ARITHMETICAL PROPERTIES OF POLYNOMIALS

P. ERDÖS*.

[*Extracted from the Journal of the London Mathematical Society, Vol.* 28, 1953.]

1. Throughout this paper $f(x)$ will denote a polynomial whose coefficients are integers with highest common factor 1, and l will denote the degree of $f(x)$. We assume that the highest coefficient in $f(x)$ is positive. It has been known for some time that if $f(x)$ is not the l-th power of a linear polynomial with integral coefficients then there are infinitely

* Received 22 April, 1952; read 15 May, 1952; revised 16 October, 1952.

many positive integers n for which $f(n)$ is l-th power free, *i.e.* $f(n)$ is not divisible by any integral l-th power greater than 1. In fact by a simple application of the Sieve of Eratosthenes and an easy limit process it can be proved (as has also been known for some time) that the integers n for which $f(n)$ is l-th power free have positive density.

Let us now assume that $l \geqslant 3$ and that $f(x)$ is not divisible by the $(l-1)$-th power of a linear polynomial with integral coefficients. It has then been conjectured that there are in general infinitely many n for which $f(n)$ is $(l-1)$-th power free, and further that the density of these n is positive.

The need for the qualification "in general" arises in the following way. It may happen that there exists an integer d such that $f(n) \equiv 0 \pmod{d}$ for every n. It is known that such an integer d must be a divisor of $l!$, and the example

$$f(x) = l!\left(\binom{x}{l}+1\right)$$

shows that d may equal $l!$. Now if l is a power of 2, $l!$ is divisible by 2^{l-1}, and if $d = l!$ it is impossible for $f(n)$ to be $(l-1)$-th power free for any n. We must therefore exclude this case, which we do by assuming from now onwards that if l is a power of 2 there exists some n (and therefore infinitely many n) such that $f(n) \not\equiv 0 \pmod{2^{l-1}}$.

In the present paper we shall prove the following

THEOREM. *If $l \geqslant 3$ and $f(x)$ satisfies the conditions stated above, then there are infinitely many positive integers n for which $f(n)$ is $(l-1)$-th power free.*

It seems very likely that the integers n with the property in question have positive density, but this I have not been able to prove.

It will be clear from the proof that the following result holds in the exceptional case when $f(n) \equiv 0 \pmod{2^{l-1}}$ for every n: *there exist infinitely many n for which $f(n) = 2^{l-1}u_n$, where u_n is odd and is $(l-1)$-th power free.*

2. In this section we dispose of the case in which $f(x)$ is reducible. We can express $f(x)$ as $g(x)h(x)$, where $g(x)$ and $h(x)$ are polynomials with integral coefficients. Moreover, there exists such a representation with $g(x)$ and $h(x)$ relatively prime, except in the case when $f(x) = \big(\phi(x)\big)^k$, where $\phi(x)$ is irreducible and $k \geqslant 2$. In the latter case, by the result first mentioned in § 1, there are infinitely many n for which $\phi(n)$ is l/k-th power free, and therefore $f(n)$ is $(l-1)$-th power free.

Suppose then that $f(x) = g(x)h(x)$, where $g(x)$ and $h(x)$ are relatively prime. By hypothesis, neither of them is the $(l-1)$-th power of a polynomial with integral coefficients. For every prime p there exists a residue class $a_p \pmod{p}$ such that $f(a_p) \not\equiv 0 \pmod{p^{l-1}}$; for if $p > 2$ this follows from a result stated in § 1, since p^{l-1} does not divide $l!$, and if $p = 2$ this

was the assumption made. Let t be a large positive integer, and let

$$T = \left(\prod_{p \leqslant t} p\right)^{l-1}.$$

Then there exists a residue class $a_T \pmod{T}$ such that $f(a_T) \not\equiv 0 \pmod{p^{l-1}}$ for each $p \leqslant t$.

