Comment. Math. Helv. 99 (2024), 111–148
DOI 10.4171/CMH/566

Non-planarity of Markoff graphs mod p

Matthew de Courcy-Ireland

Abstract. We prove the non-planarity of a family of 3-regular graphs constructed from the solutions to the Markoff equation $x^2 + y^2 + z^2 = xyz$ modulo prime numbers greater than 7. The proof uses Euler characteristic and an enumeration of the short cycles in these graphs. Non-planarity for large primes would follow assuming a spectral gap, which was the original motivation. For primes congruent to 1 modulo 4, or congruent to 1, 2, or 4 modulo 7, explicit constructions give an alternate proof of non-planarity.

1. Introduction

For each prime number p, we consider a graph whose vertices are triples in $\mathbb{F}_p^3$, with edges connecting a vertex (x, y, z) to

$$\begin{aligned} m_1(x, y, z) &= (yz - x, y, z), \\ m_2(x, y, z) &= (x, xz - y, z), \\ m_3(x, y, z) &= (x, y, xy - z). \end{aligned}$$

The operations m_1, m_2, m_3 preserve the polynomial $x^2 + y^2 + z^2 - xyz$. Thus the graph is a disjoint union of subgraphs corresponding to solutions of a Markoff-type equation

$$x^2 + y^2 + z^2 = xyz + k$$

with $k \in \mathbb{F}_p$. An especially interesting case is $k = 0$, which Markoff investigated (over $\mathbb{Z}$ rather than $\mathbb{F}_p$) and found to be related to quadratic forms and Diophantine approximation [23]. By "the Markoff graph mod p", we mean the graph with vertices $(x, y, z) \neq (0, 0, 0)$ satisfying $x^2 + y^2 + z^2 = xyz$ in $\mathbb{F}_p$, and edges given by m_1, m_2, m_3 as above. For example, Figure 1.1 shows the Markoff graph mod 7.

Mathematics Subject Classification 2020: 11D25 (primary); 05C10, 05C50, 11F72, 37P25 (secondary).

Keywords: Markoff triples, expander graphs, planar graphs, graph embeddings, cubic surfaces, Euler characteristic, totient function.

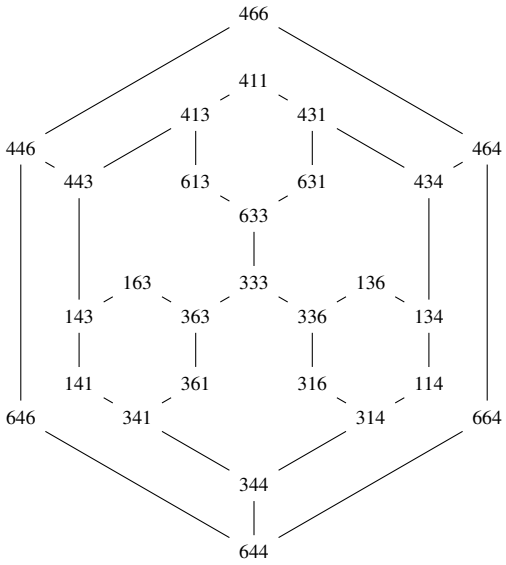


Figure 1.1. The Markoff graph mod 7 is planar. The vertices are the 28 solutions to $x^2 + y^2 + z^2 = xyz$ mod 7, excluding $(0, 0, 0)$. The labels abbreviate (x, y, z) by xyz, with edges corresponding to the moves $x \mapsto yz - x$, $y \mapsto xz - y$, and $z \mapsto xy - z$. To obtain a 3-regular graph, small loops can be drawn at the vertices of degree 2 without crossing any other edges. These vertices are fixed by a move on one of the coordinates, as for instance $(1, 6, 3)$ is fixed by changing 3 to $1 \times 6 - 3 = 3$.

Theorem 1.1. *The Markoff graph* mod p *is planar if and only if the prime* p *is* 2, 3, *or* 7.

In other words, for $p \neq 2, 3, 7$, these graphs cannot be drawn in the plane without some edges crossing. This is an indirect test of the hypothesis that the Markoff graphs form an expander family as $p \to \infty$. Indeed, by the planar separator theorem of Lipton and Tarjan [22], expansion is impossible in planar graphs. As a proof of expansion in the Markoff family remains elusive, we became interested in finding a direct proof that they are not planar. We recall this connection in Section 12. We refer to [20] for more on the spectral properties of planar graphs.

The intuition behind the proof is that a planar graph cannot have too many edges. The following folklore lemma can be shown using Euler characteristic, as we review in Section 3.

Lemma 1.2. *If a planar connected graph has* V *vertices,* E *edges, and no cycles of length less than* g*, then*

$$E \leq \frac{g}{g-2}(V - 2). \tag{1.1}$$

For a graph with 3 edges at every vertex and no self-edges, it must be that $E = 3V/2$. If there are no cycles of length less than $g = 6$, then equation (1.1) is absurd:

$$\frac{3V}{2} = E \leq \frac{6}{6-2}(V-2) < \frac{3V}{2}$$

from the strict inequality $V - 2 < V$. This shows that a finite 3-regular graph of girth 6 cannot be planar. For comparison, there is an infinite 3-regular graph of girth 6, given by tiling the plane with hexagons. Attempts to truncate this infinite graph must introduce either crossings between edges, or cycles of length less than 6, or vertices of degree different from 3.

The proof of Theorem 1.1 applies the same logic to the Markoff graphs mod p. The number of edges is not quite $3V/2$ because of a small number of self-edges whenever (x, y, z) is fixed by one of the Markoff moves m_1, m_2, m_3. This occurs for instance at $(1, 6, 3)$ in the Markoff graph mod 7, with $3 = 6 \times 1 - 3$. Moreover, there can be cycles of length shorter than 6. We will see that these are rare, and an approximate version of (1.1) still yields a contradiction for sufficiently large p provided we take $g = 7$ rather than $g = 6$, to compensate for these self-edges and short cycles. The inequalities leave only two cases unsettled, $p = 11$ and $p = 13$, whose non-planarity can be shown directly to complete the proof.

The same approach gives a bound on the Euler characteristic that would be needed for a surface to accommodate the Markoff graph mod p. To make sense of the statement, recall that the Euler characteristic of a surface is typically negative.

Theorem 1.3. *If χ is the Euler characteristic of a surface in which the Markoff graph* mod p *can be embedded, then as $p \to \infty$,*

$$\Big(\frac{1}{2} - o(1)\Big)p^2 \leq -\chi.$$

Given χ, Theorem 1.3 shows that there are only finitely many primes for which the Markoff graph mod p can be embedded in a surface of that Euler characteristic. Indeed, such an embedding is impossible for $p > (1 + o(1))\sqrt{2|\chi|}$, although our estimates on the term $o(1)$ are somewhat impractical. Theorem 1.1 is a more precise statement of this form with $\chi = 2$ for the planar case. The order of magnitude p^2 in Theorem 1.3 is correct: drawing each edge on a handle of its own gives an embedding in a surface with $-\chi = (3 + o(1))p^2$.

The exceptions $p = 2, 3, 7$ in Theorem 1.1 give the smallest Markoff graphs. The number of vertices in the Markoff graph for an odd prime p is $p^2 + 3p(-1)^{(p-1)/2}$, by a formula of Carlitz [7], which we review in Lemma 2.1. In particular, there are 28 vertices for $p = 7$ compared to 40 for $p = 5$. For $p = 3$, there are no solutions to $x^2 + y^2 + z^2 = xyz$ besides $(0, 0, 0)$, connected to itself by all three moves m_1, m_2, m_3, so the Markoff graph mod 3 is empty. However, one can obtain a more interesting

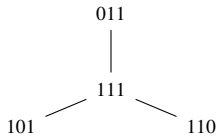


Figure 1.2. The Markoff graph for $p = 2$, with $(0, 1, 1)$ fixed by the two moves sending either coordinate 1 to $0 \times 1 - 1 = 1 \bmod 2$.

example mod 3 from the rescaling $x^2 + y^2 + z^2 = 3xyz$, as we describe in the conclusion. For $p = 2$, there are four non-zero solutions, namely $(1, 1, 1)$ connected to the permutations of $(0, 1, 1)$ with a pair of self-edges at each of the latter (Figure 1.2). These more singular examples can be drawn in the plane, but it is natural to exclude them and think of the Markoff graph mod 7 as the only non-trivial planar example.

A famous theorem of Wagner and Kuratowski [21, 28] gives another approach to non-planarity, which is useful for the finite number of primes that remain after the main strategy is executed. Their theorem characterizes planar graphs in terms of the minimal obstructions: a graph is planar if and only if it does not contain any copies of the complete bipartite graph $K_{3,3}$ or the complete graph K_5 (with different notions of "copy" in the exact formulations of Wagner and Kuratowski, as we review below). For example, we can prove the following theorems by finding explicit copies of $K_{3,3}$ inside the Markoff graph for certain primes p.

Theorem 1.4. *The Markoff graph* mod p *is not planar for any prime number congruent to* 1 mod 4.

Theorem 1.5. *If* -7 *is a non-zero quadratic residue modulo* p*, then the Markoff graph* mod p *is not planar.*

Notably, Theorem 1.5 does not apply when $p = 7$, and the Markoff graph is planar in that case. By quadratic reciprocity, -7 is a square modulo p if and only if p is a square modulo 7, that is, p is 1, 2, or 4 modulo 7. Together, Theorems 1.4 and 1.5 have the following corollary, which combines the conditions modulo 4 and modulo 7 into different possibilities modulo 28.

Corollary 1.6. *The Markoff graph* mod p *is not planar for any odd prime* $p \neq 7$*, except possibly for* $p \equiv 3, 19, 27 \bmod 28$.

In terms of density, these constructions show that the Markoff graphs are non-planar for at least a fraction $3/4$ of primes. We will use them especially to show non-planarity for $p = 11$ and $p = 13$, which are the last cases remaining in the proof of Theorem 1.1 after non-planarity for large p has been achieved by the strategy of Section 3.

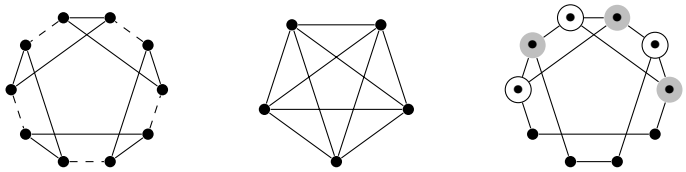


Figure 1.3. The 3-regular graph on the left contains a copy of K_5 as a graph minor, obtained by contracting the dashed edges. A subdivision of $K_{3,3}$ for the same graph is shown at right.

The method of proof is to find a copy of the complete bipartite graph on 3 pairs of vertices. The example for Theorem 1.4 uses special solutions available only when -1 has a square root modulo p, in particular the lines contained in the Markoff cubic surface, while Theorem 1.5 requires a square root of -7. These configurations are drawn in Figures 10.1 and 11.1. An interesting difference is that Theorem 1.5 is local in nature: it involves paths of bounded length, whereas Theorem 1.4 involves paths of length growing with p.

There is a subtle difference between the formulations of Wagner and Kuratowski, even though both lead to equivalent characterizations of planarity. In Kuratowski's theorem, a "copy" is simply a subdivision of $K_{3,3}$, where each edge of $K_{3,3}$ is given by a path between its endpoints in the graph of interest. The Markoff graphs are 3-regular, so that K_5 cannot occur as a subdivision. This differs from Wagner's formulation, where a "copy" refers to a graph minor. To show a graph is non-planar using Wagner's theorem, $K_{3,3}$ or K_5 may be formed by contracting edges, as well as deleting edges or isolated vertices. Contracting an edge removes it and merges its endpoints into a single vertex, which allows K_5 to occur as a minor even for graphs with only 3 edges incident to each vertex. This is illustrated in Figure 1.3. Looking for copies of K_5 might allow more flexibility in proving non-planarity, but one knows from Kuratowski's theorem that there must also be a subdivision of $K_{3,3}$ whenever K_5 occurs as a minor. In this sense, $K_{3,3}$ is the only obstruction to planarity for the Markoff graphs.

We exclude the point $(0, 0, 0)$ from the Markoff graph mod p because it is fixed by all of m_1, m_2, m_3. The rest of the level set $x^2 + y^2 + z^2 - xyz = 0$ seems to form a connected graph. This was conjectured in Baragar's thesis [4] and connectedness is now known for all sufficiently large primes p. Bourgain–Gamburd–Sarnak [6] have been able to prove connectedness for many primes p by a method that succeeds unless $p^2 - 1$ has an unusually large number of factors. In any case, their method shows that there is a "giant component": for any $\varepsilon > 0$, once p is large enough depending on ε, the Markoff graph has a connected component containing all but $O(p^\varepsilon)$ vertices. On the other hand, Chen [10] has shown that all connected components have size divisible by p. As a result, the giant component must coincide with the entire graph once p is large enough. An explicit threshold for how large p should be has been determined

by Eddy, Fuchs, Litman, Martin, and Tripeny [15]: $p > 3.448 \times 10^{392}$. We also refer to [17] for further analysis of the Markoff graphs, and cryptographic applications.

