

# BOUNDS ON EXPONENTIAL SUMS AND THE POLYNOMIAL WARING'S PROBLEM MOD $p$

TODD COCHRANE, CHRISTOPHER PINNER, AND JASON ROSENHOUSE

ABSTRACT. We give estimates for the exponential sum  $\sum_{x=1}^p \exp(2\pi i f(x)/p)$ ,  $p$  a prime and  $f$  a non-zero integer polynomial, of interest in cases where the Weil bound is worse than trivial. The results extend those of Konyagin for monomials to a general polynomial. Such bounds readily yield estimates for the corresponding polynomial Waring problem mod  $p$ ; namely the smallest  $\gamma$  such that  $f(x_1) + \cdots + f(x_\gamma) \equiv N \pmod{p}$  is solvable in integers for any  $N$ .

## 1. INTRODUCTION

For any polynomial  $f(x)$  over  $\mathbb{Z}$  and prime  $p$  let  $S(f, p)$  denote the complete exponential sum

$$(1.1) \quad S(f, p) = \sum_{x=1}^p e_p(f(x)),$$

where  $e_p(\cdot)$  is the additive character  $e_p(\cdot) = e^{\frac{2\pi i \cdot}{p}}$ . It is well known that bounds for such sums imply corresponding estimates for the number of solutions of certain congruences. For instance, if  $f$  is the monomial  $f(x) = x^d$  then from the elementary bound of Gauss,  $|S(x^d, p)| \leq (d-1)\sqrt{p}$ , one obtains (as in Lemma 9.1) that the number of solutions of the congruence

$$(1.2) \quad x_1^d + x_2^d + \cdots + x_\gamma^d \equiv N \pmod{p}$$

is at least

$$p^{\gamma-1} \left( 1 - \frac{p}{(\sqrt{p}/(d-1))^\gamma} \right)$$

for any integer  $N$ . Waring's problem mod  $p$  is the determination of the smallest integer  $\gamma = \gamma(d, p)$  such that (1.2) can be solved for all  $N$ . If  $p > (d-1)^6$  one obtains  $\gamma(d, p) \leq 3$ , that is, every integer may be represented as a sum of three  $d$ -th powers (mod  $p$ ). More generally, if  $p \geq (d-1)^{2+\epsilon}$  then  $\gamma(d, p) \leq \lceil \frac{d}{\epsilon} \rceil + 2$ . However, for  $p < (d-1)^2$  the bound

---

*Date:* November 19, 2002.

*1991 Mathematics Subject Classification.* 11L07; 11L03.

*Key words and phrases.* exponential sums.

of Gauss is trivial and therefore implies nothing about Waring's problem. For primes in this interval nontrivial bounds for Gauss sums have been obtained over the past decade by Shparlinski [20], [21], Konyagin [14], [15], Mullen and Shparlinski [18], Konyagin and Shparlinski [16] and Heath-Brown and Konyagin [11]. The bound of particular interest to us here is Konyagin's [14, Theorem 1],

$$(1.3) \quad |S(ax^d, p)| \leq p \left( 1 - \frac{c_\varepsilon}{(\log d)^{1+\varepsilon}} \right),$$

for  $d \geq 2$ ,  $p \nmid a$  and  $p \geq d \log d / (\log(\log d + 1))^{(1-\varepsilon)}$ . Although modest in strength, the importance of this bound is that it is nontrivial for values of  $p$  very small relative to  $d$ . Combining this bound with a result of Bovey [2], Konyagin was able to resolve (in the affirmative) the conjecture of Heilbronn that for the mod  $p$  Waring problem one has  $\gamma(d, p) \ll_\varepsilon d^\varepsilon$  provided that  $p/d \geq C(\varepsilon)$ . Further estimates for  $\gamma(d, p)$  have been obtained by Chowla [4], Chowla, Mann and Straus [5], Dodson [7], Dodson and Tietäväinen [8], Garcia and Voloch [9] and Tietäväinen [22].

Consider now a general polynomial  $f(x)$  of degree  $d$ , and let  $\gamma(f, p)$  be the smallest  $\gamma$  such that

$$f(x_1) + \cdots + f(x_\gamma) \equiv N \pmod{p}$$

can be solved in integers for all  $N$ . By an application of the Cauchy-Davenport Theorem, Carlitz, Lewis, Mills and Straus [3] proved that for any  $f$ ,

$$(1.4) \quad \gamma(f, p) \leq d \quad \text{for} \quad p > d,$$

generalizing the same bound of Hardy and Littlewood [10, p. 533] for monomials. The authors are unaware of any other estimates of  $\gamma(f, p)$  for general  $f$ . The classical bound of Weil [25],  $|S(f, p)| \leq (d-1)\sqrt{p}$ , implies (as above) that  $\gamma(f, p) \leq 3$  for  $p \geq (d-1)^6$  and for  $p \geq (d-1)^{2+\varepsilon}$ ,  $\gamma(d, p) \leq \lceil \frac{4}{\varepsilon} \rceil + 2$ . For  $p < (d-1)^2$  and general  $f$ , it is a major open problem to obtain nontrivial bounds for  $S(f, p)$ . Mordell [17] established the bound

$$(1.5) \quad |S(f, p)| \leq (k_1 k_2 \cdots k_r (p-1, k_1, k_2, \dots, k_r))^{\frac{1}{2r}} p^{1-\frac{1}{2r}},$$

for polynomials of the type

$$(1.6) \quad f(x) = c_1 x^{k_1} + c_2 x^{k_2} + \cdots + c_r x^{k_r}, \quad 1 \leq k_1 < \cdots < k_r < p-1, \quad p \nmid c_1 \cdots c_r,$$

but this bound is nontrivial only for a very restricted set of polynomials. Bounds of this type were also obtained by Akulinichev [1] and Karatsuba [12] for binomials and other special polynomials. Here we shall obtain an upper bound of the type (1.3) that remains nontrivial for very general  $f(x)$ .

For a prime  $p$  and vector of exponents

$$\vec{k} = (k_1, \dots, k_r), \quad 1 \leq k_1 < \dots < k_r < p - 1,$$

we are interested in obtaining bounds of the form

$$(1.7) \quad \left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \delta(\vec{k}) \right)$$

that hold for all integer polynomials of the type (1.6). (We require  $k_r$  to be strictly less than  $p-1$  to avoid the monomial  $x^{p-1}$ , which is identically 1 on the set of nonzero residues mod  $p$ .) Such bounds immediately yield (see Lemma 9.1) the estimate

$$(1.8) \quad \gamma(f, p) \leq \left\lceil \frac{\log p}{\delta(\vec{k})} \right\rceil.$$

In particular, analogous to the Heilbronn conjecture for monomials, our bound (Corollary 1.1) implies that for any  $\varepsilon > 0$ ,  $\gamma(f, p) \leq rd^\varepsilon$  for all polynomials (1.6) of degree  $d$  with  $r$  terms and  $p/d \geq C(\varepsilon)$ . It is not clear to us at present whether the dependence on  $r$  in this estimate can be removed.

We write

$$(1.9) \quad v_1 < \dots < v_J$$

for the set of distinct  $(k_i, p-1)$  that occur in  $\vec{k}$ . Then  $f(x)$  can be broken down into a sum of  $J$  polynomials

$$(1.10) \quad f(x) = f_1(x) + \dots + f_J(x)$$

where for  $j = 1, \dots, J$ , the exponents  $k_i$  occurring in  $f_j(x)$  all have the same value  $(k_i, p-1) = v_j$ . Define

$$(1.11) \quad r_j = |\{k_i : (k_i, p-1) = v_j\}|, \quad t_j = (p-1)/v_j,$$

and set

$$(1.12) \quad \sigma = \sum_{j=1}^J \phi(t_j),$$

with  $\phi$  being the Euler totient function. Our main theorem is

**Theorem 1.1.** *For any polynomial  $f$  as in (1.6) and prime  $p \geq p_0$ ,*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \frac{1}{(r \log p) \log^5(r \log p) p^{2r/\sigma}} \right).$$

The factor  $p^{2r/\sigma}$  on the right-hand side is necessary as we show in the next section, but it is an open question whether the factor  $r$  (in the denominator) can be eliminated when  $p/d \gg 1$ . The proof of the theorem follows the line of argument developed by Konyagin in [14]. We have sharpened, refined and extended each step of his work, but no doubt there is still more to be reaped from his ingenious method. Our results not only generalize his from monomials to polynomials, but in addition sharpen the results he obtained for monomials. They also encompass the bound for monomials obtained by Konyagin and Shparlinski (by a different method) in [16, Theorem 4.2].

A more precise version of Theorem 1.1, given in Theorem 6.1, reflects the dependence of the result on the currently best available lower bound for the Mahler measure of an integer polynomial having a nonzero root which is not a root of unity.

