Birational geometry and Jung's theorem, after Lamy

Christopher-Lloyd Simon

February 2018

Abstract

This document accompanies a presentation of Stephane Lamy's article [4] for our class mates of the master program in geometry. After introducing Lamy's motivations, we provide a few concepts in algebraic geometry so as to properly define morphisms between complex algebraic surfaces introduce the basics in birational geometry. We use this opportunity to set Lamy's methods in a broader perspective and describe the main results towards the classification of algebraic varieties up to birational transformations. Finally we sketch the main ideas in his proof. Most of the discussion is self contained but assumes in a few isolated places some background in analytic geometry and sheaves, which can be found in [1]. My thanks go to Etienne Ghys for letting me use the images from his book [3].

0 Introduction

The automorphism group $\mathrm{Aut}(\mathbb{C}^n)$ of the complex variety $\mathbb{C}^n$, whose elements are polynomials in the coordinates: $(P_1(x_1,\ldots,x_n),\ldots,P_1(x_1,\ldots,x_n))$, lies at the core of many open questions in complex analysis but also in Lie group theory and dynamics. For instance the famous Jacobian conjecture asks whether a polynomial map $f\colon \mathbb{C}^n \to \mathbb{C}^n$ with a non-zero constant Jacobian admits an inverse which is a polynomial map. The group of polynomial automorphisms is much better understood in dimension two thanks to the following structural result. Let us denote A the subgroup of affine elements (those extending to automorphisms of $\mathbb{CP}^2$) and by E the so called *elementary* automorphisms which are those who preserve the pencil of horizontal lines. In fromulae:

$$A = \{(x,y) \mapsto (a_1.x + b_1.y + c_1, a_2.x + b_2.y + c_2) \mid a_i, b_i, c_i \in \mathbb{C}, a_1.b_2 - a_2.b_1 \neq 0\}$$
$$E = \{(x,y) \mapsto (\alpha.x + P(y), \beta.y + \gamma) \mid \alpha, \beta \in \mathbb{C}^*, \gamma \in \mathbb{C}, P \in \mathbb{C}[X]\}$$

Theorem (Jung, 1942)**.** *$Aut(\mathbb{C}^n)$ is generated by the subgroups A and E.*

Theorem (Van der Kulk, 1953)**.** *If k is a field, the group of polynomial automorphisms $Aut(k^2)$ is the amalgamated product of A and E (over the canonical injection maps).*

There are various proofs of Jung's theorem but most of them proceed by induction on the degrees of the components of an automorphism. Lamy proposes to provide a geometric proof, based on remarks of Keller and Shafarevich, which remains in the spirit of the techniques available to the geometers in the 1900's.

Lamy's approach is to consider a polynomial automorphism as a birational transformation from the projective plane to itself and to study its indetermination points by successive blowups. This singularity at infinity contains indeed much information concerning the nature of the transformation. He is then led to investigate the behaviour of birational transformations between rational surfaces, and re-arranges the sequence of blowups and blowdowns given by Zariski's theorem to decompose the initial map into elementary and affine transformations.

1 Algebraic varieties

Affine complex algebraic varieties and regular maps

Objects An affine complex algebraic variety $V \in \mathbb{A}^n(\mathbb{C})$ is defined, after a choice of coordinates identifying the affine space with a vector space, as the vanishing locus $\{(x_1, \dots, x_n) \in \mathbb{C}^n \mid f_\alpha(x) = 0\}$ of a set of polynomial functions $f_\alpha \colon \mathbb{A}^n \to \mathbb{C}$. It is in fact the ideal generated by the polynomials f_α that defines the variety, and since the ring $A = \mathbb{C}[x_1, \dots, x_n]$ is Noetherian, this family can be chosen finite. We denote $\mathcal{V}(I)$ the common vanishing locus of the polynomial elements in the ideal I. The affine complex algebraic varieties are the closed sets of a topology on $\mathbb{A}^n$ named after Zariski. Indeed, for any family of ideals I_α and finite family J_β:

$$\bigcap_\alpha \mathcal{V}(I_\alpha) = \mathcal{V}\left(\sum_\alpha I_\alpha\right)$$
$$\bigcup_\beta \mathcal{V}(J_\beta) = \mathcal{V}\left(\prod_\beta J_\beta\right)$$

