore, a quantitative form of this result is given by Niederreiter [18]. Namely, if \(A_p(\alpha, \beta)\) is the set of \(a \in F_p^*\) with \(\alpha \leq \psi_p(a) \leq \beta\), then by the main result of Niederreiter [18], we have
\[
\max_{0 \leq \alpha < \beta \leq \pi} |\#A_p(\alpha, \beta) - \mu_{ST}(\alpha, \beta)p| \ll p^{3/4}.
\]

Combining results of Fouvry, Michel, Rivat, and Sarkozy [9] Lemma 2.3 (with \(r = 1\) and of Niederreiter [18] Lemma 3], one can show that elements of \(A_p(\alpha, \beta)\) are uniformly distributed in the following sense. For any \(\lambda \in F_p^*\) and integer \(M\) with \(1 \leq M \leq p - 1\), we put
\[
A_p(\lambda, M; \alpha, \beta) = \{ a \in A_p(\alpha, \beta) : \lambda a \in [1, M] \}.
\]
Then for \(1 \leq M \leq p - 1\), the following bound holds:
\[
\max_{\lambda \in F_p^*} \max_{0 \leq \alpha < \beta \leq \pi} |\#A_p(\lambda, M; \alpha, \beta) - \mu(\alpha, \beta)M| \ll M^{1/2}p^{1/4}(\log p)^{1/2}.
\]

Fouvry, Michel, Rivat, and Sarkozy [9] also remark that by combining a result of Fouvry and Michel [17] with the technique of Vaaler [20], one can show that
\[
\max_{0 \leq \alpha < \beta \leq \pi} |\#Q_p(\alpha, \beta) - \mu_{ST}(\alpha, \beta)p| \ll p^{3/4},
\]
where
\[
Q_p(\alpha, \beta) = \{ a \in F_p : a^2 \in A_p(\alpha, \beta) \}.
\]
The same bound can also be obtained immediately if one applies the result of Niederreiter [18] Lemma 3] to the bound of Michel [17] Corollary 2.4] (see also [7] Lemma 2.1).

Here we show that the same type of distribution is preserved when \(a\) runs through the sums \(a = u + v\) where \(u \in \mathcal{U}, v \in \mathcal{V}\) for any two sufficiently large sets \(\mathcal{U}, \mathcal{V} \subseteq F_p^*\). Namely, for any two sets \(\mathcal{U}, \mathcal{V} \subseteq F_p^*,\) we put
\[
\mathcal{W}_p(\mathcal{U}, \mathcal{V}; \alpha, \beta) = \{ (u, v) \in \mathcal{U} \times \mathcal{V} : u + v \in A_p(\alpha, \beta) \}.
\]
In particular, we obtain an asymptotic formula for \(\#\mathcal{W}_p(\mathcal{U}, \mathcal{V}; \alpha, \beta)\) which is non-trivial whenever
\[
\#\mathcal{U}\#\mathcal{V} \geq p^{3/2+\varepsilon}
\]
for any fixed \(\varepsilon > 0\) and sufficiently large \(p\).

Then, we also improve the upper bound of [19] on the nonlinearity of the Boolean function associated with the sequence of signs of Kloosterman sums; that is, for the function
\[
f(a) = \begin{cases} 0, & \text{if } K(a) > 0 \text{ or } a = 0, \\ 1, & \text{if } K(a) < 0, \end{cases} \quad a = 0, 1, \ldots, 2^n - 1,
\]
where \( n \) is defined by the inequalities
\[
2^n \leq p < 2^{n+1}.
\]

Various pseudorandom properties of the function \( f(a) \) have been studied by Fouvry, Michel, Rivat, and Sarkozy [9]. Here we estimate one more characteristic of \( f(a) \) of cryptographic interest, which in fact has already been considered in [19] whose result we now improve.

We denote by \( B_n \) the \( n \)-dimensional Boolean cube \( B_n = \{0, 1\}^n \) in a natural way identify its elements with the integers in the range \( 0 \leq a \leq 2^n - 1 \) (and thus with a subset of \( \mathbb{F}_p \)).

