ACTA ARITHMETICA
123.3 (2006)

# Mahler's measure of a polynomial in terms of the number of its monomials

by

Edward Dobrowolski (Prince George, BC)

**1. Introduction and statement of the results.** Mahler's measure $M(f)$ of a polynomial $f$ can be defined by either side of the equality

$$\exp\Big(\int_0^1 \log|f(e^{2\pi i\theta})|\,d\theta\Big) = |a_0|\prod_{i=1}^{n}\max(1,|\alpha_i|),$$

where $a_0$ is the leading coefficient of $f$, and the product runs over all its (possibly multiple) zeros.

Although the question of Lehmer, whether Mahler's measure of a polynomial with integral coefficients that is not a product of cyclotomic factors is bounded below, i.e. $M(f) > c > 1$ with an absolute constant $c$, remains still open, multiple lower bounds on $M(f)$ depending on various parameters of $f$ or valid for special classes of polynomials have been found. Here, we are dealing with bounds depending exclusively on the number $k$ of nonzero coefficients of $f$. The first result of that kind was presented in [4],

$$M(f) \ge 1 + \frac{1}{\exp_{k+1} 2k^2}.$$

Later, in [2] it was improved to

$$M(f) \ge 1 + \frac{1}{a\exp(bk^k)}$$

with explicit constants $a \le 13911$ and $b \le 2.27$. The aim of this paper is to further sharpen this result. We shall prove

2000 *Mathematics Subject Classification*: Primary 11R09; Secondary 11C08.

*Key words and phrases*: Lehmer's problem, Mahler's measure, cyclotomic polynomials.

This work originated in June 2003 at the mini-semester on "Mahler's Measure of Polynomials" sponsored by SFU and PIMS. The author takes the opportunity to express his gratitude to the organizers, especially to P. Borwein and S. Choi.

THEOREM 1. *If $f \in \mathbb{Z}[x]$ is a polynomial that is not a product of cyclotomic factors then*

$$M(f) \geq 1 + \frac{1}{\exp(a3^{\lfloor (k-2)/4 \rfloor} k^2 \log k)},$$

*where $k > 1$ is the number of monomials in $f$, and $a < 0.785$.*

This theorem is based on Theorem 2 and Lemma 1 below. Before stating these results we need to introduce some basic notation. We assume that in the expression $f(x) = \sum_{i=1}^{k} a_i x^{n_i}$ of a polynomial with $k$ nonzero coefficients, the exponents $n_1, \dots, n_k$ are strictly decreasing; $f_{\mathrm{c}}$ denotes the product of all cyclotomic factors of $f$, $f_{\mathrm{n}}$ the product of its noncyclotomic factors and possibly a constant, so that $f = f_{\mathrm{c}} f_{\mathrm{n}}$. We say that $f$ has *reciprocal exponents* if the exponents of $x$ in $x^{\deg f} f(x^{-1})$ are the same as in $f(x)$. A polynomial $f$ that satisfies a stronger condition, $x^{\deg f} f(x^{-1}) = \pm f(x)$, is called *reciprocal*. Throughout this work, $\Phi_q$ denotes the $q$th cyclotomic polynomial, and $\theta > 1.32$ denotes the real zero of the polynomial $x^3 - x + 1$. C. J. Smyth [10] proved that $M(f) \geq \theta$ for every polynomial with integral coefficients that is not reciprocal.

THEOREM 2. *Let $f \in \mathbb{Z}[x]$, $f(0) \neq 0$, be a polynomial with $k$ nonzero coefficients. There are positive constants $c_1$ and $c_2$, depending only on $k$, and polynomials $f_0, f_2 \in \mathbb{Z}[x]$ such that if*

$$\deg f_{\mathrm{c}} \geq \left(1 - \frac{1}{c_1}\right) \deg f$$

*then either*

$$f(x) = f_0(x^l), \quad \text{where } \deg f_0 \leq c_2, \tag{1}$$

*or*

$$f(x) = \Big(\prod_i \Phi_{q_i}(x^{l_i})\Big) f_2(x), \text{ where } \min_i\{l_i\} \geq \max\left\{\frac{1}{2c_1}\deg f, \deg f_2\right\}. \tag{2}$$

*The sizes of the constants are: $c_i \leq \exp(3^{\lfloor (k-2)/4 \rfloor} s_i k^2 \log k)$ with $s_1 = 0.636$ and $s_2 = 1.06$ for $f$ with reciprocal exponents; $c_i \leq \exp(3^{\lceil (k-2)/2 \rceil} t_i k^2 \log k)$ with $t_1 = 1.81$ and $t_2 = 2.841$ for $f$ that does not have reciprocal exponents.*

COROLLARY 1. *Let $f(x) = \sum_{i=1}^{k} a_i x^{n_i} \in \mathbb{Z}[x]$, $f(0) \neq 0$, be a polynomial with $k$ nonzero coefficients. If* case (2) *of Theorem 2 occurs then $f_2(x) = \pm \sum_{i=j}^{k} a_i x^{n_i}$ with some $j$, $1 < j \leq k$.*

*Proof.* Let $g(x) = \prod_i \Phi_{q_i}(x^{l_i}) = \sum_{j=1}^{N} b_j x^{m_j}$ with nonzero $b_1, \dots, b_N$. Then $m_N = 0$, $b_N = \pm 1$, and $m_{N-1} \geq \min_i l_i \geq \deg f_2$. Thus $\pm f_2(x)$ must occur in the expression of $f(x) = g(x) f_2(x)$.

Since $f_{\mathrm{n}}$ must divide $f_2$, as an immediate consequence we get

COROLLARY 2. *Suppose that $f$ satisfies the conditions of Theorem 2, including $\deg f_{\mathrm{c}} \geq (1-1/c_1)\deg f$. If the exponents of $f$ have no common factor, i.e., $f(x) \neq f_0(x^l)$ for $l > 1$, then $f_{\mathrm{n}}$ divides a polynomial with integral coefficients that has fewer than $k$ terms and the same Mahler's measure as $f$.*

The other lemma used in the proof of Theorem 1 is stated below.

LEMMA 1. *Let $f \in \mathbb{Z}[x]$ be a polynomial with $k > 1$ nonzero coefficients. If $g$ divides $f, g \in \mathbb{Z}[x]$, $g(0) \neq 0$, $g$ is monic and has no cyclotomic factors then*

$$M(g) \geq 1 + \frac{0.31 \deg g}{k! \deg f}.$$

Hence, by taking $g = f$, we have

COROLLARY 3. *If $f \in \mathbb{Z}[x]$ is a polynomial with $k > 1$ nonzero coefficients and $f$ has no cyclotomic factors then*

$$M(f) \geq 1 + \frac{0.31}{k!}.$$

It seems that a bound of that type was not previously stated in an explicit way. Lemma 8 from [4] implies only $M(f) \geq 1 + 1/\exp(2k^k)$ in that context. That is essentially the same as the bound in [2], obtained for polynomials that allow cyclotomic factors. Interestingly, the bound in Corollary 3 is slightly stronger than a bound obtained in [3] for a much easier case of an irreducible polynomial,

$$M(f) \geq 1 + \frac{\log 2e}{ek^{k+1}}.$$

This can also be slightly strengthened:

PROPOSITION 1. *Let $f \in \mathbb{Z}[x]$ be a monic irreducible polynomial with $k$ nonzero terms. If $f$ is not cyclotomic then*

$$M(f) \geq 1 + \frac{0.17}{2^m m!},$$

*where $m = \lceil k/2 \rceil$.*

Finally, it is helpful in the computations below to start with a larger $k$. For that reason we find bounds for polynomials with a very small number of terms more directly. For quadrinomials we have

PROPOSITION 2. *Let $f \in \mathbb{Z}[x]$, $f(0) \neq 0$, be a monic quadrinomial that is not a product of cyclotomic factors. Then $M(f) \geq \theta$.*

**2. Notation.** If $\alpha$ is an algebraic number then we define $M(\alpha) = M(f)$, where $f \in \mathbb{Z}[x]$ is the minimal polynomial of $\alpha$.

If $f$ is a polynomial in one variable then $|f| = \deg f$.

$\Phi_m$ denotes the $m$th cyclotomic polynomial and $\varphi(x)$ denotes Euler's totient function, so that $|\Phi_m| = \varphi(m)$.

For a vector $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, $\|\mathbf{a}\|$ denotes the ordinary Euclidean norm, $l(\mathbf{a}) = |a_1| + \cdots + |a_n|$ its length, and $\mathrm{h}(\mathbf{a}) = \max_{1 \le i \le n} |a_i|$ its height. For two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{ab}$ denotes the ordinary dot product $\mathbf{a} \cdot \mathbf{b}$.

For a fixed list of exponents $(n_1, \dots, n_k)$, there is an obvious one-to-one correspondence between vectors $\mathbf{a} = (a_1, \dots, a_k) \in \mathbb{R}^k$ and polynomials,

$$\mathbf{a} \leftrightarrow a(x) = \sum_{i=1}^{k} a_i x^{n_i}.$$

In that case a norm of a polynomial refers to the corresponding norm of the vector of its coefficients, e.g., $l(a) = l(\mathbf{a})$.

A *Laurent polynomial* in $n$ variables $\mathbf{x} = (x_1, \dots, x_n)$ is an expression of the form

$$F(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}, \tag{3}$$

where $\mathcal{J} = \mathcal{J}_F$ is a finite set of multi-exponents $\mathbf{j} = (j_1, \dots, j_n) \in \mathbb{Z}^n$, $J = |\mathcal{J}|$ is the number of elements in $\mathcal{J}$, and $\mathbf{x}^{\mathbf{j}} = \prod_{i=1}^{n} x_i^{j_i}$. Further, we write $\mathbf{j} \in \mathcal{J}_F$ only if $a_{\mathbf{j}} \neq 0$, so that $\mathcal{J}_F$ is unambiguously determined by $F$.

If $\mathbf{r} = (r_1, \dots, r_n)$ is a vector in $\mathbb{Z}^n$ then the following operation defines a (Laurent) polynomial in one variable:

$$F_{\mathbf{r}}(x) = F(x^{r_1}, \dots, x^{r_n}) = \sum_{\mathbf{j} \in \mathcal{J}} a_{\mathbf{j}} x^{\mathbf{jr}}.$$

In order to convert a Laurent polynomial $F$ into an ordinary polynomial, let

$$\mathbf{j}_F = (\min_{\mathbf{j} \in \mathcal{J}} (\mathbf{j})_1, \dots, \min_{\mathbf{j} \in \mathcal{J}} (\mathbf{j})_n),$$

where $(\mathbf{j})_i$ denotes the $i$th component of $\mathbf{j}$, and define

$$IF(\mathbf{x}) = \mathbf{x}^{-\mathbf{j}_F} F(\mathbf{x}).$$

Clearly, $IF(\mathbf{x})$ is a polynomial. Moreover, $IF(\mathbf{0}) \neq 0$, unless $F \equiv 0$ itself. We shall denote by $\mathbf{R}_0[[\mathbf{x}]]$ the ring of Laurent polynomials in $n$ variables $\mathbf{x} = (x_1, \dots, x_n)$ and with coefficients in an integral domain $\mathbf{R}$. The divisibility properties of $\mathbf{R}_0[[\mathbf{x}]]$ are almost the same as those of $\mathbf{R}_0[\mathbf{x}]$, except that in $\mathbf{R}_0[[\mathbf{x}]]$ the units are of the form $u\mathbf{x}^{\mathbf{j}}$, where $u$ is a unit of $\mathbf{R}$. For example, if $f, g \in \mathbf{R}_0[[x]]$, $g$ has no multiple zeros, and $\gamma$ is a positive integer, then

$$g(x)^{\gamma} \mid f(x) \;\Leftrightarrow\; g(x) \mid f^{(m)}(x) \text{ for } 0 \le m \le \gamma - 1, \tag{4}$$

where $f^{(0)} = f$ and $f^{(m)} = (d^m/dx^m) f$. To verify this claim, consider the relation between $f^{(m)}$ and $(If)^{(m)}$.

An *extended Laurent cyclotomic polynomial* in $n$ variables is a Laurent polynomial $\Phi_m(\mathbf{x}^{\mathbf{v}})$, where $\mathbf{x}^{\mathbf{v}} = \prod_{i=1}^{n} x_i^{v_i}$ and $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{Z}^n$. An *extended cyclotomic polynomial* is a polynomial of the form $I\Phi_m(\mathbf{x}^{\mathbf{v}})$.

For a fixed positive integer $k$, $P = \prod_{p \le k} p$, where the product runs over prime numbers. It is known that $P \le 3^k$ (see [6] for a reference).

Finally, $e(m/n) = \zeta_n^m = \exp(2\pi i m/n)$.

**3. Proof of Theorem 1.** We proceed by induction on $k$. For $k = 2$, we trivially have $M(f) \ge 2$. For $k = 3$, in the case of a reciprocal polynomial $f$, we trivially get $M(f) \ge (3 + \sqrt{5})/2$, while for nonreciprocal polynomials, $M(f) \ge \theta$ by Smyth's result [10]. Finally, by Proposition 1, we also have $M(f) \ge \theta$ for $k = 4$. This shows the statement of the theorem for $k \le 4$. Let $k \ge 5$, and suppose the theorem is true for all $k' < k$. In view of Smyth's result we also assume that $f$ is reciprocal. There are two possibilities:

Case 1: $|f_{\mathrm{c}}| \le (1 - 1/c_1)|f|$. In that case, by putting $g = f_{\mathrm{n}}$ in Lemma 1, we get

$$M(f) \ge 1 + \frac{0.31}{k! c_1} \ge 1 + \frac{1}{\exp(a 3^{\lfloor (k-2)/4 \rfloor} k^2 \log k)}$$

with $a < 0.785$ for $k \ge 5$.

