Burgess’s Bounds for Character Sums

D.R. Heath-Brown
Mathematical Institute, Oxford

1 Introduction

Let $\chi(n)$ be a non-principal character to modulus $q$. Then the well-known estimates of Burgess [2, 4, 5] say that if

$$S(N; H) := \sum_{N<n \leq N+H} \chi(n),$$

then for any positive integer $r \geq 2$ and any $\varepsilon > 0$ we have

$$S(N; H) \ll_{\varepsilon, r} H^{1-1/r} q^{(r+1)/(4r^2)+\varepsilon} \tag{1}$$

uniformly in $N$, providing either that $q$ is cube free, or that $r \leq 3$. Indeed one can make the dependence on $r$ explicit, if one so wants. Similarly the $q^\varepsilon$ factor may be replaced by a power of $d(q) \log q$ if one wishes. The upper bound has been the best-known for around 50 years. The purpose of this note is to establish the following estimate, which gives a mean-value estimate including the original Burgess bound as a special case.

**Theorem** Let $r \in \mathbb{N}$ and let $\varepsilon > 0$ be a real number. Suppose that $\chi(n)$ is a primitive character to modulus $q > 1$, and let a positive integer $H \leq q$ be given. Suppose that $0 \leq N_1 < N_2 < \ldots < N_J < q$ are integers such that

$$N_{j+1} - N_j \geq H, \quad (1 \leq j < J). \tag{2}$$

Then

$$\sum_{j=1}^J \max_{h \leq H} |S(N_j; h)|^{3r} \ll_{\varepsilon, r} H^{3r-3} q^{3/4+3/(4r)+\varepsilon}$$

under any of the three conditions

(i) $r = 1$;
(ii) $r \leq 3$ and $H \geq q^{1/(2r)+\varepsilon}$; or

(iii) $q$ is cube-free and $H \geq q^{1/(2r)+\varepsilon}$.

The case $J = 1$ reduces to the standard Burgess estimate (which would be trivial if one took $H \leq q^{1/2r}$). Moreover one can deduce that there are only $O_{\varepsilon,r}(q^{(3r+1)\varepsilon})$ points $N_j$ for which

$$\max_{h \leq H} |S(N_j; h)| \geq H^{1-1/r} q^{(r+1)/(4r^2)-\varepsilon},$$

for example. It would be unreasonable to ask for such a result without the spacing condition (2), since if $A$ and $B$ are intervals that overlap it is possible that the behaviour of both $\sum_{n \in A} \chi(n)$ and $\sum_{n \in B} \chi(n)$ is affected by $\sum_{n \in A \cap B} \chi(n)$.

There are other results in the literature with which this estimates should be compared. Friedlander and Iwaniec [8, Theorem 2’] establish a bound for $J \sum_{j=1}^J S(N_j; h)$ which can easily be used to obtain an estimate of the form

$$\sum_{j=1}^J |S(N_j; h)|^{2r} \ll_{\varepsilon,r} h^{2r-2} q^{1/2+1/(2r)+\varepsilon}.$$

This is superior to our result in that it involves a smaller exponent $2r$. However they do not include a maximum over $h$ and their result is subject to the condition that $h(N_J - N_1) \leq q^{1+1/(2r)}$.

We should also mention the work of Chang [6, Theorem 8]. The result here is not so readily compared with ours, or indeed with the Burgess estimate (1). However, with a certain amount of effort one may show that our theorem gives a sharper bound at least when $JH^3 \leq q^2$.

It would have been nice to have established a result like our theorem, but involving the $2r$-th moment. The present methods do not allow this in general. However for the special case $r = 1$ one can indeed achieve this, in the following slightly more flexible form. Specifically, suppose that $\chi(n)$ is a primitive character to modulus $q$, and let $I_1, \ldots, I_J$ be disjoint subintervals of $(0, q]$. Then for any $\varepsilon > 0$ we have

$$\sum_{j=1}^J \left| \sum_{n \in I_j} \chi(n) \right|^2 \ll_{\varepsilon} q^{1+\varepsilon} \tag{3}$$
with an implied constant depending only on \( \varepsilon \). This is a mild variant of Lemma 4 of Gallagher and Montgomery \([9]\). One can deduce the Pólya–Vinogradov as an immediate consequence of Lemma 4 (which is the same as Gallagher and Montgomery’s Lemma 4). In fact there are variants of (3) for quite general character sums. For simplicity we suppose \( q \) is a prime \( p \). Let \( f(x) \) and \( g(x) \) be rational functions on \( \mathbb{F}_p \), possibly identically zero. Then (3) remains true if we replace \( \chi(n) \) by \( \chi(f(n))e_p(g(n)) \), providing firstly that we exclude poles of \( f \) and \( g \) from the sum, and secondly that we exclude the trivial case in which \( f \) is constant and \( g \) is constant or linear. (The implied constant will depend on the degrees of the numerators and denominators in \( f \) and \( g \).) We leave the proof of this assertion to the reader.