We define the Fourier coefficients of \( f(a) \) as
\[
\hat{f}(r) = 2^{-n} \sum_{a \in B_n} (-1)^{f(a) + \langle h, r \rangle}, \quad r \in B_n,
\]
where \( \langle a, r \rangle \) denotes the inner product of \( a, r \in B_n \). Furthermore, we recall that
\[
N(f) = 2^{n-1} - 2^{n-1} \max_{r \in B_n} |\hat{f}(r)|
\]
is called the nonlinearity of \( f \) and is an important cryptographic characteristic; for example, see [5]. In particular, it is the smallest possible Hamming distance between the vector of values of \( f \) and the vector of values of a linear function in \( n \) variables over \( \mathbb{F}_2 \).

Several results about some measures of pseudorandomness of the sequence of signs of Kloosterman sums have recently been obtained by Fouvry, Michel, Rivat, and Sarkozy [9]. Motivated by (and actually using) the results of [9], the bound
\[
N(f) = 2^{n-1} \left( 1 + O \left( 2^{-n/24} n^{1/12} \right) \right)
\]
obtained in [19]. Here again we use some results of [9], but in a slightly different way, and we improve this bound.

2. Distribution of Elements of \( \mathcal{A}_p(\alpha, \beta) \)

For a sequence of \( N \) real numbers \( \gamma_1, \ldots, \gamma_N \in [0, 1) \) the discrepancy is defined by
\[
D = \max_{0 \leq \gamma \leq 1} \left| \frac{T(\gamma, N)}{N} - \gamma \right|
\]
where \( T(\gamma, N) \) is the number of \( n \leq N \) such that \( \gamma_n \leq \gamma \).

We also recall our agreement that the elements of \( \mathbb{F}_p \) have canonical representation as integers of the interval \([0, p-1]\). Thus for any field element \( c \in \mathbb{F}_p \), we interpret \( c/p \) as a rational number in the interval \([0, 1]\). Hence, for \( \lambda \in \mathbb{F}_p^* \) we can define the discrepancy \( D_p(\lambda; \alpha, \beta) \) of the sequence
\[
\frac{\lambda a}{p}, \quad a \in \mathcal{A}_p(\alpha, \beta).
\]
Then the bound (2) implies that
\[
\max_{1 \leq M \leq p-1} \max_{\lambda \in \mathbb{F}_p^*} \left| \# \mathcal{A}_p(\lambda, M; \alpha, \beta) - \mu(\alpha, \beta) M \right| \ll p^{3/4} (\log p)^{1/2},
\]
which can be reformulated in the following form:
Lemma 1. We have
\[ \max_{\lambda \in \mathbb{F}_p^*} \max_{0 \leq \alpha < \beta \leq \pi} D_p(\lambda; \alpha, \beta) \ll p^{-1/4} (\log p)^{1/2}. \]

Our main tool is a bound of exponential sums with elements of \( A_p(\alpha, \beta) \). For \( \lambda \in \mathbb{F}_p^* \) we define
\[ S_p(\lambda; \alpha, \beta) = \sum_{a \in A_p(\alpha, \beta)} e_p(\lambda a). \]

Lemma 2. We have
\[ \max_{\lambda \in \mathbb{F}_p^*} \max_{0 \leq \alpha < \beta \leq \pi} |S_p(\lambda; \alpha, \beta)| \ll p^{3/4} (\log p)^{1/2}. \]

Proof. We recall that for any real smooth function \( F(\gamma) \) defined on the interval \([0, 1]\) and any sequence of \( N \) real numbers \( \gamma_1, \ldots, \gamma_N \in [0, 1] \) of discrepancy \( D \), we have
\[ \frac{1}{N} \sum_{n=1}^{N} F(\gamma_n) = \int_{0}^{1} F(\gamma) d\gamma + O(D \max_{0 \leq \gamma \leq 1} |F'(\gamma)|); \]
see [13, Chapter 2, Theorem 5.4]. Writing
\[ S_p(\lambda; \alpha, \beta) = \sum_{a \in A_p(\alpha, \beta)} \cos \left(2\pi \frac{\lambda a}{p}\right) + i \sum_{a \in A_p(\alpha, \beta)} \sin \left(2\pi \frac{\lambda a}{p}\right) \]
and applying Lemma 2, we obtain the desired bound. \( \square \)