Case 2: $|f_{\mathrm{c}}| > (1 - 1/c_1)|f|$. By Theorem 2 and Corollary 2, we have still two possibilities:

Case 2.1: $f(x) = f_0(x^l)$ *with* $|f_0| \le c_2$. Now we have $|f_{\mathrm{n}}| < (1/c_1)|f|$, $f_{\mathrm{n}}(x) = f_{0\mathrm{n}}(x^l)$, $f_{\mathrm{c}}(x) = f_{0\mathrm{c}}(x^l)$, with suitable $f_{0\mathrm{n}}, f_{0\mathrm{c}} \in \mathbb{Z}[x]$. Consequently, $|f_{0\mathrm{n}}| < (1/c_1)|f_0|$. Hence,

$$\log M(f) = \log M(f_0) = \log M(f_{0\mathrm{n}}) \ge \frac{1}{4} \left( \frac{\log \log(c_2/c_1)}{\log(c_2/c_1)} \right)^3.$$

The first equality in this formula is an obvious consequence of the definition of Mahler's measure; the inequality is due to Voutier's result [11]. This bound is much stronger than required.

Case 2.2: $f_{\mathrm{n}}$ *divides a polynomial* $f_2$ *with fewer than* $k$ *terms and* $M(f_{\mathrm{n}}) = M(f_2)$. Now, the theorem follows by the induction hypothesis.

## 4. Proof of Theorem 2

**4.1.** *An outline of the proof.* Let $f(x) = a_k + \sum_{i=1}^{k-1} a_i x^{n_i}$. Consider the vector of nonzero exponents of $f$, $\mathbf{n} = (n_1, \dots, n_{k-1})$. An *integer relation* on $\mathbf{n}$ is a vector $\mathbf{b} \in \mathbb{Z}^{k-1}$ such that $\mathbf{bn} = 0$. Let $\Lambda \subset \mathbb{Z}^{k-1}$ be the $(k-2)$-dimensional lattice of all integer relations of $\mathbf{n}$. For a $t$-dimensional sublattice $\Gamma < \Lambda$, $\mathrm{vol}(\Gamma)$ denotes the $t$-dimensional volume of its fundamental region $\Delta_\Gamma$.

Further, for $1 \le t \le k-2$, let

$$h_t = \min_{\substack{\Gamma<\Lambda \\ \dim \Gamma = t}} \operatorname{vol}(\Gamma).$$

Thus, we have a mapping

$$f \mapsto (h_1, \dots, h_{k-2}).$$

Let $j_0$ be a positive integer, and $H_{j_0}, H_{j_0+1}, \dots, H_{k-2}$ real positive constants such that $h_{j_0} \le H_{j_0} < H_{j_0+1} < \cdots < H_{k-2}$. The exact values of $j_0$ and the constants $H_{j_0}, \dots, H_{k-2}$ will be determined later, separately for $f$ with reciprocal exponents, and for $f$ with nonreciprocal exponents. For each of these cases, define $c_2$ by putting

$$c_2 = H_{k-2}. \tag{5}$$

One of the following cases must occur:

C1: $\operatorname{vol}(\Lambda) = h_{k-2} \le c_2$, or

C2: $h_j \le H_j$, but $h_{j+1} > H_{j+1}$ for some $j$ with $j_0 \le j \le k-3$.

In Section 4.3 we show that C1 immediately implies formula (1) of Theorem 2. The treatment of C2 is more complicated. The sizes of the constants $H_{j_0}, H_{j_0+1}, \dots, H_{k-2}$ control the bounds in Theorem 2: the smaller the constants, the sharper bounds we get. However, the constants are defined recursively and form a rapidly increasing sequence. In Section 4.4 we shorten the length of this sequence, by showing the existence of a relatively large $j_0$ with not too large $h_{j_0}$. This provides a good starting point for the sequence. In Section 4.5 we show that the gap between $h_j$ and $h_{j+1}$, assumed in C2, in conjunction with the condition $|f_c| \ge (1 - 1/c_1)|f|$, implies formula (2) of Theorem 2. Finally, in Section 4.6, we compute the constants.

**4.2.** *The tools.* Here, we state the technical lemmas used in the proof. The key lemma employed when dealing with cyclotomic factors of a polynomial is Mann's result [7, Theorem 1], quoted here as

LEMMA 2. *Let $a_1, \dots, a_R$ be distinct nonzero integers, let $q$ be an integer, and suppose that $(a_1, \dots, a_R, q) = 1$. Put $a_0 = 0$. Let $b_r$, $0 \le r \le R$, be nonzero integers. Suppose that*

$$\sum_{r=0}^{R} b_r e(2\pi i a_r/q) = 0,$$

*and that no subsum of this sum vanishes. Then $q$ is square-free, and is composed entirely of primes $p \le R+1$.*

The following is a version of Siegel's lemma due to Bombieri and Vaaler.

LEMMA 3 ([1, Theorem C]). *Let $\sum_{j=1}^{n} a_{ij}x_j = 0$, $i \in \{1, \dots, m\}$, be a linear system of $m$ linearly independent equations in $n > m$ unknowns, with*

*integer coefficients $a_{ij}$. Then there are $n-m$ linearly independent solutions $\mathbf{x}_l = (x_{1l}, \dots, x_{nl}) \in \mathbb{Z}^n$, $l \in \{1, \dots, n-m\}$, with*

$$\prod_{l=1}^{n-m} \mathrm{h}(\mathbf{x}_l) \le D^{-1} \sqrt{|\det(AA^T)|},$$

*where $A$ denotes the $m \times n$ matrix $A = (a_{ij})$, $A^T$ is its transpose and $D$ is the greatest common divisor of the determinants of all minors of $A$ of order $m$.*

NOTE. In the notation of the lemma, let $\Lambda_A$ be the $m$-dimensional lattice in $\mathbb{Z}^n$ spanned by the rows of the matrix $A = \{a_{ij}\}$, and let $\Lambda_x$ be the lattice spanned by the solutions $\mathbf{x}_l$, $l \in \{1, \dots, n-m\}$. Then these solutions form a basis of $d\Lambda_A^{\perp}$, where $d = [\Lambda_A^{\perp} : \Lambda_x]$.

We also need a modified version of Lemma 9 from [4].

LEMMA 4. *Let $\mathbf{a} \in \mathbb{Z}^n$, $\mathbf{a} \neq 0$, be a vector, $B_i > 1$, $i = 1, \dots, n$, real numbers, $T = B_1 \cdots B_n$, and $l = \mathrm{h}(\mathbf{a})$. Then there are vectors $\mathbf{c}, \mathbf{r} \in \mathbb{Z}^n$ and $t \in \mathbb{Z}$ such that*

(1) $1 \le t \le T$,
(2) $t\mathbf{a} = \mathbf{r} + l\mathbf{c}$,
(3) $|r_i| \le lB_i^{-1}$,
(4) $\mathbf{c} \neq 0$ *and* $|c_i| \le B_i^{-1} + T$ *for* $i = 1, \dots, n$.

*Proof.* Consider a system of $n+1$ linear forms $|\tau| \le T$, $|\tau a_i/l - x_i| \le B_i^{-1}$, $i = 1, \dots, n$, in $n+1$ unknowns $\tau$, $x_1, \dots, x_n$. Its determinant is 1 and also $B_1^{-1} \cdots B_n^{-1} T = 1$. Hence, by Minkowski's theorem on linear forms, the system has a nontrivial integer solution $\tau = t$, $x_i = c_i$, $i = 1, \dots, n$. We can also assume that $t \ge 0$. Let $\mathbf{c} = (c_1, \dots, c_n)$ and $\mathbf{r} = t\mathbf{a} - l\mathbf{c}$. Then (1)–(3) are obvious. (1) and $|ta_i/l - c_i| \le B_i^{-1}$ imply the second part of (4). For the first part let $|a_{i_0}| = \mathrm{h}(\mathbf{a})$; then $|ta_{i_0}/l - c_{i_0}| = |ta_{i_0}/|a_{i_0}| - c_{i_0}| \le B_{i_0}^{-1}$ implies that $c_{i_0} \neq 0$, since $t \ge 1$.

We also need some facts from the proofs of Lemmas 1 and 2 from [2]. Unfortunately, the proofs presented there are not very transparent and have a number of typos. For the convenience of the reader, Lemmas 5 to 8 reproduce this material here in a more detailed way. These ideas originated in the work of Montgomery and Schinzel [8].

Let $F(\mathbf{x})$ be a polynomial of the form (3), and $\mathbf{r} \in \mathbb{Z}^n$ be a fixed vector. Define a derivative depending on $\mathbf{r}$ by

$$D_{\mathbf{r}} F(\mathbf{x}) = \sum_{i=1}^{n} r_i x_i \frac{\partial}{\partial x_i} F(\mathbf{x}).$$

The $m$th order derivative is denoted $D_{\mathbf{r}}^m F$. For convenience, we agree that

$D_{\mathbf{r}}^0 F = F$. The derivative $D_{\mathbf{r}}$ is defined in such a way that $\mathcal{J}_{D_{\mathbf{r}}^m F} \subset \mathcal{J}_F$. It also has many properties of an ordinary derivative. For example, the product rule

$$D_{\mathbf{r}}(FG) = (D_{\mathbf{r}}F)G + F(D_{\mathbf{r}}G)$$

holds. Two other important properties are shown in Lemmas 5 and 6 below.

LEMMA 5. *Let $F \in \mathbb{Z}_0[[\mathbf{x}]]$ be a Laurent polynomial in $n$ variables, $g(x) \in \mathbb{Z}[x]$ a polynomial with no multiple nonzero roots, $\gamma$ a positive integer, and $\mathbf{v}, \mathbf{r} \in \mathbb{Z}^n$ be vectors such that $\mathbf{vr} \neq 0$. If $g(\mathbf{x}^{\mathbf{v}}) \mid D_{\mathbf{r}}^m F(\mathbf{x})$ for $m = 0, \ldots, \gamma-1$, then $g(\mathbf{x}^{\mathbf{v}})^\gamma \mid F(\mathbf{x})$.*

*Proof.* Let $g(\mathbf{x}^{\mathbf{v}}) = \prod_{i=1}^N (\mathbf{x}^{\mathbf{v}} - \alpha_i)$. It suffices to prove the lemma for a single factor $\mathbf{x}^{\mathbf{v}} - \alpha$ of $g(\mathbf{x}^{\mathbf{v}})$, where $\alpha = \alpha_i \neq 0$ is one of the roots of $g$. The factors of $g(\mathbf{x}^{\mathbf{v}})$ corresponding to $\alpha_i = 0$ are units in $\mathbb{Z}_0[[\mathbf{x}]]$ and obviously divide $F(\mathbf{x})$. We shall proceed by induction on $\gamma$. For $\gamma = 1$, the claim of the lemma is assumed by its hypothesis. Suppose that the lemma is true for $\gamma = t$, and that $(\mathbf{x}^{\mathbf{v}} - \alpha) \mid D_{\mathbf{r}}^m F(\mathbf{x})$ for $m = 0, \ldots, t$. Then by induction hypothesis,

$$F(\mathbf{x}) = (\mathbf{x}^{\mathbf{v}} - \alpha)^t G(\mathbf{x}), \tag{6}$$

where $G(\mathbf{x}) \in \mathbb{C}_0[[\mathbf{x}]]$. We also have $(\mathbf{x}^{\mathbf{v}} - \alpha) \mid D_{\mathbf{r}}^t F(\mathbf{x})$. We verify easily that $D_{\mathbf{r}}(\mathbf{x}^{\mathbf{v}} - \alpha) = (\mathbf{rv})\mathbf{x}^{\mathbf{v}}$. Hence, by the product rule and successive differentiation of (6), we get

$$D_{\mathbf{r}}^t F(\mathbf{x}) \equiv t!(\mathbf{rv})^t \mathbf{x}^{t\mathbf{v}} G(\mathbf{x}) \bmod (\mathbf{x}^{\mathbf{v}} - \alpha).$$

Since $\alpha \neq 0$, $(\mathbf{x}^{\mathbf{v}} - \alpha) \mid G(\mathbf{x})$. Consequently, $(\mathbf{x}^{\mathbf{v}} - \alpha)^{t+1} \mid F(\mathbf{x})$.

LEMMA 6. *Let $F \in \mathbb{Z}_0[[\mathbf{x}]]$ be a Laurent polynomial in $n$ variables, $g(x) \in \mathbb{Z}[x]$ be a polynomial with no multiple zeros, $\gamma$ a positive integer, and $\mathbf{r} \in \mathbb{Z}^n$. If $g(x)^\gamma \mid F_{\mathbf{r}}(x)$ then $g(x) \mid (D_{\mathbf{r}}^t F)_{\mathbf{r}}(x)$ for $t = 0, 1, \ldots, \gamma - 1$.*

*Proof.* Let $F(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}$ and let $f(x) = F_{\mathbf{r}}(x) = \sum_{\mathbf{j} \in \mathcal{J}} a_{\mathbf{j}} x^{\mathbf{jr}}$. Then

$$D_{\mathbf{r}}^t F(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} (\mathbf{rv})^t a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}. \tag{7}$$

Clearly, $xf'(x) = (D_{\mathbf{r}}F)_{\mathbf{r}}(x)$. To obtain a formula for an arbitrary order of $D_{\mathbf{r}}$, put $f_0 = f$ and $f_{i+1}(x) = xf_i'(x)$ for $i \geq 0$. Hence, by (7), $(D_{\mathbf{r}}^t F)_{\mathbf{r}}(x) = f_t(x)$ for all $t \geq 0$. On the other hand, (4) implies that $g(x) \mid f^{(t)}(x)$ for $0 \leq t \leq \gamma - 1$. The definition of $f_t$ implies that $f_t(x) = \sum_{i=0}^t c_i x^i f^{(i)}$, where the coefficients $c_i$ are positive integers. Hence, we also have $g(x) \mid f_t(x)$ for $0 \leq t \leq \gamma - 1$.

