Zeitschrift: Elemente der Mathematik
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 55 (2000)

Artikel: From Fermat to Wiles: Fermat's Last Theorem Becomes a Theorem
Autor: Kleiner, Israel
DOI: https://doi.org/10.5169/seals-5627

Elem. Math. 55 (2000) 19 – 37
0013-6018/00/010019-19 $1.50+0.20/0

Elemente der Mathematik

From Fermat to Wiles: Fermat's Last Theorem Becomes a Theorem

Israel Kleiner

Israel Kleiner is professor of mathematics at York University in Toronto. He received his PhD in ring theory from Mc Gill University. His current research interests are the history of mathematics, mathematics education, and their interface. He was for many years coordinator of an in-service Master's Programme for teachers of mathematics. Recently he served as vice president of the Canadian Society for the History and Philosophy of Mathematics, and is currently on the advisory board of the International Study Group for the Relations between the History and Pedagogy of Mathematics.

1 Introduction

When historians come to judge the mathematics of the 20th century, I am confident that they will regard it as a golden age, for both the emergence of brilliant new ideas and the solution of longstanding problems (the two are, of course, not unrelated). In the latter category, Fermat's Last Theorem (FLT) is neither the most ancient nor the latest example. It has recently been announced (September 1998) that Kepler's Sphere-Packing Problem, posed by him in 1611, has been solved by Thomas Hales [22]. Of course, the Riemann Hypothesis, the Goldbach Conjecture, and other outstanding problems are still unresolved.

Here are quotations from the two main protagonists in the drama associated with FLT.

> It is impossible to separate a cube into two cubes or a fourth power into two fourth powers or, in general, any power greater than the second into powers of

Der Beweis der Fermat-Vermutung im Jahre 1995 durch Andrew Wiles und Richard Taylor ist ohne Zweifel eine der hervorragendsten Leistungen der Mathematik in diesem Jahrhundert. Zu Recht wurde denn auch die Lösung dieses über 350 Jahre alten, zahlentheoretischen Problems in weiten Kreisen als Sensation gefeiert. In diesem Beitrag gibt Israel Kleiner einerseits einen ausgezeichneten Überblick über die wesentlichen Beiträge zur Fermat-Vermutung aus den beiden letzten Jahrhunderten und der ersten Hälfte dieses Jahrhunderts, und andererseits findet der Leser eine spannende Darstellung der Ereignisse der letzten fünfzehn Jahre beginnend mit einer spektakulären Idee von Gerhard Frey im Jahre 1985 bis hin zur endgültigen Lösung des Problems durch Andrew Wiles und Richard Taylor. *jk*

> like degree. I have discovered a truly marvelous demonstration, which this margin is too narrow to contain [25, pp. 145–146].
>
> One morning in late May, Nada was out with the children and I was sitting at my desk thinking about the remaining family of elliptic equations. I was casually looking at a paper of Barry Mazur's, and there was one sentence there that just caught my attention. It mentioned a nineteenth-century construction, and I suddenly realized that I should be able to use that to make the Kolyvagin-Flach method work on the final family of elliptic equations. I went on into the afternoon and I forgot to go down for lunch, and by about three or four o'clock I was really convinced that this would solve the last remaining problem. It got to about teatime and I went downstairs and Nada was very surprised that I'd arrived so late. Then I told her – I'd solved Fermat's Last theorem [28, p. 243].

Both statements, by Fermat and Wiles, respectively – about 360 years apart – purport to have proved FLT. Wiles, as we know, published a proof, although his initial proof contained a major error which took 18 months to set right. But I'm getting ahead of myself.

The aim of this paper is to relate something of what happened in the three-and-a-half centuries between these two pronouncements, and, in particular, to describe some of the drama and the ideas connected with Wiles' proof.

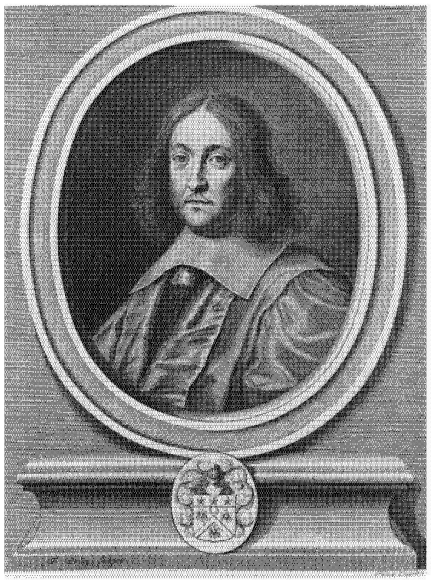

Fig. 1 Pierre de Fermat

2 The first two centuries

We begin at the beginning, with Fermat. His famous claim was that the equation $x^n + y^n = z^n$ has no (nonzero) integer solutions if $n > 2$. In the 19th century this pronouncement came to be known as Fermat's Last Theorem. (Fermat made many assertions in number theory without proof; all but one were later proved by Euler, Lagrange, and others. The exception – the last unproved "result" – was presumably the reason for the name "Fermat's Last Theorem".) Fermat made the claim in the 1630s, in the margin of Diophantus' book *Arithmetica* (250 AD), alongside his Problem 8, Book II, which said: Given a number which is a square, write it as a sum of two squares. As for Fermat's "truly marvelous demonstration", it was, of course, never published.

Fermat did publish a proof for $n = 4$, the simplest exponent to deal with. He accomplished this by introducing the *method of infinite descent*, which has turned out to be important in the proofs of many number-theoretic results. The idea is to assume that the equation $x^4 + y^4 = z^4$ does have a solution for some positive integers a, b, c, and to show that it then has a solution for positive integers u, v, w, with $w < c$. Repeating this process ad infinitum leads to a contradiction, since it introduces an infinite descending sequence of positive integers; see [13] for details.

