Journal of Number Theory 227 (2021) 265–305

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

General Section

On quadratic numbers and forms, and Markoff theory

Christophe Reutenauer

Département de mathématiques, Université du Québec à Montréal, Canada

ARTICLE INFO

Article history:
Received 19 February 2020
Received in revised form 19 March 2021
Accepted 22 March 2021
Available online 20 April 2021
Communicated by F. Pellarin

Keywords:
Markoff theory
Continued fractions
Error term for approximations of quadratic numbers
Indefinite binary quadratic forms
Minima and small values of quadratic forms
Refinement of Lagrange number of reals
Refinement of Markoff classification
Christoffel words

ABSTRACT

A formula giving the exact error term for convergents of quadratic numbers is given. This formula is applied to Markoff forms and Markoff irrationalities: characterization of the smallest values of the form (the minimum being the corresponding Markoff number), mapped bijectively to the conjugates of the corresponding Christoffel word; a new proof of the result giving the ranks of the good convergents of the Markoff irrationalities; a refinement of Markoff's classification of the latter numbers. Other applications appear in the article.

Contents

1. Introduction 266
Part 1. Preliminaries 270
2. Notations and definitions 270
3. Continuant polynomials 270
4. Convergents of the conjugate of a reduced quadratic number 272

E-mail address: Reutenauer.Christophe@uqam.ca.

https://doi.org/10.1016/j.jnt.2021.03.005
0022-314X/

5. Quadratic numbers equivalent to their conjugate . . . 273
6. A theorem of Serret . . . 274
Part 2. On quadratic numbers and quadratic forms . . . 274
7. An identity with continuant polynomials . . . 274
8. Applications to small values of quadratic forms . . . 276
9. Error term for quadratic numbers . . . 279
10. Intermezzo . . . 283
11. Lagrange number . . . 283
12. Good approximations of quadratic numbers . . . 284
13. A measure of approximation . . . 285
14. An invariant of quadratic numbers . . . 287
Part 3. Markoff theory . . . 290
15. Markoff forms . . . 290
16. A lexicographical result . . . 292
17. Small values of Markoff forms . . . 295
18. Good approximations of Markoff irrationalities . . . 297
19. Refinement of the Markoff classification . . . 298
Part 4. Study of the invariant D_0 . . . 300
20. Stabilizer and invariant . . . 300
References . . . 304

1. Introduction

By Hurwitz's $\sqrt{5}$ theorem (Theorem 1.21 in [1]), each irrational real number ξ has infinitely many rational approximations $\frac{p}{q}$ such that

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2\sqrt{5}}.$$

The fractions $\frac{p}{q}$ are necessarily convergents $\frac{p_t}{q_t}$ of ξ: $\frac{p_t}{q_t} = [a_0, a_1, \ldots, a_t]$ where the expansion of ξ into continued fractions is $[a_0, a_1, \ldots, a_n, \ldots]$. For this notation, and other standard properties of continued fractions, see [18,1].

The *Lagrange number* $L(\xi)$ of ξ is the supremum of the positive real constants C such that ξ has infinitely many rational approximations $\frac{p}{q}$ satisfying

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{Cq^2}.$$

Therefore $C \geq \sqrt{5}$.

It is not clear from the definition if for $C = L(\xi)$ infinitely many approximations satisfying the previous inequality (which we call *good approximations*) exist or not. Actually, it may happen that ξ has only finitely many good approximations, even if ξ is quadratic number, as shown by Perron (see Section 12).

For studying the Lagrange number, one has to have a determination for the *error term* $|\xi - \frac{p_t}{q_t}|$. Such an expression exists in terms of continued fractions of ξ; it is

$$\left|\xi - \frac{p_t}{q_t}\right| = \frac{1}{\lambda_t q_t^2},\ \lambda_t = [a_{t+1}, a_{t+2}\ldots] + [a_t, a_{t-1}, \ldots, a_1]^{-1}. \tag{1}$$

It follows that the Lagrange constant of ξ is the supremum of the numbers λ_t ([1] Proposition 1.22).

In the present article, we give, for quadratic numbers, an alternative expression for the error term. For *reduced* quadratic numbers $\xi = [\overline{a_0, \ldots, a_{n-1}}]$, the formula is

$$\xi - \frac{p_t}{q_t} = \frac{(-1)^t \epsilon_{t+1 \bmod n}}{q_t^2 \frac{\sqrt{d(f)}}{2}(1 + \sqrt{1 + 4(-1)^{t+1} f(1,0)\epsilon_{t+1 \bmod n}/d(f)q_t^2})}, \tag{$*$}$$

with the following notations: $f(x,y)$ is the quadratic form $cx^2 + (d-a)xy - by^2$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix}$, $d(f)$ is the discriminant of f, and the numbers ϵ_r, $r = 0, \ldots, n-1$, are defined by the *continuant polynomials* $\epsilon_r = p(a_{r+1}, \ldots, a_{n-1}, a_0, \ldots, a_{r-1})$ (in other words the denominator of the finite continued fraction $[a_r, \ldots, a_{n-1}, a_0, \ldots, a_{n-1}]$ obtained by cyclically permuting $[a_0, \ldots, a_{n-1}]$); see Corollary 9.1, and Theorem 9.1 for the more general formula involving any quadratic number (not necessarily reduced).

This formula has many applications. In particular, *for quadratic numbers*, we may characterize the Lagrange number (Corollary 11.1), characterize the ranks (defined in Section 2) of convergents which are good approximations (Theorem 12.1), characterize the quadratic numbers having infinitely many good approximations (Corollary 12.1), revisit Perron's examples and prove that the sequence of ranks of good convergents is periodic (Corollary 12.2).

Next, motivated by this formula, and by [4,16,17], we define the *second Lagrange number* $D(\xi)$ to be the supremum of all nonnegative reals d such that ξ has infinitely many rational approximations $\frac{p}{q}$ satisfying

$$\left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^2 \frac{L(\xi)}{2}(1 + \sqrt{1 + d/q^2})}.$$

As we will see in Section 13, the second Lagrange number exists if and only if ξ has infinitely many good approximations. For quadratic numbers we may characterize it exactly, as follows from the error term formula above: $D(\xi) = 4m|f(e,i)|/d(f)$, where m is the minimum of the numbers ϵ_r (Theorem 13.1; the numbers e, i are determined by the pre-period of ξ), and one has equality in the previous formula if $\frac{p}{q}$ is a good approximation.

This leads us to study independently the integer $D_0(\xi) = |f(e,i)|$. An intrinsic characterization is that $D_0(\xi)$ is, up to the sign, equal to the $2,1$-entry of the generator of the stabilizer (which is cyclic) of ξ in $PSL_2(\mathbb{Z})$ (Theorem 20.1); it implies that D and D_0 are invariant under conjugation (Corollary 20.3). They are also invariant under integer translation (but not $GL_2(\mathbb{Z})$-equivalence). This leads to a partial characterization

of the smallest values of $D_0(\xi)$ for the quadratic number ξ in a given large class (that is, the union of a class and of its conjugates); the result is somewhat too technical to be stated in the introduction, and is also partial (but will be complete in the case of Markoff irrationalities, see below); roughly speaking, the ξ satisfying $D_0(\xi) < (1/2)\sqrt{d}$ (where d depends only on the large class) lie in finitely many integral translation classes determined by the periodic pattern of the continued fraction of ξ (Theorem 14.1).

In order to prove the error term formula $(*)$, we are led to prove first a result on quadratic forms (binary, indefinite, real coefficients). Motivated by a theorem of Serret ([25] Théorème p. 79-80, see Section 6), we call *small values* of such a form f a value $f(x,y)$ such that $|f(x,y)|$ is smaller than half the discriminant of f. The theorem of Serret asserts that, under some mild assumptions, if $p,q \in \mathbb{Z}$, not both 0, satisfy $|f(p,q)| < \frac{1}{2}d(f)$, then p/q is a convergent of one of the roots of the polynomial $f(x,1)$. We give an exact formula for the small values of $f(p,q)$, in terms of the continued fractions expansion of these roots; the small values turn out to be the numbers ϵ_r, defined above (Theorem 8.1). This formula is based on a formula involving continuant polynomials (Theorem 7.1).

In the third part of the article, we apply the previous results to Markoff theory. This theory has two aspects: quadratic forms, and approximation of reals. Concerning Markoff's theory *for quadratic forms* (which is the actual content of his two papers), it starts by a theorem of Korkine and Zolotareff ([21]; this result is similar to Hurwitz's result above): for any quadratic form $f(x,y)$, with $m(f) = \min\{|f(p,q)|, p,q \in \mathbb{Z}, (p,q) \neq (0,0)\}$, one has $\frac{\sqrt{d(f)}}{m(f)} \geq \sqrt{5}$ (one has equality for x^2-xy-y^2). Markoff's theory concerns forms satisfying $\frac{\sqrt{d(f)}}{m(f)} < 3$. Such forms are proportional to forms which are $GL_2(\mathbb{Z})$-equivalent to the so-called *Markoff forms*, which will be defined below, and whose set is countable.

The Markoff theory *for approximations* concerns real numbers ξ whose Lagrange number satisfies $L(\xi) < 3$. In this case, ξ is $GL_2(\mathbb{Z})$-equivalent to the so-called *Markoff irrationalities*, whose set is also countable.

Markoff forms and Markoff irrationalities are closely related to the Markoff numbers, which form a sequence of integers. We follow the approach of [24] (see Section 15 for more details): Markoff forms, irrationalities, and numbers, are all parametrized by the so-called *Christoffel words*; these words, on the alphabet $\{a,b\}$, appear as the codings of discrete paths, associated to a given slope (see Fig. 1). To each Christoffel word w is associated bijectively a Markoff form and a Markoff irrationality. Moreover, to each Christoffel word is associated a Markoff number: this mapping is surjective; the injectivity of this mapping is open, and is called the *Markoff numbers uniqueness conjecture*, or *Frobenius conjecture*, see [1].

It is known that the minima of Markoff forms are the Markoff numbers (see e.g. [24] Theorem 9.3.1). We generalize this result: if f_w is the Markoff form associated to the Christoffel word w, of length n, then the n smallest values of f_w are exactly its small values (that is, $< \frac{1}{2}\sqrt{d(f_w)}$ in absolute value), and are bijectively and naturally associ-

ated to the n cyclic conjugates of w: precisely, for v a cyclic conjugate of w, the bijection maps v onto the number $\mu(v)_{12}$ (the 1,2-entry of the matrix $\mu(v)$), for some linear representation μ (see (14)) of the free monoid into $SL_2(\mathbb{Z})$ (in particular $\mu(w)_{12}$ is the corresponding Markoff number); this bijection is increasing for the lexicographical order (see Fig. 2 for a quick illustration to this result); this result looks like an arithmetical counterpart of several results characterizing Christoffel words by their cyclic conjugates, for example [22] (see [24] Chapter 15 for other such results). Moreover, we may characterize exactly the pairs (p, q) of integers such that $f_w(p, q)$ is equal to one of these n smallest values, in particular when $f(p, q) = m$, the corresponding Markoff number and minimum of f_w: p/q must be a convergent of the associated Markoff irrationality x_w, or of its conjugate, and the ranks are precisely determined (Theorem 17.1, Corollary 17.1 and 17.2).

Concerning Markoff irrationalities ξ, it is known that they have infinitely many good rational approximations $\frac{p}{q}$, that is, satisfying $|\xi - \frac{p}{q}| < \frac{1}{L(\xi)q^2}$. This result is not so easy to establish, and is proved as far as I know only in the book by Cassels ([7] Theorem III.B), in Bombieri's article ([2] Theorem 29) and in the author's book ([24] Theorem 8.3.3); in the two latter, the precise ranks of the convergents which are good approximations are also given. The error term formula mentioned above allows us to give a quick proof of these results, using the striking fact that this formula directly relates the Markoff irrationality and the corresponding Markoff form (see Theorem 18.1).

Finally, in the section entitled "Refinement of Markoff classification", we state and prove the result which is perhaps the main result (at least the most difficult, with a proof based on the earlier results in Section 14) of this article. For a given Christoffel word w and corresponding $GL_2(\mathbb{Z})$-equivalence class C_w of the Markoff irrationality x_w, we may classify the numbers in this class according to the value of the second Lagrange number, equivalently, to the value of the invariant D_0. Recall that $D_0(\xi)$ depends only on the integer translation class of ξ. We show that the $n = |w|$ smallest values of D_0 in C_w are the same n numbers as above: that is, the n smallest values of the Markoff quadratic form f_w, which are also equal to $\mu(v)_{12}$ for some cyclic conjugate v of w; moreover, for given v, the numbers $\xi \in C_w$ such that $D_0(\xi) = \mu(v)_{12}$ are integral translates of two numbers given by their expansion into continued fractions, or their conjugates (Theorem 19.1); in particular for $v = w$ and hence $\mu(w)_{12} = m$, the corresponding Markoff number, the numbers ξ in the class C_w, such that $D_0(\xi) = m$, are the integral translates of x_w, of another quadratic number, or the conjugates of these two numbers (Corollary 19.1). The two first cases of this corollary (with $w = a$ or b, and corresponding to $x_w =$ the golden ratio or the "silver ratio") are due to Hančl [16,17].

Acknowledgments: I thank Yann Bugeaud, François Bergeron (who draw my attention to the article [16] of Hančl), Valérie Berthé, and Hugh Thomas, for useful mail exchanges and discussions. Thank you also to Jean-Eric Pin and Jan Okninski. Thanks are due to the referee, for thorough reading and many stylistic suggestions and corrections of mistakes; moreover, he trivialized the proof of Lemma 3.2, and made an interesting

comment about the definition of what I call "second Lagrange number" (see the remark in Section 13).

This work has been partially supported by NSERC, Canada.

Part 1. Preliminaries

2. Notations and definitions

In the sequel, we consider real quadratic numbers; their expansion into continued fractions is ultimately periodic (Lagrange). We constantly use, if not specified otherwise, the notation

$$\alpha = [\overline{a_0, \dots, a_{n-1}}], \beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}],$$

where $n \geq 1, k \geq 0$ and $a_0, \dots, a_{n-1}, b_1, \dots, b_{k-1} \in \mathbb{Z}_{>0}$ and $b_0 \in \mathbb{Z}$.

If $\alpha = [a_0, a_1, \dots]$ is a continued fraction, we define as usual the convergents p_i/q_i by the equality $p_i/q_i = [a_0, \dots, a_i]$, a fraction in lowest terms with $q_i > 0$. The *rank* of this convergent is i; in particular, the ranks start at 0.

If s is an infinite word over $\mathbb{Z}_{>0}$, $[s]$ denotes the corresponding continued fraction. That is, if $s = a_0 a_1 \cdots a_n \cdots$, then $[s] = [a_0, a_1, \dots, a_n, \dots]$.

If w is a finite nonempty word, w^∞ denotes the infinite word obtained by infinitely concatenating w with itself.

If β is a quadratic number, as above, then we call *minimal periodic pattern* of β the word $a_0 \cdots a_{n-1}$ provided n is minimal, called then the *minimal period* of β. There are several minimal periodic patterns of β, which are all conjugate words; recall that two words in a free monoid are called *conjugate* if they may be written xy and yx for some words x, y. The *conjugation class* of w is the set of its conjugates.

We denote by $\widetilde{w}$ the *reversal* of the word w, obtained by reading its letters from right to left. A word is a *palindrome* if it is equal to its reversal. A *central factor* of a palindrome w is a word f such that $w = xf\tilde{x}$.

3. Continuant polynomials

Let $a_0, \dots, a_{n-1}$ be elements of a ring with unit.[1] Let $P(a) = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}$. Recall that the *continuant polynomials* $p(a_0, \dots, a_{n-1})$ are defined by the matrix equation (see [9] 2.7)

[1] Although the applications we give in this article use continuant polynomials evaluated only on the integers, one may define them for any ring with 1. Continuant polynomials are indeed useful in noncommutative ring theory, as reveals the article of Wedderburn [26], or the work of Cohn on free associative algebras and generalizations [9] 2.7. The reader may verify that Theorem 7.1 has a noncommutative version, which follows from its proof.

$$P(a_0)\cdots P(a_{n-1}) = \begin{pmatrix} p(a_0,\dots,a_{n-1}) & p(a_0,\dots,a_{n-2}) \\ p(a_1,\dots,a_{n-1}) & p(a_1,\dots,a_{n-2}) \end{pmatrix}. \tag{2}$$

One often puts an index under p, indicating the number of arguments, for example $p_n(a_0,\dots,a_{n-1})$. This is of course redundant, but useful to indicate the following conventions: $p_0() = p() = 1$ and even $p_{-1} = 0$. These conventions are consistent with the previous matrix product for $n = 1$.

We also denote $P(a_0)\cdots P(a_{n-1})$ by $P(a_0,\dots,a_{n-1})$, or even by $P(a_0\cdots a_{n-1})$; with this latter notation, we view P as a homomorphism from the free monoid generated by the ring, into the multiplicative monoid of 2 by 2 matrices over this ring. Likewise, we sometimes write $p(a_0\cdots a_{n-1})$ for $p(a_0,\dots,a_{n-1})$.

For later use, we note the following properties of continuants.

Lemma 3.1. *(i) One has the recursive formulas: for any $n \geq 1$,*

$$p(a_0,\dots,a_{n-1}) = a_0 p(a_1,\dots,a_{n-1}) + p(a_2,\dots,a_{n-1}),$$
$$p(a_0,\dots,a_{n-1}) = p(a_0,\dots,a_{n-2})a_{n-1} + p(a_0,\dots,a_{n-3}).$$

(ii) If the ring is commutative, one has $p(a_0,\dots,a_{n-1}) = p(a_{n-1},\dots,a_0)$.

(iii) If the ring is commutative, one has $P(a_0,\dots,a_{n-1})_{21} = P(a_0,a_{n-1},\dots,a_1)_{21}$.

(iv) One has $p(1,a,b,c,\dots) = p(1+a,b,c,\dots)$.

