MARKOFF UNIQUENESS, PANTS AND THE BUNYAKOVSKY CONJECTURE

GREG MCSHANE

Abstract. We show that the Bunyakovsky conjecture implies the Markov uniqueness conjecture. Explicitly: If for every positive integer m the sequence

$$k^2m^2 + 1, k \in \mathbb{Z}$$

contains a prime then Markov uniqueness conjecture is true. Surprisingly, though our result relates these two well known conjectures in number theory, the ingredients of our proof are for the most part geometric and even topological.

1. Introduction

1.1. **Markoff numbers.** A *Markoff triple* is a solution (X, Y, Z) in positive integers to the *Markoff cubic*

$$X^2 + Y^2 + Z^2 - 3XYZ = 0. \qquad (1)$$

A Markoff triple is said to be *normalised* iff $z > y > x$, Aside from the two smallest singular triples $(1, 1, 1)$ and $(1, 1, 2)$, every Markov triple consists of three distinct integers. A *Markoff number* is an integer in a Markoff triple.

1.2. **Uniqueness of Markoff Numbers.** The unicity conjecture, apparently a question asked by Frobenius [8], says that the largest number in a Markoff triple determines the remaining two numbers or alternatively that any Markoff number greater than 2 belongs to a unique normalised triple. Button and Baragar (see chapter 10 of Aigner [1]) used basic algebraic number theory to show that certain Markoff numbers satisfied the uniqueness conjecture.

Theorem 1.1 (Baragar, Button, Schmutz). *Let m be a Markoff number of the form $m = p^k$ or $m = 2p^k$ then it is unique if p is an odd prime.*

In fact Baragar and Button also proved that the uniqueness conjecture holds when m is replaced by $3m+2$ and $3m-2$ in the statement of the theorem. Our work here started as an attempt to show these results geometrically. Although we were not successful in this we think our Theorem 1.3 is of some interest in its own right.

Subsequently Aigner extended this showing:

Theorem 1.2 (Aigner). *Let m be a Markoff number of the form*

$$m = Np^k$$

where p is an odd prime and $N \leq 10^{35}$ is another Markoff number. Then m is unique.

We will not be concerned with Aigner's theorem but we will give a short proof of Button's theorem using some geometry and the fact that the ring of Gaussian integers is a unique factorisation domain. This approach will allow us to show that the unicity of Markoff numbers is a corollary of the Bunyakovsky conjecture. More precisely:

Theorem 1.3. *If for every Maroff number m the sequence*

$$k^2m^2 + 1, k \in \mathbb{Z}$$

contains a prime then the uniqueness conjecture is true.

In fact, if we drop the hypothesis that m is a Markoff number and suppose that m is any non zero integer then the polynomial $m^2X^2 + 1$ will still satisfy the hypothesis of the Bunyakovsky conjecture [3] namely:

- it has a positive leading coefficient
- it is irreducible in $\mathbb{Z}[X]$;
- its reduction is non trivial in $\mathbb{F}_p[X]$ for all primes p.

According to the conjecture, under these hypotheses, the the restriction of the polynomial $m^2X^2 + 1$ to the integers has an infinity of prime values. The simplest example is $m = 1$ and the polynomial $X^2 + 1$ and this is Landau's conjecture (studied also by Euler[5]) that there are infinitely many primes of the form $k^2 + 1$. However, the Bunyakovsky Conjecture has been proved only in the case where the polynomial has degree 1: this is Dirichlet's Theorem, that if a and b are coprime integers then there are infinitely many primes of the form $ak + b$. The interested reader can find more details and its relation to another conjecture in [15] and [9].

The proof of our theorem, though it relates these two conjectures apparently from number theory, has an essential topological ingredient in the form of Lemma 1.5. The numbers $k^2m^2 + 1$ arise as λ-lengths on a pair of pants. A *pair of pants* is a hyperbolic surface homeomorphic to a sphere minus 3 disjoint discs and is obtained as a quotient C/G where G is a Fuchsian group and $C \subset \mathbb{H}$ is a convex subset, in fact the convex hull of the limit set of G. In the text we will abuse notation and say that the pants is the quotient $\mathbb{H}/G$.

1.3. Geodesics on hyperbolic surfaces.

In the mid 20th century H. Cohn initiated a program which resulted in a correspondence between Markoff numbers and the lengths of simple closed geodesics on the modular torus. Recall that the *modular torus* is the surface obtained as the quotient of the Poincaré half space $\mathbb{H}/\Gamma'$ where $\Gamma' < P\mathrm{SL}(2,\mathbb{Z})$ is the commutator subgroup. The correspondence can be stated as follows, if γ is such a geodesic then:

$$X = \frac{2}{3}\cosh\left(\frac{\ell_\gamma}{2}\right), \tag{2}$$

is a Markoff number where ℓ_γ is the length of γ. Conversely, every Markoff number arises as the length of such a geodesic.

There is an equivalent conjecture to the Markoff conjecture (Conjecture 3 [12]) which concerns simple closed geodesics on $\mathbb{H}/\Gamma'$:
Conjecture: *The modular torus $\mathbb{H}/\Gamma'$ has the following property: if α, β is a pair of simple closed geodesics of the same length, then there is an automorphism of $\mathbb{H}/\Gamma'$ taking one to the other.*

Since the group of automorphisms of $\mathbb{H}/\Gamma'$ is isomorphic to a cyclic group of order 6 this means that the multiplicity of any number in the spectrum of lengths of simple closed geodesics is at most 6.

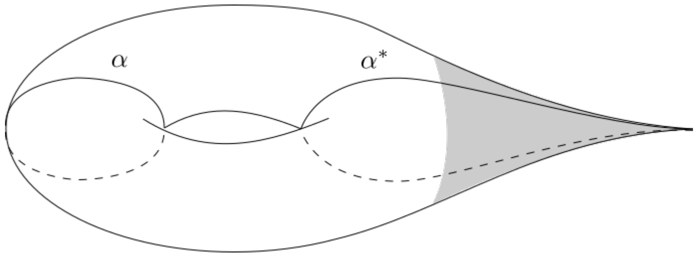

FIGURE 1. A punctured torus with a closed simple geodesic α and the arc α^* disjoint from it. The shaded area is a cusp region.

