Burgess bounds for short mixed character sums

D. R. Heath-Brown*
L. B. Pierce†

Abstract

This paper proves nontrivial bounds for short mixed character sums by introducing estimates for Vinogradov's mean value theorem into a version of the Burgess method.

1 Introduction

Let $\chi(n)$ be a non-principal character of modulus q, and consider the character sum

$$S(N,H)=\sum_{N<n\leq N+H}\chi(n). \tag{1.1}$$

The classical Pólya-Vinogradov inequality provides the bound

$$|S(N,H)|\ll q^{1/2}\log q,$$

which is nontrivial only if the length H of the character sum is longer than $q^{1/2+\varepsilon}$. In a classic series of papers, Burgess [2], [3], [4], [5] introduced a method for bounding short character sums that results in the following well-known bound: for χ a primitive multiplicative character to a prime modulus q one has

$$|S(N,H)|\ll H^{1-\frac{1}{r}}q^{\frac{r+1}{4r^2}}\log q, \tag{1.2}$$

for any integer $r\geq 1$; moreover this bound is uniform in N. This provides a nontrivial estimate for $S(N,H)$ as soon as $H>q^{1/4+\varepsilon}$; more precisely if $H=q^{1/4+\kappa}$, then the Burgess bound is of size $Hq^{-\delta}$ with $\delta\approx\kappa^2$. Indeed Burgess proved a similar bound for arbitrary moduli q when $r\leq 3$, and for general cube-free moduli for all r.

Burgess bounds have found valuable applications in a range of settings, and it would be highly desirable to develop variations of the Burgess method for mixed character sums of the form

$$\sum_{N<n\leq N+H}e_q(f_1(n)\overline{f_2(n)})\chi(f_3(n)\overline{f_4(n)}),$$

*Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, *rhb@maths.ox.ac.uk*
†Hausdorff Center for Mathematics, 62 Endenicher Allee, 53115 Bonn, *pierce@math.uni-bonn.de*

for appropriate polynomials $f_1, \dots, f_4$ and $e_q(t) = e^{2\pi it/q}$. However, it has proved difficult to handle sums involving χ evaluated at anything other than a linear function of n.

This paper will be concerned with the short mixed character sum

$$S(f; N, H) = \sum_{N<n\leq N+H} e(f(n))\chi(n), \tag{1.3}$$

for prime moduli q, where f is a real-valued polynomial and $e(t) = e^{2\pi it}$. Recall that at its heart, the Burgess method involves breaking the range of the summand n into residue classes modulo an auxiliary prime p. One then averages over a set of such primes p, and it is crucial that the argument of the characters may be made independent of p (although the range of summation may still depend on p). More explicitly, fix a prime $p \nmid q$ and split the set of $n \in (N, N+H]$ into residue classes modulo p by writing $n = aq + pm$ with $0 \leq a < p$ and $m \in (N', N'+H']$ with $N' = (N - aq)/p$, $H' = H/p$. Then, for example, the multiplicative character sum (1.1) may be written as

$$S(N,H) = \sum_{0\leq a<p} \sum_{N'<m\leq N'+H'} \chi(aq+pm) = \chi(p) \sum_{0\leq a<p} \sum_{N'<m\leq N'+H'} \chi(m),$$

so that after averaging over a set $\mathscr{P}$ of primes,

$$|S(N,H)| \leq \frac{1}{|\mathscr{P}|} \sum_{p\in\mathscr{P}} \sum_{0\leq a<p} \left| \sum_{N'<m\leq N'+H'} \chi(m) \right|. \tag{1.4}$$

The Burgess argument then proceeds by manipulating the intervals of summation in order to reach a complete character sum that may be bounded (in most cases) by the Weil bound. This reveals a fundamental barrier quickly reached by a naive application of the Burgess method to the mixed character sum (1.3): it is not trivial to make the argument of the polynomial f independent of sufficiently many primes p, and without this independence, averaging over auxiliary primes as in (1.4) cannot proceed successfully.

For the case of f linear, Burgess [6] proved that for $f(n) = an/q$ with $0 < a < q$ and q prime,

$$|S(f; N, H)| \ll H^{1-\frac{1}{r}} q^{\frac{1}{4(r-1)}} (\log q)^2, \tag{1.5}$$

for any $r \geq 2$ and $0 < N, H < q$; this was later extended in [7] to the case $r = 3$ and q an arbitrary positive integer. A similar result was also proved by Friedlander and Iwaniec [11], as a consequence of more general bounds for weighted multiplicative character sums.

In a 1995 paper, Enflo [10] reported a nontrivial bound for $S(f; N, H)$ for f a real-valued polynomial of any degree d and $H = q^{1/2}$, with q prime. His proof introduced the idea of using Weyl differencing d times before applying the Burgess method, thus stripping off the exponential factor $e(f(n))$ entirely. This insight removes the problem of dependence on the auxiliary primes, and allows the Burgess method to proceed. A careful analysis of Enflo's method gives the following result:

Theorem 1.1. *Let f be a real-valued polynomial of degree d and χ a non-principal character to a prime modulus q. Then for any $r \geq 1$ and $H < q^{\frac{3}{4}+\frac{1}{4r}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,d,\varepsilon} H^{1-\frac{1}{2^d r}} q^{\frac{r+1}{2^{d+2}r^2}+\varepsilon},$$

uniformly in N.

As this result is surpassed by new methods, we do not give a proof here. Note that this recovers the original Burgess bound (1.2) in the case $d = 0$, and for any d it proves a nontrivial bound as long as $H > q^{1/4+\varepsilon}$. Note also that it is clear that an upper bound on H is required as soon as $d \geq 1$. For example, if $f(n) = n/q$ and $H = mq$ for some $m \geq 1$ then $S(f; N, H) = mG_q(\chi)$, where $G_q(\chi)$ is the Gauss sum. Then $|S(f; N, H)| = Hq^{-1/2}$ precisely, so it is not possible to attain a generic upper bound of the form $H^\alpha q^\beta$ with $\alpha < 1$ for arbitrary H.

