Rendiconti Sem. Mat. Univ. Pol. Torino
Vol. 78, 1 (2020), 83 – 101

M. Elia

CONTINUED FRACTIONS AND FACTORING

Abstract. Legendre found that the continued fraction expansion of $\sqrt{N}$ having odd period leads directly to an explicit representation of N as the sum of two squares. Similarly, it is shown here that the continued fraction expansion of $\sqrt{N}$ having even period directly produces a factor of a composite N. Shanks' infrastructural method is then revisited, and some consequences of its application to factorization by means of the continued fraction expansion of $\sqrt{N}$ are derived.

Mathematics Subject Classification (2010): 11A55, 11A51

1. Introduction

Continued fractions have always held great fascination, for both aesthetic reasons and practical purposes. Among the many clever properties of periodic continued fractions, Legendre found how to obtain the representation of an integer N as the sum of two squares, in his own words, "*sans aucun tâtonnement*" from the continued fraction expansion of $\sqrt{N}$ when the period is odd [11]. In particular, this property holds for any prime p congruent 1 modulo 4, [11, 16]. As a kind of counterpart to Legendre's finding, this paper shows how to obtain a factor of a composite N directly from the continued fraction expansion of $\sqrt{N}$ when the period is even. In particular, this is certainly possible when both prime factors of N are congruent 3 modulo 4.

Based on this result, derived from peculiar properties of continued fraction convergents, and on an adaptation of Shanks' infrastructural machinery, a factoring algorithm is proposed whose complexity depends on the accuracy of the evaluation of certain integrals of Dirichlet's. The paper is organized as follows. Section 2 summarizes the properties of the continued fraction expansion of $\sqrt{N}$. In Section 3, some new properties of the convergents are proved, and Shanks' infrastructural method is revisited and applied to a sequence of quadratic forms generated from the convergent of the continued fraction expansion of $\sqrt{N}$. Section 4 discusses the factorization of composite numbers N when the period of the continued fraction expansion of $\sqrt{N}$ is even. Lastly, Section 5 briefly reports some conclusions.

2. Preliminaries

A regular continued fraction is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} \quad , \tag{1}$$

where a_0, a_1, $a_2,\dots,a_i,\dots$ is a sequence, possibly infinite, of positive integers. A convergent of a continued fraction is a sequence of fractions $\frac{A_m}{B_m}$, each of which is obtained by truncating the continued fraction at the m-th term. The fraction $\frac{A_m}{B_m}$ is called the m-th convergent [5, 8, 12]. The first few initial terms of the convergent of (1) are

$$\frac{A_0}{B_0}=\frac{a_0}{1},\ \frac{A_1}{B_1}=\frac{a_0a_1+1}{a_1},\ \frac{A_2}{B_2}=\frac{a_0a_1a_2+a_0+a_2}{a_1a_2+1},\dots \quad .$$

Numerators and denominators of the m-th convergent satisfy the second-order recurrences

$$\left\{\begin{array}{ll} A_m=a_mA_{m-1}+A_{m-2} & ,\ A_0=a_0,\ A_1=a_0a_1+1 \\ B_m=a_mB_{m-1}+B_{m-2} & ,\ B_0=1,\ B_1=a_1 \end{array}\right. \quad ,\ \forall\, m\geq 2; \tag{2}$$

further, we have [5, p.85] the relationships

$$A_mB_{m-1}-A_{m-1}B_m=(-1)^{m-1} \tag{3}$$

$$A_mB_{m-2}-A_{m-2}B_m=(-1)^{m-2}a_m \quad . \tag{4}$$

Equation (3) shows that numerator and denominator of the m-th convergent are relatively prime.
A continued fraction is said to be definitively periodic, with period τ, if, starting from a finite n_o, a fixed pattern a'_1, $a'_2,\dots,a'_\tau$ repeats indefinitely. Lagrange showed that any definitively periodic continued fraction represents a positive number of the form $a+b\sqrt{N}$, $a,b\in\mathbb{Q}$, i.e. an element of $\mathbb{F}=\mathbb{Q}(\sqrt{N})$, and conversely that any such positive number is represented by a definitively periodic continued fraction [5, 16]. The maximal order of $\mathbb{F}$ is denoted $\mathfrak{O}_{\mathbb{F}}$. Let $\mathcal{G}(\mathbb{F}/\mathbb{Q})=\{\iota,\sigma\}$ be the Galois group of $\mathbb{F}$ over $\mathbb{Q}$, where ι denotes the group identity, and the action of the automorphism σ, called conjugation, is defined as $\sigma(a+b\sqrt{N})=a-b\sqrt{N}$. The field norm $N_{\mathbb{F}}(\mathfrak{a})$ of $\mathfrak{a}\in\mathbb{F}$ is defined to be $N_{\mathbb{F}}(\mathfrak{a})=\mathfrak{a}\sigma(\mathfrak{a})$.
In the continued fraction expansion of $\sqrt{N}$, the period of length τ begins immediately after the first term a_0, and consists of a palindromic part formed by $\tau-1$ terms $a_1,a_2,\dots,a_2,a_1,$ followed by $2a_0$. Periodic continued fractions of this sort are conventionally written in the form

$$\sqrt{N}=\left[a_0,\overline{a_1,a_2,\dots,a_2,a_1,2a_0}\right] \quad , \tag{5}$$

where the over-lined part is the period. Note that the period of the irrational $\frac{a_0+\sqrt{N}}{N-a_0^2}$ starts immediately without anti-period; in this case, the continued fraction is called purely periodic and is denoted $\left[\overline{a_1,a_2,\dots,a_2,a_1,2a_0}\right]$.
Carr's book [2, p.70-71] gives a good collection of properties of the continued fraction expansion of $\sqrt{N}$, which are summarized in the following, with the addition of some properties taken from [5, 16, 13]:

1. Let c_n and r_n be the elements of two sequences of positive integers defined by the relation
$$\frac{\sqrt{N}+c_n}{r_n} = a_{n+1} + \frac{r_{n+1}}{\sqrt{N}+c_{n+1}}$$
with $c_0 = \lfloor \sqrt{N} \rfloor$, and $r_0 = N - a_0^2$; the elements of the sequence $a_1, a_2, \ldots, a_n \ldots$ are thus obtained as the integer parts of the left-side fraction
$$a_{n+1} = \left\lfloor \frac{\sqrt{N}+c_n}{r_n} \right\rfloor \quad . \tag{6}$$

2. Let $a_0 = \lfloor \sqrt{N} \rfloor$ be initially computed, and set $c_0 = a_0$, $r_0 = N - a_0^2$, then sequences $\{c_n\}_{n\geq 0}$ and $\{r_n\}_{n\geq 0}$ are produced by the recursions
(7)
$$a_{m+1} = \left\lfloor \frac{a_0 + c_m}{r_m} \right\rfloor \quad , \quad c_{m+1} = a_{m+1} r_m - c_m \quad , \quad r_{m+1} = \frac{N - c_{m+1}^2}{r_m} \quad .$$
These recursive equations, together with (6), allow us to compute the sequence $\{a_m\}_{m\geq 1}$ using rational arithmetical operations; however, the iterations may be stopped when $a_m = 2a_0$, having completed a period.

3. The n-th convergent to $\sqrt{N}$ can be recursively computed as
$$\frac{A_n}{B_n} = \frac{a_n A_{n-1} + A_{n-2}}{a_n B_{n-1} + B_{n-2}} \quad n \geq 1 \quad , \tag{8}$$
with initial conditions $A_{-1} = 1$, $B_{-1} = 0$, $A_0 = a_0$, and $B_0 = 1$.

4. The sequence of ratios $\frac{A_n}{B_n}$ assumes the limit value $\sqrt{N}$ as n goes to infinity, due to the inequality
$$\left| \frac{A_n}{B_n} - \sqrt{N} \right| \leq \frac{1}{B_n B_{n+1}} \quad ,$$
since A_n and B_n go to infinity along with n. Furthermore, $\frac{A_n}{B_n} < \sqrt{N}$, if n is even, and $\frac{A_n}{B_n} > \sqrt{N}$ if n is odd [8, p.132]. Therefore, any convergent of even index is smaller than any convergent of odd index.

