Birational properties of word varieties

Tatiana Bandman, Boris Kunyavskiĭ, and Alexei N. Skorobogatov

April 23, 2025

to Yuri Tschinkel, on his 60-th birtday

Abstract

We prove that the subvariety of $\mathrm{SL}(2) \times \mathrm{SL}(2)$ given by the matrix equation $w(X, Y) = \alpha$, where w is a word in two letters, is closely related to an explicit smooth conic bundle over the associated 'trace surface' in the 3-dimensional affine space. When w is the commutator word, we show that this variety can be irrational if the ground field k is not algebraically closed, answering a question of Rapinchuk, Benyash-Krivetz, and Chernousov. When k is a number field, it satisfies weak approximation with the Brauer–Manin obstruction.

Introduction

Let $w(x, y)$ be a non-trivial word in two letters, that is, an element of the free group $\mathcal{F}_2$ with two generators. Let k be a field of characteristic zero. For $\alpha \in \mathrm{SL}(2)(k)$ we define the *word variety* $S_{w,\alpha}$ as the closed subvariety of $\mathrm{SL}(2) \times \mathrm{SL}(2)$ given by the matrix equation $w(X, Y) = \alpha$. For a survey of the extensive literature on word equations, see [GKP18]. The motivating question of this paper concerns birational properties of $S_{w,\alpha}$, as well as k-rational points on $S_{w,\alpha}$.

One much studied example is the commutator variety $XYX^{-1}Y^{-1} = \alpha$, see [Tho61], [RBKC96] and, more recently, [LL21], [GS22, GMS21]. When α is semisimple and non-central, we show that the commutator variety is a dense open subset of a smooth conic bundle (a Severi–Brauer scheme of relative dimension 1) over the Markoff surface in $\mathbb{A}^3_k$ with equation

$$s^2 + t^2 + u^2 - stu = \mathrm{tr}(\alpha) + 2. \tag{1}$$

The associated Brauer class is $(s^2 - 4, \mathrm{tr}(\alpha) - 2)$; it generates the unramified Brauer group of the 'generic' Markoff surface modulo the Brauer group of the ground field. In particular, the commutator variety is birationally equivalent to the affine subvariety of $\mathbb{A}^5_k$ given by

$$s^2 + t^2 + u^2 - stu - (\mathrm{tr}(\alpha) + 2) = s^2 - 4 + (\mathrm{tr}(\alpha) - 2)x^2 - y^2 = 0.$$

We deduce that the commutator variety is k-rational if and only if the Markoff surface (1) is k-rational, which is the case if and only if either $\mathrm{tr}(\alpha)-2$ or $\mathrm{tr}(\alpha)^2-4$ is a square in k, see Theorem 3.4. In particular, for 'general' $\alpha \in \mathrm{SL}(2)(k)$ the commutator variety can be irrational over k. This gives a negative answer to a question of Rapinchuk, Benyash-Krivetz, and Chernousov [RBKC96, p. 50]. When k is a number field, we show that the Brauer–Manin obstruction is the only obstruction to weak approximation on smooth and proper models of the commutator variety, see Proposition 3.5.

Our method is based on the systematic use of the 'trace polynomial' of the word $w(x,y)$. This is very classical and goes back to Klein and Fricke. Using the Cayley–Hamilton theorem it is easy to see that the function $\mathrm{SL}(2)\times\mathrm{SL}(2)\to\mathbb{A}^1_k$ given by $\mathrm{tr}(w(X,Y))$ is a polynomial $P_w(s,t,u)$ in the variables $s=\mathrm{tr}(X)$, $t=\mathrm{tr}(Y)$, $u=\mathrm{tr}(XY)$. We define the *trace surface* $H_{w,a}$, for $a\in k$, as the surface in $\mathbb{A}^3_k$ given by $P_w(s,t,u)=a$. If $w(x,y)$ belongs to the commutator subgroup of the free group on two letters, and $\alpha\in\mathrm{SL}(2)(k)$ is semisimple and non-central with $\mathrm{tr}(\alpha)=a$, the word variety $S_{w,\alpha}$ is a torsor over the trace surface $H_{w,a}$ for the norm 1 torus associated to the splitting field $k(\sqrt{a^2-4})$ of the characteristic polynomial of α. This implies that $S_{w,\alpha}$ is a dense open subset of an explicit smooth conic bundle over $H_{w,a}$, see Theorem 3.2. This largely reduces the study of solutions of the word equation $w(X,Y)=\alpha$ to the study of rational points on the affine surface $H_{w,a}$.

We are grateful to Tim Browning, Jean-Louis Colliot-Thélène, David Harari, and Yuri Zarhin for helpful discussions and encouragement. The research of the second named author was supported by the Israel Science Foundation grant 1994/20. The work on this paper benefited from the authors' visits to the Max Planck Institute of Mathematics in Bonn and from the visit of the last named author to Pennsylvania State University as a Shapiro Visitor. The hospitality and support of these institutions are gratefully appreciated.

1 Preliminaries

1.1 Reduction of the structure group

For the reader's convenience we recall the definition of the push-forward of torsors, see [Sko01, §2.2] and [Gir71, Proposition III.3.2.1]. Let G_1 be a closed subgroup of an affine algebraic group G over a field k. Let Z be a variety over k and let $\mathcal{T}_1\to Z$ be a right Z-torsor for G_1 for the étale topology. The class of this torsor is an element $[\mathcal{T}_1]$ of the étale cohomology set $\mathrm{H}^1_{\text{ét}}(Z,G_1)$. The *push-forward* of $\mathcal{T}_1$ along the morphism $G_1\hookrightarrow G$ is the quotient $\mathcal{T}$ of $\mathcal{T}_1\times_k G$ by the diagonal action of G_1, where G_1 acts on $\mathcal{T}_1$ on the right and on G on the left. Then $\mathcal{T}\to Z$ inherits a right action of G making it a Z-torsor for G. This gives a map of pointed sets $\mathrm{H}^1_{\text{ét}}(Z,G_1)\to\mathrm{H}^1_{\text{ét}}(Z,G)$ that sends $[\mathcal{T}_1]$ to $[\mathcal{T}]$. When Z is $\mathrm{Spec}(k)$, the étale cohomology sets become Galois cohomology sets.

If a Z-torsor $\mathcal{T} \to Z$ for G is the push-forward of some Z-torsor for G_1, then one says that $\mathcal{T}$ *lifts* to a torsor for G_1, or that the structure group of $\mathcal{T}$ *reduces* to G_1.

If G_1 is commutative, every right Z-torsor for G_1 is also a left Z-torsor for G_1. (In general, a right torsor for G_1 is a left torsor for an inner form of G_1.) In this case we have the following criterion for the reduction of the structure group from G to G_1.

Proposition 1.1 *Let G_1 be a closed subgroup of an affine algebraic group G defined over a field k. Assume that G_1 is commutative. Let Z be a variety over k and let $\mathcal{T}$ be a right Z-torsor for G. The structure group of $\mathcal{T}$ reduces to G_1 if and only if there is a G-equivariant morphism of Z-schemes $\mathcal{T} \to (G_1\backslash G) \times_k Z$.*

Proof. Let $\mathcal{T}_1$ be a right Z-torsor for G_1. Let $\mathcal{T} = (\mathcal{T}_1 \times_k G)/G_1$ be the quotient by the diagonal action of G_1, which acts on $\mathcal{T}_1$ on the right and on G on the left. Thus $\mathcal{T}$ has a right action of G making it a Z-torsor for G, but also a left action of G_1. Let $G_1\backslash\mathcal{T}$ be the quotient by this left action of G_1. We have isomorphisms

$$G_1\backslash\mathcal{T} \cong G_1\backslash(\mathcal{T}_1 \times_k G)/G_1 \cong Z \times_k (G_1\backslash G),$$

since the two actions of G_1 obviously commute, and $G_1\backslash\mathcal{T}_1 = Z$.

