IC/93/326

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

LECTURES ON FERMAT'S LAST THEOREM

B. Sury

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

MIRAMARE-TRIESTE

IC/93/326

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

LECTURES ON FERMAT'S LAST THEOREM

B. Sury
International Centre for Theoretical Physics, Trieste, Italy
and
School of Mathematics, Tata Institute of Fundamental Research,
Homi Bhabha Road, Bombay 400005, India.

MIRAMARE - TRIESTE
September 1993

The aim of these two lectures is to discuss the main ideas involved in the approach towards the so-called Fermat's last theorem (FLT). The discussion leads to the point where recent work by A.Wiles starts and we do not discuss his work here.The organisation of our lectures is as follows.
♠ Mathematical history of the FLT .
♠ Present approach - in short.
♠ Elliptic curves.
♠ Modular forms.
♠ Relations between elliptic curves and modular forms.
♠ Taniyama-Shimura-Weil conjecture and FLT.

1 History of FLT

P. Fermat (1601-1665) asserted that the equation

$$X^n + Y^n = Z^n$$

has no integer solutions X, Y, Z (other than X, Y or $Z = 0$) if $n \geq 3$. In 1847, E.Kummer made the first important breakthrough. Earlier, C.F.Gauss proved the quadratic reciprocity law conjectured by L.Euler and while looking for similar laws for higher powers, he discovered thatthe calculations were easier by working over the Gaussian integers $a + b.i$ $(a, b \in \mathbf{Z}; i = \sqrt{-1})$ rather than the usual integers. He developed a theory of prime factorisation for these numbers. It is here that the algebraic numbers entered number theory.
An algebraic number is a complex number which is a solution of a polynomial equation

$$a_n x^n + \cdots\cdots + a_0 = 0$$

where $a_i \in \mathbf{Z}$ and, if $a_n = 1$ it is called an algebraic integer. Examples are $\sqrt{2}, i = \sqrt{-1}, \zeta = e^{\frac{2\pi\sqrt{-1}}{n}}$. In the setting of algebraic integers, a solution of Fermat's equation $x^n + y^n = z^n$ (if one exists), can be factorised by introducing an nth root of 1, $\zeta = e^{\frac{2\pi\sqrt{-1}}{n}}$, as

$$z^n = (x+y)(x+\zeta y)\cdot\cdots\cdot(x+\zeta^{n-1}y).$$

This factorisation takes place in the ring $\mathbf{Z}[\zeta]$, of algebraic integers of the form

$$a_0 + a_1\zeta + \cdots\cdots + a_r\zeta^r \quad ; \quad a_k \in \mathbf{Z}.$$

In 1847, Lame' announced a 'proof' of the FLT. His proposal was to show that (only the case where x, y have no common factors needs to be considered) $x+y, x+\zeta y, \cdots, x+\zeta^{n-1}y$ have no common factors. Implicitly assuming that unique factorisation holds among these numbers, he 'deduced' that each $x+\zeta^i y$ is an n-th power, and derived a contradiction from this. But, Kummer pointed out that unique factorisation can fail; e.g. $n = 23$. Kummer went on to develop the theory of 'ideal' numbers to restore unique factorisation. By these means he was able to give a proof of FLT for a class of primes called regular primes. Let us briefly describe this.

In $\mathbf{Z}[\zeta]$ (or similarly, in 'integers' in extension fields of $\mathbf{Q}$), every nonzero ideal is uniquely a product of prime ideals but the ideals may not be principal (i.e. generated by one element) unlike the usual integers. The set of ideals, modulo, the principal ideals forms a finite group, the class group, under multiplication of ideals. The problem with the above argument of Lame is that each $x + \zeta^i y$ is an n-th power of an ideal I and not of an element, in general. Thus, the ideal I, considered as an element of the class group satisfies I^n = Identity which implies that if n happens to be a prime number not dividing the order of the class group, I itself will be principal and Lame's argument goes through perfectly. The primes n which don't divide the class number (= order of class group) of the corresponding $\mathbf{Z}[\zeta]$, are called regular primes.

So, the theory of ideals, and in some sense, the subject of algebraic number theory itself was born out of these attempts to solve FLT. The solution of FLT itself would have no application, even in number theory, but the general ideas which have risen out of these attempts to solve the FLT are very rich and have many applications.

