2 CHAPTER Not the Last of Fermat

2.1 INTRODUCTION: FERMAT'S LAST THEOREM (FLT)

A lot has been written about Fermat's Last Theorem since its proof was announced in 1993.[1] What we write here will undoubtedly not be the last that's written on the subject.

As it took Andrew Wiles, a brilliant mathematician, seven years to put together the deepest mathematical proof of the last century, there's no way that we are going to be able to take you step by step through his proof. In fact, relatively few mathematicians could actually do that anyway, however strong their mathematical background. What we will try to do, instead, is to give you the background to the problem and an idea of the strategy adopted by those involved in solving it — and of the human perspiration expended in the final fall of the theorem. What follows then is a middle course between Singh's best-selling book [8] and the more mathematical, and more sharply focused, treatments to be found in Cox [1], Gouvêa [3], Mazur [5], and Ribenboim [7]. If you become interested in the topic of this chapter, we recommend you first read [8] and then, if you have some undergraduate mathematics under your belt, try [1], [3], [5], or [7]. For more background material there is also [9]. So you should think of this chapter as not the end but rather the beginning of what there is to know about

[1] The announcement was premature! There was a gap in the argument, but a complete proof did appear in 1994.

Fermat's Last Theorem. Certainly we are making no claim whatsoever that our exposition is complete.[2]

2.2 SOMETHING COMPLETELY DIFFERENT

Freda keeps some beetles and some spiders. Altogether her pets have 48 legs. How many spiders does she have?

If you play around with this problem for a while, you might suspect that it contains insufficient information. After all, if Freda has b beetles and s spiders we can only get one equation,

$$6b + 8s = 48$$

This equation holds simply because beetles have 6 legs and spiders have 8. But we have one equation that has *two* things we don't know: s and b. Normally in problems like this we need another equation. However, nobody thought to give us one. So what do we do now?

One step that might simplify things a little is to divide through by the common factor 2. Is

$$3b + 4s = 24$$

any better? Well, we want to find s, so rearrange the equation to give

$$4s = 24 - 3b$$

And then you might see that the RHS of the last equation has a factor 3:

$$4s = 3(8 - b)$$

So what? Remember that b and s are *positive integers*.

So $4s$ must also be divisible by 3. That surely means that s is divisible by 3. OK, then $s = 3k$ for some integer k. What should we do next? Perhaps we'll make some progress if we replace s by $3k$ in the equation:

$$4(3k) = 3(8 - b)$$

so

$$4k = 8 - b$$

Now it's clear that the RHS is no greater than 8, so $k = 0$, 1, or 2. Hang on! If $k = 0$, there are no spiders, and we're told that Freda has "some spiders." And if $k = 2$, then $b = 0$, but she's got "some beetles." This must mean then that $k = 1$, so $s = 3k = 3$. Freda has three spiders.

[2]Michael Rosen, in his favorable review of [7] (*Notices of the AMS*, **47**, No. 4 (2000), 474) wrote, "It is somewhat sad that no one expects any longer that an elementary proof of FLT will ever emerge."

• • • BREAK 1

Try this one. At the Post Office, Dennis spent exactly \$2 on stamps. He bought some 4¢ stamps, ten times as many 2¢ stamps and made up the balance with 10¢ stamps. How many stamps of each denomination did he buy?

Let's try one more. The other day José cashed a check at the bank. The teller accidentally interchanged the dollars and cents values. When later, José could only get 49¢ for his 1963 Cadillac he thought he was having a bad day. But then he realized that he now had twice as much cash as he had written the check out for. What was the real value of the check?

Doing this the traditional way would probably mean assuming that the original check was for x dollars and y cents. So the teller had given José y dollars and x cents. After selling the Cadillac he had y dollars and $(x + 49)$ cents. But this turned out to be twice the value of the original check — exactly $2x$ dollars and $2y$ cents. This gives us an "equation"

$$y \text{ dollars} + (x + 49) \text{ cents} = 2x \text{ dollars} + 2y \text{ cents}$$

Probably we won't get anywhere with this until we change it all into cents. At that point we get a genuine linear equation

$$100y + (x + 49) = 200x + 2y$$

Once again we have one equation in two unknowns; but once again we know that x and y are positive integers. How can we solve our equation? Probably in much the same way as we did with the insects. But let's tidy it up first. The equation then becomes

$$98y + 49 = 199x$$

There seems to be a factor of 49 on the left, so we rewrite the equation as

$$49(2y + 1) = 199x$$

Since 199 and 49 have no factors in common and 49 is a factor of the LHS of this equation, then x must be divisible by 49. So let $x = 49k$. We then get

$$2y + 1 = 199k$$

Did that help? What do we know about y? It can't be bigger than 99, because it was the cents amount on José's check. Now, since $y \leq 99$

$$199k = 2y + 1 \leq 198 + 1 = 199$$

So $k = 0$ or 1. If k were zero, then y would be negative, which is absurd. (The teller would surely have noticed a check made out for minus half a cent!) This means that $k = 1$ and $x = 49$. What's more, $2y + 1 = 199$, so $y = 99$, and José's check was for \$49.99.

• • • BREAK 2

Check the last answer just in case we've made a mistake.

The point of this section, is to show that not all the information you need is given explicitly in a problem or contained in the equation(s) you obtain from the data. More than that, the fact that some problems are about integers, and even positive integers, can be as helpful as having another equation. When you think about it, we couldn't have solved $6b + 8s = 48$ if spiders didn't come in whole numbers. And we would have been all at sea with José's problem if he had written out his check for non-integer values of dollars and cents (and if there were more than 100 cents to the dollar).

2.3 DIOPHANTUS

It turns out that we know rather little about the Greek mathematician Diophantus. We certainly don't know where he was born or when he was born. We might know how old he was when he died. Supposedly, the following was carved on his tomb.

> God granted him to be a boy for the sixth part of his life, and adding a twelfth part to this, He clothed his cheeks with down; He lit him the light of wedlock after a seventh part, and five years after his marriage He granted him a son. Alas! Late-born wretched child; after attaining the measure of half his father's full life, chill Fate took him. After consoling his grief by this science of numbers for four years he ended his life.

