# AUTOMORPHIC FORMS ON $GL(2)$ AND ITS INNER FORMS

HARUZO HIDA

CONTENTS

## 1. INTRODUCTION

We know traditionally from the time of Gauss and Eisenstein that modular forms on a congruence subgroup $\Gamma$ of $SL_2(\mathbb{Z})$ contain an amazing amount of arithmetic information. Easiest way of constructing an elliptic modular form is to make an averaging sum of its automorphic factors: an Eisenstein series. There is another explicit way of constructing modular forms. As an application of Poisson's summation formula, an infinite series attached to each quadratic form $Q(x)$ on a $\mathbb{Q}$–vector space (of dimension $m$) has been used to construct elliptic modular forms explicitly: a theta series. Since theta series is defined by

$$\theta(z) = \sum_{x \in \mathbb{Z}^m} \exp(2\pi i Q(x) z),$$

one might be able to count the number of integer solutions of $Q(x) = n$ for a given positive integer $n$ just by studying the theta series, which is a modular form of weight $\frac{m}{2}$. This is the case for the sum of squares $Q(x) = \sum_{j=1}^{m} x_j^2$ for $2 \le m \le 8$, because one can explicitely write $\theta$ down as a constant multiple of Eisenstein series and Fourier coefficients of an Eisenstein series can be computed explicitly. The idea of relating theta series and Eisenstein series to find such formula is classical going back to the days of Gauss and Jacobi and has been developed much by Siegel, Weil and Shimura more recently.

Following a recent work of Shimura [Sh], we shall give examples of the formula for the sums of squares ($2 \le m \le 8$). Write $S_m(n)$ for the number of representations of an integer $n > 0$ as sums of $m$ squares. Assuming for simplicity $n$ to be odd square-free, we have, for the quadratic residue symbol $\left(\frac{q}{p}\right)$ (primitive with respect to $q$),

(2) $S_2(n) = 2\left(1 + \left(\frac{-1}{n}\right)\right) \sum_{0<d|n} \left(\frac{-1}{d}\right)$ (Lagrange, Gauss, Jacobi);

(3) $S_3(n) = 24\pi^{-1}\sqrt{n} L(1, \left(\frac{-n}{\cdot}\right))$ (Gauss, Dirichlet, Shimura);

(4) $S_4(n) = 8 \sum_{0<d|n} d$ (Jacobi);

(5) $S_5(n) = 2^7(2\pi)^{-2}(\sqrt{n})^3 b_5(n) L(2, \left(\frac{n}{\cdot}\right))$ (Eisenstein, Smith, Minkowski).
Here $b_5(n) = 5$ if $n \equiv 3 \mod 4$, and $b_5(n) = 2^{-3}\cdot 3\cdot 5, 2^{-3}\cdot 5\cdot 7$ according as $n \equiv 1, 5 \mod 8$;

(6) $S_6(n) = (\left(\frac{-1}{n}\right) 2^4 - 4) \sum_{0<d|n} \left(\frac{-1}{d}\right) d^2$ (Jacobi);

(7) $S_7(n) = 2^9(2\pi)^{-3}(\sqrt{n})^3 b_7(n) L(3, \left(\frac{-n}{\cdot}\right))$ (Shimura);
Here $b_7(n) = 7$ if $n \equiv 1 \mod 4$, and $b_7(n) = 2^{-5}\cdot 3^2\cdot 5\cdot 7, 2^{-5}\cdot 7\cdot 37$ according as $n \equiv 3, 7 \mod 8$.

(8) $S_8(n) = 16 \sum_{0<d|n} d^3$ (Jacobi, Siegel).

One can find a slightly more complicated formula valid for all $n$ in [Sh] 3.9. When $m > 8$, there is a contribution of cusp forms; so, we cannot have such a definite formula. The contribution of cusp forms is quite subtle even when $m = 4$, which we study in this course for norm forms (of four variables) of quaternion algebras. If we start with the Hamilton quaternion algebra $H = \mathbb{Q} + \mathbb{Q}i + \mathbb{Q}j + \mathbb{Q}k$ with $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$ and $ki = -ik = j$, the norm form is exactly the sum of four squares: $N(x) = \overline{x}x = x_1^2 + x_2^2 + x_3^2 + x_4^2$ for $x = x_1 + x_2 i + x_3 j + x_4 k \in H$ and $\overline{x} = x_1 - x_2 i - x_3 j - x_4 k$.