For \( r = 1 \) the ideas of this paper are closely related to those in the article of Davenport and Erdős \([7]\), which was a precursor of Burgess’s work. For \( r \geq 2 \) the paper follows the route to Burgess’s bounds developed in unpublished notes by Hugh Montgomery, written in the 1970’s, which were later developed into the Gallagher and Montgomery article \([9]\). In particular the mean-value lemmas in §2 are essentially the same as in their paper, except that we have given the appropriate extension to general composite moduli \( q \). We reproduce the arguments merely for the sake of completeness.

After the mean-value lemmas in §2 have been established we begin the standard attack on the Burgess bounds in §3, but incorporating the sum over \( N_j \) in a non-trivial way in §4. It is this final step that involves the real novelty in the paper. This process will lead to the following key lemma.

**Lemma 1** Let a positive integer \( r \geq 2 \) and a real number \( \varepsilon > 0 \) be given. Let 0 \( \leq N_1 < N_2 < \ldots < N_J < q \) be integers such that (3) holds. Then for any primitive character \( \chi \) to modulus \( q \), and any positive integer \( H \in (q^{1/(2r)}, q] \) we have

\[
\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^{r} \ll_{\varepsilon,r} q^{1/4+1/(4r)+\varepsilon} H^{r^{-1}} \left\{ J^{2/3} + J\left(H^{-1}q^{1/(2r)} + Hq^{-1/2-1/(4r)}\right) \right\},
\]

provided either that \( r \leq 3 \) or that \( q \) is cube-free.

Throughout the paper we shall assume that \( q \) is sufficiently large in terms of \( r \) and \( \varepsilon \) wherever it is convenient. The results are clearly trivial when \( q \ll_{\varepsilon,r} 1 \). We should also point out that we shall replace \( \varepsilon \) by a small multiple from time to time. This will not matter since all our results hold for all \( \varepsilon > 0 \). Using this convention we may write \( q^{\varepsilon} \log q \ll_{\varepsilon} q^{\varepsilon} \), for example.
2 Preliminary Mean-value Bounds

Our starting point, taken from previous treatments of Burgess’s bounds, is the following pair of mean value estimates.

Lemma 2 Let $r$ be a positive integer and let $\varepsilon > 0$. Then if $\chi$ is a primitive character to modulus $q$ we have

$$\sum_{n=1}^{q} |S(n; h)|^2 \ll_{\varepsilon} q^{1+\varepsilon} h$$

for any $q$, while

$$\sum_{n=1}^{q} |S(n; h)|^{2r} \ll_{\varepsilon, r} q^r (qh^r + q^{1/2}h^{2r})$$

under any of the three conditions

(i) $q$ is cube-free; or

(ii) $r = 2$; or

(iii) $r = 3$ and $h \leq q^{1/6}$.

The case $r = 1$ is given by Norton [11, (2.8)], though the proof is attributed to Gallagher. For $r \geq 2$ the validity of the lemma under the first two conditions follows from Burgess [4, Lemma 8], using the same method as in Burgess [3, Lemma 8]. The estimate under condition (iii) is given by Burgess [5, Theorem B].

We proceed to deduce a maximal version of Lemma 2 as in Gallagher and Montgomery [9, Lemma 3].

Lemma 3 Let $r$ be a positive integer and let $\varepsilon > 0$. Then if $\chi$ is a primitive character to modulus $q$ and $H \in \mathbb{N}$ we have

$$\sum_{n=1}^{q} \max_{h \leq H} |S(n; h)|^2 \ll_{\varepsilon} q^{1+\varepsilon} H$$

for all $q$, while

$$\sum_{n=1}^{q} \max_{h \leq H} |S(n; h)|^{2r} \ll_{\varepsilon, r} q^r (qH^r + q^{1/2}H^{2r})$$

under either of the conditions.
(i) $q$ is cube-free; or
(ii) $2 \leq r \leq 3$.