There are two other ways of stating our main result which may be of more use in practice. In terms of the degree of the polynomial  $f$  we obtain:

**Corollary 1.1.** *Let  $f(x)$  be an integer polynomial of degree  $d$  as in (1.6). Then*  
*(i) Form for small  $p$ . For an arbitrary  $\delta > 0$ , if  $p \geq \left(\frac{9}{\delta^{1.06}}\right)d$  and  $p \geq C(\delta)$ , then*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \frac{1}{rp^\delta} \right).$$

*(ii) Analogue of (1.3). For any  $\varepsilon > 0$ , if  $p \geq \frac{d \log d}{(\log(r \log d))^{1-\varepsilon}}$  and  $p \geq C(\varepsilon)$  then*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \frac{1}{(r \log p)^{1+\varepsilon}} \right).$$

*(iii) Precise form for slightly larger  $p$ . Let  $c, \kappa$  be the constants defined in (4.1) below, and  $\gamma$  Euler's gamma constant. For any  $\varepsilon > 0$ , if  $p \geq \frac{(1+\varepsilon)e^\gamma}{\log 2} d \log d \log \log d$  and  $p \geq C(\varepsilon)$  then*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \frac{(1-\varepsilon)}{\frac{32}{\log 2} c(r \log p) (\log(r \log p))^{2+\kappa}} \right).$$

Instead of using the degree  $d$ , the appropriate ranges can also be expressed in terms of the minimal  $(k_i, p-1)$ ,  $k := v_1$ , occurring in  $f$  and the number of terms  $r$ ; the three bounds (i), (ii) and (iii) holding when

$$(1.13) \quad \begin{aligned} p &\geq 7.7(r/\delta)^{1.05} k, & p &\geq p(\delta), \\ p &\geq \frac{(kr) \log(kr)}{(\log(r \log(kr)))^{1-\varepsilon}}, & p &\geq p(\varepsilon), \\ p &\geq \frac{(1+\varepsilon)e^\gamma}{\log 2} (kr) \log(kr) \log \log(r \log(kr)), & p &\geq p(\varepsilon), \end{aligned}$$

respectively. The first bound can also be made asymptotically precise in terms of the  $r$  dependence; for any  $\varepsilon > 0$

$$p \geq (1 + \varepsilon)e^\gamma k \left( \frac{2r}{\delta} \right) \log \log \left( \frac{2r}{\delta} \right), \quad p \geq p(\varepsilon\delta).$$

Recalling the decomposition  $f = f_1 + f_2 + \cdots + f_J$  given in (1.10) and the definition of  $\delta(\vec{k})$  in (1.7), we note that in many cases a bound for  $S(f, p)$  can be obtained directly from bounds for the  $S(f_i, p)$ . We write  $\vec{k}_j$  for the vector of exponents occurring in  $f_j$ .

**Theorem 1.2.** *Suppose that there is a value  $L \leq J$  such that  $v_j \nmid v_L$  for all  $j \neq L$ . Then*

$$\delta(\vec{k}) \geq \frac{1}{4^{J-1}} \delta(\vec{k}_L).$$

It is not clear whether the result still holds when there are  $v_j | v_L$  - in that case the method of proof requires us to add all the  $\vec{k}_j$  with  $v_j | v_L$  to the vector  $\vec{k}_L$  on the right-hand side.

## 2. LOWER BOUND EXAMPLES

We give examples showing the need for the term  $p^{2r/\sigma}$  in Theorem 1.1 for ‘small’  $p$  (with sharpness when  $J = 1$ , or where we can reduce to the case of smallest  $r_j/\phi(t_j)$  using Theorem 1.2). Since

$$\min_j \frac{r_j}{\phi(t_j)} \leq \frac{r}{\sum_{j=1}^J \phi(t_j)} \leq \max_j \frac{r_j}{\phi(t_j)},$$

the worst bounds would seem to arise for polynomials having  $J = 1$ .

For these lower bound examples we show the existence of polynomials with  $f(x)$  appropriately small mod  $p$  for all  $x$ . For any real number  $u$  we define

$$(2.1) \quad \|u\| = \min_{k \in \mathbb{Z}} |u - kp|.$$

**Theorem 2.1.** *For a set of exponents  $(k_1, \dots, k_r)$ ,  $1 \leq k_1 < \cdots < k_r \leq p - 1$ , with  $(p - 1, k_i) = k$  for all  $i$ ,  $p - 1 = kt$ , there is an integer polynomial  $f(x) = \sum_{i=1}^r a_i x^{k_i}$ , non-zero mod  $p$ , with*

$$(2.2) \quad \|f(x)\| \leq C(t)p^{1-r/\phi(t)}, \quad C(t) = \prod_{q|t} q^{\frac{1}{2(q-1)}} \ll \sqrt{\log t},$$

for all integers  $x$ , the product being over the distinct odd prime divisors  $q$  of  $t$ .

By contrast, Lemma 5.3 below will show that for any such polynomial there must be an integer  $x$  with

$$\|f(x)\| \geq \frac{p^{1-r/\phi(t)}}{4\sqrt{m \log 4m}}, \quad m = \min\{4t, c_\varepsilon(r \log p)^{1+\varepsilon}\}.$$

**Corollary 2.1.** *For any vector of exponents  $\vec{k} = (k_1, \dots, k_r)$  and corresponding  $v_j$  as in (1.9), there are integer polynomials*

$$f_j(x) = \sum_{(p-1, k_i)=v_j} c_i x^{k_i}, 1 \leq j \leq J \quad \text{and} \quad f(x) = f_1(x) + \dots + f_J(x),$$

each nonzero mod  $p$ , satisfying

$$(2.3) \quad \left| \sum_{x=1}^p e_p(f_j(x)) \right| \geq p \left( 1 - \frac{2\pi^2 C^2(t_j)}{p^{2r_j/\phi(t_j)}} \right), \quad j = 1, \dots, J,$$

and

$$(2.4) \quad \left| \sum_{x=1}^p e_p(f(x)) \right| \geq p \left( 1 - \frac{2\pi^2 \sum_{j=1}^J C^2(t_j)}{\min_j p^{2r_j/\phi(t_j)}} \right),$$

with  $C(t_j) = O(\sqrt{\log t_j})$  as in Theorem 2.1.

There is also an integer polynomial  $f(x) = \sum_{i=1}^r a_i x^{k_i}$ , nonzero mod  $p$ , with

$$(2.5) \quad \left| \sum_{x=1}^p e_p(f(x)) \right| \geq p \left( 1 - \frac{2\pi^2}{\lfloor p^{r/T} \rfloor^2} \right), \quad T = [t_1, \dots, t_j].$$

We immediately obtain:

**Corollary 2.2.** *For any  $A > 0$ , positive integers  $k, r$  and prime  $p < \frac{1}{A}(kr) \frac{\log(kr)}{\log(r \log(kr))}$  with  $k|(p-1)$ , there is an  $f(x) = \sum_{j=1}^r c_j x^{j^k}$ , of degree  $d \leq rk$ , with*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \geq p \left( 1 - \frac{4\pi^2}{(r \log p)^A} \right).$$

Also, for any  $0 < \delta < 2$ ,  $0 < \varepsilon < \frac{1}{2}$  and prime  $p$  of the form  $p = 1 + kt$  with  $t \geq t(\varepsilon, \delta)$ , there exists a polynomial  $f(x) = \sum_{i=1}^r c_i x^{i^k}$ ,  $(j_i, t) = 1$  for all  $i$ ,  $r = \lceil \frac{1}{2}(1 + \frac{3}{2}\varepsilon)\delta\phi(t) \rceil$ , with  $p \geq (1 - 2\varepsilon) \left(\frac{2}{\delta}\right) d$ , and

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \geq p \left( 1 - \frac{1}{p^{\delta(1+\varepsilon)}} \right).$$

If  $t = p_1 \cdots p_n$  (where  $p_i$  is the  $i$ -th prime) then for these polynomials we have  $p \geq (1 - 2\varepsilon)e^\gamma k \left(\frac{2r}{\delta}\right) \log \log \left(\frac{2r}{\delta}\right)$ .

These lower bounds generalize the results of Konyagin [14], and Konyagin & Shparlinski [16] and Powell [19], for monomials. Other lower bounds occur in Karatsuba [12].

### 3. PRELIMINARY LEMMAS

We first transform the exponential sum problem into one of obtaining lower bounds for sums of the form  $\sum_{i=0}^m \|f(\alpha^i)\|^2$ , where  $\alpha$  is a primitive root mod  $p$ .

