

# 1. GEBHARD BOECKLE'S LECTURES

## Galois representations and congruences.

We first discuss profinite groups. Let  $G$  be a topological group.

**Theorem 1.1.** *The following are equivalent:*

- (a)  $G$  is compact, Hausdorff, and totally disconnected.
- (b)  $G$  is compact, and admits a neighborhood basis of the identity by open normal subgroups.
- (c) There is a directed poset  $I$  and an inverse system  $(G_i)$  of finite (discrete) groups such that  $G = \varprojlim_I G_i$ .

We say that  $G$  is *profinite* if the above conditions hold. The topology on  $\varprojlim G_i$  is that obtained by regarding it as a closed subgroup of the product  $\prod G_i$ .

Constructions:

- (a) If  $G$  is discrete, then we equip it with the profinite topology  $G^{\text{pf}} := \varprojlim G/N$ , where  $N$  runs over the finite index subgroups.
- (b) If  $G = \varprojlim G_i$  is profinite, then
  - (i) The abelianization is given by

$$G^{\text{ab}} = G/[\overline{G, G}] = \varprojlim G_i^{\text{ab}},$$

and in particular, is profinite.

- (ii) For  $H$  finite, write  $H_p$  for its maximal  $p$ -group quotient. Then

$$G_p = \varprojlim (G_i)_p$$

is a pro- $p$ -group (and in particular, profinite).

- (iii) If  $N \leq G$  is closed and normal, then  $G/N$  is profinite.

**Example 1.2.** (a) Let  $F$  be a field. Set  $G_F := \text{Aut}_F(F^{\text{sep}}) = \text{Gal}(F^{\text{sep}}/F)$  profinite. Define the poset

$$\mathcal{I}_F := \{L \subseteq F^{\text{sep}} : L \supseteq F \text{ finite Galois, } \subseteq\}.$$

Then

$$G_F \xrightarrow{\cong} \varprojlim_{L \in \mathcal{I}_F} \text{Gal}(L/F).$$

- (b) Let  $F' \subseteq F^{\text{sep}}$  be a normal extension of  $F$ . Then  $G_{F'} \leq G_F$  is closed and normal. We may thus write

$$\text{Gal}(F'/F) \cong G_F/G_{F'} = \varprojlim_{\substack{L \in \mathcal{I}_F, \\ L \subseteq F'}} \text{Gal}(L/F).$$

- (c) Let  $\mathbb{N}$  denote the natural numbers, ordered by divisibility. Then

$$\hat{\mathbb{Z}} = \varprojlim \mathbb{Z}/n = \prod_p \mathbb{Z}_p,$$

where the last step is the Chinese remainder theorem. We sometimes need a slight modification:

$$\hat{\mathbb{Z}}^{(p)} = \varprojlim_{p \nmid n} \mathbb{Z}/n = \prod_{\ell \text{ prime}, \ell \neq p} \mathbb{Z}_\ell.$$

Let's fix some notation:

- (a) Let  $K$  be a number field,  $\mathcal{O}_K$  its ring of integers. Let  $\text{Pl}_K = \text{Pl}_K^\infty \sqcup \text{Pl}_K^{\text{fin}}$  denote the set of places  $v$  of  $K$ . Let  $v$  be a finite place. We may then attach to it a maximal ideal  $\mathfrak{q}_v$  of  $\mathcal{O}_K$ , giving a bijection

$$\text{Pl}_K^{\text{fin}} \leftrightarrow \text{Max}(\mathcal{O}_K).$$

We may form the residue field  $k_v := \mathcal{O}_K/\mathfrak{q}_v$ . We denote  $q_v$  for the cardinality of  $k_v$ . We write  $\text{char}(v)$  for the characteristic of  $k_v$ . We denote by  $\mathcal{O}_v = \varprojlim \mathcal{O}/\mathfrak{q}_v^n$ , with fraction field  $K_v$ . Also, we have a short exact sequence

$$1 \rightarrow I_v \rightarrow G_v := \text{Gal}_{K_v} \rightarrow \text{Gal}_{k_v} \rightarrow 1.$$

A topological generator for  $\text{Gal}_{k_v}$  is given by

$$\text{Fr}_v : \alpha \mapsto \alpha^{q_v}.$$

We denote by  $\text{Frob}_v \in G_v$  some lift of  $\text{Fr}_v$ .

We write  $S_\infty := \text{Pl}_K^\infty$  for the set of archimedean places, so that  $K \otimes_{\mathbb{Q}} \mathbb{R} \cong \prod_{v \in S_\infty} K_v$ . For a rational prime  $p$ , we write  $S_p$  for the set of places  $v$  of  $K$  such that  $v \mid p$ .

- (b) We also need some local analogues for  $E \supseteq \mathbb{Q}_p$  a  $p$ -adic field. Let  $\mathcal{O} = \mathcal{O}_E$  denote the ring of integers,  $\pi = \pi_E$  a uniformizer, and  $\mathbb{F} = \mathcal{O}_E/\pi$  the residue field, with  $q = \#\mathbb{F}$ . Then  $E \supseteq \mathbb{Q}_q = \mathbb{Q}_p[\zeta_{q-1}] \supseteq \mathbb{Q}_p$ . We have  $W(\mathbb{F}) = \mathbb{Z}_q = \mathbb{Z}_p[\zeta_{q-1}]$ .

Continuing the examples, which may serve as exercises:

- (d) Let  $\zeta_t$  be a primitive  $t$ th root of 1. For  $k$  a finite field, we have  $G_k \cong \hat{\mathbb{Z}} = \overline{\langle \text{Fr}_k \rangle}$ , where  $\text{Fr}_k : \alpha \mapsto \alpha^{|k|}$ .
- (e) Let  $E \supseteq \mathbb{Q}_p$  (finite extension). Then  $G_E$  (Jannsen–Wingberg for  $p \geq 2$ ). Local class field theory: the Artin map  $E^\times \rightarrow G_E^{\text{ab}}$  is a continuous inclusion with dense image. Writing  $E^\times = \pi_E^{\mathbb{Z}} \times \mathcal{O}_E^\times = \pi_E^{\mathbb{Z}} \times \mathbb{F}^\times \times \mathcal{U}_E$ . Since the units are known to be a finitely generated  $\mathbb{Z}_p$ -module, we get as a corollary that

$$\text{Hom}_{\text{cts}}(G_E, \mathbb{F}_p) = H_{\text{cts}}^1(G_E, \mathbb{F}_p)$$

is finite.

- (f) We turn to the case of a number field  $K$ . We fix an embedding  $K^{\text{sep}} \subseteq K_v^{\text{sep}}$  for each place  $v$ , which gives an embedding of Galois groups  $G_v \rightarrow G_K$ . For  $S \subseteq \text{Pl}_K$  finite, we write

$$K_S := \{\alpha \in K^{\text{sep}} : K(\alpha) \text{ is unramified outside } S\},$$

which is a normal (typically infinite) extension of  $K$ . We write

$$G_{K,S} := \text{Gal}(K_S/K) = G_K/G_{K_S}$$

for its Galois group. We remark that if we take  $v \notin S$ , then since  $v$  does not ramify in  $K_S$ , we know that the map  $G_v \rightarrow G_{K,S}$  factors via the quotient  $G_v/I_v \cong G_{k_v}$ , so that  $\text{Frob}_v \in G_{K,S}$  is independent of the choice of lift. On the other hand, if  $v \in S$ , then we might ask whether the map  $G_v \hookrightarrow G_{K,S}$  (see the work of Chenievr–Clozel). The structure of  $G_{K,S}$  is unknown, but global class field theory describes  $G_{K,S}^{\text{ab}}$ . A corollary is that

$$H_{\text{cts}}^1(G_{K,S}, \mathbb{F}_p) = \text{Hom}_{\text{cts}}(G_{K,S}, \mathbb{F}_p)$$

is finite whenever  $S$  is finite. (One can appeal to Hermite–Minkowski, or class field theory.)

- (g) Consider the tame quotient of  $G_E$ , for  $E \supseteq \mathbb{Q}_p$ . Given  $E \supseteq \mathbb{Q}_p$ , we form the tower of extensions  $E^{\text{tame}}/E^{\text{unr}}/E$ , where

$$E^{\text{unr}} = \cup \{E(\zeta_n) : p \nmid n\},$$

$$E^{\text{tame}} = \cup \{E^{\text{unr}}(\sqrt[p]{\pi_E}) : p \nmid n\}.$$

It's a fact that  $G_E^{\text{tame}}$  may be expressed as the profinite completion of  $\langle st : sts^{-1} = t^q \rangle$ .

We finally come to **Galois representations**. They will typically be called  $\rho : G \rightarrow \text{GL}_n(A)$ , where  $G$  is a topological group,  $A$  is a topological ring, and  $\rho$  is a continuous map. The topology on  $\text{GL}_n(A)$  is the subspace topology coming from embedding inside  $M_n(A) \times A$  via  $g \mapsto (g, \det(g)^{-1})$ , for instance. We call  $\rho$  a Galois representation if  $G = G_F$  for some field  $F$ . The main examples of interest for  $A$  will be  $\mathbb{C}$ , finite fields, and  $p$ -adic fields, to interpolate  $\text{CNL}_{\mathcal{O}}$  (complete Noetherian local  $\mathcal{O}$ -algebras).

**Exercise 1.** Let  $G$  be profinite, and  $\rho$  as above.

- (a) If  $A = \mathbb{C}$ , then  $\rho(G)$  is finite.
- (b) If  $A = \overline{\mathbb{O}_p}$ , then there is a finite extension  $E \supseteq \mathbb{Q}_p$  such that  $\rho(G) \subseteq \text{GL}_n(E)$  up to conjugation.
- (c) If  $A = E \supseteq \mathbb{Q}_p$  (finite extension), then after conjugation, we can assume that  $\rho(G) \subseteq \text{GL}_n(\mathcal{O})$ .

In case (c), we have a  $G$ -stable lattice  $\Lambda \cong \mathcal{O}^n \subseteq E^n$ . We can apply reduction  $\mathcal{O} \rightarrow \mathbb{F}$ . This gives a reduction

$$\bar{\rho}_{\Lambda} : G \rightarrow \text{GL}_n(\mathbb{F}).$$

Let's use the notation  $\text{cp}_{\alpha}$  for the characteristic polynomial of  $\alpha \in M_n(A)$ .

**Theorem 1.3.** (a) Given a representation  $r : G \rightarrow \text{GL}_n(\mathbb{F})$ . Then there exists a semisimple representation  $r^{\text{ss}} : G \rightarrow \text{GL}_n(\mathbb{F})$  such that  $\text{cp}_r = \text{cp}_{r^{\text{ss}}}$  (Brauer–Nesbitt), where  $r^{\text{ss}}$  is unique up to isomorphism.  
 (b) We have  $\text{cp}_{\rho} \in \mathcal{O}[X]$  and  $\text{cp}_{\bar{\rho}_{\Lambda}} \in \mathbb{F}[X]$ , independent of  $\Lambda$ .

**Theorem 1.4.** For  $\rho, \rho' : G_{K,S} \rightarrow \text{GL}_n(E)$  semisimple, we have that  $\rho \sim \rho'$  (conjugate) if and only if for all  $v \in \text{Pl}_K^{\text{fin}} \setminus S$ , we have

$$\text{cp}_{\rho(\text{Frob}_v)} = \text{cp}_{\rho'(\text{Frob}_v)}.$$

**Example 1.5.** (1)  $p$ -adic cyclotomic character  $\chi_p^{\text{cyc}} : G_{\mathbb{Q}} \rightarrow \mathbb{Z}_p^{\times}$ . We have

$$G_{\mathbb{Q}} \circ \mu_{p^n} = \langle \zeta_{p^n} \rangle \cong \mathbb{Z}/p^n,$$

$(\mathbb{Z}/p^n)^{\times} = \text{Aut}_{\mathbb{Z}}(\mathbb{Z}/p^n)$ . **Facts:**

- $\chi_p^{\text{cyc}}|_{G_K}$ : unramified outside  $S_p \cup S_{\infty}$ .
- $\chi_p^{\text{cyc}}(\text{Frob}_v) = q_v \in \mathbb{Z}_p^{\times}$ .

- (2) The Tate module of an elliptic curve  $\mathcal{E}_K$ . We again have  $G_K \circ \mathcal{E}[p^n](\bar{K}) \cong (\mathbb{Z}/p^n)^{\oplus 2}$ , which gives rise to  $G_K \rightarrow \text{GL}_2(\mathbb{Z}/p^n)$ . In the limit, we get

$$\rho_{\mathcal{E},p} : G_K \rightarrow \text{GL}_2(\mathbb{Z}_p) \hookrightarrow \text{GL}_2(\mathbb{Q}_p).$$

**Facts:**

- $\rho_{\mathcal{E},p}$  is unramified outside  $S_{\infty} \cup S_p \cup \text{Bad}$ .

- For  $v$  outside those places, we have

$$\text{cp}_{\rho_{\mathcal{E},p}}(\text{Frob}_v) = X^2 - a_v(\mathcal{E})X + q_v,$$

where

$$a_v := \#\mathcal{E}(k_v).$$

This shows the geometric meaning of Frobenius.

- (3) Let  $f = q + \sum_{n \geq 2} a_n q^n$  be the  $q$ -expansion of a cuspidal Hecke eigenform  $f \in S_k(N, \varepsilon)$ ,  $k \geq 1$ ,  $\varepsilon : (\mathbb{Z}/N)^\times \rightarrow \mathbb{C}^\times$ .

**Theorem 1.6.** (a)  $E_f = \mathbb{Q}(a_n : n \geq 1)$  is a number field, with  $a_n \in \mathcal{O}_{E_f}$  for all  $n$ .

- (b) (Eichler–Shimura, Deligne, Deligne–Serre) For all finite places  $\lambda$  of  $E_f$  (with  $E_\lambda$  the completion of  $\lambda$  and  $p$  the characteristic of  $k_\lambda$ ) there exists an absolutely irreducible representation

$$\rho_{f,\lambda} : G_{\mathbb{Q}, N \cup \{p, \infty\}} \rightarrow \text{GL}_2(\bar{E}_\lambda)$$

and for all primes  $\ell \nmid Np$ , we have the relation that we just saw in Chris’s talk:

$$\text{cp}_{\rho_{f,\lambda}}(\text{Frob}_\ell) = X^2 - a_\ell X + \varepsilon(\ell)\ell^{k-1}.$$

This characterizes the representation and maybe gives the main link to Galois representation.

Now, we want to study congruences. To do this, we first go from

- cusp forms that start life over the complex numbers on the upper half plane, to
- Fourier coefficients, that live over the integers.

(For simplicity, assume that nebentypus is trivial:  $\varepsilon = 1$ .) Let  $S_k(N, \mathbb{Z})$  denote the set of all  $f = \sum_{n \geq 1} a_n q^n$  such that  $a_n \in \mathbb{Z}$  for all  $n$ .

**Fact 1.7.**  $S_k(N, \mathbb{Z})$  is a  $\mathbb{Z}$ -module of rank equal to  $\dim_{\mathbb{C}} S_k(N)$ .

For any ring  $A$ , we have

$$\underline{S}_A = S_k(N, \mathbb{Z}) \otimes_{\mathbb{Z}} A \hookrightarrow \mathbb{T}_A = \mathbb{T}(N, A).$$

**Definition 1.8.** For Hecke eigenforms  $f = \sum a_n q^n$  and  $g = \sum b_n q^n \in S_k(N, \bar{\mathbb{Z}}_p)$ , we say that  $f \equiv g \pmod{p}$  if the following equivalent conditions hold:

- for all primes  $\ell \nmid Np$ , we have  $a_\ell \equiv b_\ell \pmod{\mathfrak{m}_{\bar{\mathbb{Z}}_p}}$ .
- $\rho_f \equiv \rho_g \pmod{\mathfrak{m}_{\bar{\mathbb{Z}}_p}}$  as maps  $G_{\mathbb{Q}} \rightarrow \text{GL}_2(\bar{\mathbb{F}}_p)$ .

We pass to the same Hecke algebra. Let  $\mathbb{T}'_A$  denote the subalgebra of  $\mathbb{T}_A$  generated by  $T_\ell$  for all  $\ell \nmid N$ . This is acted on by  $\underline{S}_A$ .

**Fact 1.9.** We have bijections between the following:

- $\mathbb{T}'$ -Hecke eigensystems of forms in  $S_k(N)$ .
- $\text{Hom}_{\mathbb{C}}(\mathbb{T}'_{\mathbb{C}}, \mathbb{C})$ .
- $\text{Hom}_{\mathcal{O}}(\mathbb{T}'_{\mathcal{O}}, \mathcal{O})$ , where we choose  $E$  large enough with  $\mathbb{C} \supseteq \bar{\mathbb{Q}} \subseteq \bar{\mathbb{Q}}_p \supseteq E \supseteq \mathcal{O}$ , where  $\mathcal{O}$  always denotes the ring of integers of  $E$ .

We also have a bijection between

- $\text{Hom}_{\mathcal{O}}(\mathbb{T}_{\mathcal{O}}, \mathbb{F})$ , and

- The set of  $\bar{\rho} : G_{\mathbb{Q}} \rightarrow \mathrm{GL}_2(\mathbb{F})$  that are semisimple reductions from some  $\rho_f$ , with  $f \in S_k(N)$ .

$\mathbb{T}'$  is a finite free  $\mathcal{O}$ -algebra. For any  $\bar{\rho}$ , we get a maximal ideal  $\mathfrak{m}_{\bar{\rho}} \subseteq \mathbb{T}'_{\mathcal{O}}$ . For a given  $\bar{\rho}$ , if we take

$$(\mathbb{T}'_{\mathcal{O}})_{\mathfrak{m}_{\bar{\rho}}},$$

then this is the relevant Hecke algebra for understanding the forms congruent to  $\bar{\rho}$ .

**Theorem 1.10** (Carayol, Serre). Assume that  $\bar{\rho}$  is absolutely irreducible. Then there exists a continuous representation

$$\rho_{\bar{\rho}}^{\mathrm{mod}} : G_{\mathbb{Q}, N \cup \{p, \infty\}} \rightarrow \mathrm{GL}_2((\mathbb{T}'_0)_{\mathfrak{m}_{\bar{\rho}}}) \quad (1.1)$$

“built” from the  $\rho_f$ , for  $f \in S_k(N, \bar{\mathbb{Z}}_p)$ , with  $\rho_f \equiv \bar{\rho}$  modulo  $\mathfrak{m}_{\bar{\mathbb{Z}}_p}$ .

