J. Math. Kyoto Univ. (JMKYAZ)
15-3 (1975) 555-571

Field generators in two variables

By

Peter RUSSELL

(Communicated by Prof. Nagata, Sept. 3, 1974)

A field generator in two variables is a polynomial $f(x, y)$ such that f together with some rational function $g(x, y)$ generates the field $k(x, y)$ (k a field of constants). It was conjectured by Abhyankar and proved by Jan (for k of characteristic 0) that f has at most two points at infinity, that is, the degree form of f has at most two irreducible factors. The aim here is to give more precise results. First (see 3.7), unless $\{f=0\}$ is isomorphic to a line, there are exactly two infinitely near points of f on L, the line at infinity. (So if f has two ordinary points at infinity, no branch is tangent to L. Otherwise, branches are at most simply tangent.) More generally (see 3.6), there is a quite sharp bound on the number of infinitely near points of f on L in terms of the genus of $k(x, y)$ over $k(f)$. Secondly (see 4.5), after a suitable automorphism φ of $k[x, y]$, f is linear or has two (ordinary) points at infinity. (φ is essentially unique in the latter case (see 4.7).) φ is shown to be tame, and a proof of the result of Jung [1] and Van der Kulk [5] on the structure of automorphisms of $k[x, y]$ appears as a byproduct.

I would like to express my thanks here to S. S. Abhyankar and participants in his seminar at Purdue, where I was first introduced to these problems and learned how much there is still to be learned about polynomials in two variables.

1. Let k be a field, $k[x, y]$ the polynomial ring in two variables over k and $f \in k[x, y]$, $f \notin k$. Then $k(x, y)$ is of transcendence degree 1 over $k(f)$ and hence the function field of a complete regular curve over $k(f)$, which we denote by C_f. Let t be transcendental over k and

X, Y variables over $k(t)$. Then

1.1 $(k(t)[X, Y]/f(X, Y)-t) \xrightarrow{\sim} k(f)[x, y]$

under the map sending t to f, X to x and Y to y. We are thus led to investigate the pencil of Curves $\{V(f-\lambda) | \lambda \in k\}$, of which $V(f-t)$ is the generic member (see 2.8). (We use $V(f)$ to denote the curve, or effective divisor, defined by f.)

In order to enable us to use geometric arguments, we imbed the the affine plane $\boldsymbol{A}_k^2$ in the projective plane $\boldsymbol{P}_k^2$ in the usual way, choosing projective coordinates (X_0, X_1, X_2) on $\boldsymbol{P}_k^2$ such that $x=X_1/X_0$ and $y=X_2/X_0$. Then the *line at infinity* of $\boldsymbol{A}_k^2$ is $L=V(X_0)=\boldsymbol{P}_k^2-\boldsymbol{A}_k^2$. Let $d=\deg f$ and $F(X_0, X_1, X_2)=X_0^d f(X_1/X_0, X_2/X_0)$. Then $V(F-\lambda X_0^d)$ is the closure in $\boldsymbol{P}_k^2$ of $V(f-\lambda)$ and the points of $V(F-\lambda X_0^d)-V(f-\lambda)$, in one-one correspondence with the irreducible factors of the degree form of f, are the *points at infinity* of f. We define

1.2 $\Lambda(f)=\{V(\alpha_0 F+\alpha_1 X_0^d) | (\alpha_0, \alpha_1) \in \boldsymbol{P}_k^1\}$.

By 1.1, C_f is the normalization of $V(F-tX_0^d)$. Also, $k(f)[x, y]$ is a regular ring (it is a localization of $k[x, y]$). Hence $V(f-t)$ is an affine open subset of C_f. The points of $C_f-V(f-t)$ we will call the points at infinity of C_f.

1.3 Definition: f is a field generator if there exists $g \in k(x, y)$ such that $k(f, g)=k(x, y)$ or, equivalently, if C_f is a rational curve over $k(f)$.

Remark: It does not seem to be known whether, given that f is a field generator, g can be chosen in $k[x, y]$, or, equivalently, whether C_f has a point at infinity rational over $k(f)$.

2. Most of the contents of this section are quite well known, and its main purpose is to establish a coherent notation. We assume that k is algebraically closed.

Let S be a complete non-singular surface, $p_0 \in S$ (all points are assumed to be closed unless otherwise designated) and (see [3, II, §4, 2]).

$$\pi: S' \to S$$

the local quadratic transformation (l.q.t.) or blowing up with centre p_0. We denote by $E=\pi^{-1}(p_0)$ the exceptional fibre of π. Let C be a curve on S. We put

$$\mu(p_0, C) = \text{multiplicity of } p_0 \text{ on } C.$$

Let $\pi'(C)$ be the *proper transform* and $\pi^*(C)=\pi'(C)+\mu(p_0, C)E$ the *total transform* of C on S'. Let $(-,-)$ denote the intersection product on both S and S'.

Then (see [3, IV, § 3, 2])

2.1 (i) π^* preserve linear equivalence and the intersection product.

(ii) $E \simeq \boldsymbol{P}_k^1$ and $(E, E) = -1$.

(iii) If C is a curve on S,

$$(\pi'(C), E) = \mu(p_0, C).$$

(iv) If C, D are curves on S,

$$(C, D) = (\pi'(C), \pi'(D)) + \mu(p_0, C)\mu(p_0, D).$$

Let C be an irreducible curve on S. Then the arithmetic genus of C is given by (see [2, IV, § 2, 8]) $p_a(C)=1+\frac{1}{2}(C, C+K)$, where K is a canonical divisor on S. Now $K'=\pi^*(K)+E$ is a canonical divior on S' and hence

2.2 $p_a(C)=p_a(\pi'(C))+\frac{1}{2}\mu(p_0, C)(\mu(p_0, C)-1)$.