**Lemma 3.1.** *Suppose that  $f(x)$  is an integer polynomial of the type (1.6) and  $\alpha$  is a primitive root mod  $p$ . If  $\sum_{i=0}^m \|F(\alpha^i)\|^2 \geq A$  for any integer polynomial of the form  $F(x) = \sum_{i=1}^r a_i x^{k_i}$ ,  $p \nmid a_i$  for all  $i$ , then*

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p - \frac{2A(p-1)}{(m+1)p^2}.$$

*Proof.* First note that since  $|\sin y| \geq \frac{2}{\pi}|y|$  for  $|y| \leq \frac{\pi}{2}$ , one has

$$\cos(2y) = 1 - 2\sin^2(y) \leq 1 - \frac{8}{\pi^2}y^2,$$

for  $|y| \leq \frac{\pi}{2}$ . Choosing  $x_0 \in \mathbb{R}$  suitably we have

$$\begin{aligned} \left| \sum_{x=1}^{p-1} e_p(f(x)) \right| &= \Re e \sum_{x=1}^p e_p(f(x) - x_0) \\ &= \sum_{x=1}^p \cos\left(\frac{2\pi}{p}\|f(x) - x_0\|\right) \\ &\leq \cos\left(\frac{2\pi}{p}\| -x_0\|\right) + \sum_{x=1}^{p-1} \left(1 - 8\frac{\|f(x) - x_0\|^2}{p^2}\right) \\ &\leq p - \frac{8}{p^2} \sum_{x=1}^{p-1} \|f(x) - x_0\|^2. \end{aligned}$$

Observe that for any real number  $x_0$

$$\begin{aligned} \sum_{x=1}^{p-1} \|f(x) - f(x\alpha)\|^2 &= \sum_{x=1}^{p-1} \|f(x) - x_0 + x_0 - f(x\alpha)\|^2 \\ &\leq \sum_{x=1}^{p-1} 2\|f(x) - x_0\|^2 + \sum_{x=1}^{p-1} 2\|f(x\alpha) - x_0\|^2 \\ &= \sum_{x=1}^{p-1} 4\|f(x) - x_0\|^2. \end{aligned}$$

Thus for  $f(x) = \sum_{l=1}^r c_l x^{k_l}$  we have

$$\sum_{x=1}^{p-1} \|f(x) - x_0\|^2 \geq \frac{1}{4} \sum_{x=1}^{p-1} \|f(x) - f(x\alpha)\|^2 = \frac{1}{4} \sum_{x=1}^{p-1} \left\| \sum_{l=1}^r c_l (1 - \alpha^{k_l}) x^{k_l} \right\|^2,$$

where

$$\begin{aligned} \sum_{x=1}^{p-1} \left\| \sum_{l=1}^r c_l (1 - \alpha^{k_l}) x^{k_l} \right\|^2 &= \frac{1}{m+1} \sum_{i=0}^m \sum_{y=1}^{p-1} \left\| \sum_{l=1}^r c_l (1 - \alpha^{k_l}) (y\alpha^i)^{k_l} \right\|^2 \\ &= \frac{1}{m+1} \sum_{y=1}^{p-1} \sum_{i=0}^m \|F_y(\alpha^i)\|^2 \geq \frac{(p-1)A}{(m+1)}. \end{aligned}$$

□

The following is a slightly cleaner version of Theorem 5, Chapter 2, §2 of [13]:

**Lemma 3.2.** *The Rademacher functions,  $\{\psi_n(x)\}_{n=0}^\infty$ , on  $[0, 1)$ ,*

*$\psi_n(x) = 1$ , if  $x \in [\frac{i}{2^n}, \frac{(i+1)}{2^n})$ ,  $0 \leq i < 2^n$ ,  $i$  even,  $-1$ , if  $x \in [\frac{i}{2^n}, \frac{(i+1)}{2^n})$ ,  $0 \leq i < 2^n$ ,  $i$  odd,*

*satisfy*

$$\text{meas} \left\{ x \in (0, 1) : \left| \sum_{i=0}^L a_i \psi_i(x) \right| \geq \lambda \right\} \leq 2 \exp \left( -\frac{1}{2} \lambda^2 / \sum_{i=0}^L a_i^2 \right).$$

*Proof.* Clearly if  $f$  is a non-negative function on  $(0, 1)$  then

$$\int_0^1 f(u) du \geq t \text{ meas} \{ x \in (0, 1) : f(x) \geq t \}.$$

Hence for  $p(x) = \sum_{i=0}^L a_i \psi_i(x)$  and any positive  $z$  we have

$$\begin{aligned} \text{meas} \{ x \in (0, 1) : |p(x)| \geq \lambda \} &= \text{meas} \{ x \in (0, 1) : e^{z|p(x)|} \geq e^{z\lambda} \} \\ &\leq e^{-z\lambda} \int_0^1 e^{z|p(u)|} du \\ &\leq e^{-z\lambda} \int_0^1 (e^{zp(u)} + e^{-zp(u)}) du \\ &= 2e^{-z\lambda} \frac{1}{2^{L+1}} \sum_{\pm} e^{z(\pm a_0 \pm \dots \pm a_L)} \\ &= 2e^{-z\lambda} \prod_{i=0}^L \frac{1}{2} (e^{za_i} + e^{-za_i}). \end{aligned}$$

Observing that

$$\cosh(u) = 1 + \sum_{i=1}^{\infty} \frac{u^{2i}}{(2i)!} \leq 1 + \sum_{i=1}^{\infty} \frac{u^{2i}}{2^i i!} = e^{\frac{1}{2}u^2}$$

gives

$$\text{meas} \{ x \in (0, 1) : |p(x)| \geq \lambda \} \leq 2 \exp \left( -z\lambda + \frac{1}{2} z^2 \sum_{i=0}^L a_i^2 \right),$$

and choosing  $z = \lambda / \sum_{i=0}^L a_i^2$  yields the desired bound.  $\square$

Next we give an asymptotically sharp version of Dobrowolski's [6] Lemma 3.

**Lemma 3.3.** *Suppose that  $\alpha$  is algebraic with conjugates  $\alpha_1, \dots, \alpha_n$ .*

*Let  $L(r) := |\{i : \alpha_i^r = \alpha^r\}|$ , and  $P := \{\text{primes } p : L(p) > 1\}$ . Then for  $n \geq n(\varepsilon)$*

$$|P| \leq (1 + \varepsilon) \frac{\log n}{\log \log n},$$

*and thus for any  $s \geq (1 + \varepsilon) \log n$  the interval  $(s, 2s)$  always contains a prime  $q$  such that the  $\alpha_i^q$  are all distinct.*

*Proof.* Note that the minimum polynomial  $g_r(x)$  for  $\alpha^r$  is also the minimum polynomial for the remaining  $\alpha_i^r$  (since any automorphism of the corresponding splitting field  $\sigma : \alpha \mapsto \alpha_i$  fixes  $g_r$ ). Therefore

$$f_r(x) := \prod_{i=1}^n (x - \alpha_i^r) = g_r(x)^{L(r)},$$

and  $L(r)$  remains unchanged if we replace  $\alpha$  by one of its conjugates. It follows that

$$L(r) = n/[\mathbb{Q}(\alpha^r) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}(\alpha^r)].$$

We first show that for a prime  $p$

$$L(rp) = L(r), \quad pL(r) \quad \text{or} \quad \zeta_p \in \mathbb{Q}(\alpha^r)$$

where  $\zeta_p$  denotes a  $p$ th root of unity. If  $F = x^p - \alpha^{rp}$  is irreducible over  $\mathbb{Q}(\alpha^{rp})$  then plainly

$$L(rp)/L(r) = [\mathbb{Q}(\alpha^r) : \mathbb{Q}(\alpha^{rp})] = p,$$

so we can assume that  $F$  factors. Since the roots are  $\zeta_p^j \alpha^r$ ,  $j = 0, \dots, p-1$  the constant term of the minimum polynomial (for  $\alpha^r$  over  $\mathbb{Q}(\alpha^{rp})$ ) must take the form  $\zeta_p^k \alpha^{rl}$  for some  $l < p$ . Hence taking  $Jl \equiv 1 \pmod{p}$  we have  $\zeta_p^{Jk} \alpha^r \in \mathbb{Q}(\alpha^{rp})$ . If  $\zeta_p^{Jk} = 1$  then  $L(rp) = L(r)$ , otherwise  $\mathbb{Q}(\alpha^r)$  contains the  $p$ th roots of unity.

Now, as shown in [6],

$$(r, s) = 1 \Rightarrow L(rs) \geq L(r)L(s).$$

Indeed, each of the  $L(r)$  conjugates  $\alpha_i$  with  $\alpha_i^r = \alpha^r$  will have  $L(s)$  conjugates  $\alpha_j$  with  $\alpha_j^s = \alpha_i^s$  (and thus  $\alpha_j^{rs} = \alpha^{rs}$ ) where these sets must be distinct; since  $\alpha_j^s = \alpha_i^s = \alpha_i^s$ ,  $\alpha_i^r = \alpha_{i'}^r$  would give  $\alpha_i = \alpha_i^{gcd(r,s)} = \alpha_{i'}^{gcd(r,s)} = \alpha_{i'}$ . It follows from this inequality that if  $p \nmid r$  and  $L(p) > 1$  then  $L(pr) \neq L(r)$ . Thus if  $p \in P$  with  $p \nmid r$ , then either  $L(pr) = pL(r)$  or  $\zeta_p \in \mathbb{Q}(\alpha^r)$ .

