ACTA ARITHMETICA
140.1 (2009)

Addendum to the paper: "On the number of terms of a composite polynomial"

(Acta Arith. 127 (2007), 157–167)

by

Umberto Zannier (Pisa)

1. Introduction. Let k be a field of zero characteristic and $f(x) \in k[x]$ be of the shape $f(x) = g(h(x))$, where both $g, h \in k[x]$ have degree > 1. In [Z] we proved in Theorem 1 that, provided h is not of the shape ax^n+b, the number l of terms of f satisfies $\deg f + l - 1 \le 2l(l-1)\deg h$. Essentially as a remark, we also pointed out that the method could yield similar statements for Laurent polynomials f, h (and possibly for rational functions). In [Z], for this case we did not exploit with care the full information coming from the arguments, which resulted in a weaker dependence of the estimate on $\deg h$. (For the applications we had in mind this was amply sufficient.) Now, the above estimate has as a natural consequence an upper bound for $\deg g$ only in terms of l, and this was no more implied by our rough estimate for Laurent polynomials, stated as Theorem 2 of [Z]. This defect is rather annoying, also because recently Watt and Zieve required, for certain applications in [WZ], such a bound for $\deg g$ in the case of Laurent polynomials f, h.

Therefore, to supply a bound sufficient for the applications by Watt and Zieve, we have written this further short note, as an addendum to [Z] (1). In place of Theorem 2 therein we prove, by similar but more accurate arguments, the following:

Theorem 2*. *Let $f \in k[x, x^{-1}]$ have $l > 0$ nonconstant terms. Assume that $f = g(h(x))$, where $g \in k[x]$, $h \in k[x, x^{-1}]$ and where h is not of the shape $ax^m + a'x^{-m} + b$. Then*

$$\deg f \le 2(2l-1)(l-1)\deg h, \qquad \deg g \le 2(2l-1)(l-1).$$

2000 *Mathematics Subject Classification*: 10M05, 12E05, 12Y05.
Key words and phrases: Laurent polynomial.

(1) Further applications may come by generalizing the results of [Z2] to Laurent polynomials.

DOI: 10.4064/aa140-1-6

Here by the degree of a Laurent polynomial we mean its degree as a rational function, i.e. the maximum of the degrees of the numerator and denominator. Note that the right estimate immediately follows from the one on the left. We stress that the main point for writing this addendum is indeed that the left estimate depends linearly on $\deg h$, so it yields an upper bound for $\deg g$ dependent only on l. (Instead, the version in [Z] bounded $\deg f$ quadratically in $\deg h$, which is not only unnatural but also insufficient for the said applications in [WZ].)

We also point out that, as shown by the proof, in the non-polynomial case the term $\deg h$ can be replaced with $\deg h - 2$; the dependence on l too may be possibly refined, but these are minor points for the applications. Finally, as observed in [Z], the special shape for h has to be in fact forbidden, as shown by the identity $x^n + x^{-n} = T_n(x + x^{-1})$ where T_n is the Chebyshev polynomial.

2. Proofs. In the following we suppose, as we may, that k is algebraically closed and that h is not a polynomial, for otherwise f too is a polynomial and Theorem 1 of [Z] yields the present estimates. Note that this assumption on h entails that also f is not a polynomial. We further assume (as we may on changing x into $1/x$ if necessary) that h has a pole at $x = \infty$ of order at least the order of its pole at 0; automatically, the same then holds for $f = g \circ h$.

We start with two lemmas, similar in nature to some lemmas in [Z].

Lemma 1. *Set $\lambda := h(x)$. Any conjugate y of x over $k(\lambda)$, $y \neq x$, has the following properties*:

(i) *Let $d := [k(x,y) : k(x)]$. Then $d \leq \deg h - 1$.*
(ii) *The genus of the function field $K = k(x,y)$ (over k) is $\leq (d-1)^2$.*
(iii) *Any zero or pole of x (with respect to K) is a zero or pole of y, and conversely.*

Proof. Plainly, $k(x)$ is an extension of $k(\lambda)$ of degree $\deg h$. Since y is conjugate to x over $k(\lambda)$ we have $h(y) = h(x)$; assertion (i) follows, since x, y are two (distinct) roots of the equation $h(X) = \lambda$.

