JOURNAL de Théorie des Nombres de BORDEAUX

anciennement Séminaire de Théorie des Nombres de Bordeaux

Joseph H. SILVERMAN

Lehmer's conjecture for polynomials satisfying a congruence divisibility condition and an analogue for elliptic curves

Tome 24, n° 3 (2012), p. 751-772.

<http://jtnb.cedram.org/item?id=JTNB_2012__24_3_751_0>

Journal de Théorie des Nombres
de Bordeaux **24** (2012), 751-772

Lehmer's conjecture for polynomials satisfying a congruence divisibility condition and an analogue for elliptic curves

par Joseph H. SILVERMAN

Résumé. De nombreux auteurs ont prouvé des versions explicites de la conjecture de Lehmer dans le cas particulier de polynômes dont les coefficients sont tous congrus à 1 modulo un entier $m > 1$. Nous prouvons ici un résultat similaire pour les polynômes qui sont divisibles dans l'anneau $(\mathbb{Z}/m\mathbb{Z})[X]$ par un polynôme de la forme $1 + X + \cdots + X^n$ pour un certain $n \geq \epsilon \deg(f)$. Nous prouvons également un énoncé analogue pour les courbes elliptiques.

Abstract. A number of authors have proven explicit versions of Lehmer's conjecture for polynomials whose coefficients are all congruent to 1 modulo m. We prove a similar result for polynomials $f(X)$ that are divisible in $(\mathbb{Z}/m\mathbb{Z})[X]$ by a polynomial of the form $1 + X + \cdots + X^n$ for some $n \geq \epsilon \deg(f)$. We also formulate and prove an analogous statement for elliptic curves.

Introduction

Let

$$h : \bar{\mathbb{Q}} \longrightarrow [0, \infty)$$

denote the absolute logarithmic height [9, 11]. Lehmer's conjecture [15] asserts that there is an absolute constant $C > 0$ such that if $f(X) \in \mathbb{Z}[X]$ is a monic polynomial of degree $D \geq 1$ whose roots are not roots of unity, then

$$\sum_{f(\alpha)=0} h(\alpha) \geq C. \tag{0.1}$$

This problem has a long history; see for example [2, 15, 21, 23, 24]. The best general result known, which is due to Dobrowolski [5], says that $\sum h(\alpha) \geq C(\log\log D/\log D)^3$. Various authors have considered Lehmer's problem for restricted values of α. For example, Amoroso and Dvornicich [1] show

Manuscrit reçu le 14 décembre 2011.
The author's research partially supported by NSF grants DMS-0650017 and DMS-0854755.
Mots clefs. Lehmer conjecture, elliptic curve, canonical height.
Classification math. 11G05, 11G50, 11J97, 14H52.

that if the roots of $f(X)$ generate an abelian extension of $\mathbb{Q}$, then $\sum h(\alpha) \ge D(\log 5)/12$.

An interesting class of polynomials are those whose coefficients are all odd. More generally, one can consider polynomials whose coefficients are congruent to 1 modulo m, as in the following result.

Theorem 0.1. (Borwein, Dobrowolski, Mossinghoff [3]) *Let $m \ge 2$, and let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree D with no cyclotomic factors that satisfies*

$$f(X) \equiv X^D + X^{D-1} + \cdots + X^2 + X + 1 \pmod{m}. \tag{0.2}$$

Then

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{D}{D+1} C_m,$$

where we may take

$$C_2 = \frac{1}{4}\log 5 \quad \textit{and} \quad C_m = \log \frac{\sqrt{m^2+1}}{2} \quad \textit{for } m \ge 3.$$

We mention that an earlier paper [4] does the case of non-reciprocal polynomials, and subsequent papers [6, 10] give improved values for C_m, although asymptotically they all have the form $C_m = \log(m/2) + O(1/m^2)$. We also note the papers [17, 18] which give various generalizations of Theorem 0.1, including weakening the congruence condition (0.2), working over number fields, and considering heights of points and subspaces in projective space.

Our first result is the following generalization of Theorem 0.1, albeit with less sharp constants. See Theorem 2.1 and Corollary 2.1 for our precise results.

Theorem 0.2. *For all $\epsilon > 0$ there is a constant $C_\epsilon > 0$ with the following property: Let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree D such that*

$$f(X) \text{ is divisible by } X^{n-1} + X^{n-2} + \cdots + X + 1 \text{ in } (\mathbb{Z}/m\mathbb{Z})[X] \tag{0.3}$$

for some integers

$$m \ge 2 \qquad \textit{and} \qquad n \ge \max\{\epsilon D, 2\}.$$

Suppose further that no root of $f(X)$ is a root of unity. Then

$$\sum_{f(\alpha)=0} h(\alpha) \ge C_\epsilon \log m.$$

In particular, Lehmer's conjecture (0.1) *is true for this class of polynomials.*

The elliptic analogue of Lehmer's conjecture says that if E/K is an elliptic curve defined over a number field, then there is a constant $C_{E/K} > 0$

such that for all nontorsion points $Q \in E(\bar{K})$ of degree $D_Q = [K(Q) : K]$ we have

$$D_Q \hat{h}_E(Q) \geq C_{E/K}. \tag{0.4}$$

Here $\hat{h}_E$ is the logarithmic canonical height on E. There has been considerable work on the elliptic Lehmer conjecture; see for example [8, 14, 16]. Our second main result is an elliptic analogue of Theorem 0.2. We show that if a significant number of the Galois conjugates of Q are $\mathfrak{m}$-adically close to n-torsion points, then the elliptic Lehmer estimate (0.4) holds. We now make a few comments concerning our elliptic result, but see Corollary 4.1 for the precise statement.

An initial difficulty is to find an appropriate elliptic version of the mod m divisibility condition (0.3). In Section 1 we show that (0.3) implies a lower bound for a certain sum over the roots of f, and it is this weaker property that we generalize and adapt to the elliptic setting. More precisely, the divisibility property (0.3) says that a significant number of the roots of f are m-adically close to n^{th}-roots of unity. The analogous statement for elliptic curves, as noted earlier, is that a significant number of the Galois conjugates of Q are $\mathfrak{m}$-adically close to n-torsion points.

Theorem 0.2 deals with congruences related to cyclotomic polynomials, which is natural when studying Lehmer's problem, but one might consider other sorts of congruence conditions. For example, suppose that $f(X)$ is congruent modulo m to $X^D + X^{D-1} + \cdots + X^2 + X - 1$. Samuels [17] has considered general conditions of this sort. In Section 5 we briefly reprove one of Samuels' results and use it to make a number of remarks concerning possible generalizations.

Remark 0.1. Blanksby and Montgomery [2] used Fourier methods to prove an approximation to Lehmer's conjecture. Their proof involves two steps:

Step I: (Fourier averaging) Look at weighted sums $\displaystyle\sum_{j=1}^{J} a_j \log|1 - \alpha^j|$ and prove an upper bound of the form $O(\log J)$, which improves the trivial bound of $O(J)$.

Step II: Compute a second moment in order to obtain a lower bound for the Lehmer sum.

The Fourier averaging method (Step I) of Blanksby and Montgomery was extended to elliptic curves by Hindry and the author [7, 8], but the argument for the Step II lower bound was replaced by a pigeonhole argument in [7] and by the use of Néron models in [8].

In this paper we use Fourier averaging as in [2, 7, 8] to obtain cancelation in certain weighted sums. But the various arguments used for Step II are replaced in this paper with a lower bound obtained from the congruence

divisibility condition (0.3). In particular, this explains why our final lower bound depends on the strength of the congruence condition imposed on the polynomial.