More recently, Chang [9] introduced another idea that allows one to remove the dependence of $e(f(n))$ on the auxiliary primes p. Roughly speaking, the idea is to approximate $S(f; N, H)$ by $S(\tilde{f}; N, H)$, where $\tilde{f}$ has real coefficients that are sufficiently close to those of f but are independent of p. Chang's result improves on that of Enflo, proving that as soon as $H > q^{1/4+\kappa}$,

$$\sum_{0<n\leq H} e(f(n))\chi(n) \ll Hq^{-\delta}, \tag{1.6}$$

where

$$\delta = \frac{\kappa^2}{4((d+1)^2+2)(1+2\kappa)}. \tag{1.7}$$

(In fact Chang's results in [9] apply more generally to mixed character sums over $\mathbb{F}_{q^n}$ for any $n \geq 1$.) Chang furthermore proved in [8] a result for square-free q that is similar to (1.6), but with an additional factor $\tau(q)^{4(\log d)d^{-2}}$.

A refinement of Chang's argument improves the result to:

Theorem 1.2. *Let f be a real-valued polynomial of degree $d \geq 0$ and χ a non-principal character to a prime modulus q. Set*

$$D := \frac{d(d+1)}{2}. \tag{1.8}$$

Then if $r \geq 1$ and $H < q^{\frac{1}{2}+\frac{1}{4r}}$ we have

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,d} H^{1-\frac{1}{r}} q^{\frac{r+1+D}{4r^2}} (\log q)^2,$$

uniformly in N.

We shall use the notation (1.8) throughout the paper.

We do not claim Theorem 1.2 as substantially new; the small improvement is a consequence of approximating the coefficients of monomials in f more accurately for higher degree monomials; Chang approximates the coefficients with the same accuracy for every degree. Supposing that the result of Theorem 1.1 achieves its minimum at a value r_0, we may compare it to the result of Theorem 1.2 for $r = 2^d r_0$, and see that Theorem 1.2 is as strong for $d = 1, 2$ and stronger than Theorem 1.1 for $d \geq 3$. Additionally, note that for $H < q^{\frac{1}{2}+\frac{1}{4r}}$, the bound of Theorem 1.2 is nontrivial only if $r \geq 1 + D$.

If $H = q^{\frac{1}{4}+\kappa}$ for some small $\kappa > 0$, then Theorem 1.2 yields a nontrivial bound $Hq^{-\delta}$ where δ behaves approximately like

$$\delta = \frac{\kappa^2}{D+1}, \tag{1.9}$$

for sufficiently small κ and sufficiently large d, and hence is approximately a factor of 8 better than (1.7). (See Section 3.2 for details.)

The novelty of this paper appears in the following strategy: by choosing the coefficients of $\tilde{f}$ according to a certain grid, we are able to introduce a nontrivial auxiliary averaging that leads to a bound involving the number $J_{r,d}(X)$ occuring in Vinogradov's mean value theorem. This is the number of solutions to the system of Diophantine equations given by

$$x_1^m + \cdots + x_r^m = x_{r+1}^m + \cdots + x_{2r}^m, \qquad 1 \leq m \leq d,$$

where d is the degree of f and $1 \leq x_1, \ldots, x_{2r} \leq X$. The celebrated new results of Wooley (most recently [15] [16]) on Vinogradov's mean value theorem provide exceptionally sharp bounds for $J_{r,d}(X)$ and lead to a significant improvement on Theorem 1.2.

Let us recall the main conjecture in the setting of Vinogradov's mean value theorem:

Conjecture 1.1. *For every $r \geq 1, d \geq 1$ and $\varepsilon > 0$,*

$$J_{r,d}(X) \ll_{r,d,\varepsilon} X^{\varepsilon}(X^r + X^{2r-D}). \tag{1.10}$$

Conditional on this bound for $J_{r,d}(X)$ we prove our main result:

Theorem 1.3. *Let f be a real-valued polynomial of degree $d \geq 1$ and χ a non-principal character to a prime modulus q. Assume Conjecture 1.1 holds. Then for integers $r > D$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-D)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,\varepsilon} H^{1-\frac{1}{r}} q^{\frac{r+1-D}{4r(r-D)}+\varepsilon}, \tag{1.11}$$

uniformly in N, for any $\varepsilon > 0$.

The method of proof for Theorem 1.3 also yields character sum bounds (conditional on Conjecture 1.1) in the range $r \leq D$, but it turns out that these bounds are no better than trivial. Note that the $d = 0$ case of (1.11) would recover the classical Burgess bound (1.2). For fixed d, in the limit as $r \to \infty$, the bound (1.11) is nontrivial for $H \geq q^{1/4+\varepsilon}$. A direct comparison shows that (1.11) matches Theorem 1.2 when $r = D + 1$ (though the admissible range for H is longer), and is sharper as soon as $r > D + 1$.

If $H = q^{\frac{1}{4}+\kappa}$ for some small $\kappa > 0$, then Theorem 1.3 would yield a nontrivial bound $Hq^{-\delta}$ where δ behaves approximately like

$$\delta = \left(\frac{2\kappa}{1 + \sqrt{1 + 4D\kappa}} \right)^2 . \tag{1.12}$$

(See Section 4.2 for details.) For any fixed d, as $\kappa \to 0$, this behaves like

$$\delta = \kappa^2,$$

which we note is independent of d, and is in fact as strong as the original Burgess bound for multiplicative character sums.

Note that for $d = 1, 2$, the bound of Conjecture 1.1 holds true trivially, for all $r \geq 1$. Thus the following are immediate corollaries of Theorem 1.3:

Theorem 1.4. *Let f be a linear real-valued polynomial and χ a non-principal character to a prime modulus q. Then for $r \geq 2$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-1)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,\varepsilon} H^{1-\frac{1}{r}} q^{\frac{1}{4(r-1)}+\varepsilon},$$

uniformly in N, for any $\varepsilon > 0$.

Note that this generalizes the result (1.5) since f may now be any real-valued linear polynomial.

Theorem 1.5. *Let f be a quadratic real-valued polynomial and χ a non-principal character to a prime modulus q. Then for $r \geq 4$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-3)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll H^{1-\frac{1}{r}} q^{\frac{r-2}{4r(r-3)}+\varepsilon},$$

uniformly in N, for any $\varepsilon > 0$.

Recent breakthroughs of Wooley have provided very strong results toward Conjecture 1.1. At the time of writing, the conjecture is now known to hold for all r if $d = 3$ and for $r \geq d(d-1)$ when $d \geq 4$ (see [16]), and for 100% of the critical interval $1 \leq r \leq D$ (see [15]). In our application, the results of Wooley for large r make the following cases of Theorem 1.3 unconditional.

