BURGESS BOUND FOR CHARACTER SUMS

LIANGYI ZHAO

1. Introduction

Here we give a survey on the bounds for character sums of D. A. Burgess [1]. We henceforth set

$$S_\chi(N) = \sum_{M<n\le M+N} \chi(n), \tag{1.1}$$

where χ is a non-principle character modulo a prime number p. The result of Burgess can be generalized to composite moduli, but for the sake of simplicity, we shall only concentrate on prime moduli here.

By Polya-Vinogradov inequality, we have

$$S_\chi(N) \le 6\sqrt{p}\log p. \tag{1.2}$$

See section 12.4 of [3] for a proof of the above. This bound is non-trivial whenever N is larger than $\sqrt{p}\log p$. However, the expected bound is

$$S_\chi(N) \ll \sqrt{N}p^\varepsilon,$$

which is non-trivial if $N \gg p^{3\varepsilon}$.

Burgess [1] proved the following result.

Theorem 1 (Burgess)**.** *If p is a prime number and χ is a non-principal character of order l modulo p and N, $r \in \mathbb{N}$, then*

$$|S_\chi(N)| \le cN^{1-\frac{1}{r}}p^{\frac{r+1}{4r^2}}(\log p)^{\frac{1}{r}}, \tag{1.3}$$

where c is an absolute constant ($c = 30$ should suffice.).

A consequence of Theorem 1 is that we have a non-trivial bound for a character sums whenever $N > p^{1/4+\varepsilon}$. In turn, we have the following results.

(1) The maximum number of consecutive quadratic residue or non-residue to a prime modulus is $O(p^{1/4+\varepsilon})$, provided that p is sufficiently large. Hence any sequence of consectivie integers longer than the said quantity would contain at least one quadratic residue and non-residue.

(2) For each $\varepsilon > 0$, every interval of length $N > p^{1/4+\varepsilon}$ contains

$$\frac{\varphi(p-1)}{p-1}N\left(1+O\left(p^{-\delta}\right)\right)$$

primitive roots modulo p, where δ depends on ε. Hence the least primitive root modulo p is $O(p^{1/4+\varepsilon})$.

(3) With $r = 2$, Theorem 1 also gives a subconvexity bound for Dirichlet L-functions in the conductor aspect. If χ is a character modulo a prime p, then

$$L\left(\frac{1}{2}+it,\chi\right) \ll |t|p^{3/16+\varepsilon}.$$

Considering that there is no non-principal character modulo $p = 2$ and character sums of the form in (1.1) modulo $p = 3$ is never exceeds one in modulus, it suffices in the sequal to consider only $p \ge 5$.

Date: March 20, 2007.

2. Various Tranformations and Estimates

For the benefit the combinatorists in the audience, the arguments will go by induction on N. The statement is obvious if $N = 1$. If either

$$N > p^{1/2+1/(4r)} \log p, \text{ or } N < c^r p^{1/4+1/(4r)} \log p,$$

the majorant in (1.3) exceeds that in (1.2). Hence in those cases, Theorem 1 follows from (1.2). Therefore, we henceforth assume that

$$c^r p^{1/4+1/(4r)} \log p \le N \le p^{1/2+1/(4r)} \log p. \tag{2.1}$$

Making a shift $n \to n + h$ with $1 \le h \le H < N$, we have

$$S_\chi(N) = \sum_{M < n \le M+N} \chi(n+h) + 2\theta E(H),$$

where θ is a complex number of modulus not exceeding 1, and

$$E(H) = cH^{1-1/r} p^{\frac{r+1}{4r^2}} (\log p)^{1/r},$$

by the induction hypothesis.

Note here that $E(H)$ represent characters sums of length less than N. Let $H = AB$ with $A, B \in \mathbb{N}$. Write $h = ab$ with $1 \le a \le A$ and $1 \le b \le B$, we have, after summing over all a's and b's in the said range,

$$S_\chi(N) = \frac{1}{H} \sum_{\substack{1 \le a \le A \\ 1 \le b \le B}} \sum_{M < n \le M+N} \chi(n+ab) + 2\theta E(H). \tag{2.2}$$

The first term on the right-hand side of (2.2) is

$$\frac{1}{H} \sum_a \sum_n \chi(a) \sum_b \chi(\overline{a}n + b).$$

Here $\overline{a}$ denotes the multiplicative inverse of a modulo p. Note here that by (2.1), we have $A, B \le H < N < p$. So $\gcd(a, p) = 1$ as $p \in \mathbb{P}$.

We now have

$$|S_\chi(N)| \le H^{-1} V + 2E(H), \tag{2.3}$$

where

$$V = \sum_{x \bmod p} \nu(x) \left| \sum_{1 \le b \le B} \chi(x+b) \right|,$$

and $\nu(x)$ is the number of representations of x as $\overline{a}n$ modulo p with $1 \le A$ and $M < n \le M + N$.

