A HEURISTIC SUBEXPONENTIAL ALGORITHM TO FIND PATHS IN MARKOFF GRAPHS OVER FINITE FIELDS

JOSEPH H. SILVERMAN

Abstract. Charles, Goren, and Lauter [J. Cryptology 22(1), 2009] explained how one can construct hash functions using expander graphs in which it is hard to find paths between specified vertices. The set of solutions to the classical Markoff equation $X^2 + Y^2 + Z^2 = 3XYZ$ in a finite field $\mathbb{F}_q$ has a natural structure as a tri-partite graph using three non-commuting polynomial automorphisms to connect the points. These graphs conjecturally form an expander family, and Fuchs, Lauter, Litman, and Tran [Mathematical Cryptology 1(1), 2022] suggested using this family of Markoff graphs in the CGL construction. In this note we show that in both a theoretical and a practical sense, assuming two randomness hypotheses, one can compute paths in a Markoff graph over $\mathbb{F}_q$ by factoring $q-1$ and solving three discrete logarithm problems in $\mathbb{F}_q^*$. In particular, the path problem can be solved in subexponential time.

Contents

1. Introduction

The classical Markoff surface is the affine surface given by the equation

$$\mathcal{M} : X^2 + Y^2 + Z^2 = 3XYZ. \tag{1}$$

Date: June 19, 2024.
2010 *Mathematics Subject Classification.* Primary: 11T71; Secondary: 94A60, 05C48.
Key words and phrases. Cryptographic hash function, Markoff equation.
Silverman's research supported by Simons Collaboration Grant #712332.

Submitted to Algor. Num. Th. Symp.

There are three double covers $\mathcal{M} \to \mathbb{A}^2$ that give rise to three non-commuting involutions $\sigma_1, \sigma_2, \sigma_3$. A famous theorem of Markoff [17] says that every positive integer solution of (1) can be obtained from $(1,1,1)$ by repeatedly applying the σ_i and permuting the coordinates.

In this note we consider solutions to (1) in a finite field $\mathbb{F}_q$ of characteristic at least 5. There has recently been a lot of interest in studying the orbit structure of $\mathcal{M}(\mathbb{F}_q)$ [2, 4, 5, 8, 15, 6, 9]. Baragar [1] conjectured that the reduction modulo q map

$$\mathcal{M}(\mathbb{Z}) \longrightarrow \mathcal{M}(\mathbb{F}_q)$$

is surjective for all primes q. A recent deep result of William Chen [8], building on ground-breaking work of Bourgain, Gamburd, and Sarnak [4, 5], says that Baragar's conjecture is true for all sufficiently large primes.[1]

More precisely, consider the three non-commuting automorphisms

$$\rho_1, \rho_2, \rho_3 : \mathcal{M} \longrightarrow \mathcal{M}$$

given by the formulas

$$\rho_1 = (X, Z, 3XZ - Y), \quad \rho_2 = (3XY - Z, Y, X), \quad \rho_3 = (Y, 3YZ - X, Z),$$

where ρ_1, ρ_2, ρ_3 are obtained from $\sigma_1, \sigma_2, \sigma_3$ by composing with appropriate coordinate permutations. (The ρ_i are called "rotations" in [4, 5].) Let

$$\mathcal{R} = \langle \rho_1, \rho_2, \rho_3 \rangle \subset \mathrm{Aut}(\mathcal{M})$$

be the group of automorphisms generated by the ρ_i, and let

$$\mathcal{M}^*(\mathbb{F}_q) = \mathcal{M}(\mathbb{F}_q) \smallsetminus \big\{(0,0,0)\big\}.$$

Then Chen's theorem says that $\mathcal{M}^*(\mathbb{F}_q)$ consists of a single $\mathcal{R}$-orbit for sufficiently large prime values of q.

We consider the undirected function graph (sometimes called a Schreier graph) associated to the action of $\{\rho_1, \rho_2, \rho_3\}$ on the set $\mathcal{M}^*(\mathbb{F}_q)$, i.e., we form an undirected graph $\overline{\mathcal{M}}(\mathbb{F}_q)$ whose vertices and edges are given by

$$\begin{aligned} \mathsf{Vertices}\Big(\overline{\mathcal{M}}(\mathbb{F}_q)\Big) &= \mathcal{M}^*(\mathbb{F}_q), \\ \mathsf{Edges}\Big(\overline{\mathcal{M}}(\mathbb{F}_q)\Big) &= \Big\{ [P, \rho_i(P)] : P \in \mathcal{M}^*(\mathbb{F}_q),\ i = 1,2,3 \Big\}. \end{aligned}$$

It is conjectured in [4, 5] that $\overline{\mathcal{M}}(\mathbb{F}_q)$ is a family of expander graphs; see also [9].

Charles, Goren, and Lauter [7] have explained how one can build cryptographic hash functions from expander graphs provided that it is hard to find paths in the graph connecting two given vertices. This led Fuchs, Lauter, Litman, and Tran [10] to suggest using the Markoff graph $\overline{\mathcal{M}}(\mathbb{F}_q)$ to construct a hash function. They prove, using the connectivity ideas from [4, 5], that there is a path-finding algorithm for $\overline{\mathcal{M}}(\mathbb{F}_q)$ that runs in deterministic time $O(q \log\log q)$, and they speculate that any path-finding algorithm in $\overline{\mathcal{M}}(\mathbb{F}_q)$ must take time at least $O(q)$. This leads them to suggest that "these graphs may be good candidates" for the CGL hash function construction.

Our goal in this note is to show that under some reasonable heuristic assumptions, it is possible to solve the path-finding problem in $\overline{\mathcal{M}}(\mathbb{F}_q)$ in subexponential

[1] In an updated version of [4] (private communication), the authors note that the algorithm in the present article "should allow one to check Baragar's conjecture for much larger q. Whether it is feasible to bridge the gap and verify the conjecture for all primes is an interesting question."

(in $\log q$) time on a classical computer and in polynomial (in $\log q$) time on a quantum computer. More precisely, up to small polynomial-time tasks, it suffices to factor $q-1$ and solve three discrete logarithm problems in $\mathbb{F}_q^*$, as described in the following theorem, whose proof will be given in Section 8. We note that the path-finding algorithm is a randomized algorithm, so the time complexity upper bound in Theorem 1.1 is average case, not worst case.

Theorem 1.1 (Markoff Path-Finding Algorithm)**.** *We set the following notation*:

$$\begin{aligned}
\mathsf{PATH}(q) &= \textit{time to find a path between points in } \overline{\mathcal{M}}(\mathbb{F}_q).\\
\mathsf{DLP}(q) &= \textit{time to solve the DLP in } \mathbb{F}_q^*.\\
\mathsf{FACTOR}(N) &= \textit{time to factor } N.\\
\mathsf{NEGLIGIBLE}(q) &= \textit{tasks that take negligible (polylog) time as a function of } q, \textit{ for}\\
&\qquad \textit{for example taking square roots in } \mathbb{F}_q, \textit{ or iterations performed a}\\
&\qquad \textit{small multiple of } (q-1)/\varphi(q-1) \textit{ times; see Remark } 1.5.
\end{aligned}$$

Assume that Heuristic 7.1 *is valid. Then with high probability,*

$$\mathsf{PATH}(q) \le \mathsf{FACTOR}(q-1) + 3\cdot\mathsf{DLP}(q) + \mathsf{NEGLIGIBLE}(q).$$

Remark 1.2. It is well known that there are practical subexponential-time algorithms for factoring (sieve methods) and discrete logarithm problems (index calculus) on a classical computer; see for example [14, Chapter 3]. And there are polynomial-time algorithms (Shor's algorithm) for both problems on a quantum computer; see [19, 20].

Remark 1.3. We note that the paths constructed by our Markoff path-finder algorithm (Algorithm 1) have the following two properties that are unlikely to be present in the paths generated by the CGL graph-theoretic hash function [7]:

- They are quite long, in the sense that the number of ρ_i used to connect P to Q is almost certainly larger than, say, $q^{1/2}$.
- The path connecting P to Q has long stretches in which it repeatedly applies one of the ρ_i, e.g., it is almost certainly true that somewhere in the path there is a ρ_k that is repeated at least $q^{1/2}$ times without using the other two ρ_i.

Thus one might still consider using the Markoff graph for a CGL hash function with the proviso that long or repetitive paths are disallowed. On the other hand, the fact that one can create paths and collisions, even of a disallowed type, may cause some disquiet, as well as making it more difficult to construct a security reduction proof.

Remark 1.4. In this article we have restricted attention to the classical Markoff equation (1), but we note that the method works, mutatis mutandis, for more general Markoff–Hurwitz type equations

$$a_1X^2 + a_2Y^2 + a_3Z^2 + b_1XY + b_2XZ + b_3YZ \\ + c_1X + c_2Y + c_2Z + dXYZ + e = 0 \quad (2)$$

that admit three non-commuting involutions.

Remark 1.5. There will be a number of estimates that depend on $\varphi(q-1)/(q-1)$ or its reciprocol, where we note that this quantity is the probabily that an element

of the cyclic group $\mathbb{F}_q^*$ is a generator. For most N, the ratio is $N/\varphi(N)$ is fairly small, but it can become arbitrarily large when N is very smooth. However, this can only happen if N is huge. More explicitly, there are classical estimates that can be used to prove that

$$\frac{N}{\varphi(N)} \le 2\log\log N \quad \text{for all } N \ge 5;$$

see for example [12, Sections 18.4 and 22.9] or [18, Theorem 15]. So for q of large cryptographic size, say $q < 2^{10000}$, if some task requires $10(q-1)/\varphi(q-1)$ steps, then in a worse than worst case scenario, it takes fewer than 200 steps.

We give an initial high-level description of the Markoff path-finder algorithm in Section 2 using psudo-code (Table 1) and a picture (Table 2). A more detailed description that includes the path-finder algorithm (Algorithm 1) and its subroutines is given in Appendix C. The proof that the Markoff path-finder algorithm finds a path and has the indicated running time is given in Section 8. A key observation in constructing the path-finder algorithm, as already exploited in [4, 5], is to note that the action of the ρ_i on appropriate fibers of $\mathcal{M} \to \mathbb{A}^1$ is described via repeated application of a linear transformation in SL_2. (In fancier terminology, the fibers are $\mathbb{G}_m$-torsors.) This means that if we are given two points on a fiber, then finding a power of ρ_i that links the given points can be rephrased as a discrete logarithm problem, either in $\mathbb{F}_q^*$ or in the subgroup of norm 1 elements of $\mathbb{F}_{q^2}$.