Theorem 1.6. *Let f be a real-valued polynomial of degree 3 and χ a non-principal character to a prime modulus q. Then for $r \geq 7$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-6)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,\varepsilon} H^{1-\frac{1}{r}} q^{\frac{r-5}{4r(r-6)}+\varepsilon},$$

uniformly in N, for any $\varepsilon > 0$.

For $d \geq 4$, we have:

Theorem 1.7. *Let f be a real-valued polynomial of degree $d \geq 4$ and χ a non-principal character to a prime modulus q. Then for $r \geq d(d-1)$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-D)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,\varepsilon} H^{1-\frac{1}{r}} q^{\frac{r+1-D}{4r(r-D)}+\varepsilon},$$

uniformly in N, for any $\varepsilon > 0$.

Finally, in the intermediate range $D < r < d(d-1)$, we apply the so-called approximate main conjecture of [15], which states that for all $d \geq 4$,

$$J_{r,d}(X) \ll X^{\Delta_{r,d}}(X^r + X^{2r-D})$$

where $\Delta_{r,d} = O(d)$ (see Theorem 1.5 of [15]). This results in the following:

Theorem 1.8. *Let f be a real-valued polynomial of degree $d \geq 4$ and χ a non-principal character to a prime modulus q. Then for $D < r < d(d-1)$ and $H < q^{\frac{1}{2}+\frac{1}{4(r-D+\Delta)}}$ we have*

$$\sum_{N<n\leq N+H} e(f(n))\chi(n) \ll_{r,\varepsilon} H^{1-1/r} q^{\frac{r+1-D+2\Delta}{4r(r-D+\Delta)}+\varepsilon},$$

where

$$\Delta = \Delta_{r,d} = O(d)$$

is as specified in [15].

We have stated these results in terms of polynomials $f(n)$. However it is clear in principle that one can prove estimates for suitable general real-valued functions $f(n)$ by approximating them by appropriate polynomials. Moreover, these methods can be extended to certain multi-variable sums. We intend to return to this issue in the near future.

Although in this paper we shall confine ourselves to prime moduli q, most of our results can be modified to apply to general square-free moduli. In some cases however we cannot handle the full range $r > D$ occuring in Theorem 1.3. We leave the details to the reader.

For our proofs it will be convenient to assume that $d \geq 1$. This enables us to replace the use of the Menchov-Rademacher device (originating in [13], [14]) by the simpler "partial summation by Fourier series" of Bombieri and Iwaniec [1]. Of course Theorem 1.2 remains true for $d = 0$, since it reduces to Burgess's bound (1.2).

2 The Burgess method with coefficient approximation

To begin the proof of Theorems 1.2 and 1.3, we consider

$$T_d(N,H,\chi) = T(N,H) = \sup_{\deg(f)=d}\ \sup_{K\le H}\left|\sum_{N<n\le N+K} e(f(n))\chi(n)\right|,$$

where f runs over real-valued polynomials and χ is a non-principal multiplicative character to a prime modulus q. We first note that $T(N,H)$ has period q with respect to N, so that we can assume from now on that $0 \le N < q$.

Fix a set of primes $\mathscr{P} = \{P < p \le 2P\}$ for some parameter $P \le H$ that we will choose later. Since $H = o(q)$ in all our theorems we will have $p \nmid q$ for $p \in \mathscr{P}$. Hence we can split $n \in (N, N+K]$ into residue classes modulo p by writing $n = aq + pm$ with $0 \le a < p$. This produces values $m \in (N_{a,p}, N_{a,p} + K_{a,p}]$ with $N_{a,p} = (N - aq)/p$ and $K_{a,p} = K/p \le H/P$. Then

$$\sum_{N<n\le N+K} e(f(n))\chi(n) = \sum_{0\le a<p}\ \sum_{N_{a,p}<m\le N_{a,p}+K_{a,p}} e(f(aq+pm))\chi(aq+pm),$$

and as a result

$$T(N,H) \le \sum_{0\le a<p} T(N_{a,p}, H/P).$$

We proceed to average over $\mathscr{P}$, producing

$$T(N,H) \le |\mathscr{P}|^{-1} \sum_{p\in\mathscr{P}}\ \sum_{0\le a<p} T(N_{a,p}, H/P). \tag{2.1}$$

We now use the following lemma.

Lemma 2.1. *For any real number $L \ge 1$ we have*

$$T(U,L) \le 4L^{-1} \sum_{U-L<m\le U} T(m, 2L). \tag{2.2}$$

To see this, note that

$$T(U,L) = \left|\sum_{U<n\le U+K} e(f(n))\chi(n)\right|$$

for some polynomial f and some positive real number $K \le L$. Moreover if $U - L < m \le U$ then

$$\sum_{U<n\le U+K} e(f(n))\chi(n) = \sum_{m<n\le U+K} e(f(n))\chi(n) - \sum_{m<n\le U} e(f(n))\chi(n),$$

whence

$$\left|\sum_{U<n\le U+K} e(f(n))\chi(n)\right| \le 2T(m,2L),$$

since $U+K\le m+2L$. The result then follows since the interval $(U-L,U]$ contains at least $L/2$ integers m.

Applying (2.2) to (2.1) with $U=N_{a,p}$ and $L=H/P$, we may conclude that

$$\begin{aligned} T(N,H) &\ll |\mathscr{P}|^{-1}(H/P)^{-1}\sum_{p\in\mathscr{P}}\sum_{0\le a<p}\sum_{N_{a,p}-H/P<m\le N_{a,p}} T(m,2H/P) \\ &\ll H^{-1}(\log q)\sum_{p\in\mathscr{P}}\sum_{0\le a<p}\sum_{N_{a,p}-H/P<m\le N_{a,p}} T(m,2H/P), \end{aligned}$$

on noting that $|\mathscr{P}|\gg P(\log P)^{-1}\gg P(\log q)^{-1}$. We now define

$$\mathcal{A}(m)=\#\left\{(a,p):\frac{N-aq}{p}-\frac{H}{P}<m\le\frac{N-aq}{p}\right\},$$

which allows us to write

$$T(N,H)\ll H^{-1}(\log q)\sum_{m\in\mathbb{Z}}\mathcal{A}(m)T(m,2H/P). \tag{2.3}$$

We then set

$$S_1=\sum_m \mathcal{A}(m)$$

and

$$S_2=\sum_m \mathcal{A}(m)^2,$$

and we note the following facts, which we will prove in Section 5.