(v) If $i \leq j \leq k \leq l$ and if the u_i are positive integers, then $p(u_i,\dots,u_l) \geq p(u_j,\dots,u_k)$.

(vi) If $n \geq 0$, and $u_1,\dots,u_n$ are positive integers, then $p(u_1,\dots,u_n) \geq 1$; if moreover $n \geq 2$, then $p(u_1,\dots,u_n) \geq 2$.

(vii) If $n \geq 0$, and $u_1,\dots,u_n$ are positive integers, then $p(u_1,\dots,u_n) > p(u_1,\dots,u_{n-1})$, except if $n = 1$ and $u_1 = 1$, in which case $p(u_1,\dots,u_n) = p(u_1,\dots,u_{n-1}) = 1$.

Note that for $n = 0$, $p(u_1,\dots,u_{n-1})$ means $p_{-1} = 0$.

Proof. The recursive formulas (i) are obtained by associativity of the matrix product in (2) (see [9] 2.7). Formula (ii) is deduced by transposition of the matrix product in (2), noting that the matrices $P(a)$ are symmetric; and (iii) follows from (2) and from (ii); (iv) follows by induction from (i) (see e.g. [24] p.91 for the details). And (v) follows from (2), the fact that $p(u_r,\dots,u_s)$ is the 1,1-entry of $P(u_r,\dots,u_s)$, that $P(u_j,\dots,u_k)$ is a subproduct of $P(u_i,\dots,u_l)$ and that all u_j are ≥ 1. Assertion (vi) is immediate by induction and the last assertion (vii) follows from the second recursion in (i), and the fact that by (vi) all continuant polynomials are positive, except $p_{-1} = 0$. □

We mention the *leapfrog construction* as it is stated in [9] p.117: $p(a_0,\dots,a_{n-1})$ is equal to the sum of $a_0\cdots a_{n-1}$ and of all terms obtained by omitting one or more pairs of adjacent factors $a_i a_{i+1}$. For example, $p(a) = a$, $p(a,b) = ab+1$, $p(a,b,c) = abc+a+c$, $p(a,b,c,d) = abcd + ab + ad + cd + 1$.

For later use, we quote the following lemma. I thank the referee for pointing out a very simple proof of it.

Lemma 3.2. *Let $1 \leq i \leq j$ and $u_1, \dots, u_i$, $v_1, \dots, v_j$ be positive integers such that $p(u_1, \dots, u_i) = p(v_1, \dots, v_j)$ and $p(u_1, \dots, u_{i-1}) = p(v_1, \dots, v_{j-1})$. Then*
either $i = j$ and $u_h = v_h$ for any h;
or $j = i + 1$, $v_1 = 1$, $u_1 = v_2 + 1$, and $u_h = v_{h+1}$ for $h = 2, \dots, i$.

Proof. Recall that each finite continued fraction $[x_1, \dots, x_n]$ is equal to $p(x_1, \dots, x_n)/p(x_2, \dots, x_n)$. The hypothesis and Lemma 3.1 (ii) therefore imply that $[u_i, \dots, u_1] = [v_j, \dots, v_1]$. The lemma then follows from Theorem 162 in [18]. □

4. Convergents of the conjugate of a reduced quadratic number

It is well-known that a periodic continued fraction $[\overline{a_0, \dots, a_{n-1}}]$ represents a quadratic number α which is *reduced*, that is, $\alpha > 1$ and its conjugate $\bar{\alpha}$ is in the interval $(-1, 0)$, and conversely; moreover, one has $\bar{\alpha} = -\widetilde{\alpha}^{-1}$ where $\widetilde{\alpha} = [\overline{a_{n-1}, \dots, a_0}]$ (theorems of Galois; see Proposition 1.18 and Lemma 1.28 in [1]).

In the sequel, we need to know the convergents of $\bar{\alpha}$, as a function of those of $\widetilde{\alpha}$. We use two of the twenty eight cases of Theorem 1 of [19], where the continued fraction expansion of the conjugate of any quadratic number is determined.

Proposition 4.1 *(Herzog [19]). Let $\alpha = [\overline{a_0, \dots, a_{n-1}}]$.*
If $a_{n-1} \geq 2$, then $\bar{\alpha} = [-1, 1, a_{n-1} - 1, \overline{a_{n-2}, \dots, a_0, a_{n-1}}]$.
If $a_{n-1} = 1$, then $\bar{\alpha} = [-1, a_{n-2} + 1, \overline{a_{n-3}, \dots, a_0, a_{n-1}, a_{n-2}}]$.

Corollary 4.1. *With the same notations, let u_k/v_k (resp. p_k/q_k), $k \geq 0$, be the convergents of $\widetilde{\alpha}$ (resp. $\bar{\alpha}$).*
If $a_{n-1} \geq 2$, then

$$p_0/q_0 = -1, p_1/q_1 = 0, \text{ and for } i \geq 2, p_i/q_i = -v_{i-2}/u_{i-2}.$$

If $a_{n-1} = 1$, then

$$p_i/q_i = -v_i/u_i.$$

Lemma 4.1. *If $x = [0, b, c, d, \dots]$, then $-x = [-1, 1, b - 1, c, d, \dots]$ if $b \geq 2$, and $-x = [-1, c + 1, d, \dots]$ if $b = 1$.*

Proof. It is enough to verify the following identities

$$[0, b, y] + [-1, 1, b - 1, y] = 0, [0, 1, c, y] + [-1, c + 1, y] = 0.$$

This verification is left to the reader. □

Proof of the corollary. Suppose first that $a_{n-1} \geq 2$. By Proposition 4.1, we have $\bar{\alpha} = [-1, 1, a_{n-1} - 1, \overline{a_{n-2}, \ldots, a_0, a_{n-1}}]$. The first two convergents of $\bar{\alpha}$ are clearly -1 and 0. The next ones are the numbers $[-1, 1, a_{n-1} - 1, a_{n-2}, \ldots]$. By Lemma 4.1 (first case), this is equal to $-[0, a_{n-1}, a_{n-2}, \ldots] = -[a_{n-1}, a_{n-2}, \ldots]^{-1}$ and the result follows in this case.

Suppose now that $a_{n-1} = 1$. By Proposition 4.1, we have $\bar{\alpha} = [-1, a_{n-2} + 1, \overline{a_{n-3}, \ldots, a_1, a_{n-1}, a_{n-2}}]$. Its first two convergents are the numbers $[-1] = -1 = -1/a_{n-1} = -v_0/u_0$, $[-1, a_{n-2} + 1] = -1 + 1/(a_{n-2} + 1) = -a_{n-2}/(a_{n-2} + 1) = -v_1/u_1$ (since $u_1/v_1 = 1 + 1/a_{n-2}$); the others are the numbers $[-1, a_{n-2} + 1, a_{n-3}, \ldots]$. By Lemma 4.1 (second case), this is equal to $-[0, 1, a_{n-2}, a_{n-3}, \ldots] = -[a_{n-1}, a_{n-2}, a_{n-3}, \ldots]^{-1}$ and the result follows. □

5. Quadratic numbers equivalent to their conjugate

Recall that two real numbers x, y are called $GL_2(\mathbb{Z})$-*equivalent* if they are in the same orbit under the action of $GL_2(\mathbb{Z})$ by Möbius transformation; in other words, $y = \mu_A(x)$ for some $A \in GL_2(\mathbb{Z})$, see Section 7. By a theorem of Serret, equivalently, these two numbers have ultimately the same expansion into continued fractions (Theorem 175 [18], or Proposition 5.22 [1]).

The following result should be well-known (the equivalence between (ii) and (iii) is well-known in combinatorics on words).

Proposition 5.1. *Let $\beta = [b_0, \ldots, b_{k-1}, \overline{a_0, \ldots, a_{n-1}}]$ be a quadratic number with n minimal. Then the following conditions are equivalent:*

(i) β is $GL_2(\mathbb{Z})$-equivalent to its conjugate $\bar{\beta}$;

(ii) the word $w = a_0 \cdots a_{n-1}$ is conjugate with its reversal $\widetilde{w}$;

(iii) the word $a_0 \cdots a_{n-1}$ is a product of two palindromes.

Proof. We claim that $\bar{\beta}$ is $GL_2(\mathbb{Z})$-equivalent to $\tilde{\alpha} = [\overline{a_{n-1}, \ldots, a_0}]$. Indeed, β is $GL_2(\mathbb{Z})$-equivalent to $\alpha = [\overline{a_0, \ldots, a_{n-1}}]$. Therefore $\bar{\beta}$ is $GL_2(\mathbb{Z})$-equivalent to $\bar{\alpha}$ (since conjugation is a $\mathbb{Q}$-automorphism of quadratic fields). By the theorem of Galois, $\bar{\alpha} = -\tilde{\alpha}^{-1}$. Now, $-\tilde{\alpha}^{-1}$ is $GL_2(\mathbb{Z})$-equivalent to $\tilde{\alpha}$, which implies the claim.

Suppose that (i) holds. Then by the claim, β is $GL_2(\mathbb{Z})$-equivalent to $\tilde{\alpha}$. By Serret's theorem (quoted before the statement), two quadratic numbers are $GL_2(\mathbb{Z})$-equivalent if and only if their minimal periodic patterns are conjugate words. Thus (ii) follows.

Conversely, if (ii) holds, then by Serret's theorem, β is $GL_2(\mathbb{Z})$-equivalent to $\tilde{\alpha}$. Hence the claim implies that β is $GL_2(\mathbb{Z})$-equivalent to $\bar{\beta}$. Thus (i) holds.

If (ii) holds, then for some words u, v, $w = uv$, $\tilde{w} = vu$. Therefore $vu = \tilde{w} = \widetilde{uv} = \tilde{v}\tilde{u}$, hence $v = \tilde{v}$, $u = \tilde{u}$. Thus (iii) holds.

If (iii) holds, then $w = uv$, u, v palindromes, and then $\widetilde{w} = vu$ is conjugate with w. □

6. A theorem of Serret

We give here a theorem of Serret (which is not the same as the one in the previous section). We consider *indefinite integral binary quadratic forms* $f(x,y) = ax^2+bxy+cy^2$, $a,b,c \in \mathbb{Z}$, $(a,b,c) \neq (0,0,0)$, with *discriminant* $d(f) = b^2 - 4ac > 0$. We always assume that the polynomial ax^2+bx+c has no rational roots; equivalently, that $f(x,y) = 0$ for $x,y \in \mathbb{Z}$ only if $x = y = 0$; equivalently, that $f(x,y)$ does not factorize over $\mathbb{Z}$.

We call *small value* of the form f any nonzero number $f(p,q)$, $p,q \in \mathbb{Z}$, $(p,q) \neq (0,0)$, such that $|f(p,q)| < \frac{1}{2}\sqrt{d(f)}$. We say that a small value is *proper* if moreover p,q are relatively prime. These definitions are motivated by the theorem of Serret below.

Note that such values exist. Indeed, for the *minimum* $m(f) = \min\{|f(p,q)|, (p,q) \in \mathbb{Z}^2 \setminus (0,0)\}$ of f, one has the inequality $m(f) < \frac{1}{\sqrt{5}}\sqrt{d(f)}$, by a theorem of Korkine and Zolotareff (see e.g. Corollary 9.3.4 in [24]).

Theorem 6.1 *([25] Théorème p. 79-80). Let $f(x,y) = ax^2 + bxy + cy^2$ be an integral indefinite quadratic form, with $a > 0, c < 0$. If p,q are nonzero integers such that $|f(p,q)| < \frac{1}{2}\sqrt{d(f)}$, then p/q is a convergent of one of the roots of the equation $ax^2 + bx + c = 0$.*

For the reader who wants to have a look at [25], note that the hypothesis $a > 0$ is stated there on p. 82 ("nous pouvons supposer D positif") and the hypothesis $c < 0$ is on p. 83 ('ce cas d'exception ne peut se produire que si F est positif").[2] Note also that in Serret's statement, the coefficient of xy is even; but the proof works also when this coefficient is any integer.

The converse of Serret's theorem is not true, as the example in the next section shows.

Corollary 6.1. *If $f(p,q)$ is a small value of f and if $p,q \neq 0$, then p/q is a convergent of α or of its conjugate.*

Part 2. On quadratic numbers and quadratic forms

7. An identity with continuant polynomials

Theorem 7.1. *Let $a_0, a_1, \ldots$ be a periodic infinite sequence of period n of elements of a commutative ring. Let $i \geq 0$ and $i+1 = nq+r$, $0 \leq r \leq n-1$ (Euclidean division). Then*

$$p(a_1,\ldots,a_{n-1})p(a_0,\ldots,a_i)^2$$
$$+(p(a_1,\ldots,a_{n-2}) - p(a_0,\ldots,a_{n-1}))p(a_0,\ldots,a_i)p(a_1,\ldots,a_i)$$

[2] The coefficients of the form are $D, -2E, F$.

$$-p(a_0,\dots,a_{n-2})p(a_1,\dots,a_i)^2$$
$$=(-1)^{i+1}p(a_{r+1},\dots,a_{n-1},a_0,\dots,a_{r-1}).$$

Note that the last expression $(a_{r+1},\dots,a_{n-1},a_0,\dots,a_{r-1})$ may be obtained as follows: take the cyclic conjugate of $(a_0,\dots,a_{n-1})$ beginning by a_r, and remove this a_r.

The proposition may be equivalently stated using a certain quadratic form.

Corollary 7.1. *With the same notations, let f be the quadratic form*

$$f(x,y)=p(a_1,\dots,a_{n-1})x^2+(p(a_1,\dots,a_{n-2}) \tag{3}$$
$$-p(a_0,\dots,a_{n-1}))xy-p(a_0,\dots,a_{n-2})y^2.$$

Then

$$f(p(a_0,\dots,a_i),p(a_1,\dots,a_i))=(-1)^{i+1}p(a_{r+1},\dots,a_{n-1},a_0,\dots,a_{r-1}).$$

For this proof and for later use, it is convenient to introduce some notation. Given a 2 by 2 matrix A, we associate with it the quadratic form

$$f_A(x,y)=cx^2+(d-a)xy-by^2, A=\begin{pmatrix}a&b\\c&d\end{pmatrix}. \tag{4}$$

In other words, f_A is the form which homogenizes the quadratic polynomial whose root x is a fixed point of the Möbius transformation

$$\mu_A: x\mapsto A\cdot x=\frac{ax+b}{cx+d}$$

associated to A; indeed, one has $x=\frac{ax+b}{cx+d}$, equivalently $cx^2+(d-a)x-b=0$. Note that the latter quadratic polynomial is $f_A(x,1)$.

One has

$$f_A(x,y)=(-y,x)A\begin{pmatrix}x\\y\end{pmatrix}. \tag{5}$$

For further use, note that the discriminant of f_A is given by

$$d(f_A)=\mathrm{Tr}(A)^2-4\,\mathrm{Det}(A). \tag{6}$$

In particular, conjugate or transposed matrices give the same discriminant.

Proof of Theorem 7.1. It is enough to prove the corollary. Note that since the determinant of $P(a)$ is -1, the inverse of the matrix $P(a_0)\cdots P(a_i)$ is by (2) equal to

$$(-1)^{i+1}\begin{pmatrix} p(a_1,\dots,a_{i-1}) & -p(a_0,\dots,a_{i-1}) \\ -p(a_1,\dots,a_i) & p(a_0,\dots,a_i)\end{pmatrix}. \tag{7}$$

Let $x = p(a_0,\dots,a_i), y = p(a_1,\dots,a_i)$. Then by (2), $\begin{pmatrix} x \\ y\end{pmatrix}$ is the first column of $P(a_0)\cdots P(a_i)$ and by (7), $(-y,x)$ is the second row of $(P(a_0)\cdots P(a_i))^{-1}$. Observe that $f = f_A$ with $A = P(a_0)\cdots P(a_{n-1})$. It follows, in view of (5), that $f(x,y)$ is equal to the 2,1-entry of the product of the three matrices

$$(-1)^{i+1}(P(a_0)\cdots P(a_i))^{-1}P(a_0)\cdots P(a_{n-1})P(a_0)\cdots P(a_i).$$

By periodicity of the sequence, the product $P(a_0)\cdots P(a_i)$ is equal to $P(a_0)\cdots P(a_{n-1})^q P(a_0)\cdots P(a_{r-1})$. Hence the previous product of matrices is equal to

$$\begin{aligned}(-1)^{i+1}(P(a_0)\dots P(a_{r-1}))^{-1}(P(a_0)\dots P(a_{n-1}))^{-q}P(a_0)\dots P(a_{n-1})\\ (P(a_0)\dots P(a_{n-1}))^q P(a_0)\dots P(a_{r-1})\\ = (-1)^{i+1}(P(a_0)\dots P(a_{r-1}))^{-1}P(a_0)\dots P(a_{n-1})P(a_0)\dots P(a_{r-1})\\ = (-1)^{i+1}P(a_r)\dots P(a_{n-1})P(a_0)\dots P(a_{r-1}).\end{aligned}$$

Thus the corollary follows from (2). □

8. Applications to small values of quadratic forms

In this section, we determine the exact values of the form which appear in the conclusion of Serret's theorem. This gives a partial converse to this theorem; the counterexample below shows why it is only a partial converse (however for Markoff forms, we shall obtain a complete converse, see Section 17).

We restrict our attention to the quadratic forms f_A of (4), naturally associated as below to periodic continued fractions. However, they are general enough, in the sense that each quadratic form is $GL_2(\mathbb{Z})$-equivalent to a form $f(x,y)$ such that the roots of the polynomial $f(x,1)$ are $\xi > 1$ and $-1 < \eta < 0$ (see e.g. Theorem 9.1.1 in [24]). The form f is then necessarily of the shape λf_A, with A as below and some integer λ.

Given a periodic continued fraction $\alpha = [\overline{a_0,\dots,a_{n-1}}]$, $n \geq 1$, $a_i \in \mathbb{Z}_{>0}$, associate to it the matrix $A = P(a_0)\cdots P(a_{n-1}) = \begin{pmatrix} a & b \\ c & d\end{pmatrix}$. Then, as is well-known, the real number α represented by the continued fraction is a fixed point of the Möbius transformation μ_A; in other words $f_A(\alpha,1) = 0$. The other root is the conjugate $\bar{\alpha}$.