The degrees of both $g(x)$ and $h(x)$ are less than l, and neither of these polynomials is an $(l-1)$-th power. It follows from a simple sieve and limiting process that the density of the integers $y \equiv a_T \pmod{T}$ which satisfy $g(y) \equiv 0 \pmod{q^{l-1}}$ for some prime $q > t$ is less than $\frac{1}{2}T^{-1}$. (The proof is similar to that of the result, stated in § 1, that $f(n)$ is l-th power free for infinitely many n.)

The same holds for $h(x)$, and therefore there are infinitely many $y \equiv a_T \pmod{T}$ for which both $g(y)$ and $h(y)$ are $(l-1)$-th power free. From the fact that $g(x)$ and $h(x)$ are relatively prime it follows that the highest common factor of $g(n)$ and $h(n)$ is bounded for all n, and is thus less than t for sufficiently large t. Thus we finally obtain the existence of infinitely many $y \equiv a_T \pmod{T}$ for which $f(y) = g(y)h(y)$ is $(l-1)$-th power free, as was to be proved.

3. We assume from now onwards that $f(x)$ is irreducible, and proceed to explain certain notation which will be used throughout the subsequent work.

Let $\rho(a)$ denote the number of solutions of

$$f(n) \equiv 0 \pmod{a}, \quad 0 < n \leqslant a,$$

and let $\rho_x(a)$ denote the number of solutions of

$$f(n) \equiv 0 \pmod{a}, \quad 0 < n \leqslant x.$$

Plainly
$$\rho_x(a) \leqslant \left[\frac{x}{a}\right] \rho(a) + \rho(a). \tag{1}$$

If p does not divide the discriminant of $f(x)$ it is well known that

$$\rho(p^\gamma) \leqslant l \text{ for all } \gamma > 0. \tag{2}$$

Let x be a sufficiently large positive integer, and let $u_1, u_2, \ldots$ denote the integers satisfying the following three conditions:

(i) $x(\log x)^{-3/2} < u_i < 2x(\log x)^{-3/2}$;

(ii) u_i is squarefree;

(iii) all prime factors of u_i satisfy

$$(\log x)^\beta < p < x(\log x)^{-\beta},$$

where β is a sufficiently large number which will be determined later.

Let $k_1, k_2, \ldots$ denote the integers not exceeding x for which

(i) $f(k_i)$ is divisible by some u,

(ii) $f(k_i) \not\equiv 0 \pmod{p^{l-1}}$ for all primes $p \leqslant (\log x)^{3/2}$.

Let z denote the number of k's, and enumerate the k's in order, so that $k_1 < k_2 < \ldots < k_z \leqslant x$. Let $d^+\big(f(k_i)\big)$ denote the number of divisors of $f(k_i)$ among the u's.

We use $c_1, c_2, \ldots$ to denote positive numbers which depend only on the polynomial $f(x)$.

Occasionally, for convenience of printing, we write λ for $1/(l-1)$.

4. In this section we assume that there are infinitely many positive integers x for which

$$z > x(\log\log x)^{-2}. \tag{3}$$

In this case the proof of the theorem will be comparatively simple. We shall show that if (3) holds for a certain x then the number of k's for which $f(k_i)$ is $(l-1)$-th power free is greater than $\frac{1}{2}x(\log\log x)^{-2}$, and this implies the result.

Suppose p^{l-1} divides $f(k_i)$ for some prime p. We know that for every k_i there is some u which divides $f(k_i)$. If $p > x(\log x)^{-\beta}$ then p and u are relatively prime, by condition (iii) on the u_i. Thus

$$p^{l-1}u \leqslant f(k_i) \leqslant f(x) < c_1 x^l,$$

whence

$$p < (c_1 x^l/u)^\lambda < c_2 x(\log x)^{3\lambda/2}$$

by the lower bound for the u's. In view of condition (ii) on the k's, we have

$$(\log x)^{3/2} < p < c_2 x(\log x)^{3\lambda/2}. \tag{4}$$

The number of integers $k \leqslant x$ for which $f(k) \equiv 0 \pmod{p^{l-1}}$, where p satisfies (4), does not exceed

$$\sum_{(4)} \rho_x(p^{l-1}),$$

where the summation is over primes p satisfying (4). Using (1) and (2), the latter being applicable because p is large, we have

