Solving $x^2 - Dy^2 = N$ in integers, where $D > 0$ is not a perfect square.

Keith Matthews

In 1769, Lagrange showed how to solve this equation if $|N| < \sqrt{D}$ and gave a recursive method when $|N| > \sqrt{D}$. He gave another method which has not appeared in modern number theory books until R.A. Mollin rediscovered it and included it in his 1998 textbook *Fundamental Number Theory with Applications*.

Mollin's approach is via ideal theory in quadratic fields.

The main difficulty is showing that all solutions arise from continued fractions of certain quadratic irrationalities. Lagrange's proof is not valid. We supply a simple proof.

Pell's equation

The special case $N = 1$ is known as *Pell's equation*. If (x_0, y_0) denotes the fundamental solution of $x^2 - Dy^2 = 1$, ie, the solution with least positive x and y, then the general solution (x, y) is given by

$$x + y\sqrt{D} = \pm(x_0 + y_0\sqrt{D})^n, n \in \mathbb{Z}.$$

We can calculate (x_0, y_0) by expanding $\sqrt{D}$ as a periodic continued fraction:

$$\sqrt{D} = [a_0, \overline{a_1, \ldots, a_l}].$$

Then

$$x_0/y_0 = \begin{cases} \frac{A_{l-1}}{B_{l-1}}, & \text{if } l \text{ is even} \\ \frac{A_{2l-1}}{B_{2l-1}}, & \text{if } l \text{ is odd,} \end{cases}$$

where A_n/B_n denotes the n–th convergent to $\sqrt{D}$.

Equivalence classes of solutions of $x^2 - Dy^2 = N$.

It suffices to consider only solutions (x, y) of $x^2 - Dy^2 = N$ with $\gcd(x, y) = 1$, ie. *primitive* solutions. For $d = \gcd(x, y)$, $x = dX$, $y = dY$ give $X^2 - DY^2 = N/d^2$ and $\gcd(X, Y) = 1$.

The identity

$$(x^2 - Dy^2)(u^2 - vy^2) = (xu + yvD)^2 - D(uy + vx)^2$$

shows that solutions (x, y) of $x^2 - Dy^2 = N$ and (u, v) of $u^2 - Dv^2 = 1$ produce another primitive solution $(x', y') = (xu + yvD, uy + vx)$ of $x'^2 - Dy'^2 = N$, or equivalently

$$x' + y'\sqrt{D} = (x + y\sqrt{D})(u + v\sqrt{D}), \qquad (1)$$

where $u^2 - Dv^2 = 1$.

Equation (1) defines an equivalence relation on the set of all primitive solutions of $x^2 - Dy^2 = N$.

Finiteness of the number of equivalence classes.

There are only finitely many equivalence classes: each class contains a solution (x_1, y_1) with least positive y_1 (a *fundamental* solution), where

(i) $0 < y_1 \leq y_0\sqrt{\frac{|N|}{2(x_0+\epsilon)}}$

(ii) $0 < |x_1| \leq \sqrt{\frac{(x_0+\epsilon)|N|}{2}}$,

where $\epsilon = N/|N|$.

Associating a congruence class mod $|N|$ to each equivalence class.

If (x, y) is a solution for a class C, then $(-x, y)$ is a solution for the *conjugate* class C^*.

Primitive solutions (x, y) and (x', y') are equivalent if and only if

$$xx' - yy'D \equiv 0 \pmod{|N|} \text{ and } yx' - xy' \equiv 0 \pmod{|N|}.$$

If $x^2 - Dy^2 = N$ with $\gcd(x, y) = 1$ and P is defined by $x \equiv yP \pmod{|N|}$, then $P^2 \equiv D \pmod{|N|}$.

If $x'^2 - Dy'^2 = N$ with $\gcd(x', y') = 1$ and $x' \equiv y'P' \pmod{|N|}$, then (x, y) and (x', y') are equivalent if and only if $P \equiv P' \pmod{|N|}$.

It can happen that $C^* = C$, in which case C is called an *ambiguous* class.

A class is ambiguous if and only if $P \equiv 0$ or $|N|/2 \pmod{|N|}$.

Continued fractions of quadratic irrationalities.

We need to introduce the n–th convergent x_n:

If $\omega = [a_0, a_1, \ldots]$, then

$$x_n = [a_n, a_{n+1}, \ldots,], n \geq 0,$$

is called the n–th convergent of ω.

