Chapter 3
Integral solutions to the Pythagorean Equation

In this chapter we present two different methods to find the solutions of the Pythagorean Equation, one algebraic and one geometric. We then apply the geometric method to find solutions of some other equations. The first class of non-Pythagorean Equations that we will apply this method to is Pell's equation, and the second class, equations of degree three. As an application of our solution to the Pythagorean Equation we will prove a special case of *Fermat's Last Theorem*. In the Notes, we briefly review some classical works related to Pell's Equation over integers; explain why some cubic equations are called *elliptic*; give some references related to Fermat's Last Theorem; and discuss the *abc* Conjecture.

3.1 Solutions

Suppose (a, b, c) is a triple of integer solutions to the Pythagorean Equation. Then by definition

$$a^2 + b^2 = c^2. \tag{3.1}$$

If a, b, c have a common factor λ, then

$$\left(\frac{a}{\lambda}\right)^2 + \left(\frac{b}{\lambda}\right)^2 = \left(\frac{c}{\lambda}\right)^2.$$

So without loss of generality we may assume that a, b, c have no common factors. These are the triples we called *primitive* in Chapter 1. The Pythagorean triples we consider in this chapter are all primitive. A quick computer search produces the following list of the first few Pythagorean triples:

```
(3, 4, 5)
(5, 12, 13)
(8, 15, 17)
(7, 24, 25)
```

R. Takloo-Bighash, *A Pythagorean Introduction to Number Theory*,
Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-02604-2_3

```
(20, 21, 29)
(12, 35, 37)
(9, 40, 41)
(28, 45, 53)
...
```

Our goal in this section is to find all primitive solutions of Equation (3.1). Since $\gcd(a, b, c) = 1$, it is clear that not all a, b, c are even. We recognize several possibilities.

- a, b, c odd. This is impossible, as one side will be odd and the other side even.
- a, b odd, c even. If a is odd, then $a^2 \equiv 1 \mod 8$, and $b^2 \equiv 1 \mod 8$; hence $a^2 + b^2 \equiv 2 \mod 8$. But since c is even, $4 \mid c^2$, so $c^2 \equiv 0, 4 \mod 8$. So this case is impossible as well.
- a even, b odd, c odd.
- a odd, b even, c odd.

We will see momentarily that these last two cases are in fact possible. By symmetry we may assume that a is even, and b odd. Write

$$b^2 = c^2 - a^2 = (c - a)(c + a).$$

We claim that $\gcd(c - a, c + a) = 1$. To see this, we have

$$\gcd(c - a, c + a) = \gcd(c + a, c + a - (c - a)) = \gcd(c + a, 2a).$$

But since $c + a$ is odd, $\gcd(c + a, 2a) = \gcd(c + a, a) = \gcd(c, a)$. If there is a prime number $p | \gcd(c, a)$, then $p|a^2$ and $p|c^2$ so $p|b^2 = c^2 - a^2$, and consequently, $p|b$. This statement contradicts the assumption that $\gcd(a, b, c) = 1$. Since the product of the coprime numbers $c + a$ and $c - a$ is a square b^2, by Proposition 2.21 each of them individually is a square, i.e., there are odd coprime integers x, y such that

$$c + a = x^2, c - a = y^2.$$

Solving for c, and a, gives

$$\begin{cases} a = & \frac{x^2 - y^2}{2}, \\ b = & xy, \\ c = & \frac{x^2 + y^2}{2}. \end{cases}$$

It is of course true that

$$\left(\frac{x^2 - y^2}{2}\right)^2 + (xy)^2 = \left(\frac{x^2 - y^2}{2}\right)^2$$

as one can easily check. For example, if $(x, y) = (3, 1)$ we recover the well-known triple $(4, 3, 5)$, and if $(x, y) = (5, 1)$, then we get $(12, 5, 13)$. In general, instead of writing a triple as ordered vector, we write the triple as a set. So instead of $(12, 5, 13)$ we write $\{12, 5, 13\}$, and our general solution will be written as

$$\left\{\frac{x^2-y^2}{2}, xy, \frac{x^2+y^2}{2}\right\}.$$

We summarize this discussion in the following theorem:

Theorem 3.1. *Let a, b, c be the three sides of a primitive integral right triangle. There are odd coprime integers x, y such that*

$$\{a, b, c\} = \left\{\frac{x^2-y^2}{2}, xy, \frac{x^2+y^2}{2}\right\}.$$

3.2 Geometric method to find solutions

In this section we present a geometric method to find the solutions of the Pythagorean Equation. First a piece of notation: For a point $(x, y, z) \in \mathbb{R}^3$ with $z \neq 0$ we define $R(x, y, z)$ be the point $(x/y, x/z) \in \mathbb{R}^2$. If it is clear that if $(x, y, z) \in \mathbb{Z}^3$ with $z \neq 0$, then $R(x, y, z)$ be a *rational point*, i.e., a point with coordinates that are rational numbers in $\mathbb{R}^2$. Suppose (a, b, c) is a primitive solution of the Pythagorean Equation. We have

