Fermat's Last Theorem and the Fourth Dimension

James Propp

What's the Problem?

Fermat's Last Theorem has got to be one of the most popular problems in the history of mathematics — millions of people have toyed with it, and thousands have worked up a real mental sweat trying to solve it. The problem, posed by the French mathematician Pierre de Fermat back in the seventeenth century, is usually stated in terms of the famous equation

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

where x, y, z and n represent unspecified whole numbers. When $n = 1$ the equation has too many solutions to be interesting, and when $n = 2$ there are still infinitely many ($3^2 + 4^2 = 5^2$ is the most famous). The problem Fermat bequeathed to us is to show that when n becomes bigger than 2, the situation changes dramatically: there are no solutions at all. That is: when n is a whole number bigger than 2, no number that is the nth power of a whole number can be written as the sum of two smaller nth powers.

It stands to reason that a proposition so tantalizingly simple would have a simple proof or a simple disproof. Yet for over three centuries the problem resisted the efforts of the sharpest minds that tackled it — and we still don't have a simple proof.

Fermat's Last Theorem came to light after Fermat's death, when his son Clement-Samuel was cleaning up the old man's library. An especially cherished work in the elder Fermat's collection had been a seventeenth-century Latin edition of a millennium-old Greek treatise on numbers by the mathematician

[†]James Propp is a professor of mathematics. He is writing a book on Fermat's Last Theorem for the mathematically-interested public.

[†]**David Feldman** prepared the illustrations.

Diophantus of Alexandria. On one page, Diophantus discussed the problem of writing a given square as a sum of two squares; writing in the margin of that page, Fermat made his no-go claim about higher powers and famously said he'd found a wonderful proof of this result but couldn't include it because the margin was too small.

Fermat as Publicist

The claim is not found elsewhere in Fermat's known writings, but on several occasions he did state that a third power can't be the sum of two smaller third powers, or a fourth power the sum of two fourth powers. However, in the combative fashion of the times, Fermat would often announce his results indirectly, by proposing challenges for other mathematicians to test their wits on. He thought that these challenges would give others a greater appreciation of the hidden depths surrounding his problems about numbers and lure them into doing active research on the topic, but sometimes the tactic backfired on him.

For instance, in one of his letters he challenged the English mathematician John Wallis to solve two problems:

1. given a cube, to write that cube as a sum of two cubes; and
2. given a sum of two cubes, to write that number as a sum of two cubes in a different way.

The first problem has no solution; in fact, this is just the case $n = 3$ of Fermat's Last Theorem. The second problem has many solutions; for instance, $(3/2)^3 + (5/3)^3$ can also be written as $(2)^3 + (1/6)^3$ (Fermat was concerned here with fractions as well as whole numbers). What Fermat seems to have wanted was for Wallis to demonstrate that the first problem had no solutions and then to give a systematic approach to solving the second problem. That is the real two-part challenge Fermat had in mind.

But from the way Fermat wrote the challenge, the first part seems to be asking Wallis to do something that is in fact impossible, and that Wallis probably suspected was impossible (perhaps after hours of fruitless work). It's understandable that Wallis resented this sneaky way of disguising the nature of the challenge, and later missed few opportunities to disparage Fermat's work on numbers.

Although Fermat's efforts to interest his contemporaries in problems about numbers were unsuccessful, he did find followers posthumously, starting in the century after his death. It was up to these disciples to fill in the blanks in Fermat's work, since Fermat himself had been loath to write down details. By the middle of the nineteenth century, all of Fermat's many claims had been proved (or, in a case or two, disproved), with the exception of the famous marginal note. This "Great Theorem of Fermat" also acquired the name "Fermat's Last Theorem" to mark its recalcitrance. Nowadays many people call it "FLT" for short.

What if ...?

Ironically, everyone knew how the proof of FLT should begin: "Suppose there did exist non-zero whole numbers x, y, z satisfying $x^n + y^n = z^n$, with $n > 2$. Then ...." In mathematics, to prove that something doesn't exist, it's frequently helpful to assume for argument's sake that the thing does exist, and then show that the thing, merely by existing, would have to possess mutually incompatible properties, thus demonstrating that it couldn't exist in the first place. This is the method of proof by contradiction, or *reductio ad absurdum*, and it's the method of choice for a problem like this.

