BIRKHÄUSER

Progress in Mathematics
Volume 255

Representation Theory and Automorphic Forms

Toshiyuki Kobayashi
Wilfried Schmid
Jae-Hyun Yang
Editors

Birkhäuser
Boston • Basel • Berlin

Toshiyuki Kobayashi
RIMS, Kyoto University
Sakyo-ku, Kyoto, 606-8502
Japan
toshi@kurims.kyoto-u.ac.jp

Wilfried Schmid
Department of Mathematics
Harvard University
Cambridge, MA 02138
U.S.A.
schmid@math.harvard.edu

Jae-Hyun Yang
Department of Mathematics
Inha University
Incheon 402-751
Republic of Korea
jhyang@inha.ac.kr

Mathematics Subject Classifications (2000): 11R39, 11R42, 14G35, 22E45, 22E50, 22E55

Library of Congress Control Number: 2007933203

ISBN-13: 978-0-8176-4505-2 e-ISBN-13: 978-0-8176-4646-2

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com (KeS/MP)

Contents

Preface

Over the last half century, deep connections between representation theory and automorphic forms have been established, using a wide range of methods from algebra, geometry and analyis. In light of these developments, Changho Keem, Toshiyuki Kobayashi and Jae-Hyun Yang organized an international symposium entitled "Representation Theory and Automorphic Forms", with the hope that a broad discussion of recent ideas and techniques would lead to new breakthroughs in the field. The symposium was held at Seoul National University, Republic of Korea, February 14–17, 2005.

This volume is an outgrowth of the symposium. The lectures cover a variety of aspects of representation theory and autmorphic forms, among them, a lifting of elliptic cusp forms to Siegel and Hermitian modular forms (T. Ikeda), systematic and synthetic applications of the original theory of "visible actions" on complex manifolds to "multiplicity-free" theorems, in particular, to branching problems for reductive symmetric pairs (T. Kobayashi), an adaption of the Rankin–Selberg method to the setting of automorphic distributions (S. Miller and W. Schmid), recent developments in the Langlands functoriality conjecture and their relevance to certain conjectures in number theory, such as the Ramanujan and Selberg conjectures (F. Shahidi), cuspidality-irreducibility relation for automorphic representations (D. Ramakrishnan), and applications of Borcherds automorphic forms to the study of discriminants of certain $K3$ surfaces with involution that arise from the theory of hypergeometric functions (K.-I. Yoshikawa). By presenting some of the most active topics in the field, the editors hope that this volume will serve as an up-to-date introduction to the subject.

Acknowledgments

We thank the invited speakers for their enthusiastic lectures and the articles they have contributed. The referees also deserve our gratitude for their important role.

The symposium was initiated by, and received funding from, the Brain Korea 21 Mathematical Sciences Division of Seoul National University, BK21-MSD-SNU for short. Our thanks go to Hyuk Kim, Director of BK21-MSD-SNU, and to Chong-Kyu Han, Chairman of Department of Mathematics at SNU, for their support. We are especially indebted to Sung-Hoon Park, President of the JEI Corporation, and to Jee Hoon Park, President of JEI Distribution Co., LTD – and coincidentally a friend of the last-named editor; without the very generous financial contribution of the JEI Corporation the symposium would have been far more modest in scale.

Jaeyeon Joo and Eun-Soon Hong, secretaries of BK21-MSN-SNU, and Dong-Soo Shin did a splendid job preparing the symposium and catering to the needs of the participants. We are grateful also to Ann Kostant and Avanti Paranjpye of Birkhäuser Boston for their work in publishing this volume.

January 2007

Toshiyuki Kobayashi
Wilfried Schmid
Jae-Hyun Yang

1

Irreducibility and Cuspidality

Dinakar Ramakrishnan

253-37 Caltech
Pasadena, CA 91125, USA
dinakar@caltech.edu

Summary. Suppose ρ is an n-dimensional representation of the absolute Galois group of $\mathbb{Q}$ which is associated, via an identity of L-functions, with an automorphic representation π of $\mathrm{GL}(n)$ of the adele ring of $\mathbb{Q}$. It is expected that π is cuspidal if and only if ρ is irreducible, though nothing much is known in either direction in dimensions > 2. The object of this article is to show for $n < 6$ that the cuspidality of a *regular algebraic* π is implied by the irreducibility of ρ. For $n < 5$, it suffices to assume that π is semi-regular.

Key words: irreducibility, Galois representations, cuspidality, automorphic representations, general linear group, symplectic group, regular algebraic representations

Subject Classifications: 11F70; 11F80; 22E55

Introduction

Irreducible representations are the building blocks of general, semisimple Galois representations ρ, and *cuspidal* representations are the building blocks of automorphic forms π of the general linear group. It is expected that when an object of the former type is associated to one of the latter type, usually in terms of an identity of L-functions, the irreducibility of the former should imply the cuspidality of the latter, and vice versa. It is not a simple matter to prove this expectation, and nothing much is known in dimensions > 2. We will start from the beginning and explain the problem below, and indicate a result (in one direction) at the end of the introduction, which summarizes what one can do at this point. The remainder of the paper will be devoted to showing how to deduce this result by a synthesis of known theorems and some new ideas. We will be concerned here only with the *so-called easier* direction of showing the cuspidality of π given the irreducibility of ρ, and refer to [Ra5] for a more difficult result going the other way, which uses crystalline representations as well as a

*Partially supported by the NSF through the grant DMS-0402044.

refinement of certain deep modularity results of Taylor, Skinner–Wiles, et al. Needless to say, *easier* does not mean *easy*, and the significance of the problem stems from the fact that it does arise (in this direction) naturally. For example, π could be a functorial, automorphic image $r(\eta)$, for η a cuspidal automorphic representation of a product of smaller general linear groups: $H(\mathbb{A}) = \prod_j GL(m_j, \mathbb{A})$, with an associated Galois representation σ such that $\rho = r(\sigma)$ is irreducible. If the automorphy of π has been established by using a flexible converse theorem ([CoPS1]), then the cuspidality of π is not automatic. In [RaS], we had to deal with this question for cohomological forms π on GL(6), with $H = \mathrm{GL}(2) \times \mathrm{GL}(3)$ and r the Kronecker product, where π is automorphic by [KSh1]. Besides, the main result (Theorem A below) of this paper implies, as a consequence, the cuspidality of $\pi = \mathrm{sym}^4(\eta)$ for η defined by any non-CM holomorphic newform φ of weight ≥ 2 relative to $\Gamma_0(N) \subset \mathrm{SL}(2, \mathbb{Z})$, without appealing to the criterion of [KSh2]; here the automorphy of π is known by [K] and the irreducibility of ρ by [Ri].

Write $\overline{\mathbb{Q}}$ for the field of all algebraic numbers in $\mathbb{C}$, which is an infinite, mysterious Galois extension of $\mathbb{Q}$. One could say that the central problem in algebraic number theory is to understand this extension. *Class field theory*, one of the towering achievements of the twentieth century, helps us understand the *abelian* part of this extension, though there are still some delicate, open problems even in that well traversed situation.

Let $\mathcal{G}_{\mathbb{Q}}$ denote the absolute Galois group of $\mathbb{Q}$, meaning $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. It is a profinite group, being the projective limit of finite groups $\mathrm{Gal}(K/\mathbb{Q})$, as K runs over number fields which are normal over $\mathbb{Q}$. For fixed K, the Tchebotarev density theorem asserts that every conjugacy class C in $\mathrm{Gal}(K/\mathbb{Q})$ is the *Frobenius class* for an infinite number of primes p which are unramified in K. This shows the importance of studying the *representations* of Galois groups, which are intimately tied up with conjugacy classes. Clearly, every $\mathbb{C}$-representation, i.e., a homomorphism into $\mathrm{GL}(n, \mathbb{C})$ for some n, of $\mathrm{Gal}(K/\mathbb{Q})$ pulls back, via the canonical surjection $\mathcal{G}_{\mathbb{Q}} \to \mathrm{Gal}(K/\mathbb{Q})$, to a representation of $\mathcal{G}_{\mathbb{Q}}$, which is continuous for the profinite topology.

Conversely, one can show that every *continuous* $\mathbb{C}$-representation ρ of $\mathcal{G}_{\mathbb{Q}}$ is such a pullback, for a suitable finite Galois extension $K/\mathbb{Q}$. E. Artin associated an L-function, denoted $L(s, \rho)$, to any such ρ, such that the arrow $\rho \to L(s, \rho)$ is additive and inductive. He conjectured that for any non-trivial, irreducible, continuous $\mathbb{C}$-representation ρ of $\mathcal{G}_{\mathbb{Q}}$, $L(s, \rho)$ is entire, and this conjecture is open in general. Again, one understands well the *abelian* situation, i.e., when ρ is a 1-dimensional representation; the kernel of such a ρ defines an abelian extension of $\mathbb{Q}$. By class field theory, such a ρ is associated to a character ξ of finite order of the idele class group $\mathbb{A}^*/\mathbb{Q}^*$; here, being *associated* means they have the same L-function, with $L(s, \xi)$ being the one introduced by Hecke, *albeit* in a different language. As usual, we are denoting by $\mathbb{A} = \mathbb{R} \times \mathbb{A}_f$ the topological *ring of adeles*, with $\mathbb{A}_f = \widehat{\mathbb{Z}} \otimes \mathbb{Q}$, and by $\mathbb{A}^*$ its multiplicative group of *ideles*, which can be given the structure of a locally compact abelian topological group with discrete subgroup $\mathbb{Q}^*$.

Now fix a prime number ℓ, and an algebraic closure $\overline{\mathbb{Q}}_\ell$ of the field of ℓ-adic numbers $\mathbb{Q}_\ell$, equipped with an embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_\ell$. Consider the set $\mathcal{R}_\ell(n, \mathbb{Q})$ of continuous, semisimple representations

$$\rho_\ell : \mathcal{G}_{\mathbb{Q}} \to \mathrm{GL}(n, \overline{\mathbb{Q}}_\ell),$$

up to equivalence. The image of $\mathcal{G}_{\mathbb{Q}}$ in such a representation is usually not finite, and the simplest example of that is given by the ℓ-adic cyclotomic character χ_ℓ given by the action of $\mathcal{G}_{\mathbb{Q}}$ on all the ℓ-power roots of unity in $\overline{\mathbb{Q}}$. Another example is given by the 2-dimensional ℓ-adic representation on all the ℓ-power *division points* of an elliptic curve E over $\mathbb{Q}$.

The correct extension to the non-abelian case of the *idele class character*, which appears in class field theory, is the notion of an irreducible *automorphic representation* π of $\mathrm{GL}(n)$. Such a π is in particular a representation of the locally compact group $\mathrm{GL}(n, \mathbb{A}_F)$, which is a restricted direct product of the local groups $\mathrm{GL}(n, \mathbb{Q}_v)$, where v runs over all the primes p and ∞ (with $\mathbb{Q}_\infty = \mathbb{R}$). There is a corresponding factorization of π as a tensor product $\otimes_v \pi_v$, with all but a finite number of π_p being *unramified*, i.e., admitting a vector fixed by the maximal compact subgroup K_v. At the archimedean place ∞, π_∞ corresponds to an n-dimensional, semisimple representation $\sigma(\pi_\infty)$ of the real Weil group $W_{\mathbb{R}}$, which is a non-trivial extension of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ by $\mathbb{C}^*$. Globally, by Schur's lemma, the center $Z(\mathbb{A}) \simeq \mathbb{A}^*$ acts by a quasi-character ω, which must be trivial on $\mathbb{Q}^*$ by the automorphy of π, and so defines an idele class character. Let us restrict to the central case when π is essentially unitary. Then there is a (unique) real number t such that the twisted representation $\pi_u := \pi(t) = \pi \otimes |\cdot|^t$ is unitary (with unitary central character ω_u). We are, by abuse of notation, writing $|\cdot|^t$ to denote the quasi-character $|\cdot|^t \circ \det$ of $\mathrm{GL}(n, \mathbb{A})$, where $|\cdot|$ signifies the adelic absolute value, which is trivial on $\mathbb{Q}^*$ by the Artin product formula.

Roughly speaking, to say that π is automorphic means π_u appears (in a weak sense) in $L^2(Z(\mathbb{A})\mathrm{GL}(n, \mathbb{Q})\backslash \mathrm{GL}(n, \mathbb{A}), \omega_u)$, on which $\mathrm{GL}(n, \mathbb{A}_F)$ acts by right translations. A function φ in this L^2-space whose averages over all the horocycles are zero is called a *cusp form*, and π is called *cuspidal* if π_u is generated by the right $\mathrm{GL}(n, \mathbb{A}_F)$-translates of such a φ. Among the automorphic representations of $\mathrm{GL}(n, \mathbb{A})$ are certain distinguished ones called *isobaric automorphic representations*. Any isobaric π is of the form $\pi_1 \boxplus \pi_2 \boxplus \cdots \boxplus \pi_r$, where each π_j is a cuspidal representation of $\mathrm{GL}(n_j, \mathbb{A})$, such that $(n_1, n_2, \ldots, n_r)$ is a partition of n, where $\boxplus$ denotes the Langlands sum (coming from his theory of *Eisenstein series*); moreover, every *constituent* π_j is unique up to isomorphism. Let $\mathcal{A}(n, \mathbb{Q})$ denote the set of isobaric automorphic representations of $\mathrm{GL}(n, \mathbb{A})$ up to equivalence. Every isobaric π has an associated L-function $L(s, \pi) = \prod_v L(s, \pi_v)$, which admits a meromorphic continuation and a functional equation. Concretely, one associates at every prime p where π is unramified, a conjugacy class $A(\pi)$ in $\mathrm{GL}(n, \mathbb{C})$, or equivalently, an unordered n-tuple $(\alpha_{1,p}, \alpha_{2,p}, \ldots, \alpha_{n,p})$ of complex numbers so that

$$L(s, \pi_p) = \prod_{j=1}^{n} (1 - \alpha_{j,p} p^{-s})^{-1}.$$

If π is cuspidal and non-trivial, $L(s,\pi)$ is entire; so is the incomplete one $L^S(s,\pi)$ for any finite set S of places of $\mathbb{Q}$.

Now suppose ρ_ℓ is an n-dimensional, semisimple ℓ-adic representation of $\mathcal{G}_\mathbb{Q} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *corresponds* to an automorphic representation π of $\mathrm{GL}(n,\mathbb{A})$. We will take this to mean that there is a finite set S of places including ℓ, ∞ and all the primes where ρ_ℓ or π is ramified, such that we have

$$L(s,\pi_p) = L_p(s,\rho_\ell), \quad \forall p \notin S, \tag{0.1}$$

where the Galois Euler factor on the right is given by the characteristic polynomial of Fr_p, the Frobenius at p, acting on ρ_ℓ. When (0.1) holds (for a suitable S), we will write

$$\rho_\ell \leftrightarrow \pi.$$

A natural question in such a situation is to ask if π is cuspidal when ρ_ℓ is irreducible, and *vice versa*. It is certainly what is predicted by the general philosophy. However, proving it is another matter altogether, and positive evidence is scarce beyond $n = 2$.

One can answer this question in the affirmative, for any n, if one restricts to those ρ_ℓ which have *finite* image. In this case, it also defines a continuous, $\mathbb{C}$-representation ρ, the kind studied by E. Artin ([A]). Indeed, the hypothesis implies the identity of L-functions

$$L^S(s,\rho\otimes\rho^\vee) = L^S(s,\pi\times\pi^\vee), \tag{0.2}$$

where the superscript S signifies the removal of the Euler factors at places in S, and $\rho^\vee$ (resp. $\pi^\vee$) denotes the contragredient of ρ (resp. π). The L-function on the right is the Rankin–Selberg L-function, whose mirific properties have been established in the independent and complementary works of Jacquet, Piatetski-Shapiro and Shalika ([JPSS], and of Shahidi ([Sh1, Sh2]); see also [MW]. A theorem of Jacquet and Shalika ([JS1]) asserts that the *order of pole* at $s = 1$ of $L^S(s,\pi\times\pi^\vee)$ is 1 iff π is cuspidal. On the other hand, for any finite-dimensional $\mathbb{C}$-representation τ of $\mathcal{G}_\mathbb{Q}$, one has

$$-\mathrm{ord}_{s=1} L^S(s,\tau) = \dim_\mathbb{C} \mathrm{Hom}_{\mathcal{G}_\mathbb{Q}}(\underline{1},\tau), \tag{0.3}$$

where $\underline{1}$ denotes the trivial representation of $\mathcal{G}_\mathbb{Q}$. Applying this with $\tau = \rho\otimes\rho^\vee \simeq \mathrm{End}(\rho)$, we see that the order of pole of $L^S(s,\rho\otimes\rho^\vee)$ at $s = 1$ is 1 iff the only operators in $\mathrm{End}(\rho)$ which commute with the $\mathcal{G}_\mathbb{Q}$-action are scalars, which means by Schur that ρ is irreducible. Thus, *in the Artin case*, π *is cuspidal iff* ρ_ℓ *is irreducible*.

For general ℓ-adic representations ρ_ℓ of $\mathcal{G}_\mathbb{Q}$, the order of pole at the right edge is not well understood. When ρ_ℓ comes from *arithmetic geometry*, i.e., when it is a Tate twist of a piece of the cohomology of a smooth projective variety over $\mathbb{Q}$ which is cut out by algebraic projectors, an important *conjecture of Tate* asserts an analogue of (0.3) for $\tau = \rho_\ell\otimes\rho_\ell^\vee$, but this is unknown except in a few families of examples, such as those coming from the theory of *modular curves*, *Hilbert modular surfaces* and *Picard modular surfaces*. So one has to find a different way to approach the problem, which works at least in low dimensions.

The main result of this paper is the following:

Theorem A. *Let $n \leq 5$ and let ℓ be a prime. Suppose $\rho_\ell \leftrightarrow \pi$, for an isobaric, algebraic automorphic representation π of $GL(n, \mathbb{A})$, and a continuous, ℓ-adic representation ρ_ℓ of $\mathcal{G}_{\mathbb{Q}}$. Assume*

(i) ρ_ℓ is irreducible
(ii) π is odd if $n \geq 3$
(iii) π is semi-regular if $n = 4$, and regular if $n = 5$

Then π is cuspidal.

Some words of explanation are called for at this point. An isobaric automorphic representation π is said to be *algebraic* ([Cℓ1]) if the restriction of $\sigma_\infty := \sigma(\pi_\infty(\frac{1-n}{2}))$ to $\mathbb{C}^*$ is of the form $\oplus_{j=1}^n \chi_j$, with each χ_j algebraic, i.e., of the form $z \to z^{p_j}\bar{z}^{q_j}$ with $p_j, q_j \in \mathbb{Z}$. (We do not assume that our automorphic representations are unitary, and the arrow $\pi_\infty \to \sigma_\infty$ is normalized arithmetically.) For $n = 1$, an algebraic π is an idele class character of type A_0 in the sense of Weil. One says that π is *regular* iff $\sigma_\infty|_{\mathbb{C}^*}$ is a direct sum of characters χ_j, each occurring with *multiplicity one*. And π is *semi-regular* ([BHR]) if each χ_j occurs with *multiplicity at most two*. Suppose ξ is a 1-dimensional representation of $W_{\mathbb{R}}$. Then, since $W_{\mathbb{R}}^{\mathrm{ab}} \simeq \mathbb{R}^*$, ξ is defined by a character of $\mathbb{R}^*$ of the form $x \to |x|^w \cdot sgn(x)^{a(\xi)}$, with $a(\xi) \in \{0, 1\}$; here sgn denotes the sign character of $\mathbb{R}^*$. For every w, let $\sigma_\infty[\xi] := \sigma(\pi_\infty(\frac{1-n}{2}))[\xi]$ denote the isotypic component of ξ, which has dimension at most 2 (resp. 1) if π is semi-regular (resp. regular), and is acted on by $\mathbb{R}^*/\mathbb{R}_+^* \simeq \{\pm 1\}$. We will call a semi-regular π *odd* if for every character ξ of $W_{\mathbb{R}}$, the eigenvalues of $\mathbb{R}^*/\mathbb{R}_+^*$ on the ξ-isotypic component are distinct. Clearly, any regular π is odd under this definition. See Section 1 for a definition of this concept for any algebraic π, not necessarily semi-regular.

I want to thank the organizers, Jae-Hyun-Yang in particular, and the staff, of the *International Symposium on Representation Theory and Automorphic Forms* in Seoul, Korea, first for inviting me to speak there (during February 14–17, 2005), and then for their hospitality while I was there. The talk I gave at the conference was on a different topic, however, and dealt with my ongoing work with Dipendra Prasad on *selfdual representations*. I would also like to thank F. Shahidi for helpful conversations and the referee for his comments on an earlier version, which led to an improvement of the presentation. It is perhaps apt to end this introduction at this point by acknowledging support from the National Science Foundation via the grant DMS – 0402044.

1 Preliminaries

1.1 Galois representations

For any field k with algebraic closure $\overline{k}$, denote by $\mathcal{G}_k$ the *absolute Galois group* of $\overline{k}$ over k. It is a projective limit of the automorphism groups of finite Galois extensions

E/k. We furnish $\mathcal{G}_k$ as usual with the *profinite topology*, which makes it a *compact, totally disconnected topological group*. When $k = \mathbb{F}_p$, there is for every n a unique extension of degree n, which is Galois, and $\mathcal{G}_{\mathbb{F}_p}$ is isomorphic to $\widehat{\mathbb{Z}} \simeq \lim_n \mathbb{Z}/n$, topologically generated by the *Frobenius automorphism* $x \to x^p$.

At each prime p, let $\mathcal{G}_p$ denote the local Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ with inertia subgroup I_p, which fits into the following exact sequence:

$$1 \to I_p \to \mathcal{G}_p \to \mathcal{G}_{\mathbb{F}_p} \to 1. \tag{1.1.1}$$

The fixed field of $\overline{\mathbb{Q}}_p$ under I_p is the *maximal unramified extension* $\mathbb{Q}_p^{\mathrm{ur}}$ of $\mathbb{Q}_p$, which is generated by all the roots of unity of order prime to p. One gets a natural isomorphism of $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p)$ with $\mathcal{G}_{\mathbb{F}_p}$. If $K/\mathbb{Q}$ is unramified at p, then one can lift the Frobenius element to a conjugacy class φ_p in $\mathrm{Gal}(K/\mathbb{Q})$.

All the Galois representations considered here will be continuous and finite-dimensional. Typically, we will fix a prime ℓ, and algebraic closure $\overline{\mathbb{Q}}_\ell$ of the field $\mathbb{Q}_\ell$ of ℓ-adic numbers, and consider a continuous homomorphism

$$\rho_\ell : \mathcal{G}_{\mathbb{Q}} \to \mathrm{GL}(V_\ell), \tag{1.1.2}$$

where V_ℓ is an n-dimensional vector space over $\overline{\mathbb{Q}}_\ell$. We will be interested only in those ρ_ℓ that are unramified outside a finite set S of primes. Then ρ_ℓ factors through a representation of the quotient group $\mathcal{G}_S := G(\mathbb{Q}_S/\mathbb{Q})$, where $\mathbb{Q}_S$ is the maximal extension of $\mathbb{Q}$ which is unramified outside S. One has the Frobenius classes ϕ_p in $\mathcal{G}_S$ for all $p \notin S$, and this allows one to define the L-factors (with $s \in \mathbb{C}$)

$$L_p(s, \rho_\ell) = \det(I - \varphi_p p^{-s} | V_\ell)^{-1}. \tag{1.1.3}$$

Clearly, it is the reciprocal of a polynomial in p^{-s} of degree n, with constant term 1, and it depends only on the equivalence class of ρ_ℓ. One sets

$$L^S(s, \rho_\ell) = \prod_{p \notin S} L_p(s, \rho_\ell). \tag{1.1.4}$$

When ρ_ℓ is the trivial representation, it is unramified everywhere, and $L^S(s, \rho_\ell)$ is none other than the *Riemann zeta function*. To define the *bad factors* at p in $S - \{\ell\}$, one replaces V_ℓ in (1.1.3) the subspace $V_\ell^{I_p}$ of *inertial invariants*, on which φ_p acts.

We are primarily interested in *semisimple representations* in this article, which are direct sums of *simple* (or *irreducible*) representations. Given any representation ρ_ℓ of $\mathcal{G}_{\mathbb{Q}}$, there is an associated *semisimplification*, denoted ρ_ℓ^{ss}, which is a direct sum of the simple Jordan–Holder components of ρ_ℓ. A *theorem of Tchebotarev* asserts the density of the Frobenius classes in the Galois group, and since the local p-factors of $L(s, \rho_\ell)$ are defined in terms of the *inverse roots* of φ_p, one gets the following standard, but useful result.

Proposition 1.1.5. *Let ρ_ℓ, ρ'_ℓ be continuous, n-dimensional ℓ-adic representations of $\mathcal{G}_{\mathbb{Q}}$. Then*

$$L^S(s, \rho_\ell) = L^S(s, \rho'_\ell) \implies \rho_\ell^{\mathrm{ss}} \simeq \rho_\ell'^{\mathrm{ss}}.$$

The Galois representations ρ_ℓ which have *finite image* are special, and one can view them as continuous $\mathbb{C}$-representations ρ. Artin studied these in depth and showed, using the results of Brauer and Hecke, that the corresponding L-functions admit meromorphic continuation and a functional equation of the form

$$L^*(s,\rho) = \varepsilon(s,\rho)L^*(1-s,\rho^\vee), \tag{1.1.6}$$

where $\rho^\vee$ denotes the contragredient representation on the dual vector space, where

$$L^*(s,\rho) = L(s,\rho)L_\infty(s,\rho), \tag{1.1.7}$$

with the *archimedean factor* $L_\infty(s,\rho)$ being a suitable product (shifted) gamma functions. Moreover,

$$\varepsilon(s,\rho) = W(\rho)N(\rho)^{s-1/2}, \tag{1.1.8}$$

which is an entire function of s, with the (non-zero) $W(\rho)$ being called the *root number* of ρ. The scalar $N(\rho)$ is an integer, called the *Artin conductor* of ρ, and the finite set S which intervenes is the set of primes dividing $N(\rho)$. The functional equation shows that $W(\rho)W(\rho^\vee) = 1$, and so $W(\rho) = \pm 1$ when ρ is *selfdual* (which means $\rho \simeq \rho^\vee$). Here is a useful fact:

Proposition 1.1.9 ([T]). *Let τ be a continuous, finite-dimensional $\mathbb{C}$-representation of $\mathcal{G}_\mathbb{Q}$, unramified outside S. Then we have*

$$-\mathrm{ord}_{s=1} L^S(s,\tau) = \mathrm{Hom}_{\mathcal{G}_\mathbb{Q}}(\underline{1},\tau).$$

Corollary 1.1.10. *Let ρ be a continuous, finite-dimensional $\mathbb{C}$-representation of $\mathcal{G}_\mathbb{Q}$, unramified outside S. Then ρ is irreducible if and only if the incomplete L-function $L^S(s,\rho\otimes\rho^\vee)$ has a simple pole at $s = 1$.*

Indeed, if we set

$$\tau := \rho\otimes\rho^\vee \simeq \mathrm{End}(\rho), \tag{1.1.11}$$

then Proposition 1.1.9 says that the *order of pole* of $L(s,\rho\otimes\rho^\vee)$ at $s = 1$ is the *multiplicity of the trivial representation* in $\mathrm{End}(\rho)$ is 1, i.e., iff the *commutant* $\mathrm{End}_{\mathcal{G}_\mathbb{Q}}(\rho)$ is one-dimensional (over $\mathbb{C}$), which in turn is equivalent, by Schur's lemma, to ρ being irreducible. Hence the corollary.

For general ℓ-adic representations ρ_ℓ, there is no known analogue of Proposition 1.1.9, though it is predicted to hold (at the right edge of absolute convergence) by a *conjecture of Tate* when ρ_ℓ comes from *arithmetic geometry* (see [Ra4], Section 1, for example). Tate's conjecture is only known in certain special situations, such as for *CM abelian varieties*. For the L-functions in Tate's set-up, say of motivic weight $2m$, one does not even know that they make sense at the *Tate point* $s = m+1$, let alone know its order of pole there. Things get even harder if ρ_ℓ does not arise from a geometric situation. One cannot work in too general a setting, and at a minimum, one needs to require ρ_ℓ to have some good properties, such as being unramified outside a finite set S of primes. Fontaine and Mazur conjecture ([FoM]) that ρ_ℓ is *geometric* if it has this property (of being unramified outside a finite S) and is in addition *potentially semistable*.

1.2 Automorphic representations

Let F be a number field with adele ring $\mathbb{A}_F = F_\infty \times \mathbb{A}_{F,f}$, equipped with the adelic absolute value $|\cdot| = |\cdot|_{\mathbb{A}}$. For every algebraic group G over F, let $G(\mathbb{A}_F) = G(F_\infty) \times G(\mathbb{A}_{F,f})$ denote the restricted direct product $\prod_v' G(F_v)$, endowed with the usual locally compact topology. Then $G(F)$ embeds in $G(\mathbb{A}_F)$ as a discrete subgroup, and if Z_n denotes the center of $\mathrm{GL}(n)$, the homogeneous space $\mathrm{GL}(n,F)Z_n(\mathbb{A}_F)\backslash\mathrm{GL}(n,\mathbb{A}_F)$ has finite volume relative to the relatively invariant quotient measure induced by a Haar measure on $\mathrm{GL}(n,\mathbb{A}_F)$. An irreducible representation π of $\mathrm{GL}(n,\mathbb{A}_F)$ is admissible if it admits a factorization as a restricted tensor product $\otimes_v' \pi_v$, where each π_v is admissible and for almost all finite places v, π_v is *unramified*, i.e., has a no-zero vector fixed by $K_v = \mathrm{GL}(n,\mathcal{O}_v)$. (Here, as usual, $\mathcal{O}_v$ denotes the ring of integers of the local completion F_v of F at v.)

Fixing a unitary idele class character ω, which can be viewed as a character of $Z_n(\mathbb{A}_F)$, we may consider the space

$$L^2(n,\omega) := L^2(\mathrm{GL}(n,F)Z_n(\mathbb{A}_F)\backslash\mathrm{GL}(n,\mathbb{A}_F),\omega), \tag{1.2.1}$$

which consists of (classes of) functions on $\mathrm{GL}(n,\mathbb{A}_F)$ that are left-invariant under $\mathrm{GL}(n,F)$, transform under $Z_n(\mathbb{A}_F)$ according to ω, and are square-integrable modulo $\mathrm{GL}(n,F)Z(\mathbb{A}_F)$. Clearly, $L^2(n,\omega)$ is a unitary representation of $\mathrm{GL}(n,\mathbb{A}_F)$ under the right translation action on functions. The **space of cusp forms**, denoted $L_0^2(n,\omega)$, consists of functions φ in $L^2(n,\omega)$ which satisfy the following for *every* unipotent radical U of a standard parabolic subgroup $P = MU$:

$$\int_{U(F)\backslash U(\mathbb{A}_F)} \varphi(ux) = 0. \tag{1.2.2}$$

To say that P is a standard parabolic means that it contains the *Borel subgroup* of upper triangular matrices in $\mathrm{GL}(n)$. A basic fact asserts that $L_0^2(n,\omega)$ is a subspace of the discrete spectrum of $L^2(n,\omega)$.

By a **unitary cuspidal** (automorphic) representation π of $\mathrm{GL}_n(\mathbb{A}_F)$, we will mean an irreducible, unitary representation occurring in $L_0^2(n,\omega)$. We will, by abuse of notation, also denote the underlying admissible representation by π. (To be precise, the unitary representation is on the Hilbert space completion of the admissible space.) Roughly speaking, unitary automorphic representations of $\mathrm{GL}(n,\mathbb{A}_F)$ are those which appear weakly in $L^2(n,\omega)$ for some ω. We will refrain from recalling the definition precisely, because we will work totally with the subclass of *isobaric automorphic representations*, for which one can take Theorem 1.2.10 (of Langlands) below as their definition.

If π is an admissible representation of $\mathrm{GL}(n,\mathbb{A}_F)$, then for any $z \in \mathbb{C}$, we define the *analytic Tate twist* of π by z to be

$$\pi(z) := \pi \otimes |\cdot|^z, \tag{1.2.3}$$

where $|\cdot|^z$ denotes the 1-dimensional representation of $\mathrm{GL}(n, \mathbb{A}_F)$ given by

$$g \to |\det(g)|^z = e^{z\log(|\det(g)|)}.$$

Since the adelic absolute value $|\cdot|$ takes $\det(g)$ to a positive real number, its logarithm is well defined.

In general, by a *cuspidal automorphic representation*, we will mean an irreducible admissible representation of $\mathrm{GL}(n, \mathbb{A}_F)$ for which there exists a real number w, which we will call the *weight of* π, such that the Tate twist

$$\pi_u := \pi(w/2) \tag{1.2.4}$$

is a unitary cuspidal representation. Note that the central character of π and of its unitary *avatar* π_u are related as follows:

$$\omega_\pi = \omega_{\pi_u}|\cdot|^{-nw/2}, \tag{1.2.5}$$

which is easily checked by looking at the situation at the unramified primes, which suffices.

For any irreducible, automorphic representation π of $GL(n, \mathbb{A}_F)$, there is an associated L-function $L(s,\pi) = L(s,\pi_\infty)L(s,\pi_f)$, called the *standard* L-function ([J]) of π. It has an Euler product expansion

$$L(s,\pi) = \prod_v L(s,\pi_v), \tag{1.2.6}$$

convergent in a right-half plane. If v is an archimedean place, then one knows (cf. [La1]) how to associate a semisimple n-dimensional $\mathbb{C}$-representation $\sigma(\pi_v)$ of the Weil group W_{F_v}, and $L(s,\pi_v)$ identifies with $L(s,\sigma_v)$. We will *normalize* this correspondence $\pi_v \to \sigma(\pi_v)$ in such a way that it respects algebraicity. Moreover, if v is a finite place where π_v is unramified, there is a corresponding semisimple conjugacy class $A_v(\pi)$ (or $A(\pi_v)$) in $\mathrm{GL}(n,\mathbb{C})$ such that

$$L(s,\pi_v) = \det(1 - A_v(\pi)T)^{-1}|_{T=q_v^{-s}}. \tag{1.2.7}$$

We may find a diagonal representative $\mathrm{diag}(\alpha_{1,v}(\pi), \dots, \alpha_{n,v}(\pi))$ for $A_v(\pi)$, which is unique up to permutation of the diagonal entries. Let $[\alpha_{1,v}(\pi), \dots, \alpha_{n,v}(\pi)]$ denote the unordered n-tuple of complex numbers representing $A_v(\pi)$. Since $W_{F,v}^{\mathrm{ab}} \simeq F_v^*$, $A_v(\pi)$ clearly defines an abelian n-dimensional representation $\sigma(\pi_v)$ of $W_{F,v}$. If $\underline{1}$ denotes the trivial representation of $\mathrm{GL}(1,\mathbb{A}_F)$, which is cuspidal, we have

$$L(s,\underline{1}) = \zeta_F(s),$$

the Dedekind zeta function of F. (Strictly speaking, we should take $L(s,\underline{1}_f)$ on the left, since the right-hand side is missing the archimedean factor, but this is not serious.)

The fundamental work of Godement and Jacquet, when used in conjunction with the Rankin–Selberg theory (see 1.3 below), yields the following:

Theorem 1.2.8 ([J]). *Let $n \geq 1$, and π a non-trivial cuspidal automorphic representation of $GL(n, \mathbb{A}_F)$. Then $L(s, \pi)$ is entire. Moreover, for any finite set S of places of F, the incomplete L-function*

$$L^S(s, \pi) = \prod_{v \notin S} L(s, \pi_v)$$

is holomorphic and non-zero in $\Re(s) > w + 1$ if π has weight w. Moreover, there is a functional equation

$$L(w + 1 - s, \pi^\vee) = \varepsilon(s, \pi) L(s, \pi) \tag{1.2.9}$$

with

$$\varepsilon(s, \pi) = W(\pi) N_\pi^{(w+1)/2-s}.$$

Here N_π denotes the norm of the conductor $\mathcal{N}_\pi$ of π, and $W(\pi)$ is the root number of π.

Of course when $w = 0$, i.e., when π is unitary, the statement comes to a more familiar form. When $n = 1$, a π is simply an idele class character and this result is due to Hecke.

By the theory of Eisenstein series, there is a sum operation $\boxplus$ ([La2], [JS1]):

Theorem 1.2.10 ([JS1]). *Given any m-tuple of cuspidal automorphic representations $\pi_1, \ldots, \pi_m$ of $GL(n_1, \mathbb{A}_F), \ldots, GL(n_m, \mathbb{A}_F)$ respectively, there exists an irreducible, automorphic representation $\pi_1 \boxplus \cdots \boxplus \pi_m$ of $GL(n, \mathbb{A}_F)$, $n = n_1 + \cdots + n_m$, which is unique up to equivalence, such that for any finite set S of places,*

$$L^S(s, \boxplus_{j=1}^m \pi_j) = \prod_{j=1}^m L^S(s, \pi_j). \tag{1.2.11}$$

Call such a (Langlands) sum $\pi \simeq \boxplus_{j=1}^m \pi_j$, with each π_j cuspidal, an *isobaric automorphic*, or just *isobaric* (if the context is clear), representation. Denote by $\mathrm{ram}(\pi)$ the finite set of finite places where π is ramified, and let $\mathfrak{N}(\pi)$ be its conductor.

For every integer $n \geq 1$, set

$$\mathcal{A}(n, F) = \{\pi : \text{isobaric representation of } \mathrm{GL}(n, \mathbb{A}_F)\}/\simeq, \tag{1.2.12}$$

and

$$\mathcal{A}_0(n, F) = \{\pi \in \mathcal{A}(n, F) | \ \pi \text{ cuspidal}\}.$$

Put $\mathcal{A}(F) = \cup_{n \geq 1} \mathcal{A}(n, F)$ and $\mathcal{A}_0(F) = \cup_{n \geq 1} \mathcal{A}_0(n, F)$.

Remark 1.2.13. One can also define the analogs of $\mathcal{A}(n, F)$ for local fields F, where the "cuspidal" subset $\mathcal{A}_0(n, F)$ consists of essentially square-integrable representations of $\mathrm{GL}(n, F)$. See [La3] (or [Ra1]) for details.

Given any polynomial representation

$$r : \mathrm{GL}(n, \mathbb{C}) \to \mathrm{GL}(N, \mathbb{C}), \tag{1.2.14}$$

one can associate an L-function to the pair (π, r), for any isobaric automorphic representation π of $\mathrm{GL}(n, \mathbb{A}_F)$:

$$L(s, \pi; r) = \prod_v L(s, \pi_v; r), \tag{1.2.15}$$

in such a way that at any finite place v where π is *unramified* with residue field $\mathbb{F}_{q_v}$,

$$L(s, \pi_v; r) = \det(1 - A_v(\pi; r)T)^{-1}|_{T=q_v^{-s}}, \tag{1.2.16}$$

with

$$A_v(\pi; r) = r(A_v(\pi)). \tag{1.2.17}$$

The conjugacy class $A_v(\pi; r)$ in $\mathrm{GL}(N, \mathbb{C})$ is again represented by an unordered N-tuple of complex numbers.

The **Principle of Functoriality** predicts the existence of an isobaric automorphic representation $r(\pi)$ of $\mathrm{GL}(N, \mathbb{A}_F)$ such that

$$L(s, r(\pi)) = L(s, \pi; r). \tag{1.2.18}$$

A weaker form of the conjecture, which suffices for questions like those we are considering, asserts that this identity holds outside a finite set S of places.

This conjecture is known in the following cases of $(\mathbf{n}, \mathbf{r})$:

$$\begin{aligned} &(\mathbf{2}, \mathrm{sym}^{\mathbf{2}}): \ \text{Gelbart–Jacquet ([GJ])} \\ &(\mathbf{2}, \mathrm{sym}^{\mathbf{3}}): \ \text{Kim–Shahidi ([KSh1])} \\ &(\mathbf{2}, \mathrm{sym}^{\mathbf{4}}): \ \text{Kim ([K])} \\ &(\mathbf{4}, \Lambda^{\mathbf{2}}): \ \text{Kim ([K]).} \end{aligned} \tag{1.2.19}$$

In this paper we will make use of the last instance of functoriality, namely the *exterior square* transfer of automorphic forms from GL(4) to GL(6).

1.3 Rankin–Selberg L-functions

The results here are due to the independent and partly complementary, deep works of Jacquet, Piatetski-Shapiro and Shalika, and of Shahidi. Let π, π' be isobaric automorphic representations in $\mathcal{A}(n, F)$, $\mathcal{A}(n', F)$ respectively. Then there exists an associated Euler product $L(s, \pi \times \pi')$ ([JPSS], [JS1], [Sh1, Sh2], [MW], [CoPS2]), which converges in $\{\Re(s) > 1\}$, and admits a meromorphic continuation to the whole s-plane and satisfies the functional equation, which is given in the unitary case by

$$L(s, \pi \times \pi') = \varepsilon(s, \pi \times \pi')L(1 - s, \pi^\vee \times \pi'^\vee), \tag{1.3.1}$$

with

$$\varepsilon(s, \pi \times \pi') = W(\pi \times \pi')N(\pi \times \pi')^{\frac{1}{2}-s},$$

where the *conductor* $N(\pi \times \pi')$ is a positive integer not divisible by any rational prime not intersecting the ramification loci of $F/\mathbb{Q}$, π and π', while $W(\pi \times \pi')$ is the *root number* in $\mathbb{C}^*$. As in the Galois case, $W(\pi \times \pi')W(\pi^\vee \times \pi'^\vee) = 1$, so that $W(\pi \times \pi') = \pm 1$ when π, π' are selfdual.

It is easy to deduce from this the functional equation when π, π' are not unitary. If they are cuspidal of weights w, w' respectively, the functional equation relates s to $w + w' + 1 - s$. Moreover, since $\pi^\vee, \pi'^\vee$ have respective weights $-w, -w'$, $\pi \times \pi^\vee$ and $\pi' \times \pi'^\vee$ still have weight 0.

When v is archimedean or a finite place unramified for π, π',

$$L_v(s, \pi \times \pi') = L(s, \sigma(\pi_v) \otimes \sigma(\pi'_v)). \tag{1.3.2}$$

In the archimedean situation, $\pi_v \to \sigma(\pi_v)$ is the arrow to the representations of the Weil group W_{F_v} given by [La1]. When v is an unramified finite place, $\sigma(\pi_v)$ is defined in the obvious way as the sum of one dimensional representations defined by the Langlands class $A(\pi_v)$.

When $n = 1$, $L(s, \pi \times \pi') = L(s, \pi\pi')$, and when $n = 2$ and $F = \mathbb{Q}$, this function is the usual Rankin–Selberg L-function, extended to arbitrary global fields by Jacquet.

Theorem 1.3.3 ([JS1], [JPSS]). *Let $\pi \in \mathfrak{A}_0(n, F)$, $\pi' \in \mathfrak{A}_0(n', F)$, and S a finite set of places. Then $L^S(s, \pi \times \pi')$ is entire unless π is of the form $\pi'^\vee \otimes |.|^w$, in which case it is holomorphic outside $s = w, 1 - w$, where it has simple poles.*

The *Principle of Functoriality* implies in this situation that given π, π' as above, there exists an isobaric automorphic representation $\pi \boxtimes \pi'$ of $\mathrm{GL}(nn', \mathbb{A}_F)$ such that

$$L(s, \pi \boxtimes \pi') = L(s, \pi \times \pi'). \tag{1.3.4}$$

The (conjectural) functorial product $\boxtimes$ is the automorphic analogue of the usual tensor product of Galois representations. For the importance of this product, see [Ra1], for example.

One can always construct $\pi \boxtimes \pi'$ as an *admissible* representation of $\mathrm{GL}(nn', \mathbb{A}_f)$, but the subtlety lies in showing that this product is automorphic.

The automorphy of $\boxtimes$ is known in the following cases, which will be useful to us:

$$\begin{aligned} (\mathbf{n}, \mathbf{n}') = (\mathbf{2}, \mathbf{2})&: \text{ Ramakrishnan ([Ra2])} \\ (\mathbf{n}, \mathbf{n}') = (\mathbf{2}, \mathbf{3})&: \text{ Kim–Shahidi ([KSh1]).} \end{aligned} \tag{1.3.5}$$

The reader is referred to Section 11 of [Ra4], which contains some refinements, explanations, refinements and (minor) errata for [Ra2]. It may be worthwhile remarking that Kim and Shahidi use the functorial product on GL(2)×GL(3) which they construct to prove the *symmetric cube* lifting for GL(2) mentioned in the previous section (see (1.2.11). A *cuspidality criterion* for the image under this transfer is proved in [Ra-W], with an application to the cuspidal cohomology of congruence subgroups of $\mathrm{SL}(6, \mathbb{Z})$.

1.4 Modularity and the problem at hand

The general Langlands philosophy asserts that if ρ_ℓ is an n-dimensional ℓ-adic representation of $\mathcal{G}_{\mathbb{Q}}$, then there is an isobaric automorphic representation π of $\mathrm{GL}(n, \mathbb{A})$ such that for a suitable finite set S of places (including ∞), we have an identity of the form

$$L^S(s, \rho_\ell) = L^S(s, \pi). \tag{1.4.1}$$

When this happens, we will say that ρ_ℓ is *modular*, and we will write

$$\rho_\ell \leftrightarrow \pi. \tag{1.4.2}$$

One says that ρ_ℓ is *strongly modular* if the identity (1.4.1) holds for the full L-function, i.e., with S empty.

Special cases of this conjecture were known earlier, the most famous one being the modularity conjecture for the ℓ-adic representations ρ_ℓ defined by the Galois action on the ℓ-power division points of elliptic curves E over $\mathbb{Q}$, proved recently in the spectacular works of Wiles, Taylor, Diamond, Conrad and Breuil.

We will not consider any such (extremely) difficult question in this article. Instead we will be interested in the following:

Question 1.4.3. When a modular ρ_ℓ is irreducible, is the corresponding π cuspidal? And conversely?

This seemingly reasonable question turns out to be hard to check in dimensions $n > 2$.

One thing that is clear is that the π associated to any ρ_ℓ needs to be *algebraic* in the sense of Clozel ([Cℓ1]). To define the notion of *algebraicity*, first recall that by Langlands, the archimedean component π_∞ is associated to an n-dimensional representation σ_∞, sometimes written σ_∞, of the real Weil group $W_{\mathbb{R}}$, with corresponding equality of the archimedean L-factors $L_\infty(s, \rho_\ell)$ and $L(s, \pi_\infty)$. We will normalize things so that the correspondence is algebraic. One can explicitly describe $W_{\mathbb{R}}$ as $\mathbb{C}^* \cup j\mathbb{C}^*$, with $jzj^{-1} = \overline{z}$ and $j^2 = -1$. One gets a canonical exact sequence

$$1 \to \mathbb{C}^* \to W_{\mathbb{R}} \to \mathcal{G}_{\mathbb{R}} \to 1 \tag{1.4.4}$$

which represents the unique non-trivial extension of $\mathcal{G}_{\mathbb{R}}$ by $\mathbb{C}^*$. One has a decomposition

$$\sigma_\infty|_{\mathbb{C}^*} \simeq \oplus_{j=1}^n \xi_j, \tag{1.4.5}$$

where each ξ_j is a (quasi-)character of $\mathbb{C}^*$. One says that π is *algebraic* when every one of the characters χ_j is algebraic, i.e., there are integers p_j, q_j such that

$$\chi_j(z) = z^{p_j}\overline{z}^{q_j}. \tag{1.4.6}$$

This analogous to having a *Hodge structure*, which is what one would expect if π were to be related to a geometric object.

One says that π is *regular* if for all $i \neq j$, $\chi_i \neq \chi_j$. In other words, each character χ_j appears in the restriction of $\sigma_\infty(\pi)$ to $\mathbb{C}^*$ with *multiplicity one*. We say (following [BHR]) that π is *semi-regular* if the multiplicity of each χ_j is at most 2.

When $n = 2$, any π defined by a classical *holomorphic newform* f of weight $k \geq 1$ is algebraic and semi-regular. It is regular iff $k \geq 2$. One also expects any *Maass waveform* φ of *weight* 0 and *eigenvalue* $1/4$ for the *hyperbolic Laplacian* to be algebraic; there are interesting examples of this kind coming from the work of Langlands (resp. Tunnell) on tetrahedral (resp. octahedral) Galois representations ρ which are *even*; the *odd* ones correspond to holomorphic newforms of weight 1. We will not consider the even situation in this article.

Given a holomorphic newform $f(z) = \sum_{n=1}^{\infty} a_n q^n$, $q = e^{2\pi i z}$, of weight 2, resp. $k \geq 3$, resp. $k = 1$, level N and character ω, one knows by Eichler and Shimura, resp. Deligne ([De]), resp Deligne–Serre ([DeS]), that there is a continuous, irreducible representation

$$\rho_\ell : \mathcal{G}_{\mathbb{Q}} \to \mathrm{GL}(2, \mathbb{Q}_\ell) \tag{1.4.7}$$

such that for all primes $p \nmid N\ell$,

$$\mathrm{tr}(Fr_p | \rho_\ell) = a_p$$

and

$$\det(\rho_\ell) = \omega \chi_\ell^{k-1},$$

where χ_ℓ is the ℓ-adic *cyclotomic character* of $\mathcal{G}_{\mathbb{Q}}$, given by the Galois action on the ℓ-power roots of unity in $\overline{\mathbb{Q}}$, and Fr_p is the *geometric Frobenius* at p, which is the inverse of the arithmetic Frobenius.

1.5 Parity

We will first first introduce this crucial concept over the base field $\mathbb{Q}$, as that is what is needed in the remainder of the article.

We will need to restrict our attention to those isobaric forms π on $\mathrm{GL}(n)/\mathbb{Q}$ which are odd in a suitable sense. It is instructive to first consider the case of a classical holomorphic newform f of weight $k \geq 1$ and character ω relative to the congruence subgroup $\Gamma_0(N)$. Since $\Gamma_0(N)$ contains $-I$, it follows that $\omega(-1) = (-1)^k$. One could be tempted to call a π defined by such an f to be even (or odd) according as ω is even (or odd), but it would be a wrong move. One should look not just at ω, but at the determinant of the associated ρ_ℓ, i.e., the ℓ-adic character $\omega\chi_\ell^{k-1}$, which is odd for all k ! So all such π defined by holomorphic newforms are *arithmetically* odd. The only even ones for GL(2) are (analytic Tate twists of) Maass forms of weight 0 and Laplacian eigenvalue $1/4$.

The maximal abelian quotient of $W_{\mathbb{R}}$ is $\mathbb{R}^*$, and the restriction of the abelianization map to $\mathbb{C}^*$ identifies with the norm map $z \to |z|$. So every (quasi)-character ξ of $W_{\mathbb{R}}$ identifies with one of $\mathbb{R}^*$, given by $x \to sgn(x)^a |x|^t$ for some t, with $a \in \{0, 1\}$. Clearly, ξ determines, and is determined by (t, a). If π is an isobaric automorphic representation, let $\sigma_\infty[\xi]$ denote, for each such ξ, the ξ-isotypic component of σ_∞. The *sign group* $\mathbb{R}^*/\mathbb{R}_+^*$ acts on each isotypic component. Let $m_+(\pi, \xi)$

(resp. $m_-(\pi,\xi)$) denote the multiplicity of the eigenvalue $+1$ (resp. -1), under the action of $\mathbb{R}^*/\mathbb{R}^*_+$ on $\sigma_\infty(\xi)$.

Definition 1.5.1. *Call an isobaric automorphic representation π of $GL(n,\mathbb{A})$* odd *if for every one-dimensional representation ξ of $W_\mathbb{R}$ occurring in σ_∞,*

$$|m_+(\pi,\xi) - m_-(\pi,\xi)| \leq 1.$$

Clearly, when the dimension of $\sigma_\infty[\xi]$ is even, the multiplicity of $+1$ as an eigenvalue of the sign group needs to be equal to the multiplicity of -1 as an eigenvalue.

Under this definition, all forms on $\mathrm{GL}(1)/\mathbb{Q}$ are odd. So are the π on $\mathrm{GL}(2)/\mathbb{Q}$ which are defined by holomorphic newforms of weight $k \geq 2$. The reason is that π_∞ is (for $k \geq 2$) a discrete series representation, and the corresponding $\sigma_\infty(\pi)$ is an irreducible 2-dimensional representation of $W_\mathbb{R}$ induced by the (quasi)-character $z \to z^{-(k-1)}$ of the subgroup $\mathbb{C}^*$ of index 2, and our condition is vacuous. On the other hand, if $k = 1$, $\sigma_\infty(\pi)$ is a reducible 2-dimensional representation, given by $\underline{1} \oplus sgn$. The eigenvalues are 1 on $\sigma_\infty(\underline{1})$ and -1 on $\sigma_\infty(sgn)$. On the other hand, a Maass form of weight 0 and $\lambda = 1/4$, the eigenvalue 1 (or -1) occurs with multiplicity 2, making the π it defines an even representation. So our definition is a good one and gives what we know for $n = 2$.

For any n, note that if π is algebraic and regular, it is automatically odd. If π is algebraic and semi-regular, each isotypic space is one or two-dimensional, and in the latter case, we want both eigenvalues to occur for π to be odd.

Finally, if F is any number field with a real place u, we can define, in exactly the same way, when an algebraic, isobaric automorphic representation of $\mathrm{GL}(n,\mathbb{A}_F)$ is *arithmetically odd at u*. If F is *totally real*, then we say that π is *totally odd* if it is so at *every* archimedean place.

2 The first step in the proof

Let ρ_ℓ, π be as in Theorem A. Since ρ_ℓ is irreducible, it is in particular semisimple. Suppose π is not cuspidal. We will obtain a contradiction.

Proposition 2.1. *Let ρ_ℓ, π be associated, with π algebraic, semi-regular and odd. Suppose we have, for some $r > 1$, an isobaric sum decomposition*

$$\pi \simeq \boxplus_{j=1}^r \eta_j, \tag{2.2}$$

where each η_j is a cuspidal automorphic representation of $GL(n_j,\mathbb{A})$, with $n_j \leq 2$ $(\forall j)$. Then ρ_ℓ cannot be irreducible.

Corollary 2.3. *Theorem A holds when π admits an isobaric sum decomposition such as* (2.2) *with each $n_j \leq 2$. In particular, it holds for $n \leq 3$.*

Proof of Proposition. The hypothesis that π is algebraic and semi-regular implies easily that each η_j is also algebraic and semi-regular. Let J_m denote the set of j where $n_j = m$.

First look at any j in J_1. Then the corresponding η_j is an idele class character. Its algebraicity implies that, in classical terms, it corresponds to an algebraic Hecke character ν_j. By Serre ([Se]), we may attach an abelian ℓ-adic representation $\nu_{j,\ell}$ of $\mathcal{G}_{\mathbb{Q}}$ of dimension 1. It follows that for some finite set S of places containing ℓ,

$$L^S(s, \nu_{j,\ell}) = L^S(s, \eta_j) \quad \text{whenever} \quad n_j = 1. \tag{2.4}$$

Next consider any j in J_2. If $\sigma(\eta_{j,\infty})$ is irreducible, then a twist of η_j must correspond to a classical holomorphic newform f of weight $k \geq 2$. Moreover, the algebraicity of η_j forces this twist to be algebraic. Hence by Deligne, there is a continuous representation

$$\tau_{j,\ell} : \mathcal{G}_{\mathbb{Q}} \to \mathrm{GL}(2, \overline{\mathbb{Q}}_\ell), \tag{2.5}$$

ramified only at a finite of primes such that at every $p \neq \ell$ where the representation is unramified,

$$\mathrm{tr}(Fr_p | \tau_{j,\ell}) = a_p(\eta_j), \tag{2.6}$$

and the determinant of $\tau_{j,\ell}$ corresponds to the central character ω_j of η_j. Moreover, $\tau_{j,\ell}$ is irreducible, which is not crucial to us here.

We also need to consider the situation, for any fixed $j \in J_2$, when $\sigma(\eta_{j,\infty})$ is reducible, say of the form $\chi_1 \oplus \chi_2$. Since η_j is cuspidal, by the *archimedean purity* result of Clozel ([Cℓ1]), $\chi_1\chi_2^{-1}$ must be 1 or sgn. The former cannot happen due to the oddness of π. It follows that η_j is defined by a classical holomorphic newform f of weight 1, and by a result of Deligne and Serre ([DeS]), there is a 2-dimensional ℓ-adic representation $\tau_{j,\ell}$ of $\mathcal{G}_{\mathbb{Q}}$ *with finite image*, which is irreducible, such that (2.6) holds.

Since the set of Frobenius classes Fr_p, as p runs over primes outside S, is dense in the Galois group by Tchebotarev, we must have, by putting all these cases together,

$$\rho_\ell \simeq (\oplus_{j \in J_1} \nu_{j,\ell}) \oplus (\oplus_{j \in J_2} \tau_{j,\ell}), \tag{2.7}$$

which contradicts the irreducibility of ρ_ℓ, since by hypothesis, $r = |J_1| + |J_2| \geq 2$.
□

3 The second step in the proof

Let ρ_ℓ, π be as in Theorem A. Suppose π is not cuspidal. In view of Proposition 2.1, we need only consider the situation where π is an isobaric sum $\boxplus_j \eta_j$, with an η_j being a cusp form on $\mathrm{GL}(m)/\mathbb{Q}$ for some $m \geq 3$.

Proposition 3.1. *Let ρ_ℓ, π be associated, with π an algebraic cusp form on $GL(n)/\mathbb{Q}$ which is semi-regular and odd. Suppose we have an isobaric sum decomposition*

$$\pi \simeq \eta \boxplus \eta', \tag{3.2}$$

where η is a cusp form on $GL(3)/\mathbb{Q}$ and η' is an isobaric automorphic representation of $GL(r, \mathbb{A})$ for some $r \geq 1$. Moreover, assume that there is an r-dimensional ℓ-adic representation τ'_ℓ of $\mathcal{G}_\mathbb{Q}$ associated to η'. Then we have the isomorphism of $\mathcal{G}_\mathbb{Q}$-modules:

$$\rho_\ell^\vee \oplus (\rho_\ell \otimes \tau'_\ell) \simeq \Lambda^2(\rho_\ell) \oplus \tau_\ell'^\vee \oplus \mathrm{sym}^2(\tau'_\ell). \tag{3.3}$$

Corollary 3.4. *Let ρ_ℓ, π be associated, with π algebraic, semi-regular and odd. Suppose π admits an isobaric sum decomposition such as* (3.2) *with $r \leq 2$. Then ρ_ℓ is reducible.*

Proposition 3.1 $\Longrightarrow$ Corollary 3.4. When $r \leq 2$, η' is either an isobaric sum of algebraic Hecke characters or cuspidal, in which case, thanks to the oddness, it is defined by a classical cusp form on $\mathrm{GL}(r)/\mathbb{Q}$ of weight ≥ 1. In either case we have, as seen in the previous section, the existence of the associated ℓ-adic representation τ'_ℓ, which is irreducible exactly when η_ℓ is cuspidal. Then by the proposition, the decomposition (3.3) holds. If $r = 1$ or $r = 2$ with η' Eisensteinian, (3.3) implies that a 1-dimensional representation (occurring in τ'_ℓ) is a summand of a twist of either ρ_ℓ or $\rho_\ell^\vee$. Hence the corollary.

□

Combining Corollary 2.3 and Corollary 3.4, we see that the irreducibility of ρ_ℓ forces the corresponding π to be cuspidal when $n \leq 4$ under the hypotheses of Theorem A. So we obtain the following:

Corollary 3.5. *Theorem A holds for $n \leq 4$.*

Proof of Proposition 3.1. By hypothesis, we have a decomposition as in (3.2), and an ℓ-adic representation τ'_ℓ associated to η'.

As a short digression let us note that if η were essentially selfdual and regular, we could exploit its algebraicity, and by appealing to [Pic] associate a 3-dimensional ℓ-adic representation to η. The Proposition 3.1 will follow in that case, as in the proof of Proposition 2.1. However, we cannot (and do not wish to) assume either that η is essentially selfdual or that it is regular. We have to appeal to another idea, and here it is.

Let S be a finite set of primes including the archimedean and ramified ones. At any p outside S, let π_p be represented by an unordered $(3+r)$-tuple $\{\alpha_1, \ldots, \alpha_{3+r}\}$ of complex numbers, and we may assume that η_p (resp. η'_p) is represented by $\{\alpha_1, \alpha_2, \alpha_3\}$ (resp. $\{\alpha_4, \ldots, \alpha_{3+r}\}$. It is then straightforward to check that

$$L(s, \pi_p; \Lambda^2) = L(s, \eta_p^\vee) L(s, \eta_p \times \eta'_p) L(s, \eta'; \Lambda^2). \tag{3.6}$$

One can also deduce this as follows. Let $\sigma(\beta)$ denote, for any irreducible admissible representation β of $\mathrm{GL}(m, \mathbb{Q}_p)$, the m-dimensional representation of the extended Weil group $W'_{\mathbb{Q}_p} = W_{\mathbb{Q}_p} \times \mathrm{SL}(2, \mathbb{C})$ defined by the local Langlands correspondence (cf. [HaT], [He]). For any representation σ of $W'_{\mathbb{Q}_p}$ which splits as a direct sum $\tau \oplus \tau'$, we have

$$\Lambda^2(\sigma) \simeq \Lambda^2(\tau) \oplus (\tau \otimes \tau') \oplus \Lambda^2(\tau'), \tag{3.7}$$

with $\Lambda^2(\tau') = 0$ if τ' is 1-dimensional and

$$\Lambda^2(\tau) \simeq \tau^\vee \quad \text{when} \quad \dim(\tau) = 3. \tag{3.8}$$

In fact, this shows that the identity (3.6) works at the ramified primes as well, but we do not need it.

Now, since $\pi^\vee \simeq \eta^\vee \boxplus \eta'^\vee$, we get by putting (3.2) and (3.6) together,

$$L(s, \pi_p^\vee)L(s, \pi_p \times \eta'_p)L(s, \eta'_p; \Lambda^2) = L(s, \pi_p; \Lambda^2)L(s, {\eta'_p}^\vee)L(s, \eta'_p \times \eta'_p). \tag{3.9}$$

Appealing to Tchebotarev, and using the correspondences $\pi \leftrightarrow \rho_\ell$ and $\eta' \leftrightarrow \tau'_\ell$, we obtain the following isomorphism of $\mathcal{G}_{\mathbb{Q}}$-representations:

$$\rho_\ell^\vee \oplus (\rho_\ell \otimes \tau'_\ell) \oplus \Lambda^2(\tau'_\ell) \simeq \Lambda^2(\rho_\ell) \oplus \tau_\ell'^\vee \oplus (\tau'_\ell \otimes \tau'_\ell). \tag{3.10}$$

Using the decomposition

$$\tau'_\ell \otimes \tau'_\ell \simeq \mathrm{sym}^2(\tau'_\ell) \oplus \Lambda^2(\tau'_\ell),$$

we then obtain (3.3) from (3.10). □

4 Galois representations attached to regular, selfdual cusp forms on GL(4)

A cusp form Π on $\mathrm{GL}(m)/F$, F a number field, is said to be *essentially selfdual* iff $\Pi^\vee \simeq \Pi \otimes \lambda$ for an idele class character λ; it is *selfdual* if $\lambda = 1$. We will call such a λ a *polarization*. Let us call Π *almost selfdual* if there is a polarization λ of the form $\mu^2|\cdot|^t$ for some $t \in \mathbb{C}$ and a finite order character μ; in this case, one sees that $(\Pi \otimes \mu)^\vee \simeq \Pi \otimes \mu|\cdot|^t$, or equivalently, $\Pi \otimes \mu|\cdot|^{t/2}$ is selfdual. Clearly, if Π is essentially selfdual, then it becomes, under base change ([AC]), almost selfdual over a finite cyclic extension K of F.

Note that when Π is essentially selfdual relative to λ, it is immediate that λ_∞ occurs in the isobaric sum decomposition of $\Pi_\infty \boxtimes \Pi_\infty$, or equivalently, $\sigma(\lambda_\infty)$ is a constituent of $\sigma(\Pi_\infty)^{\otimes 2}$. This implies that if Π is algebraic, then so is Λ, and thus corresponds to an ℓ-adic character λ_ℓ of $\mathcal{G}_{\mathbb{Q}}$.

Whether or not Π is algebraic, we have, for any S,

$$L^S(s, \Pi \times \Pi \otimes \lambda^{-1}) = L^S(s, \Pi, \mathrm{sym}^2 \otimes \lambda^{-1})L^S(s, \Pi; \Lambda^2 \otimes \lambda^{-1}). \tag{4.1}$$

The L-function on the left has a pole at $s = 1$, since $\Pi^\vee \simeq \Pi \otimes \lambda$ by hypothesis. Also, neither of the L-functions on the right is zero at $s = 1$ ([JS2]). Consequently, exactly one of the L-functions on the right of (4.1) admits a pole at $s = 1$. One says that Π is of *orthogonal type* ([Ra3]), resp. *symplectic type*, if $L^S(s, \Pi, \mathrm{sym}^2 \otimes \lambda^{-1})$, resp. $L^S(s, \Pi, \Lambda^2 \otimes \lambda^{-1})$ admits a pole at $s = 1$.

The following result is a consequence of a synthesis of the results of a number of mathematicians, and it will be crucial to us in the next section, while proving Theorem A for $n = 5$.

Theorem B. *Let* Π *be a regular, algebraic cusp form on* $GL(4)/\mathbb{Q}$*, which is almost selfdual. Then there exists a continuous representation*

$$R_\ell : \mathcal{G}_{\mathbb{Q}} \to \mathrm{GL}(4, \overline{\mathbb{Q}}_\ell),$$

such that

$$L^S(s, \Pi; \Lambda^2) = L^S(s, \Lambda^2(R_\ell)),$$

for a finite set S of primes containing the ramified ones. Moreover, if Π *is of orthogonal type, we can show that* R_ℓ *and* Π *are associated, i.e., have the same degree* 4 *L-functions (outside S).*

When Π admits a discrete series component Π_p at some (finite) prime p, a stronger form of this result, and in fact its generalization to $\mathrm{GL}(n)/\mathbb{Q}$, is due to Clozel ([C$\ell$2]). But in the application considered in the next section, we will not be able to satisfy such a ramification assumption at a finite place.

In the *orthogonal case*, Π descends by the work of Ginzburg–Rallis–Soudry (cf. [So]) to define a regular cusp form β on the split $\mathrm{SGO}(4)/\mathbb{Q}$, which is given by a pair (π_1, π_2) of regular cusp forms on $\mathrm{GL}(2)/\mathbb{Q}$. By Deligne, there are 2-dimensional (irreducible) ℓ-adic representations $\tau_{1,\ell}, \tau_{2,\ell}$, with $\tau_{j,\ell} \leftrightarrow \pi_j$, $j = 1, 2$. This leads to the desired 4-dimensional $\overline{\mathbb{Q}}_\ell$-representation $R_\ell := \tau_{1,\ell} \otimes \tau_{2,\ell}$ of $\mathcal{G}_{\mathbb{Q}}$ associated to Π, such that

$$L^S(s, R_\ell) = L^S(s, \Pi). \tag{4.2}$$

It may be useful to notice that since the polarization is a square (under the almost selfduality assumption), the associated Galois representation takes values in $\mathrm{SGO}(4, \overline{\mathbb{Q}}_\ell)$, which is the connected component of $\mathrm{GO}(4, \overline{\mathbb{Q}}_\ell)$, with quotient $\{\pm 1\}$. In the general case, not needed for this article, R_ℓ will need to be either of the type above or of *Asai type* (see [Ra4]), associated to a 2-dimensional $\overline{\mathbb{Q}}_\ell$-representation $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ for a quadratic extension $K/\mathbb{Q}$.