We note that in case k is not algebraically closed, 2.1 and 2.2 remain valid if p_0 is rational over k.

2.3 Definition: (i) An *infinitely near* (i.n.) point of S is a sequence

$$q=(p_i, p_{i-1}, \cdots, p_0)$$

such that $p_0 \in S_0 = S$ and for $0<j\leq i$, $p_j \in \pi_j^{-1}(p_{j-1}) = E_j \subset S_j$, where

$$\pi_j: S_j \to S_{j-1}$$

is the l.q.t. with centre p_{j-1}. We will also say that q is infinitely near to p_0. An i.n. point $q=(p_0)$ will be called an *ordinary* point of S.

(ii) Let D be a curve on S. Then

$$\mu(q, D) = \mu(p_i, D^{(i)}),$$

where $D^{(i)}$ is the proper transform of D on S_i. We say q is on D if $\mu(q, D) > 0$, i.e. $p_i \in D^{(i)}$. (Note that then all i.n. points $q_j = (p_j, \cdots, p_0)$, $0 \leq j \leq i$, are on D.)

Remark: Let $\pi_j: S_j \to S_{j-1}, j=1, \cdots, l$ be a sequence of l.q.t. and $p \in S_l$. Then p determines uniquely an i.n. point $q=(p_i, \cdots, p_0)$ (with $i<l$ in general). If there is no danger of confusion, we will call p an i.n. point of S_0.

Let Λ be a linear system of curves on S and $p \in S$. We put

$$\mu(p, \Lambda) = \min\{\mu(p, D) \mid D \in \Lambda\}.$$

(Then $\mu(p, \Lambda)$ is the multiplicity at p of a *general member* of Λ, i.e. $\mu(p, \Lambda) = \mu(p, D)$ for D ranging over a dense open subset of Λ.) Let π be the l.q.t. with centre p. Then

$$\pi^*(\Lambda) = \{\pi^*(D) \mid D \in \Lambda\}.$$

the *total transform* of Λ, is a linear system with $\mu(p, \Lambda)E$ as fixed component. We define the *proper transform* of Λ by

$$\pi'(\Lambda) = \{\pi^*(D) - \mu(p, \Lambda)E \mid D \in \Lambda\}.$$

$\pi'(\Lambda)$ is a linear system not having E as fixed component. We note that as a consequence of 2. 1 (iv)

2. 4 $\quad (\Lambda, \Lambda) = (\pi'(\Lambda), \pi'(\Lambda)) + \mu(p, \Lambda)^2$

(where $(\Lambda, \Lambda) = (D, D')$ for any $D, D' \in \Lambda$).

2. 5 Definition: Let $q=(p_i, \cdots, p_0)$ be an i.n. point of S.

(i) $\mu(q, \Lambda) = \mu(p_i, \Lambda^{(i)})$, where $\Lambda^{(i)}$ is the proper transform of Λ on S_i.

(ii) q is a *base point* of Λ if $\mu(q, \Lambda)>0$. $B=B(\Lambda)$ is the set of base points of Λ. (Note that B is finite if Λ has no fixed component.)

(iii) Let $\pi_{i+1}: S_{i+1} \rightarrow S_i$ be the l.q.t. with centre p_i and $E_{i+1}=\pi_{i+1}^{-1}(p_i)$. Suppose q is a base point. Then q is *non-terminal* if a general member of $\Lambda^{(i+1)}$ meets E_{i+1} only in base points of Λ. Otherwise, q is *terminal.*

2.6 Remark: Suppose Λ is one-dimensional and let $g=\alpha_0 f_0 + \alpha_1 f_1$, $(\alpha_0, \alpha_1) \in \boldsymbol{P}_k{}^1$ be a local equation of $\Lambda^{(i)}$ at p_i. Let F_0, F_1 be the leading forms of f_0, f_1 and G the leading form of g for general (α_0, α_1). Then there are the following possibilities:

(i) $\deg F_0 \neq \deg F_1$, say $\deg F_0 < \deg F_1$. Then $G=\alpha_0 F_0$.

(ii) $\deg F_0 = \deg F_1$ and $F_1=\beta F_0$, $\beta \in k$. Then $G=(\alpha_0+\beta\alpha_1) F_0$.

(iii) $\deg F_0 = \deg F_1$ and $H=GCD(F_0, F_1) \neq F_0$. Then $F_0=H\tilde{F}_0$, $F_1=H\tilde{F}_1$ with $GCD(\tilde{F}_0, \tilde{F}_1)=1$ and $\deg \tilde{F}_0 = \deg \tilde{F}_1 > 0$. Now $G=H(\alpha_0\tilde{F}_0+\alpha_1\tilde{F}_1)$.

The points of $\Lambda^{(i+1)}$ on E_{i+1} are given by the different irreducible factors of G. In cases (i) and (ii) these are independent of (α_0, α_1) and lead to base points of $\Lambda^{(i+1)}$ on E_{i+1}. In case (iii), factors of $\alpha_0 F_0 + \alpha_1 F_1$ depend on (α_0, α_1) and do not lead to base points. So q is terminal in that case.

2.7 Definition: Assume Λ has no fixed component. Let $p \in S$. Then

$$m(p)=m(p, \Lambda)=\sum \mu(q, \Lambda),$$

the sum extended over all base points of Λ i.n. to p. If $T \subset S$, then

$$m(T)=m(T, \Lambda)=\sum_{p \in T} m(p, \Lambda).$$

A pencil Λ on S, which we assume to be without fixed component, defines a rational map λ: $S \rightarrow \boldsymbol{P}_k{}^1$.