(One can write down a characteristic polynomial, similar to the above.) Maybe one should also say that

$$(\mathbb{T}'_{\mathcal{O}})_{\mathfrak{m}_{\bar{\rho}}} \subseteq \prod_{f, \rho_f \equiv \bar{\rho}} \mathcal{O}, \quad (1.2)$$

where the left hand side is generated by all the

$$\{\mathrm{trace} \rho_f(\mathrm{Frob}_{\ell}) : \rho_f \equiv \bar{\rho}, \ell \nmid Np\}.$$

Now, is the inclusion (1.2) strict? If you take twice the same form, then you get twice the same value.

**Vision of Mazur?** Can  $\rho_{\bar{\rho}}^{\mathrm{mod}}$  be characterized purely in terms of “Galois” theory, maybe at least once  $\bar{\rho}$  is given? You need to start somewhere. Start with the mod  $p$  representation. Then there’s a sort of  $p$ -adic representation (1.1) that you get here, which sees all forms congruent to  $\bar{\rho}$ . Can you see this sort of thing from a purely Galois-theoretic perspective? This might have been the starting point for the study of deformations of Galois representations. Mazur maybe had one example in mind: Hida had just written down his Hida families, which are much bigger than things of fixed weight and level.

Let’s now turn to *deformation functors*. In many ways, what we’re doing now at the beginning is very formal, and maybe in the next lecture it gets a bit more Galois-theoretic. Let  $G$  be a profinite group – think of some global Galois group. Let  $E \supseteq \mathcal{O} \rightarrow \mathbb{F}$  be a local field, with uniformizer  $\pi$  and residue field cardinality  $q$ . In what natural category do these  $\mathcal{O}$ -algebras  $\mathbb{T}'_{\mathcal{O}}$  live?

**Definition 1.11.**  $\mathrm{CNL}_{\mathcal{O}}$  is the category of complete noetherian  $\mathcal{O}$ -algebras  $(A, \mathfrak{m}_A)$  with residue field  $\mathbb{F}$  and with local homomorphisms.

Typical rings:  $\mathcal{O}[[X_1, \dots, X_r]]/I$ . Some structure theorem tells you that this is all you can get. There is a finite subcategory

$$\mathrm{Ar}_{\mathcal{O}} \subseteq \mathrm{CNL}_{\mathcal{O}}$$

consisting of Artin objects.

**Question 1.12.** For  $\bar{\rho} : G \rightarrow \mathrm{GL}_n(\mathbb{F})$ , when is the functor

$$D_{\bar{\rho}} : \mathrm{CNL}_{\mathcal{O}} \rightarrow \mathrm{Set}$$

$$(A, \mathfrak{m}_A) \mapsto \{\rho_A : G \rightarrow \mathrm{GL}_n(A) \mid \rho \bmod \mathfrak{m}_A \equiv \bar{\rho}\} / \sim$$

representable by  $R_{\bar{\rho}} \in \mathrm{CNL}_{\mathcal{O}}$ ?

To fill in some terminology:

**Definition 1.13.** (a) For  $A \in \text{CNL}_{\mathcal{O}}$ , set

$$F_A := \text{Hom}_{\text{CNL}_{\mathcal{O}}}(A, \bullet) : \text{CNL}_{\mathcal{O}} \rightarrow \text{Set}.$$

(b) Say that a functor  $F : \text{CNL}_{\mathcal{O}} \rightarrow \text{Set}$  is

- (i) *representable* if there exists  $A \in \text{CNL}_{\mathcal{O}}$  such that  $F \cong F_A$ , and
- (ii) *continuous* if for all  $(A, \mathfrak{m}_A) \in \text{CNL}_{\mathcal{O}}$ , the map  $F(A) \rightarrow \lim_n F(A/\mathfrak{m}_A^n)$  is an isomorphism.

**Exercise 2.** (a) The fiber product of the diagram  $\Delta$  given by

$$\begin{array}{ccc} & B & \\ & \downarrow \psi & \\ A & \xrightarrow{\varphi} & C \end{array}$$

inside  $\text{Ar}_{\mathcal{O}}$  is

$$\{(a, b) \in A \times B \mid \varphi(a) \equiv \psi(b)\}.$$

(b)  $\text{CNL}_{\mathcal{O}}$  "has no fiber products".

**Exercise 3.** Suppose  $F = F_A$  for  $A \in \text{CNL}_{\mathcal{O}}$ . Then

- (a)  $F(\mathbb{F}) = \{*\}$ , and  $F$  is continuous.
- (b) The Mayer–Vietoris property (MV) holds for  $F$ , i.e., for all diagrams  $\Delta$ , the induced map  $(*)_{\Delta}$  is bijective, where

$$(*)_{\Delta} : F(A \times_C B) \rightarrow F(A) \times_{F(C)} F(B),$$

where on the right hand side, we take the fiber product in  $\text{Set}$ .

**Notation 1.14.** •  $\mathbb{F}[\varepsilon] := \mathbb{F}[X]/(X^2)$ .

- Call  $\varphi : A \rightarrow A'$  in  $\text{CNL}_{\mathcal{O}}$  *small* if  $\mathfrak{m}_{A'} \cdot (\ker(\varphi)) = 0$ . (e.g.,  $\mathbb{F}[\varepsilon] \rightarrow \mathbb{F}$ )
- For  $F : \text{CNL}_{\mathcal{O}} \rightarrow \text{Set}$ , define the  $T_F := F(\mathbb{F}[\varepsilon])$ , the *tangent space* of  $F$ .

You can translate this back into rings. Here are some more exercises:

**Exercise 4.** If the functor  $F$  satisfies

- $F(\mathbb{F}) = \{*\}$ , and
- $(*)_{T_F} : F(\mathbb{F}[\varepsilon] \times_{\mathbb{F}} \mathbb{F}[\varepsilon]) \rightarrow T_F \times T_F$  is bijective,

then  $T_F$  is an  $\mathbb{F}$ -vector space.

**Example 1.15.** For  $F = F_A$ , we have

$$T_F = \text{Hom}_{\text{CNL}_{\mathcal{O}}}(A, \mathbb{F}[\varepsilon]) = \text{Hom}_{\mathbb{F}}(\mathfrak{m}_A/(\mathfrak{m}_A^2, \pi), \mathbb{F}) =: T_A.$$

**Exercise 5.** For  $\varphi : A \rightarrow B$  in  $\text{CNL}_{\mathcal{O}}$ , the map  $\varphi$  is surjective if and only if  $T_{\varphi} : T_B \rightarrow T_A$  is injective.

**Theorem 1.16** (Grothendieck). *Suppose  $F : \text{CNL}_{\mathcal{O}} \rightarrow \text{Set}$  is continuous, satisfies  $F(\mathbb{F}) = \{*\}$ , the MV-property holds, and  $\dim_{\mathbb{F}} T_F < \infty$ . Then  $F$  is representable.*

**Remark 1.17.** For this last theorem, there are simplifications by Schlessinger, *Functors of Artin rings*. Schlessinger realized that these axioms of Grothendieck may be hard to verify in concrete situations, so he gave a simple (but somewhat long) list of axioms to verify. See Mazur '87, and Gouvêa's survey.

Let's now turn to *Galois deformations functors*. We have our usual

$$\bar{\rho} : G \rightarrow \mathrm{GL}_n(\mathbb{F}).$$

We can do something about the conjugation.

$$\Gamma_n(A) := \ker(\mathrm{GL}_n(A) \rightarrow \mathrm{GL}_n(\mathbb{F})).$$

We define the *lifting* and *deformation* functors

$$D_{\bar{\rho}}^{\square}, D_{\bar{\rho}} : \mathrm{CNL}_{\mathcal{O}} \rightarrow \mathrm{Set},$$

$$(A, \mathfrak{m}_A) \mapsto \{\rho_A : G \rightarrow \mathrm{GL}_n(A) \mid \rho_A \equiv \bar{\rho} \pmod{\mathfrak{m}_A}\}$$

(“lifting functor” or “framed deformation functor”), and where for  $D_{\bar{\rho}}$ , we take things modulo  $\Gamma_n(A)$ -conjugacy (i.e., “deformations are lifts modulo conjugacy”).

**Theorem 1.18.** *Suppose  $\Phi_{\rho}$  holds, i.e.,  $\#\mathrm{Hom}(G, \mathbb{F}_p) < \infty$ . Then:*

- (a)  $D_{\bar{\rho}}$  always has a “hull”.
- (b) If  $\mathrm{End}_G(\bar{\rho}) = \mathbb{F}$ , then  $D_{\bar{\rho}}$  is representable (uses Schlessinger). This gives rise to  $R_{\bar{\rho}}$ .
- (c) (Always)  $D_{\bar{\rho}}^{\square}$  is representable (Kisin, Magid–Lubotzki). Gives rise to  $R_{\bar{\rho}}^{\square}$ .

Here (c) is an exercise you can do.

**Example 1.19.** Take  $n = 1$ . Let  $\bar{\rho} : G \rightarrow \mathrm{GL}_1(\mathbb{F}) = \mathbb{F}^{\times}$ . **Fact** (Teichmüller lift): there exists  $\hat{\rho} : G \rightarrow \mathrm{GL}_1(W(\mathbb{F}))$ , a lift of  $\bar{\rho}$  such that  $\hat{\rho}(G) \xrightarrow{\mathrm{red}} \bar{\rho}(G)$ .

**Exercise 6.** Suppose that  $\mathrm{Hom}_{\mathrm{cts}}(G^{\mathrm{ab}}, \mathbb{F})$  is finite. Then the universal deformation ring  $R_{\bar{\rho}}$  (which for  $n = 1$  is the same as the universal lifting ring  $R_{\bar{\rho}}^{\square}$ ) is isomorphic to

$$\mathcal{O}[(G^{\mathrm{ab}})_p] = \varprojlim \mathcal{O}[(G^{\mathrm{ab}})_p].$$

The universal lifting is

$$\hat{\rho} \otimes (G \rightarrow (G^{\mathrm{ab}})_p \hookrightarrow \mathcal{O}[(G^{\mathrm{ab}})_p]^{\times}).$$

Special cases:

- (i)  $G = G_F$ ,  $F \supseteq \mathbb{Q}_p$  finite extension, then by local class field theory,

$$(G^{\mathrm{ab}})_p \cong \mu_{p^{\infty}}(F) \times \mathbb{Z}^{1+[F:\mathbb{Q}_p]}.$$

- (ii)  $G = G_{K,S}$ ,  $S \supseteq S_p$  finite, then global class field theory gives

$$(G^{\mathrm{ab}})_p \cong \frac{A_{K,S}}{\text{finite } p\text{-group}} \times \mathbb{Z}_p^{r_2+\delta},$$

where  $r_2$  denotes the number of complex places and  $\delta = \delta_{K,p}$  is the *Leopoldt defect*, which the Leopoldt conjecture says is zero.

We turn to cohomology. For  $\bar{\rho} : G \rightarrow \mathrm{GL}_n(\mathbb{F})$ , define

$$\mathrm{ad}_{\bar{\rho}} = M_{n \times n}(\mathbb{F})$$

with the “adjoint action”: for  $g \in A$  and  $\alpha \in M_{n \times n}$ , we take

$$g \cdot \alpha := \bar{\rho}(g) \cdot \alpha \cdot \bar{\rho}(g)^{-1}.$$

(We remark that  $\mathrm{ad}_{\bar{\rho}} \cong \bar{\rho} \otimes \bar{\rho}^{\vee}$ .)

**Proposition 1.20.** (a) We have  $T_{D_{\bar{\rho}}^{\square}} = D_{\bar{\rho}}^{\square}(\mathbb{F}[\varepsilon]) \xleftarrow{\cong} Z^1(G, \text{ad}_{\bar{\rho}})$ , where the map sends a cocycle  $c : G \rightarrow \text{ad}_{\bar{\rho}}$  to

$$\rho_c : g \mapsto (1 + \varepsilon c(g))\bar{\rho}(g).$$

(b) The above map induces an isomorphism  $H^1(G, \text{ad}_{\bar{\rho}}) \xrightarrow{\cong} T_{D_{\bar{\rho}}}$ .

(c) We have  $\dim B^1(G, \text{ad}_{\bar{\rho}}) = h^2 - h^0(G, \text{ad}_{\bar{\rho}})$ . (We use the notation  $h^i := \dim H^i$  and  $z^i = \dim Z^i$ .)

**Corollary 1.21.** For  $r = h^1(G, \text{ad}_{\bar{\rho}})$ ,  $r^{\square} = z^1(G, \text{ad}_{\bar{\rho}})$ , we have surjections

$$\mathcal{O}[[X_1, \dots, X_r]] \rightarrow R_{\bar{\rho}}$$

and

$$\mathcal{O}[[X_1, \dots, X_{r^{\square}}]] \rightarrow R_{\bar{\rho}}^{\square},$$

and  $r, r^{\square}$  are optimal.

What's the relevance of  $H^2$ ? Let  $\varphi : A^1 \rightarrow A$  be a small extension, and let  $I := \ker(\varphi)$ .

**Question 1.22.** Is the map  $D_{\bar{\rho}}(A') \rightarrow D_{\bar{\rho}}(A)$  surjective?

Let  $\rho_A : G \rightarrow \text{GL}_n(A)$  be as in the right hand side.

**Fact 1.23.** There exists a set-theoretic lift  $\rho' : G \rightarrow \text{GL}_n(A')$  of  $\rho_A$  (i.e.,  $\varphi \circ \rho' = \rho_A$ ).

**Definition 1.24.** Set

$$\begin{aligned} c_{\rho'} : G \times G &\rightarrow \text{ad}_{\bar{\rho}} \otimes_{\mathbb{F}} I, \\ (g, h) &\mapsto (\rho'(g, h) \cdot \rho'(h)^{-1} \cdot \rho'(g)^{-1}), \end{aligned}$$

where this last quantity lies in  $1 + M_{n \times n}(I)$ . We note that  $(\text{ad}_{\bar{\rho}} \otimes_{\mathbb{F}} I) \cong M_{n \times n}(I)$ . To get back, subtract 1.

**Proposition 1.25.** (a)  $c_{\rho'} \in Z^2(G, \text{ad}_{\bar{\rho}} \otimes I)$ ,  $\text{Obs}(\rho_A, \varphi) = [c_{\rho'}] \in H^2(G, \text{ad}_{\bar{\rho}} \otimes I)$  is independent of  $\rho'$ !  
(b) We have  $\text{Obs}(\rho_A, \varphi) = 0$  if and only if there exists  $\rho_{A'} \in D_{\bar{\rho}}(A')$  lifting  $\rho$ .

We give a proof of (b). If  $c_{\rho'} = \partial^2 f$  for some  $f : G \rightarrow \text{ad}_{\bar{\rho}} \otimes I$ , then you can check that if you take your  $\rho'$  and multiply it on the left with  $1 + f$ , we get an element  $(1 + f)\rho' : G \rightarrow \text{GL}_n(A')$  that is a lift of  $\rho_A$  (to  $D_{\bar{\rho}}^{\square}(A')$ ).

Mazur observed that you can do something with this obstruction theory: you can prove something more about these projections on the top line here. But you have a very coarse control over the kernel ideal  $I$  of the surjection.

**Theorem 1.26.** Consider  $\mathcal{R} = \mathcal{O}[[X_1, \dots, X_r]]$  projecting via something called  $\pi$  to  $R_{\bar{\rho}}$  ( $r = h^1(G, \text{ad}_{\bar{\rho}})$ ), with kernel  $J$ . Then

$$\mathcal{R}/\mathfrak{m}_{\mathcal{R}}J \rightarrow R_{\bar{\rho}}$$

is small with kernel  $J/\mathfrak{m}_{\mathcal{R}}J$ , and

$$\begin{aligned} \text{ob} : \text{Hom}(J/\mathfrak{m}_{\mathcal{R}}J, \mathbb{F}) &\cong H^2(G, \text{ad}_{\bar{\rho}}), \\ \psi &\mapsto (1 \otimes \psi) \text{obs}(\rho_{\bar{\rho}}^{\text{univ}}, \pi) \end{aligned}$$

is injective!

**Corollary 1.27.** With  $r, r^{\square}$  from above and  $s = h^2(G, \text{ad}_{\bar{\rho}})$  have presentations  $R_{\bar{\rho}} := \mathcal{O}[[x_1, \dots, x_{r^{\square}}]]/(f_1, \dots, f_s)$ .



**Definition 1.28** (ad hoc). The *expected cohomological dimension* is

$$\mathrm{ecd}\dim(\bar{\rho}) = z^1(G, \mathrm{ad}_{\bar{\rho}}) - h^2(G, \mathrm{ad}_{\bar{\rho}}) = h^2 - \sum_{i=0}^2 (-1)^i h^i(G, \mathrm{ad}_{\bar{\rho}}).$$

**Proposition 1.29.** *If*

$$\dim_s R_{\bar{\rho}}^{\square} / \pi \leq \mathrm{ecd}\dim(\bar{\rho}), \quad (1.3)$$

*then  $R_{\bar{\rho}}^{\square}$  is 0-flat, of relative dimension  $\mathrm{ecd}\dim(\bar{\rho})$  over  $\mathcal{O}$ , and a complete intersection.*

**Example 1.30.** (a)  $G = G_F$ , where  $F \supseteq \mathbb{Q}_{\ell}$  is finite and  $\ell \neq p$ . Then (1.3) holds.  $\mathrm{ecd}\dim = 4^2$ .

(b)  $G = G_F$ ,  $F \supseteq \mathbb{Q}_p$ . Then (1.3) holds and each  $\mathrm{ecd}\dim = n^2(1 + [F : \mathbb{Q}_p])$ .

(c) If  $G = G_{K,S}$ , then we have many examples for (1.3) if  $k$  is not totally real,  $k$  is totally odd, and  $\bar{\rho}$  is totally odd, and  $n \geq 2$ . (Have to be careful. Use that  $S \supseteq S_{\infty} \cup S_{p\text{sup}} \dots$ )

(d) We have

$$\dim_{\mathrm{Krull}} R_{\bar{\rho}}^n / \pi \leq \mathrm{ecd}\dim(\bar{\rho})$$

for  $G = G_{K,S}$ , with  $G$  and  $G_{K,S}$ , if and only if  $\delta = 0$ .

**Remark 1.31.** To avoid the reliance on  $\delta = 0$  (often) by fixing a lift  $\mu : G \rightarrow \mathcal{O}^{\times}$ .

