Continued Fractions and Pell's Equation

The Mathematical Details

Hilary Term 2023

What follows below is mostly a summary of ideas from Chapters 3 and 4 of C. D. Olds, Continued Fractions, John Wiley & Sons, 1978.

1 Continued Fractions and Convergents

Every real number x can be written as a *continued fraction* in the form

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}} =: [a_0; a_1, a_2, a_3, \dots] , \tag{1}$$

where the a_k are all integers. Here a_0 may be negative or zero, but all other coefficients are positive.

In order to compute a continued fraction representation of x, define $\lfloor x \rfloor$ to be the *floor* of x (or the integer part of x, namely the closest integer to x when rounding down), and define $\{x\} = x - \lfloor x \rfloor$ to be the fractional part of x. Note that $0 \leq \{x\} < 1$. The continued fraction representation of x is $[\lfloor x \rfloor; a_1, a_2, a_3, \dots]$ where $[a_1; a_2, a_3, \dots]$ is the continued fraction representation of $1/\{x\}$.

The *convergents* of a continued fraction are the initial terms in the continued fraction, i.e.

$$a_0, \quad a_0 + \frac{1}{a_1}, \quad a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}}, \quad a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3}}} . \tag{2}$$

These can be written as rational approximations to x as p_n/q_n where we can see from (2) above that $p_0 = a_0$, $p_1 = a_0 a_1 + 1$, $q_0 = 1$, and $q_1 = a_1$.

Lemma 1. *The convergents of a continued fraction satisfy* $p_0 = a_0$, $p_1 = a_0 a_1 + 1$, $q_0 = 1$, $q_1 = a_1$ *and*

$$p_n = a_n p_{n-1} + p_{n-2} , \tag{3}$$

$$q_n = a_n q_{n-1} + q_{n-2} , \tag{4}$$

for $n \geq 2$.

Proof. The proof is by induction on n. When $n = 2$ we have, from (2),

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = \frac{a_2(a_1 a_0 + 1) + a_0}{1 + a_1 a_2} = \frac{p_2}{q_2} . \tag{5}$$

Similarly when $n = 2$, (3) and (4) give

$$p_2 = a_2 p_1 + p_0 = a_2(a_1 a_0 + 1) + a_0 \tag{6}$$

$$q_2 = a_2 q_1 + q_0 = a_2 a_1 + 1 \tag{7}$$

so (3) and (4) hold for $n = 2$.

Now assume that (3) and (4) hold for $n = 2, 3, \ldots, k$. We will show that this implies (3) and (4) hold for $n = k + 1$ and so the result is true by strong induction. Consider

$$\frac{p_{k+1}}{q_{k+1}} = [a_0; a_1, \ldots, a_k, a_{k+1}] \tag{8}$$

$$= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_k + \cfrac{1}{a_{k+1}}}}} \tag{9}$$

$$= \left[a_0; a_1, \ldots, \left(a_k + \frac{1}{a_{k+1}}\right)\right] . \tag{10}$$

Clearly changing the a_k entry to $a_k + 1/a_{k+1}$ does not change the values of $p_0, p_1, \ldots p_{k-1}$ or $q_0, q_1, \ldots q_{k-1}$ but does change p_k and q_k so we have

$$\frac{p_{k+1}}{q_{k+1}} = \left[a_0; a_1, \ldots, \left(a_k + \frac{1}{a_{k+1}}\right)\right] \tag{11}$$

$$= \frac{(a_k + 1/a_{k+1})p_{k-1} + p_{k-2}}{(a_k + 1/a_{k+1})q_{k-1} + q_{k-2}} \tag{12}$$

where we have used (3) and (4) with $n = k$ with a_k replaced by $a_k + 1/a_{k+1}$. Rearranging (12) gives

$$\frac{p_{k+1}}{q_{k+1}} = \frac{a_{k+1}(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{a_{k+1}(a_k q_{k-1} + q_{k-2}) + q_{k-1}} \tag{13}$$

$$= \frac{a_{k+1} p_k + p_{k-1}}{a_{k+1} q_k + q_{k-1}} \tag{14}$$

by the inductive hypothesis. Hence (3) and (4) hold for $n = k + 1$ as required. □

Lemma 2. *The numerators and denominators of the convergents satisfy*

$$p_{n+1} q_n - p_n q_{n+1} = (-1)^n \tag{15}$$

for $n = 0, 1, 2, \ldots$.