2.8 Definition: The *generic member* Λ_η of Λ is the fibre of λ over the generic point η of $\boldsymbol{P}_k{}^1$.

Λ_η is a curve on $S\otimes\kappa(\eta)$, where $\kappa(\eta)$ is the residue field of η. Since $\kappa(\eta)$ is purely transcendental over k, an ordinary base point of Λ on S defines, by extension of scalars, a unique point on Λ_η. We then have the following easy version (which has the advantage of being true if char $k>0$) of Bertini's theorem.

2.9 Lemma: *The generic member of a pencil without fixed component is regular outside the base points of the pencil.*

Proof: We can cover S by affine open sets U with coordinate rings A such that there exist $f_0, f_1 \in A$ with

$$\Lambda|U = \{V(\alpha_0 f_0 + \alpha_1 f_1) \mid (\alpha_0, \alpha_1) \in \boldsymbol{P}_k{}^1\}$$

and $(f_0, f_1)A$ a zero-dimensional ideal. Then there is a $t\in\kappa(\eta)$ such that $\kappa(\eta)=k(t)$ and f_0+tf_1 is an equation for Λ_η in $A\otimes k(t)$. Now generalizing 1.1

$$(A\otimes k[t]/f_0+tf_1)_{f_1} \simeq A_{f_1},$$

and hence $(A\otimes k(t)/f_0+tf_1)_{f_1} \simeq T^{-1}A_{f_1}$, where $T\subset A_{f_1}$ is the multiplicative set of all non zero polynomials over k in f_0/f_1. Let $I\subset A\otimes k(t)$ be the maximal ideal of a point p on Λ_η (i.e. $f_0+tf_1\in I$). If $f_1\notin I$, then p is a regular point of Λ_η by the above since $T^{-1}A_{f_1}$ is a regular ring. If, on the other hand, $f_1\in I$, then $f_0\in I$ and it follows that I is the extension to $A\otimes k(t)$ of a maximal ideal $I'\subset A$ such that $f_0, f_1\in I'$. Hence p is a base point of Λ.

Remark: The strong version of Bertini's theorem asserts regularity of Λ_η over the algebraic closure of $\kappa(\eta)$. This, of course, may fail if char $k>0$.

Let p be an ordinary base point of Λ. It is easily seen (for instance by the discussion in 2.6) that $\mu(p, \Lambda)=\mu(p, \Lambda_\eta)$, and it follows that the proper transform of Λ_η under the l.q.t. with center p is the generic member of the proper transform of Λ. 2.9 therefore extends to i.n. points, that is, all i.n. singular points of Λ_η are base points of Λ. In particular, they are rational over $\kappa(\eta)$.

Since Λ has no fixed component, we can find a sequence of l.q.t.

$$S^* = S_l \xrightarrow{\pi_l} S_{l-1} \to \cdots \to S_1 \xrightarrow{\pi_1} S_0 = S$$

with centres at base points of Λ and such that Λ^*, the proper transform of Λ on S^*, is free of base points. Then $(\Lambda^*, \Lambda^*) = 0$ since two distinct members of Λ^* do not meet. By repeated application of 2. 4

2. 10 $(\Lambda, \Lambda) = \sum \mu(q, \Lambda)^2, \qquad q \in B.$

Now Λ_η^* is regular and obtained from Λ_η by l.q.t. with centres rational over $\kappa(\eta)$. Hence Λ_η^* is the normalization of Λ_η and $p_a(\Lambda_\eta^*) = g$, the genus of $\kappa(\Lambda_\eta)$ over $\kappa(\eta)$, where $\kappa(\Lambda_\eta)$ is the function field of Λ_η. By repeated application of 2. 2 we obtain

2. 11 $p_a(\Lambda_\eta) = g + \frac{1}{2} \sum \mu(q, \Lambda)(\mu(q, \Lambda) - 1), \qquad q \in B.$

3. We assume that k is algebraically closed in this section. Otherwise we return to the notation of section 1.

Let $f \in k[x, y]$, $d = \deg f > 0$ and $\Lambda = \Lambda(f)$ (see 1. 2). Then $d^2 = (\Lambda, \Lambda)$ and $p_a(\Lambda_\eta) = \frac{1}{2}(d-1)(d-2)$. By 2. 10 and 2. 11 we have

3. 1 $d^2 = \sum \mu(q, \Lambda)^2, \qquad q \in B,$

3. 2 $(d-1)(d-2) = 2g + \sum \mu(q, \Lambda)(\mu(q, \Lambda) - 1), \qquad q \in B.$

Hence

3. 3 $\sum \mu(q, \Lambda) = 3d + 2(g-1), \qquad q \in B.$

Here g is the genus of C_f, or of $k(x, y)$, over $k(f)$.

Pencils of type $\Lambda(f)$ have the d-fold line at infinity as a member. In fact, $\Lambda_\infty = V(X_0^d) = dL$, where Λ_∞ is the member of Λ given by $\alpha_0 = 0, \alpha_1 = 1$ $(\infty = (0, 1) \in \boldsymbol{P}_k^1)$. We wish to exploit this special property. Let

$$S_l \xrightarrow{\pi_l} S_{l-1} \to \cdots \to S_1 \xrightarrow{\pi_1} S_0 = \boldsymbol{P}_k^2$$

be a composite of l.q.t. Let $p_j \in S_j$ be the centre of π_{j+1} and $E_{j+1} = \pi_{j+1}^{-1}(p_j)$, $j = 0, \cdots, l-1$. Put $E_0 = L$. If D is a curve on some S_i, denote by $D^{(j)}$ its proper transform on S_j, $j \geq i$. $\Lambda^{(j)}$ will be the proper transform of Λ on S_j, and $\Lambda_\infty^{(j)}$ the member of $\Lambda^{(j)}$ given by $\infty \in \boldsymbol{P}_k^1$ (to be distinguished from $(\Lambda_\infty)^{(j)}$).