If $\omega = \frac{P_0+\sqrt{D}}{Q_0}$, where $Q_0|(P_0^2 - D)$, then

$$x_n = \frac{P_n + \sqrt{D}}{Q_n}$$

and there is a simple algorithm for calculating a_n, P_n and Q_n:

$$a_n = \left\lfloor \frac{P_n+\sqrt{D}}{Q_n} \right\rfloor,$$

$$P_{n+1} = a_n Q_n - P_n,$$

$$Q_{n+1} = \frac{D-P_{n+1}^2}{Q_n}.$$

We also note the following important identity

$$G_{n-1}^2 - DB_{n-1}^2 = (-1)^n Q_0 Q_n,$$

where $G_{n-1} = Q_0 A_{n-1} - P_0 B_{n-1}$.

The classical case $|N| < \sqrt{D}$.

If $x^2 - Dy^2 = N$, with $\gcd(x, y) = 1$, $x > 0, y > 0$, then it is easy to show that

$$\left|\frac{x}{y} - \sqrt{D}\right| < \frac{1}{2y^2} \text{ if } \frac{x}{y} > \sqrt{D},$$

$$\left|\frac{y}{x} - \frac{1}{\sqrt{D}}\right| < \frac{1}{2x^2} \text{ if } \frac{x}{y} < \sqrt{D}$$

and consequently by a theorem of Lagrange, x/y is a convergent A_{n-1}/B_{n-1} of $\sqrt{D}$. Then

$$N = x^2 - Dy^2 = A_{n-1}^2 - DB_{n-1}^2 = (-1)^n Q_n.$$

If $\sqrt{D} = [a_0, \overline{a_1, \ldots, a_l}]$, it turns out that we need only check the range $n \leq l$ to see which Q_n, if any, satisfy $Q_n = (-1)^n N$. (If l is odd, we have $Q_{\frac{l-1}{2}-r} = Q_{\frac{l+1}{2}+r}$ for $r = 0, \ldots, \frac{l-3}{2}$ and the subscripts have opposite parity.)

The corresponding (A_{n-1}, B_{n-1}) will be fundamental solutions.

An example: $x^2 - 241y^2 = \pm 15$.

$\sqrt{241} = [15, \overline{1, 1, 9, 1, 5, 3, 3, 1, 1, 3, 3, 5, 1, 9, 1, 1, 30}]$

period length $= 17$

```
P[0]=0,   Q[0]=1
P[1]=15,  Q[1]=16
P[2]=1,   Q[2]=15    << A[1]^2-241*B[1]^2=15
P[3]=14,  Q[3]=3
P[4]=13,  Q[4]=24
P[5]=11,  Q[5]=5,
P[6]=14,  Q[6]=9,
P[7]=13,  Q[7]=8,
P[8]=11,  Q[8]=15,   << A[7]^2-241*B[7]^2 =15
P[9]=4,   Q[9]=15,   << A[8]^2-241*B[8]^2 =-15
P[10]=11, Q[10]=8,
P[11]=13, Q[11]=9,
P[12]=14, Q[12]=5,
P[13]=11, Q[13]=24,
P[14]=13, Q[14]=3,
P[15]=14, Q[15]=15,  << A[14]^2-241*B[14]^2=-15
P[16]=1,  Q[16]=16,
P[17]=15, Q[17]=1,   << A[16]^2-241*B[16]^2=-1
P[18]=15, Q[18]=16,
```

A[0]/B[0]=15/1
A[1]/B[1]=16/1
A[2]/B[2]=31/2
A[3]/B[3]=295/19
A[4]/B[4]=326/21
A[5]/B[5]=1925/124
A[6]/B[6]=6101/393
A[7]/B[7]=20228/1303
A[8]/B[8]=26329/1696
A[9]/B[9]=46557/2999
A[10]/B[10]=166000/10693
A[11]/B[11]=544557/35078
A[12]/B[12]=2888785/186083
A[13]/B[13]=3433342/221161
A[14]/B[14]=33788863/2176532
A[15]/B[15]=37222205/2397693
A[16]/B[16]=71011068/4574225
A[17]/B[17]=2167554245/139624443

$(A_1, B_1) = (16, 1)$ and $(A_7, B_7) = (20228, 1303)$ are fundamental solutions of $x^2 - 241y^2 = 15$;

$(A_8, B_8) = (26329, 1696)$ and
$(A_{14}, B_{14}) = (33788863, 2176532)$ are fundamental solutions of $x^2 - 241y^2 = -15$;

$(A_{16}, B_{16}) = (71011068, 4574225)$ is the smallest solution of $x^2 - 241y^2 = -1$

$(A_{23}, B_{23}) = (1008514355700 1249, 64964120504 4600)$ is the smallest solution of $x^2 - 241y^2 = 1$

Hence if $\eta_0 = A_{23} + B_{23}\sqrt{241}$, the general solution of $x^2 - 241y^2 = 15$ is given by

$$x + y\sqrt{241} = \begin{cases} \pm(\pm 16 + \sqrt{241})\eta_0^n \\ \pm(\pm 20228 + 1303\sqrt{241})\eta_0^n. \end{cases}$$

We remark that the fundamental solution $(295, 19)$ of $x^2 - 241y^2 = 24$ is produced above, but that $(378557, 24385)$ is missing.