*Mazur's vision:* for  $v \in S$ , define “geometric” subfunctors  $D_v^{\square} \subseteq D_{\bar{\rho}/G_{K,v}}^{\square}$  and the pullback in  $(\underline{S} = (\bar{\rho}, S, (D_v^0)_{v \in S}))$  to get a diagram involving, for  $G = G_{k,S}$ ,

$$D_{\bar{\rho}}^{\square} \rightarrow \sqcup_{v \in S} D_{\bar{\rho}|G_v}^{\square}$$

and

$$D_{\underline{S}}^{\square} \rightarrow \sqcup_{v \in S} D_v^{\square}.$$

Removing frames at right point, we get a

$$R_{\mathcal{I}} \leftarrow R_{\bar{\rho}}.$$

**Question 1.32.** Can we choose  $\mathcal{I}$  in such a way that some natural map

$$R_{\rho} \rightarrow \mathbb{T}_k(N, \varepsilon, \mathcal{O})_{\mathfrak{m}_{\bar{\rho}}}$$

is an isomorphism?

We turn now to *local deformation problems*. How to get a more geometric side of deformation rings? Fix

$$\bar{\rho} : G_{K,S} \rightarrow \mathrm{GL}_n(\mathbb{F}).$$

Write  $\bar{\rho}_v := \bar{\rho}|_{G_v}$ . Fix a character  $\mu : G_{K,S} \rightarrow \mathcal{O}^{\times}$  lifting  $\det \bar{\rho}$ . Our next aim is to identify interesting special loci in

$$X_{\bar{\rho}_v}^{\square} := \mathrm{Spec}(R_{\bar{\rho}_v}^{\square}).$$

**Definition 1.33** (Thorne, who gave a compact description). A *local deformation problem* is a subfunctor  $D_v \subseteq D_{\bar{\rho}_v}^{\square, \mu_v}$  (possibly with fixed determinant, i.e.,  $\det(\rho) = \mu$  – that’s what the superscripted  $\mu_v$  means) such that

- (1)  $D_v \cong F_{R_v}$  for a quotient  $R_v$  of  $R_{\bar{\rho}_v}^{\square, \mu}$ .
- (2) For all  $A \in \mathrm{CNL}_{\mathcal{O}}$  and  $g \in \Gamma_n(A)$ ,

$$\rho \in D_v(A) \implies g\rho g^{-1} \in D_v(A).$$

**Remark 1.34.**  $\Gamma_n(R_{\bar{\rho}_v}^\square)$  acts by conjugation on  $D_{\bar{\rho}_v}^\square(R_{\bar{\rho}_v}^\square)$ , and hence (by universality of  $R_{\bar{\rho}_v}^\square$ ), it acts on  $R_{\bar{\rho}_v}^\square$ .

**Proposition 1.35** (BLGHT, Potential Automorphy 2, Lemma 3.2). *Let  $R_v$  be a quotient  $R_{\bar{\rho}_v}^{\square, \mu} \xrightarrow{\alpha_v} R_v$  such that*

- (a)  $\ker(\alpha_v)$  is invariant under  $\Gamma_n(R_{\bar{\rho}_v}^\square)$ , and
- (b)  $R_v$  is reduced and not isomorphic to  $\mathbb{F}$ .

*Then  $F_{R_v} \subseteq D_{\bar{\rho}_v}^\square$  is a local deformation problem.*

**Construction of the  $R'_v$ s? (Kisin):** Let  $X_v \subseteq X_{\bar{\rho}_v}^\square(\bar{\mathbb{Q}}_p)$  be  $\Gamma_n(\bar{\mathbb{Q}}_p)$ -invariant. Each  $x \in X_{\bar{\rho}_v}^\square(\bar{\mathbb{Q}}_p)$  gives a map

$$\rho_x : G_v \rightarrow \mathrm{GL}_n(\bar{\mathbb{Q}}_p),$$

and also

$$\alpha_x : R_{\bar{\rho}_v}^\square \rightarrow \bar{\mathbb{Q}}_p.$$

Then

$$J := \bigcap_{x \in X_v} \ker(\alpha_x)$$

is  $\Gamma_n(R_{\bar{\rho}_v}^\square)$ -invariant. (Need to check whether condition (b) above holds to get the local deformation problem, but at least there is a good starting point for something.) You can define some

$$R_{X_v} := R_{\bar{\rho}_v}^\square / I.$$

Then  $\mathrm{Spec} R_{X_v}$  “is” the schematic closure of  $X_v$  in  $X_{\bar{\rho}_v}^\square$ .

For a classification of the  $\rho : G_v \rightarrow \mathrm{GL}(V)$  with  $V \simeq \bar{\mathbb{Q}}_p^n$ , with  $v \neq \infty$  and  $\ell = \mathrm{char}(v) \neq p$ , use the **Grothendieck monodromy theorem**. Choose a surjection

$$t_p : I_v \rightarrow I_v^{\mathrm{tame}} \simeq \hat{\mathbb{Z}}^{(\ell)} \xrightarrow{\mathrm{surj}} \mathbb{Z}_p.$$

- (a) There is a unique  $N = N_\rho \in \mathrm{End}_{\bar{\mathbb{Q}}_p}(V)$  and an open subgroup  $I \subseteq I_v$  so that

$$\rho|_I = \exp(t_p(\bullet)N)|_I.$$

- (b) For  $N$  from (a),

$$r = r_\rho : I_v \rightarrow \mathrm{GL}(V),$$

$$g \mapsto \rho(g) \exp(-t_p(\bullet)N)$$

is a continuous representation with finite image. Call

$$\tau = (r, N) = \mathrm{WD}(\rho)$$

the (inertial) Galois type of  $\rho$ .

Let

$$X_\tau := \left\{ x \in X_{\bar{\rho}_v}^\square(\bar{\mathbb{Q}}_p) \mid \mathrm{WD}(\rho) \simeq \tau \right\}.$$

Note that

$$\# \left\{ \tau \mid \tau \text{ arises from } X_{\bar{\rho}_v}^\square(\bar{\mathbb{Q}}_p) \right\}_{/\cong} < \infty.$$

Define

$$R_{\bar{\rho}_v}^{\square, \tau} := R_{X_\tau} \leftarrow R_{\bar{\rho}_v}^\square.$$

**Theorem 1.36** (Sholton, Proposition 3.6). (a)  $R_{\bar{\rho}_v}^{\square, \tau}$  defines a local deformation problem at  $v$ .

- (b)  $\mathrm{Spec}(R_{\bar{\rho}_v}^{\square, \tau})$  is a union of irreducible components of  $X_{\bar{\rho}_v}^\square$ , and equidimensional ( $n^2$ ).

(c)  $R_{\bar{\rho}_v}^{\square, \tau}[\frac{1}{p}]$  is reduced, contains a dense set of (generic) smooth points.

**Warning** (depending upon how much you are interested in singularities of schemes, or not):  $R_{\bar{\rho}_v}^{\square}$  is a complete intersection, but the components  $R_{\bar{\rho}_v}^{\square, \tau}$  can apparently have much worse singularities – they need not even be Cohen–Macaulay.

Let’s say a bit about the case  $v \mid p$ , where things get a bit vaguer. Now take  $K_v \supseteq \mathbb{Q}_p$  finite, and consider the set of all representations

$$\{\rho : G_v \rightarrow \mathrm{GL}_n(\bar{\mathbb{Q}}_p)\}. \quad (1.4)$$

By  $p$ -adic Hodge theory, this contains the following set of representations that are “geometric” in some sense:

$$\{\rho : G_v \rightarrow \mathrm{GL}_n(\bar{\mathbb{Q}}_p) \mid \rho \text{ is potentially semistable}\}. \quad (1.5)$$

Fontaine attaches invariants to these  $\rho$ ’s. One thing he attaches is a Galois type

$$\mathrm{WD}(\rho) = (r_\rho, N_\rho),$$

although in a very different way than before: in this case, where  $K_v$  is a  $p$ -adic field, the pro- $p$ -part of the inertia is not this harmless  $\mathbb{Z}_p$ , but is instead a huge group that can be complicated, so one really needs to use  $p$ -adic Hodge theory to get something here. Here, as before,  $N_\rho$  is a nilpotent operator and  $r_\rho : I_v \rightarrow \mathrm{GL}_n(\bar{\mathbb{Q}}_p)$  is a finite-dimensional representation. Fontaine also attaches a set of Hodge–Tate weights

$$\nu = \underline{\mathrm{HT}}(\rho) = \{\mathrm{HT}_\iota\}_{\iota : K_v \rightarrow \bar{\mathbb{Q}}_p}.$$

We attach some

$$X_{\tau, \nu} = \{\rho : G_v \rightarrow \mathrm{GL}_n(\bar{\mathbb{Q}}_p) \text{ potentially semistable} \mid \mathrm{WD}(\rho) \simeq \tau, \underline{\mathrm{HT}}(\rho) = \nu\}.$$

Kisin then attaches potentially semistable deformation rings  $R_{\bar{\rho}_v}^{\square, \tau, \nu}$ .

**Remark 1.37.** If  $n = 2$  and  $K = \mathbb{Q}$ , and if you somehow have HT-weights  $\nu = (k = 1, 0)$ ,  $k$  “small”,  $(k \geq 2)$ ,  $\tau = (\mathrm{triv}, 0)$  (giving rise to  $\rho$  crystalline), then

$$R_{\bar{\rho}_p}^{\square, \nu, \tau} \simeq \mathcal{O}[[X_1, \dots, X_5]].$$

(Alternative: use ordinary deformation rings.)

**Global deformation rings.** Fix  $f \in S_k(N, \varepsilon, \mathbb{F})$ , take  $E$  sufficiently large. We get

$$\bar{\rho} : G_{\mathbb{Q}, Np\infty} \rightarrow \mathrm{GL}_2(\mathbb{F}).$$

We found

$$\mathfrak{m}_{\bar{\rho}} \leq \mathbb{T}_{\mathfrak{o}} \rightsquigarrow \rho_{\bar{\rho}}^{\mathrm{mod}} : G_{\mathbb{Q}, Np\infty} \rightarrow \mathrm{GL}_2((\mathbb{T}_{\mathcal{O}})_{\mathfrak{m}_{\bar{\rho}}}).$$

Now, for the ramified primes  $\ell \mid Np$ , local Langlands gives a list  $T_\ell$  of Galois types that arise from (forms parametrized by)  $(\mathbb{T}'_{\mathcal{O}})_{\mathfrak{m}_{\bar{\rho}}}$ . At  $p$ , you also get HT-weights  $(k - 1, 0)$  (papers of Faltings and of Saito). From this data, you now get a local deformation problem

$$D_\ell = D_{\bar{\rho}_\ell}^{\square, T_\ell} \subseteq D_{\bar{\rho}_\ell}^{\square} \quad \text{and} \quad D_{\bar{\rho}_p}^{\square, T_p, \nu}.$$

You form the pullback functor for a list

$$\underline{S} = (\bar{\rho}, N, S, (D_\ell)_{\ell \mid Np}),$$

where  $S$  is the divisor of  $Np\infty$ . (Let's suppose  $p > 2$ .) Get a restriction from global to local:

$$\begin{array}{ccc}
 D_{\bar{\rho}}^{\square, \mu} & \xrightarrow{\text{res}} & \prod_{\ell|Np} D_{\bar{\rho}_{\ell}}^{\square, \mu} \\
 \uparrow & & \uparrow \\
 D_{\bar{\rho}}^{\square, \mu} & \xrightarrow{\text{res}} & \prod_{\ell|Np} D_{\bar{\rho}_{\ell}}^{\square, \mu} \\
 \uparrow & & \uparrow \\
 D_{\underline{S}} & \longrightarrow & \prod_{\ell|Np} D_{\ell}
 \end{array}$$

which gives rise to  $R_{\underline{S}}$  as a quotient of  $R_{\bar{\rho}}^{\square, \mu}$ , with

$$\rho_{\underline{S}}^{\text{univ}} : G_{\mathbb{Q}, Np\infty} \rightarrow \text{GL}_2(R_{\underline{S}}).$$

Universality of  $R_{\underline{S}}$  (using local Langlands and  $p$ -adic Hodge theory) gives a surjection

$$\alpha : R_{\underline{S}} \rightarrow (\mathbb{T}_{\mathcal{O}})_{\mathfrak{m}_{\bar{\rho}}}, \quad (1.6)$$

assuming that  $\bar{\rho}$  is absolutely irreducible.

**Conjecture 1.38.**  $\alpha$  is an isomorphism under suitable hypothesis.

The hope is that one can find a Galois-theoretic representation that completely describes this thing coming from automorphic representations. This was all first put into some definite form by Wiles and Taylor–Wiles in their proof of Fermat's last theorem. Their work gave the first great supply of such isomorphisms. Then there was Breuil–Conrad–Diamond–Taylor, proving Taniyama–Shimura–Weil conjecture. Then there was Clozel–Harris–Taylor, and Khare–Winterberger proving the Serre conjecture. One thing about why this thing could be important, and why it was important in many of these cases. What helps you sometimes is that being modular is “contagious”. What is often relatively easy, maybe if the representation (a  $p$ -adic Galois representation) comes from the ring on the left hand side of (1.6), and now if you have the full isomorphism (1.6), then you know it also comes from the right hand side. But in practice, this isomorphism doesn't tell you modularity of one just form in the class giving rise to the right hand side, but all of them. That's all.

## 2. CHRIS SKINNER'S LECTURES

### Integral representations, Euler systems, and multiplicity one.

My choice of these topics is motivated by my interest in special values of  $L$ -functions, and in particular problems like the BSD conjecture. We'll focus on some representation theory, that plays a role in both the analytic and the algebraic sides of these problems. You can possibly view this as a bridge between the talks at the start and at the end of the week.

Let's start by talking about *integral representations*. It's helpful to think

$$L\text{-function} = \int_{\text{symmetric space } X} (\text{automorphic form}),$$

where perhaps the automorphic form starts on some larger symmetric space  $Y \supseteq X$ . This is useful because it's our main tool for studying  $L$ -functions.

The next part of my title is *Euler systems*. This is going to seem like something different. What are Euler systems? One starts off with a continuous action

$$G_k = \text{Gal}(\bar{k}/k) \circ V,$$

where  $V$  is a  $\mathbb{Q}_p$ -space of finite dimension (with  $\mathbb{Q}_p$  acting linearly and continuously). At least conjecturally, there's a fairly general framework for producing such  $V$  from automorphic forms or representations. This Galois representation captures something about the automorphic form that can be expressed in terms of the  $L$ -function. All of these things are thus related to one another, even if they are frequently encountered separately. Here  $V$  often stabilizes in a  $\mathbb{Z}_p$ -submodule (lattice), which might yield a good exercise for later. An Euler system is a collection of classes in Galois cohomology  $c_F \in H^1(F, T)$ , where  $F/k$  are certain abelian extensions of  $k$  satisfying certain compatibilities: for  $F' \supseteq F$ ,

$$\text{cores}_{F'/F}(c_{F'}) = ? c_F,$$

where  $?$  often seems the local Euler factors of  $V$  (or some  $L$ -function attached to  $V$ , depending upon the setting).

Both of these settings have been useful for exploring special values of  $L$ -functions (Kolyvagin, Gross–Zagier, ...). What we'll focus on in these lectures is the role that multiplicity one plays in seeing these  $L$ -functions and in producing these Euler systems. We'll see that they essentially play the same role, which is further evidence for what people say, to the effect that Euler systems are some sort of algebraic incarnation of  $L$ -functions.

What do we mean by “multiplicity one”? One frequently encounters this term in the theory of automorphic forms, in various guises:

- (1) Uniqueness of a representation in some space of functions, e.g.:
  - (a) A cuspidal automorphic representation of  $\text{GL}_2$  shows up with multiplicity one  $L^2(\text{GL}_2(k) \backslash \text{GL}_2(\mathbb{A}_k))$ .
  - (b) Uniqueness of (local) Whittaker models for  $\text{GL}_2$ .
- (2) Uniqueness of some (invariant) linear functional: for  $H \leq G$  and  $\pi$  a representation of  $G$ ,

$$\dim \text{Hom}_H(\pi, \mathbb{C}) \leq 1.$$

Or, for  $\sigma$  a representation of  $H$ , as the assertion that  $\dim \text{Hom}_H(\pi, \sigma) \leq 1$ . The first examples can be understood in terms of the latter. The latter will be a useful framework for us.

Let's now turn to integral representations and give some examples. The first integral representation we see is that of the Riemann zeta function. Let

$$\psi(t) = \sum_{n=1}^{\infty} e^{-\pi n^2 t}. \quad (2.1)$$

Then for  $\Re s$  sufficiently large,

$$\int_0^{\infty} \psi(t) t^{\frac{1}{2}s-1} dt = \pi^{-s/2} \Gamma(\frac{s}{2}) \zeta(s). \quad (2.2)$$

We see this by bringing the summation outside the integral. This gives a Mellin transform.

What's the automorphic side of this? If we look at, for  $\tau = x + iy$ ,

$$\theta(\tau) = \sum_{n \in \mathbb{Z}} e^{-2\pi i n^2 \tau}.$$

This is an automorphic form, and we have

$$(\frac{1}{2}(\theta(iy) - 1)) = \psi(2y),$$

so (2.2) is an integral representation for the Riemann zeta function coming from the symmetric space for a torus embedded inside  $\mathrm{GL}_2$ . One has similar integral representations for the Dirichlet  $L$ -functions. (No multiplicity one that we can see thus far.)

This gets souped up in the work of Hecke and Iwasawa–Tate, which inspired how automorphic  $L$ -functions have been studied subsequently. Let’s recall how that goes. Let  $k$  be a number field. We have the adeles  $\mathbb{A}_k$  and the ideles  $\mathbb{A}_k^\times$ . We have a Hecke character

$$\chi : k^\times \backslash \mathbb{A}_k^\times \rightarrow \mathbb{C}^\times,$$

which factors as a product  $\chi = \prod \chi_v$  of characters  $\chi_v : k_v^\times \rightarrow \mathbb{C}^\times$  indexed by the places  $v$  of  $k$ . (This is of course very useful, but is specific for  $\mathrm{GL}_1$ , and so obscures some of the more general features.) Let  $\phi \in \mathcal{S}(\mathbb{A})$  be a Schwartz function, which could also be a product  $\phi = \prod \phi_v$  of local Schwartz functions  $\phi_v \in \mathcal{S}(k_v)$ . We recall that this means that

- when  $v$  is finite,  $\phi_v$  is smooth and compactly-supported, and
- when  $v$  is archimedean, all derivatives decay faster than any polynomial, e.g.,  $e^{-\pi t^2}$ .