The fact that FLT holds for $n = 4$ implies that it also holds for $n = 4k$, k any positive integer. For, if $x^{4k} + y^{4k} = z^{4k}$ for some integers x, y, z, then $(x^k)^4 + (y^k)^4 = (z^k)^4$ for the integers x^k, y^k, z^k. The same type of argument shows that if FLT holds for $n = p$, then it holds for $n = pk$. Since any integer is either a multiple of 4 or a multiple of an odd prime, Fermat's proof for $n = 4$ implies that it suffices to prove FLT for odd primes p.

A proof of FLT for $n = 3$ was given by Euler about 1760, over 100 years after Fermat's proof for $n = 4$. Euler's argument, however, contained a significant gap, not noticed by anyone at the time [13].

At the end of the 18th century the Paris Academy offered a prize for a proof of FLT. In 1816, Olbers, Gauss' astronomer friend, suggested that he compete for the prize. This was fifteen years after Gauss' publication of the *Disquisitiones Arithmeticae*, which established him as one of the foremost mathematicians of his time. Gauss responded as follows [24, p. 3]:

> I am very much obliged for your news concerning the Paris prize. But I confess that Fermat's theorem as an isolated proposition has very little interest for me, because I could easily lay down a multitude of such propositions, which one could neither prove nor dispose of.

It appears that Gauss did not consider FLT a fruitful problem. (But proofs of the theorem for $n = 3$ and 5 were found among his unpublished notes; see [6, pp. 90-91].) This raises an interesting question: What *is* a good mathematical problem? Of course, individual mathematicians choose a problem to work on because it interests them; but how are they going to get their colleagues interested in it?

The following two major criteria for what makes a good problem are likely not in dispute:

(i) The solution of the problem has important consequences. This is certainly the case for the Riemann Hypothesis.

(ii) New ideas are introduced in attempts to solve the problem. This, as it turned out, was undoubtedly so for FLT, but, of course, one knows that only in retrospect. In this sense, Gauss seems to have misjudged the problem, as we shall see.

The next strides in the proof of FLT were made by Legendre and Dirichlet, who, around 1825, independently established the theorem for $n = 5$. In 1839, Lamé proved it for $n = 7$. Incidentally, in 1832 Dirichlet showed that FLT holds for $n = 14$ but could not prove it for $n = 7$; the latter result, as we noted, implies the former.

Fig. 2 Sophie Germain

3 Sophie Germain

The first important breakthrough on FLT was made in 1823 by the French mathematician Sophie Germain. She proved the following useful result, using relatively elementary methods: If p and $2p + 1$ are both prime, then $x^p + y^p = z^p$ has no solutions for which xyz is not divisible by p. Largely as a result of this theorem, it has been customary to divide the proofs of FLT for various values of p into two cases, the so-called *Case I*, in which none of x, y, z is divisible by p, and *Case II*, in which at least one of x, y, z is divisible by p. (For example, it follows from Germain's result that Case I of FLT is true for the primes 5 and 11.) Case II is usually regarded as much harder than Case I [13], [24].

Legendre extended Germain's theorem to the following: Case I of FLT holds for the prime exponent p provided that one of $4p + 1, 8p + 1, 10p + 1, 14p + 1$, or $16p + 1$ is also prime. Germain and Legendre were now able to establish the first case of FLT

for all primes $p < 100$. In 1977 Terjanian showed that the first case holds for all even exponents $2p$ [24, p. 20].

An interesting problem is whether there are infinitely many "Sophie Germain primes", namely primes p for which $2p + 1$ is also prime. "This question is of the same order of difficulty as the well-known 'twin-prime' problem" [24, p. 56].

4 Lamé

In over 200 years FLT was proved for only four exponents – 3, 4, 5, and 7! On March 1, 1847, a dramatic event occurred at a meeting of the Paris Academy of Sciences. Lamé announced that he had proved FLT for all exponents, and presented a brief outline of the proof. Before describing its gist, let us consider the following simpler problem, whose solution embodies the essential elements of Lamé's proof:

Find all primitive pythagorean triples, namely all integer solutions of $x^2 + y^2 = z^2$ with x, y, z relatively prime.

While there are elementary solutions of this problem, the following method is instructive for our purposes. We factor the left side of $x^2 + y^2 = z^2$ to obtain $(x + yi)(x - yi) = z^2$. This is now an equation in the domain of "complex integers" of the form $G = \{a + bi : a, b \in Z\}$, the so-called *Gaussian integers*. It turns out that we can do number theory in G just as in Z. In particular, a "Fundamental Theorem of Arithmetic" holds in G, namely, every nonzero, noninvertible element of G is a unique product of primes. It follows that if a product of relatively prime elements in G is a square, then each element is a square. The same result holds with the exponent 2 replaced by any exponent [2], [16].

Since x, y, z are relatively prime in Z, it can be shown that $x + yi$ and $x - yi$ are relatively prime in G. Because their product is a square, each must be a square. In particular, $x + yi = (a + bi)^2$, where $a, b \in Z$. Thus $x + yi = (a^2 - b^2) + 2abi$, and comparing real and imaginary parts we get $x = a^2 - b^2, y = 2ab$. Since $z^2 = x^2 + y^2$, it follows that $z = a^2 + b^2$. So the solutions of $x^2 + y^2 = z^2$ are $x = a^2 - b^2, y = 2ab, z = a^2 + b^2, a, b \in Z$. Conversely, it can be shown that these are solutions for *every* choice of integers a and b. For x, y, z relatively prime, a and b must be relatively prime and of opposite parity. The resulting formula yields all primitive pythagorean triples. See [16].

Two important ideas are implicit in this solution:

(a) Embedding a problem about integers in a domain of "complex integers". The notion of embedding a problem formulated in a given domain in a larger domain is a common and important mathematical technique. Hadamard's dictum that the shortest path between two truths in the real domain passes through the complex domain is, indeed, illuminating.