$$\left(\frac{a}{c}\right)^2 + \left(\frac{b}{c}\right)^2 = 1.$$

This means that $R(a, b, c)$ is a point with rational coordinates on the unit circle $x^2 + y^2 = 1$. Now suppose we have a rational point $(a/b, c/d)$, $a, b, c, d \in \mathbb{Z}$, on the unit circle centered at the origin, and suppose that the rational numbers a/b and c/d are in reduced form, meaning $\gcd(a, b) = \gcd(c, d) = 1$. We wish to show that there is a primitive solution (x, y, z) of the Pythagorean Equation such that $R(x, y, z) = (a/b, c/d)$. This claim is obvious if one of the coordinates $a/b, c/d$ is zero. So we assume that $ac \neq 0$. After changing the signs if necessary we assume $a, b, c, d > 0$. Since $(a/b)^2 + (c/d)^2 = 1$, $a^2d^2 + c^2b^2 = b^2d^2$. Since $b^2|c^2b^2$ and $b^2|b^2d^2$ we conclude $b^2|a^2d^2$, but since we have assume $\gcd(a, b) = 1$, by Theorem 2.17, we have $b^2 \mid d^2$. This means $b \mid d$. Similarly, $d \mid b$. Consequently, $b = d$. As a result every rational point in the first quadrant on the unit circle will be of the form $(a/b, c/b)$ with a, b, c natural numbers and $\gcd(a, b) = 1$ and $\gcd(c, b) = 1$. Also, we have $a^2 + c^2 = b^2$, i.e., (a, c, b) is a solution of the Pythagorean Equation. It is also easy to see that $\gcd(a, c) = 1$. In fact, if u is a common factor of a and c, then $u^2 \mid a^2 + c^2 = b^2$, giving $u^2 \mid b^2$, from which it follows $u \mid b$. This implies $u \mid \gcd(a, b) = 1$. Hence $u = 1$. Summarizing, for a rational point (x, y) on the unit circle with $x, y > 0$, there are pairwise coprime natural numbers a, b, c such that $x = a/b$, $y = c/b$ and $a^2 + c^2 = b^2$. This means $R(a, c, b) = (x, y)$. Note that $R(-a, -c, -b) = (x, y)$ as well, and (a, c, b) and $(-a, -b, -c)$ are the only primitive Pythagorean triples whose R is (x, y). Finally if either of x or y is negative, we can adjust the sign of a or c to get the correct sign. The map R is always 2-to-1 from primitive Pythagorean triples to the set of rational points on the unit circle.

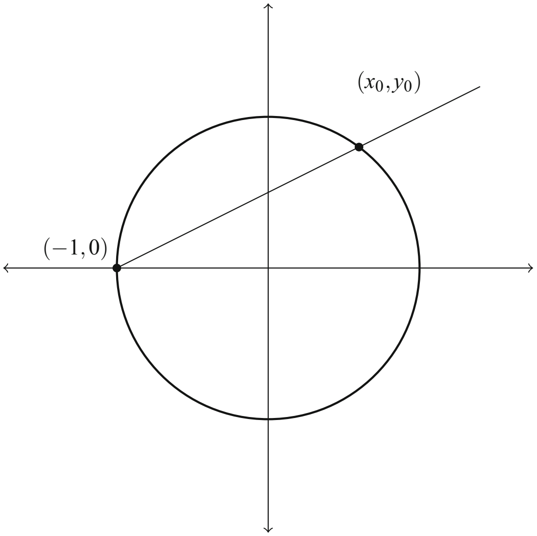

Fig. 3.1 Finding rational points on the unit circle. Here we have connected the point $(-1, 0)$ to the point (x_0, y_0)

A consequence of this discussion is that in order to find Pythagorean triples it is sufficient to determine rational points on the unit circle.

We proceed to determine the set of rational points on the unit circle. The circle $x^2 + y^2 = 1$ in Figure 3.1 has some obvious solutions, e.g., $(\pm 1, 0)$ or $(0, \pm 1)$. Let's pick one of these points, say $(-1, 0)$. The main observation is that if (x_0, y_0) is a point with rational coordinates, then the slope of the line connecting this point to the base point $(-1, 0)$ is

$$m = \frac{y_0}{x_0 + 1}$$

is a rational number.

Our idea is to do the opposite of this, i.e., pass a line with rational slope through $(-1, 0)$, look at the point of intersection of the line with the circle $x^2 + y^2 = 1$, and hope that the resulting point is a rational point. The equation of the line with slope m through $(-1, 0)$ is

$$y = m(x + 1).$$

To find the point of intersection of this line with the circle we need to solve the system of equations

$$\begin{cases} y = m(x + 1), \\ x^2 + y^2 = 1. \end{cases}$$

Inserting the value of y from the first equation in the second equation gives

$$x^2 + m^2(x + 1)^2 = 1.$$

Simplifying gives

$$(m^2 + 1)x^2 + 2m^2x + (m^2 - 1) = 0.$$

Since the product of the roots of the equation is $(m^2-1)/(m^2+1)$ and one of the roots is -1, we see that the second root is

$$x = \frac{1-m^2}{1+m^2}.$$

By using the equation $y = m(x+1)$, we see that $y = 2m/(1+m^2)$. This means that the point of intersection is

$$P_m := \left(\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2} \right). \tag{3.2}$$

Now we would like to derive a triple of integers (a, b, c) from this pair of rational numbers. Let $m = r/s$ with r, s coprime integers. Then we get

$$P_m = \left(\frac{s^2-r^2}{s^2+r^2}, \frac{2rs}{s^2+r^2} \right).$$

Now we find a primitive Pythagorean triple (u, v, w) such that

$$R(u, v, w) = \left(\frac{s^2-r^2}{s^2+r^2}, \frac{2rs}{s^2+r^2} \right).$$

We need to calculate

$$\gcd(s^2-r^2, s^2+r^2), \quad \gcd(2rs, s^2+r^2).$$

Lemma 3.2. *For coprime integers r, s, define a function*

$$\delta(r, s) = \gcd(2, s^2+r^2).$$

Then

$$\gcd(s^2-r^2, s^2+r^2) = \gcd(2rs, s^2+r^2) = \delta(r, s).$$

Proof. Since $\gcd(r, s) = 1$, we have

$$\gcd(rs, s^2+r^2) = 1.$$

Indeed, if p is a prime number and $p \mid \gcd(rs, s^2+r^2)$, then either $p \mid r$ or $p \mid s$. If $p \mid r$, then since $p \mid s^2+r^2$, we have $p \mid s^2$, and as a result $p \mid s$. So $p \mid r$, $p \mid s$, contradicting the coprimality assumption. As a result

$$\gcd(2rs, s^2+r^2) = \gcd(2, s^2+r^2).$$

Next,

$$\begin{aligned} \gcd(s^2-r^2, s^2+r^2) &= \gcd(s^2+r^2, s^2+r^2+(s^2-r^2)) \\ &= \gcd(s^2+r^2, 2s^2) = \gcd(2, s^2+r^2) = \delta(r, s), \end{aligned}$$

again as $\gcd(s^2, s^2+r^2) = 1$.

