COLLOQUIUM MATHEMATICUM

VOL. 100 2004 NO. 2

THE SOLUTION OF THE TAME GENERATORS CONJECTURE ACCORDING TO SHESTAKOV AND UMIRBAEV

BY

ARNO VAN DEN ESSEN (Nijmegen)

Abstract. The tame generators problem asked if every invertible polynomial map is tame, i.e. a finite composition of so-called elementary maps. Recently in [8] it was shown that the classical Nagata automorphism in dimension 3 is not tame. The proof is long and very technical. The aim of this paper is to present the main ideas of that proof.

Introduction. One of the most fundamental questions in the study of invertible polynomial maps is: how do they all look like?

For invertible linear maps over a field everyone knows from linear algebra that every such map is a finite composition of elementary maps. For invertible polynomial maps one also has a natural notion of an elementary map (see the next section) and the crucial question was: is every invertible polynomial map over a field a finite composition of such elementary maps? This problem is most widely known as the Tame Generators Problem. It remained open for more than 60 years and was recently solved by Shestakov and Umirbaev. The answer is no in dimension 3!

Their proof is very technical and complicated and is given in a series of two papers [7] and [8]. Together these papers are about 50 pages long and still many details are left to the reader!

Our aim here is to present the main ideas of their proof, which may be helpful in the reading of [8] (1).

1. Some history and preliminaries. Let k be any commutative ring. By $k^{[n]}$ or $k[x_1, \ldots, x_n]$ we denote the polynomial ring in n variables over k. A polynomial map $F : k^n \to k^n$ is just an n-tuple $(F_1, \ldots, F_n)$ of polynomials in $k^{[n]}$. Such a map is called *invertible over* k or a *polynomial automorphism of* k^n if there exists a polynomial map $G = (G_1, \ldots, G_n)$ such that $F \circ G = I$, the identity map. Examples of invertible polynomial maps are the so-called

2000 *Mathematics Subject Classification*: 14R10, 14E07.

Key words and phrases: Tame Generators Conjecture, Nagata conjecture, elementary map.

(1) This paper is an extended version of lectures given in Nijmegen, Toruń and Reims.

elementary polynomial maps given by

$$E_{i,c,a} := (x_1, \dots, x_{i-1}, cx_i + a, x_{i+1}, \dots, x_n),$$

where c is a unit in k and a a polynomial in $k^{[n]}$ not containing x_i. The inverse of $E_{i,c,a}$ is the elementary map $E_{i,c^{-1},-c^{-1}a}$. Of course taking finite compositions of such elementary polynomial maps we get much more examples of invertible maps: the group we obtain in this way is called the *tame group* and its elements are called *tame*. Now the crucial question is: are there any other invertible polynomial maps over k?

If k contains non-zero nilpotent elements the answer is easily seen to be yes: namely consider the case $n = 1$ and choose $e \in k$, non-zero, such that $e^2 = 0$. Since tame maps are finite compositions of the affine maps $cx + a$, with $c \in k^*$ and $a \in k$, all tame maps are affine. However the map $F := x + ex^2$ is invertible over k, with inverse $G := x - ex^2$, and clearly F is not affine, hence not tame. On the other hand if we assume that k is a domain (in fact it suffices if k is reduced) then one easily verifies that all invertible maps are affine, hence tame.

If $n = 2$ the situation is more complicated. In case k is a field of characteristic zero Jung [2] showed in 1942 that there are no other invertible polynomial maps. In other words every invertible polynomial map over k is tame. This result was extended by van der Kulk [3] in 1953 to the case of positive characteristic. Furthermore he showed that the tame group is a free amalgamated product of the groups $\mathrm{Aff}(k, 2)$ and $J(k, 2)$ over their intersection, where $\mathrm{Aff}(k, 2)$ is the *affine group* consisting of all invertible affine maps and $J(k, 2)$ is the group of *de Jonquières*, consisting of all invertible polynomial maps of the form $F = (a_1x + f_1(y), a_2y + f_2)$, where $a_1, a_2 \in k^*$, $f_2 \in k$ and $f_1(y) \in k[y]$. This last result also holds in case k is a domain (see [1, 5.1.3]). However if k is a domain which is not a field, then there do exist *wild*, i.e. non-tame invertible polynomial maps over k. Namely in 1972, Nagata [5] made the following observation: choose $0 \neq z \in k$ which is not a unit in k and define $\sigma := s_1^{-1} s_2 s_1$, where $s_1 := (x + z^{-1}y^2, y)$ and $s_2 := (x, y + z^2x)$ (z^{-1} belongs to the quotient field of k). Then the map σ has all its coefficients in k, namely

$$\sigma = (x - 2y(zx + y^2) - z(zx + y^2)^2, y + z(zx + y^2))$$

and one easily verifies that σ is invertible over k. Furthermore it follows from the free amalgamated product structure that σ is not tame over k! Applying this construction to the univariate polynomial ring $k := \mathbb{C}[z]$ Nagata conjectured that the corresponding map of 3-space given by

$$\sigma(x, y, z) = (x - 2y(zx + y^2) - z(zx + y^2)^2, y + z(zx + y^2), z)$$

is not tame over $\mathbb{C}$. Several papers appeared to give evidence to the conjectured wildness of σ, but Nagata's conjecture remained open until the recent work [8] of Shestakov and Umirbaev. On the other hand it was shown by M. Smith [9] in 1989 that σ is 1-tame, i.e., the extended map $\widetilde{\sigma} : \mathbb{C}^4 \to \mathbb{C}^4$ given by $\widetilde{\sigma}(x, y, z, t) = (\sigma(x, y, z), t)$ is tame.

