Solving the equation $ax^2 + bxy + cy^2 + dx + ey + f = 0$

Last modified 4-16-03

Introduction

This note will give some methods to solve equations

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{1}$$

in integers or rational numbers. For many of these equations, the combination of this page with the page "Solving the generalized Pell equation $x^2 - Dy^2 = N$" gives a complete, self-contained solution to the problem. For one case, we transform the equation into another equation that is well-covered in the literature. Everything here is well-known (with the possible exception of our method of "completing the square"), but no one source seems to cover methods for solving all of these equations.

We assume a, b, c, d, e, and f are given integers, not all three of a, b, c are zero, and we want to find integral or rational x and y that solve equation (1).

The general approach here begins by looking at the *discriminant* $\Delta = b^2 - 4ac$ of $ax^2 + bxy + cy^2$. If Δ is zero or a positive square, equation (1) is solved using factoring methods. Otherwise the general approach is to transform equation (1) to one of the form

$$x^2 - Dy^2 = N, \tag{2}$$

and solve equation (2). If we are looking for rational solutions, then, under (any of) our transformation(s), all solutions of (2) correspond to solutions of (1), while if we are looking for integral solutions an extra step may be needed to determine which solutions of equation (2) produce solutions of equation (1).

Finding integer solutions to (1) is not, in general, equivalent to finding rational solutions. For one thing, a given equation generally has "many more" rational solutions than integer solutions. For example, the equation $x^2 - 34y^2 = -1$ has infinitely many rational solutions, such as (3/5, 1/5) or (5/3, 1/3), but does not have any integral solutions. The equation $x^2+y^2 = 1$ has four integral solutions, and infinitely many rational solutions. Also, if an

integral solution is found to (1), it is usually easy to then find all the rational solutions. But, methods that give all rational solutions are seldom helpful in determining whether there are integral solutions, much less identifying all integral solutions. So, the problem of finding integral solutions and that of finding rational solutions are of a somewhat different character.

There are a number of topics discussed below, and as not all are relevant to every equation, here's a road map. There are four main cases. For each of the first two, 1) $\Delta = 0$ and 2) $\Delta > 0$ is square, see the section below with that heading. The last two cases are 3) $\Delta < 0$ and 4) $\Delta > 0$ is not a square. While these two are somewhat different, the starting point for each is the section, "Completing the square," which puts the equation into the form (2). Then use the appropriate subsection of the section, "Cases for $x^2 - Dy^2 = N$." For some equations, you will be done at this point. If not, for integral solutions, refer to the section, "Recovering solutions to the original equation," to complete the solution. For rational solutions, refer first to the section "Solving $Ax^2 + By^2 + Cz^2 = 0$" to get one rational solution (if there are any), and then to the section "All rational solutions from one" to finish the solution.

There are two additional sections that provide other comments. While equations of the form $Ax^2 + Bxy + Cy^2 = N$ can be put into the form (2) and then solved, there are methods that solve these equations more directly. There is a section that has some discussion of these equations and gives references for the direct methods. There are times you might want to transform an equation (1) that has $a = 0$ or $c = 0$ to an equivalent equation with $a \neq 0$ and $c \neq 0$, or just $a \neq 0$, and there is a section that shows how to do this.

If we are seeking integral solutions, the third case (above), $\Delta < 0$, is potentially not covered fully in this or the related web page. Methods here do transform any such equation into one of the form (2). If $N < 0$, then there are no solutions, while if $N = 0$ there is only the solution (to (2)) $x = 0$, $y = 0$. If $N > 0$, see the subsection "$D < 0$, $N > 0$" in the section "Cases for $x^2 - Dy^2 = N$." For integral solutions, we do show how to do a brute-force search, which is feasible when $\sqrt{-N/D}$ is not too large. But, sometimes (for integral solutions) it is better to use the methods of binary quadratic forms, and that is not covered here. Finding rational solutions when $\Delta < 0$ is fully covered here.

Methods presented here, and on the web page "Solving the generalized Pell equation $x^2 - Dy^2 = N$" are good when the coefficients of (1) are small, up to maybe 6 to 8 digits. For integral solutions to Pell equation

$x^2 - Dy^2 = \pm 1$ with large D, see Williams [13] or Lenstra [7]. For rational solutions (large D, any N) see Cremona and Rusin [3].

Self-contained approaches to solving (1) in integers are given in Chrystal, Edwards, and Mathews [2, 4, 9], with, perhaps, Edwards [4] being the most comprehensive.

Equations of the form $Ax^2 + Bxy + Cy^2 = N$

While equations of the form

$$Ax^2 + Bxy + Cy^2 = N. \tag{3}$$

can be handled by the methods discussed herein, there are methods that deal more directly with these equations.

We call $Ax^2 + Bxy + Cy^2$ a *binary quadratic form.* If $\Delta = B^2 - 4AC > 0$ is not square and $N \neq 0$, then Matthews' method [10] is the most efficient. If Δ is not a square and $N \neq 0$, you can use the methods of binary quadratic forms given in Hurwitz [6] and many other sources, e.g., [1, 8, 9]. If Δ is square or $N = 0$, then the equation is fully treated by the methods given here for (1).

