

# FROM SUM OF TWO SQUARES TO ARITHMETIC SIEGEL–WEIL FORMULAS

CHAO LI

*In loving memory of my mother Xiaoping Mao (1965–2022).*

ABSTRACT. The main goal of this expository article is to survey recent progress on the arithmetic Siegel–Weil formula and its applications. We begin with the classical sum of two squares problem and put it in the context of the Siegel–Weil formula. We then motivate the geometric and arithmetic Siegel–Weil formula using the classical example of the product of modular curves. After explaining the recent result on the arithmetic Siegel–Weil formula for Shimura varieties of arbitrary dimension, we discuss some aspects of the proof and its application to the arithmetic inner product formula and the Beilinson–Bloch conjecture. Rather than intended to be a complete survey of this vast field, this article focuses more on examples and background to provide easier access to several recent works by the author with W. Zhang [LZ22a, LZ22b] and Y. Liu [LL21, LL22].

## CONTENTS

1. Sum of two squares	1
2. Siegel–Weil formula	5
3. Geometric Siegel–Weil formula	11
4. Arithmetic Siegel–Weil formula	18
5. Local arithmetic Siegel–Weil formula	22
6. Arithmetic inner product formula	28
References	35

## 1. SUM OF TWO SQUARES

1.1. **Which prime  $p$  can be written as the sum of two squares?** For the first few primes we easily find that

$$5 = 1^2 + 2^2, \quad 13 = 2^2 + 3^2, \quad 17 = 1^2 + 4^2, \quad 29 = 2^2 + 5^2$$

are sum of two squares, while other primes like 3, 7, 11, 19, 23 are not. The answer seems to depend on the residue class of  $p$  modulo 4.

**Theorem 1.1.1.** *A prime  $p \neq 2$  is the sum of two squares if and only if  $p \equiv 1 \pmod{4}$ .*

Theorem 1.1.1 is usually attributed to Fermat, and appeared in his letter to Mersenne dated Dec 25, 1640 (hence the name *Fermat’s Christmas Theorem*), although the statement can already be found in the work of Girard in 1625. The “only if” direction is obvious, but the “if” direction is

---

*Date:* January 11, 2023.

*2010 Mathematics Subject Classification.* 11G18, 11G40 (primary), 11E25, 11F27, 14C25 (secondary).

The author would like to thank Benedict Gross, Yifeng Liu, Murilo Zanarella, Wei Zhang and the anonymous referee for helpful comments. The author’s work is partially supported by the NSF grant DMS-2101157.

far from trivial. Fermat claimed that he had an irrefutable proof, but nobody was able to find the complete proof among his work — apparently the margin was often too narrow for Fermat. The only clue (in his letters to Pascal and to Digby) is that he used a “descent argument”: if such a prime  $p$  is not of the required form, then one can construct another smaller prime and so on, until a contradiction occurs when one encounters 5, the smallest such prime. More than 100 years later, Euler (1755) gave the first rigorous proof of Theorem 1.1.1 based on infinite descent. For a detailed history of Theorem 1.1.1, see Dickson [Dic66, Ch. VI, p. 227-231].

**1.2. Which positive integer  $n$  can be written as the sum of two squares?** If  $n = x^2 + y^2$  ( $x, y \in \mathbb{Z}$ ) and  $p \mid n$ , then either  $p \mid \gcd(x, y)$  or  $(x/y)^2 \equiv -1 \pmod{p}$ , and hence either  $n/p^2$  is also the sum of two squares or  $p \not\equiv 3 \pmod{4}$  (by the quadratic reciprocity). It follows that each  $p \mid n$  with  $p \equiv 3 \pmod{4}$  must appear to an even power. On the other hand, the familiar Diophantus identity

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$$

shows that a product of integers of the form  $x^2 + y^2$  is also of the same form. Combining with Theorem 1.1.1 we obtain:

**Corollary 1.2.1.** *A positive integer  $n$  is of the form  $n = x^2 + y^2$  if and only if each prime factor  $p \equiv 3 \pmod{4}$  of  $n$  appears to an even power.*

**1.3. In how many different ways can one represent  $n$  as the sum of two squares?**

**Definition 1.3.1.** To answer this question, we naturally define the *representation number*

$$r(n) := \#\{(x, y) \in \mathbb{Z}^2 : n = x^2 + y^2\}.$$

In particular,  $n$  is of the form  $x^2 + y^2$  if and only if  $r(n) > 0$ .

**Example 1.3.2.**

$$\begin{aligned} 4 &= 0^2 + (\pm 2)^2 = (\pm 2)^2 + 0^2, & r(4) &= 4, \\ 5 &= (\pm 1)^2 + (\pm 2)^2 = (\pm 2)^2 + (\pm 1)^2, & r(5) &= 8, \\ 25 &= 0^2 + (\pm 5)^2 = (\pm 3)^2 + (\pm 4)^2 = (\pm 4)^2 + (\pm 3)^2, & r(25) &= 12. \end{aligned}$$

In his book *Fundamenta nova theoriae functionum ellipticarum* (1829), Jacobi proved the following general formula for the representation numbers.

**Theorem 1.3.3** (Jacobi).

$$r(n) = 4 \left( \sum_{\substack{d \mid n \\ d \equiv 1 \pmod{4}}} 1 - \sum_{\substack{d \mid n \\ d \equiv 3 \pmod{4}}} 1 \right).$$

As a by-product, Jacobi’s formula shows that

$$\begin{aligned} p \equiv 1 \pmod{4} &\Rightarrow r(p) = 4(2 - 0) = 8 > 0, \\ p \equiv 3 \pmod{4} &\Rightarrow r(p) = 4(1 - 1) = 0. \end{aligned}$$

which gives an immediate (and different) proof of Theorem 1.1.1!

1.4. **Jacobi's proof.** Jacobi's proof of Theorem 1.3.3 involves *Jacobi's theta series*,

$$\theta := \sum_{n \in \mathbb{Z}} q^{n^2} = 1 + 2q + 2q^4 + 2q^9 + \dots .$$

The representation numbers  $r(n)$  naturally appears as the  $n$ -th coefficients of the square of Jacobi's theta series

$$\theta^2 = \left( \sum_{n \in \mathbb{Z}} q^{n^2} \right)^2 = \sum_{n \geq 0} r(n)q^n = 1 + 4q + 4q^2 + 4q^4 + 8q^5 + 4q^8 + \dots .$$

Jacobi used his theory of elliptic functions (including his famous *Triple Product Identity*) to derive the formula ([Jac29, p. 107])

$$\theta^2 = 1 + 4 \left( \frac{q}{1-q} - \frac{q^3}{1-q^3} + \frac{q^5}{1-q^5} - \frac{q^7}{1-q^7} + \dots \right),$$

which is easily seen to be equivalent to Theorem 1.3.3.

1.5. **Another proof using modular forms.** An alternative way of evaluating  $\theta^2$  is to view  $q = e^{2\pi i\tau}$ , and  $\theta = \theta(\tau)$  as a holomorphic function on the upper half plane

$$\mathcal{H} := \{\tau = x + iy \in \mathbb{C} : y = \text{Im}(\tau) > 0\}.$$

The function  $\theta(\tau)$  satisfies two transformation rules (see [Zag08, Proposition 9])

$$\theta(\tau + 1) = \theta(\tau), \quad \theta\left(-\frac{1}{4\tau}\right) = (-2i\tau)^{1/2}\theta(\tau).$$

The first rule is clear by the periodicity of the exponential function. The second rule can be proved using the Poisson summation formula and also plays a key role in Riemann's proof the functional equation of the Riemann zeta function (see [DS05, §4.9]). These rules amounts to say that

$$\theta(\tau) \in M_{1/2}(\Gamma_1(4))$$

is a *modular form* of weight  $1/2$  and level  $\Gamma_1(4)$ . Jacobi's theta series  $\theta(\tau)$  and its variants (under the general name of *theta series*) form one of most important classes of modular forms.

It follows that

$$\theta^2(\tau) \in M_1(\Gamma_1(4))$$

is a modular form of weight  $1$  and level  $\Gamma_1(4)$ . The space  $M_1(\Gamma_1(4))$  is in fact 1-dimensional ([Zag08, Proposition 3] or [DS05, Theorem 3.6.1]), so if one can construct another a modular form of weight  $1$  and level  $\Gamma_1(4)$ , then it has to be a scalar multiple of  $\theta^2(\tau)$ . We next construct such a modular form using *Eisenstein series*, another of most important classes of modular forms.

**Definition 1.5.1.** Let  $\chi : (\mathbb{Z}/4\mathbb{Z})^\times \xrightarrow{\sim} \{\pm 1\}$  be the (unique) non-trivial character. We define an *Eisenstein series*

$$(1.5.1.1) \quad G_k^\chi(\tau) = \sum_{(0,0) \neq (c,d) \in \mathbb{Z}^2, 4|c} \frac{\chi(d)}{(c\tau + d)^k},$$

where  $\chi(d)$  is understood to be  $0$  when  $(4, d) \neq 1$ .

When  $k \geq 3$ , the series (1.5.1.1) is absolutely convergent and is nonzero only when  $k$  is odd. When  $k \geq 3$  is odd, it defines a modular form  $G_k^\chi(\tau) \in M_k(\Gamma_1(4))$  of weight  $k$ , level  $\Gamma_1(4)$  and character  $\chi$ . The constant term of the  $q$ -expansion of  $G_k^\chi(\tau)$  is nonzero and we let  $E_k^\chi(\tau)$  be a scalar

multiple of  $G_k^\chi(\tau)$  so the constant term is normalized to be 1. This *normalized Eisenstein series*  $E_k^\chi(\tau)$  then have the explicit  $q$ -expansion (see [DS05, §4.5])

$$(1.5.1.2) \quad E_k^\chi(\tau) = 1 + c_k^\chi \cdot \sum_{n \geq 1} \left( \sum_{d|n} \chi(d) d^{k-1} \right) q^n,$$

where  $c_k^\chi = 2/L(1-k, \chi)$  is related to a special value of the Dirichlet  $L$ -function  $L(s, \chi)$ .

When  $k = 1$ , the series (1.5.1.1) is not absolutely convergent, but one can still suitably modify it to obtain a modular form

$$E_1^\chi(\tau) \in M_1(\Gamma_1(4))$$

with the same formula (1.5.1.2) for its  $q$ -expansion, either using the Weierstrass  $\sigma$ -function (see [DS05, §4.8]), or using the analytic continuation of

$$(1.5.1.3) \quad G_1^\chi(\tau, s) := \sum_{(0,0) \neq (c,d) \in \mathbb{Z}^2, 4|c} \frac{\chi(d)}{(c\tau + d)|c\tau + d|^{2s}}, \quad \operatorname{Re}(s) > 1/2$$

to  $s = 0$  (see [Miy89, §7.2]). In particular, when  $k = 1$  the formula (1.5.1.2) simplifies to

$$E_1^\chi(\tau) = 1 + c_1^\chi \cdot \sum_{n \geq 1} \left( \sum_{d|n} \chi(d) \right) q^n.$$

As explained,  $E_1^\chi(\tau)$  must be a scalar multiple of  $\theta^2(\tau)$ . Since both of them have constant coefficient 1, we indeed have the equality

$$(1.5.1.4) \quad \theta^2(\tau) = E_1^\chi(\tau).$$

Comparing the coefficient before  $q$ , we obtain  $c_1^\chi = 4$  and hence

$$r(n) = 4 \sum_{d|n} \chi(d),$$

which proves Theorem 1.3.3.

**Remark 1.5.2.** As a by-product of the proof, we also obtain  $L(0, \chi) = \frac{1}{2}$  from  $c_1^\chi = 4$ . Via the functional equation of  $L(s, \chi)$ , this is equivalent to the famous Leibniz formula for  $\pi$  (1676),

$$L(1, \chi) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.$$

To summarize, Jacobi's Theorem 1.3.3 can be proved using the identity of two modular forms (1.5.1.4), namely using a relation of the form

$$\boxed{\text{theta series} \longleftrightarrow \text{Eisenstein series.}}$$

Notice that the Fourier coefficients of theta series encodes representation numbers of quadratic forms, while the Fourier coefficients of Eisenstein series are generalized divisor sums which are more explicit.

## 2. SIEGEL–WEIL FORMULA

**2.1. Siegel’s formula.** Siegel [Sie35] generalizes the formula (1.5.1.4) from the binary quadratic form  $x^2 + y^2$  to more general quadratic forms in arbitrary number of variables. Let  $\Lambda$  be a positive definite quadratic lattice over  $\mathbb{Z}$  of rank  $m$  with quadratic form  $Q : \Lambda \rightarrow \mathbb{Z}$ . Denote by  $(\ , \ )$  the associated symmetric bilinear form, defined by

$$(x, y) := Q(x + y) - Q(x) - Q(y),$$

(so  $Q(x) = \frac{1}{2}(x, x)$ ). Denote by  $\text{Sym}_n(\mathbb{Z})$  the set of symmetric matrices whose diagonal entries are in  $\mathbb{Z}$  and whose off-diagonal entries are in  $\frac{1}{2}\mathbb{Z}$ . Denote by  $\text{Sym}_n(\mathbb{Z})_{\geq 0} \subseteq \text{Sym}_n(\mathbb{Z})$  the subset of positive semi-definite matrices.

**Definition 2.1.1.** For  $T \in \text{Sym}_n(\mathbb{Z})_{\geq 0}$ , define the (generalized) *representation number*

$$r_\Lambda(T) := \#\{(x_1, \dots, x_n) \in \Lambda^n : \frac{1}{2}((x_i, x_j))_{1 \leq i, j \leq n} = T\}.$$

Define *Siegel’s theta series*

$$(2.1.1.1) \quad \theta_\Lambda(\tau) := \sum_{T \in \text{Sym}_n(\mathbb{Z})_{\geq 0}} r_\Lambda(T) q^T, \quad q^T := e^{2\pi i \text{tr} T \tau},$$

a holomorphic function on *Siegel’s half space*

$$\mathcal{H}_n := \{\tau = x + iy : x \in \text{Sym}_n(\mathbb{R}), y \in \text{Sym}_n(\mathbb{R})_{>0}\}.$$

Using the Poisson summation formula, Siegel proved that  $\theta_\Lambda(\tau)$  is a Siegel modular form on  $\mathcal{H}_n$  of weight  $m/2$ .

**Example 2.1.2.** Notice that when  $n = 1$ , Siegel’s half space  $\mathcal{H}_n$  recovers the usual upper half plane  $\mathcal{H}$ , and Siegel’s theta series  $\theta_\Lambda(\tau)$  recovers Jacobi’s theta series

$$\theta_\Lambda(\tau) := \sum_{n \geq 0} r_\Lambda(n) q^n, \quad r_\Lambda(n) := \#\{x \in \Lambda : Q(x) = n\}.$$

In general, the theta series  $\theta_\Lambda$  for a lattice  $\Lambda$  may fail to be an Eisenstein series on the nose, but Siegel’s formula shows that the weighted average of theta series within its genus class is always a Siegel Eisenstein series on  $\mathcal{H}_n$ :

$$\boxed{\text{weighted average of theta series} \longleftrightarrow \text{Siegel Eisenstein series.}}$$

More precisely, recall that two quadratic lattices  $\Lambda, \Lambda'$  are in the same *genus*, if they are isomorphic over  $\mathbb{R}$  and over  $\mathbb{Z}_p$  for all primes  $p$ . Denote by  $\text{Gen}(\Lambda)$  the set of isomorphism classes of quadratic lattices in the same genus of  $\Lambda$ . Denote by  $\text{Aut}(\Lambda)$  the automorphism group of  $\Lambda$  as a quadratic lattice.

**Theorem 2.1.3** (Siegel). *The following identity holds:*

$$(2.1.3.1) \quad \frac{\sum_{\Lambda' \in \text{Gen}(\Lambda)} \frac{1}{\#\text{Aut}(\Lambda')} \cdot \theta_{\Lambda'}(\tau)}{\sum_{\Lambda' \in \text{Gen}(\Lambda)} \frac{1}{\#\text{Aut}(\Lambda')}} = E_\Lambda(\tau).$$

Here  $E_\Lambda(\tau)$  is a certain normalized Siegel Eisenstein series on  $\mathcal{H}_n$  of weight  $m/2$ .

**Example 2.1.4.** Consider the case  $m = 2, n = 1$  and  $\Lambda = \mathbb{Z}^2$  equipped with the quadratic form  $Q = x^2 + y^2$ . Then

$$\theta_\Lambda(\tau) = \theta^2(\tau), \quad E_\Lambda(\tau) = E_1^X(\tau).$$

In this case  $\text{Gen}(\Lambda)$  is a singleton and Siegel’s formula recovers (1.5.1.4).

**Example 2.1.5** (cf. [Ser73, V.2.3]). Siegel's formula is extremely useful in studying the arithmetic of quadratic forms. For example, one can deduce his famous *mass formula* (also known as the Smith–Minkowski–Siegel mass formula), which computes the *mass* of  $\text{Gen}(\Lambda)$ , defined to be weighted size

$$\sum_{\Lambda' \in \text{Gen}(\Lambda)} \frac{1}{\#\text{Aut}(\Lambda')}$$

as an Euler product of local factors indexed by primes  $p$ .

For example, consider the simplest case when  $\Lambda$  is

- *unimodular*, i.e.,  $\det(\frac{1}{2}((x_i, x_j)_{i,j=1}^n)) \in \{\pm 1\}$  for a  $\mathbb{Z}$ -basis  $\{x_1, \dots, x_n\}$  of  $\Lambda$ ,
- and *even*, i.e., 2 divides  $Q(x)$  for all  $x \in \Lambda$ .

The rank  $m$  of any unimodular even lattice  $\Lambda$  is necessarily a multiple of 8. Siegel's mass formula computes the mass of  $\text{Gen}(\Lambda)$  explicitly as

$$\sum_{\Lambda' \in \text{Gen}(\Lambda)} \frac{1}{\#\text{Aut}(\Lambda')} = \frac{2\zeta(2)\zeta(4)\cdots\zeta(m-2)\zeta(m/2)}{\text{vol}(S^0)\text{vol}(S^1)\cdots\text{vol}(S^{m-1})} = \frac{B_{m/2}}{m} \prod_{1 \leq j < m/2} \frac{B_{2j}}{4^j}.$$

where  $\zeta(s)$  is the Riemann zeta function,  $\text{vol}(S^{k-1}) = \frac{2\pi^{k/2}}{\Gamma(k/2)}$  is the volume of the unit  $(k-1)$ -sphere and  $B_k$  is the  $k$ -th Bernoulli number.

**Example 2.1.6** (cf. [Ser73, VII.6.6]). Let  $\Lambda$  be the root lattice of type  $E_8$ , defined by

$$\Lambda := \{x = (x_1, \dots, x_8) \in \mathbb{Z}^8 \cup (\mathbb{Z} + 1/2)^8 : \sum_{i=1}^8 x_i \equiv 0 \pmod{2}\}, \quad Q = \sum_{i=1}^8 x_i^2.$$

Then  $\Lambda$  is unimodular and even. Siegel's mass formula computes that

$$\sum_{\Lambda' \in \text{Gen}(\Lambda)} \frac{1}{\#\text{Aut}(\Lambda')} = \frac{B_4 B_2 B_4 B_6}{8 \cdot 4 \cdot 8 \cdot 12} = \frac{-1/30}{8} \frac{1/6}{4} \frac{-1/30}{8} \frac{1/42}{12} = \frac{1}{696729600}.$$

The fact that  $\#\text{Aut}(\Lambda) = 696729600$  (which is also the order of the Weyl group of type  $E_8$ ) then implies that  $\Lambda$  is the *unique* unimodular and even lattice of rank 8.

In this case  $E_\Lambda(\tau)$  is related to the classical Eisenstein series  $E_4(\tau) \in M_4(\text{SL}_2(\mathbb{Z}))$  of weight 4 and level 1 by

$$E_\Lambda(\tau) = E_4(2\tau) = 1 + 240 \sum_{n \geq 1} \left( \sum_{d|n} d^3 \right) q^{2n} = 1 + 240q^2 + 2160q^4 + 6720q^6 + \cdots.$$

Here the factor 240 agrees with the constant  $c_k = 2/\zeta(1-k)$  for the weight  $k = 4$ . Siegel's formula then implies that

$$r_\Lambda(2n) := \#\left\{x = (x_i) \in \Lambda : \sum_{i=1}^8 x_i^2 = 2n\right\} = 240 \sum_{d|n} d^3.$$

In particular, we recover that there are  $r_\Lambda(2) = 240$  roots in the  $E_8$  root system (see Figure 1).

