Introduction to Hilbert modular forms

Lassina Dembélé

August 1, 2006

Notations

- F is a totally real number field of degree g.
- J_F is the set of all real embeddings of F. For each $\tau \in J_F$, we denote the corresponding embedding into $\mathbb{R}$ by $a \mapsto a^\tau$.
- $\mathcal{O}_F$ denotes the ring of integers of F, and $\mathfrak{d}$ its different.
- For an integral $\mathfrak{p}$ of F, we denote by $F_\mathfrak{p}$ and $\mathcal{O}_{F,\mathfrak{p}}$ the completions of F and $\mathcal{O}_F$, respectively, at $\mathfrak{p}$.
- $\mathbb{A}$ is the ring of adèles of F and $\mathbb{A}_f$ its finite part.
- An element $a \in F$ is **totally positive** if, for all $\tau \in J_F$, $a^\tau > 0$. We denote this by $a \gg 0$.
- Fix an integral ideal $\mathfrak{n}$ of F.

Congurence subrgroups of $\mathrm{GL}_2^+(F)$

The set J_F induces an embedding $\mathrm{GL}_2(F) \hookrightarrow \prod_{\tau \in J_F} \mathrm{GL}_2(\mathbb{R})$ by $\gamma \mapsto (\gamma^\tau)_{\tau \in J_F}$. For any subring A of F, we let

$$\mathrm{GL}_2^+(A) = \left\{ \gamma \in \mathrm{GL}_2(A) : (\gamma^\tau)_{\tau \in J_F} \in \prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R}) \right\}.$$

We have the restriction $\mathrm{GL}_2^+(F) \to \mathrm{PGL}_2^+(F)$, $\gamma \mapsto \tilde{\gamma}$, of the projection map onto $\mathrm{PGL}_2(F)$. We let $\Gamma(1) = \mathrm{GL}_2^+(\mathcal{O}_F)$.

Congruence subgroups of $\mathrm{GL}_2^+(F)$

Definition

A **congruence subgroup** *of* $\mathrm{GL}_2^+(F)$ *is a subgroup* Γ *such that* $\widetilde{\Gamma \cap \Gamma(1)}$ *has finite index in both* $\widetilde{\Gamma}$ *and* $\widetilde{\Gamma(1)}$.

As we will see later, the motivation for such a definition relies in the fact that the arithmetic of Hilbert modular forms on the field F needs to take its narrow class group into account.

Example

Let $\mathfrak{c}$ be a fractional ideal of F, and put

$$\Gamma_0(\mathfrak{c}, \mathfrak{n}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \begin{pmatrix} \mathcal{O}_F & \mathfrak{c}^{-1} \\ \mathfrak{c}\mathfrak{n} & \mathcal{O}_F \end{pmatrix} : ad - bc \in \mathcal{O}_F^{\times +} \right\}.$$

Then, $\Gamma_0(\mathfrak{c}, \mathfrak{n})$ is a congruence subgroup of $\mathrm{GL}_2^+(F)$. This is the only type of congruence subgroups that will be interested in for the rest of this lecture.

Let $\mathfrak{H}$ be the Poincaré upper-half plane and put $\mathfrak{H}_F = \mathfrak{H}^{J_F}$. Then $\prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R})$ acts on $\mathfrak{H}_F$ as follows. For any $\gamma = (\gamma_\tau)_{\tau \in J_F} \in \prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R})$ and $z = (z_\tau)_{\tau \in J_F} \in \mathfrak{H}_F$,

$$\gamma_\tau \cdot z_\tau = \frac{a_\tau z_\tau + b_\tau}{c_\tau z_\tau + d_\tau}, \text{ where } \gamma_\tau = \begin{pmatrix} a_\tau & b_\tau \\ c_\tau & d_\tau \end{pmatrix}.$$

Definition

An element $\underline{k} = (k_\tau)_\tau \in \mathbb{Z}^{J_F}$ is called a **weight vector**. *We always assume that the components $k_\tau \geq 2$ have the same parity.*

