Maths for Project Euler

Jes Modian

August 4, 2023

Abstract

Project Euler is a website that provides hundreds of problems that require programming to solve, and some of them are so difficult that there is no way to independently come up with the maths used for the problems. That's why I like to cheat a bit and read what maths are used in other people's solutions.

Contents

1 Number theory

1.1 Pythagorean triple

(Related problem: Q9, Q39, Q75, Q86)

Definition 1.1.1. A **Pythagorean triple** consists of three positive integers a, b, c such that

$$a^2 + b^2 = c^2$$

For example, $3^2 + 4^2 = 5^2$, so $(3, 4, 5)$ is a Pythagorean triple.

Definition 1.1.2. A **primitive Pythagorean triple** is a Pythagorean triple in which a, b, c are coprime. (In other words, $\gcd(a, b, c) = 1$.)

For example, $(3, 4, 5)$ is a primitive Pythagorean triple, but $(6, 8, 10)$ is not.

1.1.1 Generating a triple

Theorem 1.1.1. (**Euclid's formula**) Given a pair of integers m , n with $m > n > 0$, the formula

$$a = m^2 - n^2 \ , \ b = 2mn \ , \ c = m^2 + n^2$$

form a Pythagorean triple. The triples generated is primitive if and only if m and n are coprime and one of them is even.
Moreoever, every primitive Pythagorean triple can be expressed by Euclid's formula . The n, m are unqiue for a given triple.

Proof. [1] First, note that if $a = m^2 - n^2$, $b = 2mn$, $c = m^2 + n^2$, then

$$\begin{aligned} a^2 + b^2 &= (m^2 - n^2) + (2mn)^2 \\ &= m^4 - 2m^2n^2 + n^4 + 4m^2n^2 \\ &= m^4 + 2m^2n^2 + n^4 \\ &= (m^2 + n^2)^2 \\ &= c^2 \end{aligned}$$

Thus a, b, c form a Pythagorean triple.

Now we prove that every primitive Pythagorean triple can be expressed by Euclid's formula.

All such primitive triples can be written as (a, b, c) where $a^2 + b^2 = c^2$ and a, b, c are coprime. Thus, a, b, c are pairwise coprime (which means $\gcd(a, b) = \gcd(a, c) = \gcd(b, c) = 1$) .
To see why they are pairwise coprime, suppose two of the numbers are not coprime, say a, b , which have a common factor $k > 1$. Let $a = kr$ and $b = ks$. Then

$$\begin{aligned} a^2 + b^2 &= c^2 \\ (kr)^2 + (ks)^2 &= c^2 \\ k^2(r^2 + s^2) &= c^2 \\ k\sqrt{r^2 + s^2} &= c \end{aligned}$$

This means k is also a factor of c , so a, b, c is not a primitive triple, and we have a contradiction. Thus a, b, c must be pairwise coprime.

Now back to business. As a and b are coprime, at least one of them is odd, so we may suppose a is odd (if not, just let a and b switch place) . Then b must be even, and c must be odd.
To see why, suppose b is odd. Then c is even, and we can let $a = 2k + 1$, $b = 2l + 1$ for some non-negative integer k, l . Then

$$\begin{aligned} a^2 + b^2 &= (2k + 1)^2 + (2l + 1)^2 \\ &= 4k^2 + 4k + 1 + 4l^2 + 4l + 1 \\ &= 4(k^2 + k + l^2 + l) + 2 \end{aligned}$$

We see that $a^2 + b^2$ is congruent to 2 modulo 4, but c^2 must be a multiple of 4 (since c is even), thus it is impossible that $a^2 + b^2 = c^2$, and there is a contradiction.

This implies that b must be even, so c is odd. (odd + even = odd)

From $a^2 + b^2 = c^2$ we obtain $c^2 - a^2 = b^2$ and hence $(c - a)(c + a) = b^2$. Then $\dfrac{(c + a)}{b} = \dfrac{b}{(c - a)}$.

Since $\dfrac{(c + a)}{b}$ is rational, we set it equal to $\dfrac{m}{n}$ in lowest terms. (In other words, $\dfrac{m}{n}$ is a reduced fraction and $\gcd(m, n) = 1$. There must uniquely exist such m and n since every fraction has a reduced form.) Thus $\dfrac{(c - a)}{b} = \dfrac{n}{m}$, being the reciprocal of $\dfrac{(c + a)}{b}$. Then solving

$$\frac{c}{b} + \frac{a}{b} = \frac{m}{n}, \qquad \frac{c}{b} - \frac{a}{b} = \frac{n}{m}$$

for $\frac{c}{b}$ and $\frac{a}{b}$ gives

$$\frac{c}{b} = \frac{1}{2}(\frac{m}{n} + \frac{n}{m}) = \frac{m^2+n^2}{2mn} , \qquad \frac{a}{b} = \frac{1}{2}(\frac{m}{n} - \frac{n}{m}) = \frac{m^2-n^2}{2mn}$$

As $\frac{m}{n}$ is fully reduced, m and n are coprime, and they cannot be both even. Note that they also cannot be both odd.

To see why, suppose m and n are both odd. From $\frac{a}{b} = \frac{m^2-n^2}{2mn}$ we obtain

$$a(2mn) = b(m^2 - n^2)$$

Note that $m^2 - n^2$ is a multiple of 4 (because an odd square is congruent to 1 modulo 4), so LHS is also a multiple of 4. Note that mn is odd since both m and n is odd. This means a must be even to make LHS a multiple of 4. But this contradicts the earlier assumption that a is odd.

Thus one of m and n must be odd and the other is even, and the numerators of the two fractions $\frac{m^2+n^2}{2mn}$ and $\frac{m^2-n^2}{2mn}$ are odd ($\text{odd}^2 \pm \text{even}^2 = \text{odd} \pm \text{even} = \text{odd}$).

Then we claim that these fractions are fully reduced.

To see why, suppose they are not fully reduced. Then both the numerator and denominator are divisible by some prime p . Note that $p \neq 2$ since numerator is odd.

Since p divides the denominator $2mn$, it divides m or n (by Euclid's lemma), but it cannot divide both since m and n are coprime. Thus p can only divide either m or n . Thus, p does not divide the numerator $m^2 \pm n^2$, and there is a contradiction.

To see why p cannot divide $m^2 \pm n^2$, suppose (WLOG) p divides m and p divides $m^2 \pm n^2$. we can let $m = rp$ and $m^2 \pm n^2 = sp$.

So $(rp)^2 \pm n^2 = sp$ which means $n^2 = \pm p(s - r^2 p)$. Since p divides n^2 , p also divides n (by Euclid's lemma).

Thus it is impossible that p divides m and p divides $m^2 \pm n^2$ but p does not divide n .

Thus, when p can only divide either m or n , p doe not divide $m^2 \pm n^2$.

Since there is a contradiction, no such p exists, and the fractions $\frac{m^2 \pm n^2}{2mn}$ are reduced fractions.

Recall that we have

$$\frac{c}{b} = \frac{m^2+n^2}{2mn} , \qquad \frac{a}{b} = \frac{m^2-n^2}{2mn}$$

And since $\frac{a}{b}$ and $\frac{c}{b}$ are also reduced fractions (as a, b, c are pairwise coprime as shown in the beginning), we can equate numerators with numerators and denominators with denominators, giving Euclid's formula

$$a = m^2 - n^2 , \qquad b = 2mn , \qquad c = m^2 + n^2$$

So we have shown that every primitive triple has some m, n with $m > n > 0$ satisfying Euclid's formula.

——

Now we prove that for a given primitive triple, the m, n satisfying Euclid's formula are unique.

Suppose there are m, n and m', n' satisfying Euclid's formula, namely,

$$a = m^2 - n^2 = m'^2 - n'^2 , \qquad b = 2mn = 2m'n' , \qquad c = m^2 + n^2 = m'^2 + n'^2$$

Then $\frac{c+a}{b} = \frac{(m^2+n^2)+(m^2-n^2)}{2mn} = \frac{m}{n}$. Similarly, $\frac{c+a}{b} = \frac{m'}{n'}$. Thus, $\frac{m}{n} = \frac{m'}{n'}$

Suppose $m \neq m'$ and $n \neq n'$.

Since $\frac{m}{n}$ is a reduced fraction, $\frac{m'}{n'}$ must be a non-reduced fraction with the same value, which means $m' = km$ and $n' = kn$ for some $k > 1$.

To see why, note that we have $mn' = nm'$, so m divides nm' . Since m and n are coprime, we have m divides m' by Euclid's lemma (or called 'property of divisibility inheritence' in number thoery subsection of Toddler Probability). This means $m' = km$ for some k . Since $m \neq m'$, we have $k > 1$. Similarly we have $n' = kn$ for some $k > 1$.

But then $a = m'^2 - n'^2 = (km)^2 - (kn)^2 = k^2(m^2 - n^2)$, and
$b = 2m'n' = 2(km)(kn) = k^2(2mn)$.
We see that a and b share a common divisor $k^2 > 1$, but this is impossible since a, b, c are primitive triple, so a, b should be coprime. So there is a contradiction.
Thus, it can only be the case that $m = m'$ and $n = n'$.

——

($\Rightarrow$) Now we show that given some m, n with $m > n > 0$, if the triple generated by m, n is primitive, then m and n are coprime and one of them is even.

If the triple is primitive, then exactly one of a and b is odd, and c is even, as shown above.
Note that $b = 2mn$, so b must be even. That means $a = m^2 - n^2$ must be odd.
If m and n are both even or both odd, then a will be even (odd $-$ odd $=$ even , even $-$ even $=$ even), which cannot be true. Thus it can only be the case that exactly one of m and n is even.

To show that m and n are coprime, suppose they are not coprime. Then they have a common divisor $k > 1$. Let $m = kr$ and $n = ks$. Then

$$\begin{aligned} a &= m^2 - n^2 = (kr)^2 - (ks)^2 = k^2(r^2 - s^2) \\ b &= 2mn = 2(kr)(ks) \qquad\quad\;\, = k^2(2rs) \\ c &= m^2 + n^2 = (kr)^2 + (ks)^2 = k^2(r^2 + s^2) \end{aligned}$$

This means a , b and c have a common divisor $k^2 > 1$, so the triple (a, b, c) is not primitive, which is a contradiction.

($\Leftarrow$) Now we show that given some m, n with $m > n > 0$, if m and n are coprime and one of them is even, then the triple generated by m, n is primitive.

If m and n are coprime and (exactly) one of them is even, then first, note that $\dfrac{c}{b} = \dfrac{m^2 + n^2}{2mn}$. We repeat the argument as before to show that $\dfrac{m^2 + n^2}{2mn}$ is a reduced fraction, so $\dfrac{c}{b}$ is also reduced fraction (because $c = m^2 + n^2$ and $b = 2mn$), which means c, b are coprime, so a, b, c must be primitive triple. (because if a, b, c is not primitive, then c, b cannot be coprime.)

