ACTA ARITHMETICA
LXXXVIII.3 (1999)

Averages of short exponential sums

by

Alexandru Zaharescu (Cambridge, Mass. and Bucureşti)

1. Introduction. Let p be a large prime number, let $r(X)$ be a rational function with coefficients in $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ and consider the complete exponential sum

$$S(r,p) = \sum_{x \bmod p}^{*} e\left(\frac{r(x)}{p}\right)$$

where $\sum^*$ means that the poles of $r(X)$ are excepted.

An upper bound for $S(r,p)$ is provided by the Bombieri–Weil inequality (see [2]). In the convenient form given by Moreno and Moreno [3] it states that if r is not constant and $r(X) = f(X)/g(X)$ with $f(X), g(X) \in \mathbb{F}_p[X]$ and $(f(X), g(X)) = 1$ then:

$$|S(r,p)| \le Dp^{1/2}$$

where

$$D = \deg g(X) + \max\{\deg f(X), \deg g(X)\} - 1.$$

If we consider an incomplete sum

$$S_N(r,p) = \sum_{1 \le n \le N} e\left(\frac{r(n)}{p}\right)$$

then all that is known is an upper bound of the form

$$|S(r,p)| \ll Dp^{1/2} \log p$$

which holds true if $r(X)$ is not a linear polynomial and which is derived in a standard way (see e.g. Serre [5], Appendice) from the previous bound for complete sums.

Here we expect that for all rational functions $r(X)$ with the exception of certain polynomials for which $S_N(r,p)$ is obviously large one still has a square root type cancellation in this sum, so we expect

$$|S_N(r,p)| \ll_{\varepsilon,D} N^{1/2} p^{\varepsilon}. \tag{1.1}$$

1991 *Mathematics Subject Classification*: Primary 11L07.

To study the distribution of the sequence of fractional parts

$$\mathcal{N} = \{\{r(n)/p\} : 1 \le n \le N\}$$

one also considers the sums

$$S_N(m, r, p) = \sum_{1 \le n \le N}^{*} e\left(\frac{mr(n)}{p}\right), \qquad m = 1, 2, \dots$$

In the following we are concerned with estimating $|S_N(m, r, p)|$ in average over $m \in \{1, \dots, M\}$ via the second moment

$$M_2(N, M, r, p) = \sum_{m=1}^{M} |S_N(m, r, p)|^2.$$

Our intention is to find circumstances under which our upper bounds for $M_2(N, M, r, p)$ imply a square root type cancellation in $S_N(m, r, p)$ in average over $m \in \{1, \dots, M\}$, or at least imply in average a cancellation which is stronger than the $p^{1/2}$-bound which we know individually for any m.

Usually the quality of these upper bounds for M_2 depends on M: the larger M, the better the results. In particular, it is easy to see that

$$\sum_{m \bmod p} |S_N(m, r, p)|^2 \ll_D pN$$

so we indeed get a square root type upper bound in average for $S_N(m, r, p)$ if we allow M be as large as p.

In this paper we consider smaller values for M, usually as small as N. So we might have $N = M = p^\alpha$ where α is some fixed small positive number.

One possible way to obtain information about such short sums is to estimate higher moments, but this seems to be difficult in general.

In this paper we allow p to vary in some interval of the form $[P, 2P]$ with $P \ge N$.

In order for these things to make sense we start with a rational function $r(X) = f(X)/g(X)$ where $f(X), g(X) \in \mathbb{Z}[X]$ and $(f, g) = 1$ and with positive integers N, M, P with $N, M \le P$; then by reducing the coefficients of $f(X)$ and $g(X)$ modulo p for p prime in $[P, 2P]$ we form the above moments $M_2(N, M, r, p)$.

In the following we present a systematic method to obtain bounds for the second moment

$$M_2(N, M, P, r) = \sum_{P \le p \le 2P} M_2(N, M, r, p);$$

these bounds turn out to be essentially best possible if $M \ge N$.

First we need to exclude some trivial cases when our exponential sums are obviously large. This is the case when $r(X)$ is constant, and also when $r(X) = aX + b$ in which case we get geometric progressions which might

also be large. In fact we might get large exponential sums even if $r(X)$ is a higher degree polynomial. For example, if $r(X) = X^k$ and $N^{k+2} \le P$ then any exponential sum appearing in $M_2(N, N, P, r)$ will be as large as N.

In the following we will assume that $r(X)$ is not a polynomial.