(b) Transforming an additive problem into a multiplicative one; in this case, $x^2 + y^2 = z^2$ into $(x + yi)(x - yi) = z^2$. This, too, is an important and not uncommon device. Multiplicative problems in number theory are, in general, much easier to deal with than additive ones, especially in the presence of a Fundamental Theorem of Arithmetic.

Now to a sketch of Lamé's proof of FLT.

Assume that the equation $x^p + y^p = z^p$ has integer solutions (p a prime). Factor its left side to obtain $(x+y)(x+yw)(x+yw^2)\cdots(x+yw^{p-1}) = z^p$ (**), where w is a primitive p-th root of 1 (that is, w is a root of the equation $x^p = 1, w \neq 1$). This is now an equation in the domain $D_p = \{a_0 + a_1 w + \cdots + a_{p-1}w^{p-1} : a_i \in Z\}$ of so-called *cyclotomic integers*. Lamé claimed that if the factors on the left-hand-side of (**) are pairwise relatively prime in D_p, then, since their product is a p-th power, each must be a p-th power. From this a contradiction can be derived using Fermat's method of infinite descent by finding integers u, v, w such that $u^p + v^p = w^p$, with $w < z$. If the factors are not relatively prime, then by a suitable division by some element a, one obtains the relatively prime factors $(x+y)/a, (x+yw)/a, (x+yw^2)/a, \ldots, (x+yw^{p-1})/a$, and the proof proceeds analogously [5], [13].

After Lamé's presentation, Liouville, who was in the audience, took the floor and noted what seemed to him to be a gap in Lamé's proof, namely the latter's contention that if a product of relatively prime factors is a p-th power, each must be a p-th power. The result is, indeed, true for the integers, Liouville observed, but it remains to be shown that it is also true for the cyclotomic integers. Lamé agreed that further consideration was needed, but was convinced that he had the right approach to the proof.

What was required for the proof was a Fundamental Theorem of Arithmetic in D_p [13], [24]. This is the subject of the next section.

5 Kummer

About two months after Lamé's presentation to the Academy, Liouville received a letter from Kummer confirming the grounds for his skepticism about Lamé's proof [24, p. 7]:

> Encouraged by my friend M. Lejeune Dirichlet, I take the liberty of sending you a few copies of a dissertation which I wrote three years ago. ... In these memoirs, which I beg you to accept as a sign of my deep esteem, you will find developments concerning certain points in the theory of complex numbers composed of roots of unity, that is, roots of the equation $r^n = 1$, which have been recently the subject of some discussions at your illustrious Academy, on the occasion of an attempt by M. Lamé to prove the last theorem of Fermat.
>
> Concerning the elementary proposition for these complex numbers, that a *composite complex number may be decomposed into prime factors in only one way*, which you regret so justly in this proof, which is also lacking in some other points, I may assure you that *it does not hold in general* for complex numbers of the form $a_0 + a_1 r + a_2 r^2 + \cdots + a_{n-1}r^{n-1}$, but it is possible to rescue it, by introducing new kinds of complex numbers, which I have called *ideal complex numbers*.
>
> I considered already long ago the applications of this theory to the proof of Fermat's theorem and I succeeded in deriving the impossibility of the equation $x^n + y^n = z^n$ [for all $n < 100$].

Kummer says three fundamental things here:

(a) Unique factorization fails, in general, in D_p. He showed that it already fails for $p = 23$. It was shown in 1971 by Uchida that unique factorization fails in D_p for all $p \geq 23$.

(b) Unique factorization can be "rescued" in D_p by introducing "ideal numbers".

(c) Using unique factorization in the extended cyclotomic domains containing the ideal numbers, one can prove FLT for all primes $p < 100$. Kummer proved more. Specifically, he showed that FLT holds for all "regular" primes, a prime being *regular* if it does not divide the *class number* of D_p; equivalently, if it does not divide the numerators of the *Bernoulli numbers* $B_2, B_4, \ldots, B_{p-3}$ (for definitions of class number and Bernoulli numbers see [23], [24]). He then showed that all but three of the primes < 100 are regular; the irregular primes were handled separately. Incidentally, it was shown in 1915 that there are infinitely many irregular primes; it is not known if there are infinitely many regular primes [13], [23], [24].

The following examples of nonunique factorization into primes in various domains and its restoration by the addition of "ideal" elements illustrate some of Kummer's ideas in more elementary contexts.

(i) $D = 2Z$, the even integers. Here $100 = 2 \times 50 = 10 \times 10$, where 2, 10, 50 are primes in D (they cannot be factored in D).

(ii) D = all polynomials over the reals (say) of degree > 1. Here $x^6 = x^2 \cdot x^2 \cdot x^2 = x^3 \cdot x^3$, with x^2 and x^3 prime in D.

(iii) $D = \{a + b\sqrt{5}i : a, b \in Z\}$. Here $6 = 2 \times 3 = (1 + \sqrt{5}i)(1 - \sqrt{5}i)$, and it can be readily shown that $2, 3, 1 \pm \sqrt{5}i$ are prime in D. This example was given by Dedekind in the 1870s. In the first two examples D is not an integral domain, but its multiplicative structure illustrates well nonunique factorization [2].

As for rescuing unique factorization:

In (i) adjoin the "ideal number" 5.

In (ii) adjoin the "ideal polynomial" x.