Conversely, let $\varphi\colon \mathcal{T} \to (G_1\backslash G)\times_k Z$ be a G-equivariant morphism of Z-schemes, with G acting on the right. Let x_0 be the k-point of $G_1\backslash G$ given by the identity element of $G(k)$. Note that the right action of G_1 on $G_1\backslash G$ preserves x_0. Thus $\mathcal{T}_1 := \varphi^{-1}(x_0 \times_k Z)$ is a closed subvariety of $\mathcal{T}$ stable under the right action of G_1 so that $\mathcal{T}_1 \hookrightarrow \mathcal{T}$ is compatible with the injective homomorphism $G_1 \hookrightarrow G$. Since $\mathcal{T}$ is a right Z-torsor for G, this implies that $\mathcal{T}_1$ is a right Z-torsor for G_1 such that $\mathcal{T}$ is the push-forward of $\mathcal{T}$ along $G_1 \hookrightarrow G$. □

1.2 Norm 1 tori for quadratic extensions

Let k be a field of characteristic zero. For a quadratic extension L/k we write $R_{L/k}(\mathbb{G}_{m,L})$ for the Weil restriction of $\mathbb{G}_{m,L}$. The attached norm 1 torus $R^1_{L/k}(\mathbb{G}_{m,L})$ is defined by the exact sequence of k-tori

$$1 \to R^1_{L/k}(\mathbb{G}_{m,L}) \to R_{L/k}(\mathbb{G}_{m,L}) \xrightarrow{\mathrm{N}} \mathbb{G}_{m,k} \to 1, \qquad (2)$$

where N is induced by the norm map $L \to k$. As a variety, $R^1_{L/k}(\mathbb{G}_{m,L})$ is the affine conic $x^2 - ay^2 = 1$, where $a \in k^\times$ is such that $L = k(\sqrt{a})$. The long exact sequence of Galois cohomology groups attached to (2), by Hilbert's Theorem 90, gives rise to a canonical isomorphism

$$\mathrm{H}^1(k, R^1_{L/k}(\mathbb{G}_{m,L})) \cong k^\times/\mathrm{N}(L^\times).$$

This group classifies isomorphism classes of k-torsors for $R^1_{L/k}(\mathbb{G}_{m,L})$. Such a torsor is an affine conic $x^2 - ay^2 = b$, where $b \in k^\times$, with the obvious action of $R^1_{L/k}(\mathbb{G}_{m,L})$. Its class in $\mathrm{H}^1(k, R^1_{L/k}(\mathbb{G}_{m,L}))$ is the image of b in $k^\times/\mathrm{N}(L^\times)$.

1.3 Affine and projective conics

Let Z be a k-variety and let $\mathcal{T} \to Z$ be a right Z-torsor for $\mathrm{PGL}(n)$. To the class $[\mathcal{T}] \in \mathrm{H}^1_{\text{ét}}(Z, \mathrm{PGL}(n))$ one canonically associates an element $\partial([\mathcal{T}])$ of the Brauer group $\mathrm{Br}(Z) = \mathrm{H}^2_{\text{ét}}(Z, \mathbb{G}_m)$, where ∂ is the connecting map attached to the central extension of algebraic groups

$$1 \to \mathbb{G}_m \to \mathrm{GL}(n) \to \mathrm{PGL}(n) \to 1.$$

In fact, one can construct a Severi–Brauer scheme of relative dimension $n-1$ over Z with class $\partial([\mathcal{T}])$ directly from the torsor $\mathcal{T}$, as follows. The standard n-dimensional representation of $\mathrm{GL}(n)$ gives rise to a transitive action of $\mathrm{PGL}(n)$ on the projective space $\mathbb{P}^{n-1}_k$, which we write as a left action. Let X be the quotient of $\mathcal{T} \times_k \mathbb{P}^{n-1}_k$ by the diagonal action of $\mathrm{PGL}(n)$, acting on $\mathcal{T}$ on the right and on $\mathbb{P}^{n-1}_k$ on the left. The Z-scheme X is étale locally isomorphic to $Z \times_k \mathbb{P}^{n-1}_k$, so $X \to Z$ is a Severi–Brauer scheme of relative dimension $n-1$. We have $[X] = \partial([\mathcal{T}]) \in \mathrm{Br}(Z)$.

Lemma 1.2 *Let Z be a k-variety, let $\mathcal{T} \to Z$ be a right Z-torsor for $\mathrm{PGL}(n)$, and let $X = (\mathcal{T} \times_k \mathbb{P}^{n-1}_k)/\mathrm{PGL}(n)$ be the attached Severi–Brauer scheme over Z of relative dimension $n-1$. If $\mathcal{T}$ lifts to a Z-torsor $\mathcal{T}_1$ for a maximal k-torus $T \subset \mathrm{PGL}(n-1)$, then there is an open embedding $\mathcal{T}_1 \hookrightarrow X$.*

Proof. Since $\mathcal{T}$ is the push-forward of $\mathcal{T}_1$, we have canonical isomorphisms

$$X = (\mathcal{T} \times_k \mathbb{P}^{n-1}_k)/\mathrm{PGL}(n) \cong (\mathcal{T}_1 \times_k \mathrm{PGL}(n) \times_k \mathbb{P}^{n-1}_k)/(T \times_k \mathrm{PGL}(n)) \cong (\mathcal{T}_1 \times_k \mathbb{P}^{n-1}_k)/T,$$

where T acts on $\mathbb{P}^{n-1}_k$ on the left as a subgroup of $\mathrm{PGL}(n)$. Indeed, the actions of T and $\mathrm{PGL}(n)$ on $\mathcal{T}_1 \times_k \mathrm{PGL}(n) \times_k \mathbb{P}^{n-1}_k$ commute, as immediately follows from their definitions.

The restriction of the action of $\mathrm{PGL}(n)$ on $\mathbb{P}^{n-1}_k$ to T has a dense orbit on which T acts freely. Fixing a k-point in this orbit gives an open embedding $T \hookrightarrow \mathbb{P}^{n-1}_k$. We identify T with the image of this embedding, so that T acts on itself by translations. Twisted by $\mathcal{T}_1$, the embedding $T \hookrightarrow \mathbb{P}^{n-1}_k$ gives rise to the desired open embedding $\mathcal{T}_1 \hookrightarrow X$. □

Example 1.3 For $n = 2$ and $Z = \mathrm{Spec}(k)$ we can make Lemma 1.2 explicit. Let $T \subset \mathrm{PGL}(2)$ be a maximal k-torus and let T' be its preimage in $\mathrm{SL}(2)$. Since the centre μ_2 of $\mathrm{SL}(2)$ is the 2-torsion subgroup $T'[2] \subset T'$, the multiplication by 2 map gives an isomorphism $T' \cong T'/\mu_2 \cong T$. So the maximal k-tori in $\mathrm{PGL}(2)$ are also the maximal k-tori in $\mathrm{SL}(2)$.

For an element $g \in \mathrm{SL}(2)(k)$ we write $C_{\mathrm{SL}}(2)(g)$ for the centraliser of g in $\mathrm{SL}(2)$. The torus $T \subset \mathrm{SL}(2)$ is the centraliser $C_{\mathrm{SL}(2)}(g)$ of any non-central $g \in T(k)$. It is easy to check that for a non-central semisimple element $g \in \mathrm{SL}(2)(k)$ with trace $t = \mathrm{tr}(g)$ we have

$$C_{\mathrm{SL}(2)}(g) \simeq R^1_{L/k}(\mathbb{G}_{m,L}),$$

when $L = k(\sqrt{t^2-4})$ is a quadratic extension of k. When $t^2 - 4$ is a square in $k^\times$, we define $L = k \oplus k$. In this case $T = R^1_{L/k}(\mathbb{G}_{m,L})$ is the split torus $\mathbb{G}_{m,k}$.

