AN ELEMENTARY PROOF OF THE TAMENESS
OF POLYNOMIAL
AUTOMORPHISMS OF $k^2$

BY LUDWIK M. DRUŻKOWSKI
AND JERZY GURYCZ

Abstract. A self-contained and detailed proof of the Jung and van der Kulk
theorem on the tameness of polynomial automorphisms in two variables is
given. Following Rentschler (cf. [Ren]) this theorem is obtained as an applica-
tion of investigations of locally nilpotent derivations.

1. Introduction. In this paper $k$ denotes a field of characteristic zero,
$k^* = k \setminus \{0\}$, $\mathbb{N}$ is the set of nonnegative integers.

Definition 1.1. (i) If $F = (F_1, \ldots, F_n) : k^n \to k^n$ is a polynomial map-
ping, then $F$ determines an algebra homomorphism $F^* : k[X_1, \ldots, X_n] \to
k[X_1, \ldots, X_n]$ given by the formula

$$F^* : k[X_1, \ldots, X_n] \ni g \to g \circ F \in k[X_1, \ldots, X_n].$$

(ii) If $\Phi : k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n]$ is a homomorphism, then $\Phi$
determines a polynomial mapping $\Phi_* = (F_1, \ldots, F_n) : k^n \to k^n$ given by the formula

$$F_j[X_1, \ldots, X_n] = \Phi(X_j), \quad j = 1, \ldots, n.$$}

Obviously $(F^*)_* = F$, $(\Phi_*)_* = \Phi$, $(F \circ G)^* = G^* \circ F^*$, $(\Phi \circ \Psi)_* = (\Psi)_* \circ (\Phi)_*$
and $F^*(X_j) = F_j[X_1, \ldots, X_n]$ when $F = (F_1, \ldots, F_n)$. It is easy to check that
if $F = (F_1, \ldots, F_n) : k^n \to k^n$ is a polynomial mapping, then:

$(\ast)$ $F$ is an automorphism of $k^n \iff F^*$ is an automorphism of $k[X_1, \ldots, X_n]$.

Both authors partially supported by KBN Grant no 2 P03A 061 08
In the sequel we often will not distinguish between $F$ and $F^*$ (or between $\Phi$ and $\Phi_*$) since, in view of (§), it will not cause any confusion.

**Definition 1.2.** (i) A polynomial automorphism $F$ of $k^2$ is called a triangular automorphism if it is of the form $F(X,Y) = (X + f(Y), Y)$ or $F(X,Y) = (X, Y + g(X))$.

(ii) A polynomial automorphism $\Phi$ of $k[X,Y]$ is called triangular if $\Phi_*$ is a triangular automorphism of $k^2$.

(iii) A polynomial automorphism $F$ of $k^2$ is called tame, if it is a composition of triangular and affine (linear if $F(0) = 0$) automorphisms.

The Jung and van der Kulk theorem (cf. [Ju], [Ku]) states that a polynomial automorphism of $k^2$ is tame. This is a characterization of polynomial automorphisms of $k^2$. The structure of the group of polynomial automorphisms of $k^n$ is known only for $n = 2$. The analytic proof of the theorem can be found in [Plo] and [CK].

2. **Derivations of algebras.** Let $A$ be a commutative $k$-algebra of finite type (i.e. $A$ is finitely generated as an algebra over $k$).

**Definition 2.1.** (i) A derivation of $A$ is a $k$-linear mapping $D: A \to A$ which fulfills the Leibniz rule $D(fg) = (Df)g + fDg$.

(ii) We denote the set of derivations of $A$ by $\text{Der}(A)$.

Basic facts about derivations can be found in [Bou, chap.4] or [Mat, chap.10]; a survey of recent results and a source for the references can be found in [Now]. This gives the following formulas:

\[
D^n(fg) = \sum_{i=0}^{n} \binom{n}{i} D^i(f)D^{n-i}(g) \quad \text{and} \quad D(f^m) = mf^{m-1}D(f).
\]

Note that if $A$ is generated by the elements $\{f_1, \ldots, f_n\}$, then a derivation $D$ is determined by its values $\{Df_1, \ldots, Df_n\}$. Observe that $\text{Der}(A)$ is an $A$-module and obviously $\text{Der}(A)$ is a $k$-vectorspace, but in general the composition of two derivations is not a derivation.