As an aid to the reader, and to set up the estimates in the exact form that we need, we give the Fourier averaging arguments in full for the multiplicative group. However, we note that our Proposition 2.1 is similar to [2, Theorem 1], and our Lemma 6.1 is similar to, and plays the same role as, [2, Lemma 4]. We also refer the reader to [22], which gives a simplified proof of the main result of [2].

Remark 0.2. Our Theorem 0.2 and some of Samuels' principal results [17] are height bounds for polynomials satisfying various sorts of congruence conditions, so we conclude this introduction by briefly describing how the results differ. Our polynomials satisfy a divisibility condition modulo m, so multiplying by $X-1$, our theorem applies to polynomials $F(X)$ of the form

$$F(X) = (X^n - 1)A(X) + mB(X) \qquad \text{for some } A, B \in \mathbb{Z}[X].$$

In general, we obtain a bound for all n (see Lemma 2.1), and in particular we prove Lehmer's conjecture if $n \ge \epsilon \deg(F)$. The results in [17] apply to (factors of) polynomials that are congruent modulo m to a simpler polynomial of the same degree. For example, a typical result in [17] is a bound for (noncyclotomic factors of) polynomials $F(X)$ of degree nr satisfying

$$F(X) = (X^n - 1)^r + mB(X) \qquad \text{for some } B \in \mathbb{Z}[X].$$

Thus although there is some overlap, our result and the results in [17] apply to largely different classes of polynomials. It might be interesting to combine the methods of the two papers to prove a general result encompassing both.

Acknowledgements. The author would like to thank Michael Mossinghoff for introducing him to the topic of polynomials whose coefficients satisfy congruence conditions, and to thank Michael, Igor Shparlinski, and the referees for for their helpful comments on the initial draft of this paper and for their assistance in the phrasing of the résumé.

1. A reformulation of property (0.2)

We start by normalizing our absolute values.

Definition. We let $M_{\mathbb{Q}}$ be the usual collection of absolute values on $\mathbb{Q}$, and for any algebraic extension $K/\mathbb{Q}$, we write M_K for the set of all extensions of these absolute values to K. For $\alpha \in \bar{\mathbb{Q}}$ and $v \in M_{\bar{\mathbb{Q}}}$, we define a normalized absolute value by choosing a finite extension $K/\mathbb{Q}$ with $\alpha \in K$ and setting

$$\|\alpha\|_v = |\alpha|_v^{[K_v:\mathbb{Q}_v]/[K:\mathbb{Q}]},$$

where $|\cdot|_v$ is the v-absolute value on K extending an absolute value on $\mathbb{Q}$. We also define a normalized valuation by

$$v(\alpha) = -\log \|\alpha\|_v.$$

Then the absolute logarithmic height of α is defined by

$$h(\alpha) = \sum_{v \in M_K} \max\{\log \|\alpha\|_v, 0\}.$$

We write M_K^0, respectively M_K^∞, for the set of nonarchimedean, respectively archimedean, absolute values in M_K.

Remark 1.1. With the above normalization we have the product formula

$$\prod_{v \in M_K} \|\alpha\|_v = 1 \qquad \text{for all } \alpha \in K^*.$$

In particular, if $\alpha \in K$ is a nonzero algebraic integer, then

$$h(\alpha) = \sum_{v \in M_K^\infty} \max\{\log \|\alpha\|_v, 0\} = \sum_{v \in M_K^0} v(\alpha).$$

We also remark that

$$\prod_{v \in M_K^\infty} \|\alpha\|_v = \prod_{v \in M_K^0} \|\alpha\|_v^{-1} = \mathsf{N}_{K/\mathbb{Q}}(\alpha)^{1/[K:\mathbb{Q}]}.$$

Remark 1.2. Let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial. Then the classical Mahler measure $M(f)$ of f is related to the heights of its roots via the formula

$$\log M(f) = \sum_{f(\alpha)=0} h(\alpha).$$

In this paper we use the "sum the heights" notation because it has an obvious generalization to other algebraic groups such as elliptic curves. In such sums, we always include the roots of f with their multiplicities.

Definition. For notational convenience, we let

$$\Phi_n(X) = X^n + X^{n-1} + \cdots + X + 1.$$

If n is prime, this is the usual cyclotomic polynomial; in general it is a product of classical cyclotomic polynomials.

Property (0.2) of Theorem 0.1 says that all of the coefficients of the polynomial f are congruent to 1 modulo m. This is equivalent to saying that the monic degree D polynomial $f(X) \in \mathbb{Z}[X]$ is equal to $\Phi_D(X)$ in the ring $(\mathbb{Z}/m\mathbb{Z})[X]$. We are going to replace this equality by a divisibility condition. Note that although the ring $\mathbb{Z}/m\mathbb{Z}$ contains zero divisors if m is composite, divisibility by monic polynomials in $(\mathbb{Z}/m\mathbb{Z})[X]$ is well-behaved. The next proposition gives some properties that are weaker than Property (0.2) of Theorem 0.1.

Proposition 1.1. *Let $m, n \ge 2$, and let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree D. Consider the following three conditions*:

(i) *$f(X)$ is divisible by $\Phi_{n-1}(X)$ in $(\mathbb{Z}/m\mathbb{Z})[X]$.*

(ii) *$m^{n-1} \mid \operatorname{Res}\bigl(f(X), \Phi_{n-1}(X)\bigr)$.*

(iii) $\displaystyle\sum_{v\mid m} \sum_{f(\alpha)=0} v(\alpha^n - 1) \ge (n-1)\log m.$

Then

$$\text{(i)} \Longrightarrow \text{(ii)} \Longrightarrow \text{(iii)}.$$

(*In* (iii)*, we may work over any field in which f factors completely. The way that we have normalized our absolute values ensures that the choice of field does not matter.*)

Proof. Property (i) says that

$$f(X) = \Phi_{n-1}(X)A(X) + mB(X) \quad \text{for some } A(X), B(X) \in \mathbb{Z}[X].$$

This implies that

$$\begin{aligned}\operatorname{Res}\bigl(f(X), \Phi_{n-1}(X)\bigr) &= \operatorname{Res}\bigl(\Phi_{n-1}(X)A(X) + mB(X), \Phi_{n-1}(X)\bigr)\\ &= \operatorname{Res}\bigl(mB(X), \Phi_{n-1}(X)\bigr)\\ &= m^{n-1}\operatorname{Res}\bigl(B(X), \Phi_{n-1}(X)\bigr).\end{aligned}$$

Thus (ii) is true.

We next prove that (ii) implies (iii). For any non-archimedean absolute value v we have

$$\begin{aligned}\bigl\|\operatorname{Res}\bigl(f(X), X^n - 1\bigr)\bigr\|_v &= \bigl\|\operatorname{Res}\bigl(f(X), \Phi_{n-1}(X)\bigr)\bigr\|_v \bigl\|f(1)\bigr\|_v\\ &\le \|m\|_v^{n-1}. \qquad (1.1)\end{aligned}$$

A standard formula for the resultant [13, Section V.10] is

$$\operatorname{Res}\bigl(f(X), X^n - 1\bigr) = \prod_{f(\alpha)=0} (\alpha^n - 1). \qquad (1.2)$$

We take the v-absolute value of (1.2), use (1.1), and multiply over all $v \mid m$ to obtain the estimate

$$\prod_{v\mid m} \prod_{f(\alpha)=0} \|\alpha^n - 1\|_v \le \prod_{v\mid m} \|m\|_v^{n-1} = m^{-(n-1)}. \qquad (1.3)$$

Taking $-\log(\,\cdot\,)$ gives (iii). □

We now define a quantity that generalizes the sum appearing in Property (iii) of Proposition 1.1.

