J. Math. Soc. Japan
Vol. 60, No. 2 (2008) pp. 495–510
doi: 10.2969/jmsj/06020495

A generalization of the Shestakov-Umirbaev inequality

By Shigeru Kuroda

(Received Jul. 17, 2007)

Abstract. We give a generalization of the Shestakov-Umirbaev inequality which plays an important role in their solution of the Tame Generators Problem on the automorphism group of a polynomial ring. As an application, we give a new necessary condition for endomorphisms of a polynomial ring to be invertible, which implies Jung's theorem in case of two variables.

1. Introduction.

Let k be a field, and $k[\mathbf{x}] = k[x_1, \ldots, x_n]$ the polynomial ring in n variables over k for $n \in \boldsymbol{N}$. Assume that $\Phi = \sum_{i=0}^{l} \phi_i y^i$ is a polynomial in y over $k[\mathbf{x}]$, where $l \geq 0$ and $\phi_0, \ldots, \phi_l \in k[\mathbf{x}]$. For each $g \in k[\mathbf{x}]$, we set $\Phi(g) = \sum_{i=0}^{l} \phi_i g^i$. Then, it follows that

$$\deg^g \Phi := \max\{\deg(\phi_i g^i) \mid i = 0, \ldots, l\} \geq \deg \Phi(g)$$

in general. Here, $\deg f$ denotes the total degree of f for each $f \in k[\mathbf{x}]$. Shestakov-Umirbaev [**6**, Theorem 3] proved an inequality which estimates the difference between $\deg^g \Phi$ and $\deg \Phi(g)$. Using this result, they settled in [**7**] an important open problem on automorphisms of $k[\mathbf{x}]$ as follows.

Let $\sigma : k[\mathbf{x}] \to k[\mathbf{x}]$ be a homomorphism of k-algebras. Then, σ is an isomorphism if and only if

$$k[\sigma(x_1), \ldots, \sigma(x_n)] = k[\mathbf{x}]. \tag{1.1}$$

For example, σ is an isomorphism if there exist $(a_{i,j})_{i,j} \in GL_n(k)$ and $(b_i)_i \in k^n$ such that $\sigma(x_i) = \sum_{j=1}^{n} a_{i,j} x_j + b_i$ for each i. It also follows that σ is an isomorphism if there exists $l \in \{1, \ldots, n\}$ such that $\sigma(x_i) = x_i$ for each $i \neq l$ and

---

2000 *Mathematics Subject Classification.* Primary 14R10; Secondary 12H05.

Key Words and Phrases. polynomial automorphism, Tame Generators Problem.

Partly supported by the Grant-in-Aid for Young Scientists (Start-up) 19840041, the Ministry of Education, Culture, Sports, Science and Technology, Japan.

$\sigma(x_l) = \alpha x_l + f$ for some $\alpha \in k^\times$ and $f \in k[x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n]$. An automorphism of $k[\mathbf{x}]$ as in the former example is said to be affine, and one as in the latter example is said to be elementary. Because an invertible matrix is expressed as a product of elementary matrices, each affine automorphism can be obtained by the composition of elementary automorphisms. Then, a problem arises whether the automorphism group $\mathrm{Aut}_k\, k[\mathbf{x}]$ can be generated by elementary automorphisms. This is called the Tame Generators Problem because the subgroup of $\mathrm{Aut}_k\, k[\mathbf{x}]$ generated by elementary automorphisms is called the tame subgroup, and its elements are called tame automorphisms. If $n = 1$, then every automorphism of $k[\mathbf{x}]$ is elementary, and so tame. If $n = 2$, then $\mathrm{Aut}_k\, k[\mathbf{x}]$ is equal to the tame subgroup, which was shown by Jung [**2**] in 1942 in case k is of characteristic zero, and by van der Kulk [**3**] in 1953 for an arbitrary k. We note that this result is an immediate consequence of the following proposition.

Proposition 1.1. *If* $n = 2$, *then* $\deg \sigma(x_1) | \deg \sigma(x_2)$ *or* $\deg \sigma(x_2) | \deg \sigma(x_1)$ *holds for each* $\sigma \in \mathrm{Aut}_k\, k[\mathbf{x}]$.

Here, $a|b$ denotes that b is divisible by a for each $a, b \in \boldsymbol{N}$. Note that $\deg \sigma(x_i) > 0$ for each i, since $\sigma(x_i)$ is not contained in k by (1.1).

When $n \geq 3$, the problem becomes extremely difficult. In case of $n = 3$, Nagata [**4**] conjectured in 1972 that the automorphism $\tau \in \mathrm{Aut}_k\, k[\mathbf{x}]$ defined by

$$\begin{aligned} \tau(x_1) &= x_1 - 2(x_1x_3 + x_2^2)x_2 - (x_1x_3 + x_2^2)^2 x_3, \\ \tau(x_2) &= x_2 + (x_1x_3 + x_2^2)x_3, \\ \tau(x_3) &= x_3 \end{aligned} \tag{1.2}$$

is not tame. This conjecture was well-known, but was not settled for a long time. In 2004, however, Shestakov-Umirbaev [**7**] finally showed that Nagata's conjecture is true if k is of characteristic zero. The inequality mentioned at the beginning played a crucial role in their solution of Nagata's conjecture.

The Tame Generators Problem was thus settled for $n = 3$, but remains open for $n \geq 4$. We note that the extension $\tilde{\tau} \in \mathrm{Aut}_k\, k[\mathbf{x}]$ of Nagata's automorphism τ in (1.2) defined by $\tilde{\tau}(x_i) = \tau(x_i)$ for $i = 1, 2, 3$ and $\tilde{\tau}(x_i) = x_i$ for $i = 4, \ldots, n$ is tame (see [**5**]).

The argument in [**7**] is indeed difficult, but employs no advanced results other than those in [**6**]. Therefore, the results in [**6**] are of great importance. However, the proofs of them are also difficult, and, consequently, the proof of this landmark work of Shestakov-Umirbaev seems unfortunately not to be widely understood.

The purpose of the present paper is to generalize the results in [**6**]. Our technique is quite simple and elementary, but the obtained results are powerful.

Indeed, we derive a theorem (Theorem 4.3) which is a generalization of Proposition 1.1, and hence a generalization of Jung's theorem, as a consequence. Of course, our results are useful for a better understanding of the proof of Nagata's conjecture. Furthermore, we can improve the proof by utilizing our results. Recently, as a consequence of the main theorem of this paper, the author showed that there does not exist a tame automorphism which admits a so-called "reduction of type IV" defined by Shestakov-Umirbaev to solve Nagata's conjecture. We believe that the results in the present paper will be helpful to solve the Tame Generators Problem for $n \geq 4$.

In Section 2, we prove the main theorem (Theorem 2.1), whose proof is quite short. We derive some consequences of the main theorem in Section 3, and apply them to studying automorphisms of $k[\mathbf{x}]$ in Section 4. Section 5 is devoted to generalizing a lemma of Shestakov-Umirbaev [**6**, Lemma 5] which also plays an important role in the solution of Nagata's conjecture.

We note that Vénéreau [**8**, Theorem 1, Corollary 4] independently obtained results similar to Theorems 3.3 and 4.3 quite recently. Van den Essen–Makar-Limanov–Willems [**1**] also gave another proof of [**6**, Theorem 3] in a different fashion.

The author would like to thank Professor Eric Edo and the referee for helpful comments and suggestions.

2. Differentials.

In what follows, we always assume that k is of characteristic zero. First, we introduce some terminology concerning the grading of a polynomial ring.

