Fermat's Last Theorem

Ernst Kani

Coleman-Ellis Lecture, November 1996

Introduction

On the 23rd of June 1993, Andrew Wiles concluded a three-day lecture series in Cambridge, England, with the assertion:

Theorem. *Every semi-stable elliptic curve is modular.*

This not only electrified number theorists and mathematicians around the world, but even made the headlines of many major newspapers such as the *New York Times, Le Monde, Frankfurter Allgemeine,...*, a rare event for a mathematical theorem.

The main reason for this excitement and publicity is due to the fact that it had just been shown a few years earlier that the above theorem implies the truth of *Fermat's Last Theorem*,

$$(\mathrm{FLT}_n) \qquad x^n+y^n \neq z^n, \quad xyz \neq 0,$$

for any non-zero integers $x, y, z \in \mathbb{Z}$ and $n \geq 3$; this had been asserted by Fermat 350 years ago!

The purpose of this lecture is to relate some of the history behind FLT (= Fermat's Last Theorem[1]), to explain in simple terms how Wiles's theorem is related to FLT and, above all, to give you a glimpse of the significance of Wiles's result which, in fact, goes far beyond FLT.

1. Early History

Although FLT is an assertion about sums of n-th powers for $n \geq 3$, it was inspired by looking at the case $n = 2$, the so-called Pythagorean equation:

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

[1]So called because it was the last of Fermat's many assertions which still had to be resolved.

In high school, every student learns that $(3, 4, 5)$ and $(5, 12, 13)$ are solutions (called *Pythagorean triplets*) of this equation, but few learn that

$$12,709^2 + 13,500^2 = 18,541^2.$$

Indeed, this solution, and many others like it, had been known for almost 4000 years, and were recorded on clay tablets around the era of Hammurabi (ca. 1700 B.C.), more than 1000 years before Pythagoras (ca. 550 B.C.). In fact, from the way the following tablet (Plimpton 322, discovered by O. Neugebauer and Sachs; cf. Figure 1) is arranged, historians are convinced that the Babylonians already knew the following formula (or something close to it) for generating all Pythagorean triplets:

$$(1) \quad x = u^2 - v^2, \quad y = 2uv, \quad z = u^2 + v^2,$$

where $u, v \in \mathbb{Z}$; this formula is usually attributed to Pythagoras or Plato (ca. 400 B.C.).

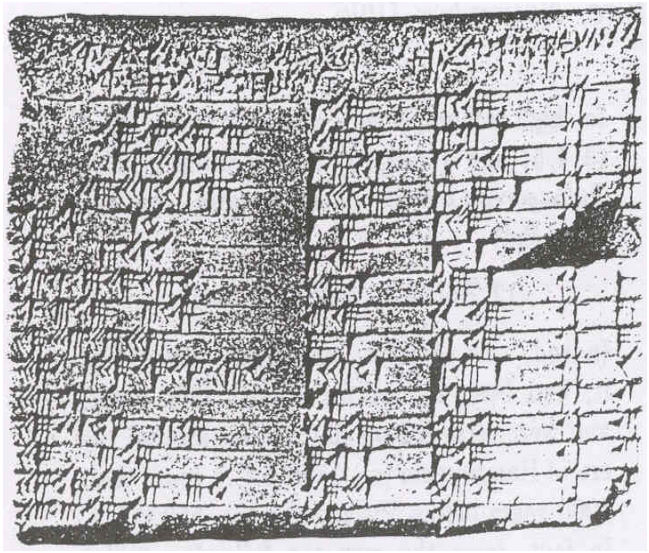
Certainly *Diophantus of Alexandria* (ca. 250 A.D.) was not only aware of this formula, but even based a large number of problems on it in his *Arithmetica*, a very remarkable collection of 13 books of which 9 have survived. (Of these, only 6 were known in the Renaissance; the other 3 were discovered only 20 years ago in a library in Iran.) Thus we find in Book II:

Problem 8: *Decompose a given square into a sum of two squares.*

Diophantus presents the numerical example $4^2 = \left(\frac{16}{5}\right)^2 + \left(\frac{12}{5}\right)^2$, but his method is perfectly general and actually leads to the formula

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

Plimpton 322
ca. 1800 - 1650 B.C.