$\sqrt{103} = [10, \overline{6, 1, 2, 1, 1, 9, 1, 1, 2, 1, 6, 20}]$.
period length = 12

```
P[0]=0, Q[0]=1
P[1]=10, Q[1]=3
P[2]=8, Q[2]=13
P[3]=5, Q[3]=6
P[4]=7, Q[4]=9
P[5]=2, Q[5]=11
P[6]=9, Q[6]=2
P[7]=9, Q[7]=11
P[8]=2, Q[8]=9
P[9]=7, Q[9]=6
P[10]=5, Q[10]=13
P[11]=8, Q[11]=3
P[12]=10, Q[12]=1
P[13]=10, Q[13]=3
convergents:
A[0]/B[0]=10/1
A[1]/B[1]=61/6
A[2]/B[2]=71/7
A[3]/B[3]=203/20
A[4]/B[4]=274/27
A[5]/B[5]=477/47
A[6]/B[6]=4567/450
A[7]/B[7]=5044/497
A[8]/B[8]=9611/947
A[9]/B[9]=24266/2391
A[10]/B[10]=33877/3338
A[11]/B[11]=227528/22419
A[12]/B[12]=4584437/451718
```

The case of general $|N|$: Necessary conditions for solubility of $x^2 - Dy^2 = N$.

Suppose $x^2 - Dy^2 = N$, with $\gcd(x, y) = 1$, $x > 0, y > 0$.

Let $x \equiv yP \pmod{|N|}$, where $-|N|/2 < P \leq |N|/2$. Then

(i) $P^2 \equiv D \pmod{|N|}$.

If $x = Py + |N|X$, then Lagrange substituted $x = Py + |N|X$ in the equation $x^2 - Dy^2 = N$ to get

$$|N|X^2 + 2PXy + \frac{(P^2-D)}{|N|}y^2 = \frac{N}{|N|}. \qquad (2)$$

He then appealed to a result on a general homogeneous equation $f(x, y) = 1$ and deduced that X/y is a convergent to a root of equation (3).

$$|N|\lambda^2 + 2P\lambda + \frac{(P^2-D)}{|N|} = 0. \qquad (3)$$

In fact

(ii) X/y is a convergent A_{n-1}/B_{n-1} to $\omega = \frac{-P+\sqrt{D}}{|N|}$ and $Q_n = (-1)^n \frac{N}{|N|}$.

(iii) If $\omega = [a_0, \ldots, a_t, \overline{a_{t+1}, \ldots, a_{t+l}}]$,

then $Q_n = 1$ for exactly one n in $t+1 \leq n \leq t+l$, where if l is even, then $(-1)^n N/|N| = 1$.

We prove part (ii) by using the following extension of Theorem 172 in Hardy and Wright's book:

Lemma. If $\omega = \frac{U\zeta+R}{V\zeta+S}$, where $\zeta > 1$ and U, V, R, S are integers such that $V > 0, S > 0$ and $US - VR = \pm 1$, then U/V is a convergent to ω.

We apply the lemma to the matrix

$$\begin{bmatrix} U & R \\ V & S \end{bmatrix} = \begin{bmatrix} X & \frac{-Px+Dy}{|N|} \\ y & x \end{bmatrix}.$$

The matrix has integer entries. For

$x \equiv yP \pmod{|N|}$ and hence

$$-Px + Dy \equiv -P^2y + Dy \pmod{|N|}$$

$$\equiv (D - P^2)y \equiv 0 \pmod{|N|}.$$

The matrix $\begin{bmatrix} X & \frac{-Px+Dy}{|N|} \\ y & x \end{bmatrix}$ has determinant

$$\begin{aligned} \Delta &= Xx - \frac{y(-Px+Dy)}{|N|} \\ &= \frac{(x-Py)x - y(-Px+Dy)}{|N|} \\ &= \frac{x^2 - Dy^2}{|N|} = \frac{N}{|N|}. \end{aligned}$$

Also if $\zeta = \sqrt{D}$ and $\omega = (-P + \sqrt{D})/|N|$, it is easy to verify that $\omega = \frac{U\zeta+R}{V\zeta+S}$.

The lemma now implies that $U/V = X/y$ is a convergent A_{n-1}/B_{n-1} to ω.

Also if $G_{n-1} = |N|A_{n-1} + PB_{n-1}$, then

$N = x^2 - Dy^2 = G_{n-1}^2 - DB_{n-1}^2 = (-1)^n |N| Q_n$.

Hence $Q_n = (-1)^n N/|N|$.