Let $\mathcal{T}_1$ be a k-torsor for T given by

$$x^2 - (t^2-4)y^2 = b,$$

where $b \in k^\times$. Let $\mathcal{T}$ be the push-forward of $\mathcal{T}_1$ along $T \hookrightarrow \mathrm{PGL}(2)$. The associated Severi–Brauer variety $X = (\mathcal{T} \times_k \mathbb{P}^1_k)/\mathrm{PGL}(2)$ is the projective conic given by the homogenised equation

$$x^2 - (t^2-4)y^2 = bz^2.$$

The open embedding $\mathcal{T}_1 \subset X$ of Lemma 1.2 is the natural embedding with complement the union of two $\bar{k}$-points.

Remark 1.4 In the above construction we can replace a maximal torus in PGL(2) by the centraliser of an element $g \in \mathrm{SL}(2)(k)$ such that $\mathrm{tr}(g) = \pm 2$ and $g \neq \pm I$. In this case $G := C_{\mathrm{PGL}(2)}(g)$ is isomorphic to the additive group $\mathbb{G}_{a,k}$. The restriction of the action of PGL(2) on $\mathbb{P}^1_k$ to G has a dense orbit isomorphic to $\mathbb{A}^1_k$ on which G acts freely. Thus we have an analogue of Lemma 1.2 in the unipotent case (with the same notation): if $\mathcal{T}$ lifts to a Z-torsor $\mathcal{T}_1$ for G, then there is an open embedding $\mathcal{T}_1 \hookrightarrow X$. When $Z = \mathrm{Spec}(k)$ we have $\mathrm{H}^1(k, \mathbb{G}_{a,k}) = 0$ by the additive version of Hilbert's Theorem 90, hence any k-torsor for G is trivial and so is isomorphic to $\mathbb{A}^1_k$. This implies that $X \cong \mathbb{P}^1_k$.

2 Simultaneous similarity of two matrices

2.1 Markoff surfaces and torsors for PGL(2)

Let k be a field of characteristic zero with algebraic closure $\bar{k}$. Write

$$F(s,t,u) = s^2 + t^2 + u^2 - stu - 4.$$

For $d \in k$ let $M_d \subset \mathbb{A}^3_k$ be the affine cubic surface given by $F(s,t,u) = d$, called a *Markoff surface*. The surface M_d is smooth if and only if $d \neq 0$ and $d \neq -4$. The singular cubic surface M_0, also called the *Cayley cubic*, has four singular points with coordinates $s, t, u = \pm 2$ with the product of signs equal to 1.

The Cayley cubic M_0 naturally arises in the problem of classification of pairs of (2×2)-matrices up to simultaneous similarity. Write

$$f\colon \mathrm{SL}(2) \times_k \mathrm{SL}(2) \to \mathbb{A}^3_k$$

for the morphism sending $A, B \in \mathrm{SL}(2)(\bar{k})$ to $(s,t,u) \in \mathbb{A}^3_k(\bar{k})$, where

$$s = \mathrm{tr}(A), \quad t = \mathrm{tr}(B), \quad u = \mathrm{tr}(AB).$$

The action of PGL(2) by simultaneous conjugation preserves the fibres of f.

The following proposition is well-known.

Proposition 2.1 *Let $A, B \in \mathrm{SL}(2)(\bar{k})$. The following properties are equivalent:*
(1) *A and B have a common eigenvector;*
(2) $\det(AB - BA) = 0$*;*
(3) $\mathrm{tr}(ABA^{-1}B^{-1}) = 2$*;*
(4) *(s, t, u) is a $\bar{k}$-point of the Cayley cubic M_0.*
These conditions are satisfied when the centralisers of A and B in $\mathrm{PGL}(2)$ *have non-trivial intersection.*

Proof. The equivalence of (1) and (2) is [She84, Theorem 3.2]. The equivalence of (2) and (3) is elementary, as both statements are equivalent to 1 being a root of the characteristic polynomial of $ABA^{-1}B^{-1}$. The equivalence of (3) and (4) follows from the Cayley–Hamilton theorem

$$A^2 = sA - I, \quad B^2 = tB - I, \quad ABAB = uAB - I, \tag{3}$$

from which one obtains the Fricke identity

$$\mathrm{tr}(ABA^{-1}B^{-1}) = s^2 + t^2 + u^2 - stu - 2. \tag{4}$$

See also [Fri83, Theorem 2.9]. For the last statement, let $g \neq \pm I$ be an element of $\mathrm{SL}(2)(\bar{k})$ commuting with A and B. The centraliser of g in $\mathrm{SL}(2)(\bar{k})$ is commutative, hence $AB = BA$, so (2) holds. □

A crucial observation is that $\mathrm{PGL}(2)$ acts freely on $f^{-1}(\mathbb{A}^3_k \setminus M_0)$.

Lemma 2.2 *Let $\mathcal{T} = f^{-1}(V)$, where $V = \mathbb{A}^3_k \setminus M_0$. The morphism $\mathcal{T} \to V$ is a torsor for* $\mathrm{PGL}(2)$ *for the étale topology.*

Proof. The fibres of $f\colon \mathcal{T} \to V$ are orbits of $\mathrm{PGL}(2)$ acting on $\mathrm{SL}(2) \times_k \mathrm{SL}(2)$ by simultaneous conjugation [Fri83, Theorem 2.2]. Since k has characteristic zero, and the varieties $\mathcal{T}$ and V are irreducible and normal, by [GIT94, Proposition 0.2] the morphism $f\colon \mathcal{T} \to V$ is a geometric quotient. By Proposition 2.1, the action of $\mathrm{PGL}(2)$ on $\mathcal{T}$ is set-theoretically free. Since $\mathcal{T}$ and $\mathrm{SL}(2)$ are affine, by Luna's étale slice theorem [GIT94, p. 199], this implies that the action of $\mathrm{PGL}(2)$ on $\mathcal{T}$ is scheme-theoretically free, thus $\mathcal{T} \to V$ is a $\mathrm{PGL}(2)$-torsor for the étale topology. □

The class of $\mathcal{T}$ is an element $[\mathcal{T}]$ of the étale cohomology set $\mathrm{H}^1_{\text{ét}}(V, \mathrm{PGL}(2))$.

2.2 Affine and projective conic bundles

As in §1.3, we associate to the torsor $\mathcal{T} \to V$ a Severi–Brauer scheme $X \to V$ of relative dimension 1, that is, a smooth conic bundle. We would like to compute $[X] \in \mathrm{Br}(V)$.

Let $K = k(s, t, u)$ be the field of functions on $\mathbb{A}^3_k$. Because of the canonical embedding $\mathrm{Br}(V) \subset \mathrm{Br}(K)$ it is enough to compute the K-conic X_K.

Lemma 2.3 *The structure group* $\mathrm{PGL}(2)_K$ *of the generic fibre* $\mathcal{T}_K$ *of* $\mathcal{T} \to V$ *can be reduced to* $T = R^1_{L/K}(\mathbb{G}_{m,L})$*, where* $L = K(\sqrt{t^2-4})$*, and* T *is embedded into* $\mathrm{PGL}(2)_K$ *as the centraliser of an element of* $\mathrm{SL}(2)_K$ *of trace* $t \in K$*.*

Proof. Let $g \in \mathrm{SL}(2)(K)$ be a matrix of trace $t \in K$. The field extension L/K is the splitting field of the characteristic polynomial of g, and thus the centraliser of g in $\mathrm{SL}(2)_K$ is isomorphic to T. As above, $T \cong T/\mu_2$ is also the centraliser of g in $\mathrm{PGL}(2)_K$, giving an embedding $T \hookrightarrow \mathrm{PGL}(2)_K$.

