CANCELLATIONS AMONGST KLOOSTERMAN SUMS

IGOR E. SHPARLINSKI AND TIANPING ZHANG

Abstract. We obtain several estimates for bilinear form with Kloosterman sums. Such results can be interpreted as a measure of cancellations amongst with parameters from short intervals. In particular, for certain ranges of parameters we improve some recent results of Blomer, Fouvry, Kowalski, Michel, and Miličević (2014) and Fouvry, Kowalski and Michel (2014).

1. Introduction

Let $p$ be a sufficiently large prime. For integers $m$ and $n$ we define the Kloosterman sum

$$K_p(m, n) = \sum_{x=1}^{p-1} e_p(mx + nx),$$

where $x_p$ is the multiplicative inverse of $x$ modulo $p$ and

$$e_p(z) = \exp(2\pi iz/p).$$

Furthermore, given two intervals

$$\mathcal{I} = [K + 1, K + M], \quad \mathcal{J} = [L + 1, L + N] \subseteq [1, p - 1],$$

and two sequences of weights $\mathcal{A} = \{\alpha_m\}_{m \in \mathcal{I}}$ and $\mathcal{B} = \{\beta_n\}_{n \in \mathcal{J}}$, we define the bilinear sums of Kloosterman sums

$$S_p(\mathcal{A}, \mathcal{B}; \mathcal{I}, \mathcal{J}) = \sum_{m \in \mathcal{I}} \sum_{n \in \mathcal{J}} \alpha_m \beta_n K_p(mn, 1).$$

2010 Mathematics Subject Classification. 11D79, 11L07.

Key words and phrases. Kloosterman sums, cancellation, bilinear form.

This work was supported in part by ARC Grant DP140100118 (for I. E. Shparlinski), by NSFC Grant 11201275 and the Fundamental Research Funds for the Central Universities Grant GK201503014 (for T. P. Zhang).

T. P. Zhang was the corresponding author (tpzhang@snnu.edu.cn).
We also consider the following special cases
\[ S_p(A; \mathcal{I}, \mathcal{J}) = S_p \left(A, \{1\}_{n=1}^N; \mathcal{I}, \mathcal{J} \right) = \sum_{m \in \mathcal{I}} \sum_{n \in \mathcal{J}} \alpha_m K_p(mn, 1), \]
\[ S_p(\mathcal{I}, \mathcal{J}) = S_p \left(\{1\}_{m=1}^M, \{1\}_{n=1}^N; \mathcal{I}, \mathcal{J} \right) = \sum_{m \in \mathcal{I}} \sum_{n \in \mathcal{J}} K_p(mn, 1), \]
\[ S_p(\mathcal{I}) = S_p \left(\{1\}_{m=1}^M, 0; \mathcal{I}, \mathcal{J} \right) = \sum_{m \in \mathcal{I}} K_p(m, 1). \]

Making the change of variable \( x \to nx (\text{mod} \; p) \), one immediately observes that \( K_p(mn, 1) = K_p(m, n) \), thus we also have
\[ S_p(\mathcal{I}, \mathcal{J}) = \sum_{m \in \mathcal{I}} \sum_{n \in \mathcal{J}} K_p(m, n). \]

We also define, for real \( \sigma > 0 \),
\[ \|A\|_\sigma = \left( \sum_{m \in \mathcal{I}} |\alpha_m|^\sigma \right)^{1/\sigma} \quad \text{and} \quad \|B\|_\sigma = \left( \sum_{n \in \mathcal{J}} |\beta_n|^\sigma \right)^{1/\sigma} \]
with the usual convention
\[ \|A\|_\infty = \max_{m \in \mathcal{I}} |\alpha_m| \quad \text{and} \quad \|B\|_\infty = \max_{n \in \mathcal{J}} |\beta_n|. \]

By the Weil bound we have
\[ |K_p(m, n)| \leq 2p^{1/2}, \]
see [7, Theorem 11.11]. Hence
\[ (1.1) \quad S_p(A, B; \mathcal{I}, \mathcal{J}) \leq 2\|A\|_1\|B\|_1 p^{1/2}. \]

We are interested in studying cancellations amongst Kloosterman sums and thus improvements of the trivial bound (1.1).

Throughout the paper, as usual \( A \ll B \) is equivalent to the inequality \( |A| \leq cB \) with some constant \( c > 0 \) (all implied constants are absolute throughout the paper).

