Continued Fractions for Square Roots

In this section I'll consider continued fractions for quadratic irrationals of the form $\sqrt{d}$, where d is a positive integer which is not a perfect square. Let's recall some earlier results we'll need.

We have an algorithm for constructing the continued fraction for a quadratic irrational. Suppose a quadratic irrational has been written in the form $x = \dfrac{m + \sqrt{d}}{s}$, where:

(a) $m, s \in \mathbb{Z}$ and $s \neq 0$.

(b) $s \mid d - m^2$.

(c) d is a positive integer which is not a perfect square.

Set

$$m_0 = m, \quad s_0 = s, \quad x_0 = x, \quad a_0 = [x_0].$$

Then for $k \geq 0$, define (in order)

$$\begin{aligned} m_{k+1} &= a_k s_k - m_k \\ s_{k+1} &= \frac{d - m_{k+1}^2}{s_k} \\ x_{k+1} &= \frac{m_{k+1} + \sqrt{d}}{s_{k+1}} \\ a_{k+1} &= [x_{k+1}] \end{aligned}$$

The continued fraction for x is $[a_0, a_1, a_2, \ldots]$, and it is known to be periodic.

In the case of $\sqrt{d}$ where d is a positive integer which is not a perfect square, I may take $m_0 = m = 0$ and $s_0 = s = 1$.

The main result establishes the form of the continued fraction for $\sqrt{d}$. To state it, I need a definition.

Definition. A sequence of numbers $b_1, \ldots b_k$ is **palindromic** if $b_i = b_{k-i+1}$ for $i = 1, \ldots k$.

The empty sequence is considered to be palindromic.

For example, the following sequences are palindromic:

$$\{10, 3, 4, 4, 3, 10\}, \quad \{17, 2, 7, 8, 7, 2, 17\}.$$

Theorem. Let d be a positive integer which is not a perfect square.

(a) The continued fraction for $\sqrt{d}$ has the form

$$\sqrt{d} = [a_0, \overline{a_1, \ldots a_{n-1}, 2a_0}].$$

(b) The sequence $\{a_1, \ldots a_{n-1}\}$ is palindromic.

To illustrate, here is the continued fraction for $\sqrt{53}$:

$$\sqrt{53} = [7, 3, 1, 1, 3, 14, 3, 1, 1, 3, 14, \ldots].$$

Note that $14 = 2 \cdot 7$, and $\{3, 1, 1, 3\}$ is palindromic.

Likewise,

$$\sqrt{52} = [7, 4, 1, 2, 1, 4, 14, 4, 1, 2, 1, 4, 14, \ldots].$$

Note that $14 = 2 \cdot 7$ and $\{4, 1, 2, 1, 4\}$ is palindromic.

The proof of the theorem will use two theorems of Galois on purely periodic continued fractions which we proved earlier:

1. Let $x = [a_0, a_1, \ldots]$ be a quadratic irrational. The continued fraction for x is purely periodic if and only if $x > 1$ and $-1 < \overline{x} < 0$.

2. Let $x = [\overline{a_0, a_1, \ldots a_{n-1}, a_n}]$ be a purely periodic quadratic irrational. Then $-\dfrac{1}{\overline{x}}$ is purely periodic, and

$$-\frac{1}{\overline{x}} = [\overline{a_n, a_{n-1}, \ldots a_1, a_0}].$$

Proof. (a) **Step 1.** We'll show: The continued fraction for $\sqrt{d} + [\sqrt{d}]$ is purely periodic.

Recall that $[\cdot]$ denotes the greatest integer function.

Since d is a positive integer and is not a perfect square, $d > 1$ and

$$\sqrt{d} + [\sqrt{d}] > 1.$$

The conjugate is $[\sqrt{d}] - \sqrt{d}$. Since $\sqrt{d}$ is not an integer,

$$[\sqrt{d}] < \sqrt{d} < [\sqrt{d}] + 1.$$

The first of these two inequalities gives $[\sqrt{d}] - \sqrt{d} < 0$. The second of these two inequalities gives

$$\begin{aligned} \sqrt{d} - [\sqrt{d}] &< 1 \\ [\sqrt{d}] - \sqrt{d} &> -1 \end{aligned}$$

.

All together, the conjugate satisfies

$$-1 < [\sqrt{d}] - \sqrt{d} < 0.$$

We showed above that $\sqrt{d} + [\sqrt{d}] > 1$, so the hypotheses of the first of Galois's theorems is satisfied. Hence, $\sqrt{d} + [\sqrt{d}]$ is purely periodic. This completes the proof of Step 1.