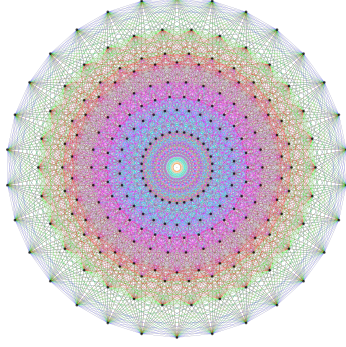


FIGURE 1. A 2-dimensional projection of the convex hull of the  $E_8$  root system. The black dots are the 240 roots. The plane the roots are projected into is one of the four planes in  $\mathbb{R}^8$  stabilized by a Coxeter element  $w$  of the Weyl group. The Coxeter element  $w$  acts on this plane via rotation through an angle of  $\frac{2\pi}{h}$ , where  $h = 30$  is the Coxeter number. It acts on the 240 roots regularly and divides them into 8 orbits of size 30 around circles of different radii. Colors of 6720 edge lines are varied based on their projected lengths to increase contrast.<sup>1</sup>

**2.2. Siegel–Weil formula.** In a series of works ([Sie36, Sie37, Sie51, Sie52]), Siegel further generalized his formula from definite to indefinite quadratic forms and from the base field  $\mathbb{Q}$  to totally real fields. In the indefinite case the formula is more difficult to state, as the theta series  $\theta_\Lambda(\tau)$  is divergent and it is necessary to introduce extra weight functions in the definition to ensure convergence. The situation was greatly clarified by Weil [Wei65] using powerful tools from representation theory, especially due to his use of the *Weil representation*.

Let  $V$  be a quadratic space over  $\mathbb{Q}$  of dimension  $m$  with bilinear form  $(\ , \ )$ . For simplicity we assume that  $m$  is *even* (so the weight  $m/2$  of the relevant Siegel modular forms is integral, see Remark 2.2.12). Consider the reductive dual pair  $(G, H) = (\mathrm{Sp}(2n), \mathrm{O}(V))$ , where  $\mathrm{Sp}(2n)$  is the symplectic group of the standard  $2n$ -dimensional symplectic space over  $\mathbb{Q}$  and  $\mathrm{O}(V)$  is the orthogonal group of  $V$ . Let  $P = MN \subseteq G = \mathrm{Sp}(2n)$  be the standard Siegel parabolic subgroup, so that under the standard basis we have

$$M(\mathbb{Q}) = \left\{ m(a) = \begin{pmatrix} a & 0 \\ 0 & {}_t a^{-1} \end{pmatrix} : a \in \mathrm{GL}_n(\mathbb{Q}) \right\},$$

$$N(\mathbb{Q}) = \left\{ n(b) = \begin{pmatrix} 1_n & b \\ 0 & 1_n \end{pmatrix} : b \in \mathrm{Sym}_n(\mathbb{Q}) \right\}.$$

**Definition 2.2.1.** Let  $\mathbb{A}$  be the ring of adèles of  $\mathbb{Q}$ . We fix the standard additive character  $\psi : \mathbb{A} \rightarrow \mathbb{C}^\times$  whose archimedean component is given by  $\psi_\infty : \mathbb{A}_\infty = \mathbb{R} \rightarrow \mathbb{C}^\times, x \mapsto e^{2\pi i x}$ . The (*Schrödinger model of the*) *Weil representation*  $\omega = \omega_{V, \psi}$  is the representation of  $G(\mathbb{A}) \times H(\mathbb{A})$  on the space of

<sup>1</sup>Computer generated picture by Jgmoxness, downloaded from <https://commons.wikimedia.org/wiki/File:E8Petrie.svg>

Schwartz functions  $\mathcal{S}(V(\mathbb{A})^n)$  such that for any  $\varphi \in \mathcal{S}(V(\mathbb{A})^n)$  and  $\mathbf{x} \in V(\mathbb{A})^n$ ,

$$(2.2.1.1) \quad \begin{aligned} \omega(m(a))\varphi(\mathbf{x}) &= \chi_V(\det a)|\det a|^{m/2}\varphi(\mathbf{x} \cdot a), & m(a) &\in M(\mathbb{A}), \\ \omega(n(b))\varphi(\mathbf{x}) &= \psi(\tfrac{1}{2} \operatorname{tr} b(\mathbf{x}, \mathbf{x}))\varphi(\mathbf{x}), & n(b) &\in N(\mathbb{A}), \\ \omega(w)\varphi(\mathbf{x}) &= \widehat{\varphi}(\mathbf{x}), & w &= \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}, \\ \omega(h)\varphi(\mathbf{x}) &= \varphi(h^{-1} \cdot \mathbf{x}), & h &\in H(\mathbb{A}). \end{aligned}$$

Here

- $\chi_V : \mathbb{A}^\times / \mathbb{Q}^\times \rightarrow \mathbb{C}^\times$  is the quadratic character corresponds to the quadratic extension  $\mathbb{Q}(\sqrt{\operatorname{disc}(V)})/\mathbb{Q}$ , and  $\operatorname{disc}(V) \in \mathbb{Q}^\times / (\mathbb{Q}^\times)^2$  is the discriminant of  $V$  defined to be

$$\operatorname{disc}(V) := (-1)^{\binom{m}{2}} \det((x_i, x_j))_{i,j=1}^m$$

for any  $\mathbb{Q}$ -basis  $\{x_1, \dots, x_m\}$  of  $V$ .

- $|\cdot| : \mathbb{A}^\times \rightarrow \mathbb{R}_{>0}$  is the normalized absolute value.
- $(\mathbf{x}, \mathbf{y}) := ((x_i, y_j))_{i,j=1}^n \in \operatorname{Mat}_n(\mathbb{A})$  for  $\mathbf{x} = (x_1, \dots, x_n) \in V(\mathbb{A})^n$  and  $\mathbf{y} = (y_1, \dots, y_n) \in V(\mathbb{A})^n$ ,
- $\widehat{\varphi}$  is the Fourier transform of  $\varphi$  using the self-dual Haar measure on  $V(\mathbb{A})^n$  with respect to  $\psi$ ,

$$\widehat{\varphi}(\mathbf{x}) := \int_{V(\mathbb{A})^n} \varphi(\mathbf{y}) \psi(\operatorname{tr}(\mathbf{x}, \mathbf{y})) d\mathbf{y}.$$

**Remark 2.2.2.** There are many nice expository articles on the Weil representation, see e.g. [Kud96, Pra93, Pra98]. See also a more general formula of the Weil representation in [Kud96, Proposition 4.3]).

**Example 2.2.3.** When  $n = 1$  we have  $G = \operatorname{Sp}(2) = \operatorname{SL}(2)$ . The standard Siegel parabolic is the standard Borel subgroup of  $\operatorname{SL}(2)$  consisting of upper triangle matrices  $P = \left\{ \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \right\}$ , and  $M = \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \right\}$ ,  $N = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \right\}$  are the diagonal and upper unipotent matrices respectively. In this case the first two formulas in (2.2.1.1) simplify to

$$(2.2.3.1) \quad \omega \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \varphi(x) = \chi_V(a)|a|^{m/2}\varphi(ax), \quad \omega \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \varphi(x) = \psi(\tfrac{1}{2}b(x, x))\varphi(x),$$

for any  $a \in \mathbb{A}^\times$ ,  $b \in \mathbb{A}$ ,  $\varphi \in \mathcal{S}(V(\mathbb{A}))$  and  $x \in V(\mathbb{A})$ .

Our next goal is to use the Weil representation  $\omega$  to construct theta series and Siegel Eisenstein series, starting from any common choice of a Schwartz function  $\varphi \in \mathcal{S}(V(\mathbb{A})^n)$ .

**Definition 2.2.4.** Associated to  $\varphi \in \mathcal{S}(V(\mathbb{A})^n)$ , define the (two-variable) *theta function*

$$\theta(g, h, \varphi) := \sum_{\mathbf{x} \in V^n} \omega(g)\varphi(h^{-1}\mathbf{x}), \quad g \in G(\mathbb{A}), h \in H(\mathbb{A}).$$

Then  $\theta(g, h, \varphi)$  is invariant under  $G(\mathbb{Q}) \times H(\mathbb{Q})$  (automorphic on both  $G$  and  $H$  in a broad sense).

**Remark 2.2.5.** Using  $\theta(g, h, \varphi)$  as an integral kernel allows one to lift automorphic forms on  $G$  to automorphic forms on  $H$  (and vice versa): for a cuspidal automorphic representation  $\pi$  of  $G(\mathbb{A})$  and  $\phi \in \pi$ , define the *theta lift*  $\theta_\varphi(\phi)$  of  $\phi$  to  $H(\mathbb{A})$  by the Petersson inner product on  $G(\mathbb{Q}) \backslash G(\mathbb{A})$ ,

$$\theta_\varphi(\phi)(h) := \langle \theta(-, h, \varphi), \phi \rangle_{\text{Pet}} = \int_{G(\mathbb{Q}) \backslash G(\mathbb{A})} \theta(g, h, \varphi) \overline{\phi(g)} dg.$$

Then  $\theta_\varphi(\phi)$  is an automorphic form on  $H(\mathbb{A})$ . This may be viewed as the starting point of the modern theory of *theta correspondence*, which is indispensable in the study of automorphic forms and the Langlands correspondence. We refer to Gan [Gan14] for an excellent recent survey on theta correspondence.



**Example 2.2.6.** Assume that  $V$  is positive definite, then the *theta integral*

$$(2.2.6.1) \quad \int_{H(\mathbb{Q}) \backslash H(\mathbb{A})} \theta(g, h, \varphi) dh$$

or in other words, the theta lift of the constant function  $\mathbf{1}$  on  $H(\mathbb{A})$  to  $G(\mathbb{A})$ , is closely related to Siegel's theta series. More precisely, for  $\Lambda \subseteq V$  a lattice over  $\mathbb{Z}$ , we take the Schwartz function  $\varphi = (\otimes_p \varphi_p) \otimes \varphi_\infty \in \mathcal{S}(V(\mathbb{A})^n)$  such that

- $\varphi_p \in \mathcal{S}(V(\mathbb{Q}_p)^n)$  is the characteristic function of  $(\Lambda \otimes \mathbb{Z}_p)^n$ ,
- $\varphi_\infty \in \mathcal{S}(V(\mathbb{R})^n)$  is the standard Gaussian function  $\varphi_\infty(\mathbf{x}) = e^{-\pi \operatorname{tr}(\mathbf{x}, \mathbf{x})}$ .

For  $\tau = \mathbf{x} + iy \in \mathcal{H}_n$ , we consider  $g_\tau = n(\mathbf{x})m(a) \in G(\mathbb{R})$ , where  $a \in \operatorname{GL}_n(\mathbb{R})$  such that  $y = a^t a$ . Then  $g_\tau \cdot i1_n = \tau$ . By (2.2.1.1) we have

$$(2.2.6.2) \quad \omega_\infty(g_\tau) \varphi_\infty(\mathbf{x}) = \chi_\infty(\det a) |\det a|^{m/2} \cdot q^{\frac{1}{2}(\mathbf{x}, \mathbf{x})}.$$

Define the *classical theta integral*

$$(2.2.6.3) \quad \chi_\infty(\det a)^{-1} |\det a|^{-m/2} \cdot \int_{H(\mathbb{Q}) \backslash H(\mathbb{A})} \theta(g_\tau, h, \varphi) dh.$$

Then it recovers the weighted average of theta series in (2.1.3.1) (see e.g., [Han13, §4.6], [KR14, §7]).

In fact, let  $K \subseteq H(\mathbb{A})$  be the stabilizer of  $\Lambda$ , then we have a bijection

$$H(\mathbb{Q}) \backslash H(\mathbb{A}) / K \xrightarrow{\sim} \operatorname{Gen}(\Lambda), \quad h \mapsto \Lambda' := h(\Lambda \otimes \widehat{\mathbb{Z}}) \cap V.$$

Let  $\{h_i\}$  be a complete set of representatives of  $H(\mathbb{Q}) \backslash H(\mathbb{A}) / K$  and let  $\{\Lambda_i\}$  be the corresponding representatives of  $\operatorname{Gen}(\Lambda)$  under this bijection. Then

$$\int_{H(\mathbb{Q}) \backslash H(\mathbb{A})} \theta(g_\tau, h, \varphi) dh = \sum_i \int_{H(\mathbb{Q}) \backslash H(\mathbb{Q}) h_i K} \theta(g_\tau, h, \varphi) dh.$$

Using  $H(\mathbb{Q}) \cap h_i K h_i^{-1} = \operatorname{Aut}(\Lambda_i)$ , each summand evaluates to

$$\int_{H(\mathbb{Q}) \backslash H(\mathbb{Q}) h_i K h_i^{-1}} \theta(g_\tau, h h_i, \varphi) dh = \frac{1}{\# \operatorname{Aut}(\Lambda_i)} \int_{h_i K h_i^{-1}} \theta(g_\tau, h h_i, \varphi) dh.$$

Unfolding the definition, the second integral equals

$$\int_K \theta(g_\tau, h_i h, \varphi) dh = \int_K \sum_{\mathbf{x} \in V^n} \omega_\infty(g_\tau) \varphi(h^{-1} h_i^{-1} \mathbf{x}) dh,$$

which by our choice of  $\{\varphi_p\}$  evaluates to

$$\operatorname{vol}(K) \sum_{\mathbf{x} \in \Lambda_i^n} \omega_\infty(g_\tau) \varphi_\infty(\mathbf{x}).$$

Thus combining with (2.2.6.2) we know that the classical theta integral (2.2.6.3) evaluates to

$$\operatorname{vol}(K) \sum_i \frac{1}{\# \operatorname{Aut}(\Lambda_i)} \cdot \sum_{\mathbf{x} \in \Lambda_i^n} q^{\frac{1}{2}(\mathbf{x}, \mathbf{x})} = \operatorname{vol}(K) \sum_i \frac{1}{\# \operatorname{Aut}(\Lambda_i)} \cdot \theta_{\Lambda_i}(\tau).$$

Finally notice that

$$\operatorname{vol}(K) \sum_i \frac{1}{\# \operatorname{Aut}(\Lambda_i)} = \operatorname{vol}(H(\mathbb{Q}) \backslash H(\mathbb{A})),$$

thus if we normalize the Haar measure  $dh$  such that  $\operatorname{vol}(H(\mathbb{Q}) \backslash H(\mathbb{A})) = 1$ , then the classical theta integral (2.2.6.3) recovers the weighted average of theta series in (2.1.3.1).

**Definition 2.2.7.** Associated to  $\varphi \in \mathcal{S}(V(\mathbb{A})^n)$ , also define the *Siegel Eisenstein series*

$$E(g, s, \varphi) := \sum_{\gamma \in P(\mathbb{Q}) \backslash G(\mathbb{Q})} \Phi_\varphi(\gamma g, s), \quad g \in G(\mathbb{A}), s \in \mathbb{C},$$

where

$$(2.2.7.1) \quad \mathcal{S}(V(\mathbb{A})^n) \rightarrow \text{Ind}_{P(\mathbb{A})}^{G(\mathbb{A})}(\chi_V |\cdot|^{s + \frac{n+1}{2}}), \quad \varphi \mapsto \Phi_\varphi(g, s) := \omega(g)\varphi(0) \cdot |\det a(g)|^{s-s_0}$$

is the *standard Siegel–Weil section* of the degenerate principal series representation of  $G(\mathbb{A})$  and

$$s_0 := \frac{m - (n + 1)}{2}.$$

Here we write  $g = nm(a)k$  under the Iwasawa decomposition  $G(\mathbb{A}) = N(\mathbb{A})M(\mathbb{A})K$  for  $K$  the standard maximal open compact subgroup of  $G(\mathbb{A})$ , and the quantity  $|\det a(g)| := |\det a|$  is well-defined.

**Example 2.2.8.** Similarly define the *classical Siegel Eisenstein series*

$$E(\tau, s, \varphi) := \chi_\infty(\det a)^{-1} |\det a|^{-m/2} \cdot E(g_\tau, s, \varphi).$$

When  $V$  is positive definite and  $\varphi$  is chosen as in Example 2.2.6, the special value  $E(\tau, s_0, \varphi)$  at  $s = s_0$  essentially recovers  $E_\Lambda(\tau)$  in (2.1.3.1).

For example, consider the case  $m = 2$ ,  $n = 1$  and  $\Lambda = \mathbb{Z}^2$  equipped with the quadratic form  $Q = x^2 + y^2$ . We have  $\text{disc}(V) = -1$  and the quadratic character  $\chi_V$  corresponds to the quadratic extension  $\mathbb{Q}(i)/\mathbb{Q}$ , and hence corresponds to the Dirichlet character  $\chi : (\mathbb{Z}/4\mathbb{Z})^\times \simeq \{\pm 1\}$ . For  $\tau = x + iy \in \mathcal{H}$ , we have

$$g_\tau = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \in \text{SL}_2(\mathbb{R}).$$

For  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{Q})$ , one can compute that the Siegel–Weil section evaluates to

$$\Phi_\varphi(\gamma g_\tau, s) = \frac{\chi(d) \cdot y^{1/2}}{c\tau + d} \cdot \text{Im}(\gamma\tau)^{s/2} = \frac{\chi(d)}{c\tau + d} \cdot \frac{y^{(1+s)/2}}{|c\tau + d|^s},$$

using (2.2.6.2) together with (2.2.3.1) (or the more general [Kud96, Proposition 4.3]). Comparing with (1.5.1.3) we see that the classical Eisenstein series

$$E(\tau, 0, \varphi) = y^{-1/2} \cdot E(g_\tau, 0, \varphi)$$

at  $s_0 = 0$  essentially recovers the Eisenstein series  $E_1^X(\tau)$  of weight 1 (up to a nonzero constant).

The Siegel Eisenstein series  $E(g, s, \varphi)$  converges absolutely when  $\text{Re}(s) > \frac{n+1}{2}$ . It has a meromorphic continuation to  $s \in \mathbb{C}$  and satisfies a functional equation relating  $s \leftrightarrow -s$  (i.e., centered at  $s = 0$ ). The Siegel–Weil formula gives a precise identity of the form

$$\boxed{\text{theta integral} \longleftrightarrow \text{value of Siegel Eisenstein series at } s = s_0.}$$

Notice that  $s = s_0$  is the unique point such that the map  $\varphi \mapsto \Phi_\varphi(g, s)$  in (2.2.7.1) is  $G(\mathbb{A})$ -equivariant, so that both sides of identity at least have the same transformation behavior with respect to the Weil representation.

**Theorem 2.2.9** (Siegel–Weil formula [Wei65, KR88a, KR88b]). *Let  $\alpha$  be the dimension of a maximal isotropic subspace of  $V$ . If  $\alpha = 0$  (i.e.,  $V$  is anisotropic) or  $\alpha > 0$  and  $m - \alpha > n + 1$ , then  $E(g, s, \varphi)$  is holomorphic at  $s_0$  and*

$$\kappa \cdot \int_{H(\mathbb{Q}) \backslash H(\mathbb{A})} \theta(g, h, \varphi) \, dh = E(g, s_0, \varphi),$$

where  $\kappa = 1$  if  $m > n + 1$  or  $\kappa = 2$  otherwise. Here the Haar measure  $dh$  is normalized so that  $\text{vol}(H(\mathbb{Q}) \backslash H(\mathbb{A})) = 1$ .

**Example 2.2.10.** When  $V$  is positive definite (so  $\alpha = 0$ ), the Siegel–Weil formula recovers Siegel’s formula (see Examples 2.2.8 and 2.2.6), and in particular recovers Jacobi’s formula as a special case of  $m = 2$ ,  $n = 1$ ,  $(G, H) = (\text{Sp}(2), \text{O}(2, 0))$  (see Example 2.1.4).

**Remark 2.2.11.** The condition in Theorem 2.2.9 is known as *Weil’s convergence condition*, which ensures the convergence of the theta integral. It is a long effort starting with the work of Kudla–Rallis [KR94] to generalize the Siegel–Weil formula outside the convergence range and for all reductive dual pairs of classical groups. We refer to Gan–Qiu–Takeda [GQT14] for the most general Siegel–Weil formula and a nice summary of the literature and history.

**Remark 2.2.12.** When  $m$  is odd, the theta series and Eisenstein series have half integral weights  $m/2$ , and are automorphic forms on the metaplectic cover  $\tilde{G} = \text{Mp}(2n)$  of  $\text{Sp}(2n)$ . In this case the Weil representation needs to be modified to be a representation of  $\tilde{G}(\mathbb{A}) \times H(\mathbb{A})$  and the Siegel–Weil formula still holds after modification.

### 3. GEOMETRIC SIEGEL–WEIL FORMULA

In this section we discuss an example of the Siegel–Weil formula in the indefinite case  $(G, H) = (\text{Sp}(4), \text{O}(2, 2))$  originating from the classical work of Hurwitz, and use it to motivate the more general geometric Siegel–Weil formula.

