C. R. Acad. Sci. Paris, Ser. I 354 (2016) 131–135

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Number theory/Group theory

Markoff triples and strong approximation

Triplets de Markoff et approximation forte

Jean Bourgain [a], Alexander Gamburd [b], Peter Sarnak [a,c]

[a] *IAS, USA*
[b] *The Graduate Center, CUNY, USA*
[c] *Princeton University, USA*

ARTICLE INFO

Article history:
Received 16 November 2015
Accepted 16 November 2015
Available online 26 January 2016

Presented by Jean Bourgain

ABSTRACT

We investigate the transitivity properties of the group of morphisms generated by Vieta involutions on the solutions in congruences to the Markoff equation as well as to other Markoff type affine cubic surfaces. These are dictated by the finite $\bar{\mathbb{Q}}$ orbits of these actions and these can be determined effectively. The results are applied to give forms of strong approximation for integer points, and to sieving, on these surfaces.

RÉSUMÉ

Nous explorons les propriétés de transitivité du groupe des morphismes engendré par les involutions de Vieta agissant sur les solutions en congruences de l'équation de Markoff ainsi que d'autres surfaces affines cubiques de type Markoff. Ces propriétés sont déterminées par les orbites finies dans $\bar{\mathbb{Q}}$ de ces actions, qui peuvent être déterminées explicitement. Les resultats permettent d'établir une forme de l'approximation forte pour les points entiers sur ces surfaces et des applications du crible.

By strong approximation, we mean the extent to which the reduction mod q of the integral points on an affine variety V over $\mathbb{Z}$ covers the points in $V(\mathbb{Z}/q\mathbb{Z})$. In a related direction and setting let $\mathcal{O} = \Gamma \cdot a$ be the orbit in $\mathbb{Z}^n$ of the action of a group Γ of polynomial morphisms of $\mathbb{A}^n$, which preserve $\mathbb{Z}^n$ and let $V = Zc\ell(\mathcal{O})$, the Zariski closure of $\mathcal{O}$. The orbit $\mathcal{O}$ is a subset of $V(\mathbb{Z})$ and strong approximation for $\mathcal{O}$ (and a fortiori $V(\mathbb{Z})$) amounts to determining the orbit of a on the induced (permutation) action of Γ on the (finite) sets $V(\mathbb{Z}/q\mathbb{Z})$. In the case where Γ acts linearly and the Levi factor of $G = Zc\ell(\Gamma)$ is semisimple, this question as well as its applications to sieving theory have been developed in [27,5,32]. We note that, on the other hand, tori pose particularly difficult problems, in terms of sparsity of elements in an orbit, strong approximation and diophantine properties (see [26] for a discussion of Artin's Conjecture in this context).

We investigate these questions in the context of Markoff's affine cubic surface $X \subset \mathbb{A}^3$ given by the equation

$$X : \Phi(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 3x_1x_2x_3 = 0. \tag{1}$$

E-mail addresses: bourgain@math.ias.edu (J. Bourgain), agamburd@gc.cuny.edu (A. Gamburd), sarnak@math.princeton.edu (P. Sarnak).

http://dx.doi.org/10.1016/j.crma.2015.12.006
1631-073X/

Recall that the set $\mathcal{M}$ of Markoff triples ([24,25]) are natural number solutions to (1) and that all of the integer solutions are of the form $(0,0,0)$, $(\varepsilon_1 x_1, \varepsilon_2 x_2, \varepsilon_3 x_3)$ with $\varepsilon_1\varepsilon_2\varepsilon_3 = 1$, $\varepsilon_j = \pm 1$, $(x_1, x_2, x_3) \in \mathcal{M}$. All members of $\mathcal{M}$ are gotten from $a = (1,1,1)$ by repeated applications of permutations of the coordinates and the 'Vieta' involutions R_1, R_2, R_3, with $R_3(x) = (x_1, x_2, 3x_1x_2 - x_3)$ and R_2, R_1 defined similarly. That is, $\mathcal{M} = \Gamma \cdot a$ where Γ is the (nonlinear) group of affine morphisms of $\mathbb{A}^3$ generated by the permutations and the R_j's. The Markoff numbers M are the coordinates of the triples $\mathcal{M}$. The first few elements of M are

