J. Math. Kyoto Univ. (JMKYAZ)
31-2 (1991) 329-358

Birational endomorphisms of the affine plane

By

D. DAIGLE*

Let A^2 be the affine plane over an algebraically closed field k. The birational endomorphisms of A^2 are those maps $A^2 \to A^2$ given by endomorphisms ϕ of the polynomial algebra $k[X, Y]$ such that $k(\phi(X), \phi(Y)) = k(X, Y)$. The simplest non-automorphic example that comes to mind is the ϕ_0 defined by $\phi_0(X) = X$ and $\phi_0(Y) = XY$. The birational endomorphisms of A^2 given by $\phi = u \circ \phi_0 \circ v$, where u and v are any automorphisms of $k[X, Y]$, are called *simple affine contractions in* A^2. The question whether every birational endomorphism of A^2 is a composite of simple affine contractions arose in the early seventies, in Abhyankar's seminar at Purdue University. That question was answered negatively by K.P. Russell who, in conversations with A. Lascu, constructed an irreducible birational endomorphism with three fundamental points. He soon exhibited a whole zoo of irreducible endomorphisms, some of them having infinitely near fundamental points. Its diversity shows that to give a reasonably complete *classification* of all birational endomorphisms of A^2 is likely to be interesting and difficult. The aim of this paper is to make some contributions to that problem.

The methods by which Russell constructed those endomorphisms and proved their irreducibility consist in a detailed analysis of the configuration of missing curves (see (1.2f) for definition). Our approach is essentially an elaboration of Russell's methods, and includes the use of some graph-theoretic techniques (weighted graphs, dual trees). Note that the last section of this paper is in fact an appendix which gathers some definitions and facts in the theory of weighted graphs.

The first three sections study birational morphisms $f: X \to Y$ of nonsingular surfaces, and the fourth section concentrates on the case $X = Y = A^2$. The most interesting results are, we think, those numbered (2.1), (2.9), (2.17), (4.3), (4.4), (4.11), (4.12) and (4.13). We point out that the last section of our paper [2] classifies the irreducible $f: A^2 \to A^2$ with two fundamental points.

Throughout this paper, our ground field is an arbitrary algebraically closed field k, all curves and surfaces are irreducible and reduced, all surfaces are nonsingular and the word "point" means "closed point". The domain (dom (f)) and codomain (codom (f)) of any birational morphism f under consideration will be tacitly assumed to be (nonsingular) surfaces. If X is a surface, Div (X) is its group of divisors and Cl (X) its divisor class group; "X is factorial" means that X is the affine spectrum of a U.F.D.; "X has

* During much of this work the author enjoyed the financial support of NSERC (Canada) and the hospitality of the University of Kentucky.
Communicated by Prof. Nagata.
Received May 16, 1988; Revised April 20, 1990

trivial units" means $\Gamma(X, \mathcal{O}_X)^* = \boldsymbol{k}^*$, where $\mathcal{O}_X$ is the structure sheaf, $\Gamma(X, \mathcal{O}_X)$ is the ring of global sections and "*" means "group of units of $\cdots$"; if $D \in \mathrm{Div}(X)$ and $\tilde{X} \to X$ is a monoidal transformation (resp. $X \hookrightarrow X'$ is an open immersion) then *the strict transform of D in $\tilde{X}$ (resp. the closure of D in X') is denoted by D* whenever no confusion seems likely to arise. $\boldsymbol{N}$, $\boldsymbol{Z}$ and $\boldsymbol{Q}$ denote respectively the sets of positive integers, integers and rational numbers.

This work grew out of our doctoral thesis. We would like to thank our professor, K.P. Russell, for having introduced us to these problems and for the numerous discussions we had about them.

1. Basic concepts

The following fact, in which the surfaces X and Y are not necessarily complete, is easily deduced from the basic properties of birational transformations of complete nonsingular surfaces.

Lemma 1.1. *Let $f: X \to Y$ be a birational morphism. Then there exists a commutative diagram*

$$\begin{array}{ccc} X & \hookrightarrow & Y_n \\ & & \downarrow \pi_n \\ f \downarrow & & \vdots \\ & & \downarrow \pi_1 \\ Y & = & Y_0 \end{array}$$

where $n \geqq 0$, $X \hookrightarrow Y_n$ is an open immersion and $\pi_i: Y_i \to Y_{i-1}$ is the blowing-up of Y_{i-1} at some point $(1 \leqq i \leqq n)$.

Two birational morphisms $f_1: X_1 \to Y_1$, $f_2: X_2 \to Y_2$ are said to be *equivalent* (write $f_1 \sim f_2$) if there are isomorphisms $x: X_1 \to X_2$, $y: Y_1 \to Y_2$ such that $f_1 = y^{-1} \circ f_2 \circ x$.

Definitions and Remarks 1.2. Let $f: X \to Y$ be a birational morphism.

(a) The least $n \geqq 0$ such that there exists a diagram as in (1.1) is denoted by $n(f)$. Clearly, $f \sim g \Rightarrow n(f) = n(g)$.

(b) A *fundamental point* of f is a point P of Y such that $f^{-1}(P)$ contains more than one point. By (1.1), there are at most $n(f)$ fundamental points.

Given a diagram as in (1.1) and $i > 0$, a fundamental point of $X \hookrightarrow Y_n \to \cdots \to Y_i$ which belongs to a curve that is contracted by $\pi_1 \circ \cdots \circ \pi_i$ is sometimes called an *infinitely near fundamental point of f*; such a point is not a fundamental point of f, according to definition (b). If f has no infinitely near (*i.n.*) fundamental points we say f has ordinary fundamental points; that is the case iff f has $n(f)$ distinct fundamental points in its codomain.

(c) Consider a diagram as in (1.1), where $n \geqq n(f)$. Let P_i be the center of π_i, let $E_i = \pi_i^{-1}(P_i)$ and let f_i be the composite $X \hookrightarrow Y_n \to \cdots \to Y_i$. Then the following are equivalent:

- $n = n(f)$;

- for $i=1, \cdots, n$, P_i is a fundamental point of f_{i-1};
- for $i=1, \cdots, n$, $E_i \cap X=\emptyset$ in $Y_n \Rightarrow E_i^2<-1$ in any nonsingular completion of Y_n (where E_i^2 is the self-intersection number of E_i).

(d) A *contracting curve* of f is a curve E in X such that $f(E)$ is a (fundamental) point. The number of contracting curves, which is an invariant of $\sim$, is denoted by $c(f)$.

(e) f is said to be *trivial* if it is an open immersion (iff $c(f)=0$, iff $n(f)=0$).

(f) The one dimensional irreducible components of the closure (in Y) of $Y \setminus f(X)$ are called the *missing curves* of f. The number of missing curves, which is an invariant of $\sim$, is denoted by $q(f)$. Given a curve C in Y, the following are equivalent:
- C is a missing curve;
- $C \cap f(X)$ is contained in the set of fundamental points;
- for some diagram as in (1.1) (equivalently for every such diagram) the strict transform of C in Y_n is disjoint from X.

(g) Let $q_0(f)$ denote the number of missing curves disjoint from $f(X)$. Clearly $q_0(f)$ is an invariant of $\sim$.

(h) A *minimal decomposition* of f is a diagram as in (1.1), with $n=n(f)$, together with an ordering of the set of missing curves (i.e., the missing curves are labelled $C_1, \cdots, C_q$ where $q=q(f) \geqq 0$). Minimal decompositions will be denoted by $\mathscr{D}$, $\mathscr{D}'$, etc. Each time we choose a minimal decomposition $\mathscr{D}$, the following notations are used:
- For the diagram, the notation is as in (1.1).
- The center of π_i is the point P_i of Y_{i-1} and the corresponding exceptional curve is E_i ($1 \leqq i \leqq n$).
- The missing curves are $C_1, \cdots, C_q$ where $q=q(f)$.
- $\mathscr{D}$ determines a subset $J=J_{\mathscr{D}}$ of $\{1, \cdots, n\}$, defined by

$$J=\{i \mid E_i \cap X=\emptyset \text{ in } Y_n\}.$$

Thus the curves of Y_n which are disjoint from X are precisely $C_1, \cdots, C_q$ and the E_i with $i \in J$. On the other hand, the contracting curves of f are the $E_i \cap X$ such that $i \in \{1, \cdots, n\} \setminus J$. We see that $|J|+c(f)=n(f)$, so $|J|$ is an invariant of $\sim$. That number will be denoted by $j(f)$. Hence

$$c(f)+j(f)=n(f).$$

- $\mathscr{D}$ determines a subset $\Delta=\Delta_{\mathscr{D}}$ of $\{1, \cdots, n\}$, defined by

$$\Delta=\{i \mid P_i \notin C_1 \cup \cdots \cup C_q \text{ in } Y_{i-1}\}.$$

The cardinality of Δ can be seen to be an invariant of $\sim$. We denote it by $\delta(f)$.

(i) By (c), if $i \in J$ then $E_i^2<-1$ in any nonsingular completion of Y_n.

(j) If $g: Y \to Z$ is a birational morphism, we denote by $\Delta c(f, g)$ the number of missing curves of f which are contracted by g.

Lemma 1.3. *Let* $f: X \to Y$ *and* $g: Y \to Z$ *be birational morphisms.*

(a) $c(g \circ f) = c(f) + c(g) - \Delta c(f, g)$ *and* $q(g \circ f) = q(f) + q(g) - \Delta c(f, g)$.

(b) $n(g \circ f) \leqq n(f) + n(g)$ *and* $j(g \circ f) \leqq j(f) + j(g) + \Delta c(f, g)$.

(c) *If* $q_0(f) = 0$ *then* $n(g \circ f) = n(f) + n(g)$ *and* $j(g \circ f) = j(f) + j(g) + \Delta c(f, g)$.

Proof. The verification of (a) is left to the reader. Choose a minimal decomposition of f and one of g, and consider the corresponding commutative diagram:

$$\begin{array}{ccccc} X & \hookrightarrow & Y_{n(f)} & \hookrightarrow & Z_n \\ f \big\downarrow & & \downarrow \vdots \downarrow & & \downarrow \vdots \downarrow \\ Y & = & Y_0 & \hookrightarrow & Z_{n(g)} \\ g \big\downarrow & & & & \downarrow \vdots \downarrow \\ Z & & = & & Z_0 \end{array}$$

where $\hookrightarrow$ means open immersion and $n = n(f) + n(g)$. By definition, $n(g \circ f) \leqq n$. The second inequality of (b) follows from this and (a), so (b) is clear. To prove (c), denote the center of $Z_i \to Z_{i-1}$ by P_i and let h_i be the composite $X \hookrightarrow Y_{n(f)} \hookrightarrow Z_n \to \cdots \to Z_i$. By (1.2c), it's enough to check that P_i is a fundamental point of h_{i-1} $(1 \leqq i \leqq n)$. If $n(g) < i \leqq n$ then that condition holds, by (1.2c) applied to the minimal decomposition of f. If $1 \leqq i \leqq n(g)$ then by (1.2c) P_i is a fundamental point of $Y_0 \hookrightarrow Z_{n(g)} \to \cdots \to Z_i$, so there is a curve Γ in Y whose image in Z_i is P_i. If $q_0(f) = 0$ then $f^{-1}(\Gamma)$ contains a curve, so P_i is a fundamental point of h_{i-1}. Hence $n(g. f) = n(f) + n(g)$, and the second equation follows from that and (a).

Remark 1.4. From the proof of (1.3), we see that if $f: X \to Y$ and $g: Y \to Z$ are birational morphisms and $q_0(f) = 0$ then each pair $(\mathcal{D}_f, \mathcal{D}_g)$ of minimal decompositions (of f, g respectively) determines a minimal decomposition $\mathcal{D}$ of $g \circ f$—the commutative diagram is as in the proof and the missing curves are labelled as follows: Let $\Gamma_1, \cdots, \Gamma_{q(f)}$ (resp. $C_1, \cdots, C_{q(g)}$) be the missing curves of f (resp. g) and let $\Gamma_{i_1}, \cdots, \Gamma_{i_k}$ be those missing curves of f which are not contracted by g, where $1 \leqq i_1 < \cdots < i_k \leqq q(f)$; if for $j = 1, \cdots, k$ we let Γ'_j be the closure in Z of $g(\Gamma_{i_j})$ then the missing curves of $g \circ f$ are $C_1, \cdots, C_{q(g)}, \Gamma'_1, \cdots, \Gamma'_k$, in that order.

Corollary 1.5. *Let* $f: X \to Y$ *and* $g: Y \to Z$ *be birational morphisms. Then*

$$c(g \circ f) - q(g \circ f) = (c(f) - q(f)) + (c(g) - q(g)),$$

i.e., the number $c - q$ *is "additive".*

2. Properties of the domain and codomain

In this section we study how the structure of a birational morphism is related to some properties of its domain and codomain. We consider properties possessed by $\boldsymbol{A}^2$, such as affineness, factoriality, the property of having trivial units, the property of having no loops at infinity, etc.

Proposition 2.1. *Let $f: X \to Y$ be a birational morphism, with missing curves $C_1, \cdots, C_q$ ($q \geqq 0$). Consider the following conditions:*
(a) *Y is affine, X is connected at infinity and no contracting curve of f is complete;*
(b) *X is affine;*
(c) *all fundamental points of f are in $C_1 \cup \cdots \cup C_q$ and the interior of $f(X)$ ($\mathrm{int}\, f(X)$) is $Y \setminus (C_1 \cup \cdots \cup C_q)$ and is affine.*
Then (a)$\Rightarrow$(b)$\Rightarrow$(c).