In (iii) adjoin the "ideal numbers" $\sqrt{2}, (1+\sqrt{5}i)/\sqrt{2}$ and $(1-\sqrt{5}i)/\sqrt{2}$. We then have: $6 = 2 \times 3 = \sqrt{2} \times \sqrt{2} \times [(1+\sqrt{5}i)/\sqrt{2}] \times [1-\sqrt{5}i)/\sqrt{2}]$ and $6 = (1+\sqrt{5}i)(1-\sqrt{5}i) = \sqrt{2} \times [(1+\sqrt{5}i)/\sqrt{2}] \times \sqrt{2} \times [(1-\sqrt{5}i)/\sqrt{2}]$. Unique factorization of the element 6 has been restored [16].

Kummer's work saw the emergence of a new subject – *algebraic number theory*, already in evidence in earlier works of Gauss, Eisenstein, and Jacobi in connection with higher reciprocity laws [13], [15]. Moreover, Kummer's work on ideal numbers was vastly extended by Dedekind through his introduction of ideals, "one of the most decisive advances of modern algebra" [6, p. 91]. Dedekind, along with Kronecker, brought algebraic number theory to maturity [6], [16], [23]. Thus FLT acted as an incentive to the introduction of important mathematical concepts and results. More generally, "elementary" number theory has inspired the construction of deep theories which have illuminated mathematics well beyond the problems which gave them birth.

6 Early decades of the 20th century

Many technical results about FLT were obtained in the period 1850-1950, but there were no major breakthroughs. Here is a very small sample of such results [24]:

(i) Case I of FLT holds for infinitely many pairwise relatively prime exponents (Maillet, 1897).

(ii) If p is a prime such that $2^{p-1} \not\equiv 1 \pmod{p^2}$, then Case I of FLT holds for p (Wieferich, 1909).

(iii) If the so-called "second factor" of the class number of D_p [22, p. 27] is not divisible by p, and if none of the Bernoulli numbers $B_{2np}(n = 1, 2, \ldots, (p-3)/2)$ is divisible by p^3, then Case II of FLT holds for p (Vandiver, 1929).

(iv) If $p \equiv 1 \pmod 4$ and p does not divide the Bernoulli numbers B_{2s} for all odd s with $2 \leq 2s \leq p-3$, then FLT holds for p (Vandiver, 1929).

Using some of these and other results, Vandiver was able to establish by the end of the 1920s that FLT holds for all primes $p < 157$ (recall that about 70 years earlier Kummer had reached $p < 100$). Using the SWAC calculating machine, Vandiver in 1954 extended the result to $p < 2521$ [24, p. 202].

At the turn of the 20th century Hilbert was asked why he never attempted to prove FLT. Here is his response [29, p. 69]:

> Before beginning I should have to put in three years of intensive study, and I haven't that much time to squander on a probable failure.

Contrast this with Gauss' statement about why he did not compete for the Paris prize offered for a proof of FLT: Gauss claimed the problem did not interest him, Hilbert that it was too difficult.

In 1908 the mathematician Paul Wolfskehl bequeathed a prize for a proof of FLT, valued at 100,000 marks (the equivalent of $1,000,000 by today's standards). This came to be known as the *Wolfskehl Prize*. His stipulation was that if the prize were not awarded by September 13, 2007, no subsequent claim would be accepted. It seems, certainly in retrospect, that Wolfskehl had a good sense of the difficulty of the problem, giving mathematicians another 100 years to come up with a proof [3].

In a lighter vein, mathematicians at the mid-20th century would likely have empathized with the following sentiments [12]:

> M. Fermat – what have you done?
> Your simple conjecture has everyone
> Churning out proofs,
> Which are nothing but goofs!
> Could it be that your statement's an erudite spoof?
> A marginal hoax
> That you've played on us folks?
> But then you're really not known for your practical jokes.
> Or is it true
> That you knew what to do

When n was greater than two?
Oh then why can't we find
That same proof ... are we blind?
You must be reproved, for I'm losing my mind.

7 Modern developments

We briefly list here a number of results about FLT – apart from those leading directly to Wiles' proof – obtained in the second half of the 20th century.

(a) *The age of the computer*

In 1973 Wagstaff proved that FLT holds for all exponents $p < 125,000$, and twenty years later Buhler, Crandall, Ernvall, and Metsänkylä pushed the result to $p < 4,000,000$. These proofs did not use merely the brute force of the computer, but were a mix of sophisticated theoretical mathematics combined with sophisticated use of computations. Specifically, methods were developed to determine the *irregular* primes up to the indicated limits, and subsequently FLT was shown to hold for these primes (recall that Kummer had established FLT for all *regular* primes) [7], [24], [32].

(b) *The Mordell Conjecture*

In 1922 Mordell conjectured that there are only finitely many points with rational coordinates on an algebraic curve of *genus* greater than one (for a definition of genus see [4], [10]). Gerd Faltings proved the conjecture in 1983 using high-powered methods of algebraic geometry, developed only in the second half of the 20th century. This was a major feat, for which Faltings was awarded the Fields Medal – the mathematical counterpart of the Nobel Prize. Now, the equation $x^n + y^n = z^n$ has genus 0 for $n = 2$ and genus greater than 1 for $n > 2$, so an immediate corollary of Mordell's Conjecture – now a theorem – is that *for each* $n > 2$ FLT has at most finitely many solutions [10], [25].

(c) *Miyoka*

Building on ideas of Faltings, and making connections between number theory and differential geometry, the Japanese mathematician Yoichi Miyoka announced in 1988 that he had proved FLT. Don Zagier, who was in the audience at the Max Planck Institute where Miyoka presented an outline of his proof, observed that "Miyoka's proof is very exciting, and some people feel that there is a very good chance that it is going to work. It's still not definite, but it looks fine so far" [28, p. 232]. Two months later Faltings found an error in the proof. Many notions in Miyoka's purported proof, however, remain important [10], [28].