#### 3.1. Hurwitz class number relation.

**Definition 3.1.1.** For any positive integer  $D$ , the *Hurwitz class number*  $H(D)$  is defined to be the weighted size of  $\text{SL}_2(\mathbb{Z})$ -equivalence classes of positive definite binary quadratic forms

$$ax^2 + bxy + cy^2 \text{ with discriminant } b^2 - 4ac = -D, \quad a, b, c \in \mathbb{Z}.$$

Here the forms equivalent to  $a(x^2 + y^2)$  ( $D = 4a^2$ ) and  $a(x^2 + xy + y^2)$  ( $D = 3a^2$ ) are counted with multiplicities  $1/2$  and  $1/3$  respectively, due to extra symmetry.

**Example 3.1.2.**

$$\begin{aligned} H(3) &= \frac{1}{3} \longleftrightarrow \{x^2 + xy + y^2\} \\ H(8) &= 1 \longleftrightarrow \{x^2 + 2y^2\} \\ H(11) &= 1 \longleftrightarrow \{x^2 + xy + 3y^2\} \\ H(12) &= \frac{4}{3} \longleftrightarrow \{2(x^2 + xy + y^2), x^2 + 3y^2\} \\ H(20) &= 2 \longleftrightarrow \{x^2 + 5y^2, 2x^2 + 2xy + 3y^2\}. \end{aligned}$$

**Example 3.1.3.** Table 1 lists the first few Hurwitz class numbers.

$D$	3	4	7	8	11	12	15	16	19	20	23	24
$H(D)$	$\frac{1}{3}$	$\frac{1}{2}$	1	1	1	$\frac{4}{3}$	2	$\frac{3}{2}$	1	2	3	2

TABLE 1. Hurwitz class numbers

**Example 3.1.4.** When  $-D$  is a fundamental discriminant and  $D > 4$ , the Hurwitz class number  $H(D)$  is equal to the class number of the imaginary quadratic field  $\mathbb{Q}(\sqrt{-D})$ .

Understanding these class numbers  $H(D)$  remains a central subject in algebraic number theory. The following remarkable formula, which we call the *Hurwitz class number relation* or the *Hurwitz formula*, gives an elementary expression for a certain sum of Hurwitz class numbers.

**Theorem 3.1.5** (Kronecker [Kro60], Gierster [Gie83], Hurwitz [Hur85]). *If  $m$  is not a perfect square, then*

$$(3.1.5.1) \quad \sum_{dd'=m} \max\{d, d'\} = \sum_{t \in \mathbb{Z}, 4m-t^2 > 0} H(4m-t^2).$$

**Example 3.1.6.** When  $m = 3$ , the Hurwitz class number relation says

$$\begin{aligned} 3 + 3 &= H(3) + H(8) + H(11) + H(12) + H(11) + H(8) + H(3) \\ &= \frac{1}{3} + 1 + 1 + \frac{4}{3} + 1 + 1 + \frac{1}{3}. \end{aligned}$$

When  $m = 5$ , the Hurwitz class number relation says

$$\begin{aligned} 5 + 5 &= H(4) + H(11) + H(16) + H(19) + H(20) + H(19) + H(16) + H(11) + H(4) \\ &= \frac{1}{2} + 1 + \frac{3}{2} + 1 + 2 + 1 + \frac{3}{2} + 1 + \frac{1}{2}, \end{aligned}$$

A quite nontrivial way to decompose the integer 6 and 10 respectively!

Hurwitz [Hur85] proved this formula using the modular  $j$ -invariant and the modular polynomial  $\Phi_m(x, y) \in \mathbb{Z}[x, y]$  of level  $m$  (which defines the modular curve  $Y_0(m)$  over  $\mathbb{Z}$ ). He recognized the LHS of (3.1.5.1) as  $\deg \Phi_m(x, x)$  and computed this degree in a different way involving Hurwitz class numbers to arrive at the formula (3.1.5.1).

**3.2. A geometric proof.** (cf. [GK93]). From the modern point of view, we have a nice geometric interpretation of Hurwitz's proof, in terms of the geometry of the modular curve

$$Y(\mathbb{C}) = \mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{H}.$$

The modular curve  $Y(\mathbb{C})$  is the moduli space of elliptic curves:

$$\begin{aligned} Y &= \{E : \text{elliptic curve over } \mathbb{C} \text{ (up to isomorphism)}\} \\ \tau &\longmapsto E_\tau = \mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau). \end{aligned}$$

which allows one to define a canonical model of  $Y$  as an algebraic curve over  $\mathbb{Q}$ . Each elliptic curve  $E_\tau$  has a Weierstrass equation

$$E_\tau : y^2 = x^3 + A_\tau x + B_\tau.$$

The  $j$ -invariant

$$j(E_\tau) := 1728 \cdot \frac{4A_\tau^3}{4A_\tau^3 + 27B_\tau^2}$$

only depends on the isomorphism class of  $E_\tau$  and gives rise to an isomorphism

$$(3.2.0.1) \quad j : Y \xrightarrow{\sim} \mathbb{A}^1, \quad \tau \longmapsto j(E_\tau).$$

**Definition 3.2.1.** Define the surface  $X$  to be the product of two modular curves,

$$X := Y \times Y = \{(E, E')\},$$

which is the moduli space of pairs of elliptic curves  $(E, E')$ . For each positive integer  $m$ , we define the *modular correspondence*  $Z(m)$  over the surface  $X$  by

$$Z(m) := \{(E, E', \varphi) : \phi : E \xrightarrow{\deg m} E'\} \rightarrow X,$$

parameterizing a pair of elliptic curves  $(E, E')$  together with a degree  $m$  isogeny  $\phi : E \rightarrow E'$ .

The isogeny  $\phi$  imposes one nontrivial condition and thus  $Z(m)$  defines a divisor on  $X$ . For example, when  $m = 1$  the modular correspondence is nothing but the diagonally embedded modular curve

$$Z(1) = \Delta(Y) \subseteq X = Y \times Y.$$

Given two divisors  $Z(m)$  and  $Z(1)$  on the surface  $X$ , one expects that the intersection  $Z(m) \cap Z(1)$  should be 0-dimensional. When this is the case (i.e., when  $Z(m)$  and  $Z(1)$  intersects *properly*), we obtain a *geometric intersection number*  $\langle Z(m), Z(1) \rangle_X$  by counting the number of intersection points weighted by intersection multiplicities (see Figure 2).

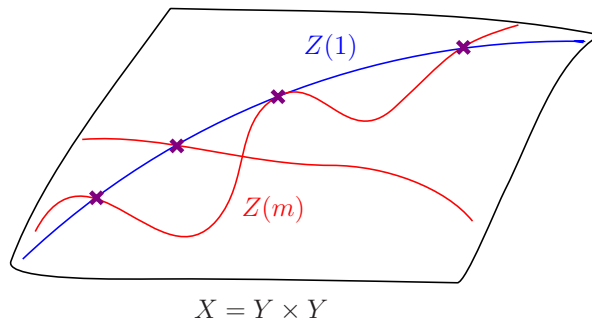


FIGURE 2. Geometric intersection number

Now a curious observation comes: the geometric intersection number  $\langle Z(m), Z(1) \rangle_X$  is equal to the LHS of (3.1.5.1),

$$(3.2.1.1) \quad \langle Z(m), Z(1) \rangle_X = \sum_{dd'=m} \max\{d, d'\}.$$

In fact, under the isomorphism (3.2.0.1) we know that  $X$  has a natural compactification

$$\overline{X} = \mathbb{P}^1 \times \mathbb{P}^1.$$

Since  $\overline{Z(1)} = \Delta(\mathbb{P}^1)$ , the well-known cohomological equivalence on  $\mathbb{P}^1 \times \mathbb{P}^1$ ,

$$\Delta(\mathbb{P}^1) \sim \mathbb{P}^1 \times \{\text{pt}\} + \{\text{pt}\} \times \mathbb{P}^1$$

then implies that

$$\langle \overline{Z(m)}, \overline{Z(1)} \rangle_{\overline{X}} = \langle \overline{Z(m)}, \mathbb{P}^1 \times \{\text{pt}\} \rangle_{\overline{X}} + \langle \overline{Z(m)}, \{\text{pt}\} \times \mathbb{P}^1 \rangle_{\overline{X}} = 2 \sum_{d|m} d,$$

where the last equality comes from counting the number of degree  $m$  isogenies with a fixed source (resp. target) elliptic curve. The desired identity (3.2.1.1) then follows from subtracting the contribution at  $\{\infty\} = \mathbb{P}^1 \setminus \mathbb{A}^1$ .

On the other hand, via the moduli interpretation we have

$$Z(m) \cap Z(1) = \{(E, E) : E \xrightarrow{\deg m} E\}.$$

When  $m$  is a perfect square, we know that  $Z(m)$  contains  $Z(1)$  (by considering the multiplication-by- $\sqrt{m}$  isogeny  $E \xrightarrow{[\sqrt{m}]} E$ ), and thus  $Z(m)$  and  $Z(1)$  do not intersect properly. However, when  $m$  is not a perfect square (as assumed in Theorem 3.1.5),  $E$  has to be an elliptic curve with complex multiplication by an imaginary quadratic order  $\mathbb{Z}[\alpha]$ , where  $\alpha\bar{\alpha} = m$ , and thus  $Z(m)$  and  $Z(1)$  do intersect properly. Using the theory of *complex multiplication*, counting the weighted number of such elliptic curves with complex multiplication exactly gives the sum of Hurwitz class numbers as the RHS of (3.1.5.1),

$$(3.2.1.2) \quad \langle Z(m), Z(1) \rangle_X = \sum_{t \in \mathbb{Z}, 4m - t^2 > 0} H(4m - t^2).$$

Combining (3.2.1.1) and (3.2.1.2) completes our sketch of the geometric proof of the Hurwitz formula (see Gross–Keating [GK93] for complete details).

**3.3. Hurwitz formula as a geometric Siegel–Weil formula.** The Hurwitz class numbers appearing in the Hurwitz formula (3.1.5.1) also naturally appear as Fourier coefficients of Siegel Eisenstein series. More precisely, consider the Siegel Eisenstein series  $E(\tau, s)$  on  $\mathrm{Sp}(4)$  of weight 2,

$$E(\tau, s) := \sum_{\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in P(\mathbb{Z}) \backslash \mathrm{Sp}_4(\mathbb{Z})} \det(C\tau + D)^{-2} \frac{\det \mathrm{Im}(\tau)^{s-1/2}}{|\det(C\tau + D)|^{2s-1}} = \sum_{T \in \mathrm{Sym}_2(\mathbb{Z})_{>0}} E_T(\tau, s) q^T + \cdots.$$

Then we have (up to a normalizing constant)

$$H(4m - t^2) = E_T(\tau, s_0 = 1/2), \quad T = \begin{pmatrix} m & t/2 \\ t/2 & 1 \end{pmatrix}.$$

Notice that the condition  $4m - t^2 > 0$  neatly translates to the condition  $T > 0$ .

We summarize our discussion by the following diagram:

$$(3.3.0.3) \quad \begin{array}{ccc} \sum_{dd'=m} \max\{d, d'\} & = & \sum_{t \in \mathbb{Z}, 4m - t^2 > 0} H(4m - t^2) \\ \parallel & & \parallel \\ \langle Z(m), Z(1) \rangle_X & = & \sum_{T = \begin{pmatrix} m & t/2 \\ t/2 & 1 \end{pmatrix} > 0} E_T(\tau, 1/2) \\ \updownarrow & & \updownarrow \\ \text{intersection number of} & = & \text{sum of Fourier coefficients of} \\ \text{2 divisors on } X & & \text{a Siegel Eisenstein series on } \mathrm{Sp}(4). \end{array}$$

In this way the Hurwitz formula can be viewed as a *geometric* Siegel–Weil formula for the pair  $(G, H) = (\mathrm{Sp}(4), \mathrm{O}(2, 2))$ , where one replaces the theta integral on  $\mathrm{O}(2, 2)$  by a *geometric theta series*, i.e., the generating series of geometric intersection numbers of modular correspondences for the surface  $X = Y \times Y$ ,

$geometric \text{ theta series on } X \longleftrightarrow value \text{ of Siegel Eisenstein series on } \mathrm{Sp}(4) \text{ at } s_0 = 1/2.$

Notice here the natural appearance of the product of modular curves due to the exceptional isomorphism  $\mathrm{SO}(2, 2) \simeq \mathrm{SL}_2 \times \mathrm{SL}_2 / \{\pm 1\}$ . This geometric Siegel–Weil formula further computes a more general geometric intersection number  $\langle Z(m), Z(n) \rangle_X$  as the sum of  $E_T(\tau, 1/2)$  with  $T = \begin{pmatrix} m & t/2 \\ t/2 & n \end{pmatrix} > 0$ .

**Remark 3.3.1.** The remarkable discovery that generating series involving intersection numbers of cycles are modular originates from the work of Hirzebruch–Zagier [HZ76] on Hilbert modular surfaces. Historically [HZ76] was the primary motivation in Kudla’s work discussed below (cf. the introduction of [KM90, Kud97a]) and also in the work of Gross–Kohnen–Zagier [GKZ87]. See Example 3.5.1 and Remark 3.5.5 (i), (ii).

**3.4. Orthogonal Shimura varieties.** Kudla proved a more general *geometric Siegel–Weil formula* by replacing the surface  $X = Y \times Y$  by an orthogonal Shimura variety of arbitrary dimension. Our next goal is to discuss Kudla’s formula.

Let  $F$  be a totally real number field. Pick a real place  $w$  of  $F$ . Let  $V$  be a quadratic space over  $F$  of dimension  $m$  such that for any place  $v | \infty$  of  $F$ ,

$$\text{the } F_v\text{-quadratic space } V_v := V \otimes_F F_v \text{ has signature } \begin{cases} (m-2, 2), & \text{if } v = w, \\ (m, 0), & \text{if } v \neq w. \end{cases}$$

Let  $G = \mathrm{GSpin}(V)$ , which sits in an exact sequence

$$1 \rightarrow \mathbb{G}_m \rightarrow G \rightarrow \mathrm{SO}(V) \rightarrow 1.$$

Associated to any open compact subgroup  $K \subseteq G(\mathbb{A}_F^\infty)$ , we have a *GSpin Shimura variety*  $X = \mathrm{Sh}_G$ , which has a smooth canonical model of dimension  $m-2$  over the reflex field  $F$  (viewed as a subfield of  $\mathbb{C}$  via the embedding induced by the place  $w$ ) and admits complex uniformization

$$X(\mathbb{C}) = G(F) \backslash [\mathbb{D} \times G(\mathbb{A}_F^\infty) / K].$$

Here  $\mathbb{D}$  is the hermitian symmetric domain of oriented negative 2-planes in  $V_w$ . Unfolding the definition we may rewrite  $X(\mathbb{C})$  as a disjoint union of quotients of  $\mathbb{D}$  by congruence subgroups of  $G(F)$  (see [Kud04, §1]). The Shimura variety  $X$  is quasi-projective, and is projective when  $V$  is anisotropic (e.g., when  $F \neq \mathbb{Q}$ , by the signature condition).

**Remark 3.4.1.** One technical reason that one prefers to work with the Shimura variety  $X$  associated to  $G$  (instead of  $\mathrm{SO}(V)$ ) is that it is of *Hodge type* (instead of *abelian type*), and admits an embedding into a Siegel modular variety (the moduli space of polarized abelian varieties) of larger dimension.

**Example 3.4.2.** (cf. [Kud04, HP14, FH00]) Consider  $F = \mathbb{Q}$ . Via accidental isomorphisms between  $G = \mathrm{GSpin}(m-2, 2)$  and classical groups of symplectic type in low ranks, the Shimura varieties  $X$  recover many classical modular varieties in low dimensions (see Table 2, where  $D$  is a quaternion algebra over  $\mathbb{Q}$  and  $E = \mathbb{Q}(\sqrt{d})$  is a real quadratic field). When  $\dim X \leq 19$ ,  $X$  is also closely related to the moduli space of polarized K3 surfaces and has proved to be useful for studying the arithmetic of K3 surfaces.

**3.5. Kudla’s generating series of special cycles and the modularity conjecture.** The Shimura variety  $X$  is equipped with special divisors  $Z(m) \rightarrow X$  generalizing in the case  $X = Y \times Y$ . Via an embedding into a Siegel modular variety,  $Z(m)$  parameterizes certain polarized abelian varieties together with a *special endomorphism* of degree  $m$  (see [MP16, §5]).

**Example 3.5.1.** The special divisor  $Z(m)$  recovers

- Heegner points on modular curves and Shimura curves when  $\dim X = 1$  (cf. [Kud04, Appendix]),

dim $X$	$G = \mathrm{GSpin}(m-2, 2)$	$X$
1	$\mathrm{GL}_2$ or $D^\times$	modular/Shimura curve
2	$\mathrm{GL}_2 \times_{\mathbb{G}_m} \mathrm{GL}_2$ or $\mathrm{GL}_{2,E}^{\det \in \mathbb{Q}^\times}$	product of modular curves or Hilbert modular surfaces
3	$\mathrm{GSp}_4$	Siegel 3-fold = moduli of abelian surfaces
4	$\mathrm{GU}(2, 2)$ (up to center)	moduli of abelian 4-folds with complex multiplication
6	$\mathrm{GSp}(4, D)$	moduli of abelian 8-folds with quaternion multiplication

TABLE 2. Examples of  $\mathrm{GSpin}$  Shimura varieties.

- modular correspondences  $Z(m)$  on  $X = Y \times Y$  considered in §3.2 when  $\dim X = 2$  and  $G$  is split,
- Hirzebruch–Zagier cycles [HZ76] on Hilbert modular surfaces when  $\dim X = 2$  and  $G$  is nonsplit.

More generally, for any  $y \in V$  with  $(y, y) > 0$ , its orthogonal complement  $V_y \subseteq V$  has rank  $n-1$ . The embedding  $G_y := \mathrm{GSpin}(V_y) \hookrightarrow G = \mathrm{GSpin}(V)$  defines a Shimura subvariety of codimension 1

$$\mathrm{Sh}_{G_y} \rightarrow X = \mathrm{Sh}_G.$$

For any  $x \in V(\mathbb{A}_F^\infty)$  with  $(x, x) \in F_{>0}$ , there exists  $y \in V$  and  $g \in G(\mathbb{A}_F^\infty)$  such that  $y = gx$ . Define the *special divisor*

$$Z(x) \rightarrow X$$

to be the  $g$ -translate of  $\mathrm{Sh}_{G_y}$ . For any  $\mathbf{x} = (x_1, \dots, x_n) \in V(\mathbb{A}_F^\infty)^n$  with  $(\mathbf{x}, \mathbf{x}) \in \mathrm{Sym}_n(F)_{>0}$ , define the *special cycle* (of codimension  $n$ )

$$Z(\mathbf{x}) = Z(x_1) \cap \dots \cap Z(x_n) \rightarrow X.$$

Here  $\cap$  denotes the fiber product over  $X$ . More generally, for a  $K$ -invariant Schwartz function  $\varphi \in \mathcal{S}(V(\mathbb{A}_F^\infty)^n)^K$  and  $T \in \mathrm{Sym}_n(F)_{>0}$ , define the *weighted special cycle*

$$Z_\varphi(T) = \sum_{\substack{\mathbf{x} \in K \backslash V(\mathbb{A}_F^\infty)^n \\ (\mathbf{x}, \mathbf{x}) = T}} \varphi(\mathbf{x}) Z(\mathbf{x}) \in \mathrm{CH}^n(X)_\mathbb{C} := \mathrm{CH}^n(X) \otimes \mathbb{C}.$$

Here  $\mathrm{CH}^n(X)$  is the Chow group of algebraic cycles of codimension  $n$  on  $X$  (up to rational equivalence). With extra care, we can also define  $Z_\varphi(T) \in \mathrm{CH}^n(X)_\mathbb{C}$  for any  $T \in \mathrm{Sym}_n(F)_{\geq 0}$  (see [Kud04, YZZ09]).

