LOCAL ROOT NUMBERS AND EPSILON DICHOTOMY

GYUJIN OH

Here we will mainly talk about local theta correspondence.

The Howe duality principle is very neat and clean but also too abstract. Perhaps the first question beyond the Howe duality is the following: given a dual reductive pair (G, G') in $\mathrm{Sp}(W)$, when is the local theta lift $\Theta(\pi)$ nonzero? This depends not only on π, but also on G', because one can embed G into many different metaplectic groups. We will try to explain that often it is the **local root number** that characterizes nonvanishing of theta lift.

From now on F is a characteristic 0 local field of residue characteristic $\neq 2$. We use $\mathrm{Mp}(W)$ to denote both $\mathbb{C}^\times$-cover and μ_2-cover of $\mathrm{Sp}(W)$ to reduce notation. For an irreducible admissible representation π of some group, $\tilde{\pi}$ means the contragredient of π.

1. What do we need?

Let us get straight into the problem. A quick recap of local theta lift is as follows.

Theorem 1.1 (Local Howe duality principle)**.** *Let (G, G') be a dual reductive pair in $\mathrm{Sp}(W)$. Let $\widetilde{G}, \widetilde{G}'$ be the inverse images of G, G' in $\mathrm{Mp}(W)$, and let $j : \widetilde{G} \times \widetilde{G}' \longrightarrow \mathrm{Mp}(W)$ be the natural homomorphism. Let ω_ψ be the Weil representation of $\mathrm{Mp}(W)$ and, for π an irreducible admissible representation of $\widetilde{G}$, let $S(\pi)$ be the π-isotypic quotient of $j^*\omega$. Then $S(\pi) = \pi \boxtimes \Theta(\pi)$ for a $\widetilde{G}'$-representation $\Theta(\pi)$, and this is either nonzero or an admissible representation of $\widetilde{G}'$ of finite length, which has a unique irreducible quotient $\theta(\pi)$ called the **local theta lift**. The local theta lift give an injective map from the set of irreducible admissible representations of $\widetilde{G}$ having nonzero theta lift to $\widetilde{G}'$ to the set of irreducible admissible representations of $\widetilde{G}'$.*

Remark 1.1. As both central $\mathbb{C}^\times$'s in $\widetilde{G}$ and $\widetilde{G}'$ act by the identity character, we necessarily need π to have the identity as its central character.

Thus we need to understand $j^*\omega$ better to say something more. We do this by using our favorite reductive dual pair, $(\mathrm{Sp}(W), \mathrm{O}(V))$ inside $\mathrm{Sp}(W \otimes V)$ (we will however eventually switch to unitary groups, which is more convenient to work with, after this motivational section is finished).

We first address some confusing aspect of Weil representation. Sometimes Weil representation is defined as a representation of $\mathrm{Mp}(W) \times \mathrm{O}(V)$ on $\mathcal{S}(V^n)$ where $\dim W = 2n$, $\mathrm{O}(V)$ acts by how it acts on V^n and $\mathrm{Mp}(W)$ acts with formula involving Fourier transform (Schrödinger model for the "degenerate Weil reprsentation"; or sometimes people call this as the Weil representation and the morally-correct $\mathrm{Mp}(W \otimes V)$-representation as the oscillator representation). But this is really just the restriction of the usual Weil representation of $\mathrm{Mp}(W \otimes V)$ defined as lifting of projective representation of $\mathrm{Sp}(W \otimes V)$ on the Heisenberg representation to $\mathrm{Mp}(W) \times \mathrm{O}(V)$. Abstractly this is it, but when you try to compute with it, the problem becomes more subtle: $\mathrm{Sp}(W), \mathrm{O}(V)$ are subgroups of $\mathrm{Sp}(W \otimes V)$ so it is not clear how to "restrict" a representation of $\mathrm{Mp}(W \otimes V)$ to these. The problem is twofold:

- $\mathrm{Mp}(W \otimes V) \longrightarrow \mathrm{Sp}(W \otimes V)$ **splits** over $\mathrm{O}(V) \subset \mathrm{Sp}(W \otimes V)$, but **noncanonically**; you can twist by any character of $\mathrm{O}(V)$.
- $\mathrm{Mp}(W \otimes V) \longrightarrow \mathrm{Sp}(W \otimes V)$ in general **does not split over** $\mathrm{Sp}(W)$ (more precisely, splits if and only if $\dim V$ is even). But there is a **canonical way** to choose a lifting $\mathrm{Mp}(W) \hookrightarrow \mathrm{Mp}(W \otimes V)$ of $\mathrm{Sp}(W) \hookrightarrow \mathrm{Sp}(W \otimes V)$. In particular the $\mathrm{O}(V)$ action as above is a result of a noncanonical choice.

Remark 1.2. This is actually the **only case** of dual reductive pairs where the metaplectic group might not split over one of G or G', so one can indeed make sense of restriction of Weil representation in other cases, after making choices.