LEMMA 7. *Let $F \in \mathbb{Z}_0[[\mathbf{x}]]$ be a Laurent polynomial in $n$ variables, and $\mathbf{r} \in \mathbb{Z}^n$ be a vector such that $F_{\mathbf{r}}$ has the same number of terms as $F$. Suppose that $\mathbf{0} \in \mathcal{J}_{\mathcal{F}}$, and that $F = F_1 F_2$ where $F_1, F_2 \in \mathbb{Z}_0[[\mathbf{x}]]$; $F_2$ is not divisible by any extended Laurent cyclotomic polynomial, but $\Phi_m(x) \mid F_{2,\mathbf{r}}(x)$. Then*

*there are linearly independent vectors* $\mathbf{v}^{(1)} = \mathbf{j}^{(1)}$ *and* $\mathbf{v}^{(2)} = \mathbf{j}^{(2)} - \mathbf{j}^{(3)}$, *where* $\mathbf{j}^{(i)} \in \mathcal{J}_{\mathcal{F}}$, $i = 1, 2, 3$, *for which*

$$m \mid (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r})P,$$

*where* $P = \prod_{p \le |\mathcal{J}_F|} p$.

*Proof.* Let $\gamma$ be the multiplicity of $\Phi_m(x)$ in $F_{\mathbf{r}}(x)$. Then by Lemma 6,

$$\Phi_m(x) \mid (D^t_{\mathbf{r}}F)_{\mathbf{r}}(x) \quad \text{for } t = 0, 1, \ldots, \gamma - 1.$$

Further, since $\mathbf{0} \in \mathcal{J}_{\mathcal{F}}$,

$$F_{\mathbf{r}}(x) = \sum_{\mathbf{j} \in \mathcal{J}_F} a_{\mathbf{j}} x^{\mathbf{jr}}, \qquad (D^t_{\mathbf{r}}F)_{\mathbf{r}}(x) = \sum_{\mathbf{j} \in \mathcal{J}_F \setminus \{\mathbf{0}\}} (\mathbf{jr})^t a_{\mathbf{j}} x^{\mathbf{jr}} \quad \text{for } t \ge 1.$$

With agreement that $(\mathbf{0r})^0$ represents 1, we can treat these formulas simultaneously. Hence,

$$\sum_{\mathbf{j} \in \mathcal{J}_F} (\mathbf{jr})^t a_{\mathbf{j}} e(\mathbf{jr}/m) = 0 \quad \text{for } 0 \le t \le \gamma - 1.$$

Put $\mathcal{J}_0 = \mathcal{J}_F$ and $\mathcal{J}_t = \mathcal{J}_F \setminus \{\mathbf{0}\}$ for $t \ge 1$. Let

$$\mathcal{J}_t = \bigcup_{s=1}^{I_t} \mathcal{J}_{ts} \quad \text{for } 1 \le t \le \gamma - 1$$

be partitions such that

$$\sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} e(\mathbf{jr}/m) = 0 \quad \text{for } 0 \le t \le \gamma - 1 \text{ and } 1 \le s \le I_t,$$

and no subsum of these sums vanishes.

One of the sets $\mathcal{J}_{0s}$, say $\mathcal{J}_{0\hat{s}}$, contains $\mathbf{0}$. Put $\mathbf{j}^*_{0\hat{s}} = \mathbf{0}$, and choose arbitrary $\mathbf{j}^*_{ts} \in \mathcal{J}_{ts}$ for all $(t, s) \ne (0, \hat{s})$. Then

$$\sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} e((\mathbf{j} - \mathbf{j}^*_{ts})\mathbf{r}/m) = 0 \quad \text{for } 0 \le t \le \gamma - 1 \text{ and } 1 \le s \le I_t. \tag{8}$$

Consider the set

$$S = \{\mathbf{j} - \mathbf{j}^*_{ts} \mid \mathbf{j} \in \mathcal{J}_{ts},\ 0 \le t \le \gamma - 1,\ 1 \le s \le I_t\}.$$

If $S = \{\mathbf{0}\}$ then the lemma is vacuously true. Suppose then that $S \ne \{\mathbf{0}\}$. We shall show that $S$ has two linearly independent vectors. Suppose to the contrary that $S$ spans a one-dimensional lattice, and let $\mathbf{v}$ be its generator. Then each vector of $S$ is of the form $\mathbf{j} - \mathbf{j}^*_{ts} = c_{\mathbf{j}}\mathbf{v}$, where $c_{\mathbf{j}}$ is an integer. Equations (8) now take form

$$\sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} e(c_{\mathbf{j}}\mathbf{vr}/m) = 0 \quad \text{for } 0 \le t \le \gamma - 1 \text{ and } 1 \le s \le I_t.$$

We have $\mathbf{vr} \neq 0$, since otherwise $F_{\mathbf{r}}$ would reduce to a single term. Let $l = m/(m, \mathbf{vr})$ and $\lambda = \mathbf{vr}/(m, \mathbf{vr})$. Then $(\lambda, l) = 1$ and

$$\sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} e(c_{\mathbf{j}} \lambda / l) = 0 \quad \text{for } 0 \leq t \leq \gamma - 1 \text{ and } 1 \leq s \leq I_t.$$

Hence,

$$\Phi_l(y) \,\Big|\, \sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} y^{c_{\mathbf{j}}} \quad \text{for } 0 \leq t \leq \gamma - 1 \text{ and } 1 \leq s \leq I_t.$$

Since we also have

$$D_{\mathbf{r}}^t F(\mathbf{x}) = \sum_{s=1}^{I_t} \mathbf{x}^{\mathbf{j}_{ts}^*} \sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} \mathbf{x}^{\mathbf{j} - \mathbf{j}_{ts}^*} \quad \text{for } 0 \leq t \leq \gamma - 1,$$

by substituting $y = \mathbf{x}^{\mathbf{v}}$, we conclude that

$$\Phi_l(\mathbf{x}^{\mathbf{v}}) \,|\, D_{\mathbf{r}}^t F(\mathbf{x}) \quad \text{for } 0 \leq t \leq \gamma - 1.$$

By Lemma 5, $\Phi_l(\mathbf{x}^{\mathbf{v}})^\gamma \,|\, F(\mathbf{x})$. Hence, by definition of $F_1$ and $F_2$, $\Phi_l(\mathbf{x}^{\mathbf{v}})^\gamma \,|\, F_1(\mathbf{x})$. Further, since $m \,|\, l\mathbf{vr}$, $\Phi_m(x) \,|\, \Phi(x^{\mathbf{vr}})$. Hence, $\Phi_m(x)^\gamma \,|\, F_{1,\mathbf{r}}(x)$. On the other hand, we have assumed that $\Phi_m(x)$ divides $F_{2,\mathbf{r}}(x)$. Therefore, $\Phi_m(x)^{\gamma+1} \,|\, F_{\mathbf{r}}(x)$. This contradicts the choice of $\gamma$, and we conclude that $S$ has two linearly independent vectors.

Clearly, in selecting a pair of linearly independent vectors from $S$, we can always start by selecting any nonzero vector from this set. By (8), the component $\mathcal{J}_{0\hat{s}}$ that contains $\mathbf{0}$ must also contain at least one nonzero vector. Let $\mathbf{v}^{(1)} = \mathbf{j}^{(1)}$ be such a vector. Select $\mathbf{v}^{(1)}$ and complete the pair by choosing any vector $\mathbf{v}^{(2)} = \mathbf{j}^{(2)} - \mathbf{j}^{(3)}$ from $S$ that is not a multiple of $\mathbf{v}^{(1)}$. Let $g_{ts} = \gcd((\mathbf{j} - \mathbf{j}_{ts}^*)\mathbf{r} \mid \mathbf{j} \in \mathcal{J}_{ts})$, $q_{ts} = m/(m, g_{ts})$, and $m_{\mathbf{j}} = (\mathbf{j} - \mathbf{j}_{ts}^*)\mathbf{r}/g_{ts}$. Then equations (8) take the form

$$\sum_{\mathbf{j} \in \mathcal{J}_{ts}} (\mathbf{jr})^t a_{\mathbf{j}} e(m_{\mathbf{j}}/q_{ts}) = 0 \quad \text{for } 0 \leq t \leq \gamma - 1 \text{ and } 1 \leq s \leq I_t.$$

Each of these equations satisfies the conditions of Lemma 2. It follows that $q_{ts} \,|\, \prod_{p \leq |\mathcal{J}_{ts}|} p$. Together with definition of $q_{ts}$ this gives

$$m \,|\, P(m, g_{ts}) \quad \text{for } 0 \leq t \leq \gamma - 1 \text{ and } 1 \leq s \leq I_t,$$

where $P = \prod_{p \leq |\mathcal{J}_{\mathcal{F}}|} p$. Consequently, $m \,|\, (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r})P$.

LEMMA 8. *Let $F \in \mathbb{Z}_0[[\mathbf{x}]]$ be a Laurent polynomial in $n$ variables and $\mathbf{r} \in \mathbb{Z}^n$ be a vector such that $F_{\mathbf{r}}$ has the same number of terms as $F$. Suppose that $\mathbf{0} \in \mathcal{J}_{\mathcal{F}}$, and that $F = F_1 F_2$, where $F_1, F_2 \in \mathbb{Z}_0[[\mathbf{x}]]$. Let $\Delta = |IF_{\mathbf{r}}|$ and $\Delta_2 = |IF_{2,\mathbf{r}}|$. If $F_2$ is not divisible by any extended Laurent cyclotomic polynomial, but the sum of the degrees of all cyclotomic factors of $IF_{2,\mathbf{r}}(x)$ counted with multiplicities exceeds $\frac{1}{2}|IF_{2,\mathbf{r}}(x)|$, then $\mathbf{vr} = 0$ for some nonzero vector $\mathbf{v}$ of the form $\mathbf{v} = a(\mathbf{j}^{(2)} - \mathbf{j}^{(3)}) - b\mathbf{j}^{(1)}$, where $\mathbf{j}^{(i)} \in \mathcal{J}_{\mathcal{F}}$, $i = 1, 2, 3$,*

*and $a, b$ are integers such that* $\max\{|a|,|b|\} < PJ^4\Delta/\Delta_2$, *where* $J = |\mathcal{J}_\mathcal{F}|$ *and* $P = \prod_{p\le J} p$.

*Proof.* Let $\gamma_m$ be the multiplicity of $\Phi_m$ in $IF_{2,\mathbf{r}}$. The conditions of the lemma imply that

$$\sum_{\Phi_m | IF_{2,\mathbf{r}}} \gamma_m \varphi(m) > \frac{1}{2}\,|IF_{2,\mathbf{r}}| = \frac{1}{2}\,\Delta_2.$$

By Lemma 7, for every factor $\Phi_m$ of $IF_{2,\mathbf{r}}$, there are linearly independent vectors $\mathbf{v}^{(1)} = \mathbf{j}^{(1)}$ and $\mathbf{v}^{(2)} = \mathbf{j}^{(2)} - \mathbf{j}^{(3)}$ such that $m \,|\, (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r})P$. For each factor $\Phi_m$, select a pair of such vectors. Clearly, only one of the pairs: $(\mathbf{v}^{(1)}, \mathbf{v}^{(2)})$ or $(\mathbf{v}^{(1)}, -\mathbf{v}^{(2)})$, has to be selected. Hence, we need to select no more than $(J-1)\binom{J}{2} \le \frac{1}{2}J^3$ distinct pairs. Let $\mathcal{P}$ be the set of all selected pairs, and let

$$g = \max\{(\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r}) \mid (\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \in \mathcal{P}\}.$$

We have $\gamma_m \le J-1$ by Hajós' lemma [5]. Hence,

$$\begin{aligned}\frac{1}{2}\,J^3(J-1)gP &> (J-1) \sum_{(\mathbf{v}^{(1)},\mathbf{v}^{(2)})\in\mathcal{P}} (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r})P \\ &= (J-1) \sum_{(\mathbf{v}^{(1)},\mathbf{v}^{(2)})\in\mathcal{P}} \sum_{m | (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r})P} \varphi(m) \ge \sum_{\Phi_m | IF_{2,\mathbf{r}}} \gamma_m \varphi(m) > \frac{1}{2}\,\Delta_2.\end{aligned}$$

Therefore, for some pair of linearly independent vectors $(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \in \mathcal{P}$,

$$g = (\mathbf{v}^{(1)}\mathbf{r}, \mathbf{v}^{(2)}\mathbf{r}) > \frac{\Delta_2}{PJ^4}.$$

Put $a = \mathbf{v}^{(2)}\mathbf{r}/g$, $b = \mathbf{v}^{(1)}\mathbf{r}/g$, and $\mathbf{v} = a\mathbf{v}^{(1)} - b\mathbf{v}^{(2)}$. Clearly, $\mathbf{v}\mathbf{r} = 0$, $\mathbf{v} \ne \mathbf{0}$, and also

$$\max\{|\mathbf{v}^{(1)}\mathbf{r}|, |\mathbf{v}^{(2)}\mathbf{r}|\} \le \max_{\mathbf{j}^{(1)},\mathbf{j}^{(2)},\mathbf{j}^{(3)}\in\mathcal{J}_F} \{|\mathbf{j}^{(1)}\mathbf{r}|, |(\mathbf{j}^{(2)} - \mathbf{j}^{(3)})\mathbf{r}|\} = \Delta.$$

Hence,

$$\max\{|a|,|b|\} < \Delta/g \le PJ^4\Delta/\Delta_2.$$

**4.3.** *The case of small* $\mathrm{vol}(\Lambda)$. We have

LEMMA 9. *Let* $f(x) = \sum_{i=1}^{k} a_i x^{n_i} \in \mathbb{Z}[x]$, $f(0) \ne 0$, *be a polynomial with $k$ nonzero terms and let $\Lambda$ be the lattice of integer relations of* $\mathbf{n} = (n_1, \dots, n_{k-1})$. *If* $|f| \le \mathrm{vol}(\Lambda)$ *then there exist a positive integer $l$ and a polynomial* $f_0 \in \mathbb{Z}[x]$ *such that*

$$f(x) = f_0(x^l) \quad \textit{and} \quad |f_0| \le \mathrm{vol}(\Lambda).$$

*Proof.* Let $A$ be a $(k-2)\times(k-1)$ matrix whose rows form a basis of $\Lambda$. Then $\sqrt{AA^T} = \mathrm{vol}(\Lambda)$. By Lemma 3, the system $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\mathbf{x} = (m_1, \dots, m_{k-1}) \in \mathbb{Z}^{k-1}$ with $\mathrm{h}(\mathbf{x}) \le \mathrm{vol}(\Lambda)$. Without loss of

generality we can assume that $\mathbf{x}$ is a basis of the one-dimensional lattice of solutions of the system. By definition of $\Lambda$, $\mathbf{n}$ is another solution of that system. Hence $\mathbf{n} = l\mathbf{x}$, where $l$ is a nonzero integer. By replacing $\mathbf{x}$ by $-\mathbf{x}$, if necessary, we can assume that $l$ is positive. The lemma follows by putting $f_0(x) = a_k + \sum_{i=1}^{k-1} a_i x^{m_i}$.

