

NORMAL SUBGROUPS IN THE CREMONA GROUP

SERGE CANTAT AND STÉPHANE LAMY

ABSTRACT. We show that the Cremona group of all birational transformations of the complex projective plane is not a simple group. The strategy is to use hyperbolic geometry and geometric group theory to produce elements in the Cremona group which generate non trivial normal subgroups.

RÉSUMÉ. Nous montrons que le groupe de Cremona, formé des transformations birationnelles du plan projectif complexe, n'est pas un groupe simple. Pour cela nous produisons des éléments engendrant des sous-groupes normaux non triviaux à l'aide de théorie géométrique des groupes et de géométrie hyperbolique.

CONTENTS

1. Introduction	1
Part A. Hyperbolic Geometry and Normal Subgroups	6
2. Gromov hyperbolic spaces	6
3. Hyperbolic spaces with constant negative curvature.	26
Part B. Algebraic Geometry and the Cremona Group	30
4. The Picard-Manin space	30
5. Tight birational maps	36
6. Complements	50
References	53

1. INTRODUCTION

1.1. The Cremona group.

The Cremona group in n variables over a field \mathbf{k} is the group $\text{Bir}(\mathbb{P}_{\mathbf{k}}^n)$ of birational transformations of the projective space $\mathbb{P}_{\mathbf{k}}^n$ or, equivalently, the group of \mathbf{k} -automorphisms of the field $\mathbf{k}(x_1, \dots, x_n)$ of rational functions of $\mathbb{P}_{\mathbf{k}}^n$. In

Date: April 2010 (preliminary version).

The second author was supported by a Marie Curie Intra European Fellowship.

dimension $n = 1$, the Cremona group coincides with the group $\text{Aut}(\mathbb{P}_{\mathbf{k}}^1)$ of regular automorphisms of $\mathbb{P}_{\mathbf{k}}^1$, that is with the group $\text{PGL}_2(\mathbf{k})$ of linear projective transformations. When $n \geq 2$, the group $\text{Bir}(\mathbb{P}_{\mathbf{k}}^n)$ is infinite dimensional; for example, it contains algebraic groups of arbitrarily large dimension.

In this article, we study the Cremona group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ over the field of complex numbers, and prove that *the group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ is not a simple group*¹. This answers one of the main open problems concerning the algebraic structure of this group (see Mumford’s questions in [Mum76]).

The proof makes use of geometric group theory and isometric actions on infinite dimensional hyperbolic spaces.

1.2. Generic elements.

Let $[x : y : z]$ be homogeneous coordinates for the projective plane $\mathbb{P}_{\mathbf{C}}^2$. Let f be a birational transformation of $\mathbb{P}_{\mathbf{C}}^2$; there are three homogeneous polynomials P, Q , and $R \in \mathbf{C}[x, y, z]$ of the same degree d , and without common factor, such that

$$f[x : y : z] = [P : Q : R].$$

The degree d is called the **degree** of f , and is denoted by $\deg(f)$. The space $\text{Bir}_d(\mathbb{P}^2(\mathbf{C}))$ of birational transformations of degree d is a quasi-projective variety. For instance, the automorphism group $\text{Aut}(\mathbb{P}_{\mathbf{C}}^2) = \text{PGL}_3(\mathbf{C})$ corresponds to the degree 1 case, and birational transformations of degree 2 form an irreducible variety of dimension 14 (see [DC09]).

We define the de Jonquières group \mathbf{J} as the group of birational transformations of the plane $\mathbb{P}_{\mathbf{C}}^2$ which preserve the pencil of lines through the point $[1 : 0 : 0]$. Let \mathbf{J}_d be the subset of \mathbf{J} made of birational transformations of degree d , and let \mathbf{V}_d be the subset of $\text{Bir}_d(\mathbb{P}^2(\mathbf{C}))$ whose elements are compositions $h_1 \circ f \circ h_2$ where h_1 and h_2 are automorphisms of $\mathbb{P}_{\mathbf{C}}^2$ and f is in \mathbf{J}_d . The dimension of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ is equal to $4d + 6$ and \mathbf{V}_d is the unique irreducible component of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ of maximal dimension (see [Ngu09]).

On an algebraic variety W , a property is said to be **generic** if it is satisfied on the complement of countably many Zariski closed subsets of W of codimension ≥ 1 .

Theorem A. *There exists an integer $k \geq 1$ with the following property. If g is a generic birational transformation of $\mathbb{P}_{\mathbf{C}}^2$ of degree $\deg(g) \geq 2$, then g^k generates a normal subgroup of the Cremona group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ that does not contain any $f \in \text{Bir}(\mathbb{P}_{\mathbf{C}}^2) \setminus \{Id\}$ with $\deg(f) < \deg(g)^k$. One can take $k = 1$ for large enough degrees d .*

¹This result should hold for any algebraically closed field \mathbf{k} , see section 6.5

As a corollary, $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ contains an uncountable number of normal subgroups. To prove this statement we choose g in the unique irreducible component of maximal dimension V_d . Thus, Theorem A and its proof do not provide any information concerning generic elements of the remaining components of the variety $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$.

1.3. Automorphisms of Kummer and Coble surfaces.

If X is a rational surface, then $\text{Aut}(X)$ is conjugate to a subgroup of $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$. In §5.2, we study two classes of examples. The first one is a (generalization of) Kummer surface. Let $\mathbf{Z}[i] \subset \mathbf{C}$ be the lattice of Gaussian integers. We start with the abelian surface $A = E \times E$ where E is the elliptic curve $\mathbf{C}/\mathbf{Z}[i]$. The group $\text{SL}_2(\mathbf{Z})$ acts by linear automorphisms on A , and commutes with the order 4 homothety $\eta(x, y) = (ix, iy)$. The quotient A/η is a (singular) rational surface on which $\text{PSL}_2(\mathbf{Z})$ acts by automorphisms. Since A/η is rational, the conjugacy by any birational map $\phi: A/\eta \dashrightarrow \mathbb{P}_{\mathbf{C}}^2$ provides an isomorphism between $\text{Bir}(A/\eta)$ and $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$, and therefore an embedding of $\text{PSL}_2(\mathbf{Z})$ into $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$.

Theorem B. *There is an integer $k \geq 1$ with the following property. Let M be an element of $\text{SL}_2(\mathbf{Z})$ with trace $|\text{tr}(M)| \geq 3$. Let g_M be the automorphism of the rational Kummer surface A/η which is induced by M . Then, g^k generates a non trivial normal subgroup of the Cremona group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2) \simeq \text{Bir}(A/\eta)$.*

Theorems A and B provide examples of normal subgroups coming respectively from generic and from highly non generic elements in the Cremona group. In §5.2.3, we also describe automorphisms of Coble surfaces that generate non trivial normal subgroups of the Cremona group; Coble surfaces are quotient of K3 surfaces, while Kummer surfaces are quotient of abelian surfaces.

1.4. An infinite dimensional hyperbolic space (see §4).

The Cremona group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ acts faithfully on an infinite dimensional hyperbolic space which is the exact analogue of the classical finite dimensional hyperbolic space \mathbb{H}^n . This action is at the heart of the proofs of Theorems A and B.

To describe this action, let us consider all rational surfaces $\pi: X \rightarrow \mathbb{P}_{\mathbf{C}}^2$ obtained from the projective plane by successive blow-ups. If $\pi': X' \rightarrow \mathbb{P}_{\mathbf{C}}^2$ is obtained from $\pi: X \rightarrow \mathbb{P}_{\mathbf{C}}^2$ by blowing up more points, then there is a natural birational morphism $q: X' \rightarrow X$, defined by $q = \pi^{-1} \circ \pi'$, and the pullback operator q^* embeds the second cohomology group $H^2(X, \mathbf{R})$ into $H^2(X', \mathbf{R})$. The inductive limit of all these cohomology groups $H^2(X, \mathbf{R})$ is called the **Picard-Manin space** of $\mathbb{P}_{\mathbf{C}}^2$. This infinite dimensional vector space comes together

with an intersection form of signature $(1, \infty)$, induced by the intersection form on each $H^2(X, \mathbf{R})$. We obtain in this way an infinite dimensional Minkowski space. The set of elements $[\alpha]$ in this space with self intersection $[\alpha]^2 = 1$ is a hyperboloid with two sheets, one of them containing Chern classes of ample line bundles of rational surfaces; this connected component is an infinite hyperbolic space for the distance \mathbf{dist} defined in terms of the intersection form by

$$\cosh(\mathbf{dist}([\alpha], [\beta])) = [\alpha] \cdot [\beta].$$

Taking the completion of this metric space, we get a complete, hyperbolic space $\mathbb{H}_{\bar{\mathcal{Z}}}$. The Cremona group acts faithfully on the Picard-Manin space, preserves the intersection form, and acts by isometries on the hyperbolic space $\mathbb{H}_{\bar{\mathcal{Z}}}$.

1.5. Normal subgroups in isometry groups.

In Part **A**, we study the general situation of a group G acting by isometries on a δ -hyperbolic space \mathcal{H} . Let us explain the content of the main Theorem **2.9** in the particular case of the Cremona group $\mathrm{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ acting by isometries on the hyperbolic space $\mathbb{H}_{\bar{\mathcal{Z}}}$. There are three kinds of isometries of such a hyperbolic space, namely elliptic, parabolic, and hyperbolic isometries. A Cremona transformation $g \in \mathrm{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ determines a hyperbolic isometry g_* of $\mathbb{H}_{\bar{\mathcal{Z}}}$ if and only if the two following equivalent properties are satisfied:

- The sequence of degrees $\deg(g^n)$ grows exponentially fast:

$$\lambda(g) := \limsup_{n \rightarrow \infty} \left(\deg(g^n)^{1/n} \right) > 1;$$

- There is a g_* -invariant plane V_g in the Picard-Manin space that intersects $\mathbb{H}_{\bar{\mathcal{Z}}}$ on a curve $\mathrm{Ax}(g_*)$ (a geodesic line) on which g_* acts by a translation; more precisely, $\mathbf{d}(x, g_*(x)) = \log(\lambda(g))$ for all x in $\mathrm{Ax}(g_*)$.

The curve $\mathrm{Ax}(g_*)$ is uniquely determined and is called the axis of g_* . We shall say that an element g of the Cremona group is **tight** if it satisfies the following three properties:

- The isometry $g_* : \mathbb{H}_{\bar{\mathcal{Z}}} \rightarrow \mathbb{H}_{\bar{\mathcal{Z}}}$ is hyperbolic;
- There exists a positive number B such that: If f is an element of the Cremona group and if $f_*(\mathrm{Ax}(g_*))$ contains two points at distance B which are at distance at most 1 from $\mathrm{Ax}(g_*)$ then $f_*(\mathrm{Ax}(g_*)) = \mathrm{Ax}(g_*)$;
- If f is a Cremona transformation and $f_*(\mathrm{Ax}(g_*)) = \mathrm{Ax}(g_*)$, then $f g f^{-1} = g$ or g^{-1} .

Applied to the Cremona group, Theorem **2.9** gives the following statement.

Theorem C. *Let g be a tight element of the Cremona group $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$. Then there exists a positive integer k such that g^k generates a non trivial normal subgroup of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$.*

Theorems A and B can now be rephrased by saying that generic elements of $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$ (resp. automorphisms g_M on the Kummer surface) are tight elements of the Cremona group (see Propositions 5.14 and 5.19).

1.6. Description of the paper. The paper starts with the proof of Theorem C in the general context of δ -hyperbolic spaces (section 2), and explains how this general statement can be used in the case of isometry groups of spaces with constant negative curvature (section 3). Section 4 provides an overview on Picard-Manin space, the associated hyperbolic space $\mathbb{H}_{\bar{z}}$, and the isometric action of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ on this space. Algebraic geometry and geometric group theory are then put together to prove theorems A (§5.1) and B (§5.2). At the end of the paper we list some remarks and comments.

1.7. Notes and references. Huge progress have been made during the past decade towards a better understanding of the algebraic structure of the Cremona group. The Cremona group can be compared to classical complex algebraic groups, and also to diffeomorphism groups of smooth or real analytic compact manifolds. Both viewpoints lead to different questions.

Recently, Jérémy Blanc, Igor Dolgachev, Vasily Iskovskikh, and Jean-Pierre Serre proved interesting results on algebraic, for instance finite, subgroups of the Cremona group (see [Ser08, Bla09] and the references therein). There is a natural topology on the Cremona group which induces the Zariski topology on $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$; it has been proved by J. Blanc [Bla10] that closed normal subgroups are trivial. Julie Déserti proved that $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ is perfect, hopfian, and co-hopfian, as most groups of diffeomorphisms of compact manifolds, and that its automorphism group is generated by inner automorphisms and the action of automorphisms of the field of complex numbers (see [Dés06b, Dés06a] for a precise result). In particular, there is no obvious reason which explains why $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ is not simple.

References [Dés06a] and [Can07] show that infinite countable subgroups of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ with Kazhdan property (T) are conjugate to subgroups of $\text{Aut}(\mathbb{P}_{\mathbb{C}}^2)$; this shows that $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ behaves like a simple Lie groups, the rank of which is somewhere in between 1 and 2. Moreover, one of the most typical properties of linear groups, namely Tits alternative, is known to be satisfied in the Cremona group (see [Can07, Fav08]).

1.8. **Acknowledgment.** Thanks to Thomas Delzant, Koji Fujiwara, and Vincent Guirardel for interesting discussions regarding hyperbolic groups and geometric group theory, and to Jérémy Blanc, Dominique Cerveau, Julie Deserti, and Jean-Pierre Serre for nice discussions on the Cremona group.

Part A. Hyperbolic Geometry and Normal Subgroups

2. GROMOV HYPERBOLIC SPACES

This section is devoted to the proof of Theorem 2.9, a result similar to Theorem C but in the general context of isometries of Gromov hyperbolic spaces.

2.1. Basic definitions.

2.1.1. *Hyperbolicity.* Let (\mathcal{H}, d) be a metric space. If x is a base point of \mathcal{H} , the Gromov product of two points $y, z \in \mathcal{H}$ is

$$(y|z)_x = \frac{1}{2} \{d(y, x) + d(z, x) - d(y, z)\}.$$

The triangle inequality implies that $(y|z)_x$ is a non negative real number. The metric space (\mathcal{H}, d) is δ -**hyperbolic** in the sense of Gromov, where $\delta \geq 0$, if

$$(x|z)_w \geq \min\{(x|y)_w, (y|z)_w\} - \delta$$

for all x, y, w , and z in \mathcal{H} . Equivalently, the space \mathcal{H} is δ -hyperbolic if, for any $w, x, y, z \in \mathcal{H}$, we have

$$d(w, x) + d(y, z) \leq \max\{d(w, y) + d(x, z), d(w, z) + d(x, y)\} + \delta. \quad (2.1)$$

Two fundamental examples of δ -hyperbolic spaces are given by simplicial metric trees, or more generally by **R**-trees, since they are 0-hyperbolic, and by the n -dimensional hyperbolic space \mathbb{H}^n , or more generally CAT(-1) spaces, since both are δ -hyperbolic for $\delta = \log(3)$ (see [GdlH90, chapter 2], [CDP90]).

2.1.2. *Geodesics.* A **geodesic segment** from x to y in \mathcal{H} is an isometry γ from an interval $[s, t] \subset \mathbf{R}$ into \mathcal{H} such that $\gamma(s) = x$ and $\gamma(t) = y$. The metric space (\mathcal{H}, d) is a **geodesic space** if any pair of points in \mathcal{H} can be joined by a geodesic segment. Let $[x, y] \subset \mathcal{H}$ be a geodesic segment, and let $\gamma: [0, l] \rightarrow \mathcal{H}$ be an arc length parametrization of $[x, y]$, with $\gamma(0) = x$ and $\gamma(l) = y$. By convention, for $0 \leq s \leq l - t \leq l$, we denote by $[x + s, y - t]$ the geodesic segment $\gamma([s, l - t])$.

A **geodesic line** is an isometry γ from \mathbf{R} to its image $\Gamma \subset \mathcal{H}$. If x is a point of \mathcal{H} and Γ is (the image of) a geodesic line, a **projection** of x onto Γ is a point $z \in \Gamma$ which realizes the distance from x to Γ . Projections exist but are not necessarily unique (in a tree, or in \mathbb{H}^n , the projection is unique). If z

and z' are two projections of a point x on a geodesic Γ , the distance between z and z' is at most 2δ . In other words, the projection is unique up to an error of magnitude 2δ . In what follows, we shall denote by $\pi_\Gamma(x)$ any projection of x on Γ .

2.2. Classical results from hyperbolic geometry.

In what follows, \mathcal{H} will be a δ -hyperbolic and geodesic metric space.

2.2.1. *Approximation by trees.* In a δ -hyperbolic space, the geometry of finite subsets is well approximated by finite subsets of metric trees; we refer to [GdlH90] chapter 2, or [CDP90], for a proof of this basic and crucial fact.

Lemma 2.1 (Approximation by trees). *Let \mathcal{H} be a δ -hyperbolic space, and (x_0, x_1, \dots, x_n) a finite list of points in \mathcal{H} . Let X be the union of n geodesic segments, one from the base point x_0 to each other x_i , $1 \leq i \leq n$. Then there exist a metric tree T and a map $\Phi : X \mapsto T$ such that:*

- (1) Φ is an isometry from $[x_0, x_i]$ to $[\Phi(x_0), \Phi(x_i)]$ for all $1 \leq i \leq n$;
- (2) For all $x, y \in X$

$$d(x, y) - 2\delta \log_2(n) \leq d(\Phi(x), \Phi(y)) \leq d(x, y).$$

The map $\Phi : X \mapsto T$ is called an **approximation tree**. For the sake of simplicity, the distance in the tree is also denoted $d(\cdot, \cdot)$. However, to avoid any confusion, we stick to the convention that a point in the tree will always be written under the form $\Phi(x)$, with $x \in X$.

Convention. When $n \leq 4$ then $2 \log_2(n) \leq 4$, choosing θ in \mathbf{R}_+ such that $\theta \geq 4\delta$, we have

$$d(x, y) - \theta \leq d(\Phi(x), \Phi(y)) \leq d(x, y). \tag{2.2}$$

From now on we fix such a θ , and we always use the approximation lemma in this way, for at most 5 points (x_0, \dots, x_4) . Note that the first point in the list will always be taken as the base point. In the figures the base point will be denoted by a white circle. Since two segments with the same extremities are δ -close, the specific choice of the segments between x_0 and the x_i is not important. We may therefore forget which segment is chosen, and refer to an approximation tree as a pair (Φ, T) associated to (x_0, \dots, x_n) . In general, the choice of the segments between x_0 and the x_i is either clear, or irrelevant.

In the remaining of §2.2 some well known facts from hyperbolic geometry are listed, with complete proofs for the convenience of the reader. First, note the following immediate consequence of the approximation lemma.

Corollary 2.2. *Let $\Phi: X \mapsto T$ be an approximation tree for at most 5 points. Then*

$$(\Phi(x)|\Phi(y))_{\Phi(z)} - \frac{\theta}{2} \leq (x|y)_z \leq (\Phi(x)|\Phi(y))_{\Phi(z)} + \theta$$

for all x, y , and z in X . In particular $(x|y)_z \leq \theta$ as soon as $\Phi(z) \in [\Phi(x), \Phi(y)]$.

Proof. By definition of the Gromov product we have

$$\begin{aligned} 2(x|y)_z &= d(x, z) + d(y, z) - d(x, y) \\ &\geq d(\Phi(x), \Phi(z)) + d(\Phi(y), \Phi(z)) - d(\Phi(x), \Phi(y)) - \theta \\ &= 2(\Phi(x)|\Phi(y))_{\Phi(z)} - \theta. \end{aligned}$$

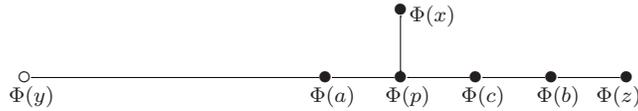
On the other hand

$$\begin{aligned} 2(x|y)_z &= d(x, z) + d(y, z) - d(x, y) \\ &\leq d(\Phi(x), \Phi(z)) + \theta + d(\Phi(y), \Phi(z)) + \theta - d(\Phi(x), \Phi(y)) \\ &= 2(\Phi(x)|\Phi(y))_{\Phi(z)} + 2\theta. \end{aligned}$$

□

Lemma 2.3 (Obtuse angle implies thinness). *Let $\Gamma \subset \mathcal{H}$ be a geodesic line; let x be a point of \mathcal{H} and $a \in \Gamma$ be a projection of x onto Γ . For all b in Γ and c in the segment $[a, b] \subset \Gamma$, we have $(x|b)_c \leq 2\theta$.*

Proof. Let y and z be two points of Γ such that a is the middle of $[y, z] \subset \Gamma$, b is contained in $[a, z]$, and $d(y, z) \geq 10\theta$. Let $\Phi: X \mapsto T$ be an approximation tree of (y, z, x) , where X is the union of the segment $[y, z] \subset \Gamma$ and a segment $[y, x]$. Let p be the point of Γ which is mapped onto the branch point of the tripod T by Φ .



We have

$$\begin{aligned} d(\Phi(x), \Phi(p)) + \theta &\geq d(x, p) \\ &\geq d(x, a) \\ &\geq d(\Phi(x), \Phi(a)) \\ &= d(\Phi(x), \Phi(p)) + d(\Phi(p), \Phi(a)), \end{aligned}$$

because $\Phi(a)$ is contained in the segment $\Phi([y, z])$. In particular $d(\Phi(p), \Phi(a))$ is bounded by θ ; from this and Corollary 2.2 we get

$$\begin{aligned} (x|b)_c &\leq (\Phi(x), \Phi(b))_{\Phi(c)} + \theta \\ &\leq d(\Phi(a), \Phi(p)) + \theta, \end{aligned}$$

so that $(x|b)_c$ is bounded by 2θ . \square

2.2.2. Shortening and weak convexity. The following lemma says that two geodesic segments are close as soon as their extremities are not too far. Lemma 3.1 makes this statement much more precise in the context of $\text{CAT}(-1)$ spaces.

