A SHORT NOTE ON THE MAHLER MEASURE AND LEHMER’S CONJECTURE

ALI H. HAKAMI

Department of Mathematics
Jazan University
P. O. Box 277
Postal Code: 45142
Saudi Arabia
e-mail: aalhakami@jazanu.edu.sa

Abstract

Let \( k \) be an algebraic number field and \( F(x) \) be a polynomial in \( k[x] \). In this short paper, we shall consider the problem of the equivalence between best possible upper bounds for the number of non-cyclotomic factors of a polynomial in \( k[x] \) and optimal lower bounds for the Mahler measure of an irreducible, non-cyclotomic polynomials in \( k[x] \).

1. Introduction and the Main Results

Mahler’s measure is a function defined on polynomials, which measures the extent to which their roots are distributed away from the unit circle. It is known to be a continuous function on polynomials with complex coefficients, however when restricted to polynomials with integer coefficients it is expected to have “gaps” in its values. This has

2010 Mathematics Subject Classification: Primary 11R18, 11R60; Secondary 10C08, 11R09.
Keywords and phrases: non-cyclotomic polynomials, heights, Mahler measure, Lehmer’s problem.
Received February 5, 2017
been conjectured by D. H. Lehmer in 1933, and is to this day one of the famous open problems in Number Theory. Mahler’s measure and Lehmer’s conjecture have fundamental connections and applications within Number Theory as well as in other areas of mathematics, for instance ergodic theory. Let $k$ be an algebraic number field and $F(x)$ be a polynomial in $k[x]$ with degree $\partial(F)$ and $F(0) \neq 0$. We shall consider here the problem of the equivalence between best possible upper bounds for the number of non-cyclotomic factors of a polynomial in $k[x]$ and optimal lower bounds for the Mahler measure (The Mahler measure $\mu(F)$ will be defined precisely in Section 2) of an irreducible, non-cyclotomic polynomial in $k[x]$. In the special case $k = \mathbb{Q}$ both question have been previously studied by Schinzel [14, 15] and by Dobrowolski [4]. Of course, we always have the trivial bound $\partial(F)$ for such quantities and in generality this is plainly all we can say. However if the “height” $H(F)$ or the number of non-zero coefficients $N(F)$ of $F(x)$ is “small” we should expect to do a bit better. Here $H(F)$ is the absolute height on the vector of coefficients of $F(x)$, to be defined precisely in Section 2.

In fact it will often be convenient to use the different height $\nu(F)$ defined in Section 2. For a polynomial $F(x)$ in $\mathbb{Z}[x]$, $\nu(F)$ is simply the supermom of $|F(x)|$ on the unit circle, after again removing any common factors from the coefficients. In view of the inequality

$$\log H(F) \leq \log \nu(F) \leq 2 \log H(F),$$

(see Lemma 2.3) these two heights will be virtually interchangeable in most of our results.

Let $\Phi_n(x)$ in $\mathbb{Z}[x]$ denote the $n$-th cyclotomic polynomial and assume that $\Phi_n$ factors in $k[x]$ as

$$\Phi_n(x) = \prod_{s=1}^{\delta(k,n)} \Phi_{n,s}(x).$$  

(1.2)
Here we suppose that each factor $\Phi_{n,s}$ is monic and irreducible in $k[x]$. If $\zeta_n$ is a primitive $n$-th root of unity, then each factor $\Phi_{n,s}$ has degree $|k(\zeta_n) : k|$. As noted in [10], Equation (1.2), the number of distinct irreducible factors of $\Phi_n$ in $k[x]$ is

$$\delta(k; n) = [k \cap \mathbb{Q}(\zeta_n) : \mathbb{Q}] \leq [k' : \mathbb{Q}], \quad (1.3)$$

where $k' \subseteq k$ is the maximum abelian subfield of $k$. Indeed, $\delta(k'; n) = \delta(k; n)$ since each factor $\Phi_{n,s}(x)$ occurs in $k'[x]$. Now suppose that $F(x)$ factors into irreducible polynomials in $k[x]$ as

$$F(x) = \left\{ \prod_{n=1}^{\infty} \prod_{s=1}^{\delta(k; n)} \Phi_{n,s}(x)^{e(n, s)} \right\} \left\{ \prod_{i=1}^{I} f_i(x)^{m(i)} \right\}. \quad (1.4)$$