Lemma 2.2. *We have $\mathcal{A}(m)=0$ unless $|m|\le 2q$. Moreover if $HP<q$ then $S_1\le S_2\ll HP$.*

From a repeated application of Hölder's inequality, it then follows from (2.3) that

$$\begin{aligned} T(N,H) &\ll H^{-1}(\log q)S_1^{1-\frac{1}{r}}S_2^{\frac{1}{2r}}\left\{\sum_{|m|\le 2q}T(m,2H/P)^{2r}\right\}^{\frac{1}{2r}} \\ &\ll H^{-\frac{1}{2r}}P^{1-\frac{1}{2r}}(\log q)\left\{\sum_{|m|\le 2q}T(m,2H/P)^{2r}\right\}^{\frac{1}{2r}}. \end{aligned}$$

As previously noted, the function $T(m,K)$ is periodic in m, with period q, so that in fact we have

$$T(N,H)\ll H^{-\frac{1}{2r}}P^{1-\frac{1}{2r}}(\log q)\left\{\sum_{m=1}^{q}T(m,2H/P)^{2r}\right\}^{\frac{1}{2r}}. \tag{2.4}$$

For any M and $K > 0$ we now define

$$T_0(M,K) = \sup_{\deg(f)=d} \left| \sum_{M<n\le M+K} e(f(n))\chi(n) \right|.$$

We can relate $T(M,K)$ to $T_0(M,K)$ using the following lemma, which is an immediate consequence of Lemma 2.2 of Bombieri and Iwaniec [1].

Lemma 2.3. *Let a_n be a sequence of complex numbers supported on the integers $n \in (A, A+B]$, and let I be any subinterval of $(A, A+B]$. Then*

$$\sum_{n\in I} a_n \ll \big(\log(B+2)\big) \sup_{\theta\in\mathbb{R}} \left| \sum_{A<n\le A+B} a_n e(\theta n) \right|.$$

Thus if $d \ge 1$ and $K \le q$ then

$$T(M,K) \ll T_0(M,K)\log(K+2) \ll T_0(M,K)\log q.$$

This is the only place in the argument where the condition $d \ge 1$ is used. We now see that (2.4) becomes

$$T(N,H) \ll H^{-\frac{1}{2r}} P^{1-\frac{1}{2r}} (\log q)^2 S_3(2H/P)^{\frac{1}{2r}}, \tag{2.5}$$

where we have set

$$S_3(K) = \sum_{m=1}^{q} T_0(m,K)^{2r}.$$

We proceed to develop a bound for $S_3(K)$, under the assumption that $K \le q$. Having removed the maximum over the length of our intervals we now handle the maximum over the polynomials f. In effect we do this by replacing the maximum by a sum over all "distinct" polynomials modulo 1. The principle here is that two polynomials will be effectively equivalent if their coefficients are sufficiently close.

Let $Q \ge K$ be an integer parameter to be chosen in due course. We partition $[0,1]^{d+1}$ into boxes B_α of side-length Q^{-j} in the j-th coordinate, for $j = 0,\ldots,d$. Note that the total number of boxes is Q^D. For each box B_α, fix $\theta_\alpha = (\theta_{\alpha,0},\ldots,\theta_{\alpha,d})$ to be the vertex of B_α with the least value in each coordinate. Thus each θ_α takes the form

$$(c_0 Q^{-0}, c_1 Q^{-1}, \ldots, c_d Q^{-d})$$

for some integers $0 \le c_j \le Q^j - 1$, $0 \le j \le d$. (Chang's original argument [9] chooses the boxes to be of side-length Q^{-d} in all coordinates, and allows θ_α to be any point in the box B_α.) Define for any $\theta \in [0,1]^{d+1}$ the polynomial

$$\theta(X) := \sum_{j=0}^{d} \theta_j X^j.$$

For any integer m, positive real number t, and index α, set

$$T(\alpha; m, t) := \left| \sum_{0<n\le t} e(\theta_\alpha(n))\chi(n+m) \right|.$$

We use these sums to approximate $T_0(m, K)$ as follows.

Lemma 2.4. *Given an integer m and real numbers $Q \ge K > 0$, there is an index α such that*

$$T_0(m, K) \ll_d T(\alpha; m, K) + K^{-1} \int_0^K T(\alpha; m, t)dt.$$

To prove this we observe that for integral m we have

$$\begin{aligned} T_0(m, K) &= \sup_{\deg(f)=d} \left| \sum_{m<n\le m+K} e(f(n))\chi(n) \right| \\ &= \sup_{\deg(f)=d} \left| \sum_{0<n\le K} e(f(n))\chi(n+m) \right|. \end{aligned}$$

Suppose then that

$$T_0(m, K) = \left| \sum_{0<n\le K} e(f(n))\chi(n+m) \right|$$

for some polynomial f of degree d, and write $f(X) = f_d X^d + \ldots + f_0$. Clearly we may assume that $0 \le f_j \le 1$ for $0 \le j \le d$. We then choose α so that $|f_j - \theta_{\alpha,j}| \le Q^{-j}$ for each index j and temporarily write $\delta_j = f_j - \theta_{\alpha,j}$ for notational convenience. Then, by summation by parts, we have

$$\begin{aligned} &\sum_{0<n\le K} e(f(n))\chi(n+m) \\ &\qquad = \sum_{n\le K} e\left(\sum_{j=0}^{d} \delta_j n^j\right) e(\theta_\alpha(n))\chi(n+m) \\ &\qquad = e\left(\sum_{j=0}^{d} \delta_j K^j\right) \sum_{n\le K} e(\theta_\alpha(n))\chi(n+m) \\ &\qquad\qquad - \int_0^K \left\{ \sum_{n\le t} e(\theta_\alpha(n))\chi(n+m) \right\} \frac{d}{dt} e\left(\sum_{j=0}^{d} \delta_j t^j\right) dt. \end{aligned}$$

Since $|\delta_j| \le Q^{-j}$ we have

$$\left|\frac{d}{dt}e\left(\sum_{j=0}^{d}\delta_j t^j\right)\right| \le 2\pi\sum_{j=1}^{d} j|\delta_j|t^{j-1} \le 2\pi\sum_{j=1}^{d} jQ^{-j}K^{j-1},$$

for $0 \le t \le K$. Thus if $Q \gg K$ we have

$$\left|\frac{d}{dt}e\left(\sum_{j=0}^{d}\delta_j t^j\right)\right| \ll_d K^{-1}$$

and hence

$$\sum_{n\le K} e(f(n))\chi(n+m) \ll_d T(\alpha;N,K) + K^{-1}\int_0^K T(\alpha;N,t)dt,$$

which proves the lemma.