Note that

$$\delta(r,s) = \begin{cases} 2 & \text{if } r \equiv s \bmod 2; \\ 1 & \text{otherwise.} \end{cases}$$

It follows from the lemma that

$$P_m = \left(\frac{\frac{s^2-r^2}{\delta(r,s)}}{\frac{s^2+r^2}{\delta(r,s)}}, \frac{\frac{2rs}{\delta(r,s)}}{\frac{s^2+r^2}{\delta(r,s)}} \right),$$

and in this representations the coordinates of P_m are in reduced form. Consequently, if we set

$$(u,v,w) = \left(\frac{s^2-r^2}{\delta(r,s)}, \frac{2sr}{\delta(r,s)}, \frac{s^2+r^2}{\delta(r,s)} \right),$$

then $R(u,v,w) = P_m$. If we do not care about the order, we may write $\{u,v,w\} = \tau(r,s)$, where

$$\tau(s,r) = \left\{ \frac{s^2-r^2}{\delta(r,s)}, \frac{2sr}{\delta(r,s)}, \frac{s^2+r^2}{\delta(r,s)} \right\}.$$

The trouble with this parametrization of Pythagorean triples is that it is not a bijection with the set of coprime integers r, s. For example, if the pairs $(r,s) = (1,2)$ and $(1,3)$ both give the famous Pythagorean triple $3, 4, 5$. In fact, in general, if r, s are both odd, we obtain

$$\{u,v,w\} = \left\{ \frac{s^2-r^2}{2}, sr, \frac{s^2+r^2}{2} \right\}.$$

So the question that we now need to answer is: What happens to the cases where either r or s is even. This has an amusing explanation.

Lemma 3.3. *Let r, s be coprime integers of different parity. Then $r+s$ and $r-s$ are coprime odd numbers and*

$$\tau(s,r) = \tau(s+r, s-r).$$

Proof. An easy check shows that

$$\frac{(s+r)^2 + (s-r)^2}{2} = s^2 + r^2;$$

$$\frac{(s+r)^2 - (s-r)^2}{2} = 2sr;$$

and

$$(s+r)(s-r) = r^2 - s^2.$$

We have proved the following theorem:

Theorem 3.4. *Let u, v, w be the three sides of a primitive integral right triangle. There are coprime integers x, y of different parity such that*

$$\{u,v,w\} = \{x^2-y^2, 2xy, x^2+y^2\}.$$

For example, if $x = 2$, $y = 1$, we obtain 3, 4, 5; if $x = 3$, $y = 2$, we have 5, 12, 13; if $x = 4$, $y = 3$, we find the triple 7, 24, 25.

Remark 3.5. It is important to compare the statement of Theorem 3.4 with Theorem 3.1.

Remark 3.6. The interesting thing about Equation (3.2) is that we do not have to assume that $m \in \mathbb{Q}$. In fact the same computation works over any field, e.g., $\mathbb{R}$, $\mathbb{C}$, or even finite fields. In general care is needed to ensure the denominator $1 + m^2$ is not zero. For fields like $\mathbb{Q}$ and $\mathbb{R}$ this is not an issue, but as soon as we work over a field like $\mathbb{C}$, then $1 + m^2$ can in fact be zero. We will return to this point in Chapters 8 and 14.

3.3 Geometric method to find solutions: Non-Pythagorean examples

It might seem superfluous to use the geometric method of §3.2 to find the solutions of the Pythagorean Equation in light of the much easier methods of §3.1. However, the geometric methods of §3.2 have applications to situations where the elementary methods of §3.1 give little or no information. To demonstrate this method we discuss two examples in this section.

The first example we discuss is Pell's Equation:

$$x^2 - Dy^2 = 1, \tag{3.3}$$

where we assume D is a square-free positive integer. Typically this equation is considered as a Diophantine equation with integral solutions where the solutions are determined using the continued fraction expansion of the quadratic surd $\sqrt{D}$, cf. [33, Ch. 7]. Here we would like to consider this equation over the rational numbers. There are some obvious solutions, namely $(+1, 0)$ and $(-1, 0)$. We will use one of these, say $(-1, 0)$, to find the other rational solutions.

The equation of the straight line passing through $(-1, 0)$ with slope m is

$$y = m(x + 1).$$

We find the points of intersection of this line with the curve with equation $x^2 - Dy^2 = 1$ by inserting the value of y from the equation of the straight line in the equation of the curve. We obtain

$$x^2 - Dm^2(x + 1)^2 = 1.$$

Expanding $(x + 1)^2$ and collecting terms gives

$$(1 - Dm^2)x^2 - 2Dm^2x - (Dm^2 + 1) = 0.$$

Since we know that $x = -1$ is a solution of this equation, we find the other solution to be

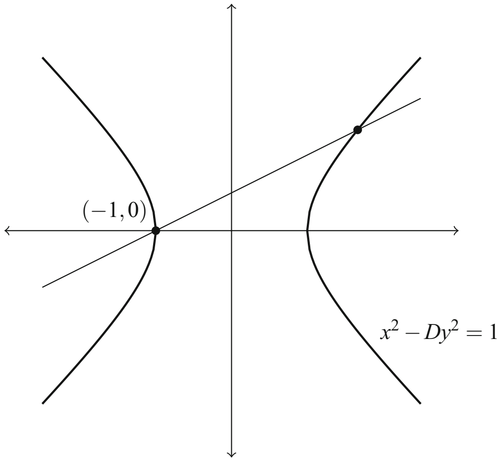

Fig. 3.2 The graph of $x^2 - Dy^2 = 1$. Here we have drawn a straight line with slope m through the point $(-1, 0)$

$$x = \frac{1 + Dm^2}{1 - Dm^2}.$$

With this at hand we find the corresponding y as

$$y = m\left(\frac{1 + Dm^2}{1 - Dm^2} + 1\right) = \frac{2m}{1 - Dm^2}.$$

Consequently, we have proved the following theorem:

Theorem 3.7. *Every solution of Pell's Equation over the rational numbers other than the pair* $(x, y) = (-1, 0)$ *is expressible as*

$$\begin{cases} x = \frac{1+Dm^2}{1-Dm^2}, \\ y = \frac{2m}{1-Dm^2} \end{cases}$$

for some $m \in \mathbb{Q}$.