Here $e(n, s) \geq 0$, $m(i) \geq 1$, and $f_i(x)$, $i = 1, 2, \ldots, I$, are distinct, irreducible, non-cyclotomic polynomials in $k[x]$. There will be no loss of generality if we assume that $F$ is monic and hence that each $f_i$ is monic and has $\partial(f_i) \geq 1$. Of course $e(n, s) = 0$ all but finitely many pairs $\{n, s\}$, has $1 \leq s \leq \delta(k; n)$. Thus, the total number of cyclotomic factors of $F$ counted with multiplicity is $\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k; n)} e(n, s)$, and the total number of non-cyclotomic factors counted with multiplicity is $\sum_{i=1}^{I} m(i)$, (see [8, 9]).

Obviously, we have the trivial bound

$$\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k; n)} e(n, s) + \sum_{i=1}^{I} m(i) \leq \partial(F), \quad (1.5)$$

and in general nothing more can be said. However, if $\partial(F)$ is large compared with $\log H(F)$ we may expect to obtain sharper bounds. In this regard, it will be convenient to set
\[ r = r(F) = \max \left\{ \frac{\varrho(F)}{\log \nu(F)}, 3 \right\}, \quad (1.6) \]

where \( \nu(F) \) is defined in (2.7) of Section 2.

We are seeking in this paper to prove the following theorem:

**Theorem 1.1.** Let \( S \) be a set of irreducible, non-cyclotomic polynomials \( f(x) \) in \( \mathbb{k}[x] \) with \( f(0) \neq 0 \). Then, the existence of a positive constant \( \theta(S, \mathbb{k}) \) satisfying any one of the following statements would imply the truth of the other two:

(i) For all polynomials \( f(x) \) in \( S \)

\[
\log \mu(f) \geq \theta(S, \mathbb{k}). \quad (1.7)
\]

(ii) For all polynomials \( F(x) \) in \( \mathbb{k}[x], F(x) \) factoring in \( \mathbb{k}[x] \) as shown in (1.4), the number of irreducible factors of \( F(x) \) in \( S \) satisfies:

\[
\sum_{i=1}^{I} m(i) \leq \frac{\log \nu(F)}{\theta(S, \mathbb{k})}. \quad (1.8)
\]

(iii) For all polynomials \( F(x) \) in \( \mathbb{k}[x], F(x) \) factoring in \( \mathbb{k}[x] \) as shown in (1.4), the maximum multiplicity of a factor of \( F(x) \) in \( S \) satisfies:

\[
\max \{ m(i) : f_i(x) \in S, 1 \leq i \leq I \} \leq \frac{\log \nu(F)}{\theta(S, \mathbb{k})}. \quad (1.9)
\]

Clearly then if such a \( \theta(S, \mathbb{k}) > 0 \) exists, each of (i), (ii), and (iii) must also hold with the plainly optimal constant

\[
\beta(S, \mathbb{k}) \leq \inf \log \mu(f). \quad (1.10)
\]
When
\[ S_0 = \{ f(x) \in \mathbb{Q}[x] : f(x) \text{ is non-reciprocal} \}, \tag{1.11} \]
we have by the well-known result of Smyth [16] that \( \beta(S_0, \mathbb{Q}) = \log \theta_0 = 0.281 \ldots \), where \( \theta_0 = 1.324 \ldots \) is the real zero of \( x^3 - x - 1 \).