To get the formula of $S_2(n)$, a key point is that the ring of Gaussian integers $\mathbb{Z}[\sqrt{-1}]$ is a PID and has four units $\{\pm 1, \pm\sqrt{-1}\}$. For an odd prime $p$,

$$\begin{gathered} p = x_1^2 + x_2^2 \text{ with } x_1, x_2 \in \mathbb{Z} \iff \\ p = \alpha\overline{\alpha} \text{ for } \alpha \in \mathbb{Z}[\sqrt{-1}] \iff \\ \left(\frac{-1}{p}\right) = 1 \ (\Leftrightarrow p \equiv 1 \mod 4). \end{gathered}$$

Thus $S_2(p) = 4, 0$ according as $p \equiv 1 \mod 4$ or not.

As for $S_4(p)$, we need to look into the order $R = \mathbb{Z} + \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}k \subset H$ and study right ideals of $R$. Again all right ideals $\mathfrak{a}$ of $R$ is principal: $\mathfrak{a} = \alpha R$ for $\alpha \in R$. Since the quaternion conjugation $x \mapsto \bar{x}$ turns right ideals into left ideals, we find $\bar{\mathfrak{a}}\mathfrak{a} = R\bar{\alpha}\alpha R$, which is a two-sided ideal generated by $N(\alpha) = \bar{\alpha}\alpha \in \mathbb{Z}$. Thus basically $S_4(p)/8$ is the number of such factorizations $p = \bar{\alpha}\alpha$, because $R$ has 8 units: $\{\pm 1, \pm i, \pm j, \pm k\}$.

We can think of another quaternion algebra $B = M_2(\mathbb{Q})$. Then a maximal order is given by $M_2(\mathbb{Z})$. The unit group of $M_2(\mathbb{Z})$ is infinite and given by $GL_2(\mathbb{Z}) = SL_2(\mathbb{Z}) \sqcup SL_2(\mathbb{Z})\varepsilon$ for $\varepsilon = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Again all right ideals of $M_2(\mathbb{Z})$ is principal (Exercise 1). We define the norm form of $M_2(\mathbb{Q})$ to be $N(x) = \det(x)$. We also have an $M_2(\mathbb{Q})$–conjugation given by $\iota : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Then $N(x) = x^\iota x$. We see easily that up to units of $M_2(\mathbb{Z})$, we have $1 + p$ elements $\alpha$ in $M_2(\mathbb{Z})$ with $N(\alpha) = p$:

$$\left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right) \text{ and } \left(\begin{smallmatrix} p & u \\ 0 & 1 \end{smallmatrix}\right) \text{ for } u = 1, \ldots, p \right\}. \tag{1.1}$$

Thus we conclude $S_4(p)/8$ gives the number of solutions $\det(\alpha) = p$ in $M_2(\mathbb{Z})$ up to units. This is the simplest example of intricate relations between different quaternion algebras, which we study in details in the course.

To motivate our study of quaternionic automorphic forms, let us continue to give examples of deep theorems whose proof rely on essentially elliptic modular forms and quaternion algebras. If one wants to solve a degree five non-soluble rational equation, what we need is a few elliptic functions in addition to classical operation of taking radicals of rationals (a proud result of F. Klein), and the solution is given as the coordinate of a 5–torsion point on a rational elliptic curve (without complex multiplication).

If one wants to find explicit generators (behaving nicely under Galois action) of an abelian extension of the rational number field $\mathbb{Q}$, we only need the exponential function $z \mapsto \mathbf{e}(z) = \exp(2\pi i z)$, which is uniformizing the multiplicative group $\mathbb{G}_m$ ($\mathbf{e} : \mathbb{C} \to \mathbb{G}_m(\mathbb{C}) = \mathbb{C}^\times$ is the universal covering). The generators are roots of unity $\{\mathbf{e}(\frac{1}{N})\}_{0 \neq N \in \mathbb{Z}}$ (a theorem of Kronecker-Weber and Hilbert).