3.4 Definition: Let D be an irreducible curve on S_l. $\varepsilon(D)$ is the multiplicity of D as a component of $\Lambda_\infty^{(l)}$, i.e. $\Lambda_\infty^{(l)}=\varepsilon(D)D+C$, where C does not have D as a component.

We note the following facts concerning $\varepsilon(D)$.

3.5.1 $D=\pi_l'(\tilde{D})$ for some $\tilde{D}\subset S_{l-1}$, then $\varepsilon(D)=\varepsilon(\tilde{D})$.

3.5.2 $\varepsilon(E_0)=d$.

3.5.3 $\varepsilon(D)\geq 0$ and if $\varepsilon(D)>0$, then $D=E_i^{(l)}$ for some $i\leq l$.

3.5.4 $\varepsilon(E_l)=\sum_{j=0}^{l-1}\varepsilon(E_j)\mu(p_{l-1},E_j^{(l-1)})-\mu(p_{l-1},\Lambda^{(l-1)})$.

In fact, by 3.5.1 and by 3.5.3,

$\sum_{j=0}^{l-1}\varepsilon(E_j)\mu(p_{l-1},E_j^{(l-1)})=\mu(p_{l-1},\Lambda_\infty^{(l-1)})=$ multiplicity of E_l in $\pi_l^*(\Lambda_\infty^{(l-1)})$.

Remark: $\mu(p_{l-1},E_j^{(l-1)})=0$ or 1, and 1 for at most two j.

3.5.5 If p_{l-1} is a terminal base point of Λ (see 2.5), then $\varepsilon(E_l)=0$.

In fact, E_l is not a fixed component of $\Lambda^{(l)}$, but $\Lambda^{(l)}$ meets E_l in infinitely many points. Hence E_l is not a component of any member of $\Lambda^{(l)}$.

3.5.6 If p_{l-1} is a base point of Λ and $\varepsilon(E_l)>0$, then all points of (a general member of) $\Lambda^{(l)}$ on E_l are base points of Λ, and there is at least one such.

In fact, since p_{l-1} is a base point, all members of $\Lambda^{(l)}$ meet E_l. But E_l is a component of $\Lambda_\infty^{(l)}$, and hence if $\Lambda_\infty^{(l)}\neq D\in\Lambda^{(l)}$, D meets E_l only in base points of Λ.

3.5.7 Let $p_l\in E_l$ be a base point of Λ. Then $\varepsilon(E_l)\leq m(p_l)$ (see 2.7).

In fact, we can find an i.n. point $(p_{l+r},\cdots,p_l)$ of S_l such that p_{l+r} is a terminal base point of Λ. If E_{l+j+1} is the exceptional fibre above p_{l+j}, then $p_{l+j}\in E_{l+j}, j=0, \cdots, r$, and

$$\varepsilon(E_{l+r+1})\geq\varepsilon(E_l)-\sum_{j=0}^{r}\mu(p_{l+j}, \Lambda^{(l+j)})$$

be repeated application of 3.5.4. Now $\varepsilon(E_{l+r+1})=0$ by 3.5.5 and hence $\varepsilon(E_l)\leq\sum_{j=0}^{r}\mu(p_{l+j}, \Lambda^{(l+j)})\leq m(p_l)$.

3.5.8 Let s be the number of i.n. base points of Λ on E_l. Then $s\varepsilon(E_l)\leq m(E_l)$ (see 2.7).

In fact, suppose $q=(p_{l+r}, \cdots, p_l)$ is a base point of Λ on E_l. Then $q_0=(p_l), q_1=(p_{l+1}, p_l), \cdots, q_{l+r}=q$ are base points of Λ on E_l. We have $\mu(p_{l+j}, E_l^{(l+j)})=1$ and $\mu(p_{l+j}, E_{l+j})=1$ for $j=0, \cdots, r$. Repeated application of 3.5.4 gives

$$\varepsilon(E_{l+r+1})\geq(r+1)\varepsilon(E_l)-\sum_{j=0}^{r}\mu(p_{l+j}, \Lambda^{(l+j)}).$$

If $\varepsilon(E_{l+r+1})=0$, let $m=0$. Otherwise there is a base point p_{l+r+1} of Λ on E_{l+r+1} by 3.5.6, and we let $m=m(p_{l+r+1})\geq\varepsilon(E_{l+r+1})$ (by 3.5.7). Hence $(r+1)\varepsilon(E_l)\leq\sum_{j=0}^{r}\mu(p_{l+j}, \Lambda^{(l+j)})+m\leq m(p_l)$. We may assume that $r+1$ is the exact number of base points of Λ on E_l i.n. to p_l (p_l determines a unique maximal sequence of them). Summing over all ordinary base points of $\Lambda^{(l)}$ on E_l we obtain the desired result.

3.6 Theorem: *Let $f\in k[x,y]$, $d=\deg f>0$, g the genus of $k(x, y)$ over $k(f)$ and s the number of points of f on the line at infinity of A_k^2, including all infinitely near points. Then*

$$(s-3)d\leq 2(g-1).$$

Proof: The i.n. points of f on L, that is the i.n. points common to $V(F)$ and L, are base points of $\Lambda=\Lambda(f)$ since $V(F)$ and L are components of different members of Λ. Also, Λ has no base points on A_k^2, and hence $m(L)=\sum\mu(q, \Lambda), q\in B$. By 3.5.2, 3.5.8 and 3.3 we have $sd\leq 3d+2(g-1)$.

