

COPYRIGHT NOTICE:

**Stephen S. Kudla, Michael Rapoport, and Tonghai Yang:
Modular Forms and Special Cycles on Shimura Curves**

is published by Princeton University Press and copyrighted, © 2006, by Princeton University Press. All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher, except for reading and browsing via the World Wide Web. Users are not permitted to mount this file on any network servers.

Follow links Class Use and other Permissions. For more information, send email to: permissions@pupress.princeton.edu

Chapter One

Introduction

In this monograph we study the arithmetic geometry of cycles on an arithmetic surface \mathcal{M} associated to a Shimura curve over the field of rational numbers and the modularity of certain generating series constructed from them. We consider two types of generating series, one for divisors and one for 0-cycles, valued in $\widehat{\text{CH}}^1(\mathcal{M})$ and $\widehat{\text{CH}}^2(\mathcal{M})$, the first and second arithmetic Chow groups of \mathcal{M} , respectively. We prove that the first type is a nonholomorphic elliptic modular form of weight $\frac{3}{2}$ and that the second type is a nonholomorphic Siegel modular form of genus two and weight $\frac{3}{2}$. In fact we identify the second type of series with the central derivative of an incoherent Siegel-Eisenstein series. We also relate the height pairing of a pair of $\widehat{\text{CH}}^1(\mathcal{M})$ -valued generating series to the $\widehat{\text{CH}}^2(\mathcal{M})$ -valued series by an inner product identity. As an application of these results we define an arithmetic theta lift from modular forms of weight $\frac{3}{2}$ to the Mordell-Weil space of \mathcal{M} and prove a nonvanishing criterion analogous to that of Waldspurger for the classical theta lift, involving the central derivative of the L -function.

We now give some background and a more detailed description of these results.

The modular curve $\Gamma \backslash \mathfrak{H}$, where $\mathfrak{H} = \{z \in \mathbb{C} \mid \text{Im } z > 0\}$ is the upper half plane and $\Gamma = \text{SL}_2(\mathbb{Z})$, is the first nontrivial example of a locally symmetric variety, and of a Shimura variety. It is also the host of the space of modular forms and is the moduli space of elliptic curves. Starting from this last interpretation, we see that the modular curve comes equipped with a set of special divisors, which, like the classical Heegner divisors, are the loci of elliptic curves with extra endomorphisms. More precisely, for $t \in \mathbb{Z}_{>0}$ let

$$(1.0.1) \quad Z(t) = \{(E, x) \mid x \in \text{End}(E) \text{ with } \text{tr}(x) = 0, x^2 = -t \cdot \text{id}_E\},$$

where E denotes an elliptic curve. The resulting divisor on the modular curve, which we also denote by $Z(t)$, is the set of points where the corresponding elliptic curve E admits an action of the order $\mathbb{Z}[x] = \mathbb{Z}[\sqrt{-t}]$ in the imaginary quadratic field $\mathfrak{k}_t = \mathbb{Q}(\sqrt{-t})$, i.e., E admits complex multiplication by this order. One may also interpret $Z(t)$ as the set of Γ -orbits in \mathfrak{H} which contain a fixed point of an element $\gamma \in \text{M}_2(\mathbb{Z})$ with $\text{tr}(\gamma) = 0$ and

$\det(\gamma) = t$.

It is a classical fact that the degree of $Z(t)$ is given by $\deg Z(t) = H(4t)$, where $H(n)$ is the Hurwitz class number. It is also known that the generating series

$$(1.0.2) \quad \sum_{t>0} \deg Z(t) q^t = \sum_{t>0} H(4t) q^t$$

is *nearly* the q -expansion of a modular form. In fact, Zagier [58] showed that the complete series, for $\tau = u + iv$,

$$(1.0.3) \quad \mathcal{E}\left(\tau, \frac{1}{2}\right) = -\frac{1}{12} + \sum_{t>0} H(4t) q^t + \sum_{n \in \mathbb{Z}} \frac{1}{8\pi} v^{-\frac{1}{2}} \int_1^\infty e^{-4\pi n^2 vr} r^{-\frac{3}{2}} dr \cdot q^{-n^2},$$

is the q -expansion of the value at $s = \frac{1}{2}$ of a nonholomorphic Eisenstein series $\mathcal{E}(\tau, s)$ of weight $\frac{3}{2}$, and hence is a modular form.

Generating series of this kind have a long and rich history. They are all modeled on the classical theta series. Recall that if (L, Q) is a positive definite quadratic \mathbb{Z} -module of rank n , one associates to it the generating series

$$(1.0.4) \quad \theta_L(\tau) = \sum_{x \in L} q^{Q(x)} = 1 + \sum_{t=1}^\infty r_L(t) q^t.$$

Here

$$(1.0.5) \quad r_L(t) = |\{x \in L \mid Q(x) = t\}|,$$

and we have set, as elsewhere in this book, $q = e(\tau) = e^{2\pi i \tau}$. It is a classical result going back to the 19th century that θ_L is the q -expansion of a holomorphic modular form of weight $\frac{n}{2}$ for some congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Similarly, Siegel considered generating series of the form

$$(1.0.6) \quad \theta_r(\tau, L) = \sum_{x \in L^r} q^{Q(x)} = \sum_{T \in \mathrm{Sym}_r(\mathbb{Z})^\vee} r_L(T) q^T,$$

where $\tau \in \mathfrak{H}_r$, and $q^T = e(\mathrm{tr}(T\tau))$, and

$$(1.0.7) \quad r_L(\tau) = |\{x \in L^r \mid Q(x) = \frac{1}{2}((x_i, x_j)) = T\}|.$$

He showed that they define Siegel modular forms of genus r and weight $\frac{n}{2}$. Generalizations to indefinite quadratic forms were considered by Hecke and Siegel, and the resulting generating series can be nonholomorphic modular forms. Hirzebruch and Zagier [20] constructed generating series whose

coefficients are given by cohomology classes of special curves on Hilbert-Blumenthal surfaces. They prove that the image under any linear functional of this generating series is an elliptic modular form. For example, they identify the modular form arising via the cup product with the Kähler class as an explicitly given Eisenstein series. One can also define special 0-cycles on Hilbert-Blumenthal surfaces and make generating functions for their degrees [25]. These can be shown to be Siegel modular forms of genus two and weight 2.

We now turn to the generating series associated to arithmetic cycles on Shimura curves. We exclude the modular curve to avoid problems caused by its noncompactness. It should be pointed out, however, that all our results should have suitable analogues for the modular curve, cf. [57]. We pay, however, a price for assuming compactness. New difficulties arise due to bad reduction and to the absence of natural modular forms.

Let B be an indefinite quaternion division algebra over \mathbb{Q} , so that

$$(1.0.8) \quad B \otimes_{\mathbb{Q}} \mathbb{R} \simeq M_2(\mathbb{R}) \quad \text{and} \quad D(B) \stackrel{\text{def.}}{=} \prod_{\substack{B \otimes_{\mathbb{Q}} \mathbb{Q}_p \\ \text{division}}} p > 1.$$

Let

$$(1.0.9) \quad V = \{x \in B \mid \text{tr}(x) = 0\},$$

with quadratic form $Q(x) = \text{Nm}(x) = -x^2$, where tr and Nm denote the reduced trace and norm on B respectively. Then V is a quadratic space over \mathbb{Q} of signature type $(1, 2)$. Let

$$(1.0.10) \quad D = \{w \in V(\mathbb{C}) \mid (w, w) = 0, (w, \bar{w}) < 0\} / \mathbb{C}^{\times},$$

where $(x, y) = Q(x+y) - Q(x) - Q(y)$ is the bilinear form associated to the quadratic form Q . Then D is an open subset of a quadric in $\mathbb{P}(V(\mathbb{C})) \simeq \mathbb{P}^2$, and $(B \otimes_{\mathbb{Q}} \mathbb{R})^{\times}$ acts on $V(\mathbb{R})$ and D by conjugation. We fix a maximal order O_B in B . Since all these maximal orders are conjugate, this is not really an additional datum. Set $\Gamma = O_B^{\times}$. The Shimura curve associated to B is the quotient

$$(1.0.11) \quad [\Gamma \backslash D].$$

Since Γ does not act freely, the quotient here is to be interpreted as an orbifold.

Let us fix an isomorphism $B \otimes_{\mathbb{Q}} \mathbb{R} = M_2(\mathbb{R})$. Then we can also identify $(B \otimes_{\mathbb{Q}} \mathbb{R})^{\times} = \text{GL}_2(\mathbb{R})$ and $(B \otimes_{\mathbb{Q}} \mathbb{R})^{\times}$ acts on $\mathfrak{H}_{\pm} = \mathbb{C} \setminus \mathbb{R}$ by fractional linear transformations. We obtain an identification

$$(1.0.12) \quad D = \mathbb{C} \setminus \mathbb{R}, \quad \text{via} \quad \begin{pmatrix} z & -z^2 \\ 1 & -z \end{pmatrix} \mapsto z,$$

equivariant for the action of $(B \otimes_{\mathbb{Q}} \mathbb{R})^{\times} = \mathrm{GL}_2(\mathbb{R})$.