Let Γ be a totally ordered additive group, and $\mathbf{w} = (w_1, \ldots, w_n)$ an element of Γ^n. We define the $\mathbf{w}$-weighted grading $k[\mathbf{x}] = \bigoplus_{\gamma\in\Gamma} k[\mathbf{x}]_\gamma$ by setting $k[\mathbf{x}]_\gamma$ to be the k-vector space generated by $x_1^{a_1} \cdots x_n^{a_n}$ for $a_1, \ldots, a_n \in \boldsymbol{Z}_{\geq 0}$ with $\sum_{i=1}^n a_i w_i = \gamma$ for each $\gamma \in \Gamma$. Here, $\boldsymbol{Z}_{\geq 0}$ denotes the set of nonnegative integers, and $l\gamma$ denotes the sum of l copies of γ for each $l \in \boldsymbol{Z}_{\geq 0}$ and $\gamma \in \Gamma$. Assume that $f = \sum_{\gamma\in\Gamma} f_\gamma$ is an element of $k[\mathbf{x}]$, where $f_\gamma \in k[\mathbf{x}]_\gamma$ for each γ. If $f \neq 0$, then the $\mathbf{w}$-degree $\deg_{\mathbf{w}} f$ of f is defined to be the maximum among $\gamma \in \Gamma$ with $f_\gamma \neq 0$. If $f = 0$, then we set $\deg_{\mathbf{w}} f = -\infty$, i.e., a symbol which is less than each element of Γ. The addition is defined by $(-\infty) + \gamma = \gamma + (-\infty) = -\infty$ for each $\gamma \in \Gamma \cup \{-\infty\}$, and the sum of l copies of $-\infty$ is denoted by $l(-\infty)$ for each $l \in \boldsymbol{Z}_{\geq 0}$. We say that f is $\mathbf{w}$-homogeneous if $f = f_\gamma$ for some γ. In case $f \neq 0$, we define $f^{\mathbf{w}} = f_\delta$, where $\delta = \deg_{\mathbf{w}} f$. Then, $\deg_{\mathbf{w}} f^{\mathbf{w}} = \deg_{\mathbf{w}} f$, $\deg_{\mathbf{w}}(f - f^{\mathbf{w}}) < \deg_{\mathbf{w}} f$, and $(f_1 f_2)^{\mathbf{w}} = f_1^{\mathbf{w}} f_2^{\mathbf{w}}$ hold for each $f, f_1, f_2 \in k[\mathbf{x}] \setminus \{0\}$. Let $\Gamma_{\geq 0}$ be the set of $\gamma \in \Gamma$ with $\gamma \geq 0$, where 0 denotes the zero element of the additive group Γ. We remark that $\deg_{\mathbf{w}} f \geq 0$ holds for each $f \in k[\mathbf{x}] \setminus \{0\}$ whenever $\mathbf{w}$ belongs to $(\Gamma_{\geq 0})^n$. If

$\Gamma = \boldsymbol{Z}$ and $\mathbf{w} = (1, \ldots, 1)$, then the $\mathbf{w}$-degree is the same as the total degree.

Now, for $\Phi \in k[\mathbf{x}][y]$ and $g \in k[\mathbf{x}]$, we define

$$\deg_{\mathbf{w}}^{g} \Phi = \max\{\deg_{\mathbf{w}}(\phi_i g^i) \mid i \in \boldsymbol{Z}_{\geq 0}\}, \tag{2.1}$$

where $\phi_i \in k[\mathbf{x}]$ for each $i \in \boldsymbol{Z}_{\geq 0}$ with $\Phi = \sum_i \phi_i y^i$. Then, $\deg_{\mathbf{w}}^{g} \Phi$ is at least $\deg_{\mathbf{w}} \Phi(g)$ in general. The purpose of this section is to prove an inequality which describes the difference between $\deg_{\mathbf{w}} \Phi(g)$ and $\deg_{\mathbf{w}}^{g} \Phi$.

Let $\partial_y^i \Phi$ denote the i-th order derivative of Φ in y for each $i \in \boldsymbol{Z}_{\geq 0}$, and $\deg_y \Phi$ the degree of Φ in y. Obviously, $\deg_{\mathbf{w}}^{g} \partial_y^i \Phi = \deg_{\mathbf{w}}(\partial_y^i \Phi)(g)$ if $i \geq \deg_y \Phi$. So, we may define a nonnegative integer $m_{\mathbf{w}}^{g}(\Phi)$ by

$$m_{\mathbf{w}}^{g}(\Phi) = \min\left\{i \in \boldsymbol{Z}_{\geq 0} \mid \deg_{\mathbf{w}}^{g} \partial_y^i \Phi = \deg_{\mathbf{w}}\left(\partial_y^i \Phi\right)(g)\right\}. \tag{2.2}$$

If $m_{\mathbf{w}}^{g}(\Phi) \geq 1$ and $g \neq 0$, then we have

$$m_{\mathbf{w}}^{g}(\Phi) = m_{\mathbf{w}}^{g}(\partial_y \Phi) + 1 \quad \text{and} \quad \deg_{\mathbf{w}}^{g} \Phi = \deg_{\mathbf{w}}^{g} \partial_y \Phi + \deg_{\mathbf{w}} g, \tag{2.3}$$

where $\partial_y \Phi = \partial_y^1 \Phi$, since k is of characteristic zero.

Let $\Omega_{k[\mathbf{x}]/k}$ be the differential module of $k[\mathbf{x}]$ over k, and $\bigwedge^r \Omega_{k[\mathbf{x}]/k}$ the r-th exterior power of the $k[\mathbf{x}]$-module $\Omega_{k[\mathbf{x}]/k}$ for $r \in \boldsymbol{N}$. Then, each $\omega \in \bigwedge^r \Omega_{k[\mathbf{x}]/k}$ is uniquely expressed as

$$\omega = \sum_{1 \leq i_1 < \cdots < i_r \leq n} f_{i_1,\ldots,i_r} dx_{i_1} \wedge \cdots \wedge dx_{i_r},$$

where $f_{i_1,\ldots,i_r} \in k[\mathbf{x}]$ for each $i_1, \ldots, i_r$. Here, df denotes the differential of f for each $f \in k[\mathbf{x}]$. We define the $\mathbf{w}$-degree of ω by

$$\deg_{\mathbf{w}} \omega = \max\{\deg_{\mathbf{w}}(f_{i_1,\ldots,i_r}) + w_{i_1} + \cdots + w_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq n\}.$$

Since $df = \sum_{i=1}^{n} (\partial f/\partial x_i) dx_i$ and k is of characteristic zero, the equality

$$\deg_{\mathbf{w}} df = \max\left\{\deg_{\mathbf{w}}\left(\frac{\partial f}{\partial x_i}\right) + w_i \,\middle|\, i = 1, \ldots, n\right\} = \deg_{\mathbf{w}} f \tag{2.4}$$

holds for each $f \in k[\mathbf{x}] \setminus k$. Obviously, $\deg_{\mathbf{w}} df < \deg_{\mathbf{w}} f$ if f is an element of k. It is easily verified that $\deg_{\mathbf{w}}(\omega + \omega') \leq \max\{\deg_{\mathbf{w}} \omega, \deg_{\mathbf{w}} \omega'\}$,

$$\deg_{\mathbf{w}}(f\omega) = \deg_{\mathbf{w}} f + \deg_{\mathbf{w}} \omega, \quad \deg_{\mathbf{w}}(\omega \wedge \eta) \leq \deg_{\mathbf{w}} \omega + \deg_{\mathbf{w}} \eta \tag{2.5}$$

for each $f \in k[\mathbf{x}]$, $\omega, \omega' \in \bigwedge^r \Omega_{k[\mathbf{x}]/k}$ and $\eta \in \bigwedge^s \Omega_{k[\mathbf{x}]/k}$, where $r, s \in \boldsymbol{N}$.

Here is our main theorem.

THEOREM 2.1. *Let $f_1, \ldots, f_r$ be elements of $k[\mathbf{x}]$ for $r \geq 1$ which are algebraically independent over k, and let $\omega = df_1 \wedge \cdots \wedge df_r$. Then, the inequality*

$$\deg_{\mathbf{w}} \Phi(g) \geq \deg_{\mathbf{w}}^g \Phi + m_{\mathbf{w}}^g(\Phi)(\deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega - \deg_{\mathbf{w}} g) \tag{2.6}$$

holds for each $\Phi \in k[f_1, \ldots, f_r][y]$, $g \in k[\mathbf{x}] \setminus \{0\}$ and $\mathbf{w} \in \Gamma^n$.