Definition. Let $\mathcal{A} \subset \bar{\mathbb{Q}}^*$ be a finite set of algebraic integers and let m and n be positive integers. We define

$$\Delta(\mathcal{A}, m, n) = \sum_{\alpha \in \mathcal{A}} \frac{1}{\log m} \sum_{v \mid m} \frac{1}{n} v(\alpha^n - 1).$$

The inner sum is over all $v \in M_K^0$ with $v \mid m$, where K is any number field containing $\mathcal{A}$. Our normalization of the valuations in $M_{\bar{\mathbb{Q}}}$ implies that the sum is independent of the choice of K.

Proposition 1.2. *Let $\mathcal{A}$ be a finite set of algebraic integers, let $j, m, n \geq 1$ be rational integers, and let $\mathcal{A}^j = \{\alpha^j : \alpha \in \mathcal{A}\}$. Then*

$$\Delta(\mathcal{A}^j, m, n) \geq \Delta(\mathcal{A}, m, n).$$

Proof. From the factorization

$$X^{jn} - 1 = (X^n - 1)\Phi_{j-1}(X^n),$$

we see that

$$v(\alpha^{jn} - 1) = v(\alpha^n - 1) + v(\Phi_{j-1}(\alpha^n)) \geq v(\alpha^n - 1)$$

for any algebraic integer α and any nonarchimedean absolute value v. Summing over $\alpha \in \mathcal{A}$ and $v \mid m$, and then dividing by $n \log m$, gives the desired result. □

Remark 1.3. We observe that if $f(X) \in \mathbb{Z}[X]$ is a monic polynomial of degree D and if we write $\mathcal{A}_f$ for the set of roots of $f(X)$, then Property (iii) of Proposition 1.1 can be succintly written as

$$\Delta(\mathcal{A}_f, m, n) \geq \frac{n-1}{n}. \tag{1.4}$$

In particular, if f satisfies the congruence

$$f(X) \equiv \Phi_D(X) \pmod{m}$$

as in the statement of Theorem 0.1, then

$$\Delta(\mathcal{A}_f, m, D+1) \geq \frac{D}{D+1}.$$

2. A height bound for polynomials satisfying congruence conditions

The next theorem is our main result for number fields. As we will see, it generalizes [3, 4, 6, 10] (Theorem 0.1), albeit with worse constants. Later we will prove an elliptic curve version of this theorem and its consequences.

Theorem 2.1. *Let $\mathcal{A} \subset \bar{\mathbb{Q}}$ be a finite set of algebraic integers that does not contain any roots of unity, and let m and n be positive integers. Then for all integers $J \ge 1$ we have*

$$\sum_{\alpha\in\mathcal{A}} h(\alpha) \ge \frac{3}{J+2}\left(\Delta(\mathcal{A},m,n)\log m - \frac{|\mathcal{A}|}{n}\frac{\log(J/2+1)+1}{J}\right). \tag{2.1}$$

Remark 2.1. It is always possible to choose a value of J to obtain a nontrivial, i.e., positive, lower bound in (2.1). The optimal choice of J depends on the relative sizes of $\Delta(\mathcal{A},m,n)\log m$ and $|\mathcal{A}|/n$. In the application most closely related to Theorem 0.1, we have

$$n = D+1 \qquad \text{and} \qquad \Delta(\mathcal{A},m,n)\log m = \frac{|\mathcal{A}|}{n} \ge \frac{D}{D+1},$$

so we get

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{D}{D+1}\cdot\frac{3}{J+2}\cdot\left(\log m - \frac{\log(J/2+1)+1}{J}\right).$$

If $m \ge 5$, then we obtain a nontrivial lower bound with $J = 1$, while for $3 \le m \le 4$ we need to take $J = 2$, and for $m = 2$ we must take $J = 3$. Of course, the bound that we obtain is not sharp. However, our goal is not to get sharp bounds in this particular case, where other authors [3, 4, 6, 10] have used intricate techniques to obtain better bounds than we could obtain even if we took more care. Instead, we aim to show how to obtain nontrivial bounds that, among other things, imply that Lehmer's conjecture is true for an interesting class of polynomials that is larger than the class considered in [3, 4, 6, 10].

The proof of Theorem 2.1 uses the following standard Fejér kernel estimate, cf. [2, Theorem 1], whose proof we relegate to Section 6.

Proposition 2.1. *For all $J \ge 1$ we have*

$$\sup_{\substack{z\in\mathbb{C}\\ |z|\le 1}} \sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\log|1-z^j| \le \frac{1}{2}\log\left(\frac{J}{2}+1\right)+\frac{1}{2}.$$

Proof of Theorem 2.1. Let K be a number field such that $\mathcal{A} \subset K$. For $\alpha \in \mathcal{A}$ and $v \in M_K$, we let

$$\alpha_v = \begin{cases} \alpha & \text{if } \|\alpha\|_v \le 1,\\ \alpha^{-1} & \text{if } \|\alpha\|_v > 1, \end{cases}$$

so in particular $\|\alpha_v\|_v \le 1$. We now compute

$$\begin{aligned}
(n\log m)\Delta(\mathcal{A},m,n) &= \sum_{\alpha\in\mathcal{A}}\sum_{v\mid m} v(\alpha^n-1) && \text{def. of } \Delta(\mathcal{A},m,n),\\
&\le \sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^0} v(\alpha^n-1) && \alpha \text{ is an algebraic integer},\\
&= \sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^\infty} \log\|\alpha^n-1\|_v && \text{product rule},\\
&\le \sum_{\alpha\in\mathcal{A}}\sum_{\substack{v\in M_K^\infty\\ \|\alpha\|_v>1}} \log\|\alpha^n\|_v + \sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^\infty}\log\|\alpha_v^n-1\|_v\\
&= n\sum_{\alpha\in\mathcal{A}} h(\alpha) + \sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^\infty}\log\|\alpha_v^n-1\|_v. && (2.2)
\end{aligned}$$

We now replace $\mathcal{A}$ with $\mathcal{A}^j$. Then using $h(\alpha^j)=jh(\alpha)$ and Proposition 1.2, which says that $\Delta(\mathcal{A},m,n)\le\Delta(\mathcal{A}^j,m,n)$, we find that

$$(n\log m)\Delta(\mathcal{A},m,n) \le nj\sum_{\alpha\in\mathcal{A}} h(\alpha) + \sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^\infty}\log\|\alpha_v^{jn}-1\|_v.$$

We multiply by the Fejér multiplier $1-\frac{j}{J+1}$ and sum over $1\le j\le J$ to obtain

$$\begin{aligned}
\frac{Jn\log m}{2}\Delta(\mathcal{A},m,n) \le{}& \frac{(J^2+2J)n}{6}\sum_{\alpha\in\mathcal{A}} h(\alpha)\\
&+\sum_{\alpha\in\mathcal{A}}\sum_{v\in M_K^\infty}\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\log\|\alpha_v^{jn}-1\|_v.
\end{aligned}$$

Note that the sum over v is over archimedean absolute values, so if we assume that α_v is chosen in the unit disk to maximize the innermost sum over j, we get the estimate

$$\begin{aligned}
&\frac{Jn\log m}{2}\Delta(\mathcal{A},m,n)\\
&\qquad\le \frac{(J^2+2J)n}{6}\sum_{\alpha\in\mathcal{A}} h(\alpha) + |\mathcal{A}|\sup_{\substack{z\in\mathbb{C}\\ |z|\le 1}}\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\log|z^j-1|.
\end{aligned}$$

We can now use Proposition 2.1 to conclude that

$$\frac{Jn\log m}{2}\Delta(\mathcal{A},m,n) \le \frac{(J^2+2J)n}{6}\sum_{\alpha\in\mathcal{A}} h(\alpha) + \frac{|\mathcal{A}|}{2}\left(\log\left(\frac{J}{2}+1\right)+1\right).$$

After a little bit of algebra, we obtain the desired result. □

We now use our main theorem to prove that Lehmer's conjecture is true for a certain interesting collection of polynomials.

