

Zeta functions of Siegel threefolds

IAS, April 5, 2001

by Gérard Laumon*

1. Introduction

The zeta functions of modular curves, or more generally of Shimura varieties associated with $GL(2)$, are now well known. What about higher dimensional Shimura varieties of PEL type?

Important progresses have been made:

- the counting of points over finite fields by Kottwitz,
- the trace formula and its stabilization by Arthur,
- the understanding of the boundary and the topological trace formula by Goresky, Harder and MacPherson, and also Franke,
- the proof by Pink of the Deligne conjecture.

But a major ingredient is still missing, the so-called Fundamental Lemma.

Some particular cases have been already treated:

- Kottwitz's simple Shimura varieties which are attached to unitary groups obtained from involutions of the second kind on division algebras over CM fields,
- Picard modular surfaces (unitary groups in 3 variables).

The first case is the easiest as the Shimura varieties are compact and all the endoscopy phenomena mysteriously disappear. The second one is very favorable as the Fundamental lemma is accessible by "hand" and the non trivial Levi or endoscopic groups of $U(3)$ are

* CNRS and Université Paris-Sud

all related to $GL(2)$ and $GL(1)$.

A similar favorable case is the $GSp(4)$ case. In 1992 Taylor obtained the first results by only using the congruence relation. In 1995, assuming all the required fundamental lemmas, I computed the parabolic part of the zeta function of the Siegel threefolds for the trivial local system. Simultaneously, Harder and Weissauer treated the general case by using the topological trace formula of Goresky and MacPherson.

In preparing this talk I checked that my approach works as well for a general local system and this is what I would like to explain today. Let me point out that the actual computation is not difficult and does not require any new idea. In some sense, the main difficulty is to understand all the quite complicated objects which are involved the very general results of Arthur, Harish-Chandra, Kottwitz, ...

2. Siegel threefolds

Let

$$G = GSp(4) = \{x \in GL(4) \mid \exists c(x) \in GL(1) \text{ such that } {}^t x J x = c(x) J\}$$

be the reductive group over \mathbb{Z} of symplectic similitudes of the form whose matrix is

$$J = \begin{pmatrix} & & & 1 \\ & & & \\ & S & & \\ -S & & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

in the standard basis of \mathbb{Z}^4 . Its center is the torus $A_G \cong GL(1)$ of scalar matrices. The kernel of the character $c : G \rightarrow GL(1)$ is the symplectic group $Sp(4)$.

For every integer $N \geq 3$, let \mathcal{S}_N be the modular scheme of quadruples

$$(A, \lambda, \zeta, \eta)$$

where A is an abelian surface, $\lambda : A \xrightarrow{\sim} \widehat{A}$ is a principal polarization ($\widehat{\lambda} = \lambda$), ζ is a primitive N -th root of unity and $\eta : (\mathbb{Z}/N\mathbb{Z})^4 \cong A[N]$ is a group scheme isomorphism which is symplectic for the Weil pairing $A[N] \times A[N] \rightarrow \mu_N \cong \mathbb{Z}/N\mathbb{Z}$ induced by λ and the choice of ζ .

It is well known that \mathcal{S}_N is a smooth, quasi-projective scheme over $\mathbb{Z}[1/N]$, of pure relative dimension 3 and that

$$\mathcal{S}_N(\mathbb{C}) = G(\mathbb{Q}) \backslash [X_\infty \times (G(\mathbb{A}_f)/K(N))]$$

where $X_\infty = \mathcal{H}_+ \cup \mathcal{H}_-$ is the union of the Siegel upper half space

$$\mathcal{H}_+ = \{\Omega \in \mathfrak{gl}(2, \mathbb{C}) \mid {}^t(S\Omega) = S\Omega \text{ and } \text{Im}(S\Omega) \gg 0\}$$

and the Siegel lower half space $\mathcal{H}_- = -\mathcal{H}_+$, and where $K(N) \subset K = G(\widehat{\mathbb{Z}})$ is the principal congruence subgroup of level N . It is also well known that

$$X_\infty = G(\mathbb{R})/K'_\infty$$

where $K'_\infty = A_G(\mathbb{R})^0 K_\infty$ and

$$K_\infty = \{x \in G(\mathbb{R}) \mid {}^t x x = 1\}$$

is a maximal compact subgroup of $\text{Sp}(4, \mathbb{R}) \subset \text{GSp}(4, \mathbb{R})$.

