PRIME POLYNOMIAL SEQUENCES

S. D. COHEN, P. ERDŐS AND M. B. NATHANSON

ABSTRACT

Let $F(x)$ be a polynomial with integral coefficients of degree $d \geqslant 2$ such that $F(n) \geqslant 1$ for all $n \geqslant 1$. Let $\mathcal{O}_F = \{F(n)\}_{n=1}^{\infty}$. Then $F(n)$ is called prime in $\mathcal{O}_F$ if $F(n)$ is not the product of strictly smaller terms of $\mathcal{O}_F$. It is proved that if $F(x)$ is not of the form $a(bx+c)^d$, then almost all terms of $\mathcal{O}_F$ are prime in $\mathcal{O}_F$.

Let $\mathcal{O} = \{a_n\}_{n=1}^{\infty}$ be a sequence of positive integers. Then $a_n \in \mathcal{O}$ is called *composite in* $\mathcal{O}$ if $a_n > 1$ and a_n can be written as a product of terms $a_i \in \mathcal{O}$ with $a_i < a_n$. If $a_n > 1$ and a_n is not composite in $\mathcal{O}$, then a_n is called *prime in* $\mathcal{O}$. In this note we consider sequences of the form $\mathcal{O}_F = \{F(n)\}_{n=1}^{\infty}$, where $F(x)$ is a polynomial with integral coefficients of degree $d \geqslant 2$ such that $F(n) \geqslant 1$ for all $n \geqslant 1$. We shall prove that if $F(x)$ is not of the form $F(x) = a(bx+c)^d$, then almost all terms of the sequence $\mathcal{O}_F$ are prime in $\mathcal{O}_F$.

Notation. Let $F(x)$ be a polynomial with integral coefficients. Let $\rho_F(m)$ denote the number of solutions of the congruence $F(n) \equiv 0 \pmod{m}$ with $1 \leqslant n \leqslant m$, and let $\theta_F(m, x)$ denote the number of solutions of the congruence $F(n) \equiv 0 \pmod{m}$ with $1 \leqslant n \leqslant x$. The polynomial $F(x)$ is $(t+1)$-free if $F(x)$ is not divisible by the $(t+1)$-st power of any non-constant polynomial. We write $f \ll g$ if $|f(x)| < c\,|g(x)|$ for some constant c and all sufficiently large x.

THEOREM. *Let $F(x)$ be a polynomial with integral coefficients of degree $d \geqslant 2$ such that $F(n) \geqslant 1$ for all $n \geqslant 1$ and such that $F(x)$ is $(t+1)$-free, where $1 \leqslant t \leqslant d-1$. Let $\mathcal{O}_F = \{F(n)\}_{n=1}^{\infty}$, and let $C(x)$ denote the number of $F(n)$ in $\mathcal{O}_F$ with $n \leqslant x$ which are composite in $\mathcal{O}_F$. Then*

$$C(x) \ll x^{(d+1)/(d+2)+\varepsilon} + x^{(t/d)(2-t/d)+\varepsilon}$$

for every $\varepsilon > 0$. In particular, if $F(x)$ is not a constant multiple of a linear polynomial, then

$$C(x) \ll x^{1-(1/d^2)+\varepsilon}$$

for every $\varepsilon > 0$.

We shall require the following result.

LEMMA. *Let $F(x)$ be a $(t+1)$-free polynomial with integral coefficients. Then*

$$\theta_F(m, x) \ll \left(1 + \frac{x}{m^{1/t}}\right) m^{\varepsilon}$$

for every $\varepsilon > 0$.

Proof. If $G(x)$ is a square-free polynomial with integral coefficients, then Nagell ([3] and [4; p. 90]) and Ore [7] have proved that, for any prime p and $k \geqslant 1$,

$$\rho_G(p^k) \leqslant d'D^2$$

Received 9 August, 1976.

[J. LONDON MATH. SOC. (2), 14 (1976), 559–562]

where d' is the degree and D the discriminant of $G(x)$. Let $\omega(m)$ denote the number of distinct primes dividing m, and let $\tau(m)$ denote the number of divisors of m. Then $2^{\omega(m)} \leqslant \tau(m) \ll m^{\varepsilon}$ for every $\varepsilon > 0$. Since $\rho_G(m)$ is a multiplicative function of m, it follows that

$$\begin{aligned} \rho_G(m) &\leqslant (d'D^2)^{\omega(m)} = 2^{\omega(m)\log_2(d'D^2)} \\ &\leqslant \tau(m)^{\log_2(d'D^2)} \ll m^{\varepsilon}. \end{aligned}$$

Now let $F(x)$ be a $(t+1)$-free polynomial. We can assume that the coefficients of $F(x)$ are relatively prime. Let $G(x)$ be the product of the irreducible polynomials dividing $F(x)$. Then $G(x)$ divides $F(x)$, and, since $F(x)$ is $(t+1)$-free, $F(x)$ divides $G(x)^t$. Let m_1 be the smallest divisor of m such that $m \mid m_1{}^t$. Then $m^{1/t} \leqslant m_1$. If $F(n) \equiv 0 \pmod{m}$, then $G(n)^t \equiv 0 \pmod{m}$. But this implies that $G(n)^t \equiv 0 \pmod{m_1{}^t}$ and so $G(n) \equiv 0 \pmod{m_1}$. Therefore,

$$\begin{aligned} \theta_F(m, x) &\leqslant \theta_G(m_1, x) \\ &\leqslant \left(1 + \frac{x}{m_1}\right) \rho_G(m_1) \\ &\ll \left(1 + \frac{x}{m^{1/t}}\right) m^{\varepsilon} \end{aligned}$$

for every $\varepsilon > 0$. This proves the lemma.