We briefly describe the heuristic assumption required by our algorithm. Let $\mathcal{H}(\mathbb{F}_q) \subset \overline{\mathcal{M}}(\mathbb{F}_q)$ be the set of points whose y-coordinate is *maximally hyperbolic*, by which we mean that the polynomial $T^2 - 3yT + 1$ has a root in $\mathbb{F}_q$ that generates the cyclic group $\mathbb{F}_q^*$, i.e., such that λ is a primitive root for $\mathbb{F}_q$. The set $\mathcal{H}(\mathbb{F}_q)$ is quite large, roughly $\varphi(q-1)/2(q-1)$ as large as $\overline{\mathcal{M}}(\mathbb{F}_q)$, so in particular (Remark 1.5) $\#\mathcal{H}(\mathbb{F}_q) \ge \frac{1}{4\log\log(q-1)}\#\overline{\mathcal{M}}(\mathbb{F}_q)$. Our heuristic assumption says that starting from a point in $P_0 \in \overline{\mathcal{M}}(\mathbb{F}_q)$ and randomly applying ρ_1 or ρ_3 with equal probability, it will not take very many iterations before we land in $\mathcal{H}(\mathbb{F}_q)$.

We illustrate the Markoff path-finding algorithm in Section 9 by executing it on a numerical example with $q = 70687$. We find paths between some randomly chosen points in $\overline{\mathcal{M}}(\mathbb{F}_q)$, and a non-trivial loop from a point back to itself. Finally, in Section 10 we briefly discuss a family of K3 surfaces that is analogous to the Markoff surface (1) and its generalizations (2) and explain how path-finding on the associated graphs can be heuristically reduced to the elliptic curve discrete logarithm problem (ECDLP).

Acknowledgements. The author would like to thank Elena Fuchs, Igor Shparlinski, and the referees for their helpful comments.

2. A High-Level Description of the Markoff Path-Finding Algorithm

For the convenience of the reader, Table 1 gives an informal description of our heuristically subexponential algorithm for finding paths in $\overline{\mathcal{M}}(\mathbb{F}_q)$. In this algorithm, we say that an element $t \in \mathbb{F}_q^*$ is *maximally hyperbolic* if the quadratic polynomial $T^2 - 3tT + 1$ has a root $\lambda \in \mathbb{F}_q$ that is a primitive root, i.e., such that λ is a generator for $\mathbb{F}_q^*$; cf. Definition 3.4. The algorithm is also illustrated by the

picture in Table 2. We refer the reader to Appendix C for a more detailed description of the Markoff path-finding algorithm, and to Section 8 for a proof that the Markoff path-finding algorithm operates successfully in subexponential time.

- **Input**: $\mathbb{F}_q$ a finite field of characteristic at least 5
  P, Q points in $\overline{\mathcal{M}}(\mathbb{F}_q)$
- Use ρ_1 and ρ_3 to randomly move P in $\overline{\mathcal{M}}(\mathbb{F}_q)$ until reaching a point P' satisfying $y(P')$ is maximally hyperbolic. This gives $i_1, \ldots, i_\alpha \in \{1, 3\}$ such that
$$P' = \rho_{i_\alpha} \circ \cdots \circ \rho_{i_1}(P).$$
- Use ρ_1^{-1} and ρ_2^{-1} to randomly move Q in $\overline{\mathcal{M}}(\mathbb{F}_q)$ until reaching a point Q' satisfying $z(Q')$ is maximally hyperbolic. This gives $j_1, \ldots, j_\beta \in \{1, 2\}$ such that
$$Q = \rho_{j_1} \circ \cdots \circ \rho_{j_\beta}(Q').$$
- Let $F(X, Y, Z) = X^2 + Y^2 + Z^2 - 3XYZ$. Randomly select maximally hyperbolic $x_0 \in \mathbb{F}_q^*$ until finding a value for which the quadratic equations
$$F\big(x_0, y(P'), Z\big) = F\big(x_0, Y, z(Q')\big) = 0$$
have a solution $(y_0, z_0) \in \mathbb{F}_q^2$. Set
$$P'' \leftarrow \big(x_0, y(P'), z_0\big) \quad \text{and} \quad Q'' \leftarrow \big(x_0, y_0, z(Q')\big).$$
We note that:
  - P'' and Q'' are on the same maximally hyperbolic x-fiber,
  - P' and P'' are on the same maximally hyperbolic y-fiber,
  - Q' and Q'' are on the same maximally hyperbolic z-fiber.
- Find a, b, c satisfying
$$P'' = \rho_2^a(P'), \quad Q' = \rho_3^b(Q''), \quad Q'' = \rho_1^c(P'').$$
As explained in Proposition 5.1, this involves solving three DLPs in $\mathbb{F}_q^*$.
- **Output**: The list of integers $(i_1, \ldots, i_\alpha)$, $(j_1, \ldots, j_\beta)$, (a, b, c) specifies the path
$$Q = \rho_{j_1} \circ \cdots \circ \rho_{j_\beta} \circ \rho_3^b \circ \rho_1^c \circ \rho_2^a \circ \rho_{i_\alpha} \circ \cdots \circ \rho_{i_1}(P).$$

TABLE 1. High-level description of the Markoff path-finding algorithm

3. Rotations on a Fiber and an Associated Matrix

The map ρ_1 may be written in matrix form as

$$\rho_1(x, y, z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3x & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{3}$$

Thus computing $\rho_1^n(x, y, z)$ amounts to taking the nth power of the matrix $\left(\begin{smallmatrix} 3x & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Similar considerations apply to ρ_2 and ρ_3. This prompts the following definitions.

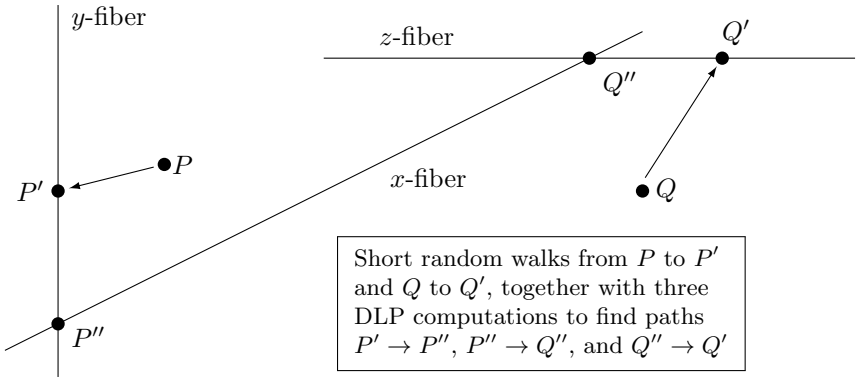

TABLE 2. Illustrating the Markoff path-finding algorithm

Definition 3.1. For $t \in \mathbb{F}_q^*$, we set the following notation:

$$L_t = \begin{pmatrix} 3t & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{F}_q), \qquad \lambda_t, \lambda_t^{-1} = \text{the eigenvalues of } L_t.$$

We note that $\lambda_t \in \mathbb{F}_{q^2}^*$, and that λ_t is in $\mathbb{F}_q^*$ if and only if $9t^2 - 4$ is a square in $\mathbb{F}_q^*$.

Formula (3) tells us that if we apply iterates of ρ_1 to a point $(x, y, z) \in \overline{\mathcal{M}}(\mathbb{F}_q)$, then

$$\rho_1^n(x, y, z) = (x, y_n, z_n) \quad \text{with} \quad \begin{pmatrix} y_n \\ z_n \end{pmatrix} = L_x^n \begin{pmatrix} y \\ z \end{pmatrix}, \tag{4}$$

and similarly for ρ_2 and ρ_3. This often allows us to find paths in fibers of $\overline{\mathcal{M}}(\mathbb{F}_q)$ by solving a DLP in $\mathbb{F}_q^*$.

Definition 3.2. In [4, 5, 10], the various L_t are separated into three cases, analogous to the classification of elements of $\mathrm{SL}_2(\mathbb{R})$. We say that $t \in \mathbb{F}_q^*$ is

$$\begin{aligned} \textit{hyperbolic}&: \text{ if } \lambda_t \in \mathbb{F}_q \smallsetminus \{\pm 1\}, \\ \textit{parabolic}&: \text{ if } \lambda_t = \pm 1, \\ \textit{elliptic}&: \text{ if } \lambda_t \in \mathbb{F}_{q^2}^* \smallsetminus \mathbb{F}_q^*. \end{aligned}$$

Remark 3.3. The characteristic polynomial of L_t is $T^2 - 3tT + 1$, whose discriminant is $9t^2 - 4$, so we see that

$$\begin{aligned} L_t \text{ is hyperbolic} &\iff 9t^2 - 4 \in \mathbb{F}_q^{*2}, \\ L_t \text{ is parabolic} &\iff 9t^2 - 4 = 0, \\ L_t \text{ is elliptic} &\iff 9t^2 - 4 \notin \mathbb{F}_q^{*2}. \end{aligned}$$

Definition 3.4. We say that $t \in \mathbb{F}_q^*$ is *maximally hyperbolic* if any one of the following equivalent conditions is true:

- L_t is hyperbolic and has order $q - 1$ in $\mathrm{SL}_2(\mathbb{F}_q)$.
- An eigenvalue λ_t of L_t is in $\mathbb{F}_q$ and generates the multiplicative group $\mathbb{F}_q^*$.
- There is a generator $\lambda \in \mathbb{F}_q^*$ such that $t = \frac{1}{3}(\lambda + \lambda^{-1})$.

4. Some Counting Results for $\mathbb{F}_q$-Rational Points

In this section we give various elementary counting result for the number of $\mathbb{F}_q$-rational points in $\overline{\mathcal{M}}$, on fibers of $\overline{\mathcal{M}}$, and on certain curves associated to $\overline{\mathcal{M}}$, in some cases with a maximal hyperbolicity requirement on one of the coordinates.

Proposition 4.1. *For $x_0 \in \mathbb{F}_q$, we denote the fiber by*

$$\overline{\mathcal{M}}_{x_0}(\mathbb{F}_q) = \big\{(x,y,z) \in \overline{\mathcal{M}}(\mathbb{F}_q) : x = x_0\big\}.$$

(a) *Let $x_0 \in \mathbb{F}_q$. Then*

$$\#\overline{\mathcal{M}}_{x_0}(\mathbb{F}_q) = \begin{cases} q-1 & \text{if } x_0 \neq 0 \text{ is hyperbolic,} \\ q+1 & \text{if } x_0 \neq 0 \text{ is elliptic,} \\ (q-1)\left[1+\left(\frac{-1}{\mathbb{F}_q}\right)\right] & \text{if } x_0 = 0, \\ 2q\left[1+\left(\frac{-1}{\mathbb{F}_q}\right)\right] & \text{if } x_0 \text{ is parabolic.} \end{cases}$$

(b) *Let $\mathbb{F}_q^{\text{elliptic}} = \{t \in \mathbb{F}_q : t \text{ is elliptic}\}$, and similarly for $\mathbb{F}_q^{\text{hyperbolic}}$ and $\mathbb{F}_q^{\text{parabolic}}$. Then*

$$\#\mathbb{F}_q^{\text{elliptic}} = \frac{q-1}{2}, \quad \#\mathbb{F}_q^{\text{hyperbolic}} = \frac{q-3}{2}, \quad \#\mathbb{F}_q^{\text{parabolic}} = 2.$$

(c) *We have*

$$\#\overline{\mathcal{M}}(\mathbb{F}_q) = q\left(1+3\left(\frac{-1}{\mathbb{F}_q}\right)\right),$$

where the count does not include the singular point $(0,0,0)$.