So, people knew what the seed of the proof should be, but there has to be some sort of soil into which a seed can be planted. Fermat himself, back in the seventeenth century, seems to have tried planting the seed in the obvious place: the study of the properties of ordinary whole numbers. This study nowadays is called elementary number theory (to distinguish it from the more abstract developments that came later). Leonhard Euler, who as the first of Fermat's posthumous disciples revived the study of numbers in the eighteenth century, was able to construct proofs of FLT for the cases $n = 3$ and $n = 4$ (proofs conceivably found earlier by Fermat), using elementary methods. But going beyond $n = 3$ and $n = 4$ was hard. Euler's successors, and their successors up till the middle of the nineteenth century, were able to settle a few more cases, but this approach petered out and couldn't even be made to handle a value of n as small as 11. It seems that the ground of elementary number theory just doesn't have the right sort of nutrients for the seed of the proof of FLT — the kernel of contradiction — to sprout and grow into a full and rigorous argument.

In the middle of the nineteenth century, mathematicians like Ernst Eduard Kummer found a different plot of land to plant the seed in: a new sub-discipline within number theory called algebraic number theory, and a sub-sub-discipline called the theory of cyclotomic number rings. Cyclotomic number rings are extensions of the ordinary arithmetic of whole numbers, in which other sorts of numbers, including imaginary numbers like the square root of minus one, are brought into the game.

With the new methods, it became possible to prove FLT for many more exponents. Kummer more or less settled FLT for all exponents under 100 (he made a few mistakes on the hard ones). When Kummer's mistakes were corrected and his methods were extended and married with the power of twentieth-century computers, it became possible to prove FLT for all exponents up into the low millions. But, for all mathematicians knew, these corroborations were a fluke; FLT might have been false not just for one exponent, but for infinitely many exponents — perhaps even for *all* prime exponents with more than a million digits, say.

Someday mathematicians might know enough about cyclotomic number rings to be able to construct a proof along the lines that Kummer envisioned; but it seems that the soil of algebraic number theory, in its current state, doesn't have the right nutrients either.

FLT's Last Century

Over the course of much of the twentieth century, professional interest in Fermat's Last Theorem as a hot research topic dwindled. The problem was still part of the lore of mathematics, and part of the field's long-term agenda, but mathematicians found it hard to come up with new plans of attack that hadn't already been tried. No one had an idea how to proceed with FLT, and some experts even began to suspect that Fermat might have guessed wrong.

But outside the academies, more people were working on the problem than ever before. Amateurs were attracted to the problem for a number of reasons. First, FLT is a simple and catchy question. Second, the fact that Fermat claimed to have found a proof raised people's hopes that a proof, indeed a simple proof, could be found. Third, there are certain people who are attracted to a problem precisely because it's hard, and here was a problem that a whole community of experts, the world's mathematicians, had despaired of solving with existing tools. Fourth, there was a cash prize for the person who solved the problem. And fifth, it's easy to *almost* prove Fermat's Last Theorem, in a certain sense.

Remember the basic strategy for proving FLT: you assume that it's false and derive a contradiction. Well, it's very easy to arrive at contradictions in mathematics — just make one mistake and, unless you inadvertently make another mistake that cancels out the first one, you're likely to hit on two assertions that don't square with each other. Even if you find your mistake, or it's pointed out to you, and you realize that your attempted proof by contradiction isn't valid, it's easy to convince yourself that, since you found a proof of FLT with only one mistake in it, you might be close to finding a proof with none. This psychological effect made Fermat's challenge a very addictive problem to work on. But despite the serious efforts of very many people, with various degrees of persistence, no one could find a proof.