To conclude this section we introduce some notation and give some results which will be used in what follows.

From now on, k will denote a field. If $f \in k^{[n]}$ the homogeneous part of the highest degree of f will be denoted by $\bar{f}$. Now let $F := (f_1, \ldots, f_n)$ be a polynomial map. If F is invertible over k, then it is well known that its Jacobian determinant is a unit in k, i.e. $\det J(f_1, \ldots, f_n) \in k^*$. It follows (see [1, 1.2.9]) that

$$f_1, \ldots, f_n \text{ are algebraically independent over } k. \tag{1.1}$$

On the other hand if F is non-linear it follows that $\det J(\bar{f}_1, \ldots, \bar{f}_n) = 0$, which implies that

$$\bar{f}_1, \ldots, \bar{f}_n \text{ are algebraically dependent over } k. \tag{1.2}$$

Furthermore we define $\deg F := \deg f_1 + \cdots + \deg f_n$.

Now let F and G be polynomial maps over k. If there exists an elementary map E such that $G = E \circ F$ we write $F \to_E G$ or $F \to G$. If furthermore $\deg G < \deg F$ we say that F *admits an elementary reduction to* G and write $F \to_{\text{red}} G$. More precisely we say that f_i is *elementarily reducible* if there exists a polynomial $a \in k^{[n]}$ not containing x_i such that $\deg(f_i - a(f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n)) < \deg f_i$.

Finally, if $f_1, \ldots, f_s$ are some elements of $k^{[n]}$ then the k-subalgebra of $k^{[n]}$ generated by the f_i's is denoted by $\langle f_1, \ldots, f_s \rangle$.

2. The two-dimensional case: Jung's theorem. Throughout this and the next sections k denotes a field of characteristic zero. To understand the work of Shestakov and Umirbaev we first consider the two-dimensional case, i.e. we prove

THEOREM 2.1 (Jung, 1942). *Every automorphism $F = (f, g)$ of k^2 is tame.*

Proof. The theorem follows by induction on $\deg F$ if we can show that

$$F \text{ admits an elementary reduction if } \deg F > 2. \tag{2.2}$$

First we prove (2.2) for the special case that $\bar{f} \in \langle \overline{g} \rangle$: namely then $\bar{f} = c\overline{g}^r$ for some $c \in k^*$ and $r \geq 1$. Consequently, $\deg(f - cg^r) < \deg f$. So if we put $E := (x - cy^r, y)$, then $\deg E \circ F < \deg F$, i.e. F admits an elementary reduction. Similarly (2.2) holds if $\overline{g} \in \langle \bar{f} \rangle$. So taking into account (1.1) and (1.2) we need to study pairs (f, g) which have the following properties: 1) f and g are algebraically independent over k, 2) $\bar{f}$ and $\overline{g}$ are algebraically dependent over k, and 3) $\bar{f} \notin \langle \overline{g} \rangle$ and $\overline{g} \notin \langle \bar{f} \rangle$. Such pairs are called $*$-reduced. They were studied by Shestakov and Umirbaev in [7]. As a consequence of their main result (on these pairs), it follows that f, g cannot form a $*$-reduced pair (see 3.4). In other words, by (1.1) and (1.2) again, condition 3) is not satisfied. So either $\bar{f} \in \langle \bar{g} \rangle$ or $\bar{g} \in \langle \bar{f} \rangle$. Hence we are done by the observations made above. ■

3. Poisson algebras and $*$-reduced pairs

DEFINITION 3.1. Let $A := k[x_1, \ldots, x_n]$ and $f, g \in A$. Then f, g is called *$*$-reduced* if

1) f, g are algebraically independent over k.
2) $\bar{f}, \bar{g}$ are algebraically dependent over k.
3) $\bar{f} \notin \langle \bar{g} \rangle$ and $\bar{g} \notin \langle \bar{f} \rangle$.

The crucial idea of [7] to study these pairs is to embed the polynomial ring $k[x_1, \ldots, x_n]$ in the so-called free Poisson algebra in $x_1, \ldots, x_n$ over k.

DEFINITION 3.2. A *Poisson algebra B* is a k-vector space endowed with two bilinear operations: $(x, y) \mapsto xy$ (multiplication) and $(x, y) \mapsto [x, y]$ (Poisson bracket) such that

(i) B is commutative and associative with respect to ".".
(ii) B is a Lie algebra with repect to $[\cdot]$.
(iii) $[a, bc] = b[a, c] + [a, b]c$ for all $a, b, c \in B$ (Leibniz' rule).

EXAMPLE 1. One easily verifies that the polynomial ring $k[x, y]$ together with the usual multiplication and Poisson bracket given by

$$[f, g] := f_x g_y - f_y g_x$$

is a Poisson algebra. Observe that the (two-dimensional) Jacobian Conjecture in terms of this bracket gets the following form: if $[f, g] = [x, y]$, then $k[f, g] = k[x, y]$.

EXAMPLE 2. (i) An important class of Poisson algebras is given by the following construction. Let L be a Lie algebra with linear basis $e_1, e_2, \ldots$. Denote by $P(L)$ the ring of polynomials in the variables $e_1, e_2, \ldots$. The operation $[x, y]$ of the algebra L can be uniquely extended to a Poisson bracket $[x, y]$ on the algebra $P(L)$ by Leibniz' rule, and $P(L)$ becomes a Poisson algebra [6].