Among the automorphisms of $\mathbb{H}/\Gamma'$ there is one in particular that we will be concerned with namely the *elliptic involution* which is the unique automorphism of order 2. The elliptic involution arises as follows: Γ' is normalised by $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and so this induces an involution of $\mathbb{H}/\Gamma'$. A *Weierstrass point* is one of the three fixed points of this involution. The interaction of this automorphisms with the various geodesics on $\mathbb{H}/\Gamma'$ is central to the proof of Theorem 1.3.

1.4. **Introducing λ-lengths.** The closed geodesics form an interesting class of curves that has been much studied not just in relation to the Markoff numbers. There is another class of curves, namely the *bicuspidal geodesics or arcs*, that have also proved to be important in many contexts for example Penner's theory of moduli space of surfaces with marked points and more recently cluster algebras. Though each arc is non compact and so has infinite length Penner [14] introduced a modified notion of length, λ-*length*, which allows one to perform useful calculations. Penner's λ-length of simple bicuspidal geodesic on a punctured surface is essentially the exponential of the length of the portion outside of some fixed system of cusp regions. Recall that a *cusp region* in a hyperbolic surface is a portion of the surface isometric to the quotient of $\{z, \operatorname{Im} z > 1\}$ by the action of the group generated by $z \mapsto z + A, A > 0$. A simple calculation shows that the area of the cusp region is A (see [11] for a discussion). On the modular torus there are two natural choices for the cusp region :

- the universal cusp region H_2 of area 2 present on any cusped hyperbolic surface.
- the maximal cusp region H_6 of area 6

(see [11] for details).

For the purposes of this paper we will uses a slightly more general notion of λ length: Let γ be a bicuspidal geodesic and H be a cusp region and choose a lift of γ, $\hat{\gamma} \subset \mathbb{H}$, joining $\gamma^{\pm} \in \partial\mathbb{H}$. The λ-*length* of γ with respect to H is defined to be the exponential of the length of the portion of $\hat{\gamma}$ outside of the pair of lifts of H tangent to $\partial\mathbb{H}$ at $\gamma^{\pm}$.

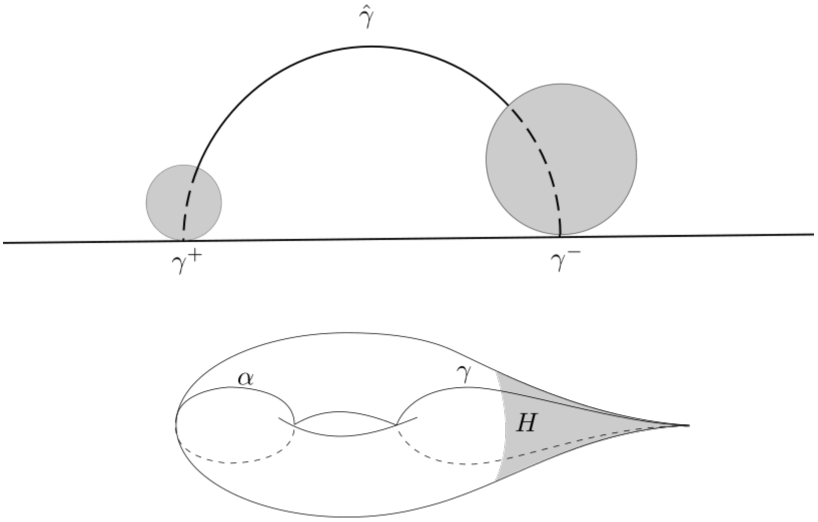

FIGURE 2. Above a lift of γ and the terminal lifts of the cusp region H of the punctured torus below. The solid arc of the demi circle is the portion used to define the λ-length with respect to H.

Using calculations in Wolpert [19] we will see that, for the maximal cusp region on the modular torus, the λ-lengths of arcs coincide with the squares of Markoff numbers. More precisely:

Theorem 1.4. *For each Markoff triple (X, Y, Z) there is a (unique) ideal triangulation of the modular torus such that the λ-lengths of the arcs are X^2, Y^2, Z^2.*

Then, using the fact that each arc is invariant under the elliptic involution one can show, by applying Lemma 2.4, that every Markoff number is the sum of two squares that is it factorises over the Gaussian integers.

1.5. **Idea of the proof.** Our proof of Button's theorem is based on the following ingredients:

- each closed simple geodesic α on $\mathbb{H}/\Gamma'$ is paired with a unique arc (simple bicuspidal geodesic) α^*.
- each arc passes through a single Weierstrass point on the punctured torus
- a reformulation (Theorem 3.3) of the unicity conjecture in terms of multiplicities for lengths of simple geodesics (see [12]).

On closer examination one sees that we don't actually need the bicuspidal geodesic to be simple only that

- it be disjoint from α,
- is invariant under the elliptic involution,
- has λ-length a prime.

1.5.1. *Key topological lemma.* It is important to note here that any pair of distinct simple closed geodesics on the punctured torus is *filling* that is the complement consists of a union of discs and a punctured disc. It follows from this that:

Lemma 1.5. *On a punctured torus any bicuspidal geodesic, whether simple or not, is disjoint from at most one closed simple geodesic.*

1.5.2. *Proving unicity for a Markoff number m.* Our argument reduces to:

if γ, γ' is a pair of bicuspidal geodesic

- each of which is invariant under the elliptic involution
- each has λ-length some prime p

then there is an automorphism of $\mathbb{H}/\Gamma'$ which takes γ to γ'. This is because the prime p splits in an essentially unique way over the Gaussian integers which means that there are at most six such geodesics on $\mathbb{H}/\Gamma'$. Now if in addition γ say is disjoint from a simple closed geodesic α then any other simple geodesic of the same length is the image of α under some automorphism of $\mathbb{H}/\Gamma'$ too. (see discussion of Conjecture 3 [12] above).

Thus we have proved Theorem 1.3 and all that is left to do is to show that $m^2k^2 + 1$ is indeed the λ-length a geodesic of the required type.