In any case, we write down the well-accepted Schrodinger model for this degenerate Weil representation to say more:

$$(\omega_{\psi,V}(m(a), z)\varphi)(x) = \chi_V^{\psi}(\det(a), z)|\det(a)|^{m/2}\varphi(xa),$$

for $a \in \mathrm{GL}_n(F)$, $m : \mathrm{GL}_n(F) \xrightarrow{\sim} M_P$, the Levi of the Siegel parabolic P, and

$$\chi_V^{\psi}(x, z) = \chi_V(x)\begin{cases} ☻ z & m \text{ odd} \\ 1 & m \text{ even} \end{cases},$$

with

$$\chi_V(x) = (x, (-1)^{m(m-1)/2}\det(V))_F$$

(one might want to call $\mathrm{disc}(V)$ instead of $\det(V)$),

$$(\omega_{\psi,V}(n(b), 1)\varphi)(x) = \psi(\frac{1}{2}\operatorname{tr}((x, x)b))\varphi(x),$$

for $b \in \mathrm{Sym}_n F$, $n : \mathrm{Sym}_n(F) \xrightarrow{\sim} N_P$, the unipotent radical of P, and

$$(\omega_{\psi,V}(w, 1)\varphi)(x) = ☻ \int_{V^n} \psi(-\operatorname{tr}((x, y)))\varphi(y)dy,$$

for $w = \begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$. In the above formulae ☻ means some 8th root of 1 (Weil index).

If we look at the formulae, we can notice that, apart from change of evaluation points (i.e. x goes to xa), the Siegel parabolic $P = M_P N_P$ acts exactly as the principal series $I_P^{\mathrm{Mp}(W)}(\chi_V^{\psi}|\det|^{\frac{m}{2}-\frac{n+1}{2}})$, where $I_P^{\mathrm{Mp}(W)}$ is the normalized induction. Thus, the map

$$\mathcal{S}(V^n) \xrightarrow{\lambda_V} I_P^{\mathrm{Mp}(W)}(\chi_V^{\psi}|\det|^{\frac{m}{2}-\frac{n+1}{2}}),$$

defined by

$$\lambda_V(\varphi)(g) = (\omega_{\psi,V}(g)\varphi)(0),$$

is $\mathrm{O}(V) \times \mathrm{Mp}(W)$-equivariant, where we give a trivial $\mathrm{O}(V)$-action on the RHS. Thus, λ_V factors through the $\mathrm{O}(V)$-coinvariant quotient of $\mathcal{S}(V^n)$, denoted as $R_n(V)$ (this should be thought as the theta lift of the trivial representation of V, $\Theta(\mathbb{1}_V, W)$). We have the following basic theorem.

Theorem 1.2 (Rallis)**.** *$\lambda_V : R_n(V) \longrightarrow I_P^{\mathrm{Mp}(W)}(\chi_V^{\psi}|\det|^{\frac{m}{2}-\frac{n+1}{2}})$ is injective.*

Proof. The proof is quite complicated. The case of $(\mathrm{Sp}(W), \mathrm{O}(V))$ is proved as Theorem II.2.1 in [R]. The general case for type I pairs is proved as Theorem 7 of Chapter 3.IV in [MVW]. □

Although we omitted the proof, this is certainly believable (what else could be in the kernel?!). Now here comes a real upshot:

Proposition 1.1. *For π a genuine irreducible admissible representation of $\mathrm{Mp}(W)$ if $\dim V$ is odd and an irreducible admissible representation of $\mathrm{Sp}(W)$ if $\dim V$ is even, we have*

$$\Theta(\pi, V) \neq 0 \iff \mathrm{Hom}_{G\times G}(R_{2n}(V), \pi \otimes \pi) \neq 0,$$

where $G = \mathrm{Mp}(W)$ or $\mathrm{Sp}(W)$.

Proof. Let's just think G as $\mathrm{Mp}(W)$ because it's annoying to write down $\mathrm{Mp}(W)$ or $\mathrm{Sp}(W)$ every time. We first explain what we mean by $R_{2n}(V)$ as a representation of $G \times G$. Note that for any two symplectic spaces W_1, W_2 of dimension $2n_1, 2n_2$, there is a canonical embedding

$$j : \mathrm{Mp}(W_1) \times \mathrm{Mp}(W_2) \hookrightarrow \mathrm{Mp}(W_1 \oplus W_2),$$

by "multiplying cocycles". Then it is easy to show that

$$j^*\omega_{\psi,W_1\oplus W_2} = \omega_{\psi,W_1} \otimes \omega_{\psi,W_2}.$$

So $R_{2n}(V)$ can be in particular seen as a representation of $G \times G$.

Now we apply this to theta lifting. Let $\mathbb{W}_i = W_i \otimes V$. If $\tau \in \mathrm{Irr}(\mathrm{O}(V))$ and if π_1 and π_2 are nonzero irreducible quotients of $\Theta(\tau, W_1), \Theta(\tau, W_2)$, respectively, then there are surjective equivariant maps

$$\omega_{\psi,\mathbb{W}_i} \longrightarrow \tau \otimes \pi_i,$$

which gives

$$j^*\omega_{\psi,\mathbb{W}_1\oplus\mathbb{W}_2} \longrightarrow \tau \otimes \pi_1 \otimes \tau \otimes \pi_2,$$

equivariant under $\mathrm{O}(V) \times \mathrm{Mp}(W_1) \times \mathrm{O}(V) \times \mathrm{Mp}(W_2)$-action. As every $\tau \in \mathrm{Irr}(\mathrm{O}(V))$ is self-dual, $\tau \cong \tilde{\tau}$, so we obtain a surjective map

