Journal of Algebra 321 (2009) 292–302

Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

Automorphisms of polynomial algebras and Dirichlet series

Vesselin Drensky [a,1], Jie-Tai Yu [b,*,2]

[a] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria
[b] Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong, China

ARTICLE INFO

Article history:
Received 18 June 2008
Available online 26 September 2008
Communicated by Changchang Xi

Keywords:
Automorphisms
Polynomial algebras
Free associative algebras
Finite fields
Coordinates
Nielsen–Schreier varieties
Dirichlet series
Generating functions

ABSTRACT

Let $\mathbb{F}_q[x,y]$ be the polynomial algebra in two variables over the finite field $\mathbb{F}_q$ with q elements. We give an exact formula and the asymptotics for the number p_n of automorphisms (f,g) of $\mathbb{F}_q[x,y]$ such that $\max\{\deg(f),\deg(g)\}=n$. We describe also the Dirichlet series generating function

$$p(s)=\sum_{n\geqslant 1}\frac{p_n}{n^s}.$$

The same results hold for the automorphisms of the free associative algebra $\mathbb{F}_q\langle x,y\rangle$. We have also obtained analogues for free algebras with two generators in Nielsen–Schreier varieties of algebras.

Introduction

Our paper is devoted to the following problem. *Let $\mathbb{F}_q[x,y]$ be the polynomial algebra in two variables over the finite field $\mathbb{F}_q$ with q elements. We would like to determine the number of $\mathbb{F}_q$-automorphisms $\varphi=(f,g)$ of $\mathbb{F}_q[x,y]$ satisfying* $\deg(\varphi):=\max\{\deg(f),\deg(g)\}=n$. Here $\varphi=(f,g)$ means that $f=\varphi(x)$, $g=\varphi(y)$.

Our consideration is motivated by Arnaud Bodin [B], who raised the question to determine the number of $\mathbb{F}_q$-automorphisms φ with $\deg(\varphi)\leqslant n$.

In the sequel all automorphisms are $\mathbb{F}_q$-automorphisms. The theorem of Jung and van der Kulk [J,K] states that the automorphisms of the polynomial algebra $K[x,y]$ over any field K are tame. In other words the group $\mathrm{Aut}(K[x,y])$ is generated by the subgroup A of affine automorphisms

* Corresponding author.
E-mail addresses: drensky@math.bas.bg (V. Drensky), yujt@hkucc.hku.hk (J.-T. Yu).
[1] The research of Vesselin Drensky was partially supported by Grant MI-1503/2005 of the Bulgarian National Science Fund.
[2] The research of Jie-Tai Yu was partially supported by an RGC-CERG grant.

0021-8693/$ – see front matter
doi:10.1016/j.jalgebra.2008.08.026

$$\alpha = (a_1x + b_1y + c_1, a_2x + b_2y + c_2), \quad a_i, b_i, c_i \in K,\ a_1b_2 \neq a_2b_1,$$

and the subgroup B of triangular automorphisms

$$\beta = \big(ax + h(y), by + b_1\big), \quad 0 \neq a, b \in K,\ b_1 \in K,\ h(y) \in K[y].$$

The proof of van der Kulk [K] gives that $\mathrm{Aut}(K[x, y])$ has the following nice structure, see e.g. [C]:

$$\mathrm{Aut}\big(K[x, y]\big) = A *_C B, \quad C = A \cap B,$$

where $A *_C B$ is the free product of A and B with amalgamated subgroup $C = A \cap B$. Using the canonical form of the elements of $\mathrm{Aut}(K[x, y])$ we have calculated explicitly the number p_n of automorphisms of degree n:

$$p_1 = q^3(q-1)^2(q+1),$$
$$p_n = \big(q(q-1)(q+1)\big)^2 \sum \left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k}, \quad n > 1,$$

where the summation is on all ordered factorizations $n = n_1 \cdots n_k$ of n, with $n_1, \ldots, n_k > 1$.