The strategy for the proof goes back to independent work of Rademacher [12] and Menchov [10], from 1922 and 1923 respectively. It clearly suffices to consider the case in which $H = 2^t$ is a power of 2. We will first prove the result under the assumption that $H \leq q^{1/(2r)}$. We will assume that $r \geq 2$, the case $r = 1$ being similar. Suppose that $|S(n; h)|$ attains its maximum at a positive integer $h = h(n) \leq H$, say. We may write

$$h = \sum_{d \in D} 2^{t-d}$$

for a certain set $D$ of distinct non-negative integers $d \leq t$. Then

$$S(n; h) = \sum_{d \in D} S(n + v_{n,d}2^{t-d}, 2^{t-d})$$

where

$$v_{n,d} = \sum_{e \in D, e < d} 2^{d-e} < 2^d.$$ 

By Hölder’s inequality we have

$$|S(n; h)|^{2r} \leq \{|D|^{2r-1} \sum_{d \in D} |S(n + v_{n,d}2^{t-d}, 2^{t-d})|^{2r},$$

We now include all possible values of $d$ and $v$ to obtain

$$|S(n; h)|^{2r} \leq (t+1)^{2r-1} \sum_{0 \leq d \leq t} \sum_{0 \leq v < 2^d} |S(n + v2^{t-d}, 2^{t-d})|^{2r},$$

and hence

$$\max_{h \leq H} |S(n; h)|^{2r} \leq (t+1)^{2r-1} \sum_{0 \leq d \leq t} \sum_{0 \leq v < 2^d} |S(n + v2^{t-d}, 2^{t-d})|^{2r}.$$
We proceed to sum over \( n \) modulo \( q \), using Lemma 2, and on recalling that \( H = 2^t \leq q^{1/(2r)} \) we deduce that

\[
\sum_{n=1}^{q} \max_{h \leq H} |S(n; h)|^{2r} \leq (t+1)^{2r-1} \sum_{0 \leq d \leq t} \sum_{0 \leq v < 2^d} q^v (q2^{r(t-d)} + q^{1/2}2^{2r(t-d)})
\]

\[
\leq (t+1)^{2r-1} \sum_{0 \leq d \leq t} \sum_{0 \leq v < 2^d} (qH^r + q^{1/2}H^{2r})2^{-d}
\]

\[
= q^r (t+1)^{2r}(qH^r + q^{1/2}H^{2r})
\]

\[
\leq (t+1)^{2r}(qH^r + q^{1/2}H^{2r}).
\]

This establishes Lemma 3 when \( H \) is a power of 2 of size at most \( q^{1/(2r)} \).

To extend this to the general case, write \( H_0 \) for the largest power of 2 of size at most \( q^{1/(2r)} \). Then

\[
\max_{h \leq H} |S(n; h)| \leq \sum_{0 \leq j \leq H/H_0} \max_{h \leq H_0} |S(n + jH_0; h)|
\]

whence

\[
\sum_{n=1}^{q} \max_{h \leq H} |S(n; h)|^{2r} \leq (H/H_0)^{2r-1} \sum_{n=1}^{q} \sum_{0 \leq j \leq H/H_0} \max_{h \leq H_0} |S(n + jH_0; h)|^{2r}
\]

\[
= (H/H_0)^{2r-1} \sum_{0 \leq j \leq H/H_0} \sum_{n=1}^{q} \max_{h \leq H_0} |S(n + jH_0; h)|^{2r}
\]

\[
= (H/H_0)^{2r-1} \sum_{0 \leq j \leq H/H_0} \sum_{n \equiv (\mod q)} \max_{h \leq H_0} |S(n; h)|^{2r}
\]

\[
\leq (H/H_0)^{2r-1} \sum_{0 \leq j \leq H/H_0} q^v (qH_0^r + q^{1/2}H_0^{2r})
\]

\[
\leq (H/H_0)^{2r} (qH_0^r + q^{1/2}H_0^{2r}).
\]

However our choice of \( H_0 \) ensures that \( qH_0^r \ll q^{1/2}H_0^{2r} \) and the lemma follows.

A variant of Lemma 3 allows us to sum over well spaced points. We will only need the case \( r = 1 \).