Applying Hölder's inequality twice, we have

$$V \le V_1^{1-1/r} V_2^{1/(2r)} W^{1/(2r)} \tag{2.4}$$

with

$$V_1 = \sum_{x \bmod p} \nu(x), \; V_2 = \sum_{x \bmod p} \nu^2(x)$$

and

$$W = \sum_{x \bmod p} \left| \sum_{1 \le b \le B} \chi(x+b) \right|^{2r}.$$

(2.4) is most easily seen as having $p = (1 - 1/(2r))^{-1}$ and $q = 2r$ in the first application of Hölder and

$$p = \frac{2r-1}{2r-2}, \; q = 2r-1$$

and re-writing

$$\nu(x)^{1+1/(2r-1)} = \nu(x)^{\frac{2r-2}{2r-1}} \nu(x)^{\frac{2}{2r-1}}$$

in the second application.

It should be noted that $\nu(x)$ is very often zero and that we could have restricted the outer sum in W only to the x for which $\nu(x) \neq 0$. But we are not able to take advantage of such a condition. We have potentially lost greatly in relaxing the said restriction, but such loss is small relative to the length of the character sum raised to the power $2r$. We hope, and indeed we do, that we shall gain some total saving in the end. It is also neccessary to have integral moments of $\nu(x)$ in the sequal. This required the second application of Hölder's inequality.

We first observe that

$$V_1 = AN, \tag{2.5}$$

and that

$$V_2 = \{(a_1, a_2, n_1, n_2) : 1 \leq a_1, a_2 \leq A, M < n_1, n_2 \leq M + N, a_1 n_2 \equiv a_2 n_2 \pmod{p}\}.$$

Fix a_1, a_2 and set $kp = a_1 n_2 - a_2 n_1$, we have

$$kp - (a_1 - a_2)M = a_1 n_2 - a_2 n_1 - a_1 M + a_2 M = a_2(n_2 - M) - a_2(n_2 - M).$$

The last difference is between two natural numbers neither of which exceeds AN. Hence we have

$$\left| k - (a_1 - a_2)\frac{M}{p} \right| \leq \frac{AN}{p}.$$

Moreover, we have $\gcd(a_1, a_2)|k$. Hence the number of possible choices for k is at most

$$\frac{2AN}{\gcd(a_1, a_2)p} + 1.$$

Moreover, by the extended Euclidean algorithm, the number of pairs (n_1, n_2) satisfying

$$a_1 n_2 - a_2 n_1 = kp$$

with a_1, a_2 and k fixed is less than both

$$\frac{2N \gcd(a_1, a_2)}{a_1}, \frac{2N \gcd(a_1, a_2)}{a_2}.$$

Hence

$$V_2 \leq 2N \sum_{a_1} \sum_{a_2} \frac{\gcd(a_1, a_2)}{\max(a_1, a_2)} \left(\frac{2AN}{\gcd(a_1, a_2)p} + 1 \right).$$

Note that

$$\sum_{a_1} \sum_{a_2} \frac{2AN}{\max(a_1, a_2)p} \leq \frac{4AN}{p} \sum_{a_1} \frac{1}{a_1} \sum_{a_2 \leq a_1} 1 = \frac{4A^2 N}{p},$$

and

$$\sum_{a_1} \sum_{a_2} \frac{\gcd(a_1, a_2)}{\max(a_1, a_2)} \leq 2 \sum_{a_1} \frac{1}{a_1} \sum_{a_2 \leq a_2} \gcd(a_1, a_2) \leq 2 \sum_{a_1} \frac{1}{a_1} \sum_{d|a_1} \sum_{\substack{d|a_2 \\ a_2 \leq a_1}} d \leq 2 \sum_{a_1} \tau(a_2) \leq 4A \log(3A),$$

where $\tau(n)$ is the number of divisors of n and the last inequality arrives upon noting that

$$\sum_{n \leq x} \tau(n) = x \log x + \theta x,$$

with $|\theta| \leq 1$. Putting the above estimates together, we have

$$V_2 \leq 8AN(ANp^{-1} + \log(3A)). \tag{2.6}$$

We now assume the following estimate for W which we shall prove in the next section.

$$W \leq (2rB)^r p + 2rB^{2r} p^{1/2}. \tag{2.7}$$

We now take

$$A = \left[\frac{N}{9rp^{1/(2r)}} \right] \text{ and } B = \left[rp^{1/(2r)} \right].$$

We first note that $A \geq 1$ by the first inequality in (2.1) and by the second inequality in the same we have

$$AN \leq \frac{N^2}{9rp^{1/(2r)}} \leq \frac{p^{1+1/r} \log^2 p}{9rp^{1/(2r)}} \leq p \log^2 p. \tag{2.8}$$