One can easily find *integral* solutions to the equation

$$X^2 - DY^2 = Z^2 \tag{3.4}$$

by using the above rational parametrization; see Exercise 3.1. As an example, let's consider the case where $D = 3$. Then Theorem 3.7 says that the rational solutions of the equation $x^2 - 3y^2 = 1$ are of the form

$$x = \frac{1 + 3m^2}{1 - 3m^2}, \quad y = \frac{2m}{1 - 3m^2}$$

for $m \in \mathbb{Q}$. If we put $m = 2$, then we get the pair $(x, y) = (-13/11, -4/11)$, from which the solution $(X, Y, Z) = (-13, -4, 11)$ for the equation $X^2 - 3Y^2 = Z^2$ is obtained. If on the other hand we put $m = 1/2$, we obtain $(x, y) = (7, 4)$. This pair gives the solution $(X, Y, Z) = (7, 4, 1)$ of $X^2 - 3Y^2 = Z^2$.

The above method works for any quadratic polynomial. Indeed, suppose $f(x, y)$ is a quadratic polynomial of degree two in the variables x, y with rational coefficients. In the examples we have discussed so far, $f(x, y) = x^2 + y^2 - 1$ in the Pythagorean case, or $f(x, y) = x^2 - Dy^2 - 1$ in the Pell case. Then the graph of $f(x, y) = 0$ either contains infinitely many points with rational coordinates, or none at all. The proof of this fact is identical to our arguments for the examples we have discussed so far.

In our next example, we consider the important case where the degree of the polynomial f is equal to 3. The most general polynomial $f(x, y)$ of degree 3 with rational coefficients can be written as

$$a_1x^3 + a_2y^3 + a_3x^2y + a_4xy^2 + a_5x^2 + a_6y^2 + a_7xy + a_8x + a_9y + a_{10}.$$

Here we assume that the a_i's are rational numbers and at least one of a_1, a_2, a_3, and a_4 is non-zero. Let C be the graph of f. If we try and imitate what we did for quadratic polynomials, we run into trouble. Indeed, suppose (a, b) is a point on the curve C. Then the equation of the line passing through (a, b) with slope m is

$$y = m(x - a) + b.$$

If we insert this expression for y in the equation $f(x, y) = 0$ we obtain a degree 3 equation in x which has three roots. By construction, one of the roots of this equation is $x = a$.. In general there is no reason that the resulting equation should have two more rational solutions. We can see this in an example as follows.

Suppose, for example, that

$$f(x, y) = y^2 + x^3 + 1.$$

Then there is an obvious solution of $(-1, 0)$. The line through this point with slope m has equation

$$y = m(x + 1).$$

We then obtain the equation

$$m^2(x + 1)^2 + x^3 + 1 = 0.$$

Expanding and simplifying give

$$x^3 + m^2x^2 + 2m^2x + (1 + m^2) = 0.$$

Since this equation a priori has a root $x = -1$, the polynomial on the left should be divisible by $(x + 1)$. One easily sees that the polynomial factors as $(x + 1)$ multiplied by

$$x^2 + (m^2 - 1)x + (m^2 + 1). \tag{3.5}$$

This quadratic equation will have rational roots if its discriminant is a rational square t^2, for some $t \in \mathbb{Q}$. We calculate the discriminant as

$$\Delta = (m^2 - 1)^2 - 4(m^2 + 1) = m^4 - 2m^2 + 1 - 4m^2 - 4 = m^4 - 6m^2 - 3.$$

So the equation we need to find rational solutions for is

$$t^2 = m^4 - 6m^2 - 3$$

which is of higher degree than the original equation $y^2 + x^3 + 1 = 0$.

The above discussion suggests the following strategy: Instead of using one rational point on the curve and a rational slope, use two rational points on the curve. Once we have two points, connect the points using a straight line; look at the intersection of the resulting line with the curve. This last point is then a new point with rational coordinates on the curve. We demonstrate this idea with a couple of examples.

Example 3.8. Consider the curve $y^2 = x^3 + 17$. An inspection reveals the points $(-1, 4)$, $(2, 5)$ with rational coordinates on the curve. The equation of the line connecting the points is

$$y = \frac{1}{3}x + \frac{13}{3}.$$

The intersection point of the line with the curve is the point determined by solving the system of equations

$$\begin{cases} y^2 = x^3 + 17, \\ y = \frac{1}{3}x + \frac{13}{3}. \end{cases}$$

To solve, we insert the value of y from the second equation in the first equation to obtain

$$\left(\frac{1}{3}x + \frac{13}{3}\right)^2 = x^3 + 17.$$

Simplifying gives

$$x^3 - \frac{1}{9}x^2 - \frac{26}{9}x - \frac{16}{9} = 0.$$

We already know two of the roots of this equation, namely -1 and 2. Since the product of the three roots of the equation is $16/9$, we find that the third root is

$$x = -\frac{8}{9}.$$

Now we use the equation of the straight line to find y:

$$y = \frac{109}{27}.$$

It is easy to check that the point $(-8/9, 109/27)$ is indeed on the curve. There are in fact infinitely many pairs of rational numbers satisfying $y^2 = x^3 + 17$, but the proof of this fact is beyond the scope of this book.

