

PYTHAGOREAN TRIPLES

PETE L. CLARK

1. PARAMETERIZATION OF PYTHAGOREAN TRIPLES

1.1. Introduction to Pythagorean triples.

By a **Pythagorean triple** we mean an ordered triple $(x, y, z) \in \mathbb{Z}^3$ such that

$$x^2 + y^2 = z^2.$$

The name comes from elementary geometry: if a right triangle has leg lengths x and y and hypotenuse length z , then $x^2 + y^2 = z^2$. Of course here x, y, z are positive real numbers. For most integer values of x and y , the integer $x^2 + y^2$ will not be a perfect square, so the positive real number $\sqrt{x^2 + y^2}$ will be irrational: e.g. $x = y = 1 \implies z = \sqrt{2}$. However, a few integer solutions to $x^2 + y^2 = z^2$ are familiar from high school algebra (and the SATs): e.g. $(3, 4, 5)$, $(5, 12, 13)$.

Remark: As soon as we have one solution, like $(3, 4, 5)$, we can find infinitely many more, however in a somewhat unsatisfying way. Namely, if (x, y, z) is a Pythagorean triple and a is any integer, then also (ax, ay, az) is a Pythagorean triple:

$$(ax)^2 + (ay)^2 = a^2(x^2 + y^2) = a^2z^2 = (az)^2.$$

This property of invariance under scaling is a characteristic feature of solutions (x_1, \dots, x_n) to **homogeneous polynomials** $P(t_1, \dots, t_n)$ in n -variables. We recall what this means: a monomial is an expression of the form $ct_1^{a_1} \cdots t_n^{a_n}$ (for a nonzero constant c), and the degree of the monomial is defined to be $a_1 + \dots + a_n$, i.e., the sum of the exponents. A polynomial is said to be **homogeneous of degree d** if each of its monomial terms has degree d , and simply **homogeneous** if it is homogeneous of some degree d . For instance, the polynomial $P(x, y, z) = x^2 + y^2 - z^2$ is homogeneous of degree 2, and indeed for any N the **Fermat polynomial**

$$P_N(x, y, z) = x^N + y^N - z^N$$

is homogeneous of degree N . Moreover, every (nonconstant) homogeneous polynomial $P(t_1, \dots, t_n)$ has zero constant term, hence $P(0, \dots, 0) = 0$. So $(0, \dots, 0)$ is a solution to any homogeneous polynomial, called the **trivial solution**.

Coming back to Pythagorean triples, these considerations show that for all $a \in \mathbb{Z}$, $(3a, 4a, 5a)$ is a Pythagorean triple (again, familiar to anyone who has studied for the SATs). For many purposes it is convenient to regard these rescaled solutions as being equivalent to each other. To this end we define a Pythagorean triple (a, b, c) to be **primitive** if $\gcd(a, b, c) = 1$. Then every nontrivial triple (a, b, c) is a positive integer multiple of a unique primitive triple, namely $(\frac{a}{d}, \frac{b}{d}, \frac{c}{d})$ where $d = \gcd(a, b, c)$.

Thanks to Katelyn Andrews and Laura Nunley for pointing out typos.

Our goal is to find all primitive Pythagorean triples. There are many ways to do so. We prefer the following method, both for its simplicity and because it motivates the study of not just integral but **rational solutions** of polynomial equations.

Namely, consider the algebraic curve $x^2 + y^2 = 1$ in \mathbb{R}^2 : i.e., the unit circle. Why? Well, suppose (a, b, c) is a nontrivial Pythagorean triple, so $a^2 + b^2 = c^2$ with $c \neq 0$ (if $c = 0$, then $a^2 + b^2 = 0 \implies a = b = 0$). So we may divide through by c , getting

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

Thus $(\frac{a}{c}, \frac{b}{c})$ is a rational point on the unit circle. Moreover, the process can be essentially reversed: suppose that $(r, s) \in \mathbb{Q}^2$ is such that $r^2 + s^2 = 1$. Then, writing $r = \frac{a}{c}$ and $s = \frac{b}{d}$ (so $cd \neq 0$), we have