1.6. **Organisation, Remarks.** In Section 2 we recall the definition of Ford circles and discuss their relation to the geodesics of the Poincaré half plane. We study the orbit of the imaginary point i under $P\mathrm{SL}(2, \mathbb{Z})(\mathbb{Z})$ and how this relates to the representation of an integer as the sum of

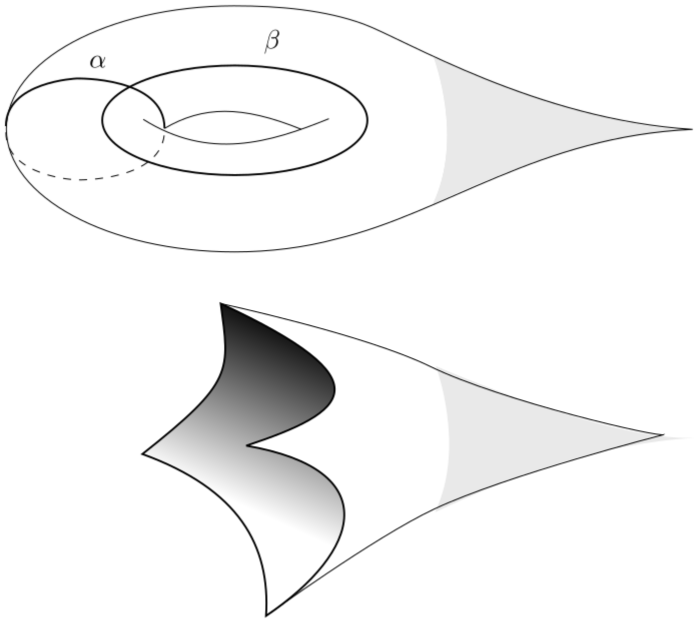

Figure 3. Above a pair of simple closed geodesics α and β on the punctured torus. When we cut along the geodesics we obtain the punctured disc below.

two squares. Section 3 is an exposition of the Markoff cubic viewed as the character variety of the free group on two generators. We then explain how to pass to λ-lengths and prove Theorem 1.4. In Section 4 we prove Button's theorem using the results of Sections 2 and 3. Finally we extend this proof to obtain Theorem 1.3 and show how the sequence k^2m^2+1 arises as a set of λ-lengths on a certain pair of pants.

The content of sections 2 and 3 is purely expository and, as such, we make no claim of originality. We will assume that the reader has some familiarity with the theory of Fuchsian groups. Almost all of the material in Sections 2 be found in Serre's book [17] and the reader should not need any other references to understand this paper.

This work grew out of an attempt to find a geometric proof of Fermat's theorem on sums of squares to understand Heath-Brown and Zagier's proofs [20]:

Theorem 1.6 (Fermat). *Let p be a prime then the equation*

$$x^2+y^2=p$$

has a solution in integers iff $p=2$ or $p-1$ is a multiple of 4.

We succeeded, the result is to be found in [10], and much of Section 2 is lifted from that paper.

Another factor was recent work by Bourgain, Gamburd and Sarnak proving that almost all Markoff numbers are highly composite. Inspiration then came from listening to Dennis Sullivan's lecture at the Abel Prize ceremony in 2022. The idea of looking at other arcs on pants is how I fancifully imagined he, as a topologist, might have tried to deal with composite Markoff numbers.

1.7. **Thanks.** It is a pleasure to thank Louis Funar, Alexis Marin, Bob Penner and Vlad Sergesciu for conversations over the years concerning this subject. We would also like to thank Yi Huang for reading early drafts of the manuscript and Boris Springborn for telling us about his work [18] which is closely related to ours.

This paper was written using vimtex for neovim and copilot.vim [23]. As such some of the text may be a little idiosyncratic in that it was composed by an AI.

2. Ford circles, lengths, midpoints

Let F denote the set $\{z, \operatorname{Im} z > 1\}$ this is a *horoball in* $\mathbb{H}$ centered at ∞. The image of F under the $\mathrm{SL}(2,\mathbb{Z})$ action consists of F and infinitely many disjoint disks, the so-called *Ford circles*, each tangent to the real line at some rational m/n. We adopt the convention that F is also a Ford circle of infinite radius. Our interest in Ford circles stems from:

Lemma 2.1. *The Ford circles are invariant under* $\mathrm{SL}(2,\mathbb{Z})$ *and so invariant under* $\Gamma' < \mathrm{SL}(2,\mathbb{Z})$ *and their projection to* $\mathbb{H}/\Gamma'$ *is the maximal cusp region* H_6 *of area* 6.

Proof. The Ford circles are invariant under Γ' and so project to a cusp region on $\mathbb{H}/\Gamma'$. The stabiliser of F in Γ' is the group generated by $z \mapsto z+6$. Now the area of the cusp region is 6 by the formula for the area of a cusp region. □

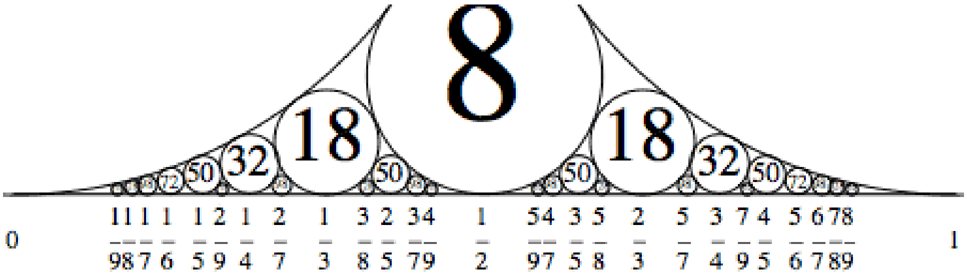

Figure 4. Ford circles with tangent points and curvatures. Recall that the curvature of a euclidean circle is twice the square of the reciprocal of its radius.

The following is well known and is easily checked:

Lemma 2.2. *The Ford circle tangent to the real line at m/n has Euclidean diameter $1/n^2$.*

We define the *length* of the vertical line $\{k/p + it,\ t \in \mathbb{R}\}$ to be the length of the sub arc joining F to the Ford circle tangent at k/p. Further we define its *mid point* to be the midpoint of this sub arc. We remark that if the projection of the line to $\mathbb{H}/\Gamma'$ is invariant by an automorphism then the midpoint is necessarily a fixed point of the automorphism.

