Journal of Number Theory **87**, 77–95 (2001)
doi:10.1006/jnth.2000.2578, available online at http://www.idealibrary.com on IDEAL®

Markoff Numbers, Principal Ideals and Continued Fraction Expansions

J. O. Button[1]

Wadham College, Oxford University, Oxford, OX1 3PN, United Kingdom
E-mail: button@maths.ox.ac.uk

Communicated by R. C. Vaughan

Received August 18, 1999; published online January 26, 2001

Given any solution triple of natural numbers to the Markoff equation $a^2+b^2+c^2=3abc$, an old problem asks whether the largest number determines the triple uniquely. We show this to be true in a range of cases by considering the factorisation of ideals in certain quadratic number fields, but also exhibit a counterexample for this approach when the question is widened to other numbers.

0. INTRODUCTION

The Diophantine equation

$$a^2+b^2+c^2=3abc$$

has a long and interesting history, having been first considered in 1879 by A. A. Markoff. He showed that all solution triples of natural numbers occur as vertices of a binary tree, with simple rules which can be applied in order to inductively build up these solutions. A problem that seems to be almost as old is referred to as the Markoff conjecture, or the unicity conjecture. It states: does the largest element of any solution triple determine the triple uniquely? This is known for prime numbers, so that if (a, b, c) and (a', b', c) are both triples of positive integers solving this equation with $a \leqslant b \leqslant c$ and $a' \leqslant b' \leqslant c$ then $a=a'$ and $b=b'$ provided c is prime.

In this paper we extend these results to find many more Markoff numbers for which the conjecture is true. We consider odd numbers and first establish this for prime powers. Then more generally we show uniqueness for numbers which are products of a prime power and a "small" factor k, where k is any number such that the conjecture is known to be true for values of c up to k^4.

[1] Current address: Selwyn College, Cambridge University, Cambridge, CB3 9DQ, United Kingdom. E-mail: jb128@dpmms.cam.ac.uk.

0022-314X/01 $35.00

Computer checks thus far have established the conjecture up to 10^{140}, so our "small" factor k can be anything less than 10^{35}.

Uniqueness proofs for particular values of c are obtained by looking at an order in the quadratic number field $\mathbb{Q}(\sqrt{9c^2-4})$ (which will be the ring of integers if $9c^2-4$ is square free, but we do not restrict ourselves only to this case), and by considering the factorisation into prime ideals of the principal ideal generated by c^2. We next look at values c where $9c^2-4$ has few factors, and we show uniqueness for a range of such numbers including the cases where either $3c-2$ or $3c+2$ is prime. Finally, however, we finish with some words of warning and a counterexample. Not a counterexample to the conjecture, but once it has been reformulated into a question about ideals in this quadratic number field then exactly the same question can be asked about any odd number c. We display such a c for which the question is false, thus showing that if the conjecture is true for all Markoff numbers then any proof would need to use explicit properties of Markoff numbers somewhere in its ingredients.

1. EXTENDING UNIQUENESS

Given a solution in positive integers of the Diophantine equation

$$a^2+b^2+c^2=3abc,$$

where $a\leqslant b\leqslant c$ and c is odd, we set $D=9c^2-4$ and $R=\mathbb{Z}+\omega\mathbb{Z}$ where $\omega=\frac{3c-2+\sqrt{D}}{2}$. Thus R is an order in the quadratic number field $\mathbb{Q}(\sqrt{D})$, which has signed norm $\|x+\sqrt{D}\,y\|=x^2-Dy^2$. Then we have (see [1] or [2])

THEOREM 1. *The odd integer c is the maximal element of just one Markoff triple (or more briefly c is unique) if and only if there exists exactly one pair of principal ideals $\{\beta R, \bar{\beta}R\}$ in R such that $\beta\bar{\beta}=-c^2$.*

As an immediate corollary we obtain uniqueness for prime Markoff numbers and also, as pointed out by Colin Maclachlan, for prime powers too:

Suppose $c=p^k$ for some prime p then we have a factorisation in R of the ideal pR which is of the form $\mathscr{P}\bar{\mathscr{P}}$, and the factorisation is unique in this order even if R is not the full ring of integers, because p is coprime to D. If there exists an element β of R with $c^2R=\beta\bar{\beta}R$ then $\beta R=\mathscr{P}^l\bar{\mathscr{P}}^{k-l}R$ by unique factorisation of ideals. But βR is a primitive ideal, so we must have $l=k$ to avoid any factors pR.

Now suppose that there exist distinct pairs of principal ideals generated by elements β and β' with $\beta\bar{\beta} = \beta'\bar{\beta'} = -c^2$ (so that c fails to be unique). Factorising c^2R gives integers p, q $(\neq 1, c)$ with $pq = c$ and

$$\beta R = \mathscr{P}^2\mathscr{Q}^2, \qquad \beta' R = \mathscr{P}^2\bar{\mathscr{Q}}^2$$

where $pR = \mathscr{P}\bar{\mathscr{P}}$ and $qR = \mathscr{Q}\bar{\mathscr{Q}}$.

Note that pR and qR do split into a unique product of prime ideals and that $(p, q) = 1$ as βR and $\beta' R$ are primitive ideals. So if c is not unique then we have an integer $p \mid c$ (and without loss of generality $1 < p^2 < c$) with $\mathscr{P}^4 \overset{+}{\sim} R$ where $\overset{+}{\sim}$ is narrow class equivalence. In order to exploit this fact we represent ideals using Hermite normal form. We know (details are in [4]) that any integral ideal $\mathscr{A}$ of the order R can be written

$$\mathscr{A} = a\mathbb{Z} + \left(\frac{-b + k\sqrt{D}}{2}\right)\mathbb{Z}$$

where $D = 9c^2 - 4$ is $1 \bmod 4$, a is the smallest positive integer in $\mathscr{A}$, b is only defined up to multiples of $2a$ and k (which has the same parity as b) divides a. If $\mathscr{A}$ is a primitive ideal (i.e. $\mathscr{A}/n$ is not an integral ideal of R for any integer $n > 1$) then we can take $k = 1$ and the norm of $\mathscr{A}$ is a, with $4a$ dividing $b^2 - D$. The invertible ideals in R are precisely the ones with $(a, b, d) = 1$, where $D = b^2 - 4ad$.