An application of Hölder's now allows us to deduce from Lemma 2.4 that

$$T_0(m,K)^{2r} \ll_d T(\alpha;m,K)^{2r} + K^{-1}\int_0^K T(\alpha;m,t)^{2r}dt$$

for some index α depending on m and K. This dependence is rather awkward, and we circumvent it in the most trivial way by summing over all available indices α, giving

$$T_0(m,K)^{2r} \ll_d \sum_{\alpha} T(\alpha;m,K)^{2r} + K^{-1}\sum_{\alpha}\int_0^K T(\alpha;m,t)^{2r}dt.$$

Thus

$$S_3(K) \ll_d S_4(K) + K^{-1}\int_0^K S_4(t)dt \tag{2.6}$$

if $0 < K \ll Q$, where we have defined

$$S_4(\tau) = \sum_{\alpha}\sum_{m=1}^{q} T(\alpha;m,\tau)^{2r}.$$

Thus we now turn our attention to bounding the sum $S_4(\tau)$. Recall the definition of the boxes B_α, and in particular the definition of the vertices θ_α. If $\mathbf{x} = (x_1,\ldots,x_{2r})$ we write

$$\Sigma_A(\mathbf{x};q) = \sum_{\alpha} e\left(\sum_{i=1}^{2r}\varepsilon(i)\theta_\alpha(x_i)\right),$$

where $\varepsilon(i) = (-1)^i$. We also set

$$\Sigma_B(\mathbf{x};\chi,q) = \sum_{m=1}^{q}\chi(F_{\mathbf{x}}(m))$$

where the polynomial $F_{\mathbf{x}}(X)$ is defined by

$$F_{\mathbf{x}}(X) = \prod_{i=1}^{2r} (X + x_i)^{\delta_q(i)}. \tag{2.7}$$

Here $\delta_q(i) = 1$ if i is even and $= \Delta(q) - 1$ if i is odd, where $\Delta(q)$ is the order of the character χ modulo q.

With this notation we then see upon expanding the sum that

$$S_4(\tau) = \sum_{\alpha} \sum_{m=1}^{q} T(\alpha; m, \tau)^{2r} = \sum_{\substack{\mathbf{x} \\ 0 < x_i \le \tau}} \Sigma_A(\mathbf{x}; q) \Sigma_B(\mathbf{x}; \chi, q). \tag{2.8}$$

We will first prove Theorem 1.2 by averaging trivially over the boxes B_α and running the Weil bound argument that is typically found in applications of the Burgess method. The key proposition for Theorem 1.2 is:

Proposition 2.1. *Suppose q is prime. Then for any $\tau \le q$ we have*

$$S_4(\tau) = \sum_{\alpha} \sum_{m=1}^{q} T(\alpha; m, \tau)^{2r} \ll_r Q^D (\tau^r q + \tau^{2r} q^{1/2}). \tag{2.9}$$

Second, we will improve on this by averaging nontrivially over the boxes B_α, resulting in the key proposition for Theorem 1.3:

Proposition 2.2. *Suppose q is prime. Then for any $\tau \le q$ we have*

$$S_4(\tau) = \sum_{\alpha} \sum_{m=1}^{q} T(\alpha; m, \tau)^{2r} \ll_r Q^D \left(\tau^r q + J_{r,d}(\tau) q^{1/2} \right). \tag{2.10}$$

The propositions will be proved and the resulting theorems deduced in Sections 3 and 4, respectively. Although Proposition 2.1 is an immediate consequence of Proposition 2.2 we have chosen to state and prove Proposition 2.1 separately, in order to highlight the different aspects of our treatment.

3 The multiplicative component

We first consider the multiplicative character sum $\Sigma_B(\mathbf{x}; \chi, q)$. The well-known Weil bound implies the following:

Lemma 3.1. *Let χ be a character of order $\Delta(q) > 1$ modulo a prime q. Suppose that $F(X)$ is a polynomial which is not a perfect $\Delta(q)$-th power over $\overline{\mathbb{F}}_q[X]$. Then*

$$\left| \sum_{m=1}^{q} \chi(F(m)) \right| \le (\deg(F) - 1)\sqrt{q}.$$

We can apply Lemma 3.1 to show that $\Sigma_B(\mathbf{x};\chi,q)$ is bounded by $O_r(q^{1/2})$, unless the polynomial $F_{\mathbf{x}}(X)$ is a perfect $\Delta(q)$-th power over $\overline{\mathbb{F}}_q$. We define $\mathbf{x}=(x_1,\ldots,x_{2r})$ to be bad if for all $i=1\ldots,2r$, there exists $j\neq i$ such that $x_j=x_i$, and $\mathbf{x}$ to be good otherwise. We take $\mathcal{B}(\tau)$ to be the collection of bad $\mathbf{x}$ with $0<x_i\leq\tau$ and similarly $\mathcal{G}(\tau)$ to be the collection of good $\mathbf{x}$ with $0<x_i\leq\tau$. The following is immediate:

Lemma 3.2. *There are at most $r^{2r+1}\tau^r$ bad $\mathbf{x}$ with $0<x_i\leq\tau$, so that*

$$\#\mathcal{B}(\tau)\ll_r\tau^r. \tag{3.1}$$

For the proof of the lemma we write the set $\{x_1,\ldots,x_{2r}\}$ without repetitions as $\{y_1,\ldots,y_t\}$, say, where $t\leq r$ since $\mathbf{x}$ is bad. We may suppose that the y_i are arranged in ascending order. There are at most rK^r choices for such a set $\{y_1,\ldots,y_t\}$, and at most r^{2r} choices for $\mathbf{x}$ which correspond to each such set. This suffices for the lemma.

Furthermore:

Lemma 3.3. *Fix $\mathbf{x}$ with $0<x_i\leq\tau$ for each $i=1,\ldots,2r$ and fix a prime q. If $\tau\leq q$ and $F_{\mathbf{x}}(X)$ is a perfect $\Delta(q)$-th power modulo q, then $\mathbf{x}$ is bad.*

This is obvious since if there were only one index i for which x_i takes a given value y say, then the factor $X+y$ occurs in $F_{\mathbf{x}}(X)$ with multiplicity either 1 or $\Delta(q)-1$, neither of which is divisible by $\Delta(q)$.