2. Definitions and Auxiliary Lemmas

At each place \( v \) of the number field \( k \), we write \( k_v \) for the completion of \( k \) at \( v \), \( \overline{k}_v \) for an algebraic closure of \( k_v \), and \( \Omega_v \) for the completion of \( k_v \). As is well-known, the field \( \Omega_v \) is then complete as a metric space and algebraically closed. Also, we introduce two normalized absolute values \( \| \cdot \|_v \) and \( \| \cdot \|_v \) on \( \Omega_v \), which are related by
\[ \| \cdot \|_v = \| \cdot \|_v^{d_v/d}, \tag{2.1} \]
where \( d = [k : \mathbb{Q}] \) and \( d_v = [k_v : \mathbb{Q}_v] \) is the local degree. If \( v | \infty \), then \( \| \cdot \|_v \) restricted to \( \mathbb{Q} \) is the usual Archimedean absolute value. If \( p \) is a prime number and \( v | p \), then \( \| \cdot \|_v \) restricted to \( \mathbb{Q} \) is the usual \( p \)-adic absolute value. Let
\[ F(x) = \sum_{l=0}^L a_l x^l = a_L \prod_{l=1}^L (x - \alpha_l), \tag{2.2} \]
be a polynomial in \( \Omega_v[x] \) and not identically zero. We define the local Mahler measure of \( F \) to be \( \mu_v(F) \), where
\[ \log \mu_v(F) = \log|a_L|_v + \sum_{l=1}^L \log^+|\alpha_l|_v. \tag{2.3} \]
We define the local height of \( F \) by
\[ H_v(F) = \max_l |a_l|_v, \quad \text{if} \ v \nmid \infty, \tag{2.4} \]
and

\[ H_v(F) = \left\{ \sum_{l=0}^{L} \|a_l\|^2_{|v|} \right\}^{d/(2d)}, \quad \text{if } v \mid \infty. \tag{2.5} \]

There is a third local measure of \( F \) which we require and define by

\[ v_v(F) = \sup \{|F(z)|_{|v|} : z \in \Omega_v \text{ and } |z|_v = 1\}. \tag{2.6} \]

Now suppose that \( F \) is given by (2.2) but has its coefficients in \( k \) and hence in \( \Omega_v \) for all places \( v \) of \( k \). In this case, we define the global Mahler measure \( \mu(F) \), the global height \( H(F) \), and the global measure \( v(F) \) by

\[ \mu(F) = \prod_v \mu_v(F), \quad H(F) = \prod_v H_v(F), \quad \text{and} \quad v(F) = \prod_v v_v(F), \tag{2.7} \]

respectively. It can easily be shown that in each of these products only finitely many factors are different from 1 (see, for example, [5], [6], and [7]). Because of the way we have normalized our absolute values \( | \cdot |_v \) and \( \| \cdot \|_v \), the global quantities \( \mu, H, \) and \( v \) do not depend on the number field \( k \) which contains the coefficients of \( F \). Thus, we may regard them as positive real valued functions defined on the not identically zero polynomials in \( \overline{Q}[x] \). For completeness we set \( \mu(F) = v(F) = H(F) = 0 \) if \( F \) is the zero polynomial.

If \( F \) is a monic, irreducible polynomial in \( Q[x] \), then by a result of Kronecker we have \( \mu(F) = 1 \) if and only if \( F(x) = x \) or \( F(x) \) is cyclotomic. If \( F \) is not \( x \) and not cyclotomic, then it is known that \( \log \mu(F) \) can be bounded away from zero by a positive quantity which depends only on the degree of \( F \). In fact, Dobrowolski [4] has shown that if \( F \) is not \( x \) and not cyclotomic, then
where \( L = \partial(F) \) is the degree of \( F \). Simpler proofs of (2.8) have been given by Cantor and Straus [3]; Rausch [13]; and Louboutin [10].