$\square$

1.2 Continued fractions

(Related: Q64, Q65)

(Most of this section and the next two subsection are copied from [2].)

Definition 1.2.1. A **simple continued fraction** is of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}} \tag{1}$$

where a_n are integers, and with the exception of a_0 , they are all positive.

For convenience, we write $[a_0; a_1, a_2, a_3, \ldots]$ to represent the continued fraction.

(In this article, when we say continued fraction, we refer to simple continued fraction, where the numerators are all 1.)

The a_n are referred to as *quotients* of the simple continued fraction.

If the sequence a_n is finite, the continued fraction is called finite continued fraction:

$$[a_0; a_1, a_2, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

If the sequence a_n is infinite, the continued fraction is called infinite continued fraction, and it can be defined as

$$[a_0; a_1, a_2, a_3, \ldots] = \lim_{n \to \infty} [a_0; a_1, a_2, \ldots, a_n]$$

1.2.1 Floor and fractional part of a number

The **floor** of a number x returns the greatest integer that is no greater than the number itself, and is denoted $\lfloor x \rfloor$.

The **fractional part** of a number x is defined as the number minus its floor, and is denoted $\{x\}$. That is

$$\{x\} = x - \lfloor x \rfloor \tag{2}$$

Note that

$$0 \le \{x\} < 1 \tag{3}$$

, and (for positive numbers) $\{x\}$ is just the part of x after the decimal point.

For example, $\{5\} = 0$, $\{2.48\} = 0.48$, $\{-8.3\} = 0.7$.

Note that dividing a number up into its floor and its fractional part is the only way that we can divide it into two parts with one part an integer and the other part equal or greater than 0 and still less than 1.

Lemma 1.2.1. If $\alpha = a + x = b + y$ where $a, b \in \mathbb{Z}$ and $0 \le x, y < 1$ then $a = b$ and $x = y$.

Proof. If

$$a + x = b + y$$

then

$$a - b = y - x$$

Note that $a - b \in \mathbb{Z}$ and $y - x \in (-1, 1)$. Because 0 is the only integer in the interval $(-1, 1)$ it follows that $a = b$ and $x = y$. ☐

1.2.2 Representations of number

To represent a number x using continued fraction, first get the floor of x and let it be a_0 . Then express the fractional part $\{x\}$ as the reciprocal of the reciprocal of the fractional part: $\frac{1}{\frac{1}{\{x\}}}$. Note that $\frac{1}{\{x\}} > 1$ (since $\{x\} < 1$) , so we can get its floor to be a_1 and the new fractional part is treated the same as the last. If x is rational, this process is repeated until there is no fractional parts left. If x is irrational, this process is repeated forever. (We'll prove this in the following theorems.)

Let's use $\frac{54}{19}$ as an example:

$$\frac{54}{19} = 2 + \frac{1}{\frac{19}{16}} = 2 + \cfrac{1}{1 + \frac{1}{\frac{16}{3}}} = 2 + \cfrac{1}{1 + \cfrac{1}{5 + \cfrac{1}{3}}} = [2; 1, 5, 3]$$

Theorem 1.2.2. Every rational number has exactly two finite simple continued fraction expansions, and every finite simple continued fraction expansion represents a rational number.

Proof. Let $\alpha \in \mathbb{Q}$ and divide it into its floor and fractional part.

If the fractional part of α is 0, α is an integer and one simple continued fraction expansion is $[\alpha] = \alpha$ and a second is $[\alpha - 1; 1] = (\alpha - 1) + \frac{1}{1}$.

These are the only two expansions. There are no more with one quotient since α only has one value. There are no more with two quotients because if the second quotient is not 1 then the expression consists of a fraction that is less than 1 and α would have a fractional part. Finally, for the same reason there are no expansions with three or more quotients. Thus there are only two ways to express an integer as a simple continued fraction:

$$[\alpha] = \alpha \qquad \text{or} \qquad [\alpha - 1; 1] = (\alpha - 1) + \frac{1}{1} \tag{4}$$

If the fractional part of α is not 0 then define the 1st *residue* of alpha, and in general, of any continued fraction, as the reciprocal of its fractional part.

$$\alpha = a_0 + \frac{1}{r_1}$$

where $a_0 \in \mathbb{Z}$ and is the floor of α , and r_1 is the 1st residue of α . Note that $1 < r_1$ because of equation (3).

Now if the k th residue is not an integer it has a fractional part and we can define the $(k+1)$ th residue recursively by the relationship

$$r_k = a_k + \frac{1}{r_{k+1}} \tag{5}$$

where $a_k = \lfloor r_k \rfloor$ and $r_{k+1} = \frac{1}{\{r_k\}}$.

Now if $1 < r_k$ then $a_k \in \mathbb{Z}^+$ and of course $1 < r_{k+1}$. Because $1 < r_1$ it follows by induction that $1 < r_n$ for each r_n which is defined.

It can be easily seen that if α is rational and not an integer then r_1 is rational, and also that if r_n is rational and not an integer then r_{n+1} is rational. Now let $\frac{b}{c} = r_k$ where r_k is not an integer and $b, c \in \mathbb{Z}^+$ with their only common factor being 1. That is, $\frac{b}{c}$ is in its lowest terms. Because r_k is not an integer $1 < c$. Then from equations (2), (3) and (5)

$$0 < r_k - a_k = \frac{b}{c} - a_k = \frac{b - a_k c}{c} = \frac{d}{c} = \frac{1}{r_{k+1}} < 1$$

where $0 < b - a_k c = d \in \mathbb{Z}$. Then

$$1 < r_{k+1} = \frac{c}{d}$$

and so

$$d < c$$

Now because r_k has a denominator of c and r_{k+1} has a denominator of d each subsequent rational residue has a smaller integer denominator when reduced to it lowest terms. So then the denominators of the residues form a decreasing sequence of integers that are all greater than zero. Eventually one of the denominators must be 1 and then the residue is an integer.

Let the first residue that is an integer be r_n . Then the recurrence relationship given in equation (5) will not hold as there is no fractional part of r_n . Instead let

$$r_n = a_n$$

Notice that again we have $a_k = \lfloor r_k \rfloor$.

At this point we have n different equations that look like

$$\begin{aligned} \alpha &= a_0 + \frac{1}{r_1} \\ &\vdots \\ r_{n-1} &= a_{n-1} + \frac{1}{r_n} \\ r_n &= a_n \end{aligned}$$

Combining these into one equation generates a simple continued fraction:

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}} = [a_0; a_1, \ldots, a_n]$$

To establish uniqueness, assume that

$$\alpha = [a_0; a_1, \ldots, a_n] = [b_0; b_1, \ldots, b_m]$$

wher each $b_i \in \mathbb{N}$. Also assume without the loss of generality that neither a_n or b_m are 1 . [1]

Now $0 < [0; b_m] < 1$. Also, because $1 < b_{m-1} + [0; b_m]$ it follows that $0 < [0; b_{m-1}, b_m] < 1$. One can prove by induction that

$$[0; b_j, b_{j+1}, \ldots, b_m] \in (0, 1)$$

for $j = 1 \ldots m$.

And because

$$a_0 + [0; a_1, \ldots, a_n] = b_0 + [0; b_1, \ldots, b_m]$$

, by Lemma 1.2.1, $a_0 = b_0$ and

$$a_1 + [0; a_2, \ldots, a_n] = b_1 + [0; b_2, \ldots, b_m]$$

. Continuing this way one finds that $a_0 = b_0$, $a_1 = b_1$, . . . , $a_n = b_m$ and also that $n = m$. So if the last partial quotient is not 1 then the simple continued fraction expansion is unique.

However, as stated earlier, an integer can be represented in two ways, given by equations (4). So α can be represented one of two ways as a simple continued fraction;

$$a = [a_0; a_1, \ldots, a_n] = [a_0; a_1, \ldots, a_n - 1, 1]$$

Proving the second part of the theorem is trivial; it is evident that any finite simple continued fraction represents a rational number as the simple continued fraction can just be simplified from the lower right-hand corner upwards to generate the rational number it represents.

□

Theorem 1.2.3. Every irrational number has an infinite simple continued fraction expansion.

[1] If $a_n = 1$ then $r_{n-1} = a_{n-1} + 1 \in \mathbb{N}$ and so we can replace $[a_0; a_1, \ldots, a_n] = [a_0; a_1, \ldots, a_{n-1}, 1]$ with $[a_0; a_1, \ldots, a_{n-1} + 1]$. Likewise for b_m .

Proof. Let $\alpha \in \mathbb{R}; \alpha \notin \mathbb{Q}$. Even though α is irrational, the residues are defined the same as in Theorem 1.2.2 In fact, because α is irrational its fractional part is not 0 and so there must be a 1st residue. Now it is obvious that r_1 is irrational, else α equals the sum of two rational numbers, and would not be irrational. Furthermore, for every irrational r_n, r_{n+1} is irrational, because of equation (5):

$$r_k = a_k + \frac{1}{r_{k+1}} \tag{5}$$

So all the residues are irrational. Thus there will never be a residue such that $r_k = a_k$ and so the irrational number has an infinite simple continued fraction expansion.

Now from the recursion formula of equation (5) an infinite number of equations are produced that look like

$$\begin{aligned} \alpha &= a_0 + \frac{1}{r_1} \\ r_1 &= a_1 + \frac{1}{r_2} \\ r_2 &= a_2 + \frac{1}{r_3} \\ &\vdots \end{aligned}$$

We can combine these into one equation and generate an infinite simple continued fraction:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}} = [a_0; a_1, a_2, a_3, \ldots]$$

□

1.2.3 Convergents

Definition 1.2.2. If we crop the (finite or infinite) expression $\alpha = [a_0; a_1, \ldots, a_n, a_{n+1}, \ldots]$ to the n th quotient we get the rational number $[a_0; a_1, \ldots, a_n]$. This is called the n th **convergent** of a number α .

Notice for α with a finite number of convergents, the last convergent is equal to α .

Let p_n and q_n denote respectively the numerator and denominator of the n th convergent in its lowest terms. We have

$$p_0 = a_0$$

$$q_0 = 1$$

The 1st convergent is

$$[a_0; a_1] = a_0 + \frac{1}{a_1} = \frac{a_1 a_0 + 1}{a_1}$$

and so

$$\begin{aligned} p_1 &= a_1 a_0 + 1 \\ q_1 &= a_1 \end{aligned}$$

Continuing in this fashion, the 2nd convergent is

$$[a_0; a_1, a_2] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = a_0 + \frac{a_2}{a_2 a_1 + 1} = \frac{a_2 a_1 a_0 + a_2 + a_0}{a_2 a_1 + 1}$$

and so

$$p_2 = a_2 a_1 a_0 + a_2 + a_0 \qquad = a_2 p_1 + p_0 \tag{6}$$
$$q_2 = a_2 a_1 + 1 \qquad = a_2 q_1 + q_0 \tag{7}$$

This can be generalized in the following theorem.