Numerical evidence obtained in [12] suggests that the Markoff graphs are not only connected, but moreover form an expander family as $p \to \infty$. This seems to demand new techniques beyond what is involved in proving connectedness. Non-planarity is a simple consequence of expansion that can be established more easily. This in turn provides some indirect evidence in favour of expansion.

The rest of this article pursues these ideas in the following sequence. In Section 2, we review how many vertices and edges are in the Markoff graph mod p, and some other basic parameters. In Section 3, we outline the strategy leading to Theorem 1.1, recall the proof of Lemma 1.2, and prove Theorem 1.3. Sections 4, 5, 6, and 7 complete the proof of Theorem 1.1 by determining the fixed points of some short words in the Markoff moves m_1, m_2, m_3. In Section 8, we show that even in the hypothetical cases where the Markoff graph is not connected, the foregoing arguments show non-planarity of the giant component of Bourgain–Gamburd–Sarnak. This relies on a lower bound for Euler's function $\phi(n)$, detailed in Section 9.

In Section 10, we prove Theorem 1.4 on non-planarity for primes congruent to 1 mod 4, which takes advantage of lines contained in the Markoff cubic surface. Section 11 proves Theorem 1.5, which applies to some primes congruent to 3 mod 4 but not all. Section 12 reviews the Lipton–Tarjan theorem and its consequence that expansion cannot occur in planar graphs, which was our motivation for investigating the question of planarity. We give a simple calculation that, based on the level of expansion observed numerically, estimates how large p must be for this method to imply non-planarity. We conclude with some examples in Sections 13 and 14, drawing the Markoff graphs for $p = 5$ and 11, with an alternative scaling for $p = 3$.

We recommend [1] as an excellent account of the Markoff surface, and cite just a few examples of recent work in addition to [6, 10, 17] already discussed above. The permutations generated by m_1, m_2, m_3 on solutions mod p have been studied in [9, 25]. Over $\mathbb{Z}$, see [24] for recent work on the fractals introduced by Markoff in Diophantine approximation, [3] for generalizations to modular billiards, and [26] for connections with hyperbolic geometry.

2. Some key counts

In this section, we record some of the fundamental counts to do with the Markoff graph mod p. How many vertices? edges? short cycles? Recall that the vertices of the graph are triples $(x, y, z) \neq (0, 0, 0)$ satisfying $x^2 + y^2 + z^2 = xyz \bmod p$. Some of the counting is best thought of more generally for surfaces of the form $x^2 + y^2 + z^2 = xyz + k$, the case $k = 0$ being somewhat degenerate.

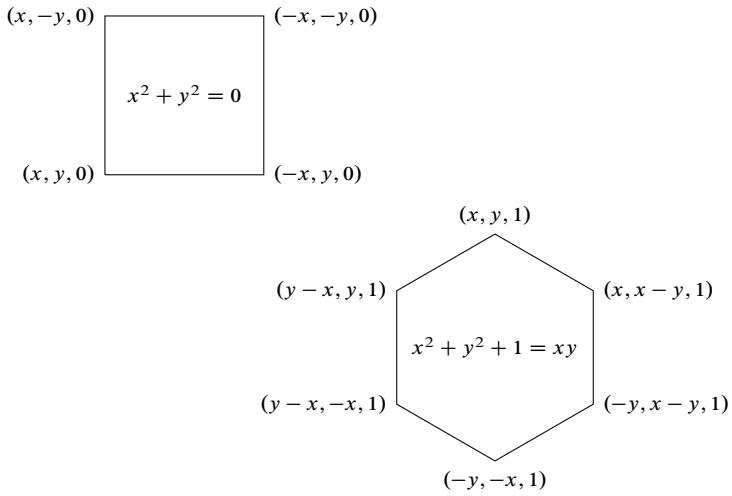


Figure 2.1. Top: the cycles of length 4 from part (3) of Lemma 2.1, which arise only when -1 is a square. Bottom: a cycle of length 6 from parts (5)–(6) of Lemma 2.1.

Lemma 2.1. *Consider a prime number $p \geq 5$.*

(1) *The number of vertices in the Markoff graph* mod p *is p^2+3p if $p \equiv 1 \bmod 4$, or p^2-3p if $p \equiv 3 \bmod 4$.*

(2) *The Markoff graph* mod p *is 3-regular, except for $3(p-3)$ vertices if $p \equiv 3 \bmod 4$ or $3(p-5)$ vertices if $p \equiv 1 \bmod 4$, which each have two neighbours and a single self-edge.*

(3) *If $p \equiv 1 \bmod 4$, then the cycles of length 4 in the Markoff graph* mod p *are of the form shown in Figure 2.1 with $z = 0$, or similarly with $x = 0$ or $y = 0$. In total, there are $3(p-1)/2$ cycles of length 4. Ignoring self-edges, there are no shorter cycles.*

(4) *If $p \equiv 3 \bmod 4$, then the shortest cycles are of length 6, ignoring self-edges.*

(5) *If $p \equiv 1 \bmod 3$, then there are $p-3$ cycles of length 6.*

(6) *If $p \equiv 2 \bmod 3$, then there are $p+1$ cycles of length 6. They are of the form shown in Figure 2.1 with $z = 1$, or similarly with any of the three coordinates equal to ± 1.*

Proof of Lemma 2.1. Part (1) is due to Carlitz [7], and part (2) to Cerbu–Gunther–Magee–Peilen [9, Lemma 2.3]. We review the arguments in Propositions 2.2 and 2.3 below, especially to confirm for part (2) that no point has multiple self-edges except $(0,0,0)$. The enumeration of short cycles is the novel aspect of Lemma 2.1. It involves

two steps: counting the number of squares and hexagons of the form described above, which we do using Proposition 2.2 in this Section; and determining whether there are any other short cycles, which we postpone to Sections 4, 5, and 6. Parts (3) and (4) follow from the enumeration in Section 4 and Corollary 7.2. Propositions 5.1 and 6.1 complete the proof of (5) and (6).

The congruences in Lemma 2.1 arise in deciding whether -1 and -3 have square roots modulo p, by quadratic reciprocity. This determines the number of solutions to the Markoff equation with $z = 0$ or $z = 1$, and hence the number of squares or hexagons of the form above. The most subtle case is when $p \equiv 1 \bmod 3$, where the fact that -3 is a quadratic residue complicates matters. We must discard solutions of the form $(x, 2x, 1)$ because the resulting cycles involve self-edges:

$$x \mapsto (2x) \cdot 1 - x = x.$$

These occur when $x^2 = -1/3$, since the Markoff equation with $y = 2x$ and $z = 1$ becomes $x^2 + 4x^2 + 1 = 2x^2$. We must discard all six of the triples

$$\left(\frac{1}{\sqrt{-3}}, \frac{2}{\sqrt{-3}}, 1\right), \quad \left(\frac{1}{\sqrt{-3}}, -\frac{1}{\sqrt{-3}}, 1\right), \quad \left(-\frac{2}{\sqrt{-3}}, -\frac{1}{\sqrt{-3}}, 1\right)$$
$$\left(-\frac{1}{\sqrt{-3}}, -\frac{2}{\sqrt{-3}}, 1\right), \quad \left(-\frac{1}{\sqrt{-3}}, \frac{1}{\sqrt{-3}}, 1\right), \quad \left(\frac{2}{\sqrt{-3}}, \frac{1}{\sqrt{-3}}, 1\right)$$

leaving only $p - 7$ solutions for $z = 1$ instead of the $p - 1$ from Proposition 2.2. These form $(p - 7)/6$ hexagons, for a total of $p - 7$ from all six level sets $x, y, z = \pm 1$. To compensate for the loss, there are four additional cycles involving $\sqrt{-3}$, as we describe in Section 6. This gives the final tally $p - 3$. ∎

In part (2), the fixed points of m_1, m_2, m_3 are the self-edges in the Markoff graph mod p, which we prefer to delete. One could also think of a self-edge as bounding a face, without changing the Euler characteristic $V - E + F$ since both E and F increase by 1. These can be drawn as small loops avoiding the other edges, so there is no difference for purposes of planarity.

The number of edges in a connected component with V vertices, after deleting self-edges, satisfies

$$E \geq \frac{3}{2}V - \frac{3}{2}\left(p - 4 - (-1)^{(p-1)/2}\right).$$

Equality holds if the component contains all the points with self-edges, for instance if the Markoff graph itself is connected.

To count the number of squares with $z = 0$, or hexagons with $z = 1$, we use the following proposition going back to Carlitz [7]. See also [6, Lemmas 3–5]. We give a proof for the sake of having all the necessary tools at hand in a common notation.

Throughout, $(\frac{\cdot}{p})$ denotes the Legendre symbol, with value 1 for (non-zero) quadratic residues mod p, -1 for non-residues, and 0 for 0.

Proposition 2.2. *Given z, the number of solutions (x, y) to*

$$x^2 + y^2 + z^2 = xyz + k$$

is as follows. If $z^2 \neq 4$ and $z^2 \neq k$, then the number of solutions over $\mathbb{F}_p$ is

$$p - \Big(\frac{z^2 - 4}{p}\Big). \tag{2.1}$$

If $z^2 = 4$, then the number is

$$\Big(1 + \Big(\frac{k-4}{p}\Big)\Big)p, \tag{2.2}$$

which is either 0, p, or $2p$. If $z^2 = k \neq 4$, then the number is

$$p + \Big(\frac{k-4}{p}\Big)(p-1), \tag{2.3}$$

either 1 or $2p - 1$. The total number of solutions (x, y, z) is

$$p^2 + \Big(\frac{k-4}{p}\Big)\Big(3 + \Big(\frac{k}{p}\Big)\Big)p + 1. \tag{2.4}$$

This count includes $(0, 0, 0)$ when $k = 0$, which leaves $p^2 \pm 3p$ vertices in the graph.

The special cases $z^2 = 4$ and $z^2 = k$ correspond to lines contained in the Markoff cubic surface:

$$z^2 = 4 \implies x = \pm y \pm \sqrt{k-4}, \quad z^2 = k \implies x = \left(\sqrt{\frac{k}{4}} + \sqrt{\frac{k-4}{4}}\right)y.$$

The conic associated to each of the four level sets $z = \pm 2, \pm\sqrt{k}$ is a pair of lines. Setting x, y, or z equal to any of these levels, we obtain up to 24 of the famous 27 lines on a cubic surface [8]. The remaining 3 lines on the Markoff surface (over an algebraic extension) are "at infinity" in projective space. Depending on whether k and $k - 4$ are quadratic residues, these lines might only become visible in an extension of $\mathbb{F}_p$. Moreover, some of the lines coalesce in the singular cases $k = 0$ and $k = 4$. For $k = 0$, a basic difference between 3 mod 4 and 1 mod 4 is that none of the lines are defined over the ground field if $p \equiv 3 \bmod 4$. In case $p \equiv 1 \bmod 4$, Figure 10.1 shows how to deduce non-planarity from the arrangement of lines in the Markoff surface.

Before turning to the proof of Proposition 2.2, it is worth noting an interpretation of the quadratic symbol in (2.1). The change of variable $z = \zeta + \zeta^{-1}$ has (one-to-two) inverse

$$\zeta = \frac{z \pm \sqrt{z^2 - 4}}{2},$$

so it is the quadratic status of $z^2 - 4$ that determines whether ζ lies in $\mathbb{F}_p$ or an extension. This change of variable plays a decisive role in the analysis of [6], as we will see in Sections 7 and 8. Conceptually, if z is the trace of a matrix in SL_2, then ζ and ζ^{-1} are the eigenvalues.

Proof of Proposition 2.2. We fix z and sum the number of solutions y for each x. There are 0, 1, or 2 solutions according to the radical that arises in solving the Markoff equation for y:

$$y = \frac{xz \pm \sqrt{x^2z^2 - 4(x^2 + z^2 - k)}}{2}.$$

The number of solutions (x, y) is then

$$\sum_x \left(1 + \left(\frac{x^2(z^2 - 4) - 4(z^2 - k)}{p}\right)\right).$$

If $z^2 = 4$, then the summand does not depend on x, and one obtains (2.2) since $-4(z^2 - k) = 4(k - 4)$ differs from $k - 4$ by a square. If $z^2 = k$, then

$$\sum_x \left(1 + \left(\frac{x^2(z^2 - 4) - 4(z^2 - k)}{p}\right)\right) = p + \left(\frac{k - 4}{p}\right) \sum_x \left(\frac{x^2}{p}\right),$$

and (2.3) follows since $(\frac{x^2}{p})$ is 0 for $x = 0$ and 1 for every other term.