$$j^*\omega_{\psi,\mathbb{W}_1\oplus\mathbb{W}_2} \longrightarrow \pi_1 \otimes \pi_2 \otimes \tau \otimes \tilde{\tau} \longrightarrow \pi_1 \otimes \pi_2 \otimes \mathbb{1}_V,$$

equivariant for $\mathrm{Mp}(W_1) \times \mathrm{Mp}(W_2) \times \mathrm{O}(V)$. As the $\mathrm{O}(V)$-action on the RHS is trivial, it factors through $\mathrm{O}(V)$-coinvariant quotient of degenerate Weil representation which is exactly $R_{n_1+n_2}(V)$. So we proved

$$(*) \cdots\cdots \ \tau \text{ has nonzero theta lifts to } W_1, W_2 \Longrightarrow \mathrm{Hom}_{\mathrm{Mp}(W_1)\times\mathrm{Mp}(W_2)}(R_{n_1+n_2}(V), \pi_1 \otimes \pi_2) \neq 0.$$

Now applying this to our situation, if π has a nonzero theta lift τ, then we can apply (*) to τ to give one side of Proposition.

For the converse, suppose $\mathrm{Hom}_{G\times G}(R_{2n}(V), \pi \otimes \pi) \neq 0$. Then, this gives a nonzero quotient map $\lambda : \omega_{\psi,\mathrm{W}} \otimes \omega_{\psi,\mathrm{W}} \longrightarrow \pi \otimes \pi \otimes \mathbb{1}_V$. Choose $f_1, f_2 \in \mathcal{S}(V^n)$ such that $\lambda(f_1 \otimes f_2) \neq 0$, and choose $\xi^\vee \in \pi^\vee$ such that $(1 \otimes \xi^\vee)(\lambda(f_1 \otimes f_2)) \neq 0$. Then

$$\omega_{\psi,\mathrm{W}} \longrightarrow \pi,$$

$$f \longmapsto (1 \otimes \xi^\vee)(\lambda(f \otimes f_2)),$$

is G-equivariant and is nonzero (at f_1). This shows that $\Theta(\pi, V) \neq 0$. □

Remark 1.3. Here we secretly used the Howe duality principle (i.e. sloppy about Θ and its irreducible quotient). A little more care will give a proof independent of Howe duality principle.

Recall that a quadratic space over a local field is determined by three invariants:

- dimension;
- discriminant = determinant, or the quadratic character associated to it (χ_V in our notation);
- and the Hasse invariant, which is valued in ± 1.

In particular, given a dimension and a quadratic character, there are two possible choices for an orthogonal space having the chosen dimension and quadratic character. We denote them as V^+ and V^-. Then, both $R_n(V^+)$ and $R_n(V^-)$ live in the same principal series, $I_P^{\mathrm{Mp}(W)}(\chi_V^\psi|\det|^{\frac{m}{2}-\frac{n+1}{2}})$, because the data used in the principal series only depend on dimension and quadratic character. We have the two general principles which lead to the theta dichotomy principle.

(DP1) Note that our principal series has exponent $\frac{m}{2} - \frac{n+1}{2}$ for the determinant character. Let $I_n(s, \chi_V^\psi) := I_P^{\mathrm{Mp}(W)}(\chi_V^\psi|\det|^s)$, where $n = \dim W$. Then, whenever two R_{2n}'s live in $I(0, \psi)$ (in this case $m = 2n+1$), each R_{2n} is irreducible, and $I_{2n}(0, \chi_V^\psi) = R_{2n}(V^+) \oplus R_{2n}(V^-)$.

(DP2) For all but finitely many s, $\dim \mathrm{Hom}_{G\times G}(I_{2n}(s, \chi), \pi \otimes \pi) = 1$, and $s = 0$ is ok.

The above two principles imply the **theta dichotomy** principle.

Principle 1.1 (Theta dichotomy). *Given $\pi \in \mathrm{Irr}(G)$, exactly one of $\Theta(\pi, V^+)$ and $\Theta(\pi, V^-)$ is nonzero.*

Proof. By (DP1) and (DP2), $\dim \mathrm{Hom}_{G\times G}(R_{2n}(V^+), \pi \otimes \pi) + \dim \mathrm{Hom}_{G\times G}(R_{2n}(V^-), \pi \otimes \pi) = 1$. So exactly one of the two Hom's is nonzero. □

We will see some reasons behind (DP1) and (DP2). Along the way, we will see that the local root number of π determines whether the nonzero map is supported on $R_{2n}(V^+)$ or $R_{2n}(V^-)$. This refined phenomenon is called the **epsilon dichotomy**.

Principle 1.2 (Epsilon dichotomy)**.** *Given the dimension and the discriminant, the V over which $\Theta(\pi, V) \neq 0$ is determined by the local root number (=ϵ-factor) of π.*

2. Local root numbers

From now on, we switch our discussion to unitary groups. This is because there are many complications in dealing with symplectic or orthogonal groups. For example, the picture of dual reductive pair is extremely asymmetric in symplectic-orthogonal case.