Corollary 2.2. *Let $f: X \to Y$ be a birational morphism and suppose that Y is affine. Then the following are equivalent:*
(a) *X is affine,*
(b) *X is connected at infinity and no contracting curve of f is complete.*

The main ingredients of the proof of (2.1) are the *Nakai-Moishezon Criterion* and the following, for which one can see [6], theorem 2, p. 168 or [7], theorem 4.2, p. 69:

Theorem 2.3 (Goodman). *Let U be an open subset of a complete nonsingular surface S. Then U is affine iff $S \setminus U$ is the support of an effective ample divisor of S.*

Before we prove the proposition, we find it convenient to define some terminologies and to state two facts. These considerations are elementary and may well exist, in one form or another, in the literature.

Let S be a complete nonsingular surface. For $D \in \mathrm{Div}(S)$, let the symbol $D \gg 0$ mean that D is effective, $D \neq 0$ and every irreducible component C of D satisfies $C \cdot D > 0$. Then the set $P(S)$ of divisors D such that $D \gg 0$ is a nonempty additive semi-group. Say that a subset Z of S is *positive* if $Z = \mathrm{supp}(D)$ for some $D \gg 0$. Then the set of positive subsets of S is stable under finite unions.

Lemma 2.4. *Let S be a complete nonsingular surface and Z a subset of S. Then the following are equivalent:*
(a) *Z is positive;*
(b) *Z is closed, $Z \neq \emptyset$ and every connected component of Z is positive;*
(c) *Z is closed $\emptyset \neq Z \neq S$ and every connected component of Z contains a positive set.*

Lemma 2.5. *Let $\pi: \tilde{S} \to S$ be the blowing-up of a nonsingular complete surface S at some point. Then:*
(a) *If Z is a positive subset of $\tilde{S}$ then $\pi(Z)$ is a positive subset of S;*
(b) *If $Z \subseteq S$, then Z is positive in S iff $\pi^{-1}(Z)$ is positive in $\tilde{S}$.*

Proof of (2.1.). Assume that (a) or (b) holds. Choose a minimal decomposition of f, with notation as in (1.2h), imbed Y_0 in a complete nonsingular surface $\bar{Y}_0$ and "complete the diagram":

$$\begin{array}{ccccc} X & \hookrightarrow & Y_n & \hookrightarrow & \bar{Y}_n \\ \Big| & & \downarrow \pi_n & & \downarrow \bar{\pi}_n \\ f \Big| & & \vdots & & \vdots \\ \downarrow & & \downarrow \pi_1 & & \downarrow \bar{\pi}_1 \\ Y & = & Y_0 & \hookrightarrow & \bar{Y}_0 \end{array}$$

where $\bar{\pi}_i$ is the blowing-up of $\bar{Y}_{i-1}$ at P_i $(1 \leqq i \leqq n)$. Then $\bar{Y}_n \setminus X$ is connected and contains a curve, hence is a nonempty union of curves. So $Y_n \setminus X$ is a (possibly empty) union of curves, i.e.,

$$Y_n \setminus X = C_1 \cup \cdots \cup C_q \cup \bigcup_{j \in J} E_j$$

$$\bar{Y}_n \setminus X = \bar{C}_1 \cup \cdots \cup \bar{C}_q \cup \bigcup_{j \in J} E_j \cup L_1 \cup \cdots \cup L_p$$

where $\bar{C}_i$ is either the closure of C_i in $\bar{Y}_0$ or its strict transform in $\bar{Y}_j$, and where $L_1, \cdots, L_p$ are the one-dimensional irreducible components of $\bar{Y}_j \setminus Y_j$ (for any $j=0, \cdots, n$). Let

$$\Lambda_0 = L_1 \cup \cdots \cup L_p \quad \text{in } \bar{Y}_0, \qquad \Lambda_n = L_1 \cup \cdots \cup L_p \quad \text{in } \bar{Y}_n,$$

$$\Gamma_0 = \bar{C}_1 \cup \cdots \cup \bar{C}_q \quad \text{in } \bar{Y}_0, \qquad \Gamma_n = \bar{C}_1 \cup \cdots \cup \bar{C}_q \quad \text{in } \bar{Y}_n,$$

$$Z_n = \Gamma_n \cup \bigcup_{j \in J} E_j \cup \Lambda_n \quad \text{in } \bar{Y}_n,$$

denote by F the set of fundamental points of f and let $\pi = \pi_1 \circ \cdots \circ \pi_n$ and $\bar{\pi} = \bar{\pi}_1 \circ \cdots \circ \bar{\pi}_n$.

CLAIMS

(1) $\bar{Y}_n \setminus X = Z_n$ is connected and $\bar{Y}_0 \setminus Y_0 = \Lambda_0 \cup$ points,

(2) $F \subseteq \Gamma_0$ and $\bar{\pi}^{-1}(F) = E_1 \cup \cdots \cup E_n$,

(3) $\bar{\pi}^{-1}(\Lambda_0) = \Lambda_n$,

(4) $\operatorname{int} f(X) = Y_0 \setminus \Gamma_0$.

We verify that $F \subseteq \Gamma_0$ and leave the rest to the reader. If $a \in F$ then $\pi^{-1}(a)$ can't contain Z_n (indeed, suppose $Z_n \subseteq \pi^{-1}(a)$ then $Z_n = \bigcup_{j \in J} E_j$ and $p=q=0$; in particular $\bar{Y}_0 \setminus Y_0$ contains no curve, so Y_0 is not affine; since (a) or (b) holds by assumption, X must be affine, so Z_n is positive by (2.3) and the Nakai-Moishezon Criterion, so is $\pi(Z_n) = \{a\}$ by (2.5) and this is absurd) and $Z_n \cap \pi^{-1}(a) \neq \emptyset$ because no contracting curve of f is complete. Thus there is an irreducible component C of Z_n such that $\emptyset \neq C \cap \pi^{-1}(a) \neq C$, by connectedness of Z_n. Clearly, $C \subseteq \Gamma_n$, so $a \in \Gamma_0$ and $F \subseteq Y_0$.

Proof of (a)⇒(b). If (a) holds then $\bar{Y}_0 \setminus Y_0 = \Lambda_0$ and Λ_0 is positive, by (2.3) and Nakai-Moishezon. Hence Λ_n is positive, by (3) and (2.5b), and Z_n is positive by connectedness of Z_n and (2.4). Let $D \in P(\bar{Y}_n)$ be such that $Z_n = \operatorname{supp}(D)$; since a straightforward argument shows that Z_n meets every curve in $\bar{Y}_n$, D is ample by Nakai-Moishezon and X is affine by (1) and (2.3). Hence (b) holds.

Proof of (b)⇒(c). Statements (2) and (4) show that f restricts to an isomorphism

$$f^{-1}(\operatorname{int} f(X)) \longrightarrow \operatorname{int} f(X),$$

and that $f^{-1}(\operatorname{int} f(X)) = X \setminus (E_1 \cup \cdots \cup E_n)$, which is just the open set obtained by removing the contracting curves from X. But if (b) holds then X is affine, thus so is X

minus the contracting curves, since removing a curve from an affine nonsingular surface yields an affine surface. Hence we are done.

The next properties (for a surface) that will interest us are the property of having a trivial divisor class group and the property of having trivial units. To begin with, we recall a well-known fact:

Lemma 2.6. *Let V be a complete nonsingular algebraic variety and $U\neq\emptyset$ an open subset of V. Among the irreducible components of $V\setminus U$, let $\Gamma_1, \cdots, \Gamma_r$, $(r\geqq 0)$ be those of codimension one in V, and let $\bar{\Gamma}_1, \cdots, \bar{\Gamma}_r$ be their images in* $\mathrm{Cl}(V)$.
(a) $\mathrm{Cl}(U)=0\Leftrightarrow\bar{\Gamma}_1, \cdots, \bar{\Gamma}_r$ *generate* $\mathrm{Cl}(V)$.
(b) $\Gamma(U, \mathcal{O}_U)^*=k^*\Leftrightarrow\bar{\Gamma}_1, \cdots, \bar{\Gamma}_r$ *are linearly independent.*

2.7. Let Y_0 be any nonsingular surface and consider

$$Y_n \xrightarrow{\pi_n} Y_{n-1} \longrightarrow \cdots \xrightarrow{\pi_1} Y_0 \qquad (n\geqq 1)$$

where $\pi_i: Y_i\to Y_{i-1}$ is the blowing-up of Y_{i-1} at some point P_i and let $E_i=\pi_i^{-1}(P_i)\in \mathrm{Div}(Y_i)$ $(1\leqq i\leqq n)$. Given integers i, ν such that $1\leqq i\leqq n$ and $0\leqq\nu\leqq n$ and given $D\in \mathrm{Div}(Y_\nu)$ we define $\mu(P_i, D)$ to be the multiplicity of P_i on the appropriate strict transform of D if $i-1\geqq\nu$, and we define it to be zero if $i-1<\nu$. Then we define

$$\mu(D)=\begin{bmatrix}\mu(P_1, D)\\ \vdots \\ \mu(P_n, D)\end{bmatrix}\in \boldsymbol{Z}^n$$

and we have the following $n\times n$ matrix:

$$\mathcal{E}=(e_{ij})=(\mu(E_1)\cdots\mu(E_n))$$

where, of course, $e_{ij}=0$ whenever $i\leqq j$. If R_i is the i^{th} row of the identity matrix I_n, define an $n\times n$ matrix $\varepsilon=(\varepsilon_{ij})$ by letting the first row be R_1 and

$$(\varepsilon_{k1}\cdots\varepsilon_{kn})=R_k+(e_{k1}\cdots e_{kk-1})(\varepsilon_{ij})_{\substack{1\leq i<k\\ 1\leq j\leq n}} \qquad (1<k\leqq n).$$

So ε is completely determined by $\mathcal{E}$, is a lower triangular matrix with $\varepsilon_{ii}=1$ $(1\leqq i\leqq n)$ and has $\det(\varepsilon)=1$. For $1\leqq i\leqq n$, define

$$\varepsilon_i: \boldsymbol{Z}^n \longrightarrow \boldsymbol{Z}$$

$$\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\longmapsto(\varepsilon_{i1}\cdots\varepsilon_{in})\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}.$$

Let $D^*\in\mathrm{Div}(Y_n)$ be the total transform of $D\in\mathrm{Div}(Y_\nu)$. One shows that, if we define $\varepsilon_i(D)=\varepsilon_i(\mu(D))\in\boldsymbol{Z}$, then

$$D^*=D+\sum_{i=1}^{n}\varepsilon_i(D)E_i \qquad (\text{in } Y_n).$$

Next, one checks that

$$\theta: \mathrm{Cl}(Y_0)\oplus\boldsymbol{Z}^n\longrightarrow\mathrm{Cl}(Y_n)$$

$$\left(\bar{D}, \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}\right) \longmapsto \bar{D}^* + \sum_{i=1}^{n} a_i \bar{E}_i$$

is an isomorphism (where $D^* \in \mathrm{Div}(Y_n)$ is the total transform of $D \in \mathrm{Div}(Y_0)$). By the above calculation, one sees that if $D \in \mathrm{Div}(Y_0)$ and if the strict transform of D in Y_n is also denoted by D, then

$$\theta^{-1}(\bar{D}) = (\bar{D}, -\varepsilon\mu(D))$$

(where the D's on the right hand side are in $\mathrm{Div}(Y_0)$). Clearly, $\theta^{-1}(\bar{E}_i) = (0, K_i)$, $1 \leqq i \leqq n$, where K_i denotes the i^{th} column of the identity matrix I_n.

Definition 2.8. Let $f: X \to Y$ be a birational morphism and write $n = n(f)$, $c = c(f)$ and $q = q(f)$. Let $\mathscr{D}$ be a minimal decomposition for f, with notation as in (1.2h). Then we define the following matrices:

$$\mu = \mu_{\mathscr{D}} = (\mu(C_1) \cdots \mu(C_q)) \qquad (n \times q)$$

$$\mathcal{E} = \mathcal{E}_{\mathscr{D}} = (\mu(E_1) \cdots \mu(E_n)) \qquad (n \times n)$$

$$\varepsilon = \varepsilon_{\mathscr{D}} = (\varepsilon_{ij}) \qquad (n \times n) \qquad \text{defined as in (2.7)},$$

and we let $\varepsilon' = \varepsilon'_{\mathscr{D}}$ be the $c \times n$ sub-matrix of ε obtained by deleting the i^{th} row whenever $i \in J$.

Observe that the product $\varepsilon'\mu$ is a $c \times q$ matrix; its q columns will be regarded as elements of $\boldsymbol{Z}^c$, even if $c = 0$ or $q = 0$. To make sense out of these extreme cases, let us *agree* that (1) the columns of a $0 \times q$ matrix generate $\boldsymbol{Z}^0$, and are linearly independent iff $q = 0$; (2) the columns of a $c \times 0$ matrix are linearly independent, and generate $\boldsymbol{Z}^c$ iff $c = 0$; and (3) the 0×0 matrix has determinant equal to 1. With these conventions, we have:

Proposition 2.9. *Let $f: X \to Y$ be a birational morphism and $\mathscr{D}$ a minimal decomposition; let the notation be as in* (2.8), *let $j = j(f)$ and $\delta = \delta(f)$.*

(a) *If* $\mathrm{Cl}(X) = 0$, *then the columns of $\varepsilon'\mu$ generate $\boldsymbol{Z}^c$,* $\mathrm{Cl}(\mathrm{int}\, f(X)) = 0$, *$q \geqq c$ and $\delta \leqq j$.*

(b) *If* $\mathrm{Cl}(Y) = 0$ *and the columns of $\varepsilon'\mu$ generate $\boldsymbol{Z}^c$, then* $\mathrm{Cl}(X) = 0$.

