Continued Fractions:

Def: A finite continued fraction is a fraction of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots \, \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}}}$$

where the a_i's are real #'s, $a_i > 0$ if $i > 0$. We say that the numbers $a_1, \dots, a_n$ are the partial denominators of this fraction. We say the fraction is simple if all the a_i's are in $\mathbb{Z}$.

It is easy to see that any finite continued fraction is a rational number. For example, we have

$$\frac{123}{36} = 3 + \cfrac{1}{2 + \cfrac{1}{2 + \frac{1}{2}}}$$

To see this, observe that

$$3 + \cfrac{1}{2 + \cfrac{1}{2 + 1/2}} = 3 + \cfrac{1}{2 + 2/5}$$

$$= 3 + \frac{5}{12}$$

$$= 41/12 = \frac{123}{36}.$$

It turns out that the converse holds as well!

<u>Thm:</u> Any number $\frac{a}{b} \in \mathbb{Q}$ can be written as a simple continued fraction.

<u>Proof:</u> Let $\frac{a}{b} \in \mathbb{Q}$ with $b > 0$. We apply the division Euclidean alg. to this to obtain

$$\begin{aligned}
a &= b a_0 + r_1 && 0 < r_1 < b, \\
b &= r_1 a_1 + r_2 && 0 < r_2 < r_1 \qquad a_1 > 0 \\
&\ \vdots \\
r_{n-2} &= r_{n-1} a_{n-1} + r_n && 0 < r_n < r_{n-1} \qquad a_{n-1} > 0 \\
r_{n-1} &= r_n a_n + 0 && \qquad a_n > 0.
\end{aligned}$$

We can rewrite these equations as

$$\frac{a}{b} = a_0 + \frac{r_1}{b}$$

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

$$\frac{b}{r_1} = a_1 + \frac{r_2}{r_1} = a_1 + \frac{1}{\frac{r_1}{r_2}}$$

$$\vdots$$

$$\vdots$$

$$\frac{r_{n-1}}{r_n} = a_n.$$

We now combine the equations to obtain the result. $\boxtimes$

<u>Example:</u> Compute the continued fraction of $\frac{13}{93}$.

Apply Euclid's algorithm to 13 and 93:

$$93 = 13 \cdot 7 + 2 \quad \rightsquigarrow \quad \frac{93}{13} = 7 + \frac{2}{13}$$

$$13 = 2 \cdot 6 + 1 \qquad \frac{13}{2} = 6 + \frac{1}{2}$$

$$2 = 2 \cdot 1 + 0$$

So we have

$$\frac{13}{93} = \frac{1}{\frac{93}{13}}$$

$$= \frac{1}{7 + \frac{2}{13}}$$

$$= \frac{1}{7 + \frac{1}{\frac{13}{2}}}$$

$$= \frac{1}{7 + \frac{1}{6 + \frac{1}{2}}}.$$

These, as you can see, are a huge pain to write out. We use the traditional shorthand

$$[a_0; a_1, a_2, \ldots, a_n]$$

to represent

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_{n-1} + \frac{1}{a_n}}}}}.$$

Thus, $\frac{13}{93} = [0; 7, 6, 2]$.

The relevant SAGE command is

continued_fraction (x).

For example,

continued_fraction (173/97)

returns

[0, 1, 7, 4, 1, 4].

Note that we have said nothing as of yet in regard to uniqueness. Observe that if $a_n > 1$, then

$$a_n = (a_n - 1) + 1 = (a_n - 1) + \frac{1}{1}$$

and so

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

If $a_n = 1$, then

$$a_{n-1} + \frac{1}{a_n} = a_{n-1} + 1$$

and so

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

So each rational number has at least 2 representations. (It turns out these are the only ones!)

Def: Let $[a_0; a_1, \ldots, a_n]$ be a continued fraction. The continued fraction $[a_0; a_1, \ldots, a_k]$ w/ $1 \le k \le n$ is called the k^{th} convergent and denoted by C_k. We set $C_0 = a_0$.

Example: $\frac{13}{93} = [0; 7, 6, 2]$.

Thus,

$$C_0 = 0$$

$$C_1 = 0 + \frac{1}{7} = \frac{1}{7}$$

$$C_2 = 0 + \frac{1}{7 + \frac{1}{6}} = \frac{6}{43}$$

$$C_3 = 0 + \frac{1}{7 + \frac{1}{6 + \frac{1}{2}}} = \frac{13}{93}$$

The SAGE command is

```
v= continued_fraction (13/93)
```

Convergent (v, k)

gives C_k.

Note: Let $C_k = [a_0; a_1, \dots, a_k]$ be the k^{th} convergent of a continued fraction. If we replace a_k w/ $a_k + \frac{1}{a_{k+1}}$, then we obtain

$$[a_0; a_1, \dots, a_k + \frac{1}{a_{k+1}}] = [a_0; a_1, \dots, a_k, a_{k+1}]$$
$$= C_{k+1}.$$

This fact will be used when we prove properties of convergents as they allow us to easily apply induction.

Define: $p_{-2} = 0$, $p_{-1} = 1$ $\quad q_{-2} = 1$, $q_{-1} = 0$.

and $\quad p_k = a_k p_{k-1} + p_{k-2}$, $q_k = a_k q_{k-1} + q_{k-2}$. $\quad 0 \le k \le n$

where $[a_0; a_1, \dots, a_n]$ is a finite simple continued fraction.

Theorem: The k^{th} convergent of the simple continued fraction $[a_0; a_1, \dots, a_n]$ has convergent

$$C_k = \frac{p_k}{q_k} \qquad 0 \le k \le n.$$

Proof: We use induction, but as the relation go back a couple of steps to calculate the next one, we must establish the first few

Cases.

$K=0$: $p_0 = a_0 p_{-1} + p_{-2} = a_0$, $q_0 = a_0 q_{-1} + q_{-2} = 1$.

Thus, $\frac{p_0}{q_0} = a_0$ ✓

$K=1$: $p_1 = a_1 p_0 + p_{-1} = a_1 a_0 + 1$, $q_1 = a_1 q_0 + q_{-1} = a_1 + 0 = a_1$

$$\frac{p_1}{q_1} = \frac{a_1 a_0 + 1}{a_1} = a_0 + \frac{1}{a_1} = c_1 .$$

~~[illegible]~~

We assume inductively that for $2 \le K < n$. we have

$$\frac{p_K}{q_K} = C_K ,$$

i.e,
$$C_K = \frac{a_K p_{K-1} + p_{K-2}}{a_K q_{K-1} + q_{K-2}} .$$

Note here that the $p_{K-1}, p_{K-2}, q_{K-1}, q_{K-2}$ depend only on the initial conditions and $a_0, .., a_{K-1}$. Thus, if we have a different value for a_K, this will not affect the equality. We replace a_K by $a_K + \frac{1}{a_{K+1}}$. Thus we have:

$$C_{K+1} = [a_0; a_1, \dots, a_{K-1}, a_K + \frac{1}{a_{K+1}}]$$

$$= \frac{\left(a_K + \frac{1}{a_{K+1}}\right) p_{K-1} + p_{K-2}}{\left(a_K + \frac{1}{a_{K+1}}\right) q_{K-1} + q_{K-2}}$$

$$= \frac{a_K a_{K+1} p_{K-1} + p_{K-1} + a_{K+1} p_{K-2}}{a_K a_{K+1} q_{K-1} + q_{K-1} + a_{K+1} q_{K-2}}$$

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

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

as desired. Thus the result holds by induction. ∎

This result allows one to calculate the convergents recursively, which can be much easier!

<u>Example:</u> Recall that $\frac{13}{93} = [0; 7, 6, 2]$.

So we have:

$p_0 = 0 \cdot p_{-1} + p_{-2} = 0$ $\qquad$ $q_0 = 0 q_{-1} + q_{-2} = 1.$

$p_1 = 7p_0 + p_{-1} = 1$ $\qquad$ $q_1 = 7q_0 + q_{-1} = 7$

$p_2 = 6p_1 + p_0 = 6$ $\qquad$ $q_2 = 6q_1 + q_0 = 43$

$p_3 = 2p_2 + p_1 = 13$ $\qquad$ $q_3 = 2q_2 + q_1 = 93$

Thus, the convergents are

$C_0 = 0$

$C_1 = \frac{1}{7}$

$C_2 = \frac{6}{43}$

$C_3 = \frac{13}{93}$

Which is exactly what we calculated last time!