**Lemma 4** Suppose that \( \chi(n) \) is a primitive character to modulus \( q > 1 \), and let a positive integer \( H \leq q \) be given. Suppose that \( 0 \leq N_1 < N_2 < \ldots <
$N_j < q$ are integers satisfying the spacing condition $[2]$. Then

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^2 \ll q (\log q)^2$$

To prove this we follow the argument in Gallagher and Montgomery $[9, \text{Lemma 4}]$. We first observe that for any $n \leq N$ we have

$$S(N; h) = S(n; N - n + h) - S(n; N - n).$$

If $h \leq H$ it follows that

$$|S(N; h)| \leq 2 \max_{k \leq 2H} |S(n; k)|$$

whenever $N - H < n \leq N$. Then, summing over integers $n \in (N - H, N]$ we find that

$$H |S(N; h)| \leq 2 \max_{n \in (N - H, N]} \sum_{k \leq 2H} |S(n; k)|$$

whence Hölder’s inequality yields

$$|S(N; h)|^{2r} \ll H^{-1} \max_{n \in (N - H, N]} |S(n; k)|^{2r}.\quad (5)$$

Since the intervals $(N_j - H, N_j]$ are disjoint modulo $q$ we then deduce that

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^{2r} \ll H^{-1} q \max_{n \in (N - H, N]} \sum_{k \leq 2H} |S(n; k)|^{2r}$$

and the result follows from Lemma $[3]$.

We can now deduce (3). By a dyadic subdivision it will be enough to prove the result under the additional assumption that there is an integer $H$ such that all the intervals $I_j$ have length between $H/2$ and $H$. Thus we may write $I_j = (M_j, M_j + h_j]$ with $h_j \leq H$ for $1 \leq j \leq J$, and $M_{j+1} - M_j \geq H/2$ for $1 \leq j < J$. We may therefore apply the case $r = 1$ of Lemma $[4]$ separately to the even numbered intervals and the odd numbered intervals to deduce (3).

### 3 Burgess’s method

In this section we will follow a mild variant of Burgess’s method. Although there are small technical differences from previous works on the subject, there is no great novelty here.
For any prime \( p < q \) which does not divide \( q \) we will split the integers \( n \in (N, N + h] \) into residue classes \( n \equiv aq \pmod{p} \), for \( 0 \leq a < p \). Then we can write \( n = aq + pm \) with \( m \in (N', N' + h'] \) say, where

\[
N' = \frac{N - aq}{p}, \quad h' = \frac{h}{p}.
\]

We then find that

\[
S(N; h) = \chi(p) \sum_{0 \leq a < p} S(N'; h')
\]

and hence

\[
|S(N; h)| \leq \sum_{0 \leq a < p} |S(N'; h')|.
\]

We now choose an integer parameter \( P \) in the range \( (\log q)^2 \leq P < q/2 \), and sum the above estimate for all primes \( p \in (P, 2P] \) not dividing \( q \). Since the number of such primes is asymptotically \( P/(\log P) \) we deduce that

\[
P/(\log P)|S(N; h)| \ll \sum_{P < p \leq 2P} \sum_{0 \leq a < p} |S(N'; h')|.
\]

(6)

We now apply the inequality (5), with \( H \) replaced by \( H/P \). Since we have \( h' \leq H/P \) we deduce that

\[
HP^{-1}|S(N'; h')| \ll \sum_{n \in (N' - H/P, N']} \max_{j \leq 2H/P} |S(n; j)|.
\]

Inserting this bound into (6) we find that

\[
|S(N; h)| \ll (\log P)H^{-1} \sum_n A(n; N) \max_{j \leq 2H/P} |S(n; j)|,
\]

where

\[
A(n, N) := \#\{(a, p) : P < p \leq 2P, 0 \leq a < p, n \leq N' < n + H/P\},
\]

\[
= \#\{(a, p) : n \leq (N - aq)/p < n + H/P\}.
\]

Since

\[
\sum_n A(n, N) = \sum_{a, p} \#\{n : n \leq N' < n + H/P\} \leq \sum_{a, p} \frac{H}{P} \ll PH
\]

we deduce from H"older’s inequality that

\[
|S(N; h)|^r \ll (\log P)^r P^{r-1}H^{-1} \sum_n A(n; N) \max_{j \leq 2H/P} |S(n; j)|^r,
\]

8
for any $h \leq H$. It should be noted that $A(n, N) = 0$ unless $|n| \leq 2q$, so that the sum over $n$ may be restricted to this range.