Theorem 8.1. *Let $f = f_A$ and define ϵ_r, $r \in \{0, \dots, n-1\}$, by*

$$\epsilon_r = p(a_{r+1}, \dots, a_{n-1}, a_0, \dots, a_{r-1}) \tag{8}$$

(i) Let p_i/q_i be the convergents of α. Then for any $i \geq 0$, $f(p_i, q_i) = (-1)^{i+1}\epsilon_{i+1 \bmod n}$.
(ii) Let p_i/q_i be the convergents of $\bar{\alpha}$.
If $a_{n-1} \geq 2$, then for any $i \geq 1$, $f(p_i, q_i) = (-1)^i \epsilon_{-i \bmod n}$.
If $a_{n-1} = 1$, then for any $i \geq 0$, $f(p_i, q_i) = (-1)^i \epsilon_{-i-2 \bmod n}$.

Note that in accordance with a previous convention, for $n = 1$, one has $\epsilon_0 = 1$.

Proof. 1. We extend the sequence of a_i by periodicity: for any i, $a_i = a_{i \bmod n}$. In view of (2) and (4), f is given by (3). Thus for $i \geq 0$, we have $f(p_i, q_i) = (-1)^{i+1}\epsilon_{i+1 \bmod n}$ by Corollary 7.1. This proves (i).

2. Consider now the periodic continued fraction $[\overline{b_0, \dots, b_{n-1}}] = [\overline{a_{n-1}, \dots, a_0}]$, representing $\widetilde{\alpha}$. Let u_i/v_i, $i \geq 0$, be the convergents of $\widetilde{\alpha}$. It follows from the previous part of the proof that $f_B(u_i, v_i) = (-1)^{i+1}\eta_{i+1 \bmod n}$, where $B = P(a_{n-1}) \cdots P(a_0)$ and $\eta_r = p(b_{r+1}, \dots, b_{n-1}, b_0, \dots, b_{r-1})$ if $r = 0, \dots, n-1$.

The matrices $P(a)$ being symmetric, we have $B = A^T$ and therefore (with the notations of (4)) $f(x, y) = -(by^2 + (d-a)(-x)y - cx^2) = -f_B(y, -x)$.

3. Suppose that $a_{n-1} \geq 2$. Then by Corollary 4.1, $p_1 = 0, q_1 = 1$, and for any $i = 2, 3 \dots$, $p_i = -v_{i-2}$ and $q_i = u_{i-2}$. Thus $f(p_1, q_1) = f(0, 1) = -p(a_0, \dots, a_{n-2}) = (-1)^1 \epsilon_{-1 \bmod n}$, since $-1 \bmod n = n-1$; and for any $i \geq 2$, $f(p_i, q_i) = -f_B(q_i, -p_i) = -f_B(u_{i-2}, v_{i-2}) = -(-1)^{i-1}\eta_{i-1 \bmod n} = (-1)^i p(b_{r+1}, \dots, b_{n-1}, b_0, \dots, b_{r-1})$, with $r = i - 1 \bmod n$. Hence $f(p_i, q_i) = (-1)^i p(b_{r-1}, \dots, b_0, b_{n-1}, \dots, b_{r+1}) = (-1)^i p(a_{n-r}, \dots, a_{n-1}, a_0, \dots, a_{n-r-2}) = (-1)^i \epsilon_{-i \bmod n}$, since $b_j = a_{n-1-j}$ and since modulo n, we have $-i \equiv n - r - 1$.

4. Suppose now that $a_{n-1} = 1$. Then by Corollary 4.1, for any $i \geq 0$, $p_i = -v_i$ and $q_i = u_i$. Thus, $f(p_i, q_i) = -f_B(q_i, -p_i) = -f_B(u_i, v_i) = -(-1)^{i+1}\eta_{i+1 \bmod n} = (-1)^i p(b_{r+1}, \dots, b_{n-1}, b_0, \dots, b_{r-1})$, with $r = i+1 \bmod n$. Hence $f(p_i, q_i) = (-1)^i p(b_{r-1}, \dots, b_0, b_{n-1}, \dots, b_{r+1}) = (-1)^i p(a_{n-r}, \dots, a_{n-1}, a_0, \dots, a_{n-r-2}) = (-1)^i \epsilon_{-i-2 \bmod n}$, since $b_j = a_{n-1-j}$ and since modulo n, we have $-i - 2 \equiv n - r - 1$. □

Corollary 8.1. *The proper small values of $f = f_A$ are in the set $\{\pm p(a_{r+1}, \dots, a_{n-1}, a_0, \dots, a_{r-1}), r = 0, \dots, n-1\}$.*

We need a lemma; it settles the case not covered by Theorem 8.1 ($a_{n-1} \geq 2, i = 0, p_0 = -1, q_0 = 1$).

Lemma 8.1. *If $a_{n-1} \geq 2$, $f(-1, 1)$ is not a small value of f.*

Proof. By (6) $d(f) = (a+d)^2 - 4(-1)^n$. We have $f(-1, 1) = c + a - d - b$ and to prove the lemma is it is enough to show that $4f(-1, 1)^2 - d(f) > 0$.

We have $4f(-1,1)^2-d(f) = (2(c+a-d-b)+a+d)(2(c+a-d-b)-a-d)+4(-1)^n = (2c-2b+3a-d)(2c-2b+a-3d)+4(-1)^n$.

Let $a_{n-1} = h+2$, $h \geq 0$. We assume first that $n \geq 3$. Then by (2), the second factor is $2c-2b+a-3d = 2p(a_1,\dots a_{n-2},h+2)-2p(a_0,\dots,a_{n-2})+p(a_0,\dots,a_{n-2},h+2)-3p(a_1,\dots,a_{n-2}) = (2h+4)p(a_1,\dots,a_{n-2})+2p(a_1,\dots,a_{n-3})-2p(a_0,\dots,a_{n-2})+(h+2)p(a_0,\dots,a_{n-2})+p(a_0,\dots,a_{n-3})-3p(a_1,\dots,a_{n-2}) = (2h+1)p(a_1,\dots,a_{n-2})+2p(a_1,\dots,a_{n-3})+hp(a_0,\dots,a_{n-2})+p(a_0,\dots,a_{n-3}) \geq 4$, since $n \geq 3$.

Now the first factor $2c-2b+3a-d$ is equal to the second one plus $2(a+d)$, hence is ≥ 8. It follows that their product is ≥ 32 and finally, $4f(-1,1)^2-d(f) > 0$.

Suppose now that $n=1$. Then $f = x^2-a_0xy-y^2$, $d(f) = a_0^2+4$, $f(-1,1) = a_0$, $4f(-1,1)^2-d(f) = 3a_0^2-4 > 0$ since $a_0 \geq 2$.

Suppose finally that $n=2$. Then $P(a_0,a_1) = \begin{pmatrix} a_0a_1+1 & a_0 \\ a_1 & 1 \end{pmatrix}$, so that the second factor is $2c-2b+a-3d = 2a_1-2a_0+a_0a_1+1-3 = 2(a_1-1)+a_0(a_1-2) \geq 2$ since $a_1 \geq 2$; the first factor is thus ≥ 6, and their product is ≥ 12, so that $4f(-1,1)^2-d(f) \geq 12-4 > 0$. □

Proof of Corollary 8.1. Let this small value be $f(p,q)$. Then the result is clear if $(p,q) = \pm(1,0)$ or $(p,q) = \pm(0,1)$, since $f(1,0) = p(a_1,\dots,a_{n-1})$ and $f(0,1) = -p(a_0,\dots,a_{n-2})$. In general, we apply Serret's theorem and the previous proposition and lemma. □

To complete the picture, note that the numbers ϵ_t are not always smaller than $\frac{1}{2}\sqrt{d(f)}$; this is shown by the example of $\alpha = [\overline{2,1}]$, $P(2,1) = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$, $f = x^2-2xy-2y^2$, $d(f) = 12$, $\frac{1}{2}\sqrt{d(f)} = \sqrt{3}$; moreover $\epsilon_1 = 2 > \sqrt{3}$.

However, they are all smaller that $\sqrt{d(f)}$ as the following result shows.

Proposition 8.1. *With the notations of the previous proposition, the numbers ϵ_i are always smaller than $\sqrt{d(f)}$.*

Note that a theorem of Hurwitz ([20] Satz p.427), which improves Serret's Theorem 6.1, states that if $|f(p,q)| < \sqrt{d(f)}$, then p/q must be a semi-convergent of one of the roots of the polynomial $f(x,1)$. Recall that a *semi-convergent* of a real number with continued fraction expansion $[a_0,a_1,a_2,\dots]$ is any rational number of the form $[a_0,\dots,a_{n-1},a]$, with $n \geq 0, 1 \leq a \leq a_n$.

Proof. By (6), the discriminant is $d(f) = \mathrm{Tr}(A)^2 - 4\,\mathrm{Det}(A) = (p(a_0,\dots,a_{n-1})+p(a_1,\dots,a_{n-2}))^2 - 4(-1)^n$, by (2). Since trace and determinant are invariant under conjugation, the discriminant is unchanged if the matrix product is conjugated. Therefore, it is enough to show that $\epsilon_0 < \sqrt{d(f)}$. By Lemma 3.1 (v), $\epsilon_0 = p(a_1,\dots,a_{n-1}) \leq p(a_0,\dots,a_{n-1})$. This implies the desired inequality when n is odd. When n is even, then

$n \geq 2$, $p(a_2, \dots, a_{n-1}) \geq 1$ and $p(a_1, \dots, a_{n-2}) \geq 1$. Thus $2 + \epsilon_0 = p(a_1, \dots, a_{n-1}) + 2 \leq a_0 p(a_1, \dots, a_{n-1}) + p(a_2, \dots, a_{n-1}) + p(a_1, \dots, a_{n-2}) = p(a_0, \dots, a_{n-1}) + p(a_1, \dots, a_{n-2})$, and consequently $\epsilon_0^2 + 4 < (\epsilon_0 + 2)^2 \leq (p(a_0, \dots, a_{n-1}) + p(a_1, \dots, a_{n-2}))^2 = \mathrm{Tr}(A)^2$, which implies $\epsilon_0^2 < \mathrm{Tr}(A)^2 - 4$, as desired. □

9. Error term for quadratic numbers

We consider in this section two quadratic numbers $\alpha = [\overline{a_0, \dots, a_{n-1}}]$ and $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$. We define the matrices $A = P(a_0, \dots, a_{n-1})$, $M = P(b_0, \dots, b_{k-1})^{-1} = \begin{pmatrix} e & h \\ i & j \end{pmatrix}$, and the quadratic form $f = f_A$. Let ϵ_i be as in (8) and let u_t/v_t and p_t/q_t be the convergents of α and β respectively. Note that for $t \geq k$,

$$\begin{pmatrix} p_t \\ q_t \end{pmatrix} = M^{-1} \begin{pmatrix} u_{t-k} \\ v_{t-k} \end{pmatrix}. \tag{9}$$

Theorem 9.1. *Let $\alpha = [\overline{a_0, \dots, a_{n-1}}]$ and $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$. The convergents p_t/q_t of β satisfy for $t \geq k$*

$$\beta - \frac{p_t}{q_t} = \frac{(-1)^t \epsilon_{t-k+1 \bmod n}}{q_t^2 \frac{\sqrt{d(f)}}{2} (1 + \sqrt{1 + 4(-1)^{t-k+1} f(e,i) \epsilon_{t-k+1 \bmod n} / d(f) q_t^2})}.$$

A particular case is when $k = 0$, that is, $\alpha = \beta$, and therefore the p_t/q_t are the convergents of α.

Corollary 9.1. *For $t \geq 0$,*

$$\alpha - \frac{p_t}{q_t} = \frac{(-1)^t \epsilon_{t+1 \bmod n}}{q_t^2 \frac{\sqrt{d(f)}}{2} (1 + \sqrt{1 + 4(-1)^{t+1} f(1,0) \epsilon_{t+1 \bmod n} / d(f) q_t^2})}.$$

Note that $f(1,0) = \epsilon_0$. Observe the dependance on t in the equations of the theorem (and similarly in the corollary): only q_t and $(-1)^{t-k+1} \epsilon_{t-k+1 \bmod n}$ depend on t, and the latter is periodic function of t, while the other quantities are independent of t, depending only on β.

We need several lemmas. The following one is well-known, and its proof is a straightforward computation. utiliser (5)

Lemma 9.1. *Let $B = M^{-1}AM$ with A, B, M being 2 by 2 matrices over $\mathbb{R}$ and M in $GL_2(\mathbb{R})$. Then $f_B = \frac{1}{\det(M)} f_A.M$.*

We have used here the classical right action of 2 by 2 matrices on quadratic form: if $f(x,y)$ is a quadratic form and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $(f \cdot M)(x,y) = f(ax + by, cx + dy)$.

Note that for a better understanding of this right action, it may be useful to view f as a function defined on columns vectors $\mathbf{x} = {}^t(x, y)$, written $f(\mathbf{x})$. Then $(f \cdot M)(\mathbf{x}) = f(M\mathbf{x})$. This implies that it is a right action, since Mv defines a left action of the matrices on the column vectors.

Given a quadratic polynomial $F(x) = ax^2 + bx + c$ over $\mathbb{R}$, with positive discriminant Δ, we call *positive root* (resp. *negative root*) of $F(x)$ the root $\frac{-b+\sqrt{\Delta}}{2a}$ (resp. $\frac{-b-\sqrt{\Delta}}{2a}$). Observe that if x is a root of F, it is the positive root if and only if $2ax + b \geq 0$, that is, $F'(x) \geq 0$.

+-root

Note that if $a > 0$ and $c < 0$, then the sign of a root of F is its actual sign as real number. Clearly, F and dF have the same positive and negative roots if $d > 0$, while they are interchanged if $d < 0$.

Given a quadratic form $f(x, y) = ax^2 + bxy + cy^2$ over $\mathbb{R}$, with positive discriminant Δ, we call *positive root* (resp. *negative root*) of f the positive root (resp. negative) root of the polynomial $f(x, 1)$.

Lemma 9.2. *If $g = f \cdot M$ and θ is the positive root of g, then $M \cdot \theta$ is the root of f whose sign is that of* $\det(M)$.

Recall that $M \cdot \theta = \mu_M(\theta)$, see Section 7. The proof of this lemma rests on a straightforward computation (see [14] Theorem 72).

Lemma 9.3. *Let β, β' be the respective positive roots of the quadratic polynomials $px^2 + qx + r$ and $px^2 + qx + r'$, of respective discriminant $\Delta > 0$ and $\Delta' \geq 0$. Then*

$$\beta - \beta' = \frac{r' - r}{\frac{\sqrt{\Delta}}{2}(1 + \sqrt{1 + \frac{4p(r-r')}{\Delta}})}.$$

Proof. We have $\beta - \beta' = \frac{-q+\sqrt{\Delta}}{2p} - \frac{-q+\sqrt{\Delta'}}{2p} = \frac{\sqrt{\Delta}-\sqrt{\Delta'}}{2p} = \frac{\Delta-\Delta'}{2p(\sqrt{\Delta}+\sqrt{\Delta'})} = \frac{4p(r'-r)}{2p\sqrt{\Delta}(1+\sqrt{\frac{\Delta'}{\Delta}})} = \frac{r'-r}{\frac{\sqrt{\Delta}}{2}(1+\sqrt{\frac{\Delta'}{\Delta}})}$. Moreover, $\frac{\Delta'}{\Delta} = 1 + \frac{\Delta'-\Delta}{\Delta} = 1 + \frac{4p(r-r')}{\Delta}$. The formula follows. □

The following crucial lemma will imply that the convergent p_t/q_t appearing in Theorem 9.1 is the *positive* root of a certain quadratic polynomial. The proof is very technical and the reader could skip it and proceed directly to the proof of the theorem. Let $B = M^{-1}AM$ and $g = f_B$.

Lemma 9.4. *For any $t \geq k$, one has $g'_x(p_t, q_t) \geq 0$.*

Proof. 1. We have by (5), $g(x, y) = (-y, x)B\begin{pmatrix} x \\ y \end{pmatrix}$. Thus $g'_x(x, y) = (0, 1)B\begin{pmatrix} x \\ y \end{pmatrix} + (-y, x)B\begin{pmatrix} 1 \\ 0 \end{pmatrix}$; thus

$$g'_x(p_t,q_t) = (0,1)M^{-1}AM\begin{pmatrix} p_t \\ q_t \end{pmatrix} + (-q_t,p_t)M^{-1}AM\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

By (9), $M(p_t,q_t)^T = (u_{t-k},v_{t-k})^T$, which is the first column of $P(a_0,\dots,a_{t-k})$, where the a_t are defined for any natural t by $a_t = a_{t \bmod n}$. Thus, the first term in $g'_x(p_t,q_t)$ is the $2,1$-entry of the product

$$\begin{aligned} &P(b_0,\dots,b_{k-1})P(a_0,\dots,a_{n-1})P(a_0,\dots,a_{t-k}) \\ &\quad = P(b_0,\dots,b_{k-1},a_0,\dots,a_{t-k+n}). \end{aligned}$$

Hence this term is $p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k+n})$ if $k \geq 1$ and $p(a_1,\dots,a_{t+n})$ if $k = 0$.