Step 2. We'll show: The continued fraction for $\sqrt{d}$ has the form

$$\sqrt{d} = [a_0, \overline{a_1, \ldots a_{n-1}, 2a_0}].$$

Suppose the period of the continued fraction for $\sqrt{d} + [\sqrt{d}]$ is n. We write

$$\sqrt{d} + [\sqrt{d}] = [\overline{a_0, a_1, \ldots, a_{n-1}}].$$

Then

$$a_0 = \left[\sqrt{d} + [\sqrt{d}]\right] = 2[\sqrt{d}].$$

Thus,

$$\sqrt{d} + [\sqrt{d}] = [\overline{2[\sqrt{d}], a_1, \ldots, a_{n-1}}] = [2[\sqrt{d}], a_1, \ldots, a_{n-1}, \overline{2[\sqrt{d}], a_1, \ldots, a_{n-1}}].$$

Note that this is purely periodic, as we knew from Step 1. Moreover, it shows that if the period for the continued fraction for $\sqrt{d}$ is n, the same is true for the continued fraction for $[\sqrt{d}] + \sqrt{d}$.

Subtract $[\sqrt{d}]$ from both sides to obtain

$$\sqrt{d} = [[\sqrt{d}], a_1, \ldots, a_{n-1}, 2[\sqrt{d}], \overline{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]}].$$

That is, the repeating part is $\{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]\}$. The continued fraction has the form claimed.

Note that the infinite continued fraction for an irrational number is unique. So $a_1, \ldots, a_{n-1}$ are the same a's we'd have obtained if we'd computed the continued fraction for $\sqrt{d}$ using the standard algorithm.

This completes the proof of part (a).

(b) By the proof of (a), the continued fraction for $\sqrt{d}+[\sqrt{d}]$ is purely periodic. So suppose that

$$\sqrt{d}+[\sqrt{d}] = [\overline{a_0, a_1, \ldots, a_{n-1}}].$$

I showed in the proof of (a) that

$$\sqrt{d} = \left[[\sqrt{d}], \overline{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]}\right].$$

Since $\sqrt{d}+[\sqrt{d}]$ is purely periodic, the second theorem of Galois shows that $\dfrac{-1}{\overline{\sqrt{d}+[\sqrt{d}]}} = \dfrac{1}{\sqrt{d}-[\sqrt{d}]}$ is purely periodic, and

$$\frac{1}{\sqrt{d}-[\sqrt{d}]} = [\overline{a_{n-1}, \ldots, a_1, a_0}].$$

But

$$\sqrt{d}-[\sqrt{d}] = [0, \overline{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]}] = \frac{1}{\overline{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]}}.$$

Hence,

$$\frac{1}{\sqrt{d}-[\sqrt{d}]} = [\overline{a_1, \ldots, a_{n-1}, 2[\sqrt{d}]}].$$

Equating the terms in two expressions for $\dfrac{1}{\sqrt{d}-[\sqrt{d}]}$, I obtain the equations

$$a_1 = a_{n-1}, \ldots, a_{n-1} = a_1.$$

This shows that the sequence $\{a_1, \ldots, a_{n-1}\}$ is palindromic. □

The following refinement of the Theorem will be used in our discussion of the Fermat-Pell equation. We refer to the algorithm for the continued fraction of a quadratic irrational — "the quadratic irrational algorithm", for short — described at the beginning of this section.

Proposition. Let d be a positive integer which is not a perfect square. Suppose the quadratic irrational algorithm is applied to:

(1) $\sqrt{d}$, producing sequences m_i, s_i, x_i, and a_i.

(2) $[\sqrt{d}]+\sqrt{d}$, producing sequences m_i', s_i', x_i', and a_i'.

Then:

(a) $m_i = m_i'$ for $i \ge 1$.

(b) $s_i = s_i'$ for $i \ge 0$.

(c) $x_i = x_i'$ for $i \ge 1$.

(d) $a_i = a_i'$ for $i \ge 1$, and $2a_0 = a_0'$.

Proof. In the proof of part (a) of the preceding theorem, we showed that

$$\sqrt{d} = \left[[\sqrt{d}], \overline{a_1, \ldots, a_{n-1}, 2a_0}\right] = \left[[\sqrt{d}], a_1, \ldots, a_{n-1}, 2[\sqrt{d}], a_1, \ldots, a_{n-1}, 2[\sqrt{d}], \ldots\right].$$

$$[\sqrt{d}]+\sqrt{d} = \left[\overline{[2\sqrt{d}], a_1, \ldots, a_{n-1}}\right] = \left[2[\sqrt{d}], a_1, \ldots, a_{n-1}, 2[\sqrt{d}], a_1, \ldots, a_{n-1}, 2[\sqrt{d}], \ldots\right].$$

Comparing terms, we see that (d) follows immediately.