The following is a restatement of Lemma 2.2 in terms of these notions:

Lemma 2.3. . *Let m/n be a rational. Then the geodesic $\{m/n+it,\ t \in \mathbb{R}\}$*

- *has length $2 \log n$.*
- *has its midpoint at $\frac{1}{n}(m+i)$.*

Finally, the key lemma that relates the $\mathrm{SL}(2,\mathbb{Z})$ action to sums of squares is and it is what we will use to calculate λ-lengths in Section 4.

Lemma 2.4. *Let n be a positive integer. The number of ways of writing n as a sum of squares*

$$n = c^2 + d^2$$

with c, d coprime integers is equal to the number the integers $0 \leq k < n-1$ coprime to n such that the line

$$\{k/n + it,\ t \in \mathbb{R}\}$$

contains a point in the $\mathrm{SL}(2,\mathbb{Z})$ orbit of i.

Proof. Suppose there is such a point which we denote w. The point w is a fixed point of some element of order 2 in $\mathrm{SL}(2,\mathbb{Z})$. Since the Ford circles are $\mathrm{SL}(2,\mathbb{Z})$ invariant this element must permute F with the Ford circle tangent to the real line at the real part of w. So, in particular, w is the midpoint of the line that it lies on and by Lemma 2.3 one has:

$$\frac{1}{n} = \mathrm{Im}\,\frac{1}{n}(k+i) = \mathrm{Im}\,\frac{ai+b}{ci+d} = \frac{\mathrm{Im}\,i}{c^2+d^2}.$$

Conversely if c, d are coprime integers then there exists a, b such that

$$ad - bc = 1 \Rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}).$$

By applying a suitable power of the parabolic transformation $z \mapsto z+1$, one can choose w such that $0 \leq \mathrm{Re}\,w < 1$. So if $n = c^2 + d^2$ then $\frac{ai+b}{ci+d}$ is on one of the lines of the family in the statement.

□

2.0.1. *Calculating λ-lengths.* Let $a/c, b/d$ be a pair of distinct rationals. We define the *length* of the arc joining these rationals to be the length, with respect to the Poincaré metric on $\mathbb{H}$, of the portion outside of the Ford circles tangent at $a/c, b/d$. and the λ-length of an arc to be the exponential of this length. It is a consequence of Lemma 2.3 below that the arcs of λ-length 1 are the edges in the so-called *Farey diagram* (see Figure 5).

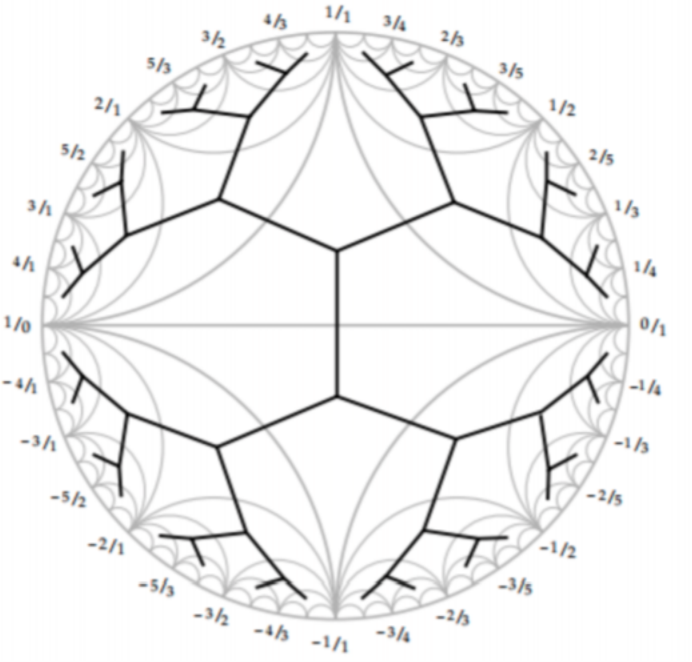

FIGURE 5. The Farey diagram.

We have:

Lemma 2.5. *Let $a/c, b/d$ be a pair of distinct extended rationals. Then the λ-length of the arc joining them is the square of the determinant of the matrix*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Further if $a/b = 1/0$ then the arc is a vertical line whose midpoint has imaginary part equal to $1/d$.

Proof. By transitivity of $P\mathrm{SL}(2,\mathbb{Z})$ on the extended rationals we may assume $a/c = 1/0$ and so the determinant of the matrix is d. The Ford circle at $1/0$ is $\{z, \operatorname{Im} z > 1\}$ and at b/d it is a circle of euclidean diameter $1/d^2$. The top of this circle is tangent to the line $\{z, \operatorname{Im} z = 1/d^2\}$ and the transformation $z \mapsto d^2 z$ maps this to the boundary of the Ford circle $\{z, \operatorname{Im} z \geq 1\}$. It follows from elementary hyperbolic geometry (the translation length of $z \mapsto d^2 z$ is $2 \log d$) that the distance between these sets, and so the length of the segment outside the Ford circles, is $2\log d$. Thus the λ-length is the square of the determinant as required.

□

3. Markoff cubic and the character variety

In this section we give a short modern exposition of the basis of H. Cohn's approach.

3.1. **Character Variety.** It is convenient to change variables and study solutions of

$$X^2+Y^2+Z^2-XYZ=0. \tag{3}$$

By the work of Fricke the set of solutions in positive real numbers can be identified with a certain slice of the *relative character variety of* $\mathbb{Z}*\mathbb{Z}$. This is the set of representations

$$\rho:\mathbb{Z}*\mathbb{Z}\to SL(2,\mathbb{R})$$

such that the trace of the image of the commutator of the generators is -2 up to conjugation. The key point in Fricke's work is that an (irreducible) representation ρ is determined up to conjugation by the three numbers

$$\begin{aligned} X &= tr\rho(\alpha), \\ Y &= tr\rho(\beta), \\ Z &= tr\rho(\alpha\beta)), \end{aligned}$$

where α,β are generators of $\mathbb{Z}*\mathbb{Z}$. Fricke calculates the trace of the commutator and shows that

$$2+tr\,(\alpha\beta\alpha^{-1}\beta^{-1})=X^2+Y^2+Z^2-XYZ. \tag{4}$$

The quotient surface $\mathbb{H}/\rho(\mathbb{Z}*\mathbb{Z})$ is invariably a once punctured torus and we identify $\mathbb{Z}*\mathbb{Z}$ with its fundamental group. The $\alpha\beta\alpha^{-1}\beta^{-1}$ is a loop around the puncture and the condition of the trace means that the monodromy around this loop is parabolic.