**Definition 2.2.** (i) If $D \in \text{Der}(A)$, then a subalgebra $\text{Ker} D$ is called the ring of constants of $D$ and also denoted by $A^D$.

(ii) We call $f \in A$ a slice (or a principal element) of $D$ if $DF = 1$.

(iii) A derivation $D$ is called locally nilpotent if for every $f \in A$ there exists an $n = n(f) \in \mathbb{N}$ such that $D^n(f) = 0$. The set of locally nilpotent derivations of $A$ is denoted by $\text{LNDer}(A)$. 
The set $\text{Ker } D$ is a $k$-vectorspace but usually it is not an ideal of $A$. Note that a slice (if it exists) is determined up to a constant from $\text{Ker } D$ (but not every derivation has a slice e.g. $D: \mathbf{k}[X] \to \mathbf{k}[X]$, $D(f) := Xf'(X)$ for $f \in \mathbf{k}[X]$). Evidently $D$ is locally nilpotent if and only if generators of the algebra are in $\text{Ker } D^p$ for some $p \in \mathbb{N}$. The sum of locally nilpotent derivations need not be locally nilpotent, e.g. $D_1 := Y \partial_X \in \text{LNDer}(A)$, $D_2 := X \partial_Y \in \text{LNDer}(A)$, but $D_1 + D_2 \notin \text{LNDer}(A)$.

We will need a simple but useful lemma (cf. [Ren]):

**Lemma 2.3.** Let $A$ be a $k$-algebra of finite type without zero divisors, $D \in \text{LNDer}(A)$ and $g \in A$ be such that $D(g) \in gA$. Then $D(g) = 0$.

**Proof.** Assume $D(g) = gh$, with $0 \neq h \in A$. Take $l, k \in \mathbb{N}$ such that $D^k(g) \neq 0$, $D^{k+1}(g) = 0$ and $D^l(h) \neq 0$, $D^{l+1}(h) = 0$. Then

$$0 = D^{k+l+1}(g) = D^{k+l}(gh) = \sum_{i=0}^{k+l} \binom{k+l}{i} D^i(g)D^{k+l-i}(h)$$

$$= \binom{k+l}{k} D^k(g)D^l(h)$$

which is a contradiction, so $D(g) = 0$. \qed

One checks that the group of automorphisms $\text{Aut}(A)$ of a $k$-algebra $A$ acts on $\text{Der}(A)$ by conjugation:

$$\tilde{\Psi}: \text{Aut}(A) \times \text{Der}(A) \ni (F, D) \mapsto FDF^{-1} \in \text{Der}(A).$$

**Definition 2.4.** We denote the result of the action $\tilde{\Psi}(F, D) := FDF^{-1}$ of the automorphism $F$ on a derivation $D$ by $D^F$.

Note that $(D + E)^F = D^F + E^F$, $(gD)^F = F(g)D^F$ for $g \in A$, $F \in \text{Aut}(A)$ and $D, E \in \text{Der}(A)$. Since $(D^F)^n = (D^n)^F$ for any $D \in \text{Der}(A)$, this action leaves $\text{LNDer}(A)$ stable i.e. if $D \in \text{LNDer}(A)$ then $D^F \in \text{LNDer}(A)$ and we have the action:

$$\Psi: \text{Aut}(A) \times \text{LNDer}(A) \ni (F, D) \mapsto FDF^{-1} \in \text{LNDer}(A).$$

We give some examples of derivations.