(c) *Consider the statements:*

(1) $\Gamma(X, \mathcal{O}_X)^* = k^*$,

(2) $\Gamma(Y, \mathcal{O}_Y)^* = k^*$,

(3) *the columns of $\varepsilon'\mu$ are linearly independent,*

(4) $\mathrm{Cl}(Y) = 0$.

Then $(1) \wedge (4) \Rightarrow (2) \wedge (3) \Rightarrow (1) \Rightarrow (2)$, *and* (3) *implies $q \leqq c$ and $\delta \leqq n - q$.*

(d) *Suppose that* $\mathrm{Cl}(X) = 0$ *and $\Gamma(X, \mathcal{O}_X)^* = \boldsymbol{k}^*$. Then $\Gamma(Y, \mathcal{O}_Y)^* = \boldsymbol{k}^*$ and $q \geqq c$, with equality iff* $\mathrm{Cl}(Y) = 0$.

Proof. Consider the minimal decomposition $\mathscr{D}$, with notation as usual. Imbed Y_0 in a complete nonsingular surface $\bar{Y}_0$ and "complete the diagram" (refer to the diagram in the proof of (2.1)). Let the closures (in $\bar{Y}_0$) of the missing curves be denoted by

$C_1, \cdots, C_q$, let the one-dimensional irreducible components of $\bar{Y}_0 \setminus Y_0$ be denoted by $L_1, \cdots, L_p$ and recall that the same notation is used for a divisor of some $\bar{Y}_i$ and its strict transform in $\bar{Y}_j$ $(j>i)$. We have

$$\bar{Y}_0 \setminus Y_0 = L_1 \cup \cdots \cup L_p \cup \text{ points}$$

$$\bar{Y}_0 \setminus \operatorname{int} f(X) = C_1 \cup \cdots \cup C_q \cup L_1 \cup \cdots \cup L_p \cup \text{ points}$$

$$\bar{Y}_n / X = \bigcup_{j \in J} E_j \cup C_1 \cup \cdots \cup C_q \cup L_1 \cup \cdots \cup L_p \cup \text{ points.}$$

Given $D \in \operatorname{Div}(\bar{Y}_i)$, let $\bar{D}$ be its image in $\operatorname{Cl}(\bar{Y}_i)$. Let $\theta : \operatorname{Cl}(\bar{Y}_0) \oplus \boldsymbol{Z}^n \to \operatorname{Cl}(\bar{Y}_n)$ be the isomorphism given in (2.7). Then

$$\theta^{-1}(\bar{L}_j) = (\bar{L}_j, -\varepsilon\mu(L_j)) = (\bar{L}_j, 0)$$

$$\theta^{-1}(\bar{C}_j) = (\bar{C}_j, -\varepsilon\mu(C_j))$$

$$\theta^{-1}(\bar{E}_j) = (0, K_j).$$

In view of that, and by (2.6), we find

(α) $\operatorname{Cl}(Y)=0$ (resp. Y has trivial units) iff $\bar{L}_1, \cdots, \bar{L}_p$ generate (resp. are linearly independent in) $\operatorname{Cl}(\bar{Y}_0)$;

(β) $\operatorname{Cl}(\operatorname{int} f(X))=0$ iff $\bar{L}_1, \cdots, \bar{L}_p, \bar{C}_1, \cdots, \bar{C}_q$ generate $\operatorname{Cl}(\bar{Y}_0)$;

(γ) $\operatorname{Cl}(X)=0$ (resp. X has trivial units) iff the set

$$\{(0, K_j) \mid j \in J\} \cup \{(\bar{C}_j, -\varepsilon\mu(C_j)) \mid 1 \leqq j \leqq q\} \cup \{(\bar{L}_j, 0) \mid 1 \leqq j \leqq p\}$$

generates (resp. is linearly independent in) the group $\operatorname{Cl}(\bar{Y}_0) \oplus \boldsymbol{Z}^n$.

On the other hand, it is clear that

(δ) $\{K_j \mid j \in J\} \cup \{-\varepsilon\mu(C_j) \mid 1 \leqq j \leqq q\}$ generates (resp. is linearly independent in) $\boldsymbol{Z}^n$ iff the columns of $\varepsilon'\mu$ generate (resp. are linearly independent in) $\boldsymbol{Z}^c$.

Now the reader can verify that, except for the inequalities $\delta \leqq j$ and $\delta \leqq n-q$, the assertions (a)-(c) of the proposition are immediate consequences of (α)-(δ). To prove the two inequalities, observe that δ is the number of zero rows in μ. Let U be the $n-\delta \times q$ sub-matrix of μ obtained by deleting the zero rows; let V be the $c \times n-\delta$ sub-matrix of ε' obtained by deleting the i^{th} column whenever the i^{th} row of μ is zero. Clearly, $VU = \varepsilon'\mu$. The matrices U, V and $VU = \varepsilon'\mu$ determine a commutative diagram of $\boldsymbol{Z}$-linear maps:

$$\begin{array}{ccccc} & & \boldsymbol{Z}^{n-\delta} & & \\ & \overset{u}{\nearrow} & & \overset{v}{\searrow} & \\ \boldsymbol{Z}^q & & \xrightarrow[w]{\hspace{3em}} & & \boldsymbol{Z}^c \end{array}$$

If the columns of $\varepsilon'\mu$ generate $\boldsymbol{Z}^c$, i.e., w is onto, then v is onto and $\delta \leqq n-c=j$. If the columns of $\varepsilon'\mu$ are linearly independent, i.e., w injective, then u is injective and $\delta \leqq n-q$.

We now prove (d). By parts (a) and (c), Y has trivial units and $q \geqq c$, with equality whenever $\operatorname{Cl}(Y)=0$. Conversely, suppose that $c=q$. Let $G = \operatorname{Cl}(\bar{Y}_0) \subseteqq \operatorname{Cl}(\bar{Y}_0) \oplus \boldsymbol{Z}^n$ and $g_i = (\bar{L}_i, 0) \in G$ $(1 \leqq i \leqq p)$. Since $n = c+j = q+j$, there are elements $e_1, \cdots, e_n$ in $\operatorname{Cl}(\bar{Y}_0)$

$\oplus Z^n$ such that $(g_1, \cdots, g_p, e_1, \cdots, e_n)$ is a basis of $\mathrm{Cl}(\bar{Y}_0)\oplus Z^n$. By elementary algebra, it follows that $(g_1, \cdots, g_p)$ is a basis of G, i.e., $(\bar{L}_1, \cdots, \bar{L}_p)$ is a basis of $\mathrm{Cl}(\bar{Y}_0)$, so $\mathrm{Cl}(Y)=0$.

Corollary 2.10. *Let* $f: X\to Y$ *be a birational morphism and suppose that* $\mathrm{Cl}(Y)=0$ *and* $\Gamma(Y, \mathcal{O}_Y)^*=k^*$. *Then*
(a) $\mathrm{Cl}(X)=0$ *iff the columns of* $\varepsilon'\mu$ *generate* Z^c;
(b) $\Gamma(X, \mathcal{O}_X)^*=k^*$ *iff the columns of* $\varepsilon'\mu$ *are linearly independent;*
(c) $\mathrm{Cl}(X)=0$ *and* $\Gamma(X, \mathcal{O}_X)^*=k^*$ *iff* $\varepsilon'\mu$ *is a square matrix with determinant* ± 1.

Remark. If we restrict ourselves to the case $j(f)=0$ then $\varepsilon'=\varepsilon$ and consequently (2.9) and (2.10) are still true when all "$\varepsilon'\mu$" are replaced by "μ".

Corollary 2.11. *Let* $f: X\to Y$ *be a birational morphism and supose that* $\Gamma(X, \mathcal{O}_X)^*=k^*$ *and* $\mathrm{Cl}(Y)=0$. *Then* $q_0(f)=0$.

Proof. $q_0(f)$ is the number of zero columns in μ. Since the columns of $\varepsilon'\mu$ are linearly independent by (2.9), $q_0(f)=0$.

Corollary 2.12. *Let* $f: X\to Y$ *and* $g: Y\to Z$ *be birational morphisms and suppose that* X, Y *and* Z *have trivial divisor class groups and trivial units. Then* $n(gf)=n(f)+n(g)$.

Proof. Immediate from (2.11) and (1.3).

Remark 2.13. Let S be a nonsingular complete surface and $U\neq\emptyset$ an open subset of S. If m is the number of curves in $S\setminus U$ and K_S is a canonical divisor of S then, clearly, the number $m+K_S^2$ is an invariant of U up to isomorphism. Let us temporarily denote that number by $\alpha(U)$. Then an easy argument shows that, if $f: X\to Y$ is a birational morphism of nonsingular surfaces, then $c(f)-q(f)=\alpha(Y)-\alpha(X)$. That gives us another proof of (1.5) and, on the other hand, shows that $c(f)=q(f)$ whenever $X=Y$. Hence $\varepsilon'\mu$ is a square matrix whenever $X=Y$, but examples show that its determinant needs not be ± 1. (By (2.9a), it is ± 1 whenever $\mathrm{Cl}(X)=0$.)

See (5.7) for the definition of $\mathcal{G}[U]$, where U is a nonsingular surface.

Definition 2.14. Let U be a nonsingular surface. We say that U *has no loops at infinity* (resp. *is linear at infinity*) if, in the equivalence class $\mathcal{G}[U]$, no graph has loops (resp. some graph is a linear tree).

Let us also say that U is *rational at infinity* if for some (equivalently, for every) open immersion $U\hookrightarrow\bar{U}$ such that $\bar{U}$ is a complete nonsingular surface, all curves in $\bar{U}\setminus U$ are rational.

Definition 2.15. Let Γ be a (not necessarely complete) curve. Let $\tilde{\Gamma}$ be the complete nonsingular model of Γ (i.e., the set of valuation rings of the function field of Γ over the ground field) and let $\tau: \tilde{\Gamma}\to\Gamma$ be the canonical birational transformation.

Then $\tilde{\Gamma}\setminus\mathrm{dom}(\tau)$ is a finite set of closed points, called *the places of Γ at infinity*. We denote the cardinality of $\tilde{\Gamma}\setminus\mathrm{dom}(\tau)$ by $P_\infty(\Gamma)$.

Lemma 2.16. *Let $f: X\to Y$ be a birational morphism.*

(a) *X is rational at infinity iff Y is rational ot infinity and all missing curves are rational.*

(b) *If X has no loops at infinity then Y has no loops at infinity.*

(c) *Suppose X has no loops at infinity and $\mathcal{D}$ is a minimal decomposition of f with notation as usual, embed Y_n in a complete nonsingular surface $\bar{Y}_n$ and let C be the closure (in $\bar{Y}_n$) of the strict transform of a missing curve. If $\tilde{C}$ is the complete nonsingular model of C then the canonical epimorphism $\tau: \tilde{C}\to C$ is bijective.*

(d) *If X has no loops at infinity and Y has $k>0$ connected components at infinity (i.e., an arbitrary member of $\mathcal{G}[Y]$ has k connected components), then*

$$\sum_{i=1}^{q} P_\infty(C_i)\leqq k+q-1\,,$$

where $C_1, \cdots, C_n$ are the missing curves of f. In particular, if Y is affine then each missing curve has exactly one place at infinity.

Proof. Most of these facts are trivial observations. Let's prove (d). Choose a smooth completion $Y\hookrightarrow\bar{Y}$ of Y (see (5.7)) and consider the graph $\mathcal{G}=(G, R)$ given by $G=\{\bar{C}_1, \cdots, \bar{C}_q, A_1, \cdots, A_k\}$, where $\bar{C}_i$ is the closure of C_i in $\bar{Y}$ and $A_1, \cdots, A_k$ are the connected components of $\bar{Y}\setminus Y$, and $R=\{\{\bar{C}_i, A_j\}\,|\,\bar{C}_i\cap A_j\neq\emptyset\}$. Since X has no loops at infinity we see that $\mathcal{G}$ doesn't have loops and that each $\bar{C}_i$ belongs to exactly $P_\infty(C_i)$ links. Thus $|R|=\sum_{i=1}^{q}P_\infty(C_i)$. On the other hand, it is a general fact that a graph $\mathcal{G}$ with no loops has at most $|\mathcal{G}|-1$ links. Hence we get the desired inequality.

See (5.5) for the notion of *strong normal crossings* (*s.n.c*).

Lemma 2.17. *Let $f: X\to Y$ be a birational morphism, where X is linear at infinity and Y is affine. Consider a minimal decomposition of f, with notation as in* (1.2h). *Then $Y_n\setminus X$ has $q=q(f)$ connected components, each one forming a linear tree*

$$C_i - E_{j1} - E_{j2} - \cdots E_{jk}\,,$$

where $\{j_1, \cdots, j_k\}\subseteqq J$ and $C_i+E_{j_1}+\cdots+E_{j_k}$ has s.n.c. in Y_n. In particular, the strict transforms of the missing curves on Y_n are nonsingular.