3. Sato–Tate conjecture for sum sets

Theorem 3. For any two sets \( U, V \subseteq \mathbb{F}_p^* \), we have
\[ \max_{0 \leq \alpha < \beta \leq \pi} |\#W_p(U, V; \alpha, \beta) - \mu_{ST}(\alpha, \beta) \#U \#V| \leq \sqrt{\#U \#V} p^{3/4} (\log p)^{1/2}. \]

Proof. Using the identity
\[ \frac{1}{p} \sum_{\lambda \in \mathbb{F}_p^*} e_p(\lambda c) = \begin{cases} 1 & \text{if } c = 0, \\ 0 & \text{if } c \in \mathbb{F}_p^*, \end{cases} \]
we write
\[ \#W_p(U, V; \alpha, \beta) = \sum_{u \in U} \sum_{v \in V} \sum_{a \in A_p(\alpha, \beta)} \frac{1}{p} \sum_{\lambda \in \mathbb{F}_p^*} e_p(\lambda(u + v - a)) \]
\[ = \frac{1}{p} \sum_{\lambda \in \mathbb{F}_p^*} S_p(-\lambda; \alpha, \beta) \sum_{u \in U} e_p(\lambda u) \sum_{v \in V} e_p(\lambda v). \]
Separating the term \( \#A_p(\alpha, \beta) \#U \#V/p \) corresponding to \( \lambda = 0 \), we derive
\[ \#W_p(U, V; \alpha, \beta) = \frac{\#A_p(\alpha, \beta) \#U \#V}{p} + O(R), \]
where
\[ R = \frac{1}{p} \sum_{\lambda \in \mathbb{F}_p^*} |S_p(-\lambda; \alpha, \beta)| \left| \sum_{u \in U} e_p(\lambda u) \right| \left| \sum_{v \in V} e_p(\lambda v) \right|. \]
Collecting the above estimates together, we obtain

\[
R \leq p^{-1/4} (\log p)^{1/2} \sum_{\lambda \in F_p^*} \left| \sum_{u \in U} e_p(\lambda u) \right| \left| \sum_{v \in V} e_p(\lambda v) \right|
\]

\[
\leq p^{-1/4} (\log p)^{1/2} \left( \sum_{\lambda \in F_p^*} \left| \sum_{u \in U} e_p(\lambda u) \right|^2 \right)^{1/2} \left( \sum_{\lambda \in F_p^*} \left| \sum_{v \in V} e_p(\lambda v) \right|^2 \right)^{1/2}.
\]

Furthermore, by (1) we see that

\[
\sum_{\lambda \in F_p^*} \left| \sum_{u \in U} e_p(\lambda u) \right|^2 \leq \sum_{\lambda \in F_p^*} \left( \sum_{u \in U} e_p(\lambda u) \right)^2 = \sum_{\lambda \in F_p^*} \sum_{u_1, u_2 \in U} e_p(\lambda (u_1 - u_2)) = p\#U.
\]

Similarly,

\[
\sum_{\lambda \in F_p^*} \left| \sum_{v \in V} e_p(\lambda v) \right|^2 \leq p\#V.
\]

Collecting the above estimates together, we obtain

\[
R \leq \sqrt{\#U\#V} p^{3/4} (\log p)^{1/2},
\]

which after substitution in (6) and using (11) leads us to the bound

\[
|\#W_p(U, V; \alpha, \beta) - \mu_{ST}(\alpha, \beta)\#U\#V| \ll \#U\#V p^{-1/4} + \sqrt{\#U\#V} p^{3/4} (\log p)^{1/2}.
\]

It remains to note that

\[
\#U\#V p^{-1/4} \leq \sqrt{\#U\#V} p^{3/4};
\]

thus the first term never dominates. \(\square\)

Clearly the asymptotic formula of Theorem 3 is nontrivial under condition (8).