It is natural to express the sequence p_n, $n = 1, 2, \ldots,$ in terms of its generating function. When the elements p_n of the sequence involve sums on the divisors of the index n it is convenient to work with the Dirichlet series generating function, i.e., with the formal series

$$p(s) = \sum_{n \geqslant 1} \frac{p_n}{n^s}.$$

For the Riemann zeta function $\zeta(s)$ the coefficients of $\zeta^k(s)$ count the number of ordered factorizations $n = n_1 \cdots n_k$ in k factors. (We have to take $(\zeta(s) - 1)^k$ if we want to count only factorizations with $n_i \geqslant 2$.) Similarly, the coefficients of the kth power $\rho^k(s)$ of the formal Dirichlet series

$$\rho(s) = \sum_{n \geqslant 2} \frac{q^n}{n^s}$$

are equal to $q^{n_1+\cdots+n_k}$ in the expression of p_n. Hence $\rho(s)$ may be considered as a q-analogue of $\zeta(s)$, although it does not satisfy many of the nice properties of the Riemann zeta function (because the sequence q^n is not multiplicative) and for $q > 1$ is not convergent for any nonzero s (because its coefficients grow faster than n^s). We have found that

$$\begin{aligned} p(s) &= \big(q(q-1)(q+1)\big)^2 \left(\sum_{k \geqslant 0} \frac{q-1}{q} \rho^k(s) - \frac{1}{q+1}\right) \\ &= \big(q(q-1)(q+1)\big)^2 \left(\frac{1}{1 - \frac{q-1}{q}\rho(s)} - \frac{1}{q+1}\right) \end{aligned}$$

and have given an estimate for the growth of p_n. For $n \geqslant 2$

$$(q-1)^3(q+1)^2q^{n+1} \leqslant p_n \leqslant (q-1)^3(q+1)^2q^{n+1} + (\log_2 n)^{\log_2 n} q^{n/2+8}.$$

Hence for a fixed q and any $\varepsilon > 0$,

$$p_n = (q-1)^3(q+1)^2q^{n+1} + \mathcal{O}\big(q^{n(1/2+\varepsilon)}\big).$$

The main contribution $(q-1)^3(q+1)^2q^{n+1}$ to p_n comes from the number of automorphisms of the form

$$\big(a_1x+b_1y+c_1+a_1h(y), a_2x+b_2y+c_2+a_2h(y)\big),$$
$$\big(b_1x+(a_1+ab_1)y+c_1+a_1h(x+ay), b_2x+(a_2+ab_2)y+c_2+a_2h(x+ay)\big),$$

where $a, a_i, b_i, c_i \in \mathbb{F}_q$, $a_1b_2 \neq a_2b_1$,

$$h(y) = h_ny^n + h_{n-1}y^{n-1} + \cdots + h_2y^2 \in y^2\mathbb{F}_q[y], \quad h_n \neq 0.$$

Based on the above result, we have also calculated explicitly the number l_n of coordinates of degree n and obtained its Dirichlet series generating function, i.e., with the formal series

$$l(s) = \sum_{n\geqslant 1} \frac{l_n}{n^s}.$$

By the theorem of Czerniakiewicz and Makar-Limanov [Cz,ML] for the tameness of the automorphisms of $K\langle x, y\rangle$ over any field K and the isomorphism $\mathrm{Aut}(K[x, y]) \cong \mathrm{Aut}(K\langle x, y\rangle)$ which preserves the degree of the automorphisms we derive immediately that the same results hold for the number of automorphisms of degree n of the free associative algebra $\mathbb{F}_q\langle x, y\rangle$.

It is easy to obtain an analogue of the formula for the number of automorphisms p_n of degree n for free algebras with two generators over $\mathbb{F}_q$ if the algebra satisfies the Nielsen–Schreier property. Examples of such algebras are the free Lie algebra and the free anti-commutative algebra (where $p_1 = q(q-1)^2(q+1)$ and $p_n = 0$ for $n > 1$), free nonassociative algebras and free commutative algebras.

In a forthcoming paper we are going to give an algebraic geometrical analogue of the main results of the present paper for infinite fields.

1. Canonical forms of automorphisms

The group G is the free product of its subgroups A and B with amalgamated subgroup $C = A \cap B$ (notation $G = A *_C B$), if G is generated by A and B and if for any $a_1, \ldots, a_{k+1} \in A$, $b_1, \ldots, b_k \in B$, $k \geqslant 1$, such that $a_2, \ldots, a_k, b_1, \ldots, b_k$ do not belong to C, the product $g = a_1b_1 \cdots a_kb_ka_{k+1}$ does not belong to C.

For the following description of $G = A *_C B$ see e.g. [MKS, p. 201, Theorem 4.4] and its corollaries.

Lemma 1.1. *Let $G = A *_C B$, $C = A \cap B$, and let*

$$A_0 = \{1, a_i \in A \mid i \in I\}, \qquad B_0 = \{1, b_j \in B \mid j \in J\}$$

be, respectively, left coset representative systems for A and B modulo C. Then each $g \in G$ can be presented in a unique way in the form

$$g = g_1 \cdots g_kc,$$

where $1 \neq g_i \in A_0 \cup B_0$, $i = 1, \ldots, k$, g_i, g_{i+1} are neither both in A_0, nor both in B_0, $c \in C$.