Proof. Since, in Y_0, each C_i has one place at infinity by (2.16d), we can choose a smooth completion $Y_0\hookrightarrow\bar{Y}_0$ of Y_0 such that, if L is the divisor of $\bar{Y}_0$ with *s.n.c.* and which satisfies $\bar{Y}_0\setminus Y_0=\mathrm{supp}(L)$, and if $C_1, \cdots, C_q$ also denote the closures of the missing curves, then $C_1, \cdots, C_q$ meet L at distinct points and $C_i\cdot L=1$ $(1\leqq i\leqq q)$. Form the diagram which appears in the proof of (2.1). Then $\bar{Y}_n\setminus Y_n=\mathrm{supp}(L)$, and (in $\bar{Y}_n$) $C_1, \cdots, C_q$ meet L at distinct points and $C_i.L=1$ $(1\leqq i\leqq q)$. Since X has no loops at infinity and L is connected, $C_1, \cdots, C_q$ belong to distinct connected components of $Y_n\setminus X$. On the other hand, if W is a connected component of $Y_n\setminus X$ and $\bar{W}$ is its closure in $\overline{Y}_n$, then $\overline{W}$ meets L, since X is connected at infinity; hence $\overline{W}$ contains a C_i, and there are exactly q connected components of $Y_n \setminus X$. We now show (by contradiction) that each one of these connected components has the desired properties. Let

$$D=\sum_{i=1}^{q} C_i+\sum_{i\in J} E_i+L\in \mathrm{Div}(\overline{Y}_n).$$

First, suppose that D does not have $s.n.c.$. By (5.19) we can consider a sequence of monoidal transformations $\overline{Y}_m\to\cdots\to\overline{Y}_n$ $(m>n)$ such that, if E_i is the exceptional curve created by $\overline{Y}_i\to\overline{Y}_{i-1}$ and

$$\begin{cases} D^n=D\in \mathrm{Div}(\overline{Y}_n), \\ D^i=(\text{strict transform of } D^{i-1})+E_i\in \mathrm{Div}(Y_i), \qquad n<i\leqq m, \end{cases}$$

then $D^m\in \mathrm{Div}(\overline{Y}_m)$ has $s.n.c.$, all centers are $i.n.$ $\mathrm{supp}(D)\cap Y_n$, $\overline{Y}_m\setminus \mathrm{supp}(D^m)\cong X$ and if $n<i\leqq m$ then

$$(*) \qquad E_i^2=-1 \text{ in } \overline{Y}_m \Longrightarrow E_i \text{ is a branch point of } \mathcal{G}_m=\mathcal{G}(\overline{Y}_m, D^m).$$

Let $\mathcal{G}_+$ be the connected subtree of $\mathcal{G}_m$ which has $C_1, \cdots, C_q$ and the irreducible components of L as vertices. Let Σ be the set of branch points v of $\mathcal{G}_m$ such that v is not in $\mathcal{G}_+$. By (*), $E_m\in\Sigma$ so $\Sigma\neq\emptyset$. If $v\in\Sigma$, then let B_v be the branch of $\mathcal{G}_m$ at v such that B_v contains $\mathcal{G}_+$. Since $\mathcal{G}_m$ is a finite tree, we can find $v\in\Sigma$ such that, if $B_v, B_1, \cdots, B_k$ are the distinct branches of $\mathcal{G}_m$ at v $(\Rightarrow k\geqq 2)$ then $\Sigma\cap(B_1\cup\cdots\cup B_k)=\emptyset$. By (*) and (1.2i), $B_i<-1$, $1\leqq i\leqq k$ (see (5.17)). Since X is linear at infinity, v can "absorb" a branch (see (5.11)); since no B_i can be absorbed, v absorbs B_v, whence $B_v\sim[-1]$. See (5.8) for the definition of $\langle\ \rangle$ and note that if $\mathcal{G}_1\subseteqq\mathcal{G}_2$ (weighted graphs) then $\langle\mathcal{G}_1\rangle\leqq\langle\mathcal{G}_2\rangle$. Whence

$$\langle\mathcal{G}(\overline{Y}_m, L)\rangle\leqq\langle B_v\rangle=\langle[-1]\rangle=0.$$

On the other hand, $\langle\mathcal{G}(\overline{Y}_0, L)\rangle>0$ since Y_0 is affine—in the terminology defined just after (2.3), $\mathrm{supp}(L)$ is a positive subset of $\overline{Y}_0$. Moreover, $\mathcal{G}(\overline{Y}_0, L)$ is just the same as $\mathcal{G}(\overline{Y}_m, L)$ since no blowing-up has center $i.n.$ L. Hence

$$\langle\mathcal{G}(\overline{Y}_m, L)\rangle>0,$$

contradiction. So $D\in \mathrm{Div}(Y_n)$ has $s.n.c.$.

Next, suppose that some connected component W of $Y_n\setminus X$ does not have the desired form; it means that either the dual tree $\mathcal{G}(\overline{Y}_n, F)$ is not linear or C_i is not a free vertex of it, where

$$F=C_i+E_{j_1}+\cdots+E_{j_k}\in \mathrm{Div}(\overline{Y}_n)$$

is the divisor (with $s.n.c.$) whose support is the closure $\overline{W}$ of W in $\overline{Y}_n$. In the first case, let v be a branch point of $\mathcal{G}(\overline{Y}_n, F)$; in the second case, let $v=C_i$. Let $B_v, B_1, \cdots, B_k$ be the distinct branches of $\mathcal{G}=\mathcal{G}(\overline{Y}_n, D)$ at v $(\Rightarrow k\geqq 2)$, where B_v is the one that contains the components of L. By (1.2i) we see that $B_i<-1$ $(1\leqq i\leqq k)$. As above, we see that v "absorbs" B_v and a contradiction follows.

Kodaira dimensions. We denote by $\bar{\kappa}(V)$ the logarithmic Kodaira dimension of a nonsingular surface V (see [7] or [13]). If C is a curve on a nonsingular surface V, we denote by $\kappa(C)$ the Kodaira dimension of the embedding $C \subset V$, in the sense of [4]. The following fact can be found in [7] or [13], where it is stated for a dominant separable morphism f.

2.18. *If* $f: X \to Y$ *is a birational morphism then* $\bar{\kappa}(Y) \leqq \bar{\kappa}(X)$.

We also point out

2.19. *If* $f: X \to Y$ *is a birational morphism and* C *is a missing curve of* f *then* $\kappa(C) \leqq \bar{\kappa}(X)$.

Proof. Choose a minimal decomposition for f (notation as usual) and embed Y_0 in a complete nonsingular surface $\bar{Y}_0$. Consider the diagram in the proof of (2.1) and, if necessary, blow-up $\bar{Y}_n$ at points of C until C is nonsingular (where C also denotes the closure of C). Since the complement of C contains X, we get $\kappa(C) \leqq \bar{\kappa}(X)$ from (2.18) and the definition of $\kappa(C)$.

3. Factorisations

Let $f: X \to Y$ be a birational morphism. A *factorization* of f is a pair (g, h) of birational morphisms such that $f = hg$; two factorizations (g, h) and (g', h') of f are *equivalent* if there is an isomorphism u such that $g' = ug$ and $h = h'u$. Let (g, h) be a factorization of f, write $W = \mathrm{dom}(h) = \mathrm{codom}(g)$ and consider $h = (W \hookrightarrow Y_{n(h)} \to \cdots \to Y_0 = Y)$ determined by some minimal decomposition of h. We say that (g, h) is *good* if $q_0(g) = 0$ and if the complement of W in $Y_{n(h)}$ is a union of curves (then $n(f) = n(g) + n(h)$ by (1.3)).

Note that if X and Y are factorial and have trivial units then by (2.11) any factorization $X \to W \to Y$ of f, such that W has the same properties, is good. For that reason, we will restrict ourselves to good factorizations.

By (1.4), all good factorizations of a given birational morphism $f: X \to Y$ can be obtained as follows. For each minimal decomposition $\mathcal{D}$ of f (with notation as usual) and for each $s \in \{0, \cdots, n\}$, let W be the open subset of Y_s obtained by removing all curves $\Gamma \in \{C_1, \cdots, C_q\} \cup \{E_j \mid j \in J \text{ and } j \leqq s\}$ such that $\forall_{i>s} \mu(P_i, \Gamma) = 0$. Then $X \to W \to Y$ is a good factorization of f. Another way to look at that procedure is to fix $\mathcal{D}$ and, for each $A \subseteq \{1, \cdots, n\}$, *try* to change the order of the blowings-up in $\mathcal{D}$ in such a way that the blowings-up at $\{P_i \mid i \in A\}$ are performed first. That point of view leads to the following ideas.

3.1. Let $f: X \to Y$ be a birational morphism and $\mathcal{D}$ and $\mathcal{D}'$ minimal decompositions of f (where the notation of (1.2h) is used for $\mathcal{D}$, and P'_i, E'_i, C'_i, etc. for $\mathcal{D}'$). Then there is a unique pair $(\sigma, \tau) = (\sigma^{\mathcal{D}, \mathcal{D}'}, \tau^{\mathcal{D}, \mathcal{D}'})$ of permutations of $\{1, \cdots, q\}$ and $\{1, \cdots, n\}$ respectively, such that $C_i = C'_{\sigma i}$ for $1 \leqq i \leqq q$ and

(a) $\mu(P_i, \Gamma) = \mu(P'_{\tau i}, \Gamma)$ for $1 \leqq i \leqq n$ and for all curves Γ in Y,

(b) $\mu(P_i, E_j)=\mu(P'_{\tau i}, E'_{\tau j})$ for all $i, j\in\{1, \cdots, n\}$.

From (b), we see that $\tau_i>\tau_j$ whenever $i>j$ and P_i is *i.n.* P_j; any permutation of $\{1, \cdots, n\}$ which satisfies this condition is called a *$\mathscr{D}$-allowable* permutation. Clearly, if τ is $\mathscr{D}$-allowable and σ is any permutation of $\{1, \cdots, q\}$ then $(\sigma, \tau)=(\sigma^{\mathscr{D},\mathscr{D}'}, \tau^{\mathscr{D},\mathscr{D}'})$ for some $\mathscr{D}'$.

A subset A of $\{1, \cdots, n\}$ is said to be *$\mathscr{D}$-closed* if, for all $i, j\in\{1, \cdots, n\}$, $i\in A$, $i>j$ and P_i *i.n.* P_j imply $j\in A$. Note that a topology on $\{1, \cdots, n\}$ is obtained and that, if $\tau=\tau^{\mathscr{D},\mathscr{D}'}$, A is $\mathscr{D}$-closed iff $\tau(A)$ is $\mathscr{D}'$-closed. It is also clear that the existence of a $\mathscr{D}'$ such that $\tau^{\mathscr{D},\mathscr{D}'}(A)=\{1, \cdots, |A|\}$ is equivalent to the $\mathscr{D}$-closedness of A.

For instance, the set $\Delta_{\mathscr{D}}$ is $\mathscr{D}$-open, so we can always find a minimal decomposition satisfying $\Delta=\{n-\delta+1, \cdots, n\}$.

Definition 3.2. Let $f: X\to Y$ be a birational morphism and $\mathscr{D}$ a minimal decomposition of f, with notation as usual. Given a $\mathscr{D}$-closed subset A of $\{1, \cdots, n\}$, define

$$Q(\mathscr{D}, A)=\{i \mid \mu(P_j, C_i)=0, \text{ all } j\notin A\},$$

$$J(\mathscr{D}, A)=\{i\in J \mid \mu(P_j, E_i)=0, \text{ all } j\notin A\}.$$

The next proposition says that to give an equivalence class of good factorizations of f is just the same thing as to give a $\mathscr{D}$-closed set. Its proof is straightforward and is left to the reader.

Proposition 3.3. *Let $f: X\to Y$ be a birational morphism and $\mathscr{D}$ a minimal decomposition of f. Then there is a unique bijection from the set of $\mathscr{D}$-closed subsets of $\{1, \cdots, n(f)\}$ to the set of equivalcnce classes of good factorizations of f, which satisfies the following condition: if C is the equivalence class assigned to the $\mathscr{D}$-closed set A, $(g, h)\in C$, $\mathscr{D}_g$ and $\mathscr{D}_h$ are minimal decompositions of g and h respectively, $\mathscr{D}'$ is the minimal decomposition of f determined by $\mathscr{D}_g$ and $\mathscr{D}_h$ as in* (1.4) *and $\tau=\tau^{\mathscr{D},\mathscr{D}'}$, then $\tau(A)=\{1, \cdots, n(h)\}$. Moreover, we then have $J_{\mathscr{D}_h}=\tau(J(\mathscr{D}, A))$ and the missing curves of h are* the *C_i's with $i\in Q(\mathscr{D}, A)$.*

Proposition 3.4. *Let $g: X\to W$ and $h: W\to Y$ be birational morphisms and suppose that X and Y are factorial and have trivial units. Then W has trivial units, $q(g)\geqq c(g)$, $q(h)\leqq c(h)$ and the following are equivalent:*
(a) *W is factorial,*
(b) *$q(g)=c(g)$ and W is connected at infinity,*
(c) *$q(h)=c(h)$ and W is connected at infinity.*

Proof. By (2.9) applied to g, W has trivial units and $c(g)\leqq q(g)$, with equality iff $\mathrm{Cl}(W)=0$. Since $c(hg)=q(hg)$, $q(h)\leqq c(h)$ and (b)$\Leftrightarrow$(c) follow from (1.5). In order to prove that (c) implies (a), assume that (c) holds and note that $\mathrm{Cl}(W)=0$ by the above remarks. There remains to check that no contracting curve of h is complete (affineness of W will then follow from (2.2)).