We will try to briefly recall the definition of local L and ϵ factors in terms of doubling method due to Piatetski-Shapiro–Rallis. We use W for a skew-Hermitian space over E/F, and V for a Hermitian space over E/F. Let W_- be the space W with the skew-Hermitian form $-\langle,\rangle_W$. Then $W \oplus W_-$ has a complete polarization $W_{-\Delta} \oplus W_\Delta$ where $W_{-\Delta}$ is the graph of minus the identity and W_Δ is the diagonal W. Then $\mathrm{U}(W + W_-) \cong \mathrm{U}(n, n)$, and we can take the Siegel parabolic P_Δ, the stabilizer of W_Δ, which has Levi $M_\Delta \cong \mathrm{GL}_n(E)$ and unipotent radical $N_\Delta \cong \mathrm{Herm}_n(K)$, via restriction to $X \cong E^n$. Let

$$I_n(s, \chi) = \{\Phi(nm(a)g, s) = \chi(\det a)|\det a|_E^{s+\frac{n}{2}}\Phi(g, s)\},$$

for any character χ.

To motivate the definitio of local L-factors, we first explain how Piatetski-Shapiro–Rallis defined integral representation of automorphic L-function for cuspidal automorphic forms of classical groups via doubling method. To be more precise, if we work over a global field k, PS–R constructed an integral representaion of the L-fuction $L(\pi \times \chi, s)$ for a cuspidal automorphic representation (π, V_π) as the Rankin–Selberg integral of π and $\tilde{\pi}$ inside the "doubled unitary group" $\mathrm{U}(W + W_-)$ against the "Siegel Eisenstein series". Namely, for $\varphi \in V_\pi$, $\tilde{\varphi} \in V_{\tilde{\pi}}$, $f_{s,\chi} \in I_n(s, \chi)$, the Rankin–Selberg integral in concern is

$$Z(\varphi, \tilde{\varphi}, f_{s,\chi}) = \int_{[\mathrm{U}(W)\times\mathrm{U}(W)]} \varphi(g_1)\tilde{\varphi}(g_2)E((g_1, g_2), f_{s,\chi})\chi^{-1}(\det g_2)\,dg_1\,dg_2,$$

where $\mathrm{U}(W) \times \mathrm{U}(W) \hookrightarrow \mathrm{U}(W + W_-)$ comes from $W + W_-$, and

$$E(h, f_{s,\chi}) = \sum_{\gamma \in P_\Delta \backslash \mathrm{U}(W+W_-)} f_{s,\chi}(\gamma h).$$

The reason why it works nicely is because it gives Eulerian integral after you unfold, just like the classical Rankin–Selberg integral. Namely,

$$\begin{aligned} Z(\varphi, \tilde{\varphi}, f_{s,\chi}) &= \int_{[\mathrm{U}(W)\times\mathrm{U}(W)]} \varphi(g_1)\tilde{\varphi}(g_2)E((g_1,g_2), f_{s,\chi})\chi^{-1}(\det g_2)\,dg_1\,dg_2 \\ &= \int_{[\mathrm{U}(W)\times\mathrm{U}(W)]} \varphi(g_1)\tilde{\varphi}(g_2) \sum_{P_\Delta(k)\backslash \mathrm{U}(W+W_-)(k)} f(\gamma(g_1,g_2))\chi^{-1}(\det g_2)\,dg_1\,dg_2 \\ &= \sum_{\gamma\in P_\Delta(k)\backslash \mathrm{U}(W+W_-)(k)/(\mathrm{U}(W)\times\mathrm{U}(W))(k)} \int_{(\mathrm{U}(W)\times\mathrm{U}(W))_\gamma(k)\backslash(\mathrm{U}(W)\times\mathrm{U}(W))(\mathbb{A})} \varphi(g_1)\tilde{\varphi}(g_2)f(\gamma(g_1,g_2))\chi^{-1}(\det g_2)\,dg_1\,dg_2, \end{aligned}$$

where $(\mathrm{U}(W)\times\mathrm{U}(W))_\gamma$ is the centralizer of γ. The double quotient parametrizes the orbits of $(\mathrm{U}(W)\times\mathrm{U}(W))(k)$ on $P_\Delta\backslash\mathrm{U}(W+W_-)$. Every orbit other than the main orbit $\gamma=1$ is **negligible**, in a sense that the stabilizer in $\mathrm{U}(W)\times\mathrm{U}(W)$ is something like a parabolic subgroup such that it has a unipotent radical of a proper parabolic as a normal subgroup. Then one can factor the integral so that it has an integral over this unipotent radical as a factor, which vanishes by cuspidality of π. Thus, only the summand for the main orbit survives, so that

$$\begin{aligned} Z(\varphi, \tilde{\varphi}, f_{s,\chi}) &= \int_{\mathrm{U}(W)^\Delta(k)\backslash(\mathrm{U}(W)\times\mathrm{U}(W))(\mathbb{A})} f_{s,\chi}((g_1,g_2))\varphi(g_1)\tilde{\varphi}(g_2)\chi^{-1}(\det g_2)\,dg_1\,dg_2 \\ &= \int_{\mathrm{U}(W)^\Delta(k)\backslash(\mathrm{U}(W)\times\mathrm{U}(W))(\mathbb{A})} f_{s,\chi}((g_2^{-1}g_1,1))\varphi(g_1)\tilde{\varphi}(g_2)\,dg_1\,dg_2 \\ &= \int_{\mathrm{U}(W)(\mathbb{A})} f_{s,\chi}((g,1))\left(\int_{\mathrm{U}(W)(k)\backslash\mathrm{U}(W)(\mathbb{A})} \varphi(g_2g)\tilde{\varphi}(g_2)\,dg_2\right)dg \\ &= \int_{\mathrm{U}(W)(\mathbb{A})} f_{s,\chi}((g,1))\langle \pi(g)\varphi \mid \tilde{\varphi}\rangle\,dg, \end{aligned}$$