3.7 Corollary: *Let $f\in k[x, y]$ be a field generator. Then there are at most two infinitely near points of f on the line at in-*

finity of A_k^2. *In particular, the degree form of* f *has at most two distinct irreducible factors.*

Proof: $k(x, y)$ is purely transcendental over $k(f)$, so $g=0$ and $(s-3)d<0$. Hence $s\leq 2$.

3.8 Proposition: *Let* k *be any field and* f *a field generator over* k. *Then the points at infinity of* f *are rational over* k, *that is, the degree form of* f *splits into linear factors over* k.

Proof: The points at infinity of f are base points of $\Lambda(f)$. We will consider them as points of $\Lambda_\eta = V(F-tX_0^d)$ and show that they are rational over $k(t)$. Now over $\bar{k}(t), \bar{k}$ an algebraic closure of k, there are at most two, and hence over $k(t)$ there are at most two with the sum of their separable degrees ≤ 2.

Note that $V(f-t)\subset C_f$ contains a point q rational over $k(t)$ since C_f is a rational curve. Also $R=k(t)[X, Y]/f-t$ has unique factorization by 1.1 and there exists $h\in R$ such that $(h)=q+\sum_{i=1}^{r} n_i\bar{q}_i$, where $\bar{q}_1, \cdots, \bar{q}_r$ are the points at infinity of C_f and (h) is the divisor of h on C_f. Hence $GCD(\deg \bar{q}_1, \cdots, \deg \bar{q}_r)=1$, and it follows that $GCD(\deg q_1, \cdots, \deg q_s)=1$ if $q_1, \cdots, q_s$ are the points at infinity of f. We conclude that there is at least one q_i rational over $k(t)$ and, possibly, one more, q_1 say, purely inseparable over $k(t)$. Let in that case κ be the residue field of q_1 and $[\kappa: k(t)]=p^n=b$, where $p=\text{char}\ k$. We note that Λ_η is not tangent to L at q_1 over $\bar{k}(t)$ since f already has two ordinary points at infinity.

Let A be the local ring of q_1 on $\boldsymbol{P}_{k(t)}^2$ and M the maximal ideal of A. Now there exist parameters u, v for A such that v is a local equation for L and $u=x^b-a$, where $a\in k-k^p$ and x mod M generates κ over $k(t)$. Now $\bar{A}=A\otimes_{k(t)}\kappa$ is the local ring of q_1 on $\boldsymbol{P}_\kappa^2$ and there exist parameters $\bar{u}, \bar{v}$ for $\bar{A}$ such that $\bar{u}^b=u\otimes 1$ and $\bar{v}=v\otimes 1$. Let $g\in A$ be a local equation for Λ_η at q_1. Then monomials appearing in the power series expansion of $\bar{g}=g\otimes 1$ are of the form $\bar{u}^{bi}\bar{v}^j$, where $u^i v^j$ appears in the power series expansion of g. Since, as stated above, $\bar{v}$ and $\bar{g}$ are not tangent, a term $\bar{u}^{be}$ appears in the leading form of $\bar{g}$. But $bi+j\geq bc$ implies $i+j>e$ if $b>1$, and the leading form of g is u^e. Hence $g-u^e\in M^{e+1}$. It follows that if $\bar{q}_1$ is a point of C_f above q_1, then $u=x^b-a$ has value ≥ 2 at $\bar{q}_1$. By [4, prop. on p. 405 and thm. 2] the genus of C_f drops if the base field is extended to κ, and this is impossible if f is a field generator.

4. An automorphism $\varphi: \boldsymbol{A}_k^2\to\boldsymbol{A}_k^2$ given by $\varphi^*: k[x,y]\to k[x,y]$ is *elementary* if either $\varphi^*(x)=x$ and $\varphi^*(y)=y+g(x)$, $g\in k[x]$, or both $\varphi^*(x)$ and $\varphi^*(x)$ are linear. φ is *tame* if it can be written as a composite of elementary automorphisms.

An automorphism φ of $\boldsymbol{A}_k^2$ determines a rational map

$$\tilde{\varphi}_0: \boldsymbol{P}_k^2=S\to\tilde{S}=\boldsymbol{P}_k^2$$

such that $\tilde{\varphi}_0|\boldsymbol{A}_k^2=\varphi$. Now either $\tilde{\varphi}_0$ is a morphism (in case φ is linear), or $\tilde{\varphi}_0$ has a unique fundamental point p_0. In fact, p_0 is the unique point of S corresponding to $\tilde{E}_0$, the line at infinity of $\tilde{S}$, which is the only irreducible curve on $\tilde{S}$ not corresponding to a curve on S. Clearly $p_0\in E_0$, the line at infinity of S. Let

$$\pi_1: S_1\to S_0=S$$

be the l.q.t. with centre p_0 and

$$\tilde{\varphi}_1: S_1\to\tilde{S}$$

the rational map such that $\tilde{\varphi}_1=\tilde{\varphi}_0\circ\pi_1$. Again, $\tilde{\varphi}_1$ is a morphism or has a unique fundamental point $p_1\in\pi_1^{-1}(p_0)=E_1$. Continuing we obtain uniquely a sequence of l.q.t.

4.1 $S_l\xrightarrow{\pi_l}S_{l-1}\to\cdots\to S_1\xrightarrow{\pi_1}S_0$

and rational maps

$$\tilde{\varphi}_j: S_j\to\tilde{S},\ j=0,\cdots,l$$

such that $\tilde{\varphi}_l$ is a morphism and for $j=0,\cdots,l-1$

(i) $\tilde{\varphi}_{j+1}=\tilde{\varphi}_j\circ\pi_{j+1}$,

(ii) p_j, the centre of π_{j+1}, is a fundamental point of $\tilde{\varphi}_j$,

(iii) $p_j\in E_j$, where $E_j=\pi_j^{-1}(p_{j-1})$ (for $j\geq 1$).