Lemma 2.4 (A little shorter, much closer. See 1.3.3 in [Del96]). *Let $[x, y]$ and $[x', y']$ be two geodesic segments of \mathcal{H} such that*

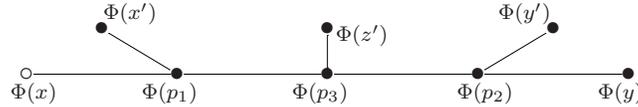
- (i) $d(x, x') \leq \beta$, and $d(y, y') \leq \beta$
- (ii) $d(x, y) \geq 2\beta + 4\theta$.

Then, the geodesic segment $[x' + \beta + \theta, y' - \beta - \theta]$ is in the 2θ -neighborhood of $[x, y]$.

Proof. Let z' be a point of $[x', y']$. Consider the approximation tree $\Phi: X \mapsto T$ of (x, y, x', z', y') . Note p_1 (resp. p_2 , resp. p_3) the points in $[x, y]$ such that $\Phi(p_1)$ (resp. $\Phi(p_2)$, resp. $\Phi(p_3)$) is the branch point of the tripod $(\Phi(x), \Phi(x'), \Phi(y)) \subset T$ (resp. $(\Phi(x), \Phi(y'), \Phi(y))$, resp. $(\Phi(x), \Phi(z'), \Phi(y))$). One easily checks that p_3 is contained in $[p_1, p_2]$ as soon as

$$\min(d(z', x'), d(z', y')) \geq \frac{\theta}{2} + \beta.$$

When z' is in $[x' + \beta + \theta, y' - \beta - \theta]$ this inequality is satisfied. In particular, the tree T looks as follows:



Therefore we have $d(\Phi(z'), \Phi(p_3)) = (\Phi(x')|\Phi(y'))_{\Phi(z')}$. By Corollary 2.2 we have $(\Phi(x')|\Phi(y'))_{\Phi(z')} \leq \frac{\theta}{2}$. We obtain

$$d(z', [x, y]) \leq d(z', p_3) \leq d(\Phi(z'), \Phi(p_3)) + \theta \leq 2\theta,$$

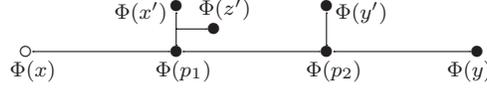
and $[x' + \beta + \theta, y' - \beta - \theta]$ is in the 2θ -neighborhood of $[x, y]$. \square

The following lemma is obvious in a $\text{CAT}(0)$ space (by convexity of the distance, see [BH99, p. 176]).

Lemma 2.5 (Weak convexity). *Let $[x, y]$ and $[x', y']$ be two geodesic segments of \mathcal{H} such that $d(x, x')$ and $d(y, y')$ are bounded from above by β . Then $[x', y']$ is in the $(\beta + 2\theta)$ -neighborhood of $[x, y]$.*

Proof. Pick $z' \in [x', y']$, and consider an approximation tree of (x, y, x', y', z') . Note $p_1 \in [x, y]$ (resp. p_2) the point such that $\Phi(p_1)$ (resp. $\Phi(p_2)$) is the branch

point of the tripod $(\Phi(x), \Phi(y), \Phi(x'))$ (resp. $(\Phi(x), \Phi(y), \Phi(y'))$).



We have

$$d(z', [x, y]) \leq d(\Phi(z'), \Phi(p_1)) + \theta \leq (\Phi(x')|\Phi(y'))_{\Phi(z')} + \beta + \theta \leq \beta + 2\theta$$

and the lemma is proved. \square

2.2.3. Canoeing to infinity. The following Lemma says that hyperbolic canoes never make loops; see 1.3.4 in [Del96] and 2.3.13 in [Hub06].

Lemma 2.6 (Canoeing). *Let y_0, \dots, y_n be a finite sequence of points in \mathcal{H} , such that*

- (i) $d(y_i, y_{i-1}) \geq 10\theta$ for all $1 \leq i \leq n$;
- (ii) $(y_{i+1}|y_{i-1})_{y_i} \leq 3\theta$ for all $1 \leq i \leq n-1$.

Then, for all $1 \leq j \leq n$,

- (1) $d(y_0, y_j) \geq d(y_0, y_{j-1}) + 2\theta$ if $j \geq 2$;
- (2) $d(y_0, y_j) \geq \sum_{i=1}^j (d(y_i, y_{i-1}) - 7\theta)$;
- (3) y_j is 5θ -close to any geodesic segment $[y_0, y_n]$.

Proof. The second assumption implies

$$d(y_{i+1}, y_{i-1}) \geq d(y_{i+1}, y_i) + d(y_{i-1}, y_i) - 6\theta \quad (2.3)$$

for all $1 \leq i \leq n-1$. Together with the first assumption, we get

$$d(y_{i+1}, y_{i-1}) \geq \max\{d(y_i, y_{i-1}), d(y_{i+1}, y_i)\} + 4\theta. \quad (2.4)$$

When applied to $i = 1$, equation (2.3) implies that

$$d(y_2, y_0) \geq d(y_2, y_1) - 7\theta + d(y_1, y_0) - 7\theta + 8\theta.$$

In particular, properties (1) and (2) are proved for $j = 2$. Now we prove (1) and (2) by induction, assuming that these inequalities have been proved for all indices up to j (included). By property (1) and equation (2.4) (for $i = j$) we have

$$\begin{aligned} d(y_0, y_j) + d(y_{j-1}, y_{j+1}) &> d(y_0, y_{j-1}) + 2\theta + d(y_j, y_{j+1}) + 4\theta \\ &\geq d(y_0, y_{j-1}) + d(y_j, y_{j+1}) + 6\theta. \end{aligned}$$

This, and the hyperbolic inequality (2.1), imply

$$d(y_0, y_j) + d(y_{j-1}, y_{j+1}) \leq d(y_0, y_{j+1}) + d(y_j, y_{j-1}) + \delta.$$

Hence

$$d(y_0, y_{j+1}) \geq d(y_0, y_j) + d(y_{j-1}, y_{j+1}) - d(y_j, y_{j-1}) - \delta.$$

From equation (2.4) with $i = j$ and the inequality $4\theta - \delta \geq 2\theta$ we get property (1); then equation (2.3) with $i = j$ implies

$$d(y_0, y_{j+1}) \geq d(y_0, y_j) + d(y_j, y_{j+1}) - (6\theta + \delta).$$

The induction hypothesis concludes the proof of (1) and (2).

We now prove assertion (3). First note that, reversing the labeling of the points, we get

$$(2') \quad d(y_n, y_j) \geq \sum_{i=j+1}^n (d(y_i, y_{i-1}) - 7\theta).$$

Let $[y_0, y_n]$ be a geodesic segment from y_0 to y_n and let $\Phi: X \mapsto T$ be an approximation tree for $(y_0, y_{j-1}, y_j, y_{j+1}, y_n)$, where X is the union of $[y_0, y_n]$ and three segments from y_0 to y_{j-1} , y_j , and y_{j+1} . First, we prove that

$$(\Phi(y_n)|\Phi(y_{j+1}))_{\Phi(y_0)} > (\Phi(y_n)|\Phi(y_j))_{\Phi(y_0)}. \quad (2.5)$$

By Corollary 2.2 it is sufficient to prove

$$(y_n|y_{j+1})_{y_0} > (y_n|y_j)_{y_0} + \frac{3\theta}{2}.$$

This inequality is equivalent to

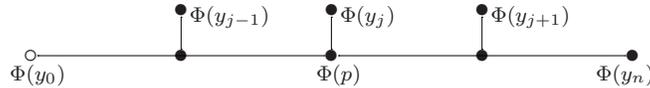
$$d(y_n, y_0) + d(y_{j+1}, y_0) - d(y_n, y_{j+1}) > d(y_n, y_0) + d(y_j, y_0) - d(y_n, y_j) + 3\theta,$$

that is, to

$$d(y_{j+1}, y_0) + d(y_n, y_j) > d(y_j, y_0) + d(y_n, y_{j+1}) + 3\theta,$$

and therefore follows from property (1) applied to $d(y_0, y_{j+1})$ and to $d(y_n, y_j)$ by reversing the labeling.

The inequality (2.5) shows that the approximation tree T has the following form:



Choose $p \in [y_0, y_n]$ such that $\Phi(p)$ is the branch point of the tripod generated by $\Phi(y_0)$, $\Phi(y_n)$, and $\Phi(y_j)$ in T . We have

$$\begin{aligned} d(y_j - p) &\leq d(\Phi(y_j), \Phi(p)) + \theta \\ &= (\Phi(y_{j+1})|\Phi(y_{j-1}))_{\Phi(y_j)} + \theta \\ &\leq (y_{j+1}|y_{j-1})_{y_j} + \frac{\theta}{2} + \theta \\ &\leq 3\theta + \frac{\theta}{2} + \theta. \end{aligned}$$

Thus the point y_j is 5θ -close to the segment $[y_0, y_n]$. \square

2.3. Rigidity of axis and non simplicity criterion.

Let G be a group of isometries of a δ -hyperbolic space $(\mathcal{H}, \mathbf{d})$, and g be an element of G . Our goal is to provide a criterion which implies that the normal subgroup $\langle\langle g \rangle\rangle$ generated by g in G does not coincide with G .

2.3.1. *Isometries.* Let $f \in \text{Isom}(\mathcal{H})$ be an isometry of a δ -hyperbolic space $(\mathcal{H}, \mathbf{d})$. The **translation length** $L(f)$ of f is the infimum of $\mathbf{d}(x, f(x))$ where x runs over all points of \mathcal{H} . We denote by $\text{Min}(f)$ the set of points that realize $L(f)$. Following [BH99, page 229] the isometry f is called

- (1) **elliptic** if f has a fixed point (in this case $L(f) = 0$ and $\text{Min}(f)$ is the set of fixed points);
- (2) **hyperbolic** if $\mathbf{d}(x, f(x))$ attains a strictly positive minimum;
- (3) **parabolic** if the infimum $L(f)$ is not realized by any point in \mathcal{H} .

A hyperbolic isometry f has an **invariant axis** if there is a geodesic line Γ such that $f(\Gamma) = \Gamma$. In this case, Γ is contained in $\text{Min}(f)$ and f acts as a translation of length $L(f)$ along Γ . If Γ and Γ' are two invariant axis, each of them is in the 2θ -tubular neighborhood of the other (apply Lemma 2.4). When f has an invariant axis, we denote by $\text{Ax}(f)$ any invariant geodesic line, even if such a line is a priori not unique.

Remark 2.7. If $(\mathcal{H}, \mathbf{d})$ is $\text{CAT}(0)$ and δ -hyperbolic, every hyperbolic isometry has an invariant axis (see [BH99, chap. II, theorem 6.8],). When \mathcal{H} is a tree or a hyperbolic space \mathbb{H}^n , the set $\text{Min}(g)$ coincides with the unique geodesic line it contains. If G is a hyperbolic group acting on its Cayley graph, any hyperbolic $g \in G$ admits a periodic geodesic line². Thus, there is a non trivial power g^k of g which has an invariant axis.

2.3.2. *Rigidity of geodesic lines.* If A and A' are two subsets of \mathcal{H} , the **intersection with precision** α of A and A' is the intersection of the tubular neighborhoods $\text{Tub}_\alpha(A)$ and $\text{Tub}_\alpha(A')$:

$$A \cap_\alpha A' = \{x \in \mathcal{H}; \mathbf{d}(x, A) \leq \alpha \text{ and } \mathbf{d}(x, A') \leq \alpha\}.$$

Let ε and B be positive real numbers. A subset A of \mathcal{H} is (ε, B) -**rigid** if $f(A) = A$ as soon as $f \in G$ satisfies

$$\text{diam}(A \cap_\varepsilon f(A)) \geq B.$$

² Sketch of proof: Consider $x, y \in \text{Min}(g)$. If Γ is an infinite geodesic in $\text{Min}(g)$, we have $x, y \in \text{Tub}_{2\theta}(\Gamma)$. By local compactness, there exists $N > 0$ such that all balls of radius 2θ in the Cayley graph \mathcal{H} contains at most N vertices. Thus there is at most N^2 segments from the ball $B(x, 2\theta)$ to the ball $B(y, 2\theta)$. Take $n = (2N)!$, we obtain that g^n fixes all these segments, and therefore also any geodesic $\Gamma \in \text{Min}(g)$.

The set A is ε -**rigid** if there exists a positive number $B > 0$ such that A is (ε, B) -rigid. If $\varepsilon' < \varepsilon$ and $A \subset \mathcal{H}$ is (ε, B) -rigid, then A is also (ε', B) -rigid. The converse holds for geodesic lines (or convex sets) when ε' is not too small:

Lemma 2.8. *Let $\varepsilon > 2\theta$. If a geodesic line $\Gamma \subset \mathcal{H}$ is $(2\theta, B)$ -rigid, then Γ is also $(\varepsilon, B + 4\varepsilon + 4\theta)$ -rigid.*

Proof. Let $B' = B + 4\varepsilon + 4\theta$. Suppose that $\text{diam}(\Gamma \cap_\varepsilon f(\Gamma)) \geq B'$. We want to show that $f(\Gamma) = \Gamma$. There exist $x, y \in \Gamma$, $x', y' \in f(\Gamma)$ such that $d(x, x') \leq 2\varepsilon$, $d(y, y') \leq 2\varepsilon$, $d(x, y) \geq B'$ and $d(x', y') \geq B'$. By Lemma 2.4, the segment

$$[u, v] := [x' + 2\varepsilon + \theta, y' - 2\varepsilon - \theta] \subset \Gamma$$

is 2θ -close to $[x, y]$, and we have $d(u, v) \geq B$. Thus $\Gamma \cap_{2\theta} f(\Gamma)$ contains the two points u and v and has diameter greater than B . Since Γ is $(2\theta, B)$ -rigid, we conclude that $\Gamma = f(\Gamma)$. \square

2.3.3. *Tight elements and the normal subgroup theorem.* We shall say that an element $g \in G$ is **tight** if

- g is hyperbolic and admits an invariant axis $\text{Ax}(g) \subset \text{Min}(g)$;
- the geodesic line $\text{Ax}(g)$ is 2θ -rigid;
- for all $f \in G$, if $f(\text{Ax}(g)) = \text{Ax}(g)$ then $fgf^{-1} = g$ or g^{-1} .

Note that if g is tight, then any iterate g^n , $n > 0$, is also tight. We now state the main theorem of Part A.

Theorem 2.9 (Normal subgroup theorem). *Let G be a group acting by isometries on a hyperbolic space \mathcal{H} . Suppose that $g \in G$ is tight, with a $(14\theta, B)$ -rigid axis $\text{Ax}(g)$. Let n be a positive integer with*

$$\frac{nL(g)}{20} \geq 60\theta + 2B.$$

Then any element $h \neq \text{Id}$ in the normal subgroup $\langle\langle g^n \rangle\rangle \subset G$ satisfies the following alternative: Either h is a conjugate of g^n or h is a hyperbolic isometry with translation length $L(h) > nL(g)$. In particular, if $n \geq 2$, the normal subgroup $\langle\langle g^n \rangle\rangle$ does not contain g .

Remark 2.10. Delzant in [Del96] gives a similar result in the case where G is a hyperbolic group and \mathcal{H} the Cayley graph of G . Our strategy of proof, presented in paragraph 2.5, follows the same lines as the one of Delzant.

In the definition of a tight element, we could impose a weaker list of hypothesis. The main point would be to replace $\text{Ax}(g)$ by a long quasi-geodesic segment (obtained, for example, by taking an orbit of a point x in $\text{Min}(g)$). Here, we take this slightly more restrictive definition because it turns out to be sufficient for our purpose, and makes the proof less technical; we refer to [Del96] or [DG10] for more general view points.

We now prove Theorem 2.9. First we need to introduce the notion of an admissible presentation.

2.4. Relators, neutral segments and admissible presentations.

Let $g \in G$ be a hyperbolic isometry with an invariant axis; let $\text{Ax}(g)$ be such an axis (see Remark 2.7), and $L = L(g)$ be the translation length of g . All isometries that are conjugate to g have an invariant axis: If $f = mgm^{-1}$, then $m(\text{Ax}(g))$ is an invariant axis for f .

2.4.1. *Axis, relators, and neutral segments.* Let $[x, y] \subset \mathcal{H}$ be an oriented geodesic segment, the length of which is at least 20θ . The segment $[x, y]$ is a **relator** if there exists an element h in G such that $[x, y]$ is contained in the tubular neighborhood $\text{Tub}_{7\theta}(h(\text{Ax}(g)))$. If $[x, y]$ is a relator, the conjugates $f = hgh^{-1}$ of g and $f^{-1} = hg^{-1}h^{-1}$ of g^{-1} have the same invariant axis $\Gamma = h(\text{Ax}(g))$ that almost contains $[x, y]$. Changing f to f^{-1} if necessary, we can assume that $\pi_\Gamma(y)$ and $f(\pi_\Gamma(x))$ lie on the same side of $\pi_\Gamma(x)$ in Γ ³. This assumption made, the isometry f is called a **support** of $[x, y]$ (and so f^{-1} is the support of the relator $[y, x]$). The segment $[x, y]$ is a **relator of size** p/q if furthermore $d(x, y) \geq \frac{p}{q}L$. We say that $[x, y]$ **contains a relator of size** p/q if there is a segment $[x', y'] \subset [x, y]$ which is such a relator.

A pair of points (x, y) is **neutral** if none of the geodesic segments $[x, y]$ between x and y contains a relator of size $11/20$; by a slight abuse of notation, we also say that the segment $[x, y]$ is neutral if this property holds (even if there are a priori several segments from x to y).

2.4.2. *Admissible presentations.* Let h be an element of the normal subgroup $\langle\langle g \rangle\rangle \subset G$. We can write h as a product $h = h_k h_{k-1} \cdots h_1$ where each h_j is a conjugate of g or its inverse:

$$\forall 1 \leq i \leq k, \quad \exists m_i \in G, \quad h_i = m_i g m_i^{-1} \text{ or } m_i g^{-1} m_i^{-1}.$$

By convention, each of the h_i comes with $m_i \in G$ and thus with an invariant axis $\text{Ax}(h_i) = m_i(\text{Ax}(g))$.

We fix a base point $x_0 \in \mathcal{H}$, and associate three sequences of points (a_i) , (b_i) , and (x_i) , $1 \leq i \leq k+1$, to the given base point x_0 and the factorization of h into the product of the $h_i = m_i g^{\pm 1} m_i^{-1}$. The definition requires several steps. First, for all $1 \leq i \leq k+1$,

- x_i is equal to $h_i(x_{i-1})$; in particular $x_{k+1} = h(x_0)$;
- a_i is the projection of x_{i-1} on the geodesic $\text{Ax}(h_i)$;
- b_i is equal to $h_i(a_i)$; in particular, both a_i and b_i are on $\text{Ax}(h_i)$.

³ Since $d(x, y) \geq 20\theta$, this assumption is meaningful: It does not depend on the choices of the projections of x and y on $\text{Ax}(f)$.

This is the first step toward the definition of the three sequences of points.

Remark 2.11. If the translation length L is larger than 480θ then, after a slight modification of the a_i and b_i , the following properties are satisfied:

- (1) Each $[a_i, b_i]$ is a subsegment of $\text{Ax}(h_i)$ of length at least $\frac{19}{20}L$;
- (2) $(x_{i-1}|b_i)_{a_i} \leq 2\theta$ and $(a_i|x_i)_{b_i} \leq 2\theta$;
- (3) Each segments $[x_{i-1}, a_i]$ and $[b_i, x_i]$ have length at least 10θ .

To obtain property (3) when $[x_{i-1}, a_i]$ is too small, we translate a_i towards b_i on a small distance (at most 12θ), and similarly we translate b_i towards a_i . The first property remains true as long as $\frac{L}{20} \geq 24\theta$. The second one follows from Lemma 2.3.

In a second step, we use this remark to move the points a_i and b_i along the geodesic line $\text{Ax}(h_i)$, and to define new sequences (a_i) , (b_i) , (x_i) that satisfy the three properties (1), (2), and (3). The definition of (x_i) has not been changed during this second step. The point a_i is no more equal to a projection of x_i on $\text{Ax}(h_i)$, but is at distance at most 12θ from such a projection; the point b_i is not equal to $h_i(a_i)$, but the distance between these two points is less than 24θ .

We say that $h_k \cdots h_1$ is an **admissible presentation** of h (with respect to the base point x_0) if one can construct the three sequences as above with the additional property that all pairs (x_{i-1}, a_i) , $1 \leq i \leq k$, are neutral.

Lemma 2.12. *Let x_0 be a base point in \mathcal{H} . Let g be a tight element of G , and h be an element of the normal subgroup generated by g . Assume that $28\theta < \frac{L}{20}$. Then h admits at least one admissible presentation.*

Remark 2.13. Thanks to this lemma, the following property can be added to the first three properties listed in Remark 2.11:

- (4) All pairs (x_{i-1}, a_i) and (b_i, x_i) are neutral.