Finally, in the last decade of the twentieth century, mathematician Andrew Wiles, aided by his former student Richard Taylor, gave a proof of Fermat's Last Theorem. The proper soil for the seed, or at least one proper soil for it, had been found: an area called the theory of elliptic curves, whose borders Fermat himself had rambled across but whose true shape didn't emerge until the nineteenth century, and whose central inner jungle, still untamed today, is being explored by number-theorists with the aid of powerful ideas from all across the spectrum of mathematics.

And here we face a paradox: Even though Fermat's problem itself is the epitome of popular mathematics — easy to state, ponder, and play with — the eventual solution has the opposite character: it's as esoteric as can be, and it uses ideas from areas of mathematics that didn't even exist in Fermat's day.

It's nonetheless possible to say some things about the proof that are accessible at a popular level. For instance, one key feature of the new approach to Fermat's Last Theorem, and some would say the reason for its success, is the way in which it brings geometry into the story. Not just any old geometry, but the *right kind* of geometry.

The Geometry of the Torus

Picture an ordinary rectangular sheet of paper. If you were to roll up such a piece of paper as shown in the left panel of Figure 1, joining the left edge to the right edge, you would get a cylinder with a vertical axis of rotational symmetry. On the other hand, if you had rolled up the piece of paper as shown in the right panel of the figure, joining the top edge to the bottom edge, you would have gotten a different cylinder, with a horizontal axis of rotational symmetry.

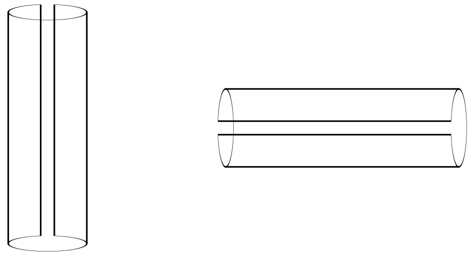

Figure 1: Two ways of bending a piece of paper to form a cylinder.

Could you have it both ways, joining left to right and top to bottom and creating two axes of rotational symmetry? The answer is No, if you limit yourself to three-dimensional space: no matter how you try to bend the two ends of a cylinder together, you can only create a new axis of symmetry by destroying the one that was there before.

But suppose that, after turning the two-dimensional rectangle into a cylinder that bends around in the third dimension, you could somehow bend that cylinder around through the fourth dimension. Then it turns out you can get a shape with two axes of full rotational symmetry. Mathematicians have known about shapes like this for over a century. But the real surprise was that Fermat's Last Theorem can be understood as making covert reference to the properties of certain shapes like this.

How can we get a handle on such a shape if we can't build it? If you're comfortable with the idea that a point in four dimensions can be specified by four numbers, or in some sense "is" just a quadruple of numbers, then the shape we're after can be described as the set of all quadruples of numbers (s, t, u, v) satisfying the two equations $s^2 + t^2 = 1$ and $u^2 + v^2 = 1$. The first equation describes a circle, and no part of a circle looks any different from any other part. The second equation also describes a circle. Combine the two circles in a four-dimensional way and you get a circle-of-circles in which every part looks the same as every other.

A more geometric way to understand this mysterious symmetrical shape is to consider distorted versions of it in ordinary three-dimensional space. If we drop the constraint that there be two axes of rotational symmetry, and brutally join the two ends of a cylinder in the third dimension rather than the fourth, we get a shape like the surface of a doughnut, as shown in Figure 2. (You can't do this with a cylinder made of paper, but a stretchable surface would work.) With its one axis of rotational symmetry, this doughnut is only an inadequate shadow of the more symmetrical shape that lives in four-dimensional space.

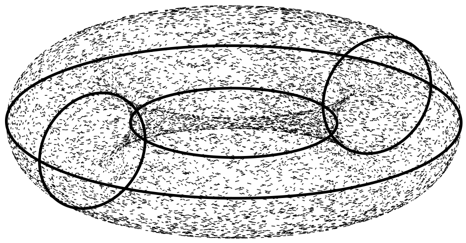

Figure 2: A torus.