3.2. **More λ lengths.** There is an embedded cusp region H_2 of area 2 on every cusped hyperbolic surface $\mathbb{H}/\rho(\mathbb{Z}*\mathbb{Z})$ and on the modular torus $\mathbb{H}/\Gamma'$ there is a maximal cusp region H_6 of area 6 (see [11] for a discussion). After possibly replacing ρ by a conjugate representation we may assume $\rho(\mathbb{Z}*\mathbb{Z})$ that

$$\rho(\alpha\beta\alpha^{-1}\beta^{-1}):z\mapsto z+6,$$

it follows that H_2 lifts to the set $\hat{H}=\{\mathrm{Im}\,z>3\}$. Let $\alpha*$ be an*arc* that is a bicuspidal geodesic without self intersetions. There is a lift of $\alpha*$ to $\mathbb{H}$ which is a vertical line which evidently meets $\hat{H}$, we claim that any lift of of $\alpha*$ which meets $\hat{H}$ is a vertical line and not a semi circle.

For, if C is a semi circle that meets $\hat{H}$ its diameter is strictly greater than 6 and it follows that C and $C+6$ meet transversely in some point x. Such a point gives rise to a self intersection on the quotient surface It follows that, the portion of α^* outside of H is connected, and we define λ-length of α^* to be the exponential of the length of this sub arc.

Lemma 3.1. *Let $\alpha*$ be an arc on a once punctured torus and α the unique simple closed geodesic disjoint it. Then the square root of the λ-length of the arc α is equal to $\frac{2}{3}\cosh \ell_\alpha/2$.*

It is possible to prove this directly using hyperbolic trigonometry following the same schema as in [11] but here we give a more conceptual proof using the computations from [19] the interested reader should also consult [7].

Given an arc α^* one may extend it to an ideal triangulation of the punctured torus: that is there is a pair of arcs β^*, γ^*, each disjoint from α^* and the complement of $\alpha^* \cup \beta^* \cup \gamma^*$ a pair of ideal triangles. Let X denote $2\cosh \ell_\alpha/2$ where α is the unique closed simple geodesic disjoint from α and likewise

$$Y = 2\cosh \ell_\beta/2,\ Z = 2\cosh \ell_\gamma/2$$

where β resp γ is the unique closed simple geodesic disjoint from β^* resp. γ^*.

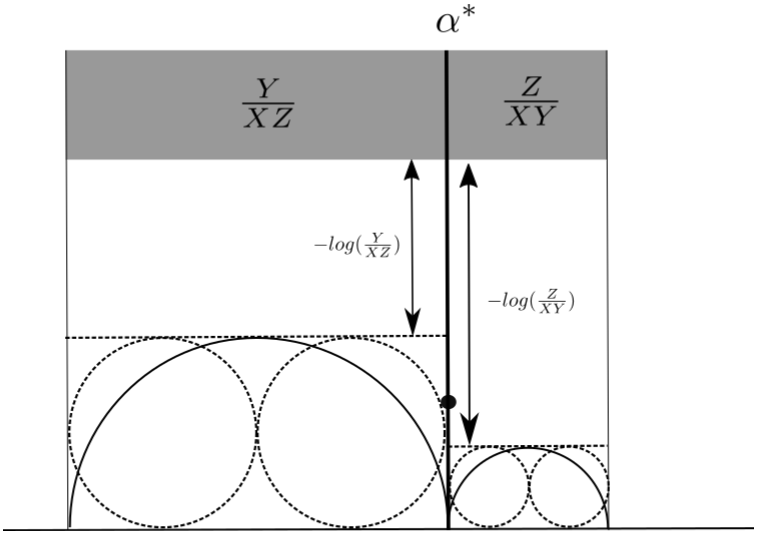

Figure 6. Calculating the hyperbolic length of α^* in the upper half plane the λ-length is the exponential of this. The mid point of α^* is marked by a circle and the two corners adjacent to α^* are shaded

In [19] Wolpert divides the Markoff cubic by XYZ to obtain

$$\frac{X}{YZ} + \frac{Y}{XZ} + \frac{Y}{XZ} = 1.$$

The three terms in this relation have a geometric interpretation which we will exploit to compute the λ-length of α^*. Let H denote the cusp region of area 2. A*corner* of an ideal triangle is one of the three components of its intersection with H. Every torus admits an *elliptic involution* which leaves each of the arcs of the ideal triangulation invariant and swaps the triangles. So, in fact, to each triangulation we can associate three numbers namely the areas of the corners of one of the ideal triangles and these coincide with Wolpert's three numbers.

Lifting the ideal triangulation to $\mathbb{H}$ as in Figure 6 one sees that α^* decomposes into two arcs of length $-\log(Y/XZ)$ and $-\log(Z/XY)$ respectively so that its is of length $2\log X$.

So, on any hyperbolic punctured torus, the λ-length of α^* wrt the cusp region of area 2 is the exponential of this, that is:

$$X^2.$$

Now on the modulaire torus $\mathbb{H}/\Gamma'$ there is an embedded cusp region of area 6 and the λ-length of α^* wrt this cusp region is

$$\frac{X^2}{9}.$$

3.3. **Sum of squares.** In the proof of Lemma 3.1 we used the fact that every torus admits an *elliptic involution* which leaves each of the arcs of the ideal triangulation invariant and swaps the triangles. For the modular torus the involution is covered by $z \mapsto -1/z$ and this means that for any arc α^* every lift contains a point of the $\mathrm{SL}(2,\mathbb{Z})$-orbit of i. In particular, by Lemma 3.1, a lift which is a vertical line ends at a rational which has as denominator a Markoff number and so this Markoff number is a sum of two squares. Conversely, every Markoff number arises as the square of a λ-length of some arc α^* and so must be the sum of two squares. By extending this reasoning slightly (compare [12] and [18]) we obtain:

Theorem 3.2. *The uniqueness conjecture is equivalent to:*
Let m be a Markoff number then exactly one of the vertical lines with endpoint k/m, where $1 \leq k \leq m-1$ is coprime to m, projects to an arc on the modular torus.

