COMPOSITE FACTORS OF BINOMIALS AND LINEAR SYSTEMS IN ROOTS OF UNITY

ROBERTO DVORNICICH AND UMBERTO ZANNIER

Abstract. In this paper we completely classify binomials in one variable which have a nontrivial factor which is *composite*, i.e. of the shape $g(h(x))$ for polynomials g, h both of degree > 1. In particular, we prove that, if a binomial has such a composite factor, then $\deg g \leq 2$ (under natural necessary conditions). This is best-possible and improves on a previous bound $\deg g \leq 24$.

This result provides evidence toward a conjecture predicting a similar bound when binomials are replaces by polynomials with any given number of terms.

As an auxiliary result, which could have other applications, we completely classify the solutions in roots of unity of certain systems of linear equations.

AMS Classification: 11B83; 11B99; 11D99.

1. Introduction

In this paper we consider *lacunary polynomials*, also said *fewnomials*, actually in a single variable; by this we mean polynomials in which the number ℓ of terms is supposed to be fixed, whereas the degrees and the coefficients may vary arbitrarily. Another notion is that of *composite polynomials*, i.e. of the shape $g(h(x))$, for polynomials g, h both of degree > 1.

With different motivations, a number of papers have been devoted to the study of polynomials which are simultaneously lacunary and composite. See for instance A. Schinzel's book [4], the papers [3] by C. Fuchs and the second author, and [5] by the second author, and the references therein.

In the paper [5], with suitable necessary assumptions, for any decomposition $f(x) = g(h(x))$ of a polynomial $f(x)$ with ℓ terms, a bound is proved for $\deg g$ depending only on ℓ.

In the paper [3] (after extending the previous result to the case of rational functions) the question was raised whether a similar bound holds more generally for any composite *divisor* $g(h(x))$ of a lacunary polynomial $f(x)$. Again in [3], after remarking that this problem seems to be very difficult in the general case, a positive evidence was provided for the case of binomials, i.e. $\ell = 2$; namely, it was proved that if $g(h(x))$ divides a binomial, then $\deg g \leq 24$ under the necessary assumption that $h(x)$ is not of the shape $ax^m + b$.

However, the problem of the optimality of the bound 24 was left open in [3], where only an example with $\deg g = 2$ was presented.

One purpose of the present paper is to prove that in fact, under the above conditions, the bound $\deg g \leq 2$ holds; this is best possible in view of the cited example. Further, we shall give a complete description of all cases of equality, namely when a binomial has a factor $g(h(x))$ with $\deg g = 2$ (and h not of the mentioned shape); see Theorem 1 below.

In the paper [3], the proof of the bound 24 depended on a result by F. Beukers and C. Smyth [1] on the number of solutions of an algebraic equation in two variables restricted to be roots of unity.

Here, to improve on the bound 24 we shall again relate our issue to equations in roots of unity, but this time we shall find it useful to study certain *systems* of linear equations instead of a single algebraic equation.

We shall prove a best-possible result about the linear systems in question, again characterizing the cases when a certain bound is attained.

Actually, such result holds not merely for solutions in roots of unity, but in the more general case when the unknowns are taken from a multiplicative subgroup Θ of $\mathbb{C}^*$ with the property that there exists a field automorphism of $\mathbb{C}$ restricting to $x \mapsto x^{-1}$ on Θ. For the group of roots of unity this automorphism may be taken the complex conjugation.

See Theorem 3 below for these results.

2. Statements of results

As in the Introduction, we start with the result on binomials. We write a binomial in the form $x^l(x^m + a)$, $a \in \mathbb{C}$, and we consider the equation

$$x^l(x^m + a) = r(x)g(h(x)), \quad l, m \geq 0, \quad \deg g \geq 2, \quad h(x) \neq bx^n + c, \tag{1}$$

where r, g, h are complex polynomials and $b, c \in \mathbb{C}$. Note that in order to avoid trivialities, the restriction that $\deg g \geq 2$ must clearly be imposed, and the same holds for the restriction that $h(x)$ is not of the shape $bx^n + c$, for any complex values of b, c. In fact, note that for any l, m we may factor $x^l(x^m + a)$ as $r(x)g(bx + c)$, where we may take $g(x)$ of any degree $\leq l + m$ and then the substitution $x \to x^n$ yields examples with $h(x) = bx^n + c$. Conversely it is not difficult to see that this procedure leads to all cases of equation (1) with such an $h(x) = bx^n + c$.