Suppose h has a complete contracting curve E. Then E has nonzero self-intersection number in any nonsingular completion of W. Indeed, consider a minimal decomposition of h and in particular write h as the composition $W \hookrightarrow Y_{n(h)} \to \cdots \to Y_0 = Y$. Then E is one of the E_j (=strict transform in $Y_{n(h)}$ of the exceptional curve of $Y_j \to Y_{j-1}$) and consequently has negative self-intersection number. On the other hand, imbed W in a complete nonsingular surface S and apply (2.6) to $W \subseteq S$; since $\mathrm{Cl}(W) = 0$, E is linearly equivalent to a divisor D supported at infinity of W, so that $E^2 = E.D = 0$, contradiction.

Let $\mathcal{P}$ be the property, for surfaces, of being factorial and having trivial units. Let $f: X \to Y$ be a birational morphism such that X and Y have $\mathcal{P}$ and such that, for some $\mathcal{D}$, the numerical data J, $\mathcal{E}$ and μ are known. Can it be decided whether f factors as $X \to W \to Y$ in a nontrivial way, where W is required to have $\mathcal{P}$? Can one list all such factorizations? The answer is *yes* and, as discussed in [1], the results (3.3) and (3.4) suggest algorithms that solve that problem. (Moreover, it follows from (4.4), below, that the problem obtained by letting $\mathcal{P}$ be the property of being isomorphic to $\boldsymbol{A}^2$ has exactly the same solution.)

Let us now consider the case where $j(f)=0$. If the domain and codomain of such an f are factorial and have trivial units then $q(f)=n(f)$, $\det \mu = \pm 1$ by (2.10) and for every good factorization (g, h) of f the surface $W = \mathrm{codom}(g) = \mathrm{dom}(h)$ is connected at infinity. So (3.3) and (3.4) yield the following result, which Russell knew in the special case where f has ordinary fundamental points.

Corollary 3.5. *Let $f: X \to Y$ be a birational morphism with $j(f)=0$, and suppose that X and Y are factorial and have trivial units. Let $\mathcal{D}$ be any minimal decomposition of f, let $\mu = \mu_{\mathcal{D}}$ and let r, s be positive integers such that $r+s=n=n(f)$. Then the following are equivalent:*

(a) *$f = hg$ for some birational morphisms $g: X \to W$ and $h: W \to Y$ such that W is factorial and has trivial units, $n(g)=r$ and $n(h)=s$.*

(b) *Modulo a permutation of the columns and a permutation of the rows, μ has the form*

$$\begin{bmatrix} H & B \\ O & G \end{bmatrix},$$

where H is an $s \times s$ matrix and O is the $r \times s$ zero matrix (hence G is an $r \times r$ matrix and B an $s \times r$ matrix).

Proof. Write $\mu = (\mu_{ij})$. By (1.4), (3.3) and (3.4), (a)⇒(b) is clear and (b)⇒(a) is almost clear; what has to be checked is that (b) implies the following (apparently) stronger statement:

(b′) Modulo a permutation of the columns and an *allowable* (see (3.1)) permutation of the rows, μ has the form described in (b).

Observe that if $1 \leqq i < n$ and $1 \leqq j \leqq n$ are such that $\mu_{ij} = 0$ and $\mu_{i+1\,j} \neq 0$, then it is allowable to interchange rows i and $i+1$. Whence (b)⇒(b′).

To conclude this section, we give a result that says that if $\delta(f)$ is the largest possible, then f factors in a nice way.

Proposition 3.6. *Let $f: X \to Y$ be a birational morphism and suppose that X and Y are factorial and have trivial units. Then $\delta(f) \leqq j(f)$, with equality iff $f=hg$ for some birational morphisms $g: X \to W$ and $h: W \to Y$ such that W is factorial and has trivial units, $n(h)=q(h)=q(f)$ and $n(g)=j(f)=\delta(f)$ (and of course $j(h)=0$).*

Proof. Let $n=n(f)$, $c=c(f)$, $q=q(f)$, $j=j(f)$ and $\delta=\delta(f)$. Then $\delta \leqq j$ by (2.9) and we have to prove that $\delta=j$ iff f factors as specified.

Suppose that $\delta=j$. By (3.1), there exists a minimal decomposition $\mathcal{D}$ of f such that $\Delta=\{n-\delta+1, \cdots, n\}$. Since $\delta=j$ and by (2.10) $c=q$, $\Delta=\{q+1, \cdots, n\}$. With notation as usual for $\mathcal{D}$, let $W=Y_q \setminus (C_1 \cup \cdots \cup C_q)$ and let $h: W \to Y$ be the birational morphism so obtained; then $n(h)=q(h)=c(h)=q$. Since the blowings-up $Y_n \to \cdots \to Y_q$ have centers away from $C_1 \cup \cdots \cup C_q$ (i.e., the centers are *i.n.* W), $f=hg$ for some $g: X \to W$. By (3.4), we conclude that W has trivial units and is factorial. We leave the converse to the reader.

4. Birational endomorphisms of $\boldsymbol{A}^2$

The set of birational endomorphisms of $\boldsymbol{A}^2$ is a monoid, under composition of morphisms. An element f of that monoid is *trivial* if it is an automorphism of $\boldsymbol{A}^2$ (this is equivalent to the definition given in (1.2e) since any open immersion $\boldsymbol{A}^2 \hookrightarrow \boldsymbol{A}^2$ is onto by, say, (2.11); it is *irreducible* if it is nontrivial and can't be decomposed as $h \circ g$ where g and h are nontrivial elements of the monoid. Observe that the "addition formula" $n(g \circ f)=n(f)+n(g)$ holds for birational endomorphisms of $\boldsymbol{A}^2$, by (2.12). In particular, if $n(f)=1$ then f is irreducible.

Two birational endomorphisms f, g of $\boldsymbol{A}^2$ are *equivalent* if $f=v^{-1} \circ g \circ u$ for some automorphisms u, v of $\boldsymbol{A}^2$; we denote that by $f \sim g$.

The interesting problem, here, is to classify all *irreducible* birational endomorphisms of $\boldsymbol{A}^2$. In view of the difficulty of the case $n(f)=2$, which we solve in the last section of [2], we believe the general problem to be very difficult—even what one should mean by "classify" is not clear at present time. In this section we solve the case $n(f)=1$ (see (4.10)) and give some general results, including (4.12), which determines the possible values of $(n(f), j(f), \delta(f))$ for irreducible f (we know that $q_0(f)=0$ and $q(f)=c(f)=n(f)-j(f)$ by (2.10) and (2.11).)

We first state a (trivial) consequence of the theory of "relatively minimal" rational surfaces [10].

Lemma 4.1. *Let S be a rational nonsingular projective surface, $D \in \mathrm{Div}(S)$ a reduced, effective divisor and $U=S \setminus \mathrm{supp}(D)$. Then the following are equivalent:*

(a) $U \cong \boldsymbol{A}^2$;

(b) *every irreducible component of D is a rational curve, $[1] \in \mathcal{G}[U]$ and $n(D)+K_S^2=10$, where $n(D)$ is the number of irreducible components of D and K_S is a canonical divisor of S.*

The following "powerful" theorem was proved by Fujita [3] and Miyanishi and

Sugie [8] in characteristic zero, and generalized by Russell [13] to arbitrary characteristic:

Theorem. *Let V be a nonsingular, factorial, rational surface with trivial units, and with logarithmic Kodaira dimension $\bar{\kappa}(V)<0$. Then $V\cong \boldsymbol{A}^2$.*

From this and (2.18), it follows immediately

Corollary 4.2. *Let $f: \boldsymbol{A}^2\to V$ be a birational morphism, where V is factorial (and nonsingular, as always). Then $V\cong \boldsymbol{A}^2$.*

Let us now consider the main results of sections 1-3 and point out what they say about the special case "$X=Y=\boldsymbol{A}^2$".

Corollary 4.3. *Let $f: \boldsymbol{A}^2\to \boldsymbol{A}^2$ be a birational morphism.*

(a) *$q_0(f)=0$, $q(f)=c(f)$ and $\delta(f)\leqq j(f)$ with equality iff f factors as $f=hg$, where g and h are birational endomorphisms of $\boldsymbol{A}^2$ such that $n(h)=q(h)=q(f)$ and $n(g)=j(f)=\delta(f)$.*
(b) *Given any minimal decomposition of f, the corresponding (square) matrix $\varepsilon'\mu$ has determinant ± 1.*
(c) *Every missing curve of f is rational and has one place at infinity. Embed $\boldsymbol{A}^2$ in $\boldsymbol{P}^2$ the standard way and let $Q_1, \cdots, Q_r$ be the points where the closures of the missing curves meet the line L at infinity; then at most one of these points Q_i does not satisfy the condition: Exactly one missing curve meets L at Q_i and that missing curve has degree one.*
(d) *All fundamental points of f are on the missing curves.*
(e) *The result numbered (2.17) is valid here.*
(f) *If C is a missing curve of f then $\kappa(C)<0$.*

Proof. $q_0(f)=0$ by (2.11), $q(f)=c(f)$ by (2.10) and the rest of (a) by (3.6) and (4.2). (b) comes from (2.10), (d) from (2.1), (e) from (2.17), (f) from (2.19) and the first two assertions of (c) from (2.16). We prove the last assertion of (c). Choose a minimal decomposition of f, with notation as usual, let $Y_0=\boldsymbol{A}^2\hookrightarrow \boldsymbol{P}^2=\bar{Y}_0$ be the embedding, consider the diagram in the proof of (2.1) and let $D=L+C_1+\cdots+C_q+\sum_{i\in J} E_i\in \mathrm{Div}(\bar{Y}_n)$. Then $\bar{Y}_n\setminus \boldsymbol{A}^2=\mathrm{supp}(D)$ and, by (5.19), D has at most one "bad" point Q_*. For any $Q\in L\setminus\{Q_*\}$, Q belongs to at most two components of D, i.e., to at most one C_j; if $Q\in C_j$ then $(L.C_j)_Q=1$ (in $\bar{Y}_n$, hence in $\bar{Y}_0$), whence $L\cdot C_j=1$ in $\boldsymbol{P}^2$.

In view of (4.2), the next two results are trivial consequences of (3.4) and (3.5) respectively.

Corollary 4.4. *Let $g: \boldsymbol{A}^2\to W$ and $h: W\to \boldsymbol{A}^2$ be birational morphisms. Then W has trivial units, $q(g)\geqq c(g)$, $q(h)\leqq c(h)$ and the following are equivalent:*

(a) *$W\cong \boldsymbol{A}^2$,*
(b) *$q(g)=c(g)$ and W is connected at infinity,*

(c) *$q(h)=c(h)$ and W is connected at infinity.*

Corollary 4.5. *Let $f: \boldsymbol{A}^2 \to \boldsymbol{A}^2$ be a birational endomorphism with $j(f)=0$, let $\mathcal{D}$ be any minimal decomposition of f, let $\mu=\mu_{\mathcal{D}}$ and let r, s be positive integers such that $r+s=n=n(f)$. Then the following are equivalent:*

(a) *$f=hg$, for some birational endomorphisms g, h of $\boldsymbol{A}^2$ such that $n(g)=r$ and $n(h)=s$.*

(b) *Modulo a permutation of the columns and a permutation of the rows, μ has the form*

$$\begin{bmatrix} H & B \\ O & G \end{bmatrix}$$

where H is an $s \times s$ matrix and O is the $r \times s$ zero matrix.

Recall that a curve C in $\boldsymbol{A}^2$ is a *line* if $C \cong \boldsymbol{A}^1$. A line is a *coordinate line* if its defining polynomial $F \in \boldsymbol{k}[X, Y]$ satisfies $\boldsymbol{k}[F, G]=\boldsymbol{k}[X, Y]$ for some G; otherwise it is a *wild line*. Theorem 2.4 of [4] says in particular that if C is a wild line then $\bar{\kappa}(C) \geqq 0$. Hence by (4.3f) it follows that *no missing curve of $f: \boldsymbol{A}^2 \to \boldsymbol{A}^2$ is a wild line.*

Corollary 4.6. *Let f be a birational endomorphism of $\boldsymbol{A}^2$, let μ be the matrix determined by some minimal decomposition of f and let C_j be a missing curve. Then the following are equivalent:*

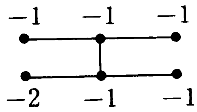
1. *C_j is a coordinate line,*
2. *C_j is nonsingular,*
3. *the j^{th} column of μ consists of 0's and 1's.*

Proof. (1)$\Rightarrow$(3) is trivial and (3)$\Rightarrow$(2) follows from (4.3e). By (4.3c) C_j is rational with one place at infinity so (2)$\Rightarrow C_j \cong \boldsymbol{A}^1$ and (1) follows from the above observations.

We now give some examples of irreducible birational endomorphisms of $\boldsymbol{A}^2$.

Example 4.7 (Russell). Let C_1 be an irreducible curve of degree two in $\boldsymbol{A}^2$, with one place at infinity (a parabola). Let P_1, P_2, P_3 be distinct points of C_1 and let C_2 (resp. C_3) be the line through P_1 and P_3 (resp. P_1 and P_2). Blow-up $\boldsymbol{A}^2$ at P_1, P_2, P_3 and remove the strict transforms of C_1, C_2, C_3 from the blown-up surface. Then the resulting open set is isomorphic to $\boldsymbol{A}^2$ and we obtain an irreducible birational morphism $f: \boldsymbol{A}^2 \to \boldsymbol{A}^2$ with $n(f)=3$.