In applying the quadratic irrational algorithm as in (1) and (2), we start with

$$m_0 = 0, \quad s_0 = 1, \quad x_0 = \sqrt{d}, \quad a_0 = [\sqrt{d}],$$

$$m'_0 = [\sqrt{d}], \quad s'_0 = 1, \quad x'_0 = [\sqrt{d}] + \sqrt{d}, \quad a'_0 = 2[\sqrt{d}].$$

In particular, we have $s_0 = s'_0 = 1$.
In addition,

$$m_1 = a_0 s_0 - m_0 = [\sqrt{d}] \cdot 1 - 0 = [\sqrt{d}],$$

$$s_1 = \frac{d - m_1^2}{s_0} = \frac{d - [\sqrt{d}]^2}{1} = d - [\sqrt{d}]^2$$

,

$$x_1 = \frac{m_1 + \sqrt{d}}{s_1} = \frac{[\sqrt{d}] + \sqrt{d}}{d - [\sqrt{d}]^2},$$

$$m'_1 = a'_0 s'_0 - m'_0 = 2[\sqrt{d}] \cdot 1 - [\sqrt{d}] = [\sqrt{d}],$$

$$s'_1 = \frac{d - m_1'^2}{s'_0} = \frac{d - [\sqrt{d}]^2}{1} = d - [\sqrt{d}]^2$$

,

$$x'_1 = \frac{m'_1 + \sqrt{d}}{s'_1} = \frac{[\sqrt{d}] + \sqrt{d}}{d - [\sqrt{d}]^2}.$$

Thus, $m_1 = m'_1$, $s_1 = s'_1$, and $x_1 = x'_1$.
Suppose inductively that for some $i \geq 1$,

$$m_i = m'_i, \quad s_i = s'_i.$$

Then

$$m_{i+1} = a_i s_i - m_i = a'_i s'_i - m'_i = m'_{i+1},$$

$$s_{i+1} = \frac{d - m_{i+1}^2}{s_i} = \frac{d - m_{i+1}'^2}{s_i i} = s'_{i+1},$$

$$x_{i+1} = \frac{m_{i+1} + \sqrt{d}}{s_{i+1}} = \frac{m'_{i+1} + \sqrt{d}}{s'_{i+1}} = x'_{i+1}.$$

By induction, $m_i = m'_i$ for $i \geq 1$, $s_i = s'_i$ for $i \geq 0$, and $x_i = x'_i$ for $i \geq 1$. This proves (a), (b), and (c).
□

Here's an example which illustrates the result. Here are the first few values of m, s, x, and a for $\sqrt{42}$:

m	s	x	a
0	1	6.48074 . . .	6
6	6	2.08012 . . .	2
6	1	12.48074 . . .	12
6	6	2.08012 . . .	2
6	1	12.48074 . . .	12
6	6	2.08012 . . .	2
6	1	12.48074 . . .	12
6	6	2.08012 . . .	2
6	1	12.48074 . . .	12
6	6	2.08012 . . .	2
6	1	12.48074 . . .	12

And here are the corresponding values of m, s, x, and a for $[\sqrt{42}] + \sqrt{42}$:

m	s	x	a
6	1	$12.48074\ldots$	12
6	6	$2.08012\ldots$	2
6	1	$12.48074\ldots$	12
6	6	$2.08012\ldots$	2
6	1	$12.48074\ldots$	12
6	6	$2.08012\ldots$	2
6	1	$12.48074\ldots$	12
6	6	$2.08012\ldots$	2
6	1	$12.48074\ldots$	12
6	6	$2.08012\ldots$	2
6	1	$12.48074\ldots$	12

You can see that the m's, x's, and a's agree beginning with the second line (index 1), and the s's agree from the start (index 0).

Proposition. Let d be a positive integer which is not a perfect square, and suppose $[\sqrt{d}] + \sqrt{d}$ has period n:

$$[\sqrt{d}] + \sqrt{d} = [\overline{a_0, a_1, \ldots, a_{n-1}}].$$

Suppose the quadratic irrational algorithm is applied to $[\sqrt{d}] + \sqrt{d}$, generating sequences m_i, s_i, and x_i. Then:

(a) x_i is purely periodic with period n.

(b) $x_0 \neq x_1, \ldots, x_{n-1}$.

Proof. (a) We saw in the derivation of the quadratic irrational algorithm that the x_i's produced by the algorithm are the same as the x_i's in the general continued fraction algorithm — that is, x_i in the quadratic irrational algorithm is the i^{th} partial quotient of $[\sqrt{d}] + \sqrt{d}$. The a_i's are purely periodic with period n:

$$x = x_0 = [a_0, a_1, \ldots a_{n-1}, a_0, a_1, \ldots a_{n-1}, \ldots].$$

But

$$[a_0, a_1, \ldots a_{n-1}, a_0, a_1, \ldots a_{n-1}, \ldots] = [a_0, a_1, \ldots a_{n-1}, x_n].$$

Comparing the two expansions, we see that

$$x_n = [a_0, a_1, \ldots a_{n-1}, \cdots] = x_0.$$

On the other hand, if the x_i's repeat in n steps, then $a_i = [x_i]$ shows that the a's repeat in n steps as well. So the a's and x's have the same period, and x_i is purely periodic with period n.