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

Multiplying through by $(cd)^2$, we get

$$(da)^2 + (bc)^2 = (bd)^2,$$

so that (da, bc, bd) is a nontrivial Pythagorean triple. If we start with a primitive Pythagorean triple (a, b, c) , pass to the rational solution $(\frac{a}{c}, \frac{b}{c})$ and then clear denominators using the above formula, we get (ca, cb, c^2) . This is not the primitive triple that we started with, but it is simply a rescaling: no big deal. At the end we will find the correct scaling that gives primitive triples on the nose.

1.2. Rational parameterization of the unit circle.

Fix any one rational point $P_\bullet = (x_\bullet, y_\bullet)$ on the unit circle. The argument that we are about to make works for any choice of P_\bullet – e.g. $(\frac{3}{5}, \frac{4}{5})$ – but let me pass along the wisdom of hindsight: the computations will be especially simple and clean if we take $P_\bullet = (-1, 0)$. So let us do so.

Now suppose $P = (x_P, y_P)$ is any other rational point on the unit circle. Then there is a unique line ℓ joining P_\bullet to P , which of course has rational coefficients:

$$\ell : y - y_P = \frac{y_P - y_\bullet}{x_P - x_\bullet} (x - x_P).$$

In particular, the slope of this line

$$m_P = \frac{y_P - y_\bullet}{x_P - x_\bullet},$$

is a rational number. This already places a limitation on the rational solutions, since “most” lines passing through the fixed point P_\bullet have irrational slope. More interesting is the converse: for any $m \in \mathbb{Q}$, let

$$\ell_m : y = (y - y_\bullet) = m(x - x_\bullet) = m(x + 1),$$

be the line passing through $P_\bullet = (-1, 0)$ with slope m . We claim that this line intersects the unit circle in precisely one additional point P_m , and that this point P_m also has rational coordinates. That is, we claim that the rational points on the unit circle are precisely the point $P_\bullet = (-1, 0)$ together with the set of points P_m

as m ranges through the rational numbers.

Why is this so? With a bit of thought, we can argue for this “in advance”. Briefly, we plug the linear equation ℓ_m into the quadratic $x^2 + y^2 = 1$ thereby getting a quadratic equation in x with rational coefficients. Because we know that this equation has at least one rational solution – namely -1 , the coordinate of P_\bullet – the other solution must be rational as well, as follows from contemplation of the quadratic formula. On the other hand, such forethought is not really necessary in this case, because we want to find the solutions explicitly anyway. In other words, let’s do it!

We have the system of equations

$$(1) \quad x^2 + y^2 = 1$$

$$(2) \quad y = m(x + 1).$$

Substituting (2) into (1) gives

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

or

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

Applying the quadratic formula, we get

$$x = \frac{-2m^2 \pm \sqrt{4m^4 - 4(1 + m^2)(m^2 - 1)}}{2(1 + m^2)}.$$

Under the radical sign we have

$$4m^4 - 4(m^2 + 1)(m^2 - 1) = 4(m^4 - (m^4 - 1)) = 4,$$

so that “luckily”¹ $\sqrt{4m^2 - 4(1 + m^2)(m^2 - 1)} = 2$, and

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

Notice that by taking the minus sign, we get the solution $x = \frac{-m^2 - 1}{1 + m^2} = -1$. That’s great, because -1 is the x -coordinate of P_\bullet , so that it had better be a solution. The other solution is the one we really want:

$$x_m = \frac{1 - m^2}{1 + m^2},$$

and then we get

$$y_m = m(1 + x_m) = m \left(1 + \frac{1 - m^2}{1 + m^2} \right) = \frac{2m}{m^2 + 1},$$

so that finally

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

¹Not really, of course: see the last paragraph above.