**Definition 3.5.2.** Define *Kudla’s generating series of special cycles*

$$(3.5.2.1) \quad Z_\varphi(\tau) = \sum_{T \in \mathrm{Sym}_n(F)_{\geq 0}} Z_\varphi(T) q^T,$$

as a formal sum valued in  $\mathrm{CH}^n(X)_\mathbb{C}$ , where

$$\tau \in \mathcal{H}_n = \{\mathbf{x} + iy : \mathbf{x} \in \mathrm{Sym}_n(F_\infty), y \in \mathrm{Sym}_n(F_\infty)_{>0}\}, \quad q^T := \prod_{v|\infty} e^{2\pi i \mathrm{tr} T \tau_v}.$$

**Remark 3.5.3.** Analogous constructions of special divisors and Kudla’s generating series also apply to Shimura varieties of unitary type associated to hermitian spaces with signature  $(m-1, 1)$  at one archimedean place and signature  $(m, 0)$  at all other archimedean places (see [Liu11a]). These Shimura varieties of orthogonal/unitary type can be naturally viewed as Shimura varieties associated to totally definite incoherent quadratic/hermitian spaces ([Zha19, Gro20]), see §6.5. In the unitary case, we obtain a generating series of the form

$$(3.5.3.1) \quad Z_\varphi(\tau) = \sum_{T \in \mathrm{Herm}_n(F)_{\geq 0}} Z_\varphi(T) q^T,$$



where we replace positive semi-definite symmetric matrices  $\text{Sym}_n(F)_{\geq 0}$  by positive semi-definite hermitian matrices  $\text{Herm}_n(F)_{\geq 0}$ , and replace Siegel's half space by the hermitian half space

$$\tau \in \mathcal{H}_n := \{x + iy : x \in \text{Herm}_n(F_\infty), y \in \text{Herm}_n(F_\infty)_{>0}\}.$$

We may view  $Z_\varphi(\tau)$  as a geometric theta series, now valued in Chow groups for cycles of arbitrary codimension  $n$ . The analogy to Siegel's theta series (2.1.1.1) and the theta integral (2.2.6.1) leads to the following Kudla's modularity conjecture.

**Conjecture 3.5.4** (Kudla's modularity). *The formal generating series  $Z_\varphi(\tau)$  converges absolutely and defines a modular form on  $\mathcal{H}_n$  of weight  $m/2$  valued in  $\text{CH}^n(X)_\mathbb{C}$ .*

**Remark 3.5.5.**

(i) The analogous modularity in Betti cohomology, i.e., the modularity of the generating series valued in  $H^{2n}(X(\mathbb{C}), \mathbb{C})$  defined by the image of  $Z_\varphi(T)$  under the cycle class map

$$\text{CH}^n(X) \rightarrow H^{2n}(X(\mathbb{C}), \mathbb{Z}),$$

is known by the classical work of Kudla–Millson [KM90]. The special case of special divisors on Hilbert modular surfaces dates back to Hirzebruch–Zagier [HZ76] (see also Funke–Millson [FM14]).

(ii) Kudla's modularity conjecture was originally formulated for orthogonal Shimura varieties over  $\mathbb{Q}$  ([Kud97a, Kud04]). In this case, Borcherds [Bor99] proved the conjecture for the divisor case  $m = 1$  (the special case of Heegner points on modular curves dates back to the classical work of Gross–Kohnen–Zagier [GKZ87]). Zhang [Zha09] proved the modularity for general  $n$  assuming the absolute convergence of the series. Bruinier–Westerholt–Raum [BWR15] proved the desired convergence and hence established Kudla's modularity conjecture for orthogonal Shimura varieties over  $\mathbb{Q}$ . More recently, Bruinier–Zemel [BZ19] has extended the modularity to toroidal compactifications of orthogonal Shimura varieties when  $n = 1$ .

(iii) For orthogonal Shimura varieties over totally real fields, Yuan–Zhang–Zhang [YZZ09] proved the modularity for  $n = 1$  (see also Bruinier [Bru12] for a different proof) and reduce the  $n > 1$  case to the convergence.

(iv) Conjecture 3.5.4 in the unitary case was formulated by Liu [Liu11a], who also proved the case  $n = 1$  and reduce the  $n > 1$  case to the convergence. Recently Xia [Xia21] proved the desired convergence when  $E = \mathbb{Q}(\sqrt{-d})$  for  $d = 1, 2, 3, 7, 11$  (in the notation of §6.5), and thus established Conjecture 3.5.4 in these cases.

(v) Kudla [Kud04, Problem 4] also proposed the modularity problem in the *arithmetic Chow group*  $\widehat{\text{CH}}^n(\mathcal{X})$  of a suitable (compactified) integral model  $\mathcal{X}$  of  $X$  (see [GS90, BGKK07] and also [Sou92]). The problem seeks to define canonically an explicit arithmetic generating series  $\widehat{Z}_\varphi(\tau)$  valued in  $\widehat{\text{CH}}^n(\mathcal{X})_\mathbb{C}$  which lifts  $Z_\varphi(\tau)$  under the restriction map

$$\widehat{\text{CH}}^n(\mathcal{X}) \rightarrow \text{CH}^n(X),$$

and such that  $\widehat{Z}_\varphi(\tau)$  is modular. When  $n = 1$ , this *arithmetic modularity* was proved by Howard–Madapusi Pera [HMP20] (orthogonal groups over  $\mathbb{Q}$ ) and Bruinier–Howard–Kudla–Rapoport–Yang [BHK<sup>+</sup>20] (unitary groups over  $\mathbb{Q}$ ). Several low dimensional cases were also proved:

- Shimura/modular curves (Kudla–Rapoport–Yang [KRY06], Sankaran [San14] and Du–Yang [DY19]),
- Hilbert modular surfaces (Bruinier–Burgos Gil–Kühn [BBGK07]),
- Product of modular curves (Berndt–Kühn [BK12a, BK12b]).

We also mention the arithmetic modularity of the difference of two arithmetic theta series by Ehlen–Sankaran [ES18] for  $n = 1$  (unitary groups over  $\mathbb{Q}$ ), the almost arithmetic modularity by Mihatsch–Zhang [MZ21, Theorem 4.3] for  $n = 1$  (unitary groups over totally real fields  $F \neq \mathbb{Q}$ ), the arithmetic modularity of Fourier–Jacobi coefficients for general  $n$  by Sankaran [San20] (anisotropic orthogonal groups), and several striking recent works involving applications of arithmetic modularity [AGHMP18, Zha21, SSTT19].

(vi) Finally, we mention several recent works on the modularity conjecture for more general classes of orthogonal and unitary Shimura varieties (indefinite at more than one archimedean places) by Rosu–Yott [RY20], Kudla [Kud19] and Maeda [Mae19, Mae21].

**3.6. Kudla’s geometric Siegel–Weil formula.** In the special case  $n = \dim X$ , the generating series  $Z_\varphi(\tau)$  in (3.5.2.1) is valued in  $\mathrm{CH}^{\dim X}(X)_\mathbb{C}$  (i.e., the Chow group of 0-cycles). When  $X$  is projective, composing with the degree map

$$\mathrm{deg} : \mathrm{CH}^{\dim X}(X)_\mathbb{C} \rightarrow \mathbb{C},$$

we obtain a generating series  $\mathrm{deg} Z_\varphi(\tau)$  valued in  $\mathbb{C}$ . Its terms encode geometric intersection numbers between special divisors on  $X$  generalizing the case  $X = Y \times Y$  considered in §3.3. Kudla proved the following remarkable geometric version of the Siegel–Weil formula (by analogy with Theorem 2.2.9 specialized to the case  $n = \dim X = m - 2$  and so  $s_0 = 1/2$ ).

**Theorem 3.6.1** (Kudla’s geometric Siegel–Weil formula [Kud97a, Corollary 10.5]). *Assume that  $X$  is projective (i.e.,  $V$  is anisotropic). Take  $n = \dim X$ . Then for any  $\varphi \in \mathcal{S}(V(\mathbb{A}_F^\infty)^n)^K$  the following identity holds (up to a nonzero constant depending only on choices of measures)*

$$(3.6.1.1) \quad \mathrm{deg} Z_\varphi(\tau) \doteq E(\tau, 1/2, \varphi \otimes \varphi_\infty).$$

Here  $\varphi_\infty \in \mathcal{S}(V(F_\infty)^n)$  is a certain Schwartz function constructed from the Kudla–Millson Schwartz form ([KM86]).

Thus Kudla’s geometric Siegel–Weil formula is a precise identity of the form

$$\boxed{\text{geometric theta series on } X \longleftrightarrow \text{value of Siegel Eisenstein series on } \mathrm{Sp}(2n) \text{ at } s_0 = 1/2.}$$

**Remark 3.6.2.** Kudla in fact proved a geometric Siegel–Weil formula for the generating series of special cycles of all dimension, i.e., without assuming  $n = \dim X$ . Here and in the next section we focus on the case of 0-cycles and refer to Kudla’s excellent surveys [Kud04, Kud02b, Kud02a] for the general case.

#### 4. ARITHMETIC SIEGEL–WEIL FORMULA

In this section we discuss an arithmetic version of the Siegel–Weil formula. Parallel to the previous section, we will use an example in the case  $(G, H) = (\mathrm{Sp}(6), \mathrm{O}(2, 2))$  considered by Gross–Keating to motivate the more general case.

**4.1. Gross–Keating formula.** Gross–Keating took the geometric point of view of the Hurwitz class number relation and found a remarkable generalization for *arithmetic* intersection numbers. As the moduli space of elliptic curves, the modular curve  $Y$  has a canonical integral model  $\mathcal{Y}$  over  $\mathrm{Spec} \mathbb{Z}$  such that  $\mathcal{Y}(\mathbb{C}) = Y$ . The integral model  $\mathcal{Y}$  is an *arithmetic surface* fibered over the arithmetic curve  $\mathrm{Spec} \mathbb{Z}$ , and its fiber above  $p$  is a smooth curve in characteristic  $p$  (Figure 3).

Analogously, the surface  $X$  has a canonical integral model  $\mathcal{X}$  over  $\mathrm{Spec} \mathbb{Z}$ , which is an *arithmetic threefold*. The modular correspondence  $Z(m) \rightarrow X$  naturally extends to a divisor  $\mathcal{Z}(m) \rightarrow \mathcal{X}$ . Now

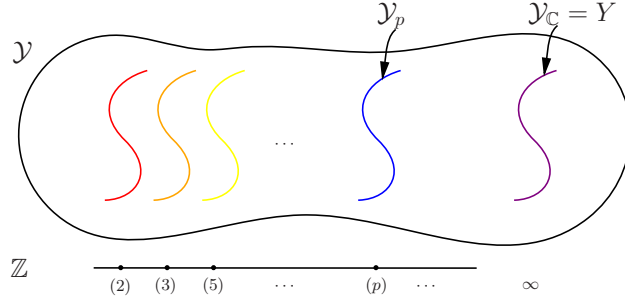


FIGURE 3. Arithmetic surface

on the arithmetic threefold  $\mathcal{X}$ , we need 3 (instead of 2) divisors so that the intersection has expected dimension 0. Define the *arithmetic intersection number* by

$$\langle \mathcal{Z}(m_1), \mathcal{Z}(m_2), \mathcal{Z}(m_3) \rangle_{\mathcal{X}} := \sum_p \langle \mathcal{Z}(m_1), \mathcal{Z}(m_2), \mathcal{Z}(m_3) \rangle_{\mathcal{X}_p} \cdot \log p,$$

where  $\langle \mathcal{Z}(m_1), \mathcal{Z}(m_2), \mathcal{Z}(m_3) \rangle_{\mathcal{X}_p}$  encodes the intersection number supported in the fiber  $\mathcal{X}_p$ . This definition makes sense when these three divisors intersect properly (or more generally when their intersection is supported in finitely many fibers  $\mathcal{X}_p$ ).

Is there a formula for this arithmetic intersection number analogous to (3.3.0.3)? The index set for  $T$  should be  $\text{Sym}_3(\mathbb{Z})$  to put the three diagonal entries  $m_1, m_2, m_3$  in  $T$ , and thus we need to look at the Siegel Eisenstein series  $E(\tau, 0)$  on  $\text{Sp}(6)$  of weight 2 (instead of on  $\text{Sp}(4)$ ). Then the relevant special point in the Siegel–Weil formula is  $s_0 = 0$ , the central point. However, it is too naive to expect  $\langle \mathcal{Z}(m_1), \mathcal{Z}(m_2), \mathcal{Z}(m_3) \rangle_{\mathcal{X}}$  to be equal to  $E(\tau, 0)$ : the arithmetic intersection number involves a  $\mathbb{Q}$ -linear combination of  $\log p$ 's and hence is no longer a rational number like  $E_T(\tau, 1/2)$  in (3.3.0.3). Moreover, the Eisenstein series  $E(\tau, s)$  turns out to have an odd functional equation at the center  $s = 0$ , and hence  $E(\tau, 0) = 0$  automatically! This automatic vanishing suggests that it is interesting to look at its first *derivative*  $E'(\tau, 0)$  at  $s = 0$ .

After these two appropriate modifications: replacing  $\text{Sp}(4)$  by  $\text{Sp}(6)$  and replacing the value at  $s_0 = 1/2$  by the central derivative at  $s_0 = 0$ , it turns out that we do have the following remarkable formula relating the arithmetic intersection numbers to  $E'(\tau, 0)$  (see a nice exposition of the proof in [VGW<sup>+</sup>07]).

**Theorem 4.1.1** (Gross–Keating [GK93], Gross–Kudla–Zagier). *Assume there is no positive definite binary quadratic form representing  $m_1, m_2, m_3$  simultaneously. Then (up to an explicit constant)*

$$(4.1.1.1) \quad \langle \mathcal{Z}(m_1), \mathcal{Z}(m_2), \mathcal{Z}(m_3) \rangle_{\mathcal{X}} \quad \doteq \quad \sum_{T = \begin{pmatrix} m_1 & * & * \\ * & m_2 & * \\ * & * & m_3 \end{pmatrix} > 0} E'_T(\tau, 0)$$

↑

↓

*arithmetic intersection number  
of 3 divisors on  $\mathcal{X}$*

↑

↓

*sum of Fourier coefficients of the derivative  
of a Siegel Eisenstein series on  $\text{Sp}(6)$*

**Remark 4.1.2.** The assumption on  $m_1, m_2, m_3$  is analogous to the assumption that  $m$  is not a perfect square in Theorem 3.1.5, which guarantees that the three divisors intersect properly.

**4.2. Arithmetic Siegel–Weil formula.** Parallel to (3.3.0.3), the Gross–Keating formula can be viewed as an *arithmetic* Siegel–Weil formula for the pair  $(G, H) = (\mathrm{Sp}(6), \mathrm{O}(2, 2))$ , where one replaces the theta integral on  $\mathrm{O}(2, 2)$  by a generating series of arithmetic intersection numbers on the arithmetic threefold  $\mathcal{X}$ ,

$$\boxed{\text{arithmetic theta series on } \mathcal{X} \longleftrightarrow \text{central derivative of Siegel Eisenstein series on } \mathrm{Sp}(6).}$$

Of course there is nothing stopping us from considering the higher dimensional case. In fact Kudla ([Kud97b]) and Kudla–Rapoport ([KR99, KR00a, KR14]) proposed vast conjectural generalizations of the Gross–Keating formula by

- (1) Taking  $\mathcal{X}$  to be a suitable regular integral model  $\mathcal{X}$  of the Shimura variety  $X$  of orthogonal/unitary type considered in §3.4–3.5 associated to quadratic/hermitian spaces.
- (2) Defining suitable integral models  $\mathcal{Z}(m) \rightarrow \mathcal{X}$  of the special divisors  $Z(m) \rightarrow X$ , and more generally integral models  $\mathcal{Z}(T)$  for special divisors  $Z(T)$ .

**Remark 4.2.1.** Such a regular integral model  $\mathcal{X}$  is known to exist for certain level subgroups  $K$  defining  $X$  (cf. [LZ22a, §14] in the unitary case). To avoid technicalities we will be vague about the choice of level structures and integral models, but see Remark 5.3.2 (iii), (iv).

Now take  $n = \dim \mathcal{X}$  ( $= \dim X + 1$ ) so that the intersection of  $n$  special divisors has expected dimension 0. We have a natural decomposition

$$\mathcal{Z}(m_1) \cap \cdots \cap \mathcal{Z}(m_n) = \bigsqcup_{T = \begin{pmatrix} m_1 & * & * \\ * & \cdots & * \\ * & * & m_n \end{pmatrix}} \mathcal{Z}(T)$$

index by symmetric/hermitian matrices  $T$  with diagonal entries  $m_1, \dots, m_n$ . Here  $\cap$  denotes the fiber product over  $\mathcal{X}$ . When  $T > 0$ , it turns out that  $\mathcal{Z}(T)$  is supported in finitely many fibers  $\mathcal{Z}(T)_p$ , and we have a well-defined  $T$ -part of the arithmetic intersection number

$$\langle \mathcal{Z}(m_1), \dots, \mathcal{Z}(m_n) \rangle_T := \sum_p \langle \mathcal{Z}(m_1), \dots, \mathcal{Z}(m_n) \rangle_{\mathcal{Z}(T)_p} \cdot \log p.$$

Now we are ready to state the conjecture on an arithmetic version of the Siegel–Weil formula (by analogy with Theorem 2.2.9 in the orthogonal case specialized to  $n = m + 1$  and so  $s_0 = 0$  is the central point), which is known as the *Kudla–Rapoport conjecture* in the unitary case ([KR14, Conjecture 11.10]).

**Conjecture 4.2.2** (Arithmetic Siegel–Weil formula, nonsingular part). *Take  $n = \dim \mathcal{X}$ . Then for any  $T \in \mathrm{Sym}_n(F)_{>0}$  (resp.  $T \in \mathrm{Herm}_n(F)_{>0}$ ) in the orthogonal (resp. unitary) case with diagonal entries  $m_1, \dots, m_n$ , the following identity holds (up to a nonzero constant depending only on choices of measures),*

(4.2.2.1)

$$\begin{array}{ccc} \langle \mathcal{Z}(m_1), \dots, \mathcal{Z}(m_n) \rangle_T & \stackrel{?}{=} & E'_T(\tau, 0) \\ \updownarrow & & \updownarrow \\ \text{arithmetic intersection number} & & \text{Fourier coefficient of the central derivative of a} \\ \text{of } n \text{ divisors on } \mathcal{X} & & \text{Siegel Eisenstein series on } \mathrm{Sp}(2n) \text{ (resp. } \mathrm{U}(n, n)) \end{array}$$

**Remark 4.2.3.** In general the  $n$  special divisors do not intersect properly, and a more sophisticated definition of the arithmetic intersection numbers is needed (cf. Definition 5.1.3). In particular, with the correct definition the conjecture works even for improper intersections.

Thus the arithmetic Siegel–Weil formula a precise conjectural identity of the form

$$\boxed{\text{arithmetic theta series on } \mathcal{X} \longleftrightarrow \text{central derivative of Siegel Eisenstein series on } \mathrm{Sp}(2n) (\mathrm{U}(n, n))}.$$

Now we can state one of the main results of [LZ22a] (see [LZ22a, Theorem 1.3.1] for more precise technical assumptions).

**Theorem 4.2.4** (with Zhang [LZ22a]). *Conjecture 4.2.2 holds for arbitrary  $n$  in the unitary case.*

Our recent work with Zhang [LZ22b] has also established a slightly weaker semi-global (at a good odd prime  $p$ ) version of Conjecture 4.2.2 in the orthogonal case. We will discuss some key ideas of the proof in §5.

**Remark 4.2.5.**

(i) When  $n = 3$  and  $G$  is split (orthogonal case), the arithmetic Siegel–Weil formula recovers the Gross–Keating formula, and also treats the case of improper intersections (see Remarks 4.1.2, 4.2.3).

(ii) Theorem 4.2.4 was previously proved when  $n = 3$  by Terstiege [Ter11, Ter13] and when the intersection is 0-dimensional ([GK93, KR99, KR00a, KR11, BY21]).

(iii) There is also an *archimedean* part of the arithmetic Siegel–Weil formula, relating archimedean arithmetic intersection numbers with the nonsingular but indefinite Fourier coefficients of  $E'(\tau, 0)$ . These Fourier coefficients are nonholomorphic, unlike the positive definite Fourier coefficients in Conjecture 4.2.2. This archimedean arithmetic Siegel–Weil formula was proved by Liu [Liu11a] (unitary case), and Garcia–Sankaran [GS19] in full generality (see also Bruinier–Yang [BY21] for an alternative proof in the orthogonal case).

(iv) Kudla conjectured that there should also be a *singular* part of the arithmetic Siegel–Weil formula, relating the singular Fourier coefficients of  $E'(\tau, 0)$  to certain arithmetic intersection numbers. However the singular part is more difficult to prove, or even to formulate precisely, cf. [Kud04, Problem 6]. As a special case, the constant term of the arithmetic Siegel–Weil formula should roughly relate the arithmetic volume of  $\mathcal{X}_n$  to logarithmic derivatives of Dirichlet  $L$ -functions. Such an explicit arithmetic volume formula was proved by Hörmann [Hör14] (orthogonal case) and Bruinier–Howard [BH21] (unitary case), though a precise comparison with the constant term of  $E'(\tau, 0)$  is yet to be formulated and established.

