Ashwin SAH

An improved bound on the least common multiple of polynomial sequences

<http://jtnb.centre-mersenne.org/item?id=JTNB_2020__32_3_891_0>
An improved bound on the least common multiple of polynomial sequences

par Ashwin SAH

1. Introduction

If \( f \in \mathbb{Z}[x] \), let \( L_f(N) = \text{lcm}\{f(n) : 1 \leq n \leq N\} \), where say we ignore values of 0 in the LCM and set the LCM of an empty set to be 1. It is a well-known consequence of the Prime Number Theorem that

\[
\log \text{lcm}(1, \ldots, N) \sim N
\]
as \( N \to \infty \). Therefore, a similar linear behavior should occur if \( f \) is a product of linear polynomials. See the work of Hong, Qian, and Tan [4] for a more precise analysis of this case. On the other hand, if \( f \) is irreducible over \( \mathbb{Q} \) and has degree \( d \geq 2 \), \( \log L_f(N) \) ought to grow as \( N \log N \) rather than linearly. In particular, Cilleruelo [2] conjectured the following growth rate.
Conjecture 1.1 ([2]). If \( f \in \mathbb{Z}[x] \) is irreducible over \( \mathbb{Q} \) and has degree \( d \geq 2 \), then
\[
\log L_f(N) \sim (d - 1)N \log N
\]
as \( N \to \infty \).

He proved this for \( d = 2 \). As noted in [7], his argument demonstrates
\[
(1.1) \quad \log L_f(N) \lesssim (d - 1)N \log N.
\]
Hong, Luo, Qian, and Wang [3] showed that \( \log L_f(N) \gg N \), which was for some time the best known lower bound. Then, very recently, Maynard and Rudnick [5] provided a lower bound of the correct magnitude.

Theorem 1.2 ([5, Theorem 1.2]). Let \( f \in \mathbb{Z}[x] \) be irreducible over \( \mathbb{Q} \) with degree \( d \geq 2 \). Then there is \( c = c_f > 0 \) such that
\[
\log L_f(N) \gtrsim cN \log N.
\]
The proof given produces \( c_f = \frac{d-1}{d^2} \), although a minor modification produces \( c_f = \frac{1}{d} \). We prove the following improved bound, which in particular recovers Conjecture 1.1 when \( d = 2 \). It also does not decrease with \( d \), unlike the previous bound.

Theorem 1.3. Let \( f \in \mathbb{Z}[x] \) be irreducible over \( \mathbb{Q} \) with degree \( d \geq 2 \). Then
\[
\log L_f(N) \gtrsim N \log N.
\]

It is also interesting to consider the problem of estimating the quantity \( \ell_f(N) = \text{rad} \ \text{lcm}(f(1), \ldots, f(n)) \). (Recall that \( \text{rad}(n) \) is the product of the distinct primes dividing \( n \).) It is easy to see that the proof of Theorem 1.2 that was given in [5] implies
\[
\log \ell_f(N) \gtrsim c_f N \log N
\]
for the same constant \( c_f = \frac{d-1}{d^2} \) (or \( c_f = \frac{1}{d} \) after slight modifications). We demonstrate an improved bound.

Theorem 1.4. Let \( f \in \mathbb{Z}[x] \) be irreducible over \( \mathbb{Q} \) with degree \( d \geq 2 \). Then
\[
\log \ell_f(N) \gtrsim \frac{2}{d} N \log N.
\]

We conjecture that the radical of the LCM should be the same order of magnitude as the LCM.

Conjecture 1.5. If \( f \in \mathbb{Z}[x] \) is irreducible over \( \mathbb{Q} \) with degree \( d \geq 2 \), then
\[
\log \ell_f(N) \sim (d - 1)N \log N
\]
as \( N \to \infty \).
Finally, we note that Theorem 1.4 proves Conjecture 1.5 for $d = 2$.

In a couple of different directions, Rudnick and Zehavi [7] have studied the growth of $L_f$ along a shifted family of polynomials $f_a(x) = f_0(x) - a$, and Cilleruelo has asked for similar bounds in cases when $f$ is not irreducible as detailed by Candela, Rué, and Serra [1, Problem 4], which may also be tractable directions to pursue.