Since $c_2 = H_{k-2} \geq \mathrm{vol}(\Lambda)$, case (1) of Theorem 2 follows.

**4.4.** *The initial lattice of integer relations of* $\mathbf{n}$. First, we need to establish a simple fact about the rank of a sparse matrix.

LEMMA 10. *Suppose that* $\mathbf{A}$ *is an* $m \times n$ *matrix with entries in an arbitrary field, such that*

(1) *Every row of* $\mathbf{A}$ *has one or two nonzero entries.*
(2) *Every column of* $\mathbf{A}$ *has at least one nonzero entry.*

*Then* $\mathrm{rank}(\mathbf{A}) \geq \lceil n/2 \rceil$.

*Proof.* We proceed by induction on $n$. Obviously, the lemma is true for $n \leq 2$. Suppose that $n \geq 3$ and that the lemma is true for $n' < n$. There are two possibilities:

1. $\mathbf{A}$ has a row with only one nonzero entry.
2. Every row of $\mathbf{A}$ has exactly two nonzero entries.

In the first case, suppose that $a_{ij}$ is the only nonzero entry of the $i$th row. Delete from $\mathbf{A}$ every row with a single nonzero entry occurring in the $j$th column as well as the $j$th column itself. This ensures that the resulting matrix $\mathbf{A}'$ with $n-1$ columns satisfies both conditions of the lemma. Hence, by induction hypothesis

$$\mathrm{rank}(\mathbf{A}) \geq 1 + \mathrm{rank}(\mathbf{A}') \geq 1 + \left\lceil \frac{n-1}{2} \right\rceil \geq \left\lceil \frac{n}{2} \right\rceil.$$

In the second case, suppose that the two nonzero entries of the first row of $\mathbf{A}$ occur in the $p$th and $r$th columns. Delete from $\mathbf{A}$ every row that has nonzero entries in both of these columns as well as the columns themselves. Again, the resulting matrix $\mathbf{A}'$, this time with $n-2$ columns, satisfies the conditions of the lemma, and by induction hypothesis

$$\mathrm{rank}(\mathbf{A}) \geq 1 + \mathrm{rank}(\mathbf{A}') \geq 1 + \left\lceil \frac{n-2}{2} \right\rceil \geq \left\lceil \frac{n}{2} \right\rceil.$$

LEMMA 11.

PART I. *Let* $k \geq 4$. *If the conditions of Theorem* 2 *are satisfied, in particular* $|f_c| \geq (1 - 1/c_1)|f|$, *then* $\mathbf{n}$ *has at least* $k_0 \geq \lfloor (k-1)/2 \rfloor$ *linearly independent integer relations* $\mathbf{b}_i$ *with* $\mathrm{h}(\mathbf{b}_i) \leq k 3^k \lfloor k/2 \rfloor^k$, $i = 1, \ldots, k_0$. *Moreover, each of the vectors* $\mathbf{b}_i$ *has exactly two or three nonzero components.*

PART II. *If in additions to the conditions listed in Part I, $f$ has reciprocal exponents, then we have*:

*For $k$ even, $\mathbf{n}$ has at least $\lceil (3k-10)/4 \rceil$ linearly independent integer relations, $(k-2)/2$ of which has height 1.*

*For $k$ odd, $\mathbf{n}$ has at least $\lceil (3k-9)/4 \rceil$ linearly independent relations, one of the relations has height 2 and $(k-3)/2$ have height 1.*

*In either case the remaining relations have height no greater than $k3^k \lfloor k/2 \rfloor^k$.*

*Proof of Part I.* For better clarity we distinguish two separate steps.

STEP 1. Recall that $f(x) = a_k + \sum_{i=1}^{k-1} a_i x^{n_i}$, and consider the set of exponents of $f$, $\mathcal{J} = \{n_1, \ldots, n_{k-1}, 0\}$.

*The condition $|f_c| \geq (1 - 1/c_1)|f|$ implies that there is a positive integer $g \geq |f|/k3^k \lfloor k/2 \rfloor^k$ and a partition $\mathcal{J} = \bigcup_{t=1}^{I} \mathcal{J}_t$ of $\mathcal{J}$ into subsets $\mathcal{J}_t$, each of cardinality at least two, such that the exponents $n_i$ within each of the components $\mathcal{J}_t$ are congruent modulo $g$.*

It is more convenient to work with the set $\{1, \ldots, k\}$ of subscripts of the exponents $n_i \in \mathcal{J}$ rather than with the set $\mathcal{J}$ itself. Suppose that $\Phi_m \mid f$, so that $\sum_{i=1}^{k} a_i e(n_i/m) = 0$. Clearly, there is a partition $\pi_m = \{J_t \mid 1 \leq t \leq I\}$ of the set $\{1, \ldots, k\}$ into subsets $J_t$ such that

$$\sum_{i \in J_t} a_i e(n_i/m) = 0 \quad \text{for } 1 \leq t \leq I, \tag{9}$$

and no subsum of these sums vanishes. Let $i_t \in J_t$ be the index such that $n_{i_t} = \min\{n_i \mid i \in J_t\}$. Put $\widehat{n}_t = n_{i_t}$, $g_t = \gcd(n_i - \widehat{n}_t \mid i \in J_t)$, $q_t = m/(m, g_t)$, and $m_i = (n_i - \widehat{n}_t)/g_t$. Then

$$\sum_{i \in J_t} a_i e(m_i/q_t) = 0 \quad \text{for } 1 \leq t \leq I.$$

Each of these sums satisfies the conditions of Lemma 2. We have $R+1 \leq k$, so that $q_t \mid P$, where $P = \prod_{p \leq k} p$. Consequently, $m \mid (g_t, m)P$. As this is true for all $t$, we conclude that

$$m \mid g_\pi P,$$

where $g_\pi = \gcd(g_1, \ldots, g_I)$. For each factor $\Phi_m$ of $f$ select a partition satisfying (9). Let $\Pi$ denote the set of all selected partitions. The equations (9) imply that every component $J_t$ of each partition from $\Pi$ has at least two elements. Consequently, $\Pi$ has at most $\lfloor k/2 \rfloor^k$ partitions. Suppose that $|f_c| \geq (1 - 1/c_1)|f|$, and let $f_c = \prod \Phi_m^{\gamma_m}$. By Hajós' lemma [5], the multiplicity of a zero in a polynomial with $k$ terms is no greater than $k-1$, thus

$\gamma_m \le k-1$ for each $m$. We have

$$(1-1/c_1)|f| \le |f_{\mathrm{c}}| = \sum_{\Phi_m | f} \gamma_m \varphi(m) \le (k-1) \sum_{\Phi_m | f} \varphi(m)$$
$$\le (k-1) \sum_{\pi \in \Pi} \sum_{m | g_\pi P} \varphi(m) = (k-1) \sum_{\pi \in \Pi} g_\pi P \le (k-1) 3^k \lfloor k/2 \rfloor^k g,$$

where $g = \max_{\pi \in \Pi} g_\pi$. Hence, there is a partition $\pi_0 = \{J_t \mid 1 \le t \le I_{\pi_0}\}$ in $\Pi$ for which

$$g = g_{\pi_0} \ge \frac{(1-1/c_1)|f|}{(k-1)3^k \lfloor k/2 \rfloor^k} \ge \frac{|f|}{k 3^k \lfloor k/2 \rfloor^k}. \tag{10}$$

The last inequality is valid by the choice of $c_1$. This proves the claim of Step 1.

Step 2. Recall that $\mathbf{n} = (n_1, \dots, n_{k-1})$ and $\mathcal{J} = \{n_1, \dots, n_{k-1}, 0\}$.

*If there is a positive integer $g$ and a partition $\mathcal{J} = \bigcup_{t=1}^{I} \mathcal{J}_t$ of $\mathcal{J}$ into subsets $\mathcal{J}_t$ of cardinality at least two such that the exponents $n_i$ within each $\mathcal{J}_t$ are congruent modulo $g$, then $\mathbf{n}$ has at least $\lfloor (k-1)/2 \rfloor$ linearly independent integer relations. The heights of the vectors representing the relations do not exceed $|f|/g$, and each vector has exactly two or three nonzero components.*

Suppose that such a partition of $\mathcal{J}$ exists. Let $\pi_0 = \bigcup_{t=1}^{I_{\pi_0}} J_t$ be the partition of the set $\{1, \dots, k\}$ of subscripts of the exponents $n_i \in \mathcal{J}$ corresponding to the partition of $\mathcal{J}$. One of the subsets of $\pi_0$, say $J_{t_0}$, must contain $k$ and at least one other element. In our notation, $\widehat{n}_{t_0} = n_k = 0$, and let $p$ be any fixed element of $J_{t_0}$ different than $k$. We have

$$g \,|\, (n_i - \widehat{n}_t, n_p) \quad \text{for } 1 \le t \le I_{\pi_0} \text{ and } i \in J_t. \tag{11}$$

By taking $a = n_p/g$ and $b_{it} = (n_i - \widehat{n}_t)/g$ we get integer relations

$$a n_i - a \widehat{n}_t - b_{it} n_p = 0 \quad \text{for } 1 \le t \le I_{\pi_0} \text{ and } i \in J_t \setminus \{i_t\}, \tag{12}$$

where the integers $a$ and $b_{it}$ have absolute values no greater than $|f|/g$. For $t \ne t_0$, vector representations of these relations are of the form

$$(0, \dots, 0, a, 0, \dots, 0, -a, 0, \dots, 0, -b_{it}, 0, \dots, 0) \in \mathbb{Z}^{k-1}, \tag{13}$$

where the only nonzero components $a$, $-a$, and $-b_{it}$ occur in the $i$th, $i_t$th and $p$th positions, respectively. For $t = t_0$, $n_{t_0} = 0$, and since we do not include that exponent in $\mathbf{n}$, the corresponding vectors are of the form

$$(0, \dots, 0, a, 0, \dots, 0, -b_{it_0}, 0, \dots, 0) \in \mathbb{Z}^{k-1}, \tag{14}$$

with $a$ in the $i$th and $-b_{it_0}$ in the $p$th positions.

We now prove that these vectors are linearly independent. For this, let $\mathbf{M}$ be the matrix whose rows consist of all the vectors (13) and (14). Let $\mathbf{M}_a$ be the submatrix of $\mathbf{M}$ formed by the columns of $\mathbf{M}$ containing the "$+a$" entries. From (12) and the fact that the sets $J_t$, $1 \le t \le I_{\pi_0}$, are disjoint, we conclude that each "$+a$" entry is the only nonzero entry in its column and its row in $\mathbf{M}_a$. Thus $\mathbf{M}_a$ is nonsingular and the rows of $\mathbf{M}$ are linearly independent. Obviously, $\operatorname{rank}\mathbf{M} = \operatorname{rank}\mathbf{M}_a$, and both ranks are equal to the number of columns in $\mathbf{M}_a$. For $t \neq t_0$, each of the sets $J_t$ corresponds to $|J_t| - 1$ columns of $\mathbf{M}_a$, while $J_{t_0}$ corresponds to only $|J_{t_0}| - 2$ columns, since it contains both indices, $p$ and $i_{t_0} = k$. Hence,

$$\operatorname{rank}(\mathbf{M}) = \Big(\sum_{t=1}^{I_{\pi_0}} (|J_t| - 1)\Big) - 1 = k - I_{\pi_0} - 1.$$

Finally, since each of the sets $J_t$ has at least two elements and their union is $\{1, \dots, k\}$, the number of such sets, $I_{\pi_0}$, is $\le \lfloor k/2 \rfloor$. Hence, there are at least $k - \lfloor k/2 \rfloor - 1 = \lfloor (k-1)/2 \rfloor$ linearly independent relations.

This, together with the bound on $g$ given in Step 1, concludes the proof of Part I of the lemma.