8 Some major ideas leading to Wiles' proof of FLT

We are about to enter the promised land. A key breakthrough, which less than ten years later would lead to a proof of FLT, came in 1985, when Gerhard Frey related FLT to elliptic curves – "a most surprising and innovative link" [11, p. 3]. Specifically, if $a^p + b^p = c^p$ holds for nonzero integers a, b, c, the associated elliptic curve – now known as the *Frey Curve* – is $y^2 = x(x - a^p)(x + b^p)$.

Fig. 3 Andrew Wiles in his childhood

Number theory and geometry, in particular diophantine equations and geometry, have been associated for about two millennia. In fact, it has been argued that the methods of Diophantus (ca. 250 AD) for the solution of diophantine equations could be viewed as geometric; these came to be known as the "tangent and secant methods" [4].

An *elliptic curve* is a plane curve given by an equation of the form $y^2 = x^3 + ax^2 + bx + c$, where a, b, c are integers or rational numbers, and the cubic polynomial on the right side of the equation has distinct roots. (The coefficients may also be taken to be real or complex numbers, in fact elements in any field, although this is not of interest in our study.) A famous result of Siegel says that this equation has finitely many *integer* solutions, although it may have infinitely many *rational* solutions.

Elliptic curves had (in effect) been studied by Diophantus and Fermat, and intensively investigated by Euler and Jacobi [4]. The name "elliptic curves" reflects their connection with elliptic functions, studied deeply in the 19th century. (See [4], [27, p. 25], [30, p. 228] for an explanation of this connection, and [4, p. 77], [25, p. 148] for reasons why elliptic curves are important in number theory; "practical" uses of elliptic curves in the factoring of large integers into primes were found in the last few decades.) The significant connection of elliptic curves with FLT came to light only in the 1980s.

Before pursuing this connection, we want to present an elementary example of the use of geometry – the secant method – to solve a diophantine equation, namely to find all

solutions in integers of $x^2 + y^2 = z^2$ (cf. the *algebraic* solution of this equation in Section 4).

Divide both sides of the equation by z^2 to get $(x/z)^2 + (y/z)^2 = 1$. Solving $x^2 + y^2 = z^2$ in Z is equivalent to solving $u^2 + v^2 = 1$ in Q. Geometrically, this requires finding all points with rational coordinates on the unit circle; they are called *rational points* [4], [26].

Suppose we are given a fixed rational point on the unit circle (see figure 4 below), say $P(-1,0)$. If $Q(a,b)$ is any other rational point on the circle, the straight line determined by P and Q has rational slope. (The slope is $b/(a+1)$.) Conversely, any straight line through $(-1,0)$ with rational slope t intersects the circle in another rational point (a,b). So, to find all rational points on the unit circle is to find the points of intersection of all straight lines PQ $(P = (-1,0))$ having rational slope t.

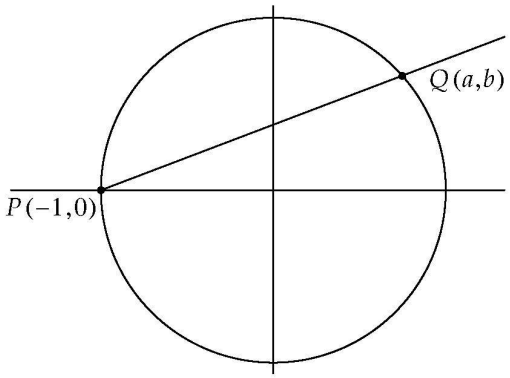

Fig. 4

We thus solve $u^2 + v^2 = 1$ and $v = t(u+1)$ simultaneously for u and v and get $u = (1-t^2)/(1+t^2), v = 2t/(1+t^2)$. Letting $t = m/n$, we find the integer solutions of $x^2 + y^2 = z^2$ to be $x = n^2 - m^2, y = 2nm, z = n^2 + m^2$.

The same method can be used to find all *rational* points on any quadratic curve, provided that we can find one rational point on the curve, and to find (at least in theory) all rational points on a *cubic* curve, provided we can find two rational points on the curve. The former problem is elementary, the latter is part of a rich theory; see [4], [26], [27].

Back to Frey's key idea, namely the association of the elliptic curve $y^2 = x(x-a^p)(x+b^p)$ with the equation $a^p + b^p = c^p$. Frey conjectured that if there are indeed integers a, b, c such that $a^p + b^p = c^p$, then the resulting elliptic curve $y^2 = x(x-a^p)(x+b^p)$ is "badly behaved": It is a counterexample to the so-called *Shimura-Taniyama Conjecture* (STC). Frey's conjecture, reformulated by Serre, is known as the *Epsilon Conjecture* (EC). Put positively, the Epsilon Conjecture says that if the Shimura-Taniyama Conjecture holds, then Fermat's Last Theorem is true.

The outline of a possible proof of FLT now emerged:

Fig. 5 Yutaka Taniyama

(a) Prove the EC, namely that the STC $\Longrightarrow$ FLT.

(b) Prove the STC.

The Shimura-Taniyama Conjecture was formulated by Taniyama in 1955 and subsequently (1960s) refined by his friend and colleague Shimura (and by Weil). It says that *every elliptic curve is modular*. The notion of modularity is technically difficult to define, "but essentially it means that there is a formula for the number of solutions of the curve's cubic equation in each finite number system" [9, p. 4]; see also [10], [14], [20], [25]. As for the STC, it "represents a deep connection between algebra and analysis" [25, p. 152]. The following two statements of Barry Mazur give a very good sense of its scope and depth:

> [The conjecture] plays a structural and deeply influential role in much of our thinking and our expectations in Arithmetic.... Although it is undeniably a conjecture 'about arithmetic', it can be phrased variously, so that in one of its guises, one thinks of it as being also deeply 'about' integral transforms in the theory of one complex variable; in another as being 'about' geometry [20, pp. 594, 596].

