

THE MARKOFF SPECTRUM AND MINIMA OF INDEFINITE BINARY QUADRATIC FORMS

MARY E. GBUR

ABSTRACT. Recently T.W. Cusick has shown that the Lagrange spectrum is the closure of the Markoff values of completely periodic doubly infinite sequences. In this note it is proved that if the Markoff value of a sequence M is attained at least three times, then M is a completely periodic sequence. We also show that if the minimum of an indefinite binary quadratic form is attained at least three times, then the form is equivalent to a rational form.

1. **Notation and definitions.** We let

$$x = [x_0; x_1, x_2, \dots] = x_0 + \frac{1}{x_1 + \frac{1}{x_2 + \dots}}$$

denote the regular simple continued fraction expansion of a positive number x . In any continued fraction of the form $[0; \overline{b_1, \dots, b_k}]$, the bar indicates that the tuple b_1, \dots, b_k is repeated infinitely often.

For any doubly infinite sequence $M = \{a_k : k = 0, \pm 1, \dots\}$ of positive integers and any integer k , we set

$$\mu_k(M) = [a_k; a_{k+1}, a_{k+2}, \dots] + [0; a_{k-1}, a_{k-2}, \dots].$$

Then $\mu(M) = \sup_k \mu_k(M)$ is called the *Markoff value* of the sequence M and $\lambda(M) = \overline{\lim}_k \mu_k(M)$ is called the *Lagrange value* of the sequence M , where both suprema are taken over all integers. The set of all Markoff values is called the *Markoff spectrum*; the set of all Lagrange values is the *Lagrange spectrum*.

It has been shown [1], [4] that the Lagrange spectrum is a proper subset of the Markoff spectrum. Hall [5] proved that in order to completely describe the Markoff spectrum it is sufficient to consider only those sequences M for which $\mu(M)$ is attained.

2. **Background and statement of results.** In [8], the first significant work on the Markoff spectrum, Markoff described the connection between an element of the Markoff spectrum and the minimum of an associated indefinite binary quadratic form. Most proofs since the work of Markoff have either relied solely on a continued fraction approach or have worked entirely with forms. In this note we exploit the intimate relationship between these two ap-

Presented to the Society, November 6, 1976; received by the editors September 23, 1976.
AMS (MOS) subject classifications (1970). Primary 10E20, 10E25, 10C25.

© American Mathematical Society 1977

proaches to obtain the following new information on the Markoff spectrum.

THEOREM 1. *If there exist three distinct integers k, m, n for which*

$$\mu_k(M) = \mu_m(M) = \mu_n(M),$$

then M is a completely periodic sequence and $\mu(M) = \lambda(M)$.

In [1], Cusick proved that the Lagrange spectrum is precisely the closure of the set of Markoff values of completely periodic doubly infinite sequences. Thus, from Theorem 1 we obtain a sufficient condition for an element of the Markoff spectrum to be a basic element of the Lagrange spectrum.

In [4], Freiman considers the sequence M defined by

$$\begin{aligned} \mu_0(M) \\ = [2; 1, 2, 2, 1, 1, 2, 2, 2, 2, 1, 2, 2, 1, 1, 2, 2, 1, 1, \overline{2, 2, 1, 1, 2, 2, 1, 2, 2}] \\ + [0; \overline{2, 2, 2, 1, 1, 2, 2, 1, 2}]. \end{aligned}$$

Freiman showed that $\mu(M)$ is not an element of the Lagrange spectrum and that $\mu(M)$ is attained exactly twice. Hence, without the three in Theorem 1 we cannot even guarantee that $\mu(M)$ is in the Lagrange spectrum.

Our method of proof also gives some information on the minima of indefinite binary quadratic forms. For any indefinite binary quadratic form $f(x, y)$ with real coefficients, we set $M(f) = \inf\{|f(m, n)|: (m, n) \text{ nonzero integral point}\}$. We prove the following.

THEOREM 2. *If $M(f) \neq 0$ is attained at least three times, then $f(x, y)$ is equivalent to a rational form and $M(f)$ is attained infinitely often.*

Using Freiman's sequence M defined above, our proof also shows that, for $\alpha = [a_0; a_1, a_2, \dots]$ and $\beta = -[0; a_{-1}, a_{-2}, \dots]$, the form

$$f(x, y) = (x - \alpha y)(x - \beta y)$$

has minimum $M(f) = 1$ which is attained exactly twice. Thus, the three in Theorem 2 also cannot be decreased.