Hence by (2.7) and the definition of B, we have

$$W \leq (2r)^{2r} p^{3/2} \tag{2.9}$$

and by (2.6) and (2.8) we have

$$V_2 \leq AN(4\log p)^2. \tag{2.10}$$

Putting (2.4), (2.5), (2.9) and (2.10) together, we get

$$V \leq 2r(AN)^{1-1/(2r)}(4\log p)^{1/r} p^{3/(4r)}.$$

Now noting that $A \leq N/(9rp^{1/(2r)})$, we get

$$V \leq N^{2-1/r} p^{\frac{r+1}{4r^2}} (\log p)^{1/r}. \tag{2.11}$$

Putting (2.3) and (2.11) together, we get

$$|S_\chi(N)| \leq \frac{V}{H} + 2E(H) \leq \frac{1}{H} N^{2-1/r} p^{\frac{r+1}{4r^2}} (\log p)^{1/r} + cH^{1-1/r} p^{\frac{r+1}{4r^2}} (\log p)^{1/r}.$$

Clearly, $H = AB \leq N/9$. It can be shown further that $H \geq (N/(9rp^{1/(2r)}) - 1)(rp^{1/(2r)} - 1) > N/10$, using the inequalities in (2.1) and recalling that $p \geq 5$ and that c will be taken to be 30. Therefore, we have

$$|S_\chi(N)| \leq \left(10 + \frac{2}{3}c\right) N^{1-1/r} p^{\frac{r+1}{4r^2}} (\log p)^{1/r}.$$

Again remembering that $c = 30$, we have the desired result.

3. The Heart of the Proof

It still remains to prove (2.7) which is at the heart of Burgess's result. Again recalling that $B < p$ from the inequalities in (2.1).

Expanding W, we get

$$W = \sum_{1 \leq b_1 \cdots b_{2r} \leq B} \cdots \sum_{x \bmod p} \chi\left((x+b_1)\cdots(x+b_r)(x+b_{r+1})^{l-1}\cdots(x+b_{2r})^{l-1}\right), \tag{3.1}$$

recalling that χ is of order l. Now we write

$$(x+b_1)\cdots(x+b_r)(x+b_{r+1})^{l-1}\cdots(x+b_{2r})^{l-1} \tag{3.2}$$

as

$$(x+m_1)^{\beta_1}\cdots(x+m_u)^{\beta_u}, \tag{3.3}$$

where $m_1, \cdots, m_u$ are all distinct. Therefore, $m_1, \cdot, m_u$ as also all distinct modulo p as $B < p$. Now we reduce $\beta_1, \cdots, \beta_u$ modulo l and have them lie between 0 and $l-1$. Now the product in (3.3) becomes, possiblly empty,

$$(x+c_1)^{\gamma_1}\cdots(x+c_v)^{\gamma_v}, \tag{3.4}$$

where $c_1, \cdots, c_v$ are all distinct and $0 < \gamma_i \leq l-1$.

If the polynomial in (3.2) or (3.3) is a perfect l-th power, then we have the trivial bound

$$\left| \sum_{x \bmod p} \chi\left((x+b_1)\cdots(x+b_r)(x+b_{r+1})^{l-1}\cdots(x+b_{2r})^{l-1}\right) \right| \leq p.$$

But this is only possible when $b_1, \cdots, b_{2r}$ can the arranged into r equal pairs. The number of such cases in which we must rely on the trivial bound does not exceed

$$r \begin{pmatrix} 2r \\ r \end{pmatrix} B^r \leq (2rB)^r.$$

Now note that these terms lead to the first term on the right-hand side of (2.7).

Now assume that (3.4) is not an empty product; i.e. the polynomials in either and hence both (3.2) and (3.3) is not an l-th power. Keep in mind that the number of such polynomials does not exceed B^{2r}. We note that

$$\sum_{x \bmod p} \chi\left((x+c_1)^{\gamma_1} \cdots (x+c_v)^{\gamma_v}\right)$$

and

$$\sum_{x \bmod p} \chi\left((x+m_1)^{\beta_1} \cdots (x+m_u)^{\beta_u}\right)$$

differ only in the terms for which the latter is zero while the first is not. There are at most $u-v$ such terms. Assuming that

$$\left|\sum_{x \bmod p} \chi\left((x+c_1)^{\gamma_1} \cdots (x+c_v)^{\gamma_v}\right)\right| \leq (v-1)p^{1/2}. \tag{3.5}$$

We have

$$\begin{aligned}
&\left|\sum_{x \bmod p} \chi\left((x+b_1)\cdots(x+b_r)(x+b_{r+1})^{l-1}\cdots(x+b_{2r})^{l-1}\right)\right| \\
= &\left|\sum_{x \bmod p} \chi\left((x+m_1)^{\beta_1}\cdots(x+m_u)^{\beta_u}\right)\right| \\
\leq &\ u-v+\left|\sum_{x \bmod p} \chi\left((x+c_1)^{\gamma_1}\cdots(x+c_v)^{\gamma_v}\right)\right| \\
\leq &\ u-v+(v-1)p^{1/2} \leq 2rp^{1/2},
\end{aligned}$$

noting that $0 < v \leq u \leq 2r$. This estimate gives rise to the second term on the right-hand side of (2.7).