$\frac{d^2}{h^2}$	w	d	n	h
$\frac{28561}{14400}$	119	169	1	120
$\frac{23280625}{11943936}$	3367	4825	2	3456
$\frac{44209201}{23040000}$	4601	6649	3	4800
$\frac{343768681}{182250000}$	12709	18541	4	13500
$\frac{9409}{5184}$	65	97	5	72
$\frac{231361}{129600}$	319	481	6	360
$\frac{12538681}{7290000}$	2291	3541	7	2700
$\frac{8667}{5120}$	799	1249	8	960
$\frac{591361}{360000}$	481	769	9	600
$\frac{66601921}{41990400}$	4961	8161	10	6480
$\frac{25}{16}$	45	75	11	60
$\frac{8579041}{5760000}$	1679	2929	12	2400
$\frac{83521}{57600}$	161	289	13	240
$\frac{10426441}{7290000}$	1771	3229	14	2700
$\frac{2809}{2025}$	56	106	15	90

Figure 1: A clay tablet and its translation:[2] Pythagorean triplets $h^2 + w^2 = d^2$

where a^2 is the square to be decomposed and m is any integer. This, of course, is just a variant of the formula (1).

While studying this problem, Pierre De Fermat (1601 -1665) wrote the following text in the margin of his copy of the *Arithmetica* (which had recently been translated from Greek to Latin by Bachet):

Cubum autem in duos cubos, aut quadrato-quadratum in duos quadrato-quadratos, et generaliter nullam in infinitum ulta quadratum postestatem in duos ejusdem nominis fas est dividere; cujus rei demonstrationem mirabilem sane dexteri. Hanc marginis exiguitas non caperet.

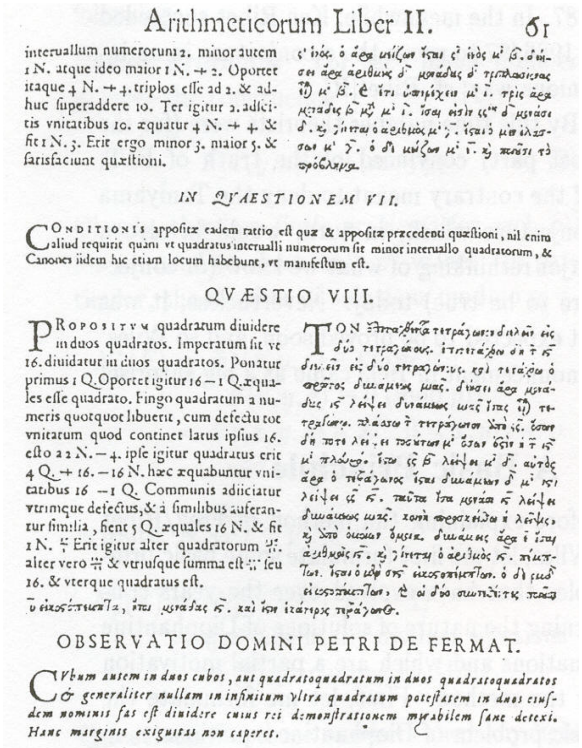
Translation[He]: On the other hand it is impossible to separate a cube into two cubes, or a biquadrate into two biquadrates, or generally any power except a square into two powers with the same exponent. I have discovered a truly marvellous proof of this, which however the margin is not large enough to contain.

[2]This translation includes some corrections. In addition, the column h, which does not appear in the original, was added for convenience.

We do not know the exact date of this entry, but in 1638 he challenged Jumeau de Saint-Croix to find two cubes whose sum is a cube (and similarly for biquadrates), so it seems likely that he became convinced of the truth of FLT around that time.

Fermat himself gave a proof of FLT for $n = 4$ which he wrote in the margin at the end of the last book of Diophantus. However, since he does not seem to refer to this conjecture in his correspondence (except in the case $n = 4$), it might well have been lost to posterity had not his son Samuel published in 1670 another edition of Diophantus, interspersed with his father's comments (cf. Figure 2).