From above, in order to look at uniqueness of various Markoff numbers, we need to be able to recognise when two invertible ideals of R are equivalent in the class group, namely the set of equivalence classes of invertible ideals of R where $\mathscr{I}$ and $\mathscr{J}$ are equivalent if there exists $\alpha \in \mathbb{Q}(\sqrt{D})$, $\alpha \neq 0$, with $\mathscr{J} = \alpha\mathscr{I}$. If D is square-free then all ideals are invertible. This can be done very neatly by representing our ideals by quadratic irrationals and using the tool of continued fractions. We take care to differentiate between equivalence in the narrow class group and in the full class group, as in our case they will be different.

THEOREM 2. *For any positive D equal to* $1 \bmod 4$ (*except* 1), *let R be the order $\mathbb{Z} + \omega\mathbb{Z}$ in the quadratic field $\mathbb{Q}(\sqrt{D})$, let the set of primitive quadratic irrationals*

$$Q = \left\{\frac{b + \sqrt{D}}{2a} : a \neq 0,\ 4a \mid D - b^2 \text{ and } \gcd(a, b, d) = 1 \text{ where } 4ad = b^2 - D\right\}$$

and let I equal the set of invertible ideals of R. Then form the quotient set

$$\mathscr{Q}/\sim,$$

where

$$\frac{b+\sqrt{D}}{2a} \sim \frac{b'+\sqrt{D}}{2a'}$$

if there exists an element $A \in PSL(2, \mathbb{Z})$ *with*

$$A\left(\frac{b+\sqrt{D}}{2a}\right) = \frac{b'+\sqrt{D}}{2a'},$$

where the action of $PSL(2, \mathbb{Z})$ *is the usual one corresponding to Möbius transformations.*

We can form the narrow and wide class groups $Cl^+(D)$ *and* $Cl(D)$ *by quotienting I in the usual way. We then have a well defined bijection between* $\mathcal{Q}/\sim$ *and* $Cl^+(D)$ *given by the map on equivalence classes induced by* Φ *where*

$$\Phi\left(\frac{b+\sqrt{D}}{2a}\right) = \left(a\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z}\right)\alpha.$$

Here α *is any element in R with the sign of the norm of* α *the same as the sign of a.*

Also the inverse in $Cl^+(D)$ *of the equivalence class represented by the ideal*

$$\mathcal{A} = a\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z}$$

is

$$\bar{\mathcal{A}} = a\mathbb{Z} + \left(\frac{-b+\sqrt{D}}{2}\right)\mathbb{Z}.$$

Proof of 2. See [4], pages 223–224. ∎

LEMMA 3. *If*

$$\mathcal{A} = a\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z},$$

$$\mathcal{B} = a'\mathbb{Z} + \left(\frac{b'+\sqrt{D}}{2}\right)\mathbb{Z}$$

are any two invertible ideals in R (where we can and do take $a, a' > 0$), then they are equivalent in $Cl^+(D)$ if the continued fractions of $\frac{b+\sqrt{D}}{2a}$ and $\frac{b'+\sqrt{D}}{2a'}$ are the same, after knocking off an even number of terms in total from the two continued fraction expansions. The ideals are equivalent in $Cl(D)$ but not in $Cl^+(D)$ (i.e. if one ideal is the product of the other one with a principle fractional ideal generated by an element of $\mathbb{Q}(\sqrt{D})$ of negative norm, for which we will use the notation $\mathscr{A} \sim \mathscr{B}$) if the continued fraction expansions agree after removing an odd number of terms.

Proof of 3. Straight from Theorem 2, along with the fact that two irrational numbers are equivalent in $PSL(2, \mathbb{Z})$ if and only if their continued fraction expansions eventually agree and the map $z \to \frac{1}{z}$ (which has determinant -1 as a Möbius transformation) has been applied an even number of times. Therefore we are counting the numbers of terms removed to keep track of this parity. ▮

For future use, we also state the following well known properties of continued fractions of quadratic irrationals.

LEMMA 4.

(1) *The continued fraction expansion is eventually periodic.*

(2) *The continued fraction expansion of $\frac{b+\sqrt{D}}{2a}$ is purely periodic if and only if*

$$\frac{b+\sqrt{D}}{2a} > 1 \qquad \text{and} \qquad 0 > \frac{b-\sqrt{D}}{2a} > -1$$

(which means that $a > 0$).

(3) *If $\tau = \frac{b+\sqrt{D}}{2a}$ has a purely periodic continued fraction expansion $[\overline{x_1, \ldots, x_n}]$ then $-1/\bar{\tau} = [\overline{x_n, \ldots, x_1}]$.*

We now apply these facts to the case where $D = 9c^2 - 4$.

LEMMA 5. *Consider the invertible ideal*

$$\mathscr{A} = a\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z} \qquad (\text{where } a > 0, (a, b, d) = 1)$$

and the continued fraction expansion of $\sigma = \frac{b+\sqrt{D}}{2a}$.

(1) *It is principal if and only if the continued fraction expansion of σ is $[\overline{3c-2, 1}]$ after removing a finite number of terms. If an even number are removed then the ideal is generated by an element of positive norm, and it is generated by an element of negative norm in the odd case.*

(2) *$\mathscr{A}^2$ is principal, generated by an element of negative norm, if and only if the periodic part of the continued fraction expansion of σ is symmetric, i.e. if the purely periodic part is $[\overline{x_1, ..., x_n}]$ then we can choose x_1 appropriately so that $x_k = x_{n+1-k}$ for $1 \leqslant k \leqslant n/2$.*

(3) *$\mathscr{A}^2$ is principal and generated by an element of positive norm if and only if the periodic part of the continued fraction expansion of σ is "anti-symmetric", namely we write the purely periodic part (with appropriate x_1) as $[\overline{x_1, ..., x_n}]$ and then we have $x_k = x_{n+2-k}$ for $2 \leqslant k \leqslant n/2$, but x_1 and $x_{(n/2)+1}$ can be chosen arbitrarily.*

Proof of 5.