Proof of Lemma 2.12. Start with a decomposition $h = h_k \cdots h_1$ with $h_i = m_i g^{\pm 1} m_i^{-1}$ and construct the sequences of points (a_i) , (b_i) , and (x_i) as above. Let \mathcal{I} be the set of indices $1 \leq i \leq k$ such that (x_{i-1}, a_i) is not neutral. Suppose \mathcal{I} is not empty (otherwise we are done). Our goal is to modify the construction in order to get a new decomposition of h which is admissible.

Changing the decomposition.– Pick $i \in \mathcal{I}$. Then there are two geodesic segments $[y, z] \subset [x_{i-1}, a_i]$ (with $y \in [x_{i-1}, z]$) and a conjugate $f = m g m^{-1}$ of g such that $[y, z]$ is 7θ -close to $\text{Ax}(f) = m(\text{Ax}(g))$ and $d(y, z) \geq \frac{11}{20}L$; we fix such a segment $[y, z]$ with a maximal length. Let y', z' be projections of y and z on $\text{Ax}(f)$; we have $d(y, y') \leq 7\theta$ and $d(z, z') \leq 7\theta$. If y' is contained in the

segment $[f(y'), z'] \subset \text{Ax}(f)$, we change f into f^{-1} . Replace h_i by the product of three conjugates of g

$$(h_i f^{-1} h_i^{-1}) h_i f.$$

This gives rise to a new decomposition of h as a product of conjugates of $g^{\pm 1}$, hence to new sequences of points. Let a'_i be a projection of x_{i-1} on $\text{Ax}(f)$. Concerning the sequence (x_i) , we have two new points $x'_i = f(x_{i-1})$ and $x'_{i+1} = h_i(x'_i)$; the point $x'_{i+2} = h_i f^{-1} h_i^{-1}(x'_{i+1})$ is equal to the old point $x_i = h_i(x_{i-1})$. Thus, the neutral pair (x_{i-1}, a_i) disappears and is replaced by three new pairs (x_{i-1}, a'_i) , (x'_i, a'_{i+1}) , and (x'_{i+1}, a'_{i+2}) (see Figure 1). Note that if any of these segments is small then it is automatically neutral; so without loss of generality we can assume that there was no need to move a'_i , a'_{i+1} or a'_{i+2} as in Remark 2.11.

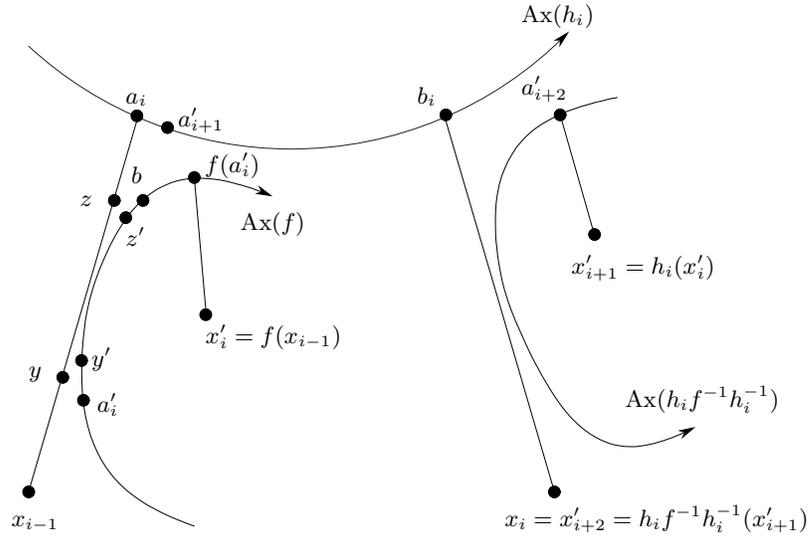


FIGURE 1.

Two estimates.— Our first claim is

$$(1) \quad d(x_{i-1}, a'_i) = d(x'_{i+1}, a'_{i+2}) \leq d(x_{i-1}, a_i) - \frac{10}{20}L.$$

To prove it, we write

$$d(x_{i-1}, a_i) \geq d(x_{i-1}, y) + d(y, z) \geq d(x_{i-1}, y) + \frac{11}{20}L,$$

and

$$d(x_{i-1}, a'_i) \leq d(x_{i-1}, y') \leq d(x_{i-1}, y) + d(y, y') \leq d(x_{i-1}, y) + 7\theta; \quad (2.6)$$

claim (1) follows because $7\theta < \frac{L}{20}$.

Our second claim is

$$(2) \quad d(x'_i, a'_{i+1}) \leq d(x_{i-1}, a_i) - \frac{1}{20}L.$$

To prove it, note b the projection of a_i on $\text{Ax}(f)$. Similarly as (2.6) we have

$$d(b, a_i) \leq d(z, a_i) + 7\theta. \quad (2.7)$$

There are two cases, according to the position of $f(a'_i)$ with respect to the segment $[a'_i, b] \subset \text{Ax}(f)$. If $f(a'_i)$ is in this segment, then

$$\begin{aligned} d(x'_i, a_i) &\leq d(f(x_{i-1}), f(a'_i)) + d(f(a'_i), b) + d(b, a_i) \\ &\leq d(x_{i-1}, a'_i) + d(a'_i, b) - L + d(b, a_i). \end{aligned}$$

Applying Lemma 2.3 for the triangles x_{i-1}, y', a'_i and a_i, z', b , and the inequalities $d(y, y'), d(z, z') \leq 7\theta$, we get $d(x_{i-1}, y) = d(x_{i-1}, a'_i) + d(y, a'_i)$ up to an error of 11θ , and similarly $d(a_i, z) = d(a_i, b) + d(z, b)$ up to 11θ . Hence, we get

$$\begin{aligned} d(x'_i, a_i) &\leq d(x_{i-1}, y) + d(y, z) + d(z, a_i) - L + 22\theta \\ &\leq d(x_{i-1}, a_i) - L + 22\theta. \end{aligned}$$

This concludes the proof of (2) in the first case, because $-L + 22\theta < -L/20$.

The second case occurs when b is in the segment $[a'_i, f(a'_i)] \subset \text{Ax}(f)$ (see Figure 1). In this case we have

$$\begin{aligned} d(x'_i, a_i) &\leq d(f(x_{i-1}), f(a'_i)) + d(f(a'_i), b) + d(b, a_i) \\ &\leq d(x_{i-1}, z) - \frac{11}{20}L + 7\theta + L - d(a'_i, b) + d(z, a_i) + 7\theta, \end{aligned}$$

and thus

$$d(x'_i, a_i) \leq d(x_{i-1}, a_i) - d(a'_i, b) + \frac{9}{20}L + 14\theta. \quad (2.8)$$

On the other hand, the triangular inequality implies

$$\begin{aligned} d(x_{i-1}, a'_i) + d(a'_i, b) + d(b, a_i) &\geq d(x_{i-1}, a_i) \\ &= d(x_{i-1}, y) + d(y, z) + d(z, a_i). \end{aligned}$$

and so

$$d(a'_i, b) \geq \frac{11}{20}L + d(x_{i-1}, y) - d(x_{i-1}, a'_i) + d(z, a_i) - d(b, a_i).$$

Using the inequalities (2.6) and (2.7) we obtain

$$d(a'_i, b) \geq \frac{11}{20}L - 14\theta.$$

Finally inequality (2.8) gives

$$d(x'_i, a'_{i+1}) \leq d(x'_i, a_i) \leq d(x_{i-1}, a_i) - \frac{2}{20}L + 28\theta$$

hence claim (2) because $28\theta < \frac{L}{20}$.

Conclusion.— Since the modification is local, and does not change the points a_j for $j \neq i$, we can perform a similar replacement for all indices i in \mathcal{I} . We obtain a new presentation for h and a new list of bad indexes \mathcal{I}' . Either this new list is empty, or by the two estimates (claim (1) and (2)) the maximum of the lengths $\mathbf{d}(x_{j-1}, a_j)$ over all the non neutral pairs (x_{j-1}, a_j) , $j \in \mathcal{I}'$, drops at least by $\frac{1}{20}L$. By induction, after a finite number of such replacements, we obtain an admissible presentation. \square

The following lemma provides a useful property of admissible presentations with a minimum number of factors h_i .

Lemma 2.14. *Let $h = h_k \cdots h_1$ be an admissible presentation with base point x_0 . If there exist two indices $j > i$ such that $h_j = h_i^{-1}$, then h admits an admissible presentation with base point x_0 and only $k - 2$ factors.*

Proof. We can assume that $j \geq i + 2$, otherwise the simplification is obvious. The decomposition for h is then

$$h = h_k \cdots h_{j+1} h_i^{-1} h_{j-1} \cdots h_{i+1} h_i h_{i-1} \cdots h_1.$$

Note that i can be equal to 1, and $j + 1$ can be equal to k . Then

$$h = h_k \cdots h_{j+1} (h_i^{-1} h_{j-1} h_i) (h_i^{-1} h_{j-2} h_i) \cdots (h_i^{-1} h_{i+1} h_i) h_{i-1} \cdots h_1$$

is an admissible presentation with base point x_0 and with $k - 2$ factors. \square

2.5. Proof of the normal subgroup theorem.

Let g be a tight element of G . Choose $B > 0$ a constant associated with the 14θ -rigidity of the axis of g . Choose $n \geq 1$ such that

$$\frac{n}{20}L(g) \geq 60\theta + 2B.$$

Let $L = nL(g)$ be the translation length of g^n . Note that $g^n \in G$ is tight, with a $(14\theta, B)$ -rigid axis $\text{Ax}(g^n) = \text{Ax}(g)$. Let h be a non trivial element of the normal subgroup $\langle\langle g^n \rangle\rangle$. Our goal is to prove $L(h) \geq L$, with equality if and only if h is conjugate to g^n .

Pick a base point x_0 such that $\mathbf{d}(x_0, h(x_0)) \leq L(h) + \theta$. Lemma 2.12 applied to g^n implies the existence of an admissible presentation with respect to the base point x_0 :

$$h = h_m \circ \cdots \circ h_1, \quad h_i = m_i g^{\pm n} m_i^{-1};$$

we assume that m is minimal among all such choices of base points and admissible presentations. Lemma 2.14 implies that h_j is different from h_i^{-1} for all $1 \leq i < j \leq m$.

Let (a_i) , (b_i) , (x_i) be the three sequences of points defined in §2.4; they satisfy the properties (1) to (4) listed in Remarks 2.11 and 2.13. Since the constructions below are more natural with segments than pair of points, and \mathcal{H} is not assumed to be uniquely geodesic, we choose geodesic segments between the points x_i , as well as geodesic segments $[a_i, b_i] \subset \text{Ax}(h_i)$.

We now introduce the following definition in order to state the key Lemma 2.15 (compare with Lemma 2.4 in [Del96]). A sequence of points $(c_{-1}, c_0, \dots, c_k, c_{k+1})$ in \mathcal{H} , and of a chain of segments $[c_i, c_{i+1}]$, $-1 \leq i \leq k$, is a **configuration of order** $k \geq 1$ for the segment $[x_0, x_j]$ if

- (i) $x_0 = c_{-1}$ and $x_j = c_{k+1}$;
- (ii) For all $0 \leq i \leq k$, and all $a \in [c_i, c_{i+1}]$, we have $(c_{i-1}|a)_{c_i} \leq 3\theta$. In particular $(c_{i-1}|c_{i+1})_{c_i} \leq 3\theta$;
- (iii) For all $0 \leq i \leq k+1$ we have $d(c_{i-1}, c_i) \geq 10\theta$;
- (iv) For all $0 \leq i \leq k$ the segment $[c_i, c_{i+1}]$ is either neutral or a relator, with the following rules:
 - (iv-a) There are never two consecutive neutral segments;
 - (iv-b) The last segment $[c_k, c_{k+1}] = [c_k, x_j]$ is neutral;
 - (iv-c) The second segment $[c_0, c_1]$ is a relator of size $18/20$ if $[c_1, c_2]$ is neutral (this is always the case when $k = 1$), and of size $17/20$ otherwise;
 - (iv-d) For the other relators $[c_{i-1}, c_i]$, with $i > 1$, the size is $5/20$ when $[c_i, c_{i+1}]$ is neutral and $4/20$ otherwise;
- (v) For all $0 \leq i \leq k$, if $[c_i, c_{i+1}]$ is a relator, then there is an index l with $1 \leq l \leq j$, such that h_l is the support of the relator $[c_i, c_{i+1}]$.

Note that properties (iv) and (v) do not concern the initial segment $[x_0, c_0]$, and that the size p/q of a relator $[c_i, c_{i+1}]$, $0 \leq i \leq k-1$, is equal to $18/20$, $17/20$, $5/20$, or $4/20$; moreover, this size is computed with respect to $L = L(g^n)$, and thus the minimum length of a relator $[c_i, c_{i+1}]$ is bounded from below by $4(60\theta + 2B)$.

Lemma 2.15. *For each $j = 1, \dots, m$, there exists $k \geq 1$ such that the segment $[x_0, x_j]$ admits a configuration of order k . Moreover if $j \geq 2$ and $k = 1$ then the first segment $[x_0, c_0]$ of the configuration has length at least $\frac{3}{20}L$.*

Proof. The proof is by induction on j , and uses the four properties that are listed in Remarks 2.11 and 2.13; we refer to them as property (1), (2), (3), and (4). Note that (2) and (3) enable us to apply Lemma 2.6; similarly, properties (ii) and (iii) for a configuration of order k show that Lemma 2.6 can be applied to the sequence of points in such a configuration.

Initialization.— When $j = 1$, we take $c_0 = a_1, c_1 = b_1$ and get a configuration of order 1. Indeed, by property (2), we have

$$(x_0|b_1)_{a_1} \leq 2\theta \text{ and } (a_1|x_1)_{b_1} \leq 2\theta,$$

and, by property (1), $[c_0, c_1]$ is a relator of size $19/20$ (it is a subsegment of $\text{Ax}(h_1)$). The segments $[x_0, c_0]$ and $[x_1, c_1]$ are neutral (property (4)) and of length at least 10θ (property (3)).

Suppose now that $[x_0, x_j]$ admits a configuration of order k . We want to find a configuration of order k' for $[x_0, x_{j+1}]$. As we shall see, the proof provides a configuration of order $k' = 1$ in one case, and of order $k' \leq k + 2$ in the other case.

Six preliminary facts.— Consider the approximation tree (Φ, T) of (x_j, x_0, x_{j+1}) . We choose $p \in [x_j, x_0]$ such that $\Phi(p) \in T$ is the branch point of the tripod T (with $p = x_j$ if T is degenerate). Let a (resp. b) be a projection of a_{j+1} (resp. b_{j+1}) on the segment $[x_j, x_{j+1}]$. By assertion (3) in Lemma 2.6 we have $d(a, a_{j+1}) \leq 5\theta$ and $d(b, b_{j+1}) \leq 5\theta$. Thus, the subsegment $[a, b] \subset [x_j, x_{j+1}]$ is contained in $[x_j + 5\theta, x_{j+1} - 5\theta]$; moreover, by Lemma 2.5, $[a, b]$ is a relator with support h_{j+1} .

For all $i \leq k + 1$, note c'_i a projection of c_i on the segment $[x_0, x_j]$. Note that, by Lemma 2.6, we have $d(c_i, c'_i) \leq 5\theta$, and the points $c'_i, -1 \leq i \leq k + 1$, form a monotonic sequence of points from x_0 to x_j along the geodesic segment $[x_0, x_j]$.

Let S_i be the interval of T defined by $S_i = [\Phi(a), \Phi(b)] \cap [\Phi(c'_i), \Phi(c'_{i+1})]$. The preimages of S_i by Φ are two intervals $I \subset [a, b]$ and $I' \subset [c'_i, c'_{i+1}]$ such that

$$\Phi(I) = \Phi(I') = S_i.$$

Fact 2.16. *If $[c_i, c_{i+1}]$ is neutral then $\text{diam}(S_i) \leq \frac{12}{20}L$.*

Since Φ is an isometry along the geodesic segments $[x_j, x_0]$ and $[x_j, x_{j+1}]$, we only have to prove $\text{diam}(I') \leq \frac{12}{20}L$. By Lemma 2.5 we know that $[c_i, c_{i+1}] \subset \text{Tub}_{7\theta}([c'_i, c'_{i+1}])$. Thanks to the triangular inequality, we can choose $J \subset [c_i, c_{i+1}]$ such that $J \subset \text{Tub}_{7\theta}(I')$ and $\text{diam}(J) \geq \text{diam}(I') - 14\theta$. The properties of the approximation tree imply $I' \subset \text{Tub}_\theta(I)$, and $I \subset \text{Tub}_{7\theta}(\text{Ax}(h_{j+1}))$. Thus $J \subset \text{Tub}_{15\theta}(\text{Ax}(h_{j+1}))$. Applying Lemma 2.4 (with $\beta = 16\theta$) and shortening J by 32θ (16θ at each end), we obtain $\text{diam}(J) \geq \text{diam}(I') - 46\theta$ and $J \subset \text{Tub}_{2\theta}(\text{Ax}(h_{j+1}))$. Thus J is a relator contained in the neutral segment $[c_i, c_{i+1}]$, hence

$$\frac{11}{20}L \geq \text{diam}(J) \geq \text{diam}(I') - 46\theta, \quad \text{and therefore} \quad \frac{12}{20}L \geq \text{diam}(I').$$

Fact 2.17. *If $[c_i, c_{i+1}]$ is a relator then $\text{diam}(S_i) \leq B$.*

By property (v), there exists an index l , with $1 \leq l \leq j$, such that $[c_i, c_{i+1}] \subset \text{Tub}_{7\theta}(\text{Ax}(h_l))$. Since $I' \subset \text{Tub}_{7\theta}([c_i, c_{i+1}])$, we get

$$I' \subset \text{Tub}_{14\theta}(\text{Ax}(h_l)).$$

On the other hand $I' \subset \text{Tub}_\theta(I)$ and $I \subset \text{Tub}_{7\theta}(\text{Ax}(h_{j+1}))$, thus

$$I' \subset \text{Tub}_{8\theta}(\text{Ax}(h_{j+1})).$$

If $\text{diam}(I') > B$, the rigidity assumption shows that h_{j+1} and h_l share the same axis, with opposite orientations; since g is tight, we obtain $h_{j+1} = h_l^{-1}$. By Lemma 2.14 this contradicts the minimality of the presentation of h . This proves that $\text{diam}(I') \leq B$.

Fact 2.18. *Suppose $[c_i, c_{i+1}]$ is a relator. Then*

$$\begin{aligned} \text{diam}([\Phi(x_j), \Phi(a)] \cap [\Phi(c'_i), \Phi(c'_{i+1})]) &\leq \frac{12}{20}L. \\ \text{diam}([\Phi(b), \Phi(x_{j+1})] \cap [\Phi(c'_i), \Phi(c'_{i+1})]) &\leq \frac{12}{20}L. \end{aligned}$$

The two inequalities are similar; we prove only the first one. By property (v), there exists an index l , with $1 \leq l \leq j$, such that h_l is the support of $[c_i, c_{i+1}]$. Let $K \subset [a, x_j]$ and $K' \subset [c'_i, c'_{i+1}]$ be two intervals such that

$$\Phi(K) = \Phi(K') = [\Phi(x_j), \Phi(a)] \cap [\Phi(c'_i), \Phi(c'_{i+1})].$$

We want to prove $\text{diam}(K) \leq \frac{12}{20}L$. Applying the triangular inequality, we can choose $J \subset [a_{j+1}, x_j]$ such that $J \subset \text{Tub}_{7\theta}(K)$ and $\text{diam}(J) \geq \text{diam}(K) - 14\theta$. Now we have $K \subset \text{Tub}_\theta(K')$, $K' \subset \text{Tub}_{7\theta}([c_i, c_{i+1}])$ and $[c_i, c_{i+1}] \subset \text{Tub}_{7\theta}(\text{Ax}(h_l))$. Thus $J \subset \text{Tub}_{22\theta}(\text{Ax}(h_l))$. Applying Lemma 2.4 with $\beta = 23\theta$, we shorten J by 46θ (23θ on each end) and obtain

$$J \subset \text{Tub}_{2\theta}(\text{Ax}(h_l)) \quad \text{and} \quad \text{diam}(J) \geq \text{diam}(K) - 60\theta.$$

The admissibility condition implies that $J \subset [a_{j+1}, x_j]$ is neutral, and therefore

$$\frac{11}{20}L \geq \text{diam}(J) \geq \text{diam}(K) - 60\theta, \quad \text{so that} \quad \frac{12}{20}L \geq \text{diam}(K).$$

Fact 2.19. *The segment $[\Phi(b), \Phi(a)]$ is not contained in the segment $[\Phi(c'_0), \Phi(x_j)]$.*

Let us prove this by contradiction, assuming $[\Phi(b), \Phi(a)] \subset [\Phi(c'_0), \Phi(x_j)]$. By Property (iv-a) and Fact 2.16, $[\Phi(b), \Phi(a)]$ intersects at least one segment $[\Phi(c'_i), \Phi(c'_{i+1})]$ for which $[c_i, c_{i+1}]$ is a relator; Fact 2.17 implies that $[\Phi(c'_i), \Phi(c'_{i+1})]$ is not contained in $[\Phi(a), \Phi(b)]$ (it must intersect the boundary points of $[\Phi(a), \Phi(b)]$). From this follows that $[\Phi(a), \Phi(b)]$ intersects at most

two relators and one neutral segment. Facts 2.16 and 2.17 now give the contradictory inequality $2B + \frac{12}{20}L \geq \frac{19}{20}L - 10\theta$.

