ISRAEL JOURNAL OF MATHEMATICS, Vol. 69, No. 2, 1990

ON GROUPS OF AUTOMORPHISMS OF A CLASS OF SURFACES

BY

L. MAKAR-LIMANOV[†]

Department of Mathematics, Wayne State University, Detroit, MI 48202 USA

ABSTRACT

In this note we describe the group of automorphisms of a commutative algebra with three generators x, y and z satisfying a relation $xy = P(z)$, where $P(z)$ is a polynomial.

The algebraic automorphisms of the projective and affine planes have attracted attention since around 1830. From the point of view of an algebraist these are the automorphisms of a field of rational functions with two variables and of a polynomial ring with two variables. The groups of automorphisms of these planes were described around 1900 and 1941, respectively, at least over the complex numbers, by algebro-geometric means. It turned out later that the description given does not really depend on the base field. So the question on the automorphisms of an affine plane is long ago settled. Nevertheless a description of automorphisms of surfaces, even surfaces birationally equivalent to the plane, is far from completion. In work [2] M. Gizatulin and V. Danilov develop an algebro-geometric approach which allows a description of the groups of automorphisms of a class of affine surfaces. It is applicable to surfaces X which may be completed by so-called zigzags. (A zigzag is a projective curve with all irreducible components isomorphic to the projective line.) It happens that under some additional assumptions the completions of X form a tree on which the group Aut(X) acts naturally. In [3] they apply this approach to the case where the corresponding zigzags are irreducible

[†] The author was visiting Bar-Ilan and Tel-Aviv Universities while working on this project. He is deeply thankful for hospitality rendered.

The author is supported by an NSF grant.

Received May 28, 1989 and in revised form August 16, 1989

(i.e. isomorphic to the projective line). One of the possible surfaces here is the 'ordinary' two-dimensional sphere $x^2 + y^2 + z^2 = 1$. This sphere over, say, an algebraically closed field may be given by the equation $uv = 1 - z^2$, in which form it was considered in [3] and corresponds to our situation with $\deg P(z) = 2$.

In 1970 the author suggested a 'purely' algebraic approach to a description of the automorphism group of a polynomial ring with two variables which worked equally well in a noncommutative setting (for the free associative algebra of rank two and for the first Weyl algebra). Here the same approach, as it is described in [1], is used for the surfaces of the form $xy = P(z)$.

Let K be an algebraically closed field and $P(z) \in K[z]$. R will denote the factor algebra $K[x, y, z]/(xy - P(z))$, where $(xy - P(z)) = (xy - P(z))K[x, y, z]$ is the principal ideal generated by $xy - P(z)$.

Let us assume in what follows that the degree d of $P(z)$ is at least two. The case $P(z)$ is a constant is trivial and for a linear $P(z)$ the algebra R is isomorphic to $K[x, y]$, also a well known situation (e.g., see [1]).

The statement which will be proved in this note is the

THEOREM. *The group* $\mathrm{Aut}(R)$ *is generated by the following automorphisms*:

(a) *Hyperbolic rotations* $H(x) = \lambda x$, $H(y) = \lambda^{-1} y$, $H(z) = z$; $\lambda \in K^* (K^* = K \setminus 0)$.

(b) *Involution* $I(x) = y$, $I(y) = x$, $I(z) = z$.

(c) *Triangular* $\Delta(x) = x$, $\Delta(y) = y + [P(z + xf(x)) - P(z)]x^{-1}$, $\Delta(z) = z + xf(x)$; $f(x) \in K[x]$.

(d) *If* $P(z) = c(z + a)^d$ *then rescalings* $R(x) = x$, $R(y) = \lambda^d y$, $R(z) = \lambda z + (\lambda - 1)a$; $\lambda \in K^*$ *should be added.*

(e) *If* $P(z) = (z + a)^i Q((z + a)^n)$ *and* $\mu \in K$ *is such that* $\mu^n = 1$, *then a symmetry* $S(x) = x$, $S(y) = \mu^i y$; $S(z) = \mu z + (\mu - 1)a$ *should be added.*

(f) *Finally, if* $\mathrm{char}\, K = \tau > 0$ *and* $P(z) = Q(z^\tau - a^{\tau - 1} z)$, *then a translation* $T(x) = x$, $T(y) = y$, $T(z) = z - a$ *should be added.*

The theorem will follow from several lemmas. (The scheme of the proof is similar to the one in [1].)