(1) Use Lemma 3 and the fact that the order

$$R = \mathbb{Z} + \left(\frac{3c - 2 + \sqrt{D}}{2}\right)\mathbb{Z}$$

with $\frac{3c-2+\sqrt{D}}{2} = [\overline{3c-2, 1}]$.

(2) Setting

$$\tau = \frac{B + \sqrt{D}}{2A} = [\overline{x_1, ..., x_n}]$$

to be the purely periodic part, if this continued fraction expansion is symmetric then Lemma 4 (3) tells us that $\tau = -1/\bar{\tau}$. Thus $D = B^2 + 4A^2$ and considering the ideal

$$\mathscr{B} = A\mathbb{Z} + \left(\frac{B + \sqrt{D}}{2}\right)\mathbb{Z}$$

with inverse in the narrow class group

$$\bar{\mathscr{B}} = A\mathbb{Z} + \left(\frac{-B + \sqrt{D}}{2}\right)\mathbb{Z}$$

we see that $\frac{B+\sqrt{D}}{2A}$ and $\frac{-B+\sqrt{D}}{2A}$ are related by the map $z \to \frac{1}{z}$. Thus $\mathscr{B} \sim \bar{\mathscr{B}}$ and thus $\mathscr{A} \sim \bar{\mathscr{A}}$ as we either have $\mathscr{A} \overset{+}{\sim} \mathscr{B}$ or $\mathscr{A} \overset{-}{\sim} \mathscr{B}$.

Conversely given $\mathscr{A}$ with $\mathscr{A} \overset{-}{\sim} \bar{\mathscr{A}}$, find the same ideal $\mathscr{B}$ as above, along with

$$\tau = \frac{B + \sqrt{D}}{2A} = [\overline{x_1, ..., x_n}]$$

which is again the purely periodic part of the continued fraction expansion. Consider the result of finding out the continued fraction expansion of τ directly by imagining two columns: the first column has as its k-th entry the quadratic irrational $\tau_k = [\overline{x_k, ..., x_{n+k-1}}]$ and the second has the quadratic irrational $\sigma_k = \text{frac } \tau_k$, so that $1 > \sigma_k > 0$ and $\tau_{k+1} = 1/\sigma_k$. Thus τ appears in the first column as τ_1 $(= \tau_{n+1} = \cdots)$.

The ideal $\bar{\mathscr{B}}$ is represented by the number $\sigma = \frac{-B+\sqrt{D}}{2A}$, with $1 > \sigma > 0$ and $-1 > \bar{\sigma}$. Thus there exists an integer m with

$$\tau' = \frac{2Am - B + \sqrt{D}}{2A} > 1 \qquad \text{and} \qquad 0 > \frac{2Am - B - \sqrt{D}}{2A} > -1$$

so that τ' has purely periodic continued fraction expansion. But τ' also represents the ideal $\bar{\mathscr{B}}$ and so τ' has the same purely periodic part of the continued fraction expansion as τ. Thus $\tau' = \tau_k$ for some k, $\sigma = \sigma_k$ and $\tau_1 = -\bar{\sigma}_k$, $\sigma_1 = -\bar{\tau}_k$. We then obtain $\tau_{1+l} = -\bar{\sigma}_{k-l}$ and $\sigma_{1+l} = -\bar{\tau}_{k-l}$ as $x = -\bar{y}$ implies $1/x = -1/\bar{y}$. Now k must be even $(= 2j$ say) since $\mathscr{B} \sim \bar{\mathscr{B}}$, so we have

$$\tau_{j+1} = \frac{B' + \sqrt{D}}{2A'} \quad \text{(say)} \qquad \text{and} \qquad \sigma_j = \frac{-B' + \sqrt{D}}{2A'}.$$

Hence $\tau_{j+1} = -1/\bar{\tau}_{j+1}$ and thus the purely periodic part of the continued fraction expansion is symmetric.

(3) We run through the same process as in (2). If

$$\tau = [\overline{x_1, ..., x_n}] = [\overline{x_1, x_n, x_{n-1}, ..., x_2}]$$

then

$$\frac{1}{\tau - x_1} = [\overline{x_n, x_{n-1}, ..., x_1}] = -\frac{1}{\bar{\tau}}$$

and so $\tau - x_1 = -\bar{\tau}$ where $\tau - x_1$ represents the ideal $\mathscr{B}$ and $-\bar{\tau}$ represents the ideal $\bar{\mathscr{B}}$. Thus $\mathscr{B} \overset{+}{\sim} \bar{\mathscr{B}}$ and $\mathscr{A} \overset{+}{\sim} \bar{\mathscr{A}}$.

In going the other way the argument in (2) applies exactly except that k must be odd this time, say $2j-1$, so that we find

$$\tau_j = \frac{B' + \sqrt{D}}{2A'}, \qquad \sigma_j = \frac{-B' + \sqrt{D}}{2A'}$$

and then $[\overline{x_k, ..., x_{n+k-1}}]$ is of the required form. Note that in this case $A' \mid B'$ and hence $A' \mid D$, so that we can find all ideal classes whose square is the identity in the narrow class group by considering all factors of D.

If D is not square-free, then if we set $B' = kA'$ so that $D = A'(k^2A' - 4C')$ for some C', the requirement that $(A', B', C') = 1$ in order to ensure an invertible ideal means that we only need consider factors of D where A' and D/A' are coprime. ∎

COROLLARY 6. *If the Markoff number c has the property that for all factorisations $c = pq$, where p and q are coprime with $p < q$, we have $q > p^3$ then c is unique.*

Proof of 6. If c is not unique then we can find coprime p and q (where $p < q$) with $c = pq$, $pR = \mathscr{P}\bar{\mathscr{P}}$ and $\mathscr{P}^4 \overset{+}{\sim} R$. Now $\mathscr{P}^4$ is a primitive ideal and is represented in Hermite normal form thus:

$$\mathscr{P}^4 = p^4\mathbb{Z} + \left(\frac{b + \sqrt{D}}{2}\right)\mathbb{Z}$$

for whatever b, and also by the quadratic irrational $\tau = \frac{b+\sqrt{D}}{2p^4}$.