(d) *We have*

$$\#\{t \in \mathbb{F}_q^* : t \text{ is maximally hyperbolic}\} = \frac{\varphi(q-1)}{2}. \tag{5}$$

(e) *We denote the set of points in $\overline{\mathcal{M}}(\mathbb{F}_q)$ whose y-coordinate is non-zero and maximally hyperbolic by*

$$\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}} := \big\{(x,y,z) \in \overline{\mathcal{M}}(\mathbb{F}_q) : y \neq 0 \text{ and } y \text{ is maximally hyperbolic}\big\}.$$

Then

$$\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}} = \frac{(q-1)\varphi(q-1)}{2}.$$

Proof. These results appear in various places, see for example [4, Lemmas 3, 4, 5], but for the convenience of the reader, we include a proof. For notational convenience, for the proof of Proposition 4.1, we let

$$Q = \left(\frac{-1}{\mathbb{F}_q}\right) = (-1)^{(q-1)/2}.$$

(a) If $x_0 \in \mathbb{F}_q^*$ is hyperbolic or elliptic, i.e., if $9x_0^2 \neq 4$, then

$$C_{x_0} : U^2 - 3x_0UV + V^2 = -x_0^2W^2$$

is a non-singular conic in $\mathbb{P}^2$. Therefore $C_{x_0} \cong_{/\mathbb{F}_q} \mathbb{P}^1$, and hence

$$\#C_{x_0}(\mathbb{F}_q) = \#\mathbb{P}^1(\mathbb{F}_q) = q+1.$$

The two points at ∞, i.e., the points with $W=0$, are defined over $\mathbb{F}_q$ if and only if $9x_0^2-4$ is a square in $\mathbb{F}_q^*$, i.e., if and only if x_0 is hyperbolic. Thus

$$\#\Big(C_{x_0}\smallsetminus\{W=0\}\Big)(\mathbb{F}_q)=\begin{cases}\#\mathbb{P}^1(\mathbb{F}_q)=q+1 & \text{if } x_0\ne0 \text{ is elliptic,}\\ \#\mathbb{P}^1(\mathbb{F}_q)-2=q-1 & \text{if } x_0\ne0 \text{ is hyperbolic.}\end{cases}$$

If $x_0=0$, then $\overline{\mathcal{M}}_{x_0}$ reduces to $y^2+z^2=0$, so

$$\#\overline{\mathcal{M}}_0(\mathbb{F}_q)=(q-1)\,[1+Q]\,.$$

(We are not counting the singular point $(0,0,0)$.) Finally, x_0 is parabolic if and only if $9x_0^2=4$, in which case the equation for C_{x_0} is $(U-3x_0V/2)^2=-x_0^2W^2$. So C_{x_0} is the union of two lines, and

$$\#C_{x_0}(\mathbb{F}_q)=\begin{cases}2q+1 & \text{if } -1 \text{ is a square in } \mathbb{F}_q,\\ 1 & \text{otherwise.}\end{cases}$$

Since there is 1 point with $W=0$, the fiber has either $2q$ points or is empty.
(b) The parabolic elements of $\mathbb{F}_q$ are those t such that $t=\pm2/3$, so there are 2 parabolic elements. The set of hyperbolic elements of $\mathbb{F}_q$ is the image of the map

$$f:\mathbb{F}_q^*\smallsetminus\{\pm1\}\longrightarrow\mathbb{F}_q,\quad \lambda\longrightarrow\frac{1}{3}(\lambda+\lambda^{-1}). \tag{6}$$

The map (6) is exactly 2-to-1 onto its image, since $f(\lambda')=f(\lambda)$ if and only if $\lambda'=\lambda^{\pm1}$, so we find that

$$\#\mathbb{F}_q^{\text{hyperbolic}}=\frac{q-3}{2}.$$

Finally, we have

$$\#\mathbb{F}_q^{\text{elliptic}}=\#\mathbb{F}_q-\#\mathbb{F}_q^{\text{hyperbolic}}-\#\mathbb{F}_q^{\text{parabolic}}=q-\frac{q-3}{2}-2=\frac{q-1}{2}.$$

(c) We use (a) and (b) to compute

$$\begin{aligned}\#\overline{\mathcal{M}}(\mathbb{F}_q)&=\#\overline{\mathcal{M}}_0(\mathbb{F}_q)+\sum_{\substack{x_0\in\mathbb{F}_q^*\\ x_0\text{ hyperbolic}}}\#\overline{\mathcal{M}}_{x_0}(\mathbb{F}_q)+\sum_{\substack{x_0\in\mathbb{F}_q^*\\ x_0\text{ elliptic}}}\#\overline{\mathcal{M}}_{x_0}(\mathbb{F}_q)+\sum_{\substack{x_0\in\mathbb{F}_q^*\\ x_0\text{ parabolic}}}\#\overline{\mathcal{M}}_{x_0}(\mathbb{F}_q)\\
&=(q-1)(1+Q)+\left(\#\mathbb{F}_q^{\text{hyperbolic}}-\frac{1+Q}{2}\right)\cdot(q-1)\\
&\qquad+\left(\#\mathbb{F}_q^{\text{elliptic}}-\frac{1-Q}{2}\right)\cdot(q+1)\\
&\qquad+\#\mathbb{F}_q^{\text{parabolic}}\cdot q(1+Q)\quad\begin{array}{l}\text{where the } \frac{1\pm Q}{2} \text{ compensates}\\ \text{for when 0 is hyperbolic or elliptic,}\end{array}\\
&=(q-1)(1+Q)+\frac{q-4-Q}{2}\cdot(q-1)+\frac{q-2+Q}{2}\cdot(q+1)+2q(1+Q)\\
&=q(q+3Q).\end{aligned}$$

(Note that we are not including $(0,0,0)$.)
(d) Let $\mathsf{Gen}(q)\subset\mathbb{F}_q^*$ denote the set of generators of $\mathbb{F}_q^*$, and consider the map

$$f:\mathbb{F}_q^*\longrightarrow\mathbb{F}_q,\quad \lambda\longrightarrow\frac{1}{3}(\lambda+\lambda^{-1}).$$

We want to count $\#f\big(\mathsf{Gen}(q)\big)$. The map f is exactly 2-to-1 onto its image, since $f(\lambda)=f(\lambda^{-1})$, except for the two points $f(\pm1)=\pm\frac{2}{3}$ that have only one

pre-image. The set of generators of $\mathbb{F}_q^*$ is invariant under inversion and does not contain ± 1, so

$$\#f\big(\mathsf{Gen}(q)\big) = \frac{1}{2}\#\,\mathsf{Gen}(q).$$

The group $\mathbb{F}_q^*$ is cyclic of order $q-1$, so $\#\,\mathsf{Gen}(q) = \phi(q-1)$, which gives (5).
(e) We compute

$$\begin{aligned}
\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}} &= \sum_{\substack{y_0\in\mathbb{F}_q^*\\ y_0 \text{ maximally hyperbolic}}} \#\big\{(x,y_0,z)\in\overline{\mathcal{M}}(\mathbb{F}_q)\big\} \\
&= \sum_{\substack{y_0\in\mathbb{F}_q^*\\ y_0 \text{ maximally hyperbolic}}} (q-1) \quad \text{from (a)}, \\
&= (q-1)\cdot\#\{y_0\in\mathbb{F}_q^* : y_0 \text{ maximally hyperbolic}\} \\
&= \frac{(q-1)\varphi(q-1)}{2} \quad \text{from (d)}.
\end{aligned}$$

□

Theorem 4.2. *Let*

$$F(X,Y,Z) = X^2+Y^2+Z^2-3XYZ,$$

let

$$a,b\in\mathbb{F}_q \quad \textit{satisfy} \quad ab(a^2-4)(b^2-4)(a-b)\neq 0,$$

and let $\mathcal{C}_{a,b}\subset\mathbb{A}^3$ *be the affine curve*

$$\mathcal{C}_{a,b} := \big\{(X,Y,Z) : F\big(X,a,Z\big) = F\big(X,Y,b\big) = 0\big\}.$$

Then

$$\begin{aligned}
&\operatorname*{Prob}_{P\in\mathcal{C}_{a,b}(\mathbb{F}_q)}\Big(x(P) \textit{ is maximally hyperbolic}\Big) \\
&\qquad := \frac{\#\big\{P=(x,y,z)\in\mathcal{C}_{a,b}(\mathbb{F}_q) : x \textit{ is maximally hyperbolic}\big\}}{\#\mathcal{C}_{a,b}(\mathbb{F}_q)} \\
&\qquad \ge \frac{\varphi(q-1)}{2q} + O_\epsilon\left(q^{-\frac{1}{2}+\epsilon}\right) \qquad (7)\\
&\qquad \ge \frac{1}{4\log\log q} + O_\epsilon\left(q^{-\frac{1}{2}+\epsilon}\right). \qquad (8)
\end{aligned}$$

Proof. We want to apply Proposition A.2 to the family of curves $\mathcal{C}_{a,b}$ parameterized by $(a,b)\in\mathbb{A}^2(\mathbb{F}_q)$. In order to apply the proposition, we need to know that the curves

$$\mathcal{C}'_{a,b}[n] := \big\{(P,\mu)\in\mathcal{C}_{a,b}\times\mathbb{P}^1 : \mu^{2n}-3x(P)\mu^n+1=0\big\}$$

are irreducible for all $n\mid q-1$. We note that these curves appear implicitly in [4] during the "Endgame", but their irreducibility is not addressed. However, an updated version of [4] now includes irreducibility proofs of the relevant curves via some intricate algebraic calculations and via a monodromy argument.[2] Proposition A.2 then gives

$$\#\big\{P\in\mathcal{C}_{a,b}(\mathbb{F}_q) : x(P) \text{ is maximally hyperbolic}\big\} \ge \frac{1}{2}\varphi(q-1) + O\left(q^{-\frac{1}{2}+\epsilon}\right),$$

[2]The author has an alternative proof that is more algebro-geometric in nature using properties of fiber products.

This combined with Weil's estimate

$$\#\mathcal{C}_{a,b}(\mathbb{F}_q) = q + O(q^{1/2})$$

gives (7), and then the discussion in Remark 1.5 gives (8). □

5. Finding Paths in Fibers by Solving the DLP

We now prove the key result that if $x \in \mathbb{F}_q$ is maximally hyperbolic, then ρ_1 acts transitively on the x-fiber of $\overline{\mathcal{M}}$, and that we can explicitly find ρ_1-paths in the x-fiber by solving a DLP in $\mathbb{F}_q^*$.