*Proof of Part II.* Part I guarantees the existence of $k_0 \ge \lfloor (k-1)/2 \rfloor$ linearly independent relations. The additional condition of the reciprocity of the exponents of $f$ provides the extra relations

$$n_1 - n_i - n_{k-i+1} = 0 \tag{15}$$

for $i = 2, \dots, \lceil k/2 \rceil$. We now show that these sets of relations combined together contain the required number of linearly independent relations. For this, consider a $(k_0 + \lceil k/2 \rceil - 1) \times (k-1)$ matrix $\mathbf{M}$ whose first $k_0$ rows are the vectors (13) and (14), constructed in the proof of Part I, and the remaining rows are the vectors corresponding to the relations (15). Thus, for $k$ even, this matrix has the form

$$\begin{bmatrix} c_{11} & \dots & c_{1,k-1} \\ \vdots & \dots & \vdots \\ c_{k_0 1} & \dots & c_{k_0,k-1} \\ 1 & -1 & -1 \\ \vdots & \ddots & \iddots \\ 1 & -1 & -1 \end{bmatrix},$$

and for $k$ odd, the form

$$\begin{bmatrix} c_{11} & \dots & c_{1,k-1} \\ \vdots & \dots & \vdots \\ c_{k_0 1} & \dots & c_{k_0,k-1} \\ 1 & -1 & -1 \\ \vdots & \ddots & \iddots \\ 1 & -1 & -1 \\ 1 & -2 & \end{bmatrix}.$$

We discuss the case of $k$ odd only. The case of $k$ even is almost identical and is left to the reader. We are going to estimate the rank of $\mathbf{M}$. For this, we decompose $\mathbf{M}$ into four blocks: Divide the row indices into $R_1 = \{1 \le i \le k_0\}$ and $R_2 = \{k_0 + 1 \le i \le k_0 + (k-1)/2\}$, and the column indices into $C_1 = \{1 \le j \le (k-1)/2\}$ and $C_2 = \{(k+1)/2 \le j \le k-1\} \setminus \{p\}$. Now, $\mathbf{M}$ decomposes into four blocks $B_{ij}$, $i, j \in \{1, 2\}$, where $B_{ij}$ has the row indices in $R_i$ and column indices in $C_j$. Let $p \in J_{t_0}$ be the index of the nonzero exponent $\widehat{n}_t$ defined in Part I. If $1 \in J_{t_0}$, delete the 1st and the $p$th columns from $\mathbf{M}$. This will produce a single row of zeros. Delete that row as well. Let $\mathbf{M}'$ be the resulting matrix. Since the $p$th column is deleted, the rows of $\mathbf{M}'$ with indices in $C_1$ have the form $(0, \dots, 0, a, 0, \dots, 0, -a, 0, \dots, 0)$ or $(0, \dots, 0, a, 0, \dots, 0)$. If $p \ge (k+1)/2$ we shall use the nonzero entries from $B_{22}$ in elementary row operations on $\mathbf{M}'$ to eliminate all nonzero entries from $B_{12}$. If $p < (k+1)/2$, we use the nonzero entries from $B_{21}$ to eliminate all nonzero entries from $B_{11}$. Since these two cases are symmetric, we assume without loss of generality that $p \ge (k+1)/2$. The process of elimination is described below.

Let $\varrho_i$ be the $i$th row of the matrix formed by $B_{11}$ and $B_{12}$. There are three possibilities:

(1) All nonzero entries of $\varrho_i$ are in $B_{11}$.
(2) All nonzero entries of $\varrho_i$ are in $B_{12}$.
(3) One nonzero entry of $\varrho_i$ is in $B_{11}$ and the other in $B_{12}$.

If (1) occurs, $\varrho_i$ has no nonzero entries in $B_{12}$, and no elimination is necessary. Suppose that (2) occurs. There are three possibilities:

(i) *$\varrho_i$ has only one nonzero entry and it occurs in the $(k+1)/2$th column.* This assumes that $p \ne (k+1)/2$, otherwise this column would be already deleted. The $(k+1)/2$th column also contains $-2$ entry in the last row of $\mathbf{M}'$. By adding an appropriate multiple of the last row to $\varrho_i$, $\varrho_i$ will change into a row of zeros. Delete that row.

(ii) *$\varrho_i$ has only one nonzero entry and it occurs in the $j$th column, where $j \neq (k+1)/2$.* In that case there is an $\widehat{i} \in R_2$ such that the $\widehat{i}$th row, $\varrho_{\widehat{i}}$, has a $-1$ entry in the $j$th column. The operation $a\varrho_{\widehat{i}} + \varrho_i \mapsto \varrho_i$ eliminates the nonzero entry $a$ of $\varrho_i$ in $B_{12}$ and produces a $-a$ entry in the $(k-j+1)$th column of $B_{11}$.

(iii) *$\varrho_i$ has two nonzero entries $a$ and $-a$ in columns $j_1$ and $j_2$ of $B_{12}$.* The elimination of these entries by two elementary operations as in the previous step will produce two nonzero entries $-a$ and $a$ in columns $k-j_1+1$ and $k-j_2+1$ of $B_{11}$.

Suppose that (3) occurs. We have two possibilities:

(i) *The nonzero entries $a$ and $-a$ of $\varrho_i$ occur in columns $j$ and $k-j+1$, corresponding to the "reciprocal" exponents of $f$.* There is a row with index in $R_2$ that has entries $-1$ in the same columns, i.e., below $a$ and $-a$. By adding a multiple of that row to eliminate the entry of $\varrho_i$ occurring in $B_{12}$ we will double the entry of $\varrho_i$ occurring in $B_{11}$.

(ii) *The nonzero entries of $\varrho_i$ occur in columns $j_1$ and $j_2 \neq k-j_1+1$.* Now, after elimination of the nonzero entry of $\varrho_i$ occurring in $B_{12}$, $\varrho_i$ will have exactly two nonzero entries in $B_{11}$.

The elimination of $B_{12}$ by this process will modify $B_{11}$ into $B'_{11}$. Clearly, each row of $B'_{11}$ will have exactly one or two nonzero entries. By (12)–(14), each column of $B_{11}$ has at least one nonzero entry. Since the elimination process may only double or create new entries in $B_{11}$, each column of $B'_{11}$ will also have at least one nonzero entry. Further, $B'_{11}$ has at least $\lfloor (k-1)/2 \rfloor - 1$ columns. By Lemma 10, $\operatorname{rank} B'_{11} \geq \lceil (k-3)/4 \rceil$. Also $\operatorname{rank} B'_{22} = (k-3)/2$, where $B'_{22}$ denotes the block obtained from $B_{22}$ by deleting the $p$th column. Since $B_{12}$ is deleted, we have

$$\operatorname{rank} \mathbf{M} \geq \operatorname{rank} B'_{11} + \operatorname{rank} B'_{22} \geq \left\lceil \frac{k-3}{4} \right\rceil + \frac{k-3}{2}.$$

In the same way, for $k$ even, we get

$$\operatorname{rank} \mathbf{M} \geq \left\lceil \frac{k-2}{4} \right\rceil + \frac{k-4}{2}.$$

In either case $\operatorname{rank} \mathbf{M} \geq \lceil (3k-10)/4 \rceil$. The claims about the heights of the relations are obvious.

**4.5.** *Consequences of the gap between $h_j$ and $h_{j+1}$.* For computational reasons it is convenient to verify the theorem for small $k$. The theorem is true for $k=2$, since its conditions are satisfied only when $f(x) = a(x^n \mp 1)$. For $k=3$, we have $f(x) = ax^n + bx^m + c$. Let $l = (m,n)$ and $f_0(x) = ax^{n/l} + bx^{m/l} + c$. By Lemma 2, the only possible cyclotomic factors of $f_0$ are $x-1$, $x+1$, and $\Phi_3$. Their multiplicity is at most 2. Hence, $|f_c| \leq 2(1+1+2)l$. This,

together with the assumption $|f_c| \geq (1-1/c_1)|f|$, gives $l \geq \frac{1}{8}(1-1/c_1)|f|$, thus implying case (1) of the theorem. Consequently, in the computation of various constants below, we assume that $k \geq 4$.

We now define $j_0$ and the constants $H_{j_0}, H_{j_0+1}, \ldots, H_{k-2}$.

*Consider first the case of nonreciprocal exponents.* Put $j_0 = \lfloor (k-1)/2 \rfloor$. By Lemma 11 and Hadamard's inequality, we have

$$h_{j_0} \leq (3^{1/2} k 3^k \lfloor k/2 \rfloor^k)^{j_0} \leq (3^{1/2} k 3^k (k/2)^k)^{(k-1)/2}.$$

Define

$$H_{j_0} = (3^{1/2} k 3^k (k/2)^k)^{(k-1)/2}. \tag{16}$$

Put $j' = k-1-j$ for $j_0 \leq j \leq k-3$, and define

$$H_{j+1} = 2\sqrt{3}\,\beta 3^k k^4 (4j')^{j'} H_j^3, \tag{17}$$

where $\beta = 1.0000000023$.

*Suppose now that $f$ has reciprocal exponents.* Put $j_0 = \lceil (3k-10)/4 \rceil$. Now, by Lemma 11 and Hadamard's inequality, we have

$$h_{j_0} \leq 5^{1/2} 3^{(k-3)/4} (k 3^k \lfloor k/2 \rfloor^k)^{j_0 - (k-1)/2}$$

for $k$ odd. For $k$ even, we have

$$h_{j_0} \leq 3^{(k-2)/4} (k 3^k \lfloor k/2 \rfloor^k)^{j_0 - (k-2)/2}.$$

In either case

$$h_{j_0} \leq 3^{(k-2)/4} (k 3^k \lfloor k/2 \rfloor^k)^{(k-3)/4}.$$

Define

$$H_{j_0} = 3^{(k-2)/4} (k 3^k (k/2)^k)^{(k-3)/4}. \tag{18}$$

For $j_0 \leq j \leq k-3$, define $H_{j+1}$ again by (17), but this time with $\beta = 1.00023$.

Suppose that

$$h_j \leq H_j, \quad \text{but} \quad h_{j+1} > H_{j+1}. \tag{19}$$

Select a lattice $\Gamma_j < \Lambda$ such that $\mathrm{vol}(\Gamma_j) = h_j$. By Lemma 3, there are linearly independent vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1-j}$ in $\mathbb{Z}^{k-1}$ satisfying

$$\prod_{i=1}^{k-1-j} \mathrm{h}(\mathbf{x}_i) \leq h_j \tag{20}$$

and forming a basis of $d\Gamma_j^{\perp}$, where $d$ is a suitable positive integer. Form a $(k-1-j) \times (k-1)$ matrix $\mathbf{M}$ by taking $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1-j}$ as its rows. By definition of $\Gamma_j$, $\mathbf{n} \in \Gamma_j^{\perp}$. Hence, there is a vector $\mathbf{r} \in \mathbb{Z}^{k-1-j}$ such that

$$d\mathbf{n} = \mathbf{r}\mathbf{M}. \tag{21}$$

Let $\mathbf{j}^{(i)}$, $i = 1, \dots, k-1$, be the columns of $\mathbf{M}$. Put

$$F(\mathbf{x}) = a_k + \sum_{i=1}^{k-1} a_i \mathbf{x}^{\mathbf{j}^{(i)}}, \tag{22}$$

where $\mathbf{x} = (x_1, \dots, x_{k-1-j})$ and $\mathbf{j}^{(i)} = (j_{1i}, \dots, j_{k-1-j,i})$. Then

$$F_{\mathbf{r}}(x) = f(x^d) = a_k + \sum_{i=1}^{k-1} a_i x^{dn_i}.$$

Factor $IF$ as follows: $IF = F_1 F_2$, $F_1, F_2 \in \mathbb{Z}[\mathbf{x}]$, $F_1$ is a product of all extended cyclotomic factors of $IF$, while $F_2$ is not divisible by any such factor. Let $\Delta = |F_{\mathbf{r}}|$ and $\Delta_2 = |I(F_{2,\mathbf{r}})|$. We shall show that $\Delta_2$ is much smaller than $\Delta$.

For this, put $n = j' = k - 1 - j$, $\mathbf{a} = \mathbf{r}$, and $B_i = (4+\varepsilon) j' \,\mathrm{h}(\mathbf{x}_i)$, where $\mathbf{x}_i$, $i = 1, \dots, j'$, are the rows of $\mathbf{M}$, and $\varepsilon > 0$, in Lemma 4. The value of $\varepsilon$ can be taken arbitrarily small and will be discussed later. We get

(i) $t\mathbf{r} = \mathbf{r}' + l\mathbf{c}$,
(ii) $|r_i'| \le ((4+\varepsilon) j' \,\mathrm{h}(\mathbf{x}_i))^{-1} l$,
(iii) $|c_i| \le B_i^{-1} + B_1 \cdots B_{j'} \le 1/8 + ((4+\varepsilon)j')^{j'} \prod_{i=1}^{j'} \mathrm{h}(\mathbf{x}_i)$
$\le 1/8 + ((4+\varepsilon)j')^{j'} h_j$,

where $t$ is a positive integer, $l = \mathrm{h}(\mathbf{r})$, and $\mathbf{c} \neq \mathbf{0}$. We claim that for some column $\mathbf{j}$ of $\mathbf{M}$ we must have

$$\mathbf{cj} \neq 0. \tag{23}$$

Suppose not; then by (21) and (i), we have

$$td\mathbf{n} = t\mathbf{r}\mathbf{M} = \mathbf{r}'\mathbf{M}.$$

Since the rows of $\mathbf{M}$ are linearly independent we have $t\mathbf{r} = \mathbf{r}'$. However, by (ii), $\mathrm{h}(\mathbf{r}') < \mathrm{h}(\mathbf{r})$. Since $t \ge 1$, this gives a contradiction.

By (21), we have $\Delta = d|f| = \mathbf{rj}$, where $\mathbf{j}$ is the first column of $\mathbf{M}$. Hence, by (i),

$$t\Delta = \mathbf{cj}l + \mathbf{r}'\mathbf{j}.$$

By the choice of the constants $B_i$ and by (ii),

$$\mathbf{r}'\mathbf{j} = \sum_{i=1}^{j'} r_i j_i < \sum_{i=1}^{j'} \mathrm{h}(\mathbf{x}_i) \frac{1}{4j' \,\mathrm{h}(\mathbf{x}_i)} = \frac{1}{4} l. \tag{24}$$

Further, by (20),

$$\mathbf{cj} \le \mathrm{h}(\mathbf{c}) h_j + j' - 1.$$

Hence,

$$t\Delta \le ((1/8 + ((4+\varepsilon)j')^{j'} H_j) H_j + j' - 1 + 1/4) \le \beta (4j')^{j'} H_j^2 l \tag{25}$$

with suitable $\beta$. Since $\varepsilon > 0$ can be arbitrarily chosen, we calculate $\beta$ by putting $\varepsilon = 0$ and rounding up the values obtained for $\beta$. This guarantees the existence of positive $\varepsilon$ satisfying (25). The resulting values of $\beta$ are given in (16) and (18).