Proof. The Markoff triples form a binary tree with a preferred vertex corresponding to the fundamental triple $(1,1,1)$. Define the multiplicity of a Markoff number to be the number of triples for which it appears as the largest integer. One can easily check that for the so-called singular Markoff numbers 1 and 2 their multiplicity is 3 and, since group of automorphisms of the tree that fix the fundamental triple of order

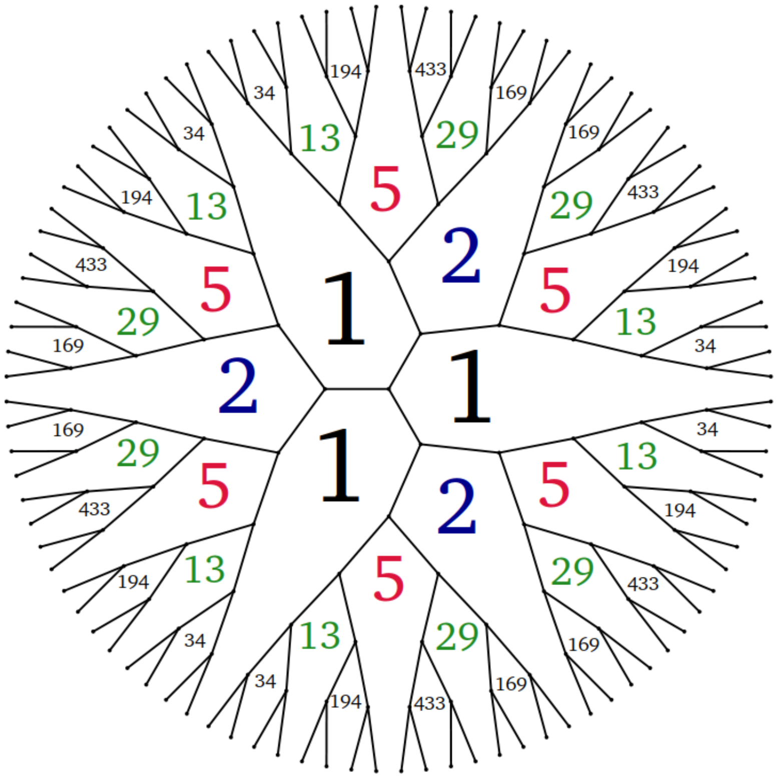

FIGURE 7. The tree of Markoff triples, the multiplicity is at most six.

6, the multiplicity of any other Markoff number is at least 6. Thus the uniqueness conjecture can be restated as: multiplicity of any other Markoff number is at most 6 (Figure 7).

Using Cohn's correspondence it follows that the uniqueness conjecture is equivalent to the number of oriented closed simple geodesics on the modular torus of any given length is at most 6. Each (unoriented) closed simple geodesic is disjoint from exactly one arc so that there can be at most three arcs of any given λ-length.

The group of orientation preserving automorphisms of the modular torus is canonically isomorphic to

$$\mathrm{SL}(2,\mathbb{Z})/\Gamma' \simeq \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\ ZZ \simeq \mathbb{Z}/6\mathbb{Z}.$$

The commutator of the generators of Γ' is $z \mapsto z+6$ and since each automorphism ϕ must leave the cusp invariant it lifts to a map of the form $\hat{\phi} : z \mapsto z + k$, $k = 0, \ldots 5$. Now consider the lift of some arc on the modular torus which, WLOG, is a vertical line. After applying (the lift of) an automorphism $\hat{\phi}$ we may assume it has its end point in $\mathbb{R}$

between 0 and 1. The statement now follows by counting multiplicities as before.

□

4. Button's theorem and its extensions

We are now ready to present a proof of Theorem 1.1.

4.1. **Proof of Button's theorem.**

Proof. : Suppose that $m = p^k$ is a Markoff number which forms a Markoff triple with x, y:

$$x^2 + y^2 + m^2 - xym = 0$$

so that

$$\bar{x}^2 + \bar{y}^2 = 0 \in \mathbb{F}_p.$$

It follows that p is either 2 or 1 mod 4 and so by Fermat's theorem there are coprime positive integers c, d, unique up to permutation, so that

$$p = c^2 + d^2 = (c + id)(c - id).$$

It is well known that the RHS is the unique factorisation of p in the Gaussian integers and it follows that the unique factorisation of m is

$$p^k = (c + id)^k (c - id)^k.$$

A consequence of this is that the pair of coprime positive integers a, b such that $p^k = a^2 + b^2$ is unique up to permutation. Explicitly we have:

$$a = \mathrm{Re}\,(c \pm id)^k \tag{5}$$

$$b = \mathrm{Im}\,(c \pm id)^k. \tag{6}$$

Since a, b are unique up to permutation then, by Lemma 2.4, there is only a single geodesic of the family of vertical lines ending at $k/mi, 0 \leq k \leq m$ which meets the $\mathrm{SL}(2, \mathbb{Z})$-orbit of i. The result then follows from the Paragraph 3.3.

Now suppose that $m = 2p^k$ is a Markoff number. By the above p^k can be written as a sum of squares $c^2 + d^2 = |c + id|^2$ essentially uniquely. Observe that 2 factors as

$$2 = i(1 + i)^2.$$

Observe that $2p^k$ can also be written as a sum of squares essentially uniquely namely

$$2p^k = |(1 + i)(c + id)|^2 = (c - d)^2 + (c + d)^2,$$

so that the result follows in this case to.

□

4.2. **Geodesics on pants.** As we said in the introduction made our proof of Button's theorem possible was that each closed simple geodesic α on $\mathbb{H}/\Gamma'$ is paired with a unique simple bicuspidal geodesic α^*. Using the fact that this bicuspidal geodesic passes through a single Weierstrass point and our reformulation (Theorem 3.3) of the unicity conjecture we concluded. On closer examination one sees that there is no real need for the geodesic to be simple merely that:

(1) it is disjoint from α;
(2) it passes through the Weierstrass point and has λ-length a prime number that can be written as a sum of two squares $a^2 + b^2$.