PROOF. Note that $h_1, \ldots, h_s$ are algebraically dependent over k if and only if $dh_1 \wedge \cdots \wedge dh_s = 0$ for each $h_1, \ldots, h_s \in k[\mathbf{x}]$ for $s \in \boldsymbol{N}$, since k is of characteristic zero. Actually, $dh_1 \wedge \cdots \wedge dh_s = 0$ if and only if each maximal minor of the s by n matrix $(\partial h_i / \partial x_j)_{i,j}$ is zero. Therefore, $\omega \wedge df_i = 0$ for $i = 1, \ldots, r$, while $\omega \neq 0$ by assumption. By chain rule, we may write $d(\Phi(g)) = (\partial_y \Phi)(g) dg + \sum_{i=1}^r \psi_i df_i$, where $\psi_i \in k[\mathbf{x}]$ for each i. Thus,

$$\omega \wedge d(\Phi(g)) = (\partial_y \Phi)(g) \omega \wedge dg + \sum_{i=1}^r \psi_i \omega \wedge df_i = (\partial_y \Phi)(g) \omega \wedge dg.$$

By means of (2.4), (2.5) and the equality above, we get

$$\begin{aligned} \deg_{\mathbf{w}} \omega + \deg_{\mathbf{w}} \Phi(g) &\geq \deg_{\mathbf{w}} \omega + \deg_{\mathbf{w}} d(\Phi(g)) \geq \deg_{\mathbf{w}}(\omega \wedge d(\Phi(g))) \\ &= \deg_{\mathbf{w}}((\partial_y \Phi)(g) \omega \wedge dg) = \deg_{\mathbf{w}}(\partial_y \Phi)(g) + \deg_{\mathbf{w}}(\omega \wedge dg). \end{aligned}$$

Subtracting $\deg_{\mathbf{w}} \omega$ from both sides of this inequality yields

$$\deg_{\mathbf{w}} \Phi(g) \geq \deg_{\mathbf{w}}(\partial_y \Phi)(g) + \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega. \tag{2.7}$$

Now, we show (2.6) by induction on $m_{\mathbf{w}}^g(\Phi)$. If $m_{\mathbf{w}}^g(\Phi) = 0$, then $\deg_{\mathbf{w}} \Phi(g) = \deg_{\mathbf{w}}^g \Phi$ by the definition of $m_{\mathbf{w}}^g(\Phi)$. In this case, (2.6) is clear. Assume that $m_{\mathbf{w}}^g(\Phi) \geq 1$. Then, $m_{\mathbf{w}}^g(\partial_y \Phi)$ is less than $m_{\mathbf{w}}^g(\Phi)$ by (2.3). By induction assumption, and by the equalities in (2.3), we obtain

$$\begin{aligned} \deg_{\mathbf{w}}(\partial_y \Phi)(g) &\geq \deg_{\mathbf{w}}^g \partial_y \Phi + m_{\mathbf{w}}^g(\partial_y \Phi) M \\ &= (\deg_{\mathbf{w}}^g \Phi - \deg_{\mathbf{w}} g) + (m_{\mathbf{w}}^g(\Phi) - 1) M, \end{aligned}$$

where $M = \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega - \deg_{\mathbf{w}} g$. Using (2.7) and the inequality above, we arrive at

$$\begin{aligned}\deg_{\mathbf{w}} \Phi(g) &\geq \deg_{\mathbf{w}}(\partial_y \Phi)(g) + \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega \\ &\geq (\deg_{\mathbf{w}}^g \Phi - \deg_{\mathbf{w}} g) + (m_{\mathbf{w}}^g(\Phi) - 1)M + \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega \\ &= \deg_{\mathbf{w}}^g \Phi + m_{\mathbf{w}}^g(\Phi)(\deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega - \deg_{\mathbf{w}} g).\end{aligned}$$

Therefore, the inequality (2.6) is true. □

3. The Shestakov-Umirbaev inequality.

In this section, we derive some consequences of Theorem 2.1.

First, we remark that the element $\deg_{\mathbf{w}}^g \Phi$ of Γ defined as in (2.1) is equal to the $(\mathbf{w}, \deg_{\mathbf{w}} g)$-degree of Φ for each $\Phi \in k[\mathbf{x}][y] \setminus \{0\}$, $g \in k[\mathbf{x}] \setminus \{0\}$ and $\mathbf{w} \in \Gamma$, where we regard Φ as a polynomial in the $n+1$ variables $x_1, \ldots, x_n$ and y over k. We denote $\Phi^{(\mathbf{w}, \deg_{\mathbf{w}} g)}$ by $\Phi^{\mathbf{w},g}$, for short.

Lemma 3.1. *Let $\Phi \in k[\mathbf{x}][y] \setminus \{0\}$, $g \in k[\mathbf{x}] \setminus \{0\}$ and $\mathbf{w} \in \Gamma$.*
(i) *The following conditions are equivalent*:
 (1) $m_{\mathbf{w}}^g(\Phi) = 0$.
 (2) $\deg_{\mathbf{w}}^g \Phi = \deg_{\mathbf{w}} \Phi(g)$.
 (3) $\Phi^{\mathbf{w},g}(g^{\mathbf{w}}) \neq 0$.
 (4) $\Phi(g) \neq 0$ *and* $\Phi(g)^{\mathbf{w}} = \Phi^{\mathbf{w},g}(g^{\mathbf{w}})$.
(ii) *It follows that* $m_{\mathbf{w}}^g(\Phi) = \min\{i \in \boldsymbol{Z}_{\geq 0} \mid (\partial_y^i(\Phi^{\mathbf{w},g}))(g^{\mathbf{w}}) \neq 0\}$.

Proof.
(i) The equivalence between (1) and (2) follows from the definition of $m_{\mathbf{w}}^g(\Phi)$. In the following, we will establish that

$$\deg_{\mathbf{w}}(\Phi(g) - \Phi^{\mathbf{w},g}(g^{\mathbf{w}})) < \deg_{\mathbf{w}}^g \Phi. \tag{3.1}$$

Assuming this, we can readily check that (2), (3) and (4) are equivalent, since

$$\Phi(g) = \Phi^{\mathbf{w},g}(g^{\mathbf{w}}) + (\Phi(g) - \Phi^{\mathbf{w},g}(g^{\mathbf{w}})),$$

and $\Phi^{\mathbf{w},g}(g^{\mathbf{w}})$ is contained in $k[\mathbf{x}]_\delta$, where $\delta = \deg_{\mathbf{w}}^g \Phi$.

Write $\Phi = \sum_i \phi_i y^i$ and $\Phi^{\mathbf{w},g} = \sum_i \phi_i' y^i$, where $\phi_i, \phi_i' \in k[\mathbf{x}]$ for each i. Then, $\deg_{\mathbf{w}}(\phi_i g^i) \leq \deg_{\mathbf{w}}^g \Phi$ for each i. Note that $\phi_i' = \phi_i^{\mathbf{w}}$ if $\deg_{\mathbf{w}}(\phi_i g^i) = \deg_{\mathbf{w}}^g \Phi$, and $\phi_i' = 0$ otherwise. We have

$$\phi_i g^i - \phi_i'(g^{\mathbf{w}})^i = \phi_i g^i - \phi_i^{\mathbf{w}}(g^{\mathbf{w}})^i = \phi_i g^i - (\phi_i g^i)^{\mathbf{w}}$$

in the former case, and $\phi_i g^i - \phi_i'(g^{\mathbf{w}})^i = \phi_i g^i$ in the latter case. In each case, $\deg_{\mathbf{w}}^g \Phi$ is greater than the $\mathbf{w}$-degree of $\phi_i g^i - \phi_i'(g^{\mathbf{w}})^i$, and hence greater than that of

$$\sum_i \bigl(\phi_i g^i - \phi_i'(g^{\mathbf{w}})^i\bigr) = \Phi(g) - \Phi^{\mathbf{w},g}(g^{\mathbf{w}}).$$

Thus, we obtain (3.1), thereby proving that (2), (3) and (4) are equivalent.

(ii) Observe that $\bigl(\partial_y^i \Phi\bigr)^{\mathbf{w},g} = \partial_y^i(\Phi^{\mathbf{w},g})$ for each $i \in \boldsymbol{Z}_{\geq 0}$. In view of this, it follows that $\deg_{\mathbf{w}}^g \partial_y^i \Phi = \deg_{\mathbf{w}} \bigl(\partial_y^i \Phi\bigr)(g)$ if and only if $\bigl(\partial_y^i(\Phi^{\mathbf{w},g})\bigr)(g^{\mathbf{w}}) \neq 0$ by the equivalence between (2) and (3) in (i). Then, the assertion immediately follows from the definition of $m_{\mathbf{w}}^g(\Phi)$. □

Now, let A be a k-subalgebra of $k[\mathbf{x}]$, and K the field of fractions of A. We define the initial algebra $A^{\mathbf{w}}$ of A for $\mathbf{w}$ to be the k-subalgebra of $k[\mathbf{x}]$ generated by $f^{\mathbf{w}}$ for $f \in A \setminus \{0\}$. Then, $\Phi^{\mathbf{w},g}$ belongs to $A^{\mathbf{w}}[y] \setminus \{0\}$ for each $\Phi \in A[y] \setminus \{0\}$ for any $g \in k[\mathbf{x}] \setminus \{0\}$. We claim that the field of fractions of $B^{\mathbf{w}}$ is equal to that of $A^{\mathbf{w}}$ whenever B is a k-subalgebra of $k[\mathbf{x}]$ whose field of fractions is equal to K. Indeed, if $fg_1 = g_2$ for $f \in A$ (resp. $f \in B$) and $g_1, g_2 \in B$ (resp. $g_1, g_2 \in A$), then we have $f^{\mathbf{w}} g_1^{\mathbf{w}} = (fg_1)^{\mathbf{w}} = g_2^{\mathbf{w}}$, so $f^{\mathbf{w}}$ belongs to the field of fractions of $B^{\mathbf{w}}$ (resp. $A^{\mathbf{w}}$). For this reason, we denote the field of fractions of $A^{\mathbf{w}}$ by $K^{\mathbf{w}}$.