(v) Ideally, putting all singular/nonsingular and archimedean/nonarchimedean parts together, one should arrive at a full arithmetic Siegel–Weil formula of the form in complete analogy to (3.6.1.1),

$$\widehat{\mathrm{deg}} \widehat{\mathcal{Z}}_\varphi(\tau) \stackrel{?}{=} E'(\tau, 0, \varphi \otimes \varphi_\infty).$$

Here  $\widehat{\mathcal{Z}}_\varphi(\tau)$  is the conjectural arithmetic theta series in Remark 3.5.5 (v),  $\widehat{\mathrm{deg}}$  is the arithmetic degree map

$$\widehat{\mathrm{deg}} : \widehat{\mathrm{CH}}^{\dim \mathcal{X}}(\mathcal{X})_{\mathbb{C}} \rightarrow \mathbb{C},$$

and  $\varphi_\infty$  is the standard Gaussian function on the totally positive definite space over  $F_\infty$ . The full arithmetic Siegel–Weil formula was established by Kudla, Rapoport and Yang ([KRY99, Kud97b, KR00b, KRY06]) for  $n = 1, 2$  (orthogonal case) in great generality. However, it remains an open problem to formulate such a precise full arithmetic Siegel–Weil formula in higher dimension.

(vi) Recently Feng–Yun–Zhang [FYZ21a] proved a higher Siegel–Weil formula over function fields for unitary groups, which relates nonsingular coefficients of the  $r$ -th derivative of Siegel Eisenstein series and intersection numbers of special cycles on moduli spaces of Drinfeld shtukas with  $r$  legs. The case  $r = 0$  (resp.  $r = 1$ ) can be viewed as an analogue of the Siegel–Weil formula (resp. the arithmetic

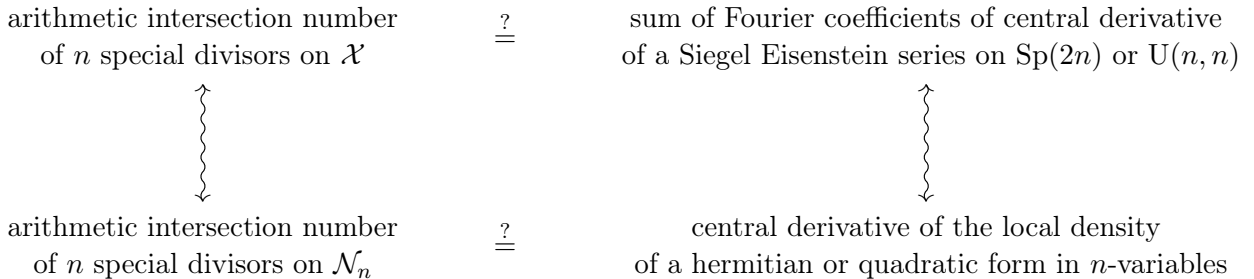
Siegel–Weil formula). Over function fields, the possibility of relating higher derivatives of analytic objects to intersection numbers was first discovered by Yun–Zhang [YZ17, YZ19] in the context of the higher Gross–Zagier formula. Over number fields, however, no analogue of such a higher Siegel–Weil formula (resp. higher Gross–Zagier formula) is currently known when  $r > 1$ . Feng–Yun–Zhang [FYZ21b] also defined higher theta series over function fields (including all singular terms) and conjectured their modularity. Previously, an arithmetic Siegel–Weil formula over function fields was proved by Wei [Wei19] for special cycles on moduli spaces of Drinfeld modules of rank 2 with complex multiplication (analogue of the special case  $n = r = 1$ ).

## 5. LOCAL ARITHMETIC SIEGEL–WEIL FORMULA

In order to prove the arithmetic Siegel–Weil formula (Conjecture 4.2.2), one first notices that it can be reduced to a local identity:

- (1) Geometric side (LHS): The arithmetic intersection numbers corresponding to nonsingular matrices  $T$  can be computed as a sum indexed by primes of  $\mathbb{Z}$  (or the ring of integers  $O_F$  in general). The local term at a finite prime  $p$  can be further reduced to an arithmetic intersection on a Rapoport–Zink space  $\mathcal{N}_n$ , which is a local analogue of Shimura varieties over  $\mathbb{Z}_p$  (or a completion of  $O_F$  in general), via the theory of  $p$ -adic uniformization of Shimura varieties ([RZ96]).
- (2) Analytic side (RHS): The nonsingular Fourier coefficients  $E_T(\tau, s)$  has a product expansion indexed by primes of  $\mathbb{Z}$ , and thus the derivative  $E'_T(\tau, 0)$  can also be written as a sum indexed by the primes of  $\mathbb{Z}$ . The term indexed by a finite prime  $p$  can be further reduced to the derivative of the local representation density of quadratic/hermitian forms over  $\mathbb{Z}_p$ .

This reduction step is illustrated in the following diagram:



The conjectural local identity on the bottom is known as the *local arithmetic Siegel–Weil formula*. The local arithmetic Siegel–Weil formula has been recently proved in our work with Zhang [LZ22a] (resp. [LZ22b]) in the unitary (resp. orthogonal) case. Next we will make this local conjecture more precise in the unitary case. In the unitary case, this reduction step was made precise by Kudla–Rapoport ([KR11], [KR14]) and the local conjecture is also known as the *local Kudla–Rapoport conjecture*.

**5.1. Geometric side.** Let  $p$  be an odd prime. Let  $F_0$  be a finite extension of  $\mathbb{Q}_p$  with residue field  $k = \mathbb{F}_q$  and uniformiser  $\varpi$ . Let  $F$  be the unramified quadratic extension of  $F_0$  (e.g.,  $F/F_0 = \mathbb{Q}_{p^2}/\mathbb{Q}_p$ ). Associated to this datum we have the *unitary Rapoport–Zink space*  $\mathcal{N}_n$ :

- It is a formal scheme over  $\mathrm{Spf}_{O_{\bar{F}}}$  of relative dimension  $n - 1$ , parameterizing deformations (up to quasi-isogenies) of a fixed  $O_F$ -hermitian formal  $p$ -divisible group  $\mathbb{X}_n/\bar{k}$  of relative height  $2n$ , dimension  $n$  and signature  $(1, n - 1)$ . Here  $O_{\bar{F}}$  is the completion of the maximal unramified extension of  $O_F$ .

- The space of *special homomorphisms*

$$\mathbb{V}_n = \mathrm{Hom}_{O_F}(\mathbb{X}_1, \mathbb{X}_n) \otimes_{O_F} F$$

has a structure of a (non-split)  $F$ -hermitian space of dimension  $n$ , coming from the principal polarization on  $\mathbb{X}_n$ .

- The unitary group  $U(\mathbb{V}_n)$  naturally acts on  $\mathcal{N}_n$  via the action on  $\mathbb{X}_n$ .
- Each vector  $x \in \mathbb{V}_n$  gives rise to a *special divisor* or *Kudla–Rapoport (KR) divisor*  $\mathcal{Z}(x) \subseteq \mathcal{N}_n$ , defined to be the locus where the homomorphism  $x$  deforms. This is the local analogue of the special divisor considered in Conjecture 4.2.2.

The Rapoport–Zink space  $\mathcal{N}_n$  is formally smooth over  $\mathrm{Spf} O_{\check{F}}$ , but its geometric structure is rather complicated. For example,  $\mathcal{N}_n$  is highly non-reduced: the reduced subscheme  $\mathcal{N}_n^{\mathrm{red}}$  has dimension  $\lfloor \frac{n-1}{2} \rfloor$ , near the middle dimension of  $\mathcal{N}_n$ . The structure of  $\mathcal{N}_n^{\mathrm{red}}$  was studied by Vollaard–Wedhorn ([VW11]), and they showed that  $\mathcal{N}_n^{\mathrm{red}}$  has a nice stratification into smooth varieties, known as the *Bruhat–Tits stratification*. Each closed stratum of the Bruhat–Tits stratification is isomorphic to a smooth projective variety  $DL_i$  of dimension  $i$  over  $\bar{k}$  ( $0 \leq i \leq \lfloor \frac{n-1}{2} \rfloor$ ), and the incidence relation between the closed strata resembles the combinatorial structure of the Bruhat–Tits building for quasi-split unitary groups over  $F_0$ . Here each  $DL_i$  is a *generalized Deligne–Lusztig variety* associated to the unitary group  $U(2i+1)$  over  $k$ .

**Example 5.1.1.** Take  $n = 1$ . Then  $\mathcal{N}_1 \cong \mathrm{Spf} O_{\check{F}}$ , and  $\mathcal{N}_1^{\mathrm{red}} = \{\mathrm{pt}\}$ .

**Example 5.1.2.** Take  $n = 3$ . Then  $\mathcal{N}_3$  has relative dimension 2 over  $\mathrm{Spf} O_{\check{F}}$ , while the reduced subscheme  $\mathcal{N}_3^{\mathrm{red}}$  has dimension 1 (Figure 4). In this case only two types of Deligne–Lusztig variety show up:

- (1)  $DL_0 = \{\mathrm{pt}\}$ , a single point.
- (2)  $DL_1 = \{x^{q+1} + y^{q+1} + z^{q+1} = 0\} \subseteq \mathbb{P}^2$ , the Fermat curve of degree  $q+1$ .

And  $\mathcal{N}_3^{\mathrm{red}}$  is an infinite tree, where

- (1) the number of  $DL_1$  containing a given  $DL_0$  is exactly  $q+1$ .
- (2) the number of  $DL_0$  contained in a given  $DL_1$  is exactly  $q^3+1$ .

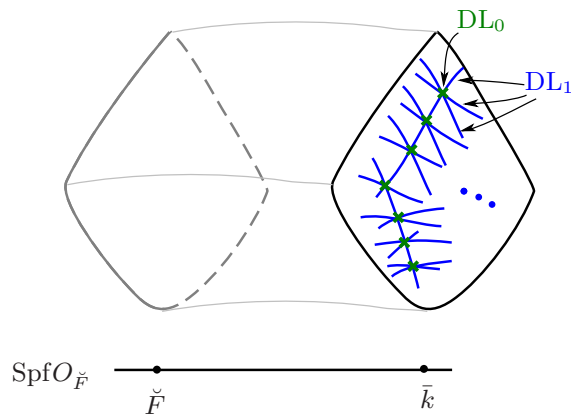


FIGURE 4.  $n = 3$

**Definition 5.1.3.** Let  $L \subseteq \mathbb{V}_n$  be an  $O_F$ -lattice of rank  $n$ . Let  $x_1, \dots, x_n$  be an  $O_F$ -basis of  $L$ . Define the *special cycle* or *Kudla–Rapoport (KR) cycle*

$$\mathcal{Z}(L) := \mathcal{Z}(x_1) \cap \dots \cap \mathcal{Z}(x_n) \subseteq \mathcal{N}_n.$$

Define the *arithmetic intersection number*

$$\text{Int}(L) := \chi(\mathcal{O}_{\mathcal{Z}(x_1)} \otimes^{\mathbb{L}} \dots \otimes^{\mathbb{L}} \mathcal{O}_{\mathcal{Z}(x_n)}),$$

where  $\chi$  denotes the Euler–Poincaré characteristic,  $\mathcal{O}_{\mathcal{Z}(x_i)}$  denotes the structure sheaf of the Kudla–Rapoport divisor  $\mathcal{Z}(x_i)$ , and  $\otimes^{\mathbb{L}}$  denotes the derived tensor product of coherent sheaves on  $\mathcal{N}_n$ . It is known (by Terstiege [Ter13] as extended in [LZ22a, Corollary 2.8.2], or by Howard [How19]) that  $\text{Int}(L)$  is independent of the choice of the basis  $x_1, \dots, x_n$  and hence is a well-defined invariant of  $L$  itself, justifying the notation.

**Remark 5.1.4.** When the intersection  $\mathcal{Z}(L)$  is 0-dimensional, we have

$$\mathcal{O}_{\mathcal{Z}(x_1)} \otimes^{\mathbb{L}} \dots \otimes^{\mathbb{L}} \mathcal{O}_{\mathcal{Z}(x_n)} = \mathcal{O}_{\mathcal{Z}(x_1)} \otimes \dots \otimes \mathcal{O}_{\mathcal{Z}(x_n)} = \mathcal{O}_{\mathcal{Z}(L)}$$

and thus  $\text{Int}(L)$  is nothing but the  $O_{\mathbb{F}}$ -length of  $\mathcal{O}_{\mathcal{Z}(L)}$  (= the sum of intersection multiplicities at all points). Even though  $\mathcal{Z}(L)$  is the intersection of  $n$  divisors in a  $n$ -dimensional formal scheme, in general  $\mathcal{Z}(L)$  may fail to have the expected dimension 0 due to improper intersection. In this case, the derived intersection is needed so that the intersection number is well-behaved.

**Example 5.1.5.** Take  $n = 3$  and  $L \cong \langle \varpi \rangle^3$  (the hermitian form with respect to an  $O_F$ -basis of  $L$  is  $\text{diag}\{\varpi, \varpi, \varpi\}$ ). In this case the intersection  $\mathcal{Z}(L)$  is not 0-dimensional, and in fact

$$\mathcal{Z}(L) \cong \text{DL}_1 = \{x^{q+1} + y^{q+1} + z^{q+1} = 0\} \subseteq \mathbb{P}^2.$$

The arithmetic intersection number turns out to be its topological Euler characteristic

$$\text{Int}(L) = \chi(\text{DL}_1) = 2 - 2g = 2 - q(q - 1) = 2 + q - q^2.$$

## 5.2. Analytic side.

**Definition 5.2.1.** Let  $L, M$  be two hermitian  $O_F$ -lattices of rank  $n, m$  respectively. Let  $S = \text{Rep}(M, L)$  be the  $O_{F_0}$ -scheme such that for any  $O_{F_0}$ -algebra  $R$ ,

$$S(R) = \text{Herm}(L \otimes_{F_0} R, M \otimes_{F_0} R),$$

where  $\text{Herm}$  denotes the set of hermitian module homomorphisms. The *local density* of representations of  $M$  by  $L$  is defined to be

$$\text{Den}(M, L) := \lim_{N \rightarrow +\infty} \frac{\#S(O_{F_0}/\varpi^N)}{q^{N \cdot \dim S_{F_0}}}.$$

It gives a quantitative measure of “how many different ways” one can embed  $L$  into  $M$  as a hermitian submodule.

**Example 5.2.2.** Consider  $L = M = \langle 1 \rangle^n$ , the rank  $n$  lattice with hermitian form given by the identity matrix. Then one can compute that

$$\text{Den}(\langle 1 \rangle^n, \langle 1 \rangle^n) = \prod_{i=1}^n (1 - (-q)^{-i}).$$

One can recognize it as the number  $\frac{G(k)}{q^{\dim G}}$ , where  $G$  is the unitary group  $U(\langle 1 \rangle^n)$  defined over  $O_{F_0}$ , or in fancier language, as the local  $L$ -factor of the Gross motive [Gro97] of the quasi-split unitary group in  $n$  variables.



The local density  $\text{Den}(M, L)$  has nice compatibility when replacing  $M$  by  $M \oplus \langle 1 \rangle^k$  for  $k \geq 0$ . More precisely, it is known ([Hir98, Theorem II]) that  $\text{Den}(\langle 1 \rangle^{n+k}, L)$  is a polynomial in  $(-q)^{-k}$  with  $\mathbb{Q}$ -coefficients.

**Example 5.2.3** ([KR11, p.677]).

$$(5.2.3.1) \quad \text{Den}(\langle 1 \rangle^{n+k}, \langle 1 \rangle^n) = \prod_{i=1}^n (1 - (-q)^{-i} X) \Big|_{X=(-q)^{-k}}.$$

**Definition 5.2.4.** Define the *normalized Siegel series*  $\text{Den}(X, L) \in \mathbb{Z}[X]$  such that

$$\text{Den}((-q)^{-k}, L) = \frac{\text{Den}(\langle 1 \rangle^{n+k}, L)}{\text{Den}(\langle 1 \rangle^{n+k}, \langle 1 \rangle^n)}.$$

These polynomials  $\text{Den}(X, L)$  plays an important role in computing the Fourier coefficients of Siegel Eisenstein series, and many works are devoted to proving more explicit formulas for them (see e.g., [Kit83, Kat99, Hir98, Hir12, IK16, CY20]). The local Siegel series satisfies a functional equation ([Hir12, Theorem 5.3])

$$(5.2.4.1) \quad \text{Den}(X, L) = (-X)^{\text{val}(L)} \cdot \text{Den}\left(\frac{1}{X}, L\right).$$

Here  $\text{val}(L) := \text{val}(\det(L)) \in \mathbb{Z}$  is the *valuation* of  $L$ . This may be viewed as a local analogue of the functional equation of Eisenstein series.

**Definition 5.2.5.** If  $\text{val}(L)$  is odd (equivalently,  $L \otimes_{O_F} F$  is a non-split hermitian space), then  $\text{Den}(1, L) = 0$  by the functional equation (5.2.4.1). In this case, define the *central derivative of the local density* by

$$\partial \text{Den}(L) := - \frac{d}{dX} \Big|_{X=1} \text{Den}(X, L).$$

**5.3. Local arithmetic Siegel–Weil formula.** Now we are ready to state the local arithmetic Siegel–Weil formula in the unitary case, originally conjectured by Kudla–Rapoport [KR11, Conjecture 1.3].

**Theorem 5.3.1** (Local arithmetic Siegel–Weil formula, with Zhang [LZ22a, Theorem 1.2.1]). *Let  $L \subseteq \mathbb{V}_n$  be an  $O_F$ -lattice of full rank  $n$ . Then*

$$\text{Int}(L) = \partial \text{Den}(L).$$

**Remark 5.3.2.**

- (i) The theorem was proved by Kudla–Rapoport [KR11] for  $n = 2$  and Terstiege [Ter13] for  $n = 3$ .
- (ii) The local arithmetic Siegel–Weil formula is proved for the orthogonal case in our work with Zhang [LZ22b] for arbitrary  $n$ . The case  $n = 3$  was previously proved by Gross–Keating [GK93] and Terstiege [Ter11].
- (iii) The local arithmetic Siegel–Weil formula is proved when the quadratic extension  $F/F_0$  is ramified for exotic smooth models in our work with Liu [LL21] for arbitrary even  $n$ , and for Krämer models<sup>2</sup> by Shi and He–Shi–Yang [Shi20, HSY20] for  $n = 2$ .

---

<sup>2</sup>During the refereeing process of this article, He–Shi–Yang [HSY22] has formulated a conjectural local arithmetic Siegel–Weil formula for Krämer models for arbitrary  $n$  and proved it for the case  $n = 3$ . The case for arbitrary  $n$  has been proved in our work with He–Shi–Yang [HLSY22].

(iv) It is more difficult to prove or formulate the local arithmetic Siegel–Weil formula in the presence of more general level structures (even when the quadratic extension  $F/F_0$  is unramified). In the unitary case [LZ22a] formulates and proves a local arithmetic Siegel–Weil formula when the level is the parahoric subgroup given by the stabilizer of an almost self-dual lattice (the case  $n = 2$  was previously proved by Sankaran [San17]). Recently Cho [Cho20] has proposed a general formulation for all minuscule parahoric levels in the unitary case.

**Example 5.3.3.** Take  $n = 3$  and  $L \cong \langle \varpi \rangle^3$ . Specializing the formula of Cho–Yamauchi [CY20] (extended to the unitary case in [LZ22a, Theorem 3.5.1]) gives

$$\begin{aligned} \text{Den}(X, L) &= (1 - X)(1 + qX)(1 - q^2X) + (q^3 + 1)(1 - X)X^2 \\ &= 1 - (1 - q + q^2)X + (1 - q + q^2)X^2 - X^3 \end{aligned}$$

It satisfies the functional equation

$$\text{Den}(X, L) = -X^3 \cdot \text{Den}\left(\frac{1}{X}, L\right).$$

It is easy to compute

$$\begin{aligned} \partial \text{Den}(L) &= -\frac{d}{dX} \Big|_{X=1} \text{Den}(X, L) \\ &= (1 - q + q^2) - 2(1 - q + q^2) + 3 \\ &= 2 + q - q^2. \end{aligned}$$

So combining Example 5.1.5 we obtain  $\text{Int}(L) = \partial \text{Den}(L)$  in this case! It is miraculous that the purely analytic quantity  $\partial \text{Den}(L)$  secretly knows about the Euler characteristic  $\chi(\text{DL}_1) = \text{Int}(L)$  of the Deligne–Lusztig curve  $\text{DL}_1$ .

**5.4. Strategy of the proof: uncertainty principle.** The previously known special cases of Theorem 5.3.1 were proved via explicit computation of both the geometric and analytic sides. Explicit computation seems infeasible for the general case. The proof in [LZ22a] instead proceeds via induction on  $n$  using the *uncertainty principle*, a standard tool from local harmonic analysis. Even for  $n = 2, 3$ , this proof is different from the previous proofs.