(ii) Now let L be the free Lie algebra with free generators $x_1, \ldots, x_n$. Then $P(L)$ is the *free Poisson algebra* with free generators $x_1, \ldots, x_n$ (see [6]). We will denote this algebra by $PL\langle x_1, \ldots, x_n \rangle$. It becomes a graded ring by putting $\deg x_i = 1$, $\deg [x_i, x_j] = 2$ if $i \neq j$, etc.

In what follows we always use the free Poisson algebra. From the Leibniz rule one easily deduces the formula

$$[f,g] = \sum_{1\le i<j\le n} (f_{x_i} g_{x_j} - f_{x_j} g_{x_i})[x_i, x_j] \quad \text{for all } f, g \in k^{[n]}.$$

This formula implies the following two facts:

$$\deg\,[f,g] \le \deg f + \deg g,$$
$$[f,g] = 0 \text{ iff } f, g \text{ are algebraically dependent over } k.$$

In particular,

$$\text{if } f, g \text{ are algebraically independent over } k, \text{ then } \deg\,[f,g] \ge 2.$$

Now let f, g be a $*$-reduced pair. Then $\bar{f}$ and $\bar{g}$ are algebraically dependent over k. So by Gordan's lemma there exists a polynomial h such that $\bar{f}, \bar{g} \in \langle h\rangle$. Since $\bar{f}$ and $\bar{g}$ are homogeneous it follows that h is homogeneous, $\bar{f} = c_1 h^p$ and $\bar{g} = c_2 h^s$ for some natural numbers p, s and $c_1, c_2 \in k^*$. We can choose h in such a way that $(p, s) = 1$. We may furthermore assume that $n := \deg f \le m := \deg g$. (The reader is warned that this n is not the same as the one used before to indicate the number of variables. However since this notation is used in [8] and will not cause any confusion, as we will be concerned with the 3 variable case only, we decided to keep the notation of [8].) Observe that 3) of 3.1 implies that $n < m$. So $m \ge 2$. Furthermore $n = p \cdot (n, m)$, i.e. $p = n/(n, m)$, and $m = s \cdot (n, m)$, i.e. $s = m/(n, m)$. Instead of saying that f, g is a $*$-reduced pair we sometimes call it a *p-reduced pair*. Also $p \ge 2$, namely if $p = 1$ then $\bar{g} \in \langle \bar{f}\rangle$, contradicting 3) of 3.1. Consequently, if we put

$$N(f,g) := pm - m - n + \deg\,[f,g]$$

then

$$N(f,g) > \deg\,[f,g]. \tag{1}$$

Now we can formulate the main theorem of [7].

THEOREM 3.3. *Let* $G(x,y) \in k[x,y]$ *with* $\deg_y G = pq + r$, *where* $0 \le r < p$. *Then*

$$\deg G(f,g) \ge qN(f,g) + mr.$$

Furthermore, if $\deg_x G(x,y) = q_1 s + r_1$ *with* $0 \le r_1 < s$, *then*

$$\deg G(f,g) \ge q_1 N(f,g) + n r_1.$$

It is this theorem which plays a crucial role in the understanding of tame maps in dimension 3. To demonstrate the power of this theorem we show how it implies the result used in the proof of Jung's theorem which asserts that there do not exist non-linear invertible polynomial maps of k^2 whose components form a $*$-reduced pair. More precisely:

Corollary 3.4. *If* $F = (f, g)$ *is invertible over* k *with* $\deg F > 2$, *then either* $\overline{f} \in \langle \overline{g} \rangle$ *or* $\overline{g} \in \langle \overline{f} \rangle$.

Proof. If the conclusion is not true then (f, g) is a $*$-reduced pair (by (1.1) and (1.2)). Let (G_1, G_2) be the inverse of (f, g). Then $x = G_1(f, g)$. Let $\deg_y G_1 = qp + r$ with $0 \le r \le p - 1$. Then by 3.3 we get

$$1 = \deg x = \deg G_1(f, g) \ge qN(f, g) + mr. \tag{2}$$

Since, as observed above, $\deg[f, g] \ge 2$, it follows from (1) that $N(f, g) > 2$. Since also $m \ge 2$ it follows from (2) that $q = r = 0$. So $\deg_y G_1 = 0$, i.e. $G_1 = G_1(x)$. Hence $x = G_1(f)$, which implies that $f = f(x)$ and $\deg f = 1$. So $\overline{f} = cx$ for some $c \in k^*$. But $\overline{f}$ and $\overline{g}$ are algebraically dependent over k (by 2) of 3.1). So $\overline{g}$ only depends on x. Hence $\overline{g} \in \langle x \rangle = \langle \overline{f} \rangle$, contradicting 3) of 3.1. ∎

To conclude this section we give some useful results concerning $*$-reduced pairs which will be used in Section 6. With the notation introduced above we have

Lemma 3.5. *Let* f, g *be a* $*$*-reduced pair. Then the elements* $f^i g^j$ *with* $j < p$ *all have different degrees.*

Proof. If $\deg f^{i_1} g^{j_1} = \deg f^{i_2} g^{j_2}$ with $j_1 \le j_2 < p$, then $i_1 n + j_1 m = i_2 n + j_2 m$, whence $(i_1 - i_2)n = (j_2 - j_1)m$. Since $(p, s) = 1$ it follows that p divides $j_2 - j_1$. However $0 \le j_2 - j_1 < p$, so $j_2 - j_1 = 0$. Hence $i_1 - i_2 = 0$, i.e. $i_1 = i_2$ and $j_1 = j_2$. ∎

Corollary 3.6. *Under the assumption of Lemma* 3.5, *if* $h = G(f, g)$ *with* $\deg_y G < p$, *then* $\overline{h} \in \langle \overline{f}, \overline{g} \rangle$.

