Theta operators on (p-adic) automorphic forms and applications

$\S\,1$. Historical overview

Ramanujan, 1910s

Theta operator was first introduced by Ramanujan around 1910s. It is the differential operator on the space of modular forms which acts on q-expansions by

$$\theta = q\frac{d}{dq},$$

so that $\theta\left(\sum_n a_n q^n\right) = \sum_n n a_n q^n$. Theta operator almost increases the weight of a given modular form by 2. More precisely, for a weight k modular form $f \in M_k(\mathbb{C})$,

$$(12\theta - kE_2)f \in M_{k+2}(\mathbb{C}),$$

where E_2 denotes the Eisenstein series of weight 2, which is only quasi-modular.

Notation: $M_k(\star)$ will always denote the space of modular forms with coefficient in $\star$ of some fixed level. For instance,

- $\star = \mathbb{C}$ - classical modular forms
- $\star = \mathbb{F}_p$ - Swinnerton-Dyer & Serre's mod p modular forms (by taking reduction mod p of q-expansions)
- $\star = R$ - Katz modular forms

If you haven't seen these before, it might be helpful to think of $M_k(\star)$ just as modular forms whose Fourier coefficients lie in $\star$; but really, Katz modular forms should be interpreted as a way of thinking about modular forms

- $\star = \mathbb{Q}_p$ - Serre's p-adic modular forms

However, $M_k(\mathbb{Q}_p)$ is something different, given by taking "p-adic completion of $M_k(\mathbb{Q})$". We will say a bit more on this in a moment.

Swinnerton-Dyer, Serre, and Katz, 1970s

Note that $E_{p+1} \equiv E_{p-1}E_2 \equiv E_2 \pmod{p}$, one can replace E_2 with a genuine modular form in the mod p setting, so that one can make sense of θ as an actural differential operator on the space of mod p modular forms. The price we pay, however, is to shift the weight even further.

Theorem (Swinnerton-Dyer, Serre; Katz, 1970s).

For $p \geq 5$, θ induces a differential operator

$$\mathrm{Ha} \cdot \theta \colon M_k(\mathbb{F}_p) \to M_{k+p+1}(\mathbb{F}_p)$$

that acts on q-expansions as $q \cdot (d/dq)$ for every k.

Here, Ha denotes the special mod p modular form called "the Hasse invariant", which has constant q-expansion 1 and can be regarded as the mod p reduction of E_{p-1}. (When evaluated on elliptic curves over finite fields, Ha indeed spits out their Hasse invariant, whence the name. In particular, Ha vanishes precisely at supersingular points.) $\mathrm{Ha} \cdot \theta$ means multiplying θ by the Hasse invariant (to clear the poles at supersingular points).

Swinnerton-Dyer & Serre obtained this result by an elementary approach via q-expansions. Around the same time, Katz reformulated the theory of modular forms geometrically, and constructed this mod p theta operator via the Gauss--Manin connection on de Rham cohomology. Let's now take a digression and discuss about this geometric perspective, since this will be crucial when defining OC forms for instance.

Roughly speaking, modular curves X are moduli spaces of (generalized) elliptic curves with level structure, and thus algebraic in nature and defined over $\mathbb{Q}$ (or even $\mathbb{Z}[1/N]$ for some N). From this point of view, modular forms are sections of automorphic line bundles on modular curves. For instance,

- $M_k(\mathbb{C}) \simeq \mathrm{H}^0(X_{\mathbb{C}}, \omega^k)$;

- $M_k(\mathbb{F}_p) \simeq \mathrm{H}^0(X_{\mathbb{F}_p}, \omega^k)$;
- $M_k(R) \simeq \mathrm{H}^0(X_R, \omega^k) \simeq \mathrm{H}^0(X, \omega^k) \otimes R$ - so we won't get anything new just by base change

However, something very interesting happens in the p-adic completion process:

- $M_k(\mathbb{Q}_p) \simeq \mathrm{H}^0(\mathcal{X}^{\mathrm{ord}}, \omega^k)$ where $\mathcal{X}$ is the rigid (or adic) analytification of $X_{\mathbb{Q}_p}$ and $\mathcal{X}^{\mathrm{ord}}$ is the ordinary locus of $\mathcal{X}$.

Here's a heuristic reason why we only look at ordinary locus: $E_{p-1}^{p^n} \to 1$ p-adically, and so

$$E_{p-1}^{-1} := \lim_{n\to\infty} E_{p-1}^{p^n-1}$$

exists as a p-adic modular form. In other words, taking p-adic limit makes E_{p-1} invertible. Remember $E_{p-1} \equiv \mathrm{Ha} \pmod{p}$, and the latter vanishes exactly at supersingular points, so inverting E_{p-1} means allowing poles at supersingular points.