In the (more subtle) *symplectic case*, this theorem is proved in my joint work [Ra-Sh] with F. Shahidi. We will start with a historical comment and then sketch the proof (for the benefit of the reader). Some years ago, Jacquet, Piatetski-Shapiro and Shalika announced a theorem, asserting that one could descend any Π (of symplectic type on $\mathrm{GL}(4)/\mathbb{Q}$) to a generic cusp form β on $\mathrm{GSp}(4)/\mathbb{Q}$ with the same (incomplete) degree 4 L-functions. Unfortunately, this work was never published, except for part of it in [JSh2]. In [Ra-Sh], Shahidi and I provide an alternate, somewhat more circuitous route, yielding something slightly weaker, but sufficient for many purposes. Here is the idea. We begin by considering the twist $\Pi_0 := \Pi \otimes \mu|\cdot|^{t/2}$ instead of Π, to make the polarization is trivial, i.e., so that Π_0 has parameter in $\mathrm{Sp}(4, \mathbb{C})$. Using the *backwards lifting* results of [GRS] (see also [So]), we get a generic cusp form Π'

on the split SO(5)/$\mathbb{Q}$, such that $\Pi_0 \to \Pi'$ is functorial at the archimedean and unramified places. Using the isomorphism of PSp(4)/$\mathbb{Q}$ with SO(5)/$\mathbb{Q}$, we may lift Π' to a generic cusp form $\tilde{\Pi}'$ on Sp(4)/$\mathbb{Q}$. By a suitable extension followed by induction, we can associate a generic cusp form Π_1 on GSp(4)/$\mathbb{Q}$, such that the following hold:

(i) The archimedean parameter of $\Pi_2 := \Pi_1 \otimes \mu^{-1}|\cdot|^{-t/2}$ is algebraic and regular; and
(ii) $L(s, \Pi_{2,p}; \Lambda^2) = L(s, \Pi_p; \Lambda^2)$ at any prime p where Π is unramified.

(4.3)

We in fact deduce a stronger statement in [Ra-Sh], involving also the ramified primes, but it is not necessary for the application considered in this paper. To continue, part (i) of (4.3) implies that Π_2 contributes to the (intersection) cohomology of (the Baily–Borel–Satake compactification over $\mathbb{Q}$ of) the 3-dimensional Shimura variety $Sh_K/\mathbb{Q}$ associated to GSp(4)/$\mathbb{Q}$, relative to a compact open subgroup K of GSp(4, $\mathbb{A}_f$); Sh_K parametrizes principally polarized abelian surfaces with level K-structure. Now by appealing to the deep (independent) works of G. Laumon ([Lau1, Lau2]) and R. Weissauer ([Wei]), one gets a continuous 4-dimensional ℓ-adic representation R_ℓ of $\mathcal{G}_{\mathbb{Q}}$ such that

$$L^S(s, \Pi_2) = L^S(s, R_\ell). \tag{4.4}$$

The assertion of Theorem B now follows by combining (4.3)(ii) and (4.4). □

5 Two useful lemmas on cusp forms on GL(4)

Let F be a number field and η a cuspidal automorphic representation of GL(4, $\mathbb{A}_F$), where $\mathbb{A}_F := \mathbb{A} \otimes_{\mathbb{Q}} F$ is the Adele ring of F. Denote by ω_η the central character of η.

First let us recall (see (1.2.11)) that by a difficult *theorem of H. Kim* ([K]), there is an isobaric automorphic form $\Lambda^2(\eta)$ on GL(6)/$\mathbb{Q}$ such that

$$L(s, \Lambda^2(\eta)) = L(s, \eta; \Lambda^2). \tag{5.1}$$

Lemma 5.2. *$\Lambda^2(\eta)$ is essentially selfdual. In fact*

$$\Lambda^2(\eta)^\vee \simeq \Lambda^2(\eta) \otimes \omega_\eta^{-1}. \tag{5.3}$$

Proof. Thanks to the strong multiplicity one theorem for isobaric automorphic representations ([JS1]), it suffices to check this at the primes p where η is unramified. Fix any such p, and represent the semisimple conjugacy class $A_p(\eta)$ by $[a, b, c, d]$. Then it is easy to check that

$$A_p(\Lambda^2(\eta)) = \Lambda^2(A_p(\eta)) = [ab, ac, ad, bc, bd, cd]. \tag{5.4}$$

Since for any automorphic representation Π, the unordered tuple representing $A_p(\Pi^\vee)$ consists of the inverses of the elements of tuple representing $A_p(\pi)$, and since $A_p(\omega_\eta) = [abcd]$, we have

$$A_p(\Lambda^2(\eta)^\vee \otimes \omega_\eta) = [(ab)^{-1}, (ac)^{-1}, (ad)^{-1}, (bc)^{-1}, (bd)^{-1}, (cd)^{-1}] \otimes [abcd], \tag{5.5}$$

which is none other than $A_p(\Lambda^2(\eta))$. The isomorphism (5.3) follows. □

Lemma 5.6. *Let η be a cusp form on $GL(4)/F$ with trivial central character. Suppose $\eta^\vee \not\simeq \eta$. Then there are infinitely many primes P in $\mathcal{O}_F$ where η_P is unramified such that* 1 *is not an eigenvalue of the conjugacy class $A_P(\Lambda^2(\eta))$ of $\Lambda^2(\eta_P)$.*

Proof of Lemma 5.6. Since $\eta^\vee \not\simeq \eta$, there exist, by the strong multiplicity one theorem, infinitely many unramified primes P where $\eta_P^\vee \not\simeq \eta_P$. Pick any such P, and write

$$A_P(\eta) = [a, b, c, d], \quad \text{with} \quad abcd = 1.$$

The fact that $\eta_P^\vee \not\simeq \eta_P$ implies that the set $\{a, b, c, d\}$ is not stable under inversion. Hence one of its elements, which we may assume to be a after renaming, satisfies the following:

$$a \notin \{a^{-1}, b^{-1}, c^{-1}, d^{-1}\}.$$

Equivalently,

$$1 \notin \{a^2, ab, ac, ad\}.$$

On the other hand, we have (5.4), which is used to conclude that the only way 1 can be in this set (attached to $\Lambda^2(\eta_P)$) is to have either bc or bd or cd to be 1. But if $bc = 1$ (resp. $bd = 1$), since $abcd = 1$, we must have $ad = 1$ (resp. $ac = 1$), which is impossible. Similarly, if $cd = 1$, we are forced to have $ab = 1$, which is also impossible. □

6 Finale

Let ρ_ℓ, π be as in Theorem A. In view of Corollary 3.5, we may assume henceforth that $n = 5$, and that π is algebraic and regular. Suppose π is not cuspidal. In view of Corollary 2.3 and Corollary 3.4, we must then have the decomposition

$$\pi \simeq \eta \boxplus \nu, \tag{6.1}$$

where η is an algebraic, regular cusp form on $\mathrm{GL}(4)/\mathbb{Q}$ and ν is an algebraic Hecke character, with associated ℓ-adic character ν_ℓ.

Note that Theorem A needs to be proved under *either* of two hypotheses. To simplify matters a bit, we will make use of the following:

Lemma 6.2. *There is a character ν_0 with $\nu_0^2 = 1$ such that for $\mu = \nu_0\nu^{-1}$, if π is almost selfdual, then so is $\pi \otimes \mu^{-1}$.*

Proof of lemma. When π is almost selfdual, there exists, by definition, an idele class character μ such that $\pi \otimes \mu$ is selfdual. But this implies, thanks to (6.1) and the cuspidality of η, that $\eta \otimes \mu$ is selfdual and $\mu\nu$ is 1 or quadratic. We are done by taking $\nu_0 = \mu\nu$. □

Consequently, we may, and we will, replace π by $\pi \otimes \mu$, ρ_ℓ by $\rho_\ell \otimes \mu_\ell$, η by $\eta \otimes \mu$ and ν by $\nu\mu$, without jeopardizing the nature of either of the hypotheses of Theorem A. In fact, the *first hypothesis simplifies to assuming that π is selfdual.* Moreover,

$$\nu^2 = 1. \tag{6.3}$$

Proof of Theorem A when π is almost selfdual. We have to rule out the decomposition (6.1), which gives (for any finite set S of places containing the ramified and unramified ones):

$$L^S(s, \pi) = L^S(s, \eta)L^S(s, \nu). \tag{6.4}$$

As noted above, we may in fact assume that π is selfdual and that $\nu^2 = 1$. Then the cusp form η will also be selfdual and algebraic. We may then apply Theorem B and conclude the existence of a 4-dimensional, semisimple ℓ-adic representation τ_ℓ associated to η. Then, expanding S to include ℓ, we see that (6.2) implies, in conjunction with the associations $\pi \leftrightarrow \rho_\ell$, $\eta \leftrightarrow \tau_\ell$,

$$L^S(s, \rho_\ell) = L^S(s, \tau_\ell)L^S(s, \nu_\ell). \tag{6.5}$$

By Tchebotarev, this gives the isomorphism

$$\rho_\ell \simeq \tau_\ell \oplus \nu_\ell, \tag{6.6}$$

which contradicts the irreducibility of ρ_ℓ. □

Proof of Theorem A for general regular π. Suppose we have the decomposition (6.1). Again, we may assume that $\nu^2 = 1$.

Let $\omega = \omega_\pi$ denote the central character of π. Then from (6.1) we obtain

$$\omega = \omega_\eta \nu. \tag{6.7}$$

Proposition 6.8. *Assume the decomposition (6.1), and denote by ω the central character of π with corresponding ℓ-adic character ω_ℓ.*

(a) We have the identity

$$L^S(s, \pi; \Lambda^2)L^S(s, \pi^\vee \otimes \omega\nu)\zeta^S(s) = L^S(s, \pi^\vee; \Lambda^2 \otimes \omega)L^S(\pi \otimes \nu)L^S(s, \omega\nu).$$

(b) There is an isomorphism of $\mathcal{G}_\mathbb{Q}$-modules

$$\Lambda^2(\rho_\ell) \oplus (\rho_\ell^\vee \otimes \omega_\ell) \oplus \underline{1} \simeq (\Lambda^2(\rho_\ell^\vee) \otimes \omega_\ell\nu_\ell) \oplus (\rho_\ell \otimes \nu_\ell) \oplus \omega_\ell\nu_\ell.$$

Proof of Proposition 6.8. (a) It is immediate, by checking at each unramified prime, that

$$L^S(s, \pi; \Lambda^2) = L^S(s, \Lambda^2(\eta))L^S(\eta \otimes \nu), \tag{6.9}$$

and (since $\nu = \nu^{-1}$)

$$L^S(s, \pi^\vee; \Lambda^2) = L^S(s, \Lambda^2(\eta^\vee))L^S(\eta^\vee \otimes \nu). \tag{6.10}$$

Since $\omega = \omega_\eta \nu$, we get from Lemma 5.2 that $\Lambda^2(\eta^\vee)$ is isomorphic to $\Lambda^2(\eta) \otimes \omega^{-1}\nu$. Twisting (6.10) by $\omega\nu$, and using the fact that

$$L^S(\eta^\vee \otimes \omega) = L^S(s, \pi^\vee \otimes \omega)/L^S(s, \omega\nu), \tag{6.11}$$

we obtain

$$L^S(s, \pi^\vee; \Lambda^2 \otimes \omega\nu)L^S(s, \omega\nu) = L^S(s, \Lambda^2(\eta))L^S(s, \pi^\vee \otimes \omega). \tag{6.12}$$

Similarly, using (6.9) and the fact that

$$L^S(s, \eta \otimes \nu) = L^S(s, \pi \otimes \nu)/\zeta^S(s),$$

we obtain the identity

$$L^S(s, \pi; \Lambda^2)\zeta^S(s) = L^S(s, \Lambda^2(\eta))L^S(s, \pi \otimes \nu). \tag{6.13}$$

The assertion of part (a) of the proposition now follows by comparing (6.12) and (6.13).

(b) Follows from part (a) by applying Tchebotarev, since $\rho_\ell \leftrightarrow \pi$. □

Proposition 6.14. *We have*

(a) $\omega\nu = 1$.
(b) $\rho_\ell^\vee \simeq \rho_\ell$.

Proof of Proposition 6.14. (a) Since ρ_ℓ is irreducible of dimension 5, it cannot admit a one-dimensional summand, and hence part (*b*) of Proposition 6.8 implies that either $\omega_\ell\nu_\ell = 1$ or

$$\omega_\ell\nu_\ell \subset \Lambda^2(\rho_\ell).$$

Since the first case gives the assertion, let us assume that we are in the second case. But then, again since ρ_ℓ is irreducible, and since $\Lambda^2(\rho_\ell)$ is a summand of $\rho_\ell \otimes \rho_\ell$, we must have

$$\rho_\ell^\vee \simeq \rho_\ell \otimes (\omega_\ell\nu_\ell)^{-1}.$$

In other words, ρ_ℓ is essentially selfdual in this case. Then so is π. More explicitly, we have (since $\nu^2 = 1$)

$$\eta^\vee \boxplus \nu \simeq \pi^\vee \simeq \pi \otimes \omega^{-1}\nu \simeq \eta \otimes \omega^{-1}\nu \boxplus \omega^{-1}.$$

As η is cuspidal, this forces the identity

$$\nu = \omega^{-1}.$$

Done.

(b) Thanks to part (a) (of this proposition), we may rewrite part (b) of Proposition 6.8 as giving the isomorphism of $\mathcal{G}_{\mathbb{Q}}$-modules

$$\Lambda^2(\rho_\ell) \oplus (\rho_\ell^\vee \otimes \omega_\ell) \simeq (\Lambda^2(\rho_\ell^\vee)) \oplus (\rho_\ell \otimes \nu_\ell). \tag{6.15}$$

We now need the following:

Lemma 6.16. *Suppose ρ_ℓ is not selfdual. Then*

$$\rho_\ell \otimes \nu_\ell \not\subset \Lambda^2(\rho_\ell). \tag{6.17}$$

Proof of Lemma 6.16. The hypothesis on ρ_ℓ implies that π is not selfdual, and since $\pi = \eta \boxplus \nu$, η is not selfdual either.

Suppose (6.17) is false. Then, since $\rho_\ell \leftrightarrow \pi$, from (6.1) we must have (since $\nu^2 = 1$)

$$A_p(\pi \otimes \nu) \subset A_p(\pi; \Lambda^2), \quad \forall\, p \notin S, \tag{6.18}$$

for a *finite* set S of primes. But we also have

$$A_p(\pi \otimes \nu) = A_p(\eta \otimes \nu) \oplus \underline{1} \tag{6.19}$$

and

$$A_p(\pi; \Lambda^2) = A_p(\Lambda^2(\eta)) \oplus A_p(\eta \otimes \nu). \tag{6.20}$$

Substituting (6.19) and (6.20) in (6.18), we obtain

$$\underline{1} \subset A_p(\Lambda^2(\eta)) \quad \forall\, p \notin S. \tag{6.21}$$

On the other hand, since η is a non-selfdual cusp form on $\mathrm{GL}(4)/\mathbb{Q}$ of trivial central character, we may apply Lemma 5.6 with $F = \mathbb{Q}$, and conclude that there is an *infinite set of primes* T such that

$$\underline{1} \not\subset A_p(\Lambda^2(\eta)) \quad \forall\, p \in T, \tag{6.22}$$

which contradicts (6.21), proving the lemma. □

In view of the identity (6.15) and Lemma 6.16, we have now proved all of Proposition 6.14. □

We are also done with the proof of Theorem A because π is selfdual when the decomposition (6.1) holds, thanks to the irreducibility of ρ_ℓ, and the selfdual case has already been established (using the algebraic regularity of π). □

References

[AC] J. Arthur and L. Clozel, *Simple Algebras, Base Change and the Advanced Theory of the Trace Formula*, Ann. Math. Studies **120**, Princeton, NJ (1989).

[BHR] D. Blasius, M. Harris and D. Ramakrishnan, *Coherent cohomology, limits of discrete series, and Galois conjugation*, Duke Math. Journal **73** (1994), no. 3, 647–685.

[Cℓ1] L. Clozel, *Motifs et formes automorphes*, in *Automorphic Forms, Shimura varieties, and L-functions*, vol. I, 77–159, Perspectives in Math. **10** (1990).

[Cℓ2] L. Clozel, *Représentations galoisiennes associées aux représentations automorphes autoduales de* $\mathrm{GL}(n)$, Publ. Math. IHES **73**, 97–145 (1991).

[CoPS1] J. Cogdell and I. Piatetski-Shapiro, *Converse theorems for* GL_n *II*, J. Reine Angew. Math. **507**, 165–188 (1999).

[CoPS2] J. Cogdell and I. Piatetski-Shapiro, *Remarks on Rankin-Selberg convolutions*, in Contributions to automorphic forms, geometry, and number theory, 255–278, Johns Hopkins Univ. Press, Baltimore, MD (2004).

[De] P. Deligne, *Formes modulaires et Repésentations ℓ-adiques*, Lecture Notes in Math. **179**, 139–179, Springer-Verlag (1971).

[De-S] P. Deligne and J.-P. Serre, *Formes modulaires de poids* 1, Ann. Sci. École Norm. Sup. (4) **7**, 507–530 (1975).

[FoM] J.-M. Fontaine and B. Mazur, *Geometric Galois representations*, in *Elliptic curves, modular forms, and Fermat's last theorem*, 41–78, Ser. Number Theory, I, Internat. Press, Cambridge, MA (1995).

[GJ] S. Gelbart and H. Jacquet, *A relation between automorphic representations of* $GL(2)$ *and* $GL(3)$, Ann. Scient. Éc. Norm. Sup. (4) **11**, 471–542 (1979).

[GRS] D. Ginzburg, S. Rallis and D. Soudry, *On explicit lifts of cusp forms from* GL_m *to classical groups*, Ann. of Math. (2) **150**, no. 3, 807–866 (1999).

[HaT] M. Harris and R. Taylor, *On the geometry and cohomology of some simple Shimura varieties*, with an appendix by V.G. Berkovich, Annals of Math Studies **151**, Princeton University Press, Princeton, NJ (2001).

[He] G. Henniart, *Une preuve simple des conjectures de Langlands pour* $\mathrm{GL}(n)$ *sur un corps p-adique*, Invent. Math. **139**, no. 2, 439–455 (2000).

[J] H. Jacquet, *Principal L-functions of the linear group*, in Automorphic forms, representations and L-functions, Proc. Sympos. Pure Math. **33**, Part 2, 63–86, Amer. Math. Soc., Providence, R.I. (1979).

[JPSS] H. Jacquet, I. Piatetski-Shapiro and J.A. Shalika, *Rankin-Selberg convolutions*, Amer. J of Math. **105**, 367–464 (1983).

[JS1] H. Jacquet and J.A. Shalika, *Euler products and the classification of automorphic forms* I & II, Amer. J of Math. **103** (1981), 499–558 & 777–815.

[JS2] H. Jacquet and J.A. Shalika, *Exterior square L-functions*, in *Automorphic forms, Shimura varieties, and L-functions*, Vol. II, 143–226, Perspectives in Math. **11** (1990), Academic Press, Boston, MA.

[K] H. Kim, *Functoriality of the exterior square of* GL_4 *and the symmetric fourth of* GL_2, with Appendix 1 by D. Ramakrishnan and Appendix 2 by Kim and P. Sarnak, J. Amer. Math. Soc. **16**, no. 1, 139–183 (2003).

[KSh1] H. Kim and F. Shahidi, *Functorial products for* $GL(2)\times GL(3)$ *and the symmetric cube for* $GL(2)$, With an appendix by Colin J. Bushnell and Guy Henniart, Annals of Math. (2) **155**, no. 3, 837–893 (2002).

[KSh2] H. Kim and F. Shahidi, *Cuspidality of symmetric powers with applications*, Duke Math. J. **112**, no. 1, 177–197 (2002).

[La1] R.P. Langlands, *On the classification of irreducible representations of real algebraic groups*, in *Representation theory and harmonic analysis on semisimple Lie groups*, 101–170, Math. Surveys Monographs **31**, AMS, Providence, RI (1989).

[La2] R.P. Langlands, *Automorphic representations, Shimura varieties, and motives. Ein Märchen*, Proc. symp. Pure Math **33**, ed. by A. Borel and W. Casselman, part 2, 205–246, Amer. Math. Soc., Providence (1979).

[La3] R.P. Langlands, *On the notion of an automorphic representation*, Proc. symp. Pure Math **33**, ed. by A. Borel and W. Casselman, part 1, 189–217, Amer. Math. Soc., Providence (1979).

[Lau1] G. Laumon, *Sur la cohomologie à supports compacts des variétés de Shimura pour* $\mathrm{GSp}(4)/\mathbb{Q}$, Compositio Math. **105** (1997), no. 3, 267–359.

[Lau2] G. Laumon, *Fonctions zétas des variétés de Siegel de dimension trois*, in *Formes automorphes II: le cas du groupe GSp(4)*, Edited by J. Tilouine, H. Carayol, M. Harris, M.-F. Vigneras, Asterisque **302**, Soc. Math. France, *Astérisuqe* (2006).

[MW] C. Moeglin and J.-L. Waldspurger, *Poles des fonctions L de paires pour* $GL(N)$, Appendice, Ann. Sci. École Norm. Sup. (4) **22** (1989), 667–674.

[Pic] *Zeta Functions of Picard Modular Surfaces*, edited by R.P. Langlands and D. Ramakrishnan, CRM Publications, Montréal (1992).

[Ra1] D. Ramakrishnan, *Pure motives and automorphic forms*, in *Motives*, (1994) Proc. Sympos. Pure Math. **55**, Part 2, AMS, Providence, RI, 411–446.

[Ra2] D. Ramakrishnan, *Modularity of the Rankin-Selberg L-series, and Multiplicity one for* $SL(2)$, Annals of Mathematics **152** (2000), 45–111.

[Ra3] D. Ramakrishnan, *Modularity of solvable Artin representations of* $GO(4)$*-type*, IMRN **2002**, No. **1** (2002), 1–54.

[Ra4] D. Ramakrishnan, *Algebraic cycles on Hilbert modular fourfolds and poles of L-functions*, in *Algebraic groups and arithmetic*, 221–274, Tata Inst. Fund. Res., Mumbai (2004).

[Ra5] D. Ramakrishnan, *Irreducibility of* ℓ*-adic associated to regular cusp forms on* $GL(4)/\mathbb{Q}$, preprint (2004), being revised.

[Ra-Sh] D. Ramakrishnan and F. Shahidi, *Siegel modular forms of genus* 2 *attached to elliptic curves*, Math Research Letters **14**, No. 2, 315–332 (2007).

[Ra-W] D. Ramakrishnan and S. Wang, *A cuspidality criterion for the functorial product on* $\mathrm{GL}(2) \times \mathrm{GL}(3)$ *with a cohomological application*, IMRN **2004**, No. **27**, 1355–1394.

[Ri] K. Ribet, *Galois representations attached to eigenforms with Nebentypus*, in *Modular functions of one variable V*, pp. 17–51, Lecture Notes in Math. **601**, Springer, Berlin (1977).

[Se] J.-P. Serre, *Abelian* ℓ*-adic representations*, Research Notes in Mathematics **7**, A.K. Peters Ltd., Wellesley, MA (1998).

[Sh1] F. Shahidi, *On the Ramanujan conjecture and the finiteness of poles for certain L-functions*, Ann. of Math. (2) **127** (1988), 547–584.

[Sh2] F. Shahidi, *A proof of the Langlands conjecture on Plancherel measures; Complementary series for p-adic groups*, Ann. of Math. **132** (1990), 273–330.

[So] D. Soudry, *On Langlands functoriality from classical groups to* GL_n, in *Automorphic forms I*, Astérisque **298**, 335–390 (2005).

[T] J. Tate, *Les conjectures de Stark sur les fonctions L d'Artin en* $s = 0$, Lecture notes edited by D. Bernardi and N. Schappacher, Progress in Mathematics **47** (1984), Birkhäuser, Boston, MA.

[Wei] R. Weissauer, *Four dimensional Galois representations*, in *Formes automorphes II: le cas du groupe GSp(4)*, Edited by J. Tilouine, H. Carayol, M. Harris, M.-F. Vigneras, Asterisque **302**, Soc. Math. France *Astérisque* (2006).

On Liftings of Holomorphic Modular Forms

Tamotsu Ikeda

Graduate School of Mathematics
Kyoto University, Kitashirakawa
Kyoto, 606-8502, Japan
ikeda@kusm.kyoto-u.ac.jp

Summary. In this article, we discuss a lifting of elliptic cusp forms to higher dimensional symmetric spaces. We will consider two cases. The first case is the Siegel modular case, and the second case is the hermitian modular case. The Fourier coefficients of our liftings are closely related to those of Eisenstein series. When the degree is 2, our lifting reduces to the classical Saito–Kurokawa lifting or hermitian Maass lifting.

Key words: Siegel modular forms, hermitian modular forms, Eisenstein series, lifting of cusp forms

Subject Classifications: 11F30, 11F46, 11F55

Introduction

In this article, we discuss a lifting of elliptic cusp forms to higher dimensional symmetric spaces. We will consider two cases. The first case is the Siegel modular case [18], and the second case is the hermitian modular case [19]. The Fourier coefficients of our liftings are closely related to those of Eisenstein series. When the degree is 2, our lifting reduces to the classical Saito–Kurokawa lifting or hermitian Maass lifting.

The finite part of the automorphic representation generated by this lifting is isomorphic to a degenerate principal series. In particular, this is a non-tempered representation.

Part I : Siegel modular case

1 Basic facts

We recall basic facts about Siegel modular forms. The Siegel upper half space $\mathfrak{H}_n$ of degree n is defined by

$$\mathfrak{H}_n = \{Z = {}^tZ \in \mathrm{M}_n(\mathbb{C}) \mid \mathrm{Im}(Z) > 0\}.$$

Here, $\mathrm{Im}(Z) > 0$ means that $\mathrm{Im}(Z)$ is positive definite. We note that $\mathfrak{H}_1 = \{\tau \in \mathbb{C} \mid \mathrm{Im}(\tau) > 0\}$ is equal to the upper half plane. The symplectic group

$$\begin{aligned}\mathrm{Sp}_n(\mathbb{R}) &= \left\{ g \in \mathrm{SL}_{2n}(\mathbb{R}) \,\middle|\, g \begin{pmatrix} \mathbf{0}_n & -\mathbf{1}_n \\ \mathbf{1}_n & \mathbf{0}_n \end{pmatrix} {}^tg = \begin{pmatrix} \mathbf{0}_n & -\mathbf{1}_n \\ \mathbf{1}_n & \mathbf{0}_n \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{SL}_{2n}(\mathbb{R}) \,\middle|\, A\,{}^tB = B\,{}^tA,\ C\,{}^tD = D\,{}^tC,\ A\,{}^tD - B\,{}^tC = \mathbf{1}_n \right\}\end{aligned}$$

acts on $\mathfrak{H}_n$ by $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)(Z) = (AZ + B)(CZ + D)^{-1}$. Put

$$\begin{aligned}\mathcal{S}'_n(\mathbb{Z}) &= \text{the set of half-integral symmetric matrices} \\ &= \left\{ B \in \frac{1}{2}\mathrm{M}_{2n}(\mathbb{Z}) \,\middle|\, B = {}^tB, B_{ii} \in \mathbb{Z}\ (1 \le i \le 2n) \right\}, \\ \mathcal{S}'_n(\mathbb{Z})^+ &= \{B \in \mathcal{S}_{2n}(\mathbb{Z}) \mid B > 0\}.\end{aligned}$$

A holomorphic function F on $\mathfrak{H}_n$ is called a Siegel modular form of weight l if

$$F((AZ + B)(CZ + D)^{-1}) = F(Z)\det(CZ + D)^l$$

for any $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{Sp}_n(\mathbb{Z})$. When $n = 1$, we need to impose that F has a Fourier expansion

$$F(Z) = \sum_{N=0}^{\infty} a_F(N) \exp(2\pi\sqrt{-1}NZ), \qquad Z \in \mathfrak{H}_1.$$

When $n \ge 2$, a Siegel modular form F automatically has a Fourier expansion

$$F(Z) = \sum_{\substack{B \in \mathcal{S}_n(\mathbb{Z}) \\ B \ge 0}} A_F(B)\mathbf{e}(BZ).$$

Here, $\mathbf{e}(X) = \exp(2\pi\sqrt{-1}\mathrm{tr}(X))$. The complex number $A_F(B)$ is called the B-th Fourier coefficient of F. A Siegel modular form F of degree n is called a cusp form if $A_F(B) = 0$ unless $B \in \mathcal{S}_n(\mathbb{Z})^+$. The space of Siegel modular (resp. cusp) forms of degree n and weight l is demoted by $M_l(\mathrm{Sp}_n(\mathbb{Z}))$ (resp. $S_l(\mathrm{Sp}_n(\mathbb{Z}))$).

2 Fourier coefficients of the Eisenstein series

Now we assume $k \equiv n \bmod 2$ and $k > n + 1$. Put

$$\Gamma_\infty^{(2n)} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{Z}), C = 0 \right\}.$$

The Siegel–Eisenstein series on $\mathfrak{H}_{2n}$ of weight $k+n$ is defined by

$$E_{k+n}^{(2n)}(Z) = \sum_{\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \Gamma_\infty^{(2n)} \backslash \mathrm{Sp}_{2n}(\mathbb{Z})} \det(CZ+D)^{-k-n}.$$

As we have assumed $k > n+1$, $E_{k+n}^{(2n)}$ is absolutely convergent. Moreover, $E_{k+n}^{(2n)}$ is a Siegel modular form of weight $k+n$. We define the normalized Eisenstein series by

$$\mathcal{E}_{k+n}^{(2n)}(Z) = 2^{-n}\zeta(1-k-n)\prod_{i=1}^{n}\zeta(1+2i-2k-2n)\cdot E_{k+n}^{(2n)}(Z).$$

For $B \in \mathcal{S}_{2n}(\mathbb{Z})^+$, we put $D_B = \det(2B)$. The absolute value of the discriminant of $\mathbb{Q}(\sqrt{(-1)^n D_B})$ is denoted by $\mathfrak{d}_B$. Put $\mathfrak{f}_B = \sqrt{D_B/\mathfrak{d}_B}$. Let χ_B be the primitive Dirichlet character modulo $\mathfrak{d}_B$ corresponding to $\mathbb{Q}(\sqrt{(-1)^n D_B})/\mathbb{Q}$.

For each prime p, let $\mathbf{e}_p : \mathbb{Q}_p \to \mathbb{C}^\times$ be the additive character of $\mathbb{Q}_p$ such that $\mathbf{e}_p(x) = \mathbf{e}(-x)$ for any $x \in \mathbb{Z}[1/p]$.

Recall that the Siegel series for $B \in \mathcal{S}'_{2n}(\mathbb{Z})^+$ is defined by

$$b_p(B,s) = \sum_{R \in \mathrm{Sym}_{2n}(\mathbb{Q}_p)/\mathrm{Sym}_{2n}(\mathbb{Z}_p)} \mathbf{e}_p(\mathrm{tr}(BR))p^{-\mathrm{ord}_p(\nu(R))s},$$

where

$$\mathrm{Sym}_{2n}(\mathbb{Q}_p) = \{R = {}^tR \mid R \in \mathrm{M}_{2n}(\mathbb{Q}_p)\},$$
$$\mathrm{Sym}_{2n}(\mathbb{Z}_p) = \{R = {}^tR \mid R \in \mathrm{M}_{2n}(\mathbb{Z}_p)\},$$
$$\nu(R) = [R\mathbb{Z}_p^{2n} + \mathbb{Z}_p^{2n} : \mathbb{Z}_p^{2n}].$$

Put

$$\gamma_p(B;X) = (1-X)(1-p^n\chi_B(p)X)^{-1}\prod_{i=1}^{n}(1-p^{2i}X^2).$$

Then there exists a polynomial $F_p(B;X) \in \mathbb{Z}[X]$ such that

$$b_p(B,s) = \gamma_p(B;p^{-s})F_p(B;p^{-s}).$$

Katsurada [20] proved the following functional equation

$$F_p(B;p^{-2n-1}X^{-1}) = (p^{2n+1}X^2)^{-\mathrm{ord}_p\mathfrak{f}_B}F_p(B;X).$$

In particular, we have $\deg F_p(B;X) = 2\mathrm{ord}_p\mathfrak{f}_B$.

It is known that for $B \in \mathcal{S}'_{2n}(\mathbb{Z})^+$, the B-th Fourier coefficient of $\mathcal{E}_{k+n}^{(2n)}(Z)$ is equal to

$$L(1-k,\chi_B)\mathfrak{f}_B^{2k-1}\prod_{p \mid D_B}F_p(B;p^{-k-n}).$$

Put $\tilde{F}_p(B;X) = X^{-\mathrm{ord}_p \mathfrak{f}_B} F_p(B; p^{-n-(1/2)}X)$. Then Katsurada's functional equation implies

$$\tilde{F}_p(B;X^{-1}) = \tilde{F}_p(B;X).$$

In terms of $\tilde{F}_p(B;X)$, the B-th Fourier coefficient of $\mathcal{E}^{(2n)}_{k+n}(Z)$ can be expressed as

$$L(1-k,\chi_B)\mathfrak{f}_B^{k-(1/2)} \prod_{p|\mathfrak{f}_B} \tilde{F}_p(B;p^{-k+(1/2)}).$$

3 Kohnen plus space

Let $\mathfrak{G}$ the group which consists of all pairs $(\gamma,\phi(\tau))$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ and $\phi(\tau)$ is a holomorphic function on $\mathfrak{H}_1$ satisfying $|\phi(\tau)| = |c\tau+d|$, with group law defined by $(\gamma_1,\phi_1(\tau))\cdot(\gamma_2,\phi_2(\tau)) = (\gamma_1\gamma_2, \phi_1(\gamma_2(\tau))\phi_2(\tau))$. If $h(\tau)$ is a function on $\mathfrak{H}_1$ and $\xi = (\gamma,\phi(\tau)) \in \mathfrak{G}$, we put

$$(h|\xi)(\tau) = (h|_{k+(1/2)}\xi)(\tau) = \phi(\tau)^{-2k-1}h(\gamma(\tau)).$$

Then $(h|\xi_1)|\xi_2 = h|(\xi_1\xi_2)$, for $\xi_1,\xi_2 \in \mathfrak{G}$. On the other hand, for $\gamma \in \mathrm{SL}_2(\mathbb{R})$, we put

$$(h||_{k+(1/2)}\gamma)(\tau) = (c\tau+d)^{-(2k+1)/2}h(\gamma(\tau)).$$

Then for $\gamma_1,\gamma_2 \in \mathrm{SL}_2(\mathbb{R})$, we have $(h||\gamma_1)||\gamma_2 = t\cdot h||(\gamma_1\gamma_2)$, for some $t \in \mathbb{C}$, $|t| = 1$.

There exists an injective homomorphism $\Gamma_0(4) \to \mathfrak{G}$ given by $\gamma \mapsto \gamma^* = (\gamma, j(\gamma,\tau))$, where

$$j(\gamma,\tau) = \left(\frac{c}{d}\right)\epsilon_d^{-1}(c\tau+d)^{1/2}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4).$$

Here,

$$\epsilon_d = \begin{cases} 1, & \text{if } d \equiv 1 \bmod 4, \\ \sqrt{-1}, & \text{if } d \equiv 3 \bmod 4. \end{cases}$$

Recall that $M_{k+(1/2)}(\Gamma_0(4))$ (resp. $S_{k+(1/2)}(\Gamma_0(4))$) consists of all holomorphic functions $h(\tau)$ on $\mathfrak{H}_1$ which satisfy $h|_{k+(1/2)}\gamma^* = h$ for every $\gamma \in \Gamma_0(4)$ and which are holomorphic (resp. which vanish) at all cusps. The Kohnen plus space $M^+_{k+(1/2)}(\Gamma_0(4))$ consists of all $h(\tau) \in M_{k+(1/2)}(\Gamma_0(4))$ whose Fourier expansion is of the form

$$h(\tau) = \sum_{\substack{N\geq 0 \\ (-1)^k N \equiv 0,1(4)}} c(N)q^N, \quad q = \exp(2\pi\sqrt{-1}\tau).$$

Similarly, the Kohnen plus space $S^+_{k+(1/2)}(\Gamma_0(4))$ is defined by

$$S^+_{k+(1/2)}(\Gamma_0(4)) = S_{k+(1/2)}(\Gamma_0(4)) \cap M^+_{k+(1/2)}(\Gamma_0(4)).$$

Kohnen [24] proved that $M^+_{k+(1/2)}(\Gamma_0(4))$ has a basis that consists of Hecke eigenforms, and that the Shimura correspondence is one-to-one, i.e., there is a one-to-one correspondence between Hecke eigenforms $f(\tau) \in M_{2k}(\mathrm{SL}_2(\mathbb{Z}))$ and Hecke eigenforms in $h(\tau) \in M^+_{k+(1/2)}(\Gamma_0(4))$, up to scalar multiplication. The form $f(\tau) \in M_{2k}(\mathrm{SL}_2(\mathbb{Z}))$ is a cusp form if and only if the corresponding $h(\tau) \in M^+_{k+(1/2)}(\Gamma_0(4))$ is a cusp form.

Let

$$f(\tau) = \sum_{N=0}^{\infty} a(N) q^N \in M_{2k}(\mathrm{SL}_2(\mathbb{Z})), \quad q = e^{2\pi\sqrt{-1}\tau}$$

be a normalized Hecke eigenform of weight $2k$. Let $\alpha_p^{\pm 1}$ be the Satake parameter of f at p, i.e.,

$$(1 - p^{k-(1/2)}\alpha_p X)(1 - p^{k-(1/2)}\alpha_p^{-1} X) = 1 - a(p)X + p^{2k-1}X^2.$$

Let $h(\tau) = \sum_{N \geq 0} c(N) q^N \in M^+_{k+(1/2)}(\Gamma_0(4))$ be a non-zero Hecke eigenform. Then $h(\tau)$ corresponds to $f(\tau)$ by the Shimura correspondence if and only if for any fundamental discriminant D such that $(-1)^k D > 0$, we have

$$c(|D| f^2) = c(|D|) \sum_{d|f} \mu(d) \chi_{|D|}(d) d^{k-1} a(f/d).$$

Here $\mu(d)$ is the Möbius function.

For any positive integer N such that $(-1)^k N \equiv 0, 1 \bmod 4$, we denote the absolute value of the discriminant of $\mathbb{Q}(\sqrt{(-1)^k N})/\mathbb{Q}$ by $\mathfrak{d}_N$ and the positive rational number such that $N = \mathfrak{d}_N \mathfrak{f}_N^2$ by $\mathfrak{f}_N$. Note that $\mathfrak{f}_N$ is an integer. Let χ_N be the primitive Dirichlet character corresponding to $\mathbb{Q}(\sqrt{(-1)^k N})/\mathbb{Q}$.

We define $\Psi_p(N; X_p) \in \mathbb{C}[X, X^{-1}]$ by

$$\Psi_p(N; X) = \frac{X^{e+1} - X^{-e-1}}{X - X^{-1}} - p^{-1/2} \chi_N(p) \frac{X^e - X^{-e}}{X - X^{-1}}.$$

Here $e = \mathrm{ord}_p \mathfrak{f}_N$. Note that $\Psi_p(N; X) = 1$ if $\mathrm{ord}_p \mathfrak{f}_N = 0$. In terms of $\Psi_p(N; X)$, we have

$$c(N) = c(\mathfrak{d}_N) \mathfrak{f}_N^{k-(1/2)} \prod_p \Psi_p(N; \alpha_p).$$

4 Lifting of cusp forms

Now we consider cusp forms. Let k be an arbitrary positive integer such that $k \equiv n \bmod 2$.

Choose a normalized Hecke eigenform

$$f(\tau) = \sum_{N=1}^{\infty} a(N) q^N \in S_{2k}(\mathrm{SL}_2(\mathbb{Z})), \qquad a(1) = 1$$

and a corresponding Hecke eigenform

$$h(\tau) = \sum_{\substack{N>0 \\ (-1)^k N \equiv 0,\, 1\ (4)}} c(N) q^N \in S^+_{k+(1/2)}(\Gamma_0(4)).$$

Let

$$\begin{aligned} L(s, f) &= \sum_{N=1}^{\infty} a(N) N^{-s} \\ &= \prod_p [(1 - p^{k-(1/2)} \alpha_p X)(1 - p^{k-(1/2)} \alpha_p^{-1} X)]^{-1} \end{aligned}$$

be the L-function of f. The set $\{\alpha_p, \alpha_p^{-1}\}$ is called the Satake parameter of f.

Put

$$A(B) = c(\mathfrak{d}_B) \mathfrak{f}_B^{k-(1/2)} \prod_p \tilde{F}_p(B; \alpha_p), \quad B \in \mathcal{S}_{2n}(\mathbb{Z})^+$$

$$F(Z) = \sum_{\substack{B \in \mathcal{S}'_{2n}(\mathbb{Z})^+ \\ B = {}^t B > 0}} A(B) \mathbf{e}(BZ), \quad Z \in \mathfrak{H}_{2n}.$$

Note that $\tilde{F}_p(B; \alpha_p)$ does not depend on the choice of α_p by Katsurada's functional equation. Then our first main theorem as as follows.

Theorem 1. *Assume $k \equiv n \bmod 2$. Then $F \in S_{k+n}(\mathrm{Sp}_{2n}(\mathbb{Z}))$ and $F \not\equiv 0$. Moreover, F is a Hecke eigenform whose standard L-function is equal to*

$$L(s, F, \mathrm{st}) = \zeta(s) \prod_{i=1}^{2n} L(s + k + n - i, f).$$

5 Outline of the proof

We consider the Fourier–Jacobi expansion

$$\begin{aligned} F\left(\begin{pmatrix} \omega & z \\ {}^t z & \tau \end{pmatrix}\right) &= \sum_{S \in \mathcal{S}_{2n-1}(\mathbb{Z})^+} \sum_{\lambda \in (2S)^{-1}\mathbb{Z}^{2n-1}/\mathbb{Z}^{2n-1}} \theta_{[\lambda]}(S; \tau, z) \mathbf{e}(\mathrm{tr}(S\omega)) \\ &\quad \times \sum_{N \in \mathbb{Z}\, N - {}^t\lambda S \lambda \geq 0} A_F\left(\begin{pmatrix} S & S\lambda \\ {}^t\lambda S & N \end{pmatrix}\right) \mathbf{e}((N - {}^t\lambda S \lambda)\tau). \end{aligned}$$

Here $\theta_{[\lambda]}(S; \tau, z) = \sum_{x \in \mathbb{Z}^{2n-1}} \mathbf{e}({}^t(x+\lambda) S (x+\lambda)\tau + 2\,{}^t(x+\lambda) S z)$. For each $S \in \mathcal{S}_{2n-1}(\mathbb{Z})^+$, $\Delta = 2 \det S$, one can show that

$$\sum_{\substack{N\in\mathbb{Z}\\ N-{}^t\lambda S\lambda\geq 0}} A_{\mathcal{E}^{(2n)}_{k'+n}}\left(\begin{pmatrix} S & S\lambda \\ {}^t\lambda S & N\end{pmatrix}\right)\mathbf{e}((N-{}^t\lambda S\lambda)\tau)$$

$$= \text{(degenerate terms)}$$

$$+\sum_{N=1}^{\infty} H(k',\mathfrak{d}_N)\mathfrak{f}_N^{k'-(1/2)}\left(\prod_{p|N}\tilde{F}_p\left(\begin{pmatrix} S & S\lambda \\ {}^t\lambda S & {}^t\lambda S\lambda+N/\Delta\end{pmatrix}\right); p^{-k'+(1/2)}\right)q^{N/\Delta}$$

is in the space generated by some translates of $\mathcal{H}_{k+(1/2)}(\tau)$. Note that k' can be arbitrarily large. From this, one can show that

$$\sum_{N=1}^{\infty} c(\mathfrak{d}_N)\mathfrak{f}_N^{k-(1/2)}\left(\prod_{p|N}\tilde{F}_p\left(\begin{pmatrix} S & S\lambda \\ {}^t\lambda S & {}^t\lambda S\lambda+N/\Delta\end{pmatrix}\right); \alpha_p\right)q^{N/\Delta}$$

is in the space generated by some translates of $h(\tau)$, and has the same K-types. It follows that $F(Z)$ is modular with respect to both the Siegel parabolic and with respect to Jacobi parabolic subgroup. Since these two parabolic subgroups generates $\mathrm{Sp}_{2n}(\mathbb{Z})$, we have the desired modularity of $F(Z)$.

6 Relation to the Saito–Kurokawa lifts

We shall show that when $n = 1$, $F(Z)$ is equal to the Saito–Kurokawa lift of $f(\tau)$. Let k be an odd integer.

Recall that a Siegel modular form $F(Z) = \sum_{B\in\mathcal{S}_2'(\mathbb{Z})} A_F(B)\mathbf{e}(BZ)$ of weight $k+1$ satisfies a Maass relation if there is a function $\beta_F : \mathbb{Z}_{\geq 0} \to \mathbb{C}$ such that

$$A_F(B) = \sum_{\substack{d>0\\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^k\cdot\beta_F\left(\frac{D_B}{d^2}\right).$$

The space of Siegel modular forms of weight $k+1$ which satisfies the Maass relation is called the Maass spezialschar. The Maass spezialschar is canonically isomorphic to the Kornen plus space $M^+_{k+(1/2)}(\Gamma_0(4))$ by

$$\Omega^{SK} : F(Z) = \sum_{B\in\mathcal{S}_2'(\mathbb{Z})} A_F(B)\mathbf{e}(BZ) \mapsto \sum_{\substack{n\geq 0\\ n\equiv 0,3(4)}} \beta_F(n)\mathbf{e}(n\tau).$$

Put $h(\tau) = \Omega^{SK}(F) \in M^+_{k+(1/2)}(\Gamma_0(4))$. Then $F(Z)$ is a Hecke eigenform if and only if $h(\tau)$ is a Hecke eigenform, and $F(Z)$ is called the Saito–Kurokawa lift of $h(\tau)$. If

$$h(\tau) = \sum_{\substack{n\geq 0\\ n\equiv 0,3(4)}} c(n)\mathbf{e}(n\tau),$$

then B-th Fourier coefficient of $F(Z)$ is equal to

$$\sum_{\substack{d>0 \\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^k c\left(\frac{D_B}{d^2}\right).$$

Let k' be a sufficiently large odd integer. It is well known that $\mathcal{E}^{(2)}_{k'+1}(Z)$ satisfies the Maass relation. Put $H(k',n)=\beta_{\mathcal{E}^{(2)}_{k'+1}}(n)$. The function

$$\Omega^{SK}(\mathcal{E}^{(2)}_{k'+1})(\tau)=\mathcal{H}_{k'+(1/2)}(\tau)=\sum_{\substack{n\geq 0 \\ n\equiv 0,3(4)}} H(k',n)\mathbf{e}(n\tau)$$

is called the Cohen–Eisenstein series (cf. Cohen [8]. [9]).

Since $\mathcal{E}^{(2)}_{k'+1}(Z)$ satisfies the Maass relation, the B-the Fourier coefficent of $\mathcal{E}^{(2)}_{k'+1}(Z)$ is equal to

$$\begin{aligned}
&\sum_{\substack{d>0 \\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^{k'} H\left(k',\frac{D_B}{d^2}\right) \\
&= H(k',\mathfrak{d}_B)\mathfrak{f}_B^{k'-(1/2)} \sum_{\substack{d>0 \\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^{1/2}\prod_p \Psi_p\left(\frac{D_B}{d^2};\, p^{k'-(1/2)}\right).
\end{aligned}$$

Since k' is arbitrary, we have

$$\prod_p \tilde{F}_p(B;X_p)=\sum_{\substack{d>0 \\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^{1/2}\prod_p \Psi_p\left(\frac{D_B}{d^2};X_p\right).$$

It follows that

$$\begin{aligned}
A(B) &= c(\mathfrak{d}_B)\mathfrak{f}_B^{k-(1/2)}\prod_p \tilde{F}_p(B;\alpha_p) \\
&= c(\mathfrak{d}_B)\mathfrak{f}_B^{k-(1/2)}\sum_{d|(m,r,l)} d^{1/2}\prod_p \Psi_p\left(\frac{D_B}{d^2};\alpha_p\right) \\
&= \sum_{\substack{d>0 \\ d^{-1}B\in\mathcal{S}_2'(\mathbb{Z})}} d^k c\left(\frac{D_B}{d^2}\right).
\end{aligned}$$

This agrees with the well-known Fourier coefficient formula for the Saito–Kurokawa lift.

Part II : Hermitian modular case

7 Hermitian modular forms and hermitian Eisensetein series

Now we consider the hermitian modular case. Let $K = \mathbb{Q}(\sqrt{-D_K})$ be an imaginary quadratic field. We denote the ring of integers of K by $\mathcal{O}$. The non-trivial automorphism of K is denoted by $x \mapsto \bar{x}$. The primitive Dirichlet character corresponding to $K/\mathbb{Q}$ is denoted by χ. We denote by $\mathcal{O}^\sharp = (\sqrt{-D})^{-1}\mathcal{O}$ the inverse different ideal of $K/\mathbb{Q}$. For each prime p, we set $K_p = K \otimes \mathbb{Q}_p$ and $\mathcal{O}_p = \mathcal{O} \otimes \mathbb{Z}_p$.

The special unitary group $\mathrm{SU}(m,m)$ is an algebraic group defined over $\mathbb{Q}$, whose group of R-valued points is given by

$$\left\{ g \in \mathrm{GL}_{2m}(R \otimes K) \,\middle|\, g \begin{pmatrix} \mathbf{0}_m & -\mathbf{1}_m \\ \mathbf{1}_m & \mathbf{0}_m \end{pmatrix} {}^t\bar{g} = \begin{pmatrix} \mathbf{0}_m & -\mathbf{1}_m \\ \mathbf{1}_m & \mathbf{0}_m \end{pmatrix},\ \det g = 1 \right\}$$

for any $\mathbb{Q}$-algebra R.

The special hermitian modular group $\Gamma_K^{(m)}$ is defined by $\mathrm{SU}(m,m)(\mathbb{Q}) \cap \mathrm{SL}_{2m}(\mathcal{O})$. Note that $\Gamma_K^{(1)} = \mathrm{SL}_2(\mathbb{Z})$.

Put

$$\Gamma_{K,\infty}^{(m)} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_K^{(m)} \,\middle|\, C = 0 \right\}.$$

We define the hermitian upper half space $\mathcal{H}_m$ by

$$\mathcal{H}_m = \{ Z \in M_m(\mathbb{C}) \mid \frac{1}{2\sqrt{-1}}(Z - {}^t\bar{Z}) > 0 \}.$$

The action of $\mathrm{SU}(2,2)(\mathbb{R})$ on $\mathcal{H}_m$ is given by

$$g\langle Z \rangle = (AZ + B)(CZ + D)^{-1}, \quad Z \in \mathcal{H}_m,\ g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

We put

$$\Lambda_m(\mathcal{O}) = \{ h = (h_{ij}) \in \mathrm{M}_m(K) \mid h_{ii} \in \mathbb{Z}, h_{ij} = \bar{h}_{ji} \in \mathcal{O}^\sharp,\ (i \neq j) \},$$
$$\Lambda_m(\mathcal{O})^+ = \{ h \in \Lambda_m(\mathcal{O}) \mid h > 0 \}.$$

For $H \in \Lambda_m(\mathcal{O})$, $\det H \neq 0$, we put

$$\gamma(H) = (-D_K)^{[m/2]} \det(H).$$

A holomorphic function F on $\mathcal{H}_m$ is called a hermitian modular form of weight l if

$$F((AZ + B)(CZ + D)^{-1}) = F(Z) \det(CZ + D)^l$$

for any $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \Gamma_K^{(m)}$. Again, we need a condition on Fourier expantion when $n = 1$. When $m \geq 2$, a hermitian modular form F automatically has a Fourier expansion

$$F(Z) = \sum_{\substack{H \in \Lambda_n(\mathcal{O}) \\ H \geq 0}} A_F(H)\mathbf{e}(HZ).$$

The complex number $A_F(H)$ is called the H-th Fourier coefficient of F. A hermitian modular form F of degree m is called a cusp form if $A_F(H) = 0$ unless $H \in \Lambda_m(\mathcal{O})^+$. The space of hermitian modular (resp. cusp) form of degree m and weight l is demoted by $S_l(\Gamma_K^{(m)})$ (resp.$S_l(\Gamma_K^{(m)})$).

The Siegel series for $H \in \Lambda_m(\mathcal{O})^+$ is defined by

$$b_p(H, s) = \sum_{R \in \mathcal{H}_m(K_p)/\mathcal{H}_m(\mathcal{O}\otimes\mathbb{Z}_p)} \mathbf{e}_p(\mathrm{tr}(HR))p^{-\mathrm{ord}_p(\nu(R))s}$$

for $\mathrm{Re}(s) \gg 0$. Here, $\mathcal{H}_m(K_p)$ (resp. $\mathcal{H}_m(\mathcal{O} \otimes \mathbb{Z}_p)$) is the additive group of all hermitian matrices with entries in K_p (resp. $\mathcal{O} \otimes \mathbb{Z}_p$).

The ideal $\nu(R) \subset \mathbb{Z}_p$ is defined as follows: Choose an element $g = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{SU}(2,2)(\mathbb{Q}_p) \cap \mathrm{SL}_{2m}(\mathcal{O} \otimes \mathbb{Z}_p)$ such that $\det D \neq 0$, $D^{-1}C = R$. Then $\nu(R) = \det(D) \in \mathbb{Z}_p$.

We define a polynomial $t_p(K/\mathbb{Q}; X) \in \mathbb{Z}[X]$ by

$$t_p(K/\mathbb{Q}; X) = \prod_{i=1}^{[(m+1)/2]} (1 - p^{2i}X) \prod_{i=1}^{[m/2]} (1 - p^{2i-1}\chi(p)X).$$

Then there exists a polynomial $F_p(H; X) \in \mathbb{Z}[X]$ such that

$$b_p(H, s) = t_p(K/\mathbb{Q}; p^{-s})F_p(H; p^{-s}).$$

$$\deg F_p(H; X) = \mathrm{ord}_p \gamma(H).$$

The functional equation of $F_p(H; X)$ is as follows:

$$F_p(H; p^{-2m}X^{-1}) = \underline{\chi}_p(\gamma(H))^{m-1}(p^m X)^{-\mathrm{ord}_p\gamma(H)} F_p(H; X).$$

Here, $\underline{\chi}_p$ is the p-component of the idele character $\mathbb{A}_{\mathbb{Q}}^\times/\mathbb{Q}^\times \to \mathbb{C}^\times$ corresponding to χ.

Put

$$\tilde{F}_p(H; X) = X^{\mathrm{ord}_p\gamma(H)} F_p(H; p^{-m}X^{-2}).$$

Then

$$\tilde{F}_p(H; X^{-1}) = \tilde{F}_p(H; X), \quad 2 \nmid m$$

$$\tilde{F}_p(H; \chi(p)X^{-1}) = \tilde{F}_p(H; X), \quad 2|m, \text{ and } \chi(p) \neq 0.$$

Assume $k \gg 0$. Put $n = [m/2]$. We define the Eisenstein series

$$E_{2k+2n}^{(m)}(Z) = \sum_{\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \Gamma_{K.\infty}^{(m)} \backslash \Gamma_K^{(m)}} \det(CZ + D)^{-2k-2n}$$

and its normalization

$$\mathcal{E}_{2k+2n}^{(m)}(Z) = 2^{-m} \prod_{i=1}^{m} L(1 + i - 2k - 2n, \chi^{i-1}) \cdot E_{2k+2n}^{(m)}(Z).$$

Then for each $H \in \Lambda_m(\mathcal{O})^+$, the H-th Fourier coefficient of $\mathcal{E}_{2k+2n}^{(m)}$ is equal to

$$|\gamma(H)|^{k+n-(m/2)} \prod_{p|\gamma(H)} \tilde{F}_p(H; p^{-k-n+(m/2)}).$$

8 The case $m = 2n + 1$

When $m = 2n + 1$, the H-th Fourier coefficient of $\mathcal{E}_{2k+2n}^{(2n+1)}(Z)$ is equal to

$$|\gamma(H)|^{k-(1/2)} \prod_{p|\gamma(H)} \tilde{F}_p(H; p^{-k+(1/2)})$$

for any $H \in \Lambda_{2n+1}(\mathcal{O})^+$.

Now let $f(\tau) = \sum_{N=1}^{\infty} a(N) q^N \in S_{2k}(\mathrm{SL}_2(\mathbb{Z}))$ be a normalized Hecke eigenform, whose L-function is given by

$$L(f, s) = \sum_{N=1}^{\infty} a(N) N^{-s} = \prod_p [(1 - p^{k-(1/2)} \alpha_p X)(1 - p^{k-(1/2)} \alpha_p^{-1} X)]^{-1}.$$

Put

$$A(H) = |\gamma(H)|^{k-(1/2)} \prod_{p|\gamma(H)} \tilde{F}_p(H, \alpha_p)$$

for $H \in \Lambda_{2n+1}(\mathcal{O})^+$ and

$$F(Z) = \sum_{H \in \Lambda_{2n+1}(\mathcal{O})^+} A(H) \mathbf{e}(HZ)$$

for $Z \in \mathcal{H}_{2n+1}$.

Then we have

Theorem 2. *Assume that $m = 2n + 1$ is odd. Then $F \in S_{2k+2n}(\Gamma_K^{(2n+1)})$ and $F \not\equiv 0$. Moreover, F is a Hecke eigenform.*

9 The case $m = 2n$

When $m = 2n$, the H-th Fourier coefficient of $\mathcal{E}^{(2n)}_{2k+2n}(Z)$ is equal to

$$|\gamma(H)|^k \prod_{p|\gamma(H)} \tilde{F}_p(H; p^{-k})$$

for any $H \in \Lambda_{2n}(\mathcal{O})^+$.

Now let $f(\tau) = \sum_{N=1}^{\infty} a(N)q^N \in S_{2k+1}(\Gamma_0(D_K), \chi)$ be a primitive form, whose L-function is given by

$$\begin{aligned} L(f, s) &= \sum_{N=1}^{\infty} a(N)N^{-s} \\ &= \prod_{p\nmid D_K} (1 - a(p)p^{-s} + \chi(p)p^{2k-2s})^{-1} \times \prod_{p|D_K} (1 - a(p)p^{-s})^{-1}. \end{aligned}$$

For each prime $p \nmid D_K$, we define the Satake parameter $\{\alpha_p, \beta_p\} = \{\alpha_p, \chi(p)\alpha_p^{-1}\}$ by

$$(1 - a(p)X + \chi(p)p^{2k}X^2) = (1 - p^k\alpha_p X)(1 - p^k\beta_p X).$$

For $p \mid D_K$, we put $\alpha_p = p^{-k}a(p)$.

Put

$$A(H) = |\gamma(H)|^k \prod_{p|\gamma(H)} \tilde{F}_p(H, \alpha_p)$$

for $H \in \Lambda_{2n}(\mathcal{O})^+$ and

$$F(Z) = \sum_{H\in\Lambda_{2n}(\mathcal{O})^+} A(H)\mathbf{e}(HZ)$$

for $Z \in \mathcal{H}_{2n}$. Then we have

Theorem 3. *Assume that $m = 2n$ is even. Then $F \in S_{2k+2n}(\Gamma_K^{(2n)})$. Moreover, F is a Hecke eigenform. $F \equiv 0$ if and only if n is odd and $f(\tau)$ comes from a Hecke character of some imaginary quadratic field.*

10 L-functions

For simplicity, we assume the class number of K is one. Then, the L-function of F is as follows.

$$\begin{aligned} L(s, F, \rho) = &\prod_{i=1}^{2n+1} L(s + k + n - i + (1/2), f) \\ &\times \prod_{i=1}^{2n+1} L(s + k + n - i + (1/2), f, \chi) \end{aligned}$$

for $m = 2n + 1$, and

$$L(s, F, \rho) = \prod_{i=1}^{2n} L(s + k + n - i + (1/2), f) \times \prod_{i=1}^{2n} L(s + k + n - i + (1/2), f, \chi)$$

for $m = 2n$ and $F \not\equiv 0$.

Here, ρ is a $2m$-dimensional representation of the L-group of $U(m, m)$.

11 The case $m = 2$

The case $m = 2$ was first considered by Kojima [27] for $K = \mathbb{Q}(\sqrt{-1})$ and later by Krieg [28] and Sugano [43] for arbitrary imaginary quadratic field. For simplicity, we assume $D_K \neq 3, 4$.

Recall that

$$G(Z) = \sum_{H \in \Lambda_2(\mathcal{O})} A_G(H)\mathbf{e}(HZ) \in M_{2k+2}(\Gamma_2^{(2)})$$

satisfies the Maass relation if and only if there is a function

$$\alpha_G^* : \mathbb{Z}_{\geq 0} \to \mathbb{C}$$

such that

$$A_G(H) = \sum_{d | \varepsilon(H)} d^{2k+1} \alpha_G^* \left(\frac{|\gamma(H)|}{d^2} \right).$$

Here

$$\varepsilon(H) = \max\{q \in \mathbb{Z}_{>0} \mid q^{-1}H \in \Lambda_2(\mathcal{O})\}.$$

We denote the space of elements of $M_{2k+2}(\Gamma_2^{(2)})$ satisfying the Maass relation by $M_{2k+2}^{\mathrm{Maass}}(\Gamma_2^{(2)})$. We put

$$\mathbf{a}_{D_K}(N) = \prod_{p | D_K} (1 + \chi_p(-N)).$$

Here, $\chi = \prod_{p|D_K} \chi_p$ is the decomposition to a product of Dirichlet characters with prime power conductors. The linear map

$$\Omega : M_{2k+2}^{\mathrm{Maass}}(\Gamma_K^{(2)}) \to M_{2k+1}(\Gamma_0(D_K), \chi)$$

is defined by

$$\Omega(G)(\tau) = \sum_{N=0}^{\infty} \mathbf{a}_{D_K}(N)\alpha_G^*(H)q^N.$$

(This definition is slightly modified by a scalar from Krieg's definition.) Then Krieg proved that the image of Ω is equal to the space

$$M^*_{2k+1}(\Gamma_0(D_K),\chi) = \left\{ f=\sum_{N\ge 0} a(N)q^N \in M_{2k+1}(\Gamma_0(D_K),\chi) \,\middle|\, a(N)=0 \text{ for } \mathbf{a}_{D_K}(N)=0 \right\}.$$

Moreover, Ω induces an isomorphism between $M^{\text{Maass}}_{2k+2}(\Gamma_2^{(2)})$ and $M^*_{2k+1}(\Gamma_0(D_K),\chi)$. For each primitive form $f \in S_{2k+1}(\Gamma_0(D_K),\chi)$, there exists an element $f^* \in S^*_{2k+1}(\Gamma_0(D_K),\chi)$ such that

$$a_{f^*}(n) = \mathbf{a}_{D_K}(n) a_f(n), \quad (n, D_K)=1.$$

When G is equal to the normalized hermitian Eisenstein series

$$\mathcal{E}^{(2)}_{2k+2}(Z) = \frac{B_{2k+2}B_{2k+1,\chi}}{8(k+1)(2k+1)} E^{(2)}_{2k+2}(Z),$$

then we have (cf. Krieg [28], p. 679)

$$\alpha^*_G(N) = \begin{cases} 0 & \mathbf{a}_{D_K}(N)=0 \\ -\frac{B_{2k+1,\chi}}{4k+2} & N=0 \\ \frac{1}{\mathbf{a}_{D_K}(N)} \sum_{d\mid N} \sum_{Q\subset Q_{D_K}} \chi_Q\left(\frac{-N}{d}\right) \chi'_Q(d) d^{2k} & N>0, \mathbf{a}_{D_K}(N)\ne 0. \end{cases}$$

Using these results, one can calculate the polynomial $F_p(H;X)$ as follows: If $p \nmid D_K$, then

$$F_p(H;X) = \sum_{i=0}^{b} p^{3i} X^i \sum_{j=0}^{a-2i} \underline{\chi}_p(p)^j p^{2j} X^j.$$

If $p \mid D_K$, then

$$F_p(H;X) = \begin{cases} \sum_{i=0}^{b} p^{3i} X^i (1+\underline{\chi}_p(\gamma(H)) p^{2(a-2i)} X^{a-2i}) & 2b<a \\ p^{3b}X^b + \sum_{i=0}^{b-1} (p^{3i}X^i + p^{4b-i}X^{2b-i}) & 2b=a. \end{cases}$$

Here $a = \operatorname{ord}_p \gamma(H)$, $b = \operatorname{ord}_p \varepsilon(H)$. When the class number of K is one, this has been already calculated by Nagaoka [35]. Using this result, one can show that $\Omega^{-1}(f^*)$ is the lift of the primitive form $f \in S_{2k+1}(\Gamma_0(D_K),\chi)$.

References

1. J. Arthur, *Unipotent automorphic representations: conjectures*, Astérisque **171–172** (1989), 13–71.

2. S. Böcherer, *Über die Fourier-Jacobi-Entwicklung Siegelscher Eisensteinreihen*, I Math. Zeitschr. **183** (1983), 21–46.
3. S. Böcherer, *Über die Fourierkoeffizienten der Siegelschen Eisensteinreihen*, Manuscripta Math. **45** (1984), 273–288.
4. S. Böcherer, *Siegel modular forms and theta series*, Proc. Symp. Pure Math., **49–2** (1989), 3–17.
5. R. E. Borcherds, E. Freitag, and R. Weissauer, *A Siegel cusp form of degree 12 and weight 12*, J. Reine Angew. Math. **494** (1998), 141–153.
6. H. Braun, *Hermitian modular functions. III*, Ann. of Math. **53**, (1951), 143–160.
7. S. Breulmann and M. Kuss, *On a conjecture of Duke-Imamoglu*, Proc. Amer. Math. Soc. **128** (2000), 1595–1604.
8. H. Cohen, *Sommes de carrés, fonctions L et formes modulaires*, C. R. Acad. Sci. Paris Sér. A-B 277 (1973), A827–A830.
9. H. Cohen, *Sums involving the values at negative integers of L-functions of quadratic characters*, Math. Ann. **217** (1975), 271–285.
10. M. Eichler and D. Zagier, *The theory of Jacobi forms*, Progress in Mathematics **55** Birkhäuser Boston, Inc., Boston, Mass. (1985).
11. P. Feit, *Poles and residues of Eisenstein series for symplectic and unitary groups*, Memoirs of the AMS. **346** (1986).
12. P. Feit, *Explicit formulas for local factors in the Euler products for Eisenstein series*, Nagoya Math. J. **113** (1989), 37–87.
13. E. Freitag, *Siegelsche Modulfunktionen*, Springer-Verlag, Berlin-New York (1983).
14. V. A. Gritsenko, *The Maass space for* $\mathrm{SU}(2,2)$. *The Hecke ring, and zeta functions* (Russian), Translated in Proc. Steklov Inst. Math. 1991, no. 4, 75–86. Galois theory, rings, algebraic groups and their applications (Russian). Trudy Mat. Inst. Steklov. **183** (1990), 68–78, 223–225.
15. T. Ibukiyama, *Conjecture on the lifting of modular forms to Siegel modular forms*, informal note.
16. A. Ichino and T. Ikeda, *On Maass lifts and the central critical values of triple product L-functions*, preprint.
17. T. Ikeda, *On the theory of Jacobi forms and the Fourier-Jacobi coefficients of Eisenstein series*, J. Math. Kyoto Univ. **34** (1994), 615–636.
18. T. Ikeda, *On the lifting of elliptic cusp forms to Siegel cusp forms of degree* $2n$, Ann. of Math. **154** (2001), 641–681.
19. T. Ikeda, *On the lifting of hermitian modular forms*, preprint.
20. H. Katsurada, *An explicit formula for Siegel series*, Amer. J. Math. **121** (1999), 415–452.
21. Y. Kitaoka, *Dirichlet series in the theory of Siegel modular forms*, Nagoya Math. J. **95** (1984), 73–84.
22. Y. Kitaoka, *A space of Siegel modular forms closed under the action of Hecke operators*, Proc. Japan Acad. Ser. A Math. Sci. **70** (1994), 194–197.
23. H. Klingen, *Über die Erzeugenden gewisser Modulgruppen*, Nachr. Akad. Wiss. Goettingen. Math.-Phys. Kl. IIa. 1956 (1956), 173–185.
24. W. Kohnen, *Modular forms of half-integral weight on* $\Gamma_0(4)$, Math. Ann. **248** (1980), 249–266.
25. W. Kohnen, *Lifting modular forms of half-integral weight to Siegel modular forms of even genus*, Math. Ann. **322** (2002), 787–809.
26. W. Kohnen and D. Zagier, *Values of L-series of modular forms at the center of the critical strip*, Invent. Math. **64** (1981), 175–198.
27. H. Kojima, *An arithmetic of Hermitian modular forms of degree two*, Invent. Math. **69** (1982), 217–227.