> It was a wonderful conjecture..., but to begin with it was ignored because it was so ahead of its time. When it was first proposed it was not taken up because it was so astounding. On the one hand you have the elliptic world, and on the other you have the modular world. Both these branches of mathematics had been

> studied intensively but separately.... Then along comes the Taniyama-Shimura conjecture, which is the grand surmise that there's a bridge between the two completely different worlds. Mathematicians love to build bridges [28, p. 190].

There are, of course, innumerable examples in mathematics of bridge building, among the best known and most important being that between algebra and geometry, viz. analytic geometry. In this paper we built bridges between number theory and algebra, and number theory and geometry.

The STC was not only most surprising, it was also very important – in the sense that if true, it had innumerable and very significant consequences [20]. So that a counterexample to the conjecture would have devastating consequences – much more severe than a counterexample to FLT! (Recall that the EC says that a counterexample to FLT would imply a counterexample to the STC.)

Enter Ken Ribet of the University of California at Berkeley. In 1986 he proved the Epsilon Conjecture. This was, of course, a big event. As Ribet relates [28, p. 201]:

> It was the crucial ingredient that I had been missing and it had been staring me in the face.... I was completely enthralled... I sort of casually mentioned to a few people [at the 1986 International Congress of Mathematicians in Berkeley] that I'd proved that the Taniyama-Shimura conjecture implies Fermat's Last Theorem. It spread like wildfire and soon large groups of people knew; they were running up to me asking, *Is it really true you've proved that Frey's elliptic equation is not modular?*

For a sketch of the ideas involved in Ribet's proof of the EC see [10]; see also [14], [28].

9 Andrew Wiles

For most of its 350-year history, Fermat's Last Theorem was not part of mainstream mathematics – in the sense that it had no *direct* link with important parts of mathematics. Ribet's proof of the Epsilon Conjecture changed all that. "What Ribet [did]", Wiles noted, "was to link Fermat's Last Theorem with a problem in mathematics [the STC] that would never go away" [8, p. 1133]. On hearing of Ribet's proof, Wiles was ecstatic [28, p. 205]:

> It was one evening at the end of the summer of 1986 when I was sipping tea at the house of a friend. Casually in the middle of a conversation he told me that Ken Ribet had proved the link between Taniyama-Shimura and Fermat's Last Theorem. I was electrified. I knew that moment that the course of my life was changing because this meant that to prove Fermat's Last Theorem all I had to do was to prove the Taniyama-Shimura conjecture. It meant that my childhood dream was now a respectable thing to work on. I just knew that I could never let that go. I just knew that I would go home and work on the Taniyama-Shimura conjecture.

Work he did on it – for the next seven years – in secret, which is most unusual in mathematics, though perhaps understandable under the circumstances. As Princeton colleague and then Chair of the Department Simon Kochen put it [17, p. 10]:

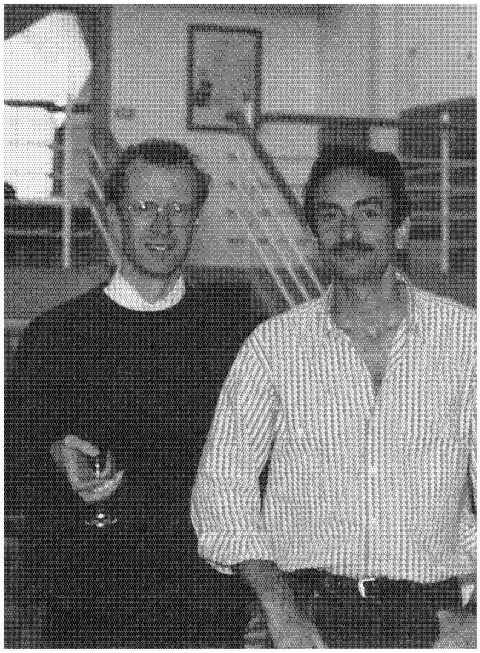

Fig. 6 Andrew Wiles and Ken Ribet

> If he [Wiles] said he was working on Fermat's Last Theorem, people would look askance. And if you start telling people who are experts, you end up collaborating with them. He wanted to do it on his own.

Here is some of what happened in the next seven years, as told by Wiles [17, p. 10]:

> I made progress in the first few years. I developed a coherent strategy. . .. Basically, I restricted myself to my work and my family. I don't think I ever stopped working on it. It was on my mind all the time. Once you're really desperate to find the answer to something, you can't let go.

Only in the 7th year did he bring into his confidence his Princeton colleague Nicholas Katz, "who agreed to serve as a sort of sounding board for Dr. Wiles" [17, p 10]. At the end of seven years, Eureka! [28, p. 244]:

> By May 1993 I was convinced that I had the whole of Fermat's Last Theorem in my hands. I still wanted to check the proof some more, but there was a conference which was coming up at the end of June in Cambridge, and I thought that would be a wonderful place to announce the proof – it's my old hometown, and I'd been a graduate student there.

The conference – on number theory – was organized by John Coates, Wiles' thesis advisor. It brought together some of the world's top experts in the subject. Wiles asked Coates to arrange for him to give a series of three lectures, one on each of the three-day

conference. The title of his proposed talks was "Elliptic curves, modular forms, and Galois representations" – no mention of FLT. Only during the third talk did it become apparent – to the experts in the audience – that a proof of FLT was the likely outcome of the talks. Ribet describes the historic event [28, p. 248]:

> I came relatively early and I sat in the front row with Barry Mazur. I had my camera with me just to record the event. There was a very charged atmosphere and people were very excited. We certainly had the sense that we were participating in a historic moment. . .. The tension had built up over the course of several days. There was this marvelous moment when we were coming close to a proof of Fermat's Last Theorem.

And from Harvard colleague Barry Mazur [28, p. 248]:

> I've never seen such a glorious lecture, full of such wonderful ideas, with such dramatic tension, and what a buildup. There was only one possible punch line.