But

$$\frac{\sqrt{D}}{p^4} > \frac{2c}{p^4} > 2$$

by the hypothesis and hence $\tau - \bar{\tau} > 2$.

So the appropriate choice of b ensures that $\tau > 1$ and $0 > \bar{\tau} > -1$, giving a purely periodic continued fraction expansion which must be $[\overline{3c-2, 1}]$ by Lemma 5 (1). But this implies that $p^4 = 1$. ∎

Thus for instance Markoff numbers of the form $5p^k$ (p prime) are unique as we are fine once $p^k > 125$, so we just need to check uniqueness up to $c = 625$. Or using the fact that we have uniqueness for $c \leqslant 10^{140}$ (see [1]) then any c of the form ap^l, where $a \leqslant 10^{35}$ and p is prime, must be unique.

Whenever we find c as the maximal element of a Markoff triple (a, b, c) then we know we have a pair of principal ideals βR and $\bar{\beta}R$ with $\beta\bar{\beta} = -c^2$. We can take

$$\beta, \bar{\beta} = \frac{X}{2} \pm \frac{Y}{2}\sqrt{D},$$

where $X = 3ac - 2b$ and $Y = a$. But the ideal cR breaks up into conjugate ideals, $cR = \mathscr{C}_+\mathscr{C}_-$, and so we can set $\beta R = \mathscr{C}_+^2$ and $\bar{\beta}R = \mathscr{C}_-^2$.

When expressed in Hermite normal form we will have

$$\mathscr{C}_\pm = c\mathbb{Z} + \left(\frac{\pm x + \sqrt{D}}{2}\right)\mathbb{Z},$$

where we know $x^2 \equiv D \bmod 4c$, and x is defined up to multiples of $2c$. However, we might have many possible choices for x if c has lots of factors. But on noting $X^2 - Da^2 = -4c^2$ we can obtain a possible value of x from X by multiplying X with α, which is the inverse of a in the group of units modulo c. Then setting $x = 3c - 2b\alpha$ we have

$$\begin{aligned} x^2 &\equiv 9c^2 + 4b^2\alpha^2 \bmod 4c \\ &\equiv D + 4(1 + b^2\alpha^2) \bmod 4c. \end{aligned}$$

But $c \mid a^2 + b^2$ and thus also divides $1 + b^2\alpha^2$ as $a\alpha \equiv D \bmod 4c$, thus we do have an appropriate candidate for x with $x^2 \equiv D \bmod 4c$. In fact this particular value of x does give us the correct ideal.

THEOREM 7. *If*

$$\mathscr{C}_+ = c\mathbb{Z} + \left(\frac{x + \sqrt{D}}{2}\right)\mathbb{Z}$$

with x as above, then $\mathscr{C}_+^2$ is principal and generated by the element

$$\beta = \frac{X}{2} + \frac{Y}{2}\sqrt{D}.$$

Proof of 7. We know that $(a, c) = 1$ and so we can find α and γ with $a\alpha + c\gamma = 1$. But α is only unique up to multiples of c and γ can be changed by adding any multiple of a. Thus we can find particular values of α and γ with c dividing γ. We choose this value of α when we take $x = 3c - 2b\alpha$ (noting that this does not change the ideal). Setting l so that $x^2 = D + 4lc$ we have

$$1 - 3\alpha bc + b^2\alpha^2 = lc.$$

We multiply through by a^2, and using $a\alpha + c\gamma = 1$ leads us to

$$3abc\gamma - c + b^2c\gamma^2 - 2b^2\gamma = la^2.$$

Thus c divides la^2 and so $l \equiv 0 \bmod c$.

We use the following lemma to find the square of an ideal which is in Hermite normal form.

LEMMA 8. *If*

$$I = a\mathbb{Z} + \frac{b + \sqrt{D}}{2}\mathbb{Z}$$

is any ideal in the order R with $b^2 = D - 4al$, *then*

$$I^2 = d\left(A\mathbb{Z} + \frac{B+\sqrt{D}}{2}\mathbb{Z}\right),$$

where, if $d = \gcd(a, b)$ *and* v, w *are such that* $va - wb = d$, *we have* $A = a^2/d^2$ *and* $B = b + 2awl/d$

Proof of 8.

See Ref. [4] pp. 240–241, where we just set the two ideals to be equal. ∎

Now we use the particular value of x chosen above when applying Lemma 8 to square the ideal $\mathscr{C}_+$, and noting $(c, x) = 1$ and $c \mid l$, we obtain

$$\mathscr{C}_+^2 = c^2\mathbb{Z} + \left(\frac{B+\sqrt{D}}{2}\right)\mathbb{Z},$$

where $B = x + 2cwl$. But B is only defined up to multiples of $2c^2$, and $2c^2 \mid 2cwl$, so

$$\mathscr{C}_+^2 = c^2\mathbb{Z} + \left(\frac{x+\sqrt{D}}{2}\right)\mathbb{Z}.$$

In order to show the ideal $\mathscr{C}_+^2$ (of norm c^2) is principal, we merely need to show that the element $\frac{X}{2} + \frac{Y}{2}\sqrt{D}$ (of norm $-c^2$) is in $\mathscr{C}_+^2$.

But

$$Y\left(\frac{x+\sqrt{D}}{2}\right) - \frac{X}{2} - \frac{Y}{2}\sqrt{D} = b(1 - a\alpha)$$

which is divisible by c^2. ∎

We can relate the above to the classical theory: with c the maximal element of any Markoff triple (a, b, c) and k defined by $ak \equiv b \pmod{c}$, $0 \leqslant k < c$, we can form the quadratic form

$$F(x, y) = \alpha x^2 + \beta xy + \gamma y^2 \qquad (x, y \in \mathbb{Z}),$$

where $\alpha = c$, $\beta = 3c - 2k$ and the discriminant $\beta^2 - 4\alpha\gamma = D$. These forms (along with their multiples, and equivalent forms thereof) are characterised by the property that if we set

$$\mu = \inf |F(x, y)| \qquad \text{with} \quad (x, y) \neq (0, 0)$$

then $9\mu^2 > \beta^2 - 4\alpha\gamma$ (see [3]).