Proof. First, we show that the surface obtained is $\boldsymbol{A}^2$. Embed $\boldsymbol{A}^2$ in $\boldsymbol{P}^2$ the standard way and let $L=\boldsymbol{P}^2 \setminus \boldsymbol{A}^2$; let P be the place of C_1 at infinity. Blow-up $\boldsymbol{P}^2$ at P_1, P_2, P_3, denote the blown-up surface by $\tilde{\boldsymbol{P}}^2$ and consider (i.e., make a picture of) the strict transforms of L, C_1, C_2, C_3 in $\tilde{\boldsymbol{P}}^2$, with self-intersection numbers 1, 1, -1, -1 respectively. To show: $U \cong \boldsymbol{A}^2$, where $U=\tilde{\boldsymbol{P}}^2 \setminus (L \cup C_1 \cup C_2 \cup C_3)$. By (4.1), enough to show that $[1] \in \mathcal{G}[U]$. Note that $(L.C_1)_P=L.C_1=2$ and blow-up $\tilde{\boldsymbol{P}}^2$ twice at $P \in C_1$; the resulting divisor, i.e., the reduced effective divisor at infinity of U, has *s.n.c.* and determines the dual graph



,

which is equivalent to [1]. So $U \cong \boldsymbol{A}^2$. To prove irreducibility, consider

$$\mu = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

and apply (4.5).

Example 4.8 (Russell). Let a be a positive integer, let C_1 and C_2 be the curves given by the polynomials

$$F_1 = X^{a+1}(X-1)^a + Y^{a+1}, \qquad F_2 = Y,$$

and let $P_1 = (0, 0)$ and $P_2 = (1, 0)$. Blow-up $\boldsymbol{A}^2$ at P_1 and P_2 and remove the strict transforms of C_1, C_2. The resulting surface is isomorphic to $\boldsymbol{A}^2$ and we get an irreducible $f : \boldsymbol{A}^2 \to \boldsymbol{A}^2$ with $n(f) = 2$. Verification left to the reader.

Example 4.9 (Russell). Let $n \geqq 3$ and let C_1 be an irreducible curve of degree $n-1$ in $\boldsymbol{A}^2$ such that
(a) C_1 has one place at infinity,
(b) C_1 has a point P_1 (in $\boldsymbol{A}^2$) of multiplicity $n-2$.
Clearly, such a curve exists. Choose distinct lines $C_2, \cdots, C_n$ such that
(c) $C_i \cap C_1 = \{P_1, P_i\}$, some $P_i \in \boldsymbol{A}^2 \setminus \{P_1\}$ $(2 \leqq i \leqq n)$.
Blow-up $\boldsymbol{A}^2$ at $P_1, \cdots, P_n$ and remove the strict transforms of $C_1, \cdots, C_n$. The resulting surface is isomorphic to $\boldsymbol{A}^2$ and we get an irreducible $f : \boldsymbol{A}^2 \to \boldsymbol{A}^2$ with $n(f) = n$. Verification left to the reader.

Let's now return to the classification problem.

Theorem 4.10. *Let f be a birational endomorphism of $\boldsymbol{A}^2$, with $n(f) = 1$. Then f is a simple affine contraction.*

Proof. Recall, from the introduction of this paper, the definition of simple affine contraction; we leave it to the reader to verify the following statement:

A birational morphism $f : \boldsymbol{A}^2 \to \boldsymbol{A}^2$ is a simple affine contraction iff $n(f) = 1$ and the missing curve of f is a coordinate line.

Now let f be any birational endomorphism of $\boldsymbol{A}^2$ with $n(f) = 1$. Then $c(f) = q(f) = 1$ and the matrix μ is the 1×1 matrix (1) by (4.3b). So the result follows from (4.6).

The above proof uses (4.6), which relies on somewhat fancy ideas (Kodaira dimension and [4]). A simpler (and longer) proof of (4.10) is given in [2]. The above theorem generalizes as follows:

Theorem 4.11. *Let f be a birational endomorphism of A^2 such that $q(f)=1$. Then $f\sim s^n$, where $n=n(f)$ and s is a simple affine contraction.*

Proof. We proceed by induction on $n=n(f)$. The case $n=1$ is just (4.10), above.

Let $n>1$ be such that the claim holds whenever $n(f)<n$. Let f be such that $n(f)=n$, let C denote the missing curve of f and choose a minimal decomposition of f, with notation as usual. Since $j(f)=n(f)-c(f)=n(f)-q(f)=n-1$ and $n\notin J$ by (1.2i),

(1) $$J=\{1, \cdots, n-1\}.$$

Again by (1.2i),

(2) $$P_{i+1}\in E_i, \qquad 1\leqq i<n.$$

Thus an elementary calculation shows that

(3) $$\varepsilon_{1j}\leqq \cdots \leqq\varepsilon_{nj}, \qquad 1\leqq j\leqq n$$

(see (2.7) and (2.8) for definitions). Since $\varepsilon_{jj}=1$, we deduce

(4) $$\varepsilon_{nj}\geqq 1, \qquad 1\leqq j\leqq n.$$

On the other hand,

$$\varepsilon'\mu=\Big(\sum_{j=1}^{n}\varepsilon_{nj}\mu(P_j, C)\Big) \qquad (1\times 1 \text{ matrix})$$

so by (4.3b)

(5) $$\sum_{j=1}^{n}\varepsilon_{nj}\mu(P_j, C)=1.$$

By (4) and (5)

$$1\geqq\sum_{j=1}^{n}\mu(P_j, C)\geqq\mu(P_1, C)=1, \qquad \text{so}$$

(6) $$\mu=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$$

and consequently $\varepsilon_{n1}=1$ by (5) and (6). If $1\leqq i<n$ then by (3) and (2)

$$1=\varepsilon_{n1}\geqq\varepsilon_{i+1\,1}=\sum_{k=1}^{i}\varepsilon_{k1}\mu(P_{i+1}, E_k)$$

$$\geqq\Big[\sum_{k=1}^{i-1}\mu(P_{i+1}, E_k)\Big]+\mu(P_{i+1}, E_i)\geqq\mu(P_{i+1}, E_i)=1$$

whence

(7) $$P_{i+1}\in E_j \text{ in } Y_i \Longleftrightarrow j=i, \qquad \text{all } i, j,$$

By (6) and (7), $P_n\notin(C\cup E_1\cup\cdots\cup E_{n-2})$ in Y_{n-1}. So the image of $A^2\hookrightarrow Y_n\to Y_{n-1}$ is contained in $W=Y_{n-1}\setminus(C\cup E_1\cup\cdots\cup E_{n-2})$, i.e., f factors as $g: A^2\to W$ followed by $h: W\to A^2$. Clearly $n(h)=n-1$ and $W\hookrightarrow Y_{n-1}\to\cdots\to Y_0$ gives a minimal decomposition of h. So $E_{n-1}\cap W$ is the only contracting curve of h and C the only missing curve, hence $q(h)=1=c(h)$. One sees that W is connected at infinity, so $W\cong A^2$ by (4.4). Thus by the inductive hypothesis we get $h\sim s^{n-1}$ and $g\sim s$. Since the missing curve of g coincides with the contracting curve of h, $f\sim s^n$.

Note that [12] contains variations of (4.10) and (4.11)—see in particular the remark (3.4).

Until recently, no example of an irreducible $f: A^2\to A^2$ with $j(f)>0$ was known. Moreover, the above theorem says that if $j(f)$ has the maximum possible value then f is reducible (unless $n(f)=1$). The hope that $j(f)>0\Rightarrow f$ *reducible* is killed by

Theorem 4.12. *Let n, j and δ be nonnegative integers. There exists an irreducible birational morphism $f: A^2\to A^2$ satisfying $n(f)=n$, $j(f)=j$ and $\delta(f)=\delta$ if and only if one of the following conditions holds:*

(a) $0=\delta=j<n$;

(b) $0\leqq\delta<j<n-1$.

Proof. The "only if" part follows from (4.3a) and (4.11). Conversely, the case (a) with $n=1$ (resp. $n=2$, $n>2$) is realized by the simple affine contractions (resp. (4.8), (4.9)). If (n, j, δ) satisfies (b), let $m=j-\delta+1\geqq 2$ and $q=n-j\geqq 2$ and choose $\delta_1\geqq 0, \cdots, \delta_{q-1}\geqq 0$ such that $\delta_1+\cdots+\delta_{q-1}=\delta$. Then example (4.13) realizes these numbers.

Example 4.13. Let $m\geqq 2$, $q\geqq 2$, $\delta_1\geqq 0, \cdots, \delta_{q-1}\geqq 0$ be integers. We construct an irreducible birational morphism $f: A^2\to A^2$ with two fundamental points and satisfying

$$n(f)=m+q-1+\delta_1+\cdots+\delta_{q-1},$$

$$q(f)=q,$$

$$\delta(f)=\delta_1+\cdots+\delta_{q-1},$$

$$j(f)=m-1+\delta(f).$$

Choose $F_1, \cdots, F_q\in \boldsymbol{k}[X, Y]$ such that if C_i is the affine plane curve $F_i(X, Y)=0$ then

- C_i is a nonsingular rational curve of degree m, with one place at infinity, with multiplicity sequence at infinity: $m-1, 1, 1, \cdots$;
- there are distinct points $P_1, P_2\in A^2$ such that

$$\forall_{i\neq j}[C_i\cap C_j=\{P_1, P_2\}, (C_i.C_j)_{P_1}=1, (C_i.C_j)_{P_2}=m-1].$$

(For instance, $F_i=a_iY^{m-1}(Y-1)+X$, where $a_1, \cdots, a_q$ are distinct elements of $\boldsymbol{k}^*$; then $P_1=(0, 1)$ and $P_2=(0, 0)$.)

We are going to embed A^2 in $\boldsymbol{F}_m$ (one of the Nagata rational surfaces). First, embed A^2 in $\boldsymbol{P}^2$ the standard way and write $\boldsymbol{P}^2\setminus A^2=L$. Let C_i also denote the closure in $\boldsymbol{P}^2$ of the curve C_i chosen above. The curves $C_1, \cdots, C_q$ all meet L at the same point P. Note that

1) $C_i \cap L = \{P\}$, $\mu(P, C_i) = m-1$, $C_i.L = m$, all i;

2) $C_i.C_j = m^2$, all i, j;

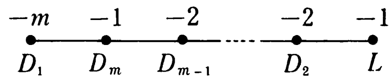
3) $(C_i.C_j)_P = m^2 - m$, all distinct i, j.

Blow-up $\boldsymbol{P}^2$ at $A_1 = P$, let D_1 be the exceptional curve, let A_2 be the point at which D_1 and L meet. Then

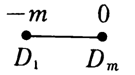
4) $C_i \cap L = \{A_2\} = C_i \cap D_1$, $C_i \cong \boldsymbol{P}^1$, $C_i^2 = 2m-1$, $C_i.L = 1$, $C_i.D_1 = m-1$, for all i;

5) $C_i \cap C_j = \{P_1, P_2, A_2\}$, $(C_i.C_j)_{A_2} = m-1$, all distinct i, j.

Blow-up $m-1$ times at the point of D_1 which is $i.n.$ A_2. Call the exceptional curves so obtained $D_2, \cdots, D_m$. On the resulting surface, the divisor $D_1 + \cdots + D_m + L$ has $s.n.c.$, its dual graph is the linear weighted tree



and the complement of that divisor is $\boldsymbol{A}^2$. Contract $L, D_2, \cdots, D_{m-1}$ and let S_0 denote the complete surface obtained. We get $\boldsymbol{A}^2 = S_0 \setminus \mathrm{supp}(D_1 + D_m)$, where $D_1 + D_m \in \mathrm{Div}(S_0)$ has $s.n.c.$ and has dual graph $\mathcal{G}(S_0, D_1 + D_m)$ as follows:

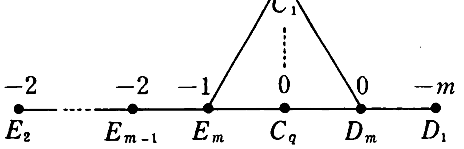


In fact, $S_0 = \boldsymbol{F}_m$ (but we don't really need to know that).

Now $C_1, \cdots, C_m$ meet D_m at distinct points and

6) $C_i \cap D_1 = \emptyset$, $C_i.D_m = 1$ and $C_i^2 = m$, all i.

We now proceed to define an equivalence class of irreducible morphisms $f: \boldsymbol{A}^2 \to \boldsymbol{A}^2$. Blow-up S_0 at P_1; blow-up $m-1$ times at P_2 (more precisely, always blow-up at the intersection point of (the strict transforms of) the C_i's). The last of these blowings-up makes $C_1, \cdots, C_q$ pairwise disjoint. If $E_1, E_2, \cdots, E_m$ are the exceptional curves so created, then on the blown up surface the divisor $E_2 + \cdots + E_m + C_1 + \cdots + C_q + D_1 + D_m$ has $s.n.c.$ and its dual graph is



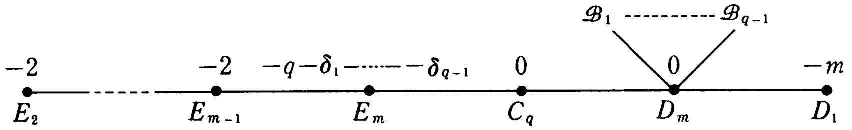
For $i = 1, \cdots, q-1$, let Q_i be the intersection point of E_m and C_i.