For an integral domain R and an element s of an integral domain S containing R, we define $I(R,s)$ to be the kernel of the substitution map $R[y] \ni f \mapsto f(s) \in S$. When $I(R,s)$ is a principal ideal of $R[y]$, a generator of $I(R,s)$, which is unique up to multiplication by units in R, is denoted by $P(R,s)$. We remark that $I(R,s)$ is always principal if R is a unique factorization domain. If R is a field and s is algebraic over R, then we may take $P(R,s)$ to be the minimal polynomial of s over R.

Proposition 3.2. *Let A be a k-subalgebra of $k[\mathbf{x}]$, and K the field of fractions of A. Then, for each $\Phi \in A[y] \setminus \{0\}$, $g \in k[\mathbf{x}] \setminus \{0\}$ and $\mathbf{w} \in \Gamma^n$, we have the following*:

(i) *If $g^{\mathbf{w}}$ is transcendental over $K^{\mathbf{w}}$, then $m_{\mathbf{w}}^g(\Phi) = 0$ and $\deg_{\mathbf{w}} \Phi(g) = \deg_{\mathbf{w}}^g \Phi$.*

(ii) *If $g^{\mathbf{w}}$ is algebraic over $K^{\mathbf{w}}$, then $m_{\mathbf{w}}^g(\Phi)$ is at most the quotient of $\deg_y \Phi^{\mathbf{w},g}$ divided by $[K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}]$. If furthermore $I(A^{\mathbf{w}}, g^{\mathbf{w}})$ is a principal ideal, then there exists $H \in A^{\mathbf{w}}[y] \setminus I(A^{\mathbf{w}}, g^{\mathbf{w}})$ such that $\Phi^{\mathbf{w},g} = P(A^{\mathbf{w}}, g^{\mathbf{w}})^m H$, where $m = m_{\mathbf{w}}^g(\Phi)$.*

Proof.

(i) If $g^{\mathbf{w}}$ is transcendental over $K^{\mathbf{w}}$, then $\Phi^{\mathbf{w},g}(g^{\mathbf{w}}) \neq 0$, since $\Phi^{\mathbf{w},g}$ is a nonzero element of $K^{\mathbf{w}}[y]$. Hence, $m^g_{\mathbf{w}}(\Phi) = 0$ and $\deg \Phi(g) = \deg^g \Phi$ by Lemma 3.1(i).

(ii) Set $P_0 = P(K^{\mathbf{w}}, g^{\mathbf{w}})$. By Lemma 3.1(ii), we have $(\partial_y^{m-1}\Phi^{\mathbf{w},g})(g^{\mathbf{w}}) = 0$ and $(\partial_y^{m}\Phi^{\mathbf{w},g})(g^{\mathbf{w}}) \neq 0$. Since k is of characteristic zero, this implies that $\Phi^{\mathbf{w},g} = P_0^m H$ for some $H \in K^{\mathbf{w}}[y]$ with $H(g^{\mathbf{w}}) \neq 0$. By the assumption that $g^{\mathbf{w}}$ is algebraic over $K^{\mathbf{w}}$, it follows that $\deg_y P_0 = [K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}]$. Thus, we get $\deg_y \Phi^{\mathbf{w},g} = m^g_{\mathbf{w}}(\Phi)[K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}] + \deg_y H$. Therefore, $m^g_{\mathbf{w}}(\Phi)$ is at most the quotient of $\deg_y \Phi^{\mathbf{w},g}$ divided by $[K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}]$. Next, assume that $I := I(A^{\mathbf{w}}, g^{\mathbf{w}})$ is a principal ideal. Set $P = P(A^{\mathbf{w}}, g^{\mathbf{w}})$, and write $\Phi^{\mathbf{w},g} = P^{m'}H'$, where $m' \in \boldsymbol{Z}_{\geq 0}$ and $H' \in A^{\mathbf{w}}[y] \setminus I$. Then, m' must be at most m, since P belongs to $P_0K^{\mathbf{w}}[y]$. On the other hand, P does not belong to $P_0^2K^{\mathbf{w}}[y]$, for otherwise $\partial_y P$ would belong to $P_0K^{\mathbf{w}}[y] \cap A^{\mathbf{w}}[y] = I = PA^{\mathbf{w}}[y]$, a contradiction. Hence, m' must be at least m, since $H'(g^{\mathbf{w}}) \neq 0$. Thus, $m' = m$. This proves the latter part. □

The inequality of Shestakov-Umirbaev [**6**, Theorem 3] which was used to solve Nagata's conjecture is generalized as follows.

Theorem 3.3. *Let $f_1, \dots, f_r$ and g be nonzero elements of $k[\mathbf{x}]$ for $r \geq 1$ such that $f_1, \dots, f_r$ algebraically independent over k, and let $A = k[f_1, \dots, f_r]$, $K = k(f_1, \dots, f_r)$ and $\omega = df_1 \wedge \cdots \wedge df_r$. Let $\mathbf{w} \in \Gamma^n$ such that $\deg_{\mathbf{w}} h \geq 0$ for each $h \in A \setminus \{0\}$, and $M = \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega - \deg_{\mathbf{w}} g$. Then, the following statements hold for each $\Phi \in A[y] \setminus \{0\}$:*

(i) *Assume that $g^{\mathbf{w}}$ is algebraic over $K^{\mathbf{w}}$. Let a and b be the quotient and residue of $\deg_y \Phi$ divided by $[K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}]$, respectively. Then, we have*

$$\begin{aligned}\deg_{\mathbf{w}} \Phi(g) &\geq (\deg_y \Phi) \deg_{\mathbf{w}} g + aM \\ &= a\big([K^{\mathbf{w}}(g^{\mathbf{w}}) : K^{\mathbf{w}}] \deg_{\mathbf{w}} g + M\big) + b \deg_{\mathbf{w}} g.\end{aligned}$$

(ii) *If $I(A^{\mathbf{w}}, g^{\mathbf{w}})$ is a principal ideal and $\deg_{\mathbf{w}} g \geq 0$, then*

$$\deg_{\mathbf{w}} \Phi(g) \geq m^g_{\mathbf{w}}(\Phi)\big(\deg^g_{\mathbf{w}} P(A^{\mathbf{w}}, g^{\mathbf{w}}) + M\big). \tag{3.2}$$

Proof.

(i) The last equality can be checked easily. We only show the first inequality. By Theorem 2.1, we get $\deg_{\mathbf{w}} \Phi(g) \geq \deg^g_{\mathbf{w}} \Phi + m^g_{\mathbf{w}}(\Phi)M$. So, it suffices to verify that $\deg^g_{\mathbf{w}} \Phi \geq (\deg_y \Phi) \deg_{\mathbf{w}} g$ and $m^g_{\mathbf{w}}(\Phi)M \geq aM$. Let $\phi_e \in A$ be the coefficient of y^e in Φ, where $e = \deg_y \Phi$. Then, $\deg^g_{\mathbf{w}} \Phi \geq \deg_{\mathbf{w}}(\phi_e g^e)$. By the assumption on $\mathbf{w}$, it follows that $\deg_{\mathbf{w}} \phi_e \geq 0$. Hence,

$$\deg^g_{\mathbf{w}} \Phi \geq \deg_{\mathbf{w}}(\phi_e g^e) = \deg_{\mathbf{w}} \phi_e + e \deg_{\mathbf{w}} g \geq (\deg_y \Phi) \deg_{\mathbf{w}} g.$$

On the other hand, we get $M \leq 0$ by (2.4) and (2.5), and $m^g_{\mathbf{w}}(\Phi) \leq a$ by Proposition 3.2(ii). Therefore, $m^g_{\mathbf{w}}(\Phi)M \geq aM$, proving the first inequality.