We can also look at the roots $\theta_\pm$ of $F(x, 1)$:

$$\theta_\pm = \frac{-(3c-2k) \pm \sqrt{D}}{2c}$$

and can see that θ_+ is precisely the quadratic irrational representing the ideal $\mathscr{C}_-$. These roots are the irrationals θ with the worst possible approximations by rationals in the following sense: we define

$$v(\theta) = \liminf_{q \to 0} q\,|q\theta - p|,$$

where for each positive integer q we choose p to be the nearest integer to $q\theta$. Then $v(\theta) > 1/3$ if and only if θ is equivalent under $PSL(2, \mathbb{Z})$ to a root of a Markoff form (see [3]). The periodic parts of the continued fraction expansion of these roots are of a special form. They are made up only of pairs of ones and of twos, which appear in patterns according to the Dixon rules. See [6, 7 and 5]. Also they must be symmetric, by Theorem 7 and Lemma 5(2).

Thus taking the quadratic irrational $\frac{x+\sqrt{D}}{2c}$ corresponding to the ideal $\mathscr{C}_+$, where we can calculate x up to multiples of $2c$ as above, we then ensure that $c < x \leqslant 3c-1$ to obtain a purely periodic continued fraction expansion (except in the case $x = c = 1$). Performing the operations of subtraction by 1 or 2 followed by $z \mapsto 1/z$ on the continued fraction, and arranging the resulting quadratic irrationals in two columns, as in the proof of Lemma 5(2), gives us a lot of information. For instance, we know we will at some point come to the number $\frac{B+\sqrt{D}}{2A}$ where $D = B^2 + 4A^2$, and this is the starting point of the symmetry of the continued fraction expansion. In fact, halfway through this expansion we will come to $\frac{B'+\sqrt{D}}{2A'}$ with $(B')^2 + 4(A')^2$ a different representation of D as a sum of two squares, as this point can also be thought of as the beginning of the symmetry. We will also come across the number $\frac{x'+\sqrt{D}}{2c}$ where x' is a perfectly good alternative to x (where we can think of x' as coming from Theorem 7 with a and b swapped round). But given the relation $\tau_{1+l} = \overline{\sigma_{k-l}}$ in the notation of Lemma 5(2) between the two columns, this means that $\frac{-x'+\sqrt{D}}{2c}$ is equal to $\frac{x+\sqrt{D}}{2c}$ minus one or two, depending on the value of the continued fraction at that point. Using the inequalities $c < x, x'$ means that unless we have $\frac{1+\sqrt{5}}{2} = [\bar{1}]$, we obtain $x + x' = 4c$ and when $2c$ appears as the denominator in the continued fraction, it is followed up by subtracting two.

We now return to the attempt to extend uniqueness of Markoff numbers by looking at the factorisation of D, and we will see that if $D = (3c-2)(3c+2)$ has few factors then in many cases the uniqueness of c will follow. Recall that if c is not unique, we have $pR = \mathscr{P}\overline{\mathscr{P}}$ with $(\mathscr{P}^2)^2 \overset{+}{\sim} R$.

Therefore the periodic part of the continued fraction expansion of $\mathscr{P}^2$ is of the form in Lemma 5(3). But as we can take $p<q$, we have $p^2<c$ so that if

$$\mathscr{P}^2 = p^2\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z}$$

for some b, the quadratic irrational

$$\tau = \frac{b+\sqrt{D}}{2p^2} \qquad \text{has} \quad \tau - \bar{\tau} = \frac{\sqrt{D}}{p^2} > 2,$$

so as in Corollary 6 the correct choice of b makes τ have a purely periodic continued fraction expansion. But the comment after Lemma 5(3) implies that there must exist a factor α of D so that if we write out the continued fraction expansion of the quadratic irrational

$$\sigma = \frac{k\alpha+\sqrt{D}}{2\alpha} \qquad \text{where} \quad D = k^2\alpha^2 + 4l\alpha \quad \text{with} \ (\alpha, l) = 1$$

then the expansion will be purely periodic, and we will come across a quadratic irrational somewhere in this repeating cycle which has a denominator equal to $2p^2$.

For which α does σ have a purely periodic continued fraction expansion? We would need to find $k \in \mathbb{N}$, which must be odd, such that

$$\frac{k\alpha+\sqrt{D}}{2\alpha} > 1 \qquad \text{and} \qquad 0 < \frac{\sqrt{D}-k\alpha}{2\alpha} < 1.$$

Hence we want $k > 2-\sqrt{D}/\alpha$ and $\sqrt{D}/\alpha > k > \sqrt{D}/\alpha - 2$. First suppose that $\alpha < \sqrt{D}$, then the largest odd integer below $\sqrt{D}/\alpha$ will work. So in this case we are guaranteed a purely periodic expansion. If however $\sqrt{D} < \alpha$ then no k will do.

The idea now is to eliminate those D whose factors α all give rise to a quadratic irrational with a short purely periodic part of its continued fraction expansion.

THEOREM 9. *Given any factorisation $\alpha\alpha'$ of $D = 9c^2 - 4$ with $\alpha < \alpha'$ and α coprime to α', set $\sigma = (k\alpha + \sqrt{D})/2\alpha$ for a suitable odd k, so that σ is the purely periodic quadratic irrational representing an invertible quadratic ideal whose square is the identity. If for every such factorisation σ has period 2 or period 4 then c is unique.*

Proof of 9. We can assume that $\sigma = [\overline{a, r, b, r}]$ for some $a, b, r \in \mathbb{N}$ (so that $a = b$ if the period is 2). This means that

$$\frac{1}{\dfrac{1}{\sigma - a} - r} = b + \frac{\sigma}{r\sigma + 1},$$

and on untangling this we find that σ satisfies the quadratic equation

$$r(rb+2)\,x^2 - ra(rb+2)\,x - rab - a - b = 0.$$

This has discriminant

$$\Delta = (rb+2)(ra+2)\,r(rab+2a+2b) = X(X-4).$$

Thus

$$\sigma = \frac{k\alpha + \sqrt{D}}{2\alpha} = \frac{ar(br+2) + \sqrt{\Delta}}{2r(br+2)}$$

so that $k = a$ and there exists $n, m \in \mathbb{N}$ with

$$\alpha n = r(br+2)\,m \qquad \text{and} \qquad Dn^2 = \Delta m^2,$$

where n and m are coprime.