This is exactly what we wanted. Before returning to the problem of Pythagorean triples, however, let us make one further observation:

$$\lim_{m \rightarrow \pm\infty} P_m = \left(\lim_{m \rightarrow \pm\infty} \frac{1-m^2}{1+m^2}, \lim_{m \rightarrow \pm\infty} \frac{2m}{1+m^2} \right) = (-1, 0) = P_\bullet.$$

The geometric interpretation of this is simple: the tangent line to the unit circle at $(-1, 0)$ is vertical, so as the slope of the line ℓ_m approaches either $+\infty$ or $-\infty$, the second intersection point P_m approaches P_\bullet and the secant lines approach the tangent line. So in fact it is true that the rational points on the unit circle correspond precisely to the set of all rational lines through P_\bullet : here we get P_\bullet itself as the double intersection point of the tangent line. Thus, instead of P_\bullet , a more appropriate name would be P_∞ , although we do not insist on this in the sequel.

1.3. Scaling to get primitive solutions.

We wish to explicitly write down all primitive Pythagorean triples (a, b, c) . As above, this is accomplished up to scaling by clearing denominators in the general rational solution $P_m = (x_m, y_m)$. Namely, put $m = \frac{u}{v}$ with $\gcd(u, v) = 1$, so

$$P_m = \left(\frac{1-u^2/v^2}{1+u^2/v^2}, \frac{2u/v}{1+u^2/v^2} \right) = \left(\frac{v^2-u^2}{v^2+u^2}, \frac{2uv}{v^2+u^2} \right).$$

Thus, multiplying through by v^2+u^2 , we get a family of integral solutions

$$(v^2-u^2, 2uv, v^2+u^2).$$

Are these solutions primitive? In other words, is $\gcd(v^2-u^2, 2uv, v^2+u^2) = 1$?

Suppose that an odd prime p divides v^2-u^2 and v^2+u^2 . Then p also divides $(v^2-u^2) + (v^2+u^2) = 2v^2$ and $(v^2+u^2) - (v^2-u^2) = 2u^2$. Since p is odd, we get $p \mid u^2$ and $p \mid v^2$ which implies $p \mid u$ and $p \mid v$, contradiction. Similarly, if $4 \mid v^2-u^2$ and $4 \mid v^2+u^2$, then $4 \mid 2v^2$ and $4 \mid 2u^2$, so $2 \mid v^2$ and $2 \mid u^2$, so 2 divides both u and v . Thus $\gcd(v^2-u^2, 2uv, v^2+u^2)$ is either 1 or 2.

Case 1: v and u have opposite parity. Then v^2-u^2 is odd, so the gcd is 1. Notice that in this case, the first coordinate v^2-u^2 is odd and the second coordinate $2uv$ is even, so this can't be the complete list of all primitive Pythagorean triples: that set is symmetric under interchanging x and y !

Case 2: u and v are both odd. Then $v^2-u^2, 2uv, v^2+u^2$ are all even, so the gcd is 2. In this case $(\frac{v^2-u^2}{2}, uv, \frac{v^2+u^2}{2})$ is the primitive integral solution we seek.

This is the answer,² but let's touch it up a bit. If $x = 2k+1$ is odd, then $x^2 = 4k^2 + 4k + 1 \equiv 1 \pmod{4}$. Thus, if u and v are both odd, not only is v^2-u^2 even, it is congruent to $v^2-u^2 \equiv 1-1 = 0 \pmod{4}$, so $\frac{v^2-u^2}{2}$ is even and uv is odd. Thus all the primitive triples arising in Case 2 are obtained by switching the first and second coordinates of a primitive triple in Case 1. To sum up:

Theorem 1. (*Classification of Pythagorean Triples*)

- The rational solutions to $x^2 + y^2 = 1$ are $\{(-1, 0) \cup \left(\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2}\right) \mid m \in \mathbb{Q}\}$.
- $(0, 0, 0)$ is a Pythagorean triple, called **trivial**. Every nontrivial Pythagorean

²Note that there is no Case 3: we assumed $\gcd(u, v) = 1$, so they can't both be even.

triple is of the form (da, db, dc) for some $d \in \mathbb{Z}^+$, where (a, b, c) is a Pythagorean triple with $\gcd(a, b, c) = 1$, called **primitive**.

c) In every primitive Pythagorean triple (a, b, c) , exactly one of a and b are even integers. Every primitive triple with a odd is of the form $(v^2 - u^2, 2uv, v^2 + u^2)$ where $u, v \in \mathbb{Z}$ are relatively prime integers of opposite parity. Conversely, all such pairs u, v yield a primitive Pythagorean triple with first coordinate odd.