(b) Suppose on the contrary that $x_0 = x_i$ where $1 \leq i \leq n - 1$. Then

$$[\overline{a_0, a_1, \ldots, a_{n-1}}] = [\overline{a_i, a_{i+1}, \ldots, a_{i+n-1}}].$$

Thus,

$$\begin{aligned} a_i &= a_0 \\ a_{i+1} &= a_1 \\ &\vdots \end{aligned}$$

Hence,
$$[\sqrt{d}] + \sqrt{d} = [\overline{a_0, a_1, \ldots, a_{i-1}}].$$

But then $[\sqrt{d}] + \sqrt{d}$ has period $i < n$, contradicting our assumption that the period is n. □

Note that by periodicity, $x_0 \neq x_i$ if $i = 1, \ldots, n-1 \pmod{n}$.

The next result will be used in our discussion of the Fermat-Pell equation.

Theorem. Let d be a positive integer which is not a perfect square. Suppose the period of $\sqrt{d}$ is n, so

$$\sqrt{d} = [a_0, \overline{a_1 \ldots, a_{n-1}, 2a_0}].$$

If the quadratic irrational algorithm is applied to $\sqrt{d}$, generating sequences m_i, s_i, x_i, then:

(a) $s_i = 1$ if and only if $i = kn$ for $k \geq 0$.

(b) $s_i \neq -1$ for all i.

Proof. (a) The continued fraction for $[\sqrt{d}] + \sqrt{d}$ is purely periodic with period n. Consequently, if the quadratic irrational algorithm is applied for $[\sqrt{d}] + \sqrt{d}$ generating sequences m_i', s_i', x_i', then $x_0' = x_{kn}'$ for $k \geq 0$.

In our discussion of the quadratic irrational algorithm, we showed that $x_i' = x_j'$ if and only if $(m_i', s_i') = (m_j', s_j')$. Hence, $s_0' = s_{kn}'$ for $k \geq 0$. But the second-to-the-last proposition above shows that $s_i = s_i'$ for $i \geq 0$. Therefore, $s_{kn} = s_0 = 1$ for $k \geq 0$.

Conversely, suppose $s_i = 1$. Then

$$x_i = \frac{m_i + \sqrt{d}}{s_i} = m_i + \sqrt{d}.$$

If $i = 0$, we're done. Otherwise, $i \geq 1$, so again by the second-to-the-last proposition above,

$$x_i' = x_i = m_i + \sqrt{d}.$$

Since x_i' is generated from $[\sqrt{d}] + \sqrt{d}$, an earlier result shows that the sequence x_i' is purely periodic. By Galois's theorem, the conjugate of $x_i' = m_i + \sqrt{d}$ satisfies

$$-1 < m_i - \sqrt{d} < 0, \quad \text{or} \quad \sqrt{d} - 1 < m_i < \sqrt{d}.$$

Thus, $m_i = [\sqrt{d}]$. It follows that
$$x_i' = [\sqrt{d}] + \sqrt{d} = x_0'.$$

Therefore, $i = kn$ by the preceding result.

(b) Suppose on the contrary that $s_i = -1$. Then $i \geq 1$ (since $s_0 = 1$) and

$$x_i = \frac{m_i + \sqrt{d}}{s_i} = -m_i - \sqrt{d}.$$

Since $i \geq 1$, I have
$$x_i' = x_i = -m_i - \sqrt{d}.$$

But x_i' is purely periodic, so our results on purely periodic continued fractions give

$$-m_i - \sqrt{d} > 1 \quad \text{and} \quad -1 < -m_i + \sqrt{d} < 0.$$

The first inequality gives $-\sqrt{d} - 1 > m_i$, while the second inequality gives $\sqrt{d} < m_i$. Together, these give $\sqrt{d} < -\sqrt{d} - 1$, a contradiction. Therefore, $s_i \neq -1$ for all i. □

It isn't true in general that for period n we have $s_0 = s_{kn}$. For instance, here's a quadratic irrational with period 3:

$$\frac{-11 + \sqrt{37}}{-4} = [1, 4, \overline{2, 1, 3}].$$

We have:

m	s	x	a
-11	-4	1.2293093674254452	1
7	3	4.36092084343274	4
5	4	2.7706906325745546	2
3	7	1.2975375043283168	1
4	3	3.3609208434327393	3
5	4	2.7706906325745546	2
3	7	1.2975375043283168	1

We don't have $s_0 = 1$ and we don't have $s_0 = s_3$, $s_0 = s_6$, and so on.