BOUND ON COMPLETE EXPONENTIAL SUMS

TODD COCHRANE
Kansas State University

Dedicated to Heini Halberstam

ABSTRACT. We give a brief survey on complete exponential sums of the type $S(g) = \sum_x e_p(g(x))$, with $g$ a polynomial in $n$ variables over the finite field $\mathbb{F}_p$, and prove several new results including the following. If $g$ is non-composite, that is, not of the type $g(x) = f(h(x))$ with $\deg f \geq 2$, and $p \geq c_1(d,n)$ then the weights of the characteristic values of $S(g)$ are all $\leq 2n - 2$. For homogeneous $g$ we can take $c_1(d,n) = 2$. A partial converse is also given. Next, if $g$ is homogeneous, absolutely irreducible of degree $d$, and has a singular locus of dimension $\ell$ in $\mathbb{F}^{n-1}$, then the characteristic values of $S(g)$ have weight $\leq n + \ell + 1$, and so $|S(g)| \leq (4d + 5)^n p^{-\frac{n+\ell+1}{2}}$.

Let $p$ be a prime, $\mathbb{F}_p$ the finite field in $p$ elements, and $g(x) \in \mathbb{F}_p[x_1,..,x_n]$, be a polynomial of degree $d = d_g$. Complete exponential sums of the type,

$$S(g) = \sum_x e_p(g(x)),$$

where the sum is over all points in $\mathbb{F}_p^n$ and, as usual $e_p(*) = e(2\pi i*/p)$, have a rich history dating back to Gauss. Gauss established the identity

$$G_p := \sum_x e_p(x^2) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod{4} \\ i\sqrt{p} & \text{if } p \equiv -1 \pmod{4}, \end{cases} \quad (1)$$

and the upper bound

$$|\sum_x e_p(x^d)| \leq ((d,p - 1) - 1)p^{1/2}$$

for any positive integer $d$. Weil (1948) proved the more general bound

$$|\sum_x e_p(g(x))| \leq (d - 1)p^{1/2}, \quad (2)$$

for any polynomial $g(x)$ in one variable over $\mathbb{F}_p$, of degree $d$ not divisible by $p$.

1991 Mathematics Subject Classification. 11L07, 11L03.

Key words and phrases. exponential sums.
For any positive integer \( j \) set

\[
S_j(g) = \sum_{x \in \mathbb{F}_{p_j}} \psi_j(g(x))
\]

where \( \psi_j() = e_p(Tr_j()) \) and \( Tr_j \) is the trace mapping from \( \mathbb{F}_{p_j} \) to \( \mathbb{F}_p \). Then, by the rationality of the \( L \)-function, there is a finite set of complex numbers \( \alpha_1, \ldots, \alpha_\ell, \beta_1, \ldots, \beta_m \), with \( \alpha_i \neq \beta_k \) for all \( i \) and \( k \), such that

\[
S_j(g) = \alpha_1^j + \cdots + \alpha_\ell^j - \beta_1^j - \cdots - \beta_m^j,
\]

for any positive integer \( j \). The values \( \alpha_i \) and \( \beta_j \) are characteristic values of the \( L \)-function associated with the \( S_j \). For simplicity we shall simply call them the characteristic values of \( S(g) \). Deligne (1980, Weil II) proved a very general version of the Riemann Hypothesis, which in this setting states that the modulus of each of the characteristic values is a power of \( \sqrt{p} \), the power being called the weight of the characteristic value. Bombieri (1978), proved that the total number of characteristic values is bounded above by \( (4d + 5)^n \), and so we have the upper bound

\[
|S(g)| \leq (4d + 5)^n p^{M/2}
\]

where \( M \) is the maximum weight of the characteristic values. Thus to bound the exponential sum \( S(g) \), it suffices to determine the maximum weight of the characteristic values.

The trivial Bound for \( S(g) \) is just \( |S(g)| \leq p^n \). If \( g \) is not of the form \( h^p - h + c \), with \( h \) a polynomial over \( \mathbb{F}_p \), then one saves a factor of \( p^{1/2} \) and obtains

\[
|S(g)| \leq (d - 1)p^{n - \frac{1}{2}}.
\]

For polynomials of the form \( h^p - h + c \) one has trivially \( |S(g)| = p^n \). The upper bound in (5) follows easily from the one variable bound of Weil (2) for polynomials of degree \( d \) not divisible by \( p \). Deligne (1977, Proposition 3.8) proved the general case. Deligne also established a sharp upper bound for polynomials nonsingular at infinity.