2 The present approach - in brief

In the modern approach, essentially three objects and their interrelationships are involved. These objects are (i) Elliptic curves, (ii) Modular forms; and (iii) Galois representations. The relationships alluded to are very special cases of what is known as Langlands's program. G.Frey had the idea of associating an elliptic curve $E_{a,b,c}$ to every solution (a, b, c) of the Fermat equation in such a way that this curve would exhibit remarkable properties which would contradict the so-called Taniyama-Shimura-Weil (T-S-W) conjecture which says, in essence, that elliptic curves over $\mathbf{Q}$ 'come from' modular forms. Frey's curve is the following. Consider (a, b, c) relatively prime such that

$$a^l + b^l + c^l = 0$$

where $l \geq 5$ is a prime. After permuting, we may suppose that b is even and that $a \equiv 1 \bmod 4$. Then $E_{a,b,c}$ is the curve given by the equation

$$y^2 = x(x - a^l)(x + b^l)$$

It turns out that there are serious difficulties in carrying out this idea, having to do with the Galois representation on points of finite order of the curve. K.Ribet succeeded in proving a result on such representations which was strong enough to show that T-S-W conjecture $\Rightarrow$ FLT. Our purpose, in these lectures is to introduce the relevant objects and study them in brief and finally indicating the proof of the implication T-S-W $\Rightarrow$ FLT. The announcement of A.Wiles is the proof of T-S-W in the particular case of semistable elliptic curves (this suffices for FLT).

3 Elliptic curves

Elliptic curves are non-singular algebraic curves of genus 1 having a specified basepoint. Any such curve can be written as the locus in $\mathbf{P}^2$ of a cubic equation of the form

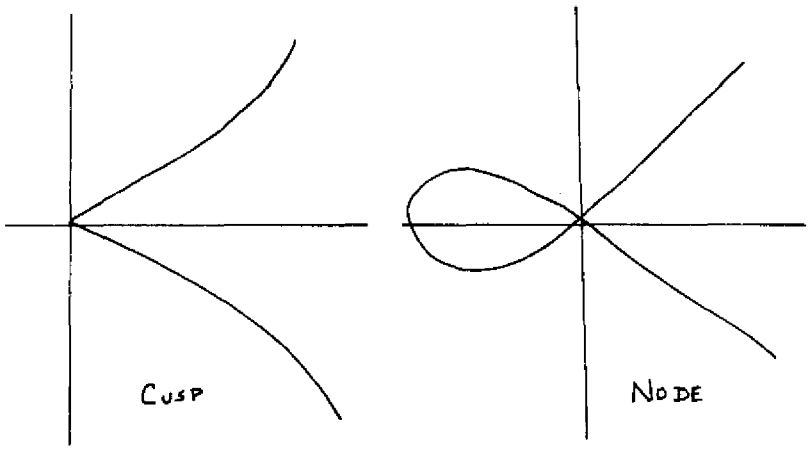
$$E \; : \; Y^2 + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

where the point $[0, 1, 0]$ in $\mathbf{P}^2$ is the base point. If K is a field such that all the a_i's belong to K, we say that E is defined over K. If Char.$K \neq 2, 3$ we

can change variables and write the equation as

$$y^2 = x^3 + ax + b$$

always remembering there is the extra point $[0,1,0]$ at infinity. The non-singularity of the curve is equivalent to the discriminant condition $4a^3 + 27b^2 \neq 0$ viz. that the polynomial $x^3 + ax + b$ have distinct roots. Moreover, if the curve is singular i.e. if the discriminant is 0, there is exactly one singular point. The singular point is called a node if there are two tangents at the singular point and called a cusp, if there is a single tangent at the singular point.