Indeed, Wiles concluded his third lecture with the sentence: "And this proves Fermat's Last Theorem; I think I'll stop here" [28, p. 249]. Specifically, what Wiles did was prove the Shimura-Taniyama Conjecture for an important class of elliptic curves, the so-called *semi-stable* elliptic curves. (Roughly speaking, an elliptic curve is semi-stable if whenever a prime p divides the discriminant of the cubic defining the curve, exactly two of its roots are congruent modulo p; see [10], [25].) That is, he showed that every semi-stable elliptic curve is modular. (The *general* STC, open when Wiles gave his proof, has recently (June 1999) been proved [33].) Ribet had earlier proved a strong form of the Epsilon Conjecture, namely that if every semi-stable elliptic curve is modular, Fermat's Last Theorem is true. For a sketch of the ideas involved in Wiles' proof see [10], [14], [18], [19].

Wiles' work is very deep and technically very demanding. "The finished proof is still rough going even for the experts" [8, p. 1134]. The following two quotations give a sense of its profundity:

> By the end of the day, it was clear to experts around the world that nearly all of the noble and grand ideas that number theory had evolved over the past three and a half centuries since the time of Fermat were ingredients in the proof (R. Murty [21, p. 17]).

So, in a sense, Wiles' proof was a grand collaborative effort of dozens of mathematicians over several centuries!

> The work is extremely deep, involving the latest ideas from a score of different fields, including the theories of group schemes, crystalline cohomology, Galois representations, deformation theory, Gorenstein rings, (geometric) Euler systems and many others (Granville [21, p. 16]).

Behold the simplicity of the question and the complexity of the answer! The problem belongs to number theory – a question about positive integers. But what area does the proof come from? It is unlikely one could give a satisfactory answer, for the proof brings together many important areas – a characteristic of recent mathematics.

Wiles' lectures at Cambridge in June 1993 were, however, not to be the end of this 350-year odyssey. Here is some of what happened next.

The news of Wiles' proof of FLT electrified the mathematical world. E-mail messages started circulating incessantly. The news also made a great splash in the media – a rare event when it comes to mathematical news. Wiles' proof made the front pages of the *New York Times*. It was also featured in *Newsweek* and *Time*, and it made the *NBC Nightly News* that evening. *People* magazine listed Wiles among "the 25 most important people of the year".

After the celebrations were over, the business of checking the proof began. Wiles submitted a 200-page paper proving FLT to *Inventiones Mathematicae*. *Six* mathematicians were assigned to referee it – most unusual (normally there are 1-3 referees), but again, warranted under the circumstances. Many errors were found; most were easily and quickly corrected. One error, however, found by Katz, could not be fixed. But it was not divulged to the mathematical community – much was at stake! After some months, when no proof or announcement of an impending proof was forthcoming, rumors began to circulate. Was Wiles' proof destined for the same fate as Fermat's? Lamé's? Miyoka's?

On December 4, 1993, five months after his extraordinary announcement at Cambridge that he had proved FLT, Wiles issued the following e-mail note on a mathematical bulletin board [28, p. 264]:

> In view of the speculation on the status of my work on the Taniyama-Shimura conjecture and Fermat's Last Theorem, I will give a brief account of the situation. During the review process a number of problems emerged, most of which have been resolved, but one in particular I have not settled. . .. I believe that I will be able to finish this in the near future using the ideas explained in my Cambridge lectures.

In January 1994, on the advice of Princeton colleague Peter Sarnak, Wiles sought the help of Cambridge mathematician Richard Taylor, his former PhD student. The preceding and ensuing months must have been most trying for Wiles, as we can surmise, and as Simon Singh confirms [28, pp. 275, 265, 273]:

> The last fourteen months [July 1993–August 1994] had been the most painful, humiliating period of [Wiles'] mathematical career. . .. The pleasure, passion, and hope that carried him through the years of secret calculations were replaced with embarrassment and despair. . .. After eight years of unbroken effort and a lifetime's obsession, Wiles was prepared to admit defeat. He told Taylor that he could see no point in continuing with attempts to fix the proof. . .. Taylor . . . suggested they persevere for one more month.

"On September 19, 1994, they [Wiles and Taylor] found the vital fix" [29, p. 73]. Wiles recalls the clinching insight [29, p. 73]:

> It was so incredibly beautiful; it was so simple and so elegant. The first night I went back home and slept on it. I checked through it again the next morning, and I went down and told my wife, 'I've got it. I think I've found it'. And it was so

> unexpected she thought I was talking about a children's toy or something, and she said, 'Got what?' I said, 'I've fixed my proof. I've got it.'

On October 25, 1994, two papers proving FLT were released for publication, one by Wiles, the other by Taylor and Wiles, as follows:

(i) A. Wiles: Modular elliptic curves and Fermat's Last theorem, *Annals of Mathematics* 142 (1995) 443–551.

(ii) R. Taylor and A. Wiles: Ring-theoretic properties of certain Hecke algebras, *Annals of Mathematics* 142 (1995) 553–572.

10 Tributes

The 1994 International Congress of Mathematicians (ICM) was held in Zürich in August. Had Wiles filled the gap in his proof of FLT prior to the start of the Congress, he would undoubtedly have received the Fields Medal at the Congress. At the next ICM in Berlin, in 1998, he was not eligible for the medal, being over 40. He was, however, awarded a one-time Special Tribute – the "International Mathematical Union Silver Plaque". On June 27, 1997, he collected the Wolfskehl prize – ten years before its expiry and now worth $50,000 (recall that in 1907 it was valued at $1,000,000).