**Example 2.5.** In the polynomial $k$-algebra $A = k[X_1, \ldots, X_n]$ we have canonical derivations $\partial_{X_i}$, where $\partial_{X_i}(X_j) = \delta_{ij}$ is the Kronecker delta.
EXAMPLE 2.6. Let \( F = (F_1, \ldots, F_n): k^n \to k^n \), where \( F_j \in A = k[X_1, \ldots, X_n] \). Put

\[
\Delta_j := \frac{\partial(F_1, \ldots, F_{j-1}, F_{j+1}, \ldots, F_n)}{\partial(X_1, \ldots, X_n)} : A \ni g \mapsto \frac{\partial(F_1, \ldots, F_{j-1}, g, F_{j+1}, \ldots, F_n)}{\partial(X_1, \ldots, X_n)} \in A,
\]

\( j = 1, \ldots, n \). Evidently \( \Delta_j \) is a derivation of \( A \).

EXAMPLE 2.7. If \( F \) is a polynomial automorphism of \( k[X_1, \ldots, X_n] \) then one can define derivations

\[
\frac{d}{dF_i} := \partial X_i F = F \circ \partial X_i \circ F^{-1}.
\]

They are locally nilpotent and follow the rule \( \frac{d}{dF_i}(F_j) = \delta_{ij} \), \( i, j = 1, \ldots, n \).

It is not difficult to check that \( \Delta_j = \frac{d}{dF_j} \), \( j = 1, \ldots, n \).

3. Graduations, bigraduations and locally nilpotent derivations.

In this section \( A = k[X, Y] \). Basic facts about bigraduations and graduations are in [Ren] and [Dix]. The set \( \text{Der}(A) \) is an \( A \)-module and can be considered as a graded \( A \)-module if there is given some graduation of \( A \).

For each \( \alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}^2 \) we define \( A_\alpha = kX^{\alpha_1}Y^{\alpha_2} \) if \( \alpha \in \mathbb{N}^2 \) and \( A_\alpha = 0 \) otherwise. Then

\[
A = \bigoplus_{\alpha \in \mathbb{Z}^2} A_\alpha.
\]

Let \( \text{Der}_\alpha(A) \) be the \( k \)-vectorspace spanned by \( X^{\alpha_1+1}Y^{\alpha_2} \partial_X, X^{\alpha_1}Y^{\alpha_2+1} \partial_Y \) (if \( \alpha \notin \mathbb{N}^2 \cup (\{-1\} \times \mathbb{N}) \cup (\mathbb{N} \times \{-1\}) \) then \( \text{Der}_\alpha(A) = 0 \)). Then

\[
\text{Der}(A) = \bigoplus_{\alpha \in \mathbb{Z}^2} \text{Der}_\alpha(A).
\]

If \( \alpha, \beta \in \mathbb{Z}^2 \), then \( A_\alpha \text{Der}_\beta(A) \subset \text{Der}_{\alpha+\beta}(A) \) and \( [\text{Der}_\alpha(A)](A_\beta) \subset A_{\alpha+\beta} \).

The following proposition says when \( D_\alpha \) is locally nilpotent.

PROPOSITION 3.1. Let \( \alpha \in \mathbb{Z}^2 \).

i) If \( \alpha_1 = -1 \) or \( \alpha_2 = -1 \) then \( \text{Der}_\alpha(A) \cap \text{LNDer}(A) = \text{Der}_\alpha(A) \).

ii) If \( \alpha \in \mathbb{N}^2 \) then \( \text{Der}_\alpha(A) \cap \text{LNDer}(A) = 0 \).

PROOF. i) obvious, ii) \( D \in \text{Der}_\alpha(A) \) is of the form \( D = c_1X^{\alpha_1+1}Y^{\alpha_2}\partial_X + c_2X^{\alpha_1}Y^{\alpha_2+1}\partial_Y \), for some \( c_1, c_2 \in k \). Then