By (7) $(-q_t,p_t)$ is the second row of $(-1)^{t+1}P(b_0,\dots,b_{k-1},a_0,\dots,a_{t-k})^{-1}$. Hence the second term is the $2,1$-entry of

$$\begin{aligned} &(-1)^{t+1}P(b_0,\dots,b_{k-1},a_0,\dots,a_{t-k})^{-1}P(b_0,\dots,b_{k-1}) \\ &\qquad P(a_0,\dots,a_{n-1})P(b_0,\dots,b_{k-1})^{-1} \\ &= (-1)^{t+1}P(a_0,\dots,a_{t-k})^{-1}P(a_0,\dots,a_{n-1})P(b_0,\dots,b_{k-1})^{-1}. \end{aligned}$$

If $t-k \geq n-1$, this matrix product is equal to

$$\begin{aligned} P(a_n,\dots,a_{t-k})^{-1}P(b_0,\dots,b_{k-1})^{-1} &= P(b_0,\dots,b_{k-1},a_n,\dots,a_{t-k})^{-1} \\ &= P(b_0,\dots,b_{k-1},a_0,\dots,a_{t-k-n})^{-1}. \end{aligned}$$

Hence the second term is by (7) equal to

$$\begin{aligned} &-(-1)^{t+1}(-1)^{k+t-k-n+1}p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k-n}) \\ &\quad = (-1)^{n+1}p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k-n}) \end{aligned}$$

in the case where $k \geq 1$, and in the case $k = 0$, it is $(-1)^{n+1}p(a_1,\dots,a_{t-n})$.

If $t-k < n-1$, the matrix product is $P(a_{t-k+1},\dots,a_{n-1})P(b_0,\dots,b_{k-1})^{-1}$. Hence the second term is by Eqs. (2) and (7) equal to

$$\begin{aligned} &(-1)^{t+1}(-1)^k(p(a_{t-k+2},\dots,a_{n-1})p(b_1,\dots,b_{k-2}) \\ &\qquad -p(a_{t-k+2},\dots,a_{n-2})p(b_1,\dots,b_{k-1})) \end{aligned}$$

if $k \geq 1$ and it is $(-1)^{t+1}p(a_{t+2},\dots,a_{n-1})$ if $k = 0$.

2. We now may show that for $t \geq k$, $g'_x(p_t,q_t) \geq 0$. Suppose first that $t-k \geq n-1$; then $g'_x(p_t,q_t)$ is equal to $p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k+n}) + (-1)^{n+1}p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k-n})$ if $k \geq 1$ and to $p(a_1,\dots,a_{t+n}) + (-1)^{n+1}p(a_1,\dots,a_{t-n})$ if $k = 0$.

It follows from Lemma 3.1(v) that in both cases, $g'_x(p_t,q_t) \geq 0$.

Suppose now that $t-k<n-1$. Then $g'_x(p_t,q_t)$ is equal to

$$p(b_1,\dots,b_{k-1},a_0,\dots,a_{t-k+n})+(-1)^{t+1+k}(p(a_{t-k+2},\dots,a_{n-1})p(b_1,\dots,b_{k-2})\\ -p(a_{t-k+2},\dots,a_{n-2})p(b_1,\dots,b_{k-1}))$$

if $k\geq 1$ and to $p(a_1,\dots,a_{t+n})+(-1)^{t+1}p(a_{t+2},\dots,a_{n-1})$ if $k=0$. In this latter case, we use Lemma 3.1(v) and deduce that $g'_x(p_t,q_t)\geq 0$. In the first case, $g'_x(p_t,q_t)$ is by (2) equal to

$$\begin{aligned}&p(b_1,\dots,b_{k-1})p(a_0,\dots,a_{t-k+n})+p(b_1,\dots,b_{k-2})p(a_1,\dots,a_{t-k+n})\\ &+(-1)^{t+1+k}(p(a_{t-k+2},\dots,a_{n-1})p(b_1,\dots,b_{k-2})\\ &-p(a_{t-k+2},\dots,a_{n-2})p(b_1,\dots,b_{k-1}))\\ &=p(b_1,\dots,b_{k-1})(p(a_0,\dots,a_{t-k+n})-(-1)^{t+1+k}(p(a_{t-k+2},\dots,a_{n-2}))\\ &+p(b_1,\dots,b_{k-2})(p(a_1,\dots,a_{t-k+n})+(-1)^{t+1+k}(p(a_{t-k+2},\dots,a_{n-1}))\end{aligned}$$

It follows from Lemma 3.1(v) that $g'_x(p_t,q_t)\geq 0$. □

Proof of Theorem 9.1. By (9) $(p_t,q_t)^T=M^{-1}(u_{t-k},v_{t-k})^T$, hence $(u_{t-k},v_{t-k})^T=M(p_t,q_t)^T$. By Theorem 8.1 (i), $f(u_t,v_t)=(-1)^{t+1}\epsilon_{t+1 \bmod n}$.

Recall that we have $B=M^{-1}AM$ and $g=f_B$. Then by Lemma 9.1 $g=(-1)^k f\cdot M$. It follows that $g(p_t,q_t)=(-1)^k(f\cdot M)(p_t,q_t)=(-1)^k f(u_{t-k},v_{t-k})=(-1)^k(-1)^{t-k+1}\epsilon_{t-k+1 \bmod n}=(-1)^{t+1}\epsilon$, with $\epsilon=\epsilon_{t-k+1 \bmod n}$.

Writing $g=a'x^2+b'xy+c'y^2$, we obtain $a'(\frac{p_t}{q_t})^2+b'\frac{p_t}{q_t}+c'+(-1)^t\epsilon/q_t^2=0$: p_t/q_t is a root of the polynomial $g(x,1)+(-1)^t\epsilon/q_t^2$.

We have $\beta=[b_0,\dots,b_{k-1},\alpha]=P(b_0)\dots P(b_{k-1})\cdot\alpha=M^{-1}\cdot\alpha$. Since $A\cdot\alpha=\alpha$, we have $B\cdot\beta=M^{-1}AM\cdot\beta=M^{-1}A\cdot\alpha=M^{-1}\cdot\alpha=\beta$ and this implies that β is a root of $g(x,1)$.

Note that α is the positive root of f (because the first coefficient of f is positive, its last coefficient is negative and α is positive). Let θ be the positive root of g. Since $g=(-1)^k f\cdot M$, Lemma 9.2 implies that $M\cdot\theta$ is the root of $(-1)^k f$ whose sign is that of $det(M)$. The latter being $(-1)^k$, $M\cdot\theta$ is the positive root of f, that is, α. Hence $\theta=M^{-1}\cdot\alpha=\beta$, which is therefore the positive root of g.

We claim that for $t\geq k$, p_t/q_t is the positive root of the polynomial $g(x,1)+(-1)^t\epsilon/q_t^2$. It amounts to prove that $2a'p_t/q_t+b'\geq 0$, by a previous remark. This follows from Lemma 9.4, which asserts that $2a'p_t+b'q_t\geq 0$.

Let $\Delta=b'^2-4a'c'$ and $\Delta'=\Delta-4a'(-1)^t\epsilon/q_t^2$, the discriminants of the polynomials $g(x,1)$ and $g(x,1)+(-1)^t\epsilon/q_t^2$.

Note that the discriminants of f and g are equal, because they are obtained each from another by $GL_2(\mathbb{Z})$-equivalence and multiplication by ± 1; hence $\Delta=d(g)=d(f)$. Moreover, a' is the first coefficient of $g=(-1)^k f\cdot M$, which is equal to $(-1)^k f(e,i)$, because $(f.M)(1,0)=f(e,i)$.

Thus the theorem follows from Lemma 9.3. □

10. Intermezzo

The simplest periodic continued fractions are the $\alpha = [\bar{a}]$, where a is a positive integer. One has $\alpha = \frac{a+\sqrt{a^2+4}}{2}$ and $\epsilon_0 = 1$. The corresponding quadratic form is $f(x,y) = x^2 - axy - y^2$, with discriminant a^2+4. For its convergents p_t/q_t, Corollary 9.1 gives the identity

$$\alpha - \frac{p_t}{q_t} = \frac{(-1)^t}{q_t^2 \frac{\sqrt{a^2+4}}{2}(1+\sqrt{1+4(-1)^{t+1}/(a^2+4)q_t^2})}. \tag{10}$$

For the particular case $a = 1$ and 2, x is the *golden ratio* and the *silver ratio*, the sequences are the Fibonacci and Pell numbers, and the corresponding identities were obtained by Hančl [16,17].

For general a, one has by (2), $p_t = p(a,\dots,a)$ ($t+1$ times) and $q_t = p(a,\dots,a)$ (t times), so that the sequences p_t and q_t differ by a shift: $q_t = p_{t-1}$. These sequences both satisfy the linear recursion of length 2: $p_{t+2} = ap_{t+1} + p_t$. The formula above is equivalent to a formula that expresses the numerator p_t as a function of the denominator q_t:

$$p_t = \frac{1}{2}aq_t + \frac{1}{2}\sqrt{(a^2+4)q_t^2 + 4(-1)^{t+1}}.$$

Since $q_t = p_{t-1}$, this gives us as a recursion of length 1 for the sequence p_t (in particular for the Fibonacci numbers and the Pell numbers). It follows also that the integers in the square root are perfect square. For $a = 1$, they are the square of the *Lucas numbers*, see (10.14.7) in [18].

11. Lagrange number

Let ξ be an irrational real number. Consider the set of real numbers L such that the inequality $|\xi - p/q| < 1/Lq^2$ holds for infinitely many rational numbers p/q.

Define $L(\xi)$ to be the supremum of all these L. It is called the *Lagrange number* of ξ .

It is known that

$$L(\xi) = \limsup_{t\to\infty} \lambda_t(\xi),$$

with $|\xi - \frac{p_t}{q_t}| = \frac{1}{\lambda_t(\xi)q_t^2}$, $\lambda_t(\xi) = [a_{t+1}, a_{t+2}\dots] + [a_t, a_{t-1},\dots,a_1]^{-1}$, where ξ has the continued fraction expansion $[a_0, a_1, a_2, \dots]$. See [1] Proposition 1.22, or [24] Theorem 5.5.1.

Corollary 11.1. *The Lagrange number of* $\beta = [b_0,\dots,b_{k-1},\overline{a_0,\cdots,a_{n-1}}]$ *is equal to* $\frac{\sqrt{d(f)}}{m}$, *where* m *is the minimum of the* n *numbers* $p(a_{r+1},\dots,a_{n-1},a_0,\dots,a_{r-1})$, $r = 0,\dots,n-1$.

This result is equivalent to known results; for example Lemma 17 in [2] and Proposition 1.29 in [1].

Proof. By Theorem 9.1 and the previous formula, $\lambda_t(\beta)$ is equal to

$$\frac{\sqrt{d(f)}}{2}\frac{1+\sqrt{1+4(-1)^{t-k+1}f(e,i)\epsilon_{t-k+1 \bmod n}/d(f)q_t^2}}{\epsilon_{t-k+1 \bmod n}}. \tag{11}$$

Since $q_t \to \infty$ and since the sequence $\epsilon_{t-k+1 \bmod n}$ is periodic, we see that the supremum of $\lambda_t(\beta)$ is equal to the maximum of the numbers $\sqrt{d(f)}/\epsilon_r$, which implies the result. □

12. Good approximations of quadratic numbers

Given a real number ξ, we say that the rational p/q is a *good approximation of* ξ if

$$|\xi - p/q| < 1/L(\xi)q^2.$$

The question is if a given ξ has infinitely many good approximations. Not all reals have this property, by an example of Perron [23] p.7-8, see below.

Note that a good approximation is necessarily a convergent (which we call then a *good convergent* and we say that its rank is *good*). Indeed, by the famous Hurwitz $\sqrt{5}$ theorem (see the Introduction), there are infinitely many approximations with $|\xi - p/q| < 1/\sqrt{5}q^2$. Hence $L(\xi) \geq \sqrt{5}$. This implies, by Legendre's theorem,[3] that a good approximation must be a convergent.

Theorem 12.1. *Let $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$. As previously, for $r \in \{0, \dots, n-1\}$, let $\epsilon_r = p(a_{r+1}, \dots, a_{n-1}, a_0, \dots, a_{r-1})$, and denote by m the minimum of these numbers. Let $P(b_0, \dots, b_{k-1})^{-1} = \begin{pmatrix} e & h \\ i & j \end{pmatrix}$ and $f = f_{P(a_0,\dots,a_{n-1})}$. Then for t large enough, a convergent p_t/q_t of β is good if and only if $\epsilon_{t-k+1 \bmod n} = m$ and $(-1)^{t-k+1}f(e,i) > 0$.*

Corollary 12.1. *The quadratic number β has infinitely many good approximations if and only if either n is odd, or n is even and there exists some $r \in \{0, \dots, n-1\}$ such that $\epsilon_r = m$ and that $(-1)^r f(e,i) > 0$.*

Using the corollary, we may revisit the example of Perron. He considers $\xi = [\overline{b, c}]$ with $b > c$. We have $n = 2$, $k = 0$, $\epsilon_0 = c$, $\epsilon_1 = b$, and therefore $m = c$ and the unique r is 0. Since $P(b,c) = \begin{pmatrix} bc+1 & b \\ c & 1 \end{pmatrix}$, the quadratic form f is $f(x,y) = cx^2 - bcxy - by^2$ and moreover $(e,i) = (1,0)$; hence $(-1)^r f(e,i) = c > 0$, so that there are by the corollary infinitely many good approximations.

[3] If $|\xi - p/q| < 1/2q^2$, then p/q is a convergent of ξ.

Suppose now that $\xi = [\overline{c, b}]$. Then $r = 1$, $f(x, y) = bx^2 - bcxy - cy^2$ and $(e, i) = (1, 0)$; then $(-1)^r f(e, i) = -b < 0$ and there are only finitely many good approximations.

Note that Perron gives other arguments.

Proof of Theorem 12.1. We know by Corollary 11.1 that $L(\beta) = \sqrt{d(f)}/m$. Note that t is the rank of a good convergent if and only if $\lambda_t(\beta) > L(\beta)$. By (11), $\lambda_t(\beta)$ is equal to

$$\frac{L(\beta)}{2}\Big(1 + \sqrt{1 + 4(-1)^{t-k+1} f(e, i)\epsilon_{t-k+1 \bmod n}/d(f)q_t^2}\Big)\frac{m}{\epsilon_{t-k+1 \bmod n}}.$$

Therefore, if $\epsilon_{t-k+1 \bmod n} = m$ and if $(-1)^{t-k+1} f(e, i) > 0$, then $\lambda_t(\beta) > L(\beta)$, and t is good. If $\epsilon_{t-k+1 \bmod n} = m$, but $(-1)^{t-k+1} f(e, i) < 0$, then $\lambda_t(\beta) < L(\beta)$. Finally, if $\epsilon_{t-k+1 \bmod n} > m$, then, since $q_t \to \infty$, we see that $\lambda_t(\beta) < L(\beta)$ for t large enough, since there are finitely many ϵ_r. □

Proof of Corollary 12.1. Suppose that n is odd. There exists $r \in 0, \dots, n-1$ such that $\epsilon_r = m$. If $t - k + 1 \equiv r$ modulo n and if t is large enough, then by the theorem, t is good as soon $(-1)^{t-k+1} f(e, i) > 0$. Suppose that this number is negative. Then replacing t by $t + n$, we find a good t, since the sequence ϵ_t has period n.

Suppose now that n is even. If there exists r such that $\epsilon_r = m$ and $(-1)^r f(e, i) > 0$, then t is good if t is large enough and if $t - k + 1 \equiv r$. Then there are infinitely many good t, since t good implies $t + n$ good.

Conversely, if there are infinitely many good t, then there exists t such that $t - k + 1 \bmod n = r$ with $\epsilon_r = m$ and $(-1)^{t-k+1} f(e, i) > 0$. Hence $(-1)^r f(e, i) > 0$, since $t-k+1$ and r have the same parity, n being even. □

We say that a subset of $\mathbb{N}$ is *ultimately periodic* if it is a union of a finite set and of finitely many arithmetic progressions.

Corollary 12.2. *The set of good t is ultimately periodic.*

Proof. This is because the sequence $\epsilon_{t-k+1 \bmod n}$ has period n, while the set of t such that $(-1)^{t-k+1} f(e, i) > 0$ has period 2. □

13. A measure of approximation

The next definition borrows some ideas of [4] Section 1.4, and Hančl [16] Definition 4.4, together with Theorem 9.1 in the present article.

Given an irrational real ξ, with Lagrange number $L(\xi)$, define $D(\xi)$ to be the supremum of the nonnegative reals d such that there are infinitely rationals p/q (p, q relatively prime) satisfying

$$\left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^2 \frac{L(\xi)}{2}\left(1 + \sqrt{1 + d/q^2}\right)}. \tag{12}$$

We call *second Lagrange number* of ξ the number $D(\xi)$. Note that if such a d exists, then there are infinitely many p/q such that $|\xi - p/q| \leq 1/L(\xi)q^2$. It is easy to verify that one cannot have equality for more than one p/q (otherwise ξ is rational), so that the inequality is strict infinitely many times, and ξ has infinitely many good rational approximations; in particular its Lagrange number must be finite. Thus, we see that $D(\xi)$ exists in $\mathbb{R}_+ \cup \infty$ if and only if ξ has infinitely many good approximations.

For a quadratic number ξ, $D(\xi)$ may be computed. We adopt the notations of Theorem 12.1.

Theorem 13.1. *Suppose that the quadratic number β has infinitely many good approximations. Then its second Lagrange number is $D(\beta) = 4m|f(e,i)|/d(f)$ and there are infinitely many approximations p/q of β satisfying the equality*

$$\left|\beta - \frac{p}{q}\right| = \frac{1}{q^2 \frac{L(\beta)}{2}\left(1 + \sqrt{1 + D(\beta)/q^2}\right)},$$

namely the good approximations of β, whereas for all other convergents p/q of sufficiently large rank,

$$\left|\beta - \frac{p}{q}\right| > \frac{1}{L(\beta)q^2} > \frac{1}{q^2 \frac{L(\beta)}{2}\left(1 + \sqrt{1 + D(\beta)/q^2}\right)}.$$

Note that the numbers $d(f), m, L(\beta)$ are not independent: they satisfy $L(\beta) = \sqrt{d(f)}/m$, see Corollary 11.1.