The Shimura curve associated to B has a modular interpretation. Namely, consider the moduli problem \mathcal{M} which associates to a scheme S over $\mathrm{Spec} \mathbb{Z}$ the category of pairs (A, ι) where

- A is an abelian scheme over S
- $\iota : O_B \rightarrow \mathrm{End}(A)$ is an action of O_B on A with characteristic polynomial

$$\mathrm{charpol}(x \mid \mathrm{Lie} A) = (T - x)(T - x^{\iota}) \in \mathcal{O}_S[T],$$

for the induced O_B -action on the Lie algebra.

Here $x \mapsto x^{\iota}$ denotes the main involution on B . If S is a scheme in characteristic zero, then the last condition simply says that A has dimension 2, i.e., that (A, ι) is a *fake elliptic curve* in the sense of Serre. This moduli problem is representable by an algebraic stack in the sense of Deligne-Mumford, and we denote the representing stack by the same symbol \mathcal{M} . We therefore have an isomorphism of orbifolds,

$$(1.0.13) \quad \mathcal{M}(\mathbb{C}) = [\Gamma \backslash D].$$

Since B is a division quaternion algebra, \mathcal{M} is proper over $\mathrm{Spec} \mathbb{Z}$ and $\mathcal{M}(\mathbb{C})$ is a compact Riemann surface (when we neglect the orbifold aspect). By its very definition, the stack \mathcal{M} is an integral model of the orbifold $[\Gamma \backslash D]$. It turns out that \mathcal{M} is smooth over $\mathrm{Spec} \mathbb{Z}[D(B)^{-1}]$ but has bad reduction at the prime divisors of $D(B)$. At the primes p with $p \mid D(B)$, the stack \mathcal{M} has semistable reduction and, in fact, admits a p -adic uniformization by the Drinfeld upper half plane $\hat{\Omega}$. In particular, the special fiber \mathcal{M}_p is connected but in general not irreducible.

In analogy with the case of the modular curve, we can define special divisors on the Shimura curve by considering complex multiplication points. More precisely, let $t \in \mathbb{Z}_{>0}$ and introduce a relative DM-stack $\mathcal{Z}(t)$ over \mathcal{M} by posing the following moduli problem. To a scheme S the moduli problem $\mathcal{Z}(t)$ associates the category of triples (A, ι, x) , where

- (A, ι) is an object of $\mathcal{M}(S)$
- $x \in \mathrm{End}(A, \iota)$ is an endomorphism such that $\mathrm{tr}(x) = 0$, $x^2 = -t \cdot \mathrm{id}_A$.

An endomorphism as above is called a *special endomorphism* of (A, ι) . The space $V(A, \iota)$ of special endomorphisms is equipped with the degree form $Q(x) = x^{\iota}x$. Note that for $x \in V(A, \iota)$ we have $Q(x) = -x^2$. We denote by the same symbol the image of $\mathcal{Z}(t)$ as a cycle in \mathcal{M} and use the

notation $Z(t) = \mathcal{Z}(t)_{\mathbb{C}}$ for its complex fiber. Note that $Z(t)$ is a finite set of points on the Shimura curve, corresponding to those fake elliptic curves which admit complex multiplication by the order $\mathbb{Z}[\sqrt{-t}]$. We form the generating series

$$(1.0.14) \quad \phi_1(\tau) = -\text{vol}(\mathcal{M}(\mathbb{C})) + \sum_{t>0} \deg(Z(t)) q^t \in \mathbb{C}[[q]].$$

Here the motivation for the constant term is as follows. Purely formally $\mathcal{Z}(0)$ is equal to \mathcal{M} with associated cohomology class in degree zero; to obtain a cohomology class in the correct degree, one forms the cup product with the natural Kähler class — which comes down to taking (up to sign) the volume of $\mathcal{M}(\mathbb{C})$ with respect to the hyperbolic volume element.

Proposition 1.0.1. *The series $\phi_1(\tau)$ is the q -expansion of a holomorphic modular form of weight $3/2$ and level $\Gamma_0(4D(B)_o)$, where $D(B)_o = D(B)$ if $D(B)$ is odd and $D(B)_o = D(B)/2$ if $D(B)$ is even.*

Just as with the theorem of Hirzebruch and Zagier, this is not proved by checking the functional equations that a modular form has to satisfy. Rather, the theorem is proved by identifying the series $\phi_1(\tau)$ with a specific Eisenstein series¹. More precisely, for $\tau = u + iv \in \mathfrak{H}$, set

$$(1.0.15) \quad \mathcal{E}_1(\tau, s, B) = v^{\frac{1}{2}(s-\frac{1}{2})} \cdot \sum_{\gamma \in \Gamma'_\infty \backslash \Gamma'} (c\tau + d)^{-\frac{3}{2}} |c\tau + d|^{-(s-\frac{1}{2})} \Phi^B(\gamma, s),$$

where $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma' = \text{SL}_2(\mathbb{Z})$, and $\Phi^B(\gamma, s)$ is a certain function depending on B . The Eisenstein series $\mathcal{E}_1(\tau, s, B)$ is the analogue for the Shimura curve of Zagier's Eisenstein series (1.0.3). It has a functional equation of the form

$$(1.0.16) \quad \mathcal{E}_1(\tau, s, B) = \mathcal{E}_1(\tau, -s, B).$$

Its value at $s = \frac{1}{2}$ is a modular form of weight $\frac{3}{2}$ and we may consider its q -expansion. Proposition 1.0.1 now follows from the following more precise result.

Proposition 1.0.2.

$$\phi_1(\tau) = \mathcal{E}_1(\tau, \frac{1}{2}, B),$$

i.e., ϕ_1 is the q -expansion of $\mathcal{E}_1(\tau, \frac{1}{2}, B)$.

¹Alternatively, $\phi_1(\tau)$ can be obtained by calculating the integral over $\mathcal{M}(\mathbb{C})$ of a theta function valued in $(1, 1)$ forms; this amounts to a very special case of the results of [33]. The analogous computation in the case of modular curves was done by Funke [11].

Proposition 1.0.2 is proved in [38] by calculating the coefficients of both power series explicitly and comparing them term by term. These coefficients turn out to be generalized class numbers. More precisely, for $t > 0$, the coefficient of q^t on either side is equal to

$$(1.0.17) \quad \deg Z(t) = 2\delta(d; D(B))H_0(t; D(B)),$$

where

$$(1.0.18) \quad \delta(d; D) = \prod_{\ell|D} (1 - \chi_d(\ell))$$

and

$$(1.0.19) \quad H_0(t; D) = \frac{h(d)}{w(d)} \sum_{\substack{c|n \\ (c,D)=1}} c \prod_{\ell|c} (1 - \chi_d(\ell)\ell^{-1}).$$

Here d denotes the fundamental discriminant of the imaginary quadratic field $k_t = \mathbb{Q}(\sqrt{-t})$ and we have written $4t = n^2d$; also, $h(d)$ denotes the class number of k_t and $w(d)$ the number of roots of unity contained in k_t . By χ_d we denote the quadratic residue character mod d . For $t = 0$, the identity in Proposition 1.0.2 reduces to the well-known formula for the volume

$$(1.0.20) \quad \text{vol}(\mathcal{M}(\mathbb{C})) = \zeta_{D(B)}(-1),$$

where in $\zeta_{D(B)}(s)$ the index means that the Euler factors for $p \mid D(B)$ have been omitted in the Riemann zeta function. Note that the fact that the generating series $\phi_1(\tau)$ is a modular form reveals some surprising and highly nonobvious coherence among the degrees of the various special cycles $Z(t)$.

In this book we will establish arithmetic analogues of Propositions 1.0.1 and 1.0.2. In contrast to the above propositions, which are statements about generating series valued in cohomology (just as was the case with the results of Hirzebruch-Zagier), our generating series will have coefficients in the arithmetic Chow groups of Gillet-Soulé [14], [48], (see also [3]). Let us recall briefly their definition in our case.