2. Previous results

First we note that by a very special case of a much more general result of Fouvry, Michel, Rivat and Sárközy [4, Lemma 2.3] we have
\[ S_p(\mathcal{I}) \ll p \log p \]
which for \( M \geq p^{1/2} \log p \) improves the trivial bound \( S_p(\mathcal{I}) \leq 2Mp^{1/2} \) following from (1.1). Recently, Fouvry, Kowalski, Michel, Raju, Rivat and Soundararajan [5, Corollary 1.6] have given the bound
\[ S_p(\mathcal{I}) \ll Mp^{1/2}(\log p)^{-\eta} \]
provided that $M \geq p^{1/2}(\log p)^{-\eta}$ with some absolute constant $\eta > 0$.

It is also easy to derive from [13, Theorem 7] that
\[
S_p(I, J) \ll MNp^{1/4} + M^{1/2}N^{1/2}p^{1+\omega(1)},
\]
which improves the trivial bound from (1.1) for $MN \geq p^{1+\varepsilon}$ for any fixed $\varepsilon > 0$.

The sums $S_p(A, B; I, J)$ and $S_p(A; I, J)$ have been estimated by Fouvry, Kowalski and Michel [3, Theorem 1.17] as a part of a much more general result about sums of so-called trace functions. Then, by [3, Theorem 1.17(2)], for initial intervals $I = [1, M]$ and $J = [1, N]$, we have
\[
S_p(A; I, J) \leq \|A\|_1 p^{1+\omega(1)}.
\]
Furthermore, by a result of Blomer, Fouvry, Kowalski, Michel, and Milićević [1, Theorem 6.1], also for an initial interval $I$ and an arbitrary interval $J$, with
\[
MN \leq p^{3/2} \quad \text{and} \quad M \leq N^2
\]
we have
\[
S_p(A; I, J) \leq (\|A\|_1\|A\|_2)^{1/2} M^{1/12}N^{7/12}p^{3/4+\omega(1)}.
\]
One can also find in [1, 3, 10] a series of other bounds on the sums $S_p(A; I, J)$ and $S_p(A, B; I, J)$ and also on more general sums.

Finally, Khan [9] has given a nontrivial estimate for the analogue of $S_p(I)$ modulo a fixed prime power which is nontrivial already for $M \geq p^\varepsilon$.

3. NEW RESULTS

We start with the sums $S_p(I, J)$ and present a bound which improves (1.1) already for $MN \geq p^{1/2+\varepsilon}$.

**Theorem 3.1.** We have,
\[
S_p(I, J) \ll (p + MN)p^{\omega(1)}.
\]

We now estimate $S_p(A; I, J)$.

**Theorem 3.2.** We have,
\[
S_p(A; I, J) \ll \|A\|_2 N^{1/2}p.
\]

We can re-write the bounds (2.1) and (2.2) in terms of the $\|A\|_\infty$ as
\[
S_p(A; I, J) \ll \|A\|_\infty M^{1+\omega(1)}
\]
and
\[
S_p(A; I, J) \ll \|A\|_\infty M^{5/6}N^{7/12}p^{3/4+\omega(1)},
\]
respectively, and the bound of Theorem 3.2 as
\[ S_p(A; I, \mathcal{J}) \ll \|A\|_\infty M^{1/2} N^{1/2} p^{1+o(1)}. \]
We now see for any fixed \( \varepsilon > 0 \) the bound (3.3) improves (3.1) and (3.2) for
\[ N < Mp^{-\varepsilon} \quad \text{and} \quad M^4 N \geq p^{3+\varepsilon} \]
respectively, and also applies to intervals \( \mathcal{I} \) and \( \mathcal{J} \) at arbitrary positions.

4. PREPARATIONS

We need the following simple result.

Lemma 4.1. For any integers \( X \) and \( Y \) with \( 1 \leq X, Y < p \), the congruence
\[ xy \equiv 1 \pmod{p}, \quad 1 \leq |x| \leq X, \ 1 \leq |y| \leq Y \]
has at most \( (XY/p + 1)p^{o(1)} \) solutions.

Proof. Writing \( xy \equiv 1 \pmod{p} \) as \( xy = 1 + kp \) for some integer \( k \) with \( |k| \leq XY/p \) and using the bound on the divisor function, see [7, Equation (1.81)], we get the desired estimate. \( \square \)

We also need the following well-known result, which dates back to Vinogradov [15, Chapter 6, Problem 14.a].