Proof. Suppose that $|\beta - \frac{p}{q}| \leq \frac{1}{q^2 \frac{L(\beta)}{2}(1+\sqrt{1+d/q^2})}$ for some $d > 0$; note that by Theorem 9.1 and Theorem 12.1 such a positive real d exists. Then p/q is a good approximation of β. It follows that $p/q = p_t/q_t$ is a convergent of β and if t is large enough, one has by Theorem 9.1 and Theorem 12.1

$$\begin{aligned}|\beta - p/q| &= \frac{m}{q^2 \frac{\sqrt{d(f)}}{2}\left(1 + \sqrt{1 + 4|f(e,i)|m/d(f)q^2}\right)} \\ &= \frac{1}{q^2 \frac{L(\beta)}{2}\left(1 + \sqrt{1 + 4|f(e,i)|m/d(f)q^2}\right)}.\end{aligned}$$

It follows that $q^2 \frac{L(\beta)}{2}\left(1+\sqrt{1 + 4|f(e,i)|m/d(f)q^2}\right) \geq q^2 \frac{L(\beta)}{2}\left(1+\sqrt{1 + d/q^2}\right)$ and finally that $4|f(e,i)|m/d(f) \geq d$. Since there are infinitely many p/q satisfying the equality above, the theorem is proved, taking in account Theorem 12.1 for the final assertion. □

It is well-known that two irrational numbers which lie in the same $GL_2(\mathbb{Z})$-equivalence class have the same Lagrange number. Observe that if their difference is an integer, then

they have the same second Lagrange number, as follows at once from the definition of the latter.

If β is quadratic, one has

$$D(\beta) = 4mD_0(\beta)/d(f),$$

where

$$D_0(\beta) = |f(e,i)|.$$

We study $D_0(\beta)$ in Section 14. Note that f, hence $d(f)$, and m depend only on the $GL_2(\mathbb{Z})$-class of β, and not on β. In a given class, D_0 is a measure of the approximation by rationals: the larger $D_0(\beta)$ is, the better are the approximations.

Remark. The referee suggested us to define $D(\xi)$ more generally as the supremum of all reals d (not only nonnegative ones) such that (12) holds for infinitely many rational numbers p/q (p, q relatively prime). With this new definition, $D(\xi)$ would exist for each quadratic number.

14. An invariant of quadratic numbers

In this section, we study the quantity $|f(e,i)|$, associated to a quadratic number β, which appeared in the previous section in relation with the second Lagrange number. Recall that for

$$\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}], \tag{13}$$

one has $f = f_{P(a_0,\dots,a_{n-1})}$ and $(e,i)^T$ is the first column of $P(b_0,\dots,b_{k-1})^{-1}$.

We begin by studying its invariance.

Proposition 14.1. *Let β be a quadratic number as in Theorem 9.1. If n is minimal, then $|f(e,i)|$ is independent of the chosen preperiodic part of the continued fraction representation (13) of β.*

Proof. The different representations of β are related by a chain of elementary transformations $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$, $\beta = [b_0, \dots, b_{k-1}, a_0, \overline{a_1, \dots, a_{n-1}, a_0}]$. Let $g = f_B, B = P(a_1, \dots, a_{n-1}, a_0)$, $P(b_0, \dots, b_{k-1}, a_0)^{-1} = \begin{pmatrix} e' & h' \\ i' & j' \end{pmatrix}$. It is enough to show that $f(e,i) = -g(e',i')$.

We have $P(b_0, \dots, b_{k-1}, a_0)^{-1} = P(a_0)^{-1}P(b_0, \dots, b_{k-1})^{-1}$, hence $(e',i')^T = P(a_0)^{-1}(e,i)^T$. Moreover, $B = P(a_0)^{-1}AP(a_0)$, so that by Lemma 9.1, $g = -f \cdot P(a_0)$ (or in other words $g(\mathbf{x}) = -f(P(a_0)\mathbf{x})$). Thus $g(e',i') = -fP(a_0)P(a_0)^{-1}(e,i)^T = -f(e,i)^T = f(e,i)$. □

We may therefore define $D_0(\beta) = |f(e,i)|$ for any quadratic number β. Note that if $k = 0$, $D_0(\beta) = f(1,0) = p(a_1, \dots, a_{n-1})$. Moreover, by (7), (e,i) does not depend on b_0; it follows that $D_0(\beta) = D_0(\beta + m)$ for any integer m.

We call *large class* of a quadratic number β the union of the $GL_2(\mathbb{Z})$-equivalence classes of β and of its conjugate $\bar\beta$ (the two classes may be equal, see Proposition 5.1). In other words, it is the smallest set of real numbers containing β, closed under $GL_2(\mathbb{Z})$-equivalence and quadratic conjugation. We call *discriminant* of this large class the discriminant of the quadratic form $f_{P(a_0,\dots,a_{n-1})}$, where $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$, with n minimal; by (6), this depends only on the large class, and not on the element β in the class.

Given a large class of quadratic numbers, we may call *small value* of $D_0(\beta)$, for β in this class, a value that is $< \frac{1}{2}\sqrt{d}$, where d is the discriminant of this class. The next result characterizes partially the small values in a given class. It will be a complete characterization in the case of Markoff irrationalities (Theorem 19.1).

Theorem 14.1. *Let $\alpha = [\overline{a_0, \dots, a_{n-1}}]$ be a reduced quadratic number. Let $\mathcal{C}$ be its large class and let d be the discriminant of $\mathcal{C}$.*

(i) Let $r \in \{0, \dots, n-1\}$. If β is an integral translate of one of the four numbers $[\overline{a_r, \dots, a_{n-1}, a_0, \dots, a_{r-1}}]$, $[\overline{a_r, \dots, a_0, a_{n-1}, \dots, a_{r+1}}]$ or their conjugates, then $D_0(\beta) = \epsilon_r$.

(ii) If $\beta \in \mathcal{C}$ satisfies $D_0(\beta) < \frac{1}{2}\sqrt{d}$, then for some $r = 0, \dots, n-1$, β is an integral translate of one of these four numbers.

Lemma 14.1. *Let $\alpha = [\overline{a_0, \dots, a_{n-1}}]$. Then $D_0(\bar\alpha) = D_0(\alpha)$.*

This will be proved independently for any quadratic number, see Corollary 20.3, where D_0 gets another characterization. One could also prove the lemma by using Proposition 4.1. In order to prove the theorem, we need the following technical lemma.

Lemma 14.2. *Let $\alpha = [a_0, \dots, a_{n-1}]$, $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$ and suppose that $D_0(\beta) < \frac{1}{2}\sqrt{d(f)}$, $f = f_{P(a_0,\dots,a_{n-1})}$. Then for some $r = 0, \dots, n-1$, β is an integral translate of $[\overline{a_r, \dots, a_{n-1}, a_0, \dots, a_{r-1}}]$ or of the conjugate of $[\overline{a_r, \dots, a_0, a_{n-1}, \dots, a_{r+1}}]$.*

Proof. 0. Define $c_0c_1c_2\cdots = (a_{n-1}\cdots a_0)^\infty$, so that for any $r \in \{0, 1, \dots, n-1\}$, one has $c_h = a_r$ as soon as $r + h \equiv -1 \bmod n$. We extend the a_i by periodicity, the period being n. Then the previous equality is true also for any $r \in \mathbb{N}$.

1. We begin by establishing the following claim: let $h \in \mathbb{Z}_{>0}$, define $r \in \{0, 1, \dots, n-1\}$ by $r + h \equiv -1 \bmod n$. Let $u = b_1 \cdots b_{k-1}$. If $u = c_{h-1}c_{h-2}\cdots c_0$, then β is an integral translate of $[\overline{a_r, \dots, a_{n-1}, a_0, \dots, a_{r-1}}]$; if $u = 1(c_{h-1} - 1)c_{h-2}\cdots c_0$ with $h \geq 1$, or if $u = (c_{h-2} + 1)c_{h-3}\cdots c_0$ and $c_{h-1} = 1$ with $h \geq 2$, then β is an integral translate of the conjugate of $[\overline{a_r, \dots, a_0, a_{n-1}, \dots, a_{r+1}}]$.

Indeed, suppose that $u = c_{h-1}c_{h-2}\cdots c_0$. Then $c_h u = a_r u = a_r a_{r+1}a_{r+2}\cdots a_{n-1}$, so that $a_r b_1 \cdots b_{k-1}(a_0\cdots a_{n-1})^\infty = (a_r\cdots a_{n-1}a_0\ldots a_{r-1})^\infty$ and it follows that β is an integral translate of $[\overline{a_r,\ldots,a_{n-1},a_0,\ldots,a_{r-1}}]$.

Suppose now that $u = 1(c_{h-1}-1)c_{h-2}\cdots c_0$ and $h \geq 1$. Then $b_1 = 1$ and $c_{h-1}-1 = b_2$, so that $a_{r+1} = c_{h-1} \geq 2$. It follows by Proposition 4.1 that the conjugate of $[\overline{a_r,\ldots,a_0,a_{n-1},\ldots,a_{r+1}}]$ is equal to $[-1,1,a_{r+1}-1,\overline{a_{r+2},\ldots,a_{n-1},a_0,\ldots,a_{r+1}}]$. Since $(-1)b_1\cdots b_{k-1}(a_0\cdots a_{n-1})^\infty = (-1)1(a_{r+1}-1)a_{r+2}\cdots a_{n-1}(a_0\cdots a_{n-1})^\infty = (-1)1(a_{r+1}-1)(a_{r+2}\cdots a_{n-1}a_0\ldots a_{r+1})^\infty$, it follows that β is an integral translate of the conjugate of $[\overline{a_r,\ldots,a_0,a_{n-1},\ldots,a_{r+1}}]$.

Finally, suppose that $u = (c_{h-2}+1)c_{h-3}\cdots c_0$ and $c_{h-1} = 1$ with $h \geq 2$. Then $a_{r+1} = c_{h-1} = 1$ and $b_1 = c_{h-2}+1 = a_{r+2}+1$. It follows by Proposition 4.1 that the conjugate of $[\overline{a_r,\ldots,a_0,a_{n-1},\ldots,a_{r+1}}]$ is equal to $[-1,a_{r+2}+1,\overline{a_{r+3},\ldots,a_{n-1},a_0,\ldots,a_{r+2}}]$. Since $(-1)b_1\cdots b_{k-1}(a_0\cdots a_{n-1})^\infty = (-1)(a_{r+2}+1)a_{r+3}\cdots a_{n-1}(a_0\cdots a_{n-1})^\infty = (-1)(a_{r+2}+1)(a_{r+3}\cdots a_{n-1}a_0\cdots a_{r+2})^\infty$, it follows that β is an integral translate of the conjugate of $[\overline{a_r,\ldots,a_0,a_{n-1},\ldots,a_{r+1}}]$. The claim is proved.

2. If $k = 0$, then $\beta = \alpha$. If $k = 1$, then β is an integral translate of $[a_{n-1},\overline{a_0,\ldots,a_{n-1}}] = [\overline{a_{n-1},a_0,\ldots,a_{n-2}}]$.

Hence we may assume that $k \geq 2$. Let $(e,i)^T$ be the first column of the matrix $P(b_0,\ldots,b_{k-1})^{-1}$. Then $D_0(\beta) = |f(e,i)| < \frac{1}{2}\sqrt{d(f)}$. Note that since $k \geq 2$, e,i have strictly opposite sign and thus $e/i < 0$. It follows from Corollary 6.1 that e/i is a convergent of $\bar{\alpha}$. Moreover, let j be the rank of this convergent. Then by Theorem 8.1:

- if $a_{n-1} \geq 2$ and $j \geq 1$, $|f(e,i)| = \epsilon_{-j \bmod n}$;
- if $a_{n-1} = 1$, $|f(e,i)| = \epsilon_{-j-2 \bmod n}$.

Next by Corollary 4.1:

- if $a_{n-1} \geq 2$, then $e/i = -1$ if $j = 0$, $e/i = 0$ if $j = 1$ and $e/i = -p(c_1,\ldots,c_{j-2})/p(c_0,\ldots,c_{j-2})$ if $j \geq 2$;
- if $a_{n-1} = 1$, then $e/i = -p(c_1,\ldots,c_j)/p(c_0,\ldots,c_j)$.

By definition of e,i, we also have $e/i = -p(b_1,\ldots,b_{k-2})/p(b_1,\ldots,b_{k-1})$.

3. Suppose that $a_{n-1} = 1$. Then $-p(c_1,\ldots,c_j)/p(c_0,\ldots,c_j) = -p(b_1,\ldots,b_{k-2})/p(b_1,\ldots,b_{k-1})$. Thus, using Lemma 3.1 (ii), $p(b_1,\ldots,b_{k-1}) = p(c_j,\ldots,c_0)$ and $p(b_1,\ldots,b_{k-2}) = p(c_j,\ldots,c_1)$. Note that $k-1$, that is, the number of argument in $p(b_1,\ldots,b_{k-1})$, is ≥ 1; also, $j+1$, that is, the number of arguments in $p(c_j,\ldots,c_0)$, is ≥ 1. We may therefore apply Lemma 3.2 and we have three cases, according to the relative values of $k-1$ and $j+1$, which by this lemma must differ by at most 1.

3.1 If $k-1 = j+1$, then $b_1 = c_j, b_2 = c_{j-1},\ldots,b_{k-1} = c_0$. Then $b_1\cdots b_{k-1} = c_j\cdots c_0$ and we apply the claim in 1.

3.2. If $k-1=j+1+1$, then the lemma (with the u's being replaced by the c's and the v's by the b's) implies that $b_1=1, c_j=b_2+1, c_{j-1}=b_3,\dots,c_0=b_{k-1}$. Thus $b_1\cdots b_{k-1}=1(c_j-1)c_{j-1}\cdots c_0, j+1\geq 1$. We then apply the claim.

3.3 If $j+1=k-1+1$, then the lemma (with the u's being replaced by the b's and the v's by the c's) implies that $c_j=1, b_1=c_{j-1}+1, b_2=c_{j-2},\dots,b_{k-1}=c_0$. Then $b_1\cdots b_{k-1}=(c_{j-1}+1)c_{j-2}\cdots c_0, c_j=1, j+1=k\geq 2$ and we apply the claim.

4. Suppose that $a_{n-1}\geq 2$. If $j=0$, then $e/i=-1$, which contradicts Lemma 8.1. If $j=1$, then $e/i=0$, and e,i have not strictly opposite sign, a contradiction.

We thus may assume that $j\geq 2$. Then $p(b_1,\dots,b_{k-1})=p(c_{j-2},\dots,c_0)$ and $p(b_1,\dots,b_{k-2})=p(c_{j-2},\dots,c_1)$. Since $k-1, j-1\geq 1$, we may apply Lemma 3.2; we have three cases, according to the relative values of $k-1$ and $j-1$, whose difference must be ≥ 1.

4.1. If $k-1=j-1$, then by the lemma $b_1=c_{j-2}, b_2=c_{j-3},\dots,b_{k-1}=c_0$. Then $b_1\cdots b_{k-1}=c_{j-2}\cdots c_0$ and we apply the claim.

4.2. If $k-1=j-1+1$, then $b_1=1, c_{j-2}=b_2+1, c_{j-3}=b_3,\dots,c_0=b_{k-1}$. Then $b_1\cdots b_{k-1}=1(c_{j-2}-1)c_{j-3}\cdots c_0$ and $j-1\geq 1$. We apply the claim.

4.3. If $j-1=k-1+1$, then $c_{j-2}=1, b_1=c_{j-3}+1, b_2=c_{j-4},\dots,b_{k-1}=c_0$. Then $b_1\cdots b_{k-1}=(c_{j-3}+1)c_{j-4}\cdots c_0$, $c_{j-2}=1$ and $j-1=k\geq 2$. We apply the claim. □

Proof of Theorem 14.1. (i) Suppose first that β is an integral translate of $[a_r,\dots,a_{n-1}, a_0,\dots,a_{r-1}]$ or $[a_r,\dots,a_0,a_{n-1},\dots,a_{r+1}]$. We may assume that $r=0$, since the result is invariant under cyclic permutation of $a_0\cdots a_{n-1}$. If $\beta=\alpha$, then $D_0(\beta)$ is equal to ϵ_0, by a previous observation. If $\beta=[a_0,a_{n-1},\dots,a_1]$, then $D_0(\beta)=f(1,0)$ where $f=f_{P(a_0,a_{n-1},\dots,a_1)}$; hence $f(1,0)=p(a_{n-1},\dots,a_1)=p(a_1,\dots,a_{n-1})=\epsilon_0$ and we are done.

Lemma 14.1 implies the two remaining cases.

(ii) We apply Lemma 14.2, since $\beta=[b_0,\dots,b_{k-1},\overline{a_0,\dots,a_{n-1}}]$ or $=[b_0,\dots,b_{k-1}, \overline{a_{n-1},\dots,a_0}]$. □

Part 3. Markoff theory

In this part, we apply the previous general results on quadratic numbers and quadratic forms to the special case of Markoff irrationalities and Markoff forms.

15. Markoff forms

Corollary 8.1 allows us to find all representations of all proper small values of an indefinite binary integral quadratic form, in particular of its minimum. However the example in Section 8 shows that the numbers ϵ_t are not all small.

For Markoff forms, one may be more precise. We begin by introducing them. We adopt here the notations of [24], where the reader will find more details.

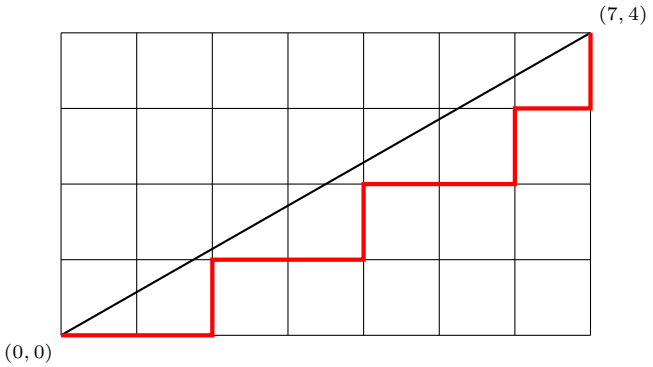

Fig. 1. The lower Christoffel words *aabaabaabab* of slope 4/7.