Corollary 2.1. *Fix $0 < \epsilon < \frac{1}{210}$. Then Lehmer's conjecture* (0.1) *is true for the set of polynomials*

$$\left\{ f(X) \in \mathbb{Z}[X] : \begin{array}{l} f(X) \text{ is monic, its roots are not roots} \\ \text{of unity, and there exist integers } m \ge 2 \\ \text{and } n \ge \max\{2, \epsilon \deg(f)\} \text{ such that} \\ \Phi_{n-1}(X) \text{ divides } f(X) \text{ in } (\mathbb{Z}/m\mathbb{Z})[X] \end{array} \right\}. \tag{2.3}$$

More precisely, if $f(X)$ is in the set (2.3)*, then*

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{\log m}{185\epsilon^{-1}\log(24\epsilon^{-1})}. \tag{2.4}$$

Remark 2.2. Igor Shparlinski has pointed out that for large m, we may take $\epsilon = (\log\log m)/(\log m)$ and conclude that if $\Phi_{n-1}(X)$ divides $f(X)$ in $(\mathbb{Z}/m\mathbb{Z})[X]$ for some $n \ge ((\log\log m)/(\log m))(\deg f)$, then

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{1}{185} + O\left(\frac{\log\log\log m}{\log\log m}\right),$$

where the big-O constant is absolute.

The proof of the corollary uses a combination of Theorem 2.1 and Proposition 1.1 as reformulated in Remark 1.3. We state the exact result that we require as a lemma.

Lemma 2.1. *Let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree D whose roots are not roots of unity, let $m, n \ge 2$ be integers, and suppose that $\Phi_{n-1}(X)$ divides $f(X)$ in $(\mathbb{Z}/m\mathbb{Z})[X]$. Then*

$$\sum_{f(\alpha)=0} h(\alpha) \ge \begin{cases} (\log m)/123341 & \text{if } \log m \ge D/16n, \\ \dfrac{\log m}{(128D/n\log m)\log(16D/n\log m)} & \text{if } \log m \le D/16n. \end{cases}$$

Proof. Let $\mathcal{A}_f$ be the set of roots of f. As noted in (1.4) of Remark 1.3, the divisibility condition on f implies that $\Delta(\mathcal{A}_f, m, n) \ge (n-1)/n$. Substituting this into (2.1) of Theorem 2.1 and using $|\mathcal{A}_f| = D$, we find that for all integers $J \ge 1$ we have

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{3}{J+2}\left(\frac{n-1}{n}\log m - \frac{D}{n}\frac{\log(J/2+1)+1}{J}\right).$$

Since we are not concerned with optimizing our constants, we observe that for $n \ge 2$ and $J \ge 2$, this implies that

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{1}{J}\left(\frac{1}{2}\log m - \frac{4D}{n}\frac{\log(J)}{J}\right). \tag{2.5}$$

We now consider two cases. First, if m is large, say

$$\log m \geq D/16n,$$

then taking $J = 1789$ gives

$$\sum_{f(\alpha)=0} h(\alpha) \geq \frac{1}{J}\left(\frac{1}{32} - 4\frac{\log(J)}{J}\right)\log m \geq \frac{\log m}{123341}. \tag{2.6}$$

Second, suppose that m is small,

$$\log m \leq D/16n.$$

Then we want to choose J to be an integer satisfying

$$\frac{J}{\log J} \geq \frac{16D}{n\log m}. \tag{2.7}$$

In particular, since $16D/n\log m \geq 256$, it suffices to take

$$J = \left\lfloor \frac{32D}{n\log m}\log\left(\frac{16D}{n\log m}\right)\right\rfloor - 1.$$

Substituting (2.7) into (2.5) and adjusting the constants yields

$$\sum_{f(\alpha)=0} h(\alpha) \geq \frac{\log m}{4J} \geq \frac{\log m}{(128D/n\log m)\log(16D/n\log m)}. \tag{2.8}$$

Combining (2.6) and (2.8) completes the proof of Lemma 2.1. □

Proof of Corollary 2.1. We are given that $n \geq \max\{\epsilon D, 2\}$. If $\log m \geq D/16n$, then Lemma 2.1 says that

$$\sum_{f(\alpha)=0} h(\alpha) \geq \frac{\log m}{123341}.$$

This is stronger than (2.4), since we have assumed that $\epsilon < \frac{1}{210}$, so we are reduced to the case that $\log m \leq D/16n$. Since $n \geq \epsilon D$, this implies that

$$\frac{D}{n\log m} \leq \frac{D}{\epsilon D\log m} \leq \frac{1}{\epsilon\log m},$$

where the upper bound is at least 16. Substituting this into Lemma 2.1, we find that

$$\sum_{f(\alpha)=0} h(\alpha) \geq \frac{\log m}{(128/\epsilon\log m)\log(16/\epsilon\log m)}.$$

Since $m \geq 2$, this gives something slightly stronger than the desired result. □

3. An elliptic analogue of $\Delta(\mathcal{A}, m, n)$

An amalgamation of Proposition 1.1, Remark 1.3, and Theorem 2.1 says that if $f(X) \in \mathbb{Z}[X]$ is a monic polynomial of degree D whose roots are not roots of unity, then

$$\begin{pmatrix} f(X) \text{ has coefficients} \\ \text{congruent to 1 modulo } m \end{pmatrix} \implies \Delta(\mathcal{A}_f, m, D+1) \ge \frac{D}{D+1}$$
$$\implies \sum_{f(\alpha)=0} h(\alpha) \ge \frac{D}{D+1} C_m.$$

The key estimate is Theorem 2.1, which gives a general lower bound for $\sum h(\alpha)$ in terms of $\Delta(\mathcal{A}, m, n)$. In this section we define an elliptic analogue of the quantity $\Delta(\mathcal{A}, m, n)$, and in the next section we prove an elliptic analogue of Theorem 2.1. We begin by recalling some basic properties of canonical height functions on elliptic curves.

Definition. Let E/K be an elliptic curve defined over a number field. We write

$$\hat{h} : E(\bar{K}) \longrightarrow \mathbb{R}$$

for the absolute logarithmic canonical height [20, VIII §9], and for each $v \in M_{\bar{K}}$ we let

$$\hat{\lambda}_v : E(\bar{K}_v) \smallsetminus \{O\} \longrightarrow \mathbb{R}$$

be a local canonical height, normalized as described in [19, Chapter VI].

Proposition 3.1. *The local canonical height satisfies the following*:

(a) *For all $v \in M_K$ there is a constant $c(v)$ such that*

$$\hat{\lambda}_v(P) \ge -c(v) \quad \text{for all } P \in E(\bar{K}_v).$$

Further, if $v \in M_K^0$ and E has good reduction at v, then we can take $c(v) = 0$.