5. The true value of $\sqrt{N}$ is the value which (8) becomes when the "approximated" quotient a_n, as defined in (6), is substituted with the complete quotient $\frac{\sqrt{N}+c_{n-1}}{r_{n-1}}$. This gives
$$\sqrt{N} = \frac{(\sqrt{N}+c_{n-1})A_{n-1} + r_{n-1}A_{n-2}}{(\sqrt{N}+c_{n-1})B_{n-1} + r_{n-1}B_{n-2}} \quad .$$

6. The value $c_0 = a_0$ is the greatest value that c_n may assume. No a_n or r_n can be greater than $2a_0$.
If $r_n = 1$ then $a_{n+1} = a_0$. For all n greater than 0, we have $a_0 - c_n < r_n$.

7. The first complete quotient that is repeated is $\frac{\sqrt{N}+c_0}{r_0}$, and a_1, r_0, and c_0 commence each cycle of repeated terms.

8. Through the first period (or cycle) of length τ, the elements $a_{\tau-j}$, $r_{\tau-j-2}$, and $c_{\tau-j-1}$ are respectively equal to a_j, r_j, and c_j.

9. The period length cannot be greater than $2a_0^2$. This bound is very loose and was tightened by Kraitchik [17, p.95], who showed that τ is upper bounded by

$$0.72\sqrt{N}\ln N \qquad N > 7 \ . \tag{9}$$

However, the period length has irregular behavior as a function of N, because it may assume any value from 1, when $N = M^2+1$, to values close to the order $O(\sqrt{N}\ln N)$ [16].

10. The element $\mathfrak{c}_m = A_m + B_m\sqrt{N} \in \mathfrak{O}_{\mathbb{F}}$ is associated to the m-th convergent.

Numerators and denominators of the convergents satisfy interesting relations [12, p.92-95]

(10)

$$\begin{cases} A_0A_{\tau-1} + A_{\tau-2} - NB_{\tau-1} = 0 \\ A_1A_{\tau-2} + A_0A_{\tau-3} - N(B_1B_{\tau-2} + B_0B_{\tau-3}) = 0 \\ A_jA_{\tau-j-1} + A_{j-1}A_{\tau-j-2} - N(B_jB_{\tau-j-1} + B_{j-1}B_{\tau-j-2}) = 0 \qquad 3 \le j \le \tau-3 \ . \end{cases}$$

Besides these properties, the following equations, [16, p.329-332], are used in the proofs:

$$\begin{cases} A_\tau = 2a_0A_{\tau-1} + A_{\tau-2} \\ B_\tau = 2a_0B_{\tau-1} + B_{\tau-2} \end{cases} \tag{11}$$

$$\begin{cases} A_\tau B_{\tau-1} - A_{\tau-1}B_\tau = (-1)^{\tau-1} \\ A_{\tau-1}B_{\tau-2} - A_{\tau-2}B_{\tau-1} = (-1)^{\tau-2} \\ A_\tau B_{\tau-2} - A_{\tau-2}B_\tau = 2a_0(-1)^\tau \end{cases} \tag{12}$$

$$\begin{cases} A_{\tau-2} = -a_0A_{\tau-1} + NB_{\tau-1} \\ B_{\tau-2} = A_{\tau-1} - a_0B_{\tau-1} \end{cases} \tag{13}$$

$$\begin{cases} A_\tau = a_0A_{\tau-1} + NB_{\tau-1} \\ B_\tau = A_{\tau-1} + a_0B_{\tau-1} \end{cases} \tag{14}$$

REMARK 1. The smallest positive solution of Pell's equation $x^2 - Ny^2 = (\pm 1)$ is $\mathfrak{c}_{\tau-1}$, whenever a solution exists. If $\{1, \sqrt{N}\}$ is an integral basis of $\mathbb{F}$, then $\mathfrak{c}_{\tau-1}$ coincides with the fundamental positive unit ε_0 of $\mathbb{F}$. If $\{1, \frac{1+\sqrt{N}}{2}\}$ is an integral basis of $\mathbb{F}$, then $\mathfrak{c}_{\tau-1}$ may be either ε_0 or ε_0^3. An easy way to check whether $\mathfrak{c}_{\tau-1} = \varepsilon_0^3$ is to solve in $\mathbb{Q}$ the equation $(x+y\sqrt{N})^3 = A_{\tau-1} + B_{\tau-1}\sqrt{N}$, which is equivalent to verifying whether some solution of the following Diophantine equation is a rational number with 2 as denominator

$$64x^9 - 48A_{\tau-1}x^6 + (27NB_{\tau-1}^2 - 15A_{\tau-1}^2)x^3 - A_{\tau-1}^3 = 0 \quad .$$

If a rational solution x_o of this equation exists, the corresponding y_o can be computed as $y_o = \sqrt{\frac{x_o^2-1}{N}}$.

The following proposition describes how to move from one period to another.

PROPOSITION 1. *The sequence $\{\mathfrak{c}_m\}_{m\geq 0}$ satisfies the relation*

$$\mathfrak{c}_{m+k\tau} = \mathfrak{c}_m \mathfrak{c}_{\tau-1}^k \quad \forall\, m,k \in \mathbb{N} \quad . \tag{15}$$

Proof. The two dependencies, with respect to m and k, are disposed of separately. The claimed equality is trivial for $m = k = 0$, and fixing $k = 1$, equation (14) allows us to write $\mathfrak{c}_\tau = a_0\mathfrak{c}_{\tau-1} + \sqrt{N}\mathfrak{c}_{\tau-1} = (a_0+\sqrt{N})\mathfrak{c}_{\tau-1} = (A_0+B_0\sqrt{N})\mathfrak{c}_{\tau-1}$. Then, by the recurrences (2) and the periodicity of the a_is, we can write

$$\mathfrak{c}_{\tau+1} = a_1\mathfrak{c}_\tau + \mathfrak{c}_{\tau-1} = a_1(A_0+B_0\sqrt{N})\mathfrak{c}_{\tau-1} + \mathfrak{c}_{\tau-1} = \mathfrak{c}_1\mathfrak{c}_{\tau-1} \quad .$$

Clearly, we can iterate by using the recurrences (2) and the symmetry of the a_is to obtain the relation $\mathfrak{c}_{\tau+m} = \mathfrak{c}_m\mathfrak{c}_{\tau-1}$, which shows that multiplication by $\mathfrak{c}_{\tau-1}$ is equivalent to a translation by τ. The conclusion is immediate by iterating on k. □

3. Convergents and quadratic forms

Let $\Delta_m = A_m^2 - NB_m^2$ denote the field norm of $\mathfrak{c}_m = A_m + \sqrt{N}B_m \in \mathfrak{O}_{\mathbb{F}}$. Several properties of convergents are better described considering, besides the sequence $\Delta = \{\Delta_m\}_{m\geq 0}$, a second sequence $\Omega = \{\Omega_m = A_mA_{m-1} - NB_mB_{m-1}\}_{m\geq 1}$. Using (3), the following relation can be shown

$$\Omega_{m+1}^2 - \Delta_m\Delta_{m+1} = N \quad \forall m \geq 0 \quad . \tag{16}$$