The *lower Christoffel word* of slope $\frac{m}{l}$ is the word that discretizes from below a segment from $[0,0]$ to $[l,m]$ in the plane, as shown in Fig. 1; here l, m are two relatively prime nonnegative integers. The *upper Christoffel word* is obtained by discretizing from above. For each given slope, there is a lower Christoffel word w and an upper Christoffel word, which turns out to be the reversal $\widetilde{w}$.

The Christoffel words a and b have slope 0 or ∞ respectively. The other Christoffel words are called *proper*.

Each proper lower Christoffel word is of the form $a\pi b$, for some palindrome π. The corresponding palindromes are called *central words* and have been extensively studied by Aldo de Luca (see [24] for results and references).

Each proper Christoffel word (lower or upper) w has a unique factorization $w = uv$, where u, v are Christoffel words (a result due to Borel and Laubie [3]): this is called its *standard factorization*. It has also a unique factorization $w = v'u'$ where u', v' are palindromes (Chuan [8]): this is called its *palindromic factorization*.

One has a precise result on the lengths of these factors: $l = |w|_a$ (the number of a's in w) and $m = |w|_b$. Moreover

$$|u| = |u'| = m^*, |v| = |v'| = l^*,$$

where x^* denotes the inverse in $\{1, \dots, l+m-1\}$ of x modulo $l+m = |w|$ (see [24] Theorem 12.1.8 and Corollary 14.1.5). Note that $l^* + m^* = l + m$ and that l, m, l^*, m^* are each relatively prime to $l + m$.

For example, $aabaabaabab = aab.aabaabab$ (standard factorization) $= aabaabaa.bab$ (palindromic factorization); moreover, the length of w is 11, $l = 7, m = 4$, and the inverses of these numbers modulo 11 are respectively $l^* = 8$ and $m^* = 3$. The central palindrome here is $\pi = abaabaaba$.

Let μ be the monoid homomorphism from the free monoid $\{a, b\}^*$ into the group $SL_2(\mathbb{Z})$ defined by

$$\mu(a) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \mu(b) = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}. \tag{14}$$

Let $w \in \{a,b\}^*$ be a lower Christoffel word. Let $\mu(w) = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$. Then $b = \frac{1}{3}(p+s)$ is a *Markoff number* m (see [10–12,2,1,24]), that we call *associated with* w. The associated *Markoff form* is then

$$f_w = f_{\mu(w)^T} = f_{\mu(\widetilde{w})} = qx^2 + (s-p)xy - ry^2. \tag{15}$$

Let χ be the monoid homomorphism $\{a,b\}^* \to \{1,2\}^*$ sending a onto 11 and b onto 22. Hence we have also

$$f_w(x,y) = p(a_1,\dots,a_{n-1})x^2 + (p(a_1,\dots,a_{n-2}) - p(a_0,\dots,a_{n-1}))xy \\ -p(a_0,\dots,a_{n-2})y^2, \tag{16}$$

with $\chi(\widetilde{w}) = a_0 \cdots a_{n-1}$, since $\mu(a) = P(1)^2, \mu(b) = P(2)^2$, and therefore $\mu(w)^T = P(\chi(\widetilde{w}))$. It is well-known, and easy to verify, that $d(f) = 9m^2 - 4$.

We denote by x_w the number whose expansion into continued fraction is $[\chi(\widetilde{w})^\infty]$, the associated *Markoff irrationality*. More generally, we call Markoff irrationality any number $GL_2(\mathbb{Z})$-equivalent to x_w for some lower Christoffel word w.

Recall that the minimum of f_w is m (see e.g. [24] Theorem 9.3.1).

16. A lexicographical result

We need this result in the sequel. The free monoid $\{a,b\}^*$ is lexicographically ordered (reading the words from left to right), with $a < b$.

Theorem 16.1. *Let w be a lower Christoffel word and let $\mathcal{C}$ be its conjugation class, lexicographically ordered. The mapping $\mathcal{C} \to \mathbb{N}$, $u \mapsto \mu(u)_{12}$, is strictly increasing. The smallest value of this function is the Markoff number m associated with w. The largest value c satisfies $2c < 3m$ and $c < \frac{1}{2}\sqrt{9m^2-4}$.*

The result that this mapping is increasing is somewhat implicit in the theory of Markoff, although not really stated as here. However, see Lemma 17 and Theorem 27 in Bombieri's article [2]; one may also prove this result by following his methods.[4] We have preferred a finitary way, which will allow us to prove also the inequalities at the end of the statement.

[4] One associates with each conjugate v of w a corresponding number $a_k + [a_k, a_{k+1}, \dots]^{-1} + [a_{k-1}, a_{k-2}, \dots]^{-1}$, where $(a_n)_{n\in\mathbb{Z}}$ is the bi-infinite word obtained by bi-infinite repetition of $w = a_0 \cdots a_{n-1}$ and $v = a_k \cdots a_{k+n-1}$. Then it is shown that this number is an increasing function of $\mu(v)_{12}$.

a	a	a	b	a	a	b	1325
a	a	b	a	a	a	b	1327
a	a	b	a	a	b	a	1715
a	b	a	a	a	b	a	1735
a	b	a	a	b	a	a	1793
b	a	a	a	b	a	a	1939
b	a	a	b	a	a	a	1949

Fig. 2. Conjugates of a Christoffel word and numbers $\mu(v)_{12}$.

The following result gives some information on the *Burrows-Wheeler tableau* of w (which is the tableau obtained by writing one above the other the conjugates of w in lexicographical order, as in Fig. 2); see [22] or [24] Section 15.2.

Lemma 16.1. *Let $w = a\pi b$ be a proper lower Christoffel word. Let $w = w_1 < w_2 < \ldots < w_{|w|} = \widetilde{w}$ be the list of its conjugates in lexicographical order. Then each pair (w_i, w_{i+1}), $i = 0, \ldots, |w|-1$, is equal to exactly one pair $(sabp, sbap)$, for some factorization $\pi = ps$.*

Note that there are $|w| - 1$ such factorizations.

Proof. See [24], Theorem 15.2.4 and its proof. □

Let $(w_i, w_{i+1}) = (sabp, sbap)$ as in the lemma. We say that this pair is of type 1 if $|s| \geq |p|$ and of type 2 if $|p| \geq |s|$. In type 1, we may write $\pi = xv\widetilde{x}$, $s = v\widetilde{x}, p = x$; in type 2, we have $\pi = xv\widetilde{x}$, $s = \widetilde{x}$, $p = xv$. Note that v is a central factor of π, hence a palindrome. Note also that if p, s have the same length (which happens if and only if w is of even length), then v is the empty word and the pair is of both types.

For example, we have $w_1 = w = a^3ba^2b$, $\pi = aabaa$, $w_3 = a\mathbf{ab}aaba$, $w_4 = a\mathbf{ba}aaba$, the type is 2, $v = aba$, $x = a$, see Fig. 2, where the numbers $\mu(w_i)_{12}$ are also written; in particular $\mu(w_1)_{12} = 1325$ is a Markoff number.

Lemma 16.2. *With this notation, one has $\mu(w_{i+1})_{12} - \mu(w_i)_{12} = 2\mu(v)_{jj}$, where j is the type of the pair.*

Proof. Suppose that the type is 1. We then have $w_i = v\widetilde{x}abx$, $w_{i+1} = v\widetilde{x}bax$. The matrix $\mu(v)$ is symmetric and $\mu(x)$ is the transpose of $\mu(\widetilde{x})$ (since the matrices μa and μb are symmetric). Let $\mu(v) = \begin{pmatrix} p & q \\ q & s \end{pmatrix}$ and $\mu(x) = \begin{pmatrix} i & j \\ k & l \end{pmatrix}$. Note that $\mu(ba) - \mu(ab) = \begin{pmatrix} 12 & 7 \\ 5 & 3 \end{pmatrix} - \begin{pmatrix} 12 & 5 \\ 7 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$. We have $\mu(w_{i+1}) - \mu(w_i) = \mu(v)\mu(\widetilde{x})\mu(ba)\mu(x) - \mu(v)\mu(\widetilde{x})\mu(ab)\mu(x) = \mu(v)\mu(\widetilde{x})(\mu(ba) - \mu(ab))\mu(x))$. This is equal to $\begin{pmatrix} p & q \\ q & s \end{pmatrix}\begin{pmatrix} i & k \\ j & l \end{pmatrix}\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} i & j \\ k & l \end{pmatrix} = \begin{pmatrix} p & q \\ q & s \end{pmatrix}\begin{pmatrix} 0 & 2(il-jk) \\ 2(jk-il) & 0 \end{pmatrix} = \begin{pmatrix} p & q \\ q & s \end{pmatrix}\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} = 2\begin{pmatrix} -q & p \\ -s & q \end{pmatrix}$. The 1,2-entry of this matrix is $2p = 2\mu(v)_{11}$.

The proof in type 2 is similar. □

Recall the conventions on continuant polynomials, see Section 3.

Lemma 16.3. *Let w be palindrome of even length over $\mathbb{Z}_{>0}$, and $a \in \mathbb{Z}_{>0}$. Then*

$$p(aw) = \sum_f xp(f),$$

where the sum is over all central factors f of w, and x is the letter before f in aw.

Proof. If w is empty, then the formula follows from the equality $p() = 1, p(a) = a$. If w is nonempty, we may write $w = bu\widetilde{u}b$, $b \in \mathbb{Z}_{>0}$. Then by Lemma 3.1 (i) and (ii), $p(aw) = p(abu\widetilde{u}b) = ap(w) + p(u\widetilde{u}b) = ap(w) + p(bu\widetilde{u})$. Then the formula follows by induction. □

Corollary 16.1. *Let $n \geq 2$ and $a_1, \ldots, a_n \in \mathbb{Z}_{>0}$. Denote by w_i the word $a_i \cdots a_n a_n \cdots a_i$ on $\mathbb{Z}_{>0}$, with $w_{n+1} =$ the empty word. Then*

$$p(w_1) = a_1^2 p(w_2) + (2a_1a_2 + 1)p(w_3) + 2a_1 \sum_{4 \leq i \leq n+1} a_{i-1}p(w_i).$$

Proof. We have by Lemma 3.1 (i) and (ii) $p(w_1) = a_1p(a_1 \ldots a_n a_n \ldots a_2) + p(a_1 \ldots a_n a_n \ldots a_3) = a_1^2 p(a_2 \ldots a_n a_n \ldots a_2) + a_1p(a_3 \ldots a_n a_n \ldots a_2) + a_1p(a_2 \ldots a_n a_n \ldots a_3) + p(a_3 \ldots a_n a_n \ldots a_3) = a_1^2 p(a_2 \ldots a_n a_n \ldots a_2) + 2a_1p(a_2 \ldots a_n a_n \ldots a_3) + p(a_3 \ldots a_n a_n \ldots a_3)$. We apply the lemma to $w = w_3$, $a = a_2$, and obtain $p(a_2 \ldots a_n a_n \ldots a_3) = \sum_{3 \leq i \leq n+1} a_{i-1}p(w_i)$. The formula follows. □

Corollary 16.2. *Let x be a palindrome of even length over $\mathbb{Z}_{>0}$. Then $p(2x2) > 4\sum_f p(f)$ where the sum is over all central factors of x.*

Proof. If x is empty, then the inequality is $p(2, 2) > 4p()$, that is $5 > 4$. Assume now that x is nonempty. We take $2x2 = a_1 \ldots a_n a_n \ldots a_1$, in particular $a_1 = 2$. Then in the formula of the previous corollary, all coefficients are ≥ 4, while the second is > 4. This proves the result. □

Proof of Theorem 16.1. The fact that this mapping is strictly increasing follows from Lemma 16.2, since the diagonal elements of $\mu(v)$ are positive. It shows also that the smallest value is attained for w_1, which is equal to w. Indeed, w is the smallest word in its conjugation class, see for example [24] Theorem 6.2.2. The smallest value is the Markoff number $m = \mu(w)_{12}$ associated to w.

It follows from Lemma 16.1 and Lemma 16.2 that the difference $c - m$ is equal to twice the sum of all $\mu(v)_{11}$ and all $\mu(v)_{22}$, v a central factor of π, except if v is the empty word in which case one takes only one of the two terms. Now, for v nonempty,

$\chi(v)$ is of even length ≥ 2 and since $\mu(v) = P(\chi(v))$, $\mu(v)_{ii}$ is by (2) equal to $p(\chi(v))$ if $i = 1$ and to $p(t)$ if $i = 2$, where t is obtained from $\chi(v)$ by removing the first and last letters. Thus the previous sum is equal to the sum of all $p(t)$, t a central factor of $\chi(\pi)$. In Corollary 16.2, let $x = \chi(\pi)$. Since $p(2x2) = m$ ([24] Theorem 10.3.5), we obtain that $2(c - m) < m$. Thus $2c < 3m$.

It follows that $2c \leq 3m-1$, $4c^2 \leq 9m^2-6m+1$; the latter is $< 9m^2-4$, since $5 < 6m$. Thus $c < \frac{1}{2}\sqrt{9m^2-4}$. □

Corollary 16.3. *Let w be a lower Christoffel word, with $|w|_a = l$ and $|w|_b = m$, and $\chi(\widetilde{w}) = a_0 \cdots a_{n-1}$, so that $n = 2|w|$. For $r = 0, \ldots, n-1$, let ϵ_r be defined by (8). Then the numbers $\mu(v)_{12}$, v a conjugate of w, coincide with the numbers ϵ_r. Moreover, let $m_1 < \ldots < m_{n/2}$ be the $|w|$ numbers $\mu(v)_{12}$. Then for $j = 1, \ldots, n/2$, $\epsilon_r = m_j$ exactly when $r = 2(j-1)l^* \bmod n$ or $r = 2jm^* - 1 \bmod n$.*

Proof. Note that the n numbers r at the end of the statement are all distinct modulo n, since l^*, m^* are additive generators modulo $|w| = n/2$. So, in view of the theorem, it is enough to show that when r is as indicated, then $\epsilon_r = m_j$.

Take the notations of Lemma 16.1. By this lemma, $\widetilde{w} = w_{|w|}$. We use the fact that $w_j = C^{(j-1)m^*}(w)$ (see [24] Theorem 15.2.4), where C is the *conjugator*, sending each word on the word obtained by removing its first letter and putting it at the end of this word. Thus, $w_{|w|} = C^{(|w|-1)m^*}(w) = C^{-m^*}(w)$, since C^N is the identity mapping, when applied to words of length N.

By (2) and the equality $\mu = P \circ \chi$, we have $\mu(v)_{12} = p(\chi(v)^-)$, where v^- denotes the word obtained from v by removing its last letter; likewise we denote ${}^-v$ the word obtained by removing the first letter; note that ${}^-v = (Cv)^-$. Hence, by Theorem 16.1, m_j is equal to $\mu(w_j)_{12} = p(\chi(w_j)^-) = p(\chi(C^{(j-1)m^*}(w))^-)$. Since $w = C^{m^*}(\widetilde{w})$ and since $\chi \circ C = C^2 \circ \chi$, this is equal to $p(\chi(C^{jm^*}(\widetilde{w}))^-) = p(C^{2jm^*}(\chi(\widetilde{w}))^-)$. The word $C^{2jm^*}(\chi(\widetilde{w}))^-$ is equal to the word ${}^-C^{2jm^*-1}(\chi(\widetilde{w}))$, thus $m_j = \epsilon_{2jm^*-1 \bmod n}$ (because $\epsilon_r = p({}^-C^r(a_0 \cdots a_{n-1}))$, see the observation after Theorem 7.1).

Moreover, letting $t = \chi(C^{(j-1)m^*}(w))$, we have $p(t^-) = p({}^-\widetilde{t})$. Next, let $t_1 = C^{(j-1)m^*}(w)$; then $t = \chi(t_1)$, $\widetilde{t} = \chi(\widetilde{t_1})$. Since for any word x, the reversal of Cx is $C^{-1}(\widetilde{x})$, and more generally the reversal of $C^k(x)$ is $C^{-k}(\widetilde{x})$, one has $\widetilde{t_1} = C^{-(j-1)m^*}(\widetilde{w}) = C^{(j-1)l^*}(\widetilde{w})$ (because $l^* + m^* = |\widetilde{w}|$), and we see that $m_j = p(t^-) = p({}^-\widetilde{t}) = p({}^-\chi(\widetilde{t_1})) = p({}^-\chi(C^{(j-1)l^*}(\widetilde{w})) = p({}^-C^{2(j-1)l^*}(\chi(\widetilde{w}))) = \epsilon_{2(j-1)l^* \bmod n}$. □

17. Small values of Markoff forms

Let w be a *proper* lower Christoffel word of length $n/2$, $x = x_w$ the associated Markoff irrationality and $f = f_w$ be the associated Markoff form. As before, denote by $m_1 < \ldots < m_{n/2}$ the numbers $\mu(v)_{12}$, v a conjugate of w, which by Corollary 16.3 coincide with the numbers ϵ_r. In particular $m_1 = m$, the Markoff number associated with w.

Theorem 17.1. *There are n small values of f_w, which are the numbers $\pm m_j$. For any pair (p,q) of integers, one has $f(p,q) = m_j$ if and only if p,q are relatively prime and if either $(p,q) = (\pm 1, 0)$ and $j = 1$, or if p/q is a convergent of x of rank congruent to $2(j-1)l^* - 1$ or a convergent of $\bar{x}$ of rank $-2 - 2(j-1)l^*$ modulo n. Moreover, $f(p,q) = -m_j$ if and only if either $(p,q) = (0, \pm 1)$ and $j = n/2$, or p/q is a convergent of x of rank congruent to $2jm^* - 2$ or a convergent of $\bar{x}$ of rank $-2jm^* - 1$ modulo n.*

The fact that opposite values of the Markoff quadratic form are attained is reminiscent of the following result, proved by Frobenius: each Markoff form is $SL_2(\mathbb{Z})$-equivalent to its opposite, see [15] p.458, [13] Theorem 2 (B) p. 20.