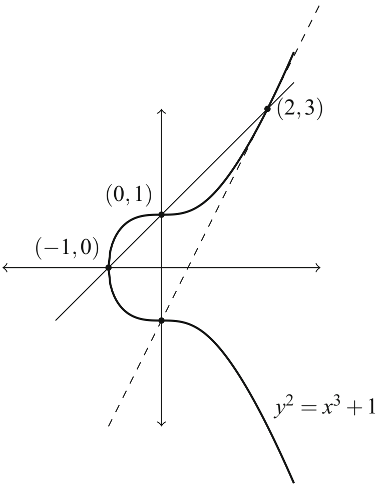

Fig. 3.3 The cubic curve $y^2 = x^3 + 1$ with the colinear points $(-1, 0)$, $(0, 1)$, and $(2, 3)$

What we did in the above example was choosing points A, B on the curve, connecting them, and looking at the point of the intersection of the resulting line with the curve. Now suppose we choose the points A, B very close to each other. As the points get close to each other, the line connecting them approaches the tangent line to the curve at the point obtained from identifying A and B. So one way to obtain rational points on a cubic curve is by starting from a rational point and drawing the tangent line to the curve at that point. The other point of intersection of the tangent line with the curve must then be a rational point. In the next example we show how this idea is used in practice.

Example 3.9. The equation $y^2 = x^3 + 1$ in Figure 3.3 has the obvious solutions $(0, 1)$ and $(-1, 0)$.

The straight line connecting the points is

$$y = x + 1.$$

The intersection of this line with the curve is the point $(2, 3)$. Now that we have a new point, we can draw the tangent line at the point $(2, 3)$ to find more points. By implicit differentiation we have

$$2yy' = 3x^2.$$

Hence the slope of the tangent line at the point $(2, 3)$ is

$$m = 2.$$

The equation of the tangent line is $y = 2x - 1$. This is the dashed line in the figure. The intersection of this line with the graph of $y^2 = x^3 + 1$ is the point $(0, -1)$ which is a new point, though not very interesting. In fact, by using more advanced techniques than what is discussed in this book, one can show that the equation $y^2 = x^3 + 1$ has only finitely many solutions in pairs of rational numbers.

See the Notes at the end of this chapter for more on these cubic curves and the connections to the theory of elliptic curves.

3.4 Application: $X^4 + Y^4 = Z^4$

At some point around 1637, Pierre de Fermat famously declared in the margin of a book that if $n \in \mathbb{N}$ is larger than 2, the Diophantine equation

$$X^n + Y^n = Z^n$$

will not have any solutions in integers X, Y, Z, except for those satisfying $XYZ = 0$. He went on to say that he had an amazing proof of the fact, but that the margin was too small to fit the proof. This claim is now known as Fermat's Last Theorem even though its proof was finally completed by Sir Andrew Wiles, then a professor at Princeton University, in a joint work with Richard Taylor in 1994. Wiles' work was a crowning achievement of modern mathematics which built on works by many, many mathematicians spanning, literally, hundreds of years. Nowadays very few mathematicians believe that Fermat actually had a proof for the general case, neither does anyone hope that one might ever be able to give a reasonably short, elementary proof of the theorem accessible to Fermat. It is, however, possible to prove many special cases of the theorem using elementary methods. Here we present a proof of the special case for $n = 4$ discovered by Fermat. The proof we give uses our knowledge of the solutions of the Pythagorean Equation.

Theorem 3.10 (Fermat). *If the integers X, Y, Z satisfy $X^4 + Y^4 = Z^2$, then $XY = 0$.*

Proof. Suppose our claim is wrong, i.e., there are solutions (X, Y, Z) with $X > 0$, $Y > 0$, and $Z > 0$. Property 2.2 allows us to choose among these solutions the triple (x, y, z) with the smallest possible z. Clearly then $\gcd(x, y, z) = 1$. By Theorem 3.4 there are coprime integers m, n such that

$$\begin{cases} x^2 = 2mn, \\ y^2 = m^2 - n^2, \\ z = m^2 + n^2. \end{cases}$$

The second equation in the list can be rewritten as $n^2 + y^2 = m^2$. Again, Theorem 3.4 tells us that there are coprime integers u, v of different parity such that

$$\begin{cases} n = 2uv, \\ y = u^2 - v^2, \\ m = u^2 + v^2. \end{cases}$$

Next, we have

$$x^2 = 2mn = 4uv(u^2 + v^2).$$

Note that since u, v are coprime and of different parity, the integers u, v, and $u^2 + v^2$ are pairwise coprime; since their product is a square, each of them individually is a square, i.e., there are integers r, s, and t such that

$$\begin{cases} u = r^2, \\ v = s^2, \\ u^2 + v^2 = t^2. \end{cases}$$

Combining these three equations gives

$$r^4 + s^4 = t^2.$$

By construction $r > 0, s > 0, t > 0$. We now observe

$$0 < t \leq t^2 = u^2 + v^2 = m < z.$$

Hence we have found a solution (r, s, t) of the equation $X^4 + Y^4 = Z^2$ with $0 < t < z$. This contradicts our assumption that (x, y, z) was the solution with the smallest possible z. □

The theorem has the following immediate corollary:

Corollary 3.11. *If the integers X, Y, Z satisfy*

$$X^4 + Y^4 = Z^4, \tag{3.6}$$

then $XY = 0$.

This is how the proof of Theorem 3.10 works: Suppose we have some integral solution (X, Y, Z) of the equation $x^4 + y^4 = z^2$ with $XYZ \neq 0$. Since z appears in the equation with even exponent, we conclude that $(X, Y, |Z|)$ will satisfy the equation as well. This means that the equation will then have solutions (X, Y, Z) with $Z \in \mathbb{N}$. Now let S be the set of all such Z's. Since S is assumed to be non-empty, Property 2.2 shows that S must have a smallest element Z_0. The main piece of the proof of the theorem consists of showing that there is another number $Z_1 \in S$ such that $Z_1 < Z_0$, and this is a contradiction as we had assumed that Z_0 was the smallest element of S.

The method used in the proof of Theorem 3.10 is called *infinite descent*. The method of infinite descent relies on the Well-ordering Principle, Property 2.2. As we saw in Theorem 2.3, the Well-ordering Principle is nothing but mathematical induction. Infinite descent was of the most powerful methods in Fermat's arsenal of tools and tricks. We will see some more applications of this method in the exercises. We will also use this method in the proof of Theorem 4.4.