More precisely, for a fixed  $O_F$ -lattice  $L^\flat \subseteq \mathbb{V} := \mathbb{V}_n$  of rank  $n - 1$ , consider functions on  $x \in \mathbb{V} \setminus L_F^\flat$ ,

$$\text{Int}_{L^\flat}(x) := \text{Int}(L^\flat + \langle x \rangle), \quad \partial \text{Den}_{L^\flat}(x) := \partial \text{Den}(L^\flat + \langle x \rangle).$$

Then it remains to show the equality of the two functions

$$\text{Int}_{L^\flat} = \partial \text{Den}_{L^\flat}.$$

By the definition of  $\text{Int}(L)$  and  $\partial \text{Den}(L)$ , both functions are easily seen to vanish when  $x$  is non-integral, i.e.,  $\text{val}(x) < 0$ . Here  $\text{val}(x)$  denotes the valuation of the norm of  $x$ . By utilizing the inductive structure of the Rapoport–Zink spaces and local densities, it is not hard to see that if  $x \perp L^\flat$  with  $\text{val}(x) = 0$ , then

$$\text{Int}_{L^\flat}(x) = \text{Int}(L^\flat), \quad \partial \text{Den}_{L^\flat}(x) = \partial \text{Den}(L^\flat)$$

for the lattice  $L^\flat \subseteq \mathbb{V}_{n-1} \cong \langle x \rangle_F^\perp$  of full rank  $n - 1$ . Thus by induction on  $n$ , the difference function  $\phi = \text{Int}_{L^\flat} - \partial \text{Den}_{L^\flat}$  already vanishes on a large subset

$$\{x \in \mathbb{V} : x \perp L^\flat, \text{val}(x) \leq 0\}.$$

We would like to deduce that  $\phi$  indeed vanishes identically. To this end, we apply the following Uncertainty Principle.

**Proposition 5.4.1** (Uncertainty Principle, [LZ22a, Proposition 8.1.6]). *Let  $\phi \in \mathcal{S}(\mathbb{V})$  be a Schwartz function on  $\mathbb{V}$ . If both  $\phi$  and its Fourier transform  $\hat{\phi}$  vanish on  $\{x \in \mathbb{V} : \text{val}(x) \leq 0\}$ . Then  $\phi = 0$ .*

In other words,  $\phi, \hat{\phi}$  cannot simultaneously have “small support” unless  $\phi = 0$ . Applying the Uncertainty Principle to the difference function  $\phi$ , then we can finish the proof as long as we get a good control over the support of  $\hat{\phi}$ . However, both functions  $\text{Int}_{L^b}, \partial\text{Den}_{L^b}$  have *singularities* along the hyperplane  $L_F^b := L^b \otimes_{O_F} F \subseteq \mathbb{V}$ . Intuitively, if  $x$  is “closer” to  $L^b$ , then  $\mathcal{Z}(x)$  and  $\mathcal{Z}(L^b)$  intersect more “improperly”, which results in the blow-up of  $\text{Int}_{L^b}(x)$  along  $L_F^b$ . These singularities cause trouble in computing the Fourier transforms or even in showing that  $\phi \in \mathcal{S}(\mathbb{V})$ .

**5.5. Strategy of the proof: decomposition and local modularity.** To overcome this difficulty, we isolate the singularities by decomposing

$$\text{Int}_{L^b} = \text{Int}_{L^b, \mathcal{H}} + \text{Int}_{L^b, \mathcal{V}}, \quad \partial\text{Den}_{L^b} = \partial\text{Den}_{L^b, \mathcal{H}} + \partial\text{Den}_{L^b, \mathcal{V}}$$

into “horizontal” and “vertical” parts. Here on the geometric side  $\text{Int}_{L^b, \mathcal{H}}$  (resp.  $\text{Int}_{L^b, \mathcal{V}}$ ) is the contribution from the horizontal (resp. vertical part) of the KR cycles, illustrated by the red (resp. blue) part in Figure 5. One can hope to understand the horizontal part explicitly using deformation theory, and the vertical part using algebraic geometry over the residue field  $\bar{k}$ .

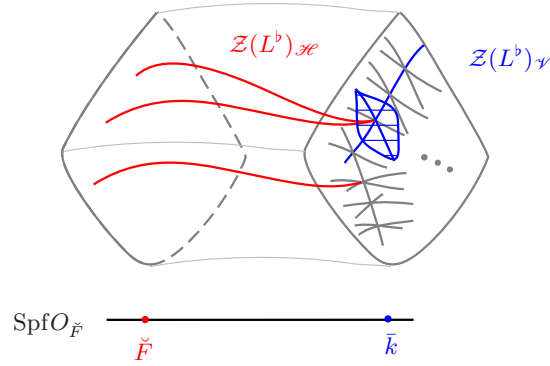


FIGURE 5. Decomposition

By the uncertainty principle and the induction on  $n$ , it remains to prove the following.

**Theorem 5.5.1** (Key Theorem).

$$(KR1) \quad \text{Int}_{L^b, \mathcal{H}} = \partial\text{Den}_{L^b, \mathcal{H}}.$$

$$(KR2) \quad \text{Int}_{L^b, \mathcal{V}} \in \mathcal{S}(\mathbb{V}) \quad \text{and} \quad \widehat{\text{Int}}_{L^b, \mathcal{V}} = -\text{Int}_{L^b, \mathcal{V}}.$$

$$(KR3) \quad \partial\text{Den}_{L^b, \mathcal{V}} \in \mathcal{S}(\mathbb{V}) \quad \text{and} \quad \widehat{\partial\text{Den}}_{L^b, \mathcal{V}} = -\partial\text{Den}_{L^b, \mathcal{V}}.$$

In other words, the horizontal part  $\partial\text{Den}_{L^b, \mathcal{H}}$  matches  $\text{Int}_{L^b, \mathcal{H}}$ , and subtracting the horizontal parts removes the singularities along  $L_F^b$  so that vertical parts are indeed in  $\mathcal{S}(\mathbb{V})$ . The extra invariance under Fourier transform

$$(5.5.1.1) \quad \widehat{\text{Int}}_{L^b, \mathcal{V}} = -\text{Int}_{L^b, \mathcal{V}}, \quad \widehat{\partial\text{Den}}_{L^b, \mathcal{V}} = -\partial\text{Den}_{L^b, \mathcal{V}}.$$

can be thought of as a *local modularity*, by analogy with the global modularity of arithmetic generating series (such as in Bruinier–Howard–Kudla–Rapoport–Yang [BHK<sup>+</sup>20] discussed in Remark 3.5.5 (v)) encoding an extra global  $\mathrm{SL}_2$ -symmetry.

**5.6. Ingredients of the Key Theorem.** Some key ingredients of the proof of the Key Theorem 5.5.1 include:

- (KR1) We describe explicitly the horizontal part of KR cycles in terms of Gross’s quasi-canonical liftings [Gro86], using the work of Tate, Grothendieck–Messing and Breuil on the deformation theory of  $p$ -divisible groups.
- (KR2) On the geometric side we show (5.5.1.1) by reducing to the case of intersection with Deligne–Lusztig curves  $\mathrm{DL}_1$ . This reduction requires the Bruhat–Tits stratification of  $\mathcal{N}^{\mathrm{red}}$  into the Deligne–Lusztig varieties  $\mathrm{DL}_i$ , as discussed in §5.1, and the Tate conjecture [Tat94] for these Deligne–Lusztig varieties. We prove the latter by reducing to a cohomological computation of Lusztig [Lus76].
- (KR3) On the analytic side we show (5.5.1.1) using Cho–Yamauchi’s explicit formula [CY20] for  $\partial\mathrm{Den}(L)$  in terms of weighted lattice counting, and reduce to a (rather subtle) lattice theoretic problem. (In fact we only show directly something weaker than (5.5.1.1) which is enough to imply Theorem 5.3.1, and we then deduce (5.5.1.1) a posteriori).

## 6. ARITHMETIC INNER PRODUCT FORMULA

In this last section we discuss an application of the arithmetic Siegel–Weil formula to the Beilinson–Bloch conjecture and the arithmetic inner product formula.

**6.1. Birch and Swinnerton-Dyer conjecture.** One long-standing problem in number theory is the determination of the rational points  $E(\mathbb{Q})$  for an elliptic curve  $E : y^2 = x^3 + ax + b$  defined over  $\mathbb{Q}$ . The celebrated *Birch and Swinnerton-Dyer (BSD) conjecture* predicts a deep link between  $E(\mathbb{Q})$  and its  $L$ -function  $L(E, s)$ . Define the *algebraic rank*

$$r_{\mathrm{alg}}(E) := \mathrm{rank} E(\mathbb{Q})$$

to be the rank of the finitely generated abelian group  $E(\mathbb{Q})$ . Define the *analytic rank*

$$r_{\mathrm{an}}(E) := \mathrm{ord}_{s=1} L(E, s)$$

to be the order of vanishing of  $L(E, s)$  at the central point  $s = 1$ . The BSD conjecture predicts the rank equality between these two notions of ranks of seemingly different nature,

$$(6.1.0.2) \quad r_{\mathrm{an}}(E) \stackrel{?}{=} r_{\mathrm{alg}}(E),$$

and a refined BSD formula

$$(6.1.0.3) \quad \frac{L^{(r)}(E, 1)}{r!} \stackrel{?}{=} \frac{\Omega(E)R(E) \prod_p c_p(E) \cdot |\mathrm{III}(E)|}{|E(\mathbb{Q})_{\mathrm{tor}}|^2}$$

for the Taylor expansion of  $L(E/\mathbb{Q}, s)$  at  $s = 1$  (here  $r = r_{\mathrm{an}}(E)$ ) in terms of various important arithmetic invariants of  $E$ . Among these invariants are the order of the (mysterious) *Tate–Shafarevich group*  $\mathrm{III}(E)$ , and the *regulator*  $R(E) := \det(\langle P_i, P_j \rangle_{\mathrm{NT}})_{r \times r}$ , where

$$\langle \cdot, \cdot \rangle_{\mathrm{NT}} : E(\mathbb{Q}) \times E(\mathbb{Q}) \rightarrow \mathbb{R}$$

is the *Néron–Tate height pairing* and  $\{P_i\}$  is a basis of the free part of  $E(\mathbb{Q})$ .

The BSD conjecture is still widely open in general, but much progress has been made in the rank 0 or 1 case. The seminal work of Gross–Zagier [GZ86], Kolyvagin [Kol90] proved the implications

$$(6.1.0.4) \quad r_{\text{an}}(E) = 0 \Rightarrow r_{\text{alg}}(E) = 0, \quad r_{\text{an}}(E) = 1 \Rightarrow r_{\text{alg}}(E) = 1,$$

confirming (6.1.0.2) when  $r_{\text{an}}(E) \leq 1$ . Due to the work of many people, many cases of (6.1.0.3) are also known when  $r_{\text{an}}(E) \leq 1$ .

The key to relate  $r_{\text{an}}(E)$  and  $r_{\text{alg}}(E)$  is the *Gross–Zagier formula*

$$(6.1.0.5) \quad L'(E, 1) \doteq \langle P, P \rangle_{\text{NT}},$$

(up to an explicit nonzero constant, including the period of  $E$  and other rational factors) relating the first derivative of the  $L$ -function at  $s = 1$  and the Néron–Tate height of certain rational points  $P$  on  $E$  known as *Heegner points* (cf. Example 6.4.1). It gives one crucial implication in (6.1.0.4),

$$(6.1.0.6) \quad r_{\text{an}}(E) = 1 \implies r_{\text{alg}}(E) \geq 1.$$

The tools of Heegner points and  $L$ -functions, linked via the Gross–Zagier formula, are indispensable in studying the arithmetic of elliptic curves. We refer to Zhang [Zha14] for an excellent recent survey on Heegner points and the Birch–Swinnerton–Dyer conjecture (see also Gross [Gro04] and Darmon [Dar04]).

**6.2. Beilinson–Bloch conjecture.** In the 1980s, Beilinson ([Bei87, Conjecture 5.9]) and Bloch ([Blo84a, Recurring Fantasy],[Blo84b, Conjecture]) proposed vast generalizations of the BSD conjecture to higher dimensional varieties.

Let  $X$  be a smooth projective variety of dimension  $n$  over a number field  $F$ . For  $1 \leq m \leq n$ , denote by  $\text{CH}^m(X)$  the Chow group of algebraic cycles of codimension  $m$  on  $X$  defined over  $F$  (up to rational equivalence), and  $\text{CH}^m(X)^0 \subseteq \text{CH}^m(X)$  the subgroup of geometrically cohomologically trivial cycles. Denote by  $L(H^{2m-1}(X), s)$  the motivic  $L$ -function associated to the  $(2m - 1)$ -th étale cohomology  $H^{2m-1}(X_{\overline{F}}, \mathbb{Q}_\ell)$ . Then the *Beilinson–Bloch (BB) conjecture* predicts the equality between analytic and algebraic ranks

$$(6.2.0.7) \quad \text{ord}_{s=m} L(H^{2m-1}(X), s) = \text{rank } \text{CH}^m(X)^0,$$

and a refined formula

$$L^{(r)}(H^{2m-1}(X), m) \doteq \det \langle \cdot, \cdot \rangle_{\text{BB}}$$

for the leading coefficient at  $s = m$  in terms of the determinant of the Beilinson–Bloch height pairing<sup>3</sup>

$$\langle \cdot, \cdot \rangle_{\text{BB}} : \text{CH}^m(X)^0 \times \text{CH}^{n+1-m}(X)^0 \rightarrow \mathbb{R}.$$

**Example 6.2.1.** When  $X/F = E/\mathbb{Q}$  and  $m = 1$  we recover the BSD conjecture as

$$\text{CH}^1(E)^0 \simeq E(\mathbb{Q}), \quad L(H^1(E), s) = L(E, s) \text{ and } \langle \cdot, \cdot \rangle_{\text{BB}} = -\langle \cdot, \cdot \rangle_{\text{NT}}.$$

The BB conjecture is even more elusive than the BSD conjecture: in general we do not know that  $\text{CH}^m(X)^0$  is finitely generated, nor do we know that  $L(H^{2m-1}(X), s)$  can be analytically continued to the central point  $s = m$ , so neither side of (6.2.0.7) is well-defined! This may be more an exciting challenge than disappointment for mathematicians — after all we were in a similar circumstance when BSD conjecture was formulated in the 1960s: we knew neither the analytic continuation of

<sup>3</sup>When  $m > 1$ , the Beilinson–Bloch height pairing is only defined assuming certain conjectures on algebraic cycles on  $X$  (see [Bei87, Conjectures 2.2.1 and 2.2.3]). This important technical issue is addressed in [LL21, LL22] so that the right-hand-side of (6.7.1.1) in Theorem 6.7.1 can be defined unconditionally, but we will intentionally ignore it for the purpose of this article.

$L(E, s)$  (except when  $E$  has complex multiplication) nor the finiteness of  $\text{III}(E)$  in order to make sense of either side of (6.1.0.3).

A good testing ground for the BB conjecture is by taking  $X$  to be Shimura varieties. By Langland's philosophy, their  $L$ -functions can be computed in terms of automorphic  $L$ -functions, so the analytic rank in (6.2.0.7) becomes accessible. Even though we do not know if  $\text{CH}^m(X)^0$  is finitely generated, it still makes sense to test if it is nontrivial. It is thus tempting to relate special cycles on Shimura varieties to automorphic  $L$ -functions, and in particular, to generalize the Gross–Zagier formula (6.1.0.5) to higher dimensional Shimura varieties and prove the analogue of (6.1.0.6) towards the BB conjecture

$$(6.2.1.1) \quad \text{ord}_{s=m} L(H^{2m-1}(X), s) = 1 \implies \text{rank CH}^m(X)^0 \geq 1.$$

Here we use the notation  $\text{rank CH}^m(X)^0 \geq 1$  to stand for the non-triviality of  $\text{CH}^m(X)_{\mathbb{Q}}^0$ .

**6.3. Arithmetic inner product formula.** As discussed in §3.5, Shimura varieties  $X$  of type  $\text{GSpin}(n-1, 2)$  and  $\text{U}(n-1, 1)$  admit a rich supply of special cycles, recovering Heegner points in the case of modular curves. As explained in [Kud04, III], when  $n$  is even, the arithmetic Siegel–Weil formula together with the doubling method of Piatetski-Shapiro–Rallis [PSR84] has important application to the *arithmetic inner product formula* of the form

$$L'(1/2, \pi) \doteq \langle \Theta_{\varphi}(\phi), \Theta_{\varphi}(\phi) \rangle_{\text{BB}}.$$

Here

- $\pi$  is a certain cuspidal automorphic representation on  $\text{Mp}(n)$  or  $\text{U}(n)$ .
- $L'(1/2, \pi)$  is the central *derivative* the standard  $L$ -function  $L(s, \pi)$  of  $\pi$  (cf. [Yam14]) with global root number  $\varepsilon(\pi) = -1$ .
- $\Theta_{\varphi}(\phi)$  is an algebraic cycle on  $X$  of codimension  $n/2$ , constructed from holomorphic forms  $\phi \in \pi$  using the method of *arithmetic theta lifting* (see Definition 6.6.3).

This arithmetic inner product formula can be viewed as an arithmetic analogue of the *Rallis inner product formula* (see [GQT14]) of the form

$$(6.3.0.2) \quad L(1/2, \pi) \doteq \langle \theta_{\varphi}(\phi), \theta_{\varphi}(\phi) \rangle_{\text{Pet}}$$

which relates the central *value*  $L(1/2, \pi)$  when  $\varepsilon(\pi) = +1$  and the Petersson inner product of the (usual) theta lift  $\theta_{\varphi}(\phi)$  discussed in Remark 2.2.5. It can also be viewed as a higher dimensional generalization of the Gross–Zagier formula (the case  $n = 2$ ). Thus the arithmetic Siegel–Weil formula is intimately linked to the Birch–Swinnerton-Dyer conjecture, and more generally the Beilinson–Bloch conjecture for higher dimensional Shimura varieties.

The conjectural arithmetic inner product formula was formulated by Kudla [Kud97b] using the Gillet–Soulé height and in more generality by Liu [Liu11a] using the Beilinson–Bloch height (in the unitary case). In the unitary case the arithmetic inner product formula has been recently proved under mild local assumptions and Kudla's modularity conjecture in our works with Liu [LL21, LL22], which has been applied to prove the first unconditional theorem for the Beilinson–Bloch conjecture for higher dimensional Shimura varieties. Our remaining goal is to explain some details about the main results of [LL21, LL22].

**6.4. Arithmetic theta lifting on modular curves.** To motivate the construction of arithmetic theta lifting, let us first consider an example of Heegner points on elliptic curves.

**Example 6.4.1** (cf. [Zag85]). Consider the curve 37a1 in Cremona's table,

$$E = 37a1 : y^2 + y = x^3 - x,$$

It is the rank one optimal curve over  $\mathbb{Q}$  of smallest conductor ( $N = 37$ ). It is isomorphic to the modular curve  $X_0^+(N) = X_0(N)/\langle w_N \rangle$  for  $N = 37$  ( $N$  is the smallest such that  $X_0^+(N)$  has positive genus). The Mordell-Weil group  $E(\mathbb{Q}) \cong \mathbb{Z}$ , generated by  $P = (0, 0)$ . Let

$$f_E = \sum_{n \geq 1} a_n q^n = q - 2q^2 - 3q^3 + 2q^4 - 2q^5 + 6q^6 - q^7 + 6q^9 + 4q^{10} - 5q^{11} - 6q^{12} + \dots \in S_2(N)$$

be the weight 2 newform of level  $N = 37$  associated to  $E$ . It gives a modular parametrization

$$\varphi_E : \mathcal{H} \rightarrow E = \mathbb{C}/\Lambda, \quad \varphi_E(\tau) = \sum_{n \geq 1} \frac{a_n q^n}{n} = q - q^2 - q^3 + \frac{1}{2}q^4 - \frac{2}{5}q^5 + q^6 - \frac{1}{7}q^7 + \dots$$

such that  $\varphi_E^*(\omega_E) = 2\pi i f_E(\tau) d\tau$ . The Shimura–Waldspurger correspondence, which can be viewed as an instance of the theta correspondence for the pair  $(G, H) = (\mathrm{Mp}(2), \mathrm{O}(3))$ ,

$$\theta : S_{3/2}^+(4N) \rightarrow S_2(N),$$

gives a weight 3/2 newform in Kohnen's plus space  $S_{3/2}^+(4N)$ ,

$$g_E = \sum_{d \geq 1} c_d q^d = -q^3 - q^4 + q^7 - q^{11} + q^{12} + 2q^{16} + 3q^{27} + \dots - 6q^{67} + \dots$$

such that  $\theta(g_E) = f_E$ .