We will not give the details of the standard method for solving (3) in integers, but we will outline here how this method works when Δ is not a square and $N \neq 0$. A form $f = f(x, y) = Ax^2 + Bxy + Cy^2$ *represents* N if there are integers x, y so that $Ax^2 + Bxy + Cy^2 = N$. In the theory of binary quadratic forms, an equivalence relation among forms is defined. Equivalent forms have the same discriminant and, for a given discriminant Δ, there are only a finite number of equivalence classes. Any two equivalent forms represent the same integers (this is not to say that a given integer cannot be represented by two inequivalent forms).

It is a theorem that if a form f represents N, there is a form g equivalent to f so that $g = g(x, y) = Nx^2 + Rxy + Sy^2$, i.e., the coefficient of the x^2 term of g is N. Of course g represents N by taking $x = 1$, $y = 0$. There is a notion of a *reduced form*, which are forms whose coefficients satisfy certain inequalities. There is a *reduction step* that, applied to a form, produces an equivalent form. Applied iteratively, this process, in a finite number of steps, produces all of the reduced forms equivalent to the original form f (if the discriminant is negative, there is only one reduced form equivalent to any given form; if the discriminant is positive, there may be several reduced forms equivalent to the original form).

There is a finite procedure for generating a list of the forms $g = g(x, y) = Nx^2+Rxy+Sy^2$ that have x^2 coefficient N, discriminant Δ, and xy coefficient R satisfying $R^2 \equiv \Delta \pmod{|4N|}$ and $-|N| < R \leq |N|$. To find all solutions to $f(x, y) = N$, it suffices to find which, if any, of these g are equivalent to f.

So the method of solving the binary quadratic form equation $Ax^2+Bxy+Cy^2 = N$ is to find the reduced forms equivalent to $f = Ax^2 + Bxy + Cy^2$, and for each $g = Nx^2 + Rxy + Sy^2$ in our list, find an equivalent reduced form, and see if this reduced form is in the list of reduced forms equivalent to f. If it is, then the solution $x = 1$, $y = 0$ to $g(x, y) = N$ can be tracked back to a solution to $f(x, y) = N$. All solutions to the original equation can be generated from the solutions found in this way.

Making $a \neq 0$ and $c \neq 0$

Some methods below require that $a \neq 0$ and $c \neq 0$, so we show how to transform an equation (1) with $a = 0$ or $c = 0$ to one with $a \neq 0$ and $c \neq 0$.

First suppose $a = 0$ and $c = 0$, so $b \neq 0$. Then the transformation $x = X + Y$, $y = X + 2Y$ makes (1)

$$b(X + Y)(X + 2Y) + d(X + Y) + e(X + 2Y) + f = 0,$$

or

$$bX^2 + 3bXY + 2bY^2 + (d + e)X + (d + 2e)Y + f = 0,$$

which is of the desired form.

Now suppose $a = 0$ and $c \neq 0$. Then the transformation $x = X$, $y = X + Y$ makes (1)

$$bX(X + Y) + c(X + Y)^2 + dX + e(X + Y) + f = 0,$$

or

$$(b + c)X^2 + (b + 2c)XY + cY^2 + (d + e)X + eY + f = 0.$$

This is of the desired form if $b + c \neq 0$. If $b + c = 0$, then the analogous transform with $x = X$, $y = -X + Y$ will have $c - b$ as the coefficient of X^2 and this will not be zero (and the coefficient of Y^2 will still be c).

The remaining possibility, $a \neq 0$ and $c = 0$, is analogous to the previous case, with the roles of a and c, x and y, and X and Y reversed.

If we just want to make $a \neq 0$, and we are indifferent to whether $c = 0$, we can proceed as follows. If $c \neq 0$, use the transformation $x = Y$, $y = X$. If $c = 0$ (so $b \neq 0$), use $x = X$, $y = X + Y$.

The Case $\Delta = 0$

First, consider the equation

$$X^2 - AY - K = 0 \tag{4}$$

where $A \neq 0$. For this to have integral solutions, it is necessary and sufficient that $X^2 \equiv K \pmod{A}$ have (obviously integral) solutions. If $\{x_i\}_{1 \leq i \leq N}$ is the set of X-values for solutions of (4) so that $0 \leq x_i < |A|$, then all X-values of solutions have $x = x_i + k|A|$ for some x_i and some integer k. For rational solutions, one can take any rational X, and $Y = (X^2 - K)/A$.

If $A = 0$, then we have

$$X^2 = K. \tag{5}$$

This only has solutions when K is square, say $K = r^2$. The solutions are $X = \pm r$, Y any integer or rational number.

Other equations (1) with $\Delta = 0$ can be reduced to one of the two equations (4, 5) just considered, as we now show.