In the remaining cases, the number of solutions is

$$p + \left(\frac{z^2 - 4}{p}\right) \sum_x \left(\frac{x^2 - 4(z^2 - k)/(z^2 - 4)}{p}\right),$$

so (2.1) follows from a convenient fact about quadratic residues: for any non-zero shift $c \neq 0$,

$$\sum_x \left(\frac{x^2 - c}{p}\right) = -1. \tag{2.5}$$

This is a standard fact that can be shown for $c = t^2$ by factoring

$$x^2 - c = (x - t)(x + t).$$

The terms $x = \pm t$ contribute 0, and if $x \neq t$, then $(x - t)$ and $(x - t)^{-1}$ are squares or not together. We may then change variable to $u = \frac{x+t}{x-t}$ and obtain a complete character

sum missing only $u = 0, 1$ since, in the projective line, $x \neq t, -t, \infty$ corresponds to $u \neq \infty, 0, 1$. This proves (2.5) in case c is a quadratic residue. The sum only depends on whether c is a quadratic residue, so the common value among non-residues can then be obtained by subtraction. Indeed, changing the order of summation gives

$$\sum_{c\neq 0}\sum_{x}\Big(\frac{x^2-c}{p}\Big) = \sum_{x}\sum_{c\neq 0}\Big(\frac{x^2-c}{p}\Big) = -(p-1),$$

so the value for c not a square must also be -1 as in (2.5).

We use this fact once again to sum over z and deduce (2.4). The total is

$$\sum_{z^2\neq 4,k}\Big(p-\Big(\frac{z^2-4}{p}\Big)\Big)+\Big(1+\Big(\frac{k}{p}\Big)\Big)\Big(p+\Big(\frac{k-4}{p}\Big)(p-1)\Big)+\Big(1+\Big(\frac{k-4}{p}\Big)\Big)p,$$

which simplifies as claimed upon collecting the terms in p^2, p, and 1. ∎

Part (2) of Lemma 2.1 restates the following result of [9, Lemma 2.3]. We give a proof to highlight a special property of $k = 0$ compared to other levels, and to confirm that only $(0, 0, 0)$ has multiple self-edges.

Proposition 2.3 (Cerbu–Gunther–Magee–Peilen [9]). *The number of non-zero fixed points of m_1 on $x^2 + y^2 + z^2 = xyz$ is*

$$p - 4 - \Big(\frac{-1}{p}\Big)$$

and only $(0, 0, 0)$ is fixed by more than one Markoff move.

In particular, the Markoff graph mod 5 has no self-edges. It is drawn in Figure 13.1.

Proof. The fixed points under $x \mapsto yz - x$ are given by $x = yz/2$. Substituting this into the Markoff equation yields

$$y^2\Big(1-\frac{z^2}{4}\Big)+z^2 = k.$$

If $z^2 = 4$, then necessarily $k = 4$. In this case, the fixed points are $(x, x, 2)$ and $(x, -x, -2)$ with x arbitrary, and symmetrically $(x, 2, x)$ or $(x, -2, -x)$ with the second and third coordinates exchanged. Assuming $z^2 \neq 4$, we solve for y as:

$$y^2 = 4\frac{z^2-k}{z^2-4}. \tag{2.6}$$

The number of solutions is then a character sum, as before

$$\sum_{z^2\neq 4}\left(1+\left(\frac{(z^2-k)(z^2-4)^{-1}}{p}\right)\right)=p-2+\sum_{z}\left(\frac{(z^2-k)(z^2-4)}{p}\right),$$

where the sum can now be taken over all z, with no contribution from $z^2=4$. For $k=0$, the factor z^2-k is always a square, with Legendre symbol 0 for $z=0$ or 1 otherwise. We account for $z=0$ separately, and the sum over all z is given by (2.5) again:

$$\sum_{z \bmod p}\left(\frac{(z^2-k)(z^2-4)}{p}\right)=\sum_{z\neq 0}\left(\frac{z^2-4}{p}\right)=-1-\left(\frac{-1}{p}\right).$$

The total number of fixed points is then $p-3-(\frac{-1}{p})$, or just $p-4-(\frac{-1}{p})$ excluding $(0,0,0)$. For $k\neq 0$ or 4, the number of fixed points is dictated by a curve (2.6) of genus 1, which in our case degenerates to a conic.

Finally, suppose (x,y,z) is fixed by both m_1 and m_2. Then $x=yz/2$ and $y=xz/2$, which implies that $y=yz^2/4$, so either $y=0$ or $z^2=4$. If $y=0$, then also $x=yz/2=0$, leaving only $(0,0,\sqrt{k})$, or just $(0,0,0)$ in the case $k=0$. If $z^2=4$, which is possible only for $k=4$, then the fixed points are $(x,x,2)$ and $(x,-x,-2)$. ∎

The proof of Theorem 1.3 uses the following result from [13, Corollary 6.2].

Theorem 2.4 (de Courcy-Ireland and Magee [13]). *There is an absolute constant $C>0$ such that any reduced word of length L in m_1, m_2, m_3 has at most $C^L p$ fixed points.*

The constant from [13] is explicit. For instance, one could take 2^{16L+10} in place of C^L. Recall that a word is reduced if there are no trivial cancellations such as m_1 appearing twice in a row.

3. Euler characteristic and main strategy

In this section, we prove Lemma 1.2 and Theorem 1.3. We begin the proof of Theorem 1.1, assuming the Markoff graph mod p is connected and using Lemma 2.1. We complete the calculations with short cycles in the following sections, and address the possibility of disconnected Markoff graphs in Section 8. These arguments prove Theorem 1.1 except perhaps for $p=11,13$. In those cases, the graph is shown to be non-planar using Theorem 1.5 for $p=11$ or Theorem 1.4 for $p=13$.

A planar drawing of a graph divides the plane into connected regions, called faces. Euler's formula states that for a connected graph with V vertices and E edges, dividing the plane into F faces,

$$V-E+F=2.$$

We recommend [11, Chapter 7] as an introduction to Euler characteristic.

To prove Lemma 1.2, we count the number of pairs (v, f) where a vertex v lies on the boundary of a face f. Let us write $v \sim f$ for this incidence relation. By hypothesis, each face has at least g vertices on its boundary. On the other hand, each vertex borders one face for each edge incident to it (with the caveat, for faces incident to a vertex whose removal would disconnect the graph, of counting with multiplicity equal to the number of edges). It follows that

$$gF \leq \sum_f \sum_v \mathbb{1}[v \sim f] = \sum_v \deg(v) = 2E$$

since every edge is counted twice, once for each endpoint. Solving Euler's formula for $F = E - V + 2$, we find $gE - g(V-2) \leq 2E$, and therefore

$$E \leq \frac{g}{g-2}(V-2).$$

This completes the proof of Lemma 1.2.

Proof of Theorem 1.3. More generally, for a graph embedded in a surface of Euler characteristic χ, we would have $V - E + F = \chi$. By the same argument as above,

$$E \leq \frac{g}{g-2}(V-\chi).$$

To prove Theorem 1.3, we think of this as a bound for χ rather than for E:

$$\frac{g-2}{g}E - V \leq -\chi.$$

Since the Markoff graphs mod p do have some short cycles, it is worth introducing some correction terms in order to take a larger value of g. Let n_L be the number of faces of length L. Counting pairs (v, f) as above leads to

$$gF - \sum_{L<g}(g-L)n_L \leq 2E$$

since most faces are incident to at least g vertices, with a deficit $g - L$ for the shorter faces. Substituting $F = E - V + \chi$ into this gives

$$(g-2)E - gV - \sum_{L<g}(g-L)n_L \leq -g\chi. \tag{3.1}$$

For the Markoff graph mod p, or a connected component of it,

$$E \geq \frac{3}{2}V - \frac{3}{2}\left(p - 4 - \left(\frac{-1}{p}\right)\right) \tag{3.2}$$

with equality if the component contains all the self-edges from Lemma 2.1, part (2). In all likelihood, the Markoff graph mod p is connected, but in any event our arguments can be applied to a sufficiently large component. This might be of interest for other surfaces $x^2 + y^2 + z^2 = xyz + k$. The component must have at least p vertices in order for (3.2) to give a positive number of edges.

With the number of edges E bounded from (3.2), inequality (3.1) becomes

$$\frac{1}{2}\Big(1 - \frac{6}{g}\Big)V - \frac{1}{2}\Big(3 - \frac{6}{g}\Big)\Big(p - 4 - \Big(\frac{-1}{p}\Big)\Big) - \sum_{L<g}\Big(1 - \frac{L}{g}\Big)n_L \leq -\chi. \tag{3.3}$$

We will choose $g = \delta \log p$ for a sufficiently small δ, or rather a nearby integer $\lfloor \delta \log p \rfloor$. This ensures that there are few short faces, by Theorem 2.4. Each face is outlined by a word in the Markoff moves m_1, m_2, m_3, up to cyclic ordering, and identifying a word and its inverse as the two orientations of the face. The word, or one of its cyclic shifts, fixes the vertices on the boundary of the face. It follows from Theorem 2.4 that

$$\sum_{L<g}\Big(1 - \frac{L}{g}\Big)n_L \leq \sum_{L<g} C^L p \lesssim p^{1+\varepsilon}$$

for any desired $\varepsilon > 0$, if δ is chosen small enough. We write $\lesssim$ for inequality up to a constant multiple, independent of p, but perhaps depending on ε. Note that the value of C is not the same as before: we multiply by the number of reduced words of length L, which is roughly 2^L, effectively enlarging the previous bound 2^{16L+10} to 2^{17L+10}.

For the giant component of Bourgain–Gamburd–Sarnak, which contains almost all the vertices, we have $V \sim p^2$. The remaining terms are negligible in comparison:

$$\frac{1}{2}\Big(1 - O\Big(\frac{1}{\log p}\Big)\Big)p^2 + O(p^{1+\varepsilon}) \leq -\chi.$$

Theorem 1.3 follows. ∎

We can now prove Theorem 1.1, assuming the Markoff graph mod p is connected. It simplifies the calculations if the graph is connected, but in any case, Section 8 gives an unconditional proof showing that the giant component is not planar. For the proof of Theorem 1.1, we take $g = 7$ rather than $g \approx \log p$, and estimate n_L directly for all $L \leq 6$. Let s be the number of square faces and h the number of hexagons – for any drawing of the graph, these are at most the number of 4-cycles or 6-cycles. As we will see in Section 4, there are no triangles or pentagons. There are then g vertices per face, with a deficit of 1 for each hexagon, and 3 for each square. The key inequality (3.3) becomes

$$\frac{1}{2}\Big(1 - \frac{6}{7}\Big)V - \frac{1}{2}\Big(3 - \frac{6}{7}\Big)\Big(p - 4 - \Big(\frac{-1}{p}\Big)\Big) - \frac{h + 3s}{7} \leq -\chi.$$

We think of this as an upper bound for V, which cannot hold once p is large enough. Recall that $\chi = 2$ for the plane:

$$V \leq 15\Big(p - 4 - \Big(\frac{-1}{p}\Big)\Big) + 2h + 6s - 28. \tag{3.4}$$

In view of Lemma 2.1, we consider four cases depending on p modulo 3 and 4. The cases are $p \equiv 1, 5, 7, 11 \bmod 12$.

If $p \equiv 1 \bmod 12$, then (at the most, for any drawing) the number of squares is $s = 3(p-1)/2$ and the number of hexagons is $h = p - 3$. The inequality (3.4) becomes

$$V \leq 26p - 118.$$

Assuming the Markoff graph mod p is connected, we can take $V = p^2 + 3p$ and solve the quadratic inequality $p^2 + 3p \leq 26p - 118$. For $p = 13$, it does hold in the form $208 \leq 220$, so this case warrants a separate argument. The next example of this form is $p = 37$, and already non-planarity follows because

$$p^2 + 3p = 1480 > 844 = 26p - 118.$$

If $p \equiv 5 \bmod 12$, the number of squares is $s = 3(p-1)/2$, and the number of hexagons is $h = p + 1$. The inequality (3.4) becomes

$$V \leq 26p - 110.$$

If one knew $V = p^2 + 3p$, planarity would be possible only for

$$6.78\ldots = \frac{23 - \sqrt{89}}{2} \leq p \leq \frac{23 + \sqrt{89}}{2} = 16.21\ldots$$

Even 5 and 17, the smallest primes of this form, therefore have non-planar Markoff graphs.

