

Continued Fraction Gems

R.A. Mollin

Dedicated to the memory of André Weil (1906–1998)

Abstract

Since the first stirrings of the notion appeared on a Babylonian tablet in 2000 B.C., continued fractions have found their way into many corners of mathematics. Some of the interrelationships are quite pretty indeed, and deserve the title of *gems*. It is the purpose of this article to take the reader from the basic notions, together with some historical data, to the modern developments, which link continued fractions with an array of mathematical notions, and display some of these gems along the way. This article is intended for anyone from the uninitiated reader to the expert, since no background in the subject is assumed.

1 Introduction

In the Berlin Museum is a Babylonian tablet from 2000 B.C. on which appears a method similar to continued fractions. This method approximates the diagonal d of a rectangle with sides $\ell \leq \omega$ by $\ell + \omega^2/\ell$. This is the first known recorded appearance of the notion of a continued fraction. In approximately 1680 B.C., a scribe named *Ahmes* recorded the approximation $\sqrt{\pi/4} \approx 8/9$ on a papyrus, which is now known as the *Rhind papyrus* after the Scottish gentleman, Henry Rhind, who purchased it in a Nile resort town in 1858. Today, we know that what Ahmes had actually found was the third convergent in the simple continued fraction expansion (described below) of $\sqrt{\pi/4}$. These early pronouncements were the first baby steps, unaware of what was to come.

The advances made by the Greeks of antiquity were major steps in the development of the theory of continued fractions. Eudexus of Cnidus (ca.

408–355 B.C.), a pupil of Plato, studied criteria for $A/B = C/D$, where $A, B, C, D \in \mathbb{R}^+$, the positive real numbers. This turned out to be equivalent to the method later developed by the Persian mathematician and poet, Omar Khayam (ca. 1048–1122 A.D.), who essentially showed that the ratios are equal if and only if they can be expanded into simple continued fractions with identical partial quotients. Of course, the most noteworthy contribution by the ancient Greeks is that of Euclid of Alexandria (ca. 365–300 B.C.), whose division algorithm calculates the greatest common divisor of two integers.

The Euclidean algorithm is a repeated application of the division algorithm, which says that if $a \in \mathbb{N} = \{1, 2, 3, 4, \dots\}$ (the natural numbers), and $b \in \mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$ (the integers), then there exist unique integers $q, r \in \mathbb{Z}$ with $0 \leq r < a$ and

$$b = aq + r.$$

If we let $a = r_0$, and $b = r_{-1}$, then by repeated application of the division algorithm, we get the recursive relation:

$$r_{j-1} = r_j q_{j+1} + r_{j+1}, \tag{1.1}$$

for all nonnegative integers $j < n$, where n is the least nonnegative integer such that $r_{n+1} = 0$. One may then conclude that the greatest common divisor of a and b is r_n , denoted $\gcd(a, b) = r_n$. For instance, if $a = 165$, and $b = 2431$, then

$$\begin{aligned} r_{-1} &= 2431 = 165 \cdot 14 + 121 = r_0 \cdot q_1 + r_1, \\ r_0 &= 165 = 121 \cdot 1 + 44 = r_1 \cdot q_2 + r_2, \\ r_1 &= 121 = 44 \cdot 2 + 33 = r_2 \cdot q_3 + r_3, \\ r_2 &= 44 = 33 \cdot 1 + 11 = r_3 \cdot q_4 + r_4, \end{aligned}$$

and

$$r_3 = 33 = 11 \cdot 3 + 0 = r_4 \cdot q_5 + r_5,$$

Thus, since $r_{n+1} = r_5 = 0$, we have $r_n = r_4 = 11 = \gcd(a, b)$. Now we show how this gives rise to continued fractions and representations of rational numbers.

In the recurrence relation (1.1), let

$$\alpha_{j-1} = r_{j-1}/r_j,$$

from which it follows that

$$\alpha_{j-1} = q_{j+1} + 1/\alpha_j, \text{ and } \alpha_{n-1} = q_{n+1}.$$

For example, in the illustration above, we get:

$$\alpha_{-1} = b/a = \frac{2431}{165} = 14 + \frac{121}{165} = 14 + \frac{1}{165/121} = q_1 + 1/\alpha_0,$$

$$\alpha_0 = \frac{165}{121} = 1 + \frac{44}{121} = 1 + \frac{1}{121/44} = q_2 + 1/\alpha_1,$$

$$\alpha_1 = \frac{121}{44} = 2 + \frac{33}{44} = 2 + \frac{1}{44/33} = q_3 + 1/\alpha_2,$$

$$\alpha_2 = \frac{44}{33} = 1 + \frac{11}{33} = 1 + \frac{1}{3},$$

so,

$$\frac{2431}{165} = 14 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{3}}}}.$$

This illustrates the fact that all rational numbers can be expressed in this way. This is formalized as follows.

If $q_j \in \mathbb{R}$ for $j = 0, 1, \dots, \ell$ where $\ell \in \mathbb{Z}$ is nonnegative and $q_j \in \mathbb{R}^+$ for $j > 0$, then an expression of the form

$$\alpha = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \dots + \frac{1}{q_{\ell-1} + \frac{1}{q_\ell}}}}$$

denoted $\langle q_0; q_1, \dots, q_\ell \rangle$, is a *finite continued fraction* of length $\ell(\alpha) = \ell$.^{1.1} A finite continued fraction is said to be *simple* if $q_j \in \mathbb{Z}$ for all $j = 0, 1, 2, \dots, \ell$.

^{1.1}There is a more general notion of a continued fraction, where the 1's in the numerators may be replaced with functions of a real or complex variable (and this comment also holds for the definition of an infinite continued fraction given below). However, we will be concerned primarily with simple continued fraction expansions in this article. For more details on other types of continued fractions, see Perron [13].

The values q_j are the *partial quotients*. The semi-colon is used after the first partial quotient q_0 to separate the integer value of α from the rest of the partial quotients, namely $q_0 = \lfloor \alpha \rfloor$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to $x \in \mathbb{R}$, or the *floor* of the real number x . Hence, the Euclidean algorithm yields that every rational number has a finite simple continued fraction expansion. In our example

$$\frac{2431}{165} = \frac{221}{15} = \langle 14; 1, 2, 1, 3 \rangle.$$

In fact, Euler first showed that $\alpha \in \mathbb{Q}$ if and only if α has a finite simple continued fraction expansion (see [5, Theorem 5.1.1, p. 223]). Thus, the finite simple continued fraction expansions provide a neat package.

Although evidence of continued fractions may be found in the works of Archimedes (ca. 287–212 B.C.), Claudius Ptolemy (ca. 85–165 A.D.), and Hindu mathematicians such as Āryabhata (ca. 476–550 A.D.) and Bhāskara (ca. 1114–1185 A.D.), among numerous others, including Chinese mathematicians of antiquity, it may be said that the foundation of the modern theory of continued fractions was laid in Bologna, Italy, by Rafael Bombelli (1526–1573) and Pietro Antonio Cataldi (1548–1626). Essentially, Bombelli gave a method for calculating \sqrt{D} for $D \in \mathbb{N}$, not a perfect square, that is equivalent to the continued fraction expansion of \sqrt{D} . He let $D = a^2 + r$ where $a = \lfloor \sqrt{D} \rfloor$, from which he deduced

$$\sqrt{D} = a + \frac{r}{a + \sqrt{D}},$$

so the expansion repeats indefinitely. Shortly, we will see that beneath the surface of this representation are the rumblings of palindromy.^{1,2}

Cataldi, who is responsible for pioneering the modern treatment of continued fractions, followed Bombelli’s method and developed symbolism for them, although his symbols differ from what we use today. This is given as a natural extension of the notion of a finite continued fraction defined above.