In the following section we develop the observations from Markoff's work [7] which we use to prove these results.

3. A chain of reduced forms. Let $f(x, y)$ be an indefinite binary quadratic form which does not represent zero nontrivially. Dickson [3] has shown that $f(x, y)$ is equivalent to an *Hermite reduced form*; that is, a form of the type $g(x, y) = a(x - \alpha y)(x - \beta y)$, with $\alpha > 1$ and $-1 < \beta < 0$.

We let $\alpha = [a_0; a_1, a_2, \dots]$ and $\beta = -[0; a_{-1}, a_{-2}, \dots]$ be the regular simple continued fraction expansions of α and β , respectively. Since $f(x, y)$ does not represent zero, then α and β are irrational numbers and $M = \{a_k: k = 0, \pm 1, \pm 2, \dots\}$ is a doubly infinite sequence of positive integers.

We define the two doubly infinite sequences $\{S_k\}, \{T_k\}$ of positive integers as follows:

$$S_0 = T_{-1} = 1, S_{-1} = T_0 = 0;$$

and, for positive k ,

$$S_k/T_k = [a_0; a_1, \dots, a_{k-1}],$$

$$S_{-k-1}/T_{-k-1} = -[0; a_{-1}, \dots, a_{-k}],$$

where these fractions are in lowest terms and we take S_{-k-1} to be positive exactly when k is odd. From the definition we obtain that, for each integer k , $S_k T_{k-1} - S_{k-1} T_k = \pm 1$.

For each integer k , we define the form

$$g_k(x, y) = g(S_k x + S_{k-1} y, T_k x + T_{k-1} y).$$

Markoff showed [8, p. 385] that, for each integer k ,

$$\mu_k(M) = \sqrt{D(g)} / |g_k(1, 0)|,$$

where $D(g)$ denotes the discriminant of the form $g(x, y)$. This fact and the following lemma are crucial to our proofs.

LEMMA. *Let (m, n) be any nonzero integral point. Then there exists an integer j such that $g(S_j, T_j)$ and $g(m, n)$ have the same sign and $|g(S_j, T_j)| \leq |g(m, n)|$, with equality only if $(m, n) = (S_k, T_k)$, for some integer k .*

Our proof of this lemma relies only on known geometric results on the Klein polygon. The referee has kindly noted that the facts used have been developed in [2]. Although an outline of a similar method recently appeared in [7, pp. 295–296], for completeness we include a proof here.

PROOF. We define a doubly infinite sequence of points $\{P_k = (p_k, q_k)\}$ by $P_{k+1} = S_k(a, 1) + T_k(-a\alpha, -\beta)$, for each integer k . From our definition of the sequences $\{S_k\}$, $\{T_k\}$, we are able to obtain the following properties:

- (i) For each integer k , p_k and p_{k+1} have different sign;
- (ii) The sequence $\{q_k\}$ of positive numbers is strictly increasing with $\lim_{k \rightarrow \infty} q_k = \infty$ and $\lim_{k \rightarrow -\infty} q_{-k} = 0$;
- (iii) For each integer k , $P_{k+1} = a_k P_k + P_{k-1}$.

Before continuing with the lemma, we shall verify these three properties. Because $p_k = a(S_{k-1} - \alpha T_{k-1})$, from the choice of the signs of S_{-k} and T_{-k} , for $k \geq 2$, we have that the sequence $\{p_{-1}, p_{-2}, \dots\}$ alternates in sign. Checking that

$$p_{-1} = a(1 + a_{-1}\alpha), \quad p_0 = -a\alpha, \quad p_1 = a,$$

and recalling that [9, pp. 42–43], for each positive j ,

$$S_{2j-1}/T_{2j-1} < \alpha < S_{2j}/T_{2j},$$

we obtain that $\{p_{-1}, p_0, \dots\}$ is also an alternating sequence and property (i) does hold.

