ACTA ARITHMETICA
XLIX (1987)

On the number of terms of a power of a polynomial

by

A. SCHINZEL (Warszawa)

To Paul Erdős
with best wishes on his
75th birthday

The conjecture made by Rényi and first published by Erdős [2], who supported it (1), asserts that if Q_k is the least number of non-zero coefficients of the square of a polynomial with exactly k non-zero complex coefficients then

$$\lim_{k=\infty} Q_k = \infty .$$

It has been proved by Erdős in the quoted paper that

$$Q_k < C_1 k^{1-C_2}$$

and the values of the positive constants C_1 and C_2 have been subsequently found by Verdenius [9] (see also Freud [3]). He also established a similar inequality for cubes. It is the principal aim of the present paper to prove an estimate for the number of non-zero coefficients, called the number of terms, of an arbitrary power of a polynomial, which contains as a special case the inequality

$$Q_k > \frac{\log\log k}{\log 2}.$$

Here is the general result.

THEOREM 1. *Let K be a field, $f \in K[x]$, $l \in \mathbb{N}$, f and f^l have $T \geqslant 2$ and t terms, respectively. If either* $\operatorname{char} K = 0$ *or* $\operatorname{char} K > l \deg f$ *then*

$$t \geqslant l + 1 + (\log 2)^{-1} \log\left(1 + \frac{\log(T-1)}{l\log 4l - \log l}\right).$$

Already for $l = 2$ there is a big gap between the obtained lower bound and Erdős's upper bound for t. Another open question concerns the number

(1) Erdős tells me that he arrived at the conjecture independently from Rényi.

of terms of $F(f(x))$, where F is a fixed non-constant polynomial. If $Q_k(F)$ is the minimal number of terms of $F(f(x))$, when f runs over all polynomials with exactly k terms then probably $\lim_{k=\infty} Q_k(F) = \infty$, but the method of this paper is insufficient to prove it.

If $\operatorname{char} K$ is positive the number of terms of f_n^l may remain bounded in spite of the fact that the number of terms of $f_n \in K[x]$ tends to infinity with n. The situation is described by the following

THEOREM 2. *Let* $\operatorname{char} K > 0$, $f \in K[x]$, $l \in N$, f *and* f^l *have* $T \geqslant 2$ *and* t *terms, respectively. If*

$$l^{T-1}(T^2 - T + 2) < \operatorname{char} K$$

then

$$t \geqslant l + 1 + (\log 2)^{-1} \log\left(1 + \frac{\log(T-1)}{l \log 4l - \log l}\right).$$

On the other hand, if $l \neq (\operatorname{char} K)^n$ $(n = 0, 1, 2, \ldots)$ *there exist polynomials* $f \in K[x]$ *with* T *arbitrarily large such that* $t \leqslant 2l$.

Finally we have

THEOREM 3. *Let* K *be a field and* $f \in K[x]$. *If in the algebraic closure of* K f *has a zero* ξ *of multiplicity exactly* n *then* f *has at least as many terms as* $(x - \xi)^n$.

The algebraic closure of K will be denoted by $\hat{K}$. The case $\operatorname{char} K = 0$ of Theorem 3 has been proved by G. Hajós [5]. The special case of Theorem 3 for $K = F_2$ and $\xi = 1$ has been given as a problem in XXVIth International Mathematical Olympiad. A. Mąkowski, the head of the Polish delegation insisted that there should be a common generalization of this problem and of Hajós's theorem. Hajós's result, slightly extended serves as the first of the three lemmata needed for the proof of Theorem 1.

LEMMA 1. *If* $g \in K[x] \setminus \{0\}$ *has in the algebraic closure of* K *a zero* $\xi \neq 0$ *of multiplicity at least* m *and either* $\operatorname{char} K = 0$ *or* $\operatorname{char} K > \deg g$, *then* g *has at least* $m+1$ *terms.*

Proof. The proof given by Hajós [5] and rediscovered by Montgomery and Schinzel [6] (Lemma 1) for $K = C$ applies without change to the case $\operatorname{char} K = 0$ or $\operatorname{char} K > \deg g$.

LEMMA 2. *If* $f(x) \in K[x]$, $f(0) \neq 0$, $f(x)^l \in K[x^d]$ *then either* $\operatorname{char} K | (l, d)$ *or* $f(x) \in K[x^d]$.

Proof. Let

$$f(x)^l = g(x^d), \qquad g(x) = \gamma_0 \prod_{\gamma \in \Gamma} (x - \gamma)^{e(\gamma)},$$

where Γ is a subset of $K \setminus \{0\}$. We get

$$f(x)^l = \gamma_0 \prod_{\gamma \in \Gamma} (x^d - \gamma)^{e(\gamma)}.$$

Since for $\gamma \neq 0$ the multiplicity of the zeros of $x^d - \gamma$ is either 1 or equal to the maximal power of $\operatorname{char} K$ dividing d, we get either $\operatorname{char} K | (l, d)$ or $l | e(\gamma)$ for all $\gamma \in \Gamma$. It follows that

$$f(x) = \gamma_1 \prod_{\gamma \in \Gamma} (x^d - \gamma)^{e(\gamma)/l} \in K[x^d].$$

Since $K[x] \cap \hat{K}[x^d] = K[x^d]$ we infer that

$$f(x) \in K[x^d].$$

LEMMA 3. *Let* $H \in K[y, z]$, $p \in Z$. *Define the sequence* $H_n = H_n(y, z; p)$ *as follows*

$$H_0 = H, \qquad H_{n+1} = \frac{\partial H_n}{\partial y} py + \frac{\partial H_n}{\partial z} z.$$