Exercises

3.1 For an integer D, find the integral solutions of Pell's Equation (3.4).

3.2 Find the rational solutions to $x^2 - y^2 = 1$ by writing $x - y = m/n$ and $x + y = n/m$.

3.3 Find every integral solution of the equation

$$a^2 + b^2 + c^2 = d^2.$$

3.4 Prove that the only integral solution to the equation $x^2 + y^2 + z^2 = 2xyz$ is $x = y = z = 0$.

3.5 Find all the rational solutions of $x^2 + y^2 = z^2 + t^2$.

3.6 Show that for all natural numbers n, the equation $x^2 - y^2 = n^3$ is solvable in integers x, y. Determine the number of solutions if n is odd.

3.7 Show that the equation

$$x^2 + (x+1)^2 + (x+2)^2 + (x+3)^2 + (x+4)^2 = y^2$$

has no solutions in integers $x, y \in \mathbb{Z}$.

3.8 Find all the solutions of the equation

$$3(x^2 + y^2) + 2xy = 664$$

in integers x, y.

3.9 Show that for every $t \in \mathbb{Z}$ the triple

$$(x, y, z) = (9t^4, 1 - 9t^3, 3t - 9t^4)$$

satisfies

$$x^3 + y^3 + z^3 = 1.$$

Also verify that for each $t \in \mathbb{Z}$

$$(x, y, z) = (1 + 6t^3, 1 - 6t^3, -6t^3)$$

is a solution of the equation $x^3 + y^3 + z^3 = 2$. Show that the equation $x^3 + y^3 + z^3 = 4$ has no solutions in $\mathbb{Z}$. It is in general not known how to solve equations of the form $x^3 + y^3 + z^3 = n$ with $x, y, z \in \mathbb{Z}$.

3.10 Find all integral right triangles whose hypotenuse is a square.

3.11 Find all right triangles one of whose legs is a square.

3.12 Find all primitive right triangles with square perimeter.

3.13 Show that for every $n \in \mathbb{N}$, there are at least n distinct primitive right triangles which share a leg.

3.14 Show that for every $n \in \mathbb{N}$, there are at least n distinct primitive right triangles which share their hypotenuse.

3.15 Find all integral right triangles whose side lengths form an arithmetic progression.

3.16 Show that for every n there are n points in the plane, not all of which are on a straight line, such that the distance between every two of them is an integer. How about infinitely many points?

3.17 Show that for every Pythagorean triple (u, v, w) we have

$$(uv)^4 + (vw)^4 + (wu)^4 = (w^4 - u^2v^2)^2.$$

Conclude that the equation

$$x^4 + y^4 + z^4 = t^2$$

has infinitely many solutions in integers x, y, z, t such that $\gcd(x, y, z) = 1$.

3.18 Solve the system of Diophantine equations

$$\begin{cases} x^2 + t = u^2, \\ x^2 - t = v^2. \end{cases}$$

3.19 Verify that the points $(1, 0)$ and $(0, 2)$ satisfy the equation

$$y^2 = x^3 - 5x + 4.$$

Use the geometric method of this chapter to find more solutions.

3.20 Verify that the point $(-3, 9)$ satisfies the equation $y^2 = x^3 - 36x$. Use this point to produce more solutions.

3.21 Use *infinite descent* to show that there is no rational number γ such that $\gamma^2 = 2$.

3.22 Show that there are no non-zero integral solutions to the following equations:

a. $2x^4 - 2y^4 = z^2$;
b. $x^4 + 2y^4 = z^2$;
c. $x^4 - y^4 = 2z^2$;
d. $8x^4 - y^4 = z^2$.

3.23 Show that the only solutions to $x^4 + y^4 = 2z^2$ in integers are $z = \pm x^2$ and $|y| = |x|$.

3.24 (✠) Find the number of solutions (x, y, z) in integers of the equation $x^2 - 5y^2 = z^2$ with $|x|, |y|, |z| < 1000$.

3.25 (✠) Find 25 pairs of integers (x, y) such that $x^2 - 2y^2 = 1$. You might want to use Equation (3.7) of the Notes.

3.26 (✠) Find ten pairs of rational numbers (x, y) such that $y^2 = x^3 + 3$.

Notes

Pell's Equation

Traditionally, Pell's Equation is Equation (3.3) with the extra assumption that x, y are integers. The equation

$$x^2 - Dy^2 = -1,$$

too, is called Pell's Equation. Calling any of these equations Pell's Equation is a famous mischaracterization by Euler. Historically these equations were of interest to mathematicians for hundreds of years before Euler and his contemporaries; see, for example, [27, Ch. 2]. This last reference states that in 628 the great Indian mathematician Brahmagupta (598–670 CE) discovered the identity

$$(a^2 - Db^2)(p^2 - Dq^2) = (ap + Dbq)^2 - D(aq + bp)^2.$$

An immediate consequence of this fact is the remarkable statement that if Pell's Equation $x^2 - Dy^2 = \pm 1$ has a non-trivial integral solution, i.e., one where $y \neq 0$, it will have infinitely many integral solutions. In fact, let (x_1, y_1) be the solution of the equation

$$x^2 - Dy^2 = \pm 1,$$

with $x_1, y_1 > 0$, and x_1 the smallest possible. We call (x_1, y_1) the *fundamental solution*. Then, there are two possibilities:

1. If $x_1^2 - Dy_1^2 = +1$, then the equation $x^2 - Dy^2 = -1$ has no solutions. Furthermore, every solution of the equation $x^2 - Dy^2 = +1$ is of the form $(\pm x_N, \pm y_N)$ with

$$x_N + \sqrt{D} y_N = (x_1 + \sqrt{D} y_1)^N \tag{3.7}$$

for some $N \in \mathbb{Z}$.

2. If $x_1^2 - Dy_1^2 = -1$, then the equation $x^2 - Dy^2 = -1$ has solutions $(\pm x_N, \pm y_N)$ determined by Equation (3.7) with $N \in \mathbb{Z}$ odd. The solutions of $x^2 - Dy^2 = +1$ are the pairs $(\pm x_N, \pm y_N)$ with $N \in \mathbb{Z}$ even.