\[
D^k(X^{\beta_1}Y^{\beta_2}) = \prod_{i=0}^{k-1} (c_1(i\alpha_1 + \beta_1) + c_2(i\alpha_2 + \beta_2))X^{\alpha_1+\beta_1}Y^{\alpha_2+\beta_2}.
\]
If $D_\alpha \in \text{LNDer}(A)$ then for any $\beta_1, \beta_2 \in \mathbb{N}$ there exists a $k \in \mathbb{N}$ such that $D^k(X^{\beta_1}Y^{\beta_2}) = 0$ and this means that $c_1(i\alpha_1 + \beta_1) + c_2(i\alpha_2 + \beta_2) = 0$ for some $i \in \mathbb{N}$. For example, if $\beta = \alpha$ one gets $c_1\alpha_1 + c_2\alpha_2 = 0$ and then $c_1\beta_1 + c_2\beta_2 = 0$ for any $\beta_1, \beta_2 \in \mathbb{N}$. This implies $c_1 = c_2 = 0$ and $D = 0$. □

Let $L_2 \cong \mathbb{R}^2$ be the set of $\mathbb{R}$-linear forms on $\mathbb{R}^2$. Each $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$ defines the following graduation of $A$:

$$A = \bigoplus_{\mu \in \mathbb{R}} A^\omega_\mu, \text{ where } A^\omega_\mu = \bigoplus_{\omega(\alpha) = \mu} A_\alpha,$$

and the graduation of $\text{Der}(A)$:

$$\text{Der}(A) = \bigoplus_{\mu \in \mathbb{R}} \text{Der}^\omega_\mu(A), \text{ where } \text{Der}^\omega_\mu(A) = \bigoplus_{\omega(\alpha) = \mu} \text{Der}_\alpha(A).$$

The elements of $\text{Der}^\omega_\mu(A)$ are called $\omega$-homogenous derivations of $A$ of degree $\mu$. When we say that $\omega \in \mathbb{R}^2$ is a graduation we mean by this the defined above graduation of $A$ and $\text{Der}(A)$. If $\omega = (0, 0)$, then this gives us no additional structure on $A$, so we usually assume

$$\omega \in \mathbb{R}^*_2 = \mathbb{R}^2 \setminus \{(0, 0)\}.$$  

We call a graduation $\omega$ positive if $A_\mu = 0$ for all $\mu < 0$ and $A_0 = k$. A graduation $\omega$ is positive if and only if $\omega_1, \omega_2 > 0$.

For each $\omega \in \mathbb{R}^2$, $h \in A$ and $D \in \text{Der}(A)$ we define:

$$\deg_\omega D := \max\{\mu \in \mathbb{R}: D_\mu \neq 0\},$$

$$\bar{D} := D_{\deg D},$$

$$\deg_\omega h = \max\{\mu \in \mathbb{R}: h_\mu \neq 0\}.$$  

The superscript (or subscript) $\omega$ in $A^\omega_\mu$, $\text{Der}^\omega_\mu$, $\deg_\omega D$, and $\deg_\omega h$ is dropped if $\omega$ is defined in the context.

By $\text{Supp}(D)$ we mean the finite set $\{\alpha \in \mathbb{Z}^2: D_\alpha \neq 0\}$.

By $\text{Newt}(D)$ we mean the convex hull of $\text{Supp}(D)$ and the point $(-1, -1)$.

**Proposition 3.2.** Let $D \in \text{LNDer}(A)$, $\omega \in \mathbb{R}^2$. Then $\bar{D} \in \text{LNDer}(A)$.

**Proof.** We have $D = D_{\mu_1} + \ldots + D_{\mu_r}$ with $0 \neq D_{\mu_i} \in \text{Der}_{\mu_i}(A)$, $\mu_1 < \mu_2 < \ldots < \mu_r$, and let $f \in A_\nu$ for some $\mu_i, \nu \in \mathbb{R}$. Then $Df = \sum D_{\mu_k} f$ with $D_{\mu_k} f \in A_{\mu_k + \nu}$ and $D^n f = D^n_{\mu_1} f + \sum_{k_1 \in \{2, \ldots, r-1\}} D_{\mu_{k_1}} \ldots D_{\mu_{k_n}} f + D^n_{\mu_r} f$.
where $D^n f \in A_{n\mu_1 + \nu}$, $D^n f \in A_{n\mu_r + \nu}$ and the rest is in $\bigoplus_{n\mu_1 + \nu < \mu < n\mu_r + \nu} A_{\mu}$. But $D^n f = 0$ for some $n \in \mathbb{N}$, which implies $D^n f = 0$, $D^n f = 0$. The generators of $A$ can be chosen homogenous and hence $\bar{D} = D_{\mu_r} \in \text{LNDer}(A)$.