Then we have the following

(1) $$\deg_y H_n \leqslant \deg_y H, \qquad \deg_z H_n \leqslant \deg_z H;$$

(2) $$H_n(x^p, x; p) = \sum_{k=1}^{n} c(k, n) x^k \frac{d^k H(x^p, x; p)}{dx^k} \qquad (n \geqslant 1)$$

for suitable coefficients $c(k, n) \in K$;

(3) *If* $\operatorname{char} K \geqslant l$ *and a polynomial* G *irreducible over* K *divides* $(H_0, H_1, \ldots, H_{l-1})$, *then either* $G^l | H$ *or for each term* $gy^\alpha z^\beta$ *of* G $(g \neq 0)$ $p\alpha + \beta$ *is the same* mod $\operatorname{char} K$ *if* $\operatorname{char} K > 0$, *has the same value if* $\operatorname{char} K = 0$, *briefly* G *is isobaric* mod $\operatorname{char} K$ *with respect to the weights* p, 1.

Proof. Directly from the definition of H_n we get

$$\deg_y H_{n+1} \leqslant \deg_y H_n, \qquad \deg_z H_{n+1} \leqslant \deg_z H_n$$

and formulae (1) follow by induction. The same method is used to prove (2) and (3).

(2) is true for $n = 1$ since

$$H_1(x^p, x; p) = \frac{\partial H}{\partial y}(x^p, x) px^p + \frac{\partial H}{\partial z}(x^p, x) x = x \frac{dH(x^p, x)}{dx}.$$

Assuming the truth of (2) for a fixed n we get

$$H_{n+1}(x^p, x; p) = \frac{\partial H_n}{\partial y}(x^p, x; p) px^p + \frac{\partial H_n}{\partial z}(x^p, x; p) x$$

$$= x \frac{dH_n(x^p, x; p)}{dx}$$

$$= x \sum_{k=1}^{n} c(k, n) \left(kx^{k-1} \frac{d^k H(x^p, x; p)}{dx^k} + x^k \frac{d^{k+1} H(x^p, x; p)}{dx^{k+1}} \right),$$

which implies (2) with n replaced by $n+1$.

In order to prove (3) let $H = G^m U$, where $U \not\equiv 0 \bmod G$. We shall show by induction on $j \leqslant m$ that

$$H_j(y, z; p) \equiv j!\binom{m}{j}\left(\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z\right)^j G^{m-j}U \bmod G^{m-j+1}. \tag{4}$$

For $j=0$ this is obviously true. Assuming it for a fixed j we get upon differentiation

$$\frac{\partial H_j}{\partial y} \equiv j!\binom{m}{j}\left(\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z\right)^j (m-j)\,G^{m-j-1}\frac{\partial G}{\partial y}U \bmod G^{m-j},$$

$$\frac{\partial H_j}{\partial z} \equiv j!\binom{m}{j}\left(\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z\right)^j (m-j)\,G^{m-j-1}\frac{\partial G}{\partial z}U \bmod G^{m-j},$$

hence

$$H_{j+1}(y, z; p) = \frac{\partial H_j}{\partial y}py+\frac{\partial H_j}{\partial z}z$$

$$\equiv (j+1)!\binom{m}{j+1}\left(\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z\right)^{j+1} G^{m-j-1}U \bmod G^{m-j}$$

and the inductive proof of (4) is complete.

Taking there $j=m$, we get

$$H_m(y; z; p) \equiv m!\left(\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z\right)^m U \bmod G,$$

hence if $m<l$ the assumption $G \mid (H_0, H_1, \ldots, H_{l-1})$ implies

$$\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z \equiv 0 \bmod G.$$

However the degree of $\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z$ does not exceed the degree of G. Hence

$$\frac{\partial G}{\partial y}py+\frac{\partial G}{\partial z}z = cG, \quad c \in \boldsymbol{K}$$

and for each term $gy^\alpha z^\beta$ $(g \neq 0)$ of G we have

$$p\alpha+\beta = c,$$

where both sides are viewed as elements of $\boldsymbol{K}$. If $\operatorname{char} \boldsymbol{K} > 0$ this means

$$p\alpha+\beta \equiv c \,(\operatorname{mod} \operatorname{char} \boldsymbol{K})$$

and if $\operatorname{char} \boldsymbol{K} = 0$

$$p\alpha+\beta = c.$$

Proof of Theorem 1. We shall prove the following equivalent inequality

$$T \leqslant 1+\left(\frac{(4l)^l}{l}\right)^{2^{t-l-1}-1}. \tag{5}$$

For $T>1$ we have $t>1$ hence (5) holds for $t=1$. For $t>1$ let

$$f(x)^l = \sum_{j=0}^{t-1} a_j x^{m_j},$$

where

$$a_j \neq 0, \quad m_0 < m_1 < \ldots < m_{t-1}, \quad (m_1-m_0, m_2-m_0, \ldots, m_{t-1}-m_0) = d.$$

We have

$$m_0 = l \operatorname{ord}_x f \equiv 0 \bmod l,$$

$$\left(f(x)\,x^{-m_0/l}\right)^l \in \boldsymbol{K}[x^d], \quad f(x)\,x^{-m_0/l}|_{x=0} \neq 0,$$

hence by Lemma 2

$$f(x)\,x^{-m_0/l} \in \boldsymbol{K}[x^d], \quad f(x) = f_0(x^d)\,x^{m_0/l}$$

and

$$f_0(x)^l = a_0+\sum_{j=1}^{t-1} a_j x^{n_j}, \tag{6}$$

where $n_j = (m_j-m_0)/d$. We get

$$0 = n_0 < n_1 < n_2 < \ldots < n_{t-1} \leqslant l \deg f, \quad (n_1, \ldots, n_{t-1}) = 1 \tag{7}$$

and since f and f_0 have the same number of terms it is enough to prove the inequality (5) for the number of terms of f_0.

