HILBERT MODULAR FORMS AND THEIR GALOIS REPRESENTATIONS

HARUZO HIDA

In this topic course, assuming basic knowledge of algebraic number theory, commutative algebra and topology, we pick topics from the theory of Hilbert modular forms and modular Galois representations into $GL(2)$. We plan to discuss the following four topics:

(1) Basics of analytic/algebraic theory of Hilbert/quaternion automorphic forms (e.g., [HMI, Chapter 2]),
(2) Relation between Quaternionic automorphic forms and Hilbert modular forms (quaternionic automorphic forms are indispensable in construction of the Galois representation though we do not go into details of construction),
(3) Description of Galois representation attached to modular forms,
(4) Description of the "big" Galois representation attached to a p-adic families of modular forms (if time allows).

Since this is a topic course, for some of the topics, we just give the results without detailed proofs. Main reference is Chapters 2 and 3 of the following book [HMI]:

[HMI] H. Hida, *Hilbert Modular Forms and Iwasawa Theory*, Oxford University Press, 2006 (a list of errata posted at `www.math.ucla.edu/~hida`).

Here are some relevant books:

[LFE] H. Hida, *Elementary Theory of L–functions and Eisenstein Series*, LMSST **26**, Cambridge University Press, Cambridge, 1993.

[MFM] T. Miyake, *Modular Forms*, Springer, New York-Tokyo, 1989.

Take a look at the **Overview** posted in the class home page to know why we need to study quaternion algebras and automorphihc forms on them even just to construct Galois representations associated to Hilbert modular forms.

Date: December 28, 2015.

The author is partially supported by the NSF grant: DMS 1464106.

Contents

1. Quaternion Algebras Over a Number Field

We recall basic structure theorems of quaternion algebras over a number field. Since our description is limited to a minimum (necessary for understanding of the later discussions), we refer to [BNT] and [AAQ] for thorough description of the arithmetic of simple algebras and quaternion algebras.

1.1. **Quaternion algebras.** Let F be a field of characteristic 0. A *quaternion* algebra D over F is a central simple algebra of dimension 4 over F. Here the word "central" means that F is the center of D and "simple" means that there are no two-sided ideals of D except for $\{0\}$ and D itself. First suppose that D is not a division algebra. Thus D has a proper left ideal $\mathfrak{a} \subsetneq D$. Since $\mathfrak{a}$ is also a vector F–subspace of D, its dimension over F is either 1, 2 or 3. If $\dim(\mathfrak{a}) = 1$, then for a generator x of the subspace $\mathfrak{a}$, bx for $b \in D$ is a constant multiple of x itself. Write this constant as $\rho(b) \in F$. Then $\rho : D \to F$ is an F–algebra homomorphism and hence surjective. Thus $\mathrm{Ker}(\rho)$ is a three dimensional two sided ideal, which contradicts to the simplicity of D. Thus $\dim(\mathfrak{a}) > 1$. Similarly, if $\dim(\mathfrak{a}) = 3$, $D/\mathfrak{a}$ is a one dimensional vector space over F on which the algebra D acts. By the same argument as above, $\dim(\mathfrak{a}) = 3$ is impossible. Thus $\dim(\mathfrak{a}) = 2$. Choose a basis x_1, x_2 of $\mathfrak{a}$ over F. Define $\rho : D \to M_2(F)$ (the 2×2 matrix algebra with entries in F) by $(x_1, x_2)\rho(b) = b(x_1, x_2)$. Then $\rho : D \to M_2(F)$ is an F–algebra homomorphism taking the identity to the identity. Thus $\mathrm{Ker}(\rho)$ is a two sided ideal of D. Since $\rho \neq 0$, the simplicity (non-existence of non-trivial two sided ideals) tells us that $\mathrm{Ker}(\rho) = \{0\}$. Thus ρ is injective. Comparing the dimension, we conclude that $\rho : D \cong M_2(F)$ is an isomorphism.

Now assume that D is a division algebra (so every non-zero element has a left and right inverse). Pick $x \in D$ which is not in the center F. Then the subalgebra $K = F[x] \subset D$ has to be a field. Thus D becomes a vector space over K via left multiplication by elements of K. Then we have

$$4 = \dim_F D = (\dim_K D) \times [K : F].$$

Thus $[K : F]$ is either 4 or 2. If $[K : F] = 4$, $K = D$, and D becomes commutative. This contradicts the centrality of D over F, and $[K : F] = 2$. Taking $y \in D - K$, we can consider the subspace $K + Ky$ in D. These two spaces D and $K + Ky$ have dimension 2 over K, and $D = K + Ky$. Then we define a representation $\rho : D \to M_2(K)$ by

$$\rho(b) \left(\begin{smallmatrix} 1 \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ y \end{smallmatrix}\right) b = \left(\begin{smallmatrix} b \\ yb \end{smallmatrix}\right).$$

This means $\alpha + \beta y = b$ and $\gamma + \delta y = yb$ when $\rho(b) = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$. Since D is simple, ρ is injective. Thus D is realized as an F–subalgebra of $M_2(K)$.

We now determine the image $\rho(D)$ of D in $M_2(K)$ explicitly. Since x is quadratic over F, we may assume that x satisfies $x^2 - ax + b = 0$ for $a, b \in F$, which is the minimal equation of x over F. Thus $X^2 - aX + b$ has two distinct roots: x, x^τ which are conjugate to each other under the generator τ of $\mathrm{Gal}(K/F)$. Then $\rho(x)$ satisfies the same equation, and the eigenvalues of $\rho(x)$ are roots of $X^2 - aX + b = 0$. Since $\rho(x)$ is not in the center of $\rho(D)$, it is not a scalar matrix, and the eigenvalues of $\rho(x)$ are two distinct roots of $X^2 - aX + b = 0$. Changing the basis $(1, y)$ suitably (we write the new basis as $(1, v)$), we may assume that $\rho(x) = \left(\begin{smallmatrix} x & 0 \\ 0 & x^\tau \end{smallmatrix}\right)$. By the definition: $\rho(b) \left(\begin{smallmatrix} 1 \\ v \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ v \end{smallmatrix}\right) b$, we see $x^\tau v = vx$ in D. Here we do not need to change 1 because 1 is already an eigenvector for $\rho(x)1 = 1x = x1$. Thus $\rho(x^\tau)\rho(v) = \rho(v)\rho(x)$. This implies that for any $a \in K$, $\rho(a^\tau)\rho(v) = \rho(v)\rho(a)$ and $\rho(v) = \left(\begin{smallmatrix} 0 & \alpha \\ \beta & 0 \end{smallmatrix}\right)$ for $\alpha, \beta \in K$. Replacing v by $v\alpha^{-\tau}$, we may assume that $\rho(v) = \left(\begin{smallmatrix} 0 & 1 \\ \xi & 0 \end{smallmatrix}\right)$ and $v^2 = \xi$. If $\xi \notin F$, then $F[v]$ in D is of degree 4 over F, which is impossible. Thus $\xi \in F$. Namely

$$\rho(D) = \left\{ \rho(a + bv) = \left(\begin{smallmatrix} a & b \\ \xi b^\tau & a^\tau \end{smallmatrix}\right) \middle| a, b \in K \right\}.$$

Thus we can always realize D as a subalgebra of $M_2(K)$ for any quadratic extension K/F embeddable into D and a suitable $\xi \in F^\times$ in the above form. We define for $\beta = a + bv \in D$,

$$N(\beta) = \det(\rho(\beta)) = aa^\tau - \xi bb^\tau \in F \quad \text{and } \mathrm{Tr}(\beta) = \mathrm{Tr}(\rho(\beta)) = a + a^\tau \in F.$$

Moreover, $\beta = a + bv$ satisfies the equation $X^2 - \mathrm{Tr}(\beta)X + N(\beta) = 0$ in D, and hence, Tr and N are independent of the choice of v and K. The map N is called the reduced norm and Tr is called the reduced trace. On the image $\rho(D) \subset M_2(K)$, the reduced norm N coincides with the determinant map of $M_2(K)$; so, N is multiplicative: $N(ab) = N(a)N(b)$ for $a, b \in D$.

Now we start with a subalgebra D_ξ ($\xi \in F^\times$) of $M_2(K)$ for a quadratic extension K/F given by

$$D_\xi = \left\{ \left(\begin{smallmatrix} a & b \\ \xi b^\tau & a^\tau \end{smallmatrix}\right) \middle| a, b \in K \right\}.$$

Since K is 2 dimensional over F, D_ξ is 4 dimensional over F. Obviously D_ξ is stable under multiplication and addition. We also see easily that the center of D_ξ is F. Moreover $D_\xi + D_\xi\delta = M_2(K)$ for any generator δ of K over F. Thus $D_\xi \otimes_F K \cong M_2(K)$, which shows that D_ξ is a central simple algebra over F.

If $\xi = \alpha\alpha^\tau$ for $\alpha \in K^\times$, for $v = \left(\begin{smallmatrix} 0 & 1 \\ \xi & 0 \end{smallmatrix}\right) \in D_\xi$, $N(a+bv) = aa^\tau - \alpha b(\alpha b)^\tau$. Choosing $a = \alpha b$, $c = \alpha b + bv$ has determinant 0 but is a nonzero matrix. Thus D_ξ is not a division algebra and hence $D_\xi \cong M_2(F)$. If $\xi \notin N_{K/F}(K^\times)$, then $N(a+bv) = aa^\tau - \xi bb^\tau = 0$ implies $a = b = 0$. Therefore $a + bv$ has always an inverse if $a + bv \neq 0$; so, D_ξ is a division algebra. We have proven

$$\xi \in N_{K/F}(K^\times) \text{ if and only if } D_\xi \cong M_2(F). \tag{1.1}$$

If $\xi = \alpha\alpha^\tau\eta$ for $\xi, \eta \in F^\times$ and $\alpha \in K^\times$, we see

$$\left(\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ \xi b^\tau & a \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & \alpha^{-1} \end{smallmatrix}\right) = \begin{pmatrix} a & \alpha^{-1}b \\ \xi(\alpha^{-1}b)^\tau & a^\tau \end{pmatrix} \in D_\xi.$$

Thus $D_\xi \cong D_\eta$ if $\xi = \alpha\alpha^\tau\eta$ for $\xi, \eta \in F^\times$ and $\alpha \in K^\times$. Therefore the map: $\xi N_{K/F}(K^\times) \mapsto D_\xi$ induces a surjection:

$$\frac{F^\times}{N_{K/F}(F^\times)} \xrightarrow{onto} \left\{\text{the isomorphism classes of } D_\xi \text{ in } M_2(K) \text{ for } \xi \in F^\times\right\}.$$

We find that by (1.1), this map is actually a bijection.

When $F = \mathbb{R}$, then the only possibility of K is $\mathbb{C}$. Since $N_{\mathbb{C}/\mathbb{R}}(\mathbb{C}^\times) = \mathbb{R}^\times_+$, we have $\mathbb{R}^\times/N_{\mathbb{C}/\mathbb{R}}(\mathbb{C}^\times) \cong \{\pm 1\}$, there is only two isomorphism classes of quaternion algebras: one is $M_2(\mathbb{R}) = D_1$ and the other is the Hamilton quaternion algebra $\mathbb{H} = D_{-1}$. When $F = \mathbb{C}$, there is no quadratic extension of $\mathbb{C}$, thus there is only one isomorphism class $M_2(\mathbb{C}) = D_1$.

Now we suppose that F is a p–adic field, that is, a finite extension of $\mathbb{Q}_p$, and we study the quaternion algebras over F. Let ϖ be the prime element of the p–adic integer ring O of F. Then

$$F^\times \cong O^\times \times \{\varpi^n | n \in \mathbb{Z}\}.$$

We define formally the logarithm map on $O^\times$ by

$$\log(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x-1)^n}{n}$$

as long as the above series converges p-adically. Since for sufficiently large r, log converges p–adically on $1 + \varpi^r O$ in $O^\times$, $O^\times \cong \mu \times {\mathbb{Z}_p}^{[F:\mathbb{Q}_p]}$ for the subgroup μ of roots of unity in O. This shows $F^\times \cong \mu \times {\mathbb{Z}_p}^{[F:\mathbb{Q}_p]} \times \mathbb{Z}$ and

$$C_2 = F^\times/(F^\times)^2 \cong \begin{cases} (\mu/\mu^2) \times \mathbb{Z}/2\mathbb{Z} & \text{if } p > 2 \\ (\mu/\mu^2) \times (\mathbb{Z}/2\mathbb{Z})^{[F:\mathbb{Q}_p]+1} & \text{if } p = 2. \end{cases} \tag{$*$}$$

For any quadratic extension K/F, we can choose a generator δ of K over F such that $\delta^2 \in F$. Since $F[\delta] = F[\alpha\delta]$ for $\alpha \in F^\times$, we have a map: $F[\delta] \mapsto \delta^2 \mod (F^\times)^2$ induces a bijection

$$\{\text{isomorphism classes of quadratic extension } K/F\} \cong C_2.$$

If $p > 2$, for any p–adic unit $u \notin (O^\times)^2$, $K = F[\sqrt{u}]$ is the unique unramified quadratic extension. Thus if we write U_K for the group of p–adic units of K, $U_F = O^\times = N_{K/F}(U_K)$ if K/F is unramified and $p > 2$. Thus $F^\times/N_{K/F}(K^\times) \cong \mathbb{Z}/2\mathbb{Z}$ if K/F is unramified and $p > 2$.

Suppose that K/F is ramified. Let O_K be the p–adic integer ring of K and $\mathfrak{P}$ be the maximal ideal of O_K. Then $N_{K/F}(x) \mod \mathfrak{P} = (x \mod \mathfrak{P})^2$ for $x \in O_K^\times$. This shows that $F^\times/N_{K/F}(K^\times)$ has a quotient group isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Since a prime element of F is a norm of the prime element of K, $F^\times/N_{K/F}(K^\times)$ is a proper quotient group of $(\mu/\mu^2) \times \mathbb{Z}/2\mathbb{Z} \cong (\mathbb{Z}/2\mathbb{Z})^2$. Thus we know from $(*)$ that $F^\times/N_{K/F}(K^\times) \cong \mathbb{Z}/2\mathbb{Z}$ even if K/F is ramified (as long as $p > 2$).

Even if $p = 2$, we can prove by local class field theory

$$F^\times/N_{K/F}(K^\times) \cong \mathrm{Gal}(K/F) \cong \mathbb{Z}/2\mathbb{Z}. \tag{1.2}$$

Thus for a given quadratic extension K/F, there are only two isomorphism classes of quaternion algebras D_ξ/F inside $M_2(K)$.

We return to unramified K/F. The unique division quaternion algebra of the form D_ξ in $M_2(K)$ is isomorphic to D_ϖ. Since $v = \left(\begin{smallmatrix} 0 & 1 \\ \varpi & 0 \end{smallmatrix}\right)$ and $v^2 = \varpi$, we know that the ramified extension $F[\sqrt{\varpi}]$ is isomorphic to $K' = F[v] \subset D_\varpi$. Thus any ramified quadratic extension K'/F is embeddable in D_ϖ. Since a division quaternion algebra embeddable in $M_2(K')$ corresponds to the generator of $F^\times/N_{K/F}(K^\times) \cong \mathbb{Z}/2\mathbb{Z}$, which is unique. There is only one isomorphism class of division quaternion algebras over F. Thus we know that any quaternion algebra over F is either isomorphic to the unique division quaternion algebra or $M_2(F)$.

Now we assume that F is a number field. For each prime ideal $\mathfrak{p}$ of the integer ring O of F, we write $F_\mathfrak{p}$ for the $\mathfrak{p}$-adic completion of F, and we put $D_\mathfrak{p} = D \otimes_F F_\mathfrak{p}$, which is a quaternion algebra over $F_\mathfrak{p}$. The prime $\mathfrak{p}$ is called ramified in $D_\mathfrak{p}$ if $D_\mathfrak{p}$ is a division quaternion algebra. We take a quadratic extension K/F inside D and take $\xi \in F^\times$ so that $D \cong D_\xi$. Then $\xi \in N_{K_\mathfrak{p}/F_\mathfrak{p}}(K_\mathfrak{p}^\times)$ for all but finitely many $\mathfrak{p}$, where $K_\mathfrak{p} = K \otimes_F F_\mathfrak{p}$. Indeed, if $K_\mathfrak{p}$ is a field extension of $F_\mathfrak{p}$, it is unramified for all but finitely many $\mathfrak{p}$ (hence $N_{K_\mathfrak{p}/F_\mathfrak{p}}(U_{K_\mathfrak{p}}) = U_{F_\mathfrak{p}}$ for almost all $\mathfrak{p}$). Since ξ is a unit for all but finitely many $\mathfrak{p}$, we know $\xi \in N_{K_\mathfrak{p}/F_\mathfrak{p}}(K_\mathfrak{p}^\times)$ for all but finitely many $\mathfrak{p}$. If $K_\mathfrak{p} = F_\mathfrak{p} \oplus F_\mathfrak{p}$, $D_\mathfrak{p} \cong M_2(F_\mathfrak{p})$ and also $\xi \in N_{K_\mathfrak{p}/F_\mathfrak{p}}(K_\mathfrak{p}^\times)$. This shows that D ramifies at only finitely many places. For each embedding $\sigma : F \hookrightarrow \mathbb{R}$, we say that D is ramified at σ, if $D_\sigma = D \otimes_{F,\sigma} \mathbb{R} \cong \mathbb{H}$. We write S for the set of all ramified places of F in D. Then global class field theory tells us the following facts found by Hasse:

(H1) *The cardinality $|S|$ is even;*

(H2) *For any given set S of places with even cardinality, there exists a unique quaternion algebra D ramifying exactly at S.*

One can find a proof of this in [BNT] XIII.3, Theorem 2 and XIII.6, Theorem 4.

Exercise 1.1.

1. *Prove that the center of D_ξ ($\xi \in F^\times$) is equal to F.*
2. *If $D \otimes_F K \cong M_2(K)$ for an F–algebra D, prove that D is a central simple algebra of dimension 4 over F.*
3. *Determine the radius of convergence of the p–adic logarithm.*
4. *For the p–adic integer ring O of a finite extension $F/\mathbb{Q}_p$, give a detailed proof of $O^\times \cong \mu \times \mathbb{Z}_p{}^{[F:\mathbb{Q}_p]}$ for the subgroup μ of roots of unity in O.*

(5) *Give a detailed proof of* $(*)$.
(6) *Let p be an odd prime, and K/F is a quadratic extension of p–adic fields. Without using class field theory, give a detailed proof of* (1.2).
(7) *Let F be a number field and D be a division quaternion algebra containing a quadratic extension K/F. Prove that if $K_{\mathfrak{p}} = F_{\mathfrak{p}} \oplus F_{\mathfrak{p}}$, $D_{\mathfrak{p}} \cong M_2(F_{\mathfrak{p}})$*

1.2. **Orders of quaternion algebras.** Let F be a field and A be a subring of F. We assume that the field of fractions of A coincides with F. Let D be a quaternion algebra over F. Let V be a finite dimensional vector space over F. An A–lattice L in V is an A–submodule of D which satisfies

(L1) L is an A-module of finite type (i.e., $L = \sum_i A\xi_i$ for finitely many ξ_i);
(L2) $L \otimes_A F = V$ (i.e. L contains a basis of V over F).

If V is an F-algebra, an A–order R of V is an A–lattice in V which is a subring of V sharing the identity. Taking a quadratic extension K/F, we realize D as

$$D_\xi = \left\{ \alpha(a,b) = \left(\begin{smallmatrix} a & b \\ \xi b^\sigma & a^\sigma \end{smallmatrix}\right) \middle| a, b \in K \right\},$$

where σ is the generator of $\mathrm{Gal}(K/F)$. Note that $D_\xi \cong D_{N_{K/F}(\eta)\xi}$ for $\eta \in K^\times$. If A' is an A–order of K, we may therefore assume that $\xi \in A'$ by replacing ξ by $N_{K/F}(\eta)\xi$ for a suitable η if necessary. Then

$$R_\xi = \left\{ \alpha(a,b) \in D_\xi \middle| a, b \in A' \right\}$$

is an A–order in D. We define, for $\alpha = \alpha(a,b) \in D_\xi$,

$$\alpha^\iota = \alpha(a^\sigma, -b) = \mathrm{Tr}(\alpha) - \alpha = N(\alpha)\alpha^{-1}.$$

Then $\alpha \mapsto \alpha^\iota$ is an F–linear involution: $(\alpha\beta)^\iota = \beta^\iota\alpha^\iota$. Recall that $N(\alpha) = \det(\alpha)$. Then in D, $N(\alpha) = \alpha\alpha^\iota \in F$ and $\mathrm{Tr}(\alpha) = \alpha + \alpha^\iota \in F$. Especially $P_\alpha(X) = X^2 - \mathrm{Tr}(\alpha)X + N(\alpha)$ is the minimal polynomial of α in D if $\alpha \notin F$, i.e. $P_\alpha(\alpha) = 0$ and P_α is monic and has minimal degree among all monic polynomials $Q(X) \in F[X]$ with $Q(\alpha) = 0$.

Now we assume F to be a p–adic field and A to be the p–adic integer ring O of F. If R is an order of D, then R is free of rank 4 over O because O is a valuation ring. Then taking a base $x = (x_1, x_2, x_3, x_4)$ of R, we define the regular representation $\rho : D \hookrightarrow M_4(F)$ by $\rho(\alpha)^t x = {}^t x\alpha = {}^t(x_1\alpha, x_2\alpha, x_3\alpha, x_4\alpha)$. Then $\rho(R)$ is contained in $M_4(O)$. Thus $Q(\alpha) = 0$ for $Q(X) = \det(X1_4 - \rho(\alpha))$. Since $P_\alpha(X)$ is the minimal polynomial of α, $P_\alpha(X)$ is a factor of $Q(X)$. Since $Q(X)$ is monic and has coefficients in O, by Gauss' lemma, $P_\alpha(X)$ has coefficients in O. Namely N and Tr induce $N : R \to O$ and $\mathrm{Tr} : R \to O$.

Suppose now that D is a division algebra. We put

$$R_0 = \left\{ \alpha \in D \middle| N(\alpha) \in O \right\}.$$

Since D is a division algebra, $\xi \notin N_{K/F}(K^\times)$. Note that $N_{K/F}(K^\times) \subset N(D^\times)$ because $aa^\sigma = N(\alpha(a,0))$. Since $F^\times/N_{K/F}(K^\times) \cong \mathbb{Z}/2\mathbb{Z}$ and D contains any quadratic extension of F, we know that $N(D^\times) = F^\times$. We may in fact assume that ξ is a prime element of O. Then it is easy to see

$$R_0 \supset \left\{ \alpha(a,b) \middle| a, b \in O_K \right\} = R$$

for the integer ring O_K of K. For $\omega = \alpha(0,1)$, we have $\omega R \omega^{-1} = R$, $\omega^2 = \xi$ and $N(\omega) = -\xi$. Thus for each $\alpha \in D^\times$, we can find the minimal exponent $w(\alpha) \in \mathbb{Z}$ such that $\alpha\omega^{-w(\alpha)} \in R$. Then w is a sort of an additive valuation: it satisfies

$$w(\alpha\beta) = w(\alpha) + w(\beta) \text{ and } w(\alpha+\beta) \geq \min(w(\alpha), w(\alpha)).$$

We put $w(0) = \infty$. Then $R = \{\alpha \in D | w(\alpha) \geq 0\} \supset R_0$, which is the (non-commutative) "valuation" ring of w. This shows that $R_0 = R$ is an order. Since on any order R', N has values in O, we know $R_0 \supset R'$. Thus R_0 is the unique maximal order O_D.

Proposition 1.2. *If F is local and D is a division algebra, then all O-orders are contained in one and only one maximal order O_D.*

In general, we call an order R in D maximal if there is no order containing R properly. Thus there may be several maximal orders. Suppose now that $D = M_2(F)$. Let L be an O–lattice of F^2. We put $R = R_L = \{\alpha \in M_2(F) | \alpha L \subset L\}$. Then R is an O–order. In fact, we can find a basis (x, y) of L. Since x and y are column vectors, we consider $X = (x, y)$ as an invertible matrix in $M_2(F)$. Then defining the regular representation $\rho : D \cong M_2(F)$ by $X\rho(\alpha)^t X = {}^t X\alpha$, we have

$$\rho(\alpha) \in M_2(O) \iff \alpha \in R_L.$$

Therefore, we have $R_L = X^{-1}M_2(O)X$. Conversely, for any given O–order R in D, we put $L = \pi(R)$ for the projection:

$$\pi : D = M_2(F) \ni \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right) \in F^2.$$

Then L is an O–lattice of F^2 and $R \cdot L \subset L$. Namely $R \subset R_L$. Thus for any order R of $M_2(F)$, there exists a maximal order R_L of the form $X^{-1}M_2(O)X$.

Proposition 1.3. *If F is a p–adic local field and $D = M_2(F)$, for any given order R, there exists a maximal order O_D containing R which is a conjugate of the standard maximal order $M_2(O)$.*

Thus all the maximal orders in $M_2(F)$ (for local F) are conjugate to each other.

Corollary 1.4. *The group $GL_2(O)$ is a maximal compact subgroup of $GL_2(F)$. If K is a maximal compact subgroup of $GL_2(F)$, then K is a conjugate of $GL_2(O)$ in $GL_2(F)$.*

Proof. Let $K_0 = GL_2(O)$. A double coset K_0xK_0 $(x \in GL_2(F))$ is of the form $K_0 \left(\begin{smallmatrix} \varpi^a & 0 \\ 0 & \varpi^b \end{smallmatrix}\right) K_0$ for a suitable integer $a \geq b$ by the theory of elementary divisors. Thus, the subgroup generated by K_0 and any x outside K_0 contains $\left(\begin{smallmatrix} \varpi^a & 0 \\ 0 & \varpi^b \end{smallmatrix}\right)$ with one of a and b nonzero. Then plainly the subgroup is not compact; so, K_0 is a maximal compact subgroup. Let $L = O^2$. Since $K \backslash K \cdot K_0$ is discrete and compact, it is finite. Then $L' = \sum_{u \in K\backslash K\cdot K_0} u(L)$ is isomorphic to O^2 as O-modules, and $R = \{x \in M_2(F) | xL' \subset L'\}$ is a maximal order of $M_2(F)$, which is compact. Thus $R^\times$ is a compact subgroup. Since $R^\times \supset K$, they are equal by the maximality of K. Since $R = g \cdot M_2(O)g^{-1}$, we have $K = R^\times = g \cdot K_0 g^{-1}$ for $g \in GL_2(F)$. $\square$

We now suppose that F is a number field. We take A to be the integer ring O of F. For any (finite dimensional) vector space V over F, we fix a base $x_1, \ldots, x_r$ and identify

V with F^r. We put $L_0 = O \cdot x_1 + \cdots + O \cdot x_r$, which is an O–lattice. Consider any O–lattice L. By definition, $L = O \cdot y_1 + \cdots + O \cdot y_m$ with $m \geq r$. Since $L_0 \otimes_O F = V$, we can find $\alpha, \beta \in F$ such that $\alpha \cdot L_0 \subset L \subset \beta \cdot L_0$. For each prime ideal $\mathfrak{p}$ of O, we write $O_\mathfrak{p}$ for the $\mathfrak{p}$–adic completion of O and $F_\mathfrak{p} = F \otimes_O O_\mathfrak{p}$. Then for almost all $\mathfrak{p}$, α and β are both $\mathfrak{p}$–adic units, and thus

$$L_{0,\mathfrak{p}} = \alpha \cdot L_{0,\mathfrak{p}} = \alpha O_\mathfrak{p} \otimes_O L_0 \subset L_\mathfrak{p} \subset \beta \cdot L_{0,\mathfrak{p}} = L_{0,\mathfrak{p}}.$$

This shows that for almost all $\mathfrak{p}$, $L_\mathfrak{p} = L_{0,\mathfrak{p}}$. Conversely, let $\{L_\mathfrak{p}\}_\mathfrak{p}$ be a family of $O_\mathfrak{p}$–lattices indexed by all prime ideals of O. We suppose that $L_\mathfrak{p} = L_{0,\mathfrak{p}}$ for almost all $\mathfrak{p}$. Such a collection $\{L_\mathfrak{p}\}_\mathfrak{p}$ is called admissible. This definition of admissibility does not depend on the choice of starting lattice L_0, because for any O–lattice, its $\mathfrak{p}$–adic completion is the same for almost all $\mathfrak{p}$. We now show that for any given admissible family $\{L_\mathfrak{p}\}_\mathfrak{p}$ of local lattices, there is a unique O–lattice L in V which gives rise to the given collection. We first take α in $F^\times$ so that

$$L_\mathfrak{p} \supset \alpha \cdot L_{0,\mathfrak{p}} \ \text{ for all } \mathfrak{p}.$$

We can always find such α, because $L_{0,\mathfrak{p}}$ and $L_\mathfrak{p}$ are different for only finitely many $\mathfrak{p}$. Since $V/\alpha \cdot L_0 \cong (F/\alpha O)^r$, we have a unique finite subgroup X in $V/\alpha \cdot L_0$ corresponding to $L_\mathfrak{p}/\alpha \cdot L_{0,\mathfrak{p}}$. Put $L = \{v \in V | v \mod \alpha \cdot L_0 \in X\}$. By definition, L satisfies the required property, and we have

$$L = \left(\prod_\mathfrak{p} L_\mathfrak{p}\right) \bigcap V$$

in $V(\mathbb{A}^{(\infty)}) = V \otimes_F \mathbb{A}^{(\infty)}$ for the finite part $\mathbb{A}^{(\infty)}$ of the adele ring $\mathbb{A}$ of $\mathbb{Q}$.