(b) *The global height is the sum of the local heights. Thus for any finite extension L/K and any $P \in E(L) \smallsetminus \{O\}$ we have*

$$\hat{h}(P) = \sum_{v \in M_L} \hat{\lambda}_v(P).$$

Proof. The first part of (a) follows from [19, Theorem VI.1.1(a)], which says in particular that $\hat{\lambda}_v$ has a logarithmic pole as $P \to O$ in the v-adic topology and that λ_v is bounded on the complement of any v-adic neighborhood of O. The second part of (a) follow from [19, Theorem VI.4.1], which says that if P reduces to a non-singular point modulo v, then

$$\hat{\lambda}_v(P) = \frac{1}{2} \max\{-v(x(P)), 0\} + \frac{1}{12} v(\mathfrak{D}_{E/K}).$$

This quantity is clearly non-negative. Finally, [19, Theorem VI.2.1] gives a proof of (b). □

Definition. Given

$K/\mathbb{Q}$ a number field,
$\mathfrak{m}$ an integral ideal of K with norm $m = \mathsf{N}_{K/\mathbb{Q}}\mathfrak{m} \geq 2$,
E/K an elliptic curve, and
$\mathcal{P}$ a finite set of nontorsion points in $E(\bar{K})$,

we define

$$\Delta(\mathcal{P}, \mathfrak{m}, n) = \sum_{P \in \mathcal{P}} \frac{1}{\log m} \sum_{v \mid \mathfrak{m}} \frac{1}{n^2} \hat{\lambda}_v(nP).$$

This quantity is the elliptic analogue of the quantity $\Delta(\mathcal{A}, m, n)$ that we defined in Section 1.

The following estimate will be used later when we do an averaging argument. It is the analogue of Proposition 1.2.

Lemma 3.1. *With notation as above, assume that E has potential good reduction at every prime dividing $\mathfrak{m}$. Let $j \geq 1$ be an integer, and let $j\mathcal{P} = \{jP : P \in \mathcal{P}\}$. Then*

$$\Delta(j\mathcal{P}, \mathfrak{m}, n) \geq \Delta(\mathcal{P}, \mathfrak{m}, n).$$

Proof. Replacing K by a finite extension, we may assume that E has good reduction at all primes dividing $\mathfrak{m}$. Let $v \in M_K^0$ be any place at which E has good reduction, and let $\pi_v \in K$ be a uniformizer at v. Further, let

$$E_0(\bar{K}_v) \subset E_1(\bar{K}_v) \subset E_2(\bar{K}_v) \subset \cdots$$

be the formal group filtration of $E(\bar{K}_v)$; see [20, Chapters IV, VII]. Here $E_0 = E$, since we have assumed good reduction, and E_1 is the formal group. The explicit formula for $\hat{\lambda}_v$ [19, Theorem VI.4.1] then has the form

$$\hat{\lambda}_v(P) = \max\{r \geq 0 : P \in E_r(\bar{K}_v)\} v(\pi_v).$$

Since the filtration consists of subgroups, it is immediate that

$$\hat{\lambda}_v(jP) \geq \hat{\lambda}_v(P) \qquad \text{for all } j \geq 1.$$

Summing over $P \in \mathcal{P}$ and $v \mid \mathfrak{m}$, and dividing by $n^2 \log m$, the desired result is immediate from the definition of Δ. □

Remark 3.1. If E has potential multiplicative reduction at v, then it is possible to have $\hat{\lambda}_v(jP) < \hat{\lambda}_v(P)$, so $\Delta(j\mathcal{P}, \mathfrak{m}, n)$ may be strictly smaller than $\Delta(\mathcal{P}, \mathfrak{m}, n)$.

4. A height lower bound for points on elliptic curves

In this section we prove our second main result, which is an elliptic analogue of the height lower bound given in Theorem 2.1. We do not explicitly keep track of the dependence on the field K or the curve E, although it would be possible to do so. We start with a Fourier averaging estimate

that is analogous to Proposition 2.1 and that has been applied in the past to the elliptic Lehmer conjecture [7], to Lang's height lower bound conjecture [8], and to Arakelov theory [12]. In order to state the result, we use the following useful definition from [11].

Definition. Let $K/\mathbb{Q}$ be a number field. An *M_K-constant* is a map

$$c : M_K \longrightarrow [0,\infty)$$

with the property that $\{v \in M_K : c(v) \neq 0\}$ is a finite set. (For convenience, we consider only non-negative M_K-constants.) A *normalized $M_{\bar{\mathbb{Q}}}$-constant* is a collection of M_K-constants

$$c_K : M_K \longrightarrow \mathbb{R},$$

one for each number field $K/\mathbb{Q}$, satisfying the compatibility condition that for all number fields L/K and all $v \in M_K$,

$$\sum_{w \in M_L,\, w|v} \frac{[L_w : K_v]}{[L : K]} c_L(w) = c_K(v).$$

Proposition 4.1. *Let $E/\bar{\mathbb{Q}}$ be an elliptic curve. There are normalized $M_{\bar{\mathbb{Q}}}$-constants c_1 and c_2, depending only on E, such that for all integers $J \geq 2$, all nontorsion points $P \in E(\bar{\mathbb{Q}})$, and all absolute values $v \in M_{\bar{\mathbb{Q}}}$ we have*

$$\sum_{j=1}^{J} \left(1 - \frac{j}{J+1}\right) \hat{\lambda}_v(jP) \geq -c_1(v)\log(J) - c_2(v). \tag{4.1}$$

(*We may, in fact, take $c_1(v) = 0$ for all nonarchimedean v.*)

Proof. If v is nonarchimedean and E has good reduction at v, then the local height $\hat{\lambda}_v$ is non-negative, so we can take $c_1(v) = c_2(v) = 0$. For nonarchimedean v of bad reduction, the inequality (4.1) with $c_1(v) = 0$ is proven in [8]. Finally, for archimedean v, the local height functions are Green's functions and the desired result follows from a general theorem of Elkies [12, Theorem 5.1] that is valid on curves of positive genus. More precisely, Elkies' theorem says that there is a constant $c = c(E, v)$ such that for any distinct points $P_0, \ldots, P_J \in E(K_v)$ we have

$$\sum_{0 \leq i < j \leq J} \hat{\lambda}_v(P_j - P_i) \geq -\frac{1}{2\pi}(J+1)\log J - cJ. \tag{4.2}$$

(We are using the fact that $\hat{\lambda}_v$ is an even function.) Taking $P_j = jP$ for $0 \leq j \leq J$, we find that

$$\sum_{0 \leq i < j \leq J} \hat{\lambda}_v(P_j - P_i) = \sum_{0 \leq i < j \leq J} \hat{\lambda}_v((j-i)P) = \sum_{j=1}^{J} (J+1-j)\hat{\lambda}_v(jP). \tag{4.3}$$

Combining (4.2) and (4.3) and dividing by $J+1$ gives (4.1). □

We have now assembled all of the tools required to prove our main result on elliptic curves.