As we were saying then, he lived to a ripe old age for someone who flourished about 250 A.D. (but he may have been living as early as 150 A.D. or as late as 360 A.D.). His fame comes from the six extant books of the thirteen-volume set he wrote called *Arithmetica*. This was a treatise that he compiled while he was in Alexandria, the center of intellectual life in the Mediterranean from about 350 B.C. until about 640 A.D. Diophantus' *Arithmetica* was to number theory what Euclid's *Elements* was to geometry. While he was in Alexandria, Diophantus collected and invented a range of number-theoretical problems. He was particularly interested in problems

whose solutions were rationals. However, problems involving integers are now known as ***Diophantine problems***. Similarly equations whose solutions are required to be whole numbers are called ***Diophantine equations***.

So the two problems of Section 1 (three if you include the one in BREAK 1) are Diophantine problems. The equations

$$6b + 8s = 100 \quad \text{and} \quad 98y + 49 = 199x$$

are Diophantine equations.

The library at Alexandria, which held hundreds of thousands of books containing all the knowledge that had been accumulated by the Greeks, was finally destroyed. The destruction began at the hands of Christians in 389 A.D. They were out to destroy all pagan monuments, and the library was housed in what was once an Egyptian temple. What books survived the Christians were largely destroyed by Moslems in 642 A.D. So it may be somewhat surprising that even six volumes of Diophantus' magnum opus survived.

2.4 ENTER PIERRE DE FERMAT

It's not clear that you would necessarily have gotten along very well with Pierre de Fermat. If you were an English mathematician, it's almost certain that you wouldn't! Pierre de Fermat was born into a rich family in the southwest of France in 1601. His family's fortune allowed him a good education, and he entered the civil service. His job was to make a preliminary assessment of petitions to the King of France. If someone wanted to petition the king, they had to go through Fermat. If he wasn't convinced of the merit of the case, then the petition went no further. Because of his role in society, Fermat had few social contacts. It was felt that, if a petition was put forward by one of his associates, Fermat (or indeed any other of the councillors at the Chamber of Petitions) might be swayed by his friendship to support it. Consequently, Fermat led a very solitary life. But this did give him the opportunity to engage in his lifelong interest of mathematics. In fact, he was so good at mathematics that he has been called the Prince of Amateur Mathematicians.

Today, most mathematicians are eager to publish their research results in journals that are readily available to their peers. In seventeenth-century Europe, though, mathematicians were more secretive. There was a tendency to throw out problems as a challenge. Mark you, these were problems that the poser already knew how to solve.

Fermat engaged in this practice and was particularly keen to embarrass the English mathematicians. At one stage he discovered that a particular square and a particular cube had only one number between them. What's more, he found that this was the only instance where a square and a cube differed by 2. Fermat gained particular satisfaction from the fact that the English mathematicians had to admit that they were unable to solve his square and cube problem.

• • • BREAK 3

Can *you* find one solution in integers of the equation $y^3 = x^2 + 2$?

It's worth remarking here that in 1918 the English mathematician L.J. Mordell was able to show that there are only a finite number of integer solutions to

$$y^3 = x^2 + k$$

for *any* integer k, positive or negative. But perhaps a lag of 280 years or so is somewhat excessive for the English to claim a victory, even if k is much more general than 2.

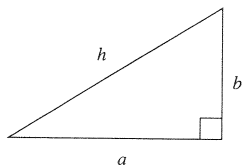
Somehow, Fermat came across a copy of Diophantus' *Arithmetica*. What he had was a translation into Latin, which, for some fortunate but unknown reason, had particularly large margins. As Fermat went through the book he frequently made comments in those margins. These may well have been lost to posterity had it not been for the fact that his son, Clément-Samuel, put together all Fermat's notes, letters, and marginal jottings and published them in a special edition of the *Arithmetica*. This volume contained many challenging problems. Eventually, only one of these remained unsolved. It was the final challenge. Because Fermat said that he had a proof, it was called Fermat's Last Theorem. However, given the fact that it remained unsolved for over 350 years, there's a very good chance that Fermat did not have a complete proof of the "theorem." Really we should have called it Fermat's Last Conjecture. But it's too late to rename it now. So what is Fermat's Last Theorem all about?

2.5 FLASHBACK TO PYTHAGORAS

Pythagoras is well known for several things. In fact we know rather more about him than we do about Diophantus. Pythagoras lived in the sixth century B.C. and wandered around the Eastern Mediterranean area until he set up his scholastic community in southern Italy. That community is

known for realizing the links between harmonics and the ratios of lengths in a lyre string. They are also supposed to have sacrificed 100 oxen when Pythagoras' theorem was proved. But no one knows who actually proved the result. Anything that Pythagoras and his followers discovered was considered the property of the community.

Pythagoras' Theorem *In a right-angled triangle, the square on the hypotenuse is equal to the sum of the squares on the other two sides.*



That is, $h^2 = a^2 + b^2$.

One important thing to note here is that Pythagoras did actually prove that this result was true *for all right-angled triangles*. A thousand years before Pythagoras, the Chinese and the Babylonians knew specific values of a, b, and h — for specific right-angled triangles. Pythagoras' contribution was proving the theorem for *any* right-angled triangle.

• • • BREAK 4

(1) Give a proof of Pythagoras' Theorem.
(2) Give another proof.

Pythagoras had a lot in common with Diophantus. Pythagoras very much preferred dealing with integers and fractions.

Pythagoras and his disciples were able to find an infinite number of right-angled triangles whose sides had integer lengths. You may remember a few of the following:

$$3^2 + 4^2 = 5^2$$
$$5^2 + 12^2 = 13^2$$
$$7^2 + 24^2 = 25^2$$
$$9^2 + 40^2 = 41^2$$

You may even be able to see a pattern here.

Three integers that can be the sides of a right-angled triangle are known as ***Pythagorean triples*** and are sometimes represented by the notation (a, b, h). We've seen that $(3, 4, 5)$, $(5, 12, 13)$, and so on, are Pythagorean triples. Now, it turns out that $(u^2 - v^2, 2uv, u^2 + v^2)$, for u, v integers with $u > v$, are Pythagorean triples. So, by using different values of u and v, we can get a number of Pythagorean triples. In fact, we can get an infinite number, as we shall soon see.