Fact 2.20. *The segment $[\Phi(c'_0), \Phi(c'_1)]$ is not contained in the segment $[\Phi(x_{j+1}), \Phi(x_j)]$.*

Since $[a_{j+1}, x_j]$ is neutral and $[c_0, c_1]$ is a relator of size $\geq 17/20$, the segment $[\Phi(c'_0), \Phi(c'_1)]$ is not contained in $[\Phi(), \Phi(x_j)]$. Assume that $[\Phi(c'_0), \Phi(c'_1)] \subset [\Phi(x_{j+1}), \Phi(x_j)]$, apply Fact 2.19 and then Fact 2.17 and Fact 2.18; this gives the contradictory inequality $B + \frac{12}{20}L \geq \frac{17}{20}L - 10\theta$.

These last two facts imply $\Phi(b) \in [\Phi(x_{j+1}), \Phi(p)]$ and $\Phi(c'_0) \in [\Phi(x_0), \Phi(p)]$. Moreover, with this new property in mind, the proofs of Facts 2.19 and 2.20 give:

Fact 2.21. *The segment $[\Phi(b), \Phi(p)]$ has length at least $\frac{6}{20}L$, and the segment $[\Phi(c'_0), \Phi(p)]$ has length at least $\frac{4}{20}L$.*

In particular the tripod T is not degenerate at $\Phi(x_0)$ nor at $\Phi(x_{j+1})$.

The induction.— We now come back to the proof by induction. We distinguish two cases, with respect to the position of $\Phi(a)$ relatively to the branch point $\Phi(p)$.

• **First case.** $(\Phi(b)|\Phi(x_0))_{\Phi(a)} \leq \frac{1}{20}L - 12\theta$.

In other words we assume that either $\Phi(a) \in [\Phi(b), \Phi(p)]$, or $\Phi(a)$ is close to $\Phi(p)$. Note that this includes the situation where $\Phi(x_j)$ is a degenerate vertex of T . By Fact 2.21, the situation is similar to Figure 2. Since the distance between a_{j+1} (resp. b_{j+1}) and a (resp. b) is at most 5θ , the triangular inequality and Corollary 2.2 imply

$$\begin{aligned} (b_{j+1}|x_0)_{a_{j+1}} &\leq (b|x_0)_a + (10\theta + 5\theta + 5\theta)/2 \\ &\leq (\Phi(b)|\Phi(x_0))_{\Phi(a)} + \theta + 10\theta, \end{aligned}$$

and the assumption made in this first case gives

$$(b_{j+1}|x_0)_{a_{j+1}} \leq \frac{1}{20}L - \theta. \quad (2.9)$$

We now consider another approximation tree $\Psi: X \mapsto T'$, for the list (a_{j+1}, b_{j+1}, x_0) , and choose a point $q \in [a_{j+1}, b_{j+1}]$ such that $\Psi(q)$ is the branch point of T' (see Fig. 3).

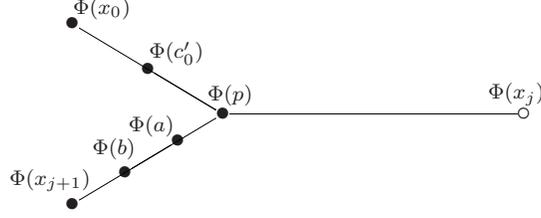


FIGURE 2. The point $\Phi(a)$ could also be on $[\Phi(p), \Phi(x_j)]$, but not far from $\Phi(p)$.

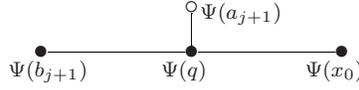


FIGURE 3.

We then define a new configuration of points $\hat{c}_{-1} = x_0$, $\hat{c}_0 = q$, $\hat{c}_1 = b_{j+1}$, $\hat{c}_2 = x_{j+1}$, and we show that this defines a configuration of order 1 for $[x_0, x_{j+1}]$. Using inequality (2.9), the fact that Ψ is an isometry along $[a_{j+1}, b_{j+1}]$, and Corollary 2.2, we obtain

$$d(q, b_{j+1}) \geq \frac{19}{20}L - (b_{j+1}|x_0)_{a_{j+1}} - \theta \geq \frac{18}{20}L.$$

Thus $[q, b_{j+1}]$ is a relator of size $18/20$ with support h_{j+1} .

By Lemma 2.3 we have $(q|x_{j+1})_{b_{j+1}} \leq 2\theta$, and by Corollary 2.2 we have $(x_0|b_{j+1})_{p_1} \leq \theta$, so we obtain property (ii); property (iii) follows from the definition of b_{j+1} (property (3) in Remark 2.11), and properties (iv) and (v) do not concern configurations of length 2. Thus, $(\hat{c}_j)_{-1 \leq j \leq 2}$ is a configuration of order 2 for the segment $[x_0, x_{j+1}]$.

Moreover, we have

$$\begin{aligned} d(x_0, q) &= (\Psi(a_{j+1})|\Psi(b_{j+1}))_{\Psi(x_0)} \\ &\geq (a_{j+1}|b_{j+1})_{x_0} - \theta \\ &\geq (a|b)_{x_0} - 11\theta \\ &\geq (\Phi(a)|\Phi(b))_{\Phi(x_0)} - 12\theta \\ &\geq d(\Phi(x_0), \Phi(p)) - 12\theta. \end{aligned}$$

Fact 2.21 then gives

$$d(x_0, q) \geq \frac{4}{20}L - 12\theta \geq \frac{3}{20}L$$

so we obtain the second assertion in Lemma 2.15.

- **Second case.** $(\Phi(b)|\Phi(x_0))_{\Phi(a)} > \frac{1}{20}L - 12\theta$.

Let i be the index such that $\Phi(p) \in [\Phi(c'_i), \Phi(c'_{i+1})]$. This index is uniquely defined if we impose $\Phi(c'_i) \neq \Phi(p)$. The assumption implies $\Phi(a) \in [\Phi(p), \Phi(x_j)]$ and $d(\Phi(p), \Phi(a)) > \frac{1}{20}L - 12\theta$. By Fact 2.21 again, the situation is similar to Figure 4. We distinguish two subcases according to the nature of $[c_i, c_{i+1}]$.

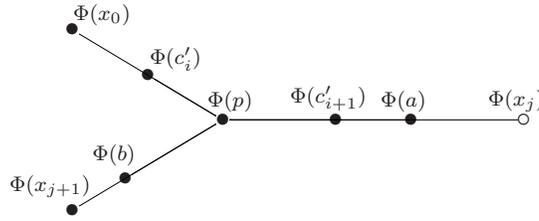


FIGURE 4. On this picture, $\Phi(c'_{i+1})$ is in between $\Phi(p)$ and $\Phi(a)$; another possible case would be $\Phi(a) \in [\Phi(p), \Phi(c'_{i+1})]$.

Second case - first subcase. Assume $[c_i, c_{i+1}]$ is neutral. Then $[c_{i+1}, c_{i+2}]$ is a relator. If $d(\Phi(p), \Phi(c'_{i+1})) \leq B$ then Fact 2.17 implies $d(\Phi(p), \Phi(a)) \leq 2B \leq \frac{1}{20}L - 12\theta$. This contradicts the assumption of the second case; as a consequence,

$$d(\Phi(p), \Phi(c'_{i+1})) > B.$$

Thus when we consider the approximation tree $\Psi: X \mapsto T'$ of $(c_{i+1}, c_i, a_{j+1}, b_{j+1})$, we obtain the pattern depicted on Figure 5, where q is a point of $[c_i, c_{i+1}]$ which is mapped to the branch point of T' . Note that $[c_i, q] \subset [c_i, c_{i+1}]$ is again a neutral segment.

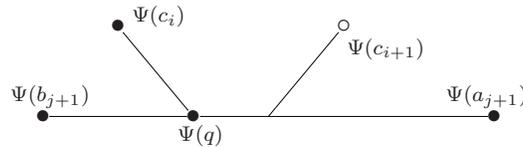


FIGURE 5.

Since the point q is 4θ close to $\text{Ax}(h_{j+1})$, the segment $[q, b_{j+1}]$ is a relator with support h_{j+1} . We have

$$\begin{aligned} d(b_{j+1}, q) &\geq d(\Psi(b_{j+1}), \Psi(q)) \\ &\geq (c_i | a_{j+1})_{b_{j+1}} - \theta \\ &\geq (c'_i | a)_b - 15\theta - \theta \\ &\geq d(\Phi(b), \Phi(p)) - 17\theta. \end{aligned}$$

Thus by Fact 2.21 we see that $[q, b_{j+1}]$ is a relator of size $5/20$, and the same conclusion holds even if one moves q along the segment $[q, b_{j+1}]$ by at most 12θ . The inequalities $(x_{j+1} | a_{j+1})_{b_{j+1}} \leq 2\theta$ and $(a_{j+1} | b_{j+1})_q \leq \theta$ imply that $(x_{j+1} | q)_{b_{j+1}} \leq 3\theta$.

In this first subcase, we define the new configuration (\hat{c}_l) by $\hat{c}_l = c_l$ for l between -1 and i , and by $\hat{c}_{i+1} = q$, $\hat{c}_{i+2} = b_{j+1}$ and $\hat{c}_{i+3} = x_{j+1}$; this defines a configuration of order $i+2$ for the segment $[x_0, x_{j+1}]$.

By construction, properties (i), (ii), and (v) are satisfied, and property (iii) is obtained after translating q along $[q, b_{j+1}]$ on a distance less than 12θ . This does not change the fact that $[q, b_{j+1}]$ is a relator of size $5/20$. Since the new configuration is obtained from the previous one by cutting it after c_i and then adding a relator $[q, b_{j+1}]$ and a neutral segment $[b_{j+1}, x_{j+1}]$, property (iv) is also satisfied.

Second case - second subcase. Assume $[c_i, c_{i+1}]$ is a relator. By Fact 2.17 the segment $[\Phi(p), \Phi(c_{i+1})]$ has length at most B , and $[c_{i+1}, c_{i+2}]$ must be neutral (otherwise again $[\Phi(p), \Phi(a)]$ would be too small). Thus the relator $[c_i, c_{i+1}]$ has size $6/20$ (or $18/20$ if $i = 0$). We consider the approximation tree $\Psi: X \mapsto T$ of $(c_{i+1}, c_i, c_{i+2}, a_{j+1}, b_{j+1})$. We obtain one of the situations depicted on Fig. 6, where now $q \in [c_{i+1}, c_i]$ or $[c_{i+1}, c_{i+2}]$.

In case (a), $[c_{i+1}, q]$ is small, and therefore neutral. We are in a case similar to the first subcase and to Figure 5. We still have

$$d(b_{j+1}, q) \geq d(\Phi(b), \Phi(p)) - 17\theta,$$

and one can define a new sequence (\hat{c}_l) by $\hat{c}_l = c_l$ for l between -1 and $i+1$, and by $\hat{c}_{i+2} = q$, $\hat{c}_{i+3} = b_{j+1}$ and $\hat{c}_{i+4} = x_{j+1}$. This new sequence cuts (c_l) after $l = i+1$, add a neutral segment after the last relator $[c_i, c_{i+1}]$, and then a relator $[q, b_{j+1}]$ and a final neutral segment $[b_{j+1}, x_{j+1}]$. Hence, (\hat{c}_l) is a configuration of order $i+3$ for the segment $[x_0, x_{j+1}]$.

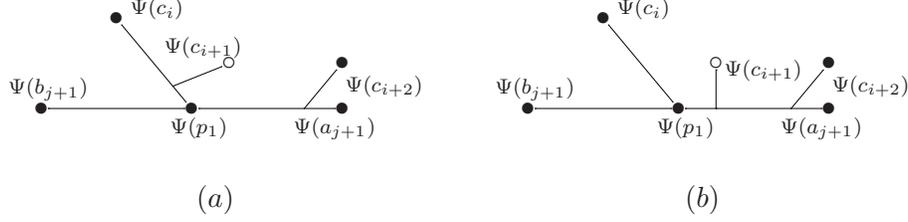


FIGURE 6.

In case (b), let us check that $d(c_{i+1}, q) \leq B + 15\theta$, so that $[c_i, q]$ is a relator of size $4/20$ (or $17/20$ if $i = 0$). We have

$$\begin{aligned} d(c_{i+1}, q) &\leq (c_i | b_{j+1})_{c_{i+1}} + 2\theta \\ &\leq (\Phi(c'_i) | \Phi(b))_{\Phi(c'_{i+1})} + 15\theta \\ &\leq B + 15\theta. \end{aligned}$$

As in the first subcase, the estimate for $d(b_{j+1}, q)$ gives

$$d(b_{j+1}, q) \geq d(\Phi(b), \Phi(p)) - 17\theta,$$

and Fact 2.21 implies that $[q, b_{j+1}]$ is a relator of size $5/20$; moreover $(x_{j+1} | p_1)_{b_{j+1}} \leq 3\theta$.

We define the new configuration (\hat{c}_l) by $\hat{c}_l = c_l$ for l between -1 and i , and by $\hat{c}_{i+1} = q$, $\hat{c}_{i+2} = b_{j+1}$ and $\hat{c}_{i+3} = x_{j+1}$; this defines a configuration of order $i + 2$ for the segment $[x_0, x_{j+1}]$.

By construction, properties (i), (ii), (iii) and (v) are satisfied. Property (iv) is also satisfied because $[\hat{c}_i, q]$ and $[q, b_{j+1}]$ are both relators of size $5/20$ (or $17/20$ if $i = 0$) and the last segment $[b_{j+1}, x_{j+1}]$ is neutral. \square

Proof of Theorem 2.9. By Lemma 2.15 there exists (c_i) a configuration of order k for $[x_0, x_m]$, where $x_m = h(x_0)$. Recall that by our choice of x_0 we have $L(h) \geq d(x_0, x_m) - \theta$.

If $k \geq 2$, then we have at least two distinct relators: the first one, $[c_0, c_1]$, has size at least $17/20$, and the last one, $[c_{k-1}, c_k]$, has size $5/20$. By Lemma 2.6 we obtain $L(h) \geq \frac{22}{20}L - 11\theta > L$.

If $k = 1$, either $m = 1$ and h is conjugate to g^n , or we have $d(x_0, c_0) \geq \frac{3}{20}L$. On the other hand $d(c_0, c_1) \geq \frac{18}{20}L$, so we obtain $L(h) \geq \frac{21}{20}L - 11\theta > L$. \square

3. HYPERBOLIC SPACES WITH CONSTANT NEGATIVE CURVATURE.

This section is devoted to the classical hyperbolic space \mathbb{H}^n , where the dimension n is allowed to be infinite. As we shall see, constant negative curvature,

which is stronger than δ -hyperbolicity, is a useful property to decide whether the axis $\text{Ax}(g)$ of an isometry is rigid.

3.1. Hyperbolic spaces.

3.1.1. *Definition.* Let H be a Hilbert space, with scalar product $(x \cdot y)_H$, and let $\|x\|_H$ denotes the norm of any element $x \in H$. Let u be a unit vector in H , and u^\perp be its orthogonal complement; each element $x \in H$ decomposes uniquely into $x = \alpha(x)u + v_x$ with $\alpha(x)$ in \mathbf{R} and v_x in u^\perp . Let $\langle \cdot | \cdot \rangle: H \times H \mapsto \mathbf{R}$ be the symmetric bilinear mapping defined by

$$\langle x|y \rangle = \alpha(x)\alpha(y) - (v_x \cdot v_y)_H.$$

This bilinear mapping is continuous and has signature equal to $(1, \dim(H) - 1)$. The set of points x with $\langle x|x \rangle = 0$ is the light cone of $\langle \cdot | \cdot \rangle$. Let \mathbb{H} be the subset of H defined by

$$\mathbb{H} = \{x \in H; \langle x|x \rangle = 1 \text{ and } \langle u|x \rangle > 0\}.$$

The space \mathbb{H} is the sheet of the hyperboloid $\langle x|x \rangle = 1$ which contains u .

The function $\text{dist}: \mathbb{H} \times \mathbb{H} \mapsto \mathbf{R}_+$ defined by

$$\cosh(\text{dist}(x, y)) = \langle x|y \rangle$$

gives a distance on \mathbb{H} , and $(\mathbb{H}, \text{dist})$ is a complete and simply connected riemannian manifold of dimension $\dim(H) - 1$ with constant scalar curvature -1 (this characterizes \mathbb{H} if the dimension is finite). As such, \mathbb{H} is a CAT(-1) space, and therefore a δ -hyperbolic space. More precisely, \mathbb{H} is δ -hyperbolic (in any dimension, even infinite) with the same constant $\delta = \log(3)$ (see [CDP90], § I.4 page 11). In particular, all properties listed in §2.1 (or even stronger ones) are satisfied.

3.1.2. *Geodesics and boundary.* The hyperbolic space \mathbb{H} is in one to one correspondence with its projection into the projective space $\mathbb{P}(H)$. The boundary of this subset of $\mathbb{P}(H)$ is the projection of the light cone of $\langle \cdot | \cdot \rangle$; we shall denote it by $\partial\mathbb{H}$ and call it the **boundary** of \mathbb{H} (or boundary at infinity).

Let Γ be a geodesic line in \mathbb{H} . Then there is a unique plane $V_\Gamma \subset H$ such that $\Gamma = V_\Gamma \cap \mathbb{H}$. The plane V_Γ intersects the light cone on two lines, and these lines determine two points of $\partial\mathbb{H}$, called the endpoints of Γ . If x and y are two points of \mathbb{H} , there is a unique geodesic segment $[x, y]$ from x to y ; this segment is contained in a unique geodesic line, namely $\text{Vect}(x, y) \cap \mathbb{H}$.

Let x be a point of \mathbb{H} and Γ be a geodesic. The projection $\pi_\Gamma(x) \in \Gamma$, which minimizes $\text{dist}(x, y)$ for y describing Γ , is unique.

3.1.3. *Distances between geodesics.* The following lemma shows that close geodesic segments are indeed exponentially close in a neighborhood of their middle points; this property is not satisfied in general δ -hyperbolic spaces because δ -hyperbolicity does not tell anything at scale $\leq \delta$.

Lemma 3.1. *There is a positive constant $c_{\mathbb{H}}$ with the following properties. Let x, x', y , and y' be four points of the hyperbolic space \mathbb{H} . Assume that*

- (i) $\text{dist}(x, x') \leq \varepsilon$, and $\text{dist}(y, y') \leq \varepsilon$,
- (ii) $\text{dist}(x, y) \geq B$,

where ε and B are positive real numbers. Let $[u, v] \subset [x, y]$ be the geodesic segment of length $\frac{3}{4}\text{dist}(x, y)$ centered in the middle of $[x, y]$. Then $[u, v]$ is contained in the tubular neighborhood $\text{Tub}_{\nu}([x', y'])$ with

$$\cosh(\nu) - 1 \leq c_{\mathbb{H}} \frac{\cosh(\varepsilon) - 1}{\cosh(B)}.$$

The constant $c_{\mathbb{H}}$ does not depend on the dimension of \mathbb{H} (the computation takes place into $\text{Vect}(x, x', y, y') \cap \mathbb{H}$), nor on B and ε ; it depends, of course, on the choice of the ratio $\text{dist}(u, v)/\text{dist}(x, y) = 3/4$.

3.1.4. *Isometries.* Let f be an isometry of \mathbb{H} ; then f is the restriction of a unique continuous linear transformation of the Hilbert space H . In particular, f extends to an homeomorphism of the boundary $\partial\mathbb{H}$. The three types of isometries (see § 2.3.1) have the following properties.

- (1) If f is elliptic there is a point x in \mathbb{H} with $f(x) = x$, and f acts as a rotation centered at x . Fixed points of f are eigenvectors of the linear extension $f: H \rightarrow H$ corresponding to the eigenvalue 1.
- (2) If f is parabolic there is a unique fixed point of f on the boundary $\partial\mathbb{H}$; this point corresponds to a line of eigenvectors with eigenvalue 1 for the linear extension $f: H \rightarrow H$. The orbit $(f^k(x))$ of all points x in \mathbb{H} converges toward this fixed point in $\mathbb{P}(H)$ when k goes to $+\infty$ and to $-\infty$.
- (3) If f is hyperbolic, then f has exactly two fixed points $\alpha(f)$ and $\omega(f)$ on the boundary $\partial\mathbb{H}$ and the orbit $f^k(x)$ of every point $x \in \mathbb{H}$ goes to $\alpha(f)$ as k goes to $-\infty$ and $\omega(f)$ as k goes to $+\infty$. The set $\text{Min}(f)$ coincides with the geodesic line from $\alpha(f)$ to $\omega(f)$. In particular, $\text{Min}(f)$ coincides with the unique f -invariant axis $\text{Ax}(f) = \text{Vect}(\alpha(f), \omega(f)) \cap \mathbb{H}$. The points $\alpha(f)$ and $\omega(f)$ correspond to eigenlines of the linear extension $f: H \rightarrow H$ with eigenvalues λ^{-1} and λ , where $\lambda > 1$; the translation length $L(f)$ is equal to the logarithm of λ (see Remark 4.5 below).

3.2. Rigidity of axis for hyperbolic spaces.

Let G be a group of isometries of the hyperbolic space \mathbb{H} , and g be a hyperbolic element of G . The following strong version of Lemma 2.8 is a direct consequence of Lemma 3.1.

Lemma 3.2. *Let ε and ε' be positive real numbers. If $A \subset \mathbb{H}$ is ε' -rigid, then A is also ε -rigid.*

Hence, in the context of the hyperbolic spaces \mathbb{H} , the ε can be dropped in the notation of ε -rigidity; we simply say that $A \subset \mathbb{H}$ is rigid.