The next theorem says what happens to the local nilpotency and homogeneity of a derivation of $A$ when it is multiplied by an element of $A$.

**Theorem 3.3.** Let $g \in A$, $\omega \in \mathbb{R}^2_*$, $D, E \in \text{Der}(A)$ be such that $D = gE$. Then:

i) $D$ is $\omega$-homogenous iff $g$ and $E$ are $\omega$-homogenous.

ii) Let $g \neq 0$. Then $D \in \text{LNDer}(A)$ iff $g \in A^D$ and $E \in \text{LNDer}(A)$.

**Proof.** i) If $g \in A_{\mu}$ and $E \in \text{Der}_{\nu}(A)$ then obviously $gE \in \text{Der}_{\mu+\nu}(A)$. Now let $gE$ be homogenous. We write homogenous decomposition of $g, E$ as $g = g_{\mu_1} + \ldots + g_{\mu_k}$, $E = E_{\nu_1} + \ldots + E_{\nu_l}$ with $\mu_1 < \ldots < \mu_k$ and $\nu_1 < \ldots < \nu_l$. We have $gE = \sum_{i,j} g_{\mu_i} E_{\nu_j}$. Since $g_{\mu_i} E_{\nu_j} \in \text{Der}_{\mu_i+\nu_j}$, we get $(gE)_{\mu_1+\nu_1} = g_{\mu_1} E_{\nu_1}$ and $(gE)_{\mu_k+\nu_1} = g_{\mu_k} E_{\nu_1}$. But $gE$ is $\omega$-homogenous, hence $k = l = 1$.

ii) Suppose $D \in \text{LNDer}(A)$, then $Dg = gEg \in gA$ gives $Dg = 0$ (by Lemma 2.3) and $Eg = 0$. Now since $Eg = 0$, we get $D^n = (gE)^n = g^n E^n$ and $D \in \text{LNDer}(A)$ if and only if $E \in \text{LNDer}(A)$.

For each $D \in \text{Der}(A)$ let

$$\sigma_1(D) = \max\{k \in \mathbb{Z}: (k, -1) \in \text{Newt}(D)\},$$

$$\sigma_2(D) = \max\{k \in \mathbb{Z}: (-1, k) \in \text{Newt}(D)\}.$$

The triangle of $D$ is the convex hull of the points $(-1, -1)$, $(\sigma_1(D), -1)$, $(-1, \sigma_2(D))$ and is denoted by $\text{Tr}(D)$. The inclusion $\text{Supp}(D) \subset \text{Tr}(D)$ is equivalent to $\text{Newt}(D) = \text{Tr}(D)$.

**Proposition 3.4.** Let $A = k[X, Y]$ and $D \in \text{LNDer}(A)$. Then $\text{Tr}(D) = \text{Newt}(D)$.

If $\text{Newt}(D) \not\subset \text{Tr}(D)$ then we can find a graduation $\omega$ such that $\bar{D} = D_{\alpha}$ for some $\alpha \in \mathbb{N}^2$. But $\bar{D}$ is locally nilpotent (3.2) and this contradicts Proposition 3.1.

**Lemma 3.5.** Let $D$ be a locally nilpotent derivation on $A = k[X, Y]$. If $D = a\partial_X + b\partial_Y \in \text{LNDer}(A)$, $a = a(X, Y), b(X, Y) \in A$ and $\gcd(a, b) = 1$, then $D$ has a slice, i.e. $D(f) = 1$ for some $f \in A$.

I. Assume additionally that $k$ is algebraically closed.

Take an $f \in A$ such that $D^2 f = 0$, $Df \neq 0$, $\deg f$ is minimal and assume that $Df \notin k^*$. 