The projection to the second factor $\mathrm{SL}(2) \times_k \mathrm{SL}(2) \to \mathrm{SL}(2)$ sends the generic fibre $\mathcal{T}_K$ to the $\mathrm{PGL}(2)_K$-orbit of $g \in \mathrm{SL}(2)(K)$, where $\mathrm{PGL}(2)_K$ acts by conjugation. The $\mathrm{PGL}(2)_K$-orbit of g is $\mathrm{PGL}(2)_K$-equivariantly isomorphic to $T\backslash\mathrm{PGL}(2)_K$ so that g is identified with the trivial coset of T. By Proposition 1.1, the preimage of g in $\mathcal{T}_K$ is a K-torsor for T, so that $\mathcal{T}_K$ is the push-forward of $\mathcal{T}_1$ along $T \hookrightarrow \mathrm{PGL}(2)_K$. $\square$

In view of Example 1.3 it remains to compute $\mathcal{T}_1$. We shall do a little bit more and describe an affine conic bundle $\pi\colon Y \to \mathbb{A}^3_k$ with generic fibre $Y_K = \mathcal{T}_1$. The trace map $\mathrm{tr}\colon \mathrm{SL}(2) \to \mathbb{A}^1_k$ has a section $\sigma\colon \mathbb{A}^1_k \to \mathrm{SL}(2)$ that sends $t \in \bar{k}$ to the companion matrix

$$g_t = \begin{pmatrix} 0 & -1 \\ 1 & t \end{pmatrix}.$$

Let Y be the preimage of $\sigma(\mathbb{A}^1_k)$ under the second projection $\mathrm{SL}(2)\times_k\mathrm{SL}(2) \to \mathrm{SL}(2)$. Thus the $\bar{k}$-points of Y are pairs of (2×2)-matrices (A, g_t), where $A \in \mathrm{SL}(2)(\bar{k})$ and $t \in \bar{k}$. As above, every fibre $\pi^{-1}(s,t,u)$ is a torsor for the centraliser of g_t in $\mathrm{PGL}(2)$. This centraliser is a torus if g_t is semisimple, otherwise it is isomorphic to $\mathbb{G}_{a,k}$. In particular, the generic fibre $Y_K = \mathcal{T}_1$ is a K-torsor for the torus $T = R^1_{L/K}(\mathbb{G}_{m,L})$.

Proposition 2.4 *All fibres of* $\pi\colon Y \to \mathbb{A}^3_k$ *are non-empty affine conics, in particular,* π *is flat and surjective. The fibres of* π *above the points of* V *are affine subsets of smooth projective conics, the fibres above the smooth points of* M_0 *are singular conics consisting of two* $\bar{k}$*-lines meeting at a point, and the fibres above the singular points of* M_0 *are double lines. The restriction of* π *to the open subset of* $\mathbb{A}^3_k$ *given by* $t^2 - 4 \neq 0$ *is isomorphic to*

$$\mu^2 - (t^2-4)\nu^2 = F(s,t,u). \tag{5}$$

The generic fibre $Y_K = \mathcal{T}_1$ *of* π *is a* K*-torsor for* $T = R^1_{L/K}(\mathbb{G}_{m,L})$ *isomorphic to the affine conic given by* (5)*.*

Proof. The fibre $\pi^{-1}\big((s,t,u)\big)$ consists of pairs (M, g_t), where

$$M = \begin{pmatrix} -x+s & y \\ tx+y-u & x \end{pmatrix}$$

is subject to $\det(M) = -x^2 - y^2 - txy + sx + uy = 1$. This shows that π is surjective with all fibres of dimension 1, hence π is flat. Homogenising this equation, we get the quadratic form

$$-x^2 - y^2 - z^2 - txy + sxz + uyz = 0$$

whose discriminant is $\frac{1}{4}F(s,t,u)$. Thus the fibres of π above the points of $V = \mathbb{A}^3_k \setminus M_0$ are affine subsets of smooth projective conics. Diagonalising this quadratic form, we set

$$\mu = \frac{t^2-4}{2}y + \frac{2u-st}{2}z, \quad \nu = x + \frac{t}{2}y - \frac{s}{2}z, \quad \xi = z,$$

and obtain $\mu^2 - (t^2-4)\nu^2 - F(s,t,u)\xi^2 = 0$. This linear change of variables is given by a matrix with determinant $t^2 - 4$. Setting $\xi = z = 1$ we obtain (5). □

Corollary 2.5 *The class of the generic fibre of the smooth conic bundle $X \to V$ is*

$$[X_K] = (t^2 - 4, F(s,t,u)) \in \mathrm{H}^2(K, \mu_2).$$

Proof. This is immediate from Proposition 2.4 and Example 1.3. □

Note that the equation of the Markoff surface M_d can be written as

$$(s^2-4)(t^2-4) = (2u-st)^2 - 4d. \tag{6}$$

For $d = 0$ we obtain that the class of $F(s,t,u)$ in $k[s,t,u]/(t \pm 2)$ is a square, so $[X_K]$ is indeed unramified on $\mathbb{A}^3_k$ away from M_0 (which of course follows from the construction of X from the V-torsor $\mathcal{T}$). It turns out that $[X_K]$ is also unramified at infinity.

Proposition 2.6 *Let $M_0' \subset \mathbb{P}^3_k$ be the Zariski closure of M_0, and let $V' = \mathbb{P}^3_k \setminus M_0'$. Then $[X] \in \mathrm{Br}(V)$ is contained in $\mathrm{Br}(V') \subset \mathrm{Br}(V)$.*

Proof. Let $F'(r,s,t,u)$ be a homogeneous cubic polynomial such that $F(s,t,u) = F'(1,s,t,u)$. The rational function $\rho = r/t$ is a uniformiser of the local ring of $\mathbb{P}^3_k$ at the divisor at infinity $r = 0$. We have

$$[X_K] = \left(\frac{t^2-4r^2}{r^2}, \frac{F'(r,s,t,u)}{r^3}\right) = \left(\left(1-4\rho^2\right)\rho^{-2}, \frac{F'(r,s,t,u)}{t^3}\rho^{-3}\right).$$

Thus $[X_K] = (1-4\rho^2, F'(r,s,t,u)t^{-3}\rho)$, hence the residue of $[X_K]$ at $r = 0$ is trivial. By the purity theorem for the Brauer group [CTS21, Theorem 3.7.1 (ii)], this implies the statement. □

It would be interesting to find a conceptual explanation for the fact that the Brauer class $[X]$ is unramified at infinity.

3 Word equations

3.1 Main theorem

Let $w(x,y)$ be a non-trivial word in two letters, i.e. a non-trivial element of the free group $\mathcal{F}_2$ with two generators. We write $w\colon \mathrm{SL}(2)\times\mathrm{SL}(2)\to\mathrm{SL}(2)$ for the PGL(2)-equivariant map defined by $w(x,y)$. By a general theorem of Borel [Bor83, Theorem B], the morphism w is dominant.

Using the Cayley–Hamilton theorem (3) we obtain a polynomial $P_w\in\mathbb{Z}[s,t,u]$ such that $\mathrm{tr}(w(A,B))=P_w(s,t,u)$, where $s=\mathrm{tr}(A)$, $t=\mathrm{tr}(B)$, $u=\mathrm{tr}(AB)$. In other words, the following diagram commutes:

$$\begin{array}{ccc} \mathrm{SL}(2)\times\mathrm{SL}(2) & \xrightarrow{w} & \mathrm{SL}(2) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathrm{tr}} \\ \mathbb{A}^3_k & \xrightarrow{P_w} & \mathbb{A}^1_k \end{array} \tag{7}$$

The composition $\mathrm{tr}\circ w$ is dominant, so the polynomial $P_w(s,t,u)$ is non-constant. For example, if $w(x,y)=xyx^{-1}y^{-1}$, then $P_w(s,t,u)=F(s,t,u)+2$. Moreover, we have the following lemma.