We apply the above argument to $V = D$ for a quaternion algebra D. Let R be an O-order of D. Then $R_\mathfrak{p}$ is an $O_\mathfrak{p}$–order of $D_\mathfrak{p}$. First suppose that $D = M_2(F)$. Then $M_2(O_\mathfrak{p})$ is maximal at every $\mathfrak{p}$, $M_2(O)$ is maximal. Thus for any order R of D, $R_\mathfrak{p} = M_2(O_\mathfrak{p})$ for almost all $\mathfrak{p}$ and for finitely many $\mathfrak{p}$ with $R_\mathfrak{p} \neq M_2(O_\mathfrak{p})$, we can find $x_\mathfrak{p} \in D_\mathfrak{p}^\times$ such that $x_\mathfrak{p} M_2(O_\mathfrak{p}) x_\mathfrak{p}^{-1} = R_\mathfrak{p}$. For other primes $\mathfrak{p}$, we simply put $x_\mathfrak{p} = 1$. Thus $x = (x_\mathfrak{p})_\mathfrak{p} \in D_\mathbb{A}^\times$ (the adelization of D). The family $\{x_\mathfrak{p} M_2(O_\mathfrak{p}) x_\mathfrak{p}^{-1}\}$ is admissible, and therefore there exists an O–lattice O_D in D such that $O_{D,\mathfrak{p}} = x_\mathfrak{p} M_2(O_\mathfrak{p}) x_\mathfrak{p}^{-1}$ for all $\mathfrak{p}$. Since $O_D = D \bigcap \prod_\mathfrak{p} x_\mathfrak{p} M_2(O_\mathfrak{p}) x_\mathfrak{p}^{-1}$ in $D_\mathbb{A}^{(\infty)}$, O_D is a subring; namely, O_D is an O–order. Since $R_{D,\mathfrak{p}}$ is maximal for all $\mathfrak{p}$, O_D has to be maximal and $O_D \supset R$.

Now suppose that D is a division algebra over a number field F. We embed D into $M_2(K)$ for a quadratic extension K/F. Let A be the integer ring of K. Then $R = M_2(A) \cap D$ is an order of D. We shall show that $R_\mathfrak{p}$ is a maximal order for almost all $\mathfrak{p}$. We may assume that $D = D_\xi$. If $K_\mathfrak{p} = K \otimes_F F_\mathfrak{p} \cong F_\mathfrak{p} \oplus F_\mathfrak{p}$, then obviously $R_\mathfrak{p} \cong M_2(A_\mathfrak{p})$ if ξ is a p–adic unit (which is true for almost all $\mathfrak{p}$). Suppose that $K_\mathfrak{p}/F_\mathfrak{p}$ is a field extension and $D_\mathfrak{p} \cong M_2(F_\mathfrak{p})$. Then ξ is an integral norm: $\xi = xx^\sigma$ for $x \in A_\mathfrak{p}^\times = (A \otimes_O O_\mathfrak{p})^\times$. This is true for almost all $\mathfrak{p}$. Conjugating D_ξ by $\alpha = \left(\begin{smallmatrix} x_\mathfrak{p} & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we may assume that $\xi = 1$. Then conjugating by $\beta = \left(\begin{smallmatrix} \delta & -\delta \\ 1 & 1 \end{smallmatrix}\right)$ for $\delta \in K$ with $\delta^2 \in F^\times$, we know $R_\mathfrak{p} \ni r \mapsto \beta\alpha r \alpha^{-1}\beta^{-1}$ induces an isomorphism of $R_\mathfrak{p}$ with $M_2(O_\mathfrak{p})$ if $\beta\alpha \in GL_2(A_\mathfrak{p})$. Since $\beta\alpha$ falls in $GL_2(A_\mathfrak{p})$ for almost all $\mathfrak{p}$, $R_\mathfrak{p}$ is maximal for almost all $\mathfrak{p}$. Thus if one gives oneself a collection of maximal orders $\{O_{D,\mathfrak{p}}\}_{\mathfrak{p}}$ such that $O_{D,\mathfrak{p}} = R_{\mathfrak{p}}$ for almost all $\mathfrak{p}$, there exists a unique maximal order O_D in D whose completion is equal to the given data $O_{D,\mathfrak{p}}$. Conversely, if R is any A-order of D, then $R_{\mathfrak{p}}$ is maximal for almost all $\mathfrak{p}$. We put $O_{D,\mathfrak{p}} = R_{\mathfrak{p}}$ if $R_{\mathfrak{p}}$ is maximal and choose a maximal order $O_{D,\mathfrak{p}}$ in $D_{\mathfrak{p}}$ containing $R_{\mathfrak{p}}$ if $R_{\mathfrak{p}}$ is not maximal. Then the family $\{O_{D,\mathfrak{p}}\}_{\mathfrak{p}}$ gives rise to a unique maximal order O_D in D containing R. We have

Proposition 1.5. *Let D be a quaternion algebra over a number field or a p–adic local field. Then for any given order R of D, there exists a maximal order containing R.*

From now on, D is any quaternion algebra over a number field F (including $M_2(F)$). Fix a maximal order O_D of D. Then for any other maximal order O'_D in D and for each prime $\mathfrak{p}$, we can find $x_{\mathfrak{p}} \in D_{\mathfrak{p}}^{\times}$ such that $x_{\mathfrak{p}} O_{D,\mathfrak{p}} x_{\mathfrak{p}}^{-1} = O'_{D,\mathfrak{p}}$. Since $O'_{D,\mathfrak{p}} = O_{D,\mathfrak{p}}$ for almost all $\mathfrak{p}$, $x_{\mathfrak{p}} \in O_{D,\mathfrak{p}}^{\times}$ for almost all $\mathfrak{p}$. Therefore $x = (x_{\mathfrak{p}})_{\mathfrak{p}} \in D_{\mathbb{A}}^{\times}$. Thus we have a bijection:

$$\{\text{maximal orders in } D\} \leftrightarrow D_{\mathbb{A}}^{\times}/U \cdot D_{\infty}^{\times}$$

for $U = \prod_{\mathfrak{p}} O_{D,\mathfrak{p}}^{\times}$ and $D_\infty = \prod_\sigma D_\sigma \cong D \otimes_{\mathbb{Q}} \mathbb{R}$ for σ running over all archimedean places of F, since $x_{\mathfrak{p}} O_{D,\mathfrak{p}} x_{\mathfrak{p}}^{-1} = O_{D,\mathfrak{p}}$ for all $\mathfrak{p}$ if $x \in U \cdot D_{\infty}^{\times}$. If $x \in D^{\times}$, then the conjugation $O_D \ni r \mapsto xrx^{-1} \in O'_D$ is well defined and hence $O_D \cong O'_D$. Thus we have a surjection:

$$D^{\times} \backslash D_{\mathbb{A}}^{\times}/U \cdot D_{\infty}^{\times} \twoheadrightarrow \{\text{isomorphism classes of maximal orders in } D\}.$$

Let $I = \mathrm{Hom}_{\text{field}}(F, \mathbb{C})$. Now we need to quote some deeper results:

Theorem 1.6 (Norm theorem)**.** *Let*

$$F_D^{\times} = \left\{ x \in F^{\times} \middle| x^{\sigma} > 0 \quad \text{if } D \otimes_{F,\sigma} F_\sigma \cong \mathbb{H} \text{ for } \sigma \in I \right\}.$$

Then $N(D^{\times}) = F_D^{\times}$. Especially $F^{\times}/N(D^{\times}) \cong \{\pm 1\}^r$, where r is the number of infinite place at which D is ramified.

A proof of this theorem is in Weil's book: [BNT] Proposition 3 in page 206. The proof given there is 4 pages long but quite elementary and can be read without reading much the material in the earlier sections of [BNT] (basic algebraic number theory suffices for that).

Theorem 1.7 (Approximation theorem)**.** *The set $D^{\times} \backslash D_{\mathbb{A}}^{\times}/U \cdot D_{\infty}^{\times}$ is a finite set. Especially, isomorphism classes of maximal orders of D are finitely many.*

A proof in the case where $D = M_2(F)$ is in [LFE] Section 9.1. An outline of the proof for division algebras D is as follows. We consider

$$D_{\mathbb{A}}^1 = \left\{ x \in D_{\mathbb{A}} \middle| |N(x)|_{\mathbb{A}} = 1 \right\}.$$

By the product formula: $|\xi|_{\mathbb{A}} = 1$ for $\xi \in F^{\times}$ (e.g., [LFE] (8.1.5)), $D^{\times}$ is a subgroup of $D_{\mathbb{A}}^1$. Then $D^{\times}$ can be shown to be a discrete subgroup of $D_{\mathbb{A}}^1$ and $D^{\times} \backslash D_{\mathbb{A}}^1$ is compact (see [MFM] Lemma 5.2.4). A similar assertion for number fields is also true, i.e.,

$$F^{\times} \text{ is discrete in } F_{\mathbb{A}}^1 = \left\{ x \in F_{\mathbb{A}} \middle| |x|_{\mathbb{A}} = 1 \right\}, \text{ and } F_{\mathbb{A}}^1/F^{\times} \text{ is compact.}$$

A proof of this fact for fields F can be found in [LFE] Theorem 8.1.1. All the arguments in the proof there works well for *division* algebras D replacing $|x|_{\mathbb{A}}$ by $|N(x)|_{\mathbb{A}}$. Then $D^1_{\mathbb{A}}/(U \cdot D^\times_\infty \cap D^1_{\mathbb{A}})$ is discrete because $U \cdot D^\times_\infty \cap D^1_{\mathbb{A}}$ is an open subgroup of $D^1_{\mathbb{A}}$. Thus $D^\times \backslash D^1_{\mathbb{A}}/(U \cdot D^\times_\infty \cap D^1_{\mathbb{A}})$ is discrete and compact and hence is finite. Note that

$$D^\times \backslash D^1_{\mathbb{A}}/\left(U \cdot D^\times_\infty \cap D^1_{\mathbb{A}}\right)$$

is the kernel of the norm map

$$N : D^\times \backslash D^\times_{\mathbb{A}}/U \cdot D^\times_\infty \to F^\times_{\mathbb{A}}/F^\times_D U_F \quad (U_F = \prod_{\mathfrak{p}} O^\times_{\mathfrak{p}}),$$

whose right-hand side is a ray class group (e.g., [LFE] Corollary 8.1.1]) which is finite. This shows the above theorem. When $D = M_2(F)$ (or more generally, if $D \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic to a product of copies of $M_2(\mathbb{R})$), by the following theorem, $|D^\times \backslash D^\times_{\mathbb{A}}/U \cdot D^\times_\infty|$ is equal to the class number of F (see (CL) below).

Theorem 1.8 (Strong approximation theorem)**.** *Let O_D be a maximal order of D. Let v be one place (either finite or archimedean) such that $D_v \cong M_2(F_v)$. Let $U^{(v)} = \prod_{\mathfrak{p} \neq v} O^\times_{D,\mathfrak{p}}$, where we put $O_{D,\sigma} = M_2(F_\sigma)$ if v is the infinite place $\sigma \in I$. Then*

$$\Gamma^{(v)} = \left\{\gamma \in D^\times \middle| N(\gamma) = 1 \text{ and } \gamma \in O^\times_{D,\mathfrak{p}} \text{ for all } \mathfrak{p} \neq v\right\}$$

is dense in $\{x \in U^{(v)} | N(x) = 1\}$.

In other words, for any given $x \in U^{(v)}$, we can find $\gamma \in \Gamma^{(v)}$ such that $\gamma \equiv x \mod \mathfrak{N} \cdot \widehat{O}^{(v)}_D$ for any ideal $\mathfrak{N}$ prime to v and $\gamma^{(v)}$ is arbitrarily close to x_σ for all infinite place $\sigma \neq v$, where $\widehat{O}^{(v)}_D = \prod_{\mathfrak{p} \neq v} O_{D,\mathfrak{p}}$. When v is a infinite place, an elementary proof can be found in Miyake's book [MFM] Theorem 5.2.10. Although Miyake gives a proof assuming that $F = \mathbb{Q}$, his argument works well for general number fields without much modification.

We see easily from the strong approximation theorem that the reduced norm $N : D \to F$ induces, for any open subgroup S of U with $S_F = N(S) \subset F^\times_{\mathbb{A}^{(\infty)}}$,

$$D^\times \backslash D^\times_{\mathbb{A}}/S \cdot D^\times_\infty = D^\times \backslash D^\times_{\mathbb{A}^{(\infty)}}/S \cong F^\times_{\mathbb{A}^{(\infty)}}/S_F \cdot F^\times_D \tag{CL}$$

if D has at least one infinite place v such that $D_v \cong M_2(F_v)$. The cardinality of the set on the left hand side of (CL) is called the class number of D and is equal to the number of the equivalence classes of the set {fractional right O_D–ideals in D} modulo the following equivalence relation $\mathfrak{a} \sim \mathfrak{b} \iff \mathfrak{a} = \alpha\mathfrak{b}$ for $\alpha \in D^\times$. Thus often the class number of D is given by that of a ray class group of F. However when $D \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic to $\mathbb{H}^I = \prod_{\sigma \in I} \mathbb{H}$ (such a quaternion algebra is called a *definite* quaternion algebra), the class number of D cannot be given by the class number of F. If D is a definite quaternion algebra over $\mathbb{Q}$ only ramifying at a prime p, one of our goals in this course is to prove that the class number of D is equal to the dimension of the space of holomorphic modular forms on $\Gamma_0(p)$ of weight 2.

Exercise 1.9.

(1) *Let F be a number field. For an O–lattice L of F^n, prove that there exists an O–ideal $\mathfrak{a}$ such that $L \cong O^{n-1} \oplus \mathfrak{a}$ as O–modules. Hint: putting $\widehat{L} = \prod_{\mathfrak{p}} L_{\mathfrak{p}}$, find $x \in GL_n(F_{\mathbb{A}})$ such that $x \cdot \widehat{L} = \widehat{O}^{n-1} \oplus \widehat{\mathfrak{a}} \subset (F^n \otimes_{\mathbb{Q}} \mathbb{A}^{(\infty)})$.*
(2) *For a given $\alpha \in GL_2(F)$, prove that $\alpha \in GL_2(O_{\mathfrak{p}})$ for almost all $\mathfrak{p}$.*
(3) *Using the outcome of* Exercise 1*, show that the class number of F gives the number $|GL_n(F)\backslash GL_n(F_{\mathbb{A}})/GL_n(\widehat{O})GL_n(F_\infty)|$.*
(4) *Prove* (CL) *using* Theorem 1.8*.*
(5) *Prove the statement just after* (CL) *about class number of D.*
(6) *Suppose that F is totally real. For each embedding $\sigma : F \hookrightarrow \mathbb{R}$, define $D_\sigma = D \otimes_{F,\sigma} \mathbb{R}$. Prove that $D \otimes_{\mathbb{Q}} \mathbb{R} \cong \prod_\sigma D_\sigma$. Describe what happens if F is not totally real.*

2. Automorphic Forms on Quaternion Algebras

In this section, we recall the definition of holomorphic automorphic forms on the multiplicative idele group of a quaternion algebra. Following the tradition of Gauss, Eisenstein, Kronecker and Hilbert, if $D = M_2(F)$ (for a totally real field F), such functions are called *modular forms*. On the other hand, general quaternionic cases are more recent, for which we use a more general term: *automorphic forms* (see [?] for the distinction of modular and automorphic forms).

2.1. **Arithmetic quotients.** Let F be a totally real field and I be the total set of embeddings of F into $\overline{\mathbb{Q}}$ (the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$). Recall the embeddings $i_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ fixed in the introduction. Since F is totally real, $i_\infty \circ \sigma$ has image in $\mathbb{R}$ for all $\sigma \in I$. As before, O denotes the integer ring of F. We fix a quaternion algebra D/F. Let I_D be a subset of I consisting of $\sigma : F \hookrightarrow \mathbb{R}$ such that at the infinite place $i_\infty \circ \sigma$ (which we write again σ for simplicity), $D_\sigma = D \otimes_{F, i_\infty \circ \sigma} \mathbb{R} \cong M_2(\mathbb{R})$. We write Σ^D for the set of all places at which D ramifies and I^D for the subset of Σ^D consisting of infinite places (thus $I = I_D \sqcup I^D$).

We fix a maximal order O_D of D. Take a quadratic extension K/F inside D such that $K_\sigma = K \otimes_F F_\sigma \cong (\mathbb{R} \oplus \mathbb{R})$ if $\sigma \in I_D$, and fix $\rho : (D \otimes_F K) \cong M_2(K)$. Then, $K_\sigma \cong \mathbb{C}$ if $\sigma \in I^D$. We may assume that $\rho(O_D)$ is contained in $M_2(O_K)$ for the integer ring O_K, because every O–order in $M_2(K)$ is contained in an adelic conjugate of $M_2(O_K)$ (Proposition 1.5). Since O_D is a $\mathbb{Z}$–lattice of $D_\infty = D \otimes_{\mathbb{Q}} \mathbb{R}$, O_D is a discrete subset of D_∞. Then we see $O_D^\times$ is a discrete subset of $D_\infty^\times$. Let $D_\infty^1 = \{x \in D_\infty | N(x) = 1\}$. We know that the natural map:

$$D_\infty^1 \times F_\infty^\times \to D_\infty^\times : (x, y) \mapsto xy$$

has finite kernel ($\cong \{\pm 1\}^I$) and cokernel. Put $O_D^{(1)} = O_D^\times \cap D_\infty^1$. Therefore $O_D^{(1)}$ is discrete in D_∞^1. The natural map from D_∞^1 to $PD_{\infty+}^\times = D_{\infty+}^\times / F_\infty^\times$ is surjective and has finite kernel (again $\{\pm 1\}^I$), where $D_{\infty+}^\times$ is the subgroup of $D_\infty^\times$ consisting of elements with totally positive norm. The image of $O_D^{(1)}$ in $PD_{\infty+}^\times$ is discrete. We have an exact sequence:

$$1 \to O_D^{(1)} \to O_D^\times \xrightarrow{N} O^\times \to \mathrm{Coker}(N) \to 1.$$

The image of the norm map N contains $(O^\times)^2$. Since $O^\times$ is a finitely generated abelian group by Dirichlet's theorem (e.g., [LFE] Theorem 1.2.3), $(O^\times : (O^\times)^2)$ is finite (a power of 2). Thus $\mathrm{Coker}(N)$ is finite. This shows that the image $\overline{O}^\times_{D+}$ of $O^\times_{D+} = O^\times_D \cap D^\times_{\infty+}$ in $PD^\times_{\infty+}$ (isomorphic to $O^\times_D/O^\times$) has the image of $O^{(1)}_D$ as a subgroup of finite index, and $\overline{O}^\times_{D+}$ is discrete in $PD^\times_{\infty+}$. Let C_∞ be the maximal compact subgroup of D^1_∞. Since

$$D^1_\infty \cong SL_2(\mathbb{R})^{I_D} \times \mathbb{H}_1{}^{I^D}$$

for $\mathbb{H}_1 = \{x \in \mathbb{H} | N(x) = 1\}$, we have $C_\infty \cong SO_2(\mathbb{R})^{I_D} \times \mathbb{H}_1{}^{I^D}$. Thus

$$D^\times_{\infty+}/F^\times_\infty C_\infty = D^1_\infty/C_\infty \cong \mathfrak{H}^{I_D}$$

via $g \mapsto g(\mathbf{i})$ for $\mathbf{i} = (\sqrt{-1}, \dots, \sqrt{-1}) \in \mathfrak{H}^{I_D}$ ($\mathfrak{H} = \{z \in \mathbb{C} | \operatorname{Im}(z) > 0\}$). For simplicity, we write $\mathfrak{Z}_D$ for $\mathfrak{H}^{I_D}$. Since $O^{(1)}_D$ is discrete, $O^{(1)}_D \cap C_\infty$ is discrete and compact and hence finite. In particular, if $I_D = \emptyset$, the reduced norm map $N : D_\infty \to \mathbb{R}$ is a positive definite quadratic form (on $\mathbb{H}_{/\mathbb{R}}$, it is sum of four squares), and $PD^\times_\infty$ is compact. Thus we have

$$\text{If } I_D = \emptyset,\ O^\times_D/O^\times \text{ is a finite group.} \tag{2.1}$$

We say that a subgroup Γ of $PD^\times_{\infty+}$ acts *properly discontinuously* on $\mathfrak{Z}_D$, if for any point $z \in \mathfrak{Z}_D$, we can find an open neighborhood U of z such that $\{\gamma \in \Gamma | \gamma(U) \cap U \neq \emptyset\}$ is finite. If this set is a singleton made up of the identity element for all z, the action is called *free*. The quotient $\Gamma\backslash\mathfrak{Z}_D$ for Γ acting freely is a complex manifold. When the action is only properly discontinuous, the quotient is a complex analytic space (locally isomorphic to a zero set of finitely many complex analytic functions in $\mathbb{C}^n$). The group $O^{(1)}_D$ acts properly discontinuously on $\mathfrak{Z}_D$ (cf. [MFM] Section 1.5) because $O^{(1)}_D \cap C_\infty$ is finite, and hence $\overline{O}^\times_{D+}$ also acts properly discontinuously on $\mathfrak{Z}_D$. We now claim:

Proposition 2.1. *For any subgroup Γ of finite index of $O^{(1)}_D$ or $\overline{O}^\times_{D+}$, the quotient analytic space $\Gamma\backslash\mathfrak{Z}_D$ is compact if D is a division algebra.*

Proof. We only need to prove the proposition for $O^{(1)}_D\backslash\mathfrak{Z}_D$ because $\Gamma\backslash\mathfrak{Z}_D$ is a covering of $O^{(1)}_D\backslash\mathfrak{Z}_D$ with finite fiber (the number of elements in the fiber is less than or equal to the index of Γ in $O^{(1)}_D$ or $\overline{O}^\times_{D+}$) and $\overline{O}^\times_{D+}\backslash\mathfrak{Z}_D$ is covered by $O^{(1)}_D\backslash\mathfrak{Z}_D$. We know that $D^\times\backslash D^1_{\mathbb{A}}$ is compact for $D^1_{\mathbb{A}} = \big\{x \in D_{\mathbb{A}} \big| |N(x)|_{\mathbb{A}} = 1\big\}$ (see 1.2). Thus $D^{(1)}\backslash D^{(1)}_A/U^{(1)}C_\infty$ is compact, where $D^{(1)}_{\mathbb{A}} = \{x \in D_{\mathbb{A}} | N(x) = 1\}$, $U = \prod_{\mathfrak{p}} O^\times_{D,\mathfrak{p}}$, $U^{(1)} = D^{(1)}_{\mathbb{A}} \cap U$ and $D^{(1)} = D^{(1)}_{\mathbb{A}} \cap D$. We consider the map $\iota : \mathfrak{Z}_D \to D^{(1)}_{\mathbb{A}}/U^{(1)}C_\infty$ given by $\iota(z) = g_\infty \mod C_\infty$ for g_∞ with $g_\infty(\mathbf{i}) = z$. Then it is easy to see that ι induces an inclusion $O^{(1)}_D\backslash\mathfrak{Z}_D$ into $D^\times\backslash D^1_{\mathbb{A}}/UC_\infty$. By the strong approximation theorem, ι is surjective and hence an isomorphism. $\square$

Exercise 2.2. *Prove that the map ι in the above proof is an embedding.*

By the above proposition, if D is a division algebra, $\Gamma\backslash\mathfrak{Z}_D$ has no cusps. Let

$$O^\times_D(\mathfrak{N}) = \big\{\gamma \in O^\times_D \big| \gamma - 1 \in \mathfrak{N}O_D\big\}$$

for each ideal $\mathfrak{N}$ of O and $O_D^{(1)}(\mathfrak{N}) = O_D^\times(\mathfrak{N}) \cap O_D^{(1)}$. We put $\Gamma(\mathfrak{N})$ to be the intersection with $\overline{O}_D^\times = \Gamma(1)$ with the image of $O_D^\times(\mathfrak{N})$ in $PD_{\infty+}^\times$. Since $O_D^{(1)}$ acts properly discontinuously on $\mathfrak{Z}_D$, for each point $z \in \mathfrak{Z}_D$, the stabilizer $O_{D,z}^{(1)} = \{\gamma \in O_D^{(1)} | \gamma(z) = z\}$ of z is a finite group. In particular, when D is definite, $\Gamma(\mathfrak{N})$ and $O_D^{(1)}$ are finite. Take a non-central element $\zeta \in O_{D,z}^{(1)}$. Then $\zeta^m = 1$ for some $m > 2$. Thus $F[\zeta]$ is a totally imaginary quadratic extension of F in D. There are only finitely many such quadratic extensions over F generated by roots of unity. Thus the order m of ζ is bounded independently of z. Since m–th roots of unity for a given $m > 2$ can be separated modulo $\mathfrak{N}$ for sufficiently small ideal $\mathfrak{N}$ in O, $O_D^{(1)}(\mathfrak{N})$ acts fixed point free on $\mathfrak{Z}_D$. Any subgroup of Γ containing $O_D^{(1)}(\mathfrak{N})$ for some ideal $\mathfrak{N}$ is called a congruence subgroup of $D^\times$. Thus

(FR) *We can find a congruence subgroup acting on $\mathfrak{Z}_D$ without fixed point.*

Actually we can give an exact lower bound for $\mathfrak{N}$ when $\Gamma(\mathfrak{N})$ acts freely on $\mathfrak{Z}_D$ (e.g., [H88] Lemma 7.1]). In particular, if $\mathfrak{N} \subset (3)$, $O_D^{(1)}(\mathfrak{N})$ acts freely on $\mathfrak{Z}_D$.

A natural question is

(CS) *Are all subgroups of finite index of $O_D^{(1)}$ congruence subgroups?*

This is the congruence subgroup problem for $D^\times$ (see [R] for a survey of the problem for general semi-simple groups). It is conjectured by Serre that if the number $r = |I_D|$ of infinite places of F unramified in D is bigger than or equal to 2, the answer should be affirmative. Serre proved this to be affirmative for $M_2(F)$ if $F \neq \mathbb{Q}$ (that is, $r = [F : \mathbb{Q}] \geq 2$). When $r = 1$, for small enough $\mathfrak{N}$, $X = \Gamma(\mathfrak{N}) \backslash \mathfrak{Z}_D$ is a compact Riemann surface of genus $g \geq 1$, and $H^1(X, \mathbb{Z}) = \mathrm{Hom}(\pi_1(X), \mathbb{Z}) \cong \mathbb{Z}^{2g}$. Thus the maximal abelian quotient $\pi_1^{ab}(X)$ of the fundamental group $\pi_1(X)$ is infinite. On the other hand, it is easy to show that the maximal abelian quotient of $\Gamma(\mathfrak{N})$ is finite if (CS) is affirmative in this case. Thus (CS) has a negative answer when $r = 1$.