Lemma 4.2. For arbitrary set \( \mathcal{U}, \mathcal{V} \subseteq \{0, \ldots, p-1\} \) and complex numbers \( \varphi_u \) and \( \psi_v \) with
\[ \sum_{u \in \mathcal{U}} |\varphi_u|^2 \leq \Phi \quad \text{and} \quad \sum_{v \in \mathcal{V}} |\psi_v|^2 \leq \Psi, \]
we have
\[ \left| \sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{V}} \varphi_u \psi_v e_p(uv) \right| \leq \sqrt{\Phi \Psi p}. \]

5. PROOF OF THEOREM 3.1

For an integer \( u \) we define
\[ \|u\|_p = \min_{k \in \mathbb{Z}} |u - kp| \]
as the distance to the closest integer. Then, changing the order of summation, we obtain
\[ S_p(\mathcal{I}, \mathcal{J}) \ll \sum_{x=1}^{p-1} \min \left\{ M, \frac{p}{\|x\|_p} \right\} \min \left\{ N, \frac{p}{\|T_p\|_p} \right\}, \]
see [7, Bound (8.6)]. We now write
\[ S_p(\mathcal{I}, \mathcal{J}) \ll MNS_1 + MpS_2 + NpS_3 + p^2S_4, \]
where
\[ S_1 = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p}, \quad S_2 = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p}, \quad S_3 = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p}, \quad S_4 = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p}. \]

By Lemma 4.1 we immediately obtain
\[ S_1 \leq (p/MN + 1)p^{o(1)}. \]

To estimate \( S_2 \), we define \( I = \lfloor \log p \rfloor \) and write
\[ S_2 \leq \sum_{i=0}^{I} S_{2,i}, \]
where
\[ S_{2,i} = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p \|x\|_p}. \]
\[ \ll e^{-i}Np^{-1} \sum_{x=1}^{p-1} 1 \ll e^{-i}Np^{-1} \sum_{x=1}^{p-1} 1. \]

Now we use Lemma 4.1 again to derive
\[ S_2 \leq \sum_{i=0}^{I} (M^{-1} + e^{-i}Np^{-1}) p^{o(1)} \]
\[ \ll (I + 1)M^{-1}p^{o(1)} + Np^{-1+o(1)} \sum_{i=0}^{I} e^{-i} \]
\[ \ll (I + 1)M^{-1}p^{o(1)} + Np^{-1+o(1)} \]
\[ \ll M^{-1}p^{o(1)} + Np^{-1+o(1)}. \]

Similarly we obtain
\[ S_3 \leq N^{-1}p^{o(1)} + Mp^{-1+o(1)}. \]
Finally, we write

\[ S_4 \leq \sum_{i,j=0}^I S_{4,i,j}, \]

where

\[ S_{4,i,j} = \sum_{x=1}^{p-1} \frac{1}{\|x\|_p \|x\|_p} \]

\[ \leq e^{-i-j} MNp^{-2} \sum_{x=1}^{p-1} \frac{1}{\|x\|_p \|x\|_p} \]

\[ \leq e^{-i-j} MNp^{-2} \sum_{x=1}^{p-1} \frac{1}{\|x\|_p \|x\|_p} \]

Applying Lemma 4.1 one more time, we obtain

\[ S_{4,i,j} \ll e^{-i-j} MNp^{-2} (e^{i+j} p/MN + 1) p^{o(1)} \]

\[ = (p^{-1} + e^{-i-j} MNp^{-2}) p^{o(1)}. \]

Hence

\[ S_4 \leq \sum_{i,j=0}^I (p^{-1} + e^{-i-j} MNp^{-2}) p^{o(1)} \]

\[ \leq (I + 1)^2 p^{-1+o(1)} + MNp^{-2+o(1)} \]

\[ \leq p^{-1+o(1)} + MNp^{-2+o(1)}. \]

Combining (5.2), (5.3), (5.4) and (5.5) we obtain the result.

