Pell's Equation I

$$x^2 - Dy^2 = 1 \qquad D \text{ square free .}$$

- Has a VERY long history: Archimedes problem, ancient Chinese and Indian mathematics.
- Approximation of real numbers by rational numbers; with great efficiency, error *very small*.
- Can be used to solve other quadratic equations, e.g. $x^2 - Dy^2 = N$.
- Has a very elegant and powerful solution introducing Continued Fractions.

Pell's Equation II

Triangular Numbers. Want $1 + 2 + \cdots + (m-1) + m = n^2$, a perfect square. This gives

$$\frac{m(m+1)}{2} = n^2 \iff m^2 + m = 2n^2 \iff (2m+1)^2 - 2n^2 = 1.$$

This is the equation $x^2 - 2y^2 = 1$. The first solution is $(3, 2)$. We can find another one by squaring $(3 + 2\sqrt{2})^2 = 17 + 12\sqrt{2}$.

Basic Results I

Theorem (1)

The set of solutions forms a group.

Proof.

$$x^2 - Dy^2 \iff (x + \sqrt{D}y)(x - \sqrt{D}y) = x^2 - Dy^2.$$

Call $x^2 - Dy^2$ the norm $N(a)$ of $x + \sqrt{D}y$. So (x, y) is a solution $\iff N(a) = 1$.

The theorem comes down to observing that $N(a_1 a_2) = N(a_1)N(a_2)$, and $a^{-1} = x - \sqrt{D}y$. □

NOTE: If (x, y) is a solution, so are $(\pm x, \pm y)$. There is a trivial solution, $(1, 0)$.

Harder Results I

Theorem (2)

There are infinitely many solutions; there are nontrivial ones with $y \neq 0$. They can be used to approximate $\sqrt{D}$:

$$x^2 - Dy^2 = 1 \Longleftrightarrow \left|\frac{x}{y} - \sqrt{D}\right| < \frac{1}{y^2}.$$

The fact that there are infinitely many solutions is left for later. The easy part is the inequality. We may as well assume $x, y > 0$.

$$x^2 - Dy^2 = 1 \Longrightarrow |x - \sqrt{D}y| = \frac{1}{x + \sqrt{D}y} \leq \frac{1}{y},$$

$$\frac{1}{x + \sqrt{D}y} \leq \frac{1}{y} \Longleftrightarrow |\frac{x}{y} - \sqrt{D}| \leq \frac{1}{y^2}$$

Harder Results II

Theorem (3)

The group of units is cyclic.

Proof. Consider just the solutions with $x, y > 0$. A solution $x + \sqrt{D}y$ will satisfy $x > 1$. Order the solutions by the size of x. There has to be a solution $x_0 + \sqrt{D}y_0$ with the smallest $x_0 > 1$. The $y_0 > 0$ is unique: $x_0^2 - Dy_0^2 = x_0^2 - Dy_1^2 \Longrightarrow y_0^2 = y_1^2$, so $y_1 = \pm y_0$.
Let $x_1 + \sqrt{D}y_1$ be another solution, with $x_1 > x_0$. Divide

$$\frac{x_1 + \sqrt{D}y_1}{x_0 + \sqrt{D}y_0} = (x_1 + \sqrt{D}y_1)(x_0 - \sqrt{D}y_0) =$$

$$= (x_0x_1 - Dy_0y_1) + (x_0y_1 - x_1y_0)\sqrt{D} = x_2 + y_2\sqrt{D}$$

We claim

Harder Results III

1. $0 < x_2 = x_1x_0 - Dy_0y_1 \iff Dy_0y_1 < x_0x_1$,
2. $0 < x_1 = x_0y_1 - x_1y_0 \iff \frac{y_0}{x_0} < \frac{y_1}{x_1}$,
3. $x_2 = x_0x_1 - Dy_0y_1 < x_1$.

For (1), $Dy_i^2 = x_i^2 - 1 < x_i^2 \iff \sqrt{D}y_i < x_i$.

For (2) use a picture; the slope of the line through (x_2, y_2) is less than the slope of the line through (x_1, y_1).

For (3), rewrite $x_1 = x_0x_2 + Dy_0y_2 > x_2$ because we know $x_2,\ y_2 > 0$ from (1) and (2).

Continued Fractions I

I tried to keep to the notation in Davenport; it is not always possible. The material is somewhat different.

Let α be any real number.

1. Write $\alpha = (a_0 = [\alpha]) + \beta_1$; $0 \leq \beta_1 < 1$.
2. If $\beta_1 = 0$, STOP. Otherwise, write

$$\alpha = a_0 + \frac{1}{\alpha_1} = a_1 + \frac{1}{a_1 = [\alpha_1] + \beta_2}.$$

Continue indefinitely, or until the first $\beta_r = 0$.