We first show that $m = 1$. We consider all quadratic irrationals that feature in the continued fraction expansion of σ and form two columns, just as in the proof of Lemma 5(2), so that the first column begins with σ and the second column starts with $\sigma - a$, which is between 0 and 1. We then fill in the remaining entries in the two columns, so that each column consists of four different quadratic irrationals until we return to σ (or only two, if σ has period two). We have two different representations of this pair of columns, one formed by writing $\sigma = \frac{k\alpha + \sqrt{D}}{2\alpha}$, which we call the D-table, and we also have the Δ-table, obtained by writing σ in terms of a, b, r and Δ.

In the D-table, four denominators appear. In the first and third stages we find twice α and then twice β, where β will also be a non-trivial factor of D, obtained halfway through the expansion. But due to the symmetry of the continued fraction expansion obtained from Lemma 5(3), the second and the fourth denominators must be equal, and both must be $2p^2$ as this is the only place where that term can appear.

Looking at the Δ-table, we find that the four denominators are $r(br+2)$ and $r(ar+2)$ corresponding to α and β, and then $2(rab+a+b)$ appears in the other two positions. It must be the case that multiplying every

denominator and numerator in the Δ-table by m, and in the D-table by n makes the corresponding denominators and numerators identical. Thus

$$\alpha n = mr(br+2) \qquad \text{and} \qquad p^2 n = m(rab+a+b).$$

But then $m \mid \alpha$ and $m \mid p^2$, thus $m = 1$, 2 or 4, but α and p^2 are odd. So $m = 1$.

We now finish off the proof in the case where $n = 1$. This gives

$$\alpha = r(br+2), \qquad \beta = r(ar+2) \qquad \text{and} \qquad p^2 = rab+a+b.$$

But α and β divide D, all of whose factors are 1 mod 4, so $r \equiv 1 \bmod 4$ and $ar+2$, $br+2 \equiv 1 \bmod 4$, giving $a, b \equiv -1 \bmod 4$. But then $p^2 \equiv -1 \bmod 4$ which cannot be true.

However this argument may not work if $n > 1$. We will show that only one such exception occurs, which can be easily dealt with. We know that $n^2 D = \Delta = X(X-4)$, so we write $n = st$ with $X = us^2$, $X-4 = vt^2$ and thus we have $(s, t) = 1$. Then $D = uv$, so from $us^2 - vt^2 = 4$ we obtain

$$u^2 s^2 - Dt^2 = 4u.$$

If this equation is satisfied in integers then we have an element

$$\gamma = \frac{us + t\sqrt{D}}{2} \in R \qquad \text{with norm } u \text{ dividing } D,$$

giving rise to a principal ideal γR. If we can show that this ideal will be invertible then we can use Lemma 5 to investigate all possibilities. First note that γR is primitive; otherwise we have a common factor k dividing us and t, which will also divide X and $X-4$. Thus $k = 1$, 2 or 4, but u is odd and $(s, t) = 1$ or 2, giving $k = 1$ or 2. Hence γR is primitive, and can be represented in Hermite normal form as

$$\gamma R = u\mathbb{Z} + \left(\frac{b+\sqrt{D}}{2}\right)\mathbb{Z} \qquad \text{with} \quad b^2 = D - 4uc$$

for some $c \in \mathbb{Z}$. As $\frac{b+\sqrt{D}}{2} \in \gamma R$, we must have $\frac{\alpha+\beta\sqrt{D}}{2} \in R$ such that

$$\left(\frac{\alpha+\beta\sqrt{D}}{2}\right)\left(\frac{us+t\sqrt{D}}{2}\right) = \frac{b+\sqrt{D}}{2}$$

and in particular, $b = \alpha us + D\beta t$.

If γR is not invertible, there will exist an odd prime p dividing u, b and c with $p^2 \mid D$. Set $u = p^k U$ and $D = p^l d$, where $(p, U) = (p, d) = 1$. Then

$$p^{2k} U^2 s^2 - p^l t^2 d = 4p^k U.$$

We must have $l \geqslant k$, or else p^{k-l} divides $t^2 d$, which cannot happen as $(u, t) = 1$ from above. But also $k \geqslant l$, because p does not divide $4U$. Thus $k = l$ and p^k divides u and D, thus p^k also divides b from the expression for b above. Putting this back into $b^2 + 4uc = D$ gives

$$B^2 p^k + 4Uc = d \qquad \text{where} \quad b = Bp^k.$$

But as $p \mid c$, p must divide d too which is not true.

Now we know we are looking at invertible ideals, so in order to find out whether there exists a principal ideal with norm u, we use Lemma 5. If so there will be a quadratic irrational with denominator $2u$ whose continued fraction expansion will have $[\overline{3c-2, 1}]$ as its periodic part. First note that if $u \leqslant 3c - 1 < \sqrt{D}$ then, from the comment before Theorem 9, the quadratic irrational will be purely periodic. However the only denominators here will give rise to elements with norm 1 (i.e. units) or elements with negative norm $-(3c-2)$. But if $3c \leqslant u \leqslant D$ then $v < \sqrt{D}$ where $D = uv$, and the quotient $\sqrt{D}/\gamma$, which has norm $-v$, is equal to $\frac{-tv + s\sqrt{D}}{2}$ and so is an element of R. This gives rise to just two pairs of possibilities for u: either 1 and $-D$, or $-(3c-2)$ and $3c+2$. But u must be positive, so γ is either a unit or has norm $3c+2$.