Remark: $P_0, \cdots, p_{l-1}$ are the i.n. base points of the linear system $\Phi = \{V(\alpha_0\varphi^*(x) + \alpha_1\varphi^*(y) + \alpha_2)\}$.

The following now is easily verified by direct calculation (computing the base points of successive proper transforms of Φ, for instance). As before, if D is a curve on S_i, $D^{(j)}$ will be its proper transform on $S_j, j \geq i$.

4. 2 Suppose $\varphi^*(x) = x$ and $\varphi^*(y) = y + a_2x^2 + \cdots + a_nx^n$, $n \geq 2$, $a_i \in k$, $a_n \neq 0$. Then the sequence 4. 2 is determined as follows

(i) $l = 2n-1$,

(ii) $p_0 \in E_0, p_1 \in E_0^{(1)} \cap E_1$,

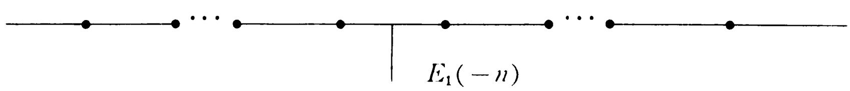
(iii) for $2 \leq j \leq n-1, p_j \in E_1^{(j)} \cap E_j$,

(iv) for $n \leq j \leq 2n-2, p_j \notin E_i^{(j)}$ for any $i < j$,

(v) p_n is in one-one correspondence with a_n, $a_n \neq 0$, and for $n+1 \leq j \leq 2n-2$, once $p_n, \cdots, p_{j-1}$ and $a_n, \cdots, a_{2n-j+1}$ are fixed, p_j is in one-one correspondence with a_{2n-j}.

The figure below gives a schematic description of the configuration of $E_0, \cdots, E_{2n-1}$ (or rather, their proper transforms) on S_{2n-1}. The number given in parenthese behind each E_i is $(E_i^{(2n-1)}, E_i^{(2n-1)})$.

$E_0(-1)$ $E_2(-2)$ $E_{n-1}(-2)$ $E_n(-2)$ $E_{n+1}(-2)$ $E_{2n-2}(-2)$ $E_{2n-1}(-1)$



Now $\tilde{\varphi}_{2n-1}$ maps E_{2n-1} isomorphically onto $\widetilde{E}_0$, and $\tilde{\varphi}_{2n-1}$ is a composite

$$\tilde{\varphi}_{2n-1} = \tilde{\pi}_1 \circ \cdots \circ \tilde{\pi}_{2n-1}$$

of l.q.t. which, looked at from the top, consist in shrinking successively the proper transforms of $E_0, E_2, \cdots, E_{2n-2}, E_1$. (This can again be verified by direct calculation. It helps to note that $(\tilde{\varphi}_0)^{-1} = \bar{\psi}_0$, where $\psi^*(x) = x$ and $\psi^*(y) = y - a_2x^2 - \cdots - a_nx^n$, so that $(\tilde{\varphi}_0)^{-1}$ has a sequence of i.n. fundamental points of the same type as $\tilde{\varphi}_0$.) Hence if $\widetilde{E}_j$ is the exceptional fibre of $\tilde{\pi}_j$

4.3 (i) $E_0^{(2n-1)} = \tilde{E}_{2n-1}$, $E_{2n-1} = \tilde{E}_0^{(2n-1)}$,

(ii) $E_1^{(2n-1)} = \tilde{E}_1^{(2n-1)}$,

(iii) $E_{2n-j}^{(2n-1)} = \tilde{E}_j^{(2n-1)}$, $j = 2, \cdots, 2n-2$.

Conditions (i) to (iv) of 4.2 allow, in view of (v), the reconstruction of ψ up to automorphisms of S and $\tilde{S}$ induced by linear automorphisms of $\boldsymbol{A}_k^2$. Hence

4.4 Lemma: *Suppose 4.1 is a sequence of l.q.t. such that conditions (i) to (iv) of 4.2 are satisfied. Let* $\pi = \pi_1 \circ \cdots \circ \pi_{2n-1}$. *Then there exists an elementary automorphism* φ: $\boldsymbol{A}_k^2 \to \boldsymbol{A}_k^2$ *and a morphism* $\tilde{\pi}$: $S_{2n-1} \to \boldsymbol{P}_k^2$ *such that*

$$\begin{tikzcd} & S_{2n-1} \arrow[dl, "\pi"'] \arrow[dr, "\tilde{\pi}"] & \\ \boldsymbol{P}_k^2 = S \arrow[rr, "\tilde{\varphi}_0"'] & & \tilde{S} = \boldsymbol{P}_k^2 \end{tikzcd}$$

commutes and $\tilde{\pi} = \tilde{\pi}_1 \circ \cdots \circ \tilde{\pi}_{2n-1}$ *is a composite of l.q.t. such that 4.3 is satisfied.*

4.5 Theorem: *Let* $f \in k[x, y]$ *be a field generator. Then there exists a tame automorphism* ψ: $\boldsymbol{A}_k^2 \to \boldsymbol{A}_k^2$ *such that either* $\psi^*(f)$ *is of degree 1 or the degree form of* $\psi^*(f)$ *has two distinct irreducible factors. Equivalently,* $V(\psi^*(f)) = \psi^{-1}(V(f))$ *is either a line or has two (ordinary) points at infinity.*

Proof: We assume first that k is algebraically closed. Our aim is to show that $\Lambda(f) = \Lambda$ has a sequence of i.n. base points as described in 4.2 (i) to (iv). We keep the notation used there and put $\mu_i = \mu(p_i, \Lambda^{(i)})$.