Let us fix some prime number ℓ . Any finite dimensional algebraic representation V of G over \mathbb{Q} defines an ℓ -adic local system \mathcal{V}_ℓ^\vee on $\mathcal{S}_N[1/\ell]$ whose restriction to $\mathcal{S}_N(\mathbb{C})^{\text{an}}$ is given by

$$\begin{array}{c} G(\mathbb{Q}) \backslash ([X_\infty \times (G(\mathbb{A}_f)/K(N))] \times (\mathbb{Q}_\ell \otimes V^\vee) \\ \downarrow \\ G(\mathbb{Q}) \backslash [X_\infty \times (G(\mathbb{A}_f)/K(N))] \end{array}$$

For example, if V is the standard four dimensional representation of $G_{\mathbb{Q}}$, the fiber of $\mathcal{V}_{\ell}^{\vee}$ at a geometric point $(A, \lambda, \zeta, \eta)$ of \mathcal{S}_N is $H^1(A, \mathbb{Q}_{\ell})$.

The goal is to compute the L function

$$L(\mathcal{S}_N, \mathcal{V}_{\ell}^{\vee}; s) = \prod_p L_p(\mathcal{S}_N, \mathcal{V}_{\ell}^{\vee}; s)$$

of the ℓ -adic cohomology of $\overline{\mathbb{Q}} \otimes \mathcal{S}_N$ with values in $\mathcal{V}_{\ell}^{\vee}$.

As \mathcal{S}_N is not proper over $\mathbb{Z}[1/N]$ there are several choices for the cohomology. In this talk I will only consider the cohomology with compact supports. (Another natural choice is the intersection cohomology of the Satake compactification.) I will also restrict myself to the computation of the local factors for the prime numbers p which does not divide ℓN . Thanks to the existence of toroidal compactifications of \mathcal{S}_N over $\mathbb{Z}[1/N]$, these are exactly the prime numbers where the ℓ -adic cohomology is unramified.

Obviously, one may assume that V is irreducible. Then $V = V_{\mu}$ has a dominant highest weight $\mu \in X^*(T)^+$ with respect to the Borel subgroup $B = TU \subset G$ of upper triangular matrices which contains the maximal torus of diagonal matrices

$$T = \{\text{diag}(t_1, t_2, t_3, t_4) \mid t_1 t_4 = t_2 t_3 = c(t)\}.$$

We may identify $X^*(T)$ to the set of triples $\mu = \mu_0 \oplus (\mu_1, \mu_2)$ of integers such that $\mu_1 + \mu_2 \equiv \mu_0 \pmod{2}$ by

$$\mu : \text{diag}(t_1, t_2, t_3, t_4) \mapsto t_1^{\mu_1} t_2^{\mu_2} c(t)^{\frac{\mu_0 - \mu_1 - \mu_2}{2}}.$$

Then the cone of dominant weights is equal to

$$X^*(T)^+ = \{\mu \in X^*(T) \mid \mu_1 \geq \mu_2 \geq 0\}.$$

The central character of the irreducible algebraic representation V_{μ} of highest weight μ is $a \mapsto a^{\mu_0}$.

In the rest of the talk, a dominant weight μ is fixed. For simplicity, I assume that μ is *regular*, i.e. that $\mu_1 > \mu_2 > 0$.

3. The main result

Let p be a prime number not dividing ℓN and let $\Phi_p \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be any Frobenius element at p . We have

$$L_p(\mathcal{S}_N, \mathcal{V}_{\mu, \ell}^\vee; s) = \exp \left(\sum_{j \geq 1} \frac{p^{-js}}{j} \text{Lef}(\Phi_p^j) \right)$$

where

$$\text{Lef}(\Phi_p^j) = \sum_i (-1)^i \text{tr}(\Phi_p^j, H_c^i(\overline{\mathbb{Q}} \otimes \mathcal{S}_N, \mathcal{V}_{\mu, \ell}^\vee))$$

for every positive integer j .

