Infinite sums, diophantine equations and Fermat's last theorem[1]

Henri DARMON and Claude LEVESQUE

Abstract. Thanks to the results of Andrew Wiles, we know that Fermat's last theorem is true. As a matter of fact, this result is a corollary of a major result of Wiles: *every semi-stable elliptic curve over* $\mathbf{Q}$ *is modular.* The modularity of elliptic curves over $\mathbf{Q}$ is the content of the Shimura-Taniyama conjecture, and in this lecture, we will restrain ourselves to explaining in elementary terms the meaning of this deep conjecture.

§1. Introduction

A few years ago, the New York Times highlighted the proof of Fermat's last theorem by Andrew Wiles, completed in collaboration with his former Ph.D. student Richard Taylor. This was the last chapter in an epic initiated around 1630, when Pierre de Fermat wrote in the margin of his Latin version of Diophantus' ARITHMETICA the following enigmatic lines, unaware of the passions they were about to unleash:

> *Cubum autem in duos cubos, aut quadrato-quadratum in duos quadrato-quadratos, et generaliter nullam in infinitum ultra quadratum, potestatem in duos ejusdem nominis fas est dividere. Cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.*

In plain English, for those unfamiliar with Latin:

> *One cannot write a cube as a sum of two cubes, a fourth power as a sum of two fourth powers, and more generally a perfect power as a sum of two like powers. I have found a quite remarkable proof of this fact, but the margin is too narrow to contain it.*

The sequel is well-known: Fermat never revealed his alleged proof. Thousands of mathematicians (from amateurs to most famous scholars) working desperately hard at refinding this proof were baffled for more than three centuries.

[1] Written English version of a lecture given in French by Henri Darmon on October 14, 1995, at CEGEP de Lévis-Lauzon on the occasion of the *Colloque des Sciences Mathématiques du Québec* and which appeared in French in the *Comptes Rendus du* 38^e *Congrès de l'Association Mathématique du Québec.*

Fermat's Last Theorem. *The equation*

$$\boxed{x^n + y^n = z^n} \quad (n \geq 3) \tag{1.1}$$

has no integral solution with $xyz \neq 0$.

Using his so-called *method of infinite descent*, Fermat himself proved the theorem when $n = 4$. Euler is credited for the proof of the case $n = 3$ (though his proof was incomplete). The list of mathematicians who worked on this problem of Fermat reads like a Pantheon of number theory: Dirichlet, Legendre, Cauchy, Lamé, Sophie Germain, Lebesgue, Kummer, Wieferich, to name but the most famous. Their results secured the proof of Fermat's last theorem for all exponents $n \leq 100$.

Though the importance of the theorem looks like being mostly symbolic, this problem of Fermat was extraordinarily fruitful for modern mathematics. Kummer's efforts generated huge bulks of mathematical theories: algebraic number theory, cyclotomic fields. In 1985, the theory of elliptic curves and modular forms threw an unexpected light on the problem. This point of view was initiated by Gerhard Frey and led ten years later to the proof of Wiles.

Here is (at last!) this famous proof of Fermat's last theorem which was so keenly sought for. Roughly! (With references quoted from the appendix.)

Proof of Fermat's Last Theorem.
By K. Ribet [**R**], *the Shimura–Taniyama conjecture* (*for semi-stable elliptic curves*) *implies the truth of Fermat's last theorem.*
Thanks to the works of Wiles [**W**] *and Taylor–Wiles* [**T**–**W**], *we know that the Shimura–Taniyama conjecture is true for semi-stable elliptic curves.* Q.E.D.

This is a very short proof and it could possibly fit in that famous margin of the book of Diophantus. Hence Fermat's proof, if it existed, was different...

Readers will point out that this last proof lacks some details! The papers of Wiles and Taylor-Wiles cover more than 130 pages of the prestigious journal "*Annals of Mathematics*", and rely on numerous previous papers which could hardly be summarized in less than one thousand pages addressed to initiated readers.

So Wiles did not succeed in making his proof contained in some narrow margin of any manuscript. In August 1995, the organizers of a conference held in Boston on Fermat's last theorem got off with printing the proof on a tee-shirt, put on by the first author during his lecture at the *Colloque des Sciences mathématiques du Québec*, and whose content is reproduced in the appendix.

In this lecture, we will refrain from dealing with the existing link between Fermat's last theorem and the Shimura–Taniyama conjecture; we refer interested readers to papers listed in the bibliography. We shall restrain ourselves to explaining in elementary terms the meaning of the Shimura–Taniyama conjecture. As a matter of fact, we would like to make readers aware of the importance of this conjecture, which goes much beyond Fermat's last theorem, and is tied to some of the deepest and most fundamental questions of number theory.

§2. Pythagoras' equation

Let us start with Pythagoras' equation

$$\boxed{x^2 + y^2 = 1} \tag{2.1}$$

whose non-zero *rational solutions* $(x, y) = (\frac{a}{c}, \frac{b}{c})$ give birth to Pythagoras' triples (a, b, c) verifying the equation $a^2 + b^2 = c^2$. This equation was highlighted in Diophantus' treatise and led Fermat to consider the case where the exponents are greater than 2. (So our starting point is the same as Fermat's one, even if we will not deal with his last theorem...)

The rational solutions of Pythagoras' equation are given in a parametric way by

$$(x, y) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right), \quad t \in \mathbf{Q} \cup \{\infty\}, \tag{2.2}$$

which provides the classification of Pythagoras' triples and leads to the complete solution of Fermat's equation for $n = 2$. Integral solutions (with $x, y \in \mathbf{Z}$) are still simpler to describe. There are 4 of them, namely $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$; hence we write

$$N_{\mathbf{Z}} = 4. \tag{2.3}$$

We can also study the equation $x^2 + y^2 = 1$ on fields other than the rational numbers; for instance, the field $\mathbf{R}$ of real numbers, or the fields $\mathbf{F}_p = \{0, 1, 2, \ldots, p-1\}$ of congruence classes modulo p, where p is a prime number.