Proof. If $f(p,q)$ is small, then p,q are relatively prime. Indeed, otherwise $(p,q) = d(p',q')$, $d \geq 2$. Then $|f(p,q)| = d^2|f(p',q')| \geq 4m > \frac{1}{2}d(f) = \frac{1}{2}\sqrt{9m^2-4}$, since $64m^2 > 9m^2 - 4$.

If $(p,q) = (\pm 1, 0)$, then $f(p,q) = m$; and if $(p,q) = (0, \pm 1)$, then $f(p,q) = -\epsilon_{n-1} = -\mu(\widetilde{w})_{12} = -m_{n/2}$, as follows from (16).

Let $\chi(\widetilde{w}) = a_0 \cdots a_{n-1}$. Suppose now that $f(p,q)$ is a small value and that $p,q \neq 0$. Then by Corollary 6.1, p/q is a convergent of x or of $\bar{x}$. Since $w \neq b$, $\widetilde{w}$ ends by a, $\chi(\widetilde{w})$ by 1, that is $a_{n-1} = 1$.

Suppose that p/q is a convergent of rank i of x. By Theorem 8.1 (i), $|f(p,q)| = \epsilon_{i+1 \bmod n}$; by Corollary 16.3, this is m_j exactly when $i+1 \equiv 2(j-1)l^*$ or $2jm^* - 1 \bmod n$, that is $i \equiv 2(j-1)l^* - 1$ or $2jm^* - 2$. Moreover, the sign of $f(p,q)$ is that of $(-1)^{i+1}$ by Theorem 8.1 (i).

We assume now that p/q is a convergent of $\bar{x}$. Then by Theorem 8.1 (ii), case $a_{n-1} = 1$, we have $|f(p,q)| = \epsilon_{-i-2 \bmod n}$. By the corollary, this is m_j exactly when $-i-2 \equiv 2(j-1)l^*$ or $2jm^* - 1 \bmod n$, that is $i \equiv -2 - 2(j-1)l^*$ or $-2jm^* - 1$. Moreover, the sign of $f(p,q)$ is that of $(-1)^i$, by Theorem 8.1 (ii), case $a_{n-1} = 1$.

To conclude for the sign, note that $n = 2|w|$ is even (so that, for example, $i+1 \equiv 2(j-1)l^*$ implies that $i+1$ is even, and $(-1)^{i+1} = 1$). Finally, note that the m_j are all $< \frac{1}{2}d(f_w)$ by Corollary 16.3 and Theorem 16.1. □

Corollary 17.1. *The Markoff form $f_w(p,q)$ attains its minimum in absolute value exactly when $(p,q) = (\pm 1, 0)$ or if p/q a convergent of rank congruent to -1 or $2m^* - 2$ modulo n of x_w, or a convergent of rank congruent to -2 or $-2m^* - 1$ modulo n of $\bar{x}_w$.*

Corollary 17.2. *There is a natural bijection between the absolute values of the small values of the Markoff form f_w and the conjugates of the Christoffel word w.*

This bijection is illustrated in Fig. 2.

When w is not proper, that is, $w = a$ or b, we have: $f_a = x^2 - xy - y^2$, $f_b = 2(x^2 - 2xy - y^2)$ and this is settled by the following result.

Proposition 17.1. *Let $f(x,y) = x^2 - kxy - y^2$, k positive integer. Let $x = \frac{k+\sqrt{k^2+4}}{2}$. The minimum of f is 1. Let p, q be relatively prime integers. Then $f(p,q) = 1$ if and only if $(p,q) = (\pm 1, 0)$, or if p/q is a convergent of x of odd rank or a convergent of $\bar{x}$ of even rank, excluding rank 0 if $k \geq 2$. Moreover $f(p,q) = -1$ if and only if $(p,q) = (0, \pm 1)$, or if p/q is a convergent of x of even rank or a convergent of $\bar{x}$ of odd rank.*

Note that $d(f) = k^2 + 4$, so that for $k = 1, 2$, one has $\frac{1}{2}\sqrt{d(f)} < 2$ and therefore the small values of f_a and $\frac{1}{2}f_b$ are ± 1.

Proof. We leave the proof to the reader, who may use directly Serret's theorem and Theorem 8.1. Note that $x = [\bar{k}]$, and that by Proposition 4.1, one has $\bar{x} = \frac{k-\sqrt{k^2+4}}{2} = [-1, 2, \bar{1}]$ if $k = 1$ and $x = [-1, 1, k-1, \bar{k}]$ if $k \geq 2$. □

18. Good approximations of Markoff irrationalities

The theory of Markoff for approximations (see for example [1,2,24]) classifies the numbers β whose Lagrange constant is smaller than 3. These numbers are all quadratic, and are exactly the numbers which are $GL_2(\mathbb{Z})$-equivalent to the number x_w, for some lower Christoffel word w (see Section 15 for this notation). It follows that $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \cdots, a_{n-1}}]$ with $a_0 \cdots a_{n-1} = \chi(\widetilde{w})$.

It is then shown that, for a number β with $L(\beta) < 3$, there are infinitely many convergents p/q of β such that

$$|\beta - p/q| < 1/L(\beta)q^2,$$

that is, p/q is a good convergent of β.

The theorem below gives another approach to the previous result, and strengthens, for Markoff irrationalities, the previous Theorem 12.1. Recall that w is a lower Christoffel word of slope m/l and m^* is the inverse of m modulo $m + l = |w|$.

Theorem 18.1. *Let $L(\beta) < 3$ and p_t/q_t be the convergents of β. Let $f = f_w$. Let $\sigma = \pm 1$ be such that $\sigma(-1)^{k-1} f(e,i) > 0$ (e, i are as in Theorem 9.1). There exists a unique $t_0 \in \{0, \dots, n-1\}$ such that $t_0 \equiv k-1$ or $2m^* + k - 2 \bmod n$ and $(-1)^{t_0} = \sigma$. Then, for t large enough, p_t/q_t is a good convergent if and only if $t \equiv t_0 \bmod n$.*

This result is a variant of a theorem of Bombieri ([2] Theorem 29; see also [5] Theorem 1.5 and [24] Theorem 8.3.3).

Proof. The equality $(-1)^{t_0} = \sigma$ determines uniquely the parity of t_0. The numbers $k-1$ and $2m^* + k - 2$ have opposite parity. Since n is even, this proves uniqueness of t_0. For existence, take the parity determined by the above equality and choose $t_0 \equiv k-1 \bmod n$ or $2m^* + k - 2 \bmod n$ depending on this parity.

mauvaise notation !

Suppose that $t \equiv t_0 \bmod n$. Since n is even, t and t_0 have the same parity, hence $(-1)^t = \sigma$. We have $t \equiv k-1$ or $2m^* + k - 2 \bmod n$, thus $t - k + 1 \equiv 0$ or $t - k + 1 \equiv 2m^* - 1 \bmod n$. Then by Corollary 16.3, $\epsilon_{t-k+1} = m$, the Markoff number associated with w, and the minimum of the numbers ϵ_r. Since $(-1)^t = \sigma$, we have $(-1)^{t-k+1} f(e,i) > 0$. Then t is good for t large enough, by Theorem 12.1. m_1

Conversely, if t is good, by Theorem 12.1 we must have $t - k + 1 \equiv r \bmod n$ with $\epsilon_r = m$ and $(-1)^{t-k+1} f(e,i) > 0$. Thus $r \equiv 0$ or $r \equiv 2m^* - 1 \bmod n$ by Corollary 16.3, and therefore $t \equiv k-1$ or $2m^* + k - 2$ and $\sigma = (-1)^t$. Thus $t \equiv t_0 \bmod n$. □

19. Refinement of the Markoff classification

By Markoff's theory, if a real β is such that $L(\beta) < 3$, then β is $GL_2(\mathbb{Z})$-equivalent to x_w, for some lower Christoffel word w. In this case, $L(\beta) = \sqrt{9 - 4/m^2}$, where m is the Markoff number associated with w. For distinct lower Christoffel words, their $GL_2(\mathbb{Z})$-equivalence classes are distinct.[5]

Theorem 19.1. *Let w be a proper lower Christoffel word of length $n/2$, let $\mathcal{C}$ be the $GL_2(\mathbb{Z})$-equivalence class of numbers containing $x_w = [\overline{a_0, \dots, a_{n-1}}]$, with $\chi(\tilde{w}) = a_0 \cdots a_{n-1}$. The $n/2 = |w|$ smallest values $D_0(\beta)$, $\beta \in \mathcal{C}$, are the numbers $\mu(v)_{12}$, where v are conjugates of w. The numbers $\beta \in \mathcal{C}$ such that $D_0(\beta) = \mu(v)_{12}$ are the integral translates of one of the two numbers $[\chi(\widetilde{v})^\infty]$, $[(C^{-1}(\chi(v))^\infty]$, or of their conjugates.*

Recall that the mapping C, called the conjugator, was defined in the proof of Corollary 16.3.

The similar results for improper Christoffel words $w = a$ or $w = b$ are due to Hančl [16,17]: if β is $GL_2(\mathbb{Z})$-equivalent to the *golden ratio* $\frac{\sqrt{5}+1}{2}$ (resp. the *silver ratio* $\sqrt{2}+1$), then $D_0(\beta) = 1$ if and only if β is an integral translate of $\frac{\sqrt{5}+1}{2}$ (resp. $\sqrt{2}+1$) or of its conjugate. One may apply (10) directly to prove this.

Proof. Note first that, since w, hence $\chi(w)$, is a product of two palindromes, x_w is equivalent to its conjugate (Proposition 5.1). Hence the large class of x_w is equal to its equivalence class.

Recall that if β is in the $GL_2(\mathbb{Z})$-equivalence class of x_w, then $D_0(\beta) = |f_w(e,i)|$ is the absolute value of some nonzero value of the quadratic form f_w (see the beginning of Section 14). Since by Theorem 17.1, the $n/2$ smallest values of f_w are the numbers $\mu(v)_{12}$, v a conjugate of w, it is enough to prove the last assertion.

Let v be a conjugate of w. Then $v = w_j$ for some $j = 0, \dots, n/2 - 1$, with the notations of Lemma 16.1, and $w_j = C^{(j-1)m^*}(w)$, $\widetilde{w} = C^{l^*}(w)$, $m^* + l^* = |w|$ (proof of Corollary 16.3). Note that the reversal of $C^k(x)$ is equal to $C^{-k}(\widetilde{x})$. Hence the rever-

[5] It is however not known if for distinct Christoffel words, the Lagrange constants are distinct: this is equivalent to the *Frobenius conjecture*, or *Markoff numbers injectivity conjecture*, see [1].

sal of $v = C^{(j-1)m^*}(w)$ is equal to $\widetilde{v} = C^{(j-1)l^*}(\widetilde{w})$; thus $\chi(\widetilde{v}) = C^{2(j-1)l^*}(\chi(\widetilde{w})) = a_r \cdots a_{n-1}a_0 \cdots a_{r-1}$, with $r = 2(j-1)l^* \bmod n$. It follows that

$$[\chi(\widetilde{v})^\infty] = [\overline{a_r, \ldots, a_{n-1}, a_0, \ldots, a_{r-1}}].$$

Thus by Theorem 14.1 (i), $D_0([\chi(\widetilde{v})^\infty]) = \epsilon_r$, which by Theorem 16.1 and Corollary 16.3 is equal to $m_j = \mu(v)_{12}$.

Note that $w = C^{m^*}(\widetilde{w})$, hence $\chi(w) = C^{2m^*}(\chi(\widetilde{w}))$. We have also $v = C^{(j-1)m^*}(w)$, hence $\chi(v) = C^{2(j-1)m^*}(\chi(w)) = C^{2jm^*}(\chi(\widetilde{w}))$. It follows, with now $r = 2jm^* - 1 \bmod n$, that $C^{-1}(\chi(v)) = C^r(\chi(\widetilde{w})) = a_r \cdots a_{n-1}a_0 \cdots a_{r-1}$. Thus

$$[C^{-1}(\chi(v))^\infty] = [\overline{a_r, \ldots, a_{n-1}, a_0, \ldots, a_{r-1}}].$$

Thus by Theorem 14.1 (i), $D_0([C^{-1}(\chi(v))^\infty]) = \epsilon_r$, which by Corollary 16.3 is equal to $m_j = \mu(v)_{12}$.

It follows from Lemma 14.1 and the fact that D_0 is invariant under integer translation, that for each β which is an integral translate of one of the two numbers $[\chi(\widetilde{v})^\infty]$, $[(C^{-1}(\chi(v))^\infty]$, or of their conjugates, $D_0(\beta) = \mu(v)_{12}$.

It remains to prove the converse. If $D_0(\beta) = \mu(v)_{12}$ for some conjugate v of w, then $D_0(\beta) < \frac{1}{2}\sqrt{d(f_w)}$ by Theorem 17.1 and Corollary 16.3. By Theorem 14.1 (ii), β is an integral translate of one of the four numbers $[\overline{a_r, \ldots, a_{n-1}, a_0, \ldots, a_{r-1}}]$, $[\overline{a_r, \ldots, a_0, a_{n-1}, \ldots, a_{r+1}}]$ or their conjugates, for some $r \in \{0, \ldots, n-1\}$. By Theorem 14.1 (i) and Corollary 16.3, we have $r \equiv 2(j-1)l^* \bmod n$ or $r \equiv 2jm^* - 1 \bmod n$, where j is defined by the equality $v = w_j$.

Thus it is enough to show that for these two values of r, the numbers $[\overline{a_r, \ldots, a_{n-1}, a_0, \ldots, a_{r-1}}]$, $[\overline{a_r, \ldots, a_0, a_{n-1}, \ldots, a_{r+1}}]$ are each equal to one of the two numbers $[\chi(\widetilde{v})^\infty]$, $[(C^{-1}(\chi(v))^\infty]$.

Observe first that $a_r \cdots a_{n-1}a_0 \cdots a_{r-1} = C^r(\chi(\widetilde{w}))$ and that $a_r \cdots a_0 a_{n-1} \cdots a_{r+1} = C^{-r-1}(\chi(w))$.

Let $r_1 = 2(j-1)l^*$ and $r_2 = 2jm^* - 1$. We have $C^{r_1}(\chi(\widetilde{w})) = C^{-r_2-1}(\chi(w))$ and $C^{r_2}(\chi(\widetilde{w})) = C^{-r_1-1}(\chi(w))$: both equalities follow from $\chi(\widetilde{w}) = C^{-r_1-r_2-1}(\chi(w))$ which is a consequence of $-r_1 - r_2 - 1 = -2j(l^* + m^*) + 2l^* \equiv 2l^* \bmod n$ and the fact that $\widetilde{w} = C^{l^*}(w)$.

Thus we conclude using the two previous displayed equations. □

Corollary 19.1. *The smallest value of $D_0(\beta)$ is the Markoff number associated to w, and it is attained when β is an integral translate of $[\chi(\widetilde{w})^\infty]$, $[(C^{-1}(\chi(w))^\infty]$, or of their conjugates.*

Let us see the example $w = ab$. Then the smallest value of $D_0(\beta)$ is the corresponding Markoff number 5, and it is attained when β is an integral translate of $[\overline{2211}]$ or $[\overline{2112}]$, or of their conjugates. The second smallest value of $D_0(\beta)$ is $\mu(ba)_{12} = 7$, and it is attained when β is an integral translate of $[\overline{1122}]$ or $[\overline{1221}]$, or of their conjugates.

In view of the previous results, one may wonder if in a given class of quadratic numbers, the numbers with a given value of D_0 lie in finitely many integral translations classes.

Part 4. Study of the invariant D_0

The main result of this part is that, for any quadratic number ξ, one may characterize the invariant $D_0(\xi)$ (and thus the second Lagrange number) as the 21-coefficient of the stabilizer of ξ in $PGL_2(\mathbb{Z})$. This gives an intrinsic characterization of $D_0(\xi)$, without use of the continued fraction expansion of ξ.

20. Stabilizer and invariant

Proposition 20.1. *Let ξ be a real quadratic number. The stabilizer of ξ in $PGL_2(\mathbb{Z})$ is an infinite cyclic group.*

The result is certainly well-known, but I could not find a reference (the similar result for $PSL_2(\mathbb{Z})$ is known, see [1] Proposition 5.2.4). We derive the proposition from a similar result for quadratic forms (Theorem 6.2.4 in [6]).

Proof. 1. In this proof only, we use an alternative action of $GL_2(\mathbb{Z})$ on quadratic forms, as it appears in [6] (2.2) et (2.4) p. 21. The action differs from the one used in the present article by the multiplication by the determinant. Let $f(x,y)$ be a quadratic form and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; then we define the quadratic form fM by $fM(x,y) = det(M)f(ax + by, cx+dy)$. In other words, $fM = det(M)f \cdot M$, with the notation $\cdot$ defined in Section 9.

2. Suppose that f is indefinite, with positive root ξ. It follows from Lemma 9.2 that if $g = fM$ and ξ is the positive root of g, then $M \cdot \xi$ is the positive root of f.

3. Let ξ be real quadratic number. Then there are two primitive integral quadratic forms having ξ as root: these two forms differ by $\pm$. Hence, by the observations in Section 9, there is a unique primitive integral quadratic form whose positive root is ξ. Denote it by f.

4. If $fM = f$, then by 2., $M \cdot \xi$ is the positive root of f, hence it is ξ. Conversely, if $M \cdot \xi = \xi$, then, with $g = fM$ and α the positive root of g, we have by 2. that $M \cdot \alpha$ is the positive root of f, thus $M \cdot \alpha = \xi$, hence $\xi = \alpha$. Since g is primitive, we must have $g = f$ by 3.