2. AN APPLICATION: FERMAT'S LAST THEOREM FOR $N = 4$

In this section we will prove Fermat's Last Theorem for $N = 4$.

Theorem 2. $X^4 + Y^4 = Z^4$ has no solutions with $X, Y, Z \in \mathbb{Z} \setminus \{0\}$.

This theorem was, in fact, first proved by Pierre de Fermat, circa 1650. Our proof follows Fermat's own. In fact, what Fermat proved was a stronger result:

Theorem 3. $X^4 + Y^4 = Z^2$ has no solutions with $X, Y, Z \in \mathbb{Z} \setminus \{0\}$.

Theorem 2 follows readily from Theorem 3, as we now explain: suppose that there exist $x, y, z \in \mathbb{Z} \setminus \{0\}$ such that $x^4 + y^4 = z^4$. By replacing $x \mapsto |x|$ and so forth, we may assume that x, y, z are all positive. Then, since $Z^4 = (Z^2)^2$, the triple (x, y, z^2) is a positive integer solution to $X^4 + Y^4 = Z^2$, contradicting Theorem 3.

Proof of Theorem 3:

Step 1: Let (x, y, z) be any positive integral solution to $X^4 + Y^4 = Z^2$. We claim that there is a positive integral solution (x', y', z') with $\gcd(x', y') = 1$ and $z' \leq z$.

Indeed, if x and y are not relatively prime, they are both divisible by some prime number p . Then $p^4 \mid X^4 + Y^4 = Z^2$, so $p^2 \mid Z$. Therefore $\frac{x}{p}, \frac{y}{p}, \frac{z}{p^2} \in \mathbb{Z}^+$ and

$$\left(\frac{x}{p}\right)^4 + \left(\frac{y}{p}\right)^4 = \frac{1}{p^4}(x^4 + y^4) = \frac{1}{p^4}(z^2) = \left(\frac{z}{p^2}\right)^2,$$

so $(\frac{x}{p}, \frac{y}{p}, \frac{z}{p^2})$ is another positive integral solution, with z -coordinate smaller than the one we started with. Therefore the process can be repeated, and since the z -coordinate gets strictly smaller each time, it must eventually terminate with a solution (x', y', z') as in the statement.

Step 2: Given a positive integral solution (x, y, z) to $X^4 + Y^4 = Z^2$ with $\gcd(x, y) = 1$, we will produce another positive integral solution (u, v, w) with $w < z$.

First, we may assume without loss of generality that x is odd and y is even. Since $\gcd(x, y) = 1$, they cannot both be even; if instead x is even and y is odd, then we can switch x and y ; so we must check is that x and y cannot both be odd. Considering $x^4 + y^4 = z^2$ modulo 4, we find $2 \equiv z^2 \pmod{4}$, which is easily seen to be impossible: $0^2 \equiv 2^2 \equiv 0 \pmod{4}$, $1^2 \equiv 3^2 \equiv 1 \pmod{4}$.