We now give several basic inequalities which relate the local and global functions we have defined on polynomials. It will be useful to define

\[
N(F) = \sum_{l=0}^{L} 1, \quad \sum_{a_i \neq 0}^{} \tag{2.9}
\]

so that \( N(F) \) is the number of nonzero coefficients of \( F \).

**Lemma 2.1** ([11], Theorem 3). Let \( F(x) \) be a polynomial in \( \Omega_v[x] \), where \( v \nmid \infty \). Then

\[
\mu_v(F) = \nu_v(F) = H_v(F). \tag{2.10}
\]

**Lemma 2.2** ([11], Theorem 4). Let \( F(x) \) be a polynomial in \( \Omega_v[x] \), where \( v \mid \infty \). Then

\[
H_v(F)^2 \leq \mu_v(F) \leq H_v(F) \leq \nu_v(F) \leq H_v(F)N(F)^{d_v/2d}. \tag{2.11}
\]

Applying Lemma 2.1 and Lemma 2.2, we have

**Lemma 2.3** ([11], Theorem 5). Let \( F(x) \) be a polynomial in \( k[x] \). Then

\[
H(F)^2 \leq \mu(F) \leq H(F) \leq \nu(F) \leq H(F)N(F)^{1/2} \leq H(F)^2. \tag{2.12}
\]

### 3. Proof of Theorem 1.1

Let

\[
S_k = \{ f(x) \in Q[x] : f(x) \text{ is irreducible, non-cyclotomic, } f(0) \neq 0 \}. \tag{3.1}
\]
Then the Lehmer conjecture amounts to the claim that there exists a positive $\theta(S_Q, \mathbb{Q})$ satisfying any of (i), (ii) or (iii). Now in general for any fields $k \subseteq L$, we have

$$\frac{\beta(S_k, k)}{[L : k]} \leq \beta(S_L, L) \leq \beta(S_k, k). \quad (3.2)$$

To see this, recall that if $\alpha$ is an algebraic number with minimal polynomial $F(x)$ in $S_L$ and $f(x)$ in $S_k$, then

$$\log \mu(F) = \hat{\nu}(F) \log \mu(x - \alpha) = \frac{\hat{\nu}(F)}{\nu(f)} \log \mu(f), \quad (3.3)$$

where clearly

$$1 \geq \hat{\nu}(F) = \frac{[L(\alpha) : L]}{[k(\alpha) : k]} = \frac{[L(\alpha) : k(\alpha)]}{[L : k]} \frac{1}{[L : k]}. \quad (3.4)$$

So in fact Lehmer’s conjecture is equivalent to the existence of a $\theta(S_k, k) > 0$ satisfying (i), (ii) or (iii) for any $k$. It is believed that when $k = \mathbb{Q}$ we should have $\beta(S_Q, \mathbb{Q}) = \log \theta_2 = 0.162\ldots$, where $\theta_2 = 1.176$ is a real zero of $x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$. What the correct value should be for a general is less certain.

It will be immediately apparent from the proof that the theorem might equally well have been stated with the $H(F)$ rather than the $\nu(F)$ height in (ii) or (iii). Indeed, the statement is still quite valid with $\mu(F)$ replacing the $\nu(F)$; though the proof in that case would be essentially a triviality. In fact the only real complexity in the proof of the theorem as stated will occur in proving that (iii) implies (i). For this we shall need the following result from Siegel’s lemma:

**Lemma 3.1.** Let $F(x)$ be a polynomial in $k[x]$. Then there exists a polynomial $G(x)$ in $k[x]$ with $\hat{\nu}(G) < \hat{\nu}(F)^2$ such that:

$$\log H(FG) = \log \mu(F) + O(\log \hat{\nu}(F)) + O(\log c_k); \quad (3.5)$$
\[ \log \nu(FG) = \log \mu(F) + O(\log \partial(F)) + O(\log c_k). \quad (3.6) \]

Here \( c_k = \frac{1}{2 \pi^{d/2} |\Delta_k|^{1/2d}} \), where \( s \) denotes the number of complex places of \( k \), \( d = [k : \mathbb{Q}] \) the degree of \( k \) and \( \Delta_k \) the discriminant of \( k \).