Next we observe that a field containing the  $r$ th and  $s$ th roots of unity with  $(r, s) = 1$ , must contain the  $(rs)$ th roots of unity. By considering the primes of  $P$  in turn, we can write  $P = P_1 \cup P_2$ ,  $r = \prod_{p \in P_1} p$ ,  $t = \prod_{p \in P_2} p$  so that

$$L(r) = r, \quad \zeta_t \in \mathbb{Q}(\alpha^r).$$

Therefore

$$n = rn/L(r) = r[\mathbb{Q}(\alpha^r) : \mathbb{Q}] \geq r[\mathbb{Q}(\zeta_t) : \mathbb{Q}] = r\phi(t),$$

and, taking logs,

$$\sum_{p \in P} \log(p-1) \leq \log n.$$

Since the sum over the first  $k$  primes  $\sum_{i=1}^k \log(p_i - 1) \sim k \log k$  we must clearly have  $|P| \leq (1+o(1)) \log n / \log \log n$ . The second claim follows from the prime number theorem  $\pi(2s) - \pi(s) \sim s / \log s$ .  $\square$

**Lemma 3.4.** *Suppose that  $\beta_1, \dots, \beta_r$  are distinct integers modulo  $p$ . If*

$$u_1 \beta_1^j + \dots + u_r \beta_r^j \equiv 0 \pmod{p}, \quad j = 0, \dots, r-1,$$

*then the  $u_i \equiv 0 \pmod{p}$ ,  $i = 1, \dots, r$ . In particular, if  $\alpha$  is a primitive root mod  $p$  then, for a polynomial of the form (1.6) with values  $t_j$  as in (1.11), the sequence  $x_i \equiv f(\alpha^i) \pmod{p}$  has period  $[t_1, \dots, t_J]$ .*

*Proof.* The first part follows immediately on observing that the matrix

$$\mathcal{A} = \begin{pmatrix} 1 & \dots & 1 & \beta_1 & \dots & \beta_r & \dots & \beta_1^{r-1} & \dots & \beta_r^{r-1} \end{pmatrix}$$

has Vandermonde determinant  $|\det \mathcal{A}| = \prod_{i < j} |\beta_i - \beta_j| \not\equiv 0 \pmod{p}$ .

For the second part, writing  $\beta_i = \alpha^{k_i}$ , the sequence  $f(\alpha^i)$  has period  $T$  if and only if  $c_1(\beta_1^T - 1)\beta_1^j + \dots + c_r(\beta_r^T - 1)\beta_r^j \equiv 0 \pmod{p}$  for all  $j$ . Thus one must have  $\beta_i^T \equiv 1 \pmod{p}$ , that is,  $\frac{(p-1)}{(k_i, p-1)} | T$  for all  $i$ .  $\square$

**Lemma 3.5.** *Suppose that  $\beta = \alpha^k$  where  $\alpha$  is a primitive root mod  $p$  and that  $t = (p-1)/(p-1, k)$ . Then the  $n$ th cyclotomic polynomial  $\Phi_n(x)$  satisfies*

$$\Phi_n(\beta) \equiv 0 \pmod{p} \iff n = p^\mu t \text{ for some } \mu \geq 0.$$

*Proof.* Observe, from the relations  $\Phi_{p^\mu s}(x) = \Phi_{ps}(x^{p^{\mu-1}})$  and  $\Phi_{ps}(x) = \Phi_s(x^p)/\Phi_s(x)$  for  $(s, p) = 1$ , that  $\Phi_{p^k s}(x) \equiv \Phi_s(x)^{\phi(p^k)} \pmod{p}$ . Thus we may assume that  $\mu = 0$ . In this case the result is just [24, Lemma 2.9].  $\square$

#### 4. MAHLER MEASURE AND LEHMER'S PROBLEM

Recalling the definition of the Mahler measure,  $\mu(F)$ , of a polynomial

$$F(x) = a_N \prod_{i=1}^N (x - \alpha_i),$$

$$\mu(F) = |a_N| \prod_{i=1}^N \max\{1, |\alpha_i|\},$$

we shall assume that  $c$  and  $\kappa$  are constants such that for any  $\varepsilon > 0$  we have a lower bound of the form

$$(4.1) \quad \log \mu(F) \geq \frac{(1 - \varepsilon)}{c(\log N)^\kappa},$$

for all polynomials  $F(x)$  in  $\mathbb{Z}[x]$  of degree  $N \geq N(\varepsilon)$  containing a nonzero root which is not a root of unity. By a result of Dobrowolski [6] we can take  $\kappa = 3$  with  $c$  as small as

we like. The unresolved problem of Lehmer amounts to the claim that we can take  $\kappa = 0$ , with perhaps  $c = 1/\log(1.17628\dots)$ .

We need a lower bound on the Mahler measure to apply the following result of Konyagin [14, Lemma 3].

**Lemma 4.1.** *Suppose that  $P$  is an irreducible integer polynomial of degree  $n \geq 2$  whose roots satisfy  $z_i^q \neq z_j^q$  for  $i \neq j$ . Suppose that  $\lambda_n$  is a lower bound for the Mahler measure of an irreducible integer polynomial of degree at most  $n$  with a nonzero non-cyclotomic root, and  $Q$  a polynomial of degree  $N$  with coefficients from  $\{0, 1, -1\}$ . If  $q > \log(N+1)/\log \lambda_n$  and  $P(z)$  divides  $Q(z^q)$  then the roots of  $P$  are all roots of unity.*

## 5. THE FAMILY OF POLYNOMIALS $\Lambda(X)$ AND THE MAIN LEMMAS

For a set of integers  $X = (x_0, \dots, x_M)$ , define  $\Lambda(X)$  to be the set of polynomials  $p(z) = \sum_{i=0}^n a_i z^i$ ,  $a_n \neq 0$ , in  $\mathbb{Z}[x]$  with  $n \leq M$  and  $\sum_{i=0}^n a_i x_{i+j} = 0$  for  $j = 0, \dots, M-n$ . We will use the following property of  $\Lambda(X)$  established by Konyagin [14, Lemma 5 and its Corollary].

**Lemma 5.1.** *If  $P_1, P_2$  are in  $\Lambda(X)$  with  $\deg(P_1 P_2) \leq M$  then  $\gcd(P_1, P_2) \in \Lambda(X)$ . In particular if  $\Lambda(X)$  contains a polynomial  $P_0$  with  $\deg P_0 \leq \frac{1}{2}M$  then the polynomial of minimal degree in  $\Lambda(X)$  is unique (to within a constant) and must divide  $P_0$ .*

In the next two lemmas we generalize and refine the main lemma of Konyagin's work [14, Lemma 6].

**Lemma 5.2.** *Suppose that  $p$  is a prime,  $n, m$  positive integers with*

$$n \leq \frac{1}{4}m,$$

*and  $x_0, \dots, x_m$  integers satisfying*

$$(5.1) \quad x_i \equiv a_1 \beta_1^i + \dots + a_r \beta_r^i \pmod{p}, \quad i = 0, \dots, m,$$

*for some integers  $a_j, \beta_j$  coprime to  $p$ . Set*

$$\begin{aligned} m_1 &:= \left\lfloor \frac{1}{2}m \right\rfloor, \\ X &:= (x_0, \dots, x_{m_1}), \\ N &= \left\lceil r \frac{\log p}{\log 2} \right\rceil, \end{aligned}$$

*and*

$$(5.2) \quad A := \frac{p^2}{8 \left\lceil \frac{1}{2}(2p^r)^{1/(n+1)} \right\rceil^2 \log(4m)}.$$

Assume that

$$\sum_{i=0}^m x_i^2 \leq A.$$

Then there exists a nonzero integer polynomial  $P_1(z) = \sum_{i=0}^n b_i z^i$  in  $\Lambda(X)$ . Further, for any positive integer  $q$  with

$$Nq \leq \frac{1}{2}m_1$$

there exists a nonzero integer polynomial  $Q(z) = \sum_{i=0}^N c_i z^i$  with the  $c_i = 0, \pm 1$  and  $Q(z^q)$  in  $\Lambda(X)$ .