Now that I have dwelt in such detail on the birth of the conjecture, I will be much briefer with subsequent early developments. The case $n = 3$ was done by L. Euler in 1753 (with some additional details furnished later by C.F. Gauss). In 1825/28 Dirichlet and Legrendre (independently) settled the case $n = 5$ and in 1832 Dirichlet also did the case $n = 14$. The latter result became superfluous when G. Lamé proved FLT for $n = 7$ in 1839. Later, in 1847, Lamé also presented



Arithmeticorum Liber II. 61

interuallum numerorum 2. minor autem 1 N. atque ideo maior 1 N. + 2. Oportet itaque 4 N. + 4. triplos esse ad 2. & adhuc superaddere 10. Ter igitur 2. adiicitis vnitatibus 10. æquatur 4 N. + 4. & fit 1 N. 3. Erit ergo minor 3. maior 5. & satisfaciunt quæstioni.

IN QVAESTIONEM VII.

CONDITIONIS appositæ eadem ratio est quæ & appositæ præcedenti quæstioni, nil enim aliud requirit quàm vt quadratus interualli numerorum sit minor interuallo quadratorum, & Canones iidem hic etiam locum habebunt, vt manifestum est.

QVÆSTIO VIII.

PROPOSITVM quadratum diuidere in duos quadratos. Imperatum sit vt 16. diuidatur in duos quadratos. Ponatur primus 1 Q. Oportet igitur 16 − 1 Q. æquales esse quadrato. Fingo quadratum a numeris quotquot libuerit, cum defectu tot vnitatum quod continet latus ipsius 16. esto 2 a N. − 4. ipse igitur quadratus erit 4 Q. + 16. − 16 N. hæc æquabuntur vnitatibus 16 − 1 Q. Communis adiiciatur vtrimque defectus, & à similibus auferantur similia, fient 5 Q. æquales 16 N. & fit 1 N. $\frac{16}{5}$. Erit igitur alter quadratorum $\frac{256}{25}$. alter vero $\frac{144}{25}$ & vtriusque summa est $\frac{400}{25}$ seu 16. & vterque quadratus est.

OBSERVATIO DOMINI PETRI DE FERMAT.

CVbum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos & generaliter nullam in infinitum vltra quadratum potestatem in duos eiusdem nominis fas est diuidere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

QVÆSTIO IX.

RVRSVS oporteat quadratum 16 diuidere in duos quadratos. Ponatur rursus primi latus 1 N. alterius verò quotcunque numerorum cum defectu tot vnitatum, quot constat latus diuidendi. Esto itaque 2 N. − 4. erunt quadrati, hic quidem 1 Q. ille verò 4 Q. + 16. − 16 N. Cæterum volo vtrumque simul æquari vnitatibus 16. Igitur 5 Q. + 16. − 16 N. æquatur vnitatibus 16. & fit 1 N. $\frac{32}{5}$. erit

Figure 2: Fermat's comment in the margin of the *Arithmetica* by Diophantus

a proof for general primes $p \geq 3$ based on the arithmetic of the number ring (called the *ring of cyclotomic integers*)

$$\mathbb{Z}[\zeta_p] = \{n_0 + n_1\zeta_p + \ldots + n_{p-2}\zeta_p^{p-2} : n_i \in \mathbb{Z}\},$$

where $\zeta_p = e^{\frac{2\pi i}{p}} = \cos(\frac{2\pi}{p}) + i\sin(\frac{2\pi}{p})$, but this proof was wrong since he assumed that the property of unique factorization holds in $\mathbb{Z}[\zeta_p]$ for all p. (In fact, it is now known that this holds if and only if $p \leq 19$.)

The most notable advance was made by Ernst Edward Kummer. Already a year before Lamé's hasty announcement, he had invented his "ideal numbers" (precursors of Dedekind's theory of ideals, which we use today) in order to rescue the unique factorization property for $\mathbb{Z}[\zeta_p]$. In 1847 (published 1850) he applied his methods to prove that (FLT_p) is true for all *regular* primes p. Furthermore, he devised a method (based on Bernoulli numbers) in order to test whether a given prime number is regular or not. Using this, he found in 1874 that of the 37 primes $p < 163$, only 8 are irregular: $p = 37, 59, 67, 101, 103, 131, 149,$ and 157.[3]

After Kummer, there were many partial results on FLT for which I refer you to P. Ribenboim's excellent book [Ri1]. By combining these with Kummer's general result, S. Wagstaff was able to verify in 1976 with the help of a computer that Fermat's Last Theorem is true for all integers $n \leq 125,000$.