$$\begin{aligned}
\sum_{(4)} \rho_x(p^{l-1}) &\leqslant \sum_{(4)} \left(\frac{x}{p^{l-1}}\rho(p^{l-1}) + \rho(p^{l-1})\right) \\
&< lx \sum_{p > (\log x)^{3/2}} p^{-l+1} + l\pi\left(c_2 x(\log x)^{3\lambda/2}\right) \\
&< c_3 x(\log x)^{-1} + c_4 x(\log x)^{(3\lambda/2)-1} \\
&< \tfrac{1}{2}x(\log\log x)^{-2}.
\end{aligned}$$

It follows from (3) that there are at least $\frac{1}{2}x(\log\log x)^{-2}$ of the k_i for which $f(k_i)$ is not divisible by the $(l-1)$-th power of any prime, and this proves the result.

5. We can now assume that for all sufficiently large x we have

$$z \leqslant x(\log\log x)^{-2}. \tag{5}$$

The proof is based on the following three lemmas, which do not themselves depend on the assumption (5).

LEMMA 1. $$\sum_{i=1}^{z} d^{+}\big(f(k_i)\big) > c_5 x(\log\log x)^{-1}.$$

The proof of this lemma is complicated, and we postpone it to §§ 7-9.

LEMMA 2. *Let p be a prime satisfying $p < (\log x)^{\beta}$. Then*

$$\sum_{(6)} d\big(f(n)\big) < c_6(x\log x)p^{-l+1},$$

where the summation is over positive integers n satisfying

$$n \leqslant x, \quad f(n) \equiv 0 \pmod{p^{l-1}}. \tag{6}$$

The proof is very similar to my proof† that

$$\sum_{n=1}^{x} d\big(f(n)\big) < c_6 x\log x,$$

and can be omitted. The lemma would in fact remain true if the condition on p were relaxed to $p < x^{\lambda-\epsilon}$, except that c_6 would then depend on ϵ.

LEMMA 3. $$\sum_{n=1}^{x} \big\{d\big(f(n)\big)\big\}^2 < x(\log x)^{c_7}.$$

This is a result of van der Corput‡.

6. We now complete the proof of the theorem. Let us assume that for every i with $1 \leqslant i \leqslant z$ we have

$$f(k_i) \equiv 0 \pmod{p^{l-1}} \tag{7}$$

for some prime p (depending on i). We shall show that this leads to a contradiction. The result then follows on giving x a suitable sequence of values. By condition (ii) on the k's, the prime p in (7) must satisfy $p > (\log x)^{3/2}$.

Denote by $k_1' < k_2' < \ldots < k_s' \leqslant x$ those k's for which $f(k)$ is divisible by more than one u. If $f(k_1')$ is divisible by u_1 and u_2 it is impossible that u_1 should divide u_2, in view of condition (i) on the u's. Hence u_2

† *Journal London Math. Soc.*, 27 (1952), 7–15

‡ *Proc. K. Neder. Akad. van Wet.*, Amsterdam, 42 (1939), 547–553.

has at least one prime factor which does not divide u_1, and it follows from conditions (i) and (iii) that

$$[u_1, u_2] > x(\log x)^{\beta-3/2}. \tag{8}$$

If (7) holds for k_i', and $p > x(\log x)^{-\beta}$, then p cannot divide u_1 or u_2, and therefore

$$p^{l-1}[u_1, u_2] \leqslant f(k_i') < c_1 x^l,$$

whence, by (8),

$$p < c_8 x\{(\log x)^{\beta-3/2}\}^{-\lambda} < c_8 x(\log x)^{-\beta/2l}, \tag{9}$$

assuming $\beta > 3$.