The pair $(\mathcal{S}_N, \mathcal{V}_{\mu, \ell}^\vee)$ is equipped of an action of the Hecke algebra

$$C_c(G(\mathbb{A}_f) // K(N), \mathbb{Q})$$

of \mathbb{Q} -valued locally constant functions with compact support on $G(\mathbb{A}_f)$ which are left and right invariant under $K(N)$. Therefore, for every positive integer j and every $f \in C_c(G(\mathbb{A}_f) // K(N), \mathbb{Q})$ which splits into $f = f^p f_p$ where $f_p = 1_{K_p}$ is the characteristic function of the hyperspecial maximal compact subgroup $K_p = G(\mathbb{Z}_p)$ of $G(\mathbb{Q}_p)$, we may also considered the traces

$$\text{Lef}(f \times \Phi_p^j) = \sum_i (-1)^i \text{tr}(f \times \Phi_p^j, H_c^i(\overline{\mathbb{Q}} \otimes \mathcal{S}_N, \mathcal{V}_{\mu, \ell}^\vee)).$$

What we really want is a formula for theses traces in terms of automorphic representations of G and of related groups.

There are two discrete series representations π_μ^W and π_μ^H of $G(\mathbb{R})$ having the same central and infinitesimal characters as the finite

dimensional representation V_μ . They form a Langlands packet. Any irreducible automorphic representation π of $G(\mathbb{A})$ which admits π_μ^W or π_μ^H as archimedean component is cuspidal and occurs with finite multiplicity $m_{\text{disc}}^G(\pi)$ in the representation of $G(\mathbb{A})$ by right translation on $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}), \mu_0)$. Here μ_0 means that we only consider functions $\varphi : G(\mathbb{Q})\backslash G(\mathbb{A}) \rightarrow \mathbb{C}$ such that $\varphi(ag) = a^{\mu_0}\varphi(g)$ for every $a \in A_G(\mathbb{R})$.

Any irreducible automorphic representation π of $G(\mathbb{A})$ is unramified at almost every prime p and its local component π_p at such a prime is completely determined (up to isomorphisms) by its Langlands parameter $t_{\pi_p} \in \widehat{T}/W$. Here \widehat{T} is a maximal torus of diagonal matrices in the Langlands dual

$$\widehat{G} = \text{GSp}(4, \mathbb{C}) \cong \text{GSpin}(5, \mathbb{C})$$

of G and W is the Weyl group of (G, T) , or equivalently of $(\widehat{G}, \widehat{T})$. The Frobenius-Hecke eigenvalues

$$(z_1(\pi_p), z_2(\pi_p), z_3(\pi_p), z_4(\pi_p))$$

of π_p are the diagonal entries of t_{π_p} . They satisfy the relation $z_1(\pi_p)z_4(\pi_p) = z_2(\pi_p)z_3(\pi_p)$, and they are well defined up to a permutation in $W \subset \mathfrak{S}_4$.

Up to equivalence, there is only one elliptic endoscopic group H for G . Its Langlands dual \widehat{H} is the centralizer of $\text{diag}(1, -1, -1, 1)$ in $\widehat{G} = \text{GSp}(4, \mathbb{C})$ and we have

$$H = [\text{GL}(2) \times \text{GL}(2)]/\text{GL}(1)$$

(central antidiagonal embedding).

For each positive integer n let σ_n be the unique discrete series representation of $\text{GL}(2, \mathbb{R})$ which has the same central and infinitesimal characters as the finite dimensional representation $\text{Sym}^{n-1}(\mathbb{C}^2)$. The twist $\sigma_n(\lambda)$ of σ_n by a complex number λ is defined by

$$\sigma_n(\lambda) = |\det|^\lambda \sigma_n.$$

The central character of $\sigma_n(\lambda)$ is equal to $a \mapsto a^{n-1}|a|^{2\lambda}$. In particular, if $n \equiv \mu_0 \pmod{2}$, the central character of $\sigma_n(\frac{\mu_0-n+1}{2})$ is equal to $a \mapsto a^{\mu_0}$.