Thm: Let $C_k = \frac{p_k}{q_k}$ be the k^{th} convergent of the finite simple continued fraction $[a_0; a_1, \ldots, a_n]$. Then

$$p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1} \qquad 1 \le k \le n.$$

Before we prove this theorem we state and prove an easy corollary.

Corl: For $1 \le k \le n$, p_k and q_k are relatively prime.

Proof: Suppose $d = \gcd(p_k, q_k)$. Then $d \mid (p_k q_{k-1} - q_k p_{k-1}) = (-1)^{k-1}$

$\Rightarrow d = 1$. $\boxtimes$

Proof (Thm): We prove this by induction on k. The base case of $k=1$ is handled easily

$$\begin{aligned} p_1 q_0 - q_1 p_0 &= (a_1 a_0 + 1)\cdot 1 - a_1 a_0 \\ &= 1 = 1^{1-1}. \end{aligned}$$

Assume the statement is true for all $1 \le j \le k$ for some k. Then

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

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

$$= -(-1)^{k-1} \quad \text{(by induction hyp.)}$$

$$= (-1)^k.$$

Thus the result holds by induction. ▣

We now investigate our first application of continued fractions. We show they can be used to investigate the Diophantine equation

$$ax + by = c.$$

We already studied this before. We can safely assume $\gcd(a,b) \mid c$ for otherwise there are no solutions. In fact, if $\gcd(a,b) \mid c$ we can divide it out of the equation so we might as well also assume that $\gcd(a,b) = 1$. Write $\frac{a}{b} = [a_0; a_1, \ldots, a_n]$. We have

$$C_{n-1} = \frac{p_{n-1}}{q_{n-1}}, \quad C_n = \frac{p_n}{q_n} = \frac{a}{b}.$$

In particular, we have $p_n b = q_n a$ and since $\gcd(a,b) = 1$, we must have $p_n = a$ and $q_n = b$. (Note if $\gcd(a,b) \neq 1$ this would not necessarily be true!)

Our previous theorem gives

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

i.e.

$$a q_{n-1} - b p_{n-1} = (-1)^{n-1}.$$

Thus, if we set $x = q_{n-1}$ and $y = -p_{n-1}$ we have

$$ax + by = (-1)^{n-1}.$$

If n is odd, then $n-1$ is even and so we have

$$ax + by = 1.$$

In particular, this gives $x = q_{n-1}$ and $y = -p_{n-1}$ as a solution to our original equation.

If n is even, then $n-1$ is odd and so

$$ax + by = -1.$$

Thus,

$$a(-x) + b(-y) = 1$$

and our solution is $x = -q_{n-1}$, $y = p_{n-1}$. Thus, we obtain a solution to the Diophantine equation arising from the $n-1^{\text{st}}$ convergent of $\frac{a}{b}$.

Example: Find solutions to the Diophantine equation

$$65x + 465y = 5$$

We begin by dividing $\gcd(65, 465) = 5$ out of the equation:

$$13x + 93y = 1.$$

The convergents of $\frac{13}{93}$ have already been computed,

$$C_2 = \frac{6}{43}.$$

Thus, $p_2 = 6$, $q_2 = 43$. Since $n = 3$ is odd,

$$x = 43, \quad y = -6$$

is a solution, i.e.

$$13(43) + 93(-6) = 1.$$

Hence,

$$13(5 \cdot 43) + 93(5 \cdot -6) = 5$$

~~[illegible]~~

~~[illegible]~~

~~[illegible]~~

~~[illegible]~~

Thus we have

$$65(43) + 465(-6) = 5$$

so $x = 43$, $y = -6$ is a solution to the original equation. To obtain all solutions, we have

$$x = 43 + 93t, \quad y = -6 - 13t.$$

<u>Lemma:</u> Let $C_k = \frac{p_k}{q_k}$ for the simple continued fraction $[a_0; a_1, \ldots, a_n]$. Then $q_{k-1} \le q_k$ for $1 \le k \le n$ with strict inequality when $k > 1$.

<u>Proof:</u> Once again we use induction on k. $q_0 = 1 \le a_1 = q_1$. Thus the base case holds. Assume it is true for $1 \le j \le k$. Then

$$q_{k+1} = a_{k+1} q_k + q_{k-1} > a_{k+1} q_k$$
$$\geq 1 \cdot q_k.$$

Thus the result is true by induction. ∎

We conclude the basics of finite simple continued fractions with the following theorem, telling us how the convergents converge to $[a_0; a_1, \ldots, a_n]$.

Thm: Let C_k be the convergents of a finite simple continued fraction.

We have

$$C_0 < C_2 < C_4 < \cdots$$

and

$$C_1 > C_3 > C_5 > \cdots .$$

Moreover, every convergent with an odd subscript is larger than every convergent with an even subscript.

Proof: We begin by observing that

$$\begin{aligned}
C_{k+2} - C_k &= (C_{k+2} - C_{k+1}) + (C_{k+1} - C_k) \\
&= \left(\frac{p_{k+2}}{q_{k+2}} - \frac{p_{k+1}}{q_{k+1}}\right) + \left(\frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k}\right) \\
&= \left(\frac{p_{k+2}q_{k+1} - p_{k+1}q_{k+2}}{q_{k+1}\,q_{k+2}}\right) + \left(\frac{p_{k+1}q_k - p_k q_{k+1}}{q_k q_{k+1}}\right) \\
&= \frac{(-1)^{k+1}}{q_{k+1}\,q_{k+2}} + \frac{(-1)^k}{q_k\,q_{k+1}} \\
&= \frac{(-1)^k\,(q_{k+2} - q_k)}{q_k\,q_{k+1}\,q_{k+2}} .
\end{aligned}$$

We know that $q_i > 0$ for all $i \geq 0$ (look at the def!),

and $q_{k+2} - q_k > 0$ by the previous lemma. Thus, we have that $C_{k+2} - C_k$ has the same sign as $(-1)^k$.

Thus, if k is even $C_{k+2} - C_k$ is positive and so we have

$$C_{2j+2} > C_{2j}$$

for all j. Hence

$$C_0 < C_2 < C_4 < \dots$$

If k is odd, $C_{k+2} - C_k$ is negative and so

$$C_{2j+1} < C_{2j-1}$$

$\Rightarrow$

$$C_1 > C_3 > C_5 > \dots$$

Now we just need to show that C_{2r-1} is greater than C_{2s}

for all r, s. Recall that

$$p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1}.$$

Thus,

$$C_k - C_{k-1} = \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}}$$

$$= \frac{p_k q_{k-1} - q_k p_{k-1}}{q_k q_{k-1}}$$

$$= \frac{(-1)^{k-1}}{q_k q_{k-1}}.$$

Thus, we have

$$C_{2j} - C_{2j-1} = \frac{(-1)^{2j-1}}{q_{2j} q_{2j-1}} < 0.$$

$\Rightarrow$ $C_{2j} < C_{2j-1}$.

Thus,

$$C_{2s} < C_{2s+2r} < C_{2s+2r-1} < C_{2r-1}.$$

This gives the result. $\blacksquare$

What this theorem is telling us is that the odd convergents are converging from above and the even convergents are converging from below.

It will be important to note that we did not use the length of the continued fraction at all in this proof. This will allow us to conclude the same property for simple infinite continued fractions.

<u>Def.:</u> An <u>infinite continued</u> fraction is a continued fraction of the form

$$a_0 + \cfrac{b_0}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cdots}}}$$

where $a_0, a_1, a_2, \ldots$, $b_0, b_1, b_2, \ldots$ are real numbers.

We will be interested in simple continued fractions. These are of the form

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

with the a_i integers and $a_1, a_2, \ldots$ all positive. Again we use the compact notation $[a_0; a_1, a_2, \ldots]$ to denote the continued fraction. The first thing we need to establish is that this actually makes sense, i.e., that

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

converges to a real number.

Using the theorem proved last time we have

$$(*)\ C_0 < C_2 < C_4 < \cdots < C_{2n} < \cdots < C_{2m+1} < \cdots < C_5 < C_3 < C_1$$

Thus, $\{C_{2k}\}$ is a monotonically increasing sequence bounded above by C_1. Similarly, $\{C_{2k+1}\}$ is a monotonically decreasing sequence bounded below by C_0. Thus, each of these must converge, say to α and β. We would like to conclude that $\alpha = \beta$ and so both sequences can be defined to have limit $[a_0; a_1, \ldots]$.

