

the fact that you encounter them (more precisely, you encounter their equations) when you compute the arc lengths of ellipses.

Since the early 1950s it has become clear to mathematicians that elliptic curves are important and fundamental objects that have connections to many areas of mathematics, including number theory, geometry, cryptography, and the mathematics of data transmission. For instance, when Andrew Wiles proved Fermat's last theorem in 1994, he did it by proving a result about elliptic curves—indeed by establishing a close connection between elliptic curves and another important part of mathematics, the theory of modular forms. (Don't even ask what those are. They cannot be described in simple terms—at least it's beyond my ability to do so.) A proof of the Birch and Swinnerton-Dyer conjecture will have ramifications throughout modern mathematics.

Although the conjecture itself is buried deep in highly advanced mathematics, we can approach it from some very humble beginnings: Pythagoras's theorem and the formula for calculating the area of a triangle.


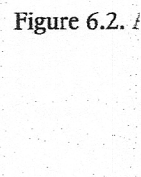
### *Half the Base Times the Height*

A classical problem, dating back to the ancient Greeks, goes as follows. Suppose you are given a positive whole number  $d$ . Is there a right triangle whose sides are rational numbers (i.e., whole numbers or fractions) for which the area is exactly  $d$ ? For example, in the case  $d = 6$ , the answer is yes. The famous Pythagorean right triangle with sides 3, 4, and 5, shown in Figure 6.1, has area

$$A = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 4 \times 3 = 6$$

In the case  $d = 5$ , there is no right triangle with whole-number sides that has area 5, but the right triangle with sides  $\frac{3}{2}$ ,  $\frac{20}{3}$ ,  $\frac{41}{6}$ , shown in Figure 6.2, has area 5.

It is a fairly straightforward piece of algebraic reasoning to show that there is a right triangle with rational sides having an area  $d$  if and only if the equation

Figure 6.1. Figure 6.2. 

has rational  
Equation

1. For those  
there is a right  
 $y = \frac{1}{2}a^2(c -$   
such that  $y^2$

is right-angled

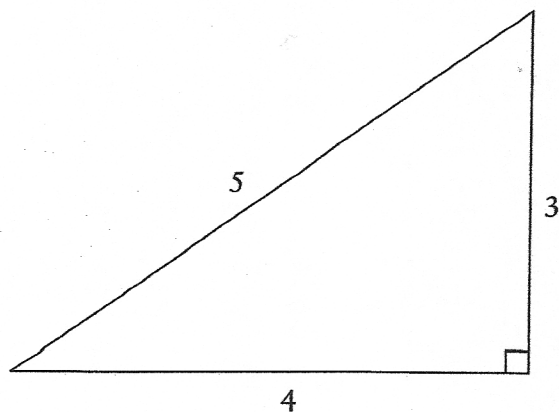


Figure 6.1. A right triangle with whole-number sides and area 6.

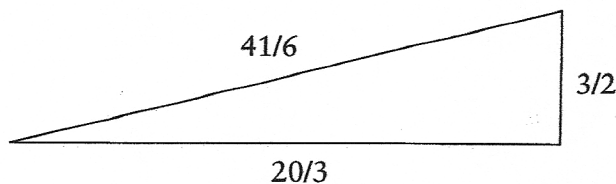


Figure 6.2. A right triangle with rational sides and area 5.

$$y^2 = x^3 - d^2x$$

has rational solutions for  $x$  and  $y$  with  $y \neq 0$ .<sup>1</sup>

Equations of the general form

$$y^2 = x^3 + ax + b$$

1. For those who are interested, here is the main idea of that argument. First, if there is a right triangle with rational sides  $a, b, c$  and area  $d$ , then  $x = \frac{1}{2}a(a-c)$ ,  $y = \frac{1}{2}a^2(c-a)$  solves the equation. Conversely, if  $x$  and  $y$  are rational numbers such that  $y^2 = x^3 - d^2x$ , with  $y \neq 0$ , then the triangle with sides

$$\left| \frac{x^2 - d^2}{y} \right|, \left| \frac{2xd}{y} \right|, \left| \frac{x^2 + d^2}{y} \right|$$

is right-angled and has area  $d$ .



where  $a$  and  $b$  are whole numbers determine what are called elliptic curves, i.e., the graph of such an equation is an elliptic curve.<sup>2</sup>

A natural question to ask is, Why there is no term involving  $x^2$ ? Why don't we allow equations of the kind  $y^2 = x^3 + ax^2 + bx + c$ ? The answer is that with a fairly simple bit of algebra you can transform such an equation to one in which there is no  $x^2$  term. Thus, to study elliptic curves, the only equations you need to look at are those of the form  $y^2 = x^3 + ax + b$ .

Some elliptic curves can look a bit strange at first. If you try to draw a graph of an equation

$$y^2 = x^3 + ax + b$$

then whenever  $x^3 + ax + b$  is negative, you won't get an answer for  $y$ . (More precisely,  $y$  will be an imaginary number.) As a consequence, elliptic curves often fall into two separate pieces, as shown in Figure 6.3. (For an elliptic curve that is in one piece, see Figure 6.4. Whether an elliptic curve is in one piece or two depends on whether the cubic expression on the right of the equation has one real root or three.)