2. Recent Results

In the 1970's and '80's, the deep and powerful methods of Algebraic Geometry (as developed by A. Grothendieck and his school) lead to many significant advances in the theory of *Diophantine Equations* (named after Diophantus). Most of these were not (or did not seem to be) directly connected with FLT. A notable exception was the *Mordell Conjecture*, which had been formulated by L.J. Mordell in 1922, and which was then established through the work of G. Faltings in 1983, for which he received the Fields Medal[4] in 1986. Specialized to Fermat equations, Falt-

[3] This evidence suggests that there are more regular primes than irregular ones, and this is borne out by further calculations (e.g. 8399 of 13848 primes $p < 150,000$ are regular; cf. [Ri2]). In addition, C.L. Siegel proved in 1964 that if certain (unproven) random distribution property of Bernoulli numbers holds, then $\frac{1}{\sqrt{e}} \approx 0.61$ of all primes are regular. However, unconditionally it is still unknown whether there exist infinitely many regular primes (whereas it was shown in 1915 by Jensen that there are infinitely many irregular primes).

[4] The Fields Medal, named after the Canadian mathematician John Charles Fields (1863 - 1932), is the most prestigious prize for mathematical research. It is awarded every 4 years at the International Congress of Mathematicians to the top 2-4 researchers under the age of 40.

ings' theorem yields the following finiteness result.

Theorem (Faltings, 1983) *For each $n \geq 4$, the set*
$\{(x, y, z) \in \mathbb{Z}^3 : x^n + y^n = z^n \text{ and } (x, y, z) = 1\}$
is finite.

While this was clearly a very significant result (particularly in its more general form), it did not convince skeptics about FLT. Indeed, there did not seem to be *any (conceptual) reason whatsoever* that the equation $x^p + y^p = z^p$ should have only the same five solutions for all primes $p > 2$!

This changed drastically in the mid 1980's when not only one but two separate reasons were advanced. On the one hand D. Masser(1985) and J. Oesterlé(1988) proposed a very remarkable general conjecture (called the *ABC-Conjecture*) from which it would follow that not only the Fermat equation but also the twisted Fermat equation $ax^n + yb^n = zc^n$ (where $a, b, c \in \mathbb{Z}$ are fixed relatively prime integers) has only "trivial" solutions (in particular, only finitely many solutions) for all *sufficiently large* exponents n. This conjecture, as well as the statement about twisted Fermat equations (called the *asymptotic Fermat Conjecture*) is still open at present.[5]

On the other hand, in a Paris seminar in 1985, G. Frey suggested a method (based on some (vague) conjectures of Serre) that a certain well-known conjecture, called the *Taniyama Conjecture* (or (TWS)- Conjecture), should imply FLT. I cannot resist the temptation of relating a personal anecdote about this discovery. Indeed, I can still remember the day (but not the date - probably in the spring of 1982) when Gerd Frey, who is a good friend of mine, phoned me up and said: "I've just proved FLT, can you find the mistake?" Of course I couldn't, but after giving me an hour lecture he himself saw that there were a number of gaps to be filled. These gaps were formulated in terms of a precise conjecture by J.P. Serre in a letter to Frey in 1985 and became known as the "ε-Conjecture"; this was published as part of a far more general conjecture by Serre[Se] in 1987. In the meanwhile, Ken Ribet succeeded in 1986/87 to prove the ε-conjecture in an ingenious way; cf. Ribet[R].

By this time number theorists were (for the most part) convinced of the truth of FLT, for the contrary meant to deny the Taniyama Conjecture which, in turn, would involve a major rethinking of what we know (or conjecture to be true) today. Nevertheless, it was not expected to be proved soon, and so Wiles' announcement in 1993 came as a big surprise!

[5]For comprehensive discussion of how these and other conjectures fit together, cf. Frey[Fr3].