Recall

$$p_{2n+1} q_{2n} - p_{2n} q_{2n+1} = (-1)^{2n} = 1.$$

Thus,

$$\beta - \alpha < C_{2k+1} - C_{2k} = \frac{p_{2k+1}}{q_{2k+1}} - \frac{p_{2k}}{q_{2k}}$$

$$= \frac{p_{2k+1} q_{2k} - p_{2k} q_{2k+1}}{q_{2k} q_{2k+1}}$$

$$= \frac{1}{q_{2k} q_{2k+1}}.$$

However, we know q_j is an increasing unbounded sequence and so as $k \to \infty$, we get $\beta = \alpha$. Thus we have that our infinite simple continued fraction actually makes sense!

<u>Thm:</u> The value of any infinite continued fraction is an irrational number.

<u>Proof:</u> Suppose not, so $[a_0; a_1, a_2, \ldots] = \frac{a}{b} \in \mathbb{Q}$. Then $\frac{a}{b} = \lim_{n \to \infty} C_n$. We know that $\frac{a}{b}$ must lie

strictly between C_n and C_{n+1} for any n (one is larger, one is smaller depending on the parity of n!)

$$0 < \left| \frac{a}{b} - C_n \right| < \left| C_{n+1} - C_n \right|$$

$$= \left| \frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n q_{n+1}}$$

i.e.,

$$\left| \frac{a}{b} - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}$$

$$\Rightarrow \quad \left| q_n a - b p_n \right| < \frac{b}{q_{n+1}} \qquad (\text{assume wlog } b > 0)$$

Since q_i increase w/o bound, we choose large enough n so that $b < q_{n+1} \Rightarrow 0 < \frac{b}{q_{n+1}} < 1$.

$$\Rightarrow \quad 0 < \left| q_n a - b p_n \right| < 1$$

This is a contradiction because $q_n a - b p_n \in \mathbb{Z}$. ∎

__Thm:__ If $[a_0; a_1, a_2, \ldots] = [b_0; b_1, b_2, \ldots]$ then $a_n = b_n \;\; \forall\, n \geq 0$.

__Proof:__ We begin by noting the following 2 facts:

① Since $[a_0; a_1, a_2, \dots, a_n] = a_0 + \dfrac{1}{[a_1; a_2, \dots, a_n]}$,

we have

$$\lim_{n\to\infty} [a_0; a_1, \dots, a_n] = a_0 + \frac{1}{\lim_{n\to\infty} [a_1; a_2, \dots, a_n]}$$

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

Thus,

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

② Let $x = [a_0; a_1, a_2, \dots]$. Then $C_0 < x < C_1$, i.e.,

$$a_0 < x < a_0 + \frac{1}{a_1}.$$

Since $a_1 \geq 1$, we have

$$a_0 < x \leq a_0 + 1.$$

Thus, we have $\lfloor x \rfloor = a_0$.

Now assume

$$[a_0; a_1, a_2, \dots] = [b_0; b_1, b_2, \dots]. = x$$

Thus, from ① above we have

$$a_0 + \frac{1}{[a_1; a_2, \dots]} = b_0 + \frac{1}{[b_1; b_2, \dots]} = x$$

$\Rightarrow$ (by ②) $a_0 = \lfloor x \rfloor = b_0$. Then we must have

$$[a_1; a_2, \dots] = [b_1; b_2, \dots].$$

Proceed now by induction to finish the proof for all n. ∎

Since all our ~~fo~~ continued fractions are now infinite, it becomes more difficult to write them down explicitly. This is much like the case when one does decimal expansion. As in that situation, it can happen that we are lucky and get a block of integers that repeat. Namely, we may have

$$[a_0; a_1, \ldots, a_m, b_1, \ldots, b_n, b_1, \ldots, b_n, \ldots]$$

In this case we write

$$[a_0; a_1, \ldots, a_m, \overline{b_1, \ldots, b_n}].$$

If $b_1, \ldots, b_n$ is the smallest block that repeats we say that $b_1, \ldots, b_n$ is the <u>period</u> of the expansion and the <u>length</u> of the period is n.

We finally compute some examples.

<u>Example:</u> What number does the continued fraction $[1; \overline{5,6}]$ represent?

Note that since 5, 6 repeat, we can set $x = [\overline{5,6}]$ and write

$$[1; \overline{5,6}] = 1 + \cfrac{1}{5 + \cfrac{1}{6 + \cfrac{1}{x}}}.$$

So we really need to figure out x. We have

$$x = 5 + \frac{1}{6+\frac{1}{x}} = 5 + \frac{x}{6x+1}$$

$$= \frac{30x + 5 + x}{6x+1}$$

i.e, we have that x satisfies:

$$6x^2 + x = 31x + 5 \iff 6x^2 - 30x - 5 = 0$$

$$\iff x = \frac{30 \pm \sqrt{900 - (-5)(6)}}{12}$$

$$= \frac{5}{2} \pm \frac{\sqrt{930}}{12}$$

Thus, we have

$$[1; \overline{5,6}] = 1 + \frac{1}{\frac{5}{2} + \frac{\sqrt{930}}{12}} \qquad \left(\text{since } x > 0 \text{ and } \tfrac{5}{2} - \tfrac{\sqrt{930}}{12} < 0\right)$$

$$= 1 + \frac{12}{30+\sqrt{930}}\left(\frac{30-\sqrt{930}}{30-\sqrt{930}}\right)$$

$$= 1 + \frac{360 - 12\sqrt{930}}{900 - 930} = 1 + \frac{360 - 12\sqrt{930}}{-30}$$

$$= 1 + \frac{60 - 2\sqrt{930}}{5} = \frac{65 - 2\sqrt{930}}{5}.$$

Recall when we were working with finite continued fractions first we saw each finite simple continued fraction is a rational number. We then proved that each rational number could be expressed as a finite simple continued fraction. We have shown that a simple infinite continued fraction must be an irrational number. We now show each irrational number is a simple infinite continued fraction.

Thm: Every irrational number has a unique representation as a simple infinite continued fraction.

Proof: We have already seen that if

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

then $a_i = b_i \ \forall i \geq 0$, so if a number has a representation as an infinite continued (simple) fraction, it is necessarily unique! We actually give an algorithm to associate the continued fraction to the real numbers.

Let x_0 be our irrational real number. Set

$$x_1 = \frac{1}{x_0 - \lfloor x_0 \rfloor},$$

$$x_2 = \frac{1}{x_1 - \lfloor x_1 \rfloor}, \quad \ldots \quad \text{etc.}$$

Set $a_0 = \lfloor x_0 \rfloor$, $a_1 = \lfloor x_1 \rfloor$, etc.

The a_k are defined inductively by

$$a_k = \lfloor x_k \rfloor \quad \text{where} \quad x_{k+1} = \frac{1}{x_k - a_k} \qquad k \geq 0.$$

It is clear that if x_i is irrational, so is x_{i+1}. Since x_0 is irrational, so are all the x_i's. Thus we have

$$0 < x_k - a_k = x_k - \lfloor x_k \rfloor < 1.$$

$$\Rightarrow \qquad \frac{1}{x_k - a_k} > 1 \quad \text{for all } k$$

$$\Rightarrow \qquad x_{k+1} > 1 \quad \text{for all } k \quad \left(\text{since } x_{k+1} = \frac{1}{x_k - a_k} \right).$$

Thus, $a_{k+1} = \lfloor x_{k+1} \rfloor \geq 1$ for all $k \geq 0$. This gives an infinite sequence of integers so that $a_k \geq 1$ except possibly for $k = 0$. By making substitutions we obtain

$$x_0 = a_0 + \frac{1}{x_1}$$

$$= a_0 + \cfrac{1}{a_1 + \cfrac{1}{x_2}}$$

$$= \cdots = [a_0; a_1, \ldots, a_n, x_{n+1}] \qquad \forall n \geq 0$$

We now show that $x_0 = [a_0; a_1, a_2, \ldots]$. We have

$$x_0 = C'_{n+1} = [a_0; a_1, \ldots, a_n, x_{n+1}]$$

$$= \frac{p'_{n+1}}{q'_{n+1}} = \frac{x_{n+1} p'_n + p'_{n-1}}{x_{n+1} q'_n + q'_{n-1}}.$$

However, we see that $C_k = C'_k$ for $0 \leq k \leq n$ were C_k are the convergents of $[a_0; a_1, a_2, \ldots]$ so that

$$x_0 = \frac{x_{n+1} p_n + p_{n-1}}{x_{n+1} q_n + q_{n-1}}.$$

Thus,

$$x_0 - C_n = \frac{x_{n+1} p_n + p_{n-1}}{x_{n+1} q_n + q_{n-1}} - \frac{p_n}{q_n}$$

$$= \frac{\cancel{q_n p_n x_{n+1}} + p_{n-1} q_n - \cancel{x_{n+1} p_n q_n} - p_n q_{n-1}}{q_n (x_{n+1} q_n + q_{n-1})}$$

$$= \frac{-1 (p_n q_{n-1} - p_{n-1} q_n)}{q_n (x_{n+1} q_n + q_{n-1})}$$

$$= \frac{-1 (-1)^{n-1}}{q_n (x_{n+1} q_n + q_{n-1})}$$

$$= \dots = [a_0; a_1, \dots, a_n, x_{n+1}]. \qquad \forall\ n > 0.$$

We now show that in fact $x_0 = [a_0; a_1, a_2, \dots]$.