Exercise 2.3. *Prove that the maximal abelian quotient of* $\Gamma(\mathfrak{N})$ *is finite, assuming* (CS) *is affirmative.*

We associate with the algebra D an algebraic group $\underline{D}^\times$ defined over F. As a group functor, $\underline{D}^\times(R) = (D \otimes_F R)^\times$ for all F-algebra R. Then we consider an algebraic group G^D defined over $\mathbb{Q}$ given by $G^D = \mathrm{Res}_{F/\mathbb{Q}} \underline{D}^\times$. If we fix a maximal order O_D, we can extend $\underline{D}^\times$ to a group scheme defined over the integer ring O of F by $\underline{D}^\times(R) = (O_D \otimes_O R)^\times$ for all O-algebras R. Thus G^D extends to a group scheme over $\mathbb{Z}$ by taking $G_{/\mathbb{Z}}^D = \mathrm{Res}_{O/\mathbb{Z}} \underline{D}_{/O}^\times$ (see [HMI, Theorem 2.16]). The center Z of G^D is an algebraic group satisfying $Z(A) = (O \otimes_{\mathbb{Z}} A)^\times$; so, it is independent of D and $Z = \mathrm{Res}_{O/\mathbb{Z}} \mathbb{G}_{m/O}$.

Exercise 2.4. *Prove that the functor $\underline{D}_{/F}^\times$ is an affine algebraic group over F.*

Let $G^D(\mathbb{R})^+$ be the identity connected component of the real Lie group $G^D(\mathbb{R})$; then, $G^D(\mathbb{R})^+ = \{x \in G^D(\mathbb{R}) | N(x) \gg 0\}$. We let $g \in G^D(\mathbb{R})$ act on $\mathfrak{Z}_D = \mathfrak{H}^{I_D}$ by the linear fractional transformation of

$$g_\sigma = \sigma(g) \in GL_2(K \otimes_{K,\sigma} \mathbb{R}) = GL_2(\mathbb{R})$$

component-wise. Write $C_{\sigma+}$ for the stabilizer of $\sqrt{-1}$ in $(D\otimes_{F,\sigma}\mathbb{R})^\times$ and define a closed subgroup $C^D_{\mathbf{i}} \subset G^D(\mathbb{R})$ by $Z(\mathbb{R})\cdot\left(\prod_{\sigma\in I_D} C_{\sigma+}\times\prod_{\sigma\in I^D}\mathbb{H}^\times\right)$, which is the stabilizer of $\mathbf{i}=(\sqrt{-1},\dots,\sqrt{-1})\in\mathfrak{H}^{I_D}$ in the connected component $G^D(\mathbb{R})^+$. Thus we have $\mathfrak{Z}_D=\mathfrak{H}^{I_D}\cong G^D(\mathbb{R})^+/C^D_{\mathbf{i}}$ by $g(\mathbf{i})\leftrightarrow g$ for the identity-connected component $G^D(\mathbb{R})^+$ of $G^D(\mathbb{R})$. Write simply $G=G^{M_2(F)}=\mathrm{Res}_{O/\mathbb{Z}}GL(2)$. Since G and G^D have the common center Z canonically isomorphic to $\mathrm{Res}_{F/\mathbb{Q}}\mathbb{G}_m$, we use the same symbol Z to indicate the center of G^D independently of D. For any open compact subgroup $S\subset G^D(\mathbb{A}^{(\infty)})$ under the adele topology, we think of the *automorphic manifold* associated with the subgroup S:

$$Y(S)=Y^D(S):=G^D(\mathbb{Q})\backslash G^D(\mathbb{A})/Z(\mathbb{A})S\cdot C^D_{\mathbf{i}}.$$

Exercise 2.5. *Let* $\Gamma_g(S)=\left(g^{-1}S\cdot G^D(\mathbb{R})^+g\right)\cap D^\times$.

(1) *Show that* $\Gamma_g(S)$ *is a discrete subgroup of* $G^D(\mathbb{R})$*;*
(2) *Prove that* $Y(S)\cong\bigsqcup_g\Gamma_g(S)\backslash\mathfrak{Z}_D$ *via the isomorphism*

$$G^D(\mathbb{Q})\backslash G^D(\mathbb{Q})gZ(\mathbb{A})S\cdot G^D(\mathbb{R})^+/Z(\mathbb{A})S\cdot C^D_{\mathbf{i}}\cong\Gamma_g(U)\backslash\mathfrak{Z}_D$$

given by $gx_\infty\mapsto(gx)_\infty(\mathbf{i})\in\mathfrak{Z}_D$ *if* $D^\times_{\mathbb{A}}=\bigsqcup_g G^D(\mathbb{Q})g\cdot Z(\mathbb{A})S\cdot D^\times_{\infty+}$ *by* Theorem 1.7.
(3) *Prove that* $Y(S)$ *is a complex analytic space of dimension* $r=|I_D|$ *and is a complex manifold if* S *is sufficiently small.*

2.2. **Archimedean Hilbert modular forms.** Let us recall the definition of the adelic Hilbert modular forms and their Hecke ring of level $\mathfrak{N}$ for an integral ideal $\mathfrak{N}$ of F (cf. [HMI, §2.3.2]). Thus in this subsection, $D=M_2(F)$ for a totally real field F, and $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)^\iota=\left(\begin{smallmatrix}d&-b\\-c&a\end{smallmatrix}\right)$.

Let $G=\mathrm{Res}_{O/\mathbb{Z}}GL(2)$ as an algebraic group over $\mathbb{Z}$; $G(A)=GL_2(A\otimes_{\mathbb{Z}}O)$ for each commutative ring A. Let $T_0=\mathbb{G}^2_{m/O}$ be the diagonal torus of $GL(2)_{/O}$, and put, $T=\mathrm{Res}_{O/\mathbb{Z}}\mathbb{G}_m$ and $T_G=\mathrm{Res}_{O/\mathbb{Z}}T_0$. Then T_G contains the center Z of G. Write $I=\mathrm{Hom}_{\text{field}}(F,\overline{\mathbb{Q}})$. Then the set of algebraic characters $X(T_G)=\mathrm{Hom}_{\text{alg gp}}(T_{G/\overline{\mathbb{Q}}},\mathbb{G}_{m/\overline{\mathbb{Q}}})$ can be identified with $\mathbb{Z}[I]^2$ so that $\kappa=(\kappa_1,\kappa_2)\in\mathbb{Z}[I]^2$ induces the following character on $T_G(\mathbb{Q})=F^\times\times F^\times$

$$T_G(\mathbb{Q})\ni(\xi_1,\xi_2)\mapsto\kappa(\xi_1,\xi_2)=\xi_1^{\kappa_1}\xi_2^{\kappa_2}\in\overline{\mathbb{Q}}^\times,$$

where $\xi^{\kappa_j}=\prod_{\sigma\in I}\sigma(\xi_j)^{\kappa_{j,\sigma}}\in\overline{\mathbb{Q}}^\times$. We consider the following set of continuous "Neben" characters

$$\varepsilon=(\varepsilon_1,\varepsilon_2:T(\widehat{\mathbb{Z}})\to\mathbb{C}^\times,\varepsilon_+:Z(\mathbb{A})/Z(\mathbb{Q})\to\mathbb{C}^\times).$$

If a character $\psi:T(\widehat{\mathbb{Z}})\to\mathbb{C}^\times$ is continuous, it is of finite order, and we have an ideal $\mathfrak{c}(\psi)$ maximal among integral ideals $\mathfrak{c}$ satisfying $\psi(x)=1$ for all $x\in T(\widehat{\mathbb{Z}})=\widehat{O}^\times$ with $x-1\in\mathfrak{c}\widehat{O}$. We call $\mathfrak{c}(\psi)$ the conductor of ψ.

Exercise 2.6. *Prove that a continuous character* $\psi:T(\widehat{\mathbb{Z}})\to\mathbb{C}^\times$ *is of finite order* (see [MFG] Proposition 2.2).

The character $\varepsilon_+ : Z(\mathbb{A})/Z(\mathbb{Q}) \to \mathbb{C}^\times$ is an arithmetic Hecke character such that $\varepsilon_+(z) = \varepsilon_1(z)\varepsilon_2(z)$ for $z \in Z(\widehat{\mathbb{Z}})$ and $\varepsilon_+(x_\infty) = x^{-(\kappa_1+\kappa_2)+I}$. We can define the conductor $\mathfrak{c}(\varepsilon_+)$ in the same manner as above taking the restriction of ε_+ to $Z(\widehat{\mathbb{Z}}) \cong T(\widehat{\mathbb{Z}})$. We define $\mathfrak{c}(\varepsilon) = \mathfrak{c}(\varepsilon_1)\mathfrak{c}(\varepsilon_2) \subset \mathfrak{c}(\varepsilon_+)$.

The standard level group of $\Gamma_0(\mathfrak{N})$-type for an integral ideal $\mathfrak{N}$ is given by

$$\widehat{\Gamma}_0(\mathfrak{N}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in G(\widehat{\mathbb{Z}}) \middle| c \in \mathfrak{N}\widehat{O} \right\}. \tag{2.2}$$

We also define the *principal congruence subgroup*

$$\widehat{\Gamma}(\mathfrak{N}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in G(\widehat{\mathbb{Z}}) \middle| a-1, b, c, d-1 \in \mathfrak{N}\widehat{O} \right\}.$$

Exercise 2.7. *Show that for any integral ideal $0 \neq \mathfrak{N}$ and any $a \in G(\mathbb{A}^{(\infty)})$, there exists an integral ideal $0 \neq \mathfrak{N}' \subset \mathfrak{N}$ such that $\widehat{\Gamma}(\mathfrak{N}') \subset a\widehat{\Gamma}(\mathfrak{N})a^{-1}$.*

Let $\mathfrak{d}$ be the absolute different of F, and choosing an idele $\delta \in \widehat{O}$ with $\mathfrak{d}\widehat{O} = \delta\widehat{O}$, we then define a variant of $\widehat{\Gamma}_0(\mathfrak{N})$:

$$S_0(\mathfrak{N}) = \left(\begin{smallmatrix} \delta & 0 \\ 0 & 1 \end{smallmatrix} \right)^{-1} \widehat{\Gamma}_0(\mathfrak{N}) \left(\begin{smallmatrix} \delta & 0 \\ 0 & 1 \end{smallmatrix} \right). \tag{2.3}$$

This type of level groups has been often used by Shimura. Hilbert modular forms on the level group $S_0(\mathfrak{N})$ have a simpler form of Fourier expansion than those of level $\widehat{\Gamma}_0(\mathfrak{N})$.

If $\mathfrak{c}(\varepsilon^-) \supset \mathfrak{N}$ for $\varepsilon^- = \varepsilon_2^{-1}\varepsilon_1$, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \mapsto \varepsilon\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) = \varepsilon_2(ad-bc)\varepsilon^-(a_{\mathfrak{N}})$ is a continuous character of the compact group $S_0(\mathfrak{N})$ (this type of "Neben" character was first considered in [H89]). Here $a_{\mathfrak{N}}$ is the projection of a to the product $F_{\mathfrak{N}}$ of $F_{\mathfrak{l}}$ over all prime factors $\mathfrak{l}$ of $\mathfrak{N}$.

Exercise 2.8. *Show that $\varepsilon(uu') = \varepsilon(u)\varepsilon(u')$ for $u, u' \in S_0(\mathfrak{N})$ if $\mathfrak{c}(\varepsilon^-) \supset \mathfrak{N}$.*

We define the automorphy factor $J_\kappa(g, z)$ of weight κ for $z \in \mathfrak{Z} = \mathfrak{H}^I$ by

$$J_\kappa(g, z) = \det(g)^{\kappa_1 - I} j(g, z)^{\kappa_2 - \kappa_1 + I} \quad \text{for } g \in G(\mathbb{R}) \text{ and } z \in \mathfrak{Z}. \tag{2.4}$$

Here $j(g, z) = (c_\sigma z_\sigma + d_\sigma)_{\sigma \in I} \in \mathbb{C}^I = F \otimes_{\mathbb{R}} \mathbb{C}$, writing $g = (g_\sigma) \in GL_2(\mathbb{R})^I = GL_2(F_\infty)$ and $z = (z_\sigma) \in \mathfrak{Z}$. The power $j(g, z)^{\kappa_2 - \kappa_1 + I}$ is an abbreviation of $\prod_\sigma (c_\sigma z_\sigma + d_\sigma)^{\kappa_{2,\sigma} - \kappa_{1,\sigma} + 1}$, and similarly $\det(g)^{\kappa_1 - I} = \prod_\sigma \det(g_\sigma)^{\kappa_{1,\sigma} - 1}$. Then we define $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ to be the space of functions $f : G(\mathbb{A}) \to \mathbb{C}$ satisfying the following three conditions. A function in $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ is called a Hilbert *cusp form* of level $\mathfrak{N}$ and with character ε.

(SA1) We have the following automorphy

$$f(\alpha x u z) = \varepsilon_+(z)\varepsilon(u) f(x) J_\kappa(u_\infty, \mathbf{i})^{-1}$$

for all $\alpha \in G(\mathbb{Q})$, $z \in Z(\mathbb{A})$, and $u \in S_0(\mathfrak{N})C_{\mathbf{i}}$ for the stabilizer $C_{\mathbf{i}}$ in $G(\mathbb{R})^+$ of $\mathbf{i} = (\sqrt{-1}, \ldots, \sqrt{-1}) \in \mathfrak{Z} = \mathfrak{H}^I$, where $G(\mathbb{R})^+$ is the identity-connected component of $G(\mathbb{R})$;

(SA2) Choosing $u \in G(\mathbb{R})$ with $u(\mathbf{i}) = z$ for each $z \in \mathfrak{H}^I$, define a function $f_g : \mathfrak{Z} \to \mathbb{C}$ by $f_g(z) = f(g u_\infty) J_\kappa(u_\infty, \mathbf{i})$ for each $g \in G(\mathbb{A}^{(\infty)})$. Then f_g is a holomorphic function on $\mathfrak{Z}$ for all g;

(SA3) $\int_{F_{\mathbb{A}}/F} f\left(\left(\begin{smallmatrix}1 & u\\ 0 & 1\end{smallmatrix}\right)x\right)du = 0$ for all $x \in GL_2(F_{\mathbb{A}}^{(\infty)})$. This is equivalent to the following statement: The function: $z \mapsto f_g(z)|\prod_{\sigma\in I}\mathrm{Im}(z_\sigma)^{k_\sigma/2}|$ is bounded independently of $z \in \mathfrak{Z}$, where $k_\sigma = \kappa_{2,\sigma} - \kappa_{1,\sigma} + 1$. Here "$du$" is an additive Haar measure on $F_{\mathbb{A}}/F$.

Replacing the word "bounded" in (SA3) by "slowly increasing" with polynomial growth as $\mathrm{Im}(z) \to \infty$, we define a larger space of *modular forms* $G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$.

The function f_g in (SA2) satisfies the classical automorphy condition

$$f_g(\gamma(z)) = \varepsilon^{-1}(g^{-1}\gamma g)f_g(z)J_\kappa(\gamma, z) \quad \text{for all } \gamma \in \Gamma_{0,g}(\mathfrak{N}), \tag{2.5}$$

where $\Gamma_{0,g}(\mathfrak{N}) = g \cdot S_0(\mathfrak{N})g^{-1}G(\mathbb{R})^+ \cap G(\mathbb{Q})$, and $G(\mathbb{R})^+$ is the subgroup of $G(\mathbb{R})$ made up of matrices with totally positive determinant. Indeed, we have under the notation in (SA2),

$$\begin{aligned} f_g(\gamma(z)) &= f(g(\gamma u)_\infty)J_\kappa((\gamma u)_\infty, \mathbf{i}) = f(\gamma\gamma^{(\infty)^{-1}}g \cdot u_\infty)J_\kappa((\gamma u)_\infty, \mathbf{i}) \\ &= \varepsilon^{-1}(g^{-1}\gamma g)f(gu_\infty)J_\kappa(\gamma, u_\infty(\mathbf{i}))J_\kappa(u_\infty, \mathbf{i}) = \varepsilon^{-1}(g^{-1}\gamma g)f_g(z)J_\kappa(\gamma, z). \end{aligned}$$

The same computation applied to $\alpha \in G(\mathbb{Q})^+$ yields

$$f_g(\alpha(z))J_\kappa(\alpha, z)^{-1} = f_{\alpha^{(\infty)^{-1}}g}(z). \tag{2.6}$$

By (SA3) combined with (2.6), we conclude f_g is decreasing rapidly towards all cusps of $\Gamma_{0,g}(\mathfrak{N})$. Since we have $g\widehat{\Gamma}_0(\mathfrak{N})g^{-1} \supset \widehat{\Gamma}(\mathfrak{N}')$ for a suitable integral ideal $\mathfrak{N}'$, the discrete congruence subgroup $\Gamma_{0,g}(\mathfrak{N})$ contains

$$\Gamma_\infty(\mathfrak{a}) = \left\{\left(\begin{smallmatrix}1 & a\\ 0 & 1\end{smallmatrix}\right) \middle| a \in \mathfrak{a}\right\}$$

for a suitable ideal $\mathfrak{a} \supset \mathfrak{N}'$. The action of $\left(\begin{smallmatrix}1 & a\\ 0 & 1\end{smallmatrix}\right)$ on $\mathfrak{Z}$ is given by $z \mapsto z + a$; so, f_g in (2.5) satisfies $f_g(z + a) = f_g(z)$ for all $a \in \mathfrak{a}$. Since $F_\infty/\mathfrak{a}$ is a compact abelian group isomorphic to $(\mathbb{R}/\mathbb{Z})^d$, we can apply the standard Fourier analysis for the group $F_\infty/\mathfrak{a}$ (e.g., [LFE] Section 8.4), and we get the following Fourier expansion of f_g locally uniformly (and absolutely) convergent over $\mathfrak{Z}$:

$$f_g(z) = \sum_{a\in\mathfrak{a}^*} a(\xi, f_g)q^\xi,$$

where q^ξ is an abbreviation of $\exp(2\pi i\sum_\sigma \xi^\sigma z_\sigma)$ and

$$\mathfrak{a}^* = \mathfrak{a}^{-1}\mathfrak{d}^{-1} = \{\xi \in F | \mathrm{Tr}_{F/\mathbb{Q}}(\xi O) \subset \mathbb{Z}\}.$$

Since f_g decreases as $\mathrm{Im}(z_\sigma) \to \infty$ uniformly for $\sigma \in I$, we have $a(\xi, f) = 0$ if $\sigma(\xi) \le 0$ for one embedding $\sigma : F \hookrightarrow \mathbb{R}$. This shows that f_g decreases exponentially as $\mathrm{Im}(z_\sigma) \to \infty$ uniformly for $\sigma \in I$. Writing $\mathfrak{a}^*_+$ for the subset of $\mathfrak{a}^*$ made up of totally positive elements (that is, elements ξ with $\sigma(\xi) > 0$ for all $\sigma \in I$), we thus have

$$f_g(z) = \sum_{a\in\mathfrak{a}^*_+} a(\xi, f_g)q^\xi \tag{2.7}$$

for $f \in S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$. If $f \in G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$, the expansion can have a nontrivial constant term $a(0, f_g)$. Again by $g\widehat{\Gamma}_0(\mathfrak{N})g^{-1} \supset \widehat{\Gamma}(\mathfrak{N}')$, for a subgroup E of finite index in $O^\times$ (on which ε is trivial), $\Gamma_{0,g}(\mathfrak{N})$ contains diagonal matrices $\left(\begin{smallmatrix}\epsilon & 0\\ 0 & \epsilon^{-1}\end{smallmatrix}\right)$ for all $\epsilon \in E$. The effect of these matrices on f_g is given by $f_g(\epsilon^2 z) = \epsilon^{\kappa_1-\kappa_2-I} f_g(z)$ by (2.5). Thus we have

$$a(\epsilon^{-2}\xi, f_g) = \epsilon^{\kappa_1-\kappa_2-I} a(\xi, f_g) \tag{2.8}$$

for a sufficiently small subgroup $E \subset O^\times$ of finite index, and we conclude

$$G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}) = S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C}) \ \text{ unless } \kappa_1 - \kappa_2 \in \mathbb{Z} \cdot I. \tag{2.9}$$

Indeed, E contains a basis over $\mathbb{R}$ of the kernel $\mathrm{Ker}(N : F^\times_{\infty+} \to \mathbb{R}^\times)$ of the norm map N by Dirichlet's theorem. Here $F_\infty = F \otimes_\mathbb{Q} \mathbb{R}$, and $F^\times_{\infty+}$ is the identity connected component of the multiplicative group $F^\times_\infty$.

Similarly we have $G_\kappa = 0$ unless $\kappa_1 + \kappa_2 = [\kappa_1 + \kappa_2]I$ for $[\kappa_1 + \kappa_2] \in \mathbb{Z}$. To see this, we note by (SA1), for scalar matrices $\left(\begin{smallmatrix}\alpha & 0\\ 0 & \alpha\end{smallmatrix}\right)$ with $\alpha \in F^\times$, $f \in G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ satisfies $f(x) = f(\alpha x) = \varepsilon_+(\alpha^{(\infty)})\alpha^{\kappa_1+\kappa_2-I} f(x)$ by (2.5). Thus $\varepsilon_+(\alpha^{(\infty)})\alpha^{\kappa_1+\kappa_2-I}$ has to be equal to 1 for all $\alpha \in F^\times$, which implies the infinity type of the Hecke character ε_+ is $\kappa_1+\kappa_2-I$. On $\widehat{O}^\times$, ε_+ is of finite order m, and hence for $\epsilon \in O^\times$, we have $\varepsilon_+(\epsilon^{(\infty)})^m = 1$ and hence $\epsilon^{m(\kappa_1+\kappa_2-I)} = 1$ for all $\epsilon \in O^\times$. Again by Dirichlet's theorem, we get

$$G_\kappa \neq 0 \Rightarrow \kappa_1 + \kappa_2 \in \mathbb{Z} \cdot I \ \text{ and } \ \varepsilon_+(\alpha^{(\infty)})\alpha^{\kappa_1+\kappa_2-I} = 1 \ \text{ for all } \alpha \in F^\times. \tag{2.10}$$

We hereafter simply write $[\kappa]$ for $[\kappa_1 + \kappa_2] \in \mathbb{Z}$ if $\kappa_1 + \kappa_2 \in \mathbb{Z} \cdot I$. Also, by (2.9) and (2.10) combined, $G_\kappa \neq S_\kappa$ implies that $\kappa_1 \in \mathbb{Z} \cdot I$; so, in this case, we write $\kappa_1 = [\kappa_1]I$ for $[\kappa_1] \in \mathbb{Z}$.

We define the level $\mathfrak{N}$ semi-group $\Delta_0(\mathfrak{N}) \subset M_2(\widehat{O}) \cap G(\mathbb{A}^{(\infty)})$ by

$$\Delta_0(\mathfrak{N}) = \left\{ \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in M_2(\widehat{O}) \cap G(\mathbb{A}^{(\infty)}) \middle| a_\mathfrak{N} \in O^\times_\mathfrak{N},\ c \in \mathfrak{N}\widehat{O} \right\}. \tag{2.11}$$

Here $O_\mathfrak{N}$ is the product of $O_\mathfrak{l}$ over all prime factors $\mathfrak{l}$ of $\mathfrak{N}$. The opposite semi-group $\Delta^*_0(\mathfrak{N})$ is defined to be the image of $\Delta_0(\mathfrak{N})$ for the involution ι of $M_2(F)$ with $x + x^\iota = \mathrm{Tr}(x)$. Thus

$$\Delta^*_0(\mathfrak{N}) = \left\{ \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in M_2(\widehat{O}) \cap G(\mathbb{A}^{(\infty)}) \middle| d_\mathfrak{N} \in O^\times_\mathfrak{N},\ c \in \mathfrak{N}\widehat{O} \right\}. \tag{2.12}$$

We fix a prime p, and write v for one of the rational places p or ∞. Let $\mathcal{W}$ be a p-adic valuation ring of a number field inside $\mathbb{C}$ containing the values of ε on $T(\widehat{\mathbb{Z}})$ and $Z(\widehat{\mathbb{Z}})$ and all the conjugates of O in $\overline{\mathbb{Q}}$. We fix once and for all a prime element $\varpi_\mathfrak{q}$ for each prime ideal $\mathfrak{q}$. Here we choose $\varpi_\mathfrak{q}$ inside $\mathcal{W} \cap F$. We extend our Neben character $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_+)$ to $T(\mathbb{A}^{(\infty)})$ and $\Delta_0(\mathfrak{N})$ as follows.

(ex0) *For each $w \in \mathbb{Z}[I]$, we extend the character $x_p \mapsto x_p^w$ of $T(\mathbb{Z}_p)$ to $T(\mathbb{Q}_p)$ trivially on $\bigoplus_{\mathfrak{p}|p} \varpi_\mathfrak{p}^\mathbb{Z}$ and write this extension as $x_p \mapsto x_p^{w_p}$, where $x_p^w = \prod_{\sigma \in I} \sigma(x_p)^{w_\sigma}$ for $x \in O_p^\times = T(\mathbb{Z}_p)$. Thus p^{w_p} is a p-adic unit. This only applies to the place $v = p$, because $x \in T(\widehat{\mathbb{Z}})$ has trivial component $x_\infty = 1$ at the infinity. In particular, $x_\infty^{w_\infty} := x_\infty^w = 1$ for any $x \in \Delta_0(\mathfrak{N})$ and any $w \in \mathbb{Z}[I]$ in the following conditions.*

(ex1) *We extend ε_2 to the idele group $T(\mathbb{A}^{(\infty)})$ trivially on $\bigoplus_\mathfrak{q} \varpi_\mathfrak{q}^\mathbb{Z}$ and then extend ε_1 to $T(\mathbb{A}^{(\infty)})$ by $\varepsilon_1\varepsilon_2(x) = \varepsilon_+(x^{(\infty)})$. We put $\varepsilon^-(a) = \varepsilon_2^{-1}(a)\varepsilon_1(a)$ for $a \in T(\mathbb{A}^{(\infty)})$. Thus identifying $T^2 \cong T_G$ by $(a, d) \mapsto \left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)$, we get an extension of ε to $T_G(\mathbb{A})$.*

(ex2) *We extend the character* ε *of* $T_G(\widehat{\mathbb{Z}})\cap S_0((v)\cap\mathfrak{N})$ *to a character of the semi-group* $\Delta_0((v)\cap\mathfrak{N})$ *by*

$$\varepsilon^v_\Delta(\delta)=\det(\delta)_v^{-I_v}\varepsilon_2(\det(\delta))\varepsilon_v^-(a_{(v)\cap\mathfrak{N}})$$

for $\delta=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\Delta_0((v)\cap\mathfrak{N})\cup\Delta_0((v)\cap\mathfrak{N})^\iota$. *Here* $(v)\cap\mathfrak{N}=(p)\cap\mathfrak{N}$ *if* $v=p$ *and* $(v)\cap\mathfrak{N}=\mathfrak{N}$ *if* $v=\infty$.

(ex3) *Since the character* $Z(\mathbb{A}^{(\infty)})/Z(\mathbb{Q})\ni z\mapsto|z^{(\infty)}|_{\mathbb{A}}^{-2}z_v^{-2I_v}\varepsilon_+(z)$ *and* ε^v_Δ *on* $S_0((v)\cap\mathfrak{N})$ *coincide on* $S_0((v)\cap\mathfrak{N})\cap Z(\mathbb{A}^{(\infty)})$, *we may extend* $\varepsilon^v_\Delta|_{S_0((v)\cap\mathfrak{N})}$ *to* $\varepsilon_S=\varepsilon^v_S$: $Z(\mathbb{A})S_0((v)\cap\mathfrak{N})=Z(\mathbb{A})S_1((v)\cap\mathfrak{N})\to\mathcal{B}^\times$ *by* $\varepsilon^v_S(zs)=|z^{(\infty)}|_{\mathbb{A}}^{-2}z_v^{-2I_v}\varepsilon_+(z)\varepsilon^v_\Delta(s)$ *for* $z\in Z(\mathbb{A})$ *and* $s\in S_1((v)\cap\mathfrak{N})$.