There's a different way to try to understand the symmetrical surface, by ignoring the space that it sits inside and imagining instead what it would be like to be confined to it. Picture an ant on the original rectangle. If the left and right edges of the rectangle are joined, forming a cylinder, the ant's point of view is that when it crosses the join, it's transported from the left edge of the rectangle to the right edge, or vice versa. Now suppose we had a kind of magic paper which, without the need for any bending, would miraculously transport the ant from the left edge to the right edge and vice versa. To us, this scenario would look very different from the non-magic scenario, in which ordinary paper is rolled up into a cylinder; but from the ant's point of view, the two are the same.

Now suppose we had a fancier brand of magic paper which, without any bending, would miraculously transport the ant from the left edge to the right edge and vice versa, *and* from the bottom edge to the top edge and vice versa. From the ant's point of view, there's no difference between the rectangular universe of magic paper that it inhabits and the symmetrical surface in four dimensions that we're trying to understand.

Magic paper doesn't exist, but computers that can simulate it do. There's even money to be made from creating such simulations, as was discovered twenty years ago by the inventors of the popular video arcade game Asteroids. Nowadays, thanks to the World Wide Web, you don't even need a stack of quarters to experience what it would be like to live on such a surface; you can play Mike Hall's web-version [Hal]. Other games situated in the magic-paper universe can be found at the TorusGames site [Key]. With the help of such programs, this magical, un-makable surface, this creature whose natural habitat is the fourth dimension, can be made amenable to (virtual) exploration. (For other four-dimensional fun, try Rich Schwartz's game Lucy and Lily [Sch].)

Elliptic Curves

As the name of the second web-site might lead you to guess, a surface of this kind is also known as a torus. There's another name for tori: they're often called elliptic curves. This nomenclature is unfortunate, since the uninitiated are apt to think that the term "elliptic curve" refers to the kind of curve that's called an ellipse (shown in Figure 3.) Elliptic curves are totally different from ellipses. But there are good historical reasons for this confusing terminology.

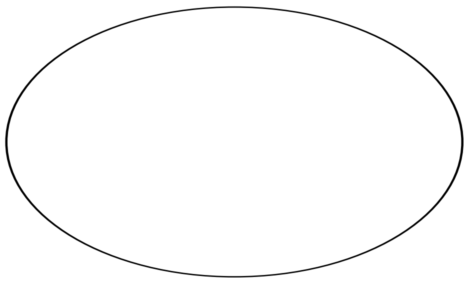

Figure 3: An ellipse.

Let's go back and talk about the ellipse a bit. An ellipse is what you get when you view a circle from a slant. The ancient Greeks studied the ellipse, but the shape didn't come into its own until the seventeenth-century astronomer Johannes Kepler discovered that the orbits of the planets are better approximated by ellipses than by circles.

Mathematicians of the seventeenth century, developing ideas that would later become the calculus, tried to find a formula for the circumference of an ellipse, and failed. It turned out that new kinds of mathematical functions had to be invented, just as the ancient Greeks had had to invent the sine and cosine functions in order to solve *their* problems about triangles.

These new functions turned out to be useful for lots of real-world problems: for instance, studying the behavior of swinging pendulums or buckling beams. But because the functions rose to prominence from their use in measuring the circumference of ellipses, they became known as elliptic functions.

Now, if you studied trigonometry, you probably did it twice: the first time trig was about properties of triangles, and the second time it was about properties of the circle of radius 1. In fact, trig functions like sine and cosine are sometimes called "circular functions" to honor the way they relate to properties of the circle. Something analogous happens with elliptic functions: curves like the one shown in Figure 4 give you a good way of thinking about elliptic functions. So curves like this became known as elliptic curves.

Just as the circle is given by the equation $x^2 + y^2 = 1$, this particular curve is given by the equation $x^3 + y^3 = 1$. Figure 5 shows another elliptic curve, with the equation $y^2 = -(x^2 - 1)(x^2 - 9)$. Algebraists call the locus one curve, even though it has two components, because it's given by a single equation.