1.1. Commentary and setup. Interestingly, we avoid analysis of what is known as “Chebyshev’s problem” regarding the greatest prime factor $P^+(f(n))$ of $f(n)$, which is an essential element of the argument in [5]. Our approach is to study the product

$$Q(N) = \prod_{n=1}^{N} |f(n)|.$$  

We first analyze the contribution of small primes and linear-sized primes, which we show we can remove and retain a large product. Then we show that each large prime appears in the product a fixed number of times, hence providing a lower bound for the LCM and radical of the LCM. For convenience of our later analysis we write

$$Q(N) = \prod_{p} p^{\alpha_p(N)}.$$ 

Note that $\log Q(N) = dN \log N + O(N)$ by Stirling’s approximation, if $d$ is the degree of $f$. Finally, let $\rho_f(m)$ denote the number of roots of $f$ modulo $m$.

Remark on notation. Throughout, we use $g(n) \ll h(n)$ to mean $|g(n)| \leq ch(n)$ for some constant $c$, $g(n) \lesssim h(n)$ to mean for every $\epsilon > 0$ we have $|g(n)| \leq (1 + \epsilon)h(n)$ for sufficiently large $n$, and $g(n) \sim h(n)$ to mean $\lim_{n \to \infty} \frac{g(n)}{h(n)} = 1$. Additionally, throughout, we will fix a single $f \in \mathbb{Z}[x]$ that is irreducible over $\mathbb{Q}$ and has degree $d \geq 2$. We will often suppress the dependence of constants on $f$. We will also write

$$f(x) = \sum_{i=0}^{d} f_i x^i.$$ 

Acknowledgements. I thank Ze’ev Rudnick, Juanjo Rué, and Mark Shusterman for helpful comments, suggestions, and references.

2. Bounding small primes

The analysis in this section is very similar to that of [5, Section 3], except that we do not use the resulting bounds to study the Chebyshev problem.
We define
\[ Q_S(N) = \prod_{p \leq N} p^{\alpha_p(N)}, \]
the part of \( Q(N) = \prod_{n=1}^{N} |f(n)| \) containing small prime factors. The main result of this section is the following asymptotic.

**Proposition 2.1.** We have \( \log Q_S(N) \sim N \log N \).

**Remark 2.2.** This asymptotic directly implies the earlier stated Equation (1.1).

The argument is a simple analysis involving Hensel’s Lemma and the Chebotarev density theorem. The Hensel-related work has already been done in [5].

**Lemma 2.3 ([5, Lemma 3.1]).** Fix \( f \in \mathbb{Z}[x] \) and assume that it has no rational zeros. Let \( \rho_f(m) \) denote the number of roots of \( f \) modulo \( m \). Then if \( p \nmid \text{disc}(f) \) we have
\[ \alpha_p(N) = N \frac{\rho_f(p)}{p-1} + O\left( \frac{\log N}{\log p} \right) \]
and if \( p | \text{disc}(f) \) we have
\[ \alpha_p(N) \ll \frac{N}{p}, \]
where the implicit constant depends only on \( f \).

**Proof of Proposition 2.1.** We use Lemma 2.3. Noting that the deviation of the finitely many ramified primes from the typical formula is linear-sized, we will be able to ignore them with an error of \( O(N) \). We thus have
\[
\log Q_S(N) = \sum_{p \leq N} \alpha_p(N) \log p = \sum_{p \leq N} N \frac{\log p}{p-1} \rho_f(p) + O\left( N + \sum_{p \leq N} \log N \right) = N \sum_{p \leq N} \frac{\log p}{p-1} \rho_f(p) + O(N) = N \log N + O(N),
\]
using the Chebotarev density theorem alongside the fact that \( f \) is irreducible over \( \mathbb{Q} \) in the last equation (see e.g. [6, Equation (4)]). \( \square \)

### 3. Removing linear-sized primes

We define
\[ Q_{LI}(N) = \prod_{N < p \leq DN} p^{\alpha_p(N)}, \]
for appropriately chosen constant \( D = D_f \). We will end up choosing \( D = 1 + d[f_d] \) or so, although any greater constant will also work for the final argument. The result main result of this section is the following.
Proposition 3.1. We have $\log Q_{LI}(N) = O(N)$.

In order to prove this, we show that all large primes appear in the product $Q(N)$ a limited number of times.