Proposition 3.3. *Let G be a group of isometries of \mathbb{H} , and $g \in G$ be a hyperbolic isometry. Let ε and B be positive real numbers. Let n be a positive integer with $nL(g) \leq B/6$. Let p be a point of $Ax(g)$. If $Ax(g)$ is not (B, ε) -rigid, there exists an element f of G such that $f(Ax(g)) \neq Ax(g)$ and $\text{dist}(f(x), x) < \varepsilon$ for all $x \in [g^{-n}(p), g^n(p)]$.*

Proof. Assume that $B \geq 6nL(g) \geq 6L(g)$. Since $Ax(g)$ is not (B, ε) -rigid, there exists an isometry $h \in G$ such that $h(Ax(g))$ is different from $Ax(g)$, but $Ax(g)$ contains a segment J of length B which is mapped by h into the tubular neighborhood $\text{Tub}_\varepsilon(Ax(g))$. Changing h into $g^j \circ h$, we assume that the point p is near the middle of the segment $h(J)$.

Changing h into $h \circ g^k$ moves J to $g^{-k}(J)$. We can thus change h into $h_1 = h \circ g^k$ for some $k \in \mathbf{Z}$, B into $B_1 \geq B - 4L(g)$, and find two points x and y on $Ax(g)$ that satisfy

- (1) $\text{dist}(x, y) \geq B_1$, $p \in [x, y]$, and $\text{dist}(p, x) \geq B/3$, $\text{dist}(p, y) \geq B/3$;
- (2) either $h_1(x) \in \text{Tub}_\varepsilon([g^{-1}(x), x])$ and $h_1(y) \in \text{Tub}_\varepsilon([g^{-1}(y), y])$, or $h_1(y) \in \text{Tub}_\varepsilon([g^{-1}(x), x])$ and $h_1(x) \in \text{Tub}_\varepsilon([g^{-1}(y), y])$;
- (3) $g^i(x)$ and $g^i(y)$, with $-2 \leq i \leq 2$, are at distance at most ε from $h(Ax(g))$.

This does not change the axis $h(Ax(g))$ and the value of ε .

We now change h_1 into the commutator $h_2 = h_1^{-1}g^{-1}h_1g$. We still have $h_2(Ax(g)) \neq Ax(g)$, because the uniqueness of the axis of a hyperbolic isometry of \mathbb{H} implies $h_2(Ax(g)) = h_1^{-1}(Ax(g))$. Moreover, there is an $\varepsilon_2 \leq 2\varepsilon$ such that property (2) above is replaced by $\text{dist}(x, h_2(x)) \leq \varepsilon_2$ and $\text{dist}(y, h_2(y)) \leq \varepsilon_2$, or by $\text{dist}(g^2(x), h_2(x)) \leq \varepsilon_2$ and $\text{dist}(g^2(y), h_2(y)) \leq \varepsilon_2$; similar properties are then satisfied by the points $g^i(x)$ and $g^i(y)$, $-2 \leq i \leq 2$, in place of x and y .

Changing, once again, h_2 into $h_3 = h_2 \circ g^{-2}$ if necessary, we can assume that $\text{dist}(x, h_2(x)) \leq \varepsilon_2$ and $\text{dist}(y, h_2(y)) \leq \varepsilon_2$.

Consider the arc length parametrization

$$m: t \in [-\infty, +\infty] \mapsto Ax(g)$$

such that $m(0) = p$ and $g(m(t)) = m(t + L(g))$. Since p is in the interval $[x + B/3, y - B/3]$, we obtain $\text{dist}(m(t), h_3(m(t))) \leq \varepsilon_2$ for all t in $[-B/3, B/3]$. Defining $f = h_3$ we get

$$\text{dist}(z, f(z)) \leq 2\varepsilon, \quad \forall z \in [g^{-n}(p), g^n(p)]$$

as soon as $nL(g) \leq B/6$. □

Corollary 3.4. *Let G be a group of isometries of \mathbb{H} , g be a hyperbolic element of G , and p a point of $Ax(g)$. Let ε and B be positive numbers and n a positive integer, with $nL(g) \leq B/8$. If $Ax(g)$ is not rigid (B, ε) -rigid, then there exists $f \in G$ and $\nu > 0$ such that*

- (1) $\cosh(\nu) - 1 \leq c_{\mathbb{H}}(\cosh(\varepsilon) - 1)/(\cosh(B))$;
- (2) $f(Ax(g)) \neq Ax(g)$ and $\text{dist}(f(z), z) < \nu$ for all $z \in [g^{-n}(p), g^n(p)]$.

Proof. Assume that $Ax(g)$ is not (B, ε) -rigid, then it is not $(3B/4, \nu)$ -rigid, with ν as in Lemma 3.1. Proposition 3.3 shows the existence of f in G with $f(Ax(g)) \neq Ax(g)$ and $\text{dist}(f(z), z) \leq 2\nu$ for all z in $[g^{-n}(p), g^n(p)]$ as soon as $nL(g) \leq B/8$. □

Part B. Algebraic Geometry and the Cremona Group

4. THE PICARD-MANIN SPACE

4.1. Néron-Severi groups and rational morphisms.

The Néron-Severi group of a smooth projective rational surface is the group

$$\mathbf{N}^1(X) = H^2(X, \mathbf{Z}) \cap H^{1,1}(X, \mathbf{R}),$$

of Chern classes of holomorphic line bundles of X ; it coincides with the space of Cartier divisors modulo numerical equivalence (see [Laz04]). The rank $\rho(X)$ of this abelian group is called the Picard number of X . If D is a divisor, we denote by $[D]$ its class in $\mathbf{N}^1(X)$. The intersection form defines an integral quadratic form

$$([D_1], [D_2]) \mapsto [D_1] \cdot [D_2]$$

on $\mathbf{N}^1(X)$, the signature of which is equal to $(1, \rho(X) - 1)$ by Hodge index theorem. We also note $\mathbf{N}^1(X)_{\mathbf{R}} = \mathbf{N}^1(X) \otimes \mathbf{R}$.

Example 4.1. The Néron-Severi group of the plane is isomorphic to $\mathbf{Z}[H]$ where $[H]$ is the class of a line H and $[H] \cdot [H] = 1$.

If $\pi: X \rightarrow Y$ is a birational morphism, then the pullback morphism

$$\pi^*: \mathbf{N}^1(Y) \rightarrow \mathbf{N}^1(X)$$

is injective and preserves the intersection form. For example, if π is just the blow-up of a point with exceptional divisor $E \subset X$, then $\mathbf{N}^1(X)$ is isomorphic to $\pi^*(\mathbf{N}^1(Y)) \oplus \mathbf{Z}[E]$, this sum is orthogonal with respect to the intersection form, and $[E] \cdot [E] = -1$.

4.2. Indeterminacies and dynamical degrees of birational transformations.

Let $f: X \dashrightarrow Y$ be a rational map between smooth projective surfaces. The indeterminacy set $\text{Ind}(f)$ is finite, and the curves which are blown down by f form a codimension 1 analytic subset of X , called the exceptional set $\text{Exc}(f)$.

Let H be an ample line bundle on Y . Consider the pullback by f of the linear system of divisors which are linearly equivalent to H ; when $X = Y = \mathbb{P}_{\mathbb{C}}^2$ and H is a line, this linear system is called the homaloidal net of f . The base locus of this linear system is supported on $\text{Ind}(f)$, but can of course include infinitely near points. We shall call it the **base locus** of f . To resolve the indeterminacies of f , one blows up the base locus (see [Laz04] for base loci, base ideal, and their blow ups); in other words, one blows up $\text{Ind}(f)$, obtaining $\pi: X' \rightarrow X$, then one blows up $\text{Ind}(f \circ \pi)$, and so on; the process stops in a finite number of steps.

Remark 4.2. If f is a birational transformation of a projective surface with Picard number one, then $\text{Ind}(f)$ is contained in $\text{Exc}(f)$.⁴

The rational map f determines a linear map $f^*: H^2(Y, \mathbf{Z}) \rightarrow H^2(X, \mathbf{Z})$ which preserves the Hodge decomposition and maps the Néron-Severi group $\mathbf{N}^1(Y)$ into the Néron-Severi group of X (see [DF01] for example).

Assume now that f is a birational selfmap of X . The **dynamical degree** $\lambda(f)$ of f is the spectral radius of the sequence of linear maps $((f^n)^*)_{n \geq 0}$:

$$\lambda(f) = \lim_{n \rightarrow +\infty} \left(\|(f^n)^*\|^{1/n} \right)$$

where $\|\cdot\|$ is an operator norm on $\text{End}(H^*(X, \mathbf{R}))$; the limit exists because the sequence $\|(f^n)^*\|$ is submultiplicative (see [DF01]). The number $\lambda(f)$ is invariant under conjugacy: $\lambda(f) = \lambda(gfg^{-1})$ if $g: X \dashrightarrow Y$ is a birational map.

Example 4.3. Let $[x : y : z]$ be homogeneous coordinates for the projective plane $\mathbb{P}_{\mathbb{C}}^2$. Let f be an element of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$; there are three homogeneous polynomials $P, Q,$ and $R \in \mathbf{C}[x, y, z]$ of the same degree d , and without common

⁴ One can prove this by considering the factorization of f as a sequence of blow-ups followed by a sequence of blow-downs. Any curve in $\text{Exc}(f)$ corresponds to a -1 curve at some point in the sequence of blow-down, but is also the strict transform of a curve of positive self-intersection from the source (this is where we use $\rho(X) = 1$).

factor of degree ≥ 1 , such that

$$f[x : y : z] = [P : Q : R].$$

The degree d is called the degree of f , and is denoted by $\deg(f)$. The action f^* of f on $\mathbf{N}^1(\mathbb{P}^2(\mathbf{C}))$ is multiplication by $\deg(f)$. The dynamical degree of f is thus equal to the limit of $\deg(f^n)^{1/n}$.

4.3. Picard-Manin classes.

We follow the presentation given in [BFJ08, Can07] which, in turn, is inspired by the fifth chapter of Manin's book [Man86]. See also the survey paper [Fav08].

4.3.1. *Models and Picard-Manin space.* A model of $\mathbb{P}_{\mathbf{C}}^2$ is a smooth projective surface X with a birational morphism $X \rightarrow \mathbb{P}_{\mathbf{C}}^2$. Let \mathcal{B} be the set of all models of $\mathbb{P}_{\mathbf{C}}^2$. If $\pi: X \rightarrow \mathbb{P}_{\mathbf{C}}^2$, and $\pi': X' \rightarrow \mathbb{P}_{\mathbf{C}}^2$ are two models, we say that X' dominates X if the induced birational map $\pi^{-1} \circ \pi': X' \dashrightarrow X$ is a morphism. In this case, $\pi^{-1} \circ \pi'$ contracts a finite number of exceptional divisors and induces an injective map $\pi^*: \mathbf{N}^1(X)_{\mathbf{R}} \hookrightarrow \mathbf{N}^1(X')_{\mathbf{R}}$.

Let $\mathcal{B}_X \subset \mathcal{B}$ be the set of models that dominate X . The space $\mathcal{Z}(X)$ of **(finite) Picard-Manin classes**⁵ is the injective limit

$$\mathcal{Z}(X) = \lim_{\rightarrow \mathcal{B}_X} \mathbf{N}^1(X')_{\mathbf{R}}.$$

The Néron-Severi group $\mathbf{N}^1(X')_{\mathbf{R}}$ of any model $X' \rightarrow X$ embeds in $\mathcal{Z}(X)$ and can be identified to its image into $\mathcal{Z}(X)$; the same property holds for the lattice $\mathbf{N}^1(X')$. Thus, a Picard-Manin class is just a (real) Néron-Severi class of some model dominating X . The Picard-Manin class of a divisor D is still denoted by $[D]$, as for Néron-Severi classes.

4.3.2. *Basis of $\mathcal{Z}(X)$.* A basis of the real vector space $\mathcal{Z}(X)$ is constructed as follows. On the set of models $\pi: Y \rightarrow X$ together with marked points $p \in Y$, we introduce the equivalence relation: $(p, Y) \sim (p', Y')$ if the induced birational map $\pi'^{-1} \circ \pi: Y \dashrightarrow Y'$ is an isomorphism in a neighborhood of p that maps p onto p' . Let \mathcal{V}_X be the quotient space; to denote points of \mathcal{V}_X we just write $p \in \mathcal{V}_X$, without any further reference to a model $\pi: Y \rightarrow X$ with $p \in Y$. Then to each $p \in \mathcal{V}_X$ we can associate a Picard-Manin class in $\mathcal{Z}(X)$. For this, consider $\bar{Y} \rightarrow Y$ the blow-up of p and $E_p \subset \bar{Y}$ the exceptional divisor; the Néron-Severi class $[E_p]$ determines a Picard-Manin class and one easily verifies that this class depends only on the class $p \in \mathcal{V}_X$ (not on the model (Y, π) , see

⁵ Picard-Manin classes can be thought of as classes of divisors on the so-called Riemann-Zariski space; one can also define Weil classes by mean of a projective limit. These spaces are not useful for our purposes.

[Man86]). The classes $[E_p]$, $p \in \mathcal{V}_X$, have self-intersection -1 , are mutually orthogonal, and are orthogonal to $\mathbf{N}^1(X)_{\mathbf{R}}$. Moreover,

$$\mathcal{Z}(X) = \mathbf{N}^1(X)_{\mathbf{R}} \oplus \text{Vect}([E_p]; p \in \mathcal{V}_X)$$

and this sum is orthogonal with respect to the intersection form. To get a basis of $\mathcal{Z}(X)$, we then fix a basis $([H_i])_{1 \leq i \leq \rho(X)}$ of $\mathbf{N}^1(X)_{\mathbf{R}}$ (where the H_i are Cartier divisors) and complete it with the family $([E_p])_{p \in \mathcal{V}_X}$.

4.3.3. *Completion.* The **(completed) Picard-Manin space** $\bar{\mathcal{Z}}(X)$ of X is the L^2 -completion of $\mathcal{Z}(X)$ (see [BFJ08, Can07] for details); hence

$$\bar{\mathcal{Z}}(X) = \left\{ [D] + \sum_{p \in \mathcal{V}_X} a_p [E_p]; [D] \in \mathbf{N}^1(X)_{\mathbf{R}}, a_p \in \mathbf{R} \text{ and } \sum a_p^2 < +\infty \right\}$$

whereas $\mathcal{Z}(X)$ corresponds to the case where the a_p vanish for all but a finite number of $p \in \mathcal{V}_X$. For the projective plane $\mathbb{P}_{\mathbf{C}}^2$, the Néron-Severi group $\mathbf{N}^1(\mathbb{P}_{\mathbf{C}}^2)$ is isomorphic to $\mathbf{Z}[H]$, where H is a line; elements of $\bar{\mathcal{Z}}(X)$ are then given by sums

$$a_0[H] + \sum_{p \in \mathcal{V}_{\mathbb{P}_{\mathbf{C}}^2}} a_p [E_p]$$

with $\sum a_p^2 < +\infty$. We shall call this space the Picard-Manin space (without further reference to $\mathbb{P}_{\mathbf{C}}^2$ or to the completion).

4.3.4. *Integral points.* Note that $\mathcal{Z}(X)$ contains the injective limit of the lattices $\mathbf{N}^1(X')$ (with integer coefficients). This provides an integral structure for $\mathcal{Z}(X)$: The basis of $\mathcal{Z}(X)$ constructed in §4.3.2 is made of integral points.

4.4. Action of $\text{Bir}(X)$ on $\bar{\mathcal{Z}}(X)$.

If $\pi: X' \rightarrow X$ is a morphism, then π induces an isomorphism $\pi^*: \mathcal{Z}(X) \rightarrow \mathcal{Z}(X')$. Let us describe this fact when π is the (inverse of the) blow-up of a point $q \in X$. In this case we have

$$\mathbf{N}^1(X') = \pi^* \mathbf{N}^1(X) \oplus \mathbf{Z}[E_p] \quad \text{and} \quad \mathcal{V}_{X'} = \mathcal{V}_X \cup \{(q, X)\}$$

where $(q, X) \in \mathcal{V}_X$ denotes the point of \mathcal{V}_X given by $q \in X$. Thus the bases of $\mathcal{Z}(X)$ and $\mathcal{Z}(X')$ are in bijection, the only change being that $[E_p]$ is viewed as the class of an exceptional divisor in $\mathcal{Z}(X)$, and as an element of $\mathbf{N}^1(X')$ in $\mathcal{Z}(X')$. Note that π^* extends uniquely as a continuous isomorphism $\pi^*: \bar{\mathcal{Z}}(X') \rightarrow \bar{\mathcal{Z}}(X)$ that preserves the intersection form. Since any birational morphism π is a composition of blow-downs of exceptional curves of the first kind, this proves that π acts by isometries on both $\mathcal{Z}(X)$ and $\bar{\mathcal{Z}}(X)$.

Now consider $f: \mathbb{P}_{\mathbb{C}}^2 \dashrightarrow \mathbb{P}_{\mathbb{C}}^2$ a birational map. There exist a surface X and two morphisms $\pi: X \rightarrow \mathbb{P}_{\mathbb{C}}^2$, $\sigma: X \rightarrow \mathbb{P}_{\mathbb{C}}^2$, such that $f = \sigma \circ \pi^{-1}$. Defining f^* by $f^* = (\pi^*)^{-1} \circ \sigma^*$, and f_* by $f_* = (f^{-1})^*$, we get a representation

$$f \mapsto f_*$$

of the Cremona group in the orthogonal group of $\mathcal{Z}(\mathbb{P}_{\mathbb{C}}^2)$ (resp. $\bar{\mathcal{Z}}(\mathbb{P}_{\mathbb{C}}^2)$) with respect to the intersection form. This representation is faithful, because $f_*[E_p] = [E_{f(p)}]$ for all points p in the domain of definition of f , and preserves the integral structure.

4.5. Action on an infinite dimensional hyperbolic space.

4.5.1. *The hyperbolic space $\mathbb{H}_{\bar{\mathcal{Z}}}$.* We define

$$\mathbb{H}_{\bar{\mathcal{Z}}} = \{[D] \in \bar{\mathcal{Z}}; [D]^2 = 1, [H] \cdot [D] > 0\}$$

and a distance dist on $\mathbb{H}_{\bar{\mathcal{Z}}}$ by the following formula

$$\cosh(\text{dist}([D_1], [D_2])) = [D_1] \cdot [D_2].$$

Since the intersection form is of Minkowski type, this gives to $\mathbb{H}_{\bar{\mathcal{Z}}}$ the structure of an infinite dimensional hyperbolic space, as in §3.1.

4.5.2. *Cremona isometries.* By the previous paragraph the action of the Cremona group preserves the two-sheeted hyperboloid $\{[D] \in \bar{\mathcal{Z}}(\mathbb{P}_{\mathbb{C}}^2); [D]^2 = 1\}$ and since the action also preserves the nef cone, we obtain a faithful representation of the Cremona group into the group of isometries of $\mathbb{H}_{\bar{\mathcal{Z}}}$:

$$\text{Bir}(\mathbb{P}_{\mathbb{C}}^2) \hookrightarrow \text{Isom}(\mathbb{H}_{\bar{\mathcal{Z}}}).$$

In the context of the Cremona group, the classification of isometries into three types (see § 3.1.4) has an algebraic-geometric meaning.

Theorem 4.4 ([Can07], [Giz80], [DF01]). *Let f be an element of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$. The isometry f_* of $\mathbb{H}_{\bar{\mathcal{Z}}}$ is*

- *elliptic if, and only if, there exists a positive iterate f^k of f and a birational map $\varepsilon: \mathbb{P}_{\mathbb{C}}^2 \rightarrow X$ such that $\varepsilon \circ f \circ \varepsilon^{-1}$ is an element of $\text{Aut}(X)^0$ (the connected component of the identity in $\text{Aut}(X)$);*
- *parabolic if, and only if, one of the two following cases occur: f preserves a pencil of elliptic curves and $\deg(f^n)$ grows quadratically with n , or f preserves a pencil of rational curves and $\deg(f^n)$ grows linearly with n ;*
- *hyperbolic if and only if the dynamical degree $\lambda(f)$ is > 1 .*

As a consequence, when $\lambda(f) > 1$ then f_* preserves a unique geodesic line $\text{Ax}(f) \subset \mathbb{H}_{\bar{\mathcal{Z}}}$; this line is the intersection of a plane $V_f \subset \bar{\mathcal{Z}}$ which intersects the isotropic cone of $\bar{\mathcal{Z}}$ on two lines R_f^+ and R_f^- such that

$$f_*(a) = \lambda(f)^{\pm 1} a$$

for all $a \in R_f^\pm$ (the lines R_f^+ and R_f^- correspond to $\omega(f)$ and $\alpha(f)$ in the notation of § 3.1.4).

Remark 4.5. The translation length $L(f_*)$ is therefore equal to $\log(\lambda(f))$. Indeed, take $[\alpha] \in R_f^-$, $[\omega] \in R_f^+$ normalized such that $[\alpha] \cdot [\omega] = 1$. The point $[P] = \frac{1}{\sqrt{2}}([\alpha] + [\omega])$ is on the axis $\text{Ax}(f_*)$. Since $f_*[P] = \frac{1}{\sqrt{2}}(\lambda(f)^{-1}[\alpha] + \lambda(f)[\omega])$, we get

$$e^{L(f_*)} + \frac{1}{e^{L(f_*)}} = 2 \cosh(\text{dist}([P], f_*[P])) = 2([P] \cdot f_*[P]) = \lambda(f) + \frac{1}{\lambda(f)}.$$

4.5.3. *Automorphisms.* Assume that $f \in \text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ can be lifted, via a birational transformation π , as an automorphism g on a smooth rational surface X :

$$\begin{array}{ccc} X & \xrightarrow{g} & X \\ \downarrow \pi & & \downarrow \pi \\ \mathbb{P}_{\mathbb{C}}^2 & \xrightarrow{f} & \mathbb{P}_{\mathbb{C}}^2 \end{array}$$

Then we have an isomorphism $\pi_*: \bar{\mathcal{Z}}(X) \rightarrow \bar{\mathcal{Z}}(\mathbb{P}_{\mathbb{C}}^2)$ and an orthogonal decomposition

$$\bar{\mathcal{Z}}(X) = \mathbf{N}^1(X)_{\mathbf{R}} \oplus \mathbf{N}^1(X)_{\mathbf{R}}^\perp$$

where $\mathbf{N}^1(X)_{\mathbf{R}}^\perp$ is spanned by the classes $[E_p]$, $p \in \mathcal{V}_X$. This orthogonal decomposition is g_* -invariant. In particular, f_* preserves the finite dimensional subspace $\pi_* \mathbf{N}^1(X)_{\mathbf{R}} \subset \bar{\mathcal{Z}}(\mathbb{P}_{\mathbb{C}}^2)$.