• • • BREAK 5

Show that $(u^2 - v^2, 2uv, u^2 + v^2)$ is indeed a Pythagorean triple.

Before we get started on the following arguments, let's agree to use the abbreviation PT for Pythagorean triple. Now, we will say that a PT (a, b, h) is *primitive* if a, b, h are pairwise coprime. We now prove two key results.

Theorem 1 *Every PT is a multiple of a primitive PT. Moreover, if*

$$(a, b, h) = \lambda(a_1, b_1, h_1)$$

with (a_1, b_1, h_1) *primitive, then*

$$\lambda = \gcd(a, b) = \gcd(a, h) = \gcd(b, h) = \gcd(a, b, h)$$

Thus (a_1, b_1, h_1) *is primitive if and only if* $\gcd(a_1, b_1, h_1) = 1$.

Proof Let $\lambda = \gcd(a, b)$. Then $\lambda | a$, $\lambda | b$, so $\lambda^2 | a^2$, $\lambda^2 | b^2$, and thus $\lambda^2 | h^2$. But if $\lambda^2 | h^2$, then certainly $\lambda | h$. It follows that

$$\gcd(a, b) = \gcd(a, b, h)$$

Similarly, we can prove that

$$\gcd(a, h) = \gcd(a, b, h)$$

and

$$\gcd(b, h) = \gcd(a, b, h) = \lambda,$$

say. Set

$$a = \lambda a_1, \quad b = \lambda b_1, \quad h = \lambda h_1$$

Then $\gcd(a_1, b_1, h_1) = 1$, so (a_1, b_1, h_1) is certainly a primitive PT. Moreover, if (a, b, h) is primitive, then

$$\gcd(a, b) = \gcd(a, h) = \gcd(b, h) = 1,$$

so $\gcd(a, b, h) = 1$. This completes the proof. □

Now let (a, b, h) be a primitive PT. We claim that one of a, b is odd and the other even, and that h is odd. For we certainly can't have a, b, h all even, since (a, b, h) is primitive; and we can't have a, b both odd, since $a^2 \equiv 1 \pmod 4$ and $b^2 \equiv 1 \pmod 4$ implies $h^2 \equiv 2 \pmod 4$, which is clearly impossible. Suppose, then, that a is odd, b is even, and h is odd.

Theorem 2 *Let* (a, b, h) *be a primitive PT with a odd, b even, h odd. There are then unique integers* u, v *such that* $u > v$ *and*

$$a = u^2 - v^2, \quad b = 2uv, \quad h = u^2 + v^2$$

moreover, u, v *are coprime and of opposite parity.*[3]

Proof We have $a^2 + b^2 = h^2$, so

$$\left(\frac{b}{2}\right)^2 = \left(\frac{h+a}{2}\right)\left(\frac{h-a}{2}\right) \tag{1}$$

but $\frac{h+a}{2}$, $\frac{h-a}{2}$ are coprime, for any factor of $\frac{h+a}{2}$ and $\frac{h-a}{2}$ is a factor of h and of a (why?). But (a, b, h) is primitive, so $\gcd(a, h) = 1$. Thus it follows from (1) that each of $\frac{h+a}{2}$, $\frac{h-a}{2}$ is a perfect square, say

$$\frac{h+a}{2} = u^2, \quad \frac{h-a}{2} = v^2,$$

u, v coprime. But then $h = u^2 + v^2$, $a = u^2 - v^2$, and $(\frac{b}{2})^2 = u^2 v^2$, so $b = 2uv$. Since a is odd, u and v must have opposite parity. Finally, we prove uniqueness. For if we also have $h = u_1^2 + v_1^2$, $a = u_1^2 - v_1^2$, $b = 2u_1 v_1$, then

$$u^2 + v^2 = u_1^2 + v_1^2, \quad u^2 - v^2 = u_1^2 - v_1^2,$$

showing that $u^2 = u_1^2$, $v^2 = v_1^2$, so $u = u_1$, $v = v_1$. □

Notice that we can extend the uniqueness statement to arbitrary PTs. Thus, if (a, b, h) is a PT, it is uniquely expressible as

$$(a, b, h) = \lambda(a_1, b_1, h_1)$$

where (a_1, b_1, h_1) is primitive, since λ is determined as $\gcd(a, b, h)$; hence (a, b, h) is uniquely expressible as

$$\lambda(u^2 - v^2, 2uv, u^2 + v^2) \tag{2}$$

where u, v are coprime of opposite parity, and we assume a_1 odd.

[3]This means that one is odd and the other is even.

• • • **BREAK 6**

(1) Which of the following are primitive PTs and which are not? For those that are, find the values of u and v that yield the given PT.

$(3, 4, 5)$; $(5, 12, 13)$; $(231, 108, 255)$; $(99, 20, 101)$.

(2) Write those that are not primitive PTs as a product $\lambda(a, b, h)$ as in Theorem 1, and then find the corresponding u and v.

2.6 SCRIBBLES IN MARGINS

Apparently, Fermat was working on Diophantus' *Arithmetica* in about 1637. He was fascinated by Pythagoras' Theorem and Pythagorean triples. While working in this area of Diophantus' book, Fermat was inspired to write (in Latin)

> To divide a cube into two cubes, a fourth power into two fourth powers, and in general any power above the second into two powers of the same denomination, is impossible. Of this I have assuredly found a marvelous proof, but this margin is too narrow to contain it.

What Fermat had first done was to do what mathematicians continually try to do. He had attempted to generalize. Pythagorean triples are positive-integer solutions to the equation

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

Fermat had asked the obvious mathematical question: Do the Diophantine equations

$$x^3 + y^3 = z^3,$$
$$x^4 + y^4 = z^4,$$
$$x^5 + y^5 = z^5,$$

etc. have any positive-integer solutions? He had worked on them and decided that they did not. He then thought that he could prove it.

Fermat's Last Theorem *For $n \geq 3$ an integer, there are no positive-integer solutions of*

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

But did he really have a proof? We think the answer has to be "most likely not." There are at least two compelling reasons for this. The first is

that Fermat had made other claims of this nature that were doubtful. For instance, the problem of the cube and square that differ by 2, which Fermat used to torture the English, appears not to have been satisfactorily solved by Fermat. (This may be little consolation to the poor English mathematicians of the seventeenth century.)