The algebra R may be considered as a subalgebra of the algebra $S = K[x, x^{-1}, z]$. Let us introduce a degree function on S (and therefore on R) by $\deg_{n,m} x^i z^j = ni + mj$ where n and m is a pair of nonnegative real numbers at least one of which is positive. Any element of S may be represented as a sum of homogeneous components relative to this degree. For $s \in S$ let us denote by $d(s)$ the highest degree of any homogeneous component of s and by $|s|_{n,m}$ the component with degree equal to $d(s)$. Let T be a subalgebra of S. Let us denote by $|T|_{n,m}$ the subalgebra of S generated by $|t|_{n,m}$ for $t \in T$. Generally speaking $|T|_{n,m}$ is not a subalgebra of T. We will be dropping subscripts n and m whenever possible.

LEMMA 1. *$R = K[x,z] \oplus yK[y,z]$ and if $m > 0$ then $|R| = K[x, x^{-1}z^d, z]$ where $d = \deg P(z)$.*

PROOF. R is linearly generated by monomials $x^iy^jz^k$, and $x^iy^jz^k$ is equal to either $P(z)^jx^{i-j}z^k$ or $P(z)^iy^{j-i}z^k$. Therefore $R = K[x,z] + K[y,z]$. If $r \in R$ and $r = r_x + r_y$, where $r_x \in K[x,z]$, $r_y \in K[y,z]$, then we may assume that $r_y \in yK[y,z]$. Now $|K[x,z]| = K[x,z]$ and $|yK[y,z]| = x^{-1}z^dK[x^{-1}z^d, z]$ since $y = x^{-1}P(z)$ and $|R| = K[x, x^{-1}z^d, z]$. But $K[x,z] \cap x^{-1}z^dK[x^{-1}z^d, z] = 0$ and hence $R = K[x,z] \oplus yK[y,z]$. □

LEMMA 2. *For any $p, q \in K(x,z)$ such that $d(p) \neq 0$ there exist homogeneous $p_1, q_1 \in K(x,z)$ such that $|K[p, p^{-1}, q]| \subset K[p_1, p_1^{-1}, q_1]$ and $|p| = \lambda p_1^a$ where $\lambda \in K^*$ and a is a natural number.*

PROOF. This is Lemma 6.8.3 from [1]. □

Let $g \in \mathrm{Aut}(R)$ and $p = g(x)$, $q = g(z)$. Note p, p^{-1} and q generate an algebra containing $g(R) = R$.

LEMMA 3. *If n/m is irrational then $|p|$ is either x^a or $(x^{-1}z^d)^a$.*

PROOF. By Lemma 2 there exist homogeneous p_1, q_1 such that $K[p_1, p_1^{-1}, q_1] \supset |K[p, p^{-1}, q]| \supset K[x, z, x^{-1}z^d]$ where $d = \deg_z P(z)$. It is clear from the choice of n and m that p_1 and q_1 are monomials. It is also clear that the mapping D of the multiplicative semigroup $T = |K[p_1, p_1^{-1}, q_1]|$ to the integer vectors $Z \times Z$ given by $x^iz^j \to (i,j)$ will embed T into the halfplane bounded by $D(p_1)$. Therefore $D(p_1)$ should not lie inside of the angle between the vectors $(1,0) = D(x)$ and $(-1,d) = D(x^{-1}z^d)$. On the other hand $|p| \in |R|$ and therefore $D(|p|) \in$ linear span $\{(1,0), (0,1), (-1,d)\}$. Therefore $D(p_1)$ which is collinear with $D(|p|)$ is either $(1,0)$ or $(-1,d)$ which proves the lemma. □

LEMMA 4. *If $|p| = cx^a$ for any choice of positive n and m then $p = c_1x + c_2$ $(c_i \in K, c_1 \neq 0)$.*

PROOF. $p = p_0(x,z) + yp_1(y,z)$. It follows from our assumption that $p_1 = 0$, $p_0 \in K[x]$. (Otherwise $|p|_{\delta,1} \neq cx^a$ for a sufficiently small δ.) Since $g(R) \not\subset$

$K[x]$ it is clear that $\deg_z q > 0$. Now $R = K[p, q] \oplus p^{-1}P(q)K[p^{-1}P(q), q]$ which implies that $x \in K[p]$ because $\deg_z$ of any element from $R \setminus K[p]$ is bigger than zero. Therefore p is a linear polynomial. □

If $p = c_1 x + c_2$ we may assume that $c_1 = 1$ because we can take an automorphism $x \to cx$, $y \to c^{-1}x$, $z \to z$.