By Hodge index theorem, the intersection form has signature $(1, \rho(X) - 1)$ on $\mathbf{N}^1(X)$, so that $\pi_* \mathbf{N}^1(X)_{\mathbf{R}}$ intersects $\mathbb{H}_{\bar{\mathcal{Z}}}$ on a f_* -invariant hyperbolic subspace of dimension $\rho(X) - 1$. This proves the following lemma.

Lemma 4.6. *If f is conjugate to an automorphism $g \in \text{Aut}(X)$ by a birational map $\pi: X \rightarrow \mathbb{P}_{\mathbb{C}}^2$, then:*

- (1) *The isometry $f_*: \mathbb{H}_{\bar{\mathcal{Z}}} \rightarrow \mathbb{H}_{\bar{\mathcal{Z}}}$ is hyperbolic (resp. parabolic, resp. elliptic) if and only if the isometry $g_*: \mathbf{N}^1(X)_{\mathbf{R}} \rightarrow \mathbf{N}^1(X)_{\mathbf{R}}$ is hyperbolic (resp. parabolic, resp. elliptic) for the intersection form on $\mathbf{N}^1(X)_{\mathbf{R}}$;*
- (2) *The translation length of f_* is equal to the translation length of g_* ;*
- (3) *If f_* is hyperbolic then, modulo π_* -conjugacy, the plane V_f corresponds to V_g , which is contained in $NS(X)_{\mathbf{R}}$, and $\text{Ax}(f_*)$ corresponds to $\text{Ax}(g_*)$.*

4.5.4. *Quadratic mappings.* Let $f: \mathbb{P}_{\mathbb{C}}^2 \dashrightarrow \mathbb{P}_{\mathbb{C}}^2$ be a quadratic birational map. The base locus of f (resp. f^{-1}) is made of three points p_1, p_2, p_3 (resp. q_1, q_2, q_3), where infinitely near points are allowed.

Example 4.7. The set of birational quadratic maps is an irreducible algebraic variety of dimension 14. A Zariski open subset is made of birational transformations $f = h_1 \circ \sigma \circ h_2$, with $h_i \in \text{Aut}(\mathbb{P}_{\mathbb{C}}^2)$, $i = 1, 2$, and σ the standard quadratic involution $\sigma[x : y : z] = [yz : zx : xy]$. The set $\{p_1, p_2, p_3\}$ is the image of $\text{Ind}(\sigma) = \{[1 : 0 : 0]; [0 : 1 : 0]; [0 : 0 : 1]\}$ by h_2^{-1} , and $\{q_1, q_2, q_3\} = h_1(\text{Ind}(\sigma))$.

We have

$$f_*([H]) = 2[H] - [E_{q_1}] - [E_{q_2}] - [E_{q_3}]$$

with $[H]$ the class of a line in $\mathbb{P}_{\mathbb{C}}^2$. If f is an isomorphism on a neighborhood of p , and $f(p) = q$, then

$$f_*([E_p]) = [E_q].$$

The exceptional set $\text{Exc}(f^{-1})$ is made of three lines through the points q_i (see [DC09]). Assume, for example, that the line through q_1 and q_2 is contracted on p_1 by f^{-1} , then we have

$$f_*([E_{p_1}]) = [H] - [E_{q_2}] - [E_{q_3}].$$

Note that these formulas are correct even when there are some infinitely near base points. For instance if f is the Hénon map

$$[x : y : z] \dashrightarrow [yz : y^2 - xz : z^2]$$

then q_2 is infinitely near q_1 , and q_3 is infinitely near to q_2 , but nevertheless the formula for the image of H is still the same. In such a situation it is crucial to work with total transforms (as opposed to strict transforms).

If one takes a generic quadratic map, then f induces a hyperbolic isometry on $\mathbb{H}_{\bar{\mathbb{Z}}}$, and it is possible to compute explicitly the points on the axis of f (see §5.1.4 for precise statements). Note that $[H]$ is *not* on the axis; in fact in this situation the axis of f does not contain any finite class.

5. TIGHT BIRATIONAL MAPS

5.1. Generic Cremona transformations.

In this section we prove Theorem A concerning normal subgroups generated by generic Cremona transformations.

5.1.1. *De Jonquières transformations.* Let d be a positive integer. As mentioned in the introduction, the set $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ of plane birational transformations of degree d is quasi-projective: it is a Zariski open subset in a subvariety of the projective space made of triples of homogeneous polynomials modulo scalar multiplication.

Recall that J_d denotes the set of de Jonquières transformations of degree d , defined as the birational transformations of degree d of $\mathbb{P}_{\mathbf{C}}^2$ that preserve the pencil of lines through $q_0 = [1 : 0 : 0]$. Then we define V_d as the image of the composition map

$$(h_1, f, h_2) \mapsto h_1 \circ f \circ h_2$$

where (h_1, f, h_2) describes $\text{PGL}_3(\mathbf{C}) \times J_d \times \text{PGL}_3(\mathbf{C})$. As the image of an irreducible algebraic set by a regular map, V_d is an irreducible subvariety of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$. The dimension of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ is equal to $4d + 6$ and V_d is the unique irreducible component of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ of maximal dimension (in that sense, generic elements of $\text{Bir}_d(\mathbb{P}_{\mathbf{C}}^2)$ are contained in V_d). In degree 2, i.e. for quadratic Cremona transformations, V_2 coincides with $\text{Bir}_2(\mathbb{P}_{\mathbf{C}}^2)$.

Let f be an element of J_d . In affine coordinates,

$$f(x, y) = (B_y(x), A(y))$$

where A is in $\text{PGL}_2(\mathbf{C})$ and B_y takes values into $\text{Aut}(\mathbb{P}_{\mathbf{C}}^1)$. More precisely, B_y is given by a function $B: y \mapsto B(y)$ with values in $\text{GL}_2(\mathbf{C})$:

$$B(y) = \begin{pmatrix} a(y) & b(y) \\ c(y) & d(y) \end{pmatrix} \in \text{GL}_2(\mathbf{C}[y])$$

with coefficients a, c and b, d of degree $d - 1$ and d respectively. The degree of the function $\det(B(y))$ is equal to $2d - 2$; if B is generic, $\det(B(y))$ has $2d - 2$ roots y_i , $1 \leq i \leq 2d - 2$, and $B(y_i)$ is a rank 1 complex matrix for each of these roots. The image of $B(y_i)$ is a line, and this line corresponds to a point x_i in $\mathbb{P}(\mathbf{C}^2)$. The birational transformation f contracts each horizontal line corresponding to a root y_i onto a point $q_i = (x_i, A(y_i))$. This provides $2d - 2$ points of indeterminacy for f^{-1} ; again, if B is generic, the $2d - 2$ points q_i are distinct, generic points of the plane. The same conclusion holds if we change f into its inverse, and gives rise to $2d - 2$ indeterminacy points p_1, \dots, p_{2d-2} for f . One more indeterminacy point (for f and f^{-1}) coincides with $p_0 = q_0 = [1 : 0 : 0]$.

An easy computation shows that the base locus of f is made of

- (1) the point p_0 itself, with multiplicity $d - 1$;
- (2) the $2d - 2$ single points p_1, \dots, p_{2d-2} .

Moreover, all sets of distinct points $\{p_0, p_1, \dots, p_{2d-2}\}$ with $p_0 = [1 : 0 : 0]$ and no three of them on a line through p_0 can be obtained as the indeterminacy set of a de Jonquières transformation of degree d . In particular, on the complement of a strict Zariski closed subset of J_d , the points p_i form a set of $2d - 1$ distinct points in the plane: There are no infinitely near points in the list. Thus, for a generic element of V_d , we have

$$f_*[H] = d[H] - (d - 1)[E_{p_0}] - \sum_{i=1}^{2d-2} [E_{p_i}]$$

where the p_i are generic distinct points of the plane.

Remark 5.1. If $\Sigma \subset \mathbb{P}_{\mathbb{C}}^2$ is a generic set of cardinal k , and h is an automorphism of $\mathbb{P}_{\mathbb{C}}^2$, then h is the identity map as soon as $h(\Sigma) \cap \Sigma$ contains five points. Applied to $\text{Ind}(f)$, we obtain the following: Let g be a generic element of V_d , and h be an automorphism of $\mathbb{P}_{\mathbb{C}}^2$; if h is not the identity map, then $h(\text{Ind}(g)) \cap \text{Ind}(g)$ contains at most four points.

5.1.2. *Tightness is generic.* Our goal is to prove the following statement.

Theorem 5.2. *For all integers $d \geq 2$ there exists a positive integer $k(d)$ such that the following properties are satisfied by a generic element $g \in V_d$:*

- (1) g is a tight Cremona transformation;
- (2) If $n \geq k(d)$, then g^n generates a normal subgroup of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ all non trivial elements f of which satisfy $\lambda(f) \geq d^n$.

Moreover, for d sufficiently large one can take $k(d) = 1$.

Since V_d is irreducible, tight Cremona transformations are dense in V_d ; since V_d is the unique component of $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$ of maximal dimension, properties (1) and (2) are also generic in $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$. Thus Theorem 5.2 implies theorem A.

5.1.3. *Strong algebraic stability is generic.* Given a surface X , a birational transformation $f \in \text{Bir}(X)$ is said to be **algebraically stable** if one of the following equivalent properties hold:

- (1) if x is a point of $\text{Ind}(f)$, then $f^k(x) \notin \text{Ind}(f^{-1})$ for all $k \geq 0$;
- (2) if y is a point of $\text{Ind}(f^{-1})$ then $f^k(y) \notin \text{Ind}(f)$ for all $k \leq 0$;
- (3) the action f^* of f on the Néron-Severi group $N^1(X)$ satisfies $(f^k)^* = (f^*)^k$ for all $k \in \mathbf{Z}$.

If f is an element of $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$, algebraic stability is equivalent to $\lambda(f) = \deg(f)$ (if f is not algebraically stable, property (3) implies $\lambda(f) < \deg(f)$).

Note that condition (2) can be rephrased by saying that f is algebraically stable if

$$f^k(\text{Ind}(f^{-1})) \cap \text{Ind}(f) = \emptyset$$

for all $k \geq 0$. We now prove that generic elements of V_d satisfy a property which is stronger than algebraic stability (recall from Remark 4.2 that $\text{Ind}(f) \subset \text{Exc}(f)$).

Lemma 5.3. *If g is a generic element of V_d then*

$$g^k(\text{Ind}(g^{-1})) \cap \text{Exc}(g) = \emptyset$$

for all $k \geq 0$. In particular g is algebraically stable.

Proof. For a fixed integer $k \geq 0$, the condition

$$g^k(\text{Ind}(g^{-1})) \cap \text{Exc}(g) \neq \emptyset \tag{5.1}$$

is equivalent to the existence of a point $m \in \text{Ind}(g^{-1})$ such that $g^k(m) \in \text{Exc}(g)$, and hence to $\text{Jac}_g(g^k(m)) = 0$, where Jac_g , the Jacobian determinant of g , is the equation of $\text{Exc}(g)$. This is an algebraic condition, that defines an algebraic subvariety I_k in $\text{Bir}_d(\mathbb{P}_{\mathbb{C}}^2)$.

Fact 5.4. *There exists $g \in V_d$ such that $g \notin I_k$ for any $k \geq 0$.*

Since V_d is irreducible, this fact implies that the codimension of I_k is positive and shows that equation (5.1) is satisfied on the intersection of countably many Zariski dense open subsets of V_d .

We are reduced to prove Fact 5.4. For this, let f be the Hénon mapping defined by

$$f: [x : y : z] \dashrightarrow \left[\frac{yz^{d-1}}{10} : y^d + xz^{d-1} : z^d \right].$$

The affine plane $\{z \neq 0\}$ is f -invariant, and f restricts to a polynomial automorphism of this plane. The exceptional set $\text{Exc}(f)$ is the line at infinity $\{z = 0\}$, and its image $\text{Ind}(f^{-1})$ is the point $q = [0 : 1 : 0]$. This point is fixed by f : In the affine chart $\{y \neq 0\}$, with affine coordinates $(x, z) = [x : 1 : z]$ around q , the map f is given by

$$(x, z) \mapsto \left(\frac{z^{d-1}}{10(1 + xz^{d-1})}, \frac{z^d}{1 + xz^{d-1}} \right).$$

In particular, q is an attracting fixed point. Let U be the neighborhood of q defined by

$$U = \{[x : 1 : z] ; |x| < 1/10, |z| < 1/10\}.$$

Then, $f(U)$ is contained $\{[x : 1 : z] ; |x| < 1/90, |z| < 1/90\}$. Let h be the linear transformation of the plane which, in the affine coordinates (x, z) , is the translation $(x, z) \mapsto (x + 1/20, z + 1/20)$. We have $h(f(U)) \subset U \setminus \{z = 0\}$, and $f(U) \cap h(f(U)) = \emptyset$. We now take $g = h \circ f$. Its exceptional set is the line at infinity $\{z = 0\}$; the unique indeterminacy point of g^{-1} is $h(q) = [1/20 : 1 : 1/20]$.

By construction, for all $k \geq 0$, $g^k(\text{Ind}(g^{-1}))$ is a point of $U \setminus \text{Exc}(g)$. This proves that g is not in I_k , for any $k \geq 0$. \square

We say that a Cremona transformation g is **strongly algebraically stable** if g and g^{-1} satisfy the property stated in Lemma 5.3. This lemma shows that generic element of \mathbf{V}_d are strongly algebraically stable.

Note that algebraic stability implies that g is well defined along the forward orbit $g^k(\text{Ind}(g^{-1}))$, $k \geq 0$; similarly, g^{-1} is well defined along the backward orbit of $\text{Ind}(g)$.

Lemma 5.5. *Let g be strongly algebraically stable. Then we have:*

- (1) For all $k > j \geq 0$, $g^k(\text{Ind}(g^{-1})) \cap g^j(\text{Ind}(g^{-1})) = \emptyset$;
- (2) For all $k > j \geq 0$, $g^{-k}(\text{Ind}(g)) \cap g^{-j}(\text{Ind}(g)) = \emptyset$;
- (3) For all $k \geq 0$ and $j \geq 0$, $g^k(\text{Ind}(g^{-1})) \cap g^{-j}(\text{Ind}(g)) = \emptyset$.

Proof. Suppose that there is a point $q \in g^k(\text{Ind}(g^{-1})) \cap g^j(\text{Ind}(g^{-1}))$. Since g is well defined along the positive orbit of $\text{Ind}(g^{-1})$, we can apply g^{-j} and assume that $j = 0$ and $q \in g^l(\text{Ind}(g^{-1})) \cap \text{Ind}(g^{-1})$ with $l = k - j > 0$. This means that there exists $p \in \text{Ind}(g^{-1})$ such that $g^l(p) = q$. But then $g^{l-1}(p) \in \text{Exc}(g)$, contradicting the assumption.

Property (2) is proven similarly, replacing g by g^{-1} .

Suppose (3) is false. Then there exists $p \in \text{Ind}(g^{-1})$, $q \in \text{Ind}(g)$, $k, j \geq 0$ such that $g^k(p) = g^{-j}(q)$. This contradicts the algebraic stability of g . \square

5.1.4. *Rigidity is generic.* Let g be an element of \mathbf{V}_d . Consider the isometry g_* of $\bar{\mathcal{Z}}$. If g is algebraically stable, its dynamical degree is equal to d and the translation length of g_* is equal to $\log(d)$. The Picard-Manin classes $[\alpha]$ and $[\omega]$ corresponding to the end points of the axis of g_* satisfy $g_*[\alpha] = [\alpha]/d$ and $g_*[\omega] = d[\omega]$. To compute explicitly such classes, we start with the class $[H]$ of a line in $\mathbb{P}_{\mathbb{C}}^2$, and use the fact that

$$\frac{1}{d^n} g_*^n [H] \rightarrow c^{ste}[\omega]$$

when n goes to $+\infty$ (see § 4.5.2).

Assuming that g is generic in \mathbf{V}_d , its base locus is made of one point p_0 of multiplicity $d - 1$ and $2d - 2$ points p_i , $1 \leq i \leq 2d - 2$, of multiplicity 1, and similarly for the base locus of g^{-1} . Note $[E^+]$ (resp. $[E^-]$) the sum of the classes of the exceptional divisors, with multiplicity $d - 1$ for the first one, obtained by blowing-up the $2d - 1$ distinct points in $\text{Ind}(g)$ (resp. $\text{Ind}(g^{-1})$). We have

$$g_*[H] = d[H] - [E^-], \quad g_*^2[H] = d^2[H] - d[E^-] - g_*[E^-], \quad \text{etc.}$$

Thus, the lines R_g^- and R_g^+ of the Picard-Manin space generated by

$$[\alpha] = [H] - \sum_{i=1}^{\infty} \frac{g_*^{-i+1}[E^+]}{d^i} \quad \text{and} \quad [\omega] = [H] - \sum_{i=1}^{\infty} \frac{g_*^{i-1}[E^-]}{d^i}$$

correspond to the end points of the axis of g_* . By Lemma 5.5, both infinite sums appearing in these formulas are sums of classes of the exceptional divisors obtained by blowing up the backward (resp. forward) orbit of $\text{Ind}(g)$ (resp. $\text{Ind}(g^{-1})$). By construction, $[\alpha]$ and $[\omega]$ satisfy $[\alpha] \cdot [\omega] = 1$ and $[\alpha]^2 = [\omega]^2 = 0$, because both of them are on the boundary of $\mathbb{H}_{\bar{z}}$. All points on $\text{Ax}(g_*)$ are linear combinations $u[\alpha] + v[\omega]$ with the condition

$$1 = (u[\alpha] + v[\omega])^2 = 2uv.$$

The intersection of $[H]$ with a point on $\text{Ax}(g_*)$ is minimal for $u = v = \frac{1}{\sqrt{2}}$; note $[P] = \frac{1}{\sqrt{2}}(\alpha + \omega)$ the point which realizes the minimum. We have

$$[P] = \sqrt{2}[H] - \frac{1}{\sqrt{2}}[R] \quad \text{with} \quad [R] = \frac{[E^+] + [E^-]}{d} + \frac{g_*^{-1}[E^+] + g_*[E^-]}{d^2} + \dots$$

Once again, Lemma 5.5 implies the following fact.

Fact 5.6. *The class $[R]$ is a sum of classes of exceptional divisors obtained by blowing up distinct points of $\mathbb{P}_{\mathbb{C}}^2$ (there is no blow-up of infinitely near points).*

Proposition 5.7. *There exists $\varepsilon_0 > 0$ such that the following property is satisfied for any $d \geq 2$, and for any generic transformation $g \in V_d$. Let $[P]$ be the Picard-Manin class defined above. If f is a birational transformation of the plane such that $\text{dist}(f_*[Q], [Q]) \leq \varepsilon_0$ for all $[Q]$ in $\{g_*^{-1}[P], [P], g_*[P]\}$, then f is the identity map.*

Proof. We proceed in two steps.

First step.— We show that if $\text{dist}(f_*[P], [P]) \leq \varepsilon$ with a small enough ε , then f is linear.

By the triangular inequality

$$\begin{aligned} \text{dist}(f_*[H], [H]) &\leq \text{dist}(f_*[H], f_*[P]) + \text{dist}(f_*[P], [P]) + \text{dist}([P], [H]) \\ &\leq 2\text{dist}([P], [H]) + \varepsilon. \end{aligned}$$

Recall that $\cosh(\text{dist}([D_1], [D_2])) = [D_1] \cdot [D_2]$ for all pairs of points $[D_1], [D_2]$ in $\mathbb{H}_{\bar{z}}$ and that the degree of f is given by $\deg(f) = f_*[H] \cdot [H]$. Using $[H] \cdot [P] = \sqrt{2}$ and $\cosh(2x) = 2\cosh^2(x) - 1$, we get

$$\deg(f) \leq \cosh(t + \varepsilon),$$

where $\cosh(t) = 3$, i.e. $t = \log((3 + \sqrt{5})/2)$. If $\cosh(t + \varepsilon) < 4$, we obtain $\deg(f) \leq 3$.