For example when $D = 2$, the fundamental solution to $x^2 - 2y^2 = \pm 1$ is $(1, 1)$ which satisfies $1^2 - 2 \cdot 1^2 = -1$. If $N = 2$, we compute

$$(1 + \sqrt{2})^2 = 3 + 2\sqrt{2},$$

and it is clear that $(3, 2)$ satisfies $3^2 - 2.2^2 = +1$. If $N = 3$,

$$(1 + \sqrt{2})^3 = 7 + 5\sqrt{2},$$

and $7^2 - 2 \cdot 5^2 = -1$.

Because of these observations, finding the solutions of Pell's Equation reduces to the search for the fundamental solution. Note that even though the fundamental solution (x_1, y_1) is the smallest solution of the equation, it does not have to be *small* in any reasonable sense. For example, the smallest solution of $x^2 - 61y^2 = 1$ is $(x, y) = (1766319049, 226153980)$. The most effective way to write down the fundamental solution is via *continued fractions*. This method was originally discovered by the Indian mathematicians Jayadeva (c. 950–$\sim$ 1000 CE) and Bhaskara (c. 1114 –1185 CE) who completed Brahmagupta's method, though they gave no formal proof of this. The formal proof was provided by Lagrange in the 18th century. For a complete history of this subject we refer the reader to Weil's book [57]. For details of this method, see [27, Ch. 3] or [33, Ch. 7], especially §7.6.3.

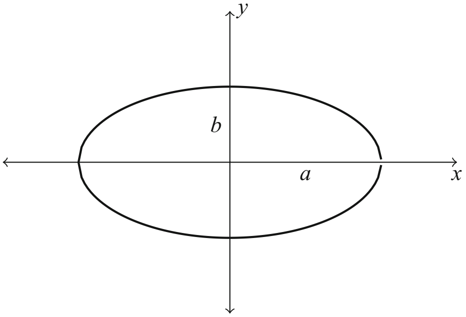

Fig. 3.4 Ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

Elliptic curves

The cubic curves considered in §3.3 are called *elliptic curves*. These are some of the most important objects in all of mathematics, and they have been the subject of intense research for a few hundred years. The genesis of the adjective in the name of these curves goes back to 17th and 18th centuries. Let us briefly explain the connection; see [92] for details and references.

Consider the ellipse with the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

with $a > b$. It is an easy integration exercise to show that the area of the ellipse is equal to πab. Now suppose we want to compute the perimeter of the ellipse.

A parametrization for the ellipse is given by

$$\begin{cases} x = a \sin t \\ y = b \cos t \end{cases} \qquad 0 \leq t \leq 2\pi.$$

By the arc length formula, itself an application of the Pythagorean Theorem, the perimeter ℓ of the ellipse is equal to

$$\begin{aligned} \ell &= \int_0^{2\pi} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \\ &= 4\int_0^{\pi/2} \sqrt{a^2 \cos^2 t + b^2 \sin^2 t}\, dt \\ &= 4a \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 t}\, dt, \end{aligned}$$

with $k^2 = 1 - b^2/a^2$. A change of variables with $u = \sin t$ gives

$$\ell = 4a \int_0^1 \frac{\sqrt{1-k^2u^2}}{\sqrt{1-u^2}}\, du. \tag{3.8}$$

This is a special value of an *elliptic integral of second kind*. In general, *elliptic integrals of the second kind* are defined as follows: For $0 \leq w \leq 1$ we define

$$E(w) = \int_0^w \frac{\sqrt{1-k^2u^2}}{\sqrt{1-u^2}}\, du.$$

Elliptic integrals are in general not expressible in terms of elementary functions. Because of their many applications in mathematical physics these types of integrals attracted a lot of attention starting in the 18th century. It was Abel in the 19th century who realized that the correct object of study is the inverse of the function E. The motivation for this point of view is the $\sin^{-1}$ integral: We know

$$\sin^{-1} w = \int_0^w \frac{du}{\sqrt{1-u^2}},$$

but the more natural function to work with is the inverse function of $\sin^{-1}$, the ubiquitous sine. Going back to Equation (3.8), we make one more change of variable $z = 1 - k^2u^2$ to obtain

$$\ell = 2a \int_\lambda^1 \frac{z}{\sqrt{z(1-z)(z-\lambda)}}\, dz,$$

with $\lambda = 1 - k^2$. Upon setting $z = q + \frac{1+\lambda}{3}$ the integral transforms to

$$\ell = 2a \int_{\frac{2\lambda-1}{3}}^{\frac{2-\lambda}{3}} \frac{q + \frac{1+\lambda}{3}}{\sqrt{-q^3 + \frac{1}{3}(\lambda^2+\lambda-1)q + \frac{1}{27}(2\lambda^3+3\lambda^2-3\lambda-2)}}\, dq.$$

Finally (!), set $q = -\sqrt[3]{4}v$ to get

$$\ell = 2\sqrt[3]{4}a \int_{-\frac{2-\lambda}{3\sqrt[3]{4}}}^{-\frac{2\lambda-1}{3\sqrt[4]{3}}} \frac{-\sqrt[3]{4}v + \frac{1+\lambda}{3}}{\sqrt{4v^3 - \frac{\sqrt[3]{4}}{3}(\lambda^2+\lambda-1)v + \frac{1}{27}(2\lambda^3+3\lambda^2-3\lambda-2)}}\, dv.$$

Let

$$g_2 = \frac{\sqrt[3]{4}}{3}(\lambda^2+\lambda-1),$$

and

$$g_3 = -\frac{1}{27}(2\lambda^3+3\lambda^2-3\lambda-2).$$

Karl Weierstrass defined a function $\wp(u)$ with the property that

$$u = \int_{\wp(u)}^{\infty} \frac{dv}{\sqrt{4z^3 - g_2z - g_3}}.$$

So clearly ℓ is related to the function $\wp$. A remarkable property of the $\wp$-function is that it satisfies the functional equation

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3,$$

i.e., the point $(\wp(z), \wp'(z))$ lies on the curve

$$y^2 = 4x^3 - g_2x - g_3. \tag{3.9}$$

In fact, the points $(\wp(z), \wp'(z))$ give a full parametrization for the points with complex coordinates on the curve. Furthermore,

$$\wp(u+v) = -\wp(u) - \wp(v) + \frac{1}{4}\left(\frac{\wp'(u) - \wp'(v)}{\wp(u) - \wp(v)}\right)^2,$$

and

$$\wp(-u) = \wp(u), \quad \wp'(-u) = -\wp'(u).$$

These formulae have an interesting interpretation for the points on the curve. We define a group law $\oplus$ on the set of points of the curve as follows: For a point A on the curve, define $-A$ to be the reflection of A with respect to the x-axis; for three points A, B, C, we say $A \oplus B = C$ if A, B, and $-C$ are colinear; and O, the identity point, is the point at infinity in the direction of the y axis, i.e. $A \oplus (-A)$ for any point A.

