A tour of modular forms and quaternions

Alex Ghitza (University of Melbourne)

Variously joint with:

Samuele Anni	**Yiannis Fam**	**Anna Medvedovsky**
Marseille	Melbourne and LSGNT	MPIM Bonn

Why?

Maybe congruences?

D. H. Lehmer: $\tau(n) \equiv n\sigma_9(n) \pmod 7$ for all $n \geqslant 1$.

Here

$$\Delta(q) = q\prod_{n=1}^{\infty}\bigl(1-q^n\bigr)^{24} = \sum_{n=1}^{\infty}\tau(n)q^n \equiv q + 4q^2 + 5q^4 + 4q^8 + 2q^9 + \dots$$

and $\sigma_9(n) = \sum_{d|n} d^9$ appears in

$$E_{10}(q) = 1 - 264\sum_{n=1}^{\infty}\sigma_9(n)q^n \equiv 1 + 2\bigl(q + 2q^2 + 3q^4 + q^7 + 4q^8 + q^9 + \dots\bigr)$$

Elliptic curves

Let $\mathbb{F}$ be an algebraically closed field.

Elliptic curve E over $\mathbb{F}$: smooth, genus one, projective curve with distinguished point $\mathcal{O}$.

It has affine Weierstraß equations of the form

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \qquad a_j \in \mathbb{F},$$

an abelian group structure, and a one-dimensional *space of invariant differentials*

$$\omega_E = \mathrm{Span}_{\mathbb{F}} \left(\frac{dx}{2y + a_1x + a_3} \right).$$

Elliptic curve E over ring R: proper smooth curve over R with a section $\mathcal{O}$ and all of whose geometric fibres are elliptic curves as described above. It's a group scheme over R.

Has invertible sheaf $\underline{\omega}_{E/R}$ encapsulating the spaces of invariant differentials of the fibres.

Tate curve

Tate_q is an elliptic curve over $\mathbb{Z}((q))$ given by the equation

$$y^2 + xy = x^3 + B(q)x + C(q),$$

where

$$B(q) = -5\sum_{n=1}^{\infty} \sigma_3(n) q^n$$

$$C(q) = -\sum_{n=1}^{\infty} \frac{5\sigma_3(n) + 7\sigma_5(n)}{12} q^n.$$

Its canonical differential is

$$\omega_{\mathrm{can}} = \frac{dx}{2y + x}.$$

For any ring R_0 get Tate_q as an elliptic curve over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$.

Modular forms mod p

Modular form of weight k, level 1, over ring R_0, is a mapping

$$\big(E/R,\omega\big):\ \ R \text{ an } R_0\text{-algebra, } \omega \text{ non-vanishing section of } \underline{\omega}_{E/R} \mapsto f\big(E/R,\omega\big) \in R$$

that is R-isomorphism invariant, commutes with R_0-base change and is homogeneous of degree $-k$:

$$f\big(E/R,\lambda\omega\big) = \lambda^{-k} f\big(E/R,\omega\big) \qquad \text{for all } \lambda \in R^{\times}.$$

Also *holomorphic at infinity*, a condition on the Tate curve:

$$f\big(\mathrm{Tate}_q,\omega_{\mathrm{can}}\big) \in \mathbb{Z}[\![q]\!] \otimes_{\mathbb{Z}} R_0.$$

The space of all such f is

$$M_k\big(1;R_0\big).$$

When $R_0 = \overline{\mathbb{F}}_p$, we speak of *modular forms mod p*.

Higher level

Level $\Gamma_1(N)$, where N is invertible in the ring R_0: pairs $\big(E/R, \omega\big)$ replaced by triples

$$\big(E/R, \alpha, \omega\big), \qquad \alpha : \ \mu_N \hookrightarrow E.$$

If $N \geqslant 4$ there is a moduli space $Y_1(N)$ of elliptic curves with level structure.

Compactification $X_1(N)$ is a smooth projective curve over $\mathbb{Z}[1/N]$.