Since $\varepsilon^\infty_\Delta$ is defined over $\Delta_0(\mathfrak{N})$ and coincides with $z\mapsto\varepsilon_+(z_\mathfrak{N})\varepsilon_2(\det(z^{(\mathfrak{N})}))$ on $Z(\mathbb{A}^{(\infty)})\cap\Delta_0(\mathfrak{N})$, we may extend it to the subgroup generated by $Z(\mathbb{A}^{(\infty)})\Delta_0(\mathfrak{N})$ (which contains $\Delta_0^*(\mathfrak{N})$) so that $\underline{\varepsilon}(z\delta)=\varepsilon_+(z_\mathfrak{N})\varepsilon_2(\det(z^{(\mathfrak{N})}))\varepsilon^\infty_\Delta(\delta)$ for $\delta\in\Delta_0(\mathfrak{N})$ and $z\in Z(\mathbb{A}^{(\infty)})$.

For each $y\in\Delta_0(\mathfrak{N})$, we can decompose

$$S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})=\bigsqcup_{u,t}utS_0(\mathfrak{N}) \tag{2.13}$$

for finitely many $u\in U(\widehat{\mathbb{Z}})$ and $t\in T_G(\mathbb{A}^{(\infty)})$ with $\det(t)=\det(y)$ (see (2.33)).

$$f|[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})](g)=\sum_{u,t}\varepsilon^\infty_\Delta((ut)^\iota)f(gut)\overset{(*)}{=}\underline{\varepsilon}(\det(y))\sum_h\underline{\varepsilon}(ut)^{-1}f(gut). \tag{2.14}$$

The second identity $(*)$ follows from $\det(t)=\det(y)$, $(ut)^\iota=\det(t)(ut)^{-1}$ and multiplicativity of $\underline{\varepsilon}$, and by this, the sum is independent of the choice of u and t as long as $\det(t)=\det(y)$. It is easy to verify that the *Hecke operator* defined by (2.14) preserves the space $G_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ and $S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ by confirming (SA1–3) for $f|[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]$. Fix an isomorphism $\overline{\mathbb{Q}}_p\cong\mathbb{C}$ (compatible with the two embeddings $\overline{\mathbb{Q}}_p\overset{i_p}{\hookleftarrow}\overline{\mathbb{Q}}\overset{i_\infty}{\hookrightarrow}\mathbb{C}$). Then assuming $\mathfrak{N}\subset(p)$, we can also define the action of $[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]_p$ taking $v=p$ in place of ∞ and replacing $\varepsilon^\infty_\Delta(y)$ by $\det(y_p)^{-\kappa_1}\varepsilon^p_\Delta(y)$. Since $\det(y_p)^{-\kappa_1}\varepsilon^p_\Delta(y)=\det(y_p)^{-I_p-\kappa_1}\varepsilon^\infty_\Delta(y)$ (under the convention of (ex0)), the only difference of the two normalization is $\det(y_p)^{-I_p-\kappa_1}$, and in particular, we have

$$\det(y_p)^{-\kappa_1}[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]=[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]_p$$

for $y=\left(\begin{smallmatrix}1&0\\0&\varpi_\mathfrak{q}\end{smallmatrix}\right)$ for our chosen uniformizer $\varpi_\mathfrak{q}$, because $\varpi_\mathfrak{q}^{I_p}=1$ (but $\varpi_\mathfrak{p}^{-\kappa_1}\neq1$ if $\kappa_1\neq0$). We simply write therefore $T(\varpi_\mathfrak{q})$ (resp. $T_p(\varpi_\mathfrak{q})$) for the operator $[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]$ (resp. $[S_0(\mathfrak{N})y^\iota S_0(\mathfrak{N})]_p$) with $y=\left(\begin{smallmatrix}1&0\\0&\varpi_\mathfrak{q}\end{smallmatrix}\right)$. When $\mathfrak{q}$ is a factor of $\mathfrak{N}$, writing $\mathfrak{N}^{(\mathfrak{q})}$ for the prime-to–$\mathfrak{q}$ part of $\mathfrak{N}$ and assuming that the conductor of ε^- is a factor of $\mathfrak{N}^{(\mathfrak{q})}$, we find that the operator $T(\varpi_\mathfrak{q})$ on $S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ does not preserve $S_\kappa(\mathfrak{N}^{(\mathfrak{q})},\varepsilon;\mathbb{C})\subset S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ and does not induce the operator $T(\varpi_\mathfrak{q})$ on $S_\kappa(\mathfrak{N}^{(\mathfrak{q})},\varepsilon;\mathbb{C})$. Thus if it is necessary to distinguish two operators $T(\varpi_\mathfrak{q})$ of level $\mathfrak{N}$ and of level $\mathfrak{N}^{(\mathfrak{q})}$, we write $U(\varpi_\mathfrak{q})$ (resp. $U_p(\varpi_\mathfrak{q})$) for $T(\varpi_\mathfrak{q})$ (resp. $T_p(\varpi_\mathfrak{q})$) of level $\mathfrak{N}$ if $\mathfrak{q}\supset\mathfrak{N}$. Note here $T_p(\varpi_\mathfrak{q})=T(\varpi_\mathfrak{q})$ and $U_p(\varpi_\mathfrak{q})=U(\varpi_\mathfrak{q})$ if $\kappa_1=0$ or $\mathfrak{q}\nmid p$.

2.3. **Hilbert modular forms with integral coefficients.** Let $\mathbf{e}_F : F_{\mathbb{A}}/F \to \mathbb{C}^\times$ be the standard additive character determined by the condition: $\mathbf{e}_F(x_\infty) = \exp(2\pi i \sum_\sigma x_\sigma)$ for $x_\infty \in F_\infty$ (e.g., [LFE] Theorem 8.3.1). We start with

Proposition 2.9. *Each member f of $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ has a Fourier expansion of the following form,*

$$f\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right) = |y|_{\mathbb{A}} \sum_{0 \ll \xi \in F} \mathbf{a}_\infty(\xi y, f)(\xi y_\infty)^{-\kappa_1} \mathbf{e}_F(i\xi y_\infty)\mathbf{e}_F(\xi x). \tag{2.15}$$

More generally, each modular form f of $G_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ with $\kappa_1 = [\kappa_1]I$ for $[\kappa_1] \in \mathbb{Z}$ can be expanded into

$$f\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right) = |y|_{\mathbb{A}} \left\{ \mathbf{a}_0(y, f)|y|_{\mathbb{A}}^{-[\kappa_1]} + \sum_{0 \ll \xi \in F} \mathbf{a}_\infty(\xi y, f)(\xi y_\infty)^{-\kappa_1} \mathbf{e}_F(i\xi y_\infty)\mathbf{e}_F(\xi x) \right\}.$$

Here $y \mapsto \mathbf{a}_\infty(y, f)$ is a function defined on $y \in F_{\mathbb{A}}^\times$ only depending on its finite part $y^{(\infty)}$ and satisfies $\mathbf{a}_\infty(uy, f) = \varepsilon_1(u)\mathbf{a}_\infty(y, f)$ for $u \in T(\widehat{\mathbb{Z}})$. The function $\mathbf{a}_\infty(y, f)$ is supported by the set $(\widehat{O} \times F_\infty) \cap F_{\mathbb{A}}^\times$ of integral ideles. The function $y \mapsto \mathbf{a}_0(y, f)$ factors through the class group $Cl_F = F_{\mathbb{A}^{(\infty)}}^\times / \widehat{O}^\times F^\times$.

We note that the function $\left(\begin{smallmatrix} y_\infty & 0 \\ 0 & 1 \end{smallmatrix}\right) \mapsto y_\infty^{-\kappa_1}\mathbf{e}_F(iy_\infty)$ is the restriction of the canonical Whittaker function of $G(\mathbb{R})^+$ to matrices of the form $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (whose Mellin transform gives the optimal Γ-factor of the standard L-function of f). Here is a sketch of a proof.

Proof. Since the proof is basically the same for cusp forms and modular forms, we give an argument for cusp forms (see [H88] Section 4 for modular forms). We consider the unipotent subgroup

$$U(R) = \left\{ \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right) \middle| u \in (R \otimes_{\mathbb{Q}} F) \right\}$$

of G. Then for $\left(\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right) \in U(\mathbb{Q})$,

$$f\left(\left(\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right) = f\left(\left(\begin{smallmatrix} y & x+\alpha \\ 0 & 1 \end{smallmatrix}\right)\right) = f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right).$$

Thus the function $x \mapsto f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right)$ for a fixed y is a function on $F_{\mathbb{A}}/F$, which is a compact abelian group. Applying the standard Fourier analysis to this group $F_{\mathbb{A}}/F$ (e.g., [LFE] 8.3–4), we can expand

$$f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right) = \sum_{\xi \in F} c(\xi, y, f)\mathbf{e}_F(\xi x).$$

Taking $\alpha \in F^\times$, by (SA1), we have

$$\sum_{\xi \in F} c(\xi, \alpha y, f)\mathbf{e}_F(\xi \alpha x) = f\left(\left(\begin{smallmatrix} \alpha y & \alpha x \\ 0 & 1 \end{smallmatrix}\right)\right)$$
$$= f\left(\left(\begin{smallmatrix} \alpha & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right) = f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right)\right) = \sum_{\xi \in F} c(\xi, y, f)\mathbf{e}_F(\xi x).$$

Thus $c(\xi,y,f)=c(\xi\alpha^{-1},\alpha y,f)$ for all $\alpha\in F^\times$. In other words, $c(\xi,y,f)$ only depends on ξy; so, writing $c(y,f)=c(1,y,f)$, we have $c(\xi,y,f)=c(\xi y,f)$. Taking $g=\left(\begin{smallmatrix} y^{(\infty)} & x^{(\infty)} \\ 0 & 1\end{smallmatrix}\right)$ and $u_\infty=\left(\begin{smallmatrix} y_\infty & x_\infty \\ 0 & 1\end{smallmatrix}\right)$ in (SA2), as in (2.7) we have

$$f_g(x_\infty+iy_\infty)=y_\infty^{\kappa_1-I}f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1\end{smallmatrix}\right)\right)=\sum_{\xi\in\mathfrak{a}^*}a(\xi,f_g)\exp(2\pi i\sum_\sigma \xi^\sigma(x_\sigma+iy_\sigma)).$$

This shows

$$c(\xi y,f)=y_\infty^{I-\kappa_1}a(\xi,f_g)\exp(-2\pi\sum_\sigma\xi^\sigma y_\sigma).$$

Since $\Gamma_{0,g}(\mathfrak{N})=g\cdot S_0(\mathfrak{N})g^{-1}G(\mathbb{R})^+\cap G(\mathbb{Q})$, $\mathfrak{a}$ is given by $y^{(\infty)}\delta^{-1}\widehat{O}\cap F$; so, $\mathfrak{a}^*=(y^{(\infty)})^{-1}\widehat{O}\cap F$. This shows that $c(\xi y,f)=0$ unless $(\xi y)^{(\infty)}\in\widehat{O}$, and hence $c(y,f)=0$ unless $y^{(\infty)}\in\widehat{O}$. Thus we may define

$$\mathbf{a}_\infty(y,f)=c(y,f)\exp(2\pi\sum_\sigma y_\sigma)|y|_{\mathbb{A}}^{-1}y_\infty^{\kappa_1}=|y^{(\infty)}|_{\mathbb{A}}^{-1}a(1,f_g)$$

for $g=\left(\begin{smallmatrix} y^{(\infty)} & x^{(\infty)} \\ 0 & 1\end{smallmatrix}\right)$, which satisfies the desired properties. In particular, from $f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix} u & 0 \\ 0 & 1\end{smallmatrix}\right)\right)=\varepsilon_1(u)f\left(\left(\begin{smallmatrix} y & x \\ 0 & 1\end{smallmatrix}\right)\right)$, we have $\mathbf{a}_\infty(uy,f)=\varepsilon_1(u)\mathbf{a}_\infty(y,f)$ for $u\in T(\widehat{\mathbb{Z}})$. $\square$

In view of the well known decomposition (e.g., [HMI, Lemma 2.46])

$$S_0(\mathfrak{N})\left(\begin{smallmatrix} \varpi_\mathfrak{q} & 0 \\ 0 & 1\end{smallmatrix}\right)S_0(\mathfrak{N})=\begin{cases}\left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right)S_0(\mathfrak{N})\sqcup\bigsqcup_{u\in\delta_\mathfrak{q}^{-1}O_\mathfrak{q}/\delta_\mathfrak{q}^{-1}\mathfrak{q}}\left(\begin{smallmatrix} \varpi_\mathfrak{q} & u \\ 0 & 1\end{smallmatrix}\right)S_0(\mathfrak{N}) & \text{if } \mathfrak{q}\nmid\mathfrak{N},\\ \bigsqcup_{u\in\delta_\mathfrak{q}^{-1}O_\mathfrak{q}/\delta_\mathfrak{q}^{-1}\mathfrak{q}}\left(\begin{smallmatrix} \varpi_\mathfrak{q} & u \\ 0 & 1\end{smallmatrix}\right)S_0(\mathfrak{N}) & \text{if } \mathfrak{q}|\mathfrak{N},\end{cases}$$

we can directly verify by computation the following fact

$$\begin{aligned}&\mathbf{a}_\infty(y,f|T(\varpi_\mathfrak{q}))=\mathbf{a}_\infty(y\varpi_\mathfrak{q},f)+N(\mathfrak{q})\varepsilon_+(\varpi_\mathfrak{q})\mathbf{a}_\infty(\frac{y}{\varpi_\mathfrak{q}},f)\ \text{ if } \mathfrak{q}\nmid\mathfrak{N},\\ &\mathbf{a}_\infty(y,f|U(\varpi_\mathfrak{q}))=\mathbf{a}_\infty(y\varpi_\mathfrak{q},f)\ \text{ if } \mathfrak{q}|\mathfrak{N}.\end{aligned}\tag{2.16}$$

Remark 2.1. The Hecke operator $T(\varpi_\mathfrak{q})$ for $\mathfrak{q}|\mathfrak{N}$ acts slightly differently from other $T(\varpi_\mathfrak{l})$ for $\mathfrak{l}$ prime to $\mathfrak{N}$. If $\mathfrak{N}^{(\mathfrak{q})}$ is the prime-to-$\mathfrak{q}$ part of $\mathfrak{N}$ and if the conductor of ε is prime to $\mathfrak{q}$, the action of $T(\varpi_\mathfrak{q})$ on $S_\kappa(\mathfrak{N}^{(\mathfrak{q})},\varepsilon;\mathbb{C})$ is not the restriction of the action of $U(\varpi_\mathfrak{q})$ on $S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ to $S_\kappa(\mathfrak{N}^{(\mathfrak{q})},\varepsilon;\mathbb{C})$ though $S_\kappa(\mathfrak{N}^{(\mathfrak{q})},\varepsilon;\mathbb{C})$ is canonically a subspace of $S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$.

Let us prove (2.16). We call an idele $y\in F_\mathbb{A}^\times$ *integral* if $y^{(\infty)}\in\widehat{O}$. Since the computation in the cases where $\mathfrak{q}|\mathfrak{N}$ and $\mathfrak{q}\nmid\mathfrak{N}$ is the same, we treat the case of $\mathfrak{q}\nmid\mathfrak{N}$. By the above decomposition of $S_0(\mathfrak{N})\left(\begin{smallmatrix} \varpi_\mathfrak{q} & 0 \\ 0 & 1\end{smallmatrix}\right)S_0(\mathfrak{N})$ (into a disjoint union of right cosets) combined with $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi_\mathfrak{q}\end{smallmatrix}\right)=\varpi_\mathfrak{q}\begin{pmatrix} \varpi_\mathfrak{q}^{-1} & 0 \\ 0 & 1\end{pmatrix}$, we have for integral y

$$\begin{aligned}|y|_\mathbb{A}\mathbf{a}_\infty(y,f|T(\varpi_\mathfrak{q}))\mathbf{e}_F(x)&=\varepsilon_+(\varpi_\mathfrak{q})|\varpi_\mathfrak{q}^{-1}y|_\mathbb{A}\mathbf{a}_\infty(\frac{y}{\varpi_\mathfrak{q}},f)\mathbf{e}_F(x)\\ &\quad+|\varpi_\mathfrak{q}y|_\mathbb{A}\mathbf{a}_\infty(\varpi_\mathfrak{q}y,f)\mathbf{e}_F(x)\left(\sum_{u\in\delta_\mathfrak{q}^{-1}O_\mathfrak{q}/\delta_\mathfrak{q}^{-1}\mathfrak{q}O_\mathfrak{q}}\mathbf{e}_F(yu)\right)\end{aligned}$$

Note that $\sum_{u\in\delta_{\mathfrak{q}}^{-1}O_{\mathfrak{q}}/\delta_{\mathfrak{q}}^{-1}\mathfrak{q}O_{\mathfrak{q}}}\mathbf{e}_F(yu)=|\varpi_{\mathfrak{q}}|_{\mathbb{A}}^{-1}=N(\mathfrak{q})$, because $y\in\widehat{O}$. From this we get the desired formula.

For each $\mathbb{Q}$-algebra $R\subset\mathbb{C}$ containing the values of characters $\varepsilon=(\varepsilon_+,\varepsilon_1,\varepsilon_2)$ and κ on $T_G(\mathbb{Q})$, we define

$$\begin{aligned} S_\kappa(\mathfrak{N},\varepsilon;R) &= \left\{f\in S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})\middle|\mathbf{a}_\infty(y,f)\in R\ \text{ for all } y\right\}, \\ G_\kappa(\mathfrak{N},\varepsilon;R) &= \left\{f\in G_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})\middle|\mathbf{a}_0(y,f),\mathbf{a}_\infty(y,f)\in R\ \text{ for all } y\right\}. \end{aligned} \tag{2.17}$$

Recall the p-adic valuation ring $\mathcal{W}$ of a number field inside $\mathbb{C}$ containing the values of ε on $T(\widehat{\mathbb{Z}})$ and $Z(\widehat{\mathbb{Z}})$ and all the Galois conjugates of O in $\overline{\mathbb{Q}}$. As in [HMI, (4.3.7)], for a $\mathcal{W}$-algebra $R\subset\overline{\mathbb{Q}}$ with $i_\infty:\overline{\mathbb{Q}}\hookrightarrow\mathbb{C}$:

$$G_\kappa(\mathfrak{N},\varepsilon;R)\otimes_{R,i_\infty}\mathbb{C}=G_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})\ \text{ and }\ S_\kappa(\mathfrak{N},\varepsilon;R)\otimes_{R,i_\infty}\mathbb{C}=S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C}). \tag{2.18}$$

We recall the embedding $i_p:\overline{\mathbb{Q}}\hookrightarrow\overline{\mathbb{Q}}_p$. Then for any $\overline{\mathbb{Q}}_p$-algebra R, we define, consistently with (2.18),

$$G_\kappa(\mathfrak{N},\varepsilon;R)=G_\kappa(\mathfrak{N},\varepsilon;\overline{\mathbb{Q}})\otimes_{\overline{\mathbb{Q}},i_p}R \text{ and } S_\kappa(\mathfrak{N},\varepsilon;R)=S_\kappa(\mathfrak{N},\varepsilon;\overline{\mathbb{Q}})\otimes_{\overline{\mathbb{Q}},i_p}R.$$

There is a more intrinsic definition of these spaces as global sections of the automorphic line bundle $\underline{\omega}_{\kappa,\varepsilon}$ over Hilbert modular Shimura varieties (see [HMI, Chapter 4]). By linearity, $(y,f)\mapsto\mathbf{a}_\infty(y,f)$ extends to functions on $F_{\mathbb{A}}^\times\times G_\kappa(\mathfrak{N},\varepsilon;R)$ with values in R. Then we define the p-adic q-expansion coefficients $\mathbf{a}_p(y,f)$ (originally given in [PAF, (4.63)]) of $f\in G_\kappa(\mathfrak{N},\varepsilon;R)$ by

$$\mathbf{a}_p(y,f)=y_p^{-\kappa_1}\mathbf{a}_\infty(y,f). \tag{2.19}$$

Here $y_p^{-\kappa_1}=\prod_\sigma\sigma(y_p)^{-\kappa_{1,\sigma}}$ does not follow the convention in (ex0). Even if we have divided by possibly a nonunit $y_p^{\kappa_1}$, the coefficients $\mathbf{a}_p(y,f)$ reflects faithfully the p-integrality coming from the q-expansion. Indeed, writing $y=\xi c$ for $\xi\in F_+^\times$ and an idele c with $c_p=c_\infty=1$, we have $\mathbf{a}_p(y,f)=|c^{(\infty)}|_{\mathbb{A}}^{-1}a(\xi,f_g)$ for $g=\left(\begin{smallmatrix}c&0\\0&1\end{smallmatrix}\right)$, which is the q-expansion coefficients of the classical modular form f_g up to p-adic unit $|c^{(\infty)}|_{\mathbb{A}}^{-1}$. By Proposition 2.9, we have

$$\mathbf{a}_p(uy,f)=\varepsilon_1(u)u_p^{-\kappa_1}\mathbf{a}_p(y,f)\ \text{ for } u\in T(\widehat{\mathbb{Z}}). \tag{2.20}$$

The coefficient $\mathbf{a}_p$ behaves better then $\mathbf{a}_\infty$ under the action of Hecke operators, because we have the following p-integral formula without any conditions on κ which follows from (2.16) and (2.19) combined (see also Proposition **??** in the text and [PAF] (4.65–66) and (4.79)):

$$\begin{aligned} &\mathbf{a}_p(y,f|T_p(\varpi_{\mathfrak{q}}))=\mathbf{a}_p(y\varpi_{\mathfrak{q}},f)+N(\mathfrak{q})\varepsilon_+(\varpi_{\mathfrak{q}})\mathbf{a}_p(\frac{y}{\varpi_{\mathfrak{q}}},f)\ \text{ if } \mathfrak{q}\nmid p\mathfrak{N}, \\ &\mathbf{a}_p(y,f|U_p(\varpi_{\mathfrak{q}}))=\mathbf{a}_p(y\varpi_{\mathfrak{q}},f)\ \text{ if } \mathfrak{q}|p\mathfrak{N}. \end{aligned} \tag{2.21}$$

The formal q-expansion of f has values in the space of functions on $T(\mathbb{A}^{(\infty)})$ with values in the formal monoid algebra $R[[q^\xi]]_{\xi\in F_+^\times}$ of the multiplicative semi-group $F_+^\times$, which is given by

$$f(y) = \sum_{\xi \gg 0} \mathbf{a}_p(\xi y, f) q^{\xi}. \tag{2.22}$$

Here $\xi \gg 0$ implies $\sigma(\xi) > 0$ for all $\sigma \in I$. This is the p-adic analogue of the Archimedean Fourier expansion (2.15). In particular, if $\mathbf{a}_p(y, f) = 0$ for all integral ideles y, the modular form f vanishes. In other words, the q-expansion: $y \mapsto f(y)$ determines f uniquely (for any algebra R for which the space of R-integral modular forms is well-defined).

Let $W = \varprojlim_n \mathcal{W}/p^n\mathcal{W}$ be the p-adic completion of $\mathcal{W}$. From the q-expansion principle (see [HMI, Corollary 4.16]), for any p-adically complete W-algebra R in $\widehat{\mathbb{Q}}_p$ (the p-adic completion of $\overline{\mathbb{Q}}_p$), we conclude that the space

$$S_\kappa(\mathfrak{N}, \varepsilon; R) = \left\{ f \in S_\kappa(\mathfrak{N}, \varepsilon; \widehat{\mathbb{Q}}_p) \middle| \mathbf{a}_p(y, f) \in R \text{ for all integral } y \right\} \tag{2.23}$$

coincides with the space of R-integral cusp forms defined algebro-geometrically (see [HMI, §4.3.3]). For the moment, we take the above description (2.23) as the definition of the space $S_\kappa(\mathfrak{N}, \varepsilon; R)$ for W-algebras R.

As a direct consequence of (2.21), under the following condition

$$\text{either } p|\mathfrak{N} \text{ or } [\kappa] \geq 0, \tag{2.24}$$

the space of R-integral modular forms $S_\kappa(\mathfrak{N}, \varepsilon; R)$ is stable under Hecke operators. We then define the *Hecke algebra* $h_\kappa(\mathfrak{N}, \varepsilon; R)$ by the subalgebra of $\mathrm{End}_R(S_\kappa(\mathfrak{N}, \varepsilon; R))$ generated over R by $T_p(\varpi_\mathfrak{l})$ for all prime ideal $\mathfrak{l}$. Actually we have well-defined linear operators $T_p(y)$ for general integral ideles y as we will see in the following section, and $T_p(y)$ is an integral polynomial of $T_p(\varpi_\mathfrak{l})$ and the operators given by the action of $S_0(\mathfrak{N})$ (via ε) (see Lemma 2.21 in the text and [MFG] Lemma 3.9). Thus $h_\kappa(\mathfrak{N}, \varepsilon; R)$ is the R-subalgebra of $\mathrm{End}_R(S_\kappa(\mathfrak{N}, \varepsilon; R))$ generated by $T_p(y)$ for all integral ideles y.

Since we have chosen $\varpi_\mathfrak{q}$ inside $\mathcal{W}$, writing $\widehat{A} = A \otimes_\mathcal{W} W$ for any $\mathcal{W}$-algebra $A \subset \overline{\mathbb{Q}}_p$, we have a well defined A-integral subspace $S_\kappa(\mathfrak{N}, \varepsilon; A)$ of $S_\kappa(\mathfrak{N}, \varepsilon; \widehat{A})$ given by

$$\{f \in S_\kappa(\mathfrak{N}, \varepsilon; \widehat{A}) | \mathbf{a}_p(y, f) \in A \ \text{ if } y \text{ is a product of } \varpi_\mathfrak{q} \text{ for primes } \mathfrak{q}\}.$$

Then we have

$$S_\kappa(\mathfrak{N}, \varepsilon; \widehat{A}) = S_\kappa(\mathfrak{N}, \varepsilon; A) \otimes_A \widehat{A}.$$

2.4. **Duality and Hecke algebras.** The elementary duality theorem between the Hecke algebra and the space of modular forms we state now is quite useful in many different applications.

Let

$$T(y) = \left\{ g \in \left(\begin{smallmatrix} \delta & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} \Delta_0(\mathfrak{N}) \left(\begin{smallmatrix} \delta & 0 \\ 0 & 1 \end{smallmatrix}\right) \middle| \det g = y \right\}.$$

Decompose the double coset $S_0(\mathfrak{N})T(y)S_0(\mathfrak{N})$ for an integral finite idele y into a disjoint union of right cosets $\bigsqcup_{\alpha:\det(\alpha)=y} \alpha S_0(\mathfrak{N})$. Then we extend the definition of the *Hecke operator* $T(\varpi_\mathfrak{l})$ to general integral ideles y by $f|T_p(y)(g) = \sum_\alpha y_p^{-\kappa_1} \varepsilon_\Delta^p(\alpha^\iota) f(g\alpha)$. Here we have taken α so that $\det(\alpha) = y$ (which tells us that the operator depends on y). Then by (2.21), we find, for each integral finite idele y,

$$\mathbf{a}_p(1, f|T_p(y)) = \mathbf{a}_p(y, f). \tag{2.25}$$

Since we have chosen $\varpi_{\mathfrak{q}}$ inside $\mathcal{W}$, the operator $T_p(y)$ for $y = \prod_{\mathfrak{q}} \varpi_{\mathfrak{q}}^{e(\mathfrak{q})}$ with $e(\mathfrak{q}) \geq 0$ preserves $S_\kappa(\mathfrak{N}, \varepsilon; A)$ for any $\mathcal{W}$-algebra A. Thus we may extend the definition of the Hecke algebra $h_\kappa(\mathfrak{N}, \varepsilon; A)$ to $\mathcal{W}$-algebras A so that it is the A-subalgebra of $\mathrm{End}_A(S_\kappa(\mathfrak{N}, \varepsilon; A))$ generated by $T_p(y)$ for all y of the form $y = \prod_{\mathfrak{q}} \varpi_{\mathfrak{q}}^{e(\mathfrak{q})}$. Then we define an A–bilinear pairing

$$(\ ,\ ) : h_\kappa(\mathfrak{N}, \varepsilon; A) \times S_\kappa(\mathfrak{N}, \varepsilon; A) \to A \ \text{ by } \ (h, f) = \mathbf{a}_p(1, f|h).$$

Theorem 2.10 (Duality)**.** *Let A be a $\mathcal{W}$–algebra. Then $(\ ,\ )$ induces isomorphisms*

$$\mathrm{Hom}_A(S_\kappa(\mathfrak{N}, \varepsilon; A), A) \cong h_\kappa(\mathfrak{N}, \varepsilon; A), \quad \mathrm{Hom}_A(h_\kappa(\mathfrak{N}, \varepsilon; A), A) \cong S_\kappa(\mathfrak{N}, \varepsilon; A),$$

and the latter isomorphism is given by sending ϕ to f with $\mathbf{a}_p(y, f) = \phi(T_p(y))$.