We write the automorphisms of $K[x, y]$ over any field K as functions. If $\varphi = (f_1(x, y), g_1(x, y))$, $\psi = (f_2(x, y), g_2(x, y))$, then $\varphi \circ \psi(u) = \varphi(\psi(u))$, $u \in K[x, y]$, and hence

$$\varphi \circ \psi = \big(f_2\big(f_1(x, y), g_1(x, y)\big), g_2\big(f_1(x, y), g_1(x, y)\big)\big).$$

The following presentation of the automorphisms of $K[x, y]$ is well known, see e.g. Wright [Wr]. We include the proof for self-containness of the exposition.

Proposition 1.2. *Define the sets of automorphisms of* $K[x,y]$

$$A_0 = \{\iota = (x,y),\ \alpha = (y, x+ay) \mid a \in K\},$$
$$B_0 = \{\beta = (x+h(y), y) \mid h(y) \in y^2K[y]\}.$$

Every automorphism φ *of* $K[x,y]$ *can be presented in a unique way as a composition*

$$\varphi = (f,g) = \alpha_1 \circ \beta_1 \circ \alpha_2 \circ \beta_2 \circ \cdots \circ \alpha_k \circ \beta_k \circ \lambda,$$

where $\alpha_i \in A_0$, $\alpha_2, \dots, \alpha_k \neq \iota$, $\beta_i \in B_0$, $\beta_1, \dots, \beta_k \neq \iota$, $\lambda \in A$. *If* $\beta_i = (x + h_i(y), y)$ *and* $\deg(h_i(y)) = n_i$, *then the degree of* φ

$$n = \deg(\varphi) = \max\{\deg(f), \deg(g)\} = n_1 \cdots n_k$$

is equal to the product of the degrees of β_i.

Proof. First we shall show that A_0 and B_0 are, respectively, left coset representative systems for the group A of affine automorphisms and the group B of triangular automorphisms modulo the intersection $C = A \cap B$.

Since ι is the identity automorphism, $\alpha = (y, x+ay) \notin C$ and

$$\alpha_1^{-1} \circ \alpha_2 = (y, x + a_1y)^{-1} \circ (y, x+a_2y) = (x, (-a_1+a_2)x + y)$$

does not belong to C when $a_1 \neq a_2$, to show the statement for A_0 it is sufficient to verify that for any $\lambda \in A$ there exist $\alpha \in A_0$ and $\gamma \in C$ such that $\lambda = \alpha \circ \gamma$. We choose $\alpha = \iota$ and $\gamma = \lambda$ if $\lambda \in C$ and have the presentation

$$\begin{aligned}\lambda &= (a_1x + b_1y + c_1, a_2x + b_2y + c_2)\\ &= \left(y, x + \frac{b_2}{a_2}y\right) \circ \left(\left(b_1 - \frac{a_1b_2}{a_2}\right)x + a_1y + c_1, a_2y + c_2\right)\end{aligned}$$

when $\lambda \notin C$ (and hence $a_2 \neq 0$). Similarly, for B_0 it is sufficient to use that $\beta = (x + h(y), y) \notin C$ when $0 \neq h(y) \in y^2K[y]$,

$$\beta_1^{-1} \circ \beta_2 = (x + h_1(y), y)^{-1} \circ (x + h_2(y), y) = (x - h_1(y) + h_2(y), y) \notin C$$

for $h_1(y) \neq h_2(y)$ (because h_1 and h_2 have no monomials of degree < 2) and to see that

$$(ax + h(y), by + b_1) = (x + k(y), y) \circ (ax + h_1y + h_0, by + b_1) \in B_0 \circ C,$$

where $a, b, b_1 \in K$, $a, b \neq 0$, $h(y) = h_ny^n + \cdots + h_1y + h_0$ and

$$k(y) = \frac{1}{a}(h_ny^n + \cdots + h_3y^3 + h_2y^2).$$

Lemma 1.1 gives that every automorphism φ of $K[x,y]$ has a unique presentation

$$\varphi = (f,g) = \alpha_1 \circ \beta_1 \circ \alpha_2 \circ \beta_2 \circ \cdots \circ \alpha_k \circ \beta_k \circ \alpha_{k+1} \circ \gamma,$$

where $\alpha_i \in A_0$, $\beta_i \in B_0$, $\beta_1, \alpha_2, \dots, \alpha_k, \beta_k \neq \iota$, $\gamma \in C$. Since $\lambda = \alpha_{k+1} \circ \gamma \in A$, we obtain the presentation.