**Theorem.** Deligne (1973, Theorem 8.4). Let \( g \) be a polynomial of degree \( d \) not divisible by \( p \), whose maximal homogeneous part is nonsingular. Then \( S(g) \) has \( (d - 1)^n \) characteristic values each of modulus \( p^{n/2} \). In particular,

\[
|S(g)| \leq (d - 1)^n p^{n/2}.
\]

When \( n = 1 \) the inequality in (6) is just the result of Weil (2). The theorem of Deligne was generalized by Katz (1980, Theorem 5.1.1), Adolphson and Sperber (1989) and Denef and Loeser (1991, Theorem 9.2). The latter two papers deal with exponential sums over Tori, and give more general conditions under which one can reach the conclusion of Deligne's theorem.
Example 1. (Adapted from an example in Adolphson and Sperber (1989)). Let
\[ g(x) = a_1 x_1^{d_1} + \cdots + a_n x_n^{d_n} + h(x), \]
and suppose that none of the exponents \( d_i \) are divisible by \( p \), and that every nonzero monomial \( c_i x_1^{i_1} \cdots x_n^{i_n} \) in \( h \) satisfies \( \sum_{j=1}^n \frac{i_j}{d_j} < 1 \). Then \( S(g) \) has \( \Pi_{i=1}^n (d_i - 1) \) characteristic values, each of weight \( n \). Since the exponents \( d_i \) may be distinct the polynomial \( g \) may have bad singularities at infinity.

There is still much middle ground to be covered between the upper bounds of (5) and (6). Can one characterize those polynomials \( g \) for which one saves an extra factor of \( p^{1/2} \) in (5)? It would appear that this question has not been addressed in the literature, although the result may be known. An obvious case where one cannot save the factor of \( p^{1/2} \) is when \( g \) degenerates to a polynomial in a single variable, that is, \( g(x) = f(L(x)) \) for some polynomial \( f \) in a single variable of degree greater than one, and linear polynomial \( L(x) \). Davenport and Lewis (1962) showed that for a nondegenerate cubic polynomial \( g \), \( S(g) \ll p^{5/4} \) when \( n = 2 \), whence by the Riemann Hypothesis (and the analogous upper bound for any extension of \( \mathbb{F}_p \)) we now obtain \( S(g) \ll p \), and \( S(g) \ll p^2 \) when \( n = 3 \). Does one always obtain the extra savings of \( p^{1/2} \) for nondegenerate polynomials?

Example 2.
\[ \sum_{x,y} e_p(x^2 y^2) = p + \sum_{y \neq 0} \left( \sum_x e_p(x^2 y^2) \right) = p + (p - 1)G_p, \]
where \( G_p \) is the Gaussian sum in (1). This sum has three characteristic values, of weights 1, 2 and 3 respectively.

The polynomial in Example 2 is nondegenerate and yet does not admit an extra savings of \( p^{1/2} \). However, \( x^2 y^2 \) is a composite polynomial. As we shall see in Theorem 1, this is the characteristic that identifies when one cannot obtain the extra savings. We call the polynomial \( g \) composite if it can be expressed in the manner \( g(x) = f(h(x)) \) with \( f \) a polynomial in one variable over \( \overline{\mathbb{F}}_p \) of degree at least 2 and \( h \) any polynomial in \( n \) variables over \( \mathbb{F}_p \).

**Theorem 1.** If \( g \) is non-composite and \( p \geq c_1(d,n) \) then the weights of the characteristic values of \( S(g) \) are all \( \leq 2n - 2 \), and so
\[ |S(g)| \leq (4d_f + 5)^n p^{n-1}. \quad (7) \]

If \( g \) is homogeneous then we may take \( c_1(d,n) = 2 \). Conversely, if \( g \) is a composite polynomial of the form \( g(x) = f(h(x)) \), with \( h \) homogeneous and \( p \nmid d_f := \text{deg}(f) \), then there are at least \( d_f - 1 \) characteristic values of weight \( 2n - 1 \).