(ii) Note that $\deg^g_{\mathbf{w}} \Psi \geq 0$ if Ψ is a nonzero element of $A[y] \cup A^{\mathbf{w}}[y]$. Actually, we may write $\deg^g_{\mathbf{w}} \Psi = \deg_{\mathbf{w}} \psi + l \deg_{\mathbf{w}} g$, where $\psi \in A \setminus \{0\}$ and $l \in \boldsymbol{Z}_{\geq 0}$, and it follows that $\deg_{\mathbf{w}} \psi \geq 0$ and $\deg_{\mathbf{w}} g \geq 0$ by assumption. First, assume that $g^{\mathbf{w}}$ is transcendental over $K^{\mathbf{w}}$. Then, $m^g_{\mathbf{w}}(\Phi) = 0$ and $\deg_{\mathbf{w}} \Phi(g) = \deg^g_{\mathbf{w}} \Phi$ by Proposition 3.2(i). Hence, the right-hand side of (3.2) is zero, and $\deg_{\mathbf{w}} \Phi(g) = \deg^g_{\mathbf{w}} \Phi \geq 0$ as noted. Therefore, (3.2) is true if $g^{\mathbf{w}}$ is transcendental over $K^{\mathbf{w}}$. Next, assume that $g^{\mathbf{w}}$ is algebraic over $K^{\mathbf{w}}$. By Proposition 3.2(ii), we have $\Phi^{\mathbf{w},g} = P^m H$ for some $H \in A^{\mathbf{w}}[y]$, where $P = P(A^{\mathbf{w}}, g^{\mathbf{w}})$ and $m = m^g_{\mathbf{w}}(\Phi)$. Since $\deg^g_{\mathbf{w}} H \geq 0$ as noted, we obtain

$$\deg^g_{\mathbf{w}} \Phi = \deg^g_{\mathbf{w}} \Phi^{\mathbf{w},g} = m \deg^g_{\mathbf{w}} P + \deg^g_{\mathbf{w}} H \geq m^g_{\mathbf{w}}(\Phi) \deg^g_{\mathbf{w}} P.$$

With the aid of this inequality, (3.2) follows from Theorem 2.1. □

The following lemma is well-known. For the sake of completeness, we include a proof at the end of this section.

LEMMA 3.4. *Assume that $f, g \in k[\mathbf{x}]$ are $\mathbf{w}$-homogeneous for some $\mathbf{w} \in \Gamma^n$. If $\deg_{\mathbf{w}} f > 0$, $\deg_{\mathbf{w}} g > 0$, and f, g are algebraically dependent over k, then there exist mutually prime natural numbers $l(f,g)$ and $l(g,f)$ as follows*:

(i) $g^{l(f,g)} = \alpha f^{l(g,f)}$ *for some* $\alpha \in k \setminus \{0\}$.
(ii) $I(k[f], g) = \big(y^{l(f,g)} - \alpha f^{l(g,f)}\big) k[f][y]$.
(iii) $[k(f)(g) : k(f)] = l(f,g)$.
(iv) $l(f,g) = (\deg_{\mathbf{w}} f) \gcd(\deg_{\mathbf{w}} f, \deg_{\mathbf{w}} g)^{-1}$ *in case* $\Gamma = \boldsymbol{Z}$.

As a corollary to Theorem 3.3, we obtain the inequality of Shestakov-Umirbaev [**6**, Theorem 3] as follows.

COROLLARY 3.5 (Shestakov-Umirbaev). *Assume that $\Gamma = \boldsymbol{Z}$, and $f, g \in k[\mathbf{x}] \setminus k$ satisfy $\deg_{\mathbf{w}} f > 0$ and $\deg_{\mathbf{w}} g > 0$ for some $\mathbf{w} \in \Gamma^n$. Then, for each $\Phi \in k[f][y] \setminus \{0\}$, it follows that*

$$\deg_{\mathbf{w}} \Phi(g) \geq a(\operatorname{lcm}(\deg_{\mathbf{w}} f, \deg_{\mathbf{w}} g) + M) + b \deg_{\mathbf{w}} g. \tag{3.3}$$

Here, $M = \deg_{\mathbf{w}}(df \wedge dg) - \deg_{\mathbf{w}} f - \deg_{\mathbf{w}} g$*, and* a *and* b *are the quotient and residue of* $\deg_y \Phi$ *divided by* $(\deg_{\mathbf{w}} f)\gcd(\deg_{\mathbf{w}} f, \deg_{\mathbf{w}} g)^{-1}$*, respectively.*

Proof. We remark that $k[f]^{\mathbf{w}} = k[f^{\mathbf{w}}]$, and $\deg_{\mathbf{w}} h \geq 0$ holds for each $h \in k[f] \setminus \{0\}$. In fact, if $h = \sum_{i=0}^{e} c_i f^i$, where $e \in \boldsymbol{Z}_{\geq 0}$ and $c_0, \dots, c_e \in k$ with $c_e \neq 0$, then $\deg_{\mathbf{w}} h = e \deg_{\mathbf{w}} f \geq 0$ and $h^{\mathbf{w}} = c_e(f^{\mathbf{w}})^e$, since $\deg_{\mathbf{w}} f > 0$ by assumption. Consequently, we have $k(f)^{\mathbf{w}} = k(f^{\mathbf{w}})$. First, assume that $f^{\mathbf{w}}$ and $g^{\mathbf{w}}$ are algebraically dependent over k. Put $N = [k(f^{\mathbf{w}})(g^{\mathbf{w}}) : k(f^{\mathbf{w}})]$. Then, Theorem 3.3(i) gives that

$$\deg_{\mathbf{w}} \Phi(g) \geq a'(N \deg_{\mathbf{w}} g + M) + b' \deg_{\mathbf{w}} g, \tag{3.4}$$

where a' and b' are the quotient and residue of $\deg_y \Phi$ divided by N, respectively. By Lemma 3.4, we have

$$N = \frac{\deg_{\mathbf{w}} f^{\mathbf{w}}}{\gcd(\deg_{\mathbf{w}} f^{\mathbf{w}}, \deg_{\mathbf{w}} g^{\mathbf{w}})} = \frac{\deg_{\mathbf{w}} f}{\gcd(\deg_{\mathbf{w}} f, \deg_{\mathbf{w}} g)} = \frac{\operatorname{lcm}(\deg_{\mathbf{w}} f, \deg_{\mathbf{w}} g)}{\deg_{\mathbf{w}} g}.$$

This implies that the right-hand side of (3.4) is equal to that of (3.3), thus proving (3.3). If $f^{\mathbf{w}}$ and $g^{\mathbf{w}}$ are algebraically independent over k, then $\deg_{\mathbf{w}} \Phi(g) = \deg_{\mathbf{w}}^g \Phi$ by Proposition 3.2(i). As in the proof of Theorem 3.3, we get $\deg_{\mathbf{w}}^g \Phi \geq (\deg_y \Phi) \deg_{\mathbf{w}} g$. The right-hand side of (3.3) is equal to $(\deg_y \Phi) \deg_{\mathbf{w}} g + aM$, in which $M \leq 0$. Therefore, (3.3) is true. □

In the original statement of [**6**, Theorem 3], the "Poisson bracket" $[f, g]$ is used instead of $df \wedge dg$. The degrees of $[f, g]$ and $df \wedge dg$ are defined in the same way.

To conclude this section, we prove Lemma 3.4. The assertions (ii), (iii) and (iv) easily follow from (i). We only show that $f^l g^{-m}$ belongs to k for some $l, m \in \boldsymbol{N}$ with $\gcd(l, m) = 1$. Without loss of generality, we may assume that k is algebraically closed. Indeed, $f^l g^{-m}$ necessarily belongs to k if $f^l g^{-m}$ is algebraic over k, since the field of fractions of $k[\mathbf{x}]$ is a regular extension of k.

By the assumption that f and g are algebraically dependent over k, we may find a nontrivial algebraic relation $\sum_{i,j} \beta_{i,j} f^i g^j = 0$, where $\beta_{i,j} \in k$ for each $i, j \in \boldsymbol{Z}_{\geq 0}$. Let J be the set of $(i, j) \in (\boldsymbol{Z}_{\geq 0})^2$ such that $\beta_{i,j} \neq 0$. Take $(i_0, j_0), (i_1, j_1) \in J$ such that i_0 and i_1 are the maximum and minimum of $\{i \mid (i, j) \in J\}$, respectively. Since f and g are $\mathbf{w}$-homogeneous, we may assume that $i \deg_{\mathbf{w}} f + j \deg_{\mathbf{w}} g$ are the same for any $(i, j) \in J$. Hence, $(i_1 - i_0) \deg_{\mathbf{w}} g = (j_0 - j_1) \deg_{\mathbf{w}} f$. Then, we have $i_1 - i_0 > 0$, for otherwise $J = \{(i_0, j_0)\}$, and so $0 = \sum_{(i,j) \in J} \beta_{i,j} f^i g^j = \beta_{i_0, j_0} f^{i_0} g^{j_0} \neq 0$, a contradiction.