The elements of the sequences Δ and Ω satisfy the recurrent relations

$$\left\{ \begin{array}{l} \Delta_{m+1} = a_{m+1}^2\Delta_m + 2a_{m+1}\Omega_m + \Delta_{m-1} \\ \Omega_{m+1} = \Omega_m + a_{m+1}\Delta_m \end{array} \right. \qquad m \geq 1 \tag{17}$$

with initial conditions $\Delta_0 = a_0^2 - N$, $\Delta_1 = (1+a_0a_1)^2 - Na_1^2$ and $\Omega_1 = (1+a_0a_1)a_0 - Na_1$. Using (17), it is immediate to see that $c_{m+1} = |\Omega_m|$ and $r_{m+1} = |\Delta_m|$.
Introducing the matrix

$$T(a_m) = \begin{bmatrix} a_m^2 & a_m & 1 \\ 2a_m & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad , \tag{18}$$

and defining the column vector $\Lambda_m = [\Delta_m, 2\Omega_m, \Delta_{m-1}]^T$, equations (17) can be written as

$$\Lambda_{m+1} = T(a_{m+1})\Lambda_m \quad \forall\, m \geq 1 \ . \tag{19}$$

Iterating this relation, we have

$$\Lambda_{m+n} = T(a_{m+n})T(a_{m+n-1})\cdots T(a_{m+2})T(a_{m+1})\Lambda_m = T_{(m,n)}\Lambda_m \quad \forall\, m,n \geq 1 \ , \tag{20}$$

where $T_{(m,n)} = \prod_{j=m+1}^{m+n} T(a_j)$ is a matrix that only depends on the sequence of coefficients a_t. Furthermore, from (16) we may derive the relation

$$\Omega_{m+1}^2 - \Omega_m^2 = \Delta_m(\Delta_{m+1} - \Delta_{m-1}) \ \ \forall\, m \geq 1,$$

which allows us to write equation (17) as

$$\begin{cases} \Delta_{m+1} = \Delta_{m-1} + a_{m+1}(\Omega_{m+1} + \Omega_m) \\ \Omega_{m+1} = \Omega_m + a_{m+1}\Delta_m \end{cases} \quad \forall\, m \geq 1 \ . \tag{21}$$

DEFINITION 1. *Let Υ be the sequence of quadratic forms $f_m(x,y) = \Delta_m x^2 + 2\Omega_m xy + \Delta_{m-1}y^2$, $m \geq 1$, defined by means of the sequences Δ and Ω.*

Note that it may sometimes be convenient to denote a quadratic form simply with the triple of coefficients, i.e. the 3-dimensional vector Λ_m; further, due to equation (16), all quadratic forms in Υ have the same discriminant $4N$.

REMARK 2. The absolute values of Δ_m and Ω_m are bounded as

$$|\Delta_m| < 2\frac{1}{a_{m+1}}\sqrt{N} \leq 2\sqrt{N} \qquad , \qquad |\Omega_m| < \sqrt{N} \quad \forall\, m \geq 1 \ .$$

The bound $2\sqrt{N}$ for Δ_m is well known, [8, Theorem 171, p.140], and can be slightly tightened considering the following chain of inequalities

$$\begin{aligned} |A_m^2 - NB_m^2| &= B_m^2\left|\frac{A_m}{B_m} - \sqrt{N}\right|(\frac{A_m}{B_m} + \sqrt{N}) \leq \frac{B_m}{B_{m+1}}\left|\frac{A_m}{B_m} - \sqrt{N} + 2\sqrt{N}\right| \\ &\leq \frac{B_m}{B_{m+1}}\left|\frac{A_m}{B_m} - \sqrt{N}\right| + 2\sqrt{N}\frac{B_m}{B_{m+1}} \leq \frac{1}{B_{m+1}^2} + 2\frac{B_m}{a_{m+1}B_m + B_{m-1}}\sqrt{N} \\ &= 2\frac{1}{a_{m+1}}\sqrt{N} + \frac{1}{B_{m+1}^2} - 2\sqrt{N}\frac{B_{m-1}}{a_{m+1}(a_{m+1}B_m + B_{m-1})} < 2\frac{1}{a_{m+1}}\sqrt{N} \ . \end{aligned}$$

The bound for $|\Omega_m|$ is an immediate consequence of equation (16), we have $\Delta_m\Delta_{m+1} < 0$ since the signs in the sequence Δ alternate; consequently

$$\Omega_m^2 = N + \Delta_m\Delta_{m+1} < N \ ,$$

thus taking the positive square root of both sides, the inequality $|\Omega_m| < \sqrt{N}$ is obtained.

3.1. Periodicity and Symmetry

The sequences Δ and Ω are periodic in the same way as the sequence of coefficients a_m, although their periods are even, and may be τ or 2τ depending on whether τ is even or odd. Further, within a period, there exist interesting symmetries.

THEOREM 1 (Periodicity of Δ). *Starting with $m = 1$, the sequence $\Delta = \{\Delta_m\}_{m\geq 0}$ is periodic with period τ or 2τ depending on whether τ is even or odd. The elements of the first block $\{\Delta_m\}_{m=0}^{\tau} \subset \Delta$ satisfy the symmetry relation $\Delta_m = (-1)^\tau \Delta_{\tau-m-2}$, $\forall\, 0 \leq m \leq \tau-2$.*

Proof. The period of the sequence Δ is τ or 2τ, as a consequence of equation (15), because the norm of $A_{\tau-1} + \sqrt{N}B_{\tau-1}$ is $(-1)^\tau$.
The symmetry of the sequence Δ within the τ elements of the first period follows from the relations

$$(22)\qquad \left\{\begin{array}{l} A_{\tau-m-2} = (-1)^{m-1}A_{\tau-1}A_m + (-1)^m NB_{\tau-1}B_m \\ B_{\tau-m-2} = (-1)^m A_{\tau-1}B_m + (-1)^{m-1}B_{\tau-1}A_m \end{array}\right. ,\quad 0 \leq m \leq \tau-2\,,$$

which are proved using the recurrences (2) together with (13) and (14) [16, p.329-330]; the transformation defined by (22) is identified by the matrix

$$(23)\qquad M_{\tau-1} = \left[\begin{array}{cc} -A_{\tau-1} & NB_{\tau-1} \\ -B_{\tau-1} & A_{\tau-1} \end{array}\right] .$$

We have

$$\left\{\begin{array}{rcl} A^2_{\tau-m-2} - NB^2_{\tau-m-2} & = & (A_{\tau-1}A_m - NB_{\tau-1}B_m)^2 - N(-A_{\tau-1}B_m + B_{\tau-1}A_m)^2 \\ & = & (A_m^2 - NB_m^2)(A^2_{\tau-1} - NB^2_{\tau-1}) = (-1)^\tau (A_m^2 - NB_m^2) \end{array}\right.$$

that is $\Delta_{\tau-m-2} = (-1)^\tau \Delta_m$. Actually, equation (22) can be written in the form

$$(24)\qquad A_{\tau-m-2} + \sqrt{N}B_{\tau-m-2} = (-1)^{m-1}(A_{\tau-1} + \sqrt{N}B_{\tau-1})(A_m - \sqrt{N}B_m)$$

or more compactly as $\mathfrak{c}_{\tau-m-2} = (-1)^{m-1}\mathfrak{c}_{\tau-1}\sigma(\mathfrak{c}_m)$. □

THEOREM 2 (Periodicity of Ω). *The sequence $\Omega = \{\Omega_m\}_{m\geq 1}$ is periodic of period τ or 2τ depending on whether τ is even or odd. The elements of the first block $\{\Omega_m\}_{m=1}^{\tau} \subset \Omega$ satisfy the symmetry relation*
$\Omega_{\tau-m-1} = (-1)^{\tau+1}\Omega_m$, $\forall\, m \leq \tau-2$.

Proof. The periodicity of the sequence Ω follows from the property expressed by equation (15), noting that

$$\Omega_j = \frac{1}{2}\left((A_j + \sqrt{N}B_j)((A_{j-1} - \sqrt{N}B_{j-1}) + (A_j - \sqrt{N}B_j)((A_{j-1} + \sqrt{N}B_{j-1})\right) .$$

The symmetry property of the sequence Ω within a period follows from (22) in the same way as does that of the sequence Δ; we have

$$\begin{array}{rcl} A_{\tau-1-j}A_{(\tau-1)-j-1} - NB_{\tau-1-j}B_{(\tau-1)-j-1} & = & -(A_{\tau-1}A_j - NB_{\tau-1}B_j)(A_{\tau-1}A_{j-1} - NB_{\tau-1}B_{j-1}) \\ & & +N(A_{\tau-1}B_j - B_{\tau-1}A_j)(A_{\tau-1}B_{j-1} - B_{\tau-1}A_{j-1}) \\ & = & -(A^2_{\tau-1} - NB^2_{\tau-1})(A_jA_{j-1} - NB_jB_{j-1}) \end{array}$$

that is, $\Omega_{\tau-j-1} = (-1)^{\tau+1}\Omega_j$. □

The two quadratic forms $f_n(x,y) = \Delta_n x^2 + 2\Omega_n xy + \Delta_{n-1}y^2$ and $f_{\tau-1-n}(x,y) = \Delta_{n-1}x^2 - 2\Omega_n xy + \Delta_n y^2$ are associated respectively to the positions n and $\tau-1-n$, as a consequence of the symmetries of the sequences Δ and Ω shown by Theorems 1 and 2, within the first block of length τ in Υ. It should be noted that $f_m(x,y)$ and $f_{\tau-1-m}(x,y)$ are improperly equivalent.