28. A. Krieg, *The Maass spaces on the Hermitian half-space of degree* 2, Math. Ann. **289** (1991), 663–681.
29. S. Kudla and W. J. Sweet Jr., *Degenerate principal series representations for* $\mathrm{U}(n,n)$, Israel J. Math. **98** (1997), 253–306.
30. N. Kurokawa, *Examples of eigenvalues of Hecke operators on Siegel cusp forms of degree two*, Invent. Math. **49** (1978), 149–165.
31. H. Maass, *Über eine Spezialschar von Modulformen zweiten Grades*, Invent. Math **52** (1979), 95–104.
32. H. Maass, *Über eine Spezialschar von Modulformen zweiten Grades, II* Invent. Math. **53** (1979), 249–253.
33. H. Maass, *Über eine Spezialschar von Modulformen zweiten Grades, III* Invent. Math. **53** (1979), 255–265.
34. T. Miyake, *Modular Forms*, Springer, (1989).
35. S. Nagaoka, *On Eisenstein series for the Hermitian modular groups and the Jacobi groups*, Abh. Math. Sem. Univ. Hamburg **62** (1992), 117–146.
36. S. Raghavan and J. Sengupta, *A Dirichlet series for Hermitian modular forms of degree* 2, Acta Arith. **58** (1991), 181–201.
37. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publ. Math. Soc. Japan **11**, Iwanami Shoten and Princeton University Press, (1971).
38. G. Shimura, *On modular forms of half integral weight*, Ann. of Math. **97** (1973), 440–481.
39. G. Shimura, *On certain reciprocity-laws for theta functions and modular forms*, Acta Math. **141** (1978), 35–71.
40. G. Shimura, *On Eisenstein series*, Duke Math. J. **50** (1983), 417–476.
41. G. Shimura, *Euler products and Fourier coefficients of automorphic forms on symplectic groups*, Invent. Math. **116** (1994), 531–576.
42. G. Shimura, *Euler products and Eisenstein series*, CBMS Regional Conference Series in Mathematics **93** the American Mathematical Society, Providence, RI, (1997).
43. T. Sugano, *On Maass space for* $\mathrm{SU}(2,2)$ (Japanese), Surikaisekikenkyusho Kokyuroku (RIMS Kokyuroku) **546** (1985), 1–16.
44. J. Tate, *Number theoretic background*, Proc. Sympos. Pure Math., XXXIII, patr 2 (1979), 3–26.
45. J.-L. Waldspurger, *Correspondances de Shimura et quaternions*, Forum Math. **3** (1991), 219–307.
46. D. Zagier, *Sur la conjecture de Saito-Kurokawa* (d'après H. Maass), Seminar on Number Theory 1979–80, Paris, Prog. Math., 12, Birkhäuser, Boston, Mass. 371–394 (1981).

3

Multiplicity-free Theorems of the Restrictions of Unitary Highest Weight Modules with respect to Reductive Symmetric Pairs

Toshiyuki Kobayashi

RIMS, Kyoto University
Kyoto 606-8502, Japan
toshi@kurims.kyoto-u.ac.jp

Summary. The complex analytic methods have found a wide range of applications in the study of multiplicity-free representations. This article discusses, in particular, its applications to the question of restricting highest weight modules with respect to reductive symmetric pairs. We present a number of multiplicity-free branching theorems that include the multiplicity-free property of some of known results such as the Clebsh–Gordan–Pieri formula for tensor products, the Plancherel theorem for Hermitian symmetric spaces (also for line bundle cases), the Hua–Kostant–Schmid K-type formula, and the canonical representations in the sense of Vershik–Gelfand–Graev. Our method works in a uniform manner for both finite and infinite dimensional cases, for both discrete and continuous spectra, and for both classical and exceptional cases.

Key words: multiplicity-free representation, branching rule, symmetric pair, highest weight module, Hermitian symmetric space, reproducing kernel, semisimple Lie group

Subject Classifications: Primary 22E46, Secondary 32A37, 05E15, 20G05, 53C35

1 Introduction and statement of main results

The purpose of this article is to give a quite detailed account of the theory of multiplicity-free representations based on a non-standard method (*visible actions* on complex manifolds) through its application to branching problems. More precisely, we address the question of restricting irreducible highest weight representations π of reductive Lie groups G with respect to symmetric pairs (G, H). Then, our main goal is to give a simple and sufficient condition on the triple (G, H, π) such that the restriction $\pi|_H$ is multiplicity-free. We shall see that our method works in a uniform way for infinite- and finite-dimensional representations, for classical and exceptional cases, and for continuous and discrete spectra.

This article is an outgrowth of the manuscript [44] which I did not publish, but which has been circulated as a preprint. Since then, we have extended the theory, in particular, to the following three directions:

1) the generalization of our main machinery (Theorem 2.2) to the vector bundle case ([49]),
2) the theory of 'visible actions' on complex manifolds ([50, 51, 52]),
3) 'multiplicity-free geometry' for coadjoint orbits ([53]).

We refer the reader to our paper [47] for a precise statement of the general results and an exposition of the related topics that have recently developed.

In this article we confine ourselves to the line bundle case. On the one hand, this is sufficiently general to produce many interesting consequences, some of which are new and others which may be regarded as prototypes of various multiplicity-free branching theorems (e.g., [5, 10, 46, 54, 58, 66, 68, 81, 90, 92]). On the other hand, the line bundle case is sufficiently simple, so that we can illustrate the essence of our main ideas without going into technical details. Thus, keeping the spirit of [44], we have included here the proof of our method (Theorem 2.2), its applications to multiplicity-free theorems (Theorems A–F), and the explicit formulae (Theorems 8.3, 8.4, and 8.11), except that we referred to another paper [50] for the proof of some algebraic lemmas on the triple of involutions of Lie algebras (Lemmas 3.6 and 7.5).

1.1 Definition of multiplicity-free representations

Let us begin by recalling the concept of the multiplicity-free decomposition of a unitary representation.

Suppose H is a Lie group of type I in the sense of von Neumann algebras. Any reductive Lie group is of type I as well as any algebraic group. We denote by $\widehat{H}$ the unitary dual of H, that is, the set of equivalence classes of irreducible unitary representations of H. The unitary dual $\widehat{H}$ is endowed with the Fell topology.

Suppose that $(\pi, \mathcal{H})$ is a unitary representation of H defined on a (second countable) Hilbert space $\mathcal{H}$. By a theorem of Mautner, π is decomposed uniquely into irreducible unitary representations of H in terms of the direct integral of Hilbert spaces:

$$\pi \simeq \int_{\widehat{H}} m_\pi(\mu)\mu \, d\sigma(\mu), \tag{1.1.1}$$

where $d\sigma(\mu)$ is a Borel measure on $\widehat{H}$, and the multiplicity function $m_\pi : \widehat{H} \to \mathbb{N} \cup \{\infty\}$ is uniquely defined almost everywhere with respect to the measure $d\sigma$.

Let $\mathrm{End}(\mathcal{H})$ be the ring of continuous operators on $\mathcal{H}$, and $\mathrm{End}_H(\mathcal{H})$ the subring of H-intertwining operators, that is, the commutant of $\{\pi(g) : g \in H\}$ in $\mathrm{End}(\mathcal{H})$.

Definition 1.1. *We say that the unitary representation* $(\pi, \mathcal{H})$ *is* multiplicity-free *if the ring* $\mathrm{End}_H(\mathcal{H})$ *is commutative.*

It is not difficult to see that this definition is equivalent to the following property:

$$m_\pi(\mu) \le 1 \text{ for almost all } \mu \in \widehat{H} \text{ with respect to the measure } d\sigma(\mu)$$

by Schur's lemma for unitary representations. In particular, it implies that any irreducible unitary representation μ of H occurs at most once as a subrepresentation of π.

1.2 Multiplicities for inductions and restrictions

With regard to the question of finding irreducible decompositions of unitary representations, there are two fundamental settings: one is the induced representation from smaller groups (e.g., harmonic analysis on homogeneous spaces), and the other is the restriction from larger groups (e.g., tensor product representations).

To be more rigorous, suppose G is a Lie group, and H is a closed subgroup of G. The G-irreducible decomposition of the induced representation $L^2\text{-}\mathrm{Ind}_H^G\,\tau$ $(\tau \in \widehat{H})$ is called the *Plancherel formula*, while the H-irreducible decomposition of the restriction $\pi|_H$ $(\pi \in \widehat{G})$ is referred to as the *branching law*.

This subsection examines multiplicities in the irreducible decomposition of the induction and the restriction for reductive symmetric pairs (G, H) (see Subsection 3.1 for the definition).

Let us start with the induced representation. Van den Ban [2] proved that the multiplicity in the Plancherel formula for $L^2\text{-}\mathrm{Ind}_H^G\,\tau$ is finite as far as $\dim \tau < \infty$. In particular, this is the case if τ is the trivial representation **1**. Over the past several decades, the induced representation $L^2\text{-}\mathrm{Ind}_H^G\,\mathbf{1}$ has developed its own identity (harmonic analysis on reductive symmetric spaces G/H) as a rich and meaningful part of mathematics.

In contrast, the multiplicities of the branching law of the restriction $\pi|_H$ $(\pi \in \widehat{G})$ are usually infinite. For instance, we saw in [36] that this is the case if $(G, H) = (GL(p+q, \mathbb{R}), GL(p, \mathbb{R}) \times GL(q, \mathbb{R}))$ where $\min(p, q) \geq 2$, for any tempered representation π of G. In this article, we illuminate by Example 6.3 this wild behavior.

In light of such a wild phenomenon of branching laws for reductive symmetric pairs (G, H) with H non-compact, we proposed in [38, 40] to seek a 'nice' class of the triple (G, H, π) in which a systematic study of the restriction $\pi|_H$ could be launched.

Finiteness of multiplicities is a natural requirement for this program. By also imposing discrete decomposability on the restriction $\pi|_H$, we established the general theory for *admissible restriction* in [38, 40, 41] and found that there exist fairly rich triples (G, H, π) that enjoy this nice property. It is noteworthy that new interesting directions of research in the framework of admissible restrictions have been recently developed by M. Duflo, D. Gross, J.-S. Huang, J.-S. Li, S.-T. Lee, H.-Y. Loke, T. Oda, P. Pandžić, G. Savin, B. Speh, J. Vargas, D. Vogan, and N. Wallach (see [45, 48] and references therein).

Multiplicity-freeness is another ideal situation in which we may expect an especially simple and detailed study of the branching law of $\pi|_H$. Thus, we aim for principles that lead us to an abundant family of multiplicity-free cases. Among them, a well-known one is the dual pair correspondence, which has given fruitful examples in the infinite-dimensional theory in the following setting:

a) G is the metaplectic group, and π is the Weil representation.
b) $H = H_1 \cdot H_2$ forms a dual pair, that is, H_1 is the commutant of H_2 in G, and vice versa.

This paper uses a new principle that generates multiplicity-free representations. The general theory discussed in Section 2 brings us to uniformly bounded multiplicity theorems (Theorems B and D) and multiplicity-free theorems (Theorems A, C, E and F) in the following setting:

a) π is a unitary highest weight representation of G (see Subsection 1.3).
b) (G, H) is a symmetric pair (see Subsection 1.4).

We note that we allow the case where continuous spectra occur in the branching law, and consequently, irreducible summands are not always highest weight representations.

We remark that our bounded multiplicity theorems for the restriction $\pi|_H$ (π: highest weight module) may be regarded as the counterpart of the bounded multiplicity theorem for the induction L^2-$\mathrm{Ind}_H^G\,\tau$ (τ: finite-dimensional representation) due to van den Ban.

1.3 Unitary highest weight modules

Let us recall the basic notion of highest weight modules.

Let G be a non-compact simple Lie group, θ a Cartan involution of G, and $K := \{g \in G : \theta g = g\}$. We write $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ for the Cartan decomposition of the Lie algebra $\mathfrak{g}$ of G, corresponding to the Cartan involution θ.

We assume that G is of *Hermitian type*, that is, the Riemannian symmetric space G/K carries the structure of a Hermitian symmetric space, or equivalently, the center $\mathfrak{c}(\mathfrak{k})$ of $\mathfrak{k}$ is non-trivial. The classification of simple Lie algebras $\mathfrak{g}$ of Hermitian type is given as follows:

$$\mathfrak{su}(p,q)\,,\ \mathfrak{sp}(n,\mathbb{R})\,,\ \mathfrak{so}(m,2)\ (m \neq 2)\,,\ \mathfrak{e}_{6(-14)}\,,\ \mathfrak{e}_{7(-25)}\,.$$

Such a Lie algebras $\mathfrak{g}$ satisfies the rank condition:

$$\operatorname{rank} G = \operatorname{rank} K, \tag{1.3.1}$$

or equivalently, a Cartan subalgebra of $\mathfrak{k}$ becomes a Cartan subalgebra of $\mathfrak{g}$. By a theorem of Harish-Chandra, the rank condition (1.3.1) is equivalent to the existence of (relative) discrete series representations of G. Here, an irreducible unitary representation $(\pi, \mathcal{H})$ is called a *(relative) discrete series representation* of G if the matrix coefficient $g \mapsto (\pi(g)u, v)$ is square integrable on G (modulo its center) for any $u, v \in \mathcal{H}$.

If $\mathfrak{g}$ is a simple Lie algebra of Hermitian type, then there exists a characteristic element $Z \in \mathfrak{c}(\mathfrak{k})$ such that

$$\mathfrak{g}_{\mathbb{C}} := \mathfrak{g} \otimes \mathbb{C} = \mathfrak{k}_{\mathbb{C}} \oplus \mathfrak{p}_+ \oplus \mathfrak{p}_- \tag{1.3.2}$$

is the eigenspace decomposition of $\mathrm{ad}(Z)$ with eigenvalues 0, $\sqrt{-1}$ and $-\sqrt{-1}$, respectively. We note that $\dim \mathfrak{c}(\mathfrak{k}) = 1$ if $\mathfrak{g}$ is a simple Lie algebra of Hermitian type, and therefore $\mathfrak{c}(\mathfrak{k}) = \mathbb{R}Z$.

Suppose V is an irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module. We set

$$V^{\mathfrak{p}_+} := \{v \in V : Yv = 0 \text{ for any } Y \in \mathfrak{p}_+\}\,. \tag{1.3.3}$$

Since K normalizes $\mathfrak{p}_+$, $V^{\mathfrak{p}_+}$ is a K-submodule. Further, $V^{\mathfrak{p}_+}$ is either zero or an irreducible finite-dimensional representation of K. We say V is a *highest weight module* if $V^{\mathfrak{p}_+} \neq \{0\}$.

Definition 1.3. *Suppose π is an irreducible unitary representation of G on a Hilbert space $\mathcal{H}$. We set $\mathcal{H}_K := \{v \in \mathcal{H} : \dim_{\mathbb{C}} \mathbb{C}\text{-span}\{\pi(k)v : k \in K\} < \infty\}$. Then, $\mathcal{H}_K$ is a dense subspace of $\mathcal{H}$, on which the differential action $d\pi$ of the Lie algebra $\mathfrak{g}$ (and consequently that of its complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$) and the action of the compact subgroup K is well defined. We say $\mathcal{H}_K$ is* the underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module of $(\pi, \mathcal{H})$. *We say $(\pi, \mathcal{H})$ is a* unitary highest weight representation *of G if $\mathcal{H}_K^{\mathfrak{p}_+} \neq \{0\}$. Then, π is* of scalar type *(or* of scalar minimal K-type*) if $\mathcal{H}_K^{\mathfrak{p}_+}$ is one-dimensional; π is a* (relative) holomorphic discrete series representation *for G if the matrix coefficient $g \mapsto (\pi(g)u, v)$ is square integrable on G modulo its center for any $u, v \in \mathcal{H}$. Lowest weight modules and anti-holomorphic discrete series representations are defined similarly with $\mathfrak{p}_+$ replaced by $\mathfrak{p}_-$.*

This definition also applies to G which is not simple (see Subsection 8.1).

The classification of irreducible unitary highest weight representations was accomplished by Enright–Howe–Wallach [12] and H. Jakobsen [30] independently; see also [13]. There always exist infinitely many (relative) holomorphic discrete series representations of scalar type for any non-compact simple Lie group of Hermitian type.

1.4 Involutions on Hermitian symmetric spaces

Suppose G is a non-compact simple Lie group of Hermitian type. Let τ be an involutive automorphism of G commuting with the Cartan involution θ. We use the same letter τ to denote its differential. Then τ stabilizes $\mathfrak{k}$ and also $\mathfrak{c}(\mathfrak{k})$. Because $\tau^2 = \mathrm{id}$ and $\mathfrak{c}(\mathfrak{k}) = \mathbb{R}Z$, we have the following two possibilities:

$$\tau Z = Z\,, \tag{1.4.1}$$

$$\tau Z = -Z\,. \tag{1.4.2}$$

The geometric meanings of these conditions become clear in the context of the embedding $G^\tau/K^\tau \hookrightarrow G/K$, where $G^\tau := \{g \in G : \tau g = g\}$ and $K^\tau := G^\tau \cap K$ (see [14, 27, 28, 35]). The condition (1.4.1) implies:

1-a) τ acts **holomorphically** on the Hermitian symmetric space G/K,
1-b) $G^\tau/K^\tau \hookrightarrow G/K$ defines a complex submanifold,

whereas the condition (1.4.2) implies:

2-a) τ acts **anti-holomorphically** on G/K,
2-b) $G^\tau/K^\tau \hookrightarrow G/K$ defines a totally real submanifold.

Definition 1.4. *We say the involutive automorphism τ is* of holomorphic type *if* (1.4.1) *is satisfied, and is of* anti-holomorphic type *if* (1.4.2) *is satisfied. The same terminology will be applied also to the symmetric pair (G, H) (or its Lie algebras $(\mathfrak{g}, \mathfrak{h})$) corresponding to the involution τ.*

Here we recall that (G, H) is called a *symmetric pair* corresponding to τ if H is an open subgroup of G^τ (see Subsections 3.1 and 3.2). We note that the Lie algebra $\mathfrak{h}$ of H is equal to $\mathfrak{g}^\tau := \{X \in \mathfrak{g} : \tau X = X\}$. The classification of symmetric pairs $(\mathfrak{g}, \mathfrak{g}^\tau)$ for simple Lie algebras $\mathfrak{g}$ was accomplished by M. Berger [6]. The classification of symmetric pairs $(\mathfrak{g}, \mathfrak{g}^\tau)$ of holomorphic type (respectively, of anti-holomorphic type) is regarded as a subset of Berger's list, and will be presented in Table 3.4.1 (respectively, Table 3.4.2).

1.5 Multiplicity-free restrictions — infinite-dimensional case

We are ready to state our main results. Let G be a non-compact simple Lie group of Hermitian type, and (G, H) a symmetric pair.

Theorem A (multiplicity-free restriction). *If π is an irreducible unitary highest weight representation of scalar type of G, then the restriction $\pi|_H$ is multiplicity-free.*

The branching law of the restriction $\pi|_H$ may and may not contain discrete spectra in Theorem A. If (G, H) is of holomorphic type, then the restriction $\pi|_H$ is discretely decomposable (i.e., there is no continuous spectrum in the branching law); see Fact 5.1. Besides, the following theorem asserts that the multiplicities are still uniformly bounded even if we drop the assumption that π is of scalar type.

Theorem B (uniformly bounded multiplicities). *We assume that the symmetric pair (G, H) is of holomorphic type. Let π be an irreducible unitary highest weight representation of G.*
1) The restriction $\pi|_H$ splits into a discrete Hilbert sum of irreducible unitary representations of H:

$$\pi|_H \simeq \sum_{\mu \in \widehat{H}}^{\oplus} m_\pi(\mu)\mu\,,$$

and the multiplicities are uniformly bounded:

$$C(\pi) := \sup_{\mu \in \widehat{H}} m_\pi(\mu) < \infty\,.$$

2) $C(\pi) = 1$ if π is of scalar type.

The second statement is a direct consequence of Theorems A and B (1). As we shall see in Section 6, such uniform boundedness theorem does not hold in general if π is not a highest weight representation (see Examples 6.2 and 6.3).

Here are multiplicity-free theorems for the decomposition of tensor products, which are parallel to Theorems A and B:

Theorem C (multiplicity-free tensor product). *Let π_1 and π_2 be irreducible unitary highest (or lowest) weight representations of scalar type. Then the tensor product $\pi_1 \widehat{\otimes} \pi_2$ is multiplicity-free as a representation of G.*

Here, $\pi_1 \widehat{\otimes} \pi_2$ stands for the tensor product representation of two unitary representations $(\pi_1, \mathcal{H}_1)$ and $(\pi_2, \mathcal{H}_2)$ realized on the completion $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$ of the pre-Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. (We do not need to take the completion if at least one of $\mathcal{H}_1$ or $\mathcal{H}_2$ is finite dimensional.) Theorem C asserts that multiplicities in the direct integral of the irreducible decomposition are not greater than one in both discrete and continuous spectra. We note that continuous spectra appear in the irreducible decomposition of the tensor product representation $\pi_1 \widehat{\otimes} \pi_2$ only if

$$\begin{cases} \pi_1 \text{ is a highest weight representation, and} \\ \pi_2 \text{ is a lowest weight representation,} \end{cases}$$

or in reverse order.

If π_1 and π_2 are simultaneously highest weight representations (or simultaneously lowest weight representations), then the tensor product $\pi_1 \widehat{\otimes} \pi_2$ decomposes discretely. Dropping the assumption of 'scalar type', we have still a uniform estimate of multiplicities:

Theorem D (uniformly bounded multiplicities). *Let π_1 and π_2 be two irreducible unitary highest weight representations of G.*
1) The tensor product $\pi_1 \widehat{\otimes} \pi_2$ splits into a discrete Hilbert sum of irreducible unitary representations of G:

$$\pi_1 \widehat{\otimes} \pi_2 \simeq \sum_{\mu \in \widehat{G}}^{\oplus} m_{\pi_1,\pi_2}(\mu)\mu \,,$$

and the multiplicities $m_{\pi_1,\pi_2}(\mu)$ are uniformly bounded:

$$C(\pi_1, \pi_2) := \sup_{\mu \in \widehat{G}} m_{\pi_1,\pi_2}(\mu) < \infty \,.$$

2) $C(\pi_1, \pi_2) = 1$ if both π_1 and π_2 are of scalar type.

Remark 1.5. For classical groups, we can relate the constants $C(\pi)$ and $C(\pi_1, \pi_2)$ to the *stable constants* of branching coefficients of finite-dimensional representations in the sense of F. Sato [77] by using the see-saw dual pair correspondence due to R. Howe [23].

Our machinery that gives the above multiplicity-free theorems is built on complex geometry, and we shall explicate the general theory for the line bundle case in Section 2. The key idea is to transfer properties on representations (e.g., unitarity, multiplicity-freeness) into the corresponding properties of reproducing kernels, which we analyze by geometric methods.

1.6 Multiplicity-free restrictions — finite-dimensional case

Our method yields multiplicity-free theorems not only for infinite-dimensional representations but also for finite-dimensional representations.

This subsection presents multiplicity-free theorems that are regarded as a 'finite dimensional version' of Theorems A and C. They give a unified explanation of the multiplicity-free property of previously known branching formulae obtained by combinatorial methods such as the Littlewood–Richardson rule, Koike–Terada's Young diagrammatic methods, Littelmann's path method, minor summation formulae, etc. (see [25, 55, 62, 68, 73, 80] and references therein). They also contain some 'new' cases, for which there are, to the best of our knowledge, no explicit branching formulae in the literature.

To state the theorems, let $\mathfrak{g}_{\mathbb{C}}$ be a complex simple Lie algebra, and $\mathfrak{j}$ a Cartan subalgebra. We fix a positive root system $\Delta^+(\mathfrak{g}_{\mathbb{C}}, \mathfrak{j})$, and write $\alpha_1, \ldots, \alpha_n$ for the simple roots. Let $\omega_1, \ldots, \omega_n$ be the corresponding fundamental weights. We denote by $\pi_\lambda \equiv \pi_\lambda^{\mathfrak{g}_{\mathbb{C}}}$ the irreducible finite-dimensional representation of $\mathfrak{g}_{\mathbb{C}}$ with highest weight λ.

We say π_λ is of **pan type** if λ is a scalar multiple of some ω_i such that the nilradical of the maximal parabolic subalgebra corresponding to α_i is abelian (see Lemma 7.3.1 for equivalent definitions).

Theorem E (multiplicity-free restriction — finite-dimensional case). *Let π be an arbitrary irreducible finite-dimensional representation of $\mathfrak{g}_{\mathbb{C}}$ of pan type, and let $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}})$ be any symmetric pair. Then, the restriction $\pi|_{\mathfrak{h}_{\mathbb{C}}}$ is multiplicity-free.*

Theorem F (multiplicity-free tensor product — finite-dimensional case). *The tensor product $\pi_1 \otimes \pi_2$ of any two irreducible finite-dimensional representations π_1 and π_2 of pan type is multiplicity-free.*

Theorems E and F are the counterpart to Theorems A and C for finite-dimensional representations. The main machinery of the proof is again Theorem 2.2.

Alternatively, one could verify Theorems E and F by a classical technique: finding an open orbit of a Borel subgroup. For example, Littelmann [61] and Panyushev independently classified the pair of maximal parabolic subalgebras $(\mathfrak{p}_1, \mathfrak{p}_2)$ such that the diagonal action of a Borel subgroup B of a complex simple Lie group $G_{\mathbb{C}}$ on $G_{\mathbb{C}}/P_1 \times G_{\mathbb{C}}/P_2$ has an open orbit. Here, P_1, P_2 are the corresponding maximal parabolic subgroups of $G_{\mathbb{C}}$. This gives another proof of Theorem F.

The advantage of our method is that it enables us to understand (or even to discover) the multiplicity-free property simultaneously for both infinite- and finite-dimensional representations, for both continuous and discrete spectra, and for both classical and exceptional cases by the single principle. This is because our main machinery (Theorem 2.2) uses only a *local* geometric assumption (see Remark 2.3.2 (2)). Thus, we can verify it at the same time for compact and non-compact complex manifolds, and in turn get finite and infinite-dimensional results, respectively.

Once we tell a priori that a representation is multiplicity-free, we may be tempted to find explicitly its irreducible decomposition. Recently, S. Okada [68] found explicit branching laws for some classical cases that arise in Theorems E and F by using minor summation formulae, and H. Alikawa [1] for $(\mathfrak{g}, \mathfrak{h}) = (\mathfrak{e}_6, \mathfrak{f}_4)$ corresponding to Theorem E. We note that the concept of pan type representations includes *rectangular-shaped* representations of classical groups (see [58, 68]).

There are also some few cases where $\pi_1 \otimes \pi_2$ is multiplicity-free even though neither π_1 nor π_2 is of pan type. See the recent papers [46] or [81] for the complete list of such pairs (π_1, π_2) for $\mathfrak{g}_{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C})$. The method in [46] to find all such pairs is geometric and based on the 'vector bundle version' of Theorem 2.2 proved in [49], whereas the method in [81] is combinatorial and based on case-by-case argument.

We refer the reader to our papers [50, 51, 52] for some further results relevant to Theorems E and F along the same line of argument here.

1.7 SL_2 examples

We illustrate the above theorems by SL_2 examples.

Example 1.7. 1) We denote by π_n the holomorphic discrete series representation of $G = SL(2, \mathbb{R})$ with minimal K-type χ_n ($n \geq 2$), where we write χ_n for the character of $K = SO(2)$ parametrized by $n \in \mathbb{Z}$. Likewise π_{-n} denotes the anti-holomorphic discrete series representation of $SL(2, \mathbb{R})$ with minimal K-type χ_{-n} ($n \geq 2$). We note that any holomorphic discrete series of $SL(2, \mathbb{R})$ is of scalar type.

We write $\pi^{\varepsilon}_{\sqrt{-1}\nu}$ ($\varepsilon = \pm 1, \nu \in \mathbb{R}$) for the unitary principal series representations of $SL(2, \mathbb{R})$. We have a unitary equivalence $\pi^{\varepsilon}_{\sqrt{-1}\nu} \simeq \pi^{\varepsilon}_{-\sqrt{-1}\nu}$. We write χ_ζ for the unitary character of $SO_0(1, 1) \simeq \mathbb{R}$ parametrized by $\zeta \in \mathbb{R}$.

Let $m \geq n \geq 2$. Then, the following branching formulae hold. All of them are multiplicity-free, as is 'predicted' by Theorems A and C:

$$\pi_n|_{SO_0(1,1)} \simeq \int_{-\infty}^{\infty} \chi_\zeta \, d\zeta \, , \tag{1.7.1 (a)}$$

$$\pi_n|_{SO(2)} \simeq \sum_{k \in \mathbb{N}}^{\oplus} \chi_{n+2k} \, , \tag{1.7.1 (b)}$$

$$\pi_m \widehat{\otimes} \pi_{-n} \simeq \int_0^{\infty} \pi^{(-1)^{m-n}}_{\sqrt{-1}\nu} d\nu \quad \oplus \quad \sum_{\substack{k \in \mathbb{N} \\ 0 \leq 2k \leq m-n-2}} \pi_{m-n-2k} \, , \tag{1.7.1 (c)}$$

$$\pi_m \widehat{\otimes} \pi_n \simeq \sum_{k \in \mathbb{N}}^{\oplus} \pi_{m+n+2k} \, . \tag{1.7.1 (d)}$$

The key assumption of our main machinery (Theorem 2.2) that leads us to Theorems A and C is illustrated by the following geometric results in this SL_2 case:
i) Given any element z in the Poincaré disk D, there exists $\varphi \in \mathbb{R}$ such that $e^{\sqrt{-1}\varphi}z = \overline{z}$. In fact, one can take $\varphi = -2\arg z$. This is the geometry that explains the multiplicity-free property of (1.7.1) (b).
ii) Given any two elements $z, w \in D$, there exists a linear fractional transform T on D such that $T(z) = \overline{z}$ and $T(w) = \overline{w}$. This is the geometry for (1.7.1) (d).

These are examples of the geometric view point that we pursued in [50] for symmetric pairs.

2) Here is a 'finite-dimensional version' of the above example. Let π_n be the irreducible $n+1$-dimensional representation of $SU(2)$. Then we have the following branching formulae: For $m, n \in \mathbb{N}$,

$$\pi_n|_{SO(2)} \simeq \chi_n \oplus \chi_{n-2} \oplus \cdots \oplus \chi_{-n}\,, \tag{1.7.1 (e)}$$

$$\pi_m \otimes \pi_n \simeq \pi_{n+m} \oplus \pi_{n+m-2} \oplus \cdots \oplus \pi_{|n-m|}\,. \tag{1.7.1 (f)}$$

The formula (1.7.1) (e) corresponds to the character formula, whereas (1.7.1) (f) is known as the Clebsch–Gordan formula. The multiplicity-free property of these formulae is the simplest example of Theorems E and F.

1.8 Analysis on multiplicity-free representations

Multiplicity-free property arouses our interest in developing beautiful analysis on such representations, as we discussed in Subsection 1.6 for finite-dimensional cases. This subsection picks up some recent topics about detailed analysis on multiplicity-free representations for infinite-dimensional cases.

Let G be a connected, simple non-compact Lie group of Hermitian type. We begin with branching laws without continuous spectra, and then discuss branching laws with continuous spectra.

1) (Discretely decomposable case) Let (G, H) be a symmetric pair of holomorphic type. Then, any unitary highest weight representation π of G decomposes discretely when restricted to H (Fact 5.1).

1-a) Suppose now that π is a holomorphic discrete series representation. L.-K. Hua [26], B. Kostant, W. Schmid [78] and K. Johnson [32] found an explicit formula of the restriction $\pi|_K$ (K-type formula). This turns out to be multiplicity-free. Alternatively, the special case of Theorem B (2) by setting $H = K$ gives a new proof of this multiplicity-free property.

1-b) Furthermore, we consider a generalization of the Hua–Kostant–Schmid formula from compact H to non-compact H, for which Theorem B (2) still ensures that the generalization will be multiplicity-free. This generalized formula is stated in Theorem 8.3, which was originally given in [39, Theorem C]. In Section 8, we give a full account of its proof. W. Bertram and J. Hilgert [7] obtained some special cases independently, and Ben Saïd [5] studied a quantative estimate of this multiplicity-free H-type formula (see also [90, 91] for some *singular* cases).

1-c) The branching formulae of the restriction of *singular* highest weight representations π are also interesting. For instance, the restriction of the Segal–Shale–Weil representation ϖ of $Mp(n, \mathbb{R})$ with respect to $U(p, n-p)$ (more precisely, its double covering) decomposes discretely into a multiplicity-free sum of the so-called *ladder representations* of $U(p, n-p)$ (e.g., [33, Introduction]). This multiplicity-free property is a special case of Howe's correspondence because $(U(p, n-p), U(1))$ forms a dual pair in $Mp(n, \mathbb{R})$, and also is a special case of Theorem A because $(\mathfrak{sp}(n, \mathbb{R}), \mathfrak{u}(p, n-p))$ forms a symmetric pair. Explicit branching laws for most of classical cases corresponding to Theorems B (2) and D (2) (see Theorems 8.3, 8.4, 8.11) can be obtained by using the 'see-saw dual pair', which we hope to report in another paper.

2) (Branching laws with continuous spectra) Suppose π_1 is a highest weight module and π_2 is a lowest weight module, both being of scalar type.

2-a) If both π_1 and π_2 are discrete series representations in addition, then the tensor product $\pi_1 \widehat{\otimes} \pi_2$ is unitarily equivalent to the regular representation on $L^2(G/K, \chi)$, the Hilbert space of L^2-sections of the G-equivalent line bundle $G \times_K \mathbb{C}_\chi \to G/K$ associated to some unitary character χ of K (R. Howe [23], J. Repka [74]). In particular, Theorem C gives a new proof of the multiplicity-free property of the Plancherel formula for $L^2(G/K, \chi)$. Yet another proof of the multiplicity-free property of $L^2(G/K, \chi)$ was given in [47, Theorem 21] by still applying Theorem 2.2 to the *crown domain* (equivalently, the Akhiezer–Gindikin domain) of the Riemannian symmetric space G/K. The explicit decomposition of $L^2(G/K, \chi)$ was found by J. Heckman [20] and N. Shimeno [79] that generalizes the work of Harish-Chandra, S. Helgason, and S. Gindikin–F. Karpelevich for the trivial bundle case.

In contrast to Riemannian symmetric spaces, it is known that the 'multiplicity-free property' in the Plancherel formula fails for (non-Riemannian) symmetric spaces G/H in general (see [3, 8] for the description of the multiplicity of the most continuous series representations for G/H in terms of Weyl groups).

2-b) Similar to the case 2-a), the restriction $\pi|_H$ for a symmetric pair (G, H) of non-holomorphic type is multiplicity-free and is decomposed into only continuous spectra if π is a holomorphic discrete series of scalar type. This case was studied by G. Ólafsson–B. Ørsted ([69]).

2-c) Theorem C applied to non-discrete series representations π_1 and π_2 (i.e., tensor products of *singular* unitary highest weight representations) provides new settings of multiplicity-free branching laws. They might be interesting from the view point of representation theory because they construct 'small' representations as discrete summands. (We note that irreducible unitary representations of reductive Lie groups have not been classified even in the spherical case. See [4] for the split case.) They might be interesting also from the view point of spectral theory and harmonic analysis which is relevant to the *canonical representation* in the sense of Vershik–Gelfand–Graev. Once we know the branching law is a priori multiplicity-free, it is promising to obtain its explicit formula. Some special cases have been worked on in this direction so far, for $G = SL(2, \mathbb{R})$ by V. F. Molchanov [64]; for $G = SU(2, 2)$ by B. Ørsted and G. Zhang [70]; for $G = SU(n, 1)$ by G. van Dijk and S. Hille

[10]; for $G = SU(p, q)$ by Y. Neretin and G. Ol'shanskiĭ [66, 67]. See also G. van Dijk–M. Pevzner [11], M. Pevzner [72] and G. Zhang [92]. Their results show that a different family of irreducible unitary representations (sometimes, spherical complementary series representations) can occur in the same branching laws and each multiplicity is not greater than one.

1.9 Organization of this article

This paper is organized as follows: In Section 2, we give a proof of an abstract multiplicity-free theorem (Theorem 2.2) in the line bundle setting. This is an extension of a theorem of Faraut–Thomas [15], whose idea may go back to Gelfand's proof [17] of the commutativity of the Hecke algebra $L^1(K\backslash G/K)$. Theorem 2.2 is a main method in this article to find various multiplicity-free theorems. In Section 3, we use Theorem 2.2 to give a proof of Theorem A. The key idea is the reduction of the geometric condition (2.2.3) (*strongly visible action* in the sense of [47]) to the existence problem of a 'nice' involutive automorphism σ of G satisfying a certain rank condition. Section 4 considers the multiplicity-free theorem for the tensor product representations of two irreducible highest (or lowest) weight modules and gives a proof of Theorem C. Sections 5 and 6 examine our assumptions in our multiplicity-free theorems (Theorems A and C). That is, we drop the assumption of 'scalar type' in Section 5 and prove that multiplicities are still uniformly bounded (Theorems B and D). We note that multiplicities can be greater than one in this generality. In Section 6, we leave unchanged the assumption that (G, H) is a symmetric pair, and relax the assumption that π is a highest weight module. We illustrate by examples a wild behavior of multiplicities without this assumption. In Section 7, analogous results of Theorems A and C are proved for finite-dimensional representations of compact groups. In Section 8, we present explicit branching laws that are assured a priori to be multiplicity-free by Theorems A and C. Theorem 8.4 generalizes the Hua–Kostant–Schmid formula. In Section 9 (Appendix) we present some basic results on homogeneous line bundles for the convenience of the reader, which give a sufficient condition for the assumption (2.2.2) in Theorem 2.2.

2 Main machinery from complex geometry

J. Faraut and E. Thomas [15], in the case of trivial twisting parameter, gave a sufficient condition for the commutativity of $\mathrm{End}_H(\mathcal{H})$ by using the theory of reproducing kernels, which we extend to the general, twisted case in this preliminary section. The proof parallels theirs, except that we need just find an additional condition (2.2.2) when we formalize Theorem 2.2 in the line bundle setting.

2.1 Basic operations on holomorphic line bundles

Let $\mathcal{L} \to D$ be a holomorphic line bundle over a complex manifold D. We denote by $\mathcal{O}(\mathcal{L}) \equiv \mathcal{O}(D, \mathcal{L})$ the space of holomorphic sections of $\mathcal{L} \to D$. Then $\mathcal{O}(\mathcal{L})$

carries a Fréchet topology by the uniform convergence on compact sets. If a Lie group H acts holomorphically and equivariantly on the holomorphic line bundle $\mathcal{L} \to D$, then H defines a (continuous) representation on $\mathcal{O}(\mathcal{L})$ by the pullback of sections.

Let $\{U_\alpha\}$ be trivializing neighborhoods of D, and $g_{\alpha\beta} \in \mathcal{O}^\times(U_\alpha \cap U_\beta)$ the transition functions of the holomorphic line bundle $\mathcal{L} \to D$. Then an anti-holomorphic line bundle $\overline{\mathcal{L}} \to D$ is a complex line bundle with the transition functions $\overline{g_{\alpha\beta}}$. We denote by $\overline{\mathcal{O}}(\overline{\mathcal{L}})$ the space of anti-holomorphic sections for $\overline{\mathcal{L}} \to D$.

Suppose σ is an anti-holomorphic diffeomorphism of D. Then the pullback $\sigma^*\mathcal{L} \to D$ is an anti-holomorphic line bundle over D. In turn, $\overline{\sigma^*\mathcal{L}} \to D$ is a holomorphic line bundle over D (see Appendix for more details).

2.2 Abstract multiplicity-free theorem

Here is the main machinery to prove various multiplicity-free theorems of branching laws including Theorems A and C (infinite-dimensional representations) and Theorems E and F (finite-dimensional representations).

Theorem 2.2. *Let $(\pi, \mathcal{H})$ be a unitary representation of a Lie group H. Assume that there exist an H-equivariant holomorphic line bundle $\mathcal{L} \to D$ and an anti-holomorphic involutive diffeomorphism σ of D with the following three conditions:*

There is an injective (continuous) H-intertwining map $\mathcal{H} \to \mathcal{O}(\mathcal{L})$. (2.2.1)

There exists an isomorphism of H-equivariant holomorphic line bundles
$\Psi : \mathcal{L} \xrightarrow{\sim} \overline{\sigma^*\mathcal{L}}$. (2.2.2)

Given $x \in D$, there exists $g \in H$ such that $\sigma(x) = g \cdot x$. (2.2.3)

Then, the ring $\operatorname{End}_H(\mathcal{H})$ of continuous H-intertwining operators on $\mathcal{H}$ is commutative. Consequently, $(\pi, \mathcal{H})$ is multiplicity-free (see Definition 1.1).

2.3 Remarks on Theorem 2.2

This subsection gives brief comments on Theorem 2.2. First, we consider a special case, and also a generalization.

Remark 2.3.1 *(specialization and generalization).* 1) Suppose $\mathcal{L} \to D$ is the trivial line bundle. Then, the condition (2.2.2) is automatically satisfied. In this case, Theorem 2.2 was proved in [15].
2) An extension of Theorem 2.2 to the equivariant vector bundle $\mathcal{V} \to D$ is the main subject of [49], where a more general multiplicity-free theorem is obtained under an additional condition that the isotropy representation of $H_x = \{h \in H : h \cdot x = x\}$ on the fiber $\mathcal{V}_x$ is multiplicity-free for generic $x \in D$. Obviously, the H_x-action on $\mathcal{V}_x$ is multiplicity-free for the case $\dim \mathcal{V}_x = 1$, namely, for the line bundle case.

Next, we examine the conditions (2.2.2) and (2.2.3).

Remark 2.3.2. 1) In many cases, the condition (2.2.2) is naturally satisfied. We shall explicate how to construct the bundle isomorphism Ψ in Lemma 9.4 for a Hermitian symmetric space D.
2) As the proof below shows, Theorem 2.2 still holds if we replace D by an H-invariant open subset D'. Thus, the condition (2.2.3) is **local**. The concept of **visible action** (see [46, 49, 51]) arises from the condition (2.2.3) on the base space D.
3) The condition (2.2.3) is automatically satisfied if H acts transitively on D. But we are interested in a more general setting where each H-orbit has a positive codimension in D. We find in Lemma 3.3 a sufficient condition for (2.2.3) in terms of rank condition for a symmetric space D.

2.4 Reproducing kernel

This subsection gives a quick summary for the reproducing kernel of a Hilbert space $\mathcal{H}$ realized in the space $\mathcal{O}(\mathcal{L})$ of holomorphic sections for a holomorphic line bundle $\mathcal{L}$ (see [49] for a generalization to the vector bundle case). Since the reproducing kernel $K_{\mathcal{H}}$ contains all the information on the Hilbert space $\mathcal{H}$, our strategy is to make use of $K_{\mathcal{H}}$ in order to prove Theorem 2.2.

Suppose that there is an injective and continuous map for a Hilbert space $\mathcal{H}$ into the Fréchet space $\mathcal{O}(\mathcal{L})$. Then the point evaluation map

$$\mathcal{O}(\mathcal{L}) \supset \mathcal{H} \to \mathcal{L}_z \simeq \mathbb{C}, \quad f \mapsto f(z)$$

is continuous with respect to the Hilbert topology on $\mathcal{H}$.

Let $\{\varphi_\nu\}$ be an orthonormal basis of $\mathcal{H}$. We define

$$K_{\mathcal{H}}(x, y) \equiv K(x, y) := \sum_\nu \varphi_\nu(x)\overline{\varphi_\nu(y)} \in \mathcal{O}(\mathcal{L})\widehat{\otimes}\overline{\mathcal{O}}(\overline{\mathcal{L}}) .$$

Then, $K(x, y)$ is well defined as a holomorphic section of $\mathcal{L} \to D$ for the first variable, and as an anti-holomorphic section of $\overline{\mathcal{L}} \to D$ for the second variable. The definition is independent of the choice of an orthonormal basis $\{\varphi_\nu\}$. $K(x, y)$ is called the reproducing kernel of $\mathcal{H}$.

Lemma 2.4. 1) *For each $y \in D$, $K(\cdot, y) \in \mathcal{H} \otimes \overline{\mathcal{L}_y}$ ($\simeq \mathcal{H}$) and $(f(\cdot), K(\cdot, y))_{\mathcal{H}} = f(y)$ for any $f \in \mathcal{H}$.*
2) Let $K_i(x, y)$ be the reproducing kernels of Hilbert spaces $\mathcal{H}_i \subset \mathcal{O}(\mathcal{L})$ with inner products $(\ ,\ )_{\mathcal{H}_i}$, respectively, for $i = 1, 2$. If $K_1 \equiv K_2$, then $\mathcal{H}_1 = \mathcal{H}_2$ and $(\ ,\ )_{\mathcal{H}_1} = (\ ,\ )_{\mathcal{H}_2}$.
3) If $K_1(x, x) = K_2(x, x)$ for any $x \in D$, then $K_1 \equiv K_2$.

Proof. (1) and (2) are standard. We review only how to recover $\mathcal{H}$ together with its inner product from a given reproducing kernel. For each $y \in D$, we fix an isomorphism $\mathcal{L}_y \simeq \mathbb{C}$. Through this isomorphism, we can regard $K(\cdot, y) \in \mathcal{H} \otimes \overline{\mathcal{L}_y}$ as an element of $\mathcal{H}$. The Hilbert space $\mathcal{H}$ is the completion of the $\mathbb{C}$-span of $\{K(\cdot, y) : y \in D\}$ with pre-Hilbert structure

$$(K(\cdot, y_1), K(\cdot, y_2))_{\mathcal{H}} := K(y_2, y_1) \in \mathcal{L}_{y_2} \otimes \overline{\mathcal{L}_{y_1}} \ (\simeq \mathbb{C}) . \tag{2.4.1}$$

This procedure is independent of the choice of the isomorphism $\mathcal{L}_y \simeq \mathbb{C}$. Hence, the Hilbert space $\mathcal{H}$ together with its inner product is recovered.
3) We denote by $\overline{D}$ the complex manifold endowed with the conjugate complex structure on D. Then, $\overline{\mathcal{L}} \to \overline{D}$ is a holomorphic line bundle, and $K(\cdot, \cdot) \equiv K_{\mathcal{H}}(\cdot, \cdot)$ is a holomorphic section of the holomorphic line bundle $\mathcal{L} \boxtimes \overline{\mathcal{L}} \to D \times \overline{D}$. As the diagonal embedding $\iota : D \to D \times \overline{D}, z \mapsto (z, z)$ is totally real, $(K_1 - K_2)|_{\iota(D)} \equiv 0$ implies $K_1 - K_2 \equiv 0$ by the unicity theorem of holomorphic functions. □

2.5 Construction of *J*

Suppose we are in the setting of Theorem 2.2. We define an anti-linear map

$$J : \mathcal{O}(\mathcal{L}) \to \mathcal{O}(\mathcal{L}) , \quad f \mapsto Jf$$

by $Jf(z) := \overline{f(\sigma(z))}$ $(z \in D)$. Jf is regarded as an element of $\mathcal{O}(\mathcal{L})$ through the isomorphism $\Psi_* : \mathcal{O}(\mathcal{L}) \simeq \mathcal{O}(\overline{\sigma^* \mathcal{L}})$ (see (2.2.2)).

Lemma 2.5. *In the setting of Theorem 2.2, we identify $\mathcal{H}$ with a subspace of $\mathcal{O}(\mathcal{L})$. Then, the anti-linear map J is an isometry from $\mathcal{H}$ onto $\mathcal{H}$.*

Proof. We put $\widetilde{\mathcal{H}} := J(\mathcal{H})$, equipped with the inner product

$$(Jf_1, Jf_2)_{\widetilde{\mathcal{H}}} := (f_2, f_1)_{\mathcal{H}} \quad \text{for } f_1, f_2 \in \mathcal{H} . \tag{2.5.1}$$

If $\{\varphi_\nu\}$ is an orthonormal basis of $\mathcal{H}$, then $\widetilde{\mathcal{H}}$ is a Hilbert space with orthonormal basis $\{J\varphi_\nu\}$. Hence, the reproducing kernel of $\widetilde{\mathcal{H}}$ is given by $K_{\widetilde{\mathcal{H}}}(x, y) = K_{\mathcal{H}}(\sigma(y), \sigma(x))$ because

$$K_{\widetilde{\mathcal{H}}}(x, y) = \sum_\nu J\varphi_\nu(x)\overline{J\varphi_\nu(y)} = \sum_\nu \overline{\varphi_\nu(\sigma(x))}\, \overline{\overline{\varphi_\nu(\sigma(y))}} = K_{\mathcal{H}}(\sigma(y), \sigma(x)) . \tag{2.5.2}$$

We fix $x \in D$ and take $g \in H$ such that $\sigma(x) = g \cdot x$ (see (2.2.3)). Substituting x for y in (2.5.2), we have

$$K_{\widetilde{\mathcal{H}}}(x, x) = K_{\mathcal{H}}(\sigma(x), \sigma(x)) = K_{\mathcal{H}}(g \cdot x, g \cdot x) = K_{\mathcal{H}}(x, x) .$$

Here, the last equality holds because $\{\varphi_\nu(g \cdot)\}$ is also an orthonormal basis of $\mathcal{H}$ as $(\pi, \mathcal{H})$ is a unitary representation of H. Then, by Lemma 2.4, the Hilbert space $\widetilde{\mathcal{H}}$ coincides with $\mathcal{H}$ and

$$(Jf_1, Jf_2)_{\mathcal{H}} = (f_2, f_1)_{\mathcal{H}} \quad \text{for } f_1, f_2 \in \mathcal{H} . \tag{2.5.3}$$

This is what we wanted to prove. □

2.6 Proof of $A^* = JAJ^{-1}$

Lemma 2.6 (see [15]). *Suppose* $A \in \mathrm{End}_H(\mathcal{H})$. *Then the adjoint operator* A^* *of* A *is given by*

$$A^* = JAJ^{-1}. \tag{2.6.1}$$

Proof. We divide the proof into three steps.

Step 1 *(positive self-adjoint case).* Assume $A \in \mathrm{End}_H(\mathcal{H})$ is a positive self-adjoint operator. Let $\mathcal{H}_A$ be the Hilbert completion of $\mathcal{H}$ by the pre-Hilbert structure

$$(f_1, f_2)_{\mathcal{H}_A} := (Af_1, f_2)_{\mathcal{H}} \quad \text{for } f_1, f_2 \in \mathcal{H}. \tag{2.6.2}$$

If $f_1, f_2 \in \mathcal{H}$ and $g \in H$, then

$$\begin{aligned}(\pi(g)f_1, \pi(g)f_2)_{\mathcal{H}_A} &= (A\pi(g)f_1, \pi(g)f_2)_{\mathcal{H}} \\ &= (\pi(g)Af_1, \pi(g)f_2)_{\mathcal{H}} = (Af_1, f_2)_{\mathcal{H}} = (f_1, f_2)_{\mathcal{H}_A}.\end{aligned}$$

Therefore, $(\pi, \mathcal{H})$ extends to a unitary representation on $\mathcal{H}_A$. Applying (2.5.3) to both $\mathcal{H}_A$ and $\mathcal{H}$, we have

$$\begin{aligned}(Af_1, f_2)_{\mathcal{H}} &= (f_1, f_2)_{\mathcal{H}_A} = (Jf_2, Jf_1)_{\mathcal{H}_A} = (AJf_2, Jf_1)_{\mathcal{H}} \\ &= (Jf_2, A^*Jf_1)_{\mathcal{H}} = (Jf_2, JJ^{-1}A^*Jf_1)_{\mathcal{H}} = (J^{-1}A^*Jf_1, f_2)_{\mathcal{H}}.\end{aligned}$$

Hence, $A = J^{-1}A^*J$, and (2.6.1) follows.

Step 2 *(self-adjoint case).* Assume $A \in \mathrm{End}_H(\mathcal{H})$ is a self-adjoint operator. Let $A = \int \lambda dE_\lambda$ be the spectral decomposition of A. Then every projection operator $E_\lambda \in \mathrm{End}(\mathcal{H})$ also commutes with $\pi(g)$ for all $g \in H$, namely, $E_\lambda \in \mathrm{End}_H(\mathcal{H})$. We define

$$A_+ := \int_{\lambda \geq 0} \lambda dE_\lambda, \qquad A_- := \int_{\lambda < 0} \lambda dE_\lambda.$$

Then $A = A_+ + A_-$. Let I be the identity operator on $\mathcal{H}$. As a positive self-adjoint operator $A_+ + I$ is an element of $\mathrm{End}_H(\mathcal{H})$, we have $(A_+ + I)^* = J(A_+ + I)J^{-1}$ by Step 1, whence $A_+^* = JA_+J^{-1}$. Applying Step 1 again to $-A_-$, we have $A_-^* = JA_-J^{-1}$. Thus,

$$A^* = A_+^* + A_-^* = JA_+J^{-1} + JA_-J^{-1} = J(A_+ + A_-)J^{-1} = JAJ^{-1}.$$

Step 3 *(general case).* Suppose $A \in \mathrm{End}_H(\mathcal{H})$. Then A^* also commutes with $\pi(g)$ $(g \in H)$ because π is unitary. We put $B := \frac{1}{2}(A + A^*)$ and $C := \frac{\sqrt{-1}}{2}(A^* - A)$. Then, both B and C are self-adjoint operators commuting with $\pi(g)$ $(g \in H)$. It follows from Step 2 that $B^* = JBJ^{-1}$ and $C^* = JCJ^{-1}$. As J is an anti-linear map, we have

$$(\sqrt{-1}\,C)^* = -\sqrt{-1}\,C^* = -\sqrt{-1}\,JCJ^{-1} = J(\sqrt{-1}\,C)J^{-1}.$$

Hence, $A = B + \sqrt{-1}\,C$ also satisfies $A^* = JAJ^{-1}$. □

2.7 Proof of Theorem 2.2

We are now ready to complete the proof of Theorem 2.2. Let $A, B \in \mathrm{End}_H(\mathcal{H})$. By Lemma 2.6, we have

$$AB = J^{-1}(AB)^*J = (J^{-1}B^*J)(J^{-1}A^*J) = BA\,.$$

Therefore, $\mathrm{End}_H(\mathcal{H})$ is commutative. □

3 Proof of Theorem A

This section gives a proof of Theorem A by using Theorem 2.2. The core of the proof is to reduce the geometric condition (2.2.3) to an algebraic condition (the existence of a certain involution of the Lie algebra). This reduction is stated in Lemma 3.3. The reader who is familiar with symmetric pairs can skip Subsections 3.1, 3.2, 3.4 and 3.5.

3.1 Reductive symmetric pairs

Let G be a Lie group. Suppose that τ is an involutive automorphism of G. We write

$$G^\tau := \{g \in G : \tau g = g\}$$

for the fixed point subgroup of τ, and denote by G^τ_0 its connected component containing the unit element. The pair (G, H) (or the pair $(\mathfrak{g}, \mathfrak{h})$ of their Lie algebras) is called a *symmetric pair* if the subgroup H is an open subgroup of G^τ, that is, if H satisfies

$$G^\tau_0 \subset H \subset G^\tau.$$

It is called a *reductive symmetric pair* if G is a reductive Lie group; a *semisimple symmetric pair* if G is a semisimple Lie group. Obviously, a semisimple symmetric pair is a reductive symmetric pair.

We shall use the same letter τ to denote the differential of τ. We set

$$\mathfrak{g}^{\pm\tau} := \{Y \in \mathfrak{g} : \tau Y = \pm Y\}\,.$$

Then, it follows from $\tau^2 = \mathrm{id}$ that we have a direct sum decomposition

$$\mathfrak{g} = \mathfrak{g}^\tau \oplus \mathfrak{g}^{-\tau}.$$

Suppose now that G is a semisimple Lie group. It is known that there exists a Cartan involution θ of G commuting with τ. Take such θ, and we write $K := G^\theta = \{g \in G : \theta g = g\}$. Then, K is compact if G is a linear Lie group. The direct sum decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p} \equiv \mathfrak{g}^\theta \oplus \mathfrak{g}^{-\theta}$$

is called a Cartan decomposition. Later, we shall allow G to be non-linear, in particular, K is not necessarily compact. The *real rank* of $\mathfrak{g}$, denoted by $\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}$, is defined to be the dimension of a maximal abelian subspace of $\mathfrak{g}^{-\theta}$.

As $(\tau\theta)^2 = \mathrm{id}$, the pair $(\mathfrak{g}, \mathfrak{g}^{\tau\theta})$ also forms a symmetric pair. The Lie group

$$G^{\tau\theta} = \{g \in G : (\tau\theta)(g) = g\}$$

is a reductive Lie group with Cartan involution $\theta|_{G^{\tau\theta}}$, and its Lie algebra $\mathfrak{g}^{\tau\theta}$ is reductive with Cartan decomposition

$$\mathfrak{g}^{\tau\theta} = \mathfrak{g}^{\tau\theta,\theta} \oplus \mathfrak{g}^{\tau\theta,-\theta} = \mathfrak{g}^{\tau,\theta} \oplus \mathfrak{g}^{-\tau,-\theta}. \tag{3.1.1}$$

Here, we have used the notation $\mathfrak{g}^{-\tau,-\theta}$ and alike, defined as follows:

$$\mathfrak{g}^{-\tau,-\theta} := \{Y \in \mathfrak{g} : (-\tau)Y = (-\theta)Y = Y\}.$$

Then the dimension of a maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{g}^{-\tau,-\theta}$ is equal to the real rank of $\mathfrak{g}^{\tau\theta}$, which is referred to as the *split rank* of the semisimple symmetric space G/H. We shall write $\mathbb{R}\text{-}\operatorname{rank} G/H$ or $\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}/\mathfrak{g}^{\tau}$ for this dimension. Thus,

$$\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\theta\tau} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}/\mathfrak{g}^{\tau}. \tag{3.1.2}$$

In particular, we have $\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}/\mathfrak{k}$ if we take τ to be θ.

The Killing form on the Lie algebra $\mathfrak{g}$ is non-degenerate on $\mathfrak{g}$, and is also non-degenerate when restricted to $\mathfrak{h}$. Then, it induces an $\mathrm{Ad}(H)$-invariant non-degenerate bilinear form on $\mathfrak{g}/\mathfrak{h}$, and therefore a G-invariant pseudo-Riemannian structure on the homogeneous space G/H, so that G/H becomes a symmetric space with respect to the Levi-Cività connection and is called a *semisimple symmetric space*. In this context, the subspace $\mathfrak{a}$ has the following geometric meaning: Let $A := \exp(\mathfrak{a})$, the connected abelian subgroup of G with Lie algebra $\mathfrak{a}$. Then, the orbit $A \cdot o$ through $o := eH \in G/H$ becomes a flat, totally geodesic submanifold in G/H. Furthermore, we have a (generalized) Cartan decomposition:

Fact 3.1 *(see [16, Section 2]).* $G = KAH$.

Sketch of Proof. The direct sum decomposition of the Lie algebra

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{g}^{-\tau,-\theta} \oplus \mathfrak{g}^{\tau,-\theta}$$

lifts to a diffeomorphism:

$$\mathfrak{g}^{-\tau,-\theta} + \mathfrak{g}^{\tau,-\theta} \xrightarrow{\sim} K\backslash G, \quad (X, Y) \mapsto Ke^X e^Y.$$

Since $\exp(\mathfrak{g}^{\tau,-\theta}) \subset H$, the decomposition $G = KAH$ follows if we show

$$\mathrm{Ad}(H \cap K)\mathfrak{a} = \mathfrak{g}^{-\tau,-\theta}. \tag{3.1.3}$$

The equation (3.1.3) is well known as the key ingredient of the original Cartan decomposition $G^{\tau\theta} = K^{\tau} A K^{\tau}$ in light of (3.1.1). □

Furthermore, suppose that σ is an involutive automorphism of G such that σ, τ and θ commute with one another. We set

$$G^{\sigma,\tau} := G^\sigma \cap G^\tau = \{g \in G : \sigma g = \tau g = g\}.$$

Then $(G^\sigma, G^{\sigma,\tau})$ forms a reductive symmetric pair, because σ and τ commute. The commutativity of σ and θ implies that the automorphism $\sigma : G \to G$ stabilizes K and induces a diffeomorphism of G/K, for which we use the same letter σ.

3.2 Examples of symmetric pairs

This subsection presents some basic examples of semisimple (and therefore, reductive) symmetric pairs.

Example 3.2.1 *(group manifold).* Let G' be a semisimple Lie group, and $G := G' \times G'$. We define an involutive automorphism τ of G by $\tau(x, y) := (y, x)$. Then, $G^\tau = \{(g, g) : g \in G'\}$ is the diagonal subgroup, denoted by $\operatorname{diag}(G')$, which is isomorphic to G'. Thus, $(G' \times G', \operatorname{diag}(G'))$ forms a semisimple symmetric pair.

We set

$$I_{p,q} := \begin{pmatrix} \overbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}^{p} & 0 \\ 0 & \overbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}^{q} \end{pmatrix}$$

$$J := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

Example 3.2.2. Let $G = SL(n, \mathbb{C})$, and fix p, q such that $p + q = n$. Then,

$$\tau(g) := I_{p,q}\, g^* I_{p,q} \quad (g \in G)$$

defines an involutive automorphism of G, and $G^\tau = SU(p, q)$ (the indefinite unitary group). Thus, $(SL(n, \mathbb{C}), SU(p, q))$ forms a semisimple symmetric pair.

Example 3.2.3. Let $G = SL(n, \mathbb{C})$, and $\sigma(g) := \overline{g}$. Then σ is an involutive automorphism of G, and $G^\sigma = SL(n, \mathbb{R})$. We note that σ commutes with the involution τ in the previous example, and

$$\begin{aligned} G^{\sigma,\tau} &= \{g \in SL(n, \mathbb{C}) : \overline{g} = g = I_{p,q}\, {}^t\overline{g}\, I_{p,q}\} \\ &= SO(p, q). \end{aligned}$$

Thus, $(SL(n, \mathbb{C}), SL(n, \mathbb{R}))$, $(SU(p, q), SO(p, q))$, $(SL(n, \mathbb{R}), SO(p, q))$ are examples of semisimple symmetric pairs.

Example 3.2.4. Let $G := SL(2n, \mathbb{R})$, and $\tau(g) := J\, {}^t g^{-1} J^{-1}$. Then, $G^\tau = Sp(n, \mathbb{R})$ (the real symplectic group). Thus, $(SL(2n, \mathbb{R}), Sp(n, \mathbb{R}))$ forms a semisimple symmetric pair.

3.3 Reduction of visibility to the real rank condition

The following lemma gives a sufficient condition for (2.2.3). Then, it plays a key role when we apply Theorem 2.2 to the branching problem for the restriction from G to G^τ (with the notation of Theorem 2.2, $D = G/K$ and $H = G_0^\tau$). This lemma is also used in reducing 'visibility' of an action to an algebraic condition ([50, Lemma 2.2]).

Lemma 3.3. *Let σ and τ be involutive automorphisms of G. We assume that the pair (σ, τ) satisfies the following two conditions:*

$$\sigma,\ \tau \text{ and } \theta \text{ commute with one another.} \tag{3.3.1}$$

$$\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\tau\theta} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\sigma,\tau\theta}. \tag{3.3.2}$$

Then for any $x \in G/K$, there exists $g \in G_0^\tau$ such that $\sigma(x) = g \cdot x$.

Proof. It follows from the condition (3.3.1) that $\theta|_{G^\sigma}$ is a Cartan involution of a reductive Lie group G^σ and that $\tau|_{G^\sigma}$ is an involutive automorphism of G^σ commuting with $\theta|_{G^\sigma}$. Take a maximal abelian subspace $\mathfrak{a}$ in

$$\mathfrak{g}^{-\theta,\sigma,\tau\theta} := \{Y \in \mathfrak{g} : (-\theta)Y = \sigma Y = \tau\theta Y = Y\}.$$

From the definition, we have $\dim\mathfrak{a} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\sigma,\tau\theta}$, which in turn equals $\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\tau\theta}$ by the condition (3.3.2). This means that $\mathfrak{a}$ is also a maximal abelian subspace in

$$\mathfrak{g}^{-\theta,\tau\theta} = \{Y \in \mathfrak{g} : (-\theta)Y = \tau\theta Y = Y\}.$$

Let $A = \exp(\mathfrak{a})$. Then it follows from Fact 3.1 that we have a generalized Cartan decomposition

$$G = G_0^\tau A K. \tag{3.3.3}$$

Let $o := eK \in G/K$. Fix $x \in G/K$. Then, according to the decomposition (3.3.3), we find $h \in G_0^\tau$ and $a \in A$ such that

$$x = ha \cdot o.$$

We set $g := \sigma(h)\,h^{-1}$. We claim $g \in G_0^\tau$. In fact, by using $\sigma\tau = \tau\sigma$ and $\tau h = h$, we have

$$\tau(g) = \tau\sigma(h)\,\tau(h^{-1}) = \sigma\tau(h)\,\tau(h)^{-1} = \sigma(h)\,h^{-1} = g.$$

Hence, $g \in G^\tau$. Moreover, since the image of the continuous map

$$G_0^\tau \to G, \quad h \mapsto \sigma(h)\,h^{-1}$$

is connected, we have $g \in G_0^\tau$.

On the other hand, we have $\sigma(a) = a$ because $\mathfrak{a} \subset \mathfrak{g}^{-\theta,\sigma,-\tau} \subset \mathfrak{g}^\sigma$. Therefore we have

$$\sigma(x) = \sigma(h)\,\sigma(a) \cdot o = \sigma(h)\,h^{-1}ha \cdot o = g \cdot x,$$

proving the lemma. □

3.4 Hermitian symmetric space G/K

Throughout the rest of this section, we assume that G is a simple, non-compact, Lie group of Hermitian type. We retain the notation of Subsection 1.3.

Let $G_{\mathbb{C}}$ be a connected complex Lie group with Lie algebra $\mathfrak{g}_{\mathbb{C}}$, and Q^- the maximal parabolic subgroup of $G_{\mathbb{C}}$ with Lie algebra $\mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_-$. Then we have an open embedding $G/K \hookrightarrow G_{\mathbb{C}}/Q^-$ because $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} + (\mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_-)$. Thus, a G-invariant complex structure on G/K is induced from $G_{\mathbb{C}}/Q^-$. (We remark that the embedding $G/K \hookrightarrow G_{\mathbb{C}}/Q^-$ is well defined, even though G is not necessarily a subgroup of $G_{\mathbb{C}}$.)

Suppose τ is an involutive automorphism of G commuting with θ. We recall from Subsection 1.4 that we have either

$$\tau Z = Z \qquad \text{(holomorphic type)}, \tag{1.4.1}$$

or

$$\tau Z = -Z \qquad \text{(anti-holomorphic type)}. \tag{1.4.2}$$

Here is the classification of semisimple symmetric pairs $(\mathfrak{g}, \mathfrak{g}^\tau)$ with $\mathfrak{g}$ simple such that the pair $(\mathfrak{g}, \mathfrak{g}^\tau)$ satisfies the condition (1.4.1) (respectively, (1.4.2)). Table 3.4.2

Table 3.4.1.