From now on, we fix a weight $\underline{k}$. We define an action of $\Gamma_0(\mathfrak{c}, \mathfrak{n})$ on the space of functions $f : \mathfrak{H}_F \to \mathbb{C}$ by putting

$$f\|_{\underline{k}}\gamma = \left(\prod_{\tau \in J_F} \det(\gamma_\tau)^{k_\tau/2}(c_\tau z_\tau + d_\tau)^{-k_\tau}\right) f(\gamma z),\ \gamma \in \Gamma_0(\mathfrak{c}, \mathfrak{n}).$$

Definition

A **classical Hilbert modular form** *of level $\Gamma_0(\mathfrak{c}, \mathfrak{n})$ and weight $\underline{k}$ is a holomorphic function $f : \mathfrak{H}_F \to \mathbb{C}$ such that $f\|_{\underline{k}}\gamma = f$, for all $\gamma \in \Gamma_0(\mathfrak{c}, \mathfrak{n})$. The space of all classical Hilbert modular forms of level $\Gamma_0(\mathfrak{c}, \mathfrak{n})$ and weight $\underline{k}$ is denoted by $M_{\underline{k}}(\mathfrak{c}, \mathfrak{n})$.*

The Fourier expansion

Let $f : \mathfrak{H}_F \to \mathbb{C}$ be a Hilbert modular form. Since it is $\Gamma_0(\mathfrak{c}, \mathfrak{n})$-invariant, we have in particular

$$f(z+\mu) = f(z), \quad \text{for all } z \in \mathfrak{H}_F,\ \mu \in \mathfrak{c}^{-1}.$$

Therefore, it admits a Fourier expansion of the form

$$f(z) = \sum_{\mu \in \mathfrak{d}^{-1}} a_\mu e^{2\pi i \mathrm{Tr}(\mu z)},$$

where $\mathrm{Tr}(\mu z) = \sum_{\tau \in J_F} \mu^\tau z_\tau$.

Koecher's principle

When $g > 1$, every Hilbert modular form is automatically holomorphic at cusps as the next lemma shows.

Lemma (Koecher's principle)

Assume that $g > 1$. Then, f is **holomorphic** *at the cusp ∞ (hence at all cusps $\in \Gamma_0(\mathfrak{c}, \mathfrak{n})\backslash \mathbf{P}^1(F)$) in the following sense:*

$$a_\mu \neq 0 \Rightarrow \mu = 0 \text{ or } \mu \gg 0.$$

Proof

Let $\varepsilon \in \mathcal{O}_F^{\times +}$ be a totally positive unit. Then $\gamma(\varepsilon) = \begin{pmatrix} \varepsilon & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(\mathfrak{c}, \mathfrak{n})$, which means that $f\|_{\underline{k}}\gamma(\varepsilon) = f$. Equating the q–expansion of both members of this equality, it follows that

$$a_{\varepsilon\mu} = \mathrm{N}(\varepsilon)^{\underline{k}/2} a_\mu, \quad \text{for all } \mu \in \mathfrak{c}\mathfrak{d}^{-1},$$

where we use the notation $\mathrm{N}(\varepsilon)^{\underline{k}} = \prod_{\tau \in J_F} (\varepsilon^\tau)^{k_\tau}$.

Proof

Now, let us assume that there is a non–zero $\mu_0 \in \mathfrak{c}\mathfrak{d}^{-1}$ not totally positive such that $a_{\mu_0} \neq 0$. We choose τ_0 such that $\mu_0^{\tau_0} < 0$. By the Dirichlet units theorem, we can find $\varepsilon \in \mathcal{O}_F^{\times +}$ such that

$$\varepsilon^{\tau_0} > 1 \quad \text{and} \quad \varepsilon^{\tau} < 1, \text{ for all } \tau \neq \tau_0.$$

We now consider the subseries of $f(z) = \sum_{\mu \in \mathfrak{c}\mathfrak{d}^{-1}} a_\mu e^{2\pi i \mathrm{Tr}(\mu z)}$ index by the set $\{\mu_0 \varepsilon^m, \ m \in \mathbb{N}\}$, in which we put $z = \underline{i}$. Then

$$a_{\mu_0 \varepsilon^m} e^{-2\pi \mathrm{Tr}(\mu_0 \varepsilon^m)} = \mathrm{N}(\varepsilon)^{m\underline{k}/2} a_{\mu_0} e^{-2\pi \mathrm{Tr}(\mu_0 \varepsilon^m)}.$$