3. A Basic Principle

Before explaining the method of Frey/Ribet/Wiles, let me first formulate some basic principles that have evolved over the years concerning the nature of solutions of Diophantine equations and which are a partial motivation for the method. First, let me formulate the basic problem of Diophantine equations:

Problem: Find all the integer solutions $(x, y, z) \in \mathbb{Z}^3$ of a given Diophantine equation

$$F(x, y, z) = 0, \tag{2}$$

where $F \in \mathbb{Z}[x, y, z]$ is an integral polynomial.

Examples: 1) Fermat polynomials:

$$F(x, y, z) = F_n(x, y, z) = x^n + y^n - z^n.$$

2) Elliptic curves:

$$F_{a,b}(x, y, z) = y^2 z - x^3 + axz^2 + bz^3,$$

where $a, b \in \mathbb{Z}$ and the discriminant $\Delta(F_{a,b}) = 16(4a^3 + 27b^3) \neq 0$.

To give you an impression of the difficulty of this problem, let me remark that at present **no general algorithm is known** which decides in a finite amount of time whether a given polynomial $F(x, y, z)$ has at least one non-trivial integer solution $(x, y, z) \neq (0, 0, 0)$

or not,[6] let alone an algorithm that finds all the solutions! Let us, therefore, consider the following

Easier Problem: For each prime number p, solve the congruence

$$F(x,y,z) \equiv 0 \pmod{p}. \tag{3}$$

Clearly, this is a *finite problem* (for each p), for we need to check only p^3 values. In particular, the number of solutions modulo p,

$$\begin{aligned} N_p^*(F) &= \#\{(x,y,z) \in (\mathbb{Z}/p\mathbb{Z})^3 : \\ &\qquad F(x,y,z) \equiv 0 \pmod{p}\} \\ &= \#\{(x,y,z) \in \mathbb{Z}^3 : 0 \le x,y,z < p \\ &\qquad \text{and } p | F(x,y,z)\}, \end{aligned}$$

is finite: $N_p^*(F) \le p^3$. Put:

$$\begin{aligned} N_p(F) &= (N_p^*(F)-1)/(p-1) \\ &= \#\text{of essentially distinct solutions} \\ &\quad \text{of (3) (excluding (0,0,0))}. \end{aligned}$$

Question: Do these numbers shed any light on the solutions of equation (2)?

The naive interpretation of this question is blatantly false: there exist polynomials $F(x,y,z)$ with only trivial integral solutions, yet $N_F(p) \neq 0$ for all primes p. In addition, it follows from a theorem due to H. Hasse and A. Weil that $N_F(p) \approx p$, for p large, so the mere existence of solutions modulo p cannot yield any information about the existence of integral solutions. Nevertheless, we have the following

Basic (Conjectural) Principle: the sequence of numbers

$$a_p(F) \stackrel{def}{=} (p+1) - N_p(F), \text{ as } p \to \infty, \tag{4}$$

should determine the nature of the solutions of (2).

For elliptic curves, this principle assumes the form of two very precise conjectures which have been partly verified:

[6] In fact, it is known that for integer polynomials $F(x_1,\dots,x_r)$ in $r \ge 13$ variables, no such algorithm can exist, as was shown by Matijasevič in 1970, thereby supplying a negative answer to Hilbert's 10th problem; cf. [DMR].

(TWS)–Conjecture: - due to Y. Taniyama (1955), A. Weil (1967), G. Shimura (1971)

(B/SwD)–Conjecture: - B. Birch, H.P.F. Swinnerton–Dyer (1960's)

The (TWS)-Conjecture will be explained in the next section. I will not discuss the (B/SwD)- conjecture in detail here, but only mention the following recent result (which at the same time shows the importance of the (TWS)-conjecture):

Theorem 1 (V. A. Kolyvagin (1988), K. Murty, R. Murty (1991)).[7] *Let E : $F_{a,b}(x,y,z) = 0$ be an elliptic curve satisfying (TWS). Then the sequence of numbers*

$$a_p(E) = p + 1 - N_p(F_{a,b}), \quad p \to \infty,$$

determines a ("computable") real constant $L_E(1) \in \mathbb{R}$. If

$$L_E(1) \neq 0,$$

then the equation $F_{a,b}(x,y,z) = 0$ has only finitely many integral solutions $(x,y,z) \in \mathbb{Z}^3$ with $gcd(x,y,z) = 1$, and these can be explicitly calculated.

Note. The above theorem constitutes an explicit algorithm which has been implemented on a MAPLE package called APECS.