In (1) a and c cannot both be zero, because if they were b would also have to be zero for Δ to be zero. If $a = 0$, switch x and y, a and c, and d and e, so a is not zero.

Now, after multiplying through by $4a$, (1) can always be rearranged as

$$(2ax + by)^2 - (b^2 - 4ac)y^2 + 4adx + 4aey + 4af = 0. \tag{6}$$

As $\Delta = 0$, this is

$$(2ax + by)^2 + 4adx + 4aey + 4af = 0,$$

or

$$(2ax + by + d)^2 - (2bd - 4ae)y - (d^2 - 4af) = 0.$$

Let $X = 2ax + by + d$, $A = 2bd - 4ae$, and $K = d^2 - 4af$, and we are reduced to one of the equations (4, 5). We recover x from $x = (X - by - d)/2a$. If we are seeking rational solutions, we are done. If we seek integral solutions, it is necessary that $2a$ divide $X - by - d$. If $A \neq 0$, for each x_i satisfying $x_i^2 \equiv K \pmod{A}$, $0 \leq x_i < |A|$, test x_i, $x_i + |A|$, $x_i + 2|A|$, $\ldots$, $x_i + 2a|A|$, and see which give solutions. For any $x_i + k|A|$ that gives an integral x, $x_i + k|A| + 2aj|A|$ will also be a solution, and all solutions arise this way. If $A = 0$, there are integral solutions, as above, only if K is a square, say $K = r^2$. For all y so that $x = (\pm r - by - d)/2a$ is an integer, x, y is a solution to the original equation.

$\Delta > 0$ is Square

First, consider the situation where $a = c = 0$. Of course $b \neq 0$, as otherwise the equation would be linear, so $\Delta = b^2 > 0$.

We have

$$bxy + dx + ey + f = 0. \tag{7}$$

Multiply through by b, and note that the equation factors as

$$(bx + e)(by + d) = ed - bf.$$

Write $ed - bf = N$. If $N \neq 0$, for every factorization of $N = rs$ into integers (and there are finitely many such factorizatons), solve the simultaneous equations

$$bx + e = r, \quad by + d = s$$

to get

$$x = (r - e)/b, \quad y = (s - d)/b.$$

If there are integral solutions to (7), they will be among these x and y. To find all rational solutions, let r be an arbitrary nonzero rational number, set $s = N/r$, and take the solution given just above.

If $N = 0$ there are two infinite sets of solutions, namely $x = -e/b$, any y, and $y = -d/b$, any x.

As an example, let's find the integer solutions to $3xy - 10x - y - 2 = 0$. Then $N = ed - bf = 16$, and the factors of 16 are ± 1, ± 2, ± 4, ± 8, ± 16. Taking $r = 2$, $s = 8$, gives $x = 1$, $y = 6$; $r = 8$, $s = 2$, gives $x = 3$, $y = 4$; $r = -1$, $s = -16$, gives $x = 0$, $y = -2$; $r = -4$, $s = -4$, gives $x = -1$, $y = 2$; $r = -16$, $s = -1$, gives $x = -5$, $y = 3$. All other r, s so that $rs = 16$ give non-integral solutions. So there are 5 integral solutions to this equation.

Returning to the general equation (1), we can now assume not both a and c are zero. If a is zero, switch x and y, etc. so that a is not zero. Now rewrite (1) as (6), substitute r^2 for $b^2 - 4ac$, and rewrite as

$$(2ax + by + ry)(2ax + by - ry) + 4adx + 4aey + 4af = 0. \tag{8}$$

Set $S = 2ax + by + ry$, $T = 2ax + by - ry$. We can write x and y as $x = (bT + rS + rT - bS)/(4ar)$, $y = (S - T)/(2r)$. Substituting in the obvious way, and clearing denominators puts the equation into the form

$$BST + DS + ET + F = 0. \tag{9}$$

This is in the form (7), and is solved by the methods above. Every rational solution of (9) gives a rational solution of (1). If there are finitely many integral solutions to (9), each can be checked to see if it produces a solution to (1). If there are infinitely many integral solutions to (9), then only finitely many equivalence classes modulo $4ar$ need to be checked to determine the integral solutions to (1).

Completing the Square

The purpose of the methods in this section is to transform equation (1) to equation (2). The methods of this section are most useful if the discriminant, $\Delta = b^2 - 4ac$, is not a square. In this case, note that $a \neq 0$ and $c \neq 0$. When the discriminant is a square, sections above show how to find solutions. But, the methods in this section apply to any equation (1) that has $a \neq 0$, and $c \neq 0$. If you want to transform an equation (1) to an equation (2), and $a = 0$ or $c = 0$, first use the methods in the section "Making $a \neq 0$ and $c \neq 0$."