The second reason is that, whatever proof he thought he had, it would have been elementary from a modern standpoint. Consequently, it would be extremely surprising that no one in the last 300 years had been able to find Fermat's proof.

What seems to be quite possible is that Fermat made a mistake. He thought he had a proof, but there was an error. Mathematicians do make errors. As you will see in Chapter 5, Kempe made an error when he thought he had proved the Four Color Theorem. As you will see in Section 9 of this chapter, Wiles, who ultimately put the final nail in the coffin of Fermat's Last Theorem, made an error in his first "proof" which was made public in 1993.

It is unlikely, however, that Fermat was attempting to mislead. There is no evidence of his doing anything like that on any other occasion. And he seems to have mentioned only the cases $n = 3$ and $n = 4$ in correspondence with other mathematicians. It is perhaps likely that he meant to delete his marginal comment but forgot to do so. What's more, if he had a conjecture rather than a theorem he would probably have admitted it. For instance, he stated once that he thought that $2^{2^n} + 1$ was always a prime number for every positive integer n. But he added that while he was pretty sure that this was the case, he didn't know how to prove it. (See Chapter 4 for a counterexample.)

On balance, we probably have to believe in Fermat's integrity. It is reasonable to believe that he thought he had a proof of FLT but that somehow he had made an error and his proof was false.

2.7 $n = 4$

The case $n = 3$, that is, the fact that $x^3 + y^3 = z^3$ has no integer solutions, was known to Arabian mathematicians some 700 years before Fermat scribbled his cryptic margin message. However, they were unable to prove it to the rigorous level required by present day mathematics. Fermat thought he had a proof, but it is not likely that he did. Surprisingly, then, the value $n = 4$ was the first case of Fermat's Last Theorem to be solved, though perhaps "solved" is not quite the right word. Fermat himself put down

enough details in 1640 that we can reconstruct a complete proof. This proof was not complete by today's standards. However, it did introduce a nice new method. The cunning idea that Fermat used for his "proof" of the case $n = 4$ (and which was completed in detail by Leibniz in 1678) was the ***method of infinite descent***. The concept here is to assume that you have a solution of the equation and show that you can produce a smaller solution. Hence you can keep getting smaller and smaller solutions. This is obviously not possible when you are dealing with positive integers, so the method of infinite descent will show that there are no solutions. Another way of thinking of this is to suppose that you have the *smallest* solution. By showing a smaller solution you have a contradiction.

Now we can consider the equation $x^4 + y^4 = z^4$. Just to be perverse, though, we will apply the method of infinite descent to $x^4 + y^4 = z^2$. Let's think about this for a moment. How does the logic go? Suppose $x^4 + y^4 = z^4$ has a solution in integers $x = p$, $y = q$, $z = r$. Then $p^4 + q^4 = r^4 = (r^2)^2$. So a solution to $x^4 + y^4 = z^4$ will give a solution to $x^4 + y^4 = z^2$ (via $p^4 + q^4 = (r^2)^2$). Logically, then, if we can show that $x^4 + y^4 = z^2$ has no solutions in positive integers, then the last sentence tells us that neither does $x^4 + y^4 = z^4$.

If we then start with $x^4 + y^4 = z^2$, we can immediately relate this to Pythagoras. So $(x^2)^2 + (y^2)^2 = z^2$. We now know that we can concentrate on primitive solutions and hence suppose that

$$x^2 = u^2 - v^2, \quad y^2 = 2uv, \quad \text{and} \quad z = u^2 + v^2,$$

where u and v have no common factor and they are not both odd. Suppose now that u is even and v is odd. Then $u^2 = 4s$ and $v^2 = 4t + 1$ for some s and t. Hence $x^2 = 4(s - t) - 1 = 4(s - t - 1) + 3$. But squares can only have remainders of 0 or 1 when divided by 4. Hence u is odd and v is even. We now have

$$x^2 + v^2 = u^2$$

We then have, since (x, v, u) must be a primitive PT (remember that $\gcd(u, v) = 1$),

$$x = p^2 - q^2, \quad v = 2pq, \quad \text{and} \quad u = p^2 + q^2$$

where p, q have no factors in common and p, q are not both odd. Recall that $y^2 = 2uv$. So now we have

$$y^2 = 4pq(p^2 + q^2)$$

But p and q have no common factors, so neither do p and $p^2 + q^2$, or q and $p^2 + q^2$. This means that p, q, $p^2 + q^2$ are all squares. Hence

$$p = r^2, \quad q = s^2, \quad p^2 + q^2 = t^2, \quad \text{and} \quad \gcd(r, s) = 1.$$

Now here's the coup de grâce! This shows that

$$(r^2)^2 + (s^2)^2 = t^2$$

or

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

Now, we can measure the "size" of a primitive solution of $x^4 + y^4 = z^2$ by the size of z. If we could show that $t < z$, we'd have a primitive solution r, s, t smaller than the original x, y, z. So how is t related to z?

Now

$$z = u^2 + v^2 > u^2 = (p^2 + q^2)^2 = t^4.$$

Then $z > t^4$. So $t < \sqrt[4]{z} < z$, since $z > 1$, obviously. A solution x, y, z of $x^4 + y^4 = z^2$ has led us to a smaller solution r, s, t. Infinite descent comes into play. So Fermat's Last Theorem is true for $n = 4$.

• • • BREAK 7

Why is it that squares have only remainders of 0 and 1 on division by 4?

It's worth thinking about where we are right now. It may seem that we have only covered one of an infinite number of cases of Fermat's Last Theorem, the case $n = 4$. In fact, though, we have settled an infinite number of cases. Why is this so? Have a look at $x^8 + y^8 = z^8$. Suppose r, s, t is a solution to this. Then

$$r^8 + s^8 = t^8$$

and

$$\left(r^2\right)^4 + \left(s^2\right)^4 = \left(t^2\right)^4$$

A solution to $x^8 + y^8 = z^8$ would then give us a solution to $x^4 + y^4 = z^4$. We have just shown that the last equation has no integer solutions. So $x^8 + y^8 = z^8$ has no integer solutions. It's the same little trick that we used in going from knowing that $x^4 + y^4 = z^2$ has no solutions to knowing that $x^4 + y^4 = z^4$ has no solutions.