4. Nonlinearity

**Theorem 4.** For the nonlinearity \(N(f)\) of the Boolean function \(f(h)\) given by (1), we have

\[
N(f) = 2^{n-1} \left( 1 + O \left( 2^{-n/16} n^{1/8} \right) \right).
\]

**Proof.** We estimate the Fourier coefficients \(\hat{f}(k)\) of \(f\) by using the result that for any integers \(M, h_1, h_2\) with \(0 \leq M \leq M + c_1 < M + c_2 < 2^n\) we have

\[
\sum_{b=0}^{M-1} (-1)^{f(b+c_1)+f(b+c_2)} \ll M^{2/3} p^{1/6} (\log p)^{1/3} + p^{1/2} \log p,
\]

which is a combination of [9, Lemma 2.3] with a special case \(r = 2\) of [9, Lemma 4.4]. In fact, the above bound can be simplified as

\[
\sum_{b=0}^{M-1} (-1)^{f(b+c_1)+f(b+c_2)} \ll M^{2/3} p^{1/6} (\log p)^{1/3}
\]

(since for \(M \leq p^{1/2} \log p\) the bound (7) is trivial and for \(M > p^{1/2} \log p\) we also have \(M^{2/3} p^{1/6} (\log p)^{1/3} > p^{1/2} \log p\).
We now fix some \( m \leq n \) and write \( a, r \in \mathcal{B}_n \) as

\[ a = b + 2^m c \quad \text{and} \quad r = s + 2^m t, \]

with \( 0 \leq b, s < 2^m \) and \( 0 \leq c, t < 2^{n-m} \). In particular

\[ \langle a, r \rangle = \langle b, s \rangle + \langle c, t \rangle. \]

Therefore,

\[
|\hat{f}(r)| = |\hat{f}(s + 2^m t)| = \left| 2^{-n} \sum_{b=0}^{2^{n-1}} \sum_{c=0}^{2^{n-m-1}} (-1)^{f(b+2^m c)+(b,s)+(c,t)} \right| \leq 2^{-n} \sum_{b=0}^{2^{n-1}} \sum_{c=0}^{2^{n-m-1}} (-1)^{f(b+2^m c)+(c,t)}.
\]

By the Cauchy inequality we obtain

\[
|\hat{f}(r)|^2 \leq 2^{m-2n} \sum_{b=0}^{2^{n-1}} \sum_{c=0}^{2^{n-m-1}} (-1)^{f(b+2^m c)+(c,t)} \left| 2^{n-m-1} \sum_{j=0}^{2^{n-m}-1} (-1)^{f(b+2^m c)+(c,t)} \right|^2
\]

\[
= 2^{m-2n} \sum_{b=0}^{2^{n-1}} \sum_{c_1, c_2=0}^{2^{n-m-1}} (-1)^{f(b+2^m c_1)+f(b+2^m c_2)+(c_1,t)+(c_2,t)}
\]

\[
\leq 2^{m-2n} \sum_{b=0}^{2^{n-1}} \sum_{c_1, c_2=0}^{2^{n-m-1}} (-1)^{f(b+2^m c_1)+f(b+2^m c_2)}.
\]

For \( 2^{n-m} \) choices of \( c_1 = c_2 \), the sums over \( b \) are equal to \( 2^m \). For the other choices of \( c_1 \) and \( c_2 \) we can use the bound (7), getting

\[
|\hat{f}(r)|^2 = O \left( 2^{m-2n} \left( 2^{n-m} 2^m + 2^{2(n-m)} 2^{2m/3} 2^{n/6} n^{1/3} \right) \right)
\]

\[
= O \left( 2^{m-n} + 2^{n/6-m/3} n^{1/3} \right).
\]

We now define \( m \) by the inequalities \( 2^m \leq 2^{7n/8} n^{1/4} < 2^{m+1} \), and after simple calculations conclude the proof. \( \square \)

5. Comments

It seems very plausible that [17] Corollary 2.4 can be used to derive a nontrivial estimate for sums

\[ T_p(\chi; \alpha, \beta) = \sum_{a \in \mathcal{A}_p(\alpha, \beta)} \chi(a), \]

with a nonprincipal multiplicative character \( \chi \) of \( \mathbb{F}_p^* \). In this case one can obtain a multiplicative analogue of our results and study the set

\[ \mathcal{Z}_p(\mathcal{U}, \mathcal{V}; \alpha, \beta) = \{ (u, v) \in \mathcal{U} \times \mathcal{V} : uv \in \mathcal{A}_p(\alpha, \beta) \}. \]