If $t \leqslant l+1$ we apply Lemma 1. Since $\operatorname{char} \boldsymbol{K} = 0$ or $\operatorname{char} \boldsymbol{K} > n_{t-1}$ the lemma is applicable with $g = f_0^l$, $m = l$ and it gives $t \geqslant l+1$, hence $t = l+1$. Every zero ξ of f_0^l is of multiplicity $\geqslant l$, hence on differentiation

$$a_0+\sum_{j=1}^{l} a_j \xi^{n_j} = 0, \quad \sum_{j=1}^{l} a_j \binom{n_j}{i}\xi^{n_j} = 0 \quad (1 \leqslant i < l).$$

Since $\operatorname{char} \boldsymbol{K} = 0$ or $\operatorname{char} \boldsymbol{K} > n_{t-1}$ we have

$$\left|\binom{n_j}{i}\right|_{\substack{0 \leqslant i \leqslant l \\ 1 \leqslant j \leqslant l}} = \prod_{0 \leqslant q < r < l} \frac{n_r-n_q}{r-q} \neq 0,$$

hence $a_j \xi^{n_j}$ are uniquely determined by a_0. Since $a_j \neq 0$ and $(n_1, \ldots, n_{t-1}) = 1$ there is only one possible value for ξ. Then

$$f_0(x) = c(x-\xi)^{\deg f_0}, \quad c \in \boldsymbol{K}, \ \xi \neq 0$$

and Lemma 1 applies with $g = f_0$, $m = l\deg f_0$. It gives $l\deg f_0 + 1 \leqslant l+1$, $\deg f_0 = 1$, $T = 2$, hence (5).

The further proof proceeds by induction for fields K algebraically closed. Assume that (5) holds for lth powers with less than $t \geqslant l+2$ terms and consider again the conditions (6) and (7).

By Dirichlet's theorem there exist integers $p_1, p_2, \ldots, p_{t-1}$ such that

$$\left|\frac{n_j}{n_{t-1}} - \frac{p_j}{p_{t-1}}\right| < \frac{1}{4lp_{t-1}} \quad (j = 1, 2, \ldots, t-2) \tag{8}$$

and

$$0 < p_{t-1} \leqslant (4l)^{t-2}.$$

The inequality $p_i < 0$ or $p_i > p_{t-1}$ would imply

$$\frac{1}{p_{t-1}} < \left|\frac{n_i}{n_{t-1}} - \frac{p_i}{p_{t-1}}\right| < \frac{1}{4lp_{t-1}},$$

a contradiction; hence we have

$$0 \leqslant p_j \leqslant p_{t-1} \leqslant (4l)^{t-2} \quad (j = 1, 2, \ldots, t-2). \tag{9}$$

Setting

$$p_{t-1}[n_1, \ldots, n_{t-1}] = n_{t-1}[p_1, \ldots, p_{t-1}] + [r_1, \ldots, r_{t-1}] \tag{10}$$

we get from (8)

$$|r_j| < \frac{n_{t-1}}{4l} \quad (j = 1, 2, \ldots, t-2), \quad r_{t-1} = 0.$$

If $\max\limits_{1 \leqslant i \leqslant t-2} |r_i| = 0$, then by (9), (7) and (10)

$$(4l)^{t-2} \geqslant p_{t-1} = (p_{t-1}n_1, \ldots, p_{t-1}n_{t-1}) \geqslant n_{t-1},$$

hence

$$T \leqslant 1 + \deg f_0 = 1 + \frac{n_{t-1}}{l} \leqslant 1 + \frac{(4l)^{t-2}}{l} \leqslant 1 + \left(\frac{(4l)^l}{l}\right)^{2^{t-l-1}-1}.$$

Therefore, assume that

$$0 < \max_{1 \leqslant j \leqslant t-1} |r_j| < \frac{n_{t-1}}{4l}, \quad r_{t-1} = 0 \tag{11}$$

and put

$$r = \min_{1 \leqslant j \leqslant t-1} r_j, \quad F(y, z) = z^{-r}\left(a_0 + \sum_{j=1}^{t-1} a_j y^{p_j} z^{r_j}\right).$$

By (9) and the choice of r we have

$$F(y, z) \in K[y, z], \quad (F(y, z), yz) = 1.$$

(Note that by (7) and (8) no two terms of F are similar.) By (6) and (8) we have

$$f_0(x^{p_t-1})^l = x^r F(x^{n_t-1}, x). \tag{12}$$

Let

$$F(y, z) = F_0(y, z)^l H(y, z); \quad F_0, H \in K[y, z], \tag{13}$$

where H is not divisible by the lth power of any polynomial in $K[y, z] \setminus K$. It follows from (12) and (13) that every zero of $H(x^{n_t-1}, x)$ except possibly $x = 0$ is at least l-tuple. Hence for any $\xi \in \hat{K} \setminus \{0\}$

$$\operatorname{ord}_{x-\xi} H(x^{n_t-1}, x) \leqslant l \operatorname{ord}_{x-\xi} \frac{d^k}{dx^k} H(x^{n_t-1}, x) \quad (k < l)$$

and by (2) with $p = n_{t-1}$

$$\operatorname{ord}_{x-\xi} H(x^{n_t-1}, x) \leqslant l \operatorname{ord}_{x-\xi} H_m(x^{n_t-1}, x; n_{t-1}) \quad (m < l).$$