Note $[D]$ the point on the Picard-Manin space such that $[P]$ is the middle point of the geodesic segment from $[H]$ to $[D]$; explicitly

$$[D] = 2\sqrt{2}[P] - [H] = 3[H] - 2[R].$$

Suppose that $\deg(f) = 3$. Then $\text{dist}([H], f_*[H]) = \cosh^{-1}(3) = \text{dist}([H], [D])$. Since $\text{dist}(f_*[P], [P]) \leq \varepsilon$, we must have $\text{dist}(f_*[H], [D]) \leq \varepsilon'$ where ε' goes to 0 with ε . On the other hand, since f has degree 3,

$$f_*[H] = 3[H] - 2[E_1] - [E_2] - [E_3] - [E_4] - [E_5]$$

for some exceptional divisors E_i above $\mathbb{P}_{\mathbb{C}}^2$ (they can come from blow-ups of infinitely near points). By Fact 5.6, all exceptional classes in the infinite sum defining $[R]$ come from blow-ups of distinct points; hence

$$[2E_1 + E_2 + E_3 + E_4 + E_5] \cdot [R] \geq \frac{1}{d}(-2(d-1) - 1 - 1 - 1 - 1) \geq -3.$$

This implies that

$$\begin{aligned} f_*[H] \cdot [D] &= f_*[H] \cdot (3[H] - 2[R]) \\ &= 9 + 2[2E_1 + E_2 + E_3 + E_4 + E_5] \cdot [R] \\ &\geq 3. \end{aligned}$$

So $f_*[H] \cdot [D] \geq 3$ and the distance between $f_*[H]$ and $[D]$ cannot be smaller than ε' if ε is small enough, contradicting $\deg(f) = 3$.

Suppose that $\deg(f) = 2$. We have $f_*[H] = 2[H] - [E_1] - [E_2] - [E_3]$ where the $[E_i]$ are classes of exceptional divisors above $\mathbb{P}_{\mathbb{C}}^2$. The product $f_*[H] \cdot [P]$ is close to $[H] \cdot [P] = \sqrt{2}$, and is given by

$$\begin{aligned} f_*[H] \cdot [P] &= [2H - E_1 - E_2 - E_3] \cdot \left(\sqrt{2}[H] - \frac{1}{\sqrt{2}}[R] \right) \\ &= 2\sqrt{2} + \frac{1}{\sqrt{2}}[E_1 + E_2 + E_3] \cdot [R]. \end{aligned}$$

As a consequence, the intersection product of $[E_1 + E_2 + E_3]$ and $[R]$ must be close to -2 . By Fact 5.6, the curves E_1 , E_2 , and E_3 appear at most once in $[R]$, so that $[E_1 + E_2 + E_3] \cdot [R] \geq -1 - \frac{1}{d} \geq -\frac{3}{2}$; this is a contradiction and $\deg(f) \neq 2$.

Second step.— Now we show that if f is linear and $\text{dist}(f_*[Q], [Q]) \leq \varepsilon_0$ for all $[Q]$ in the segment $[g_*^{-1}[P], g_*[P]]$, then f is the identity map.

Applying the assumption to $[Q] = g_*^{-1}[P]$, we get

$$\text{dist}((gfg^{-1})_*[P], [P]) \leq \varepsilon_0$$

and the first step shows that gfg^{-1} must be linear. This implies that $f(\text{Ind}(g))$ coincides with $\text{Ind}(g)$. The same argument with $[Q] = g_*[P]$ gives $f(\text{Ind}(g^{-1})) = \text{Ind}(g^{-1})$. If $\text{deg}(g) \geq 3$, the set $\text{Ind}(g)$ is a generic set of points in $\mathbb{P}_{\mathbb{C}}^2$ with cardinal at least 5, so $f(\text{Ind}(g)) = \text{Ind}(g)$ implies that f is the identity map. Finally, if $\text{deg}(g) = 2$, the set $\text{Ind}(g) \cup \text{Ind}(g^{-1})$ is a generic set of 6 points in the plane, so if $f(\text{Ind}(g)) = \text{Ind}(g)$ and $f(\text{Ind}(g^{-1})) = \text{Ind}(g^{-1})$ then again f is the identity map. \square

Corollary 5.8. *If $g \in V_d$ is a generic birational map, then $\text{Ax}(g_*)$ is rigid.*

Proof. Suppose $\text{Ax}(g_*)$ is not rigid. Choose $\varepsilon > 0$. By Proposition 3.3 there exists f which does not preserve $\text{Ax}(g_*)$ such that $d(f_*[Q], [Q]) < \varepsilon$ for all $[Q]$ in the segment $[g_*^{-1}[P], g_*[P]]$. For small enough ε , this contradicts Proposition 5.7. \square

Proof of Theorem 5.2. Let g be a generic element of V_d , with $d \geq 2$. Suppose that $f_*(\text{Ax}(g_*)) = \text{Ax}(g_*)$. If f_* preserves the orientation on $\text{Ax}(g_*)$, then $(fgf^{-1}g^{-1})_*$ fixes each point in $\text{Ax}(g)$, and Proposition 5.7 gives $fgf^{-1} = g$. Similarly, if f_* reverses the orientation, considering $(fgf^{-1}g)_*$ we obtain $fgf^{-1} = g^{-1}$.

Since we know by Corollary 5.8 that $\text{Ax}(g_*)$ is rigid, we obtain that g_* is tight, hence by Theorem 2.9 g_*^k generates a non trivial subgroup of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ for large enough k .

We can be more precise on k by reconsidering the proof of Corollary 5.8. The rigidity of the axis of g_* follows from Proposition 5.7, where the condition on ε is independent of d . On the other hand Lemma 3.1 relates ε to B by

$$\cosh(\varepsilon) - 1 \leq c^{ste}(\cosh(B))^{-1}.$$

Thus, we obtain the rigidity property if B is larger than a constant independent of d . Then, the proof of Theorem 2.9 only requires $L(g_*^k) \geq 40B + c^{ste}$. Thus, if $L(g_*) = \log(d)$ is sufficiently large, one can take $k = 1$. \square

5.2. Automorphisms of rational surfaces.

5.2.1. *Preliminaries.* Let g be an automorphism of a rational surface X . If $\pi: X \dashrightarrow \mathbb{P}_{\mathbf{C}}^2$ is a birational map, conjugation by π provides a natural isomorphism between $\text{Bir}(X)$ and $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ and π_* conjugates the actions of $\text{Bir}(X)$ on $\bar{\mathcal{Z}}(X)$ and $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ on $\bar{\mathcal{Z}}(\mathbb{P}_{\mathbf{C}}^2)$. We thus identify $\text{Bir}(X)$ to $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ and $\bar{\mathcal{Z}}(X)$ to $\bar{\mathcal{Z}}(\mathbb{P}_{\mathbf{C}}^2)$ without further reference to the choice of π . This paragraph provides a simple criterion to check whether g is a tight element of the Cremona group $\text{Bir}(X) = \text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$.

As explained in §4.5.3, the Néron-Severi group $\mathbf{N}^1(X)$ embeds into the Picard-Manin space $\bar{\mathcal{Z}}$ and g_* preserves the orthogonal decomposition

$$\bar{\mathcal{Z}} = \mathbf{N}^1(X)_{\mathbf{R}} \oplus \mathbf{N}^1(X)_{\mathbf{R}}^{\perp},$$

where $\mathbf{N}^1(X)_{\mathbf{R}}^{\perp}$ is the orthogonal complement with respect to the intersection form. Assume that g_* is hyperbolic, with axis $\text{Ax}(g_*)$ and translation length $\log(\lambda(g))$. Lemma 4.6 shows that the plane V_g is contained in $\mathbf{N}^1(X)_{\mathbf{R}}$ and the axis of g_* : $\mathbb{H}_{\bar{\mathcal{Z}}} \rightarrow \mathbb{H}_{\bar{\mathcal{Z}}}$ is equal to $V_g \cap \mathbb{H}_{\bar{\mathcal{Z}}}$.

Remark 5.9. Let h be a birational transformation of X such that h_* preserves V_g . The restriction of h_* to V_g satisfies one of the following two properties.

- (1) h_* and g_* commute on V_g ;
- (2) h_* is an involution of V_g and h_* conjugates g_* to its inverse on V_g .

This is because the group of isometries of a hyperbolic quadratic form in two variables is isomorphic to the semi-direct product $\mathbf{R} \rtimes \mathbf{Z}/2\mathbf{Z}$.

Lemma 5.10. *Let g be a hyperbolic automorphism of a rational surface X . Assume that*

- (i) g_* is the identity on the orthogonal complement of V_g in $\mathbf{N}^1(X)$;
- (ii) the action of $\text{Aut}(X)$ on $\mathbf{N}^1(X)$ is faithful.

Let h be an automorphism of X such that h_ preserves V_g . Then either $hgh^{-1} = g$ or g^{-1} .*

Proof. Let us study the action of h_* and g_* on $\mathbf{N}^1(X)$. Since h_* preserves V_g , it preserves its orthogonal complement V_g^{\perp} and, by assumption (i), commutes to g_* on V_g^{\perp} . If h_* commutes to g_* on V_g , then h_* and g_* commute (on $\mathbf{N}^1(X)$), and the conclusion follows from the second assumption. If h_* does not commute to g , Remark 5.9 implies that h_* is an involution on V_g , and that $h_*g_*(h_*)^{-1} = (g_*)^{-1}$ on V_g and therefore also on $\mathbf{N}^1(X)$. Once again, assumption (ii) implies that h conjugates g to its inverse. \square

Remark 5.11. The plane V_g is a subspace of $\mathbf{N}^1(X)_{\mathbf{R}}$ and it may very well happen that this plane does not intersect the lattice $\mathbf{N}^1(X)$ (except at the origin). But, if g_* is the identity on V_g^{\perp} , then V_g^{\perp} and V_g are defined over

the integers. In that case, $V_g \cap \mathbf{N}^1(X)$ is a rank 2 lattice in the plane V_g . This lattice is g_* -invariant, and the spectral values of the linear transformation $g_* \in \mathrm{GL}(V_g)$ are quadratic integers. From this follows that $\lambda(g)$ is a quadratic integer. In the opposite direction, if $\lambda(g)$ is a quadratic integer, then V_g is defined over the integers, and so is V_g^\perp . The restriction of g_* to V_g^\perp preserves the lattice $V_g^\perp \cap \mathbf{N}^1(X)$ and a negative definite quadratic form. This implies that a positive iterate of g_* is the identity on V_g^\perp .

Let us now assume that there exists an ample class $[D'] \in \mathbf{N}^1(X) \cap V_g$. In that case, $[D']$ and $g_*[D']$ generate a rank 2 subgroup of $\mathbf{N}^1(X) \cap V_g$ and $\lambda(g)$ is a quadratic integer (by Remark 5.11). In what follows, we shall denote by $[D]$ the class

$$[D] = \frac{1}{\sqrt{[D'] \cdot [D']}} [D']. \quad (5.2)$$

This is an ample class with real coefficients which determines a point $[D]$ in $\mathbb{H}_{\bar{\mathbb{Z}}}$.

Lemma 5.12. *Let h be a birational transformation of a projective surface X . Let $[D'] \in \mathbf{N}^1(X)$ be an ample class, and $[D] = [D']/\sqrt{[D'] \cdot [D']} \in \mathbb{H}_{\bar{\mathbb{Z}}}$. If*

$$\cosh(\mathrm{dist}(h_*[D], [D])) < 1 + ([D'] \cdot [D'])^{-1}$$

then h_ fixes $[D']$ and is an automorphism of X .*

Proof. Write $h_*[D] = [D] + [F] + [R]$ where $[F]$ is in $\mathbf{N}^1(X)_{\mathbf{R}}$ and $[R]$ is in $\mathbf{N}^1(X)_{\mathbf{R}}^\perp$. More precisely, F is an element of $\mathbf{N}^1(X)$ divided by the square root of the self intersection $[D']^2$, and $[R]$ is a sum $\sum m_i [E_i]$ of exceptional divisors obtained by blowing up points of X , coming from indeterminacy points of h ; the m_i are integers divided by the square root of $[D']^2$. We get

$$1 \leq [D] \cdot h_*[D] = [D] \cdot ([D] + [F] + [R]) = 1 + [D] \cdot [F]$$

because $[D]$ does not intersect the $[E_i]$. The number $[D] \cdot [F]$ is a non negative integer divided by $[D']^2$. By assumption, this number is less than $([D'] \cdot [D'])^{-1}$ and so it must be zero. In other words, the distance between $[D]$ and $h_*[D]$ vanishes, and $[D]$ is fixed. Since $[D]$ is ample, h is an automorphism of X . \square

Proposition 5.13. *Let g be a hyperbolic automorphism of a rational surface X . Assume that (i) V_g contains an ample class $[D']$ and (ii) g_* is the identity on $V_g^\perp \cap \mathbf{N}^1(X)$. Then $Ax(g_*)$ is rigid. Moreover if we assume that*

(iii) if $h \in \mathrm{Aut}(X)$ and $h_(Ax(g_*)) = Ax(g_*)$, then either h commutes with g or $hgh^{-1} = g^{-1}$;*

Then any $h \in \mathrm{Bir}(X)$ which preserves $Ax(g_)$ is an automorphism of X , and g is a tight element of $\mathrm{Bir}(X)$. Thus, for sufficiently large n the iterate g^n generates a non trivial normal subgroup in the Cremona group $\mathrm{Bir}(\mathbb{P}_{\mathbf{C}}^2) = \mathrm{Bir}(X)$.*

Note that if the action of $\text{Aut}(X)$ on $\mathbf{N}^1(X)$ is faithful then by Lemma 5.10 condition (iii) is automatically satisfied.

Proof. Let $[D] \in \text{Ax}(g_*)$ be the ample class defined by equation (5.2). If the axis of g_* is not rigid, Proposition 3.3 provides a birational transformation f of X such that the distances between $f_*[D]$ and $[D]$ and between $f_*(g_*[D])$ and $g_*[D]$ are bounded by $([D']^2)^{-1}$, and, moreover, $f_*(\text{Ax}(g_*)) \neq \text{Ax}(g_*)$. Lemma 5.12 implies that f is an automorphism of X fixing $[D]$ and $g_*[D]$. This contradicts $f_*(\text{Ax}(g_*)) \neq \text{Ax}(g_*)$ and shows that $\text{Ax}(g_*)$ is rigid.

Assume now that $h \in \text{Bir}(X)$ preserves the axis of g_* . Then $h_*[D']$ is an ample class, hence h is an automorphism of X . Property (iii) then implies that g is a tight element of $\text{Bir}(X)$. The conclusion follows from Theorem 2.9. \square

In the next two paragraphs we construct two families of examples which satisfy the assumption of Proposition 5.13. Note that the surfaces X that we shall consider have quotient singularities; if we blow-up X to get a smooth surface, then the class $[D']$ is big and nef but is not ample.

5.2.2. *Generalized Kummer surfaces.* Let $\Lambda \subset \mathbf{C}$ be the lattice of Gaussian integers $\mathbf{Z}[i]$, and Y be the abelian surface $\mathbf{C}/\Lambda \times \mathbf{C}/\Lambda$. The group $\text{GL}_2(\mathbf{Z}[i])$ acts by linear transformations on \mathbf{C}^2 , preserving the lattice Λ^2 ; this provides an embedding $\text{GL}_2(\mathbf{Z}[i]) \rightarrow \text{Aut}(Y)$. Let X be the quotient of Y by the action of the group of order 4 generated by $\eta(x, y) = (ix, iy)$. This surface is rational, with sixteen singularities; all of them are quotient singularities that can be resolved by a single blow-up. Such a surface is a so called (generalized) **Kummer surface**; classical Kummer surfaces are quotient of tori by $(x, y) \mapsto (-x, -y)$, and are K3-surfaces (these surfaces are not rational).

The linear map η generates the center of the group $\text{GL}_2(\mathbf{Z}[i])$. As a consequence, $\text{GL}_2(\mathbf{Z}[i])$, or more precisely $\text{PGL}_2(\mathbf{Z}[i])$, acts by automorphisms on X . Let M be an element of $\text{SL}_2(\mathbf{Z})$ such that

- (i) the trace $\text{tr}(M)$ of M is at least 3;
- (ii) M is in the level 2 congruence subgroup of $\text{SL}_2(\mathbf{Z})$, i.e. M is equal to the identity modulo 2.

Let \hat{g} be the automorphism of Y defined by M and g be the automorphism of X induced by \hat{g} .

Proposition 5.14. *The automorphism $g: X \rightarrow X$ satisfies properties (i), (ii), and (iii) of Proposition 5.13. In particular, g determines a tight element of $\text{Bir}(\mathbb{P}^2(\mathbf{C}))$; hence, if n is large enough, g^n generates a non trivial normal subgroup of $\text{Bir}(\mathbb{P}^2(\mathbf{C}))$.*

Proof. The Néron-Severi group of Y has rank 4, and is generated by the following classes: The class $[A]$ of horizontal curves $\mathbf{C}/\mathbf{Z}[i] \times \{*\}$, the class $[B]$ of vertical curves $\{*\} \times \mathbf{C}/\mathbf{Z}[i]$, the class $[\Delta]$ of the diagonal $\{(z, z) \in Y\}$, and the class $[\Delta_i]$ of the graph $\{(z, iz) \in Y\}$.

The vector space $H^2(Y, \mathbf{R})$ is isomorphic to the space of bilinear alternating two forms on the 4-dimensional real vector space \mathbf{C}^2 . The action of g_* is given by the action of M^{-1} on this space. The dynamical degree of \hat{g} is equal to the square of the spectral radius of M , i.e. to the quadratic integer

$$\lambda(\hat{g}) = \frac{1}{2} \left(a + \sqrt{a^2 - 4} \right)$$

where $a = \text{tr}(M)^2 - 2 > 2$. Thus, the plane $V_{\hat{g}}$ intersects $\mathbf{N}^1(Y)$ on a lattice. Let $[F]$ be an element of $V_{\hat{g}} \cap \mathbf{N}^1(X)$ with $[F]^2 > 0$. Since Y is an abelian variety, then $[F]$ (or its opposite) is ample. Since M has integer coefficients, the linear map \hat{g}_* preserves the three dimensional subspace W of $\mathbf{N}^1(Y)$ generated by $[A]$, $[B]$ and $[\Delta]$. The orthogonal complement of $V_{\hat{g}}$ intersects W on a line, on which \hat{g}_* must be the identity, because $\det(M) = 1$. The orthogonal complement of W is also a line, so that \hat{g}_* is the identity on $V_{\hat{g}}^\perp \subset \mathbf{N}^1(Y)$.

Transporting this picture in $\mathbf{N}^1(X)$, we obtain: The dynamical degree of g is equal to the dynamical degree of \hat{g} (see [Gue07], for more general results), the plane $V_{\hat{g}}$ surjects onto V_g , the image of $[F]$ is an ample class $[D']$ contained in $V_g \cap \mathbf{N}^1(X)$, and g_* is the identity on V_g^\perp .

Automorphisms of X permute the sixteen singularities of X . The fundamental group Γ of $X \setminus \text{Sing}(X)$ is the affine group $\mathbf{Z}/4\mathbf{Z} \ltimes (\mathbf{Z}[i] \times \mathbf{Z}[i])$ where $\mathbf{Z}/4\mathbf{Z}$ is generated by η . The abelian group $\mathbf{Z}[i] \times \mathbf{Z}[i]$ is the unique maximal free abelian subgroup of rank 4 in Γ and, as such, is invariant under all automorphisms of Γ . This implies that all automorphisms of X lift to (affine) automorphisms of Y .

Let h be an automorphism of X which preserves the axis $\text{Ax}(g_*)$. Then h_* conjugates g_* to g_* or $(g_*)^{-1}$ (Remark 5.9), and we must show that h conjugates g to g or g^{-1} . Let \hat{h} be a lift of h to Y . There exists a linear transformation $N \in \text{GL}_2(\mathbf{C})$ and a point $(a, b) \in Y$ such that

$$h(x, y) = N(x, y) + (a, b).$$

The lattice $\mathbf{Z}[i] \times \mathbf{Z}[i]$ is N -invariant, and N conjugates M to M or its inverse M^{-1} , because h_* conjugates g_* to g_* or its inverse. Then

$$\hat{h} \circ \hat{g} \circ \hat{h}^{-1} = M^\pm(x, y) + (\text{Id} - M^\pm)(a, b).$$

On the other hand, since \hat{h} is a lift of an automorphism of X , the translation $t : (x, y) \mapsto (x, y) + (a, b)$ is normalized by the cyclic group generated by η . Thus a and b are in $(1/2)\mathbf{Z}[i]$. Since M is the identity modulo 2, we have

$M(a, b) = (a, b)$ modulo $\mathbf{Z}[i] \times \mathbf{Z}[i]$. Hence $\hat{h} \circ \hat{g} \circ \hat{h}^{-1} = \hat{g}^{\pm}$ and, coming back to X , h conjugates g to g or its inverse. \square

Remark 5.15. The lattice of Gaussian integers can be replaced by the lattice of Eisenstein integers $\mathbf{Z}[j] \subset \mathbf{C}$, with $j^3 = 1$, $j \neq 1$, and the homothety η by $\eta(x, y) = (jx, jy)$. This leads to a second rational Kummer surface with an action of the group $\mathrm{PSL}_2(\mathbf{Z})$, and a statement similar to Theorem B can be proved for this example.

5.2.3. *Coble surfaces.* Let $S \subset \mathbb{P}_{\mathbf{C}}^2$ be a rational sextic with ten double points m_i , $1 \leq i \leq 10$, such that (i) no three of the m_i are on a line (ii) no six of them are on a conic and (iii) there is no cubic containing all the m_i . Such sextic curves exist, and have been studied by Halphen (see [Hal82], or the appendix of [Giz80]). Let X be the surface obtained by blowing up the ten double points of S , and let $\pi: X \rightarrow \mathbb{P}_{\mathbf{C}}^2$ be the natural projection. By definition X is the **Coble surface** defined by S .