Theorem 1.2.4. The numerator and denominator of the n th convergent of a real number $\alpha = [a_0; a_1, \ldots]$ where $0 < a_i \in \mathbb{R}$ is given by

$$\begin{aligned} p_n &= a_n p_{n-1} + p_{n-2} \\ q_n &= a_n q_{n-1} + q_{n-2} \end{aligned} \tag{8}$$

and so the n th convergent is

$$\frac{p_n}{q_n} = [a_0; a_1, \ldots, a_n] \tag{9}$$

(This is not necessarily a simple continued fraction so each a_i is not necessarily an integer.)

Proof. From equations (6) and (7) we can see that the theorem holds for $n = 2$, for all $0 < a_0, a_1 \in \mathbb{R}$. Let us now assume that the theorem holds for $n = 1 \ldots k$ and then we will prove that it holds for $n = k + 1$ and so the result follows by strong induction. Specifically, assume

$$\begin{aligned} p_k &= a_k p_{k-1} + p_{k-2} \\ q_k &= a_k q_{k-1} + q_{k-2} \end{aligned}$$

and

$$\frac{p_k}{q_k} = [a_0; a_1, \ldots, a_k] \tag{10}$$

Now note that

$$\begin{aligned} [a_0; a_1, \ldots, a_k, a_{k+1}] &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{\left(a_k + \cfrac{1}{a_{k+1}}\right)}}} \\ &= \left[a_0; a_1, \ldots, a_k + \frac{1}{a_{k+1}}\right] \end{aligned} \tag{11}$$

Because $0 < a_k + \dfrac{1}{a_{k+1}} \in \mathbb{R}$, the only difference between the $(k+1)$ th convergent and the k th convergent is that $a_k + \dfrac{1}{a_{k+1}}$ is in place of a_k . So if we replace a_k in equation (10) we get an expression for $(k+1)$ th convergent. This is valid because it is evident from the recursion relationships of the numerators and denominators of the convergents, given from the induction assumption, that changing the value of a_k to $a_k + \dfrac{1}{a_{k+1}}$ will have no effect on the value of $p_{n-1}, p_{n-2}, \ldots$ or $q_{n-1}, q_{n-2}, \ldots$. Fromthe equations (10) and (11):

$$\begin{aligned} [a_0; a_1, \ldots, a_k, a_{k+1}] &= \frac{\left(a_k + \frac{1}{a_{k+1}}\right) p_{k-1} + p_{k-2}}{\left(a_k + \frac{1}{a_{k+1}}\right) q_{k-1} + q_{k-2}} \\ &= \frac{a_{k+1} a_k p_{k-1} + p_{k-1} + a_{k+1} p_{k-2}}{a_{k+1} a_k q_{k-1} + q_{k-1} + a_{k+1} q_{k-2}} \\ &= \frac{a_{k+1}(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{a_{k+1}(a_k q_{k-1} + q_{k-2}) + q_{k-1}} \\ &= \frac{a_{k+1} p_k + p_{k-1}}{a_{k+1} q_k + q_{k-1}} \\ &= \frac{p_{k+1}}{q_{k+1}} \end{aligned}$$

and the theorem is proved by the induction principle. □

It is common practice to define $p_{-2} = 0$, $p_{-1} = 1$, $q_{-2} = 1$ and $q_{-1} = 0$. Then the previous theorem can apply to $n = 0$ and $n = 1$, as

$$\begin{aligned}
p_0 &= a_0(1) + 0 = a_0 \\
q_0 &= a_0(0) + 1 = 1 \\
p_1 &= a_1(a_0) + 1 = a_1 a_0 + 1 \\
q_1 &= a_1(1) + 0 = a_1
\end{aligned}$$

Because in this section there are no requirements that a_n are integers, we can use the idea of residues to prove this corollary to the previous theorem.

Corollary 1.2.5. If r_n is the n th residue of α and the numerators and denominators of the convergents of α are defined as in Theorem 1.2.4 then

$$\alpha = \frac{r_n p_{n-1} + p_{n-2}}{r_n q_{n-1} + q_{n-2}} \tag{12}$$

Proof. When assembling a simple continued fraction from the recursion relationships given by equation (5), if we stop at the n th equation we will end up with the equation

$$\begin{aligned}
\alpha &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{r_n}}}} \\
&= [a_0; a_1, \ldots, a_{n-1}, r_n]
\end{aligned}$$

While this may not be a simple continued fraction because r_n may not be an integer, all the terms of the continued fraction are real, so Theorem 1.4 still applies. So in equation (9) we can replace a_n with r_n to make a modified form of the last convergent, and thus

$$\alpha = \frac{r_n p_{n-1} + p_{n-2}}{r_n q_{n-1} + q_{n-2}}$$

□

The formula derived in Theorem 1.2.4 lead us to a very important relation that is central in solving Pell's equation.

Theorem 1.2.6. If the numerators and denominators of the convergents of a continued fraction are defined as in Theorem 1.2.4 then

$$p_{n+1}q_n - p_n q_{n+1} = (-1)^n \tag{13}$$

Proof. Using the extended definitions of p_{-2}, p_{-1}, q_{-2} and q_{-1} we can establish the relation true for $n = -2$.

$$p_{-1}q_{-2} - p_{-2}q_{-1} = 1 \cdot 1 - 0 \cdot 0 = 1 = (-1)^{-2}$$

Now assume the relation true for $n = k$. That is

$$p_{k+1}q_k - p_k q_{k+1} = (-1)^k \tag{14}$$

Then consider when $n = k + 1$.

$$p_{k+2}q_{k+1} - p_{k+1}q_{k+2} = (a_{k+2}p_{k+1} + p_k)q_{k+1} - p_{k+1}(a_{k+2}q_{k+1} + q_k)$$

from equations (8)

$$\begin{aligned}
&= p_k q_{k+1} - p_{k+1} q_k \\
&= (-1)(p_{k+1}q_k - p_k q_{k+1}) \\
&= (-1)^{k+1}
\end{aligned}$$

by the induction assumption of equation (14). So the theorem is proved by induction. □

In a simple continued fraction the a_n are all integers and it is easy to see from the recnrrence relations given in equations (8) that all the p_n and q_n will be integers. They also have an important property described by a corollary of Theorem 1.2.6.

Corollary 1.2.7. When defined by equations (8), the numerators of each convergent of a simple continued fraction share no common factors with their corresponding denominator other than 1, and so the convergents are in their lowest terms.

Proof. Any common factor of p_n and q_n could be factored out of the left-hand side of equation (13) and so must also be a factor of the right-hand side. However, the only factors of the right-hand side are -1 and 1. Thus they share no common factors other than 1, and the convergents are in their lowest terms. ☐

1.2.4 (Extra) Convergence of infinite continued fractions

[3]

Let
$$c_n = [a_0; \ldots, a_n] = \frac{p_n}{q_n}$$
denote the n th convergent.

Theorem 1.2.8. (Extra of Theorem 1.2.6) If the numerators and denominators of the convergents of a continued fraction are defined as in Theorem 1.2.4 then

$$p_n q_{n-1} - q_n p_{n-1} = (-1)^{n-1}$$
$$p_n q_{n-2} - q_n p_{n-2} = (-1)^n a_n$$

Equivalently,
$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = (-1)^{n-1} \cdot \frac{1}{q_n q_{n-1}}$$
$$\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}} = (-1)^n \cdot \frac{a_n}{q_n q_{n-2}}$$

Proof. The first equation is just Theorem 1.2.6 with shifted indices.

For the second equation, we have

$$\begin{aligned} p_n q_{n-2} - q_n p_{n-2} &= (a_n p_{n-1} + p_{n-2}) q_{n-2} - p_{n-2}(a_n q_{n-1} + q_{n-2}) \\ &= a_n (p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) \\ &= (-1)^n a_n \end{aligned}$$

☐

Theorem 1.2.9. The even indexed convergents c_{2n} increase strictly with n , and the odd indexed convergents decrease strictly with n . Also, the odd index convergents c_{2n+1} are greater than all of the even indexed convergents c_{2m} .

Proof. The a_n are positive for $n \geq 1$, so the q_n are positive. By Theorem 1.2.8, for $n \geq 2$,

$$c_n - c_{n-2} = (-1)^n \cdot \frac{a_n}{q_n q_{n-2}}$$

which proves the first claim (because RHS will be positive if n is even, and negative if n is odd).

Suppose for the sake of contradiction that there exists integers r, m such that $c_{2m+1} < c_{2r}$. Theorem 1.2.8 implies that for $n \geq 1$,

$$c_n - c_{n-1} = (-1)^{n-1} \cdot \frac{1}{q_n q_{n-1}}$$

has sign $(-1)^{n-1}$, so for all $s \geq 0$ we have $c_{2s+1} > c_{2s}$. Thus it is impossible that $r = m$. If $r < m$, then by what we proved in the first paragraph, $c_{2m+1} < c_{2r} < c_{2m}$, a contradiction (with $s = m$). If $r > m$, then $c_{2r+1} < c_{2m+1} < c_{2r}$, which is also a contradiction (with $s = r$).

☐

Theorem 1.2.10. Let $a_0, a_1, \ldots$ be a sequence of integers such that $a_n > 0$ for all $n \geq 1$, and for each $n \geq 0$, set $c_n = [a_0; a_1, \ldots, a_n]$. Then $\lim_{n\to\infty} c_n$ exists.

Proof. For any $m \geq n$, the number c_n is a partial convergent of $[a_0; \ldots, a_m]$. By Theorem 1.2.9 the even convergents c_{2n} form a strictly increasing sequence and the odd convergents c_{2n+1} form a strictly decreasing sequence. Moreoever, the even convergents are all $\leq c_1$ and the odd convergents are all $\geq c_0$. Hence (let) $\alpha_0 = \lim_{n\to\infty} c_{2n}$ and $\alpha_1 = \lim_{n\to\infty} c_{2n+1}$ both exist (because bounded monotonic sequence converges), and $\alpha_0 \leq \alpha_1$. Finally, by Theorem 1.2.8,

$$|c_{2n} - c_{2n-1}| = \frac{1}{q_{2n}q_{2n-1}}m \leq \frac{1}{2n(2n-1)} \to 0$$

so $\alpha_0 = \alpha_1$. $\square$

Theorem 1.2.11. Let $\alpha \in \mathbb{R}$ be a real number. Then α is the value of the (possibly infinite) simple continued fraction $[a_0, a_1, a_2, \ldots]$ produced by the continued fraction procedure.