We proceed to sum over the values $N = N_j$ in Lemma 1, finding that

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^r \ll (\log P)^r P^{r-1} H^{-1} \sum_n A(n) \max_{j \leq 2H/P} |S(n; j)|^r,$$

where

$$A(n) := \#\{(a, p, N_j) : n \leq (N_j - aq)/p < n + H/P\}.$$  

From Cauchy’s inequality we then deduce that

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^r \ll (\log P)^r P^{r-1} H^{-1} N^{1/2} \left\{ \sum_{|n| \leq 2q} \max_{j \leq 2H/P} |S(n; j)|^{2r} \right\}^{1/2},$$

where

$$N := \sum_n A(n)^2 \leq H^{-1} M,$$

with

$$M := \#\{(a_1, a_2, p_1, p_2, N_j, N_k) : |(N_j - a_1q)/p_1 - (N_k - a_2q)/p_2| \leq H/P\}.$$  

Thus

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^r \ll (\log P)^r P^{r-3/2} H^{-1/2} M^{1/2} \left\{ \sum_{|n| \leq 2q} \max_{j \leq 2H/P} |S(n; j)|^{2r} \right\}^{1/2}.$$

The second sum on the right may be bounded via Lemma 3 giving

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^r \ll \varepsilon, r q^r P^{r-3/2} H^{-1/2} (q^{1/2}(H/P)^{r/2} + q^{1/4}(H/P)^r) M^{1/2},$$

on replacing $\varepsilon$ by $\varepsilon/2$.

Naturally, in order to apply Lemma 3 we will need to have $q$ cube-free, or $r \leq 3$. The natural choice for $P$ is to take $2Hq^{-1/(2r)} \leq P \ll Hq^{-1/(2r)}$ so that $q^{1/2}(H/P)^{r/2}$ and $q^{1/4}(H/P)^r$ have the same order of magnitude. The conditions previously imposed on $P$ are then satisfied provided that $H \geq q^{1/(2r)}$. With this choice for $P$ we deduce that

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j; h)|^r \ll \varepsilon, r q^{1/4+3/(4r)+\varepsilon} H^{r-2} M^{1/2}. \quad (7)$$
4 Estimating $\mathcal{M}$

In this section we will estimate $\mathcal{M}$ and complete the proof of Lemma 1. It is the treatment of $\mathcal{M}$ which represents the most novel part of our argument.

We split $\mathcal{M}$ as $\mathcal{M}_1 + \mathcal{M}_2$ where $\mathcal{M}_1$ counts solutions with $p_1 = p_2$ and $\mathcal{M}_2$ corresponds to $p_1 \neq p_2$. When $p_1 = p_2$ the defining condition for $\mathcal{M}$ becomes
\[
| (N_j - N_k) - q(a_1 - a_2) | \leq p_1 H / P \leq 2H.
\]
Thus
\[
|a_1 - a_2| \leq q^{-1}(|N_j - N_k| + 2H) \leq 3.
\]
Moreover, given $N_k$ and $a_1 - a_2$, there will be at most 5 choices for $N_j$, in view of the spacing condition (2). Thus we must allow for $O(P)$ choices for $p_1$, for $O(P)$ choices for $a_1$ and $a_2$, and $O(J)$ choices for $N_j$ and $N_k$, so that
\[
\mathcal{M}_1 \ll P^2 J. \quad (8)
\]

To handle $\mathcal{M}_2$ we begin by choosing a prime $\ell$ in the range
\[
q/H < \ell \leq 2q/H.
\]
This is possible, by Bertrand’s Postulate. We then set
\[
M_j := \left\lfloor \frac{N_j \ell}{q} \right\rfloor, \quad (1 \leq j \leq J)
\]
so that the $M_j$ are non-negative integers in $[0, \ell)$. Moreover the spacing condition (2) implies that
\[
M_{j+1} > \frac{N_{j+1} \ell}{q} - 1 \geq \frac{(N_j + H) \ell}{q} - 1 > \frac{N_j \ell}{q} \geq M_j,
\]
so that the integers $M_j$ form a strictly increasing sequence. Since
\[
|N_j - qM_j / \ell| \leq q/\ell
\]
we now see that if $(a_1, a_2, p_1, p_2, N_j, N_k)$ is counted by $\mathcal{M}_2$ then
\[
\left| \frac{qM_j / \ell - a_1 q}{p_1} - \frac{qM_k / \ell - a_2 q}{p_2} \right| \leq \frac{H}{P} + \frac{q}{\ell p_1} + \frac{q}{\ell p_2},
\]
whence
\[
|p_2 M_j - p_1 M_k - \ell \delta| \leq \frac{H \ell p_1 p_2}{P q} + p_1 + p_2 \leq 12P,
\]
with \( \delta = a_1 p_2 - a_2 p_1 \). If \( p_1, p_2 \) and \( \delta \) are given, there is at most one pair of integers \( a_1, a_2 \) with \( 0 \leq a_1 < p_1, 0 \leq a_2 < p_2 \) and \( a_1 p_2 - a_2 p_1 = \delta \). Thus