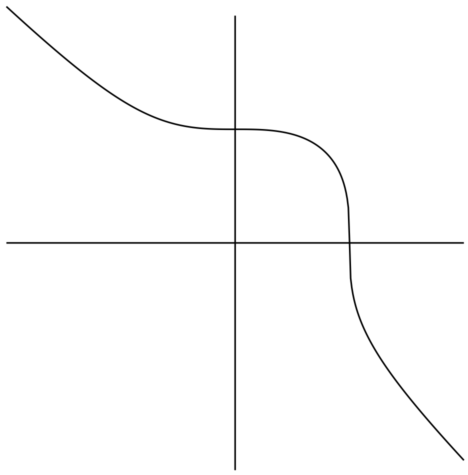

Figure 4: The curve $x^3 + y^3 = 1$.

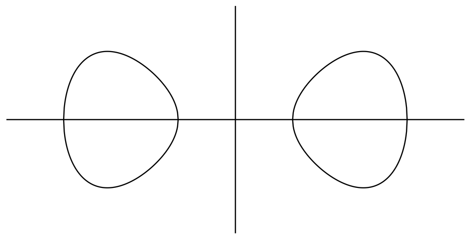

Figure 5: The curve $y^2 = -(x^2 - 1)(x^2 - 9)$.

So this is what mathematicians call an elliptic curve (and why) — but how do we get from here to doughnut-shaped surfaces in four-dimensional space?

The answer is: instead of looking at pairs of real numbers x, y that satisfy the algebraic relation $y^2 = -(x^2-1)(x^2-9)$, look at pairs of *complex* numbers that satisfy that relation, such as the pair $x = 0$ and $y = \sqrt{-9}$.

Complex Numbers

Every complex number can be written as $a + bi$, where i is the square root of minus one and where a and b are ordinary real numbers, called the real and imaginary parts of that complex number. So we can plot a complex number in two dimensions, by plotting the point (a, b), as shown in Figure 6. (For instance, $\sqrt{-9} = 0 + 3i$ would be represented by the point $(0, 3)$ on the vertical axis.) But to plot a *pair* of complex numbers, or to draw the graph of an equation involving two complex variables, you need two plus two dimensions — that is, you need to draw a surface in a four-dimensional space.

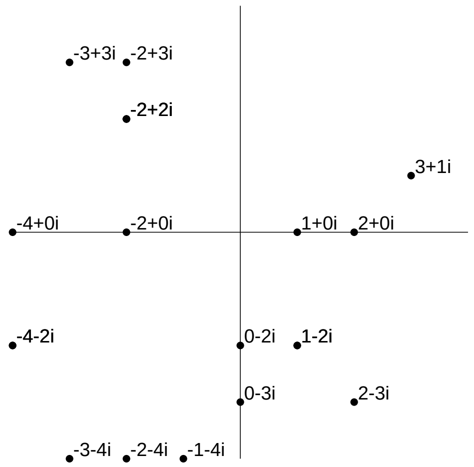

Figure 6: Some complex numbers plotted in the plane.

To get a peek at the four-dimensional surface that's latent in our two-dimensional picture Figure 5, let's step halfway into the fourth dimension, by stopping at the third. We're going to keep x a real number, but we're going to let y be a real number or an imaginary number, according to whether $-(x^2-1)(x^2-9)$ is positive or negative. To plot the value of y when it's imaginary, we'll use a third dimension, as in Figure 7.

Notice that where before we had just two closed components of the curve, we now have three: an extra component appears in the middle, touching the first two, protruding into the third dimension. Our new picture of the elliptic curve has two other components as well, one on the right and one on the left. They go off to infinity, but the nineteenth-century mathematicians who invented

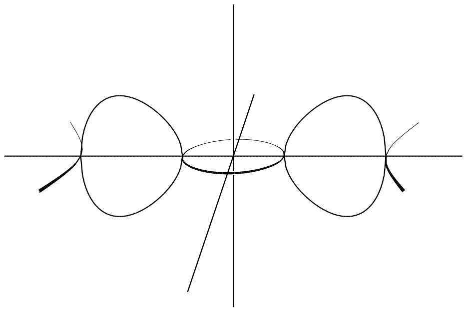

Figure 7: The plot of $y^2 = -(x^2 - 1)(x^2 - 9)$ with x real and y arbitrary.

projective geometry advised us that in contexts like this, it's appropriate to add in extra points at infinity. If you take their advice, you'll view these two unbounded components of the curve as actually meeting at a point at infinity.