Proof. Since $S_\kappa(\mathfrak{N}, \varepsilon; A) = S_\kappa(\mathfrak{N}, \varepsilon; \mathcal{W}) \otimes_{\mathcal{W}} A$, we may assume that $A = \mathcal{W}$. Since W is p-adic completion of $\mathcal{W}$, it is faithfully flat over $\mathcal{W}$; so, we may assume that $A = W$. Actually we prove the duality first for the quotient field K of W. The space $S_\kappa(\mathfrak{N}, \varepsilon; K)$ is finite dimensional over $\mathbb{Q}_p$; so, we need to prove nondegeneracy of the pairing. By (2.25), $\mathbf{a}_p(1, f|T_p(y)) = \mathbf{a}_p(y, f)$; so, if $(h, f) = 0$ for all h, $\mathbf{a}_p(y, f) = (T_p(y), f) = 0$ for all integral y, and hence $f = 0$. If $(h, f) = 0$ for all f, then $0 = (h, f|T_p(y)) = \mathbf{a}_p(1, f|T_p(y)h) = (T_p(y), f|h) = \mathbf{a}_p(y, f|h)$; so, $f|h = 0$ for all f, which implies $h = 0$. If $\phi \in \mathrm{Hom}_W(h_\kappa(\mathfrak{N}, \varepsilon; W), W)$, then we find $f \in S_\kappa(\mathfrak{N}, \varepsilon; K)$ with $(h, f) = \phi(h)$, and $\mathbf{a}_p(y, f) = (T_p(y), f) = \phi(T_p(y)) \in W$; so, $f \in S_\kappa(\mathfrak{N}, \varepsilon; W)$. This shows $S_\kappa(\mathfrak{N}, \varepsilon; W) = \mathrm{Hom}_W(h_\kappa(\mathfrak{N}, \varepsilon; W), W)$. Since W is a discrete valuation ring, we also have $\mathrm{Hom}_W(S_\kappa(\mathfrak{N}, \varepsilon; W), W) \cong h_\kappa(\mathfrak{N}, \varepsilon; W)$. □

This tells us

Corollary 2.11. *Let $H = h_\kappa(\mathfrak{N}, \varepsilon; A)$. Let V and V' be H–modules free of finite rank over A with an A–bilinear pairing $\langle\ ,\ \rangle : V \times V' \to A$. Define a formal q–expansion of $\Theta(v \otimes v')$ by $\mathbf{a}_p(y, \Theta(v \otimes v')) = \langle v|T(y), v'\rangle$. Then Θ gives an H–linear map of $V \otimes_A V'$ into $S_\kappa(\mathfrak{N}, \varepsilon; A)$ regarding $V \otimes_A V'$ as an H–module through V. If V is H–free of rank 1 and $\mathrm{Hom}_A(V, A) \cong V'$ by $\langle\ ,\ \rangle$ and $\langle hv, v'\rangle = \langle v, hv'\rangle$ for $h \in H$, Θ induces an isomorphism $V \otimes_H V' \cong S_\kappa(\mathfrak{N}, \varepsilon; A)$.*

Proof. Just apply the theorem to $\Theta(v \otimes v') \in \mathrm{Hom}_A(H, A) = S_\kappa(\mathfrak{N}, \varepsilon; A)$ given by $\Theta(v \otimes v')(h) = \langle hv, v'\rangle$. □

We can give an analytic proof of this fact when V is the space of quaternionic automorphic firms (see the discussions after [HMI, Proposition 2.51]).

2.5. **Quaternionic automorphic forms.** We now generalize the definition (SA1–3) of Hilbert modular forms to automorphic forms on a quaternion algebra $D_{/F}$.

We first define the standard level (open compact) subgroup in $G^D(\mathbb{A}^{(\infty)})$. We assume that $D_p = D \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong M_2(F_p)$ for the fixed prime p. For each prime ideal $\mathfrak{l}$ outside Σ^D, we fix an isomorphism $O_{D,\mathfrak{l}} \cong M_2(O_{\mathfrak{l}})$ so that for the p-adic place $\mathfrak{p}|p$ coming from $i_p \circ \sigma$, this isomorphism is given by the one already fixed: $O_D \hookrightarrow M_2(O_K) \xrightarrow{i_p\circ\sigma} M_2(O_{\mathfrak{p}})$. By means of these isomorphisms, we identify $D_{\mathfrak{l}}$ with $M_2(F_{\mathfrak{l}})$. Let $d(D)$ be the product of prime ideals in Σ^D. For an integral ideal $\mathfrak{N}_0$ of F prime to $d(D)$, putting $\mathfrak{N} = \mathfrak{N}_0 d(D)$, we define

$$\widehat{\Gamma}_0^D(\mathfrak{N}) = \left\{ x \in \widehat{O}_D^\times \middle| x_{\mathfrak{N}_0} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \quad \text{with} \quad c \in \mathfrak{N}_0 O_{\mathfrak{N}_0} \right\}, \tag{2.26}$$

where $\widehat{O}_D = O_D \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$, and $O_{\mathfrak{N}_0} = \prod_{\mathfrak{l}|\mathfrak{N}_0} O_{\mathfrak{l}}$. We set $S_0^D(\mathfrak{N}) = s^{-1}\widehat{\Gamma}_0^D(\mathfrak{N})s$ for $s = \left(\begin{smallmatrix} \delta^{(d(D))} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with a finite idele δ such that $\delta\widehat{O} \cap F = \mathfrak{d}$. Similarly we define $\Delta_0^D(\mathfrak{N}) \subset D^\times_{\mathbb{A}^{(\infty)}}$ so that it is the product of local components $\Delta_{\mathfrak{l}}$ which coincide with the local components of $\Delta_0(\mathfrak{N})$ as long as $\mathfrak{l} \nmid d(D)$ and $\Delta_{\mathfrak{l}} = O_{D,\mathfrak{l}}$ if $\mathfrak{l}|d(D)$.

We can think of the *double coset ring* $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N}))$ (which is a collection of formal linear combinations of double cosets $S_0^D(\mathfrak{N})xS_0^D(\mathfrak{N})$ for $x \in \Delta_0^D(\mathfrak{N})$ with multiplication given by convolution product; see 2.7 for a precise definition of the ring). We have $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})) \cong R(S_0^D(\mathfrak{N}), \Delta_0^{D*}(\mathfrak{N}))$. Here $\Delta_0^{D*}(\mathfrak{N})$ is the image of $\Delta_0^D(\mathfrak{N})$ under the involution ι ($x^\iota = N(x)x^{-1}$ for the reduced norm map $N : D \to F$). We call modules over these isomorphic double coset rings *Hecke modules*.

We have $\mathcal{T}(\mathfrak{l}) = S_0^D(\mathfrak{N}) \left(\begin{smallmatrix} \varpi_{\mathfrak{l}} & 0 \\ 0 & 1 \end{smallmatrix}\right) S_0^D(\mathfrak{N})$ and $\mathcal{T}(\mathfrak{l},\mathfrak{l}) = S_0^D(\mathfrak{N})\varpi_{\mathfrak{l}} S_0^D(\mathfrak{N})$ in $R(S_0^D(\mathfrak{N}), \Delta_0^{D*}(\mathfrak{N}))$ for $\mathfrak{l} \nmid d(D)$, because the local component $\Delta_0^D(\mathfrak{N})_{\mathfrak{l}}$ at $\mathfrak{l}$ of $\Delta_0^D(\mathfrak{N})$ is identical to $\Delta_0(\mathfrak{N})_{\mathfrak{l}}$. For $\mathfrak{l}|d(D)$, we take $\alpha_{\mathfrak{l}} \in O_{D,\mathfrak{l}}$ so that its reduced norm generates $\mathfrak{l}O_{\mathfrak{l}}$. Then we define $\mathcal{T}(\mathfrak{l}) = S_0^D(\mathfrak{N})\alpha_{\mathfrak{l}} S_0^D(\mathfrak{N})$ for $\mathfrak{l}|d(D)$, and we have

$$R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N})) \cong R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N})). \tag{2.27}$$

These elements $\mathcal{T}(\mathfrak{l})$ and $\mathcal{T}(\mathfrak{l},\mathfrak{l})$ (indexed by primes $\mathfrak{l}$) are generators of the commutative ring $R(S_0^D(\mathfrak{N}), \Delta_0^{D*}(\mathfrak{N}))$ over $\mathbb{Z}$ (see Lemma 2.21). The above isomorphism brings $\mathcal{T}(\mathfrak{l})$ and $\mathcal{T}(\mathfrak{l},\mathfrak{l})$ to the corresponding elements in the right-hand side. As an operator on the space of automorphic forms, $\mathcal{T}(\mathfrak{l},\mathfrak{l})$ induces the central action $\langle \mathfrak{l} \rangle$ of $\varpi_{\mathfrak{l}}$.

A particular feature of quaternionic automorphic forms is that they are often vector valued, though Hilbert modular forms defined in (SA1–3) are scalar valued. Here we define the space in which quaternionic automorphic forms have values. For a given ring R, we consider the following module $L(\kappa^*; R)$ of the multiplicative semi-group $M_2(R)$. Let $\kappa^* = (\kappa_1 + I, \kappa_2)$ and put $n = \kappa_2 - \kappa_1 - I \in \mathbb{Z}[I]$, which is the restriction of $\kappa^* \in X(T_G)$ to $T \subset G_1$, and we confirm $(\kappa|_T)^* = k - 2I = n$. We suppose that $n \geq 0$ (i.e., $n_\sigma \geq 0$ for all $\sigma \in I$), and we consider polynomials with coefficients in R of $(X_\sigma, Y_\sigma)_{\sigma \in I}$ homogeneous of degree n_σ for each pair (X_σ, Y_σ). The collection of all such polynomials forms an R-free module $L(\kappa^*; R)$ of rank $\prod_\sigma (n_\sigma + 1)$.

As before, we write v for the fixed place p or ∞; so, the base ring $\mathcal{B}$ is W if $v = p$ and $\mathbb{C}$ if $v = \infty$. Suppose that R is a $\mathcal{B}$-algebra. Then $i_v(\sigma(\delta_v))$ (which we write simply $\sigma(\delta_v)$) for $\delta \in G^D(\mathbb{A})$ can be regarded as an element in $M_2(R)$. Take a Neben character ε as in (ex1–4) of Section 2.2 with $\varepsilon^-|_{T(\widehat{\mathbb{Z}})}$ factoring through $(O/\mathfrak{N}_0)^\times$.

We define ε_j $(j = 1, 2)$ as in (ex1) and extend ε to $\Delta_0^D(\mathfrak{N})$ by $\varepsilon_D^\Delta(\delta) = \varepsilon_2(N(\delta))\varepsilon^-(a)$ if $\delta_{\mathfrak{N}_0} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Since $s \mapsto \varepsilon_D^\Delta(s)$ and $z \mapsto \varepsilon_+(z)$ coincide on $Z(\mathbb{A}^{(\infty)}) \cap S_0^D(\mathfrak{N})$, we may extend ε to a character $\varepsilon_D^S : S_0^D(\mathfrak{N})Z(\mathbb{A}^{(\infty)}) \to R^\times$ by $\varepsilon_D^S(zu) = \varepsilon_D^\Delta(u)\varepsilon_+(z)$ for $z \in Z(\mathbb{A}^{(\infty)})$ and $u \in S_0^D(\mathfrak{N})$. We let $\Delta_0^D(\mathfrak{N})$ and $Z(\mathbb{A})S_0^D(\mathfrak{N})G^D(\mathbb{R})^+$ act on $L(\kappa^*; R)$

as follows.

$$\begin{aligned}\delta\cdot\Phi\left(\left(\begin{smallmatrix}X_\sigma\\ Y_\sigma\end{smallmatrix}\right)\right)&=\varepsilon_D^\Delta(\delta)^{-1}N(\delta_v)^{\kappa_1}\Phi\left(\sigma(({}^s\delta_v)^\iota)\left(\begin{smallmatrix}X_\sigma\\ Y_\sigma\end{smallmatrix}\right)\right),\\ (zu)\cdot\Phi\left(\left(\begin{smallmatrix}X_\sigma\\ Y_\sigma\end{smallmatrix}\right)\right)&=\varepsilon_D^S(z^{(\infty)}u^{(\infty)})^{-1}N(z_vu_v)^{\kappa_1}\Phi\left(\sigma({}^s(z_vu_v)^\iota)\left(\begin{smallmatrix}X_\sigma\\ Y_\sigma\end{smallmatrix}\right)\right).\end{aligned}\tag{2.28}$$

Here $z\in Z(\mathbb{A})$, $u\in S_0^D(\mathfrak{N})G^D(\mathbb{R})$, and $\delta^\iota=N(\delta)\delta^{-1}$ and ${}^s\delta=s\delta s^{-1}$ for s given just below (2.26). We write $L(\kappa^*\varepsilon;R)$ for the module $L(\kappa^*;R)$ with this $(\Delta_0^D(\mathfrak{N}),Z(\mathbb{A})S_0^D(\mathfrak{N})G^D(\mathbb{R}))$-action.

By computation, $z\in Z(\mathbb{A})$ acts on $L(\kappa^*\varepsilon;R)$ through scalar multiplication by $\varepsilon_+(z)^{-1}z_v^{\kappa_1+\kappa_2-I}$; in particular, $\epsilon\in O_+^\times\subset Z(\mathbb{A})$ acts trivially on $L(\kappa^*\varepsilon;R)$. If S is a sufficiently small open compact subgroup so that $S\cap Z(\mathbb{Q})\subset O_+^\times$ in $G(\mathbb{A}^{(\infty)})$, central elements in $\Gamma_x=xSx^{-1}\cap G^D(\mathbb{Q})$ act trivially on $L(\kappa^*\varepsilon;R)$.

Assume that $v=\infty$. We state a definition of automorphic forms on $G^D(\mathbb{A})$ now for a division algebra D. For each κ, we define $\kappa^D\in\mathbb{Z}[I]^2$ by $\kappa^D=(\kappa_1^D,\kappa_2^D)$ for $\kappa_j^D=\sum_{\sigma\in I^D}\kappa_{j,\sigma}\sigma$. Thus κ^D is the projection of κ to $\mathbb{Z}[I^D]^2$. Similarly, we define κ_D by the projection of κ to $\mathbb{Z}[I_D]^2\subset\mathbb{Z}[I]^2$.

With each $\kappa\in\mathbb{Z}[I]^2=X(T_G)$, we associate an automorphy factor,

$$J_\kappa^D(g,z)=\det(g)^{\kappa_{D,1}-I_D}j(g,z)^{\kappa_{D,2}-\kappa_{D,1}+I_D},\tag{2.29}$$

for $g\in G^D(\mathbb{R})$ and $z\in\mathfrak{Z}_D$. We write $\kappa^{*,D}$ for the projection of κ^* to $\mathbb{Z}[I^D]$. We take a subset $\Theta\subset I_D$ and split I_D as $I_D=\Theta\sqcup\overline{\Theta}$. Define for $z\in\mathfrak{Z}_D$

$$z_\sigma^\Theta=\begin{cases}\overline{z}_\sigma & \text{if }\sigma\in\overline{\Theta},\\ z_\sigma & \text{if }\sigma\in\Theta\ .\end{cases}$$

Then, if $I_D\neq\emptyset$, we define $S^D_{\kappa,\Theta}(\mathfrak{N},\varepsilon;\mathbb{C})$ to be the space of functions f on $G^D(\mathbb{A})$ with values in the left $C_{\mathbf{i}}^D$-module $L(\kappa^{*,D},\mathbb{C})$ satisfying the following conditions.

(SB1) We have the following automorphy

$$f(\alpha xuz)=\varepsilon_+(z)\varepsilon_D^\Delta(u^{(\infty)})u_\infty^{-1}\cdot f(x)J_\kappa^D(u_\infty,\mathbf{i}^\Theta)^{-1}$$

for all $\alpha\in G^D(\mathbb{Q})$, $z\in Z(\mathbb{A})$, and $u\in S_0^D(\mathfrak{N})C_{\mathbf{i}}^D$, where $G^D(\mathbb{R})^+$ is the identity-connected component of $G^D(\mathbb{R})$. Here $f(x)\mapsto u_\infty\cdot f(x)$ is the action of the I^D-component u_∞^D of u_∞ on $L(\kappa^{*,D};\mathbb{C})$;

(SB2) Choosing $u\in G^D(\mathbb{R})$ with $u^D=1$ and $u(\mathbf{i})=z$ for each $z\in\mathfrak{Z}_D$, define a function $f_g:\mathfrak{Z}_D\to\mathbb{C}$ by $f_g(z)=f(gu_\infty)J_\kappa(u_\infty,\mathbf{i})$ for each $g\in G(\mathbb{A}^{(\infty)})$. Then, for all g, f_g is a function on $\mathfrak{Z}_D$ holomorphic in z_σ for $\sigma\in\Theta$ and antiholomorphic in z_σ for $\sigma\in\overline{\Theta}$.

If $\Theta=I_D$, we simply write $\mathcal{S}_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})$ for $\mathcal{S}^D_{\kappa,I_D}(\mathfrak{N},\varepsilon;\mathbb{C})$, which is the space of holomorphic *automorphic forms* on $G^D(\mathbb{A})$ of level $S_0(\mathfrak{N})$ and of weight (κ,ε).

When $I^D=I$ ($\Leftrightarrow I_D=\emptyset$), the variety $Y^D(S)$ is a finite set of points; so, the condition (SB2) is empty, and we may replace $\mathbb{C}$ in (SB1) by any ring $A\subset\mathbb{C}$ with values of ε. Writing $\mathcal{M}_\kappa^D(\mathfrak{N},\varepsilon;A)$ for the space of functions satisfying (SB1) in this definite case, we need to take $S_\kappa^D(\mathfrak{N},\varepsilon;A)$ to be the following quotient: $S_\kappa^D(\mathfrak{N},\varepsilon;A)=\mathcal{M}_\kappa^D(\mathfrak{N},\varepsilon;A)/Iv(\mathfrak{N},\varepsilon;A)$, where $Iv(\mathfrak{N},\varepsilon;A)$ is the subspace made up of functions in $\mathcal{S}_\kappa^D(\mathfrak{N},\varepsilon;A)$ factoring through the reduced norm map $N:G^D(\mathbb{A})\to T(\mathbb{A})$. If $\kappa_2-\kappa_1\neq I$

or ε is nontrivial for some $x \in G^D(\mathbb{A})$ with $N(x) = 1$, $Iv(\mathfrak{N}, \varepsilon; \mathbb{C}) = 0$; so, no modification is necessary. Decomposing $S_0^D(\mathfrak{N})y^\iota S_0^D(\mathfrak{N})$ for $y \in \Delta_0^D(\mathfrak{N})$ into $\bigsqcup_h hS_0^D(\mathfrak{N})$, we shall define the action of $R(S_0^D(\mathfrak{N}), \Delta_0^D(\mathfrak{N}))$ on $S_\kappa^D(\mathfrak{N}, \varepsilon; A)$ by

$$f|[S_0^D(\mathfrak{N})y^\iota S_0^D(\mathfrak{N})](g) = \sum_h h \cdot f(gh) = \sum_h h^{-\iota} \cdot f(gh^{-\iota}). \tag{2.30}$$

When $D = M_2(F)$, we simply write $S_{\kappa,\Theta}(\mathfrak{N}, \varepsilon; \mathbb{C})$ for the space of functions on $GL_2(F_\mathbb{A})$ satisfying (SB1-2) and (SA3). As we have stated for holomorphic Hilbert modular forms, the Hecke algebra $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$ naturally acts on $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$.

We now extend the definition of automorphic forms on $G^D(\mathbb{A})$ and on $\mathfrak{Z}_D$ to more general level groups $U \subset G^D(\widehat{\mathbb{Z}})$. Let $D_\mathbb{H} = \prod_{\sigma \in I^D} D_\sigma \cong \mathbb{H}^{I^D}$. For simplicity, we write V for $L(\kappa^{*,D}; \mathbb{C})$. Let $U \subset \widehat{O}_D^\times$ be an open subgroup. Thus the image $\Gamma(U)$ of $U \cap G^D(\mathbb{Q})$ in $PG^D(\mathbb{R})^+$ is a congruence subgroup of $\Gamma(1)$. An automorphic form of weight κ and of analytic type $(\Theta, \overline{\Theta})$ (of level U) is a function $f : G^D(\mathbb{A}) \to V$ satisfying the following two conditions (sb1-2).

(sb1) $f(\gamma x s) = s_\infty^{-1} \cdot f(x) J_\kappa^D(s_\infty, \mathbf{i}^\Theta)^{-1}$ *for* $\gamma \in G^D(\mathbb{Q})$ *and* $s \in U \cdot C_\mathbf{i}^D$. *The action of* $s_\infty : f(x) \mapsto s_\infty \cdot f(x)$ *is through the action of* $D_\mathbb{H}^\times$ *on* V.

For each $g \in G^D(\mathbb{A})$ with $g_\infty = 1$, as before, we define a function $f_g : \mathfrak{Z}_D \to V$ by

$$f_g(z) = s_\infty f(g s_\infty) J_\kappa^D(s_\infty, \mathbf{i}^\Theta)$$

for $s_\infty \in G^D(\mathbb{R})^+$ with $s_\infty(\mathbf{i}) = z$. If $z = s_\infty(\mathbf{i}) = s'_\infty(\mathbf{i})$, we can write $s'_\infty = s_\infty c$ with $c \in F_\infty^\times C_\infty$. Thus

$$\begin{aligned} s'_\infty f(g s'_\infty) J_\kappa^D(s'_\infty, \mathbf{i}^\Theta) &= s_\infty c f(g s_\infty c) J_\kappa^D(s_\infty c, \mathbf{i}^\Theta) \\ &= s_\infty c c^{-1} f(g s_\infty c) J_\kappa^D(c, \mathbf{i}^\Theta)^{-1} J_\kappa^D(c, \mathbf{i}^\Theta) J_\kappa^D(s_\infty, \mathbf{i}^\Theta) = f_g(z). \end{aligned}$$

This shows that f_g is well defined independently of the choice of s_∞. For

$$\gamma \in \Gamma_g(U) = \left(g^{-1} U G^D(\mathbb{R})^+ g\right) \cap G^D(\mathbb{Q})$$

which is a congruence subgroup of the unit group of another maximal order $O_D^g = g^{-1} O_D g \cap D$, we see

$$\begin{aligned} f_g(\gamma(z)) &= \gamma_\infty s_\infty \cdot f(g \gamma_\infty s_\infty) J_\kappa^D(\gamma_\infty s_\infty, \mathbf{i}^\Theta) \\ &= \gamma_\infty s_\infty \cdot f(\gamma g \gamma^{(\infty)-1} g^{-1} g s_\infty) J_\kappa^D(\gamma_\infty s_\infty, \mathbf{i}^\Theta) \\ &= \gamma_\infty s_\infty \cdot f(\gamma g s_\infty) J_\kappa^D(\gamma_\infty s_\infty, \mathbf{i}^\Theta) \quad (g \gamma^{(\infty)-1} g^{-1} \in U) \\ &= \gamma_\infty s_\infty \cdot f(g s_\infty) J_\kappa^D(s_\infty, \mathbf{i}^\Theta) J_\kappa^D(\gamma, z^\Theta) \\ &= \gamma_\infty \cdot f_g(z) J_\kappa^D(\gamma, z^\Theta). \end{aligned}$$

Thus f_g is a modular form on $\Gamma_g(U)$. Now, for $\sigma \in I_B$, we impose the holomorphy or the anti-holomorphy condition:

(sb2) $\frac{\partial f_g}{\partial z_\sigma} = 0$ *for* $\sigma \in \overline{\Theta}$ *and* $\frac{\partial f_g}{\partial \overline{z}_\sigma} = 0$ *for* $\sigma \in \Theta$.

Except for the cases where D is definite and $\dim V = 1$ and $D = M_2(F)$, the space of functions satisfying the above conditions (sb1-2) will be denoted as $S^D_{\kappa,\Theta}(U;\mathbb{C})$. If $\Theta = I_D$, we drop the subscript Θ. When $D = M_2(F)$, we write $S_{\kappa,\Theta}(U;\mathbb{C})$ (dropping the superscript $D = M_2(F)$) for the space of functions satisfying the above conditions (sb1-2) and (SA3). We write $G_{\kappa,\Theta}(U;\mathbb{C})$ for the space of functions satisfying (sb1-2) with polynomial growth towards the cusps of $Y^D(U)$ when $D = M_2(F)$.

In order to define the space $S^D_\kappa(U;\mathbb{C})$ as before, when D is definite and $\dim V = 1$, we write $\mathcal{M}(U;\mathbb{C})$ for the space of functions: $G^D(\mathbb{A}) \to V$ satisfying (sb1-2). Then writing $Iv(U;\mathbb{C})$ for the subspace of functions in $\mathcal{M}(U;\mathbb{C})$ factoring through the reduced norm map $N : G^D(\mathbb{A}) \to T(\mathbb{A})$, we define $S^D_\kappa(U;\mathbb{C}) = \mathcal{M}(U;\mathbb{C})/Iv(U;\mathbb{C})$. We let each double coset $[UxU]$ of $x \in G^D(\mathbb{A}^{(\infty)})$ act on $S_\kappa(U;\mathbb{C})$ as a Hecke operator. In other words, decomposing $UxU = \bigsqcup_y yU$, the action is given by

$$f|[UxU](g) = \sum_g f(gy).$$

2.6. The Jacquet–Langlands correspondence.

As before, we identify $D^{(d(B)\infty)}_{\mathbb{A}}$ with $M_2(F^{(d(B)\infty)}_{\mathbb{A}})$. We state the theorem of Jacquet–Langlands and Shimizu in the following way.

Theorem 2.12. *For an open compact subgroup $U^{(d(D))}$ of $GL_2(F^{(d(D)\infty)}_{\mathbb{A}})$, define open subgroups of $G^D(\mathbb{A}^{(\infty)})$ and $GL_2(F^{(\infty)}_{\mathbb{A}})$ by*

$$U_D = U^{(d(D))} \times \prod_{\mathfrak{q}|d(D)} O^\times_{D,\mathfrak{q}} \subset G^D(\mathbb{A}^{(\infty)}),$$
$$U_0(d(D)) = U^{(d(D))} \times S_0(d(D))_{d(D)} \subset GL_2(F^{(\infty)}_{\mathbb{A}}).$$

Then we have a $\mathbb{C}$-linear embedding $i : S^D_\kappa(U_D;\mathbb{C}) \hookrightarrow S_\kappa(U_0(d(B));\mathbb{C})$ for all weights κ with $k_\sigma \geq 2$ and for all $\sigma \in I^D$, where $k_\sigma = \kappa_{2,\sigma} - \kappa_{1,\sigma} + 1$. For all $x \in GL_2(F^{(d(D)\infty)}_{\mathbb{A}})$, we have $i \circ [U_D x U_D] = [U_0(d(B))xU_0(d(B))] \circ i$. The image of this embedding only depends on $d(D)$ and is made up of cusp forms in $S_\kappa(U_0(d(B));\mathbb{C})$ new at all primes $\mathfrak{q}|d(D)$. In particular, if $d(D) = 1$, the above morphism is a surjective isomorphism.