5. It follows from 4. that the stabilizer in $GL_2(\mathbb{Z})$ of f coincides with the stabilizer in $GL_2(\mathbb{Z})$ of ξ. By Theorem 6.2.4 in [6] p.133, the stabilizer of f is of the form $\{\pm T^k, k \in \mathbb{Z}\}$; hence the stabilizer in $PGL_2(\mathbb{Z})$ is a cyclic group. It is infinite, as follows from the expansion of ξ as a continued fraction, which is periodic. □

We also need the following lemma, which could be well-known, but I could not find a reference. The proof of freeness was shortened thanks to an idea of Jean-Eric Pin.

Lemma 20.1. *The set $\mathcal{P}$ of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $GL_2(\mathbb{N})$ such that $a \geq b \geq d, a \geq c \geq d$ is equal to the multiplicative semigroup generated by the matrices $P(x) = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}$, $x \in \mathbb{Z}_{>0}$, and they generate it freely.*

Proof. 0. Each nonempty product of matrices $P(x)$ satisfies the indicated inequalities: this follows from Lemma 3.1 (v) and from (2).

1. We prove the converse. Suppose that some matrix as in the statement satisfies the indicated inequalities. If $d = 0$, then we must have $b = c = 1$, and $a \geq 1$, and the matrix is equal to $P(a)$.

Suppose that $d \geq 1$. By Euclidean division we have $b = dq + r$, $0 \leq r < d$. We have $q \geq 1$ since $b \geq d$. Then $P(q)^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a - cq & r \end{pmatrix}$, and by induction, all we have to show is that $c \geq a - cq \geq r$.

We have $ad - bc = \epsilon$, $\epsilon = \pm 1$. Hence $a - cq = \frac{\epsilon + bc}{d} - cq = \frac{\epsilon}{d} + \frac{c(b-dq)}{d} = \frac{\epsilon}{d} + \frac{cr}{d}$. We have $0 \leq \frac{c(b-dq)}{d} = \frac{cr}{d} < c$. Thus $\frac{\epsilon}{d} \leq a - cq < \frac{\epsilon}{d} + c \leq c + 1$. Hence $a - cq \leq c$.

We show now that $a - cq \geq r$. Suppose first that $c > d$, hence $c \geq d + 1$. We have (by the previous calculation)

$$a - cq = \frac{\epsilon + cr}{d} \geq \frac{(d+1)r + \epsilon}{d} \geq r + \frac{r + \epsilon}{d}.$$

The latter is $\geq r$ for $r \geq 1$.

If $r = 0$, then $b = dq$ and $\epsilon = d(a - cq)$ hence $d = 1$ and $a = cq + \epsilon$. We have two subcases: $\epsilon = 1$ and then our matrix is $\begin{pmatrix} qc + 1 & q \\ c & 1 \end{pmatrix} = P(q, c)$; $\epsilon = -1$ and then our matrix is $\begin{pmatrix} qc - 1 & q \\ c & 1 \end{pmatrix} = P(q - 1, 1, c - 1)$ and we have $cq - 1 \geq q, c$ so that $c, q \geq 2$, and the matrix is in $\mathcal{P}$ in both cases.

Suppose now that $c = d$. Then $\epsilon = ac - bc = c(a - b)$ so that $c = 1, \epsilon \geq 0, a = b + 1$ and our matrix is $\begin{pmatrix} b + 1 & b \\ 1 & 1 \end{pmatrix} = P(b, 1)$.

2. Freeness: suppose that $M = P(a_1) \cdots P(a_n), n \geq 1, a_i \in \mathbb{Z}_{>0}$. Let $A = P(0)P(1) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = P(1)P(0) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. It is well-known that A, B generate freely the monoid $SL_2(\mathbb{N})$. Moreover $P(0)$ conjugates A and B, and $P(n) = A^n P(0)$. Hence $M = A^{a_1}P(0) \cdots A^{a_n}P(0)$. If n is even, equivalently $\det(M) = 1$, then $M = A^{a_1}B^{a_2} \cdots B^{a_n}$; hence the sequence $a_1, \ldots, a_n$ is completely determined by M, since $SL_2(\mathbb{N})$ is free on A, B. If n is odd, equivalently $\det(M) = -1$, then $M = A^{a_1}B^{a_2} \cdots A^{a_n}P(0)$ and we argue similarly. □

Corollary 20.1. *Suppose that some power A^n of $A \in GL_2(\mathbb{Z})$ with positive exponent is in the semigroup $\mathcal{P}$ of the previous lemma. Suppose also that A has two distinct real fixed points and that $tr(A) \geq 1$. Then $A \in \mathcal{P}$.*

Proof. 1. We may assume that $n \geq 2$. Let $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, $A^n = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since A has distinct real fixed points, the discriminant of $\gamma x^2 + (\delta - \alpha)x - \beta$ is positive, that is $t^2 - 4\epsilon > 0$, with $t = tr(A)$ and $\epsilon = det(A) = \pm 1$. Hence either $\epsilon = -1$, or $\epsilon = 1$ and $t \geq 3$.

Define the sequence s_k, $k \geq -1$, by $s_{-1} = -\epsilon$, $s_0 = 0$ and the recursion $s_k = ts_{k-1} - \epsilon s_{k-2}$ for any $k \geq 1$. Then it follows from the identity $A^2 = tA - \epsilon I$ that $A^k = s_k A - \epsilon s_{k-1} I$ for any $k \geq 0$.

We have therefore $a = s_n\alpha - \epsilon s_{n-1}$, $b = s_n\beta$, $c = s_n\gamma$, $d = s_n\delta - \epsilon s_{n-1}$.

2. We claim that for $k \geq 1$, $s_k \geq 1$, and that for $k \geq 2$, $s_k/s_{k-1} > 1$, except if $k = 2, t = 1, \epsilon = -1$. The claim is proved below.

3. We assume that the general case holds. Then for $n \geq 1$, $s_n \geq 1$; and for $n \geq 2$, $0 < s_{n-1}/s_n < 1$, hence $\pm\epsilon s_{n-1}/s_n > -1$. Recall that $a \geq b \geq d, a \geq c \geq d \geq 0$. We deduce that $\beta, \gamma \geq 0$. We have $d = s_n(\delta - \epsilon s_{n-1}/s_n) \geq 0$, hence $\delta \geq \epsilon s_{n-1}/s_n > -1$. We deduce that $\delta \geq 0$.

We have $a - b = s_n\alpha - \epsilon s_{n-1} - s_n\beta = s_n(\alpha - \beta) - \epsilon s_{n-1}$ and thus $\alpha - \beta = (a-b)/s_n + \epsilon s_{n-1}/s_n > -1$, hence $\alpha \geq \beta$.

We have $b - d = s_n\beta - s_n\delta + \epsilon s_{n-1}$, hence $\beta - \delta = (b-d)/s_n - \epsilon s_{n-1}/s_n$. We conclude as before that $\beta \geq \delta$. The other inequalities $\alpha \geq \gamma \geq \delta$ are proved similarly. Thus $A \in \mathcal{P}$.

4. In the exceptional case $n = 2, t = 1, \epsilon = -1$, we have $A^n = A^2 = tA - \epsilon I = A + I = \begin{pmatrix} \alpha + 1 & \beta \\ \gamma & \delta + 1 \end{pmatrix}$. By hypothesis, $\alpha + 1 \geq \beta \geq \delta + 1 \geq 0$ and $\alpha + 1 \geq \gamma \geq \delta + 1$. Then $\delta \geq -1$. Suppose that $\delta = -1$; it follows that $A^2 = \begin{pmatrix} \alpha + 1 & \beta \\ \gamma & 0 \end{pmatrix}$, hence $\beta = \gamma = 1$ and $\alpha = 1 - \delta = 2$; then $A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$, whose determinant is -3, a contradiction. Hence $\delta \geq 0$. Thus $\alpha \geq 0$, $\gamma \geq 1$, $\beta \geq \delta$, $\gamma \geq \delta$ by the previous inequalities.

Suppose that $\alpha < \beta$. Then $\alpha + 1 = \beta$. Thus $A = \begin{pmatrix} \alpha & \alpha + 1 \\ \gamma & 1 - \alpha \end{pmatrix}$, whose determinant is $-1 = \alpha - \alpha^2 - \alpha\gamma - \gamma$; thus $0 = \alpha^2 + \alpha(\gamma - 1) + (\gamma - 1)$; since $\gamma \geq 1$, we must have $\gamma = 1$ and $\alpha = 0$; but $\alpha + 1 \geq \delta + 1 = 2 - \alpha$, hence $2\alpha \geq 1$, $\alpha > 0$, a contradiction. Thus $\alpha \geq \beta$ and similarly $\alpha \geq \gamma$. Thus $A \in \mathcal{P}$.

5. For the claim, suppose first that $\epsilon = 1$. Then $t \geq 3$. We have $s_0 = 0, s_1 = 1$, and for $k \geq 2$, $s_k/s_{k-1} = t - s_{k-2}/s_{k-1}$; hence $s_2/s_1 = t > 1$, and for $k \geq 3$, $s_k/s_{k-1} > 1$ by induction. It follows also that $s_k \geq 1$ for $k \geq 1$.

Suppose now that $\epsilon = -1$. Then $s_0 = 0, s_1 = 1$ and $s_k = ts_{k-1} + s_{k-2}$. It follows easily that $s_k \geq 1$ for any $k \geq 1$. We have $s_2 = ts_1 + s_0 = t \geq 1$ and therefore $s_2/s_1 = t$. We have $s_k/s_{k-1} = t + s_{k-2}/s_{k-1} > t$ if $k \geq 3$. We have also $s_2/s_1 > 1$ if $t \neq 1$. □

Corollary 20.2. *Let $\xi = [\overline{a_0, \dots, a_{n-1}}]$ be a reduced real quadratic number, with n minimal. Its stabilizer in $PGL_2(\mathbb{Z})$ is generated by $P(a_0) \dots P(a_{n-1})$.*

Proof. It is well-known that the matrix $P(a_0) \dots P(a_{n-1})$ stabilizes ξ. Let $A \in GL_2(\mathbb{Z})$ such that $A \bmod \{I_2, -I_2\}$ is a generator of the stabilizer of ξ in $PGL_2(\mathbb{Z})$ (Proposition 20.1). In $PGL_2(\mathbb{Z})$, some nontrivial power of A is equal to $P(a_0, \dots, a_{n-1})$. We may assume that the exponent is positive, by replacing A by A^{-1} if necessary. We may assume that $tr(A) \geq 0$, by replacing A by $-A$ if necessary. The matrix A has two different real fixed points ξ and $\bar{\xi}$, and its trace is ≥ 1 (otherwise it is 0, and then $A^2 = \pm I$ and the stabilizer is finite, a contradiction). Then $A^h = \pm P(a_0, \dots, a_{n-1})$.

Taking the notations of the proof of Corollary 20.1, we have $tr(A^h) = ts_h - 2\epsilon s_{h-1} = s_{h+1} - \epsilon s_h$. For $h \geq 1$, this is ≥ 1. Indeed, this is clear if $\epsilon = -1$ and if $\epsilon = 1$, we have $s_{h+1}/s_{h-1} > 1$ for $h \geq 2$, and therefore $s_{h+1} - s_{h-1} > 0$. Since the trace of $P(a_1, \dots, a_{h-1})$ is positive, we see that $A^h = P(a_0, \dots, a_{n-1})$.

It then follows from Corollary 20.1 that $A \in \mathcal{P}$, that is $A = P(b_0, \dots, b_{k-1})$. Thus $(P(b_0) \cdots P(b_{k-1}))^h = P(a_0) \cdots P(a_{n-1})$. Since $\mathcal{P}$ is a monoid freely generated by the matrices $P(a)$, we see that the sequence $b_0, \dots, b_{k-1}$ repeated h times is equal to the sequence $a_0, \dots, a_{n-1}$. By minimality of n, $h = 1$, $k = n$, $b_i = a_i$ and finally $A = P(a_0, \dots, a_{n-1})$. □

Theorem 20.1. *The invariant $D_0(\xi)$ of a real quadratic number is equal to $|c|$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the generator of the stabilizer in $PGL_2(\mathbb{Z})$ of ξ.*

Proof. Let $\beta = [b_0, \dots, b_{k-1}, \overline{a_0, \dots, a_{n-1}}]$ be a quadratic number with n minimal. We know that $D_0(\beta) = |f(e, i)|$ with the notations of Section 9. In this section, one finds also the following equalities: $f(e, i) = g(1, 0) = f_B(1, 0) = 2, 1$-entry of B, with $B = M^{-1}AM$, $A = P(a_0, \dots, a_{n-1})$; moreover, $\beta = M^{-1} \cdot \alpha$, so that the stabilizers G and H of α and β are related by $H = M^{-1}GM$. Since by Corollary 20.2, G is generated by A, we see that H is generated by B and the theorem follows. □

Remark 20.1. Note that by the same calculation, one has that ξ is a root of a quadratic equation of the form $D_0(\xi)\xi^2 + b\xi + c$, with integer coefficients. However, it is not true that $D_0(\xi)$ is equal to the smallest positive dominating coefficient a of an integral quadratic equation for ξ. This is shown by the example $\xi = [\overline{2, 4}]$; one has $D_0(\xi) = 4$ (since $P(2, 4) = \begin{pmatrix} 9 & 2 \\ 4 & 1 \end{pmatrix}$), and ξ satisfies the equation $4\xi^2 - 8\xi - 2 = 0$, whose coefficients are not relatively prime. Hence, although a always divides $D_0(\xi)$, they may be unequal.

Observe also that Theorem 20.1 is in general not true if one replaces PGL by PSL, as the example of the positive fixed point of the matrix $P(2)$ shows: one has $D_0(\sqrt{2}+1) = 1$ and its stabilizer in $PSL_2(\mathbb{Z})$ is generated by $P(2)^2 = \mu(a)$, whose $2,1$-coefficient is 2.

Corollary 20.3. *$D_0(\xi) = D_0(\bar{\xi})$ for any real quadratic number.*

Proof. Indeed, ξ and $\bar{\xi}$ have the same stabilizer in $PGL_2(\mathbb{Z})$. □

Finally, we leave the proof of the following result to the reader. One may use the link between the automorphism group of a form and the Pell equation, see for example [14] Theorem 87 p.112.

Corollary 20.4. *$D_0(\xi)$ is equal to $|ay|$, where the minimal polynomial of ξ over the integers is $ax^2 + bx + c$ and where for some integer x, the pair (x, y) is the minimum solution of the generalized Pell equation $x^2 - (b^2 - 4ac)y^2 = \pm 4$.*

As an application, for d non square natural number, $D_0(\sqrt{d})$ is equal to the smallest positive y such that $x^2 - dy^2 = \pm 1$.

References

[1] M. Aigner, Markov's Theorem and 100 Years of the Uniqueness Conjecture, Springer Verlag, 2013.
[2] E. Bombieri, Continued fractions and the Markoff tree, Expo. Math. 25 (2007) 187–213.
[3] J.-P. Borel, F. Laubie, Quelques mots sur la droite projective réelle, J. Théor. Nr. Bordx. 5 (1993) 23–51.
[4] J. Borwein, A. van der Poorten, J. Shallit, W. Zudilin, Neverending Fractions: An Introduction to Continued Fractions, Cambridge University Press, 2014.
[5] E.B. Burger, A. Folsom, A. Pekker, R. Roengpyta, J. Snyder, On a quantitative refinement of the Lagrange spectrum, Acta Arith. 102 (2002) 55–82.
[6] J. Buchmann, U. Vollmer, Binary Quadratic Forms, An Algorithmic Approach, Springer Verlag, 2007.
[7] J.W.S. Cassels, An Introduction to Diophantine Approximation, Cambridge University Press, 1957.
[8] W.-F. Chuan, α-words and factors of characteristic sequences, Discrete Math. 177 (1997) 33–50.
[9] P.M. Cohn, Free Rings and Their Relations, Academic Press, 1985.
[10] H. Cohn, Approach to Markoff's minimal forms through modular functions, Ann. Math. 61 (1955) 1–12.
[11] H. Cohn, Markoff's forms and primitive words, Math. Ann. 61 (1972) 8–22.
[12] H. Cohn, Growth types of Fibonacci and Markoff, Fibonacci Q. 17 (1979) 178–183.
[13] T.W. Cusick, M.E. Flahive, The Markoff and Lagrange Spectra, AMS, 1989.
[14] L.E. Dickson, Introduction to the Theory of Numbers, Dover, 1957.
[15] G. Frobenius, Über die Markoffschen Zahlen, Sitz.ber. K. Preuss. Akad. Wiss. Berl. 26 (1913) 458–487.
[16] J. Hančl, Sharpening of theorems of Vahlen and Hurwitz and approximation properties of the golden ratio, Arch. Math. 105 (2015) 129–137.
[17] J. Hančl, Second basic theorem of Hurwitz, Lith. Math. J. 56 (2016) 72–76.
[18] G.H. Hardy, E.M. Wright, An Introduction to Theory of Numbers, 5th edition, Oxford University Press, 1979.
[19] F. Herzog, On the continued fractions of conjugate quadratic irrationalities, Can. Math. Bull. 23 (1980) 199–206.
[20] A. Hurwitz, Ueber die angenäherte Darstellung der Zahlen durch rationale Brüche, Math. Ann. 44 (1894) 417–436.

[21] A. Korkine, G. Zolotareff, Sur les formes quadratiques, Math. Ann. 6 (1873) 366–389.
[22] S. Mantaci, A. Restivo, M. Sciortino, Burrows-Wheeler transform and Sturmian words, Inf. Process. Lett. 86 (2001) 241–246.
[23] O. Perron, Über die Approximationen irrationaler Zahlen durch rationale, Sitz.ber. Heidelb. Akad. Wiss. 8 (1921) 2–12.
[24] C. Reutenauer, From Christoffel Words to Markoff Numbers, Springer, 2019.
[25] J.-A. Serret, Cours d'algèbre supérieure, Tome I, Edition Jacques Gabay, 1992 (first edition Gauthiers-Villars 1877).
[26] H.M. Wedderburn, Noncommutative domains of integrity, J. Reine Angew. Math. 167 (1932) 129–141.