A divisor on \mathcal{M} is an element of the free abelian group generated by the closed irreducible reduced substacks which are, locally for the étale topology, Cartier divisors. A Green function for the divisor \mathcal{Z} is a function g on $\mathcal{M}(\mathbb{C})$ with logarithmic growth along the complex points of $Z = \mathcal{Z}_{\mathbb{C}}$ and which satisfies the Green equation of currents on $\mathcal{M}(\mathbb{C})$,

$$(1.0.21) \quad dd^c g + \delta_Z = [\eta],$$

where η is a smooth $(1, 1)$ -form. Let $\hat{Z}_{\mathbb{Z}}(\mathcal{M})$ be the group of pairs (\mathcal{Z}, g) , where g is a Green function for the divisor \mathcal{Z} . The first arithmetic Chow group $\widehat{\text{CH}}_{\mathbb{Z}}^1(\mathcal{M})$ is the factor group of $\hat{Z}_{\mathbb{Z}}(\mathcal{M})$ by the subgroup generated by the Arakelov principal divisors $\widehat{\text{div}} f$ associated to rational functions on \mathcal{M} . For us it will be more convenient to work instead with the \mathbb{R} -linear version $\widehat{\text{CH}}^1(\mathcal{M})$. In its definition one replaces \mathbb{Z} -linear combinations of divisors by \mathbb{R} -linear combinations and divides out by the \mathbb{R} -subspace generated by the Arakelov principal divisors. Such groups were introduced by Gillet-Soulé [15]; for the case relevant to us, see [3]. Note that restriction to the generic fiber defines the degree map

$$(1.0.22) \quad \text{deg}_{\mathbb{Q}} : \widehat{\text{CH}}^1(\mathcal{M}) \longrightarrow \text{CH}^1(\mathcal{M}_{\mathbb{C}}) \otimes \mathbb{R} \xrightarrow{\sim} \mathbb{R}.$$

The group $\widehat{\text{CH}}^2(\mathcal{M})$ is defined in an analogous way, starting with 0-cycles on \mathcal{M} . Since the fibers of \mathcal{M} over $\text{Spec } \mathbb{Z}$ are geometrically connected of dimension 1, the arithmetic degree map yields an isomorphism

$$(1.0.23) \quad \widehat{\text{deg}} : \widehat{\text{CH}}^2(\mathcal{M}) \xrightarrow{\sim} \mathbb{R}.$$

Finally we mention the Gillet-Soulé arithmetic intersection pairing,

$$(1.0.24) \quad \langle , \rangle : \widehat{\text{CH}}^1(\mathcal{M}) \times \widehat{\text{CH}}^1(\mathcal{M}) \longrightarrow \widehat{\text{CH}}^2(\mathcal{M}) = \mathbb{R}.$$

It will play the role of the cup product in cohomology in this context.

We now define a generating series with coefficients in $\widehat{\text{CH}}^1(\mathcal{M})$ using the divisors $\mathcal{Z}(t)$. For $t > 0$, we equip the divisor $\mathcal{Z}(t)$ with the Green function $\Xi(t, v)$ depending on a parameter $v \in \mathbb{R}_{>0}$, constructed in [24]. Let $\hat{Z}(t, v)$ be the corresponding class in $\widehat{\text{CH}}^1(\mathcal{M})$. For $t < 0$ note that $\mathcal{Z}(t) = \emptyset$. However, the function $\Xi(t, v)$ is still defined and is smooth for $t < 0$, hence it is a Green function for the trivial divisor, and we may define again $\hat{Z}(t, v)$ to be the class of $(\mathcal{Z}(t), \Xi(t, v)) = (0, \Xi(t, v))$. To define $\hat{Z}(0, v)$, we take our lead from the justification of the absolute term in the generating series (1.0.14).

Let ω be the Hodge line bundle on \mathcal{M} , i.e., the determinant bundle of the dual of the relative Lie algebra of the universal family (\mathcal{A}, ι) over \mathcal{M} ,

$$(1.0.25) \quad \omega = \wedge^2(\text{Lie } \mathcal{A})^*.$$

The complex fiber of this line bundle comes equipped with a natural metric. This metric is well defined up to scaling.² We denote by $\hat{\omega}$ the class of this

²The normalization of the metric we use differs from the standard normalization.

metrized line bundle under the natural map from $\widehat{\text{Pic}}(\mathcal{M})$ to $\widehat{\text{CH}}^1(\mathcal{M})$ and set

$$(1.0.26) \quad \hat{\mathcal{Z}}(0, v) = -\hat{\omega} - (0, \log(v) + \mathbf{c}),$$

where \mathbf{c} is a suitable constant.

The DM-stack $\mathcal{Z}(t)$ is finite and unramified over \mathcal{M} . It is finite and flat, i.e., a relative divisor, over $\text{Spec } \mathbb{Z}[D(B)^{-1}]$ but may contain irreducible components of the special fiber \mathcal{M}_p when $p \mid D(B)$. This integral extension of the 0-cycles $Z(t)$ is therefore sometimes different from the extension obtained by flat closure in \mathcal{M} . Its nonflatness depends in a subtle way on the p -adic valuation of t . Our definition of $Z(t)$ is a consequence of our insistence on a thoroughly modular treatment of our special cycles, which is essential to our method. We strongly suspect that in fact the closure definition does not lead to (variants of) our main theorems and that therefore our definition is the ‘right one’. We do not know this for sure since the closure definition is hard to work with.

We form the generating series,

$$(1.0.27) \quad \hat{\phi}_1 = \sum_{t \in \mathbb{Z}} \hat{\mathcal{Z}}(t, v) q^t \in \widehat{\text{CH}}^1(\mathcal{M})[[q^{\pm 1}]],$$

where the coefficients depend on the parameter $v \in \mathbb{R}_{>0}$ via the Green function $\Xi(t, v)$. The first main result of this book, proved in Chapter 4, may now be formulated as follows:

Theorem A. *For $\tau = u + iv$, $\hat{\phi}_1(\tau)$ is a (nonholomorphic) modular form of weight $\frac{3}{2}$ and level $\Gamma_0(4D(B)_o)$ with values in $\widehat{\text{CH}}^1(\mathcal{M})$.*

To explain the meaning of the statement of the theorem, recall that the \mathbb{R} -version $\widehat{\text{CH}}^1(\mathcal{M})$ of the arithmetic Chow group splits canonically into a direct sum of a finite-dimensional \mathbb{C} -vector space $\widehat{\text{CH}}^1(\mathcal{M}, \mu)$, the classical Arakelov Chow group with respect to the hyperbolic metric, and the vector space $C^\infty(\mathcal{M}(\mathbb{C}))_0$ of smooth functions on $\mathcal{M}(\mathbb{C})$ orthogonal to the constant functions. Correspondingly, the series $\hat{\phi}_1$ is the sum of a series $\hat{\phi}_1^0$ in q with coefficients in $\widehat{\text{CH}}^1(\mathcal{M}, \mu)$ and a series $\hat{\phi}_1^\infty$ in q with coefficients in $C^\infty(\mathcal{M}(\mathbb{C}))_0$. The assertion of the theorem should be interpreted as follows. There is a smooth function on \mathfrak{H} with values in the finite-dimensional vector space $\widehat{\text{CH}}^1(\mathcal{M}, \mu)$ which satisfies the usual transformation law for a modular form of weight $\frac{3}{2}$ and of level $\Gamma_0(4D(B)_o)$ whose q -expansion is equal to $\hat{\phi}_1^0$, and there is a smooth function on $\mathfrak{H} \times \mathcal{M}(\mathbb{C})$ which satisfies the usual transformation law for a modular form of weight $\frac{3}{2}$ and of level

$\Gamma_0(4D(B)_o)$ in the first variable and whose q -expansion in the first variable is equal to $\hat{\phi}_1^\infty$. Obviously, the series $\hat{\phi}_1^0$ satisfies the above condition if for any linear form $\ell : \widehat{\text{CH}}^1(\mathcal{M}, \mu) \rightarrow \mathbb{C}$ the series $\ell(\hat{\phi}_1)$ with coefficients in \mathbb{C} is a nonholomorphic modular form of weight $\frac{3}{2}$ and level $\Gamma_0(4D(B)_o)$ in the usual sense.

Let us explain briefly what is involved in the proof of Theorem A. The structure of $\widehat{\text{CH}}^1(\mathcal{M}, \mu)$ is encapsulated in the following direct sum decomposition

$$(1.0.28) \quad \widehat{\text{CH}}^1(\mathcal{M}, \mu) = \widetilde{\text{MW}} \oplus \mathbb{R}\hat{\omega} \oplus \text{Vert}.$$

Here

$$(1.0.29) \quad \widetilde{\text{MW}} \simeq \text{MW}(\mathcal{M}_{\mathbb{Q}}) := \text{Pic}^0(\mathcal{M}_{\mathbb{Q}})(\mathbb{Q}) \otimes \mathbb{R}$$

is the orthogonal complement to $(\mathbb{R}\hat{\omega} \oplus \text{Vert})$, and the subspace Vert is spanned by the elements $(Y, 0)$, where Y is an irreducible component of a fiber \mathcal{M}_p for some p . Also, $\text{MW}(\mathcal{M}_{\mathbb{Q}})$ is the Mordell-Weil group of $\mathcal{M}_{\mathbb{Q}}$, tensored with \mathbb{R} . By the above remark, we have to prove the modularity of $\ell(\hat{\phi}_1^0)$ for linear functionals ℓ on each of the summands of (1.0.28).

For the summand $\widetilde{\text{MW}}$, this is done by comparing the restriction to the generic fiber of our generating series $\hat{\phi}_1$ with the generating series considered by Borcherds [2], for which he proved modularity. Proposition 1.0.1 is used to produce divisors of degree 0 in the generic fiber from our special divisors.

For the summand $\mathbb{R}\hat{\omega}$, the modularity follows from the following theorem which is the main result of [38]. Note that this theorem not only gives modularity but even identifies the modular form explicitly. We form the generating series with coefficients in \mathbb{C} obtained by cupping with $\hat{\omega}$,

$$(1.0.30) \quad \langle \hat{\omega}, \phi_1 \rangle = \sum_t \langle \hat{\omega}, \hat{\mathcal{Z}}(t, v) \rangle q^t.$$

Theorem 1.0.3. *The series above coincides with the q -expansion of the derivative at $s = \frac{1}{2}$ of the Eisenstein series (1.0.15),*

$$\langle \hat{\omega}, \hat{\phi}_1 \rangle = \mathcal{E}'_1(\tau, \frac{1}{2}, B).$$

Next, consider the pairings of the generating series $\hat{\phi}_1$ with the classes $(Y, 0) \in \text{Vert}$, where Y is an irreducible component of a fiber with bad reduction \mathcal{M}_p , i.e., $p \mid D(B)$. The corresponding series can be identified with classical theta functions for the positive definite ternary lattice associated to the definite quaternion algebra $B^{(p)}$ with $D(B^{(p)}) = D(B)/p$. This

is based on the theory of p -adic uniformization and uses the analysis of the special cycles at primes of bad reduction [36].