Lemma 5. *If $p = x + c$ then $c = 0$.*

Proof. Let $q = q_0(x, z) + yq_1(y, z)$ where q_i are polynomials. If $q_1 \neq 0$, then $|q|_{0,1} = c_2 x^i z^j$, where $j \geqq d$ and $|R|_{0,1} \subset |K[p, p^{-1}, q]| \not\ni z$. So $q_1 = 0$ and $q \in K[x, z]$. Now $g(y) = (x + c)^{-1}P(q)$. There exists an i such that

$$x^i g(y) = x^i (x + c)^{-1} P(q) \in K[x, z]$$

which means that $P(q)$ is divisible by $x + c$ if $c \neq 0$. Then $(x + c)^{-1}P(q) \in K[x, z]$ and $g(R) \subset K[x, z]$ which is impossible. □

Lemma 6. *If $g(x) = x$, then $g(z) = q_0(x) + cz$, where $q_0(x) \in K[x]$ and $c \in K^*$.*

Proof. As it was shown in the proof of the previous lemma $g(x) = x$ implies $g(z) \in K[x, z]$. Since $g^{-1}(x) = x$, $g^{-1}(z) \in K[x, z]$ and therefore g is an automorphism of $K[x, z]$. Let $g(z) = q_0(x) + zq_1(x, z)$. Then $f(x) = x$, $f(z) = zq_1$ is also an automorphism of $K[x, z]$ and $z = f^{-1}(zq_1) = f^{-1}(z) \cdot f^{-1}(q_1)$ is possible only if $q_1 \in K$. □

Lemma 7. *The automorphism Δ of $K[x, z]$ which is given by $\Delta(x) = x$, $\Delta(z) = z + xr(x)$, where $r(x) \in K[x]$ induces an automorphism of R.*

Proof.

$$\Delta(y) = x^{-1}P(z + xr(x)) = y + x^{-1}[P(z + xr) - P(z)] = y + S(x, z)$$

where $S \in K[x, z]$. So Δ is defined on R. It is clear that Δ is invertible. □

Lemma 8. *If $g(x) = x$, $g(z) = c_0 + c_1 z$ then either*
(a) *c_1 is not a root of 1 and $P(z) = c_2(z - c_0(1 - c_1)^{-1})^d$ or*
(b) *c_1 is a root of 1 of degree k ($k \neq 1$) and $P(z) = Q((z - c_0(1 - c_1)^{-1})^k)$ or*
(c) *$c_1 = 1$, $c_0 \neq 0$, $P(z) = Q(z^\tau - c_0^{\tau - 1} z)$ where $\tau = \operatorname{char} K$ or*
(d) *$c_1 = 1$, $c_0 = 0$.*

Proof. Let $g(y) = r_0(x, z) + yr_1(y, z)$. Then $P(c_1 z + c_0) = xr_0(x, z) + P(z)r_1(y, z) \in K[z]$. Therefore $r_0 = 0$, $r_1 \in K[z]$ and $P(c_1 z + c_0) = P(z)r_1(z)$.

Hence $r_1(z) \in K$, $P(c_1 z + c_0) = c_3 P(z)$ and the finite set of roots of $P(z)$ is invariant under transformation $\alpha(z) = c_1 z + c_0$. If c_1 is not a root of 1 then α has a unique fixed point $z_0 = c_0(1 - c_1)^{-1}$ and all other trajectories of α are infinite. So in this case $P(z) = c_2(z - z_0)^d$. If $c_1^k = 1$ where $k > 1$ is minimal possible, then α has order k and preserves $(z - z_0)^k$; hence in this case $P(z) = (z - z_0)^i Q((z - z_0)^k)$. If $c_1 = 1$ and $c_0 \neq 0$ then α does not have fixed points and trajectories are finite only if char $K = \tau > 0$. In this case α has order τ and $P(z) = Q(z^\tau - c_0^{\tau-1} z)$. Finally, if $c_1 = 1$ and $c_0 = 0$ polynomial $P(z)$ may be arbitrary. □

So we have settled the case $|g(x)|_{n,m} = cx^a$ for any choice of $n, m > 0$. The complete list of possible automorphisms is given in Lemmas 7 and 8.

If $|g(x)|_{n,m} = c(x^{-1}z^d)^a$ then since $I(x) = y$, $I(y) = x$, $I(z) = z$ gives an automorphism of R, Ig is also an automorphism and $|Ig(x)|_{n',m'} = cx^a$ with an appropriate choice of n', m'. Hence it remains to consider the case when $|g(x)| = cx^a$ for some choices of n, m and $|g(x)| = c(x^{-1}z^d)^a$ for other choices. It is clear that then there are uniquely defined relatively prime integers n and m for which

$$|g(x)| = c_1 x^\alpha + \cdots + c_2(x^{-1}z^d)^\beta$$

where $c_1, c_2 \in K^*$, $\alpha = \deg_x g(x)$, $d\beta = \deg_z g(x)$. Let us denote these particular n and m by ρ and σ.