Any irreducible automorphic representation σ of $\mathrm{GL}(2, \mathbb{A})$, which admits $\sigma_n(\frac{\mu_0-n+1}{2})$ as archimedean component for some integer $n \equiv \mu_0 \pmod{2}$, is cuspidal and occurs with multiplicity 1 in $L^2(\mathrm{GL}(2, \mathbb{Q}) \backslash \mathrm{GL}(2, \mathbb{A}), \mu_0)$.

Any irreducible automorphic representation σ of $\mathrm{GL}(2, \mathbb{A})$ is unramified at almost every prime p and its local component σ_p at such a prime is completely determined (up to isomorphisms) by its unordered pair $(z'(\sigma_p), z''(\sigma_p))$ of Frobenius-Hecke eigenvalues.

THÉORÈME. — We assume all the required fundamental lemmas for our group G . In particular we assume the ordinary transfer of a function $f^p \in C_c(G(\mathbb{A}_f^p))$ to a function $h^p \in C_c(H(\mathbb{A}_f^p))$ for a suitable normalization of the transfer factors.

Then, for every $f = f^p 1_{K_p} \in C_c(G(\mathbb{A}_f) // K(N))$ and every integer j , the trace $\mathrm{Lef}(f \times \Phi_p^j)$ is equal to the sum of a main term

$$\begin{aligned} & -\frac{1}{2} \sum_{\pi} m_{\mathrm{disc}}^G(\pi) \mathrm{tr} \pi_f^p(f^p) p^{\frac{3j}{2}} (z_1(\pi_p)^j + z_2(\pi_p)^j + z_3(\pi_p)^j + z_4(\pi_p^j)) \\ & + \frac{1}{4} \sum_{\rho_1, \rho_2} \mathrm{tr}(\rho_{1,f}^p \times \rho_{2,f}^p)(h^p + h^p \circ \iota) \\ & \quad \times p^{\frac{3j}{2}} (z'(\rho_{1,p})^j + z''(\rho_{1,p})^j - z'(\rho_{2,p})^j - z''(\rho_{2,p}^j)) \end{aligned}$$

where

- π runs through a system of representatives of equivalence classes of cuspidal automorphic irreducible representations of $G(\mathbb{A})$ such that

$$\pi_{\infty} \cong \pi_{\mu}^{\mathrm{W}} \text{ or } \pi_{\mu}^{\mathrm{H}},$$

- ρ_1, ρ_2 runs through a system of representatives of equivalence classes of cuspidal automorphic irreducible representations of

$\mathrm{GL}(2, \mathbb{A})$ such that ρ_1 and ρ_2 have the same central character $\omega_{\rho_1} = \omega_{\rho_2}$ and that

$$\rho_{1, \infty} \cong \sigma_{\mu_1 + \mu_2 + 3} \left(\frac{\mu_0 - \mu_1 - \mu_2 - 2}{2} \right)$$

and

$$\rho_{2, \infty} \cong \sigma_{\mu_1 - \mu_2 + 1} \left(\frac{\mu_0 - \mu_1 + \mu_2}{2} \right),$$

- ι is the automorphism of H which permutes the two $\mathrm{GL}(2)$ factors,

and of an explicit complementary term which is entirely expressible in terms of cuspidal automorphic irreducible representations of $\mathrm{GL}(1, \mathbb{A})$ and $\mathrm{GL}(2, \mathbb{A})$.

Moreover there exists an explicit virtual ℓ -adic representation W_ℓ of the group algebra $C_c(G(\mathbb{A}_f) // K(N), \mathbb{Q}) [\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})]$ such that, for every $f = f^p 1_{K_p} \in C_c(G(\mathbb{A}_f) // K(N), \mathbb{Q})$ and every integer j , this complementary term is equal to

$$\mathrm{tr}(f \times \Phi_p^j, W_\ell).$$

4. Deligne conjecture

In order to compute $\mathrm{Lef}(f \times \Phi_p^j)$ we first need to apply the Grothendieck-Lefschetz fixed point formula. For a general Hecke operator $f = f^p 1_{K_p}$, this is not directly possible as there may be hidden contributions of fixed points at infinity. Fortunately, the things become much simpler if we take the integer j *large enough with respect to f* as it was conjectured by Deligne and it has been proved by Pink (and also Fujiwara).