Furthermore,  $\phi_v = 1_{\mathcal{O}_{k_v}}$  for almost all finite  $v$ . We then form

$$\theta(x) = \sum_{\alpha \in k} \phi(\alpha x).$$

(It’s a good exercise to see how to specialize this to obtain something like (2.1).) We then form the integral

$$\int_{k^\times \backslash \mathbb{A}_k^\times} \chi(x) |x|^s \theta(x) d^\times x.$$

These integrals converge absolutely for  $\Re s$  sufficiently large and unfold in the usual way, giving (at least for  $\chi$  not a power of the absolute value, so that we don’t need to worry about the contribution of  $\alpha = 0$ )

$$\int_{\mathbb{A}_k^\times} \chi(x) |x|^s \phi(x) dx.$$

If  $\phi = \prod \phi_v$ , then these factor further as

$$\prod \int_{k_v^\times} \chi_v(x) |x|_v^s \phi_v(x) dx. \quad (2.3)$$

One can show that the local integrals at non-archimedean places are rational functions, form the greatest common divisor of their denominators, and this turns out to be the way you can define the local  $L$ -function. This is Tate’s thesis. We haven’t yet really made any reference to multiplicity one. This shows up when you try to generalize to other settings.

We may think of  $\mathbb{A}_k^\times$  as  $\mathrm{GL}_1(\mathbb{A}_k)$ . Let’s now consider  $\mathrm{GL}_n(\mathbb{A}_k)$ . We discuss Godement–Jacquet theory, which is a generalization of what Tate did to  $\mathrm{GL}_n$ . Let  $\pi$  be a cuspidal automorphic representation (by convention, irreducible). Abstractly, this is isomorphic to a restricted tensor product  $\otimes \pi_v$  of irreducible local representations  $\pi_v$  of  $\mathrm{GL}_n(k_v)$ . We can thus identify an element  $\varphi \in \pi$  with a sum of tensor products of vectors (although, unlike in the case of characters, it will not

pointwise be a product of local functions). Now, mimicking what was done before, we take a Schwartz function  $\phi \in \mathcal{S}(M_n(\mathbb{A}_k))$ , and form a theta function

$$\theta(x) = \sum_{\alpha \in M_n(k)} \phi(\alpha x).$$

We then form

$$\int_{\mathrm{GL}_n(k) \backslash \mathrm{GL}_n(\mathbb{A}_k)} \varphi(x) |\det(x)|^s \theta(x) d^\times x.$$

This unfolds to

$$\int_{\mathrm{GL}_n(\mathbb{A})} \varphi(x) |\det x|^s \phi(x) d^\times x.$$

But does it factor? Not obviously.

Let's now form

$$\theta(h, g) = \sum_{\alpha \in M_n(k)} \phi(h^{-1} \alpha g)$$

and consider

$$\int_{[\mathrm{GL}_n]} \varphi(g) |\det g|^s \theta(h, g) dg.$$

This is now automorphic as a function of  $h$ , so we can decompose it with respect to the automorphic spectrum. To compute the coefficients in that decomposition, we consider, for  $\tilde{\varphi}$  in the contragredient (or dual)  $\tilde{\pi}$  of  $\pi$ , the iterated integral

$$\int_{[\mathrm{GL}_n^1]} \left( \int_{[\mathrm{GL}_n]} \varphi(g) |\det g|^s \theta(h, g) dg \right) \tilde{\varphi}(h) dh,$$

where  $\mathrm{GL}_n^1$  means either that we mod out by the center or that we restrict to  $|\det| = 1$ . Then, reordering terms and unfolding, we obtain

$$\int_{\mathrm{GL}_n(\mathbb{A})} \phi(g) |\det g|^s \left( \int_{[\mathrm{GL}_n^1]} \tilde{\varphi}(h) \varphi(hg) dh \right) dg.$$

We can understand the parenthetical inner integral as

$$\langle \tilde{\varphi}, \pi(g)\varphi \rangle,$$

where

$$\langle \varphi_1, \varphi_2 \rangle = \int_{[\mathrm{GL}_n^1]} \varphi_1(h) \varphi_2(h) dh.$$

This pairing defines a  $G$ -invariant functional  $\langle, \rangle : \tilde{\pi} \times \pi \rightarrow \mathbb{C}$ , which is locally unique, hence factors as a product of local invariant functionals  $\langle, \rangle_v : \tilde{\pi}_v \times \pi_v \rightarrow \mathbb{C}$ , thus

$$\langle, \rangle = (*) \prod_v \langle, \rangle_v.$$

The leading constant  $(*)$  will depend upon our normalizations of the local and global integrals, and our normalization of the comparison between  $\pi$  and  $\otimes \pi_v$ .

This is the first example where multiplicity one shows up in what we've discussed. In the afternoon talk, we'll very quickly describe a few other automorphic  $L$ -function settings where we see multiplicity one, and then start to move to the Euler system side of things.

**Afternoon talk.**

Factoring the integral is only one step towards understanding the  $L$ -functions. The next thing one needs to do is to compute these local integrals. This of course seems more tractable than working globally, which is the point. For instance, in the Iwasawa–Tate setting, when all of the data is unramified (meaning  $\chi_v$  is an unramified character and  $\phi_v$  is the characteristic function of the ring of integers), then the local factor in (2.3) is easy to compute, and gives a local zeta function. This is more complicated in the Godement–Jacquet setting, but still doable. Then in the ramified situations, there is the question of how to choose good vectors so that one gets the  $L$ -function on the nose. This is useful for many of the applications that the speaker makes of these kinds of functions. In some settings it’s still much of an art and there are lots of interesting questions.

Let’s turn to Rankin–Selberg convolutions. We’ll begin classically, say with holomorphic modular eigenforms  $f$  and  $g$  of weights  $k_f \geq k_g$ , say of level 1, i.e., on  $\mathrm{SL}_2(\mathbb{Z})$ . Write

$$f = \sum a_n q^n, \quad g = \sum b_n q^n.$$

We’ll consider the Dirichlet series

$$\sum_n \overline{a_n} b_n n^{-s}. \quad (2.4)$$

The integral representation is

$$\int_{\mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}} \overline{f(\tau)} g(\tau) E_k(\tau, s) y^{k_f} d \mathrm{vol}(\tau), \quad (2.5)$$

where  $k := k_f - k_g$  and

$$E_k(\tau, s) := \sum_{\gamma \in \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \backslash \mathrm{SL}_2(\mathbb{Z})} j(\gamma, \tau)^{-k} |y(\gamma(\tau))|^k. \quad (2.6)$$

The way you get from (2.5) to (2.4) is to unfold the sum in (2.6) with the integral in (2.5), which yields an integral over  $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \backslash \mathbb{H}$ . This is a nice exercise if you haven’t done it.

What does this have to do with multiplicity one? Maybe we try to set this up a bit more automorphically. Let  $\pi_1$  and  $\pi_2$  be cuspidal automorphic representations of  $\mathrm{GL}_2(\mathbb{A}_k)$ . Let  $\chi_1, \chi_2$  be Hecke characters  $\chi_i : k^\times \backslash \mathbb{A}_k^\times \rightarrow \mathbb{C}$  such that  $\chi_1 \chi_2 = (\chi_{\pi_1} \chi_{\pi_2})^{-1}$ . Let  $I_s(\chi_1, \chi_2)$  denote the space of functions  $f : \mathrm{GL}_2(\mathbb{A}) \rightarrow \mathbb{C}$  that are smooth (meaning the usual thing at archimedean places and “fixed by an open subgroup” at finite places) and “ $K$ -finite” and satisfying

$$f_s \left( \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} g \right) = \chi_1(a) \chi_2(d) \left| \frac{a}{d} \right|^{s+\frac{1}{2}} f_s(g).$$

We then define an Eisenstein series by averaging:

$$E(f_s, g) := \sum_{\gamma \in B(\mathbb{Q}) \backslash \mathrm{GL}_2(\mathbb{Q})} f_s(\gamma g).$$

The Rankin–Selberg integral is now just the integral

$$\int_{[\mathrm{GL}_2]} \varphi_1(g) \varphi_2(g) E(f_s, g) dg.$$



We can unfold this to obtain

$$\int_{Z(\mathbb{A})B(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A})} \varphi_1(g)\varphi_2(g)f_s(g) dg.$$

Let's replace  $\varphi_i$  by its Whittaker expansion

$$\varphi_1(g) = \sum_{a \in k^\times} W_\psi \left( \begin{pmatrix} a & \\ & 1 \end{pmatrix} g \right).$$

Then, unfolding a bit further, we arrive at

$$\int_{Z(\mathbb{A})N(\mathbb{A})\backslash\mathrm{GL}_2(\mathbb{A})} W_\psi(g)\varphi_2(g)f_s(g) dg = \int_{Z(\mathbb{A})N(\mathbb{A})\backslash\mathrm{GL}_2(\mathbb{A})} W_\psi(g)W'_{\bar{\psi}}(g)f_s(g) dg,$$

where  $W'_{\bar{\psi}}$  is the Whittaker function for  $\varphi_2$  with respect to the conjugate character. This last integrand is a product of local functions, so the integral factors, assuming that all vectors in our representations are pure tensors.

Where's the multiplicity one? It's hidden, because, just like in the case of  $\mathrm{GL}_1$  characters, we have taken for one of our automorphic forms something particularly special, namely an Eisenstein series. The picture that might be better is, to give something slightly more complicated, there's something called the **triple product integral**. Now we'll take  $\varphi_i \in \pi_i$  for  $i = 1, 2, 3$ , where it'll be slightly simpler to assume that at least one is cuspidal. Look at the function

$$\int_{[Z\backslash\mathrm{GL}_2]} \varphi_1(g)\varphi_2(g)\varphi_3(g) dg.$$

This integral defines a trilinear form on the product of the three representations that is invariant by the diagonal action of  $\mathrm{GL}_2(\mathbb{A})$ , or equivalently, a  $\mathrm{GL}_2(\mathbb{A})$ -invariant functional  $\Lambda$  on the tensor product  $\pi_1 \otimes \pi_2 \otimes \pi_3$ , i.e., an element

$$\Lambda \in \mathrm{Hom}_{\mathrm{GL}_2(\mathbb{A})}(\pi_1 \otimes \pi_2 \otimes \pi_3, \mathbb{C}).$$

This space is one-dimensional, as are its local avatars:

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(k_v)}(\pi_{1,v} \otimes \pi_{2,v} \otimes \pi_{3,v}) \leq 1.$$

We can thus factor  $\Lambda = \prod \Lambda_v$ , where each  $\Lambda_v \in \mathrm{Hom}_{\mathrm{GL}_2(k_v)}(\pi_{1,v} \otimes \pi_{2,v} \otimes \pi_{3,v})$ . Suppose  $v$  is a place for which  $\pi_{i,v}$  is unramified for each  $i$ . Then we can take

$$\Lambda_v = \int_{ZN(k_v)\backslash\mathrm{GL}_2(k_v)} W_{\psi_v} W_{\bar{\psi}_v}(g) f_v(g) dg.$$

Here, locally, we're realizing the unramified representation  $\pi_{3,v}$  as an induced representation  $\pi_{3,v} = \pi_v(\chi_{1,v}, \chi_{2,v})$ . We didn't unfold to this computation; this was all local.

There's one more example we'd like to emphasize: *toric integrals*. Let  $K/k$  be a quadratic extension. We can then think of  $K$  as a two-dimensional  $k$ -space, which gives a way to identify

$$\mathrm{GL}_2/k \cong \mathrm{Aut}_k(K) \leftrightarrow K^\times.$$

We might for simplicity that  $\pi$  is a cuspidal automorphic representation of  $\mathrm{GL}_2(\mathbb{A}_k)$ , and let  $\varphi \in \pi$  be a cusp form. Let  $\chi : K^\times \backslash \mathbb{A}_K^\times \rightarrow \mathbb{C}^\times$  be a Hecke character for our extension. Let's assume that  $\chi|_{\mathbb{A}_k^\times} = \chi_\pi^{-1}$ , i.e., the restriction is the inverse of the central character. We are then going to think of the integral

$$\int_{\mathbb{A}_k^\times K^\times \backslash \mathbb{A}_K^\times} \varphi(t)\chi(t) d^\times t$$

as defining an element of the space  $\text{Hom}_{\mathbb{A}_k^\times}(\pi, \mathbb{C}(\chi^{-1}))$ , which has dimension  $\leq 1$ . Such integrals will thus factor as products of local linear functionals.

Some of these can be computed quickly in terms of integrals we already know. Suppose  $\pi_v$  is unramified and  $v$  splits in  $K/k$ . Then

$$K_v^\times = k_v^\times \times k_v^\times \cong \left\{ \begin{pmatrix} a & \\ & b \end{pmatrix} \right\} \subseteq \text{GL}_2(k).$$

We may thus identify  $\chi_v$  with a pair of characters  $(\chi_{1,v}, \chi_{2,v})$ . We can define the local functional

$$\int_{k_v^\times} W_\psi \begin{pmatrix} t & \\ & 1 \end{pmatrix} \chi_{1,v}(t) d^\times t,$$

which defines an element of  $\text{Hom}_{K_v^\times}(\pi_v, \mathbb{C}(\chi_v^{-1}))$ . For unramified data, this evaluates to  $L(\pi_v, \chi_{1,v})$ . It turns out that (Waldspurger's formula)

$$\left| \int \varphi(t) \chi(t) d^\times t \right|^2 \sim L(\text{BC}_{K/k}(\pi) \otimes \chi, \tfrac{1}{2}).$$

There's a similar relation in the triple product case, which we can guess using the Rankin–Selberg unfolding that we saw earlier.

**Remark 2.1.** This doesn't make sense for general arguments other than  $s = \frac{1}{2}$ , except in some form when  $\varphi$  is an Eisenstein series (“formula of Damerell”). One needs to be able to vary  $\varphi$  (or  $\chi$ ) in a family, preserving the central character compatibility condition.

All of these examples are special cases of the Gan–Gross–Prasad conjectures, which we'll hear more about later in the week.

We'll now begin by giving one example of an Euler system. Tomorrow, we'll explore this in greater generality and more detail. The simplest case is that of cyclotomic units. Let  $F$  be a number field or a local field. Kummer theory gives an isomorphism

$$F^\times / F^{\times N} \rightarrow H^1(F, \mu_N),$$

as follows. Let  $\alpha \in F^\times$ . Choose an  $n$ th root  $\alpha^{1/N}$ . The ambiguity in this choice is an element of  $\mu_N$ , i.e., an  $n$ th root of unity. For any  $\sigma \in G_F$ , we can look at  $\sigma(\alpha^{1/N})/\alpha^{1/N}$ . Since the numerator and denominator are both  $N$ th roots of 1, the ratio must lie in  $\mu_N$ . This gives us a way of constructing elements of  $H^1$  very concretely. The classes we obtain are unramified away from  $N\alpha$ .

Let's now restrict ourselves to the field  $\mathbb{Q}[\mu_N]$ . Restrict the above map to the unit group of that field:

$$\mathcal{O}_{\mathbb{Q}[\mu_N]}^\times \rightarrow H^1(\mathbb{Q}(\mu_N), \mu_{p^n}).$$

Taking inverse limits, we obtain, with  $\mathbb{Z}_p(1) = \varprojlim \mu_{p^n}$  (where the Galois group acts via the cyclotomic character), a map

$$\mathcal{O}_{\mathbb{Q}[\mu_N]}^\times \otimes \mathbb{Z}_p \rightarrow H^1(\mathbb{Q}(\mu_N), \mathbb{Z}_p(1)).$$

Thus, for any  $(a, N) = 1$ , we get, say with  $\zeta_N = e^{2\pi i/N}$ ,

$$\frac{\zeta_N^a - 1}{\zeta_N - 1} \mapsto z_N,$$

say. If we look at the corestriction

$$\text{cores}_{\mathbb{Q}[\mu_{N\ell}]/\mathbb{Q}[\mu_N]}(z_{N\ell}) = \begin{cases} z_N & \text{if } \ell \mid N \\ (1 - \text{Fr}_\ell^{-1})z_N & \text{if } \ell \nmid N. \end{cases}$$

We look at

$$\det(1 - \text{Fr}_\ell^{-1} X \mid \mathbb{Z}_p(1)^*) \mid X = \text{Fr}_\ell^{-1}.$$

Some idea for checking this:  $(\mathbb{Z}/\ell\mathbb{Z})^\times \subset (\mathbb{Z}/\ell N\mathbb{Z})^\times \cong \text{Gal}(\mathbb{Q}(\mu_{N\ell})/\mathbb{Q})$ , where the last map sends  $a \mapsto \sigma_a(\zeta_{N\ell}) = \zeta_{N\ell}^a$ .

**Remark 2.2.** Given  $f$  a holomorphic modular form, say of weight 2, we have  $\omega_f = f(\tau) d\tau \in H^0(X, \Omega_X^1) \subseteq H^1(X, \mathbb{C})$ . Then  $\|f\|^2 = \int_X \omega_f \wedge \overline{\omega}_f$ .

Similarly, for  $X \hookrightarrow X \times X \times X$ , we get some particular values of the Rankin–Selberg convolution by looking at  $\int_X \overline{\omega}_f \wedge \omega_g \wedge \omega_{E_k}$ . Then from the rational structure on cohomology, one can get something like rational structure on the  $L$ -values.

**Remark 2.3.** Let's talk about the corestriction map in this setting. Abstractly, it's a map

$$H^1(\mathbb{Q}[\mu_{N\ell}], \mathbb{Z}_p(1)) \rightarrow H^1(\mathbb{Q}[\mu_N], \mathbb{Z}_p(1)).$$

How are we going to understand this? We could of course write it down at the level of cocycles, or something. But what's it's really doing is that if we restrict back, i.e., compose with the restriction map

$$H^1(\mathbb{Q}[\mu_N], \mathbb{Z}_p(1)) \rightarrow H^1(\mathbb{Q}[\mu_{N\ell}], \mathbb{Z}_p(1)), \quad (2.7)$$

then the composition is just the trace map, given by

$$\sum_{\sigma \in \text{Gal}(\mathbb{Q}[\mu_{N\ell}]/\mathbb{Q}[\mu_N])} \sigma.$$

The restriction map (2.7) is actually an injection because there are no Galois invariants of  $\mathbb{Z}_p(1)$ .