If one wants to generalize this to abelian extensions of an imaginary quadratic field $K$, one need to consider (all) torsion points of an elliptic curve $E$ with complex multiplication by $K$. Thus the desired generator is given again by an elliptic function. This is a famous "Kronecker's dream of his youth" and the origin of Hilbert's twelfth problem.

Since modular functions (that is, modular forms of weight 0) $f : \mathfrak{H} \to \mathbb{C}$ on a subgroup $\Gamma$ of $SL_2(\mathbb{Z})$ can be considered as classifying functions of "all" elliptic curves with some extra structures (for example, a point on the curve of a given order $N$), because over $\mathbb{C}$, any elliptic curve $E$ can be uniformized as $E(\mathbb{C}) = \mathbb{C}/\mathbb{Z}z + \mathbb{Z}$ for a point $z \in \mathfrak{H} = \{z \in \mathbb{C} | i(\bar{z} - z) > 0\}$. Thus all information we get as above can be formulated more naturally using elliptic modular forms and functions (this is the point of view actually taken by F. Klein). Among elliptic modular forms, those forms $f$ which are eigenforms of all Hecke operators $T(n)$ are most important. As was shown by Hecke and Shimura, the eigenvalues $a_n$ of $T(n)$: $f|T(n) = a_n f$ generate a number field $\mathbb{Q}(f)$ (that is a finite extension of $\mathbb{Q}$ called a *Hecke field*). When $\mathbb{Q}(f) = \mathbb{Q}$, we call $f$ a *rational Hecke eigenform*.

One of the most spectacular achievements in the recent history of Number theory is the proof of the Shimura-Taniyama conjecture by Wiles and Taylor et al. This could be (rather in an over-simplified way) formulated as follows. Starting from a rational Hecke eigenform $f$ of weight 2 on the congruence subgroup $\Gamma_0(N)$ of $SL_2(\mathbb{Z})$, Eichler (one example) and Shimura (in general) in the 1950's created a rational elliptic curve $E_{f/\mathbb{Q}}$ whose

$L$–function $L(s,E_f)$ is identical to $L(s,f)$ (so $L(s,E_f)$ has analytic continuation to whole $s$–plane, proving the conjecture of Hasse-Weil for this particular $E_f$). If we use the classical definition of $L$–functions of elliptic curve, this could be formulated as $1+p-a_p=|E(\mathbb{F}_p)|$ as long as

(U1) the equation of the curve modulo $p$ gives an elliptic curve over the finite field $\mathbb{F}_p$.

If we use a slightly modern formulation, $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts naturally and continuously on the étale cohomology group $H^1(E_{f/\overline{\mathbb{Q}}},\mathbb{Z}_\ell)\cong\mathbb{Z}_\ell^2$ (for any prime $\ell$), and the Galois action is characterized so that $\mathrm{Tr}(Frob_p)=a_p$ for almost all primes $p\neq\ell$ (independently of $\ell$ different from $p$), where $Frob_p$ is the Frobenius element of $p$ in the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Thus the Galois action on $H^1(E_{f/\overline{\mathbb{Q}}},\mathbb{Z}_\ell)$ gives a family of Galois representations $\{\rho_\ell\}_\ell$ indexed by primes $\ell$ with independent trace $\mathrm{Tr}(\rho_\ell(Frob_p))=a_p\in\mathbb{Z}$ as long as

(U2) the image of the inertia group at $p$ under $\rho_\ell$ is trivial ($\rho_\ell$ is called unramified at $p$ in this case).

The condition (U2) is actually a consequence of (U1) (a result of Hasse-Deuring and Shimura) and (U2) implies (U1) (a later result of Serre-Tate).

The conjecture then states that any rational elliptic curve $E$ is isogenous over $\mathbb{Q}$ to $E_f$ for a suitably chosen rational Hecke eigenform $f$. An isogeny is a morphism of schemes: $E\to E_f$ which is surjective and having finite kernel. The $L$–function is an isogeny invariant.