Finally, if

$$\lambda = (a_1x + b_1y + c_1, a_2x + b_2y + c_2), \qquad \beta = \bigl(x + h(y), y\bigr), \qquad \alpha = (y, x + ay),$$

$h = h_ny^n + \cdots + h_2y^2$, $h_n \neq 0$, then the degree of $\beta \circ \lambda$ and $\alpha \circ \beta \circ \lambda$ is n and homogeneous components of maximal degree of these automorphisms are, respectively,

$$\overline{\beta \circ \lambda} = \bigl(a_1h_ny^n, a_2h_ny^n\bigr), \qquad \overline{\alpha \circ \beta \circ \lambda} = \bigl(a_1h_n(x + ay)^n, a_2h_n(x + ay)^n\bigr).$$

They are different from 0 because $(a_1, a_2) \neq (0, 0)$. If the homogeneous component of maximal degree of $\psi = \alpha_i \circ \beta_i \circ \cdots \circ \alpha_k \circ \beta_k \circ \lambda$ is

$$\overline{\psi} = \bigl(d_1(x + a_0y)^m, d_2(x + a_0y)^m\bigr),$$

then for $\beta \circ \psi$ and $\alpha \circ \beta \circ \psi$ we obtain

$$\overline{\beta \circ \psi} = \bigl(d_1h_n^m y^{mn}, d_2h_n^m y^{mn}\bigr),$$
$$\overline{\alpha \circ \beta \circ \psi} = \bigl(d_1h_n^m(x + ay)^{mn}, d_2h_n^m(x + ay)^{mn}\bigr)$$

and we apply induction. □

2. The main results

Now we give a formula for the number of automorphisms of degree n for $\mathbb{F}_q[x, y]$ and the corresponding Dirichlet series generating function.

Theorem 2.1.

(i) *The number p_n of automorphisms $\varphi = (f, g)$ of degree n of $\mathbb{F}_q[x, y]$, i.e., such that*

$$n = \max\bigl\{\deg(f), \deg(g)\bigr\}$$

is given by the formulas

$$p_1 = q^3(q-1)^2(q+1),$$
$$p_n = \bigl(q(q-1)(q+1)\bigr)^2 \sum \left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k}, \quad n > 1,$$

where the summation is on all ordered factorizations $n = n_1 \cdots n_k$ of n, with $n_1, \ldots, n_k > 1$.

(ii) *The Dirichlet series generating function $p(s)$ of the sequence p_n, $n = 1, 2, \ldots$, is*

$$p(s) = \sum_{n \geqslant 1} \frac{p_n}{n^s} = \bigl(q(q-1)(q+1)\bigr)^2 \left(\frac{1}{1 - \frac{q-1}{q}\rho(s)} - \frac{1}{q+1}\right),$$

where

$$\rho(s) = \sum_{n \geqslant 2} \frac{q^n}{n^s}.$$

Proof. (i) Applying Proposition 1.2, p_n is a sum on all ordered factorizations $n = n_1 \cdots n_k$, $n_i > 1$, of the number of automorphisms of the form

$$\varphi = (f, g) = \alpha_1^\delta \circ \beta_1 \circ \alpha_2 \circ \beta_2 \circ \cdots \circ \alpha_k \circ \beta_k \circ \lambda,$$

where: $\alpha_i = (y, x + a_i y)$, $a_i \in \mathbb{F}_q$, δ is 1 or 0, depending on whether or not α_1 participates in the decomposition of φ; $\beta_i = (x + h_i(y)y)$, $h_i(y) \in y^2\mathbb{F}_q[y]$, $\deg(h_i) = n_i$; λ is an affine automorphism. We have $q + 1 = |\mathbb{F}_q| + 1$ possibilities for α_1^δ and $q = |\mathbb{F}_q|$ possibilities for the other α_i. The number of polynomials $h_{n_i}(y) = h_{n_i,i}y^{n_i} + h_{n_i-1,i}y^{n_i-1} + \cdots + h_{n_i,2}y^2$ of degree n_i is $(q-1)q^{n_i-2}$ (because $h_{n_i,i} \neq 0$). Finally, the cardinality of the affine group is