If $q_j \in \mathbb{R}$ where $j \in \mathbb{Z}$ is nonnegative and $q_j \in \mathbb{R}^+$ for $j > 0$, then an expression of the form

^{1,2}A *palindrome* is a *word* that reads the same forwards and backwards. For instance, a classical palindrome is: “able was I ere I saw elba”. A modern palindrome is: “a toyota”. A numerical palindrome is 1881. The etymology of *palindrome* is from the Greek *palindromos* or *running back again*. This derives from *palin*, meaning *back*, plus *dramein*, meaning *to run*.

$$\alpha = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \dots + \frac{1}{q_k + \frac{1}{q_{k+1} + \dots}}}}$$

denoted $\langle q_0; q_1, \dots, q_k, q_{k+1}, \dots \rangle$, is an *infinite continued fraction*. If $q_k \in \mathbb{Z}$ for all nonnegative integers k , then the infinite continued fraction is called *simple*.

The development of the theory of continued fractions into a field in its own right must be credited to John Wallis (1616–1703), who was a contemporary of Newton and one of the founding members of the Royal Society. In his book *Opera Mathematica*, published in 1695, the term *continued fraction* was first introduced. In this book, he laid the foundations of continued fractions as we know them today.

In a paper published in 1737, titled *De Fractionibus Continuis*, Euler showed that irrational numbers yield infinite simple continued fraction expansions. For instance, he gave:

$$e = \langle 2; 1, 2, 1, 1, 4, 1, 1, 6, \dots, 1, 1, 2k, \dots \rangle,$$

for $k \in \mathbb{N}$, $k > 1$.

Also, by 1869, J.J. Sylvester had found continued fraction representations related to π , which were equivalent to expressions found by Wallis. In particular,^{1.3}

$$\pi = \langle 3; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, 1, 84, 2, 1, 1, 15, \dots \rangle,$$

and there is no known pattern in the sequence of partial quotients as there is for e . However, as is the case with \sqrt{D} , there are real numbers that *do* have some repeated patterns; in fact, some very special repeated patterns.

An infinite simple continued fraction $\alpha = \langle q_0; q_1, q_2, \dots \rangle$ is called *periodic* (sometimes called *eventually periodic*), if there exists an integer $k \geq 0$ and

^{1.3}An amusing identity, which is an excellent problem to assign for calculus students to solve, is: $\int_0^1 \frac{t^4(1-t)^4}{1+t^2} dt = 22/7 - \pi$, with a reminder for the more recalcitrant that $\int_0^1 \frac{t^4(1-t)^4}{1+t^2} dt \neq 0$.

$\ell \in \mathbb{N}$ such that $q_n = q_{n+\ell}$ for all integers $n \geq k$. We use the notation

$$\alpha = \langle q_0; q_1, \dots, q_{k-1}, \overline{q_k, q_{k+1}, \dots, q_{\ell+k-1}} \rangle,$$

as a convenient abbreviation. The smallest such natural number $\ell = \ell(\alpha)$ is called the *period length* of α , and q_0, q_1, \dots, q_{k-1} is called the *pre-period* of α . If k is the *least* nonnegative integer such that $q_n = q_{n+\ell}$ for all $n \geq k$, then $q_k, q_{k+1}, \dots, q_{k+\ell-1}$ is called the *fundamental period* of α . If $k = 0$ is the least such value, then α is said to be *purely periodic*, namely

$$\alpha = \langle \overline{q_0; q_1, \dots, q_{\ell-1}} \rangle.$$

For instance,

$$\frac{1 + \sqrt{7}}{2} = \langle \overline{1; 1, 4, 1} \rangle$$

is purely periodic.

Euler proved, in the above cited paper, that a periodic continued fraction is the zero of a quadratic equation. To understand how this works, we need the following notion.

Definition 1.2 (Quadratic Irrationals)

Suppose that $D \in \mathbb{N}$ is not a perfect square. Then a quadratic irrational is a number of the form

$$\alpha = \frac{P + \sqrt{D}}{Q}, \quad (P, Q \in \mathbb{Z})$$

where $Q \neq 0$ and $P^2 \equiv D \pmod{Q}$. Also, the algebraic conjugate of α is

$$\alpha' = \frac{P - \sqrt{D}}{Q}.$$

It is easily checked that α and α' of Definition 1.2 are the roots of

$$x^2 - T(\alpha)x + N(\alpha) = 0, \tag{1.3}$$

where

$$T(\alpha) = \alpha + \alpha' = \frac{P + \sqrt{D}}{Q} + \frac{P - \sqrt{D}}{Q} = \frac{2P}{Q},$$

is the *trace* of α , and

$$N(\alpha) = \alpha\alpha' = \left(\frac{P + \sqrt{D}}{Q}\right) \left(\frac{P - \sqrt{D}}{Q}\right) = \frac{P^2 - D}{Q}$$

is the *norm* of α . Conversely, it can be shown that any root of

$$Ax^2 + Bx + C = 0,$$

where $A \neq 0$, and $B^2 - 4AC$ is not a perfect square must be a quadratic irrational. There are special kinds of quadratic irrationals, which are of primary interest. The following is attributable to Lagrange, who along with Euler and Lambert,^{1,4} made significant advances in the theory of continued fractions in the eighteenth century.

Theorem 1.4 *Let $\alpha \in \mathbb{R}$. Then α has a periodic infinite simple continued fraction expansion if and only if α is a quadratic irrational. Furthermore, if $\alpha = \langle q_0; q_1, \dots \rangle$ is an infinite simple continued fraction, with $\ell(\alpha) = \ell \in \mathbb{N}$, then α is purely periodic if and only if $\alpha > 1$ and $-1 < \alpha' < 0$. Any quadratic irrational which satisfies these two conditions is called reduced.*

Proof. See [5, Theorem 5.3.2, p. 241]. □

Corollary 1.5 *If $D \in \mathbb{N}$ is not a perfect square, then*

$$\sqrt{D} = \langle q_0; \overline{q_1, q_2, \dots, q_{\ell-1}} \rangle,$$

where $q_j = q_{\ell-j}$ for $j = 1, 2, \dots, \ell - 1$ and $q_0 = \lfloor \sqrt{D} \rfloor$.^{1,5}

^{1,4}Johann Heinrich Lambert (1728–1777) was a Swiss-German mathematician, and an associate of Euler for a couple of years at the Berlin Academy. Among his results was the first proof that π is irrational. By this time, Euler had already demonstrated the irrationality of e .

^{1,5}An open conjecture is that for any $n \in \mathbb{N}$, there exist infinitely many primes p such that $\ell(\sqrt{p}) = n$. Resolution of this conjecture would have some rather serious consequences. Among them would be resolution of the longstanding conjecture that there are infinitely many primes of the form $n^2 + 1$. However, it is known that for any $n \in \mathbb{N}$, there exist infinitely many *squarefree* $D \in \mathbb{N}$ such that $\ell(\sqrt{D}) = n$. See [2] and [3].

From Corollary 1.5, we see that in the simple continued fraction of \sqrt{D} ,

$$q_1 q_2 \cdots q_{\ell-1}$$

is a palindrome. Later we will see how this palindromy is tied with other concepts.

Lagrange was the first to give a proof that the Pell equation:^{1.6}

$$x^2 - Dy^2 = 1$$

has an integer solution with $y \neq 0$ and D not a perfect square by using continued fractions. In order to show how all solutions of this equation and its cousin $x^2 - Dy^2 = -1$ are generated from continued fractions, we need the following notion, the properties of which were presented in Wallis's book, discussed above.