If $p \equiv 7 \bmod 12$, there are no squares and the number of hexagons is $p - 3$. The inequality (3.4) becomes

$$V \leq 17p - 79. \tag{3.5}$$

Assuming the Markoff graph mod p is connected, we can take $V = p^2 - 3p$ since $p \equiv 3 \bmod 4$ in this case. The inequality is already impossible for $p = 19$, the first candidate after $p = 7$ in this progression mod 12. Indeed, the larger root of $p^2 - 3p = 17p - 79$ is

$$10 + \sqrt{21} \approx 14.58\ldots$$

If $p \equiv 11 \bmod 12$, there are no squares and the number of hexagons is $p + 1$. The inequality (3.4) becomes

$$V \leq 17p - 71. \tag{3.6}$$

Assuming connectedness, this inequality shows non-planarity for

$$p > 10 + \sqrt{29} = 15.38\ldots$$

The first prime $p = 11$ in this progression requires special treatment. For example, modulo 11, we have $-7 = 4 = 2^2$, so non-planarity follows from Theorem 1.5. The other case left after the arguments above is $p = 13$, which has a non-planar Markoff graph by Theorem 1.4.

4. Short words

In this section, we advance the proof of Lemma 2.1 by identifying which words in m_1, m_2, m_3 can possibly bound a face of 6 sides or less. Up to a permutation of the coordinates, we may assume the word's first move is m_1, followed by m_2. We need only consider reduced words, where none of the involutions m_1, m_2, or m_3 occurs twice in a row. It is convenient to omit the m's, simply writing j for m_j. The words of length up to 6 are then

$$\begin{aligned}
&1\\
&21\\
&121, 321\\
&2121, 3121, 1321, 2321\\
&12121, 32121, 13121, 23121, 21321, 31321, 12321, 32321\\
&212121, 232121, 313121, 323121, 321321, 231321, 312321, 132321\\
&312121, 132121, 213121, 123121, 121321, 131321, 212321, 232321.
\end{aligned}$$

We can immediately discard words where some move occurs only once, such as 232321. These do not give new faces in the Markoff graph, but simply add a self-edge somewhere along a face that has already been counted.

Likewise, there is no contribution from words that have a shorter conjugate. The fixed points of $w^{-1}w_0w$ do not yield new faces in the Markoff graph. Instead, one applies w to the fixed point, traverses a face bounded by w_0, and returns along the same path.

After deleting words with a lone letter or a shorter conjugate, we are left with

$$2121, \quad 212121, \quad 321321, \quad 323121, \quad 231321, \quad 312321.$$

The last three are equivalent to each other under cyclic shifts and permutations of the coordinates:

$$312321 \sim 131232$$
$$231321 \sim 123132 \sim 212313,$$

where the rightmost words have the same structure as 323121 up to a permutation. We will see in Section 5 that these words do not bound any faces. The remaining cases 2121, 212121, and 321321 will be treated in Sections 6 and 7, completing the proof of Lemma 2.1.

5. Fixed points of 323121

Proposition 5.1. *The fixed points of $m_3m_2m_3m_1m_2m_1$ on the Markoff surface $x^2 + y^2 + z^2 = xyz$ are the triples (x, y, z) satisfying*

$$x^4 - 5x^2 + 8 = 0, \quad z = \pm x, \quad y = \frac{xz}{x^2 - 2},$$

together with $(0, 0, 0)$. These do not correspond to faces in the Markoff graph mod p. *Instead, there are self-edges m_2 at both neighbours of the fixed point (x, y, z) under m_1 and m_3.*

For example, this occurs in the Markoff graph mod 11 with $x = 3$ (Figure 13.3). The self-edges correspond to $6 = 3 \times 4 - 6$ at $(3, 6, 4)$ or $(4, 6, 3)$. Over other fields, one needs to have an element

$$x = \sqrt{\frac{5 + \sqrt{-7}}{2}}.$$

Proof. To lower the degree of the fixed point system, note that m_{323121} fixes (x, y, z) if and only if

$$m_3m_2m_3(x, y, z) = m_1m_2m_1(x, y, z),$$
$$\begin{pmatrix} x \\ x(xy - z) - y \\ x(x(xy - z) - y) - xy + z \end{pmatrix} = \begin{pmatrix} z(z(yz - x) - y) - yz + x \\ z(yz - x) - y \\ z \end{pmatrix}.$$

This simplifies to

$$\begin{cases} z(z(yz - x) - 2y) = 0, \\ y(z^2 - x^2) = 0, \\ x(x(xy - z) - 2y) = 0. \end{cases}$$

Assuming $xyz \neq 0$, we find that $z^2 = x^2$ and solve for y from $(x^2-2)y = xz$. Substituting this into the Markoff equation $x^2 + y^2 + z^2 = xyz + k$, we are left with a single-variable sextic for x:

$$\frac{(x^4 - 5x^2 + 8)x^2}{(x^2-2)^2} = k.$$

For $k = 0$, this reduces to a biquadratic equation $x^4 - 5x^2 + 8 = 0$ as claimed, assuming $x \neq 0$. If $x = 0$, and likewise if y or z vanishes, then the system implies that at least two variables must vanish. The only such solutions of the Markoff equation are $(0,0,0)$ for $k \neq 0$, or more generally the permutations of $(0, 0, \pm\sqrt{k})$ for other levels.

From $z(yz - x) - 2y = 0$, we see that m_2 fixes $(yz - x, y, z)$, and similarly for $(x, y, xy - z)$. This shows that there are self-edges at the neighbours of (x, y, z), as claimed and completing the proof. ∎

6. Fixed points of 321321

The situation here depends on whether -3 is a quadratic residue modulo p. If so, then the next proposition shows that there are four hexagons fixed by 321321 and its cyclic shifts. Permutations of the coordinates do not lead to any further hexagons: the permuted words are either cyclic shifts 132132 and 213213, or their inverses, which all bound the same faces. This case accounts for the four hexagons visible in the Markoff graph mod 7 (Figure 1.1).

Proposition 6.1. *The fixed points of $m_3m_2m_1m_3m_2m_1$ on $x^2 + y^2 + z^2 = xyz$ are the triples (x, y, z) satisfying*

$$y^2 + 3y + 3 = 0, \quad x^2 = \frac{y^2}{(y+1)^2}, \quad z = -x,$$

or

$$y^2 - 3y + 3 = 0, \quad x^2 = \frac{y^2}{(y-1)^2}, \quad z = x.$$

Proof. The fixed points are given by $m_{321321}(x, y, z) = (x, y, z)$, or equivalently

$$m_1m_2m_3(x, y, z) = m_3m_2m_1(x, y, z),$$

$$\begin{pmatrix} (xy - z)(x(xy - z) - y) - x \\ x(xy - z) - y \\ xy - z \end{pmatrix} = \begin{pmatrix} yz - x \\ z(yz - x) - y \\ (yz - x)(z(yz - x) - y) - z \end{pmatrix}.$$

This simplifies to

$$\begin{cases} x((xy-z)^2-y^2)=0, \\ y(x^2-z^2)=0, \\ z((yz-x)^2-y^2)=0. \end{cases}$$

Suppose that $xyz \neq 0$. Then $x^2 = z^2$ from the middle equation, and this leads to a redundancy. Since $z = \pm x$, we have $xy - z = \pm(yz - x)$ and the remaining two equations become equivalent. We consider the two cases $z = \pm x$ separately and solve for x from $y^2 = (xy - z)^2 = x^2(y \mp 1)^2$. Substituting this relation and $z = \pm x$ into the Markoff equation, one finds

$$x^2 + y^2 + z^2 = xyz + k \implies 2\frac{y^2}{(y \mp 1)^2} + y^2 = \pm\frac{y^3}{(y \mp 1)^2} + k.$$

For $k = 0$, assuming $y \neq 0$, we divide by y^2 and obtain the two quadratics from the statement of the proposition.

It remains to consider the possibility that some of x, y, z could be 0. If $x = 0$, then $y(x^2 - z^2) = 0$ implies that either y or z must also be 0. This leaves only $(0, 0, 0)$ as a fixed point on the original surface $x^2 + y^2 + z^2 = xyz$, or more generally permutations of $(0, 0, \sqrt{k})$ on $x^2 + y^2 + z^2 = xyz + k$. ∎

The solutions for y are $\frac{1}{2}(\pm 3 \pm \sqrt{-3})$. Modulo 7, we choose $\sqrt{-3} = \pm 2$ and obtain for example $(x, y, z) = (4, 1, 3)$ from one of the hexagons of Figure 1.1.

7. Alternating words

The remaining words 2121 and 212121 do not change the third coordinate z, and act linearly on (x, y). By diagonalizing this action, one can determine the fixed points of any alternating word $(21)^L$.

Proposition 7.1. *The only fixed points of* $(m_2 \circ m_1)^L$ *on* $x^2 + y^2 + z^2 = xyz + k$ *are* $(0, 0, \sqrt{k})$ *unless* L *is divisible by* p, *or* L *is a factor of* $(p - 1)/2$ *or* $(p + 1)/2$.

(a) *If* L *is divisible by* p, *then the fixed points are* $(x, y, \pm 2)$ *together with* $(0, 0, \sqrt{k})$. *The former lie on the lines* $(x \mp y)^2 = k - 4$.

(b) *If* $(p \pm 1)/2$ *is divisible by* L, *then the fixed points of* $(m_2 \circ m_1)^L$ *are* (x, y, z) *where* $z = \zeta + \zeta^{-1}$ *with* $\zeta \in \mathbb{F}_{p^2}^{\times}$ *a solution of*

$$\zeta^{2L} = 1, \quad \zeta \neq \pm 1,$$

together with $(0, 0, \sqrt{k})$ *for the level* $x^2 + y^2 + z^2 = xyz + k$.

The form of the fixed points in (a) and (b) does not depend on k. One simply imposes $x^2 + y^2 + z^2 = xyz + k$ in addition to the fixed-point system, and includes also the exceptional points $(0, 0, \sqrt{k})$ where two coordinates equal 0.

The case 2121, where $L = 2$, corresponds to $z = 0$ and $\zeta = \sqrt{-1}$. However, if $k = 0$, even though we allow ζ in a quadratic extension, the value $z = 0$ is only possible for $p \equiv 1 \bmod 4$. Substituting $z = 0$ in the Markoff equation $x^2 + y^2 + z^2 = xyz$ gives $x^2 + y^2 = 0$. If $p \equiv 3 \bmod 4$, then the only solution is $(0, 0, 0)$, or else $(x/y)^2 = -1$.

Corollary 7.2. *The fixed points of* 2121 *are* $(x, y, 0)$ *with*

$$x^2 + y^2 = 0.$$

If $p \equiv 3 \bmod 4$*, the only fixed point on* $x^2 + y^2 + z^2 = xyz \bmod p$ *is* $(0, 0, 0)$.

The case 212121, where $L = 3$, corresponds to $z = \pm 1$ with

$$\pm\zeta = \frac{1 + \sqrt{-3}}{2}.$$

Corollary 7.3. *The fixed points of* 212121 *are* $(x, y, \pm 1)$ *with*

$$x^2 + y^2 + 1 = \pm xy.$$

Proof of Proposition 7.1. The action of $m_2 \circ m_1$ is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \xrightarrow{m_1} \begin{pmatrix} yz - x \\ y \\ z \end{pmatrix} \xrightarrow{m_2} \begin{pmatrix} yz - x \\ z(yz - x) - y \\ z \end{pmatrix},$$

that is,

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -1 & z \\ -z & z^2 - 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

This matrix has determinant 1 and trace $z^2 - 2$. In terms of a change of variable

$$z = \zeta + \zeta^{-1}$$

the eigenvalues are then ζ^2 and ζ^{-2}. Here, ζ may lie in a quadratic extension of $\mathbb{F}_p$ if need be. In order to have $z = \zeta + \zeta^{-1}$ belong to $\mathbb{F}_p$, it must be that either $\zeta^{p+1} = 1$ or $\zeta^{p-1} = 1$. The order of ζ^2 in $\mathbb{F}_{p^2}^{\times}$, is therefore a divisor of $(p - 1)/2$ or $(p + 1)/2$.

For $z \neq \pm 2$, the matrix representing $m_2 m_1$ can be diagonalized, and its order is the order of ζ^2. After computing the eigenvectors, we find

$$\begin{pmatrix} -1 & z \\ -z & z^2 - 1 \end{pmatrix} = \begin{pmatrix} 1 & \zeta \\ \zeta & 1 \end{pmatrix} \begin{pmatrix} \zeta^2 & 0 \\ 0 & \zeta^{-2} \end{pmatrix} \begin{pmatrix} 1 & \zeta \\ \zeta & 1 \end{pmatrix}^{-1}$$

and

$$\begin{pmatrix} -1 & z \\ -z & z^2-1 \end{pmatrix}^L = \begin{pmatrix} 1 & \zeta \\ \zeta & 1 \end{pmatrix} \begin{pmatrix} \zeta^{2L} & 0 \\ 0 & \zeta^{-2L} \end{pmatrix} \begin{pmatrix} 1 & \zeta \\ \zeta & 1 \end{pmatrix}^{-1}.$$

If $\zeta^{2L} = 1$, then every vector (x, y) is fixed, as claimed. Conversely, if $\zeta^{2L} \neq 1$, only $(0, 0)$ is fixed. This gives only $(0, 0, 0)$ in the Markoff surface, or more generally $(0, 0, \sqrt{k})$ for other level sets $x^2 + y^2 + z^2 = xyz + k$.