Let n_0 be a fixed positive integer. Then the first n convergent $C_k = p_k/q_k$ $0 \le k \le n$ are the same for $[a_0; a_1, \dots]$ as $[a_0; a_1, a_2, \dots, a_n, x_{n+1}]$. Denote the convergents of $[a_0; a_1, a_2, \dots, a_n, x_{n+1}]$ by C_k' and C_k for the convergents of $[a_0; a_1, a_2, \dots]$. Recall that

$$C_{n+1} = [a_0; a_1, \dots, a_n, a_{n+1}]$$

$$= [a_0; a_1, \dots, a_n + \frac{1}{a_{n+1}}].$$

Thus,

$$x_0 = C_{n+1}' = [a_0; a_1, \dots, a_n, x_{n+1}]$$

$$= [a_0; a_1, \dots, a_n + \frac{1}{x_{n+1}}]$$

$$= \frac{p_{n+1}'}{q_{n+1}'} = \frac{a_{n+1}' p_n + p_{n-1}}{a_{n+1}' q_n + q_{n-1}}$$

$$= \left(a_n + \frac{1}{x_{n+1}}\right)$$

By definition $x_{nn} > a_{n+1}$, so

$$|x_0 - C_n| = \frac{1}{q_n(x_{nn}q_n + q_{n-1})} < \frac{1}{q_n(a_{nn}q_n + q_{n-1})}$$

$$= \frac{1}{q_n q_{n+1}}.$$

Since the q_i are integers increasing w/o bound as $n \to \infty$, we see that

$$\lim_{n \to \infty} C_n = x_0$$

as desired. ∎

Cor: If $C_n = p_n/q_n$ is the n^{th} convergent for x, then $\left|x - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{nn}} \leq \frac{1}{q_n^2}$.

Example: Compute the continued fraction of $\sqrt{2}$.

Note that $1^2 = 1 < 2 < 2^2 = 4$, so $\lfloor \sqrt{2} \rfloor = 1$.

Thus, $x_0 = 1 + (\sqrt{2} - 1)$. $a_0 = 1$.

$$x_1 = \frac{1}{x_0 - \lfloor x_0 \rfloor} = \frac{1}{\sqrt{2} - 1}$$

$$= \frac{\sqrt{2}+1}{1} = \sqrt{2} + 1$$

Thus, $a_1 = \lfloor \sqrt{2} + 1 \rfloor = \lfloor \sqrt{2} \rfloor + 1$
$= 2$.

$$x_2 = \frac{1}{x_1 - \lfloor x_1 \rfloor} = \frac{1}{\sqrt{2} + 1 - 2}$$

$$= \frac{1}{\sqrt{2}-1} = \sqrt{2}+1 = x_1$$

Thus, $a_1 = a_2$. We see that now we will get

$x_3 = x_2$, etc. so we have

$$\sqrt{2} = [1; \bar{2}].$$

Example: Give the first few convergents for π.

$$x_0 = 3 + (\pi - 3). \qquad a_0 = 3$$

$$x_1 = \frac{1}{x_0 - \lfloor x_0 \rfloor} = \frac{1}{0.1415\ldots} = 7.067\ldots \qquad a_1 = 7$$

$$x_2 = \frac{1}{x_1 - \lfloor x_1 \rfloor} = \frac{1}{7.067\ldots - 7} \qquad a_2 = 15$$

$$= 15.9965\ldots$$

$$x_3 = \frac{1}{x_2 - \lfloor x_2 \rfloor} = \frac{1}{15.9965\ldots - 15} = \qquad a_3 = 1.$$

$$= 1.0034\ldots$$

Thus,

$$\pi \doteq [3; 7, 15, 1, \ldots].$$

We could continue to compute more convergents, but we would never end up with a repeating pattern. In fact, we already must know the decimal expansion of π in order to compute its convergents.

Continued-fraction (pi) gives

$[3; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1]$.

One must be careful with the last few digits computed by SAGE. For instance, we know $\sqrt{2} = [1; \overline{2}]$, but SAGE gives

$$\sqrt{2} = [1; 2, 2, \ldots, 2, 1, 1].$$

One of the nice applications of continued fractions is to rational approximations to irrational numbers. Rational numbers are easy to deal with. However, irrational numbers can be very difficult to deal with in which case having a good rational approximation is important. For instance, whenever you do a numerical computation with π you are really using a rational approximation. When one tries to obtain a good rational approximation to an irrational # x one usually means finding the closest rational # $\frac{a}{b}$ where b is bounded by some fixed number.

Given any irrational number x, $\exists\ \frac{c}{b} \in \mathbb{Q}$ with

$$\frac{c}{b} < x < \frac{c+1}{b}.$$

Thus we have

$$|x - \frac{c}{b}| < \frac{1}{b}.$$

If we let $a = c$ or $c+1$ depending on which is closer to x, $\frac{c}{b}$ or $\frac{c+1}{b}$, we have

$$|x - \frac{a}{b}| < \frac{1}{2b}.$$

We saw before in our corollary that if $[a_0; a_1, \ldots]$ is the continued fraction of x, then

$$|x - C_k| = |x - \frac{p_k}{q_k}| < \frac{1}{q_k^2}.$$

So continued fractions give us significantly better approximations than naively we would expect to get.

<u>Thm:</u> If $1 \le b \le q_n$, then the rational number $\frac{a}{b}$ satisfies

$$|x - \frac{p_n}{q_n}| \le |x - \frac{a}{b}|.$$

This theorem is saying that the continued fraction convergents are the best possible rational number approximations to irrational #'s.

Proof (Thm): Suppose

$$\left|x - \frac{p_n}{q_n}\right| > \left|x - \frac{a}{b}\right|.$$

Then

$$|q_n x - p_n| = q_n \left|x - \frac{p_n}{q_n}\right|$$
$$> q_n \left| x - \frac{a}{b}\right|$$
$$> b \left| x - \frac{a}{b}\right|$$
$$= |bx - a|.$$

This contradicts the lemma. ∎

Example: This theorem allows us to conclude our convergents to π are very good approximations. Recall

$$\pi = [3; 7, 15, 1, 292, 1, \ldots]$$

Then

$$C_1 = \text{convergent}(\pi, 1) = \frac{22}{7}$$

$$C_2 = \frac{333}{106}$$

$$C_3 = \frac{355}{113}$$

$$C_4 = \frac{103993}{33102}$$

We know that $|\pi - C_k| < \frac{1}{q_k^2}$.

Before we prove this theorem we need the following technical lemma.

Lemma: Let $C_n = p_n/q_n$ be the n^{th} convergent of $x = [a_0; a_1, \ldots]$. If a and b are integers, w/ $1 \le b < q_{n+1}$, then

$$|q_n x - p_n| \le |bx - a|.$$

Proof: Consider the system of equations

$$p_n y + p_{n+1} z = a$$
$$q_n y + q_{n+1} z = b.$$

The determinant of this system is $p_n q_{n+1} - p_{n+1} q_n = (-1)^{n+1}$, so it has a unique integral solution. In particular,

$$y = (-1)^{n+1}(a q_{n+1} - b \cdot p_{n+1})$$
$$z = (-1)^{n+1}(b p_n - a q_n).$$

Claim: $y \neq 0$.

Pf. If $y = 0$, then $a q_{n+1} = b p_{n+1}$. However, $\gcd(p_{n+1}, q_{n+1}) = 1$
$\Rightarrow b \mid q_{n+1}$. But we assumed $b < q_{n+1}$. #.