Concerning $\Delta_2 = |I(F_{2,\mathbf{r}})|$, we note that

$$IF(\mathbf{x}) = \sum_{i=1}^{k} a_i \mathbf{x}^{\mathbf{j}^{(i)} - \mathbf{j}_F},$$

where $\mathbf{j}^{(k)} = \mathbf{0}$, so that

$$(\mathbf{j}^{(i)} - \mathbf{j}_F)_j \le 2\,\mathrm{h}(\mathbf{x}_j) \tag{26}$$

for any component of the exponents. The same must hold for the exponents of $F_2$, which is a polynomial factor of $IF$. If $F_2(\mathbf{x}) = \sum_{\mathbf{j}\in\mathcal{J}_{F_2}} b_{\mathbf{j}}\mathbf{x}^{\mathbf{j}}$ then $I(F_{2,t\mathbf{r}}) = \sum_{\mathbf{j}\in\mathcal{J}_{F_2}} b_{\mathbf{j}}\mathbf{x}^{t(\mathbf{j}-\mathbf{j}_0)\mathbf{r}}$, where $t\mathbf{j}_0\mathbf{r} = j_{F_{2,t\mathbf{r}}}$. Hence

$$t\Delta_2 = |I(F_{2,t\mathbf{r}})| = \max_{\mathbf{j}\in\mathcal{J}_{F_2}} (\mathbf{c}(\mathbf{j}-\mathbf{j}_0)l + (\mathbf{j}-\mathbf{j}_0)\mathbf{r}'). \tag{27}$$

We claim that

$$\mathbf{c}(\mathbf{j}-\mathbf{j}_0) = 0 \quad \text{for all } \mathbf{j} \in \mathcal{J}_{F_2}. \tag{28}$$

Suppose the contrary. Since $F_2 \in \mathbb{Z}[\mathbf{x}]$, neither $\mathbf{j}$ nor $\mathbf{j}_0$ has negative components, and again we have

$$(\mathbf{j}-\mathbf{j}_0)_i \le 2\,\mathrm{h}(\mathbf{x}_i).$$

Similarly to (24), this gives $(\mathbf{j}-\mathbf{j}_0)\mathbf{r}' < l/2$. Hence, by (27),

$$t\Delta_2 > l - l/2 = l/2. \tag{29}$$

From this and (25), we get

$$\Delta/\Delta_2 < 2\beta(4j')^{j'} H_j^2 \le 2\beta 64 H_{k-3}^2. \tag{30}$$

The last inequality is valid due to the formula (17). Define

$$c_1 = \beta 64 H_{k-3}^2. \tag{31}$$

Then $\Delta_2 > (2/c_1)\Delta$. Put $\Delta_{2,\mathrm{c}} = |(I(F_{2,\mathbf{r}}))_{\mathrm{c}}|$ and $\Delta_{\mathrm{c}} = |(I(F_{\mathbf{r}}))_{\mathrm{c}}|$. The condition $|f_{\mathrm{c}}| > (1-1/c_1)|f|$ of the theorem implies that

$$\Delta_{\mathrm{c}} = (\Delta - \Delta_2) + \Delta_{2,\mathrm{c}} > (1-1/c_1)\Delta.$$

Hence, $\Delta_{2,\mathrm{c}} > \Delta_2 - (1/c_1)\Delta > \frac{1}{2}\Delta_2$. Thus $F, F_1$, and $F_2$ satisfy the conditions of Lemma 8. We have $J = k$ and $P \le 3^k$. Therefore, there is a nonzero vector of the form $\mathbf{v} = a(\mathbf{j}^{(2)} - \mathbf{j}^{(3)}) - b\mathbf{j}^{(1)}$, where $\mathbf{j}^{(i)} \in \mathcal{J}_{\mathcal{F}}$, $i = 1, 2, 3$, and $a, b$ are integers with $\max\{|a|, |b|\} < PJ^4\Delta/\Delta_2 < 2\beta 3^k k^4 (4j')^{j'} H_j^2$, such that $\mathbf{v}\mathbf{r} = 0$.

On the other hand, we have

LEMMA 12. *Let* $f(x) = a_k + \sum_{i=1}^{k-1} a_i x^{n_i}$ *be a polynomial with* $k \geq 4$ *nonzero terms and let* $j$ *be a positive integer,* $1 \leq j \leq k-3$. *Further, let* $F$ *and* $\mathbf{r}$ *be defined by* (22) *and* (21). *If there is a nonzero vector* $\mathbf{v}$ *of the form* $\mathbf{v} = a(\mathbf{j}^{(i_2)} - \mathbf{j}^{(i_3)}) + b\mathbf{j}^{(i_1)}$ *or* $\mathbf{v} = a\mathbf{j}^{(i_2)} + b\mathbf{j}^{(i_1)}$, *where* $a$ *and* $b$ *are integers, and* $\mathbf{j}^{(i)}$, $1 \leq i \leq 3$, *are exponents of* $F$ *such that* $\mathbf{vr} = 0$, *then* $h_{j+1} \leq \sqrt{3}\max\{|a|,|b|\}h_j$.

*Proof.* Let $\widehat{\mathbf{v}}$ be the vector whose $i_1$th component is $b$, $i_2$th component is $a$, $i_3$th component is $a$ or 0, according to the form of $\mathbf{v}$, and all other components are 0. Then $\mathbf{v} = \mathbf{M}\widehat{\mathbf{v}}$ and

$$d\mathbf{n}\widehat{\mathbf{v}} = \mathbf{r}\mathbf{M}\widehat{\mathbf{v}} = 0. \tag{32}$$

Since the rows of $\mathbf{M}$ lie in $\Gamma_j^{\perp}$ and $\mathbf{M}\widehat{\mathbf{v}} = \mathbf{v} \neq \mathbf{0}$, we conclude that $\widehat{\mathbf{v}} \notin \Gamma_j$. Hence

$$h_{j+1} = \mathrm{vol}(\Gamma_{j+1}) \leq \mathrm{vol}(\mathrm{span}(\Gamma_j \cup \{\widehat{\mathbf{v}}\})) \leq \|\widehat{\mathbf{v}}\|\,\mathrm{vol}(\Gamma_j) \leq \sqrt{3}\max\{|a|,|b|\}h_j.$$

Returning to the proof of (28), by (19) and Lemma 12, we get

$$H_{j+1} < h_{j+1} \leq \sqrt{3}\max\{|a|,|b|\}h_j < 2\sqrt{3}\,\beta 3^k k^4 (4j')^{j'} H_j^3.$$

This contradicts the definition (17) of $H_{j+1}$.

Therefore, (28) holds. Since (23) is also true, we have

$$\mathbf{c}(\mathbf{j} - \mathbf{j}_0) = 0 \quad \text{for all } \mathbf{j} \in \mathcal{J}_{F_2}, \quad \text{but } \mathbf{cj} \neq 0 \quad \text{for some } \mathbf{j} \in \mathcal{J}_F.$$

This implies that $F_1$ has at least one extended cyclotomic factor $I\Phi_q(\mathbf{x}^{\mathbf{v}})$ for which $\mathbf{vc} \neq 0$. Let $\widetilde{F}_1$ be the product of all such factors of $F_1$, and let $\widetilde{F}_2 = IF/\widetilde{F}_1$. Thus, $\widetilde{F}_2$ is the product of $F_2$ and all factors $I\Phi_q(\mathbf{x}^{\mathbf{v}})$ of $F_1$ for which $\mathbf{vc} = 0$. Let $\widetilde{F}_1(\mathbf{x}) = \prod_i I\Phi_{q_i}(\mathbf{x}^{\mathbf{v}_i})$. We have

$$f(x^{td}) = F_{t\mathbf{r}}(x) = I(\widetilde{F}_{1,t\mathbf{r}}(x))I(\widetilde{F}_{2,t\mathbf{r}}(x)).$$

Let $\psi(\mathbf{x}) = \Phi_q(\mathbf{x}^{\mathbf{v}})$. Clearly, $I((I\psi)_{t\mathbf{r}}(x)) = I(\psi_{t\mathbf{r}}(x))$. Hence,

$$I(\widetilde{F}_{1,t\mathbf{r}}(x)) = \prod_i I(\Phi_{q_i}(x^{\mathbf{v}_i t\mathbf{r}})).$$

Put $l_i = |\mathbf{v}_i t\mathbf{r}|$. Then $I(\Phi_{q_i}(x^{\mathbf{v}_i t\mathbf{r}})) = \Phi_q(x^{l_i})$. Hence,

$$I(\widetilde{F}_{1,t\mathbf{r}}(x)) = \prod_i \Phi_{q_i}(x^{l_i}), \quad \text{where} \quad l_i = |\mathbf{v}_i t\mathbf{r}|.$$

By definition of $\widetilde{F}_1$, $\mathbf{v}_i\mathbf{c} \neq 0$. Also, since $\Phi_{q_i}(\mathbf{x}^{\mathbf{v}_i})$ are factors of $IF$, by (26) we have $(\mathbf{v})_i \leq 2\,\mathrm{h}(\mathbf{x}_i)$. Hence, $\mathbf{vr}' < l/2$. Similarly to (29) we get

$$l_i = |\mathbf{v}_i t\mathbf{r}| = |\mathbf{vc}l + \mathbf{v}_i\mathbf{r}'| > l - l/2 = l/2. \tag{33}$$

On the other hand,

$$|I(\widetilde{F}_{2,t\mathbf{r}}(x))| = \max_{\mathbf{j} \in \mathcal{J}_{\widetilde{F}_2}} (\mathbf{j} - \mathbf{j}_0)\mathbf{r}' \leq l/2.$$

Put $\widetilde{f}_2(x) = I(\widetilde{F}_{2,t\mathbf{r}}(x))$. Then

$$f(x^{td}) = \Big(\prod_i \Phi_{q_i}(x^{l_i})\Big)\widetilde{f}_2(x). \tag{34}$$

Thus, we have a decomposition of type (2) from Theorem 2, but with $f(x^{td})$ in place of $f$. Fortunately, this obstacle can be removed. We have

LEMMA 13. *Let $f \in \mathbb{Z}[x]$, $f(0) \neq 0$, and $d$ be a positive integer. If $f(x^d)$ has a decomposition*

$$f(x^d) = \Big(\prod_i \Phi_{q_i}(x^{l_i})\Big)\widetilde{f}_2(x),$$

*where $\widetilde{f}_2 \in \mathbb{Z}[x]$ and $l_i$ are positive integers such that $\min_i l_i > |\widetilde{f}_2|$, then $\widetilde{f}_2(x) = f_2(x^d)$ with $f_2 \in \mathbb{Z}[x]$, and*

$$f(x) = \Big(\prod_i \Phi_{\widetilde{q}_i}(x^{\widetilde{l}_i})\Big)f_2(x)$$

*with positive integers $\widetilde{l}_i$ such that $\min_i \widetilde{l}_i > |f_2|$. Moreover, $\min_i \widetilde{l}_i \geq (1/d)\min_i l_i$.*

*Proof.* First, note that if for some $q_i$, $p^t \,|\, q_i$, where $p$ is a prime and $t > 1$, then $\Phi_{q_i}(x^{l_i}) = \Phi_{q_i/p^{t-1}}(x^{l_i p^{t-1}})$. Hence, without loss of generality, we can assume that the numbers $q_i$ are squarefree.

The condition $\min_i l_i > |\widetilde{f}_2|$ implies that $\pm\widetilde{f}_2$ occurs in the expression of $f(x^d)$. Consequently, $\widetilde{f}_2(x) = f_2(x^d)$ with suitable $f_2 \in \mathbb{Z}[x]$. Divide both sides of the equation

$$f(x^d) = \Big(\prod_i \Phi_{q_i}(x^{l_i})\Big)f_2(x^d)$$

by $f_2(x^d)$ and by the product of those of $\Phi_{q_i}(x^{l_i})$ for which $d \,|\, q_i$. We get

$$g(x^d) = \prod_{i:\, d\nmid q_i} \Phi_{q_i}(x^{l_i}) \tag{35}$$

with suitable $g \in \mathbb{Z}[x]$. Let $\Omega(d)$ be the number of prime factors of $d$ counted with multiplicities. We now show by induction on $\Omega(d)$ that by deleting some of the factors on the right-hand side of (35) and replacing $l_i$ by their suitable multiples $\widetilde{l}_i$, we can get

$$g(x^d) = \prod_{i:\, d\nmid q_i}{}^{*} \Phi_{q_i}(x^{\widetilde{l}_i}),$$

where $d \,|\, \widetilde{l}_i$ for all $i$. The asterisk indicates that the set of indices in this product might be a proper subset of the set of indices in (35).

If $\Omega(d) = 0$ there is nothing to prove. Suppose that $\Omega(d) > 0$ and let a prime $p$ divide $d$. If $p \,|\, l_i$ in all the exponents in (35) then by dividing $d$ and

all $l_i$ by $p$ we are reduced to the case of $\Omega(d/p) = \Omega(d) - 1$. The statement is then true by induction hypothesis. Otherwise, divide both sides of (35) by those of the factors $\Phi_{q_i}(x^{l_i})$ in which $p \,|\, q_i$. We get

$$h(x^p) = \prod_{i:\, p \nmid q_i} \Phi_{q_i}(x^{l_i}) \tag{36}$$

with a suitable $h \in \mathbb{Z}[x]$. In order to investigate this equation, factor both sides into products of irreducible polynomials. Let $h(x) = \prod_i \Phi_{m_i}(x)$. If $p \,|\, m_i$ then $\Phi_{m_i}(x^p) = \Phi_{pm_i}(x)$, so that $p^2 \,|\, pm_i$; if $p \nmid m_i$ then $\Phi_{m_i}(x^p) = \Phi_{m_i}(x)\Phi_{pm_i}(x)$. To factor $\Phi_{q_i}(x^{l_i})$, let $l_i = l_{i1}l_{i2}$, where each prime factor of $l_{i1}$ divides $q_i$, while $(l_{i2}, q_i) = 1$. Then

$$\Phi_{q_i}(x^{l_i}) = \prod_{\delta \mid l_{i2}} \Phi_{\delta l_{i1} q_i}(x).$$

Since $p \nmid l_i$ and $q_i$ is squarefree, $p^2$ does not divide any of the indices of polynomials $\Phi_{\delta l_{i1} q_i}(x)$ in this product. Consequently, $p \nmid m_i$ for every $m_i$. Hence,

$$h(x^p) = \prod_i (\Phi_{m_i}\Phi_{pm_i}(x)) = \prod_i \prod_{\delta \mid l_{i2}} \Phi_{\delta l_{i1} q_i}(x).$$

By comparing both sides of this equation and because of uniqueness of factorization, we deduce that each factor $\Phi_{\delta l_{i1} q_i}(x)$ with $p \nmid q_i$ in the product on the right-hand side can be matched with a factor $\Phi_{\delta l_{i1} q_{i'}}(x)$, where $q_{i'} = pq_i$, also occurring in this product. Thus (36) can be written as

$$h(x^p) = \prod_{i:\, p \nmid q_i}^{*} \Phi_{q_i}(x^{pl_i}).$$

Again, the asterisk indicates that the set of indices in this product is not the same as in (36). This proves the induction step. The inequality $\min_i \widetilde{l_i} \geq (1/d) \min_i l_i$ follows by the construction of $\widetilde{l_i}$.