As the proof is non-constructive the nature of \( G(x) \) is not all clear although since \( \log \mu(G) = O(\log \partial(F)) \) we can say that such a polynomial ought to contain relatively few non-cyclotomic factors. Now for any \( f(x) \) in \( S \), we can use this lemma to show the existence of a polynomial highly divisible by \( f(x) \) whose height differs little from its Mahler measure. In particular, applying the lemma with \( F(x) = (x^3 - x - 1)^m \) and \( k = \mathbb{Q} \) produces the set of extremal polynomials \( Q_m(x) = F(x)G(x) \) in \( \mathbb{Q}[x] \) with

\[
\log H(Q_m) = m \log \theta_0 + O(m); \quad (3.7)
\]
\[
\log \nu(Q_m) = m \log \theta_0 + O(m), \quad (3.8)
\]
described in [11, 12].

**Proof of Lemma 3.1.** We shall need to briefly introduce yet another height \( h(F) \) on polynomials in \( \mathbb{Q}[x] \). For

\[
F(x) = \sum_{i=0}^{L} a_i x^i \in k[x], \quad (3.9)
\]

\( h(F) \) is defined by

\[
h(F) = \prod_{v \in V_k} h_v(F), \quad (3.10)
\]

where

\[
h_v(F) = \max_{0 \leq i \leq L} |a_i|_v. \quad (3.11)
\]
Now plainly at the finite places

\[ h_v(F) = H_v(F), \quad (3.12) \]

while at the infinite places the two local heights are related by

\[ h_v(F) = \left( \max_{0 \leq i < L} \| a_i \| \right)^{d_v/2d} \]

\[ \leq \sum_{i=0}^{L} (\| a_i \|)^{d_v/2d} = H_v(F) \]

\[ \leq (N(F) \max_{0 \leq i < L} \| a_i \|)^{d_v/2d} = N(F)^{d_v/2d} h_v(F). \quad (3.13) \]

Hence summing over all the places we obtain at once the global inequality:

\[ \log h(F) \leq \log H(F) \leq \log h(F) + \frac{1}{2} \log N(F) \leq \log h(F) + \frac{1}{2} \log(\delta(F) + 1). \quad (3.14) \]

Now if \( F(x) \) is a polynomial in \( \mathbb{k}[x] \) and \( N \) is a positive integer with

\[ M = N - \delta(F) > 0, \quad (3.15) \]

we can define a set

\[ S_{N,F} = \{ f(x) \in \mathbb{k}[x] : \delta(f) < N \text{ and } F(x)|f(x) \}. \quad (3.16) \]

Then by a result of Bombieri-Vaaler ([2], Theorem 1) (derived from their form of Siegel’s lemma ([1], Theorem 14)), there exist polynomials \( P_1(x), \ldots, P_M(x) \) in \( \mathbb{k}[x] \) forming a basis for \( S_{N,F} \) which satisfy

\[ \sum_{i=1}^{M} \log h(P_i) \leq M \log \mu(F) + M \log(c_k) + N^2 u \left( \frac{\delta(F)}{N} \right). \quad (3.17) \]
The function $u(\theta)$ has a fairly complicated definition but possesses the simple upper bound for $0 < \theta < 1$

$$u(\theta) < \frac{1}{2} \theta^2 \log \left( \frac{1}{4\theta} \right) + \frac{3}{4} \theta^2 \leq \theta^2 \log \left( \frac{2}{\theta} \right),$$

(3.18)

which is all that we shall use here. In particular then there is a polynomial $P(x)$ in $S_{N,F}$ with

$$\log h(P) \leq \log \mu(F) + \log c_k + O \left( \frac{\partial(F)^2}{M} \log \left( \frac{2N}{\partial(F)} \right) \right).$$