4. Automorphisms admitting a reduction of type I–IV and the main results. In the previous section we saw that Jung's theorem is a consequence of (2.2), i.e. the assertion that every automorphism of k^2 admits an elementary reduction. This immediately leads to the following question:

Question. Does every non-linear tame automorphism of k^3 admit an elementary reduction?

For several years the authors of [8] believed that the answer to this question was affirmative. However in 2001 they discovered the following "exotic" tame automorphism of k^3, i.e. one which does not admit an elementary reduction. It was this discovery which formed the real starting point for their solution of the tame generators problem. Here is their example.

Example. Let $h_1 = x_1$, $h_2 = x_2 + x_1^2$, $h_3 = x_3 + 2x_1x_2 + x_1^3$, $g_1 = 4h_2 + h_3^2$, $g_2 = 6h_1 + 6h_3h_2 + h_3^3$, $g_3 = h_3$. Then $h := (h_1, h_2, h_3)$ and $g := (g_1, g_2, g_3)$ are tame. Let $f = g_2^2 - g_1^3$. Finally, put $f_1 = g_1$, $f_2 = g_2 + (g_3 + f)$, $f_3 = g_3 + f$

and $F = (f_1, f_2, f_3)$. Then F is tame, but does not admit an elementary reduction.

Namely one easily verifies that $\bar{f}_1 = x_1^6$, $\bar{f}_2 = x_1^9$ and $\bar{f}_3 = 12x_1^7x_3 - 12x_1^6x_2^2$. If f_1 is elementarily reducible then $\bar{f}_1 \in \overline{\langle f_2, f_3\rangle}$, which, since $\bar{f}_2$ and $\bar{f}_3$ are algebraically independent over k, implies that $\bar{f}_1 \in \langle \bar{f}_2, \bar{f}_3\rangle$, but this is clearly not the case. Similarly $\bar{f}_2 \notin \overline{\langle f_1, f_3\rangle}$, i.e. f_2 is not elementarily reducible. It remains to see that f_3 is not elementarily reducible. So suppose it is. Then there exists $G(f_1, f_2)$ such that $\bar{f}_3 = \overline{G(f_1, f_2)}$. Observe that f_1, f_2 is a 2-reduced pair. So by 3.3 we get

$$8 = \deg f_3 = \deg G(f_1, f_2) \geq q(2.9 - 9 - 6 + \deg\,[f_1, f_2]) + 9r.$$

Since $\deg\,[f_1, f_2] = 14$ we get $8 \geq q \cdot 17 + 9r$, so $q = r = 0$, i.e. $\deg_y G = 0$. So $\bar{f}_3 = \overline{G(f_1)} \in \langle \bar{f}_1\rangle$, a contradiction.

In fact this example is a special case of the following class of "exotic" automorphisms of k^3 introduced in [8], which all do not admit an elementary reduction.

Definition 4.1. Let $F = (f_1, f_2, f_3) \in \mathrm{Aut}_k\, k^3$. We say that F *admits a reduction of type I* (with active element f_3) if the following conditions are satisfied:

(a) $\deg f_1 = 2n$, $\deg f_2 = sn$, s odd ≥ 3, $2n < \deg f_3 \leq sn$ and $\bar{f}_3 \notin \langle \bar{f}_1, \bar{f}_2\rangle$.

(b) There exists $\alpha \in k^*$ such that $g_1 := f_1$ and $g_2 := f_2 - \alpha f_3$ satisfy:

  (i) g_1, g_2 is a $*$-reduced pair with $\deg g_1 = 2n$ and $\deg g_2 = sn$.

  (ii) $(g_1, g_2, f_3) \to_{\mathrm{red}} (g_1, g_2, g_3)$ with $\deg\,[g_1, g_3] < sn + \deg\,[g_1, g_2]$.

Observe that each f_i has degree > 1. Furthermore such a map F has the property that after a preliminary linear transformation L of the form $L = (x_1, x_2 - \alpha x_3, x_3)$ with $\alpha \in k^*$ the map $L \circ F = (g_1, g_2, g_3)$ admits an elementary reduction, where $\deg g_1 = \deg f_1$ and $\deg g_2 = \deg f_2$. More precisely there exists $a \in \langle g_1, g_2\rangle$ such that $\deg(g_3 - a(g_1, g_2)) < \deg g_3$. In other words, if $E = (x_1, x_2, x_3 - a(x_1, x_2))$ then $\deg E \circ L \circ F < \deg F$. So if F admits a reduction of type I, then it admits a reduction to an automorphism of lower degree (than F) by a sequence of two elementary transformations.