Let $H(X,Y) \in k[X,Y]$ be an irreducible polynomial such that $H(x,y) = 0$. Since $h(x) = h(y)$ and since the degree is multiplicative in towers, we have $\deg(x) = \deg(y)$, considering x, y as rational functions in the function field $k(x,y)/k$. Hence $\deg_X H = \deg_Y H$, and certainly $\deg_Y H = \deg(x) = [k(x,y) : k(x)] = d$. Hence, by a theorem of Castelnuovo we obtain the estimate $g \leq (d-1)^2$. (Viewing our curve embedded in $\mathbb{P}_1^2$ rather than $\mathbb{P}_2$, one may recover this bound from the well-known formula $2g-2 = X.(X+K)$ for the genus of a curve X on a surface S, where K is the canonical class: see e.g. [S, IV(20)]. See also [St, Thm. III.10.3].)

Finally, assertion (iii) follows at once from the fact that h is supposed to be a polynomial neither in x nor in $1/x$, so any zero or pole of x is a pole of $\lambda = h(x) = h(y)$, and thus is a zero or pole of y. ∎

LEMMA 2. *Let K/k be a function field in one variable, of genus g, and let $z_1, \ldots, z_s \in K$ be not all constant and such that $1 + z_1 + \cdots + z_s = 0$. Suppose also that no proper subsum of the left side vanishes. Then*

$$\max(\deg(z_i)) \leq \binom{s}{2}(\#S + 2g - 2),$$

where S is any set of points of K containing all zeros and poles of all the z_i.

Here and below by $\deg(z)$ we mean the degree with respect to K, i.e. $[K : k(z)]$; equivalently, this is the number of poles (or zeros) of z counted with multiplicity. This lemma is an immediate consequence of Corollary I of [BM] (as improved after Thm. B therein); we have just used the fact that the "K-height" of the projective point $(1 : z_1 : \ldots : z_s)$ is bounded below by the maximum degree. (Actually, [BM] gives a bound with $2g - 2$ replaced by $\max(0, 2g - 2)$, but the same proof yields in fact the above estimate. This may also be recovered immediately from Theorem 1 in [Z3], without any modification, and anyway for the present purposes this would make no difference.)

Proof of Theorem 2.* The Laurent polynomial $f(x)$ will be written as follows:

$$f(x) = c_0 + c_1 x^{m_1} + \cdots + c_l x^{m_l}, \tag{1}$$

where no m_i is zero and $c_i \in k$, $c_1 \cdots c_l \neq 0$, $m_1 < \cdots < m_l$. Recall that we are assuming that f is not in $k[x]$ and that its pole-order at $x = \infty$ is at least the pole-order at $x = 0$; this entails that $m_l \geq -m_1 > 0$. Hence

$$m_l < \deg f = m_l - m_1 \leq 2m_l.$$

Suppose first that $h \in k(x^n)$ for some integer $n > 1$, so $h = \tilde{h}(x^n)$, and $f = \tilde{f}(x^n)$, where $\tilde{f}(x) = g(\tilde{h}(x))$. Since $\tilde{f}$ has the same number of (nonconstant) terms as f, we may argue with $\tilde{h}(x)$ in place of $h(x)$. Note in fact that $\tilde{h}(x)$ cannot be of the forbidden shape for otherwise $h(x)$ would also be. If we assume the inequality to be proved with $\tilde{h}$ in place of h, $\tilde{f}$ in place of f and the same g, we find the sought estimate and more.

Therefore, we may suppose that, for any $n > 1$, $h \notin k(x^n)$.

Secondly, suppose that h is decomposable in the form $h(x) = p(q(x))$ for $p \in k[x]$ a polynomial and $q \in k[x, x^{-1}]$ a Laurent polynomial. Note that if $q(x) = ax^m + b + a'x^{-m}$ then $m = \pm 1$ by the previous assumption. If q is not of this form, we may now write $f(x) = r(q(x))$, where $r(x) = g(p(x))$.

As before, if we assume the sought inequality with $q(x)$ in place of $h(x)$ and $g(p(x))$ in place of $g(x)$ we again obtain the inequality we want to prove. Hence it will suffice to prove the theorem on replacing g, h with $g \circ p, q$ respectively.

Therefore, by suitably iterating this argument, we may assume from now on that the only possible decomposition $h(x) = p(q(x))$ has either $q = ax + b + a'/x$ (with $a, b, a' \in k$) or $\deg p = 1$.

We also suppose $\deg h > 2$, for otherwise, on the present assumptions, $h(x)$ would necessarily be of the forbidden shape $ax + b + a'/x$.