If $\mathbf{x}$ is bad, we will apply the trivial bound $O(q)$ to $\Sigma_B(\mathbf{x};\chi,q)$; we may conclude from (3.1) that

$$\sum_{\mathbf{x}\in\mathcal{B}(\tau)}\left|\sum_{m=1}^{q}\chi(F_{\mathbf{x}}(m))\right|\ll_r\tau^r q. \tag{3.2}$$

For good $\mathbf{x}$ we may apply Lemmas 3.1 and 3.3 to obtain the following standard result.

Lemma 3.4. *If q is prime and $\tau\leq q$ then*

$$\sum_{\mathbf{x}\in\mathcal{G}(\tau)}|\sum_{m=1}^{q}\chi(F_{\mathbf{x}}(m))|\ll_r\tau^{2r}q^{1/2}. \tag{3.3}$$

3.1 Proof of Theorem 1.2

At this point we may prove Proposition 2.1. Using the trivial bound

$$|\Sigma_A(\mathbf{x};q)|\leq Q^D$$

in (2.8), we observe that

$$\begin{aligned}&\sum_{\alpha}\sum_{m=1}^{q}T(\alpha;m,\tau)^{2r}\\ &\quad\leq\; Q^D\left(\sum_{\mathbf{x}\in\mathcal{G}(\tau)}\left|\sum_{m=1}^{q}\chi(F_{\mathbf{x}}(m))\right|+\sum_{\mathbf{x}\in\mathcal{B}(\tau)}\left|\sum_{m=1}^{q}\chi(F_{\mathbf{x}}(m))\right|\right).\end{aligned}$$

We substitute the bounds (3.3) and (3.2) to complete the proof of Proposition 2.1. Applying Proposition 2.1 to $S_4(K)$ and $S_4(t)$ in (2.6), we may conclude that for any $K \leq q$ we have

$$S_3(K) \ll_{r,d} Q^D(K^{2r}q^{1/2} + K^r q)$$

so long as the integer Q is at least K. We apply this in (2.5) with $K = 2H/P$ and $Q = \lceil 2H/P \rceil$, obtaining

$$T(N,H) \ll_{r,d} H^{-\frac{1}{2r}} P^{1-\frac{1}{2r}} (\log q)^2 (H/P)^{\frac{D}{2r}} \left((H/P)^{2r} q^{1/2} + (H/P)^r q \right)^{\frac{1}{2r}}.$$

We then extract the best result by choosing P such that

$$\frac{1}{2} H q^{-1/(2r)} \leq P \leq H q^{-1/(2r)}.$$

The restriction $HP < q$ of Lemma 2.2 is then satisfied when $H < q^{\frac{1}{2}+\frac{1}{4r}}$, and we will also have $2H/P \leq q$ for sufficiently large q. We therefore obtain the result of Theorem 1.2 in the form

$$T(N,H) \ll_{r,d} H^{1-\frac{1}{r}} q^{\frac{r+1+D}{4r^2}} (\log q)^2.$$

3.2 Optimal choice of r

Recall that we have set

$$D = \frac{1}{2} d(d+1).$$

We observe that if $H = q^{\frac{1}{4}+\kappa}$ for small $\kappa > 0$, then the bound of Theorem 1.2 is of the form $Hq^{-\delta}$ where

$$\delta = \frac{\kappa r - \frac{1}{4}(D+1)}{r^2}.$$

As a function of r, this attains a maximum at the real value

$$r(\kappa, d) := \frac{\frac{1}{2}(D+1)}{\kappa}.$$

Upon choosing the closest integer $r = r(\kappa, d) + \theta$ where $-1/2 < \theta \leq 1/2$, we compute that for this choice of r we have

$$\delta = \kappa^2 \left(\frac{\frac{1}{4}(D+1) + \kappa\theta}{\frac{1}{4}(D+1)^2 + (d+1)\kappa\theta + \kappa^2\theta^2} \right).$$

For sufficiently small κ this behaves like

$$\delta = \frac{\kappa^2}{D+1}.$$

4 Introduction of the Vinogradov bounds

We improve on the strategy of Theorem 1.2 by treating the additive character sum $\Sigma_A(\mathbf{x};q)$ in (2.8) nontrivially. Recalling the definition of the vector $\theta_\alpha = (\theta_{\alpha,1}, \theta_{\alpha,2}, \ldots, \theta_{\alpha,d})$, we see that

$$\begin{aligned}\sum_\alpha e\left(\sum_{i=1}^{2r} \varepsilon(i)\theta_\alpha(x_i)\right) &= \sum_\alpha e\left(\theta_{\alpha,1}\sum_{i=1}^{2r}\varepsilon(i)x_i + \cdots + \theta_{\alpha,d}\sum_{i=1}^{2r}\varepsilon(i)x_i^d\right) \\ &= \prod_{s=1}^{d}\left(\sum_{c=1}^{Q^s} e\left(\frac{c\sum_{i=1}^{2r}\varepsilon(i)x_i^s}{Q^s}\right)\right) \\ &= Q^D \Xi_Q(\mathbf{x}),\end{aligned}$$

say, where $\Xi_Q(\mathbf{x})$ is the indicator function for the set

$$\{\mathbf{x} = (x_1, \ldots, x_{2r}) \in \mathbb{N}^{2r} \cap (0,\tau]^{2r} : \sum_{i=1}^{2r} \varepsilon(i)x_i^s \equiv 0 \pmod{Q^s},\ \forall s \leq d\}.$$

Our application has $0 \leq \tau \leq K$ in (2.6), and $Q \geq K$ in Lemma 2.4. Moreover we will be taking $K = 2H/P$ in (2.5). Any integer $Q \geq 2H/P$ is therefore acceptable. In the definition of $\Xi_Q(\mathbf{x})$ we will have

$$\left|\sum_{i=1}^{2r} \varepsilon(i)x_i^s\right| < 2r\tau^s \leq 2rK^s \leq (2rK)^s = (4rH/P)^s.$$

