COLLOQUIUM MATHEMATICUM

VOL. 145 2016 NO. 2

ON THE IRREDUCIBILITY OF A POLYNOMIAL ASSOCIATED WITH THE STRONG FACTORIAL CONJECTURE

BY

MICHAEL FILASETA and BRADY ROCKS (Columbia, SC)

Abstract. Asymptotically, more than 2/3 of the polynomials from a sequence of polynomials in $\mathbb{Z}[x]$, arising from an example associated with the Strong Factorial Conjecture, are shown to be irreducible in $\mathbb{Z}[x]$.

1. Introduction. The Strong Factorial Conjecture of E. Edo and A. van den Essen [3] is concerned with the linear functional L on the space of complex polynomials defined by sending a monomial generator $z_1^{a_1} \cdots z_n^{a_n}$ to $(a_1!) \cdots (a_n!)$. The conjecture asserts that for a non-zero multi-variable complex polynomial F, the maximum number of consecutive zeroes that may appear in the sequence $\{L(F^n) : n \geq 1\}$ is $N(F) - 1$, where $N(F)$ is the number of monomials appearing in F with non-zero coefficient.

In [12], the second author considered the irreducibility in $\mathbb{Z}[x]$ of the polynomials

$$f_{n,m}(x) = \sum_{j=0}^{n} \binom{n}{j} (mj)!\, x^j$$

in connection with his studies on the Strong Factorial Conjecture, specifically in the case $F = 1 + \lambda z^m$ where $\lambda \in \mathbb{C}$. Among other results, $f_{n,m}(x)$ was established in [12] to be irreducible when $n = p^r$ where p is a prime $> m$ and r is a positive integer.

In this paper, we prove the following.

THEOREM 1.1. *Fix a positive integer m. Then*

$$\liminf_{X \to \infty} \frac{|\{n \leq X : f_{n,m}(x) \text{ is irreducible}\}|}{X} \geq \log 2.$$

As $\log 2 = 0.693147\ldots$, we deduce that more than 2/3 of the polynomials $f_{n,m}(x)$ are irreducible in $\mathbb{Z}[x]$ for a fixed positive integer m. We do not know

2010 *Mathematics Subject Classification*: Primary 11R09; Secondary 11C08, 12E05.
Key words and phrases: Strong Factorial Conjecture, irreducible, Newton polygon.
Received 14 September 2015.
Published online 11 July 2016.

DOI: 10.4064/cm6761-3-2016

of an instance where $f_{n,m}(x)$ is reducible, so presumably a much stronger result than Theorem 1.1 holds.

2. Preliminaries on Newton polygons. Let

$$f(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{Z}[x] \quad \text{with } a_0 a_n \neq 0.$$

Let p be a prime. For an integer $m \neq 0$, we denote by $\nu_p(m)$ the exponent in the largest power of p dividing m. We define $\nu_p(0) = \infty$. Let S be the set of lattice points $(j, \nu_p(a_{n-j}))$, for $0 \leq j \leq n$, in the extended plane. We consider the lower edges along the convex hull of these points. The left-most edge has an endpoint $(0, \nu_p(a_n))$ and the right-most edge has $(n, \nu_p(a_0))$ as an endpoint. The polygonal path along the lower edges of the convex hull from $(0, \nu_p(a_n))$ to $(n, \nu_p(a_0))$ is called the *Newton polygon* of $f(x)$ with respect to the prime p. The endpoints of every edge belong to the set S, and each edge has a distinct slope that increases as we move along the Newton polygon from left to right.

The following important theorem due to G. Dumas [2] connects the Newton polygon of $f(x)$ with respect to a prime p with the Newton polygon of its factors with respect to the same prime.