Proof

But, as $m \to \infty$, $e^{-2\pi \mathrm{Tr}(\mu_0 \varepsilon^m)} \sim e^{-2\pi \mu_0^{\tau_0} (\varepsilon^{\tau_0})^m}$, and the exponential growth ensures that $\mathrm{N}(\varepsilon)^{m\underline{k}/2} a_{\mu_0} e^{-2\pi \mathrm{Tr}(\mu_0 \varepsilon^m)} \to \infty$. Therefore the series does not converge, which is a contradiction. So we must have $a_{\mu_0} = 0$.

Defintion of cusp forms

Definition

We say that f is a **cusp form** *if the constant term a_0 in the Fourier expansion is equal to 0 for any $f\|_{\underline{k}}\gamma$, $\gamma \in \mathrm{GL}_2^+(F)$ (i.e., if f vanishes at all cusps). We will denote by $S_{\underline{k}}(\mathfrak{c}, \mathfrak{n})$) the space of cusp forms of weight $\underline{k}$ and level $\Gamma_0(\mathfrak{c}, \mathfrak{n})$.*

Corollary

$S_{\underline{k}}(\mathfrak{c}, \mathfrak{n}) = M_{\underline{k}}(\mathfrak{c}, \mathfrak{n})$ *unless* $k_\tau = k_{\tau'}$ *for all* $\tau, \tau' \in J_F$.

Proof. Let assume that there is $f \in M_{\underline{k}}(\mathfrak{c}, \mathfrak{n})$ that is not a cusp form. Then at some cusp σ, the q–expansion must give $a_0 \neq 0$. From

$$a_0 = \mathrm{N}(\varepsilon)^{\underline{k}/2} a_0, \quad \text{for all } \varepsilon \in \mathcal{O}_F^{\times +},$$

it follows that we must have $\mathrm{N}(\varepsilon)^{\underline{k}/2} = 1$ for all $\varepsilon \in \mathcal{O}_F^{\times +}$. But this is possible only if we have $k_\tau = k_{\tau'}$ for all $\tau, \tau' \in J_F$.

Proposition

(i) $M_{\underline{k}}(\mathfrak{c}, \mathfrak{n}) = 0$ *unless* $k_\tau \geq 0$ *for all* $\tau \in J_F$.
(ii) $M_0(\mathfrak{c}, \mathfrak{n}) = \mathbb{C}$ *and* $S_0(\mathfrak{c}, \mathfrak{n}) = 0$.

Proof. van der Geer [**?**, Chap. I. sec. 6.]

Example: Eiseinstein series

Let $\mathfrak{c}$ be an ideal in F and $k \geq 2$ an even integer. Put

$$G_{k,\mathfrak{c}}(z) = \mathrm{N}(\mathfrak{c})^k \sum_{(c,d)\in \mathbf{P}^1(\mathfrak{c}\times\mathfrak{a}\mathfrak{c})} \mathrm{N}(cz+d)^{-k},$$

where $\mathbf{P}^1(\mathfrak{c}\times\mathfrak{a}\mathfrak{c}) = \{(c,d)\in\mathfrak{c}\times\mathfrak{a}\mathfrak{c} | (c,d)\neq(0,0)\}/\mathcal{O}_F^\times$. It can be shown that $G_{k,\mathfrak{c}}$ is a modular form of weight $\underline{k} = (k,\cdots,k)$ and level $\Gamma_0(\mathfrak{c},\mathfrak{a})$. We call $G_{k,\mathfrak{c}}$ a **Eiseinstein series** of weight k and level $\Gamma_0(\mathfrak{c},\mathfrak{a})$. The Eisenstein series $G_{k,\mathfrak{c}}$ only depends on the ideal class of $\mathfrak{c}$.