Blow-up δ_1+1 times at Q_1,
then δ_2+1 times at Q_2,
$\vdots$
and $\delta_{q-1}+1$ times at Q_{q-1};

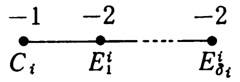
more precisely, always blow-up the point of E_m which is *i.n.* Q_i. Denote by $E_1^1, \cdots, E_{\delta_1+1}^1, E_1^2, \cdots, E_{\delta_{q-2}+1}^{q-2}, E_1^{q-1}, \cdots, E_{\delta_{q-1}+1}^{q-1}$ the exceptional curves so created. On the resulting surface, call it S_n, consider the divisor

$$D=E_2+\cdots+E_m+(E_1^1+\cdots+E_{\delta_1}^1)+\cdots+(E_1^{q-1}+\cdots+E_{\delta_{q-1}}^{q-1})$$
$$+C_1+\cdots+C_q+D_m+D_1,$$

whose dual graph $\mathcal{G}(S_n, D)$ is



where, for $i=1, \cdots, q-1$, $\mathcal{B}_i$ is

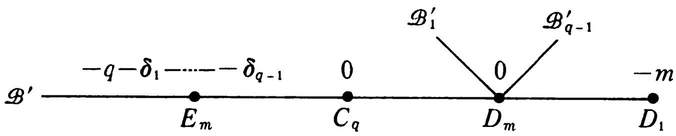


C_i being linked to D_m. We claim that the complement of supp(D) is isomorphic to A^2. By (4.1), enough to show that $\mathcal{G}(S_n, D)\sim[1]$. Now $\mathcal{G}(S_n, D)$ contracts to

$$[-2, \cdots, -2, -q-\delta_1-\cdots-\delta_{q-1}, 0, \delta_1+\cdots+\delta_{q-1}+q-1, -m]$$
$$\sim[-2, \cdots, -2, -1, 0, 0, -m]\sim[m-1, 0, -m]\sim[-1, 0, 0]\sim[1],$$

where we use the notation for linear weighted trees defined in (5.13) and the fact pointed out in (5.14).

So we get an equivalence class of birational morphisms $f: A^2\to A^2$ with $n(f)$, $q(f)$, $\delta(f)$ and $j(f)$ as desired. We leave it to the reader to verify that if $f=h\circ g$ with $0<n(h)<n(f)$, then h gives rise to a sub weighted tree $\mathcal{G}'$ of $\mathcal{G}=\mathcal{G}(S_n, D)$ such that $\mathcal{G}'$ contains D_1, D_m and at least one more vertex, $\mathcal{G}'\neq\mathcal{G}$ and $\mathcal{G}'\sim[1]$. We claim that $\mathcal{G}$ does not contain such a $\mathcal{G}'$. To see that, suppose $\mathcal{G}'$ exists. Then C_q is in $\mathcal{G}'$, otherwise $\mathcal{G}'$ would contract to $[p, -m]$ for some $p>0$, and $[p, -m]\not\sim[1]$ by (5.16). Next, E_m is in $\mathcal{G}'$, for otherwise $\mathcal{G}'$ contracts to $[0, p, -m]$ for some $p\geqq 0$, and this is not equivalent to $[1]$. So $\mathcal{G}'$ has the form



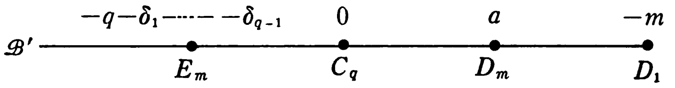
.

where each $\mathscr{B}'$, $\mathscr{B}_i'$ is either empty or a linear branch, and

$$\mathscr{B}_i'=[-1, -2, \cdots, -2] \qquad \text{if not empty,}$$

$$\mathscr{B}'=[-2, \cdots, -2] \qquad \text{if not empty.}$$

Note that, if $\mathscr{B}_i'$ is not empty then the vertex of weight -1 is there and is the neighbour of D_m. Hence we see that all (nonempty) $\mathscr{B}_i'$ can be absorbed by D_m, and that the absorption of $\mathscr{B}_i'$ increase the weight of D_m by the number $|\mathscr{B}_i'|$. Let $a=|\mathscr{B}_1'|+\cdots+|\mathscr{B}_{q-1}'|$. Then $\mathcal{G}'$ contracts to the minimal weighted tree



By (5.16), $a-q-\delta_1-\cdots-\delta_{q-1}=-1$, so $|\mathscr{B}_1'|+\cdots+|\mathscr{B}_{q-1}'|=|\mathscr{B}_1|+\cdots+|\mathscr{B}_{q-1}|$, i.e., $\mathscr{B}_i'=\mathscr{B}_i$ for all i.

Let $b=|\mathscr{B}'|$. Then

$$\mathcal{G}'\sim[-2, \cdots, -2, -q-\delta_1-\cdots-\delta_{q-1}, 0, a, -m]$$

$$\sim[-2, \cdots, -2, -1, 0, 0, -m]\sim[b+1, 0, -m].$$

By (5.16) again, $b+1-m=-1$, i.e., $b=m-2$ and $\mathcal{G}'=\mathcal{G}$. Hence f is irreducible.

5. Appendix: weighted graphs

Although most of the material contained in this section appeared at several other places, we include it to establish notation and to make reference easier. We used [11] as our main reference. Note that (5.18) and the last three assertions of (5.19) didn't seem to be known before this.

For our purposes a graph consists of finitely many vertices, some of them being connected by links, such that the links are not oriented and at most one can exist between two given vertices. So let us say that a *graph* is a pair $\mathcal{G}=(G, R)$ where G is a finite set and R is a set of subsets of G, such that every $a\in R$ contains *exactly* two elements. The elements of G are called the *vertices* of $\mathcal{G}$ and those of R are the *links*. Two vertices u, v of $\mathcal{G}$ are said to be *linked* if $\{u, v\}\in R$; we also say that u is a *neighbour* of v, and vice-versa. The set of neighbours of v is denoted by $\mathcal{N}_{\mathcal{G}}(v)$. A vertex v of $\mathcal{G}$ is *free* (resp. *linear, a branch point*) if it has at most one (resp. at most two, at least three) neighbour(s). $|\mathcal{G}|$ denotes the number of vertices of $\mathcal{G}$.

Given vertices u, v a *chain* from u to v is a sequence $(x_0, \cdots, x_q)$ of vertices such that $q>0$, $u=x_0$, $v=x_q$ and $\{x_i, x_{i+1}\}\in R$ for $0\leqq i<q$. The chain is *simple* if the links $\{x_0, x_1\}, \cdots, \{x_{q-1}, x_q\}$ are distinct. It is a *loop* if it is simple and if $x_0=x_q$. The *connected components* of $\mathcal{G}$ are defined in the obvious way. A *tree* is a connected graph without loops. A *linear tree* is a tree without branch points.

If $\mathcal{G}=(G, R)$ is a graph and $V\subseteqq G$ then $\mathcal{G}\setminus V$ is the graph (G', R') where $G'=G\setminus V$ and $R'=\{a\in R \mid a\cap V=\emptyset\}$. If $\mathcal{G}$ is a tree and v is a vertex of $\mathcal{G}$ then the connected components of $\mathcal{G}\setminus\{v\}$ are called the *branches* of $\mathcal{G}$ at v; clearly, the tree $\mathcal{G}$ has $|\mathcal{N}_{\mathcal{G}}(v)|$ branches at v.

Definition 5.1. A *weighted graph* is a triple $\mathcal{G}=(G, R, \Omega)$ where (G, R) is a graph and Ω is some set map $G\to\boldsymbol{Z}$. If $v\in G$, $\Omega(v)$ is called the *weight* of v.

A weighted graph can be blown up at a link or at a vertex:

Definition 5.2. Let $\mathcal{G}=(G, R, \Omega)$ be a weighted graph and let x be either a link or a vertex of $\mathcal{G}$. A *blowing-up of* $\mathcal{G}$ at x is a weighted graph $\mathcal{G}'=(G', R', \Omega')$ together with an injective map $G\hookrightarrow G'$, such that if G is identified with its image in G' then $G'=G\cup\{e\}$ for some $e\notin G$ and the following conditions are satisfied:

1. if $x=\{u, v\}\in R$ then $R'=(R\setminus\{\{u, v\}\})\cup\{\{e, u\}, \{e, v\}\}$ and

$$\Omega'(w)=\begin{cases}\Omega(w) & \text{if } w\notin\{e, u, v\}\\ \Omega(w)-1 & \text{if } w\in\{u, v\}\\ -1 & \text{if } w=e;\end{cases}$$

2. if $x\in G$ then $R'=R\cup\{\{e, x\}\}$ and

$$\Omega'(w)=\begin{cases}\Omega(w) & \text{if } w\notin\{e, x\}\\ \Omega(w)-1 & \text{if } w=x\\ -1 & \text{if } w=e.\end{cases}$$

Because the blowing-up of $\mathcal{G}$ at x always exists and is essentially unique, "blowing-up" is usually thought of as the operation by which $\mathcal{G}'$ is obtained from $\mathcal{G}$ and x. We sometimes refer to e as *the vertex which is created in the blowing-up*; that vertex is clearly a superfluous vertex of $\mathcal{G}'$:

Definition 5.3. Let $\mathcal{G}$ be a weighted graph. A *superfluous vertex* of $\mathcal{G}$ is a linear vertex e of weight -1 such that $\mathcal{N}_{\mathcal{G}}(e)\neq\emptyset$ and if $u, v\in\mathcal{N}_{\mathcal{G}}(e)$ then u and v are not linked to each other.

Definition 5.4. Let $\mathcal{G}=(G, R, \Omega)$ be a weighted graph and e a superfluous vertex of $\mathcal{G}$. A *blowing-down of* $\mathcal{G}$ *at* e is a weighted graph $\mathcal{G}'=(G', R', \Omega')$ together with an injective map $G'\hookrightarrow G$ that makes $\mathcal{G}$ a blowing-up of $\mathcal{G}'$ at some vertex or link and e the vertex which is created in that blowing-up. The blowing-down of $\mathcal{G}$ at e always exists and is essentially unique. We say that e *disappears* in the blowing-down.

We say that $\mathcal{G}$ *contracts* to $\mathcal{G}'$ if either $\mathcal{G}'$ is isomorphic to $\mathcal{G}$ or if $\mathcal{G}'$ can be obtained from $\mathcal{G}$ by performing finitely many blowings-down (define *isomorphism* the obvious way, i.e., a bijection of the sets of vertices that preserves links and weights). A weighted graph is said to be *minimal* if it has no superfluous vertex. Two weighted graphs $\mathcal{G}$ and $\mathcal{G}'$ are *equivalent* ($\mathcal{G}\sim\mathcal{G}'$) if one can be obtained from the other by a finite sequence of blowings-up and blowings-down. Clearly, if $\mathcal{G}$ and $\mathcal{G}'$ are equivalent then $\mathcal{G}$ is connected (resp. has no loops, is a tree) iff $\mathcal{G}'$ has the same property.

For convenience, let's give a name to the conditions (i) and (ii) that appear on page 70 of [11]:

Definition 5.5. Let D be a divisor of a nonsingular surface S. We say that D has *strong normal crossings* (*s.n.c.*) if D is effective, reduced, and if the following conditions hold:

1. every irreducible component of D is a nonsingular curve;
2. if C and C' are distinct irreducible components of D such that $C\cap C'\neq\emptyset$, then $C\cap C'$ consists of a single point where C and C' meet transversally;
3. if C, C' and C'' are distinct irreducible components of D then $C\cap C'\cap C''=\emptyset$.

Note that if D has *s.n.c.* and S is not complete then there is an open embedding $S\hookrightarrow\bar{S}$ such that $\bar{S}$ is a complete nonsingular surface and the closure of D in $\bar{S}$ has *s.n.c.*.

Definition 5.6. Let S be a nonsingular complete surface and let D be a divisor of S with *s.n.c.*. The *dual graph* $\mathcal{G}(S, D)$ associated to the pair (S, D) is the weighted graph which has the irreducible components of D as vertices, two of them linked iff they intersect in S, and such that each vertex C has weight C^2 (self-intersection number in S).

Clearly, if (S, D) is as above, $\pi:\tilde{S}\to S$ is the blowing-up of S at some point $P\in\operatorname{supp}(D)$, $E=\pi^{-1}(P)$, $\tilde{D}$ is the strict transform of D and $D'=\tilde{D}+E\in\operatorname{Div}(\tilde{S})$ then D' has *s.n.c.* and $\mathcal{G}(\tilde{S}, D')$ is a blowing-up of $\mathcal{G}(S, D)$ in a natural way.

Definition 5.7. Let X be a nonsingular surface. A *smooth completion* of X is an open immersion $X\hookrightarrow S$ such that S is a nonsingular complete surface and $S\setminus X=\operatorname{supp}(D)$ for some $D\in\operatorname{Div}(S)$ with *s.n.c.*. The weighted graph $\mathcal{G}(S, D)$ is therefore determined by $X\hookrightarrow S$; it is easily verified that the equivalence class of $\mathcal{G}(S, D)$ depends only on X. That equivalence class is denoted by $\mathcal{G}[X]$. Note that smooth completions exist for any X.