Theorem 4.1. *Suppose that we are given the following quantities*:

$K/\mathbb{Q}$	*a number field.*
E/K	*an elliptic curve.*
n	*a positive integer.*
$\mathfrak{m}$	*an integral ideal of K with norm $m = \mathsf{N}_{K/\mathbb{Q}}\mathfrak{m} \ge 2$.*
$\mathcal{P}$	*a finite set of nontorsion points in $E(\bar{K})$.*

Suppose further that E has potential good reduction at every prime dividing $\mathfrak{m}$. Then there is a constant C_E, depending only on E, such that for all integers $J \ge 1$ we have

$$\sum_{P\in\mathcal{P}} \hat{h}(P) \ge \frac{6}{(J+1)(J+2)} \left(\Delta(\mathcal{P},\mathfrak{m},n)\log m - C_E \frac{|\mathcal{P}|}{n^2} \cdot \frac{\log(J+1)}{J} \right).$$

Proof. To ease notation, we let

$$f_j = 1 - \frac{j}{J+1} \qquad \text{and} \qquad F_k = \sum_{j=1}^{J} j^k f_j.$$

Earlier in the proof of Proposition 2.1 we used the values of F_0 and F_1. In this section we will use the values

$$F_0 = \frac{J}{2} \qquad \text{and} \qquad F_2 = \frac{J(J+1)(J+2)}{12}. \tag{4.4}$$

Replacing K by a finite extension, we may assume that $\mathcal{P} \subset E(K)$. We let

$$M_K^{\mathrm{bad}} = M_K^{\infty} \cup \{v \in M_K^0 : E \text{ has bad reduction at } v\}.$$

Then

$$v \in M_K \smallsetminus M_K^{\mathrm{bad}} \quad \Longrightarrow \quad \lambda_v(Q) \ge 0 \quad \text{for all } Q \in E(\bar{K}_v).$$

We compute

$$\begin{aligned} n^2 \sum_{P\in\mathcal{P}} \hat{h}(P) &= \sum_{P\in\mathcal{P}} \hat{h}(nP) \\ &= \sum_{P\in\mathcal{P}} \sum_{v\in M_K} \hat{\lambda}_v(nP) \\ &\ge \sum_{P\in\mathcal{P}} \sum_{v\mid\mathfrak{m}} \hat{\lambda}_v(nP) + \sum_{P\in\mathcal{P}} \sum_{v\in M_K^{\mathrm{bad}}} \hat{\lambda}_v(nP) \\ &= (n^2 \log m)\Delta(\mathcal{P},\mathfrak{m},n) + \sum_{P\in\mathcal{P}} \sum_{v\in M_K^{\mathrm{bad}}} \hat{\lambda}_v(nP). \end{aligned}$$

Replacing $\mathcal{P}$ with $j\mathcal{P} = \{jP : P \in \mathcal{P}\}$ and using Lemma 3.1, which says that $\Delta(j\mathcal{P}, \mathfrak{m}, n) \ge \Delta(\mathcal{P}, \mathfrak{m}, n)$ (this is where we use the assumption that E has potential good reduction at the primes dividing $\mathfrak{m}$), we find that

$$n^2 j^2 \sum_{P\in\mathcal{P}} \hat{h}(P) \ge (n^2 \log m)\Delta(\mathcal{P}, \mathfrak{m}, n) + \sum_{P\in\mathcal{P}} \sum_{v\in M_K^{\mathrm{bad}}} \hat{\lambda}_v(njP).$$

Multiplying both sides by f_j and summing $j = 1$ to J gives

$$\begin{aligned} F_2 n^2 \sum_{P\in\mathcal{P}} \hat{h}(P) & \\ &\ge F_0 n^2 \log(m)\Delta(\mathcal{P}, \mathfrak{m}, n) + \sum_{j=1}^{J} \sum_{P\in\mathcal{P}} \sum_{v\in M_K^{\mathrm{bad}}} f_j \hat{\lambda}_v(njP) \\ &\ge F_0 n^2 \log(m)\Delta(\mathcal{P}, \mathfrak{m}, n) + |\mathcal{P}| \sum_{v\in M_K^{\mathrm{bad}}} \inf_{Q\in E(K)} \sum_{j=1}^{J} f_j \hat{\lambda}_v(jQ). \end{aligned}$$

Proposition 4.1 says that there are normalized $M_{\bar{\mathbb{Q}}}$-constants c_1 and c_2, depending only on E, such that

$$\inf_{Q\in E(K)} \sum_{j=1}^{J} f_j \hat{\lambda}_v(jQ) \ge -c_1(v)\log(J) - c_2(v).$$

Summing over $v \in M_K^{\mathrm{bad}}$ gives constants that depend only on E, so adjusting the constants and using the assumption that $J \ge 1$, we find that there is a constant C_E, depending only on E, such that

$$F_2 n^2 \sum_{P\in\mathcal{P}} \hat{h}(P) \ge F_0 n^2 \log(m)\Delta(\mathcal{P}, \mathfrak{m}, n) - C_E|\mathcal{P}|\log(J+1).$$

Using the formulas (4.4) for F_0 and F_2, dividing by $F_2 n^2$, and adjusting the constant gives the desired result. □

Corollary 2.1 says roughly that the classical Lehmer's conjecture is true for polynomials $f(X)$ such that

(4.5) $\quad \Phi_{n-1}(X)$ divides $f(X)$ in $(\mathbb{Z}/m\mathbb{Z})[X]$ for some $n \ge \epsilon \deg(f)$.

As noted in Remark 1.3, the divisibility condition in (4.5) is stronger than the assertion that $\Delta(\mathcal{A}_f, m, n) \ge (n-1)/n$, where $\mathcal{A}_f$ denotes the set of roots of f. Since we assume that $n \ge 2$, this implies in particular that $\Delta(\mathcal{A}_f, m, n)$ is uniformly bounded away from 0. Thus the following result is an elliptic version of a strengthening of Corollary 2.1.

Definition. Let E/K be an elliptic curve and let $Q \in E(\bar{K})$. We let

$$\mathcal{P}_Q = \{\sigma(Q) : \sigma \in \mathrm{Gal}(\bar{K}/K)\} \quad \text{and} \quad D_Q = [K(Q) : K] = |\mathcal{P}_Q|.$$

We remark that all of the points in $\mathcal{P}_Q$ have the same canonical height; cf. [20, Theorem VIII.5.10].

Corollary 4.1. *Let E/K be an elliptic curve defined over a number field. Fix constants $\delta, \epsilon > 0$. Then the elliptic Lehmer conjecture* (0.4) *is true for the following set of points*:

$$\left\{ Q \in E(\bar{K}) : \begin{array}{l} \text{there exists an integral ideal } \mathfrak{m} \text{ of } K \\ \text{with } \mathsf{N}_{K/\mathbb{Q}}\mathfrak{m} \geq 2 \text{ such that } E \text{ has good} \\ \text{reduction at all primes dividing } \mathfrak{m} \text{ and} \\ \text{an integer } n \geq \max\{\sqrt{\epsilon D_Q}, 2\} \text{ with} \\ \Delta(\mathcal{P}_Q, \mathfrak{m}, n) \geq \delta \end{array} \right\}. \tag{4.6}$$

For points in the set (4.6)*, the constant in* (0.4) *has the form $C_{E/K,\delta,\epsilon} \log m$, where $C_{E/K,\delta,\epsilon}$ is positive and depends only on the indicated quantities.*

Proof. We are given that $n \geq \sqrt{\epsilon D_Q}$ and $\Delta(\mathcal{P}_Q, \mathfrak{m}, n) \geq \delta$. Using these values in Theorem 4.1 together with some trivial estimates yields

$$\sum_{P \in \mathcal{P}_Q} \hat{h}(P) \geq \frac{1}{J^2} \left(\delta \log m - \frac{C_E}{\epsilon} \cdot \frac{\log(J+1)}{J} \right).$$

We now choose J to be the smallest integer satisfying

$$\frac{\log(J+1)}{J} \leq \min\left\{ \frac{\epsilon\delta}{2C_E} \log m, \frac{1}{2} \right\}.$$

This yields an estimate of the desired form

$$D_Q \hat{h}(Q) \geq C_{E/K,\delta,\epsilon} \log m,$$

where we are using the fact, noted earlier, that every point in $\mathcal{P}_Q$ has canonical height equal to $\hat{h}(Q)$. □

5. Other congruence conditions on the coefficients

Cyclotomic polynomials play a key role in Lehmer's conjecture, so the congruence condition (0.2) and the more general congruence divisibility condition (0.3) are natural ones to consider. However, there is no reason not to consider similar congruences in which the cyclotomic polynomial is replaced by some other polynomial. This was done in considerable generality by Samuels [17]. To illustrate, we reprove a special case of one of Samuels' result and use it to make some remarks.