Solutions in real numbers of the equation $x^2 + y^2 = 1$ correspond to points on a circle of radius 1. Let us give the set of real solutions a quantitative measure by writing

$$N_{\mathbf{R}} = 2\pi, \tag{2.4}$$

the circonference of the circle.

The solutions of $x^2 + y^2 = 1$ on $\mathbf{F}_p$ form a finite set, and we set

$$N_p = \#\{(x, y) \in \mathbf{F}_p^2 : x^2 + y^2 = 1\}. \tag{2.5}$$

To calculate N_p, we let x run between 0 and $p - 1$ and look for solutions whose first coordinate is x. There will be 0, 1, or 2 solutions according to whether $1 - x^2$ is not a square modulo p, is equal to 0, or is a non-zero square modulo p, respectively. Since half of the non-zero integers modulo p are squares, it is expected that N_p is roughly equal to p; this prompts us to define a_p as the "error term" of this rough estimate:

$$a_p = p - N_p. \tag{2.6}$$

In so doing, we arrive at the main problem which, as will be seen later, leads directly to the Shimura–Taniyama conjecture.

Problem 1. *Does there exist a simple formula for the numbers N_p as a function of p (or, which in the same, for the numbers a_p)?*

Experimental methods play an important role in the theory of numbers, probably to a greater extent than in other fields of pure mathematics. Gauss was a prodigious calculator, and found his quadratic reciprocity law in some empiric way, before giving it many rigorous proofs. Following in the footsteps of the master, let us give a list of the values of N_p for some values of p.

p	N_p	a_p
2	2	0
3	4	−1
5	4	1
7	8	−1
11	12	−1
13	12	1
17	16	1
19	20	−1
23	24	−1
29	28	1
31	32	−1
37	36	1
41	40	1
⋮	⋮	⋮
10007	10008	−1
⋮	⋮	⋮

Table 1: $x^2 + y^2 = 1$

A look at the table leads at once to the following conjecture.

Conjecture 2. *The value of N_p is 2 if $p = 2$ and we have*

$$N_p = \begin{cases} p-1 & \text{if } p \equiv +1 \pmod 4, \\ \\ p+1 & \text{if } p \equiv -1 \pmod 4. \end{cases} \tag{2.7}$$

(In particular, we see that $p \neq N_p$, which might be of interest to our computer science colleagues: $\mathbf{P} \neq \mathbf{NP}$!)

How can we prove Conjecture 2? Let us come back to the parametrization

$$(x,\, y) = \left(\frac{1-t^2}{1+t^2},\, \frac{2t}{1+t^2} \right). \tag{2.8}$$

The values $t = 0, 1, \ldots, p-1, \infty$ give birth to a complete list of $p+1$ distinct solutions, excepted when -1 is a square j^2 modulo p. In the latter case, the denominator vanishes for the two values $t = j,\, -j$, so these values are not admissible. Therefore, when p is odd,

$$N_p = \begin{cases} p-1 & \text{if } -1 \text{ is a square modulo } p, \\ \\ p+1 & \text{if } -1 \text{ is not a square modulo } p. \end{cases} \tag{2.9}$$

The condition that -1 be a square modulo p may *a priori* look subtle, but we are fortunate to be able to count on the following theorem proved by Fermat.

Theorem 3 (Fermat). *The integer -1 is a square modulo p if and only if $p = 2$ or $p \equiv 1 \pmod 4$.*

Here is a proof, slightly different from that of Fermat. The multiplicative group $\mathbf{F}_p^\times$ is cyclic of order $p-1$, and the element -1 of order 2 has a square root if and only if $\mathbf{F}_p^\times$ possesses some elements of order 4.

Theorem 3 (that we just proved) together with formula (2.9) provides a proof of Conjecture 2 about the value of N_p. What is the purpose of such an explicit formula for N_p? Let us consider, for instance, the following infinite product (taken over all the primes p):

$$\begin{aligned} \prod_p \frac{p}{N_p} \quad &= \quad \prod_p \left(1 - \frac{a_p}{p}\right)^{-1} \\ \text{“}=\text{”} \quad & \left\{ \prod_{p\equiv 1(4)} \left(1 - \frac{1}{p}\right)^{-1} \right\} \cdot \left\{ \prod_{p\equiv -1(4)} \left(1 + \frac{1}{p}\right)^{-1} \right\} \\ \text{“}=\text{”} \quad & \left\{ \prod_{p\equiv 1(4)} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots\right) \right\} \cdot \left\{ \prod_{p\equiv -1(4)} \left(1 - \frac{1}{p} + \frac{1}{p^2} - \frac{1}{p^3} + \cdots\right) \right\} \end{aligned} \tag{2.10}$$

$$“=” \quad 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \frac{1}{13} - \cdots \tag{2.11}$$

$$= \quad \frac{\pi}{4} \qquad \text{(by Leibniz's formula)}, \tag{2.12}$$

where the equality (2.11) is (formally) a consequence of the unique factorization of integers as products of powers of primes. We then deduce

$$\prod_p \frac{N_p}{p} = \frac{4}{\pi}. \tag{2.13}$$

To tell the truth, our proof of the equality (2.13) is a fallacy, because of the off-hand way the convergence questions were dealt with (this contempt would give analysts the shivers). This is why some equalities were used within inverted commas. Eighteenth century mathematicians like Euler were quite at ease with such formal series manipulations, guided by their instinct to reach the right conclusion by avoiding traps. As a matter of fact, it is true that

$$\prod_p \frac{N_p}{p} \quad \text{converges to} \quad \frac{4}{\pi},$$

though the convergence is very slow.