Concerning the coefficients of $r(X)$ our method is very flexible and we may allow them to be larger than P. However, an upper bound for them needs to be assumed, otherwise one can use the Chinese Remainder Theorem to find (huge) coefficients for $r(X)$ which produce images mod p at our disposal for any prime $p \in [P, 2P]$ and so we will not get anything by averaging over p: it is like working with each p individually. For example, if $r(X) = (X + a)/(X + 1)$ say, where $a - 1$ is divisible by the product of primes p from $[P, 2P]$ then all our sums will be trivial.

In the following we assume that all the coefficients of $r(X)$ are bounded by P^{K_1} for some fixed number K_1. So our upper bounds will depend on K_1. They will also depend on $\deg f$ and $\deg g$ which are also assumed to be bounded by some fixed positive number K_2, say.

Now we can state our main results.

THEOREM 1. *Let $\varepsilon > 0$, $N \le P$ and M be positive integers and let $r(X) = f(X)/g(X)$ be a rational function with $\deg f, \deg g \le K_2$ with integer coefficients bounded by P^{K_1}, $r(X)$ not a polynomial. Then*

$$M_2(N, M, P, r) \ll_{\varepsilon, K_1, K_2} N(N + M)P^{1+\varepsilon}.$$

As a consequence, for $M \ge N$ we get the desired square root type cancellation in average for our exponential sums:

COROLLARY 2. *Under the hypotheses of Theorem 1, if $M \ge N$ then for almost all pairs (p, m) with p prime, $p \in [P, 2P]$, $m \in \{1, \dots, M\}$ (in the sense that the exceptional set has density $< P^{-\varepsilon}$) one has*

$$\Big| \sideset{}{^*}\sum_{1 \le n \le N} e\Big(\frac{mr(n)}{p}\Big) \Big| \ll_{\varepsilon, K_1, K_2} N^{1/2} P^{\varepsilon}.$$

As an application of the above results we note that almost all the sets $\mathcal{N}_{p,m}$ defined mod 1 by $\{mr(n)/p\}_{1 \le n \le N}$ are uniformly distributed in $[0, 1]$. To measure how far is a given finite sequence $\mathcal{N} = \{x_n : 1 \le n \le N\}$ of points in $[0, 1]$ from being uniformly distributed one defines the discrepancy of $\mathcal{N}$ by:

$$D(\mathcal{N}) = \sup_{0 \le \alpha < \beta \le 1} |\#(\mathcal{N} \cap [\alpha, \beta]) - N(\beta - \alpha)|.$$

Concerning the discrepancy of our sequences we have the following result:

THEOREM 3. *Under the hypotheses of Theorem 1 for almost all pairs (p, m) with p prime, $p \in [P, 2P]$ and $m \in \{1, \dots, M\}$ we have*

$$D(\mathcal{N}_{p,m}) \ll_{\varepsilon, K_1, K_2} (NM^{-1/2} + N^{1/2})P^{\varepsilon}.$$

In particular, if $N = P^\alpha$ and $M \geq N$ then we get a square root type saving in average in the discrepancy, no matter how small the fixed positive number α is.

A problem originally proposed by Hardy and Littlewood and intensively studied later concerns the distribution of fractional parts of polynomials (see Schmidt [4] and Baker [1]). In particular, one is interested in finding small values of such fractional parts.

If we consider the analogous problem in which instead of a polynomial we have a rational function $r(X)$, where by $\{mr(n)/p\}$ we understand that $r(n)$ is to be computed in $\mathbb{F}_p$ and not in $\mathbb{Q}$, then Theorem 3 implies

COROLLARY 4. *Under the hypotheses of Theorem* 1 *assume also that* $M \geq N$. *Then for almost all pairs* (p, m) *with* p *prime,* $p \in [P, 2P]$ *and* $m \in \{1, \dots, M\}$ *the following holds true*:

For any $\beta \in [0, 1]$ *there exists* $1 \leq n \leq N$ *such that*

$$\left| \left\{ \frac{mr(n)}{p} \right\} - \beta \right| \ll_{\varepsilon, K_1, K_2} N^{-1/2} P^{\varepsilon}. \tag{1.2}$$

2. Exponential sums and pair correlations. There are two main ideas in the proof of Theorem 1.

The first idea is to bring into play the pair correlations for the original sequence $\mathcal{N}_p = \{\{r(n)/p\} : 1 \leq n \leq N\}$ to bound the second moments $M_2(N, M, r, p)$. The other idea, which will be explained in the next section, provides us with an alternative way to estimate these pair correlations when we vary the modulus p.