$$|A| = (q^2 - 1)(q^2 - q)q^2 = q^3(q-1)^2(q+1).$$

If $\lambda = (a_1x + b_1y + c_1, a_2x + b_2y + c_2)$, then we have $q^2 - 1$ possibilities for the nonzero element $a_1x + b_1y$, $q^2 - q$ possibilities for $a_2x + b_2y$ which is linearly independent with $a_1x + b_1y$ and q^2 possibilities for c_1, c_2. Hence, for $n = 1$

$$\begin{aligned} p_1 &= |A| = q^3(q-1)^2(q+1), \\ p_n &= \sum_{n_1\cdots n_k = n} (q+1)q^{k-1}\left(\prod_{i=1}^k (q-1)q^{n_i-2}\right)q^3(q-1)^2(q+1) \\ &= (q+1)^2 \sum_{n_1\cdots n_k = n} (q-1)^{k+2} q^{n_1+\cdots+n_k-k+2} \\ &= \bigl(q(q-1)(q+1)\bigr)^2 \sum_{n_1\cdots n_k=n} \left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k} \end{aligned}$$

when $n > 1$.

(ii) If a_n, b_n, $n = 1, 2, \ldots$, are two sequences, then, see e.g. [W], the product of their Dirichlet series generating functions $a(s)$, $b(s)$ is

$$a(s)b(s) = \left(\sum_{n\geqslant 1} \frac{a_n}{n^s}\right)\left(\sum_{n\geqslant 1} \frac{b_n}{n^s}\right) = \sum_{n\geqslant 1}\left(\sum_{i=1}^{n-1} a_i b_{n-i}\right)\frac{1}{n^s}.$$

Applied to $\rho^k(s)$ this gives

$$\rho^k(s) = \sum_{n\geqslant 2}\left(\sum_{n_1\cdots n_k=n} q^{n_1+\cdots+n_k}\right)\frac{1}{n^s}.$$

Hence

$$\begin{aligned} \sum_{n\geqslant 2}\frac{p_n}{n^s} &= \bigl(q(q-1)(q+1)\bigr)^2 \sum_{k\geqslant 1}\left(\frac{q-1}{q}\right)^k \left(\sum_{n\geqslant 2}\sum_{n_1\cdots n_k=n} q^{n_1+\cdots+n_k}\frac{1}{n^s}\right) \\ &= \bigl(q(q-1)(q+1)\bigr)^2 \sum_{k\geqslant 1}\left(\frac{q-1}{q}\right)^k \rho^k(s), \end{aligned}$$

$$\begin{aligned}
p(s) &= \sum_{n\geqslant 1}\frac{p_n}{n^s} = p_1 + \sum_{n\geqslant 2}\frac{p_n}{n^s} \\
&= \big(q(q-1)(q+1)\big)^2\left(-\frac{1}{q+1} + 1 + \sum_{k\geqslant 1}\left(\frac{q-1}{q}\right)^k \rho^k(s)\right) \\
&= \big(q(q-1)(q+1)\big)^2\left(\sum_{k\geqslant 0}\left(\frac{q-1}{q}\right)^k \rho^k(s) - \frac{1}{q+1}\right) \\
&= \big(q(q-1)(q+1)\big)^2\left(\frac{1}{1-\frac{q-1}{q}\rho(s)} - \frac{1}{q+1}\right). \qquad \square
\end{aligned}$$

As a consequence of the above theorem, we also give a formula for the number of coordinates with degree n in $\mathbb{F}_q[x, y]$ and the corresponding Dirichlet series generating function.

Theorem 2.2.

(i) *The number l_n of coordinates with degree n in $\mathbb{F}_q[x, y]$, is given by the formulas*

$$l_1 = (q-1)q(q+1),$$

$$l_n = \frac{p_n}{(q-1)q(q+1)} = q(q-1)(q+1)\sum\left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k}, \quad n > 1,$$

where the summation takes on all ordered factorizations $n = n_1 \cdots n_k$ of n, with $n_1, \ldots, n_k > 1$.