Definition 1.6 Let $n \in \mathbb{N}$ and let α have a continued fraction expansion

$$\langle q_0; q_1, \dots, q_n, \dots \rangle$$

with $q_j \in \mathbb{R}^+$ when $j > 0$. Then

$$C_k = \langle q_0; q_1, \dots, q_k \rangle$$

is the k^{th} convergent of α for any nonnegative integer $k \leq n$.

There is an important means for representing convergents, which we will need to determine the solutions of Diophantine equations such as Pell's equations.

Theorem 1.7 Let $\alpha = \langle q_0; q_1, \dots, q_n, \dots \rangle$ for $n \in \mathbb{N}$ be a continued fraction expansion. Define two sequences for $k \in \mathbb{Z}$ nonnegative:

$$A_{-2} = 0, A_{-1} = 1, A_k = q_k A_{k-1} + A_{k-2},$$

^{1.6}Although the study of the equation $x^2 - Dy^2 = 1$ has been around since the time of Archimedes, historians generally agree that the first systematic method for solving this Pell equation was given by William Brouncker (1620–1684) who was a Lord and the first president of the Royal Society. Euler's misapprehension that John Pell had developed the method stemmed from the method of solution given by Wallis in his book *Algebra*. Euler misinterpreted this as having been given originally by Pell.

and

$$B_{-2} = 1, B_{-1} = 0, B_k = q_k B_{k-1} + B_{k-2}.$$

Then

$$C_k = A_k/B_k = \frac{q_k A_{k-1} + A_{k-2}}{q_k B_{k-1} + B_{k-2}},$$

is the k^{th} convergent of α for any nonnegative integer $k \leq n$. Also, if $\alpha = \sqrt{D}$, $P_0 = 0$, $Q_0 = 1$, and for $k \geq 0$,

$$P_{k+1} = q_k Q_k - P_k, \quad Q_{k+1} = (D - P_{k+1}^2)/Q_k,$$

then

$$A_{k-1}^2 - DB_{k-1}^2 = (-1)^k Q_k. \text{ }^{1.7}$$

After Lagrange's aforementioned proof of the solvability of $x^2 - Dy^2 = 1$ using continued fractions, the solvability of $x^2 - Dy^2 = -1$ was determined via the parity of $\ell(\sqrt{D})$. In fact, the sequences of A_j and B_j given in Theorem 1.7 are the key to solving the Pell equations as follows.

Theorem 1.8 (\square) *Suppose that $D \in \mathbb{N}$ is not a perfect square and*

$$\alpha = \sqrt{D} = \langle q_0; \overline{q_1, q_2, \dots, q_\ell} \rangle$$

with $\ell = \ell(\sqrt{D})$. If ℓ is even, then all positive solutions of

$$x^2 - y^2 D = 1 \tag{1.9}$$

are given by

$$x = A_{k\ell-1} \text{ and } y = B_{k\ell-1}$$

for $k \in \mathbb{N}$, whereas there are no solutions to

$$x^2 - y^2 D = -1. \tag{1.10}$$

If ℓ is odd, then all positive solutions of Equation 1.9 are given by

$$x = A_{2k\ell-1} \text{ and } y = B_{2k\ell-1}$$

for $k \in \mathbb{N}$, whereas all positive solutions of Equation 1.10 are given by

$$x = A_{(2k-1)\ell-1} \text{ and } y = B_{(2k-1)\ell-1}$$

for $k \in \mathbb{N}$.

^{1.7}See the more general equations, (2.2)–(2.4), in Section 2.

Proof. See [5, Corollary 5.3.3, p. 249]. \square

The elegant solution of Pell's equations given in Theorem 1.8 deserves consideration as a gem in the theory of continued fractions (and all such gems will be signified by the symbol \square).

Example 1.11 Let $D = 65$. Then $\sqrt{D} = \langle 8; \overline{16} \rangle$, so $\ell(\sqrt{65}) = 1$. We calculate that $A_0 = 8$, $A_1 = 129$, $A_2 = 2072$, $A_3 = 33281, \dots$, and $B_0 = 1$, $B_1 = 16$, $B_2 = 257$, $B_3 = 4128, \dots$, so by Theorem 1.8, all of the infinitely many solutions to (1.9) are given by

$$A_{2k-1}^2 - B_{2k-1}^2 \cdot 65 = 1,$$

for all $k \in \mathbb{N}$. In particular,

$$A_1^2 - B_1^2 \cdot 65 = 129^2 - 16^2 \cdot 65 = A_3^2 - B_3^2 \cdot 65 = 33281^2 - 4128^2 \cdot 65 = 1.$$

Also, all of the infinitely many solutions of (1.10) are given by

$$A_{2k-2}^2 - B_{2k-2}^2 \cdot 65 = -1,$$

for all $k \in \mathbb{N}$, so in particular,

$$A_0^2 - B_0^2 \cdot 65 = 8^2 - 1^2 \cdot 65 = A_2^2 - B_2^2 \cdot 65 = 2072^2 - 257^2 \cdot 65 = -1.$$

From Theorems 1.7–1.8 we see that if $x^2 - y^2D = \pm 1$, for nonsquare D , has a solution $x, y \in \mathbb{Z}$, then $x/y = A_j/B_j = C_j$ is a convergent in the simple continued fraction expansion of \sqrt{D} . There is a more general result.

Theorem 1.12 Let $D \in \mathbb{N}$ not a perfect square, and $n \in \mathbb{Z}$ with $|n| < \sqrt{D}$. If $x^2 - Dy^2 = n$, then there exists a $j \in \mathbb{N}$ such that $x/y = A_j/B_j = C_j$ is a convergent in the simple continued fraction expansion of \sqrt{D} .

There is a related result for the case where $n > \sqrt{D}$, but this involves the introduction of continued fraction expansions with negative partial quotients (see [5, pp. 334–340]).

The use of continued fractions to solve Diophantine equations is not new. In fact, F.E. Lucas (1842–1891) showed that every divisor of a sum of two relatively prime squares is a sum of two squares using continued fractions (see [14]). We will explore the interrelationships between continued fractions and sums of squares in Section Three. First, we look at the relationship between ideals and continued fractions in Section Two.

2 Continued Fractions and Ideals

In this section, we will see how ideal theory is linked to continued fractions via what the late Dan Shanks (see [7]) called the *infrastructure of a real quadratic field*.

Let $D_0 > 1$ be a square-free positive integer and set:

$$\sigma_0 = \begin{cases} 2 & \text{if } D_0 \equiv 1 \pmod{4}, \\ 1 & \text{otherwise.} \end{cases}$$

Define

$$\omega_0 = (\sigma_0 - 1 + \sqrt{D_0})/\sigma_0,$$

and

$$\Delta_0 = (\omega_0 - \omega'_0)^2 = 4D_0/\sigma_0^2.$$

The value Δ_0 is called a *fundamental discriminant* or *field discriminant* with associated *radicand* D_0 , and ω_0 is called the *principal fundamental surd associated with* Δ_0 . Let

$$\Delta = f_\Delta^2 \Delta_0$$

for some $f_\Delta \in \mathbb{N}$. If we set $g = \gcd(f_\Delta, \sigma_0)$, $\sigma = \sigma_0/g$,

$$D = (f_\Delta/g)^2 D_0,$$

and

$$\Delta = 4D/\sigma^2,$$

then Δ is called a *discriminant* with associated *radicand* D . Furthermore, if we let

$$\omega_\Delta = (\sigma - 1 + \sqrt{D})/\sigma,$$

then ω_Δ is called the *principal surd* associated with the discriminant $\Delta = (\omega_\Delta - \omega'_\Delta)^2$. This will provide the canonical basis element for certain rings that we now define.