For every subset $U \subset \mathbb{C}^n$, we can define the ideal $\mathcal{I}(U)$ of vanishing polynomials on U. The equality $\mathcal{V}(\mathcal{I}(U)) = U$ holds if and only if U is Zarisky closed (an algebraic variety). Moreover, Hilbert's Nullstellensatz says that for any ideal J of A, the equality $\mathcal{I}(\mathcal{V}(J)) = J$ holds if and only if the ideal J is radical, that is $J = \sqrt{J}$. This amounts to asking for the ring A/J to be free from nilpotent elements (called a reduced ring). A closed set V in $\mathbb{A}^n$ inherits the Zariski topology, and subvarieties of V correspond to ideals containing $\mathcal{I}(V)$. The variety V is called irreducible if its ideal I is prime. On the complex numbers this means that V is connected as an analytic variety. It is said to be smooth if the underlying complex variety is smooth. The coordinate ring $\mathbb{C}[V]$ is the ring of restrictions to V of polynomial functions in A. It is isomorphic to A/I. When V is irreducible, its function field is the field of fractions $\mathbb{C}(V) = \mathrm{Frac}(A/I)$, or equivalently the field of restrictions to V of those rational functions on $\mathbb{C}^n$ defined somewhere on V.

Morphisms So far we have defined some objets: affine complex algebraic varieties. Let us define the appropriate notion of morphism between them called regular maps [1]. Those morphisms should be such that the group of automorphisms of the algebraic variety $\mathbb{A}^n$ is the group of polynomial maps with a polynomial inverse. A map $f \colon V \to W$ between two closed subsets of $\mathbb{A}^n$ is called regular (or a morphism) if it is the restriction of a polynomial map on $\mathbb{A}^n$ sending V into W. A morphism is an isomorphism if it has an inverse, which means in our case that the polynomial map on $\mathbb{A}^n$ was an automorphism.

Projective complex algebraic varieties

Objects An affine variety $V \subset \mathbb{A}^n$ has a natural completion in the projective plane $\mathbb{P}^n = \mathbb{A}^n \sqcup \mathbb{P}^{n-1}$. We could define a projective complex algebraic variety as the completion of an affine variety. As the affine varieties were well defined up to invertible affine transformations (the choice of coordinates), the projective varieties are well defined up to projective transformation.

Analogously to the affine case, a projective complex algebraic variety can be defined, after a choice of coordinates identifying $\mathbb{P}^n$ with the set of lines in $\mathbb{C}^{n+1}$, as the vanishing locus of a family of homogenous polynomial functions $f_\alpha \colon \mathbb{C}^{n+1} \to \mathbb{C}$. There is also a correspondence between algebraic varieties and graded ideals of the graded algebra of homogeneous polynomial functions in $n+1$ variables. However this point of view is a little more subtle than it is in the affine case, and it is more convenient to see $\mathbb{P}^n$ as an "algebraic gluing" of affine spaces in order to talk about its subvarieties. In fact,

[1] Note that we use the term function when the value is in the field of definition $\mathbb{C}$ as opposed to the words map and transformation whose target can be varieties. A map is defined everywhere at the source as opposed to a transformation.

the modern point of view is to define algebraic varieties abstractly using local models which we glue together, as we do in differential and analytical geometry. But this formulation, although conceptually clean, involves knowledge of scheme theory.

Morphisms If we tried to define morphisms between projective complex algebraic varieties as we did in the affine case, as restrictions of endomorphisms of the projective space, we would only end up with constant maps and the identity. This is because a non constant holomorphic map is open, and if it is defined everywhere it is an isomorphism of complex projective space. Moreover, we defined our varieties up to projective linear tranformations and the automorphism group of $\mathbb{CP}^n$ as a complex variety is $\mathrm{PGL}_{n+1}(\mathbb{C})$ (a finite dimensional complex Lie group, whereas $\mathrm{Aut}(\mathbb{C}^n)$ is infinite dimensional). To avoid the use of schemes and reasonably define our morphisms, we rely on Chow's and Serre's GAGA theorems relating the categories of analytic spaces in $\mathbb{P}^n$ and of projective algebraic varieties.