We evidently have

$$\sum_{i=1}^{z} d^+\big(f(k_i)\big) \leqslant \sum_{i=1}^{s} d^+\big(f(k_i')\big) + z,$$

since $d^+\big(f(k_i)\big) = 1$ if k_i is not among the set k'. Hence, by (5) and Lemma 1,

$$\sum_{i=1}^{s} d^+\big(f(k_i')\big) > c_5 x(\log\log x)^{-1} - x(\log\log x)^{-2} > c_9 x(\log\log x)^{-1}. \tag{10}$$

Denote now by $t_1 < t_2 < \ldots < t_w \leqslant x$ those numbers k' with the property that $f(t_i)$ is divisible by p^{l-1} for some $p > (\log x)^{\beta}$. Then p satisfies (9). If k_i' is not one of the numbers t then $f(k_i')$ satisfies (6) for some prime p with $(\log x)^{3/2} < p < (\log x)^{\beta}$. Hence

$$\begin{aligned}\sum_{i=1}^{s} d^+\big(f(k_i')\big) &\leqslant \sum_{p}\sum_{(6)} d\big(f(n)\big) + \sum_{i=1}^{w} d\big(f(t_i)\big)\\ &= \Sigma_1 + \Sigma_2,\end{aligned} \tag{11}$$

say, where p in Σ_1 satisfies the inequalities just stated. Using Lemma 2, we obtain

$$\Sigma_1 < c_6 x \log x \sum_{p} p^{-l+1} < c_{10} x(\log x)^{-1/2}. \tag{12}$$

Also, in view of the definition of the numbers t and the inequality (9), we have

$$w \leqslant \sum_{(13)} \rho_x(p^{l-1}),$$

where the summation here is over

$$(\log x)^{\beta} \leqslant p < c_8 x(\log x)^{-\beta/2l}. \tag{13}$$

Thus, by (1) and (2),

$$\begin{aligned} w &\leqslant \sum_{(13)} (xp^{-l+1}+1)\rho(p^{l-1})\\ &\leqslant lx \sum_{(13)} p^{-l+1} + l \sum_{(13)} 1\\ &< c_{11} x(\log x)^{-\beta} + c_{12} x(\log x)^{-1-\beta/2l}\\ &< c_{13} x(\log x)^{-1-c_7}\end{aligned}$$

provided $\beta > 2l(1+c_7)$. Consequently, by Lemma 3 and Cauchy's inequality, we have

$$\Sigma_2 = \sum_{i=1}^{w} d\left(f(t_i)\right) \leqslant \{wx(\log x)^{c_7}\}^{1/2} < c_{14}\, x(\log x)^{-1/2}. \tag{14}$$

But now the estimates for Σ_1 and Σ_2, when substituted from (12) and (14) in (11), give a contradiction to (10). As was seen earlier, this contradiction proves the theorem.

7. We now come to the proof of Lemma 1. We recall that $d^+\left(f(k_i)\right)$ denotes the number of divisors of $f(k_i)$ among the u's. Hence, interchanging the order of summation, we have

$$\sum_{i=1}^{z} d^+\left(f(k_i)\right) = \sum_j \rho_x'(u_j), \tag{15}$$

where $\rho_x'(u_j)$ denotes the number of positive integers $n \leqslant x$ satisfying

$$f(n) \equiv 0 \pmod{u_j}, \quad f(n) \not\equiv 0 \pmod{p^{l-1}} \text{ for all } p \leqslant (\log x)^{3/2}. \tag{16}$$

The proof of Lemma 1 falls into two stages.

LEMMA 1A. $\rho_x'(u_j) > c_{15}\rho_x(u_j)$ *for all* j.

LEMMA 1B. $\sum_j \rho_x(u_j) > c_{16}\, x(\log\log x)^{-1}$.

In view of (15), these two results imply Lemma 1.

8. *Proof of Lemma* 1A. Let D denote the discriminant of $f(x)$, and let t be a large but fixed positive integer greater than D. Defining T as in § 2, we recall that there exists a residue-class $a_T \pmod T$ such that $f(a_T) \not\equiv 0 \pmod{p^{l-1}}$ for all $p \leqslant t$. Let $\rho_x''(u_j)$ denote the number of positive integers $n \leqslant x$ satisfying

$$f(n) \equiv 0 \pmod{u_j}, \quad n \equiv a_T \pmod T. \tag{17}$$

On comparing (17) with (16), we see that

$$\rho_x'(u_j) \geqslant \rho_x''(u_j) - \sum_p \rho_x''(p^{l-1}u_j), \tag{18}$$

where the summation is over primes p satisfying $t \leqslant p \leqslant (\log x)^{3/2}$.