It is quite possible that the hypothesis \( p \geq c_1(d,n) \) is only due to our method of proof. We do not know at present if the converse statement in Theorem 1 holds when \( h \) is not homogeneous.
Theorem 1 is of particular interest when \( n = 2 \). In this case we obtain \( |S(g)| \ll p \), for non-composite \( g \), which is best possible in general. Another interesting special case is when \( g \) is of prime degree. In this case \( g \) is composite if and only if \( g \) degenerates to a polynomial in a single variable of degree at least 2. Thus, the result of Davenport and Lewis mentioned above is just a special case of Theorem 1.

It is worth pointing out that one occasionally saves a factor of \( p^{1/2} \) (or more) at the ground field level (due to cancellation in the sum (3)), even though all of the characteristic values have weight \( 2n - 1 \). Indeed, if \((d, p - 1) = 1\) then \( \sum_x e_p(x^d) = 0 \)! But if \( p \nmid d \) then this sum has \( d - 1 \) characteristic values each of weight 1. Theorem 2 provides further examples where this phenomenon occurs.

Is it possible to characterize those polynomials for which one saves yet another factor of \( p^{1/2} \)? We start by considering a special class of homogeneous polynomials.

**Theorem 2.** Suppose that \( g \) is a homogeneous polynomial over \( \mathbb{F}_p \) of degree \( d \) with \((d, p - 1) = 1\), and having the factorization

\[
g = P_1^{e_1} P_2^{e_2} \ldots P_s^{e_s},
\]

with the \( P_i \) distinct (nonassociate) irreducible polynomials over \( \mathbb{F}_p \). If exactly \( r \) of the factors \( P_i \) are absolutely irreducible then

\[
S(g) = (r - 1)p^{n-1} + O(p^{n-\frac{3}{2}}).
\]

In particular, if exactly one of the factors is absolutely irreducible then,

\[
|S(g)| \ll \frac{1}{p} p^{n-\frac{3}{2}}.
\]

It is important to point out that one cannot conclude from (9) that all of the weights of the characteristic values are \( \leq 2n - 3 \). Indeed, \( g \) could be a composite polynomial of the form \( g = P^e \) with \( P \) absolutely irreducible, in which case some of the weights are \( 2n - 1 \).

It is reasonable to ask whether (9) holds unconditionally for any absolutely irreducible homogeneous polynomial. The answer is no. One need only take a sum of two composite polynomials in distinct variables to obtain a counterexample.

**Example 3.** Let \( g(x, y, u, v) = (x^2 + y^2)^3 + (u^3 + v^3)^2 \) and assume \( p \neq 2, 3 \). Since \( X^3 + Y^2 \) is absolutely irreducible it follows from Schinzel's decomposability theorem (see Schinzel (1980), section 8, Theorem 11), that \( g \) is absolutely irreducible over any field. Now, since the number of solutions of the equation \( x^2 + y^2 = \lambda \) is \( p - \left( \frac{-1}{p} \right) \) for any nonzero \( \lambda \) we obtain,

\[
\sum_{x,y} e_p((x^2 + y^2)^3) = p \left( \frac{-1}{p} \right) \sum_\lambda e_p(\lambda^3).
\]

\( \dagger \)The constants in the big "Oh's" and \( \ll \) depend only on \( n \) and the degrees of the polynomials.
Also, by (21) we have
\[ \sum_{u,v} c_p((u^3 + v^3)^2) = p \sum_{\lambda} e_p(\lambda^2) + O(p). \]
Thus
\[ S(g) = p^2 G_p \sum_{\lambda} e_p(\lambda^3) + O(p^{\frac{5}{3}}). \]

The polynomial in the previous example is absolutely irreducible but it has a singular locus of dimension 1 in projective 3-space. In the next Theorem we give an upper bound for \( S(g) \) in terms of the dimension of the singular locus. Let \( g \) be any polynomial over \( \mathbb{F}_p \), not necessarily homogeneous, with maximal homogeneous part \( g_d \). The singular locus of \( g \) at infinity is the algebraic subset of projective \((n-1)\)-space \( \mathbb{P}^{n-1} \) defined by,
\[ \frac{\partial g_d}{\partial x_i} = 0, \quad 1 \leq i \leq n \quad g_d = 0. \]
Let \( \ell \) denote the dimension of the singular locus, putting \( \ell = -1 \) if \( g \) is nonsingular at infinity. Birch (1962) proved that
\[ |S(g)| \ll p^{n-2-\ell(n-\ell)}. \quad (10) \]
We sharpen the upper bound in (10) and at the same time generalize the result of Deligne (6).