We work here with a general sequence $\mathcal{N} = \{x_n\}_{1 \leq n \leq N}$ of points in $[0, 1]$ and then we apply the results to the above sets $\mathcal{N}_p$ for all our places $p \in [P, 2P]$.

We let $T \geq 2$ be a parameter to be chosen later and define the function h periodic mod 1 and which on $[-1/2, 1/2]$ is given by

$$h(t) = \begin{cases} T(1 - T|t|), & |t| \leq 1/T, \\ 0, & 1/T \leq |t| \leq 1/2. \end{cases}$$

We expand h in a Fourier series:

$$h(t) = \sum_{m \in \mathbb{Z}} c_m e(mt).$$

Here $c_0 = 1$ and for any $m \neq 0$ we have

$$c_m = \frac{T^2}{\pi^2 m^2} \sin^2 \left(\frac{\pi m}{T} \right).$$

Note that for $|m| \leq T/2$ we have $|c_m| \gg 1$.

Now let

$$E(T) = \#\{1 \le n_1, n_2 \le N : |x_{n_1} - x_{n_2}| \le 1/T\}$$

and

$$E_h = \sum_{1 \le n_1, n_2 \le N} h(x_{n_1} - x_{n_2}).$$

Then obviously $E_h \le TE(T)$. On the other hand, we have

$$E_h = \sum_{1 \le n_1, n_2 \le N} h(x_{n_1} - x_{n_2}) = \sum_{1 \le n_1, n_2 \le N} \sum_{m \in \mathbb{Z}} c_m e(m(x_{n_1} - x_{n_2}))$$
$$= \sum_{m \in \mathbb{Z}} c_m \sum_{1 \le n_1, n_2 \le N} e(m(x_{n_1} - x_{n_2})) = \sum_{m \in \mathbb{Z}} c_m \Big| \sum_{1 \le n \le N} e(mx_n) \Big|^2.$$

We now let $T = 2M$. Since $c_m \ge 0$ for any m and $c_m \gg 1$ for $|m| \le M$, we derive

$$\sum_{1 \le m \le M} \Big| \sum_{1 \le n \le N} e(mx_n) \Big|^2 \ll E_h \ll ME(2M). \tag{2.1}$$

We note in passing that here we could add the term $m = 0$ to the left hand side. This would contribute an N^2 and so we see that by this method it is not possible to obtain a square root type cancellation in Corollary 2 above if M is smaller than N since in that case all the terms $m = 1, \dots, M$ together will be dominated by the single term $m = 0$.

3. Averaging over the modulus P. For any $p \in [P, 2P]$ we consider the set $\mathcal{N}_p = \{\{r(n)/p\}_{1 \le n \le N}\}$ for which we apply (2.1), where $E(2M) = E_p(2M)$ that appears on its right hand side depends on p also. We add these inequalities to get

$$\sum_{P \le p \le 2P} \sum_{1 \le m \le M} \Big| \sum_{1 \le n \le N} e\Big(\frac{mr(n)}{p}\Big) \Big|^2 \ll M \sum_{P \le p \le 2P} E_p(2M). \tag{3.1}$$

Here for any p we have

$$E_p(2M) = \#\{(n, n') : 1 \le n, n' \le N, r(n) - r(n') = h \bmod p,\ |h| \le p/(2M)\}. \tag{3.2}$$

Therefore

$$\sum_p E_p(2M) = \#\{(n, n', h, p) : 1 \le n, n' \le N,\ P \le p \le 2P, \ |h| \le p/(2M);\ r(n) - r(n') = h \bmod p\}. \tag{3.3}$$

The above congruence is equivalent to

$$g(n')f(n) - g(n)f(n') = hg(n')g(n) \bmod p. \tag{3.4}$$

Now the point is that for any given n, n' and h the integer $A = g(n')f(n) - g(n)f(n') - hg(n')g(n)$ has very few prime divisors p unless $A = 0$ in which case any prime is a divisor and so any prime $P \le p \le 2P$ will contribute to a tuple (n, n', h, p) as above.

So we distinguish two types of admissible tuples, according as $A = 0$ or not.