Key matrix. Clearly, the column vectors Λ_m and $\Lambda_{\tau-m-1}$ are transformed one into the other by an involutory matrix J of determinant 1

$$\begin{bmatrix} \Delta_{m-1} \\ -2\Omega_m \\ \Delta_m \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta_m \\ 2\Omega_m \\ \Delta_{m-1} \end{bmatrix} .$$

Using the matrices $T(a_n)$ and equation (20), and applying to Λ_m the sequence of matrices $T(a_{m+1}), T(a_{m+2}), \ldots, T(a_{\tau-1-m})$ in reverse order, we obtain $\Lambda_{\tau-1-m}$

$$\Lambda_{\tau-1-m} = T(a_{\tau-1-m})\cdots T(a_{m+1})\Lambda_m \quad \Rightarrow \quad \Lambda_m = JT(a_{\tau-1-m})\cdots T(a_{m+1})\Lambda_m \ . \tag{25}$$

Assuming τ is even, this equation implies that Λ_m is an eigenvector of eigenvalue 1 of the matrix

$$E_m = JT(a_{\tau-1-m})\cdots T(a_{m+1}) = JT(a_{m+1})T(a_m)\cdots T(a_{\frac{\tau}{2}-1})T(a_{\frac{\tau}{2}})T(a_{\frac{\tau}{2}+1})\cdots T(a_{m+1})$$

since $T(a_{\tau-1-n}) = T(a_{n+1})$ by the symmetry of the sequence $\{a_n\}_{n=1}^{\tau-1}$. Observing that $JT(a_m)J = T(a_m)^{-1}$ and $J^2 = I$, we have

$$\begin{aligned} E_m &= (JT(a_{n+2})J)(JT(a_{n+3})J)J\cdots\cdots(JT(a_{\frac{\tau}{2}-1})J)JT(a_{\frac{\tau}{2}})T(a_{\frac{\tau}{2}-1})\cdots T(a_{n+2}) \\ &= T(a_{n+2})^{-1}\cdots T(a_{\frac{\tau}{2}-1})^{-1}JT(a_{\frac{\tau}{2}})T(a_{\frac{\tau}{2}-1})\cdots T(a_{n+2}) \\ &= (T(a_{\frac{\tau}{2}-1})\cdots T(a_{n+2}))^{-1}JT(a_{\frac{\tau}{2}})(T(a_{\frac{\tau}{2}-1})\cdots T(a_{n+2})) \ . \end{aligned} \tag{26}$$

It follows that the matrix E_m has the same characteristic polynomial $z^3 - z^2 - z + 1$ as $JT(a_{\frac{\tau}{2}})$, i.e. E_m has eigenvalue -1 with multiplicity 1, and eigenvalue 1 with geometric multiplicity 2.

Assuming τ is odd, the symmetries of the sequences $\{a_n\}_{n=1}^{\tau-1}$, $\{\Delta_n\}_{n=1}^{\tau-1}$, and $\{\Omega_n\}_{n=1}^{\tau-1}$, refer to an even number $\tau-1$ of terms, and equation (26) is written as

$$\begin{aligned} D_n &= (JT(a_{n+2})J)(JT(a_{n+3})J)J\cdots\cdots(JT(a_{\frac{\tau-3}{2}})J)JT(a_{\frac{\tau-3}{2}})\cdots T(a_{n+2}) \\ &= T(a_{n+2})^{-1}\cdots T(a_{\frac{\tau-3}{2}})^{-1}JT(a_{\frac{\tau-3}{2}})\cdots T(a_{n+2}) \\ &= (T(a_{\frac{\tau-3}{2}})\cdots T(a_{n+2}))^{-1}J(T(a_{\frac{\tau-3}{2}})\cdots T(a_{n+2})) \ . \end{aligned} \tag{27}$$

It follows that the matrix D_n has the same characteristic polynomial $z^3 + z^2 - z - 1$ of J, i.e. D_n has eigenvalue 1 with multiplicity 1, and eigenvalue -1 with geometric multiplicity 2.

An example may clarify the method.

EXAMPLE 1. Consider the continued fraction expansion of $\sqrt{386}$, which has period $\tau = 12$

$$[[19],[1,1,1,4,1,18,1,4,1,1,1,38]]$$

Consider the vector $\Lambda_3 = [7,-30,-23]$, since $\tau - 1 - 3 = 8$ the vector Λ_8 by symmetry is $[-23,30,7]$, i.e. $\Lambda_8 = J\Lambda_3$. However, Λ_8 may be obtained by multiplying Λ_3 by a convenient sequence of matrices

$$T(a) = \begin{bmatrix} a^2 & a & 1 \\ 2a & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\Lambda_8 = T(4)T(1)T(18)T(1)T(4)\Lambda_3$$

Since $\Lambda_3 = J\Lambda_8$, we have the equation $\Lambda_3 = JT(4)T(1)T(18)T(1)T(4)\Lambda_3$, that is

$$\Lambda_3 = \begin{bmatrix} 9801 & 1980 & 400 \\ -97020 & -19601 & -3960 \\ 240100 & 48510 & 9801 \end{bmatrix} \Lambda_3 \quad \Rightarrow \quad \Lambda_3 = E_3\Lambda_3 \ ,$$

i.e. Λ_3 is an eigenvector of E_3 for the eigenvalue 1.
The characteristic polynomial of E_3 is found to be $Z^3 - Z^2 - Z + 1 = (Z+1)(Z-1)^2$ which is the same of the matrix $JT(a_6)$, with

$$T_{\frac{\tau}{2}} = T(18) = \begin{bmatrix} 324 & 18 & 1 \\ 36 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad ;$$

note that $\frac{\tau}{2} = 6$, and in position 5 we find the vector $\Lambda_5 = [2,-36,-31]$ whose first entry gives the factor 2 of 386.

THEOREM 3. *The correspondence $m \leftrightarrow \Lambda_m$ is one-to-one for $1 \le m \le \tau$, i.e. all quadratic forms $f_m(x,y)$ within a period are distinct.*

Proof. The proof is by contradiction. Suppose, contrary to the theorem's claim, that $\Lambda_{n_1} = \Lambda_{n_2} = X$ for some $n_1 < n_2$, then equation (20) implies the existence of a matrix $P_{n_2n_1} = \prod_{j=n_1+1}^{n_2} T(a_j)$ such that $\Lambda_{n_2} = P_{n_2n_1}\Lambda_{n_1}$. Thus X must be an eigenvector, for the eigenvalue 1, of the non-negative (positive whenever $n_2 - n_1 \ge 2$) matrix $P_{n_2n_1}$ which is the product of non-negative matrices.
If $n_2 = n_1 + 1$, it is direct to compute the characteristic polynomial $p(x)$ of $P_{n_2n_1} = T(a_{n_1})$

$$p(x) = x^3 - (a_{n_1}^2 + 1)x^2 - (a_{n_1}^2 + 1)x + 1 \ ,$$

which is a 3-degree reciprocal polynomial which has a single root -1, and the remaining roots are certainly different from 1, because $a_{n_1} \neq 0$; thus, in this case, X cannot exist.