Since $a, b > 0$ are unique up to permutation then, by Lemma 2.4, there is a single geodesic of the family of vertical lines ending at k/p which meets the $\mathrm{SL}(2,\mathbb{Z})$-orbit of i. We now show how to compute the λ-length of any such geodesic which finishes the proof of Theorem 1.3.

By cutting along the closed geodesic α (see [11]) we obtain a (degenerate) pair of pants with a single cusp and boundary consisting of a pair of closed geodesics of lengths $\ell(\alpha)$. Any curve disjoint from α, for example α^*, survives this cutting process and we can compute its length on this pair of pants. Since we are only interested in λ-lengths it is more convenient to work with a slightly different pair of pants which has the same set of λ-lengths but which is uniformised by a particularly nice Fuchsian group G_m. This coincidence is due to the fact that the pants obtained by cutting along α and this new pair of pants have a common (double) cover Figure 9 covers of a certain orbifold.

Let

$$P = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

so that

$$Q = SPS^{-1} = \begin{pmatrix} 1 & 0 \\ -m & 1 \end{pmatrix}.$$

Let $G_m < \mathrm{SL}(2,\mathbb{Z})$ denote the group generated by P, Q. One has

$$QP = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -m & 1 \end{pmatrix} = \begin{pmatrix} 1 - m^2 & m \\ -m & 1 \end{pmatrix}$$

which admits a square root in $P\mathrm{SL}(2,\mathbb{Z})$ that is:

$$C^2 = \begin{pmatrix} m & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} m^2 - 1 & -m \\ m & 1 \end{pmatrix}.$$

The matrices

$$C = \begin{pmatrix} m & -1 \\ 1 & 0 \end{pmatrix}, C' = \begin{pmatrix} 0 & 1 \\ -1 & -m \end{pmatrix}$$

generate a free group $G_m^* < P\mathrm{SL}(2,\mathbb{Z})$ and

$$CC' = \begin{pmatrix} 1 & 2m \\ 0 & 1 \end{pmatrix} = P^2.$$

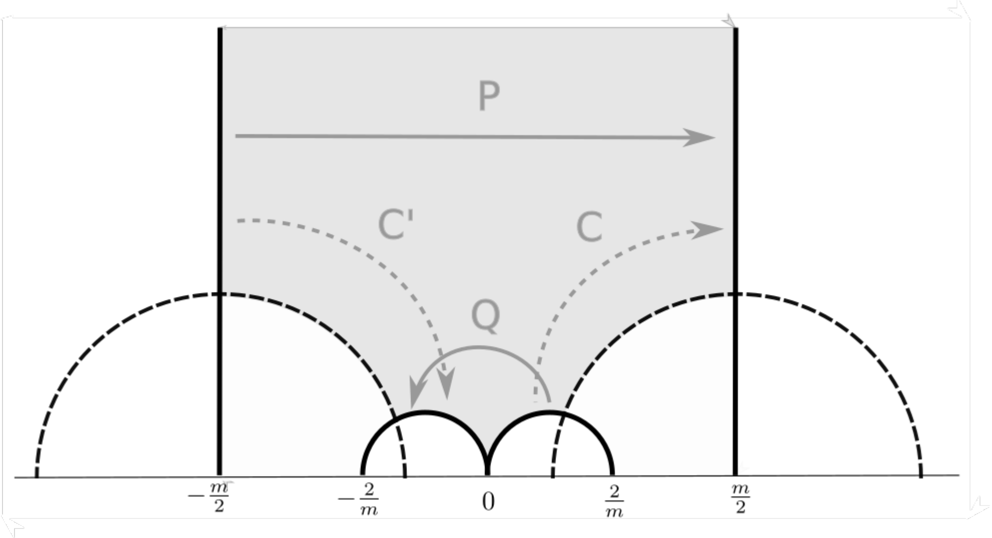

FIGURE 8. A fundamental domain for the group generated by P, Q and C, C' with the side pairings indicated by solid arrows for P, Q and dashed arrows for C, C'.

Lemma 4.1. *The groups G_m and G^*_m have a common fundamental domain (see Figure 8).*

- *The quotient $\mathbb{H}/G_m$ is a surface with two cusps and a single geodesic boundary component of length $2\ell(\alpha)$.*
- *The quotient $\mathbb{H}/G_m*$ is a (degenerate) pair of pants with a single cusp and boundary consisting of a pair of closed geodesics each of lengths $\ell(\alpha)$.*
- *The orbit of i under G_m and G^*_m is the same.*

Proof. The lemma follows from the matrix calculations above and studying Figure 8. □

Lemma 4.2. *For every integer k there is a bicuspidal geodesic on $\mathbb{H}/G^*_m$, and so on $\mathbb{H}/G_m$, which has λ-length k^2m^2+1, and is invariant under the involution induced by S.*

Proof. Consider the image of i under the sequence of elements of G_m given by:

$$Q^kP = \begin{pmatrix}1 & m\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ -km & 1\end{pmatrix} = \begin{pmatrix}1-km^2 & m\\ -km & 1\end{pmatrix}$$

By Lemma 2.4 the imaginary part of $Q^kP.i$ is given by

$$\frac{1}{k^2m^4+1}$$

So there is a half infinite geodesic in $\mathbb{H}$ joining the ideal point ∞ to $Q^kP.i$. The projection of this segment to $\mathbb{H}/G_m$ together with its