Since the u's are composed of large primes, we have $(u_j, T) = 1$. Since also $u_j \leqslant x$, it follows easily from the definition of $\rho_x''(u_j)$ in (17) that

$$\rho_x''(u_j) \geqslant \frac{x\rho(u_j)}{2Tu_j}. \tag{19}$$

In the terms of the sum on the right of (18), we have $(p, u_j) = 1$ by condition (iii) on the u's, provided $\beta > \frac{3}{2}$. It follows easily that

$$\rho_x''(p^{l-1}u_j) \leqslant \left[\frac{x}{p^{l-1}u_j T}\right]\rho(p^{l-1}u_j) + \rho(p^{l-1}u_j)$$

$$\leqslant \frac{lx\rho(u_j)}{p^{l-1}u_j T} + l\rho(u_j), \tag{20}$$

since $\rho(p^{l-1}) \leqslant l$ by (2), in view of the fact that $p > D$.

Using (19) and (20) in (18), we obtain

$$\rho_x'(u_j) \geqslant \frac{x}{2Tu_j}\rho(u_j)\left\{1 - 2l\sum_{p>l} p^{-l+1}\right\} - l\rho(u_j)\pi\left((\log x)^{3/2}\right)$$

$$> \frac{x\rho(u_j)}{4Tu_j} - c_{17}\rho(u_j)(\log x)^{3/2}(\log\log x)^{-1}$$

$$> \frac{x\rho(u_j)}{5Tu_j},$$

since $u_j < 2x(\log x)^{-3/2}$. Since

$$\rho_x(u_j) \leqslant \frac{2x}{u_j}\rho(u_j)$$

by (1), Lemma 1A now follows.

9. *Proof of Lemma* 1B. We have

$$\sum_j \rho_x(u_j) \geqslant \sum_j \frac{x}{2u_j}\rho(u_j) > \tfrac{1}{4}(\log x)^{3/2}\sum_j \rho(u_j).$$

Hence to prove Lemma 1B it suffices to show that

$$\sum_j \rho(u_j) > c_{18}x(\log x)^{-3/2}(\log\log x)^{-1}. \tag{21}$$

Put $y = x(\log x)^{-3/2}$. Consider all numbers of the form vp, where v is a squarefree positive integer not exceeding $x^{1/3}$ whose prime factors all lie between $(\log x)^{\beta}$ and x^{ϵ}, and p is a prime satisfying

$$y/v < p < 2y/v.$$

Here ϵ is a sufficiently small positive number which will be chosen later. The numbers vp are distinct and each of them is a u. By the multiplicative property of $\rho(m)$, we have

$$\sum_j \rho(u_j) \geqslant \sum_v \rho(v)\sum_p \rho(p), \tag{22}$$

where the summations are over v and p as just defined.

By the prime ideal theorem*, we have

$$\sum_{p<z} \rho(p) = z(\log z)^{-1} + O\left(z(\log z)^{-2}\right); \tag{23}$$

and from this it follows that the inner sum on the right of (22) satisfies

$$\sum_{p} \rho(p) > c_{19} y v^{-1} (\log y)^{-1}.$$

Thus, by (22),

$$\sum_{j} \rho(u_j) > c_{19} y (\log y)^{-1} \sum_{v} \rho(v)/v.$$

Hence in order to prove (21) it will suffice to show that

$$\sum_{v} \rho(v)/v > c_{20} \log x (\log\log x)^{-1}. \tag{24}$$

The following simple proof of (24) was suggested to me by the referee. We have, for $s > 1$,

$$\sum_{v} \rho(v) v^{-s} = \prod_{p} \left(1+\rho(p) p^{-s}\right) - \sum_{w} \rho(w) w^{-s}, \tag{25}$$

where p runs through all primes satisfying $(\log x)^{\beta} < p < x^{\alpha}$, and w runs through the squarefree integers greater than $x^{1/3}$ which are entirely composed of such primes. Plainly

$$\prod_{p<(\log x)^{\beta}} \left(1+\rho(p) p^{-s}\right) \sum_{w} \rho(w) w^{-s} < \sum_{n>x^{1/3}} \rho(n) n^{-s}. \tag{26}$$