**Lemma 3.2.** Let $N$ be sufficiently large depending on $f$, and let $p > N$ be prime. Then

$$\alpha_p(N) \leq d^2.$$ 

**Proof.** Note that $f \equiv 0 \pmod{p}$ has at most $d$ solutions, hence at most $d$ values of $n \in [1, N]$ satisfy $p|f(n)$ since $p > N$. For those values, we see $p^{d+1} > N^{d+1} \geq |f(n)|$ for all $n \in [1, N]$ if $N$ is sufficiently large, and $f$ is irreducible hence has no roots. Thus $p^{d+1}$ does not divide any $f(n)$ when $n \in [1, N]$.

Therefore $\alpha_p(N)$ is the sum of at most $d$ terms coming from the values $f(n)$ that are divisible by $p$. Each term, by the above analysis, has multiplicity at most $d$. This immediately gives the desired bound. \qed

**Proof of Proposition 3.1.** Using Lemma 3.2 we find

$$\log Q_{LI}(N) \leq d^2 \sum_{N < p \leq DN} \log p = O(N)$$

by the Prime Number Theorem. \qed

4. Multiplicity of large primes

Note that Lemma 3.2 is already enough to recreate Theorem 1.2. Indeed, we see that

$$\log \frac{Q(N)}{Q_S(N)} = (d - 1)N \log N + O(N)$$

from $Q(N) = dN \log N + O(N)$ and Proposition 2.1. Furthermore, by definition and by Lemma 3.2,

$$\frac{Q(N)}{Q_S(N)} = \prod_{p > N} p^{\alpha_p(N)} \leq \prod_{p > N, p|Q(N)} p^{d^2} \leq \ell_f(N)^{d^2} \leq L_f(N)^{d^2}.$$ 

This immediately gives the desired result (and recreates the constant $\frac{d-1}{d^2}$ appearing in the proof given in [5]).

In order to improve this bound, we will provide a more refined analysis of the multiplicity of large primes. More specifically, we will show that we have a multiplicity of $\frac{d(d-1)}{2}$ for primes $p > DN$, with $D$ chosen as in Section 3.

**Lemma 4.1.** Let $N$ be sufficiently large depending on $f$, and let $p > DN$ be prime, where $D = 1 + d|f_d|$. Then

$$\alpha_p(N) \leq \frac{d(d-1)}{2}.$$
Proof. Fix prime $p > DN$. As in the proof of Lemma 3.2, when $N$ is large enough in terms of $f$, we have that $p^{d+1}$ never divides any $f(n)$ for $n \in [1, N]$. Thus for $1 \leq i \leq d + 1$ let $b_i = \# \{ n \in [1, N] : p^i | f(n) \}$, where we see $b_{d+1} = 0$. Note that

$$\alpha_p(N) = \sum_{i=1}^{d} i(b_i - b_{i+1}) = \sum_{i=1}^{d} b_i.$$  

We claim that $b_i \leq d - i$ for all $1 \leq i \leq d$, which immediately implies the desired result.

Suppose for the sake of contradiction that $b_i \geq d - i + 1$ for some $1 \leq i \leq d$. Then let $m_1, \ldots, m_{d-i+1}$ be distinct values of $m \in [1, N]$ such that $p^i | f(m)$. Consider the value

$$A = A_i = \sum_{j=1}^{d-i+1} \frac{f(m_j)}{\prod_{k \neq j}(m_j - m_k)}.$$  

We have from the standard theory of polynomial identities that

$$A = \sum_{\ell=0}^{d} f_\ell \sum_{j=1}^{d-i+1} \frac{m_\ell^j}{\prod_{k \neq j}(m_j - m_k)} = \sum_{\ell=d-i}^{d} f_\ell \sum_{a_1+\cdots+a_{d-i+1}=\ell-(d-i)} \prod_{j=1}^{d-i+1} m_j^{a_j},$$  

where the inner sum is over all tuples $(a_1, \ldots, a_{d-i+1})$ of nonnegative integers that sum to $\ell - (d-i)$. Therefore $A \in \mathbb{Z}$. Furthermore, since $p^i | f(m_j)$ for all $1 \leq j \leq d - i + 1$, we have from the definition of $A$ that

$$p^i | A \prod_{1 \leq j < k \leq d-i+1} (m_j - m_k).$$  

Note that each $m_j - m_k$ is nonzero and bounded in magnitude by $N < p$, hence we deduce $p^i | A$.