$$1, 2, 5, 13, 29, 34, 89, 169, 194, \ldots \tag{2}$$

M^s the Markoff sequence is, the set of largest coordinates of an $x \in \mathcal{M}$ counted with multiplicity and Frobenius Uniqueness conjecture [15] asserts that $M = M^s$. The sequence M^s is very sparse, as shown in [36]:

$$\sum_{\substack{m \in M^s \\ m \le T}} 1 \sim c(\log T)^2 \ \text{ as } \ T \to \infty, \ (c > 0). \tag{3}$$

Markoff triples and numbers arise in many different contexts: see, for example, [2] and [1] and references therein.

The fundamental strong approximation conjecture for X is the following transitivity:

Conjecture 1. *For p a prime, Γ acts on $X(p) := X(\mathbb{Z}/p\mathbb{Z})$ with two orbits: $\{0\}$ and $X^*(p) = X(p)\backslash\{0\}$.*

Remark 1. Numerical experiments indicate that not only are the Cayley graphs of the action of Γ on $X^*(p)$ (with respect to a fixed set of generators of Γ) connected, but that they also form an expander family.

The Conjecture implies that the reduction mod p from $\mathcal{M}$ to $X^*(p)$ is onto. This in turn implies that the only congruence constraints on Markoff numbers $m \bmod p$ are those first noted in [15], namely that $m \neq 0, \pm 2/3 \bmod p$, if $p \equiv 3(4)$ and $p \neq 3$.

Our first result is that $X^*(p)$ has a giant orbit and that no orbit is small.

Theorem 1. *For $\varepsilon > 0$ and p large there is an orbit $\mathcal{C}(p)$ of $X^*(p)$ for which*

$$|X^*(p)\backslash\mathcal{C}(p)| \le p^{\varepsilon} \quad (\text{note } |X^*(p)| \sim p^2),$$

and all Γ orbits $\mathcal{D}(p)$ in $X^*(p)$ satisfy $|\mathcal{D}(p)| \gg (\log p)^{\frac{1}{3}}$.

Let E be the set of primes for which Conjecture 1 fails. This set is very small, basically we can prove the Conjecture unless $p^2 - 1$ is very smooth.

Theorem 2. *For $\varepsilon > 0$ the number of $p \in E$ with $p \le x$ for which the Conjecture fails, is $O_\varepsilon(x^\varepsilon)$.*

There is an extension of Theorem 2 to composite moduli q, at least with suitable restrictions on its prime factors. Applying this together with some sieving (cf. [20], Chapter 7) on $\mathcal{M}$ and M allows us to say some things about the divisors of the sparse Markoff sequence M^s. For example,

Theorem 3. *Almost all Markoff numbers are composite, that is*

$$\sum_{\substack{p \in M^s \\ p \text{ prime}, \ p \le T}} 1 = o\Big(\sum_{\substack{m \in M^s \\ m \le T}} 1\Big).$$

Our methods can be used to prove results similar to Theorems 1 and 2 for more general Markoff type cubic surfaces. Namely $X_k : \Phi(x_1, x_2, x_3) = k$, the family of surfaces $S_{A,B,C,D}$ in [8], those in [14], and even the general such non-degenerate cubic surface

$$Y = Y(\alpha, \beta, \gamma, \delta) : \sum_{i,j=1}^{3} \alpha_{ij} x_i x_j + \sum_{j=1}^{3} \beta_j x_j + \gamma = \delta x_1 x_2 x_3 \tag{4}$$

with α_{ij}, β_j, γ, δ integers.