$(\mathfrak{g}, \mathfrak{g}^\tau)$ is of holomorphic type	
$\mathfrak{g}$	$\mathfrak{g}^\tau$
$\mathfrak{su}(p, q)$	$\mathfrak{s}(\mathfrak{u}(i, j) + \mathfrak{u}(p-i, q-j))$
$\mathfrak{su}(n, n)$	$\mathfrak{so}^*(2n)$
$\mathfrak{su}(n, n)$	$\mathfrak{sp}(n, \mathbb{R})$
$\mathfrak{so}^*(2n)$	$\mathfrak{so}^*(2p) + \mathfrak{so}^*(2n-2p)$
$\mathfrak{so}^*(2n)$	$\mathfrak{u}(p, n-p)$
$\mathfrak{so}(2, n)$	$\mathfrak{so}(2, p) + \mathfrak{so}(n-p)$
$\mathfrak{so}(2, 2n)$	$\mathfrak{u}(1, n)$
$\mathfrak{sp}(n, \mathbb{R})$	$\mathfrak{u}(p, n-p)$
$\mathfrak{sp}(n, \mathbb{R})$	$\mathfrak{sp}(p, \mathbb{R}) + \mathfrak{sp}(n-p, \mathbb{R})$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{so}(10) + \mathfrak{so}(2)$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{so}^*(10) + \mathfrak{so}(2)$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{so}(8, 2) + \mathfrak{so}(2)$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{su}(5, 1) + \mathfrak{sl}(2, \mathbb{R})$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{su}(4, 2) + \mathfrak{su}(2)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{e}_{6(-78)} + \mathfrak{so}(2)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{e}_{6(-14)} + \mathfrak{so}(2)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{so}(10, 2) + \mathfrak{sl}(2, \mathbb{R})$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{so}^*(12) + \mathfrak{su}(2)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{su}(6, 2)$

Table 3.4.2.

$(\mathfrak{g}, \mathfrak{g}^\tau)$ is of anti-holomorphic type	
$\mathfrak{g}$	$\mathfrak{g}^\tau$
$\mathfrak{su}(p,q)$	$\mathfrak{so}(p,q)$
$\mathfrak{su}(n,n)$	$\mathfrak{sl}(n,\mathbb{C}) + \mathbb{R}$
$\mathfrak{su}(2p,2q)$	$\mathfrak{sp}(p,q)$
$\mathfrak{so}^*(2n)$	$\mathfrak{so}(n,\mathbb{C})$
$\mathfrak{so}^*(4n)$	$\mathfrak{su}^*(2n) + \mathbb{R}$
$\mathfrak{so}(2,n)$	$\mathfrak{so}(1,p) + \mathfrak{so}(1,n-p)$
$\mathfrak{sp}(n,\mathbb{R})$	$\mathfrak{gl}(n,\mathbb{R})$
$\mathfrak{sp}(2n,\mathbb{R})$	$\mathfrak{sp}(n,\mathbb{C})$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{f}_{4(-20)}$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{sp}(2,2)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{e}_{6(-26)} + \mathfrak{so}(1,1)$
$\mathfrak{e}_{7(-25)}$	$\mathfrak{su}^*(8)$

is equivalent to the classification of totally real symmetric spaces G^τ/K^τ of the Hermitian symmetric space G/K (see [14, 27, 28, 35]).

3.5 Holomorphic realization of highest weight representations

It is well known that an irreducible highest weight representation π of G can be realized as a subrepresentation of the space of global holomorphic sections of an equivariant holomorphic vector bundle over the Hermitian symmetric space G/K. We supply a proof here for the convenience of the reader in a way that we shall use later.

Lemma 3.5. *Let $(\pi, \mathcal{H})$ be an irreducible unitary highest weight module. We write χ for the representation of K on $U := \mathcal{H}_K^{\mathfrak{p}_+}$ (see Definition 1.3). Let $\mathcal{L} := G \times_K U \to G/K$ be the G-equivariant holomorphic vector bundle associated to χ. Then, there is a natural injective continuous G-homomorphism $\mathcal{H} \to \mathcal{O}(\mathcal{L})$.*

Proof. Let $(\ ,\ )_{\mathcal{H}}$ be a G-invariant inner product on $\mathcal{H}$. We write $(\ ,\ )_U$ for the induced inner product on U. Then, K acts unitarily on $\mathcal{H}$, and in particular on U. We consider the map

$$G \times \mathcal{H} \times U \to \mathbb{C}, \quad (g, v, u) \mapsto (\pi(g)^{-1}v, u)_{\mathcal{H}} = (v, \pi(g)u)_{\mathcal{H}}.$$

For each fixed $g \in G$ and $v \in \mathcal{H}$, the map $U \to \mathbb{C}$, $u \mapsto (\pi(g)^{-1}v, u)_{\mathcal{H}}$ is an anti-linear functional on U. Then there exists a unique element $F_v(g) \in U$ by the Riesz representation theorem for the finite-dimensional Hilbert space U such that

$$(F_v(g), u)_U = (\pi(g)^{-1}v, u)_{\mathcal{H}} \quad \text{for any } u \in U.$$

Then it is readily seen that $F_v(gk) = \chi(k)^{-1} F_v(g)$ and $F_{\pi(g')v}(g) = F_v(g'^{-1}g)$ for any $g, g' \in G$, $k \in K$ and $v \in \mathcal{H}$. As u is a smooth vector in $\mathcal{H}$, $(F_v(g), u)_U = (v, \pi(g)u)_{\mathcal{H}}$ is a C^∞-function on G. Then $F_v(g)$ is a C^∞-function on G with value in U for each fixed $v \in \mathcal{H}$. Thus, we have a non-zero G-intertwining operator given by

$$F : \mathcal{H} \to C^\infty(G \times_K U), \quad v \mapsto F_v .$$

As U is annihilated by $\mathfrak{p}_+$, F_v is a holomorphic section of the holomorphic vector bundle $G \times_K U \to G/K$, that is, $F_v \in \mathcal{O}(G \times_K U)$. Then, the non-zero map $F : \mathcal{H} \to \mathcal{O}(G \times_K U)$ is injective because $\mathcal{H}$ is irreducible. Furthermore, F is continuous by the closed graph theorem. Hence Lemma 3.6 is proved. □

3.6 Reduction to the real rank condition

The next lemma is a stepping stone to Theorem A. It becomes also a key lemma to the theorem that the action of a subgroup H on the bounded symmetric domain G/K is 'strongly visible' for any symmetric pair (G, H) (see [50]).

Lemma 3.6. *Suppose $\mathfrak{g}$ is a real simple Lie algebra of Hermitian type. Let τ be an involutive automorphism of $\mathfrak{g}$, commuting with a fixed Cartan involution θ. Then there exists an involutive automorphism σ of $\mathfrak{g}$ satisfying the following three conditions:*

σ, τ and θ commute with one another. (3.6.1)

$\mathbb{R}\text{- rank}\,\mathfrak{g}^{\tau\theta} = \mathbb{R}\text{- rank}\,\mathfrak{g}^{\sigma,\tau\theta}.$ (3.6.2)

$\sigma Z = -Z.$ (3.6.3)

Proof. We shall give a proof in the special case $\tau = \theta$ in Subsection 4.1. For the general case, see [50, Lemma 3.1] or [44, Lemma 5.1]. □

3.7 Proof of Theorem A

Now, we are ready to complete the proof of Theorem A.

Without loss of generality, we may and do assume that G is simply connected. Let $(\pi, \mathcal{H})$ be an irreducible unitary highest weight representation of scalar type. We define a holomorphic line bundle by $\mathcal{L} := G \times_K \mathcal{H}_K^{\mathfrak{p}_+}$ over the Hermitian symmetric space $D := G/K$. Then it follows from Lemma 3.5 that there is an injective continuous G-intertwining map $\mathcal{H} \to \mathcal{O}(\mathcal{L})$.

Suppose (G, H) is a symmetric pair. We first note that for an involutive automorphism τ of G, there exists $g \in G$ such that $\tau^g\theta = \theta\tau^g$ if we set

$$\tau^g(x) := g\tau(g^{-1}xg)g^{-1}$$

for $x \in G$. Then $G^{\tau_g} = gHg^{-1}$ is θ-stable. Since the multiplicity-free property of the restriction $\pi|_H$ is unchanged if we replace H by gHg^{-1}, we may and do assume that $\theta H = H$, in other words, $\theta\tau = \tau\theta$.

Now, by applying Lemma 3.6, we can take σ satisfying (3.6.1), (3.6.2) and (3.6.3). We use the same letter σ to denote its lift to G. It follows from (3.6.3) that the induced involutive diffeomorphism $\sigma : G/K \to G/K$ is anti-holomorphic (see Subsection 1.4). In light of the conditions (3.6.1) and (3.6.2), we can apply Lemma 3.3 to see that for any $x \in D$ there exists $g \in H$ such that $\sigma(x) = g \cdot x$.

Moreover, by using Lemma 9.4 in the Appendix, we have an isomorphism $\overline{\sigma^*\mathcal{L}} \simeq \mathcal{L}$ as G-equivariant holomorphic line bundles over G/K. Therefore, all the assumptions of Theorem 2.2 are satisfied. Thus, we conclude that the restriction $\pi|_H$ is multiplicity-free by Theorem 2.2. □

4 Proof of Theorem C

In this section we give a proof of Theorem C.

Throughout this section, we may and do assume that G is simply connected so that any automorphism of $\mathfrak{g}$ lifts to G. We divide the proof of Theorem C into the following cases:
Case I. Both π_1 and π_2 are highest weight modules.
Case I′. Both π_1 and π_2 are lowest weight modules.
Case II. π_1 is a highest weight module, and π_2 is a lowest weight module.
Case II′. π_1 is a lowest weight module, and π_2 is a highest weight module.

4.1 Reduction to the real rank condition

The following lemma is a special case of Lemma 3.6 with $\tau = \theta$. We shall see that Theorem C in Case I (likewise, Case I′) reduces to this algebraic result.

Lemma 4.1.1. *Suppose $\mathfrak{g}$ is a real simple Lie algebra of Hermitian type. Let θ be a Cartan involution. Then there exists an involutive automorphism σ of $\mathfrak{g}$ satisfying the following three conditions:*

$$\sigma \text{ and } \theta \text{ commute.} \tag{4.1.1}$$

$$\mathbb{R}\text{-}\operatorname{rank}\mathfrak{g} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}^{\sigma}. \tag{4.1.2}$$

$$\sigma Z = -Z. \tag{4.1.3}$$

Proof. We give a proof of the lemma based on the classification of simple Lie algebras $\mathfrak{g}$ of Hermitian type.

We recall that for any involutive automorphism σ of G, there exists $g \in G$ such that $\sigma^g\theta = \theta\sigma^g$. Thus, (4.1.1) is always satisfied after replacing σ by some σ^g. The remaining conditions (4.1.2) and (4.1.3) (cf. Table 3.4.2) are satisfied if we choose $\sigma \in \operatorname{Aut}(G)$ in the following Table 4.1.2 for each simple non-compact Lie group G of Hermitian type:

Table 4.1.2.

$(\mathfrak{g}, \mathfrak{g}^\sigma)$ satisfying (4.1.2) and (4.1.3)		
$\mathfrak{g}$	$\mathfrak{g}^\sigma$	$\mathbb{R}$- rank $\mathfrak{g} = \mathbb{R}$- rank $\mathfrak{g}^\sigma$
$\mathfrak{su}(p,q)$	$\mathfrak{so}(p,q)$	$\min(p,q)$
$\mathfrak{so}^*(2n)$	$\mathfrak{so}(n,\mathbb{C})$	$[\frac{1}{2}n]$
$\mathfrak{sp}(n,\mathbb{R})$	$\mathfrak{gl}(n,\mathbb{R})$	n
$\mathfrak{so}(2,n)$	$\mathfrak{so}(1,n-1)+\mathfrak{so}(1,1)$	$\min(2,n)$
$\mathfrak{e}_{6(-14)}$	$\mathfrak{sp}(2,2)$	2
$\mathfrak{e}_{7(-25)}$	$\mathfrak{su}^*(8)$	3

Here, we have proved lemma. □

Remark 4.1.3. The choice of σ in Lemma 4.1.1 is not unique. For example, we may choose $\mathfrak{g}^\sigma \simeq \mathfrak{e}_{6(-26)} \oplus \mathbb{R}$ instead of the above choice $\mathfrak{g}^\sigma \simeq \mathfrak{su}^*(8)$ for $\mathfrak{g} = \mathfrak{e}_{7(-25)}$.

4.2 Proof of Theorem C in Case I

Let G be a non-compact simply-connected, simple Lie group such that G/K is a Hermitian symmetric space.

Let $(\pi_1, \mathcal{H}_1)$ and $(\pi_2, \mathcal{H}_2)$ be two irreducible unitary highest weight representations of scalar type. By Lemma 3.5, we can realize $(\pi_i, \mathcal{H}_i)$ in the space $\mathcal{O}(\mathcal{L}_i)$ of holomorphic sections of the holomorphic line bundle $\mathcal{L}_i := G \times_K (\mathcal{H}_i)_K^{\mathfrak{p}_+}$ $(i = 1, 2)$ over the Hermitian symmetric space G/K. We now define a holomorphic line bundle $\mathcal{L} := \mathcal{L}_1 \boxtimes \mathcal{L}_2$ over $D := G/K \times G/K$ as the outer tensor product of $\mathcal{L}_1$ and $\mathcal{L}_2$. Then we have naturally an injective continuous $(G \times G)$-intertwining map $\mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2 \to \mathcal{O}(\mathcal{L})$.

Let us take an involution σ' of $\mathfrak{g}$ as in Lemma 4.1.1 (but we use the letter σ' instead of σ), and lift it to G. We set $\sigma := \sigma' \times \sigma'$. Then it follows from (4.1.3) that σ' acts anti-holomorphically on G/K, and so does σ on D. Furthermore, we have isomorphisms of holomorphic line bundles $\overline{(\sigma')^* \mathcal{L}_i} \simeq \mathcal{L}_i$ $(i = 1, 2)$ by Lemma 9.4 and thus $\overline{\sigma^* \mathcal{L}} \simeq \mathcal{L}$.

We now introduce another involutive automorphism τ of $G \times G$ by $\tau(g_1, g_2) := (g_2, g_1)$. Then $(G \times G)^\tau = \operatorname{diag}(G) := \{(g, g) : g \in G\}$. We shall use the same letter θ to denote the Cartan involution $\theta \times \theta$ on $G \times G$ (and $\theta \oplus \theta$ on $\mathfrak{g} \oplus \mathfrak{g}$). Then we observe the following isomorphisms:

$$(\mathfrak{g} \oplus \mathfrak{g})^{\tau\theta} = \{(X, \theta X) : X \in \mathfrak{g}\} \qquad \simeq \mathfrak{g}\,,$$

$$(\mathfrak{g} \oplus \mathfrak{g})^{\sigma,\tau\theta} = \{(X, \theta X) : X \in \mathfrak{g}^{\sigma'}\} \simeq \mathfrak{g}^{\sigma'}.$$

Thus, the condition (4.1.2) implies

$$\mathbb{R}\text{- rank}(\mathfrak{g} \oplus \mathfrak{g})^{\tau\theta} = \mathbb{R}\text{- rank}(\mathfrak{g} \oplus \mathfrak{g})^{\sigma,\tau\theta}\,.$$

Therefore, given $(x_1, x_2) \in D \simeq (G \times G)/(K \times K)$, there exists $(g, g) \in (G \times G)^\tau$ satisfying $(g \cdot x_1, g \cdot x_2) = (\sigma'(x_1), \sigma'(x_2))$ $(= \sigma(x_1, x_2))$ by Lemma 3.3.

Let us apply Theorem 2.2 to the setting $(\mathcal{L} \to D, \mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2, \mathrm{diag}(G), \sigma)$. Now that all the assumptions of Theorem 2.2 are satisfied, we conclude that the tensor product $\pi_1 \widehat{\otimes} \pi_2$ is multiplicity-free as a G-module, that is, Theorem C holds in the case I. □

4.3 Proof of Theorem C in Case II

Let us give a proof of Theorem C in the case II. We use the same τ as in Subsection 4.2, that is, $\tau(g_1, g_2) := (g_2, g_1)$ and define a new involution σ by $\sigma := \tau\theta$, that is, $\sigma(g_1, g_2) = (\theta g_2, \theta g_1)$ for $g_1, g_2 \in G$. Obviously, σ, τ and the Cartan involution θ of $G \times G$ all commute.

We write M for the Hermitian symmetric space G/K, and $\overline{M}$ for the conjugate complex manifold. Then σ acts anti-holomorphically on $D := M \times \overline{M}$ because so does τ and because θ acts holomorphically.

By the obvious identity $(\mathfrak{g} \oplus \mathfrak{g})^{\tau\theta} = (\mathfrak{g} \oplus \mathfrak{g})^{\sigma,\tau\theta}$, we have $\mathbb{R}\text{-}\operatorname{rank}(\mathfrak{g} \oplus \mathfrak{g})^{\tau\theta} = \mathbb{R}\text{-}\operatorname{rank}(\mathfrak{g} \oplus \mathfrak{g})^{\sigma,\tau\theta}$ $(= \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g})$. Therefore, it follows from Lemma 3.3 that for any $(x_1, x_2) \in D$ there exists $(g, g) \in (G \times G)^\tau$ such that $\sigma(x_1, x_2) = (g, g) \cdot (x_1, x_2)$.

Suppose π_1 (respectively, π_2) is a unitary highest (respectively, lowest) weight representation of scalar type. We set $\mathcal{L}_1 := G \times_K (\mathcal{H}_1)_K^{\mathfrak{p}_+}$ and $\mathcal{L}_2 := G \times_K (\mathcal{H}_2)_K^{\mathfrak{p}_-}$. Then, $\mathcal{L}_1 \to M$ and $\mathcal{L}_2 \to \overline{M}$ are both holomorphic line bundles, and we can realize π_1 in $\mathcal{O}(M, \mathcal{L}_1)$, and π_2 in $\mathcal{O}(\overline{M}, \mathcal{L}_2)$, respectively. Therefore, the outer tensor product $\pi_1 \boxtimes \pi_2$ is realized in a subspace of holomorphic sections of the holomorphic line bundle $\mathcal{L} := \mathcal{L}_1 \boxtimes \mathcal{L}_2$ over $D = M \times \overline{M}$.

Now, we apply Theorem 2.2 to $(\mathcal{L} \to D, \mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2, \mathrm{diag}(G), \sigma)$. The condition (2.2.2) holds by Lemma 9.4. Hence, all the assumptions of Theorem 2.2 are satisfied, and therefore, Theorem C holds in the case II. □

Hence, Theorem C has been proved.

5 Uniformly bounded multiplicities — Proof of Theorems B and D

This section gives the proof of Theorems B and D. Since the proof of Theorem B parallels that of Theorem D, we deal mostly with Theorem D here. Without loss of generality, we assume G is a non-compact simple Lie group of Hermitian type.

5.1 General theory of restriction

A unitary representation $(\pi, \mathcal{H})$ of a group L is *discretely decomposable* if π is unitarily equivalent to the discrete Hilbert sum of irreducible unitary representations of L:

$$\pi \simeq \sum_{\mu \in \widehat{L}}^{\oplus} m_\pi(\mu)\mu \,.$$

Furthermore, we say π is *L-admissible* ([38]) if all the multiplicities $m_\pi(\mu)$ are finite. In this definition, we do not require $m_\pi(\mu)$ to be uniformly bounded with respect to μ.

Suppose L' is a subgroup of L. Then, the restriction of π to L' is regarded as a unitary representation of L'. If π is L'-admissible, then π is L-admissible ([38, Theorem 1.2]).

We start by recalling from [42] a discrete decomposability theorem of branching laws in the following settings:

Fact 5.1. 1) Suppose τ is of holomorphic type (see Definition 1.4) and set $H := G_0^\tau$. If π is an irreducible unitary highest weight representation of G, then π is $(H \cap K)$-admissible. In particular, π is H-admissible. The restriction $\pi|_H$ splits into a discrete Hilbert sum of irreducible unitary highest weight representations of H:

$$\pi|_H \simeq \sum_{\mu \in \widehat{H}}^{\oplus} m_\pi(\mu)\mu \qquad \text{(discrete Hilbert sum)}, \tag{5.1.1}$$

where the multiplicity $m_\pi(\mu)$ is finite for every μ.
2) Let π_1, π_2 be two irreducible unitary highest weight representations of G.Then the tensor product $\pi_1 \widehat{\otimes} \pi_2$ is K-admissible under the diagonal action. Furthermore, $\pi_1 \widehat{\otimes} \pi_2$ splits into a discrete Hilbert sum of irreducible unitary highest weight representations of G, each occurring with finite multiplicity. Furthermore, if at least one of π_1 or π_2 is a holomorphic discrete series representation for G, then any irreducible summand is a holomorphic discrete series representation.

Proof. See [42, Theorem 7.4] for the proof. The main idea of the proof is taking normal derivatives of holomorphic sections, which goes back to S. Martens [63]. The same idea was also employed in a number of papers including Lipsman ([60, Theorem 4.2]) and Jakobsen–Vergne ([31, Corollary 2.3]). □

Remark 5.1. Fact 5.1 (1) holds more generally for a closed subgroup H satisfying the following two conditions:

1) H is θ-stable.
2) The Lie algebra $\mathfrak{h}$ of H contains Z.

Here, we recall that Z is the generator of the center of $\mathfrak{k}$. The proof is essentially the same as that of Fact 5.1 (1).

Theorem B (2) follows from Theorem A and Fact 5.1 (1). Likewise, Theorem D (2) follows from Theorem C and Fact 5.1 (2). What remains to show for Theorems B and D is the uniform boundedness of multiplicities.

5.2 Remarks on Fact 5.1

Some remarks about Fact 5.1 are in order.

Remark 5.2.1. A Cartan involution θ is clearly of holomorphic type because $\theta Z = Z$. If $\theta = \tau$, then $H = K$ and any irreducible summand μ is finite dimensional. In this case, the finiteness of $m_\pi(\mu)$ in Fact 5.1 (1) is a special case of Harish-Chandra's admissibility theorem (this holds for any irreducible unitary representation π of G).

Remark 5.2.2. Fact 5.1 asserts in particular that there is no continuous spectrum in the irreducible decomposition formula. The crucial assumption for this is that (G, H) is of holomorphic type. In contrast, the restriction $\pi|_H$ is not discretely decomposable if (G, H) is of anti-holomorphic type and if π is a holomorphic discrete series representation of G ([38, Theorem 5.3]). In this setting, R. Howe, J. Repka, G. Ólafsson, B. Ørsted, G. van Dijk, S. Hille, M. Pevzner, V. Molchanov, Y. Neretin, G. Zhang and others studied irreducible decompositions of the restriction $\pi|_H$ by means of the L^2-harmonic analysis on Riemannian symmetric spaces $H/H \cap K$ ([9, 10, 11, 23, 64, 66, 69, 70, 74]). The key idea in Howe and Repka [23, 74] is that a holomorphic function on G/K is uniquely determined by its restriction to the totally real submanifold $H/H \cap K$ (essentially, the unicity theorem of holomorphic functions), and that any function on $H/H \cap K$ can be approximated (in a sense) by holomorphic functions on G/K (essentially, the Weierstrass polynomial approximation theorem).

Remark 5.2.3. A finite multiplicity theorem of the branching law (5.1.1) with respect to semisimple symmetric pairs (G, H) holds for more general π (i.e., π is not a highest weight module), under the assumption that π is discretely decomposable as an $(\mathfrak{h}_{\mathbb{C}}, H \cap K)$-module (see [41, Corollary 4.3], [45]). However, the multiplicity of the branching law can be infinite if the restriction is not discretely decomposable (see Example 6.3).

Remark 5.2.4. Theorems B and D assert that multiplicities $m_\pi(\mu)$ in Fact 5.1 are **uniformly bounded** when we vary μ. This is a distinguished feature for the restriction of highest weight representations π. A similar statement may fail if π is not a highest weight module (see Example 6.2).

5.3 Reduction to the scalar type case

In order to deduce Theorem D (1) from Theorem D (2), we use the idea of 'coherent family' of representations of reductive Lie groups (for example, see [85]). For this, we prepare the following Lemma 5.3 and Proposition 5.4.1.

Lemma 5.3. *Suppose that $(\pi, \mathcal{H})$ is an irreducible unitary highest weight representation of G. Then there exist an irreducible unitary highest weight representation π' of scalar type and a finite-dimensional representation F of G such that the underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module π_K occurs as a subquotient of the tensor product $\pi'_K \otimes F$.*

Proof. Without loss of generality, we may and do assume that G is simply connected. Since G is a simple Lie group of Hermitian type, the center $\mathfrak{c}(\mathfrak{k})$ of $\mathfrak{k}$ is one-dimensional. We take its generator Z as in Subsection 1.4, and write C for the connected subgroup with Lie algebra $\mathfrak{c}(\mathfrak{k})$. Then, K is isomorphic to the direct product group of C and a semisimple group K'.

As $(\pi, \mathcal{H})$ is an irreducible unitary highest weight representation of G, $\mathcal{H}_K^{\mathfrak{p}_+}$ is an irreducible (finite-dimensional) unitary representation of K. The K-module $\mathcal{H}_K^{\mathfrak{p}_+}$ has an expression $\sigma \otimes \chi_0$, where $\sigma \in \widehat{K}$ such that $\sigma|_C$ is trivial and χ_0 is a unitary character of K.

Let χ' be a unitary character of K such that χ' is trivial on the center Z_G of G (namely, χ' is well defined as a representation of $\mathrm{Ad}_G(K) \simeq K/Z_G$). For later purposes, we take χ' such that $-\sqrt{-1}\, d\chi'(Z) \gg 0$. There exists an irreducible finite-dimensional representation F of G such that $F^{\mathfrak{p}_+} \simeq \sigma \otimes \chi'$ as K-modules because $\sigma \otimes \chi'$ is well defined as an algebraic representation of $\mathrm{Ad}_G(K)$.

We set $\chi := \chi_0 \otimes (\chi')^*$ of K. Because $-\sqrt{-1}\, d\chi(Z) \ll 0$, the irreducible highest weight $(\mathfrak{g}_{\mathbb{C}}, K)$-module V' such that $(V')^{\mathfrak{p}_+} \simeq \chi$ is unitarizable. Let $(\pi', \mathcal{H}')$ denote the irreducible unitary representation of G whose underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module $\mathcal{H}'_K$ is isomorphic to V'. Since $\mathcal{H}'_K$ is an irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module, $\mathcal{H}'_K \otimes F$ is a $(\mathfrak{g}_{\mathbb{C}}, K)$-module of finite length. Furthermore, as $\mathcal{H}'_K$ is a highest weight module, so are all subquotient modules of $\mathcal{H}'_K \otimes F$. Then, $\mathcal{H}_K$ arises as a subquotient of $\mathcal{H}'_K \otimes F$ because the K-module $\mathcal{H}_K^{\mathfrak{p}_+}$ occurs as a subrepresentation of $(\mathcal{H}'_K \otimes F)^{\mathfrak{p}_+}$ in view of

$$\mathcal{H}_K^{\mathfrak{p}_+} \simeq \sigma \otimes \chi_0 \simeq \chi \otimes (\sigma \otimes \chi') \simeq (\mathcal{H}'_K)^{\mathfrak{p}_+} \otimes F^{\mathfrak{p}_+} \subset (\mathcal{H}'_K \otimes F)^{\mathfrak{p}_+} .$$

Hence, we have shown Lemma 5.3. □

5.4 Uniform estimate of multiplicities for tensor products

Let (π, X) be a $(\mathfrak{g}_{\mathbb{C}}, K)$-module of finite length. This means that π admits a chain of submodules

$$0 = Y_0 \subset Y_1 \subset \cdots \subset Y_N = X \tag{5.4.1}$$

such that Y_i/Y_{i-1} is irreducible for $i = 1, \ldots, N$. The number N is independent of the choice of the chain (5.4.1), and we will write

$$m(\pi) := N .$$

That is, $m(\pi)$ is the number of irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-modules (counted with multiplicity) occurring as subquotients in π. Here is a uniform estimate of $m(\pi)$ under the operation of tensor products:

Proposition 5.4.1. *Let F be a finite-dimensional representation of a real reductive connected Lie group G. Then there exists a constant $C \equiv C(F)$ such that*

$$m(\pi \otimes F) \le C$$

for any irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module π.

Before entering the proof, we fix some terminologies:

Definition 5.4.2. *We write* $\mathcal{F}(\mathfrak{g}_{\mathbb{C}}, K)$ *for the category of* $(\mathfrak{g}_{\mathbb{C}}, K)$*-modules of finite length. The* Grothendieck group $\mathcal{V}(\mathfrak{g}_{\mathbb{C}}, K)$ *of* $\mathcal{F}(\mathfrak{g}_{\mathbb{C}}, K)$ *is the abelian group generated by* $(\mathfrak{g}_{\mathbb{C}}, K)$*-modules of finite length, modulo the equivalence relations*

$$X \sim Y + Z$$

whenever there is a short exact sequence

$$0 \to Y \to X \to Z \to 0$$

of $(\mathfrak{g}_{\mathbb{C}}, K)$*-modules. Then*

$$m : \mathcal{F}(\mathfrak{g}_{\mathbb{C}}, K) \to \mathbb{N}$$

induces a group homomorphism of abelian groups:

$$m : \mathcal{V}(\mathfrak{g}_{\mathbb{C}}, K) \to \mathbb{Z}\,.$$

The Grothendieck group $\mathcal{V}(\mathfrak{g}_{\mathbb{C}}, K)$ is isomorphic to the free abelian group having irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-modules as its set of finite generators.

Suppose (π, X) is a $(\mathfrak{g}_{\mathbb{C}}, K)$-module of finite length. Then, in the Grothendieck group $\mathcal{V}(\mathfrak{g}_{\mathbb{C}}, K)$, we have the relation

$$X = \bigoplus_{Y} m_\pi(Y) Y\,, \tag{5.4.2}$$

where the sum is taken over irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-modules. Then we have

$$m(\pi) = \sum_{Y} m_\pi(Y)\,. \tag{5.4.3}$$

Suppose (π', X') is also a $(\mathfrak{g}_{\mathbb{C}}, K)$-modules of finite length. We set

$$[\pi : \pi'] := \dim \operatorname{Hom}_{(\mathfrak{g}_{\mathbb{C}}, K)}\left(\bigoplus_{Y} m_\pi(Y)Y, \bigoplus_{Y} m_{\pi'}(Y)Y\right) \tag{5.4.4}$$

$$= \sum_{Y} m_\pi(Y) m_{\pi'}(Y)\,. \tag{5.4.5}$$

The definition (5.4.4) makes sense in a more general setting where one of Xor X' is not of finite length. To be more precise, we recall from [41, Definition 1.1]:

Definition 5.4.3. *Let* $\mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$ *be the category of* $(\mathfrak{g}_{\mathbb{C}}, K)$*-modules* (π, X) *having the following properties:*
1) (K-admissibility) $\dim \operatorname{Hom}_K(\tau, \pi) < \infty$ *for any* $\tau \in \widehat{K}$.
2) (discretely decomposability, see [41, Definition 1.1]) X admits an increasing filtration

$$0 = Y_0 \subset Y_1 \subset Y_2 \subset \cdots$$

of $\mathfrak{g}_{\mathbb{C}}$*-modules such that* Y_i / Y_{i-1} *is of finite length and that* $X = \bigcup_{i=1}^{\infty} Y_i$.

We refer the reader to [41] for algebraic results on discretely decomposable $(\mathfrak{g}_{\mathbb{C}}, K)$-modules such as:

Lemma 5.4.4. *Suppose $X \in \mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$.*
1) Any submodule or quotient of X is an object of $\mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$.
2) The tensor product $X \otimes F$ is also an object of $\mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$ for any finite-dimensional $(\mathfrak{g}_{\mathbb{C}}, K)$-module.

For $X \in \mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$, we can take the filtration $\{Y_i\}$ such that Y_i/Y_{i-1} is irreducible as a $(\mathfrak{g}_{\mathbb{C}}, K)$-module for any i. Then, for any irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module,

$$\#\{i : Y_i/Y_{i-1} \text{ is isomorphic to } Y\}$$

is finite and independent of the filtration, which we will denote by $m_\pi(Y)$.

Definition 5.4.5. *Suppose $X \in \mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$. We say the $(\mathfrak{g}_{\mathbb{C}}, K)$-module X is multiplicity-free if*

$$m_\pi(Y) \leq 1 \quad \textit{for any irreducible } (\mathfrak{g}_{\mathbb{C}}, K)\textit{-module } Y.$$

This concept coincides with Definition 1.1 if X is the underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module of a unitary representation of G. The point of Definition 5.4.5 is that we allow the case where X is not unitarizable.

Generalizing (5.4.5), we set

$$[\pi : \pi'] := \sum_Y m_\pi(Y) m_{\pi'}(Y)$$

for $\pi, \pi' \in \mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$. Here are immediate results from the definition:

Lemma 5.4.6. *Let $\pi, \pi' \in \mathcal{A}(\mathfrak{g}_{\mathbb{C}}, K)$.*
1) $[\pi : \pi'] < \infty$ if at least one of π and π' belongs to $\mathcal{F}(\mathfrak{g}_{\mathbb{C}}, K)$.
2) $\dim \mathrm{Hom}_{(\mathfrak{g}_{\mathbb{C}}, K)}(\pi, \pi') \leq [\pi : \pi']$.
3) $[\pi : \pi'] = [\pi' : \pi]$.
4) $m_\pi(Y) = [\pi : Y]$ if Y is an irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module.
5) $[\pi : \pi'] \leq m(\pi)$ if π' is multiplicity-free.

Now, we return to Proposition 5.4.1.

Proof of Proposition 5.4.1. We divide the proof into three steps:

Step 1 *(π is a finite-dimensional representation).* We shall prove

$$m(\pi \otimes F) \leq \dim F \tag{5.4.6}$$

for any finite-dimensional representation π of G.

Let $\mathfrak{b} = \mathfrak{t} + \mathfrak{u}$ be a Borel subalgebra of $\mathfrak{g}_{\mathbb{C}}$ with $\mathfrak{u}$ nilradical. We denote by $H^j(\mathfrak{u}, V)$ the jth cohomology group of the Lie algebra $\mathfrak{u}$ with coefficients in a $\mathfrak{u}$-module V. Since the Lie algebra $\mathfrak{b}$ is solvable, we can choose a $\mathfrak{b}$-stable filtration

$$F = F_k \supset F_{k-1} \supset \cdots \supset F_0 = \{0\}$$

such that $\dim F_i/F_{i-1} = 1$.

Let us show by induction on i that

$$\dim H^0(\mathfrak{u}, \pi \otimes F_i) \le i\,. \tag{5.4.7}$$

This will imply $m(\pi \otimes F) = \dim H^0(\mathfrak{u}, \pi \otimes F) \le k = \dim F$.

The inequality (5.4.7) is trivial if $i = 0$. Suppose (5.4.7) holds for $i-1$. The short exact sequence of $\mathfrak{b}$-modules

$$0 \to \pi \otimes F_{i-1} \to \pi \otimes F_i \to \pi \otimes (F_i/F_{i-1}) \to 0$$

gives rise to a long exact sequence

$$\begin{aligned} 0 \to H^0(\mathfrak{u}, \pi \otimes F_{i-1}) &\to H^0(\mathfrak{u}, \pi \otimes F_i) \to H^0(\mathfrak{u}, \pi \otimes (F_i/F_{i-1})) \\ &\to H^1(\mathfrak{u}, \pi \otimes F_{i-1}) \to \dots \end{aligned}$$

of $\mathfrak{t}$-modules. In particular, we have

$$\dim H^0(\mathfrak{u}, \pi \otimes F_i) \le \dim H^0(\mathfrak{u}, \pi \otimes F_{i-1}) + \dim H^0(\mathfrak{u}, \pi \otimes (F_i/F_{i-1}))\,. \tag{5.4.8}$$

Because F_i/F_{i-1} is trivial as a $\mathfrak{u}$-module, we have

$$H^0(\mathfrak{u}, \pi \otimes (F_i/F_{i-1})) = H^0(\mathfrak{u}, \pi) \otimes (F_i/F_{i-1})\,. \tag{5.4.9}$$

By definition $H^0(\mathfrak{u}, \pi)$ is the space of highest weight vectors, and therefore the dimension of the right-hand side of (5.4.9) is one. Now, the inductive assumption combined with (5.4.8) implies $\dim H^0(\mathfrak{u}, \pi \otimes F_i) \le i$, as desired.

Step 2 *(π is a principal series representation).* In this step, we consider the case where π is a principal series representation. We note that π may be reducible here.

Let $P = LN$ be a Levi decomposition of a minimal parabolic subgroup P of G, W an irreducible (finite-dimensional) representation of L, and $\operatorname{Ind}_P^G(W)$ the underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module of a principal series representation induced from the representation $W \boxtimes \mathbf{1}$ of $P = LN$ (without ρ-shift). Then, the socle filtration is unchanged so far as the parameter lies in the equisingular set, and thus, there are only finitely many possibilities of the socle filtration of $\operatorname{Ind}_P^G(W)$ for irreducible representations W of L. We denote by $m(G)$ the maximum of $m(\operatorname{Ind}_P^G(W))$ for irreducible representations W of L.

Let F be a finite-dimensional representation of G. Then we have an isomorphism of $(\mathfrak{g}_{\mathbb{C}}, K)$-modules

$$\operatorname{Ind}_P^G(W) \otimes F \simeq \operatorname{Ind}_P^G(W \otimes F)\,,$$

where F is regarded as a P-module on the right-hand side. We take a P-stable filtration

$$W_n := W \otimes F \supset W_{n-1} \supset \cdots \supset W_0 = \{0\}$$

such that each W_i/W_{i-1} is irreducible as a P-module. We notice that $n \le \dim F$ by applying Step 1 to the L-module $F|_L$. As $\operatorname{Ind}_P^G(W \otimes F)$ is isomorphic to $\bigoplus_{i=1}^n \operatorname{Ind}_P^G(W_i/W_{i-1})$ in the Grothendieck group $\mathcal{V}(\mathfrak{g}_{\mathbb{C}}, K)$, we have shown that

$$m(\operatorname{Ind}_P^G(W) \otimes F) \le n\, m(G) \le (\dim F)\, m(G)$$

for any irreducible finite-dimensional representation W of L.

Step 3 *(general case).* By Casselman's subrepresentation theorem (see [87, Chapter 3]), any irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module π is realized as a subrepresentation of some induced representation $\operatorname{Ind}_P^G(W)$. Then

$$m(\pi \otimes F) \le m(\pi \otimes \operatorname{Ind}_P^G(W)) \le C$$

by step 2. Thus, Proposition 5.4.1 is proved. □

5.5 Proof of Theorem D

Now let us complete the proof of Theorem D.

Let $\pi = \pi_1 \boxtimes \pi_2$ be an irreducible unitary highest weight representation of $G' := G \times G$. It follows from Lemma 5.3 that there exist an irreducible unitary highest weight representation $\pi' = \pi_1' \boxtimes \pi_2'$ of scalar type and a finite-dimensional representation F of G' such that π_K occurs as a subquotient of $\pi_K' \otimes F$.

By using the notation (5.4.4), we set $[V_1 : V_2] := [(V_1)_K : (V_2)_K]$ for G-modules V_1 and V_2 of finite length. Then, for $\mu \in \widehat{G}$, we have

$$\begin{aligned} m_{\pi_1,\pi_2}(\mu) &= \dim \operatorname{Hom}_G(\mu, \pi|_{\operatorname{diag}(G)}) \\ &\le [\mu : \pi|_{\operatorname{diag}(G)}] \\ &\le [\mu : (\pi' \otimes F)|_{\operatorname{diag}(G)}] \\ &= [\mu \otimes (F^*|_{\operatorname{diag}(G)}) : \pi'|_{\operatorname{diag}(G)}] \\ &\le m(\mu \otimes (F^*|_{\operatorname{diag}(G)})) \qquad (5.5.1) \\ &\le C(F^*)\,. \end{aligned}$$

Here the inequality (5.5.1) follows from Lemma 5.4.6 (5) because $\pi'|_{\operatorname{diag}(G)} \simeq \pi_1' \widehat{\otimes} \pi_2'$ is multiplicity-free (see Theorem D (2)). In the last inequality, $C(F^*)$ is the constant in Proposition 5.4.1. This completes the proof of Theorem D (1). □

Remark 5.5. The argument in Subsections 8.8 and 8.9 gives a different and more straightforward proof of Theorem D.

6 Counterexamples

In this section, we analyze the assumptions in Theorems A and B by counterexamples, that is, how the conclusions fail if we relax the assumptions on the representation π.

Let (G, H) be a reductive symmetric pair corresponding to an involutive automorphism τ of G, and π an irreducible unitary representation of G. We shall see that the multiplicity of an irreducible summand occurring in the restriction $\pi|_H$ can be:
1) **greater than one** if π is not of scalar type (but we still assume that π is a highest weight module);
2) **finite but not uniformly bounded** if π is not a highest weight module (but we still assume that $\pi|_H$ decomposes discretely);
3) **infinite** if $\pi|_H$ contains continuous spectra.

Although our concern in this paper is mainly with a non-compact subgroup H, we can construct such examples for (1) and (2) even for $H = K$ (a maximal compact subgroup modulo the center of G).

Case (1) will be discussed in Subsection 6.1, (2) in Subsection 6.2, and (3) in Subsection 6.3, respectively. To construct an example for (3), we use those for (1) and (2).

6.1 Failure of multiplicity-free property

Let $G = Sp(2, \mathbb{R})$. Then, the maximal compact subgroup K is isomorphic to $U(2)$. We take a compact Cartan subalgebra $\mathfrak{t}$. Let $\{f_1, f_2\}$ be the standard basis of $\sqrt{-1}\,\mathfrak{t}^*$ such that $\Delta(\mathfrak{g}, \mathfrak{t}) = \{\pm f_1 \pm f_2, \pm 2f_1, \pm 2f_2\}$, and we fix a positive system $\Delta^+(\mathfrak{k}, \mathfrak{t}) := \{f_1 - f_2\}$. In what follows, we shall use the notation (λ_1, λ_2) to denote the character $\lambda_1 f_1 + \lambda_2 f_2$ of $\mathfrak{t}$.

Given $(p, q) \in \mathbb{Z}^2$ with $p \geq q$, we denote by $\pi^{U(2)}_{(p,q)}$ the irreducible representation of $U(2)$ with highest weight $(p, q) = pf_1 + qf_2$. Then $\dim \pi^{U(2)}_{(p,q)} = p - q + 1$.

The set of holomorphic discrete series representations of G is parametrized by $\lambda := (\lambda_1, \lambda_2) \in \mathbb{N}^2$ with $\lambda_1 > \lambda_2 > 0$. We set $\mu \equiv (\mu_1, \mu_2) := (\lambda_1 + 1, \lambda_2 + 2)$ and denote by $\pi^G_\mu \equiv \pi^{Sp(2,\mathbb{R})}_{(\mu_1,\mu_2)}$ the holomorphic discrete series representation of G characterized by

$$Z(\mathfrak{g})\text{-infinitesimal character } = (\lambda_1, \lambda_2) \quad \text{(Harish-Chandra parameter)},$$

$$\text{minimal } K\text{-type } = \pi^{U(2)}_{(\mu_1,\mu_2)} \quad \text{(Blattner parameter)}.$$

We note that π^G_μ is of scalar type if and only if $\mu_1 = \mu_2$.

We know from Theorem B that multiplicities of K-type τ occurring in π^G_μ are uniformly bounded for fixed $\mu = (\mu_1, \mu_2)$. Here is the formula:

Example 6.1 *(upper bound of K-multiplicities of holomorphic discrete series).*

$$\sup_{\tau \in \widehat{K}} \dim \operatorname{Hom}_K(\tau, \pi^G_\mu|_K) = \left[\frac{\mu_1 - \mu_2 + 2}{2}\right]. \tag{6.1.1}$$

The right side of (6.1.1) $= 1$ if and only if either of the following two cases holds:

$$\mu_1 = \mu_2 \quad \text{(i.e., } \pi^G_\mu \text{ is of scalar type)}, \qquad (6.1.2)\,(a)$$

$$\mu_1 = \mu_2 + 1 \quad \text{(i.e., } \pi^G_\mu \text{ is of two-dimensional minimal } K\text{-type)}. \qquad (6.1.2)\,(b)$$

Thus, the branching law of the restriction $\pi_\mu^G|_K$ is multiplicity-free if and only if $\mu_1 = \mu_2$ or $\mu_1 = \mu_2+1$. The multiplicity-free property for $\mu_1 = \mu_2$ (i.e., for π_μ^G of scalar type) follows from Theorem A. The multiplicity-free property for $\mu_1 = \mu_2+1$ is outside of the scope of this paper, but can be explained in the general framework of the 'vector bundle version' of Theorem 2.2 (see [47, Theorem 2], [49]).

Proof. It follows from the Blattner formula for a holomorphic discrete series representation ([32], [78]) that the K-type formula of π_μ^G is given by

$$\begin{aligned} \pi_\mu^G|_K &\simeq \pi_{(\mu_1,\mu_2)}^{U(2)} \otimes S(\mathbb{C}^3) \\ &= \pi_{(\mu_1,\mu_2)}^{U(2)} \otimes \bigoplus_{\substack{a\geq b\geq 0 \\ (a,b)\in\mathbb{N}^2}} \pi_{(2a,2b)}^{U(2)}, \end{aligned} \tag{6.1.3}$$

where $K = U(2)$ acts on $\mathbb{C}^3 \simeq S^2(\mathbb{C}^2)$ as the symmetric tensor of the natural representation. We write $n_\mu(p,q)$ for the multiplicity of the K-type $\pi_{(p,q)}^{U(2)}$ occurring in $\pi_\mu^G \equiv \pi_{(\mu_1,\mu_2)}^{Sp(2,\mathbb{R})}$, that is,

$$n_\mu(p,q) := \dim \mathrm{Hom}_K(\pi_{(p,q)}^K, \pi_\mu^G|_K).$$

Then, applying the Clebsch–Gordan formula (1.7.1) (f) to (6.1.3), we obtain

$$n_\mu(p,q) = \#\{(a,b) \in \mathbb{N}^2 : (a,b) \text{ satisfies } a \geq b \geq 0, \text{(6.1.4) and (6.1.5)}\},$$

where

$$p + q = \mu_1 + \mu_2 + 2a + 2b, \tag{6.1.4}$$

$$\max(2a + \mu_2, 2b + \mu_1) \leq p \leq 2a + \mu_1. \tag{6.1.5}$$

In particular, for fixed (μ_1, μ_2) and (p,q), the integer b is determined by a from (6.1.4), whereas the integer a satisfies the inequalities $p - \mu_1 \leq 2a \leq p - \mu_2$. Therefore,

$$n_\mu(p,q) \leq \left[\frac{(p-\mu_2)-(p-\mu_1)}{2}\right] + 1 = \left[\frac{\mu_1 - \mu_2 + 2}{2}\right]. \qquad \square$$

6.2 Failure of uniform boundedness

We continue the setting of Subsection 6.1. Let B be a Borel subgroup of $G_\mathbb{C} \simeq Sp(2,\mathbb{C})$. Then, there exist 4 closed orbits of $K_\mathbb{C} \simeq GL(2,\mathbb{C})$ on the full flag variety $G_\mathbb{C}/B$. (By the Matsuki duality, there exist 4 open orbits of $G = Sp(2,\mathbb{R})$ on $G_\mathbb{C}/B$. This observation will be used in the proof of Example 6.3.) By the Beilinson–Bernstein correspondence, we see that there are 4 series of discrete series representations of G. Among them, two are holomorphic and anti-holomorphic discrete series representations, that is, π_μ^G and $(\pi_\mu^G)^*$ (the contragredient representation) with notation as in Subsection 6.1. The other two series are non-holomorphic discrete series representations. Let us parametrize them. For $\lambda := (\lambda_1, \lambda_2) \in \mathbb{Z}^2$ $(\lambda_1 > -\lambda_2 > 0)$, we write W_λ for the discrete series representation of G characterized by

$$\begin{aligned} &Z(\mathfrak{g})\text{-infinitesimal character } = (\lambda_1, \lambda_2) && \text{(Harish-Chandra parameter)},\\ &\text{minimal } K\text{-type } = \pi^{U(2)}_{(\lambda_1+1,\lambda_2)} && \text{(Blattner parameter)}. \end{aligned}$$

Then, non-holomorphic discrete series representations are either W_λ or its contragredient representation W_λ^* for some $\lambda \in \mathbb{Z}^2$ with $\lambda_1 > -\lambda_2 > 0$. We define a θ-stable Borel subalgebra $\mathfrak{q} = \mathfrak{t}_{\mathbb{C}} + \mathfrak{u}$ of $\mathfrak{g}_{\mathbb{C}} = \mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_{\mathbb{C}}$ such that

$$\Delta(\mathfrak{u} \cap \mathfrak{p}_{\mathbb{C}}, \mathfrak{t}) := \{2f_1, f_1 + f_2, -2f_2\}, \quad \Delta(\mathfrak{u} \cap \mathfrak{k}_{\mathbb{C}}, \mathfrak{t}) := \{f_1 - f_2\}.$$

Then, the Harish-Chandra module $(W_\lambda)_K$ is isomorphic to the cohomological parabolic induction $\mathcal{R}^1_{\mathfrak{q}}(\mathbb{C}_{(\lambda_1,\lambda_2)})$ of degree 1 as $(\mathfrak{g}_{\mathbb{C}}, K)$-modules with the notation and the normalization as in [86]. We set $\mu_1 := \lambda_1 + 1$ and $\mu_2 := \lambda_2$.

Example 6.2 *(multiplicity of K-type of non-holomorphic discrete series W_λ).* We write $m_\lambda(p, q)$ for the multiplicity of the K-type $\pi^{U(2)}_{(p,q)}$ occurring in W_λ, that is,

$$m_\lambda(p, q) := \dim \operatorname{Hom}_K(\pi^{U(2)}_{(p,q)}, W_\lambda|_K).$$

Then, $m_\lambda(p, q) \neq 0$ only if $(p, q) \in \mathbb{Z}^2$ satisfies

$$p \geq \mu_1, \ \ p - q \geq \mu_1 - \mu_2 \text{ and } p - q \in 2\mathbb{Z} + \mu_1 + \mu_2. \tag{6.2.1}$$

Then,

$$m_\lambda(p, q) = 1 + \min\left(\left[\frac{p - \mu_1}{2}\right], \frac{p - q - \mu_1 + \mu_2}{2}\right). \tag{6.2.2}$$

In particular, for each fixed λ, the K-multiplicity in W_λ is not uniformly bounded, namely,

$$\sup_{\tau \in \widehat{K}} \dim \operatorname{Hom}_K(\tau, W_\lambda|_K) = \sup_{(p,q) \text{ satisfies (6.2.1)}} m_\lambda(p, q) = \infty.$$

Proof. For $p, q \in \mathbb{Z}$, we write $\mathbb{C}_{(p,q)}$ for the one-dimensional representation of $\mathfrak{t}_{\mathbb{C}}$ corresponding to the weight $pf_1 + qf_2 \in \mathfrak{t}^*_{\mathbb{C}}$. According to the $\mathfrak{t}_{\mathbb{C}}$-module isomorphism:

$$\mathfrak{u} \cap \mathfrak{p}_{\mathbb{C}} \simeq \mathbb{C}_{(2,0)} \oplus \mathbb{C}_{(1,1)} \oplus \mathbb{C}_{(0,-2)},$$

the symmetric algebra $S(\mathfrak{u} \cap \mathfrak{p}_{\mathbb{C}})$ is decomposed into irreducible representations of $\mathfrak{t}_{\mathbb{C}}$ as

$$\begin{aligned} S(\mathfrak{u} \cap \mathfrak{p}_{\mathbb{C}}) &\simeq \bigoplus_{a,b,c \in \mathbb{N}} S^a(\mathbb{C}_{(2,0)}) \otimes S^b(\mathbb{C}_{(1,1)}) \otimes S^c(\mathbb{C}_{(0,-2)}) \\ &\simeq \bigoplus_{a,b,c \in \mathbb{N}} \mathbb{C}_{(2a+b,b-2c)}. \end{aligned} \tag{6.2.3}$$

We denote by $H^j(\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}, \pi)$ the jth cohomology group of the Lie algebra $\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}$ with coefficients in the $\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}$-module π. If π is a $\mathfrak{k}_{\mathbb{C}}$-module, then $H^j(\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}, \pi)$ becomes naturally a $\mathfrak{t}_{\mathbb{C}}$-module. Then, Kostant's version of the Borel–Weil–Bott theorem (e.g., [85, Chapter 3]) shows that

$$H^j(\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}, \pi^{U(2)}_{(p,q)}) = \begin{cases} \mathbb{C}_{(p,q)} & (j=0)\,, \\ \mathbb{C}_{(q-1,p+1)} & (j=1)\,, \\ \{0\} & (j \neq 0,1)\,. \end{cases} \tag{6.2.4}$$

By using the Blattner formula due to Hecht–Schmid (e.g., [85, Theorem 6.3.12]), the K-type formula of W_λ is given by

$$\begin{aligned} m_\lambda(p,q) &= \dim \operatorname{Hom}_K(\pi^{U(2)}_{(p,q)}, W_\lambda|_K) \\ &= \sum_{j=0}^{1}(-1)^j \dim \operatorname{Hom}_{\mathfrak{t}_{\mathbb{C}}}(H^j(\mathfrak{u}\cap\mathfrak{k}_{\mathbb{C}}, \pi^{U(2)}_{(p,q)}), S(\mathfrak{u}\cap\mathfrak{p}_{\mathbb{C}})\otimes\mathbb{C}_{(\mu_1,\mu_2)})\,. \end{aligned}$$

Now, comparing (6.2.3) with the above formula (6.2.4) as $\mathfrak{t}_{\mathbb{C}}$-modules, we see

$$\begin{aligned} m_\lambda(p,q) &= \#\{(a,b,c)\in\mathbb{N}^3 : p = 2a+b+\mu_1, q = b-2c+\mu_2\} \\ &\quad - \#\{(a,b,c)\in\mathbb{N}^3 : q-1 = 2a+b+\mu_1, p+1 = b-2c+\mu_2\} \\ &= \#\{(a,b,c)\in\mathbb{N}^3 : p = 2a+b+\mu_1, q = b-2c+\mu_2\} \\ &= 1+\min\left(\left[\frac{p-\mu_1}{2}\right], \frac{p-q-\mu_1+\mu_2}{2}\right)\,. \end{aligned}$$

Thus, the formula (6.2.2) has been verified. □

6.3 Failure of finiteness of multiplicities

Multiplicities of the branching laws can be infinite in general even for reductive symmetric pairs (G, H). In this subsection, we review from [43, Example 5.5] a curious example of the branching law, in which the multiplicity of a discrete summand is non-zero and finite and that of another discrete summand is infinite. Such a phenomenon happens only when continuous spectra appear.

Example 6.3 *(infinite and finite multiplicities).* Let $(G_{\mathbb{C}}, G)$ be a reductive symmetric pair $(Sp(2,\mathbb{C}), Sp(2,\mathbb{R}))$. We note that $(G_{\mathbb{C}}, G)$ is locally isomorphic to the symmetric pair $(SO(5,\mathbb{C}), SO(3,2))$. We take a Cartan subgroup $H = TA$ of $G_{\mathbb{C}}$. We note that $T \simeq \mathbb{T}^2$ and $A \simeq \mathbb{R}^2$, and identify $\widehat{T}$ with $\mathbb{Z}^2$.

Let $\varpi \equiv \varpi^{Sp(2,\mathbb{C})}_{(a,b)}$ be the unitary principal series representation of $G_{\mathbb{C}}$ induced unitarily from the character χ of a Borel subgroup B containing $H = TA$ such that

$$\chi|_H \simeq \mathbb{C}_{(a,b)} \boxtimes \mathbf{1}\,.$$

We assume $a, b \geq 0$ and set

$$c(\mu_1, \mu_2; a, b) := \#\{(s, t, u) \in \mathbb{N}^3 : a = \mu_1 + 2s + t,\ b = \mu_2 + t + 2u\}.$$

Then, the discrete part of the branching law of the restriction $\varpi_{(a,b)}^{Sp(2,\mathbb{C})}|_{Sp(2,\mathbb{R})}$ is given by the following spectra:

$$\bigoplus_{\mu_1 \geq \mu_2 \geq 3} c(\mu_1, \mu_2; a, b) \left(\pi_{(\mu_1,\mu_2)}^{Sp(2,\mathbb{R})} \oplus \left(\pi_{(\mu_1,\mu_2)}^{Sp(2,\mathbb{R})}\right)^*\right) \oplus \sum_{\lambda_1 > -\lambda_2 > 0}^{\oplus} \infty(W_\lambda \oplus W_\lambda^*), \tag{6.3.1}$$

with the notation as in Examples 6.1 and 6.2.

The first term of (6.3.1) is a finite sum because there are at most finitely many (μ_1, μ_2) such that $c(\mu_1, \mu_2; a, b) \neq 0$ for each fixed (a, b). For instance, the first term of (6.3.1) amounts to

$$\bigoplus_{\substack{3 \leq \mu_1 \leq a \\ \mu_1 \equiv a \bmod 2}} \pi_{(\mu_1,3)}^{Sp(2,\mathbb{R})} \oplus \bigoplus_{\substack{3 \leq \mu_1 \leq a \\ \mu_1 \equiv a \bmod 2}} \left(\pi_{(\mu_1,3)}^{Sp(2,\mathbb{R})}\right)^* \quad \text{(multiplicity-free)}$$

if $b = 3$.

The second term of (6.3.1) is nothing other than the direct sum of all non-holomorphic discrete series representations of $G = Sp(2, \mathbb{R})$ with infinite multiplicities for any a and b.

Sketch of Proof. There exist 4 open G-orbits on $G_{\mathbb{C}}/B$, for which the isotropy subgroups are all isomorphic to $T \simeq \mathbb{T}^2$. By the Mackey theory, the restriction $\varpi_{(a,b)}^{G_{\mathbb{C}}}|_G$ is unitarily equivalent to the direct sum of the regular representations realized on L^2-sections of G-equivariant line bundles $G \times_T \mathbb{C}_{(\pm a, \pm b)} \to G/T$. That is,

$$\varpi_{(a,b)}^{G_{\mathbb{C}}}|_G \simeq \bigoplus_{\varepsilon_1, \varepsilon_2 = \pm 1} L^2(G/T, \mathbb{C}_{(\varepsilon_1 a, \varepsilon_2 b)}).$$

Therefore, an irreducible unitary representation σ of G occurs as a discrete spectrum in $\varpi_{(a,b)}^{G_{\mathbb{C}}}|_G$ if and only if σ occurs as a discrete summand in $L^2(G/T, \mathbb{C}_{(\varepsilon_1 a, \varepsilon_2 b)})$ for some $\varepsilon_1, \varepsilon_2 = \pm 1$. Further, the multiplicity is given by

$$\dim \operatorname{Hom}_G(\sigma, \varpi_{(a,b)}^{G_{\mathbb{C}}}|_G) = \sum_{\varepsilon_1, \varepsilon_2 = \pm 1} \dim \operatorname{Hom}_{\mathbb{T}^2}(\mathbb{C}_{(\varepsilon_1 a, \varepsilon_2 b)}, \sigma|_{\mathbb{T}^2}),$$

by the Frobenius reciprocity theorem.

Since T is compact, σ must be a discrete series representation of $G = Sp(2, \mathbb{R})$ if σ occurs in $L^2(G/T, \mathbb{C}_{(\varepsilon_1 a, \varepsilon_2 b)})$ as a discrete summand. We divide the computation of multiplicities into the following two cases:

Case I. σ is a holomorphic series representation or its contragredient representation. Let $\sigma = \pi_\mu^{Sp(2,\mathbb{R})}$. Combining (6.1.3) with the weight formulae

$$S(\mathbb{C}^3)|_{\mathbb{T}^2} \simeq \bigoplus_{s,t,u\in\mathbb{N}} S^s(\mathbb{C}_{(2,0)}) \otimes S^t(\mathbb{C}_{(1,1)}) \otimes S^u(\mathbb{C}_{(0,2)}) \simeq \bigoplus_{s,t,u\in\mathbb{N}} \mathbb{C}_{(2s+t,t+2u)}\,,$$

$$\pi^{U(2)}_{(\mu_1,\mu_2)}|_{\mathbb{T}^2} \simeq \bigoplus_{\substack{p+q=\mu_1+\mu_2\\ \mu_2\le p\le\mu_1}} \mathbb{C}_{(p,q)}\,,$$

we have

$$\dim \mathrm{Hom}_{\mathbb{T}^2}(\mathbb{C}_{(a,b)}, \pi^{Sp(2,\mathbb{R})}_{\mu}|_{\mathbb{T}^2}) = c(\mu_1, \mu_2; a, b)\,.$$

Case II. σ is a non-holomorphic discrete series representation. Let $\sigma = W_\lambda$. It follows from the K-type formula (6.2.2) of W_λ that we have

$$\dim \mathrm{Hom}_{\mathbb{T}^2}(\mathbb{C}_{(a,b)}, W_\lambda|_{\mathbb{T}^2}) = \sum_{p\ge q} m_\lambda(p,q) \dim \mathrm{Hom}_{\mathbb{T}^2}(\mathbb{C}_{(a,b)}, \pi^{U(2)}_{(p,q)}) = \infty\,.$$

Likewise for $\sigma = W^*_\lambda$ (the contragredient representation).

Hence, the discrete part of the branching law is given by (6.3.1). □

7 Finite-dimensional cases — Proof of Theorems E and F

7.1 Infinite v.s. finite-dimensional representations

Our method applied to infinite-dimensional representations in Sections 3 and 4 also applies to **finite**-dimensional representations, leading us to multiplicity-free theorems, as stated in Theorems E and F in Section 1, for the restriction with respect to symmetric pairs.

The comparison with multiplicity-free theorems in the infinite-dimensional case is illustrated by the following correspondence:

a non-compact simple group G	$\leftrightarrow$	a compact simple group G_U
a unitary highest weight module	$\leftrightarrow$	a finite-dimensional module
scalar type (Definition 1.3)	$\leftrightarrow$	"pan type" (Definition 7.3.3)
Theorems A and B	$\leftrightarrow$	Theorems E and F.

The main goal of this section is to give a proof of Theorems E and F by using Theorem 2.2. Geometrically, our proof is built on the fact that the H_U action on the Hermitian symmetric space is strongly visible if (G_U, H_U) is a symmetric pair (see [50]).

7.2 Representations associated to maximal parabolic subalgebras

Let $\mathfrak{g}_\mathbb{C}$ be a complex simple Lie algebra. We take a Cartan subalgebra $\mathfrak{j}$ of $\mathfrak{g}_\mathbb{C}$, and fix a positive system $\Delta^+(\mathfrak{g}_\mathbb{C}, \mathfrak{j})$. We denote by $\{\alpha_1, \dots, \alpha_n\}$ the set of simple roots, and by $\{\omega_1, \dots, \omega_n\}$ $(\subset \mathfrak{j}^*)$ the set of the fundamental weights.

We denote by $\pi_\lambda^{\mathfrak{g}_\mathbb{C}}$ irreducible finite-dimensional representation of $\mathfrak{g}_\mathbb{C}$ with highest weight $\lambda = \sum_{i=1}^n m_i\omega_i$ for $m_1, \dots, m_n \in \mathbb{N}$. It is also regarded as a holomorphic representation of $G_\mathbb{C}$, a simply connected complex Lie group with Lie algebra $\mathfrak{g}_\mathbb{C}$.

We fix a simple root α_i, and define a maximal parabolic subalgebra

$$\mathfrak{p}_{i\mathbb{C}}^- := \mathfrak{l}_{i\mathbb{C}} + \mathfrak{n}_{i\mathbb{C}}^-$$

such that the nilradical $\mathfrak{n}_{i\mathbb{C}}^-$ and the Levi part $\mathfrak{l}_{i\mathbb{C}}$ ($\supset \mathfrak{j}$) are given by

$$\Delta(\mathfrak{l}_{i\mathbb{C}}, \mathfrak{j}) = \mathbb{Z}\text{-span of } \{\alpha_1, \dots, \widehat{\alpha_i}, \dots, \alpha_n\} \cap \Delta(\mathfrak{g}_\mathbb{C}, \mathfrak{j}),$$
$$\Delta(\mathfrak{n}_{i\mathbb{C}}^-, \mathfrak{j}) = \Delta^-(\mathfrak{g}_\mathbb{C}, \mathfrak{j}) \setminus \Delta(\mathfrak{l}_{i\mathbb{C}}, \mathfrak{j}).$$

We shall see that irreducible finite-dimensional representations realized on generalized flag varieties $G_\mathbb{C}/P_\mathbb{C}$ is multiplicity-free with respect to any symmetric pairs if $P_\mathbb{C}$ has an abelian unipotent radical.

We write $P_{i\mathbb{C}}^- = L_{i\mathbb{C}} N_{i\mathbb{C}}^-$ for the corresponding maximal parabolic subgroup of $G_\mathbb{C}$.

Let $\mathrm{Hom}(\mathfrak{p}_{i\mathbb{C}}^-, \mathbb{C})$ be the set of Lie algebra homomorphisms over $\mathbb{C}$. Since any such homomorphism vanishes on the derived ideal $[\mathfrak{p}_{i\mathbb{C}}^-, \mathfrak{p}_{i\mathbb{C}}^-]$, $\mathrm{Hom}(\mathfrak{p}_{i\mathbb{C}}^-, \mathbb{C})$ is naturally identified with

$$\mathrm{Hom}(\mathfrak{p}_{i\mathbb{C}}^-/[\mathfrak{p}_{i\mathbb{C}}^-, \mathfrak{p}_{i\mathbb{C}}^-], \mathbb{C}) \simeq \mathbb{C}\omega_i .$$

Next, let $\mathrm{Hom}(P_{i\mathbb{C}}^-, \mathbb{C}^\times)$ be the set of complex Lie group homomorphisms. Then, we can regard $\mathrm{Hom}(P_{i\mathbb{C}}^-, \mathbb{C}^\times) \subset \mathrm{Hom}(\mathfrak{p}_{i\mathbb{C}}^-, \mathbb{C})$. As its subset, $\mathrm{Hom}(P_{i\mathbb{C}}^-, \mathbb{C}^\times)$ is identified with $\mathbb{Z}\omega_i$ since $G_\mathbb{C}$ is simply connected.

For $k \in \mathbb{Z}$, we write $\mathbb{C}_{k\omega_i}$ for the corresponding character of $P_{i\mathbb{C}}^-$, and denote by

$$\mathcal{L}_{k\omega_i} := G_\mathbb{C} \times_{P_{i\mathbb{C}}^-} \mathbb{C}_{k\omega_i} \to G_\mathbb{C}/P_{i\mathbb{C}}^- \tag{7.2.1}$$

the associated holomorphic line bundle. We naturally have a representation of $G_\mathbb{C}$ on the space of holomorphic sections $\mathcal{O}(\mathcal{L}_{k\omega_i})$. Then, by the Borel–Weil theory, $\mathcal{O}(\mathcal{L}_{k\omega_i})$ is non-zero and irreducible if $k \geq 0$ and we have an isomorphism of representations of $G_\mathbb{C}$ (also of $\mathfrak{g}_\mathbb{C}$):

$$\pi_{k\omega_i}^{\mathfrak{g}_\mathbb{C}} \simeq \mathcal{O}\left(\mathcal{L}_{k\omega_i}\right). \tag{7.2.2}$$

7.3 Parabolic subalgebra with abelian nilradical

A parabolic subalgebra with abelian nilradical is automatically a maximal parabolic subalgebra. Conversely, the nilradical of a maximal parabolic subalgebra is not necessarily abelian. We recall from Richardson–Röhrle–Steinberg [75] the following equivalent characterization of such parabolic algebras:

Lemma 7.3.1. *Retain the setting of Subsection 7.2. Then the following four conditions on the pair* $(\mathfrak{g}_{\mathbb{C}}, \alpha_i)$ *are equivalent:*
i) *The nilradical* $\mathfrak{n}_{i\mathbb{C}}^-$ *is abelian.*
ii) $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{l}_{i\mathbb{C}})$ *is a symmetric pair.*
iii) *The simple root* α_i *occurs in the highest root with coefficient one.*
iv) $(\mathfrak{g}_{\mathbb{C}}, \alpha_i)$ *is in the following list if we label the simple roots* $\alpha_1, \dots, \alpha_n$ *in the Dynkin diagram as in Table 7.3.2.*

Type A_n $\alpha_1, \alpha_2, \dots, \alpha_n$ (7.3.1)
Type B_n α_1 (7.3.2)
Type C_n α_n (7.3.3)
Type D_n $\alpha_1, \alpha_{n-1}, \alpha_n$ (7.3.4)
Type E_6 α_1, α_6 (7.3.5)
Type E_7 α_7 (7.3.6)

For types G_2*,* F_4*,* E_8*, there are no maximal parabolic subalgebras with abelian nilradicals.*

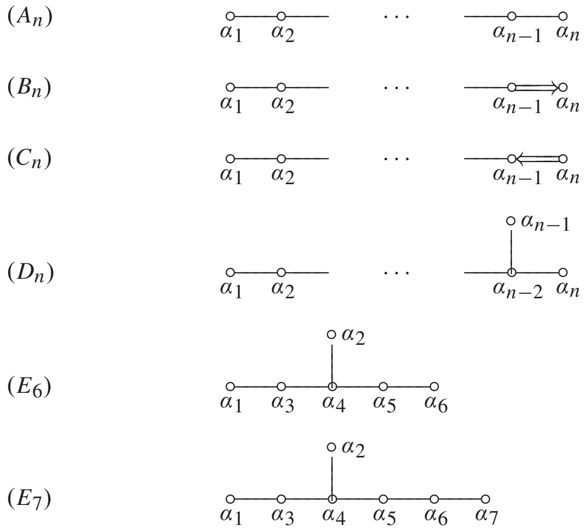
Proof. See [75] for the equivalence (i) ⇔ (iii) ⇔ (iv). The implication (iv) ⇒ (ii) is straightforward. For the convenience of the reader, we present a table of the symmetric pairs $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{l}_{i\mathbb{C}})$ corresponding to the index i in (iv).