which factors as a product of local zeta integrals. Namely, if $\varphi=\otimes\xi_v$, $\tilde{\varphi}=\otimes\widetilde{\xi_v}$ and $f_{s,\chi}=\otimes f_{s,\chi_v}$, then $Z(\varphi,\tilde{\varphi},f_{s,\chi})$ is the product of local zeta integrals

$$Z_v(\xi_v,\tilde{\xi}_v,f_{s,\chi_v}) = \int_{G(k_v)} f_{s,\chi_v}((g,1))\langle \pi_v(g)\xi_v, \tilde{\xi}_v\rangle\,dg.$$

Indeed, this gives what we will call as L-function for $L(s,\pi\times\chi)$ because at unramified places for a suitable choice of $f_{s,\chi}$, $Z_v(\xi_v^\circ,\tilde{\xi}_v^\circ,f_{s,\chi}) = L(s,\mathrm{BC}_{K_v/k_v}(\pi_v)\otimes\chi_v)$ where $\circ$ means spherical vectors.

More pedantically, we can define local L-factor and local ϵ-factor via the following process, just as in the classical Rankin–Selberg case. Below we drop v and go back to our original local setting. Let q be the cardinality of the residue field of F.

(1) The local L-factor is the normalized generator of the $\mathbb{C}[q^{-s},q^s]$-fractional ideal $\mathcal{I}(\pi\times\chi)$ which is the $\mathbb{C}$-span of local zeta integrals for "good sections" $f_{s,\chi}\in I_n(s,\chi_v)$. In the classical Rankin–Selberg situation, the analogous ideal is the span over the "standard sections", i.e. twists of a section in $I_n(s_0,\chi_v)$ (for a fixed s_0 with large enough real part) along s. Here, one cannot use only the standard sections, partly because the "suitable choice" to realize the expected L-factor is not a standard section. One needs to define a consistent notion

of "good sections," which contains standard sections and more; they should contain the "suitable choices", and they should be stable under the normalized intertwining operators which will be recalled shortly.

In short, $\mathcal{I}(\pi \times \chi) = \left(\frac{1}{P(q^{-s})}\right)$ where $P(X) \in \mathbb{C}[X]$ with $P(0) = 1$, and then we let $L(s+\frac{1}{2}, \pi \times \chi) := P(q^{-s})^{-1}$ (shift by 1/2 is merely cosmetic), analogous to the classical Rankin–Selberg case.

(2) There are accordingly normalized intertwining operators $M^*(s,\chi) : I_n(s,\chi) \longrightarrow I_n(-s,\chi^\dagger)$ where $\chi^\dagger(x) = \chi(\overline{x})^{-1}$, which arises in the functional equation of L-factors. Recall that an unnormalized (i.e. easy-to-write-down) intertwining operator $M(s,\chi)$ can be defined as

$$M(s,\chi)\Phi(g) = \int_{\mathrm{Herm}_n(E)} \Phi(wn(x)g, s)dx.$$

Then $M^*(s,\chi)$ is some constant times $M(s,\chi)$, consistently chosen, so that

$$M^*(-s,\chi^\dagger) \circ M^*(s,\chi) = 1.$$

Then the local functional equation involves $(s,\Phi) \leftrightarrow (-s, M^*(s,\chi)\Phi)$.

(3) The **local γ-factor** is the factor arising in the local functional equation for $Z(\xi,\tilde{\xi},f_s)$ and the **local ϵ-factor** is the factor arising in the local functional equation for $\frac{Z(\xi,\tilde{\xi},f_s)}{L(s+\frac{1}{2},\pi\times\chi)}$. Namely, we define local ϵ-factor as

$$\frac{Z(-s,\xi,\tilde{\xi},M^*(s,\chi)f)}{L(\frac{1}{2}-s,\pi,\chi^\dagger)} = \epsilon(s+\frac{1}{2},\pi,\chi)\frac{Z(s,\xi,\tilde{\xi},f)}{L(s+\frac{1}{2},\pi,\chi)}.$$

Remark 2.1. We are secretly suppressing the choice of additive character ψ that is used to choose normalization $M^*(s,\chi)$ consistently. Everything depends on this ψ (except the local L-factor).

3. Epsilon dichotomy

Now all these foundational discussions are not in vain, as these will be a part of the proof for the epsilon dichotomy principle.

We start by stating some facts in the unitary case analogous to what we've observed for the symplectic-orthogonal reductive pair. Let $\dim W = n, \dim V = m$. The metaplectic group for either $W \otimes V$ or $(W + W_-) \otimes V$ splits over $\mathrm{U}(W)$, $\mathrm{U}(V)$, etc., but the choice of splitting homomorphism is not unique, and in particular involves the choice of characters χ_n, χ_m of $E^\times$ where $\chi_n|_{F^\times} = \epsilon_{E/F}^m$, and $\epsilon_{E/F}$ is the quadratic character. This choice has to appear in our formulae, because for example if we pullback the Weil representation of $\mathrm{Mp}((W + W_-) \otimes V)$ to $\mathrm{U}(W) \times \mathrm{U}(W)$ using this splitting, then it becomes $\omega_{V,\chi} \otimes (\chi_m \cdot \widetilde{\omega_{V,\chi}})$, where $\omega_{V,\chi}$ is the pullback of Weil representation of $\mathrm{Mp}(W \otimes V)$ via the same choice of splitting. This is why even defining a theta lift involves the choice of χ. We denote $\chi = (\chi_n, \chi_m)$ and $\omega_{V,W,\chi} = \omega_{V,\chi} \otimes \omega_{W,\chi}$.