As a result of all that, then, we are now able to assert confidently that $x^{12} + y^{12} = z^{12}$ has no integer solutions, that $x^{16} + y^{16} = z^{16}$ has no integer

solutions, that.... Fine, so we now know that, for any positive integer m, $x^{4m} + y^{4m} = z^{4m}$ has no integer solutions.

This same argument now puts us in a position where we only have to show that Fermat's Last Theorem is true when n is an odd prime. If we could prove that $x^p + y^p = z^p$ has no integer solutions for all odd primes p, then the truth of Fermat's Last Theorem for any composite number n would follow. This is because, if $n \geq 3$, then either n is a power of 2 divisible by 4 or $n = pq$ for some odd prime p. Since $x^4 + y^4 = z^4$ has no integer solutions, we have already settled the first case. On the other hand, if $n = pq$, then $x^{pq} + y^{pq} = z^{pq}$ gives $(x^q)^p + (y^q)^p = (z^q)^p$. The usual argument now reduces the cases we have to consider for n. We thus have to settle FLT only for n an odd prime.

2.8 EULER ENTERS THE FRAY

Mathematics is a hard taskmaster. We have seen how excited mathematicians get when they prove a theorem — remember the story of Pythagoras and the 100 oxen. We have noted the incompleteness of Fermat's attempts at producing proofs. But mathematics has standards. The full proof must be foolproof.

Although Fermat thought that he had settled the case $n = 3$ by the same infinite descent approach that he had used with the case $n = 4$, there is no evidence of this. In fact, it is generally accepted that the first person to prove FLT for $n = 3$ was the Swiss mathematician Leonhard Euler.

Euler is known for his contributions to many areas of mathematics. You will see his work on the Königsberg bridges problem in Section 5.4. Then there is the result that $a^{\phi(n)} \equiv 1 \pmod{n}$, where ϕ is Euler's totient function[4] (see Chapter 2 of [4]). And there is his polyhedral formula $V - E + F = 2$, which links the number of vertices, edges, and faces of a polyhedron homeomorphic to a sphere. Euler did much, much more. He was extremely prolific.

In 1753, Euler was able to settle the case $n = 3$ of Fermat's Last Theorem. This was the first development in Fermat's Last Theorem in over 100 years. However, there was a step in the proof that happened to be correct for the situation where Euler applied it, but it was a step that would not work in general. Strictly speaking, Euler should have justified this step. Now the general approach to Fermat's Last Theorem seemed to be to try to knock off a case at a time. This was a bit like proving that any map with

[4]In fact, $\phi(n)$ is the number of positive integers m such that $m \leq n$ and m is prime to n.

10 countries is 4-colorable, and then that a map with 11 is 4-colorable, and so on (see Chapter 5). While everyone knew that this case-by-case approach was never going to solve either problem, there was always the hope that, after enough cases had been dealt with, someone would see a general approach or find a counterexample.

One person who tried to move away from this case-by-case strategy was one of the outstanding pre-twentieth-century women mathematicians, Sophie Germain. Her attack lay in looking at a special class P of primes, namely, those primes p for which $2p + 1$ is also a prime. This class obviously includes 5 and 11 but not 7 or 13. For primes of this class, Germain was able to show that if there is a solution of $x^p + y^p = z^p$, then one of x, y, or z would be a multiple of p. This result could then be used to restrict the set of possible solutions.

As a result of the high interest in the problem generated partly by Germain's new idea, the French Academy established a series of prizes for a solution to Fermat's problem. This was not to be the last such prize.

In 1825, as a result of Germain's work, two famous mathematicians, Peter Gustav Lejeune Dirichlet and Adrien-Marie Legendre, working independently, managed to settle the case $n = 5$.

But still the method used by Fermat and Euler was popular. Gabriel Lamé thought that he would be able to use this approach to solve Fermat's Last Theorem not just for some specific cases, but for all odd primes. This was in 1847. And it was the German mathematician Ernst Kummer who first was able to point out that the approach would not work in general, and second was able to show that it was valid for a large class of primes called the regular primes (for a definition of a regular prime, see [1]).

What Lamé (and Euler) had done was to make a very simple oversight. He had assumed that, in the realms in which he had been working, unique factorization occurred. This is a perfectly natural assumption. In ordinary arithmetic this always holds. For instance, $1050 = 2 \times 3 \times 5^2 \times 7$, and this is the only way to factorize 1050 expressing the result as a product of primes (irreducibles). Taking a simpler example, $6 = 2 \times 3$, the factorization is unique. It can only be done in one way among the positive integers, assuming that one ignores the order of the factors.

But the way that the proofs had been developing required calculations to take place among the complex numbers. A complex number is one of the form $a + bi$, where $i = \sqrt{-1}$. For instance, $3 + 4i$ and $2 - 3i$ are two complex numbers. If you want to do arithmetic with them you just treat

the i as an algebraic quantity except that, if you ever get i^2, you convert it to -1. So

$$(3+4i)+(2-3i)=5+i$$
$$(3+4i)-(2-3i)=1+7i$$

and

$$\begin{aligned}(3+4i)(2-3i) &= 6+8i-9i-12i^2\\ &= 6-i-12(-1)\\ &= 6-i+12\\ &= 18-i\end{aligned}$$

Furthermore,

$$\left(1+\sqrt{5}i\right)\left(1-\sqrt{5}i\right)=1-5i^2=6$$

Ah! $(1+\sqrt{5}i)(1-\sqrt{5}i)=2\times 3$, where each of the four (complex) numbers is irreducible in the set of numbers $a+b\sqrt{5}i$, with a, b ordinary integers. So, using the complex numbers that were relevant for the Euler–Fermat proof of FLT, it was possible for numbers to be factorized in more than one way. It is possible for 6 to be factorized as 2×3 or as $(1+\sqrt{5}i)(1-\sqrt{5}i)$.