(3.19)

So by (3.14), we are guaranteed the existence of a polynomial $G(x) = P(x)/F(x)$ in $k[x]$ with $\partial(G) < M = N - (F)$ and

$$\log H(FG) \leq \log \mu(F) + \log c_k + O \left( \frac{\partial(F)^2}{M} \log \left( \frac{2N}{\partial(F)} \right) \right) + \frac{1}{2} \log N.$$  

(3.20)

Choosing $M = \partial(F)^2$ gives us the required upper bound

$$\log H(FG) \leq \log \mu(F) + \log c_k + O(\log \partial(F)).$$

(3.21)

Now from inequality (2.12) and the non-negativity and additivity of $\log \mu$, we always have the lower bound

$$\log H(FG) \geq \log \mu(FG) \geq \log \mu(F),$$

(3.22)

so that (3.5) holds as claimed.

The second expression follows at once on observing that by inequality (2.12)

$$\log H(FG) \leq \log \nu(FG) \leq \log H(FG) + \frac{1}{2} \log N(FG)$$

$$\leq \log H(FG) + \frac{1}{2} \log N(\partial(F)^2 + \partial(F)).$$

(3.23)
Proof of Theorem 1.1. The proof of (i) $\Rightarrow$ (ii) will follow closely that of Schinzel ([14], Theorem 1). Suppose that all the polynomials $f(x) \in S$ satisfy (1.7)

$$1 \leq \frac{\log \mu(f)}{\theta(S, k)},$$

and suppose that $F(x) \in k[x]$ factors in $k[x]$ as shown in (1.4). Then by the non-negativity and additivity of $\log \mu$ and the height inequality (2.12) discussed in Lemma 2.3, we have

$$\sum_{i=1}^{I} m(i) \leq \sum_{i=1}^{I} m(i) \frac{\log \mu(f_i)}{\theta(S, k)} \leq \frac{1}{\theta(S, k)} \sum_{i=1}^{I} \log \mu(f_i^{m(i)})$$

$$= \frac{\log \mu(F)}{\theta(S, k)} \leq \frac{\log H(F)}{\theta(S, k)} \leq \frac{\log \nu(F)}{\theta(S, k)},$$

(ii) $\Rightarrow$ (i) is immediate. Finally, we use the above lemma to show that (iii) $\Rightarrow$ (i). Let $f(x) \in k[x]$ be a polynomial in $S$. Then for each positive integer $m$, there must be a polynomial $F_m(x) = (x - f)^m G_m(x) \in k[x]$ with the comparatively low height

$$\log H(F_m) = \log \mu(f^m) + O(\log(m\hat{c}(f)) + O(\log c_k);$$

$$\log \nu(F_m) = \log \mu(f^m) + O(\log(m\hat{c}(f)) + O(\log c_k).$$

Since by construction $F_m(x)$ is divisible by $f(x)$ with multiplicity at least $m$ applying condition (iii) to $F_m(x)$ gives

$$m \leq \frac{\log \nu(F_m)}{\theta(S, k)} = \frac{m \log \mu(f)}{\theta(S, k)} + O \left( \frac{\log(m\hat{c}(f))}{m} \right) + O \left( \frac{\log c_k}{m} \right).$$

In the other words, for any $\varepsilon > 0$, we can, by taking $m > m_0(\varepsilon, \hat{c}(f), c_k)$ suitably large, ensure that

$$1 \leq \frac{\log \mu(f)}{\theta(S, k)} + \varepsilon$$

-condition (i) is therefore plainly assured.
Acknowledgements

The author is grateful to Professor M. B. Bhandari (NCU) and Professor R. Prakash (JU) for their useful suggestions during the preparation of the paper. He would also like to thank the referees and the editors for their corrections and comments which improved the presentation of the results of this paper.

References