Now it finally remains *only* to prove (3.5) which lies at the heart's core of Burgess's result.

Let $\mathbb{F}_p$ be the field with p elements. Let $K = \mathbb{F}_p[X]$ and Z be the algebraic extension of K by adjoining y to K, where

$$y^l = f$$

where f is the polynomial in (3.4). Here we note that K is a principal ideal domain. We define a character, *á la* H. Hasse [2], as follows.

If $\mathfrak{a}$ is an ideal in K, then

$$\chi(\mathfrak{a}) = \begin{cases} \chi\left(N_{\mathfrak{a}}(f)\right), & \text{if } \mathfrak{a} \text{ is prime to } \mathfrak{f}, \\ 0, & \text{otherwise.} \end{cases}$$

Here $N_{\mathfrak{a}}(f)$ is the norm of f in the residue class ring $K/\mathfrak{a}$ [1] and

$$\mathfrak{f} = \prod_{\mathfrak{p}} \mathfrak{p},$$

where the product is over prime ideals $\mathfrak{p}$ that divide the ideal generated by f.

Now we define an L-function for Z/K as

$$L(s,\chi) = \prod_{\mathfrak{p}} \left(1 - \frac{\chi(\mathfrak{p})}{R(\mathfrak{p})^s}\right)^{-1} = \sum_{\mathfrak{a}} \frac{\chi(\mathfrak{a})}{R(\mathfrak{a})^s},$$

where the product is over the prime ideals $\mathfrak{p}$ of K and the sum is over all ideals $\mathfrak{a}$ of K.

H. Hasse [2] showed that

$$p^{(d-1)s} L(s,\chi),$$

[1] It suffices to think of this as the following in this case. If $\mathfrak{a}$ is a prime ideal, then $K/\mathfrak{a}$ is a field and vector space over $\mathbb{F}_p$. If M is the matrix representation of the linear transformation $g \to fg$ in the said vector space, then the norm in question is the norm of the matrix M. This norm is then extended to all ideal $\mathfrak{a}$, prime or not, by multiplicity.

where d is the degree of $\mathfrak{f}$, is a polynomial of degree $d-1$ in p^s with roots $\varpi_1, \cdots, \varpi_{d-1}$ and if $\sigma(\chi)$ is the coefficient of p^{-s} in $L(s,\chi)$, then

$$\sigma(\chi) = \sum_{\mathfrak{p}} \chi(\mathfrak{p}) = -\sum_{k=1}^{d-1} \varpi_k, \tag{3.6}$$

where the sum over $\mathfrak{p}$ is over all primes ideals of degree one in K. In our case, these ideals are exactly those generated by polynomials in K that are linear and monic. Moreover, if $\mathfrak{p}$ is generated by such a polynomial, $x-a$ say, then it is easy to compute that

$$\chi(\mathfrak{p}) = \chi(N_{\mathfrak{a}}(f)) = \chi(f(a)).$$

Therefore, we have

$$\sum_{\mathfrak{p} \text{ of degree } 1} \chi(\mathfrak{p}) = \sum_{a \bmod p} \chi(f(a)) \tag{3.7}$$

Now it is due to A. Weil that $L(s,\chi)$ satisfies the Riemann hypothesis; i.e. all the zeros of $L(s,\chi)$ have real part $1/2$. This is to say that $|\varpi_k| = p^{1/2}$. From this, we infer that

$$\left| \sum_{\mathfrak{p} \text{ of degree } 1} \chi(\mathfrak{p}) \right| \leq (d-1)p^{1/2}. \tag{3.8}$$

Now, noting that $d = v$ and putting together (3.7) and (3.8), we get (3.5) and hence the desired theorem.

References

[1] D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. **12** (1962), no. 3, 179–192.
[2] H. Hasse, *Theorie der relativ-zyklischen algebraischen Funktionköper, insbesondere bei endlichen konstantenkörper*, J. reine angew. Math. **172** (1935), 37–54.
[3] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, 2004.

Department of Mathematics, Royal Institute of Technology(KTH)
Lindstedtsvägen 25, Stockholm 10044 Sweden
Email: lzhao@math.kth.se