(ii) *The Dirichlet series generating function $l(s)$ of the sequence l_n, $n = 1, 2, \ldots,$ is*

$$l(s) = \sum_{n\geqslant 1}\frac{l_n}{n^s} = q(q-1)(q+1)\left(\frac{1}{1-\frac{q-1}{q}\rho(s)} - \frac{1}{q+1}\right),$$

where

$$\rho(s) = \sum_{n\geqslant 2}\frac{q^n}{n^s}.$$

Proof. (i) According to the well-known theorem of Jung and van der Kulk [J,K], for a coordinate $f \in K[x, y]$ with $\deg(f) > 1$, two automorphisms (f, g) and (f, g_1) with $\deg(g) < \deg(f)$ and $\deg(g_1) < \deg(f)$ if and only if $g_1 = cg + d$ where $c \in K - \{0\}$, $d \in K$ (so (c, d) has $(q-1)q$ choices), hence for a fixed coordinate $f \in \mathbb{F}_q[x, y]$ with $\deg(f) > 1$, there are $(q-1)q$ automorphisms (f, g) with $\deg(g) < \deg(f)$, therefore $l_n = \frac{z_n}{(q-1)q}$, where z_n is the number of automorphisms (f, g) with $\deg(f) = n > \deg(g)$.

Now we can determine z_n as follows. Tracing back to the proof of Theorem 2.1, we can see in the current case $\alpha_1^\delta = (y, x + a_1 y)^\delta$ only has one choice (i.e. $\delta = 0$) instead of $(q+1)$ choices in the decomposition as we have $\deg(f) > \deg(g)$ now, so $z_n = \frac{p_n}{q+1}$. Therefore,

$$\begin{aligned}
l_n &= \frac{z_n}{(q-1)q} = \frac{1}{(q-1)q(q+1)}p_n \\
&= q(q-1)(q+1)\sum\left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k}.
\end{aligned}$$

When $n=1$, for coordinates $ax+by+c$ $(a,b,c\in\mathbb{F}_q,\ (a,b)\neq(0,0))$, we have $(q-1)q$ choices for (a,b), q choices for c, hence

$$l_1=q\big(q^2-1\big)=(q-1)q(q+1).$$

(ii) It follows from (i), and by similar calculation in the proof of Theorem 2.1(ii). $\square$

The following theorem gives an estimate for the growth of p_n.

Theorem 2.3. *For $n\geqslant 2$ the number of automorphisms of degree n of $\mathbb{F}_q[x,y]$ satisfies the inequalities*

$$(q-1)^3(q+1)^2q^{n+1}\leqslant p_n\leqslant (q-1)^3(q+1)^2q^{n+1}+(\log_2 n)^{\log_2 n}q^{n/2+8}.$$

For a fixed q and any $\varepsilon>0$,

$$p_n=(q-1)^3(q+1)^2q^{n+1}+\mathcal{O}\big(q^{n(1/2+\varepsilon)}\big).$$

Proof. By Theorem 2.1(i), for $n\geqslant 2$

$$p_n=\big(q(q-1)(q+1)\big)^2\sum\left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k},$$

where the summation is on all ordered factorizations $n=n_1\cdots n_k$ of n, with $n_1,\dots,n_k>1$. For $k=1$ we obtain the summand $(q-1)^3(q+1)^2q^{n+1}$. Hence it is sufficient to show that for the number $f(n)$ of ordered factorizations $n=n_1\cdots n_k$ with $k\geqslant 2$ satisfies

$$f(n)\leqslant(\log_2 n)^{\log_2 n}=\mathcal{O}\big(n^{\varepsilon}\big),$$

and for $k\geqslant 2$

$$\big(q(q-1)(q+1)\big)^2\left(\frac{q-1}{q}\right)^k q^{n_1+\cdots+n_k}\leqslant q^{n/2+8}.$$

If $n=p_1\cdots p_m$ is the factorization of n in primes, then the number k in the ordered factorizations is bounded by m. The number of factorizations $f(n)$ is bounded by the number of factorizations $n=n_1\cdots n_m$ in m parts, allowing $n_i=1$ for some i. Hence $f(n)\leqslant m^m$ (because each p_j may participate as a factor of any n_i). Since

$$2^m\leqslant p_1\cdots p_m=n,\quad m\leqslant\log_2 n,$$

we derive the inequality for $f(n)$. Let $\varepsilon>0$. Writing n in the form $n=2^t$, $t=\log_2 n$, we obtain $t^t=\log_2 n^{\log_2 n}$. For $a>1$ and for n sufficiently large (and hence t sufficiently large)

$$\begin{aligned}\log_a f(n)&\leqslant\log_a\big(t^t\big)=t\log_a t\\&\leqslant t^2\leqslant 2^t\varepsilon\log_a q=\log_a\big(q^{2^t\varepsilon}\big)=\log_a\big(q^{n\varepsilon}\big)\end{aligned}$$