As noted above, we will transform equation (1) to equation (2). Given a, b, c, d, e and f, we will find D and N. The transformation will be such that there are matrices $A = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$ and $B = \begin{pmatrix} v \\ w \end{pmatrix}$ so that if x, y is an integral solution to (1) then there are integral X, Y that solve

$$X^2 - DY^2 = N \tag{10}$$

and

$$\begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} + B \tag{11}$$

While r, s, t, u, v, w will be rational numbers, they will not, in general, be integers. So there may be integral solutions to (10) that do not correspond to integral solutions to (1) under the transformation (11). How to use the set of integral solutions to (10) to find the integral solutions to (1) in a finite number of steps is discussed in the section "Recovering solutions to the original equation." Any rational solution to the transformed equation corresponds to a rational solution of the original equation.

There are a number of ways to transform (1) into (2). For alternatives to what we present here, see Edwards [4, p. 339], Chrystal [2, p. 486], or Gauss [5, §216, p. 205].

First we transform (1) so as to make $b = 0$. If b is already 0, then skip this step. Note that we also assume $a \neq 0$. The idea is to find T, A, B, C so that

$$aTx^2 + bTxy + \ldots = A(Bx + Cy)^2 + \ldots = AB^2x^2 + 2ABCxy + \ldots$$

So, set $aT = AB^2$ and $bT = 2ABC$. Consider the ratio $aT/bT = AB^2/2ABC = B/2C$, which gives $2a/b = B/C$. Set B equal to the numerator of $2a/b$, reduced, and set C equal to the denominator. Also, $a/B^2 = A/T$, so set A equal to the numerator of a/B^2, reduced, and set T equal to the denominator. With $X = Bx + Cy$ and $Y = y$, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{B} & \frac{-C}{B} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Substituting in (1) for x, y in terms of X, Y gives

$$AX^2 + (cT - AC^2)Y^2 + (dT/B)X + (eT - dTC/B)Y + fT = 0,$$

or

$$AB'X^2 + B'(cT - AC^2)Y^2 + (dTB'/B)X + (eTB' - dTCB'/B)Y + fTB' = 0,$$

where B' is the least positive integer that makes all the coefficients integers. This last equation is of the form

$$ax^2 + cy^2 + dx + ey + f = 0,$$

i.e., of the form (1) with $b = 0$. The discriminant of the transformed equation is

$$-4(AB')(B')(cT - AC^2) = 4A^2B'^2C^2 - 4cAB'^2T$$
$$= (bTB'/B)^2 - 4ca(TB'/B)^2 = (TB'/B)^2(b^2 - 4ac).$$

So the discriminant of the transformed equation is a rational square times the discriminant of the original equation.

It is always possible to use $T = 4a$, $A = 1$, $B = 2a$, and $C = b$ to effect the above transformation, but the method as presented sometimes results in smaller coefficients for the transformed equation.

Now assume we have an equation (1) with $a \neq 0$, $b = 0$, and $d \neq 0$ (if $d = 0$ then skip this step). We will transform this equation into one with

$d = 0$ by a method similar to that used to make $b = 0$. We want T, A, B, C so that

$$aTx^2 + dTx + \ldots = A(Bx + C)^2 + \ldots = AB^2x^2 + 2ABCx + \ldots$$

where the "..." represent the "y" terms and the constant term.

Similarly to the step above, set $aT = AB^2$ and $dT = 2ABC$. Consider the ratio $aT/dT = AB^2/2ABC = B/2C$, which gives $2a/d = B/C$. Set B equal to the numerator of $2a/d$, reduced, and set C equal to the denominator of $2a/d$, reduced. Also, $a/B^2 = A/T$, so set A equal to the numerator of a/B^2, reduced, and set T equal to the denominator. With $X = Bx + C$ and $Y = y$, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{B} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} \frac{-C}{B} \\ 0 \end{pmatrix}.$$

Substituting we get

$$AX^2 + cTY^2 + eTY + fT - AC^2 = 0.$$

This last equation is of the form

$$ax^2 + cy^2 + ey + f = 0$$

i.e., of the form (1) with $b = d = 0$. As above, the discriminant of the transformed equation is a rational square times the discriminant of the equation we began with.

Making $e = 0$ is similar to making $d = 0$. Assume we have an equation (1) with $c \neq 0$, $b = d = 0$, and $e \neq 0$ (if $e = 0$ then skip this step). We will transform this equation into one with $e = 0$. We want T, A, B, C so that

$$cTy^2 + eTy + \ldots = A(By + C)^2 + \ldots = AB^2y^2 + 2ABCy + \ldots$$

where the "..." represent the "x" terms and the constant term.

Set $cT = AB^2$ and $eT = 2ABC$. Consider the ratio $cT/eT = AB^2/2ABC = B/2C$, which gives $2c/e = B/C$. Set B equal to the numerator of $2c/e$, reduced, and set C equal to the denominator of $2c/e$, reduced. Also, $c/B^2 = A/T$, so set A equal to the numerator of c/B^2, reduced, and set T equal to the denominator. With $Y = By + C$ and $X = x$, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{B} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{-C}{B} \end{pmatrix}.$$

Substituting we get

$$aTX^2 + AY^2 + fT - AC^2 = 0.$$

This last equation is of the form

$$ax^2 + cy^2 + f = 0$$

i.e., of the form (1) with $b = d = e = 0$. As above, the discriminant of the transformed equation is a rational square times the discriminant of the equation we began with.