**Theorem 3.** Suppose that \( p > c_2(d,n) \). Let \( g \) be a polynomial of degree \( d \) over \( \mathbb{F}_p \) such that its maximal homogeneous part is absolutely irreducible and has a singular locus of dimension \( \ell \). Then all of the characteristic values of \( S(g) \) have weight \( \leq n + \ell + 1 \) and we have
\[ |S(g)| \leq (4d + 5)^n p^{\frac{n+\ell+1}{2}}. \quad (11) \]
If \( g \) is homogeneous, then we may take \( c_2(d,n) = 2 \).

If \( n \geq 3 \) and the maximal homogeneous part of \( g \) is nonsingular, then \( g_d \) is necessarily absolutely irreducible and we obtain the best possible exponent \( \frac{n}{2} \) in (11). In this case the upper bound in (11) coincides with the upper bound of Deligne in (6), up to the constant factor. Skorobogatov (1992, Theorem 3.2) obtained (11) for the case of homogeneous \( g \) with \((d, p - 1) = 1\).

The upper bound in (11) is best possible for polynomials \( g \) that degenerate to polynomials nonsingular at \( \infty \) in fewer variables. A nondegenerate polynomial is given in Example 3 for which the upper bound in (11) is best possible. The results of Adolphson and Sperber, and Denef and Loeser provide special classes of polynomials for which one can improve on the upper bound in (11) when singularities are present at infinity; see Example 1.
An upper bound very similar to (10) can be given in terms of a factor called the $h$-
invariant of $g$, the smallest positive integer $h$ such that $g_d$ may be written as
\[ g_d = A_1 B_1 + A_2 B_2 + \cdots + A_h B_h \]
for some forms $A_i, B_i \in \mathbb{F}_p[x]$ of positive degree. Schmidt (1984, Theorem 1) established
\[ |S(g)| \ll p^{n-2^{1-\delta} [h/\Phi(d)]}, \]
where $[\alpha]$ denotes the smallest integer $> \alpha$, and one can take $\Phi(2) = \Phi(3) = 1$, $\Phi(4) = 3$, $\Phi(5) = 13$ and $\Phi(d) < (\log 2)^{-d} d!$ in general. Davenport and Lewis (1962) much earlier obtained (13) for the case of cubic polynomials. In general (13) is weaker than (11). Are there examples where the upper bound in (13) is sharper than (11)? When $d = 3$ this would require $2\ell + h + 3 > 2n$. We note that one always has $h > (n - \ell - 1)/2$ since any solution of $A_1 = B_1 = \cdots = A_h = B_h = 0$ is a singular point of $g_d$.

The author wishes to express his thanks to Shreeram Abhyankar for fruitful discussions on algebraic geometry and to the many others who helped steer him to the relevant sources.

**Section 2. Lemmas.**

**Lemma 1.** If $h$ is a non-composite polynomial defined over $\mathbb{F}_p$ of degree $d_h$ then for all but $c_3(d_h, n)$ values of $\tau \in \mathbb{F}_p$, the polynomial $h(x) - \tau$ is absolutely irreducible. If moreover $h$ is homogeneous (and non-composite) then $h(x) - \tau$ is absolutely irreducible for all nonzero values of $\tau$.

**Proof.** Since $h$ is non-composite, polynomial, it follows from a theorem of Bertini (see Shafarevich (1994), Chapter II, section 6.1, Theorem 1) that for all but finitely many values of $\tau$, the fibres $h^{-1}(\tau)$, in affine $n$-space, are irreducible. That the number depends only on $d_h$ and $n$ follows from elimination theory. By a theorem of Noether (see Schmidt (1976), Chapter V, Theorem 2A), we can take $c_3(d_h, n) = m^{2^m}$ where $m = \binom{n + d_h - 1}{n}$. For homogenous $h$ it is easy to see that for any nonzero $\tau$, $h - \tau$ is reducible over $\mathbb{F}_p$ if and only if $h - 1$ is reducible over $\mathbb{F}_p$. Since $h - \tau$ is reducible for at most finitely many values of $\tau$ it follows that it is absolutely irreducible for all nonzero $\tau$.