Example (Frey). The above leads to a *computer proof* (a true proof!) of FLT_3 and FLT_4, using only *four* short computer commands.

4. The TWS–Conjecture

Roughly speaking, the TWS-Conjecture may be viewed as stating that the numbers $a_p(E)$ possess many "hidden symmetries"; in particular, the knowledge of the a_p's for the first few p's determines *all the others*.

Before explaining this more precisely, let us look at the elliptic curve E defined by the equation

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

[7] This theorem was first proven by Kolyvagin under an additional hypothesis, which was then later removed by Murty-Murty and, independently, by D. Bump, S. Friedberg and J. Hoffstein.

The Elliptic Curve $E : y^2 + y = x^3 - x^2$

The number $N_p(E)$ of solutions of E over $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ and the number $a_p(E) = p+1-N_p(E)$ are given by:

p	2	3	5	7	11	13	17	19	23	29	31	37	41
$N_p(E)$	5	5	5	10	11	10	20	20	25	30	25	35	50
$a_p(E)$	-2	-1	1	-2	1	4	-2	0	-1	0	7	3	-8

On the other hand, the unique newform $f(z) \in S_2(\Gamma_0(11))$ of level 11 is:

$$\begin{aligned} f(z) = q\prod_{n=1}^{\infty}(1-q^n)^2(1-q^{11n})^2 \;\; &= \;\; \sum_{n=1}^{\infty} a_n(f)q^n \\ = q - 2\,q^2 - q^3 + 2\,q^4 + q^5 + 2\,q^6 - 2\,q^7 - 2\,q^9 - 2\,q^{10} + q^{11} - 2\,q^{12} + 4\,q^{13} + 4\,q^{14} \\ -q^{15} - 4\,q^{16} - 2\,q^{17} + 4\,q^{18} + 2\,q^{20} + 2\,q^{21} - 2\,q^{22} - q^{23} - 4\,q^{25} - 8\,q^{26} + 5\,q^{27} \\ -4\,q^{28} + 2\,q^{30} + 7\,q^{31} + 8\,q^{32} - q^{33} + 4\,q^{34} - 2\,q^{35} - 4\,q^{36} + 3\,q^{37} - 4\,q^{39} - 8\,q^{41} + \ldots \end{aligned}$$

Its first few Fourier coefficients at prime indices are:

p	2	3	5	7	11	13	17	19	23	29	31	37	41
$a_p(f)$	-2	-1	1	-2	1	4	-2	0	-1	0	7	3	-8

In this case, the numbers $a_p(E)$ have a very remarkable interpretation: each turns out to be equal to the p-th Fourier coefficient of the function f defined by product expansion

$$f(z) = q\prod_{n=1}^{\infty}(1-q^n)^2(1-q^{11n})^2,$$

where $q = e^{2\pi iz}$ (see the insert on the top of this page). Now it can be shown that this function has many "hidden symmetries", i.e. it satisfies the transformation law (5) below (with $N = 11$), and that this characterizes the function f uniquely.

This phenomenon can be generalized to arbitrary elliptic curves, but for this we need the following two concepts:

1) The *conductor* $N = N_E$ of an elliptic curve $E = E_{a,b}$: this is a positive integer

$$N \mid \Delta_{a,b}$$

which is closely related to the discriminant $\Delta_{a,b}$ (and which is explicitly computable).

2) The space $S(N) = S_2(\Gamma_0(N))$ of *modular forms of level* N: this consists of (complex-valued) functions of the form

$$f(z) \;\; = \;\; \sum_{n=1}^{\infty} a_n(f)q^n, \quad \text{with } q = e^{2\pi iz},$$

where the $a_n(f) \in \mathbb{C}$ and the sum converges for $Im(z) > 0$; these are to satisfy certain additional properties such as the rule

$$(5) \qquad f\left(\frac{az+b}{cz+d}\right) \;\; = \;\; (cz+d)^2 f(z),$$

where $a, b, c, d \in \mathbb{Z}$ are any integers with $ad - bc = 1$ and $N|c$.

Properties: 1) $S(N)$ is a finite-dimensional $\mathbb{C}$-vector space. There is an explicit formula for its dimension $g_N := \dim_{\mathbb{C}} S(N)$, which is approximately $g_N \approx \frac{N}{12}$.