Let $\mathrm{Fix}(f \times \Phi_p^j)$ be the fixed point set of the correspondence $f \times \Phi_p^j$ on $\mathbb{F}_p \otimes \mathcal{S}_N$. If j is large enough with respect to f , this set is finite and each fixed point occurs with multiplicity 1. For each point in

$\text{Fix}(f \times \Phi_p^j)$, the correspondence induces an endomorphism of the stalk of the local system $\mathcal{V}_{\mu,\ell}^\vee$ at that point. The *naive local term* is by definition the trace of this endomorphism.

THÉORÈME (Pink). — *If j is large enough with respect to f , $\text{Lef}(f \times \Phi_p^j)$ is precisely the sum over $\text{Fix}(f \times \Phi_p^j)$ of the naive local terms.*

5. Kottwitz's results

Kottwitz has given the following group theoretical expression for the sum over $\text{Fix}(f \times \Phi_p^j)$ of the naive local terms

$$\sum_{\gamma_0} \sum_{\substack{(\gamma, \delta) \\ \alpha(\gamma_0; \gamma, \delta) = 1}} c(\gamma_0; \gamma, \delta) \text{Tr}(\gamma_0, V_\mu) O_\gamma^G(f^p) \text{TO}_\delta^G(\varphi_j).$$

Here

$$\varphi_j = 1_{K_{p^j} \text{diag}(p, p, 1, 1) K_{p^j}}$$

is the characteristic function of $K_{p^j} \text{diag}(p, p, 1, 1) K_{p^j} \subset G(\mathbb{Q}_{p^j})$ where \mathbb{Q}_{p^j} is the unramified extension of degree j of \mathbb{Q}_p , \mathbb{Z}_{p^j} is its ring of integers and $K_{p^j} = G(\mathbb{Z}_{p^j})$.

Assuming fundamental lemmas he has also stabilized the above expression and he has proved that it is equal to

$$T_e^G(f^G) - \frac{1}{4} T_e^H(f^H) = \sum_{\gamma} \tau(G_\gamma) O_\gamma^G(f^G) - \frac{1}{4} \sum_{\delta} \tau(H_\delta) O_\delta^H(f^H)$$

where γ and δ run through systems of representatives of the elliptic semi-simple conjugacy classes in $G(\mathbb{Q})$ and $H(\mathbb{Q})$, $\tau(G_\gamma)$ and $\tau(H_\delta)$ are volumes and $O_\gamma(f^G)$ and $O_\delta^H(f^H)$ are the orbital integrals at γ and δ of some smooth functions

$$f^G = f_\infty^G f^p b_j^G(\varphi_j) \text{ and } f^H = f_\infty^H h^p b_j^H(\varphi_j)$$

on $G(\mathbb{A})$ and $H(\mathbb{A})$. Here

$$f_\infty^G = -\frac{1}{2}(f_{\pi_\mu^W} + f_{\pi_\mu^H})$$

where $f_{\pi_\mu^W}$ and $f_{\pi_\mu^H}$ are pseudo-coefficients of the discrete series representations π_μ^W and π_μ^H of $G(\mathbb{R})$,

$$\begin{aligned} f_\infty^H = & -f_{\mu_1+\mu_2+3}\left(\frac{\mu_0-\mu_1-\mu_2-2}{2}\right) \times f_{\mu_1-\mu_2+1}\left(\frac{\mu_0-\mu_1+\mu_2}{2}\right) \\ & + f_{\mu_1-\mu_2+1}\left(\frac{\mu_0-\mu_1+\mu_2}{2}\right) \times f_{\mu_1+\mu_2+3}\left(\frac{\mu_0-\mu_1-\mu_2-2}{2}\right) \end{aligned}$$

where $f_n(\lambda)$ is a pseudo-coefficient of the discrete series representation $\sigma_n(\lambda)$ of $\mathrm{GL}(2, \mathbb{R})$, and

$$b_j^G : C_c(G(\mathbb{Q}_{p^j})//K_{p^j}) \rightarrow C_c(G(\mathbb{Q}_p)//K_p)$$

and

$$b_j^H : C_c(G(\mathbb{Q}_{p^j})//K_{p^j}) \rightarrow C_c(H(\mathbb{Q}_p)//K_p^H)$$

are base change homomorphisms between unramified Hecke algebras ($K_p^H = [\mathrm{GL}(2, \mathbb{Z}_p) \times \mathrm{GL}(2, \mathbb{Z}_p)]/\mathbb{Z}_p^\times$).