Recalling that $N_{\mathbf{R}} = 2\pi$ and that $N_{\mathbf{Z}} = 4$, we conclude that

$$\left(\prod_p \frac{N_p}{p}\right) \cdot N_{\mathbf{R}} \;=\; 2N_{\mathbf{Z}}. \tag{2.14}$$

This magical formula unveils a mysterious relation between the solutions of the equation $x^2 + y^2 = 1$ on finite fields $\mathbf{F}_p$, on the real numbers $\mathbf{R}$, and on the ring $\mathbf{Z}$ of integers. In particular, the numbers N_p which depend only on the solutions of the equation $x^2+y^2 = 1$ on $\mathbf{Z}_p$, “know” the behaviour of the equation over the real numbers: thanks to these numbers N_p, we recover the number π, related to the circumference of the circle. Fundamentally, this is only a simple reinterpretation of Leibniz's formula, but in fact this is quite a fruitful one. At the beginning of the twenty-first century, number theory had not yet digested the deep meaning of this formula and of its generalizations, as will be seen later.

§3. The Fermat–Pell equation

In his abundant correspondence with his colleagues from Europe, Fermat liked to send them mathematical challenges. By doing so, he invited the English mathematicians Wallis and Brouncker to find the integer solutions of the equation

$$\boxed{x^2 - 61y^2 = 1}\,. \tag{3.1}$$

This is a particular case of the so-called Fermat–Pell equation $x^2 - Dy^2 = 1$. Fermat had a crush for this equation and had developed a general method to solve it, based on continued fractions. When $D = 61$, the smallest non-trivial solution is

$$(x,\, y) = (1766319049,\, 226153980)\,. \tag{3.2}$$

It is the odd size of this smallest solution that led Fermat to take $D = 61$, although he pretended (with a bit of maliciousness) that this value of D was taken at random. This Fermat–Pell equation, of degree 2, is a conic in the plane, as is Pythagoras' equation. Let us denote by N_p the number of solutions modulo p, and let us give once more the list of the numbers N_p for some values of p.

p	N_p	a_p
2	2	0
3	2	1
5	4	1
7	8	−1
11	12	−1
13	12	1
17	18	−1
19	18	1
23	24	−1
29	30	−1
31	32	−1
37	38	−1
41	40	1
43	44	−1
47	46	1
53	54	−1
59	60	−1
61	122	−61
67	68	−1
71	72	−1
73	72	1
⋮	⋮	⋮
10007	10006	1
10009	10008	1
⋮	⋮	⋮

Table 2: $x^2 - 61y^2 = 1$

Using the parametrization

$$(x,\, y) = \left(\frac{1+61t^2}{1-61t^2},\, \frac{2t}{1-61t^2}\right), \quad t \in \mathbf{Q} \cup \{\infty\}, \tag{3.3}$$

of the conic (3.1), we find as before that $N_2 = 2$, that $N_p = 2p$ if $p = 61$, and that otherwise

$$N_p = \begin{cases} p-1 & \text{if 61 is a square modulo } p, \\ \\ p+1 & \text{if 61 is not a square modulo } p. \end{cases} \tag{3.4}$$

Let us now use Gauss reciprocity law which for our purposes asserts that for p-odd, 61 is a square modulo p if and only if p is a square modulo 61. So for $p \neq 2$, 61, we find

$$N_p = \begin{cases} p-1 & \text{if } p \text{ is a square modulo 61}, \\ \\ p+1 & \text{if } p \text{ is not a square modulo 61}. \end{cases} \tag{3.5}$$

This simple formula (which is periodic since it depends only on p modulo 61) for the numbers N_p allows to deduce, with formal calculations closely copied on those of equations (2.10) to (2.12), the identity

$$\prod_p \frac{p}{N_p} \quad \text{“}=\text{”} \quad \frac{1}{2}\sum_n \frac{a_n}{n}, \tag{3.6}$$

where

$$a_n = \begin{cases} 0 & \text{if } 61|n, \text{ or if } n \text{ is even}, \\ +1 & \text{if } n \text{ odd is a non-zero square modulo 61}, \\ -1 & \text{if } n \text{ odd is not a square modulo 61}. \end{cases} \tag{3.7}$$

One verifies (with the help of Abel's summation formula, for instance) that the infinite sum in (3.6) converges (conditionally). Some kind of heroic calculations (which we invite the readers to do) lead to an identity analoguous to the formula (2.12) of Leibniz,

$$\sum_n \frac{a_n}{n} = \frac{\log(1766319049 + 226153980\sqrt{61})}{2\sqrt{61}}. \tag{3.8}$$

One recognizes in this expression the coefficients which appeared in the solution (3.2) of (3.1). In conclusion, the knowledge of the numbers N_p allowed us to “recover” a (fundamental) solution of a Fermat–Pell equation.

As a matter of fact, the identity (3.6) can be formally rewritten as

$$\left(\prod_p \frac{N_p}{p}\right) \cdot N_{\mathbf{R}} \quad \text{“}=\text{”} \quad 4\sqrt{61}N_{\mathbf{Z}}. \tag{3.9}$$

The quantities $N_{\mathbf{R}}$ and $N_{\mathbf{Z}}$ are both infinite, since the hyperbola defined by the equation $x^2 - 61y^2 = 1$ has no finite length and the Fermat–Pell equation possesses an infinity of integral solutions. It is all the same natural to define the quotient $\frac{N_{\mathbf{R}}}{N_{\mathbf{Z}}}$ as

$$\frac{N_{\mathbf{R}}}{N_{\mathbf{Z}}} := \log(1766319049 + 226153980\sqrt{61}), \tag{3.10}$$

namely, as the quantity appearing in the numerator of the right hand side of (3.8). As a matter of fact, the set of integral solutions of (3.1) is an abelian group isomorphic to $\mathbf{Z}\times\mathbf{Z}/2\mathbf{Z}$ and the application

$$(x,\, y) \mapsto \log(|x + y\sqrt{61}|) \tag{3.11}$$

sends this group into a discrete subgroup G of $\mathbf{R}$ which is isomorphic to $\mathbf{R}$. It is therefore natural to define $N_{\mathbf{R}}/N_{\mathbf{Z}}$ as the volume of $\mathbf{R}$, *i.e.*, as in (3.10).