The process produces a sequence $\{a_0, \ldots, a_r, \ldots\}$. It stops if and only if α is rational. In general it associates a sequence $\{a_0, \ldots, a_r, \ldots\}$ with a_i positive integers for $i > 0$ to each real number.

The CF expansion is UNIQUE. Exercise, see Davenport.

Continued Fractions II

Example (Euclidean Algorithm)

$$\frac{8}{3} = 2 + \frac{2}{3} = 2 + \frac{1}{\frac{3}{2}} = 2 + \frac{1}{1 + \frac{1}{2}}.$$

Definition

We denote by $[a_0, \ldots, a_r]$ the value of the continued fraction. In the example it is $[2, 1, 2]$.

In Davenport, the symbol $[a_0, \ldots, a_r]$ denotes the numerator of the CF. He then writes $\frac{[a_0, \quad a_r]}{[a_1, \ldots, a_r]}$ for the value of the CF.
I will write $\langle a_0, \ldots a_r \rangle$ for the value of the CF.

Continued Fractions III

Define $p_{-1} = 1, q_{-1} = 0$, and p_r, q_r by

$$\begin{pmatrix} p_r & q_r \\ p_{r-1} & q_{r-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdots\cdots \begin{pmatrix} a_r & 1 \\ 1 & 0 \end{pmatrix}$$

Equivalently, by taking transposes, and noting that the matrices on the right are symmetric,

$$\begin{pmatrix} p_r & p_{r-1} \\ q_r & q_{r-1} \end{pmatrix} = \begin{pmatrix} a_r & 1 \\ 1 & 0 \end{pmatrix} \cdots\cdots \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $p_r = p_{r-1}a_r + p_{r-2},\ q_r = q_{r-1}a_r + q_{r-2}$.This follows by a simple induction from the first formula. The first equation can be written

$$\begin{pmatrix} p_r & p_{r-1} \\ q_r & q_{r-1} \end{pmatrix} = \begin{pmatrix} p_{r-1} & p_{r-2} \\ q_{r-1} & q_{r-2} \end{pmatrix} \cdot \begin{pmatrix} a_r & 1 \\ 1 & 0 \end{pmatrix}.$$

Continued Fractions IV

Write $\alpha = \langle a_0, \ldots, a_n, \alpha_{n+1} \rangle$ for the value of the continued fraction formed of these numbers; In other words, $a_0 + \frac{1}{a_1 + \cdots + \frac{1}{\alpha_{n+1}}}$. This could be an arbitrary sequence with a_i integers ≥ 1 for $i > 0$. Or it could be obtained by carrying out the CF algorithm; $a_0 = [\alpha]$, $\alpha = a_0 + \frac{1}{\alpha_1} \ldots$.

Theorem

Let p_k, q_k be the numbers obtained by the algorithm above. Then $\alpha = \dfrac{p_{r-1}\alpha_r + p_{r-2}}{q_r\alpha_r + q_{r-1}}, \qquad \alpha_r = \dfrac{p_r\alpha + p_{r-1}}{q_r\alpha + q_{r-1}}.$

The values $\frac{p_k}{q_k}$ are called the convergents of the CF.

Proof: Do an induction. The hypothesis is that $\langle a_0, \ldots, a_n \rangle$ and $\langle a_0, \ldots a_{n-1} \rangle$ are given by the ratios of the entries in the

Continued Fractions V

columns of the product

$$\begin{pmatrix} p_r & p_{r-1} \\ q_r & q_{r-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix},$$

and the similar product going up to $n-1$ only.

$\langle a_0 \rangle = a_0$ can be encoded as $\begin{pmatrix} p_0 & q_0 \\ p_{-1} & q_{-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}$. At the next step, $\langle a_0, a_1 \rangle = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{p_0 a_1 + p_{-1}}{q_0 a_1 + q_{-1}} = \frac{p_1}{q_1}$.

$$\begin{pmatrix} p_1 & p_0 \\ q_1 & q_0 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_0 a_1 + 1 & a_0 \\ a_1 & 1 \end{pmatrix}$$

The induction step follows from the formula $\langle a_0, \ldots, a_r \rangle = a_0 + \frac{1}{\langle a_1, \ldots, a_r \rangle}$ rewritten in matrix notation

$$\begin{pmatrix} p_r & q_r \\ p_{r-1} & q_{r-1} \end{pmatrix} = \begin{pmatrix} p'_{r-1} & q'_{r-1} \\ p'_{r-2} & q'_{r-2} \end{pmatrix} \cdot \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Continued Fractions VI

The $'$ refer to the p, q for $\langle a_1, \ldots, a_n \rangle$. You can compare $\frac{a_0 p'_{r-1} + q'_{r-1}}{p'_{r-1}} = a_0 + \frac{1}{p'_{r-1}/q'_{r-1}}$.
Use the induction hypothesis for $\langle a_1, \ldots, a_n \rangle$, and take the transpose. □

Davenport calls the expressions for the numerator and denominator $[a_0, \ldots, a_k]$ and $[a_1, \ldots, a_k]$.