Note also that "cutting out the ordinary locus" forces us to delve into the realm of formal or rigid analytic geometry intuitively because requiring that "some section is invertible (non-vanishing mod p)" is not legal in algebraic geometry.

Now, the key player of today is

- $M_k^{\dagger,v}(\mathbb{Q}_p) := \mathrm{H}^0(\mathcal{X}^{\leq v}, \omega^k) \subset M_k(\mathbb{Q}_p)$, where $\mathcal{X}^{\leq v}$ is a strict neighborhood of the ordinary locus where elliptic curves are "not too supersingular".

The theory of p-adic modular forms already proves fruitful, for instance, Serre himself derived a new construction of the p-adic zeta function immediately using p-adic families of Eisenstein series.
You may then wonder why we bother introducing the weird-looking space $M_k^{\dagger,v}(\mathbb{Q}_p)$. Well, one important reason is that $M_k(\mathbb{Q}_p)$ is a huge space of uncountable dimension, which contains a lot of modular forms which whose arithmetic significance is less apparent. The OC space $M_k^{\dagger,v}(\mathbb{Q}_p)$, on the other hand, admits a well-behaved spectral theory of the Hecke operator U_p, which is very interesting in its own right. However, slope (the p-adic valuation of the U_p-eigenvalue) of modular forms is an exciting topic for another day, unfortunately.

P.s. a very high level reason: OC forms appears naturally as certain locally analytic vectors in the completed cohomology of modular curves...

Coleman, 1990s

Even though $\mathrm{Ha}\cdot\theta$ is not divisible by Ha, $(\mathrm{Ha}\cdot\theta)^{k+1}$ suddenly becomes divisible by Ha^{k+1}, and thus

$$\theta^{k+1}\colon M_{-k}(\mathbb{F}_p)\to M_{k+2}(\mathbb{F}_p)$$

makes sense (we have to start from weight $-k$ for reasons that will hopefully be clear later). Of course, this then gives rise to a map $\theta^{k+1}\colon M_{-k}(\mathbb{Q}_p)\to M_{k+2}(\mathbb{Q}_p)$ between p-adic modular forms.

In his celebrated classicality paper (1996), Coleman gave a cohomological construction of $\theta^{k+1}\colon\omega^{-k}\to\omega^{k+1}$ on the level on sheaves, and as a result θ^{k+1} preserves the subspace OC forms, which is certainly not the case for θ itself. From there on, his ingenious dimension counting arguments allowed him to deduce that OC forms small slope are classical. At the heart of his work lies a deep comparison result between rigid and coherent cohomology, which we will now move towards.

Applications

- The weight part of Serre's conjecture (Edixhoven)
- Classicality (Coleman): Let f be an OC form of weight $k+2$ of slope $\leq k+1$. If f does not lie in the image of θ^{k+1}, then f is classical (of Iwahori level at p for the experts).
- p-adic -functions over CM fields (Katz)

$\S\,2$. Perspectives on theta operators

The dual BGG complex

In 1980s, in order to understand singular cohomology of locally symmetric spaces, Faltings introduced the dual BGG complex and represented the former in terms of coherent cohomology groups which are certain spaces of automorphic forms.

To begin with, let me mention the classical BGG complex in representation theory of Lie algebras.

- $(,X)$ Shimura datum
- Lie algebra of
- $\subset$ split Cartan subalgebra (the eigenvalues of ad are in the field, for all $\in$)
- $\subset$ root system of relative to
- $\subset$ fixed root basis
- $^{+}$ positive roots
- Weyl group

It is well-known that finite-dimensional irreducible representations of are parametrized by the set of dominant integral weights . One way of constructing the highest weight representation $()$ of weight is to realize it as the quotient of the Verma module $M()$ by its unique maximal submodule,

$$M() \to () \to 0$$

General theory of Verma modules show that there exists non-trivial morphisms between Verma modules of two different weights if and only if these two weights are "strongly linked", and in which case the morphisms are injective and the Hom space actually has rank 1. It follows that the above sequence can be extended to

$${}_{\in()=1} M(\,\cdot\,) \to M() \to () \to 0,$$

where $\cdot$ stands for the dot action.