We have the following characterization:

Theorem 1. *Assume that the conditions in* (1) *are satisfied. Then* $a \neq 0$ *and* $\deg g = 2$. *Also, after a substitution of type* $x \mapsto \lambda x$, $g \mapsto g \circ \sigma$, $h \mapsto \sigma^{-1} \circ h$, *where* σ *is an affine automorphism (i.e. a polynomial of degree* 1*), all solutions are given by* $g(x) = x(x-1)$, $h(x) = \frac{x^{(q+1)\delta}-1}{x^\delta-1}$, *where* q, δ *are positive integers with* $q\delta > 1$, *and where* $q(q+1)\delta$ *divides* m *and* $l \geq \delta \geq 1$.

In particular, this shows that binomials $x^m + a$ cannot have *non-trivial* composite factors; namely, the exponent l in equation (1) must be positive if there is a composite factor $g(h(x))$ with $h(x)$ not of the shape $bx^n + c$. (Note as above that omitting this restriction leads to plenty of 'trivial' examples.)

Note that if $h(x)$ is as in the last part of the statement, then $h(x) - 1 = x^\delta \frac{x^{q\delta}-1}{x^\delta-1}$, so $g(h(x)) = x^\delta \frac{x^{(q+1)\delta}-1}{x^\delta-1} \cdot \frac{x^{q\delta}-1}{x^\delta-1}$ is a polynomial all of whose roots are 0 or roots of unity (of order dividing $q(q+1)\delta$) and which therefore is indeed a factor of some binomial.

Hence, in particular, this result is best-possible and strengthens the bound in [3] where $\deg g$ was estimated by 24.

Also, this shows that the example produced in [3] (see Remark 7.2 therein) is essentially unique.

We proceed to illustrate our other result, which shall be auxiliary for the former.

We let Θ be a subgroup of $\mathbb{C}^*$ such that the following assumption holds:

Assumption: There exists an automorphism τ of $\mathbb{C}$ such that

$$\theta^\tau = \theta^{-1}, \quad \text{for every } \theta \in \Theta. \tag{2}$$

This assumption is verified in particular for any subgroup of the unit circle (on taking τ the complex conjugation) and hence for the group of roots of unity (to which we have later the said application). Other simple examples arise on taking Θ as the group generated by numbers of norm 1 from linearly disjoint real quadratic fields.

For a subgroup Θ of $\mathbb{C}^*$ we denote as usual by $\mathbb{Z}[\Theta]$ the group algebra over Θ (which as an additive abelian group is free on the set of all $\theta \in \Theta$).

We note that in our context τ acts as an involution on Θ and hence on $\mathbb{Z}[\Theta]$.

To avoid confusion, we shall denote generators of $\mathbb{Z}[\Theta]$ by inserting parentheses, e.g. (-1) denotes the free generator corresponding to -1 in case -1 lies in Θ.

For an element $D = \sum_{\theta\in\Theta} m_\theta(\theta) \in \mathbb{Z}[\Theta]$, where the m_θ are integers (almost all equal to 0) and θ runs through all pairwise distinct elements of Θ (a notation which we shall tacitly adopt throughout), we set, for any integer k,

$$D(k) := \sum_{\theta\in\Theta} m_\theta \theta^k \in \mathbb{C}. \tag{3}$$

We define a norm on $\mathbb{Z}[\Theta]$ by putting, for a $D = \sum m_\theta(\theta) \in \mathbb{Z}[\Theta]$,

$$||D|| := \sum |m_\theta|. \tag{4}$$

Definition 2. We say that $D \sim_\Theta D'$, where $D, D' \in \mathbb{Z}[\Theta]$, if there exists a $\xi \in \Theta$ such that $D = \pm(\xi)D'$. Plainly this is an equivalence relation which preserves the norm.

We now assume to have a nonzero $D \in \mathbb{Z}[\Theta]$ as above and an integer $k_0 \geq 1$ such that

$$D(k) = 0 \quad \text{for } k = 1, \dots, k_0 \quad \text{and} \qquad ||D|| \leq 2k_0 + 1. \tag{5}$$

This is of course invariant by the above defined equivalence.