*Proof.* Let  $\vec{b}$  be a vector  $\vec{b} = (b_0, \dots, b_n)$  in  $\mathbb{N}^{n+1}$  with  $1 \leq b_i \leq B := \left\lceil \frac{1}{2}(2p^r)^{1/(n+1)} \right\rceil$ . For  $0 \leq j \leq m - n$  we have by Lemma 3.2 that

$$\begin{aligned} \text{meas} \left\{ x \in (0, 1) : \left| \sum_{i=0}^n b_i x_{i+j} \psi_i(x) \right| \geq \frac{1}{2}p \right\} &\leq 2 \exp \left( -\frac{1}{8}p^2 / \sum_{i=0}^n b_i^2 x_{i+j}^2 \right) \\ &\leq 2 \exp \left( -\frac{1}{8}p^2 / AB^2 \right) \leq \frac{1}{2m}. \end{aligned}$$

Hence

$$\text{meas} \left\{ x \in (0, 1) : \left| \sum_{i=0}^n b_i x_{i+j} \psi_i(x) \right| < \frac{1}{2}p, j = 0, \dots, m - n \right\} \geq \frac{1}{2},$$

and there are at least  $2^n$  distinct  $(\psi_0(x), \dots, \psi_n(x))$  with

$$\left| \sum_{i=0}^n b_i x_{i+j} \psi_i(x) \right| < \frac{1}{2}p, j = 0, \dots, m - n.$$

Since there are  $2^n B^{n+1} > p^r$  choices of  $\vec{b}$  and  $x$  we can find two pairs  $\vec{b}', x'$  and  $\vec{b}'', x''$  with

$$\sum_{i=0}^n b'_i \psi_i(x') \beta_i^l \equiv \sum_{i=0}^n b''_i \psi_i(x'') \beta_i^l \pmod{p}, \quad l = 1, \dots, r.$$

Thus taking  $b_i = b'_i \psi_i(x') - b''_i \psi_i(x'')$  we have

$$\left| \sum_{i=0}^n b_i x_{i+j} \right| < p, \quad j = 0, \dots, m - n,$$

with

$$\sum_{i=0}^n b_i x_{i+j} \equiv \sum_{l=1}^r a_l \beta_l^j \left( \sum_{i=0}^n b_i \beta_i^l \right) \equiv 0 \pmod{p}, \quad j = 0, \dots, m - n.$$

Hence all the  $\sum_{i=0}^n b_i x_{i+j} = 0$ ,  $j = 0, \dots, m - n$  and the non-zero polynomial  $P(z) = \sum_{i=0}^n b_i z^i$  is in  $\Lambda(X)$  (since  $m_1 - n_1 \leq m/2 < m - n$  where  $n_1 \leq n$  is the degree of  $P$ ).

The construction of the  $0, \pm 1$  polynomial is similar; since

$$\text{meas} \left\{ x \in (0, 1) : \left| \sum_{i=0}^N x_{iq+j} \psi_i(x) \right| < \frac{1}{2}p, j = 0, \dots, m - Nq \right\} \geq \frac{1}{2},$$

we have at least  $2^N > p^r$  distinct choices of  $(\psi_0(x), \dots, \psi_N(x))$  with  $\left| \sum_{i=0}^N x_{iq+j} \psi_i(x) \right| < \frac{1}{2}p$ ,  $j = 0, \dots, m - Nq$ . Thus we are guaranteed two  $(\psi_0(x), \dots, \psi_N(x))$ ,  $(\psi_0(x'), \dots, \psi_N(x'))$  with

$$\sum_{i=0}^N \psi_i(x) \beta_l^{iq} \equiv \sum_{i=0}^N \psi_i(x') \beta_l^{iq} \pmod{p}, \quad l = 1, \dots, r.$$

Hence, taking  $c_i = \frac{1}{2}(\psi_i(x) - \psi_i(x'))$ , we similarly obtain  $\sum_{i=0}^N c_i x_{j+iq} = 0$ ,  $j = 0, \dots, m - Nq$ , and  $Q(z^q) = \sum_{i=0}^N c_i x^{qi}$  is in  $\Lambda(X)$  (since  $m_1 - N_1q \leq m - Nq$ , where  $N_1q \leq Nq$  is the degree of  $Q(z^q)$ ).  $\square$

**Lemma 5.3.** *Suppose that  $\beta_i = \alpha^{k_i}$ ,  $i = 1, \dots, r$ , where  $\alpha$  is a primitive root mod  $p$ , and the  $1 \leq k_1 < \dots < k_r < p - 1$  are distinct integers, with  $t_1, \dots, t_J$  the distinct values of  $(p - 1)/(p - 1, k_i)$ .*

*Suppose that  $m$  is a positive integers satisfying  $m \geq 4T$ ,  $T = [t_1, \dots, t_J]$ , or for some  $\varepsilon > 0$*

$$m \geq \frac{8c}{\log 2} (1 + \varepsilon) (r \log p) (\log(r \log p))^{1+\kappa}, \quad \text{and} \quad p \geq C(\varepsilon),$$

*with  $\kappa, c$  as given in (4.1), and that  $n$  is a positive integer with*

$$n < \min \left\{ \frac{m}{4}, \sum_{j=1}^J \phi(t_j) \right\}.$$

*Then, with the  $x_i$  and  $A$  as defined in (5.1) and (5.2) of Lemma 5.2, we must have*

$$\sum_{i=0}^m x_i^2 > A.$$

*Proof.* We suppose on the contrary that  $\sum_{i=0}^m x_i^2 \leq A$ . Let  $P(z) = \sum_{i=0}^t w_i z^i$ ,  $t \leq n$ , be a polynomial of smallest degree in  $\Lambda(X)$ .

We first show that all the nonzero roots of  $P(z)$  must be roots of unity. If  $m \geq 4T$  then, since  $T$  is by Lemma 3.4 the period of the sequence  $x_i$ , plainly  $z^T - 1$  is in  $\Lambda(X)$  with  $T \leq \frac{1}{2}m_1$ , and by Lemma 5.1 we must have  $P(z) | z^T - 1$ . Note that  $T \geq \sum_{j=1}^J \phi(t_j) \geq r$  (the first inequality follows since the roots of  $z^T - 1$  include the  $\phi(t_j)$  primitive  $t_j$ th roots of unity for  $j = 1, \dots, J$ , the second since the number of exponents  $k_i$  with  $t_j = (p - 1)/(p - 1, k_i)$  is at most  $\phi(t_j)$ ).

Otherwise, suppose that  $P(z)$  has a nonzero non-cyclotomic root  $z_1$  with conjugates  $z_1, \dots, z_{n'}$ . Take  $s = (1 + \frac{1}{2}\varepsilon)c(\log m)^{1+\kappa}$  then plainly  $s > (1 + o(1)) \log n$  for sufficiently large  $p$  and by Lemma 3.3 we can find a prime  $q$  in  $(s, 2s)$  such that the  $z_i^q$  are distinct. Now

$$Nq < 2sN \leq \frac{2c}{\log 2} \left( 1 + \frac{3}{4}\varepsilon \right) (r \log p) (\log m)^{1+\kappa} < \frac{1}{2}m_1,$$

so by Lemma 5.2 there is a  $0, \pm 1$  polynomial  $Q(z^q)$  in  $\Lambda(X)$  with  $\deg(Q(z^q)) \leq \frac{1}{2}m_1$  and hence  $P(z)|Q(z^q)$ . But for  $p \geq C(\varepsilon)$

$$\frac{\log(N+1)}{\log \lambda_n} \leq \left(1 + \frac{1}{8}\varepsilon\right) \log(r \log p) c(\log n)^\kappa < \left(1 + \frac{1}{4}\varepsilon\right) c(\log m)^{1+\kappa} < s < q$$

and Lemma 4.1 implies that the non-zero roots of  $P(z)$  are in fact all roots of unity.

Now since  $P(z) \in \Lambda(X)$  we have

$$\sum_{i=0}^t w_i \left( \sum_{l=1}^r a_l \beta_l^{i+j} \right) = a_1 P(\beta_1) \beta_1^j + \cdots + a_r P(\beta_r) \beta_r^j \equiv 0 \pmod{p},$$

for  $j = 0, \dots, m_1 - t$ , and, since  $m_1 - t \geq m_1 - n \geq m/4 - 1 \geq r - 1$  and  $p \nmid a_i$ , by Lemma 3.4 we must have  $P(\beta_i) \equiv 0 \pmod{p}$  for each  $i = 1, \dots, r$ . Thus from Lemma 3.5, since it consists only of cyclotomics (or powers of  $x$ ),  $P(z)$  must contain a cyclotomic factor of the form  $\Phi_{t_j p^{\mu_j}}(x)$ ,  $\mu_j \geq 0$ , for each  $j = 1, \dots, J$ , and

$$n \geq \deg(P) \geq \sum_{j=1}^J \phi(t_j p^{\mu_j}) \geq \sum_{j=1}^J \phi(t_j).$$

For  $n < \sum_{j=1}^J \phi(t_j)$  this is a contradiction. So  $\sum_{i=0}^m x_i^2 > A$  as claimed.  $\square$

## 6. PROOF OF THEOREM 1.1 AND COROLLARY 1.1

We shall prove the following more precise form of Theorem 1.1. Let  $f(x) = \sum_{i=1}^r c_i x^{k_i}$  be a polynomial over  $\mathbb{Z}$  with  $p \nmid c_1 c_2 \cdots c_r$  and define  $t_i = (p-1)/(p-1, k_i)$  and  $\sigma = \sum_{j=1}^J \phi(t_j)$  as before (1.11), (1.12). Let  $c, \kappa$  be values such that the Mahler measure lower bound (4.1) holds as indicated.