Proposition 5.1. *Let $x \in \mathbb{F}_q^*$ be maximally hyperbolic, and let*

$$P = (x, y, z) \in \overline{\mathcal{M}}(\mathbb{F}_q) \quad \text{and} \quad P' = (x, y', z') \in \overline{\mathcal{M}}(\mathbb{F}_q) \tag{9}$$

be any two points on the x-fiber of $\overline{\mathcal{M}}$. Then there exists an $n \ge 0$ such that

$$P' = \rho_1^n(P), \tag{10}$$

and we can compute an exponent n satisfying (10) by solving a quadratic equation in $\mathbb{F}_q$ and then solving a DLP in the group $\mathbb{F}_q^*$. (See Algorithm 2 in Table 7 for an explicit algorithm.)

Proof. The assumption that x is maximally hyperbolic means, by definition, that the eigenvalues $\lambda_x, \lambda_x^{-1}$ of the matrix L_x are elements of order $q-1$ in $\mathbb{F}_q^*$. Further, since we always assume that $q > 3$, we know that $\lambda_x \ne \pm 1$. This allows us to diagonalize L_x working over $\mathbb{F}_q$. Explicitly,

$$U = \begin{pmatrix} 1 & -\lambda_x^{-1} \\ -1 & \lambda_x \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_q) \quad \text{satisfies} \quad UL_xU^{-1} = \begin{pmatrix} \lambda_x & 0 \\ 0 & \lambda_x^{-1} \end{pmatrix}. \tag{11}$$

(Note that $\lambda_x \ne \pm 1$ implies that U is invertible, since $\det(U) = \lambda_x - \lambda_x^{-1}$.)

We first prove that ρ_1 acts transitively on the x-fiber of $\overline{\mathcal{M}}$. To do this, we characterize the integers m such that ρ_1^m fixes $P = (x, y, z)$. Thus

$$\begin{aligned}
\rho_1^m(P) = P \quad &\iff \quad \begin{pmatrix} y \\ z \end{pmatrix} = L_x^m \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{from (4)}, \\
&\iff \quad U \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \lambda_x^m & 0 \\ 0 & \lambda_x^{-m} \end{pmatrix} U \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{from (11)}, \\
&\implies \quad \left(\begin{array}{l} 1 \text{ is an eigenvalue of } L_x^m, \text{ since we know that } \left(\begin{smallmatrix} y \\ z \end{smallmatrix}\right) \ne \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \text{ and} \\ \text{that } U \text{ is invertible, so } U \left(\begin{smallmatrix} y \\ z \end{smallmatrix}\right) \ne \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \text{ is an eigenvector} \\ \text{of } \left(\begin{smallmatrix} \lambda_x^m & 0 \\ 0 & \lambda_x^{-m} \end{smallmatrix}\right) \text{ with eigenvalue } 1, \end{array} \right) \\
&\iff \quad \lambda_x^m = 1 \quad \text{since the eigenvalues of } \left(\begin{smallmatrix} \lambda_x^m & 0 \\ 0 & \lambda_x^{-m} \end{smallmatrix}\right) \text{ are } \lambda_x^{\pm m}, \\
&\iff \quad q-1 \mid m \quad \text{since } \lambda_x \text{ has order } q-1 \text{ in } \mathbb{F}_q^*.
\end{aligned}$$

This proves in particular that the $q-1$ points

$$P,\ \rho_1(P),\ \rho_1^2(P),\ \rho_1^3(P),\ \ldots,\ \rho_1^{q-2}(P) \tag{12}$$

are distinct. On the other hand, Propostion 4.1 and the assumption that x is hyperbolic tell us that the x-fiber of $\overline{\mathcal{M}}$ has exactly $q-1$ points with coordinates in $\mathbb{F}_q$, so (12) is the complete list of such point. This completes the proof that ρ_1 acts transitively on the x-fiber.

Now let P and P' be points (9) on the x-fiber of $\overline{\mathcal{M}}$. We have just proven that ρ_1 acts transitively, so we know that there exists an $n \geq 0$ such that $P' = \rho_1^n(P)$, and we want to describe how to compute such an n. The first step is to compute λ_x, which requires solving a quadratic equation. We then repeat our earlier calculation,

$$\begin{aligned} P' = \rho_1^n(P) \quad &\iff \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = L_x^n \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{from (4)}, \\ &\iff \quad U \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} \lambda_x^n & 0 \\ 0 & \lambda_x^{-n} \end{pmatrix} U \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{from (11)}. \end{aligned} \tag{13}$$

We note that the vectors

$$U \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} y' - \lambda_x^{-1} z' \\ -y' + \lambda_x z' \end{pmatrix} \quad \text{and} \quad U \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} y - \lambda_x^{-1} z \\ -y + \lambda_x z \end{pmatrix}$$

are non-zero (since $(y,z) \neq (0,0)$ and $(y',z') \neq (0,0)$ and U is invertible) and that they involve only the known quantities $\lambda_x, y, z, y', z' \in \mathbb{F}_q$. Hence the only unknown quantity in (13) is n. Since the vectors are non-zero, at least one of the coordinates of (13) gives an equation of the form $\alpha\lambda_x^n = \beta$ with known $\alpha, \beta \in \mathbb{F}_q^*$, so we can find n by solving a DLP in $\mathbb{F}_q^*$. □

Remark 5.2. We have focused on points of $\overline{\mathcal{M}}(\mathbb{F}_q)$ that have a coordinate that is maximally hyperbolic for reasons of computational and expositional simplicity. But we note that we could also use points with a maximally elliptic coordinate, where $t \in \mathbb{F}_q^*$ is said to be *maximally elliptic* if L_t has order $q+1$ in $\mathrm{SL}_2(\mathbb{F}_q)$, or equivalently if the roots of the polynomial $T^2 - 3tT + 1$ generate the subgroup of $\mathbb{F}_{q^2}^*$ of index $q-1$. Checking for maximal ellipticity requires a factorization of $q+1$, which could be an advantage if $q+1$ is easier than $q-1$ to factor. And using maximal elliptic points would more-or-less double the probability of finding a point having a fiber on which the associated rotation acts transitively. On the other hand, we would then need to solve the DLP in the order $q+1$ subgroup of $\mathbb{F}_{q^2}^*$, which is more difficult than working in $\mathbb{F}_q$, although still a subexponential problem. In any case, the algorithm still requires solving three DLPs.

6. Checking If $t \in \mathbb{F}_q^*$ Is Maximally Hyperbolic

We analyze the running time of Algorithm 3 in Table 8, which checks whether a given $t \in \mathbb{F}_q^*$ is maximally hyperbolic, i.e., whether the matrix $L_t = \left(\begin{smallmatrix} 3t & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order $q-1$ in $\mathrm{SL}_2(\mathbb{F}_q)$. Algorithm 3 is then invoked in Steps 5, 13, and 22 of the Markoff path-finder algorithm (Algorithm 1 in Table 6).

The first step in Algorithm 3 is to find a non-zero root of

$$T^2 - 3tT + 1 = 0 \tag{14}$$

in $\mathbb{F}_q$, or show there is no such root. This is done by checking if the discriminant $9t^2 - 4$ of (14) is a square, and if it is, using a practical polynomial-time square-root algorithm.

Assuming that the equation (14) has a root $\lambda \in \mathbb{F}_q$, it remains to check whether λ generates $\mathbb{F}_q^*$, i.e., whether λ is a primitive root. The most straightforward way to check this is to first factor $q-1$,

$$q - 1 = \prod_{i=1}^{r} p_i^{e_i},$$

which need only be done once, and then use the elementary fact:

$$\lambda \text{ is a primitive root} \quad\Longleftrightarrow\quad \lambda^{(q-1)/p_i} \ne 1 \text{ for all } 1 \le i \le r. \tag{15}$$

Hence once $q-1$ has been factored, we have:

$$\begin{pmatrix}\text{time to check if } t \in \mathbb{F}_q\\ \text{is maximally hyperbolic}\end{pmatrix} = \begin{pmatrix}\text{time to compute}\\ \text{a square root in } \mathbb{F}_q\end{pmatrix} + \begin{pmatrix}\text{time to compute } r\\ \text{exponentiations in } \mathbb{F}_q\end{pmatrix}.$$

Since taking square roots and doing exponentiations take practical polynomial time, and since $r < \log_2(q)$, the time to check if an element of $\mathbb{F}_q^*$ is maximally hyperbolic is negligible.

Remark 6.1. We note that the factorization of $q-1$ is used to make it easy to check if an element of $\mathbb{F}_q^*$ is a primitive root. However, rather than completely factoring $q-1$, we could instead use Lenstra's elliptic curve factorization algorithm [16] to find all moderately small prime factors. This is very efficient, since the running time for Lenstra's algorithm to factor an integer N depends on the size of the smallest prime factor of N. We can then use the partial factorization to create a probabilistic primitive root algorithm that has a high success rate. Thus for example, if $q \approx 2^{4000}$ and we use Lenstra's algorithm to find all primes $p < 2^{100}$ that divide $q-1$, we can consider the algorithm

$$\lambda \text{ is probably a primitive root} \quad\Longleftrightarrow\quad \lambda^{(q-1)/p} \ne 1 \text{ for all } p \mid q-1,\ p < 2^{100}. \tag{16}$$

The probability that (16) misidentifies an element of $\mathbb{F}_q^*$ as a primitive root when $q \approx 2^{4000}$ is less than

$$1 - \prod_{p \mid q-1,\, p > 2^{100}} \left(1 - \frac{1}{p}\right) < 1 - \left(1 - \frac{1}{2^{100}}\right)^{40} \approx \frac{1}{2^{94}},$$

so the probability is negligible. And even if (16) returns a false positive, the path-finding algorithm can simply restart. Finally, we note the since factoring $q-1$ and solving the DLP in $\mathbb{F}_q^*$ are of roughly the same order of difficulty using the best known algorithms, the saving in only partially factoring $q-1$ is minimal at best.

7. A Heuristic Assumption

In this section we describe the heuristic assumption that underlies our Markoff path-finding algorithm. It says that a random walk in $\overline{\mathcal{M}}(\mathbb{F}_q)$ using the rotations ρ_1 and ρ_3 quickly simulates choosing points in $\overline{\mathcal{M}}(\mathbb{F}_q)$ randomly uniformly. We will apply this heuristic specifically to the assertion that such a random walk has the expected probability of landing in the set of points of $\overline{\mathcal{M}}(\mathbb{F}_q)$ whose y-coordinates are maximally hyperbolic. We refer the reader to Section B for data that supports Heuristic 7.1.