Also, by (2)

$$\operatorname{ord}_x H(x^{n_t-1}, x) \leqslant \operatorname{ord}_x H_m(x^{n_t-1}, x; n_{t-1}).$$

Thus finally

$$H(x^{n_t-1}, x) | H_m(x^{n_t-1}, x; n_{t-1})^l \quad (1 \leqslant m < l)$$

and for indeterminates $u_1, \ldots, u_{t-1}$

$$H(x^{n_t-1}, x) \Big| \sum_{m=1}^{l-1} u_m H_m(x^{n_t-1}, x; n_{t-1})^l. \tag{14}$$

Suppose first that $(H, H_1, \ldots, H_{t-1}) \neq 1$, where H_m stands for $H_m(y, z; n_{t-1})$. Then by the choice of H and the assertion (3) of Lemma 3 H, hence also F, has a factor $G \notin K$ isobaric mod char K with respect to the weights n_{t-1}, 1. Since $(F, yz) = 1$ G has at least two terms. Let

$$F/G = \sum_{i=1}^{n} G_i,$$

where G_i are polynomials isobaric mod char K with respect to the weights n_{t-1}, 1 and n is minimal. Since G is isobaric mod char K with respect to the weights n_{t-1}, 1

$$F = \sum_{i=1}^{n} GG_i$$

is the corresponding representation of F. Since G has at least two terms, the same is true for GG_1 hence F has at least two terms with weights congruent mod char K, if char $K > 0$, equal if char $K = 0$. However the weights of the terms of F are $p_j n_{t-1} + r_j - r = p_{t-1} n_j - r$ $(0 \leqslant j < t)$. Since n_k are distinct the equality $p_{t-1} n_i - r = p_{t-1} n_j - r$, with $i \neq j$, is impossible. The congruence $p_{t-1} n_i - r \equiv p_{t-1} n_j - r (\mathrm{mod\ char}\ K)$ implies $p_{t-1} \equiv 0 (\mathrm{mod\ char}\ K)$ or $n_i \equiv n_j (\mathrm{mod\ char}\ K)$. Since char $K = 0$ or char $K > n_{t-1}$ the latter case with $i \neq j$ is impossible and we get

$$0 < \mathrm{char}\ K \leqslant p_{t-1}.$$

Hence by (9)

$$T \leqslant 1 + \deg f < 1 + \frac{\mathrm{char}\ K}{l} \leqslant 1 + \frac{p_{t-1}}{l} \leqslant 1 + \frac{(4l)^{t-2}}{l} \leqslant 1 + \left(\frac{(4l)^l}{l}\right)^{2^{t-l-1}-1}$$

and (5) holds.

Suppose now that $(H, H_1, \ldots, H_{t-1}) = 1$. Then

$$(H, \sum_{m=1}^{t-1} u_m H_m^l) = 1.$$

Therefore the resultant R of H and $\sum_{m=1}^{t-1} u_m H_m^l$ with respect to y is non-zero and in view of (14)

$$H(x^{n_t - 1}, x) \mid R(x).$$

Now, the degree of R does not exceed

$$\deg_y H \deg_z \sum_{m=1}^{t-1} u_m H_m^l + \deg_z H \deg_y \sum_{m=1}^{t-1} u_m H_m^l.$$

In virtue of (1) we get

$$\deg R \leqslant 2l \deg_y H \deg_z H.$$

On the other hand, if there is no cancellation in $H(x^{n_t-1}, x)$ we have

$$\deg H(x^{n_t-1}, x) \geqslant \max(n_{t-1} \deg_y H, \deg_z H).$$

It follows that either

$$\deg_y H = \deg_z H = 0 \tag{15}$$

or

$$n_{t-1} \leqslant 2l \deg_z H \leqslant 2l \deg_z F \leqslant 2l(\max r_i - \min r_i) < n_{t-1}$$

by (11), a contradiction.

If there is a cancellation in $H(x^{n_t-1}, x)$ then $\deg_y H \neq 0$ and

$$n_{t-1} \leqslant \deg_z H \leqslant \deg_z F < n_{t-1},$$

a contradiction again. Thus we have (15), i.e. $H \in K$ and so by (13)

$$F(y, z) = \mathrm{const}\, F_0(y, z)^l; \tag{16}$$

by (12)

$$f_0(x^{p_t-1})^l = \mathrm{const}\, x^r F_0(x^{n_t-1}, x)^l;$$

$$F_0(x^{n_t-1}, x) = \mathrm{const}\, x^{-r/l} f_0(x^{p_t-1}).$$

The number of terms of $F(y, z)$ is t, the number of terms of $F_0(y, z)$ is $T_0 \geqslant T$. Let

$$F_0(y, z) = \sum_{\tau=1}^{T_0} b_\tau y^{\alpha_\tau} z^{\beta_\tau}, \qquad \langle \alpha_\tau, \beta_\tau \rangle \text{ all different}, \ b_\tau \neq 0.$$

By (11) there exists an index $i < t-1$ such that

$$\begin{vmatrix} p_i & p_{t-1} \\ r_i & r_{t-1} \end{vmatrix} = -p_{t-1} r_i \neq 0,$$

hence

$$T_0 = \mathrm{card}\,\{\langle \alpha_\tau r_i - \beta_\tau p_i, \alpha_\tau r_{t-1} - \beta_\tau p_{t-1} \rangle : \tau \leqslant T_0\}.$$

Now, for $j = i$ or $t-1$ let

$$T_j = \mathrm{card}\,\{\alpha_\tau r_j - \beta_\tau p_j : \tau \leqslant T_0\}.$$