A more direct proof can be given for homogeneous $h$. We may assume $\tau = 1$. Let $k = \overline{\mathbb{F}}_p$, and $H(z, x) = z^d - h(x)$, with $d = d_h$. Certainly $h - 1$ is reducible over $k$ if and only if $H$ is reducible over $k$. View $H$ as a polynomial in $z$ with coefficients in $k(x)$. By the theorem of Capelli (1898) and Rédei (1959), (see also Schinzel (1982), section 13, Theorem 21), $H$ is reducible over $k(x)$ if and only if for some prime divisor $p$ of $d$, $h = b^p$, with $b \in k(x)$. But this implies that $h$ is composite. \(\square\)

It would be of interest to find the smallest possible value for $c_3(d_h, n)$. When $n = 2$ the best one can hope for is $c_3(d_h, 2) = d_h - 1$ as the following example illustrates.

**Example 4.** The polynomial $h(x, y) = y(x - r_1)(x - r_2) \cdots (x - r_d) + x - \tau$ is reducible if and only if $x = r_i$ for some $i$.

Y. Stein (1989) was able to establish this best possible value for the case $n = 2$ over fields of characteristic 0. One should also see the work of Ruppert (1986) and section 11 of Schinzel's book (1982) on polynomials for related material.
Lemma 2. Let \( g(x) \) be an absolutely irreducible homogeneous polynomial over \( \mathbb{F}_p \) with singular locus of dimension \( \ell \) in \( \mathbb{P}^{n-1} \). Then for any \( \tau \in \mathbb{F}_p \) the number \( N(\tau) \) of solutions of the equation \( g(x) = \tau \) with coordinates in \( \mathbb{F}_p \) satisfies

\[
|N(\tau) - p^{n-1}| \leq \begin{cases} 
p^{\frac{n+4\ell}{2}}, & \tau \neq 0 
p^{\frac{n+\ell+1}{2}}, & \tau = 0.
\end{cases}
\]

(14)

Proof. Let \( N_{\mathbb{P}^n}(\tau) \) denote the number of projective solutions of the equation \( g(x) = \tau x_0^d \) with coordinates in \( \mathbb{F}_p \), and \( N_{\mathbb{P}^{n-1}}(0) \) denote the number of projective solutions of the equation \( g(x) = 0 \), with coordinates in \( \mathbb{F}_p \). Then it is easy to see that for any nonzero \( \tau \),

\[
N(\tau) = N_{\mathbb{P}^n}(\tau) - N_{\mathbb{P}^{n-1}}(0)
\]

\[
= (p^{n-1} + \cdots + p + 1) + O(p^{n+4\ell}) - (p^{n-2} + \cdots + p + 1) + O(p^{n+\ell+1}) = p^{n-1} + O(p^{\frac{n+4\ell}{2}}),
\]

by Theorem 1 of Hooley (1991). If \( \tau = 0 \) then we use \( N(\tau) = (p - 1)N_{\mathbb{P}^{n-1}}(0) + 1 \) together with Hooley’s theorem to obtain (14).

Let \( \chi \) be any multiplicative character on \( \mathbb{F}_p^* \) of exponent \( d \); \( \chi^d = \chi_0 \), where \( \chi_0 \) is the principal character. Then for any homogeneous polynomial \( h(x) \) of degree \( d \) the character sum

\[
\sum_{x \in \mathbb{P}^{n-1}} \chi(h(x)),
\]

(15)

over all points of \( \mathbb{P}^{n-1} \) with coordinates in \( \mathbb{F}_p \), is well defined and has an associated set of characteristic values of integral weight.

Lemma 3. (i) Let \( h \) be any non-composite homogeneous polynomial in \( n \geq 2 \) variables over \( \mathbb{F}_p \) of degree \( d \). Then for any nonprincipal character \( \chi \) of exponent \( d \), the weights of the characteristic values of the sum in (15) are all \( \leq 2n - 3 \).