Some examples arise as follows. Let

$$\Delta_m := \sum_{\zeta^m=1} (\zeta) \in \mathbb{Z}[\mathbb{C}^*] \tag{6}$$

denote the element formed with all complex m-th roots of unity. We clearly have $\Delta_m(k) = 0$ for all k which are not multiples of m. Therefore, setting for integers $q, \delta \geq 1$, $q\delta > 1$,

$$D := \Delta_{(q+1)\delta} - \Delta_{q\delta},$$

we have $D(k) = 0$ for $k = 1, \dots, q\delta - 1 \geq 1$ and $||D|| = (2q-1)\delta$; hence the conditions (5) hold with $k_0 = q\delta - 1$. In particular, on setting $q = 1$ and $\delta = m$, we have $\Delta_m \sim \Delta_{2m} - \Delta_m$.

Of course, the same holds on replacing D with an equivalent element.

Conversely, we have the following

Theorem 3. *Let Θ be a subgroup of $\mathbb{C}^*$ satisfying the above assumption as in (2). Assume that for integers m_θ, $\theta \in \Theta$, $k_0 \geq 1$, setting $D := \sum_{\theta\in\Theta} m_\theta(\theta)$, we have (5) above. Then there are integers $q, \delta > 0$ with $q\delta > 1$ such that $D \sim_\Theta \Delta_{(q+1)\delta} - \Delta_{q\delta}$.*

Remark 1. Moreover, it will follow from the proof that Θ contains all roots of unity of order δ if $q = 1$ and of order $q(q+1)\delta$ if $q > 1$. .

The main point in this result concerns the equivalence within $\mathbb{C}^*$. However we shall also prove that this may be in fact realized within the group Θ.

Remark 2. As will be clear from the proof, one may relax the assumption $||D|| \leq 2k_0+1$ to $||D|| \leq 2k_0+r$, for any fixed r, to obtain similar conclusions. However, these conclusions depend opn r, and become more and more complicated as r increases.

3. Proofs

We start with Theorem 3, which shall be a tool for Theorem 1.

Proof of Theorem 3. Let us denote by D_+, D_- the parts of D formed with positive and negative coefficients m_θ respectively. To these sums we associate the polynomials

$$f_+(t) := \prod_{m_\theta>0} (t-\theta)^{m_\theta}, \qquad f_-(t) = \prod_{m_\theta<0} (t-\theta)^{-m_\theta}. \tag{7}$$

As to their degrees, we have

$$a_+ := ||D_+|| = \sum_{m_\theta>0} m_\theta = \deg f_+, \qquad a_- := ||D_-|| = \sum_{m_\theta>0} -m_\theta = \deg f_-. \tag{8}$$

By symmetry we may assume that $a_+ \geq a_-$, hence

$$\delta := a_+ - a_- \geq 0. \tag{9}$$

Finally, let us consider

$$H(t) := f_+(t) - t^\delta f_-(t). \tag{10}$$

We know that $D(k) = 0$ for $k = 1, \dots, k_0$, hence the first k_0 power-sums (with strictly positive exponents) associated to the roots of f_+, counted with multiplicity, coincide with the same quantities associated to f_-. However, by Newton's identity, power-sums and elementary symmetric functions of order up to k_0 can be expressed from each other as universal polynomials with rational coefficients. Therefore, taking into account that f_+, f_- are monic, the first k_0+1 leading coefficients of f_+ and f_- coincide. We resume this conclusion in the following inequality:

$$\deg H \leq a_+ - k_0 - 1. \tag{11}$$

Note that when $\delta = 0$ this yields $\deg H \leq \frac{||D||}{2} - k_0 - 1 < 0$ and in fact the above shows that $H = 0$. But this in turn implies $f_+ = f_-$, which is a contradiction with the opening definitions.

Therefore we assume in the sequel that $\delta > 0$.