Since $\{S_1, S_2, \dots\}$ and $\{T_1, T_2, \dots\}$ are strictly increasing sequences of positive integers and β is negative, then $\{q_1, q_2, \dots\}$ is a strictly increasing unbounded sequence of positive numbers. As above, for each positive integer j ,

$$S_{-2j}/T_{-2j} < \beta < S_{-2j-1}/T_{-2j-1}.$$

We chose T_{-k-1} to be negative exactly when k is odd and so we obtain that, for each positive integer k , $q_{-k} = S_{-k-1} - \beta T_{-k-1}$ is a positive number. Also, by [9, p. 52] it follows that $\{q_0, q_{-1}, \dots\}$ is a strictly decreasing sequence converging to 0. Observing that

$$q_0 = [0; a_{-1}, a_{-2}, \dots] < 1 = q_1 < q_2 = a_0 + q_0,$$

we have thus shown that property (ii) holds.

Finally, for each integer $k \geq 3$, we have

$$|S_{-k-1}| = a_{-k}|S_{-k}| + |S_{-k+1}|, \quad |T_{-k-1}| = a_{-k}|T_{-k}| + |T_{-k+1}|.$$

The signs of $\{S_{-2}, S_{-3}, \dots\}$ and $\{T_{-2}, T_{-3}, \dots\}$ have been chosen to alternate so that this implies, for each integer $k \geq 3$,

$$S_{-k-1} = -a_{-k}S_{-k} + S_{-k+1}, \quad T_{-k-1} = -a_{-k}T_{-k} + T_{-k+1}.$$

From the definition of the continued fractions, this also holds for each integer $k \leq 2$. Hence, for each integer k ,

$$P_{k+1} = a_k P_k + P_{k-1},$$

which is property (iii).

Returning to the proof of the lemma, we consider any nonzero integral point (m, n) . If L denotes the lattice in R^2 generated by the points $P_1 = (a, 1)$ and $P_0 = (-a\alpha, -\beta)$, then the lattice point $P = (p, q)$ defined by

$$P = mP_1 + nP_0 = (ma - na\alpha, m - n\beta)$$

has the property that $g(m, n) = pq$. Since we are concerned only with the product $pq = g(m, n)$ and both P and $-P$ are lattice points, we may assume that q is positive. From properties (i) and (ii) given above, it follows that there exists an integer j such that p and p_j have the same sign and $q_j \leq q < q_{j+2}$. We shall show that, for this choice of j , either j or $j+2$ satisfies the conditions of the lemma. From our choice of j we have that $g(m, n) = pq$ has the same sign as $g(S_{k-1}, T_{k-1}) = p_k q_k$, for $k = j, j+2$.

Since, for each integer k , $S_k T_{k-1} - S_{k-1} T_k = \pm 1$, then the points P_j and P_{j+1} also generate the lattice L . Hence, there is no lattice point (other than vertices) in the fundamental parallelogram formed by the vectors P_j and P_{j+1} . From

$$P_{j+2} = a_{j+1} P_{j+1} + P_j$$

it thus follows that there is no lattice point inside the triangle T formed by $(0, 0)$, P_j and P_{j+2} .

We set $A = \min\{|p_j q_j|, |p_{j+2} q_{j+2}|\}$. Since the lattice point P cannot lie interior to the triangle T , by convexity P cannot lie in the hyperbolic region $|xy| \leq A$ unless $P = P_j$. Therefore, if $P \neq P_j$, we have

$$|g(m, n)| = pq > A = \min\{|g(S_{j-1}, T_{j-1})|, |g(S_{j+1}, T_{j+1})|\}.$$

This completes the proof of the lemma.

4. Proofs of Theorem 1 and Theorem 2.

PROOF OF THEOREM 1. We index the sequence M so that $\mu_0(M) = \mu_m(M) = \mu_n(M)$, for at least two distinct positive integers m, n and we define the real numbers

$$\alpha = [a_0; a_1, a_2, \dots] \quad \text{and} \quad \beta = -[0; a_{-1}, a_{-2}, \dots].$$

