Linear Diophantine Equations

A *diophantine equation* is any equation in which the solutions are restricted to integers.

The word *diophantine* is derived from the name of the ancient Greek mathematician Diophantus, who was one of the first people to consider such problems systematically. Diophantus lived in Alexandria around 250C.E. and he wrote a textbook called *Arithmetica*, one of the earliest known manuscripts on algebra. The most famous diophantine equation is the equation

$$x^n + y^n = z^n$$

which was studied by Fermat and is the subject of the notorious problem known as *Fermat's Last Theorem*. One solution is always possible, just by taking x to be zero and setting $y = z$ to any integer, but this is a trivial solution and one desires to find non-trivial solutions in which all of x, y, and z are non-zero.

If $n = 0$, 1, or 2 there are many non-trivial solutions (the solutions have to be integers) but for any integer $n > 2$ there are no non-trivial solutions at all. This was stated by Fermat in the year 1637, and the first correct proof was published in 1995 by Richard Taylor and Andrew Wiles. The excellent Wikipedia article on Fermat's Last Theorem is highly recommended; there is also a PBS documentary on it. In this lecture we consider only the *linear* diophantine equations, which are easy to solve using our knowledge of the Euclidean algorithm.

We solve the linear diophantine equation $ax = b$ in a *single variable* x, for given integers a, b. Obviously if $ax = b$ and a, x, b are integers then $a \mid b$ and $x = b/a$. If $a \nmid b$ (a does not divide b) then the diophantine equation $ax = b$ has no solution.

Don't forget: Solutions to diophantine equations must be integers. Now that we have solved the linear diophantine equation in one variable, let us consider the *two-variable* linear equation, which is considerably more interesting. The linear diophantine equation in two variables x, y is the equation

$$ax + by = c$$

where a, b, c are given integers. Let's put $g = \gcd(a, b)$ for ease of notation. Observe that g must divide the left hand side of the equation, so if $g \nmid c$ (g does not divide c) then there are no solutions. Now assume that $g \mid c$. Then g is a common divisor of all integers a, b, c and so we can simplify the given equation by dividing through by g. This produces a new equation, that we may as well call the *reduced* equation:

$$Ax + By = C \qquad \text{with } A = \tfrac{a}{g}, B = \tfrac{b}{g}, \text{ and } C = \tfrac{a}{g}.$$

In the reduced equation, we have that $\gcd(A, B) = 1$. It now suffices to solve the reduced equation, because the integer solutions to the reduced equation are clearly the same as the integer solutions to the original one.

We can easily find ONE solution to the reduced equation by the Euclidean algorithm, which gives integers s, t such that $As + Bt = 1$. Then multiply both sides by C to get $A(sC) + B(tC) = C$. This shows that $x_0 = sC$, $y_0 = tC$ is a solution of the reduced equation; it will also be a solution of the original equation. We are nearly done solving the two-variable problem. The remaining question is how to find *all* solutions, now that we have found one solution. It is pretty easy to find more solutions from the one we already have: just put $x = x_0 + Bn$, $y = y_0 - An$ for any integer n and check that this solves the reduced linear equation $Ax + By = C$:

$$\begin{aligned} A(x_0 + Bn) + B(y_0 - An) &= Ax_0 + ABn + By_0 - ABn \\ &= Ax_0 + By_0 = C. \end{aligned}$$

At this point we have found an infinite number of solutions. Still, there might be other solutions that we just haven't noticed. It turns out, however, that there are no other solutions: this procedure produces them all. We need to *prove* the last claim. Suppose that the pair x, y is *any* solution to the reduced equation $Ax + By = C$. Compare this with the solution pair x_0, y_0 produced by the Euclidean algorithm. This gives two equations:

$$\begin{aligned} Ax + By &= C \\ Ax_0 + By_0 &= C \end{aligned}$$

and by subtracting the second equation from the first we obtain the equation

$$A(x - x_0) + B(y - y_0) = 0.$$

Since A and B are relatively prime, the *only* solutions to the above are the obvious ones, $x - x_0 = Bn$, $y - y_0 = -An$. Thus $x = x_0 + Bn$, $y = y_0 - An$,

as stated above. This completes the proof. It is perhaps worth stating explicitly the fact that we just used to finish the proof that we have found all the solutions to the linear diophantine equation in two variables.

Lemma. *If* $\gcd(A, B) = 1$ *then the only solutions to the diophantine equation* $Au + Bv = 0$ *are of the form* $u = Bn$, $v = -An$ *where* n *is an arbitrary integer.*

The proof is left as an exercise for you.

We summarize the results on linear diophantine equations in two variables, in the form of a theorem, all the parts of which are now proved (once you finish the proof of the lemma).

Theorem. *Given integers* a, b, c *the linear diophantine equation* $ax + by = c$ *has no solution unless* $g = \gcd(a, b)$ *divides* c. *If* $g \mid c$ *then we can reduce the equation to* $Ax + By = C$ *where* $A = a/g$, $B = b/g$, *and* $C = c/g$. *Now* $\gcd(A, B) = 1$ *and we can find an integer solution* $x_0 = sC$, $y_0 = tC$ *by first finding* s, t *such that* $As + Bt = 1$ *by the Euclidean algorithm. Then* $x = x_0 + Bn$, $y = y_0 - An$ *(where* n *is an arbitrary integer) gives the complete solution.*

What about linear diophantine equations in more than two variables? With a bit of care, they can be solved as well. Consider the linear equation in three variables:

$$ax + by + cz = d. \tag{$*$}$$

If $\gcd(a, b, c)$ does not divide d then there are no solutions. So assume that $\gcd(a, b, c)$ divides d. Then we can reduce the given equation by dividing through by the gcd, just as we did before. This gives a new equation of the form $Ax + By + Cz = D$. To solve it let $G = \gcd(A, B)$ and solve

$$Ax + By = G \tag{1}$$

$$Gw + Cz = D. \tag{2}$$

Notice that (1) is not reduced, but you can replace it by the reduced equation $(A/G)x + (B/G)y = 1$. If (w_0, z_0) is a solution to (2) and (x_0, y_0) is a solution to (1) then $(x_0 w_0, y_0 w_0, z_0)$ is a solution to the original problem $(*)$, and all solutions to $(*)$ are obtained in this way. With even more variables you can iterate this idea. Thus it is possible to solve linear diophantine equations in any number of variables, so long as the gcd of the coefficients divides the constant term. If that is so, there are infinitely many solutions except in the one variable case.