Our first crucial observation was that the Siegel parabolic of Sp acts like a principal series for the distribution δ_0 (i.e. the δ-distribution at 0). The analogous fact is that the degenerate Weil representation for the dual pair $(\mathrm{U}(W+W_-), \mathrm{U}(V))$ has the same property. Namely, $\omega_{V,W+W_-,\chi}$ can be realized on $\mathcal{S}(V^n)$ as a Schrödinger model, and there is a natural map

$$\omega_{V,W+W_-,\chi} \longrightarrow I_n(\frac{m-n}{2}, \chi),$$

$$\varphi \longmapsto \omega_{V,\chi}(g)\varphi(0),$$

which factors through the quotient of $\omega_{V,W+W_-,\chi}$ where $\mathrm{U}(V)$ acts by the character χ_n. We denote this quotient by $R_n(V,\chi)$. Again, by [MVW], $R_n(V,\chi) \longrightarrow I_n(\frac{m-n}{2}, \chi)$ is an injection. Also, we can argue similarly as before that

$$\Theta_\chi(\pi, V) \neq 0 \neq 0 \Longleftrightarrow \mathrm{Hom}_{\mathrm{U}(W)\times\mathrm{U}(W)}(R_n(V,\chi), \pi \otimes (\chi \cdot \tilde{\pi})) \neq 0.$$

Remark 3.1. Actually that we can argue similarly as before is a lie, and the proof of this is more difficult, because we don't have self-duality. We need to exploit some kind of extra involution which amounts to transforming a representation to its contragredient, which is usually called the **MVW involution** in the literature. See [HKS] for the proof of this statement.

Given the dimension m, there are two Hermitian spaces of dimension m over E, determined by the sign

$$\epsilon(V) = \epsilon_{E/F}((-1)^{\frac{m(m-1)}{2}} \det V) \in \{\pm 1\}.$$

We denote V_m^ϵ to be the E-Hermitian space of dimension m of sign ϵ. Then both $R_n(V_m^+, \chi)$ and $R_n(V_m^-, \chi)$ lie inside $I_n(\frac{m-n}{2}, \chi)$.

Now we are back in the game, and we can prove the theta dichotomy principle for unitary groups.

Proof. We have following facts (proofs can be found in [KR] and [KS]).

- $R_n(V_m^+, \chi)$ is always not isomorphic to $R_n(V_m^-, \chi)$. This is believable.
- The Siegel principal series $I_n(s,\chi)$ can have at most 2 irreducible submodules in any case. Indeed, let π be any invariant submodule of $I_n(s,\chi)$. Then by Frobenius reciprocity,

  $$\mathrm{Hom}_{\mathrm{U}(n,n)}(\pi, I_n(s,\chi)) = \mathrm{Hom}_{\mathrm{GL}_n}(\pi_N, \chi|-|^{\text{some power}+s}),$$

  so the problem is really about counting the dimension of some eigenspace of the Jacquet module of $I_n(s,\chi)$. But the Jacquet module $I_n(s,\chi)_N$ has a very explicit description, namely it has a filtration

  $$I_n(s,\chi)_N = I^0 \supset \cdots \supset I^n \supset I^{n+1} = 0,$$

  where the successive quotients $I^r/I^{r+1} \cong \mathrm{Ind}_{P_{n-r,r}}^{\mathrm{GL}_n}(\xi_r)$, for $P_{n-r,r}$ the maximal parabolic corresponding to the partition $n = n-r+r$ and ξ_r is some very explicit character. From the knowledge of explicit characters one deduces the desired result.

Then, (DP1), or $I_n(0,\chi) = R_n(V_n^+,\chi) \oplus R_n(V_n^-,\chi)$, just follows from the above two facts, because $I_n(0,\chi)$ is a unitary representation, so it is completely reducible (!).

For (DP2), it is pretty standard that $\dim \mathrm{Hom}_{\mathrm{U}(W)\times\mathrm{U}(W)}(I_n(s,\chi), \pi \otimes (\chi\cdot\tilde{\pi})) \le 1$, for all s, if π is supercuspidal. Note on the other hand that our local zeta integral gives precisely the desired intertwining map:

$$\begin{aligned} Z^*(\frac{m-n}{2}) : I_n(\frac{m-n}{2},\chi) &\longrightarrow \tilde{\pi}\otimes(\chi\cdot\pi), \\ \Phi &\mapsto \left(\xi\otimes\tilde{\xi} \mapsto \frac{Z(s,\xi,\tilde{\xi},\Phi)}{L(s+\frac{1}{2},\pi,\chi)}\Big|_{s=\frac{m-n}{2}}\right), \end{aligned}$$

where $\Phi \in I_n(\frac{m-n}{2},\chi)$ is extended to the standard section. In our choice of $s=0$ ($m=n$), $Z^*(0)$ gives a nonzero element and spans $\mathrm{Hom}_{\mathrm{U}(W)\times\mathrm{U}(W)}(I_n(s,\chi),\tilde{\pi}\otimes(\chi\cdot\pi))$. Switching π to $\tilde{\pi}$ we get the theta dichotomy. ☐

Remark 3.2. Again there is a slight lie, because $Z^*(s)$ might vanish as good sections are not standard sections. In that case, the leading coefficient of the Laurent expansion of $\frac{Z}{L}$ will do the trick.