Recall that a subset M of an open $U \in \mathbb{C}^n$ for the usual topology, is called a complex *analytic set* if it is closed in U, and locally defined at a point $x \in M$ by the vanishing locus of a family $f_\alpha \colon V \to \mathbb{C}$ of analytic functions defined on a neighborhood V of x in U. Since U is a ringed space for the sheaf $\mathcal{O}_U$ of analytic functions on U, M is endowed with a natural sheaf whose stalk at a point x is the ring of restrictions to M of the germs $\mathcal{O}_{U,x}$ of analytic functions on U at this point. This ring is isomorphic to $\mathcal{O}_V / (f_\alpha)$. Analytic sets as ringed spaces $(M, \mathcal{O}_M)$ for variable n, U and M make up our local models. A complex *analytic space* is a ringed space $(X, \mathcal{O}_X)$ over a topological Hausdorff space X which is locally isomorphic to one of our models. If it is everywhere smooth, it is called [2] an *analytic manifold*. Morphisms and isomorphims of analytic spaces are defined as morphisms and isomorphisms of the associated ringed spaces. An analytic subspace Y of X is a closed subset locally defined around $y \in Y$ as the vanishing locus of some analytic germs in $\mathcal{O}_{X,y}$.

The projective space $\mathbb{CP}^n$ is an analytic space and Chow's theorem asserts that every analytic subset of $\mathbb{P}^n$ is a projective complex algebraic variety. In fact Serre's GAGA theorems [5] establish an equivalence between the categories of closed complex projective subanalytic spaces and of complex projective algebraic varieties. Thus we can finally take our morphisms between projective varieties as given by analytic morphisms of the underlying analytic spaces.

2 Birational geometry

Birational transformations between surfaces

Let X and Y be projective complex algebraic varieties. A *rational transformation* between them, denoted $\varphi \colon X \dashrightarrow Y$, is given by a morphism from a non empty (Zarisky) open set $U \subset X$ to Y. Such a map can be written in coordinates with rational functions. When $U = X$, the transformation is defined everywhere so it is a morphism. When X is a surface, we can always extend U to the complement of a finite set of points, and consistently define the *strict transform* by φ of an algebraic curve C as the closure of $\varphi(C \cap U)$; this could be a single point.

A *birational transformation* is a rational transformation φ admitting an inverse. Such a transformation induces an isomorphism between a non empty open set $U' \subset U$ and an open subset of Y. This time U' and its image are the complement of algebraic varieties with codimension at least one. So loosely speaking, birational surfaces are isomorphic surfaces outside of codimension one subspaces. A variety is called *rational* if it is birationally equivalent to the complex plane. Note that a birational transformation which is a morphism may not be an isomorphism if it collapses curves to points, as we shall see next for blowdown maps.

[2] In English terminology, an analytic space (and set) can have singularities (as for an algebraic variety) whereas a manifold is smooth. In French we respectively use the terms "espace (et ensemble) analytique" and "variété analytique". The french word "variété" (if not followed by "singulière") usually assumes smoothness except in the algebraic category.

Quadratic transforms and the Cremona group

Let us describe the group of birational transformations from $\mathbb{P}^2$ to itself called the complex Cremona group. It obviously contains all projective linear transformations. Here's another one, called the *standard quadratic transform*, defined in homogeneous coordinates by:

$$\sigma : [x : y : z] \dashrightarrow [yz : zx : xy]$$

This transformation is well defined outside of $A = [1 : 0 : 0]$, $B = [0 : 1 : 0]$, $C = [0 : 0 : 1]$ and induces an automorphism of the plane minus the three lines $x = 0$, $y = 0$, $z = 0$. It sends the line $z = 0$ to the point C, any line passing through C to another such line, and a generic line to a conic passing by the three points A, B and C. It also sends the pencil of lines through its fixed point $M = [1 : 1 : 1]$ to the pencil of conics passing through A, B, C and M, as shown in the following (real version) pictures:

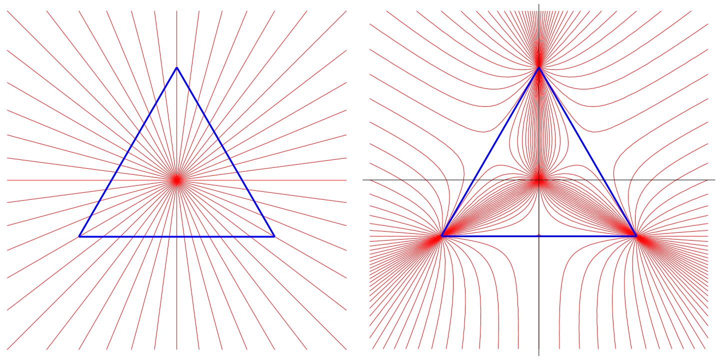

Figure 1: Pencils of lines and conics

Max Noether proved that the Cremona group is generated by the subgroup of projective linear transformations and the standard quadratic transform. He used this to resolve the singularities of projective curves and showed that any singular curve in the projective plane can be transformed by a Cremona automorphism into a curve whose only singularities are transversal double points.

Divisors and intersection numbers

The group of divisors $\mathrm{Div}(V)$ of a complex projective algebraic variety V is the free abelian group over the base of codimension one irreducible subvarieties of V. On a complex surface S, its elements are thus formal linear combinations of connected (possibly singular) algebraic curves $D = \sum_i \lambda_i C_i$. The fundamental class of a subvariety is a 2-cycle, so the group of divisors maps into the second homology group. Moreover, the cup product on the cohomology ring of the closed complex manifold S gives rise, by Poincaré duality, to an intersection form on its second homology group. This intersection form can thus be defined on the group of divisors, it is denoted $(D_1 \cdot D_2)$. For transversal curves C_1, C_2 this is simply their number of intersections. Note that the self intersection of a curve might be negative, this means that it cannot be continuously deformed along its normal bundle. To a meromorphic function $f \colon S \to \mathbb{P}^1$ we can associate a so called *principal divisor* (f) defined as the vanishing locus of f (with multiplicities for coefficients) minus its pole locus. The set of principal divisors form a normal subgroup of $\mathrm{Div}(S)$, and the intersection form descends to the corresponding quotient: the Picard group of S.

Blowing up surfaces

Let S be a surface containing a point p. There exists a surface $\tilde{S}$ and a blowdown map $\pi \colon \tilde{S} \to S$ such that $E = \pi^{-1}(p)$ is isomorphic to $\mathbb{P}^1$ and π induces an isomorphism from $\tilde{S} \setminus E$ to S. Moreover $\tilde{S}$ and

π are unique up to isomorphism. The map π is called the blowdown map, or the contraction of E, the surface $\tilde{S}$ is the blowup of S at p and E is called the exceptional divisor.

The blowup of $\mathbb{C}^2$ at the origin can be defined as an algebraic surface in $\mathbb{C}^2 \times \mathbb{CP}^1$: it is the closure of the graph of the meromorphic function $(x, y) \to u/v$. This yields the local model for a smooth surface at p. Topologically, $\tilde{S}$ is obtained by removing a small polydisc arround p and gluing back $\mathbb{P}^1 \times \Delta$ along its boundary. Figure 2 is a visual description in the real situation.