Let $[\alpha, \beta] = \alpha\mathbb{Z} + \beta\mathbb{Z}$ be a \mathbb{Z} -module. Then

$$\mathfrak{O}_\Delta = [1, \omega_\Delta],$$

is an *order* in $K = \mathbb{Q}(\sqrt{\Delta}) = \mathbb{Q}(\sqrt{D_0})$ with conductor f_Δ . If $f_\Delta = 1$, then \mathfrak{O}_Δ is called the *maximal order in* K , or *ring of integers of* K .

It may be shown that any \mathbb{Z} -module $I \neq (0)$ of \mathfrak{D}_Δ has a representation of the form $[a, b + c\omega_\Delta]$, where $a, c \in \mathbb{N}$ with $0 \leq b < a$. We will only be concerned with *primitive* ones, namely those for which $c = 1$. In other words, I is a primitive \mathbb{Z} -submodule of \mathfrak{D}_Δ if whenever $I = (z)J$ for some $z \in \mathbb{Z}$ and some \mathbb{Z} -submodule J of \mathfrak{D}_Δ , then $|z| = 1$. Thus, a canonical representation of a primitive \mathbb{Z} -submodule of \mathfrak{D}_Δ is obtained by setting $\sigma a = Q$ and $b = (P - 1)/2$ if $\sigma = 2$, while $b = P$ if $\sigma = 1$ for $P, Q \in \mathbb{Z}$, namely

$$I = [Q/\sigma, (P + \sqrt{D})/\sigma]. \quad (2.1)$$

Now we bring ideal theory onto the stage by giving a criterion for a primitive \mathbb{Z} -module to be a primitive ideal in \mathfrak{D}_Δ . A nonzero \mathbb{Z} -module I as given in (2.1) is called a primitive \mathfrak{D}_Δ -ideal if and only if $P^2 \equiv D \pmod{Q}$ (see [5, Theorem 3.5.1, p. 173]). *Henceforth, when we refer to an \mathfrak{D}_Δ -ideal it will be understood that we mean a primitive \mathfrak{D}_Δ -ideal.* Also, the value Q/σ is called the *norm of I* , denoted by $N(I)$. Hence, from the discussion in Section 1, we see that I is an \mathfrak{D}_Δ -ideal if and only if $\alpha = (P + \sqrt{D})/Q$ is a quadratic irrational. We denote this ideal by $I = [\alpha]$. This yields a mapping

$$\alpha \mapsto [\alpha],$$

from the set of quadratic irrationals with underlying discriminant Δ to the set of \mathfrak{D}_Δ -ideals. However, the mapping is not one-to-one since

$$I = [Q/\sigma, (P + \sqrt{D})/\sigma] = [Q/\sigma, (nQ \pm (P + \sqrt{D}))/\sigma],$$

for any $n \in \mathbb{Z}$ (see [5, Theorem 3.5.2, p. 175]).

For instance,

$$I = [3, (5 + \sqrt{85})/2] = [3, (11 + \sqrt{85})/2] = J,$$

so

$$\alpha = \frac{5 + \sqrt{85}}{6} \mapsto I = [\alpha],$$

and

$$\beta = \frac{11 + \sqrt{85}}{6} \mapsto J = [\beta] = [\alpha],$$

but $\alpha \neq \beta$.

Given the notion of a reduced quadratic irrational in Theorem 1.4, it is not surprising that we define a *reduced ideal* I to be one which contains an

element $\beta = (P + \sqrt{D})/\sigma$ such that $I = [N(I), \beta]$, where $\beta > N(I)$ and $-N(I) < \beta' < 0$, since this corresponds exactly to the reduced quadratic irrational $\alpha = \beta/N(I) > 1$ with $-1 < \alpha' < 0$. In fact, it can be shown that $I = [Q/\sigma, (P + \sqrt{D})/\sigma]$ is reduced if $Q/\sigma < \sqrt{\Delta}/2$; and that if I is reduced, then $Q/\sigma < \sqrt{\Delta}$ (see [5, Corollaries 5.5.1–5.5.2, p. 259]). Also, observe that if α is reduced, then the map $\alpha \mapsto [\alpha]$ becomes one-to-one when considered as a map from the set of *reduced* quadratic irrationals to the set of *reduced* \mathfrak{D}_Δ -ideals.

Now we link continued fractions to the ideals defined above. Let I be an \mathfrak{D}_Δ -ideal given by (2.1). Define

$$P_0 = P, Q_0 = Q, \text{ and recursively for } j \geq 0,$$

$$q_j = \left\lfloor \frac{P_j + \sqrt{D}}{Q_j} \right\rfloor. \quad (2.2)$$

$$P_{j+1} = q_j Q_j - P_j, \quad (2.3)$$

and

$$D = P_{j+1}^2 + Q_j Q_{j+1}. \quad (2.4)$$

It follows that we have the simple continued fraction expansion of α given by,

$$\alpha = \frac{P + \sqrt{D}}{Q} = \frac{P_0 + \sqrt{D}}{Q_0} = \langle q_0; q_1, \dots, q_j, \dots \rangle.$$

Now the stage is set for the appearance of the result that formally merges ideals and continued fractions. We only need the notion of the equivalence of two \mathfrak{D}_Δ -ideals I and J , denoted $I \sim J$ to proceed. We write $I \sim J$ to denote the fact that there exist nonzero integers $\alpha, \beta \in \mathfrak{D}_\Delta$ such that $(\alpha)I = (\beta)J$, where (x) denotes the principal \mathfrak{D}_Δ -ideal generated by $x \in \mathfrak{D}_\Delta$. For a given discriminant Δ , the *class group* of \mathfrak{D}_Δ determined by these equivalence classes, denoted by \mathfrak{C}_Δ , is finite with order, denoted h_Δ , called the *class number* of \mathfrak{D}_Δ .

Theorem 2.5 (□) The Continued Fraction Algorithm

Suppose that $\Delta \in \mathbb{N}$ is a discriminant, P_j, Q_j are given by (2.2)–(2.4), and

$$I_j = [Q_{j-1}/\sigma, (P_{j-1} + \sqrt{D})/\sigma]$$

for nonnegative $j \in \mathbb{Z}$. Then $I_1 \sim I_j$ for all $j \in \mathbb{N}$. Furthermore, there exists a least natural number n such that I_{n+j} is reduced for all $j \geq 0$, and these I_{n+j} are all of the reduced ideals equivalent to I_1 .

Proof. See [5, Theorem 5.5.2, pp. 261–266]. \square

The remarkable link between ideals and continued fractions that lies below the surface of Theorem 2.5 is illustrated by the following.

Example 2.6 Let $\Delta = D = 385$, and set

$$I_1 = [7, (7 + \sqrt{385})/2],$$

with $Q_0 = 14$, $P_0 = 7$ and $\sigma = 2$. Then we may tabulate the following data for

$$\alpha = \frac{7 + \sqrt{385}}{14} = \frac{P_0 + \sqrt{385}}{Q_0} = \alpha_0.$$

Table 2.7

i	0	1	2	3	4	5	6	7	8	9	10
P_i	7	7	17	19	17	15	15	17	19	17	7
Q_i	14	24	4	6	16	10	16	6	4	24	14
q_i	1	1	9	6	2	3	2	6	9	1	1

Here $\ell(\alpha) = 10 = \ell(\alpha_0)$, and

$$\alpha = \langle \overline{1; 1, 9, 6, 2, 3, 2, 6, 9, 1} \rangle,$$

which is purely periodic since α is reduced. In fact, $\ell(\alpha_j) = 10$ for all $j \geq 0$ since we need merely appropriately permute the columns of Table 2.7 to get the continued fraction expansion of α_j .

From Example 2.6, we see that the continued fraction expansions of the reduced quadratic irrational α and the associated reduced ideal $[\alpha]$ form a related cycle. This is the essence of the continued fraction algorithm.