**Theorem 6.1.** *Suppose that  $\varepsilon > 0$  and that*

$$(6.1) \quad m = \min \left\{ 4[t_1, \dots, t_J], \quad \left\lceil (1 + \varepsilon) \frac{8c}{\ln 2} (r \log p) (\log(r \log p))^{1+\kappa} \right\rceil \right\},$$

where in the latter case we also assume that  $p \geq C(\varepsilon)$ . Then

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p - \frac{(p-1)}{4(m+1) \log(4m) \left\lceil \frac{1}{2} (2p^r)^{1/\sigma} \right\rceil^2}.$$

*Proof.* From Lemma 5.3 we have

$$\sum_{i=0}^m \|F(\alpha^i)\|^2 > \frac{p^2}{8 \left\lceil \frac{1}{2} (2p^r)^{\frac{1}{n+1}} \right\rceil^2 \log(4m)}$$

for any polynomial  $F(x) = \sum_{i=1}^r a_i x^{k_i}$ ,  $p \nmid a_i$ , and primitive root  $\alpha \pmod{p}$ , and positive integer  $n < \min\{m/4, \sigma\}$ . Hence from Lemma 3.1

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p - \frac{(p-1)}{4(m+1) \log(4m) \left\lceil \frac{1}{2} (2p^r)^{\frac{1}{n+1}} \right\rceil^2}.$$

When  $\sigma \leq m/4$  taking  $n + 1 = \sigma$  gives the desired result. If  $m/4 < \sigma = \sum_{j=1}^J \phi(t_j) \leq [t_1, \dots, t_J]$  then  $m = \lceil (1 + \varepsilon) \frac{8c}{\ln 2} (r \log p) (\log(r \log p))^{1+\kappa} \rceil$  and choosing  $n + 1 = \lceil m/4 \rceil$  gives

$$\left\lceil \frac{1}{2} (2p^r)^{\frac{1}{n+1}} \right\rceil \leq \left\lceil \frac{1}{2} (2p)^{4r/m} \right\rceil \leq \left\lceil \frac{1}{2} (1 + \varepsilon_1) \right\rceil = 1 \leq \left\lceil \frac{1}{2} (2p^r)^{\frac{1}{\sigma}} \right\rceil$$

for  $p \geq C(\varepsilon_1)$ , and the result is plain.  $\square$

*Proof of Corollary 1.1.* From (6.1) we have

$$\left| \sum_{x=1}^p e_p(f(x)) \right| \leq p \left( 1 - \frac{(1 - \varepsilon)}{\frac{32c}{\ln 2} (r \log p) (\log(r \log p))^{2+\kappa} \lceil \frac{1}{2} (2p)^{r/\sigma} \rceil^2} \right)$$

for  $p \geq C(\varepsilon)$ . The three bounds of parts (i), (ii) and (iii) of the corollary follow when  $(2p)^{r/\sigma} < 2p^{\frac{\delta}{2}(1-\varepsilon)}$ ,  $2(r \log p)^{\varepsilon/4}$ ,  $2$  respectively (using the rough bounds  $rp^{\delta\varepsilon}$  and  $(r \log p)^{1+\frac{1}{2}\varepsilon}$  for  $\frac{32c}{(1-\varepsilon)\ln 2} (r \log p) (\log(r \log p))^{2+\kappa}$  in (i) and (ii)). Thus it is enough to obtain  $\sigma = \sum_{j=1}^J \phi(t_j) \geq Ar$  with  $A = \frac{2}{\delta(1-\varepsilon)}$ ,  $\frac{\log p}{\log((r \log p)^{\varepsilon/4})}$ ,  $\frac{\log(2p)}{\log 2}$  respectively.

We first show that if  $p - 1 \geq Bd$  with  $B$  sufficiently large then  $\sum_{j=1}^J \phi(t_j) \geq Ar$ . Observe that, since the  $k_i \leq d \leq (p - 1)/B$ ,

$$r_j = \{k_i = \lambda_i v_j : (k_i, p - 1) = v_j\} \leq |\{\lambda_i \leq t_j/B : (\lambda_i, t_j) = 1\}|.$$

For the large  $t_j \geq B^2$ , with  $B \geq B(\varepsilon)$ , we use that

$$\sum_{n \leq x, (n, t) = 1} 1 = \sum_{d|t} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \leq x \frac{\phi(t)}{t} + 2^{\omega(t)}$$

to obtain

$$(6.2) \quad r_j \leq \frac{\phi(t_j)}{B} + 2^{\omega(t_j)} \leq (1 + \varepsilon) \frac{\phi(t_j)}{B}.$$

For the small  $t_j \leq B^2$  and  $B \geq B(\varepsilon)$  we have

$$(6.3) \quad r_j \leq \frac{t_j}{B} \leq \frac{\phi(t_j)}{B} \max_{t \leq B^2} \frac{t}{\phi(t)} \leq (1 + \varepsilon) e^\gamma \frac{\log \log B}{B} \phi(t_j),$$

and hence

$$r = \sum_{j=1}^J r_j \leq (1 + \varepsilon) e^\gamma \frac{\log \log B}{B} \sum_{j=1}^J \phi(t_j) \leq \frac{1}{A} \sum_{j=1}^J \phi(t_j),$$

as long as  $B \geq B(\varepsilon)$  and  $B \geq (1 + \varepsilon) e^\gamma A \log \log A$ .

To obtain a more explicit dependence that avoids the  $B(\varepsilon)$ , observe that for  $u = \log 5 / \log 11$  we have  $\phi(t)/2^{\omega(t)} \geq 3(t/210)^u$  and for  $v = \log(11/10) / \log(11)$  we have  $t/\phi(t) \leq \frac{35}{8}(t/210)^v$ . We assume that  $A > 1$  since trivially  $r_j \leq \phi(t_j)$ . For the large  $t_j$  with  $2^{\omega(t_j)} < C\phi(t_j)/B$ ,  $C > 1$ , we have as in (6.2) that

$$r_j \leq (1 + C) \frac{\phi(t_j)}{B} < 2C \frac{\phi(t_j)}{B},$$

while for the small  $t_j$  with  $3(t/210)^u \leq \phi(t)/2^{\omega(t)} \leq B/C$ , as in (6.3),

$$r_j \leq \frac{\phi(t_j)}{B} \max_{t \leq 210(B/3C)^{1/u}} \frac{35}{8} \left( \frac{t}{210} \right)^v \leq \frac{35}{8} \left( \frac{B}{3C} \right)^{v/u} \frac{\phi(t_j)}{B}.$$

Choosing  $C = \frac{35}{16}(16B/105)^{v/(u+v)}$  we thus have

$$r_j \leq \frac{2}{3} \left( \frac{105}{16B} \right)^{u/(u+v)} \phi(t_j) < \frac{\phi(t_j)}{A}$$

and  $\sum_{j=1}^J \phi(t_j) \geq Ar$  for  $B \geq \frac{105}{16}(2A/3)^{\log(11/2)/\log 5}$ .