To prove in general that X does not exist, we observe that any $P_{n_2n_1}$ has a reciprocal characteristic polynomial $q(x)$ of degree 3, because we have

$$q(x)=\det\left(\lambda I_3-\prod_{j=n_1+1}^{n_2}T(a_j)\right)=\det\left(\lambda I_3-J\prod_{j=n_1+1}^{n_2}T(a_j)J\right)=\det\left(\lambda I_3-\prod_{j=n_1+1}^{n_2}T(a_j)^{-1}\right),$$

$$q(x)=\det\left(\lambda I_3-\prod_{j=n_1+1}^{n_2}T(a_j)\right)=\det\left(\lambda I_3-\left(\prod_{j=n_1+1}^{n_2}T(a_j)\right)^{-1}\right),$$

where the last equality is justified by [9, Theorem 1.3.20, p.53]. The reciprocal polynomial $q(x)$ has an eigenvalue equal to either -1 or 1. If the eigenvalue is -1, which occurs when n_2-n_1 is odd, the eigenvector X does not exist. If the eigenvalue is 1, which occurs when n_2-n_1 is even, there is a second eigenvector for the same eigenvalue, because we have

$$J\Lambda_{n_2}=JP_{n_2n_1}\Lambda_{n_1}=JP_{n_2n_1}J\cdot J\Lambda_{n_1}=\left(\prod_{j=n_1+1}^{n_2}T(a_j)\right)^{-1}J\Lambda_{n_1}\ \Rightarrow\ \left(\prod_{j=n_1+1}^{n_2}T(a_j)\right)J\Lambda_{n_2}=J\Lambda_{n_2}.$$

Then, X and JX should be distinct eigenvectors (because $\Omega_{n_2}\neq 0$ for every n_2) of the same eigenvalue 1 of multiplicity one, which is impossible.
In conclusion, the eigenvector X of eigenvalue 1 does not exist, so $m\leftrightarrow\Lambda_m^T$ is a one-to-one mapping within each period. □

3.2. Odd period

In [11, p.59-60], Legendre describes a constructive method for computing the representation of a positive (square-free) N as the sum of two squares, by means of the continued fraction expansion of $\sqrt{N}$. This result is stated as a theorem with a different proof from that of Legendre [11, p.60].

THEOREM 4. *Let N be a positive integer such that the continued fraction expansion of $\sqrt{N}$ has odd period τ. The representation of $N=x^2+y^2$ is given by $x=\Delta_{\frac{\tau-1}{2}}$ and $y=\Omega_{\frac{\tau-1}{2}}$.*

PROOF. Since τ is odd, by the anti-symmetry in the sequence $\{\Delta_n\}_{n=0}^{\tau-2}$, we have $\Delta_{\frac{\tau-1}{2}-1}=-\Delta_{\frac{\tau-1}{2}}$, so that the quadratic form $\Delta_{\frac{\tau-1}{2}}X^2+2\Omega_{\frac{\tau-1}{2}}XY+\Delta_{\frac{\tau-3}{2}}Y^2$ has discriminant $4\Delta^2_{\frac{\tau-1}{2}}+4\Omega^2_{\frac{\tau-1}{2}}=4N$, which shows the assertion. □

3.3. Even period

Let N be a square-free composite integer such that the continued fraction of $\sqrt{N}$ has even period. We say that $\mathfrak{c}_{\tau-1}=A_{\tau-1}+B_{\tau-1}\sqrt{N}$ splits N whenever $A_{\tau-1}+1$ and $A_{\tau-1}-1$ are divisible by proper factors, say m_1 and m_2, of $N=m_1m_2$, respectively.

LEMMA 1. *If the period τ of the continued fraction expansion of $\sqrt{N}$ is even, we have*

$$\Delta_\tau = \Delta_{\tau-2} \quad \textit{and} \quad \Omega_\tau = -\Omega_{\tau-1}$$

with $\Omega_{\tau-1} = -a_0$.

Proof. Since $\Delta_{\tau-1} = 1$, we have $\Omega_{\tau-1}^2 - \Delta_{\tau-2} = N$, thus $\Omega_{\tau-1} = -\sqrt{N+\Delta_{\tau-2}}$ because $\tau-1$ is odd. Considering the Taylor series around the origin for the square root, we have

$$\Omega_{\tau-1} = -\sqrt{N+\Delta_{\tau-2}} = -\sqrt{N}\left(1 - \frac{\Delta_{\tau-2}}{2N} + \frac{\Delta_{\tau-2}^2}{8N^2} + \cdots\right) = -\left\lfloor \sqrt{N} \right\rfloor = -a_0 \quad .$$

Using equation (17) with $m = \tau - 1$ we have

$$\Delta_\tau = \Delta_{\tau-2} + a_\tau \left(2\Omega_{\tau-1} + a_\tau \Delta_{\tau-1}\right) = \Delta_{\tau-2} \quad .$$

Thus, equation (21) finally gives $\Omega_\tau = -\Omega_{\tau-1}$. □

LEMMA 2. *Let τ be even, and define the integer $\gamma \in \mathfrak{O}_{\mathbb{F}}$ by the product*

$$\gamma = \prod_{m=0}^{\tau-1} \left(\sqrt{N} + (-1)^m \Omega_m\right) \quad ,$$

then $\frac{\gamma}{\sigma(\gamma)} = \left(A_{\tau-1} + B_{\tau-1}\sqrt{N}\right)^2 = \mathfrak{c}_{\tau-1}^2$.

Proof. The norm of $\frac{\gamma}{\sigma(\gamma)}$ is patently 1, thus it remains to prove that $\frac{\gamma}{\sigma(\gamma)}$ lies in $\mathfrak{O}_{\mathbb{F}}$. We have

$$\frac{\gamma}{\sigma(\gamma)} = \prod_{m=0}^{\tau-1} \frac{\sqrt{N} + (-1)^m \Omega_m}{-\sqrt{N} + (-1)^m \Omega_m} = \prod_{m=0}^{\tau-1} \frac{(\sqrt{N} + (-1)^m \Omega_m)^2}{\Omega_m^2 - N} = \prod_{m=0}^{\tau-1} \frac{(\sqrt{N} + (-1)^m \Omega_m)^2}{\Delta_m \Delta_{m-1}} \quad .$$

Observing that $\prod_{m=0}^{\tau-1}(\Delta_m \Delta_{m-1}) = \prod_{m=0}^{\tau-1} \Delta_m^2$ by the periodicity of the sequence $\{\Delta_m\}_m$, it follows that $\frac{\gamma}{\sigma(\gamma)}$ is a perfect square. Considering the following identity

$$\frac{\sqrt{N} + (-1)^m \Omega_m}{\Delta_m} = (-1)^m \frac{A_{m-1} - B_{m-1}\sqrt{N}}{A_m - B_m\sqrt{N}} \quad ,$$

we have that the base of the square giving $\frac{\gamma}{\sigma(\gamma)}$ is

$$\prod_{m=0}^{\tau-1} \frac{(\sqrt{N} + (-1)^m \Omega_m)}{\Delta_m} = \prod_{m=0}^{\tau-1} (-1)^m \frac{A_{m-1} - B_{m-1}\sqrt{N}}{A_m - B_m\sqrt{N}} = (-1)^{\frac{\tau}{2}} \frac{A_{-1} - B_{-1}\sqrt{N}}{A_{\tau-1} - B_{\tau-1}\sqrt{N}} \quad . \quad .$$

Now, $A_{-1} = 1$ and $B_{-1} = 0$ by definition, thus

$$(28) \qquad \prod_{m=0}^{\tau-1} \frac{(\sqrt{N} + (-1)^m \Omega_m)}{\Delta_m} = (-1)^{\frac{\tau}{2}} (A_{\tau-1} + B_{\tau-1}\sqrt{N}) = (-1)^{\frac{\tau}{2}} \mathfrak{c}_{\tau-1} \quad ,$$

and in conclusion $\frac{\gamma}{\sigma(\gamma)} = \mathfrak{c}_{\tau-1}^2$, which shows the claimed property. □

The close connection between the continued fraction expansion of $\sqrt{N}$ and the factorization of N is proved using the matrix $M_{\tau-1}$ defined in equation (23). Note that the matrix $M_{\tau-1}$ is involutory, or neg-involutory, since its square is either plus or minus the identity matrix I_2, i.e. $M_{\tau-1}^2=(-1)^{\tau}I_2$. If τ is even, the eigenvalues of matrix $M_{\tau-1}$ are ± 1, and $M_{\tau-1}$ is involutory. If τ is odd, the eigenvalues are $\pm i$, and $M_{\tau-1}$ is neg-involutory.