Now the next question is: does every non-linear automorphism of k^3 admit either an elementary reduction or a reduction of type I? It turns out that the situation is much more complicated: in their paper Shestakov and Umirbaev introduce 3 more classes of "exotic" automorphisms, admitting a reduction of type II, III, or IV. Just as in the type I case the components f_i of these automorphisms have very special restrictions on their degrees. In particular it follows that $\deg f_i > 1$ for all i. Furthermore, without going into details we just mention that if an automorphism F admits a reduction of type II it can be reduced to an automorphism G with $\deg G < \deg F$ by a sequence of three elementary transformations, two of which are linear. Similarly F admitting a reduction of type III can be reduced to an automorphism G with $\deg G < \deg F$ by a sequence of three elementary transformations, one of which is linear and another is of the form $(x_1, x_2, x_3) \mapsto (x_1, x_2 - \gamma x_3 - \alpha x_3^2, x_3)$, i.e. quadratic. The type IV reduction is even more complicated since it consists of a sequence of four elementary transformations one of which is linear and two are quadratic. During the reduction process in the type III and IV cases the degree may go up at the intermediate steps, but finally becomes lower than $\deg F$. Now the main theorem of [8] is:

THEOREM 4.2. *Every non-linear tame automorphism of* k^3 *admits either an elementary reduction or a reduction of one of the types* I–IV.

COROLLARY 4.3. *Let* $F = (f_1, f_2, f_3)$ *be a non-linear tame automorphism with* $f_3 = x_3$. *Then* F *admits an elementary reduction.*

Proof. If F admits a reduction of one of the types I–IV then, as observed before, $\deg f_i > 1$ for all i. Since $\deg f_3 = 1$ Theorem 4.2 implies that F admits an elementary reduction. ∎

COROLLARY 4.4. *Let* F *be as in* 4.3. *Then* F *is tame iff* (f_1, f_2) *is tame over* $k[x_3]$.

COROLLARY 4.5. *The Nagata automorphism* $\sigma : \mathbb{C}^3 \to \mathbb{C}^3$ *is not tame.*

5. Sketch of the proof of Theorem 4.2. To prove Theorem 4.2 we introduce the class of so-called simple automorphisms.

DEFINITION 5.1. By induction on $\deg F$ we define *simple* automorphisms of k^3. First, all automorphisms of degree 3, i.e. linear ones, are simple; and if $\deg F > 3$ then F is called simple if it admits either an elementary reduction or a reduction of one of the types I–IV to a simple automorphism G (with $\deg G < \deg F$).

Theorem 4.2 can then be reformulated as

THEOREM 5.2. *If* F *is tame*, *then* F *is simple.*

The converse is obvious since every reduction is done by a sequence of elementary transformations.

We are going to prove this theorem by contradiction. So suppose that there exists a tame automorphism F of k^3 which is not simple. Then we have a sequence

$$F_0 := F \to F_1 \to \cdots \to F_l = I,$$

where F_0 is not simple, but F_l is simple. Let r be maximal such that F_r is not simple. So $r \le l-1$. Then $F_r \to_E F_{r+1}$ with F_r not simple and F_{r+1} simple. Hence $F_{r+1} \to_{E^{-1}} F_r$. So $\theta := F_{r+1}$ and $\tau := F_r$ are tame automorphisms, which satisfy: 1) $\theta \to \tau$, 2) θ is simple, and 3) τ is not simple.

Amongst all pairs θ, τ satisfying 1), 2) and 3) we choose (once and for all) one pair θ_0, τ_0 such that $\deg \theta_0$ is minimal and we write $\theta_0 = (f_1, f_2, f_3)$. Since $\theta_0 \to \tau_0$ we have 3 cases:

(1) $\tau_0 = (f_1 + a(f_2, f_3), f_2, f_3)$,
(2) $\tau_0 = (f_1, f_2 + a(f_1, f_3), f_3)$,
(3) $\tau_0 = (f_1, f_2, f_3 + a(f_1, f_2))$.

Since θ_0 is simple there are 5 cases for θ_0, namely θ_0 admits either an elementary reduction or a reduction of one of the types I–IV to a simple automorphism.

The whole proof consists of showing that in each of the 15 cases τ_0 is simple, which is a contradiction since by definition it is not!

To get an idea of how the simplicity of τ_0 is obtained, we consider from these 15 cases only a relatively easy case, namely when τ_0 is of the form (2) and θ_0 admits a reduction of type I to a simple automorphism. More precisely we show

PROPOSITION 5.3. *If* $\theta_0 = (f_1, f_2, f_3)$ *admits a reduction of type* I *and* $\tau_0 = (f_1, f_2 + a(f_1, f_3), f_3)$, *then* τ_0 *is simple.*

Proof. First we claim

$$\deg a(f_1, f_3) \le \deg f_2. \tag{5.3.1}$$

Namely, suppose $\deg a(f_1, f_3) > \deg f_2$. Then $\deg \tau_0 > \deg \theta_0$. Also $\theta_0 \to_E \tau_0$, whence $\tau_0 \to_{E^{-1}} \theta_0$. So τ_0 admits an elementary reduction to the simple automorphism θ_0 which has lower degree. So by the definition of simplicity this implies that τ_0 is simple, a contradiction. So (5.3.1) holds.

Now *the point is that the estimation* (5.3.1) *gives a very strong restriction on the form of the polynomial* $a(f_1, f_3)$. More precisely we get

LEMMA 5.4. $a(f_1, f_3) = \beta f_3 + T(f_1)$, $\beta \in k^*$, $\deg T(f_1) < \deg f_2$.

The proof of this lemma is the most technical part. Therefore we postpone its proof until the next section.