Now we use our knowledge of Pythagorean triples: since $(x^2)^2 + (y^2)^2 = z^2$ and $\gcd(x, y) = 1$, (x^2, y^2, z) is a primitive Pythagorean triple with first coordinate odd. So by Theorem 1 there exist coprime integers m and n of opposite parity such that

$$(3) \quad \begin{aligned} x^2 &= m^2 - n^2 \\ y^2 &= 2mn \\ z &= m^2 + n^2. \end{aligned}$$

Now rewrite (3) as $n^2 + x^2 = m^2$. Since $\gcd(m, n) = 1$, this is again a primitive Pythagorean triple. Moreover, since x is odd, n must be even. So we can use our parameterization again (!) to write

$$\begin{aligned}x &= r^2 - s^2, \\n &= 2rs, \\m &= r^2 + s^2,\end{aligned}$$

for coprime integers r, s of opposite parity. Now observe

$$m \left(\frac{n}{2}\right) = \frac{2mn}{4} = \frac{y^2}{4} = \left(\frac{y}{2}\right)^2.$$

Since m and $\frac{n}{2}$ are coprime integers whose product is a perfect square, they must both be perfect squares. Similarly,

$$rs = \frac{2rs}{2} = \frac{n}{2} = \square,$$

so r and s must both be squares. Let us put $r = u^2, s = v^2, m = w^2$, and substitute these quantities into $m = r^2 + s^2$ to get

$$u^4 + v^4 = w^2.$$

Here $w \geq 1$, so

$$w \leq w^4 < w^4 + v^4 = m^2 + n^2 = z,$$

so that as promised, we found a new positive integral solution (u, v, w) with $w < z$.

Step 3: Steps 1 and 2 together lead to a contradiction, as follows: if we have any positive integral solution (x, y, z) to $X^4 + Y^4 = Z^2$, then by Step 1 we have one (x', y', z') with $z' \leq z$ with $\gcd(x', y') = 1$. Then by Step 2 we have another positive integral solution (u, v, w) with $w < z' \leq z$. Then by Step 1 we have another positive integral solution (u', v', w') with $w' \leq w < z' \leq z$ with $\gcd(u', v') = 1$, and then by Step 2 we get another solution whose final coordinate is strictly smaller than w . And so on. In other words, the assumption that there are any positive integer solutions at all leads to the construction of an infinite sequence of positive integer solutions (x_n, y_n, z_n) with $z_{n+1} < z_n$ for all n . But that's impossible: there are no infinite strictly decreasing sequences of positive integers. Contradiction!

This technique is known as **infinite descent**. It has been developed over the years, and the modern **theory of descent** is still a mainstay of Diophantine geometry.

3. RATIONAL POINTS ON CONICS

The method of drawing lines that we used to find all rational points on the unit circle has further applicability. Namely, we consider an arbitrary **conic curve**

$$(4) \quad aX^2 + bY^2 = cZ^2,$$

for $a, b, c \in \mathbb{Q} \setminus \{0\}$.

Remark: More generally, one calls a plane conic any curve given by an equation

$$aX^2 + bXY + cXZ + dY^2 + eYZ + fZ^2 = 0.$$

for $a, b, c, d, e, f \in \mathbb{Q}$, not all zero. But as one learns in linear algebra, by making a linear change of variables, new coordinates can be found in which the equation is

diagonal, i.e., in the form (4), and one can easily relate integral/rational points on one curve to those on the other. So by considering only diagonalized conics, we are not losing out on any generality.

Now, as in the case $a = b = c = 1$, we have a bijective correspondence between primitive integral solutions to $aX^2 + bY^2 = cZ^2$ and rational points on

$$(5) \quad ax^2 + by^2 = c.$$

If we can find any one rational point $P_\bullet = (x_\bullet, y_\bullet)$ on (5) then our previous geometric method works: by taking the set of all lines through P_\bullet with rational slope, together with the line $x = x_\bullet$ with slope ∞ and intersecting with the conic (5), we get all rational solutions.

In the exercises the reader is invited to try this in certain cases where there are “obvious” rational solutions. For instance, if $a = c$ then an obvious rational solution is $(1, 0)$. The reader is asked to carry this out in a particular case – and also to investigate the structure of the primitive integral solutions – in the exercises.