Type	$\mathfrak{g}_{\mathbb{C}}$	$\mathfrak{l}_{i\mathbb{C}}$	i
A_n	$\mathfrak{sl}(n+1, \mathbb{C})$	$\mathfrak{sl}(i, \mathbb{C}) + \mathfrak{sl}(n+1-i, \mathbb{C}) + \mathbb{C}$	$i = 1, 2, \dots, n$
B_n	$\mathfrak{so}(2n+1, \mathbb{C})$	$\mathfrak{so}(2n-1, \mathbb{C}) + \mathbb{C}$	$i = 1$
C_n	$\mathfrak{sp}(n, \mathbb{C})$	$\mathfrak{gl}(n, \mathbb{C})$	$i = n$
D_n	$\mathfrak{so}(2n, \mathbb{C})$	$\mathfrak{so}(2n-2, \mathbb{C}) + \mathbb{C}$	$i = 1$
	$\mathfrak{so}(2n, \mathbb{C})$	$\mathfrak{gl}(n, \mathbb{C})$	$i = n-1, n$
E_6	$\mathfrak{e}_6$	$\mathfrak{so}(10, \mathbb{C}) + \mathbb{C}$	$i = 1, 6$
E_7	$\mathfrak{e}_7$	$\mathfrak{e}_6 + \mathbb{C}$	$i = 1$

If $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{l}_{i\mathbb{C}})$ is a symmetric pair, then $[\mathfrak{n}_{i\mathbb{C}}^-, \mathfrak{n}_{i\mathbb{C}}^-] \subset \mathfrak{n}_{i\mathbb{C}}^- \cap \mathfrak{l}_{i\mathbb{C}} = \{0\}$, whence (ii) ⇒ (i). □

Definition 7.3.3. *We say the representation* $\pi_{k\omega_i}^{\mathfrak{g}_{\mathbb{C}}}$ $(k = 0, 1, 2, \dots)$ *is of* pan type*, or a pan representation if* $(\mathfrak{g}_{\mathbb{C}}, \alpha_i)$ *satisfies one of (therefore, all of) the equivalent conditions of Lemma 7.3.1. Here,* **pan** *stands for a* **p***arabolic subalgebra with* **a***belian* **n***ilradical.*

Table 7.3.2.



7.4 Example of pan representations

Example 7.4. Let $\mathfrak{g}_{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C})$ and $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. (This $\mathfrak{g}_{\mathbb{C}}$ is not a simple Lie algebra, but the above concept is defined similarly.) Then, π_λ is of pan type if and only if

$$\lambda_1 = \cdots = \lambda_i \geq \lambda_{i+1} = \cdots = \lambda_n,$$

for some i $(1 \leq i \leq n-1)$. Then, $(\mathfrak{l})_{i\mathbb{C}} \simeq \mathfrak{gl}(i, \mathbb{C}) + \mathfrak{gl}(n-i, \mathbb{C})$.

In particular, the kth symmetric tensor representations $S^k(\mathbb{C}^n)$ $(k \in \mathbb{N})$ and the kth exterior representations $\Lambda^k(\mathbb{C}^n)$ $(0 \leq k \leq n)$ are examples of pan representations since their highest weights are given by $(k, 0, \dots, 0)$ and $(\underbrace{1, \dots, 1}_{k}, \underbrace{0, \dots, 0}_{n-k})$, respectively.

S. Okada [68] studied branching laws for a specific class of irreducible finite-dimensional representations of classical Lie algebras, which he referred to as ‘rectangular-shaped representations’. The notion of ‘pan representations’ is equivalent to that of rectangular-shaped representations for type (A_n), (B_n), and (C_n). For type (D_n), $\pi_{k\omega_{n-1}}$, $\pi_{k\omega_n}$ $(k \in \mathbb{N})$ are rectangular-shaped representations, while $\pi_{k\omega_1}$ $(k \in \mathbb{N})$ are not.

7.5 Reduction to rank condition

Suppose $(\mathfrak{g}_{\mathbb{C}}, \alpha_i)$ satisfies the equivalent conditions in Lemma 7.3.1. Let θ be the complex involutive automorphism of the Lie algebra $\mathfrak{g}_{\mathbb{C}}$ that defines the symmetric pair $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{l}_{i\mathbb{C}})$. We use the same letter θ to denote the corresponding holomorphic involution of a simply connected $G_{\mathbb{C}}$. We take a maximal compact subgroup G_U of $G_{\mathbb{C}}$ such that $\theta G_U = G_U$. Then $K := G_U^{\theta} = G_U \cap L_{i\mathbb{C}}$ becomes a maximal compact subgroup of $L_{i\mathbb{C}}$.

Let τ be another complex involutive automorphism of $\mathfrak{g}_{\mathbb{C}}$, and $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}})$ the symmetric pair defined by τ. We also use the same letter τ to denote its lift to $G_{\mathbb{C}}$. We recall from Subsection 3.7 that the 'twisted' involution τ^g for $g \in G_{\mathbb{C}}$ is given by

$$\tau^g(x) = g\tau(g^{-1}xg)g^{-1} \quad (x \in G_{\mathbb{C}}) .$$

Lemma 7.5. *Let (θ, τ) be as above.*
1) *There exist an involutive automorphism σ of G_U and $g \in G_{\mathbb{C}}$ satisfying the following three conditions (by an abuse of notation, we write τ for τ^g):*

$$\tau\mathfrak{g}_U = \mathfrak{g}_U, \ \sigma\theta = \theta\sigma, \ \sigma\tau = \tau\sigma. \tag{7.5.1}$$

The induced action of σ on G_U/K is anti-holomorphic. (7.5.2)

$(\mathfrak{g}_U)^{\sigma,-\tau,-\theta}$ *contains a maximal abelian subspace in* $(\mathfrak{g}_U)^{-\tau,-\theta}$. (7.5.3)

2) *For any $x \in G_U/K$, there exists $h \in (G_U^{\tau})_0$ such that $\sigma(x) = h \cdot x$. In particular, each $(G_U^{\tau})_0$-orbit on G_U/K is preserved by σ.*

Proof. 1) See [50, Lemma 4.1] for the proof.
2) The second statement follows from the first statement and a similar argument of Lemma 3.3.

7.6 Proof of Theorem E

We are now ready to complete the proof of Theorem E in Section 1.

Let $\pi = \pi_{k\omega_i}^{\mathfrak{g}_{\mathbb{C}}}$ be a representation of pan type. As in Subsection 7.2, we consider the holomorphic line bundle $\mathcal{L}_{k\omega_i} \to G_{\mathbb{C}}/P_{i\mathbb{C}}^-$ and realize π on the space of holomorphic sections $\mathcal{O}(\mathcal{L}_{k\omega_i})$. We fix a G_U-invariant inner product on $\mathcal{O}(\mathcal{L}_{k\omega_i})$. With notation as in Subsection 7.5, we have a diffeomorphism

$$G_U/K \simeq G_{\mathbb{C}}/P_{i\mathbb{C}}^- ,$$

through which the holomorphic line bundle $\mathcal{L}_{k\omega_i} \to G_{\mathbb{C}}/P_{i\mathbb{C}}^-$ is naturally identified with the G_U-equivariant holomorphic line bundle $\mathcal{L} \to D$, where we set $\mathcal{L} := G_U \times_K \mathbb{C}_{k\omega_i}$ and $D := G_U/K$ (a compact Hermitian symmetric space).

Now, applying Lemma 7.5, we take σ and set $H := (G_U^{\tau})_0$. We note that the complexification of the Lie algebra of H is equal to $\mathfrak{h}_{\mathbb{C}}$ up to a conjugation by $G_{\mathbb{C}}$. By Lemma 7.5, the condition (2.2.3) in Theorem 2.2 is satisfied. Furthermore, we see the condition (2.2.2) holds by a similar argument of Lemma 9.4. Therefore, the restriction $\pi|_{(G_U)_0^{\tau}}$ is multiplicity-free by Theorem 2.2. Hence, Theorem E holds by Weyl's unitary trick. □

7.7 Proof of Theorem F

Suppose π_1 and π_2 are representations of pan type. We realize π_1 and π_2 on the space of holomorphic sections of holomorphic line bundles over compact symmetric spaces G_U/K_1 and G_U/K_2, respectively. We write θ_i for the corresponding involutive automorphisms of G_U that define K_i $(i = 1, 2)$. In light of Lemma 7.3.1 (iv), we can assume that $\theta_1\theta_2 = \theta_2\theta_1$. Then, applying Lemma 7.5 to (θ_1, θ_2) we find an involution $\sigma' \in \mathrm{Aut}(G_U)$ satisfying the following three conditions:

$\sigma'\theta_i = \theta_i\sigma'$ $(i = 1, 2)$. (7.7.1)

The induced action of σ' on G_U/K_i $(i = 1, 2)$ is anti-holomorphic. (7.7.2)

$(\mathfrak{g}_U)^{\sigma',-\theta_1,-\theta_2}$ contains a maximal abelian subspace of $(\mathfrak{g}_U)^{-\theta_1,-\theta_2}$. (7.7.3)

We remark that the condition (7.7.2) for $i = 2$ is not included in Lemma 7.5, but follows automatically by our choice of σ.

We define three involutive automorphisms τ, θ and σ on $G_U \times G_U$ by $\tau(g_1, g_2) := (g_2, g_1)$, $\theta := (\theta_1, \theta_2)$ and $\sigma := (\sigma', \sigma')$, respectively. Then $(G_U \times G_U)^\tau = \mathrm{diag}(G_U)$. By using the identification

$$(\mathfrak{g}_U \oplus \mathfrak{g}_U)^{-\tau} = \{(X, -X) : X \in \mathfrak{g}_U\} \xrightarrow{\sim} \mathfrak{g}_U\,, \quad (X, -X) \mapsto X\,,$$

we have isomorphisms

$$(\mathfrak{g}_U \oplus \mathfrak{g}_U)^{-\tau,-\theta} \simeq (\mathfrak{g}_U)^{-\theta_1,-\theta_2}\,,$$
$$(\mathfrak{g}_U \oplus \mathfrak{g}_U)^{\sigma,-\tau,-\theta} \simeq (\mathfrak{g}_U)^{\sigma',-\theta_1,-\theta_2}\,.$$

Thus, the condition (7.7.3) implies that $(\mathfrak{g}_U \oplus \mathfrak{g}_U)^{\sigma,-\tau,-\theta}$ contains a maximal abelian subspace of $(\mathfrak{g}_U \oplus \mathfrak{g}_U)^{-\tau,-\theta}$. Then, by Lemma 7.5 and by a similar argument of Lemma 3.3 again, for any $(x, y) \in G_U/K_1 \times G_U/K_2$ there exists a $g \in G_U$ such that $\sigma'(x) = g \cdot x$ and $\sigma'(y) = g \cdot y$ simultaneously. Now, Theorem F follows readily from Theorem 2.2. □

7.8 List of multiplicity-free restrictions

For the convenience of the reader, we present the list (see Table 7.8.1) of the triple $(\mathfrak{g}_\mathbb{C}, \mathfrak{h}_\mathbb{C}, i)$ for which we can conclude from Theorem E that the irreducible finite-dimensional representation $\pi^{\mathfrak{g}_\mathbb{C}}_{k\omega_i}$ of a simple Lie algebra $\mathfrak{g}_\mathbb{C}$ is multiplicity-free when restricted to $\mathfrak{h}_\mathbb{C}$ for any $k \in \mathbb{N}$ by Theorem E.

Some of the above cases were previously known to be multiplicity-free by case-by-case argument, in particular, for the case $\mathrm{rank}\,\mathfrak{g}_\mathbb{C} = \mathrm{rank}\,\mathfrak{h}_\mathbb{C}$. Among them, the corresponding explicit branching laws have been studied by S. Okada [68] and H. Alikawa [1].

There are some few representations π that are not of pan type, but are multiplicity-free when restricted to symmetric subgroups H. Our method still works to capture such cases, but we do not go into details here (see [46, 51, 52]).

Table 7.8.1.

$\mathfrak{g}_{\mathbb{C}}$	$\mathfrak{h}_{\mathbb{C}}$	i
$\mathfrak{sl}(n+1,\mathbb{C})$	$\mathfrak{sl}(p,\mathbb{C})+\mathfrak{sl}(n+1-p,\mathbb{C})+\mathbb{C}$	$1,2,\dots,n$
$\mathfrak{sl}(n+1,\mathbb{C})$	$\mathfrak{so}(n+1,\mathbb{C})$	$1,2,\dots,n$
$\mathfrak{sl}(2m,\mathbb{C})$	$\mathfrak{sp}(m,\mathbb{C})$	$1,2,\dots,2m-1$
$\mathfrak{so}(2n+1,\mathbb{C})$	$\mathfrak{so}(p,\mathbb{C})+\mathfrak{so}(2n+1-p,\mathbb{C})$	1
$\mathfrak{sp}(n,\mathbb{C})$	$\mathfrak{sp}(p,\mathbb{C})+\mathfrak{sp}(n-p,\mathbb{C})$	n
$\mathfrak{sp}(n,\mathbb{C})$	$\mathfrak{gl}(n,\mathbb{C})$	n
$\mathfrak{so}(2n,\mathbb{C})$	$\mathfrak{so}(p,\mathbb{C})+\mathfrak{so}(2n-p,\mathbb{C})$	$1,n-1,n$
$\mathfrak{so}(2n,\mathbb{C})$	$\mathfrak{gl}(n,\mathbb{C})$	$1,n-1,n$
$\mathfrak{e}_6$	$\mathfrak{so}(10,\mathbb{C})+\mathfrak{so}(2,\mathbb{C})$	$1,6$
$\mathfrak{e}_6$	$\mathfrak{sl}(6,\mathbb{C})+\mathfrak{sl}(2,\mathbb{C})$	$1,6$
$\mathfrak{e}_6$	$\mathfrak{f}_4$	$1,6$
$\mathfrak{e}_6$	$\mathfrak{sp}(4,\mathbb{C})$	$1,6$
$\mathfrak{e}_7$	$\mathfrak{e}_6+\mathfrak{so}(2,\mathbb{C})$	7
$\mathfrak{e}_7$	$\mathfrak{so}(12,\mathbb{C})+\mathfrak{sl}(2,\mathbb{C})$	7
$\mathfrak{e}_7$	$\mathfrak{sl}(8,\mathbb{C})$	7

8 Generalization of the Hua–Kostant–Schmid formula

This section discusses an explicit irreducible decomposition formula of the restriction $\pi|_H$ where the triple (π, G, H) satisfies the following two conditions:

1) π is a holomorphic discrete series representation of scalar type (Definition 1.3).

2) (G, H) is a symmetric pair defined by an involution τ of holomorphic type (Definition 1.4).

We know a priori from Theorem B (1) that the branching law is discrete and multiplicity-free. The main result of this section is Theorem 8.3, which enriches this abstract property with an explicit multiplicity-free formula. The formula for the special case $H = K$ corresponds to the Hua–Kostant–Schmid formula ([26, 32, 78]). We also present explicit formulas for the irreducible decomposition of the tensor product representation (Theorem 8.4) and of the restriction $U(p,q) \downarrow U(p-1,q)$ (Theorem 8.11).

Let us give a few comments on our proof of Theorem 8.3. Algebraically, our key machinery is Lemma 8.7 which assures that the irreducible G-decomposition is determined only by its K-structure. Geometrically, a well-known method of taking normal derivatives (e.g., S. Martens [63], Jakobsen–Vergne [31]) gives a general algorithm to obtain branching laws for highest weight modules. This algorithm yields explicit formulae by using the observation that the fiber of the normal bundle for $G^\tau/K^\tau \subset G/K$ is the tangent space of another Hermitian symmetric space $G^{\tau\theta}/K^\tau$. The key ingredient of the geometry here is the following nice properties of the two symmetric pairs (G, G^τ) and $(G, G^{\tau\theta})$:

a) $K \cap G^\tau = K \cap G^{\tau\theta}$,

b) $\mathfrak{p} = (\mathfrak{p} \cap \mathfrak{g}^\tau) \oplus (\mathfrak{p} \cap \mathfrak{g}^{\tau\theta})$.

Unless otherwise mentioned, we shall assume H is connected, that is, $H = G_0^\tau$ throughout this section.

8.1 Notation for highest weight modules

We set up the notation and give a parametrization of irreducible highest weight modules for both finite and infinite-dimensional cases.

First, we consider finite-dimensional representations. Let us take a Cartan subalgebra $\mathfrak{t}$ of a reductive Lie algebra $\mathfrak{k}$ and fix a positive system $\Delta^+(\mathfrak{k}, \mathfrak{t})$. We denote by $\pi_\mu^{\mathfrak{k}}$ the irreducible finite-dimensional representation of $\mathfrak{k}$ with highest weight μ, if μ is a dominant integral weight. A $\mathfrak{k}$-module $\pi_\mu^{\mathfrak{k}}$ will be written also as π_μ^K if the action lifts to K.

Next, let G be a connected reductive Lie group, θ a Cartan involution, $K = \{g \in G : \theta g = g\}$, $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ the corresponding Cartan decomposition and $\mathfrak{g}_{\mathbb{C}} = \mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_{\mathbb{C}}$ its complexification. We assume that there exists a central element Z of $\mathfrak{k}$ such that

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_+ + \mathfrak{p}_- \tag{8.1.1}$$

is the eigenspace decomposition of $\frac{1}{\sqrt{-1}} \operatorname{ad}(Z)$ with eigenvalues 0, 1, and -1, respectively. This assumption is satisfied if and only if G is locally isomorphic to a direct product of connected compact Lie groups and non-compact Lie groups of Hermitian type (if G is compact, we can simply take $Z = 0$).

We set

$$\widetilde{Z} := \frac{1}{\sqrt{-1}} Z\,. \tag{8.1.2}$$

As in Definition 1.3, we say an irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module V is a *highest weight module* if

$$V^{\mathfrak{p}_+} = \{v \in V : Yv = 0 \quad \text{for all } Y \in \mathfrak{p}_+\}$$

is non-zero. Then, $V^{\mathfrak{p}_+}$ is irreducible as a K-module, and the $(\mathfrak{g}_{\mathbb{C}}, K)$-module V is determined uniquely by the K-structure on $V^{\mathfrak{p}_+}$. If μ is the highest weight of $V^{\mathfrak{p}_+}$, we write V as $\pi_\mu^{\mathfrak{g}}$. That is, the irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module $\pi_\mu^{\mathfrak{g}}$ is characterized by the K-isomorphism

$$(\pi_\mu^{\mathfrak{g}})^{\mathfrak{p}_+} \simeq \pi_\mu^{\mathfrak{k}}\,. \tag{8.1.3}$$

An irreducible unitary highest weight representation π of G will be denoted by π_μ^G if the underlying $(\mathfrak{g}_{\mathbb{C}}, K)$-module of π is isomorphic to $\pi_\mu^{\mathfrak{g}}$. Let Λ_G be the totality of μ such that $\pi_\mu^{\mathfrak{g}}$ lifts to an irreducible unitary representation of G. For simply connected G, irreducible unitary highest weight representations were classified, that is, the set Λ_G ($\subset \sqrt{-1}\mathfrak{t}^*$) was explicitly found in [12] and [30] (see also [13]). In particular, we recall from [12] that

$$\lambda(\widetilde{Z}) \in \mathbb{R} \quad \text{for any } \lambda \in \Lambda_G$$

and

$$c_G := \sup_{\lambda \in \Lambda_G} \lambda(\widetilde{Z}) < \infty \tag{8.1.4}$$

if G is semisimple.

The highest weight module $\pi_\mu^{\mathfrak{g}}$ is the unique quotient of the generalized Verma module

$$N^{\mathfrak{g}}(\mu) := U(\mathfrak{g}_{\mathbb{C}}) \otimes_{U(\mathfrak{k}_{\mathbb{C}}+\mathfrak{p}_+)} \pi_\mu^{\mathfrak{k}}, \tag{8.1.5}$$

where $\pi_\mu^{\mathfrak{k}}$ is regarded as a module of the maximal parabolic subalgebra $\mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_+$ by making $\mathfrak{p}_+$ act trivially. Furthermore, $\pi_\mu^{\mathfrak{g}}$ has a $Z(\mathfrak{g}_{\mathbb{C}})$-infinitesimal character $\mu + \rho_{\mathfrak{g}} \in \mathfrak{t}_{\mathbb{C}}^*$ via the Harish-Chandra isomorphism

$$\mathrm{Hom}_{\mathbb{C}\text{-algebra}}(Z(\mathfrak{g}_{\mathbb{C}}), \mathbb{C}) \simeq \mathfrak{t}_{\mathbb{C}}^*/W,$$

where $Z(\mathfrak{g}_{\mathbb{C}})$ is the center of the enveloping algebra $U(\mathfrak{g}_{\mathbb{C}})$, W is the Weyl group of the root system $\Delta(\mathfrak{g}, \mathfrak{t})$, and $\rho_{\mathfrak{g}}$ is half the sum of positive roots $\Delta^+(\mathfrak{g}, \mathfrak{t}) := \Delta^+(\mathfrak{k}, \mathfrak{t}) \cup \Delta(\mathfrak{p}_+, \mathfrak{t})$.

8.2 Strongly orthogonal roots

Let G be a non-compact simple Lie group of Hermitian type, and τ an involution of holomorphic type which commutes with the Cartan involution θ.

We take a Cartan subalgebra $\mathfrak{t}^\tau$ of the reductive Lie algebra

$$\mathfrak{k}^\tau := \{X \in \mathfrak{k} : \tau X = X\}$$

and extend it to a Cartan subalgebra $\mathfrak{t}$ of $\mathfrak{k}$. We note that $\mathfrak{t}^\tau = \mathfrak{k}^\tau \cap \mathfrak{t}$. The pair $(\mathfrak{k}, \mathfrak{k}^\tau)$ forms a reductive symmetric pair, and $\mathfrak{t}$ plays an analogous role to the fundamental Cartan subalgebra with respect to this symmetric pair. Thus, using the same argument as in [84], we see that if $\alpha \in \Delta(\mathfrak{k}, \mathfrak{t})$ satisfies $\alpha|_{\mathfrak{t}^\tau} = 0$ then $\alpha = 0$. Thus, we can take positive systems $\Delta^+(\mathfrak{k}, \mathfrak{t})$ and $\Delta^+(\mathfrak{k}^\tau, \mathfrak{t}^\tau)$ in a compatible way such that

$$\alpha|_{\mathfrak{t}^\tau} \in \Delta^+(\mathfrak{k}^\tau, \mathfrak{t}^\tau) \quad \text{if } \alpha \in \Delta^+(\mathfrak{k}, \mathfrak{t}). \tag{8.2.1}$$

Since τ is of holomorphic type, we have $\tau Z = Z$, and therefore $\tau\mathfrak{p}_+ = \mathfrak{p}_+$. Hence, we have a direct sum decomposition $\mathfrak{p}_+ = \mathfrak{p}_+^\tau \oplus \mathfrak{p}_+^{-\tau}$, where we set

$$\mathfrak{p}_+^{\pm\tau} := \{X \in \mathfrak{p}_+ : \tau X = \pm X\}.$$

Let us consider the reductive subalgebra $\mathfrak{g}^{\tau\theta}$. Its Cartan decomposition is given by

$$\mathfrak{g}^{\tau\theta} = (\mathfrak{g}^{\tau\theta} \cap \mathfrak{g}^\theta) + (\mathfrak{g}^{\tau\theta} \cap \mathfrak{g}^{-\theta}) = \mathfrak{k}^\tau + \mathfrak{p}^{-\tau},$$

and its complexification is given by

$$\mathfrak{g}_{\mathbb{C}}^{\tau\theta} = \mathfrak{k}_{\mathbb{C}}^\tau \oplus \mathfrak{p}_+^{-\tau} \oplus \mathfrak{p}_-^{-\tau}. \tag{8.2.2}$$

The Cartan subalgebra $\mathfrak{t}^\tau$ of $\mathfrak{k}^\tau$ is also a Cartan subalgebra of $\mathfrak{g}^{\tau\theta}$.

Let $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau)$ be the set of weights of $\mathfrak{p}_+^{-\tau}$ with respect to $\mathfrak{t}^\tau$. The roots α and β are said to be *strongly orthogonal* if neither $\alpha + \beta$ nor $\alpha - \beta$ is a root. We take a maximal set of strongly orthogonal roots $\{\nu_1, \nu_2, \dots, \nu_l\}$ in $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau)$ such that

i) ν_1 is the lowest root among the elements in $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau)$,

ii) ν_{j+1} is the lowest root among the elements in $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau)$ that are strongly orthogonal to $\nu_1, \dots, \nu_j$.

A special case applied to $\tau = \theta$ shows $\mathfrak{k}^\tau = \mathfrak{k}$, $\mathfrak{t}^\tau = \mathfrak{t}$, $\mathfrak{p}^{-\tau} = \mathfrak{p}$, and $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau) = \Delta(\mathfrak{p}_+, \mathfrak{t})$. In this case, we shall use the notation $\{\bar{\nu}_1, \bar{\nu}_2, \dots, \bar{\nu}_{\bar{l}}\}$ for a maximal set of strongly orthogonal roots in $\Delta(\mathfrak{p}_+, \mathfrak{t})$ such that

i) $\bar{\nu}_1$ is the lowest root among $\Delta(\mathfrak{p}_+, \mathfrak{t})$,

ii) $\bar{\nu}_{j+1}$ is the lowest root among the elements in $\Delta(\mathfrak{p}_+, \mathfrak{t})$ that are strongly orthogonal to $\bar{\nu}_1, \dots, \bar{\nu}_j$ $(1 \le j \le \bar{l})$.

Then, $\bar{l} = \mathbb{R}\text{-}\operatorname{rank} \mathfrak{g}$ by [57]. Likewise, in light of (8.2.2) for the Hermitian symmetric space $G^{\tau\theta}/G^{\tau\theta} \cap K = G^{\tau\theta}/G^{\tau,\theta}$, we have $l = \mathbb{R}\text{-}\operatorname{rank} \mathfrak{g}^{\tau\theta}$. In general, $l \le \bar{l}$.

8.3 Branching laws for semisimple symmetric pairs

It follows from (8.1.3) that the highest weight module $\pi_\mu^{\mathfrak{g}}$ is of scalar type, namely, $(\pi_\mu^{\mathfrak{g}})^{\mathfrak{p}_+}$ is one-dimensional, if and only if

$$\langle \mu, \alpha \rangle = 0 \qquad \text{for any } \alpha \in \Delta(\mathfrak{k}, \mathfrak{t}). \tag{8.3.1}$$

Furthermore, the representation π_μ^G is a (relative) holomorphic discrete series representation of G if and only if

$$\langle \mu + \rho_{\mathfrak{g}}, \alpha \rangle < 0 \qquad \text{for any } \alpha \in \Delta(\mathfrak{p}_+, \mathfrak{t}). \tag{8.3.2}$$

We are now ready to state the branching law of holomorphic discrete series representations π_μ^G of scalar type with respect to semisimple symmetric pairs (G, H):

Theorem 8.3. *Let G be a non-compact simple Lie group of Hermitian type. Assume that $\mu \in \sqrt{-1}\,\mathfrak{t}^*$ satisfies* (8.3.1) *and* (8.3.2)*. Let τ be an involutive automorphism of G of holomorphic type, $H = G_0^\tau$ (the identity component of G^τ), and $\{\nu_1, \dots, \nu_l\}$ be the set of strongly orthogonal roots in $\Delta(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau)$ as in Subsection 8.2. Then, π_μ^G decomposes discretely into a multiplicity-free sum of irreducible H-modules:*

$$\pi_\mu^G|_H \simeq \sum_{\substack{a_1 \ge \cdots \ge a_l \ge 0 \\ a_1, \dots, a_l \in \mathbb{N}}}^{\oplus} \pi^H_{\mu|_{\mathfrak{t}^\tau} - \sum_{j=1}^{l} a_j \nu_j} \qquad \textit{(discrete Hilbert sum)}. \tag{8.3.3}$$

The formula for the case $H = K$ (that is, $\tau = \theta$) was previously found by L.-K. Hua (implicit in the classical case), B. Kostant (unpublished) and W. Schmid [78] (see also Johnson [32] for an algebraic proof). In this case, each summand in the right side is finite dimensional.

For $\tau \neq \theta$, some special cases have been also studied by H. Jakobsen, M. Vergne, J. Xie, W. Bertram and J. Hilgert [7, 30, 31, 89]. Further, quantitative results by means of reproducing kernels are obtained in [5]. The formula (8.3.3) in the above generality was first given by the author [39].

We shall give a proof of Theorem 8.3 in Subsection 8.8.

8.4 Irreducible decomposition of tensor products

As we saw in Example 3.2.1, the pair $(G \times G, \operatorname{diag}(G))$ forms a symmetric pair. Correspondingly, the tensor product representation can be regarded as a special (and easy) case of restrictions of representations with respect to symmetric pairs. This subsection provides a decomposition formula of the tensor product of two holomorphic discrete series representations of scalar type. This is regarded as a counterpart of Theorem 8.3 for tensor product representations.

We recall from Subsection 8.2 that $\{\bar{\nu}_1, \dots, \bar{\nu}_{\bar{l}}\}$ is a maximal set of strongly orthogonal roots in $\Delta(\mathfrak{p}_+, \mathfrak{t})$ and $\bar{l} = \mathbb{R}\text{-}\operatorname{rank}\mathfrak{g}$.

Theorem 8.4. *Let G be a non-compact simple Lie group of Hermitian type. Assume that $\mu_1, \mu_2 \in \sqrt{-1}\,\mathfrak{t}^*$ satisfy the conditions* (8.3.1) *and* (8.3.2)*. Then, the tensor product representation $\pi^G_{\mu_1} \widehat{\otimes} \pi^G_{\mu_2}$ decomposes discretely into a multiplicity-free sum of irreducible G-modules:*

$$\pi^G_{\mu_1} \widehat{\otimes} \pi^G_{\mu_2} \simeq \sum_{\substack{a_1 \geq \cdots \geq a_{\bar{l}} \geq 0 \\ a_1, \dots, a_{\bar{l}} \in \mathbb{N}}} \pi^G_{\mu_1 + \mu_2 - \sum_{j=1}^{\bar{l}} a_j \bar{\nu}_j} .$$

The proof of Theorem 8.4 will be given in Subsection 8.9.

8.5 Eigenvalues of the central element Z

Our proof of Theorems 8.3 and 8.4 depends on the algebraic lemma that the K-type formula determines the irreducible decomposition of the whole group (see Lemma 8.7). This is a very strong assertion, which fails in general for non-highest weight modules. This subsection collects some nice properties peculiar to highest weight modules that will be used in the proof of Lemma 8.7.

For a K-module V, we define a subset of $\mathbb{C}$ by

$$\operatorname{Spec}_{\widetilde{Z}}(V) := \{\text{eigenvalues of } \widetilde{Z} \text{ on } V\},$$

where we set

$$\widetilde{Z} := \frac{1}{\sqrt{-1}} Z .$$

For instance, $\operatorname{Spec}_{\widetilde{Z}}(V)$ is a singleton if V is an irreducible K-module. We also note that $\operatorname{Spec}_{\widetilde{Z}}(\mathfrak{g}_{\mathbb{C}}) = \{0, \pm 1\}$ by (8.1.1).

Lemma 8.5. *Suppose V is an irreducible $(\mathfrak{g}_{\mathbb{C}}, K)$-module. Then,*
1) $\operatorname{Spec}_{\widetilde{Z}}(V) \subset a_0 + \mathbb{Z}$ *for some* $a_0 \in \mathbb{C}$.
2) *If* $\sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V) < \infty$*, then V is a highest weight module.*
3) *If V is a highest weight module $\pi^{\mathfrak{g}}_{\lambda}$, then* $\operatorname{Spec}_{\widetilde{Z}}(V) \subset -\mathbb{N} + \lambda(\widetilde{Z})$ *and* $\sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V) = \operatorname{Re} \lambda(\widetilde{Z})$.
4) *If V is a unitary highest weight module, then* $\operatorname{Spec}_{\widetilde{Z}}(V) \subset (-\infty, c_G]$*, where c_G is a constant depending on G.*
5) *If both V and F are highest weight modules of finite length, then any irreducible subquotient W of $V \otimes F$ is also a highest weight module.*

Proof. 1) For $a \in \mathbb{C}$, we write the eigenspace of $\widetilde{Z}$ as $V_a := \{v \in V : \widetilde{Z}v = av\}$. Then, it follows from the Leibniz rule that

$$\mathfrak{p}_+ V_a \subset V_{a+1}\,, \quad \mathfrak{k}_{\mathbb{C}} V_a \subset V_a\,, \quad \text{and} \quad \mathfrak{p}_- V_a \subset V_{a-1}\,.$$

An iteration of this argument shows that

$$\operatorname{Spec}_{\widetilde{Z}}(U(\mathfrak{g}_{\mathbb{C}})V_a) \subset a + \mathbb{Z}\,.$$

Now we take a_0 such that $V_{a_0} \neq \{0\}$. Since V is irreducible, we have $V = U(\mathfrak{g}_{\mathbb{C}})V_{a_0}$, and therefore $\operatorname{Spec}_{\widetilde{Z}}(V) \subset a_0 + \mathbb{Z}$.
2) Suppose $\sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V) < \infty$. Since $\operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V)$ is discrete by (1), there exists $a \in \operatorname{Spec}_{\widetilde{Z}}(V)$ such that $\operatorname{Re} a$ attains its maximum. Then

$$\mathfrak{p}_+ V_a \subset V_{a+1} = \{0\}\,.$$

Thus, $V_a \subset V^{\mathfrak{p}_+}$. Hence, V is a highest weight module.
3) The highest weight module $\pi_\lambda^{\mathfrak{g}}$ is isomorphic to the unique irreducible quotient of the generalized Verma module $N^{\mathfrak{g}}(\lambda) = U(\mathfrak{g}_{\mathbb{C}}) \otimes_{U(\mathfrak{k}_{\mathbb{C}}+\mathfrak{p}_+)} \pi_\lambda^{\mathfrak{k}}$. By the Poincaré–Birkhoff–Witt theorem, $N^{\mathfrak{g}}(\lambda)$ is isomorphic to $S(\mathfrak{p}_-) \otimes \pi_\lambda^{\mathfrak{k}}$ as a $\mathfrak{k}$-module. Thus, any $\mathfrak{k}$-type $\pi_\mu^{\mathfrak{k}}$ occurring in $\pi_\lambda^{\mathfrak{g}}$ is of the form

$$\mu = \lambda + \sum_{\alpha \in \Delta(\mathfrak{p}_-, \mathfrak{t})} m_\alpha \alpha,$$

for some $m_\alpha \in \mathbb{N}$. As $\alpha(\widetilde{Z}) = -1$ for any $\alpha \in \Delta(\mathfrak{p}_-, \mathfrak{t})$, we have

$$\mu(\widetilde{Z}) = \lambda(\widetilde{Z}) - \sum_{\alpha \in \Delta(\mathfrak{p}_-, \mathfrak{t})} m_\alpha\,. \tag{8.5.1}$$

In particular, we have the following equivalence:

$$\operatorname{Re} \mu(\widetilde{Z}) = \operatorname{Re} \lambda(\widetilde{Z}) \iff \mu = \lambda\,, \tag{8.5.2}$$

and we also have

$$\operatorname{Spec}_{\widetilde{Z}}(\pi_\lambda^{\mathfrak{g}}) \subset \{\lambda(\widetilde{Z}) - \sum_{\alpha \in \Delta(\mathfrak{p}_-, \mathfrak{t})} m_\alpha : m_\alpha \in \mathbb{N}\} = -\mathbb{N} + \lambda(\widetilde{Z})\,. \tag{8.5.3}$$

Furthermore, since the $\mathfrak{k}$-type $\pi_\lambda^{\mathfrak{k}}$ occurs in $\pi_\lambda^{\mathfrak{g}}$, we have $\lambda(\widetilde{Z}) \in \operatorname{Spec}_{\widetilde{Z}}(\pi_\lambda^{\mathfrak{g}})$. Here, $\sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(\pi_\lambda^{\mathfrak{g}}) = \operatorname{Re} \lambda(\widetilde{Z})$.
4) This statement follows from (8.1.4) and from (3).
5) For two subsets A and B in $\mathbb{C}$, we write $A + B := \{a + b \in \mathbb{C} : a \in A, b \in B\}$. Then, $\operatorname{Spec}_{\widetilde{Z}}(V \otimes F) \subset \operatorname{Spec}_{\widetilde{Z}}(V) + \operatorname{Spec}_{\widetilde{Z}}(F)$. Therefore,

$$\begin{aligned} \sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(W) &\le \sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V) \\ &\le \sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(V) + \sup \operatorname{Re} \operatorname{Spec}_{\widetilde{Z}}(F) < \infty. \end{aligned}$$

Hence, W is also a highest weight module by (2). $\square$

8.6 Bottom layer map

The following lemma finds an irreducible summand ('bottom layer') from the K-type structure.

Lemma 8.6. *Let V be a $(\mathfrak{g}_{\mathbb{C}}, K)$-module. We assume that V decomposes into an algebraic direct sum of (possibly, infinitely many) irreducible highest weight modules. We set*

$$\mathrm{Supp}_{\mathfrak{k}}(V) := \{\mu \in \sqrt{-1}\mathfrak{t}^* : \mathrm{Hom}_{\mathfrak{k}}(\pi^{\mathfrak{k}}_{\mu}, V) \neq \{0\}\}.$$

If the evaluation map

$$\mathrm{Supp}_{\mathfrak{k}}(V) \to \mathbb{R}, \quad \mu \mapsto \mathrm{Re}\, \mu(\widetilde{Z})$$

attains its maximum at μ_0, then

$$\mathrm{Hom}_{(\mathfrak{g}_{\mathbb{C}},K)}(\pi^{\mathfrak{g}}_{\mu_0}, V) \neq \{0\}.$$

Proof. Take a non-zero map $q \in \mathrm{Hom}_{\mathfrak{k}}(\pi^{\mathfrak{k}}_{\mu_0}, V)$. As V is an algebraic direct sum of irreducible highest weight modules, there exists a projection $p : V \to \pi^{\mathfrak{g}}_{\lambda}$ for some λ such that $p \circ q \neq 0$. This means that $\pi^{\mathfrak{k}}_{\mu_0}$ occurs in $\pi^{\mathfrak{g}}_{\lambda}$, and therefore we have

$$\mathrm{Re}\, \mu_0(\widetilde{Z}) \leq \sup \mathrm{Re}\, \mathrm{Spec}_{\widetilde{Z}}(\pi^{\mathfrak{g}}_{\lambda}) = \mathrm{Re}\, \lambda(\widetilde{Z}).$$

Here, the last equality is by Lemma 8.5 (3).

Conversely, the maximality of μ_0 implies that $\mathrm{Re}\, \mu_0(\widetilde{Z}) \geq \mathrm{Re}\, \lambda(\widetilde{Z})$. Hence, $\mathrm{Re}\, \mu_0(\widetilde{Z}) = \mathrm{Re}\, \lambda(\widetilde{Z})$, and we have then $\mu_0 = \lambda$ by (8.5.2). Since $\pi^{\mathfrak{g}}_{\lambda}$ is an irreducible summand of V, we have $\mathrm{Hom}_{(\mathfrak{g}_{\mathbb{C}},K)}(\pi^{\mathfrak{g}}_{\mu_0}, V) \neq \{0\}$. □

8.7 Determination of the $\mathfrak{g}_{\mathbb{C}}$-structure by K-types

In general, the K-type formula is not sufficient to determine the irreducible decomposition of a unitary representation even in the discretely decomposable case. However, this is the case if any irreducible summand is a highest weight module. Here is the statement that we shall use as a main machinery of the proof of Theorems 8.3 and 8.4.

Lemma 8.7. *Suppose $(\pi, \mathcal{H})$ is a K-admissible unitary representation of G, which splits discretely into a Hilbert direct sum of irreducible unitary highest weight representations of G. Let $\mathcal{H}_K$ be the space of K-finite vectors of $\mathcal{H}$. Assume that there exists a function $n_\pi : \mathfrak{t}^*_{\mathbb{C}} \to \mathbb{N}$ such that $\mathcal{H}_K$ is isomorphic to the following direct sum as $\mathfrak{k}$-modules:*

$$\mathcal{H}_K \simeq \bigoplus_{\lambda} n_\pi(\lambda) \pi^{\mathfrak{g}}_{\lambda} \quad \textit{(algebraic direct sum)}. \tag{8.7.1}$$

Then, $n_\pi(\lambda) \neq 0$ only if $\lambda \in \Lambda_G$, that is, $\pi^{\mathfrak{g}}_{\lambda}$ lifts to an irreducible unitary representation π^G_{λ} of G. Furthermore, the identity (8.7.1) *holds as a $(\mathfrak{g}_{\mathbb{C}}, K)$-module* *isomorphism, and the unitary representation* π *has the following decomposition into irreducible unitary representations of* G*:*

$$\pi \simeq \sum_{\lambda}^{\oplus} n_{\pi}(\lambda)\pi_{\lambda}^{G} \quad \textit{(discrete Hilbert sum).} \tag{8.7.2}$$

Proof. We write an abstract irreducible decomposition of $\mathcal{H}$ as

$$\mathcal{H} \simeq \sum_{\lambda \in \Lambda_G}^{\oplus} m_{\lambda}\pi_{\lambda}^{G} \quad \text{(discrete Hilbert sum).}$$

Since $\mathcal{H}$ is K-admissible, the multiplicity $m_\lambda < \infty$ for all λ, and we have an isomorphism of $(\mathfrak{g}_{\mathbb{C}}, K)$-modules with the same multiplicity m_λ (see [43, Theorem 2.7]):

$$\mathcal{H}_K \simeq \bigoplus_{\lambda \in \Lambda_G} m_{\lambda}\pi_{\lambda}^{\mathfrak{g}} \quad \text{(algebraic direct sum).} \tag{8.7.3}$$

Let us show $n_\pi(\lambda) = m_\lambda$ for all λ. For this, we begin with an observation that

$$\mathrm{Spec}_{\widetilde{Z}}(\mathcal{H}_K) = \bigcup_{\lambda | m_\lambda \neq 0} \mathrm{Spec}_{\widetilde{Z}}(\pi_{\lambda}^{\mathfrak{g}})$$

is a subset in $\mathbb{R}$ and has an upper bound. This follows from Lemma 8.5 (4) applied to each irreducible summand in (8.7.3).

First, we consider the case where there exists $a \in \mathbb{R}$ such that

$$\lambda(\widetilde{Z}) \equiv a \bmod \mathbb{Z} \quad \text{for any } \lambda \text{ satisfying } n_\pi(\lambda) \neq 0. \tag{8.7.4}$$

Then, the set

$$\{\lambda(\widetilde{Z}) : \lambda \in \mathfrak{t}_{\mathbb{C}}^{*}, n_\pi(\lambda) \neq 0\} \tag{8.7.5}$$

is contained in $\mathrm{Spec}_{\widetilde{Z}}(\mathcal{H}_K)$ by (8.7.1), and is discrete by (8.7.4). Hence, it is a discrete subset of $\mathbb{R}$ with an upper bound. Thus, we can find $\mu_0 \in \mathfrak{t}_{\mathbb{C}}^{*}$ such that $n_\pi(\mu_0) \neq 0$ and that $\mu_0(\widetilde{Z})$ attains its maximum in (8.7.5). In turn, the evaluation map $\mathrm{Supp}_{\mathfrak{k}}(\mathcal{H}_K) \to \mathbb{R}$, $\mu \mapsto \mu(\widetilde{Z})$ attains its maximum at $\mu_0 \in \mathrm{Supp}_{\mathfrak{k}}(\mathcal{H}_K)$ by (8.7.1) and Lemma 8.5 (3). Therefore, $\mathrm{Hom}_{(\mathfrak{g}_{\mathbb{C}}, K)}(\pi_{\mu_0}^{\mathfrak{g}}, \mathcal{H}_K) \neq \{0\}$ by Lemma 8.6. Thus, we have shown $m_{\mu_0} \neq 0$, that is, $\pi_{\mu_0}^{G}$ occurs as a subrepresentation in $\mathcal{H}$.

Next, we consider the unitary representation π' on

$$\mathcal{H}' := \sum_{\lambda \neq \mu_0}^{\oplus} m_\lambda \pi_{\lambda}^{G} \oplus (m_{\mu_0} - 1)\pi_{\mu_0}^{G},$$

the orthogonal complement of a subrepresentation $\pi_{\mu_0}^{G}$ in $\mathcal{H}$. Then, the K-type formula (8.7.1) for $(\pi', \mathcal{H}')$ holds if we set

$$n_{\pi'}(\lambda) := \begin{cases} n_\pi(\lambda) - 1 & (\lambda = \mu_0), \\ n_\pi(\lambda) & (\lambda \neq \mu_0). \end{cases}$$

Hence, by the downward induction on $\sup \mathrm{Spec}_{\widetilde{Z}}(\mathcal{H}_K)$, we have $n_\pi(\lambda) = m_\lambda$ for all λ.

For the general case, let A be the set of complete representatives of $\{\lambda(\widetilde{Z}) \in \mathbb{C} \bmod \mathbb{Z} : n_\pi(\lambda) \neq 0\}$. For each $a \in A$, we define a subrepresentation $\mathcal{H}_a$ of $\mathcal{H}$ by

$$\mathcal{H}_a := \sum_{\lambda(\widetilde{Z}) \equiv a \bmod \mathbb{Z}}^{\oplus} m_\lambda \pi_\lambda^G \quad \text{(discrete Hilbert sum)}.$$

Then, we have an isomorphism of unitary representations of G:

$$\mathcal{H} \simeq \sum_{a \in A}^{\oplus} \mathcal{H}_a \,.$$

Since $\mathrm{Spec}_{\widetilde{Z}}(\pi_\lambda^{\mathfrak{g}}) \subset a + \mathbb{Z}$ if and only if $\lambda(\widetilde{Z}) \equiv a \bmod \mathbb{Z}$ by Lemma 8.5 (3), we get from (8.7.1) the following K-isomorphism

$$(\mathcal{H}_a)_K \simeq \bigoplus_{\lambda(\widetilde{Z}) \equiv a \bmod \mathbb{Z}} n_\pi(\lambda) \pi_\lambda^{\mathfrak{g}}, \tag{8.7.6}$$

for each $a \in A$. Therefore, our proof for the first step assures $n_\pi(\lambda) = m_\lambda$ for any λ such that $\lambda \equiv a \bmod \mathbb{Z}$. Since $a \in A$ is arbitrary, we obtain the lemma in the general case. □

8.8 Proof of Theorem 8.3

In this section, we give a proof of Theorem 8.3. This is done by showing a more general formula in Lemma 8.8 without the scalar type assumption (8.3.1). Then, Theorem 8.3 follows readily from Lemma 8.8 because the assumption (8.3.1) makes $\dim \pi_\mu^{\mathfrak{k}} = 1$ and $\mathbb{S}_{(a_1,\dots,a_l)}(\mu) = \{\mu - \sum_{j=1}^{l} a_j \nu_j\}$ (see (8.8.1) for notation).

For the discussion below, it is convenient to use the concept of a multiset. Intuitively, a multiset is a set counted with multiplicities; for example, $\{a, a, a, b, c, c\}$. More precisely, a multiset $\mathbb{S}$ consists of a set S and a function $m : S \to \{0, 1, 2, \dots, \infty\}$. If $\mathbb{S}' = \{S, m'\}$ is another multiset on S such that $m'(x) \le m(x)$ for all $x \in S$, we say $\mathbb{S}'$ is a *submultiset* of $\mathbb{S}$ and write $\mathbb{S}' \subset \mathbb{S}$.

Suppose we are in the setting of Subsection 8.2 and recall τ is an involution of holomorphic type. For a $\Delta^+(\mathfrak{k}, \mathfrak{t})$-dominant weight μ, we introduce a multiset $\mathbb{S}(\mu)$ consisting of $\Delta^+(\mathfrak{k}^\tau, \mathfrak{t}^\tau)$-dominant weights:

$$\mathbb{S}(\mu) := \bigcup_{\substack{a_1 \ge \cdots \ge a_l \ge 0 \\ a_1, \dots, a_l \in \mathbb{N}}} \mathbb{S}_{(a_1,\dots,a_l)}(\mu) \,,$$

where we define the multiset $\mathbb{S}_{(a_1,\dots,a_l)}(\mu)$ by

$$\begin{aligned} &\{\text{highest weight of irreducible } \mathfrak{k}^\tau\text{-modules occurring in} \\ &\quad \pi^{\mathfrak{k}^\tau}_{-\sum_{j=1}^{l} a_j \nu_j} \otimes \pi_\mu^{\mathfrak{k}}|_{\mathfrak{k}^\tau} \text{ counted with multiplicities}\}. \end{aligned} \tag{8.8.1}$$

Because the central element $\widetilde{Z} = \frac{1}{\sqrt{-1}} Z$ of $\mathfrak{k}_{\mathbb{C}}$ acts on the irreducible representation $\pi_\mu^{\mathfrak{k}}$ by the scalar $\mu(\widetilde{Z})$ and because $\nu_j(\widetilde{Z}) = 1$ for all j $(1 \le j \le l)$, any element ν in $\mathbb{S}_{(a_1,\dots,a_l)}(\mu)$ satisfies $\nu(\widetilde{Z}) = -\sum_{j=1}^{l} a_j + \mu(\widetilde{Z})$. Therefore, the multiplicity of each element of the multiset $\mathbb{S}(\mu)$ is finite.

Lemma 8.8. *Let τ be an involution of G of holomorphic type, and $H = G_0^\tau$. If π_μ^G is a (relative) holomorphic discrete series representation of G, then it decomposes discretely into irreducible H-modules as*

$$\pi_\mu^G|_H \simeq \sum_{\nu \in \mathbb{S}(\mu)}^{\oplus} \pi_\nu^H \quad \textit{(discrete Hilbert sum).}$$

Proof of Lemma 8.8. It follows from Fact 5.1 (1) that π_μ^G is $(H \cap K)$-admissible, and splits discretely into a Hilbert direct sum of irreducible unitary representations of H.

Applying Lemma 8.7 to $H = G_0^\tau$, we see that Lemma 8.8 is deduced from the following $\mathfrak{k}^\tau$-isomorphism:

$$\pi_\mu^{\mathfrak{g}} \simeq \bigoplus_{\nu \in \mathbb{S}(\mu)} \pi_\nu^{\mathfrak{g}^\tau} \quad \text{(algebraic direct sum).} \tag{8.8.2}$$

The rest of the proof is devoted to showing (8.8.2).

Since π_μ^G is a holomorphic discrete series, $\pi_\mu^{\mathfrak{g}}$ is isomorphic to the generalized Verma module $N^{\mathfrak{g}}(\mu) = U(\mathfrak{g}_{\mathbb{C}}) \otimes_{U(\mathfrak{k}_{\mathbb{C}}+\mathfrak{p}_+)} \pi_\mu^{\mathfrak{k}}$ as a $\mathfrak{g}$-module, which in turn is isomorphic to the $\mathfrak{k}$-module $S(\mathfrak{p}_-) \otimes \pi_\mu^{\mathfrak{k}}$.

According to the decomposition $\mathfrak{p}_- = \mathfrak{p}_-^\tau \oplus \mathfrak{p}_-^{-\tau}$ as $\mathfrak{k}^\tau$-modules, we then have the following $\mathfrak{k}^\tau$-isomorphism:

$$\pi_\mu^{\mathfrak{g}} \simeq S(\mathfrak{p}_-) \otimes \pi_\mu^{\mathfrak{k}} \simeq S(\mathfrak{p}_-^\tau) \otimes S(\mathfrak{p}_-^{-\tau}) \otimes \pi_\mu^{\mathfrak{k}}. \tag{8.8.3}$$

Now, we consider the Hermitian symmetric space $G^{\tau\theta}/G^{\tau,\theta}$, for which the complex structure is given by the decomposition $\mathfrak{g}_{\mathbb{C}}^{\tau\theta} = \mathfrak{k}_{\mathbb{C}}^\tau \oplus \mathfrak{p}_+^{-\tau} \oplus \mathfrak{p}_-^{-\tau}$ (see (8.2.2)). Then, the Hua–Kostant–Schmid formula ([78, Behauptung c]) applied to $G^{\tau\theta}/G^{\tau,\theta}$ decomposes the symmetric algebra $S(\mathfrak{p}_-^{-\tau})$ into irreducible $\mathfrak{k}^\tau$-modules:

$$S(\mathfrak{p}_-^{-\tau}) \simeq \bigoplus_{\substack{a_1 \ge \cdots \ge a_l \ge 0 \\ a_1,\dots,a_l \in \mathbb{N}}} \pi^{\mathfrak{k}^\tau}_{-\sum_{j=1}^{l} a_j \nu_j}. \tag{8.8.4}$$

It follows from the definition of $\mathbb{S}(\mu)$ that we have the following irreducible decomposition as $\mathfrak{k}^\tau$-modules:

$$S(\mathfrak{p}_-^{-\tau}) \otimes \pi_\mu^{\mathfrak{k}} \simeq \bigoplus_{\nu \in \mathbb{S}(\mu)} \pi_\nu^{\mathfrak{k}^\tau}.$$

Combining this with (8.8.3), we get a $\mathfrak{k}^\tau$-isomorphism

$$\pi_\mu^{\mathfrak{g}} \simeq \bigoplus_{\nu \in \mathbb{S}(\mu)} S(\mathfrak{p}_-^\tau) \otimes \pi_\nu^{\mathfrak{k}^\tau} .$$

Next, we consider the Verma module $N^{\mathfrak{g}^\tau}(\nu) = U(\mathfrak{g}_{\mathbb{C}}^\tau) \otimes_{U(\mathfrak{k}_{\mathbb{C}}^\tau + \mathfrak{p}_+^\tau)} \pi_\nu^{\mathfrak{k}^\tau}$ of the subalgebra $\mathfrak{g}^\tau$. Then, $\pi_\nu^{\mathfrak{g}^\tau}$ is the unique irreducible quotient of $N^{\mathfrak{g}^\tau}(\nu)$. We shall show later that $N^{\mathfrak{g}^\tau}(\nu)$ is irreducible as a $\mathfrak{g}^\tau$-module, but at this stage we denote by $\pi_\nu^{\mathfrak{g}^\tau}, \pi_{\nu'}^{\mathfrak{g}^\tau}, \pi_{\nu''}^{\mathfrak{g}^\tau}, \dots$ the totality of irreducible subquotient modules of $N^{\mathfrak{g}^\tau}(\nu)$. (There are at most finitely many subquotients, and all of them are highest weight modules.) Then, as $\mathfrak{k}^\tau$-modules, we have the following isomorphisms:

$$\begin{aligned} S(\mathfrak{p}_-^\tau) \otimes \pi_\nu^{\mathfrak{k}^\tau} &\simeq N^{\mathfrak{g}^\tau}(\nu) \\ &\simeq \pi_\nu^{\mathfrak{g}^\tau} \oplus \pi_{\nu'}^{\mathfrak{g}^\tau} \oplus \pi_{\nu''}^{\mathfrak{g}^\tau} \oplus \cdots . \end{aligned}$$

Therefore, we get a $\mathfrak{k}^\tau$-isomorphism:

$$\pi_\mu^{\mathfrak{g}} \simeq \bigoplus_{\nu \in \mathbb{S}(\mu)} (\pi_\mu^{\mathfrak{g}^\tau} \oplus \pi_{\nu'}^{\mathfrak{g}^\tau} \oplus \pi_{\nu''}^{\mathfrak{g}^\tau} \oplus \cdots) .$$

Accordingly, the restriction $\pi_\mu^G|_H$ splits discretely into irreducible unitary representations of H by Lemma 8.7:

$$\pi_\mu^G|_H \simeq \sum_{\nu \in \mathbb{S}(\mu)}^{\oplus} (\pi_\nu^H \oplus \pi_{\nu'}^H \oplus \pi_{\nu''}^H \oplus \cdots) .$$

Since π_μ^G is a (relative) holomorphic discrete series representation of G, all irreducible summands in the right-hand side must be (relative) holomorphic discrete series representations of H by Fact 5.1 (1). Therefore, $N^{\mathfrak{g}^\tau}(\nu)$ is irreducible, and the other subquotients $\pi_{\nu'}^{\mathfrak{g}^\tau}, \pi_{\nu''}^{\mathfrak{g}^\tau}, \dots$ do not appear. Hence, the $\mathfrak{k}^\tau$-structures of the both sides of (8.8.2) are the same. Thus, Lemma 8.8 is proved. □

8.9 Proof of Theorem 8.4

For two irreducible representations $\pi_{\mu_1}^{\mathfrak{k}}$ and $\pi_{\mu_2}^{\mathfrak{k}}$, we define a multiset $\mathbb{S}(\mu_1, \mu_2)$ consisting of $\Delta^+(\mathfrak{k}, \mathfrak{t})$-dominant weights by

$$\mathbb{S}(\mu_1, \mu_2) := \bigcup_{\substack{a_1 \geq \cdots \geq a_{\bar{l}} \geq 0 \\ a_1, \dots, a_{\bar{l}} \in \mathbb{N}}} \mathbb{S}_{(a_1, \dots, a_{\bar{l}})}(\mu_1, \mu_2) ,$$

where $\mathbb{S}_{(a_1, \dots, a_{\bar{l}})}(\mu_1, \mu_2)$ is the multiset consisting of highest weights of irreducible $\mathfrak{k}$-modules occurring in $\pi^{\mathfrak{k}}_{-\sum_{j=1}^{\bar{l}} a_j \bar{\nu}_j} \otimes \pi_{\mu_1}^{\mathfrak{k}} \otimes \pi_{\mu_2}^{\mathfrak{k}}$ counted with multiplicities.

Theorem 8.4 is derived from the following more general formula:

Lemma 8.9. *The tensor product of two (relative) holomorphic discrete series representations $\pi^G_{\mu_1}$ and $\pi^G_{\mu_2}$ decomposes as follows:*

$$\pi^G_{\mu_1} \widehat{\otimes} \pi^G_{\mu_2} \simeq \sum_{\nu \in \mathbb{S}(\mu_1,\mu_2)}^{\oplus} \pi^G_\nu .$$

Proof. We define two injective maps by

$$\begin{aligned} \mathrm{diag} &: \mathfrak{p}_+ \to \mathfrak{p}_+ \oplus \mathfrak{p}_+ , \quad X \mapsto (X, X) , \\ \mathrm{diag}' &: \mathfrak{p}_+ \to \mathfrak{p}_+ \oplus \mathfrak{p}_+ , \quad X \mapsto (X, -X) . \end{aligned}$$

It then follows that we have $\mathfrak{k}$-isomorphisms:

$$\begin{aligned} S(\mathfrak{p}_-) \otimes S(\mathfrak{p}_-) &\simeq S(\mathfrak{p}_- \oplus \mathfrak{p}_-) \\ &\simeq S(\mathrm{diag}(\mathfrak{p}_-)) \otimes S(\mathrm{diag}'(\mathfrak{p}_-)) \\ &\simeq \bigoplus_{\substack{a_1 \ge \cdots \ge a_{\bar{l}} \ge 0 \\ a_1, \dots, a_{\bar{l}} \in \mathbb{N}}} S(\mathrm{diag}(\mathfrak{p}_-)) \otimes \pi^{\mathfrak{k}}_{-\sum_{j=1}^{\bar{l}} a_j \bar{\nu}_j} . \end{aligned}$$

This brings us the following $\mathfrak{k}$-isomorphisms:

$$\begin{aligned} \pi^{\mathfrak{g}}_{\mu_1} \otimes \pi^{\mathfrak{g}}_{\mu_2} &\simeq S(\mathfrak{p}_-) \otimes \pi^{\mathfrak{k}}_{\mu_1} \otimes S(\mathfrak{p}_-) \otimes \pi^{\mathfrak{k}}_{\mu_2} \\ &\simeq \bigoplus_{\nu \in \mathbb{S}(\mu_1,\mu_2)} S(\mathrm{diag}(\mathfrak{p}_-)) \otimes \pi^{\mathfrak{k}}_\nu \\ &\simeq \bigoplus_{\nu \in \mathbb{S}(\mu_1,\mu_2)} N^{\mathfrak{g}}_\nu . \end{aligned}$$

The rest of the proof goes similarly to that of Lemma 8.8. □

8.10 Restriction $U(p,q) \downarrow U(p-1,q)$ and $SO(n,2) \downarrow SO(n-1,2)$

Suppose (G, H) is a reductive symmetric pair whose complexification $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{h}_{\mathbb{C}})$ is one of the following types:

$(\mathfrak{sl}(n,\mathbb{C}), \mathfrak{gl}(n-1,\mathbb{C}))$ (or $(\mathfrak{gl}(n,\mathbb{C}), \mathfrak{gl}(1,\mathbb{C}) + \mathfrak{gl}(n-1,\mathbb{C}))$),

$(\mathfrak{so}(n,\mathbb{C}), \mathfrak{so}(n-1,\mathbb{C}))$.

As is classically known (see [83]), for compact (G, H) such as $(U(n), U(1) \times U(n-1))$ or $(SO(n), SO(n-1))$, any irreducible finite-dimensional representation π of G is multiplicity-free when restricted to H. For non-compact (G, H) such as $(U(p,q), U(1) \times U(p-1,q))$ or $(SO(n,2), SO(n-1,2))$, an analogous theorem still holds for highest weight representations π.

Theorem 8.10. *If* $(\mathfrak{g}, \mathfrak{h}) = (\mathfrak{u}(p, q), \mathfrak{u}(1) + \mathfrak{u}(p-1, q))$ *or* $(\mathfrak{so}(n, 2), \mathfrak{so}(n-1, 2))$, *then any irreducible unitary highest weight representation of* G *decomposes discretely into a multiplicity-free sum of irreducible unitary highest weight representations of* H.

In contrast to Theorem A, the distinguishing feature of Theorem 8.10 is that π is not necessarily of scalar type but an arbitrary unitary highest weight module. The price to pay is that the pair (G, H) is very special. We do not give the proof here that uses the vector bundle version of Theorem 2.2 (see [49]). Instead, we give an explicit decomposition formula for holomorphic discrete series π. The proof of Theorem 8.10 for the case $(G, H) = (SO_0(n, 2), SO_0(n-1, 2))$ can be also found in Jakobsen and Vergne [31, Corollary 3.1].

8.11 Branching law for $U(p, q) \downarrow U(p-1, q)$

This subsection gives an explicit branching law of a holomorphic discrete series representation π_μ^G of $G = U(p, q)$ when restricted to $H = U(1) \times U(p-1, q)$. Owing to (8.3.2), such π_μ^G is parametrized by $\mu = (\mu_1, \ldots, \mu_{p+q}) \in \mathbb{Z}^{p+q}$ satisfying

$$\mu_1 \geq \cdots \geq \mu_p, \mu_{p+1} \geq \cdots \geq \mu_{p+q}, \mu_{p+q} \geq \mu_1 + p + q .$$

Here is the formula.

Theorem 8.11 (Branching law $U(p, q) \downarrow U(p-1, q)$**).** *Retain the above setting. Then, the branching law of* π_μ^G *of the restriction to the subgroup* H *is multiplicity-free for any* μ; *it is given as follows:*

$$\pi_\mu^G|_H \simeq \sum_{a=0}^{\infty}{}^{\oplus} \sum_{\substack{\mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_p \geq \mu_p \\ \lambda_{p+1} \geq \mu_{p+1} \geq \cdots \geq \lambda_{p+q} \geq \mu_{p+q} \\ \sum_{i=1}^{q} (\lambda_{p+i} - \mu_{p+i}) = a}}{}^{\oplus} \mathbb{C}_{\sum_{i=1}^{p} \mu_i - \sum_{i=2}^{p} \lambda_i - a} \boxtimes \pi^{U(p-1,q)}_{(\lambda_2, \ldots, \lambda_p, \lambda_{p+1}, \ldots, \lambda_{p+q})} . \tag{8.11.1}$$

Proof. For $(G, H) \equiv (G, G^\tau) = (U(p, q), U(1) \times U(p-1, q))$, we have

$$G^{\tau\theta} \simeq U(1, q) \times U(p-1),$$

$$H \cap K \; (= K^\tau = K^{\tau\theta}) \simeq U(1) \times U(p-1) \times U(q),$$

$\mathfrak{t}^\tau = \mathfrak{t}$, and

$$\Delta^+(\mathfrak{p}_+^{-\tau}, \mathfrak{t}^\tau) = \{e_1 - e_{p+i} : 1 \leq i \leq q\},$$

by using the standard basis of $\Delta(\mathfrak{g}, \mathfrak{t}) = \{\pm(e_i - e_j) : 1 \leq i < j \leq p+q\}$. Thus, $l = \mathbb{R}$- rank $G^{\tau\theta} = 1$ and $\nu_1 = e_1 - e_{p+1}$. Hence, the $\mathfrak{k}^\tau$-type formula (8.8.4) amounts to

$$S(\mathfrak{p}_-^{-\tau}) \simeq \bigoplus_{a=0}^{\infty} \pi^{H\cap K}_{-a(e_1-e_{p+1})}$$
$$\simeq \bigoplus_{a=0}^{\infty} \mathbb{C}_{-a} \boxtimes \mathbf{1} \boxtimes \pi^{U(q)}_{(a,0,\dots,0)} \qquad (8.11.2)$$

as $H \cap K \simeq U(1) \times U(p-1) \times U(q)$ modules. Here, $\mathbf{1}$ denotes the trivial one-dimensional representation of $U(p-1)$.