Let us apply Lemma 2 to $g(x)$ and $g(z)$ and denote the p_1 and q_1 obtained there by φ and ψ. As we know $|g(x)| = c_1\varphi^a$ and $|R| \subset |K[g(x), g(x)^{-1}, g(z)]| \subset K[\varphi, \varphi^{-1}, \psi]$. With our choice of n and m, $\varphi = x^\gamma + \cdots + \tilde{c}(x^{-1}z^d)^\delta$ for suitable γ, δ.

Lemma 9. *If $|\psi|_{1,0} = cx^i$ then $\sigma \equiv 0 (\operatorname{mod} \rho)$ and there exists a triangular automorphism $\Delta(x) = x, \Delta(z) = z + c_3 x^r$ where $r = \sigma/\rho\, (\Delta(y) = \Delta(x)^{-1} P(\Delta(z)))$, such that $|\Delta g(x)|_{\rho,\sigma} = c_2(x^{-1}z^d)^\beta$.*

Proof. $z \in K[\varphi, \varphi^{-1}, \psi]$. Therefore $\sigma = \deg z$ is a linear combination of $\gamma\rho = \deg$ and $i\rho = \deg \psi$. Now $\Delta|g(x)| = c_1 x^\alpha + \cdots + c_2(x^{-1}(z + c_3 x^r)^d)^\beta$. By the choice of r, $\Delta(z)$ is (ρ, σ) homogeneous, and c_3 may be chosen in such a way that the coefficient with x^α in $\Delta g(x)$ is zero. But then $|\Delta g(x)| = |c_2(x^{-1}z^d)^\beta|$ because otherwise

$$|\Delta g(x)| = c_4 x^j (x^{-1}z^d)^k + \cdots + c_2(x^{-1}z^d)^\beta$$

where $c_4, j, k \neq 0$ and $|\Delta g(x)|_{n,m} = c_4 x^j (x^{-1}z^d)^k$ for some choice of n and m, contradicting the conclusion of Lemma 3. □

LEMMA 10. *If* $|\psi|_{0,1} = c(x^{-1}z^d)^i$ *then there exists a triangular automorphism* Δ *such that* $|\Delta Ig(x)|_{d\sigma-\rho,\sigma} = c_1(x^{-1}z^d)^\alpha$.

PROOF. It is clear that $|Ig(x)|_{d\sigma-\rho,\sigma} = c_2x^\beta + \cdots + c_1(x^{-1}z^d)^\alpha$ and $|I(\psi)|_{1,0} = cx^i$, so we may apply the previous lemma. □

It follows from Lemmas 9 and 10 that if either $|\psi|_{1,0} = cx^i$ or $|\psi|_{0,1} = c(x^{-1}z^d)^i$, then with the help of a triangular automorphism or involution and a triangular automorphism we can decrease $\deg_x g(x) + \deg_z g(x)$. So after a finite number of such steps we obtain an automorphism g' for which either $g'(x) = c_1x$ (and then use Lemma 8) or for corresponding ψ' both $|\psi'|_{1,0} = cx^iz^j$ where $j \neq 0$ and $|\psi'|_{0,1} = c(x^{-1}z^d)^kz^l$ where $l \neq 0$.

LEMMA 11. *If* $|\psi|_{1,0} = cx^iz^j$ *where* $j \neq 0$, *then* $\gamma = j = 1$.

PROOF. $|R| \subset K[x^\gamma, x^{-\gamma}, x^iz^j]$, therefore $K[x^\gamma, x^{-\gamma}]\ni x$ implying $\gamma = 1$. So $K[x, x^{-1}, x^iz^j] = K[x, x^{-1}, z^j]\ni z$ and $j = 1$. □

REMARK. Similar considerations show that if $|\psi|_{0,1} = c'(x^{-1}z^d)^iz^j$ where $j \neq 0$, then $\delta = j = 1$.

So from now on we may assume that $|\psi|_{1,0} = cx^iz$, $|\psi|_{0,1} = c'(x^{-1}z^d)^kz$ and $\varphi = x + \cdots + \tilde{c}x^{-1}z^d$ where $\tilde{c} \neq 0$.