The space of modular forms mod p of level N and weight k is given briefly by

$$M_k\big(N; \overline{\mathbb{F}}_p\big) = \mathrm{H}^0\big(X_1(N)_{\overline{\mathbb{F}}_p}, \underline{\omega}^{\otimes k}\big).$$

Supersingular elliptic curves

Let E be an elliptic curve over $\overline{\mathbb{F}}_p$. We say that E is

- *ordinary* if $\mathrm{End}(E)$ is an order in an imaginary quadratic field, iff $E[p]$ is a group of order p;
- *supersingular* if $\mathrm{End}(E)$ is an order in the quaternion algebra D ramified at p and ∞, iff $E[p]$ is the trivial group.

The supersingular case is of particular interest to us.

Given $E/\overline{\mathbb{F}}_p$ supersingular, there exists a unique (up to $\mathbb{F}_{p^2}$-isomorphism) elliptic curve $E_0/\mathbb{F}_{p^2}$ such that E is isomorphic over $\overline{\mathbb{F}}_p$ to $E_0 \times \overline{\mathbb{F}}_p$, and the p^2-power Frobenius on E_0 is the multiplication by $-p$ map.

We call E_0 the *canonical* $\mathbb{F}_{p^2}$*-structure* on E.

Modular forms mod p as reductions

Good source of modular forms mod p: take a normalised Hecke eigenform in characteristic zero and reduce its Fourier coefficients modulo p.

Extreme(ly useful) special case: Eisenstein series of weight $p-1$

$$E_{p-1}(q) = 1 - \frac{2p-2}{B_{p-1}} \sum_{n=1}^{\infty} \sigma_{p-2}(n)\, q^n, \qquad \sigma_{p-2}(n) = \sum_{d|n} d^{p-2}.$$

Its reduction modulo p is the *Hasse invariant* $A \in M_{p-1}\big(1; \overline{\mathbb{F}}_p\big)$ that satisfies

$$A(q) = 1.$$

Viewed as a global section over $X_1(N)_{\overline{\mathbb{F}}_p}$, all the zeros of A are simple and occur precisely at the supersingular elliptic curves.

Hecke operators

The spaces $M_k\big(N;\overline{\mathbb{F}}_p\big)$ are equipped with a family of Hecke operators $\big\{T_\ell\big\}$ indexed by primes $\ell \nmid Np$.

They can be defined by explicit formulas on q-expansions: if

$$f(q) = \sum_{n=0}^{\infty} a_n q^n \qquad \text{and} \qquad \big(\langle \ell \rangle f\big)(q) = \sum_{n=0}^{\infty} b_n q^n,$$

then

$$\big(T_\ell f\big)(q) = \sum_{n=0}^{\infty} a_{\ell n} q^n + \ell^{k-1} \sum_{n=0}^{\infty} b_n q^{\ell n}.$$

Can also be given by decomposing the double coset

$$\mathrm{GL}_2\big(\mathbb{Z}_\ell\big) \begin{pmatrix} 1 & 0 \\ 0 & \ell \end{pmatrix} \mathrm{GL}_2\big(\mathbb{Z}_\ell\big)$$

or in terms of degree ℓ isogenies between elliptic curves.

The quotients W_k

Multiplication by the Hasse invariant is Hecke-equivariant, injective

$$f \longmapsto A \cdot f : \; M_{k-(p-1)}\big(N; \overline{\mathbb{F}}_p\big) \longrightarrow M_k\big(N; \overline{\mathbb{F}}_p\big).$$

We consider the Hecke module structure of the quotient

$$W_k(N) = M_k\big(N; \overline{\mathbb{F}}_p\big) / A \cdot M_{k-(p-1)}\big(N; \overline{\mathbb{F}}_p\big)$$

This behaves very regularly once $k \geqslant p+1$:

- $W_{k+p^2-1}(N) \cong W_k(N)$.

- $W_{k+p+1}(N) \cong W_k(N)[1]$.
  Tate twist of the Hecke action: T_ℓ acts as ℓT_ℓ.

- $W_{pk}(N) \cong W_k(N)$.

How does one prove such isomorphisms?

- G. Robert (1980): multiplication by E_{p+1} induces $W_k(N)[1] \cong W_{k+p+1}(N)$.

- Serre (1987–1996) uses geometry of the modular curve $X_1(N)_{\overline{\mathbb{F}}_p}$: much more soon.

- Trace formula: slightly more, later.