2) Each $f \in S(N)$ is uniquely described by its first $2g_N \approx \frac{N}{6}$ Fourier coefficients $a_1(f), \ldots, a_{2g_N}(f)$.

3) The space $S(N)$ has a *distinguished* $\mathbb{C}$-basis $\mathfrak{B}(N) = \mathfrak{B}^+(N) \cup \mathfrak{B}^-(N)$. The functions in $\mathfrak{B}^+(N)$ are called *newforms*, those in $\mathfrak{B}^-(N)$ *oldforms*. For each N, these forms are explicitly computable (and have been computed for $N \leq 10^6$).

The above properties show that for each N, the set of functions $\mathfrak{B}(N)$ is determined by a finite amount of data, and hence may be viewed as being explicitly known. The (TWS)-Conjecture relates the Diophantine numbers $a_p(E)$ to these functions as follows.

Conjecture (TWS): *For every elliptic curve E of conductor N, there is a (unique) newform $f(z) = \sum a_n(f)q^n \in \mathfrak{B}^+(N)$ of level N such that*

$$(6)\quad a_p(E) = a_p(f), \quad \text{for all primes } p \not| N.$$

At first sight, this seems to be a rather daring and mysterious conjecture: why should the numbers $a_p(E)$ have anything to do with modular forms?

The first major piece of evidence for this conjecture was provided by A. Weil who showed in 1967 that its falsity would contradict a main principle of Number Theory (the principle that certain arithmetically defined functions (called L-functions) should have a functional equations). Shortly thereafter, G. Shimura[Sh] showed that the converse to the conjecture is in fact true:

Theorem 2 (Shimura, 1971). *For each $f \in \mathfrak{B}^+(N)$ with integral Fourier coefficients there is an elliptic curve E (of conductor N) such that (6) holds.*

Although this result provides us with many explicit (numerical) examples for which the (TWS)-Conjecture is true, it is too weak to prove that there are infinitely many elliptic curves which satisfy (TWS), for there is no way to guarantee that there any modular forms *with integral coefficients* for large N. This, however, and much more, follows from the important theorem proven by Wiles[W] (with the help of R. Taylor[8]):

Theorem 3 (Wiles, 1995). *The conjecture (TWS) is true if N_E is squarefree.*[9]

As should be evident from the above discussion, Wiles's result goes much further than merely proving (FLT): it should be viewed as an important step towards realizing the goal of finding a general algorithm for solving Diophantine problems involving elliptic curves.

5. $\mathbf{TWS}_{ss} \Rightarrow \mathbf{FLT}$

Although the work of Wiles[10] clearly advances our understanding of the arithmetic of elliptic curves, it is less evident how it relates to FLT, and indeed, the deduction of FLT from Theorem 3 constitutes another major step in the proof of FLT. Here is a brief sketch of the ideas involved:

Proof of $\mathbf{TWS}_{ss} \Rightarrow$ FLT: Since FLT_3 and FLT_4 are known to be true, it is elementary to see that we can restrict attention to primes $p \geq 5$.

Suppose, therefore, that FLT_p is false, i.e. that there exist $a, b, c \in \mathbb{Z}$ with $abc \neq 0$ such that

$$a^p + b^p = c^p.$$

[8] The original proof of Wiles and Taylor is 130 pages long, and fills an entire issue of the Annals. Since its publication, a number of simplifications have been suggested by a number of people such as G. Faltings, H. Lenstra and F. Diamond; cf. [Di]. For an overview of the original proof, together with a lot of background information, the reader is encouraged to consult [DDT].

[9] Recently (February, 1997), Conrad, Diamond and Taylor have announced that they can prove that (TWS) is true as long as 27 does not divide N_E.

[10] Due to the age restriction, Wiles just missed getting the prestigious Fields Medal for his work. However, he has received many other awards, including an *Honourary Doctorate* from Queen's University in May 1997.

By interchanging a and b we may suppose without loss of generality that $2|a$, and so we have in particular that $16|a^p$. Consider the elliptic curve

$$E: \qquad y^2z \;=\; x(x-a^pz)(x+b^pz),$$

called a *Frey curve.*[11] Then:

1) $\Delta = (abc)^{2p}$

2) N_E is squarefree (this uses the fact that $16|a^p$).