The work we did in §3.3 shows that if $g_2, g_3 \in \mathbb{Q}$, then the $\oplus$ of any two points with coordinates in $\mathbb{Q}$ will be again a point with coordinates in $\mathbb{Q}$. Clearly, also, for a point A with rational coordinates, $-A$ will have rational coordinates. This means that the collection of points on the curve with rational coordinates forms a group. It is a truly surprising fact that, by a theorem of Mordell, this group is finitely generated. We refer the reader to [48, Ch. 3] or [47, Ch. VIII] for details.

Fermat's Last Theorem

Fermat's Last Theorem is an esoteric statement with no applications as such, but despite its obscurity it has given rise to an enormous amount of mathematics. Edwards [19] presents algebraic number theory as it was originally motivated by false proofs of Fermat's Last Theorem. "The Proof", a NOVA documentary [114] on Wiles' work, is an excellent account of the last steps toward the proof. Charles Mozzochi's endearing photo essay "The Fermat Diary" [36] is a photo album of all those whose works contributed to the proof of the theorem in the last fifty years. Finally, even though it is written for experts, Sir Andrew Wiles' introduction to his masterful paper in the Annals of Mathematics [110] is a delight to read.

The abc Conjecture

The *abc* Conjecture is an easy to state conjecture with many surprising consequences in number theory. The conjecture was formulated by D. W. Masser and J. Oesterlé in the 80's. This is the statement:

Conjecture 3.12 (The abc Conjecture). If $\varepsilon > 0$, then the number of triples (a, b, c) of coprime natural numbers such that $c = a + b$ and

$$c > \left(\prod_{p|abc} p\right)^{1+\varepsilon}$$

is finite.

The conjecture could also be formulated as follows: For every $\varepsilon > 0$, there is a constant $\kappa_\varepsilon > 0$ such that for every triple (a, b, c) of coprime natural numbers satisfying $c = a + b$ we have

$$c \le \kappa_\varepsilon \left(\prod_{p|abc} p\right)^{1+\varepsilon} .$$

To see a quick application, let us apply the *abc* Conjecture to Fermat's Last Theorem. Suppose we have three coprime natural numbers x, y, z such that $x^n + y^n = z^n$. If $\varepsilon > 0$ is given, then applying the *abc* Conjecture with $a = x^n$, $b = y^n$, and $c = z^n$ shows that with the exception of finitely many choices of x, y, z we have

$$z^n \le \left(\prod_{p|x^n y^n z^n} p\right)^{1+\varepsilon}$$

Next, $p \mid x^n y^n z^n$ if and only if $p \mid xyz$. So we have

$$\prod_{p|x^n y^n z^n} p = \prod_{p|xyz} p.$$

Now we observe that if n is a natural number, $\prod_{p|n} p \le n$. Using this observation we have

$$\prod_{p|xyz} p \le xyz < z^3.$$

In the last step we have used the fact that $x < z$ and $y < z$. Putting everything together, we conclude that except for finitely many choices of x, y, z we have

$$z^n < (z^3)^{1+\varepsilon} = z^{3(1+\varepsilon)}.$$

This implies that $n < 3(1 + \varepsilon)$. Since the choice of ε is arbitrary, this means $n \le 3$. What we have proved is the following:

Corollary 3.13 (Assuming *abc* Conjecture). *For each $n > 3$, Fermat's equation $x^n + y^n = z^n$ has at most finitely many solutions in coprime natural numbers x, y, z.*

The statement of the *abc* Conjecture is ineffective. This means that for a fixed $\varepsilon > 0$ the conjecture does not provide any estimate for the number or the size of triples (a, b, c) satisfying the conditions of the conjecture. There are several explicit versions of the *abc* Conjecture in literature. Here we state one of these explicit conjectures which is due to Alan Baker [63].

To state Baker's *abc* Conjecture we need some notation. For a natural number n, we set rad (n) to be the product of the prime divisors of n, i.e.,

$$\text{rad}\,(n) = \prod_{p|n} p.$$

For example, rad $(1) = 1$, rad $(12) = 2 \times 3 = 6$ and rad $(25) = 5$. We also let $\omega(n) = \sum_{p|n} 1$, i.e., the number of prime divisors of n. With this definition we have $\omega(1) = 0$, $\omega(12) = 2$, $\omega(25) = 1$. Using this notation, the original *abc* Conjecture asserts that for $\varepsilon > 0$, there is $\kappa_\varepsilon > 0$ such that for a triple (a, b, c) of coprime natural numbers, we have

$$c < \kappa_\varepsilon (\text{rad}\,(abc))^{1+\varepsilon}.$$

Conjecture 3.14 (Baker's abc Conjecture). Let (a, b, c) be a triple of coprime natural numbers such that $c = a + b$. Let $N = \text{rad}\,(abc)$ and $r = \omega(N)$. Then

$$c < \frac{6}{5} N \frac{(\log N)^r}{r!}.$$

We leave it to the reader to verify that Baker's *abc* Conjecture in fact implies the *abc* Conjecture. The papers by Granville and Tucker [77] and Waldschmidt [107] outline various applications of the *abc* Conjecture. In April of 2012, Shinichi Mochizuki of Kyoto University announced a proof of the *abc* Conjecture occupying hundreds of pages. At the time of this writing it is still not known if Mochizuki's proof is correct, and for that reason the *abc* Conjecture is still considered open.