Finally, for the series $\hat{\phi}_1^\infty$, we show that the coefficients of the spectral expansion of $\hat{\phi}_1$ are Maass forms. More precisely, if f_λ is an eigenfunction of the Laplacian with eigenvalue λ , then the coefficient of f_λ in $\hat{\phi}_1$ is up to an explicit scalar the classical theta lift $\theta(f_\lambda)$ to a Maass form of weight $\frac{3}{2}$ and level $\Gamma_0(4D(B)_o)$.

To formulate the second main result of this book, Theorem B, we form a generating series for 0-cycles on \mathcal{M} instead of divisors on \mathcal{M} . The idea is to impose a *pair* of special endomorphisms, i.e., ‘twice as much CM’. Let $\text{Sym}_2(\mathbb{Z})^\vee$ denote the set of half-integral symmetric matrices of size 2, and let $T \in \text{Sym}_2(\mathbb{Z})^\vee$. We define a relative DM-stack $\mathcal{Z}(T)$ over \mathcal{M} by posing the following moduli problem. To a scheme S the moduli problem $\mathcal{Z}(T)$ associates the category of triples (A, ι, \mathbf{x}) where

- (A, ι) is an object of $\mathcal{M}(S)$
- $\mathbf{x} = [x_1, x_2] \in \text{End}(A, \iota)^2$ is a pair of endomorphisms with $\text{tr}(x_1) = \text{tr}(x_2) = 0$, and $\frac{1}{2}(\mathbf{x}, \mathbf{x}) = T$.

Here $(\mathbf{x}, \mathbf{x}) = ((x_i, x_j))_{i,j}$. It is then clear that $\mathcal{Z}(T)$ has empty generic fiber when T is positive definite, since in characteristic 0 a fake elliptic curve cannot support linearly independent complex multiplications. However, perhaps somewhat surprisingly, $\mathcal{Z}(T)$ is not always a 0-divisor on \mathcal{M} .

To explain the situation, recall from [24] that any $T \in \text{Sym}_2(\mathbb{Z})^\vee$ with $\det(T) \neq 0$ determines a set of primes $\text{Diff}(T, B)$ of odd cardinality. More precisely, let $\mathcal{C} = (\mathcal{C}_p)$ be the (*incoherent*) collection of local quadratic spaces where $\mathcal{C}_p = V_p$ for $p < \infty$ and where \mathcal{C}_∞ is the positive definite quadratic space of dimension 3. If $T \in \text{Sym}_2(\mathbb{Q})$ is nonsingular, we let V_T be the unique ternary quadratic space over \mathbb{Q} with discriminant $-1 = \text{discr}(V)$ which represents T . We denote by B_T the unique quaternion algebra over \mathbb{Q} such that its trace zero subspace is isometric to V_T and define

$$(1.0.31) \quad \text{Diff}(T, B) = \{p \leq \infty \mid \text{inv}_p(B_T) \neq \text{inv}(\mathcal{C}_p)\}.$$

Note that $\infty \in \text{Diff}(T, B)$ if and only if T is not positive definite.

If $|\text{Diff}(T, B)| > 1$ or $\text{Diff}(T, B) = \{\infty\}$, then $\mathcal{Z}(T) = \emptyset$. Assume now that $\text{Diff}(T, B) = \{p\}$ with $p < \infty$. If $p \nmid D(B)$, then $\mathcal{Z}(T)$ is a 0-cycle on \mathcal{M} with support in the fiber \mathcal{M}_p , as desired. In fact, the cycle is concentrated in the supersingular locus of \mathcal{M}_p . If, however, $p \mid D(B)$, then $\mathcal{Z}(T)$ is (almost always) a vertical *divisor* concentrated in \mathcal{M}_p .

Our goal now is to form a generating series with coefficients in $\widehat{\text{CH}}^2(\mathcal{M})$,

$$(1.0.32) \quad \hat{\phi}_2 = \sum_{T \in \text{Sym}_2(\mathbb{Z})^\vee} \hat{\mathcal{Z}}(T, v) q^T.$$

Here the coefficients $\hat{\mathcal{Z}}(T, v) \in \widehat{\text{CH}}^2(\mathcal{M})$ will in general depend on $v \in \text{Sym}_2(\mathbb{R})_{>0}$. How to define them is evident from the above only in the case when T is positive definite and $\text{Diff}(T, B) = \{p\}$ with $p \nmid D(B)$. In this case we set

$$(1.0.33) \quad \hat{\mathcal{Z}}(T, v) = (\mathcal{Z}(T), 0) \in \widehat{\text{CH}}^2(\mathcal{M}),$$

independent of v . Then $\hat{\mathcal{Z}}(T, v)$ has image $\log |\mathcal{Z}(T)| \in \mathbb{R}$ under the arithmetic degree map (1.0.23). If $T \in \text{Sym}_2(\mathbb{Z})^\vee$ is nonsingular with $|\text{Diff}(T, B)| > 1$, we set $\hat{\mathcal{Z}}(T, v) = 0$. In the remaining cases, the definition we give of the coefficients of (1.0.32) is more subtle. If $\text{Diff}(T, B) = \{\infty\}$, then $\hat{\mathcal{Z}}(T, v)$ does depend on v ; its definition is purely archimedean and depends on the rotational invariance of the $*$ -product of two of the Green functions in [24], one of the main results of that paper. If $\text{Diff}(T, B) = \{p\}$ with $p \mid D(B)$, then the definition of $\hat{\mathcal{Z}}(T, v)$ (which is independent of v) relies on the $\text{GL}_2(\mathbb{Z}_p)$ -invariance of the degenerate intersection numbers on the Drinfeld upper half plane, one of the main results of [36]. Finally, for singular matrices $T \in \text{Sym}_2(\mathbb{Z})_{\geq 0}^\vee$ we are, in effect, imposing only a ‘single CM’, and the naive cycle is a divisor, so that its class lies in the wrong degree; we again use the heuristic principle that was used in the definition of the constant term of (1.0.14) and in the definition of $\hat{\mathcal{Z}}(0, v)$ in (1.0.26). In these cases we are guided in our definitions by the desire to give a construction that is on the one hand as natural as possible, and on the other hand to obtain the modularity of the generating series. We refer to Chapter 6 for the details.

Our second main theorem identifies the generating series (1.0.32) with an explicit (nonholomorphic) Siegel modular form of genus two. Recall that such a modular form admits a q -expansion as a Laurent series in

$$(1.0.34) \quad q^T = e(\text{tr}(T\tau)), \quad T \in \text{Sym}_2(\mathbb{Z})^\vee,$$

and that the coefficients may depend on the imaginary part $v \in \text{Sym}_2(\mathbb{R})_{>0}$ of $\tau = u + iv \in \mathfrak{H}_2$. We introduce a Siegel Eisenstein series $\mathcal{E}_2(\tau, s, B)$ which is *incoherent* in the sense of [24]. In particular, 0 is the center of symmetry for the functional equation, and $\mathcal{E}_2(\tau, 0, B) = 0$. The derivative at $s = 0$ is a nonholomorphic Siegel modular form of weight $\frac{3}{2}$.

Theorem B. *The generating function $\hat{\phi}_2$ is a Siegel modular form of genus two and weight $\frac{3}{2}$ of level $\Gamma_0(4D(B)_o) \subset \mathrm{Sp}_2(\mathbb{Z})$. More precisely,*

$$\hat{\phi}_2(\tau) = \mathcal{E}'_2(\tau, 0, B),$$

i.e., the q -expansion of the Siegel modular form on the right-hand side coincides with the generating series $\hat{\phi}_2$.

Here we are identifying implicitly $\widehat{\mathrm{CH}}^2(\mathcal{M})$ with \mathbb{R} via $\widehat{\mathrm{deg}}$, cf. (1.0.23). Theorem B is proved in Chapter 6 by explicitly comparing the coefficients of the q -expansion of $\mathcal{E}'_2(\tau, 0, B)$ with the coefficients $\hat{\mathcal{Z}}(T, v)$. This amounts to a series of highly nontrivial identities, one for each T in $\mathrm{Sym}_2(\mathbb{Z})^\vee$. Let us explain what is involved.

First let T be positive definite with $\mathrm{Diff}(T, B) = \{p\}$ for $p \nmid D(B)$. The calculation of the coefficient of $\mathcal{E}'_2(\tau, 0, B)$ corresponding to T comes down to the determination of derivatives of Whittaker functions or of certain representation densities. This determination is based on the explicit formulas for such densities due to Kitaoka [22] for $p \neq 2$. For $p = 2$, corresponding results are given in [55]. The determination of the arithmetic degree of $\mathcal{Z}(T)$ boils down to the problem of determining the length of the formal deformation ring of a 1-dimensional formal group of height 2 with two special endomorphisms. This is a special case of the theorem of Gross and Keating [17]. We point out that for both sides the prime number 2 ('the number theorist's nightmare') complicates matters considerably.