(ii) If moreover \( h \) is absolutely irreducible with singular locus of dimension \( \ell \) then the weights are all \( \leq n + \ell \). In particular

\[
\sum_{x \in \mathbb{P}^{n-1}} \chi(h(x)) \ll p^{\frac{n+4\ell}{2}}.
\]

(16)

Proof. For any value \( \tau \) in \( \mathbb{F}_p \) let \( N(\tau) \) denote the number of solutions of the equation \( h(x) = \tau \) with coordinates in \( \mathbb{F}_p \), and write

\[
N(\tau) = p^{n-1} + E(\tau).
\]

(17)

Since \( h \) is homogeneous it is easy to see that \( N(\tau) \) and \( E(\tau) \) depend only on the residue class \( \overline{\tau} \) of \( \tau \) in the quotient group \( \mathbb{F}_p^*/(\mathbb{F}_p^*)^d \), where here \( (\mathbb{F}_p^*)^d \) denotes the set of \( d^{th} \) powers. Let \( \{\tau_1, \ldots, \tau_{(p-1,d)}\} \) be a set of coset representatives for these residue classes. Then for any nonprincipal character \( \chi \) of exponent \( d \),

\[
\sum_{x \in \mathbb{P}^{n-1}} \chi(h(x)) = \frac{1}{p-1} \sum_{\tau \neq 0} \chi(\tau)N(\tau) = \frac{1}{p-1} \sum_{\tau \neq 0} \chi(\tau)E(\tau),
\]
and so
\[
\sum_{x \in \mathbb{F}^{n-1}} \chi(h(x)) = \frac{1}{(p-1,d)} \sum_{i=1}^{(p-1,d)} \chi(\tau_i)E(\tau_i). 
\] (18)

Now, if \( h \) is non-composite, then by Lemma 1 \( h(x) - \tau \) is absolutely irreducible for any nonzero \( \tau \) whence by the theorem of Lang and Weil (1954), \(|E(\tau)| \ll p^{n-3/2}\). It follows from (18) that
\[
|\sum_{x \in \mathbb{F}^{n-1}} \chi(h(x))| \ll p^{n-3/2}.
\]
Since the analogous upper bound holds when \( \mathbb{F}_p \) is replaced by \( \mathbb{F}_p^j \) and \( \chi \) by \( \chi_j = \chi(\mathcal{N}_{\mathbb{F}_p^j/\mathbb{F}_p}) \), we conclude by the theorem of Bombieri (1978, Theorem 3) that all of the characteristic values have weight \( \leq 2n - 3 \). The second part of the lemma follows from (18) and Lemma 2 in exactly the same manner. \( \Box \)

Section 3. Formulae for \( S(g) \).

Let \( g \) be a polynomial of the form \( g(x) = f(h(x)) \) with \( f \) and \( h \) polynomials over \( \mathbb{F}_p \) and \( h \) homogeneous of degree \( d = d_h \). Then using the notation of section 2 and (18) we have,
\[
S(g) = \sum_{\tau} e_p(f(\tau))N(\tau) = e_p(f(0))N(0) + \sum_{\tau \neq 0} N(\tau) \sum_{\tau \neq 0} e_p(f(\tau)) \\
= e_p(f(0))N(0) + \frac{1}{(p-1,d)} \sum_{i=1}^{(p-1,d)} N(\tau_i) \sum_{\tau \neq 0} \sum_{\chi^d=1} \chi(\tau/\tau_i) e_p(f(\tau)) \\
= e_p(f(0))N(0) + \sum_{\chi^d=1} \sum_{x \in \mathbb{F}^{n-1}} \chi(h(x)) \sum_{\tau \neq 0} \chi(\tau) e_p(f(\tau)).
\]

Pulling off the \( \chi = \chi_o \) term and writing
\[
N_p(h) = \frac{p^{n-1} - 1}{p - 1} + E_p(h),
\] (19)

for the number of projective zeros of \( h \) with coordinates in \( \mathbb{F}_p \), we obtain
\[
S(g) = pe_p(f(0))E_p(h) + (p^{n-1} + E_p(h)) \sum_{\tau} e_p(f(\tau)) \\
+ \sum_{\chi^d=\chi_0} \sum_{x \in \mathbb{F}^{n-1}} \chi(h(x)) \sum_{\tau \neq 0} \chi(\tau) e_p(f(\tau)).
\] (20)