Heuristic Assumption 7.1. *Let $P_0 \in \overline{\mathcal{M}}(\mathbb{F}_q)$, and as in Proposition* 4.1(e)*, we consider the set*

$$\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}} := \bigl\{(x,y,z) \in \overline{\mathcal{M}}(\mathbb{F}_q) : y \ne 0 \text{ and } y \text{ is maximally hyperbolic}\bigr\}.$$

For $n \ge 1$ and for $\boldsymbol{i} = (i_1, i_2, \ldots, i_n) \in \{1,3\}^n$, define the $\boldsymbol{i}$-path of P_0 to be the set of points

$$\mathsf{Path}(\boldsymbol{i}, P_0) := \Bigl\{\rho_{i_1}(P_0),\ \rho_{i_2}\rho_{i_1}(P_0),\ \ldots,\ \rho_{i_n}\cdots\rho_{i_2}\rho_{i_1}(P_0)\Bigr\}.$$

Then for $n > 4\log_2(q)$,

$$\begin{aligned}\operatorname*{Prob}_{\boldsymbol{i}\in\{1,3\}^n}\Big(\mathsf{Path}(\boldsymbol{i},P_0)\cap\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}\neq\emptyset\Big)&\\ \gtrapprox 1-\left(1-\frac{\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}}{\#\overline{\mathcal{M}}(\mathbb{F}_q)}\right)^{n-2\log_2(q)}&\\ \approx 1-\left(1-\frac{\varphi(q-1)}{2(q-1)}\right)^{n-2\log_2(q)}\qquad&\textit{for large } q.\end{aligned}$$

Hence for large q, it is highly likely that a path of length roughly $4\log_2(q)$ will include a point whose y-coordinate is maximally hyperbolic.

Justification. To ease notation, we let

$$\mathcal{T}_q := \overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}} \quad\text{and}\quad \mathcal{S}_q := \overline{\mathcal{M}}(\mathbb{F}_q).$$

It is conjectured that the collection of Markoff graphs $\{\overline{\mathcal{M}}(\mathbb{F}_q)\}$ is an expander family. In any case, it is reasonable to assume that if we take a random path of length greater than (say)

$$n_0(q) := \log_2\big(\overline{\mathcal{M}}(\mathbb{F}_q)\big) \approx 2\log_2(q),$$

then the probability that we land in $\mathcal{T}_q$ is roughly $\#\mathcal{T}_q/\#\mathcal{S}_q$, i.e., the same as if we randomly chose a point in $\mathcal{S}_q$.

Hence as we randomly use ρ_1 and ρ_3 to "rotate" on the x-fiber and the z-fiber, after we get to the $n_0(q)$th point in the path, it is reasonable to view the subsequent points in the path as being random points in $\mathcal{S}_q$, at least insofar as to whether they lie in $\mathcal{T}_q$. Hence in a random path of length $n > n_0(q)$, the path contains $n - n_0(q)$ points that each have probabilty $\#\mathcal{T}_q/\#\mathcal{S}_q$ of being in $\mathcal{T}_q$.

We then use the standard results from probability that if $\mathcal{T}\subseteq\mathcal{S}$ are finite sets and if we randomly choose m points in $\mathcal{S}$ (with replacement), then the probability that we get at least one point in $\mathcal{T}$ is

$$1-\left(1-\frac{\#\mathcal{T}}{\#\mathcal{S}}\right)^m,$$

and that the average number of samples taken from $\mathcal{S}$ before getting an element of $\mathcal{T}$ is $\#\mathcal{S}/\#\mathcal{T}$. Applying this with $m = n - n_0(q)$ gives the first estimate. For the second, Proposition 4.1(c,e) gives us the values of $\#\overline{\mathcal{M}}(\mathbb{F}_q)$ and $\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}$, which after a bit of algebra yields

$$1-\left(1-\frac{\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}}{\#\overline{\mathcal{M}}(\mathbb{F}_q)}\right)^n = 1-\left(1-\frac{\varphi(q-1)}{q-1}\cdot\frac{(1-q^{-1})^2}{2\pm 3q^{-1}}\right)^n.$$

For large q, the q^{-1} terms are negligible, which gives the second estimate. □

Remark 7.2. We note that Remark 1.5 tells us that the probability estimate appearing in Heuristic 7.1 satisfies

$$1-\left(1-\frac{\varphi(q-1)}{2(q-1)}\right)^{n-2\log_2(q)} \geq 1-\left(1-\frac{1}{4\log\log q}\right)^{n-2\log_2(q)}.$$

So for example,the probability in Heuristic 7.1 is very close to 100% for (say) $2^{1000} < q < 2^{10000}$, even in the worst case scenario that $q-1$ is very smooth (which is when the ratio $\varphi(q-1)/(q-1)$ is smallest), provided that we take $n \approx 4\log_2(q)$. We also note that this analysis is very pessimistic; for most initial points P_0, the

mixing will begin almost immediately, leading to many very short paths to points in $\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}$.

8. The Markoff Path-Finder Algorithm

We give the proof of our main result (Theorem 1.1), which we restate for the convenience of the reader.

Theorem 8.1 (Theorem 1.1)**.** *We set the following notation*:

$$\begin{aligned}
\mathsf{PATH}(q) &= \textit{time to find a path between points in } \overline{\mathcal{M}}(\mathbb{F}_q).\\
\mathsf{DLP}(q) &= \textit{time to solve the DLP in } \mathbb{F}_q^*.\\
\mathsf{FACTOR}(N) &= \textit{time to factor } N.\\
\mathsf{NEGLIGIBLE}(q) &= \textit{tasks that take negligible (polylog) time as a function of } q,\\
&\quad \textit{for example taking square roots in } \mathbb{F}_q, \textit{ or iterations}\\
&\quad \textit{performed } (q-1)/\varphi(q-1) \le 2\log\log(q) \textit{ times.}
\end{aligned}$$

Assume that Heuristics 7.1 *is valid. Then with high probability, the Markoff path-finder Algorithm described in detail as Algorithm* 1 *in Table* 6 *will find a path between randomly given points in the graph* $\overline{\mathcal{M}}(\mathbb{F}_q)$ *in time*

$$\mathsf{PATH}(q) \le \mathsf{FACTOR}(q-1) + 3\cdot\mathsf{DLP}(q) + \mathsf{NEGLIGIBLE}(q). \tag{17}$$

Proof. The Markoff path-finder algorithm (Algorithm 1 in Table 6) terminates with a list of positive integers

$$(i_1,\dots,i_\alpha),\ (j_1,\dots,j_\beta),\ (a,b,c)$$

satisfying

$$\begin{aligned}
P' &= \rho_{i_\alpha}\circ\rho_{i_{\alpha-1}}\circ\cdots\circ\rho_{i_2}\circ\rho_{i_1}(P) && \text{Steps 3–10}\\
Q &= \rho_{j_1}\circ\rho_{j_2}\circ\cdots\circ\rho_{j_{\beta-1}}\circ\rho_{j_\beta}(Q') && \text{Steps 11–18}\\
P'' &= \rho_2^a(P'),\ Q' = \rho_3^b(Q''),\ Q'' = \rho_1^c(P'') && \text{Steps 26–30}
\end{aligned}$$

We use these to compute

$$\begin{aligned}
Q &= \rho_{j_1}\circ\cdots\circ\rho_{j_\beta}(Q')\\
&= \rho_{j_1}\circ\cdots\circ\rho_{j_\beta}\circ\rho_3^b(Q'')\\
&= \rho_{j_1}\circ\cdots\circ\rho_{j_\beta}\circ\rho_3^b\circ\rho_1^c(P'')\\
&= \rho_{j_1}\circ\cdots\circ\rho_{j_\beta}\circ\rho_3^b\circ\rho_1^c\circ\rho_2^a(P')\\
&= \rho_{j_1}\circ\cdots\circ\rho_{j_\beta}\circ\rho_3^b\circ\rho_1^c\circ\rho_2^a\circ\rho_{i_\alpha}\circ\cdots\circ\rho_{i_1}(P).
\end{aligned}$$

Hence Algorithm 1 gives the a path in $\overline{\mathcal{M}}(\mathbb{F}_q)$ from P to Q.

We next consider the running time of each step of the algorithm. In Step 2 we factor the integer $q-1$. Once this is done, the time to check whether an element $t\in\mathbb{F}_q$ is maximally hyperbolic is negligible; see Section 6.

In Steps 3–10 and Steps 11–18, we randomly move a point around $\overline{\mathcal{M}}(\mathbb{F}_q)$ and check whether one of its coordinates is maximally hyperbolic. Heuristic 7.1 says that each of these loops needs to look at an average of $2(q-1)/\varphi(q-1)$ points before terminating, and as already noted, checking maximal hyperbolicity takes negligible time once we have factored $q-1$. Similarly, Theorem 4.2 says that the

loop in Step 22 is executed no more than (roughly) $4\log\log q$ times, with the maximal hyperbolicity and the square root computations taking negligible time. Hence Steps 3–25 take average time $12(q-1)/\varphi(q-1)$ multiplied by some small power of $\log(q)$. As explained in Remark 1.5, we have $12(q-1)/\varphi(q-1) \le 24\log\log(q)$, which allows us to conclude that Steps 3–25 take a neglible amount of time.

Steps 26–30 use the MarkoffDLP algorithm three times, and the MarkoffDLP algorithm (Algorithm 2 in Table 7) requires taking a square root in $\mathbb{F}_q^*$ (neglible time) and computing a discrete logarithm in $\mathbb{F}_q^*$. Hence the time to execute Steps 26–30 is essentially the time it takes to compute three DLPs in $\mathbb{F}_q^*$.

Adding these time estimates yields (17), which completes the proof that the Markoff path-finding algorithm terminates in the specified time. □

9. The Markoff Path Finder Algorithm in Action: An Example

We illustrate the Markoff path-finder algorithm (Algorithm 1 in Table 6) by computing a numerical example. We take

$$\begin{aligned} q &= 70687, \qquad q-1 = 2\cdot 3^3\cdot 7\cdot 11\cdot 17,\\ P &= (45506, 13064, 18) \in \overline{\mathcal{M}}(\mathbb{F}_q),\\ Q &= (11229, 5772, 56858) \in \overline{\mathcal{M}}(\mathbb{F}_q). \end{aligned}$$

We use a simplified version of the algorithm in which $i_\alpha = 1$ for all α and $j_\beta = 1$ for all β, since in practice this almost always works. Thus Steps 3–10 say to apply ρ_1 to P until the y-coordinate is maximally hyperbolic. We do a similar computation in Steps 11–18, except now we apply iterates of ρ_1^{-1} to Q and stop when we reach an iterate whose z-coordinate is maximally hyperbolic. Table 3 show our computations. It lists the iterates and indicates whether the appropriate coordinate is hyperbolic or elliptic; and if the coordinate is hyperbolic, it lists $o(\lambda)$, the order of an associated eigenvalue in $\mathbb{F}_q^*$. We find that

$$y\bigl(\rho_1^2(P)\bigr) \quad\text{and}\quad z\bigl(\rho_1^{-15}(Q)\bigr) \quad\text{are maximally hyperbolic,}$$

so the output from Steps 3–18 are

$$\begin{aligned} &\alpha = 2, && i_1 = i_2 = 1, && P' = \rho_1^2(P) = (45506, 40902, 10340),\\ &\beta = 15, && j_1 = \cdots = j_{15} = 1, && Q' = \rho_1^{-15}(Q) = (11229, 2424, 19535). \end{aligned}$$