\[
\mathcal{M}_2 \leq \sum_{M_j, M_k} \# \{(p_1, p_2, m) : |m| \leq 12P, p_2 M_j - p_1 M_k \equiv m (\text{mod } \ell)\}.
\]

We now consider how many pairs \( p_1, p_2 \) there may be for each choice of \( M_j, M_k \). We define the set

\[
\Lambda := \{(x, y, z) \in \mathbb{Z}^3 : xM_j - yM_k \equiv z (\text{mod } \ell)\},
\]

which will be an integer lattice of determinant \( \ell \). Admissible pairs \( p_1, p_2 \) produce points \( x \bar{=} (x, y, z) \in \Lambda \) with \( x \neq y \) both prime and \( |x| \leq 12P \), where

\[
|x| := \max(|x|, |y|, |z|).
\]

The lattice \( \Lambda \) has a \( \mathbb{Z} \)-basis \( b_1, b_2, b_3 \) such that

\[
|b_1| \leq |b_2| \leq |b_3| \tag{9}
\]

and

\[
\det(\Lambda) \ll |b_1|, |b_2|, |b_3| \ll \det(\Lambda) = \ell, \tag{10}
\]

and with the property that there is an absolute constant \( c_0 \) such that if \( x \in \Lambda \) is written as \( \lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 b_3 \) then

\[
|\lambda_i| \leq c_0 |x|/|b_i|, \quad (1 \leq i \leq 3).
\]

The existence of such a basis is a standard fact about lattices, see Browning and Heath-Brown [11, Lemma 1, (ii)], for example. When \( |b_3| \leq 12c_0 P \) we now see that the number of lattice elements of size at most \( 12P \) is

\[
\leq \left( 1 + \frac{12c_0 P}{|b_1|} \right) \left( 1 + \frac{12c_0 P}{|b_2|} \right) \left( 1 + \frac{12c_0 P}{|b_3|} \right)
\]

\[
\ll \frac{|b_1|, |b_2|, |b_3|}{P^3}
\]

\[
\ll \frac{\det(\Lambda)}{P^3} \ll HP^3 q^{-1}
\]

by (9) and (10). If \( |b_1| > 12c_0 P \) the only vector in \( \Lambda \) of norm at most \( 12P \) is the origin, while if \( |b_1| \leq 12c_0 P < |b_2| \) the only possible vectors are of the form \( \lambda_1 b_1 \). In this latter case \( (p_2, p_1, m) = \lambda_1 b_1 \) so that \( \lambda_1 \) divides \( \text{h.c.f.}(p_2, p_1) = 1 \). Hence there is at most 1 solution in this case.
There remains the situation in which $|b_2| \leq 12c_0P < |b_3|$, so that the admissible vectors are linear combinations $\lambda_1 b_1 + \lambda_2 b_2$. In this case we write $b_i = (x_i, y_i, z_i)$ for $i = 1, 2$ and set $\Delta = x_1y_2 - x_2y_1$. If $\Delta = 0$ then $(x_1, y_1)$ and $(x_2, y_2)$ are proportional, and hence are both integral scalar multiples of some primitive vector $(x, y)$ say. However we then see that if $(p_2, p_1, m) = \lambda_1 b_1 + \lambda_2 b_2$ then $(p_2, p_1)$ is a scalar multiple of $(x, y)$, so that $b_1$ and $b_2$ determine $p_1$ and $p_2$. Thus when $\Delta = 0$ the primes $p_1$ and $p_2$ are determined by $M_j$ and $M_k$. In order to summarize our conclusions up to this point we write $\mathcal{M}_3$ for the contribution to $\mathcal{M}_2$ corresponding to all cases except that in which $|b_2| \leq 12c_0P < |b_3|$ and $\Delta \neq 0$. With this notation we then have

$$\mathcal{M}_3 \ll (HP^3q^{-1} + 1)J^2. \tag{11}$$

Suppose now that $|b_2| \leq 12c_0P < |b_3|$ and $\Delta \neq 0$. We will write $\mathcal{M}_4$ for the corresponding contribution to $\mathcal{M}$. In this case we must have $\lambda_3 = 0$, and the number of choices for $\lambda_1$ and $\lambda_2$ will be

$$\leq \left(1 + \frac{12c_0P}{|b_1|}\right) \left(1 + \frac{12c_0P}{|b_2|}\right) \ll \frac{P^2}{|b_1|.|b_2|}.$$ 

Thus if $L < |b_1|.|b_2| \leq 2L$, say, the contribution to $\mathcal{M}_4$ will be $O(P^2L^{-1})$ for each pair $M_j, M_k$. 