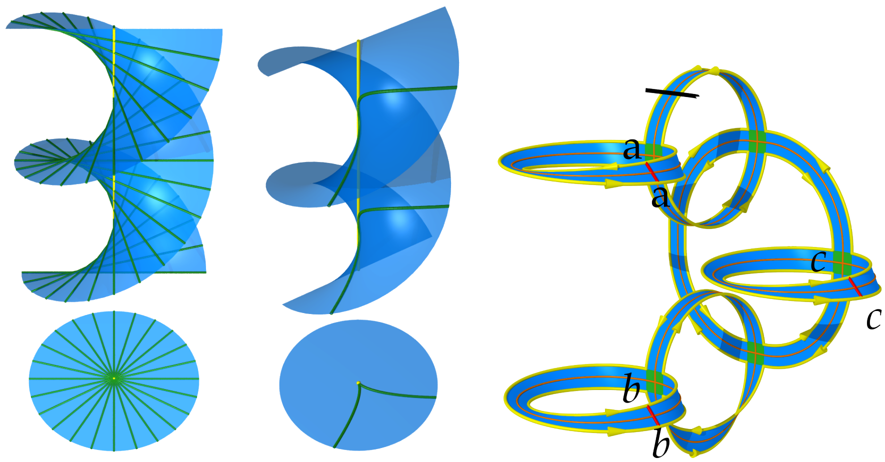

Figure 2: Blowing up a pencil of lines, a singular curve, and several times

If C is a curve in S passing through p, we denote $\tilde{C}$ its strict transform by π^{-1}, that is the closure of $\pi^{-1}(C)$; and by $\pi^* C$ its *total transform* which is $\tilde{C} + E$ if C is smooth at p. As the 2 suggests, blowups also enable to resolve singularities at p. In fact, any projective singular complex algebraic curve has a strict transform, lying in a surface obtained by several successive blowups of $\mathbb{P}^2$, which is a smooth curve. The blowup map preserves intersection numbers: for divisors D_1, D_2 of S we have $(\pi^* D_1 \cdot \pi^* D_2) = (D_1 \cdot D_2)$ and $(E \cdot \pi^* D_1) = 0$. From this we deduce $E^2 = -1$ and for a smooth curve through p we have $\tilde{C}^2 = C^2 - 1$.

Classification of varieties up to birationality

Lamy's methods can be set in the more general context of classifying surfaces up to birationality. We introduced projective varieties because according to another lemma due to Chow's, they suffice for the classification: every algebraic variety is birational to a projective variety. Moreover and much more difficult is the resolution of singularities theorem we owe to Hironaka, which says that we can concentrate on *smooth* varieties: every complex variety is biratianally equivalent to a smooth projective variety.

Now smooth projective complex algebraic varieties and rational transformations form a category which is equivalent to the category of function fields over the complex numbers and their morphisms. The equivalence is obtained by associating to a variety its field of rational functions. The reciprocal construction associates to a function field its set of valuations. It was first given for curves by Dedekind and Weber and later on for more general varieties by Weil. In particular, two surfaces are birationally equivalent if and only if their function fields are isomorphic.

For curves, birationality actually implies isomorphism, so the classification boils down to the study of moduli spaces of complex structures on Riemann surfaces of a given genus. Those are well understood since the final achievement of Koebe and Poincaré with their unifomization theorem [2].

Lamy considers smooth projective complex algebraic surfaces, and in fact he only needs to work within rational surfaces. For surfaces, birational transformations are well described by Zarisky's lemma.

Lemma (Zarisky, 1944)**.** *Every birational transformation between surfaces $\varphi\colon X \dashrightarrow Y$ can be obtained as a composition ob blowups and blowdowns of curves: there exists a surface M and sequences of blowdown maps π_1 and π_2 such that the following diaram commutes:*

$$\begin{tikzcd} & M \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ X \arrow[rr, dashed, "\varphi"] & & Y \end{tikzcd}$$

Moreover, there is a unique minimal choice of (M, π_1, π_2) up to isomorphism, satisfying the universal property that every other (M', π_1', π_2') factors uniquely through the minimal one with $f\colon M' \to M$.

The collection of points blown up by π_1 is the set $\mathrm{ind}(\varphi)$ of *indetermination points* of g, and those points belonging to X form the the set of *proper* indetermination points.

A minimal model of a birational class of surfaces is a smooth surface which is not the blowup of another smooth surface. By Castelnuovo's contraction theorem of self intersection -1 curves, this is equivalent to saying it does not contain such a curve. Every surface is birational to (at least one) minimal surface; so it is enough to classify those. This is the objective of the minimal model program for surfaces and it is well understood. In fact, surfaces up to birationality are completely described by Enriques Kodaira's classification, but we shall not go into so much detail. Let's just mention that many theorems, such as Zarisky's lemma, fall in defect in higher dimensions; and the general classification up to birationality (as well as the minimal model program) is still an active field of research.