After a few months, Wallis and Brouncker gave an answer to Fermat's question, sending him the solution (3.2) of (3.1), together with a general method (essentially similar to the method of Fermat based on continued fractions) to solve the Fermat–Pell equation $x^2 - Dy^2 = 1$. We do not know what was the reaction of the Toulouse mathematician, but one can imagine he felt some secret resentment... This shows that Wiles and Taylor are not the first two English mathematicians to brilliantly take up Fermat's challenges.

§4. The equation $x^3 + y^3 = 1$

Let us keep the same momentum, and after having dealt with conics let us switch to equations of degree 3. As a tribute to Fermat, let us study for instance

$$\boxed{x^3 + y^3 = 1}\,. \tag{4.1}$$

Does there exist as before a simple formula for the number N_p of solutions of this equation modulo p? Once more, let us give a table.

p	N_p	a_p
2	2	0
3	3	0
5	5	0
7	6	1
11	11	0
13	6	7
17	17	0
19	24	−5
23	23	0
29	29	0
31	33	−2
37	24	13
41	41	0
43	10	33
47	47	0
53	53	0
⋮	⋮	⋮
10007	10007	0
10009	9825	184

Table 3: $x^3 + y^3 = 1$

Contrary to the case of the degree 2 equations, the integers a_p are not all 0 or ± 1, and seem to behave rather randomly. However, one may guess by inspection a few properties of these integers a_p. For example, it looks like a_p always vanishes when 3 divides $p+1$. But what is going on when $p \equiv 1 \pmod 3$? Once more, it is Gauss himself who provided the answer by proving the following theorem.

Theorem 4 (Gauss).

(1) *If* $p \equiv -1 \pmod 3$, *then* $a_p = 0$.

(2) *If* $p \equiv 1 \pmod 3$, *then the number* $4p$ *can be written as* $4p = A^2 + 27B^2$ *with* $A \equiv -1 \pmod 3$, *which makes* A *unique, so we have* $a_p = A + 2$.

The following table allows us to verify this theorem for a few values of p:

p	N_p	a_p	$4p = A^2 + 27B^2$
2	2	0	$---$
3	3	0	$---$
5	5	0	$---$
7	6	1	$28 = (-1)^2 + 27 \cdot 1^2$
11	11	0	$---$
13	6	7	$52 = 5^2 + 27 \cdot 1^2$
17	17	0	$---$
19	24	-5	$76 = (-7)^2 + 27 \cdot 1^2$
23	23	0	$---$
29	29	0	$---$
31	33	-2	$124 = (-4)^2 + 27 \cdot 2^2$
37	24	13	$148 = 11^2 + 27 \cdot 1^2$
41	41	0	$---$
43	10	33	$172 = 8^2 + 27 \cdot 2^2$
47	47	0	$---$
53	53	0	$---$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10007	10007	0	$---$
10009	9825	184	$40036 = 182^2 + 27 \cdot 16^2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 4: $x^3 + y^3 = 1$ (sequel)

§5. Elliptic curves

An elliptic curve is a diophantine equation of degree 3 having at least one rational solution. For example, the equation $x^3+y^3 = 1$. One can prove that any elliptic curve over the rational numbers $\mathbf{Q}$ may be written, after a proper change of variables, in the form

$$y^2 = x^3 + ax + b, \tag{5.1}$$

where a, b are rational numbers.

As before, denote by N_p the number of solutions of the equation (5.1) over the finite field $\mathbf{F}_p$ of p elements.

Question 5. *Is there an explicit formula for the numbers N_p associated to an elliptic curve like the equation $x^3 + y^3 = 1$?*

Said otherwise, we would like to generalize the result of Gauss for the equation $x^3+y^3 = 1$ to the case of any given elliptic curve. This is exactly the scope of the Shimura–Taniyama conjecture proved by Wiles for a very large class of elliptic curves.

Before giving explicit statements, let us see how the land lies by considering the elliptic curve

$$\boxed{y^2 + y = x^3 - x^2} \tag{5.2}$$

studied by Eichler. Here are some values of N_p as calculated by a computer:

p	N_p	a_p
2	4	-2
3	4	-1
5	4	1
7	9	-2
11	10	1
13	9	4
17	19	-2
19	19	0
23	24	-1
29	29	0
31	24	7
$\vdots$	$\vdots$	$\vdots$
10007	9989	18
$\vdots$	$\vdots$	$\vdots$

Table 5: $y^2 + y = x^3 - x^2$

This time, it is more difficult to guess a structure for the values of the integers a_p which again seem to behave rather randomly. Hasse proved the deep inequality

$$|a_p| \le 2\sqrt{p} \tag{5.3}$$

(valid for all elliptic curves), but this is far from providing an *exact formula* for the numbers N_p.

Eichler, building on deep results of Hecke, was however successfull in obtaining an exact formula. The starting point is to extend the definition of the coefficient a_p (valid for the prime index p) to any index n by setting

$$\begin{cases} a_1 &= 1, \\ a_p &= p - N_p, \\ a_{p^r} &= a_p a_{p^{r-1}} - p a_{p^{r-2}}, \\ a_n &= \prod_{i=1}^{r} a_{{p_i}^{e_i}}, \qquad \text{where} \qquad n = \prod_{i=1}^{r} {p_i}^{e_i}. \end{cases} \tag{5.4}$$