which gives the estimate for $f(n)$. For the second inequality, we have

$$\big(q(q-1)(q+1)\big)^2=q^2\big(q^2-1\big)^2<q^6,\quad \frac{q-1}{q}<1,$$

hence we have to show that

$$n_1 + \cdots + n_k \leqslant \frac{n}{2} + 2$$

for $n = n_1 \cdots n_k$, where $n_i \geqslant 2$ and $k \geqslant 2$. We consider the function

$$u(t_1, \ldots, t_{k-1}) = t_1 + \cdots + t_{k-1} + \frac{n}{t_1 \cdots n_{k-1}},$$
$$2 \leqslant t_1 \leqslant \cdots \leqslant t_{k-1} \leqslant t_k = \frac{n}{t_1 \cdots t_{k-1}}.$$

Hence

$$t_1(t_1 \cdots t_{k-1}) \leqslant n,$$
$$\frac{\partial u}{\partial t_1} = 1 - \frac{n}{t_1(t_1 \cdots t_{k-1})} \leqslant 0$$

with $\partial u/\partial t_1 = 0$ for $t_1 = \cdots = t_{k-1} = n/k$ only. Considered as a function of t_1, the function $u(t_1, \ldots, u_{k-1})$ decreases for $t_1 \in [2, t_2]$ and has its maximal value for $t_1 = 2$. Hence

$$u(t_1, \ldots, t_{k-1}) = t_1 + t_2 + \cdots + t_k \leqslant 2 + (t_2 + \cdots + t_k), \quad t_2 \cdots t_k = \frac{n}{2}.$$

If $k = 2$ we already have

$$t_1 + t_2 \leqslant \frac{n}{2} + 2.$$

For $k \geqslant 3$ we have $n \geqslant 8$, $n/4 \geqslant 2$ and by induction

$$t_2 + \cdots + t_k \leqslant \frac{n}{2 \cdot 2} + 2 = \frac{n}{4} + 2,$$
$$t_1 + (t_2 + \cdots + t_k) \leqslant 2 + \left(\frac{n}{4} + 2\right) \leqslant \frac{n}{4} + \left(\frac{n}{4} + 2\right) = \frac{n}{2} + 2. \qquad \square$$

The main contribution $(q-1)^3(q+1)^2q^{n+1}$ to p_n comes from the case when $k = 1$. By Proposition 1.2 this means that such automorphisms are of the form

$$\varphi = (f, g) = \alpha \circ \beta \circ \lambda,$$

where $\alpha = \iota = (x, y)$ or $\alpha = (y, x + ay)$, $a \in \mathbb{F}_q$, $\beta = (x + h(y), y)$, $h(y) \in y^2\mathbb{F}_q[y]$ is a polynomial of degree n and λ is an affine automorphism, i.e., φ has the form from the introduction.

3. Free Nielsen–Schreier algebras

Recall that a variety of algebras over a field K is the class of all (maybe nonassociative) algebras satisfying a given system of polynomial identities. Examples of varieties are the classes of all commutative–associative algebras, all associative algebras, all Lie algebras, all nonassociative algebras, etc. The variety satisfies the Nielsen–Schreier property if the subalgebras of its free algebras are free in the same variety. See, for instance, [MSY]. If a Nielsen–Schreier variety is defined by a homogeneous (with respect to each variable) system of polynomial identities, then the automorphisms of the finitely generated free algebras are tame, see Lewin [L]. We have the following analogue of Theorem 2.1:

Theorem 3.1. *Let $F(x, y)$ be the free $\mathbb{F}_q$-algebra with two generators in a Nielsen–Schreier variety defined by a homogeneous system of polynomial identities and let $F(x, y) \neq 0$. Let c_n be the dimension of all homogeneous polynomials $u(x)$ in one variable of degree n in $F(x, y)$.*

(i) *The number p_n of automorphisms $\varphi = (f, g)$ of degree n of $F(x, y)$ is given by the formulas*

$$p_1 = q^3(q-1)^2(q+1) \quad \textit{for unitary algebras},$$

$$p_1 = q(q-1)^2(q+1) \quad \textit{for nonunitary algebras},$$

$$p_n = p_1(q+1)\sum q^{k-1}\prod_{i=1}^{k}\big(\big(q^{c_{n_i}}-1\big)q^{c_2+\cdots+c_{n_i-1}}\big), \quad n > 1,$$

where the summation is on all ordered factorizations $n = n_1 \cdots n_k$ of n, with $n_1, \ldots, n_k > 1$.