The word "new" means that a cusp form in the image of i cannot be in the space generated by $\sum_{\mathfrak{q}|d(D)} S_\kappa(U_0(d(D)/\mathfrak{q});\mathbb{C})$ under the action $f(h) \mapsto x \cdot f(h) = f(hx)$ of $x = \left(\begin{smallmatrix} \varpi_\mathfrak{q} & 0 \\ 0 & 1 \end{smallmatrix}\right) \in G(F_{d(D)})$; in other words, any $f \in \mathrm{Im}(i)$ is orthogonal to $x \cdot g(h) = g(hx)$ for any $g \in \sum_{\mathfrak{q}|d(D)} S_\kappa(U_0(d(D)/\mathfrak{q});\mathbb{C})$ and $x \in G(F_{d(D)})$, that is, $\int_{G(\mathbb{Q})\backslash G(\mathbb{A})/Z(\mathbb{A})} \overline{f(h)} x \cdot g(h) dh = 0$ for a right invariant Haar measure dh on $G(\mathbb{Q})\backslash G(\mathbb{A})/Z(\mathbb{A})$. We write $S^{new}_\kappa(U_0(d(D));\mathbb{C})$ for the subspace of new forms. Some more comments on the new forms will be given after the classification lemma (Lemma 2.23) of local representations.

Applying the above theorem to $U = S^D_0(\mathfrak{N})$, we have

Corollary 2.13. *Suppose $\mathfrak{N} = \mathfrak{N}_0 d(D)$ for an integral ideal $\mathfrak{N}_0$ prime to $d(D)$. Identify $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$ and $R(S^D_0(\mathfrak{N}), \Delta^D_0(\mathfrak{N}))$ as in (2.27). Then we have an $R(S_0(\mathfrak{N}), \Delta_0(\mathfrak{N}))$-linear embedding $\iota : S^D_\kappa(\mathfrak{N}, \varepsilon;\mathbb{C}) \hookrightarrow S_\kappa(\mathfrak{N}, \varepsilon;\mathbb{C})$ for all weights κ with $k_\sigma \geq 2$ for all*

$\sigma \in I^D$. The image of this embedding only depends on $d(D)$ and is made up of cusp forms in $S^{new}_{\kappa}(\mathfrak{N}, \varepsilon; \mathbb{C})$ new at all primes $\mathfrak{q}|d(D)$. In particular, if $d(D) = 1$, the above morphism is a surjective isomorphism.

Since known proofs of this theorem require harmonic analysis and a good knowledge of representation theory of adele groups (especially the theta correspondence), we do not give a full proof of this fact, and we quote here some references where one can find a proof. The above formulation of the theorem was given in [PAF] Section 4.3, [H81] Section 2 and [H88] Theorem 2.1, where one can find an exposition of how to deduce this result from the original result of Jacquet–Langlands (whose exposition can be found in [AAG] Section 10). The result of Jacquet–Langlands is more general than the above theorem in the following points:

- (JL1) It covers any real-analytic automorphic forms including Maass forms and any level group U (not necessarily under the condition $U_{\mathfrak{q}} = O^{\times}_{D,\mathfrak{q}}$ for $\mathfrak{q}|d(D)$);
- (JL2) It gives also a description of the relation of the local representations of $D^{\times}_{\mathfrak{q}}$ and $GL_2(F_{\mathfrak{q}})$ for $\mathfrak{q}|d(D)$;
- (JL3) Their approach is more representation theoretic: It gives a precise correspondence of automorphic representations π_D realized on the L^2–space of functions on $G^D(\mathbb{Q})\backslash G^D(\mathbb{A})$ (with a given central character χ) and π realized on the L^2–space of functions on $GL_2(F)\backslash GL_2(F_{\mathbb{A}})$ with the same central character χ. If we factorize $\pi_D = \otimes_v \pi_v(D)$ and $\pi = \otimes_v \pi_v$, $\pi_v(D) \cong \pi_v$ as long as $D_v \cong M_2(F_v)$. For v with division D_v, the correspondence $\pi_v(D) \mapsto \pi_v$ is given by the Weil representation π_v associated to $\pi_v(D)$ with respect to the norm form of D/F.

We will give a slightly more detailed description of the correspondence $\pi_D \mapsto \pi$ in (JL3) at the end of the following subsection.

Shimizu originally proved a result close to this theorem when $d(D) = 1$ (and a weaker assertion when $d(D) \neq 1$). Later he explictly realized the correspondence using theta series.

The method of proof in the original work of Shimizu and Jacquet–Langlands is to compute the trace of the operator $[UxU]$ on both sides by means of the Eichler-Selberg trace formula. Eichler initiated the comparison of such traces (and he proved results slightly weaker than the above corollary for definite quaternion algebras over $\mathbb{Q}$; [E]). We will give an exposition later of this for $F = \mathbb{Q}$ along a line closer to the treatment of Eichler, Shimizu and Shimura.

We can actually embed $S^D_{\kappa,\Theta}(U; \mathbb{C})$ into $S_{\kappa}(U_0(d(D)); \mathbb{C})$ in a way compatible with the Hecke operator action. This fact follows from Theorem 2.12 and the following result:

Proposition 2.14. *For each $\Theta \subset I_D$, there is an isomorphism: $S^D_{\kappa}(U; \mathbb{C}) \cong S^D_{\kappa;\Theta}(U; \mathbb{C})$ which is linear under the action of $[UxU]$ for all $x \in G^D(\mathbb{A}^{(\infty)})$.*

Proof. Let $j = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in GL_2(\mathbb{R})$. We define $t(\Theta) \in D^{\times}_{\infty}$ by

$$t(\Theta)_{\sigma} = \begin{cases} j & \text{if } s \in \overline{\Theta}, \\ 1 & \text{otherwise.} \end{cases}$$

Then $f(x) \mapsto f(xt(\Theta))$ gives the isomorphism i_Θ. Left multiplication by $t(\Theta)$ only takes effect at infinite places and hence commutes with $[UxU]$. □

In the previous section, we have only defined the space $S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})$ of quaternionic automorphic forms with coefficients in $\mathbb{C}$. However, we have a good integral structure on $S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})$ over $\mathcal{W}$ (see [HMI, Chapter 3]). Now we assume to have such an integral structure, and for each $\mathcal{W}$-algebra A, we write $S_\kappa^D(\mathfrak{N},\varepsilon;A)$ for the space of A-integral automorphic forms. We thus suppose

(R1) $S_\kappa^D(\mathfrak{N},\varepsilon;A) = S_\kappa^D(\mathfrak{N},\varepsilon;\mathcal{W}) \otimes_{\mathcal{W}} A$ for every $\mathcal{W}$-algebra R;
(R2) $S_\kappa^D(\mathfrak{N},\varepsilon;\mathcal{W}) \subset S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})$ is stable under the action of $R(S_0^D(\mathfrak{N}),\Delta_0^D(\mathfrak{N}))$.

We then extends the Hecke action of $R(S_0^D(\mathfrak{N}),\Delta_0^D(\mathfrak{N}))$ to $S_\kappa^D(\mathfrak{N},\varepsilon;A)$ by the identity (R1). Thus $S_\kappa^D(\mathfrak{N},\varepsilon;W)$ becomes a module over the Hecke algebra $h_\kappa(\mathfrak{N},\varepsilon;W)$. Let $S_\kappa^D(\mathfrak{N},\varepsilon;A)^*$ be the A-linear dual of $S_\kappa^D(\mathfrak{N},\varepsilon;A)$. We regard $S_\kappa^D(\mathfrak{N},\varepsilon;A)^*$ as $h_\kappa(\mathfrak{N},\varepsilon;A)$-module by the adjoint action. Then by Proposition **??**, we have a morphism of $h_\kappa(\mathfrak{N},\varepsilon;W)$–modules

$$f_A : S_\kappa^D(\mathfrak{N},\varepsilon;A) \otimes_A S_\kappa^D(\mathfrak{N},\varepsilon;A)^* \to S_\kappa(\mathfrak{N},\varepsilon;A).$$

For a suitable choice of $\phi \in S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C})^*$, $g \mapsto f_{\mathbb{C}}(g \otimes \phi)$ induces an embedding $\iota : S_\kappa^D(\mathfrak{N},\varepsilon;\mathbb{C}) \to S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ in Corollary 2.13. By the faithful flatness of $\mathbb{C}$ over the field of fractions $\mathcal{K}$ of $\mathcal{W}$, ι induces an embedding $\iota : S_\kappa^D(\mathfrak{N},\varepsilon;\mathcal{K}) \to S_\kappa(\mathfrak{N},\varepsilon;\mathcal{K})$ which is $h_\kappa(\mathfrak{N},\varepsilon;\mathcal{K})$–linear. Then extending scalar to $\mathcal{K}$-algebra A, we find

Corollary 2.15. *Suppose $\mathfrak{N} = \mathfrak{N}_0 d(D)$ for an integral ideal $\mathfrak{N}_0$ prime to $d(D)$. Let A be a $\mathcal{K}$-algebra. Then we have an $h_\kappa(\mathfrak{N},\varepsilon;A)$-linear embedding*

$$\iota_A : S_\kappa^D(\mathfrak{N},\varepsilon;A) \hookrightarrow S_\kappa(\mathfrak{N},\varepsilon;A)$$

for all weights κ with $k_\sigma \geq 2$ for all $\sigma \in I^D$. In particular, if $d(D) = 1$, the above morphism is a surjective isomorphism.

Exercise 2.16.

(1) *If $\chi : \mathbb{H}^\times \to \mathbb{C}^\times$ is a continuous character, prove that χ factors through the reduced norm $N : \mathbb{H}^\times \to \mathbb{R}^\times$.*
(2) *Prove that $v \in \mathbb{Z}[I]$ is an integer multiple of $\sum_\sigma \sigma$ if $\epsilon^v = 1$ for all ϵ in a subgroup of finite index in $O^\times$.*
(3) *Prove that if U is an open compact subgroup of $G^D(\mathbb{A}^{(\infty)})$, the double coset UyU for $y \in G^D(\mathbb{A}^{(\infty)})$ contains only finitely many left (and right) cosets*
(4) *Give a detailed proof that $(f|j)_x(z) = f_x(-\overline{z})$ for any $x \in G^D(\mathbb{A}^{(\infty)})$ if $D = M_2(\mathbb{Q})$ and $f \in S_\kappa(U;\mathbb{C})$, where $f|j(g) = f(gj)$ and j is as in the proof of* Proposition 2.14.
(5) *Prove that for a given κ and U, there are only finitely many Hecke characters $\chi : F^\times U_F \backslash F_{\mathbb{A}}^\times \to \mathbb{C}^\times$ with $\chi(z_\infty) = z_\infty^{\kappa_1+\kappa_2-I}$, where $U_F = U \cap Z(\mathbb{A}^{(\infty)})$ regarded as an open subgroup of $(F_{\mathbb{A}}^{(\infty)})^\times$.*

2.7. **Local representations of $GL(2)$.** We summarize here well-known results from representation theory of the local group $\mathcal{G} = GL_2(F_{\mathfrak{q}})$ for a prime ideal $\mathfrak{q}$ in order to supplement the representation theoretic description in (JL1-3) of the Jacquet–Langlands correspondence.

A representation V of $\mathcal{G}$ with coefficients in a field K of characteristic 0 is called *smooth* if for each $v \in V$, we find an open subgroup $S \subset \mathcal{G}$ that leaves v invariant. If furthermore, $H^0(S, V)$ is finite-dimensional for all open subgroups S, V is called *admissible*. Let π be an admissible semi-simple representation of $\mathcal{G}$ on a vector space $V = V(\pi)$ over a field K of characteristic 0.

Let $\mathcal{U}(F_{\mathfrak{q}})$ be the maximal upper unipotent subgroup of $\mathcal{G}$, and write

$$\mathcal{U}(O_{\mathfrak{q}}) = \left\{ \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right) \middle| u \in O_{\mathfrak{q}} \right\} \subset \mathcal{U}(F_{\mathfrak{q}}).$$

Let $\mathcal{B}(F_{\mathfrak{q}})$ (resp. $\mathcal{B}(O_{\mathfrak{q}})$) be the normalizer of $\mathcal{U}(F_{\mathfrak{q}})$ in $\mathcal{G}$ (resp. $GL_2(O_{\mathfrak{q}})$). Then $\mathcal{B}(A) = \mathcal{T}(A) \ltimes \mathcal{U}(A)$ for the subgroup $\mathcal{T}(A)$ of diagonal matrices (for $A = F_{\mathfrak{q}}$ and $O_{\mathfrak{q}}$). They are made up of upper triangular matrices. We have a Haar measure du of $\mathcal{U}(F_{\mathfrak{q}})$ with $\int_{\mathcal{U}(O_{\mathfrak{q}})} du = 1$. We define

$$V(\mathcal{B}) = V(\mathcal{B}, \pi) = \left\{ v - \pi(n)v \in V(\pi) \middle| v \in V(\pi),\ n \in \mathcal{U}(F_{\mathfrak{q}}) \right\}$$

as a $\mathcal{T}(F_{\mathfrak{q}})$-module, and put $V_{\mathcal{B}} = V_{\mathcal{B}}(\pi) = V/V(\mathcal{B})$, which is called the *Jacquet module* of V. The notion of Jacquet module is useful in the classification of admissible representations of a local group like $\mathcal{G}$ (see [BZ]). By this definition, $V \mapsto V_{\mathcal{B}}$ is a right exact functor. For each $w = v - \pi(n)v \in V(\mathcal{B})$, we take a sufficiently large open compact subgroup $\mathcal{U}_w \subset \mathcal{U}(F_{\mathfrak{q}})$ containing n. Then we see that $\int_{\mathcal{U}} \pi(u)v du = 0$ for every open subgroup $\mathcal{U}$ with $\mathcal{U}_w \subset \mathcal{U} \subset \mathcal{U}(F_{\mathfrak{q}})$. If $\int_{\mathcal{U}} \pi(u)v du = 0$ for every sufficiently large open subgroup $\mathcal{U}$ of $\mathcal{U}(F_{\mathfrak{q}})$, for the stabilizer $\mathcal{U}'$ of v in $\mathcal{U}$, we find

$$\left(\int_{\mathcal{U}'} du\right) v = \left(\int_{\mathcal{U}'} du\right) v - \int_{\mathcal{U}} \pi(u)v du = \left(\int_{\mathcal{U}'} du\right) \sum_{n \in \mathcal{U}/\mathcal{U}'} (v - \pi(n)v)$$

which is in $V(\mathcal{B})$. Thus $V(\mathcal{B})$ is the collection of v with $\int_{\mathcal{U}} \pi(u)v du = 0$ for every sufficiently large open subgroup $\mathcal{U}$ of $\mathcal{U}(F_{\mathfrak{q}})$. By this fact, the functor $V \mapsto V_{\mathcal{B}}$ is left exact, and we conclude that the association is an exact functor.

Exercise 2.17. *Explain why $V \mapsto V_{\mathcal{B}}$ is exact (sending exact sequences of representations to exact sequences of K-vector spaces).*

For a smooth character λ of $\mathcal{T}(F_{\mathfrak{q}})$ (regarding it as a character of $\mathcal{B}$ via the projection $\mathcal{B} \twoheadrightarrow \mathcal{T} \cong \mathcal{B}/\mathcal{U}$), the *smooth induction* from $\mathcal{B}$ of λ is defined by

$$\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}(\lambda) = \{f : \mathcal{G} \to V(\lambda) | f\text{: smooth},\ f(bg) = \lambda(b)g(g)\ \forall b \in \mathcal{B}(F_{\mathfrak{q}})\}, \tag{2.31}$$

on which we let $\mathcal{G}$ act by $g \cdot f(x) = f(xg)$. Here the word "smooth" means that for each $f \in \mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} V(\lambda)$, we find an open compact subgroup S such that $f(xk) = f(x)$ for all $k \in S$. Thus $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \lambda$ is smooth as a representation of $\mathcal{G}$ by definition. Since the smooth induction preserves admissibility (see [BZ] 2.3), $V = \mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \lambda$ has composition series $\{0\} \subset V_1 \subset V_2 \subset \cdots \subset V$ stable under the action of $\mathcal{G}$, and hence its semi-simplification $(\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \lambda)^{ss} = \bigoplus_j V_{j+1}/V_j$ is well-defined as an admissible $\mathcal{G}$-module.

Put $\widetilde{\lambda} = \delta_{\mathcal{B}}^{1/2}\lambda$ for the module character $\delta_{\mathcal{B}}$ of $\mathcal{B}$:

$$\int_{\mathcal{U}(F_{\mathfrak{q}})} \phi(u)du = \delta_{\mathcal{B}}(b) \int_{\mathcal{U}(F_{\mathfrak{q}})} \phi(bxb^{-1})du \quad \text{(for all } \phi).$$

Exercise 2.18. *Prove* $\delta_{\mathcal{B}}\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right) = |a^{-1}d|_{\mathfrak{q}}$ *for the standard absolute value* $|x|_{\mathfrak{q}} = |O_{\mathfrak{q}}/xO_{\mathfrak{q}}|^{-1}$ $(x \in O_{\mathfrak{q}} - \{0\})$.

If $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\widetilde{\lambda}$ is irreducible, we call the induced representation *principal* (or in the principal series). If $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\widetilde{\lambda}$ is reducible, only one factor of the composition series is infinite dimensional, which is called *special* (or a Steinberg representation). Since $\mathcal{T}$ is diagonal, we can write $\lambda\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right) = \lambda_1(a)\lambda_2(d)$. Then a principal (resp. special) representation $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\widetilde{\lambda}$ is denoted by $\pi(\lambda_1, \lambda_2)$ (resp. $\sigma(\lambda_1, \lambda_2)$).

The following results are well known (e.g., [BZ]),

(π1) $\mathrm{Hom}_{\mathcal{B}}(V_{\mathcal{B}}, V(\lambda)) \cong \mathrm{Hom}_{\mathcal{G}}(V, \mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} V(\lambda))$ [BZ] 1.9 (reciprocity);

(π2) If π is absolutely irreducible, then $\dim_K V_{\mathcal{B}} \leq |\mathfrak{W}| = 2$, where $\mathfrak{W}$ is the Weyl group of $\mathcal{T}$ in $\mathcal{G}$ (see [BZ] 2.9);

(π3) If π is absolutely irreducible and $V_{\mathcal{B}} \neq 0$, then we have a surjective linear map $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\lambda \twoheadrightarrow V$ of $\mathcal{G}$-modules for a character $\lambda : \mathcal{T} \to K^{\times}$ (see [BZ] 2.4–5);

(π4) $\left(\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\widetilde{\lambda}\right)^{ss} \cong \left(\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}\widetilde{\lambda^w}\right)^{ss}$ for $w \in \mathfrak{W}$ (see [BZ] 2.9),

where we have written $\lambda^w(t) = \lambda(wtw^{-1})$.

On $V^{\mathcal{U}}$, we have the Hecke operator $T(\varpi)$ for a prime element ϖ of $O_{\mathfrak{q}}$ given by $T(\varpi)v = \sum_{\alpha\mathcal{U}\subset\mathcal{U}\xi\mathcal{U}/\mathcal{U}} \pi(\alpha)v$ for $\xi = \left(\begin{smallmatrix} \varpi & 0 \\ 0 & 1 \end{smallmatrix}\right)$, where $\mathcal{U} - \mathcal{U}(O_{\mathfrak{q}})$. We generalize this action of Hecke operators as follows. Write $\mathcal{B} = \mathcal{B}(O_{\mathfrak{q}})$ and $\mathcal{U} = \mathcal{U}(O_{\mathfrak{q}})$ for simplicity. Let

$$\begin{aligned} \mathcal{D} &= \left\{x \in \mathcal{T}(F_{\mathfrak{q}}) \middle| x\mathcal{U}x^{-1} \supset \mathcal{U}\right\} \\ \mathcal{D}^* &= \left\{x \in \mathcal{T}(F_{\mathfrak{q}}) \middle| x\mathcal{U}x^{-1} \subset \mathcal{U}\right\}. \end{aligned} \tag{2.32}$$

The semi-group $\mathcal{D}$ is called the *expanding semi-group* in $\mathcal{T}(F_{\mathfrak{q}})$, and $\mathcal{D}^*$ is called *shrinking semi-group*. The set $\Delta_{\mathcal{U}} = \mathcal{U} \cdot \mathcal{D} \cdot \mathcal{U}$ and $\Delta_{\mathcal{U}}^* = \mathcal{U} \cdot \mathcal{D}^* \cdot \mathcal{U}$ are also multiplicative semi-groups.

Exercise 2.19. *Show that* $\mathcal{D}$ *(resp.* $\mathcal{D}^*$*) is generated by* $\mathcal{T}(O_{\mathfrak{q}})$ *and* $\left(\begin{smallmatrix} \varpi^{e_1} & 0 \\ 0 & \varpi^{e_2} \end{smallmatrix}\right)$ *for integers* $e_1 \leq e_2$ *(resp.* $e_1 \geq e_2$*).*

Define the so-called Iwahori subgroups by

$$I_0(r) = \left\{u \in GL_2(O_{\mathfrak{q}}) \middle| u \mod \mathfrak{q}^r \in \mathcal{B}(O/\mathfrak{q}^r)\right\},$$
$$I_1(r) = \left\{u \in I_0(r) \middle| u \mod \mathfrak{q}^r \in \mathcal{U}(O/\mathfrak{q}^r)\right\}.$$

These subgroups S have the Iwahori decomposition $S = \mathcal{U}'T'\mathcal{U} \cong \mathcal{U}' \times T' \times \mathcal{U}$ for open compact subgroups $T' \subset \mathcal{T}(O_{\mathfrak{q}})$ and $\mathcal{U}'$ in the opposite unipotent ${}^t\mathcal{U}$. Each $x \in \mathcal{D}$ shrinks ${}^t\mathcal{U}$: $x{}^t\mathcal{U}x^{-1} \subset {}^t\mathcal{U}$. From this, $\Delta_S = S \cdot \mathcal{D} \cdot S$ and $\Delta_S^* = S \cdot \mathcal{D}^* \cdot S$ are again multiplicative sub-semigroups of $\mathcal{G}$ (this statement includes $\Delta_{\mathcal{B}} = \mathcal{B} \cdot \mathcal{D} \cdot \mathcal{B} = \Delta_{I_0(\infty)}$ and $\Delta_{\mathcal{B}}^* = \mathcal{B} \cdot \mathcal{D}^* \cdot \mathcal{B}$). We call Δ_S (resp. Δ_S^*) the *expanding* (resp. *shrinking*) semi-group with respect to $(S,\mathcal{U})$. When $G = GL(2)$, Δ_S for $S = I_0(r)$ $(r > 0)$ is almost equal to the $\mathfrak{q}$-part of the image of the semi-group $\Delta_0(\mathfrak{q}^r)$ introduced in (2.11) under the involution ι (strictly speaking, we have $\Delta_0(\mathfrak{q}^r)_p = \Delta_S$ modulo center).

Exercise 2.20. *Prove, as sets,*

$$I_0(r) = {}^tU(\mathfrak{q}O_\mathfrak{q})\mathcal{T}(O_\mathfrak{q})U(O_\mathfrak{q}) = {}^tU(\mathfrak{q}O_\mathfrak{q}) \times \mathcal{T}(O_\mathfrak{q}) \times U(O_\mathfrak{q}),$$

where $U(\mathfrak{a}) = \left\{\left(\begin{smallmatrix}1&u\\0&1\end{smallmatrix}\right) \middle| u \in \mathfrak{a}\right\}$ *for an ideal* $\mathfrak{a}$ *of* $O_\mathfrak{q}$.

By the Iwahori decomposition, we have, for $X = \mathcal{U}, \mathcal{B}$, and an Iwahori subgroup S,

$$\begin{aligned} X\xi X &= \bigsqcup_{u\in\xi^{-1}\mathcal{U}\xi\backslash\mathcal{U}} X\xi u = \bigsqcup_{u\in\mathcal{U}\backslash\xi\mathcal{U}\xi^{-1}} Xu\xi \quad \text{for } \xi\in\mathcal{D} \\ X\xi X &= \bigsqcup_{u\in\xi\mathcal{U}\xi^{-1}\backslash\mathcal{U}} u\xi X = \bigsqcup_{u\in\mathcal{U}\backslash\xi^{-1}\mathcal{U}\xi} \xi uX \quad \text{for } \xi\in\mathcal{D}^*. \end{aligned} \tag{2.33}$$

By this fact, the *double coset ring* (or the Hecke ring; e.g., [?] 3.1) generated additively over $\mathbb{Z}$ by double cosets of X in Δ_X have the following homomorphic relations as algebras:

$$\begin{array}{ccccccccc} \mathcal{R} & = & R(\mathcal{U},\Delta_\mathcal{U}) & \twoheadrightarrow & R(\mathcal{B},\Delta_\mathcal{B}) & \cong & R(S,\Delta_S) & = & \mathcal{R}_\mathcal{B} \\ & & \wr\downarrow & & \wr\downarrow & & \wr\downarrow & & \\ \mathcal{R}^* & = & R(\mathcal{U},\Delta^*_\mathcal{U}) & \twoheadrightarrow & R(\mathcal{B},\Delta^*_\mathcal{B}) & \cong & R(S,\Delta^*_S) & = & \mathcal{R}^*_\mathcal{B} \end{array}$$

via $\mathcal{U}\xi\mathcal{U} \mapsto \mathcal{B}\xi\mathcal{B} \mapsto S\xi S$ for $\xi \in \Delta_\mathcal{U}$ (resp. $\xi \in \Delta^*_\mathcal{U}$) and $\mathcal{R} \ni \mathcal{U}\xi\mathcal{U} \mapsto \mathcal{U}\xi^\iota\mathcal{U} \in \mathcal{R}^*$ for $x^\iota x = \det(x)$. These algebras are commutative: $T(\xi)T(\eta) = T(\xi\eta)$ for $T(\xi) = \mathcal{U}\xi\mathcal{U}$ and $\xi, \eta \in \mathcal{D}$ or $\mathcal{D}^*$. Here is a proof of these facts. Since we only need to deal with $\mathcal{D}$ or $\mathcal{D}^*$, in the following lemma, we prove the result for $\mathcal{D}^*$ (and the result for $\mathcal{D}$ follows by applying the involution "ι").

Lemma 2.21. *Let the notation and the assumption be as above. The algebras* $R(S,\Delta^*_S)$ *for Iwahori subgroups* S *are commutative, and if* $S \supset B(\mathbb{Z}_p)$, *they are all isomorphic to the polynomial ring* $\mathbb{Z}[t_1,t_2,t_2^{-1}]$ *(with* t_2 *inverted) for* $t_1 = S\left(\begin{smallmatrix}\varpi&0\\0&1\end{smallmatrix}\right)S$ *and* $t_2 = \left(\begin{smallmatrix}\varpi&0\\0&\varpi\end{smallmatrix}\right)$. *If* $S \supset \mathcal{U}$, *we have* $R(S,\Delta^*_S) \cong \mathbb{Z}[T_S][t_1,t_2,t_2^{-1}]$ *for the quotient group* $T_S = \mathcal{T}(O_\mathfrak{q})S/S$, *where* $\mathbb{Z}[T_S]$ *is the group algebra of* T_S.