This problem of unique factorization had actually been a problem for Fermat, though he hadn't realized it. Let's go back to the equation $x^3=y^2+2$, which Fermat claimed had a unique solution in positive integers. His method of proving this was to note that

$$x^3=y^2+2=\left(y+\sqrt{2}i\right)\left(y-\sqrt{2}i\right)$$

Now, since the left-hand side of the equation is a cube and since the two factors on the right have no factors in common, then, given unique factorization, these factors must both be cubes. Hence

$$\begin{aligned}y+\sqrt{2}i=\left(p+q\sqrt{2}i\right)^3 &= p^3+3p^2q\sqrt{2}i+6pq^2i^2+2q^3\sqrt{2}i^3\\ &= p^3-6pq^2+\sqrt{2}\left(3p^2q-2q^3\right)i\end{aligned}$$

where p and q are integers. This implies that $3p^2q-2q^3=1$, and so $q(3p^2-2q)=1$. Since p and q are integers, $q=1$ or -1. Hence, if $q=1$, $3p^2-2=1$. So $p=\pm 1$. This means that $y=p^3-6pq^2=\pm 5$. Since y is a positive integer, $y=5$, and hence $x=3$.

• • • **BREAK 8**

What happens if $q=-1$ in the above argument?

The possible lack of unique factorization was something that Fermat overlooked in his $x^3=y^2+2$ problem. (Could it have been the reason that the English mathematicians couldn't do it?) However, it was also an insurmountable difficulty for a certain class of odd primes, in one approach to FLT. For the rest of the primes, the regular primes, Kummer was able to show that the Euler–Fermat style of proof would go through. But how to tackle the irregular primes?

Apparently, Lamé was devastated when Kummer's letter about the regular and irregular primes was read to the French Academy of Sciences in 1847. It was no consolation that Lamé had proved the case $n=7$ in 1839. He had hoped that he could get much further.

Again there was a long period of little development. The next significant progress occurred in 1908 when Dickson proved that FLT was true for all n up to 7000.

In 1908, Paul Wolfskehl died and left 100,000 marks for anyone who could prove the truth of FLT. There is a fascinating story about this. Apparently, Wolfskehl had been spurned in love and decided to commit suicide. He nominated a time and place for himself and put everything in readiness for the final event. Having got everything nicely organized ahead of time, he had some time to kill before he pulled the trigger. Being an amateur mathematician of some note, he started to read a book on number theory. He became interested in Kummer's proof for regular primes and was surprised to find that Kummer had made a mistake. Wolfskehl tried to patch up Kummer's proof and became so engrossed in the problem that the appointed time of his suicide passed unnoticed. As it happened, he was able to rectify Kummer's proof. Elated by this, he decided that life was worth living after all. So he gave up the idea of suicide and changed his will to include the very substantial prize of 100,000 marks for the first person to prove FLT.

Unfortunately, the prize did not have the reaction he had hoped for. Mathematicians generally seemed to have decided that FLT was too hard or not sufficiently significant to bother with. On the other hand, the prize, worth just under 2 million US dollars in today's money, attracted a great deal of interest outside professional mathematical circles. The committee responsible for overseeing the prize in the German town of Göttingen was inundated with entries. The prize was not awarded (until recently).

2.9 I HAD TO SOLVE IT

In 1963, the ten-year-old Andrew Wiles wandered into a library in Cambridge, England, and found a copy of E.T. Bell's *The Last Problem*. This presented the long history of FLT including its Greek roots. Wiles was fascinated by what he read. In [7] he says

> It looked so simple, and yet all the great mathematicians in history couldn't solve it. Here was a problem that I, a ten-year-old, could understand and I knew from that moment that I would never let it go. I had to solve it.

How many other ten-year-olds have said they had to solve it? How many other people had said they had to solve it? Possibly thousands. But none had succeeded and at that time, not much seemed to be happening on the FLT front. With the advent of computers after the Second World War, it soon became known that the theorem was true for progressively larger values of n. By 1993 it had been shown to be true for n up to 4,000,000. However, proving the theorem by computer (in general, using the computer to prove anything in mathematics) is not just a matter of starting out with a particular value of n, and checking all possible values of x, y, and z. That approach is doomed to failure from the start.[5] There are, after all, an infinite number of numbers to try for x, y, and z, so the process is, in principle, endless. So there have to be ways found to reduce the problem to a finite search. This means that some relevant mathematical arguments have to be developed and applied. Tricks like that produced by Sophie Germain are needed to make the problem a finite one.

There is, of course, another problem with the computer. It may perhaps help you to show that a particular value of n satisfies FLT, but it won't enable you to solve the entire problem. At the end of the computations you will just know that the theorem is true for another value of n. There will still remain an infinite number of numbers n to be checked. In fact, it may be better to solve for a particular value of n by hand. This way you may get some insight into the problem that you would not get by using the computer. On the other hand, to solve for a particular n by computer requires various reduction techniques. These techniques might be developed into a more general argument. But these are the hands-on parts of the process and in some sense are independent of the computer.

So there is still the need for mathematicians to invent proofs. Of course, this is the position at the end of the twentieth century. It may be that, in the future, machines will have been developed to the stage where they can find proofs of theorems. Penrose in [6] suggests that this is unlikely, but who knows for sure what the future will bring?

[5] Unless FLT turned out to be false.

Anyway, while we have been chatting, Wiles has achieved his Bachelor's degree and become a graduate student under John Coates at Cambridge. In order to get a Ph.D. (Doctor of Philosophy degree) at Cambridge, and most other universities operating under the British system, it's necessary to work on a problem for 3 years or so and then submit your results in the form of a book called a thesis. The thesis is examined by a small group of people. They are looking for original material; if they don't find it, the candidate is unlikely to pass and be granted the Ph.D.

Coates decided that Wiles should work on ***elliptic curves***. These are equations of the form $y^2 = Ax^3 + Bx^2 + Cx + D$, where A, B, C, D are whole numbers. We've already seen one of these in Section 3, namely, $y^2 = x^3 - 2$. You remember that Fermat claimed to be able to show that the only solution in positive integers of this equation is $x = 3$, $y = 5$.

While it may seem that Wiles was putting his ambition on hold, it is important to have a Ph.D. to be an academic mathematician; and Wiles wanted a career in academia. As it turned out, elliptic curves are fundamental to his proof of FLT. But there is another concept fundamental to his proof — that of modular functions.