Thus, by Wiles's theorem, there is an $f = f_E \in \mathfrak{B}^+(N_E)$ such that (6) holds.

Claim: Such an f_E does not exist!

The verification of this claim is really the heart of the proof. For this, Ribet[R] proves the following "Lowering the Level Principle" (also known as Serre's ε-Conjecture) which is a special case of Serre's general conjecture (cf. [Se]):

Theorem 4 ("Lowering the Level" - Ribet, 1991). *Suppose $f = f_E \in \mathfrak{B}^+(N)$ is a newform of level N. For a fixed prime number $p > 3$ let M_p denote the product of the prime numbers $q > 2$ such that $p|expt_q(\Delta_E)$. Then there exists $g \in \mathfrak{B}^+(N/M_p)$ such that*

$$a_n(g) \equiv a_n(f) \pmod{p},$$

for all $n \geq 1$ with $\gcd(n, N) = 1$.[12]

Conclusion. Apply this to f_E as above. Then by 1) we obtain that $M_p = \frac{N}{2}$, so by Ribet's theorem there is a newform $g \in \mathfrak{B}^+(2)$. But this is impossible since $\dim S(2) = 0$. Thus, no such modular form f_E can exist, so neither can E and hence no such Fermat triplet (a, b, c) exists!

[11] In his fundamental paper, Frey[Fr1](see also [Fr2]) showed how many Diophantine statements can be reduced to the study of elliptic curves by means of certain elliptic curves now called Frey curves.

[12] This theorem should be read with a grain of salt, for one cannot assume that g has coefficients in $\mathbb{Z}$. Thus, while the precise statement of the theorem is somewhat more technical, the basic flavour is the same.

References

[DDT] H. Darmon, F. Diamond, R. Taylor: Fermat's Last Theorem. In: *Current Developments in Mathematics, 1995* (R. Bott et al., eds) International Press Inc., Cambridge, 1995, pp. 1–154.

[DMR] M.Davis, Y.Matijasevič, J.Robinson: Hilbert's Tenth Problem. Diophantine Equations: positive aspects of a negative solution; in: *Mathematical Developments arising from Hilbert Problems*, Proc. Symp. Pure Math. 28 (1976), pp. 323–378.

[Di] F. Diamond: The Taylor-Wiles construction and multiplicity one. *Invent. math.* **128** (1997), 379–391.

[Fr1] G.Frey: Links between stable elliptic curves and certain Diophantine equations. *Ann. Univ. Sarav.* **1** (1986), 1–40.

[Fr2] G. Frey: Links between solutions of $A-B=C$ and elliptic curves. In: *Number Theory, Ulm 1987* (H.P. Schlickewei, E. Wirsing, eds.) Springer Lecture Notes 1380 (1989), pp. 31–62.

[Fr3] G. Frey: On ternary equations of Fermat type and relations with elliptic curves (23pp.) In: *Proceedings of the Conference on Fermat's Last Theorem, Boston University, August 9–18, 1995.* (G. Cornell, J. Silverman, G. Stevens, eds.) (to appear).

[He] T. L. Heath: Diophantus of Alexandria. (2nd Edition). Cambridge U. Press, Cambridge, 1910; Dover Reprint, 1964.

[Ri1] P. Ribenboim: 13 Lectures on Fermat's Last Theorem. Springer-Verlag, New York - Berlin - Heidelberg, 1979.

[Ri2] P. Ribenboim: The Book of Prime Number Records. (Second Edition.) Springer-Verlag, New York-Berlin-Heidelberg, 1989.

[R] K. Ribet: On modular representations of $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms. *Invent. math.* **100** (1990), 431–476.

[Se] J.-P. Serre: Sur les représentations modulaires de degré 2 de $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$. *Duke Math. J.* **54** (1987), 179–230.

[Sh] G. Shimura: Introduction to the Arithmetic Theory of Automorphic Functions. Iwanami Shoten and Princeton U. Press, Princeton, 1971.

[W] A. Wiles: Modular elliptic curves and Fermat's Last Theorem. *Annals of Math.* **141** (1995), 443–551.