We continue by noting that applying the main assumption (2) to the equations (5) we also obtain

$$D(k) = 0 \quad \text{for } k = -1, \dots, -k_0. \tag{12}$$

Then, setting

$$\tilde{f}_+(t) := \prod_{m_\theta>0} (t-\theta^{-1})^{m_\theta}, \qquad \tilde{f}_-(t) = \prod_{m_\theta<0} (t-\theta^{-1})^{-m_\theta}, \tag{13}$$

$$\tilde{H}(t) := \tilde{f}_+(t) - t^\delta \tilde{f}_-(t), \tag{14}$$

we obtain by the same argument that

$$\deg \tilde{H} \leq a_+ - k_0 - 1. \tag{15}$$

We want now to relate these polynomials and inequalities. We note that

$$t^{a_+} \tilde{f}_+(t^{-1}) \prod_{m_\theta>0} \theta^{m_\theta} = t^{a_+} \prod_{m_\theta>0} (\theta t^{-1} - 1)^{m_\theta} = \prod_{m_\theta>0} (\theta - t)^{m_\theta}.$$

Noting that $\prod_{m_\theta>0}\theta^{m_\theta}=(-1)^{a_+}f_+(0)$ and arguing similarly with f_-, we obtain

$$f_\pm(t)=t^{a_\pm}\tilde{f}_\pm(t^{-1})\cdot f_\pm(0). \tag{16}$$

Substituting in (10) and using (9), we get

$$H(t)=t^{a_+}\left(\tilde{f}_+(t^{-1})\cdot f_+(0)-\tilde{f}_-(t^{-1})\cdot f_-(0)\right). \tag{17}$$

Also, by (14) and (16), we obtain

$$K(t):=t^{a_+}\tilde{H}(t^{-1})=t^{a_+}\tilde{f}_+(t^{-1})-t^{a_-}\tilde{f}_-(t^{-1}). \tag{18}$$

We note that K is a polynomial and that by (15) it is divisible by t^{k_0+1}, i.e.

$$K(t)=t^{k_0+1}K^*(t), \tag{19}$$

for some polynomial K^*.

Multiplying (18) by $f_+(0)$ and subtracting (10), we derive

$$f_+(0)\cdot K(t)-H(t)=t^{a_-}\tilde{f}_-(t^{-1})(t^\delta f_-(0)-f_+(0)),$$

whence, letting

$$\rho:=f_+(0)/f_-(0), \tag{20}$$

we find, on recalling (19) and (16),

$$f_+(0)\cdot t^{k_0+1}K^*(t)-H(t)=\left(t^\delta-\rho\right)f_-(t). \tag{21}$$

Now, note that $||D||=a_++a_-$, whence our assumption $||D||\le 2k_0+1$ may be written as

$$a_++a_-\le 2k_0+1.$$

Also, from our definition $\delta:=a_+-a_-$ and from (11), this yields

$$\deg H\le a_+-k_0-1\le\frac{\delta-1}{2},$$

whereas (15) entails the same inequality for the degree of K^*, i.e.

$$\deg K^*\le a_+-k_0-1\le\frac{\delta-1}{2}.$$

Let now $d:=\lfloor\frac{\delta-1}{2}\rfloor$ be the integral part of $\frac{\delta-1}{2}$. Then we have

$$\deg H\le d,\qquad \deg K^*\le d. \tag{22}$$

Let us now write $k_0+1=q\delta+r$, for integers r,q with $-\delta/2<r\le\delta/2$. This yields

$$t^{k_0+1}=(t^{q\delta}-\rho^q)t^r+\rho^qt^r\equiv t^r\rho^q\pmod{(t^\delta-\rho)}.$$

Then (21) implies the congruence

$$H(t)-f_+(0)\rho^qt^rK^*(t)\equiv 0\pmod{(t^\delta-\rho)}. \tag{23}$$

Now, if $r\ge 0$ the left-hand side is a polynomial of degree at most $r+d\le\frac{\delta}{2}+d<\delta$, and hence it must vanish. Similarly, if $r<0$, on multiplying the left-hand side by t^{-r}, we obtain a polynomial of degree $<\delta$ and divisible again by $t^\delta-\rho$; hence we find that the left-hand side is identically zero in all cases, i.e.