THEOREM 2.1. *Let $g(x)$ and $h(x)$ be in $\mathbb{Z}[x]$ with $g(0)h(0) \neq 0$, and let p be a prime. Let k be a non-negative integer such that p^k divides the leading coefficient of $g(x)h(x)$ but p^{k+1} does not. Then the edges of the Newton polygon for $g(x)h(x)$ with respect to p can be formed by constructing a polygonal path beginning at $(0, k)$ and using translates of the edges in the Newton polygons for $g(x)$ and $h(x)$ with respect to the prime p, using exactly one translate for each edge of the Newton polygons for $g(x)$ and $h(x)$. Necessarily, the translated edges are translated in such a way as to form a polygonal path with the slopes of the edges increasing.*

As a particular consequence of Theorem 2.1, we have the following. Let $f(x) \in \mathbb{Z}[x]$ with $f(0) \neq 0$. Let

$$(x_0, y_0), (x_1, y_1), \ldots, (x_r, y_r), \qquad \text{with } 0 = x_0 < x_1 < \cdots < x_r = \deg f,$$

denote the lattice points along the edges of the Newton polygon of $f(x)$ with respect to a prime p. Set $d_j = x_j - x_{j-1}$ for $1 \leq j \leq r$. Then the set $\{1, \ldots, r\}$ can be written as a disjoint union of sets $S_1, \ldots, S_t$ where t is the number of irreducible factors of $f(x)$ (counted with multiplicities) and the t numbers $\sum_{u \in S_j} d_u$, for $1 \leq j \leq t$, are the degrees of the irreducible factors of $f(x)$. Note that it is important here to consider all lattice points along the edges of the Newton polygon of $f(x)$ with respect to p and not just lattice points of the form $(j, \nu_p(a_{n-j}))$ used in the construction of the Newton polygon.

Before applying Theorem 2.1 to obtain information about the factorization of $f_{n,m}(x)$, we first obtain information on Newton polygons of $f_{n,m}(x)$. We begin with a classical result on the largest power of a prime dividing a binomial coefficient that we use to compute $\nu_p(a_j)$ where $a_j = \binom{n}{j}(mj)!$ is the coefficient of x^j in $f_{n,m}(x)$.

Lemma 2.2. *Let n and j be non-negative integers with $n > 0$, and let p be a prime. If b is the number of borrows needed when j is subtracted from n in base p, then*

$$\nu_p\left(\binom{n}{j}\right) = b.$$

Lemma 2.2 is due to E. E. Kummer [7] but originally stated in the form of carries when adding j and $n-j$ in base p. Kummer uses another classical result due to A.-M. Legendre [8], connecting the largest power of p dividing $n!$ with the sum of the base p digits of n.

The next lemma can be found in [12]. The proof given here is based on a somewhat different analysis.

Lemma 2.3. *Let k, m and r be positive integers, and let q be a prime $> mk$. Let $n = kq^r$. Then the Newton polygon of $f_{n,m}(x)$ with respect to q consists of a single edge which has slope $-m(q^r-1)/(q^r(q-1))$.*

Proof. For $0 \le j \le n$, we set $a_j = \binom{n}{j}(mj)!$ so that $f_{n,m}(x) = \sum_{j=0}^{n} a_j x^j$. In particular,

$$\nu_q(a_0) = \nu_q(1) = 0.$$

Since $q > mk$, we have

$$\nu_q(a_n) = \nu_q((mn)!) = \sum_{u=1}^{\infty}\left\lfloor \frac{mn}{q^u}\right\rfloor = \sum_{u=1}^{r}\left\lfloor \frac{mkq^r}{q^u}\right\rfloor = \sum_{u=1}^{r}\frac{mkq^r}{q^u} = \frac{mk(q^r-1)}{q-1}.$$

We deduce that the line through $(0, \nu_q(a_n))$ and $(n, \nu_q(a_0))$ has slope equal to $-m(q^r-1)/(q^r(q-1))$ and equation

$$y = \frac{-m(q^r-1)}{q^r(q-1)}\cdot x + \frac{mk(q^r-1)}{q-1}.$$

We want to prove that, for $0 < j < n$, the point $(n-j, \nu_q(a_j))$ is above this line, that is,

$$\nu_q(a_j) \ge \frac{-m(q^r-1)}{q^r(q-1)}\cdot(n-j) + \frac{mk(q^r-1)}{q-1} = \frac{mj(q^r-1)}{q^r(q-1)}.$$