Suppose Λ has only one ordinary base point on E_0, p_0 say. Since $d = \deg f = (E_0, \Lambda)$ ($(E_0, \Lambda) = (E_0, D)$ for any $D \in \Lambda$) and since a general member of Λ is irreducible ($k(f)$ is algebraically closed in $k(x, y)$) either $d = 1$ or Λ is tangent to E_0 at p_0, i.e. there is a second i.n. base point (p_1, p_0) on E_0. By 3.7 there are at most two, and hence

(assuming $d>1$)

(1) $$d=\mu_0+\mu_1$$

by 2.1 (iv). Arguing as in the proof of 3.6, we have $m(E_0)=3d-2$. Hence $m(E_2)\leq m(E_0)-\mu_0-\mu_1=2d-2$. Now $\varepsilon(E_1)=d-\mu_0$ and $\varepsilon(E_2)=\varepsilon(E_0)+\varepsilon(E_1)-\mu_1=d$ by 3.5.4. If s is the number of i.n. base points of Λ on E_2, we therefore have $s\leq 1$ by 3.5.8. On the other hand, $s\geq 1$ by 3.5.6. Suppose now there is a unique i.n. base point p_j of Λ on E_j and $p_j\in E_1^{(j)}, j=2,\cdots,r$. Then $\mu_r=(E_r, \Lambda^{(r)})$ by 2.1 (iv) and $(E_r, \Lambda^{(r)})=\mu_{r-1}$ by 2.1 (iii). Also $m(E_r)\leq m(E_2)\leqq 2d-2$. Hence $\mu_r=\mu_1$ and $\varepsilon(E_r)=d$ by induction on r, and we see as before that there is a unique base point p_{r+1} of Λ on E_{r+1}.

Let then n be the first integer such that p_n, the unique base point of Λ on E_n, is not on $E_1^{(n)}$. We note that

(2) $$n\geq 2,$$

(3) $$\mu_j=\mu_1,\quad j=2,\cdots,n,$$

(4) $$\varepsilon(E_n)=d,$$

(5) $\mu_0=(E_1, \Lambda^{(1)})=(n-1)\mu_1+\nu$, where $\nu\geq 0$ is the contribution to $(E_1, \Lambda^{(1)})$ arising from base points on E_1 other than p_1. In particular, $\mu_0\geq(n-1)\mu_1$.

We now show by induction on r that for $r<n-2$

(i) there is a unique base point p_{n+r+1} of Λ on E_{n+r+1}, and $p_{n+r+1}\notin E_j^{(n+r+1)}$ for $j\leq n+r$,

(ii) $\varepsilon(E_{n+r+1})=\mu_0-r\mu_1>0$

(iii) $\mu_{n+r+1}=\mu_1$.

In fact, this is true for $r=-1$. So let $-1<r<n-2$ and assume it is true for $r'<r$. Then $p_{n+r}\in E_j^{(n+r)}$ for $j=n+r$ only and $\varepsilon(E_{n+r+1})=\varepsilon(E_{n+r})-\mu_{n+r}=\mu_0-r\mu_1$, and by (5), $\mu_0-r\mu_1>0$. This proves (ii) for r. Now

$$\sum_{i=0}^{n+r}\mu_i+m(E_{n+r+1})\leq 3d-2$$

and $\sum_{i=0}^{n+r}\mu_i=\mu_0+(n+r)\mu_1$ by (3) and (iii). Hence if s is the number of i.n. base points of Λ on E_{n+r+1},

$$\mu_0+(n+r)\mu_1+s(\mu_0-r\mu_1)\leq 3d-2,$$

or, using (1), $(s-2)\mu_0+(n+r-sr-3)\mu_1\leq -2$. So if $s\geq 2$, we have, using (5), $s(n-r-1)-n+r-1\leq -(2/\mu_1)<0$. Since the left hand side is an integer, $s(n-r-1)\leq n-r$ and $s\leq (n-r)/(n-r-1)<2$ (since $n-r-1\geq 2$). Hence $s\geq 2$ leads to a contradiction and $s\leq 1$. On the other hand, $s\geq 1$ in view of (ii). We have

$$\mu_{n+r+1}=(E_{n+r+1}, \Lambda^{(n+r+1)})=\mu_{n+r},$$

repeating an earlier argument. Also, $p_{n+r}\notin E_j^{(n+r)}$ for $j<n+r$ implies $p_{n+r+1}\notin E_j^{(n+r+1)}$ for $j<n+r$, and by uniqueness of p_{n+r} as base point of Λ on E_{n+r}, $p_{n+r+1}\notin E_{n+r}^{(n+r+1)}$. This proves (i) and (iii). We note that $\Lambda^{(2n-1)}$ has base points only on $E_1^{(2n-1)}$ and E_{2n-1}, and no other $E_j^{(2n-1)}$. Hence

$$(E_j^{(2n-1)}, \Lambda^{(2n-1)})=0, \quad j=0,2,3,\cdots,2n-2. \tag{6}$$

In view of (5)

$$(E_1^{(2n-1)}, \Lambda^{(2n-1)})=\nu<\mu_0 \tag{7}$$

and, using (3) and (iii)

$$(\Lambda^{(2n-1)}, \Lambda^{(2n-1)})=d^2-\sum_{i=0}^{2n-2}\mu_i^2=d^2-(\mu_0^2+(2n-2)\mu_1^2). \tag{8}$$

We have arrived at a sequence of l.q.t. as required to apply 4. 4. Let φ be as in 4. 4 and put $\psi=\varphi^{-1}$. Then $\varphi(V(f))=V(\psi^*(f))$, and if $\tilde{\Lambda}=\Lambda(\psi^*(f))$ (considered as pencil on $\tilde{S}$), then $\Lambda^{(2n-1)}=\tilde{\Lambda}^{(2n-1)}$. Put $\tilde{d}=\deg\psi^*(f)$. Then

$$\tilde{d}^2=(\tilde{\Lambda}, \tilde{\Lambda})=(\Lambda^{(2n-1)}, \Lambda^{(2n-1)})+\nu^2 \tag{9}$$

in view of (6), (7) and 4. 3 (the salient point there is that E_1 is the last curve to be shrunk under $\tilde{\pi}$). Finally, combining (8) and (9), we conclude that $\tilde{d}^2<d^2$, and the degree of f has been reduced by an elementary automorphism. We can continue until either $\deg f=1$ or f has two (ordinary) points at infinity.