If $z = \pm 2$, then $\zeta = \pm 1$ so there is a repeated eigenvalue $\zeta^2 = \zeta^{-2} = 1$, and the eigenvectors above become multiples of each other by ± 1. In this case, $m_2 m_1$ has order p in view of the following Jordan form:

$$\begin{pmatrix} -1 & z \\ -z & z^2-1 \end{pmatrix} = \begin{pmatrix} -1 & \pm 2 \\ \mp 2 & 3 \end{pmatrix} = \begin{pmatrix} \pm 2 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \pm 2 & 0 \\ 2 & 1 \end{pmatrix}^{-1}.$$

The powers of $m_2 m_1$ are given by

$$\begin{pmatrix} -1 & z \\ -z & z^2-1 \end{pmatrix}^L = \begin{pmatrix} -1 & \pm 2 \\ \mp 2 & 3 \end{pmatrix} = \begin{pmatrix} \pm 2 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \pm 2 & 0 \\ 2 & 1 \end{pmatrix}^{-1}.$$

If L is divisible by p, then every vector (x, y) is fixed. If L is not divisible by p, then the fixed points are given by $x = \pm y$ with the same sign as in $z = \pm 2$. Substituting this into $x^2 + y^2 + z^2 = xyz + k$ gives $2x^2 + 4 = 2x^2 + k$. There are no such fixed points, unless $k = 4$. ∎

8. Non-planarity of the cage

In this section, we show that even if the Markoff graph mod p is disconnected, it has a large non-planar component. This is the giant component constructed by Bourgain–Gamburd–Sarnak from what they call *the cage* [6, Section 3.2]. The cage consists of triples (x, y, z) where at least one of the coordinates has maximal order with respect to the analysis from Section 7. It is shown in [6] that all of these points belong to the same connected component.

Recall the change of variable

$$z = \zeta + \zeta^{-1}, \quad \zeta^{p+1} = 1 \text{ or } \zeta^{p-1} = 1.$$

The maximal order is therefore $p + 1$. The number of elements of order $p + 1$ in the cyclic group $\mathbb{F}_{p^2}^\times$ is given by Euler's totient function $\phi(p + 1)$. These correspond to $\frac{1}{2}\phi(p + 1)$ values of $z = \zeta + \zeta^{-1}$. We ignore the possibility that $\zeta = \zeta^{-1}$, since then $z = \pm 2$. This arises only for $p \equiv 1 \bmod 4$, in which case we might as well conclude

non-planarity from Theorem 1.4. The configuration used to prove Theorem 1.4 meets every level set where a coordinate x, y, or z takes a given value, as will be clear from (10.1), and in particular it lies in the same component as the cage.

For each of these maximal values of z, there are $p+1$ solutions (x, y), by equation (2.1). Indeed, in these cases, $\zeta^p = \zeta^{-1}$ so ζ is "imaginary" and there are $p+1$ solutions (x, y) rather than $p-1$. There are then $\frac{1}{2}p\phi(p+1)$ triples (x, y, z) where z has maximal order, and similarly for the first or second coordinate. Of course, more than one coordinate could have maximal order.

An interesting example is $p = 7$, where the cage encloses the entire graph. In this case, $p+1 = 8$ so let ζ be an eighth root of unity. Write $i^2 = -1$ in the quadratic extension of $\mathbb{F}_7$, and observe that $3^2 = 9 \equiv 2 \bmod 7$. The maximal order therefore occurs for z equal to

$$\frac{1+i}{\sqrt{2}} + \frac{1-i}{\sqrt{2}} = \frac{2}{\sqrt{2}} = \sqrt{2} = \pm 3.$$

The Markoff graph mod 7 (Figure 1.1) has four vertices such as $(3, 3, 3)$ up to sign changes, where all coordinates have maximal order. The twelve neighbours of those, such as $(6, 3, 3)$, have two coordinates of maximal order. Another twelve points, such as $(1, 6, 3)$, have only one maximal coordinate. These account for all solutions in the form $28 = 4 + 12 + 12 = p\phi(p+1)$. There is a cycle of length 8 at every vertex, with three such octagons meeting at $(3, 3, 3)$; two octagons and a hexagon at $(6, 3, 3)$; or an octagon, a hexagon, and a self-edge at $(1, 6, 3)$.

The points in the cage show that there is a connected component of size at least

$$V \geq \frac{1}{2}p\phi(p+1) > \frac{p^2}{1000 \log\log p}, \tag{8.1}$$

where we have used a loose estimate for Euler's totient function ϕ. Asymptotically, a formula of Mertens gives

$$\phi(n) \geq (e^{-\gamma} + o(1))\frac{n}{\log\log n},$$

where γ is the Euler–Mascheroni constant and $e^{-\gamma} \approx 0.5614$; see [19, Theorem 7] or [18, Theorem 429]. The correct constant is much larger than the underestimate $1/500$ from (8.1), but perhaps only applicable for large n. The rougher form (8.1) is valid for all p and follows from Chebyshev-style estimates for prime numbers. We discuss these in Section 9.

We substitute (8.1) in (3.4), where the number of squares is $s = 0$ since we are now interested only in $p \equiv 3 \bmod 4$. The number of hexagons is at most $p+1$, as in (3.6). If the connected component of the cage is planar, it follows that

$$\frac{1}{2}p\phi(p+1) \leq 17p - 71, \quad \phi(p+1) < 34. \tag{8.2}$$

Even with a crude bound for ϕ, this implies

$$\frac{p+1}{\log\log(p+1)} \leq 500\phi(p+1) < 17000. \tag{8.3}$$

The solution to $x/\log\log x = 17000$, using Newton's method for instance, is $x = 40134.5\ldots$ In particular, the inequality in (8.3) is reversed if $p > 40133$ (which factors as 67×599, the nearest prime being 40129). One could certainly narrow the search further using better estimates, but it is already feasible to compute $\phi(p+1)$ for all primes up to 40129 (and we only need those congruent to 3 mod 4). The criterion (8.2) is satisfied for $p \leq 101$, but no larger primes. The congruences modulo 28 from Corollary 1.6 show non-planarity for several of these, leaving only

$$p = 7, 19, 31, 47, 59, 83.$$

Of the remaining cases, $p = 7$ does in fact have a planar Markoff graph, which coincides with the cage. The others are small enough that one can check the graph is connected, either by enumerating enough triples (x, y, z), or by a spectral method (Section 12; see also [12] for connectedness up to $p \leq 2999$). The true value V is then even larger than the lower bound from the cage, and non-planarity follows from the reasoning in Section 3.

Figure 13.2 shows part of the Markoff graph for $p = 19$, the smallest example where neither Theorem 1.4 nor Theorem 1.5 nor the lower bound from the cage is enough to deduce non-planarity.

9. Lower bound for Euler's totient function

In this section, we prove the estimate (8.1) for Euler's function $\phi(n)$, given by

$$\phi(n) = n\prod_{p|n}\Big(1 - \frac{1}{p}\Big) \geq \frac{1}{500}\frac{n}{\log\log n}. \tag{9.1}$$

This is a standard topic, with excellent expositions available in [19, Theorem 7], [18, Theorem 429], and [16, Section 2.2]. We follow them closely, and simply keep track of the implicit constants.

Consider the contributions to (9.1) from large primes $p > L$ and small primes $p \leq L$. Eventually, a good choice will be $L = \log n$. There are not too many large factors of n, because $n > L^k$ if there are k large primes among the factors of n. Therefore, $k < \log n/\log L$, and each $p > L$ contributes at least $1 - 1/L$ to the product. For the small primes, we obtain a lower bound by extending the product to

all $p \le L$, regardless of whether they divide n, since each term $1 - 1/p$ is less than 1. It follows that

$$\prod_{p|n}\Big(1-\frac{1}{p}\Big) > \Big(1-\frac{1}{L}\Big)^{\log n/\log L} \prod_{p\le L}\Big(1-\frac{1}{p}\Big).$$

The exponent $\log n/\log L$ is greater than 1, so the binomial expansion gives

$$\Big(1-\frac{1}{L}\Big)^{\log n/\log L} \ge 1 - \frac{\log n}{L\log L},$$

which will be bounded below if one chooses $L \asymp \log n$ or larger. With $L = C\log n$, the contribution of large primes is at least

$$1 - \frac{1}{C\log L} \ge \frac{1}{2} \quad \text{for } n \ge e^{e^{2/C}}. \tag{9.2}$$

The decisive contribution, that of small primes, is given asymptotically by a formula of Mertens:

$$\prod_{p\le L}\Big(1-\frac{1}{p}\Big) \sim \frac{e^{0.5772\ldots}}{\log L},$$

where the value in the exponent is the Euler–Mascheroni constant [19, Theorem 7]. For our purposes, it is better to have a less precise estimate that applies already for small values of L.

We first take logarithms to convert the product to a sum, and then extract the leading term from the power series $\log(1-x) = -x + \cdots$ obtaining:

$$\prod_{p\le L}\Big(1-\frac{1}{p}\Big) = \exp\Big(-\sum_{p\le L}\frac{1}{p} + \sum_{p\le L}\Big(\frac{1}{p} + \log\Big(1-\frac{1}{p}\Big)\Big)\Big). \tag{9.3}$$

The second sum converges, since its terms are dominated by p^{-2}. Numerically,

$$\sum_{p\le L}\Big(\frac{1}{p} + \log\Big(1-\frac{1}{p}\Big)\Big) \ge \sum_{p}\Big(\frac{1}{p} + \log\Big(1-\frac{1}{p}\Big)\Big) = -0.3157\ldots$$

The main term in (9.3) is therefore $\sum_{p\le L} 1/p$, which is well known to be of order $\log\log L$ (as discussed in the same reference [19, Theorem 7] for instance). For an explicit bound of this form, we first sum by parts:

$$\sum_{p\le L}\frac{1}{p} = \sum_{p\le L}\frac{\log p}{p}\frac{1}{\log p} = \int_2^L \frac{S(t)}{t(\log t)^2}\,dt + \frac{S(L)}{\log L}, \tag{9.4}$$

where we have differentiated $1/\log p$ and integrated $\log p/p$. The summatory function can be bounded by extending the range to include prime powers:

$$S(t) = \sum_{p \le t} \frac{\log p}{p} \le \sum_{m \le t} \frac{\Lambda(m)}{m}, \tag{9.5}$$

where $\Lambda(m) = \log p$ if m is a power of a prime p, and 0 otherwise. These weights are more convenient because of the identity

$$\sum_{d \mid n} \Lambda(d) = \log n$$

and its sum

$$\sum_{d \le L} \Lambda(d) \Big\lfloor \frac{L}{d} \Big\rfloor = \sum_{\ell \le L} \log \ell.$$

We multiply and divide by L, noting that $L/d \le \lfloor L/d \rfloor + 1$,

$$\sum_{d \le L} \frac{\Lambda(d)}{d} \le \frac{1}{L} \sum_{\ell \le L} \log \ell + \frac{1}{L} \sum_{d \le L} \Lambda(d).$$

The first sum can be estimated by an integral:

$$\int_1^x \log t \, dt \le \sum_{\ell \le x} \log \ell \le \int_1^{x+1} \log t \, dt = (x+1)\log(x+1) - x. \tag{9.6}$$

For the remainder, we claim that the following Chebyshev-style estimate holds already for any $x \ge 2$:

$$\sum_{d \le x} \Lambda(d) \le x \log 4 + (\log x + 2) \frac{\log x}{\log 2}. \tag{9.7}$$

Assuming this for the moment, we continue with (9.5):

$$\begin{aligned} S(t) &\le \sum_{m \le t} \frac{\Lambda(m)}{m} \le \frac{1}{t} \sum_{\ell \le t} \log \ell + \frac{1}{t} \sum_{d \le t} \Lambda(d) \\ &\le \log t - 1 + \log 4 + \frac{(\log t + 2)\log t}{t \log 2} \le \log t + 2. \end{aligned}$$

Finally, we substitute this into (9.4) and find

$$\sum_{p \le L} \frac{1}{p} \le \int_2^L \frac{\log t + 2}{t(\log t)^2} \, dt + \frac{\log L + 2}{\log L}.$$

The integral can be computed exactly by a substitution $u = \log t$ with $du = dt/t$, whence

$$\sum_{p \le L} \frac{1}{p} \le \log\log L - \log\log 2 + 2\Big(\frac{1}{\log 2} - \frac{1}{\log L}\Big) + \frac{\log L + 2}{\log L}$$
$$\le \log\log L + 5.$$

The original product from (9.3) is then, with $L = \log n$,

$$\prod_{p \le L} \Big(1 - \frac{1}{p}\Big) \ge \exp(-\log\log L - 5 - 0.3157) \ge \frac{1}{250 \log\log n}.$$

The loose estimate (8.1) gives up an extra factor of 2 from the large primes. This is guaranteed by (9.2) for $n \ge e^{e^2} \approx 1618$, and one can check the smaller values of n to be sure (with room to spare) that $\phi(n) > \frac{1}{500} n / \log\log n$ for all $n \ge 2$.