We notice that this extension is a rather natural one: if we denote by N_{p^r} the number of solutions of the elliptic curve over the finite field $\mathbf{F}_{p^r}$ of p^r elements, then we have

$$a_{p^r} = p^r - N_{p^r}. \tag{5.5}$$

Theorem 6 (Eichler). *The formal series* $\sum_{n=1}^{\infty} a_n q^n$ *is given by the formula:*

$$\begin{aligned} q\prod_{n=1}^{\infty}(1-q^n)^2\cdot(1-q^{11n})^2 \quad = \quad & q - \mathbf{2q^2} - \mathbf{q^3} + 2q^4 + \mathbf{q^5} + 2q^6 - \mathbf{2q^7} \\ & -2q^9 - 2q^{10} + \mathbf{q^{11}} - 2q^{12} + \mathbf{4q^{13}} + 4q^{14} \\ & -q^{15} - 4q^{16} - \mathbf{2q^{17}} + 4q^{18} + 2q^{20} + 2q^{21} \\ & -2q^{22} - \mathbf{q^{23}} - 4q^{25} - 8q^{26} + 5q^{27} - 4q^{28} \\ & +2q^{30} + \mathbf{7q^{31}} + \ldots + \mathbf{18q^{10007}} + \ldots \end{aligned}$$

The reader can at leisure verify the truth of Eichler's theorem for a few values of p, by comparing the coefficients of q^p written in boldface, with the values from Table 5.

The Shimura–Taniyama conjecture, proved by Wiles, is a direct generalization of Eichler's theorem, in the sense that Wiles gave a *very precise description* of the generating function $\sum_n a_n q^n$, where the integers a_n are the coefficients associated to any given elliptic curve.

More precisely, let

$$f(z) \;=\; \sum_{n=1}^{\infty} a_n e^{2\pi i n z} \tag{5.6}$$

be a Fourier series with coefficients $a_n \in \mathbf{R}$, and let N be a positive integer. We say that $f(z)$ is a *modular form* of level N if the following conditions are satisfied:

(1) The series defining f converges for $Im(z) > 0$, *i.e.*, when $|e^{2\pi iz}| < 1$. The series f then represents a holomorphic function on the Poincaré upper half plane of complex numbers having a strictly positive imaginary part.

(2) For all $\begin{pmatrix} a & b \\ Nc & d \end{pmatrix} \in SL_2(\mathbf{Z})$, we have

$$f\left(\frac{az+b}{Ncz+d}\right) = (Ncz+d)^2 f(z), \tag{5.7}$$

where $SL_2(\mathbf{Z})$ is the group of 2×2 matrices of determinant 1 with coefficients in $\mathbf{Z}$.

Here is at last the famous Shimura–Taniyama conjecture.

Conjecture 7 (Shimura–Taniyama). *Let $y^2 = x^3 + ax + b$ be an elliptic curve over the rational numbers $\mathbf{Q}$, and let a_n $(n = 1, 2, \ldots)$ be the integers defined for this curve by the equations of* (5.4). *Then the generating function*

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi inz} \tag{5.8}$$

is a modular form.

In fact, the conjecture is more precise:

(1) It predicts the value of the level N of the modular form associated to the elliptic curve. This level would be equal to the *arithmetic conductor* of the curve, which depends only on the primes having "bad reduction". The exact definition of N will not be used in our treatment.

(2) The space of modular forms of a given level N is a vector space over $\mathbf{R}$ whose dimension, a finite number, can easily be calculated out of the value of N. This space is equipped with certain natural linear operators defined by Hecke. The conjecture also states that the modular form f is an eigenform (*i.e.*, a characteristic vector) for all Hecke operators.

One shows that there is but a finite number of modular forms of level N which are eigenforms for all Hecke operators, and whose first Fourier coefficient a_1 is equal to 1. So once the conductor N of an elliptic curve has been calculated, we are led to a finite list of possibilities for the sequence $\{a_n\}_{n\in\mathbf{N}}$ associated to this curve. From this point of view, the Shimura–Taniyama conjecture gives an explicit formula for the numbers N_p of rational points on the elliptic curve modulo p.

Thanks to the works of Wiles and Taylor–Wiles, we now know that the Shimura–Taniyama conjecture is true for a very large class of elliptic curves. As a matter of fact, Diamond proved, improving upon the results of Wiles and Taylor-Wiles, that it suffices that the elliptic curve has good reduction, or in the worst case has only one double point modulo 3 or 5.

The formula of Wiles for the integers N_p associated to an elliptic curve looks at first less explicit than that of Fermat (Conjecture 2) for the equation $x^2 + y^2 = 1$, or than that of Theorem 4 of Gauss for the equation $x^3 + y^3 = 1$. Nevertheless it allows one to give a meaning to the expression $\prod_p \frac{p}{N_p}$, or to be more precise[2], to the quantities

$$\prod_p \frac{p}{N_p + 1}.$$

This is achieved by introducing the L-series associated to the elliptic curve E:

$$L(E, s) \; = \; \prod_p \left(1 - \frac{a_p}{p^s} + \frac{1}{p^{2s-1}}\right)^{-1} \; = \; \sum_n \frac{a_n}{n^s}. \tag{5.9}$$

One notes that formally,

$$L(E, 1) \quad \text{“} = \text{”} \quad \prod_p \frac{p}{N_p + 1}, \tag{5.10}$$

though the series defining $L(E, s)$ converges only for $\mathrm{Re}(s) > \frac{3}{2}$. In order to make $L(E, 1)$ meaningfull, one needs to know that the series defining $L(E, s)$ admits an analytic continuation at least up to the value $s = 1$.

The following fundamental result of Hecke will then prove useful.

Theorem 8 (Hecke). *If the sequence $\{a_n\}_{n \in \mathbf{N}}$ comes from a modular form, then the function $L(E, s)$ admits an analytic continuation to the whole complex plane, and in particular, the value of $L(E, 1)$ is well defined.*

If one knows that the elliptic curve E is modular, then the result of Hecke allows one to define

$$\prod_p \frac{p}{N_p + 1} := L(E, 1). \tag{5.11}$$

[2]In our naïve definition of N_p, we systematically omitted to count the solution which corresponds to the “point at infinity” and which naturally comes into play when one considers an equation of the elliptic curve in the Desargues projective plane. It is therefore natural to replace N_p by $N_p + 1$.