Since $\deg_{\mathbf{w}} f > 0$ and $\deg_{\mathbf{w}} g > 0$ by assumption, we have $j_0 - j_1 > 0$. Set $l' = i_1 - i_0$, $m' = j_0 - j_1$ and $l = l'/e$, $m = m'/e$, where $e = \gcd(l', m')$. Then, J is contained in $\{(i_0, j_0) + p(l, -m) \mid p = 0, \ldots, e\}$. By putting $\beta'_p = \beta_{i_0+lp, j_0-mp}$ for each p, we get

$$0 = \sum_{(i,j)\in J} \beta_{i,j} f^i g^j = f^{i_0} g^{j_0} \sum_{p=0}^{e} \beta'_p (f^l g^{-m})^p = \beta'_e f^{i_0} g^{j_0} \prod_{p=1}^{e} (f^l g^{-m} - \alpha_p),$$

where $\alpha_1, \ldots, \alpha_e \in k$ are the solutions of the equation $\sum_{p=0}^{e} \beta'_p y^p = 0$. Thus, $f^l g^{-m} = \alpha_p$ for some p. Therefore, $f^l g^{-m}$ is contained in k. This completes the proof of Lemma 3.4.

4. Polynomial automorphisms.

As an application of our result, we study features of elements of $\mathrm{Aut}_k k[\mathbf{x}]$. We give a necessary condition for n-tuples $\mathbf{f} = (f_1, \ldots, f_n)$ of elements of $k[\mathbf{x}]$ to satisfy $k[f_1, \ldots, f_n] = k[\mathbf{x}]$.

First, we recall an elementary fact about initial algebras.

Lemma 4.1. *If $g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}$ are algebraically independent over k for $\mathbf{w} \in \Gamma$ and $g_1, \ldots, g_r \in k[\mathbf{x}]$, where $r \in \boldsymbol{Z}_{\geq 0}$, then $k[g_1, \ldots, g_r]^{\mathbf{w}} = k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$.*

Proof. Clearly, $k[g_1, \ldots, g_r]^{\mathbf{w}}$ contains $k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$. We show the reverse inclusion by induction on r. The assertion is obvious if $r = 0$. Assume that $r \geq 1$. It suffices to verify that $h^{\mathbf{w}}$ belongs to $k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$ for each $h \in k[g_1, \ldots, g_r] \setminus \{0\}$. Take $H \in A[y]$ such that $h = H(g_r)$, where $A = k[g_1, \ldots, g_{r-1}]$. By induction assumption, we have $A^{\mathbf{w}} = k[g_1^{\mathbf{w}}, \ldots, g_{r-1}^{\mathbf{w}}]$. Besides, $H^{\mathbf{w}, g_r}$ belongs to $A^{\mathbf{w}}[y] \setminus \{0\}$. Hence, $H^{\mathbf{w}, g_r}(g_r^{\mathbf{w}})$ is contained in $k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$. Moreover, $H^{\mathbf{w}, g_r}(g_r^{\mathbf{w}})$ is not zero, since $g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}$ are algebraically independent over k by assumption. Hence, $H(g_r)^{\mathbf{w}} = H^{\mathbf{w}, g_r}(g_r^{\mathbf{w}})$ by Lemma 3.1(i). Since $h = H(g_r)$, we get $h^{\mathbf{w}} = H(g_r)^{\mathbf{w}}$. Thus, $h^{\mathbf{w}}$ belongs to $k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$. Therefore, $k[g_1, \ldots, g_r]^{\mathbf{w}}$ is contained in $k[g_1^{\mathbf{w}}, \ldots, g_r^{\mathbf{w}}]$. □

The following proposition is an immediate consequence of Lemma 4.1.

Proposition 4.2. *Assume that $f_1, \ldots, f_n \in k[\mathbf{x}]$ satisfy $k[f_1, \ldots, f_n] = k[\mathbf{x}]$. Then, for each $\mathbf{w} \in \Gamma^n$, it follows that $f_1^{\mathbf{w}}, \ldots, f_n^{\mathbf{w}}$ are algebraically independent over k if and only if $k[f_1^{\mathbf{w}}, \ldots, f_n^{\mathbf{w}}] = k[\mathbf{x}]$.*

Proof. The "if" part is clear, for $k[\mathbf{x}]$ has transcendence degree n over k. Assume that $f_1^{\mathbf{w}}, \ldots, f_n^{\mathbf{w}}$ are algebraically independent over k. Then,

$k[f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}}] = k[f_1,\dots,f_n]^{\mathbf{w}}$ by Lemma 4.1. Since $k[f_1,\dots,f_n] = k[\mathbf{x}]$, we have $k[f_1,\dots,f_n]^{\mathbf{w}} = k[\mathbf{x}]^{\mathbf{w}} = k[\mathbf{x}]$. Thus, $k[f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}}] = k[\mathbf{x}]$, proving the "only if" part. □

Next, we consider the case where $k(f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}})$ has transcendence degree $n-1$ over k for some $\mathbf{w} \in \Gamma^n$. We define an element $\Delta_{\mathbf{f}}^{\mathbf{w}}$ of Γ as follows: Let $\lambda_{\mathbf{f}}^{\mathbf{w}} : k[\mathbf{x}] \to k[\mathbf{x}]$ be the homomorphism defined by $\lambda_{\mathbf{f}}^{\mathbf{w}}(x_i) = f_i^{\mathbf{w}}$ for $i = 1,\dots,n$. Then, $\ker \lambda_{\mathbf{f}}^{\mathbf{w}}$ is a prime ideal of $k[\mathbf{x}]$ of hight one. Since $k[\mathbf{x}]$ is a unique factorization domain, there exists $Q \in k[\mathbf{x}] \setminus \{0\}$ such that $\ker \lambda_{\mathbf{f}}^{\mathbf{w}} = Qk[\mathbf{x}]$. We define $\Delta_{\mathbf{f}}^{\mathbf{w}}$ to be the $\mathbf{w}_{\mathbf{f}}$-degree of Q, where

$$\mathbf{w}_{\mathbf{f}} = (\deg_{\mathbf{w}} f_1,\dots,\deg_{\mathbf{w}} f_n).$$

Note that $\Delta_{\mathbf{f}}^{\mathbf{w}}$ is uniquely determined by $\mathbf{f}$ and $\mathbf{w}$, since Q is unique up to multiplication by elements in $k \setminus \{0\}$.

The following theorem is an extension of Proposition 1.1.

THEOREM 4.3. *Let $f_1,\dots,f_n$ be elements of $k[\mathbf{x}]$ such that $k[f_1,\dots,f_n] = k[\mathbf{x}]$, and $\mathbf{w} = (w_1,\dots,w_n)$ an element of $(\Gamma_{\geq 0})^n$. If $k(f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}})$ has transcendence degree $n-1$ over k, then*

$$\sum_{i=1}^{n} \deg_{\mathbf{w}} f_i \geq \Delta_{\mathbf{f}}^{\mathbf{w}} + \sum_{i=1}^{n} w_i - \max\{w_i \mid i = 1,\dots,n\}, \tag{4.1}$$

where $\mathbf{f} = (f_1,\dots,f_n)$.

PROOF. Since $k(f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}})$ has transcendence degree $n-1$ over k, we may find l such that x_l is not contained in $k[f_1^{\mathbf{w}},\dots,f_n^{\mathbf{w}}]$. Moreover, we may assume that $f_1^{\mathbf{w}},\dots,f_{n-1}^{\mathbf{w}}$ are algebraically independent over k by changing the indices of $f_1,\dots,f_n$ if necessary. Set $A = k[f_1,\dots,f_{n-1}]$ and $g = f_n$. Then, there exists $\Phi \in A[y]$ such that $\Phi(g) = x_l$, since $A[g] = k[\mathbf{x}]$ by assumption. Furthermore, $A^{\mathbf{w}} = k[f_1^{\mathbf{w}},\dots,f_{n-1}^{\mathbf{w}}]$ by Lemma 4.1, and so $A^{\mathbf{w}}$ is a polynomial ring over k. Accordingly, $I(A^{\mathbf{w}}, g^{\mathbf{w}})$ is a principal ideal of $A^{\mathbf{w}}[y]$. Since $w_i \geq 0$ for $i = 1,\dots,n$ by assumption, $\deg_{\mathbf{w}} h \geq 0$ holds for each $h \in k[\mathbf{x}] \setminus \{0\}$. Then, we can easily check that $f_1,\dots,f_{n-1}$, g and $\mathbf{w}$ satisfy the assumptions of Theorem 3.3(ii). Therefore, we obtain

$$\deg_{\mathbf{w}} \Phi(g) \geq m_{\mathbf{w}}^{g}(\Phi)\bigl(\deg_{\mathbf{w}}^{g} P + M\bigr), \tag{4.2}$$

where $P = P(A^{\mathbf{w}}, g^{\mathbf{w}})$, $M = \deg_{\mathbf{w}}(\omega \wedge dg) - \deg_{\mathbf{w}} \omega - \deg_{\mathbf{w}} g$ and $\omega = df_1 \wedge \cdots \wedge$

df_{n-1}. We show that

$$M \geq \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} \deg_{\mathbf{w}} f_i. \tag{4.3}$$

Let α be the determinant of the n by n matrix $(\partial f_i/\partial x_j)_{i,j}$. Then, $\omega \wedge dg = df_1 \wedge \cdots \wedge df_n = \alpha dx_1 \wedge \cdots \wedge dx_n$. Since $k[f_1, \ldots, f_n] = k[\mathbf{x}]$ by assumption, α belongs to $k \setminus \{0\}$. Hence, we have

$$\deg_{\mathbf{w}}(\omega \wedge dg) = \deg_{\mathbf{w}}(\alpha dx_1 \wedge \cdots \wedge dx_n) = \deg_{\mathbf{w}} \alpha + \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} w_i. \tag{4.4}$$

In view of (2.4) and (2.5), it follows that

$$\deg_{\mathbf{w}} \omega = \deg_{\mathbf{w}}(df_1 \wedge \cdots \wedge df_{n-1}) \leq \sum_{i=1}^{n-1} \deg_{\mathbf{w}} df_i = \sum_{i=1}^{n-1} \deg_{\mathbf{w}} f_i. \tag{4.5}$$

Since $g = f_n$, the inequality (4.3) follows from (4.4) and (4.5).