3 Lamy's work

Sketch of his proof of Jung's theorem

Lamy's point of view is to consider a polynomial automorphism $g \in \mathrm{Aut}(\mathbb{C}^2)$ as birational transformation of the projective plane:

$$[x : y : z] \mapsto [z^n g_1(x/z, y/z) : z^n g_2(x/z, y/z) : z^n]$$

and reasons by induction on its number of indetermination points $\#\mathrm{ing}(g)$.

He considers more generally birational transformations $g\colon X \dashrightarrow \mathbb{P}^2$ which are obtained after a sequence of blow ups and contractions of a polynomial automorphism. This means that there is a decomposition $X = \mathbb{C}^2 \sqcup D$ where D is a union of irreducible curves, called the divisor at infinity, and $\mathbb{P}^2 = \mathbb{C}^2 \sqcup L$ where L is the line at infinity. In this precise configuration, much more information can be deduced from Zarisky's lemma.

Lemma (Lamy)**.** *Let $g\colon X \dashrightarrow \mathbb{P}^2$ be a birational transformation coming from a polynomial automorphism. Suppose g is not a morphism. Then*

1 The transformation g has only one proper indetermination point, it is located on the divisor at infinity.

2 The indetermination points $\{p_1, \dots, p_s\}$ of g satisfy, choosing p_1 to be the proper one, that each p_j is on the exceptional divisor obtained after blowing up p_{j-1}.

3 Every irreducible curve contained in the divisor at infinity of X is contracted to a point by g.

4 The first contracted curve by π_2 in the Zarisky diagram associated to g is the strict transform of a curve contained in the divisor at infinity of X.

5 In particular if $X = P^2$, the first curve contracted by π_2 is the strict transform of the line at infinity.

Let's get back to g: it has only one proper indetermination point, located on the line at infinity and precomposing with an affine transformation we can suppose it is $[0:0:1]$. His strategy is then to find an elementary automorphism φ such that $g \circ \varphi^{-1}$ has strictly less indtermination points (when there are no more indetermination points we are left with an affine map). To do so he performs multiple blowups and blowdows in the divisor at infinity and reorders them to get a composition of transformations preserving a pencil of lines through the indetermination point such that the composition has less indetermination points: this ensures that φ is indeed conjugate by an affine map to an elementary automorphism.

Amalgamated product structure and generalisation over a general field

As announced in the introduction, we can easily deduces Van der Kulk's theorem over the complex numbers from Jung's theorem. One has to show (without loss of generality) that a product $h = a_1 \circ e_1 \circ \cdots \circ a_n \circ e_n$ with $a_j \in A \setminus E$ and $e_j \in E \setminus A$ cannot be the identity. This is because every e_j contracts the line at infinty of $\mathbb{P}^2$ to the point $[0:0:1]$, so h contracts it to its image by a_1, thus $h \neq id$. Another proof providing immediately the amalgamated product structure can be derived from Bass-Serre theory. First consider all smooth rational surfaces (they are the algebraic models having $\mathbb{C}(X_1, X_2)$ as field of rational functions) and construct an oriented graph whose verticies given by blowup and blowdown transformations. The graph is connected by Zarisky's lemma. The natural action by the automorphism group is transitive on edges and non edge-reversing. One just needs to compute the stabiliser of an edge and of its vertices to get the amalgamated product structure. Finally Lamy notes that his proof is valid on any algebraically closed field as the techniques he uses (mainly the properties of the intersection form and Zarisky's lemma), are still available. He therefore deduces it over any field by proving inertia of the property by field extension.

References

[1] Henry Cartan, *Sur les fonctions de plusieurs variables complexes: les espaces analytiques*, Congrès International des Mathématiciens, Edinburgh (1958), 33–52.

[2] Henri-Paul de Saint-Gervais, *Uniformisation des surfaces de riemann - retour sur un théorème centenaire*, ENS Editions, 2011.

[3] Etienne Ghys, *A singular mathematical promenade*, ENS Editions, 2017.

[4] Stéphane Lamy, *Une preuve géométrique du théorème de Jung*, L'Enseignement Mathematique **48** (2002), 291–315.

[5] Jean-Pierre Serre, *Géometrie algébrique et géometrie analytique*, Anales de l'institu Fourier **6** (1956), 1–42.