Of course, we apply this lemma to (34) with $td$ in place of $d$. It remains to prove that

$$\min_i l_i > \frac{1}{2c_1} |f|.$$

For this, by combining (33) and (25), we get

$$tl_i > \frac{1}{2}\, l \geq \frac{t\Delta}{\beta(4j')^{j'} H_j^2} \geq \frac{t\Delta}{2c_1}.$$

Here $\Delta$ and $l_i$ refer to $f(x^d)$. By Lemma 13, the bound obtained carries over to $f(x)$.

**4.6.** *Computation of the constants.* Recall that we assume $k \geq 4$. By (31), we have $c_1 = \beta 64 H_{k-3}^2$, and by (5), $c_2 = H_{k-2}$. By (17),

$$H_{j+1} = 2\sqrt{3}\,\beta 3^k k^4 (4j')^{j'} H_j^3 \leq 2\sqrt{3}\,\beta 3^k k^4 (4(k-1-j_0))^{k-1-j_0} H_j^3.$$

Put $B = 2\sqrt{3}\,\beta 3^k k^4 (4(k-1-j_0))^{k-1-j_0}$. Then

$$H_{j_0+m} \leq B^{(3^m-1)/2} H_{j_0}^{3^m}. \tag{37}$$

We need to consider two cases separately:

*The case of reciprocal exponents.* We have $j_0 = \lceil (3k-10)/4 \rceil$. Put $m = k-3-j_0 = \lfloor (k-2)/4 \rfloor$. By (18), $H_{j_0} = 3^{(k-2)/4}(k3^k(k/2)^k)^{(k-3)/4}$. Together with (37) this gives

$$H_{k-3} = \exp(3^{\lfloor (k-2)/4 \rfloor} c(k) k^2 \log k),$$

where $c(k)$ is defined by the equality and approaches $1/4$ as $k$ tends to infinity. Hence,

$$c_1 = \beta 64 H_{k-3}^2 = \exp(3^{\lfloor (k-2)/4 \rfloor} a(k) k^2 \log k),$$

where $a(k)$ is defined by the equality and approaches $1/2$ as $k$ tends to infinity. Now we find a decreasing function $b(k)$ with the same limit value $1/2$, and such that $b(k) \geq a(k)$. By a somewhat tedious, but simple calculation we find that

$$b(k) = 0.5 + \frac{0.203}{\log k} + \frac{1.27}{k \log k} + \frac{3.25}{k^2} + \frac{1.53}{k^2 \log k} + \frac{\log 64.1}{3^{\lfloor (k-2)/4 \rfloor} k^2 \log k}$$

satisfies these conditions. We find that $b(14) < 0.632$. On the other hand, the maximum value of $a(k)$ on the interval $[4, 14]$ is approximately 0.6359. This proves that

$$c_1 = \exp(3^{\lfloor (k-2)/4 \rfloor} 0.636(k) k^2 \log k)$$

for $k \geq 4$. Finally, we check that this estimate is also valid for smaller values of $k$. In a similar way we find that

$$c_2 = H_{k-2} = 2\sqrt{3}\,\beta 3^k k^4 (64) H_{k-3}^3 \leq \exp(3^{\lfloor (k-2)/4 \rfloor} 1.06 k^2 \log k).$$

*The case of nonreciprocal exponents.* Now, we have $j_0 = \lfloor (k-1)/2 \rfloor$, $m = k-1-j_0 = \lceil (k-1)/2 \rceil$, and $H_{j_0} = (3^{1/2} k 3^k (k/2)^k)^{(k-1)/2}$, defined by (16). Similarly to the previous case, we get

$$c_1 = \beta 64 H_{k-3}^2 \leq \exp(3^{\lceil (k-2)/2 \rceil} 2.84 k^2 \log k),$$
$$c_2 = H_{k-2} \leq \exp(3^{\lceil (k-2)/2 \rceil} 2.841 k^2 \log k).$$

To prove Lemma 1, we will need yet another version of Lemma 9 from [4].

LEMMA 14. *Let $\mathbf{a} \in \mathbb{Z}^n$ be a vector, $B > 1$ a real number, $T = n! B^n$, and $q > T$ a rational integer. Then there are vectors $\mathbf{c}, \mathbf{r} \in \mathbb{Z}^n$ and $t \in \mathbb{Z}$ such that*

(1) $1 \le t \le T$,
(2) $t\mathbf{a} = \mathbf{r} + q\mathbf{c}$,
(3) $l(\mathbf{r}) \le qB^{-1}$,
(4) *for* $\mathbf{a} \ne 0$ *and* $q = l(\mathbf{a})$ *we also have* $0 < l(\mathbf{c}) \le B^{-1} + T$.

*Proof.* Let

$$C = \Big\{(\tau, x_1, \dots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; |\tau| \le T,\ \sum_{i=1}^{n} |\tau a_i/q - x_i| \le B^{-1}\Big\}.$$

The set $C$ is closed, convex, symmetric, and has volume $(2T)(2^n B^{-n}/n!) = 2^{n+1}$. By Minkowski's convex body theorem, $C$ contains a nonzero vector $(t, c_1, \dots, c_n) \in \mathbb{Z}^{n+1}$. Since $C$ is symmetric, we can assume that $t \ge 0$. Then $B > 1$ implies that $t \ge 1$; this proves (1). Set $\mathbf{r} = t\mathbf{a} - q\mathbf{c}$. Then (2) and (3) hold trivially. Further, if $\mathbf{a} \ne 0$ and $q = l(\mathbf{a})$ then (3) gives $l(\mathbf{r}) \le l(\mathbf{a})B^{-1} < l(\mathbf{a}) \le tl(\mathbf{a})$. From this, and by (2), $0 < l(\mathbf{c})$. Finally, $\sum_{i=1}^{n} |c_i| \le B^{-1} + \sum_{i=1}^{n} t|a_i|/l(\mathbf{a}) < B^{-1} + t \le B^{-1} + T$ shows (4).

Following [4], we deduce Lemma 1 from an analogous result for an algebraic integer rather than for a polynomial $g$:

LEMMA 15. *If* $\alpha$ *is a nonzero algebraic integer, not a root of unity, with* $\deg(\alpha) = n$, *and* $f$ *is a polynomial with integer coefficients that has* $k$ *nonzero terms and* $f(\alpha) = 0$ *then*

$$M(\alpha) \ge 1 + \frac{0.31n}{k!|f|}.$$

*Proof of Lemma 1.* We show that Lemma 15 immediately implies Lemma 1. This argument is given in [4] (Lemma 7 implies Lemma 8 there). We repeat it for the convenience of the reader.

Suppose that the conditions of Lemma 1 are satisfied and that $\gamma_i$ is the multiplicity of $\alpha_i$ in $g$, where $\alpha_1, \dots, \alpha_m$ are zeros of $g$ representing all classes of conjugate zeros of $g$. We have

$$\begin{aligned} M(g) = \prod_{i=1}^{m} M(\alpha_i)^{\gamma_i} &\ge \prod_{i=1}^{m} \left(1 + \frac{0.31 \deg \alpha_i}{k!|f|}\right)^{\gamma_i} \\ &\ge 1 + \frac{0.31 \sum_{i=1}^{m} \gamma_i \deg \alpha_i}{k!|f|} = 1 + \frac{0.31|g|}{k!|f|}. \end{aligned}$$

*Proof of Lemma 15.* Let $f(x) = \sum_{i=1}^{k} a_i x^{n_i}$. If $\deg \alpha \le 2$ then $M(\alpha) \ge 2$. The lemma is obviously true in these cases. Assume that $\deg \alpha \ge 3$. If $\alpha$ is not reciprocal then $M(\alpha) \ge \theta$. The lemma is then true, because we must have $k \ge 3$, $|f| \ge \deg \alpha$, and $\theta > 1 + 0.31/3!$. Consequently, in what follows, we assume that $\alpha$ is reciprocal.

We proceed by induction on $k$. The case of $k = 1$ is vacuous. For $k = 2$, the conditions of the lemma are satisfied only when $\alpha$ is not a

unit; then $M(\alpha) \geq 2$. For $k = 3$ and $f$ reciprocal, we easily find that $M(\alpha) \geq (3+\sqrt{5})/2$. The same holds for $k = 3$ and nonreciprocal $f$. To see this, let $f(x) = ax^n + bx^m + c$. We have $a\alpha^n + b\alpha^m + c = 0$. Hence, also $c\alpha^n + b\alpha^{n-m} + a = 0$, because $\alpha$ is reciprocal. By eliminating $\alpha^n$ from these equations, we get $bc\alpha^m - ab\alpha^{n-m} + c^2 - a^2 = 0$. The left-hand side of this equation is not identically 0, because $f$ was not reciprocal. Since $\alpha$ is not a root of unity, none of the three coefficients in this equation can be 0. Thus, $\alpha$ is a zero of a trinomial which has lower degree than $f$. Clearly, we can continue that process until we obtain a reciprocal trinomial vanishing at $\alpha$. Hence, again $M(\alpha) \geq (3+\sqrt{5})/2$.

Let now $k \geq 4$, and suppose that the lemma is true for all $k' < k$. Further, let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ be the conjugates of $\alpha$. Without loss of generality, at each step of the induction process we can assume the following:

(1) $f(0) \neq 0$,
(2) $\gcd(n_1, \ldots, n_k) = 1$,
(3) $l(f) \leq k!B^k + B^{-1}$, where $B = 1 + 0.31n/k!|f|$.

For (1), replace $f$ by $If$ if necessary. Then $If(0) \neq 0$ and $|If| \leq |f|$, so that the bound improves.

For (2), suppose that $\gcd(n_1, \ldots, n_k) = q$. Then $f(x) = f_0(x^q)$, $f_0$ satisfies (2), and $|f_0| = |f|/q$. Replace $f$ by $f_0$, and $\alpha$ by $\alpha^q$, and suppose that the lemma holds in the new situation. Let $\deg \alpha^q = m$. Then each conjugate of $\alpha^q$ occurs $n/m$ times among the numbers $\alpha_1^q, \ldots, \alpha_n^q$, and we have

$$M(\alpha) = M(\alpha^q)^{n/mq} \geq \left(1 + \frac{0.31mq}{k!|f|}\right)^{n/mq} \geq 1 + \frac{0.31n}{k!|f|}.$$

For (3), suppose that $l(f) > k!B^k + B^{-1}$. Apply Lemma 14 by taking as $\mathbf{a}$ the vector of coefficients of $f$, $q = l(\mathbf{a})$, and $B$ as above. We get $t\mathbf{a} = \mathbf{r} + q\mathbf{c}$, $1 \leq t \leq T$, $l(\mathbf{r}) \leq qB^{-1}$, and $0 < l(\mathbf{c}) \leq B^{-1} + T$, where $T = k!B^k$. The equation for the corresponding polynomials is $tf(x) = r(x) + qc(x)$. At $x = \alpha$,

$$r(\alpha) + qc(\alpha) = 0.$$

If $c(\alpha) = 0$ then we can replace $f$ by $c$, since $l(\mathbf{c})$ satisfies (3). If $c(\alpha) \neq 0$ then also $r(\alpha) \neq 0$, and $r(\alpha) = -qc(\alpha)$. Hence, we get

$$l(\mathbf{r})^n M(\alpha)^{|f|} \geq \Big|\prod_{i=1}^{n} r(\alpha_i)\Big| = \Big|\prod_{i=1}^{n} qc(\alpha_i)\Big| \geq q^n.$$

Since $l(\mathbf{r}) \leq qB^{-1}$, this gives $M(\alpha) > B = 1 + 0.31n/k!|f|$.

Consequently, in what follows we assume that (1)–(3) hold. Let $p$ be a prime such that

$$2(B^{-1} + k!B^k) \leq p < 4(B^{-1} + k!B^k).$$

Since $f(\alpha)=0$, we have $p \mid f(\alpha^p)$. Two cases may occur.

CASE 1: $f(\alpha^p) \neq 0$. Since (3) holds, the standard argument works:

$$l(f)^n M(\alpha)^{p|f|} \geq \Big| \prod_{i=1}^{n} f(\alpha_i^p) \Big| \geq p^n.$$

Hence, by the choice of $p$, we get

$$m(\alpha) = \log M(\alpha) \geq \frac{n(\log p - \log l(f))}{|f|p} \geq \frac{n \log 2}{2|f|(B^{-1}+k!B^k)},$$

where $B = 1 + 0.31n/k!|f|$. For $k \geq 4$, this gives $m(\alpha) \geq 0.31n/k!|f|$.