Bernstein--Gelfand--Gelfand completed this to an exact sequence,

$$0 \to M(_0 \cdot\,) \to {}_{\in()=^{+}-1} M(\,\cdot\,) \to \ \to {}_{\in()=1} M(\,\cdot\,) \to M() \to () \to 0$$

One way to think about this is to view it as a resolution of the highest weight representation $()$ in terms of Verma modules, whose structure is much simpler and thus is easier to understand. For instance, one can deduce Weyl's character formula of $()$ from the BGG complex, and in fact the BGG complex gives a categorification (in the so-called BGG category) of Weyl's character formula.

Now, there is a (contra-variant in nature) bundle functor that translates this BGG complex into a complex of automorphic vector bundles on the Shimura variety associated to $(,X)$. In particular, Faltings proved a deep result relating morphisms between Verma modules and differential operators between the corresponding automorphic vector bundles. For example, in the case of modular curves, the dual BGG complex simply (with respect to weight k) reads

$$\omega^{-k} \to \omega^{k+2}$$

and it gives a resolution of the de Rham cohomology $\mathrm{m}^k\mathrm{H}^1_{\mathrm{d}}(X)$ of the modular curve. The differential operator here in this complex coincides with Coleman's θ^{k+1}.

Remark.

1. For practical purposes, one works with certain parabolic subgroups instead of just Borel and use "generalized Verma modules".
2. The dual BGG complex for a wide class of Shimura varieties has been well-understood by the work of Faltings, Faltings--Chai, Lan--Polo, and other people. It works integrally, has a mod p analogue, and can be extended to rigid analytic settings, etc.
3. It follows that rigid cohomology of Shimura varieties can be computed by the dual BGG complex, and the latter involves various spaces of overconvergent modular forms of different weights and image of theta operators. Upshot: Coleman's classicality theorem is really a comparison theorem between systems of Hecke eigenvalues appearing in rigid cohomology and those in classical coherent cohomology.

Some recent developments

- unitary Shimura varieties (de Shalit, Goren, ...)
- PEL Shimura varieties of type A and C (Eischen, Mantovan, ...)
- classicality in the Hilbert case generalizing Coleman's approach (Tian--Xiao)
- applications to p-adic -functions and Iwasawa theory, the theory of OC forms, ...

Other approaches

- Group cohomology avatar (Serre, Stevens, Edixhoven--Khare 2000s)
- Theta operator as the Fontaine operator (Jiang 2023)

- Theta operator as differential of a -action on the big Igusa tower (Howe 2020, 2025)
-

§ . What's next?

Ortiz's thesis

- $(, X)$ be a Shimura datum of Hodge type
- integral model of the Shimura variety of some hyperspecial level at p
- special fiber
- $^{\mathrm{or}}$ toroidal compactification
- $\mathrm{Fl} \to$ flag variety parametrizing full flags of the Hodge bundle ω on

Theorem (Ortiz, 2024; 2026).

Let $\in {}^{+}$ *be a positive root. For each* $\in X()$ *there exists a basic theta operator* $\theta\colon () \to (+)$ *which is a differential operator over* $\mathrm{Fl}_{\mathbb{F}_p}$*, and is an explicit weight depending on . These basic theta operators extend to toroidal compactifications and satisfy the following properties:*

1. *Hecke equivariance away from* $\mathbf{p}$*;*
2. *commutation relations - for* ${}_1, {}_2 \in {}^{+}$*,* $[\theta_1, \theta_2] = \theta_{1+2}$ *up to an explicit collection of Hasse invariants and a constant scalar; (in particular, they do not naïvely commute with each other in general)*
3. *differential operators coming from representation theory (e.g. the ones in the dual BGG complex) can be written as non-commutative polynomials on basic theta operators + Hasse invariants;*
4.

In the Siegel 3-fold case, there are 4 basic theta operators, and they can be constructed geometrically first over the ordinary locus (and then extended to the full flag variety by multiplying a suitable collection of Hasse invariants). Hence, they can also be viewed as theta operators on the space of p-adic Siegel modular forms.

Question A & B: Can one also construct geometrically the theta operators that preserve OC forms and characterize them?

- On the one hand, it is rare that theta operators would preserve OC forms for various reasons. Hence, it is not clear at all what they would "look like", nor which linear combinations of the basic theta operators do have this property.
- On the other hand, theta operators in the dual BGG complex do have this property a priori. I expect that these are precisely the only possible ones.
- One way of attacking this question is to move to the Galois side and analyze the effect of theta operators on Galois representations attached to p-adic automorphic forms.
  Question C: Can one extend Coleman's approach to classicality beyond the Hilbert case?
  Question D: Do we have "boundary classicality" criterion similar to that of Coleman?