In Steps 19–25 we randomly choose $x \in \mathbb{F}_q^*$ and check whether x is maximally hyperbolic and whether the quadratic equations

$$F(x, 40902, Z) = 0 \quad\text{and}\quad F(x, Y, 19535) = 0$$

have solutions $y, z \in \mathbb{F}_q$. It took 5 tries, as listed in Table 4. So the output from Steps 19–25 consists of the two points

$$P'' = (29896, 40902, 935) \quad\text{and}\quad Q'' = (29896, 595, 19535).$$

In Steps 26–30 we find a path on the y-fiber from P' to P'', a path on the z-fiber from Q'' to P', and a path on the x fiber from Q'' to P''. This is done using the Markoff DLP Algorithm (Algorithm 2 in Table 7), which uses Proposition 5.1 to convert the path problem in a maximal hyperbolic fiber into a discrete logarithm problem in $\mathbb{F}_q^*$. Implementing this algorithm, we find that

$$P'' = \rho_2^{26986}(P'), \quad Q' = \rho_3^{65193}(Q''), \quad Q'' = \rho_1^{30287}(P'').$$

$i=0$	$P=(45506,13064,18)$	$y=13064$	hyperbolic, $o(\lambda)=1683$
$i=1$	$\rho_1(P)=(45506,18,40902)$	$y=18$	hyperbolic, $o(\lambda)=4158$
$i=2$	$\rho_1^2(P)=(45506,40902,10340)$	$y=40902$	hyperbolic, $o(\lambda)=70686$

$j=0$	$Q=(11229,5772,56858)$	$z=56858$	elliptic
$j=1$	$\rho_1^{-1}(Q)=(11229,65943,5772)$	$z=5772$	hyperbolic, $o(\lambda)=35343$
$j=2$	$\rho_1^{-2}(Q)=(11229,6407,65943)$	$z=65943$	hyperbolic, $o(\lambda)=10098$
$j=3$	$\rho_1^{-3}(Q)=(11229,29942,6407)$	$z=6407$	elliptic
$j=4$	$\rho_1^{-4}(Q)=(11229,16944,29942)$	$z=29942$	elliptic
$j=5$	$\rho_1^{-5}(Q)=(11229,35748,16944)$	$z=16944$	elliptic
$j=6$	$\rho_1^{-6}(Q)=(11229,2200,35748)$	$z=35748$	elliptic
$j=7$	$\rho_1^{-7}(Q)=(11229,66363,2200)$	$z=2200$	elliptic
$j=8$	$\rho_1^{-8}(Q)=(11229,21119,66363)$	$z=66363$	elliptic
$j=9$	$\rho_1^{-9}(Q)=(11229,46109,21119)$	$z=21119$	elliptic
$j=10$	$\rho_1^{-10}(Q)=(11229,47313,46109)$	$z=46109$	hyperbolic, $o(\lambda)=594$
$j=11$	$\rho_1^{-11}(Q)=(11229,7133,47313)$	$z=47313$	elliptic
$j=12$	$\rho_1^{-12}(Q)=(11229,47632,7133)$	$z=7133$	hyperbolic, $o(\lambda)=5049$
$j=13$	$\rho_1^{-13}(Q)=(11229,47838,47632)$	$z=47632$	elliptic
$j=14$	$\rho_1^{-14}(Q)=(11229,19535,47838)$	$z=47838$	hyperbolic, $o(\lambda)=7854$
$j=15$	$\rho_1^{-15}(Q)=(11229,2424,19535)$	$z=19535$	hyperbolic, $o(\lambda)=70686$

TABLE 3. ρ_1 iterates of P until reaching a maximally hyperbolic y-fiber and ρ_1^{-1} iterates of Q until reaching a maximally hyperbolic z-fiber

x	$F(x,40902,Z)$	$F(x,Y,19535)$
29628	irreducible	irreducible
19562	irreducible	$(Y-42621)(Y-57310)$
43036	irreducible	irreducible
6057	$(Z-27506)(Z-70305)$	irreducible
29896	$(Z-935)(Z-45089)$	$(Y-595)(Y-6503)$

TABLE 4. Finding a point on $F(x,40902,Z)=F(x,Y,19535)=0$

Finally, the algorithm outputs

$$(1,1),\ (1,1,\dots,1),\ (26986,30287,65193),$$

where the second item is a 15-tuple. We check that this gives a path from P to Q by computing

$$\begin{aligned} P &= (45506, 13064, 18) \\ \rho_1^2(P) &= (45506, 40902, 10340) \\ \rho_2^{26986} \circ \rho_1^2(P) &= (29896, 40902, 935) \\ \rho_1^{30287} \circ \rho_2^{26986} \circ \rho_1^2(P) &= (29896, 595, 19535) \\ \rho_3^{65193} \circ \rho_1^{30287} \circ \rho_2^{26986} \circ \rho_1^2(P) &= (11229, 2424, 19535) \\ \rho_1^{15} \circ \rho_3^{65193} \circ \rho_1^{30287} \circ \rho_2^{26986} \circ \rho_1^2(P) &= (11229, 5772, 56858) = Q. \end{aligned}$$

If we run the algorithm a second time, the randomness in Steps 19–25 means that we are likely to obtain a different path. (And if we hadn't simplified the choices of the i_α and j_β, that randomness would also lead to different paths.) For example, using the same (q, P, Q) as input and running the algorithm again, we obtained the output

$$(a, c, b) = (26703, 52102, 29583),$$

which gives the path

$$Q = \rho_1^{15} \circ \rho_3^{29583} \circ \rho_1^{52102} \circ \rho_2^{26703} \circ \rho_1^2(P).$$

We also note that we can combine a path from P to Q with a path from Q to P to find a non-trivial loop that starts and returns to P, since it is unlikely that the two paths will be exact inverses of one another. Indeed, running the algorithm to find a path from Q to P, we found

$$(1,1,1), \underbrace{(1,1,\ldots,1)}_{\text{11-tuple}}, (389, 14491, 39906),$$

which gives the path

$$P = \rho_1^{11} \circ \rho_3^{39906} \circ \rho_1^{14491} \circ \rho_2^{389} \circ \rho_1^3(Q).$$

Combining this with the first path from P to Q that we found earlier gives the loop

$$P = \rho_1^{11} \circ \rho_3^{39906} \circ \rho_1^{14491} \circ \rho_2^{389} \circ \rho_1^{18} \circ \rho_3^{65193} \circ \rho_1^{30287} \circ \rho_2^{26986} \circ \rho_1^2(P)$$

where we have combined the middle $\rho_1^3 \circ \rho_1^{15}$ into a single ρ_1^{18}.

10. Markoff-Type K3 Surfaces and the ECDLP

In this section we briefly discuss K3 surfaces that are analogous to the Markoff surface. These surfaces, which were dubbed tri-involutive K3 (TIK3) surfaces in [11], are surfaces

$$\mathcal{W} \subset \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$$

given by the vanishing of a $(2,2,2)$ form. With appropriate non-degeneracy conditions, the three double covers $\mathcal{W} \to \mathbb{P}^1 \times \mathbb{P}^1$ give three non-commuting involutions $\sigma_1, \sigma_2, \sigma_3 \in \mathrm{Aut}(\mathcal{W})$. If the $(2,2,2)$-form is symmetric, then $\mathcal{W}$ also admits coordinate permutation automorphisms, in which case we can define the analogues of the rotations $\rho_1, \rho_2, \rho_3 \in \mathrm{Aut}(\mathcal{W})$. Fuchs, Litman, Tran, and the present author studied the orbit structure of $\mathcal{W}(\mathbb{F}_q)$ for various groups of autormorphisms. In view of [10], one might consider using the graph structure on $\mathcal{W}(\mathbb{F}_q)$ induced by $\{\sigma_1, \sigma_2, \sigma_3\}$ or $\{\rho_1, \rho_2, \rho_3\}$ to implement the CGL [7] hash function algorithm. However, the three fibrations $\mathcal{W} \to \mathbb{P}^1$ have genus 1 fibers, the Jacobians of these

fibrations are elliptic surfaces of rank at least 1, and the action of the automorphisms on fibers can be described in terms of translation by a section to the elliptic surface. See for example [3], where this geometry is explained and explicit formulas are provided.

Thus the Markoff path-finder algorithm, with suitable tweaks, yields a K3 path-finder algorithm whose running time is determined primarily by how long it takes to solve three instances of the elliptic curve discrete logarithm problem. Thus on a classical computer, the algorithm currently takes exponential time to find paths in $\mathcal{W}(\mathbb{F}_q)$, but that is reduced to polynomial time on a quantum computer. However, since the algorithm can look at many elliptic curves lying in the fibration $\mathcal{W}(\mathbb{F}_q) \to \mathbb{P}^1(\mathbb{F}_q)$, it may well be possible to find one whose order is fairly smooth, in which case the ECDLP becomes easier to solve. We have not pursued this further, but it might be interesting to see if under reasonable heuristic assumptions, one can solve the path-finding problem in $\mathcal{W}(\mathbb{F}_q)$ in subexponential time on a classical computer.

Appendix A. Proof Sketch of a General Inclusion/Exclusion Argument

In this section we sketch the inclusion-exclusion argument described in [4] and isolate the crucial irreducibility assumption. For simplicity we restrict to hyperbolic points, but it would not be hard to do something similar for elliptic points.

Remark A.1. We have stated Proposition A.2 in the language of modern algebraic geometry, since that seems the most natural context. But we note that in the special case that $\mathcal{T} \subseteq \mathbb{A}^k$ and $\mathcal{C} \subset \mathbb{A}^{k+n}$ is a family of affine curves, we may view $\mathcal{C}$ as being given by a set of equations

$$F_1(T_1,\dots,T_k,X_1,\dots,X_n) = \cdots = F_r(T_1,\dots,T_k,X_1,\dots,X_n) = 0$$

for some polynomials

$$F_i(T_1,\dots,T_k,X_1,\dots,X_n) \in \mathbb{Z}[T_1,\dots,T_k,X_1,\dots,X_n].$$

Then for any (finite) field $\mathbb{F}$ and any point $\boldsymbol{t} = (t_1,\dots,t_k) \in \mathbb{A}^k(\mathbb{F})$, the set of $\mathbb{F}$-rational points on the fiber of $\mathcal{C}$ over $\boldsymbol{t}$ is the set

$$\mathcal{C}_{\boldsymbol{t}}(\mathbb{F}) = \Big\{(x_1,\dots,x_n) \in \mathbb{A}^n(\mathbb{F}) : F_i(t_1,\dots,t_k,x_1,\dots,x_n) = 0 \text{ for } 1 \le i \le r\Big\}.$$

It is in this setting that we apply Propsostion A.2 to prove Theorem 4.2, using the parameter space

$$\mathcal{T} = \big\{(a,b) \in \mathbb{A}^2 : ab(a^2-4)(b^2-4)(a-b) \ne 0\big\}$$

and the curve $\mathcal{C} \subset \mathbb{A}^2 \times \mathbb{A}^3$ given by the two equations

$$X^2 + a^2 + Z^2 - 3XaZ = X^2 + Y^2 + b^2 - 3XYb = 0.$$

Proposition A.2 (après [4])**.** *Let $\mathcal{T}/\mathbb{Z}$ be a quasi-projective regular scheme over $\mathbb{Z}$, let $\mathcal{C}/\mathbb{Z} \to \mathcal{T}/\mathbb{Z}$ be an irreducible flat[3] family of affine curves, and let $x : \mathcal{C} \to \mathbb{A}^1 \times \mathcal{T}$ be a non-constant $\mathcal{T}$-morphism. Thus for any field $\mathbb{F}$ and any point $\boldsymbol{t} \in \mathcal{T}(\mathbb{F})$, the fiber $\mathcal{C}_{\boldsymbol{t}}$ is a curve defined over $\mathbb{F}$, and there is a map*

$$x : \mathcal{C}_{\boldsymbol{t}}(\mathbb{F}) \longrightarrow \mathbb{A}^1(\mathbb{F}) = \mathbb{F}.$$

[3]Flatness implies that the fibers over any field $\mathbb{F}$ are 1-dimensional and have constant geometric genus; cf. Remark A.1.