If k is not algebraically closed, we repeat the preceding argument over an algebraic closure $\bar{k}$ of k. The ordinary base points of Λ on

E_0 are rational over k by 3.8. If there is only one, p_0, we have inductively $p_1=E_0^{(1)}\cap E_1$, $p_2=E^{(2)}\cap E_2$, $\cdots$, $p_{n-1}=E_1^{(n-1)}\cap E_{n-1}$ rational over k. Also, since over $\bar{k}$ there is a unique base point p_{n+r} on E_{n+r}, $r=0$, $\cdots$, $n-2$, p_{n+r} is purely inseparable over k. Then the argument given at the end of the proof of 3.8 shows that p_{n+r} is rational over k. Hence φ has coefficients in k.

4.6 Corollary (*Jung, Van der Kulk*): *Every automorphism φ of A_k^2 is tame.*

Proof: $\varphi^*(x)$ is a field generator and $V(\varphi^*(x))\simeq V(x)=A_k^1$ has only one (ordinary) point at infinity. By the theorem, we can find a tame automorphism ψ such that $\deg\psi^*(\varphi^*(x))=1$. So we may assume, applying a linear automorphism, that $\psi^*(\varphi^*(x))=x$. Then $\rho=\varphi\circ\psi$ is elementary and $\varphi=\rho\circ\psi^{-1}$ is tame.

4.7 Corollary: *Let f be a field generator and suppose $V(f)\not\simeq A_k^2$. Then the automorphism ψ of 4.5 is unique up to linear automorphism and characterized by the fact that $\deg\psi^*(f)$ is minimal.*

Proof: If ψ is as constructed in 4.5, then $V(\psi^*(f))$ has two ordinary points at infinity, q_0 and q_1 say, and $\deg\psi^*(f)\leq\deg f$. Let φ be a non-linear automorphism of A_k^2. Then an initial segment p_0, $\cdots$, p_{2n-2} of the sequence of i.n. fundamental points of $\tilde{\varphi}_0$ will satisfy 4.2 (i) to (iv) for some n (since $\varphi^*(x),\varphi^*(y)$ are field generators, for instance). $\Lambda=\Lambda(\psi^*(f))$ is not tangent to E_0 at q_0 and q_1 by 3.7, and hence $\Lambda^{(2n-1)}$ meets $E_0^{(2n-1)}$ and possibly $E_1^{(2n-1)}$, but no other $E_j^{(2n-1)}$. But $\tilde{\varphi}_{2n-1}$ is a morphism on $E_0^{(2n-1)}$ and $E_1^{(2n-1)}$ and contracts $E_0^{(2n-1)}$ and $E_1^{(2n-1)}$ to the same point on $\tilde{S}$, and this is the only ordinary point at infinity of $\varphi(V(\psi^*(f)))$.

4.8 Corollary: Let f be a field generator and ψ as in 4.5. Then an irreducible factor of $\psi^*(f)$ has two ordinary points at infinity or is an irreducible polynomial in a linear form in x and y.

Proof: We may assume that $\psi^*(f)$ is not linear. Then $\psi^*(f)$ has two ordinary points at infinity, p and q say, which are rational over k by 3.8. Also, an irreducible factor g of $\psi^*(f)$ is not tangent to E_0 by 3.7, and if g does not pass through q, the multiplicity of p on g is equal to deg g. In that case g splits in $\bar{k}[x, y]$ into linear irreducible factors passing through p. Hence $g \in k[u]$, where u is the linear form in $k[x, y]$ vanishing at p.

4.9 Remarks: (i) An irreducible factor of a field generator is not, in general, a field generator.

(ii) Let $g \in k[x, y]$ such that $k[x, y]/g$ is a polynomial ring in one variable over k. If char $k=0$, Abhyankar and Moh [6] have shown that g is a ring generator, that is, ther exists $h \in k[x, y]$ such that $k[x, y]=k[g, h]$. It follows from 4.8 that the same is true without restriction on the characteristic if it is assumed that g is an irreducible factor of a field generator. This, of course, is a much weaker result than that of Abhyankar and Moh, but nevertheless has useful applications over fields of positive characteristic, where the stronger theorem fails.

McGILL UNIVERSITY

References

[1] H. W. E. Jung, Über ganze rationale Transformationen des Eberre, J. Reine Angew. Math. 184 (1942), 161-174.

[2] J.-P. Serre, Groupes algébriques et corps de classes, Herman, Paris, 1959.

[3] I. R. Schafarewitsch, Grundzüge der algebraischen Geometrie, Vieweg, Braunschweig, 1972.

[4] J. Tate, Genus change in inseparable extensions of function fields, Proc. Amer. Math. Soc. 3 (1952), 400-406.

[5] W. Van der Kulk, On polynomial rings in two variables, Nieuw. Arch. Wisk. (3) I (1953), 33-41.

[6] S. S. Abhyankar and T.-T. Moh, Embeddings of the line in the plane, to appear in J. reine angew. Math.