To complete the proof, it remains only to show that $m_{\mathbf{w}}^g(\Phi) \geq 1$ and $\deg_{\mathbf{w}}^g P = \Delta_{\mathbf{f}}^{\mathbf{w}}$. Actually, (4.2) implies $\deg_{\mathbf{w}} \Phi(g) \geq \deg_{\mathbf{w}}^g P + M$ whether or not $\deg_{\mathbf{w}}^g P + M \geq 0$, since $\deg_{\mathbf{w}} \Phi(g) \geq 0$. So, (4.1) follows from (4.3) and

$$\max\{w_i \mid i = 1, \ldots, n\} \geq w_l = \deg_{\mathbf{w}} x_l = \deg_{\mathbf{w}} \Phi(g).$$

First, suppose to the contrary that $m_{\mathbf{w}}^g(\Phi) = 0$. Then, $\Phi^{\mathbf{w},g}(g^{\mathbf{w}}) = \Phi(g)^{\mathbf{w}} = x_l^{\mathbf{w}} = x_l$ by Lemma 3.1. Recall that x_l does not belong to $k[f_1^{\mathbf{w}}, \ldots, f_n^{\mathbf{w}}]$, and $k[f_1^{\mathbf{w}}, \ldots, f_n^{\mathbf{w}}] = A^{\mathbf{w}}[g^{\mathbf{w}}]$. Since $\Phi^{\mathbf{w},g}$ is in $A^{\mathbf{w}}[y]$, it follows that $\Phi^{\mathbf{w},g}(g^{\mathbf{w}})$ belongs to $A^{\mathbf{w}}[g^{\mathbf{w}}]$. This is a contradiction. Thus, we get $m_{\mathbf{w}}^g(\Phi) \geq 1$. Next, take $Q \in k[\mathbf{x}]$ so that $\ker \lambda_{\mathbf{f}}^{\mathbf{w}} = Qk[\mathbf{x}]$. Let $\iota : k[\mathbf{x}] \to A^{\mathbf{w}}[y]$ be the homomorphism defined by $\iota(x_i) = f_i^{\mathbf{w}}$ for $i = 1, \ldots, n-1$ and $\iota(x_n) = y$. Then, ι is an isomorphism, since we are assuming that $f_1^{\mathbf{w}}, \ldots, f_{n-1}^{\mathbf{w}}$ are algebraically independent over k. This assumption implies further that the $\mathbf{w_f}$-degree of Q is equal to the $(\mathbf{w}, \deg_{\mathbf{w}} g)$-degree of $\iota(Q)$. It is equal to $\deg_{\mathbf{w}}^g \iota(Q)$ as mentioned at the beginning of Section 3. Thus, we get $\Delta_{\mathbf{f}}^{\mathbf{w}} = \deg_{\mathbf{w}}^g \iota(Q)$. By definition, $\lambda_{\mathbf{f}}^{\mathbf{w}}$ is equal to the composite of ι and the substitution map $A^{\mathbf{w}}[y] \ni \psi \mapsto \psi(g^{\mathbf{w}}) \in k[\mathbf{x}]$. Hence, we have

$$\iota(Qk[\mathbf{x}]) = \iota(\ker \lambda_{\mathbf{f}}^{\mathbf{w}}) = I(A^{\mathbf{w}}, g^{\mathbf{w}}) = PA^{\mathbf{w}}[y].$$

Since ι is an isomorphism, we get $\iota(Q) = \alpha P$ for some $\alpha \in k \setminus \{0\}$. Thus, $\deg^g_{\mathbf{w}} \iota(Q) = \deg^g_{\mathbf{w}} P$. Therefore, we obtain $\Delta^{\mathbf{w}}_{\mathbf{f}} = \deg^g_{\mathbf{w}} P$. □

In case of $n = 2$, we have the following corollary to Theorem 4.3, which implies Proposition 1.1.

COROLLARY 4.4. *Assume that* $f_1, f_2 \in k[x_1, x_2]$ *satisfy* $k[f_1, f_2] = k[x_1, x_2]$. *If* $f_1^{\mathbf{w}}$ *and* $f_2^{\mathbf{w}}$ *are algebraically dependent over* k *for* $\mathbf{w} \in (\boldsymbol{Z}_{\geq 0})^2$, *then* $\deg_{\mathbf{w}} f_1$ *and* $\deg_{\mathbf{w}} f_2$ *are positive integers which satisfy*

$$\deg_{\mathbf{w}} f_1 + \deg_{\mathbf{w}} f_2 \geq \operatorname{lcm}(\deg_{\mathbf{w}} f_1, \deg_{\mathbf{w}} f_2) + \min\{w_1, w_2\}, \tag{4.6}$$

where $\mathbf{w} = (w_1, w_2)$. *In particular,* $\deg_{\mathbf{w}} f_1 | \deg_{\mathbf{w}} f_2$ *or* $\deg_{\mathbf{w}} f_2 | \deg_{\mathbf{w}} f_1$.

PROOF. Since $w_i \geq 0$ for $i = 1, 2$ by assumption, $\deg_{\mathbf{w}} f_i \geq 0$ for $i = 1, 2$. We show that $\deg_{\mathbf{w}} f_i \neq 0$ for $i = 1, 2$ by contradiction. Suppose the contrary, say $\deg_{\mathbf{w}} f_1 = 0$. Then, $w_i = 0$ for some $i \in \{1, 2\}$, since f_1 does not belong to k. We claim that $\mathbf{w} \neq 0$, for otherwise $f_i^{\mathbf{w}} = f_i$ for $i = 1, 2$, which is impossible because $k[f_1, f_2] = k[x_1, x_2]$ and $f_1^{\mathbf{w}}$, $f_2^{\mathbf{w}}$ are algebraically dependent over k by assumption. Hence, we have $w_j > 0$ for $j \in \{1, 2\} \setminus \{i\}$. Since we suppose that $\deg_{\mathbf{w}} f_1 = 0$, this implies that f_1 belongs to $k[x_i]$, and $f_1^{\mathbf{w}} = f_1$. Then, $f_2^{\mathbf{w}}$ also belongs to $k[x_i]$, since $f_1^{\mathbf{w}}$ and $f_2^{\mathbf{w}}$ are algebraically dependent over k. Consequently, f_2 belongs to $k[x_i]$ because $w_i = 0$ and $w_j > 0$. Thus, $k[f_1, f_2]$ is contained in $k[x_i]$, a contradiction. Therefore, $\deg_{\mathbf{w}} f_i \neq 0$ for $i = 1, 2$.

Put $P = P(k[f_1^{\mathbf{w}}], f_2^{\mathbf{w}})$ and $\mathbf{f} = (f_1, f_2)$. Then, we have $\Delta^{\mathbf{w}}_{\mathbf{f}} = \deg^{f_2}_{\mathbf{w}} P$ as in the proof of Theorem 4.3. By Lemma 3.4, $P = \beta\big(y^{l(f_1,f_2)} - \alpha(f_1^{\mathbf{w}})^{l(f_2,f_1)}\big)$, where $\alpha, \beta \in k \setminus \{0\}$. Then, we have $\deg^{f_2}_{\mathbf{w}} P = \operatorname{lcm}(\deg_{\mathbf{w}} f_1, \deg_{\mathbf{w}} f_2)$. Thus, $\Delta^{\mathbf{w}}_{\mathbf{f}} = \operatorname{lcm}(\deg_{\mathbf{w}} f_1, \deg_{\mathbf{w}} f_2)$. By Theorem 4.3, we obtain

$$\begin{aligned} \deg_{\mathbf{w}} f_1 + \deg_{\mathbf{w}} f_2 &\geq \Delta^{\mathbf{w}}_{\mathbf{f}} + w_1 + w_2 - \max\{w_1, w_2\} \\ &= \operatorname{lcm}(\deg_{\mathbf{w}} f_1, \deg_{\mathbf{w}} f_2) + \min\{w_1, w_2\}. \end{aligned}$$

The last statement is a consequence of the first statement, since $a + b \geq \operatorname{lcm}(a, b)$ implies $a|b$ or $b|a$ for each $a, b \in \boldsymbol{N}$. □

5. A lemma of Shestakov-Umirbaev.

In this section, we generalize a lemma of Shestakov-Umirbaev [**6**, Lemma 5], which was also used in the proof of Nagata's conjecture.