Also, we have the Kummer map

$$\mathbb{Q}[\mu_{N\ell}]^\times \xrightarrow{\text{Kummer}} H^1(\mathbb{Q}[\mu_{N\ell}], \mathbb{Z}_p(1)).$$

And this construction is Galois-invariant. We have the trace map  $\alpha \mapsto \prod_\sigma \sigma(\alpha)$  from  $\mathbb{Q}[\mu_{N\ell}]^\times \rightarrow \mathbb{Q}[\mu_N]^\times$ . We claim that this induces the corestriction map via the Kummer map. This remains the case when we tensor with  $\mathbb{Z}_p$ .

So when we're dealing with corestriction, what we really want to understand is what is the norm of the particular  $\alpha$  that we're working with.

All of this is fairly formal applied to a specific setting. The Kummer map is also a connecting map in a long exact sequence of Galois cohomology, associated to

$$0 \rightarrow \mu_{p^n} \rightarrow \bar{F}^\times \xrightarrow{\alpha \mapsto \alpha^{p^n}} \bar{F}^\times \rightarrow 0.$$

From the formal stuff, we have an arithmetic question. We take

$$\alpha = \frac{\zeta_{N\ell}^a - 1}{\zeta_N - 1},$$

and we want to know, what is the norm of this element? Look at

$$\text{norm}_{\mathbb{Q}(\mu_{N\ell})/\mathbb{Q}(\mu_N)}(\alpha).$$

We have

$$\text{Gal}(\mathbb{Q}(\mu_{N\ell})/\mathbb{Q}(\mu_N)) \subseteq \text{Gal}(\mathbb{Q}(\mu_{N\ell})/\mathbb{Q}) \cong (\mathbb{Z}/N\ell)^\times,$$

The subgroup here will be identified with congruence classes  $b$  satisfying  $b \equiv 1 \pmod{N}$ . The last isomorphism is given by  $\sigma_c \mapsto c$ , where  $\sigma_c(\zeta_{N\ell}) = \zeta_{N\ell}^c$ . Now, suppose for instance that  $\ell \nmid N$ . Then  $\zeta_{N\ell}^{b\ell} = \zeta_N^b = \zeta_N$ . As  $b \in (\mathbb{Z}/N\ell)^\times$  runs over  $\ell - 1$  residue classes modulo  $\equiv 1 \pmod{N}$ , then  $\zeta_{N\ell}^b$  runs over  $\ell$ th roots of  $\zeta_N$ , but excluding  $\zeta_N^{\ell-1}$ . A short calculation then gives the claim corestriction formula. In the other case where  $\ell \mid N$ ,  $b$  runs over  $\ell$  classes, and we get the other answer.

We're going to continue our discussion of Euler systems. We turn to the example given by Heegner points. Let  $K/\mathbb{Q}$  be an imaginary quadratic field, with ring of integers  $\mathcal{O}_K$ . Let  $N$  be a positive integer such that all primes  $\ell \mid N$  split in  $K/\mathbb{Q}$ . We take

$$X_0(N) = \Gamma_0(N) \backslash [\mathfrak{h} \sqcup \mathbb{P}^1(\mathbb{Q})],$$

and write  $\tau \in \mathfrak{h}$ . This classifies elliptic curves together with an isogeny of order  $N$ , e.g.,

$$E := \mathbb{C}/\mathbb{Z} \oplus \mathbb{Z}\tau \rightarrow E' := \mathbb{C}/\frac{1}{N}(\mathbb{Z} + N\tau\mathbb{Z}) \cong \mathbb{C}/\mathbb{Z} + N\tau\mathbb{Z},$$

whose kernel is  $\frac{1}{N}\mathbb{Z}/\mathbb{Z}$ .

We're going to produce some points on  $X_0(N)$ . Let  $c$  be positive integer. We can have an order

$$\mathcal{O}_c := \mathbb{Z} + c\mathcal{O}_K \subseteq \mathcal{O}_K.$$

This gives us a lattice inside the complex numbers (having chosen a complex embedding of  $K$ ). We can then form the quotient  $\mathbb{C}/\mathcal{O}_c$ . To produce a lattice that is slightly larger, we will use that each  $\ell \mid N$  splits to *choose* an ideal  $\mathfrak{n} \subseteq \mathcal{O}_K$  such that  $\mathcal{O}_K/\mathfrak{n} \cong \mathbb{Z}/N\mathbb{Z}$ . We then obtain an isogeny of elliptic curves

$$[\mathbb{C}/\mathcal{O}_c \rightarrow \mathbb{C}/\mathfrak{n}^{-1}\mathcal{O}_c] \in X_0(N)(K[c]),$$

where  $K[c]$  denotes the *ring class field* of  $K$  of conductor  $c$ . To explain what this means, we introduce some notation. For a module  $M$ , write  $\hat{M} := M \otimes \hat{\mathbb{Z}}$ , where  $\hat{\mathbb{Z}} = \prod_\ell \mathbb{Z}_\ell$ . (For instance, if  $M$  is a  $\mathbb{Q}$ -module, then we may also write  $\hat{M}$  as  $M \otimes_{\mathbb{Q}} \hat{\mathbb{A}}_f$ .) Class field theory tells us that

$$\begin{aligned} \text{Pic}(\mathcal{O}_c) &= K^\times \backslash \hat{K}^\times / \hat{\mathcal{O}}_c^\times \xrightarrow{\text{rec}} \text{Gal}(K[c]/K), \\ [\mathfrak{b}] &\mapsto \sigma_{\mathfrak{b}}. \end{aligned}$$

More generally, for any fractional ideal  $\mathfrak{a}$  of  $\mathcal{O}_c$ , we get a point

$$x_c(\mathfrak{a}) := [\mathbb{C}/\mathfrak{a} \rightarrow \mathbb{C}/\mathfrak{n}^{-1}\mathfrak{a}] \in X_0(N)(K[c]).$$

The action of  $\text{Gal}(K[c]/K)$  on  $X_0(N)(K[c])$  on these points is described by the relation

$$\sigma_{\mathfrak{b}} x_c(\mathfrak{a}) = x_c(\mathfrak{b}^{-1}\mathfrak{a}). \quad (2.8)$$

This is the content of CM theory.

Now we're going to look at the divisors

$$y_c(\mathfrak{a}) := x_c(\mathfrak{a}) - \infty$$

obtained by subtracting off the cusp at infinity, which is defined over  $\mathbb{Q}$ . This difference defines an element of the Jacobian  $J_0(N) := \text{Jac}(X_0(N))$ , defined over  $K[c]$ . The Jacobian is an abelian variety. For abelian varieties, there is a construction very similar to Kummer theory that gives us an analogue of the Kummer map defined yesterday,

$$J_0(N)(K[c]) \rightarrow H^1(K[c], J_0(N)[p^m]),$$

where  $J_0(N)[p^m]$  is the  $p^m$ -torsion subgroup of the Jacobian. The map is defined as follows. Given  $y \in J_0(N)(K[c])$ , let us choose  $y'$  so that  $p^m y' = y$ . We send this to the class of the cocycle  $\sigma \mapsto \sigma(y') - y'$ . Applying this map to  $y_c(\mathfrak{a})$  gives us a cohomology class

$$z_c(\mathfrak{a}) \in H^1(K[c], J_0(N)[p^m]).$$

In the special case where  $\mathfrak{a} = \mathcal{O}_c$ , we drop it from the notation and write simply  $z_c$ .

Suppose now that  $\ell \nmid cD_K N$ . Then we're going to look at what happens when we take the point  $z_{c\ell}$  (attached as above to the order  $\mathcal{O}_{c\ell}$ ) and form the norm

$$\sum_{\sigma \in \text{Gal}(K[c\ell]/K[c])} \sigma z_{c\ell}.$$

We can write this as the sum

$$\sum_{\mathfrak{b} \in \ker(\text{Pic } \mathcal{O}_{c\ell} \rightarrow \text{Pic } \mathcal{O}_c)} \sigma_{\mathfrak{b}} z_{c\ell}.$$

We can in turn rewrite  $\sigma_{\mathfrak{b}} z_{c\ell}$  as  $\sum_{\mathfrak{b}} y_{c\ell}(\mathfrak{b}^{-1}) - \infty$ , summing over the same  $\mathfrak{b}$  as before. Now, how many  $\mathfrak{b}$ 's are there, and what is this sum?

Suppose that  $\ell$  is inert in  $K$ . Comparing what happens with  $c$  and  $c\ell$ , we get

$$\frac{(\mathcal{O}_K \otimes \mathbb{Z}_{\ell})^{\times}}{((\mathbb{Z} + \ell \mathcal{O}_K) \otimes \mathbb{Z}_{\ell})^{\times}},$$

which is cyclic of order  $\ell + 1$ . Now,  $\mathfrak{b}$  lying in the kernel of the above map means that  $\mathfrak{b}\mathcal{O}_c$  is a principal ideal, say  $\beta\mathcal{O}_c$  with  $\beta \in K^{\times}$ . If we look at  $\beta\mathfrak{b}^{-1}\mathcal{O}_c$ , then we see that inside  $\mathcal{O}_c$ , it has index  $\ell$ . As  $\mathfrak{b}$  varies, this exhausts all lattices of index  $\ell$ . By definition of the Hecke operator  $T_{\ell}$ , we see that

$$\sum_{\mathfrak{b}} (y_{c\ell}(\mathfrak{b}^{-1}) - \infty) = T_{\ell}(y_c - \infty).$$

Under the Kummer map, this tells us that

$$\text{cor}_{K[c\ell]/K[c]} z_{c\ell} = T_{\ell} z_c.$$

Kolyvagin used these relations as follows.

$$y_K = \text{trace}_{K[1]/K} y_1 \in J_0(N)(K) \xrightarrow{\phi_E} E(K),$$

where  $E$  is an elliptic curve of conductor  $N$  with a modular parametrization  $\phi_N : X_0(N) \rightarrow E$  sending  $\infty$  to 0. Using relations, Kolyvagin could show that if the point is not torsion, then the rank of the elliptic curve is 1. Around the same time, Gross–Zagier showed that the Néron–Tate height of this point is nonzero if and only if  $L$ -function for  $E/K$  vanishes exactly to order one. This gave some of the first theoretical evidence for the Birch and Swinnerton-Dyer conjecture. This was a spectacular application by Kolyvagin that got people interested in Euler systems.

You can also run Kolyvagin's argument in another way. You can look at the primes  $\ell$  that are *split* in  $K$ , say  $\ell = \lambda\bar{\lambda}$ . Then

$$\frac{(\mathcal{O}_K \otimes \mathbb{Z}_{\ell})^{\times}}{((\mathbb{Z} + \ell \mathcal{O}_K) \otimes \mathbb{Z}_{\ell})^{\times}} = \frac{\mathcal{O}_{\lambda}^{\times} \times \mathcal{O}_{\bar{\lambda}}^{\times}}{\{(a, b) : a \equiv b(\ell)\}},$$

where the numerator is really  $\mathbb{Z}_\ell^\times \times \mathbb{Z}_\ell^\times$ . Arguing as above, we're no longer summing over all the lattices, but instead we miss two of them:  $\lambda\mathcal{O}_c$  and  $\bar{\lambda}\mathcal{O}_c$ . One obtains

$$\sum_{\mathfrak{b}} (y_{c\ell}(\mathfrak{b}^{-1}) - \infty) = T_\ell(y_c - \infty) - (\mathrm{Fr}_\lambda^{-1} + \mathrm{Fr}_{\bar{\lambda}}^{-1})(y_c - \infty).$$

(The inverses come from (2.8).)

Work of the speaker and Jetchev and Wan used the split primes to do something like what Kolyvagin did.

Let  $\mathfrak{c} = \lambda_1 \cdots \lambda_r$  be a squarefree product of ideals  $\lambda_i \nmid ND_K$  of degree 1 in  $K$ , with  $\lambda_i = \bar{\lambda}_i$ . Let  $c$  be the integer such that  $(c) = \mathfrak{c} \cap \mathbb{Z}$ . Let  $z(\mathfrak{c})$  be basically the class that we were just analyzing, but let's modify it slightly:

$$z(\mathfrak{c}) := \prod_{i=1}^r (-\mathrm{Fr}_{\lambda_i}) z_c \in H^1(K[c], T_p J_0(N)), \quad (2.9)$$

where  $T_p J_0(N) = \varprojlim J_0(N)[p^m]$ . We obtain now, for  $\lambda \mid \ell$ ,

$$\mathrm{cores}_{K[c\ell]/K[c]} z(\mathfrak{c}\lambda) = (\mathrm{Fr}_\lambda^{-2} - T_\ell \mathrm{Fr}_\lambda^{-1} + 1) z(\mathfrak{c}). \quad (2.10)$$

Let's think about this last expression in terms of the Hecke polynomial  $X^2 - T_\ell X + \ell$ . You can see that if we evaluate this at  $\lambda^{-1}$ , we almost get the quantity appearing on the right hand side of (2.10):

$$(X^2 - T_\ell X + \ell) \big|_{X=\mathrm{Fr}_\lambda^{-1}}.$$

They are congruent modulo  $\ell - 1$ , which is  $N(\lambda) - 1$ . When working with Euler systems, it's acceptable to work with congruences modulo  $\ell - 1$ . There's a general way to massage those classes so that they give the relations on the nose, but there's no need to do so. Later, we'll pose a question that will suggest that this comparison is a feature of some integral representation theory, once we see where these norm relations come from in those terms. In any event, once you have these relations, you can take these objects and run Kolyvagin's argument and reprove Kolyvagin's theorem.

We next want to explain how to set up the construction of the Kolyvagin system in such a way that some representation theory naturally shows up, with this Galois relation (2.10) showing up as something like a Hecke module, which in turn is closely related to representation theory. The argument we're going to give generalizes quite significantly, for instance, it produces an Euler system in the sense of these split primes for the diagonal cycles coming from the arithmetic Gan–Gross–Prasad settings, and one can generalize Kolyvagin's statement to a rank one statement about Selmer groups for certain Rankin–Selberg convolutions of  $\mathrm{GL}_n \times \mathrm{GL}_{n+1}$ . With Euler systems, we're trying to bound the orders of Selmer groups of elliptic curves, something like  $H_f^1(\mathbb{Q}, E[p^N])$ , where the subscript  $f$  denotes some sort of Block–Kato condition coming from the geometry of the elliptic curve. One way to bound a Selmer group like this is to use global duality to give classes in the arithmetic dual, i.e.,  $H_f^1(\mathbb{Q}, E[p^N]^*)$ , where

$$E[p^N]^* = \mathrm{Hom}(E[p^N], \mathbb{Q}_p/\mathbb{Z}_p(1)). \quad (2.11)$$

Here one wants to allow ramification at good primes  $\ell$ , chosen to capture the orders of classes that we want to control. One can then reduce to looking at

$$H_{\mathrm{ur}}^1(\mathbb{Q}_\ell, E[p^N]). \quad (2.12)$$

The local duality relates this to  $H^1(I_\ell, E[p^N]^*)^{G_{Q_\ell}}$ . So if we can produce a class in the latter that is highly ramified (of large order), then we can force the class in (2.12) to have small order. That, in a nutshell, is how Selmer groups are bounded. The whole trick is to produce ramified classes in (2.11) that you can measure the size of in some way. Now, producing things that are provably ramified is a hard problem. There are few cases where we can do this. It's much easier to check that something is unramified (think of the criterion of Néron–Ogg–Shafarevich or something like that). What Kolyvagin's argument does is, he says well, let's start off with classes (2.9) that are over ramified extensions. If they are truly over that ramified extension and not, say, defined over some extension with less ramification, then Kolyvagin can use that to produce ramified classes, and it is exactly these relations (2.10) that allow him to understand how ramified these classes actually are. These norm relations are thus crucial for producing ramified classes with controlled or measurable ramification. That's sort of the algebra background for why one is interested in Euler systems.

We'll now describe a more representation-theoretic picture that produces these Kolyvagin classes. Recall the situation. We have the modular curve  $X_0(N)$  and the Heegner point  $x_c \in X_0(N)(K[c])$ , corresponding to an isogeny  $[\mathbb{C}/\mathcal{O}_c \rightarrow \mathbb{C}/\mathfrak{n}\mathcal{O}_c]$ . The modular curve is a curve defined over  $\mathbb{Q}$ , so we can think of this point as defining a map

$$\mathrm{Spec} K[c] \rightarrow X_0(N).$$

We can base our curve to a curve over  $K[c]$ , yielding the following map of varieties over  $K[c]$ :

$$\mathrm{Spec} K[c] \xrightarrow{x_c} X_0(N) \times \mathrm{Spec} K[c].$$

The left hand side has dimension 0, while the right hand side has dimension 1. There is thus a cycle class map

$$H^0(\mathrm{Spec} K[c]) \xrightarrow{\mathrm{cyc}} H^2(X_0(N)/\mathrm{Spec} K[c], \mathbb{Z}_p(1)).$$

(In general, the exponent is  $2d$ , where  $d$  is the codimension.)

This is a single point. What we did before was to take the difference between two points,

$$\mathrm{cyc}(x_c) - \mathrm{cyc}(\infty) \in H^2(X_0(N) \times \mathrm{Spec} K[c], \mathbb{Z}_p(1))_0, \quad (2.13)$$

but it's now homologically trivial in the sense that if we pass to the algebraic closure, i.e., we consider the degree map

$$H^2(X_0(N) \times \mathrm{Spec} K[c], \mathbb{Z}_p(1))_0 \xrightarrow{\mathrm{deg}} H^2(X_0(N) \times \mathrm{Spec} \bar{K}[c], \mathbb{Z}_p(1))_0 \cong \mathbb{Z}_p$$

which the difference (2.13) lies in the kernel of, then we obtain the element

$$y_c \in H^1(K[c], H^1(X_0(N)/\bar{K}, \mathbb{Z}_p(1))) = H^1(K[c], T_p J_0(N)).$$

Now, we don't want to keep tracking the field of definition, so let's think of  $x_c$  as a map over  $K$ . Starting with

$$\mathrm{Spec} K[c](\bar{K}) = \{\sigma : K[c] \hookrightarrow \bar{K}\} \simeq \mathrm{Gal}(K[c]/K) = K^\times \backslash \hat{K}^\times / \hat{\mathcal{O}}_c^\times,$$

we can think of

$$H^0(K^\times \backslash \hat{K}^\times / \hat{\mathcal{O}}_c^\times) \xrightarrow{\mathrm{cyc}} H^2(X_0(N) \times K^\times \backslash \hat{K}^\times / \hat{\mathcal{O}}_c^\times, \mathbb{Z}_p(1)).$$

There is an action of  $\mathrm{Gal}(K[c]/K)$  on  $X_0(N) \times K^\times \backslash \hat{K}^\times / \hat{\mathcal{O}}_c^\times$  that we may reinterpret as multiplication by  $\hat{K}^\times$ .