The group Γ_Y is, again, the one generated by the corresponding Vieta involutions R_1, R_2, R_3. For such a Y and action Γ_Y we show first that there are only finitely many finite orbits in $Y(\bar{\mathbb{Q}})$, and that these may be determined effectively. The analogue of Conjecture 1 for Y is that for p large, Γ_Y has one big orbit on $Y(\mathbb{Z}/p\mathbb{Z})$ and that the remaining orbits, if there are any, correspond to one of the finite $\bar{\mathbb{Q}}$ orbits determined above.

The determination of the finite orbits of Γ on $X_k(\bar{\mathbb{Q}})$ and on $S_{A,B,C,D}(\bar{\mathbb{Q}})$ has been carried out in [13] and [22] respectively. Remarkably, for these the Γ action on an affine 3-space corresponds to the (nonlinear) monodromy group for Painlevé VI equations on their parameter spaces. In this way, the finite orbits in question turn out to correspond bijectively to those Painlevé VI's, which are algebraic functions of their independent variable. Applying this to X_k shows that our version of Conjecture 1 for these is equivalent to the "Q-conjectures" of [28] that concern the transitivity systems for Nielsen moves on pairs of generators[1] of $SL_2(\mathbb{F}_p)$ (at least if p is large).

In this setting of the more general surfaces Y in (4), strong approximation for $Y(\mathbb{Z}_S)$, where S is the set of primes dividing $A_{11}A_{22}A_{33}$ (so that Γ_Y preserves the S-integers $\mathbb{Z}_S$), will follow from Conjecture 1 for Y (and the results we can prove towards it, as in Theorem 2) once we have a point of infinite order in $Y(\mathbb{Z}_S)$. If there is no such point we can increase S or replace $\mathbb{Z}$ by $\mathcal{O}_K$ the ring of integers in a number field $K/\mathbb{Q}$ to produce such a point and with it strong approximation for $Y\big((\mathcal{O}_K)_S\big)$.

Vojta's Conjectures and the results proven towards them ([34,11]) assert that cubic and higher degree affine surfaces typically have few S-integral points. In the rare cases where these points are Zariski dense such as tori (e.g., $N(x_1, x_2, x_3) = k$, where N is the norm form of a cubic extension of $\mathbb{Q}$) strong approximation fails. So these Markoff surfaces appear to be rather special affine cubic surfaces in not only having a Zariski dense set of integral points, but also a robust strong approximation. The story for rational points on projective cubic surfaces is very different from the affine integral one. Once there are points, there are many of them (see [23] for a detailed study).

We give a brief overview of our proof of Theorems 1 and 2 and some comments about their extensions. Theorem 1 in the weaker form that $|\mathcal{C}(p)| \sim |X^*(p)|$ as $p \to \infty$, can be viewed as the finite field analogue of [18], where it is shown that the action of Γ on the compact real components of the character variety of the mapping class group of the once punctured torus is ergodic. As in [18], our proof makes use of the rotations $\tau_{ij} \circ R_j$, $i \neq j$ where τ_{ij} permutes x_i and x_j. These preserve the conic sections gotten by intersecting $X^*(p)$ with the plane $y_k = x_k$ (k different from i and j). If $\tau_{ij} \circ R_j$ has order t_1 $\big($here $t_1 | p(p-1)(p+1)\big)$. Then x and these t_1 points of the conic section are connected (i.e. are in the same Γ orbit). If t_1 is maximal (i.e. is p, $p-1$ or $p+1$), then this entire conic section is connected and such conic sections in different planes that intersect are also connected. This leads to a large component, which we denote by $\mathcal{C}(p)$.

If our starting rotation has order t_1, which is not maximal, then the idea's to ensure that among the t_1 points to which it is connected, at least one has a corresponding rotation of order $t_2 > t_1$, and then to repeat. To ensure that one can progress in this way a critical equation over $\mathbb{F}_p$ intervenes:

$$\left.\begin{aligned} &x + \frac{b}{x} = y + \frac{1}{y}, b \neq 1 \\ &\text{with } x \in H_1,\ y \in H_2 \text{ with } H_1, H_2 \text{ subgroups of } \mathbb{F}_p^* \text{ (or } \mathbb{F}_{p^2}^*). \end{aligned}\right\} \tag{5}$$

If $t_1 = |H_1| \geq p^{1/2+\delta}$ (with δ small and fixed), one can apply the proven RH (Riemann Hypothesis) for curves over finite fields [35] to count the number of solutions to (5). Together with a simple inclusion/exclusion argument this shows that one of the t_1 points connected to our starting x has a corresponding maximal rotation and hence x is connected to $\mathcal{C}(p)$.