Definition 2.8 Suppose that $\Delta \in \mathbb{N}$ is a discriminant with radicand D and $I = I_1 = [Q_0/\sigma, (P + \sqrt{D})/\sigma]$ is a reduced \mathfrak{D}_Δ -ideal. If $\ell \in \mathbb{N}$ is the least value such that

$$I_1 = I_{\ell+1} = [Q_\ell/\sigma, (P_\ell + \sqrt{D})/\sigma],$$

then

$$\alpha_j = \frac{P_j + \sqrt{D}}{Q_j} \quad (2.9)$$

for $j \geq 0$ all have the same period length

$$\ell(\alpha_j) = \ell(\alpha_0) = \ell(\alpha) = \ell,$$

and

$$[\alpha_j] = [I_{j+1}] = [Q_j/\sigma, (P_j + \sqrt{D})/\sigma].$$

We denote this common value by $\ell = \ell(\mathcal{C})$, where \mathcal{C} is the class of I in \mathfrak{C}_Δ , called the period length of the cycle of reduced ideals equivalent to I . If we wish to track a specific ideal I in the cycle, then we write $\ell(I)$ for ℓ .

From the continued fraction algorithm, we see that if

$$I = [Q/\sigma, (P + \sqrt{D})/\sigma]$$

is a reduced \mathfrak{D}_Δ -ideal, then the set

$$\{Q_1/\sigma, Q_2/\sigma, \dots, Q_\ell/\sigma\}$$

represents the norms of all reduced ideals equivalent to I . This is achieved via the simple continued fraction expansion of $\alpha = (P + \sqrt{D})/Q$. For instance, in Example 2.6 the norms of all ideals equivalent to $[7, (7 + \sqrt{385})/2]$ are given by $\{Q_0, Q_1, Q_2, Q_3, Q_4, Q_5\} = \{7, 12, 2, 3, 8, 5\}$. These norms correspond to the ideals $I_j = [Q_{j-1}/\sigma, (P_{j-1} + \sqrt{385})/2]$ for $j = 1, 2, 3, 4, 5, 6$, given by Table 2.7.

Another manifestation of the continued fraction algorithm, given in Theorem 2.5, is the relation between the continued fraction expansions of $\alpha = (P + \sqrt{D})/Q$ and units in the ring of integers of $\mathbb{Q}(\sqrt{D})$. First, we remind the reader of the following.

Definition 2.10 *If $\Delta \in \mathbb{N}$ is a discriminant, then an element $u \in \mathfrak{D}_\Delta$ is called a unit of \mathfrak{D}_Δ if $|N(u)| = 1$. In particular, the smallest positive unit satisfying (1.3) is called the fundamental unit of \mathfrak{D}_Δ , denoted ε_Δ .*

For example,

$$\varepsilon_\Delta = \frac{1 + \sqrt{5}}{2} = \omega_\Delta$$

is the fundamental unit of

$$\mathfrak{D}_\Delta = \mathfrak{D}_5 = [1, (1 + \sqrt{5})/2],$$

since $x = \varepsilon_5$ is the smallest positive solution of $x^2 - x - 1 = 0$ in \mathfrak{D}_Δ . It turns out that ε_5 is the smallest fundamental unit in any real quadratic order.^{2.8}

We are now in a position to provide the application of the continued fraction algorithm to units. The following was proved by Lagrange in 1769 for the principal class, but easily generalizes to the stated result for an arbitrary discriminant.

Theorem 2.11 (\square) (Lagrange)

If Δ , I_j , P_j , and Q_j are given as in Theorem 2.5, and $\ell = \ell(I_1)$, then

$$\varepsilon_\Delta = \prod_{k=1}^{\ell} \frac{P_k + \sqrt{D}}{Q_k},$$

and

$$N(\varepsilon_\Delta) = (-1)^\ell.$$

Proof. See [5, Theorem 5.5.3, p. 269]. \square

Example 2.12 Consider Example 2.6 again. From Table 2.7 we have:

$$\begin{aligned} \varepsilon_\Delta &= \left(\frac{7 + \sqrt{385}}{24} \right) \cdot \left(\frac{17 + \sqrt{385}}{4} \right) \cdot \left(\frac{19 + \sqrt{385}}{6} \right) \cdot \left(\frac{17 + \sqrt{385}}{16} \right) \\ &\quad \cdot \left(\frac{15 + \sqrt{385}}{10} \right) \cdot \left(\frac{15 + \sqrt{385}}{16} \right) \cdot \left(\frac{17 + \sqrt{385}}{6} \right) \cdot \left(\frac{19 + \sqrt{385}}{4} \right) \end{aligned}$$

^{2.8}The value $\tau=(1+\sqrt{5})/2$ is called the *golden mean*. It has a long and rich history. For instance, the Romans of antiquity used it in their architectural designs. Also, the value is related to the Fibonacci numbers via $F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$. See [5, pp. 26–33] for more information. Furthermore, we have the fascinating continued fraction expansion of the infinite series: $\sum_{n=1}^{\infty} \frac{1}{2^{\lfloor n/\tau \rfloor}} = (0, 2^0, 2^1, 2^1, 2^2, 2^3, \dots, 2^{F_n}, \dots)$. See [1].

$$\cdot \left(\frac{17 + \sqrt{385}}{24} \right) \cdot \left(\frac{7 + \sqrt{385}}{14} \right) = 95831 + 4884\sqrt{385},$$

and

$$N(\varepsilon_\Delta) = 1 = (-1)^\ell = (-1)^{10}.$$

The next result makes clearer the connection between Theorems 2.5 and 2.11.

Theorem 2.13 *If $\Delta \in \mathbb{N}$ is a discriminant with associated radicand D and $\ell = \ell(\sqrt{D})$, with $A = A_{\ell-1}$ and $B = B_{\ell-1}$ defined in Theorem 1.7, then*

$$\varepsilon_\Delta = A + B\sqrt{D},$$

or

$$\varepsilon_\Delta^3 = A + B\sqrt{D},$$

where the latter can only occur if $D \equiv 5 \pmod{8}$.

Example 2.14 *In Example 2.12, $\varepsilon_\Delta = \varepsilon_D = \varepsilon_{385} = 95831 + 4884\sqrt{385}$, where $A_{\ell-1} = A_9 = 95831$ and $B_{\ell-1} = B_9 = 4884$ in the simple continued fraction expansion of $\sqrt{385}$.*

However, if we look at the example where $\Delta = D = 85$, then $\varepsilon_\Delta = (9 + \sqrt{85})/2$ and $\varepsilon_\Delta^3 = 378 + 41\sqrt{85}$. In the latter example, $\ell = 5$, $A_{\ell-1} = A_4 = A = 378$ and $B_{\ell-1} = B_4 = B = 41$, where these last facts may be gleaned from the following table, which represents the data for the simple continued fraction expansion of $\sqrt{85}$.

Table 2.15

i	0	1	2	3	4	5
P_i	0	9	7	2	7	9
Q_i	1	4	9	9	4	1
q_i	9	4	1	1	4	18

Now that we have explored the relationship between continued fraction expansions and ideals, and have linked the results with what we learned in Section 1, we are now ready to explore palindromy and sums of squares from the perspective of the interlinked notions of continued fractions, ideals, and norms of quadratic irrationals from real quadratic fields.