Serre's approach

Look at multiplication by the Hasse invariant A **at the level of sheaves** on $X_1(N)_{\overline{\mathbb{F}}_p}$

$$0 \longrightarrow \underline{\omega}^{\otimes k-(p-1)} \longrightarrow \underline{\omega}^{\otimes k} \longrightarrow \mathcal{V}_k \longrightarrow 0.$$

Take global sections, apply Serre duality etc. to get

$$0 \longrightarrow W_k(N) \longrightarrow V_k(N) \longrightarrow S_{(p+1)-k}\big(N; \overline{\mathbb{F}}_p\big)^{\vee} \longrightarrow 0.$$

So

$$W_k(N) \cong V_k(N) \qquad \text{for } k \geqslant p+1.$$

$\mathcal{V}_k$ is supported on the supersingular locus

This simplifies things considerably.

Since any supersingular elliptic curve E has a canonical $\mathbb{F}_{p^2}$-structure E_0, so does its space of invariant differentials $\omega_E \cong \omega_{E_0} \otimes_{\mathbb{F}_{p^2}} \overline{\mathbb{F}}_p$, so $\omega_E^{\otimes p^2-1}$ has a canonical basis.

Gives Hecke isomorphism

$$V_{k+p^2-1} \cong V_k.$$

V_k as functions on the quaternion algebra D

Serre pushes this further and identifies V_k with the space of functions

$$f:\ U_1(N)\backslash G\big(\mathbb{A}^\infty\big)/G(\mathbb{Q}) \longrightarrow \overline{\mathbb{F}}_p, \quad f(\lambda x) = \lambda^{-k} f(x) \text{ for all } \lambda \in \mathcal{O}_p^\times/\mathcal{O}_p^\times(1) \cong \mathbb{F}_{p^2}^\times,$$

where $G = D^\times$ is the algebraic group over $\mathbb{Q}$ given by the multiplicative group of the quaternion algebra D, and $U_1(N)$ is an appropriately chosen level structure.

[Actually, Serre worked with full level structure $\Gamma(N)$.
The case $\Gamma_1(N)$ is sketched in Edixhoven's Serre weights paper, and worked out in full detail in Yiannis Fam's MPhil thesis.
Yiannis also gives a refinement of this for fixed Dirichlet character, in particular proving the $\Gamma_0(N)$ case.]

Serre's main result

Theorem (Serre)**.** Let $N \geqslant 4$ be prime to p.
The systems of Hecke eigenvalues coming from the spaces of modular forms mod p on $X_1(N)_{\overline{\mathbb{F}}_p}$ (all weights k put together) are the same as the systems of Hecke eigenvalues coming from the spaces of locally constant functions $G\big(\mathbb{A}^\infty\big)/G(\mathbb{Q}) \longrightarrow \overline{\mathbb{F}}_p$, where G is the multiplicative group of the quaternion algebra ramified at p and ∞.

Yiannis Fam's main result

Theorem (Fam)**.** Let B be an indefinite quaternion algebra over $\mathbb{Q}$, of discriminant $\delta > 1$ relatively prime to p. Let $N \geqslant 4$ be prime to $p\delta$.
The systems of Hecke eigenvalues coming from the spaces of modular forms mod p on the Shimura curve defined by B and of level structure N (all weights k put together) are the same as the systems of Hecke eigenvalues coming from the spaces of locally constant functions $G\big(\mathbb{A}^\infty\big)/G(\mathbb{Q}) \longrightarrow \overline{\mathbb{F}}_p$, where G is the multiplicative group of the quaternion algebra ramified at $p\delta$ and ∞.

Shimura curves and false elliptic curves

Complex analytically, we have $B \hookrightarrow B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$, so picking a maximal order $\mathcal{O}_B$ and looking at the group of units of reduced norm 1, we get a discrete subgroup Γ^B of $\mathrm{SL}_2(\mathbb{R})$ and then the quotient $\Gamma^B \backslash \mathcal{H} = X^B$.

This turns out to be compact already, so no need to compactify, but also no cusps (hence no q-expansions to rely on, and no Eisenstein series).