Now let us show how Lemma 5.4 enables us to prove Proposition 5.3. First, since θ_0 admits a reduction of type I we have

$$\theta_0 = (f_1, f_2, f_3) \to (f_1, \underbrace{f_2 - \alpha f_3}_{:=g_2}, f_3) \to (g_1, g_2, \underbrace{f_3 - g(g_1, g_2)}_{:=g_3})$$

with $g_1 := f_1$, $\deg g_2 = \deg f_2$, $\deg g_3 < \deg f_3$, g_1, g_2 is a $*$-reduced pair, $G := (g_1, g_2, g_3)$ is simple and $\deg [g_1, g_3] < sn + \deg [g_1, g_2]$. By Lemma 5.4 we have $\tau_0 = (f_1, f, f_3)$, where $f := f_2 + \beta f_3 + T(f_1)$ with $T(f_1) < \deg f_2$. Now we distinguish two cases.

CASE 1: $\alpha + \beta \neq 0$. We show that τ_0 admits a reduction of type I to a simple automorphism (hence τ_0 is simple!) namely

$$\begin{aligned}\tau_0 &= (f_1, f, f_3) \to (f_1, f - (\alpha+\beta) f_3, f_3) = (f_1, (f_2 - \alpha f_3) + T(f_1), f_3) \\ &= (g_1, \underbrace{g_2 + T(g_1)}_{:=g_2'}, f_3) \to (g_1, g_2', \underbrace{f_3 - g(g_1, g_2)}_{=g_3}) = (g_1, g_2', \underbrace{f_3 - \widetilde{g}(g_1, g_2')}_{=g_3}).\end{aligned}$$

To see that $\tau_0 = (f_1, f, f_3) \to (f_1, f - (\alpha+\beta) f_3, f_3) \to (g_1, g_2', g_3)$ is a reduction of type I one also needs to check that $\deg g_2' = \deg f$, (g_1, g_2') is a $*$-reduced pair and $\deg [g_1, g_3] < sn + \deg [g_1, g_2']$. We leave this easy verification to the reader. It remains to see that $G' := (g_1, g_2', g_3)$ is simple. Assume that G' is not simple. Since θ_0 admits a reduction of type I to the simple automorphism $G = (g_1, g_2, g_3)$ we see that G is simple and $\deg G < \deg \theta_0$. Also $G \to G'$, since $g_2' = g_2 + T(g_1)$. But G is simple and G' is not simple. Since $\deg G < \deg \theta_0$ this gives a contradiction with the minimal choice of θ_0. So G' is simple, which completes the proof of case 1.

CASE 2: $\alpha + \beta = 0$, i.e. $\beta = -\alpha$. Now we will show that τ_0 admits an elementary reduction to a simple automorphism (hence τ_0 is simple). Namely we get

$$\begin{aligned}\tau_0 = (f_1, f, f_3) = (f_1, \underbrace{f_2 - \alpha f_3}_{=g_2} + T(f_1), f_3) &= (g_1, \underbrace{g_2 + T(g_1)}_{:=g_2'}, f_3) \\ \to (g_1, g_2', \underbrace{f_3 - g(g_1, g_2)}_{=g_3}) &= (g_1, g_2', f_3 - \widetilde{g}(g_1, g_2')).\end{aligned}$$

So $\tau_0 = (f_1, f, f_3) \to (g_1, g_2', g_3)$ is an elementary reduction. By the same argument as above (g_1, g_2', g_3) is simple, which completes the proof of Proposition 5.3. ∎

6. The proof of Lemma 5.4. The aim of this section is to give the complete proof of Lemma 5.4, thereby clearly demonstrating how the fundamental estimates given in Theorem 3.3 play a crucial role. So it suffices to show

THEOREM 6.1. *Let (f_1, f_2, f_3) be an automorphism of k^3 which admits a reduction of type* I. *If $a \in \langle f_1, f_3 \rangle$ satisfies $\deg a \le sn$, then $a = \beta f_3 + T(f_1)$ for some $\beta \in k^*$ and $T(f_1) \in \langle f_1 \rangle$ with $\deg T(f_1) < sn$.*

The main ingredient in the proof is

PROPOSITION 6.2. *Let* (f_1, f_2, f_3) *be as in* 6.1. *Then*

(i) $\deg\,[f_1, f_3] > sn$.
(ii) *If* $a \in \langle f_1, f_3\rangle$, *then either* $\overline{a} \in \langle \overline{f}_1, \overline{f}_3\rangle$ *or* $\deg a > sn$.

It is the second statement which gives sufficient control over the highest degree part of a polynomial in $\langle f_1, f_3\rangle$! Before we prove this result let us first show how it implies 6.1.

Proof of Theorem 6.1. Since $\deg a \leq sn$ it follows from 6.2(ii) that $\overline{a} \in \langle \overline{f}_1, \overline{f}_3\rangle$, so $\overline{a} = \sum c_{ij}\overline{f}_1^i\overline{f}_3^j$ with

$$i \deg f_1 + j \deg f_3 \leq sn. \tag{6.1.1}$$

First we show that terms $\overline{f}_1^i\overline{f}_3^j$ with $j \geq 2$ cannot appear in $\overline{a}$: namely if $j \geq 2$ then $\deg \overline{f}_1^i\overline{f}_3^j \geq \deg \overline{f}_3^2 > \deg f_1 + \deg f_3$ (since $\deg f_3 > 2n = \deg f_1$) $\geq \deg\,[f_1, f_3] > sn$ (by 6.2(i)), which contradicts (6.1.1).

Also the terms with $j = 1$ and $i \geq 1$ cannot appear in $\overline{a}$: namely, for such a term we have $\deg \overline{f}_1^i\overline{f}_3^j \geq \deg f_1 + \deg f_3 \geq \deg\,[f_1, f_3] > sn$ (by 6.2(i)), contradicting (6.1.1) again.