CASE 2: $f(\alpha^p)=0$. Let $g$ and $g_p$ be the minimal polynomials of $\alpha$ and $\alpha^p$, respectively. Then $g_p \neq g$, since $\alpha$ is not a root of unity. The polynomial $f$ factors as $f = gg_ph$. Hence, $f' = g'g_ph + gg_p'h + gg_ph'$. Since $g_p(\alpha) \equiv 0 \bmod p$, we get

$$f'(\alpha) = f'(\alpha) f_p(\alpha) h(\alpha) \equiv 0 \bmod p.$$

In our notation, $f'(x) = \sum_{i=1}^{k-1} n_i a_i x^{n_i - 1}$. Put $\mathbf{a}' = (n_1 a_1, \ldots, n_{k-1} a_{k-1})$, $q = p$, and $B_1 = k^{1/(k-1)}$ in Lemma 14. We get $t\mathbf{a}' = \mathbf{r} + p\mathbf{c}$, $l(\mathbf{r}) \leq B_1^{-1} p$, and $t \leq (k-1)! B_1^{k-1}$. For the corresponding polynomials we get

$$tf'(\alpha) = r(\alpha) + pc(\alpha) \equiv 0 \bmod p.$$

Hence,

$$r(\alpha) \equiv 0 \bmod p.$$

This relation can be useful only if $r(x)$ is not identically 0. We shall show that it is indeed so. Suppose to the contrary that $r(x) \equiv 0$. The relation $t\mathbf{a}' = \mathbf{r} + p\mathbf{c}$ implies that $p \mid t\mathbf{a}'$. By our choice of $p, B$ and $B_1$,

$$p \geq 2(B^{-1} + k!B^k) > (k-1)! B_1^{k-1} \geq t.$$

Thus $p \nmid t$. Similarly, by (3), $p \geq 2(B^{-1} + k!B^k) > l(\mathbf{a})$, so that $p \nmid a_i$ for $1 \leq i \leq k-1$. Also, by (2), $p \nmid (n_1, \ldots, n_{k-1})$. Consequently, $p \nmid n_i a_i$ for some $i$, a contradiction with $p \mid t\mathbf{a}'$.

Therefore, $r(x) \not\equiv 0$. Again two cases are possible.

CASE 2.1: $r(\alpha) = 0$. It suffices to notice that $\mathbf{r}$ was obtained from $f'$, so that $r(x)$ has at most $k-1$ terms. The lemma holds by induction hypothesis.

CASE 2.2: $r(\alpha) \neq 0$. Now, we use the fact that $r(\alpha) \equiv 0 \bmod p$. We have

$$(l(\mathbf{r}) M(\alpha))^{|r|} \geq \Big| \prod_{i=1}^{n} r(\alpha) \Big| \geq p^n.$$

Since $l(\mathbf{r}) \leq B_1^{-1} p$ and $|r| < |f|$, this gives

$$M(\alpha) \geq B_1^{n/|f|} \geq 1 + \frac{n \log k}{k-1} > 1 + \frac{0.31n}{k!|f|}.$$

**5. Proof of Proposition 1.** Without loss of generality, we can assume that $f$ is reciprocal. Obviously, we can also assume that all zeros of $f$ that are not roots of unity have degree at least 3. The polynomial $f$ has the form

$$f(x) = x^{n+m} + \varepsilon\eta a x^n + \varepsilon a x^m + \eta,$$

where $n, m$, and $a$ are positive integers, $n > m$, $\varepsilon = \pm 1$, and $\eta = \pm 1$. We can also assume that $(m, n) = 1$. If $a = 1$ then $f(x) = (x^n \pm 1)(x^m \pm 1)$ is a product of cyclotomic polynomials. Hence, $a \geq 2$. If $f(\alpha) = 0$ then

$$\alpha^{n+m} + \eta = -\varepsilon\eta a(x^{n-m} + \eta). \tag{38}$$

Hence,

$$a \,|\, (\alpha^{n+m} + \eta). \tag{39}$$

*Consider the case of* $a \geq 3$. First, we determine which roots of unity can be among zeros of $f$. For this, suppose that $\xi$ is a root of unity and $f(\xi) = 0$. Since $a \geq 3$, (39) implies that $\xi^{n+m} + \eta = 0$. Hence, by (38), $\xi^{n-m} + \eta = 0$ as well. Since $(m, n) = 1$, both equations together imply that $\xi = \pm 1$. The multiplicity of $\xi$ is at most 3. Hence, $|f_{\mathrm{c}}| \leq 6$. By (38), (39), and because $f$ has at least one zero of degree at least 3 which is not a root of unity,

$$2^{m+n-|f_{\mathrm{c}}|} M(f)^{m+n} \geq \prod_{\alpha} |(\alpha^{m+n} + \eta)| \geq a^{\max\{3, m+n-|f_{\mathrm{c}}|\}}, \tag{40}$$

where the product runs over all zeros of $f$ that are not roots of unity. Hence,

$$M(f) \geq \max\{(a/2)^{3/|f|}, (a/2)^{1-6/|f|}\} \geq (a/2)^{1/3}.$$

For $a \geq 5$, this gives $M(f) > \theta$. On the other hand, the inequality $M(f) \geq (a/2)^{1-6/|f|}$ implies that $M(f) > \theta$: for $a = 4$ if $|f| \geq 11$, and for $a = 3$ if $|f| \geq 20$. This leaves out a finite number of polynomials for which the statement of the proposition was checked by direct computation of $M(f)$.

*Consider now the case of* $a = 2$. We need a slightly better use of (38). We have

$$2 \,|\, (\alpha^{m+n} + \eta) \;\Rightarrow\; 2 \,|\, (\alpha^{m+n} - \eta),$$

so that

$$4 \,|\, (\alpha^{2|f|} - 1). \tag{41}$$

By Lemma 2, the possible cyclotomic zeros of $f$ are $\pm 1$, $\zeta_3$, and $\zeta_6$. Consider their multiplicity. If $\zeta$ is a root of unity and $f(\zeta) = f'(\zeta) = 0$ then

$$(n+m)\zeta^n + \varepsilon\eta n a \zeta^{n-m} + \varepsilon m a = 0.$$

This is possible only if $n + m = 2(n - m)$. Since $(n, m) = 1$, we must have $m = 1$ and $n = 3$. Hence,

$$f(x) = x^3 + 2\varepsilon\eta x^2 + 2\varepsilon x + \eta.$$

For such polynomials $M(f) > 2.6$. Consequently, we can assume that the roots are single. Suppose now that $\zeta = \zeta_3$ or $\zeta = \zeta_6$. By (38), $2 \mid (\zeta^{n+m} + \eta)$. However, we cannot have $\zeta^{n+m} + \eta = 0$, since this would imply that $3 \mid (m, n)$. Hence, $\zeta^{n+m} + \eta = 2\eta$. Now, again by (38), we get

$$\zeta^n + \eta\zeta^m + \varepsilon = 0. \tag{42}$$

This is only possible if $3 \nmid nm$. Suppose that $\zeta = \zeta_3$. By taking traces in (42), we get $-1 - \eta + 2\varepsilon = 0$. Hence, $\varepsilon = \eta = 1$. Suppose now that $\zeta = \zeta_6$. Since $(m, n) = 1$, $m$ and $n$ cannot be both even. Now, by taking traces we get $\pm 1 \pm \eta + 2\varepsilon = 0$, where the combination of $\pm$ signs does not allow $-1-\eta+2\varepsilon = 0$. Hence, we cannot have $\varepsilon = \eta = 1$ in this case. Consequently, $\zeta_3$ or $\zeta_6$ is not a zero of $f$. Hence, $|f_c| \le 4$, and (41) gives

$$2^{|f|-|f_c|} M(f)^{2|f|} \ge \prod_{\alpha} |\alpha^{2|f|} - 1| \ge 4^{|f|-|f_c|},$$

where the product runs over the zeros of $f$ that are not roots of unity. Thus, $M(f) \ge 2^{(1-4/|f|)/2}$. For $|f| \ge 22$, this gives $M(f) > \theta$. Again, this leaves out a finite number of polynomials for which the statement of the proposition was checked by direct calculation of $M(f)$.

**6. Proof of Proposition 2.** Without loss of generality we can assume that $f$ is reciprocal. The improvement of the bound is due to the fact that the vector of coefficients of a reciprocal polynomial is symmetric, so it suffices to work with half of them only. Let $f(x) = \sum_{i=1}^{k} a_i x^{n_i}$. If $k = 2m - 1$ is odd then $l(f) = |a_m| + 2\sum_{i=1}^{m-1} |a_i|$, if $k = 2m$ is even then $l(f) = 2\sum_{i=1}^{m} |a_i|$. Put $B = 2(1 + 0.17/2^m m!)$, $n = m$, $\mathbf{a} = (2a_1, \dots, 2a_{m-1}, a_m)$ or $\mathbf{a} = (2a_1, \dots, 2a_{m-1}, 2a_m)$, according to the parity of $k$, and $q = l(\mathbf{a})$, in Lemma 14. Then $T = m!B^m$ and $t\mathbf{a} = \mathbf{r} + l\mathbf{c}$.

Suppose first that $\mathbf{r} = \mathbf{0}$. Since $f$ is monic, the content of $\mathbf{a}$ is either 1 or 2. Hence, $\mathbf{a} = \mathbf{c}_1$ or $\mathbf{a} = 2\mathbf{c}_1$, and $\mathbf{c}$ is a multiple of $\mathbf{c}_1$. In either case $l(f) = l(\mathbf{a}) \le 2(B^{-1} + T)$. Following [3], choose a prime $p$ such that $T \le p \le 2T$. We have

$$l(f)^{|f|} M(f)^{p|f|} \ge \Big| \prod_{\alpha:\, f(\alpha)=0} f(\alpha^p) \Big| \ge p^{|f|}.$$

The product does not vanish because $f$ is irreducible and not cyclotomic. From this, we get

$$M(f) \ge 1 + \frac{\log 2}{4(B^{-1} + T)} \ge 1 + \frac{0.17}{2^m m!},$$

where the last inequality is valid for $m \ge 3$. This can be assumed due to Proposition 1.

Now, suppose that $\mathbf{r} \ne \mathbf{0}$. Let $f_r(x) = \sum_{i=1}^{m} r_i(x^{n_i} + x^{n_{k-i+1}})$ or $f_r(x) =$

$2r_m x^{n_m} + \sum_{i=1}^{m-1} r_i(x^{n_i} + x^{n_{k-i+1}})$, for $k$ even or odd, respectively. Define in an analogous way a polynomial $f_c$, with respect to $\mathbf{c}$. Then $2f(x) = f_r(x) + lf_c(x)$, and $l(f_r) = 2l(\mathbf{r}) \le 2B^{-1}l < l(\mathbf{a}) = l(f)$. Hence, $f$ does not divide $f_r$, and $f_r(\alpha) = -lf_c(\alpha)$ if $\alpha$ is a zero of $f$. Therefore,

$$(2l(\mathbf{r})M(f))^{|f|} \ge \Big| \prod_{\alpha: f(\alpha)=0} f_r(\alpha) \Big| \ge l^{|f|}.$$

This gives $m(f) \ge \frac{1}{2}B = 1 + 0.17/2^m m!$.

**7. Final remarks.** The proof of Proposition 2 shows that the smallest measure of a reciprocal quadrinomial that is not a product of cyclotomic factors is larger than the smallest measure of a nonreciprocal quadrinomial. Numerical evidence suggests that the former occurs for $f(x) = x^7 - 2x^5 - 2x^2 + 1$, $M(f) \cong 1.55603$, while the latter occurs for $f(x) = x^4 - x^3 - x^2 + 1 = (x-1)(x^3 - x - 1)$, $M(f) = \theta$.

In view of Theorem 1, two questions seem to be interesting:

QUESTION 1. Let $\alpha$ be an algebraic integer and suppose that $f(\alpha) = 0$, $f \in \mathbb{Z}[x]$. What is the minimum number of nonzero coefficients of $f$?

QUESTION 2. What is the minimum number of nonzero coefficients of $f$ that satisfy an extra condition $M(f) = M(\alpha)$?

Concerning Question 1, it was kindly pointed to the author by A. Schinzel that already M. Schacher and E. G. Straus [9] noticed, that when $\alpha$ has many real conjugates then the Descartes rule of signs implies that the number of terms in $f$ cannot be small. However, in the most interesting case, i.e., when $M(\alpha)$ is small, it is known that $\alpha$ cannot have many real conjugates.

Concerning Question 2, suppose that a polynomial $f$ with $k$ terms is optimal. If $f(x) \ne f_0(x^l)$ then Corollary 2 immediately provides a bound $|f_c| \le (c_1 - 1)\deg\alpha$. If $f(x) = f_0(x^l)$ then $M(\alpha) = M(\alpha^l)$ and $l$ divides $\deg\alpha$. By applying Corollary 2 to $f_0$ we obtain the same bound again. Thus the number of possible cyclotomic factors of $f$ is finite. This provides an algorithm allowing one to answer Question 2 for a given $\alpha$. Unfortunately, the bounds involved are still too large for practical use.

## References


[1] E. Bombieri and J. Vaaler, *On Siegel's Lemma*, Invent. Math. 73 (1983), 11–32; Addendum, ibid. 75 (1984), 377.

[2] E. Dobrowolski, *Mahler's measure of a polynomial in function of the number of its coefficients*, Canad. Math. Bull. 34 (1991), 186–195.

[3] —, *On a question of Lehmer* , Mém. Soc. Math. France (N.S.) 1980/81, no. 2, 35–39.

[4] E. Dobrowolski, W. Lawton and A. Schinzel, *On a problem of Lehmer*, in: Studies in Pure Mathematics, Birkhäuser, Basel, 1983, 135–144.

[5] G. Hajós, *Solution of problem 41*, Mat. Lapok 4 (1953), 40–41.
[6] D. Hansen, *On the product of primes*, Canad. Math. Bull. 15 (1972), 33–37.
[7] H. B. Mann, *On linear relations between roots of unity*, Mathematika 12 (1965), 107–117.
[8] H. L. Montgomery and A. Schinzel, *Some arithmetic properties of polynomials in several variables*, in: Transcendence Theory: Advances and Applications, Academic Press, London, 1977, 195–203.
[9] M. Schacher and E. G. Straus, *Some applications of a non-Archimedean analogue of Descartes' rule of signs*, Acta Arith. 25 (1974), 353–357.
[10] 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.
[11] P. Voutier, *An effective lower bound for the height of algebraic numbers*, Acta Arith. 74 (1996), 81– 95.

College of New Caledonia
Prince George, BC
Canada
E-mail: dobrowolski@cnc.bc.ca

*Received on 13.8.2004*
*and in revised form on 16.12.2005* (4816)