★ We're afraid that from here we are going to have to skim over many of the details of Wiles' proof. In that vein we give an approximation to the definition of modular functions but refer the reader who would like something more precise to [1] or [3]. Approximately, a *modular function of level N* is a complex function f such that

$$f\left(\frac{az+b}{cz+d}\right) = f(z)$$

for all integers a, b, c, d with $ad - bc = 1$ and $N|c$. An example of a modular function is

$$16e^{\pi iz}\prod_{n=1}^{\infty}\left(\frac{1+e^{2\pi inz}}{1+e^{2\pi i(n-1)z}}\right)^8$$

In the 1950s, Yutaka Taniyama and Goro Shimura began to be interested in these relatively complicated functions of complex variables. It is easy to define complex functions. For instance, $f(z) = z^2$ is an example. This means that $f(i) = i^2 = -1$,

$$f(1+2i) = (1+2i)^2 = 1 + 4i + 4i^2 = -3 + 4i,$$

and so on.

Taniyama and Shimura began to see a link between elliptic curves and modular functions. In every case they tried, they could get two modular functions to fit into a given elliptic curve. So they made the following conjecture.

The Shimura–Taniyama Conjecture: *Given an elliptic curve*

$$y^2 = Ax^3 + Bx^2 + Cx + D$$

over $\mathbb{Q}$ *(the rationals) there exist non-constant modular functions* $f(z)$ *and* $g(z)$*, of the same level* N*, such that* $f(z)^2 = Ag(z)^3 + Bg(z)^2 + Cg(z) + D$.

(For reasons we won't go into here, André Weil's name is sometimes also attached to this conjecture, and Taniyama's name is sometimes written first.)

For experts in the field, this was an astounding and exciting conjecture. It is rare that two such apparently disparate areas of mathematics as the theory of elliptic curves and the theory of modular functions turn out to be so closely linked. If the conjecture were true, it would undoubtedly mean advances in both areas. These would come because results in one area could be applied to the other.

Rather tragically, Taniyama committed suicide before his work with Shimura could be advanced very far.

The Shimura–Taniyama Conjecture attracted a considerable amount of attention. Many mathematicians were confident that it was true and even wrote papers with results worded "Assuming the Shimura–Taniyama conjecture, such and such follows." This is not an unusual thing to happen in mathematics. At one stage there was a conjecture about something called the Classification of Finite Simple Groups. Many group theorists wrote papers based on the assumption that the classification was correct. What is the value of such work? Well, first, if the conjecture turns out to be true, then we know that all the results that had assumed it are also true, so the field has progressed. And, secondly, the results based on the conjecture may lead to a contradiction. In that case the conjecture would be shown to be false. So there are potential gains both ways.

Now there is one final type of function to be introduced, whose significance is explained below. It is called a ***cusp form of weight 2 and level N***. Such a function F satisfies the identity

$$F\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 F(z)$$

where a, b, c, d, N are as in the definition of a modular function, but we will not attempt to define it fully here.

The next development was somewhat surprising. Gerhard Frey decided to assume that FLT was *false*, in order to produce a contradiction. Where would that lead and what had that got to do with the Shimura–Taniyama Conjecture?

In 1982 Frey assumed that u, v, w satisfy $x^p + y^p = z^p$, where p is an odd prime. In other words, $u^p + v^p = w^p$. From this equation he produced the elliptic curve

$$y^2 = x\,(x - u^p)\,(x + v^p)$$

But this was a strange elliptic curve, so strange that it probably wasn't related to a modular form. If that was so, it would contradict the Shimura–Taniyama Conjecture; you see, Frey was feeling his way towards a proof of FLT by contradiction.

So Frey went about setting up a proof that the Shimura–Taniyama Conjecture implied FLT. He recognized that a complete proof would require the expertise of a specialist in the arithmetic of modular forms; he thought such an expert would find it fairly easy to design a proof, and was surprised that it turned out to be more difficult than he had supposed. However, Jean-Pierre Serre succeeded in showing that "Shimura–Taniyama implies FLT" would follow from a certain "level-lowering" conjecture of his own, which experts could indeed start working on.

So at this stage it required two conjectures to prove FLT. This was reduced to one conjecture when Kenneth Ribet proved Serre's level conjecture in 1986. An outline of the proof went like this. Assume that the equation $x^p + y^p = z^p$ has a solution. Then produce the Frey elliptic curve. If the Shimura–Taniyama Conjecture is true, there exists a cusp form of weight 2 and level N, suitably related to the Frey elliptic curve. By the Serre level conjecture, this implies the existence of a cusp form of weight 2 and level smaller than N. Repeating this step eventually leads to the existence of a non-zero cusp form of weight 2 and level 2. It is known that the vector space of such cusp forms is the zero space, so we have a contradiction. Hence the original assumption is false and FLT is true.

So the only hurdle left to be overcome was finding a proof of the Shimura–Taniyama Conjecture.

You can imagine that the announcement of Frey's elliptic curve, and the consequent arguments linking Shimura–Taniyama and Fermat, caused a great deal of excitement in the mathematical community. However, surprisingly few mathematicians were prepared to make the attempt to prove the Shimura–Taniyama conjecture. They felt that it was an impossible task and turned their attention to other things.

But this was the spur that Wiles needed. He had long since finished his Ph.D. and had moved to a job in the Mathematics Department at Princeton University. Now he devoted himself almost entirely to settling the Shimura–Taniyama Conjecture. He didn't expect that it would be easy. He thought it might take 10 years or so. But he decided to concentrate all his research efforts into solving this one problem.

As part of his strategy he locked himself away from his colleagues and worked single-mindedly on his own. This was actually an unusual strategy. Nowadays mathematicians prefer to meet with their peers at regular intervals to discuss their work and the work of others. They also regularly go to conferences to learn what the latest techniques and results in their area are. But Wiles, very naturally, wanted the scalp of FLT for himself, and he also did not want to be distracted by others. So working on his own in his attic, he started to learn all there was to know about the subject he was about to tackle.

In 1988, Yoichi Miyaoka described a proof of FLT to a mathematics seminar in Bonn. Wiles must surely have been very upset by this news. The "proof," however, lasted only a short period before an unpatchable difficulty was discovered.