Associated to E are the two quantities

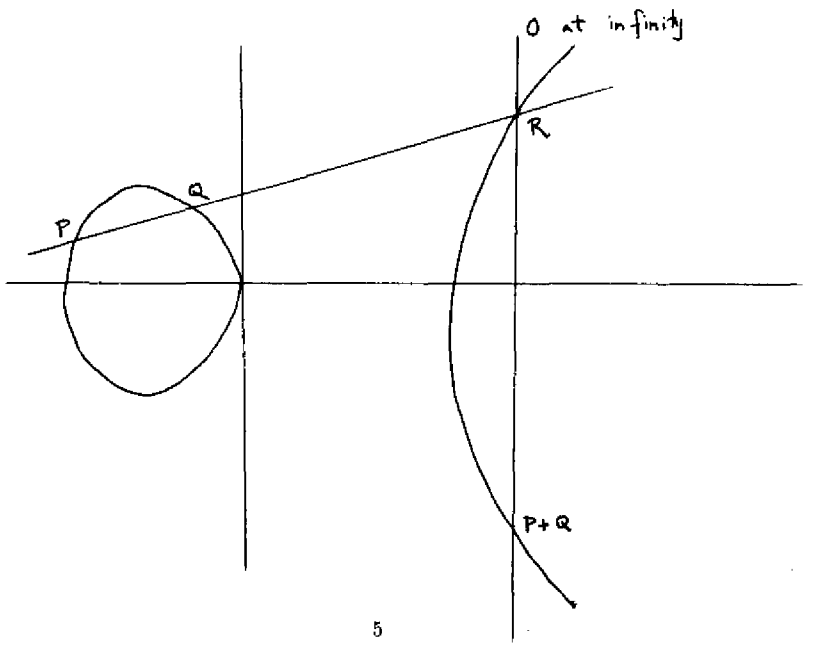
$$\Delta = -16(4a^3 + 27b^2)$$

$$j = 1728\frac{(4a)^3}{\Delta}$$

Two elliptic curves are isomorphic over the algebraic closure $\overline{K}$ iff they have the same j - invariant.

Group Law

The points of E satisfy a group law. Let $P, Q \in E$ and let $L \subset \mathbf{P}^2$ be the line joining them. Since the equation for E is a cubic, L intersects E at a third point R. Let L' be the line joining R and $O = [0,1,0]$. Then $P + Q$ is the third point of intersection of L' with E. This makes E into an abelian group with O as the identity.



If E is defined over K, then the set

$$E(K) = \{O\} \cup \{(x,y) \in K^2 : y^2 = x^3 + ax + b\}$$

is a subgroup of E. Over the field $\mathbf{C}$ of complex numbers, $E(\mathbf{C})$ is juct a complex torus of dimension 1 i.e. $\mathbf{C}/\Lambda$ where Λ is a lattice in $\mathbf{C}$. This follows from the classical theory of elliptic functions. Over arithmetic fields like number fields K, $E(K)$ is more subtle. If K is a number field (i.e. a finite extension of $\mathbf{Q}$), the theorem of Mordell-Weil states that $E(K)$ is a finitely generated abelian group i.e. $E(K) \cong \mathbf{Z}^r \bigoplus F$ for some finite abelian group F. The possible torsion subgroups F have been determined by Mazur but the rank r of the Mordell-Weil group is more subtle. The conjecture of Birch and Swinnerton - Dyer states that this rank r is also the order at $s = 1$ of the (so-called) L-function $L(s,E)$ of E. At the present time, the conjecture is known to be true for $r \leq 2$.

Isogenies

Since there is a group law on E, $\forall m$, there is a morphism $[m] : E \to E$, the multiplication-by-m map. If $m \neq 0$, this is surjective, as E is a curve, and hence it is a finite map. In fact, $[m]$ is a homomorphism of groups and has a finite kernel. These maps are examples of isogenies.

An *isogeny* $\phi : E \to E$ is, by definition, a morphism which preserves O. It can be shown that isogenies are necessarily group homomorphisms. $\mathrm{Hom}(E_1, E_2)$, the set of isogenies from E_1 to E_2, forms a free abelian group. (The freeness is not obvious). In the case $E_1 = E_2 = E$, $\mathrm{End}(E) = \mathrm{Hom}(E,E)$ is, indeed, an integral domain of characteristic 0. If $\mathrm{End}(E)$ contains isogenies other than the $[m]$, then $\mathrm{End}(E) \supset \mathbf{Z}$; $\mathrm{End}(E) \neq \mathbf{Z}$. In this case, E is said to have complex multiplication. If $m \in \mathbf{Z}$, $m \neq 0$ and E is an elliptic curve over an algebraically closed field K, then the group $E[m] := \mathrm{Ker}\ [m]$ of m-division points has the following structure.

If, either Char.$K = 0$ or (m,Char.K) = 1, then

$$E[m] \cong \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$$

If Char. $K = p > 0$, then either

$$E[p^n] \cong \mathbf{Z}/p^n\mathbf{Z} \quad \forall \quad n \geq 1$$

or

$$E[p^n] = \{0\} \quad \forall \quad n \geq 1$$

Tate Module

Let E be an elliptic curve over a field K, and let l be a prime number not dividing the characteristic of K. Note that the natural action of $\mathrm{Gal}(\overline{K}/K)$ on E gives an action on each $E[l^n] \cong \mathbf{Z}/l^n\mathbf{Z} \times \mathbf{Z}/l^n\mathbf{Z}$. So, we have representations

$$Gal(\overline{K}/K) \to Aut\ E[l^n] \cong GL_2(\mathbf{Z}/l^n\mathbf{Z})$$

Since it is easier to study representations on a ring of char. 0, we construct the *Tate module*