Clearly $T_i T_{t-1} \geqslant T_0$, hence for a suitable $k \in \{i, t-1\}$

$$T_k^2 \geqslant T_0. \tag{17}$$

Now, let us choose elements η, ζ of K such that all non-empty sums

$$\sum_{\alpha_\tau r_k - \beta_\tau p_k = \mathrm{const}} b_\tau \eta^{\alpha_\tau} \zeta^{\beta_\tau}$$

are non-zero. Then T_k is the number of terms of $F_0(\eta x^{r_k}, \zeta x^{-p_k})$. Let

$$s = \mathrm{ord}_x F_0(\eta x^{r_k}, \zeta x^{-p_k}), \qquad G(x) = x^{-s} F_0(\eta x^{r_k}, \zeta x^{-p_k}) \in K[x].$$

We have by (16)

$$G(x)^l = \mathrm{const}\, x^{-ls} F(\eta x^{r_k}, \zeta x^{-p_k}) \tag{18}$$

$$= \mathrm{const}\, \zeta^{-r} x^{p_k r - ls} (a_0 + \sum_{j=1}^{t-1} a_j \eta^{p_j} \zeta^{r_j} x^{p_j r_k - r_j p_k})$$

and the number of terms of $G(x)^l$ is at most $t-1$ since two terms in the

parenthesis on the right hand side of (18), namely a_0 and $a_k \eta^{p_k} \zeta^{r_k} x^{p_k r_k - r_k p_k}$, coalesce. Moreover we have

$$p_j r_k - r_j p_k = p_{t-1}(p_j n_k - n_j p_k) \quad \text{for all } j < t;$$

thus

$$G(x)^l x^{ls - p_k r} \in \boldsymbol{K}(x^{p_{t-1}}).$$

Since $G(0) \neq 0$ we get from (18) and the above

$$\min_{1 \leqslant j < t} (p_j r_k - r_j p_k) \leqslant ls - p_k r \equiv 0 \pmod{p_{t-1}}, \tag{19}$$

hence

$$G(x)^l \in \boldsymbol{K}[x^{p_{t-1}}].$$

In virtue of Lemma 2

$$G(x) \in \boldsymbol{K}[x^{p_{t-1}}]; \quad G(x) = G_0(x^{p_{t-1}}), \quad G_0 \in \boldsymbol{K}[y].$$

The number of terms of $G_0(x)^l$ is the same as that of G^l, hence at most $t-1$. Moreover by (18), (19), (9) and (11)

$$l \deg G_0 = \frac{l \deg G}{p_{t-1}} \leqslant \frac{1}{p_{t-1}} \Big(\max_{1 \leqslant j < t} (p_j r_k - r_j p_k) - \min_{1 \leqslant j < t} (p_j r_k - r_j p_k) \Big)$$

$$< \frac{4n_{t-1}}{4l} \leqslant n_{t-1} < \operatorname{char} \boldsymbol{K},$$

unless $\operatorname{char} \boldsymbol{K} = 0$. The inductive assumption applies and since the number of terms of G_0 is equal to that of G we get

$$T_k \leqslant 1 + \left(\frac{(4l)^l}{l} \right)^{2^{t-l-2}-1}.$$

Hence by (17)

$$T \leqslant T_0 \leqslant T_k^2 \leqslant 1 + 2\left(\frac{(4l)^l}{l} \right)^{2^{t-l-2}-1} + \left(\frac{(4l)^l}{l} \right)^{2^{t-l-1}-2} < 1 + \left(\frac{(4l)^l}{l} \right)^{2^{t-l-1}-1}$$

and the inductive proof is complete. The assumption that $\boldsymbol{K}$ is algebraically closed does not diminish the generality.

LEMMA 4. *Let $\boldsymbol{K}$ be any field, $\boldsymbol{U}$ a finite subset of $\boldsymbol{K}$ and $P \in \boldsymbol{K}[t_1, \ldots, t_r] \setminus \{0\}$. The equation $P(t_1, \ldots, t_r) = 0$ has no more than $\deg P (\operatorname{card} \boldsymbol{U})^{r-1}$ solutions $(t_1, \ldots, t_r) \in \boldsymbol{U}^r$.*

Proof. This is Lemma 8 in [8], p. 302.

LEMMA 5. *Let p be a prime, $N = \sum_{\nu=0}^{n} c_\nu p^\nu$, where $0 \leqslant c_\nu < p$. The number of coefficients of $(x+1)^N$ non-divisible by p equals $\prod_{\nu=0}^{n} (c_\nu + 1)$.*

Proof. This is an immediate consequence of a theorem of Lucas about binomial coefficients (see [1], p. 114).

Proof of Theorem 2. Put

$$f(x) = \sum_{j=1}^{T} A_j x^{N_j}, \quad N_j \text{ all different}, \ A_j \neq 0 \ (1 \leqslant j \leqslant T)$$

and let us assign two vectors $[i_1, i_2, \ldots, i_l], [j_1, j_2, \ldots, j_l] \in \{1, 2, \ldots, T\}^l$ to the same class if

$$\sum_{\lambda=1}^{l} N_{i_\lambda} = \sum_{\lambda=1}^{l} N_{j_\lambda}.$$

Let $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_s$ be all distinct classes, so that

$$\{1, 2, \ldots, T\}^l = \bigcup_{r=1}^{s} \boldsymbol{C}_r.$$

We have

$$f(x)^l = \sum_{r=1}^{s} \sum_{[i_1, i_2, \ldots, i_l] \in \boldsymbol{C}_r} x^{\sum_{\lambda=1}^{l} N_{i_\lambda}} \prod_{\lambda=1}^{l} A_{i_\lambda}.$$