On the other hand, we recall a classical branching formula $U(p) \downarrow U(p-1)$:

$$\pi^{U(p)}_{(\mu_1,\dots,\mu_p)}|_{U(1)\times U(p-1)} \simeq \bigoplus_{\mu_1\ge\lambda_2\ge\mu_2\ge\cdots\ge\lambda_p\ge\mu_p} \mathbb{C}_{\sum_{i=1}^p \mu_i - \sum_{i=2}^p \lambda_i} \otimes \pi^{U(p-1)}_{(\lambda_2,\dots,\lambda_p)},$$

whereas the classical Pieri rule says that

$$\pi^{U(q)}_{(a,0,\dots,0)} \otimes \pi^{U(q)}_{(\mu_{p+1},\dots,\mu_{p+q})} \simeq \bigoplus_{\substack{\lambda_{p+1}\ge\mu_{p+1}\ge\cdots\ge\lambda_{p+q}\ge\mu_{p+q} \\ \sum_{i=1}^q(\lambda_{p+i}-\mu_{p+i})=a}} \pi^{U(q)}_{(\lambda_{p+1},\dots,\lambda_{p+q})}.$$

These two formulae together with (8.11.2) yield the following $\mathfrak{k}^\tau$-isomorphisms:

$$S(\mathfrak{p}_-^{-\tau}) \otimes \pi^{\mathfrak{k}}_{\mu}|_{\mathfrak{k}^\tau}$$
$$\simeq \bigoplus_{a=0}^{\infty} ((\mathbb{C}_{-a} \boxtimes \mathbf{1}) \otimes \pi^{U(p)}_{(\mu_1,\dots,\mu_p)}|_{U(1)\times U(p-1)}) \boxtimes (\pi^{U(q)}_{(a,0,\dots,0)} \otimes \pi^{U(q)}_{(\mu_{p+1},\dots,\mu_{p+q})})$$
$$\simeq \bigoplus_{a=0}^{\infty} \bigoplus_{\substack{\mu_1\ge\lambda_2\ge\mu_2\ge\cdots\ge\lambda_p\ge\mu_p \\ \lambda_{p+1}\ge\mu_{p+1}\ge\cdots\ge\lambda_{p+q}\ge\mu_{p+q} \\ \sum_{i=1}^q(\lambda_{p+i}-\mu_{p+i})=a}} \mathbb{C}_{\sum_{i=1}^p \mu_i - \sum_{i=2}^p \lambda_i - a} \boxtimes \pi^{U(p-1)}_{(\lambda_2,\dots,\lambda_p)} \boxtimes \pi^{U(q)}_{(\lambda_{p+1},\dots,\lambda_{p+q})}.$$

In view of the $\mathfrak{k}^\tau$-isomorphisms

$$\pi^{\mathfrak{g}}_{\mu} \simeq S(\mathfrak{p}_-^{\tau}) \otimes S(\mathfrak{p}_-^{-\tau}) \otimes \pi^{\mathfrak{k}}_{\mu}|_{\mathfrak{k}^\tau}$$

and $N^{\mathfrak{g}^\tau}(\nu) \simeq S(\mathfrak{p}_-^{\tau}) \otimes \pi^{\mathfrak{k}^\tau}_{\nu}$, we have now shown that the $\mathfrak{k}^\tau$-structure of $\pi^{\mathfrak{g}}_{\mu}$ coincides with that of

$$\bigoplus_{a=0}^{\infty} \bigoplus_{\substack{\mu_1\ge\lambda_2\ge\mu_2\ge\cdots\ge\lambda_p\ge\mu_p \\ \lambda_{p+1}\ge\mu_{p+1}\ge\cdots\ge\lambda_{p+q}\ge\mu_{p+q} \\ \sum_{i=1}^q(\lambda_{p+i}-\mu_{p+i})=a}} N^{\mathfrak{g}^\tau}\left(\sum_{i=1}^p \mu_i - \sum_{i=2}^p \lambda_i - a, \lambda_2, \dots, \lambda_{p+q}\right).$$

As in the last part of the proof of Theorem 8.3, we see that any generalized Verma module occurring in the right-hand side is irreducible (and is isomorphic to the underlying $(\mathfrak{g}^\tau_{\mathbb{C}}, H \cap K)$-module of a holomorphic discrete series of H). Therefore, Theorem follows from Lemma 8.7. □

9 Appendix: Associated bundles on Hermitian symmetric spaces

In this appendix, we explain standard operations on homogeneous vector bundles. The results are well-known and elementary, but we recall them briefly for the convenience of the reader. The main goal is Lemma 9.4 which is used to verify the condition (2.2.2) in Theorem 2.2.

9.1 Homogeneous vector bundles

Let M be a real manifold, and V a (finite-dimensional) vector space over $\mathbb{C}$. Suppose that we are given an open covering $\{U_\alpha\}$ of M and transition functions

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL_{\mathbb{C}}(V)$$

satisfying the following compatibility conditions:

$$g_{\alpha\beta}\ g_{\beta\gamma}\ g_{\gamma\alpha} \equiv \mathrm{id} \quad \text{on } U_\alpha \cap U_\beta \cap U_\gamma\ ; \qquad g_{\alpha\alpha} \equiv \mathrm{id} \quad \text{on } U_\alpha\,.$$

A complex vector bundle $\mathcal{V}$ over M with typical fiber V is constructed as the equivalence class of $\coprod_\alpha (U_\alpha \times V)$, where $(x, v) \in U_\beta \times V$ and $(y, w) \in U_\alpha \times V$ are defined to be equivalent if $y = x$ and $w = g_{\alpha\beta}(x)v$. Then, the space of sections $\Gamma(M, \mathcal{V})$ is identified with the collection

$$\{(f_\alpha) : f_\alpha \in C^\infty(U_\alpha, V),\ f_\alpha(x) = g_{\alpha\beta}(x) f_\beta(x),\ \text{for } x \in U_\alpha \cap U_\beta\}\,. \tag{9.1.1}$$

If M is a complex manifold and if every $g_{\alpha\beta}$ is holomorphic (or anti-holomorphic), then $\mathcal{V} \to M$ is a holomorphic (or anti-holomorphic, respectively) vector bundle.

Next, let G be a Lie group, K a closed subgroup of G, and $M := G/K$ the homogeneous manifold. Then, we can take an open covering $\{U_\alpha\}$ of M such that for each α there is a local section $\varphi_\alpha : U_\alpha \to G$ of the principal bundle $G \to G/K$. Given a representation $\chi : K \to GL_{\mathbb{C}}(V)$, we define the homogeneous vector bundle, $\mathcal{V} := G \times_K (\chi, V)$. Then $\mathcal{V}$ is associated with the transition functions:

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL_{\mathbb{C}}(V), \quad g_{\alpha\beta}(x) := \chi(\varphi_\alpha(x)^{-1}\varphi_\beta(x))\,.$$

The section space $\Gamma(M, \mathcal{V})$ is identified with the following subspace of $C^\infty(G, V)$:

$$\{f \in C^\infty(G, V) : f(gk) = \chi^{-1}(k) f(g),\ \text{for } g \in G, k \in K\}\,. \tag{9.1.2}$$

9.2 Pullback of vector bundles

Let G' be a Lie group, K' a closed subgroup of G', and $M' := G'/K'$ the homogeneous manifold. Suppose that $\sigma : G' \to G$ is a Lie group homomorphism such that $\sigma(K') \subset K$. We use the same letter σ to denote by the induced map $M' \to M$, $g'K' \mapsto \sigma(g')K$. Then the pullback of the vector bundle $\mathcal{V} \to M$, denoted by $\sigma^*\mathcal{V} \to M'$, is associated to the representation

$$\chi \circ \sigma : K' \to GL_{\mathbb{C}}(V)\,.$$

Then we have a commuting diagram of the pullback of sections (see (9.1.2)):

$$\begin{array}{cccccc} \sigma^*: & \Gamma(M,\mathcal{V}) & \to & \Gamma(M',\sigma^*\mathcal{V}), & (f_\alpha)_\alpha \mapsto (f_\alpha\circ\sigma)_\alpha\,, \\ & \cap & & \cap & \\ \sigma^*: & C^\infty(G,V) & \to & C^\infty(G',V), & f \quad\mapsto\quad f\circ\sigma\,. \end{array}$$

9.3 Push-forward of vector bundles

Suppose that V and W are complex vector spaces and that $\xi : V \to W$ is an anti-linear bijective map. Then, we have an anti-holomorphic group isomorphism

$$GL_{\mathbb{C}}(V) \to GL_{\mathbb{C}}(W), \quad g \mapsto g^\xi := \xi\circ g\circ\xi^{-1}\,.$$

Let $\mathcal{V} \to M$ be a complex vector bundle with transition functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL_{\mathbb{C}}(V)$. Then, one constructs a complex vector bundle $\xi_*\mathcal{V} \to M$ with the transition functions $g^\xi_{\alpha\beta} : U_\alpha\cap U_\beta \to GL_{\mathbb{C}}(W)$. We have a natural homomorphism

$$\xi_* : \Gamma(M,\mathcal{V}) \to \Gamma(M,\xi_*\mathcal{V}), \quad (f_\alpha) \mapsto (\xi\circ f_\alpha)\,,$$

which is well defined because the compatibility condition in (9.1.1) is satisfied as follows: If $x \in U_\alpha \cap U_\beta$, then

$$g^\xi_{\alpha\beta}(x)(\xi\circ f_\beta)(x) = (\xi\circ g_{\alpha\beta}(x)\circ\xi^{-1})(\xi\circ f_\beta)(x) = \xi\circ g_{\alpha\beta}(x) f_\beta(x) = \xi\circ f_\alpha(x)\,.$$

If $\mathcal{V}$ is the homogeneous vector bundle $G\times_K(\chi, V)$ associated to a representation $\chi : K \to GL_{\mathbb{C}}(V)$, then $\xi_*\mathcal{V}$ is isomorphic to the homogeneous vector bundle $G\times_K(\chi^\xi, W)$ associated to the representation

$$\chi^\xi : K \to GL_{\mathbb{C}}(W), \quad k \mapsto \chi^\xi(k) := \xi\circ\chi(k)\circ\xi^{-1}\,.$$

9.4 A sufficient condition for the isomorphism $\xi_*\sigma^*\mathcal{V} \simeq \mathcal{V}$

We are particularly interested in the case where $G' = G$, $K' = K$, $V = W = \mathbb{C}$ and $\xi(z) := \bar{z}$ (the complex conjugate of z) in the setting of Subsections 9.2 and 9.3.

By the identification of $GL_{\mathbb{C}}(\mathbb{C})$ with $\mathbb{C}^\times$, we have $g^\xi = \overline{g}$ for $g \in GL_{\mathbb{C}}(V) \simeq \mathbb{C}^\times$. Then, χ^ξ coincides with the conjugate representation

$$\overline{\chi} : K \to GL_{\mathbb{C}}(W) \simeq \mathbb{C}^\times, \quad k \mapsto \overline{\chi(k)}$$

for $\chi \in \mathrm{Hom}(K, \mathbb{C}^\times)$. Thus, we have an isomorphism of G-equivariant holomorphic line bundles:

$$\xi_*\sigma^*\mathcal{V} \simeq G\times_K(\overline{\chi\circ\sigma}, \mathbb{C}) \tag{9.4.1}$$

with the following correspondence of sections:

$$\xi_*\circ\sigma^* : \Gamma(M,\mathcal{V}) \to \Gamma(M,\xi_*\sigma^*\mathcal{V}), \quad (f_\alpha) \mapsto (\overline{f_\alpha\circ\sigma})\,.$$

We now apply the formula (9.4.1) to the setting where $M = G/K$ is an irreducible Hermitian symmetric space.

Lemma 9.4. *Let $\chi : K \to \mathbb{C}^\times$ be a unitary character. We denote by $\mathcal{V}$ the homogeneous line bundle $G \times_K (\chi, \mathbb{C})$. Suppose σ is an involutive automorphism of G of anti-holomorphic type (see Definition 1.4). Then we have an isomorphism of G-equivariant holomorphic line bundles: $\xi_* \sigma^* \mathcal{V} \simeq \mathcal{V}$.*

Proof. In view of (9.4.1), it suffices to show $\overline{\chi \circ \sigma} = \chi$. As the character χ of K is unitary, we have $\overline{\chi(k)} = \chi(k^{-1})$ for any $k \in K$. Let Z be a generator of the center $\mathfrak{c}(\mathfrak{k})$ of $\mathfrak{k}$. Since σ is of anti-holomorphic type, we have $\sigma Z = -Z$, and then

$$\overline{\chi \circ \sigma(\exp tZ)} = \overline{\chi(\exp(-tZ))} = \chi(\exp tZ) \qquad (t \in \mathbb{R}).$$

On the other hand, if $k \in [K, K]$, then $\overline{\chi \circ \sigma(k)} = 1 = \chi(k)$ because $[K, K]$ is a connected semisimple Lie group. As $K = \exp \mathfrak{c}(\mathfrak{k}) \cdot [K, K]$, we have shown $\overline{\chi \circ \sigma} = \chi$. Hence the lemma. □

Acknowledgement. I owe much to J. Faraut for enlightening discussions, in particular, for clarifying the idea of his joint work [15] with E. Thomas at an early stage on the occasion of the CIMPA school on Invariant Theory in Tunisia organized by P. Torasso in 1996, when I formalized Theorem A and gave its proof in the classical case.

Some of the results of this article were presented in various places including the conferences ‘Representation Theory at Saga’ in 1997 (see [39]) organized by K. Mimachi, at ‘Summer Solstice Days at Paris University VI and VII’ in 1999, at Copenhagen organized by B. Ørsted and H. Schlichtkrull in 1999, and at Oberwolfach organized by A. Huckleberry and K.-H. Neeb and J. Wolf in 2000 and 2004, at the MSRI program ‘Integral Geometry’ organized by S. Gindikin, L. Barchini, and R. Zierau in 2001, at the workshop ‘Representation theory of Lie groups, harmonic analysis on homogeneous spaces and quantization’ at the Lorentz center in Leiden in 2002 (see [46]) organized by G. van Dijk and V. F. Molchanov, and at the workshop ‘Representation theory and automorphic forms’ in Korea in 2005 coorganized by J. Yang and W. Schmid with the author. I express my deep gratitude to the organizers of these conferences and to the participants for helpful and stimulating comments on various occasions. I also thank an anonymous referee who read the manuscript very carefully. Special thanks are due to Ms. Suenaga for her help in preparing for the final manuscript.

References

1. Alikawa H, Multiplicity free branching rules for outer automorphisms of simple Lie algebras, (2007) *J. Math. Soc. Japan* 59:151–177
2. van den Ban EP (1987), Invariant differential operators on a semisimple symmetric space and finite multiplicities in a Plancherel formula. *Ark. Mat.* 25:175–187
3. van den Ban EP, Schlichtkrull H (1997), The most continuous part of the Plancherel decomposition for a reductive symmetric space. *Ann. of Math.* 145:267–364
4. Barbasch D, The spherical unitary dual of split and p-adic groups, preprint
5. Ben Saïd, S (2002), Weighted Bergman spaces on bounded symmetric domains. *Pacific J. Math.* 206:39–68
6. Berger M (1957), Les espaces symétriques non-compacts. *Ann. Sci. École Norm. Sup. (3)* 74:85–177

7. Bertram W, Hilgert J (1998), Hardy spaces and analytic continuation of Bergman spaces. *Bull. Soc. Math. France* 126:435–482
8. Delorme P (1998), Formule de Plancherel pour les espaces symétriques réductifs. *Ann. of Math.* 147:417–452
9. van Dijk G (1999), Canonical representations associated to hyperbolic spaces. II. *Indag. Math.* 10:357–368
10. van Dijk G, Hille SC (1997), Canonical representations related to hyperbolic spaces. *J. Funct. Anal.* 147:109–139
11. van Dijk G, Pevzner M (2001), Berezin kernels of tube domains. *J. Funct. Anal.* 181:189–208
12. Enright T, Howe R, Wallach N (1983), A classification of unitary highest weight modules. In *Representation theory of reductive groups*, Progr. Math. 40:97–143, Birkhäuser
13. Enright T, Joseph A (1990), An intrinsic classification of unitary highest weight modules. *Math. Ann.* 288:571–594
14. Faraut J, Ólafsson G (1995), Causal semisimple symmetric spaces, the geometry and harmonic analysis. *Semigroups in Algebra, Geometry and Analysis*, Walter de Gruyter 3–32
15. Faraut J, Thomas E (1999), Invariant Hilbert spaces of holomorphic functions. *J. Lie Theory* 9:383–402
16. Flensted-Jensen M (1980), Discrete series for semisimple symmetric spaces. *Ann. of Math.* 111:253–311
17. Gelfand IM (1950), Spherical functions on symmetric spaces. *Dokl. Akad. Nauk.* SSSR 70:5–8
18. Gutkin E (1979), Coefficients of Clebsch-Gordan for the holomorphic discrete series. *Lett. Math. Phys.* 3:185–192
19. Harish-Chandra (1953), (1954), Representations of semi-simple Lie groups. I, III. *Trans. Amer. Math. Soc.* 75:185–243; 76:234–253; (1955) IV. *Amer. J. Math.* 77:743–777
20. Heckman G, Schlichtkrull H (1994), Harmonic analysis and special functions on symmetric spaces. *Perspectives in Mathematics* 16, Academic Press
21. Helgason S (1970), (1976), A duality for symmetric spaces with applications to group representations. I, II. *Adv. Math.* 5:1–154; 22:187–219
22. Hilgert J, Reproducing kernels in representation theory, in prepration
23. Howe R (1983), Reciprocity laws in the theory of dual pairs. *Representation Theory of Reductive Groups*, Trombi PC (ed), Progr. Math. 40:159–175, Birkhäuser
24. Howe R (1989), Transcending classical invariant theory. *J. Amer. Math. Soc.* 2:535–552
25. Howe R (1995), Perspectives on invariant theory: Schur duality, multiplicity-free actions and beyond. *Israel Math. Conf. Proc.* 8:1–182
26. Hua LK (1963), Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains. Amer. Math. Soc.
27. Jaffee H (1975), Real forms of Hermitian symmetric spaces. *Bull. Amer. Math. Soc.* 81:456–458
28. Jaffee H (1978), Anti-holomorphic automorphisms of the exceptional symmetric domains. *J. Differential Geom.* 13:79–86
29. Jakobsen HP (1979), Tensor products, reproducing kernels, and power series. *J. Funct. Anal.* 31:293–305
30. Jakobsen HP (1983), Hermitian symmetric spaces and their unitary highest weight modules. *J. Funct. Anal.* 52:385–412
31. Jakobsen HP, Vergne M (1979), Restrictions and expansions of holomorphic representations. *J. Funct. Anal.* 34:29–53

32. Johnson K (1980), On a ring of invariant polynomials on a Hermitian symmetric space. *J. Algebra* 67:72–81
33. Kashiwara M, Vergne M (1978), On the Segal–Shale–Weil representations and harmonic polynomials. *Invent. Math.* 44:1–47
34. Kobayashi S (1968), Irreducibility of certain unitary representations. *J. Math. Soc. Japan* 20:638–642
35. Kobayashi S, Nagano T (1964), On filtered Lie algebras and geometric structures I. *J. Math. Mech.* 13:875–908
36. Kobayashi T (1989), Unitary representations realized in L^2-sections of vector bundles over semisimple symmetric spaces. Proceedings of the Joint Symposium of Real Analysis and Functional Analysis (cosponsored by the Mathematical Society of Japan), 39–54, in Japanese
37. Kobayashi T (1992), A necessary condition for the existence of compact Clifford-Klein forms of homogeneous spaces of reductive type. *Duke Math. J.* 67:653–664
38. Kobayashi T (1994), Discrete decomposability of the restriction of $A_{\mathfrak{q}}(\lambda)$ with respect to reductive subgroups and its applications. *Invent. Math.* 117:181–205
39. Kobayashi T (1997), Multiplicity free theorem in branching problems of unitary highest weight modules. *Proceedings of the Symposium on Representation Theory* held at Saga, Kyushu 1997, Mimachi K (ed), 9–17
40. Kobayashi T (1998), Discrete decomposability of the restriction of $A_{\mathfrak{q}}(\lambda)$ with respect to reductive subgroups, Part II — micro-local analysis and asymptotic K-support. *Ann. of Math.* 147:709–729
41. Kobayashi T (1998), Discrete decomposability of the restriction of $A_{\mathfrak{q}}(\lambda)$ with respect to reductive subgroups, Part III — restriction of Harish-Chandra modules and associated varieties. *Invent. Math.* 131:229–256
42. Kobayashi T (1998), Discrete series representations for the orbit spaces arising from two involutions of real reductive Lie groups. *J. Funct. Anal.* 152:100–135
43. Kobayashi T (2000), Discretely decomposable restrictions of unitary representations of reductive Lie groups — examples and conjectures. Adv. Stud. Pure Math., *Analysis on Homogeneous Spaces and Representation Theory of Lie Groups*, Okayama–Kyoto, 26:98–126
44. Kobayashi T (2000), Multiplicity-free restrictions of unitary highest weight modules for reductive symmetric pairs, UTMS Preprint Series 2000–1, University of Tokyo, Mathematical Sciences, 39 pages
45. Kobayashi T (2002), Branching problems of unitary representations. *Proc. I.C.M.* 2002, Beijing, vol. 2, Higher Ed. Press, Beijing, 615–627
46. Kobayashi T (2004), Geometry of multiplicity-free representations of $GL(n)$, visible actions on flag varieties, and triunity. *Acta Appl. Math.* 81:129–146
47. Kobayashi T (2005), Multiplicity-free representations and visible actions on complex manifolds. *Publ. Res. Inst. Math. Sci.* 41:497–549
48. Kobayashi T (2005), Theory of discretely decomposable restrictions of unitary representations of semisimple Lie groups and some applications. *Sugaku Expositions* 18:1–37, Amer. Math. Soc.
49. Kobayashi T, Propagation of multiplicity-free property for holomorphic vector bundles, math. RT/0607004
50. Kobayashi T (2007), Visible actions on symmetric spaces, to appear in Transformation Group 12, math. DE/0607005
51. Kobayashi T (2007), A generalized Cartan decomposition for the double coset space $(U(n_1) \times U(n_2) \times U(n_3))\backslash U(n)/(U(p) \times U(q))$, J. Math Soc. Japan 59:669–691

52. Kobayashi T, Visible actions on flag varieties of $SO(n)$ and multiplicity-free representations, preprint.
53. Kobayashi T, Nasrin S (2003), Multiplicity one theorem in the orbit method. *Lie Groups and Symmetric Spaces*: In memory of F. I. Karpelevič, S. Gindikin (ed), Translation Series 2, Amer. Math. Soc. 210:161–169
54. Kobayashi T, Ørsted B (2003), Analysis on the minimal representation of $\mathrm{O}(p, q)$. II. Branching laws. *Adv. Math.* 180:513–550
55. Koike K, Terada I (1987), Young diagrammatic methods for the representation theory of the classical groups of type B_n, C_n, D_n. *J. Algebra* 107:466–511
56. Koike K, Terada I (1990), Young diagrammatic methods for the restriction of representations of complex classical Lie groups to reductive subgroups of maximal rank. *Adv. Math.* 79:104–135
57. Korányi A, Wolf JA (1965), Realization of Hermitian symmetric spaces as generalized half-planes. *Ann. of Math.* 81:265–285
58. Krattenthaler C (1998), Identities for classical group characters of nearly rectangular shape. *J. Algebra* 209:1–61
59. Lang S (1998), $SL_2(\mathbb{R})$. Springer
60. Lipsman R (1979), On the unitary representation of a semisimple Lie group given by the invariant integral on its Lie algebra. *Adv. Math. Suppl.* 6:143–158; (1977) II, *Canad. J. Math.* 29:1217–1222
61. Littelmann P (1994), On spherical double cones. *J. Algebra* 166:142–157
62. Macdonald IG (1997), *Symmetric Functions and Hall Polynomials*. Oxford University Press
63. Martens S (1975), The characters of the holomorphic discrete series. *Proc. Natl. Acad. Sci. USA* 72:3275–3276
64. Molchanov VF (1980), Tensor products of unitary representations of the three-dimensional Lorentz group. *Math. USSR, Izv.* 15:113–143
65. Neeb K-H (1997), On some classes of multiplicity free representations. *Manuscripta Math.* 92:389–407
66. Neretin YA (2002), Plancherel formula for Berezin deformation of L^2 on Riemannian symmetric space. *J. Funct. Anal.* 189:336–408
67. Neretin YA, Ol'shanskiĭ GI (1997), Boundary values of holomorphic functions, singular unitary representations of the groups $\mathrm{O}(p, q)$ and their limits as $q \to \infty$. *J. Math. Sci.* (New York) 87:3983–4035
68. Okada S (1998), Applications of minor summation formulas to rectangular-shaped representations of classical groups. *J. Algebra* 205:337–367
69. Ólafsson G, Ørsted B (1996), Generalizations of the Bargmann transform. *Lie Theory and Its Application in physics* (Clausthal, 1995) Dobrev and Döbner (eds), World Scientific 3–14
70. Ørsted B, Zhang G (1997), Tensor products of analytic continuations of holomorphic discrete series. *Canad. J. Math.* 49:1224–1241
71. Pevsner M (1996), Espace de Bergman d'un semi-groupe complexe. *C. R. Acad. Sci. Paris Sér. I Math.* 322:635–640
72. Pevzner M (2005), Représentations des groupes de Lie conformes et quantification des espaces symétriques. Habilitation, l'université de Reims, 36pp.
73. Proctor RA (1983), Shifted plane partitions of trapezoidal shape. *Proc. Amer. Math. Soc.* 89:553–559
74. Repka J (1979), Tensor products of holomorphic discrete series representations. *Canad. J. Math.* 31:836–844

75. Richardson R, Röhrle G, Steinberg R (1992), Parabolic subgroup with abelian unipotent radical. *Invent. Math.* 110:649–671
76. Rossmann W (1979), The structure of semisimple symmetric spaces. *Canad. J. Math.* 31:156–180
77. Sato F (1993), On the stability of branching coefficients of rational representations of reductive groups. *Comment. Math.* Univ. St. Paul 42:189–207
78. Schmid W (1969–70), Die Randwerte holomorphe Funktionen auf hermetisch symmetrischen Raumen. *Invent. Math.* 9:61–80
79. Shimeno N (1994), The Plancherel formula for spherical functions with a one-dimensional K-type on a simply connected simple Lie group of Hermitian type. *J. Funct. Anal.* 121:330–388
80. Stembridge JR (1990), Hall–Littlewood functions, plane partitions, and the Rogers–Ramanujan identities. *Trans. Amer. Math. Soc.* 319:469–498
81. Stembridge JR (2001), Multiplicity-free products of Schur functions. Ann. Comb. 5:113–121
82. Vinberg ÉB (2001), Commutative homogeneous spaces and co-isotropic symplectic actions. Russian Math. Surveys 56:1–60
83. Vinberg ÉB, Kimelfeld BN (1978), Homogeneous domains on flag manifolds and spherical subgroups of semisimple Lie groups. *Funct. Anal. Appl.* 12:168–174
84. Vogan, Jr. D (1979), The algebraic structure of the representation of semisimple Lie groups. I. *Ann. of Math.* 109:1–60
85. Vogan, Jr. D (1981), Representations of Real Reductive Lie groups. Progr. Math. 15, Birkhäuser
86. Vogan, Jr. D (1987), *Unitary Representations of Reductive Lie Groups*. Ann. Math. Stud. 118, Princeton University Press
87. Wallach N (1988), *Real Reductive Groups I*, Academic Press
88. Wolf J (1980), Representations that remain irreducible on parabolic subgroups. Differential Geometrical Methods in Mathematical Physics (Proc. Conf., Aix-en-Provence/Salamanca, 1979), pp. 129–144, Lecture Notes in Math. 836, Springer
89. Xie J (1994), Restriction of discrete series of $SU(2,1)$ to $S(U(1) \times U(1,1))$. *J. Funct. Anal.* 122:478–518
90. Yamashita H, Wachi A, Isotropy representations for singular unitary highest weight modules, in preparation.
91. Zhang G (2001), Tensor products of minimal holomorphic representations. *Represent. Theory* 5:164–190
92. Zhang G (2002), Branching coefficients of holomorphic representations and Segal-Bargmann transform. *J. Funct. Anal.* 195:306–349

4

The Rankin–Selberg Method for Automorphic Distributions

Stephen D. Miller and Wilfried Schmid

[1] Institute of Mathematics
The Hebrew University
Jerusalem 91904, Israel
and
Department of Mathematics
Hill Center-Busch Campus
Rutgers University, Piscataway, NJ 08854
miller@math.huji.ac.il
[2] Department of Mathematics
Harvard University
Cambridge, MA 02138
schmid@math.harvard.edu

Summary. This paper describes our method of pairing automorphic distributions. We present a third technique for obtaining the analytic properties of automorphic L-functions, in addition to the existing methods of integral representations (Rankin–Selberg) and Fourier coefficients of Eisenstein series (Langlands–Shahidi). We recently used this technique to establish new cases of the full analytic continuation of the exterior square L-functions. The paper here gives an exposition of our method in two special, yet representative cases: the Rankin–Selberg tensor product L-functions for $PGL(2,\mathbb{Z})\backslash PGL(2,\mathbb{R})$, as well as for the exterior square L-functions for $GL(4,\mathbb{Z})\backslash GL(4,\mathbb{R})$.

Key words: analytic L-functions, integral representations, Rankin–Selberg, exterior square, automorphic distributions

Subject Classifications: 11M99, 22E30, 32N10

1 Introduction

We recently established the holomorphic continuation and functional equation of the exterior square L-function for $GL(n,\mathbb{Z})$, and more generally, the archimedean

*Partially supported by NSF grant DMS-0301172 and an Alfred P. Sloan Foundation Fellowship.
**Partially supported by NSF grant DMS-0500922.

theory of the $GL(n)$ exterior square L-function over $\mathbb{Q}$. We refer the reader to our paper [15] for a precise statement of the results and their relation to previous work on the subject. The purpose of this note is to give an account of our method in the simplest non-trivial cases, which can be explained without the technical overhead necessary for the general case.

Let us begin by recalling the classical results, about standard L-functions and Rankin–Selberg L-functions of modular forms. We consider a cuspidal modular form F, of weight k, on the upper half plane $\mathcal{H}$. To simplify the notation, we suppose that it is automorphic for $\Gamma = SL(2,\mathbb{Z})$, though the arguments can be adapted to congruence subgroups of $SL(2,\mathbb{Z})$. Like all modular forms, F has a Fourier expansion,

$$F(z) = \sum_{n\geq 1} a_n \, e(nz)\,, \quad \text{with } \; e(z) \;=_{\text{def}}\; e^{2\pi i z}\,. \tag{1.1}$$

For a general modular form, the Fourier series may involve a non-zero constant term a_0; it is the hypothesis of cuspidality that excludes the constant term. The Dirichlet series

$$L(s,F) = \sum_{n\geq 1} a_n \, n^{-\frac{k-1}{2}-s} \tag{1.2}$$

converges for $\operatorname{Re} s \gg 0$, extends holomorphically to the entire s-plane, and satisfies a functional equation relating $L(s,F)$ to $L(1-s,F)$. This is the standard L-function of the modular form F.

Hecke proved the holomorphic continuation and functional equation by expressing $L(s,F)$ in terms of the Mellin transform of F along the imaginary axis,

$$\begin{aligned} \int_0^\infty F(iy)\, y^{s-1}\, dy &= \sum_{n\geq 1} a_n \int_0^\infty e^{-2\pi n y}\, y^{s-1}\, dy \\ &= \sum_{n\geq 1} a_n \, n^{-s} \int_0^\infty e^{-2\pi y}\, y^{s-1}\, dy \\ &= (2\pi)^{-s}\, \Gamma(s)\, L\left(s - \frac{k-1}{2}, F\right), \end{aligned} \tag{1.3}$$

at least for $\operatorname{Re} s \gg 0$. The transformation law for the modular form F under $z \mapsto -1/z$,

$$F(-1/z) = (-z)^k \, F(z)\,, \tag{1.4}$$

implies that $F(iy)$ decays rapidly not only as $y \to \infty$, but also as $y \to 0$. That makes the Mellin transform, and hence also $\Gamma\big(s + \frac{k-1}{2}\big)L(s,F)$, globally defined and holomorphic. The Gamma function has no zeroes, so $L(s,F)$ is entire as well. The transformation law (1.4), coupled with the change of variables $y \mapsto 1/y$ and the shift $s \mapsto s + \frac{k-1}{2}$, gives the functional equation

$$\begin{aligned} &(2\pi)^{-s-\frac{k-1}{2}}\, \Gamma\left(s + \frac{k-1}{2}\right) L(s,F) \\ &\quad = i^k \, (2\pi)^{s-1-\frac{k-1}{2}}\, \Gamma\left(1 - s + \frac{k-1}{2}\right) L(1-s,F)\,. \end{aligned} \tag{1.5}$$

The factor i^k comes up naturally in the computation, yet might be misleading since $\Gamma = SL(2,\mathbb{Z})$ admits only modular forms of even weights.

In addition to F, we now consider a second modular form of weight k, which need not be cuspidal,

$$G(z) = \sum_{n\geq 0} b_n\, e(nz)\,. \tag{1.6}$$

The Rankin–Selberg L-function of the pair F, $\overline{G}$ = complex conjugate of G, is the Dirichlet series

$$L(s, F\otimes\overline{G}) = \zeta(2s) \sum_{n\geq 1} a_n\, \overline{b}_n\, n^{1-k-s}\,. \tag{1.7}$$

Its analytic continuation and functional equation were established separately by Rankin [17] and Selberg [18]. The proof depends on properties of the non-holomorphic Eisenstein series

$$E_s(z) = \pi^{-s}\,\Gamma(s)\,\zeta(2s) \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\mathrm{Im}\,(\gamma z))^s \qquad (\Gamma_\infty = \{\gamma\in\Gamma \mid \gamma\infty=\infty\}). \tag{1.8}$$

This sum is well defined since Γ_∞ acts on $\mathcal{H}$ by integral translations. It converges for $\mathrm{Re}\, s > 1$ and extends meromorphically to the entire s-plane with only one pole, of first order, at $s = 1$. The function $E_s(z)$ is Γ-invariant by construction, has moderate growth as $\mathrm{Im}\, z \to \infty$, and satisfies the functional equation

$$E_s(z) = E_{1-s}(z)\,. \tag{1.9}$$

Both $F(z)$ and $G(z)$ transform according to a factor of automorphy under the action of Γ, but $(\mathrm{Im}\, z)^k F(z)\overline{G}(z)$ is Γ-invariant, as is the measure $y^{-2}dxdy$. Since $G(z)$ and $E_s(z)$ have moderate growth as $\mathrm{Im}\, z \to \infty$, and since $F(z)$ decays rapidly, the integral

$$I(s) = \int_{\Gamma\backslash\mathcal{H}} (\mathrm{Im}\, z)^{k-2} F(z)\,\overline{G}(z)\,E_s(z)\,dx\,dy \tag{1.10}$$

converges. From $E_s(z)$, the function $I(s)$ inherits both the functional equation

$$I(s) = I(1-s) \tag{1.11}$$

and the analytic properties: it is holomorphic, with the exception of a potential first order pole at $s = 1$.

The definition (1.8) of $E_s(z)$ involves a sum of $\mathrm{Im}\,\gamma z$, with γ ranging over $\Gamma_\infty\backslash\Gamma$. But the rest of the integrand in (1.10) is Γ-invariant. That justifies the process known as "unfolding",

$$\begin{aligned} &\pi^s(\Gamma(s)\,\zeta(2s))^{-1} I(s) \\ &\quad = \int_{\Gamma\backslash\mathcal{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\mathrm{Im}\, z)^{k-2} F(z)\,\overline{G}(z)\,(\mathrm{Im}\,(\gamma z))^s\,dx\,dy \\ &\quad = \int_{\Gamma_\infty\backslash\mathcal{H}} (\mathrm{Im}\, z)^{s+k-2} F(z)\,\overline{G}(z)\,dx\,dy\,, \end{aligned} \tag{1.12}$$

at least for Re $s > 1$, in which case the integral on the right converges. Since Γ_∞ acts on $\mathcal{H}$ by integral translations, the strip $\{0 \le \operatorname{Re} z \le 1\}$ constitutes a fundamental domain for this action. Substituting the series (1.1, 1.6) for $F(z)$ and $G(z)$, one finds

$$\begin{aligned}
&\pi^s (\Gamma(s)\,\zeta(2s))^{-1} I(s) \\
&\quad = \int_0^\infty \int_0^1 \sum_{\substack{n>0\\ m\ge 0}} a_n\,\overline{b}_m\, e((n-m)x)\, e^{-2\pi(n+m)y} y^{s+k-2} dx\,dy \\
&\quad = \sum_{n\ge 1} a_n\,\overline{b}_n \int_0^\infty e^{-4\pi n y}\, y^{s+k-2}\, dy \\
&\quad = (4\pi)^{-s-k+1}\,\Gamma(s+k-1) \sum_{n\ge 1} a_n\,\overline{b}_n\, n^{-s-k+1}\,,
\end{aligned} \tag{1.13}$$

again for Re $s > 1$. Equivalently,

$$I(s) = 2^{1-k}\,(2\pi)^{1-k-2s}\,\Gamma(s)\,\Gamma(s+k-1)\,L(s, F\otimes\overline{G})\,. \tag{1.14}$$

The Gamma factors have no zeroes, so $L(s, F\otimes\overline{G})$ extends holomorphically to all of $\mathbb{C}$, except possibly for a first order pole at $s = 1$. In effect, (1.11) is the functional equation for the Rankin–Selberg L-function. With some additional effort one can modify these arguments, to make them work even when F and G have different weights.

Maass [12] extended the proofs of the analytic continuation and functional equation for the standard L-function to the case of Maass forms, i.e., Γ-invariant eigenfunctions of the hyperbolic Laplacian on $\mathcal{H}$; see section 2 below. Jacquet [6] treats the Rankin–Selberg L-function for Maass forms. We just saw how the Gamma factors in (1.3) and (1.13) arise directly from the standard integral representation of the Gamma function. In contrast, for Maass forms, the Gamma factor for the standard L-function arises from the Mellin transform of the Bessel function $K_\nu(y)$,

$$\int_0^\infty K_\nu(y)\, y^{s-1} = 2^{s-2}\,\Gamma\left(\frac{s-\nu}{2}\right)\Gamma\left(\frac{s+\nu}{2}\right) \qquad (\operatorname{Re} s \gg 0)\,, \tag{1.15}$$

and for the Rankin–Selberg L-function of a pair of Maass forms, from the integral

$$\begin{aligned}
&\int_0^\infty K_\mu(y)\, K_\nu(y)\, y^{s-1}\, dy \\
&= 2^{s-3}\, \frac{\Gamma\left(\frac{s-\mu-\nu}{2}\right)\Gamma\left(\frac{s+\mu-\nu}{2}\right)\Gamma\left(\frac{s-\mu+\nu}{2}\right)\Gamma\left(\frac{s+\mu+\nu}{2}\right)}{\Gamma(s)} \quad (\operatorname{Re} s \gg 0)\,.
\end{aligned} \tag{1.16}$$

Though (1.16) can be established by elementary means, it is still complicated and its proof lacks a conceptual explanation.

In the case of Rankin–Selberg L-functions of higher rank groups, the integrals analogous to (1.16) become exceedingly difficult, or even impossible, to compute.

In fact, it is commonly believed that such integrals may not always be expressible in terms of Gamma functions [1, §2.6]. If true, this would not contradict Langlands' prediction that the functional equations involve certain definite Gamma factors [10, 11]: the functional equations pin down only the ratios of the Gamma factors on the two sides, which can of course be expressed also as ratios of other functions.

Broadly speaking, the existing approaches to the L-functions for higher rank groups overcome the problem of computing these so-called *archimedean integrals* in one of two ways. Even if the integrals cannot be computed explicitly, it may be possible to establish a functional equation with unknown coefficients; it may then be possible to identify the coefficients in some special case, or by an analysis of their zeroes and poles. The Langlands-Shahidi method, on the other hand, often exhibits the functional equation with precisely the Gamma factors predicted by Langlands. Both methods have one difficulty in common: ruling out poles – other than those at the expected places – of the L-functions in question requires considerable effort, and is not always possible.

We are approaching the analytic continuation and functional equation of L-functions from a different point of view. Instead of working with automorphic forms – i.e., the higher dimensional analogues of modular forms and Maass forms – we attach the L-functions to *automorphic distributions*. In the case of modular forms and Maass forms, the automorphic distributions can be described quite concretely as boundary values. Alternatively but equivalently, they can be described abstractly; see [14, §2] or section three below. Computing with distributions presents some technical difficulties. What we gain in return are explicit formulas for the archimedean integrals that arise in the setting of automorphic distributions. This has led us to some new results.

In the next section we show how our method works in the simplest case, for the standard L-functions of modular forms and Maass forms. We treat the Rankin–Selberg L-function in section four, following the description of our main analytic tool in section three. Section five, finally, is devoted to the exterior square L-function for $GL(4, \mathbb{Z})$. That is the first not-entirely-trivial case of the main result of [15]. It can be explained more transparently than the general case for two reasons: the main analytic tool is the pairing of distributions, which for $GL(4)$ reduces to a variant of the Rankin–Selberg method for $GL(2)$. Also, the general case involves a somewhat subtle induction, with $GL(4)$ representing merely the initial step.

2 Standard L-functions for $SL(2)$

Holomorphic functions on the disk or the upper half plane have hyperfunction boundary values, essentially by definition of the notion of hyperfunction. Holomorphic functions of moderate growth, in particular modular forms, have distribution boundary values:

$$\tau(x) = \lim_{y \to 0^+} F(x + iy) \tag{2.1}$$

is the automorphic distribution corresponding to a modular form F for $SL(2,\mathbb{Z})$, of weight k. The limit exists in the strong distribution topology. From F, the distribution τ inherits its $SL(2,\mathbb{Z})$-automorphy property

$$\tau(x) = (cx+d)^{-k}\,\tau\left(\frac{ax+b}{cx+d}\right) \quad \text{for all} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})\,. \tag{2.2}$$

In terms of Fourier expansion (1.1) of the cuspidal modular form $F(z)$, the limit (2.1) can be taken term-by-term,

$$\tau(x) = \sum_{n>0} a_n\, e(nx)\,. \tag{2.3}$$

We shall argue next that it makes sense to take the Mellin transform of the distribution τ, and that this Mellin is an entire function of the variable s. The argument will be a special case of the techniques developed in our paper [13].

Note that the periodic distribution τ has no constant term. It can therefore be expressed as the ℓ-th derivative of a continuous, periodic function ϕ_ℓ, for every sufficiently large integer ℓ,

$$\begin{gathered}\tau(x) = \phi_\ell^{(\ell)}(x)\,, \quad \text{with} \;\; \phi_\ell \in C(\mathbb{R}/\mathbb{Z}) \\ \left(\phi_\ell(x) = \sum_{n>0} (2\pi i n)^{-\ell}\, a_n\, e(nx)\right).\end{gathered} \tag{2.4}$$

Using the formal rule for pairing the "test function" x^{s-1} against the derivative of a distribution, we find

$$\int_0^\infty x^{s-1}\,\tau(x)\,dx = \int_0^\infty x^{s-1}\frac{d^\ell}{dx^\ell}\,\phi_\ell(x)\,dx = (-1)^\ell \int_0^\infty \phi_\ell(x)\,\frac{d^\ell}{dx^\ell}\,x^{s-1}\,dx\,. \tag{2.5}$$

As a continuous, periodic function, ϕ_ℓ is bounded. That makes the expression on the right in (2.5) integrable away from $x = 0$, provided $\ell > \operatorname{Re} s$. Indeed, if we multiply the Mellin kernel x^{s-1} by a cutoff function $\psi \in C^\infty(\mathbb{R})$, with $\psi(x) \equiv 1$ near $x = \infty$ and $\psi(x) \equiv 0$ near $x = 0$, the resulting integral is an entire function of the variable s – we simply choose ℓ larger than the real part of any particular s. Increasing the value of ℓ further does not affect the integral, as can be seen by a legitimate application of integration by parts. The identity (2.2), with $a = d = 0$, $b = -c = 1$, gives

$$\tau(x) = (-x)^{-k}\,\tau(-1/x)\,, \tag{2.6}$$

so the behavior of $\tau(x)$ near zero duplicates its behavior near ∞, except for the factor $(-x)^k$ which can be absorbed into the Mellin kernel. The expression on the right in (2.5) is therefore integrable even down to zero, and

$$s \mapsto \int_0^\infty \tau(x)\,x^{s-1}\,dx \;\; \text{is a well defined, entire holomorphic function.} \tag{2.7}$$

The change of variables $x \mapsto 1/x$ and the transformation law (2.6) imply

$$\int_0^\infty \tau(x)\, x^{s-1}\, dx = (-1)^k \int_0^\infty \tau(-x)\, x^{k-s-1}\, dx\,. \tag{2.8}$$

The integral on the right is of course well defined, for the same reason as the integral (2.7).

In view of the argument we just sketched, it is entirely legitimate to replace $\tau(x)$ by its Fourier series and to interchange the order of integration and summation: for $\operatorname{Re} s \gg 0$,

$$\begin{aligned} \int_0^\infty \tau(x)\, x^{s-1}\, dx &= \int_0^\infty \sum_{n>0} a_n\, e(nx)\, x^{s-1}\, dx \\ &= \sum_{n>0} a_n \int_0^\infty e(nx)\, x^{s-1}\, dx \\ &= L\left(s - \frac{k-1}{2}, F\right) \int_0^\infty e(x)\, x^{s-1}\, dx\,; \end{aligned} \tag{2.9}$$

recall (1.2). The integral on the right makes sense for $\operatorname{Re} s > 0$ if one regards $e(x)$ as a distribution and applies integration by parts, as was done in the case of $\tau(x)$. In the range $0 < \operatorname{Re} s < 1$ it converges conditionally. This integral is well known,

$$\int_0^\infty e(x)\, x^{s-1}\, dx = (2\pi)^{-s}\, \Gamma(s)\, e(s/4) \qquad (0 < \operatorname{Re} s < 1). \tag{2.10}$$

Since $\Gamma(s)e(s/4)$ has no zeroes, (2.7) and (2.9–10) imply that $L(s, F)$ is entire. Replacing $\tau(x)$ by $\tau(-x)$ in (2.9) has the effect of replacing $e(x)$ by $e(-x)$, and accordingly the factor $e(s/4)$ by $e(-s/4)$ in (2.10). Thus (2.7–10) imply

$$\begin{aligned} &(2\pi)^{-s}\, e(s/4)\, \Gamma(s)\, L\left(s - \frac{k-1}{2}, F\right) \\ &\quad = (-1)^k\, (2\pi)^{s-k}\, e((s-k)/4)\, \Gamma(k-s)\, L\left(1 - s + \frac{k+1}{2}, F\right). \end{aligned} \tag{2.11}$$

Since $e(-k/4) = i^{-k}$, this functional equation is equivalent to the functional equation stated in (1.5).

A Maass form is a Γ-invariant eigenfunction $F \in C^\infty(\mathcal{H})$ for the hyperbolic Laplacian Δ, of moderate growth towards the boundary of $\mathcal{H}$. It is convenient to express the eigenvalue as $(\lambda^2 - 1)/4$, so that

$$y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) F = \frac{\lambda^2 - 1}{4}\, F\,. \tag{2.12}$$

Near the real axis, the Maass form F has an asymptotic expansion,

$$F(x+iy) \sim y^{\frac{1-\lambda}{2}} \sum_{k\geq 0} \tau_{\lambda,k}(x)\, y^{2k} + y^{\frac{1+\lambda}{2}} \sum_{k\geq 0} \tau_{-\lambda,k}(x)\, y^{2k} \tag{2.13}$$

as y tends to zero from above, with distribution coefficients $\tau_{\pm\lambda,k}$. In the exceptional case $\lambda = 0$, the leading terms $y^{(1-\lambda)/2}$, $y^{(1+\lambda)/2}$ must be replaced by, respectively, $y^{1/2}$ and $y^{1/2}\log y$. The leading coefficients

$$\tau_\lambda =_{\text{def}} \tau_{\lambda,0}\,, \qquad \tau_{-\lambda} =_{\text{def}} \tau_{-\lambda,0} \tag{2.14}$$

determine the others recursively. They are the *automorphic distributions* corresponding to the Maass form F. Each of the two also determines the other – in a way we shall explain later – unless λ is a negative odd integer, in which case the $\tau_{-\lambda,k}$ all vanish identically. To avoid making statements with trivial counterexamples, we shall not consider $\tau_{-\lambda}$ when $\lambda \in \mathbb{Z}_{<0} \cap (2\mathbb{Z}+1)$, and for $\lambda = 0$, we shall only consider the coefficient of $y^{1/2}$, not the coefficient of $y^{1/2}\log y$.

Unlike modular forms, Maass forms are Γ-invariant as functions, i.e., without a factor of automorphy. However, because of the nature of the asymptotic expansion (2.13), the Γ-invariance of F translates into an automorphy condition on the automorphic distributions,

$$\tau_\lambda(x) = |cx+d|^{\lambda-1}\, \tau_\lambda\!\left(\frac{ax+b}{cx+d}\right) \quad \text{for all} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma\,. \tag{2.15}$$

To simplify the discussion, we suppose $\Gamma = SL(2,\mathbb{Z})$, as before. Then (2.15), with $a = b = d = 1$, $c = 0$ implies $\tau_\lambda(x) \equiv \tau_\lambda(x+1)$, so τ_λ has a Fourier expansion

$$\tau_\lambda(x) = \sum_{n\in\mathbb{Z}} a_n\, e(nx)\,. \tag{2.16}$$

From the point of view of L-functions, cuspidal Maass forms are more interesting than non-cuspidal forms. The condition of cuspidality on F is equivalent to two conditions on the automorphic distribution τ_λ, namely

$$a_0 = 0\,, \quad \text{and } \tau_\lambda \text{ vanishes to infinite order at } x = 0 \tag{2.17}$$

[13]. To explain the meaning of the second condition, we note that the discussion leading up to (2.4) applies also in the present context, since $a_0 = 0$. The automorphy condition (2.15), with $a = d = 0$, $b = -c = -1$, asserts

$$\tau_\lambda(x) = |x|^{\lambda-1}\, \tau_\lambda(-1/x)\,. \tag{2.18}$$

Combined with (2.4) and the chain rule for the change of variables $x \mapsto -1/x$, this implies

$$\tau_\lambda(x) = |x|^{\lambda-1} \left(x^2 \frac{d}{dx}\right)^{\ell} (\phi_\ell(-1/x)) \ \text{ on } \mathbb{R} - \{0\}\,, \tag{2.19}$$

for every sufficiently large $\ell \in \mathbb{N}$, with some $\phi_\ell \in C(\mathbb{R} - \{0\})$ which remains bounded as $|x| \to \infty$. Moving the factor $|x|^{\lambda-1}$ across the differential operator and keeping track of the powers of x shows that the right hand side of (2.19) defines a distribution even on a neighborhood of the origin – a distribution with the remarkable property that for each $m \in \mathbb{N}$ it can be expressed, locally near $x = 0$, as

$$x^m P_m\left(x\,\frac{d}{dx}\right)\psi_m(x), \text{ with } \psi_m \text{ defined and continuous near the origin;} \qquad (2.20)$$

here P_m denotes a complex polynomial, whose coefficients depend on m and λ. In [13] we introduced the terminology *vanishing to infinite order* at $x = 0$ for the property (2.20) of a distribution defined on a neighborhood of the origin in $\mathbb{R}$.

To summarize the discussion so far, we have shown that a distribution τ_λ satisfying the automorphy condition (2.15) for $\Gamma = SL(2,\mathbb{Z})$, and additionally the condition $a_0 = 0$, agrees on $\mathbb{R} - \{0\}$ with a distribution that vanishes to infinite order at $x = 0$. Thus either τ_λ itself vanishes to infinite order at $x = 0$ – this is the meaning of the second condition in (2.17), of course – or else differs from such a distribution by one with support at the origin. A distribution supported at the origin is a linear combination of the delta function and its derivatives, and cannot vanish to infinite order at $x = 0$ unless it is identically zero. If, contrary to our standing hypothesis, Γ is a congruence subgroup of $SL(2,\mathbb{Z})$, the conditions (2.17) must be imposed at each of the cusps of Γ. In that case the second condition (2.17) must also be stated slightly differently.

If $F(x+iy)$ is a Maass form, then so is $F(-x+iy)$. It therefore makes sense to speak of even and odd Maass forms, i.e., Maass forms such that $F(-x+iy) = \pm F(x+iy)$. Every Maass form can be expressed uniquely as the sum of an even and an odd Maass form. If F is cuspidal, then so are the even and odd parts. The parity of F affects the Gamma factors in the functional equation of $L(s,F)$. We shall therefore suppose that F, and hence also τ_λ, has a definite parity,

$$\begin{aligned} \tau_\lambda(-x) &= (-1)^\eta\,\tau_\lambda(x)\,, \quad \text{or equivalently} \\ a_{-n} &= (-1)^\eta\,a_n \quad \text{for all } n \quad (\eta \in \mathbb{Z}/2\mathbb{Z})\,. \end{aligned} \qquad (2.21)$$

We also suppose that F is cuspidal, so that τ_λ satisfies (2.17). As one consequence of the parity condition, the L-function

$$L(s,F) = \sum_{n\geq 1} a_n\, n^{-s+\frac{\lambda}{2}} \qquad (\operatorname{Re} s \gg 0) \qquad (2.22)$$

completely determines all the a_n, and therefore also τ_λ and F. We had remarked earlier that τ_λ and $\tau_{-\lambda}$ play essentially symmetric roles unless λ is a negative odd integer or $\lambda = 0$. Outside of those exceptional cases, the Fourier coefficients of τ_λ and $\tau_{-\lambda}$ are related by the factor of proportionality $c_\lambda |n|^\lambda$, with $c_\lambda \neq 0$. Switching τ_λ and $\tau_{-\lambda}$ has the minor effect of renormalizing the L-function (2.22) by the non-zero constant c_λ. It is not difficult to eliminate the remaining ambiguity in normalizing $L(s,F)$, but we shall not pursue the matter here.

Arguing exactly as in the case of a modular form, we see that the *signed Mellin transform*

$$M_\eta(s,\tau_\lambda) = \int_{\mathbb{R}} \tau_\lambda(x)\,(\operatorname{sgn} x)^\eta\,|x|^{s-1}\,dx \tag{2.23}$$

is a well defined entire holomorphic function. It is legitimate to substitute the Fourier series (2.16) for τ_λ and to interchange the order of summation and integration, again for the same reasons as in the case of modular forms, hence

$$\begin{aligned} M_\eta(s,\tau_\lambda) &= 2\sum_{n\geq 1} a_n\, n^s \int_{\mathbb{R}} e(x)\,(\operatorname{sgn} x)^\eta\,|x|^{s-1}\,dx \\ &= 2\,G_\eta(s)\,L\left(s+\frac{\lambda}{2},F\right), \end{aligned} \tag{2.24}$$

with

$$G_\eta(s) = \int_{\mathbb{R}} e(x)\,(\operatorname{sgn} x)^\eta\,|x|^{s-1}\,dx = \begin{cases} \frac{2\,\Gamma(s)}{(2\pi)^s}\cos(\frac{\pi s}{2}) & \text{if } \eta = 0 \\ \frac{2\,i\,\Gamma(s)}{(2\pi)^s}\sin\left(\frac{\pi s}{2}\right) & \text{if } \eta = 1\,; \end{cases} \tag{2.25}$$

the explicit formula for $G_\eta(s)$ follows from (2.10). Since $M_\eta(s,F)$ is entire, (2.24–25) show that $L(s,F)$ extends meromorphically to the entire s-plane.

The change of variables $x \mapsto -1/x$ in (2.23), combined with the transformation rule (2.18), gives the functional equation

$$M_\eta(s,\tau_\lambda) = (-1)^\eta\,M_\eta(1-s-\lambda,\tau_\lambda)\,, \tag{2.26}$$

which in turn implies the functional equation

$$G_\eta\left(s-\frac{\lambda}{2}\right)\,L(s,F) = (-1)^\eta\,G_\eta\left(1-s-\frac{\lambda}{2}\right)\,L(1-s,F) \tag{2.27}$$

for $L(s,F)$. Standard Gamma identities establish the equivalence between Maass' version of the functional equation and (2.27).

Though we know that the product $G_\eta\big(s-\frac{\lambda}{2}\big)L(s,F)$ is entire, we cannot yet conclude that $L(s,F)$ is also entire: unlike $\Gamma(s)$, $G_\eta(s)$ has zeroes. To deal with this problem, we consider the Fourier transform $\widehat{\tau_\lambda}$ of the tempered distribution τ_λ. We use the normalization $\widehat{f}(y) = \int_{\mathbb{R}} f(x)e(-xy)dx$. Then $e(nx)$, considered as tempered distribution, has Fourier transform $\mathcal{F}e(nx) = \delta_n(x) =$ Dirac delta function at $x = n$, and

$$\widehat{\tau_\lambda}(x) = \sum_{n\neq 0} a_n\,\delta_n(x)\,. \tag{2.28}$$

This distribution visibly vanishes in a neighborhood of the origin, in particular vanishes to infinite order at $x = 0$. According to [13, theorem 3.19], the fact that τ_λ vanishes to infinite order at $x = 0$ – cf. (2.17) – implies that $\widehat{\tau_\lambda}(1/x)$ extends across the origin to a distribution that vanishes there to infinite order. Since both $\widehat{\tau_\lambda}(x)$ and $\widehat{\tau_\lambda}(1/x)$ have this property, the signed Mellin transform

$$\begin{aligned} M_\eta(s, \widehat{\tau_\lambda}) &= 2 \sum_{n>0} a_n\, n^{s-1} \qquad (\operatorname{Re} s \ll 0) \\ &= 2\, L\left(1 - s + \frac{\lambda}{2}, F\right) \end{aligned} \tag{2.29}$$

is a well defined, entire holomorphic function. In other words, $L(s, F)$ is entire, as was to be shown.

The preceding argument essentially applies also to the case of modular forms, except that one is then dealing with automorphic distributions that are neither even nor odd, but have only positive Fourier coefficients. In fact, if one considers modular forms and Maass forms not for $SL(2)$ but for $GL(2)$, a single argument treats both types of automorphic distributions absolutely uniformly. However, the case of modular forms is simpler in one respect: the fact that the L-function has no poles requires no special argument.

3 Pairings of automorphic distributions

In the last section we encountered automorphic distributions as distributions on the real line, obtained by a limiting process. For higher rank groups, it is necessary to take a more abstract point of view, which we shall now explain.

Initially in this section G shall denote a reductive Lie group, Z_G^0 the identity component in the center Z_G of G, and $\Gamma \subset G$ an arithmetically defined subgroup. Note that G acts unitarily on $L^2(\Gamma\backslash G/Z_G^0)$, via right translation. We consider an irreducible unitary representation (π, V) of G which occurs discretely in $L^2(\Gamma\backslash G/Z_G^0)$,

$$j : V \hookrightarrow L^2(\Gamma\backslash G/Z_G^0)\,. \tag{3.1}$$

Recall the notion of a C^∞ vector for π: a vector $v \in V$ such that $g \mapsto \pi(g)v$ is a C^∞ map from G to the Hilbert space V. The space of C^∞ vectors $V^\infty \subset V$ is dense, G-invariant, and gets mapped to $C^\infty(\Gamma\backslash G/Z_G^0)$ by the embedding (3.1). That makes

$$\tau = \tau_j : V^\infty \longrightarrow \mathbb{C}\,, \qquad \tau(v) = jv(e)\,, \tag{3.2}$$

a well defined linear map. It is Γ-invariant because $jv \in C^\infty(\Gamma\backslash G/Z_G^0)$, and is continuous with respect to the natural topology on V^∞. One should therefore think of τ as a Γ-invariant distribution vector for the dual representation (π', V') – i.e., $\tau \in ((V')^{-\infty})^\Gamma$. Very importantly, τ determines j completely. Indeed, j is G-invariant, so the defining identity (3.2) specifies the value of jv, $v \in V^\infty$, not only at the identity, but at any $g \in G$. Since V^∞ is dense in V, knowing the effect of j on V^∞ means knowing j.

The space $L^2(\Gamma\backslash G/Z_G^0)$ is self-dual, hence if V occurs discretely, so does its dual V'. Since we shall be working primarily with τ, we switch the roles of V and V'. From now on,

$$\tau \in (V^{-\infty})^\Gamma \tag{3.3}$$

shall denote a Γ-invariant distribution vector corresponding to a discrete embedding $V' \hookrightarrow L^2(\Gamma\backslash G/Z_G^0)$. Not all Γ-invariant distribution vectors correspond to embeddings into $L^2(\Gamma\backslash G/Z_G^0)$; some correspond to Eisenstein series, and others not even to those.

The arithmetically defined subgroup Γ is arithmetic with respect to a particular $\mathbb{Q}$-structure on G. If $P \subset G$ is a parabolic subgroup, defined over $\mathbb{Q}$, with unipotent radical U, then $\Gamma \cap U$ is a lattice in U; in other words, the quotient $U/(\Gamma \cap U)$ is compact. One calls $\tau \in (V^{-\infty})^\Gamma$ cuspidal if

$$\int_{U/(\Gamma\cap U)} \pi(u)\tau\, du = 0\,, \tag{3.4}$$

for the unipotent radical U of any parabolic subgroup P that is defined over $\mathbb{Q}$. Since there exist only finitely many Γ-conjugacy classes of such parabolics, cuspidality amounts to only finitely many conditions. Essentially by definition, cuspidal embeddings $V' \hookrightarrow L^2(\Gamma\backslash G/Z_G^0)$ correspond to cuspidal distribution vectors $\tau \in (V^{-\infty})^\Gamma$, and conversely every cuspidal automorphic τ arises from a cuspidal embedding of V' into $L^2(\Gamma\backslash G/Z_G^0)$.

To get a handle on $\tau \in (V^{-\infty})^\Gamma$, we realize the space of C^∞ vectors V^∞ as a subspace $V^\infty \hookrightarrow V^\infty_{\lambda,\delta}$ of the space of C^∞ vectors for a not-necessarily-unitary principal series representation $(\pi_{\lambda,\delta}, V_{\lambda,\delta})$. The Casselman embedding theorem [3] guarantees the existence of such an embedding. For the moment, we leave the meaning of the subscripts λ, δ undefined. They are the parameters of the principal series, which we shall explain presently in the relevant cases. A theorem of Casselman-Wallach [3, 22] asserts that the inclusion $V^\infty \hookrightarrow V^\infty_{\lambda,\delta}$ extends continuously to an embedding of the space of distribution vectors,

$$V^{-\infty} \hookrightarrow V^{-\infty}_{\lambda,\delta}\,. \tag{3.5}$$

This allows us to consider the automorphic distribution τ as a distribution vector for a principal series representation,

$$\tau \in (V^{-\infty}_{\lambda,\delta})^\Gamma\,. \tag{3.6}$$

When $G = SL(2,\mathbb{R})$, cuspidal modular forms correspond to embeddings of discrete series representations into $L^2(\Gamma\backslash G)$, and cuspidal Maass forms to embeddings of unitary principal series or complementary series representations. The realization of discrete series representations of $SL(2,\mathbb{R})$ as subrepresentations of principal series representations is very well known, making (3.6) quite concrete. For general groups, the Casselman embeddings cannot be described equally explicitly, nor do they need to be unique. Those are not obstacles to using (3.6) in studying L-functions. In fact, the non-uniqueness is sometimes helpful in ruling out poles of L-functions.

Our tool in studying Rankin–Selberg and related L-functions is the pairing of automorphic distributions. In this paper, we shall only discuss Rankin–Selberg L-functions for $GL(2)$ and the exterior square L-function for $GL(4)$. Both involve the pairing of automorphic distributions of $GL(2)$. To minimize notational effort, we shall work with the group

$$G = PGL(2,\mathbb{R}) \cong SL^{\pm}(2,\mathbb{R})/\{\pm 1\}$$
$$(SL^{\pm}(2,\mathbb{R}) = \{g \in GL(2,\mathbb{R}) \mid \det g = \pm 1\})\,, \tag{3.7}$$

rather than $G = GL(2,\mathbb{R})$, for the remainder of this section. We let $B \subset G$ denote the lower triangular subgroup. For $\lambda \in \mathbb{C}$ and $\delta \in \mathbb{Z}/2\mathbb{Z}$, we define

$$\chi_{\lambda,\delta} : B \to \mathbb{C}^*\,, \quad \chi_{\lambda,\delta}\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \left(\operatorname{sgn}\frac{a}{d}\right)^{\delta} \left|\frac{a}{d}\right|^{\frac{\lambda}{2}}\,. \tag{3.8}$$

The parameterization of the principal series involves a "ρ-shift", i.e., a shift by the half-sum of the positive roots. In our concrete setting

$$\rho = 1\,, \tag{3.9}$$

and we shall write $\chi_{\lambda-\rho,\delta}$ instead of $\chi_{\lambda-1,\delta}$ to be consistent with the usual notation in the subject. The space of C^∞ vectors for the principal series representation $\pi_{\lambda,\delta}$ is

$$V^{\infty}_{\lambda,\delta} = \{F \in C^\infty(G) \mid F(gb) = \chi_{\lambda-\rho,\delta}(b^{-1})F(g) \text{ for all } g \in G,\ b \in B\}, \tag{3.10}$$

with action

$$(\pi_{\lambda,\delta}(g)F)(h) = F(g^{-1}h) \qquad (F \in V^{\infty}_{\lambda,\delta}\,,\ \ g,h \in G)\,. \tag{3.11}$$

Quite analogously

$$V^{-\infty}_{\lambda,\delta} = \{\tau \in C^{-\infty}(G) \mid \tau(gb) = \chi_{\lambda-\rho,\delta}(b^{-1})\,\tau(g) \text{ for all } g \in G,\ b \in B\} \tag{3.12}$$

is the space of distribution vectors, on which G acts by the same formula as on $V^{\infty}_{\lambda,\delta}$.

The tautological action of $GL(2,\mathbb{R})$ on $\mathbb{R}^2$ induces a transitive action of $G = PGL(2,\mathbb{R})$ on $\mathbb{RP}^1$; in fact $\mathbb{RP}^1 \cong G/B$, since B is the isotropy subgroup at the line spanned by the second standard basis vector of $\mathbb{R}^2$. According to the so-called "fundamental theorem of projective geometry", the action of G on $\mathbb{RP}^1$ induces a simply transitive, faithful action on the set of triples of distinct points in $\mathbb{RP}^1 \times \mathbb{RP}^1 \times \mathbb{RP}^1$. Put differently, G has a dense open orbit in

$$\mathbb{RP}^1 \times \mathbb{RP}^1 \times \mathbb{RP}^1 \ \cong\ G/B \times G/B \times G/B\,, \tag{3.13}$$

and can be identified with that dense open orbit once a base point has been chosen. The three matrices

$$f_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad f_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{3.14}$$

lie in distinct cosets of B, so

$$G \hookrightarrow\ G/B \times G/B \times G/B\,, \quad g \mapsto\ (gf_1B,\ gf_2B,\ gf_3B)\,, \tag{3.15}$$

gives a concrete identification of G with its open orbit in $\mathbb{RP}^1 \times \mathbb{RP}^1 \times \mathbb{RP}^1$.

Formally at least, the existence of the open orbit can be used to define a G-invariant trilinear pairing

$$\begin{aligned} &V^{\infty}_{\lambda_1,\delta_1} \times V^{\infty}_{\lambda_2,\delta_2} \times V^{\infty}_{\lambda_3,\delta_3} \longrightarrow \mathbb{C}\,, \\ &(F_1, F_2, F_3) \mapsto P(F_1, F_2, F_3) =_{\mathrm{def}} \int_G F_1(gf_1)\, F_2(gf_2)\, F_3(gf_3)\, dg\,, \end{aligned} \tag{3.16}$$

between any three principal series representations. Although the G-invariance of the pairing is obvious from this formula, it is not clear that the integral converges. Before addressing the question of convergence, we should remark that the "fundamental theorem of projective geometry" is field-independent. The same ideas have been used to construct triple pairings for representations of $PGL(2, \mathbb{Q}_p)$. We should also point out that a different choice of base points f_j would have the effect of multiplying the pairing by a non-zero constant.