6. Arthur-Selberg trace formula

We first work with the group G .

The non-invariant Arthur trace formula for G is a relation

$$J_{\mathrm{geom}}^G(f^G) = J_{\mathrm{spec}}^G(f^G).$$

The geometric side $J_{\mathrm{geom}}^G(f^G)$ is the sum of a main term $T_e^G(f^G)$ and of complementary terms which are sums of weighted orbital integrals of f^G . The spectral side $J_{\mathrm{spec}}^G(f^G)$ is also the sum of a main term

$$\sum_{\pi} m_{\mathrm{disc}}(\pi) \mathrm{tr} \pi(f^G)$$

where π runs through a system of representatives of the isomorphism classes of irreducible representations of $G(\mathbb{A})$ which occur discretely in $L^2(A_G(\mathbb{R})^0 G(\mathbb{Q}) \backslash G(\mathbb{A}))$, and of complementary terms which are expressible in terms of weighted characters.

Arthur trace formula is quite complicated and before to use it we need to simplify it. We will do it by adapting to our problem a trick which was found by Flicker and Kazhdan and which has some similarity with the Deligne conjecture.

First we choose f_∞^G *very cuspidal* (a non-invariant property which reinforces Arthur's cuspidality). Then we pick an auxiliary prime number $q \neq p$ such that $f = 1_{K_q} f^q$ and we replace our Hecke operator f by $\tilde{f} = f_q f^q$ where $f_q \in C_c(G(\mathbb{Q}_q) // K_q)$ is a varying Hecke function which is general enough with respect to f^q . Finally we assume that j is large enough with respect to \tilde{f} .

Then all the weighted orbital integrals disappear and the spectral side simplifies drastically.

We may do similar computations for H . As a consequence we get a computable spectral expression for $\text{Lef}(\tilde{f} \times \Phi_p^j)$ which a priori holds only for general enough f_q and large enough j . But then it also holds for arbitrary f_q and arbitrary integer j .

7. A first consequence

Let $\pi \in G(\mathbb{A})$. Following Piatetski-Shapiro we say that an irreducible automorphic representation π almost comes from a Levi subgroup $M \supset T$ of G if there exists an irreducible automorphic representation σ de $M(\mathbb{A})$ such that, for almost all prime number p where π_p and σ_p are unramified, the Langlands parameters $t_{\pi_p} \in \hat{T}/W$ and $t_{\sigma_p} \in \hat{T}/W^M$ match. In this definition we may replace M by the endoscopic group H .

THÉORÈME. — *Let π be an irreducible cuspidal automorphic representation of $G(\mathbb{A})$ whose archimedean component is isomorphic*

either to π_μ^{W} or π_μ^{H} . Let us assume that π almost comes neither from a proper Levi subgroup $T \subset M \subsetneq G$ nor from the endoscopic group H . Then there exist a finite set S of prime numbers, a number field E and, for each non archimedean place λ of E of characteristic ℓ , a semi-simple λ -adic representation $W_\lambda(\pi)$, which is pure of weight $\mu_0 + 3$, such that

$$\begin{aligned} \text{tr}(\Phi_p^j, W_\lambda(\pi)) = & \left[\frac{m_{\text{disc}}^G(\pi_\mu^{\text{W}} \otimes \pi_{\text{f}}) + m_{\text{disc}}^G(\pi_\mu^{\text{H}} \otimes \pi_{\text{f}})}{2} \right] \\ & \times p^{\frac{3j}{2}} (z_1(\pi_p)^j + z_2(\pi_p)^j + z_3(\pi_p)^j + z_4(\pi_p)^j) \end{aligned}$$

for every prime number $p \in \{\ell\} \cup S$ and every integer j .