Using the well known result†

$$s_m = \sum_{n=1}^{m} \rho(n) < c_{21} m,$$

we obtain by partial summation

$$\sum_{n>x^{1/3}} \rho(n) n^{-s} \leqslant \sum_{m>x^{1/3}} s_m \left(m^{-s} - (m+1)^{-s}\right)$$

$$< c_{22} \frac{s}{s-1} x^{(1-s)/3}. \tag{27}$$

Now put $s = 1+(\epsilon \log x)^{-1}$. Using (23) we find that, for $z < x^{\alpha}$,

$$\log\log z - c_{23} < \sum_{p<z} \rho(p) p^{-s} < \log\log z + c_{24},$$

where c_{23} and c_{24} are independent of ϵ. It follows that

$$c_{25} \log z < \prod_{p<z} \left(1+\rho(p) p^{-s}\right) < c_{26} \log z.$$

Using this, with (27), in (26), we obtain

$$\sum \rho(w) w^{-s} < c_{27} (\epsilon \log x) e^{-1/(3\epsilon)} (\log\log x)^{-1}.$$

---

* See Lemma 7 of my paper cited earlier.

† This is a consequence of the prime ideal theorem.

Also the product on the right of (25) satisfies

$$\prod_p \left(1+\rho(p)p^{-s}\right) > \frac{c_{25}\,\epsilon \log x}{c_{26}\beta \log\log x}.$$

Hence, if ϵ is sufficiently small, we have from (25)

$$\sum_v \rho(v)/v \geqslant \sum_v \rho(v)\,v^{-s} > c_{28}(\epsilon)\log x(\log\log x)^{-1}.$$

This proves (24), and so completes the proof of Lemma 1B.

10. My original proof of Lemma 1B was very much more complicated. It depended on the following lemma, which may be of some interest in itself and is therefore stated here without proof.

Let Z be large and let $q_1, q_2, \ldots$ be any primes not exceeding Z. Let $w_1, w_2, \ldots$ be all the squarefree integers composed of these primes. Let there correspond to each q_i a real number α_i satisfying $1 \leqslant \alpha_i \leqslant l$. For each w_i define

$$g(w_i) = \prod_{q_j | w_i} \alpha_j.$$

Then, if $k > 1$,

$$\sum_{w_i > Z^k} g(w_i)/w_i < \frac{1}{[c_{29}k]!}\sum_{w_i} g(w_i)/w_i = \frac{1}{[c_{29}k]!}\prod_j (1+\alpha_j/q_j),$$

where c_{29} depends only on l.

I conclude with two remarks bearing on the present paper. It was proved by Estermann* that every sufficiently large positive integer is the sum of a square and a squarefree integer, and by similar methods one can prove that every sufficiently large positive integer is the sum of an l-th power and an l-th power free integer. The methods of the present paper would doubtless allow one to prove that every sufficiently large positive integer is the sum of an l-th power and an $(l-1)$-th power free integer.

It is possible to prove that there are infinitely many primes p for which $f(p)$ is l-th power free, provided of course that $f(x)$ is not the l-th power of a linear polynomial. It is reasonable to conjecture that there are infinitely many primes p for which $f(p)$ is $(l-1)$-th power free, assuming that $f(x)$ satisfies the conditions of § 1; but the methods of the present paper do not seem to be powerful enough to prove this. I have also not been able to prove, for example, that n^4+2 is squarefree for infinitely many n.

Department of Mathematics,
University College, London.

* *Math. Annalen*, 105 (1931), 653–662.

Printed by C. F. Hodgson & Son, Ltd., Pakenham Street, London, W.C.1.