Proof. For $\xi \in \mathcal{D}^*$, we consider the double coset $\mathcal{B}\xi\mathcal{B}$. Decompose

$$\mathcal{B} = \bigsqcup_{\eta\in\Xi(\xi)} \eta(\xi\mathcal{B}\xi^{-1}\cap\mathcal{B}).$$

Multiplying by $\xi\mathcal{B}\xi^{-1}$ from the right, we get $\mathcal{B}\xi\mathcal{B}\xi^{-1} = \bigsqcup_{\eta\in\Xi(\xi)}\eta\xi\mathcal{B}\xi^{-1} \Leftrightarrow \mathcal{B}\xi\mathcal{B} = \bigsqcup_{\eta\in\Xi(\xi)}\eta\xi\mathcal{B}$. If $\xi = \left(\begin{smallmatrix}\varpi^{a_1}&0\\0&\varpi^{a_2}\end{smallmatrix}\right)$, we have

$$\xi\mathcal{B}\xi^{-1}\cap\mathcal{B} = \left\{\left(\begin{smallmatrix}a&b\\0&d\end{smallmatrix}\right)\in\mathcal{B}\middle| b\in\varpi^{a_1-a_2}O_\mathfrak{q}\right\}.$$

We may choose the subset $\Xi(\xi)$ inside $\mathcal{U}$ so that

$$\Xi(\xi) \ni \left(\begin{smallmatrix}1&b\\0&1\end{smallmatrix}\right) \mapsto \left(b \mod \varpi^{a_1-a_2}O_\mathfrak{q}\right) \in O/\varpi^{a_1-a_2}O$$

is a bijection. Then we have $\mathcal{B}\xi\mathcal{B} = \bigsqcup_{\eta\in\Xi(\xi)}\eta\xi\mathcal{B}$ and a formula of the index: $(\mathcal{B} : \xi\mathcal{B}\xi^{-1}\cap\mathcal{B}) = |\varpi^{[\xi]}|_\mathfrak{q}^{-1}$ with $[\xi] = (a_1 - a_2)$. Writing $\deg(\mathcal{B}\xi\mathcal{B})$ for the number of right cosets of $\mathcal{B}$ in $\mathcal{B}\xi\mathcal{B}$, we find $\deg(\mathcal{B}\xi\zeta\mathcal{B}) = \deg(\mathcal{B}\xi\mathcal{B})\deg(\mathcal{B}\zeta\mathcal{B})$, because $[\xi\zeta] = [\xi] + [\zeta]$ for $\xi, \zeta \in \mathcal{D}^*$. Since $\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B} \supset \mathcal{B}\xi\zeta\mathcal{B}$, if we can show that $\deg(\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B}) = \deg(\mathcal{B}\xi\zeta\mathcal{B})$, we get $\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B} = \mathcal{B}\xi\zeta\mathcal{B}$ and $(\mathcal{B}\xi\mathcal{B}) \cdot (\mathcal{B}\zeta\mathcal{B}) = \mathcal{B}\xi\zeta\mathcal{B}$ in the double coset ring $R(\mathcal{B}, \Delta^*_{\mathcal{B}})$, which in particular shows the commutativity of $R(\mathcal{B}, \Delta^*_{\mathcal{B}})$. To see $\deg(\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B}) = \deg(\mathcal{B}\xi\zeta\mathcal{B})$, we note $\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B} = \bigcup_{\eta\in\Xi(\xi)} \bigcup_{\eta'\in\Xi(\zeta)} \eta\xi\eta'\zeta\mathcal{B}$. This implies

$$\deg(\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B}) \le \deg(\mathcal{B}\xi\mathcal{B})\deg(\mathcal{B}\zeta\mathcal{B}) = \deg(\mathcal{B}\xi\zeta\mathcal{B}),$$

and hence we get the identity $\deg(\mathcal{B}\xi\mathcal{B}\zeta\mathcal{B}) = \deg(\mathcal{B}\xi\zeta\mathcal{B})$. Since the α_js give independent generators of $\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}})$, the monoid algebra $\mathbb{Z}[\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}})]$ is isomorphic to a polynomial ring with two variables $\mathbb{Z}[\alpha, \varpi 1_2, (\varpi 1_2)^{-1}]$ with $\alpha = \left(\begin{smallmatrix}\varpi & 0\\ 0 & 1\end{smallmatrix}\right)$ (because $\varpi 1_2 \in Z(O_{\mathfrak{q}})$ is invertible in $\mathcal{D}^*$). The association $\alpha \mapsto T(\varpi) = \mathcal{B}\alpha\mathcal{B}$ and $\varpi 1_2 \mapsto \mathcal{B}\varpi\mathcal{B}$ therefore induces a surjective algebra homomorphism $\mathbb{Z}[\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}})] \to R(\mathcal{B}, \Delta^*_{\mathcal{B}})$, which can be easily seen to be an isomorphism. Replacing $\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}})$ by $\mathcal{D}^*/(\mathcal{T}(O_{\mathfrak{q}}) \cap S) = T_S \times (\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}}))$ in the above argument, the same proof works well for any S with $I_0(r) \supset S \supset I_1(r)$ and yields $R(S, \Delta^*_S) \cong \mathbb{Z}[T_S][\mathcal{D}^*/\mathcal{T}(O_{\mathfrak{q}})] \cong \mathbb{Z}[T_S][t_1, t_2, t_2^{-1}]$, where $\mathbb{Z}[T_S]$ is embedded into $R(S, \Delta^*_S)$ by sending $t \in T_S$ to StS. □

We let $\mathcal{R}^*$ act on $v \in V^{\mathcal{U}} = H^0(\mathcal{U}, V)$ by

$$T(\xi)v = [\mathcal{U}\xi\mathcal{U}]v = \sum_{u\in\xi\mathcal{U}\xi^{-1}\backslash\mathcal{U}} \pi(u\xi)v = \int_{\xi^{-1}\mathcal{U}\xi} \pi(\xi)\pi(u)v\,du, \tag{2.34}$$

and similarly for $\overline{v} \in V_{\mathcal{B}}$ in place of $v \in V^{\mathcal{U}}$; then the projection: $V^{\mathcal{U}} \to V_{\mathcal{B}}$ is $\mathcal{R}^*$-linear. To see the last identity of (2.34), it is sufficient to recall that we have normalized the measure du so that $\int_{\mathcal{U}} du = 1$. We may regard the above action as an action of $\mathcal{R}$ via the isomorphism $\mathcal{R} \cong \mathcal{R}^*$:

$$v|T(\xi) = v|[\mathcal{U}\xi\mathcal{U}] = \sum_{u\in\xi^{-1}\mathcal{U}\xi\backslash\mathcal{U}} \pi((\xi u)^\iota)v = \int_{\xi\mathcal{U}\xi^{-1}} \pi(\xi^\iota)\pi(u^\iota)v\,du. \tag{2.35}$$

For $\alpha = \left(\begin{smallmatrix}\varpi & 0\\ 0 & 1\end{smallmatrix}\right)$, we have $\mathcal{U}(F_{\mathfrak{q}}) = \cup_{j=0}^{\infty} \alpha^{-j}\mathcal{U}\alpha^j$. Thus writing $T(\varpi) = [\mathcal{U}\alpha\mathcal{U}]$ and $T(\xi) = [\mathcal{U}\xi\mathcal{U}]$ for $\xi \in \mathcal{D}^*$ as an operator on $V^{\mathcal{U}} = H^0(\mathcal{U}, V)$, we see easily from (2.34) that $T(\alpha^j) = T(\varpi)^j$ and for each finite-dimensional subspace $X \subset V(\mathcal{B})$, $T(\varpi)|_X$ is nilpotent on $X^{\mathcal{U}}$ by (2.34).

For any $\mathcal{R}$-eigenvector $v \in V^{\mathcal{U}}$ with $t\overline{v} = \lambda(t)\overline{v}$ ($t \in \mathcal{T}(F_{\mathfrak{q}})$, $\overline{v} = v \mod V(\mathcal{B})$), we get

$$v|\left[\mathcal{U}\left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)\mathcal{U}\right] = \left[\mathcal{U} : \left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)\mathcal{U}\left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)^{-1}\right]\lambda\left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)v = |a^{-1}d|_{\mathfrak{q}}\lambda\left(\begin{smallmatrix}a & 0\\ 0 & d\end{smallmatrix}\right)v, \tag{2.36}$$

where $|\ |_{\mathfrak{q}}$ is the standard absolute value of $O_{\mathfrak{q}}$ such that $|\varpi|_{\mathfrak{q}}^{-1} = N(\mathfrak{q})$.

Lemma 2.22. *If $V = V(\pi)$ is admissible, we have a canonical splitting $V^{\mathcal{U}} \cong V_{\mathcal{B}} \oplus V(\mathcal{B})^{\mathcal{U}}$ as Hecke modules, where $V^{\mathcal{U}} = H^0(\mathcal{U}(O_{\mathfrak{q}}), V)$.*

An absolutely irreducible admissible representation π is called *supercuspidal* if $V_{\mathcal{B}} = 0$. In other words, by (π1), an absolutely irreducible supercuspidal representation can never appear in a subquotient of an induced representation $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \lambda$.

Proof. We have by definition, $V^{\mathcal{U}} = V^{\mathcal{U}(\mathbb{Z}_p)} = \bigcup_r V^{I_1(r)}$. The subspace $V_r = V^{S_1(r)}$ is finite-dimensional and stable under $\mathcal{R}$. By the Jordan decomposition applied to $T(\varpi)$, we can decompose uniquely that $V_r = V_r^\circ \oplus V^{nil}$ so that $T(\varpi)$ is an automorphism on V_r° and is nilpotent on V^{nil}. We may replace $T(\varpi)$ by $T(\alpha^a) = T(\varpi)^a$ for any positive a in the definition of the above splitting. Since $T(\varpi)$ is nilpotent over any finite-dimensional subspace of $V(\mathcal{B})$, V_r° injects into $V_{\mathcal{B}}$; so, $\dim V_r^\circ$ is bounded by $\dim V_{\mathcal{B}} \leq |\mathfrak{W}| = 2$. For any $\mathcal{T}$-eigenvector $\overline{v} \in V_{\mathcal{B}}$, lift it to $v \in V$. Then for sufficiently large j, $\pi(\alpha^{-j})v$ is in $V^{\mathcal{U}}$. Since $\pi(\alpha^{-j})\overline{v}$ is a constant multiple of $\overline{v}$, we may replace $\overline{v}$ and v by $\pi(\alpha^{-j})\overline{v}$ and $\pi(\alpha^{-j})v$, respectively. Then for sufficiently large k, $w = T(\alpha^k)v \in V_r^\circ$, and $T(\varpi)^{-k}\overline{w}$ is equal to $\overline{v}$ for the image $\overline{w}$ in $V_{\mathcal{B}}$. This finishes the proof when the action of $\mathcal{T}$ on $V_{\mathcal{B}}$ is semi-simple. In general, take a sufficiently large r so that V_r surjects down to $V_{\mathcal{B}}$. We apply the above argument to the semi-simplification of V_r under the action of the Hecke algebra. Thus $V^\circ = \bigcup_r V_r^\circ \cong V_{\mathcal{B}}$, and this finishes the proof of $V^{\mathcal{U}} = V_{\mathcal{B}} \oplus V(\mathcal{B})^{\mathcal{U}}$ as $\mathcal{R}$-modules. □

Lemma 2.23. *We have*

1. *If an absolutely irreducible admissible representation π of $\mathcal{G}$ is finite dimensional, it is one dimensional and is a character.*
2. *If $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \lambda$ is absolutely reducible, λ_j has to satisfy $\lambda_1/\lambda_2(x) = |x|_{\mathfrak{q}}^{\pm 1}$.*
3. *If $\lambda_1/\lambda_2(x) = |x|_{\mathfrak{q}}^{\pm 1}$, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ is reducible, the length of the composition series of $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ is 2. The infinite dimensional irreducible subquotient in $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ is denoted by $\sigma(\lambda_1, \lambda_2)$ and is called special (or Steinberg) representation of $\mathcal{G}$.*
4. *If $\lambda_1/\lambda_2(x) = |x|_{\mathfrak{q}}^{-1}$, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ contains the subspace on which $\mathcal{G}$ acts by the character $x \mapsto \lambda_1(\det(x))|\det(x)|_{\mathfrak{q}}^{1/2}$, and if $\lambda_1/\lambda_2(x) = |x|_{\mathfrak{q}}$, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}^w$ contains the quotient on which $\mathcal{G}$ acts by the character $x \mapsto \lambda_2(\det(x))|\det(x)|_{\mathfrak{q}}^{1/2}$. If the conductor of λ_1 is given by $\mathfrak{q}^r$, the subspace of the infinite dimensional irreducible subquotient on which $I_0(r)$ acts by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \lambda_1(a)\lambda_2(d)$ is one dimensional.*

Each vector in the subspace of the infinite dimensional subquotient (the Steinberg representation) described by the fourth assertion is called a *minimal* vector, which is uniquely determined up to scalar multiple. Indeed, for any induced representation $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$, we can have two type of minimal vectors v_1 and v_2 on which $I_0(r)$ acts by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) v_1 = \lambda_1(a)\lambda_2(d)v_1$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) v_2 = \lambda_2(a)\lambda_1(d)v_2$ if $\mathfrak{q}^r$ is the smallest ideal among the conductors of λ_1 and λ_2. When λ_1 is unramified for a Steinberg representation $\sigma(\lambda_1, \lambda_2)$, the minimal vector $v_1 \in \sigma(\lambda_1, \lambda_2)$ coincides with the *new* vector. A cusp form in the image of the Jacquet–Langlands correspondence in Theorem 2.12 regarded as an element of the representation space of $GL_2(F_{\mathfrak{q}})$ for $\mathfrak{q} \in \Sigma^D$ is a linear combination of such new vectors. This is another aspect of the word "new" in Theorem 2.12. In particular, a nonzero new vector is not a linear combination of translations by elements of $\mathcal{G}$ of vectors fixed by a maximal compact subgroup of $\mathcal{G}$ (because in this Steinberg representation, there is no vector fixed by any maximal compact subgroup).

Proof. Suppose that the representation space $V(\pi)$ of π is finite dimensional. Let $\alpha = \left(\begin{smallmatrix}\varpi & 0\\ 0 & 1\end{smallmatrix}\right)$ for a prime element ϖ of $O_{\mathfrak{q}}$, and define

$$\mathcal{U}_n = \alpha^n \mathcal{U} \alpha^{-n} = \left\{ \left(\begin{smallmatrix}1 & u\\ 0 & 1\end{smallmatrix}\right) \middle| u \in \varpi^n O_{\mathfrak{q}} \right\}.$$

Since $\{\mathcal{U}_n\}_n$ gives a fundamental system of neighborhoods of $\mathcal{U}(F_{\mathfrak{q}})$, we find a nonzero vector v fixed by $\mathcal{U}_n$ for some $n \in \mathbb{Z}$. Then $\pi(\alpha^{-m})v$ is fixed by $\alpha^{-m}\mathcal{U}_n\alpha^m = \mathcal{U}_{n-m}$. This shows that $H^0(\mathcal{U}_n, V(\pi)) \neq 0$ for all n. Since $\mathcal{U}_n \subset \mathcal{U}_{n-1}$, we have an infinite sequence of nontrivial subspace

$$H^0(\mathcal{U}_0, V(\pi)) \supset H^0(\mathcal{U}_{-1}, V(\pi)) \supset \cdots \supset H^0(\mathcal{U}_{-n}, V(\pi)) \supset \cdots.$$

From $\dim V < \infty$, we conclude $H^0(\mathcal{U}(F_{\mathfrak{q}}), V(\pi)) = \bigcap_n H^0(\mathcal{U}_n, V(\pi)) \neq 0$, which is stable under $\mathcal{B}(F_{\mathfrak{q}})$ because $\mathcal{B}$ normalizes $\mathcal{U}$. Since π is admissible, the stabilizer of $0 \neq v \in H^0(\mathcal{U}(F_{\mathfrak{q}}), V(\pi))$ contains an open subgroup S of $\mathcal{G}$. In particular, the orbit $S(\infty)$ of the infinity under S in the one dimensional projective space $\mathbf{P}^1(F_{\mathfrak{q}}) = F_{\mathfrak{q}} \sqcup \{\infty\}$ is open. Since $\mathcal{U}(F_{\mathfrak{q}})$ acts transitively on $\mathbf{P}^1(F_{\mathfrak{q}}) - \{\infty\} = F_{\mathfrak{q}}$, the subgroup $\mathcal{H}$ of $\mathcal{G}$ generated by $\mathcal{U}(F_{\mathfrak{q}})$ and S acts $\mathbf{P}^1(F_{\mathfrak{q}})$ transitively. In particular, $\mathcal{H}$ contains all conjugates of $\mathcal{U}(F_{\mathfrak{q}})$ and hence all unipotent elements. As is well known, $SL_2(F_{\mathfrak{q}})$ is generated by unipotent elements (e.g., [PAF] Lemma 4.46). Thus π factors through $\mathcal{G}/SL_2(F_{\mathfrak{q}}) \cong F_{\mathfrak{q}}^{\times}$. Then by Schur's lemma, π is one dimensional. This proves (1).

Suppose that $W = \mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ is reducible. Take a proper subspace $V \subset W$ stable under $\mathcal{G}$. Then $V_{\mathcal{B}}$ is a proper subspace of $W_{\mathcal{B}}$, because $V \mapsto V_{\mathcal{B}}$ is exact and (π1). Since $\dim W_{\mathcal{B}} = 2$ by (π4), if $\widetilde{\lambda^w} \neq \widetilde{\lambda} \Leftrightarrow \lambda_1/\lambda_2 \neq |\cdot|_{\mathfrak{q}}^{\pm 1}$ for $w = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ (cf. Exercise 2.18), we may assume that $V_{\mathcal{B}} = V(\lambda)$. Since the action of w preserves V, $V_{\mathcal{B}}$ also has a nonzero $\mathcal{B}$-eigenvector belonging to λ^w; so, $V_{\mathcal{B}} = W_{\mathcal{B}}$, a contradiction. Thus if W is reducible, $\lambda_1/\lambda_2 = |\cdot|_{\mathfrak{q}}^{\pm 1}$.

If $\lambda_1/\lambda_2 = |\cdot|_{\mathfrak{q}}^{-1}$, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}(\widetilde{\lambda})$ contains $x \mapsto \lambda_1(\det(x))|\det(x)|_{\mathfrak{q}}^{1/2}$; so, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}(\widetilde{\lambda})$ is reducible. If $\lambda_1/\lambda_2 = |\cdot|_{\mathfrak{q}}$, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}(\widetilde{\lambda^w})$ contains $x \mapsto \lambda_2(\det(x))|\det(x)|_{\mathfrak{q}}^{1/2}$; so, $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}}(\widetilde{\lambda})$ is reducible by (π4).

As for the last assertion, by an explicit computation, the subspace of minimal vectors in $\mathrm{Ind}_{\mathcal{B}}^{\mathcal{G}} \widetilde{\lambda}$ (on which $I_0(r)$ acts as specified) is two dimensional. Indeed, by Lemma 2.22, the space is isomorphic to $V_{\mathcal{B}}$ under the projection $V \mapsto V_{\mathcal{B}}$, and hence at most two dimensional. One can easily create two linearly independent vectors in the space; so, it is two dimensional. Then by the third assertion, the one dimensional subquotient takes one dimensional subquotient of this two dimensional subspace. The rest gives the desired one dimensional subspace of the infinite dimensional irreducible subquotient. □

All admissible absolutely irreducible representations of $\mathcal{G}$ are classified into four disjoint classes of representations: characters, principal representations, Steinberg representations and supercuspidal representations.

To give a more precise description of the correspondence $\pi_D \mapsto \pi$ in (JL3), we first note a consequence of Proposition 1.2:

Lemma 2.24. *If $D_{\mathfrak{q}}$ is a division algebra, each admissible irreducible representation of $D_{\mathfrak{q}}^{\times}$ over $\mathbb{C}$ is finite dimensional. If it has nonzero vector fixed by the unit group of* *the maximal order, it is isomorphic to a character $\lambda \circ N$ for an admissible character $\lambda : F_{\mathfrak{q}}^{\times} \to \mathbb{C}^{\times}$, where $N : D_{\mathfrak{q}} \to F_{\mathfrak{q}}^{\times}$ is the reduced norm map.*

Proof. By Proposition 1.2, the unique maximal order $R = O_{D_{\mathfrak{q}}}$ of $D_{\mathfrak{q}}$ has a unique maximal two-sided ideal $\mathfrak{m}$ with $R/\mathfrak{m} = O/\mathfrak{q}$. Then $\mathfrak{m}^n$ are all two-sided ideals, and hence $1+\mathfrak{m}^n$ is a normal subgroup of $D_{\mathfrak{q}}^{\times}$. Let $V = V(\pi)$ be an admissible irreducible representation of $D_{\mathfrak{q}}^{\times}$. Since $1+\mathfrak{m}^n$ give a fundamental system of open neighborhoods of the identity of $D_{\mathfrak{q}}^{\times}$, for sufficiently large n, $H^0(1+\mathfrak{m}^n, V) \neq 0$. Since $1+\mathfrak{m}^n$ is normal in $D_{\mathfrak{q}}^{\times}$, $H^0(1+\mathfrak{m}^n, V)$ is stable under $D_{\mathfrak{q}}^{\times}$. Then by the irreducibility of V, $H^0(1+\mathfrak{m}^n, V) = V$. Since V is admissible, $H^0(1+\mathfrak{m}^n, V)$ is finite dimensional.

If $H^0(R^{\times}, V) \neq 0$, by the above argument, we have $H^0(R^{\times}, V) = V$, and hence the action of $D_{\mathfrak{q}}^{\times}$ on V factors through the abelian group $D_{\mathfrak{q}}^{\times}/R^{\times} \cong \mathbb{Z}$. Thus by the irreducibility, V is one dimensional. Since the reduced norm map N has kernel inside $R^{\times}$, the character π factors through $F_{\mathfrak{q}}^{\times}$; so, $\pi = \chi \circ N$. This finishes the proof. □

Secondly, we note that by the strong multiplicity one theorem (e.g., [AAG] Theorems 5.14 and 10.10), for a given quaternion algebra D over F, the representation Π_D of $G^D(\mathbb{A}^{(\infty)})$ on the space generated by $\Pi_D(g)h(x) = h(xg)$ for all $g \in G^D(\mathbb{A}^{(\infty)})$ and all $h \in S_{\kappa}^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ is a direct sum of finitely many irreducible representations π_D with multiplicity one.

Here is a slightly more precise description of the correspondence $\pi_D \mapsto \pi$ in (JL3). Start with an eigenform $0 \neq f \in S_{\kappa}^D(\mathfrak{N}, \varepsilon; \mathbb{C})$ of all Hecke operators in $h_{\kappa}(\mathfrak{N}, \varepsilon; \mathbb{C})$. By the multiplicity one theorem, the representation π_D of $G^D(\mathbb{A}^{(\infty)})$ on the space generated by $\pi_D(g)f(x) = f(xg)$ for all $g \in G^D(\mathbb{A}^{(\infty)})$ is irreducible. We can factor $\pi_D = \otimes_{\mathfrak{q}} \pi_{\mathfrak{q}}(D)$ for local irreducible representations $\pi_{\mathfrak{q}}(D)$ of $D_{\mathfrak{q}}^{\times}$ (for the localization-completion $D_{\mathfrak{q}} = D \otimes_F F_{\mathfrak{q}}$). Thus the space $V(\mathfrak{N}, \varepsilon; \pi_D)$ is the tensor product of the corresponding subspace $V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}}(D))$, and if $\mathfrak{q}^r$ exactly divides $\mathfrak{N}$ and $\mathfrak{q}$ is prime to $d(D)$,

$$V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}}(D)) = \left\{ v \in V(\pi_{\mathfrak{q}}(D)) \middle| u \cdot v = \varepsilon(u)v \text{ for all } u \in I_0(r) \right\}.$$

We decompose $V(\mathfrak{N}, \varepsilon; \pi) = \otimes_{\mathfrak{q}} V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}})$ for subspaces $V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}}) \subset V(\pi_{\mathfrak{q}})$ similarly defined for $\pi = \pi_{M_2(F)}$. If $D_{\mathfrak{q}}$ is a division algebra, our definition of $S_0^D(\mathfrak{N})$ implies $S_0(\mathfrak{N})_{\mathfrak{q}} = O_{D,\mathfrak{q}}^{\times}$; so, $\pi_{\mathfrak{q}}(D) = \lambda \circ N$ as above. Then the associated automorphic representation $\pi = \pi_{M_2(F)}$ of $GL_2(F_{\mathbb{A}}^{(\infty)})$ is given as follows:

$$\pi_{\mathfrak{q}} \cong \begin{cases} \pi_{\mathfrak{q}}(D) & \text{if } D_{\mathfrak{q}} \cong M_2(F_{\mathfrak{q}}) \\ \sigma(\lambda, \lambda|\cdot|_{\mathfrak{q}}^{-1}) & \text{if } D_{\mathfrak{q}} \text{ is division and } \pi_{\mathfrak{q}}(D) \text{ is a character } \lambda \circ N, \end{cases} \tag{2.37}$$

where $N : D_{\mathfrak{q}}^{\times} \to F_{\mathfrak{q}}^{\times}$ is the reduced norm map. Then writing $\mathfrak{N} = d(D)\mathfrak{N}_0$, $d(D)$ is prime to $\mathfrak{N}_0$. Then by Lemma 2.24, $\pi_{\mathfrak{q}}(D) = \lambda \circ N$ for $\mathfrak{q}|d(D)$, and $U(\varpi_{\mathfrak{q}})$ acts on the one dimensional space $V(\pi_{\mathfrak{q}}(D))$ via the multiplication by $\lambda(\varpi_{\mathfrak{q}})$. Since $\pi_{\mathfrak{q}} = \sigma(\lambda, \lambda|\cdot|_{\mathfrak{q}}^{-1})$, $H^0(I_0(1), V(\pi_{\mathfrak{q}}))$ is one dimensional, and the eigenvalue of $U(\varpi_{\mathfrak{q}})$ on this space is again given by $\lambda(\varpi_{\mathfrak{q}})$ as easily computable by the expression of the representation as an induced representation in Lemma 2.23. If $\mathfrak{q}$ is prime to $d(D)$, the representation $\pi_{\mathfrak{q}}$ and $\pi_{\mathfrak{q}}(D)$ are isomorphic, identifying $D_{\mathfrak{q}}$ with $M_2(F_{\mathfrak{q}})$; thus, $V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}}(D)) \cong V(\mathfrak{N}, \varepsilon_{\mathfrak{q}}; \pi_{\mathfrak{q}})$ as Hecke modules. Thus the action of $T(\varpi_{\mathfrak{q}})$ of the corresponding representations match, and hence we have an $h_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$-linear isomorphism $V(\mathfrak{N},\varepsilon;\pi_D)\cong V(\mathfrak{N},\varepsilon;\pi_{M_2(F)})$. The representation π associated to π_D appears as an automorphic representation spanned by $\pi(g)f_0$ for a Hecke eigenform $f_0\in S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ unique up to scalar multiples. In other words, fixing the isomorphism $V(\mathfrak{N},\varepsilon;\pi_D)\cong V(\mathfrak{N},\varepsilon;\pi_{M_2(F)})$, we get an inclusion $S^D_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})=\bigoplus_{\pi_D}V(\mathfrak{N},\varepsilon;\pi_D)\cong\bigoplus_{\pi_{M_2(F)}}V(\mathfrak{N},\varepsilon;\pi_{M_2(F)})\hookrightarrow S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ which is the linear map in Corollary 2.13, and the correspondence $f\mapsto f_0$ gives rise to the embedding $S^D_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})\hookrightarrow S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$. A Hecke eigenform in $S^D_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ or $S_\kappa(\mathfrak{N},\varepsilon;\mathbb{C})$ is called $\mathfrak{q}$-*new* (resp. $\mathfrak{q}$-*minimal*) for a prime $\mathfrak{q}$ of F, if it gives a new vector (resp. a minimal vector) of the local $\mathfrak{q}$-component of the automorphic representation generated by the automorphic form.

2.8. **Galois representations.** For almost all ideals $\mathfrak{a}$ of the Hecke algebra $h=h_\kappa(\mathfrak{N},\varepsilon;W)$, we can associate a modular two-dimensional Galois representation $\rho_{\mathfrak{a}}$. Thus h may be considered to be a deformation ring parameterizing all "modular" deformations of given level and given "Neben" characters. By the techniques invented by Wiles (and Taylor), under suitable assumptions, we can prove that a local ring of h is the universal deformation ring we studied in Fall 2015. First we describe the representation ρ_P for prime ideals P of h. In the following description, we normalize the local Artin symbol $[u,F_{\mathfrak{q}}]$ so that $[\varpi_{\mathfrak{q}},F_{\mathfrak{q}}]$ modulo the inertia subgroup is the arithmetic Frobenius element in $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_{N(\mathfrak{q})})$ for $q>0$ with $(q)=\mathbb{Z}\cap\mathfrak{q}$.