In what follows we adopt the notation of Lemma 1, letting in particular $K = k(x, y)$. Since $f \in k(h(x))$ we have $f(x) - f(y) = 0$, where y is as in Lemma 1. In view of (1) this reads

$$c_1 x^{m_1} - c_1 y^{m_1} + \cdots + c_l x^{m_l} - c_l y^{m_l} = 0. \tag{2}$$

We shall exploit (2) by means of Lemma 2. Before applying it, we deal with possible vanishing subsums of the left side of (2). We partition the terms on the left of (2) into minimal subsets with vanishing sum. (*A priori* this partition may be done in several ways; we can choose freely one of them.) Among such subsets we pick the one containing the term $c_l x^{m_l}$. We denote the corresponding terms by $w_0, \ldots, w_s$ agreeing that $w_s = c_l x^{m_l}$. We shall then obtain a relation $w_0 + \cdots + w_s = 0$, without proper vanishing subsums, where $w_s = c_l x^{m_l}$ and where $w_0, \ldots, w_s$ are distinct terms taken from the left side of (2). Also, we may clearly write such a vanishing relation in the form

$$p(x) = q(y),$$

where p and q are nonzero Laurent polynomials obtained as certain nonempty subsums of terms $\pm w_j$.

This equation says that $p(x)$ lies in the intersection $k(x) \cap k(y)$, which is a field intermediate between $k(\lambda)$ and $k(x)$ (we are using throughout the notation of Lemma 1 above). By the Lüroth Theorem (see e.g. [Sc]) the field $k(x) \cap k(y)$ is of the shape $k(u(x))$, where $u \in k(x)$ is such that $\lambda = h(x) = t(u(x))$ for some $t \in k(X)$. Note that we may change t, u to $t \circ \phi$, $\phi^{-1} \circ u$ for any homography $\phi \in \mathrm{PGL}_2(k)$.

Now, h is a Laurent polynomial, not a polynomial, and so $\infty \in \mathbf{P}_1(k)$ has precisely the preimages $0, \infty$ under the map $P \mapsto h(P)$. Hence ∞ has either two or one preimage under the map $t(x)$.

If ∞ has two preimages under $t(x)$, we may assume, for a suitable choice of the said homography ϕ, that they are $0, \infty$. Then $0, \infty$ must each have a single preimage (in $\{0, \infty\}$) under the map $u(x)$. This implies that $u(x) = cx^\rho$, $c \in k^*$, $\rho \in \mathbb{Z}$, and then we must have $\rho = \pm 1$ by the present normalization of h and in particular $\deg u = 1$.

If ∞ has just one preimage under $t(x)$, we may assume by suitable choice of ϕ that it is ∞. Then ∞ has just $0, \infty$ as preimages under $u(x)$. So $t(x)$ is a polynomial and $u(x)$ a Laurent polynomial. By our normalization, either $u(x) = ax + b + a'/x$ with $b \in k$, $a, a' \in k^*$, or $\deg t = 1$.

Let us treat these three cases separately.

In the first case, we have $\deg u = 1$, so $k(x) \cap k(y) = k(x)$, i.e., $k(x) \subset k(y)$ and since they have the same degree over $k(\lambda)$ we have $k(x) = k(y)$ and $y = L(x)$ for a linear fractional $L \in \mathrm{PGL}_2(k)$; note that L must be of finite order, because $h(x) = h(L(x))$. Since $h(x) = h(y)$ (or by Lemma 1(iii)) we deduce that L either fixes both $0, \infty$ or exchanges them. Hence either $y = \alpha x$ with a root of unity α of order n, whence $h \in k(x^n)$ contrary to the assumptions (if $n = 1$ we have $y = x$), or $y = \beta/x$ for a $\beta \in k^*$. Note that the latter holds for at most one conjugate: if another conjugate y' of x equals β'/x, then $y' = \gamma y$ for a $\gamma \in k^*$ which is necessarily a root of unity, which case we have just excluded. Since we are working on the assumption $\deg h > 2$, we can start with another conjugate y if necessary (note that $h(X) = \lambda$ has no multiple roots!), and so we can suppose that this case does not occur at all.

Take now the second case, i.e. $u(x) = ax + b + a'/x$. We have $p(x) = v(u(x))$, where $v \in k(X)$. But p is a Laurent polynomial, so necessarily $v \in k[X]$ must be in fact a polynomial: if not, v has some finite pole, whence p has a pole which is neither 0 nor ∞. In turn, we conclude that $p(x)$ has equal pole orders at 0 and ∞; since the pole order at ∞ is m_l, which is largest among the $|m_i|$, we conclude that $m_1 = -m_l$ and that $p(x)$ contains the term $c_1 x^{m_1}$.