In the former case $u = 1$. However we know how to find all units in R; by starting with $s_0 = 2$, $t_0 = 0$ and $s_1 = 3c$, $t_1 = 1$ we obtain all units $\frac{s_i + t_i\sqrt{D}}{2}$ with $s_i, t_i \geqslant 0$ by the second order recurrence relation $s_{i+1} = 3cs_i - s_{i-1}$ and $t_{i+1} = 3ct_i - t_{i-1}$, so that $s_2 = 9c^2 - 2$ and $t_2 = 3c$. Taking the fundamental unit $\frac{3c + \sqrt{D}}{2}$ for $i = 1$, we would have $9c^2 D = X(X-4)$, so that $X = (ar+2)(br+2) = 9c^2$. Let us look at this case in more detail, as it occurs frequently. We consider the Δ-table and D-table as before, so that by multiplying every numerator and denominator in the D-table by $n = 3c$ we obtain the Δ-table. Thus n divides the denominators $r(br+2)$ and $r(ar+2)$. Converting this into statements purely about a, b and r gives us

$$(ar+2)(br+2) \mid r^2(br+2)^2$$

$$\text{and} \quad (ar+2)(br+2) \mid r^2(ar+2)^2.$$

Thus $ar+2$ divides $r^2(br+2)$, but if r is odd then $ar+2$ and r^2 will be coprime, so that $ar+2 \mid br+2$. However, if r is even then $X = (ar+2)(br+2)$ is even but equal to $9c^2$, which is not true. So $ar+2$ divides $br+2$, but the

argument is symmetric in a and b. Therefore we must have $a=b$ and we are dealing with the period two case. We end up with

$$X=(ar+2)^2=9c^2 \qquad \text{and so} \quad X-4=ar(ar+4),$$

giving $3c-2=ar$. Thus there are only two denominators featuring in the D-table, both of which are twice factors of D and so cannot be $2p^2$. We can divide out by $ar+2$ throughout in the Δ-table, and we see that these two denominators are $2r$ and $2a$, coming from factors of $3c-2$ and giving rise to the quadratic irrational $[\overline{a, r}]$. All such expansions will be of this form if $3c+2$ is prime.

We now have to eliminate the cases where $n=s_i t_i$ is obtained from the unit $\frac{s_i+t_i\sqrt{D}}{2}$ for $i\geqslant 2$. We do this by obtaining estimates for the size of the terms in the continued fraction expansion of σ, both from the D-table and from the Δ-table. As $s_2=9c^2-2$ and $t_2=3c$, we have $n\geqslant(9c^2-2)\,3c$, with

$$X(X-4)=n^2(9c^2-4)\geqslant 9c^2(9c^2-2)^2\,(9c^2-4),$$

where $X=(ar+2)(br+2)$. Now if $2p^2$ appears as the next denominator in the D-table, we would have $D=a^2\alpha^2+4p^2\alpha=b^2\beta^2+4p^2\beta$ where α, β divide D.

Also, as in the second row of the D-table we have

$$\frac{a\alpha+\sqrt{D}}{2p^2}-r=\frac{-b\beta+\sqrt{D}}{2p^2},$$

we obtain $2p^2r=a\alpha+b\beta$. Therefore we have

$$a\alpha,\ b\beta<\sqrt{D} \qquad \text{and} \qquad 2r<2p^2r<2\sqrt{D}.$$

Thus r, a and b are all less than $\sqrt{D}$, so that $ar+2$ and $br+2$ are each less than $D+2=9c^2-2$. Hence $X<(9c^2-2)^2$, giving $X(X-4)<(9c^2-2)^2\,9c^2(9c^2-4)$, which is a contradiction to the inequality above.

We finally have to eliminate $u=3c+2$. This is done in a similar way; the element $\gamma_0=\frac{3c+2+\sqrt{D}}{2}$ has norm $3c+2$ and any other possibilities for γ of the same norm must be a product of this and a unit, because we know from Lemma 5(1) that there is only one invertible ideal which is principal and generated by an element of norm $-(3c-2)$. Given any γ of norm $3c+2$, we have $\sqrt{D}/\gamma=\frac{-tv+s\sqrt{D}}{2}$ which is an element δ of norm $-(3c-2)$. But δR is an invertible ideal, as this is proved exactly as for γR, with u exchanged for v and s for t.

Writing

$$\gamma_i = \frac{(3c+2)\, s_i + t_i \sqrt{D}}{2},$$

we obtain the same recurrence relation as before by multiplying γ_0 by units, and we have $(s_0, t_0) = (1, 1)$ and $(s_1, t_1) = (3c-1, 3c+2)$. The case $i=0$ gives $n=1$ which has been dealt with, and the case $i=1$ gives $n = (3c-1)(3c+2)$ which is eliminated in the same way as $n=1$ by looking modulo 4. We have $\alpha n = r(br+2)$ and $p^2 n = rab + a + b$, where $n \equiv 2 \bmod 4$. Thus $\alpha n \equiv 2 \bmod 4$ means that r is not even, so we have $b \equiv 0 \bmod 4$ and similarly $a \equiv 0 \bmod 4$ too. Then $p^2 n$ ($\equiv 2 \bmod 4$) would be divisible by 4.

Finally, if $i \geqslant 2$ then certainly s_i and t_i are both greater than $(3c+1)(3c-2)$, thus $n > (3c+1)^2 (3c-2)^2 > (9c^2-2)\, 3c$ for $c > 1$, and so we will definitely have the lower bound for n which worked before, and which works now in exactly the same way. ∎

As mentioned before, we see that if $3c+2$ is prime then c is unique, as all continued fraction expansions of invertible quadratic ideals with square the identity will have purely periodic part $[\overline{\beta, \alpha}]$ where $\alpha\beta = 3c-2$. Also if $3c-2$ is prime then all such expansions will have the form $[\overline{a, r, b, r}] = [\overline{\beta-2, 1, \alpha-2, 1}]$ where $3c+2 = \alpha\beta$, so that again c is unique (see also [1]). However not all expansions of period four have come from a Markoff number c with $3c-2$ prime. For instance, the Markoff number $c = 5741$ (which appears in a triple with 2 and 985) has $3c-2$ factorising as $17 \cdot 1013$ and $3c+2 = 5^2 \cdot 13 \cdot 53$. Then taking $\alpha = 17 \cdot 53$, so that α is neither a factor of $3c-2$ nor of $3c+2$, the ideal represented by $\frac{k\alpha + \sqrt{D}}{2\alpha}$ (for k odd) has the periodic part of its continued fraction expansion equal to $[\overline{19, 17, 3, 17}]$. Thus we can eliminate the possibility of finding a denominator of the form $2p^2$ anywhere in this continued fraction expansion.