What about the triangle area equation  $y^2 = x^3 - d^2x$ , where  $d$  is a whole number? As we noted, this equation has rational-number solutions for  $x$  and  $y$  precisely when  $d$  is the area of a right triangle with rational sides. For this equation the discriminant is  $\Delta = -16(-4d^2) = 64d^2$ , which is nonzero, so the graph of the equation is an elliptic curve. (Put  $a = -d$  and  $b = 0$  in the formula for the discriminant.) Thus, the ancient Greek problem of finding whole numbers  $d$  that are areas of right triangles with rational sides is equivalent to the problem of finding rational points (i.e., points whose coefficients are rational numbers) on certain elliptic curves. That is the problem Birch and Swinnerton-Dyer set out to investigate.

2. Strictly speaking, for the graph to be an elliptic curve, the equation should satisfy an additional condition: Its discriminant should be nonzero. The discriminant is the quantity  $\Delta = -16(4a^3 + 27b^2)$ .

Figure 6.  
pieces, it

The  
some w  
curves.  
could b  
metaph  
On  
times v  
what B  
we first

We are  
less (an  
minute  
on fore  
through  
After w  
say this

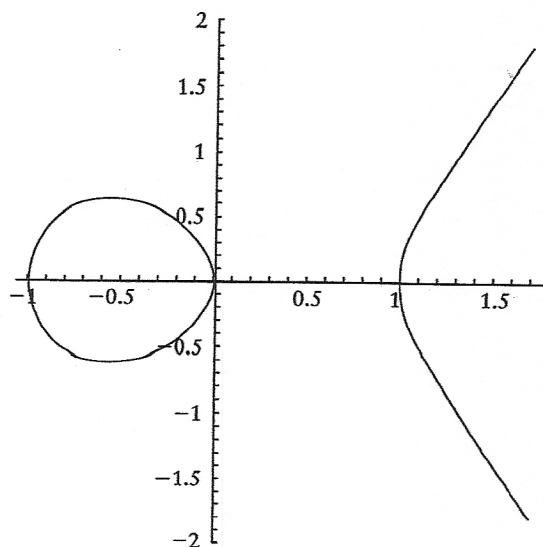


Figure 6.3. The elliptic curve  $y^2 = x^3 - x$ . Although it splits into two separate pieces, it is a single curve, determined by a single equation.

The two researchers approached the task by trying to find some way of “counting” the number of rational points on elliptic curves. Of course, since they were dealing with collections that could be infinite, the word “counting” has to be taken somewhat metaphorically.

One way to count a possibly infinite collection that sometimes works is to carry out a series of finite subcounts. This is what Birch and Swinnerton-Dyer did. To describe their method, we first have to break off and talk about finite arithmetic.

### *Counting by the Clock: Finite Arithmetic*

We are all familiar with one situation where we count an endless (and hence potentially infinite) collection: the way we count minutes. There is (let's be optimistic) no last minute—time goes on forever. And yet we count minutes using just sixty numbers, 0 through 59. What we do, of course, is keep restarting the count. After we reach 59 minutes, we start again at 0. Another way to say this is that we treat 60 as if it were 0. Mathematicians would



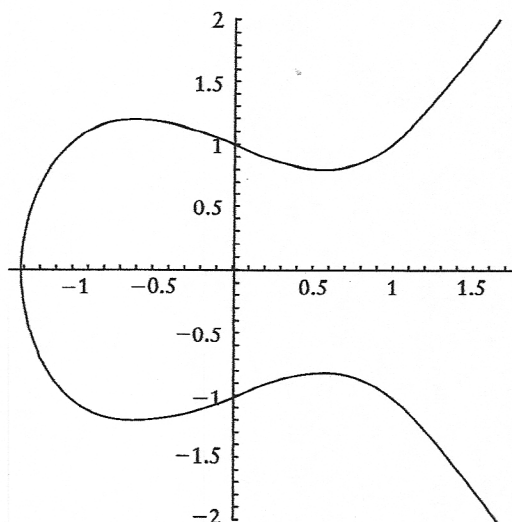


Figure 6.4. The elliptic curve  $y^2 = x^3 - x + 1$ . An example of an elliptic curve that does not split into two pieces.

say that we count minutes *modulo* 60. (The number 60 is the *modulus* for this counting.) Similarly, we count hours modulo 12 (or modulo 24).

For any positive whole number  $N$ , we can count modulo  $N$ . The numbers we use for this counting are  $0, 1, 2, \dots, N - 1$ . After  $N - 1$ , we start again at 0.

We can then do arithmetic modulo  $N$ . To explain how, let's take the case  $N = 7$ . For this modulus, the counting numbers are  $0, 1, 2, 3, 4, 5, 6$ . When we add any two numbers in this range, we discard 7 whenever it arises. For example, in arithmetic modulo 7,

$$2 + 3 = 5, \quad 3 + 4 = 0, \quad 4 + 5 = 2, \quad 6 + 6 = 5$$

Of course, this looks strange, and could be confusing, so mathematicians don't write it this way. Instead, they express the above additions like this:

$$2 + 3 \equiv 5 \pmod{7}$$

They  
four  
modu  
You c  
M  
the tw  
7. Th

Simila  
Sir  
subtra

(To ch  
ond te  
Ho  
can alv