(i) Therefore $Df$ has some prime factor $p$, i.e. $Df = ph$, where $p, h \in A \setminus k$ or $p \in A \setminus k$, $h = 1$. Since $0 = D^2 f = D(ph)$, we get $Dp \in pA$, so by Lemma 2.3 we get $Dp = 0$. It follows that $D[(p)] \subset (p)$ where $(p) := pA$ is the principal ideal of the polynomial $p$. Put $R := A/(p)$, $[f] := f + (p) \in R$ for $f \in A$ and consider the derivation

$$\hat{D} : R \to R \text{ defined by the formula } \hat{D}[f] := [Df]$$

Since $(p)$ is prime, $R$ has no zero-divisors, so we can take $K = \text{the field of fractions of } R$, and then we extend the derivation $\hat{D}$ to $K$ (we denote it again by $\hat{D}$). Note that $\hat{D} \in \text{Der } K$.

(ii) Set $\hat{K} := \text{Ker } \hat{D} \subset K$; then $\hat{K}$ is a subfield of $K$ and $k \subset \hat{K}$. Note that $\hat{K}$ is an algebraically closed subfield of $K$.

(Proof: If $t \in K$ is an algebraic element over $\hat{K}$ and if $\varphi = T^{k1} + \ldots + c_1 T + c_0 \in \hat{K}[T]$ is a minimal polynomial of $t$, then $0 = (k t^{k1} + \ldots + c_1) \hat{D} t$ and because deg $\varphi$ is minimal we get: $\hat{D} t = 0$, i.e. $t \in \hat{K}$.)

(iii) Since $R = k[[X],[Y]]$ and $p([X],[Y]) = 0$ ($[X] := X + (p)$, $[Y] := Y + (p)$), we have tr. deg$_k K = 1$ and by Noether's theorem (cf. e.g. [L]) $R$ is an integral extension of $k[g]$ for some $g \in R$. Because $\hat{K}$ is a quotient field of $R$ and $\hat{K}$ is algebraically closed in $K$.

(*) $\hat{K} = k$ or $\hat{K} = K$ or $D$ has a slice.

(iv) The case $\hat{K} = K$ is not possible.

Since $[a] = [DX] := \hat{D}[X] = [0] = \hat{D}[Y] =: [DY] = [b]$, we get $a, b \in (p)$, which is a contradiction, because gcd$(a, b) = 1$.

(v) The case $\hat{K} = k$ is not possible.

Since $\hat{D}[f] = [Df] = [hp] = [0]$, we have $[f] = f + (p) \in \hat{K} = k$, i.e. $f = tp + c$ for some $c \in k$ and $t \in A$. Hence $Df = p \hat{D} t$ and $0 = D^2 f = pD^2 t$, so $D^2 t = 0$, contradicting the minimality of $deg f$ (because $deg t < deg p + deg t = deg f$).

(vi) Thus $Df \in k^*$ and $D(\frac{1}{Df}.f) = 1$ as desired.

II. The proof in the case where $k$ is not algebraically closed follows easily by taking the algebraic closure $\bar{k}$ instead of $k$.

The next theorem is a version of Theorem 5 in [Ren]. We present here an elementary proof of it.

**Theorem 3.6.** Let $A = k[X,Y]$. For each $D \in L\text{nder}(A)$ there exist a $g \in A$ and an $E \in L\text{nder}(A)$ such that $D = gE$, $Eg = 0$ and $E$ has a slice ($Ef = 1$ for some $f \in A$).