$$T_l(E) = lim_{\leftarrow_n} E[l^n]$$

the projective limit with respect to the maps

$$E[l^{n+1}] \xrightarrow{[l]} E[l^n]$$

In other words, $T_l(E)$ consists of sequences $(a_1, a_2, \cdots\cdots)$ of points $a_i \in E$ such that $la_1 = 0$ and $la_{i+1} = a_i$. In fact, $T_l(E) \cong \mathbf{Z}_l \times \mathbf{Z}_l$ as a $\mathbf{Z}_l$ - module. Since the Galois action commutes with the maps

$$E[l^{n+1}] \xrightarrow{[l]} E[l^n]$$

we have an action of $Gal(\overline{K}/K)$ on $T_l(E)$ i.e. the so-called l-adic representation

$$\rho_l \ : \ Gal\ (\overline{K}/K) \ \to \ Aut\ T_l(E) \ \cong \ GL(2, \mathbf{Z}_l)$$

The l-adic representation ρ_l is a continuous representation. Let us suppose now that K is a number field and let us recall the definition of a Frobenius conjugacy class. For a finite extension L of K, corresponding to any discrete valuation v of K (equivalently, to any prime ideal of the ring $\mathcal{O}_K$ of integers of K), we have the decomposition groups G_{w_i} for the various extensions w_i of v to K. These are the subgroups of $\mathrm{Gal}(L/K)$ stabilising the w_i and are, hence, conjugate since $\mathrm{Gal}(L/K)$ acts transitively on the w_i. There are natural homomorphisms from the G_{w_i} to $\mathrm{Gal}(f_{w_i}/f_v)$, the Galois group of the finite residue field extensions. The kernel I_{w_i}, which is called the Inertia group at w_i, is trivial for almost all (i.e. all but finitely many) v. For such v (i.e. when the Inertia groups are trivial), the decomposition groups G_{w_i} being isomorphic to $\mathrm{Gal}(f_{w_i}/f_v)$, have a special element Fr_i viz. the inverse image of the Frobenius homomorphism $\sigma_i \in \mathrm{Gal}(f_{w_i}/f_v)$, defined by $\sigma_i x = x^q$ where $q = \mathrm{Card}\ (f_v)$. Since the G_{w_i} are conjugate, we have a well- defined conjugacy

class Fr_v in $\mathrm{Gal}(L/K)$. In particular, Tr $\rho_l(\mathrm{Fr}_v)$ and det $\rho_l(\mathrm{Fr}_v)$ make sense where ρ_l is the l-adic representation defined above. Now, composing with the natural map : $\mathbf{Z}_l \to \mathbf{F}_l$ one gets a representation

$$\rho_l \ : \ Gal(\overline{K}/K) \ \to \ GL(2, \mathbf{F}_l)$$

We call ρ_l *unramified at* v if $\rho_l(I_w) = \{Id\}$ for some (and hence all) extensions w of v to finite extensions of K.

Reduction mod p and minimal model
If a cubic curve C satisfying Weierstrass's equation is singular, it is easy to show that it has exactly one singular point which is either a node or a cusp and the non- singular points C_{ns} of C satisfy the group law. If it is a node, $E_{ns} \cong \overline{K}^{\times}$, the multiplicative group, and, if it is a cusp, $E_{ns} \cong \overline{K}^{+}$, the additive group of $\overline{K}$. All this was over a general algebraically closed field $\overline{K}$. If E is an elliptic curve defined over $\mathbf{Q}$, then E can be described by Weierstrass equations with coefficients in $\mathbf{Z}$. Moreover, there is a model over $\mathbf{Z}$ for which the discriminant is minimal in some sense. For E in such a minimal form, we have the reduced curves E_p over $\mathbf{F}_p$ for each prime p. These are plane cubics. We make the following definitions.
E has *good reduction* at p if E_p is non-singular.
E has *semistable (or multiplicative) reduction* at p if E_p has a node.
E has *unstable (or additive) reduction* at p if E_p has a cusp.
In the latter two cases, E is said to have bad reduction at p. One can see easily that if E has bad reduction at p, then p divides the minimal discriminant Δ.

4 Modular forms

Let $\mathcal{H}$ denote the upper half-plane $\{z : \mathrm{Im} z > 0\}$. The group $\Gamma = SL(2, \mathbf{Z})$ acts as usual, by fractional linear transformations

$$z \mapsto \gamma z = \frac{az+b}{cz+d} \ ; \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Definition : A meromorphic function f on $\mathcal{H}$ is called a modular function of weight k for Γ if :
(i) $f(\gamma z) = (cz+d)^k f(z) \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$;
(ii) The Fourier expansion of f in the variable $q = e^{2\pi i z}$ has the form

$$f(z) = \sum_{n=n_0}^{\infty} c(n) q^n$$

for some integer n_0.