Definition 5.8. An arbitrary weighted graph $\mathcal{G}=(G, R, \Omega)$ determines a bilinear form $B(\mathcal{G})$, on the real vector space $\boldsymbol{R}^G$ which has G as a basis, defined by

$$v_i.v_i=\Omega(v_i),\qquad \text{all } i,$$

$$v_i.v_j=\begin{cases}1 & \text{if } \{v_i, v_j\}\in R\\ 0 & \text{if } i\neq j \text{ and } \{v_i, v_j\}\notin R,\end{cases}$$

where $G=\{v_1, v_2, \cdots\}$. The *discriminant* of $B(\mathcal{G})$ is denoted by $d(\mathcal{G})$ (i.e., $d(\mathcal{G})$ is the determinant of the $|\mathcal{G}|\times|\mathcal{G}|$ matrix (v_i, v_j)). One can check that if $\mathcal{G}'$ is a blowing-up of $\mathcal{G}$ then $d(\mathcal{G}')=-d(\mathcal{G})$. Thus the number $(-1)^{|G|-1}d(\mathcal{G})$ depends only on the equivalence class of $\mathcal{G}$. We let $\langle\mathcal{G}\rangle$ denote the nonnegative integer $\max\dim W$, where W runs in the set of linear subspaces $W\subseteq\boldsymbol{R}^G$ such that $x.x\geqq 0$, all $x\in W$. One can check that $\langle\mathcal{G}\rangle$ depends only on the equivalence class of $\mathcal{G}$. The following (elementary) fact can be found at p. 78 of [11]:

5.9. *If $\mathcal{G}$ has no loops and $\langle \mathcal{G} \rangle \leqq 1$ then there can be at most two vertices with non-negative weights, and if there are two of them then these two vertices are linked and one of the weights is actually zero.*

Of particular interest for us is the equivalence class $\mathcal{G}[\boldsymbol{A}^2]$, which consists of the weighted trees equivalent to $[1]$, where

Definition 5.10. Given $n \in \boldsymbol{Z}$, the symbol $[n]$ will denote any weighted tree which has one vertex, say v, and such that v has weight n.

More generally, we are also interested in those weighted trees which are equivalent to a linear tree. In this respect, we make the following observations.

Lemma 5.11. *Let $\mathcal{G}$ be a weighted tree equivalent to a linear tree, and suppose that b is a branch point of $\mathcal{G}$. Then for some branch $\mathcal{B}$ of $\mathcal{G}$ at b, "b can absorb $\mathcal{B}$", i.e., there exists a unique weighted tree $\mathcal{G}'$ such that:*

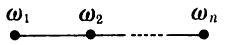
1. *$\mathcal{G}$ contracts to $\mathcal{G}'$,*
2. *$\mathcal{G}' = \mathcal{G} \setminus \mathcal{B}$ as graphs,*
3. *$\mathcal{G}' \setminus \{b\} = \mathcal{G} \setminus (\{b\} \cup \mathcal{B})$ as weighted graphs.*

Corollary 5.12. *Every minimal weighted tree equivalent to a linear tree is linear.*

It doesn't seem to be possible to give a reasonable description of all weighted trees equivalent to $[1]$. However, those which are minimal must be linear (5.12) and it turns out that they can be listed. Before we do that, we need to introduce some notations.

Definition 5.13.

1. Given integers $\omega_1, \cdots, \omega_n$, let $[\omega_1, \cdots, \omega_n]$ be the linear weighted tree



If s_i is either an integer or a finite sequence of integers (for each $i=1, \cdots, k$), let $[s_1, \cdots, s_k]$ be the linear weighted tree $[\omega_1, \cdots, \omega_n]$, where $(\omega_1, \cdots, \omega_n)$ is the sequence obtained by concatenating $s_1, \cdots, s_k$.

2. Given $p, q \in \boldsymbol{Z}$ with $p \geqq 0$, let R_p^q be the $p+1$-tuple $(-q-2, -2, \cdots, -2)$, and let L_p^q be the $p+1$-tuple $(-2, \cdots, -2, -q-2)$. To be precise, $R_0^q = (-q-2) = L_0^q$.

Example 5.14. The tree $[L_2^1, 0, 2, R_0^2]$ is just the same as $[-2, -2, -3, 0, 2, -4]$ which is, by the way, equivalent to $[1]$. To see this, observe that if A, B are (possibly empty) finite sequences of integers and $a, b \in \boldsymbol{Z}$ then

$$[A, a, 0, b, B] \sim [A, a+i, 0, b-i, B]$$

for any $i \in \boldsymbol{Z}$. In our case,

$$[-2,-2,-3,0,2,-4]\sim[-2,-2,-1,0,0,-4]\sim[3,0,-4]\sim[0,0,-1]\sim[0,1]$$

which is equivalent to [1]; indeed, if $n\in\boldsymbol{Z}$ then $[0, n]\sim[-1, -1, n]\sim[0, n+1]$ and consequently $[0, n]\sim[0, -1]\sim[1]$.

Proposition 5.15. *The following is a list of all minimal weighted trees equivalent to* [1].

1. [1]
2. $[0, \alpha]$, $\alpha\in\boldsymbol{Z}\setminus\{-1\}$
3. $[\cdots, L_{\alpha_5}^{\alpha_4+1}, L_{\alpha_3}^{\alpha_2+1}, L_{\alpha_1}^{\alpha_0}, 0, \alpha_0+1, R_{\alpha_2}^{\alpha_1}, R_{\alpha_4}^{\alpha_3+1}, R_{\alpha_6}^{\alpha_5+1}, \cdots]$ *where* $\alpha_0, \cdots, \alpha_k$ *is a finite sequence of nonnegative integers, with* $k\geq 1$.

Remark. The above list appeared in theorem 9 of [9], with a different notation. However, geometry is very much involved in the cited result (i.e., in both the assertion and its proof) while this proposition is purely graph-theoretic. A graph-theoretic proof is given in [1].

Corollary 5.16. *Let* $\mathcal{G}$ *be a minimal weighted tree equivalent to* [1]. *Then* $\mathcal{G}$ *is linear and:*

1. *If* $|\mathcal{G}|=1$ *then* $\mathcal{G}=[1]$.
2. *If* $|\mathcal{G}|=2$ *then* $\mathcal{G}=[0, \alpha]$, *some* $\alpha\in\boldsymbol{Z}\setminus\{-1\}$.
3. *If* $|\mathcal{G}|>2$ *then* $\mathcal{G}$ *has exactly two vertices with nonnegative weights, these vertices are linked and exactly one of them, say u, has weight zero. Moreover, u has two neighbours, say x and y, and* $\Omega(x)+\Omega(y)=-1$.

Definition 5.17.

1. For a weighted graph $\mathcal{G}$, the symbol $\mathcal{G}<-1$ is an abbreviation for the statement "every vertex of $\mathcal{G}$ has weight less than -1".
2. Let $\mathcal{G}$ be a weighted tree and v a vertex of $\mathcal{G}$. We say that v is a *special vertex* if the number of branches $\mathcal{B}$ of $\mathcal{G}$ at v such that $\mathcal{B}<-1$ is at least two.

Corollary 5.18. *Let* $\mathcal{G}\sim[1]$ *and suppose that* v *is a special vertex of* $\mathcal{G}$. *Then*

$$\Omega(v)+|\mathcal{N}_{\mathcal{G}}(v)|\leq 1\,.$$

Proof. Let $n=|\mathcal{N}_{\mathcal{G}}(v)|$ and let $\mathcal{B}_1$, $\mathcal{B}_2$ be branches of $\mathcal{G}$ at v such that $\mathcal{B}_1<-1$ and $\mathcal{B}_2<-1$. By (5.11), we may consider the tree $\mathcal{G}'$ obtained from $\mathcal{G}$ by letting v absorb all branches other than $\mathcal{B}_1$ and $\mathcal{B}_2$. Clearly, the weight $\Omega'(v)$ of v in $\mathcal{G}'$ satisfies $\Omega'(v)\geq\Omega(v)+n-2$, since $n-2$ branches of $\mathcal{G}$ at v disappeared in the contraction. Since $|\mathcal{G}'|>2$, $\mathcal{G}'$ is not minimal by (5.16), i.e., $\Omega'(v)=-1$ and we get the desired inequality.

Desigularization of Divisors. Say that a nonsingular surface X *has no loops at infinity* if no element of $\mathcal{G}[X]$ has loops (see (5.7)). We now make a simple observation that turns out to be very useful.

Lemma 5.19. *Let S be a complete nonsingular surface and suppose that $D\in\mathrm{Div}(S)$ is reduced and effective. Then there exists a sequence $S_m\to\cdots\to S_0=S$ of monoidal transformations such that, if E_i is the exceptional curve created in $S_i\to S_{i-1}$ and if we define for $G\in\mathrm{Div}(S)$*

$$\begin{cases} G^0=G\in\mathrm{Div}(S_0) \\ G^i=(\textit{strict transform of } G^{i-1})+E_i\in\mathrm{Div}(S_i), \qquad 1\leqq i\leqq m, \end{cases}$$

then $D^m\in\mathrm{Div}(S_m)$ has s.n.c.. Assume that m is minimal with respect to that property. Then the centers of the monoidal transformations are infinitely near (i.n.) D, $S_m\setminus\mathrm{supp}(D^m)\cong S\setminus\mathrm{supp}(D)$ and, if $S\setminus\mathrm{supp}(D)$ has no loops at infinity, every E_i such that $E_i^2=-1$ in S_m is a branch point of $\mathcal{G}(S_m, D^m)$. Moreover, if $S\setminus\mathrm{supp}(D)\cong \boldsymbol{A}^2$ then:

1. *if $m\geqq 2$ then $P_i\in E_{i-1}$ $(2\leqq i\leqq m)$;*
2. *if $m\geqq 1$ then P_1 belongs to at least two irreducible components of D;*
3. *if $m\geqq 1$ and $D=A+B$, where A and B are effective divisors and B has s.n.c. in S, then P_i belongs to the strict transform of A in S_{i-1} $(1\leqq i\leqq m)$.*

Proof. Everything before the "Moreover" is very well known, except perhaps the last assertion (the verification of which we leave to the reader). We prove (1), (2), (3). Let's use the same notation for a curve and for its strict transform in any blown up surface. Since $\boldsymbol{A}^2$ has no loops at infinity,

$$\textit{every } E_i \textit{ such that } E_i^2=-1 \textit{ in } S_m \textit{ is a branch point of } \mathcal{G}(S_m, D^m).$$

If (1) doesn't hold then $\mathcal{G}(S_m, D^m)$ contains two branch points u, v of weight -1 such that u, v are not neighbours of each other. Contract $\mathcal{G}(S_m, D^m)$ to a linear weighted tree $\mathcal{L}$ (5.12); then u and v are still in $\mathcal{L}$ and one of the following holds:

- $\mathcal{L}$ contains vertices u, v with positive weights;
- $\mathcal{L}$ contains vertices u, v with nonnegative weights and not neighbours of each other.

Thus $\langle\mathcal{L}\rangle>1$ by (5.9), and this is absurd since every $\mathcal{G}\in\mathcal{G}[\boldsymbol{A}^2]$ has $\langle\mathcal{G}\rangle=1$. Hence (1) holds.

Proof of (2). By the above, E_m is a branch point of $\mathcal{G}(S_m, D^m)$, of weight -1, and no other E_i has weight $\geqq -1$ (in S_m). If P_1 belongs to only one component of D, all components of D are in the same branch of $\mathcal{G}(S_m, D^m)$ at E_m. Thus E_m is a "special vertex" (5.17) and we get a contradiction with (5.18).

Proof of (3). Since E_m is a branch point of $\mathcal{G}(S_m, D^m)$, P_m belongs to at least three components of $D^{m-1}=A+B^{m-1}$. Since B has s.n.c. in S, B^i has s.n.c. in S_i $(0\leqq i\leqq m)$ and P_m belongs to at most two components of B^{m-1}. Thus P_m belongs to (the strict transform of) A and, by (1), so do $P_1, \cdots, P_{m-1}$.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OTTAWA
OTTAWA, ONT.
CANADA, K1N6N5

References

[1] D. Daigle, Birational endomorphisms of the affine plane, Thesis, McGill University, Montréal, Québec, Canada, 1987.
[2] D. Daigle, Local trees in the theory of affine plane curves, to appear in this journal.
[3] T. Fujita, On Zariski problem, Proc. Japan Acad. **55A** (1979), 106-110.
[4] R. Ganong, Kodaira dimension of embeddings of the line in the plane, J. Math. Kyoto U., **25**-4 (1985), 649-657.
[5] J.E. Goodman, Affine open subsets of algebraic varieties and ample divisors, Ann. of Math., **89** (1969), 160-183.
[6] R. Hartshorne, Ample subvarieties of algebraic varieties. Lectures Notes in Math. **156**, Springer-Verlag, Heidelberg, 1970.
[7] S. Iitaka, On logarithmic Kodaira dimension of algebraic varieties, Complex Analysis and Algebraic Geometry, Iwanami (1977), 175-189.
[8] M. Miyanishi and T. Sugie, Affine surfaces containing cylinderlike open sets, J. Math. Kyoto U., **20** (1980), 11-42.
[9] J.A. Morrow, Minimal normal compactifications of $\mathbf{C}^2$, Rice U. Studies, **59**-1 (1973), 97-111.
[10] M. Nagata, On rational surfaces I, Mem. Coll. Sci. Kyoto, (A) **32** (1960), 351-370.
[11] C.P. Ramanujam, A topological characterization of the affine plane as an algebraic variety, Ann. of Math., **94** (1971), 69-88.
[12] K.P. Russell, Simple birational extensions of two dimensional affine rational domains, Compositio Math., **33** (1976), 197-208.
[13] K.P. Russell, On affine-ruled rational surfaces, Math. Ann., **255** (1981), 287-302.