In the split of Shimura and Langlands, we may generalize this conjecture to general compatible family $\{\rho_\lambda\}_\lambda$ of Galois representations. Here $\lambda$ runs over all prime ideals (or henselian (non-archimedan) places) of a number field $E$, and $\rho_\lambda:\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})\to GL_2(E_\lambda)$ with $\mathrm{Tr}(\rho(Frob_p))\in E$ is independent of $\ell$ as long as $\rho_\ell$ is unramified at $p$. Such a family can be created for any given Hecke eigenform $f$ so that $\mathrm{Tr}(\rho(Frob_p))=a_p$ (so $E=\mathbb{Q}(f)$): This is a result of Shimura when the weight $k$ is equal to 2, of Deligne if $k>2$ (although Shimura also obtained a little weaker form of Deligne's result (of 1969) for more general automorphic forms: see [68c] in [CPS]) and of Deligne-Serre for $k=1$. Thus if $\det(\rho_\ell)/N_\ell^{k-1}$ is of finite order for the $\ell$–adic cyclotomic character, one expects to have a Hecke eigenform $f$ of weight $k$ giving rise to the compatible family $\{\rho_\lambda\}_\lambda$. This generalized form of the conjecture is also known if $a_p$ is a $p$–adic unit and $k\geq 2$ (under a mild assumption) by a result of Skinner-Wiles, and some cases of $k=1$ has been successfully attacked by Langlands and R. Taylor et al.

So it looks sufficient only to study elliptic modular forms. This is not the case for a general base field $F$. We can consider an arbitrary base field $F$ and a compatible family $\rho=\{\rho_\lambda\}_\lambda$ of representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$. We can formulate the conjecture that there should exist a Hecke eigen automorphic form $f:GL_2(F)\backslash GL_2(F_{\mathbb{A}})\to\mathbb{C}$ giving rise to the family, because we can naturally lift each elliptic Hecke eigenform to an adelic Hecke eigenform on $GL_2(\mathbb{A})$. This direction of the conjecture has been also proven when $F$ is totally real, by K. Fujiwara and Skinner-Wiles (under mild assumptions). The "direction" is to find a modular form on $GL_2(F_{\mathbb{A}})$ out of a given (arithmetic family of) Galois representation.

However, as Shimura noticed in the 1960's, there is no way to create Galois representation out of a Hecke eigenform on the split $GL(2)$. If $F$ is not totally real, the modular variety $GL_2(F)\backslash GL_2(F_{\mathbb{A}})$ is just a differential manifold (not an algebraic variety); so, there is no way to have subtle arithmetic on the manifold to create Galois representations. Even if $F$ is totally real, the Hilbert modular variety does not yield a desired 2–dimensional Galois representations (as can be checked for real quadratic cases). *Creating Galois representation (or even creating an elliptic curve from a given Hilbert modular rational Hecke*

*eigenform of weight* 2*) could be more difficult than finding modular forms out of arithmetic Galois representations or elliptic curves.*

The only known systematic way of creating algebro-geometric object (see, for example, [67b] in [CPS] and [H81]) out of an automorphic form is to study Shimura curves obtained from quaternion algebras over a totally real field $F$ whose automorphic manifold is an algebraic curve defined canonically over $F$, although the above question is still open in general. The utility of such quaternion algebras was first noticed and studied by Shimura. They are not only useful in creating out of quaternionic Hecke eigenforms elliptic curves defined over $F$ (in the rational weight 2 case: [H81]) and families of Galois representations (cf. [68c] in [CPS] and [LFE] Chapter 7) but also in solving (cyclotomic and anticyclotomic) Hilbert's twelfth problem for CM fields ([67b] in [CPS]), using quaternionic automorphic functions. There are some other significant applications of quaternionic automorphic forms (for example, the geometric proof of local Langlands' conjecture for $GL(2)$ and the study of $p$–adic property of Hilbert modular forms).

If we start with a quaternionic Hecke eigen automorphic form $f_B$ for a quaternion algebra $B_{/F}$, we have the associated family $\rho$ of Galois representations by the result of Shimura. Then by Fujiwara or Skinner-Wiles, we find a Hilbert modular form $f$ having the same eigenvalue as $f_B$. This suggests a natural question if *the Hecke eigenvalues of each quaternionic automorphic form would be realized by a Hilbert modular form*. In other words, as Langlands pointed out, the non-abelian reciprocity law in rough form (if exists) depends only on the $\overline{\mathbb{Q}}$–points of the starting algebraic group defined over $F$ (not on its $F$–form). A genesis of this question (formulated as an after-thought of many fundmental works) can be found in a problem Eichler studied in the 1950's (Eichler's basis problem, which he conceived with no definite knowledge of the non-abelian reciprocity law). Here is how Eichler reached this question (out of my guess): As Gauss and Jacobi knew, positive definite quadratic forms $Q(x)$ of four variable with coefficients in $\mathbb{Q}$ give rise to modular forms of weight 2:

$$\theta = \sum_{n \in \mathbb{Z}^4} \mathbf{e}(Q(n)).$$

Eichler studied the norm form of an ideal $\mathfrak{a}$ of a definite quaternion algebra $B$ over $\mathbb{Q}$ and asked what subspace of elliptic modular forms can be spanned by such theta series $\theta(\mathfrak{a})$, and more generally, he asked himself to find a natural base of the space. His result in a special case is as follows: Suppose that $B_p = B \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong M_2(\mathbb{Q}_p)$ for all but one prime $\ell$. Take a maximal order $R$ of $B$ with $R \otimes_{\mathbb{Z}} \mathbb{Z}_p = M_2(\mathbb{Z}_p)$ for $p \neq \ell$. In this case, the automorphic variety $B^\times \backslash B_{\mathbb{A}}^\times / \widehat{R}^\times B_\infty^\times$ for $B_\infty = B \otimes_{\mathbb{Q}} \mathbb{R}$ and $B_{\mathbb{A}} = B \otimes_{\mathbb{Q}} \mathbb{A}$ is zero-dimensional; so, it is a set in bijection to the $R$–right ideal classes: $\{$right $R$–ideals$\}$ modulo left multiplication by $B^\times$. Take a Hecke eigenform $f : B^\times \backslash B_{\mathbb{A}}^\times / \widehat{R}^\times \to \mathbb{C}$ with eigenvalue $a_p$ for $T(p)$, and form $\theta(f) = \sum_{\mathfrak{a}} f(\mathfrak{a})\theta(\mathfrak{a})$. Then $\{\theta(f)\}_f$ gives a base of $S_2(\Gamma_0(\ell))$ and $\theta(f)|T(p) = a_p\theta(f)$, as expected.

This fact has been proven in a far more general setting by Jacquet and Langlands for all quaternion algebras $B$ over any number field $F$, and any Hecke eigenform on $f : B^\times \backslash B_{\mathbb{A}}^\times \to \mathbb{C}$ gives $\theta(f) : GL_2(F) \backslash GL_2(F_{\mathbb{A}}) \to \mathbb{C}$ with the same Hecke eigenvalues. This association is now called the *Jacquet-Langlands correspondence* (or Jacquet-Langlands-Shimizu correspondence including Shimizu who studied the basis problem over totally real fields $F$). We shall study this basis problem of Eichler in a very special case of $B_{/\mathbb{Q}}$ and try to give a sketch of a proof of the correspondence (*integral* over $\mathbb{Z}$).

My hope is that after going through this introductory account of quaternionic automorphic forms over $\mathbb{Q}$, students can start his/her own learning of more general cases of such

automorphic forms which have proven record of utility in establising important arithmetic facts over general base fields $F$.

We will spend first four to five weeks describing arithmetic of quaternion algebras and automorphic forms on them. Later part will be devoted to the sketch of the proof. Main reference is Miyake's book [MFM] Chapters 5 and 6.

**Exercises**

1. Prove that all right ideals of $M_2(\mathbb{Z})$ with an element of non-zero determinant is principal.
2. Prove that the $p+1$ elements in (1.1) gives all solutions of $\det(\alpha) = p$ in $M_2(\mathbb{Z})$ up to right multiplication by units in $M_2(\mathbb{Z})$.

## REFERENCES

**Books**

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

[BNT] A. Weil, *Basic Number theory*, Springer, 1974

[CPS] G. Shimura, *Collected Papers*, I-IV, Springer, 2002

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

[MFM] T. Miyake, *Modular Forms*, Springer, 1989

[NTH] A. Weil, *Number Theory, An approach through history*, Birkhäuser, 1984

**Articles**

[H81] H. Hida, On abelian varieties with complex multiplication as factors of the jacobians of Shimura curves, Amer. J. Math. **103** (1981), 727–776

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

[S] J.-P. Serre, Le probleme des groups de congruence pour $SL_2$, Ann. of Math. **92** (1970), 487–527

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