So, schematically, what we get are four closed curves forming a kind of necklace, with each curve touching two of the others. But this is exactly what we have when we draw four circles on the surface of a doughnut, as shown in Figure 2!

So far, we've required x to be a real number. When you let x and y be any old complex numbers satisfying $y^2 = -(x^2 - 1)(x^2 - 9)$, the three-dimensional backdrop of Figure 7 becomes the spine of a larger four-dimensional backdrop, and the four closed curves become a kind of skeleton that the rest of the elliptic curve hangs on. I won't show you the details, but at least I hope you can see that the doughnut shape doesn't come out of nowhere. Moreover, it turns out that when you take this complexified four-dimensional picture and bend it in the right way, the curves of Figure 7 becomes four perfect circles. Likewise, when the curve $x^3 + y^3 = 1$ shown in Figure 5 is complexified and suitably bent, then the original curve, lying inside the surface that arose from it, forms a perfect circle.

This discussion barely scratches the surface of elliptic curves and their symmetries. For instance, consider the seemingly unrelated, recreational problem of placing 12 dots in the plane so that every line that goes through two of them goes through exactly one other point. How big could the number of such lines be? The answer is 19, and if you want to draw such a picture with 12 points and 19 lines, the prettiest way is to choose all sixteen points to lie on the curve given by the algebraic relation $(x - 1)((x + 2)^2 - 3y^2) = 8$. This is an elliptic curve with threefold symmetry, but its four-dimensional unfolding has even more symmetry; in fact, if you lift the curve up into the fourth dimension, so that it becomes a circle on a torus, then the twelve points are evenly spaced! (See the article "Planting Trees" by Stefan Burr [Bur81].)

Back to Fermat

How did these symmetrical surfaces, these visitors from the fourth dimension, get involved with Fermat's Last Theorem?

This part of the story can be traced back to the second half of Fermat's challenge to Wallis: if you've got some non-zero number c that can be written as $a^3 + b^3$ in one way (with a and b rational numbers — that is to say, whole numbers or fractions), then, leaving aside the case where $a = b$ or $a = 0$ or $b = 0$, Fermat said that there must be another way to write c as $a'^3 + b'^3$ with a' and b' two other rational numbers. That is, in addition to the point (a, b) and its twin (b, a), there must be another "rational point" (a', b') on the curve $x^3 + y^3 = c$, where a point (x, y) is called rational if both of its coordinates x, y are rational numbers.

The general notion of an elliptic curve didn't exist in Fermat's day, but the curves of the form $x^3 + y^3 = c$, along with other specific curves studied by Fermat, are in fact examples of elliptic curves. Fermat had some tricks for finding rational points on his elliptic curves, and these were developed further by later mathematicians, notably Isaac Newton. At the end of the nineteenth century a beautiful picture emerged in the work of mathematician Henri Poincaré, which combined Fermat's interest in finding rational solutions to algebraic equations with the new four-dimensional view of elliptic curves: in the four-dimensional picture, the rational points are perfectly evenly spaced.

For instance, the rational points on the curve $x^3+y^3 = 1$ are the point $(1, 0)$, the point $(0, 1)$, and an extra honorary point at infinity. But when you carry the picture up to the symmetrical torus in four-dimensional space, these three points become the vertices of an equilateral triangle.

Similarly, the rational points on other elliptic curves will give you squares, or regular pentagons, or regular hexagons, or pairs of regular pentagons, or various other things. And sometimes you get infinitely many rational points on the elliptic curve, and they're smeared out to appear to fill up a circle or a pair of circles.

The mathematician Louis Mordell, in the first half of this century, enlarged on Poincaré's vision. In this, he drew inspiration from Fermat's ideas, including the method that Fermat used in proving FLT for the case $n = 4$.