Note that n in base q consists of the single digit mk followed by r zeroes. Fix $j \in (0,n)$, and let $t = \nu_q(j)$. Then $j < n$ implies $t \in [0,r]$, and j in base q ends with exactly t digits that are zero. It follows that when j is subtracted from n in base q, exactly $r-t$ borrows are required. Hence,

$$\nu_q\left(\binom{n}{j}\right) = r - t.$$

Using the fact that $q^t \mid j$, we now deduce that

$$\begin{aligned}
\nu_q(a_j) &\geq \nu_q\left(\binom{n}{j}(mj)!\right) = \nu_q\left(\binom{n}{j}\right) + \nu_q((mj)!) \\
&= r - t + \sum_{u=1}^{\infty} \left\lfloor \frac{mj}{q^u} \right\rfloor = r - t + \sum_{u=1}^{t} \left\lfloor \frac{mj}{q^u} \right\rfloor + \sum_{u=t+1}^{r} \left\lfloor \frac{mj}{q^u} \right\rfloor \\
&= r - t + \sum_{u=1}^{t} \frac{mj}{q^u} + \sum_{u=t+1}^{r} \left\lfloor \frac{mj}{q^u} \right\rfloor \\
&\geq r - t + \sum_{u=1}^{t} \frac{mj}{q^u} + \sum_{u=t+1}^{r} \left(\frac{mj}{q^u} - 1\right) = \sum_{u=1}^{r} \frac{mj}{q^u} = \frac{mj(q^r-1)}{q^r(q-1)}.
\end{aligned}$$

The lemma follows. ∎

Lemma 2.4. *Let k and m be positive integers, and let q be a prime number $\geq (m+1)^2/(km)$. Let p be a prime in the interval $(kqm/(m+1), kq]$, and let $n = kq$. Then the Newton polygon of $f_{n,m}(x)$ with respect to p has an edge with slope $-m/p$.*

Comment. Though not needed for this paper, the statement of Lemma 2.4 seemingly holds for a larger range of primes p.

Proof of Lemma 2.4. Again, we set $f_{n,m}(x) = \sum_{j=0}^{n} a_j x^j$ where $a_j = \binom{n}{j}(mj)!$ for $0 \leq j \leq n$. Observe that

$$2p > \frac{2kqm}{m+1} \geq kq \geq n,$$

so $\nu_p(n!) = 1$. One checks that

$$\nu_p\left(\binom{n}{j}\right) = \begin{cases} 1 & \text{if } n-p < j < p, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

If the expression $(mj)!$ is divisible by p, then $j \geq p/m$. On the other hand, the condition $p > kqm/(m+1)$ is equivalent to $p/m > n - p$. Thus,

$$\nu_p\left(\binom{n}{j}(mj)!\right) = 0 \qquad \text{for } 0 \leq j \leq n-p.$$

The inequality $q \geq (m+1)^2/(km)$ implies

$$p^2 > \left(\frac{mn}{m+1}\right)^2 \geq mn.$$

From $p \in (kqm/(m+1), kq]$, we have

$$m \le mn/p < m+1.$$

Hence,

$$\nu_p(a_n) = \nu_p((mn)!) = \lfloor mn/p \rfloor + \lfloor mn/p^2 \rfloor + \cdots = \lfloor mn/p \rfloor = m.$$

We justify that the Newton polygon of $f_{n,m}(x)$ with respect to p consists of the segment s from $(0, m)$ to $(p, 0)$ together with the segment from $(p, 0)$ to $(n, 0)$. What is left to establish is that the points $(n-j, \nu_p(a_j))$, for $n-p < j < n$, lie on or above the segment s. Since the line through $(0, m)$ and $(p, 0)$ has equation $y = (-m/p)x + m$, we want to prove

$$\nu_p(a_j) \ge -m(n-j)/p + m \tag{2.2}$$

for $n-p < j < n$. As $p \le n$, we have

$$-m(n-j)/p + m = -mn/p + mj/p + m \le -m + mj/p + m = mj/p.$$

Thus, for $j \in (n-p, n)$, it suffices to show that either (2.2) holds or

$$\nu_p(a_j) \ge mj/p. \tag{2.3}$$

For $n-p < j < p$, using (2.1), we see that

$$\nu_p(a_j) = \nu_p\left(\binom{n}{j}(mj)!\right) = 1 + \nu_p((mj)!) \ge 1 + \lfloor mj/p \rfloor > mj/p,$$

so that (2.3) holds for such j. For $p \le j < n$, we have

$$\nu_p(a_j) = \nu_p((mj)!) \ge \lfloor mj/p \rfloor \ge \lfloor mp/p \rfloor = m,$$

implying (2.2) for these j. The lemma follows. ∎

3. Proof of Theorem 1.1. H. Cramér [1] showed that if the Riemann Hypothesis holds and p_n is the nth prime number, then $p_{n+1} - p_n = O(\sqrt{p_n} \log p_n)$. According to C. J. Moreno [10], P. Erdős posed the related problem of establishing that, for every $\varepsilon > 0$, almost all numbers n are a distance $\le n^{(1/2)+\varepsilon}$ from a prime. More specifically, Erdős asked whether there is a constant $c < 1$ such that

$$\sum_{\substack{p_{n+1}-p_n > x^{(1/2)+\varepsilon} \\ p_{n+1} \le x}} (p_{n+1} - p_n) \ll x^c.$$