Proof. The theorem is obvious if the sequence is finite. Suppose the sequence is infinite. Let r_{n+1} be the $(n+1)$ th residue. Then

$$\alpha = [a_0, a_1, \ldots, a_n, r_{n+1}]$$

By Corollary 1.2.5, we have

$$\alpha = \frac{r_{n+1}p_n + p_{n-1}}{r_{n+1}q_n + q_{n-1}}$$

Then if $c_n = [a_0, a_1, \ldots, a_n]$, then

$$\begin{aligned}
\alpha - c_n &= \frac{r_{n+1}p_n + p_{n-1}}{r_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} \\
&= \frac{(r_{n+1}p_n + p_{n-1})q_n - (r_{n+1}q_n + q_{n-1})p_n}{q_n(r_{n+1}q_n + q_{n-1})} \\
&= \frac{p_{n-1}q_n - p_nq_{n-1}}{q_n(r_{n+1}q_n + q_{n-1})} \\
&= \frac{(-1)^n}{q_n(r_{n+1}q_n + q_{n-1})} \qquad \text{(by Theorem 1.2.8)}
\end{aligned}$$

$\square$

Thus

$$\begin{aligned}
|\alpha - c_n| &= \frac{1}{q_n(r_{n+1}q_n + q_{n-1})} \\
&< \frac{1}{q_n(a_{n+1}q_n + q_{n-1})} \qquad \text{(since } r_{n+1} > a_{n+1}\text{)} \\
&= \frac{1}{q_n \cdot q_{n+1}} \qquad \text{(by Theorem 1.2.4)} \\
&\leq \frac{1}{n(n+1)} \to 0
\end{aligned}$$

1.3 Quadratic irrationals

Definition 1.3.1. A **quadratic irrational** is an irrational number in the form $F + G\sqrt{M}$ where $F, G \in \mathbb{Q}$ and $M \in \mathbb{Z}^+$ and is not a perfect square.

We'll prove that every number in this form is indeed irrational.

Theorem 1.3.1. $\sqrt{M}$ is irrational if $M \in \mathbb{Z}^+$ and is not a perfect square.

Proof. (This is madness.) If $M \in \mathbb{Z}^+$ and is not a perfect square then there exists $n \in \mathbb{Z}^+$ such that $n < \sqrt{M} < n+1$ and so

$$n^2 < M < (n+1)^2 \tag{15}$$

If $\sqrt{M}$ is rational then it can be expressed as $\sqrt{M} = \frac{c}{d}$ where $c, d \in \mathbb{Z}^+$ and $\frac{c}{d}$ is in its lowest terms. So

$$\begin{aligned} M &= \frac{c^2}{d^2} \\ c^2 &= Md^2 \end{aligned} \tag{16}$$

Now from equations (15) and (16) we get

$$\begin{aligned} n^2d^2 < Md^2 &< (n+1)^2d^2 \\ n^2d^2 < c^2 &< (n+1)^2d^2 \\ nd < c &< (n+1)d \\ 0 < c - nd &< d \\ 0 < f &< d \end{aligned} \tag{17}$$

where $c - nd = f \in \mathbb{Z}^+$.

Again from equations (15) and (16) we get

$$\begin{aligned} n^2c^2 < Mc^2 &< (n+1)^2c^2 \\ n^2c^2 < M^2d^2 &< (n+1)^2c^2 \\ nc < Md &< (n+1)c \\ 0 < Md - nc &< c \\ 0 < g &< c \end{aligned} \tag{18}$$

where $Md - nc = g \in \mathbb{Z}^+$.

Then

$$\begin{aligned} g^2 - Mf^2 &= M^2d^2 - 2Mncd + n^2c^2 - M(c^2 - 2ncd + n^2d^2) \\ &= (M^2d^2 - Mc^2) - (2Mncd - 2Mncd) + (n^2c^2 - Mn^2d^2) \\ &= 0 \end{aligned}$$

So

$$M = \frac{g^2}{f^2} \tag{19}$$

But because of equations (17) and (18) in conjunction with equation (19) above, $\frac{c^2}{d^2}$ was not in its lowest terms, contrary to assumption. Hence there is a contradiction and $\sqrt{M}$ must be irrational. □

It is obvious that adding a rational number to an irrational number will produce an irrational number; if the result were rational then an irrational number could be produced by subtracting a rational from a rational, which is clearly impossible. Furthermore, multiplying an irrational by a rational produces an irrational. Else au irrational could be produced by dividing a rational by a rational, which we know only produces rational numbers. Thus, numbers of the form $F + G\sqrt{M}$ as described previously are also irrational.

Theorem 1.3.2. If $a + b\sqrt{M} = c + d\sqrt{M}$ where $a, b, c, d \in \mathbb{Q}$, $M \in \mathbb{Z}^+$ and is not a perfect square, then $a = c$ and $b = d$.

Proof. If the above conditions are true and $d - b \neq 0$ then

$$\sqrt{M} = \frac{a-c}{d-b}$$

and is rational, contrary to Theorem 1.3.1. So $d - b = 0$ and hence $a = c$ also. $\square$

Lemma 1.3.3. All quadratic irrationals $\alpha = F + G\sqrt{M}$ where $F, G \in \mathbb{Q}$ and $M \in \mathbb{Z}^+$ and is not a perfect square, solve a quadratic equation with integer coefficients and also have a conjugate of the form $\alpha' = F - G\sqrt{M}$ which satisfies the same quadratic equation.

Proof. The quadratic equation

$$\begin{aligned} 0 &= (x - (F + G\sqrt{M}))(x - (F - G\sqrt{M})) \\ &= x^2 - 2Fx + F^2 - G^2M \end{aligned}$$

has roots α and α' , and has rational coefficients. Now let H be the common denominator of $2F$ and $F^2 - G^2M$. That is, let H be the least number that makes both $2FH$ and $F^2H - G^2MH$ integers. Then α and α' are roots of the quadratic equation

$$Hx^2 - 2FHx + F^2H - G^2MH = 0$$

which has integer coefficients.

$\square$

From now on, let the symbol $'$ be the symbol for conjugate. So α' means the conjugate of α and $(\alpha + \beta)'$ means the conjugate of the sum of α and β .

Lemma 1.3.4. When applying one of the following operations;

1. addition
2. subtraction
3. multiplication
4. division

between two quadratic irrationals involving the same integer as the subject of the square root, conjugating the quadratic irrationals before the operation is equivalent to conjugating the result after the operation.

Proof. Let $\alpha = F + G\sqrt{M}$ be a quadratic irrational and $\beta = H + I\sqrt{M}$ be another. So α and β have the same integer under the square root. Now consider each operation:

1)

$$\begin{aligned} \alpha' + \beta' &= (F + G\sqrt{M})' + (H + I\sqrt{M})' \\ &= F - G\sqrt{M} + H - I\sqrt{M} \\ &= (F + H) - (G + I)\sqrt{M} \\ &= (F + G\sqrt{M} + H + I\sqrt{M})' \\ &= (\alpha + \beta)' \end{aligned}$$

2)

$$\begin{aligned} \alpha' - \beta' &= (F + G\sqrt{M})' - (H + I\sqrt{M})' \\ &= F - G\sqrt{M} - H + I\sqrt{M} \\ &= (F - H) - (G - I)\sqrt{M} \\ &= ((F - H) + (G - I)\sqrt{M})' \\ &= (F + G\sqrt{M} - H - I\sqrt{M})' \\ &= (\alpha - \beta)' \end{aligned}$$

3)

$$\begin{aligned}
\alpha'\beta' &= (F - G\sqrt{M})(H - I\sqrt{M}) \\
&= (FH + GIM) - (FI + HG)\sqrt{M} \\
&= (FH + GIM + FI\sqrt{M} + HG\sqrt{M})' \\
&= ((F + G\sqrt{M})(H + I\sqrt{M}))' \\
&= (\alpha\beta)'
\end{aligned}$$

4)

$$\begin{aligned}
\frac{\alpha'}{\beta'} &= \frac{(F - G\sqrt{M})}{(H - I\sqrt{M})} \times \frac{(H + I\sqrt{M})}{(H + I\sqrt{M})} \\
&= \frac{(FH - GIM) - (HG - FI)\sqrt{M}}{H^2 - I^2M} \\
&= \left(\frac{(FH - GIM) + (HG - FI)\sqrt{M}}{H^2 - I^2M}\right)' \\
&= \left(\frac{(F + G\sqrt{M})}{(H + I\sqrt{M})} \times \frac{(H - I\sqrt{M})}{(H - I\sqrt{M})}\right)' \\
&= \left(\frac{\alpha}{\beta}\right)'
\end{aligned}$$

□

Lemma 1.3.5. Every expression in the form $F \pm G\sqrt{M}$ where $F, G \in \mathbb{Q}$; $M \in \mathbb{Z}^+$ and is not a perfect square and $G > 0$, has an equivalent expression in the form

$$\frac{A \pm \sqrt{D}}{B}$$

where $A \in \mathbb{Z}$; $B, D \in \mathbb{Z}^+$ and D is not a perfect square.

Proof. If $F, G \in \mathbb{Q}$ then they can be expressed $F = \frac{f}{h}$; $G = \frac{g}{h}$ where h is the lowest common denominator of F and G , and $f \in \mathbb{Z}$; $g, h \in \mathbb{Z}^+$. Then

$$\begin{aligned}
F \pm G\sqrt{M} &= \frac{f \pm g\sqrt{M}}{h} \\
&= \frac{f \pm \sqrt{g^2M}}{h} \\
&= \frac{A \pm \sqrt{D}}{B}
\end{aligned}$$

where A, B and D satisfy the previously stated conditions and the two forms are equivalent. □

1.3.1 Reduced quadratic irrationals

Definition 1.3.2. A quadratic irrational α is **reduced** if $\alpha > 1$ and $-1 < \alpha' < 0$.

If α is reduced and in regular form $\frac{A + \sqrt{D}}{B}$ then $\alpha' = \frac{A - \sqrt{D}}{B}$ and we can deduce

$$\begin{aligned}
1 &< \frac{A + \sqrt{D}}{B} \\
B &< A + \sqrt{D}
\end{aligned} \tag{20}$$

and

$$
\begin{aligned}
-1 &< \frac{A-\sqrt{D}}{B} \\
-B &< A-\sqrt{D} \\
\sqrt{D}-A &< B
\end{aligned} \tag{21}
$$

Also because $B > 0$ we can deduce

$$
\begin{aligned}
\frac{A-\sqrt{D}}{B} &< 0 \\
A-\sqrt{D} &< 0 \\
A &< \sqrt{D}
\end{aligned} \tag{22}
$$

and

$$
\begin{aligned}
-1+1 &< \alpha' + \alpha \\
0 &< \frac{2A}{B} \\
0 &< A
\end{aligned} \tag{23}
$$

From equations (22) and (23) we find

$$0 < A < \sqrt{D} \tag{24}$$

and from equations (20), (21), (24) we conclude

$$0 < \sqrt{D} - A < B < \sqrt{D} + A < 2\sqrt{D} \tag{25}$$

This set of inequalities will allow us to state the next theorem.