Euler's Rule: $[a_0, \ldots, a_n]$ is defined as:
First take the product of all the terms. Then take every product that can be obtained by omitting any pair of consecutive terms. Then take every product that can be obtained by omitting any two separate pairs of consecutive terms, and so on. If there are no elements left, count it as a 1.
The conclusion is that

Continued Fractions VII

$$[a_0, \ldots, a_n] = [a_n, \ldots, a_0] \qquad \langle a_0, \ldots, a_n \rangle = \frac{p_n}{q_n} = \frac{[a_0, \ldots, a_n]}{[a_1, \ldots, a_n]}.$$

Continued Fractions VIII

Corollary

$p_r q_{r-1} - q_r p_{r-1} = (-1)^{r-1}$. *In particular,* $gcd(p_r, q_r) = 1$.

Proof.

The determinants of any of the matrices are (-1). □

Example

$\alpha = \frac{16}{6}$ has the same continued fraction as $\frac{8}{3}$, namely $[2, 1, 2]$. The matrix that comes out is $\begin{pmatrix} 8 & 3 \\ 2 & 1 \end{pmatrix}$ If you want the *gcd*, you have to write $\frac{16}{3} = \frac{8}{3}$, and adjust.

Periodic Continued Fractions I

Example. The continued fraction of $\sqrt{6}+2$ is

$$\begin{aligned}\sqrt{6}+2 =& 4+\frac{1}{\frac{1}{\sqrt{6}-2}} = 4+\frac{1}{\frac{\sqrt{6}+2}{2}} = 4+\frac{1}{2+\frac{\sqrt{6}-2}{2}} = \\ =& 4+\frac{1}{2+\frac{1}{\sqrt{6}+2}}.\end{aligned}$$

This implies that the CF of $\sqrt{6}+2$ is $\langle\overline{4,2}\rangle$ where the meaning of the underlined is that it repeats periodically.
We can write this as $\sqrt{6}+2=\langle 4,2,\sqrt{6}+2\rangle$. We find

$$\begin{pmatrix} p_3 & p_2 \\ q_3 & q_2 \end{pmatrix} = \begin{pmatrix} \sqrt{6}+2 & 1 \\ 1 & 0 \end{pmatrix}\cdot\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\cdot\begin{pmatrix} 4 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 9\sqrt{6}+22 & 9 \\ 2\sqrt{6}+5 & 2 \end{pmatrix}.$$

Periodic Continued Fractions II

Then $\sqrt{6} = \frac{9\sqrt{6}+22}{2\sqrt{6}+5} - 2 = \frac{5\sqrt{6}+12}{2\sqrt{6}+5}$. The matrix $\begin{pmatrix} 5 & 6\cdot 2 \\ 2 & 5 \end{pmatrix}$ has determinant $1 = 5\cdot 5 - 6\cdot 2\cdot 2 = 1$, a solution to the Pell equation $x^2 - 6y^2 = 1$. □

The key fact is that $\sqrt{D} = \langle a_0, \overline{a_1, \ldots, a_r, a_{r+1} = 2a_0} \rangle$, and $\sqrt{D} = \langle a_0, \ldots, a_r, \alpha_{r+1}, \sqrt{D} + a_0 \rangle$. Note that $a_0 = p_0$.
The recursion relation is

$$\begin{pmatrix} p_{r+1} & q_{r+1} \\ p_r & q_r \end{pmatrix} = \begin{pmatrix} \sqrt{D} + a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} p_r & q_r \\ p_{r-1} & q_{r-1} \end{pmatrix}$$

We get an identity $\sqrt{D} = \frac{p_r(\sqrt{D}+a_0)+p_{r-1}}{q_r(\sqrt{D}+a_0)+q_{r-1}}$ which yields

$$q_r\sqrt{D}(\sqrt{D} + a_0) + \sqrt{D}q_{r-1} = (\sqrt{D} + a_0)p_r + p_{r-1}.$$

Periodic Continued Fractions III

This implies

$$\begin{cases} Dq_r = a_0 p_r + p_{r-1}, \\ a_0 q_r + q_{r-1} = p_r. \end{cases}$$

Plug into the relationship $p_r q_{r-1} - p_{r-1} q_r = (-1)^{r-1}$:

$$p_r(p_r - a_0 q_r) - q_r(Dq_r - a_0 p_r) = (-1)^{r-1}.$$

The conclusion is $p_r^2 - Dq_r^2 = (-1)^{r-1}$. So depending on whether r is even of odd, we get a solution to $x^2 - Dy^2 = 1$ or $x^2 - Dy^2 = -1$.

In the second case, just go to the next period.

The argument can be carried out with either of $\sqrt{D} \pm a_0$.