To estimate the number of pairs of vectors $b_1, b_2$ with $L < |b_1|.|b_2| \leq 2L$ we observe that there are $O(B_1^3B_2^3)$ possible choices with $B_1 < |b_1| \leq 2B_1$ and $B_2 < |b_2| \leq 2B_2$. A dyadic subdivision then shows that we will have to consider $O(L^3\log L)$ pairs $b_1, b_2$. Writing $b_i = (x_i, y_i, z_i)$ for $i = 1, 2$ as before we will have

$$x_1M_j - y_1M_k \equiv z_1 \pmod{\ell}, \quad x_2M_j - y_2M_k \equiv z_2 \pmod{\ell}.$$ 

These congruences determine $\Delta M_j$ and $\Delta M_k$ modulo $\ell$, and since $\ell$ is prime and $0 \leq M_j, M_k < \ell$ we see that $b_1$ and $b_2$ determine $M_j, M_k$ precisely, providing that $\ell \nmid \Delta$. However

$$|\Delta| \leq 2|b_1||b_2| \leq 2(|b_1|.|b_2|.|b_3|)^{2/3} \ll \det(\Lambda)^{2/3} = \ell^{2/3}$$

by (2) and (10). Since $\Delta \neq 0$ we then see that $\ell \nmid \Delta$ providing that $q/H$, or equivalently $\ell$, is sufficiently large. Under this assumption we therefore conclude that there are $O(L^3\log L)$ pairs $M_j, M_k$ for which $|b_2| \leq 12c_0P < |b_3|$ and $\Delta \neq 0$ and for which $L < |b_1|.|b_2| \leq 2L$. Thus each dyadic range $(L, 2L]$ contributes $O(P^2L^{-1}\min(J^2, L^3\log L))$ to $\mathcal{M}_4$. Since

$$P^2L^{-1}\min(J^2, L^3) \leq P^2L^{-1}(J^2)^{2/3}(L^3)^{1/3} = P^2J^{4/3}$$

12
we deduce that
\[ M_4 \ll P^2 J^{2/3} \log q, \]
and comparing this with the bounds (8) and (11) we then see that
\[ M \ll (HP^3 q^{-1} + 1) J^2 + P^2 J^{4/3} \log q. \]

We may now insert this bound into (7), recalling that \( P \) is of order \( Hq^{1/(2r)} \) to deduce, after replacing \( \varepsilon \) by \( \varepsilon / 2 \) that
\[ \sum_{j=1}^J \max_{h \leq H} |S(N_j; h)|^r \ll \varepsilon, r \cdot q^{1/4 + 1/(4r) + \varepsilon} H^{-1} \left\{ J^{2/3} + J (H^{-1} q^{1/(2r)} + Hq^{-1/2 - 1/(4r)}) \right\}, \]
as required for Lemma 1.

5 Deduction of the theorem

We will prove the theorem by induction on \( r \). The result for \( r = 1 \) is an immediate consequence of Lemma 4, together with the Pólya–Vinogradov inequality.

For \( r \geq 2 \) we will use a dyadic subdivision, classifying the \( N_j \) according to the value \( V = 2^r \) for which
\[ V/2 < \max_{h \leq H} |S(N_j; h)|^r \leq V. \] (12)
Clearly numbers \( N_j \) for which the corresponding \( V \) is less than 1 make a satisfactory contribution in our theorem, and so it suffices to assume that (12) holds for all \( N_j \).