We can furthermore decide whether it is V_n^+ or V_n^- to which π lifts to a nonzero representation.

Theorem 3.1 (Epsilon dichotomy, [HKS])**.** *Let π be a supercuspidal representation of* $\mathrm{U}(W)$*, and let χ be a character of $E^\times$ such that $\chi|_{F^\times} = \epsilon_{E/F}^n$. Then $\Theta_\chi(\tilde{\pi},V)\neq 0$ if and only if*

$$\epsilon(\frac{1}{2},\pi,\chi) = \epsilon_{E/F}(-2)^n\epsilon_{E/F}(\det V).$$

Proof. Note that our choice of χ implies $\chi=\chi^\dagger$, so there is extra structure on $I_n(0,\chi)$, namely we have $M^*(0,\chi) : I_n(0,\chi)\longrightarrow I_n(0,\chi)$ with $M^*(0,\chi)^2=1$. This will help us detect the difference between $R_n(V_n^+,\chi)$ and $R_n(V_n^-,\chi)$.

(1) The functional equation for local zeta integral implies that

$$Z^*(0)\circ M^*(0,\chi) = \epsilon(\frac{1}{2},\pi,\chi)Z^*(0).$$

(2) One way of seeing $R_n(V_n^+,\chi)\not\cong R_n(V_n^-,\chi)$ is by the fact that they have different Whittaker functionals. Namely, for $\Phi\in I_n(s,\chi)$, we can define the Whittaker integral

$$W_T(s)\Phi(g) := \int_{\mathrm{Herm}_n(E)} \Phi(wn(x)g,s)\psi(\mathrm{tr}(Tx))dx,$$

for $T\in\mathrm{Herm}_n(E)$. Then, $W_{-2T}(0)$ is nonzero on $R_n(V_T,\chi)$, where V_T is the Hermitian space of dimension n with Hermitian form defined by T, and the same Whittaker functional is zero on the other Hermitian space of same dimension.

(3) On the other hand, if $\det T\neq 0$, we have a functional equation

$$W_T(-s)\circ M^*(s,\chi) = \chi(\det T)^{-1}|\det T|^{-s}\epsilon_{E/F}(\det T)^{n-1}W_T(s).$$

Applying this to $s = 0$, we have

$$W_T(0) \circ M^*(0, \chi) = \epsilon_{E/F}^{-1}(\det T) W_T(0).$$

(4) For $V = V_n^\pm$, take T to be a Hermitian matrix representing the Hermitian form for V. Then we have

$$W_{-2T}(0) \circ M^*(0, \chi)|_{R_n(V,\chi)} = \epsilon_{E/F}(\det(-2T)) W_{-2T}(0),$$

which implies that $M^*(0, \chi)$ acts on $R_n(V, \chi)$ as the scalar $\epsilon_{E/F}(-2)^n \epsilon_{E/F}(\det V)$. Combined with the functional equation for local zeta integral, we get the epsilon dichotomy.

□

4. Witt tower and first occurrence

In this section we describe, without proof, some more refined phenomena related to the theta dichotomy.

- **Witt tower.** We can consider lifting a fixed supercuspidal representation π of $\mathrm{U}(W)$ to various different V's. In particular, we can consider sending it to various entries in a **Witt tower**, which is a tower of Hermitian spaces with a fixed sign. To be more precise, if $\dim W$ is even, then we can consider two Witt towers $\{V_{2r}^+\}$ and $\{V_{2r}^-\}$; $V_{2r}^+ =: V_{r,r}$ is the split space (r copies of hyperbolic planes), and $V_{2r}^- = V_2^- \oplus V_{r-1,r-1}$. Similarly, if $\dim W$ is odd, then the two Witt towers are $\{V_{2r+1}^+\}$ and $\{V_{2r+1}^-\}$. Here it also holds that $V_{2r+1}^\pm = V_1^\pm \oplus V_{r,r}$, and $V_1^\pm$ is one-dimensional Hermitian space over E with Hermitian form $(x, y) = \alpha \overline{x} y$ with $\alpha \in F^\times$ with $\epsilon_{E/F}(\alpha) = \pm 1$.
- **Persistence.** The basic property of theta lifts to Witt towers is **persistence**, which means that if $\Theta_\chi(\pi, V_m^\epsilon) \neq 0$, then $\Theta_\chi(\pi, V_{m+2r}^\epsilon) \neq 0$ for all $r \geq 0$. This is quite easy to see, if you use that

$$\omega_{V_{m+2r}^\epsilon, W, \chi} \cong \omega_{V_m^\epsilon, W, \chi} \otimes \omega_{V_{r,r}, W, 1}.$$