Proof of the Theorem. Let $F(x)$ be a $(t+1)$-free polynomial of degree $d \geqslant 2$. Fix $0 < \lambda < 1$. Let $C_1(x)$ denote the number of $n \leqslant x$ such that

$$F(n) = F(u_1)(Fu_2) \dots F(u_s),$$

where $1 \leqslant u_i \leqslant x^{\lambda}$ and $1 < F(u_i) < F(n)$ for $i = 1, 2, \dots, s$. Let $C_2(x)$ denote the number of $n \leqslant x$ such that $F(n)$ is divisible by some $F(u)$ with $x^{\lambda} < u \leqslant x$ and $1 < F(u) < F(n)$. Then

$$C(x) \leqslant C_1(x) + C_2(x). \tag{1}$$

We first estimate $C_1(x)$. Let $x^{t/d} < n \leqslant x$, and suppose that

$$F(n) = F(u_1) F(u_2) \dots F(u_s),$$

where $1 \leqslant u_1 \leqslant u_2 < \dots \leqslant u_s \leqslant x^{\lambda}$ and $1 < F(u_i) < F(n)$. Choose constants $0 < \alpha < \beta$ such that

$$\alpha n^d < F(n) < \beta n^d$$

for all $n \geqslant 1$. Then $F(n) > \alpha n^d > \alpha x^t$ for $n > x^{t/d}$, and so

$$F(u_1) \dots F(u_{r-1}) \leqslant \alpha x^t < F(u_1) \dots F(u_{r-1}) F(u_r) = m \tag{2}$$

for some $r \leqslant s$. Since $2^{r-1} \leqslant F(u_1) \dots F(u_{r-1}) \leqslant \alpha x^t$, it follows that $r < \gamma \log x$ for some $\gamma > 0$ and all $x > x_0$. Moreover, $m \mid F(n)$. For fixed m of the form (2), the number of $n \leqslant x$ such that $F(n)$ is divisible by $m = F(u_1) \dots F(u_r)$ is, by the lemma,

$$\theta_F(m, x) \ll \left(1 + \frac{x}{m^{1/t}}\right) m^{\varepsilon} \ll x^{\varepsilon}$$

for every $\varepsilon > 0$.

We must now estimate the number of m of the form (2). Since

$$F(u_r) < \beta u_r^d \leqslant \beta x^{\lambda d},$$

it follows from (2) that

$$\alpha^r (u_1 u_2 \dots u_r)^d \leqslant m < \alpha\beta\, x^{\lambda d + t}$$

and so

$$u_1 u_2 \dots u_r < (\alpha^{1-r}\beta)^{1/d}\, x^{\lambda + (t/d)}.$$

Given $\varepsilon > 0$, choose $\delta > 0$ such that $-\gamma \log(1-\delta) < \varepsilon$. There exists $N(\delta) = N > 1$ such that $F(n) > (1-\delta)\, n^d$ for all $n \geqslant N$. Suppose $m = F(u_1) \dots F(u_r)$, where

$$u_1 \leqslant \dots \leqslant u_p \leqslant N < u_{p+1} \leqslant \dots \leqslant u_r.$$

Let $m_0 = F(u_1) \dots F(u_p)$ and $m_1 = F(u_{p+1}) \dots F(u_r)$. Then $m = m_0 m_1$. Since $F(u_i) \geqslant 2$, it follows that the number of possible integers m_0 is $\ll (\log x)^N \ll x^\varepsilon$. Moreover,

$$m_1 > (1-\delta)^{r-p} (u_{p+1} \dots u_r)^d \geqslant (1-\delta)^r (u_{p+1} \dots u_r)^d$$

and so

$$\begin{aligned}(u_{p+1} \dots u_r)^d &< (1-\delta)^{-r} m_1 \\ &< (1-\delta)^{-\gamma \log x} m_1 \\ &< x^\varepsilon m \\ &\ll x^{\lambda d + t + \varepsilon}.\end{aligned}$$