Then for every $\epsilon > 0$ there is a constant $C_1(\mathcal{C} \to \mathcal{T}, x, \epsilon)$ such that the following holds:

For each prime power q, each point $\boldsymbol{t} \in \mathcal{T}(\mathbb{F}_q)$, and each $n \geq 1$, define a curve $\mathcal{C}'_{\boldsymbol{t}}[n]/\mathbb{F}_q$ by

$$\mathcal{C}'_{\boldsymbol{t}}[n] := \big\{(P,\mu) \in \mathcal{C}_{\boldsymbol{t}} \times \mathbb{A}^1 : \mu^{2n} - 3x(P)\mu^n + 1 = 0\big\}. \tag{18}$$

We make the following irreducibility assumption:

$$\textbf{Assumption}\text{: For all } n \mid q-1, \text{ the curve } \mathcal{C}'_{\boldsymbol{t}}[n] \text{ is irreducible.} \tag{19}$$

Then[4]

$$\#\big\{P \in \mathcal{C}_{\boldsymbol{t}}(\mathbb{F}_q) : x(P) \text{ is maximally hyperbolic}\big\} \geq \frac{1}{2}\varphi(q-1) - C_1 \cdot q^{\frac{1}{2}+\epsilon}.$$

Proof. We consider the curve family of curves $\mathcal{C}' \to \mathcal{T}$ given informally as the set of points

$$\mathcal{C}' := \big\{(P,\lambda) \in \mathcal{C} \times \mathbb{A}^1 : \lambda^2 - 3x(P)\lambda + 1 = 0\big\},$$

and more generally, for each $n \geq 1$ we define $\mathcal{C}'[n] \to \mathcal{T}$ by[5]

$$\mathcal{C}'[n] := \big\{(P,\mu) \in \mathcal{C} \times \mathbb{A}^1 : \mu^{2n} - 3x(P)\mu^n + 1 = 0\big\},$$

so in particular $\mathcal{C}' = \mathcal{C}'[1]$.

There are natural maps

$$\begin{array}{ccccc} \mathcal{C}'[n] & \xrightarrow[\text{degree } n]{F_n} & \mathcal{C}' & \xrightarrow[\text{degree } 2]{} & \mathcal{C}, \\ (P,\mu) & \longmapsto & (P,\mu^n),\ (P,\lambda) & \longmapsto & P, \end{array}$$

and for each $\boldsymbol{t} \in \mathcal{T}(\mathbb{F}_q)$, we get curves $\mathcal{C}'_{\boldsymbol{t}}[n]$, $\mathcal{C}'_{\boldsymbol{t}}$, and $\mathcal{C}_{\boldsymbol{t}}$ and corresponding maps.

We note that for $n \mid q-1$, the field $\mathbb{F}_q$ contains a primitive nth root of unity, so the map F_n, which is essentially the nth power map on one of the coordinates, has the property that

$$F_n : \mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q) \longrightarrow \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q) \quad \text{is exactly } n\text{-to-1,} \tag{20}$$

except for a negligible set of points whose cardinality is bounded independent of q.[6]

We first estimate the size of

$$\mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q)^{\mathrm{prim}} := \big\{(P,\lambda) \in \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q) : \lambda \text{ generates } \mathbb{F}_q^*\big\}.$$

We replace our curves with non-singular projective models, which again introduces a negligible set of points. Then the Riemann–Hurwitz genus formula [13, Corollary 2.4] tells us that the genera of the $\mathcal{C}'_{\boldsymbol{t}}[n]$ satisfy

$$g\big(\mathcal{C}'_{\boldsymbol{t}}[n]\big) = 1 + n\big(g(\mathcal{C}'_{\boldsymbol{t}}) - 1\big) + \frac{1}{2}\sum_{\gamma \in \mathcal{C}'_{\boldsymbol{t}}} (e_\gamma - 1) = O\big(n\big),$$

[4]We remark that an analogous calculation gives the same estimate for the number of P such that $x(P)$ is maximally elliptic.

[5]Scheme-theoretically, we define an affine scheme $\mathcal{D}[n] = \big\{(\mu,\xi) \in \mathbb{A}^2 : \mu^{2n} - 3\xi\mu^n + 1 = 0\big\}$, and then $\mathcal{C}'[n]$ is the fiber product fitting into the diagram

$$\begin{array}{ccc} \mathcal{C}'[n] & \longrightarrow & \mathcal{C} \\ \downarrow & & \downarrow x \\ \mathcal{D}[n] \times \mathcal{T} & \xrightarrow{\mathrm{proj}_1 \times \mathbf{1}_{\mathcal{T}}} & \mathbb{A}^1 \times \mathcal{T} \end{array}$$

[6]More precisely, points where $\mathcal{C}'_{\boldsymbol{t}}[n]$ or $\mathcal{C}'_{\boldsymbol{t}}$ is singular and points where F_n is ramified over $\mathbb{F}_q$, which may vary to some extent depending on q, but are bounded in terms of the geometry of $F_n : \mathcal{C}'[n] \to \mathcal{C}'$ over $\mathbb{Z}$.

where the big-O constant depends only on the genus of $\mathcal{C}_{\boldsymbol{t}}$ and the degree of the map $x : \mathcal{C}_{\boldsymbol{t}} \to \mathbb{P}^1$. (We also note that the map F_n is separable over $\mathbb{F}_q$, since $\deg(F_n) = n < q$.) Then Weil's estimate for the number of points on *irreducible* curves over finite fields yields[7]

$$\#\mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q) = q + O\Big(g\big(\mathcal{C}'_{\boldsymbol{t}}[n]\big)\sqrt{q}\Big) = q + O(n\sqrt{q}). \tag{21}$$

We note that $(P,\lambda) \in \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q)$ satisfies:

$$(P,\lambda) \in F_n\big(\mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q)\big) \quad\Longleftrightarrow\quad \text{the order of } \lambda \text{ in } \mathbb{F}_q^* \text{ divides } \frac{q-1}{n}. \tag{22}$$

Ignoring a negligible number of points that are singular, "at infinity," or ramified, we have

$$\mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q)^{\mathrm{prim}} = \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q) \smallsetminus \bigcup_{\substack{n\mid q-1\\ n\neq 1}} F_n\big(\mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q)\big). \tag{23}$$

We use this to calculate (omitting an $O(1)$ term coming from the "negligible" points)

$$\begin{aligned}
\#\mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q)^{\mathrm{prim}} &\approx \sum_{n\mid q-1} \mu\left(\frac{q-1}{n}\right) \#F_n\big(\mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q)\big) \quad \text{using inclusion/exclusion with (22) and (23),}\\
&\approx \sum_{n\mid q-1} \mu\left(\frac{q-1}{n}\right) \cdot \frac{1}{n}\#\mathcal{C}'_{\boldsymbol{t}}[n](\mathbb{F}_q) \quad \text{from (20),}\\
&= \sum_{n\mid q-1} \mu\left(\frac{q-1}{n}\right) \cdot \Big(\frac{q}{n} + O(\sqrt{q})\Big) \quad \text{from (21),}\\
&= q \cdot \frac{\varphi(q-1)}{q-1} + O\big(d(q-1)\sqrt{q}\big),
\end{aligned} \tag{24}$$

where $d(N)$ is the number of divisors of N. We note that $d(N) \le N^{2/\log\log N}$ for all $N \ge 3$ (and indeed, the 2 can be improved), so in particular $d(N) \le C_2(\epsilon)N^{\epsilon}$. We compute

$$\begin{aligned}
&\#\big\{P \in \mathcal{C}_{\boldsymbol{t}}(\mathbb{F}_q) : x(P) \text{ is maximally hyperbolic}\big\}\\
&\qquad \ge \frac{1}{2}\#\big\{(P,\lambda) \in \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q) : \lambda \text{ generates } \mathbb{F}_q^*\big\}\\
&\qquad\qquad \text{since the map } \mathcal{C}'_{\boldsymbol{t}} \to \mathcal{C}_{\boldsymbol{t}} \text{ has degree 2, so the}\\
&\qquad\qquad \text{map } \mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q) \to \mathcal{C}_{\boldsymbol{t}}(\mathbb{F}_q) \text{ is at most 2-to-1,}\\
&\qquad = \frac{1}{2}\#\mathcal{C}'_{\boldsymbol{t}}(\mathbb{F}_q)^{\mathrm{prim}}\\
&\qquad = \frac{q\varphi(q-1)}{2(q-1)} + O\big(d(q-1)\sqrt{q}\big) \quad \text{from (24),}\\
&\qquad = \frac{1}{2}\varphi(q-1) + O_\epsilon(q^{\frac{1}{2}+\epsilon}) \quad \text{since } d(N) \le C_3(\epsilon)N^{\epsilon}.
\end{aligned}$$

[7]We stress that this is where we use the assumption (19) that $\mathcal{C}'_{\boldsymbol{t}}[n]$ is irreducible, since if it were not, then the number of $\mathbb{F}_q$-rational points would be roughly q times the number of irreducible components. Extra factors of this sort would invalidate the inclusion-exclusion argument, which involves summing both positive and negative terms.

q	$q-1$	$\frac{\#\overline{\mathcal{M}}(\mathbb{F}_q)^{y\text{-max-hyp}}}{\#\overline{\mathcal{M}}(\mathbb{F}_q)}$	Heuristic 7.1 Experiment
17389	$2^2 \cdot 3^3 \cdot 7 \cdot 23$	7.320	10.032
48611	$2 \cdot 5 \cdot 4861$	5.001	6.185
55163	$2 \cdot 27581$	4.000	4.662
70687	$2 \cdot 3^3 \cdot 7 \cdot 11 \cdot 17$	8.181	11.507
104729	$2^3 \cdot 13 \cdot 19 \cdot 53$	4.662	5.654
200560490131	$2 \cdot 3 \cdot 5 \cdot \cdots \cdot 29 \cdot 31$	13.085	20.230

TABLE 5. Experiments to test Heuristic 7.1 (100000 samples)

This completes our sketch[8] of the proof of Proposition A.2. □

APPENDIX B. COMPUTATIONS TO CHECK HEURISTIC 7.1

Remark B.1 (**Testing Heuristic 7.1**)**.** We choose a random point P in $\overline{\mathcal{M}}(\mathbb{F}_q)$ and randomly apply ρ_1 or ρ_3 until the y-coordinate of the resulting point is maximally hyperbolic. For each prime in Table 5, we compute the average value of n for 10^5 randomly chosen points. We compare this with the theoretical value $2(q-1)/\varphi(q-1)$, which is the theoretical expected number of trials to find a maximally hyperbolic element in $\mathbb{F}_q^*$.