But there need not be any rational solutions at all! An easy example of this is

$$x^2 + y^2 = -1,$$

where indeed there are clearly no \mathbb{R} -solutions. But this is not the only obstruction. Consider for instance

$$3x^2 + 3y^2 = 1,$$

whose real solutions form a circle of radius $\frac{1}{\sqrt{3}}$. We claim that there are however no rational points on this circle. Equivalently, there are no integral solutions to $3X^2 + 3Y^2 = Z^2$ with $\gcd(x, y, z) = 1$.

For suppose there is such a primitive integral solution (x, y, z) . Then, since $3 \mid 3x^2 + 3y^2 = z^2$, we have $3 \mid z$. So we may put $z = 3z'$, getting $3x^2 + 3y^2 = 9(z')^2$, or

$$x^2 + y^2 = 3z'^2.$$

Now reducing mod 3, we get

$$x^2 + y^2 \equiv 0 \pmod{3}.$$

Since the squares mod 3 are 0 and 1, the only solution mod 3 is $x \equiv y \equiv 0 \pmod{3}$, but this means $3 \mid x$, $3 \mid y$, so that the solution (x, y, z) is not primitive after all: 3 is a common divisor.

Thus we also have techniques for showing that the conic $aX^2 + bY^2 = cZ^2$ *does not* have nontrivial integral solutions: namely, **sign considerations** and **congruence considerations**, the latter being a fancy name for working modulo n to get a contradiction, as we did above. On the other hand we have a technique for showing that the conic *does* have nontrivial integral solutions: simple brute force searching. For any fixed a, b, c , there either are or are not nontrivial \mathbb{Z} -solutions, so clearly at most one of these techniques can succeed. The big question is: is it possible for them both to fail?? I.e., are there $a, b, c \in \mathbb{Z} \setminus \{0\}$ such that $aX^2 + bY^2 = cZ^2$ has no nontrivial integral solutions but it is not possible to show this by sign considerations and congruence obstructions?

The answer is provided by the following elegant theorem of Legendre.

Theorem 4. *Let a, b, c be nonzero integers, squarefree, relatively prime in pairs, and neither all positive nor all negative. Then*

$$(6) \quad ax^2 + by^2 + cz^2 = 0$$

has a solution in nonzero integers (x, y, z) iff all of the following hold:

(i) There exists $u \in \mathbb{Z}$ such that $-ab \equiv u^2 \pmod{c}$.

(ii) There exists $v \in \mathbb{Z}$ such that $-bc \equiv v^2 \pmod{a}$.

(iii) there exists $w \in \mathbb{Z}$ such that $-ca \equiv w^2 \pmod{b}$.

Since we can compute all of the squares modulo any integer n by a direct, finite calculation, one can use Legendre's Theorem to write a computer program that, upon being inputted (a, b, c) , determines whether or not (6) has nontrivial solutions.

But there is more to Theorem 6 than this: it also gives a negative answer to our big question: in every case in which there is no nontrivial integer solution, we can detect this by considerations of sign and/or congruence.

To see this, let us now prove one implication of Theorem 4, namely that if there exists a nonzero integral solution (x, y, z) , then conditions (i) through (iii) do hold. The point is that, exactly as in the above example, the conditions arise naturally upon reduction modulo a suitable integer n . Indeed:

Suppose $x, y, z \in \mathbb{Z}$ with $\gcd(x, y, z) = 1$ and $ax^2 + by^2 + cz^2 = 0$. Observe that if any prime p divides both c and x , then also $p \mid by^2$, and since $\gcd(b, c) = 1$, $p \mid y$. Thus $\gcd(c, x) = 1 \iff \gcd(c, y) = 1$. So if $\gcd(c, x) > 1$, there exists a prime p dividing x , y and c , so $p^2 \mid ax^2 + by^2 = -cz^2$. Since c is squarefree, $p \mid z^2$, hence $p \mid z$, contradicting $\gcd(x, y, z) = 1$. So we may assume $\gcd(c, x) = \gcd(c, y) = 1$, which means that a, b, x, y are all units in $\mathbb{Z}/c\mathbb{Z}$. So reducing (6) modulo c gives:

$$ax^2 + by^2 \equiv 0 \pmod{c},$$

or

$$\left(\frac{x}{y}\right)^2 \equiv \frac{-b}{a}.$$

Multiplying through by a^2 we get

$$\left(\frac{ax}{y}\right)^2 \equiv -ab \pmod{c},$$

so that $-ab$ is a square modulo c : this is precisely condition (i). Similarly (indeed, symmetrically), we get (ii) by reducing modulo a and (iii) by reducing modulo b .