6. Proof of Theorem 3.2

Changing the order of summation, as in the proof of Theorem 3.1 we obtain

\[ S_p(A; \mathcal{I}, \mathcal{J}) = \left| \sum_{m \in \mathcal{I}} \sum_{x=1}^{p-1} \alpha_m \gamma_x e_p(m \overline{x}_p) \right|, \]

where

\[ |\gamma_x| \leq \min \left\{ N, \frac{p}{\|x\|_p} \right\}. \]
Thus, similarly to the proof of Theorem 3.1 we define \( I = \lceil \log p \rceil \) and write

\[
S_p(I, J) \ll |S_0| + \sum_{i=1}^{I} |S_i|,
\]

where

\[
S_0 = \sum_{m \in I} \sum_{x=1}^{p-1} \alpha_m \gamma_x e_p(m \overline{r}_p),
\]

\[
S_i = \sum_{m \in I} \sum_{x=1}^{p-1} \alpha_m \gamma_x e_p(m \overline{r}_p), \quad i = 1, \ldots, I.
\]

Now use Lemma 4.2, we have

\[
|S_0| \ll \|A\|_2 N \sqrt{(p/N + 1)p} \leq \|A\|_2 N^{1/2} p.
\]

Also, for \( i = 1, \ldots, I \), using that if \( e^{i+1} p/N \geq \|x\|_p > e^i p/N \) then \( \gamma_x \ll N e^{-i} \), hence, by Lemma 4.2, we obtain

\[
S_i \ll \|A\|_2 (N^2 e^{-2i} e^i p/N)^{1/2} p^{1/2} = e^{-i/2} \|A\|_2 N^{1/2} p.
\]

Therefore,

\[
\sum_{i=1}^{I} |S_i| \ll \|A\|_2 N^{1/2} p \sum_{i=1}^{I} e^{-i/2} \ll \|A\|_2 N^{1/2} p.
\]

Combining (6.2) and (6.3), we obtain the result.

7. Comments

It is also natural to consider cancellations between some other exponential and character sums. For example, in [14] one can find some bound on the following sums

\[
S_p(f, \mathcal{A}, \mathcal{B}; \mathcal{C}) = \sum_{(u,v) \in \mathcal{C}} \alpha_u \beta_v e_p(v/f(u)),
\]

\[
T_p(f, \mathcal{A}, \mathcal{B}; \mathcal{C}) = \sum_{(u,v) \in \mathcal{C}} \alpha_u \beta_v \chi(v + f(u)),
\]

(where \( \chi \) is a multiplicative character modulo \( p \), over a convex set \( \mathcal{C} \subseteq [1, U] \times [1, V] \), with some integer \( 1 \leq U, V < p \).
Here we also note that one can also obtain a nontrivial cancellation for sums
\[ H_{k,p}(a; \mathcal{I}) = \sum_{m \in \mathcal{I}} G_{k,p}(am) \]
of Gaussian sums
\[ G_{k,p}(a) = \sum_{x=0}^{p-1} e_p(ax^k) \]
with a positive integer \( k \mid p - 1 \). Indeed, we define
\[ \tau_p(a; \chi) = \sum_{x=1}^{p-1} \chi(x) e_p(ax), \]
where \( \chi \) is a multiplicative character; we refer to [7, Chapter 3] for a background on multiplicative characters. Then by the orthogonality of characters, we have
\[ G_{k,p}(a) = \sum_{\chi^k=\chi_0} \chi(a) \tau_p(a; \chi), \]
where the summation is over all nonprincipal multiplicative characters \( \chi \) modulo \( p \), such that \( \chi^k \) is the principal character \( \chi_0 \), see also [11, Theorem 5.30]. Using that \(|\tau_p(a; \chi)| = p^{1/2}\) for any nonprincipal multiplicative characters \( \chi \) and integer \( a \) with \( \gcd(a, p) = 1 \), we derive
\[ |H_{k,p}(a; \mathcal{I})| = p^{1/2} \sum_{\chi^k=\chi_0} \chi(a) \sum_{m \in \mathcal{I}} \chi(m). \]
Thus applying the Burgess bound, see [7, Equation (12.58)], we derive that
\[ H_{k,p}(a; \mathcal{I}) \ll M^{1-1/\nu} p^{1/2+\nu+1}/(4\nu^2) (\log p)^{1/\nu} \]
\[ = M^{1-1/\nu} p^{(2\nu^2+\nu+1)/4(\nu^2)} (\log p)^{1/\nu} \]
for any fixed \( k \mid p - 1 \) and \( \nu = 1, 2, \ldots \).