So $\overline{a} = \beta\overline{f}_3 + \lambda\overline{f}_1^r$ with $r \cdot 2n \leq sn$. Now observe that $2r$ is even and s is odd, so $2rn < sn$. Then consider $a_1 := a - \beta f_3 - \lambda f_1^r$. So $\deg a_1 < \deg a$. Repeating the above argument with a_1 instead of a we obtain $a_1 = a_1(f_1)$ and $\deg a_1 < sn$. Hence $a = \beta f_3 + T(f_1)$ with $\deg T(f_1) < sn$. ∎

Proof of Proposition 6.2. Since f_3 is reducible by g_1, g_2 there exists $G(g_1, g_2)$ such that $g_3 := f_3 - G(g_1, g_2)$ satisfies $\deg g_3 < \deg f_3$. So $\overline{f}_3 = \overline{G(g_1, g_2)}$.

CLAIM. $\overline{f}_3 \notin \langle \overline{g}_1, \overline{g}_2\rangle$ (so $\overline{G(g_1, g_2)} \notin \langle \overline{g}_1, \overline{g}_2\rangle$).

CASE 1: $\deg f_3 = sn$. Then $\deg \overline{f}_3 = \deg \overline{f}_2 = \deg \overline{g}_2 = sn$. So $\overline{g}_2 = \overline{f}_2 - \alpha\overline{f}_3$. Now suppose that $\overline{f}_3 \in \langle \overline{g}_1, \overline{g}_2\rangle = \langle \overline{f}_1, \overline{f}_2 - \alpha\overline{f}_3\rangle$. Then, since $\deg \overline{f}_3 = \deg(\overline{f}_2 - \alpha\overline{f}_3) = sn$ (s odd) and $\deg \overline{f}_1 = 2n$, it follows that $\overline{f}_3 = c(\overline{f}_2 - \alpha\overline{f}_3)$ for some $c \in k^*$. So $\overline{f}_2$ and $\overline{f}_3$ are linearly dependent over k. In particular $\overline{f}_3 \in \langle \overline{f}_2\rangle$, contradicting the hypothesis that $\overline{f}_3 \notin \langle \overline{f}_1, \overline{f}_2\rangle$.

CASE 2: $\deg f_3 < sn$. Then $\overline{g}_2 = \overline{f}_2$, so $\langle \overline{g}_1, \overline{g}_2\rangle = \langle \overline{f}_1, \overline{f}_2\rangle$. Since by hypothesis $\overline{f}_3 \notin \langle \overline{f}_1, \overline{f}_2\rangle$ we get $\overline{f}_3 \notin \langle \overline{g}_1, \overline{g}_2\rangle$, which completes the proof of the claim.

(i) Observe that g_1, g_2 is 2-reduced. Write $\deg_y G = q \cdot 2 + r$ with $0 \leq r \leq 1$. If $q = 0$ then $\deg_y G = r \leq 1 < 2\ (= p)$. So it follows from 3.6 that $\overline{G} \in \langle \overline{g}_1, \overline{g}_2\rangle$, which contradicts the claim. So $q \geq 1$. Then by 3.3 we get

$$sn \geq f_3 = \deg G(g_1, g_2) \geq q(2 \cdot sn - sn - 2n + \deg\,[g_1, g_2]) + sn \cdot r.$$

Since $q \geq 1$ it follows that $r = 0$. So $\deg_y G = 2q$ is even. Hence $\deg_y \partial G/\partial y = 2(q-1)+1$. So applying 3.3 to $\partial G/\partial y$ (whose "r" is 1) we get

$$\deg \frac{\partial G}{\partial y} \geq (q-1)N(g_1, g_2) + sn \cdot 1 \geq sn. \tag{6.2.1}$$

Now observe that

$$[f_1, f_3] = [g_1, g_3 + G(g_1, g_2)] = [g_1, g_3] + \frac{\partial G}{\partial y}(g_1, g_2)[g_1, g_2].$$

Since $\deg\,[g_1, g_3] < \deg\,[g_1, g_2] + sn$ by 4.1(b)(ii), and $\deg \frac{\partial G}{\partial y}(g_1, g_2)[g_1, g_2] \geq sn + \deg\,[g_1, g_2]$ (by (6.2.1)), we get

$$\deg\,[f_1, f_3] \geq sn + \deg\,[g_1, g_2] > sn.$$

(ii) Now let $a \in \langle f_1, f_3 \rangle$, say $a = G(f_1, f_3)$. If $\overline{f}_1, \overline{f}_3$ are algebraically independent over k, then $\overline{a} \in \langle \overline{f}_1, \overline{f}_3 \rangle$. So assume that $\overline{f}_1, \overline{f}_3$ are algebraically dependent over k. Then f_1, f_3 is a 2-reduced pair (because $\deg f_1 = 2n$ and $\deg f_3 = sn$, with s odd). Write $\deg_y G = q \cdot 2 + r$ with $0 \leq r \leq 1$. If $q = 0$ then $\deg_y G = r \leq 1 < 2$ $(= p)$, so by 3.6, $\overline{a} \in \langle \overline{f}_1, \overline{f}_3 \rangle$. So let $q \geq 1$. Then by 3.3,

$$\deg a = \deg G(f_1, f_3) \geq qN(f_1, f_3) + sn \cdot r \geq N(f_1, f_3) > \deg\,[f_1, f_3]$$

(by (1) in Section 3). Since $\deg\,[f_1, f_3] > sn$ by (i), this completes the proof of 6.2. ∎

7. Final comments. The method described in [8] even gives an algorithm to decide if a given polynomial automorphism of k^3 is tame. More precisely it decides if a given automorphism F admits an elementary reduction or a reduction of one of the types I–IV. To decide if F admits a reduction of one of the types I–IV one needs various technical parts of the proof. Therefore we only show how one can decide if F admits an elementary reduction.