THEOREM 5. *If the period τ of the continued fraction expansion of $\sqrt{N}$ is even, the element $\mathfrak{c}_{\tau-1}$ in $\mathbb{Q}(\sqrt{N})$ splits $2N$, and a factor of $2N$ is located at positions $\frac{\tau-2}{2}+j\tau$, $j=0,1,\ldots$, in the sequence $\Delta=\{\mathfrak{c}_m\sigma(\mathfrak{c}_m)\}_{m\geq 1}$.*

Proof. It is sufficient to consider $j=0$, due to the periodicity of Δ. Since τ is even, $M_{\tau-1}$ is involutory and has eigenvalues ± 1 with corresponding eigenvectors

$$X^{(h)}=\left[\frac{A_{\tau-1}-(-1)^h}{d},\frac{B_{\tau-1}}{d}\right]^T \quad \text{with} \quad d=\gcd\{A_{\tau-1}-(-1)^h,B_{\tau-1}\} \quad h=0,1 \ .$$

Considering equation (22) written as

$$\begin{bmatrix} A_{\tau-j-2} \\ B_{\tau-j-2} \end{bmatrix}=(-1)^{j-1}M_{\tau-1}\begin{bmatrix} A_j \\ B_j \end{bmatrix} \ ,$$

we see that $Y^{(j)}=[A_j,B_j]^T$ is an eigenvector of $M_{\tau-1}$, of eigenvalue $(-1)^{j-1}$ if and only if j satisfies the condition $\tau-j-2=j$, that is $j=\frac{\tau-2}{2}=\tau_0$. From the comparison of $X^{(h)}$ and $Y^{(\tau_0)}$, we have

$$A_{\tau_0}=\frac{A_{\tau-1}-(-1)^{\tau_0-1}}{d} \qquad B_{\tau_0}=\frac{B_{\tau-1}}{d} \ , \tag{29}$$

where the equalities are fully motivated because $\gcd\{A_{\tau_0},\ B_{\tau_0}\}=1$. Direct computation yields

$$\Delta_{\tau_0}=\frac{(A_{\tau-1}-(-1)^{\tau_0-1})^2-NB_{\tau-1}^2}{d^2}=2\frac{(-1)^{\tau_0-1}A_{\tau-1}+1}{d^2} \ , \tag{30}$$

which can be written as $A_{\tau_0}^2-NB_{\tau_0}^2=2(-1)^{\tau_0-1}\frac{A_{\tau_0}}{d}$; dividing this equality by $2\frac{A_{\tau_0}}{d}$ we have

$$\frac{dA_{\tau_0}}{2}-N\frac{1}{\frac{2A_{\tau_0}}{d}}B_{\tau_0}^2=(-1)^{\tau_0-1} \ .$$

Noting that $\gcd\{A_{\tau_0},\ B_{\tau_0}\}=1$, it follows that $\frac{2A_{\tau_0}}{d}$ is certainly a divisor of $2N$, i.e. $\Delta_{\tau_0}|2N$. □

EXAMPLE 2. Consider $N=3\cdot 5\cdot 7\cdot 11\cdot 19=21945$, the period of the continued fraction of $\sqrt{21945}$ is 10, and is fully shown in the following table for the sequences Δ and Ω

j	Δ_j	Ω_j
-1	1	
0	-41	148
1	64	-139
2	-129	117
3	16	-141
4	-21	147
5	16	-147
6	-129	141
7	64	-117
8	-41	139
9	1	-148
10	-41	148

In position $j = \frac{\tau-2}{2} = 4$ we find 21, a factor of N, as expected. The same factor 21 can be found by considering the fundamental unit $\epsilon_9 = 3004586089 + 20282284\sqrt{21945}$, in fact we have $3004586089 - 1 = 2^3 \cdot (3 \cdot 7) \cdot 4229^2$, and the second factor $5 \cdot 11 \cdot 19$ may be obtained from $3004586089 + 1 = 2 \cdot (5 \cdot 11^3 \cdot 19) \cdot 109^2$.

In principle, in many cases the above Theorem 5 yields a factor of N; however there are examples in which only the factor 2 appears.

EXAMPLE 3. Let $N = 8527 \times 8537 = 72794999$ be a composite number. The period of $\sqrt{N}$ is $\tau = 3864$ and in position 1931 we do not find a factor of N but $\Delta_{1931} = 2$ which is a factor of $2N$.

It would be interesting to find a general condition that can discriminate the various situations, i.e. whether a factor of N is found or not. This objective can be achieved almost in full when $N = pq$ is the product of two primes, a case that cleverly shows the difficulty of the whole problem.

3.4. Factoring $N = pq$

When $N = pq$ is the product of two distinct primes, the analysis of section 3.3 may be further pursued, leading to the following remarkable property:

PROPOSITION 2. *If $p \equiv q \equiv 3 \bmod 4$, the fundamental unit ε_0 (or the cube ε_0^3) splits $N = pq$, then $\Delta_{\frac{\tau-2}{2}}$ is equal to $(q|p)p$, with $p < q$.*

This proposition is given without the proof, which uses units and splitting of primes in quadratic number fields (see [6, 4, 10]); further, the complete classification in terms of residues of p and q modulo 8, proved in [6],. is reported in Table 7.1 for easy reference.

4. Factorization

Gauss recognized that the factoring problem was important, although very difficult,

> *... Problema, numeros primos a compositis dignoscendi, hosque in factores suos primos resolvendi, ad gravissima ac utilissima totius arithmeticae pertinere, et geometrarum tum veterum tum recentiorum industriam ac sagacitatem occupavisse, tam notum est, ut de hac re copiose loqui superfluum foret. ...* C. F. GAUSS [*Disquisitiones Arithmeticae* ART. 329]

and, in spite of much effort, various different approaches, and the problem's increased importance due to the large number of cryptographic applications, no satisfactorily factoring method has yet been found.

Many factorizations make use of the regular continued fraction expansion of $\sqrt{N}$, combined with the idea of using quadratic forms [7, 13]. The infrastructure method, proposed by Shanks [15], considers the subset $\Psi = \{f_m(x,y)\}_{1\leq m\leq \tau-1}$ in the periodic sequence $\Upsilon = \{f_m(x,y)\}_{m\geq 1}^{\infty}$ of reduced principal quadratic forms. It should be remarked that the forms $f_m(x,y) = \Delta_m x^2 + 2\Omega_m xy + \Delta_{m-1}y^2$ in Υ are reduced following a different convention from that commonly adopted [1].

DEFINITION 2. *A real quadratic form $f(x,y) = ax^2+2bxy+cy^2$ of discriminant $4N$ is said to be reduced if, defining $\kappa = \min\{|a|,|c|\}$, b is the sole integer such that $\sqrt{N}-|b| < \kappa < \sqrt{N}+|b|$, with the sign of b chosen opposite to the sign of a.*

DEFINITION 3. *The distance between $f_{m+1}(x,y)$ and $f_m(x,y)$ is defined to be*

$$d(f_{m+1},f_m) = \frac{1}{2}\ln\left(\frac{\sqrt{N}+(-1)^m\Omega_m}{\sqrt{N}-(-1)^m\Omega_m}\right) \quad . \tag{31}$$

The distance between two quadratic forms $f_m(x,y)$ and $f_n(x,y)$, with $m > n$, is defined to be the sum

$$d(f_m,f_n) = \sum_{j=n}^{m-1} d(f_{j+1},f_j) \quad . \tag{32}$$

Taking the above definitions, Shanks showed that, by the Gauss composition law of quadratic forms with the same determinant, followed by reduction, the set Ψ equipped with the distance $d(f_{m+1},f_m)$ modulo $R = \ln \mathfrak{c}_{\tau-1}$ resembles a cyclic group, with $f_{\tau-1}(x,y)$ playing the role of identity. Composition followed by reduction affords big steps (giant steps) within Ψ, thus two operators were further defined [3, p.259] to allow small steps (baby steps), precisely

1. One-step forward: The operator ρ^+ that transforms one reduced quadratic form into the next in the sequence Υ, is defined as

$$\rho^+([a,2b,c]) = [\frac{b_1^2-N}{a}, 2b_1, a] \quad ,$$

where b_1 is $2b_1 = [2b \bmod (2a)] + 2ka$ with k chosen in such a way that $-|a| < b_1 < |a|$.