Proof. The proof is again by induction on n. For the base case we have $p_0 = a_0$, $p_1 = a_0a_1 + 1$, $q_0 = 1$, $q_1 = a_1$ so when $n = 0$

$$p_{n+1}q_n - p_nq_{n+1} = p_1q_0 - p_0q_1 = a_0a_1 + 1 - a_0a_1 = 1 = (-1)^0 , \quad (16)$$

so (15) holds when $n = 0$. Now suppose (15) holds for $n = k$ then, by definition of p_{k+2} from (3) and of q_{k+2} from (4), we have

$$\begin{aligned} p_{k+2}q_{k+1} - p_{k+1}q_{k+2} &= (a_{k+2}p_{k+1} + p_k)q_{k+1} - p_{k+1}(a_{k+2}q_{k+1} + q_k) && (17)\\ &= p_kq_{k+1} - p_{k+1}q_k && (18)\\ &= -(-1)^k \quad \text{by the inductive hypothesis} && (19)\\ &= (-1)^{k+1} . && (20)\end{aligned}$$

Hence the result follows by induction. □

Lemma 3. *For each value of k, the integers p_k and q_k are coprime.*

Proof. Suppose that p_k and q_k have a common integer factor t so we may write $p_k = t\tilde{p}_k$ and $q_k = t\tilde{q}_k$ for some integers $\tilde{p}_k$ and $\tilde{q}_k$. By Lemma 2 we have

$$\begin{aligned} (-1)^k &= p_{k+1}q_k - p_kq_{k+1} && (21)\\ &= t(p_{k+1}\tilde{q}_k - \tilde{p}_kq_{k+1}) . && (22)\end{aligned}$$

Since t and $p_{k+1}\tilde{q}_k - \tilde{p}_kq_{k+1}$ are integers, the only way that their product can be $(-1)^k$ is if both terms are 1 or -1. Hence $t = \pm 1$ and p_k and q_k are coprime. □

2 Quadratic Irrationals

A *quadratic irrational* is an irrational real root of a quadratic equation with integer coefficients. Every quadratic irrational may be written in the form

$$x = \frac{P + \sqrt{D}}{Q} , \quad (23)$$

where $P, Q, D \in \mathbb{Z}$, $D > 0$ is not a perfect square and Q divides $P^2 - D$. Note that if Q does not divide $P^2 - D$ we may re-write

$$\begin{aligned} x &= \frac{PQ + \sqrt{DQ^2}}{Q^2} && (24)\\ x &= \frac{\tilde{P} + \sqrt{\tilde{D}}}{\tilde{Q}} , && (25)\end{aligned}$$

and then

$$\tilde{P}^2 - \tilde{D} = (P^2 - D)\tilde{Q} , \quad (26)$$

so $\tilde{Q}$ divides $\tilde{P}^2 - \tilde{D}$.

With this form, x is a root of the polynomial

$$\left(x - \frac{P + \sqrt{D}}{Q}\right)\left(x - \frac{P - \sqrt{D}}{Q}\right) = 0 , \quad (27)$$

which is equivalent to

$$Qx^2 - 2Px + \frac{P^2 - D}{Q} = 0 . \tag{28}$$

This polynomial has integer coefficients since Q divides $P^2 - D$.

2.1 Reduced Quadratic Irrationals

We say that x is a *reduced quadratic irrational* if x is a quadratic irrational satisfying $x > 1$ and $-1 < x' < 0$ where $x' = (P - \sqrt{D})/Q$.

Lemma 4. *For a fixed D there is a finite number of reduced quadratic irrationals.*

Proof. If x is a reduced quadratic irrational we find the following conditions on P and Q:

1. Since $x > 1$ and $x' < 0$ we have $x > x'$ which means $Q > 0$.
2. Since $x > 1$ and $x' > -1$ we have $x + x' > 0$ which means $P > 0$.
3. Since $x > 1$ we have $P + \sqrt{D} > Q$.
4. Since $x' < 0$ we have $P < \sqrt{D}$.
5. Since $x' > -1$ we have $\sqrt{D} - P < Q$.

We can combine these to get $0 < P < \sqrt{D}$ and $0 < Q < P + \sqrt{D} < 2\sqrt{D}$. Hence, for a fixed D, there are finitely many integer values of P satisfying $0 < P < \sqrt{D}$ and finitely many integer values of Q satisfying $0 < Q < 2\sqrt{D}$ so we can conclude there is a finite number of reduced quadratic irrationals associated with any given D. □

Lemma 5. *If α_n is a reduced quadratic irrational and we write $\alpha_n = \lfloor \alpha_n \rfloor + 1/\alpha_{n+1}$ then α_{n+1} is also a reduced quadratic irrational with the same subject of the square root.*

Proof. First we show that $\alpha_{n+1} > 1$ and $-1 < \alpha'_{n+1} < 0$. We have

$$\frac{1}{\alpha_{n+1}} = \alpha_n - \lfloor \alpha_n \rfloor , \tag{29}$$

and since $0 < \alpha_n - \lfloor \alpha_n \rfloor < 1$ we have $0 < 1/\alpha_{n+1} < 1$ which gives $\alpha_{n+1} > 1$.