$$H(t)=f_+(0)\rho^q\cdot t^rK^*(t). \tag{24}$$

Substituting into (21) we obtain

$$f_-(t)=f_+(0)t^rK^*(t)\cdot\frac{t^{q\delta}-\rho^q}{t^\delta-\rho}.$$

Also, using (10) in this equation we find

From each of these equations we deduce that $t^r K^*(t)$ is a polynomial. However, $f_+(t), f_-(t)$ are coprime polynomials, whence $t^r K^*(t)$ is constant, which is easily found to be ρ^{-q} on setting $t = 0$ in the last displayed equation. On the other hand, both $f_+(t), f_-(t)$ are monic (which yields $\rho^q = f_+(0)$). Summing up, we find

$$f_+(t) = \frac{t^{(q+1)\delta} - \rho^{q+1}}{t^\delta - \rho}, \qquad f_-(t) = \frac{t^{q\delta} - \rho^q}{t^\delta - \rho}. \tag{25}$$

Conversely, we note that $q\delta = 1$ is impossible (since we would have $q = \delta = 1$ and $D(1)$ would not vanish), so $q\delta > 1$, and if $q\delta > 1$ this definition yields a solution to our conditions.

Let us now distinguish two cases.

First case: $q = 1$. Necessarily $\delta > 1$ now. We find that f_- is constant $= 1$, whence D_- is the empty sum, whereas $f_+(t) = t^\delta + \rho$. Let θ_0 be a root of this equation, so $\theta_0 \in \Theta$. The other roots are of the shape $\theta_0\zeta$, where $\zeta^\delta = 1$, and we have all of these roots. Therefore the roots of unity of order dividing δ lie in Θ, so $D_+ = \sum_{\zeta^\delta=1}(\theta_0\zeta) \sim_\Theta \sum_{\zeta^\delta=1}(\zeta)$, as required.

Second case: $q > 1$. Let now θ_1 be such that $\theta_1^\delta = \rho$. We do not still know if $\theta_1 \in \Theta$. In any case, the roots of f_+ are precisely the elements $\theta_1\zeta$, where $\zeta^{(q+1)\delta} = 1$ but $\zeta^\delta \neq 1$.

Since $q > 1$, among these ζ we find both a primitive $(q+1)\delta$-th root of unity, and its square. Their quotient lies in Θ, which therefore contains a primitive $(q+1)\delta$-th root of unity, and hence contains all roots of unity of order dividing $(q+1)\delta$. Then also $\theta_1 \in \Theta$, and hence the roots of f_- are of the shape $\theta_1\gamma$, where γ is a $q\delta$-th root of unity (such that $\gamma^\delta \neq 1$). Since both $\theta_1\gamma$ and θ_1 lie in Θ, it follows that $\gamma \in \Theta$.

Therefore, we may write

$$D_+ = (\theta_1)(\Delta_{(q+1)\delta} - \Delta_\delta), \qquad D_- = (\theta_1)(\Delta_{q\delta} - \Delta_\delta).$$

and

$$D = D_+ - D_- = (\theta_1)(\Delta_{(q+1)\delta} - \Delta_{q\delta}) \sim_\Theta \Delta_{(q+1)\delta} - \Delta_{q\delta},$$

proving the theorem. □

Proof of Theorem 1. Let us consider equation (1), where we put $d := \deg h \geq 2$. We first make a few reduction steps.

First, if $a = 0$ then $g(h(x))$ must be a power of x, whence the same must hold for g and h, against the present assumptions on h.

Therefore from now on we assume $a \neq 0$.

Note that after a suitable rescaling, in order to prove the theorem we may assume that $a = -1$ and that h, g are both monic.

Note that, for any root $\xi \in \mathbb{C}$ of $g(x)$ of multiplicity $\mu \geq 1$, the polynomial $(h(x) - \xi)^\mu$ divides $x^l(x^m - 1)$. If $\mu > 1$, this forces $h(x) - \xi$ to be a power of x, against the assumptions. Hence $\mu = 1$ for each root of g.

Also, the unique possible multiple root of $h_\xi = h(x) - \xi$ is $x = 0$, so we may write

$$h(x) - \xi = x^{j_\xi} h_\xi(x),$$

where $j_\xi \geq 0$ and where h_ξ has only nonzero roots, which therefore must be simple roots.

The next observation is that j_ξ may vanish for at most one ξ. In fact, suppose that $j_\xi = j_\eta = 0$, where $\xi \neq \eta$ are distinct roots of g. Then, letting $h_\xi - h_\eta = \eta - \xi$ is constant.