Henceforth in this section we assume $b = d = e = 0$, i.e., our equation is of the form

$$ax^2 + cy^2 + f = 0 \tag{12}$$

with $a, c \neq 0$ and $\gcd(a, c, f) = 1$. If $\gcd(a, c)$ does not divide f, then there are no solutions to the transformed equation or the original equation, and we are done.

It may be possible to transform (12) so as to reduce the size of the coefficients a, c, f, which generally makes the equation easier to solve. For example, if a prime p divides a and f (and there is a solution), p must divide y. So we could write $x = X$, $y = pY$, $a = pA$, $cp = C$, $f = pF$, and we have

$$pAX^2 + c(pY)^2 = pF,$$

or (canceling a p)

$$AX^2 + CY^2 = F.$$

The variables transform as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

If a power of p divides a and f, the above can be repeated until p does not divide both.

A similar transformation can be made if a prime divides both c and f.

If a and c are odd, $a \equiv c \pmod 4$, and $f \equiv 0 \pmod 4$ then for any solution to (12) a and c must be even, as can be checked by considering the equation modulo 4. So we can write

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

and use the equation

$$aX^2 + cY^2 + f/4 = 0.$$

It might be possible to repeat this reduction.

Now we are ready to make our final transformation to the general Pell equation (2). If $|a| = 1$ or $|c| = 1$, the final transformation should be clear.

Otherwise, we can multiply through by a to get $(ax)^2 + acy^2 + af = 0$. Then take $X = ax$, $Y = y$, $D = -ac$, and $M = -af$ to have $X^2 - DY^2 = M$.

Rather than multiply through by a, it may be possible to obtain smaller coefficients in the final equation as follows. Find nonzero r of smallest magnitude (r might be negative) so that ra is a square, and find nonzero s of smallest magnitude so that sc is square. If $|rc| < |sa|$, use the transformation $X = x\sqrt{ra}$, $Y = y$, to get a Pell equation with $D = -rc$, $M = -rf$. If $|rc| > |sa|$, switch the roles of r and s and of x and y in the transformation of the previous sentence. If $|rc| = |sa|$, pick the transform that minimizes the magnitude of M.

To produce final matrices for the transform, combine the various individual transformations in the obvious way.

Cases for $x^2 - Dy^2 = N$

- $D < 0$, $N = 0$
  The only integral or rational solution is $x = y = 0$.

- $D < 0$, $N < 0$
  There are no integral or rational solutions.

- $D < 0$, $N > 0$
  There may or may not be solutions. If there are integer solutions, there will be finitely many, while if there are rational solutions, there will be infinitely many.

  A simple way to find integer solutions is to search on $0 \le y \le \sqrt{-N/D}$. For any such y for which $N + Dy^2$ is a square, set $x = \sqrt{N + Dy^2}$ and then solutions are $(\pm x, \pm y)$.

  Often a better method to find integer solutions is to use the method of binary quadratic forms for positive definite forms, given, for instance, in Hurwitz [6]. While we outline this method in the section "Equations of the

form $Ax^2 + Bxy + Cy^2 = N$," this is the one case where we do not give full details of the method. Alternatively, see [1, 8, 9].

For rational solutions, see the section "Solving $Ax^2 + By^2 + Cz^2 = 0$."

- $D = 0$

The equation reduces to $x^2 = N$. If $N < 0$, or if $N > 0$ is not a square, then there are no integral or rational solutions. If $N \geq 0$ is a square, then there are infinitely many solutions given by $x = \pm\sqrt{N}$, and any y (integer or rational as appropriate).

- $D > 0$, D is a square, $N = 0$

There are infinitely many integral or rational solutions. Write $D = r^2$. All solutions are $x = ry$, or $x = -ry$ for arbitrary y (integer or rational as appropriate).

- $D > 0$, D is a square, $N \neq 0$

There are finitely many integral solutions. Write $D = r^2$. Write $N = st$ in every possible way (including ways with $s < 0$). For each pair, s, t solve the simultaneous equations $x - ry = s$, $x + ry = t$ to get $x = (s + t)/2$, $y = (t - s)/2r$. Select the integral solutions.

Or, you can search on $0 \leq y \leq (N - 1)/2r$ when $N > 0$, or $0 \leq y \leq (1 - N)/2r$ when $N < 0$, taking $x = \sqrt{Dy^2 + N}$ when $Dy^2 + N$ is a square, and taking $(\pm x, \pm y)$ as the solutions.

There are always infinitely many rational solutions. One is $x = (N + 1)/2$, $y = (N - 1)/2r$. From this, all of the infinitely many solutions can be generated using the method discussed below in the section "All rational solutions from one."

- $D > 0$, D not a square, $N = 0$

The only integral or rational solution is $x = y = 0$.