Since $\deg x = \deg x^{-1}z^d$ and ψ may be replaced by $\psi\varphi^i$ where i is any integer, we may assume that $|\psi|_{1,0} = cz$, $|\psi|_{0,1} = c'z$ and $\psi = z \cdot \chi(w)$ where $w = x^{-\sigma}z^\rho$ and $\chi(w) \in K(w)$ is a ratio of two polynomials of the same degree. We may also assume that $\rho > 1$, because otherwise $\sigma \equiv 0 \pmod{\rho}$ and there exists a 'reducing' triangular automorphism (see proof of Lemma 9).

LEMMA 12. *For* φ *and* ψ *as above*, $K[\varphi, \varphi^{-1}, \psi] \not\supset |R|$.

PROOF. If $K[\varphi, \varphi^{-1}, \psi]\ni z$ then $z = \psi \cdot f(\varphi^{-\sigma}\psi^\rho)$ (where $f(t) \in K[t]$) because $\deg z = \deg \psi$. Now $\varphi = x\xi(w)$ where $\xi(w)$ is a polynomial of degree at most two since $2\rho = d\sigma$ and $2\rho \geqq d$.

So $1 = \chi f(w\xi^{-\sigma}\chi^\rho)$. It is clear that $\sigma \leqq \rho$ because $d \geqq 2$. Therefore if $\chi(w_0) = 0$ for some $w_0 \in K$ then ξ must have w_0 as a root of multiplicity two and $\rho - 2\sigma < 0$. But then $x^{-2}z^d = w^2$ and $\sigma = 1$, so $\rho < 2$ contrary to our assumption on ρ. Consequently $\chi(w)$ does not have roots and thus is a constant, since it is a ratio of two polynomials of the same degree. We may assume then that $\psi = z$.

If $K[\varphi, \varphi^{-1}, \psi] \ni x$, then $x = \varphi \cdot h(\varphi^{-\sigma}\psi^{\rho})$ and $1 = \xi h(w\xi^{-\sigma})$ where $h(t) \in K[t]$.

Since $\xi(w) = 1 + \cdots + \tilde{c}x^{-2}z^{d}$, $\xi(0) \neq 0$. Let w_0 be a root of $\xi(w)$ of multiplicity n. Then $\xi h(w\xi^{-\sigma})$ has a pole in w_0 if $n - \sigma kn < 0$ (here $k = \deg h(t)$). So $1 - \sigma k \geqq 0$ and $\sigma = k = 1$. Therefore $1 = \xi(h_0 + h_1 w\xi^{-1}) = h_0\xi + h_1 w$ and ξ is a linear polynomial. But then $x^{-2}z^{d} = w$ and $\sigma = 2$.

Since σ cannot be 1 and 2 simultaneously, the lemma is proved. □

We reduced our assumptions on $|\psi|_{1,0}$ and $|\psi|_{0,1}$ to a contradiction. Therefore any automorphism may be presented as a product of involutions, triangular automorphisms, a hyperbolic rotation and an automorphism from Lemma 8.

Remark. (1) Though we assumed that K is algebraically closed it is not really essential. It is not difficult to show that all roots necessary in Lemma 9 belong to the field itself.

(2) The groups described are subgroups of the group of automorphisms of the polynomial ring $K[x, y, z]$. It would be of considerable interest to describe all factor algebras of $K[x, y, z]$ with this property.

(3) The structure of $\mathrm{Aut}(R)$ is especially nice when $P(z)$ has degree at least 3 and is a 'general' polynomial (i.e. not of the form (d), (e) or (f) of the theorem). Let T be the group of all triangular transformations and let G be the group of hyperbolic rotations. Then $G_0 = \mathrm{span}(T, ITI)$ is the free product of these groups and is a normal subgroup of $\mathrm{Aut}(R)$. Further $\mathrm{Aut}(R)$ is a semi-direct product of G_0 and G_1 where G_1 is a semidirect product of G and $\mathrm{span}(I)$. When $\deg P(z) = 2$ there is a big 'linear' part which makes the structure more complicated.

References

1. P. M. Cohn, *Free Rings and their Relations*, second edition, Academic Press, New York, 1985.

2. V. I. Danilov and M. H. Gizatulin, *Automorphisms of affine surfaces, I*, Izv. Akad. Nauk SSSR, Ser. Mat. **39** (1975), 523–565. English translation: Math. USSR, Izv. **9** (1975), 493–534.

3. V. I. Danilov and M. H. Gizatulin, *Automorphisms of affine surfaces, II*, Izv. Akad. Nauk SSSR, Ser. Mat. **41** (1977), 54–103. English translation: Math. USSR, Izv. **11** (1977), 51–98.