If $|H_1| \leq p^{1/2+\delta}$ then RH for these curves is of little use (their genus is too large) and we have to proceed using other methods. We assume that $|H_1| \geq |H_2|$ so that the trivial upper bound for the number of solutions to (5) is $2|H_2|$. What we need is a power saving in this upper bound in the case where $|H_2|$ is close to $|H_1|$; that is a bound of the form $C_\tau |H_1|^\tau$, with $\tau < 1$, $C_\tau < \infty$ (both fixed). We know of three methods to achieve this. The first is combinatorial and while it is special to the equation (5) and it produces poor exponents τ, it is otherwise robust and in fact we use it specifically in the composite cases q needed for Theorem 3. It uses the expansion theory (cf. [16]) in $SL_2(\mathbb{F}_p)$ ([4]) as well as the "projective Szemeredi–Trotter Theorem" proved in [3] for pairs of points in $\mathbb{P}^1(\mathbb{F}_p)$, which are incident by a subset of $PGL_2(\mathbb{F}_p)$.

The second and third methods are related to "elementary" proofs of *RH* for curves. One can use auxiliary polynomials as in Stepanov's [33] proof of *RH* for curves to give the desired power saving with an explicit τ (cf. [19] who deal with $x + y = 1$ and $|H_1| = |H_2|$ in (5)). The third method gives the best upper bound, namely

$$20 \max\left\{(|H_1|.|H_2|)^{1/3}, \frac{|H_1|.|H_2|}{p}\right\}$$

and is due to Corvaja and Zannier [12]. It uses their method for estimating the g.c.d. of $u - 1$ and $v - 1$ in terms of the degrees of u and v and their supports, as well as (hyper) Wronskians. As they show, their technique is also robust and can be used to give an elementary proof of *RH* for curves.

The above lead to a proof of part 1 of Theorem 1. To continue one needs to deal with t_1, which is very small (here $|H_1| = t_1$ which divides $p^2 - 1$).

[1] The connectedness [17] and expansion ([16,7]) for T-systems of $SL_2(\mathbb{F}_p)$ on 4 or more generators are known.

To handle these, we lift to characteristic zero and examine the finite orbits of Γ in $X(\bar{\mathbb{Q}})$. In fact, by the Chebotarev Density Theorem, a necessary condition for Conjecture 1 to hold is that there are no such orbits other than $\{0\}$. Again using the rotations in the conic sections by planes one finds that any such finite orbit must be among the solutions to

$$(t_1+t_1^{-1})^2+(t_2+t_2^{-1})^2+(t_3+t_3^{-1})^2=(t_1+t_1^{-1})(t_2+t_2^{-1})(t_3+t_3^{-1})$$
$$\text{with } t_j\text{'s roots of unity.} \tag{6}$$

For this particular surface X one can show using the inequality between the geometric and arithmetic means, that (6) has no nontrivial solutions for complex numbers with $|t_j|=1$ (pointed out to us by Bombieri). For the more general surfaces X_k, $S_{A,B,C,D}$ and those in (4), there is a variety of solutions with $|t_j|=1$. However, Lang's G_m Conjecture, which is established effectively (see [21,31]), yields that there are only finitely many solutions to these equations in roots of unity. This allows for an explicit determination of the finite orbits of Γ_Y in $Y(\bar{\mathbb{Q}})$ (as noted earlier for the cubic surfaces $S_{A,B,C,D}$, the long list of these orbits [22] correspond to the algebraic Painlevé VI solutions). This $\bar{\mathbb{Q}}$ analysis leads to part 2 of Theorem 1 and, combined with the discussion above, it yields a proof of Conjecture 1, at least if p^2-1 is not very smooth. To prove Theorem 1, we need to show that there are very few primes for which the above arguments fail. This is done by extending the arguments and results in [9] and [10] concerning points (x, y) on irreducible curves over $\mathbb{F}_p$ for which $\operatorname{ord}(x)+\operatorname{ord}(y)$ is small (here $\operatorname{ord}(x)$ is the order of x in $\mathbb{F}_p^*$).