- $D > 0$, D not a square, $N \neq 0$

First consider the case where we want integral solutions. This is the generalized Pell equation. Either there are no solutions, or there are infinitely many. If $N = +1$, there are always infinitely many solutions.

For $N = \pm 1$, $N = \pm 4$, or $N^2 < D$, use the continued fraction method. If N does not fall into one of these categories use LMM, brute-force search, Lagrange reduction, the method of binary quadratic forms, or the cyclic

method. The page, "Solving the generalized Pell equation $x^2 - Dy^2 = N$" gives references for all of these methods, and gives detailed descriptions of all but the last three of these methods.

For rational solutions, see the section "Solving $Ax^2 + By^2 + Cz^2 = 0$." If there are rational solutions, there will be infinitely many.

Recovering solutions to the original equation

We assume we have transformed equation (1) to equation (10), and that we have completely solved equation (10). We also assume we have matrices with rational entries $A = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$ and $B = \begin{pmatrix} v \\ w \end{pmatrix}$ so that if (x, y) is an integral solution to (1), then there is an integral solution (X, Y) to (10) with

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} v \\ w \end{pmatrix}$$

If we are interested in rational solutions to (1), then all rational solutions of (10) correspond to rational solutions to (1), and we are done.

From now on, for this section, assume we are interested in integral solutions to (1). If $D < 0$, or $D > 0$ is square and $N \neq 0$, then there are only finitely many solutions to (10), and we can check each to see whether it corresponds to a solution to (1).

Now assume $D > 0$. If r, s, t, u, v, and w are integers, then all solutions to (10) correspond to solutions to (1), and we are done. If at least one of r, s, t, u, v, and w is rational without being an integer, then we need a finite procedure to generate the integral solutions to (1) from the integral solutions to (10). Let L be the least common multiple of the denominators of r, s, t, u, v, and w.

If $D > 0$ is square, say $D = k^2$, and $N = 0$, then all of the solutions to (10) are given by $X = \pm kY$. Separately for each sign ± 1, the congruence class of Y determines whether $X = kY$ or $X = -kY$ corresponds to a solution to (1). So one only needs to check $Y = 0, 1, 2, \ldots, L - 1$ in each of these two equations, see which of these Y correspond to integer solutions to (1), and take the whole congruence class for any Y that does correspond to solutions to (1).

Now assume $D > 0$ is not square. Let (T, U) be the minimal positive solution to $x^2 - Dy^2 = 1$. Let k be the least positive integer so that $(T +$

$U\sqrt{D})^k \equiv 1 \pmod{L}$, i.e., writing $T_k + U_k\sqrt{D} = (T + U\sqrt{D})^k$, so that $T_k \equiv 1 \pmod{L}$ and $U_k \equiv 0 \pmod{L}$. [Exercise: show that k exists. Hint: $x = 1$, $y = 0$ is a solution to $x^2 - Dy^2 = 1$.]

Then for (X, Y) any solution to (10), (X, Y) generates an integral solution to (1) if and only if (X', Y') generates and integral solution to (1), where $X' + Y'\sqrt{D} = (X + Y\sqrt{D})(T + U\sqrt{D})^k$. So, it suffices to consider all (X, Y) that are fundamental solutions to (10), and for each, determine which of $(X + Y\sqrt{D})(T + U\sqrt{D})^i$ for $0 \le i \le k-1$ correspond to solutions to (1). For any such (X, Y) and i that correspond to a solution to (1), and any integer j, $(X + Y\sqrt{D})(T + U\sqrt{D})^{i+jL}$ also corresponds to a solution to (1), and all solutions to (1) arise in this manner.

Solving $Ax^2 + By^2 + Cz^2 = 0$

In the context of this article, the material in this section is used to find rational solutions to $x^2 - Dy^2 = N$ either when $D > 0$ is not a square and $N \neq 0$, or when $D < 0$ and $N > 0$. If Z is the common denominator of x and y, rational, then by writing $x = X/Z$ and $y = Y/Z$, our equation becomes $X^2 - DY^2 - NZ^2 = 0$, and so is of the form considered here.

We seek solutions of

$$Ax^2 + By^2 + Cz^2 = 0 \tag{13}$$

in rational numbers. We assume A, B, and C are nonzero integers.

The following method of solution, due to Lagrange, is given in Weil [12, pp. 100-101]. Or see Serre [11, p. 42]. If the coefficients of (13) are large, see the method of Cremona and Rusin [3].

For (13) to have nontrivial solutions, i.e., solutions other than $x = y = z = 0$, it is necessary and sufficient that the signs of A, B, and C are not all the same, and that there are solutions w to each of the following three equations:

$$\begin{aligned} w^2 &\equiv -BC \pmod{|A|}, \\ w^2 &\equiv -AC \pmod{|B|}, \\ w^2 &\equiv -AB \pmod{|C|}. \end{aligned}$$

Assuming (13) satisfies these conditions, we can multiply through by C, replace Cz with z, and rearrange and re-label coefficients to get an equation

of the form

$$z^2 = Ax^2 + By^2. \tag{14}$$

Here, A and B cannot both be negative. We can also assume both are squarefree, for if, say, $A = A'm^2$, we could set $X = mx$ and have the equivalent equation $z^2 = A'X^2 + By^2$. And we can assume that $|A| \le |B|$. (We do not assume that A and B are relatively prime.)