If $z = 0$, then we have ~~[illegible]~~ $p_n y = a$, $q_n y = b$ and so

$$|bx - a| = |q_n y\, x - p_n y|$$
$$= |y|\,|q_n x - p_n| \ge |q_n x - p_n|$$

since $y \in \mathbb{Z}$. Thus, if $z = 0$ the result is true.

We now assume $z \neq 0$.

So for example we have

$$|\pi - c_1| < \frac{1}{49}$$

$$|\pi - c_2| < \frac{1}{11236}$$

$$|\pi - c_3| < \frac{1}{12769}$$

$$|\pi - c_4| < \frac{1}{1095742404}$$

Our previous theorem tells us for instance that we cannot find a rational number with denominator less than or equal to 106 closer to π than $\frac{333}{106}$.

It would be nice if whenever one had $|x - \frac{a}{b}| < \frac{1}{b^2}$ that $\frac{a}{b}$ must be a convergent of x. This isn't quite true. What we do have is the following theorem.

<u>Thm:</u> Let $x \in \mathbb{R} \setminus \mathbb{Q}$. Let $\frac{a}{b} \in \mathbb{Q}$ w/ $\gcd(a,b) = 1$, $b \geq 1$. If

$$|x - \frac{a}{b}| < \frac{1}{2b^2}.$$

then $\frac{a}{b}$ is one of the convergents $\frac{p_n}{q_n}$.

Proof: Suppose $\frac{a}{b}$ is not a convergent. Since the q_n are increasing integers, $\exists$ n so that $q_n \leq b < q_{n+1}$. We have

$$|q_n x - p_n| \underbrace{\leq}_{\text{by lemma}} |bx - a| = b\left|x - \frac{a}{b}\right| < b\left(\frac{1}{2b^2}\right) = \frac{1}{2b}.$$

Thus,

$$\left|x - \frac{p_n}{q_n}\right| < \frac{1}{2bq_n}.$$

Since $\frac{a}{b} \neq \frac{p_n}{q_n}$, we have $aq_n - bp_n \neq 0$ and is in $\mathbb{Z}$.

$\Rightarrow$ $1 \leq |bp_n - aq_n|$. Thus,

$$\frac{1}{bq_n} < \left|\frac{bp_n - aq_n}{bq_n}\right| = \left|\frac{p_n}{q_n} - \frac{a}{b}\right| \overset{\Delta\text{-inequality.}}{\leq} \left|\frac{p_n}{q_n} - x\right| + \left|x - \frac{a}{b}\right|$$

$$< \frac{1}{2bq_n} + \frac{1}{2b^2}.$$

$\Rightarrow$ $2b < q_n$, i.e., $b < q_n$. #. ■

Our remaining goal with continued fractions is to use them to characterize solutions to Pell's equation

$$x^2 - dy^2 = 1.$$

Before we can do this, we need to spend a bit of time on periodic continued fractions. Recall a continued fraction is periodic if $\exists\, m, n$ so that

$$[a_0; a_1, a_2, a_3, \ldots] = [a_0; a_1, \ldots, a_m, \overline{b_1, \ldots, b_n}]$$

where the overline represents the fact that the $b_1, \ldots, b_n$ repeat forever.

We have seen for example that $\sqrt{2} = [1; \overline{2}]$. It turns out that irrational numbers of the form $a + b\sqrt{D}$ w/ $a, b \in \mathbb{Q}$, $D \in \mathbb{Z}_{>0}$, all have periodic continued fractions and if a continued fraction is periodic then it must represent a number of this form.

<u>Def:</u> Let $x \in \mathbb{R} \setminus \mathbb{Q}$. We say x is a <u>quadratic irrational number</u> if $\exists\, a, b \in \mathbb{Q}$, $D \in \mathbb{Z}_{>0}$ so that $x = a + b\sqrt{D}$.

<u>Thm:</u> Any periodic simple continued fraction is a quadratic irrational number and vice versa.

Proof: Let $x = [a_0; a_1, \ldots, a_m, \overline{b_1, \ldots, b_n}]$ and

$$y = [\overline{b_1, \ldots, b_n}].$$

Note that $y = [b_1, \ldots, b_n, y] = \dfrac{p_{n+1}}{q_{n+1}} = \dfrac{y p_n + p_{n-1}}{y q_n + q_{n-1}}$.

So $y = \dfrac{y p_n + p_{n-1}}{y q_n + q_{n-1}}$. By clearing the denominator we obtain a quadratic equation in y. Thus, we must have that either y is a quadratic irrational number or a rational number. However, we know an infinite continued fraction cannot be rational so we must have that y is a quadratic irrational number.

We have

$$x = [a_0; a_1, \ldots, a_m; y] = \frac{p'_{m+1}}{q'_{m+1}} = \frac{y p'_{m} + p'_{m-1}}{y q'_{m} + q'_{m-1}}.$$

where we use $\frac{p_i'}{q_i'}$ to denote the convergents of x. We know $x \notin \mathbb{Q}$ because it is an infinite continued fraction. We showed above that y is a quadratic irrational, i.e. $\exists\, a, b, c, d \in \mathbb{Z}, D \in \mathbb{Z}_{>0}$ s.t. $y = \frac{a}{b} + \frac{c}{d}\sqrt{D}$. Plugging this in and clearing the denominator by multiplying by the conjugate shows x is a quadratic irrational as well. This proves one direction of the theorem. The "easy" direction.

We now wish to show that if we have a quadratic irrational then its continued fraction is periodic. To do this, we need to have a reasonable way to compute the continued fraction of $\sqrt{D}$.

First, observe that we can write any quadratic irrational in the form $x = \frac{a + \sqrt{D}}{b}$. (For example, if $x = \frac{\alpha}{\beta} + \frac{\gamma}{\delta}\sqrt{D}$, then

$$x = \frac{\alpha\delta + \sqrt{(\beta\gamma)^2 D}}{\beta\gamma}$$

). Note that D cannot be a perfect square since $x \notin \mathbb{Q}$. Multiplying the top and bottom by b we have

$$x = \frac{ab + \sqrt{b^2 D}}{b^2}.$$

This shows we can write

$$x = \frac{m_0 + \sqrt{D}}{b_0}$$

where $b_0 \mid (D - m_0^2)$, D, b_0, m_0 are integers, $b_0 \neq 0$ and D is not a perfect square. We put it in this form so we can easily compute the continued fraction expansion.

Recall that when we showed $x \in \mathbb{R} \setminus \mathbb{Q}$ had a continued fraction expansion, we used that we could define

$$a_i = \lfloor x_i \rfloor$$

$$x_{i+1} = \frac{1}{x_i - a_i} \qquad x_0 = x$$

and defined this way we had $x = [a_0; a_1, \dots]$. We now show that if we set $x_0 = x$, $x_i = \dfrac{m_i + \sqrt{D}}{\cancel{q_i}\, b_i}$ ~~and~~,

$a_i = \lfloor x_i \rfloor$, $m_{i+1} = a_i \overset{b_i}{\cancel{q_i}} - m_i$, $\overset{b_{i+1}}{\cancel{q_{i+1}}} = \dfrac{D - m_{i+1}^2}{\cancel{q_i}\, b_i}$, then these

satisfy the above equations and so give the continued fraction expansion of x.

Let $a_0 = \lfloor x_0 \rfloor$. We define the sequences x_i, a_i, m_i, and $\overset{b_i}{\cancel{q_i}}$ as above.

<u>Claim 1:</u> $m_i, \overset{b_i}{\cancel{q_i}} \in \mathbb{Z}$, $\overset{b_i}{\cancel{q_i}} \neq 0$ and $\overset{b_i}{\cancel{q_i}} \mid (D - m_i^2)$.