To prove (9.7), recall the notation $\psi(x) = \sum_{n \le x} \Lambda(n)$. In terms of ψ,

$$\sum_{m \le x} \psi\Big(\frac{x}{m}\Big) = \sum_{d \le x} \Lambda(d) \Big\lfloor \frac{x}{d} \Big\rfloor = \sum_{n \le x} \log n.$$

Subtraction gives

$$\sum_{n \le x} \log n - 2 \sum_{n \le x/2} \log n = \sum_{m \le x} \psi\Big(\frac{x}{m}\Big) - 2 \sum_{m \le x/2} \psi\Big(\frac{x}{2m}\Big)$$
$$\ge \psi(x) - \psi\Big(\frac{x}{2}\Big)$$

because each difference $\psi(x/(2j-1)) - \psi(x/(2j))$ is non-negative. This can be simplified using (9.6) for the logarithms:

$$\psi(x) \le \psi(x/2) + x\log 2 + \log x + (x+1)\log\Big(1 + \frac{1}{x}\Big)$$
$$< \psi(x/2) + x\log 2 + \log x + 2.$$

This can be iterated to bound $\psi(x)$ in terms of $\psi(x/2)$, then $\psi(x/4)$, $\psi(x/8)$, and so on. After roughly $k \sim \log x / \log 2$ iterations, we reach an empty sum $\psi(x/2^k) = 0$, leaving only a geometric progression:

$$\psi(x) < x\Big(1 + \frac{1}{2} + \dots\Big)\log 2 + (\log x + 2)\frac{\log x}{\log 2}$$
$$< x\log 4 + (\log x + 2)\frac{\log x}{\log 2},$$

as required.

For comparison, although it is only the upper bound that is relevant in our context, Niven [27] gives a lower bound of the same character, bounding $\sum_{p\le L} 1/p$ from below by $\log\log L$ less an explicit constant. That argument does not require Chebyshev's estimates for $\psi(x)$. For the correct constant in the asymptotic as $L\to\infty$, see [19, Theorem 7, p. 22].

10. Proof of Theorem 1.4

To prove Theorem 1.4, we produce a complete bipartite graph joining the permutations of $(2+2i,2,2)$ and $(2-2i,2,2)$. One can check as follows that $(m_1\circ m_2)^{(p-1)/2}$ takes $(2+2i,2,2)$ to $(2,2-2i,2)$. By definition,

$$m_2(2+2i,2,2) = (2+2i,2+4i,2)$$

and then

$$m_1\circ m_2(2+2i,2,2) = (2+6i,2+4i,2).$$

Inductively, one finds that for each $k\ge 0$,

$$(m_1\circ m_2)^k(2+2i,2,2) = (2+(4k+2)i,2+4ki,2). \tag{10.1}$$

In particular, with $k=(p-1)/2$, the claim follows since we work modulo p:

$$(m_1\circ m_2)^{(p-1)/2}(2+2i,2,2) = (2+2pi,2+2(p-1)i,2) = (2,2-2i,2).$$

We emphasize that the only coordinate not equal to 2, in addition to moving from the x-coordinate to the y-coordinate, has changed from $2+2i$ to $2-2i$.

In the same way, we find that $(m_1\circ m_3)^{(p-1)/2}$ takes $(2+2i,2,2)$ to $(2,2,2-2i)$, while $(m_2\circ m_3)^{(p-1)/2}$ takes $(2,2+2i,2)$ to $(2,2,2-2i)$. Thus the Markoff graph contains the configuration drawn in Figure 10.1. We abbreviate $m_j\circ m_k$ by m_jm_k. The top half of the figure has an outer curve connecting $(2+2i,2,2)$ to $(2,2-2i,2)$ and $(2,2,2-2i)$ via

$$(m_1m_3)^{(p-1)/2}(2+2i,2,2) = (2,2,2-2i),$$
$$(m_1m_2)^{(p-1)/2}(2+2i,2,2) = (2,2-2i,2),$$

and an inner curve connecting the points $(2,2+2i,2)$ and $(2,2,2+2i)$ to $(2-2i,2,2)$. The bottom half shows the analogous relation for the second and third coordinates, namely,

$$(m_2m_3)^{(p-1)/2}(2,2+2i,2) = (2,2,2-2i),$$

but because of the choices we have already made in drawing the top half, $(2,2+2i,2)$ is part of the inner circle while $(2,2,2-2i)$ is on the outside. Likewise, $(2,2-2i,2)$

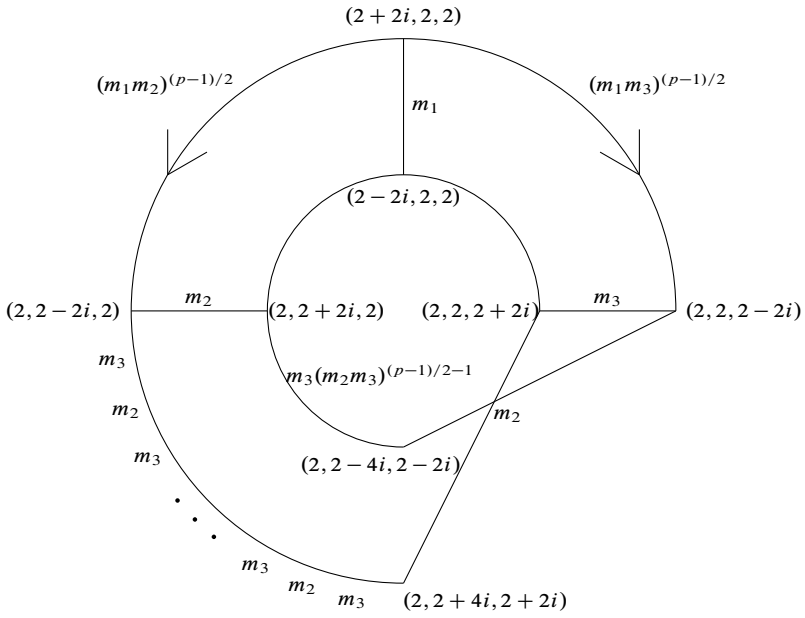


Figure 10.1. For any prime $p \equiv 1 \bmod 4$, the Markoff graph contains a subdivision of the complete bipartite graph connecting $(2, 2, 2 + 2i)$ and its permutations to $(2, 2, 2 - 2i)$ and its permutations.

on the outside is connected to $(2, 2, 2 + 2i)$ on the inside. To connect the outer and inner circles in this way requires a crossing of edges. In Figure 10.1, this crossing corresponds to the final move m_2 in the paths

$$(2, 2, 2 - 2i) = m_2(2, 2 - 4i, 2 - 2i) = m_2 \circ m_3(m_2m_3)^{(p-1)/2-1}(2, 2 + 2i, 2),$$
$$(2, 2, 2 + 2i) = m_2(2, 2 + 4i, 2 + 2i) = m_2(m_3m_2)^{(p-1)/2}(2, 2 + 2i, 2).$$

Starting from this configuration, we contract edges as follows to produce a minor isomorphic to $K_{3,3}$. First, contract the edges forming the inner and outer quarter-circles in the top half of the figure. This connects $(2 + 2i, 2, 2)$ to $(2, 2 - 2i, 2)$ and to $(2, 2, 2 - 2i)$, as well as $(2 - 2i, 2, 2)$ to $(2, 2 + 2i, 2)$ and $(2, 2, 2 + 2i)$. Second, contract the edges in the "outer third quadrant" from $(2, 2 - 2i, 2)$ to $(2, 2 + 4i, 2 + 2i)$, leaving a path from $(2, 2 - 2i, 2)$ to $(2, 2, 2 + 2i)$. Third, contract the "inner third quadrant" to obtain a path from $(2, 2 + 2i, 2)$ to $(2, 2, 2 - 2i)$. The resulting graph has six vertices, namely the permutations of $(2 \pm 2i, 2, 2)$, such that all vertices with a coordinate $2 + 2i$ are connected to all vertices with a coordinate $2 - 2i$. This is a complete bipartite graph $K_{3,3}$ as required.

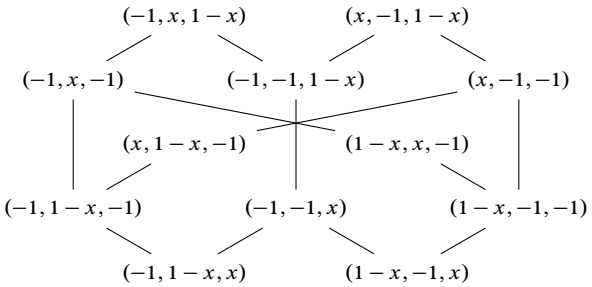


Figure 11.1. If -7 is a non-zero square modulo p, then $x^2 + y^2 + z^2 = xyz \bmod p$ has solutions $(x, -1, -1)$ and $(1 - x, -1, -1)$ and their permutations. From a solution with coordinate x, one can reach any solution with coordinate $1 - x$ by either a single move or two moves.

In this construction, two moves such as m_1 and m_2 alternate between the lines $x - y = \pm 2i$ contained in the Markoff surface for $p \equiv 1 \bmod 4$. For $p \equiv 3 \bmod 4$, all the lines are imaginary.

11. Proof of Theorem 1.5

In this section, we prove Theorem 1.5 by producing a copy of Figure 11.1 inside the Markoff graph mod p, so long as -7 is a quadratic residue. Consider the Markoff equation with $y = z = -1$. The final coordinate must satisfy

$$x^2 + 2 = x, \quad x = \frac{1 \pm \sqrt{-7}}{2},$$

so there are two solutions x and $1 - x$ if -7 is a non-zero square modulo p; a single solution for $p = 7$; and no solutions otherwise. Suppose -7 is a square. Then there are six solutions $(x, -1, -1)$, $(1 - x, -1, -1)$ and their permutations. A single move m_1 connects $(x, -1, -1)$ to $(1 - x, -1, -1)$. A duo of moves $m_1 \circ m_2$ leads to

$$(x, -1, -1) \mapsto (x, 1 - x, -1) \mapsto (-1, 1 - x, -1).$$

In the same way, $m_1 \circ m_3$ takes $(x, -1, -1)$ to $(-1, -1, 1 - x)$, and one has similar paths $m_j \circ m_k$ starting from $(-1, x, -1)$ and $(-1, -1, x)$. These paths starting from $(x, -1, -1)$ are illustrated in Figure 11.1. Together with their counterparts at $(-1, x, -1)$ and $(-1, -1, x)$, they form a copy of $K_{3,3}$.

The same configuration occurs in other level sets $x^2 + y^2 + z^2 = xyz + k$ whenever there are solutions with two coordinates equal to -1. Setting $y = z = -1$, the solution for x is

$$x = \frac{1 \pm \sqrt{4k - 7}}{2}.$$

It would seem to follow that the Markoff graph of level k has a non-planar component provided $4k-7$ is a non-zero square modulo p. However, the configuration might not be as shown in Figure 11.1 if there are self-edges. For example, if $k=4$, one has $\sqrt{4k-7}=\sqrt{9}=3$, so the special values are $x=2$ and $1-x=-1$. Each point of the form $(2,-1,-1)$ has a pair of self-edges because $-1\mapsto 2(-1)-1=-1$. The whole construction lies in a planar component consisting of $(-1,-1,-1)$ and its neighbours, which is essentially the cluster from Figure 1.2 since $k=4\equiv 0 \bmod 2$. Self-edges occur in Figure 11.1 only for $x=-1$ or $x=2$, but these can equal $(1+\sqrt{4k-7})/2$ only for $k=4$.

12. Planar graphs do not expand

This section reviews why planar graphs cannot form an expander family, which was our motivation for studying planarity of Markoff graphs (or, more hopefully, their non-planarity). The failure of expansion in planar graphs is a consequence of a celebrated theorem of Lipton and Tarjan.

Theorem 12.1 (Lipton–Tarjan planar separator theorem [22]). *For any planar graph on n vertices, the vertex set can be partitioned into three sets A, B, and C such that no vertex in A is connected to any vertex in B, each of A and B contains at most $2n/3$ vertices, and C contains at most $2\sqrt{2n}$ vertices.*

Moreover, Lipton and Tarjan give an algorithm for computing such a partition in $O(n)$ steps.