Theorem 1.3.6. There are only a finite number of reduced quadratic irrationals associated with any given D .

Proof. We know from equations (24) and (25) that for α to be a reduced quadratic irrational it is necessary that A must be between 0 and $\sqrt{D}$ and B must be between 0 and $2\sqrt{D}$. Furthermore, both A and B must be integers. Thus, if we fix D then there are a finite number of pairs of integers which meet this criteria. Thus there are only a finite number of potential candidates of A and B that make α reduced when D is fixed. □

Theorem 1.3.7. If α_n is reduced and $\alpha_n = a_n + \frac{1}{\alpha_{n+1}}$ where a_n is the floor of α_n then α_{n+1} is reduced and has the same integer as the subject of the square root.

Proof. First we show α_{n+1} is reduced.

$$
\begin{aligned}
0 &< \alpha_n - a_n < 1 \\
0 &< \frac{1}{\alpha_{n+1}} < 1
\end{aligned}
$$

so

$$\alpha_{n+1} > 1 \tag{26}$$

Furthermore, from Theorem 1.3.4

$$
\begin{aligned}
(\alpha_n - a_n)' &= \left(\frac{1}{\alpha_{n+1}}\right)' \\
\alpha_n' - a_n &= \frac{1}{\alpha_{n+1}'}
\end{aligned}
$$

then because $1 < a_n$ and $-1 < \alpha_n' < 0$

$$1 < a_n - \alpha_n' = \frac{-1}{\alpha_{n+1}'}$$

so

$$-1 < \alpha'_{n+1} < 0 \tag{27}$$

Equations (26) and (27) fulfill the requirements for α_{n+1} being reduced.

Now we show the subject of the square root is the same for α_{n+1} as it is for α_n . Let

$$\alpha = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{A_n+\sqrt{D}}{B_n}$$

So

$$\begin{aligned}
0 &= a\alpha_n^2 + b\alpha_n + c \\
&= a\left(a_n + \frac{1}{\alpha_{n+1}}\right)^2 + b\left(a_n + \frac{1}{\alpha_{n+1}}\right) + c \\
&= aa_n^2 + \frac{2aa_n}{\alpha_{n+1}} + \frac{a}{\alpha_{n+1}^2} + ba_n + \frac{b}{\alpha_{n+1}} + c \\
&= (aa_n^2 + ba_n + c)\alpha_{n+1}^2 + (2aa_n + b)\alpha_{n+1} + a
\end{aligned} \tag{28}$$

Note that the coefficients of this last quadratic equation are integers. Thus when we solve it for the positive root $\alpha_{n+1} = \dfrac{A_{n+1}+\sqrt{D_{n+1}}}{B_{n+1}}$ we will get A_{n+1}, B_{n+1}, D_{n+1} all integers. Now solve for D_{n+1}:

$$\begin{aligned}
D_{n+1} &= (2aa_n + b)^2 - 4(aa_n^2 + ba_n + c)a \\
&= 4a^2a_n^2 + 4aa_nb + b^2 - 4a^2a_n^2 - 4aba_n - 4ac \\
&= b^2 - 4ac \\
&= D
\end{aligned}$$

Thus the theorem is proved. $\square$

Finally we come to the focal point of this section.

Theorem 1.3.8. If α is a reduced quadratic irrational then its simple continued fraction expansion is purely periodic.

Proof. Theorem 1.2.3 proves that the simple continued fraction of α is infinite. However, Theorem 1.3.7 implies that every residue of α is reduced. As a consequence of these two seemingly opposite statements, at some point there occurs some residue, r_k, that is a repetition of a previous residue, r_j .

Consider the simple continued fraction expansion of α , because a_j and a_k are the largest integers less than r_j and r_k respectively and $r_j = r_k$, it follows that $a_j = a_k$ and

$$\begin{aligned}
r_j &= r_k \\
a_j + \frac{1}{r_{j+1}} &= a_k + \frac{1}{r_{k+1}} \\
r_{j+1} &= r_{k+1}
\end{aligned}$$

Furthermore, the same reasoning can be applied to show that $r_{j+2} = r_{k+2}$, $r_{j+3} = r_{k+1}$ and so on.

Now because $r_{n-1} = a_{n-1} + \frac{1}{r_n}$ we can manipulate two expressions about r_{j-1} and r_{k-1} side by side to eventually show that they are equal.

$$\begin{aligned}
r_{j-1} &= a_{j-1} + \frac{1}{r_j} & \qquad r_{k-1} &= a_{k-1} + \frac{1}{r_k} \\
r'_{j-1} &= a_{j-1} + \frac{1}{r'_j} & \qquad r'_{k-1} &= a_{k-1} + \frac{1}{r'_k}
\end{aligned}$$

from Theorem 1.3.4. Because $r_j = r_k$ it follows that $r'_j = r'_k$, and

$$r'_{j-1} - a_{j-1} = r'_{k-1} - a_{k-1}$$

so

$$a_{j-1} - r'_{j-1} = a_{k-1} - r'_{k-1}$$

Because r_{j-1} and r_{k-1} are reduced, we have $-1 < r_{j-1}, r_{k-1} < 0$ so $0 < |r_{j-1}|, |r_{k-1}| < 1$, and it follows from Lemma 1.2.1 that $a_{j-1} = a_{k-1}$ and $r'_{j-1} = r'_{k-1}$ and thus, $r_{j-1} = r_{k-1}$. As before, the same method shows that $r_{j-2} = r_{k-2}$, $r_{j-3} = r_{k-3}$ and so on, up to $r_{j-(j-1)} = r_{k-(j-1)}$ which means $r_1 = r_{k-j+1}$, and $\alpha = r_{k-j}$ (α can be seen as r_0).

Let m be the value where r_m is the first residue where the value equals α . Then $r_i = r_{m+i}$ for all $i \in \mathbb{Z}^+$. Furthermore, taking the unique integer a_n for α and each of r_n we get $a_0 = a_m$ and $a_i = a_{m+i}$ for all $i \in \mathbb{Z}^+$.

Thus $\alpha = [\overline{a_0; a_1, \ldots, a_{m-1}}]$. [2]

$\square$

The variable m is known as the length of the period and will be quite important in solving Fell's equation.

1.4 Continued fractions of square roots

Square roots of natural numbers are not reduced quadratic irrationals because their conjugates are never between -1 and 0. Thus they are never purely periodic. However, they do have a special form.

Theorem 1.4.1. Simple continued fractions of square roots take the form

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

when $D \in \mathbb{Z}^+$ and is not a perfect square.

Proof. If $D \in \mathbb{Z}^+$ and is not a perfect square then $1 < \sqrt{D}$. This means its conjugate $-\sqrt{D} < -1$ and so $\sqrt{D}$ is not reduced. However, if a_0 is the greatest integer less than $\sqrt{D}$ then

$$1 < a_0 + \sqrt{D}$$

and its conjugate

$$-1 < a_0 - \sqrt{D} < 0$$

So $a_0 + \sqrt{D}$ is reduced and by Theorem 1.3.8 its simple continued fraction representation is purely periodic.

So

$$a_0 + \sqrt{D} = 2a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{m-1} + \cfrac{1}{2a_0 + \cfrac{1}{\ddots}}}}}$$

and

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

$\square$

Before we go on to solving Pell's equation we will find the simple continued fraction representation of $\sqrt{7}$ and show that it takes the form above.

First note that $a_0 = \lfloor \sqrt{7} \rfloor = 2$. So $\sqrt{7} = 2 + \frac{1}{r_1}$ and

$$\begin{aligned} r_1 &= \frac{1}{\sqrt{7}-2} \cdot \frac{\sqrt{7}+2}{\sqrt{7}+2} \\ &= \frac{\sqrt{7}+2}{3} \end{aligned}$$

[2] The overhead bar is the standard mathematical notation used when the content below it repeats forever.

Now $a_1 = \lfloor r_1 \rfloor = \lfloor \dfrac{\sqrt{7}+2}{3} \rfloor = 1$. So $r_1 = \dfrac{\sqrt{7}+2}{3} = 1 + \dfrac{1}{r_2}$ and

$$\begin{aligned} r_2 &= \frac{1}{\left(\dfrac{\sqrt{7}+2}{3} - 1\right)} \\ &= \frac{3}{\sqrt{7}-1} \cdot \frac{\sqrt{7}+1}{\sqrt{7}+1} \\ &= \frac{\sqrt{7}+1}{2} \end{aligned}$$

Now $a_2 = \lfloor r_2 \rfloor = \lfloor \dfrac{\sqrt{7}+1}{2} \rfloor = 1$. So $r_2 = \dfrac{\sqrt{7}+1}{2} = 1 + \dfrac{1}{r_3}$ and

$$\begin{aligned} r_3 &= \frac{1}{\left(\dfrac{\sqrt{7}+1}{2} - 1\right)} \\ &= \frac{2}{\sqrt{7}-1} \cdot \frac{\sqrt{7}+1}{\sqrt{7}+1} \\ &= \frac{\sqrt{7}+1}{3} \end{aligned}$$

Now $a_3 = \lfloor r_3 \rfloor = \lfloor \dfrac{\sqrt{7}+1}{3} \rfloor = 1$. So $r_3 = \dfrac{\sqrt{7}+1}{3} = 1 + \dfrac{1}{r_4}$ and

$$\begin{aligned} r_4 &= \frac{1}{\left(\dfrac{\sqrt{7}+1}{3} - 1\right)} \\ &= \frac{3}{\sqrt{7}-2} \cdot \frac{\sqrt{7}+2}{\sqrt{7}+2} \\ &= \sqrt{7}+2 \end{aligned}$$

Now $a_4 = \lfloor r_4 \rfloor = \lfloor \sqrt{7}+2 \rfloor = 4$. So $r_4 = \sqrt{7}+2 = 4 + \dfrac{1}{r_5}$ and

$$\begin{aligned} r_5 &= \frac{1}{\sqrt{7}-2} \\ &= r_1 \end{aligned}$$

As a consequence, $a_5 = a_1$, $a_6 = a_2$, ... Thus the simple continued fraction representation is

$$\sqrt{7} = [2; \overline{1,1,1,4}] \tag{29}$$

which is consistent with Theorem 1.4.1.

1.5 Pell's equation

(Related: Q66, Q94)

Definition 1.5.1. Pell's equation is an equation of the form

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

where D is a given non-square positive integer, and integer solutions are sought for x and y .

The trivial solution is $x = 1$, $y = 0$ or $x = -1, y = 0$, but that is not very interesting. So we want x, y to be positive integers.