Let $D = f_1 \partial X + f_2 \partial Y$ for some $f_1, f_2 \in A$. As $A$ is a factorial ring there exists greatest common divisor of $f_1$ and $f_2$; denote it by $g$. Hence $f_1 = gh_1$, $f_2 = rh_2$, $h_1h_2 \in A$.
$f_2 = gh_2$ and $h_1, h_2$ have no common non-constant factors. Let $E = h_1 \partial_X + h_2 \partial_Y$. If $f_1 = f_2 = 0$ there is nothing to prove, otherwise $g \neq 0$ and 3.3.ii gives $Eg = 0$ and $E \in \text{LNDER}(A)$. It remains to prove that $E$ has a slice but it is just done in 3.5. \hfill \Box

4. Locally Nilpotent Derivations of $k[X, Y]$. A simple example of a locally nilpotent derivation of $k[X, Y]$ is $f \partial_X$, where $f \in k[Y]$. Theorem 4.1 states that each locally nilpotent derivation of $A$ is of this kind up to conjugation by a polynomial automorphism of simple type. It is an analogue of the Jordan representation theorem, which states that the orbits of the action of $\text{GL}_n(k)$ on the ring $M_n$ of $n$ by $n$ matrices by the conjugation $\text{GL}_n \times M_n \ni (A, M) \mapsto A^{-1}MA \in M_n$ have Jordan representants. Rentschler’s theorem states that the orbits of the defined earlier group action $\Psi$ of $\text{Aut}(k[X, Y])$ on the set of locally nilpotent derivations LNDER(A) have representants of the form $f \partial_X$ with $f \in k[Y]$. An endomorphism $F = (F_1, \ldots, F_n)$ of $k[X_1, \ldots, X_n]$ is called triangular if $F_i - X_i \in k[X_1, \ldots, X_{i+1}]$ for $i = 1, \ldots, n-1$ and $F_n = X_n$, one checks that a triangular endomorphism is invertible (a construction of the inverse mapping is obvious). An automorphism is called tame if it is a product of invertible linear transformations and triangular automorphisms. One checks that tame automorphisms form a subgroup of $\text{Aut}(A)$.

**Theorem 4.1 [Rentschler, 1968].** Let $A = k[X, Y]$. If $D \in \text{LNDER}(A)$ then there exist a tame polynomial automorphism $F: A \to A$ and a polynomial $h \in k[Y]$ such that

$$DF = h \partial_X.$$ 

**Proof.** Let $D = a \partial_X + b \partial_Y \neq 0$ with $a, b \in A$, $\sigma := \sigma_1(D)$ and $\tau := \sigma_2(D)$. If $\sigma = -1$ then there is nothing to prove ($a \in k[Y], b = 0$). If $\tau = -1$ then $F: (X) \mapsto (Y)$ is a desired automorphism. Now we have to consider the case where $\sigma$ and $\tau$ are nonnegative. If we take the positive graduation $\omega = (\tau + 1, \sigma + 1)$ then we get $\deg D = \sigma \tau - 1$ because $\text{Tr}(D) = \text{Newt}(D)$.

There exist an $E \in \text{LNDER}(A)$ and a $g \in A$ such that $\bar{D} = gE$ and $E$ has a slice (by 3.6 and 3.3), $g$ and $E$ are $\omega$-homogenous (by 3.3.i). Then $\deg E < 0$ (because $E_\mu(A_\nu) \subset A_{\mu+\nu}$ and $E(f) = 1$ for some $f \in A$). Let $E = \sum_{\omega(r,s) = \deg(E)} D_{r,s}$ if $D_{r,s} \neq 0$ then either $r$ or $s$ must be negative because $\omega$ is positive. Hence $E = \xi \partial_X + \zeta X^r \partial_Y$ or $E = \xi Y^s \partial_X + \zeta \partial_Y$. Suppose $\sigma \geq \tau$, then

$$E = \xi \partial_X + \zeta X^r \partial_Y,$$

with $r = \frac{\sigma - \tau}{\tau + 1}$ and $\xi, \zeta \neq 0$. 