Sufficiency.

Suppose $P^2 \equiv D \pmod{|N|}$ and let

$\omega = \frac{-P+\sqrt{D}}{|N|} = [a_0, \ldots, a_t, \overline{a_{t+1}, \ldots, a_{t+l}}]$.

Suppose we have $Q_n = 1$ for some n in $t+1 \leq n \leq t+l$. Then

$G_{n-1}^2 - DB_{n-1}^2 = (-1)^n |N|$,

where $G_{n-1} = |N|A_{n-1} + PB_{n-1}$.

(i) If l is even and $(-1)^n N/|N| = 1$, the equation $x^2 - Dy^2 = N$ has a solution (G_{n-1}, B_{n-1}),

(ii) If l is odd, then (G_{n-1}, B_{n-1}) is a solution of $x^2 - Dy^2 = (-1)^n |N|$, while (G_{n+l-1}, B_{n+l-1}) will be a solution of $x^2 - Dy^2 = (-1)^{n+1} |N|$;

(iii) one of the (G_{k-1}, B_{k-1}) with least B_{k-1} satisfying $G_{k-1}^2 - DB_{k-1}^2 = N$ and arising from the continued fraction expansions of $(-P + \sqrt{D})/|N|$ and $(P + \sqrt{D})/|N|$, will be a fundamental solution of $x^2 - Dy^2 = N$.

An example: $x^2 - 221y^2 = \pm 217$.

We find the solutions of $P^2 \equiv 221 \pmod{217}$ are ± 2 and ± 33.

(a) $\frac{2+\sqrt{221}}{217} = [0, 12, \overline{1, 6, 2, 6, 1, 28}]$.

i	0	1	2	3	4	5	6	7
P_i	2	-2	14	11	13	13	11	14
Q_i	217	1	25	4	13	4	25	1
A_i	0	1	1	7	15	97	112	3233
B_i	1	12	13	90	193	1248	1441	41596

The period length is 6 and $Q_1 = Q_7 = 1$ and $(-1)^1 = (-1)^7 = -1$.

Hence $(G_0, B_0) = (-2, 1)$ is a solution of $x^2 - 221y^2 = -217$ and this is clearly a fundamental one, so there is no need to examine the continued fraction expansion of $\frac{-2+\sqrt{221}}{217}$.

(b) $\frac{33+\sqrt{221}}{217} = [0, 4, 1, 1, \overline{6, 1, 28, 1, 6, 2}]$.

i	0	1	2	3	4	5	6	7	8	9
P_i	33	-33	17	0	13	11	14	14	11	13
Q_i	217	-4	17	13	4	25	1	25	4	13
A_i	0	1	1	2	13	15	433	448	3121	6690
B_i	1	4	5	9	59	68	1963	2031	14149	30329

We observe that $Q_6 = 1$. The period length is even and $(-1)^6 = 1$. Hence $(G_5, B_5) = (1011, 68)$ is a solution of $x^2 - 221y^2 = 217$.

c) $\frac{-33+\sqrt{221}}{217} = [-1, 1, 10, \overline{1, 28, 1, 6, 2, 6}]$.

i	0	1	2	3	4	5	6	7	8
P_i	-33	-184	29	11	14	14	11	13	13
Q_i	217	-155	4	25	1	25	4	13	4
A_i	-1	0	-1	-1	-29	-30	-209	-448	-2897
B_i	1	1	11	12	347	359	2501	5361	34667

We observe that $Q_4 = 1$. The period length is even and $(-1)^4 = 1$. Hence $(G_3, B_3) = (179, 12)$ is a solution of $x^2 - 221y^2 = 217$.

It follows from (b) and (c) that $(179, 12)$ is a fundamental solution.

Summarising: The fundamental solutions for $x^2 - 221y^2 = -217$ are $(\pm 2, 1)$, while those for $x^2 - 221y^2 = 217$ are $(\pm 179, 12)$.

Extensions

Lagrange discussed the general equation $ax^2 + bxy + cy^2 = N$, where $D = b^2 - 4ac > 0$ and not a perfect square and $\gcd(a, N) = 1 = \gcd(a, b, c)$.

The above analysis goes through with suitable modifications. However an exceptional case, not noted by Lagrange, arises when $D = 5$ and $aN < 0$. Then there are solutions not arising directly via convergents. This was pointed out by Serret in 1877 and quantified in 1984 by M. Pavone.

An example is $x^2 - xy - y^2 = -1$ where the solution $(0, 1)$ is such an exception.

Using an extension to our lemma, we are able to avoid using Pavone's theorem, whose proof is not simple.

If $\gcd(a, N) \neq 1$, then a suitable unimodular change of variables exists which will ensure that $\gcd(a, N) = 1$.