Note that D must be non-square for x, y to have non-trivial solutions, because if D is perfect square, then we can let $D = c^2$ for some positive integer c and we have

$$x^2 - c^2y^2 = 1$$
$$(x + cy)(x - cy) = 1$$

Because the only possible factors of 1 are -1 and 1, this leaves us with two options.

The first is that

$$x + cy = x - cy = 1$$

then $x = 1$ and $cy = 0$, so $y = 0$. This is the trivial solution.

The second is that

$$x + cy = x - cy = -1$$

then $x = -1$, and $cy = 0$ so $y = 0$ which is another trivial solution.

1.5.1 Solving Pell's equation with continued fractions

Theorem 1.5.1. Let $D \in \mathbb{Z}^+$ and not be a perfect square, so $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$. Also let p_n and q_n be defined as in Theorem 1.2.4 .
If the length of the period, m, is even then $(x, y) = (p_{m-1}, q_{m-1})$ solves the Pell equation $x^2 - Dy^2 = 1$ for integers. If the length of the period, m, is odd then $(x, y) = (p_{m-1}, q_{m-1})$ solves the negative Pell equation $x^2 - Dy^2 = -1$ for integers and $(x, y) = (p_{2m-1}, q_{2m-1})$ solves the Pell equation $x^2 - Dy^2 = 1$ for integers.

Proof. Because $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$ it follows that

$$\sqrt{D} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{m-1} + \cfrac{1}{a_0 + \sqrt{D}}}}} \tag{30}$$

Now from Theorem 1.2.5 we get

$$\begin{aligned} \sqrt{D} &= \frac{(a_0 + \sqrt{D})p_{m-1} + p_{m-2}}{(a_0 + \sqrt{D})q_{m-1} + q_{m-2}} \\ (a_0 + \sqrt{D})q_{m-1}\sqrt{D} + q_{m-2}\sqrt{D} &= (a_0 + \sqrt{D})p_{m-1} + p_{m-2} \\ q_{m-1}D + (a_0q_{m-1} + q_{m-2})\sqrt{D} &= a_0p_{m-1} + p_{m-2} + p_{m-1}\sqrt{D} \end{aligned}$$

and from theorem 1.3.2 it follows that

$$q_{m-1}D = a_0p_{m-1} + p_{m-2} \qquad \text{and} \qquad p_{m-1} = a_0q_{m-1} + q_{m-2}$$

so

$$p_{m-2} = q_{m-1}D - a_0p_{m-1} \qquad \text{and} \qquad q_{m-2} = p_{m-1} - a_0q_{m-1} \tag{31}$$

We can adjust the formula given in Theorem 1.2.6 by letting $n = m - 2$ to show

$$\begin{aligned}(-1)^{m-2} &= p_{m-1}q_{m-2} - p_{m-2}q_{m-1} \\ &= p_{m-1}(p_{m-1} - a_0 q_{m-1}) - (q_{m-1}D - a_0 p_{m-1})q_{m-1}\end{aligned}$$

from equations (31). So

$$p_{m-1}^2 - Dq_{m-1}^2 = (-1)^m \tag{32}$$

So when m is even $(x, y) = (p_{m-1}, q_{m-1})$ solve the Pell equation, and when m is odd it solves the negative Pell equation.

Note that when setting up this proof in equation (30) we did not need to stop at the end of the first period. Instead, we could have stopped at the end of any period. If we stopped at the end of the second period, equation (30) would look like

$$\sqrt{D} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{2m-1} + \cfrac{1}{a_0 + \sqrt{D}}}}}$$

From this equation the previous logic can be carried out the same with the only difference being that $m - 1$ is replaced by $2m - 1$ and $m - 2$ is replaced by $2m - 2$. The new version of equation (32) will then look like

$$p_{2m-1}^2 - Dq_{2m-1}^2 = (-1)^m \tag{33}$$

Thus when m is odd and it is not sufficient to stay in the first period to solve the Pell equation, $(x, y) = (p_{2m-1}, q_{2m-1})$ will give a solution in integers.

□

The general form of equation (33) for the k th period is

$$p_{km-1}^2 - Dq_{km-1}^2 = (-1)^{km}$$

for $k \in \mathbb{Z}^+$. This shows when m is even, $(x, y) = (p_{km-1}, q_{km-1})$ will solve Pell's equation for all k , and when m is odd, $(x, y) = (p_{km-1}, q_{km-1})$ will solve Pell's equation for all even k , and will solve the negative Pell equation for all odd k .

A direct consequence of this is that for any given non-square positive integer D , there are an infinite number of solutions for (x, y) .

Example

Let's solve the Pell equation $x^2 - 7y^2 = 1$.

We know from equation (29) that $\sqrt{7} = [2; \overline{1, 1, 1, 4}]$. We can see that the length of the period is 4 so Theorem 1.4.2 tells us that the numerator and the denominator of the 3rd convergent will solve the Pell equation for $D = 7$. The numerator and denominators of the first 7 convergents are calculated from Theorem 1.2.4:

$$\begin{aligned}p_n &= a_n p_{n-1} + p_{n-2} \\ q_n &= a_n q_{n-1} + q_{n-2}\end{aligned}$$

which are as follows:

n	-2	-1	0	1	2	3	4	5	6	7
a_n			2	1	1	1	4	1	1	1
p_n	0	1	2	3	5	8	37	45	82	127
q_n	1	0	1	1	2	3	14	17	31	48

We can see that $p_3 = 8$ and $q_3 = 3$. This indeed is a solution to Pell's equation with $D = 7$ as $8^2 - 7 \times 3^2 = 1$. Furthermore, calculating into the second period shows that $p_7 = 127$ and $q_7 = 48$. A quick check with a calculator shows that it is true that $127^2 - 7 \times 48^2 = 1$. If more solutions are required this is easily extended into the 3rd period or further.

1.5.2 Generating more solutions given a solution

But there is a faster method for finding other solutions once we find the first solution.

Note that we can 'factorize' the Pell's equation:

$$\begin{aligned} x^2 - Dy^2 &= 1 \\ (x + y\sqrt{D})(x - y\sqrt{D}) &= 1 \end{aligned}$$

Then the LHS is the product of a pair of quadratic irrationals which are conjugate of each other.

(Now we copy from another source [4]. So there is a slight change of notation.)

Definition 1.5.2. Let $\mathbb{Z}(\sqrt{D})$ denote the set of all numbers in the form $x + y\sqrt{D}$, $x, y \in \mathbb{Z}$. Symbolically,

$$\mathbb{Z}(\sqrt{D}) = \{x + y\sqrt{D} \mid x, y \in \mathbb{Z}\}$$

Definition 1.5.3. The **norm** of a number $u = x + y\sqrt{D}$ is defined as

$$N(u) = uu' = (x + y\sqrt{D})(x - y\sqrt{D}) = x^2 - Dy^2$$

Theorem 1.5.2. Let $u, v \in \mathbb{Z}(\sqrt{D})$, then $u + v$, $uv \in \mathbb{Z}(\sqrt{D})$. If $u \in \mathbb{Z}(\sqrt{D})$ and $N(u) = 1$, then $u^{-1} \in \mathbb{Z}(\sqrt{D})$.

Proof. Let $u = p + q\sqrt{D}$ and $v = r + s\sqrt{D}$, where $p, q, r, s \in \mathbb{Z}$. Then

$$u + v = (p + q\sqrt{D}) + (r + s\sqrt{D}) = (p + r) + (q + s)\sqrt{D} \quad \in \mathbb{Z}(\sqrt{D})$$

$$uv = (p + q\sqrt{D})(r + s\sqrt{D}) = (pr + qsD) + (ps + qr)\sqrt{D} \quad \in \mathbb{Z}(\sqrt{D})$$

For the if-then statement, first note that

$$(p + q\sqrt{D})^{-1} = \frac{1}{(p + q\sqrt{D})(p - q\sqrt{D})}(p - q\sqrt{D}) \tag{2}$$

If $u \in \mathbb{Z}(\sqrt{D})$ and $N(u) = 1$, then

$$u^{-1} = \frac{1}{N(u)}(p - q\sqrt{D}) = \frac{1}{1}(p - q\sqrt{D}) = u' \quad \in \mathbb{Z}(\sqrt{D})$$

□

Theorem 1.5.3. The norm and the conjugate are multiplicative in u:

$$N(uv) = N(u)N(v) \qquad \text{and} \qquad (uv)' = u'v'$$

Proof.

$$\begin{aligned} u'v' = (p - q\sqrt{D})(r - s\sqrt{D}) &= (pr + qsD) - (ps + qr)\sqrt{D} \\ &= (uv)' \end{aligned}$$

Then

$$N(uv) = (uv)(uv)' = uvu'v' = (uu')(vv') = N(u)N(v)$$

□

Note that for $u = x + y\sqrt{D}$, a pair of integers (x, y) is a solution to Pell's equation if and only if $N(u) = 1$.

Lemma 1.5.4. Let (x, y) be an integer solution to Pell's equation and $u = x + y\sqrt{D}$.

1. If $x > 0$ and $y > 0$, then $u > 1$;
2. If $x > 0$ and $y < 0$, then $0 < u < 1$;
3. If $x < 0$ and $y > 0$, then $-1 < u < 0$;

4. If $x < 0$ and $y < 0$, then $u < -1$;

Proof. Suppose $x > 0$ and $y > 0$. Since $(x - y\sqrt{D})(x + y\sqrt{D}) = 1$, we have $x - y\sqrt{D} > 0$ and $x + y\sqrt{D} > x - y\sqrt{D}$. Hence $u > 1$ and $u' < 1$. This proves the first two statements. The third and the fourth statements follow from the first two.

□

Definition 1.5.4. Let (a, b) be a non-trivial solution to Pell's equation with positive integer components $a > 0,\ b > 0$. We say that this solution is **fundamental** if the number $u = a + b\sqrt{D}$ takes the minimal possible value.

Note that the number u is uniquely determined by Theorem 1.3.2 . Let's also note that $u > 1$ by Lemma 1.5.4.

Theorem 1.5.5. Let (x_1, y_1) be a fundamental solution to Pell's equation and $u = x_1 + y_1\sqrt{D}$. Let

$$u^n = x_n + y_n\sqrt{D}\,, \qquad n = 0, 1, 2, \ldots$$

Then $(\pm x_n, \pm y_n)$, $n = 0, 1, 2, \ldots$, is the complete set of solutions to Pell's equation.

In other words, every positive solution is a pair of coefficients of the expanded form of u^n for $n = 0, 1, 2, \ldots$.

Proof. The trivial solution $(1, 0)$ is in this set and we get it for $n = 0$. Let (x, y) be an arbitrary non-trivial solution to Pell's equation, and $v = x + y\sqrt{D}$. We may assume $x > 0$ and $y > 0$. All we need to show is that $v = x + y\sqrt{D}$ can be represented as u^n for some positive integer n .