Definition : f is a *modular form* of weight k for Γ if f is holomorphic on $\mathcal{H}$ and $n_0 = 0$ (in which case we say f is holomorphic at ∞ and set $f(\infty) = c(0)$).
If further, $f(\infty) = 0$, f is called a *cusp form* for Γ.
Example : The Eisenstein series

$$G_{2k}(z) = \sum_{(m,n) \neq (0,0)} \frac{1}{(m+nz)^{2k}}$$

is a modular form of weight $2k$.
Let $M_k = \{$Space of modular forms of weight $2k\}$;
$M_{k,0} = \{$Space of cusp forms of weight $2k\}$.
$M_k, M_{k,0}$ are $\mathbf{C}$ - vector spaces and $M = \sum_{k=0}^{\infty} M_k$ is a graded $\mathbf{C}$ - algebra because $f \in M_k, g \in M_l \Rightarrow fg \in M_{k+l}$.

Proposition
Every modular form is a polynomial in G_4 and G_6 i.e.

$$M \cong \mathbf{C}[G_4, G_6]$$

Further, $M_k = \{0\}$ if $k < 0$ and, for $k \geq 0$,
$dim_{\mathbf{C}} M_k = [\frac{k}{6}](\mathrm{resp}[\frac{k}{6}]+1)$ if $k \equiv 1$ (resp. $\not\equiv 1$) mod 6.
Define $\mathcal{H}^* = \mathcal{H} \cup \mathbf{P}^1(\mathbf{Q})$. Since $SL(2, \mathbf{Z})$ acts transitively on $\mathbf{P}^1(\mathbf{Q})$, the quotient space $\mathcal{H}^*/\Gamma$ which is a Riemann surface, can be thought of as $\mathcal{H}/\Gamma$ to which a single point (called a cusp) has been added.

Modular forms for congruence subgroups
Let us consider an integer $N \geq 1$ and the subgroup

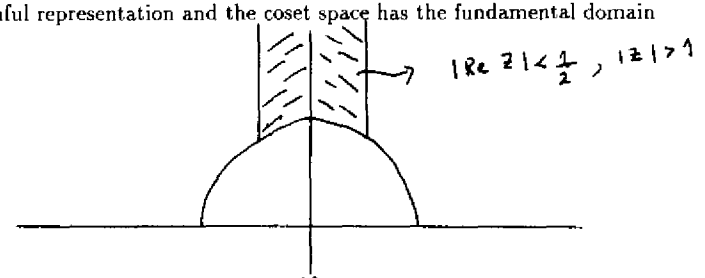
$$\Gamma_0(N) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \ : \ c \equiv 0 \bmod N\}$$

of finite index in Γ.
The cusps of $\Gamma_0(N)$ are the orbits of $\Gamma_0(N)$ on $\mathbf{P}^1(\mathbf{Q}) \subset \mathcal{H}^*$.
Define a modular function of weight k for $\Gamma_0(N)$ to be one for Γ such that :
(i) $f(\gamma z) = (cz+d)^k f(z) \ \forall \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$; and
(ii) f is meromorphic at each cusp of $\mathcal{H}^*/\Gamma_0(N)$.
As before, f as above is a *modular form* if it is holomorphic everywhere and is a *cusp form* if it vanishes at all cusps.

Cusp forms of weight 2
If f is a modular form of weight 2 for $\Gamma_0(N)$, then the differential form $f(z)dz$ on $\mathcal{H}$ is $\Gamma_0(N)$-invariant (check!). If, further, f is a cusp form, then $f(z)dz$ therefore defines a holomorphic 1-form on $\mathcal{H}^*/\Gamma_0(N)$. In fact,
{Space of weight 2 cusp forms for $\Gamma_0(N)$} $\cong$ {Holomorphic 1-forms on $\mathcal{H}^*/\Gamma_0(N)$}.
Thus, we can calculate the genus of $\mathcal{H}^*/\Gamma_0(N)$, and thereby find the dimension of the space of weight 2 cusp forms for $\Gamma_0(N)$.
For instance, for $N = 2$, this genus $= 0$.