Thus, by taking $Q = \lceil 4rH/P \rceil$, the congruences in the set above can hold only if they are actually equalities in $\mathbb{Z}$. We may then replace $\Xi_Q(\mathbf{x})$ by the indicator function $\Xi(\mathbf{x})$ of the set

$$V_{r,d}(\tau) := \{\mathbf{x} = (x_1, \ldots, x_{2r}) \in \mathbb{N}^{2r} \cap (0,\tau]^{2r} : \sum_{i=1}^{2r} \varepsilon(i)x_i^s = 0,\ \forall s \leq d\}.$$

Then we see that (2.8) may be bounded by

$$\sum_\alpha \sum_{m=1}^{q} T(\alpha; m, \tau)^{2r} \leq Q^D\{\Sigma(\mathcal{G}) + \Sigma(\mathcal{B})\},$$

where

$$\Sigma(\mathcal{G}) = \sum_{\mathbf{x}\in\mathcal{G}(\tau)\cap V_{r,d}(\tau)} \left|\sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m))\right|$$

and

$$\Sigma(\mathcal{B}) = \sum_{\mathbf{x}\in\mathcal{B}(\tau)\cap V_{r,d}(\tau)} \left|\sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m))\right|.$$

We now prove Proposition 2.2. Lemma 3.3 shows that $F_{\mathbf{x}}(X)$ is not a perfect $\Delta(q)$-th power modulo q for $\mathbf{x} \in \mathcal{G}(\tau)$ and $\tau \leq q$, and then Lemma 3.1 yields

$$\sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m)) \ll_r q^{1/2}.$$

We expect $\mathbf{x}$ to be good generically, so we will apply the upper bound

$$\#(\mathcal{G}(\tau) \cap V_{r,d}(\tau)) \leq \#V_{r,d}(\tau) = J_{r,d}(\tau),$$

whence

$$\Sigma(\mathcal{G}) = \sum_{\mathbf{x} \in \mathcal{G}(\tau) \cap V_{r,d}(\tau)} \left| \sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m)) \right| \ll_r J_{r,d}(\tau) q^{1/2}. \tag{4.1}$$

For $\mathbf{x} \in \mathcal{B}(K)$ we use (3.2) to deduce that

$$\Sigma(\mathcal{B}) = \sum_{\mathbf{x} \in \mathcal{B}(\tau) \cap V_{r,d}(\tau)} \left| \sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m)) \right| \leq \sum_{\mathbf{x} \in \mathcal{B}(\tau)} \left| \sum_{m=1}^{q} \chi(F_{\mathbf{x}}(m)) \right| \ll_r \tau^r q.$$

Proposition 2.2 then follows.

4.1 Proof of Theorem 1.3

We proceed to prove Theorem 1.3. Assuming that Conjecture 1.1 holds, we see from Proposition 2.2 that

$$S_4(\tau) \ll_{r,d,\varepsilon} Q^D \left\{ (\tau^r + \tau^{2r-D}) q^{1/2} + \tau^r q \right\} q^{\varepsilon}. \tag{4.2}$$

If $r \leq D$, the contribution of bad $\mathbf{x}$ dominates, and we cannot obtain a nontrivial bound. Thus from now on we only consider $r > D$. Since d is then bounded in terms of r, the implied constant in the $\ll_{r,d,\varepsilon}$ notation may be bounded as a function of r and ε alone. We now apply (4.2) to (2.6) to conclude that for any $1 \leq K \leq q$ we have

$$S_3(K) \ll_{r,\varepsilon} Q^D (K^{2r-D} q^{1/2} + K^r q) q^{\varepsilon}.$$

We apply this to (2.5) to obtain

$$T(N,H) \ll_{r,\varepsilon} H^{-\frac{1}{2r}} P^{1-\frac{1}{2r}} Q^{\frac{D}{2r}} \left(K^{2r-D} q^{1/2} + K^r q \right)^{\frac{1}{2r}} q^{\varepsilon}.$$

As before we take $K = 2H/P$ and $Q = \lceil 4rH/P \rceil$. It is optimal to choose P to balance the last two terms by taking

$$\frac{1}{2} H q^{-\frac{1}{2(r-D)}} \leq P < H q^{-\frac{1}{2(r-D)}}. \tag{4.3}$$

We may then satisfy the requirement $HP < q$ of Lemma 2.2 by restricting $H < q^{\frac{1}{2}+\frac{1}{4(r-D)}}$; the requirement $2H/P \leq q$ holds for sufficiently large q. Then

$$T(N,H) \ll_{r,\varepsilon} H^{1-1/r} q^{\frac{r+1-D}{4r(r-D)}+\varepsilon}.$$

This completes the proof of Theorem 1.3.

As already noted, Theorems 1.4 through 1.6 hold because Conjecture 1.1 is trivially true for $d = 1, 2$ and is now known to be true for $d = 3$ by recent results of Wooley [16]. For $d \geq 4$ Wooley [16], [15] has proved the following results towards Conjecture 1.1:

Proposition 4.1. *For $d \geq 4$ and $r \geq d(d-1)$,*

$$J_{r,d}(X) \ll_{r,\varepsilon} X^{\varepsilon}(X^r + X^{2r-D}). \tag{4.4}$$

For $d \geq 4$ and $D < r < d(d-1)$ then

$$J_{r,d}(X) \ll_r X^{2r-D+\Delta}, \tag{4.5}$$

where the order of magnitude of $\Delta = \Delta(r,d)$ is $O(d)$, as specified in [15].

The result (4.4) immediately implies Theorem 1.7. Theorem 1.8 follows from applying (4.5) in Proposition 2.2 to deduce that

$$\sum_{\alpha} \sum_{m=1}^{q} T(\alpha; m, \tau)^{2r} \ll_{r,\varepsilon} Q^D \left(\tau^{2r-D+\Delta} q^{1/2} + \tau^r q\right) q^{\varepsilon}.$$

The argument then proceeds as before, after choosing P such that

$$\frac{1}{2} H q^{-\frac{1}{2(r-D+\Delta)}} \leq P < H q^{-\frac{1}{2(r-D+\Delta)}}$$

in place of (4.3).

4.2 A note on δ

We remark that if $H = q^{1/4+\kappa}$ for some small $\kappa > 0$ then Theorem 1.3 would give a nontrivial bound $Hq^{-\delta}$ where

$$\delta = \frac{4\kappa(r-D) - 1}{4r(r-D)}.$$

As a function of r this attains a maximum at the real value

$$r_{\kappa,d} := D + \frac{1 + \sqrt{4D\kappa + 1}}{4\kappa}.$$

We choose r to be an integer $r = r_{\kappa,d} + \theta$ with $-1/2 < \theta \leq 1/2$, and for this choice, δ is approximately

$$\delta = \left(\frac{2\kappa}{1 + \sqrt{1 + 4D\kappa}}\right)^2.$$

For any fixed d, as $\kappa \to 0$, this behaves like

$$\delta = \kappa^2,$$

which we note is independent of d.