Let us assume the contrary. As $x > 0$ and $y > 0$, we know that $v > 1$. Since $u > 1$ the terms of the sequence $1, u, u^2, \ldots, u^n, \ldots$ get arbitrary large, thus there exists n such that $u^n < v < u^{n+1}$. Let us multiply this inequality by $(u^n)^{-1}$. We get

$$1 < v(u^n)^{-1} < u \tag{3}$$

Let's make a numbe of observations.

Firstly, note that $(u^n)^{-1} = (u^{-1})^n = (x_1 - y_1\sqrt{D})^n$. By Theorem 1.5.2, this means $(u^n)^{-1} \in \mathbb{Z}(D)$ and hence $v(u^n)^{-1} \in \mathbb{Z}(D)$. We can let $v(u^n)^{-1} = \bar{x} + \bar{y}\sqrt{D}$ for some $\bar{x}, \bar{y} \in \mathbb{Z}$.

Secondly, by Theorem 1.5.3, $N(v(u^n)^{-1}) = N(v)N((u^{-1})^n) = N(v)(N(u^{-1}))^n = N(v)(N(u'))^n = (1)(1)^n = 1$ by so $(\bar{x}, \bar{y})$ is a solution to Pell's equation.

Thirdly, by equation (3), $1 < v(u^n)^{-1}$, so by Lemma 1.5.4 we get $\bar{x} > 0, \bar{y} > 0$ (the if-then condition also works in reverse since any other possibilities of $\bar{x}, \bar{y}$ will make $v(u^n)^{-1} < 1$) .

Finally, this contradicts to equation (3), namely to $v(u^n)^{-1} < u$, since u was fundamental. To elaborate, we have a value $v(u^n)^{-1} = \bar{x} + \bar{y}\sqrt{D}$ that is smaller than value of the fundamental solution $u = x_1 + y_1\sqrt{D}$. This is impossible since u is supposed to be the minimum value.

Thus the theorem is proved.

□

Example

Let's use this theorem in an example. Let's say for the Pell's equation

$$x^2 - 2y^2 = 1$$

we find the fundamental solution $(3, 2)$. (check: $3^2 - 2(2)^2 = 1$)

Then every positive solution is the coefficient of the expanded form of $u^n = (3 + 2\sqrt{2})^n$. To find the next solution, let $n = 2$ and expand:

$$\begin{aligned}(3 + 2\sqrt{2})^2 &= 9 + (2)(3)(2\sqrt{2}) + 8 \\ &= 17 + 12\sqrt{2}\end{aligned}$$

Then $(17, 12)$ is the next solution to the equation (check: $17^2 - 2(12^2) = 289 - 288 = 1$).

To find the third solution, let $n = 3$ and expand:

$$\begin{aligned}(3+2\sqrt{2})^3 &= (17+12\sqrt{2})(3+2\sqrt{2}) \\ &= 99+70\sqrt{2}\end{aligned}$$

Then $(99, 70)$ is the third solution to the equation (check: $99^2 - 2(70^2) = 9801 - 9800 = 1$). We can find every solution by repeating this process.

And to find the fundamental solution, we can find use the continued fraction method in section 1.5.1. The first solution is guaranteed to be the fundamental solution.

Theorem 1.5.6. Let $D \in \mathbb{Z}^+$ and not be a perfect square, so $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_{m-1}, 2a_0}]$. Also let p_n and q_n be defined as in Theorem 1.2.4 .

If the length of the period, m, is even then $(x, y) = (p_{m-1}, q_{m-1})$ is the fundamental solution to Pell's equation. If the length of the period, m, is odd then $(x, y) = (p_{2m-1}, q_{2m-1})$ is the fundamental solution.

Proof. From Theorem 1.2.4,

$$\begin{aligned}p_n &= a_n p_{n-1} + p_{n-2} \\ q_n &= a_n q_{n-1} + q_{n-2}\end{aligned}$$

We want to show that p_n is a strictly increasing sequence for $n \geq 0$.

Note that $p_{-1} = 1$ and $p_0 = a_0 > 0$, and $p_1 = a_1 p_0 + 1 > p_0$.

Assume that $p_0, \ldots, p_k$ is a strictly increasing sequence, where $k \geq 1$. Then $p_k > p_0 > 0$ and $p_{k-1} \geq p_0 > 0$, and

$$\begin{aligned}p_{k+1} - p_k &= a_{k+1} p_k + p_{k-1} - p_k \\ &= p_k(a_{k+1} - 1) + p_{k-1}\end{aligned}$$

Since $p_k(a_{k+1} - 1) \geq 0$ and $p_{k-1} > 0$, we have $p_{k+1} - p_k > 0$, which means $p_0, \ldots, p_{k+1}$ is also strictly increasing. By induction, p_n is a strictly increasing sequence for $n \geq 0$.

Since $q_0 = 1$, $q_1 = a_1 > 0$ and $q_2 = a_2 q_1 + 1 > q_1$, we can use similar argument to show that q_n is a strictly increasing sequence for $n \geq 1$.

Now, since $n = m - 1$ (if m is even) or $n = 2m - 1$ (if m is odd) is the minimum n such that (p_n, q_n) solves Pell's equation, and p_n is strictly increasing, we have

$$p_{m-1} < p_k \quad (\text{or } p_{2m-1} < p_k) \qquad \text{and} \qquad q_{m-1} < q_k \quad (\text{or } q_{2m-1} < q_k)$$

for any $k > m - 1$ (or $k > 2m - 1$) such that (p_k, q_k) solve Pell's equation.

This means

$$\begin{aligned}p_{m-1} + q_{m-1}\sqrt{D} &< p_k + q_k\sqrt{D} \\ (\text{or}) \qquad p_{2m-1} + q_{2m-1}\sqrt{D} &< p_k + q_k\sqrt{D}\end{aligned}$$

so (p_{m-1}, q_{m-1}) (or (p_{2m-1}, q_{2m-1})) is the fundamental solution.

□

Now we give the 'general' formula for x and y respectively.

Theorem 1.5.7. Let (x_1, y_1) be a fundamental solution to Pell's equation and $u = x_1 + y_1\sqrt{D}$, $u' = x_1 - y_1\sqrt{D}$. Let

$$u^n = x_n + y_n\sqrt{D}, \qquad n = 0, 1, 2, \ldots$$

Then

$$x_n = \frac{u^n + (u')^n}{2}, \quad y_n = \frac{u^n - (u')^n}{2\sqrt{D}}$$

Proof. For $n = 0$ we have $(x_0, y_0) = (1, 0)$, $u^0 = 1 + 0\sqrt{D} = 1$. So

$$\frac{1^0 + 1^0}{2} = 1 = x_0 \ , \quad \frac{1^n - 1^n}{2\sqrt{D}} = 0 = y_0$$

For $n = 1$,

$$\frac{(x_1 + y_1\sqrt{D}) + (x_1 - y_1\sqrt{D})}{2} = x_1 \ , \quad \frac{(x_1 + y_1\sqrt{D}) - (x_1 - y_1\sqrt{D})}{2\sqrt{D}} = y_1$$

Assume that

$$x_k = \frac{u^k + (u')^k}{2} \ , \quad y_k = \frac{u^k - (u')^k}{2\sqrt{D}}$$

When $n = k + 1$,

$$\begin{aligned} u^{k+1} &= (x_k + y_k\sqrt{D})(x_1 + y_1\sqrt{D}) \\ &= x_k x_1 + y_k y_1 D + (x_k y_1 + y_k x_1)\sqrt{D} \end{aligned}$$

So for LHS, $x_{n+1} = x_k x_1 + y_k y_1 D$ and $y_{n+1} = x_k y_1 + y_k x_1$. By induction hypothesis,

$$\begin{aligned} x_{n+1} &= (\frac{u^k + (u')^k}{2})x_1 + (\frac{u^k - (u')^k}{2\sqrt{D}})y_1 D \\ &= \frac{(u^k + (u')^k)x_1 + (u^k - (u')^k)y_1\sqrt{D}}{2} \\ &= \frac{u^k(x_1 + y_1\sqrt{D}) + (u')^k(x_1 - y_1\sqrt{D})}{2} \\ &= \frac{u^{k+1} + (u')^{k+1}}{2} \\ &= \text{RHS} \end{aligned}$$

And

$$\begin{aligned} y_{n+1} &= (\frac{u^k + (u')^k}{2})y_1 + (\frac{u^k - (u')^k}{2\sqrt{D}})x_1 \\ &= \frac{(u^k + (u')^k)y_1\sqrt{D} + (u^k - (u')^k)x_1}{2\sqrt{D}} \\ &= \frac{u^k(x_1 + y_1\sqrt{D}) - (u')^k(x_1 - y_1\sqrt{D})}{2\sqrt{D}} \\ &= \frac{u^{k+1} - (u')^{k+1}}{2\sqrt{D}} \\ &= \text{RHS} \end{aligned}$$

Thus the formula holds for all non-negative integers n .

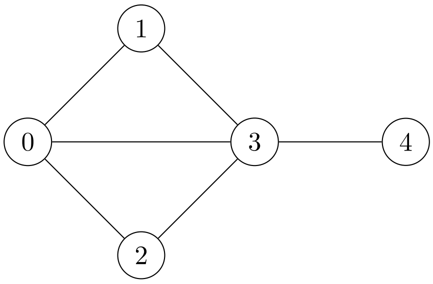
□

2 Graph theory

(Related: Q107, Q83)

Definition 2.0.1. A graph $G = (V, E)$ is a set of vertices and edges E where each edge (u, v) is a connection between vertices. $u, v \in V$.

For example, in the following graph,



$$V = \{0, 1, 2, 3, 4\}$$
$$E = \{(0, 1), (0, 2), (0, 3), (1, 3), (2, 3), (3, 4)\}$$

2.1 Prim's algorithm

(Q107)[5]

Definition 2.1.1. $G(V, E, w)$ is an **edge weighted graph** if there exists a weight function $w : E \to \mathbb{R}$ that assigns a weight to every edge $e \in E$.

Definition 2.1.2. $G(V, E)$ is **connected** if there exists a path between any two vertices. $G(V, E)$ is a **tree** if it is connected and has no cycles.

Definition 2.1.3. Let $H(V', E')$ be a subgraph of $G(V, E)$. We say that H is a **spanning tree** of G if H is a tree and $V' = V$.
In other words, a spanning tree of G is a tree that contains every vertex of G .