Theorem 2.25. *Suppose $k=\kappa_2-\kappa_1+I\geq I$. Let P be a prime ideal of $h=h_\kappa(\mathfrak{N},\varepsilon;W)$ and write $k(P)$ for the quotient field of h/P. We assume that $k(P)$ has characteristic different from 2. Then we have a semisimple Galois representation $\rho_P:\mathrm{Gal}(\overline{F}/F)\to GL_2(k(P))$ unramified outside $p\mathfrak{N}$ such that*

(1) *ρ_P is continuous with respect to the profinite topology on $k(P)$ induced from the profinite subring h/P.*
(2) *We have $\mathrm{Tr}(\rho_P(Frob_{\mathfrak{l}}))=T(\varpi_{\mathfrak{l}})\mod P$ for all prime ideal $\mathfrak{l}$ prime to $p\mathfrak{N}$ and $\det(\rho_P)=\varepsilon_+\mathcal{N}^{[\kappa]}$ for the p-adic cyclotomic character $\mathcal{N}$, where we regard ε_+ as a Galois character by global class field theory.*
(3) *Let $\mathfrak{m}$ be the unique maximal ideal containing P. Either if $k\geq 2I$ and $T_p(p)\notin\mathfrak{m}$ or if $k=2I$ and $T_p(\varpi_{\mathfrak{p}})\notin\mathfrak{m}$ for a prime factor $\mathfrak{p}$ of p in F, we have $\rho_P|_{D_{\mathfrak{p}}}\cong\left(\begin{smallmatrix}\epsilon_{\mathfrak{p}}&*\\0&\delta_{\mathfrak{p}}\end{smallmatrix}\right)$ for the decomposition subgroup $D_{\mathfrak{p}}$ at $\mathfrak{p}$, and $\delta_{\mathfrak{p}}([y,F_{\mathfrak{p}}])=U_p(y)\mod P$ for the local Artin symbol $[y,F_{\mathfrak{p}}]$; in particular, $\delta_{\mathfrak{p}}([u,F_{\mathfrak{p}}])=\varepsilon_{1,\mathfrak{p}}(u)u^{-\kappa_1}$ for $u\in O_{\mathfrak{p}}^\times$. Here we have written $U_p(y)=[S_0(p^\infty\mathfrak{N})\left(\begin{smallmatrix}y&0\\0&1\end{smallmatrix}\right)S_0(p^\infty\mathfrak{N})]$ for $y\in F_{\mathfrak{p}}^\times$.*
(4) *Write $\mathfrak{N}=\mathfrak{N}_0\mathfrak{c}(\varepsilon^-)$. If the prime to p-part $\mathfrak{N}_0^{(p)}$ of $\mathfrak{N}_0$ is square-free and is prime to $\mathfrak{c}(\varepsilon^-)$, for each prime factor $\mathfrak{l}$ of $\mathfrak{N}_0$ prime to p, we have $\rho_P|_{D_{\mathfrak{l}}}\cong\left(\begin{smallmatrix}\epsilon_{\mathfrak{l}}&*\\0&\delta_{\mathfrak{l}}\end{smallmatrix}\right)$ for the decomposition subgroup $D_{\mathfrak{l}}$ at $\mathfrak{l}$, and $\delta_{\mathfrak{l}}([y,F_{\mathfrak{l}}])=U(y)\mod P$ for the local Artin symbol $[y,F_{\mathfrak{l}}]$ ($y\in F_{\mathfrak{l}}^\times$); in particular, $\delta_{\mathfrak{l}}([u,F_{\mathfrak{l}}])=\varepsilon_{1,\mathfrak{l}}(u)$ for $u\in O_{\mathfrak{l}}^\times$.*

The representation ρ_P can be realized to have values in $GL_2(h/P)$ which is unique up to isomorphisms over h/P if $\rho_{\mathfrak{m}}$ for the maximal ideal $\mathfrak{m}$ containing P is absolutely irreducible.

Now we describe local structure of automorphic representations via local Galois properties.

Proposition 2.26. *Let the notation and the assumption be as in the theorem. Suppose* $P = \mathrm{Ker}(\lambda)$ *for the algebra homomorphism* $\lambda : h_\kappa(\mathfrak{N}, \varepsilon; W) \to \overline{\mathbb{Q}}_p$ *satisfying* $f|T(\varpi_\mathfrak{q}) = \lambda(T(\varpi_\mathfrak{q}))f$ *for all prime ideals* $\mathfrak{q}$ *for a Hecke eigenform* $f \in S_\kappa(\mathfrak{N}, \varepsilon; \overline{\mathbb{Q}})$. *Write* $\pi = \otimes_\mathfrak{q} \pi_\mathfrak{q}$ *for the irreducible representation generated by* f. *Let* $\mathfrak{l} \nmid p$ *be a prime.*

(1) *If* $\pi_\mathfrak{l} \cong \pi(\eta_1, \eta_2)$ *for characters* $\eta_j : F_\mathfrak{l}^\times \to \overline{\mathbb{Q}}^\times$, *then the restriction of* ρ_P *to the decomposition group* $D_\mathfrak{l}$ *is isomorphic to* $\left(\begin{smallmatrix} \eta_2 & 0 \\ 0 & \eta_1 \end{smallmatrix}\right)$, *where we abused notation so that* η_j *is identified with the Galois character* $D_\mathfrak{l} \to \overline{\mathbb{Q}}^\times$ *inducing the local characters* η_j *via local class field theory.*
(2) *If* $\pi_\mathfrak{l} \cong \sigma(\eta, \eta|\cdot|_\mathfrak{l}^{-1})$, *the restriction of* ρ_P *to the decomposition group* $D_\mathfrak{l}$ *is isomorphic to a non-semisimple* $\left(\begin{smallmatrix} \eta\mathcal{N}_\mathfrak{l} & * \\ 0 & \eta \end{smallmatrix}\right)$ *for the cyclotomic character* $\mathcal{N}_\mathfrak{l} : D_\mathfrak{l} \to \mathbb{Z}_p^\times$ *with* $\mathcal{N}_\mathfrak{l}([u, F_\mathfrak{l}]) = |u|_\mathfrak{l}^{-1}$.
(3) *If* $\pi_\mathfrak{l}$ *is supercuspidal and* $\mathfrak{N}_0 = \mathfrak{N}/\mathfrak{c}(\varepsilon^-)$ *is prime to* $\mathfrak{c}(\varepsilon^-)$, *then* $\mathfrak{l}^2|\mathfrak{N}_0$ *and* ρ_P *restricted to* $D_\mathfrak{l}$ *is absolutely irreducible.*

Here is a brief outline of the proof of the above results. When $F = \mathbb{Q}$, all these follows from the theorems of [GME, §4.2.3-7]. First suppose that $k(P)$ has characteristic 0. By Theorem 2.10, the projection $\pi : h \to h/P$ (regarded as a linear form $\pi : h \to h/P$) gives rise to a Hecke eigenform $f \in S_\kappa(\mathfrak{N}, \varepsilon; k(P))$ with $\mathbf{a}_\infty(\varpi_\mathfrak{l}, f) = \pi(T(\varpi_\mathfrak{l}))$ for all primes $\mathfrak{l}$; so, we are in the setting of Proposition 2.26.

Let $\mathfrak{N}_0 = \mathfrak{N}/\mathfrak{c}(\varepsilon^-)$. By looking at Fourier expansion, if a prime factor $\mathfrak{l}$ of $\mathfrak{N}$ satisfies $\mathfrak{l}^2 \nmid \mathfrak{N}_0$ and $\mathfrak{l} \nmid \mathfrak{c}(\varepsilon^-)$, we can show that the Hecke operator $T(\varpi_\mathfrak{l})$ on $S_\kappa(\mathfrak{N}, \varepsilon; \mathbb{C})$ is invertible (e.g., [MFM] Theorem 4.6.17 or [H88] Lemma 12.2). Then by Lemma 2.22, $\pi_\mathfrak{l}$ cannot be supercuspidal, which implies the fact $\mathfrak{l}^2|\mathfrak{N}_0$ when $\pi_\mathfrak{l}$ is supercuspidal (in (3) of Proposition 2.26).

If there exists a quaternion algebra $D_{/F}$ with $\dim \mathfrak{Z}_D = 1$ such that f is in the image of the Jacquet–Langlands correspondence $\iota_{h/P}$, the existence of ρ_P satisfying the conditions in Theorem 2.25 and Proposition 2.26 follows from the work of Carayol [C86b] (see also [68c] in [CPS] and [H81] Theorem 4.12). In particular, if $[F : \mathbb{Q}]$ is odd, we have ρ_P by the work of Carayol. Even if $[F : \mathbb{Q}]$ is even, Blasius-Rogawski [BR] realized the representation ρ_P in the p-adic étale cohomology group on the Shimura variety of a unitary group of dimension 3, and at the same time, R. Taylor [T89] (and [T95]) generalized the method of Wiles in [W] to obtain the representation ρ_P by gluing together Galois representations coming from quaternion algebras D as above, using Wiles' observation (see [T89]) that the image of the Jacquet–Langlands correspondence is p-adically dense in the space of p-adic modular forms (if we vary quaternion algebras D). Once ρ_P is constructed, for the reduced part h^{red} (i.e., $h^{red} = h$ modulo the nilradical), we can find a pseudo-representation $\pi : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to h^{red}$ unramified outside $p\mathfrak{N}$ in the sense of Wiles with $\mathrm{Tr}(\pi(Frob_\mathfrak{l})) = T(\varpi_\mathfrak{l})$ for all $\mathfrak{l}$ outside $p\mathfrak{N}$ (by the gluing technique described in [GME] 4.2.5). Out of the pseudo-representation $\pi \mod P$, we can always create a representation as in (1) and (2). Since Taylor's construction comes from gluing mod p^n representations of Carayol which satisfies (4), it supplies us the assertion (4) of the theorem for all ρ_P. As for the assertion (3) of the theorem, it is verified in [W] and [H89].

It remains to prove the existence of $\rho_{\mathfrak{m}}$ and the assertions (3) and (4) of the theorem for each maximal ideal $\mathfrak{m}$ of h. Take a prime ideal $P \subset \mathfrak{m}$ (so, $k(P)$ has characteristic 0). Then by conjugating ρ_P, we may assume that ρ_P has values in the maximal compact subgroup $GL_2(W_P)$ in $GL_2(k(P))$ (Corollary 1.4), where W_P is the p-adic integer ring of $k(Q)$. Let $\rho'_{\mathfrak{m}} = \rho_P \mod \mathfrak{m}_P$ for the maximal ideal $\mathfrak{m}_P$ of W_P. The representation $\rho'_{\mathfrak{m}}$ satisfies the four conditions for the coefficient ring $k(\mathfrak{m}_P) = W_P/\mathfrak{m}_P$ in place of $k(\mathfrak{m})$. Since $\mathrm{Tr}(\rho_P) = \mathrm{Tr}(\pi) \mod P$, the trace of $\rho'_{\mathfrak{m}}$ has values in $h/\mathfrak{m}$. Since the characteristic of $k(\mathfrak{m})$ is odd, by the theory of pseudo representations, if $\rho'_{\mathfrak{m}}$ is irreducible, the representation $\rho' : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to GL_2(h/\mathfrak{m})$ constructed out of the pseudo representation $\pi \mod \mathfrak{m}$ has the same trace as $\rho'_{\mathfrak{m}}$. Since ρ_P is also constructed by the same pseudo representation π, if we follow the same procedure as described in Proposition **??** to construct ρ_P and ρ', we find that $\rho' = \rho_P \mod \mathfrak{m}_P = \rho'_{\mathfrak{m}}$. Thus in this case, $\rho'_{\mathfrak{m}}$ satisfies the assertions (3) and (4); so, we put $\rho_{\mathfrak{m}} = \rho'_{\mathfrak{m}}$.

If $\rho'_{\mathfrak{m}}$ is reducible, we define $\rho_{\mathfrak{m}}$ by the semi-simplification of $\rho'_{\mathfrak{m}}$. Thus $\rho_{\mathfrak{m}} \cong \epsilon \oplus \delta$ for two global characters ϵ and δ. Suppose that $T(\varpi_{\mathfrak{p}}) \notin \mathfrak{m}$ for all $\mathfrak{p}|p$. Then we may define a character $\overline{\delta}_{\mathfrak{p}} : D_{\mathfrak{p}} \to k(\mathfrak{m})^\times$ by $\overline{\delta}_{\mathfrak{p}}([y, F_{\mathfrak{p}}]) = U_p(y) \mod \mathfrak{m}$ for $y \in F_{\mathfrak{p}}^\times$. Then $\rho'_{\mathfrak{m}}$ for the maximal ideal $\mathfrak{m}_P$ of W_P satisfies the assertions (3) and (4) for $k(\mathfrak{m}_P)$ in place of $k(P)$. Put $\overline{\epsilon}_{\mathfrak{p}} = \det(\rho_P)\delta_{\mathfrak{p}}^{-1} \mod \mathfrak{m}_P$. Since $\det \rho_P$ has values in h/P, $\overline{\epsilon}_{\mathfrak{p}}$ has values in $k(\mathfrak{m})^\times$. In particular, we have $\rho_{\mathfrak{m}}|_{D_{\mathfrak{p}}} \cong \overline{\epsilon}_{\mathfrak{p}} \oplus \overline{\delta}_{\mathfrak{p}}$ for the decomposition subgroup $D_{\mathfrak{p}}$ at $\mathfrak{p}$. By the Brauer-Nesbitt theorem (e.g., [MFG] Corollary 2.8; see also Proposition **??** in the text), we have $\{\overline{\epsilon}_{\mathfrak{p}}, \overline{\delta}_{\mathfrak{p}}\}$ coincides with $\{\epsilon|_{D_{\mathfrak{p}}}, \delta|_{D_{\mathfrak{p}}}\}$ as sets; so, the assertion (3) follows for our choice of $\rho_{\mathfrak{m}}$. The assertion (4) of the theorem can be proven similarly.

The uniqueness under absolute irreducibility follows again from the Brauer-Nesbitt theorem.

We consider the following condition,

$$\text{(sf)} \qquad \mathfrak{N}/\mathfrak{c}(\varepsilon^-) \text{ is square-free and is prime to } \mathfrak{c}(\varepsilon^-).$$

We can generalize the result in the above theorem from a prime ideal to any ideal of a local ring of h under mild assumptions including (sf):

Corollary 2.27. *Suppose* (sf). *Let* $\mathbf{T}$ *be the localization of* $h_\kappa(\mathfrak{N}, \varepsilon; W)$ *at a maximal ideal* $\mathfrak{m}$. *If* $\rho_{\mathfrak{m}}$ *is absolutely irreducible, for any ideal* $\mathfrak{a}$ *of* $\mathbf{T}$ *containing the nilradical of* $\mathbf{T}$, *we have a unique Galois representation* $\rho_{\mathfrak{a}} : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to GL_2(\mathbf{T}/\mathfrak{a})$ *up to isomorphisms over* $\mathbf{T}/\mathfrak{a}$ *such that*

(1) $\rho_{\mathfrak{a}}$ *is continuous with respect to the profinite topology on* $\mathbf{T}/\mathfrak{a}$.
(2) *We have* $\mathrm{Tr}(\rho_{\mathfrak{a}}(Frob_{\mathfrak{l}})) = T(\varpi_{\mathfrak{l}}) \mod \mathfrak{a}$ *for all prime ideal* $\mathfrak{l}$ *prime to* $p\mathfrak{N}$ *and* $\det(\rho_{\mathfrak{a}}) = \varepsilon_+\mathcal{N}^{[\kappa]}$ *for the p-adic cyclotomic character* $\mathcal{N}$, *where we regard* ε_+ *as a Galois character by global class field theory.*

Moreover, if we assume further that $T(\varpi_{\mathfrak{p}}) \notin \mathfrak{m}$ *for all prime factor* $\mathfrak{p}$ *of* p *in* F *and that* $\overline{\epsilon}_{\mathfrak{p}} \neq \overline{\delta}_{\mathfrak{p}}$ *for all* $\mathfrak{p}|p$, *then* $\rho_{\mathfrak{a}}|_{D_{\mathfrak{p}}} \cong \left(\begin{smallmatrix} \epsilon_{\mathfrak{p}} & * \\ 0 & \delta_{\mathfrak{p}} \end{smallmatrix}\right)$ *for the decomposition subgroup* $D_{\mathfrak{p}}$ *at* $\mathfrak{p}$, *and* $\delta_{\mathfrak{p}}([y, F_{\mathfrak{p}}]) = U_p(y) \mod \mathfrak{a}$ *for the local Artin symbol* $[y, F_{\mathfrak{p}}]$; *in particular,* $\overline{\epsilon}_{\mathfrak{p}} \equiv \epsilon_{\mathfrak{p}}$

mod $\mathfrak{m}$ *and* $\overline{\delta}_{\mathfrak{p}} \equiv \delta_{\mathfrak{p}} \mod \mathfrak{m}$. *As for the restriction to* $D_{\mathfrak{l}}$ *for a prime* $\mathfrak{l}$ *outside* p, *if* $\rho_{\mathfrak{m}}|_{D_{\mathfrak{l}}} \cong \begin{pmatrix} \overline{\epsilon}_{\mathfrak{l}} & * \\ 0 & \overline{\delta}_{\mathfrak{l}} \end{pmatrix}$ *with* $\overline{\epsilon}_{\mathfrak{l}} \neq \overline{\delta}_{\mathfrak{l}}$ *for a prime factor* $\mathfrak{l}$ *of* $\mathfrak{N}_0$ *prime to* p, *we have* $\rho_{\mathfrak{a}}|_{D_{\mathfrak{l}}} \cong \begin{pmatrix} \epsilon_{\mathfrak{l}} & * \\ 0 & \delta_{\mathfrak{l}} \end{pmatrix}$ *for the decomposition subgroup* $D_{\mathfrak{l}}$ *at* $\mathfrak{l}$, *and* $\delta_{\mathfrak{l}}([y, F_{\mathfrak{l}}]) = U_p(y) \mod \mathfrak{a}$ *for the local Artin symbol* $[y, F_{\mathfrak{l}}]$; *in particular,* $\delta_{\mathfrak{l}}([u, F_{\mathfrak{l}}]) = \varepsilon_{1,\mathfrak{l}}(u)$ *for* $u \in O_{\mathfrak{l}}^{\times}$.

Proof. Let $\pi : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to h^{red}$ be the pseudo representation as in the above sketch of the proof of Theorem 2.25. Let $\pi_{\mathbb{T}}$ be the projection of π to $\mathbb{T}^{red}$ for the reduced part $\mathbb{T}^{red}$ of $\mathbb{T}$. Let $t(\varpi_{\mathfrak{l}})$ and $u(y)$ for the projection to $\mathbb{T}^{red}$ of the Hecke operator $T(\varpi_{\mathfrak{l}})$ and $U_p(y)$. By the irreducibility of $\rho_{\mathfrak{m}}$, by the method of Wiles, we have a representation $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to GL_2(\mathbb{T}^{red})$ unramified outside $p\mathfrak{N}$ and $\mathrm{Tr}(\rho(Frob_{\mathfrak{l}}) = T(\varpi_{\mathfrak{l}})$ for all prime ideals $\mathfrak{l}$ outside $p\mathfrak{N}$. The uniqueness under absolute irreducibility follows from Chebotarev's density theorem and the theorem of Carayol and Serre. By this uniqueness, we have $\rho_{\mathfrak{m}} = \rho \mod \mathfrak{m}$.

Since $\mathbb{T}^{red} \otimes_{\mathbb{Z}} \mathbb{Q} = \prod_P k(P)$, where P running through prime ideals of $\mathbb{T}$ of residual characteristic P. We have $\rho \cong \bigoplus_P \rho_P$ over $\mathbb{T}^{red} \otimes_{\mathbb{Z}} \mathbb{Q}$. Let $\delta_{\mathfrak{p}} : D_{\mathfrak{p}} \to \mathbb{T}^{red}$ be the character given by $\delta([y, F_{\mathfrak{p}}]) = u(y)$, and define $\epsilon_{\mathfrak{p}} = \det(\rho)\delta_{\mathfrak{p}}^{-1}$. Writing the representation space of ρ as $V(\rho)$, we define $V(\delta_{\mathfrak{p}})$ by the quotient of $V(\rho)$ by $\{x \in V(\rho) | \sigma v = \epsilon(\sigma) v \ \forall \sigma \in D_{\mathfrak{p}}\}$. By (3) of Theorem 2.25, $V(\delta_{\mathfrak{p}}) \otimes_{\mathbb{Z}} \mathbb{Q}$ is free of rank 1 over $\mathbb{T}^{red} \otimes_{\mathbb{Z}} \mathbb{Q}$. Since the semi-simplification of $\rho_{\mathfrak{m}}|_{D_{\mathfrak{p}}}$ is isomorphic to $\overline{\epsilon}_{\mathfrak{m}} \oplus \overline{\delta}_{\mathfrak{p}}$, we find from $\overline{\epsilon}_{\mathfrak{m}} \neq \overline{\delta}_{\mathfrak{p}}$ that $V(\delta_{\mathfrak{p}})$ is free of rank 1 over $\mathbb{T}^{red}$. This shows the second assertion.

The third assertion follows from a similar argument. □

References

Books

[AAG] S. S. Gelbart, *Automorphic Forms on Adele Groups*, Annals of Math. Studies **83**, Princeton University Press, Princeton, NJ, 1975.

[AAQ] M.-F. Vignéras, *Arithmétique des algébres de quaternions*, Lecture Notes in Mathematics **800**, Springer, New York, 1980.

[ACM] G. Shimura, *Abelian Varieties with Complex Multiplication and Modular Functions*, Princeton University Press, Princeton, NJ, 1998.

[BCG] R. P. Langlands, *Base Change for $GL(2)$*, Annals of Math. Studies **96**, Princeton University Press, 1980

[BNT] A. Weil, *Basic Number Theory*, Springer, New York, 1974.

[CPS] G. Shimura, *Collected Papers*, I, II, III, IV, Springer, New York, 2002.

[GME] H. Hida, *Geometric Modular Forms and Elliptic Curves*, 2000, World Scientific Publishing Co., Singapore (a list of errata posted at www.math.ucla.edu/˜hida)

[HMI] H. Hida, *Hilbert Modular Forms and Iwasawa Theory*, Oxford University Press, 2006 (a list of errata posted at `www.math.ucla.edu/~hida`).

[IAT] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press and Iwanami Shoten, 1971, Princeton-Tokyo

[ICF] L. C. Washington, *Introduction to Cyclotomic Fields*, Graduate Text in Mathematics, **83**, Springer, 1980

[LFE] H. Hida, *Elementary Theory of L–functions and Eisenstein Series*, LMSST **26**, Cambridge University Press, Cambridge, 1993

[MFG] H. Hida, *Modular Forms and Galois Cohomology*, Cambridge Studies in Advanced Mathematics **69**, 2000, Cambridge University Press

[MFM] T. Miyake, *Modular Forms*, Springer, New York-Tokyo, 1989.

[PAF] H. Hida, p-Adic Automorphic Forms on Shimura Varieties, Springer Monographs in Mathematics, 2004

Articles

[B] D. Blasius, Elliptic curves, Hilbert modular forms, and the Hodge conjecture, in *Contributions to Automorphic Forms, Geometry, and Number Theory*, pp.83–103, a supplement of the American Journal of Mathematics, Johns Hopkins University Press, Baltimore, MD, 2004.

[BCDT] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$ or Wild 3-adic exercises, Journal AMS **14** (2001), 843–939.

[BDST] K. Buzzard, M. Dickinson, N. Shepherd-Barron, and R. Taylor, On icosahedral Artin representations, Duke Math. J. **109** (2001), 283–318.

[BL] J.-L. Brylinski and J.-P. Labesse, Cohomologie d'intersection et fonctions L de certaines variétés de Shimura, Ann. Sci. École Norm. Sup. (4) **17** (1984), 361–412.

[BR] D. Blasius and J. D. Rogawski, Galois representations for Hilbert modular forms, Bull. Amer. Math. Soc. (New Series) **21** (1989), 65–69.

[BZ] I. N. Bernstein and A. V. Zelevinsky, Induced representations of reductive $\mathfrak{p}$–adic groups, I, Ann. Sci. Ec. Norm. Sup. 4th series **10** (1977), 441–472.

[C86a] H. Carayol, Sur la mauvaise réduction des courbes de Shimura, Compositio Math. **59** (1986), 151–230.

[C86b] H. Carayol, Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert, Ann. Sci. Ec. Norm. Sup. 4-th series, **19** (1986), 409–468.

[D] W. Duke, Continued Fractions and Modular Functions, Bulletin AMS **42** (2005), 137–162.

[E] M. Eichler, The basis problem for modular forms and the traces of the Hecke operators, in "Modular functions of one variable, I" Lecture Notes in Math., **320** (1973), 75–151.

[H81] H. Hida, On abelian varieties with complex multiplication as factors of the Jacobians of Shimura curves (Doctor's Thesis at Kyoto University, 1980), Amer. J. Math. **103** (1981), 727–776.

[H86] H. Hida, Galois representations into $GL_2(\mathbb{Z}_p[[X]])$ attached to ordinary cusp forms, Inventiones Math. **85** (1986), 545–613.

[H88] H. Hida, On p-adic Hecke algebras for GL_2 over totally real fields, Ann. of Math. **128** (1988), 295–384.

[H89] H. Hida, Nearly ordinary Hecke algebras and Galois representations of several variables, Proc. JAMI Inaugural Conference, Supplement to Amer. J. Math. (1989), 115–134.

[Hl] D. Hilbert, Mathematical problems, Bull. Amer. Math. Soc. **37** (2000), 407–436.

[Hz] A. Hurwitz, Über die Zhalentheorie der Quaternionen, Göttingen Nachr. Akad. Wiss. 1896, 313–340 (Werke II, No. LXIV).

[KW] C. Khare and J.-P. Wintenberger, Serre's modularity conjecture. I, II. I: Invent. Math. **178** (2009), 485–504; II. Invent. Math. **178** (2009), 505–586.

[K] M. Kisin, Moduli of finite flat group schemes and modularity, Annals of Math. **170** (2009), 1085–1180

[M] B. Mazur, Deforming Galois representations, in "Galois group over $\mathbb{Q}$", MSRI publications **16**, (1989), 385–437.

[R] A. S. Rapinchuk, The congruence subgroup problem, Contemporary Math. **243** (1999), 175–188.

[SeT] J.-P. Serre and J. Tate, Good reduction of abelian varieties, Ann. of Math. **88** (1968), 452–517 (Serre's Œuvres II, , 472–497, No. 79).

[Se] J.-P. Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Duke Math. J. **54** (1987), 179–230 (Œuvres IV, 107–158, No. 143).

[Sh] G. Shimura, The representation of integers as sums of squares, Amer. J. Math. **124** (2002), 1059–1081.

[SW00] C. Skinner and A. Wiles, Residually reducible representations and modular forms, Inst. Hautes Études Sci. Publ. Math. No. **89** (2000), 5–126.

[SW01] C. Skinner and A. Wiles, Nearly ordinary deformations of irreducible residual representations, Ann. Fac. Sc. Toulouse Math. **10** (2001), 185–215.

[T89] R. Taylor, On Galois representations associated to Hilbert modular forms, Inventiones Math. **98** (1989), 265–280.

[T95] R. Taylor, On Galois representations associated to Hilbert modular forms II, in Series in Number Theory **1** (1995): "Elliptic curves, Modular forms, & Fermat's last theorem", pp.185–191.

[T02] R. Taylor, Remarks on a conjecture of Fontaine and Mazur, Journal of the Inst. of Math. Jussieu **1** (2002), 125–143.

[T03] R. Taylor, On icosahedral Artin representations II, Amer. J. Math. **125** (2003), 549–566.

[T04] R. Taylor, Galois representations. Ann. Fac. Sci. Toulouse Math. (6) **13** (2004), 73–119.

[W] A. Wiles, On ordinary Λ-adic representations associated to modular forms, Inventiones Math. **94** (1988), 529–573.