Let  $d$  be a positive integer such that  $-d \equiv 0, 1 \pmod{4}$ . If  $N = 37$  splits in  $K = \mathbb{Q}(\sqrt{-d})$ , then one can construct a rational point  $P_d \in E(\mathbb{Q})$  using Heegner points on  $X_0(N)$  associated to quadratic orders of discriminant  $-d$ . For example, when  $-d$  is a fundamental discriminant, we have a Heegner point  $y_K \in X_0(H_K)$  defined over the Hilbert class field  $H_K$  of  $K$ , and  $P_d = \mathrm{tr}_{H_K/K} \varphi_E(y_K - \infty)$ . The point  $P_d$  may depend on the choice of  $d$ , even when  $E(\mathbb{Q}) \cong \mathbb{Z}$ . In Table 3, we compute a list of Heegner points  $P_d \in E(\mathbb{Q})$  for small  $d$ 's, and also the integers multiples  $n_d$  such that  $P_d = n_d \cdot P$  for the generator  $P = (0, 0)$ .

$d$	3	4	7	11	12	16	27	...	67	...
$P_d$	(0, -1)	(0, -1)	(0, 0)	(0, -1)	(0, 0)	(1, 0)	(-1, -1)	...	(6, -15)	...
$n_d$	-1	-1	1	-1	1	2	3	...	-6	...

TABLE 3. Heegner points on  $E = X_0^+(37)$

From Table 3, we observe the miraculous coincidence that the integer  $n_d$  exactly matches the coefficient  $c_d$  of  $g_E$ ! In other words, the generating series

$$Z(\tau) = \sum_{d \geq 1} P_d q^d = g_E \cdot P \in S_{3/2}^+(4 \cdot 37) \otimes E(\mathbb{Q})$$

is a *modular form* valued in  $E(\mathbb{Q}) \otimes \mathbb{C}$ . This allows us to define the *arithmetic theta lift* by taking the Petersson inner product of  $Z(\tau)$  with  $g_E$ ,

$$\Theta(g_E) := \langle Z(\tau), g_E \rangle_{\mathrm{Pet}} = \langle g_E, g_E \rangle_{\mathrm{Pet}} \cdot P \in E(\mathbb{Q}) \otimes \mathbb{C},$$

which is now a canonical element in  $E(\mathbb{Q}) \otimes \mathbb{C}$  (i.e., no need to choose any particular  $d$ ). The arithmetic inner product formula in this case asserts the identity

$$(6.4.1.1) \quad L'(E, 1) \doteq \frac{\langle \Theta(g_E), \Theta(g_E) \rangle_{\mathrm{NT}}}{\langle g_E, g_E \rangle_{\mathrm{Pet}}}.$$

In fact, in this case we have  $\frac{\langle \Theta(g_E), \Theta(g_E) \rangle_{\text{NT}}}{\langle g_E, g_E \rangle_{\text{Pet}}} = \langle g_E, g_E \rangle_{\text{Pet}} \cdot \langle P, P \rangle_{\text{NT}}$  and we can compute explicitly

$$\begin{aligned}\langle g_E, g_E \rangle_{\text{Pet}} &= 0.7146356107 \dots = \frac{3\omega^+(E)}{4\pi}, \\ \langle P, P \rangle_{\text{NT}} &= 0.0511114082 \dots, \\ L'(E, 1) &= 0.3059997738 \dots.\end{aligned}$$

Here  $\omega^+(E)$  is the real period of  $E$ . Therefore

$$\frac{L'(E, 1)}{\langle g_E, g_E \rangle_{\text{Pet}} \cdot \langle P, P \rangle_{\text{NT}}} = 8.37758040956 \dots = \frac{8\pi}{3}$$

and the equality (6.4.1.1) indeed holds up to an elementary constant  $\frac{8\pi}{3}$ .

**6.5. Unitary Shimura varieties.** Now let us come to the setting of unitary Shimura varieties. Let  $E/F$  be a CM extension of a totally real number field. Let  $\mathbb{V}$  be a totally definite *incoherent hermitian space* over  $\mathbb{A}_E$  of rank  $n$ . Here *incoherent* means that  $\mathbb{V}$  is not the base change of a global  $E/F$ -hermitian space, or equivalently the product of the local Hasse invariants of the local hermitian spaces  $\mathbb{V}_v := \mathbb{V} \otimes_{\mathbb{A}_F} F_v$  is

$$\prod_v \varepsilon(\mathbb{V}_v) = -1.$$

Picking any place  $w | \infty$  of  $F$  gives a nearby global  $E/F$ -hermitian space  $V$  such that

$$V_v \cong \mathbb{V}_v, \text{ if } v \neq w, \quad \text{but } V_w \text{ has signature } (n-1, 1).$$

Associated to any open compact subgroup  $K \subseteq \text{U}(V)(\mathbb{A}_F^\infty)$ , we have a *unitary Shimura variety*  $X = \text{Sh}_{\text{U}(V)}$  (cf. [Zha19, Gro20]), which has a smooth canonical model of dimension  $n-1$  over  $E$  (viewed as a subfield of  $\mathbb{C}$  via the embedding induced by a place above  $w$ ) and admits complex uniformization

$$X(\mathbb{C}) = \text{U}(V)(F) \backslash [\mathbb{D} \times \text{U}(V)(\mathbb{A}_F^\infty) / K],$$

where

$$\mathbb{D} := \{z \in \mathbb{C}^{n-1} : |z| < 1\} \cong \frac{\text{U}(n-1, 1)}{\text{U}(n-1) \times \text{U}(1)}$$

is the hermitian symmetric domain associated to  $\text{U}(V_\infty)$ .

We remark that  $X$  is a Shimura variety *of abelian type* (rather than of PEL or Hodge type). Unlike Shimura varieties of PEL type associated to unitary similitude groups, it lacks a good moduli description in terms of abelian varieties with additional structures and thus it is technically more difficult to study. Nevertheless, its étale cohomology and  $L$ -function will be computed in terms of automorphic forms in the forthcoming work of Kisin–Shin–Zhu [KSZ], under the help of the endoscopic classification for unitary groups due to Mok [Mok15] and Kaletha–Minguez–Shin–White [KMSW14]. In particular, the analytic side of (6.2.1.1) indeed makes sense.

**6.6. Arithmetic theta lifting.** From now on assume that  $n = 2m$  is even. Let  $W = E^n$  be the standard  $E/F$ -skew-hermitian space with matrix  $\begin{pmatrix} 0 & 1_m \\ -1_m & 0 \end{pmatrix}$ . Its unitary group  $\text{U}(W)$  is a quasi-split unitary group of rank  $n$ . Let  $\pi$  be a cuspidal automorphic representation of  $\text{U}(W)(\mathbb{A}_F)$ .

**Assumption 6.6.1.** We impose the following (mild) local assumptions on  $E/F$  and  $\pi$ .

- (1)  $E/F$  is split at all 2-adic places and  $F \neq \mathbb{Q}$ . If  $v \nmid \infty$  is ramified in  $E$ , then  $v$  is unramified over  $\mathbb{Q}$ . Assume that  $E/\mathbb{Q}$  is Galois or contains an imaginary quadratic field (for simplicity).
- (2) For  $v | \infty$ ,  $\pi_v$  is the holomorphic discrete series with Harish-Chandra parameter  $\{\frac{n-1}{2}, \frac{n-3}{2}, \dots, \frac{-n+3}{2}, \frac{-n+1}{2}\}$ .



- (3) For  $v \nmid \infty$ ,  $\pi_v$  is tempered.
- (4) For  $v \nmid \infty$  ramified in  $E$ ,  $\pi_v$  is spherical with respect to the stabilizer of  $O_{E_v}^{2m}$ .
- (5) For  $v \nmid \infty$  inert in  $E$ ,  $\pi_v$  is unramified or almost unramified with respect to the stabilizer of  $O_{E_v}^{2m}$ . If  $\pi_v$  is almost unramified, then  $v$  is unramified over  $\mathbb{Q}$ .

**Remark 6.6.2.** In Assumption 6.6.1 (5),  $\pi_v$  is *almost unramified* means that it has a nonzero Iwahori-fixed vector and its Satake parameter contains  $\{q_v, q_v^{-1}\}$  and  $2m - 2$  complex numbers of norm 1. Equivalently, the theta lift of  $\pi_v$  to the non-quasi-split unitary group of same rank is spherical with respect to the stabilizer of an almost self-dual lattice (see [Liu21]).

Let  $S_\pi = \{v \text{ inert} : \pi_v \text{ almost unramified}\}$ . Then under Assumption 6.6.1, the global root number for the (complete) standard  $L$ -function  $L(s, \pi)$  (cf. [Yam14]) equals

$$\varepsilon(\pi) = (-1)^{|S_\pi|} \cdot (-1)^{m[F:\mathbb{Q}]}$$

by the epsilon dichotomy for unitary groups due to Harris–Kudla–Sweet [HKS96] and Gan–Ichino [GI14, Theorem 11.1]. When  $\text{ord}_{s=1/2} L(s, \pi) = 1$ , we have  $\varepsilon(\pi) = -1$ , and hence there is a canonical choice of totally definite incoherent space  $\mathbb{V} = \mathbb{V}_\pi$  of rank  $n$  such that for  $v \nmid \infty$ ,

$$\varepsilon(\mathbb{V}_v) = -1 \text{ exactly for } v \in S_\pi.$$

Let  $X$  be the associated unitary Shimura variety of dimension  $n - 1 = 2m - 1$  over  $E$ . The assumption  $F \neq \mathbb{Q}$  implies that  $X$  is projective.

**Definition 6.6.3.** Assuming Kudla’s modularity Conjecture 3.5.4, then Kudla’s generating series  $Z_\varphi(\tau)$  of codimension  $m$  special cycles is a hermitian modular form on the hermitian half plane  $\mathcal{H}_m$  valued in  $\text{CH}^m(X)_\mathbb{C}$ . For holomorphic forms  $\phi \in \pi$ , define the *arithmetic theta lift* by taking the Petersson inner product on  $\mathcal{H}_m$ ,

$$\Theta_\varphi(\phi) := (Z_\varphi(\tau), \phi)_{\text{Pet}} \in \text{CH}^m(X)_\mathbb{C},$$

Moreover, under Assumption 6.6.1,  $\Theta_\varphi(\phi)$  is in fact cohomologically trivial and lies in the  $\pi$ -isotopic Chow group  $\text{CH}^m(X)_\pi^0$  (see [LL21, Proposition 6.10]).

**6.7. Arithmetic inner product formula.** Now we are ready to state the arithmetic inner product formula for unitary Shimura varieties.

**Theorem 6.7.1** (with Liu [LL21, LL22]). *Let  $\pi$  be a cuspidal automorphic representation of  $\text{U}(W)(\mathbb{A}_F)$  satisfying Assumption 6.6.1. Assume that  $\varepsilon(\pi) = -1$ . Assume that Kudla’s modularity Conjecture 3.5.4 holds. Then for any holomorphic  $\phi \in \pi$  and  $\varphi \in \mathcal{S}(\mathbb{V}(\mathbb{A}_F^\infty)^m)$ , the following identity holds (up to simpler factors depending on  $\phi$  and  $\varphi$ ),*

$$(6.7.1.1) \quad L'(1/2, \pi) \doteq \langle \Theta_\varphi(\phi), \Theta_\varphi(\phi) \rangle_{\text{BB}}.$$

**Remark 6.7.2.** The simpler factors can further be made explicit. For example, if

- $\pi$  is unramified or almost unramified at all finite places,
- $\phi \in \pi$  is a holomorphic newform such that  $(\phi, \bar{\phi})_\pi = 1$ ,
- $\varphi_v$  is the characteristic function of self-dual or almost self-dual lattices at all finite places  $v$ .

Then we have

$$(6.7.2.1) \quad \frac{L'(1/2, \pi)}{\prod_{i=1}^{2m} L(i, \eta_{E/F}^i)} C_m^{[F:\mathbb{Q}]} \prod_{v \in S_\pi} \frac{q_v^{m-1}(q_v + 1)}{(q_v^{2m-1} + 1)(q_v^{2m} - 1)} = (-1)^m \langle \Theta_\varphi(\phi), \Theta_\varphi(\phi) \rangle_{\text{BB}},$$

where  $C_m = 2^{-2m} \pi^{m^2} \frac{\Gamma(1) \cdots \Gamma(m)}{\Gamma(m+1) \cdots \Gamma(2m)}$ .

Notice that the Grand Riemann Hypothesis predicts that  $L'(1/2, \pi) \geq 0$ , while Beilinson's Hodge index conjecture ([Bei87, Conjecture 5.5]) predicts that  $(-1)^m \langle \Theta_\varphi(\phi), \Theta_\varphi(\phi) \rangle_{\text{BB}} \geq 0$ . It is a good reality check that these two (big) conjectures are compatible with (6.7.2.1).

Without assuming Kudla's modularity conjecture, we cannot define  $\Theta_\varphi(\phi)$  but we may still obtain unconditional nonvanishing results on Chow groups as predicted by the Beilinson–Bloch conjecture in (6.2.1.1).

**Theorem 6.7.3** (with Liu [LL21, LL22]). *Let  $\pi$  be a cuspidal automorphic representation of  $U(W)(\mathbb{A}_F)$  satisfying Assumption 6.6.1. Let  $\text{CH}^m(X)_{\mathfrak{m}_\pi}^0$  the localization of  $\text{CH}^m(X)_{\mathbb{C}}^0$  at the maximal ideal  $\mathfrak{m}_\pi$  of the spherical Hecke algebra (away from all ramification) associated to  $\pi$ . Then the implication*

$$(6.7.3.1) \quad \text{ord}_{s=1/2} L(s, \pi) = 1 \implies \text{rank } \text{CH}^m(X)_{\mathfrak{m}_\pi}^0 \geq 1$$

*holds when the level subgroup  $K \subseteq U(V)(\mathbb{A}_F^\infty)$  is sufficiently small.*

**Remark 6.7.4.** The implication analogous to (6.7.3.1) was previously known for several low dimensional  $X$  including:

- modular curves (Gross–Zagier [GZ86]),
- Shimura curves (Zhang [Zha01], Kudla–Rapoport–Yang [KRY06], Yuan–Zhang–Zhang [YZZ09], Liu [Liu11b]),
- certain  $U(2) \times U(3)$  Shimura threefolds when  $\pi$  is endoscopic (Xue [Xue19]).

**6.8. Symmetric power  $L$ -functions of elliptic curves.** We illustrate Theorem 6.7.3 by an example coming from symmetric power  $L$ -functions of elliptic curves, which is particularly attractive in view of recent progress on the symmetric power functionality by Newton–Thorne [NT21].

**Example 6.8.1.** Let  $A/F$  be a modular elliptic curve without complex multiplication such that

- $\text{Sym}^{2m-1} A$  is modular.
- $A$  has bad reduction only at places  $v$  split in  $E$ .

Assume that  $E/F$  satisfies Assumption 6.6.1. Then there exists  $\pi$  satisfying Assumption 6.6.1 such that

$$L(s + 1/2, \pi) = L(\text{Sym}^{2m-1} A_E, s + m).$$

As  $S_\pi = \emptyset$  and  $\varepsilon(\pi) = (-1)^{m[F:\mathbb{Q}]}$ , Theorem 6.7.3 applies to  $\pi$  when  $m[F:\mathbb{Q}]$  is odd (e.g., when  $m = 3$  and  $F$  is a totally real cubic field).

**6.9. Summary.** We end our discussion by the analogy between geometric and arithmetic formulas in Table 4.

Automorphic/Geometric	Arithmetic
Hurwitz formula (3.1.5.1)	Gross–Keating formula (4.1.1.1)
Geometric Siegel–Weil formula (3.6.1.1) (Kudla's formula)	Arithmetic Siegel–Weil formula (4.2.2.1) (Kudla–Rapoport Conjecture)
Rallis inner product formula (6.3.0.2)	Arithmetic inner product formula (6.7.1.1) (Gross–Zagier formula in higher dimension)

TABLE 4. Summary

## REFERENCES

- [AGHMP18] Fabrizio Andreatta, Eyal Z. Goren, Benjamin Howard, and Keerthi Madapusi Pera. Faltings heights of abelian varieties with complex multiplication. *Ann. of Math. (2)*, 187(2):391–531, 2018. 18
- [BBGK07] Jan H. Bruinier, José I. Burgos Gil, and Ulf Kühn. Borcherds products and arithmetic intersection theory on Hilbert modular surfaces. *Duke Math. J.*, 139(1):1–88, 2007. 17
- [Bei87] A. A. Beilinson. Height pairing between algebraic cycles. In *K-theory, arithmetic and geometry (Moscow, 1984–1986)*, volume 1289 of *Lecture Notes in Math.*, pages 1–25. Springer, Berlin, 1987. 29, 34
- [BGKK07] J. I. Burgos Gil, J. Kramer, and U. Kühn. Cohomological arithmetic Chow rings. *J. Inst. Math. Jussieu*, 6(1):1–172, 2007. 17
- [BH21] Jan Hendrik Bruinier and Benjamin Howard. Arithmetic volumes of unitary Shimura varieties. *arXiv e-prints*, page arXiv:2105.11274, May 2021. 21
- [BHK<sup>+</sup>20] Jan H. Bruinier, Benjamin Howard, Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. Modularity of generating series of divisors on unitary Shimura varieties. *Astérisque*, (421, Diviseurs arithmétiques sur les variétés orthogonales et unitaires de Shimura):7–125, 2020. 17, 28
- [BK12a] Rolf Berndt and Ulf Kuehn. On Kudla’s Green function for signature (2,2), part I. *arXiv e-prints*, page arXiv:1205.6417, May 2012. 17
- [BK12b] Rolf Berndt and Ulf Kuehn. On Kudla’s Green function for signature (2,2) Part II. *arXiv e-prints*, page arXiv:1209.3949, September 2012. 17
- [Blo84a] Spencer Bloch. Algebraic cycles and values of  $L$ -functions. *J. Reine Angew. Math.*, 350:94–108, 1984. 29
- [Blo84b] Spencer Bloch. Height pairings for algebraic cycles. In *Proceedings of the Luminy conference on algebraic K-theory (Luminy, 1983)*, volume 34, pages 119–145, 1984. 29
- [Bor99] Richard E. Borcherds. The Gross-Kohnen-Zagier theorem in higher dimensions. *Duke Math. J.*, 97(2):219–233, 1999. 17
- [Bru12] Jan Hendrik Bruinier. Regularized theta lifts for orthogonal groups over totally real fields. *J. Reine Angew. Math.*, 672:177–222, 2012. 17
- [BWR15] Jan Hendrik Bruinier and Martin Westerholt-Raum. Kudla’s modularity conjecture and formal Fourier-Jacobi series. *Forum Math. Pi*, 3:e7, 30, 2015. 17
- [BY21] Jan Hendrik Bruinier and Tonghai Yang. Arithmetic degrees of special cycles and derivatives of Siegel Eisenstein series. *J. Eur. Math. Soc. (JEMS)*, 23(5):1613–1674, 2021. 21
- [BZ19] Jan Bruinier and Shaul Zemel. Special Cycles on Toroidal Compactifications of Orthogonal Shimura Varieties. *arXiv e-prints*, page arXiv:1912.11825, December 2019. 17
- [Cho20] Sungyoon Cho. Special cycles on unitary Shimura varieties with minuscule parahoric level structure. *arXiv e-prints*, page arXiv:2002.00172, February 2020. 26
- [CY20] Sungmun Cho and Takuya Yamauchi. A reformulation of the Siegel series and intersection numbers. *Math. Ann.*, 377(3-4):1757–1826, 2020. 25, 26, 28
- [Dar04] Henri Darmon. *Rational points on modular elliptic curves*, volume 101 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2004. 29
- [Dic66] Leonard Eugene Dickson. *History of the theory of numbers. Vol. II: Diophantine analysis*. Chelsea Publishing Co., New York, 1966. 2
- [DS05] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. 3, 4
- [DY19] Tuoping Du and Tonghai Yang. Arithmetic Siegel-Weil formula on  $X_0(N)$ . *Adv. Math.*, 345:702–755, 2019. 17
- [ES18] Stephan Ehlen and Siddarth Sankaran. On two arithmetic theta lifts. *Compos. Math.*, 154(10):2090–2149, 2018. 18
- [FH00] Eberhard Freitag and Carl Friedrich Hermann. Some modular varieties of low dimension. *Adv. Math.*, 152(2):203–287, 2000. 15
- [FM14] Jens Funke and John Millson. The geometric theta correspondence for Hilbert modular surfaces. *Duke Math. J.*, 163(1):65–116, 2014. 17
- [FYZ21a] Tony Feng, Zhiwei Yun, and Wei Zhang. Higher Siegel-Weil formula for unitary groups: the non-singular terms. *arXiv e-prints*, page arXiv:2103.11514, March 2021. 21