Definition 2.1.4. The **weight** of a graph is the sum of the weights of all edges.

$$\text{Weight} = \sum_{e \in E} w(e)$$

Definition 2.1.5. Let $G(V, E, w)$ be an edge-weighted graph where $w : E \to \mathbb{N}$. $H(V, E')$ is a **minimum spanning tree** (MST) of G if it is a spanning tree with weights less than or equal to the weight of any other spanning tree of G , i.e., $\sum_{e \in E'} w(e) \leq \sum_{e \in E''} w(e)$ for all other spanning trees $H'(V, E'')$ of G .

2.1.1 The procedure

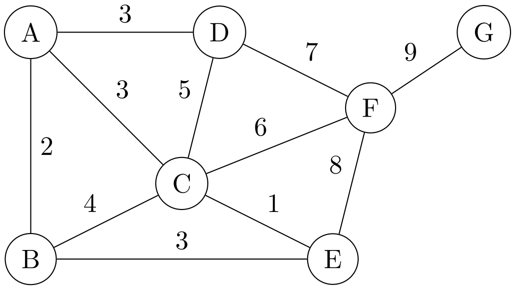
The objective of Prim's algorithm is to find the minimum spanning tree of a weighted undirected graph.

The algorithm may informally be described as performing the following steps:

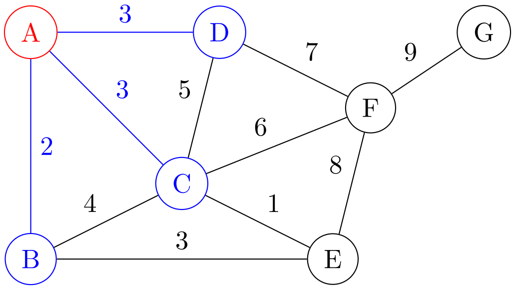
1. Initialize a tree with a single vertex, chosen arbitrarily from the graph.
2. Grow the tree by one edge: Of the edges that connect the tree to vertices not yet in the tree, find the minimum-weight edge, and transfer it to the tree.

3. Repeat step 2 (until all vertices are in the tree).

Example [6]

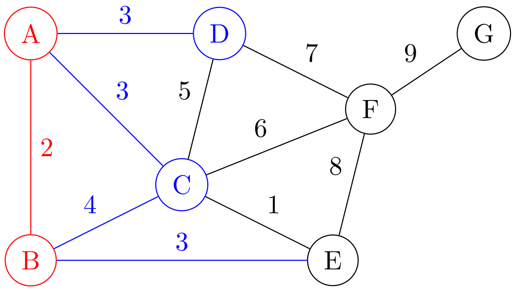


Let's start with vertex A. First add A to the visited list and examine all vertices reachable from A.



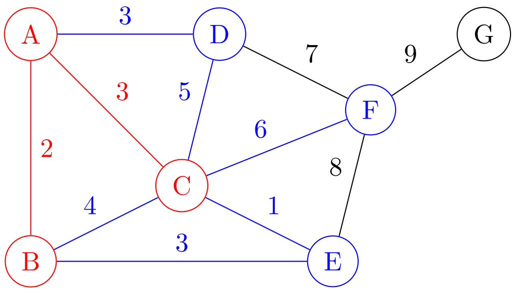
$$\text{Visited} = \{A\}$$

Choose the smallest edge that connects to an unvisited vertex, which is B . Add B to the visited list. Now look at all the vertices reachable from A and B .



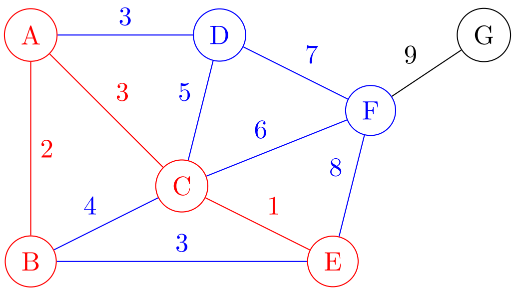
$$\text{Visited} = \{A, B\}$$

The three edges all have a weight of 3. Picking any one of these will work. Let's pick AC .

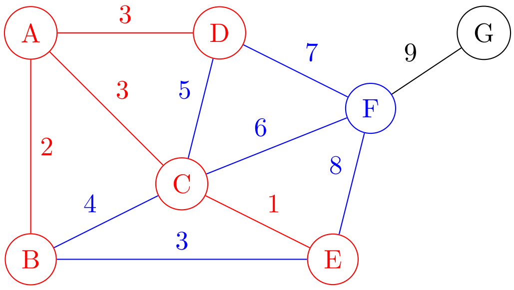


$$\text{Visited} = \{A, B, C\}$$

Continue in this manner, each time picking the smallest edge that connects to an unvisited vertex.

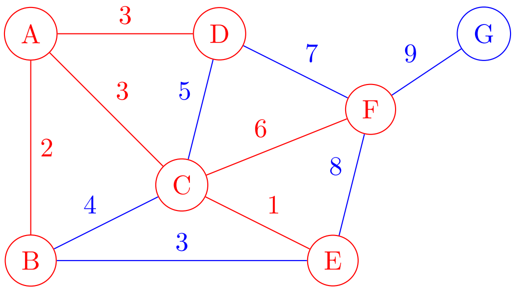


Visited = $\{A, B, C, E\}$



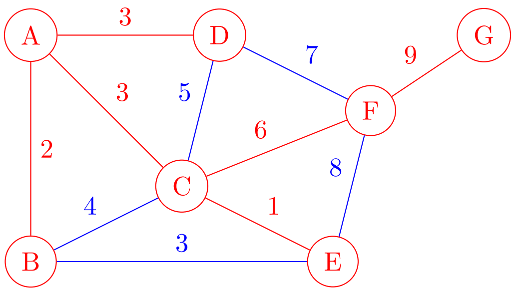
Visited = $\{A, B, C, D, E\}$

Notice at this point the edge BE with weight of 3 is the smallest edge but both vertices are already in MST so we do not pick it. Instead we will choose to add F to the MST.



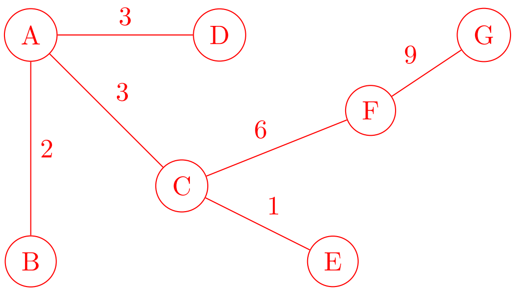
Visited = $\{A, B, C, D, E, F\}$

The only unvisited vertex remaining is G so we add this to the minimum spanning tree.



Visited = $\{A, B, C, D, E, F\}$

All the vertices are now connected in a tree. If we drop the blue edges, we get the minimum spanning tree with a total weight of $2 + 3 + 3 + 1 + 6 + 9 = 24$.



2.1.2 Proof of optimality

[5] Let's recall two facts about spanning trees.

1. Let T be a spanning tree of G . If you add any edge $e \notin T$ to T then $T' = T \cup \{e\}$ contains a cycle. This is easy to see. Let $e = (u, v)$ be an edge not in T . Since T is a spanning tree, there is already a path between any two vertices of G; in particular there is a path between u and v . Adding e to T will thus close the cycle from u to v .

2. Consider $T' = T \cup \{e\}$ as defined above, and let $\mathcal{C}$ be the cycle created by adding the edge e . If you remove any edge $e' \in \mathcal{C}$ from the cycle, you will get a new spanning tree of G; since removing an edge from a cycle will not disconnect the graph.

Now let's get into the proof.

Proof. Let T be the spanning tree return by the algorithm, and suppose there doesn't exist any MST of G consistent with T (which means T is not minimum spanning tree). Consider an optimal MST O of G:

Algorithm 1 Prim's Algorithm

Input: A weighted graph $G(V, E, w)$

Output: A spanning tree that minimizes $\sum_{e \in E'} w(e)$

1: $U \leftarrow V$ $\triangleright$ U is the set of unvisited vertices
2: Pick an arbitrary start vertex $s \in V$
3: $T = \{s\}$
4: $U \leftarrow U \backslash \{s\}$
5: **while** $U \neq \emptyset$ **do**
6: Choose $u \in U$ adjacent to a $v \in T$ such that $w(u, v)$ is the smallest out of such vertices
7: $T \leftarrow T \cup \{u\}$
8: $U \leftarrow U \backslash \{u\}$
9: **end while**
return T

Let $e = (x, y)$ be the first edge chosen by the algorithm that is inconsistent with any MST of G (which means e is not contained in any MST), and let T' be the subtree of T created by the algorithm just before the edge e was chosen. Let P_{xy} be the path between x and y in O. Let $\mathcal{C}$ be the cycle created in O from adding e to P_{xy} (Fact 1).

P_{xy} must contain an edge $e' = (a, b)$ that has an end point in T' and end point outside of T' . Why? Since otherwise, P_{xy} would be fully contained in T' , and choosing $e(x, y)$ next would create a cycle in $T' \cup e$, a contradiction to step 6 of the Algorithm.

Therefore both $e(x, y)$ and $e'(a, b)$ have an end point in T' and an end point outside of T' . Since the algorithm chose e instead of e' , it follows that $w(e) \leq w(e')$. By Fact (2), we can break $\mathcal{C}$ by removing $e'(a, b)$ and obtaining a new MST of G, call it O' . Notice that the total weight of O' is $w(O') = w(O) + w(e) - w(e')$. Since $w(e) \leq w'(e)$, it follows that $w(O') \leq w(O)$. Thus there exists an optimal MST of G , O' , that contains the edge $e(x, y)$ (which is a contradiction).

Therefore, in order to show that Prim's Algorithm does indeed produce an optimal MST for G , it suffices to repeat this argument for every new edge $\hat{e}$ chosen by the algorithm, such that $\hat{e}$ doesn't appear in any optimal solution.

$\square$

References

[1] Wikipedia, "Pythagorean triple." [Online]. Available: https://en.wikipedia.org/wiki/Pythagorean_triple#Generating_a_triple

[2] Jesse Unger, "Solving pell's equation with continued fractions." [Online]. Available: https://ir.canterbury.ac.nz/bitstream/handle/10092/10158/unger_2009_report.pdf?sequence=1&isAllowed=y

[3] William, "The sequence of partial convergents." [Online]. Available: https://wstein.org/edu/2007/spring/ent/ent-html/node60.html#prop:dets

[4] The University of University Auckland, "Pell's equation (handout for maths 714)." [Online]. Available: https://www.math.auckland.ac.nz/class714/Pell.pdf

[5] Lalla Mouatadid, "Greedy algorithms: Minimum spanning tree (csc 373 - algorithm design, analysis, and complexity)." [Online]. Available: http://www.cs.toronto.edu/~lalla/373s16/notes/MST.pdf

[6] Michael Sambol, "Prim's algorithm in 2 minutes." [Online]. Available: https://www.youtube.com/watch?v=cplfcGZmX7I&ab_channel=MichaelSambol