Letting S_ξ be the set of roots of h_ξ, and S_θ be the set of roots of roots of h_η, the fact that $h_\xi - h_\eta$ is constant implies that the first d elementary symmetric functions $e_0, \dots, e_{d-1}$ of the roots respectively in S_ξ, S_η coincide. This implies that

$$\sum_{u \in S_\xi} u^r = \sum_{u \in S_\eta} u^r, \qquad r = 0, 1, \dots, d-1.$$

Due to the hypothesis that $h(x) \neq bx^d + c$, there is some integer m with $0 < r < d$ for which the elementary symmetric function e_r coincides for S_ξ and $S\eta$ and is different from 0. Pick one such r.

Observe also that, since S_ξ and S_η are both sets of m-th roots of unity, applying complex conjugation we deduce that the last displayed equation holds even for $r = -1, \dots, -(d-1)$. But e_r/e_d is the elementary $(m-d)$-th symmetric function of the reciprocals of the elements of S_ξ (resp S_η). This yields that also the function e_d coincides for the two sets S_ξ, S_η. But then $h_\xi = h_\eta$, a contradiction.

We have proved that j_ξ may vanish for at most one ξ; on the other hand, we cannot have j_ξ and j_η both positive for $\xi \neq \eta$, for otherwise both $h(x) - \xi$, $h(x) - \eta$ would vanish at $x = 0$.

We conclude, in view of $\deg g \geq 2$, that $\deg g = 2$ and that g has precisely two distinct roots ξ, η, where we may assume that $j_\xi = 0$, $j_\eta > 0$.

With this new information, let us consider the formal sum

$$D := \sum_{u \in S_\xi} (u) - \sum_{u \in S_\eta} (u).$$

We have, in the notation of this paper, $||D|| = 2d - j_\eta$.

Also, since $(h(x) - \xi) - (h(x) - \eta)$ is a (nonzero) constant, the symmetric functions $e_0, \dots, e_{d-1}$ of the respective roots (counted with multiplicity) coincide. But then also the power sums coincide from order 1 to order $d-1$.

This implies that

$$D(k) = 0, \qquad k = 1, \dots, d-1.$$

Putting $k_0 = d-1$, and taking into account that S_ξ, S_η are sets of roots of unity, we may then apply Theorem 3 (on taking the required automorphism may be taken complex conjugation). We deduce that $D \sim_\Theta \Delta_{(q+1)\delta} - \Delta_{q\delta}$, where Θ is the group of m-th roots of unity.

Explicitly in terms of polynomials, formula (25) (where we can take $\rho = 1$ in view of the present normalization) yields

$$h(x) - \xi = \frac{x^{(q+1)\delta} - 1}{x^\delta - 1}, \qquad h(x) - \eta = x^\delta \frac{x^{q\delta} - 1}{x^\delta - 1}. \tag{26}$$

Note that the difference $(h(x) - \xi) - (h(x) - \eta)$ of these polynomials (on the right) is 1, hence $\eta - \xi = 1$.

As remarked above, the group Θ contains the roots of unity of order $q(q+1)\delta$, hence this number divides m.

□

Acknowledgement. We wish to thank the referee for the very helpful comments and suggestions.

References

[1] - F. Beukers, C. Smyth, Cyclotomic points on curves. In: Number Theory for the Millennium, I, A. K. Peters, Natick, MA, 67–85, 2002.

[2] - R. Dvornicich, U. Zannier, On Sums of Roots of Unity, Monatsh. Math., 129 (2000), 97–108.

[3] - C. Fuchs, U. Zannier, Composite rational functions expressible with few terms, JEMS, 14 (2012), 175–208.

[4] - A. Schinzel, Polynomials with special regard to reducibility, Enc. of Math. Appl. 77, Cambridge Univ. Press, 2000.

[5] - U. Zannier, On the number of terms of a composite polynomial, Acta Arith., 127 (2007), 157–167, Addendum, ibid. 140, 93–99.

(R. Dvornicich) Dipartimento di Matematica, Università di Pisa, Largo Pontecorvo 5, 56127 Pisa, Italia

E-mail address: dvornic@dm.unipi.it

(U. Zannier) Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italia

E-mail address: u.zannier@sns.it