Now $|A|$ is clearly bounded by $P^{2k}N^{2k'}(P/M+1)$ which is $< P^{2k+2k'+1}$ so if $A \neq 0$ then the triplet (n, n', h) can appear in at most $2k+2k'+1$ tuples (n, n', h, p). Therefore the total number of such tuples is $\ll_{k,k'} N^2P/M$. It remains to estimate the number of triplets (n, n', h) for which $A = 0$, in other words, to estimate the number of pairs (n, n') for which $r(n) - r(n')$ is an integer h which belongs to the interval $[-P/M, P/M]$.

First of all, we have the N diagonal solutions $n = n'$. Assume now that $n \neq n'$.

Let us first treat the special case $r(X) = aX + b/X$, i.e. the case of so-called Kloosterman sums. We have $r(n) - r(n') = a(n-n') + b/n - b/n'$. Put $z = n' - n$. Then $r(n') - r(n) = az - bz/(n(n+z))$. If this is an integer then n has to divide bz. Now for any fixed z the number bz cannot have too many divisors. More precisely, since $|bz| \le P^K N \le P^{K+1}$ the number of divisors of bz will be $\ll_{\varepsilon,K} P^{\varepsilon}$.

Therefore the number of such pairs (n, n') is $\ll_{\varepsilon,K} NP^{\varepsilon}$. It follows that

$$\sum_p E_p(2M) \ll_{\varepsilon,K,K'} \left(\frac{N^2P}{M} + NP^{1+\varepsilon}\right). \tag{3.5}$$

Therefore the right hand side of (3.1) is $\ll_{\varepsilon,K,K'} P^{1+\varepsilon}N(N+M)$.

The left hand side of (3.1) equals $M_2(N, M, P, r)$. Thus Theorem 1 is proved for Kloosterman sums.

Actually in this case we can weaken the hypothesis on the coefficients a and b; more precisely, it is enough to assume that $|a|, |b| \le \exp(p^{\varepsilon})$.

The point is that if one of $|a|, |b|$ is large (larger than p^K, say) then the other has to be large too, otherwise $az - bz/(n(n+z))$ will be itself too large for the constraint that we have for it, i.e. that it lies in $[-P/M, P/M]$. Now if both a and b are large then for any z there will be at most one admissible n, in the sense that for at most one value of n the quantity $|az - bz/(n(n+z))|$ is smaller than P/M. ∎

It would be nice to have this part of the proof generalized such that it can be applied to a general $r(X)$.

For $r(X) = aX + b/X$ there is no upper bound for $|a|$, $|b|$ needed to be assumed in this last part of the proof, i.e. when $A = 0$. But when $A \neq 0$ we need the number of prime divisors p of A with $P \le p \le 2P$ to be $< P^{\varepsilon}$ so that their contribution in Theorem 1 could be neglected, in the sense

that it could be swallowed in the factor P^ε and that is why we assume $|a|, |b| < \exp(p^\varepsilon)$.

We now return to the proof of Theorem 1. Take a general $r(X) = f(X)/g(X)$ with $(f,g) = 1$. Let $f(X) = a_r X^r + a_{r-1}X^{r-1} + \ldots + a_0$ and $g(X) = b_s g^s + \ldots + b_0$. Then the resultant $R(f,g)$ of f and g is a nonzero integer. From its expression as a determinant whose nonzero entries are coefficients of $f(X)$ and $g(X)$ we obtain

$$|R(f,g)| \ll_{K'} p^{2KK'}.$$

On the other hand, we have $f(X)F(X) - g(X)G(X) = R(f,g)$ for some polynomials $F(X), G(X) \in \mathbb{Z}[X]$.

From the assumption that

$$\frac{f(n)}{g(n)} - \frac{f(n')}{g(n')} \in \mathbb{Z}$$

we derive in order the following:

$$g(n)\left(\frac{f(n)}{g(n)} - \frac{f(n')}{g(n')}\right) \in \mathbb{Z}; \quad \frac{g(n)f(n')}{g(n')} \in \mathbb{Z}; \quad \frac{g(n)f(n')F(n')}{g(n')} \in \mathbb{Z};$$
$$\frac{g(n)(R(f,g) + g(n')G(n'))}{g(n')} \in \mathbb{Z}; \quad \frac{g(n)R(f,g)}{g(n')} \in \mathbb{Z}.$$

Now let us fix an n and see how many n' satisfy the last relation. The numerator is a nonzero integer, as n was not a pole of $r(x)$, which is bounded by $p^{2KK'}N^s p^K$ and therefore its number of divisors is $O_{\varepsilon,K,K'}(p^\varepsilon)$. Moreover for any such divisor d there are at most s distinct values for n' such that $g(n') = d$.