Lemma 3.1 *If $w\in[\mathcal{F}_2,\mathcal{F}_2]$, then there is a polynomial $Q_w(s,t,u)\in k[s,t,u]$ such that we have*

$$P_w(s,t,u)=Q_w(s,t,u)F(s,t,u)+2. \tag{8}$$

Proof. Take any point $P=(s,t,u)\in M_0(\bar{k})$. Let $\lambda,\mu\in\bar{k}$ be such that $\lambda^2-s\lambda+1=0$ and $\mu^2-t\mu+1=0$. The Fricke identity (4) implies that $f^{-1}(P)$ contains

$$A=\begin{pmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{pmatrix},\quad B=\begin{pmatrix}\mu & 0\\ 0 & \mu^{-1}\end{pmatrix}.$$

Since $w(x,y)\in[\mathcal{F}_2,\mathcal{F}_2]$ and $AB=BA$, we have $w(AB)=I$ so that $P_w(s,t,u)=\mathrm{tr}(I)=2$. Thus, the restriction of $P_w(s,t,u)-2$ to the Cayley cubic M_0 is zero. Since $F(s,t,u)$ is irreducible, we obtain (8). $\square$

Let $\alpha\in\mathrm{SL}(2)(k)$ be a non-central element. Let $S_{w,\alpha}\subset\mathrm{SL}(2)\times\mathrm{SL}(2)$ be the closed subset given by $w(X_1,X_2)=\alpha$. The isomorphism class of $S_{w,\alpha}$ depends only on the similarity class of α, and so it is determined by $\mathrm{tr}(\alpha)$.

For $a\in k$ let $H_{w,a}\subset\mathbb{A}^3_k$ be the affine surface given by $P_w(s,t,u)=a$. Thus we have a natural morphism $f\colon S_{w,\alpha}\to H_{w,a}$ with $a=\mathrm{tr}(\alpha)$. As was pointed out in [BZ16, Proposition 2.2], the variety $S_{w,\alpha}$ is non-empty when α is non-central and semisimple. Indeed, P_w is non-constant, hence surjective on $\bar{k}$-points. Thus the composition $P_w\circ f$ is surjective. Hence $S_{w,\alpha'}$ is non-empty for some $\alpha'\in\mathrm{SL}(2)$ with $\mathrm{tr}(\alpha')=\mathrm{tr}(\alpha)$. If α is semisimple and non-central, then α' and α are conjugate in $\mathrm{SL}(2)(\bar{k})$, so that $S_{w,\alpha}$ is non-empty.

Theorem 3.2 *Let k be a field of characteristic zero. Let $w(x, y)$ be a non-trivial word in two letters. Let $\alpha \in \mathrm{SL}(2)(k)$ be a non-central semisimple element with trace $a = \mathrm{tr}(\alpha)$. Let $X_{w,a} \to H_{w,a} \cap V$ be the restriction of the smooth conic bundle $X \to V$ to $H_{w,a} \cap V$. Then we have the following statements.*

(i) *The open subset $S_{w,\alpha} \cap f^{-1}(V) \subset S_{w,\alpha}$ is isomorphic to a dense open subset of $X_{w,a}$ as a variety over $H_{w,a} \cap V$. The complement to $S_{w,\alpha} \cap f^{-1}(V)$ in $X_{w,a}$ consists of two sections of $X_{w,a} \to H_{w,a} \cap V$.*

(ii) *Assume that $w(x, y)$ is contained in the commutator subgroup $[\mathcal{F}_2, \mathcal{F}_2] \subset \mathcal{F}_2$. Then $H_{w,a} \subset V$ so that $S_{w,\alpha}$ is isomorphic to a dense open subset of $X_{w,a}$ as a variety over $H_{w,a} \cap V$. The complement $X_{w,a} \setminus S_{w,\alpha}$ is the union of two sections of $X_{w,a} \to H_{w,a}$. In particular, we have $\dim(S_{w,\alpha}) = 3$. If, moreover, the polynomial $P_w(s, t, u) - a$ is geometrically irreducible, then $S_{w,a}$ is geometrically integral.*

Proof. (i) The centraliser of α in PGL(2) is the norm 1 torus $T = R^1_{k'/k}(\mathbb{G}_{m,k'})$, where $k' = k(\sqrt{a^2 - 4})$.

Write $H'_{w,a} = H_{w,a} \cap V$ and let $\mathcal{T}_{w,a} = f^{-1}(H'_{w,a})$ be the restriction to $H'_{w,a}$ of the V-torsor $f \colon \mathcal{T} \to V$ for PGL(2) defined in Lemma 2.2. We claim that the structure group of $\mathcal{T}_{w,a} \to H'_{w,a}$ reduces to T.

Since α is semisimple, all elements of $\mathrm{SL}(2)(\bar{k})$ of trace a are conjugate to α. Thus the PGL(2)-equivariant morphism $w \colon \mathrm{SL}(2) \times \mathrm{SL}(2) \to \mathrm{SL}(2)$ maps $\mathcal{T}_{w,a}$ to the PGL(2)-orbit of α in SL(2), where PGL(2) acts by conjugation. There is a unique PGL(2)-equivariant isomorphism of this orbit and $T\backslash\mathrm{PGL}(2)$ that sends α to the trivial coset x_0 of T. We obtain a PGL(2)-equivariant morphism

$$\mathcal{T}_{w,a} \to (T\backslash\mathrm{PGL}(2)) \times_k H'_{w,a}$$

of schemes over $H'_{w,a}$, so we can apply Proposition 1.1 (using that T is commutative). The desired $H'_{w,a}$-torsor for T lifting $\mathcal{T}_{w,a}$ is the preimage of $x_0 \times_k H'_{w,a}$ in $\mathcal{T}_{w,a}$, but this is exactly $S_{w,\alpha} \cap f^{-1}(V)$. Now (i) follows from Lemma 1.2.

(ii) Since $a \neq 2$, from Lemma 3.1 we obtain that $M_0 \cap H_{w,a} = \emptyset$, so that $H_{w,a} \subset V$ and $H'_{w,a} = H_{w,a}$, hence $S_{w,\alpha} \cap f^{-1}(V) = S_{w,\alpha}$. Since $H_{w,a} \subset \mathbb{A}^3_k$ is the zero set of a non-constant polynomial, every irreducible component of $H_{w,a}$ has dimension 2. Thus $\dim(S_{w,\alpha}) = \dim(X_{w,a}) = 3$.

When $P_w(s, t, u) - a$ is geometrically irreducible, the surface $H_{w,a}$ is geometrically integral. Since $X_{w,a}$ is a smooth conic bundle over $H_{w,a}$, we see that $X_{w,a}$ is geometrically integral, so $S_{w,\alpha}$ is geometrically integral too. □

To illustrate the practical aspect of Theorem 3.2 (ii) we state the following

Corollary 3.3 *Let k be a field of characteristic zero. Let $w(x, y)$ be a non-trivial word in two letters contained in the commutator subgroup $[\mathcal{F}_2, \mathcal{F}_2] \subset \mathcal{F}_2$. Suppose that $(s, t, u) \in V(k)$ is such that $t^2 - 4$ is a non-zero square in k. Then for any $\alpha \in \mathrm{SL}(2)(k)$ with $\mathrm{tr}(\alpha) = P_w(s, t, u)$ the equation $w(x, y) = \alpha$ has a solution $(A, B) \in \mathrm{SL}(2)(k) \times \mathrm{SL}(2)(k)$ such that $\mathrm{tr}(A) = s$, $\mathrm{tr}(B) = t$, $\mathrm{tr}(AB) = u$.*

Proof. By Corollary 2.5, the fibre of $X \to V$ over the k-point (s,t,u) is isomorphic to $\mathbb{P}^1_k$. Now Theorem 3.2 (ii) implies that the fibre of $f\colon S_{w,\alpha} \to \mathbb{A}^3_k$ over (s,t,u) is isomorphic to $\mathbb{P}^1_k$ with two k-points removed, so it has a k-point. $\square$