- [FYZ21b] Tony Feng, Zhiwei Yun, and Wei Zhang. Higher theta series for unitary groups over function fields. *arXiv e-prints*, page arXiv:2110.07001, October 2021. 22
- [Gan14] Wee Teck Gan. Theta correspondence: recent progress and applications. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. II*, pages 343–366. Kyung Moon Sa, Seoul, 2014. 8
- [GI14] Wee Teck Gan and Atsushi Ichino. Formal degrees and local theta correspondence. *Invent. Math.*, 195(3):509–672, 2014. 33
- [Gie83] Joseph Gierster. Ueber Relationen zwischen Klassenzahlen binärer quadratischer Formen von negativer Determinante. *Math. Ann.*, 21(1):1–50, 1883. 12
- [GK93] Benedict H. Gross and Kevin Keating. On the intersection of modular correspondences. *Invent. Math.*, 112(2):225–245, 1993. 12, 14, 19, 21, 25
- [GKZ87] B. Gross, W. Kohlen, and D. Zagier. Heegner points and derivatives of  $L$ -series. II. *Math. Ann.*, 278(1-4):497–562, 1987. 15, 17
- [GQT14] Wee Teck Gan, Yannan Qiu, and Shuichiro Takeda. The regularized Siegel-Weil formula (the second term identity) and the Rallis inner product formula. *Invent. Math.*, 198(3):739–831, 2014. 11, 30
- [Gro86] Benedict H. Gross. On canonical and quasicanonical liftings. *Invent. Math.*, 84(2):321–326, 1986. 28
- [Gro97] Benedict H. Gross. On the motive of a reductive group. *Invent. Math.*, 130(2):287–313, 1997. 24
- [Gro04] Benedict H. Gross. Heegner points and representation theory. In *Heegner points and Rankin  $L$ -series*, volume 49 of *Math. Sci. Res. Inst. Publ.*, pages 37–65. Cambridge Univ. Press, Cambridge, 2004. 29
- [Gro20] Benedict H. Gross. Incoherent definite spaces and Shimura varieties. *arXiv e-prints*, page arXiv:2005.05188, May 2020. 16, 32
- [GS90] Henri Gillet and Christophe Soulé. Arithmetic intersection theory. *Inst. Hautes Études Sci. Publ. Math.*, (72):93–174 (1991), 1990. 17
- [GS19] Luis E. Garcia and Siddarth Sankaran. Green forms and the arithmetic Siegel-Weil formula. *Invent. Math.*, 215(3):863–975, 2019. 21
- [GZ86] Benedict H. Gross and Don B. Zagier. Heegner points and derivatives of  $L$ -series. *Invent. Math.*, 84(2):225–320, 1986. 29, 34
- [Han13] Jonathan Hanke. Quadratic forms and automorphic forms. In *Quadratic and higher degree forms*, volume 31 of *Dev. Math.*, pages 109–168. Springer, New York, 2013. 9
- [Hir98] Yumiko Hironaka. Local zeta functions on Hermitian forms and its application to local densities. *J. Number Theory*, 71(1):40–64, 1998. 25
- [Hir12] Yumiko Hironaka. Spherical functions on  $U(2n)/(U(n) \times U(n))$  and Hermitian Siegel series. In *Geometry and analysis of automorphic forms of several variables*, volume 7 of *Ser. Number Theory Appl.*, pages 120–159. World Sci. Publ., Hackensack, NJ, 2012. 25
- [HKS96] Michael Harris, Stephen S. Kudla, and William J. Sweet. Theta dichotomy for unitary groups. *J. Amer. Math. Soc.*, 9(4):941–1004, 1996. 33
- [HLSY22] Qiao He, Chao Li, Yousheng Shi, and Tonghai Yang. A proof of the Kudla-Rapoport conjecture for Krämer models. *arXiv e-prints*, page arXiv:2208.07988, August 2022. 25
- [HMP20] Benjamin Howard and Keerthi Madapusi Pera. Arithmetic of Borcherds products. *Astérisque*, (421, Diviseurs arithmétiques sur les variétés orthogonales et unitaires de Shimura):187–297, 2020. 17
- [Hör14] Fritz Hörmann. *The geometric and arithmetic volume of Shimura varieties of orthogonal type*, volume 35 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2014. 21
- [How19] Benjamin Howard. Linear invariance of intersections on unitary Rapoport-Zink spaces. *Forum Math.*, 31(5):1265–1281, 2019. 24
- [HP14] Benjamin Howard and Georgios Pappas. On the supersingular locus of the  $GU(2, 2)$  Shimura variety. *Algebra Number Theory*, 8(7):1659–1699, 2014. 15
- [HSY20] Qiao He, Yousheng Shi, and Tonghai Yang. The Kudla-Rapoport conjecture at a ramified prime. *arXiv e-prints*, page arXiv:2011.01356, November 2020. 25
- [HSY22] Qiao He, Yousheng Shi, and Tonghai Yang. Kudla-Rapoport conjecture at ramified primes. *arXiv e-prints*, page arXiv:2208.07992, August 2022. 25
- [Hur85] Adolf Hurwitz. Ueber Relationen zwischen Classenzahlen binärer quadratischer Formen von negativer Determinante. *Math. Ann.*, 25(2):157–196, 1885. 12
- [HZ76] F. Hirzebruch and D. Zagier. Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus. *Invent. Math.*, 36:57–113, 1976. 15, 16, 17

- [IK16] Tamotsu Ikeda and Hidenori Katsurada. An explicit formula for the Siegel series of a quadratic form over a non-archimedean local field. *arXiv e-prints*, page arXiv:1602.06617, February 2016. 25
- [Jac29] Carl Gustav Jacob Jacobi. *Fundamenta nova theoriae functionum ellipticarum. Auctore D. Carolo Gustavo Iacobo Iacobi... sumtibus fratrum Borntæger*, 1829. 3
- [Kat99] Hidenori Katsurada. An explicit formula for Siegel series. *Amer. J. Math.*, 121(2):415–452, 1999. 25
- [Kit83] Yoshiyuki Kitaoka. A note on local densities of quadratic forms. *Nagoya Math. J.*, 92:145–152, 1983. 25
- [KM86] Stephen S. Kudla and John J. Millson. The theta correspondence and harmonic forms. I. *Math. Ann.*, 274(3):353–378, 1986. 18
- [KM90] Stephen S. Kudla and John J. Millson. Intersection numbers of cycles on locally symmetric spaces and Fourier coefficients of holomorphic modular forms in several complex variables. *Inst. Hautes Études Sci. Publ. Math.*, (71):121–172, 1990. 15, 17
- [KMSW14] Tasho Kaletha, Alberto Minguez, Sug Woo Shin, and Paul-James White. Endoscopic Classification of Representations: Inner Forms of Unitary Groups. *arXiv e-prints*, page arXiv:1409.3731, September 2014. 32
- [Kol90] V. A. Kolyvagin. Euler systems. In *The Grothendieck Festschrift, Vol. II*, volume 87 of *Progr. Math.*, pages 435–483. Birkhäuser Boston, Boston, MA, 1990. 29
- [KR88a] Stephen S. Kudla and Stephen Rallis. On the Weil-Siegel formula. *J. Reine Angew. Math.*, 387:1–68, 1988. 10
- [KR88b] Stephen S. Kudla and Stephen Rallis. On the Weil-Siegel formula. II. The isotropic convergent case. *J. Reine Angew. Math.*, 391:65–84, 1988. 10
- [KR94] Stephen S. Kudla and Stephen Rallis. A regularized Siegel-Weil formula: the first term identity. *Ann. of Math. (2)*, 140(1):1–80, 1994. 11
- [KR99] Stephen S. Kudla and Michael Rapoport. Arithmetic Hirzebruch-Zagier cycles. *J. Reine Angew. Math.*, 515:155–244, 1999. 20, 21
- [KR00a] Stephen S. Kudla and Michael Rapoport. Cycles on Siegel threefolds and derivatives of Eisenstein series. *Ann. Sci. École Norm. Sup. (4)*, 33(5):695–756, 2000. 20, 21
- [KR00b] Stephen S. Kudla and Michael Rapoport. Height pairings on Shimura curves and  $p$ -adic uniformization. *Invent. Math.*, 142(1):153–223, 2000. 21
- [KR11] Stephen S. Kudla and Michael Rapoport. Special cycles on unitary Shimura varieties I. Unramified local theory. *Invent. Math.*, 184(3):629–682, 2011. 21, 22, 25
- [KR14] Stephen S. Kudla and Michael Rapoport. Special cycles on unitary Shimura varieties II: Global theory. *J. Reine Angew. Math.*, 697:91–157, 2014. 9, 20, 22
- [Kro60] L. Kronecker. Ueber die Anzahl der verschiedenen Classen quadratischer Formen von negativer Determinante. *J. Reine Angew. Math.*, 57:248–255, 1860. 12
- [KRY99] Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. On the derivative of an Eisenstein series of weight one. *Internat. Math. Res. Notices*, (7):347–385, 1999. 21
- [KRY06] Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. *Modular forms and special cycles on Shimura curves*, volume 161 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2006. 17, 21, 34
- [KSZ] Mark Kisin, Sug Woo Shin, and Yihang Zhu. Cohomology of certain Shimura varieties of abelian type (temporary title). in preparation. 32
- [Kud96] Stephen Kudla. Notes on the local theta correspondence. *unpublished notes, available online*, 1996. <http://www.math.toronto.edu/skudla/castle.pdf>. 8, 10
- [Kud97a] Stephen S. Kudla. Algebraic cycles on Shimura varieties of orthogonal type. *Duke Math. J.*, 86(1):39–78, 1997. 15, 17, 18
- [Kud97b] Stephen S. Kudla. Central derivatives of Eisenstein series and height pairings. *Ann. of Math. (2)*, 146(3):545–646, 1997. 20, 21, 30
- [Kud02a] Stephen S. Kudla. Derivatives of Eisenstein series and arithmetic geometry. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)*, pages 173–183. Higher Ed. Press, Beijing, 2002. 18
- [Kud02b] Stephen S. Kudla. Derivatives of Eisenstein series and generating functions for arithmetic cycles. Number 276, pages 341–368. 2002. Séminaire Bourbaki, Vol. 1999/2000. 18

- [Kud04] Stephen S. Kudla. Special cycles and derivatives of Eisenstein series. In *Heegner points and Rankin L-series*, volume 49 of *Math. Sci. Res. Inst. Publ.*, pages 243–270. Cambridge Univ. Press, Cambridge, 2004. 15, 16, 17, 18, 21, 30
- [Kud19] Stephen Kudla. Remarks on generating series for special cycles. *arXiv e-prints*, page arXiv:1908.08390, August 2019. 18
- [Liu11a] Yifeng Liu. Arithmetic theta lifting and  $L$ -derivatives for unitary groups, I. *Algebra Number Theory*, 5(7):849–921, 2011. 16, 17, 21, 30
- [Liu11b] Yifeng Liu. Arithmetic theta lifting and  $L$ -derivatives for unitary groups, II. *Algebra Number Theory*, 5(7):923–1000, 2011. 34
- [Liu21] Yifeng Liu. Theta correspondence for almost unramified representations of unitary groups. *arXiv e-prints*, page arXiv:2105.14223, May 2021. 33
- [LL21] Chao Li and Yifeng Liu. Chow groups and  $L$ -derivatives of automorphic motives for unitary groups. *Ann. of Math. (2)*, 194(3):817–901, 2021. 1, 25, 29, 30, 33, 34
- [LL22] Chao Li and Yifeng Liu. Chow groups and  $L$ -derivatives of automorphic motives for unitary groups, II. *Forum of Mathematics, Pi*, 10:e5, 71pp, 2022. 1, 29, 30, 33, 34
- [Lus76] G. Lusztig. Coxeter orbits and eigenspaces of Frobenius. *Invent. Math.*, 38(2):101–159, 1976. 28
- [LZ22a] Chao Li and Wei Zhang. Kudla-Rapoport cycles and derivatives of local densities. *J. Amer. Math. Soc.*, 35(3):705–797, 2022. 1, 20, 21, 22, 24, 25, 26, 27
- [LZ22b] Chao Li and Wei Zhang. On the arithmetic Siegel-Weil formula for  $\mathrm{GSpin}$  Shimura varieties. *Invent. Math.*, 228(3):1353–1460, 2022. 1, 21, 22, 25
- [Mae19] Yota Maeda. The modularity of special cycles on orthogonal Shimura varieties over totally real fields under the Beilinson-Bloch conjecture. *arXiv e-prints*, page arXiv:1908.08063, August 2019. 18
- [Mae21] Yota Maeda. Modularity of special cycles on unitary Shimura varieties over CM-fields. *arXiv e-prints*, page arXiv:2101.09232, January 2021. 18
- [Miy89] Toshitsune Miyake. *Modular forms*. Springer-Verlag, Berlin, 1989. Translated from the Japanese by Yoshitaka Maeda. 4
- [Mok15] Chung Pang Mok. Endoscopic classification of representations of quasi-split unitary groups. *Mem. Amer. Math. Soc.*, 235(1108):vi+248, 2015. 32
- [MP16] Keerthi Madapusi Pera. Integral canonical models for spin Shimura varieties. *Compos. Math.*, 152(4):769–824, 2016. 15
- [MZ21] Andreas Mihatsch and Wei Zhang. On the Arithmetic Fundamental Lemma conjecture over a general  $p$ -adic field. *arXiv e-prints*, page arXiv:2104.02779, April 2021. 18
- [NT21] James Newton and Jack A. Thorne. Symmetric power functoriality for holomorphic modular forms, II. *Publ. Math. Inst. Hautes Études Sci.*, 134:117–152, 2021. 34
- [Pra93] Dipendra Prasad. Weil representation, Howe duality, and the theta correspondence. In *Theta functions: from the classical to the modern*, volume 1 of *CRM Proc. Lecture Notes*, pages 105–127. Amer. Math. Soc., Providence, RI, 1993. 8
- [Pra98] Dipendra Prasad. A brief survey on the theta correspondence. In *Number theory (Tiruchirapalli, 1996)*, volume 210 of *Contemp. Math.*, pages 171–193. Amer. Math. Soc., Providence, RI, 1998. 8
- [PSR84] I. Piatetski-Shapiro and S. Rallis.  $L$ -functions of automorphic forms on simple classical groups. In *Modular forms (Durham, 1983)*, Ellis Horwood Ser. Math. Appl.: Statist. Oper. Res., pages 251–261. Horwood, Chichester, 1984. 30
- [RY20] Eugenia Rosu and Dylan Yott. Generating series of a new class of orthogonal Shimura varieties. *Algebra Number Theory*, 14(10):2743–2772, 2020. 18
- [RZ96] Michael Rapoport and Thomas Zink. *Period spaces for  $p$ -divisible groups*, volume 141 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1996. 22
- [San14] Siddarth Sankaran. Unitary cycles on Shimura curves and the Shimura lift II. *Compos. Math.*, 150(12):1963–2002, 2014. 17
- [San17] Siddarth Sankaran. Improper intersections of Kudla-Rapoport divisors and Eisenstein series. *J. Inst. Math. Jussieu*, 16(5):899–945, 2017. 26
- [San20] Siddarth Sankaran. Arithmetic special cycles and Jacobi forms. *arXiv e-prints*, page arXiv:2002.03499, February 2020. 18
- [Ser73] J.-P. Serre. *A course in arithmetic*. Graduate Texts in Mathematics, No. 7. Springer-Verlag, New York-Heidelberg, 1973. Translated from the French. 6

- [Shi20] Yousheng Shi. Special cycles on unitary Shimura curves at ramified primes. *arXiv e-prints*, page arXiv:2004.07158, April 2020. 25
- [Sie35] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36(3):527–606, 1935. 5
- [Sie36] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. II. *Ann. of Math. (2)*, 37(1):230–263, 1936. 7
- [Sie37] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. III. *Ann. of Math. (2)*, 38(1):212–291, 1937. 7
- [Sie51] Carl Ludwig Siegel. Indefinite quadratische Formen und Funktionentheorie. I. *Math. Ann.*, 124:17–54, 1951. 7
- [Sie52] Carl Ludwig Siegel. Indefinite quadratische Formen und Funktionentheorie. II. *Math. Ann.*, 124:364–387, 1952. 7
- [Sou92] C. Soulé. *Lectures on Arakelov geometry*, volume 33 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1992. With the collaboration of D. Abramovich, J.-F. Burnol and J. Kramer. 17
- [SSTT19] Ananth N. Shankar, Arul Shankar, Yunqing Tang, and Salim Tayou. Exceptional jumps of Picard ranks of reductions of K3 surfaces over number fields. *arXiv e-prints*, page arXiv:1909.07473, September 2019. 18
- [Tat94] John Tate. Conjectures on algebraic cycles in  $l$ -adic cohomology. In *Motives (Seattle, WA, 1991)*, volume 55 of *Proc. Sympos. Pure Math.*, pages 71–83. Amer. Math. Soc., Providence, RI, 1994. 28
- [Ter11] Ulrich Terstiege. Intersections of arithmetic Hirzebruch-Zagier cycles. *Math. Ann.*, 349(1):161–213, 2011. 21, 25
- [Ter13] Ulrich Terstiege. Intersections of special cycles on the Shimura variety for  $\mathrm{GU}(1, 2)$ . *J. Reine Angew. Math.*, 684:113–164, 2013. 21, 24, 25
- [VGW<sup>+</sup>07] Gunther Vogel, Ulrich Görtz, Torsten Wedhorn, Volker Meusers, Eva Viehmann, Konstantin Ziegler, Stefan Wewers, Inken Vollaard, Irene I Bouw, and Michael Rapoport. Argos Seminar on Intersections of Modular Correspondences. *Astérisque*, (312):vii–xiv, 2007. Held at the University of Bonn, Bonn, 2003–2004. 19
- [VW11] Inken Vollaard and Torsten Wedhorn. The supersingular locus of the Shimura variety of  $\mathrm{GU}(1, n - 1)$  II. *Invent. Math.*, 184(3):591–627, 2011. 23
- [Wei65] André Weil. Sur la formule de Siegel dans la théorie des groupes classiques. *Acta Math.*, 113:1–87, 1965. 7, 10
- [Wei19] Fu-Tsun Wei. On derivatives of Siegel-Eisenstein series over global function fields. *J. Lond. Math. Soc. (2)*, 100(2):518–544, 2019. 22
- [Xia21] Jiacheng Xia. Some cases of Kudla’s modularity conjecture for unitary Shimura varieties. *arXiv e-prints*, page arXiv:2101.06304, January 2021. 17
- [Xue19] Hang Xue. Arithmetic theta lifts and the arithmetic Gan-Gross-Prasad conjecture for unitary groups. *Duke Math. J.*, 168(1):127–185, 2019. 34
- [Yam14] Shunsuke Yamana.  $L$ -functions and theta correspondence for classical groups. *Invent. Math.*, 196(3):651–732, 2014. 30, 33
- [YZ17] Zhiwei Yun and Wei Zhang. Shtukas and the Taylor expansion of  $L$ -functions. *Ann. of Math. (2)*, 186(3):767–911, 2017. 22
- [YZ19] Zhiwei Yun and Wei Zhang. Shtukas and the Taylor expansion of  $L$ -functions (II). *Ann. of Math. (2)*, 189(2):393–526, 2019. 22
- [YZZ09] Xinyi Yuan, Shou-Wu Zhang, and Wei Zhang. The Gross-Kohnen-Zagier theorem over totally real fields. *Compos. Math.*, 145(5):1147–1162, 2009. 16, 17, 34
- [Zag85] D. Zagier. Modular points, modular curves, modular surfaces and modular forms. In *Workshop Bonn 1984 (Bonn, 1984)*, volume 1111 of *Lecture Notes in Math.*, pages 225–248. Springer, Berlin, 1985. 31
- [Zag08] Don Zagier. Elliptic modular forms and their applications. In *The 1-2-3 of modular forms*, Universitext, pages 1–103. Springer, Berlin, 2008. 3
- [Zha01] Shou-Wu Zhang. Gross-Zagier formula for  $\mathrm{GL}_2$ . *Asian J. Math.*, 5(2):183–290, 2001. 34
- [Zha09] Wei Zhang. *Modularity of generating functions of special cycles on Shimura varieties*. ProQuest LLC, Ann Arbor, MI, 2009. Thesis (Ph.D.)—Columbia University. 17

- [Zha14] Wei Zhang. The Birch–Swinnerton-Dyer conjecture and Heegner points: a survey. In *Current developments in mathematics 2013*, pages 169–203. Int. Press, Somerville, MA, 2014. 29
- [Zha19] Shouwu Zhang. Linear forms, algebraic cycles, and derivatives of L-series. *Sci. China Math.*, 62(11):2401–2408, 2019. 16, 32
- [Zha21] Wei Zhang. Weil representation and Arithmetic Fundamental Lemma. *Ann. of Math. (2)*, 193(3):863–978, 2021. 18

COLUMBIA UNIVERSITY, DEPARTMENT OF MATHEMATICS, 2990 BROADWAY, NEW YORK, NY 10027, USA  
*E-mail address:* `chaoli@math.columbia.edu`