In conclusion there are at most $O_{\varepsilon,K,K'}(p^\varepsilon N)$ admissible triplets (n, n', h) of second type ($A = 0$). Therefore

$$\sum_p E_p(2M) \ll_{\varepsilon,K,K'} \left(\frac{N^2 P}{M} + NP^{1+\varepsilon}\right)$$

as in the case $r(X) = aX + b/X$ and this is enough to finish the proof of Theorem 1 as before. ∎

4. Proof of Theorem 3. To prove Theorem 3 we use for each set $\mathcal{N}_{p,m}$ the Erdős–Turán inequality which provides an upper bound for the discrepancy of the set in terms of exponential sums:

$$D(\mathcal{N}_{p,m}) \le \sum_{1 \le l \le L} \frac{3}{l} |S_N(ml, r, p)| + \frac{N}{L+1},$$

which is valid for any positive integer L. We sum these inequalities over the

corresponding ranges for $m,\ p$:

$$\sum_{P\le p\le 2P}\sum_{1\le m\le M} D(\mathcal{N}_{p,m}) \le \sum_{P\le p\le 2P}\sum_{1\le m\le M}\Big(\sum_{1\le l\le L}\frac{3}{l}|S_N(ml,r,p)| + \frac{N}{L+1}\Big)$$

where L is a parameter to be chosen later. Now we change the order of summation and then use Cauchy's inequality to bound the inner sum:

$$\begin{aligned}\sum_{P\le p\le 2P}\sum_{1\le m\le M} & D(\mathcal{N}_{p,m})\\ &\le \frac{PMN}{L+1} + \sum_{1\le l\le L}\frac{3}{l}\sum_{P\le p\le 2P}\sum_{1\le m\le M}|S_N(ml,r,p)|\\ &\le \frac{PMN}{L+1} + \sum_{1\le l\le L}\frac{3}{l}(MP)^{1/2}\Big(\sum_{P\le p\le 2P}\sum_{1\le m\le M}|S_N(ml,r,p)|^2\Big)^{1/2}.\end{aligned}$$

Note that $S_N(ml,r,p) = S_N(m,lr,p)$. Thus

$$\sum_{P\le p\le 2P}\sum_{1\le m\le M}|S_N(ml,r,p)|^2 = M_2(N,M,P,lr),$$

which is bounded by $N(N+M)P^{1+\varepsilon}$ by Theorem 1 applied to $lr(X)$. We obtain

$$\begin{aligned}\sum_{P\le p\le 2P}\sum_{1\le m\le M} & D(\mathcal{N}_{p,m})\\ &\ll_{\varepsilon,k,\deg r} \frac{PMN}{L} + \sum_{1\le l\le L}\frac{M^{1/2}N^{1/2}(N+M)^{1/2}P^{1+\varepsilon}}{l}.\end{aligned}$$

If we take $L = N$, the right hand side above is

$$\ll_{\varepsilon,k,\deg r} M^{1/2}N^{1/2}(N+M)^{1/2}P^{1+\varepsilon}.$$

It now follows that for almost all pairs p, m we have

$$D(\mathcal{N}_{p,m}) \ll_{\varepsilon,k,\deg r} (NM^{-1/2} + N^{1/2})P^{\varepsilon},$$

which completes the proof of Theorem 3.

References

[1] R. C. Baker, *Diophantine Inequalities*, Clarendon Press, Oxford, 1986.

[2] E. Bombieri, *On exponential sums in finite fields*, Amer. J. Math. 88 (1966), 71–105.

[3] C. J. Moreno and O. Moreno, *Exponential sums and Goppa codes*, Proc. Amer. Math. Soc. 111 (1991), 523–531.

[4] W. M. Schmidt, *Small fractional parts of polynomials*, in: Regional Conference Series in Math. 32, Amer. Math. Soc., Providence, 1977.

[5] J. P. Serre, *Majoration de sommes exponentielles*, Astérisque 41–42 (1977), 111–126.

Department of Mathematics
Room 2-267
Massachusetts Institute of Technology
77 Mass. Avenue
Cambridge, Massachusetts 02139
U.S.A.

Mathematics Research Institute
of the Romanian Academy
P.O. Box 1-764
Bucureşti, Romania

Received on 27.1.1998
and in revised form on 21.10.1998 (3330)