Since $f(x)^l$ has t terms we have for all but t classes $\boldsymbol{C}_r$, say for all $r > t$

$$\sum_{[i_1, i_2, \ldots, i_l] \in \boldsymbol{C}_r} x^{\sum_{\lambda=1}^{l} N_{i_\lambda}} \prod_{\lambda=1}^{l} A_{i_\lambda} = 0. \tag{20}$$

Let us consider the system of linear equations

$$\sum_{\lambda=1}^{l} x_{i_\lambda} = \sum_{\lambda=1}^{l} x_{j_\lambda} \quad \text{for} \quad [i_1, \ldots, i_l], [j_1, \ldots, j_l] \in \boldsymbol{C}_r \tag{21}$$

and all $r \leqslant s$.

This system with T unknowns has at least two linearly independent solutions namely $[1, 1, \ldots, 1]$ and $[N_1, \ldots, N_T]$. Hence the matrix $\boldsymbol{M}$ of the system is of rank $\varrho \leqslant T-2$. The linear space of solutions has a basis consisting of $T-\varrho$ vectors: $v_1, v_2, \ldots, v_{T-\varrho}$ the components of which are minors of $\boldsymbol{M}$ of order ϱ (see R. Fricke [4], p. 81). Since in each row of the matrix $\boldsymbol{M}$ the sum of the positive elements and the sum of the negative elements is at most l, by the result of [7] the minors in question are in absolute value at most l^ϱ. Hence

$$v_i = [v_{i1}, v_{i2}, \ldots, v_{iT}], \quad \text{where} \quad |v_{ij}| \leqslant l^\varrho \quad (1 \leqslant i \leqslant T-\varrho). \tag{22}$$

Since every solution of (21) is a linear combination of $v_1, \ldots, v_{T-\varrho}$ we have for suitable $u_i^0 \in \boldsymbol{Q}$ $(1 \leqslant i \leqslant T-\varrho)$

$$N_j = \sum_{i=1}^{T-\varrho} u_i^0 v_{ij} \quad (1 \leqslant j \leqslant T)$$

and since N_j are distinct

$$\prod_{\substack{j,k=1\\ j<k}}^{T}\sum_{i=1}^{T-\varrho} u_i^0(v_{ik}-v_{ij})\neq 0.$$

Since the polynomial

$$\prod_{\substack{j,k=1\\ j<k}}^{T}\sum_{i=1}^{T-\varrho} u_i(v_{ik}-v_{ij})\in Q[u_1,\ldots,u_{T-\varrho}]$$

does not vanish identically and is of degree $\binom{T}{2}$ it follows from Lemma 4 with $U=\left\{u\in\boldsymbol{Z}\colon |u|\leqslant \frac{1}{2}\binom{T}{2}+\frac{1}{2}\right\}$ that it does not vanish on the set $U^{T-\varrho}$. Hence there exist integers $u_1,\ldots,u_{T-\varrho}$ such that

$$|u_i^1|\leqslant \tfrac{1}{2}\binom{T}{2}+\tfrac{1}{2} \quad (1\leqslant i\leqslant T-\varrho) \tag{23}$$

and

$$\prod_{\substack{j,k=1\\ j<k}}^{T}\sum_{i=1}^{T-\varrho} u_i^1(v_{ik}-v_{ij})\neq 0. \tag{24}$$

Let us put

$$N_j^1=\sum_{i=1}^{T-\varrho} u_i^1 v_{ij}-\min_{1\leqslant j\leqslant T}\sum_{i=1}^{T-\varrho} u_i^1 v_{ij} \quad (1\leqslant j\leqslant T).$$

By (23) and (24) we have for all $j\leqslant T$

$$0\leqslant N_j^1\leqslant (T-\varrho)\left(\binom{T}{2}+1\right)l^{\varrho}\leqslant (T^2-T+2)\,l^{T-2}. \tag{25}$$

By (24) N_j^1 are all distinct. Since $[N_1^1,\ldots,N_T^1]$ is a solution of (21) we have for all $r\leqslant s$ and suitable integers $\nu(r)$

$$\sum_{\lambda=1}^{l} N_{i_\lambda}^1=\nu(r) \tag{26}$$

for all vectors $[i_1,\ldots,i_l]\in\boldsymbol{C}_r$. Let us put

$$f_1(x)=\sum_{j=1}^{T} A_j x^{N_j^1}.$$

The polynomial f_1 has T terms and in virtue of (25)

$$l\deg f_1\leqslant l^{T-1}(T^2-T+2)<\operatorname{char}\boldsymbol{K}.$$

Moreover by (26) and (20)

$$f_1(x)^l=\sum_{r=1}^{s} x^{\nu(r)}\sum_{[i_1,i_2,\ldots,i_l]\in\boldsymbol{C}_r}\prod_{\lambda=1}^{l} A_{i_\lambda}$$
$$=\sum_{r=1}^{t} x^{\nu(r)}\sum_{[i_1,i_2,\ldots,i_l]\in\boldsymbol{C}_r}\prod_{\lambda=1}^{l} A_{i_\lambda}.$$

Hence $f_1(x)^l$ has at most t terms and by Theorem 1

$$t\geqslant l+1+(\log 2)^{-1}\log\left(1+\frac{\log(T-1)}{l\log 4l-\log l}\right).$$

This shows the first part of the theorem.