Consequently,

$$u_{p+1} \dots u_r \ll x^{\lambda + (t/d) + \varepsilon}. \tag{3}$$

By a result of Oppenheim [**5**, **6**] and Szekeres and Turán [**8**], the number of products of the form (3) is $\ll x^{\lambda + (t/d) + \varepsilon}$. Therefore, the number of integers m of the form (2) is $\ll x^{\lambda + (t/d) + \varepsilon}$, and so

$$C_1(x) \ll x^{\lambda + (t/d) + \varepsilon}. \tag{4}$$

We shall now estimate $C_2(x)$. The number of $F(n)$ with $n \leqslant x$ which are divisible by some $F(u)$ with $x^\lambda < u \leqslant x^{(d+1)/(d+2)}$ does not exceed

$$\begin{aligned}\sum_{u=x^\lambda}^{x^{(d+1)/(d+2)}} \theta_F(F(u), x) &< \sum_{u=x^\lambda}^{x^{(d+1)/(d+2)}} \left(1 + \frac{x}{F(u)^{1/t}}\right) x^\varepsilon \\ &< x^{(d+1)/(d+2)+\varepsilon} + x^{1+\varepsilon} \sum_{u=x^\lambda}^{\infty} F(u)^{-1/t} \\ &< x^{(d+1)/(d+2)+\varepsilon} + x^{1+\varepsilon} \int_{x^\lambda - 1}^{\infty} (\alpha u^d)^{-1/t}\, du \\ &\ll x^{(d+1)/(d+2)} + x^{1-\lambda(d/t-1)+\varepsilon}.\end{aligned}$$

Moreover, Anderson, Cohen, and Stothers [**1**, **2**] have proved that the number of $F(n)$ with $n \leqslant x$ which are divisible by some $F(u)$ with $u > x^{(d+1)/(d+2)}$ is

$$\ll x^{(d+1)/(d+2)}.$$

L

Therefore,

$$C_2(x) \ll x^{(d+1)/(d+2)+\varepsilon} + x^{1-\lambda(d/t-1)+\varepsilon}. \tag{5}$$

Combining (1), (4), and (5), we obtain

$$C(x) \ll x^{(d+1)/(d+2)+\varepsilon} + x^{1-\lambda(d/t-1)+\varepsilon} + x^{\lambda+t/d+\varepsilon}.$$

The minimum of the right-hand side of this inequality occurs when

$$\lambda = (t/d)\left(1-(t/d)\right).$$

This yields

$$C(x) \ll x^{(d+1)/(d+2)+\varepsilon} + x^{(t/d)(2-t/d)+\varepsilon}.$$

This completes the proof of the theorem.

COROLLARY. *Let $F(x)$ be a square-free quadratic polynomial. Then*

$$C(x) \ll x^{\frac{3}{4}+\varepsilon}.$$

Remarks. If $F(x) = x^2+bx+c$, then the polynomial identity

$$F(x)\,F(x+1) = F(x^2+(b+1)x+c)$$

implies that the number $C(x)$ of composite numbers in $\mathcal{O}_F$ satisfies

$$x^{\frac{1}{2}} \ll C(x) \ll x^{\frac{3}{4}+\varepsilon}.$$

The exact order of magnitude of $C(x)$ is unknown. One can conjecture that if $F(x)$ is a polynomial of degree $d \geqslant 2$ that is not of the form $a(bx+c)^d$, then $C(x) \ll x^{(1/d)+\varepsilon}$, but this is unknown even for $d = 2$. On the other hand, it is not difficult to construct monic polynomials $F(x)$ for which $C(x) = 0$ for all x. For example, let p be prime and let $F(x) = \left(x(x+1)\ldots(x+p-1)\right)^{2l}+p^k$ for $1 \leqslant k \leqslant l$. Then $F(n) \equiv p^k \pmod{p^{2k}}$ for every n, but $F(u_1)\ldots F(u_r) \equiv 0 \pmod{p^{2k}}$ whenever $r \geqslant 2$, and so no $F(n)$ in $\mathcal{O}_F$ is composite in $\mathcal{O}_F$. It is an open problem to determine those polynomials $F(x)$ for which the sequence $\mathcal{O}_F = \{F(n)\}_{n=1}^{\infty}$ contains infinitely many numbers composite in $\mathcal{O}_F$.

References

1. I. Anderson, S. D. Cohen and W. W. Stothers, " Primitive polynomial subsequences ", *Mathematika*, 21 (1974), 239–247.
2. S. D. Cohen, "Dense primitive polynomial sequences ", *Mathematika*, 22 (1975), 89–91.
3. T. Nagell, " Généralisation d'un théorème de Tchebycheff ", *J. de Mathématiques*, 4 (1921), 343–356.
4. ———, *Introduction to Number Theory* (Chelsea, New York, 1951).
5. A. Oppenheim, " On an arithmetic function ", *J. London Math. Soc.*, 1 (1926), 205–211.
6. ———, " On an arithmetic function (II) ", *J. London Math. Soc.*, 2 (1927), 123–130.
7. O. Ore, " Anzahl der Wurzeln höherer Kongruenzen ", *Norsk Matematisk Tidskrift*, 3 (1921), 63–66.
8. G. Szekeres and P. Turán, " Über das zweite Hauptproblem der 'Factorisatio Numerorum' ", *Acta Scientiarum Mathematicarum*, 6 (1932), 143–154.

S. D. Cohen
Department of Mathematics
University of Glasgow
Glasgow G12 8QW, Scotland

M. B. Nathanson
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901, U.S.A.

P. Erdős
Mathematical Institute
Hungarian Academy of Sciences
Budapest, Hungary