A standard way to quantify expansion is the Cheeger constant. For a graph G, the Cheeger constant $h(G)$ is defined as

$$h(G)=\min \frac{|\partial A|}{\min(|A|,|G\setminus A|)}, \tag{12.1}$$

where the minimum is taken over all non-empty, proper subsets A of the vertices of G, and ∂A is the set of edges joining a vertex in A to another vertex in its complement $G\setminus A$. If $h(G)=0$, then G is disconnected since there is a subset A with $|\partial A|=0$, that is, no edges from A to its complement. Expansion refers to a sequence of graphs with a growing number of vertices, but $h(G)$ bounded strictly away from 0.

Theorem 12.1 implies that, for any sequence of planar graphs with a growing number of vertices, $h(G)\to 0$. Indeed, given a planar graph G on n vertices, consider sets A, B, and C as in Theorem 12.1. We use one of the large parts, say A, as a candidate for the ratio $|\partial A| \div \min(|A|,|G\setminus A|)$ in the definition (12.1) of $h(G)$. There are no edges between A and B, so

$$|\partial A|\le |C|\le 2\sqrt{2n}.$$

On the other hand, $|A| \geq n/3 - 2\sqrt{2n}$ because B and C together account for at most $2n/3 + 2\sqrt{2n}$ vertices. Likewise, $|G \setminus A| \geq n/3$ because $|A| \leq 2n/3$. It follows that

$$h(G) \leq \frac{|\partial A|}{\min(|A|, |G \setminus A|)} \leq \frac{2\sqrt{2n}}{n/3 - 2\sqrt{2n}} \lesssim n^{-1/2}. \tag{12.2}$$

In particular, $h(G) \to 0$ as $n \to \infty$.

In contrast, numerical evidence [12] suggests that $h(G)$ is bounded away from 0 for Markoff graphs with $p \to \infty$. It is easier to compute a different measure of expansion, namely the next-largest eigenvalue of the adjacency matrix of G. For a d-regular graph, the largest eigenvalue is d and we denote the next-largest absolute value among the eigenvalues by λ. The *Cheeger inequality* for d-regular graphs (see [2, 14]) states that

$$\frac{1}{2}(d - \lambda) \leq h(G) \leq \sqrt{2d(d - \lambda)} \tag{12.3}$$

In particular, $h \to 0$ if and only if $\lambda \to d$. The constant function equal to 1 at every vertex is an eigenvector for the eigenvalue d. The multiplicity of this eigenvalue is the number of connected components of the graph. This is a practical way to check connectedness of Markoff graphs.

In the Markoff case, $d = 3$ and, from the data in [12], λ appears to converge to different values as $p \to \infty$ along the subsequences of primes congruent to 1 mod 4 or 3 mod 4. In both cases, λ seems to remain bounded away from 3, in which case h must remain bounded away from 0. Once the number of vertices $n = p^2 \pm 3p$ is large enough, the inequality that would follow from (12.2) and (12.3), namely

$$\frac{1}{2}(3 - \lambda) \leq \frac{2\sqrt{2n}}{n/3 - 2\sqrt{2n}}$$

must therefore fail, and then the Markoff graph mod p cannot be planar.

For example, consider once again $p = 19$, the first instance where neither Theorem 1.4 nor Theorem 1.5 applies. The number of vertices in this case is

$$n = p^2 - 3p = 304,$$

so that

$$\frac{2\sqrt{2n}}{n/3 - 2\sqrt{2n}} = 0.948\ldots \tag{12.4}$$

It is feasible to compute all the eigenvalues on a personal computer equipped with Pari [5], and the next largest in modulus is approximately $\lambda = 2.873\ldots$ This gives $(3 - \lambda)/2 = 0.0634\ldots$ which is well below (12.4). Thus the spectral method does not apply to $p = 19$. Assuming that a similar spectral gap persists for larger primes congruent to 3 mod 4, the comparison would become favourable to deducing non-planarity once $p \geq 163$, at which point $2\sqrt{2n} \div (n/3 - 2\sqrt{2n}) < 0.06$.

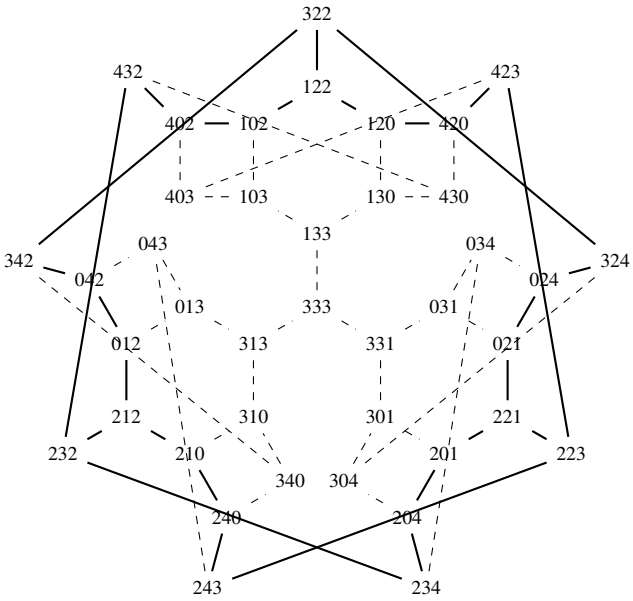


Figure 13.1. The Markoff graph mod 5. The 40 vertices are labelled xyz, where $x^2 + y^2 + z^2 = xyz$ mod 5. The thicker edges illustrate the construction proving Theorem 1.4 for $p = 5$. One can choose $i = \pm 2$ and have $i^2 \equiv -1$ mod 5. Points of the form $(3, 2, 2)$ and $(1, 2, 2)$ play the role of $(2 \pm 2i, 2, 2)$ from Figure 10.1

13. Examples

The first non-planar Markoff graph occurs for $p = 5$. It is drawn (with crossings) in Figure 13.1. The number of squares is

$$s = 3(p - 1)/2 = 6,$$

the number of hexagons is

$$h = p + 1 = 6,$$

and there are no self-edges. Therefore, $V = 40$ and $E = 3V/2 = 60$. The construction of Theorem 1.4 gives cycles of length $2p = 10$. By inspection, there are no cycles of length 7, so one can take $g = 8$ to improve the bounds on the Euler characteristic. There are cycles of length 8, for instance traversing a hexagon and one of its adjacent squares, but these do not bound their own faces. Nevertheless, taking $g = 8$ gives $1/2 \leq -\chi$, which rounds to $1 \leq -\chi$.

For a non-orientable surface, formed from a sphere with n cross-caps, the Euler characteristic is $\chi = 2 - n$, and the bound $1 \leq -\chi$ amounts to $n \geq 3$. For more on

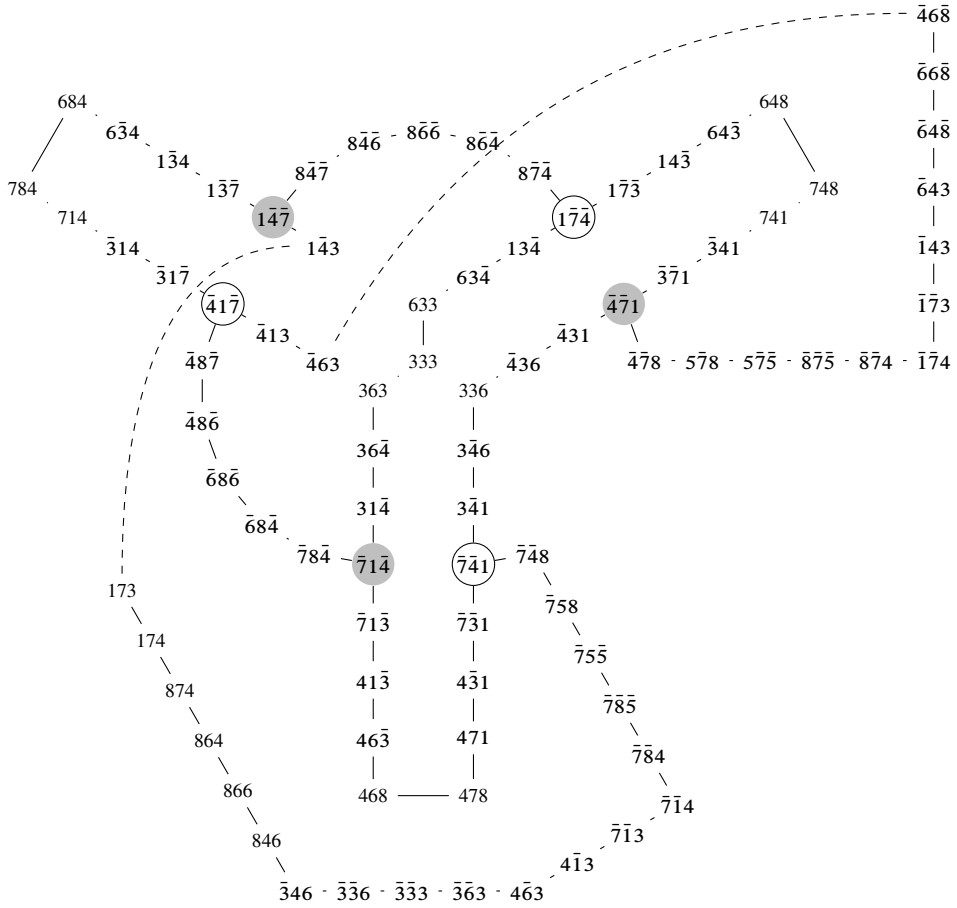


Figure 13.2. The Markoff graph modulo 19 contains a subdivision of the complete bipartite graph joining the even permutations of $(1, -4, -7)$, circled in grey, to the odd permutations, circled in white. Triples (x, y, z) are abbreviated as xyz, and $\overline{x}$ denotes $-x$ mod 19.

cross-caps, see [11, pp. 94–103]. To draw the Markoff graph mod 5 on a surface with $n = 3$, imagine a cross-cap attached over each of the three hexagons in Figure 13.1. For an orientable surface of genus γ, we would have $\chi = 2 - 2\gamma$, hence $\gamma \geq 2$, but this bound does not seem to be attainable.

The example $p = 11$ illustrates both Theorem 1.5 and Proposition 5.1. Carlitz's formula $p^2 - 3p$ (Lemma 2.1) gives 88 vertices in total, which can be thought of as four signed copies of $22 = 16 + 6$. The 16 vertices in this partition form a tree following three steps from any of $(3, 3, 3)$ or its sign changes such as $(8, 8, 3)$. The self-edges from Proposition 5.1 occur at $(3, 4, 6)$, as well as its permutations and sign changes. This limits the branching so that there are only 16 vertices per tree. The

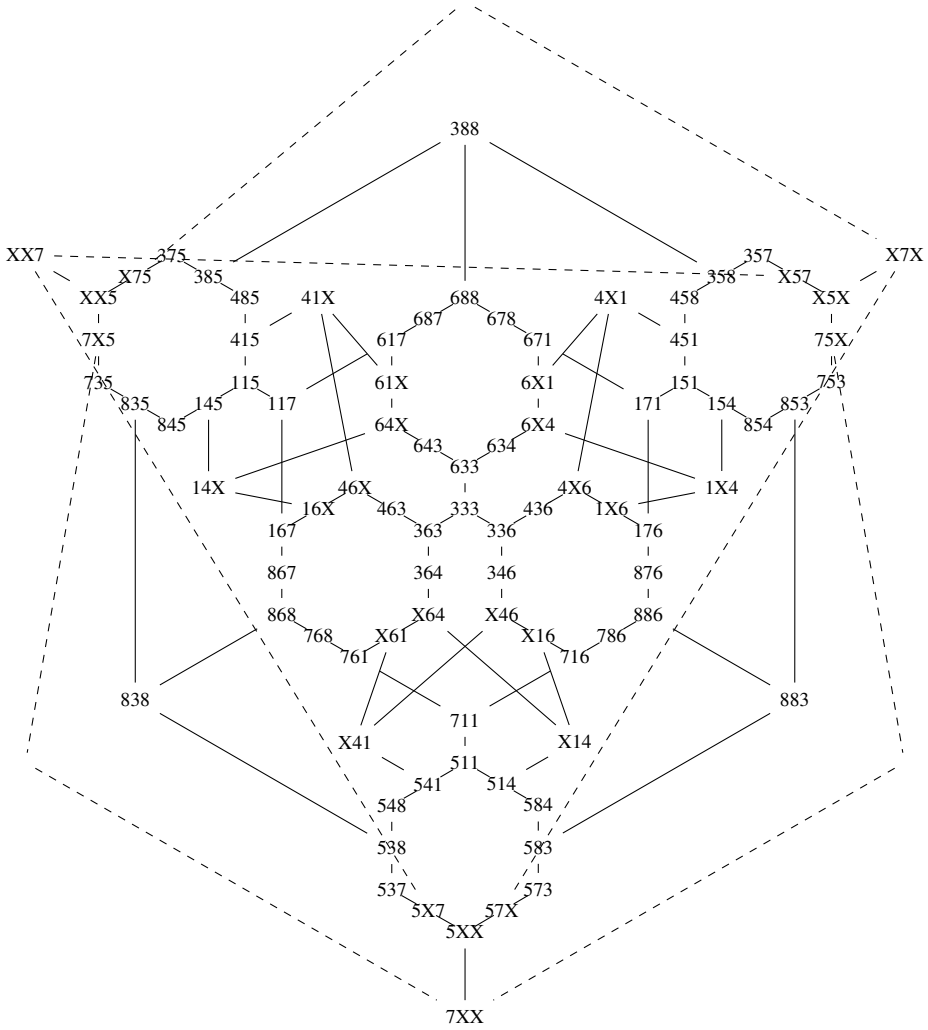


Figure 13.3. The Markoff graph mod 11. Labels xyz abbreviate (x, y, z), and X denotes $10 = -1 \bmod 11$. The dashed edges outline a copy of Figure 11.1. The cage consists of triples with some coordinate equal to 5 or 6. Self-edges occur at permutations and sign changes of $(3, 4, 6)$, for instance $(8, 7, 6)$ or $(3, 7, 5)$.

remaining 6 vertices in $4 \times (16 + 6)$ come from four copies of Figure 11.1, one for each sign change. We have

$$\sqrt{-7} = \sqrt{4} = 2 \bmod 11,$$

and $1/2 = 6$, so the special values $\frac{1}{2}(1 \pm \sqrt{-7})$ are $x = 5$ and $1 - x = 7$.