As in the previous example, one may expect some useful pieces of arithmetic information about the curve E from the value of $L(E,1)$ (or more generally, from the behaviour of $L(E,s)$ at the neighbourhood of $s=1$).

This is exactly the content of the Birch–Swinnerton-Dyer conjecture, of which a particular case is the following.

Weak Birch–Sinnerton-Dyer conjecture. *The elliptic curve E possesses a finite number of rational points if and only if $L(E,1) \neq 0$.*

This conjecture is far from being proved, and is still one of the most important open questions in the theory of elliptic curves. One can count although on some partial results, for instance, the following one, which is a consequence of the works of Gross–Zagier, Kolyvagin, together with an analytic result due to Bump–Friedberg–Hoffstein and Murty–Murty.

Theorem 9 (Gross–Zagier, Kolyvagin). *Let E be a modular elliptic curve. If the function $L(E,s)$ possesses a zero of order 0 or 1 at $s=1$, then the weak Birch–Swinnerton-Dyer conjecture is true for E.*

The case where the function $L(E,s)$ has a zero of order >1 still remains very mysterious. One expects in this case that the equation of the curve E has always rational solutions, but we still ignore how to find (or build) them in a systematic way, or even whether or not there is an algorithm to determine in all cases the set of all rational solutions. Despite spectacular progresses over the past few years, several number theorists, in love with elliptic curves, will be kept very busy.

Appendix: The t-shirt of the Boston University Conference

On the front of the above-mentioned t-shirt, one can read the following.

FERMAT'S LAST THEOREM: *Let* $n, a, b, c \in \mathbf{Z}$ *with* $n > 2$. *If* $a^n + b^n = c^n$ *then* $abc = 0$.

Proof. The proof follows a program formulated around 1985 by Frey and Serre [F,S]. By classical results of Fermat, Euler, Dirichlet, Legendre and Lamé, we may assume that $n = p$, an odd prime ≥ 11. Suppose that $a, b, c \in \mathbf{Z}, abc \neq 0$, and $a^p + b^p = c^p$. Without loss of generality we may assume $2|a$ and $b \equiv 1 \pmod 4$. Frey [F] observed that the elliptic curve $E : y^2 = x(x - a^p)(x + b^p)$ has the following "remarkable" properties:
(1) E is semistable with conductor $N_E = \prod_{\ell | abc} \ell$; and
(2) $\bar{\rho}_{E,p}$ is unramified outside $2p$ and is flat at p.
By the modularity theorem of Wiles and Taylor–Wiles [W,T–W], there is an eigenform $f \in S_2(\Gamma_0(N_E))$ such that $\rho_{f,p} = \bar{\rho}_{E,p}$. A theorem of Mazur implies that $\bar{\rho}_{E,p}$ is irreducible, so Ribet's theorem [R] produces a Hecke eigenform $g \in S_2(\Gamma_0(2))$ such that $\rho_{g,p} \equiv \rho_{f,p} \pmod{\mathcal{P}}$ for some $\mathcal{P}|p$. But $X_0(2)$ has genus zero, so $S_2(\Gamma_0(2)) = 0$. This is a contradiction and Fermat's Last Theorem follows. Q.E.D.

On the back of the t-shirt, one finds the following bibliography.

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

[**R**] Ribet, K: On modular representations of Gal($\bar{\mathbf{Q}}/\mathbf{Q}$)) arising from modular forms. *Invent. Math.* **100** (1990), 431-476.

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

[**T–W**] Taylor, R.L., Wiles, A.: Ring-theoretic properties of certain Hecke algebras. *Annals of Math.* **141** (1995), 553-572.

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

Annoted bibliography

The references appear under seven headings, each one dealing with a given theme. Readers interested only by easily understood survey papers will appreciate references 1 to 4, 8 to 11, 14 to 18 of Section B.

(A) Fermat's last theorem

The following references provide historic informations about Fermat's last theorem or about methods not dealing with elliptic curves

1. E.T. Bell, *The Last Problem*, 2^e édition, MAA Spectrum, Mathematical Association of America, Washington, DC, 1990, 326 pages.

2. H.M. Edwards, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Graduate Texts in Math. **50**, Springer–Verlag, New York, Berlin, Heidelberg, 1977, 410 pages.

3. C. Houzel, *De Diophante à Fermat*, in *Pour la Science* **220**, January 1996, 88–96.

4. P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer–Verlag, New York, Berlin, Heidelberg, 1979, 302 pages.

5. L.C. Washington, *Introduction to Cyclotomic Fields*, Graduate Texts in Math. **83**, Springer–Verlag, New York Berlin 1982, 389 pages.

(B) Elliptic curves and Fermat's last theorem

To learn more on the links between Fermat's last theorem and elliptic curves, we suggest the following references.

1. N. Boston, *A Taylor-made Plug for Wiles' Proof*, College Math. J. **26**, No. 2, 1995, 100–105.

2. B. Cipra, *"A Truly Remarkable Proof"*, in *What's happening in the Mathematical Sciences*, AMS Volume **2**, 1994, 3–7.

3. J. Coates, *Wiles Receives NAS Award in Mathematics*, Notices of the AMS **43**, 7, 1994, 760–763.

4. D.A. Cox, *Introduction to Fermat's Last Theorem*, Amer. Math. Monthly **101**, No. 1, 1994, 3–14.

5. B. Edixoven, *Le rôle de la conjecture de Serre dans la preuve du théorème de Fermat*, Gazette des mathématiciens **66**, Oct. 1995, 25–41. Addendum: idem **67**, Jan. 1996, 19.