<u>Pf:</u> We use induction on i. The case $i = 0$ was handled above. Assume the statement for $i = n$. Then

$$m_{n+1} = a_n \overset{b_n}{\cancel{q_n}} - m_n \in \mathbb{Z}$$

by induction hypothesis.

$$\overset{b_{n+1}}{\cancel{q_{n+1}}} = \frac{D - m_{n+1}^2}{\cancel{q_n}\, b_n} = \frac{D - m_n^2}{\cancel{q_n}\, b_n} + 2a_n m_n - a_n^2 \overset{b_n}{\cancel{q_n}} \in \mathbb{Z}$$

by our induction hypothesis and so $\overset{b_n}{\cancel{q_n}} \mid (D - m_{n+1}^2)$. ~~Thus~~

We have that $\overset{b_{n+1}}{\cancel{q_{n+1}}} \neq 0$ for otherwise we would have $D = m_{n+1}^2$, but D is not a perfect square. Finally,

$\overset{b_n}{\cancel{q_n}} = \dfrac{D - m_{n+1}^2}{\underset{b_{n+1}}{\cancel{q_{n+1}}}}$ and $\overset{b_n}{\cancel{q_n}} \in \mathbb{Z}$, so $\overset{b_{n+1}}{\cancel{q_{n+1}}} \mid (D - m_{n+1}^2)$

as desired. $\square$

<u>Claim 2</u>: $x_{i+1} = \dfrac{1}{x_i - a_i}$

<u>Pf.</u> Note this is equivalent to showing $x_i - a_i = \dfrac{1}{x_{i+1}}$.

$$x_i - a_i = \frac{m_i + \sqrt{D}}{b_i q_i} - a_i$$

$$= \frac{m_i + \sqrt{D} - a_i q_i b_i}{q_i b_i}$$

$$= \frac{\sqrt{D} - m_{i+1}}{q_i b_i}$$

$$= \frac{D - m_{i+1}^2}{b_i q_i (\sqrt{D} + m_{i+1})}$$

$$= \frac{q_{i+1} b_{i+1}}{\sqrt{D} + m_{i+1}}$$

$$= \frac{1}{x_{i+1}}$$

as desired. $\square$

So we have shown that $x = [a_0; a_1, \ldots]$. Of course, the goal is to show this continued fraction is periodic so we still have a lot of work to do!

Write $\overline{x_i} = \dfrac{m_i - \sqrt{D}}{q_i b_i}$.

Now recall that when expanding x in a continued fraction, we have $x = [a_0; a_1, \ldots, a_{n-1}, x_n]$

$$= \frac{x_n p_{n-1} + p_{n-2}}{x_n q_{n-1} + q_{n-2}}.$$

Taking the conjugate of each side yields:

$$\overline{x} = \frac{\overline{x_n}\, p_{n-1} + p_{n-2}}{\overline{x_n}\, q_{n-1} + q_{n-2}}.$$

(Exercise: $\overline{x_n p_{n-1} + p_{n-2}} = \overline{x_n}\, p_{n-1} + p_{n-2}$ and $\overline{\left(\frac{a}{b}\right)} = \frac{\overline{a}}{\overline{b}}$.).

Solving the above equation for $\overline{x_n}$ we have

$$\overline{x_n} = \frac{-q_{n-2}}{q_{n-1}} \left(\frac{\overline{x} - p_{n-2}/q_{n-2}}{\overline{x} - p_{n-1}/q_{n-1}} \right).$$

As $n \to \infty$, p_{n-2}/q_{n-2} and p_{n-1}/q_{n-1} both converge to x, which is <u>not</u> equal to $\overline{x}$ since $\sqrt{D} \notin \mathbb{Q}$ and so

$$\left(\frac{\overline{x} - p_{n-2}/q_{n-2}}{\overline{x} - p_{n-1}/q_{n-1}} \right) \longrightarrow 1 \quad \text{as } n \to \infty.$$

Thus, for large enough n we have $\left(\dfrac{\bar{x} - p_{n-2}/q_{n-2}}{\bar{x} - p_{n-1}/q_{n-1}} \right) > 0$

and so $\bar{x}_n < 0$. Say for $n > N$ with N fixed.

We know that $x_n > 0$ for $n \geq 1$. (ratio of positive integers)

and so for $n > N$ we have $x_n - \bar{x}_n > 0$. Thus,

using that $x_n = \dfrac{m_n + \sqrt{D}}{b_n}$ we have

$$\frac{m_n + \sqrt{D}}{b_n} - \frac{m_n - \sqrt{D}}{b_n} > 0$$

i.e.

$$\frac{2\sqrt{D}}{b_n} > 0. \quad \text{for } n > N.$$

Thus, $b_n > 0$ for $n > N$.

We also have since $b_{n+1} = \dfrac{D - m_{n+1}^2}{b_n}$

that

$$b_n b_{n+1} = D - m_{n+1}^2 \leq D$$

$$b_n \leq b_n b_{n+1} \leq D$$

and

$$m_{n+1}^2 < m_{n+1}^2 + b_n b_{n+1} = D \qquad (\text{since } b_n > 0 \text{ for } n > N)$$

for $n > N$.

$\Rightarrow$ for $n>N$, $|m_{n+1}|<\sqrt{D}$.

We know D is a fixed positive integer, and so for $n>N$ we know that b_n and m_{n+1} can only take on finitely many values. Thus, for $n>N$, the pairs (m_n, b_n) can only take on finitely many values. So $\exists$ distinct integers j,k so that $(m_j, b_j) = (m_k, b_k)$. wlog $j<k$.

Thus,
$$X_j = \frac{m_j + \sqrt{D}}{q_j} = \frac{m_k + \sqrt{D}}{q_k} = X_k.$$

Thus,

$$x = [a_0; a_1, \ldots, a_{j-1}, \overline{a_j, a_{j+1}, \ldots, a_{k-1}}]$$

because $a_j = \lfloor X_j \rfloor = \lfloor X_k \rfloor = a_k$ and

$$\begin{aligned} m_{k+1} &= a_k b_k - m_k \\ &= a_j b_j - m_j \\ &= m_{j+1} \end{aligned}$$

$$\begin{aligned} b_{k+1} &= \frac{D - m_{k+1}^2}{b_k} = \frac{D - m_{j+1}^2}{b_j} \\ &= b_{j+1}. \end{aligned}$$

∎

In fact, one can actually show that if $\sqrt{D} \notin \mathbb{Q}$, then

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

We will not need this however and so will not spend time ~~passing~~ proving it.

We now return to Pell's equation

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

We can immediately remove several cases.

$D < -1$: Then $x^2 - Dy^2 \geq 0$ unless $x = y = 0$. The only solution in this case are $(\pm 1, 0)$.

$D = -1$: We are looking at $x^2 + y^2 = 1$. So we have the solutions ~~*~~ $(\pm 1, 0)$, $(0, \pm 1)$.

$D = 0$: We are looking at $x^2 = 1$, so $(\pm 1, y)$ is a solution for any y. Not real interesting though, as y does not even appear in the equation now!

Suppose D is a perfect square, $D = N^2$. Then we can write the equation

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

as

$$(x + Ny)(x - Ny) = 1.$$

Since we are only interested in integral x, y, this gives

$$x + Ny = \pm 1, \quad x - Ny = \pm 1. \quad \text{and must be equal}$$

Thus,

$$x = \frac{(x+Ny) + (x-Ny)}{2} = \pm 1.$$

So in this case the only solutions are $(\pm 1, 0)$.

We now get to the interesting part of the problem, namely when D is a positive integer that is not a square.

<u>Thm:</u> Let (p,q) be a positive solution of $x^2 - Dy^2 = 1$. Then p/q is a convergent of the continued fraction expansion of $\sqrt{D}$.

Before we prove the theorem we note that if (x_0, y_0) is a solution, so is $(-x_0, \pm y_0)$, $(x_0, \pm y_0)$. Thus we can always obtain all of the solutions by studying only those for $x_0, y_0 > 0$.

<u>Proof:</u> Since (p,q) is a solution, we have

$$(*) \qquad (p - \sqrt{D}q)(p + \sqrt{D}q) = 1.$$

Thus, we must have $p > q\sqrt{D}$. (~~otherwise~~ actually, this follows from $p^2 - Dq^2 = 1$. and $D \in \mathbb{Z}_{>1}$.

The equation $(*)$ gives

$$\frac{p}{q} - \sqrt{D} = \frac{1}{q(p + \sqrt{D}q)}.$$

Thus,

$$0 < \frac{p}{q} - \sqrt{D} < \frac{\sqrt{D}}{q(p+\sqrt{D}q)}$$

$$< \frac{\sqrt{D}}{q(q\sqrt{D} + q\sqrt{D})} = \frac{\sqrt{D}}{q^2 2\sqrt{D}} = \frac{1}{2q^2}.$$

Thus, our theorem before that if $\left|\frac{a}{b} - x\right| < \frac{1}{2b^2} \Rightarrow \frac{a}{b}$ is a convergent of x implies p/q is a convergent of $\sqrt{D}$. ∎

Be careful with this theorem! It says if we have a solution then it is a convergent. It does <u>NOT</u> say if we have a convergent then it provides a solution! We do have a partial result in this direction however.