Now, let's take an embedding  $K \hookrightarrow M_2(\mathbb{Q})$  such that  $\mathcal{O}_K = K \cap M_2(\mathbb{Z})$ . This induces  $K^\times \hookrightarrow \mathrm{GL}_2(\mathbb{Q})$ . We fix  $\tau_0 \in \mathfrak{h}$  such that  $\mathrm{Stab}_{\mathrm{GL}_2^+(\mathbb{Q})}(\tau) = K^\times$ . For a compact open subgroup  $U$  of  $\mathrm{GL}_2(\mathbb{A}_f)$ , we set

$$X_U := \mathrm{GL}_2(\mathbb{Q}) \backslash [\mathfrak{h}^{\pm,*} \times \mathrm{GL}_2(\mathbb{A}_f)/U],$$

where  $*$  means “add the cusps” so as to obtain a compactification. The above contains the set of “CM points”

$$\{[\tau_0, g] : g \in \mathrm{GL}_2(\mathbb{A}_f)\},$$

which comes with with an action of  $\hat{K}^\times$  via the reciprocity map  $\hat{K}^\times \xrightarrow{\mathrm{rec}} G_K^{\mathrm{ab}}$  defined by

$$\mathrm{rec}(\alpha)[\tau_0, g] = [\tau_0, \alpha g].$$

We might as well take

$$U = U_0(N) = \mathrm{GL}_2(\hat{\mathbb{Z}}) \cap \begin{pmatrix} * & * \\ N* & * \end{pmatrix}.$$

In that case,  $x_c = [\tau_0, g_c]$  for a particular choice of  $g_c$ .

This is the classical picture of Heegner points. We're going to change the picture just a little bit. Think

$$H := \mathbb{Q}^\times \backslash K^\times \rightarrow G := \frac{\mathrm{GL}_2 \times K^\times}{\mathbb{Q}^\times}. \quad (2.14)$$

(The right hand side may be thought of as  $\mathrm{GU}(1, 1)$ , if you'd like.) We get

$$X_{U_0(N)} \rightarrow \mathrm{Sh}_G(U) := G(\mathbb{Q}) \backslash G(\mathbb{A}) / K_\infty U, \quad U := U_0(N) \times V.$$

We also get

$$\mathrm{Sh}_H(V) := \hat{\mathbb{Q}}^\times K^\times \backslash \hat{K}^\times / V.$$

There is a map

$$\mathrm{Sh}_G(U) \rightarrow \mathrm{Sh}_H(V).$$

In the interest of time, we're going to cut to the chase here. Look at compactly supported functions on  $H(\mathbb{A}_f) \backslash G(\mathbb{A}_f)$ , valued in  $\mathbb{Z}_p$ . We'll produce elements in the Shimura variety of  $G$  (over  $K$ ) valued in  $\mathbb{Z}_p(1)$ :

$$C_c(H(\mathbb{A}_f) \backslash G(\mathbb{A}_f), \mathbb{Z}_p) \rightarrow H_{\mathrm{ct}}^2(\mathrm{Sh}_{G/K}, \mathbb{Z}_p(1)) \quad (2.15)$$

$$\mathbf{1}(gU) \mapsto U_H = \hat{H}^\times gUg^{-1}.$$

Now we have

$$\mathrm{Sh}_H(U_H) \xrightarrow{\mathrm{cyc}} H^2(\mathrm{Sh}_G(gUg^{-1}), \mathbb{Z}_p(1)) \xrightarrow{g} H^2(\mathrm{Sh}_G(U), \mathbb{Z}_p(1)).$$

These maps are  $G(\mathbb{A}_f)$ -equivariant, or “Hecke equivariant”.

**Exercise 7.** Recover the Heegner point Euler system that we wrote down before by taking the right open compact and the right elements  $g_c$ , with  $V$  coming from  $\mathcal{O}_c$ .

The point is that through this process, the Galois action has been subsumed into some kind of group action. What's really going on in some ways is that the Shimura varieties for these unitary groups are just some copies of modular curves, where the number of copies is just the class group of some ring class extension. Once you have



this picture, the Euler system should just be the image of a bunch of elements on the left hand side of (2.15), and you just have to choose the elements:

$$\text{Euler system} = \text{image of elements in } C_c((H \backslash G)(\mathbb{A}_f)/\mathbb{Z}_p).$$

This can factor. Suppose  $c = \ell_1 \cdots \ell_r$  is ap roduct of primes. We might then take

$$\varphi_c = \varphi_S \otimes \bigotimes_{\ell_i} \varphi_{\ell_i} \otimes \bigotimes_{\substack{\ell \neq \ell_i \\ \ell \notin S}} \varphi_\ell^0 \in \left( C_c(X(\mathbb{Z}_S), \mathbb{Z}_p) \otimes C_c(X(\mathbb{A}_f^S), \mathbb{Z}_p) \right)^{U_c},$$

where  $U_c := U_0(N) \times \hat{\mathcal{O}}_c^\times$ , the  $\varphi_{\ell_i}$  are chosen well and  $\varphi_\ell^0 := \mathbf{1}(G(\mathbb{Z}_\ell))$  is the unramified choice. By choosing  $\varphi_S$  appropriately, we can even arrange that we're in the homologically trivial classes, hence giving rise to

$$z_c = \text{cyc}(\varphi_c)$$

in Galois cohomology; this is easy to arrange by just choosing the right section for  $\varphi_S$ , which amounts to acting by a Hecke operator of degree zero or something like that. The norm relations are all about the  $\varphi_{\ell_i}$ . We have

$$\text{cor}_{K[c\ell]/K[c]} z_{c\ell} = \text{cyc} \left( \sum_{t \in (\mathcal{O}_c \otimes \mathbb{Z}_\ell)^\times / (\hat{\mathcal{O}}_{c\ell} \otimes \mathbb{Z}_\ell)^\times} (1, t) \varphi_{c\ell} \right) \in C_c(X(\mathbb{A}_f), \mathbb{Z}_p)^{U_c}.$$

Suppose  $\ell \nmid c$ . Then we have

$$\varphi_\ell^0 = \mathbf{1}(X(\mathbb{Z}_\ell))$$

and

$$\varphi'_\ell := \text{trace } \varphi_\ell = \sum_t \varphi_\ell(x(1, t)).$$

What we're interested in is, how to compare these two? Both of the above define elements of

$$C_c(H \backslash G(\mathbb{Q}_\ell), \mathbb{Z}_p)^{G(\mathbb{Z}_\ell)} = C_c(H \backslash G(\mathbb{Q}_\ell)/G(\mathbb{Z}_\ell), \mathbb{Z}_p).$$

This space is cyclic for the action of the Hecke algebra (at least after tensoring by  $\mathbb{Q}_p$ )

$$\mathcal{H}(G // G(\mathbb{Z}_\ell), \mathbb{Q}_p)$$

generated with respect to that action by the  $\varphi_\ell^0$ . So we can write

$$\varphi'_\ell = P_\ell * \varphi_\ell^0.$$

In the remaining minutes, let's say something about how that calculation might go.

Let  $\pi$  and  $\chi$  be unramified representations of  $\text{GL}_2(\mathbb{Q}_\ell)$  and  $K_\ell^\times$ , respectively, with  $\chi_\pi \chi|_{\mathbb{Q}_\ell^\times} = 1$ . This pair defines a representation of the group  $G$  as in (2.14). Take a spherical vector  $\theta \in \pi$ . Maybe we realize  $\pi$  in the Whittaker model. Let's also assume that  $\ell$  splits. The sort of functional we're looking at is

$$\int_{(g, t_0) \in H \backslash G(\mathbb{Q}_\ell)} \int_{t \in \mathbb{Q}_\ell^\times \backslash K_\ell^\times} \theta(tg) \chi(tt_0) \varphi(g, t_0) d^\times t = \Lambda(\varphi).$$

Taking

$$\varphi = \varphi'_\ell = P_\ell * \varphi_\ell^0,$$

we obtain

$$\Lambda(\varphi'_\ell) = \xi_{(\pi, \chi)}(P_\ell) \cdot \Lambda(\varphi_\ell^0),$$

where

$$\Lambda(\varphi_\ell^0) = \int_{\mathbb{Q}_\ell^\times \setminus K_\ell^\times} \theta(t) \chi(t) d^\times t.$$

We might want to arrange that  $1 = \Lambda(\varphi'_\ell)$ . Since  $\Lambda(\varphi_\ell^0)$  is the  $L$ -factor, this motivates taking  $\xi(P_\ell)$  to be the Hecke polynomial of which the  $L$ -factor is the inverse. Here we're assuming that  $\ell$  splits in  $K/\mathbb{Q}$ .

Locally, we can think

$$K_\ell^\times = \mathcal{O}_\ell^\times \times \mathcal{O}_\ell^\times \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell),$$

where the inclusion is the diagonal embedding. We can take

$$g_\ell = \begin{pmatrix} 1 & \frac{1}{\ell} \\ 0 & 1 \end{pmatrix}.$$

Then

$$\varphi'_\ell = \sum_{a \in (\mathbb{Z}/\ell)^\times} \varphi_\ell^0 \left( \begin{pmatrix} 1 & \frac{a}{\ell} \\ 0 & 1 \end{pmatrix} \right).$$

You end up with something that's not quite what you want, but is *congruent* to what you want modulo  $\ell - 1$ . Anyway, the norm relation from before falls out.

You can do this more generally for  $H \subset G$  defining Shimura varieties. One can is to take

$$G = G(\mathrm{U}(n, 1) \times \mathrm{U}(n - 1, 1) \times \mathrm{U}(1)),$$

which contains

$$H = G(\mathrm{U}(n - 1, 1) \times \mathrm{U}(1)).$$

(This belongs to the “GGP” setting.) We have a corresponding map of Shimura varieties

$$\mathrm{Sh}_H \rightarrow \mathrm{Sh}_G$$

of dimensions  $n - 1$  and  $n + n - 1 = 2n - 1$ , so the codimension is  $n$ , and we get a cycle class map

$$H^0(\mathrm{Sh}_H) \xrightarrow{\mathrm{cyc}} H^{2n}(\mathrm{Sh}_G, n).$$

This leads to

$$H^1(\text{ring class fields}, H^{2n-1}(\mathrm{Sh}_{\mathrm{U}(n-1,1) \times \mathrm{U}(n,1)})).$$

See the work of the speaker and Xinwen Wan.

**Remark 2.4.** We consider

$$C_c(X(\mathbb{Q}_\ell), \mathbb{Z}_p)^{G(\mathbb{Z}_\ell)} \supseteq \mathrm{trace} C_c(X(\mathbb{Q}_\ell), \mathbb{Z}_p)^{G_\ell(\mathbb{Z}_\ell)},$$

for some subgroup  $G_\ell(\mathbb{Z}_\ell) \subseteq G(\mathbb{Z}_\ell)$ . What is the right hand side? Contains

$$\left( P_\ell, (\ell - 1) C_c(\dots)^{G(\mathbb{Z}_\ell)} \right).$$

**Remark 2.5.** See paper of David Loeffler regarding the ramified primes

## 3. JAYCE GETZ'S LECTURES

**3.1. Langlands  $L$ -functions and converse theorem.** Let  $F$  be a number field,  $\mathbb{A}_F$  the ring of adeles, and  $G$  a split reductive group over  $F$ . Let  $A_G$  denote the neutral component in the real topology of the greatest  $\mathbb{Q}$ -split torus in  $\text{Res}_{F/\mathbb{Q}} Z_G$ . (For example,  $A_{\text{GL}_r} = \mathbb{R}_{>0} I_r$ .)

**Goal:** given an  $L$ -map  $\rho : {}^L G \rightarrow \text{GL}_n(\mathbb{C})$ , construct a transfer of automorphic representations of  $G(\mathbb{A}_F)$  to  $\text{GL}_n(\mathbb{A}_F)$ .

Given an automorphic representation  $\pi$  of  $G(\mathbb{A}_F)$ , we can form the Langlands  $L$ -function

$$L(s, \pi, \rho) := \prod_v L(s, \pi_v, \rho).$$

If  $\pi$  is unramified at  $v$ , then

$$L(s, \pi_v, \rho) = \frac{1}{\det(1 - \rho(c(\pi_v))q_v^{-s})},$$

where  $c(\pi_v) \in {}^L G$  is the Langlands class of  $\pi_v$ .

Let  $\psi : F \backslash \mathbb{A}_F \rightarrow \mathbb{C}^\times$  be a nontrivial character. Then we obtain  $\varepsilon$ -factors  $\varepsilon(s, \pi_v, \rho, \psi_v)$ , defined via local Langlands and the known definition in the case of  $\text{GL}_n$ . There is also the  $\gamma$ -factor

$$\gamma(s, \pi_v, \rho, \psi_v) := \frac{\varepsilon(s, \pi_v, \rho, \psi_v) L(1 - s, \pi_v^\vee, \rho)}{L(s, \pi_v, \rho)}.$$

**Conjecture 3.1** (Langlands).  $L(s, \pi, \rho)$  admits a meromorphic continuation to  $\mathbb{C}$ , is bounded in vertical strips, and satisfies a functional equation

$$L(s, \pi, \rho) = \varepsilon(s, \pi, \rho, \psi) L(1 - s, \pi^\vee, \rho).$$

If  $L(s, \pi, \rho)$  satisfies the conjecture, then we say that it is *nice*.

One case where we know this in complete generality is the case of the standard representation

$$\rho := \rho_{\text{st}} = \text{triv} \times \text{Id} : {}^L \text{GL}_n = \text{Gal}_F \times \text{GL}_n(\mathbb{C}) \rightarrow \text{GL}_n(\mathbb{C}),$$

in which case we write simply

$$L(s, \pi) := L(s, \pi, \rho_{\text{st}}).$$

Then the conjecture is known, and you can prove it in three different ways that are largely distinct:

- Godement–Jacquet (which preceded the general formulation of this conjecture).
- The Rankin–Selberg method.
- The Langlands–Shahidi method.

At first glance, none of these methods will work to address the general conjecture, but we'll see that there is a sort of modification of the first two that might be applicable.

Now, what's the converse theorem?

**Theorem 3.2** (Cogdell–Piatetski-Shapiro; Hecke, Weil). *Let  $\sigma$  be an admissible irreducible representation of  $\text{GL}_n(\mathbb{A}_F)$ . (In order to get this sort of thing, you have to choose a bunch of local data and outside finitely many places you declare that it is unramified.) Assume that*

- the central character  $\omega_\sigma$  is  $Z_{\mathrm{GL}_n(F)}$ -invariant,
- $\sigma_v$  is generic for all  $v$ , and
- $L(s, \sigma)$  converges for  $\Re s$  sufficiently large.

Let  $S$  be a finite set of finite primes. Assume that  $L(s, \sigma \times \sigma')$  is nice and holomorphic for all cuspidal automorphic representations  $\sigma'$  of  $\mathrm{GL}_m(\mathbb{A}_F)$  unramified at  $S$ , with  $1 \leq m \leq n-2$  (for  $n \geq 2$ ; for  $n = 2$ , just take  $m = 1$ ). Then there exists an automorphic representation  $\pi'$  of  $\mathrm{GL}_n(\mathbb{A}_F)$  that is nearly equivalent to  $\sigma$  in the sense that  $\sigma_v \cong \pi'_v$  for all  $s \notin S$ .

To see the relevance of this, take  $\pi$  on  $G(\mathbb{A}_F)$ , and take  $\sigma = \rho(\pi)$ , giving an admissible representation for almost all places (assuming the local Langlands correspondence).

This gives some overplay of the interplay between  $L$ -functions and automorphic representations and exactly how the two are interrelated via functoriality. This sort of thing is in the background when you read things on the Braverman–Kazhdan program, Beyond Endoscopy, etc.

Now let's talk about a particular case.

**3.2. Rankin–Selberg products and triple products.** Let  $\underline{r} = (r_1, r_2, r_3)$ . Set  $\mathrm{GL}_{\underline{r}} := \mathrm{GL}_{r_1} \times \mathrm{GL}_{r_2} \times \mathrm{GL}_{r_3}$ . We have an  $L$ -map that we call the Rankin–Selberg product (taking complex points, say)

$$\otimes^2 : {}^L(\mathrm{GL}_{r_1} \times \mathrm{GL}_{r_2}) \rightarrow \mathrm{GL}_{r_1 r_2}, \quad (3.1)$$

$$\otimes^3 : {}^L\mathrm{GL}_{\underline{r}} \rightarrow \mathrm{GL}_{r_1 r_2 r_3}. \quad (3.2)$$

We'll think of these as hypothetical Langlands transfers.

**Theorem 3.3** (Chevalley). *Given any faithful representation  $\rho : H \rightarrow \mathrm{GL}_n$  (say with  $H$  reductive), any other representation is a subrepresentation of  $\rho^{\otimes m} \otimes \rho^{\vee \otimes n}$  for some  $m$  and  $n$ .*

Now,  ${}^L G$  is not a reductive group, but ignore that fact for a moment. Say we're given an  $L$ -map

$$\rho : {}^L G \rightarrow \mathrm{GL}_n. \quad (3.3)$$

Suppose we know the Rankin–Selberg transfer (3.1) as well as transfer for (3.3). Then, any transfer from  $G$  to  $\mathrm{GL}(V)$  is a “subtransfer” of some  $\rho^{\otimes m} \otimes \rho^{\vee \otimes n}$ .

**Question 3.4** (Not an exercise). Given functorial transfers  $\rho^{\otimes m} \otimes \rho^{\vee \otimes n}$ , how to pick out the transfers associated to subrepresentations?

By analogy to representation theory, some of the basic ways to construct new representations from old ones is to take direct sums, induced representations and tensor products. In the automorphic world, we know how to do the first two (isobaric sum and Eisenstein series), but not the third.

The importance of triple product  $L$ -functions (attached to (3.2)) is that if they are nice, then, using the converse theorem, we get that Rankin–Selberg transfers exist. This is the motivation for what we're about to do.