There is a moduli interpretation though: X^B is the moduli space of *false elliptic curves*, aka abelian surfaces with quaternionic multiplication by B

$$\big(E/R, \iota\big), \qquad \iota : \mathcal{O}_B \hookrightarrow \mathrm{End}_R(E) \text{ ring homomorphism.}$$

There is a notion of supersingular false elliptic curve, and a purely algebraic-geometric definition of the Hasse invariant, etc.

A happy consequence

How could we get the isomorphism of Hecke modules

$$W_k^B(N)[1] \cong W_{k+p+1}^B(N) \qquad \text{for } k \geqslant p+1?$$

Already Serre indicated the possibility of mimicking Robert's multiplication by E_{p+1} purely in the quaternionic context.

This also works in the Shimura curve setting; we construct a function

$$\chi^B : \ G\big(\mathbb{A}^\infty\big)/G(\mathbb{Q}) \to \overline{\mathbb{F}}_p$$

such that multiplication by χ^B gives the desired isomorphism.

We are also optimistic about showing that any system of Hecke eigenvalues arising from X^B in some weight already appears, possibly up to twist, in weight $\leqslant p+1$.
(The modular curve version of this was proved by Edixhoven.)

Many other generalisations

At first sight, Serre's result may seem just an instance of the law of small numbers: the behaviour of global sections of sheaves on a curve is determined here by their restriction to a codimension one subvariety.

But the phenomenon turns out to be much more general than that:

- G (2003): Siegel modular varieties of any dimension
- Reduzzi (2013): certain Shimura varieties of PEL type
- Goldring–Koskivirta (2019), Terakado–Yu (2022): Shimura varieties of Hodge type

In each case, despite the dimension of the moduli spaces being arbitrarily large, the restriction of modular forms mod p to a natural finite set of points retains all the systems of Hecke eigenvalues.

The wild case (level Np)

Back in the modular curve setting, we now allow one factor of p to sneak into the level of the modular forms, in other words we work with the curve $X_1(Np)$.

Computationally, we still see many relations for $k \geqslant p+1$ (ss means semisimplification)

- $W_{k+p^2-p}(Np) \cong W_k(Np)$
- $W_{k+2}(Np)^{\mathrm{ss}} \cong W_k(Np)[1]^{\mathrm{ss}}$
- $W_k(Np)[(p-1)/2]^{\mathrm{ss}} \cong W_k(Np)^{\mathrm{ss}}$

How to prove these isomorphisms?

We haven't been able to find a replacement for Robert's multiplication by E_{p+1}.

The modular curve $X_1(Np)_{\overline{\mathbb{F}}_p}$ is singular, so Serre's approach becomes trickier.

Theorem (Anni–G–Medvedovsky 202?)**.** Let M_1, M_2, N_1, N_2 be free $\mathbb{Z}_p$-modules of finite rank, each with an action of an operator T. Let $\overline{M_1} = M_1 \otimes \overline{\mathbb{F}}_p$, etc.
Suppose we have T-equivariant embeddings $\iota_1 : \overline{N_1} \hookrightarrow \overline{M_1}$ and $\iota_2 : \overline{N_2} \hookrightarrow \overline{M_2}$ and consider the quotients

$$W_1 = \overline{M_1}/\iota_1\big(\overline{N_1}\big), \qquad W_2 = \overline{M_2}/\iota_2\big(\overline{N_2}\big).$$

Then $W_1^{\mathrm{ss}} \cong W_2^{\mathrm{ss}}$ as $\overline{\mathbb{F}}_p[T]$-modules if and only if for every $n \geqslant 0$ we have

$$\big(\mathrm{tr}(T^n|M_1) - \mathrm{tr}(T^n|N_1)\big) - \big(\mathrm{tr}(T^n|M_2) - \mathrm{tr}(T^n|N_2)\big) \equiv 0 \pmod{p^{1+v_p(n)}}.$$

Using the Eichler–Selberg trace formula and the previous theorem, we prove

Theorem (Anni–G–Medvedovsky 202?)**.** For $k \geqslant p+3$ we have

$$W_{k+2}(Np)^{\mathrm{ss}} \cong W_k(Np)[1]^{\mathrm{ss}} \qquad \text{and} \qquad W_k(Np)[(p-1)/2]^{\mathrm{ss}} \cong W_k(Np)^{\mathrm{ss}}.$$