Lecture 32: Fermat's Last Theorem and Modularity of Elliptic Curves

William Stein

Math 124 Harvard University **Fall 2001**

In this lecture I will sketch an outline of the proof of Fermat's last theorem, then give a rigorous account of what it means for an elliptic curve to be "modular".

The are several exercises below. They are optional, but if you do them and give them to Grigor, I suspect that he would look at them (whether or not you do the exercises will not directly affect your course grade in any way).

1 Fermat's Last Theorem

Theorem 1.1. *Let $n > 2$ be an integer. If $a, b, c \in \mathbb{Z}$ and*

$$a^n + b^n = c^n,$$

then $abc = 0$.

Proof (sketch). First reduce to the case when $n = \ell$ is a prime greater than 3 (see Exercise 1.2). Suppose that

$$a^\ell + b^\ell = c^\ell$$

with $a, b, c \in \mathbb{Z}$ and $abc \neq 0$. Permuting (a, b, c), we may suppose that b is even and that we have $a \equiv 3 \pmod 4$. Following Gerhard Frey, consider the elliptic curve E defined by

$$y^2 = x(x - a^\ell)(x + b^\ell).$$

The discriminant of E is $2^4(abc)^{2\ell}$ (see Exercise 1.3 below).

Andrew Wiles and Richard Taylor [Annals of Math., May 1995] proved that E must be "modular". This means that there is a "modular form"

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

of "level $N = abc$" such that for all primes $p \nmid abc$,

$$a_p = p + 1 - \#E(\mathbb{Z}/p\mathbb{Z}).$$

Ken Ribet [Inventiones Math., 1991] used that the discriminant of E is a perfect ℓth power (away from 2) to prove that there is a cuspidal modular form

$$g(z) = \sum_{n=1}^{\infty} b_n e^{2\pi i n z}$$

of "level 2" such that

$$a_p \equiv b_p \pmod{\ell} \qquad \text{for all } p \nmid abc.$$

This is a contradiction because the space of "cuspidal modular forms" of level 2 has dimension 0 (see Section 3.1). □

Exercise 1.2. Reduce to the prime case. That is, show that if Fermat's last theorem is true for prime exponents, then it is true.

Exercise 1.3. Prove that $y^2 = x(x - a^\ell)(x + b^\ell)$ has discriminant $2^4(abc)^{2\ell}$.

The rest of this lecture is about the words in the proof that are in quotes.

2 Holomorphic Functions

The complex *upper half plane* is the set

$$\mathfrak{h} = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}.$$

A *holomorphic function* $f : \mathfrak{h} \to \mathbb{C}$ is a function such that for all $z \in \mathfrak{h}$ the derivative

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

exists. Holomorphicity is a very strong condition because $h \in \mathbb{C}$ can approach 0 in many ways.

Example 2.1. Let $\mathrm{SL}_2(\mathbb{Z})$ denote the set of 2×2 integers matrices with determinant 1. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, then the corresponding *linear fractional transformation*

$$\gamma(z) = \frac{az + b}{cz + d}$$

is a holomorphic function on $\mathfrak{h}$. (Note that the only possible pole of γ is $-\frac{d}{c}$, which is not an element of $\mathfrak{h}$.)

For future use, note that if $f : \mathfrak{h} \to \mathbb{C}$ is a holomorphic function, and $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, then

$$f|_\gamma(z) = f(\gamma(z))(cz + d)^{-2}$$

is again a holomorphic function.

Example 2.2. Let $q(z) = e^{2\pi i z}$. Then q is a holomorphic function on $\mathfrak{h}$ and $q' = 2\pi i q$. Moreover, q defines a surjective map from $\mathfrak{h}$ onto the punctured open unit disk $D = \{z \in \mathbb{C} : 0 < |z| < 1\}$.