We have $E = \frac{d}{dF_1} = \Delta_1^F$ for a polynomial automorphism $F$ of $A$ ($\frac{\partial F_2}{\partial x} = -\frac{\xi}{r+1} X^r$, $\frac{\partial F_2}{\partial y} = \xi$ and $F_2 = \xi Y - \frac{\xi}{\xi(r+1)} X^{r+1}$, $F_1 := \frac{1}{\xi} X$). If $F \in \text{Aut}(A)$ then $\text{Ker} \frac{d}{dF_1} = k[F_2]$. This means that $g \in k[F_2]$. Moreover $g = \nu F_2^l$ for some $\nu \in k$ and $l \in \mathbb{N}$, because $g$ and $F_2$ are $\omega$-homogenous. Define $G := F^{-1} = (\xi X, \frac{1}{\xi} Y + \frac{\xi}{\xi(r+1)} X^{r+1})$. Now $D^G = G(g)E^G = G(g)\partial_X = \nu Y^l \partial_X$ (because $E^G(X) = G(E(\frac{1}{\xi} X)) = 1$, $E^G(Y) = G(E(\xi Y - \frac{\xi}{\xi(r+1)} X^{r+1})) = 0$ and $G(g) = \nu[G(F_2)]^l = \nu Y^l$). The functions $G$ and $F$ respect graduation $\omega$ (i.e. $G(A_\mu) \subset A_\mu$), because $F_i, G_i \in A_{\omega_i}$ (i=1,2). We get $\text{Der}_\mu(A)^G \subset \text{Der}_\mu(A)$. This means that $D^G = \tilde{D}^G$ and $\sigma_1(D^G) < \sigma_1(D)$. By induction $\sigma_1(D) = -1$ or $\sigma_2(D) = -1$. □

Now we are able to prove the main theorem that was proved by H. Jung in 1942 (but "Jung's proof is of very delicate matter", cf. [Na]) and W. van der Kulk (1953).

**Theorem 4.5 (Main Theorem).** Each polynomial automorphism $F = (F_1, F_2): k^2 \to k^2$ is tame.

*Proof.* (cf. [Ess, Ren]). Take the derivation $\frac{d}{dF_1} := F \circ \partial_X \circ F^{-1}$ of $k$-algebra $A := k[X, Y]$. The derivation $\frac{d}{dF_1}$ is locally nilpotent, and it has a slice: $\frac{d}{dF_1} F_1 = 1$ and $\frac{d}{dF_1} F_2 = 0$. From Rentschler's theorem one gets a tame automorphism $G$ of $A$ and a polynomial $p \in k[Y]$ such that

$$G \circ \frac{d}{dF_1} \circ G^{-1} = p(Y) \partial_X.$$

Taking the value of this derivation on the polynomial $G(F_1)$ we get:

$$1 = G \circ \frac{d}{dF_1} \circ G^{-1}[G(F_1)] = p(Y) \partial_X [G(F_1)]$$

and hence $p(Y) = \alpha$, $F_1 \circ G_\ast(X, Y) := G(F_1) = \frac{1}{\alpha} X + q(Y)$ for some $\alpha \in k^\ast$.

On the other hand, computing the kernel of the derivation $D := G \circ \frac{d}{dF_1} \circ G^{-1} = p(Y) \partial_X$ we conclude that the equation $D f = 0$ is equivalent to the following

$$p(Y) \partial_X f = 0 = G \circ \frac{d}{dF_1} \circ G^{-1}(f) \iff f \in k[Y] \iff G^{-1}(f) \in k[F_2],$$

i.e. $k[Y] = k[G(F_2)]$. Note that if $R, S \in A$, then:

$$k[R] = k[S] \iff R = \beta S + \gamma \text{ for } \beta \in k^\ast, \gamma \in k.$$

Hence $G(F_2) = \beta Y + \gamma$ for some $\beta \in k^\ast$, $\gamma \in k$. Therefore the automorphism $G(F) = (\frac{1}{\alpha} X + q(Y), \beta Y + \gamma)$ is tame so is $F$ (since $G$ is tame). □

After finishing the proof one can check that it will work in the case of an arbitrary field $k$. 
References


Received February 24, 1997

Jagiellonian University
Instytut Matematyki
Reymonta 4
PL 30–059 Kraków

e-mail: druzkows@im.uj.edu.pl
e-mail: gurycz@im.uj.edu.pl