As mathematicians continued to study elliptic curves, a dichotomy emerged. Every elliptic curve either had infinitely many rational points or else had sixteen or fewer (counting the point at infinity as an honorary rational point). But no one had a proof of this.

In 1969, mathematician Yves Hellegouarch realized that if you could get an elliptic curve whose rational points, when lifted up to the symmetrical four-dimensional picture of the curve, formed a regular p-sided polygon ("p-gon"), with p a large prime number, you'd be well on your way to getting a counterexample to FLT for exponent p. Turning that around: if you knew that FLT was true for the exponent p, you'd know that you couldn't have p rational points on an elliptic curve arranged in a regular p-gon.

Two other mathematicians working in the 1970s and early 1980s, Vadim

Demjanenko and Gerhard Frey, partly independently but with some mutual influence, studied the problem and came to similar conclusions about the connection between elliptic curves and Fermat's Last Theorem. But for these researchers, the link was initially seen as bad news. Elliptic curves were what they wanted to understand; reducing their question to a notoriously difficult problem didn't seem like progress. As Frey wryly put it, "To try to solve a question and to come to Fermat's problem is not encouraging."

But at least there was a link! And in fact Hellegouarch had done some work back in the 1970s, suggesting that the link went both ways — that is, if you could prove the claim about elliptic curves, you might be able to use that information to get a proof (or partial proof) of FLT. Somewhat later, but independently, Frey noticed the same thing: if $A^n + B^n = C^n$ is a counterexample to FLT, then the elliptic curve $y^2 = x(x-A)(x+B)$ has properties that look very fishy to someone conversant with modern number theory. Maybe (Hellegouarch and Frey thought), using known facts about elliptic curves, one could prove that this elliptic curve had self-contradictory properties. Then one would know that no such triple (A, B, C) existed. That is, one could use facts about elliptic curves as a way of tackling FLT.

In the mid-1970s, Barry Mazur found a proof of the result about elliptic curves that Hellegouarch, Demjanenko, and Frey had tried and failed to prove. When Frey heard the news, he was electrified. Maybe Mazur's result, or Mazur's methods, could be applied to prove FLT!

The Road to Wiles' Attic

Frey began to look for more ways to relate FLT to what was known about elliptic curves, as well as what wasn't known but was strongly believed. Like Hellegouarch, Frey studied the properties of an elliptic curve $y^2 = x(x-A)(x+B)$ derived from a putative counterexample to FLT. He worked at the problem for several years, studying it from various angles. Finally, he found the angle that seemed most promising. In 1984 he startled the mathematical community by announcing strong reasons for thinking that such an elliptic curve couldn't be modular.

I haven't told you what it means for an elliptic curve to be "modular", and I'll only explain it here in the vaguest of terms: it means that there's an entirely different way to think about that specific elliptic curve using a different, even weirder geometry, called hyperbolic geometry. By the 1980s most number-theorists believed that all elliptic curves — to be precise, all rational elliptic curves (elliptic curves given by equations involving only rational numbers as coefficients, and having at least one rational solution) — were modular. This proposition had become known as the Shimura-Taniyama-Weil Conjecture, in honor of Yutaka Taniyama, who had first proposed a preliminary version of it, and Goro Shimura and André Weil, who had made the claim more specific and more testable. Subsequent researchers had obtained abundant evidence in favor of the proposition. So for Frey to announce that he'd found a way to construct an elliptic curve that seemed to be non-modular was quite dramatic — even if his construction hinged on FLT being false.

This announcement gave Frey's work an impact that the work of Hellegouarch had lacked. The seemingly non-modular elliptic curves of Hellegouarch and Frey were dubbed "Frey curves", and number-theorists began to study them with the hope of proving that they were as non-modular as they seemed to be (assuming that they existed at all).

Frey's work wasn't a theorem, but more of a sketch, with some key ideas missing. In 1986 Kenneth Ribet, building on work of Jean-Pierre Serre, showed that Frey was right: if there were a counterexample to FLT, then the associated Frey curve would have to be a non-modular elliptic curve.

And *this* convinced many experts, who'd hitherto been agnostic about FLT, that FLT must be was true — because there was so much evidence that every rational elliptic curve was modular.