Next let T be positive definite with $\mathrm{Diff}(T, B) = \{p\}$ for $p \mid D(B)$. In this case, the corresponding derivatives of representation densities are determined in [54] for $p \neq 2$ and in [55] for $p = 2$. The determination of the corresponding coefficient of $\hat{\phi}_2$ depends on the calculation of the intersection product of special cycles on the Drinfeld upper half space. This is done in [36] for $p \neq 2$. These calculations are completed here for $p = 2$.

Now let T be nonsingular with $\mathrm{Diff}(T, B) = \infty$. Then the calculation of the corresponding coefficients of $\mathcal{E}'_2(\tau, 0, B)$ and of $\hat{\phi}_2$ is given in [24] in the case where the signature of T is $(1, 1)$. The remaining case, where the signature is $(0, 2)$, is given here, using the method of [24].

Next, we consider the coefficients corresponding to singular matrices T of rank 1. For such a matrix

$$(1.0.35) \quad T = \begin{pmatrix} t_1 & m \\ m & t_2 \end{pmatrix} \in \mathrm{Sym}_2(\mathbb{Z})^\vee,$$

with $\det(T) = 0$ and $T \neq 0$, we may write $t_1 = n_1^2 t$, $t_2 = n_2^2 t$, and $m = n_1 n_2 t$ for the relatively prime integers n_1 and n_2 and $t \in \mathbb{Z}_{\neq 0}$. The pair n_1, n_2 is unique up to simultaneous change in sign, and t is uniquely

determined. Also, note that, if $t_1 = 0$, then $n_1 = 0$, $n_2 = 1$, and $t = t_2$, while if $t_2 = 0$, then $n_1 = 1$, $n_2 = 0$, and $t = t_1$. Then the comparison between the corresponding singular coefficients of $\hat{\phi}_2$ and $\mathcal{E}'_2(\tau, 0, B)$ in this case is based on the following result, proved in Chapter 5. It relates the singular Fourier coefficients of the derivative of the genus two Eisenstein series occurring in Theorem B with the Fourier coefficients of the genus one Eisenstein series occurring in Theorem A.

Theorem 1.0.4. (i) *Let $T \in \text{Sym}_2(\mathbb{Z})^\vee$, with associated $t \in \mathbb{Z}_{\neq 0}$ as above. Then*

$$\begin{aligned} \mathcal{E}'_{2,T}(\tau, 0, B) &= -\mathcal{E}'_{1,t}(t^{-1}\text{tr}(T\tau), \frac{1}{2}, B) \\ &\quad - \frac{1}{2} \cdot \mathcal{E}_{1,t}(t^{-1}\text{tr}(T\tau), \frac{1}{2}, B) \cdot \left(\log\left(\frac{\det v}{t^{-1}\text{tr}(Tv)}\right) + \log(D(B)) \right). \end{aligned}$$

(ii) *For the constant term*

$$\mathcal{E}'_{2,0}(\tau, 0, B) = -\mathcal{E}'_{1,0}(i \det v, \frac{1}{2}, B) - \frac{1}{2} \mathcal{E}_{1,0}(i \det(v), \frac{1}{2}, B) \cdot \log D(B).$$

It is this theorem that motivated our definition of the singular coefficients of the generating series $\hat{\phi}_2$. Just as for Proposition 1.0.1, we see that the modularity of the generating function $\hat{\phi}_2$ is not proved directly but rather by identifying it with an explicit modular form.

The coherence in our definitions of the generating series $\hat{\phi}_1$ and $\hat{\phi}_2$ is displayed by the following arithmetic inner product formula, which relates the inner product of the generating series $\hat{\phi}_1$ with itself under the Gillet-Soulé pairing with the generating series $\hat{\phi}_2$. Let

$$(1.0.36) \quad \mathfrak{H} \times \mathfrak{H} \longrightarrow \mathfrak{H}_2 \quad (\tau_1, \tau_2) \longmapsto \text{diag}(\tau_1, \tau_2) = \begin{pmatrix} \tau_1 & 0 \\ 0 & \tau_2 \end{pmatrix}$$

be the natural embedding into the Siegel space of genus two.

Theorem C. *For $\tau_1, \tau_2 \in \mathfrak{H}$*

$$\langle \hat{\phi}_1(\tau_1), \hat{\phi}_1(\tau_2) \rangle = \hat{\phi}_2(\text{diag}(\tau_1, \tau_2)).$$

Explicitly, for any $t_1, t_2 \in \mathbb{Z}$ and $v_1, v_2 \in \mathbb{R}_{>0}$,

$$\langle \hat{\mathcal{Z}}(t_1, v_1), \hat{\mathcal{Z}}(t_2, v_2) \rangle = \sum_{\substack{T \in \text{Sym}_2(\mathbb{Z})^\vee \\ \text{diag}(T) = (t_1, t_2)}} \hat{\mathcal{Z}}(T, \text{diag}(v_1, v_2)).$$

Theorem C, which is proved in Chapter 7, is the third main result of this book and provides the arithmetic analogue of Theorem 6.2 in [23], which relates to the cup product of two generating series with values in cohomology. Let us explain what is involved here, first assuming that $t_1 t_2 \neq 0$.

The proof distinguishes two cases. In the first case $t_1 t_2 \notin \mathbb{Q}^{\times, 2}$. In this case all matrices T occurring in the sum on the right-hand side are automatically nonsingular; at the same time the divisors $\mathcal{Z}(t_1)$ and $\mathcal{Z}(t_2)$ have empty intersection in the generic fiber, so that the Gillet-Soulé pairing decomposes into a sum of local pairings, one for each prime of \mathbb{Q} . Consider the case when $t_i > 0$ for $i = 1, 2$. Then the key to the formula above is the decomposition of the intersection (fiber product) of the special cycles $\mathcal{Z}(t_i)$ according to ‘fundamental matrices’,

$$(1.0.37) \quad \mathcal{Z}(t_1) \times_{\mathcal{M}} \mathcal{Z}(t_2) = \coprod_{\substack{T \\ \text{diag}(T)=(t_1, t_2)}} \mathcal{Z}(T).$$

Here $\mathcal{Z}(T)$ appears as the locus of objects $((A, \iota), x_1, x_2)$ in the fiber product where $\mathbf{x} = [x_1, x_2]$ satisfies $\frac{1}{2}(\mathbf{x}, \mathbf{x}) = T$. Note that, by the remarks preceding the statement of Theorem B, the intersection of the $\mathcal{Z}(t_i)$ need not be proper since these divisors can have common components in the fibers of bad reduction \mathcal{M}_p for $p \mid D(B)$. Of course, all matrices T occurring in the disjoint sum in (1.0.37) are positive definite. The occurrence in the sum of Theorem C of summands corresponding to matrices T which are not positive definite is due to the Green functions component of the $\mathcal{Z}(t, v)$. Similar archimedean contributions occur in the cases where one of the t_i is negative.

In the second case $t_1 t_2 \in \mathbb{Q}^{\times, 2}$. In this case, $\mathcal{Z}(t_1)$ and $\mathcal{Z}(t_2)$ intersect in the generic fiber. In addition to the contribution of the nonsingular T to the sum in Theorem C, there is also a contribution of the two singular matrices T , where T is given by (1.0.35) with $m = \pm\sqrt{t_1 t_2}$. In this case the Gillet-Soulé pairing does not localize. Instead we use the arithmetic adjunction formula from Arakelov theory [10], [40]. To calculate the various terms in this formula we must, among other things, go back to the proof of the Gross-Keating formula and use the fine structure of the deformation locus of a special endomorphism of a p -divisible group of dimension 1 and height 2.

We stress that the proof of Theorem C sketched so far has nothing to do with Eisenstein series. However, the modularity of both sides of the identity in Theorem C allows us to deduce from the truth of the statement for all $t_1 t_2 \neq 0$ first the value of the constant c in (1.0.26) and then the truth of the statement for all (t_1, t_2) . In this way we can also prove our conjecture [38] on the self-intersection of the Hodge line bundle.

Theorem 1.0.5. *Let $\hat{\omega}_0$ be the Hodge line bundle on \mathcal{M} metrized with the normalization of Bost [3]. Then*

$$\langle \hat{\omega}_0, \hat{\omega}_0 \rangle = 2 \cdot \zeta_{D(B)}(-1) \left[\frac{\zeta'(-1)}{\zeta(-1)} + \frac{1}{2} - \frac{1}{4} \sum_{p|D(B)} \frac{p+1}{p-1} \log p \right].$$

Formally, this result specializes for $D(B) = 1$ to the formula of Bost [4] and Kühn [39] in the case of the modular curve (note that due to the stacks aspect our quantity is half of theirs). In their case they use the section Δ of $\omega^{\otimes 6}$ to compute the self-intersection of $\hat{\omega}_0$ explicitly from its definition. For Shimura curves there is no such natural modular form and our result comes about only indirectly. We note that the general form of this formula is related to formulas given by Maillot and Roessler [42].