The following appreciations of Wiles' work come from some of the foremost experts in the subject:

> To complete his [proof] Wiles needed to draw on and further develop many modern ideas in mathematics. In particular, he had to tackle the Shimura-Taniyama conjecture, an important 20th-century insight into both algebraic geometry and complex analysis. In doing so, Wiles forged a link between these major branches of mathematics. Henceforth insights from either field are certain to inspire new results in the other. Moreover, now that this bridge has been built, other connections between distant mathematical realms may emerge (Singh and Ribet [29, p. 68]).

> In mathematical terms, the final proof is the equivalent of splitting the atom or finding the structure of DNA. A proof of Fermat is a great intellectual triumph, and one shouldn't lose sight of the fact that it has revolutionized number theory in one fell swoop. For me, the charm and beauty of Andrew's work has been that it has been a tremendous step for number theory (Coates [28, p. 279]).

> Fermat's Last Theorem deserves a special place in the history of civilization. By its simplicity it has tantalized amateurs and professionals alike, and with remarkable fecundity led to the development of many areas of mathematics such as algebraic geometry, and more recently the theory of elliptic curves and representation theory. It is truly fitting that the proof crowns an edifice composed of the greatest insights of modern mathematics (R. Murty [21, p. 20]).

This statement surely belies Gauss' claim that FLT was not an interesting problem to work on! Even the greatest among mathematicians can misjudge.

The last word belongs to Wiles [28, p. 285]:

> I had this very rare privilege of being able to pursue in my adult life what had been my childhood dream. I know it's a rare privilege, but if you can tackle something in adult life that means that much to you, then it's more rewarding than anything imaginable. Having solved this problem, there's certainly a sense of loss, but at the same time there is this tremendous sense of freedom. I was so obsessed by this problem that for eight years I was thinking about it all the time – when I woke up in the morning to when I went to sleep at night. That's a long time to think about one thing. That particular odyssey is now over. My mind is at rest.

References

[1] A. D. Aczel, *Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem*, Four Walls Eight Windows, 1996.

[2] W. W. Adams and L. J. Goldstein, *Introduction to Number Theory*, Prentice-Hall, 1976.

[3] K. Barner, Paul Wolfskehl and the Wolfskehl Prize, *Notices of the AMS* 44 (1997) 1294–1303.

[4] I. G. Bashmakova, *Diophantus and Diophantine Equations*, translated from the Russian by A. Shenitzer, Math. Assoc. of America, 1997.

[5] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press, 1966.

[6] N. Bourbaki, *Elements of the History of Mathematics*, Springer-Verlag, 1994.

[7] J. Buhler, R. Crandell, R. Ernvall, and T. Metsänkylä, Irregular primes and cyclotomic invariants to four million, *Math. Comp.* 61 (1993) 151–153.

[8] B. Cipra, Princeton mathematician looks back on Fermat proof, *Science* 268 (26 May 1995) 1133–1134.

[9] B. Cipra, "A truly remarkable proof", in *What is Happening in the Mathematical Sciences*, Amer. Math. Society, Vol. 2, 1994, pp. 3–7.

[10] D. A. Cox, Introduction to Fermat's Last Theorem, *Amer. Math. Monthly* 101 (1994) 3–14.

[11] K. Devlin, F. Gouvêa, and A. Granville, Fermat's Last Theorem, a theorem at last, *MAA Focus* 13 (August 1993) 3–4.

[12] J. P. Dowling, Fermat's Last Theorem, *Math. Mag.* 59 (1986) 76.

[13] H. M. Edwards, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer-Verlag, 1977.

[14] F. Gouvêa, "A marvellous proof", *Amer. Math. Monthly* 101 (1994) 203–222.

[15] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag, 1982.

[16] I. Kleiner, The roots of commutative algebra in algebraic number theory, *Math. Mag.* 68 (1995) 3–15.

[17] G. Kolata, Andrew Wiles: A math whiz battles 350-year-old puzzle, *Math. Horizons* (Winter 1993), 8–11.

[18] J. Kramer, Über die Fermat-Vermutung, *El. Math.* 50 (1995) 12–25.

[19] J. Kramer, Über den Beweis der Fermat-Vermutung II, *El. Math.* 53 (1998) 45–60.

[20] B. Mazur, Number theory as gadfly, *Amer. Math. Monthly* 98 (1991) 593–610.

[21] R. Murty, A long-standing mathematical problem is solved: Fermat's Last Theorem, *Can. Math. Soc. Notes* 25 (Sept. 1993) 16–20.

[22] I. Peterson, Cracking Kepler's sphere-packing problem, *Science News* 154 (August 15, 1998) 103.

[23] H. Pollard and H. G. Diamond, *The Theory of Algebraic Numbers*, 2nd ed., Math. Assoc. of America, 1975.

[24] P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, 1979.

[25] K. A. Ribet and B. Hayes, Fermat's Last Theorem and modern arithmetic, *American Scientist* 82 (March-April, 1994) 144–156.

[26] J. H. Silverman, *A Friendly Introduction to Number Theory*, Prentice-Hall, 1997.

[27] J. H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, 1992.

[28] S. Singh, *Fermat's Enigma: The Quest to Solve the World's Greatest Mathematical Problem*, Penguin, 1997.

[29] S. Singh and K. Ribet, Fermat's last stand, *Scien. Amer.* 277 (November 1997) 68–73.

[30] J. Stillwell, *Mathematics and Its History*, Springer-Verlag, 1989.

[31] A. van der Poorten, *Notes on Fermat's Last Theorem*, Wiley, 1996.

[32] S. M. Wagstaff, The irregular primes to 125000, *Math. Comp.* 32 (1978) 583–591.

[33] H. Darmon, A proof of the full Taniyama-Shimura-Weil Conjecture is announced, *Notices of the Amer. Math. Soc.* 46 (1999), 1397–1401.

Israel Kleiner
Department of Mathematics and Statistics
York University
Toronto, Ontario MJ3 1P3
Canada
e-mail: kleiner@home.com