Similarly, for general quadratic polynomials \( f(X) = aX^2 + bX \), with \( \gcd(a, p) = 1 \), we can define the double sums
\[ F_p(a, b; \mathcal{I}) = \sum_{m \in \mathcal{I}} \sum_{x=0}^{p-1} e_p(m (ax^2 + bx)). \]
It is easy to see that
\[
\sum_{x=0}^{p-1} e_p(ax^2 + bx) = e_p \left( -\frac{b^2}{4a} \right) \sum_{x=0}^{p-1} e_p \left( a \left( x + b/(2a) \right)^2 \right)
\]
\[
= e_p \left( -\frac{b^2}{4a} \right) \sum_{x=0}^{p-1} e_p (ax^2) = \left( \frac{a}{p} \right) e_p \left( -\frac{b^2}{4a} \right) G_{2,p}(1),
\]
where \((a/p)\) is the Legendre symbol of \(a\) modulo \(p\). Hence
\[
\sum_{m \in \mathcal{I}} \sum_{x=0}^{p-1} e_p (m(ax^2 + bx)) = G_{2,p}(1) \sum_{m \in \mathcal{I}} \left( \frac{am}{p} \right) e_p \left( -\frac{(bm)^2}{4am} \right)
\]
\[
= \left( \frac{a}{p} \right) G_{2,p}(1) \sum_{m \in \mathcal{I}} \left( \frac{m}{p} \right) e_p \left( -\frac{b^2m}{4a} \right).
\]
Now, using the bound of Burgess [2] on short mixed sums (see [6, 8, 12] for various generalisations) we easily derive that for any fixed \(\nu = 2, 3, \ldots\) we have
\[
\mathcal{F}_p(a, b; \mathcal{I}) \ll M^{1-1/\nu} p^{1/2+1/(4(\nu-1))} \left( \log p \right)^2
\]
\[
= M^{1-1/\nu} p^{(2\nu-1)/(4(\nu-1))} \left( \log p \right)^2,
\]
where the implied constant may depend on \(\nu\).

We note that the bounds (7.1) and (7.2) are nontrivial provided that \(M \geq p^{1/4+\varepsilon}\) for any fixed \(\varepsilon > 0\) and sufficiently large \(p\).

ACKNOWLEDGEMENT

The second author gratefully acknowledge the support, the hospitality and the excellent conditions at the School of Mathematics and Statistics of UNSW during his visit.

This work was supported in part by ARC Grant DP140100118 (for I. E. Shparlinski), by NSFC Grant 11201275 and the Fundamental Research Funds for the Central Universities Grant GK201503014 (for T. P. Zhang).

REFERENCES

[1] V. Blomer, É. Fouvry, E. Kowalski, P. Michel, and D. Miličević, ‘On moments of twisted L-functions’, Preprint, 2014 (available from http://arxiv.org/abs/1411.4467).
[2] D. A. Burgess, ‘Partial Gaussian sums’, Bull. London Math. Soc., 20 (1988), 589–592.
[3] É. Fouvry, E. Kowalski and P. Michel, ‘Algebraic trace functions over the primes’, Duke Math. J. 163 (2014), 1683–1736.
[4] 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.

[5] É. Fouvry, E. Kowalski, P. Michel, C. S. Raju, J. Rivat and K. Soundararajan, ‘On short sums of trace functions’, *Preprint*, 2015 (available from http://arxiv.org/abs/1508.00512).

[6] D. R. Heath-Brown and L. Pierce, ‘Burgess bounds for short mixed character sums’, *J. Lond. Math. Soc.*, 91 (2015) 693–708.

[7] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society, Providence, RI, 2004.

[8] B. Kerr, ‘Some mixed character sums’, *Preprint*, 2014 (available from http://arxiv.org/abs/1410.3587).

[9] R. Khan, ‘The divisor function in arithmetic progressions modulo prime powers’, *Preprint*, 2015 (available from http://arxiv.org/abs/1510.03377).

[10] E. Kowalski, P. Michel, and W. Sawin, ‘Bilinear forms with Kloosterman sums and applications’, *Preprint*, 2015 (available from http://arxiv.org/abs/1511.01636).

[11] R. Lidl and H. Niederreiter, *Finite Fields*, Cambridge Univ. Press, Cambridge, 1997.

[12] L. B. Pierce, ‘Burgess bounds for multi-dimensional short mixed character sums’, *J. Number Theory*, (to appear).

[13] I. E. Shparlinski, ‘Distribution of inverses and multiples of small integers and the Sato–Tate conjecture on average’, *Michigan Math. J.*, 56 (2008), 99–111.

[14] I. E. Shparlinski, ‘On bilinear exponential and character sums with reciprocals of polynomials’, *Preprint*, 2015, (available from http://arxiv.org/abs/1504.03192).

[15] I. M. Vinogradov, *Elements of number theory*, Dover Publ., NY, 1954.