The above three theorems are the main results in this book. As an application of these results, we introduce an arithmetic version of the Shimura-Waldspurger correspondence and obtain analogues of results of Waldspurger [53] and of Gross-Kohnen-Zagier [18].

If f is a cusp form of weight $\frac{3}{2}$ for $\Gamma_0(4D(B)_o)$, we can define the arithmetic theta lift of f by

$$(1.0.38) \quad \hat{\theta}(f) := C \cdot \int_{\Gamma_0(4D(B)_o) \backslash \mathfrak{H}} f(\tau) \overline{\hat{\phi}_1(\tau)} v^{-\frac{1}{2}} du dv \in \widehat{\mathbf{CH}}^1(\mathcal{M}^B),$$

for a constant C given in section 3 of Chapter 9. Of course, this is the analogue of the classical theta lift from modular forms of weight $\frac{3}{2}$ to modular forms of weight 2, but with $\hat{\phi}_1(\tau)$ replacing the classical theta kernel of Niwa [43] and Shintani [47]. By the results discussed above, it follows that

$$(1.0.39) \quad \langle \hat{\theta}(f), \mathbb{1} \rangle = \langle f, \mathcal{E}_1(\tau, \frac{1}{2}; B) \rangle_{\text{Pet}} = 0,$$

$$(1.0.40) \quad \langle \hat{\theta}(f), \hat{\omega} \rangle = \langle f, \mathcal{E}'_1(\tau, \frac{1}{2}; B) \rangle_{\text{Pet}} = 0,$$

and

$$(1.0.41) \quad \langle \hat{\theta}(f), a(\phi) \rangle = \langle f, \theta(\phi) \rangle_{\text{Pet}} = 0, \quad \text{for all } \phi \in C^\infty(\mathcal{M}(\mathbb{C}))_0,$$

since f is a holomorphic cusp form. Here, for $\phi \in C^\infty(\mathcal{M}(\mathbb{C}))_0$, we denote by $a(\phi)$ the corresponding class in $\widehat{\mathbf{CH}}^1(\mathcal{M})$ and by $\theta(\phi)$ the corresponding Maass cusp form of weight $\frac{3}{2}$. Thus $\hat{\theta}(f)$ lies in the space of $\widehat{\mathbf{MW}} \oplus \text{Vert}^0$, where Vert^0 is the subspace of Vert orthogonal to $\hat{\omega}$.

In order to obtain information about the nonvanishing of $\hat{\theta}(f)$, we con-

sider the height pairing $\langle \widehat{\phi}_1(\tau_1), \widehat{\theta}(f) \rangle$. Using Theorems B and C, we obtain

$$\begin{aligned}
 \langle \widehat{\phi}_1(\tau_1), \widehat{\theta}(f) \rangle &= \langle f, \langle \widehat{\phi}_1(\tau_1), \widehat{\phi}_1 \rangle \rangle \\
 (1.0.42) \qquad &= \langle f, \widehat{\phi}_2(\text{diag}(\tau_1, \cdot)) \rangle \\
 &= \langle f, \mathcal{E}'_2(\text{diag}(\tau_1, \cdot), 0; B) \rangle \\
 &= \frac{\partial}{\partial s} \left\{ \langle f, \mathcal{E}_2(\text{diag}(\tau_1, \cdot), s; B) \rangle \right\} \Big|_{s=0}.
 \end{aligned}$$

We then consider the integral $\langle f, \mathcal{E}_2(\text{diag}(\tau_1, \cdot), s; B) \rangle$ occurring in the last expression. This integral is essentially the doubling integral of Piatetski-Shapiro and Rallis [45] (see also [41]), except that we only integrate against one cusp form.

Theorem 1.0.6. *Let F be a normalized newform of weight 2 on $\Gamma_0(D(B))$ and let f be the good newvector, in the sense defined in section 3 of Chapter 8, corresponding to F under the Shimura-Waldspurger correspondence. Then*

$$\langle f, \mathcal{E}_2(\text{diag}(\tau_1, \cdot), s; B) \rangle = \mathcal{C}(s) \cdot L(s+1, F) \cdot f(\tau_1),$$

where

$$\mathcal{C}(s) = \frac{3}{2\pi^2} \prod_{p|D(B)} (p+1)^{-1} \cdot \left(\frac{D(B)}{2\pi} \right)^s \Gamma(s+1) \cdot \prod_{p|D(B)} \mathcal{C}_p(s),$$

with

$$\mathcal{C}_p(s) = (1 - \epsilon_p(F) p^{-s}) - \frac{p-1}{p+1} (1 + \epsilon_p(F) p^{-s}) B_p(s).$$

Here $L(s, F)$ is the standard Hecke L -function of F , $\epsilon_p(F)$ is the Atkin-Lehner sign of F ,

$$F|W_p = \epsilon_p(F) F,$$

and $B_p(s)$ is a rational function of p^{-s} with

$$B_p(0) = 0 \quad \text{and} \quad B'_p(0) = \frac{1}{2} \cdot \frac{p+1}{p-1} \log(p).$$

Note that $\mathcal{C}_p(0) = 2$ if $\epsilon_p(F) = -1$ and $\mathcal{C}_p(0) = \mathcal{C}'_p(0) = 0$ if $\epsilon_p(F) = 1$. As a consequence, we have the following analogue of Rallis's inner product formula [46], which characterizes the nonvanishing of the arithmetic theta lift.

Corollary 1.0.7. *For F with associated f as in Theorem 1.0.6,*

$$\langle \widehat{\phi}_1(\tau_1), \widehat{\theta}(f) \rangle = \mathcal{C}(0) \cdot L'(1, F) \cdot f(\tau_1).$$

In particular,

$$\langle \widehat{\theta}(f), \widehat{\theta}(f) \rangle = \mathcal{C}(0) \cdot L'(1, F) \cdot \langle f, f \rangle,$$

and hence

$$\widehat{\theta}(f) \neq 0 \quad \iff \quad \begin{cases} \epsilon_p(F) = -1 & \text{for all } p \mid D(B), \text{ and} \\ L'(1, F) \neq 0. \end{cases}$$

Let $S_2^{\text{new}}(D(B))^{(-)}$ be the space of normalized newforms of weight 2 for $\Gamma_0(D(B))$ for which all Atkin-Lehner signs are -1 . Note that, for $F \in S_2^{\text{new}}(D(B))^{(-)}$, the root number of $L(s, F)$ is given by

$$\epsilon(1, F) = - \prod_{p \mid D(B)} \epsilon_p(F) = -1.$$

Since the vertical part of $\widehat{\phi}_1(\tau)$ is a linear combination of theta functions for the anisotropic ternary spaces $V^{(p)}$, for $p \mid D(B)$, and since the classical theta lift of a form F with $\epsilon(1, F) = -1$ to such a space vanishes by Waldspurger's result [50], [53], it follows that $\widehat{\theta}(f) \in \widetilde{\text{MW}}$. Recall from (1.0.29) that this space is isomorphic to $\text{MW}(\mathcal{M}_{\mathbb{Q}})$ via the restriction map $\text{res}_{\mathbb{Q}}$.

Corollary 1.0.8. *For each $F \in S_2^{\text{new}}(D(B))^{(-)}$, let f be the corresponding good newvector of weight $\frac{3}{2}$. Then*

$$\text{res}_{\mathbb{Q}}(\widehat{\phi}_1^B(\tau)) = \mathcal{E}_1(\tau, \frac{1}{2}; B) \cdot \frac{\omega_{\mathbb{Q}}}{\deg \omega_{\mathbb{Q}}} + \sum_{\substack{F \in S_2^{\text{new}}(D(B))^{(-)} \\ L'(1, F) \neq 0}} \frac{f(\tau) \cdot \text{res}_{\mathbb{Q}} \widehat{\theta}(f)}{\langle f, f \rangle},$$

where $\omega_{\mathbb{Q}}$ is the restriction of the Hodge bundle to $\mathcal{M}_{\mathbb{Q}}$.

Next, for each $t \in \mathbb{Z}_{>0}$, write $Z(t)(F)$ for the component³ of the cycle $Z(t) = \mathcal{Z}(t)_{\mathbb{Q}}$ in the F -isotypic part $\text{CH}^1(\mathcal{M}_{\mathbb{Q}})(F)$ of the Chow group $\text{CH}^1(\mathcal{M}_{\mathbb{Q}})$. Note that $Z(t)(F)$ has zero image in $H^2(\mathcal{M}_{\mathbb{C}})$ and hence defines a class in $\text{MW}(\mathcal{M}_{\mathbb{Q}})$.

³Here we transfer F to a system of Hecke eigenvalues for the quaternion algebra B via the Jacquet-Langlands correspondence.