The cage for $p = 11$ consists of triples with a coordinate equal to ± 5. Indeed, since $z = 0, \pm 2$ do not occur for $p \equiv 3 \bmod 4$, the possible values are $\pm z = 1, 3, 4, 5$. Both signs lead to the same order for ζ^2, where $z = \zeta + \zeta^{-1}$. For $z = 5$, writing $i^2 = -1$, we have

$$\zeta = \frac{z + \sqrt{z^2 - 4}}{2} = \frac{5 + \sqrt{-1}}{2} = 3(-1 + 2i),$$
$$\zeta^2 = 6 + 8i, \quad \zeta^4 = 5 + 8i, \quad \zeta^8 = 5 - 8i = \overline{\zeta^4}.$$

From these values, it follows that $\zeta^{12} = 1$ and no smaller exponent works. Thus ζ has order $12 = p + 1$, which is as large as possible, putting $z = 5$ in the cage.

The first case not covered by either Theorem 1.4 or Theorem 1.5 is $p = 19$, since $-7 \equiv 2^2 \bmod 11$ and the other primes $5 \le p \le 17$ are congruent to $1 \bmod 4$. Figure 13.2 exhibits a copy of $K_{3,3}$ showing that the Markoff graph mod 19 is not planar, as we already know from (3.5). The various paths in Figure 13.2 were found by trial and error after drawing part of the graph.

14. Conclusion

We have shown that the Markoff graph mod 7 is the last of its kind: for $p \neq 2, 3, 7$, these graphs are not planar. Moreover, the Euler characteristic of a surface in which they can be embedded is, in absolute value, at least roughly $p^2/2$ as $p \to \infty$. This non-planarity is consistent with the conjecture that the Markoff graphs form an expander family as $p \to \infty$. For p in various arithmetic progressions, non-planarity can be seen by explicit constructions involving $\sqrt{-1}$ or $\sqrt{-7}$. The general argument is based on variations of the classical Lemma 1.2. These can be thought of either as an upper bound for the number of vertices V of a 3-regular graph embedded in a surface of given Euler characteristic χ, or as a bound for χ given V. The methods are applicable to other examples beyond the Markoff graphs mod p, as long as we have some knowledge of the short cycles.

Lemma 1.2 has some interesting sharp cases for graphs of higher degree. If every vertex has degree d, then $E = dV/2$. We can always take $g = 3$ for graphs without repeated edges. For a planar graph, Euler's formula then implies

$$E \le \frac{g(V-2)}{g-2}, \quad \text{or} \quad V \ge \frac{12}{6-d}.$$

This is achieved for $d = 3$ by the tetrahedron with $V = 4$; for $d = 4$ by the octahedron with $V = 6$; and for $d = 5$ by the icosahedron with $V = 12$. All of the Markoff graphs have the symmetries of a tetrahedron, both rotations and reflections. These act by the four sign changes such as

$$(x, y, z) \mapsto (-x, -y, z),$$

together with permutations of the coordinates. The Markoff moves themselves give further symmetries, which are closely related to the projective linear group $\mathrm{PGL}(2, p)$. Especially for small primes such as $p = 5, 7, 11$, there might be good ways to draw the Markoff graphs on Platonic solids with cross-caps or handles attached.

For example, Figure 1.1 could be folded into a tetrahedron with $(3, 3, 3)$ or one of its sign changes as the centre of each face, or as vertices. One can also see the outermost hexagon in Figure 1.1 or 13.3 as a cross-section of a cube. Another natural home for the Markoff graphs mod p, in view of $\mathrm{PGL}(2, p)$, would be the hyperbolic surfaces defined from related subgroups of $\mathrm{PGL}(2, \mathbb{R})$, or perhaps 3-dimensional hyperbolic models with respect to $\mathrm{PGL}(2, \mathbb{C})$, or the non-congruence modular curves from [10]. It would be interesting to compare Theorem 1.3 with embeddings of minimal complexity, and find the optimal c for which there are sequences of embeddings with $-\chi = (c + o(1))p^2$ as $p \to \infty$.

Finally, we comment on another scaling of the Markoff equation, which is the usual form over the integers:

$$x^2 + y^2 + z^2 = 3xyz.$$

Multiplying each variable by 3 transforms this to the Markoff graphs studied here. The moves are scaled in a compatible way, for instance

$$x \mapsto 3yz - x = \frac{(3y)(3z) - 3x}{3}.$$

For $p \neq 3$, the scaling is invertible and either equation leads to the same Markoff graph mod p. For $p = 3$, the $3xyz$-version of the Markoff equation reduces to

$$x^2 + y^2 + z^2 = 0,$$

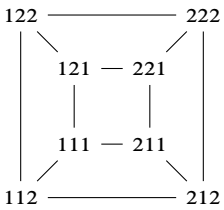
where the cubic term has disappeared. Whereas $(0, 0, 0)$ is the only solution to

$$x^2 + y^2 + z^2 = xyz,$$

this form has 8 other solutions $(\pm 1, \pm 1, \pm 1)$. The moves collapse to sign changes:

$$x \mapsto 3yz - x = -x \bmod 3.$$

In this way, the rescaled Markoff graph mod 3 can be drawn as a cube, giving a more interesting planar example than the empty graph we dismissed earlier:



Acknowledgements. We would like to thank Peter Sarnak, Elena Fuchs, Michael Magee, Eva Bayer-Fluckiger, Martin Stoller, and Maryna Viazovska for their encouragement in this project. We thank Will Sawin for suggesting another approach in case the graph is disconnected (apply the arguments of Section 3 to each component separately and conclude at least one of them is non-planar, though not necessarily the component containing the cage), and Will Chen for suggesting that the surfaces from [10] might give embeddings with $-\chi$ asymptotically as small as possible. We are very grateful to the anonymous referees for their feedback, including the possibility of K_5 occurring as a graph minor. Many thanks also to Tim Browning for advice on the manuscript.

References

[1] M. Aigner, *Markov's theorem and 100 years of the uniqueness conjecture*. Springer, Cham, 2013 Zbl 1276.00006 MR 3098784

[2] N. Alon and V. D. Milman, λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B* **38** (1985), no. 1, 73–88 Zbl 0549.05051 MR 782626

[3] N. Andersen and W. Duke, Markov spectra for modular billiards. *Math. Ann.* **373** (2019), no. 3-4, 1151–1175 Zbl 1433.37027 MR 3953124

[4] A. Baragar, The Markoff equation and equations of Hurwitz. PhD thesis, Brown University, 1991

[5] C. Batut, K. Belabas, D. Benardi, H. Cohen, and M. Olivier, User's guide to PARI-GP, 1998, https://pari.math.u-bordeaux.fr/, visited on 2 February 2024

[6] J. Bourgain, A. Gamburd, and P. Sarnak, Markoff surfaces and strong approximation: 1. 2016, arXiv:1607.01530

[7] L. Carlitz, The number of points on certain cubic surfaces over a finite field. *Boll. Un. Mat. Ital. (3)* **12** (1957), 19–21 Zbl 0077.26105 MR 87673

[8] A. Cayley, On the triple tangent planes of surfaces of the third order. *Cambridge and Dublin Math. J.* **4** (1849), 118–138

[9] A. Cerbu, E. Gunther, M. Magee, and L. Peilen, The cycle structure of a Markoff automorphism over finite fields. *J. Number Theory* **211** (2020), 1–27 Zbl 1459.37096 MR 4074547

[10] W. Chen, Nonabelian level structures, Nielsen equivalence, and Markoff triples. *Ann. of Math. (2)* **199** (2024), no. 1, 301–443 Zbl 07782633 MR 4681147

[11] J. H. Conway, H. Burgiel, and C. Goodman-Strauss, *The symmetries of things*. A K Peters, Wellesley, MA, 2008 Zbl 1173.00001 MR 2410150

[12] M. de Courcy-Ireland and S. Lee, Experiments with the Markoff surface. *Exp. Math.* **31** (2022), no. 3, 814–829 Zbl 1523.11058 MR 4477406

[13] M. de Courcy-Ireland and M. Magee, Kesten–McKay law for the Markoff surface mod p. *Ann. H. Lebesgue* **4** (2021), 227–250 Zbl 07480704 MR 4213160

[14] J. Dodziuk, Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.* **284** (1984), no. 2, 787–794 Zbl 0512.39001 MR 743744

[15] J. Eddy, E. Fuchs, M. Litman, D. Martin, and N. Tripeny,Connectivity of Markoff mod-p graphs and maximal divisors. 2023, arXiv:2308.07579

[16] J. Friedlander and H. Iwaniec, *Opera de cribro*. Amer. Math. Soc. Colloq. Publ. 57, American Mathematical Society, Providence, RI, 2010 Zbl 1226.11099 MR 2647984

[17] E. Fuchs, K. Lauter, M. Litman, and A. Tran, A cryptographic hash function from Markoff triples. *Mathematical Cryptology* **1** (2021), no. 1, 103–121, https://journals.flvc.org/mathcryptology/article/view/129266, visited on 2 February 2024

[18] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*. Fourth edn., The Clarendon Press, Oxford University Press, New York, 1979 Zbl 0086.25803 MR 568909

[19] A. E. Ingham, *The distribution of prime numbers*. Camb. Tracts Math. Math. Phys. 30, Cambridge University Press, London, 1932 Zbl 0006.39701 MR 184920

[20] A. J. Kollár and P. Sarnak, Gap sets for the spectra of cubic graphs. *Comm. Amer. Math. Soc.* **1** (2021), 1–38 Zbl 07750987 MR 4328114

[21] C. Kuratowski, Sur le problème des courbes gauches en topologie. *Fund. Math.* **15** (1930), 271—283 Zbl 56.1141.03

[22] R. J. Lipton and R. E. Tarjan, A separator theorem for planar graphs. *SIAM J. Appl. Math.* **36** (1979), no. 2, 177–189 Zbl 0432.05022 MR 524495

[23] A. Markoff, Sur les formes quadratiques binaires indéfinies. *Math. Ann.* **17** (1880), no. 3, 379–399 Zbl 12.0143.02 MR 1510073

[24] C. Matheus and C. G. Moreira, Fractal geometry of the complement of Lagrange spectrum in Markov spectrum. *Comment. Math. Helv.* **95** (2020), no. 3, 593–633 Zbl 1465.11165 MR 4152626

[25] C. Meiri and D. Puder, The Markoff group of transformations in prime and composite moduli. *Duke Math. J.* **167** (2018), no. 14, 2679–2720 Zbl 1447.11049 MR 3859362

[26] M. Mirzakhani, Counting mapping class group orbits on hyperbolic surfaces. 2016, arXiv:1601.03342

[27] I. Niven, Mathematical notes: A proof of the divergence of $\sum 1/p$. *Amer. Math. Monthly* **78** (1971), no. 3, 272–273 Zbl 0207.35104 MR 1536249

[28] K. Wagner, Über eine Eigenschaft der ebenen Komplexe. *Math. Ann.* **114** (1937), no. 1, 570–590 Zbl 63.0550.01 MR 1513158

Received 24 March 2022.

Matthew de Courcy-Ireland
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; matthew.decourcy-ireland@epfl.ch, matthew.decourcyireland@gmail.com