It suffices to solve (14) in integers. In fact, we can assume that x, y, and z are pairwise relatively prime. If a prime p were to divide, say, z and x, then p^2 would divide By^2 and, as B is squarefree, p would divide y. So the common factor of p could be removed from x, y, and z. Arguments for z and y, and for x and y are similar.

Now, x and B are relatively prime, as any common divisor would divide z, so x has an inverse x' modulo $|B|$. From $Ax^2 \equiv z^2 \pmod{|B|}$, we get $A \equiv (x'z)^2 \pmod{|B|}$. Note that this shows that for there to be a solution to (14), A must be a quadratic residue, not necessarily prime to B, modulo $|B|$ (which is one of our tests above).

Write $A \equiv a^2 \pmod{|B|}$, where $0 \le a \le \frac{1}{2}|B|$. Because A is squarefree, $A \ne a^2$ unless $A = 1$ and $a = \pm 1$. This case will be treated below. Write $a^2 - A = BB_1d^2$, where d is selected so that B_1 is squarefree. Then

$$|B_1| \le \left|\frac{a^2 - A}{B}\right| \le \frac{1}{4}|B| + 1.$$

So $|B_1| < |B|$ except when $B = \pm 1$ and $A = \pm 1$, which is easily solved. Otherwise,

$$B_1(Bdy)^2 = (a^2 - A)(z^2 - Ax^2) = (az \pm Ax)^2 - A(ax \pm z)^2.$$

Putting $X = ax \pm z$, $Y = Bdy$, $Z = az \pm Ax$, we have that (X, Y, Z) is a solution of $Z^2 = AX^2 + B_1Y^2$. Now $A \equiv a^2 \pmod{|B_1|}$, so this reduction can be repeated until we reach an equation $Z^2 = AX^2 + B_nY^2$ with $|B_n| < |A|$. Then put $A' = |B_n|$, $B' = A$, and continue. As the magnitudes of the coefficients A and B decrease steadily, one of two things must occur. One is that we eventually reach an equation of the form $Z^2 = MX^2 + NY^2$ where M is not a quadratic residue modulo $|N|$. In this case (13) does not have a solution. Note that this cannot occur if (13) passed the tests that we set above. The other possibility is that we eventually reach an equation of the

form $Z^2 = X^2 + NY^2$, which has the solution $(1, 0, 1)$. We obtain a solution to the original equation (13) by following our steps backwards.

If we have found one solution to our equation (2), then the next section shows how to find all rational solutions.

All rational solutions from one

If the equation (2) has a rational solution, then it has infinitely many rational solutions, when either $D > 0$ and $N \neq 0$ or $D < 0$ and $N > 0$. In both these cases, the equation (2), taken over the reals, is a non-degenerate conic section. In the first case, it is a hyperbola, and in the second case it is an ellipse.

Clearly, if we have two rational points on a curve, the slope of the line connecting them is rational, or undefined (if they have the same x–coordinate). Conversely, if you have a rational point on a conic section, and draw a line with rational slope through it, that line either intersects the conic at exactly one more rational point, or is tangent to the curve at the original point.

More formally, suppose we have a non-degenerate curve (2), with the rational point $(x_0,\ y_0)$ on it. Draw a line with rational slope α through this point. This line has equation

$$\frac{y - y_0}{x - x_0} = \alpha.$$

Solving this for x gives

$$x = \frac{1}{\alpha}(y - y_0) + x_0. \tag{15}$$

Plugging this in to (2), and substituting $x_0^2 - Dy_0^2$ for N (because $x_0^2 - Dy_0^2 = N$), and collecting all terms on the left hand side, gives an expression that factors as

$$(y - y_0)\left(\frac{1}{\alpha^2}(y - y_0) - D(y + y_0) + \frac{2}{\alpha}x_0\right) = 0.$$

This gives two solutions for y, namely $y = y_0$ (which we already knew) and

$$y = \frac{y_0 + D\alpha^2 y_0 - 2\alpha x_0}{1 - D\alpha^2}.$$

This second solution is rational whenever x_0 and α are rational, and $1 - D\alpha^2 \neq 0$. If $\alpha \neq 0$, use (15) to get x, which is also rational. If $\alpha = 0$,

the second solution is $(-x_0,\ y_0)$. For the vertical line through $(x_0,\ y_0)$, the second solution is $(x_0,\ -y_0)$.

This method generates all of the rational solutions from any one.