Theorem 1.0.9. *The F -isotypic component of the generating function*

$$\text{res}_{\mathbb{Q}}(\widehat{\phi}_1^B(\tau)) = \sum_{t \geq 0} Z(t) q^t,$$

is

$$\text{res}_{\mathbb{Q}}(\widehat{\phi}_1^B(\tau))(F) = \sum_{t \geq 0} Z(t)(F) q^t = \frac{f(\tau) \cdot \text{res}_{\mathbb{Q}} \widehat{\theta}(f)}{\langle f, f \rangle}.$$

In particular,

$$Z(t)(F) = \frac{a_t(f) \cdot \text{res}_{\mathbb{Q}} \widehat{\theta}(f)}{\langle f, f \rangle},$$

where

$$f(\tau) = \sum_{t > 0} a_t(f) q^t$$

is the Fourier expansion of f . Moreover, for t_1 and $t_2 \in \mathbb{Z}_{>0}$, the height pairing of the F -components of $Z(t_1)$ and $Z(t_2)$ is given by

$$\langle Z(t_1)(F), Z(t_2)(F) \rangle = \mathcal{C}(0) \cdot L'(1, F) \cdot \frac{a_{t_1}(f) \cdot a_{t_2}(f)}{\langle f, f \rangle}.$$

This result is the analogue in our case of the result of Gross-Kohnen-Zagier [18], Theorem C, p.503. The restriction to newforms in $S_2^{\text{new}}(D(B))$ with all Atkin-Lehner signs equal to -1 is due to the fact that our cycles are invariant under all Atkin-Lehner involutions. To remove this restriction, one should use ‘weighted’ cycles, see section 4 of Chapter 3.

In fact, we construct an arithmetic theta lift of automorphic representations σ in the space $\mathcal{A}_{00}(G')$ on the metaplectic extension $G'_{\mathbb{A}}$ of $\text{SL}_2(\mathbb{A})$. This theta lift, which is only defined for representations corresponding to holomorphic cusp forms of weight $\frac{3}{2}$, is the analogue of the classical theta lift considered by Waldspurger [50], [51], [53]. We formulate a conjectural analogue of Waldspurger’s nonvanishing criterion and prove it in certain cases as an application of Theorem 1.0.6 and Corollary 1.0.7. For forms F with $\epsilon(1, F) = +1$, Waldspurger proved that the classical theta lift is nonzero if and only if (i) certain local conditions (theta dichotomy) are satisfied at every place, and (ii) $L(1, F) \neq 0$. In the arithmetic case, we show that for (certain) forms F of weight 2 with $\epsilon(1, F) = -1$, the arithmetic theta lift is nonzero if and only if (i) the local theta dichotomy conditions are satisfied, and (ii) $L'(1, F) \neq 0$. A more detailed discussion can be found in section 1 of Chapter 9 as well as in [29]. Our construction is similar in spirit

to that of [16], where Gross formulates an arithmetic analogue of another result of Waldspurger [52] and shows that, in certain cases, this analogue can be proved using the results of Gross-Zagier [19] and their extension by Zhang [60].

We now mention some previous work on such geometric and arithmetic-geometric generating functions. The classic work of Hirzebruch-Zagier mentioned above inspired much work on modular generating functions valued in cohomology. Kudla and Millson considered modular generating functions for totally geodesic cycles in Riemannian locally symmetric spaces for the classical groups $O(p, q)$, $U(p, q)$, and $Sp(p, q)$ [31], [32], [33]. Such cycles were also considered by Oda [44] and Tong-Wang [49]. In the case of symmetric spaces for $O(n, 2)$, the generating function of Kudla-Millson [33] and Kudla [23] for the cohomology classes of algebraic cycles of codimension r is a Siegel modular form of weight $\frac{n}{2} + 1$ and genus r . In the case $r = n$, i.e., for 0-cycles, the generating function was identified in [23] as a special value of an Eisenstein series via the Siegel-Weil formula. A similar relation to Eisenstein series occurs in the work of Gross and Keating [17] for the generating series associated to the graphs of modular correspondences in a product of two modular curves. Borchers [2] used Borchers products to construct modular generating series with coefficients in $\mathbb{C}H^1$ for divisors on locally symmetric varieties associated to $O(n, 2)$ and proved that they are holomorphic modular forms. We also mention recent related work of Bruinier [5], [6], Bruinier-Funke [8], Funke [11], and Funke-Millson [12], [13].

The results in the arithmetic context are all inspired by the theorem of Gross and Zagier [19]. Part of a generating series for triple arithmetic intersections of curves on the product of two modular curves was implicitly considered in the paper by Gross and Keating [17], where the ‘good nonsingular’ coefficients are determined explicitly, cf. also [1]. For Shimura curves, Kudla [24] considered the generating series obtained from the Gillet-Soulé height pairing of special divisors. It was proved that this generating series coincided for ‘good’ nonsingular coefficients with the diagonal pull-back of the central derivative of a Siegel Eisenstein series of genus two. The ‘bad’ nonsingular coefficients were determined in [36]. However, the singular coefficients were left out of this comparison. In [37] we considered the 0-dimensional case, where the ambient space is the moduli space of elliptic curves with complex multiplication. In this case we were able to determine the generating series completely and to identify it with the derivative of a special value of an Eisenstein series. Another generating series is obtained in [38] by pairing special divisors on arithmetic models of Shimura curves, equipped with Green functions, with the metrized dualizing line bundle. Again this can be determined completely and identified with a

special value of a derivative of an Eisenstein series. A generating series in a higher-dimensional case is constructed by Bruinier, Burgos, and Kühn [7]. They consider special divisors on arithmetic models of Hilbert-Blumenthal surfaces whose generic fibers are Hirzebruch-Zagier curves, equip them with (generalized) Green functions [9], and obtain a generating series by taking the pairing with *the square* of the metrized dualizing line bundle. They identify this series with a special value of an Eisenstein series. Finally we mention partial results in higher-dimensional cases (Hilbert-Blumenthal surfaces, Siegel threefolds) in [34], [35].

This monograph is not self-contained. Rather, we make essential use of our previous papers. We especially need the results in [24] about the particular Green functions we use, as well as the results on Eisenstein series developed there. We also use the results on representation densities from [54], [55]. Furthermore, for the analysis of the situation at the fibers of bad reduction we use the results contained in [36]. These are completed in [38], which is also essential for our arguments in other ways. Finally, we need some facts from [27] in order to apply the results of Borcherds. These papers are not reproduced here. Still, we have given here all the definitions necessary for following our development and have made an effort to direct the reader to the precise reference where he can find the proof of the statement in question. We also have filled in some details in the proof of other results in the literature. Most notable here are our exposition in section 6 of Chapter 3 of the special case of the theorem of Gross and Keating [17] that we use, and the exposition in Chapter 8 of the doubling method of Piatetski-Shapiro and Rallis [45] in the special case relevant to us. In the first instance, we were aided by a project with a similar objective, namely to give an exposition of the general result of Gross and Keating, undertaken by the ARGOS seminar in Bonn [1]. In the second instance, we use precise results about nonarchimedean local Howe duality for the dual pair $(\overline{\mathrm{SL}}_2, \mathrm{O}(3))$ from [30].

We have structured this monograph in the following manner. In Chapter 2 we provide the necessary background from Arakelov geometry. The key point here is to show that the theory of Gillet-Soulé [14], [3] continues to hold for the DM-stacks of the kind we encounter. We also give a version of the arithmetic adjunction formula. It turns out that among the various versions of it the most naive form, as presented in Lang's book [40], is just what we need for our application of it in Chapter 7. In Chapter 3 we define the special cycles on Shimura curves and review the known facts about them. Here we also give a proof of the special case of the Gross-Keating formula which we need. In Chapter 4 we prove Theorem A, along the lines sketched above. In Chapter 5 we introduce the Eisenstein series of genus one and two which are relevant to us and calculate their Fourier

expansion. In particular, we prove Theorem 1.0.4. In Chapter 6 we define the generating series $\hat{\phi}_2$ and prove Theorem B by comparing term by term this series with the Fourier coefficients of the Siegel Eisenstein series of genus two determined in the previous chapter. For the ‘bad nonsingular’ coefficients of $\hat{\phi}_2$, the calculation in the case $p = 2$ had been left out in [36]. In the appendix to Chapter 6 we complete the calculations for $p = 2$. Chapter 7 is devoted to the proof of the inner product formula, Theorem C. In Chapter 8 we give an exposition of the doubling method in our case. The point is to determine explicitly all local zeta integrals for the kind of good test functions that we use. The case $p = 2$ again requires additional efforts. In Chapter 9 we give applications of our results to the arithmetic theta lift and to L-functions and prove Theorems 1.0.6 and 1.0.9 and Corollaries 1.0.7 and 1.0.8 above.

This book is the result of a collaboration over many years. The general idea of forming the arithmetic generating series and relating them to modular forms arising from derivatives of Eisenstein series is due to the first author. The other two authors joined the project, each one contributing a different expertise to the undertaking. In the end, we can honestly say that no proper subset of this set of authors would have been able to bring this project to fruition. While the book is thus the product of a joint enterprise, some chapters have a set of *principal authors* which are as follows:

Chapter 2: SK, MR

Chapter 4: SK

Chapter 5: SK, TY

Appendix to Chapter 6: SK, MR

Chapter 7: SK, MR

Chapter 8: SK, TY

The material of this book, as well as its background, has been the subject of several survey papers by us individually: [25], [26], [28], [29], [56], [57]. It should be pointed out, however, that in the intervening time we made progress and that quite a number of question marks which still decorate the announcements of our results in these papers have been removed.