Definition. The *length* of a polynomial $g(X) = \sum a_i X^i \in \mathbb{Z}[X]$ is the quantity

$$L(g) = \sum |a_i|.$$

Theorem 5.1. (Special Case of [17, Corollary 5.3]) *Let $m \ge 2$ and let $f(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree D satisfying $f(1) \ne 0$. Further let $u(X) \in \mathbb{Z}[X]$ be a polynomial of degree at most $D-1$, and suppose that*

$$f(X) \equiv \Phi_D(X) + u(X) \pmod{m},$$

but that $f(X)$ has no roots in common with $\Phi_D(X) + u(X)$. Then

$$\sum_{f(\alpha)=0} h(\alpha) \ge \frac{D}{D+1} \log\left(\frac{m}{L\big(X^{D+1} - 1 + (X-1)u(X)\big)} \right).$$

Proof. We are given that

$$f(X) = \Phi_D(X) + u(X) + mr(X) \qquad \text{for some } r(X) \in \mathbb{Z}[X].$$

Then

$$\begin{aligned}
&\operatorname{Res}\big(f(X), X^{D+1} - 1 + (X-1)u(X)\big) \\
&\qquad = \operatorname{Res}\big(f(X), \Phi_D(X) + u(X)\big) \cdot \operatorname{Res}\big(f(X), X-1\big) \\
&\qquad = \operatorname{Res}\big(mr(X), \Phi_D(X) + u(X)\big) f(1) \\
&\qquad = m^D \operatorname{Res}\big(r(X), \Phi_D(X) + u(X)\big) f(1).
\end{aligned}$$

By assumption, the resultants are nonzero integers, so we find that

$$\begin{aligned}
D \log m &\le \log \Big| \operatorname{Res}\big(f(X), X^{D+1} - 1 + (X-1)u(X)\big) \Big| \\
&= \sum_{f(\alpha)=0} \sum_{v \in M_K^\infty} \log \| \alpha^{D+1} - 1 + (\alpha - 1)u(\alpha) \|_v \\
&= \sum_{v \in M_K^\infty} \sum_{\substack{f(\alpha)=0 \\ \|\alpha\|_v > 1}} \log \| \alpha^{D+1} \|_v \\
&\qquad + \sum_{v \in M_K^\infty} \sum_{\substack{f(\alpha)=0 \\ \|\alpha\|_v > 1}} \log \left\| \frac{\alpha^{D+1} - 1 + (\alpha-1)u(\alpha)}{\alpha^{D+1}} \right\|_v \\
&\qquad + \sum_{v \in M_K^\infty} \sum_{\substack{f(\alpha)=0 \\ \|\alpha\|_v \le 1}} \log \| \alpha^{D+1} - 1 + (\alpha-1)u(\alpha) \|_v \\
&\le h(\alpha^{D+1}) + D \sup_{|z|=1} \log \big| z^{D+1} - 1 + (z-1)u(z) \big| \qquad (5.1) \\
&\le (D+1)h(\alpha) + D \log L\big(X^{D+1} - 1 + (X-1)u(X)\big).
\end{aligned}$$

(We note that in (5.1), it suffices to take the supremum over $|z| = 1$, since $\log|w|$ is harmonic inside the unit disk.) □

Remark 5.1. The upshot of Theorem 5.1 is that if m is sufficiently large, then we obtain a Lehmer-type lower bound. However, in the cyclotomic

case, i.e., $u = 0$, a Fourier averaging argument allowed us to prove nontrivial estimates for all values of $m \geq 2$. We do not know how to perform such an averaging argument in the general case. It would also be interesting to prove a version of Theorem 5.1 under the weaker assumption that $f(X)$ is divisible modulo m by $\Phi_{n-1}(X)$ for, say, $n \geq \epsilon D$. Again we do not have the requisite averaging lemma.

Remark 5.2. If we take $u = 0$ in Theorem 5.1, we obtain the estimate

$$\sum_{f(\alpha)=0} h(\alpha) \geq \frac{D}{D+1} \log \frac{m}{2}.$$

This has the same form as Theorem 0.1, although the constant in Theorem 0.1 is a little bit better than ours (and our estimate is trivial for $m = 2$). On the other hand, it is interesting that such an elementary argument produces a lower bound that agrees with the best known lower bounds [3, 4, 6, 10] up to an additional $O(1/m^2)$.

Remark 5.3. The estimate proven in Theorem 5.1 is nontrivial if and only if $m > L\big(X^{D+1} - 1 + (X-1)u(X)\big)$. As Samuels does in [17], it is sometimes possible to improve the estimate a little bit. We illustrate with the case $u(X) = -2$, so

$$f(X) \equiv X^D + X^{D-1} + \cdots + X^2 + X - 1 \pmod{m}$$

and

$$L\big(X^{D+1} - 1 - (X-1)u(X)\big) = L(X^{D+1} - 2X + 1) = 4.$$

This gives a nontrivial height bound for $m \geq 5$. If D is odd, then the supremum in (5.1) occurs at $z = -1$ and is equal to $\log 4$, but if D is even, then the supremum is strictly smaller than $\log 4$ and we can obtain a small improvement in the theorem. However, for large (even) values of D we have

$$\sup_{|z|=1} |z^{D+1} - 2z + 1| = 4 + O(D^{-2}) \quad \text{at } z \approx -e^{\pi i/(D+1)},$$

so for $m = 4$ we only obtain $\sum h(\alpha) \gg D^{-2}$, which is weaker than Lehmer's conjecture.

6. Proof of proposition 2.1

In this section we give the proof of Proposition 2.1, for which we need the following standard lemma, cf. [2, Lemma 4].

Lemma 6.1. *For all $\theta \in \mathbb{R} \smallsetminus 2\pi i\mathbb{Z}$ and all $t \geq 0$ we have*

$$\log|1 - e^{i\theta}| \leq \log|1 - e^{-t}e^{i\theta}| + \frac{1}{2}t.$$

Proof. For notational convenience we let

$$F_t(\theta) = \log|1 - e^{-t}e^{i\theta}| + \frac{1}{2}t - \log|1 - e^{i\theta}|,$$

so we need to prove that $F_t(\theta) \ge 0$. We have

$$F_0(\theta) = 0 \qquad \text{for all } \theta \in \mathbb{R} \smallsetminus 2\pi i\mathbb{Z}.$$

For $t > 0$ we observe that

$$\log|1 - e^{-t}e^{i\theta}| = \operatorname{Re}\Bigl(\sum_{n=1}^{\infty} -\frac{e^{-nt}e^{in\theta}}{n}\Bigr)$$

is given by an absolutely convergent series, so we may differentiate it term-by-term. Hence

$$\begin{aligned}\frac{\partial F_t}{\partial t}(\theta) &= \operatorname{Re}\Bigl(\sum_{n=1}^{\infty}(e^{-t+i\theta})^n\Bigr) + \frac{1}{2} = \operatorname{Re}\left(\frac{e^{-t+i\theta}}{1 - e^{-t+i\theta}}\right) + \frac{1}{2} \\ &= \frac{1}{2}\cdot\frac{e^{2t}-1}{(e^t - \cos\theta)^2 + \sin^2\theta} > 0 \qquad \text{for all } t > 0.\end{aligned}$$