Set $m_i = \deg_{\mathbf{w}} f_i + \deg_{\mathbf{w}}(df_j \wedge df_l)$ for $f_1, f_2, f_3 \in k[\mathbf{x}] \setminus k$ and $i \in \{1, 2, 3\}$, where $\Gamma = \boldsymbol{Z}$, $\mathbf{w} = (1, \ldots, 1)$, and $j, l \in \boldsymbol{N} \setminus \{i\}$ with $1 \leq j < l \leq 3$.

LEMMA 5.1 (Shestakov-Umirbaev). *In the notation above, it follows that* $m_1 \leq \max\{m_2, m_3\}$. *If* $m_2 \neq m_3$, *then* $m_1 = \max\{m_2, m_3\}$.

We note that the statement of Lemma 5.1 is equivalent to the statement that there exist $1 \leq i_1 < i_2 \leq 3$ such that $m_{i_1} = m_{i_2} \geq m_i$ for $i = 1, 2, 3$. Therefore, the following theorem is a generalization of Lemma 5.1.

THEOREM 5.2. *Let* $\eta_1, \ldots, \eta_l$ *be elements of* $\Omega_{k[\mathbf{x}]/k}$, *where* $l \geq 2$. *Then, there exist* $1 \leq i_1 < i_2 \leq l$ *such that*

$$\deg_{\mathbf{w}} \eta_{i_1} + \deg_{\mathbf{w}} \tilde{\eta}_{i_1} = \deg_{\mathbf{w}} \eta_{i_2} + \deg_{\mathbf{w}} \tilde{\eta}_{i_2} \geq \deg_{\mathbf{w}} \eta_i + \deg_{\mathbf{w}} \tilde{\eta}_i$$

for $i = 1, \ldots, l$, *where* $\tilde{\eta}_i = \eta_1 \wedge \cdots \wedge \eta_{i-1} \wedge \eta_{i+1} \wedge \cdots \wedge \eta_l$ *for each* i.

Lemma 5.1 is obtained as a special case of Theorem 5.2 where $l = 3$ and $\eta_i = df_i$ for $i = 1, 2, 3$. Indeed, for each distinct i, j and l, it follows that

$$m_i = \deg_{\mathbf{w}} f_i + \deg_{\mathbf{w}}(df_j \wedge df_l) = \deg_{\mathbf{w}} df_i + \deg_{\mathbf{w}}(df_j \wedge df_l) = \deg_{\mathbf{w}} \eta_i + \deg_{\mathbf{w}} \tilde{\eta}_i.$$

We prove Theorem 5.2 by contradiction. Suppose to the contrary that there exists i_0 such that $\deg_{\mathbf{w}} \eta_{i_0} + \deg_{\mathbf{w}} \tilde{\eta}_{i_0} > \deg_{\mathbf{w}} \eta_i + \deg_{\mathbf{w}} \tilde{\eta}_i$ for each $i \neq i_0$. Write $\eta_i = \sum_{j=1}^{n} f_{i,j} x_j^{-1} dx_j$ for each i, where $f_{i,j} \in x_j k[\mathbf{x}]$ for each j. Set $d\mathbf{x}_I = dx_{i_1} \wedge \cdots \wedge dx_{i_{l-1}}$ and $\mathbf{x}_I = x_{i_1} \cdots x_{i_{l-1}}$ for each $i_1, \ldots, i_{l-1}$, where $I = (i_1, \ldots, i_{l-1})$. Then, we may write $\tilde{\eta}_i = \sum_J \tilde{f}_{i,J} (\mathbf{x}_J)^{-1} d\mathbf{x}_J$, where the sum is taken over $J = (j_1, \ldots, j_{l-1})$ with $1 \leq j_1 < \cdots < j_{l-1} \leq n$, and $\tilde{f}_{i,J} \in \mathbf{x}_J k[\mathbf{x}]$ for each J. By definition, there exist j_0 and J_0 such that $\deg_{\mathbf{w}} \eta_{i_0} = \deg_{\mathbf{w}} f_{i_0,j_0}$ and $\deg_{\mathbf{w}} \tilde{\eta}_{i_0} = \deg_{\mathbf{w}} \tilde{f}_{i_0,J_0}$. Then, $\deg_{\mathbf{w}}(f_{i,j} \tilde{f}_{i,J}) < \deg_{\mathbf{w}}(f_{i_0,j_0} \tilde{f}_{i_0,J_0})$ for each j and J if $i \neq i_0$ by the choice of i_0. In particular, $f_{i_0,j_0} \neq 0$ and $\tilde{f}_{i_0,J_0} \neq 0$. By changing the indices of $\eta_1, \ldots, \eta_l$ and $x_1, \ldots, x_n$ if necessary, we may assume that $i_0 \neq 1$ and $J_0 = (1, \ldots, l-1)$. Note that the (i, l)-cofactor of the l by l matrix

$$M = \begin{pmatrix} f_{1,1} \cdots f_{1,l-1} \; f_{1,j_0} \\ f_{2,1} \cdots f_{2,l-1} \; f_{2,j_0} \\ \cdots\cdots\cdots \\ f_{l,1} \cdots f_{l,l-1} \; f_{l,j_0} \end{pmatrix}$$

is equal to $(-1)^{i+l} \tilde{f}_{i,J_0}$ for $i = 1, \ldots, l$. Hence, $\det M = \sum_{i=1}^{l} (-1)^{i+l} f_{i,j_0} \tilde{f}_{i,J_0}$.

Since $\deg_{\mathbf{w}}(f_{i,j_0}\tilde{f}_{i,J_0}) < \deg_{\mathbf{w}}(f_{i_0,j_0}\tilde{f}_{i_0,J_0})$ if $i \neq i_0$, we get $\deg_{\mathbf{w}}(\det M) = \deg_{\mathbf{w}}(f_{i_0,j_0}\tilde{f}_{i_0,J_0})$. On the other hand, the $(1,u)$-cofactor of M is equal to $(-1)^{u+1}\tilde{f}_{1,J_u}$ for $u = 1,\ldots,l$, where $J_u = (1,\ldots,u-1,u+1,\ldots,l-1,j_0)$ for $1 \leq u < l$ and $J_l = J_0$. Hence, $\det M = \sum_{u=1}^{l}(-1)^{u+1}f_{1,u}\tilde{f}_{1,J_u}$. Since we assume that $i_0 \neq 1$, it follows that $\deg_{\mathbf{w}}(f_{1,u}\tilde{f}_{1,J_u}) < \deg_{\mathbf{w}}(f_{i_0,j_0}\tilde{f}_{i_0,J_0})$ for each u. Thus, $\deg_{\mathbf{w}}(\det M) < \deg_{\mathbf{w}}(f_{i_0,j_0}\tilde{f}_{i_0,J_0})$, and we are led to a contradiction. Therefore, Theorem 5.2 is true.

References

[1] A. van den Essen, L. Makar-Limanov and R. Willems, Remarks on Shestakov-Umirbaev, Report 0414, Radboud University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands, 2004.

[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), **1** (1953), 33–41.

[4] M. Nagata, On Automorphism Group of $k[x,y]$, Lectures in Mathematics, Department of Mathematics, Kyoto University, **5**, Kinokuniya Book-Store Co. Ltd., Tokyo, 1972.

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

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

[7] I. Shestakov and U. Umirbaev, The tame and the wild automorphisms of polynomial rings in three variables, J. Amer. Math. Soc., **17** (2004), 197–227.

[8] S. Vénéreau, A parachute for the degree of a polynomial in algebraically independent ones, arXiv:math.AC/07041561.

Shigeru Kuroda
Department of Mathematics and Information Sciences
Tokyo Metropolitan University
1-1 Minami-Ohsawa, Hachioji
Tokyo 192-0397, Japan
E-mail: kuroda@tmu.ac.jp