Let E_i , $1 \leq i \leq 10$, be the exceptional divisors of π . The canonical class of X is

$$[K_X] = -3[H] + \sum_{i=1}^{10} [E_i]$$

where $[H]$ is the pullback (by π) of the class of a line in the Néron-Severi group of $\mathbb{P}_{\mathbf{C}}^2$. The strict transform S' of S is an irreducible divisor of X , and its class coincides with $-2[K_X]$; more precisely, there is a meromorphic section Ω of $2K_X$ that does not vanish and has simple poles along S' . The self-intersection of S' is -2 , and S' can be blown down: Blowing down S' , one constructs a birational morphism $q: X \rightarrow X_0$. The surface X_0 has a unique singularity, at $m = q(S')$, and is locally isomorphic to a quadratic cone ($uv = w^2$ in \mathbf{C}^3) in the neighborhood of m . The section Ω defines a holomorphic sections of $K_{X_0}^{\otimes 2}$ that trivializes $2K_{X_0}$; in particular, $H^0(X, -2K_X)$ has dimension 1, and the base locus of $-2K_X$ coincides with S' . The automorphism group $\mathrm{Aut}(X)$ acts linearly on the space of sections of $-2K_X$, and preserves its base locus S' . It follows that q conjugates $\mathrm{Aut}(X)$ and $\mathrm{Aut}(X_0)$.

The rank of the Néron-Severi group $\mathbf{N}^1(X)$ is equal to 11. Let W be the orthogonal complement of $[K_X]$ with respect to the intersection form. The linear map $q^*: \mathbf{N}^1(X_0) \rightarrow \mathbf{N}^1(X)$ provides an isomorphism between $\mathbf{N}^1(X_0)$ and its image W . Let $O'(\mathbf{N}^1(X))$ be the group of isometries of the lattice $\mathbf{N}^1(X)$ which preserve the canonical class $[K_X]$ and the cone of ample classes.

Theorem 5.16 (Coble's theorem, see [DO88, Dol08]). *Let X be a Coble surface. The morphism $\mathrm{Aut}(X) \rightarrow O'(\mathbf{N}^1(X))$, $f \mapsto (f^*)^{-1}$, is injective and its image is the level 2 congruence subgroup of $O'(\mathbf{N}^1(X))$.*

Remark 5.17. Let $\text{Amp}(W)$ be the set of ample classes in $W \otimes \mathbf{R}$. This convex cone is invariant under the action of $\text{Aut}(X)$, hence under the action of a finite index subgroup of the orthogonal group $O(W)$. As such, $\text{Amp}(W)$ is equal to $\{[D] \in W \mid [D]^2 > 0, [D] \cdot [H] > 0\}$.

The theorem and the remark give a recipe to construct automorphisms of X . Let ψ be an isometry of the lattice W . If ψ is equal to the identity modulo 2, and $\psi[H] \cdot [H] > 0$, then $\psi = g_*$ for a unique automorphism of X . Let us apply this idea to cook up a tight automorphism of X .

Lemma 5.18 (Pell -Fermat equation). *Let $Q(u, v) = au^2 + buv + cv^2$ be a quadratic binary form with integer coefficients. Assume that Q is non degenerate, indefinite, and does not represent 0. Then there exists an isometry ϕ in the orthogonal group $O_Q(\mathbf{Z})$ with eigenvalues $\lambda(Q) > 1 > 1/\lambda(Q) > 0$.*

Sketch of the proof. Isometries in $O_Q(\mathbf{Z})$ correspond to units in the quadratic field defined by the polynomial $Q(t, 1) \in \mathbf{Z}[t]$; finding units, or isometries, is a special case of Dirichlet's units theorem and amounts to solve a Pell-Fermat equation (see [FT93], chapter V.1). \square

Let $[D_1]$ and $[D_2]$ be the classes

$$\begin{aligned} [D_1] &= 6[H] - [E_9] - [E_{10}] - \sum_{i=1}^8 2[E_i], \\ [D_2] &= 6[H] - [E_7] - [E_8] - \sum_{i \neq 7,8} 2[E_i]. \end{aligned}$$

Both of them have self intersection 2, are contained in W , and intersect $[H]$ positively; as explained in Remark 5.17, $[D_1]$ and $[D_2]$ are ample classes of X_0 . Let V be the plane containing $[D_1]$ and $[D_2]$, and Q be the restriction of the intersection form to V . If u and v are integers, then

$$Q(u[D_1] + v[D_2]) = (u[D_1] + v[D_2]) \cdot (u[D_1] + v[D_2]) = 2u^2 + 8uv + 2v^2$$

because $[D_1] \cdot [D_2] = 4$. This quadratic form does not represent 0, because the discriminant, 12, is not a perfect square. From Lemma 5.18, there is an isometry ϕ of V with an eigenvalue $\lambda(\phi) > 1$. An explicit computation shows that the group of isometries of Q is the semi-direct product of a cyclic group \mathbf{Z} , generated by ϕ , with

$$\phi([D_1]) = 4[D_1] + [D_2], \quad \phi([D_2]) = -[D_1],$$

and the group $\mathbf{Z}/2\mathbf{Z}$ generated by involution which permutes $[D_1]$ and $[D_2]$. The second iterate of ϕ is equal to the identity modulo 2. Coble's theorem (Theorem 5.16) now implies that there exists a unique automorphism g of X such that

- (1) g_* coincides with ϕ^2 on V ;
- (2) g_* is the identity on the orthogonal complement V^\perp .

The dynamical degree of g is the square of $\lambda(\phi)$, and is equal to $7 + 4\sqrt{3}$. Properties (i), (ii), and (iii) of Proposition 5.13 are satisfied. Thus, large powers of g generate non trivial normal subgroups of the Cremona group:

Proposition 5.19. *Let X be a Coble surface. There are hyperbolic automorphisms of X that generate non trivial normal subgroups of the Cremona group.*

6. COMPLEMENTS

6.1. Polynomial automorphisms and monomial transformations.

The group of polynomial automorphisms of the affine plane, and the group of monomial transformations of $\mathbb{P}_{\mathbb{C}}^2$ were both sources of inspiration for the results in this paper. We now use these two groups to construct hyperbolic elements g of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ for which $\langle\langle g \rangle\rangle$ coincides with $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$.

6.1.1. *Monomial transformations.* Consider the group of monomial transformations of $\mathbb{P}_{\mathbb{C}}^2$. By definition, this group is isomorphic to $\text{GL}_2(\mathbb{Z})$, acting by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : (x, y) \mapsto (x^a y^b, x^c y^d)$$

in affine coordinates (x, y) . The matrix $-\text{Id}$ corresponds to the standard quadratic involution $\sigma(x, y) = (1/x, 1/y)$.

If one considers $\text{PSL}_2(\mathbb{Z})$ as a subgroup of $\text{PSL}_2(\mathbb{R}) \simeq \text{Isom}(\mathbb{H}^2)$, it is an interesting exercise to check that all hyperbolic matrices $\text{PSL}_2(\mathbb{Z})$ are tight elements of $\text{PSL}_2(\mathbb{Z})$. However, when we see $\text{GL}_2(\mathbb{Z})$ as a subgroup of the Cremona group, we obtain the following striking remark.

Proposition 6.1. *The normal subgroup of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ generated by any monomial transformation is never proper: If g is not the identity, then $\langle\langle g \rangle\rangle = \text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$.*

Remark 6.2 (Gizatullin and Noether). If N is a normal subgroup of $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ containing a non trivial automorphism of $\mathbb{P}_{\mathbb{C}}^2$, then N coincides with $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$. The proof is as follows. Since $\text{Aut}(\mathbb{P}_{\mathbb{C}}^2)$ is the simple group $\text{PGL}_3(\mathbb{C})$ and N is normal, N contains $\text{Aut}(\mathbb{P}_{\mathbb{C}}^2)$. In particular, N contains the automorphism h defined by

$$h(x, y) = (1 - x, 1 - y)$$

in affine coordinates. An easy calculation shows that the standard quadratic involution σ satisfies $\sigma = (h\sigma)h(h\sigma)^{-1}$; hence, σ is conjugate to h , and σ is contained in N . The conclusion follows from Noether's theorem, which states that σ and $\text{Aut}(\mathbb{P}_{\mathbb{C}}^2)$ generate $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$.

Proof. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be any non trivial monomial map in $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$. The commutator of g with the diagonal map $f(x, y) = (\alpha x, \beta y)$ is the diagonal map

$$g^{-1} f^{-1} g f : (x, y) \mapsto (\alpha^{1-d} \beta^b x, \alpha^c \beta^{1-a} y). \quad (6.1)$$

Thus, the normal subgroup $\langle\langle g \rangle\rangle$ contains an element of $\text{Aut}(\mathbb{P}_{\mathbf{C}}^2) \setminus \{Id\}$ and Remark 6.2 concludes the proof. \square

6.1.2. *Polynomial automorphisms.* Danilov proved that the group of polynomial automorphisms of the affine plane \mathbf{C}^2 with Jacobian determinant one is not simple (see [Dan74] and [FL10]). Denote this group by $\text{Aut}[\mathbf{C}^2]_{\Omega}$, where Ω stands for the 2-form $dx \wedge dy$. Danilov's proof uses an action on a tree; in fact, Jung's theorem shows that $\text{Aut}[\mathbf{C}^2]_{\Omega}$ is a non trivial amalgamated product, and Bass-Serre theory provides a faithful action of $\text{Aut}[\mathbf{C}^2]_{\Omega}$ on a tree (see [Lam02] and [Lam01]). Since $\text{Aut}[\mathbf{C}^2]_{\Omega}$ is a subgroup of $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$, we also have the action of $\text{Aut}[\mathbf{C}^2]_{\Omega}$ on the hyperbolic space $\mathbb{H}_{\bar{z}}$. It is a nice observation that the notion of hyperbolic isometry coincides in both contexts, and correspond to the exponential growth of the sequence of degrees $\deg(g^n)$.

Fix a complex number $a \in \mathbf{C}^*$ and a polynomial $p \in \mathbf{C}[y]$ of degree $d \geq 2$, and consider the automorphism of \mathbf{C}^2 defined by $h(x, y) = (y, p(y) - ax)$. This automorphism determines an algebraically stable birational transformation of $\mathbb{P}_{\mathbf{C}}^2$, namely

$$h[x : y : z] = [yz^{d-1} : P(y, z) - axz^{d-1} : z^d], \quad (6.2)$$

where $P(y, z) = p(y/z)z^d$. There is a unique indeterminacy point $\text{Ind}(h) = \{[1 : 0 : 0]\}$, and a unique indeterminacy point for the inverse, $\text{Ind}(h^{-1}) = \{[0 : 1 : 0]\}$. This Cremona transformation is hyperbolic, with translation length $L(h_*) = \log(d)$; in particular, the translation length goes to infinity with d .

Proposition 6.3. *For all integers $d \geq 2$, equation (6.2) defines a subset $H_d \subset \mathbf{V}_d$ which depends on $d + 2$ parameters and satisfies: For all h in H_d , h is a hyperbolic, algebraically stable Cremona transformation, but the normal subgroup generated by h coincides with $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$.*

Proof. The automorphism h is the composition of the de Jonquière transformation

$$(x, y) \mapsto (P(y) - ax, y)$$

and the linear map $(x, y) \mapsto (y, x)$. As such, h is an element of \mathbf{V}_d . If f denotes the automorphism $f(x, y) = (x, y + 1)$, then

$$(h^{-1} \circ f \circ h)(x, y) = (x - 1/a, y)$$

is linear (thus, the second step in the proof of proposition 5.7 does not work for h). As a consequence, the commutator $f^{-1} h^{-1} f h$ is linear and non trivial, and

$\langle\langle h \rangle\rangle$ intersects $\text{Aut}(\mathbb{P}_{\mathbf{C}}^2)$ non trivially. The conclusion follows from Remark 6.2. \square

6.2. Projective surfaces.

The reason why we focussed on the group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ comes from the fact that $\text{Bir}(X)$ is small compared to $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ when X is an irrational complex projective surface. The proof of the following proposition illustrates this property.

Proposition 6.4. *Let X be a complex projective surface. If the group $\text{Bir}(X)$ is infinite and simple, then X is birationally equivalent to $C \times \mathbb{P}^1(\mathbf{C})$ where C is a curve with trivial automorphism group.*

Sketch of the proof. Assume, first, that the Kodaira dimension of X is non negative. Replace X by its unique minimal model, and identify $\text{Bir}(X)$ with $\text{Aut}(X)$. The group $\text{Aut}(X)$ acts on the homology of X , and the kernel is equal to $\text{Aut}(X)^0$ up to a finite index group. The action on the homology group provides a morphism to $\text{GL}_n(\mathbf{Z})$ for some $n \geq 1$. Reducing modulo p for large primes p , one sees that $\text{GL}_n(\mathbf{Z})$ is residually finite. Since $\text{Aut}(X)$ is assumed to be simple, this implies that $\text{Aut}(X)$ coincides with $\text{Aut}(X)^0$. This group is abelian (see [Akh95]) for surfaces with non negative Kodaira dimension; thus $\text{Bir}(X) = \text{Aut}(X)$ is not both infinite and simple.

Assume now that X is ruled (and not rational). Up to a birational change of coordinates, X is a product $\mathbb{P}_{\mathbf{C}}^1 \times C$ where C is a smooth curve of genus $g(C) \geq 1$. The group $\text{Bir}(X)$ projects surjectively onto $\text{Aut}(C)$. By simplicity, $\text{Aut}(C)$ must be trivial. In that case, $\text{Bir}(X)$ coincides with the infinite simple group $\text{PGL}_2(\mathcal{M}(C))$ where $\mathcal{M}(C)$ is the field of meromorphic functions of C .

The remaining case is when X is rational, and Theorem A concludes the proof. \square

6.3. SQ-universality and the number of quotients.

A group is said to be **SQ-universal** (or SubQuotient-universal) if any countable group can be embedded into one of its quotients. A pioneering work was [HNN49] where it was proved that the free group over two generators is SQ-universal (even if this terminology seems to have been coined much later by Neumann).

If G is a non elementary hyperbolic group, then G is SQ-universal. This result has been obtained by Delzant and Olshanskii in [Del96] and [OI'95]. It seems reasonable to expect that the Cremona group $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ is also SQ-universal, with a proof similar to Delzant and Olshanskii's proofs, but for $\text{Bir}(\mathbb{P}_{\mathbf{C}}^2)$ acting on $\mathbb{H}_{\bar{\mathbf{Z}}}$ instead of a hyperbolic group acting on itself. Dahmani and Guirardel [DG10] have announced a general result which, thanks to what

is proved in our paper, should easily imply the SQ-universality of the Cremona group. We thus refer to a forthcoming preprint of Dahmani and Guirardel for the SQ-universality.

SQ-universality implies the existence of an uncountable number of non isomorphic quotients. The following weaker result is a direct consequence of Theorem A.

Proposition 6.5. *The Cremona group $\text{Bir}(\mathbb{P}_{\mathbb{C}}^2)$ has an uncountable number of distinct normal subgroups.*

6.4. Quasi-morphisms.

Koji Fujiwara informed us that our proof of Theorems A and B also implies the existence of an infinite dimensional vector space of homogenous quasi-morphisms on the Cremona group. Unfortunately, we are not able to construct any non trivial quasi-morphism with a nice dynamical or (algebro-)geometric flavour.

6.5. Positive characteristic.

The use of the field of complex numbers is probably not essential, and we expect that the Cremona group over any field \mathbf{k} (any characteristic, algebraically closed or not) is not simple, with essentially the same proof. The construction of the Picard-Manin space $\mathbb{H}_{\bar{\mathbb{Z}}}$, as described in section 4, only requires to work with an algebraically closed field. Thus, Theorem C applies for any field \mathbf{k} . The problem is to construct examples of tight elements in the group $\text{Bir}(\mathbb{P}_{\mathbf{k}}^2)$. The first construction, with generic elements, requires a new proof: For genericness, we remove a countable number of hypersurfaces in \mathbb{V}_d , and countably many strict Zariski-closed subsets can cover \mathbb{V}_d when \mathbf{k} is countable. For Kummer and Coble surfaces, the proof seems to extend to arbitrary algebraically closed fields, but we need to check all details. This is the reason why we postpone the proof to a forthcoming version which, we hope, will include examples of tight elements in all \mathbb{V}_d , $d \geq 2$, and on Coble surfaces as well.

REFERENCES

- [Akh95] Dmitri N. Akhiezer. *Lie group actions in complex analysis*. Friedr. Vieweg & Sohn, 1995.
- [BFJ08] Sébastien Boucksom, Charles Favre, and Mattias Jonsson. Degree growth of meromorphic surface maps. *Duke Math. J.*, 141(3):519–538, 2008.
- [BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319. Springer-Verlag, Berlin, 1999.
- [Bla09] Jérémy Blanc. Sous-groupes algébriques du groupe de Cremona. *Transform. Groups*, 14(2):249–285, 2009.
- [Bla10] Jérémy Blanc. Groupes de cremona, connexité et simplicité. *To appear in Ann. Sci. ENS*, 2010.

- [Can07] Serge Cantat. Groupes de transformations birationnelles du plan. *Preprint*, pages 1–50, 2007.
- [CDP90] M. Coornaert, T. Delzant, and A. Papadopoulos. *Géométrie et théorie des groupes*, volume 1441 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1990.
- [Dan74] V. I. Danilov. Non-simplicity of the group of unimodular automorphisms of an affine plane. *Mat. Zametki*, 15:289–293, 1974.
- [DC09] Julie Déserti and Dominique Cerveau. Transformations birationnelles de petit degré. *Preprint*, 2009.
- [Del96] Thomas Delzant. Sous-groupes distingués et quotients des groupes hyperboliques. *Duke Math. J.*, 83(3):661–682, 1996.
- [Dés06a] Julie Déserti. Groupe de Cremona et dynamique complexe: une approche de la conjecture de Zimmer. *Int. Math. Res. Not.*, pages Art. ID 71701, 27, 2006.
- [Dés06b] Julie Déserti. Sur les automorphismes du groupe de Cremona. *Compos. Math.*, 142(6):1459–1478, 2006.
- [DF01] J. Diller and C. Favre. Dynamics of bimeromorphic maps of surfaces. *Amer. J. Math.*, 123(6):1135–1169, 2001.
- [DG10] F. Dahmani and V. Guirardel. *Preprint ???*, 2010.
- [DO88] Igor Dolgachev and David Ortland. Point sets in projective spaces and theta functions. *Astérisque*, (165), 1988.
- [Dol08] Igor V. Dolgachev. Reflection groups in algebraic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 45(1):1–60, 2008.
- [Fav08] Charles Favre. Sous-groupes de type fini du groupe de cremona. *Exposé au séminaire Bourbaki*, 2008.
- [FL10] Jean-Philippe Furter and Stéphane Lamy. Normal subgroup generated by a plane polynomial automorphism. *To appear in Transformation Groups*, 2010.
- [FT93] A. Fröhlich and M. J. Taylor. *Algebraic number theory*. Cambridge University Press, 1993.
- [GdlH90] Étienne Ghys and Pierre de la Harpe. *Sur les groupes hyperboliques d’après Mikhael Gromov*, volume 83 of *Progr. Math.* Birkhäuser Boston, 1990.
- [Giz80] M. H. Gizatullin. Rational G -surfaces. *Izv. Akad. Nauk SSSR Ser. Mat.*, 44(1):110–144, 239, 1980.
- [Gue07] Vincent Guedj. Propriétés ergodiques des applications rationnelles. *Panorama de la Société mathématique de France*, to appear, 30:1–130, 2007.
- [Hal82] G. Halphen. Sur les courbes planes du sixième degré à neuf points doubles. *Bull. Soc. Math. France*, 10:162–172, 1882.
- [HNN49] Graham Higman, B. H. Neumann, and Hanna Neumann. Embedding theorems for groups. *J. London Math. Soc.*, 24:247–254, 1949.
- [Hub06] John Hamal Hubbard. *Teichmüller theory and applications to geometry, topology, and dynamics. Vol. 1*. Matrix Editions, Ithaca, NY, 2006.
- [Lam01] Stéphane Lamy. L’alternative de Tits pour $\text{Aut}[\mathbb{C}^2]$. *J. Algebra*, 239(2):413–437, 2001.
- [Lam02] Stéphane Lamy. Une preuve géométrique du théorème de Jung. *Enseign. Math. (2)*, 48(3-4):291–315, 2002.
- [Laz04] Robert Lazarsfeld. *Positivity in algebraic geometry. I*, volume 48. Springer-Verlag, Berlin, 2004. Classical setting: line bundles and linear series.
- [Man86] Yu. I. Manin. *Cubic forms*, volume 4 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1986.
- [Mum76] David Mumford. Algebraic geometry in mathematical developments arising from hilbert problems. In *(Proc. Sympos. Pure Math., Vol. XXVIII, Northern Illinois Univ., De Kalb, Ill., 1974)*, pages 431–444. Amer. Math. Soc., Providence, R. I., 1976.
- [Ngu09] Dat Dang Nguyen. Composantes irréductibles des transformations de Cremona de degré d . *Thèse de Mathématique, Univ. Nice, France*, 2009.

- [Ol'95] A. Yu. Ol'shanskiĭ. SQ-universality of hyperbolic groups. *Mat. Sb.*, 186(8):119–132, 1995.
- [Ser08] Jean-Pierre Serre. Le groupe de cremona et ses groupes finis. *Exposé au séminaire Bourbaki*, 2008.

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE RENNES, RENNES, FRANCE
E-mail address: `serge.cantat@univ-rennes1.fr`

MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY, UNITED KINGDOM
E-mail address: `s.lamy@warwick.ac.uk`

(on leave from INSTITUT CAMILLE JORDAN, UNIVERSITÉ LYON 1, FRANCE
E-mail address: `lamy@math.univ-lyon1.fr`)