As an example, consider the curve $x^2 + y^2 = 1$, which has the rational point $(-1, 0)$. Take the line through this point with slope $\alpha = b/a$. Then the other point of intersection has

$$x = \frac{1-\alpha^2}{1+\alpha^2} = \frac{a^2-b^2}{a^2+b^2}$$

and

$$y = \frac{2\alpha}{1+\alpha^2} = \frac{2ab}{a^2+b^2}.$$

References

[1] D. A. Buell, *Binary Quadratic Forms, Classical Theory and Modern Computations*, Springer-Verlag, New York, 1989. As the title says, including solution of binary quadratic form equations, summarized in Theorems 4.26 and 4.27 on page 75.

[2] G. Chrystal, *Algebra, An Elementary Text-Book* (a.k.a. *Textbook of Algebra*), Parts I and II, Dover, NY, 1961. Other editions include Adam and Charles Black, 1900, and Chelsea, 1964, and AMS currently. Data for the current AMS edition is: *Algebra, an Elementary Text-Book for the Higher Classes of Secondary Schools and for Colleges*: Seventh Edition - G. Chrystal - AMS — CHEL, 1964, 1212 pp., Hardcover, ISBN 0-8218-1931-3, List: $40, All AMS Members: $36, CHEL/84.H. Part II, Chapter XXXIII, §15 – §20, pages 478 to 490, covers the general equation (1). Part I, Chapter VII, §13, pages 140 to 142, shows how to factor (1) when this is possible, and it is not difficult to modify the given approach to complete the square when factorization is not possible.

[3] John E. Cremona and Dave Rusin, Efficient solution of rational conics, *Math. Comp.*, posted on December 18, 2002, PII S 0025-5718(02)01480-1 (to appear in print). Finds rational solutions when coefficients are large. Also available at http://www.maths.nott.ac.uk/personal/jec/papers/index.html. See also http://www.math.niu.edu/~rusin/papers/research-math/legendre/legendre.dvi, which includes some associated Maple code.

[4] Harold M. Edwards, *Fermat's Last Theorem*, Springer-Verlag, NY, 1977. Chapters 7 (pp. 245 to 304) and 8 (pp. 305 to 341), especially sections 8.2 (pp. 313-318) and 8.7 (pp. 339-341) apply the cyclic method to solve equations of the form $ax^2 + bxy + cy^2 + dx + ey + f = 0$.

[5] Carl Friedrich Gauss, *Disquisitiones Arithmeticae*, English edition, translated by Arthur A. Clarke, revised by William C. Waterhouse, Springer-Verlag, New York, 1986. A classic.

[6] Adolf Hurwitz, *Lectures on Number Theory*, Springer-Verlag, New York, 1986. Chapter 6 is a wonderful exposition of methods to solve binary quadratic form equations, $Ax^2+Bxy+Cy^2 = N$. Much theory of simple continued fractions is also developed.

[7] H. W. Lenstra Jr., Solving the Pell equation, *Notices of the American Mathematical Society*, **49** No. 2 (February 2002), pp. 182–192. The equation $x^2 - Dy^2 = \pm 1$ for large D.

[8] William Judson Leveque, *Topics in Number Theory*, Volumes I and II, Dover, New York , 2002. Chapter 1 of Volume II gives the standard method of solution of binary quadratic form equations for positive and negative discriminants.

[9] G. B. Mathews, *Theory of Numbers*, Chelsea, New York, not dated. Chapters III to VI discuss the theory of binary quadratic forms. In Chapter IX a method for solving equations $ax^2+bxy+cy^2+dx+ey+f = 0$ is outlined.

[10] Keith Matthews, The diophantine equation $ax^2 + bxy + cy^2 = N$, $D = b^2 - 4ac > 0$, *J. Théor. Nombres Bordeaux* 14 (2002), 257-270. The most efficient method for solving binary quadratic form equations when the discriminant is positive. See www.numbertheory.org/papers.html#jntb for additions and www.maths.uq.edu.au/˜krm/ for related material.

[11] Jean-Pierre Serre, *A Course in Arithmetic*, Springer-Verlag, New York, 4th corrected printing, 1993. Page 42 gives the reduction method for finding rational solutions.

[12] André Weil, *Number Theory – An approach through history from Hammurapi to Legendre*, Birkhaüser, Boston, 2001. Pages 100–101 give the reduction method for finding rational solutions.

[13] H. C. Williams, Solving the Pell equation, in Bruce Berndt et al., *Surveys in Number Theory: Papers from the Millennial Conference on Number Theory*, A. K. Peters, 2002. Also included in *Number Theory for the Millennium*, Volumes 1, 2, 3, M. A. Bennett et al. editors, A. K. Peters, 2002. Williams' web page gives this last reference as H. C. Williams, Solving the Pell equation, *Proc. Millennial Conference on Number Theory*, A. K. Peters, Natick MA, 2002, pp. 397-435. Discusses the equation $x^2 - Dy^2 = \pm 1$. Terrific overview, including discussion when D is large.

Send comments to John Robertson at jpr2718@aol.com