The identity in (20) provides a way of expressing the characteristic values of \( S(g) \) in terms of the characteristic values associated with the terms on the righthand side. Of course,
one must first pass to an extension $\mathbb{F}_q$ of $\mathbb{F}_p$ with $d|q - 1$ in order to pick up a full set of $d$ multiplicative characters of exponent $d$. An alternate way of writing (20) is

$$S(g) = p^{n-1} \sum_{\tau} e_p(f(\tau)) + \frac{p}{p-1} e_p(f(0))E(0)$$

$$+ (p-1,d)^{-1} \sum_{i=1}^{(p-1,d)} E(\tau_i) \sum_{\lambda} e_p(f(\tau_i \lambda^d)). \quad (21)$$

If $g$ itself is homogeneous then we may take $f(x) = x$ and $h = g$. The identities in (20) and (21) simplify to

$$S(g) = pE_p(g) + \sum_{\chi^d=\chi_0} \sum_{x \in \mathbb{F}^{n-1}} \chi(g(x)) \sum_{\tau \neq 0} \chi(\tau)e_p(\tau), \quad (22)$$

and

$$S(g) = \frac{p}{p-1} E(0) + (p-1,d)^{-1} \sum_{i=1}^{(p-1,d)} E(\tau_i) \sum_{\lambda} e_p(\tau_i \lambda^d), \quad (23)$$

respectively.

**Section 4. Proof of Theorem 1.**

Suppose first that $g$ is any non-composite homogeneous polynomial. Then by (22), Lemma 3(i) and the fact that all of the characteristic values associated with $E_p$ have weight $\leq 2n - 4$, it follows that the characteristic values of $S(g)$ all have weight $\leq 2n - 2$.

Suppose next that $g$ is a composite polynomial of the form $g(x) = f(h(x))$ with $f$ and $h$ polynomials over some extension $\mathbb{F}_q$ of $\mathbb{F}_p$, with $h$ homogeneous and $d_f \geq 2$. By modifying $f$ and $h$ if need be we may assume that $h$ is non-composite. Furthermore, we claim that by multiplying $f$ and $h$ by constants if need be we may assume that $f$ and $h$ have coefficients in $\mathbb{F}_p$. Indeed, the maximal homogeneous part of $g$, $ah(x)^{d_f}$, for some constant $a \in \mathbb{F}_q$, has coefficients in $\mathbb{F}_p$. Thus, if $P(x)$ is any irreducible factor of $h(x)$ over $\mathbb{F}_q$ then all of its conjugates must also be factors of $h(x)$. But the product of the distinct conjugates is a polynomial over $\mathbb{F}_p$, and so $h$ can be expressed as a product of polynomials with coefficients in $\mathbb{F}_p$ up to a constant factor.

If $p \nmid d_f$ then the sum $S(f) = \sum_{\tau} e_p(f(\tau))$ has $d_f - 1$ characteristic values, each of weight 1. Thus the term $p^{n-1} \sum_{\tau} e_p(f(\tau))$ in (20) gives rise to $d_f - 1$ characteristic values of $S(g)$, each of weight $2n - 1$, while the weights associated with the remaining terms in (20) are all $\leq 2n - 2$.

Suppose now that $g$ is any non-composite polynomial. We shall apply Hooley’s method of bounding weights by looking at second moments. For $\lambda \in \mathbb{F}_p^*$, let

$$S_f(\lambda) = \sum_{x} \psi_f(\lambda g(x)), \quad (24)$$
so that $S_j(1) = S_j$, and let $M_j$ denote the second moment,

$$M_j = \sum_{\lambda \in \mathbb{F}_{p^j}, \lambda \neq 0} |S_j(\lambda)|^2.$$

The following proposition may be gleaned from Hooley’s (1982) work; the proof is given in Cochrane (1994, section 4.4).

**Proposition.** Suppose that $N$ and $C$ are positive real numbers such that $M_j \leq Cp^{N_j}$ for all $j$ sufficiently large, and that $p > \frac{C}{B} + 1$, where $B$ is the number of characteristic values of maximal weight. Then all of the characteristic values associated with the sums $S_j$ have weights strictly less than $N$.

