Elliptic curves and Fermat's last theorem

Abhiram Kumar

28th April, 2025

Outline

- Fermat's Last Theorem (FLT)
- A Remarkable Elliptic Curve
- Galois Representations
- Modularity of Elliptic Curves
- Proof overview of FLT

Fermat's Last Theorem

Fermat's last theorem states that there are no integers a, b and c satisfying $a^n + b^n = c^n$ with $abc \neq 0$ and $n > 2$.

Fermat's Last Theorem

Fermat's last theorem states that there are no integers a, b and c satisfying $a^n + b^n = c^n$ with $abc \neq 0$ and $n > 2$.
In order to prove this theorem, it is enough to the case when the exponent n is a prime.

Theorem (FLT)

If $p \geq 5$ is prime, and $a, b, c \in \mathbb{Z}$, then

$$a^p + b^p + c^p = 0 \implies abc = 0$$

Fermat's Last Theorem

Suppose (a^p, b^p, c^p) is a hypothetical solution to Fermat's equation with $abc \neq 0$.
Not all of a, b, c can be odd, so we suppose b is even. The integers a and c have to be odd and hence $\pm 1 \pmod 4$. Both cannot be $-1 \pmod 4$ since that would mean $b^p \equiv 2 \pmod 4$, contradicting $p \geq 5$.
Therefore, we may assume without losing generality that $a \equiv -1 \pmod 4$ and $2 \mid b$.

To derive a contradiction, the idea is to transform the remarkable triple (a^p, b^p, c^p) into a remarkable elliptic curve E, so remarkable that it doesn't exist.

Fermat's Last Theorem

To derive a contradiction, the idea is to transform the remarkable triple (a^p, b^p, c^p) into a remarkable elliptic curve E, so remarkable that it doesn't exist.
The way we show this is by associating to E, a modular form f whose associated Galois representation has strange properties. This would prove that such an f cannot exist.

A Remarkable Elliptic Curve

Let $p \geq 5$ be prime and let a, b, c be coprime integers satisfying $abc \neq 0$, $a \equiv -1 \pmod 4$, $2 \mid b$, and $a^p + b^p + c^p = 0$. Gerhard Frey considered the following elliptic curve:

$$E_{a^p,b^p,c^p} : y^2 = x(x - a^p)(x + b^p)$$

A Remarkable Elliptic Curve

Let $p \geq 5$ be prime and let a, b, c be coprime integers satisfying $abc \neq 0$, $a \equiv -1 \pmod 4$, $2 \mid b$, and $a^p + b^p + c^p = 0$. Gerhard Frey considered the following elliptic curve:

$$E_{a^p,b^p,c^p} : y^2 = x(x - a^p)(x + b^p)$$

Proposition

The elliptic curve E_{a^p,b^p,c^p} is semistable whose minimal discriminant and conductor are given by the formulas

- $\Delta_{a^p,b^p,c^p} = 2^{-8} \cdot (abc)^{2p}$, *and*
- $N_{a^p,b^p,c^p} = \prod_{l \mid abc} l$

The absolute Galois group of $\mathbb{Q}$ defined as $G_{\mathbb{Q}} := \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ can be endowed with a topology in which a basis of neighborhoods of the origin is given by the collection of subgroups $H \subset G_{\mathbb{Q}}$ of finite index. This makes $G_{\mathbb{Q}}$ a compact topological group.

The absolute Galois group of $\mathbb{Q}$ defined as $G_{\mathbb{Q}} := \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ can be endowed with a topology in which a basis of neighborhoods of the origin is given by the collection of subgroups $H \subset G_{\mathbb{Q}}$ of finite index. This makes $G_{\mathbb{Q}}$ a compact topological group.
A two dimensional **Galois representation** over a topological (local) ring A is defined as a continuous group homomorphism

$$\rho : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_2(A)$$

The **residual representation** $\bar{\rho} : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_2(k)$ is obtained by composing ρ with the restriction map $\mathrm{GL}_2(A) \longrightarrow \mathrm{GL}_2(k)$ where k is the residue field of A.

The local Galois group $G_{\mathbb{Q}_l}$ at a prime l is a subgroup of $G_{\mathbb{Q}}$ if we fix an embedding of $\bar{\mathbb{Q}}$ into $\bar{\mathbb{Q}}_l$.
The kernel of the natural map $G_{\mathbb{Q}_l} \longrightarrow \mathrm{Gal}(\bar{\mathbb{F}}_l/\mathbb{F}_l)$ is called the **inertia group** I_l at l
We say that a Galois representation ρ is **unramified** at l if $I_l \subset \ker \rho_{|_{G_{\mathbb{Q}_l}}}$

Tate Module

For an elliptic curve E over $\mathbb{Q}$, we have the Tate module defined as

$$T_p(E) := \varprojlim E[p^n] \cong \mathbb{Z}_p^2$$

The group $G_{\mathbb{Q}}$ acts on $T_p(E)$ and we obtain the p-adic Galois representation

$$\rho_{E,p} : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_2(\mathbb{Z}_p)$$

associated to E.