Moreno establishes this asymptotic in a weaker form with x^c replaced nevertheless by a function which is small compared to x as x tends to infinity. D. Wolke [13] resolved the problem of Erdős in the affirmative, and a number of other authors (cf. [5, 6, 9, 11]) have since improved on the value of c

in the asymptotic. In particular, K. Matomäki's work [9] implies that

$$\sum_{\substack{p_{n+1}-p_n>\sqrt{p_n}\\ p_n\le x}} (p_{n+1}-p_n) \ll x^{2/3}. \tag{3.1}$$

For our purposes, the weaker result of Moreno would suffice, but we use (3.1).

Fix a positive integer m. Let $M=(m+1)^2/m$. Note that $M\ge 4$. Let $\mathcal{A}$ be the set of positive integers n that have a prime factor $q>\sqrt{Mn}$. Let $\mathcal{B}$ be the set of positive integers n for which there exists a prime p satisfying $n-\sqrt{n}<p\le n$. Set $\mathcal{C}=\mathcal{A}\cap\mathcal{B}$. We next obtain the asymptotic densities of the sets $\mathcal{A}$ and $\mathcal{B}$ in the set of integers, that is, the values of

$$\lim_{x\to\infty}\frac{|\{n\le x : n\in\mathcal{A}\}|}{x} \quad\text{and}\quad \lim_{x\to\infty}\frac{|\{n\le x : n\in\mathcal{B}\}|}{x}.$$

The asymptotic density of $\mathcal{A}$ is connected with the distribution of smooth numbers (numbers with only small prime factors) and is easily explained. Using the notation $\pi(x)$ for the number of primes $\le x$, and p to represent a prime, observe that

$$\begin{aligned}
&|\{x<n\le 2x : n\in\mathcal{A}\}|\\
&\quad=\sum_{\sqrt{Mx}<p\le 2x}\left(\left\lfloor\frac{2x}{p}\right\rfloor-\left\lfloor\frac{x}{p}\right\rfloor\right)+O\bigg(\sum_{\sqrt{Mx}<p\le\sqrt{2Mx}}\left(\left\lfloor\frac{2x}{p}\right\rfloor-\left\lfloor\frac{x}{p}\right\rfloor\right)\bigg)\\
&\quad=\bigg(\sum_{\sqrt{Mx}<p\le 2x}\frac{x}{p}\bigg)+O(\pi(2x))+O\bigg(\sum_{\sqrt{Mx}<p\le\sqrt{2Mx}}\frac{x}{p}\bigg).
\end{aligned}$$

Using Mertens' estimate for the sum of the reciprocals of the primes (cf. [4, Theorem 427]) and a Chebyshev estimate (cf. [4, Theorem 7]), we can deduce from the above that

$$\lim_{x\to\infty}\frac{|\{n\le x : n\in\mathcal{A}\}|}{x}=\log 2. \tag{3.2}$$

For the asymptotic density of $\mathcal{B}$, we consider first the asymptotic density of the complement of $\mathcal{B}$ in the set of positive integers. Fix a positive integer n in the complement of $\mathcal{B}$. Let p' and p'' be the consecutive primes for which $p'\le n<p''$. Since $n\notin\mathcal{B}$, we have $p'\le n-\sqrt{n}$. Thus,

$$p''-p'>n-(n-\sqrt{n})=\sqrt{n}\ge\sqrt{p'}.$$

Therefore, such n lie in an interval $[p',p'')$ where p' and p'' are consecutive primes for which $p''-p'>\sqrt{p'}$. By (3.1), the n in the complement of $\mathcal{B}$ have asymptotic density 0. Therefore,

$$\lim_{x\to\infty}\frac{|\{n\le x : n\in\mathcal{B}\}|}{x}=1. \tag{3.3}$$

Combining (3.2) and (3.3), we deduce that

$$\lim_{x\to\infty} \frac{|\{n \le x : n \in \mathcal{C}\}|}{x} = \log 2.$$

Thus, to establish Theorem 1.1, it suffices to show that if n is a sufficiently large element of $\mathcal{C}$, then $f_{n,m}(x)$ is irreducible.