2. LOOKING FOR COUNTEREXAMPLES

We now see that there are limits to the approach of examining the continued fraction expansion of each invertible ideal whose square is the identity and trying to find a denominator $2p^2$, where p divides c. Consider ideals $\mathscr{I}$ whose periodic part of the continued fraction expansion is of order six and of the form

$$\sigma = [\overline{a, r, 1, b, 1, r}];$$

these will have $\mathscr{I}^2 \sim R$ by Lemma 5(3) and do occur in practice, for instance taking c to be the Markoff number 7561, so that $3c-2=37\cdot 613$ and $3c+2=5\cdot 13\cdot 349$, we put $\alpha=5\cdot 613$ and we get

$$\frac{7\alpha+\sqrt{D}}{2\alpha}=[\overline{7, 4, 1, 121, 1, 4}].$$

Therefore it seems natural that these should be the next ideals to be tackled in the same manner as Theorem 9.

We find on untangling the continued fraction expansion that

$$\sigma=\frac{a(r+1)(br+b+2r)+\sqrt{\Delta}}{2(r+1)(br+b+2r)},$$

where now

$$\Delta=(br+b+2r)(ar+a+2)(r+1)[a(br+b+2r)+2b+4].$$

We find by evaluating $\frac{1}{\sigma-a}$ that we have a denominator $2[a(br+b+2r)+a+b+2]$ which we would certainly want to show cannot be equal to $2p^2$, where p is a non-trivial factor of c. By examining possibilities for the different residue classes of r and a modulo small numbers, we find that we can eliminate cases unless $r\equiv 0\ (24)$ and $a\equiv 3\ (12)$. Actually putting $r=24$ and $a=3$ we suppose that we have some odd number c with $D=9c^2-4$ and α dividing D, giving rise to an ideal $\mathscr{I}$ whose square is the identity and which is represented by $\sigma=(k\alpha+\sqrt{D})/2\alpha$ with

$$\sigma=\frac{75(25b+48)+\sqrt{\Delta}}{2\cdot 25(25b+48)}=[\overline{3, 24, 1, b, 1, 24}]$$

with $D=\Delta$, thus we obtain

$$3c=(br+b+2r)(ar+a+2)+2=25\cdot 77b+48\cdot 77+2$$

and denominator on the next step

$$a(br+b+2r)+a+b+2=76b+149.$$

We ask: is it possible for $76b+149$ to be a square, say p^2, where p divides $3c$?

Let us set $3c=kp$, then eliminating b from our two equations involving p, we obtain

$$77^2+2\cdot 76=25\cdot 77p^2-76kp. \tag{1}$$

Thus we would need 76 dividing $25p^2-1$. Looking mod 19 we need $(5p)^2 \equiv 1$. All solutions of $x^2 \equiv 1$ (19) are of the form $x = 19s \pm 1$ for $s \in \mathbb{Z}$. Trying $s = 5t+1$, $x = 19s+1$ or $s = 5t-1$, $x = 19s-1$ in order to obtain a solution x divisible by 5, we try some small values of t. But $t = 3$ would give $p = 53$, and this fits the equation (1) with $k = 1341$. Moreover b would be 35 and this satisfies both conditions. Therefore, taking $c = 447 \cdot 53 = 23691$ we obtain $D = 9c^2 - 4$ and $\alpha = 923$ dividing D, where the ideal $\mathscr{I}$ represented by σ has $\mathscr{I}^2 \overset{+}{\sim} R$, and σ is equal to the continued fraction expansion above with $b = 35$. Moreover we see that $p = 53$ splits in R, giving rise to an ideal $\mathscr{P}$ with $\mathscr{P}\bar{\mathscr{P}} = pR$, by showing $\left(\frac{D}{p}\right) = 1$. As p divides c, we have $\left(\frac{D}{p}\right) = \left(\frac{-4}{p}\right)$, but 4 is clearly a quadratic residue mod p, and so is -1, as $p \equiv 1 \bmod 4$. Thus we obtain:

THEOREM 10. *There exists $D = 9c^2 - 4$ with c odd, a non-trivial factorisation $c = pq$, and an ideal $\mathscr{P}$ with $\mathscr{P}\bar{\mathscr{P}} = pR$ where $\mathscr{P}^4 \overset{+}{\sim} R$.*

Of course, in this case c is not a Markoff number. But once we transfer the original problem on the uniqueness of a Markoff number c into the question of whether there exists a factor of c which splits into ideals in the order $\mathbb{Z} + \left(\frac{3c-2+\sqrt{D}}{2}\right)\mathbb{Z}$ whose fourth power is equivalent to the identity in the ideal class group, then this problem has now been rephrased into a form which can be applied to any odd number c, and which is not true in this setting. Therefore this implies that an approach which would work for the general unicity conjecture must use the appearance of the Markoff numbers in the Markoff equation in a fundamental way, rather than reformulating the problem into a question which can also be asked about other numbers.

REFERENCES

1. A. Baragar, On the unicity conjecture for Markoff numbers, *Canad. Math. Bull.* **39** (1996), 3–9.
2. J. O. Button, The uniqueness of the prime Markoff numbers, *Bull. London Math. Soc.* **58** (1998), 9–17.
3. J. W. S. Cassels, "An Introduction to Diophantine Approximation," Chap. 2, Cambridge Univ. Press, Cambridge, UK, 1957.
4. H. Cohen, "A Course in Computational Algebraic Number Theory," Graduate Texts in Mathematics, Vol. 138, Springer-Verlag, New York, 1993.
5. T. W. Cusick and M. E. Flahive, "The Markoff and Lagrange Spectra," Math. Surveys and Monographs, Vol. 30, Amer. Math. Soc., Providence, RI, 1989.
6. L. E. Dickson, "Studies in the Theory of Numbers," Chicago, 1930.
7. C. Series, The geometry of Markoff numbers, *Math. Intell.* **7**, 3 (1985), 20–29.