Of course the meat of Legendre's theorem is the converse implication, i.e., that conditions (i) through (iii) are sufficient to ensure the existence of a nontrivial integer solution. This argument requires tools that we do not yet possess – especially, the **quadratic reciprocity law** – so we will have to come back to it later.

Once we know whether there are any integral solutions, we can search by brute force until we find one. The following result of Holzer puts an explicit upper bound on our search:

Theorem 5. *If the equation $ax^2 + by^2 + cz^2 = 0$ has any solutions in nonzero integers, it has such a solution (x, y, z) with $|x| \leq \sqrt{|bc|}$, $|y| \leq \sqrt{|ac|}$, $|z| \leq \sqrt{|ab|}$.*

A proof of Theorem 4 will be given later on, after we have discussed quadratic reciprocity. As for Theorem 5, we choose not to give a proof, but rather we alert the interested reader to a very short and elementary proof due to L.J. Mordell.³

Thus the study of homogeneous quadratic equations over \mathbb{Z} (or, what comes to the same, over \mathbb{Q}) is admirably complete. The same cannot be said for polynomial equations of higher degree, as we will soon see.

4. TRANSITION TO ALGEBRAIC NUMBER THEORY

Let us not leave behind the homogeneous quadratic equation just yet. Suppose that p is an odd prime. Are there any nontrivial \mathbb{Z} -solutions to

$$(7) \quad pX^2 + pY^2 = Z^2?$$

When $p = 3$, congruence considerations showed that there are no nontrivial solutions. When $p = 5$, a little searching turns up $(1, 2, 5)$, i.e., $5(1)^2 + 5(2)^2 = 5^2$. When $p = 7$, going modulo 7 one can again see that there are no nontrivial solutions.

What does Legendre's Theorem have to say? Nothing directly, because in Theorem 4 the coefficients are assumed pairwise relatively prime. But with only a little retouching we can get an equation to which Legendre's Theorem applies. Namely, for x, y, z satisfying (7), we must have $p \mid z^2 \implies p \mid z$. Writing $z = pz'$, we get

$$pX^2 + pY^2 = p^2(Z')^2,$$

or

$$(8) \quad X^2 + Y^2 - pZ'^2 = 0.$$

Legendre's Theorem does apply to (8) and gives the following result.

Corollary 6. *For any odd prime p , the equation $pX^2 + pY^2 = Z^2$ has nontrivial integer solutions if and only if -1 is a square modulo p .*

We don't yet have a general criterion that tells us "all at once" for which primes the congruence $-1 \equiv x^2 \pmod{p}$ is solvable. However, we can easily compile the following data: among odd primes $p < 100$, -1 is a square modulo p for

$$p = 5, 13, 17, 29, 37, 41, 53, 61, 73, 89, 97$$

whereas -1 is **not** a square modulo p for

$$p = 3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, 79, 83.$$

Detecting patterns in data can be quite challenging, but here it seems pretty clear: all the primes in the first list are $1 \pmod{4}$, and all the primes in the second list are $-1 \pmod{4}$. This was already known to Fermat.

Lemma 7. *(Fermat's Lemma) For an odd prime p , the following are equivalent:*

- (i) *There exists $x \in \mathbb{Z}$ such that $x^2 \equiv -1 \pmod{p}$ (i.e., -1 is a square mod p).*
- (ii) *There exists $k \in \mathbb{Z}$ such that $p = 4k + 1$ (i.e., p is $1 \pmod{4}$).*

³On the magnitude of the integer solutions of the equation $ax^2 + by^2 + cz^2 = 0$, J. Number Theory 1 (1969), 1-3.