3.2 Brauer groups of projective Markoff surfaces

For $d \neq 0$, $d \neq -4$ the Markoff surface M_d is a dense open subset of the smooth cubic surface $M'_d \subset \mathbb{P}^3_k$ given by

$$r(s^2+t^2+u^2) - stu - (d+4)r^3 = 0.$$

The Brauer group of M'_d was computed in [CTWX20, Proposition 3.2] and [LM21, Lemmas 3.1, 3.2]. Let us recall this computation. Projection from the line $r=t=0$ contained in M'_d to $\mathbb{P}^1_k$ with coordinates $(r:t)$ is a conic bundle $\phi\colon M'_d \to \mathbb{P}^1_k$ with five geometric singular fibres. The k-fibre $\phi^{-1}(\infty)$ at the point at infinity $\infty = (0:1)$ is the union of two k-lines given by $r=s=0$ and $r=u=0$. The remaining singular fibres are above the k-points $t/r = \pm 2$ and above $(t/r)^2 = d+4$, the latter being either a degree 2 closed point or a union of two k-points. For these fibres, the quadratic extension over which the components are defined is given by adjoining $\sqrt{d}$. The element

$$\mathcal{A}_0 := ((t/r)^2 - 4, d) = (1 - 4(r/t)^2, d) \in \mathrm{Br}(k(\mathbb{P}^1_k))$$

has residue d at the points $t/r = \pm 2$ and is unramified at all other points of $\mathbb{P}^1_k$. We also note that the value of $\mathcal{A}_0$ at the point at infinity $(r:t) = (0:1)$ is trivial. Let $\mathcal{A}$ be the image of $\mathcal{A}_0$ in $\mathrm{Br}(k(M'_d))$. We deduce that $\mathcal{A}$ is unramified on M'_d and hence $\mathcal{A} \in \mathrm{Br}(M'_d)$. Moreover, the value of $\mathcal{A}$ at any k-point of $\phi^{-1}(\infty)$ is trivial.

For 'general' d, namely, when $[k(\sqrt{d}, \sqrt{d+4}) : k] = 4$, we have $\mathrm{Br}(M'_d)/\mathrm{Br}(k) \cong \mathbb{Z}/2$ with generator $\mathcal{A}$. If $d+4$ is a square in k, but d is not, then $\mathrm{Br}(M'_d)/\mathrm{Br}(k) \cong (\mathbb{Z}/2)^2$ is generated by $\mathcal{A}$ and another element. Finally, if d or $d(d+4)$ is a square in k, then $\mathrm{Br}(M'_d)/\mathrm{Br}(k) = 0$.

3.3 Commutator word variety

In this section $w(x,y)$ is the commutator word $xyx^{-1}y^{-1}$. For $\alpha \in \mathrm{SL}(2)(k)$ we denote the variety $S_{w,\alpha} \subset \mathrm{SL}(2) \times \mathrm{SL}(2)$ defined by $w(X_1, X_2) = \alpha$ simply by S_α.

Theorem 3.4 *Let k be a field of characteristic zero. Let $\alpha \in \mathrm{SL}(2)(k)$ be a non-central semisimple element, and let $d = \mathrm{tr}(\alpha) - 2$. Let $X_d \to M_d$ be the restriction of the smooth conic bundle $X \to V$ to $M_d \subset V$, $d \neq 0$, $d \neq -4$. Then we have the following statements.*

(i) *The class $[X_d] \in \mathrm{Br}(M_d)$ equals $\mathcal{A} \in \mathrm{Br}(M'_d) \subset \mathrm{Br}(M_d)$.*

(ii) *S_α is isomorphic to a dense open subset of X_d.*

(iii) *S_α is geometrically integral.*

(iv) *S_α is k-unirational. In particular, S_α has a Zariski dense set of k-points.*

(v) *S_α is k-rational if and only if the Markoff surface M_d is k-rational. This is the case if and only if d or $d(d+4)$ is a square in k.*

Proof. (i) The Markoff surface M_d is given by $F(s,t,u)=d$, so for $d\neq 0$ we have $M_d\subset V$. Corollary 2.5 implies that $[X_d]=(t^2-4,d)=\mathcal{A}$.

(ii) For $w(x,y)=xyx^{-1}y^{-1}$ we have $P_w(s,t,u)=F(s,t,u)+2$. The trace surface $H_{w,a}$, where $a=\mathrm{tr}(\alpha)=d+2$, is the zero set of $P_w(s,t,u)-a=F(s,t,u)-d$, hence $H_{w,a}=M_d$. Now (ii) follows from Theorem 3.2 (ii).

(iii) This follows from Theorem 3.2 (ii) as M_d is geometrically integral for any d.

(iv) To prove that S_α is k-unirational, by (ii) it is enough to prove that X_d is k-unirational. We start by recalling the well-known k-unirationality of M'_d. For this it is enough to construct a rational curve $C\subset M'_d$ which is a double section of the conic bundle $\phi\colon M'_d\to\mathbb{P}^1_k$, because the pullback of $\phi\colon M'_d\to\mathbb{P}^1_k$ to C is a conic bundle $D\to C$ with a section, hence D is a k-rational variety dominating M'_d. We can take C to be the line $r=t=0$. Indeed, the $\bar{k}$-fibres of ϕ are residual conics to this line, so C is a double section of ϕ.

The k-point $x_0=(0:0:0:1)$ is contained in $C\cap\phi^{-1}(\infty)$. As we have seen above, this implies $\mathcal{A}(x_0)=0$. Since x_0 is in $C(k)$, it lifts to a k-point y_0 on D. Thus we have a dominant morphism $g\colon D\to M'_d$ of generic degree 2, where D is a smooth, projective, and k-rational surface with a k-point y_0 such that $g(y_0)=x_0$. The k-rationality of D implies that the natural map $\mathrm{Br}(k)\to\mathrm{Br}(D)$ is an isomorphism. Thus $g^*\mathcal{A}\in\mathrm{Br}(k)$. But $(g^*\mathcal{A})(y_0)=\mathcal{A}(x_0)=0$, hence $g^*\mathcal{A}=0$. By (i) this implies that the pullback of the conic bundle $X_d\to M_d$ to $g^{-1}(M_d)\subset D$ has a section, hence it is birationally equivalent to $D\times_k\mathbb{P}^1_k$, and therefore is k-rational.

(v) If d or $d(d+4)$ is a square in k, then M'_d is k-rational. (In the first case M'_d contains two skew lines defined over k and in the second case M'_d contains two skew lines conjugate over k and individually defined over $k(\sqrt{d+4})$, see [CTWX20, Remark 3.3].) This implies $\mathrm{Br}(M'_d)/\mathrm{Br}(k)=0$. Since $\mathcal{A}$ vanishes at a k-point of M'_d, we have $\mathcal{A}=0$. Thus X_d is birationally equivalent to $M_d\times_k\mathbb{P}^1_k$, and hence is a k-rational variety. (If $d(d+4)$ is a square in k, then the k-rationality of X_d is proved in [RBKC96, Lemma 4] by elementary but seemingly involved calculations.)

Finally, assume that neither d nor $d(d+4)$ is a square in k. Define $K=k(\sqrt{d+4})$ if $d+4$ is not a square in k, otherwise let $K=k$. We note that $d+4$ is a square in K, but d is not, and this implies that $\mathrm{Br}(M'_{d,K})/\mathrm{Br}(K)\simeq(\mathbb{Z}/2)^2$ is generated by the images of $\mathcal{A}$ and some other element $\mathcal{A}_1\in\mathrm{Br}(M'_{d,K})$. By Lichtenbaum's theorem, the kernel of the restriction map $\mathrm{Br}(K(M_d))\to\mathrm{Br}(K(X_d))$ is generated by $\mathcal{A}$, see [CTS21, Proposition 7.1.3]. Thus the image of $\mathcal{A}_1$ in $\mathrm{Br}(K(X_d))$ is unramified over K, and is non-zero modulo $\mathrm{Br}(K)$. The unramified Brauer group is a birational invariant of smooth, proper, regular varieties [CTS21, Corollary 6.2.11], hence $X_{d,K}$ is not K-rational. Thus X_d is not k-rational. Since $\mathrm{Br}(M'_d)/\mathrm{Br}(k)\neq 0$, the Markoff surface M_d is not k-rational too. □

The existence of a k-point in S_α is a particular case of a general result of R.C.

Thompson [Tho61, Theorem 2]. When k is a number field, our method can be used to prove the following local-to-global statement for rational points on S_α. Recall that a variety over a field k is *split* if it contains an open geometrically integral k-subscheme, see [CTS21, Definition 10.1.3].