Since M is a doubly infinite sequence of positive integers, then α and β are well-defined irrational numbers and the binary quadratic form

$$g(x, y) = (x - \alpha y)(x - \beta y)$$

is an indefinite Hermite reduced form with $g(1, 0) = 1$. We recall that, for any positive integer k ,

$$\mu_k(M) = \sqrt{D(g)} / |g(S_k, T_k)|,$$

where $S_k/T_k = [a_0; a_1, \dots, a_{k-1}]$. Therefore, from

$$\mu_0(M) = \mu_m(M) = \mu_n(M),$$

for the positive integers m, n , it follows that

$$(*) \quad g(S_m, T_m) = \pm 1 \quad \text{and} \quad g(S_n, T_n) = \pm 1.$$

If we let b and c denote the real numbers for which

$$g(x, y) = x^2 + bxy + cy^2,$$

then the equations of $(*)$ form a system of two equations in the indeterminants b and c . The determinant of this system is

$$T_m^2 T_n^2 (S_m/T_m - S_n/T_n).$$

Since S_m/T_m and S_n/T_n are different convergents of α , this determinant is nonzero and b, c are rational numbers.

Perron [9, p. 80] has shown that the roots of a rational Hermite reduced form have completely periodic continued fraction expansions. Hence, there exists an integer R such that $\alpha = [a_0; a_1, \dots, a_R]$. Since $g_{-R-1}(x, y)$ is also a rational Hermite reduced form with positive root given by $[a_{-R-1}; a_{-R}, \dots, a_{-1}, \alpha]$, this continued fraction must also be completely periodic. This implies that this root is α . Therefore, $\beta = [0; a_R, \dots, a_1, a_0]$ and M is a completely periodic sequence.

PROOF OF THEOREM 2. Since $f(x, y)$ is equivalent to an Hermite reduced form, we may assume that

$$f(x, y) = a(x - \alpha y)(x - \beta y),$$

with $\alpha > 1$ and $-1 < \beta < 0$. We let the continued fraction expansions of α and β be given by

$$\alpha = [a_0; a_1, a_2, \dots] \quad \text{and} \quad \beta = -[0; a_{-1}, a_{-2}, \dots].$$

Since $f(x, y)$ does not represent zero, then α and β are irrational numbers and the sequence $M = \{a_k\}$ is a doubly infinite sequence of positive numbers.

By hypothesis, the minimum $M(f)$ is attained at least three times. Therefore, by the lemma of §3, there exist three distinct integers $j = k, m, n$ with

$M(f) = |f(S_j, T_j)|$ and so $\mu_k(M) = \mu_m(M) = \mu_n(M)$. Hence, by Theorem 1, M is a completely periodic sequence and $f(x, y)$ is a rational form.

ADDED IN PROOF. G. A. Freiman has pointed out that on p. 148 in [6] there is a statement of a corollary to our Theorem 1; viz., if $\mu(M) = \mu_i(M)$, for three or more integers i , then M is completely periodic. No indication of a proof is given.

REFERENCES

1. T.W. Cusick, *The connection between the Lagrange and Markoff spectra*, Duke Math. J. **42** (1975), 507–517. MR **51** #10240.
2. B.N. Delone and D.K. Faddeev, *The theory of irrationalities of the third degree*, Trudy Mat. Inst. Steklov. **11** (1940); English transl., Transl. Math. Monographs, vol. 10, Amer. Math. Soc., Providence, R.I., 1964, Supplement I, Sections 3,5. MR **2**, 349d; **28** #3955.
3. L. E. Dickson, *Introduction to the theory of numbers*, Univ. of Chicago Press, Chicago, 1929, Chap. VII.
4. G. A. Freĭman, *Noncoincidence of the Markov and Lagrange Spectra*, Mat. Zametki **3** (1968), 195–200 = Math. Notes **3** (1968), 125–128. MR **37** #2695.
5. M. Hall, Jr., *The Markoff spectrum*, Acta Arith. **18** (1971), 387–399. MR **45** #5084.
6. A. G. Kuroš, *Matematika v SSSR za sorok let 1917–1957*, vol. 1, State Publications, Physical-Mathematical Literature, Moscow, 1959.
7. Å. Lindgren, *One-sided minima of indefinite binary quadratic forms and one-sided diophantine approximations*, Ark. Mat. **13** (1975), 287–302. MR **52** #8042.
8. A. Markoff, *Sur les formes binaires indéfinies*, Math. Ann. **15** (1879), 381–406.
9. O. Perron, *Die Lehre von der Kettenbrüche*, Chelsea, New York, 1929.

DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS, OHIO 43210