5 Proof of Lemma 2.2

This is merely a generalization of the proof in Section 4 of [12]. The first property in Lemma 2.2 is a direct result of the definition of $\mathcal{A}(m)$, on using our assumption that $0 \le N \le q$.

For the second property we first note that $\mathcal{A}(m) \le \mathcal{A}(m)^2$ since $\mathcal{A}(m)$ is a non-negative integer. It follows that $S_1 \le S_2$.

We now observe that $\mathcal{A}(m)^2$ counts quadruples (p, p', a, a') for which

$$m \le \frac{N - aq}{p} < m + H/P, \quad m \le \frac{N - a'q}{p'} < m + H/P.$$

For such a quadruple we must have

$$\left| \frac{N - aq}{p} - \frac{N - a'q}{p'} \right| \le H/P.$$

Under this condition there are $O(H/P)$ corresponding values of m. It follows that

$$\begin{aligned} \sum_m \mathcal{A}(m)^2 &\ll HP^{-1} \# \{p, p', a, a' : 0 \le \left| \frac{N - aq}{p} - \frac{N - a'q}{p'} \right| \le H/P\} \\ &\ll HP^{-1} \sum_{p, p' \in \mathscr{P}} \mathcal{M}(p, p'), \end{aligned} \tag{5.1}$$

where

$$\mathcal{M}(p, p') = \# \{a \pmod{p}, a' \pmod{p'} : 0 \le \left| \frac{N - aq}{p} - \frac{N - a'q}{p'} \right| \le H/P\}.$$

First consider the case $p = p'$. Then

$$|a - a'| \le \frac{Hp}{Pq} \le \frac{2H}{q} < 1,$$

since $H = o(q)$ in all our theorems. Thus $a = a'$ so that $\mathcal{M}(p, p) \ll P$ and hence $\sum_{p = p' \in \mathscr{P}} \mathcal{M}(p, p') \ll P^2$, which makes an satisfactory contribution to (5.1).

Next, consider the case $p \ne p'$. We choose (by Bertrand's postulate) a prime l such that

$$\frac{q}{H} < l \le \frac{2q}{H}.$$

(Here we use the fact that $H < q$ for large enough q.) Let $M = \left[\frac{Nl}{q}\right]$ or $1+\left[\frac{Nl}{q}\right]$ be chosen so that $l \nmid M$. Then $|Nl/q - M| \leq 1$ implies that $|N - qM/l| \leq q/l$, so that

$$\left|\frac{qM/l - aq}{p} - \frac{qM/l - a'q}{p'}\right| \leq \frac{H}{P} + \frac{q}{lp} + \frac{q}{lp'}$$

for every pair a, a' counted by $\mathcal{M}(p, p')$. Thus

$$|M(p'-p) - (ap' - a'p)l| \leq \frac{pp'Hl}{qP} + p' + p \leq \frac{2pp'}{P} + p' + p \leq 12P.$$

For a given δ there is at most one way to choose a, a' with $0 \leq a < p$ and $0 \leq a' < p'$ which satisfy $ap' - a'p = \delta$. Thus

$$\sum_{p \neq p' \in \mathscr{P}} \mathcal{M}(p, p') \ll \#\{p \neq p' \in \mathscr{P}, |m| \leq 12P : M(p'-p) \equiv m \pmod{l}\}.$$

We chose M so that $l \nmid M$, and hence the condition $M(p'-p) \equiv m \pmod{l}$ determines $p' - p$ uniquely modulo l. Since by hypothesis $P < q/H < l$ this suffices to determine at most two values for $p' - p$ in $\mathbb{Z}$. So we may choose p freely and there are then at most two possibilities for p'. As a result, after counting up the possible choices for m, we conclude that

$$\sum_{p \neq p' \in \mathscr{P}} \mathcal{M}(p, p') \ll P^2.$$

Applying this in (5.1), we conclude that

$$\sum_{m} \mathcal{A}(m)^2 \ll HP,$$

as required.

Acknowledgements

Pierce was partially supported during this work by a Marie Curie Fellowship funded by the European Commission and the National Science Foundation on grant DMS-0902658.

References

[1] E. Bombieri and H. Iwaniec, *On the order of* $\zeta(1/2+it)$, Ann. Suola Norm. Sup. Pisa Cl. Sci. (4) **13** (1986), 449–472.

[2] D. A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematatika **4** (1957), 106–112.

[3] ———, *On character sums and L-series*, J. Reine Angew. Math. **3** (1962), 193–206.

[4] ———, *On character sums and L-series II*, Proc. London Math. Soc. **3** (1963), 524–536.

[5] ———, *The character sum estimate with $r = 3$*, J. London Math. Soc. (2) **33** (1986), 219–226.

[6] ———, *Partial Gauss sums*, Bull. London Math. Soc. **20** (1988), 589–592.

[7] ———, *Partial Gauss sums II*, Bull. London Math. Soc. **21** (1989), 153–158.

[8] M.-C. Chang, *Short character sums for composite moduli*, arXiv:1201.0229.

[9] ———, *An estimate of incomplete mixed character sums*, An Irregular Mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 243–250.

[10] P. Enflo, *Some problems in the interface between number theory, harmonic analysis and geometry of Euclidean space*, Quaestiones Mathematicae **18** (1995), 309–323.

[11] J. Friedlander and H. Iwaniec, *Estimates for character sums*, Proc. American Math. Soc. **119** (1993), 365–372.

[12] D. R. Heath-Brown, *Burgess's bounds for character sums*, Proceedings in Mathematics and Statistics, Springer, New York **43** (2012), 199–213.

[13] D. Menchov, *Sur les séries de fonctions orthogonales*, Fund. Math. **1** (1923), 82–105.

[14] H. Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonal-Funktionen*, Math. Ann. **87** (1922), 112–138.

[15] T. Wooley, *Approximating the Main Conjecture in Vinogradov's Mean Value Theorem*, arXiv:1401.2932.

[16] ———, *The cubic case of the Main Conjecture in Vinogradov's Mean Value Theorem*, arXiv:1401.3150.