Consider such an n. Then $n \in \mathcal{A}$ implies that we can write $n = kq$ where k is a positive integer and q is a prime satisfying

$$q > \sqrt{Mn} = \sqrt{Mkq} \ \Rightarrow\ q > Mk > mk.$$

By Lemma 2.3, we deduce that the Newton polygon of $f_{n,m}(x)$ with respect to the prime q consists of a single edge with slope $-m/q$. Since q is a prime $> m$, the fraction $-m/q$ is reduced. As a consequence of Theorem 2.1, we can deduce that each irreducible factor of $f_{n,m}(x)$ has degree divisible by q (as noted in [12]).

Next, we apply Lemma 2.4. Since $q > Mk$ where $M = (m+1)^2/m$, we see that

$$q > \frac{(m+1)^2 k}{m} \ge \frac{(m+1)^2}{km}.$$

We set p to be the largest prime $\le n$. To apply Lemma 2.4, we want to show that

$$p > \frac{nm}{m+1}.$$

Since n is sufficiently large and m is fixed, this inequality is an easy consequence of the Prime Number Theorem (i.e., that there is a prime number in the interval $((1-\varepsilon)n, n]$, where $\varepsilon = 1/(m+1)$). Lemma 2.4 implies that the Newton polygon of $f_{n,m}(x)$ with respect to the prime p has an edge with slope $-m/p$. Theorem 2.1 now shows that $f_{n,m}(x)$ has an irreducible factor of degree $\ge p$.

To establish that $f_{n,m}(x)$ is irreducible, it is now sufficient to show that the smallest multiple of q that is $\ge p$ is $n = kq$. This is equivalent to establishing that $n - q < p$. Since $q > \sqrt{Mn} > \sqrt{n}$, we need only show that $n - \sqrt{n} < p$. The latter inequality follows from $n \in \mathcal{B}$, completing the proof of Theorem 1.1.

Acknowledgments. The first author expresses his gratitude to the NSA for grant support during the period of this research.

REFERENCES

[1] H. Cramér, *Some theorems concerning prime numbers*, Ark. Mat. Astronom. Fys. 15 (1920), no. 5, 33 pp.

[2] G. Dumas, *Sur quelques cas d'irréductibilité des polynomes à coefficients rationnels*, J. Math. Pures Appl. 2 (1906), 191–258.

[3] E. Edo and A. van den Essen, *The Strong Factorial Conjecture*, J. Algebra 397 (2014), 443–456.
[4] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford Univ. Press, New York, 1979.
[5] D. R. Heath-Brown, *The differences between consecutive primes*, J. London Math. Soc. (2) 18 (1978), 7–13.
[6] M. N. Huxley, *A note on large gaps between prime numbers*, Acta Arith. 38 (1980/81), 63–68.
[7] E. E. Kummer, *Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen*, J. Reine Angew. Math. 44 (1852), 93–146.
[8] A.-M. Legendre, *Théorie des Nombres*, Firmin Didot Frères, Paris, 1830.
[9] K. Matomäki, *Large differences between consecutive primes*, Q. J. Math. 58 (2007), 489–518.
[10] C. J. Moreno, *The average size of gaps between primes*, Mathematika 21 (1974), 96–100.
[11] A. S. Peck, *Differences between consecutive primes*, Proc. London Math. Soc. (3) 76 (1998), 33–69.
[12] B. J. Rocks, *Incompatibility of Diophantine equations arising from the Strong Factorial Conjecture*, Arts & Sciences Electronic Theses and Dissertations, Washington Univ. in St. Louis, paper 439 (2015).
[13] D. Wolke, *Grosse Differenzen zwischen aufeinanderfolgenden Primzahlen*, Math. Ann. 218 (1975), 269–271.

Michael Filaseta, Brady Rocks
Mathematics Department
University of South Carolina
Columbia, SC 29208, U.S.A.
E-mail: filaseta@math.sc.edu
ROCKS@math.sc.edu