In the third and last case, we have $\deg t = 1$ so $k(x) \cap k(y) = k(\lambda)$, so $p(x) \in k(h(x))$, which means that $p(x) = z(h(x))$ for some rational function $z \in k(x)$. Again, as in the second case, we reach the conclusion that $p(x)$ contains the term $c_1 x^{m_1}$: in fact, the ratio between the orders of the poles of $p(x)$ at $\infty, 0$ is the same as for $h(x)$, which in turn is the same as for $f(x)$. (Apart from a factor 2 in the final estimate, it would suffice in these last two cases to deduce that $p(x)$ contains a term $c_\ell x^{m_\ell}$ with $m_\ell \leq 0$.)

We may now renumber the indices to assume that the said term $c_\ell x^{m_\ell}$ is w_0.

Dividing the relation $w_0 + \cdots + w_s = 0$ by $w_0 = c_\ell x^{m_\ell}$ and setting $z_j := w_j/w_0$ we find

$$1 + z_1 + \cdots + z_s = 0.$$

Note that $z_s = (c_l/c_\ell)x^{m_l - m_\ell}$ is nonconstant of degree in x equal to $m_l + |m_\ell| \geq m_l - m_1 = \deg f$. We are then in a position to apply Lemma 2. We proceed to estimate the relevant quantities.

Note that $\deg(z_s) = \deg(x)(m_l - m_\ell) = d(m_l - m_\ell)$ (where the degree is meant, as above, relative to $k(x, y)$). Further, x, y have altogether at most d distinct zeros each. Each zero or pole of x is a zero or pole of y and conversely. We can then bound $\#S$ by $2d$. Finally, by Lemma 1 the genus g of $K = k(x, y)$ satisfies $g \le (d - 1)^2$.

Combining these estimates and using Lemma 2 we find (using now $\deg f$ in the usual way, as the degree of a rational function in $k(x)$, not in $k(x, y)$)

$$d \deg f \le d(m_l - m_\ell) \le \binom{s}{2}(2d + 2(d-1)^2 - 2) = \binom{s}{2}(2d^2 - 2d).$$

Now, we have $s + 1 \le 2l$, whence $\deg f \le 2(2l - 1)(l - 1)(d - 1)$. Finally, recall that $d \le \deg h - 1$ by Lemma 3. All of this gives

$$\deg f \le 2(2l - 1)(l - 1)(\deg h - 2),$$

which proves Theorem 2* (and slightly more). ∎

Remark. In principle the method also applies to decompositions of rational functions $f, g, h \in k(x)$, where by the *number of terms* of a rational function we mean the maximum of the number of terms of the numerator and denominator in a reduced fraction. However, working out such an extension, with suitable new assumptions and conclusions, seems not to be entirely straightforward, and so we leave this out of the present note, which is conceived just as an addendum.

Acknowledgements. I wish to thank Mike Zieve for pointing out to me some issues motivating this paper, as well as for several very helpful comments.

References

[BM] D. Brownawell and D. Masser, *Vanishing sums in function fields*, Math. Proc. Cambridge Philos. Soc. 100 (1986), 427–434.

[Sc] A. Schinzel, *Polynomials with Special Regard to Reducibility*, Encyclopedia Math. Appl. 77, Cambridge Univ. Press, Cambridge, 2000.

[S] J.-P. Serre, *Algebraic Groups and Class Fields*, Grad. Texts in Math. 117, Springer, 1988.

[St] H. Stichtenoth, *Algebraic Function Fields and Codes*, Springer, 1993.

[WZ] S. Watt and M. Zieve, *Functional composition of symbolic Laurent polynomials*, in preparation.

[Z] U. Zannier, *On the number of terms of a composite polynomial*, Acta Arith. 127 (2007), 157–167.

[Z2] —, *On composite lacunary polynomials and the proof of a conjecture of Schinzel*, Invent. Math. 174 (2008), 127–138.

[Z3] U. Zannier, *Some remarks on the S-unit equation in function fields*, Acta Arith. 64 (1993), 87–98.

Scuola Normale Superiore
Piazza dei Cavalieri, 7
56126 Pisa, Italy
E-mail: u.zannier@sns.it

Received on 2.2.2009
and in revised form on 14.4.2009 (5922)