6. G. Faltings, *The Proof of Fermat's Last Theorem by R. Taylor and A. Wiles*, Notices AMS **42**, No. 7, 743–746.

7. G. Frey, *Links Between Stable Elliptic Curves and Certain Diophantine Equations*, Ann. Univ. Sarav. **1**, 1986, 1–40.

8. G. Frey, *Links Between Elliptic Curves and Solutions of $A - B = C$*, Indian Math. Soc. **51**, 1987, 117–145.

9. G. Frey, *Links Between Solutions of $A - B = C$ and Elliptic Curves*, dans *Number Theory, Ulm, 1987, Proceedings*, Lecture Notes in Math. **1380**, Springer–Verlag, New York, 1989, 31–62.

10. D. Goldfeld, *Beyond the last theorem*, in *The Sciences* **1996**, March/April, 34–40.

11. C. Goldstein, *Le théorème de Fermat*, La Recherche **263**, Mars 1994, 268–275.

12. C. Goldstein, *Un théorème de Fermat et ses lecteurs*, Presses Universitaires de Vincennes, 1995.

13. F.Q. Gouvêa, *A Marvelous Proof*, Amer. Math. Monthly **101**, No. 3,1994, 203–222.

14. B. Hayes and K. Ribet, *Fermat's Last Theorem and Modern Arithmetic*, Amer. Scientist **82**, 1994, 144–156.

15. Y. Hellegouarch, *Points d'ordre $2p^h$ sur les courbes elliptiques*, Acta Arith. **26**, 1974/75, 253–263.

16. Y. Hellegouarch, *Fermat enfin démontré*, in *Pour la Science* **220**, February 1996, 92–97.

17. S. Lang, *Old and New Conjectured Diophantine Inequalities*, Bull. AMS (New Series) **23**, No. 1, 1990, 37–75.

18. B. Mazur, *Number Theory as Gadfly*, Amer. Math. Monthly **98**, No. 7, 1991, 593–610.

19. B. Mazur, *Questions about Number*, in New Directions in Mathematics, Cambridge Univ. Press, Cambridge, à paraître.

20. M.R. Murty, *Fermat's Last Theorem: an Outline*, Gazette Sc. Math. Québec, Vol. **XVI**, No. 1, 1993, 4–13.

21. M.R. Murty, *Reflections on Fermat's Last Theorem*, Elem. Math. **50** (1995) no. 1, 3–11.

22. J. Oesterlé, *Nouvelles approches du "théorème" de Fermat"*, Séminaire Bourbaki No. **694** (1987-88), Astérisque **161–162**, 1988, 165–186.

23. K. Ribet, *On Modular Representations of Gal*($\bar{\mathbf{Q}}/\mathbf{Q}$) *Arising from Modular Forms*, Invent. Math. **100**, 1990, 431–476.

24. K. Ribet, *From the Taniyama–Shimura Conjecture to Fermat's Last Theorem*, Ann. Fac. Sci. Toulouse (5) **11** (1990) no. 1, 116–139.

25. K. Ribet, *Wiles Proves Taniyama's Conjecture; Fermat's Last Theorem Follows*, Notices Amer. Math. Soc. **40**, 1993, 575–576.

26. K. Ribet, *Galois Representations and Modular Forms*, Bull. AMS (New Series) **32**, No. 4, 1995, 375–402.

27. M. Rosen, *New Results on the Arithmetic of Elliptic Curves*, Gazette Sc. Math. Québec, Vol. **XIV**, No. 1, 1993, 30–43.

28. K. Rubin and A. Silverberg, *A Report on Wiles' Cambridge Lectures*, Bull Amer. Math. Soc. (New Series) **31**, 1994, 15–38.

29. R. Schoof, *Proof of Taniyama–Weil Conjecture for Semi-stable Elliptic Curves over* **Q**, Duke Math. J. **54**, 1987, 179–230.

30. J-P. Serre, *Sur les représentations modulaires de degré* 2 *de Gal*($\bar{\mathbf{Q}}/\mathbf{Q}$), Duke Math. J. **54**, 1987, 179–230.

31. J-P. Serre, *Lettre à J.-F. Mestre*, in Current Trends in Arithmetical Algebraic Geometry, ed. by K. Ribet, Contemporary Mathematics **67**, AMS, 1987.

32. A. van der Poorten, *Notes on Fermat's Last Theorem*, Canadian Math. Society Series of Monographs and Advanced Texts, Wiley Interscience, Jan. 1996.

33. A. Wiles, *Modular Forms, Elliptic Curves, and Fermat's Last Theorem*, Proc. International Congress of Math., 1994, Birkhauser Verlag, Basel, 1995, 243–245.

(C) About the works of Wiles and Taylor

The following references concentrate on the work of Wiles and his per se proof of the Shimura–Taniyama conjecture.

1. J. Coates and S.T. Yau, *Elliptic Curves and Modular Forms*, in Proceedings of a conference in Hong Kong in 1993, International Press, Cambridge (MA) and Hong Kong, 1995.

2. H. Darmon, F. Diamond et R. Taylor, *Fermat's Last Theorem*, Current Developments in Math. **1**, International Press, 1995, 1–154.

3. H. Darmon, *The Shimura–Taniyama Conjecture, (d'après Wiles)*, (en Russe) Uspekhi Mat. Nauk **50** (1995), no. 3(303), pages 33–82. (Version anglaise à paraître dans Russian Math Surveys).

4. V.K. Murty, ed., *Elliptic Curves, Galois Representations and Modular Forms*, CMS Conference Proc., AMS, Providence RI, 1996.

5. J. Oesterlé, *Travaux de Wiles (et Taylor...), Partie II*, Séminaire Bourbaki 1994-95, exposé No. **804**, 20 pages.

6. K. Ribet, *Galois Representations and Modular Forms*, Bull. AMS (New Series) **32**, 1995, No. **4**, 375–402.

7. J-P. Serre, *Travaux de Wiles (et Taylor...), Partie I*, Séminaire Bourbaki 1994–95, exposé No. **803**, 13 pages.

8. R.L. Taylor and A. Wiles, *Ring Theoretic Properties of Certain Hecke Algebras*, Annals of Math. **141**, 1995, 553–572.