Our methods fall short of dealing with all p, specifically for those rare p's for which p^2-1 is very smooth. The following hypothesis, which is a strong variant of the conjectures of M.C. Chang and B. Poonen [9], would suffice to deal with all large p's.

Hypothesis. Given $d\in\mathbb{N}$ there is $\delta>0$ and $K=K(\delta,d)$ such that for p large and $f(x,y)$ absolutely irreducible over $\mathbb{F}_p$ and of degree d and $f(x,y)=0$ is not a translate of a subtorus of $(\bar{\mathbb{F}}_p^*)^2$, the set of $(x,y)\in(\mathbb{F}_p^*)^2$ for which $f(x,y)=0$ and $\max(\operatorname{ord}x,\operatorname{ord}y)\le p^{\delta}$, is at most K.

For the extension of Theorem 2 to composite moduli, we take $q=p_1p_2\dots p_\nu$ with $p_\ell\equiv 1(4)$ and for which Theorem 2 holds, and make use of the special conic sections $x_j=2(\bmod p_\ell)$, which consists of two lines. This allows us to bypass the difficulties connected with maximal orders of elements in $(\mathbb{Z}/q\mathbb{Z})^*$ and Charmichael numbers. The proof of Theorem 3 also necessitates extending Zagier's result (4) to counting such m's subject to a congruence $\bmod q$ (cf. [6]), which is accomplished using the methods in [29] or [30].

Acknowledgements

It is a pleasure to thank E. Bombieri, S. Cantat, M.C. Chang, P. Corvaja, W. Goldman, E. Hrushovski, Yu.I. Manin, M. Mirzakhani, B. Poonen, I. Shparlinski, U. Zannier for insightful discussions and stimulating correspondence.

While working on this paper, the authors were supported, in part, by the following NSF DMS awards: 1301619 (Bourgain), 064507 (Gamburd), 1302952 (Sarnak).