The bounds involving  $k$  and  $r$  follow from making  $t_1 = (p-1)/k$  large enough that the first term of the sum gives  $\phi(t_1) \geq (1-\varepsilon)e^{-\gamma}t_1/\log \log t_1 \geq Ar$ ,  $t_1 \geq t_1(\varepsilon)$ , or more explicitly  $\phi(t_1) \geq \frac{8}{35}(210)^v t_1^{1-v} > Ar$  for  $t_1 > 210(Ar/48)^{\log 11/\log 10}$ .  $\square$

## 7. PROOF OF THEOREM 1.2

The proof uses a method that was applied to binomials by Akuliničev [1]. Let  $f = f_1 + f_2 + \dots + f_J$  as in (1.10). We proceed inductively on  $J$ . Suppose that  $v_j \nmid v_L$ . Write  $F(x) = f(x) - f_j(x) = \sum_{(k_i, p-1) \neq v_j} c_i x^{k_i}$ , and set

$$Y = \{1 \leq y < p : y^{v_j} \equiv 1 \pmod{p}\}, \quad |Y| = v_j.$$

Applying the Cauchy-Schwartz inequality gives

$$\begin{aligned} v_j \left| \sum_{x=1}^p e_p(f(x)) \right| &= \left| \sum_{y \in Y} \sum_{x=1}^p e_p(f(xy)) \right| \\ &= \left| \sum_{x=1}^p e_p(f_j(x)) \sum_{y \in Y} e_p(F(xy)) \right| \\ &\leq \left( \sum_{x=1}^p 1 \right)^{\frac{1}{2}} \left( \sum_{x=1}^p \left| \sum_{y \in Y} e_p(F(xy)) \right|^2 \right)^{\frac{1}{2}} \\ &= p^{\frac{1}{2}} \left( \sum_{y_1, y_2 \in Y} \sum_{x=1}^p e_p(F(xy_1) - F(xy_2)) \right)^{\frac{1}{2}}. \end{aligned}$$

Now if  $y_1^{v_L} \not\equiv y_2^{v_L} \pmod{p}$ , then  $F(xy_1) - F(xy_2) = \sum_{(k_i, p-1) \neq v_j} c_i (y_1^{k_i} - y_2^{k_i}) x^{k_i}$  is a polynomial containing exponents  $\vec{k}_L$ . By induction

$$\left| \sum_{x=1}^p e_p(F(xy_1) - F(xy_2)) \right| \leq p \left( 1 - \frac{1}{4^{J-2}} \delta(\vec{k}_L) \right).$$

For the remaining  $y_1^{v_L} \equiv y_2^{v_L} \pmod{p}$  we use the trivial bound  $p$  for this sum. Now

$$|\{y_1, y_2 \in Y : y_1^{v_L} \equiv y_2^{v_L} \pmod{p}\}| = v_j(v_j, v_L),$$

and thus

$$v_j \left| \sum_{x=1}^p e_p(f(x)) \right| \leq p^{\frac{1}{2}} \left( (v_j^2 - v_j(v_j, v_L)) p \left( 1 - \frac{1}{4^{J-2}} \delta(\vec{k}_L) \right) + p v_j(v_j, v_L) \right)^{\frac{1}{2}}.$$

Hence, since  $v_j \nmid v_L$ ,

$$\begin{aligned} \left| \sum_{x=1}^p e_p(f(x)) \right| &\leq p \left( 1 - \left( 1 - \frac{(v_j, v_L)}{v_j} \right) \frac{\delta(\vec{k}_L)}{4^{J-2}} \right)^{\frac{1}{2}} \\ &\leq p \left( 1 - \frac{1}{2} \frac{\delta(\vec{k}_L)}{4^{J-2}} \right)^{\frac{1}{2}} \leq p \left( 1 - \frac{\delta(\vec{k}_L)}{4^{J-1}} \right). \end{aligned}$$

### 8. PROOF OF THEOREM 2.1 AND ITS COROLLARIES

Suppose that  $1 \leq k_1 < \dots < k_r \leq p-1$  with  $(k_i, p-1) = k$  for all  $i$  and  $p-1 = kt$ .

We write

$$\psi_t(x) = \frac{x^t - 1}{\Phi_t(x)},$$

where  $\Phi_t(x)$  is the  $t$ -th cyclotomic polynomial, and

$$S = \left\{ f(x) = \sum_{i=1}^r a_i x^{k_i} \in \mathbb{Z}[x] \right\}.$$

Take  $L, M$  to be the  $s = \phi(t)$  dimensional lattices

$$L = \{ \vec{b} = (b_0, \dots, b_{t-1}) \in \mathbb{Z}^t : b_0 + b_1 x + \dots + b_{t-1} x^{t-1} = \psi_t(x) \theta(x) \text{ for some } \theta \in \mathbb{Z}[x] \},$$

$$M = \{ \vec{b} \in L : b_i \equiv f(\alpha^i) \pmod{p}, i = 0, \dots, t-1, \text{ for some } f \in S \},$$

where  $\alpha$  is a primitive root mod  $p$ . Clearly the coefficients of  $x^j \psi_t(x)$ ,  $j = 0, \dots, s-1$  form a basis for  $L$  and  $pL \subset M \subset L$ .

We assume that

$$M = \{ u_1 \vec{a}_1 + \dots + u_s \vec{a}_s : (u_1, \dots, u_s) \in \mathbb{Z}^s \},$$

and consider the convex symmetric subset of  $\mathbb{R}^s$

$$V_R = \{ (u_1, \dots, u_s) \in \mathbb{R}^s : \|u_1 \vec{a}_1 + \dots + u_s \vec{a}_s\|_\infty \leq R \}.$$

Writing  $A$  for the matrix whose  $i$ th column is  $\vec{a}_i$ , by Theorem 1 of Vaaler [23]

$$\begin{aligned} \text{Vol}(V_R) &= (2R)^s \text{Vol}(V_{\frac{1}{2}}) \\ &= (2R)^s \int_{\mathbb{R}^s} \chi_{[-\frac{1}{2}, \frac{1}{2}]^t}(A\vec{x}) d\mu_s(\vec{x}) \geq (2R)^s d(M)^{-1}, \end{aligned}$$

where  $d(M)$  is the lattice constant  $d(M) = |\det(A^T A)|^{\frac{1}{2}}$  (this is independent of the basis  $\vec{a}_1, \dots, \vec{a}_s$  of  $M$ ). Hence, by Minkowski's Theorem, if  $R = d(M)^{1/s}$  we are guaranteed a non-zero  $(u_1, \dots, u_s) \in \mathbb{Z}^s$  in  $V_R$ , giving a non-zero element of  $M$  with entries bounded by  $R$ . Hence  $\|f(\alpha^i)\| \leq R$  for  $i = 0, \dots, t-1$  for some  $f \in S$  (and, since  $f(\alpha^i) \equiv f(\alpha^{i'}) \pmod{p}$  for  $i \equiv i' \pmod{t}$ , we have  $\|f(x)\| \leq R$  for all integers  $x$ ).

Now from the canonical basis theorem for sublattices of lattices,

$$\frac{d(M)}{d(L)} = |L/M| = \frac{|L/pL|}{|M/pL|}.$$

We have  $|L/pL| = p^s$  since the coefficients of the  $x^j \psi_t(x)$  are still linearly independent mod  $p$ . By Lemma 3.5 we have  $\Phi_t(\tilde{\alpha}^{k_i}) \equiv 0 \pmod{p}$  for  $\alpha \tilde{\alpha} \equiv 1 \pmod{p}$ , giving  $\Phi_t(x) \equiv (\alpha^{k_i} x - 1) H_i(x) \pmod{p}$  for some integer polynomials  $H_i$ ,  $i = 1, \dots, r$ , and

$$1 + \alpha^{k_i} x + \dots + \alpha^{k_i(t-1)} x^{t-1} = \frac{\alpha^{k_i t} x^t - 1}{\alpha^{k_i} x - 1} \equiv \psi_t(x) H_i(x) \pmod{p}.$$

Thus, writing  $\beta_i = \alpha^{k_i}$ , we are guaranteed vectors  $\vec{b}_i$ ,  $i = 1, \dots, r$ , in  $M$  with  $\vec{b}_i \equiv (1, \beta_i, \dots, \beta_i^{t-1}) \pmod{p}$ . Plainly these  $\vec{b}_i$  generate  $M/pL$ . Moreover, since  $t - 1 > r - 1$ , they are linearly independent over  $\mathbb{Z}/p\mathbb{Z}$  (if  $u_1 \vec{b}_1 + \dots + u_r \vec{b}_r \equiv 0 \pmod{p}$  then Lemma 3.4 gives  $u_i \equiv 0 \pmod{p}$ ).

Thus  $|M/pL| = p^r$  and  $R = d(L)^{1/s} p^{1-r/s}$ . The result follows since  $d(L) = C(t)^s$  (see the proof of Theorem 6 in [19]).

Finally, for the estimate on the size of  $C(t)$  observe, using  $q$  to denote an odd prime, that for any  $X > 0$

$$\begin{aligned} 2 \log C(t) &= \sum_{q|t} \frac{\log q}{q-1} \leq \sum_{q \leq X} \frac{\log q}{q-1} + \sum_{q > X, q|t} \frac{\log q}{q-1} \\ &\leq \log X + O(1) + \frac{1}{X} \sum_{q|t} \log q \leq \log X + O(1) + \frac{\log t}{X}. \end{aligned}$$

Taking  $X = \log t$  thus gives  $\log C(t) \leq \frac{1}{2} \log \log t + O(1)$  and  $C(t) = O(\sqrt{\log t})$ .