But from the above formula and the triangle inequality we have

$$|A| = \left| \sum_{\ell=d-i}^{d} f_\ell \sum_{a_1+\cdots+a_{d-i+1}=\ell-(d-i)} \prod_{j=1}^{d-i+1} m_j^{a_j} \right| \leq \sum_{\ell=d-i}^{d} |f_\ell| \left( \frac{\ell}{d-i} \right) N^{\ell-(d-i)} \leq (1 + |f_d|d^i)N^i$$  

for sufficiently large $N$ in terms of $f$, using the fact that there are $(\binom{d}{d-i})$ tuples of nonnegative integers $(a_1, \ldots, a_{d-i+1})$ with sum $\ell - (d-i)$ and that $|m_j| \leq N$ for all $1 \leq j \leq d - i + 1$.  

Thus, as \( p > DN \geq (1 + |f_d|d)N \), we have
\[
|A| \leq (1 + |f_d|d^i)N^i \leq (1 + |f_d|d)^iN^i < p^i.
\]
Combining this with \( p^i|A| \), we deduce \( A = 0 \).

However, we will see that this leads to a contradiction as the “top-degree” term of \( A \) is too large in magnitude for this to occur. First, we claim that if \( 1 \leq i \leq d \) and \( d - i \leq \ell \leq d \), then
\[
\sum_{a_1 + \ldots + a_{d-i+1} = \ell - (d-i)} a_1 \ldots a_{d-i+1} m_{d-i+1}^{d-i+1} \prod_{j=1}^{d-i+1} m_j^{a_j} \in [1, 2^d].
\]
Indeed, note that each \( m_j > 0 \) and the denominator occurs as a subset of the terms in the numerator, hence the desired fraction is always at least 1.

For an upper bound, simply use the well-known AM-GM inequality. As it
the terms in the numerator, hence the desired fraction is always at least 1. If we take the terms
of this to occur. First, we claim that
\[
\sum_{a_1 + \ldots + a_{d-i+1} = \ell - (d-i)} a_1 \ldots a_{d-i+1} m_{d-i+1}^{d-i+1} \prod_{j=1}^{d-i+1} m_j^{a_j} \in [1, 2^d].
\]
Indeed, note that each \( m_j > 0 \) and the denominator occurs as a subset of the terms in the numerator, hence the desired fraction is always at least 1. For an upper bound, simply use the well-known AM-GM inequality. As it
turns out, a sharp upper bound for the above is \( \frac{1}{d-i+1} \binom{d}{d-i} \), which does not exceed \( 2^d \) for the given range of \( i \) and \( \ell \).

Next, we see that, using Equation (4.1) and the triangle inequality,
\[
|A| = \left| \sum_{\ell = d-i}^d f_\ell \prod_{a_1 + \ldots + a_{d-i+1} = \ell - (d-i)} a_1 \ldots a_{d-i+1} m_{d-i+1}^{d-i+1} \prod_{j=1}^{d-i+1} m_j^{a_j} \right| 
\geq |f_d| \sum_{a_1 + \ldots + a_{d-i+1} = d-i} \prod_{j=1}^{d-i+1} m_j^{a_j} - \sum_{\ell = d-i}^{d-1} |f_\ell| \sum_{a_1 + \ldots + a_{d-i+1} = \ell - (d-i)} \prod_{j=1}^{d-i+1} m_j^{a_j} 
\geq |f_d| \sum_{j=1}^{d-i+1} m_j^i - 2^d \sum_{\ell = d-i}^{d-1} |f_\ell| \sum_{j=1}^{d-i+1} m_j^{\ell-(d-i)} 
= \sum_{j=1}^{d-i+1} f^*(m_j),
\]
where we define \( f^*(x) = |f_d|x^i - 2^d \sum_{\ell = d-i}^{d-1} |f_\ell|x^{\ell-(d-i)} \). But since \( A = 0 \) and \( f^* \) clearly has a global minimum over the positive integers, we immediately deduce that \( |m_j| \) for all \( 1 \leq j \leq d - i + 1 \) is bounded in terms of some constant depending only on \( f \) and \( d = \deg f \).