5 Modular curves and the T-S-W conjecture

Consider as before the action of $\Gamma = SL(2,\mathbf{Z})$ on $\mathcal{H}$. The group $\Gamma/\{\pm I\}$ has a faithful representation and the coset space has the fundamental domain



There is a classical function, holomorphic on $\mathcal{H}$ and Γ - invariant, called the j-function, which gives a holomorphic isomorphism

$$j \ : \ \mathcal{H}/\Gamma \ \to \mathbf{P}^1(\mathbf{C}) \setminus \{\infty\} \ = \ \mathbf{A}^1(\mathbf{C})$$

If one takes $q = e^{2\pi i \tau}$ as a local parameter at infinity, then one can compactify $\mathcal{H}/\Gamma$ by adjoining the point at infinity, thus obtaining a compact Riemann surface isomorphic to $\mathbf{P}^1(\mathbf{C})$. In terms of q, the function j has a Laurent expansion

$$j(q) = \frac{1}{q} + 744 + 196884q + O(q^2)$$

In fact, $j : \mathcal{H}^*/\Gamma \to \mathbf{P}^1(\mathbf{C})$ gives an analytic isomorphism. A better way to conceive of j is in terms of complex tori as follows. Let $\Lambda = \mathbf{Z}\omega_1 \bigoplus \mathbf{Z}\omega_2$ be a lattice in $\mathbf{C}$ where $\{\omega_1, \omega_2\}$ is a $\mathbf{Z}$-basis. We can suppose that $\frac{\omega_1}{\omega_2} = \tau \in \mathcal{H}$. Since j is Γ-invariant, the value $j(\tau)$ is independent of the choice of basis $\{\omega_1, \omega_2\}$ and $j(\tau)$ remains same also if we replace $\{\omega_1, \omega_2\}$ by $\{c\omega_1, c\omega_2\}$ for any $c \in \mathbf{C}^*$. Thus, we can define $j(\Lambda) = j(\tau)$, and we have $j(c\Lambda) = j(\Lambda)$. Since $\mathbf{C}/\Lambda$ is a complex torus (of dimension 1) this shows that j is the single invariant for isomorphism classes of such tori. Now, let N be a positive integer and let

$$\Gamma_0(N) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \ : \ c \equiv 0 \bmod N\}$$

Let C_N denote the cyclic group generated by $1/N$ considered as a point of the torus $\mathbf{C}/[\tau, 1]$. We have :
The association $\tau \mapsto (\mathbf{C}/[\tau,1] \ , \ C_N)$ gives a bijection between $\mathcal{H}/\Gamma_0(N)$ and isomorphism classes of tori with a cyclic subgroup of order N. Note that though the map $\mathcal{H}/\Gamma \to \mathbf{A}^1$ is just a classification over $\mathbf{C}$, $\mathbf{A}^1$ is defined over $\mathbf{Q}$. In fact, (from the theory of the so-called Shimura varieties) there exists an affine curve $Y_0(N)$ defined over $\mathbf{Q}$ such that $Y_0(N)(\mathbf{C}) \cong \mathcal{H}/\Gamma_0(N)$, and such that $Y_0(N)$ parametrizes isomorphism classes of pairs (E, C) **algebraically** in the following sense, where E is an elliptic curve and C is a cyclic subgroup of order N. If $k \supset \mathbf{Q}$ is a field, then a point $Y_0(N)(k)$ corresponds to such a pair (E, C) with E defined over k and C invariant under the Galois group $\mathrm{Gal}(\overline{\mathbf{Q}}/k)$. The affine curve $Y_0(N)$ can be compactified by adjoining the cusps i.e. the points that lie over $j = \infty$. This completion is denoted by $X_0(N)$.

Consider the curve $X_0(N)$ defined over $\mathbf{Q}$. Now, we state the following
(Weaker version of the) T-S-W conjecture :
Every elliptic curve over $\mathbf{Q}$ *is parametrised by modular functions i.e.* $\exists N$ *and a surjective morphism* $X_0(N) \to E$ *defined over* $\mathbf{Q}$.

To make it more precise, let $\mathcal{S}(N)$ be the $\mathbf{C}$-vector space of differentials of the first kind on $X_0(N)$. Recall that for a curve C, a differential ω is said to be of the first kind if $\mathrm{Ord}_P\omega \geq 0 \ \forall \ P \in C$. The space of such differentials has dimension g = Genus of C. In terms of the complex variable on $\mathcal{H}$ covering $Y_0(N)(\mathbf{C})$ by the map $\mathcal{H}/\Gamma_0(N) \to Y_0(N)(\mathbf{C})$, such a differential can be expressed as $\omega = f(\tau)d\tau$, where f is holomorphic on $\mathcal{H}$. In terms of the parameter $q = e^{2\pi i\tau}$, we can write