9. A. Wiles, *Modular Elliptic Curves and Fermat's Last Theorem*, Annals of Math. **141**, 1995, 443–551.

(D) Videos

Some readers may enjoy the numerous videos dealing with Fermat's last theorem and its proof.

1. Fermat Fest, *Fermat's Last Theorem. The Theorem and Its Proof: an Exploration of Issues and Ideas.* Shown on the occasion of a "Fermat Fest" in San Francisco, CA, on July 28, 1993, Video, *Selected Lectures in Mathematics*, AMS, Providence, RI, 1994, (98 min.)

2. B. Mazur, *Modular Elliptic Curves and Fermat's Last Theorem*, CMS meeting in Vancouver, August 1993, Video, *Selected Lectures in Mathematics*, AMS, Providence, RI, 1995, (50 min.)

3. K. Ribet, *Modular Elliptic Curves and Fermat's Last Theorem*, Lecture given at George Washington U. , Washington DC, 1993, Video, *Selected Lectures in Mathematics*, AMS, Providence, RI, 1993, (100 min.)

(E) Fermat and Gauss

To learn more on the works of Fermat and Gauss, in particular on the proof of Theorem 3 of Fermat, on the Fermat–Pell equation, and on the equation $x^3 + y^3 = 1$:

1. L.E. Dickson, *History of the Theory of Numbers*, Vol. II, Chelsea Publ. Co., New York, 1971.

2. K. Ireland et M. Rosen, *A Classical Introduction to Modern Number Theory*, 2^{nd} edition, Graduate Texts in Math. **84** Springer–Verlag, New York, 1990, 389 pages.

3. W. Scharlau et H. Opolka, *From Fermat to Minkowski. Lectures on the Theory of Numbers and Its Historical Development,* Translated from the german by Walter K. Bühler and G. Cornell, Undergraduate Texts in Math., Springer–Verlag, New York-Berlin, 1985, 184 pages.

4. A. Weil, *Fermat et l'équation de Pell*, in *Collected Papers*, Vol. III, Springer–Verlag, New York, 1979, 413–420.

5. A. Weil, *Number Theory. An Approach Through History. From Hammurapi to Legendre*, Birkhauser Boston Inc., Boston, MA, 1984, 375 pages.

(F) Elliptic curves

There is plenty of choice for the readers keen to learn more on elliptic curves.

1. J.W.S. Cassels, *Lectures on Elliptic Curves*, London Math. Society Student Texts **24**, Cambridge University Press, 1991, 137 pages.

2. H. Darmon, *Wiles' Theorem and the Arithmetic of Elliptic Curves*, in *Modular Forms and Fermat's Last Theorem*, Springer–Verlag, New York, 1997, 549-569.

3. D. Husemöller, *Elliptic Curves*, Graduate Texts in Math. **111**, Springer–Verlag, New York, 1987, 350 pages.

4. H. Kisilevsky and M.R. Murty, *Elliptic Curves and Related Topics*, CRM Proceedings and Lecture Notes, AMS, 1994, 195 pages.

5. A.W. Knapp, *Elliptic Curves*, Mathematical Notes **40**, Princeton U. Press, Princeton, NJ, 1992, 427 pages.

6. S. Lang, *Elliptic Curves: Diophantine Analysis*, Springer–Verlag, New York, 1978, 261 pages.

7. M.R. Murty and V.K. Murty, *Lectures on Elliptic Curves*, Lectures given at Andhra U., India, 1989, 92 pages.

8. M.R. Murty, *Topics in Number Theory*, Lectures given at the Mehta Research Institute, India, 1993, 117 pages.

9. J.H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Undergraduate Texts in Math., Springer–Verlag, New York, 1992, 281 pages.

10. J.H. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Math. **106**, Springer–Verlag, New York, 1992, 400 pages.

11. J.H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Graduate Texts in Math., vol. **151**, Springer–Verlag, New York, 1994, 525 pages.

12. J. Tate, *Rational Points on Elliptic Curves*, Philips Lectures, Haverford College, 1961, unpublished notes.

(G) Modular forms and functions and the Shimura–Taniyama conjecture

1. T. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, Graduate Texts in Math. **41**, Springer–Verlag, New York, 1976, 248 pages.

2. J. Cremona, *Algorithms for Modular Elliptic Curves*, Cambridge Univ. Press, Cambridge, 1992, 343 pages.

3. N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, 2^{nd} edition, Graduate Texts in Math. **97**, Springer–Verlag, New York, 1993, 248 pages.

4. S. Lang, *Introduction to Modular Forms*, Springer–Verlag, New York, 1976, 261 pages.

5. T. Miyake, *Modular Forms*, Springer–Verlag, New York, 1989.

6. M.R. Murty, *Elliptic Curves and Modular Forms*, Can. Math. Bull. **34** (3), 1991, 375–384.

7. A. Ogg, *Modular Forms and Dirichlet Series*, Benjamin, New York, 1969.

8. J-P. Serre, *A Course in Arithmetic*, 2^{nd} edition, Graduate Texts in Math. **7**, Springer–Verlag, New York, Berlin, Heidelberg, 1973, 115 pages.

Added in proof. *In the December* 1999 *issue of the Notices of the AMS, H. Darmon reported on the recent proof by C. Breuil, B. Conrad, F. Diamond and R. Taylor of the full Shimura–Taniyama–Weil conjecture for all elliptic curves over* $\mathbf{Q}$.

Henri DARMON, CICMA, Mathematics Dept., McGILL University, Montréal, Canada H3A 2K6

Claude LEVESQUE, CICMA, Dép. de Mathématiques et de Statistique, Université LAVAL, Québec, Canada G1K 7P4