Adelic Hilbert modular forms

We recall that $\prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R})$ acts transitively on $\mathfrak{H}_F$ by linear fractional transforms and that the stabilizer of $\underline{i} = (i, \ldots, i)$ is given by $K_\infty^+ = (\mathbb{R}^\times \mathrm{SO}_2(\mathbb{R}))^{J_F}$.

We consider the unique action of $\prod_{\tau \in J_F} \mathrm{GL}_2(\mathbb{R})$ on $\mathfrak{H}_F$ that extends the action of $\prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R})$. Namely, one each copy of $\mathfrak{H}$, we let the element $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ acts by $z \mapsto -\bar{z}$.

Level structure

We consider the following compact open subgroup of $\mathrm{GL}_2(\mathbb{A}_f)$:

$$K_0(\mathfrak{n}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\widehat{\mathcal{O}_F}) : c \in \mathfrak{n} \right\},$$

where $\widehat{\mathcal{O}_F} = \prod_{\mathfrak{p}} \mathcal{O}_{F,\mathfrak{p}}$.

Automorphy factor

We set $\underline{t} = (1, \ldots, 1)$ and $\underline{m} = \underline{k} - 2\underline{t}$, then choose $\underline{v} \in \mathbb{Z}^{J_F}$ such that each $v_\tau \geq 0$, $v_\tau = 0$ for some τ, and $\underline{m} + 2\underline{v} = n\underline{t}$ for some non-negative $n \in \mathbb{Z}$.

Definition

For any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \prod_\tau \mathrm{GL}_2(\mathbb{R})$ *and* $z \in \mathfrak{H}_F$*, put*

$$j(\gamma, z) = \prod_{\tau \in J_F} (c_\tau z_\tau + d_\tau).$$

The map $(\gamma, z) \mapsto j(\gamma, z)$ *is called an* **automorphy factor**.

Definition

Definition

An **adelic Hilbert modular form** *of weight* $\underline{k}$ *and level* $\mathfrak{n}$ *is a function* $f : \mathrm{GL}_2(\mathbb{A}) \to \mathbb{C}$ *satisfying the following conditions:*

(i) $f(\gamma g u) = f(g)$ *for all* $\gamma \in \mathrm{GL}_2(F)$, $u \in K_0(\mathfrak{n})$ *and* $g \in \mathrm{GL}_2(\mathbb{A})$.

(ii) $f(gu) = \det(u)^{\underline{k}-\underline{v}-\underline{t}} j(u, \underline{i})^{-\underline{k}} f(g)$ *for all* $u \in K_\infty^+$ *and* $g \in \mathrm{GL}_2(\mathbb{A})$.

Definition (con't)

Definition

For all $x \in \mathrm{GL}_2(\mathbb{A}_f)$, *define* $f_x : \mathfrak{H}_F \to \mathbb{C}$ *by* $z \mapsto \det(g)^{\underline{t}-\underline{v}-\underline{k}} j(g, \underline{i}) f(xg)$, *where we choose* $g \in \prod_{\tau \in J_F} \mathrm{GL}_2^+(\mathbb{R})$ *such that* $z = g \cdot \underline{i}$. *By (ii)* f_x *does not depend on the choice of* g.

(iii) f_x *is holomorphic (when* $F = \mathbb{Q}$, *an extra holomorphy condition at cusps is needed).*

(iv) *In addition, when* $\int_{U(\mathbb{A})/U(\mathbb{Q})} f(ux)du = 0$ *for all* $x \in \mathrm{GL}_2(\mathbb{A})$ *and all additive Haar measures* du *on* $U(\mathbb{A})$, *where* U *is the unipotent radical of* GL_2/F, *we say that* f *is an* **adelic cusp form**.

We will denote the space of all Hilbert modular forms (resp. cusp forms) of weight $\underline{k}$ and level $\mathfrak{n}$ by $M_{\underline{k}}(\mathfrak{n})$ (resp. $S_{\underline{k}}(\mathfrak{n})$).

There is a relation between classical and adelic Hilbert modular forms which proves important when dealing with questions that relate to the arithmetic of these forms.