**3.3. Poisson summation conjecture.** In the exercises, we'll see the Godement–Jacquet Poisson summation formula for  $M_{n \times n}$ , which leads to the functional equation and analytic continuation of the standard  $L$ -functions  $L(s, \pi)$ . Bravmern–Kazhdan suggest replacing  $M_{n \times n}$  by a certain  $M_\rho$ , and try to run the argument to study  $L(s, \pi, \rho)$ . This proposal was refined by Ngô. Sakellaridis suggested working with some spherical varieties. Need a little more flexibility.

We'll start with the geometric setup.

**3.3.1. Global setting.** Let  $F$  be a number field, let  $R$  be an  $F$ -algebra, and let  $G/F$  be a reductive group. Assume given an action

$$\mathcal{R} : X \times G \rightarrow X$$

on an affine  $G$ -scheme  $X$ . Assume that there is a unique open orbit  $X^\circ \subseteq X$ . Let  $Z_G^+$  denote the greatest split torus in  $Z_G$  (e.g., for  $G = \mathrm{GL}_n$ ,  $Z_G^+ = \mathbb{G}_m$ ). Fix an isomorphism  $\mathbb{G}_m^n \xrightarrow{\cong} Z_G^+$ . Assume that the pullback  $\mathcal{R} : X \times \mathbb{G}_m^n \rightarrow X$  extends to  $\mathcal{R} : X \times \mathbb{G}_a^n \rightarrow X$ .

**Example 3.5.**  $X = \mathbb{G}_a$  (which we don't write as  $\mathbb{A}^1$ , because that denotes the adeles). Then the map  $X \times \mathbb{G}_m \rightarrow X$  given by  $(t, a) \mapsto ta$  extends to  $\mathbb{G}_a \supset \mathbb{G}_m$  (at 0, but not at  $\infty$ ).

Assume that there exists an automorphic  $\iota : G \rightarrow G$  such that  $\iota(z) = z^{-1}$  for  $z \in Z_G^+(\mathbb{R})$ .

**Example 3.6.** Take  $X$  to be a vector space,  $G = \mathrm{GL}_X$ ,  $\iota(g) = g^{-\mathrm{T}}$ , and  $X(R) := \{x \in R^n \mid Q(x) = 0\}$ , where  $Q$  is an anisotropic quadratic form. Or take  $G = \mathrm{GO}_Q$ , and  $\iota(g) = \mathrm{simil}(g^{-1})g$ .

We need a generalization. In integral representations, you often need some sort of Whittaker or Fourier coefficients to integrate against. The reason is that if you don't have it, then when you integrate against cusp forms, you get zero. If we don't have this in the cases of interest, then we end up getting zero. There's a geometric way of phrasing this:

**Definition 3.7.** Suppose given a  $\mathbb{G}_a^n$ -torsor  $p : V^\circ \rightarrow X^\circ$  equipped with a  $\mathbb{G}_a^n \rtimes G$ -action. Assume that  $p$  is  $\mathbb{G}_a^n \rtimes G$ -equivariant. Consider the map

$$V^\circ(R) \times R^n \times G(R) \rightarrow V^\circ(R) \times R^n,$$

$$(v, z, g) \mapsto (vg, (0 \times g)^{-1}z(0 \times g)).$$

Assume given  $\Psi : V^\circ \times \mathbb{G}_a^n \rightarrow \mathbb{G}_a$  such that

$\Psi 1$   $\Psi$  factors through  $p \times \mathrm{id} : V^\circ \times \mathbb{G}_a^n \rightarrow X^\circ \times \mathbb{G}_a^n$ .

$\Psi 2$   $\Psi$  is  $G$ -invariant.

$\Psi 3$  For each  $w \in V^\circ(R)$ , the map  $\Psi(w, \bullet) : \mathbb{G}_{a,R}^n \rightarrow \mathbb{G}_{a,R}$  is a group homomorphism.

An *affine  $\Psi$ -bundle* is a pair  $(p : V^\circ \rightarrow X^\circ, \Psi)$  as above.

Given  $w \in V^\circ(R)$ , set  $x := p(w)$ . We obtain  $R^n \xrightarrow{\cong} V_x^0(R)$ ,  $z \mapsto wz$ . For  $g \in G_\lambda(R)$ , there is a unique  $z_w(g) \in R^n$  such that  $wz_w(g) = wg$ . Then  $\Psi_w(g) = \Psi(w, z_w(g))$ .

**Lemma 3.8.**  $\Psi_w : G_x \rightarrow \mathbb{G}_a$  is a character, depending only upon  $x$ .

This is the geometry that underlies the Poisson summation conjecture. In the remaining time today, we'll start doing some of the harmonic analysis. We'll enrich the geometry by introducing some function spaces on the  $F$ -points of everything. We'll start that process now and continue it in the next lecture.

**3.3.2. Local harmonic analysis.** Start with some global character  $\psi : F \backslash \mathbb{A}_F \rightarrow \mathbb{C}^\times$ . Let  $v$  be a place of  $F$ . Define  $F = F_v$ ,  $\psi = \psi_v$ . Thus, we'll work locally and omit  $v$  from notation. We ask for a Schwartz space sitting as follows:

$$C_c^\infty(X^\circ(F)) \subseteq \mathcal{S}(X(F)) \subseteq C^\infty(X^\circ(F)) \cap L^2(X^\circ(F)).$$

(This assumes that we have an invariant measure to define  $L^2$ . It's better in some sense to work instead with half-densities.) We want the Schwartz space to be stable under  $G(F)$ . One can do a bit more: one can form an  $\varepsilon_\psi$ -hermitian bundle over  $X^\circ(F)$ , using the affine  $\Psi$ -bundle just mentioned. The sections are

$$\{f \in C^\infty(V^\circ(F)) \mid \mathcal{R}(a \times I)f(w) = \psi(\Psi(w, a))f(w)\}.$$

The motivation is that if you don't incorporate such sections rather than just functions, then you'll get zero when you try to integrate against cusp forms. Consider for instance a cusp form  $\varphi$  on  $\mathrm{GL}_2(\mathbb{A}_F)$ . Then

$$\int_{[\mathbb{G}_a]} \varphi \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} dx = 0,$$

while

$$\int_{[\mathbb{G}_a]} \varphi \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \psi(x) dx = W_\psi^\varphi(1).$$

Actually, it's better, and we do it in the paper (see the course references (TODO: add link here)), and you need it. We twist by half-densities:  $\varepsilon_\psi \otimes |\Lambda|^{1/2} = \mathcal{L}_\psi$ . These are canonical – no choice of measure. They are helpful for normalization. On the other hand, you can't evaluate at points. But you can work around this. Using these makes the formulation of the theory much simpler. Another benefit is that with half-densities, Schwartz spaces behave well under restriction. Maybe the reason behind this might be an exercise:

**Exercise 8.**  $L^2(X^\circ(F), |\Lambda|^{1/2})$  is a unitary  $G(F)$ -representation.

#### 4. CHEN WAN'S LECTURES

We'll focus on the *period* aspect of the relative Langlands duality of Ben-Zvi–Sakellaridis–Venkatesh. Recall from Chris's talk that we saw something about integral representations of  $L$ -functions, which involves integrating automorphic forms  $\phi$  on some larger space  $Y$  over some smaller space  $X$ . In nice cases, such integrals represent  $L$ -functions:

$$\int_X (\phi, \text{automorphic form on } Y \supseteq X) = L\text{-function.} \quad (4.1)$$

We've seen in previous lectures that if you have this sort of integral representation of  $L$ -functions, then you can use it to study the meromorphic continuation of  $L$ -functions, or their special values. Such identities are thus very useful.

The examples might seem a bit *ad hoc*. Why should we integrate automorphic forms on  $Y$  over  $X$ , for certain pairs  $(X, Y)$ ? Is there some way to systematically find, for a given  $L$ -function, which integrals represent them? Can we determine

systematically which  $L$ -functions are represented by some integral? The theory of relative Langlands duality at least gives some explanation of this in some cases. This is what we'll talk about in our lectures.

Let's first say the rough idea. If you look at this kind of integral (4.1), it's kind of an automorphic integral involving the quotient variety  $\mathcal{M} = Y/X$  (whatever this means). Relative Langlands duality generalizes the classical Langlands correspondence to varieties: one passes from  $\mathcal{M}$  to some dual variety  $\hat{\mathcal{M}}$ , and the proposal in relative Langlands duality is that

$$(\text{automorphic integral on } \mathcal{M}) = (L\text{-function associated to } \hat{\mathcal{M}}).$$

Moreover, it's a duality, which means that the same holds with the roles of  $\mathcal{M}$  and  $\hat{\mathcal{M}}$  reversed:

$$(\text{automorphic integral on } \hat{\mathcal{M}}) = (L\text{-function associated to } \mathcal{M}).$$

Everything is general at this point – later, we'll say precisely what it means. But the idea is that at least in terms of periods integrals of automorphic forms, you have certain nice Hamiltonian spaces and dualities between them so that the integrals attached to one space are given by the  $L$ -functions attached to the other.

Let's now set up some notation. We'll first define the automorphic integrals associated to this data, and if there's time, the  $L$ -functions and the duality. Let  $K$  be a global field, and  $\mathbb{A} = \mathbb{A}_K$  the ring of adeles. Sometimes we'll also work over a local field  $F$ , (e.g.,  $\mathbb{R}, \mathbb{C}, \mathbb{Q}_p, \dots$ ). Let  $G$  be a connected reductive group defined over  $K$ . (If you're not familiar with this language, you can just think of the matrix groups that you're familiar with, such as  $\mathrm{GL}_n$ ,  $\mathrm{SL}_n$ ,  $\mathrm{PGL}_n$ ,  $\mathrm{SO}_n$ ,  $\mathrm{Sp}_{2n}$ .) Throughout these lectures, we'll mostly focus on the groups that are *split*, i.e., that have a maximal split torus.

The foundation of the Langlands program is based on the duality between these reductive groups. For each such group  $G$ , we can write down its dual group  $\hat{G}$ . Reductive groups are parametrized by root data. For each root data, you can take

| $G$                  | $\hat{G}$          |
|----------------------|--------------------|
| $\mathrm{GL}_n$      | $\mathrm{GL}_n$    |
| $\mathrm{SL}_n$      | $\mathrm{PGL}_n$   |
| $\mathrm{SO}_{2n+1}$ | $\mathrm{Sp}_{2n}$ |
| $\mathrm{SO}_{2n}$   | $\mathrm{SO}_{2n}$ |

its dual root data, which then defines the dual group:

The *Langlands conjecture* says roughly that automorphic representations of  $G(\mathbb{A})$  should correspond to Langlands parameters  $\phi : L_K \rightarrow \hat{G}(\mathbb{C})$  valued in the dual group, and similarly with  $G$  and  $\hat{G}$  swapped.

**Definition 4.1.** A *BZSV quadruple* for  $G$  is  $\Delta = (G, H, \iota, \rho_H)$ , where

- (1)  $H \subset G$  is a split connected reductive subgroup,
- (2)  $\iota : \mathrm{SL}_2 \rightarrow G$  is a homomorphism (possibly trivial) such that  $H$  commutes with  $\mathrm{image}(\iota)$ , so that we may extend to  $\iota : H \times \mathrm{SL}_2 \rightarrow G$ .
- (3)  $\rho_H$  is a symplectic representation of  $H$ , i.e., a homomorphism  $H \rightarrow \mathrm{Sp}(V)$ .

Define the automorphic quotient  $[G] := G(K) \backslash G(\mathbb{A})$ , and let  $\phi : [G] \rightarrow \mathbb{C}$  be an automorphic form. The goal today is to define the period integral  $\mathcal{P}_\Delta(\phi)$  associated to a quadruple  $\Delta$  as above.

Let's start with some easy special cases, before moving to the general case.

- (1) Suppose  $\iota = 1$  (the trivial representation,  $\mathbb{C}$ ) and  $\rho_H = 0$  (the *zero-dimensional* symplectic representation  $\emptyset$ ), thus  $\Delta = (G, H, 1, 0)$ . Then

$$\mathcal{P}_\Delta(\phi) := \int_{[H]} \phi(h) dh.$$

(In general, this does not converge and must be understood via truncation, etc.)

- (2) Suppose still that  $\iota = 1$  is trivial, but allow  $\rho_H$  to be general. Then

$$\mathcal{P}_\Delta(\phi) := \int_{[H]} \phi(h) \cdot \Theta_{\rho_H}(h) dh,$$

where  $\Theta_{\rho_H}$  is a suitable theta function, whose definition depends also upon the choice of a Schwartz function:

- (a) It's easiest to define in the special *polarized* case that  $\rho_H = \tau \oplus \tau^\vee$  for some  $\tau : H \rightarrow \mathrm{GL}(W)$ . In that case, after choosing  $\varphi \in \mathcal{S}(W(\mathbb{A}))$ , we define

$$\Theta_{\rho_H}(h) := \sum_{x \in W(F)} \varphi(\tau(h)x).$$

To illustrate that case, we consider a couple examples:

- (i) Suppose

$$\Delta = (\mathrm{GL}_n \times \mathrm{GL}_n, \mathrm{GL}_n \times \mathrm{GL}_n, 1, \mathrm{std} \otimes \mathrm{std} \oplus (\mathrm{std} \otimes \mathrm{std})^\vee)$$

Then, writing  $\phi = \phi_1 \otimes \phi_2$  with each  $\phi_i : [\mathrm{GL}_n] \rightarrow \mathbb{C}$ , we may take  $W = \mathrm{Mat}_{n \times n}$  above (the space of  $n \times n$  matrices), and we obtain

$$\mathcal{P}_\Delta(\phi) := \int_{[\mathrm{GL}_n] \times [\mathrm{GL}_n]} \phi_1(h_1) \phi_2(h_2) \sum_{x \in \mathrm{Mat}_{n \times n}(F)} \varphi(h_1^{-1} x h_2) dh_1 dh_2.$$

- (ii) Take  $\Delta = (\mathrm{GL}_n \times \mathrm{GL}_n, \mathrm{GL}_n^{\mathrm{diag}}, 1, \mathrm{std} \oplus \mathrm{std}^\vee)$ . Then, taking  $W$  to be the  $n$ -dimensional vector space on which  $\mathrm{GL}_n$  acts by right multiplication, we obtain

$$\mathcal{P}_\Delta(\phi) := \int_{[\mathrm{GL}_n]} \phi_1(h) \phi_2(h) \sum_{x \in W(K)} \varphi(xh) dh,$$

which is the Rankin–Selberg integral for  $\mathrm{GL}_n \times \mathrm{GL}_n$ .

- (b) In general,  $\rho_H : H \rightarrow \mathrm{Sp}(V)$  induces the so-called Weil representation, which may be realized on  $\mathcal{S}(Y(\mathbb{A}))$  for a maximal isotropic subspace  $Y$  of  $V$ . As you may know, the Weil representation is not on  $\mathrm{Sp}(V)$ , but on its double cover. So you can't do it for an arbitrary symplectic representation – need a condition called “anomaly-free” that allows you to resolve the ambiguity, but let's not go into the details there. There is anyway some condition under which we can split the Weil representation over  $H$ , giving a well-defined representation. The splitting is not necessarily unique, and perhaps BZSV do not currently choose a canonical splitting, but it's expected that there is some natural splitting. Of course there is such a natural splitting in the polarized case. In any event, one then sets, for  $h \in H(\mathbb{A})$  and  $\varphi \in \mathcal{S}(Y(\mathbb{A}))$ ,

$$\Theta_{\rho_H}(h) := \sum_{x \in Y(K)} \Omega_{\rho_H}(h) \varphi(x).$$



(3) In the fully general case, we define

$$\mathcal{P}_\Delta(\phi) := \int_{[H]} \mathcal{P}_\iota(\phi)(h) \Theta_{\rho_H}(h) dh. \quad (4.2)$$

Here  $\mathcal{P}_\iota$  is the Fourier coefficient of  $\phi$  associated to  $\iota$ , defined as follows. We first recall that  $\iota : \mathrm{SL}_2 \rightarrow G$ . In particular, we can let  $L := \mathrm{Cent}_G(\mathrm{image}(\iota(T_{\mathrm{SL}_2})))$ , with  $T_{\mathrm{SL}_2} \leq \mathrm{SL}_2$  the diagonal maximal torus consisting of elements of the form  $\mathrm{diag}(t, t^{-1})$ . Then  $L$  is a Levi subgroup of  $G$ . We also know, since  $H$  commutes with the image of  $\iota$ , that  $H$  is a subgroup of  $L$ . Now let  $U := \exp(\mathfrak{u})$ , where  $\mathfrak{u}$  is the positive weight space for  $\mathrm{Ad}(\iota(\mathrm{diag}(t, t^{-1})))$ . We also define  $\bar{U} := \exp(\bar{\mathfrak{u}})$ , with  $\bar{\mathfrak{u}}$  the negative weight space.

**Example 4.2.** Take  $G = \mathrm{GL}_{2n}$  and  $\iota : \mathrm{SL}_2 \rightarrow \mathrm{GL}_{2n}$  given by

$$\begin{aligned} \begin{pmatrix} t & \\ & t^{-1} \end{pmatrix} &\mapsto \begin{pmatrix} tI_n & \\ & t^{-1}I_n \end{pmatrix}, \\ \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} &\mapsto \begin{pmatrix} I_n & I_n \\ & I_n \end{pmatrix}, \\ \begin{pmatrix} 1 & \\ 1 & 1 \end{pmatrix} &\mapsto \begin{pmatrix} I_n & \\ I_n & I_n \end{pmatrix}. \end{aligned}$$

Then

$$L = \left\{ \begin{pmatrix} a & \\ & b \end{pmatrix} : a, b \in \mathrm{GL}_n \right\}, \quad U = \left\{ \begin{pmatrix} I_n & * \\ & I_n \end{pmatrix} \right\}, \quad \bar{U} = \left\{ \begin{pmatrix} I_n & \\ * & I_n \end{pmatrix} \right\}.$$

If  $\iota$  is even (i.e., all the weights of  $\mathrm{Ad}(\iota(\mathrm{diag}(t, t^{-1})))$  are even), then the Fourier coefficient is given by a very easy formula:

$$\mathcal{P}_\iota(\phi)(h) := \int_{[U]} \phi(uh) \xi(u)^{-1} du, \quad (4.3)$$

where  $\xi : [U] \rightarrow \mathbb{C}^\times$  is a character defined, for  $x \in \mathfrak{u}(\mathbb{A})$ , by

$$\xi(\exp(x)) := \psi(\langle x, \iota \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \rangle),$$

where  $\psi : K \backslash \mathbb{A} \rightarrow \mathbb{C}^\times$  is some fixed character.