Then, in 1993, with a great sense of theater, Wiles announced his proof at a meeting in the Isaac Newton Institute in Cambridge, England. As with Miyaoka, the news hit the front pages of the world's leading newspapers! But, once again, an error was found. Wiles tried to find a way round his mistake, but there seemed to be no way that he could patch up his original argument. It was back to the attic.

Another blow came when rumours circulated that a counterexample to FLT had been found. There was, it was claimed, some n, suitably enormous, for which integer values of x, y, z could be found such that $x^n + y^n = z^n$. An advantage of e-mail is that news like this sweeps the mathematical world very quickly. Another advantage is that the original e-mail is dated. The message that started the excitement was sent on April 1!

At this stage Wiles was not working alone. A former research student of his, Richard Taylor, came and worked with him in an effort to provide mathematical support in areas where Wiles felt he would like to have some expert cooperation. It took some time to work around the error of the 1993 proof, but Wiles and Taylor came up with a proof of a restricted form of the Shimura–Taniyama Conjecture in late 1994.[6] This time the referees who pored over the two submitted manuscripts could find no error. In 1995 these

[6] In fact, they confined themselves to the so-called semi-stable elliptic curves; this, however, is enough to prove FLT. Subsequently, Henri Darmon reported in the *NOTICES of the American Mathematical Society* [2], that Cristophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor have announced a full proof of what Darmon refers to as the Shimura–Taniyama–Weil conjecture.

manuscripts were published under the titles "Modular elliptic curves and Fermat's Last Theorem," by Andrew Wiles, and "Ring-theoretic properties of certain Hecke algebras," by Richard Taylor and Andrew Wiles. They took up 130 pages in the prestigious mathematics research journal the *Annals of Mathematics*.

Although the Wolfskehl prize started off at the equivalent of about $2 million, it lost considerable value in the deep recession that Germany experienced in the 1930s. However, in 1997, when presented to Wiles, it was worth a healthy $50,000.

But, from the international mathematical community's perspective, Wiles missed out on the big prize — the Fields Medal. The will of John Charles Fields, a Canadian mathematician, established a fund to provide a medal that would play the role of the Nobel Prize in mathematics. The International Congress of Mathematicians in 1932 adopted Fields' proposal, and the medal was first awarded at the next congress in 1936. It is almost certain that Wiles would have been awarded a Fields Medal. However, there is a restriction to the award. A recipient has to be under 40, and Wiles was 41 when he and Taylor completed the proof of Fermat's Last Theorem! However, the International Congress of Mathematicians, held in Berlin in 1998, awarded Wiles a special International Mathematical Union ★ silver plaque in recognition of his work.

• • • BREAK 9

(1) Sally Jones bought some pigs, goats, and sheep. Altogether she purchased 100 animals and spent $600. Now, the pigs cost $21 each, the goats $8, and the sheep $3. If there was an even number of pigs, how many of each animal did Sally buy?

(2) Show that the equation

$$x^4 + 251 = 5y^4$$

has no integral solutions for x and y.

(3) Show that the equation

$$x^2 + y^2 = 80z + 102$$

has no integral solutions for x, y, and z.

(4) How old was Diophantus when he died? (See the inscription from his tomb at the beginning of Section 3.)

(5) Show that the equation

$$x^n + y^n = z^n, \quad n \geq 3$$

has no solution in positive rational numbers x, y, z.

(6) Show that the equation $x^{n/2} + y^{n/2} = z^{n/2}$, where n is a positive integer $\neq 1$, 2, or 4, has no solution in positive integers x, y, z.

Significant Dates for Fermat's Last Theorem (FLT)

1636? Fermat "proves" FLT for $n = 3$.
1637? Fermat writes a marginal note.
1640 Fermat proves FLT for $n = 4$.
1753 Euler proves FLT for $n = 3$.
1825? Germain considers the case where p and $2p + 1$ are both prime.
1825 Dirichlet and Legendre independently prove FLT for $n = 5$.
1839 Lamé proves FLT for $n = 7$.
1847 Lamé tries to prove FLT for all n assuming unique factorization. Kummer points out Lamé's error and proves FLT for all regular primes n.
1908 Dickson proves FLT for all $n \leq 7000$.
1908 Establishment of Wolfskehl prize.
1955 Beginnings of the Shimura–Taniyama Conjecture.
1985 Frey and Serre link the Shimura–Taniyama Conjecture with FLT via a potentially non-modular elliptic curve.
1986 Ribet completes the connection between the Shimura–Taniyama Conjecture and FLT.
1988 Miyaoka claims a proof of FLT.
1993 Wiles announces a proof of FLT that turns out to be incomplete.
1994 Wiles and Taylor complete Wiles' 1993 proof of FLT by replacing the defective part of the argument.

REFERENCES

1. Cox, David, Introduction to Fermat's Last Theorem, *Amer. Math. Monthly*, **101** (1994), 3–14.
2. Darmon, Henri, A proof of the full Shimura–Taniyama–Weil conjecture is announced, *NOTICES of the American Mathematical Society*, **46** (1999), 1397–1401.
3. Gouvêa, Fernando Q., "A Marvellous Proof," *Amer. Math. Monthly*, **101** (1994), 203–222.
4. Hilton, Peter, Derek Holton, and Jean Pedersen, *Mathematical Reflections — In a Room with Many Mirrors*, Springer-Verlag, New York, 2nd printing, 1998.

5. Mazur, Barry, Number Theory as Gadfly, *Amer. Math. Monthly*, **98** (1991), 593–610.
6. Penrose, Roger, *The Emperor's New Mind*, Vintage, London, 1989.
7. Ribenboim, Paulo, *Fermat's Last Theorem for Amateurs*, Springer-Verlag, 1999.
8. Singh, Simon, *Fermat's Last Theorem*, Fourth Estate, London, 1998.
9. van der Poorten, Alf, *Notes on Fermat's Last Theorem*, Wiley, Chichester, 1996.

Information on the web can be found regarding

- Fermat's Last Theorem at `http://www.ams.org/new-in-math/fermat.html` and `http://www.ams.org/mathweb/mi-mathbytopic,html#fermat`
- Andrew Wiles at `http://www.sigmaxi.org/prizes&awards/awiles.html` and `http://www.auburn.edu/~wakefju/wiles.html`
- The Fields Medal at `http://www.math.utoronto.ca/fields`