Remark B.2. We note that the experimental values in Table 5 are somewhat larger than expected, especially when $q-1$ is quite smooth. We are not sure what is causing the discrepency, possibly the "random walks" using two rotations are somewhat less random than expected due the presence of loops caused by collisions. In any case, the experimental numbers are small enough that even for q of cryptographic size, the number of iterations of Steps 3–10 and Steps 11–18 in the Markoff path-finder algorithm (Algorithm 1 in Table 6) will be practical.

APPENDIX C. THE MARKOFF PATH-FINDER ALGORITHM AND SUBROUTINES

The following algorithms are described in Tables 6–8 in this section.

Algorithm 1 - MarkoffPathFinder**:** Returns a path in $\overline{\mathcal{M}}(\mathbb{F}_q)$ from P to Q.

Algorithm 2 - MarkoffDLP**:** Returns an integer $n \geq 0$ so that $P = \rho_k^n(Q)$ in $\overline{\mathcal{M}}(\mathbb{F}_q)$.

Algorithm 3 - MaximalEllipticQ**:** Returns **true** if t is maximal hyperbolic in $\mathbb{F}_q^*$, i.e., if the matrix $\left(\begin{smallmatrix} 3t & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order $q-1$ in $\mathrm{SL}_2(\mathbb{F}_q)$; otherwise returns **false**. It assumes that a factorization of $q-1$ is known; but see Remark 6.1.

[8]The reason that we call this a sketch is we have omitted rigorously tracking the sets of singular, ramification, and at infinity points that we have asserted are negligible in the calculation.

Algorithm 1 MarkoffPathFinder

Input: q, P, Q with $P, Q \in \overline{\mathcal{M}}(\mathbb{F}_q)$

comment: Use a factorization algorithm to factor $q-1$ and store it so that it is accessible by subroutines.
PrimeFactorList $\leftarrow$ {primes that divide $q-1$}
comment: Randomly move P using ρ_1 and ρ_3 until the y-coordinate is maximally hyperbolic
$P' \leftarrow P,\ \alpha \leftarrow 0$
while MaximalEllipticQ$\big(q, y(P')\big) =$ **false do**
  $\alpha \leftarrow \alpha + 1$
  $i_\alpha \xleftarrow{\$} \{1,3\}$
  $P' \leftarrow \rho_{i_\alpha}(P')$
end while
comment: $P' = \rho_{i_\alpha} \circ \rho_{i_{\alpha-1}} \circ \cdots \circ \rho_{i_2} \circ \rho_{i_1}(P)$
comment: Randomly move Q using ρ_1^{-1} and ρ_2^{-1} until the z-coordinate is maximally hyperbolic
$Q' \leftarrow Q,\ \beta \leftarrow 0$
while MaximalEllipticQ$\big(q, z(Q')\big) =$ **false do**
  $\beta \leftarrow \beta + 1$
  $j_\beta \xleftarrow{\$} \{1,2\}$
  $Q' \leftarrow \rho_{j_\beta}^{-1}(Q')$
end while
comment: $Q = \rho_{j_1} \circ \rho_{j_2} \circ \cdots \circ \rho_{j_{\beta-1}} \circ \rho_{j_\beta}(Q')$
comment: Find random points with the same maximally hyperbolic x-coordinate that can be used to connect P' to Q'
repeat
  $x \xleftarrow{\$} \mathbb{F}_q^*$
until MaximalEllipticQ$(q,x) =$ **true and** $F\big(x, y(P'), Z\big)$ has a root $z \in \mathbb{F}_q$ **and** $F\big(x, Y, z(Q')\big)$ has a root $y \in \mathbb{F}_q$
$P'' \leftarrow \big(x, y(P'), z\big)$
$Q'' \leftarrow \big(x, y, z(Q')\big)$
comment:
- P'' and Q'' are on the same maximally hyperbolic x-fiber.
- P' and P'' are on the same maximally hyperbolic y-fiber.
- Q' and Q'' are on the same maximally hyperbolic z-fiber.

comment:
- Find fibral paths $P'' \to P$ and $P' \to Q''$ and $Q'' \to P''$.
- Proposition 5.1 ensures that such paths exist.

$a \leftarrow$ MarkoffDLP$(q, P', P'', 2)$
$b \leftarrow$ MarkoffDLP$(q, Q'', P', 3)$
$c \leftarrow$ MarkoffDLP$(q, P'', Q'', 1)$
comment: $P'' = \rho_2^a(P'),\ Q' = \rho_3^b(Q''),\ Q'' = \rho_1^c(P'')$

Output: $(i_1, \ldots, i_\alpha),\ (j_1, \ldots, j_\beta),\ (a, b, c)$

TABLE 6. Returns a path in $\overline{\mathcal{M}}(\mathbb{F}_q)$ from P to Q

Algorithm 2 MarkoffDLP

Input: $q, P, Q,\ k$ with $P, Q \in \overline{\mathcal{M}}(\mathbb{F}_q)$ and $k \in \{1, 2, 3\}$ and the kth coordinate of P maximally hyperbolic

comment: if $k = 2$ (y-fiber) or $k = 3$ (z-fiber), swap coordinates to use the x-fiber
if $k = 2$ **then**
  $P \leftarrow \big(y_P, z_P, x_P\big)$
  $Q \leftarrow \big(y_Q, z_Q, x_Q\big)$
else if $k = 3$ **then**
  $P \leftarrow \big(z_P, x_P, y_P\big)$
  $Q \leftarrow \big(z_Q, x_Q, y_Q\big)$
end if
comment: Now x_P is maximally hyperbolic.
$\lambda \leftarrow \big(3x_P + \sqrt{9x_P^2 - 4}\big)/2$ in $\mathbb{F}_q$
comment: The maximal hyperbolicity of x_P says that λ generates $\mathbb{F}_q^*$.
$b \leftarrow \big(y_Q - z_Q/\lambda\big)\big/\big(y_P - z_P/\lambda\big)$
Use a DLP algorithm to find n so that $\lambda^n = b$ in $\mathbb{F}_q^*$.

Output: n

TABLE 7. Returns an integer $n \geq 0$ so that $P = \rho_k^n(Q)$ in $\overline{\mathcal{M}}(\mathbb{F}_q)$. See Proposition 5.1 for an explanation of why this algorithm works.

Algorithm 3 MaximalEllipticQ

Input: q, t

result $\leftarrow$ **false**
comment: Check if $T^2 - 3tT + 1$ has two distinct roots in $\mathbb{F}_q$
if $(9t^2 - 4)^{(q-1)/2} = 1$ in $\mathbb{F}_q$ **then**
  $\lambda \leftarrow (3t + \sqrt{9t^2 - 4})/2$ in $\mathbb{F}_q$
  result $\leftarrow$ **true**
  for $p \in$ PrimeFactorList **do**
    if $\lambda^{(q-1)/p} = 1$ in $\mathbb{F}_q^*$ **then**
      result $\leftarrow$ **false**
    end if
  end for
end if

Output: *result*

TABLE 8. Check whether t is maximally hyperbolic, or equivalently, whether the matrix $L_t \leftarrow \left(\begin{smallmatrix} 3t & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order $q-1$ in $\mathrm{SL}_2(\mathbb{F}_q)$, or equivalently, whether L_t has an eigenvalue in $\mathbb{F}_q^*$ that is a primitive root.

References

[1] A. Baragar. *The Markoff equation and equations of Hurwitz*. ProQuest LLC, Ann Arbor, MI, 1991. Thesis (Ph.D.)–Brown University.

[2] E. Bellah, S. Chen, E. Fuchs, and L. Ye. Bounding lifts of markoff triples mod p, 2023. arXiv:2311.11468.

[3] M. Bhargava, W. Ho, and A. Kumar. Orbit parametrizations for K3 surfaces. *Forum Math. Sigma*, 4:Paper No. e18, 86, 2016.

[4] J. Bourgain, A. Gamburd, and P. Sarnak. Markoff surfaces and strong approximation, 1, 2016. arXiv:1607.01530, (updated Dec 2023, private communication).

[5] J. Bourgain, A. Gamburd, and P. Sarnak. Markoff triples and strong approximation. *C. R. Math. Acad. Sci. Paris*, 354(2):131–135, 2016.

[6] A. Cerbu, E. Gunther, M. Magee, and L. Peilen. The cycle structure of a Markoff automorphism over finite fields. *J. Number Theory*, 211:1–27, 2020.

[7] D. X. Charles, K. E. Lauter, and E. Z. Goren. Cryptographic hash functions from expander graphs. *J. Cryptology*, 22(1):93–113, 2009.

[8] W. Chen. Nonabelian level structures, nielsen equivalence, and Markoff triples, 2020. arXiv:2011.12940.

[9] M. de Courcy-Ireland and M. Magee. Kesten-McKay law for the Markoff surface mod p. *Ann. H. Lebesgue*, 4:227–250, 2021.

[10] E. Fuchs, K. Lauter, M. Litman, and A. Tran. A cryptographic hash function from Markoff triples. *Mathematical Cryptology*, 1(1):103–121, Jan. 2022.

[11] E. Fuchs, M. Litman, J. H. Silverman, and A. Tran. Orbits on K3 surfaces of Markoff type. *Experimental Math.*, published online 03 Aug 2023. doi.org/10.1080/10586458.2023.2239265.

[12] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[13] R. Hartshorne. *Algebraic geometry*. Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977.

[14] J. Hoffstein, J. Pipher, and J. H. Silverman. *An introduction to mathematical cryptography*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2014.

[15] S. V. Konyagin, S. V. Makarychev, I. E. Shparlinski, and I. V. Vyugin. On the structure of graphs of Markoff triples. *Q. J. Math.*, 71(2):637–648, 2020.

[16] H. W. Lenstra, Jr. Factoring integers with elliptic curves. *Ann. of Math. (2)*, 126(3):649–673, 1987.

[17] A. Markoff. Sur les formes quadratiques binaires indéfinies. *Math. Ann.*, 17(3):379–399, 1880.

[18] J. B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.*, 6:64–94, 1962.

[19] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *35th Annual Symposium on Foundations of Computer Science (Santa Fe, NM, 1994)*, pages 124–134. IEEE Comput. Soc. Press, Los Alamitos, CA, 1994.

[20] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997.

Email address: joseph_silverman@math.brown.edu

Mathematics Department, Box 1917 Brown University, Providence, RI 02912 USA. ORCID: 0000-0003-3887-3248