2. One-step backward: The operator ρ^- that transforms a reduced quadratic form into the immediately preceding quadratic form in the sequence Υ is defined as

$$\rho^-([a,2b,c]) = [c,2b_1,\frac{b_1^2 - N}{c}] \ ,$$

where b_1 is $2b_1 = [-2b \bmod (2c)] + 2kc$ with k chosen such that $-|c| < b_1 < |c|$.

The infrastructure machinery was used to compute the fundamental unit, the regulator, and the class number [3], with complexity smaller than $O(\sqrt{N})$, although not of polynomial complexity in N. From a different perspective, by Theorem 5, in many cases a factor of N is exactly positioned in the middle of a period of the sequence Δ. Therefore, instead of trying to find special quadratic forms randomly located in Ψ (the principal genus), or some ambigue form in some non-principal genus, we may try to localize the position of some factor of N within a period whose length τ is unknown. Then, it is shown that, by extending the infrastructure machinery to the whole sequence Υ, some factors of N can be computed with a complexity substantially bounded by the complexity required to evaluate an integral of Dirichlet's at a given accuracy: the more precise the evaluation of the integral, the less complex the factorization; at the limit, it is of polynomial complexity; clearly, to be more accurate in the integral evaluation, greater complexity is required. To pursue this idea, we briefly review and adapt the previous definitions of the infrastructure components to the new task. Let us recall that the quadratic forms $f_m(x,y)$ are primitive, i.e. $\gcd\{\Delta_m, 2\Omega_m, \Delta_{m-1}\} = 1$, and at least one between $|\Delta_m|$ and $|\Delta_{m-1}|$ is less than $\sqrt{N}$ and $0 < |\Omega_m| < \sqrt{N}$. Further, since $\mathfrak{c}_{\tau-1}$ is either equal to the positive fundamental unit of $\mathbb{F} = \mathbb{Q}(\sqrt{N})$ or equal to its cube, the regulator of $\mathfrak{O}_\mathbb{F}$ is either $R_\mathbb{F} = \ln \mathfrak{c}_{\tau-1}$, or $R_\mathbb{F} = \frac{1}{3} \ln \mathfrak{c}_{\tau-1}$. The following observations are instrumental to motivate the procedure:

1. The sign of Δ_{m-1} is the same as that of Ω_m, which is opposite to that of Δ_m, thus in the sequence Υ the two triplets of signs $(-,+,+)$ and $(+,-,-)$ alternate.

2. The distance of $f_m(x,y)$ from the beginning of Υ is defined by referring to a properly selected hypothetical quadratic form, i.e. $f_0(x,y) = f_\tau(x,y) = f_0(x,y) = \Delta_0 x^2 - 2\sqrt{N-\Delta_0}xy + y^2$, which is located before $f_1(x,y)$, that is $d(f_m, f_0)$ is given by (32) if $m < \tau$, and by $d(f_m, f_0) = d(f_{m \bmod \tau}, f_0) + kR_\mathbb{F}$ if $k\tau \le m < (k+1)\tau$.

3. Let "$\bullet$" denote the form composition $f_m(x,y) \bullet f_n(x,y)$ in Υ, that is the Gauss composition [3] of $f_m(x,y)$ and $f_n(x,y)$ followed by a reduction performed with the minimum number of steps, ending with a reduced form whose triplet of signs is $(-,+,+)$ if m and n have the same parity, and $(+,-,-)$ otherwise. This distance defined by (31) holds in Υ with good approximation, and is compatible with the "$\bullet$" operation, that is we have

$$f_{\ell(m,n)}(x,y) = f_m(x,y) \bullet f_n(x,y) \Rightarrow d(f_{\ell(m,n)}, f_0) \approx d(f_m, f_0) + d(f_n, f_0) \ .$$

It is remarked that the error affecting this distance estimation is of order $O(\ln N)$ as shown by Schoof in [14].

4. Shanks [15] observed that, within the first period, the composition law "$\bullet$" induces a structure similar to a cyclic group for the addition of distances modulo the "regulator".

5. Between the elements of Υ the distance is nearly maintained by the giant steps, and is rigorously maintained by the baby steps.

THEOREM 6. *The distance $d(f_\tau, f_0)$ is exactly equal to $\ln \mathfrak{c}_{\tau-1}$, i.e. this distance $d(f_\tau, f_0)$ is either the regulator $R_{\mathbb{F}}$ or $3R_{\mathbb{F}}$. The distance $d(f_{\frac{\tau}{2}}, f_0)$ is exactly equal to $\frac{1}{2}d(f_\tau, f_0)$.*

Proof. The distance between f_τ and f_0 is the summation

$$d(f_\tau, f_0) = \sum_{j=0}^{\tau-1} d(f_{j+1}, f_j) = \sum_{j=0}^{\tau-1} \frac{1}{2}\ln\left(\sum_{j=0}^{\tau-1} \frac{\sqrt{N}+(-1)^j\Omega_j}{\sqrt{N}-(-1)^j\Omega_j}\right) = \frac{1}{2}\ln\left(\prod_{j=0}^{\tau-1} \frac{\sqrt{N}+(-1)^j\Omega_j}{\sqrt{N}-(-1)^j\Omega_j}\right) .$$

Recalling that $N - \Omega_j^2 = -\Delta_j\Delta_{j-1} > 0$, and taking into account the periodicity of the sequence Δ, the last expression can be written with rational denominator as

$$\frac{1}{2}\ln\left(\prod_{j=0}^{\tau-1} \frac{(\sqrt{N}+(-1)^j\Omega_j)^2}{-\Delta_j\Delta_{j-1}}\right) = \frac{1}{2}\ln\left(\prod_{j=0}^{\tau-1} \frac{(\sqrt{N}+(-1)^j\Omega_j)^2}{\Delta_j^2}\right) = \ln\left(\prod_{j=0}^{\tau-1} \frac{\sqrt{N}+(-1)^j\Omega_j}{(-1)^{j-1}\Delta_j}\right) .$$

The conclusion follows from Lemma 2, showing that the product $\prod_{j=0}^{\tau-1} \frac{\sqrt{N}+(-1)^j\Omega_j}{(-1)^{j-1}\Delta_j}$, which has field norm one and is an element of the order $\mathfrak{O}_{\mathbb{F}}$, is actually the unit $\mathfrak{c}_{\tau-1}$ by equation (28). The connection between $\ln \mathfrak{c}_{\tau-1}$ and the regulator is motivated by Remark 1.
The equality $d(f_{\frac{\tau}{2}}, f_0) = \frac{1}{2}d(f_\tau, f_0)$ is an immediate consequence of the symmetry of the sequence $f_m(x,y)$ within a period. □

Since Theorem 5 guarantees that, when τ is even, a factor of N is located in the positions $\frac{\tau-2}{2} + k\tau$ of the sequence Υ, Shanks' method allows us to find such a factor, if $\ln(\mathfrak{c}_{\tau-1})$, or an odd multiple of it, is exactly known. Now, a formula of Dirichlet's gives the product

$$h_{\mathbb{F}}R_{\mathbb{F}} = \frac{\sqrt{D}}{2}L(1,\chi) = -\sum_{n=1}^{\lfloor \frac{D-1}{2} \rfloor} \left(\frac{D}{n}\right) \ln\left(\sin\frac{n\pi}{D}\right) \tag{33}$$

where $R_{\mathbb{F}}$ is the regulator, $L(1,\chi)$ is a Dedekind L-function, $D = N$ if $N \equiv 1 \bmod 4$ or $D = 4N$ otherwise, and character χ is the Jacobi symbol in this case. If the product $h_{\mathbb{F}}R_{\mathbb{F}}$ is known exactly (computed), for example using equation (33), the distance from the beginning of the sequence where the quadratic form can be found $[1, 2\Omega_{\tau-1}, \Delta_{\tau-2}]$ is known. Since this distance is an integer multiple of the regulator, and our target is to find a quadratic form that is located in the middle of some period, then

1. if $h_{\mathbb{F}}$ is odd, a factor of N is found in the position at distance $\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{2}$, or $3\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{2}$, from the beginning;

2. If $h_{\mathbb{F}}$ is even, in a position at distance $\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{2}$, or $3\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{2}$ the quadratic form $[1, 2\Omega_{\tau-1}, \Delta_{\tau-2}]$ is found, (which reveals a posteriori that $h_{\mathbb{F}}$ is even); in this case, the procedure can be repeated with target the position at distance $\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{4}$, or $3\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{4}$; again, either a factor of N is found or $h_{\mathbb{F}}$ is found to be a multiple of 4. Clearly the process can be iterated ℓ times until $\frac{h_{\mathbb{F}}R_{\mathbb{F}}}{2^\ell}$ is an odd multiple of $R_{\mathbb{F}}$, and a factor of N is found.