The question of convergence of the integral (3.16) is most easily understood in terms of the "unbounded realization" of the principal series, which we discuss next. The subgroup

$$N = \left\{ n_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R} \right\} \cong \mathbb{R} \tag{3.17}$$

of G acts freely on G/B, and its image omits only a single point, the coset of

$$s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.18}$$

It follows that any $F \in V^{\infty}_{\lambda,\delta}$ is completely determined by its restriction to $N \cong \mathbb{R}$; the defining identities (3.8–10) imply that ϕ_0 = restriction of F to $\mathbb{R}$ is related to ϕ_∞ = restriction of $\pi_{\lambda,\delta}(s)F$ to $\mathbb{R}$ by the identity $\phi_\infty(x) = |x|^{\lambda-1}\phi(-1/x)$. This leads naturally to the identification

$$V^{\infty}_{\lambda,\delta} \cong \{\phi \in C^\infty(\mathbb{R}) \mid |x|^{\lambda-1}\,\phi(-1/x) \in C^\infty(\mathbb{R})\}\,, \tag{3.19}$$

with action

$$\begin{aligned} (\pi_{\lambda,\delta}(g)\phi)(x) = (\mathrm{sgn}(ad-bc))^{\delta} \left(\frac{|cx+d|}{\sqrt{|ad-bc|}} \right)^{\lambda-1} \phi\left(\frac{ax+b}{cx+d} \right) \\ \text{for } g^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G\,. \end{aligned} \tag{3.20}$$

If $(\phi_1, \phi_2, \phi_3) \in (C^\infty(\mathbb{R}))^3$ correspond to $(F_1, F_2, F_3) \in V^{\infty}_{\lambda_1,\delta_1} \times V^{\infty}_{\lambda_2,\delta_2} \times V^{\infty}_{\lambda_3,\delta_3}$ via the unbounded realization (3.19),

$$\begin{aligned} &P(F_1, F_2, F_3) = \int_{\mathbb{R}^3} \phi_1(x)\, \phi_2(y)\, \phi_3(z)\, k(x,y,z)\, dx\, dy\, dz\,, \quad \text{with} \\ &k(x,y,z) = \mathrm{sgn}((x-y)(y-z)(z-x))^{\delta_1+\delta_2+\delta_3} \\ &\qquad \times |x-y|^{\frac{-\lambda_1-\lambda_2+\lambda_3-1}{2}}\, |y-z|^{\frac{\lambda_1-\lambda_2-\lambda_3-1}{2}}\, |x-z|^{\frac{-\lambda_1+\lambda_2-\lambda_3-1}{2}}\,. \end{aligned} \tag{3.21}$$

This can be seen from the explicit form of the isomorphism (3.19), coupled with the definition (3.10) of $V^{\infty}_{\lambda,\delta}$. We should point out that in the setting of Maass forms, δ plays the role of the parity η in (2.21).

Contrary to appearance, the integral (3.21) is really an integral over the compact space $\mathbb{RP}^1 \times \mathbb{RP}^1 \times \mathbb{RP}^1$: the integral retains the same general form when one or more of the coordinates x, y, z are replaced by their reciprocals; this follows from the behavior of the ϕ_j at ∞ specified in (3.19). The convergence of the integral is therefore a purely local matter. Near points where exactly two of the coordinates coincide, absolute convergence is guaranteed when the real part of the corresponding exponent is greater than -1. To analyze the convergence near points of the triple diagonal $\{x = y = z\}$, it helps to "blow up" the triple diagonal in the sense of real algebraic geometry – or equivalently, to use polar coordinates in the normal directions. One then sees that absolute convergence requires not only the earlier condition

$$\operatorname{Re}(\lambda_i - \lambda_j - \lambda_k) > -1 \quad \text{if } i \neq j,\ j \neq k,\ k \neq i\,, \tag{3.22}$$

but also

$$\operatorname{Re}(\lambda_1 + \lambda_2 + \lambda_3) < 1\,. \tag{3.23}$$

Both conditions certainly hold when the V_{λ_i,δ_i} belong to the *unitary principal series*, i.e., when all the λ_j are purely imaginary.

The argument we have sketched establishes the existence of an invariant trilinear pairing between the spaces of C^∞ vectors of any three unitary principal series representations. The pairing is known to be unique up to scaling [16]. Even when the λ_i are not purely imaginary, one can use (3.21) to exhibit an invariant trilinear pairing by meromorphic continuation. Indeed, for compactly supported functions of one variable, the functional $f \mapsto \int_{\mathbb{R}} f(x)|x|^{s-1}dx$ extends meromorphically to $s \in \mathbb{C}$, with first order poles at the non-positive integers, but no other poles. As was just argued, the integral kernel in (3.21) can be expressed as $|u|^s$ or $|u|^{s_1}|v|^{s_2}$, in terms of suitable local coordinates, after blowing up when necessary. Localizing the problem as before, by means of a suitable partition of unity, one can therefore assign a meaning to the integral (3.21) for all triples $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{C}^3$ outside certain hyperplanes, where the integral has poles. Even for parameters $(\lambda_1, \lambda_2, \lambda_3)$ on these hyperplanes one can exhibit an invariant triple pairing by taking residues.

Let us now consider the datum of distribution vectors $\tau_j \in V^{-\infty}_{\lambda_j,\delta_j}$ for three principal series representations V_{λ_j,δ_j}, $1 \leq j \leq 3$. The unbounded realization of the $V^{-\infty}_{\lambda_j,\delta_j}$ is slightly more complicated than the C^∞ case (3.19): unlike a C^∞ function, a distribution is not determined by its restriction to a dense open subset of its domain. The distribution analogue of (3.19),

$$V^{-\infty}_{\lambda,\delta} \cong \{(\sigma_0, \sigma_\infty) \in (C^{-\infty}(\mathbb{R}))^2 \mid \sigma_\infty(x) = |x|^{\lambda-1}\,\sigma_0(-1/x)\}\,, \tag{3.24}$$

therefore involves a pair of distributions on $\mathbb{R}$ that determine each other on $\mathbb{R} - \{0\}$. Suppose now that $\tau_j \cong (\sigma_{j,0}, \sigma_{j,\infty})$ via (3.24). Then

$$\begin{aligned}(x, y, z) \mapsto\ & \sigma_{1,0}(x)\,\sigma_{2,0}(y)\,\sigma_{3,0}(z)\,\operatorname{sgn}((x-y)(y-z)(z-x))^{\delta_1+\delta_2+\delta_3} \\ & \times |x-y|^{\frac{-\lambda_1-\lambda_2+\lambda_3+1}{2}}\,|y-z|^{\frac{\lambda_1-\lambda_2-\lambda_3+1}{2}}\,|x-z|^{\frac{-\lambda_1+\lambda_2-\lambda_3+1}{2}}\end{aligned} \tag{3.25}$$

extends naturally to a distribution on $\{(x,y,z) \in (\mathbb{RP}^1)^3 \mid x \neq y \neq z \neq x\}$; as one or more of the coordinates tend to ∞, one replaces those coordinates by the negative of their reciprocals, and simultaneously the corresponding $\sigma_{j,0}$ by $\sigma_{j,\infty}$. Since $\{(x,y,z) \in (\mathbb{RP}^1)^3 \mid x \neq y \neq z \neq x\} \cong G$ via the identification (3.15), we may regard (3.25) as a distribution on G. In fact, this distribution is

$$\{g \;\mapsto\; \tau_1(gf_1)\,\tau_2(gf_2)\,\tau_3(gf_3)\} \;\in\; C^{-\infty}(G)\,, \tag{3.26}$$

although the latter description has no immediately obvious meaning without the steps we have just gone through. The apparent discrepancy between the signs in the exponents in (3.21) and (3.25) reflects the fact that

$$|x-y|^{-1}\,|y-z|^{-1}\,|z-x|^{-1}\,dx\,dy\,dz \;\cong\; dg = \text{Haar measure on } G \tag{3.27}$$

via the identification (3.15). Let us formally record the substance of our discussion:

3.28 Observation. *For* $\tau_j \in V^{-\infty}_{\lambda_j,\delta_j}\,,\ 1 \leq j \leq 3\,,$

$$g \;\mapsto\; \tau_1(gf_1)\,\tau_2(gf_2)\,\tau_3(gf_3)$$

is a well defined distribution on G.

To motivate our result on pairings of automorphic distributions, we temporarily deviate from our standing assumption that $\Gamma \subset G$ be arithmetically defined; instead we suppose that $\Gamma \subset G$ is a discrete, cocompact subgroup. In that case, if $\tau_j \in (V^{-\infty}_{\lambda_j,\delta_j})^\Gamma$, $1 \leq j \leq 3$, are Γ-invariant distribution vectors, (3.26) defines a distribution on the compact manifold $\Gamma\backslash G$. As such, it can be integrated against the constant function 1, and

$$\int_{\Gamma\backslash G} \tau_1(gf_1)\,\tau_2(gf_2)\,\tau_3(gf_3)\,dg \tag{3.29}$$

has definite meaning. The value of the integral remains unchanged when the variable of integration g is replaced by gh, for any particular $h \in G$. Thus, if $\psi \in C_c^\infty(G)$ has total integral one,

$$\begin{aligned}
&\int_{\Gamma\backslash G} \tau_1(gf_1)\,\tau_2(gf_2)\,\tau_3(gf_3)\,dg \\
&\quad = \int_G \int_{\Gamma\backslash G} \tau_1(ghf_1)\,\tau_2(ghf_2)\,\tau_3(ghf_3)\,\psi(h)\,dg\,dh \\
&\quad = \int_{\Gamma\backslash G} \left(\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\tau_3(ghf_3)\,\psi(h)\,dh \right) dg\,.
\end{aligned} \tag{3.30}$$

The implicit use of Fubini's theorem at the second step can be justified by a partition of unity argument. In short, we have expressed the integral (3.29) as the integral over $\Gamma\backslash G$ of the Γ-invariant function

$$g \;\mapsto\; \int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\tau_3(ghf_3)\,\psi(h)\,dh\,. \tag{3.31}$$

This function is smooth, like any convolution of a distribution with a compactly supported C^∞ function. Note that the integral (3.31) is well defined even for parameters $(\lambda_1,\lambda_2,\lambda_3)\in\mathbb{C}^3$ which correspond to poles of the integral (3.21).

We now return to our earlier setting, of an arithmetically defined subgroup $\Gamma\subset G=PGL(2,\mathbb{R})$, specifically a congruence subgroup

$$\Gamma \;\subset\; PGL(2,\mathbb{Z})\,. \tag{3.32}$$

In this context, the integral (3.29) has no obvious meaning, since we would have to integrate a distribution over the noncompact manifold $\Gamma\backslash G$. The "smoothed" integral, however, potentially makes sense: if the integrand (3.31) can be shown to decay rapidly towards the cusps of $\Gamma\backslash G$, it is simply an ordinary, convergent integral. That is the case, under appropriate hypotheses:

3.33 Theorem. *Let $\tau_j\in(V^{-\infty}_{\lambda_j,\delta_j})^\Gamma$, $1\le j\le 3$, be Γ-automorphic distributions, and $\psi\in C_c^\infty(G)$ a test function, subject to the normalizing condition*

$$\int_G \psi(g)\,dg = 1\,.$$

If at least one of the τ_j is cuspidal, the Γ-invariant C^∞ function

$$F(g) = \int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\tau_3(ghf_3)\,\psi(h)\,dh$$

decays rapidly along the cusps of Γ; in particular $\int_{\Gamma\backslash G}F(g)\,dg$ converges absolutely. This integral does not depend on the specific choice of ψ. If, in addition, one of the τ_j depends holomorphically on a complex parameter,

$$\int_{\Gamma\backslash G}F(g)\,dg = \int_{\Gamma\backslash G}\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\tau_3(ghf_3)\,\psi(h)\,dh\,dg$$

also depends holomorphically on that parameter.

Why does F decay rapidly? It is not a modular form – the Casimir operator of G does not act on it finitely. Nor does F satisfy the condition of cuspidality. However, F can be expressed as the restriction to the diagonal of a modular form in three variables:

$$(g_1,g_2,g_3)\mapsto \int_G \tau_1(g_1hf_1)\,\tau_2(g_2hf_2)\,\tau_3(g_3hf_3)\,\psi(h)\,dh \tag{3.34}$$

is a C^∞ function on $G\times G\times G$; this follows from the fact that the cosets f_jB lie in general position. Since $\tau_j\in(V^{-\infty}_{\lambda_j,\delta_j})^\Gamma$, (3.34) is a Γ-invariant eigenfunction of the Casimir operator in each of the variables separately. It is cuspidal in the variable corresponding to the cuspidal factor τ_j, hence decays rapidly in this one direction. It has at worst moderate growth in the other directions, and therefore decays rapidly when restricted to the diagonal. The remaining assertions of the lemma are relatively straightforward.

We shall need a variant of the theorem in the last section, for the analysis of the exterior square L-function for $GL(4)$. Two of the τ_j then occur coupled, as a distribution vector for a principal series representation of $G \times G$, Γ-invariant only under the diagonal action, not separately. These two τ_j arise from a single cuspidal automorphic distribution τ for $GL(4,\mathbb{R})$. In this situation the rapid decay of F reflects the cuspidality of τ.

4 The Rankin–Selberg L-function for $GL(2)$

The argument we are about to sketch parallels the classical arguments of Rankin [17] and Selberg [18], and of Jacquet [6] in the case of Maass forms. We shall pair two automorphic distributions against an Eisenstein series. In our setting, of course, the Eisenstein series is also an automorphic distribution.

We recall the construction of the distribution Eisenstein series from [15], specialized to the case of $G = PGL(2,\mathbb{R})$. To simplify the discussion, we only work at full level – in other words, with

$$\Gamma = PGL(2,\mathbb{Z}) \simeq SL^{\pm}(2,\mathbb{Z})/\{\pm 1\}\,. \tag{4.1}$$

We define $\delta_\infty \in V_{\nu,0}^{-\infty}$ in terms of the unbounded realization (3.24): δ_∞ corresponds to $(\sigma_0, \sigma_\infty)$, with $\sigma_0 = 0$ and $\sigma_\infty =$ Dirac delta function at 0. Then $\pi_{\nu,0}(\gamma)\delta_\infty = \delta_\infty$ for all $\gamma \in \Gamma_\infty = \{\gamma \in \Gamma \mid \gamma\infty = \infty\}$. In particular, the series

$$E_\nu \in V_{\nu,0}^{-\infty}\,, \quad E_\nu = \zeta(\nu+1) \sum_{\gamma \in \Gamma/\Gamma_\infty} \pi_{\nu,0}(\gamma)\delta_\infty\,, \tag{4.2}$$

makes sense at least formally. It is Γ-invariant by construction. Hence, when we describe E_ν in terms of the unbounded realization (3.24), it suffices to specify the first member σ_0 of the pair $(\sigma_0, \sigma_\infty)$. This allows us to regard E_ν as a distribution on the real line,

$$E_\nu \simeq \sum_{p,q\in\mathbb{Z},\ q>0} q^{-\nu-1}\,\delta_{p/q}(x)\,. \tag{4.3}$$

To see the equivalence of (4.2) and (4.3), we note that $\delta_{p/q}(x)$, with $p,q \in \mathbb{Z}$ relatively prime, corresponds to the translate of δ_∞ under

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix} \in \Gamma\,, \quad \text{with } r,s \in \mathbb{Z} \text{ chosen so that } sp - rq = 1\,. \tag{4.4}$$

The disappearance of the factor $\zeta(\nu+1)$ in (4.3) reflects the fact that we now sum over all pairs of integers p, q, with $q > 0$, not over relatively prime pairs.

The integral of the series (4.3) against a compactly supported test function converges uniformly and absolutely when $\operatorname{Re}\nu > 1$. Hence $E_\nu \in V_{\nu,0}^{-\infty}$ is well defined for $\operatorname{Re}\nu > 1$, and depends holomorphically on ν in this region. The periodic distribution (4.3) has a Fourier expansion,

$$E_\nu \simeq \sum_{n\in\mathbb{Z}} a_n\, e(nx)\,. \tag{4.5}$$

To calculate the Fourier coefficients, we reinterpret the sum as a distribution on $\mathbb{R}/\mathbb{Z}$. Then

$$\begin{aligned} a_n &= \int_{\mathbb{R}/\mathbb{Z}} e(-nx) \sum_{p,q\in\mathbb{Z},\ q>0} q^{-\nu-1}\,\delta_{p/q}(x)\,dx \\ &= \sum_{q>0}\ \sum_{0\le p<q} q^{-\nu-1}\,e(-np/q) = \begin{cases} \sum_{d|n} d^{-\nu} & \text{if } n\neq 0 \\ \zeta(\nu) & \text{if } n=0\,. \end{cases} \end{aligned} \tag{4.6}$$

The a_n, $n \neq 0$, are entire functions of ν, whereas $a_0 = \zeta(\nu)$ has a pole at $\nu = 1$, so

$$\begin{gathered} E_\nu \ \text{ extends meromorphically to the entire complex plane,} \\ \text{with a single pole at } \nu = 1, \text{ of order one.} \end{gathered} \tag{4.7}$$

We should remark that δ_∞ is even with respect to the involution $x \mapsto -x$. This is the reason why at full level there is no Eisenstein series of odd parity – i.e., no Eisenstein series in $V_{\nu,1}^{-\infty}$.

The Eisenstein series (4.2) satisfies a functional equation, which relates $E_{-\nu} \in V_{-\nu,0}^{-\infty}$ to $E_\nu \in V_{\nu,0}^{-\infty}$ via the intertwining operator

$$J_\nu : V_{-\nu,0}^{-\infty} \longrightarrow V_{\nu,0}^{-\infty}\,. \tag{4.8}$$

On the level of C^∞ vectors, and in terms of the unbounded realization (3.19), the operator is given by the formula

$$(J_\nu\phi)(x) = \int_{\mathbb{R}} \phi(y)\,|y-x|^{\nu-1}\,dy\,. \tag{4.9}$$

Because of the condition on ϕ at infinity, this integral has no singularity at $y = \infty$. At $y = x$, the integral converges when $\operatorname{Re}\nu > 0$, but continues meromorphically to the entire complex plane. It is known that the integral transform (4.9) extends continuously from an operator $J_\nu : V_{-\nu,0}^{\infty} \to V_{\nu,0}^{\infty}$ between the spaces of C^∞ vectors, to the operator (4.8). Alternatively and equivalently, (4.8) can be defined as the adjoint of $J_\nu : V_{-\nu,0}^{\infty} \to V_{\nu,0}^{\infty}$, using the natural duality[1] between $V_{\nu,0}^{\infty}$ and $V_{-\nu,0}^{-\infty}$. Either way one sees that

$$V_{-\nu,0}^{-\infty} \ni e(nx) \xrightarrow{\ J_\nu\ } G_0(\nu)\,|n|^{-\nu}\,e(nx) \in V_{\nu,0}^{-\infty} \quad (n\neq 0)\,. \tag{4.10}$$

[1] The duality which extends the G-invariant pairing $V_{\nu,0}^{\infty} \times V_{-\nu,0}^{\infty} \to \mathbb{C}$ given by integration over $\mathbb{R}$, in terms of the unbounded realization.

Here $G_0(\nu)$ refers to the Gamma factor described in (2.25), and $e(nx)$ is shorthand for the pair $(e(nx), |x|^{\mp\nu-1}e(-n/x))$ – cf. (3.24); the second member of the pair can be given a definite meaning even at the origin, using the notion of *vanishing to infinite order* that was discussed in section 2.

In view of the relation (4.10), J_ν maps the Fourier series (4.5) for $E_{-\nu}$ to $G_0(\nu)$ times the corresponding series for E_ν, except possibly for the constant term and a distribution supported at infinity. However, no non-zero linear combination of a constant function and a distribution supported at infinity can be Γ-invariant. This proves

$$J_\nu E_{-\nu} = G_0(\nu)\, E_\nu\,. \tag{4.11}$$

That is the functional equation satisfied by the Eisenstein series. The parameter ν is natural from the point of view of representation theory. In the eventual application, we shall work with

$$s = (\nu + 1)/2 \tag{4.12}$$

instead. Note that $\nu \mapsto -\nu$ corresponds to $s \mapsto 1 - s$.

We now fix two automorphic distributions, either of which may arise from a modular form or a Maass form,

$$\tau_1 \in (V^{-\infty}_{\lambda_1,\delta_1})^\Gamma \;\text{ and }\; \tau_2 \in (V^{-\infty}_{\lambda_2,\delta_2})^\Gamma\,, \tag{4.13}$$

of which at least one is cuspidal. According to (4.7) and theorem 3.33, the integral

$$P^\Gamma_\nu(\tau_1, \tau_2, E_\nu) = \int_{\Gamma\backslash G}\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,E_\nu(ghf_3)\,\psi(h)\,dh\,dg \tag{4.14}$$

depends meromorphically on $\nu \in \mathbb{C}$, with a potential first order pole at $\nu = 1$ but no other singularities. The subscript ν is meant to emphasize the fact that the third argument lies in the space $(V^{-\infty}_{\nu,0})^\Gamma$, and the superscript Γ distinguishes this pairing of Γ-invariant distribution vectors from the pairing (3.21) between spaces of C^∞ vectors.

We shall derive the Rankin–Selberg functional equation from the functional equation (4.11) of the Eisenstein series. Since the latter involves the intertwining operator, we need to know how J_ν relates $P^\Gamma_{-\nu}$ to P^Γ_ν. First the analogous statement about the pairing (3.21): for $F_1 \in V^\infty_{\lambda_1,\delta_1}$, $F_2 \in V^\infty_{\lambda_2,\delta_2}$, $F_3 \in V^\infty_{-\nu,0}$,

$$\begin{aligned} &P(F_1, F_2, J_\nu F_3) \\ &\quad = (-1)^{\delta_1+\delta_2}\,\frac{G_{\delta_1+\delta_2}\left(\frac{\lambda_1-\lambda_2-\nu+1}{2}\right) G_{\delta_1+\delta_2}\left(\frac{-\lambda_1+\lambda_2-\nu+1}{2}\right)}{G_0(1-\nu)}\,P(F_1, F_2, F_3)\,. \end{aligned} \tag{4.15}$$

Note that $P(\dots)$ on the left and the right side of the equality refer to the pairing $V^\infty_{\lambda_1,\delta_1} \times V^\infty_{\lambda_2,\delta_2} \times V^\infty_{\nu,0} \to \mathbb{C}$, respectively $V^\infty_{\lambda_1,\delta_1} \times V^\infty_{\lambda_2,\delta_2} \times V^\infty_{-\nu,0} \to \mathbb{C}$. The Gamma factors $G_\delta(\dots)$ have the same meaning as in (2.25). Since both sides of the equality depend meromorphically on ν, it suffices to establish it for values of ν in some non-empty open region. In view of (3.21) and (4.9), the assertion (4.15) reduces to the identity

$$\begin{aligned}&\int_{\mathbb{R}} \frac{(\operatorname{sgn}(y-t)(t-x))^{\delta_1+\delta_2}}{(\operatorname{sgn}(y-z)(z-x))^{\delta_1+\delta_2}}\,|x-t|^{\alpha-1}\,|y-t|^{\beta-1}\,|z-t|^{-\alpha-\beta}\,dt \\ &\quad = (-1)^{\delta_1+\delta_2}\frac{G_{\delta_1+\delta_2}(\alpha)G_{\delta_1+\delta_2}(\beta)}{G_0(\alpha+\beta)}\,|x-y|^{\alpha+\beta-1}\,|x-z|^{-\beta}\,|y-z|^{-\alpha}\,,\end{aligned} \tag{4.16}$$

with $\alpha = (-\lambda_1+\lambda_2-\nu+1)/2$, $\beta = (\lambda_1-\lambda_2-\nu+1)/2$. The integral converges in the region $\operatorname{Re}\alpha > 0$, $\operatorname{Re}\beta > 0$, $\operatorname{Re}(\alpha+\beta) < 1$. The uniqueness of the triple pairing ensures that (4.15) must be correct up to a multiplicative constant. But then (4.16) must also be correct, except possibly for the specific constant of proportionality. That constant can be pinned down in a variety of ways; see, for example, [15, Lemma 4.32].

A partition of unity argument shows that the quantities $P_\nu^\Gamma(\tau_1, \tau_2, J_\nu E_{-\nu})$ and $P_{-\nu}^\Gamma(\tau_1, \tau_2, E_{-\nu})$ are related by the same Gamma factors as the global pairings in (4.15). Combining this information with (4.11) and the standard Gamma identity $G_\delta(\nu)G_\delta(1-\nu) = (-1)^\delta$, we find

$$\begin{aligned}&P_\nu^\Gamma(\tau_1, \tau_2, E_\nu) \\ &\quad = (-1)^{\delta_1+\delta_2} G_{\delta_1+\delta_2}\left(\frac{\lambda_1-\lambda_2-\nu+1}{2}\right) \\ &\qquad \times G_{\delta_1+\delta_2}\left(\frac{-\lambda_1+\lambda_2-\nu+1}{2}\right) P_{-\nu}^\Gamma(\tau_1, \tau_2, E_{-\nu})\,.\end{aligned} \tag{4.17}$$

Once we relate $P_\nu^\Gamma(\tau_1, \tau_2, E_\nu)$ to the Rankin–Selberg L-function, this identity will turn out be the functional equation.

We begin by substituting the expression (4.2) for E_ν in (4.14). Initially we argue formally; the unfolding step will be justified later, when we see that the resulting integral converges absolutely:

$$\begin{aligned}&P_\nu^\Gamma(\tau_1, \tau_2, E_\nu) \\ &\quad = \int_{\Gamma\backslash G}\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,E_\nu(ghf_3)\,\psi(h)\,dh\,dg \\ &\quad = \zeta(\nu+1)\sum_{\Gamma/\Gamma_\infty}\int_{\Gamma\backslash G}\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\delta_\infty(\gamma^{-1}ghf_3)\,\psi(h)\,dh\,dg \\ &\quad = \zeta(\nu+1)\int_{\Gamma_\infty\backslash G}\int_G \tau_1(ghf_1)\,\tau_2(ghf_2)\,\delta_\infty(ghf_3)\,\psi(h)\,dh\,dg\,.\end{aligned} \tag{4.18}$$

The integrand for the outer integral on the right is no longer Γ-invariant, but it is $(\Gamma \cap N)$-invariant, of course, and has all the other properties of the integrand in (4.14). Those are the properties used in the proof of theorem 3.33 to establish rapid decay. In other words, the same argument shows that the integrand in (4.18) decays rapidly in the direction of the cusp. However, $\Gamma_\infty\backslash G$ is not "compact in the directions opposite to the cusp", and we still need to argue that the integral converges in those directions as well.

Together with the upper triangular unipotent subgroup $N \subset G$, the two subgroups

$$K = SO(2)/\{\pm 1\}\,, \quad A = \left\{ a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\} \tag{4.19}$$

determine the Iwasawa decomposition

$$G^0 = NAK \tag{4.20}$$

of the identity component $G^0 \simeq SL(2,\mathbb{R})/\{\pm 1\}$ of G. Since Γ_∞ meets both components of G, and since $\Gamma_\infty \cap G^0 = \Gamma \cap N$, we can make the identification $\Gamma_\infty \backslash G \simeq (\Gamma \cap N)\backslash G^0$. Hence, and because

$$dg = e^{-2\rho}(a)\, dn\, da\, dk\,, \quad \text{with} \quad e^\rho(a_t) = e^t\,, \tag{4.21}$$

the identity (4.18) can be rewritten as

$$\begin{aligned} &P_\nu^\Gamma(\tau_1, \tau_2, E_\nu) \\ &\quad = \zeta(\nu+1) \int_K \int_A \int_{(\Gamma\cap N)\backslash N} \int_G e^{-2\rho}(a)\, \tau_1(nakhf_1)\, \tau_2(nakhf_2) \\ &\qquad \times \delta_\infty(nakhf_3)\, \psi(h)\, dh\, dn\, da\, dk\,. \end{aligned} \tag{4.22}$$

As the t tends to $+\infty$, the point $g = na_tk$ moves towards the cusp. In the opposite direction, as $t \to -\infty$, the integrand in (4.22) grows at most like a power of e^{-t}. To see this, and to determine the rate of growth or decay, we temporarily regard the three instances of the argument nak as independent of each other, as in the discussion around (3.34). In the case of the τ_j, the maximum rate of growth is $e^{(-|\mathrm{Re}\,\lambda_j|+1)t}$, and in the case of δ_∞, it is $e^{(\mathrm{Re}\,\nu+1)t}$, without absolute value sign around Re ν. The reason for the latter assertion is that we know the behavior of $\delta_\infty(g)$ when g is multiplied on the left by any $n \in N$ – unchanged – and when g is multiplied on the left by any $a_t \in A$ – by the factor $e^{(\mathrm{Re}\,\nu+1)t}$; cf. (4.28) below. In short, the integrand in (4.22) can be made to decay as $t \to -\infty$ by choosing Re ν large enough. That makes the integral converge absolutely and justifies the unfolding process.

The smoothing function $\psi \in C_c^\infty(G)$ in theorem 3.33 is arbitrary so far, except for the normalization $\int_G \psi(g)dg = 1$. We can therefore require ψ to have support in G^0, and also impose the condition

$$\psi(kg) = \psi(g) \;\text{ for all } k \in K,\, g \in G\,; \tag{4.23}$$

the latter can be arranged by averaging the original function ψ over K. The analogue of (4.21) for the KAN decomposition is $dg = e^{2\rho}(a)\, dk\, da\, dn$. Hence

$$\begin{aligned} &\int_A \int_N e^{2\rho}(a)\, \psi(an)\, dn\, da = 1\,, \quad \text{or equivalently} \int_A \psi_A(a)\, da = 1\,, \\ &\quad \text{with} \quad \psi_A(a) = e^{2\rho}(a) \int_N \psi(an)\, dn = \int_N \psi(na)\, dn\,, \end{aligned} \tag{4.24}$$

restates the normalization condition for the K-invariant function ψ.

We had argued earlier that the function $e(\ell x)$, for $\ell \neq 0$, has a canonical extension – now viewed as distribution – across infinity. That allows us to regard $e(\ell x)$ as a well defined element of the unbounded model (3.24). We can also make sense of the constant function 1 as element of the unbounded model for $\operatorname{Re}\lambda > 0$, and for other values of λ by meromorphic continuation. Whether or not ℓ equals zero, we let $B_{\ell,\lambda,\delta} \in V_{\lambda,\delta}^{-\infty}$ denote the distribution vector that corresponds to $e(\ell x)$. Then

$$\pi_{\lambda,\delta}(n_x) B_{\ell,\lambda,\delta} = e(-\ell x)\, B_{\ell,\lambda,\delta}\,, \quad \text{and} \quad B_{\ell,\lambda,\delta}(n_x) = e(\ell x)\,. \tag{4.25}$$

The latter equation has meaning since $N \subset G/B$ is open and $B_{\ell,\lambda,\delta}$, like any vector in $V_{\lambda,\delta}^{-\infty}$, transforms according to a character under right translation by elements of B. We had assumed that at least one among τ_1 and τ_2 is cuspidal – τ_1, say, for definiteness. Then

$$\tau_1 = \sum_{\ell \neq 0} a_\ell \, B_{\ell,\lambda_1,\delta_1}\,, \quad \tau_2 = \sum_{\ell \in \mathbb{Z}} b_\ell \, B_{\ell,\lambda_2,\delta_2} + \cdots \tag{4.26}$$

are the Fourier expansions of τ_1 and τ_2. Here $\ldots$ stands for a vector in $V_{\lambda_2,\delta_2}^{-\infty}$ that is N-invariant and supported on $sB \subset G/B$; recall (3.18) for the definition of $s \in G$. The series for τ_1 has no such singular contribution on sB, as was explained in (2.17) and the passage that follows it.

In (4.22), the process of averaging over $\Gamma \backslash \Gamma_\infty$ from the left and smoothing from the right commute. Thus, using the fact that δ_∞ and $\ldots$ in (4.26) are N-invariant, we find

$$\begin{aligned}
&\int_{(\Gamma\cap N)\backslash N} \int_G \tau_1(nakhf_1)\, \tau_2(nakhf_2)\, \delta_\infty(nakhf_3)\, \psi(h)\, dh\, dn \\
&\quad = \sum_{\ell \neq 0} a_\ell\, b_{-\ell} \int_G B_{\ell,\lambda_1,\delta_1}(akhf_1)\, B_{-\ell,\lambda_2,\delta_2}(akhf_2)\, \delta_\infty(akhf_3)\, \psi(h)\, dh \\
&\quad = \sum_{\ell \neq 0} a_\ell\, b_{-\ell} \int_G B_{\ell,\lambda_1,\delta_1}(ahf_1)\, B_{-\ell,\lambda_2,\delta_2}(ahf_2)\, \delta_\infty(ahf_3)\, \psi(h)\, dh \\
&\quad = \sum_{\ell \neq 0} a_\ell\, b_{-\ell} \int_G B_{\ell,\lambda_1,\delta_1}(ah)\, B_{-\ell,\lambda_2,\delta_2}(ahn_1)\, \delta_\infty(ahs)\, \psi(h)\, dh\,;
\end{aligned} \tag{4.27}$$

at the second step we have used the K-invariance of ψ, and at the last step, we have inserted the concrete values $f_1 = e$, $f_2 = n_1$, $f_3 = s$ – cf. (3.14) and (3.17).

When we substitute (4.27) into (4.22), we can make several simplifications. The expression on the right in (4.27) no longer depends on the variable k, so the integral over K in (4.22) can be omitted. The distribution δ_∞ is supported on $sB \subset G$. Hence, when the variable h in (4.22) is written as $h = kn\tilde{a}$, with $k \in K$, $n \in N$, $\tilde{a} \in A$, and $dh = dk\, dn\, d\tilde{a}$, the k-integration reduces to evaluation at $k = e$. Since A acts via $e^{2\rho}$ on the cotangent space at $sB \in G/B$,

$$\delta_\infty(aks)\, dk = e^{2\rho}(a)\, \delta_\infty(ksa^{-1})\, dk = \chi_{\nu+\rho}(a)\, \delta_\infty(ks)\, dk \quad \text{for } a \in A\,. \tag{4.28}$$

It follows that $\delta_\infty(ahs)\,dk = \delta_\infty(akn\tilde{a}s)\,dk = \delta_\infty(aks(s^{-1}ns)\tilde{a}^{-1})\,dk$ contributes the factor $\chi_{\nu+\rho}(a)\,\chi_{\nu-\rho}(\tilde{a}) = e^{-2\rho}(\tilde{a})\,\chi_{\nu+\rho}(a\tilde{a})$ when it is integrated over K. Effectively we have replaced the integrals over $h \in G$ in (4.22) and (4.27) by integrals over NA. But the integrand being smoothed in (4.27) is already N-invariant. Thus, instead of smoothing over G with respect to ψ, we only need to smooth over A with respect to ψ_A, as defined in (4.24):

$$\begin{aligned} P_\nu^\Gamma(\tau_1,\tau_2,E_\nu) = \zeta(\nu+1) \sum_{\ell\neq 0} a_\ell\, b_{-\ell} \int_A \int_A e^{-2\rho}(a\tilde{a}) \\ \times B_{\ell,\lambda_1,\delta_1}(a\tilde{a})\, B_{-\ell,\lambda_2,d_2}(a\tilde{a}n_1)\, \chi_{\nu+\rho}(a\tilde{a})\, \psi_A(\tilde{a})\, d\tilde{a}\, da\,. \end{aligned} \tag{4.29}$$

We parametrize $a, \tilde{a} \in A$ as $a = a_t$, $\tilde{a} = a_{\tilde{t}}$, as in (4.19), with $t, \tilde{t} \in \mathbb{R}$ and $da = dt$, $d\tilde{a} = d\tilde{t}$. Then, in view of the definition (3.12) of $V_{\lambda,\delta}^{-\infty}$ and the characterization (4.25) of $B_{\ell,\lambda,\delta}$,

$$\begin{aligned} B_{\ell,\lambda_1,\delta_1}(a_t a_{\tilde{t}}) &= e^{(1-\lambda_1)(t+\tilde{t})}\, B_{\ell,\lambda_1,\delta_1}(e) = e^{(1-\lambda_1)(t+\tilde{t})}\,, \\ B_{-\ell,\lambda_2,\delta_2}(a_t \tilde{a}_{\tilde{t}} n_1) &= e^{(1-\lambda_2)(t+\tilde{t})}\, B_{-\ell,\lambda_2,\delta_2}(a\tilde{a}n_1 a^{-1}\tilde{a}^{-1}) \\ &= e^{(1-\lambda_2)(t+\tilde{t})}\, e(-\ell\, e^{2(t+\tilde{t})})\,, \\ \chi_{\nu+\rho}(a_t a_{\tilde{t}}) &= e^{(\nu+1)(t+\tilde{t})}\,, \qquad e^{-2\rho}(a_t a_{\tilde{t}}) = e^{-2(t+\tilde{t})}\,. \end{aligned} \tag{4.30}$$

This leads to the equation

$$\begin{aligned} P_\nu^\Gamma(\tau_1,\tau_2,E_\nu) = \zeta(\nu+1) \sum_{\ell\neq 0} a_\ell\, b_{-\ell} \\ \times \int_{\mathbb{R}} \int_{\mathbb{R}} e^{(\nu+1-\lambda_1-\lambda_2)(t+\tilde{t})}\, e(-\ell\, e^{2(t+\tilde{t})})\, \psi_A(a_{\tilde{t}})\, d\tilde{t}\, dt\,. \end{aligned} \tag{4.31}$$

To simplify this expression further, we set $x = e^{2t}$, $y = e^{2\tilde{t}}$, and

$$\psi_A(a_{\tilde{t}}) = \psi_{\mathbb{R}}(y) \qquad (y = e^{2\tilde{t}})\,. \tag{4.32}$$

Then $dx = 2e^{2t}dt$, $dy = 2e^{2\tilde{t}}d\tilde{t}$, and the normalization (4.24) becomes

$$\int_0^\infty \psi_{\mathbb{R}}(y)\frac{dy}{y} = 2\,. \tag{4.33}$$

Putting all the pieces together, we find

$$\begin{aligned} &P_\nu^\Gamma(\tau_1,\tau_2,E_\nu) \\ &\quad = \frac{\zeta(\nu+1)}{4} \sum_{\ell\neq 0} a_\ell\, b_{-\ell} \int_0^\infty \int_0^\infty (xy)^{\frac{\nu+1-\lambda_1-\lambda_2}{2}}\, e(-\ell\, x\, y)\, \psi_{\mathbb{R}}(y)\, \frac{dy}{y}\, \frac{dx}{x}\,. \end{aligned} \tag{4.34}$$

We know from the derivation of this formula that the integral and the sum must converge for $\mathrm{Re}\,\nu \gg 0$, and indeed they do. Since the smoothing function $\psi_{\mathbb{R}}$ has compact support in $(0, \infty)$, the inner integral is the Fourier transform of a compactly supported C^∞ function on $\mathbb{R}$. The resulting function of x is smooth at the origin and decays rapidly at infinity. That makes the outer integral converge, provided $\mathrm{Re}\,\nu$ is large enough. A change of variables then shows that the double integral has order of growth $O(|\ell|^{\mathrm{Re}(\lambda_1+\lambda_2-\nu-1)/2})$, so the sum does converge, again for $\mathrm{Re}\,\nu \gg 0$.

If we regard $e(-\ell x)$, $\ell \neq 0$, not as a function, but as a distribution that vanishes to infinite order at infinity, the integral $\int_0^\infty e(-\ell x)\, x^{\frac{\nu+1-\lambda_1-\lambda_2}{2}}\, dx$ converges for $\mathrm{Re}\,\nu \gg 0$, and the smoothing process in (4.34) becomes unnecessary. Taking this approach, we make the change of variables $x \mapsto x/y$, which splits off the integral (4.33). Hence

$$\begin{aligned}
&P_\nu^\Gamma(\tau_1, \tau_2, E_\nu) \\
&\quad = \frac{\zeta(\nu+1)}{2} \sum_{\ell\neq 0} a_\ell\, b_{-\ell} \int_0^\infty x^{\frac{\nu+1-\lambda_1-\lambda_2}{2}}\, e(-\ell\, x)\, \frac{dx}{x} \\
&\quad = \frac{\zeta(\nu+1)}{2} \sum_{\ell\neq 0} a_\ell\, b_{-\ell}\, |\ell|^{\frac{\lambda_1+\lambda_2-\nu-1}{2}} \int_0^\infty x^{\frac{\nu+1-\lambda_1-\lambda_2}{2}}\, e(-(\operatorname{sgn}\ell)x) \frac{dx}{x} \qquad (4.35) \\
&\quad = \frac{\zeta(2s)}{2} \sum_{\ell\neq 0} a_\ell\, b_{-\ell}\, |\ell|^{\frac{\lambda_1+\lambda_2}{2}-s} \int_0^\infty x^{s-\frac{\lambda_1+\lambda_2}{2}}\, e(-(\operatorname{sgn}\ell)x)\, \frac{dx}{x}\,.
\end{aligned}$$

At the last step, we have expressed ν in terms of s, as in (4.12).

By definition, the Rankin–Selberg L-function of the pair of automorphic distributions τ_1, τ_2 is

$$L(s, \tau_1 \otimes \tau_2) = \zeta(2s) \sum_{n>0} a_n\, b_n\, n^{\frac{\lambda_1+\lambda_2}{2}-s}\,. \qquad (4.36)$$

Recall that the Fourier coefficients a_n, b_n depend on the choice of the embedding parameter λ_j over $-\lambda_j$. The standard L-function (2.22), and (1.2) in the case of modular forms, with $\lambda = 1-k$, are defined in terms of the renormalized coefficients $a_n|n|^{\lambda/2}$. For the same reason the renormalized coefficients appear in the Rankin–Selberg L-function. To make the connection between (4.35) and the L-function, notice that translation by the matrix

$$r = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (4.37)$$

transforms $\tau_j \in (V_{\lambda_j,\delta_j}^{-\infty})^\Gamma$, realized as $\tau_j(x)$ in terms of the unbounded model, to $(-1)^{\delta_j}\tau_j(-x)$. Since $r \in \Gamma$, that means $\tau_j(-x) = (-1)^{\delta_j}\tau_j(x)$, i.e.,

$$a_{-n} = (-1)^{\delta_1}\, a_n\,, \quad b_{-n} = (-1)^{\delta_2}\, b_n\,. \qquad (4.38)$$

Hence

$$\begin{aligned} \zeta(2s) \sum_{\ell>0} a_\ell\, b_{-\ell}\, |\ell|^{\frac{\lambda_1+\lambda_2}{2}-s} &= (-1)^{\delta_2}\, L(s,\tau_1\otimes\tau_2)\,, \\ \zeta(2s) \sum_{\ell<0} a_\ell\, b_{-\ell}\, |\ell|^{\frac{\lambda_1+\lambda_2}{2}-s} &= (-1)^{\delta_1}\, L(s,\tau_1\otimes\tau_2)\,. \end{aligned} \tag{4.39}$$

This allows us to re-write (4.35) as

$$\begin{aligned} &2\, P_\nu^\Gamma(\tau_1,\tau_2,E_\nu) \\ &\quad = L(s,\tau_1\otimes\tau_2) \\ &\qquad \times \left\{ (-1)^{\delta_2} \int_{-\infty}^0 |x|^{s-\frac{\lambda_1+\lambda_2}{2}}\, e(x)\, \frac{dx}{x} \;+\; (-1)^{\delta_1} \int_0^\infty |x|^{s-\frac{\lambda_1+\lambda_2}{2}}\, e(x)\, \frac{dx}{x} \right\} \\ &\quad = (-1)^{\delta_1}\, G_{\delta_1+\delta_2}\!\left(s - \frac{\lambda_1+\lambda_2}{2}\right) L(s,\tau_1\otimes\tau_2)\,; \end{aligned} \tag{4.40}$$

recall (2.25), and also the relationship $\nu = 2s-1$ between ν and s.

To complete the proof of the functional equation, we combine (4.40) with (4.17) and appeal to the standard Gamma identity $G_\delta(s)G_\delta(1-s) = (-1)^\delta$:

4.41 Proposition. *The Rankin–Selberg L-function satisfies the functional equation*

$$L(1-s,\tau_1\otimes\tau_2) = \prod\nolimits_{\varepsilon_1,\varepsilon_2=\pm1} G_{\delta_1+\delta_2}\!\left(s + \varepsilon_1\frac{\lambda_1}{2} + \varepsilon_2\frac{\lambda_2}{2}\right) L(s,\tau_1\otimes\tau_2)\,.$$

We have shown that (4.40) has a holomorphic continuation to $\mathbb{C}-\{1\}$, with at most a simple pole at $s=1$. Traditionally one states the functional equation and analytic continuation not for the expression in (4.40), but rather for Langlands' completed L-function

$$\Lambda(s,\tau_1\otimes\tau_2) = L_\infty(s,\tau_1\otimes\tau_2)\, L(s,\tau_1\otimes\tau_2)\,, \tag{4.42}$$

whose "component at infinity" is a product of Gamma factors that depend on the type of the τ_j. If both τ_1 and τ_2 correspond to Maass forms, then

$$\text{Maass case:}\quad L_\infty(s,\tau_1\otimes\tau_2) = \prod\nolimits_{\varepsilon_1,\varepsilon_2=\pm1} \Gamma_{\mathbb{R}}\!\left(s + \varepsilon_1\frac{\lambda_1}{2} + \varepsilon_2\frac{\lambda_2}{2} + \eta\right), \quad \text{with } \eta\in\{0,1\}\,,\ \ \eta\equiv\delta_1+\delta_2 \pmod 2\,. \tag{4.43}$$

Here $\Gamma_{\mathbb{R}}$ denotes the Artin Γ-factor $\pi^{-s/2}\Gamma(s/2)$. If one of the τ_j, say τ_2 for definiteness, corresponds to a holomorphic cusp form of weight k, then

$$\text{mixed case:}\quad L_\infty(s,\tau_1\otimes\tau_2) \;=\; \Gamma_{\mathbb{C}}\!\left(s + \frac{\lambda_1}{2} + \frac{k-1}{2}\right) \Gamma_{\mathbb{C}}\!\left(s - \frac{\lambda_1}{2} + \frac{k-1}{2}\right), \tag{4.44}$$

where $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$. Finally, when both τ_1 and τ_2 correspond to holomorphic cusp forms, of weights k_1 and k_2, respectively,

$$\text{modular forms case:} \quad L_\infty(s, \tau_1 \otimes \tau_2) = \Gamma_{\mathbb{C}}\left(s + \frac{k_1 + k_2}{2} - 1\right)\Gamma_{\mathbb{C}}\left(s + \frac{|k_1 - k_2|}{2}\right). \tag{4.45}$$

In all cases, the functional equation of the previous proposition directly implies the equality of $\Lambda(s, \tau_1 \otimes \tau_2)$ and $\Lambda(1 - s, \tau_1 \otimes \tau_2)$, up to a sign; this follows from standard Gamma identities, in particular the identity $G_\delta(s)G_\delta(1 - s) = (-1)^\delta$ and the Legendre duplication formula.

Just as important as the functional equation is the assertion of holomorphy: both $L(s, \tau_1 \otimes \tau_2)$ and $\Lambda(s, \tau_1 \otimes \tau_2)$ are holomorphic except for potential first order poles at $s = 0$ and $s = 1$. For the uncompleted L-function this follows from a classical argument of Jacquet [6, Lemma 14.7.5]. His argument does not require any detailed calculations, and holds in great generality.

Once $L(s, \tau_1 \otimes \tau_2)$ is known to be holomorphic on $\mathbb{C} - \{0, 1\}$, one can deduce the holomorphy of $\Lambda(s, \tau_1 \otimes \tau_2)$ on $\mathbb{C} - \{0, 1\}$ from the results of this section, as follows. Because of the functional equation, it suffices to rule out poles in the region $\{\operatorname{Re} s \geq 1/2, s \neq 1\}$. In effect, we must show that all poles of $L_\infty(s, \tau_1 \otimes \tau_2)$ with $\operatorname{Re} s \geq 1/2$ are compensated by zeroes of $L(s, \tau_1 \otimes \tau_2)$. This is an issue only in the Maass case: modular forms have weights at least 2, and the parameter λ of a Maass form necessarily lies in the region $\{|\operatorname{Re} \lambda| < 1/2\}$. In the Maass case, only one of the four Gamma factors in (4.43) can have a pole with $\operatorname{Re} s \geq 1/2$. Maass forms correspond to irreducible principal series representations, which involve λ_j and $-\lambda_j$ symmetrically. We can therefore assume that $\operatorname{Re} \lambda_j \geq 0$, in which case the pole can only come from the factor $\Gamma_{\mathbb{R}}\big(s - \frac{\lambda_1+\lambda_2}{2}\big)$, with $\eta = 0$, and must occur at $s = \frac{\lambda_1+\lambda_2}{2}$. But then $\delta_1 = \delta_2$, and $G_{\delta_1+\delta_2}\big(s - \frac{\lambda_1+\lambda_2}{2}\big) = G_0\big(s - \frac{\lambda_1+\lambda_2}{2}\big)$ also has a pole at $s = \frac{\lambda_1+\lambda_2}{2}$. We know that (4.40) is holomorphic on $\mathbb{C} - \{0, 1\}$, thus forcing $L(s, \tau_1 \otimes \tau_2)$ to vanish at $s = \frac{\lambda_1+\lambda_2}{2}$, as was to be shown.

5 Exterior Square on $GL(4)$

Recall that if F is a Hecke eigenform on $GL(n, \mathbb{Z})\backslash GL(n, \mathbb{R})$, or more generally, on the quotient of $GL(n, \mathbb{R})$ by a congruence subgroup, the standard L-function of F has an Euler product

$$L(s, F) = \prod_p \prod_{j=1}^{n} (1 - \alpha_{p,j}\, p^{-s})^{-1} . \tag{5.1}$$

The exterior square L-function is then defined as an Euler product

$$L(s, F, Ext^2) = \prod_p L_p(s, F, Ext^2) , \tag{5.2}$$

whose factor at any unramified prime p equals

$$L_p(s, F, Ext^2) = \prod_{1 \le j < k \le n} (1 - \alpha_{p,j}\, \alpha_{p,k}\, p^{-s})^{-1} \,. \tag{5.3}$$

The appropriate definition of the factors $L_p(s, F, Ext^2)$ corresponding to the finitely many ramified primes is still a subtle issue. Harris and Taylor recently exhibited local factors for the ramified primes that are consistent with Langlands functoriality principles, in their proof of the local Langlands conjectures for $GL(n)$. However, Shahidi had much earlier given a separate definition, which by all expectations agrees with the one provided by Harris-Taylor, though the agreement of the two definitions is not obvious. Shahidi furthermore proved that the L-function with his definition of the ramified factors satisfies a functional equation of the type Langlands predicted. Since there can only be one definition which obeys this functional equation, the potential discrepancy between the Harris-Taylor and Shahidi definitions poses no problem from the point of view of L-functions, though it still is a problem for the group-theoretic definition of the Langlands conjectures. In any case, an argument which produces the analytic continuation and functional equation of $L(s, F, Ext^2)$ must give a definition which agrees with Shahidi's.

In our paper [15], we carry out the *archimedean analysis* of the exterior square L-function for $GL(n)$; we establish the holomorphy of the partial L-function $L_S(s, F, Ext^2)$ and its completion at infinity $\Lambda_S(s, F, Ext^2)$, in both cases with the factors in (5.2) corresponding to the set S of ramified primes omitted. To keep the discussion simple, we avoid the problem of ramification in the present paper by treating only the full level subgroup $GL(4, \mathbb{Z}) \subset GL(4, \mathbb{R})$.

By necessity, the notation in this section will not completely agree with that of the earlier sections; in particular, we now set

$$G = GL(4, \mathbb{R})\,, \;\; G_0 = SL^{\pm}(2, \mathbb{R})\,, \;\; \Gamma = GL(4, \mathbb{Z})\,, \;\; \Gamma_0 = SL^{\pm}(2, \mathbb{Z}). \tag{5.4}$$

We shall also work with the subgroups

$$\begin{aligned} G_1 &= \left\{ \begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix} \,\middle|\, g_1,\, g_2 \in GL(2, \mathbb{R}) \right\} \;\subset\; G\,, \\ \Gamma_1 &= \left\{ \begin{pmatrix} \gamma & 0 \\ 0 & \gamma \end{pmatrix} \,\middle|\, \gamma \in GL(2, \mathbb{Z}) \right\} \;\subset\; \Gamma\,, \\ U &= \left\{ \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \,\middle|\, u \in M_{2\times 2}(\mathbb{R}) \right\} \;\subset\; G\,. \end{aligned} \tag{5.5}$$

Note that $G_1 \simeq GL(2, \mathbb{R}) \times GL(2, \mathbb{R})$ contains $\Gamma_1 \simeq GL(2, \mathbb{Z})$, but not as an arithmetic subgroup.

Again we let $B \subset G$ denote the lower triangular Borel subgroup, and we define $B_1 = G_1 \cap B$. Each pair

$$(\mu, \eta) \in \mathbb{C}^4 \times (\mathbb{Z}/2\mathbb{Z})^4 \tag{5.6}$$

determines a character $\chi_{\mu,\eta} : B \to \mathbb{C}^*$,

$$\chi_{\mu,\eta}(a_{i,j}) = \prod_{1 \le i \le 4} |a_{i,i}|^{\mu_i} \, (\operatorname{sgn} a_{i,i})^{\eta_i} \,, \tag{5.7}$$

and by restriction also a character $\chi_{\mu,\eta} : B_1 \to \mathbb{C}^*$. For $G = GL(4,\mathbb{R})$,

$$\rho = \left(\frac{3}{2}, \, \frac{1}{2}, \, -\frac{1}{2}, \, -\frac{3}{2} \right) \tag{5.8}$$

represents the half sum of the positive roots. In analogy to (3.12),

$$W_{\mu,\eta}^{-\infty} = \{ \tau \in C^{-\infty}(G) \mid \tau(gb) = \chi_{\mu-\rho,\eta}(b^{-1}) \, \tau(g) \text{ for all } g \in G, \ b \in B \} \tag{5.9}$$

is the space of distribution vectors for a generic principal series representation of G. Principal series representations of $G_1 \simeq GL(2,\mathbb{R}) \times GL(2,\mathbb{R})$ are induced from B_1, and hence also parameterized by pairs $(\mu, \eta) \in \mathbb{C}^4 \times (\mathbb{Z}/2\mathbb{Z})^4$,

$$V_{\mu,\eta}^{-\infty} = \{ \tau \in C^{-\infty}(G_1) \mid \tau(gb) = \chi_{\mu-\rho,\eta}(b^{-1}) \tau(g) \text{ for all } g \in G_1, b \in B_1 \}. \tag{5.10}$$

Our current use of the notation $V_{\mu,\eta}^{-\infty}$ is not consistent with (3.12). Not only is G_1 a product of two copies of $GL(2,\mathbb{R})$, but the representations we consider need not be trivial on the center of $GL(2,\mathbb{R})$, in contrast to the situation in section 3, where we considered only automorphic distributions for $PGL(2,\mathbb{R})$. However, the ρ-shift in (5.10) *is* consistent with (3.12): the quantity ρ defined in (5.8) restricts to the corresponding quantities for the two factors of $G_1 \simeq GL(2,\mathbb{R}) \times GL(2,\mathbb{R})$.

The arithmetic group Γ intersects $U \simeq \mathbb{R}^4$ in a lattice, so $(\Gamma \cap U)\backslash U$ is compact. That makes it possible to define the operator

$$\begin{aligned} A : (W_{\mu,\eta}^{-\infty})^{\Gamma} \ &\longrightarrow \ (V_{\mu,\eta}^{-\infty})^{\Gamma_1} \,, \\ A\tau(g) = \int_{(\Gamma \cap U)\backslash U} &\tau(u\,g) \, e(-\operatorname{tr} u) \, du \qquad (g \in G_1) \,. \end{aligned} \tag{5.11}$$

What matters is the fact that the $U \cdot G_1$-orbit of the identity coset in G/B is open. One can therefore restrict any $\tau \in (W_{\mu,\eta}^{-\infty})^{\Gamma}$ to this open subset, and then further to G_1, once the dependence on the variable $u \in U$ has been smoothed out by taking a single Fourier component. The restriction to G_1 still transforms according to $\chi_{\mu,\eta}^{-1}$ under right translation by elements of $B_1 = G_1 \cap B$. This makes $A\tau$ lie in $V_{\mu,\eta}^{-\infty}$. Conjugation by any $\gamma \in \Gamma_1$ preserves the character $u \mapsto e(-\operatorname{tr} u)$ of U and the lattice $\Gamma \cap U$. Since $\Gamma_1 \subset \Gamma$, the Γ-invariance of τ ensures the Γ_1-invariance of $A\tau$.

We now consider a particular cuspidal $\tau \in (W_{\mu,\eta}^{-\infty})^{\Gamma}$. Since Γ contains the center of $SL^{\pm}(4,\mathbb{R})$, any such τ must vanish identically unless

$$\sum_{1 \le j \le 4} \eta_j = 0 \ \text{ in } \mathbb{Z}/2\mathbb{Z} \,. \tag{5.12}$$

We shall also suppose that

$$\sum_{1\le j\le 4} \mu_j = 0\,. \tag{5.13}$$

This is not a serious restriction: it holds automatically when τ arises from a discrete summand of $L^2(\Gamma\backslash G/Z_G^0)$, as in (3.1). Even when that is not the case, we can make (5.13) hold by twisting τ with an appropriate character of Z_G^0, without destroying the Γ-invariance.

In section 3, we described the pairing of three $PGL(2,\mathbb{Z})$-automorphic distributions on $PGL(2,\mathbb{R})$. By limiting ourselves to the case of $PGL(2,\mathbb{R})$ we avoided some notational complications in (3.21) and (3.24), without essential loss of generality: in the case of full level, $-1 \in GL(2,\mathbb{Z})$ must act trivially on any automorphic distribution. In the current setting, we do need the pairing for triples of automorphic distributions on $GL(2,\mathbb{R})$. Theorem 3.33 remains correct as stated in this more general situation, provided the integration is performed over $SL^{\pm}(2,\mathbb{Z})\backslash SL^{\pm}(2,\mathbb{R})$ – the center of $GL(2,\mathbb{R})$ is noncompact and remains noncompact even modulo $GL(2,\mathbb{Z})$. The statement requires the Γ-invariance of all three of the arguments τ_j of the pairing P. Formally, at least, invariance under the diagonal action of Γ on the three arguments suffices to produce a Γ-invariant integrand for the outer integral in theorem 3.33. It is the proof of rapid decay that forces us to assume Γ-invariance of each factor. In the present setting, $A\tau$ arises from a cuspidal automorphic distribution τ on $GL(4,\mathbb{R})$. It is not difficult to adapt the proof of theorem 3.33 to this case: after smoothing by some $\psi \in C_c^\infty(G_0)$, the product of $A\tau$ with the Eisenstein series E_ν does decay rapidly along the cusp.

We again define the Eisenstein series E_ν by the formula (4.2), but now summing over $\Gamma_0/(\Gamma_0)_\infty$; since $-1 \in (\Gamma_0)_\infty$, (4.5–7) remain correct. We should remark that the pairing of three automorphic distributions on $GL(2,\mathbb{R})$ vanishes identically unless $-1 \in GL(2,\mathbb{R})$ acts trivially under the diagonal action. The parity condition (5.12) implies that -1 acts trivially under the diagonal action on $A\tau$. But -1 also acts trivially on delta function δ_∞, and hence on the Eisenstein series E_ν. In short, the parity condition imposed by the action of the center is satisfied in our situation. We have assembled all ingredients to make sense of

$$P_\nu^{\Gamma_0}(A\tau, E_\nu) = \int_{\Gamma_0\backslash G_0}\int_{G_0} A\,\tau\begin{pmatrix} ghf_1 & 0 \\ 0 & ghf_2\end{pmatrix} E_\nu(ghf_3)\,\psi(h)\,dh\,dg\,. \tag{5.14}$$

As a function of ν this is holomorphic, except for a potential first order pole at $\nu = 1$. What we said in section 3 about the intertwining operator J_ν and its interaction with the pairing remains valid, except for the parity subscripts of the Gamma factors in (4.15) and (4.17), since we now work on $GL(2,\mathbb{R})$. The roles of λ_1 and λ_2 are played by, respectively, $\mu_1 - \mu_2$ and $\mu_3 - \mu_4$, as can be seen by comparing the definition (5.10) of $V_{\mu,\eta}^{-\infty}$ to the definition (3.12). Thus, and because of (5.13), $\frac{\lambda_1-\lambda_2}{2}$ corresponds to $\mu_1 + \mu_4$ and $\frac{\lambda_2-\lambda_1}{2}$ corresponds to $\mu_2 + \mu_3$. This explains the arguments of the Gamma factors in the identity

$$P_\nu^{\Gamma_0}(A\tau, E_\nu) = (-1)^{\eta_2+\eta_3}$$
$$\times\, G_{\eta_1+\eta_4}\left(\mu_1+\mu_4-\frac{\nu-1}{2}\right) G_{\eta_2+\eta_3}\left(\mu_2+\mu_3-\frac{\nu-1}{2}\right) P_{-\nu}^{\Gamma_0}(A\tau, E_{-\nu})\,, \tag{5.15}$$

which takes the place of (4.17) in the current setting. In the special case when $\eta_1 = \eta_2$ and $\eta_3 = \eta_4$ – i.e, when the action of $G_1 \cong GL(2,\mathbb{R}) \times GL(2,\mathbb{R})$ on $A\tau$ drops to $PGL(2,\mathbb{R}) \times PGL(2,\mathbb{R})$ – (5.15) agrees with in (4.17), as it must. In the remaining cases the identity is deduced from the appropriate variant of (4.16); for details see [15].

The identity (5.15) is the source of the functional equation of the exterior square L-function, just as (4.17) was the source of the functional equation for the Rankin–Selberg L-function $L(s, \tau_1 \otimes \tau_2)$. To make the connection between the identity (5.15) and the exterior square L-function, we need to consider the Fourier expansion of τ on

$$N = \left\{ n(x,u,v) = \begin{pmatrix} 1 & x_1 & u_1 & v \\ 0 & 1 & x_2 & u_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle|\; x \in \mathbb{R}^3,\ u \in \mathbb{R}^2, v \in \mathbb{R} \right\}. \tag{5.16}$$

Since the N-orbit through the identity coset in G/B is open, it is legitimate to restrict τ to N. This restriction is $(\Gamma \cap N)$-invariant, which allows us to regard τ as lying in $C^{-\infty}((\Gamma \cap N)\backslash N)$. Every $(\Gamma \cap N)$-invariant smooth function on N, and dually every $(\Gamma \cap N)$-invariant distribution, has a Fourier expansion with components indexed by – roughly speaking – the irreducible unitary representations of N. For the one dimensional, or *abelian*, representations this is literally true, but typically infinite dimensional representation contribute more than once, but finitely often. The *non-abelian* Fourier components will turn out not to matter for our purposes. Thus we write

$$\tau(n(x,u,v)) = \sum_{1\le j\le 3} a_{n_1,n_2,n_3}\, e(n_j x_j) + \dots\,, \tag{5.17}$$

with $\dots$ denoting the sum of the non-abelian Fourier components of τ. The a_{n_1,n_2,n_3} with positive indices n_j determine all the others:

$$a_{\epsilon_1 n_1, \epsilon_2 n_2, \epsilon_3 n_3} = \epsilon_1^{\eta_1}\, \epsilon_2^{\eta_1+\eta_2}\, \epsilon_3^{\eta_1+\eta_2+\eta_3}\, a_{n_1,n_2,n_3} \quad (\epsilon_j \in \{\pm 1\})\,. \tag{5.18}$$

Indeed, τ is invariant under the action of all diagonal matrices with entries ± 1, since Γ contains these. Each of them acts on N by conjugation, which has the effect of reversing the signs of some of the coordinates. One can then use (5.7) to determine how the a_{n_1,n_2,n_3} change when the signs of one or more of the indices is flipped.

When τ is a Hecke eigendistribution, the Fourier coefficients a_{n_1,n_2,n_3} are related to the Hecke eigenvalues. Specifically, $k^{\mu_1+\mu_2} a_{1,k,1}$ is the eigenvalue of the Hecke operator $T_{1,k,1}$. The eigenvalues for Hecke operators indexed by unramified primes can be expressed in terms of the $\alpha_{j,p}$ in (5.1) [20]. Jacquet and Shalika [7, §2] have used this expression to identify the factors $L_p(s, \tau, Ext^2)$ for unramified primes p in terms of the Hecke eigenvalues – in complete generality for all n, not just $n = 4$. In the case of $GL(4)$,

$$L_p(s,\tau,Ext^2) = (1-p^{-2s})^{-1}\sum_{k\geq 0} a_{1,p^k,1}\, p^{k(\mu_1+\mu_2-s)}\,. \tag{5.19}$$

At full level, when there are no ramified primes, the Euler product of the local factors for all primes, as in (5.2), expresses the exterior square L-function as

$$L(s,\tau,Ext^2) = \zeta(2s)\sum_{n\geq 1} a_{1,n,1}\, n^{\mu_1+\mu_2-s}\,. \tag{5.20}$$

One can use this as the definition of the exterior square L-function whether or not τ is a Hecke eigendistribution.

5.21 Lemma. *When s and ν are related by the equation $2s = \nu + 1$,*

$$P_\nu^{\Gamma_0}(A\tau, E_\nu) = 2\,(-1)^{\eta_2}\, G_{\eta_1+\eta_2}(s-\mu_1-\mu_2)\, G_{\eta_1+\eta_3}(s-\mu_1-\mu_3)\, L(s,\tau,Ext^2)\,.$$

Since the proof is lengthy, we shall first deduce the functional equation, which follows from the lemma in combination with (5.15), (5.12–13), and the standard Gamma identity $G_\delta(s)G_\delta(1-s) = (-1)^\delta$:

5.22 Proposition. $L(1-s,\tau,Ext^2) = \prod_{1\leq i<j\leq 4} G_{\eta_i+\eta_j}(s-\mu_i-\mu_j)\, L(s,\tau,Ext^2)$.

This result is originally due to Kim [9] and, in the special case when $W_{\mu,\eta}$ belongs to the spherical principal series, to Stade [21]. We refer the reader to our paper [15] for a discussion of the history of the exterior square L-function for $GL(n)$.

The usual statement of functional equation relates the exterior square L-function $L(s,\tau,Ext^2)$ for $GL(n)$ to that of the dual automorphic distribution $\widetilde{\tau}$. In our case, with $G = GL(4,\mathbb{R})$, these two L-functions coincide; that makes it possible to state the functional equation without reference to $\widetilde{\tau}$.

Just as in the case of the Rankin–Selberg L-function for $GL(2)$, Jacquet's general argument implies that $L(s,\tau,Ext^2)$ is holomorphic, except for possible first order poles at $s = 0$ and $s = 1$ [15]. The fact that $P_\nu^{\Gamma_0}(A\tau, E_\nu)$ is holomorphic, together with an analysis of the poles and zeros of the Gamma factors, establishes the holomorphy of the completed exterior square L-function, again with the possible exception of first order poles at 0 and 1. We conclude our paper with the proof of the lemma.

Proof of Lemma 5.21. Recall the notational conventions (5.4); in particular $G_0^0 = SL(2,\mathbb{R})$ denotes the identity component of $G_0 = SL^\pm(2,\mathbb{R})$. We shall suppose that the smoothing function ψ is supported on G_0^0, as we did in section 4. We also impose the K-invariance condition (4.23) and define ψ_A as we did in (4.24). In section 3 we had pointed out that the expression (3.34) is smooth as function of all three variables. For the same reason

$$(g_1,g_2,g_3) \mapsto \int_{G_0^0} A\tau\begin{pmatrix} g_1hf_1 & 0\\ 0 & g_2hf_2\end{pmatrix}\delta_\infty(g_3hf_3)\,\psi(h)\,dh \tag{5.23}$$

is a C^∞ function on $G_0 \times G_0 \times G_0$. It is also an eigenfunction of the Casimir operator in each of the three variables, of moderate growth since τ and δ_∞ are distribution vectors. The cuspidality of τ implies that the restriction of this function to the triple diagonal decays rapidly in the cuspidal directions. We can therefore set $g_1 = g_2 = g_3 = g$ and integrate with respect to g over the quotient $\Gamma_{0,\infty}\backslash G_0$, with $\Gamma_{0,\infty} = \{\gamma \in \Gamma_0 \mid \gamma\infty = \infty\}$.