Proposition 3.5 *Let k be a number field. Let $\alpha \in \mathrm{SL}(2)(k)$ be a non-central semisimple element. The Brauer–Manin obstruction is the only obstruction to weak approximation on any smooth and proper variety birationally equivalent to S_α. Moreover, if we exclude the case when $\mathrm{tr}(\alpha)+2$ is a square in k but $\mathrm{tr}(\alpha)-2$ is not, then S_α satisfies weak approximation.*

Proof. Using Hironaka's theorem, we can find a smooth, proper, geometrically integral variety X'_d over k that contains X_d as an open subset. Moreover, we can choose X'_d such that there is a morphism $\varphi\colon X'_d \to M'_d$ which extends $X_d \to M_d$. The irreducibility of X'_d implies that the fibres of $\varphi\colon X'_d \to M'_d$ above points of codimension 1 are curves. Thus the restriction of φ to $\mathrm{Spec}(R)$, where R is the local ring of a codimension 1 point of X'_d, is proper and flat. By [CTS21, Lemma 10.2.1], using that $\mathcal{A} \in \mathrm{Br}(M'_d) \subset \mathrm{Br}(R)$, we obtain that the generic fibre of $X_d \to M_d$ extends to a smooth conic bundle over $\mathrm{Spec}(R)$, so its closed fibre is split. By [CTS21, Proposition 10.1.12], this implies that the closed fibre of the restriction of φ to $\mathrm{Spec}(R)$ is split. We conclude that all fibres of $\varphi\colon X'_d \to M'_d$ above codimension 1 points of M'_d are split.

Let Ω be the set of places of k and let k_v be the completion of k at $v \in \Omega$. Since $\mathcal{A} \in \mathrm{Br}(M'_d)$ and M'_d is proper, there exists a finite subset $S_0 \subset \Omega$ such that $\mathcal{A}(P_v) = 0 \in \mathrm{Br}(k_v)$ for every $v \notin S_0$ and every $P_v \in M'_d(k_v)$, see [CTS21, Proposition 13.3.1 (iii)].

Let S be a finite set of places of k and let $P_v \in X'_d(k_v)$, $v \in S$, be local points coming from an adelic point $(P_v)_{v\in\Omega}$ in the Brauer–Manin set $X'_d(\mathbf{A}_k)^{\mathrm{Br}}$. We want to approximate $(P_v)_{v\in S}$ by a k-point. Without loss of generality we can assume that $S_0 \subset S$. After a small deformation we can arrange that $P_v \in X_d$, for $v \in S$.

By functoriality of the Brauer–Manin set we have $(\varphi(P_v))_{v\in\Omega} \in M'_d(\mathbf{A}_k)^{\mathrm{Br}}$. Since M'_d is birationally equivalent to a smooth projective surface that is a conic bundle over $\mathbb{P}^1_k$ with four degenerate $\bar{k}$-fibres, by a theorem of Salberger and Colliot-Thélène [CT90] the Brauer–Manin obstruction is the only obstruction to weak approximation on M'_d. Thus there is a point $Q \in M'_d(k)$ arbitrarily close to each $\varphi(P_v)$ in the local topology of k_v, for $v \in S$. Since $\varphi(P_v) \in M_d(k_v)$ for $v \in S$, we can ensure that $Q \in M_d(k)$. Then the fibre $\varphi^{-1}(Q)$ is a smooth projective conic with class $\mathcal{A}(Q) \in \mathrm{Br}(k)$. Since $\varphi(P_v)$ is close to Q, this conic has a k_v-point for all $v \in S$. Since $S_0 \subset S$, it also has k_v-points for all $v \notin S$. By the Minkowski–Hasse theorem, $\varphi^{-1}(Q)$ has a k-point. Then $\varphi^{-1}(Q) \simeq \mathbb{P}^1_k$, so we can find a k-point on $\varphi^{-1}(Q)$ arbitrarily close to P_v for $v \in S$.

To prove the last statement note that the conditions guarantee that $\mathrm{Br}(M'_d)/\mathrm{Br}(k)$ is generated by the image of $\mathcal{A}$. Since the generic fibre of φ is a conic, the Brauer group $\mathrm{Br}(X'_d)$ is vertical. As the fibres of $\varphi\colon X'_d \to M'_d$ over the codimension 1 points of M_d' are split, we have $\mathrm{Br}(X_d') = \mathrm{Br}(k)$. Thus there is no Brauer–Manin obstruction on X_d' in this case. □

3.4 Complements

Part (i) of the following proposition can be compared to a theorem of R.C. Thompson that $S_{-I}(k) \neq \emptyset$ if and only if -1 is a sum of two squares in k, see [Tho61, Theorem 1]. Part (ii), together with Theorem 3.4 (iv), gives a proof in the case of $\mathrm{char}(k) = 0$ of [Tho61, Theorem 2] which says that if k has more than three elements, every non-central element of $\mathrm{SL}(2)(k)$ is a commutator [Tho61, Theorem 2].

Proposition 3.6 *Let k be a field of characteristic zero.*

(i) *The variety S_{-I} is k-rational of dimension 3 if -1 is a sum of two squares in k, otherwise $S_{-I} = \emptyset$ so that S_{-I} is not k-rational.*

(ii) *If $\alpha \in \mathrm{SL}(2)(k)$ is a non-central element such that $\mathrm{tr}(\alpha) = \pm 2$, then S_α is k-rational of dimension 3.*

Proof. (i) If $A, B \in \mathrm{SL}(2)(\bar{k})$ are such that $ABA^{-1}B^{-1} = -I$, then $\mathrm{tr}(B) = \mathrm{tr}(-B)$, hence $\mathrm{tr}(B) = 0$. Similarly, we obtain $\mathrm{tr}(A) = 0$. The Fricke identity (4) implies that $\mathrm{tr}(AB) = 0$. Thus $f(S_{-I})$ is a subset of $\{(0,0,0)\}$. By Lemma 2.2, $f^{-1}\big((0,0,0)\big)$ is a k-torsor for PGL(2). In particular, it is not empty with transitive action of PGL(2), thus $f^{-1}\big((0,0,0)\big) = S_{-I}$.

By Corollary 2.5 the class of this torsor in the Brauer group $\mathrm{Br}(k)$ is $(-4,-4) = (-1,-1)$. By a basic property of quaternion algebras, $(-1,-1) = 0$ if and only if -1 is a norm of the quadratic extension $k(\sqrt{-1})$, that is, if and only if -1 a sum of two squares in k, cf. [CTS21, Proposition 1.1.8].

In the k-torsor S_{-I} is trivial, it is isomorphic to PGL(2), in particular, it is k-rational. In the opposite case, it has no k-points, and hence is not k-rational.

(ii) The case $\mathrm{tr}(\alpha) = 2$, $\alpha \neq I$, can be dealt with by an explicit computation. We can find a basis in which

$$\alpha = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Suppose that $A, B \in \mathrm{SL}(2)(k)$ are (2×2)-matrices such that $ABA^{-1}B^{-1} = \alpha$. We have $\mathrm{tr}(B) = \mathrm{tr}(ABA^{-1}) = \mathrm{tr}(\alpha B)$, and this implies that B is upper-triangular. Likewise, we obtain that A is also upper-triangular. Writing

$$A = \begin{pmatrix} \lambda & x \\ 0 & \lambda^{-1} \end{pmatrix}, \quad B = \begin{pmatrix} \mu & y \\ 0 & \mu^{-1} \end{pmatrix},$$

we obtain that the matrix equation $ABA^{-1}B^{-1} = \alpha$ is equivalent to

$$x\lambda(1-\mu^2) - y\mu(1-\lambda^2) = 1.$$

We conclude that S_α is birationally equivalent to $\mathbb{A}^3_k$.