$$\omega = \sum_{n=1}^{\infty} a_n q^n \frac{dq}{q} = f_\infty(q)\frac{dq}{q}$$

Note that this shows that what we are actually concerned with are *cusp forms* for $\Gamma_0(N)$. The coefficients a_n are called the Fourier coefficients of the form ω. We define the Hecke operators, for our purposes, as follows. (In our situation, we are concerned only with cusp forms of weight 2 but the action of Hecke operators can, in general, be defined on forms of any weight).
For every prime p such that $N \not\equiv 0 \bmod p$, there is an operator T_p on $\mathcal{S}(N)$ whose effect on the Fourier coefficients is given by

$$T_p : \sum a_n q^n \mapsto \sum a_{pn} q^n + \sum a_n q^{pn}$$

If p/N, there is again an operator T_p such that $T_p : \sum a_n q^n \mapsto a_{pn} q^n$. A non-zero form in $\mathcal{S}(N)$ which is a common eigenvector for all the T_p is called an eigenform. The precise form of the T-S-W conjecture is the following.

Conjecture (Taniyama-Shimura-Weil)
Let E be an elliptic curve over $\mathbf{Q}$, and N its conductor. Then, there exists an eigenform in $\mathcal{S}(N)$ such that for each prime p not dividing N the eigenvalue a_p of T_p for this form satisfies $|E(\mathbf{F}_p)| = 1 + p - a_p$. (Thus, a_p is also the trace of the Frobenius in the l-adic representations). In addition, there exists a rational map $\pi : X_0(N) \to E$ defined over $\mathbf{Q}$, such that if $\omega_{E,\pi}$ is a suitably normalised differential of the first kind on E, then $\pi^*\omega_{E,\pi}$ is the

above eigenform for the Hecke operators, and

$$\pi^*\omega_{E,\pi} = \sum_{n=1}^{\infty} a_n q^n \frac{dq}{q}$$

where $a_1 = 1$ and a_p is the eigenvalue as above.

In the above, recall that the conductor N of E is a measure of its bad reduction and is a number of the form $\prod_{p/\Delta} p^{f_p}$, where Δ is the minimal discriminant of E. Also, recall that $\omega_{E,\pi}$ is chosen from a 1-dimensional space. Finally, note that a_p are integers since $a_p = p + 1 - |E(\mathbf{F}_p)|$. Here, and in above, $|X|$ refers to the cardinality of a set X.

6 Modular representations

Let $N \in \mathbf{Z}$. Let $T = T_N$ be the subring of $\mathrm{End}_{\mathbf{C}}\mathcal{S}(N)$ generated by the operators T_p over $\mathbf{Z}$. Then, it is known that T is a free $\mathbf{Z}$-module of rank equal to the genus $g(N)$ of $X_0(N)$, which is also equal to the dimension of $\mathcal{S}(N)$. Let $\mathcal{M}$ be a maximal ideal of T. Then the residue field $T/\mathcal{M} = k_{\mathcal{M}}$ is a finite field, say of characteristic l. By a theorem of Deligne-Serre, there exists a semisimple continuous homomorphism $\rho : G_{\mathbf{Q}} = Gal(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2, k_{\mathcal{M}})$ such that
(i) $\det \rho = \chi_l$, the l-th cyclotomic character;
(ii) ρ is unramified at all primes not dividing N; and
(iii) $\mathrm{tr}\ \rho(Fr_p) = T_p \bmod \mathcal{M}$ for all p not dividing N.

Recall that $\chi_l : G_{\mathbf{Q}} \to \mathbf{F}_l^* \subset k_{\mathcal{M}}^*$ is the character such that $\forall \sigma \in G_{\mathbf{Q}}$ and an lth root ζ of unity, we have $\sigma\zeta = \zeta^{\chi(\sigma)}$. Also, Fr_p denotes the Frobenius conjugacy class detyermined by p. The representation ρ is unique upto isomorphism. This follows from the Cebotarev density theorem, which implies that all elements of Image ρ are conjugate to $\rho(Fr_p)$ for some p, together with the fact that trace and determinant determine a 2-dimensional representation. (We remark that for an elliptic curve over $\mathbf{Q}$, the corresponding l-adic representation ρ_l of $G_{\mathbf{Q}}$ on $GL(2, \mathbf{F}_l)$ has for its determinant, the lth cyclotomic character (as in (i) above!) and the trace of $\rho_l(Fr_p)$ is $a_p = p + 1 - |E(\mathbf{F}_p)|$ which is the eigenvalue predicted in the T-S-W conjecture!)