3 Palindromy, Sums of Squares, Ambiguity

Diophantus (ca. 250 A.D.) may be credited with the first systematic study of those integers that can be represented as a sum of two integer squares. He observed that 15, for instance, is not the sum of two integer squares. Fermat, first proved that every prime $p \equiv 1 \pmod{4}$ is a sum of two squares in essentially one way. Today, a course in elementary number theory teaches that $D \in \mathbb{N}$ is a sum of two integer squares if and only if the canonical prime factorization of D (via the fundamental theorem of arithmetic), does not contain a prime $p \equiv 3 \pmod{4}$ to an odd exponent.^{3,9} There is a criterion for representation of any $D \in \mathbb{N}$ as a sum of two integer squares via palindromy of the simple continued fraction expansion of \sqrt{D} . To state this result, we need the following notion.

Definition 3.1 *A reduced quadratic irrational*

$$\alpha = \langle \overline{q_0; q_1, \dots, q_{\ell-1}} \rangle = \frac{P + \sqrt{D}}{Q}$$

is said to have pure symmetric period if

$$q_j = q_{\ell-j-1} \text{ for all nonnegative } j \leq \ell - 1.$$

In other words,

$$q_0 q_1 \dots q_{\ell-1}$$

is a palindrome.

Example 3.2 *If $D = \Delta = 145$, and $\alpha = (9 + \sqrt{145})/8$, then one calculates via (2.2)–(2.4)*

$$\alpha = \langle \overline{2; 1, 1, 1, 2} \rangle,$$

which has pure symmetric period.

Now we link the notion of pure symmetric period with a special kind of ideal that will merge the notions of continued fractions and sums of squares with ideal theory.

^{3,9}The *number of ways* that D can be represented as a sum of two integer squares (with which we will not be explicitly concerned here), is a little more complicated. See [5, Theorem 6.1.3, pp. 279–286].

Definition 3.3 If $I = [Q/\sigma, (P + \sqrt{D})/\sigma]$ is an \mathfrak{O}_Δ -ideal where Δ is a discriminant, then

$$I' = [Q/\sigma, (P - \sqrt{D})/\sigma]$$

is called the conjugate ideal of I . If $I = I'$, then I is said to be an ambiguous ideal. If $I \sim I'$, then the class of I in \mathfrak{C}_Δ is called ambiguous.

Certainly all ambiguous ideals are in an ambiguous class, but an ambiguous class need not contain an ambiguous ideal.

Example 3.4 Let $D = 221 = \Delta$, and consider $I = [5, (11 + \sqrt{221})/2]$. Then the continued fraction expansion for I , determined by the continued fraction algorithm, is given via:

Table 3.5

i	0	1	2	3
P_i	11	9	5	9
Q_i	10	14	14	10
q_i	2	1	1	2

Since $[\alpha] = I$ where $\alpha = (11 + \sqrt{221})/10$ and α is reduced, then I is reduced, and none of the ideals $I_1 = [5, (11 + \sqrt{221})/2]$, $I_2 = [7, (9 + \sqrt{221})/2]$, $I_3 = [7, (5 + \sqrt{221})/2]$, and $I_4 = [5, (9 + \sqrt{221})/2]$ is ambiguous. Hence, the class of I has no ambiguous ideal in it since it can be shown that if a class has an ambiguous ideal, then it has an ambiguous reduced ideal in it (see [9, Theorem 2.3]).^{3,10} Notice as well that $\alpha = (11 + \sqrt{221})/10 = \langle 2; 1, 1, 2 \rangle$ has pure symmetric period.

Theorem 3.6 (\square) Let $\alpha = (P + \sqrt{D})/Q = \langle q_0; q_1, \dots, q_\ell \rangle$ be a reduced quadratic irrational with radicand $D \in \mathbb{N}$. Then the following are equivalent.

- (a) α has pure symmetric period.
- (b) $N(\alpha) = \alpha\alpha' = -1$.
- (c) $D = P^2 + Q^2$.

^{3,10}For complete details on ambiguous classes of ideals with no ambiguous ideals in them, see [4, Chapter 6, pp. 187–199].

- (d) The ideal class of $[\alpha]$ has at most one reduced ambiguous ideal in it, where $[\alpha'] = [\alpha_{\ell-1}]$.
- (e) For all nonnegative integers $j \leq \ell - 1$,

$$\alpha'_j \alpha_{\ell-j} = -1,$$

where α_j is given by (2.9).

Proof. See [9, Theorem 2.2]. \square

Example 3.7 In Example 3.4, we saw an instance where there were no ambiguous ideals in the class of $[\alpha] = [(11 + \sqrt{221})/10]$. Here α has pure symmetric period. Notice that $D = 221 = 11^2 + 10^2 = P^2 + Q^2$, $N(\alpha) = \alpha\alpha' = -1$, and it may be verified that $\alpha'_3\alpha_1 = -1 = \alpha'_2\alpha_2 = \alpha'_1\alpha_3$.

An instance where there is exactly one ambiguous class with an ambiguous ideal is given by the quadratic irrational $\alpha = (5 + \sqrt{145})/10$, where the only ambiguous ideal in the class $[\alpha]$ is $[5, (5 + \sqrt{145})/2]$. To see this, consider the continued fraction data for α via the continued fraction algorithm.

Table 3.8

i	0	1	2	3	4
P_i	5	5	7	9	7
Q_i	10	12	8	8	12
q_i	1	1	2	2	1

Here $\ell(\alpha) = 5$. Also, among the ideals

$$I_2 = [6, (5 + \sqrt{145})/2], \quad I_3 = [4, (7 + \sqrt{145})/2],$$

$$I_4 = [4, (9 + \sqrt{145})/2], \quad \text{and} \quad I_5 = [6, (7 + \sqrt{145})/2],$$

none is ambiguous. Thus, as above, since the existence of an ambiguous ideal in the class implies the existence of reduced one, then $[5, (5 + \sqrt{145})/2]$ is the only ambiguous reduced ideal in its class. However, $\alpha = \langle \overline{1}; 1, 2, 2, \overline{1} \rangle$ does not have pure symmetric period, and

$$[\alpha'] = [\alpha] \neq [\alpha_{\ell-1}] = [6, (7 + \sqrt{145})/2].$$

This last example motivates another notion that will allow us to link the results developed thus far with another type of palindromy illustrated in Corollary 1.5, and fulfills the promise made in the remark made after that corollary.

Definition 3.9 If $\alpha = \langle \overline{q_0; q_1, \dots, q_\ell} \rangle = (P + \sqrt{D})/Q$ is a reduced quadratic irrational where $D \in \mathbb{N}$ is a radicand, then α is said to have pure skew-symmetric period if

$$q_j = q_{\ell-j} \text{ for all natural numbers } j \leq \ell - 1.$$

In other words,

$$q_1 q_2 \dots q_{\ell-1}$$

is a palindrome.

Theorem 3.10 (\square) Let $\alpha = (P + \sqrt{D})/Q$ be a reduced quadratic irrational with underlying radicand $D \in \mathbb{N}$ and let α_j , P_j , Q_j and ℓ be given as in Definition 2.8. Then the following are equivalent.

- (a) α has pure skew-symmetric period.
- (b) $\alpha\alpha'_1 = -1$.
- (c) $D = P_o^2 + Q_0Q_1$.
- (d) The ideal $[\alpha]$ is ambiguous.
- (e) For all natural numbers $j \leq \ell - 1$,

$$\alpha_{\ell-j+1}\alpha'_j = -1.$$

Proof. See [9]. \square

Example 3.11 In Example 3.7, $\alpha = (5 + \sqrt{145})/10 = \langle \overline{1; 1, 2, 2, 1} \rangle$ has pure skew-symmetric period, $D = P_o^2 + Q_0Q_1 = 5^2 + 10 \cdot 12$, $[\alpha] = [\alpha']$ is ambiguous,

$$\alpha\alpha'_1 = \left(\frac{5 + \sqrt{145}}{10} \right) \left(\frac{5 - \sqrt{145}}{12} \right) = -1,$$

and $\alpha_{\ell-j+1}\alpha'_j = \alpha_{6-j}\alpha'_j = -1$ for all $j \in \mathbb{N}$ with $j \leq 4$ (where $\alpha_5 = \alpha_0$).