*Proof of Corollary 2.1.* Observe that if the integer polynomial  $F$  satisfies  $\|F(x)\| \leq \Delta p$  for all  $x$  then

$$\begin{aligned} \left| \sum_{x=1}^p e_p(F(x)) \right| &\geq \left| \sum_{x=1}^p \cos \left( \frac{2\pi \|F(x)\|}{p} \right) \right| \\ &\geq \sum_{x=1}^p \left( 1 - \frac{1}{2} \left( \frac{2\pi \|F(x)\|}{p} \right)^2 \right) \\ &\geq p(1 - 2\pi^2 \Delta^2). \end{aligned}$$

From Theorem 2.1 we have polynomials  $f_j$ ,  $j = 1, \dots, J$ , with  $\|f_j(x)\| \leq C(t_j) p^{1-r_j/\phi(t_j)}$  for all  $x$ , and hence an  $f = f_1 + \dots + f_J$  with  $\|f(x)\| \leq \sum_{j=1}^J C(t_j) \max_j p^{1-r_j/\phi(t_j)}$  for all  $x$ , and thus (2.3), (2.4) are plain.

For (2.5) take  $\alpha$  to be a primitive root mod  $p$  and  $\Delta$  such that  $1/\Delta = \lfloor p^{r/T} \rfloor$ . For each  $\vec{a} = (a_1, \dots, a_r)$ ,  $0 \leq a_i < p$ , consider the values of  $\sum_{i=1}^r a_i \alpha^{k_i l} \pmod{p}$ ,  $l = 0, \dots, T-1$ . Since  $p^r \geq \lceil 1/\Delta \rceil^T$  we are, by the box principle, guaranteed two  $\vec{a}, \vec{a}'$ , whose values lie within  $p\Delta$  of each other for all  $l = 0, \dots, T-1$ . Taking  $c_i = a_i - a'_i$  thus gives a non-trivial

polynomial  $f(x) = \sum_{i=1}^r c_i x^{k_i}$  with  $\|f(\alpha^l)\| \leq p\Delta$  for  $l = 0, \dots, T$ . As  $\alpha^{k_i l} \equiv \alpha^{k_i l'} \pmod{p}$  for  $l \equiv l' \pmod{T}$  we have  $\|f(x)\| \leq p\Delta$  for all integers  $x$ .  $\square$

9. THE POLYNOMIAL WARING PROBLEM

**Lemma 9.1.** *Suppose that  $F_1(x), \dots, F_\gamma(x)$  are integer polynomials with  $\sum_{x=1}^p e_p(jF_i(x)) \leq p(1 - \delta_i)$  for any  $(j, p) = 1$ . If  $\delta_1 + \dots + \delta_\gamma \geq \log p$  then*

$$F_1(x_1) + \dots + F_\gamma(x_\gamma) \equiv N \pmod{p},$$

has an integer solution for all  $N$ .

*Proof.* Since  $\sum_{j=1}^p e_p(ju) = p$  if  $u \equiv 0 \pmod{p}$  and 0 otherwise, the number of solutions to  $F_1(x_1) + \dots + F_\gamma(x_\gamma) \equiv N \pmod{p}$  in  $S = \{\vec{x} = (x_1, \dots, x_\gamma) : 1 \leq x_i \leq p\}$  is just

$$\begin{aligned} \frac{1}{p} \sum_{j=1}^p \sum_{\vec{x} \in S} e_p(j(F_1(x_1) + \dots + F_\gamma(x_\gamma) - N)) &= p^{\gamma-1} + \frac{1}{p} \sum_{j=1}^{p-1} e_p(-jN) \prod_{i=1}^{\gamma} \sum_{x_i=1}^p e_p(jF_i(x_i)) \\ &\geq p^{\gamma-1} - p^\gamma \prod_{i=1}^{\gamma} (1 - \delta_i) \\ &> p^{\gamma-1} (1 - p \exp(-(\delta_1 + \dots + \delta_\gamma))) \geq 0, \end{aligned}$$

if  $\delta_1 + \dots + \delta_\gamma \geq \log p$ .  $\square$

The above lemma immediately gives the bound  $\gamma(f, p) \leq \lceil \log p / \delta(\vec{k}) \rceil$  mentioned in the introduction. Hence from Corollary 1.1 (i), we have  $\gamma(f, p) \leq rp^\epsilon$  once  $p \geq C(\epsilon)d$ , and  $\gamma(f, p) \leq (r \log^2 p)^{1+\epsilon}$  once  $p \geq C(\epsilon)d \log d / (\log \log d)^{1-\epsilon}$ . Once  $p > 2d^2$  using the Weil bound in the proof of Lemma 9.1 yields a better bound, as indicated in the introduction.

REFERENCES

- [1] N. M. Akulinichev, *Estimates for rational trigonometric sums of a special type*, Doklady Acad. Sci. USSR 161 (1965), 743-745. English trans. in Doklady 161, no. 4 (1965), 480-482.
- [2] J. D. Bovey, *A new upper bound for Waring's problem (mod p)*, Acta Arith. 32 (1977), 157-162.
- [3] L. Carlitz, D. J. Lewis, W. H. Mills, E. G. Straus, *Polynomials over finite fields with minimal value sets*, Mathematika 8 (1961), 121-130.
- [4] I. Chowla, *On Waring's problem (mod p)*, Proc. Ind. Nat. Acad. Sci. India Sect. A 13 (1943), 195-200.
- [5] S. Chowla, H. B. Mann and E. G. Straus, *Some applications of the Cauchy-Davenport theorem*, Norske Vid. Selsk. Forh. Trondheim 32 (1959), 74-80.
- [6] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*, Acta Arith. 34 (1979), 391-401.
- [7] M. M. Dodson, *On Waring's problem in  $GF[p]$* , Acta Arith. 19 (1971), 147-173.

- [8] M. M. Dodson and A. Tietäväinen, *A note on Waring's problem in  $GF[p]$* , Acta Arith. 30 (1976), 159-167.
- [9] A. Garcia and J. F. Voloch, *Fermat curves over finite fields*, J. Number Theory 30 (1988), 345-356.
- [10] G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum". VIII: The number  $\Gamma(k)$  in Waring's problem*, Proc. London Math. Soc. (2) 28 (1927), 518-542.
- [11] D. R. Heath-Brown and S. Konyagin, *New bounds for Gauss sums derived from  $k$ th powers, and for Heilbronn's exponential sum*, Quart. J. Math. 51 (2000), 221-235.
- [12] A. A. Karatsuba, *On estimates of complete trigonometric sums*, Matem. Zametki. 1 (1967), 199-208 (Russian); translation in Math. Notes (1968), 133-139.
- [13] B. S. Kashin and A. A. Saakyan, *Orthogonal Series*, "Nauka" Moscow, 1984; English transl., Amer. Math. Soc, Providence, RI, 1989.
- [14] S. V. Konyagin, *On estimates of Gaussian sums and Waring's problem for a prime modulus*, Trudy Mat. Inst. Steklov 198 (1992), 111-124; translation in Proc. Steklov Inst. Math. 1994, 105-107.
- [15] ——— *Exponential sums over multiplicative groups of residues*, preprint, (2000), Moscow State University.
- [16] S. V. Konyagin and I. E. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999.
- [17] L. J. Mordell, *On a sum analogous to a Gauss's sum*, Q.J. Math., 3 (1932), 161-167.
- [18] G. L. Mullen and I. E. Shparlinski, *Open problems and conjectures in finite fields*, in Finite Fields and Applications (Glasgow, 1995), London Math. Soc. Lecture Note Series, No. 233, S. Cohen and H. Niederreiter, eds., Cambridge University Press, Cambridge, 1996, 243-268.
- [19] C. Powell, *Bounds for multiplicative cosets over fields of prime order*, Mathematics of Computation 218 (1997), 807-822.
- [20] I. E. Shparlinski, *On bounds of Gaussian sums*, Matem. Zametki, 50 (1991), 122-130 (in Russian).
- [21] ——— *On Gaussian sums for finite fields and elliptic curves*, Proc. 1-st French-Soviet Workshop on Algebraic Coding, Paris, 1991, Lect. Notes in Computer Sci., 537 (1992), 5-15.
- [22] A. Tietäväinen, *Proof of a conjecture of S. Chowla*, J. Number Theory 7 (1975), 353-356.
- [23] J. D. Vaaler, *A geometric inequality with applications to linear forms*, Pacific J. of Mathematics 83 (1979), 543-553.
- [24] L. C. Washington, *Introduction to Cyclotomic Fields*, Springer-Verlag Graduate Texts in Math. 83, New York, 1982.
- [25] A. Weil, *On some exponential sums*, Proc. Nat. Acad. Sci. U.S.A. 34 (1948), 204-207.

DEPARTMENT OF MATHEMATICS, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506  
*E-mail address:* `cochrane@math.ksu.edu`

DEPARTMENT OF MATHEMATICS, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506  
*E-mail address:* `pinner@math.ksu.edu`

DEPARTMENT OF MATHEMATICS, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506  
*E-mail address:* `jasonr@math.ksu.edu`