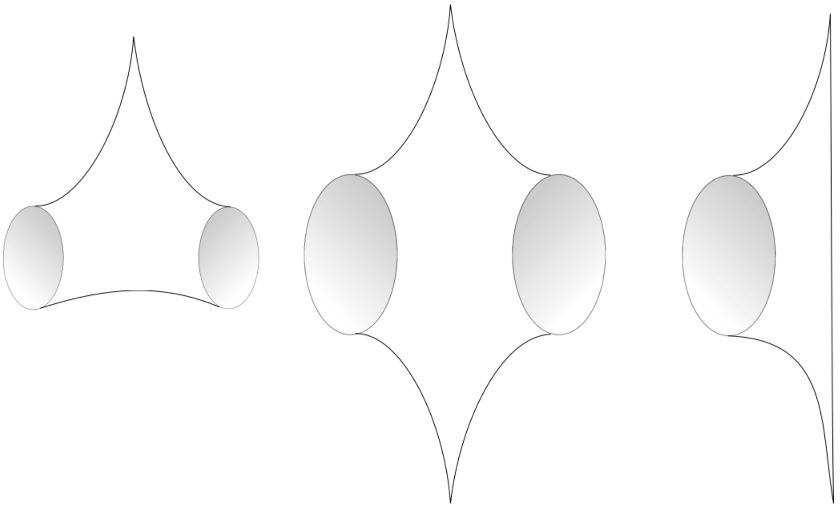

FIGURE 9. On the left the pair of pants $\mathbb{H}/G_m$, on the right the pair of pants $\mathbb{H}/G_m^*$, and in the middle the surface $\mathbb{H}/(G_m \cap G_m^*)$.

image under the involution form a bicuspidal geodesic of the required λ-length.

□

4.3. **Afterword: confusing coefficients.** Both G_m and G_m^* are subgroups of the principal congruence subgroup $\Gamma(n)$ and so the off diagonal coefficients of elements are always divisible by m. The modular group acts and Γ' act transitively on the rationals since the corresponding quotient orbifold and surface respectively have a single cusp. The pair of pants $\mathbb{H}/G_m^*$ has a pair of cusps, each cusp corresponding to the orbit of ∞ or 0 under G_m^* and so there are three distinct types of bicuspidal geodesic:

(1) type 1 has a lift to $\mathbb{H}$ with ideal endpoints ∞ and some other point of the orbit of ∞;
(2) type 2 has a with ideal endpoints 0 and some other point of the orbit of 0;
(3) type 3 has a with ideal endpoints ∞ and some point of the orbit of 0.

If $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G_m^*$ then the the image of ∞ by this element is a/b. It follows that the terminal Ford circles are F and another one tangent at a/b of radius c^2. One sees in this way that the λ-length of the corresponding bicuspidal geodesic which is of type 1 above is always divisible by m. A similar argument shows that the λ-length of any type 2 geodesic is divisible by m too. This curious observation led us to believe falsely that no useful information could be extracted from the bicuspidal geodesics. Fortunately, the third type of geodesic is not divisible by m, and is always congruent to 1 modulo m.

References

[1] M. Aigner *Markov's Theorem and 100 Years of the Uniqueness Conjecture*, Springer(2013)

[2] A. Baragar, *On the Unicity Conjecture for Markoff Numbers* Canadian Mathematical Bulletin , Volume 39 , Issue 1 , 01 March 1996 , pp. 3 - 9

[3] V. Bouniakowsky, *Sur les diviseurs numériques invariables des fonctions rationnelles entières,* Mem. Acad. Sci. St. Peteresbourg, 6e série, vol. VI (1857), 305–329.1
Available at: `https://books.google.fr/books?hl=fr&id=wXIhAQAAMAAJ&pg=PA305#v=onepage&q&f=false`

[4] J. O. Button, *The uniqueness of the prime Markoff numbers*, J. London Math. Soc. (2) 58 (1998), 9–17.

[5] L. Euler, letter to Goldbach, 28th October 1752 (letter CXLIX), available at `http://eulerarchive.maa.org/correspondence/letters/OO0877.pdf`

[6] Lester R. Ford *Automorphic Functions* Chelsea Publishing Company, 1951

[7] Yi Huang *Moduli Spaces of Surfacese* University of Melbourne Doctoral Thesis 2014 `https://www.yihuang.site/pdfs/yi%20huang_moduli%20spaces%20of%20surfaces%20(phd%20thesis).pdf`

[8] G. Frobenius *Über die Markoffschen Zahlen* S. B. Preuss Akad. Wiss., Berlin (1913) pp. 458–487

[9] Gareth A. Jones, Alexander K. Zvonkin *Block designs and prime values of polynomials* `https://arxiv.org/abs/2105.03915`

[10] G. McShane, V. Sergesciu *Geometry of fermat's sum of squares*, 2021 `https://macbuse.github.io/squares.pdf`

[11] G. McShane, *Simple geodesics and a series constant over Teichmuller space* Invent. Math. (1998)

[12] Greg McShane, Hugo Parlier, *Multiplicities of simple closed geodesics and hypersurfaces in Teichmüller space,* Geom. Topol. Volume 12, Number 4 (2008), 1883-1919.

[13] M.L. Lang, S.P Tan, *A simple proof of the Markoff conjecture for prime powers* Geometriae Dedicata volume 129, pages15–22 (2007)

[14] R. C. Penner, *The decorated Teichmueller space of punctured surfaces*, Communications in Mathematical Physics 113 (1987), 299–339.

[15] Paul Pollack *Hypothesis H and an impossibility theorem of Ram Murty* Rendiconti del Seminario Matematico 68(2) 2010

[16] Paul Schmutz Schaller. *Geometry of Riemann surfaces based on closed geodesics.* Bull. Amer. Math. Soc. (N.S.), 1998.

[17] J-P. Serre, *A Course in Arithmetic*, Graduate Texts in Mathematics, Springer-Verlag New York 1973

[18] Boris Springborn, *The hyperbolic geometry of Markov's theorem on Diophantine approximation and quadratic forms.* Enseign. Math. 63 (2017)

[19] Scott Wolpert, *On the Kahler form of the moduli space of once-punctured tori*, Comment. Math. Helv. 58(1983)246-256

[20] D. Zagier, *A one-sentence proof that every prime $p = 1$ (mod 4) is a sum of two squares*, American Mathematical Monthly, 97 (2): 144

[21] Y. Zhang, *An elementary proof of uniqueness of Markoff numbers* preprint, arXiv:math.NT/0606283

[22] Y. Zhang, *Congruence and uniqueness of certain Markoff numbers* Acta Arithmetica, Volume: 128, Issue: 3, page 295-301

[23] Tim Pope Copilot for vim `https://github.com/github/copilot.vim`

Institut Fourier 100 rue des maths, BP 74, 38402 Gieres. France
Email address: `mcshane at univ-grenoble-alpes.fr`