At this point Andrew Wiles, energized by Ribet's result, decided to try to prove that all rational elliptic curves were modular, or at least all elliptic curves in a broad class that included the Frey curves. This result would give the needed contradiction. That's because Ribet's work had shown that the Frey curve, constructed from a putative counterexample to FLT, *wasn't* modular. If Wiles could show that Frey's curve *was* modular, this contradiction would show that no such curve could exist. That is, no such counterexample A, B, C could exist, and Fermat's Last Theorem would be established!

The Nut, Cracked

So we can say (at last!) that the theory of elliptic curves was the soil in which Wiles wanted to plant the seed of the proof. But the soil wasn't exactly easy to till. Many people badly wanted to know whether all rational elliptic curves were modular, as seemed to be the case, but the experts were convinced that a proof was a long way off. Wiles, in attempting to prove some version of the modularity conjecture (as the linchpin of a proof of FLT), took an odd sort of consolation from the notorious difficulty of the modularity conjecture: at least he wouldn't have to worry that a lot of people were trying the same thing he was working on. The smart money said that it was too soon to try to prove Shimura-Taniyama-Weil.

It turned out that the smart money was, in a way, right: the tools that Wiles needed didn't all exist in the 1980s. But during the period when Wiles was doing his work, other researchers created some of the tools he needed, not realizing the use to which they could be put. Wiles' timing, with hindsight, can be judged to have been nearly perfect: the new tools gave him the leverage he needed just when he needed it. After seven years of hard work, plus an eighth, excruciating year of announcement, retraction, collaboration, and revision, Wiles finally proved in 1994 that the (with hindsight, fictitious) Frey curve was modular. In combination with Ribet's work, this proved Fermat's Last Theorem at last.

I want to stress that the twentieth-century proof of Fermat's Last Theorem uses not just algebra and calculus and elliptic curves but all kinds of modern math. So you shouldn't get the idea that the fourth dimension is *the* magic key to the problem; it's one of dozens of magic keys, all of which played crucial roles.

It might seem unjust that such a huge amount of machinery, whose scope I have barely hinted at, should be required for the solution of as simple-sounding a problem as Fermat's Last Theorem. But we all know the principle of leverage that makes a nutcracker work, and it makes a kind of sense that when one is trying to crack as eminently tough a nut as FLT, it might be necessary to apply force at a point far removed from where the nut itself is, or seems to be.

The image of the nutcracker, with its suggestions of strength judiciously applied, is meant to convey a sense of both the effort and the elegance behind Wiles' accomplishment, but the analogy leaves out something important that I tried to convey earlier with a different agricultural metaphor. We shouldn't forget that a nut in the end is just another kind of seed. Perhaps when Fermat's problem is planted in the soil of some still-unknown mathematical country, it will open in the way seeds are designed to open, from the inside out. Then we may have a more accessible answer to this most delightfully accessible and wonderfully difficult of mathematical riddles.

Thanks to Joe Buhler, Henry Cohn, John Conway, Noam Elkies, Gerhard Frey, Yves Hellegouarch, Franz Lemmermeyer, Barry Mazur, and Kenneth Ribet for commenting on earlier versions of this article and offering general technical advice. All mistakes are the responsibility of the author.

Bibliography

[Bur81] Stefan Burr. Planting trees. In David A. Klarner, editor, *The Mathematical Gardner*, pages 90–99. Prindle, Weber, and Schmidt, Boston, 1981.

[Hal] `http://kresch.com/online_games/asteroids/`. A web version of the game Asteroids. Copyright by Mike Hall (1998).

[Key] `http://humber.northnet.org/weeks/TorusGames/`. TorusGames copyrighted by Key Curriculum Press (2001).

[Pro] Jim Propp. `http://www.math.wisc.edu/~propp/flt4d.html`. Some notes and references associated with this article.

[Sch] Rich Schwartz. Lucy and Lily. `http://www.math.umd.edu/~res/`. A brilliant four-dimensional “video game”. Copyright by Rich Schwartz (2000).