So let $F = (f_1, f_2, f_3)$. We show how to decide if f_3 is elementarily reducible by f_1, f_2. If $\overline{f}_1, \overline{f}_2$ are algebraically independent over k, then f_3 is reducible iff $\overline{f}_3 \in \langle \overline{f}_1, \overline{f}_2 \rangle$ and this question is easy to decide either by Gröbner basis methods or using the homogeneity of the $\overline{f}_i$. So assume that $\overline{f}_1, \overline{f}_2$ are algebraically dependent over k. If $\overline{f}_2 \in \langle \overline{f}_1 \rangle$, then $\overline{f}_2 = c\overline{f}_1^t$ for some $c \in k^*$ and $t \geq 1$. Observe that f_3 is reducible in F iff it is reducible in $F' := (f_1, f_2 - cf_1^t, f_3)$. Since $\deg F' < \deg F$ the desired result follows by induction on the degree. A similar argument holds if $\overline{f}_1 \in \langle \overline{f}_2 \rangle$. So we may assume that f_1, f_2 is a $*$-reduced pair and that $\deg f_1 < \deg f_2$.

Now suppose that f_3 is reducible by f_1 and f_2. Then there exists a polynomial $G(x, y) \in k[x, y]$ such that $\overline{f}_3 = \overline{G(f_1, f_2)}$. Write $\deg_y G = qp + r$

with $0 \leq r < p$. Then by 3.3 and the fact that $N(f_1, f_2) \geq 1$ we get

$$\deg f_3 = \deg G(f_1, f_2) \geq qN(f_1, f_2) + mr \geq q.$$

So $q \leq \deg f_3$. Also $r < p \leq \deg f_1$. So $\deg_y G = qp + r \leq C := \deg f_3 \cdot \deg f_1 + \deg f_1$. Similarly using the second degree estimate in 3.3 involving $\deg_x G$ we get $\deg_x G \leq C$. Hence $G(f_1, f_2)$ belongs to the finite-dimensional k-vector space V generated by the monomials $f_1^i f_2^j$ with $i, j \leq C$. So if we define $\overline{V}$ to be the finite-dimensional k-vector space generated by the highest degree homogeneous parts of the elements of V, then we infer that f_3 is reducible by f_1 and f_2 iff $\bar{f}_3$ belongs to $\overline{V}$, and this question is easy to decide by linear algebra.

To conclude this paper let us mention some interesting open problems.

Problem 1. Do there exist tame automorphisms of type II–IV?

Problem 2. What happens if k has positive characteristic? Do there exist non-tame automorphisms of k^3?

In this respect the following 2001 result of Stefan Maubach [4] is interesting. If k is a finite field and F an automorphism of k^n, then obviously it induces a bijection on k^n, which we denote by $\mathcal{E}(F)$. So this bijection has a sign, i.e. it is either odd or even.

Theorem 7.1. *Let $k = \mathbb{F}_{2^m}$ with $m \geq 2$. If $F \in \operatorname{Aut}_k k^n$ is tame, then $\mathcal{E}(F)$ is even.*

This leads to the following problem:

Problem 3. Let $k = \mathbb{F}_{2^m}$ with $m \geq 2$. Does there exist $F \in \operatorname{Aut}_k k^3$ with $\mathcal{E}(F)$ odd?

To formulate the last problem we make the following observations: in 1942 Jung proved the 2-dimensional case of the tame generators problem, in 1972 Nagata constructed his candidate counterexample and finally in 2002 Shestakov and Umirbaev solved the 3-dimensional case. This leads to

The 30-years cycle problem. What happens in 2032?

REFERENCES

[1] A. van den Essen, *Polynomial Automorphisms and the Jacobian Conjecture*, Progr. Math. 190, Birkhäuser, 2000.

[2] H. Jung, *Über ganze birationale Transformationen der Ebene*, J. Reine Angew. Math. 184 (1942), 161–174.

[3] W. van der Kulk, *On polynomial rings in two variables*, Nieuw Arch. Wisk. 3 (1953), no. 1, 33–41.

[4] S. Maubach, *Polynomial automorphisms over finite fields*, Serdica Math. J. 27 (2001), 343–350.

[5] M. Nagata, *On Automorphism Group of* $k[X, Y]$, Kyoto Univ. Lectures in Math. 5., Kyoto Univ., Kinokuniya, Tokyo, 1972.

[6] I. Shestakov, *Quantization of Poisson superalgebras and speciality of Jordan Poisson superalgebras*, Algebra and Logic 32 (1993), 309–317.

[7] I. Shestakov and U. Umirbaev, *Poisson brackets and two-generated subalgebras of rings of polynomials*, J. Amer. Math. Soc. 17 (2004), 181–196.

[8] —, —, *The tame and wild automorphisms of polynomial rings in three variables*, ibid., 197–227.

[9] M. Smith, *Stably tame automorphisms*, J. Pure Appl. Algebra 58 (1989), 209–212.

Department of Mathematics
Postbus 9010
6500 GL Nijmegen
The Netherlands
E-mail: essen@math.kun.nl

Received 17 March 2004;
revised 21 May 2004 (4441)