Thm: Let p/q be a convergent of the continued fraction expansion of $\sqrt{D}$.

Then: (p,q) is a solution of one of the equations

$$x^2 - Dy^2 = k$$

for $|k| < 1 + 2\sqrt{D}$.

Proof: Let p/q be a convergent of $\sqrt{D}$. Then we know

$$\left| \frac{p}{q} - \sqrt{D} \right| < \frac{1}{q^2}.$$

i.e.,
$$|p - q\sqrt{D}| < \frac{1}{q}.$$

Thus,

$$\begin{aligned} |p + q\sqrt{D}| &= |p - q\sqrt{D} + 2q\sqrt{D}| \\ &\le |p - q\sqrt{D}| + 2q\sqrt{D} \\ &< \frac{1}{q} + 2q\sqrt{D} \\ &\le q(1 + 2\sqrt{D}) \qquad \left(\frac{1}{q} \le q\right). \end{aligned}$$

Combining the fact that

$$|p - q\sqrt{D}| < \frac{1}{q}$$

and
$$|p + q\sqrt{D}| < q(1 + 2\sqrt{D})$$

gives
$$|p^2 - Dq^2| < 1 + 2\sqrt{D},$$

as desired. ∎

Example: We have that $\sqrt{29} = [5; \overline{2, 1, 1, 2, 10}]$. The first few convergents are:

$$C_1 = \frac{11}{2}$$

$$C_2 = \frac{16}{3}$$

$$C_3 = \frac{27}{5}$$

$$C_4 = \frac{70}{13}$$

$$\vdots$$

$$C_9 = \frac{9801}{1820}$$

$$\vdots$$

$$C_{19} = \frac{192119201}{35675640}$$

Then we have $p_i^2 - 29q_i^2 =$:

1 : 5

2 : -5

3 : 4

$\vdots$

9 : 1

$\vdots$

19 1

Thus, the first two solutions of $x^2 - 29y^2 = 1$ we encounter are $(9801, 1820)$ and $(192119201, 35675640)$.

We have shown if $x^2 - Dy^2 = 1$ has a solution then this solution must be a convergent of $\sqrt{D}$. Of course, we still need to show that there actually are solutions! The result we proved before that the continued fraction of $\sqrt{D}$ is periodic will be essential!

Recall that when proving $\sqrt{D}$ has a periodic continued fraction we defined m_i and b_i so that:

$$m_0 = 0, \quad b_0 = 1$$

$$m_{i+1} = a_i b_i - m_i, \quad b_{i+1} = \frac{D - m_{i+1}^2}{b_i}.$$

with $m_i, b_i \in \mathbb{Z}$, $b_i \neq 0$, $b_{i+1} \mid (D - m_{i+1}^2)$, and $x_k = \dfrac{m_k + \sqrt{D}}{b_k}$ for $k \geq 0$.

<u>Thm:</u> Let p_k/q_k be convergents of the continued fraction expansion of $\sqrt{D}$. Then

$$p_k^2 - Dq_k^2 = (-1)^{k+1} b_{k+1}$$

where $b_{k+1} > 0$ for $t = 0, 1, \ldots$

Proof: Write

$$\sqrt{D} = [a_0; a_1, \ldots, a_n, x_{n+1}].$$

Then we have

$$\sqrt{D} = \frac{x_{n+1} p_n + p_{n-1}}{x_{n+1} q_n + q_{n-1}}.$$

We know we can write $x_{n+1} = \dfrac{m_{n+1} + \sqrt{D}}{b_{n+1}}$, so we substitute this in

$$\sqrt{D} = \frac{\frac{m_{n+1} + \sqrt{D}}{b_{n+1}} p_n + p_{n-1}}{\frac{m_{n+1} + \sqrt{D}}{b_{n+1}} q_n + q_{n-1}}$$

$$= \frac{(m_{n+1} + \sqrt{D}) p_n + b_{n+1} p_{n-1}}{(m_{n+1} + \sqrt{D}) q_n + b_{n+1} q_{n-1}}$$

Thus,

$$\sqrt{D}(m_{n+1} q_n + b_{n+1} q_{n-1} - p_n) = m_{n+1} p_n + b_{n+1} p_{n-1} - D q_n.$$

However, the RHS $\in \mathbb{Q}$ and $\sqrt{D} \notin \mathbb{Q}$ so we must have

$$m_{n+1} q_n + b_{n+1} q_{n-1} = p_n \qquad (1)$$

and

$$m_{n+1} p_n + b_{n+1} p_{n-1} = D q_n \qquad (2)$$

Multiply (1) by $\cancel{q_n}$ p_n and the (2) by $-q_n$ and add them together to get:

$$p_n^2 - Dq_n^2 = b_{k+1}(p_n q_{n-1} - p_{n-1} q_n)$$

$$= (-1)^{n-1} b_{k+1}$$

as desired. It only remains to show $b_{k+1} > 0$. We showed this before for large enough n. We know that

$$C_{2k} < \sqrt{D} < C_{2k+1} \qquad k \geq 0.$$

Thus, $$\frac{p_{2k}}{q_{2k}} < \sqrt{D} \Rightarrow p_{2k} < \sqrt{D}\, q_{2k}$$

$$\Rightarrow \quad p_{2k}^2 - Dq_{2k}^2 < 0.$$

Similarly we have

$$p_{2k+1}^2 - Dq_{2k+1}^2 > 0.$$

Thus,

$$\frac{p_k^2 - Dq_k^2}{p_{k-1}^2 - Dq_{k-1}^2} = -\frac{b_{k+1}}{b_k} \qquad k \geq 1.$$

And this is always negative since the top and bottom of the LHS are opposite signs. Thus, $\frac{b_{k+1}}{b_k} > 0$.

$b_1 = D - a_0^2 > 0 \Rightarrow b_2 > 0$ and inductively that $b_k > 0$ for $k \geq 1$. ∎

At this point we know any solution to $x^2 - Dy^2 = 1$ must be a convergent of $\sqrt{D}$ and that the convergents satisfy

$$p_k^2 - Dq_k^2 = (-1)^{k+1} b_{k+1}.$$

Thus, for (p_k, q_k) to have any chance of being a solution of Pell's equation we must have $b_{k+1} = 1$ (remember $b_{k+1} > 0$!).

<u>Cor:</u> Let n be the length of the period of $\sqrt{D}$. Then $b_k = 1$ iff $n \mid k$.

<u>Proof:</u> Write $\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots, a_n}]$. We know that for any k we have

$$x_{kn+1} = x_1.$$

Thus,

$$\frac{m_{kn+1} + \sqrt{D}}{b_{kn+1}} = \frac{m_1 + \sqrt{D}}{b_1}$$

i.e.

$$\sqrt{D}(b_{kn+1} - b_1) = m_{kn+1} b_1 - m_1 b_{kn+1}.$$

Since $\sqrt{D} \notin \mathbb{Q}$ we have $b_{kn+1} = b_1$

and $m_{kn+1} = m_1$.

Then we have

$$b_1 = D - m_1^2 = D - m_{kn+1}^2 = b_{kn} b_{kn+1} = b_{kn} b_1$$

$$\uparrow$$

$$b_0 = 1$$

Thus, $b_{kn} = 1$. Hence, we have $b_k = 1$ if $n \mid k$.

Now we need the other direction. Let k be a positive integer with $b_k = 1$. Then

$$X_k = m_k + \sqrt{D}$$

$\Rightarrow$

$$\lfloor x_k \rfloor = \lfloor m_k + \sqrt{D} \rfloor$$

$$= m_k + \lfloor \sqrt{D} \rfloor$$

$$= m_k + a_0$$

Now

$$X_{k+1} = \frac{1}{X_k - \lfloor x_k \rfloor}$$

$\Rightarrow$

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

$$= m_k + a_0 + \frac{1}{x_{k+1}}$$

We know that

$$x_1 = \frac{1}{x_0 - a_0} \Rightarrow a_0 + \frac{1}{x_1} = x_0 = \sqrt{D}$$

$$= x_k - m_k$$

$$= a_0 + \frac{1}{x_{k+1}}$$

$\Rightarrow X_1 = X_{k+1}$. But then the set $[a_1, \dots, a_k]$

repeats in the expansion of $\sqrt{D}$. $\Rightarrow n \mid k$. ∎

Thm: Let p_k/q_k be the convergents of the continued fraction expansion of $\sqrt{D}$. Let n be the length of the period.