Now, for any fixed constant $c$,

$$S_j(\lambda) = \sum_{\tau} \psi_j(\lambda \tau)(N_j(\tau) - c), \quad (25)$$

where $N_j(\tau)$ is the number of solutions of the equation $g(x) = \tau$ with coordinates in $\mathbb{F}_{p^j}$. Defining $S_j(0)$ formally by the expression in (25), it follows from Parsevals identity that

$$M_j \leq \sum_{\lambda \in \mathbb{F}_{p^j}} |S_j(\lambda)|^2 = p^j \sum_{\tau \in \mathbb{F}_{p^j}} |N_j(\tau) - c|^2. \quad (26)$$

By Lemma 1, the polynomial $g(x) - \tau$ is absolutely irreducible for all but $c_3(d, n)$ values of $\tau$. Putting $c = p^j(n-1)$ in (26) it follows from the estimate of Lang and Weil (1954) and the elementary upper bound $N_j(\tau) \leq p^{j(n-1)}$ for any $\tau$, that

$$M_j \ll p^j \sum_{\tau, g(x) = \tau \atop \text{abs. irred.}} p^{2j(n-\frac{d}{2})} + O(p^{j(2n-1)}) \ll p^{j(2n-1)}.$$

Thus, by the Proposition, all of the characteristic values have weight strictly less than $2n - 1$, for $p$ sufficiently large.

**Section 5. Proof of Theorem 2.**

If $(p - 1, d) = 1$ then (22) is just the well known identity,

$$S(g) = pE_{\overline{\varphi}}(g) = (p - 1)^{-1}(pN(g) - p^n), \quad (27)$$

where $N(g)$ is the number of zeros of $g$ with coordinates in $\mathbb{F}_p$, and $E_{\overline{\varphi}}(g)$ is as defined in (19). The earliest reference that I know of to (27) is the work of Chalk and Williams (1965, Lemma 15). Skorobogatov (1992) more recently used a version of (27) in his work on exponential sums.

Suppose that $g$ admits the factorization

$$g = p_1^{e_1} \cdots p_s^{e_s},$$
with the $P_i$ irreducible polynomials over $\mathbb{F}_p$, and say that exactly $r$ of the polynomials $P_i$ are absolutely irreducible. We first remark that by (27) the value of $S(g)$ depends only on the irreducible factors $P_i$ and not at all on the exponents $e_i$, (of course the $e_i$ must be such that the degree of $g$ is relatively prime to $p - 1$.) Let $N(P_i)$ denote the number of zeros of $P_i$ with coordinates in $\mathbb{F}_p$. If $P_i$ is absolutely irreducible then by the theorem of Lang and Weil (1954),

$$N(P_i) = p^{n-1} + O(p^{n-\frac{3}{2}}).$$

If $P_i$ is not absolutely irreducible, then the set of zeros of $P_i$ with coordinates in $\mathbb{F}_p$ is contained in an algebraic subset of codimension 2 and so $N(P_i) \ll p^{n-2}$. Also, the number of common zeros of two or more of the factors $P_i$ is $O(p^{n-2})$. Thus

$$N(g) = r(p^{n-1} + O(p^{n-\frac{3}{2}})) + O(p^{n-2}) = rp^{n-1} + O(p^{n-\frac{3}{2}}),$$

and so by (27),

$$S(g) = (r - 1)p^{n-1} + O(p^{n-\frac{3}{2}}).$$

(28)

Section 6. Proof of Theorem 3.

Suppose that $g$ is homogeneous and absolutely irreducible, with singular locus of dimension $\ell$ in $\mathbb{P}^{n-1}$. Then by Lemma 2, all of the characteristic values associated with $E_p(g)$ have weight $\leq n + \ell - 1$. It follows from Lemma 3(ii) and (22) that all of the characteristic values of $S(g)$ have weight $\leq n + \ell + 1$, proving Theorem 3 for the case of homogeneous polynomials.

For nonhomogeneous $g$ we again turn to the method of Hooley and the Proposition in section 4. The details have been worked out in section 4.6 of Cochrane (1994).

REFERENCES


A. Capelli, Sulla riduttibilità delle equazioni algebriche, Nota secunda, ibidem (3) 4 (1898), 84–90.


———, On the number of points on a complete intersection over a finite field, J. of Number Theory 38 (1991), 338–358.

N.M. Katz, Sommes exponentielles, Astérisque 79 (1980).


A.N. Skorobogatov, *Exponential sums, the geometry of hyperplane sections, and some diophantine problems*, Israel J. of Math. 80 (1992), 359–379.