Multidimensional analogues of our results which involve joint distributions of Kloosterman sums can be obtained as well.
Also, as a curiosity, we mention that Theorem 3 can be combined with the techniques of [2, 3, 4] to study sets of elements of the Beatty sequence \( \lfloor \vartheta m + \rho \rfloor \) (where \( \vartheta > 0 \) and \( \rho \) are real) which belong to \( A_p(\alpha, \beta) \); that is, sets of the form

\[
B_p(\vartheta, \rho; M; \alpha, \beta) = \{ m \in [1, M] : \lfloor \vartheta m + \rho \rfloor \in A_p(\alpha, \beta) \}.
\]

**References**

[1] A. Adolphson, ‘On the distribution of angles of Kloosterman sums’, *J. Reine Angew. Math.* 395 (1989), 214–220. MR983069 (90k:11109)

[2] W. Banks and I. E. Shparlinski, ‘Non-residues and primitive roots in Beatty sequences’, *Bull. Aust. Math. Soc.* 73 (2006), 433–443. MR2230651 (2007a:11118)

[3] W. Banks and I. E. Shparlinski, ‘Short character sums with Beatty sequences’, *Math. Res. Lett.* 13 (2006), 539–547. MR2250489

[4] W. Banks and I. E. Shparlinski, ‘Prime divisors in Beatty sequences’, *J. Number Theory* 123 (2007), 413–425.

[5] C. Carlet and C. Ding, ‘Highly nonlinear mappings’, *J. Compl.*, 20 (2004), 205–244. MR2067428 (2006d:94043)

[6] C.-L. Chai and W.-C. W. Li, ‘Character sums, automorphic forms, equidistribution, and Ramanujan graphs. I: The Kloosterman sum conjecture over function fields’, *Forum Math.* 15 (2003), 679–699. MR2010030 (2005h:11136)

[7] É. Fouvry and P. Michel, ‘Sommes de modules de sommes d’exponentielles’, *Pacific J. Math.* 209 (2003), 261–288. MR1978371 (2004d:11072)

[8] É. Fouvry and P. Michel, ‘Sur le changement de signe des sommes de Kloosterman’, *Ann. Math.* (to appear).

[9] É. Fouvry, P. Michel, J. Rivat and A. Sárközy, ‘On the pseudorandomness of the signs of Kloosterman sums’, *J. Aust. Math. Soc.* 77 (2004), 425–436. MR2099811 (2005h:11165)

[10] H. Iwaniec and E. Kowalski, *Analytic number theory*. American Mathematical Society, Providence, RI, 2004. MR2061214 (2005k:11005)

[11] N. M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Princeton Univ. Press, Princeton, NJ, 1988. MR955052 (91a:11028)

[12] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, Amer. Math. Soc, Providence, RI, 1999. MR1659828 (2000b:11070)

[13] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York-London-Sydney, 1974. MR0419394 (54:7415)

[14] G. Laumon, ‘Exponential sums and \( l \)-adic cohomology: A survey’, *Israel J. Math.* 120 (2000), 225–257. MR1815377 (2002m:11075)

[15] R. Lidl and H. Niederreiter, *Finite fields*, Cambridge University Press, Cambridge, 1997. MR1429394 (97i:11115)

[16] P. Michel, ‘Autour de la conjecture de Sato-Tate pour les sommes de Kloosterman’, *Duke Math. J.* 92 (1998), 221–254. MR1612781 (99d:11094)

[17] P. Michel, ‘Minoration de sommes d’exponentielles’, *Duke Math. J.* 95 (1998), 227–240. MR1652095 (99i:11069)

[18] H. Niederreiter, ‘The distribution of values of Kloosterman sums’, *Arch. Math.* 56 (1991), 270–277. MR1091880 (92b:11057)

[19] I. E. Shparlinski, ‘On the nonlinearity of the sequence of signs of Kloosterman sums’, *Bull. Aust. Math. Soc.* 71 (2005), 405–409. MR2150929 (2006e:11115)

[20] J. Vaaler, ‘Some extremal functions in Fourier analysis’, *Bull. Amer. Math. Soc.* 12 (1985), 183–216. MR776471 (86g:32005)

**Department of Computing, Macquarie University, Sydney, NSW 2109, Australia**

E-mail address: igor@ics.mq.edu.au