When the factor m_1 of N is found, the second factor is $m_2 = \frac{N}{m_1}$, thus the procedure can be iterated to find all factors of N. Mimicking Shanks' infrastructure, giant steps are performed to get close to forms at distance $\frac{kR_{\mathbb{F}}}{2}$, or $3\frac{kR_{\mathbb{F}}}{2}$, for some $1 \leq k \leq h_{\mathbb{F}}$, then baby steps are performed to get the exact position.

5. Conclusions

It has been shown that the complexity of factoring a composite number $4N$ is upper bounded by the complexity of evaluating, at a certain degree of accuracy, the product $h_{\mathbb{F}}R_{\mathbb{F}}$, as defined by Dirichlet using the $L(1, \chi_N)$ function, and also that is not necessary to know $h_{\mathbb{F}}$ and $R_{\mathbb{F}}$ separately. The more precise the evaluation of the product $h_{\mathbb{F}}R_{\mathbb{F}}$, the less complex the factoring $2N$; if we are lucky, the complexity could be polynomial in N. It is an open problem to find which is the best compromise between the approximate evaluation of $h_{\mathbb{F}}R_{\mathbb{F}}$ and the computational complexity for obtaining such approximation. In this context, the following expression, taken from [3, p.262], may be useful for efficiently evaluating the product $h_{\mathbb{F}}R_{\mathbb{F}}$ as a function of N

$$h_{\mathbb{F}}R_{\mathbb{F}} = \frac{1}{2}\sum_{x\geq 1}\left(\frac{N}{x}\right)\left(\frac{\sqrt{N}}{x}\mathrm{erfc}\left(x\sqrt{\frac{\pi}{N}}\right) + E_1\left(\frac{\pi x^2}{N}\right)\right) \quad , \tag{34}$$

where the complementary error function $\mathrm{erfc}(x)$, and the exponential integral function $E_1(x)$, can be closely approximated [18, p.297-299]

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-t^2}dt = 1 - \mathrm{erf}(z) = 1 - \frac{2}{\sqrt{\pi}}\sum_{n=0}^\infty \frac{(-1)^n z^{2n+1}}{n!(2n+1)}$$

$$E_1(z) = \int_1^\infty \frac{e^{-tz}}{t}dt = -\gamma - ln(z) - \sum_{n=1}^\infty \frac{(-1)^n z^n}{n \cdot n!} \quad .$$

As a last observation, the arguably, a fast (how fast is open) algorithm for factoring is achievable by combining results of Dirichlet, Shanks, and the above observations, which were suggested by Legendre's finding that continued fractions permit the representation of primes as the sum of two squares explicitly computed.

Acknowledgement. The very useful and constructive suggestions of the unknown referee, who pointed out several misprints, incompleteness, and provided the Example 3 are gratefully acknowledged. I also thank Karan Khathuria and Simran Tinani, PhD students at the University of Zurich, for their careful reading of a preliminary version of the paper, and for pointing out several misprints, errors, and imprecisions.

References

[1] D.A. Buell, *Binary Quadratic Forms*, New York: Springer-Verlag, 1989.

[2] G.S. Carr, *Formulas and Theorems in Mathematics*, Chelsea: New York, 1970.

[3] H. Cohen, *A Course in Computational Algebraic Number Theory*, New York: Springer-Verlag, 1993.

[4] **H. Cohn**, *A dvanced in Number Theory*, New York: Dover, 1980.

[5] H. Davenport, *The Higher Arithmetic*, New York: Dover, 1983.

[6] M. Elia, Relative Densities of Ramified Primes in $\mathbb{Q}(\sqrt{pq})$, *International Mathematical Forum*, vol. 3, n.8, 2008, p.375-384.

[7] C.F. Gauss, *Disquisitiones Arithmeticae*, New York: Springer-Verlag, 1986.

[8] G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers*, Oxford: Clarendon Press, 1971.

[9] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge: Cambridge Univ. Press, 1999.

[10] Hua Loo Keng, *Introduction to Number Theory*, New York: Springer, 1981.

[11] A-M. Legendre, *Essai sur la Théorie des Nombres*, Chez Courcier, Paris, 1808, reissued by Cambridge University Press, 2009.

[12] O. Perron, *Die Lehre von den Kettenbrüchen, BandI:Elementare Kettenbrüche*, Springer, 1977.

[13] H. Riesel, *Prime Numbers and Computer Methods for Factorization*, Boston: Birkhäuser, 1984.

[14] R. Schoof, Quadratic Fields and Factorization, *Computational methods in number theory*, Mathematical Centre Tracts 154, Amsterdam, (1982), p.235-286. (p 129 of .pdf file)

[15] D. Shanks, The infrastructure of a real quadratic field and its applications, *Proc. 1972 Number Theory Conference*, Boulder (1972), p.217-224. New York, 1985.

[16] W. Sierpinski, *Elementary Theory of Numbers*, North Holland, New York, 1988.

[17] J. Steuding, *Diophantine Analysis*, Chapman & Hall, New York, 2003.

[18] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1968.

AMS Subject Classification: 11A55, 11A51

M. Elia

Lavoro pervenuto in redazione il 19.10.2019.

p mod 8	q mod 8	Split?	$(p\|q)$	$\Delta_{\tau/2-1}$	T mod 4
3	3	Yes	±1	$-(p\|q)p$	$1+(p\|q)$
3	7	Yes	±1	$-(p\|q)p$	$1+(p\|q)$
7	3	Yes	±1	$-(p\|q)p$	$1+(p\|q)$
7	7	Yes	±1	$-(p\|q)p$	$1+(p\|q)$
5	3	Yes	1	p	0
3	5	Yes	1	$-p$	2
5	3	Yes	-1	$2p$	0
3	5	Yes	-1	$-2p$	2
5	7	Yes	1	p	0
7	5	Yes	1	$-p$	2
5	7	Yes	-1	$-2p$	2
7	5	Yes	-1	$2p$	0
1	3	No	-1	-2	2
1	3	Yes	1	p	AND 0
1	3	No/Yes	1	$-2,-2p$	2
3	1	No	-1	-2	2
3	1	Yes	1	$2p$	AND 0
3	1	No/Yes	1	$-2,-p$	2
7	1	No	-1	2	0
7	1	No	1	2	AND 0
7	1	Yes	1	$-p,-2p$	2
1	7	No	-1	2	0
1	7	No/Yes	1	$2,p,2p$	0
5	1	No	-1		1,3
5	1	No	1		AND 1,3
5	1	Yes	1	$-p$	AND 2
5	1	Yes	1	p	AND 0
1	5	No	-1		1,3
1	5	No	1		AND 1,3
1	5	Yes	1	$-p$	AND 2
1	5	Yes	1	p	AND 0
5	5	No	-1		1,3
5	5	No	1		AND 1,3
5	5	Yes	1	$-p$	AND 2
5	5	Yes	1	p	AND 0
1	1	No	-1		1,3
1	1	No	1		AND 1,3
1	1	Yes	1	$-p$	AND 2
1	1	Yes	1	p	AND 0

Table 7.1: $p < q$