But then, in particular, we also have \( |f(m_1)| < C_f \) for some constant \( C_f \) depending only on \( f \), yet it is divisible by \( p > DN \). For \( N \) sufficiently large in terms of \( f \), this can only happen if \( f(m_1) = 0 \), but since \( f \) is irreducible over \( \mathbb{Q} \) and \( \deg f = d \geq 2 \) this is a contradiction! Therefore we conclude that in fact \( b_i \leq d - i \) for all \( 1 \leq i \leq d \), which as remarked above finishes the proof.

We have actually proven something stronger, namely that for this range of \( p \) we have at most \( d - i \) values \( n \in [1, N] \) with \( p^i|f(n) \). In particular, this
implies that for \( p > DN \) we have

\[(4.2) \quad \#\{n \in [1, N] : p|f(n)\} \leq d - 1.\]

5. Finishing the argument

Proof of Theorem 1.3. The argument is similar to the one at the beginning of Section 4, but refined. We have

\[
\log \frac{Q(N)}{Q_S(N)Q_{LI}(N)} = (d - 1)N \log N + O(N)
\]

by \( Q(N) = dN \log N + O(N) \) and Propositions 2.1 and 3.1. Furthermore, by definition and by Equation (4.2),

\[
\frac{Q(N)}{Q_S(N)Q_{LI}(N)} = \prod_{p > DN} p^{\alpha_p(N)} \leq L_f(N)^{d-1}.
\]

The inequality comes from the fact that for \( p > DN > N \), there are at most \( d - 1 \) values of \( n \in [1, N] \) with \( p|f(n) \) from Equation (4.2), and the LCM takes the largest power of \( p \) from those involved hence has a power of at least \( \frac{\alpha_p(N)}{d-1} \) on \( p \). Taking logarithms, we deduce

\[(d - 1) \log L_f(N) \geq (d - 1)N \log N + O(N),\]

which immediately implies the result since \( d \geq 2 \). \( \square \)

Proof of Theorem 1.4. The argument is essentially identical to the one at the beginning of Section 4, but with a better multiplicity bound from Lemma 4.1. We have

\[
\log \frac{Q(N)}{Q_S(N)Q_{LI}(N)} = (d - 1)N \log N + O(N)
\]

by \( Q(N) = dN \log N + O(N) \) and Propositions 2.1 and 3.1. Furthermore, by definition and by Lemma 4.1,

\[
\frac{Q(N)}{Q_S(N)Q_{LI}(N)} = \prod_{p > DN} p^{\alpha_p(N)} \leq \prod_{p > DN, p|Q(N)} p^{\frac{d(d-1)}{2}} \leq \ell_f(N)^{\frac{d(d-1)}{2}}.
\]

Taking logarithms, we deduce

\[
\frac{d(d-1)}{2} \log \ell_f(N) \geq (d - 1)N \log N + O(N),
\]

which immediately implies the result since \( d \geq 2 \). \( \square \)
6. Discussion

We see from our approach that the major obstruction to proving Conjecture 1.1 is the potential for large prime factors $p > N$ to appear multiple times in the product $Q(N)$. In particular, it is possible to show that Conjecture 1.5 is equivalent to the assertion that

$$\lim_{N \to \infty} \frac{\# \{ p \text{ prime} : p^2 | Q(N) \}}{\# \{ p \text{ prime} : p | Q(N) \}} = 0.$$ 

Indeed, the bounds we have given are sufficient to show that there are $\Theta(N)$ prime factors of $Q(N)$, of which only $O(\frac{N}{\log N})$ are less than $DN$. Therefore the asymptotic size of the LCM is purely controlled by whether multiplicities for large primes in $[2, \frac{d(d-1)}{2}]$ appear a constant fraction of the time or not (noting that $\log p = \Theta(\log N)$ for these large primes, so that the sizes of their contributions are the same up to constant factors).

Similarly, Conjecture 1.1 is equivalent to the assertion that

$$\lim_{N \to \infty} \frac{\# \{ p \text{ prime} : \exists 1 \leq m < n \leq N : p | f(m), p | f(n) \}}{\# \{ p \text{ prime} : p | Q(N) \}} = 0.$$ 

Our bound for Conjecture 1.5 corresponds to using the fact that we can upper bound the multiplicities for all primes $p > DN$ by $\frac{d(d-1)}{2}$. In general, smaller multiplicities other than 1 could be possible but infrequent, which may be a direction to further approach Conjecture 1.1 and Conjecture 1.5.

References


Ashwin Sah
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

E-mail: asah@mit.edu

URL: www.mit.edu/~asah