① If $2 \mid n$, then the positive solutions of $x^2 - Dy^2 = 1$ are given by

$$x = p_{kn-1} \ , \quad y = q_{kn-1} \qquad k = 1, 2, 3, \dots$$

② If $2 \nmid n$, then all positive solutions of $x^2 - Dy^2 = 1$ are given by

$$x = p_{2kn-1} \ , \quad y = q_{2kn-1} \qquad k = 1, 2, \dots$$

Proof: We have basically already shown this. We know all solutions are convergents, and

$$p_k^2 - Dq_k^2 = (-1)^{k+1} b_{k+1}$$

w/ $b_{k+1} = 1$ iff $n \mid k+1$. So if p_k, q_k is a solution then $b_{k+1} = 1$ and $(-1)^{k+1} = 1$. So $k+1$ must be even and $n \mid k+1$. So $\exists\, t$ s.t $k+1 = tn$. and $2 \mid k+1$. If n is even, then

$$k = tn - 1$$

and if n is odd, then t must be even so

$$k = 2sn - 1.$$

∎

Example: Find a solution to $x^2 - 31y^2 = 1$.

$$\sqrt{31} = [5; \overline{1, 1, 3, 5, 3, 1, 1, 10}]$$

So $n = 8$ in this case. Since this is even, the theorem says p_{8k-1}, q_{8k-1} are solutions for $k = 1, 2, \ldots$

Thus, we need p_7, q_7, for our first solution.

$$C_7 = \frac{1520}{273}, \text{ so } (1520, 273) \text{ is a solution}$$

$$C_{15} = \frac{4620799}{829920} \text{ so } (4620799, 829920) \text{ is}$$

another solution.

The problem with this method is that one has to keep computing convergents to get solutions. To get C_{15}, we have to compute $C_8, \ldots, C_{14}$ first. This is not an ideal way to obtain solutions. We would like to be able to use our first solution to give us all other solutions.

Def: A solution (x_0, y_0) of Pell's equation is a fundamental solution if given another positive solution (x_1, y_1), then $x_0 < x_1$, $y_0 < y_1$.

Thm: Let (x_ϕ, y_ϕ) be the fundamental solution of $x^2 - dy^2 = 1$. Then every pair (x_n, y_n) defined by

$$x_n + y_n\sqrt{D} = (x_\phi + y_\phi\sqrt{D})^n \qquad n = 1, 2, \ldots$$

is also a solution.

Proof: Observe that we have

$$\begin{aligned} x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n &= (x_1 + y_1\sqrt{D})(x_{1} + y_1\sqrt{D})^{n-1} \\ &= (x_1 + y_1\sqrt{D})(x_{n-1} + y_{n-1}\sqrt{D}) \\ &= (x_1 x_{n-1} + y_1 y_{n-1} D) + (x_{n-1} y_1 + x_1 y_{n-1})\sqrt{D}. \end{aligned}$$

Thus, $x_n = x_1 x_{n-1} + y_1 y_{n-1}D$, $y_n = x_{n-1} y_1 + x_1 y_{n-1}$.

Claim: $(x_1 - y_1\sqrt{D})^n = x_n - y_n\sqrt{D}$.

Pf. Use induction on n. $n=1$ trivial. Assume the result for $n-1$. Then

$$\begin{aligned} (x_1 - y_1\sqrt{D})^n &= (x_1 - y_1\sqrt{D})(x_1 - y_1\sqrt{D})^{n-1} \\ &= (x_1 - y_1\sqrt{D})(x_{n-1} - y_{n-1}\sqrt{D}) \quad \text{(Ind. hyp)} \\ &= (x_1 x_{n-1} + y_1 y_{n-1} D) - (x_{n-1} y_1 + x_1 y_{n-1})\sqrt{D} \\ &= x_n - y_n\sqrt{D}, \text{ as desired.} \end{aligned}$$

Since x_1, y_1 are both positive, it is clear x_n and y_n must be positive. We have:

$$x_n^2 - Dy_n^2 = (x_n + y_n\sqrt{D})(x_n - y_n\sqrt{D})$$

$$= (x_1 + y_1\sqrt{D})^n (x_1 - y_1\sqrt{D})^n$$

$$= (x_1^2 - Dy_1^2)^n$$

$$= 1^n = 1$$

$\blacksquare$

<u>Example:</u> Recall that we found $(1520, 273)$ as a fundamental solution to $x^2 - 31y^2 = 1$. Thus,

$$(1520 + 273\sqrt{31})^n$$

~~are rules solutions~~ give solutions x_n, y_n as well:

$n = 2$ $\quad x_2 = 4620799 \quad y_2 = 829920$

3 $\quad x_3 = 14047227440 \quad y_3 = 2522956527$

~~4~~

Finally, we showed this method gives all solutions.

Thm: Let (x_1, y_1) be a fundamental solution of $X^2 - Dy^2 = 1$. Then every positive solution of the equation ~~g~~ is given by x_n, y_n where x_n, y_n are obtained by

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n. \qquad n = 1, 2, \ldots$$

Proof: Suppose $\exists$ u, v that is a solution but not one of the x_n, y_n.

We know that $x_1 + y_1\sqrt{D} > 1$, so $\lim_{n\to\infty} (x_1 + y_1\sqrt{D})^n = \infty$

so we can find n s.t

$$(x_1 + y_1\sqrt{D})^n < u + v\sqrt{D} < (x_1 + y_1\sqrt{D})^{n+1}.$$

i.e.,

$$x_n + y_n\sqrt{D} < u + v\sqrt{D} < (x_1 + y_1\sqrt{D})(x_n + y_n\sqrt{D}).$$

Multiply this by $x_n - y_n\sqrt{D}$ (this is positive!) we get

$$1 = x_n^2 - y_n^2 D < (x_n - y_n\sqrt{D})(u + v\sqrt{D}) < x_1 + y_1\sqrt{D}.$$

Set $\quad r = x_n u - D y_n v$, $s = x_n v - y_n u$ so that

$$(x_n - y_n\sqrt{D})(u + v\sqrt{D}) = r + s\sqrt{D}. \quad (*)$$

Then we have

$$r^2 - Ds^2 = (x_n u - D y_n v)^2 - D(x_n v - y_n u)^2$$

$$= (x_n u)^2 - \cancel{2x_n y_n Duv} + D y_n^2 v^2$$
$$- D(x_n v)^2 + \cancel{2x_n uv y_n D} + D(y_n u)^2$$

$$= (x_n^2 - Dy_n^2)(u^2 - Dv^2) = 1 \cdot 1 = 1.$$

Thus, (r,s) is a solution to $x^2 - Dy^2 = 1$ w/

$$1 < r + s\sqrt{D} < x_1 + y_1\sqrt{D}.$$

prod of two things bigger than 1. by (*)

If we can show r and s are both positive we will have a contradiction.

Since $1 < r + s\sqrt{D}$ and $(r + s\sqrt{D})(r - s\sqrt{D}) = r^2 - Ds^2 = 1$,

we must have $r - s\sqrt{D} > 0$ and $r - s\sqrt{D} < 1$.

Thus,

$$2r = (r + s\sqrt{D}) + (r - s\sqrt{D}) > 1 + 0 > 0$$

$$2s\sqrt{D} = (r + s\sqrt{D}) - (r - s\sqrt{D}) > 1 - 1 = 0.$$

$\Rightarrow r, s > 0$. #. Thus it must be that all solutions arise from $(x_1 + y_1\sqrt{D})^n$.

$\blacksquare$

Finally, consider the equation $x^2 - Dy^2 = -1$. We saw before that

$$p_k^2 - Dq_k^2 = (-1)^{k+1} b_{k+1}$$

and $b_{k+1} = 1$ iff $n \mid k+1$ where $n =$ period of $\sqrt{D}$. This allows us to conclude that if we want solutions to $x^2 - Dy^2 = -1$, we should look at those k where $k+1$ is odd and $n \mid k+1$.

So we can write $K+1 = nt$ for some t.

~~As~~

Suppose n is even. Then $K+1$ = even. This can't give us any solutions!

Suppose n is odd. Then $K+1$ = odd if t is also odd! So our solutions

look like

n odd

$$p_{n(2t+1)-1} \quad q_{n(2t+1)-1} \quad \text{for } t = 0, 1, 2, \ldots$$