We now give three separate arguments, depending on which of the three terms on the right of (11) dominates. If
\[ \sum_{j=1}^J \max_{h \leq H} |S(N_j; h)|^r \ll_{\varepsilon, r} q^{1/4 + 1/(4r) + \varepsilon} H^{r-1} J^{2/3} \]
then
\[ JV^r \ll_{\varepsilon, r} q^{1/4 + 1/(4r) + \varepsilon} H^{r-1} J^{2/3}, \]
whence
\[ JV^{3r} \ll_{\varepsilon, r} q^{3/4 + 3/(4r) + 3\varepsilon} H^{3r-3}, \]
which suffices for the theorem. If the second term dominates we will have

$$\sum_{j=1}^{J} \max_{h \leq H} |S(N_j, h)| r \ll_{\epsilon, r} q^{1/4 + 1/(4r) + \epsilon} H^{r-1} J H^{-1} q^{1/(2r)},$$

so that

$$J V^r \ll_{\epsilon, r} q^{1/4 + 3/(4r) + \epsilon} H^{r-2} J.$$  

In this case it follows that

$$V^r \ll_{\epsilon, r} q^{1/4 + 3/(4r) + \epsilon} H^{r-2}. \quad (13)$$

We now use Lemma 4, which implies that

$$J V^{2r} \ll_{\epsilon, r} q^\epsilon (q H^{r-1} + q^{1/2} H^{2r-1}) \ll_{\epsilon, r} q^{1/2 + \epsilon} H^{2r-1} \quad (14)$$

since $H \geq q^{1/(2r)}$. Coupled with (13) this yields

$$J V^{3r} \ll_{\epsilon, r} q^{3/4 + 3/(4r) + 2\epsilon} H^{3r-3}$$

which again suffices for the theorem. Finally, if the third term on the right of (4) dominates we must have

$$J V^r \ll_{\epsilon, r} q^{-1/4 + \epsilon} H^r J$$

whence $V \ll_{\epsilon} H q^{-1/(4r) + \epsilon/r}$. Here we shall use the inductive hypothesis, which tells us that

$$J V^{3r-3} \ll_{\epsilon, r} q^{3/4 + 3/(4r - 4) + \epsilon} H^{3r-6}$$

if either $r = 2$ or $H \geq q^{1/(2r-2)}$ and $r \geq 3$. Under this latter assumption we therefore deduce that

$$J V^{3r} \ll_{\epsilon, r} q^{3/4 + \phi + 4\epsilon} H^{3r-3}$$

with

$$\phi = \frac{3}{4r - 4} - \frac{3}{4r} \leq \frac{3}{4r}$$

for $r \geq 2$. It therefore remains to consider the case in which $r \geq 3$ and $q^{1/(2r)} \leq H \leq q^{1/(2r-2)}$. However for such $H$ we may again use the bound (14), whence

$$J V^{3r} \ll_{\epsilon, r} q^{1/2 + \epsilon} H^{2r-1} q^{-1/4 + \epsilon} H^r = q^{1/4 + 2\epsilon} H^{3r-3} \left( H q^{-1/(2r-2)} \right)^2 q^{1/(r-1)} \leq q^{1/4 + 1/(r-1) + 2\epsilon} H^{3r-3}.$$  

To complete the proof of this final case it remains to observe that $1/4 + 1/(r-1) \leq 3/4 + 3/(4r)$ for $r \geq 3$. 

14
References

[1] T.D. Browning and D.R. Heath-Brown, Equal sums of three powers, *Invent. Math.*, 157 (2004), 553–573.

[2] D.A. Burgess, On character sums and primitive roots, *Proc. London Math. Soc. (3)*, 12 (1962), 179–192.

[3] D.A. Burgess, On character sums and $L$-series, *Proc. London Math. Soc. (3)*, 12 (1962), 193–206.

[4] D.A. Burgess, On character sums and $L$-series. II, *Proc. London Math. Soc. (3)*, 13 (1963), 524–536.

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

[6] M.-C. Chang, On a question of Davenport and Lewis and new character sum bounds in finite fields, *Duke Math. J.*, 145 (2008), 409–442.

[7] H. Davenport and P. Erdös, The distribution of quadratic and higher residues, *Publ. Math. Debrecen* 2 (1952), 252–265.

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

[9] P. X. Gallagher and H. L. Montgomery, A Note on Burgess's estimate, *Mathematical Notes*, 88 (2010), 321–329.

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

[11] K.K. Norton, On character sums and power residues, *Trans. Amer. Math. Soc.*, 167 (1972), 203–226.

[12] H. Rademacher, Eigen Staze über Reihn von allgemeinen Orthogonal-Funktionen, *Math. Ann.*, 87 (1922), 112–138.

---

Mathematical Institute, 24–29, St. Giles’, Oxford OX1 3LB UK

rhb@maths.ox.ac.uk