Let $\mathfrak{c}_\lambda$, $\lambda = 1, \ldots, h^+$, be representatives of the narrow ideal classes of F. For each $\lambda = 1, \ldots, h^+$, take $x_\lambda \in \mathrm{GL}_2(\mathbb{A})$, so that $t_\lambda = \det(x_\lambda)$ generates the ideal $\mathfrak{c}_\lambda$. Then, by the strong approximation theorem,

$$\mathrm{GL}_2(\mathbb{A}) = \coprod_{\lambda=1}^{h^+} \mathrm{GL}_2(F) x_\lambda \left(\prod_\tau \mathrm{GL}_2^+(\mathbb{R}) \times K_0(\mathfrak{n}) \right),$$

and we see that

$$\Gamma_\lambda = \Gamma_0(\mathfrak{c}_\lambda, \mathfrak{n}) = x_\lambda \left(\prod_\tau \mathrm{GL}_2^+(\mathbb{R}) \times K_0(\mathfrak{n}) \right) x_\lambda^{-1} \cap \mathrm{GL}_2(F).$$

To each adelic Hilbert modular form f, we associated the h^+-tuple $(f_1, \ldots, f_{h^+}) \in \bigoplus_{\lambda=1}^{h^+} S_{\underline{k}}(\mathfrak{c}_\lambda, \mathfrak{n})$, where $f_\lambda = f_{x_\lambda}$ is given by Definition 6. Then, we have

Proposition

The map

$$\begin{aligned} S_{\underline{k}}(\mathfrak{n}) &\rightarrow \bigoplus_{\lambda=1}^{h^+} S_{\underline{k}}(\mathfrak{c}_\lambda, \mathfrak{n}) \\ f &\mapsto (f_1, \ldots, f_{h^+}) \end{aligned}$$

is an isomorphism of complex vector spaces.

Proof.

The converse of the map is given by the $\mathbb{C}$-valued function f on $\mathrm{GL}_2(\mathbb{A})$ defined by

$$f(\gamma x_\lambda g) = (f_\lambda \|_{\underline{k}} g_\infty)(\underline{i}), \quad \gamma \in \mathrm{GL}_2(F) \text{ and } g \in \mathrm{GL}_2^+(\mathbb{R}) \times K_0(\mathfrak{n}).$$

The theory of Hilbert modular forms has a wide range of applications. Here we list few of them.

In the wake of Wiles proof of the Fermat Last Theorem, a strategy was outlined by Darmon in order to solve the generalized Fermat equation $x^p + y^q = z^r$, for p, q, r a set of arbitrary primes. In his framework, Hilbert modular play a central rôle. For example, to solve the generalized Fermat equation $x^p + y^p = z^5$ one is led to the natural consideration of Galois representations associated to Hilbert modular forms over the real quadratic field $\mathbb{Q}(\sqrt{5})$.

R. Livné, K. Lauter et al. have constructed Ramanujan graphs using Hilbert modular forms. Their works find some application to the construction of robust networks.

Many conjectures relating to classical modular forms find their natural generalization to the setting of Hilbert modular forms. One such conjecture is the Serre conjecture. In this case it is stated as follows.

Conjecture

Let $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_2(\overline{\mathbb{F}_\ell})$ *be a continous irreducible Galois representation such that* $\det(\rho(c_\tau)) = -1$*, where* c_τ *is complex conjugation at* $\tau \in J_F$*, ane which is unramified outside a finite set of primes. Then* ρ *comes form a Hilbert cusp form.*

The Serre conjecture for Hilbert modular forms is still far from a complete proof as the key ingredient used by Khare and other breaks down in this case.

The Goal of the next three lectures

- Relate Hilbert modular forms to Brandt module using the Eichler-Shimizu or Jacquet-Langlands correspondence.
- Show how to compute this Brandt module in a more efficient way (in the case of real quadratic fields).
- The Eichler-Shimura construction for Hilbert modular forms. This is mainly a conjecture, but we hope that the construction of a database of modular modular elliptic curves and abelian surfaces will provide more evidence in instances where one cannot use Shimura curves.