Tate Module

For an elliptic curve E over $\mathbb{Q}$, we have the Tate module defined as

$$T_p(E) := \varprojlim E[p^n] \cong \mathbb{Z}_p^2$$

The group $G_\mathbb{Q}$ acts on $T_p(E)$ and we obtain the p-adic Galois representation

$$\rho_{E,p} : G_\mathbb{Q} \longrightarrow \mathrm{GL}_2(\mathbb{Z}_p)$$

associated to E.
The residual representation $\bar{\rho}_{E,p} : G_\mathbb{Q} \longrightarrow \mathrm{GL}_2(\mathbb{F}_p)$ describes the action of $G_\mathbb{Q}$ on $E[p] \cong \mathbb{F}_p^2$.

A Remarkable Galois Representation

Gerhard Frey and Jean-Pierre Serre noted that the residual representation $\bar{\rho}_{E,p}$ coming from the Tate module of the Frey curve E has some remarkable local properties.

Theorem

The following is true for the Frey curve E:

- $\bar{\rho}_{E,p}$ *is absolutely irreducible*
- $\bar{\rho}_{E,p}$ *is odd*
- $\bar{\rho}_{E,p}$ *is unramified outside $2p$, flat at p, and semistable at 2.*

Modularity of Elliptic Curves

A **modular form** is a holomorphic function on the upper-half plane satisfying certain symmetry relations and growth conditions.

Modularity of Elliptic Curves

A **modular form** is a holomorphic function on the upper-half plane satisfying certain symmetry relations and growth conditions.

The theory of Eichler and Shimura associates to a modular form an odd two dimensional Galois representation.

A **modular form** is a holomorphic function on the upper-half plane satisfying certain symmetry relations and growth conditions.
The theory of Eichler and Shimura associates to a modular form an odd two dimensional Galois representation.
A crucial step in the proof of FLT is the following theorem due to Ribet

Theorem (Ribet)

Let f be a weight two newform of level Nl where $l \nmid N$ is a prime. Suppose $\bar{\rho}_f$ is absolutely irreducible and that one of the following is true:

- *$\bar{\rho}_f$ is unramified at l; or*
- *$l = p$ and $\bar{\rho}_f$ is flat at p.*

Then there is a weight two newform g of level N such that $\bar{\rho}_f \cong \bar{\rho}_g$.

Modularity of Elliptic Curves

A Galois representation is called **modular** if it "comes from" a modular form.
An Elliptic curve E over $\mathbb{Q}$ is called modular if $\rho_{E,p}$ is modular for all primes p.

Modularity of Elliptic Curves

A Galois representation is called **modular** if it "comes from" a modular form.
An Elliptic curve E over $\mathbb{Q}$ is called modular if $\rho_{E,p}$ is modular for all primes p.

Theorem (Wiles)

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

FLT proof overview

Consider the semistable Frey curve $E = E_{a^p,b^p,c^p}$ with conductor $N = N_{a^p,b^p,c^p}$. By Wiles's modularity theorem, we know that $\rho_{E,p}$ is modular and there is a modular form f (which will be a weight two newform of level N) such that $\rho_f \cong \rho_{E,p}$.

FLT proof overview

Consider the semistable Frey curve $E = E_{a^p,b^p,c^p}$ with conductor $N = N_{a^p,b^p,c^p}$. By Wiles's modularity theorem, we know that $\rho_{E,p}$ is modular and there is a modular form f (which will be a weight two newform of level N) such that $\rho_f \cong \rho_{E,p}$.
But $\bar{\rho}_{E,p}$ is absolutely irreducible and unramified outside $2p$ and flat at p. Ribet's theorem then implies that there is a weight two newform g of level 2 such that $\bar{\rho}_g \cong \bar{\rho}_{E,p}$.

FLT proof overview

Consider the semistable Frey curve $E = E_{a^p,b^p,c^p}$ with conductor $N = N_{a^p,b^p,c^p}$. By Wiles's modularity theorem, we know that $\rho_{E,p}$ is modular and there is a modular form f (which will be a weight two newform of level N) such that $\rho_f \cong \rho_{E,p}$.
But $\bar{\rho}_{E,p}$ is absolutely irreducible and unramified outside $2p$ and flat at p. Ribet's theorem then implies that there is a weight two newform g of level 2 such that $\bar{\rho}_g \cong \bar{\rho}_{E,p}$.
The dimension of the space of such modular forms can be computed easily and it turns out to be 0.

Consider the semistable Frey curve $E = E_{a^p,b^p,c^p}$ with conductor $N = N_{a^p,b^p,c^p}$. By Wiles's modularity theorem, we know that $\rho_{E,p}$ is modular and there is a modular form f (which will be a weight two newform of level N) such that $\rho_f \cong \rho_{E,p}$.
But $\bar{\rho}_{E,p}$ is absolutely irreducible and unramified outside $2p$ and flat at p. Ribet's theorem then implies that there is a weight two newform g of level 2 such that $\bar{\rho}_g \cong \bar{\rho}_{E,p}$.
The dimension of the space of such modular forms can be computed easily and it turns out to be 0.
Therefore such a g cannot exist, which by Ribet's theorem means that such an f cannot exist, which by Wiles's modularity theorem means that the Frey curve cannot exist, which means that Fermat's last theorem is indeed true.

Reference

- *Modular forms and Fermat's last theorem*, Springer-Verlag, New York, 1997; MR1638473