In order to prove the second part, let us put $\operatorname{char}\boldsymbol{K}=p$, $l=p^{\alpha}m$, where $m\not\equiv 0 \bmod p$, $m>1$. Take

$$f_n(x)=(1+x)^{(p^{\varphi(m)n}+m-1)/m}$$

and let T_n, t_n be the number of terms of f_n and f_n^l, respectively. We have

$$f_n(x)^l=(1+x)^{(p^{\varphi(m)n}+m-1)p^{\alpha}}=(1+x^{p^{\varphi(m)n+\alpha}})(1+x^{p^{\alpha}})^{m-1}$$

hence

$$t_n\leqslant 2m\leqslant 2l.$$

On the other hand, if

$$\frac{p^{\varphi(m)}+m-1}{m}=\sum_{i=0}^{k} c_i p^i, \qquad \frac{p^{\varphi(m)}-1}{m}=\sum_{i=0}^{k} d_i p^i \quad (0\leqslant c_i, d_i<p,\ c_k\neq 0)$$

then $k<\varphi(m)$; hence

$$\frac{p^{\varphi(m)n}+m-1}{m}=\sum_{i=0}^{k} c_i p^i+\sum_{\nu=1}^{n-1}\sum_{i=0}^{k} d_i p^{\varphi(m)\nu+i}$$

is a reduced representation of $(p^{\varphi(m)n}+m-1)/m$ to the base p and, by Lemma 5

$$T_n=\prod_{i=0}^{k}(c_i+1)\Big(\prod_{i=1}^{k}(d_i+1)\Big)^{n-1}\geqslant 2^n.$$

LEMMA 6. *If $\boldsymbol{K}$ is a field of characteristic p, $\xi\in\hat{\boldsymbol{K}}$*

$$(x-\xi)^{p^m}\Big|\sum_{j=0}^{p-1} x^j f_j(x^p), \quad \textit{where} \quad f_j\in\hat{\boldsymbol{K}}[y] \tag{27}$$

then

$$(y-\xi^p)^m\,|\,f_j(y) \quad \textit{for all } j<p.$$

Proof by induction on m. For $m=1$, we have

$$f_j(x^p) \equiv f_j(\xi^p) \bmod (x-\xi)^p,$$

hence

$$(x-\xi)^p \Big| \sum_{j=0}^{p-1} x^j f_j(\xi^p)$$

and on comparing the degrees we get $f_j(\xi^p)=0$ for all $j<p$; thus

$$y-\xi^p \mid f_j(y).$$

Assuming that the lemma is true with m replaced by $m-1$ we get first by applying the case $m=1$, that

$$f_j(y) = (y-\xi^p) g_j(y), \quad g_j \in \hat{\boldsymbol{K}}[y],$$

hence by (27)

$$(x-\xi)^{p(m-1)} \Big| \sum_{j=0}^{p-1} x^j g_j(x^p)$$

and by the inductive assumption

$$(y-\xi^p)^{m-1} \mid g_j(y) \quad (0 \leqslant j < p),$$

which gives the assertion.

LEMMA 7. *Let $\boldsymbol{K}$ be a field of characteristic p,*

$$f(x) = \sum_{j=0}^{p-1} x^j f_j(x^p) \in \boldsymbol{K}[x], \quad n \equiv r \bmod p, \quad 0 \leqslant r < p.$$

If $\xi \in \hat{\boldsymbol{K}}$ is a zero of f of multiplicity exactly n, then

(28) *for all nonnegative $j \leqslant p$*

$$f_j(x) = (x-\xi^p)^{(n-r)/p} g_j(x), \quad g_j \in \hat{\boldsymbol{K}}[x];$$

(29) *for all nonnegative $s < r$*

$$\sum_{j=s}^{p-1} \binom{j}{s} \xi^{j-s} g_j(\xi^p) = 0;$$

(30)
$$\sum_{j=r}^{p-1} \binom{j}{r} \xi^{j-r} g_j(\xi^p) \neq 0.$$

Proof. Since $(x-\xi)^{n-r} \mid f(x)$, (28) follows from Lemma 6. Now the condition $(x-\xi)^n \| f(x)$ reduces to [2]

$$(x-\xi)^r \| g(x), \quad \text{where} \quad g(x) = \sum_{j=0}^{p-1} x^j g_j(x^p).$$

[2] $a \| b$ means that $a \mid b$ and $(a, b/a) = 1$.

If $r=0$ the condition (29) is void and (30) follows from $g(\xi) \neq 0$. If $r>0$ we write

$$g(x) = (x-\xi)^r h(x), \quad h(\xi) \neq 0$$

and differentiating $s \leqslant r$ times we find that

$$g^{(s)}(\xi) = 0 \quad \text{for} \quad s < r, \quad g^{(r)}(\xi) = r!\, h(\xi) \neq 0,$$

which gives (29) and (30).

Remark. The implication given in Lemma 7 is, in fact, an equivalence.

Proof of Theorem 3. For $\xi = 0$ the theorem is clear. For $\xi \neq 0$ in view of Lemma 1 we may assume char $\boldsymbol{K} = p$. We proceed by induction on n. For $n=1$ the theorem is obviously true. Assume it is true for all multiplicities less than $n \geqslant 2$ and let f have a zero $\xi \in \boldsymbol{K}$ of multiplicity exactly n. Let

(31)
$$f(x) = \sum_{j=0}^{p-1} x^j f_j(x^p), \quad f_j \in \boldsymbol{K}[y]$$

and

(32)
$$n = \sum_{i=1}^{k} c_i p^{n_i}, \quad 0 < c_i < p, \quad 0 \leqslant n_1 < n_2 < \ldots < n_k.$$

If $n_1 > 0$, then by Lemma 6

$$(y-\xi^p)^{n/p} \mid f_j(y) \quad (0 \leqslant j < p)$$

and for at least one j

$$(y-\xi)^{n/p} \| f_j(y).$$

Hence, by the inductive assumption the number of terms of f_j is at least that of $(y-\xi^p)^{n/p}$, i.e. that of $(x-\xi)^n$.