Also

$$(\alpha_n - \lfloor \alpha_n \rfloor)' = \left(\frac{1}{\alpha_{n+1}} \right)' , \tag{30}$$

and so

$$-\frac{1}{\alpha'_{n+1}} = \lfloor \alpha_n \rfloor - \alpha'_n . \tag{31}$$

Now $-1 < \alpha' < 0$ and $\lfloor \alpha_n \rfloor \geq 1$ (since $\alpha_n > 1$) and so

$$-\frac{1}{\alpha'_{n+1}} = \lfloor \alpha_n \rfloor - \alpha'_n > 1 , \tag{32}$$

which gives $-1 < \alpha'_{n+1} < 0$.

Now we show that α_{n+1} takes the form of a quadratic irrational. Write $\alpha_n = (P_n + \sqrt{D})/Q_n$ so that the solutions of

$$Q_n x^2 - 2P_n x + \frac{P_n^2 - D}{Q_n} = 0 \qquad (33)$$

are $x = \alpha_n$ and $x = \alpha'_n$. Substitute $x = \alpha_n == \lfloor \alpha_n \rfloor + 1/\alpha_{n+1}$ into (33) to get

$$Q_n \left(\lfloor \alpha_n \rfloor + 1/\alpha_{n+1}\right)^2 - 2P_n \left(\lfloor \alpha_n \rfloor + 1/\alpha_{n+1}\right) + \frac{P_n^2 - D}{Q_n} = 0\,. \qquad (34)$$

We can rearrange this to get a quadratic equation in α_{n+1}:

$$\alpha_{n+1}^2 \left(\frac{(\lfloor \alpha_n \rfloor Q_n - P_n)^2}{Q_n} - \frac{D}{Q_n} \right) + 2\alpha_{n+1}\left(Q_n \lfloor \alpha_n \rfloor - P_n\right) + Q_n = 0\,. \qquad (35)$$

This has the root

$$\alpha_{n+1} = \frac{P_n - Q_n \lfloor \alpha_n \rfloor + \sqrt{D}}{\lfloor \alpha_n \rfloor^2 Q_n - 2\lfloor \alpha_n \rfloor P_n + (P_n^2 - D)/Q_n} \qquad (36)$$

$$= \frac{P_{n+1} + \sqrt{D}}{Q_{n+1}}\,, \qquad (37)$$

where we took the positive square root in the quadratic equation formula. Taking the negative square root would give α'_{n+1}.

In (37) we have $P_{n+1} = P_n - Q_n \lfloor \alpha_n \rfloor$ which is an integer. Also

$$Q_{n+1} = \lfloor \alpha_n \rfloor^2 Q_n - 2\lfloor \alpha_n \rfloor P_n + \frac{P_n^2 - D}{Q_n} \qquad (38)$$

is an integer since Q_n divides $P_n^2 - D$. We can rewrite (38) as

$$Q_{n+1} = \frac{(\lfloor \alpha_n \rfloor Q_n - P_n)^2 - D}{Q_n} \qquad (39)$$

$$= \frac{P_{n+1}^2 - D}{Q_n}\,. \qquad (40)$$

Thus we see that Q_{n+1} divides $P_{n+1}^2 - D$ so α_{n+1} is a quadratic irrational with D as the subject of the square root. □

Lemma 6. *If x is a reduced quadratic irrational, then its continued fraction expansion is purely periodic, i.e.* $x = [\overline{a_0; a_1, \ldots, a_{m-1}}]$.

Proof. Recall that to compute the continued fraction form of x we perform the following steps:

1. Set $x_0 = x$
2. `for` $k = 0, 1, 2, \ldots$

$$a_k = \lfloor x_k \rfloor$$

$$x_{k+1} = \frac{1}{\{x_k\}}$$

`end`

Since $x_0 = x$ is a reduced quadratic irrational, Lemma 5 tells us that all x_k are reduced quadratic irrationals with the same subject of the square root. The Lemma 4 tells us that there are finitely many such reduced quadratic irrationals and so there must be integers j and k with $j < k$ such that $x_j = x_k$. Clearly then $a_j = a_k$ and $x_{j+1} = x_{k+1}$ etc so that the sequence of a's repeats.