Bibliography

- [1] ARGOS (Arithmetische Geometrie Oberseminar), Proceedings of the Bonn seminar 2003/04, forthcoming.
- [2] R. Borcherds, *The Gross-Kohnen-Zagier theorem in higher dimensions*, Duke Math. J., **97** (1999), 219–233.

- [3] J.-B. Bost, *Potential theory and Lefschetz theorems for arithmetic surfaces*, Ann. Sci. École Norm. Sup., **32** (1999), 241–312.
- [4] _____, Lecture, Univ. of Maryland, Nov. 11, 1998.
- [5] J. H. Bruinier, *Borcherds products and Chern classes of Hirzebruch–Zagier divisors*, Invent. Math., **138** (1999), 51–83.
- [6] _____, *Borcherds Products on $O(2, l)$ and Chern Classes of Heegner Divisors*, Lecture Notes in Math., **1780**, Springer-Verlag, New York, 2002.
- [7] J. H. Bruinier, J. I. Burgos Gil, and U. Kühn, *Borcherds products in the arithmetic intersection theory of Hilbert modular surfaces*, preprint, 2003.
- [8] J. H. Bruinier and J. Funke, *On two geometric theta lifts*, preprint, 2003.
- [9] J. I. Burgos Gil, J. Kramer, and U. Kühn, *Cohomological arithmetic Chow rings*, preprint, 2003.
- [10] G. Faltings, *Calculus on arithmetic surfaces*, Annals of Math., **119** (1984), 387–424.
- [11] J. Funke, *Heegner divisors and nonholomorphic modular forms*, Compositio Math., **133** (2002), 289–321.
- [12] J. Funke and J. Millson, *Cycles in hyperbolic manifolds of non-compact type and Fourier coefficients of Siegel modular forms*, Manuscripta Math., **107** (2003), 409–444.
- [13] _____, *Cycles with local coefficients for orthogonal groups and vector valued Siegel modular forms*, preprint, 2004.
- [14] H. Gillet and C. Soulé, *Arithmetic intersection theory*, Publ. Math. IHES, **72** (1990), 93–174.
- [15] _____, *Arithmetic analogues of standard conjectures*, in Proc. Symp. Pure Math., **55**, Part 1, 129–140, AMS, Providence, R.I., 1994.
- [16] B. H. Gross, *Heegner points and representation theory*, in Heegner points and Rankin L-Series, Math. Sci. Res. Inst. Publ., **49**, Cambridge Univ. Press, Cambridge, 2004.

- [17] B. Gross and K. Keating, *On the intersection of modular correspondences*, *Invent. math.*, **112** (1993), 225–245.
- [18] B. H. Gross, W. Kohlen, and D. Zagier, *Heegner points and derivatives of L -functions. II*, *Math. Annalen*, **278** (1987), 497–562.
- [19] B. H. Gross and D. Zagier, *Heegner points and the derivatives of L -series*, *Invent. math.*, **84** (1986), 225–320.
- [20] F. Hirzebruch and D. Zagier, *Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus*, *Invent. math.*, **36** (1976), 57–113.
- [21] Y. Kitaoka, *A note on local densities of quadratic forms*, *Nagoya Math. J.*, **92** (1983), 145–152.
- [22] _____, *Arithmetic of Quadratic Forms*, *Cambridge Tracts in Mathematics*, **106**, Cambridge Univ. Press, 1993.
- [23] S. Kudla, *Algebraic cycles on Shimura varieties of orthogonal type*, *Duke Math. J.*, **86** (1997), 39–78.
- [24] _____, *Central derivatives of Eisenstein series and height pairings*, *Annals of Math.*, **146** (1997), 545–646.
- [25] _____, *Derivatives of Eisenstein series and generating functions for arithmetic cycles*, *Séminaire Bourbaki 876*, *Astérisque*, **276**, 341–368, Soc. Math. France, Paris, 2002.
- [26] _____, *Derivatives of Eisenstein series and arithmetic geometry*, in *Proc. Intl. Cong. Mathematicians, Vol II (Beijing, 2002)*, 173–183, Higher Education Press, Beijing, 2002.
- [27] _____, *Integrals of Borcherds forms*, *Compositio Math.*, **137** (2003), 293–349.
- [28] _____, *Special cycles and derivatives of Eisenstein series*, in *Heegner points and Rankin L -Series*, *Math. Sci. Res. Inst. Publ.*, **49**, 243–270, Cambridge Univ. Press, Cambridge, 2004.
- [29] _____, *Modular forms and arithmetic geometry*, in *Current Developments in Mathematics, 2002*, 135–179, International Press, Somerville, MA, 2003.
- [30] _____, *Notes on the local theta correspondence for $(\widetilde{\mathrm{SL}}_2, O(3))$* , preprint, 2005.

- [31] S. Kudla and J. Millson, *The theta correspondence and harmonic forms I*, Math. Annalen, **274** (1986), 353–378.
- [32] _____, *The theta correspondence and harmonic forms II*, Math. Annalen, **277** (1987), 267–314.
- [33] _____, *Intersection numbers of cycles on locally symmetric spaces and Fourier coefficients of holomorphic modular forms in several complex variables*, Publ. Math. IHES, **71** (1990), 121–172.
- [34] S. Kudla and M. Rapoport, *Arithmetic Hirzebruch–Zagier cycles*, J. reine angew. Math., **515** (1999), 155–244.
- [35] _____, *Cycles on Siegel 3-folds and derivatives of Eisenstein series*, Ann. Sci. École. Norm. Sup., **33** (2000), 695–756.
- [36] _____, *Height pairings on Shimura curves and p -adic uniformization*, Invent. math., **142** (2000), 153–223.
- [37] S. Kudla, M. Rapoport, and T. Yang, *On the derivative of an Eisenstein series of weight 1*, Int. Math. Res. Notices (1999), 347–385.
- [38] _____, *Derivatives of Eisenstein series and Faltings heights*, Compositio Math., **140** (2004), 887–951.
- [39] U. Kühn, *Generalized arithmetic intersection numbers*, J. reine angew. Math., **534** (2001), 209–236.
- [40] S. Lang, *Introduction to Arakelov Theory*, Springer-Verlag, New York, 1988.
- [41] J.-S. Li, *Nonvanishing theorems for the cohomology of certain arithmetic quotients*, J. reine angew. Math., **428** (1992), 177–217.
- [42] V. Maillot and D. Roessler, *Conjectures sur les dérivées logarithmiques des fonctions L d’Artin aux entiers négatifs*, Math. Res. Lett., **9** (2002), no. 5-6, 715–724.
- [43] S. Niwa, *Modular forms of half integral weight and the integral of certain theta-functions*, Nagoya Math. J., **56** (1975), 147–161.
- [44] T. Oda, *On modular forms associated to quadratic forms of signature $(2, n - 2)$* , Math. Annalen., **231** (1977), 97–144.
- [45] I. I. Piatetski-Shapiro and S. Rallis, *L -functions for classical groups*, Lecture Notes in Math., **1254**, 1–52, Springer-Verlag, New York, 1987.

- [46] S. Rallis, *Injectivity properties of liftings associated to Weil representations*, *Compositio Math.*, **52** (1984), 139–169.
- [47] T. Shintani, *On construction of holomorphic cusp forms of half integral weight*, *Nagoya Math. J.*, **58** (1975), 83–126.
- [48] C. Soulé, D. Abramovich, J.-F. Burnol, and J. Kramer, *Lectures on Arakelov Theory*, *Cambridge Studies in Advanced Math.* **33**, Cambridge Univ. Press, Cambridge, 1992.
- [49] Y. Tong and S. P. Wang, *Construction of cohomology of discrete groups*, *Trans. AMS*, **306** (1988), 735–763.
- [50] J.-L. Waldspurger, *Correspondance de Shimura*, *J. Math. Pures Appl.*, **59** (1980), 1–132.
- [51] _____, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, *J. Math. Pures Appl.*, **60** (1981), 375–484.
- [52] _____, *Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie*, *Compositio Math.*, **54** (1985), 173–242.
- [53] _____, *Correspondances de Shimura et quaternions*, *Forum Math.*, **3** (1991), 219–307.
- [54] T. H. Yang, *An explicit formula for local densities of quadratic forms*, *J. Number Theory*, **72** (1998), 309–356.
- [55] _____, *Local densities of 2-adic quadratic forms*, *J. Number Theory*, **108** (2004), 287–345.
- [56] _____, *The second term of an Eisenstein series*, *Proc. Intl. Cong. Chinese Mathematicians*, forthcoming.
- [57] _____, *Faltings heights and the derivative of Zagier’s Eisenstein series*, in *Heegner points and Rankin L-Series*, *Math. Sci. Res. Inst. Publ.*, **49**, 271–284, Cambridge Univ. Press, Cambridge, 2004.
- [58] D. Zagier, *Nombres de classes et formes modulaires de poids $3/2$* , *C. R. Acad. Sci. Paris*, **281** (1975), 883–886.
- [59] _____, *Modular points, modular curves, modular surfaces and modular forms*, in *Lecture Notes in Math.* **1111**, 225–248, Springer-Verlag, Berlin, 1985.
- [60] Shou-Wu Zhang, *Gross–Zagier formula for GL_2* , *Asian J. Math.*, **5** (2001), 183–290.