(ii) *The Dirichlet series generating function $p(s)$ of the sequence p_n, $n = 1, 2, \ldots,$ is*

$$p(s) = \frac{p_1}{q}\left(\frac{q+1}{1-q\sigma(s)} - 1\right),$$

where

$$\sigma(s) = \sum_{n \geqslant 2}\big(q^{c_n}-1\big)q^{c_2+\cdots+c_{n-1}}\frac{1}{n^s}.$$

Proof. One of the important properties of free algebras in Nielsen–Schreier varieties defined by homogeneous polynomial identities is the following. If several homogeneous elements in the free algebra are algebraically dependent, then one of them is a polynomial of the others. This fact implies that the automorphisms of finitely generated free algebras are tame. Applied to the free algebra $F(x, y)$ with two generators, this gives that $\mathrm{Aut}(F(x, y)) = A *_C B$, where A is the affine group if we consider unitary algebras and the general linear group when we allow nonunitary algebras, B is the group of triangular automorphisms and $C = A \cap B$. Hence we have an analogue of Proposition 1.2. Counting the elements of degree n in $\mathrm{Aut}(F(x, y))$ we obtain that p_1 is the number $q^3(q-1)^2(q+1)$ of elements of the affine group if the algebra $F(x, y)$ has 1 and the number $q(q-1)^2(q+1)$ of elements of the general linear group $GL_2(\mathbb{F}_q)$ for nonunitary algebras. Then the proof follows the steps of the proof of Theorem 2.1, taking into account that the number of polynomials $h(y) = h_n y^n + \cdots + h_2 y^2$ of degree n is $(q^{c_n} - 1)q^{c_{n-1}} \cdots q^{c_2}$. $\square$

The free Lie algebra and the free anti-commutative algebra in two variables have no elements in one variable of degree > 1 and all automorphisms are linear. Hence

$$p_1 = q(q-1)^2(q+1), \qquad p_n = 0, \quad n > 1.$$

For the free nonassociative algebra the number c_n is equal to the number of nonassociative and noncommutative monomials in one variable, or to the Catalan number and

$$c_n = \frac{1}{n}\binom{2n-2}{n-1}, \quad n \geqslant 1.$$

No explicit expression is known for the number c_n of nonassociative commutative monomials of degree n.

Acknowledgment

The authors would like to thank Arnaud Bodin for bringing their attention to the problem and for helpful comments and suggestions.

References

[B] A. Bodin, Private communication, May 20, 2008.
[C] P.M. Cohn, Free Rings and Their Relations, second ed., London Math. Soc. Monogr. Ser., vol. 19, Academic Press, Inc., London, 1985.
[Cz] A.J. Czerniakiewicz, Automorphisms of a free associative algebra of rank 2. I, Trans. Amer. Math. Soc. 160 (1971) 393–401; A.J. Czerniakiewicz, Automorphisms of a free associative algebra of rank 2. II, Trans. Amer. Math. Soc. 171 (1972) 309–315.
[J] H.W.E. Jung, Über ganze birationale Transformationen der Ebene, J. Reine Angew. Math. 184 (1942) 161–174.
[K] W. van der Kulk, On polynomial rings in two variables, Nieuw Arch. Wiskd. (3) 1 (1953) 33–41.
[L] J. Lewin, On Schreier varieties of linear algebras, Trans. Amer. Math. Soc. 132 (1968) 553–562.
[MKS] W. Magnus, A. Karrass, D. Solitar, Combinatorial Group Theory. Presentations of Groups in Terms of Generators and Relations, second ed., Dover Publications, Inc., New York, 1976.
[MSY] A.A. Mikhalev, V. Shpilrain, J.-T. Yu, Combinatorial Methods. Free Groups, Polynomials, and Free Algebras, CMS Books Math./Ouvrages Math. SMC, vol. 19, Springer-Verlag, New York, 2004.
[ML] L.G. Makar-Limanov, On automorphisms of free algebra with two generators, Funktsional. Anal. i Prilozhen. 4 (3) (1970) 107–108 (in Russian); translation: Funct. Anal. Appl. 4 (1970) 262–263.
[W] H.S. Wilf, Generatingfunctionology, second ed., Academic Press, Inc., Boston, MA, 1994; third edition, A.K. Peters, Ltd., Wellesley, MA, 2006.
[Wr] D. Wright, The amalgamated free product structure of $GL_2(k[X_1,\dots,X_n])$ and the weak Jacobian theorem for two variables, J. Pure Appl. Algebra 12 (3) (1978) 235–251.