Now we need to show that the repeating pattern starts at a_0. We have

$$x_j = \frac{1}{\{x_{j-1}\}} = \frac{1}{x_{j-1} - \lfloor x_{j-1} \rfloor} = \frac{1}{x_{j-1} - a_{j-1}} , \tag{41}$$

and so

$$x_{j-1} - a_{j-1} = \frac{1}{x_j} . \tag{42}$$

The same equation also holds for x_{k-1} so, using the fact that $x_j = x_k$ we have

$$x_{j-1} - a_{j-1} = x_{k-1} - a_{k-1} , \tag{43}$$

$$x'_{j-1} - a_{j-1} = x'_{k-1} - a_{k-1} . \tag{44}$$

Since x_{j-1} and x_{k-1} are reduced quadratic irrationals, it follows that $x'_{j-1}, x'_{k-1} \in (-1, 0)$ and $a_{j-1}, a_{k-1} \in \mathbb{Z}$. Thus $x'_{j-1} = x'_{k-1}$ and $a_{j-1} = a_{k-1}$. We can then repeat this argument to see $x_{j-2} = x_{k-2}$ and finally $x_0 = x_{k-j}$. Hence if $m > 0$ is the smallest positive integer such that $x_m = x_0$, we have $x_{m+i} = x_i$ and $a_{m+i} = a_i$ for all $i \in \mathbb{N}$. So $x = [\overline{a_0; a_1, \ldots, a_{m-1}}]$. □

Lemma 7. *If $D \in \mathbb{N}$ and D is not a perfect square then $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$.*

Proof. Since D is not a perfect square, $D > 1$ so $\sqrt{D} > 1$ and $-\sqrt{D} < -1$ so $\sqrt{D}$ is not reduced. However, if we set $x = a_0 + \sqrt{D}$ where $a_0 = \lfloor \sqrt{D} \rfloor$, then x is reduced. Hence by Lemma 6

$$a_0 + \sqrt{D} = [\overline{2a_0; a_1, a_2, \ldots, a_{m-1}}] \tag{45}$$

and hence $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$. □

3 Pell's Equation

Pell's equation is

$$x^2 - Dy^2 = 1 . \tag{46}$$

We are interested in finding integer solutions x, y for $D \in \mathbb{N}$ in the case where D is not a perfect square.

Theorem 1. *Let the continued fraction expansion of $\sqrt{D}$ be $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$. If the length of the period, m, is even then $(x, y) = (p_{m-1}, q_{m-1})$ is a solution of Pell's equation. If m is odd then $(x, y) = (p_{2m-1}, q_{2m-1})$ is a solution of Pell's equation.*

Proof. We have

$$\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}] \tag{47}$$

$$= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{m-1} + \cfrac{1}{a_0 + \sqrt{D}}}}} \tag{48}$$

$$= [a_0; a_1, a_2, \ldots, a_{m-1}, a_0 + \sqrt{D}] \tag{49}$$

$$= \frac{(a_0 + \sqrt{D})p_{m-1} + p_{m-2}}{(a_0 + \sqrt{D})q_{m-1} + q_{m-2}}, \tag{50}$$

using the same idea as in the proof of Lemma 1. We can rearrange (50) to get

$$Dq_{m-1} + \sqrt{D}(a_0 q_{m-1} + q_{m-2}) = a_0 p_{m-1} + p_{m-2} + \sqrt{D} p_{m-1} . \tag{51}$$

Now decompositions of the form $\alpha + \beta\sqrt{D}$ are unique so (51) gives

$$Dq_{m-1} = a_0 p_{m-1} + p_{m-2} , \tag{52}$$

$$a_0 q_{m-1} + q_{m-2} = p_{m-1} . \tag{53}$$

We can rearrange to get

$$p_{m-2} = Dq_{m-1} - a_0 p_{m-1} , \tag{54}$$

$$q_{m-2} = p_{m-1} - a_0 q_{m-1} . \tag{55}$$

Now recall from Lemma 2 that $p_{n+1}q_n - p_n q_{n+1} = (-1)^n$ for all $n \geq 0$. Set $n = m - 2$ to get

$$p_{m-1}q_{m-2} - p_{m-2}q_{m-1} = (-1)^m . \tag{56}$$

Using (54) and (55) gives

$$(-1)^m = p_{m-1}(p_{m-1} - a_0 q_{m-1}) - q_{m-1}(Dq_{m-1} - a_0 p_{m-1}) \tag{57}$$

$$= p_{m-1}^2 - Dq_{m-1}^2 . \tag{58}$$

Hence if m is even $p_{m-1}^2 - Dq_{m-1}^2 = 1$ and $(x, y) = (p_{m-1}, q_{m-1})$ is a solution of Pell's equation. If m is odd, we have $p_{m-1}^2 - Dq_{m-1}^2 = -1$.

Note that we could have written (48) by going to the end of the second period so

$$\sqrt{D} = [a_0; a_1, a_2, \ldots, a_{2m-1}, a_0 + \sqrt{D}] . \tag{59}$$

Then the same argument as above gives $p_{2m-1}^2 - Dq_{2m-1}^2 = (-1)^{2m} = 1$ and so $(x, y) = (p_{2m-1}, q_{2m-1})$ is a solution of Pell's equation. □

In fact the ideas in the proof can be generalised to give $p_{km-1}^2 - Dq_{km-1}^2 = (-1)^{km}$, so Pell's equation has infinitely many solutions.