The reader may wonder if there are reduced quadratic irrationals α with *both* pure symmetric *and* pure skew-symmetric period. It is an easy exercise to show that the answer is *yes*, but only if $\ell(\alpha) = 1$. For instance, $\alpha = 8 + \sqrt{65} = \langle 16 \rangle$ has both such periods. Hence, $65 = P_0^2 + Q_0^2 = 8^2 + 1^2$. For our purposes, this special case exhausts the consideration of palindromy and continued fractions. In Section 4, we will look at some connections between ambiguous ideals, continued fractions and Diophantine equations, which complement some of the results on Diophantine equations already considered herein.

4 Continued Fractions, Ideals, and Diophantine Equations

Eisenstein looked for criteria concerning the solvability of Diophantine equations of the form

$$|x^2 - Dy^2| = 4, \text{ where } \gcd(x, y) = 1, \quad (4.1)$$

which was a problem first considered by Gauss. It can be shown that if $x^2 - Dy^2 = -4$ is solvable for relatively prime integers x, y , then $\varepsilon_D \notin [1, \sqrt{D}]$, where ε_D is the fundamental unit of $[1, (1 + \sqrt{D})/2]$. The converse is false since, for example, $\varepsilon_{21} = (5 + \sqrt{21})/2 \notin [1, \sqrt{21}]$, where ε_{21} is the fundamental unit of $[1, (1 + \sqrt{21})/2]$. Yet, the Diophantine equation $x^2 - 21y^2 = -4$ is not solvable. However, if $D \equiv 5 \pmod{8}$, then (4.1) is solvable if and only if $\varepsilon_D \notin [1, \sqrt{D}]$. The following generalizes this notion by looking at more general equations.

Theorem 4.2 (\square) *Suppose that $\Delta = 4D$ is a discriminant with associated radicand D , I is a (primitive) \mathfrak{D}_Δ -ideal with $1 < N(I) < \sqrt{\Delta}$, and $N(I) \mid \Delta$. If $D = ab$ with $a, b \in \mathbb{N}$ and $ac < b$, then the Diophantine equation*

$$|ax^2 - by^2| = c,$$

where $c \in \{1, 2, 4\}$, has a solution $x, y \in \mathbb{Z}$ with $\gcd(x, y) = 1$ if and only if, in the simple continued fraction expansion of \sqrt{D} , $ac = N(I) = Q_{\ell/2}$ for $c = 1, 2$ and if $c = 4$, then $4a = N(I) = Q_f$ for some $f < \ell/2$.

Proof. See [10].^{4.11}

□

Example 4.3 Let $\Delta = 4D = 4 \cdot 805 = 2^2 \cdot 5 \cdot 7 \cdot 23$. Then

$$5x^2 - 161y^2 = -4$$

has the solution $x = 17$ and $y = 3$. Here $a = 5$, and $b = 161$ with $\ell = 18$, and

$$Q_{\ell/2} = a = 5 = Q_9, \quad Q_f = 4a = 20 = Q_3,$$

where $f = \ell/6$. This illustrates the fact that the $4a$ value, when it occurs, is about a sixth of the way along the period (not always exactly one-sixth as in this example). Notice as well that when $Q_j = 4$ for some j , then we have uncovered the fundamental unit. To see this, recall from Theorem 1.7, we have,

$$A_{j-1}^2 - B_{j-1}^2 D = A_5^2 - B_5^2 D = 1447^2 - 51^2 \cdot 805 = 4 = (-1)^6 Q_6,$$

and indeed

$$\varepsilon_D = \frac{A_5 + B_5 \sqrt{D}}{2} = \frac{1447 + 51\sqrt{805}}{2}$$

is the fundamental unit of $\mathfrak{D}_\Delta = [1, (1 + \sqrt{805})/2]$. The reader may now compare this with Theorem 1.8, where the solutions of Pell's equation was given in terms of continued fraction expansions. The interrelationships are now more transparent.

It is worthy of note that $Q_f = 4a$ for some $f < \ell$ and $D = P_f^2 + Q_f Q_{f-1}$ means that we have a means of finding a nontrivial factorization of D . For instance, in Example 4.3, $Q_3 = 20$ and $D = 805 = P_3^2 + Q_3 Q_2 = 15^2 + 20 \cdot 29 = 5(3^2 \cdot 5 + 4 \cdot 29) = 5 \cdot 161$, so we have a nontrivial factorization of D . Although finding a nontrivial factor of 805 is obvious, the method illustrated here generalizes to a wide range of integers. We study continued fractions applied to factoring in the next section.

^{4.11}There are also results linking continued fractions and Jacobi symbols in [10], and [8], which we do not have the room to develop here.

5 Continued Fractions and Factoring

In Section 4, we saw how the continued fraction expansion of \sqrt{D} may be used to determine a factor of D . Factoring is difficult, and this has implications for modern day cryptography, namely the design and implementation of secrecy codes or systems. Numerous cryptographic schemes, such as the *RSA cryptosystem*, depend upon the difficulty of factoring for their security (see [4]–[5] for details on the numerous applications of both elementary and algebraic number theory to cryptography). Given the importance of the secure transmission of data in our information-based society, and the role played by cryptography in ensuring that this data exchange remains private, it is fitting that we conclude this article with the part played by continued fractions.

Prior to 1970, a 25-digit integer was considered difficult to factor. In 1970, John Brillhart and Michael Morrison developed the continued fraction factoring method. This method raised the difficulty of factoring to fifty digits.^{5,12} Once the continued fraction method was up and running in 1970, legions of 20- to 45-digit numbers were being factored, which could not be factored before. The first major success was the factorization of the seventh Fermat number

$$F_7 = 2^{2^7} + 1 = 2^{128} + 1,$$

a thirty-nine digit number with two prime factors (see [11]–[12]).

Underlying the continued fraction method is an idea due to Maurice Kraitchik in the 1920's. Kraitchik's idea was that it might suffice to find a *multiple* of D as a difference of squares, namely

$$x^2 \equiv y^2 \pmod{D}, \tag{5.1}$$

so that one of $x - y$ or $x + y$ *could* be a nontrivial factor of D . We say *could* here since we fail to get a *nontrivial* factor of D when $x \equiv \pm y \pmod{D}$. However, it can be shown that if D is divisible by at least two distinct odd primes, then for at least half of the pairs x modulo D and y modulo D , satisfying (5.1) with $\gcd(x, y) = 1$, we have

$$1 < \gcd(x - y, D) < D.$$

^{5,12}Since then the number field sieve has supplanted this method, at least for a certain class of numbers to which it is best suited, but this is in the realm of algebraic number theory. For an elementary discussion of this method and applications of algebraic number theory to a variety of other cryptographic methods, including elliptic curves, see [6].

This classical idea has seeds in the work of Gauss, but Kraitchik introduced it into a new century in the pre-dawn of the computer age. Below we will see how this idea manifests itself in the continued fraction method.

(□) **The Continued Fraction Factoring Method**

Let $D \in \mathbb{N}$, not a perfect square, and let $C_j = A_j/B_j$ be a convergent in the simple continued fraction expansion of \sqrt{D} where A_j, B_j are given in Theorem 1.7, from which we know that for $j \in \mathbb{N}$,

$$A_{j-1}^2 - DB_{j-1}^2 = (-1)^j Q_j.$$

If j is even and $Q_j = m^2$ for some $m \in \mathbb{N}$, then $D \mid (A_{j-1}^2 - m^2)$. If $\gcd(A_{j-1} \pm m, D) > 1$, namely if $A_{j-1} \pm m \neq 0, 1$, then we have a nontrivial factor of D .