Let us first prove that (i) \implies (ii). For this, if $x \in \mathbb{Z}/p\mathbb{Z}$ is such that $x^2 \equiv -1 \pmod{p}$, the order of x in the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$ is 4. By [Corollary 10, Algebra Handout 1], this implies $4 \mid \#(\mathbb{Z}/p\mathbb{Z})^\times = p - 1$, so $p \equiv 1 \pmod{4}$.

The other implication is more difficult. From a modern perspective, it is natural to use the fact that the group of units $(\mathbb{Z}/p\mathbb{Z})^\times$ of $\mathbb{Z}/p\mathbb{Z}$ is cyclic. But probably we should give at least a taste of more traditional arguments, which are also simple and appealing. Especially, this gives us an excuse to discuss the following result.

Proposition 8. (*Wilson's Theorem*) *If p is a prime, then $(p - 1)! \equiv -1 \pmod{p}$.*

Proof: We must show that if we multiply all the nonzero elements of the finite field $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, we get -1 . This is certainly true for $p = 2$, so we may assume that p is an odd prime. Consider the map $\iota : \mathbb{F}_p^\times \rightarrow \mathbb{F}_p^\times$ given by $x \mapsto x^{-1}$ (inversion). This map is a bijection: every element has a unique inverse. Moreover, the “fixed points” of this map are those elements $x = \iota(x)$ which are equal to their own inverse: $x = x^{-1}$; this is equivalent to the equation $x^2 = 1$. There are two obvious solutions: namely $x = \pm 1$ (which are different since $p \neq 2$). On the other hand, the polynomial $x^2 - 1$ can have at most two roots in the field \mathbb{F}_p . Thus every element different from ± 1 in \mathbb{F}_p^\times is distinct from its inverse, so multiplying the nonzero elements of \mathbb{F}_p^\times gives $1 \cdot -1 \cdot (x_1 \cdot x_1^{-1}) \cdots (x_{\frac{p-3}{2}} \cdot x_{\frac{p-3}{2}}^{-1})$ – i.e., all the elements apart from ± 1 can be paired with their distinct and unique multiplicative inverse, qed!

Now we use Wilson's Theorem to prove Fermat's Lemma. We showed that for a prime p , the product of the first $p - 1$ integers is -1 modulo p . But this can be immediately generalized as follows: suppose that S is any **reduced residue system** modulo p , i.e., a set of $p - 1$ integers such that

$$\bar{S} = \{\bar{s} \mid s \in S\} = \mathbb{F}_p^\times,$$

i.e., a set that represents every nonzero residue class modulo p exactly once. Then the conclusion of Wilson's theorem must be true for S just as it is for the particular choice $S_0 = \{1, \dots, p - 1\}$: namely, $\prod_{i \in S} i \equiv -1 \pmod{p}$. Consider the set

$$S = \left\{ \frac{-(p-1)}{2}, \frac{-(p-1)}{2} + 1, \dots, -1, 1, \dots, \frac{(p-1)}{2} \right\}.$$

Observe that S has $p - 1$ elements, and the difference between the least element and the greatest element is precisely $p - 1 < p$, so that no two of them can differ by a multiple of p : they are too close together for that. Thus S is a reduced residue system modulo p , so that Wilson's theorem says that

$$-1 \equiv \prod_{i \in S} i \pmod{p}.$$

On the other hand, by “folding in half” – i.e., by multiplying each element in S with its partner of opposite sign – we directly compute that

$$\prod_{i \in S} i = (-1)^{\frac{p-1}{2}} \cdot \left(\left(\frac{p-1}{2} \right)! \right)^2.$$

If now $p \equiv 1 \pmod{4}$, then $\frac{p-1}{2}$ is even, so that $(-1)^{\frac{p-1}{2}} = 1$. Putting everything together we get $-1 \equiv x^2 \pmod{p}$, i.e., $p \mid x^2 + 1$. This proves Fermat's Lemma.