Let $\mathbf{F}$ be a finite field, and let $\gamma : G_{\mathbf{Q}} \to GL(2, \mathbf{F})$ be a continuous, semisimple representation. γ is said to be *modular of level N* if $\exists$ a maximal ideal $\mathcal{M}$ of T and an embedding $i : T/\mathcal{M} \to \overline{\mathbf{F}}$ such that the representations

$$G_{\mathbf{Q}} \xrightarrow{\gamma} GL(2, \mathbf{F}) \to GL(2, \overline{\mathbf{F}})$$

$$G_{\mathbf{Q}} \xrightarrow{\rho} GL(2, T/\mathcal{M}) \xrightarrow{i} GL(2, \overline{\mathbf{F}})$$

are isomorphic. Equivalently, one requires a homomorphism $\alpha : T \to \overline{\mathbf{F}}$ such that $\mathrm{tr}\ \gamma(Fr_p) = \alpha(T_p)$ and $\det \gamma(Fr_p) = p$, for all but finitely many primes p. Finally, there is the notion of a representation γ being finite. Let p be a prime $\neq l$. Then, γ is *finite at p* if γ is unramified at p. (In general, finiteness is the following notion. Now, the decomposition group $G_p \cong Gal(\overline{\mathbf{Q}_p}/\mathbf{Q}_p)$. γ is finite at p, if there exists a finite, flat group scheme $\mathcal{G}$ over $\mathbf{Z}_p$, with an action of $\mathbf{F}$ on $\mathcal{G}$ making $\mathcal{G}$ of rank 2 over $\mathbf{F}$, such that γ is isomorphic to the representation of $Gal(\overline{\mathbf{Q}_p}/\mathbf{Q}_p)$ on $\mathcal{G}(\overline{\mathbf{Q}_p})$).

Remark
By the uniqueness of the Deligne-Serre representation, we can restate the T-S-W conjecture as follows. *If ρ is the l-adic representation on E, then ρ is modular of level N (= equal to the conductor of E).*

7 T-S-W ⇒ FLT

Proposition (Serre)
Let E be a semistable elliptic curve over $\mathbf{Q}$, put in minimal model over $\mathbf{Z}$. Let ρ be the representation of $G_{\mathbf{Q}}$ on $E[l]$ for some prime l. Let p be a prime. Let Δ be the minimal discriminant.
Then, the representation ρ is finite at $p \Leftrightarrow Ord_p\Delta \equiv 0 \mod l$.

Finally, a result of Ribet shows
Theorem
Let γ be an irreducible 2-dimensional representation of $G_{\mathbf{Q}}$ over a finite field of characteristic $l > 2$. Assume that γ is modular of square-free level N, and that there is a prime q/N, $q \neq l$ such that γ is **not** *finite at q. Suppose further that p/N such that γ is finite at p.*
Then γ is modular of level $\frac{N}{p}$.

Using these results, we outline the proof of
T-S-W for semistable curves ⇒ FLT
Suppose T-S-W is true for semistable elliptic curves. It suffices to show that there is no triple (a, b, c) of relatively prime non-zero integers such that $a^l + b^l + c^l = 0$ where $l \geq 5$ is a prime. Suppose that there is such a triple. Without loss of generality, we may assume that a, b, c are relatively prime. After permuting them, we may even suppose that b is even and that $a \equiv 1 \mod 4$. Consider the Frey elliptic curve

$$E : y^2 = x(x - a^l)(x + b^l)$$

One can compute its conductor to be $N = abc$, and the minimal discriminant to be $\Delta = \frac{(abc)^{2l}}{2^8}$. We then obtain the representation ρ of $G_{\mathbf{Q}}$ on $E[l]$ which is known to be irreducible by results of Mazur and Serre. Frey had already noted that if $p \neq l$ and p/N but $p \neq 2$, then ρ is unramified at p, so finite at p. By Serre's proposition above, one sees therefore that ρ is **not** finite at 2. By T-S-W, the representation ρ is modular. Applying Ribet's theorem repeatedly, we deduce that ρ is modular of level 2. This is impossible, because $\mathcal{S}(2)$ has dimension 0. This proves the implication that T-S-W ⇒ FLT.

Acknowledgements
This is based on two lectures that I delivered at the I.C.T.P, Trieste in August, 1993. I would like to thank Professors M.S.Narasimhan, A.Verjovsky and H.Assadi for their encouragement and the (resilient) audience for sitting through the lectures. Finally, the support (financial and otherwise) provided by the I.C.T.P is gratefully acknowledged.