Thus for any fixed $\theta \in \mathbb{R} \smallsetminus 2\pi i\mathbb{Z}$, the function $F_t(\theta)$ as a function of $t \ge 0$ satisfies $F_0(\theta) = 0$ and $(\partial F_t/\partial t)(\theta) \ge 0$. Hence $F_t(\theta) \ge 0$ for all $t \ge 0$. □

Here is the elementary algebraic verification of the trigonometric identity that was used in the above calculation.

$$\begin{aligned}\operatorname{Re}\left(\frac{e^{-t+i\theta}}{1 - e^{-t+i\theta}}\right) + \frac{1}{2} &= \operatorname{Re}\left(\frac{e^{i\theta}}{e^t - e^{i\theta}}\right) + \frac{1}{2} \\ &= \operatorname{Re}\left(\frac{e^{i\theta}(e^t - e^{-i\theta})}{e^{2t} - 2(\cos\theta)e^t + 1}\right) + \frac{1}{2} \\ &= \operatorname{Re}\left(\frac{e^{i\theta}e^t - 1}{e^{2t} - 2(\cos\theta)e^t + 1}\right) + \frac{1}{2} \\ &= \frac{e^t\cos\theta - 1}{e^{2t} - 2(\cos\theta)e^t + 1} + \frac{1}{2} \\ &= \frac{1}{2}\cdot\frac{e^{2t}-1}{e^{2t} - 2(\cos\theta)e^t + 1} \\ &= \frac{1}{2}\cdot\frac{e^{2t}-1}{(e^t - \cos\theta)^2 + \sin^2\theta}.\end{aligned}$$

Proof of Proposition 2.1. The functions $\log|1 - z^j|$ are harmonic on the open unit disk $|z| < 1$, so the maximum occurs on the boundary. For

$z = e^{i\theta}$ on the unit circle we estimate

$$\begin{aligned}
&\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\log|1-z^j| \\
&\quad\le \sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\left(\log|1-e^{-t}e^{ij\theta}|+\frac{1}{2}t\right) \qquad \text{from Lemma 6.1,} \\
&\quad= \mathrm{Re}\Bigg(\sum_{j=1}^{J}\left(1-\frac{j}{J+1}\right)\sum_{k=1}^{\infty}-\frac{e^{-kt}e^{ijk\theta}}{k}\Bigg)+\frac{Jt}{4} \\
&\quad= \mathrm{Re}\Bigg(\sum_{k=1}^{\infty}\frac{e^{-kt}}{k}\sum_{j=1}^{J}-\left(1-\frac{j}{J+1}\right)e^{ijk\theta}\Bigg)+\frac{Jt}{4} \\
&\quad= \sum_{k=1}^{\infty}\frac{e^{-kt}}{k}\left(\frac{1}{2}-\frac{1}{2J+2}\left|\sum_{j=0}^{J}e^{ijk\theta}\right|^2\right)+\frac{Jt}{4} \\
&\quad\le \sum_{k=1}^{\infty}\frac{e^{-kt}}{2k}+\frac{Jt}{4} \\
&\quad= -\frac{1}{2}\log(1-e^{-t})+\frac{Jt}{4}.
\end{aligned}$$

This estimate holds for all $t > 0$, so in particular we can set $t = \log(1 + 2J^{-1})$, which (after some algebra) gives the estimate

$$\sum_{j=1}^{J} f_j \log|1-z^j| \le \frac{1}{2}\log\left(\frac{J}{2}+1\right)+\frac{J}{4}\log\left(1+\frac{2}{J}\right).$$

Finally we observe that $x\log(1+x^{-1}) \le 1$ for all $x > 0$, which gives the desired result. □

References

[1] F. Amoroso and R. Dvornicich, *A lower bound for the height in abelian extensions*. J. Number Theory **80(2)** (2000), 260–272.

[2] P. E. Blanksby and H. L. Montgomery, *Algebraic integers near the unit circle*. Acta Arith. **18** (1971), 355–369.

[3] P. Borwein, E. Dobrowolski, and M. J. Mossinghoff, *Lehmer's problem for polynomials with odd coefficients*. Ann. of Math. (2) **166(2)** (2007), 347–366.

[4] P. Borwein, K. G. Hare, and M. J. Mossinghoff, *The Mahler measure of polynomials with odd coefficients*. Bull. London Math. Soc. **36(3)** (2004), 332–338.

[5] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*. Acta Arith. **34** (1979), 391–401.

[6] A. Dubickas and M. J. Mossinghoff, *Auxiliary polynomials for some problems regarding Mahler's measure*. Acta Arith. **119(1)** (2005), 65–79.

[7] M. Hindry and J. H. Silverman, *The canonical height and integral points on elliptic curves*. Invent. Math. **93(2)** (1988), 419–450.

[8] M. Hindry and J. H. Silverman, *On Lehmer's conjecture for elliptic curves*. In Séminaire de Théorie des Nombres, Paris 1988–1989, volume **91** of Progr. Math., pages 103–116. Birkhäuser Boston, Boston, MA, 1990.
[9] M. Hindry and J. H. Silverman, *Diophantine Geometry: An Introduction*. Volume **201** of Graduate Texts in Mathematics. Springer-Verlag, New York, 2000.
[10] M. I. M. Ishak, M. J. Mossinghoff, C. Pinner, and B. Wiles, *Lower bounds for heights in cyclotomic extensions*. J. Number Theory **130(6)** (2010), 1408–1424.
[11] S. Lang, *Fundamentals of Diophantine Geometry*. Springer-Verlag, New York, 1983.
[12] S. Lang, *Introduction to Arakelov Theory*. Springer-Verlag, New York, 1988.
[13] S. Lang, *Algebra*. Volume **211** of Graduate Texts in Mathematics. Springer-Verlag, New York, third edition, 2002.
[14] M. Laurent, *Minoration de la hauteur de Néron-Tate*. In Séminaire de Théorie des Nombres, Progress in Mathematics, pages 137–151. Birkhäuser, 1983. Paris 1981–1982.
[15] D. H. Lehmer, *Factorization of certain cyclotomic functions*. Ann. of Math. (2) **34(3)** (1933), 461–479.
[16] D. W. Masser, *Counting points of small height on elliptic curves*. Bull. Soc. Math. France **117(2)** (1989), 247–265.
[17] C. L. Samuels, *The Weil height in terms of an auxiliary polynomial*. Acta Arith. **128(3)** (2007), 209–221.
[18] C. L. Samuels, *Estimating heights using auxiliary functions*. Acta Arith. **137(3)** (2009), 241–251.
[19] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*. Volume **151** of Graduate Texts in Mathematics. Springer-Verlag, New York, 1994.
[20] J. H. Silverman *The Arithmetic of Elliptic Curves*. Volume **106** of Graduate Texts in Mathematics. Springer, Dordrecht, second edition, 2009.
[21] C. J. Smyth, *On the product of the conjugates outside the unit circle of an algebraic integer*. Bull. London Math. Soc. **3** (1971), 169–175.
[22] C. J. Smyth, *Some inequalities for certain power sums*. Acta Arith. **28(3–4)** (1976), 271–273.
[23] C. J. Smyth, *The Mahler measure of algebraic numbers: a survey*. In Number theory and polynomials, volume **352** of London Math. Soc. Lecture Note Ser., pages 322–349. Cambridge Univ. Press, Cambridge, 2008.
[24] C. L. Stewart, *Algebraic integers whose conjugates lie near the unit circle*. Bull. Soc. Math. France **106(2)** (1978), 169–176.

Joseph H. Silverman
Mathematics Department, Box 1917
Brown University, Providence, RI 02912 USA
E-mail: jhs@math.brown.edu
URL: http://www.math.brown.edu/~jhs