3 Cuspidal Modular Forms

Let N be a positive integer and consider the set

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \; : \; N \mid c \right\}.$$

Definition 3.1 (Cuspidal Modular Form). A *cuspidal modular form* of level N is a holomorphic function $f : \mathfrak{h} \to \mathbb{C}$ such that

1. $f|_\gamma = f$ for all $\gamma \in \Gamma_0(N)$,
2. for every $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,
$$\lim_{z \to \infty} f(\gamma(z)) = 0,$$
and
3. f has a Fourier expansion:
$$f = \sum_{n=1}^{\infty} a_n q^n.$$

Exercise 3.2. Prove that condition 3 is implied by conditions 1 and 2, so condition 3 is redundant. [Hint: Since $\gamma = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, condition 1 implies that $f(z+1) = f(z)$, so there is a function $F(q)$ on the open punctured unit disc such that $F(q(z)) = f(z)$. Condition 2 implies that $\lim_{q \to 0} F(q) = 0$, so by complex analysis F extends to a holomorphic function on the full open unit disc.]

Definition 3.3. The *q-expansion* of f is the Fourier expansion $f = \sum_{n=1}^{\infty} a_n q^n$.

Exercise 3.4. Suppose that $f \in S_2(\Gamma_0(N))$. Prove that

$$f(z)dz = f(\gamma(z))d(\gamma(z))$$

for all $\gamma \in \Gamma_0(N)$. [Hint: This is simple algebraic manipulation.]

Exercise 3.5. Let $S_2(\Gamma_0(N))$ denote the set of cuspidal modular forms of level N. Prove that $S_2(\Gamma_0(N))$ forms a $\mathbb{C}$-vector space under addition.

3.1 The Dimension of $S_2(\Gamma_0(N))$

The dimension of $S_2(\Gamma_0(N))$ is

$$\dim_{\mathbb{C}} S_2(\Gamma_0(N)) = 1 + \frac{\mu}{12} - \frac{\nu_2}{4} - \frac{\nu_3}{3} - \frac{\nu_\infty}{2},$$

where $\mu = N \prod_{p \mid N}(1 + 1/p)$, and $\nu_2 = \prod_{p \mid N}\left(1 + \left(\frac{-4}{p}\right)\right)$ unless $4 \mid N$ in which case $\nu_2 = 0$, and $\nu_3 = \prod_{p \mid N}\left(1 + \left(\frac{-3}{p}\right)\right)$ unless $2 \mid N$ or $9 \mid N$ in which case $\nu_3 = 0$, and $\nu_\infty = \sum_{d \mid N} \varphi(\gcd(d, N/d))$. For example,

$$\dim_{\mathbb{C}} S_2(\Gamma_0(2)) = 1 + \frac{3}{12} - \frac{1}{4} - \frac{0}{3} - \frac{2}{2} = 0,$$

and

$$\dim_{\mathbb{C}} S_2(\Gamma_0(11)) = 1 + \frac{12}{12} - \frac{0}{4} - \frac{0}{3} - \frac{2}{2} = 1.$$

One can prove that the vector space $S_2(\Gamma_0(11))$ has basis

$$f = q \prod_{n=1}^{\infty} (1-q^n)^2(1-q^{11n})^2 = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 + \cdots .$$

Exercise 3.6. Compute the dimension of $S_2(\Gamma_0(25))$.

4 Modularity of Elliptic Curves

Let E be an elliptic curve defined by a Weierstrass equation $y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Q}$. For each prime $p \nmid \Delta = -16(4a^3 + 27b^2)$, set

$$a_p = p + 1 - \#E(\mathbb{Z}/p\mathbb{Z}).$$

Definition 4.1 (Modular). E is *modular* if there exists a cuspidal modular form

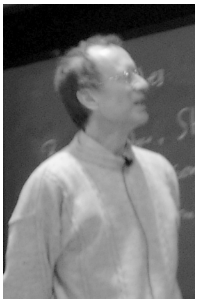
$$f(z) = \sum_{n=1}^{\infty} b_n q^n \in S_2(\Gamma_0(\Delta))$$

such that $b_p = a_p$ for all $p \nmid \Delta$.

At first glance, modularity appears to be a bizarre and unlikely property for an elliptic curve to have. When poor Taniyama (and Shimura) first suggested in 1955 that every elliptic curve is modular, people were dubious. But Taniyama was right. The proof is that conjecture is one of the crowning achievements of number theory.

Theorem 4.2 (Breuil, Conrad, Diamond, Taylor, Wiles).

Every elliptic curve over $\mathbb{Q}$ is modular.



Wiles