Example 5.2 *Let $D = 649$ and $\alpha = \sqrt{D}$. Then the simple continued fraction expansion of α is given in the following via the continued fraction algorithm.*

j	0	1	2	3	4
P_j	0	25	23	22	11
Q_j	1	24	5	33	16
q_j	25	2	9	1	2
A_j	25	51	484	535	
B_j	1	2	19	21	

We stop at $j = 4$ since $Q_4 = 16 = 4^2$ and we only need the value of A_{j-1} . Here $A_{j-1} = A_3 = 535$, so $535^2 \equiv 4^2 \pmod{649}$. Since

$$535 - 4 = 3^2 \cdot 59 \text{ and } 535 + 4 = 7^2 \cdot 11,$$

we have a nontrivial factor of 649. In fact, $649 = 11 \cdot 59$. Also, the reader will find, upon completion of the table, that $Q_6 = 5^2$, $A_5 = 5197$, $5197 - 5 = 2^3 \cdot 11 \cdot 59$ and $5197 + 5 = 2 \cdot 3^2 \cdot 17^2$, so we have another means of factoring D at $j = 6$. There is even more since $Q_{10} = 9$, from which the reader may show similarly that we can factor D .

Example 5.2 illustrates several means that yield nontrivial factorizations of D . However, there are cases where there are *none*.

Example 5.3 Let $D = 731$ and $\alpha = \sqrt{731}$. Then the table describing the continued fraction expansion of α is given by the following.

j	0	1	2
P_j	0	27	27
Q_j	1	2	1
q_j	27	27	54

We can never find a nontrivial square Q_j for any j , since $\alpha = \sqrt{731} = \langle 27; \overline{27, 54} \rangle$.

There is still hope of finding a nontrivial factor of D in such cases as that illustrated in Example 5.3. However, we need to modify the continued fraction method somewhat. If the continued fraction method fails for \sqrt{D} , then we may use that same algorithm on \sqrt{kD} where we may suitably choose m , usually as the product of the first few primes to avoid introducing squares under the surd. If we get $Q_j = h^2$ for even j in the continued fraction expansion of \sqrt{kD} , then

$$A_{j-1}^2 - B_{j-1}^2 kD = (-1)^j Q_j,$$

so $A_{j-1}^2 \equiv h^2 \pmod{kD}$ and we may have a factor of D , since we may get $\gcd(A_{j-1} \pm h, D) > 1$.

Example 5.4 Let $\alpha = \sqrt{6 \cdot 731} = \sqrt{4386}$. (The reader may verify that $\sqrt{2 \cdot 731}$ fails to lead to a factorization, so we go to the product of the first two primes for m , namely $m = 6$.) We need not go far in the tabular description of α to get the desired factorization in this case.

j	0	1	2
P_j	0	66	54
Q_j	1	30	49
q_j	66	4	
A_j	66	265	

Here, $Q_j = Q_2 = 7^2$, $A_{j-1} = A_1 = 265$, so $265^2 \equiv 7^2 \pmod{4386}$. Therefore,

$$265 - 7 = 2 \cdot 3 \cdot 43 \text{ and } 265 + 7 = 2^4 \cdot 17.$$

Indeed, since

$$4386 = 6 \cdot 731 = 2 \cdot 3 \cdot 17 \cdot 43,$$

we have the complete factorization $731 = 17 \cdot 43$.

Example 5.4 shows that the continued fraction algorithm may be extended to cases where the original method fails for D itself. This modified approach may also be used as an alternative if we find that we have gone through numerous values of j in the tabular expansion of the continued fraction expansion as above, without finding any square Q_j .

It is worthy of note here that, in the factorization of F_7 , the above idea was extended as follows. If one can find a *set* of values Q_k , say $\{Q_{k_1}, \dots, Q_{k_n}\}$ such that

$$\prod_{j=1}^n (-1)^{k_j} Q_{k_j} = x^2,$$

then for $A \equiv \prod_{j=1}^n A_{k_{j-1}} \pmod{D}$, we have

$$A^2 - x^2 \equiv 0 \pmod{D},$$

and (possibly) $\gcd(A \pm x, D) > 1$. It was this idea, together with the inspired method of implementing it using linear algebra over F_2 (the field of two elements), that was used by Morrison and Brillhart [11] in 1970. At that time, this was the biggest integer ever factored by such a general purpose technique, which is called *CFRAC*, for the *Continued Fraction Factoring Algorithm*, which was the leader in factorization methods for over a decade. Since then more powerful methods such as the *quadratic sieve* and *number field sieve* have taken over (see [5]–[6]).

For a given $D \in \mathbb{N}$, and $\ell = \ell(\sqrt{D})$, it is an open problem to do the following. Given knowledge of some $n \in \mathbb{N}$ such that $Q_j = n$ for some $j \leq \ell/2$ in the simple continued fraction expansion of \sqrt{D} , provide an efficient method for finding $i \in \mathbb{N}$ such that $n = Q_i$ where $i \neq j \leq \ell/2$. To solve this problem would be tantamount to solving the factoring problem. In other words, this is virtually intractable at this point in time.

Acknowledgments: The author's research is supported by NSERC Canada grant # A8484. Also, the author welcomes the opportunity to thank the referee for valuable comments.

References

- [1] P.G. Anderson, T. C. Brown, and P. J.-S. Shiue, *A Simple Proof of a Remarkable Continued Fraction Identity*, *Proceed. Amer. Math. Soc.* **123** (1995), 2005–2009.
- [2] P. Chowla and S. Chowla, *Problems on Periodic Simple Continued Fractions*, *Proc. Nat. Acad. Sci. U.S.A.* **69** (1972), 37–45.
- [3] C. Friesen, *On Continued Fractions of Given Period*, *Proceed. Amer. Math. Soc.* **103**, (1988), 9–14.
- [4] R.A. Mollin, **Quadratics**, CRC Press, Boca Raton, New York, London, Tokyo (1996).
- [5] R.A. Mollin, **Fundamental Number Theory with Applications**, CRC Press, Boca Raton, New York, London, Tokyo (1998).
- [6] R.A. Mollin, **Algebraic Number Theory**, CRC Press, Boca Raton, New York, London, Tokyo (1999).
- [7] R.A. Mollin, *Prime-Producing Quadratics*, *Amer. Math. Monthly* **104** (1997), 529–544.
- [8] R.A. Mollin, *Jacobi Symbols, Ambiguous Ideals, and Continued Fractions*, *Acta. Arith.* **LXXXV** (1998), 331–349.
- [9] R.A. Mollin, and K. Cheng, *Palindromy and Ambiguous Ideals Revisited*, (to appear: *J. Number Theory*).
- [10] R.A. Mollin, and A. J. van der Poorten, *Continued Fractions, Jacobi Symbols, and Quadratic Diophantine Equations*, (to appear: *Canad. Math. Bulletin*).
- [11] M.A. Morrison and J. Brillhart, *The Factorization of F_7* , *Bull. Amer. Math. Soc.*, **77**, 264 (1971).
- [12] M.A. Morrison and J. Brillhart, *A Method of Factoring and the Factorization of F_7* , *Math. Comp.*, **29**, 183–205 (1975).
- [13] O. Perron, **Die Lehre von den Kettenbrüchen**, Stuttgart, Teubner (1913); Reprinted, Chelsea, New York (1977).

- [14] H.C. Williams, **Édouard Lucas and Primality Testing**, Toronto (1998).

Mathematics Department
University of Calgary
Calgary, Alberta
T2N 1N4 Canada
ramollin@math.ucalgary.ca
<http://www.math.ucalgary.ca/~ramollin/>