**Example 4.3.** In the above example,

$$\xi \begin{pmatrix} I_n & x \\ & I_n \end{pmatrix} = \psi \left( \left\langle \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\rangle \right) = \psi(\mathrm{trace}(x)).$$

**Exercise 9.** Let  $G := \mathrm{GL}_n$  and  $\iota : \mathrm{SL}_2 \rightarrow \mathrm{GL}_n$  be  $\rho_{\mathrm{sym}}^{n-1}$ . Then write down the Fourier coefficient  $\mathcal{P}_\iota(\phi)$ .

If  $\iota$  is *not* even, then we have a Weil representation  $\Omega_\iota$  of  $U(\mathbb{A})$  on  $\mathcal{S}(Z(\mathbb{A}))$ , where  $Z$  is a maximal isotropic subspace of the weight 1 space of  $\mathrm{Ad}(\iota(\mathrm{diag}(t, t^{-1})))$ . Then we build another theta kernel  $\Theta_\iota(u) := \sum_{x \in Z(K)} \Omega_\iota(u) \varphi(x)$ , where  $\varphi \in \mathcal{S}(Z(\mathbb{A}))$ . The period integral (depending upon  $\varphi$ ) is then given by integrating not just against a character as in (4.3), but against the theta function:

$$\mathcal{P}_\iota(\phi) := \int_{[U]} \phi(u) \Theta_\iota(u) du. \quad (4.4)$$

So, in the integral  $\mathcal{P}_\Delta(\phi)$ , there are possibly two different choices of Schwartz functions: one in the theta kernel in (4.2), and another in the

theta kernel in (4.4). But in some special cases ( $\iota$  trivial or even), we don't need these.

**Example 4.4.** Take  $\Delta = (G, H, \iota, \rho_H)$ , where

$$G = \mathrm{GL}_{2n},$$

$$H = \mathrm{GL}_n = \left\{ \begin{pmatrix} h & \\ & h \end{pmatrix} \mid h \in \mathrm{GL}_n \right\},$$

$\rho_H = 0$  and

$$\iota : \begin{pmatrix} t & \\ & t^{-1} \end{pmatrix} \mapsto \begin{pmatrix} tI_n & \\ & t^{-1}I_n \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \mapsto \begin{pmatrix} I_n & I_n \\ & I_n \end{pmatrix}.$$

Then

$$\mathcal{P}_\Delta(\phi) = \int_{[\mathrm{GL}_n]/\mathrm{GL}_1(\mathbb{A})} \int_{[\mathrm{Mat}_{n \times n}]} \phi \left( \begin{pmatrix} I_n & x \\ & I_n \end{pmatrix} \begin{pmatrix} h & \\ & h \end{pmatrix} \right) \psi(\mathrm{trace}(x))^{-1} dx dh,$$

which is relevant for the *Shalika model*.

**Exercise 10.** For  $G = \mathrm{GL}_6$ , find  $\Delta$  for which  $\mathcal{P}_\Delta(\phi)$  is given by

$$\int_{[\mathrm{GL}_2]} \int_{[\mathrm{Mat}_{2 \times 2}^3]} \phi \left( \begin{pmatrix} I_2 & X & Z \\ & I_2 & Y \\ & & I_2 \end{pmatrix} \begin{pmatrix} h & & \\ & h & \\ & & h \end{pmatrix} \right) \psi(\mathrm{trace}(X + Y))^{-1} dX dY dZ dh.$$

**Example 4.5.**  $\iota : \mathrm{SL}_2 \rightarrow \mathrm{GL}_3$  given by  $h \mapsto \mathrm{diag}(h, 1)$  is *not* even, because we get odd weight spaces for  $\mathrm{diag}(t, t^{-1}, 1)$  acting via the adjoint representation for  $\mathrm{GL}_3$ .

⚡⚡⚡

## 5. RAPHAEL BEUZART-PLESSIS'S LECTURES

### Introduction to Relative Trace Formulas (RTF)

**5.1. RTF in general.** Let  $K$  be a global field. Let  $G$  be a connected reductive group over  $K$ . Let  $A(G)$  denote the space of automorphic forms on the automorphic quotient  $[G] := G(K) \backslash G(\mathbb{A}_K)$ .

**5.1.1. Periods.** Let  $H \subset G$  be an algebraic group. Let  $\chi : [H] \rightarrow \mathbb{C}^\times$  be a character. By the  $(H, \chi)$ -period, we mean the functional

$$\mathcal{P}_{H, \chi} : A(G) \rightarrow \mathbb{C}$$

$$\varphi \mapsto \int_{[H]} \varphi(h) \chi(h) dh.$$

**Remark 5.1.** The integral often diverges. We'll explain in some examples how one might regularize them. It's usually convergent if the automorphic form  $\varphi$  is cuspidal. (Usually, but not always – there might be issues with the center, or if one takes  $H$  to be the Borel subgroup of  $G = \mathrm{GL}_2$ , then it never converges.)

**Remark 5.2.** We can replace  $\chi$  by some “small” automorphic representation of  $H$  (e.g., “degenerate” Eisenstein series).

**Empirical fact:**  $\mathcal{P}_{H, \chi}$  is often related to  $L$ -functions and functoriality.

**Example 5.3.** (a) (Hecke)  $G = \mathrm{GL}_2 \supset \begin{pmatrix} * & \\ & 1 \end{pmatrix} := \mathbb{G}_m$ . For  $\varphi \in \pi \subset \mathcal{A}_{\mathrm{cusp}}(G)$ ,

$$\mathcal{P}_{H, |\cdot|^s}(\varphi) \sim L(\tfrac{1}{2} + s, \pi).$$

- (b) (Rankin–Selberg)  $G = \mathrm{GL}_2 \times \mathrm{GL}_2 \supset H = \mathrm{GL}_2^{\mathrm{diag}}$ ,  $\int_{[H/Z]} \varphi_1(h) \varphi_2(h) E(h, f, s) dh \sim L(s, \pi_1 \times \pi_2)$ , where  $\varphi_i \in \pi_i \subset A_{\mathrm{cusp}}(\mathrm{GL}_2)$  and  $E$  is a suitable Eisenstein series.
- (c) (Harder–Langlands–Rapoport) Let  $L/K$  be a quadratic extension,  $G = \mathrm{GL}_{2,L}$ ,  $H = \mathrm{GL}_{2,K}$ . Take  $\pi \subset A_{\mathrm{cusp}}(\mathrm{GL}_{2,K})$  and  $\chi : \mathbb{A}_K^\times / K^\times \rightarrow \mathbb{C}^\times$  such that  $\omega_\pi = \chi^{-1} \circ N_{L/K}$ . Then the following are equivalent:
- There exists  $\varphi \in \pi$  with  $\mathcal{P}_{H,\chi}(\varphi) \neq 0$ .
  - There exists  $\sigma \subset A_{\mathrm{cusp}}(\mathrm{GL}_{2,K})$  such that  $\pi$  is the base change  $\sigma_L$  and  $\omega_\sigma = \chi^{-1}$ .
- (d) (Jacquet–Shalika) Take

$$G := \mathrm{GL}_{2n} \supseteq H = \left\{ \begin{pmatrix} g & \\ & g \end{pmatrix} \begin{pmatrix} 1 & X \\ & 1 \end{pmatrix} \mid g \in \mathrm{GL}_n, X \in \mathrm{Mat}_n \right\}.$$

Let

$$\begin{aligned} \chi : [H] &\rightarrow \mathbb{C}^\times, \\ \begin{pmatrix} g & \\ & g \end{pmatrix} \begin{pmatrix} 1 & X \\ & 1 \end{pmatrix} &\mapsto \psi(\mathrm{trace}(X)). \end{aligned}$$

For  $\pi \subset A_{\mathrm{cusp}}(\mathrm{GL}_{2n}/Z)$ , the following are equivalent:

- $\mathcal{P}_{H,\chi}|_\pi$  is not identically zero.
- $\pi$  is a functorial lift from  $\mathrm{SO}_{2n+1}$ .

5.1.2. *Theta series.* Take  $X := H \backslash G$ . Define  $\Theta^X : C_c^\infty(X(\mathbb{A})) \rightarrow C^\infty([G])$  by

$$f \mapsto \Theta_f^X(g) := \sum_{x \in X(K)} f(xg).$$

For  $\varphi \in A(G)$ , define

$$\mathcal{P}_X(\varphi, f) := \int_{[G]} \varphi(g) \Theta_f^X(g) dg.$$

Provided that  $f$  is supported inside  $H(\mathbb{A}) \backslash G(\mathbb{A})$  (in general, a proper subset of  $X(\mathbb{A})$ , with the obstruction measured by Galois cohomology), we have

$$\mathcal{P}_X(\varphi, f) = \int_{H(\mathbb{A}) \backslash G(\mathbb{A})} \mathcal{P}_H(R(x)\varphi) f(x) dx.$$

In particular, writing  $\Theta_{f,\pi}^X$  for the orthogonal projection of  $\Theta_f^X$  to  $\pi$ , we see that

$$\mathcal{P}_H|_\pi \neq 0 \iff \Theta_{f,\pi}^X \neq 0.$$

This makes sense for any (smooth)  $G$ -variety  $X$ , not necessarily homogeneous. Put this way, it contains many basic examples.

**Example 5.4.** (a) Take  $X = \mathbb{G}_a \circ \mathbb{G}_m$ . Then an automorphic form is just a Hecke character  $\chi$ , and by Tate,

$$\int_{[\mathrm{GL}_1]}^{\mathrm{reg}} \chi(t) \Theta_f^X(t) dt \sim L\left(\frac{1}{2}, \chi\right).$$

(We'll talk about the regularization later.)

(b) Take  $X = \mathrm{Mat}_n \circ G = \mathrm{GL}_n \times \mathrm{GL}_n$ . Take two cusps forms in the *same* cuspidal automorphic representation, say  $\varphi_1, \varphi_2 \in \pi \subset A_{\mathrm{cusp}}(\mathrm{GL}_n)$ . Then

$$\int_{[G]}^{\mathrm{reg}} \varphi_1(g) \varphi_2(g) \Theta_f^X(g) dg \sim L\left(\frac{1-n}{2}, \pi\right).$$

**5.1.3. RTF formally.** Before defining an RTF, we need to talk about an *automorphic kernel*. Take  $f \in C_c^\infty(G(\mathbb{A}))$ . It acts on the space of automorphic forms by the right convolution  $R(f)$ . This action is given by a kernel, which is like a continuous matrix, given by

$$(R(f)\varphi)(x) := \int_{[G]} K_f(x, y) \varphi(y) dy.$$

The kernel has  $K_f(x, y)$  two different expansions. One, which we call the *geometric* expansion, is as a sum over rational points over the group:

$$\sum_{\gamma \in G(K)} f(x^{-1}\gamma y).$$

The other, which is a kind of *spectral* expansion, is given by

$$\sum_{\substack{\varphi \in A_{\text{cusp}}(G) \\ \text{O.N.B.}}} (R(f)\varphi)(x) \overline{\varphi(y)} + \dots$$

where  $\dots$  denotes the Eisenstein contribution.

We now consider the following setting. Let  $H_1 \subset G \supset H_2$  be two subgroups. Let  $\chi_i : [H_i] \rightarrow \mathbb{C}^\times$ . Define  $\mathfrak{X}$  to be, informally, the quotient

$$"H_1, \chi_1 \backslash G / H_2, \chi_2".$$

More precisely:

**Definition 5.5.** For  $f \in C_c^\infty(G(\mathbb{A}))$ ,

$$\text{RTF}_{\mathfrak{X}}(f) := \int_{[H_1] \times [H_2]} K_f(h_1, h_2) \chi_1(h_1) \chi_2(h_2) dh_1 dh_2.$$

Why is this related to period integrals? Well, one way to compute the RTF is via the *spectral expansion*

$$\text{RTF}_{\mathfrak{X}}(f) := \sum_{\varphi \in A_{\text{cusp}}(G)} \mathcal{P}_{H_1, \chi_1}(R(f)\varphi) \mathcal{P}_{H_2, \chi_2}(\overline{\varphi}) + \dots$$

Another is the *geometric expansion*

$$\text{RTF}_{\mathfrak{X}}(f) = \sum_{\gamma \in H_1(K) \backslash G(K) / H_2(K)} \text{vol}([ (H_1 \times H_2)_\gamma ]) \text{Orb}_\gamma(f),$$

where

$$\text{Orb}_\gamma(f) := \int_{(H_1 \times H_2)_\gamma(\mathbb{A}) \backslash (H_1 \times H_2)(\mathbb{A})} H(h_1^{-1}\gamma h_2) \chi_1(h_1) \chi_2(h_2) dh_1 dh_2.$$

Here  $(H_1 \times H_2)_\gamma$  is the stabilizer with respect to the action  $\gamma \cdot (h_1, h_2) = h_1^{-1}\gamma h_2$ .

**Remark 5.6.** Two sources of divergence on the geometric side:

- the volumes can be infinite, and
- the orbital integrals can diverge.

However, one advantage of these distributions over the original ones is that they are factorizable, being now taken over an adelic space. At least formally,

$$\text{Orb}_\gamma(f) = \prod_v \text{Orb}_\gamma(f_v) \quad \text{for } f = \prod_v f_v,$$

where  $\text{Orb}_\gamma(f_v)$  is a local orbital integral.

5.1.4. *RTF and theta series.* Assume for simplicity that  $\chi_1 = \chi_2 = 1$ .

There is one first obvious relation between RTF and theta series. Note that

$$K_f = \Theta_f^G \quad \text{for } G \circ G \times G.$$

Thus

$$\text{RTF}_{\mathfrak{X}}(f) = \int_{[H_1 \times H_2]} \Theta_f^G = \int_{[H_2]} \Theta_{\varphi}^{H_2 \setminus G},$$

where  $\varphi(x) = \int_{H_1(\mathbb{A})} f(h_1 x) dh_1$ , which lies in  $C_c^\infty(H_1(\mathbb{A}) \setminus G(\mathbb{A}))$ . We can also write the above as the Petersson norm

$$\left\langle \Theta_{\varphi_1}^{H_1 \setminus G}, \Theta_{\varphi_2}^{H_2 \setminus G} \right\rangle_{\text{Pet}},$$

where for  $f = f_1 * f_2^\vee$  with  $f_1, f_2 \in C_c^\infty(G(\mathbb{A}))$ , we set

$$\varphi_i(x) := \int_{H_i(\mathbb{A})} f_i(h_i x) dh_i.$$

More generally, for  $X \circ G$ , we can consider the distributions

$$\text{RTF}_{X/G}(f) := \int_{[G]} \Theta_f^X \quad \text{for } f \in C_c^\infty(X(\mathbb{A})).$$

You can write, in a similar way, a geometric expansion of this in terms of  $G(K)$  on  $X(K)$ . Let's try to make sense of the convergence. These  $\Theta_f^X$  are not arbitrary functions – they have a very special behavior. Here's a fact that we'll write in vague terms, and then give one example.

**Fact 5.7.**  $\Theta_f^X$  is “asymptotically finite”, i.e., is asymptotic to eigenfunctions with respect to the action of some tori (where we pluralize “torus” because the precise one depends upon in which direction you are tending to infinity). Informally, by declaring the integrals of such asymptotics to be zero, you can define regularized integrals. Provided that  $j++\delta$  the asymptotics are nontrivial, there is a canonical way to regularize  $\text{RTF}_{X/G}$ .

**Example 5.8.** Take  $G = \mathbb{G}_m$ . Let  $\chi_0, \chi_\infty : [\mathbb{G}_m] \rightarrow \mathbb{C}^\times$ . Define  $\mathcal{F}_{\chi_0, \chi_\infty}([\mathbb{G}_m])$  to be the set of all  $f \in C^\infty([\mathbb{G}_m])$  such that there are  $c_0 = c_0(f)$  and  $c_\infty = c_\infty(f) \in \mathbb{C}$  such that

$$\begin{aligned} |f(t) - c_\infty \chi_\infty(t)| &\ll_N |t|^{-N} & \text{for } |t| \geq 1, \\ |f(t) - c_0 \chi_0(t)| &\ll_N |t|^N & \text{for } |t| \leq 1. \end{aligned}$$

We have a short exact sequence of  $\mathbb{A}^\times$ -representations

$$0 \rightarrow \mathcal{S}([\mathbb{G}_m]) \rightarrow \mathcal{F}_{\chi_0, \chi_\infty} \rightarrow \mathbb{C}_{\chi_0} \oplus \mathbb{C}_{\chi_\infty} \rightarrow 0,$$

where  $\mathcal{S}([\mathbb{G}_m])$  consists in particular of functions that are rapidly decaying. The map is given by  $f \mapsto (c_0(f), c_\infty(f))$ .

(a) If  $\chi_0 \neq 1 \neq \chi_\infty$ , then there is a unique functional

$$\int_{[\mathbb{G}_m]}^{\text{reg}} : \mathcal{F}_{\chi_0, \chi_\infty} \rightarrow \mathbb{C}$$

that is  $\mathbb{A}^\times$ -invariant and such that

$$\int_{[\mathbb{G}_m]}^{\text{reg}} f = \int_{[\mathbb{G}_m]} f(t) dt \quad \text{for } f \in \mathcal{S}([\mathbb{G}_m]).$$

(b) Take  $X_1 = \mathbb{G}_a \circ \mathbb{G}_m$ . Then, for all  $f \in C_c^\infty(\mathbb{A})$ , we have  $\Theta_f^{X_1} \in \mathcal{F}_{|\cdot|^{-1}, 1}$ .

- (c) Define  $X_2 = \mathbb{G}_a^2 \circ \mathbb{G}_m$  by  $t \cdot (x, y) := (tx, t^{-1}y)$ . Then, for all  $f \in C_c^\infty(\mathbb{A}^2)$ , we have  $\Theta_f^{X_2} := \mathcal{F}_{|\cdot|^{-1}, |\cdot|}$ .
- (d) Take  $\mathbb{G}_m = A = \begin{pmatrix} * & \\ & 1 \end{pmatrix} \subset \mathrm{PGL}_2$ , then for all  $f \in C_c^\infty(X_3(\mathbb{A}))$ , we have  $\Theta_f^{X_3} \in \mathcal{F}_{|\cdot|^{-1}, |\cdot|}$ .

## REFERENCES