If $n_1 = 0$ we apply Lemma 7 and infer (28), (29), (30) with $r = c_1$. (30) implies that at least one of the elements $g_j(\xi^p)$ $(c_1 < j < p)$ is not zero.

We assert that among the numbers $g_j(\xi^p)$ $(0 \leqslant j < p)$ there are at least c_1+1 different from 0. Indeed, otherwise there would be, at least $p-c_1$ indices j with $g_j(\xi^p) = 0$. Let the remaining indices be $j_1, \ldots, j_{c_1}$. The system of equations (29) gives

$$\sum_{t=1}^{c_1} \binom{j_t}{s} \xi^{j_t} g_{j_t}(\xi^p) = 0 \quad (0 \leqslant s < c_1).$$

However

$$\left| \binom{j_t}{s} \right|_{\substack{0 \leqslant s < c_1 \\ 1 \leqslant t \leqslant c_1}} = \prod_{0 \leqslant q < r < c_1} \frac{j_r - j_q}{r-q} \neq 0,$$

hence $g_{j_t}(\xi^p) = 0$ for all t and thus $g_j(\xi^p) = 0$ for all $j < p$, contrary to (30).

Let now $g_j(\xi^p) \neq 0$ for $j \in S$, where S is a set of cardinality c_1+1. We have for $j \in S$

$$(y-\xi^p)^{(n-c_1)/p} \| f_j(y),$$

hence by the inductive assumption $f_j(y)$ has at least as many terms as $(y-\xi^p)^{(n-c_1)/p}$, i.e. by Lemma 5 and by (32) at least $\prod_{i=2}^{k}(c_i+1)$ terms. It follows that $f(x)$ has at least $\prod_{i=1}^{k}(c_i+1)$ terms, but this is exactly by Lemma 5 the number of terms of $(x-\xi)^n$.

References

[1] E. R. Berlekamp, *Algebraic coding theory*, New York 1968.

[2] P. Erdős, *On the number of terms of the square of a polynomial*, Nieuw Arch. Wiskunde (2) 23 (1949), pp. 63–65.

[3] R. Freud, *On the minimum number of terms in the square of a polynomial* (Hungarian), Mat. Lapok 24 (1973), pp. 95–98.

[4] R. Fricke, *Lehrbuch der Algebra*, Band I, Braunschweig 1924.

[5] G. Hajós, *Solution of Problem 41* (Hungarian). Mat. Lapok 4 (1953), pp. 40–41.

[6] H. L. Montgomery and A. Schinzel, *Some arithmetic properties of polynomials in several variables*, in: *Transcendence Theory: Advances and Applications*, London-New York-San Francisco 1977, pp. 195–203.

[7] A. Schinzel, *An inequality for determinants with real entries*, Colloq. Math. 38 (1978), pp. 319–321.

[8] — *A relation between two conjectures on polynomials*, Acta Arith. 38 (1980), pp. 285–322.

[9] W. Verdenius, *On the number of terms of the square and the cube of polynomials*, Indag. Math. 11 (1949), pp. 546–565.

Received on 22.11.1985
and in revised form on 24.3.1986 (1566)

ACTA ARITHMETICA
XLIX(1987)

Perfect powers in products of integers from a block of consecutive integers

by

T. N. Shorey (Bombay)

To Professor P. Erdős on his 75th birthday

1. Erdős and Selfridge [5] confirmed an old conjecture by proving that the product of two or more consecutive positive integers is never a power. We consider a more general question. For an integer $v > 1$, we define $P(v)$ to be the greatest prime factor of v and we write $P(1) = 1$. Let $m \geqslant 0$ and $k \geqslant 2$ be integers. Let $d_1, \ldots, d_t$ with $t \geqslant 2$ be distinct integers in the interval $[1, k]$. For integers $l \geqslant 2$, $y > 0$ and $b > 0$ with $P(b) \leqslant k$, we consider the equation

$$(m+d_1)\ldots(m+d_t) = by^l. \tag{1}$$

For $l \geqslant 2$, let v_l be a real number satisfying $0 < v_l \leqslant 1$. If $\alpha > 1$ and $k^\alpha < m \leqslant k^l$, then equation (1) implies that $P(m+d_i) \leqslant k$ for $1 \leqslant i \leqslant t$ and hence

$$t < \alpha^{-1}k + \pi(k).$$

See Erdős and Turk [6], Lemma 2.1. For $m > k^l$, we have

THEOREM 1. *Let $\varepsilon > 0$ and $0 \leqslant u < 1$. Suppose that equation* (1) *with*

$$l \geqslant 3, \quad m > k^l, \quad t \geqslant v_l k \tag{2}$$

is satisfied. Then the inequalities

$$v_l \geqslant \frac{1}{2-u} + \varepsilon, \quad v_l \geqslant \frac{1}{2}\left(1 + \frac{2l-3+u}{(2l-4+u)(l-1)}\right) \tag{3}$$

imply that k is bounded by an effectively computable number depending only on ε.

We observe that (3) with an optimal choice of u is somewhat stronger than

$$v_l \geqslant \frac{1}{2}\left(1 + \frac{1}{l-1}\right).$$

We apply Theorem 1 together with Lemma 6 of [9] to derive