References

[1] Martin Aigner, Markov's Theorem and 100 Years of the Uniqueness Conjecture, Springer, 2013.
[2] E. Bombieri, Continued fractions and the Markoff tree, Expo. Math. 25 (3) (2007) 187–213.
[3] J. Bourgain, A modular Szemeredi–Trotter theorem for hyperbolas, C. R. Acad. Sci. Paris, Ser. I 350 (2012) 793–796.
[4] J. Bourgain, A. Gamburd, Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$, Ann. Math. 167 (2008) 625–642.
[5] J. Bourgain, A. Gamburd, P. Sarnak, Affine linear sieve, expanders and sum product, Invent. Math. 179 (2010) 559–644.
[6] Jean Bourgain, Alex Gamburd, Peter Sarnak, Generalization of Selberg's $\frac{3}{16}$ theorem and affine sieve, Acta Math. 207 (2) (2011) 255–290.
[7] E. Breuillard, A. Gamburd, Strong uniform expansion in $SL_2(\mathbb{F}_p)$, Geom. Funct. Anal. 20 (5) (2010) 1201–1209.
[8] S. Cantat, F. Loray, Dynamics on character varieties and Malgrange irreducibility of Painlevé VI equation, Ann. Inst. Fourier (Grenoble) 59 (2009) 2927–2978.
[9] Chang Mei-Chu, Elements of large order in prime finite fields, Bull. Aust. Math. Soc. 88 (2013) 169–176.
[10] M.-C. Chang, B. Kerr, I. Shparlinski, U. Zannier, Elements of large orders on varieties over prime finite fields, J. Théor. Nr. Bordx. 26 (2014) 579–594.
[11] P. Corvaja, U. Zannier, On integral points on surfaces, Ann. Math. (2) 160 (2004) 705–726.
[12] P. Corvaja, U. Zannier, Greatest common divisors of $u-1$, $v-1$ in positive characteristic and rational points on curves over finite fields, J. Eur. Math. Soc. 15 (2013) 1927–1942.
[13] B. Dubrovin, M. Mazzocco, Monodromy of certain Painlevé-VI transcendents and reflection groups, Invent. Math. 141 (2000) 55–147.
[14] M.H. El-Huti, Cubic surfaces of Markov type, Math. USSR Sb. 22 (3) (1974) 333–348.
[15] G. Frobenius, Über die Markoffschen Zahlen, Preuss. Akad. Wiss. Sitzungsbericht (1913) 458–487.
[16] A. Gamburd, I. Pak, Expansion of product replacement graphs, Combinatorica 26 (4) (2006) 411–429.
[17] R. Gilman, Finite quotients of the automorphism group of a free group, Can. J. Math. 29 (1977) 541–551.
[18] W. Goldman, The modular group action on real $SL(2)$-characters of a one-holed torus, Geom. Topol. 7 (2003) 443–486.
[19] R. Heath-Brown, S. Konyagin, New bounds for Gauss sums derived from k-th powers and for Heilbronn's exponential sum, Q. J. Math. (2000) 221–235.
[20] C. Hooley, Applications of Sieve Methods to the Theory of Numbers, Cambridge Tracts in Mathematics, vol. 70, 1976.
[21] M. Laurent, Exponential Diophantine equations, C. R. Acad. Sci. Paris, Ser. I 296 (1983) 945–947.
[22] O. Lisovyy, Y. Tykhyy, Algebraic solutions of the sixth Painlevé equation, J. Geom. Phys. 85 (2014) 124–163.
[23] Yu.I. Manin, Cubic Forms, 1974.

[24] A. Markoff, Sur les formes quadratiques binaires indéfinies, Math. Ann. 15 (1879) 381–409.
[25] A. Markoff, Sur les formes quadratiques binaires indéfinies, Math. Ann. 17 (1880) 379–399.
[26] C.R. Matthews, Counting points modulo p for some finitely generated subgroups of algebraic groups, Bull. Lond. Math. Soc. (3) 14 (1982) 149–154.
[27] C. Matthews, L. Vaserstein, B. Weisfeiler, Congruence properties of Zariski dense groups, Proc. Lond. Math. Soc. (3) 48 (1984) 514–532.
[28] D. McCullough, M. Wanderley, Nielsen equivalence of generating pairs in $SL(2, q)$, Glasg. Math. J. 55 (2013) 481–509.
[29] Greg McShane, Igor Rivin, Simple curves on hyperbolic tori, C. R. Acad. Sci. Paris, Ser. I 320 (1995) 1523–1528.
[30] M. Mirzakhani, Counting mapping class group orbits on hyperbolic surfaces, arXiv:1601.03342v1, 2015.
[31] P. Sarnak, S. Adams, Betti numbers of congruence groups, with an appendix by Z. Rudnick Isr. J. Math. 88 (1994) 31–72.
[32] P. Sarnak, A. Saleh-Golsefidy, The affine sieve, J. Amer. Math. Soc. 26 (4) (2013) 1085–1105.
[33] S.A. Stepanov, The number of points of a hyperelliptic curve over a prime field, Math. USSR, Izv. 3 (5) (1969) 1103–1114.
[34] P. Vojta, A generalization of theorems of faltings and Thue–Siegel–Roth–Wirsing, J. Amer. Math. Soc. 25 (1992) 763–804.
[35] A. Weil, On the Riemann hypothesis in function fields, Proc. Natl. Acad. Sci. USA 27 (1941) 345–347.
[36] D. Zagier, On the number of Markoff numbers below a given bound, Math. Comput. 39 (160) (1982) 709–723.