If $\mathrm{tr}(\alpha) = -2$, $\alpha \neq -I$, then our method still works. Let us explain how to adjust the proof of Theorem 3.2 to this case. We have $a = \mathrm{tr}(\alpha) = d+2$ which gives $d = -4$, so $H_{w,a} = M_{-4} \subset V$. Let $M_\sharp := M_{-4} \setminus \{(0,0,0)\}$. Let $\mathcal{T}_\sharp \to M_\sharp$ be the restriction of $\mathcal{T} \to V$ and let $X_\sharp \to M_\sharp$ be the restriction of $X \to V$.

The commutator gives a PGL(2)-equivariant map $w\colon \mathrm{SL}(2) \times \mathrm{SL}(2) \to \mathrm{SL}(2)$ sending $\mathcal{T}_\sharp$ to the subvariety of SL(2) whose $\bar{k}$-points are matrices of trace -2. All such matrices, except $-I$, are conjugate to α. The preimage $w^{-1}(-I) = S_{-I}$ is $f^{-1}\big((0,0,0)\big)$, by the proof of (i). Thus $w(\mathcal{T}_\sharp)$ is contained in the PGL(2)-orbit of α in SL(2), where PGL(2) acts by conjugation. Let G be the centraliser of α in PGL(2). Thus $G \cong \mathbb{G}_{a,k}$, in particular, G is commutative. There is a unique PGL(2)-equivariant isomorphism of this orbit of α and $G\backslash\mathrm{PGL}(2)$ that sends α to the trivial coset x_0 of G. We obtain a PGL(2)-equivariant morphism

$$\mathcal{T}_\sharp \to (G\backslash\mathrm{PGL}(2)) \times_k M_\sharp$$

of schemes over $M_\sharp$, so we can apply Proposition 1.1 (using that G is commutative). The desired $M_\sharp$-torsor for G lifting $\mathcal{T}_\sharp$ is the preimage of $x_0 \times_k M_\sharp$ in $\mathcal{T}_\sharp$, but this is exactly S_α because $f(S_\alpha)$ does not contain $(0,0,0)$. We know that S_α is a dense open subset of $X_\sharp$ by Remark 1.4. Moreover, by the same remark, the generic fibre of $X_{-4} \to M_{-4}$ is isomorphic to the projective line over the function field of M_{-4}, so S_α is birationally equivalent to $M_{-4} \times_k \mathbb{P}^1_k$. Finally, M_{-4} is k-rational as a cubic surface with a double k-point. Thus S_α is birationally equivalent to $\mathbb{A}^3_k$. □

When n is prime, for any field k, Larsen and Lu showed that the commutator morphism $w\colon \mathrm{SL}(n) \times_k \mathrm{SL}(n) \to \mathrm{SL}(n)$ is flat over the complement to the identity in $\mathrm{SL}(n)$, see [LL21]. Without using [LL21] we have the following by-product of our method.

Corollary 3.7 *Let k be a field of characteristic zero.*

(i) *Let $w(x,y)$ be a non-trivial word in two letters contained in the commutator subgroup $[\mathcal{F}_2, \mathcal{F}_2] \subset \mathcal{F}_2$. Then the restriction of the morphism*

$$w\colon \mathrm{SL}(2) \times_k \mathrm{SL}(2) \to \mathrm{SL}(2)$$

to the open subset of non-central semisimple elements in SL(2) *is faithfully flat.*

(ii) *The commutator map* $\mathrm{SL}(2) \times_k \mathrm{SL}(2) \to \mathrm{SL}(2)$ *is faithfully flat over the complement to the identity in* SL(2).

Proof. Since the source and the target are smooth, by miracle flatness it is enough to check that all fibres have the same dimension. Thus (i) holds by Theorem 3.2 (ii), whereas (ii) holds by Theorem 3.4 and Proposition 3.6. □

It is well-known that S_I has dimension 4, so the result of Corollary 3.7 (ii) is best possible. For the sake of completeness we note that S_I is k-rational.

Proposition 3.8 *The variety S_I of commuting pairs of elements of* SL(2) *is k-rational of dimension* 4.

Proof. T.S. Motzkin and O. Taussky proved that the variety of commuting pairs of $(n \times n)$-matrices contains the set of commuting pairs of regular semisimple matrices as a dense open subset [MT55, Theorems 5, 6]. This implies that S_I is geometrically integral. Hence S_I is birationally equivalent to the centraliser of the generic point of SL(2). The generic point is a regular semisimple element, so its centraliser is a maximal torus in SL(2) defined over the function field of SL(2). On the one hand, SL(2) is k-rational of dimension 3. On the other hand, any maximal torus of $\mathrm{SL}(2)_K$ defined over a field extension K/k is K-rational of dimension 1. This finishes the proof. □

References

[BZ16] T. Bandman and Yu.G. Zarhin. Surjectivity of certain word maps on $PSL(2, \mathbb{C})$ and $SL(2, \mathbb{C})$. *Eur. J. Math.* **2** (2016) 614–643.

[Bor83] A. Borel. On free subgroups of semisimple groups. *Enseign. Math.* **29** (1983) 151–164; reproduced in Œuvres - Collected Papers, vol. IV, Springer-Verlag, Berlin–Heidelberg, 2001, pp. 41–54.

[CT90] J.-L. Colliot-Thélène. Surfaces rationnelles fibrées en coniques de degré 4. *Sém. théorie des nombres Paris* 1988–89. Progress in Math., Birkhäuser, 1990, p. 43–55.

[CTS21] J.-L. Colliot-Thélène and A.N. Skorobogatov, *The Brauer–Grothendieck Group*, Springer, Cham, 2021.

[CTWX20] J.-L. Colliot-Thélène, D. Wei, and F. Xu, Brauer–Manin obstruction for Markoff surfaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **21** (2020) 1257–1313.

[Fri83] S. Friedland. Simultaneous similarity of matrices. *Adv. Math.* **50** (1983) 189–265.

[Gir71] J. Giraud. *Cohomologie non abélienne.* Springer-Verlag, 1971.

[GMS21] A. Ghosh, C. Meiri, and P. Sarnak. Commutators in SL_2 and Markoff surfaces I. *New Zealand J. Math.* **52** (2021 [2021–2022]) 773–819.

[GS22] A. Ghosh and P. Sarnak. Integral points on Markoff type cubic surfaces. *Invent. Math.* **229** (2022) 689–749.

[GKP18] N.L. Gordeev, B.È. Kunyavskiĭ, and E.B. Plotkin. Geometry of word equations in simple algebraic groups over special fields. *Uspekhi Mat. Nauk* **73** (2018) 3–52; English translation: *Russian Math. Surveys* **73** (2018) 753–796.

[LL21] M. Larsen and Z. Lu. Flatness of the commutator map over SL_n. *Int. Math. Res. Notices* (2021) 5605–5622.

[LM21] D. Loughran and V. Mitankin. Integral Hasse principle and strong approximation for Markoff surfaces. *Int. Math. Res. Notices* (2021), no. 18, 14086–14122.

[MT55] T.S. Motzkin and O. Taussky. Pairs of matrices with property L. II. *Trans. Amer. Math. Soc.* **80** (1955) 387–401.

[GIT94] D. Mumford, J. Fogarty, and F. Kirwan. *Geometric invariant theory.* 3rd ed., Springer-Verlag, 1994.

[RBKC96] A.S. Rapinchuk, V.V. Benyash-Krivetz, and V.I. Chernousov. Representation varieties of the fundamental groups of compact orientable surfaces. *Israel J. Math.* **93** (1996) 29–71.

[She84] D. Shemesh. Common eigenvectors of two matrices. *Linear Algebra Appl.* **62** (1984) 11–18.

[Sko01] A. Skorobogatov. *Torsors and rational points.* Cambridge University Press, 2001.

[Tho61] R.C. Thompson. Commutators in the special and general linear groups. *Trans. Amer. Math. Soc.* **101** (1961) 16–33.

Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, Israel
bandman@macs.biu.ac.il

Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, Israel
kunyav@macs.biu.ac.il

Department of Mathematics, South Kensington Campus, Imperial College London SW7 2AZ, United Kingdom – and – Institute for the Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russia
a.skorobogatov@imperial.ac.uk