  Namely $\omega_{V_{r,r}, W, 1}$, similar to the symplectic-orthogonal case, as a $\mathrm{U}(W)$-representation is just $\mathcal{S}(W^r)$ with its natural $\mathrm{U}(W)$-action. Thus any $\mathrm{U}(W)$-equivariant map

$$\lambda : \omega_{V_m^\epsilon, W, \chi} \longrightarrow \pi,$$

  can be paired with obvious nonzero $\mathrm{U}(W)$-equivariant map $\delta_0 : \mathcal{S}(W^r) \longrightarrow \mathbb{1}_W, f \mapsto f(0)$, which gives a nonzero $\mathrm{U}(W)$-equivariant map

$$\lambda \otimes \delta_0 : \omega_{V_{m+2r}^\epsilon, W, \chi} \longrightarrow \pi.$$

  In particular, by the persistence property, we can talk about the **first occurrences** of nonzero theta lift on each Witt tower. We denote it by $m_0^\epsilon(\pi, \chi)$. Theta dichotomy is then precisely that

$$\min\{m_0^+(\pi, \chi), m_0^-(\pi, \chi)\} \leq n,$$

$$\max\{m_0^+(\pi, \chi), m_0^-(\pi, \chi)\} > n.$$

- **Stable range.** A priori it might be the case that every theta lift to a Witt tower can be zero. But this does not happen; namely, if $V_m^\epsilon \supset V_{n,n}$, then $\Theta_\chi(\pi, V_m^\epsilon) \neq 0$. So in particular $m_0^\epsilon(\pi,\chi) \le 2n+2$. A proof of this also uses that $\omega_{V_{n,n},W,1} \cong \mathcal{S}(W^n)$. Given a fixed standard basis $\mathbf{e}$ of W over E (n-dimensional), we can view it as an element of W^n, and thus $g \mapsto \mathbf{e}\cdot g$ gives an inclusion $\mathrm{U}(W) \hookrightarrow W^n$, which is right-$\mathrm{U}(W)$-equivariant. Thus the restriction gives a surjective equivariant map $\mathcal{S}(W^n) \twoheadrightarrow \mathcal{S}(\mathrm{U}(W))$, and any irreducible admissible representation appears as a quotient of $\mathcal{S}(\mathrm{U}(W))$.
- **First occurrence.** It is a general phenomenon that, if we start with π supercuspidal, then the first occurence of a nonzero theta lift in a Witt tower is supercuspdial, whereas any later occurrence is not supercuspidal.
- **Early lifts and poles.** In fact, "early nonzero theta lifts" can be accounted for poles at local L-factor. Specifically, if $\min\{m_0^+(\pi,\chi), m_0^-(\pi,\chi)\} = n-2r$, then the local L-factor $L(s,\tilde{\pi},\chi)$ has simple poles at $s = -\frac{1}{2}, -\frac{3}{2}, \cdots, -r+\frac{1}{2}$, i.e. of form
$$L(s,\tilde{\pi},\chi) = \frac{1}{(1-q^{-(s+\frac{1}{2})})(1-q^{-(s+\frac{3}{2})})\cdots(1-q^{-(s+r-\frac{1}{2})})}.$$
From this we can easily construct counterexample to local analogue of naive Ramanujan conjecture. Namely, the later-occuring Witt tower has first occurrence index $\ge n+2$. Let this first occurrence nonzero theta lift be denoted as π'. Then, for π', π is an "early lift," so the local L-factor for π' has a pole outside critical strip, even though it is supercuspidal. This says that π' is supercuspidal but not tempered.
- **Conservation relation.** Finally, there is a very precise relation between $m_0^+(\pi,\chi)$ and $m_0^-(\pi,\chi)$, which is called the **conservation relation**:
$$m_0^+(\pi,\chi) + m_0^-(\pi,\chi) = 2n+2.$$
Note that the theta dichotomy is a corollary of this.

Remark 4.1. There is a global analogue of this picture. For example, the cuspidal spectrum of $\mathrm{O}(Q)(\mathbb{A})$ has an orthogonal decomposition $R_1 \oplus \cdots \oplus R_m$, where R_i is consisted of those having its first occurrence at $\mathrm{Sp}_{2i}(\mathbb{A})$ among all symplectic groups. The first occurrences are cuspidal, and the automorphic multiplicity of $\pi \subset \mathscr{A}(\mathrm{O}(Q))$ is equal to the automorphic multiplicty of the first occurrence in $\mathscr{A}(\mathrm{Sp}_{2i})$. The first occurrence can be read off from poles of L-functions.

References

[HKS] Michael Harris, Stephen Kudla, William Sweet, *Theta dichotomy for unitary groups*, JAMS 9, 1996.

[K] Stephen Kudla, Notes on the local theta correspondence.

[KR] Stephen Kudla, Stephen Rallis, *Ramified degenerate principal series representations for* Sp(n), Israel J. Math. 78, 1992.

[KS] Stephen Kudla, William Sweet, *Degenerate principal series representations for* U(n, n), Israel J. Math. 98, 1997.

[LR] Erez Lapid, Stephen Rallis, *On the local factors of representations of classical groups.*

[MVW] Colette Moeglin, Marie-France Vigneras, Jean-Loup Waldspurger, *Correspondances de Howe sur un corps p-adique*, LNM 1291.

[R] Stephen Rallis, *On the Howe duality conjecture*, Compositio Math. 51, 1984.