In analogy with (4.18), we insert the definition (4.2) of E_ν into (5.14) and unfold: for Re $\nu \gg 0$,

$$P_\nu^{\Gamma_0}(A\tau, E_\nu) = \zeta(\nu+1) \int_{\Gamma_{0,\infty}\backslash G_0} \int_{G_0} A\tau \begin{pmatrix} ghf_1 & 0 \\ 0 & ghf_2 \end{pmatrix} \delta_\infty(ghf_3)\psi(h)dh\,dg\,, \tag{5.24}$$

The justification of this step hinges on two facts. First of all, the function (5.23) has moderate growth, as was just pointed out Secondly, we know the behavior of δ_∞ under left translation by elements of A. From here on we can justify the unfolding exactly as in section 4. In (5.24) we can replace G_0 in the inner integral by the identity component G_0^0 on which ψ is supported. Since $\Gamma_{0,\infty}$ meets both connected components of G_0, we can also replace G_0 by G_0^0 in the outer integral, provided we simultaneously replace $\Gamma_{0,\infty}$ by $\Gamma_{0,\infty}^0 = \Gamma_{0,\infty} \cap G_0^0$. We parameterize G_0^0 by the Iwasawa decomposition $g = n_x a k$ – recall (3.17) and (4.20–21) . To avoid confusion, we now let N_0, A_0, K_0 denote the subgroups of $G_0^0 = SL(2,\mathbb{R})$ analogous to N, A, K in sections 3 and 4. Note that $\Gamma \cap N_0$ has index 2 in $\Gamma_{0,\infty}^0$, which also contains -1, so $(\Gamma \cap N_0)\backslash N_0 A_0 K_0$ covers $\Gamma_{0,\infty}^0\backslash G_0^0$ twice. Thus

$$\begin{aligned} P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1) \\ &\times \int_{A_0}\int_0^1\int_{G_0^0} e^{-2\rho}(a)\, A\tau \begin{pmatrix} n_x a h f_1 & 0 \\ 0 & n_x a h f_2 \end{pmatrix} \delta_\infty(n_x a h f_3)\, \psi(h)\, dh\, dx\, da\,; \end{aligned} \tag{5.25}$$

we have legitimately omitted the integration over the Iwasawa component k because ψ is K-invariant.

Recall the definition (5.11) of $A\tau$. It will be convenient to replace τ by τ^0, defined by the formula

$$\tau^0(g) = \int_{(\Gamma\cap Z_N)\backslash Z_N} \tau(ng)\,dn = \int_0^1 \tau\left(\begin{pmatrix} 1&0&0&v \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} g\right) dv\,, \tag{5.26}$$

with Z_N = center of N. Then τ^0 is invariant under left translation by elements of Z_N, and by elements of $\Gamma \cap N$. We shall also need to know that

$$\tau^0(s_{2,3}\,g) = \tau^0(g) \quad \text{for all } g \in G, \quad \text{with } s_{2,3} = \begin{pmatrix} 1&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{pmatrix}. \tag{5.27}$$

Indeed, $s_{2,3}$ is contained in Γ and commutes with the one parameter group over which τ is averaged to produce τ^0. The passage from τ^0 to $A\tau$ involves averaging over three more variables,

$$A\tau(g) = \int_{\mathbb{R}^3/\mathbb{Z}^3} \tau^0\left(\begin{pmatrix} 1 & 0 & u_1 & 0 \\ 0 & 1 & x_2 & u_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} g\right) e(-u_1 - u_2)\, dx_2\, du_1\, du_2\,. \tag{5.28}$$

Since

$$\begin{pmatrix} 1 & 0 & u_1 & 0 \\ 0 & 1 & x_2 & u_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & x & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x & u_1 - xx_2 & u_1 x \\ 0 & 1 & 0 & u_2 + xx_2 \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x_2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{5.29}$$

the equations (5.25–26) and (5.28) imply

$$\begin{aligned} P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1) \\ &\times \int_{A_0}\int_0^1\int_{G_0^0}\int_{\mathbb{R}^3/\mathbb{Z}^3} \tau^0\left(\begin{pmatrix} 1 & x & u_1 & 0 \\ 0 & 1 & 0 & u_2 \\ 0 & 0 & 1 & x \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x_2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} ahf_1 & 0 \\ 0 & ahf_2 \end{pmatrix}\right) \\ &\times\ e^{-2\rho}(a)\, e(-u_1 - u_2)\, \delta_\infty(n_x ahf_3)\, \psi(h)\, dx_2\, du_1\, du_2\, dh\, dx\, da\,. \end{aligned} \tag{5.30}$$

We appeal to the invariance of τ^0 under the center of N to justify setting the $(1,4)$-entry of the first matrix in the argument of τ^0 equal to zero.

The variable of integration x occurs three times in (5.30). Since δ_∞ is N_0-invariant, we may as well drop the factor n_x in its argument. When we omit the integration with respect to x and treat the remaining instances of x as two separate variables, the integrand – after averaging over $\mathbb{R}^3/\mathbb{Z}^3$ and smoothing with respect to ψ – is a C^∞ function of those two variables; this follows from the fact that (5.23) is separately smooth in all three arguments. We can therefore replace the single integral with respect to x by a double integral, provided we multiply the integrand by the delta function, evaluated on the difference of the two variables. Since

$$\begin{pmatrix} 1 & k & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \ell \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & x_1 & u_1 & 0 \\ 0 & 1 & 0 & u_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 & x_1+k & u_1 & 0 \\ 0 & 1 & 0 & u_2 \\ 0 & 0 & 1 & x_3+\ell \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.31}$$

modulo the center of N, the integrand in (5.30) is separately periodic when the remaining instances of the variable x are uncoupled. The sum

$$\delta_0(x_1 - x_3) = \sum_{\ell\in\mathbb{Z}} e(\ell(x_1 - x_3)) \tag{5.32}$$

represents the "delta function along the diagonal" in $\mathbb{R}^2/\mathbb{Z}^2$. Thus, in view of what we just said,

$$\begin{aligned} P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1) \sum_{\ell\in\mathbb{Z}} \int_{A_0}\int_{\mathbb{R}^2/\mathbb{Z}^2}\int_{G_0^0}\int_{\mathbb{R}^3/\mathbb{Z}^3} e(\ell(x_1 - x_3)) \\ &\times\ \tau^0\left(\begin{pmatrix} 1 & x_1 & u_1 & 0 \\ 0 & 1 & 0 & u_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x_2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} ahf_1 & 0 \\ 0 & ahf_2 \end{pmatrix}\right) e^{-2\rho}(a) \\ &\times\ e(-u_1 - u_2)\, \delta_\infty(ahf_3)\, \psi(h)\, dx_2\, du_1\, du_2\, dh\, dx_1\, dx_3\, da\,. \end{aligned} \tag{5.33}$$

We use the matrix identity

$$\begin{pmatrix} 1&0&0&0\\ 0&1&\ell&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&x_1&u_1&0\\ 0&1&0&u_2\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&1&x_2&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix} = \begin{pmatrix} 1&x_1&u_1-\ell x_1&0\\ 0&1&0&u_2+\ell x_3\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&1&x_2+\ell&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}, \tag{5.34}$$

the $(\Gamma \cap N)$-invariance of τ^0, and the change of variables $u_1 \mapsto u_1 + \ell x_1$, $u_2 \mapsto u_2 - \ell x_3$ to eliminate the factor $e(\ell(x_1 - x_3))$ in (5.33) while simultaneously replacing x_2 by $x_2 + \ell$. We then combine the x_2-integral over $\{0 \le x_2 \le 1\}$ with the sum over ℓ into a single integral over $\mathbb{R}$:

$$\begin{aligned} P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1) \\ &\times \int_{A_0}\int_{\mathbb{R}^2/\mathbb{Z}^2}\int_{G_0^0}\int_{\mathbb{R}}\int_{\mathbb{R}^2/\mathbb{Z}^2} \tau^0\left(\begin{pmatrix} 1&x_1&u_1&0\\ 0&1&0&u_2\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&1&y&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} ahf_1&0\\ 0&ahf_2 \end{pmatrix}\right) \\ &\times e^{-2\rho}(a)\, e(-u_1-u_2)\,\delta_\infty(ahf_3)\,\psi(h)\,du_1\,du_2\,dy\,dh\,dx_1\,dx_3\,da\,. \end{aligned} \tag{5.35}$$

The symbol y instead of x_2 is meant to emphasize the new role of this variable.

Recall the invariance of τ^0 under $s_{2,3}$, as defined in (5.27). Conjugating $s_{2,3}$ across the first matrix in the argument of τ^0 has the effect of switching the roles of the x_i and the u_j,

$$\begin{aligned} P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1) \\ &\times \int_{A_0}\int_{\mathbb{R}^2/\mathbb{Z}^2}\int_{G_0^0}\int_{\mathbb{R}}\int_{\mathbb{R}^2/\mathbb{Z}^2} \tau^0\left(\begin{pmatrix} 1&x_1&u_1&0\\ 0&1&0&u_2\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&y&0\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} ahf_1&0\\ 0&ahf_2 \end{pmatrix}\right) \\ &\times e^{-2\rho}(a)\, e(-x_1-x_3)\,\delta_\infty(ahf_3)\,\psi(h)\,dx_1\,dx_3\,dy\,dh\,du_1\,du_2\,da\,. \end{aligned} \tag{5.36}$$

The congruence

$$\begin{pmatrix} 1&x_1&u_1&0\\ 0&1&0&u_2\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix} \equiv \begin{pmatrix} 1&0&u_1&0\\ 0&1&0&u_2\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&x_1&0&0\\ 0&1&0&0\\ 0&0&1&x_3\\ 0&0&0&1 \end{pmatrix} \tag{5.37}$$

modulo the center of N implies that we can view the integral with respect to $du_1\,du_2$ as projecting τ^0 to the trivial Fourier components with respect to those two variables, whereas the other integrations operate from the right. Right translation commutes with projection onto the trivial Fourier components, thus allowing us to shift the integration with respect to $du_1\,du_2$ all the way to the inside. The passage from τ to τ^0 already involves a projection. Together with the $du_1\,du_2$-integral, this gives us the projection

$$\tau \;\mapsto\; \tau_{\mathrm{abelian}}\,, \quad \tau_{\mathrm{abelian}}(g) = \int_{(\Gamma\cap[N,N])\backslash N} \tau(n\,g)\,dn\,, \tag{5.38}$$

onto the sum of the abelian Fourier coefficients – equivalently of invariants for the derived group $[N,N] \subset N$. Thus (5.37) reduces to

$$
\begin{aligned}
P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 2\,\zeta(\nu+1)\,\times \\
&\times \int_{A_0}\int_{G_0^0}\int_{\mathbb{R}}\int_{\mathbb{R}^2/\mathbb{Z}^2} \tau_{\text{abelian}}\left(\begin{pmatrix} 1 & x_1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & x_3\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&1&y&0\\0&0&0&1\end{pmatrix}\begin{pmatrix} ahf_1 & 0\\ 0 & ahf_2\end{pmatrix}\right) \qquad (5.39)\\
&\times\; e^{-2\rho}(a)\,e(-x_1-x_3)\,\delta_\infty(ahf_3)\,\psi(h)\,dx_1\,dx_3\,dy\,dh\,da\,.
\end{aligned}
$$

Now we argue as we did in the passage from (4.27) to (4.35). First we substitute e, n_1, s for f_1, f_2, f_3 as in (3.14). We then parameterize $h \in G_0^0$ as $h = k\tilde{a}n_{\tilde{x}}$, and observe that the argument of δ_∞ must lie in $N_0 A_0 s\{\pm 1\}$ to give a non-zero contribution. At this point the argument diverges slightly from our earlier argument, where we worked modulo the center of $SL(2,\mathbb{R})$. There are three instances of the variable h in (5.39). When h is replaced by $(-1)\cdot h$, δ_∞ remains unchanged, and the other two instances of h effect a hypothetical sign change of $(-1)^{\eta_1+\eta_2+\eta_3+\eta_4}$ – hypothetical only since $\sum_{1\le j\le 4}\eta_j = 0$; cf. (5.12). Thus $k = e$ and $k = -1$ contribute equally, in effect doubling the factor 2 in (5.39). Since

$$
\begin{pmatrix} 1 & x_1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & x_3\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&1&y&0\\0&0&0&1\end{pmatrix}\begin{pmatrix} 1&\tilde{x}&0&0\\0&1&0&0\\0&0&1&\tilde{x}\\0&0&0&1\end{pmatrix} \equiv \begin{pmatrix} 1 & -\tilde{x}y & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & \tilde{x}y\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & x_1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & x_3\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&1&y&0\\0&0&0&1\end{pmatrix} \qquad (5.40)
$$

modulo a left factor lying in $[N, N]$, the variable $n_{\tilde{x}}$ can simply be absorbed into the $dx_1\,dx_3$-integration. We can therefore replace $\psi(h)\,dh$ by $\psi_A(\tilde{a})\,d\tilde{a}$ and the other instances of h by $\tilde{a}$, as in (4.29). The smoothing by ψ has now been replaced by smoothing with respect to ψ_A, in the single variable a. This reflects the fact that the A-direction is the only non-compact direction for the integral (5.30), aside from the smoothing integral over $h \in G_0^0$, of course[2]. Just as in section 4, the smoothing in the variable a will turn out to be unnecessary when we interpret the integrand – in effect, a Fourier series in one variable, without constant term – as a distribution which can be made convergent by integration by parts, under our standing assumption that $\operatorname{Re}\nu \gg 0$. To summarize, we can eliminate the integration over h and the factor $\psi(h)$ in (5.39), provided we double the factor 2, set $h = e$ in the argument of τ^0, and replace $\delta_\infty(ahf_3)$ by $\chi_{\nu+\rho}(a)$, in analogy to (4.28) and the comment that follows it. Finally we combine the factors $e^{-2\rho}(a)$ and $\chi_{\nu+\rho}(a)$ into the single expression $\chi_{\nu-\rho}(a)$:

$$
\begin{aligned}
P_\nu^{\Gamma_0}(A\tau, E_\nu) &= 4\,\zeta(\nu+1)\,\times \\
&\times \int_{A_0}\int_{\mathbb{R}}\int_{\mathbb{R}^2/\mathbb{Z}^2} \tau_{\text{abelian}}\left(\begin{pmatrix} 1 & x_1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & x_3\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&1&y&0\\0&0&0&1\end{pmatrix}\begin{pmatrix} a & 0\\ 0 & an_1\end{pmatrix}\right) \qquad (5.41)\\
&\times\; \chi_{\nu-\rho}(a)\,e(-x_1-x_3)\,dx_1\,dx_3\,dy\,da\,.
\end{aligned}
$$

For each $n \in (\mathbb{Z} - \{0\})^3$, there exists a unique $B_{n,\mu,\eta} \in W_{\mu,\eta}^{-\infty}$ characterized by the properties

[2] The integration with respect to $y \in \mathbb{R}$ in the equivalent integral (5.39) was obtained by unfolding an integral over $\mathbb{R}/\mathbb{Z}$.

$$\begin{aligned} \pi_{\mu,\eta}(n(x,u,v))B_{n,\mu,\eta} &= e(-n_1x_1 - n_2x_2 - n_3x_3)\, B_{n,\mu,\eta}\,, \\ B_{n,\mu,\eta}(n(x,u,v)) &= e(n_1x_1 + n_2x_2 + n_3x_3) \end{aligned} \tag{5.42}$$

[5]; these identities are analogous to (4.25) and use the notation (5.16). The $B_{n,\mu,\eta}$ corresponding to different values of n are related by the action of the diagonal subgroup $A \subset G$, but this need not concern us here. The cuspidality of τ implies that the Fourier coefficients in (5.17) vanish whenever one or more of the indices are zero. Explicitly,

$$a_n \neq 0 \implies n \in (\mathbb{Z} - \{0\})^3\,. \tag{5.43}$$

Comparing (5.42) to (5.16) and the definition (5.38) of τ_{abelian}, one finds

$$\tau_{\text{abelian}} = \sum_{n\in(\mathbb{Z}-\{0\})^3} a_n\, B_{n,\mu,\eta}\,. \tag{5.44}$$

The inner integral in (5.41) picks out the terms in the sum corresponding to $n_1 = n_3 = 1$. Hence

$$\begin{aligned} &P_\nu^{\Gamma_0}(A\tau, E_\nu) \\ &\quad = 4\,\zeta(\nu+1)\sum_{\ell\neq 0} a_{1,\ell,1} \\ &\qquad \times \int_{A_0}\int_{\mathbb{R}} B_{1,\ell,1;\mu,\eta}\left(\begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&y&0\\ 0&0&0&1\end{pmatrix}\begin{pmatrix} a & 0\\ 0 & an_1\end{pmatrix}\right)\chi_{\nu-\rho}(a)\,dy\,da\,. \end{aligned} \tag{5.45}$$

We parameterize A_0 as in (4.19), $A_0 = \{a_t \mid t \in \mathbb{R}\}$. Then $\chi_{\nu-\rho}(a_t) = e^{(\nu-1)t}$; cf. (4.30). Conjugating a_t across n_1 and using the transformation rule (5.9) that defines $W^{-\infty}_{\mu,\eta}$, we can re-write (5.45) as follows:

$$\begin{aligned} &P_\nu^{\Gamma_0}(A\tau, E_\nu) \\ &\quad = 4\,\zeta(\nu+1)\sum_{\ell\neq 0} a_{1,\ell,1} \\ &\qquad \times \int_{\mathbb{R}^2} B_{1,\ell,1;\mu,\eta}\left(\begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&y&0\\ 0&0&0&1\end{pmatrix}\begin{pmatrix} 1 & 0\\ 0 & a_tn_1a_{-t}\end{pmatrix}\right)e^{(\nu+1-2\mu_1-2\mu_3)t}\,dy\,dt\,. \end{aligned} \tag{5.46}$$

The passage from (5.45) to (5.46) also depends on the identity (5.13), which implies $(1-\mu_1+\mu_2)+(1-\mu_3+\mu_4) = 2-2(\mu_1+\mu_3)$. Note that $a_tn_1a_{-t} = n_{e^{2t}}$; cf. (3.17). We appeal to the matrix identity

$$\begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&y&0\\ 0&0&0&1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&z\\ 0&0&0&1\end{pmatrix} = \begin{pmatrix} 1&0&0&0\\ 0&1&1/y&z\\ 0&0&1&yz\\ 0&0&0&1\end{pmatrix}\begin{pmatrix} 1&0&0&0\\ 0&-1/y&0&0\\ 0&1&y&0\\ 0&0&0&1\end{pmatrix}, \tag{5.47}$$

with $z = e^{2t}$, the characterization (5.42) of $B_{n,\mu,\eta}$, and once more to (5.9), to conclude

$$\begin{aligned} &P_\nu^{\Gamma_0}(A\tau, E_\nu) \\ &\quad = 4\,(-1)^{\eta_2}\,\zeta(\nu+1)\,\sum_{\ell\neq 0} a_{1,\ell,1} \\ &\qquad \times \int_{\mathbb{R}^2} e(\ell/y + y\,e^{2t})\,(\operatorname{sgn} y)^{\eta_2+\eta_3}\,|y|^{\mu_2-\mu_3-1}\,e^{(\nu+1-2\mu_1-2\mu_3)t}\,dy\,dt\,. \end{aligned} \tag{5.48}$$

We simplify the integrand by making the change of variables $y \mapsto \ell/y$, followed by the substitution $x = |\ell|\,|y|^{-1}\,e^{2t}$. Then $dx = 2\,x\,dt$, hence

$$\begin{aligned} &P_\nu^{\Gamma_0}(A\tau, E_\nu) \\ &\quad = 2\,(-1)^{\eta_2}\,\zeta(\nu+1)\,\sum_{\ell\neq 0} a_{1,\ell,1}\,|\ell|^{\,\mu_1+\mu_2-\frac{\nu+1}{2}} \\ &\qquad \times \int_0^\infty \int_{\mathbb{R}} \frac{e(y + (\operatorname{sgn}\ell y)x)}{(\operatorname{sgn}\ell y)^{\eta_2+\eta_3}}\; x^{\frac{\nu-1}{2}-\mu_1-\mu_3}\,|y|^{\frac{\nu-1}{2}-\mu_1-\mu_2}\,dy\,dx\,. \end{aligned} \tag{5.49}$$

Recall the definition (5.20) of the exterior square L-function. We now separate the terms in (5.49) corresponding to positive and negative values of ℓ. Appealing to (5.12) and (5.18), we find

$$\begin{aligned} &P_\nu^{\Gamma_0}(A\tau, E_\nu) \\ &\quad = 2\,L\left(\frac{\nu+1}{2},\ \tau\ ,\ Ext^2\right) \\ &\qquad \times \left\{ (-1)^{\eta_2} \int_0^\infty \int_{\mathbb{R}} \frac{e(y+(\operatorname{sgn} y)x)}{(\operatorname{sgn} y)^{\eta_2+\eta_3}}\,|x|^{\frac{\nu-1}{2}-\mu_1-\mu_3}\,|y|^{\frac{\nu-1}{2}-\mu_1-\mu_2}\,dy\,dx \right. \\ &\qquad \left. + (-1)^{\eta_4} \int_{-\infty}^0 \int_{\mathbb{R}} \frac{e(y+(\operatorname{sgn} y)x)}{(\operatorname{sgn} y)^{\eta_2+\eta_3}}\,|x|^{\frac{\nu-1}{2}-\mu_1-\mu_3}\,|y|^{\frac{\nu-1}{2}-\mu_1-\mu_2}\,dy\,dx \right\}. \end{aligned} \tag{5.50}$$

The factor in curly parentheses equals

$$\begin{aligned} &\int_{\mathbb{R}^2} \frac{(-1)^{\eta_2}\,e(x+y)}{(\operatorname{sgn} x)^{\eta_1+\eta_3}\,(\operatorname{sgn} y)^{\eta_1+\eta_2}}\,|x|^{\frac{\nu-1}{2}-\mu_1-\mu_3}\,|y|^{\frac{\nu-1}{2}-\mu_1-\mu_2}\,dy\,dx \\ &\quad = \int_{\mathbb{R}} \frac{(-1)^{\eta_2}\,e(x)}{(\operatorname{sgn} x)^{\eta_1+\eta_3}}\,|x|^{\frac{\nu-1}{2}-\mu_1-\mu_3} dx \ \times \int_{\mathbb{R}} \frac{e(y)}{(\operatorname{sgn} y)^{\eta_1+\eta_2}}\,|y|^{\frac{\nu-1}{2}-\mu_1-\mu_2} dy \\ &\quad = (-1)^{\eta_2}\,G_{\eta_1+\eta_3}\left(\frac{\nu+1}{2}-\mu_1-\mu_3\right)\,G_{\eta_1+\eta_2}\left(\frac{\nu+1}{2}-\mu_1-\mu_2\right)\,. \end{aligned} \tag{5.51}$$

That completes the proof of the lemma. □

References

[1] Daniel Bump, *The Rankin–Selberg method: a survey*, Number theory, trace formulas and discrete groups (Oslo, 1987), 1989, pp. 49–109.
[2] Daniel Bump and Solomon Friedberg, *The exterior square automorphic L-functions on* $\mathrm{GL}(n)$, Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, Part II (Ramat Aviv, 1989), 1990, pp. 47–65.
[3] W. Casselman, *Jacquet modules for real reductive groups*, Proceedings of the International Congress of Mathematicians (Helsinki, 1978), 1980, pp. 557–563.
[4] ———, *Canonical extensions of Harish-Chandra modules to representations of G*, Canad. J. Math. **41** (1989), no. 3, 385–438.
[5] William Casselman, Henryk Hecht, and Dragan Miličić, *Bruhat filtrations and Whittaker vectors for real groups*, The mathematical legacy of Harish-Chandra (Baltimore, MD, 1998), 2000, pp. 151–190.
[6] Hervé Jacquet, *Automorphic forms on* GL(2). *Part II*, Springer-Verlag, Berlin, 1972. Lecture Notes in Mathematics, Vol. 278.
[7] Hervé Jacquet and Joseph Shalika, *Exterior square L-functions*, Automorphic forms, Shimura varieties, and L-functions, Vol. II (Ann Arbor, MI, 1988), 1990, pp. 143–226.
[8] Henry H. Kim, *Langlands-Shahidi method and poles of automorphic L-functions: application to exterior square L-functions*, Canad. J. Math. **51** (1999), no. 4, 835–849.
[9] ———, *Functoriality for the exterior square of* GL_4 *and the symmetric fourth of* GL_2, J. Amer. Math. Soc. **16** (2003), no. 1, 139–183. With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.
[10] R. P. Langlands, *Problems in the theory of automorphic forms*, Lectures in modern analysis and applications, III, 1970, pp. 18–61. Lecture Notes in Math., Vol. 170.
[11] Robert P. Langlands, *Euler products*, Yale University Press, New Haven, Conn., 1971.
[12] Hans Maass, *Über eine neue Art von nichtautomorphen analytische Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen*, Math. Annalen **121** (1949), 141–183.
[13] Stephen D. Miller and Wilfried Schmid, *Distributions and analytic continuation of Dirichlet series*, J. Funct. Anal. **214** (2004), no. 1, 155–220.
[14] ———, *Automorphic Distributions, L-functions, and Voronoi Summation for GL(3)*, Annals of Mathematics. To appear.
[15] ———, *The archimedean theory of the Exterior Square L-functions over* $\mathbb{Q}$. preprint, 2005.
[16] A. I. Oksak, *Trilinear Lorentz invariant forms*, Comm. Math. Phys. **29** (1973), 189–217.
[17] R. A. Rankin, *Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions. I. The zeros of the function $\sum_{n=1}^{\infty} \tau(n)/n^s$ on the line $\Re s = 13/2$. II. The order of the Fourier coefficients of integral modular forms*, Proc. Cambridge Philos. Soc. **35** (1939), 351–372.
[18] Atle Selberg, *Bemerkungen über eine Dirichletsche Reihe, die mit der Theorie der Modulformen nahe verbunden ist*, Arch. Math. Naturvid. **43** (1940), 47–50 (German).
[19] Freydoon Shahidi, *A proof of Langlands' conjecture on Plancherel measures; complementary series for p-adic groups*, Ann. of Math. (2) **132** (1990), no. 2, 273–330.
[20] Takuro Shintani, *On an explicit formula for class-*1 *"Whittaker functions" on GL_n over P-adic fields*, Proc. Japan Acad. **52** (1976), no. 4, 180–182.

[21] Eric Stade, *Mellin transforms of* $\mathrm{GL}(n,\mathbb{R})$ *Whittaker functions*, Amer. J. Math. **123** (2001), no. 1, 121–161.

[22] Nolan R. Wallach, *Asymptotic expansions of generalized matrix entries of representations of real reductive groups*, Lie group representations, I (College Park, Md., 1982/1983), 1983, pp. 287–369.

5

Langlands Functoriality Conjecture and Number Theory

Freydoon Shahidi

Department of Mathematics
Purdue University
West Lafayette, Indiana
shahidi@math.purdue.edu

Summary. We discuss several applications of the recent developments in the Langlands functoriality conjecture such as the automorphy of the symmetric powers of 2-dimensional complex representations of Galois groups of number fields, lattice point problems, Ramanujan–Selberg and Sato–Tate conjectures. We conclude by explaining how these recent developments are established.

Key words: Langlands functoriality, Ramanujan–Selberg conjectures, automorphic representations, L-functions

Subject Classifications: 11F70, 11R42, 22E50, 22E55

1 Introduction

Recent developments in establishing Langlands functoriality conjecture has led to breakthroughs towards certain important conjectures in number theory, most notably those of Ramanujan and Selberg. The present cases of functoriality [4, 5, 16, 17, 44, 46, 49] are proved by applying appropriate converse theorems of Cogdell and Piatetski-Shapiro [14, 15] to analytic properties of certain automorphic L-functions mainly obtained from the Langlands–Shahidi method [10, 25, 41, 55, 56, 70, 72, 73, 74, 75, 79]. We refer to [1, 11, 12, 13, 24, 31, 40, 65, 66, 77, 78, 80, 81] for some recent expository articles and book chapters on the subject.

The purpose of this paper is to explain these conjectures and the progress made towards them, as well as an exposition of the techniques involved in establishing functoriality.

Here is an overview of the paper. Section 2 is devoted to a review of modular forms, Galois representations and Artin L-functions with the goal of producing new instances of the Artin conjecture for these L-functions [42, 48]. Section 3 covers certain applications of recent progress (cf. [48]) made in the Selberg conjecture [67]. More precisely, we show how the estimate already proved in [48] regarding the conjecture resolves completely the problem of asymptotically counting the number of lattice points inside a hyperbolic circle and that of estimating shifted sums of Fourier coefficients of holomorphic cuspidal eigenforms. The best estimates in the conjecture and that of Ramanujan are due to Kim and Sarnak [46]. The Ramanujan conjecture for Maass forms themselves is treated in Section 4. Section 5 is devoted to a treatment of the Sato–Tate conjecture [68] and explains the progress that follows from the analytic properties of higher symmetric power L-functions which are now available [47] as a consequence of the new cases of functoriality [43, 48], following Serre's criterion ([69], appendix to [76]). Sections 6 and 7 cover an exposition of the functoriality conjecture through its recent instances established in [4, 17, 43, 48], including both symmetric powers and the transfer from classical groups to appropriate general linear groups. On the other hand, Section 8 explains what the general Ramanujan conjecture for classical groups is and how it can be deduced from the conjecture for appropriate GL_n's, when cuspidal representations are globally generic [62, 74, 80, 83]. Finally, we have devoted the last section of this paper to explain the techniques involved in proving the functoriality. This we have done by explaining very briefly the steps of the proof in the case of functoriality for $GL(2) \times GL(3)$ from which the functoriality for the third symmetric powers of cusp forms on $GL(2)$ are deduced.

This paper is based on two lectures given at the International Symposium on Representation Theory and Automorphic Forms at the Seoul National University, Seoul, Korea, over the period of February 14–17, 2005. I would like to thank the organizers and particularly Professor Jae-Hyun Yang for their kind invitation and hospitality. The author is partially supported by the NSF grant DMS–0200325.

2 Modular forms, Galois representations and Artin L-functions

Let $\mathfrak{h}$ denote the Poincaré upper half plane of complex numbers z for which $\mathrm{Im}(z) > 0$. For a positive integer N let $\Gamma_0(N)$ be the Hecke subgroup of level N, i.e.,

$$\Gamma_0(N) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) | c \equiv 0(N) \right\}.$$

This is an example of a congruence subgroup of $SL_2(\mathbb{Z})$, i.e., any subgroup Γ with $\Gamma_N \subset \Gamma \subset SL_2(\mathbb{Z})$, where

$$\Gamma_N = \{\gamma \in SL_2(\mathbb{Z}) | \gamma \equiv I(N)\}.$$

Let f be a (holomorphic) cusp form of weight $k \geq 1$ with respect to $\Gamma_0(N)$ of a given character (Nebentypus) ε, where $\varepsilon\colon (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*$ is a character extended to $\mathbb{Z}/N\mathbb{Z}$ by vanishing on integers m with $(m, N) \neq 1$. Then

$$f(\gamma \cdot z) = \varepsilon(a^{-1})(cz + d)^k f(z),$$

where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ and

$$\gamma \cdot z = \frac{az+b}{cz+d},$$

i.e., defined through the action of $\Gamma_0(N)$ on $\mathfrak{h}$ by fractional linear transformations. Similar actions hold for any congruence subgroup of $SL_2(\mathbb{Z})$. We shall finally assume that f is a new form and an eigenform for all the Hecke operators [21, 84], normalized by $a_1 = 1$, where

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$$

denotes its Fourier expansion at infinity. We recall that $a_{mn} = a_m a_n$ if $(m,n) = 1$ and point out that if ε is primitive, then every member of $S_k(\Gamma_0(N), \varepsilon)$ is a new form.

Next let $\Gamma = \mathrm{Gal}(\mathbb{Q}_a/\mathbb{Q})$ be the Galois group of an algebraic closure $\mathbb{Q}_a$ of $\mathbb{Q}$. Let $F/\mathbb{Q}$ be a finite Galois extension, and for a rational prime p in $\mathbb{Q}$ which remains unramified in F, write $pO_F = \mathfrak{p}_1 \ldots \mathfrak{p}_r$, with distinct prime ideals in O_F, the ring of integers in F. Let $\mathfrak{p} = \mathfrak{p}_1$ and let $F_{\mathfrak{p}}$ be the completion of F at $\mathfrak{p}$. Let $G = \mathrm{Gal}(F/\mathbb{Q})$. The decomposition group $G_{\mathfrak{p}}$, i.e., the set of all the $\sigma \in G$ which fix $\mathfrak{p}$ is canonically isomorphic to $\mathrm{Gal}(F_{\mathfrak{p}}/\mathbb{Q}_p)$. The inertia group

$$I_{\mathfrak{p}} = \{\sigma \in G_{\mathfrak{p}} | \sigma(x) \equiv x(\mathfrak{p}),\ x \in O_{\mathfrak{p}}\}$$

is trivial since p is unramified in F. If $\overline{F}_{\mathfrak{p}}$ and $\overline{\mathbb{Q}}_p$ are the corresponding residue fields $O_{\mathfrak{p}}/\mathfrak{p}$ and $\mathbb{Z}_p/p\mathbb{Z}_p$, then

$$G_{\mathfrak{p}} = G_{\mathfrak{p}}/I_{\mathfrak{p}} = \mathrm{Gal}(\overline{F}_{\mathfrak{p}}/\overline{\mathbb{Q}}_p).$$

The isomorphism is defined by $\sigma \mapsto \overline{\sigma}$, $\overline{\sigma}(\overline{x}) = \overline{\sigma(x)}$, $\overline{x} \in \overline{F}_{\mathfrak{p}}$ representing $x \in O_{\mathfrak{p}}$. The preimage of a fixed generator of the cyclic Galois group $\mathrm{Gal}(\overline{F_{\mathfrak{p}}}/\overline{\mathbb{Q}}_p)$, its Frobenius element $\overline{\sigma}$, is called a Frobenius element. It is defined by

$$Fr_{\mathfrak{p}}(x) \equiv x^p (\mathrm{mod}\ \mathfrak{p})$$

which is the preimage of $\overline{\sigma} \in \mathrm{Gal}(\overline{F}_{\mathfrak{p}}/\overline{\mathbb{Q}}_p)$ satisfying $\overline{\sigma}\colon x \mapsto x^p,\ x \in \overline{F}_{\mathfrak{p}}$. Changing $\mathfrak{p}$ to another prime dividing p, one then obtains a well-defined conjugacy class of Frobenius elements.

Let $\rho\colon \Gamma \to GL_2(\mathbb{C})$ be a two-dimensional continuous representation of $\Gamma = \mathrm{Gal}(\mathbb{Q}_a/\mathbb{Q})$. By continuity it factors through a finite Galois extension $F/\mathbb{Q}$. Then for each prime p unramified in F

$$\mathrm{Trace}\ \rho(Fr_{\mathfrak{p}})$$

is independent of the choice of $\mathfrak{p}$ dividing p.

Conjecture 2.1. Assume $\det \rho(c) = -1$, where c denotes the image of complex conjugation in Γ. Then there exists a new eigenform f of weight 1 such that for all unramified p

$$\text{Trace}\ \rho(Fr_{\mathfrak{p}}) = a_p,$$

where

$$f(z) = \sum_{n \geq 1} a_n e^{2\pi i n z}.$$

The conjecture is a theorem when the image of ρ in $PGL_2(\mathbb{C})$ is solvable, i.e., is isomorphic to either a dihedral group (D_{2n}), A_4 (tetrahedral type), or S_4 (octahedral type) (cf. [22, 47, 52, 86]).

On the other hand, when the image is A_5 (icosahedral type), the conjecture is known only in special cases (cf. [9]). The even case, i.e., when $\det \rho(c) = 1$, is completely out of reach. This should correspond to Maass forms with $\lambda = \frac{1}{4}$ (cf. Section 3).

The converse of the conjecture is due to Deligne and Serre [19]. In fact, if

$$f \in S_1(\Gamma_0(N), \varepsilon)$$

with character ε, then the two-dimensional irreducible representation ρ is of conductor N and $\det(\rho) = \varepsilon$. We note that ε should be considered as a character of Γ through class field theory.

Given a continuous representation r of Γ acting on a finite-dimensional representation V, one can define the corresponding Artin L-function by

$$L(s, r) = \prod_p \det(I_V - r(Fr_{\mathfrak{p}})|V^{I_{\mathfrak{p}}} \cdot p^{-s})^{-1},$$

where $s \in \mathbb{C}$. Here I_V is the identity matrix for V and $V^{I_{\mathfrak{p}}}$ is the subspace of V fixed by the inertia subgroup $I_{\mathfrak{p}}$ for some $\mathfrak{p}|p$. Consequently, the choice of $Fr_{\mathfrak{p}}$ in its $I_{\mathfrak{p}}$-coset is irrelevant which matters if p ramifies. Also observe that $L(s, r)$ is insensitive to conjugation which comes in as we change $\mathfrak{p}$ to another prime dividing p. In particular, for each p the corresponding local factor is independent of the choice of prime $\mathfrak{p}|p$, giving a well-defined Artin L-function (cf. [53]).

Conjecture 2.2 *(Artin conjecture).* The Artin L-function $L(s, r)$ is entire unless r contains the trivial representation.

When r is odd and two dimensional, then clearly Conjecture 2.1 implies Conjecture 2.2 since $L(s, r) = L(s, f)$, where f is the form attached to r. It follows from Booker's recent converse theorem [6] that for $GL(2)$, Artin Conjecture 2.2 also implies its strong form (Conjecture 2.1). Thus Conjecture 2.2 is already valid for all the two-dimensional solvable representations (cf. [22, 47, 86]).

It is an important problem to prove Conjecture 2.2 for different representations. Given an integer m, let

$$\text{Sym}^m : GL_2(\mathbb{C}) \to GL_{m+1}(\mathbb{C})$$

be the $(m+1)$-dimensional irreducible representation of $GL_2(\mathbb{C})$ on the space of symmetric tensors of rank m. Simply said, given $g \in GL_2(\mathbb{C})$, $\mathrm{Sym}^m g$ is the matrix of change of coefficients in a homogeneous polynomial of degree m in two variables (x, y) under the change of coordinates $(x, y) \to (x, y)g$. Up to a twist every finite-dimensional irreducible representation of $GL_2(\mathbb{C})$ is a Sym^m for some non-negative integer m.

Now let ρ be a two-dimensional representation of $\Gamma = \mathrm{Gal}(\mathbb{Q}_a/\mathbb{Q})$. Then $\mathrm{Sym}^m\rho = \mathrm{Sym}^m \cdot \rho$ defines a $(m+1)$-dimensional representation of Γ which is of quite a bit of interest and one likes to know whether $L(s, \mathrm{Sym}^m\rho)$ satisfies the Artin conjecture. We have the following result proved in [48].

Theorem 2.3 ([48]). *Suppose ρ is odd and of either octahedral type or icosahedral type for which Conjecture 2.1 is valid. Then $L(s, Sym^3\rho)$ is entire. It is primitive if and only if ρ is icosahedral. Here primitive simply means that it is not induced from a proper subgroup.*

Similar results are proved for certain Artin L-functions of degrees 6 and 12 in [48] (see also [64].)

A much more general result has been proved in [42] which we shall now explain.

Let $\mathbb{A}_\mathbb{Q}$ be the ring of adeles of $\mathbb{Q}$. Theory of automorphic forms for $GL_N(\mathbb{A}_\mathbb{Q})$ concerns irreducible constituents of $L^2(\mathbb{A}_\mathbb{Q}^* GL_N(\mathbb{Q})\backslash GL_N(\mathbb{A}_\mathbb{Q}))$. One of the cornerstones of Langlands program is that one must be able to attach in a canonical way to every N-dimensional continuous representation of Γ an irreducible automorphic representation of $GL_N(\mathbb{A}_\mathbb{Q})$, i.e., an irreducible constituent of $L^2(\mathbb{A}_\mathbb{Q}^* GL_N(\mathbb{Q})\backslash GL_N(\mathbb{A}_\mathbb{Q}))$, either discretely or continuously.

Theorem 2.4 ([42]). *Assume the two-dimensional continuous representation ρ satisfies Conjecture 2.1. Fix $m \in \mathbb{Z}^+$. Given a rational prime number p not dividing the conductor of ρ, let $\prod_p^{(m)}$ be the $GL_{m+1}(\mathbb{Z}_p)$-spherical irreducible representation of $GL_{m+1}(\mathbb{Q}_p)$ defined by the conjugacy class of $Sym^m\rho(Fr_\mathfrak{p})$, $\mathfrak{p}|p$, in $GL_{m+1}(\mathbb{C})$. Then for every other prime p including $p=\infty$, $\mathbb{Q}_p = \mathbb{R}$, there exist an irreducible admissible representation $\prod_p^{(m)}$ of $GL_{m+1}(\mathbb{Q}_p)$ such that $\prod^{(m)} = \otimes_p \prod_p^{(m)}$ appears in $L^2(\mathbb{A}_\mathbb{Q}^* GL_{m+1}(\mathbb{Q})\backslash GL_{m+1}(\mathbb{A}_\mathbb{Q}))$. If the form f attached to ρ by Conjecture 2.1 generates the cuspidal automorphic representation $\pi_f = \otimes_p \pi_p$ of $GL_2(\mathbb{A}_\mathbb{Q})$, we set $\prod_p^{(m)} = Sym^m\pi_p$ and $\Pi^{(m)} = Sym^m\pi$. In this notation $Sym^m\pi$ is an automorphic representation of $GL_{m+1}(\mathbb{A}_\mathbb{Q})$.*

To prove Theorem 2.4, one basically needs to consider the icosahedral cases. Then one needs to show that irreducible representations of $\tilde{G} = SL_2(F_5)$, the 2-fold cover of A_5, the image of ρ in $PGL_2(\mathbb{C})$, are all automorphic. The group $\tilde{G}$ has two 2-dimensional irreducible representations σ_1 and σ_2. Its three-dimensional irreducible representations are $\mathrm{Sym}^2\sigma_1$ and $\mathrm{Sym}^2\sigma_2$ and thus are automorphic by [23]. The other irreducible representations are as follows. The 4-dimensional irreducible representations are $\mathrm{Sym}^3\sigma_1 \cong \mathrm{Sym}^3\sigma_2$ and $\sigma_1 \otimes \sigma_2$. They are both automorphic by recent results in [48] and [63]. There is only one 6-dimensional representation of $\tilde{G}$. It is isomorphic to $\mathrm{Sym}^2\sigma_1 \otimes \sigma_2 \cong \sigma_1 \otimes \mathrm{Sym}^2\sigma_2$ which was recently proved to be automorphic in [48]. The remaining irreducible representation is of dimension 5. It is basically the induced representation $\mathrm{Ind}_{A_4}^{A_5}\chi$ from a character χ. Its automorphy is then proved in [42].

3 Lattice point problems and the Selberg conjecture

Counting the number of integral points inside a circle of radius $\sqrt{X}$ in the euclidean plane is an estimation of the sum

$$\sum_{m\leq X} r(m),$$

where $r(m)$ is the number of ways a positive number m can be written as a sum of squares of two integers, i.e., the cardinality of the set

$$\{(a,b)\in\mathbb{Z}^2 | a^2+b^2=m\}.$$

It is then clear, using an area calculation, that

$$\sum_{m\leq X} r(m) = \pi X + O(X^{1/2}),$$

solving Gauss's circle problem. It is conjectured that the error term can be replaced by $O\left(X^{\frac{1}{4}+\varepsilon}\right)$. The best estimate so far is $O\left(X^{\frac{23}{73}+\varepsilon}\right)$ due to Huxley.

Given a congruence subgroup Γ, the hyperbolic Riemann surface $\Gamma\backslash\mathfrak{h}$ carries the metric $|dz|/y$. This leads to the distance function

$$d(z,w) = \log \frac{|z-\overline{w}|+|z-w|}{|z-\overline{w}|-|z-w|}, \tag{3.1}$$

whose cosh can be written as

$$\cosh d(z,w) = 1+2u(z,w),$$

where

$$u(z,w) = |z-w|^2/4\mathrm{Im}(z)\mathrm{Im}(w). \tag{3.2}$$

Clearly $u(\gamma\cdot z,\gamma\cdot w)=u(z,w)$.

The corresponding lattice problem can be formulated as:

Problem 3.1. Fix $X>0$, z and w in $\mathfrak{h}$. Consider the lattice generated by Γ acting at z, i.e., the orbit $\Gamma\cdot z$. Count the number of lattice points in $\Gamma\cdot z$ inside the hyperbolic circle of radius X centered at w, i.e., find the cardinality $P(X)$ of

$$\{\gamma\in\Gamma | 2+4u(\gamma\cdot z,w)\leq X\}.$$

Example. Suppose $z = w = \sqrt{-1}$ and $\Gamma = SL_2(\mathbb{Z})$. Then $P(X)$ will be the cardinality of

$$\{(a,b,c,d) \in \mathbb{Z}^4 | ad - bc = 1,\ a^2 + b^2 + c^2 + d^2 \leq X\}.$$

For arbitrary Γ, this counts the number of elements in Γ for which $a^2+b^2+c^2+d^2 \leq X$.

One is able to find an asymptotic formula for $P(X)$ in terms of eigenvalues of the hyperbolic Laplacian $\Delta = -y^2(\partial^2/\partial x^2 + \partial^2/\partial y^2)$ on $L_0^2(\Gamma\backslash\mathfrak{h})$, the L^2-space of Maass cusp forms on $\Gamma\backslash\mathfrak{h}$ (cf. [33, 81]).

Let λ denote an eigenvalue of Δ. For convenience write $\lambda = s(1-s)$ for $0 \leq \lambda < 1/4$ or $1/2 < s \leq 1$. We have

Conjecture 3.2 *(Selberg [67]).* Assume Γ is a congruence subgroup. Then there are no Maass forms which are eigenfunctions of Δ having an eigenvalue λ with $0 < \lambda < 1/4$, or equivalently for $\frac{1}{2} < s < 1$.

This is a very hard conjecture and even the small improvements are of particular interest (cf. [33, 48, 65, 66, 81]). It is well known that there are only a finite number of normalized eigenfunctions for which $0 < \lambda < 1/4$. Let $\frac{1}{2} < s_j < 1$ denote the corresponding s_j, $\lambda_j = s_j(1-s_j)$ and let $\{u_j\}$ denote an orthonormal set of (Maass) cuspidal eigenfunctions for all the Hecke operators and Δ, where the norm is that of Petersson. In particular, if $|F|$ denotes the volume of the fundamental domain $F = \Gamma\backslash\mathfrak{h}$, then the constant Maass forms u_0, corresponding to $s_0 = 1$, $\lambda_0 = 0$, is simply $u_0(z) = |F|^{-1/2}$, $z \in \mathfrak{h}$.

As explained in [33], one can then calculate, using an appropriate kernel function for a certain trace formula, to get

$$P(X) = P_{z,w}(X) = \sum_{1/2 < s_j \leq 1} c\pi^{1/2} \frac{\Gamma(s_j - \frac{1}{2})}{\Gamma(s_j + 1)} u_j(z)\overline{u}_j(w) X^{s_j} + O(X^{2/3}), \quad (3.3)$$

where $c = 2$, if $-1 \in \Gamma$, and $c = 1$, otherwise.

It is now clear that if there are no cuspidal eigenfunctions for $2/3 < s_j < 1$ or equally no $0 < \lambda_j < 2/9 \cong 0.222\ldots$, then

$$P(X) = \frac{c\pi}{|F|} X + O(X^{2/3}). \quad (3.4)$$

It is a consequence of the recent results of Kim–Shahidi [48] that:

Theorem 3.4. *The eigenvalues* $\lambda_j \geq \frac{66}{289} \cong 0.228\ldots$. *Thus (3.4) is valid.*

One can then conclude

$$\text{card}\{(a,b,c,d) \in \mathbb{Z}^4 | ad - bc = 1,\ a^2 + b^2 + c^2 + d^2 \leq X\} = 6X + O(X^{2/3})$$

since $|F| = \frac{\pi}{3}$ which is of course not new since the validity of Conjecture 3.2 is known for $SL_2(\mathbb{Z})$; but the fact that (3.4) is valid for an arbitrary congruence subgroup Γ is quite new. We refer to [33] for other number-theoretic consequences of the lattice point problem such as

$$\sum_{m \leq X} r(m)r(m+1) = 8X + O(X^{2/3}).$$

Another problem whose solution follows from the eigenvalue estimate stated in Theorem 3.4 is that of *shifted sums for Fourier coefficients of cusp forms* [26, 48].

More precisely, let

$$f(z) = \sum_{n=1}^{\infty} a_f(n) e^{2\pi i n z}$$

be a new normalized eigenform as discussed in Section 2. Let $h \in \mathbb{Z}$, $h \neq 0$ and consider

$$D_{f,h}(X) = \sum_{n \leq X} a_f(n) a_f(n+h), \tag{3.5}$$

where X is a large positive real number. Then again this sum can be asymptotically estimated by means of eigenvalues of Maass cusp forms in $L^2(\Gamma \backslash \mathfrak{h})$ and the estimate stated in Theorem 3.4 shows that

$$D_{f,h}(X) = O_{f,h}(X^{2/3}). \tag{3.6}$$

It is conjectured that in both problems the error term can be improved to $X^{1/2}$, but this is presently out of reach.

We conclude this section by stating the best present estimate for λ_j due to Kim and Sarnak [46].

Theorem 3.5 ([46]). $\lambda_j \geq \frac{1}{4} - \left(\frac{7}{64}\right)^2 = 0.2380371.$

4 Ramanujan conjecture for Maass forms

Let f be a normalized Maass cusp form for $\Gamma_0(N)$ which is an eigenfunction for all the Hecke operators and $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$. Write

$$\Delta f = \frac{1}{4}(1 - s^2) f. \tag{4.1}$$

Then (cf. [34], pg. 132 or [33], equations (1.26), (1.34) and (3.4))

$$f(x+iy) = \sum_{n \neq 0} (|n|y)^{1/2} a_n K_{s/2}(2\pi |n| y) e^{2\pi i n x}, \tag{4.2}$$

where K_ν is the Whittaker–Bessel function satisfying

$$z^2 \frac{d^2 K_\nu}{dz^2} + z \frac{dK_\nu}{dz} - (z^2 + \nu^2) K_\nu = 0 \tag{4.3}$$

and

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}}\, e^{-z} \qquad (z \in \mathbb{R}). \tag{4.4}$$

Here the a_n's are the Fourier coefficients of f, although f is only real analytic.

The Ramanujan conjecture which was proved by Deligne for holomorphic forms [18] is expected to be valid for these functions so that (cf. [34], pg. 132):

Conjecture 4.1. Suppose p is a prime. Then $|a_p| \leq 2p^{-1/2}$.

If we write

$$a_p = p^{-1/2}(\alpha_p + \alpha_p^{-1}),$$

with $\alpha_p \in \mathbb{C}$, then Conjecture 4.1 is equivalent to

$$|\alpha_p| = 1. \tag{4.5}$$

It should be pointed out that the Selberg's conjecture 3.2 is the archimedean analogue of that of Ramanujan, i.e., Conjecture 4.1, which can easily be shown to be equivalent to s being pure imaginary.

Over $\mathbb{Q}$, i.e., for classical Maass forms, the best estimate at present is due to Kim–Sarnak as follows (cf. [46]). It is again a consequence of recent progress in Langlands functoriality (cf. [1, 43, 48, 54]).

Theorem 4.2 ([46]). *Suppose p is a prime. Then*

$$|a_p| \leq p^{-1/2}(p^{7/64} + p^{-7/64}) \tag{4.6}$$

or equivalently,

$$p^{-7/64} \leq |\alpha_p| \leq p^{7/64}. \tag{4.7}$$

We will discuss the best estimates over an arbitrary number field later in Section 6.

Remark 4.3. The lower bound in the inequality (3.12) in page 390 of [81] is neither expected, nor implied by (4.7). It should be considered as a typo and disregarded. It is not clear to us how it appeared! It was noticed only when the volume was already in press.

5 Sato–Tate conjecture

Let f be a holomorphic cuspidal normalized new eigenform of weight $k > 1$ and level N. Assuming that f satisfies the Ramanujan conjecture, we may even include Maass forms. Moreover, assume that the Nebentypus ε of f is trivial and that f is not of CM-type. Then, given a rational prime $p \nmid N$, let a_p denote the corresponding Fourier coefficient of f. Write

$$a_p = p^{\frac{k-1}{2}}(\alpha_p + \alpha_p^{-1}), \tag{5.1}$$

with $\alpha_p \in \mathbb{C}$. Then, the Ramanujan conjecture is equivalent to $|\alpha_p| = 1$. Write

$$\alpha_p = e^{i\theta_p} \tag{5.2}$$

where $0 \le \theta_p \le \pi$. Then the conjecture is equivalent to

$$a_p/p^{\frac{k-1}{2}} \in [-2, 2]. \tag{5.3}$$

Note that the matrix

$$e(\theta_p) = \begin{pmatrix} e^{i\theta_p} & 0 \\ 0 & e^{-i\theta_p} \end{pmatrix} \in SU(2) \tag{5.4}$$

for all $p \nmid N$, where

$$SU(2) = \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \in GL_2(\mathbb{C}) \middle| |a|^2 + |b|^2 = 1 \right\}. \tag{5.5}$$

Every conjugacy class in $SU(2)$ has a unique element of the form $e(t)$ for some $t \in \mathbb{R}$. The Haar measure on $SU(2)$ gives

$$d\mu = \frac{2}{\pi} \sin^2 t dt \tag{5.6}$$

as the measure on the conjugacy classes X of $SU(2)$, $0 \le t \le \pi$ (cf. [69, 76]).

A sequence $\{t_i\}$, $t_i \in \mathbb{R}$, is equidistributed with respect to measure $d\mu$ defined by (5.6) if for any $\varphi \in C(X)$

$$\int_0^\pi \varphi(e(t)) \frac{2}{\pi} \sin^2 t dt = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \varphi(e(t_i)). \tag{5.7}$$

One can then conjecture:

Conjecture 5.1 *((Sato–Tate)* **(cf. [69])**). The sequence $\{\theta_p\}$ is equidistributed with respect to $d\mu$.

As explained in [68, 76], to prove these conjectures one needs to know certain properties of the symmetric power L-functions

$$L_m(s, f) = \prod_{p \nmid N} \prod_{j=0}^m (1 - \alpha_p^j \beta_p^{m-j} p^{-s})^{-1} \tag{5.8}$$

for all $m \in \mathbb{Z}^+$, where $\beta_p = \alpha_p^{-1}$. More precisely, one needs to know the holomorphy and non-vanishing for each $L_m(s, f)$ on $Re(s) = 1$. Observe that by validity of the Ramanujan conjecture $|\alpha_p| = |\beta_p| = 1$, each infinite product in (5.8) is absolutely convergent for $Re(s) > 1$. Incidentally, as explained by Langlands in [54], this absolute convergence for all m will by itself imply the validity of the Ramanujan conjecture (cf. [76] for a survey of both these conjectures).

While some progress is made in [47, 76] on proving the non-vanishing of $L_m(s, f)$ on $Re(s) = 1$ for $m \leq 9$, again based on recent progress on functoriality [43, 48], we are still quite far from proving Conjecture 5.1. (We also refer to [47, 82] for the holomorphy on the line which is no longer guaranteed if f is of CM-type or $k = 1$.) On the other hand, following certain ideas of Serre [69], explained as an appendix in [76], the following result is proved in a recent paper of Kim–Shahidi [47], using the holomorphy and non-vanishing results for $L_m(s, f)$ on $Re(s) = 1$ for $m \leq 9$ discussed above. It may be interpreted as evidence towards the validity of Conjecture 5.1.

Proposition 5.2 ([47]). *For every $\varepsilon > 0$, there are sets S_1 and S_2 of prime numbers of positive lower density such that $a_p > c - \varepsilon$ for $p \in S_1$ and $a_p < -c + \varepsilon$ for $p \in S_2$, where $c = 2\cos(2\pi/11) \simeq 1.68\ldots$*

Remark. Richard Taylor now seems to have a proof of this conjecture for elliptic curves (weight two modular forms) with at least one multiplicative reduction over $\mathbb{Q}$ (or any totally real field).

6 Functoriality for symmetric powers

Theorems 2.3, 2.4, 3.4, 3.5, 4.2 and Proposition 5.2, which are all of arithmetic significance, are all consequences of recent cases of functoriality established in [43, 48] which we shall now explain.

It is well known [21, 35] that each normalized cuspidal newform f which is an eigenform for all the Hecke operators and the hyperbolic Laplacian, holomorphic or not, can be realized as a subrepresentation $\pi_f = \pi$ of $L^2(GL_2(\mathbb{Q})\backslash GL_2(\mathbb{A}_{\mathbb{Q}}), \omega)$, for a *grössencharacter* ω of $\mathbb{Q}$. It is therefore natural to consider the constituents of $L^2(GL_2(F)\backslash GL_2(\mathbb{A}_F), \omega)$ for any number field F. Infinite-dimensional irreducible subrepresentations of these L^2-spaces (as ω varies) are the so-called cuspidal automorphic representations of $GL_2(\mathbb{A}_F)$. Given such a representation π, one can write $\pi = \otimes_v \pi_v$ in which π determines the class of each π_v, an irreducible unitary representation of $GL_2(F_v)$, uniquely. Here F_v is the completion of F at a place v and if O_v is the ring of integers of F_v for each finite place v, then π_v will have a vector fixed by $GL_2(O_v)$ for almost all such v. Such a representation is called an *unramified* or *spherical* (sometimes *class one*) representation of $GL_2(F_v)$. If π_v is unramified, then

$$\pi_v = \mathrm{Ind}_{B_2(F_v)}^{GL_2(F_v)} \chi_{1v} \otimes \chi_{2v} \otimes \mathbf{1},$$

where $\chi_{iv}, i = 1, 2$ are unramified quasi-characters of F_v^*, with $\chi_{1v} \otimes \chi_{2v}$ considered as a character of diagonal subgroup of the upper triangular elements $B(F_v)$ of $GL_2(F_v)$. Set $\alpha_v = \chi_{1v}(\varpi_v)$ and $\beta_v = \chi_{2v}(\varpi_v)$, where ϖ_v is a uniformizing parameter of O_v or a generator of its maximal ideal. Then the (semisimple) conjugacy class of

$$t_v = \begin{pmatrix} \alpha_v & 0 \\ 0 & \beta_v \end{pmatrix}$$

in $GL_2(\mathbb{C})$ determines the class of π_v uniquely. Note that when $F = \mathbb{Q}$ and f is a normalized new eigenform of weight k, then $a_p = p^{\frac{k-1}{2}}(\alpha_p + \beta_p)$.

Now, let Sym^m denote the m-th symmetric power representation of $GL_2(\mathbb{C})$ discussed in Section 2. Then

$$\mathrm{Sym}^m(t_v) = \mathrm{diag}(\alpha_v^m, \alpha_v^{m-1}\beta_v, \ldots, \beta_v^m).$$

The conjugacy class of $\mathrm{Sym}^m(t_v)$ in $GL_{m+1}(\mathbb{C})$ then determines a unique class of unramified representations $\mathrm{Sym}^m\pi_v$ of $GL_{m+1}(F_v)$.

Now, let S be a finite set of places of F including archimedean ones, such that for every $v \notin S$, π_v is unramified. The *Langlands functoriality conjecture* in this case can be formulated as

Conjecture 6.1. There exists an automorphic representation $\mathrm{Sym}^m\pi = \prod = \otimes_v \prod_v$ of $GL_{m+1}(\mathbb{A}_F)$, i.e., an irreducible constituent (either discrete or continuous) of $L^2(\mathbb{A}_F^* GL_{m+1}(F)\backslash GL_{m+1}(\mathbb{A}_F))$ such that $\Pi_v = \mathrm{Sym}^m\pi_v$ for all $v \notin S$.

The transfer from π to $\mathrm{Sym}^m\pi$ corresponds to the homomorphism

$$\mathrm{Sym}^m\colon GL_2(\mathbb{C}) \to GL_{m+1}(\mathbb{C})$$

from the L-group $GL_2(\mathbb{C})$ of GL_2 to the L-group $GL_{m+1}(\mathbb{C})$ of GL_{m+1} as usually used to formulate the conjecture. Functoriality for the pair (GL_2, GL_{m+1}) predicts that every homomorphism between $GL_2(\mathbb{C})$ and $GL_{m+1}(\mathbb{C})$ must lead to a transfer of automorphic forms from $GL_2(\mathbb{A}_F)$ to those of $GL_{m+1}(\mathbb{A}_F)$.

The present state-of-the-art allows us to prove a very precise version of functoriality conjecture for $m = 3$ and 4. More precisely, let at each place v, W_v' be the Weil–Deligne group at v, a natural extension of $\mathrm{Gal}(\overline{F}_v/F_v)$ (cf. [85]).

The local Langlands conjecture proved in [28, 31, 50, 57] attaches a two-dimensional representation

$$\phi_v\colon W_v' \to GL_2(\mathbb{C})$$

to each π_v (in fact, every irreducible admissible representation of $GL_2(F_v)$), preserving Artin root numbers and L-functions.

The composite

$$\mathrm{Sym}^m\phi_v\colon W_v' \to GL_{m+1}(\mathbb{C})$$

is then a $(m+1)$-dimensional representation of W_v' and the recent results of Harris–Taylor [28] and Henniart [30] then attaches an irreducible admissible representation $\mathrm{Sym}^m\pi_v$ of $GL_{m+1}(F_v)$. The precise version of functoriality conjecture then demands that

$$\mathrm{Sym}^m\pi = \otimes_v \mathrm{Sym}^m\pi_v$$

be an automorphic representation of $GL_{m+1}(\mathbb{A}_F)$.

While this conjecture was proved for $m = 2$ in [23] in 1978, the following theorem is quite recent (2002) and is proved in [43, 48].

Theorem 6.2 (Kim–Shahidi [48] and Kim [43]). *The representation $\mathrm{Sym}^m\pi$ is an automorphic representation of $GL_{m+1}(\mathbb{A}_F)$ for $m = 3$ and $m = 4$.*

In the case $m = 3$, it is proved in [47, 48] that $\mathrm{Sym}^3\pi$ is cuspidal unless π is dihedral or tetrahedral in which case both the Ramanujan and the Selberg conjectures are valid for π.

When $\mathrm{Sym}^3\pi$ is cuspidal, one can use the estimates in [58] to prove that

$$q_v^{-\left(\frac{1}{2}-\frac{1}{17}\right)} \leq |\alpha_v^3| = |\beta_v^{-3}| \leq q_v^{\frac{1}{2}-\frac{1}{17}}$$

leading to

$$q_v^{-5/34} \leq |\alpha_v| \text{ and } |\beta_v| \leq q_v^{5/34}$$

at the finite places.

When $F = \mathbb{Q}$ and $F_v = \mathbb{R}$, the archimedean analogue of these estimates is

$$\lambda_j \geq \frac{1}{4} - \left(\frac{5}{34}\right)^2 = \frac{66}{289},$$

stated as Theorem 3.4 in Section 3. As explained in that section, this leads to a proof of (3.4).

The best estimate for $|\alpha_v|$ and $|\beta_v|$ for forms over an arbitrary number field requires the automorphy of both $\mathrm{Sym}^3\pi$ and $\mathrm{Sym}^4\pi$. It gives [47]

$$q_v^{-1/9} < |\alpha_v| \text{ and } |\beta_v| < q_v^{1/9}$$

with similar estimates at archimedean places [44] which is slightly weaker than 7/64 in [46] proved for Maass forms (over $\mathbb{Q}$).

7 Functoriality for classical groups

Langlands functoriality conjecture is expected to be valid in the generality of every pair (G, G') of connected reductive groups over a local or global field F. For simplicity of exposition, let us assume that $G' = GL_N$ for some positive integer N. Let ${}^LG$ denote the (complex) L-group of G. Then ${}^LG' = GL_N(\mathbb{C})$. Let us recall that ${}^LG = Sp_{2n}(\mathbb{C})$, $SO_{2n}(\mathbb{C})$ and $SO_{2n+1}(\mathbb{C})$, for $G = SO_{2n+1}$, SO_{2n} and Sp_{2n} with all the groups split over F, respectively (cf. [7, 12, 54]).

An irreducible admissible representation $\pi = \otimes_v \pi_v$ of $G(\mathbb{A}_F)$ is called automorphic if it appears in $L^2(G(F)\backslash G(\mathbb{A}_F))$ (either continuously or discretely). Again, as in the case of GL_N, classes of almost all π_v's (called unramified representations), are parametrized by semisimple conjugacy classes in ${}^LG_v$, the L-group of G as a group over F_v. Let t_v denote this conjugacy class for each such π_v. (They will have vectors fixed under the action of $G(O_v)$ which makes sense since G will be quasisplit and defined over O_v for almost all v). To make matters even easier let us assume that G is split over F. Let ${}^LG^0$ denote the connected component of ${}^LG$. We shall now formulate the functoriality conjecture as follows. Let S be a non-empty set of places such that for each $v \notin S$, π_v is unramified.

Conjecture 7.1. Given a homomorphism

$$\rho\colon {}^LG^0 \to GL_N(\mathbb{C}),$$

there exists an automorphic representation $\Pi = \otimes_v \Pi_v$ of $GL_N(\mathbb{A}_F)$ such that Π_v is the unramified representation of $GL_N(F_v)$ attached to the conjugacy class of $GL_N(\mathbb{C})$ generated by $\rho(t_v)$ for all $v \notin S$. In other words, ρ is functorial.

Now, let $G = SO_{2n+1}, SO_{2n}$ or Sp_{2n}, all split over F; or even $G\mathrm{Spin}_m$, the general spin group of semisimple rank $[\frac{m}{2}]$; a group whose derived group is Spin_m, the two sheet algebraic covering of SO_m, defined by means of a particular central element (cf. [3, 4]). Then ${}^LG\mathrm{Spin}_m = GSO_m$ or $GSp_{2[\frac{m}{2}]}$ according as m is even or odd, respectively. In particular, in each case we have an embedding

$$i\colon {}^LG^0 \to GL_N(\mathbb{C}). \tag{7.1}$$

We shall assume N is minimal.

Finally assume π is *globally generic*, i.e., it has a non-vanishing Fourier coefficient with respect to a generic character of the unipotent radical of the subgroup of upper triangulars (a Borel subgroup). We refer to [62, 74, 75, 80] for appropriate definitions. We then have:

Theorem 7.2 ([4, 5, 16, 17]). *Assume F is a number field. The embedding i is functorial whenever π is globally generic.*

The case of classical groups is due to Cogdell–Kim–Piatetski-Shapiro–Shahidi [16, 17] and that of general spin groups was proved by Asgari–Shahidi [4].

While the transfer in [17] is strong in the sense that it is accomplished for even π_v with $v \in S$, the case proved in [4] is still weak.

Remark 7.3. When $G = G\mathrm{Spin}_5 = GSp_4$, this establishes the generic transfer of automorphic forms from $GSp_4(\mathbb{A}_F)$ to $GL_4(\mathbb{A}_F)$. This is proved in [5] in its strongest form where the cuspidality of the image is also determined. This, in particular, proves the holomorphy of the 4-dimensional spinor L-functions for GSp_4 on all of $\mathbb{C}$ for generic forms which is of considerable interest in number theory.

Remark 7.4. The general transfer for the classical groups where π is not necessarily assumed to be generic is expected to follow from Arthur's trace formula. This requires different forms of fundamental lemma whose proofs are still in progress.

8 Ramanujan conjecture for classical groups

Let G be a connected reductive group over a global field F. Let π be a cuspidal representation of $G(\mathbb{A}_F)$. Write $\pi = \otimes_v \pi_v$. The representation π is tempered if and only if each π_v is. This means that its matrix coefficients are in $L^{2+\varepsilon}(Z(F_v)\backslash G(F_v))$ for all $\varepsilon > 0$.

When $G = GL_2$, the Ramanujan and the Selberg conjectures are tantamount to cuspidal representations π of $GL_2(\mathbb{A}_{\mathbb{Q}})$ being tempered; similarly for $GL_2(\mathbb{A}_F)$ for any number field F. More generally, it is expected that cuspidal representations of $GL_n(\mathbb{A}_F)$ are all tempered for any n and any number field F. (For a function field this is a theorem due to Lafforgue [51] for any n. The case of $n = 2$ was established earlier by Drinfeld [20].)

For a general reductive group this may not be true. Counter-examples exist already for Sp_4 and $U(2,1)$ (cf. [32, 49]). None are generic. On the other hand one expects the following version of the Ramanujan conjecture to be true (cf. [17, 62, 74, 80]).

Conjecture 8.1. Let G be a quasisplit connected reductive group over a global field F. Let π be a globally generic cuspidal representation of $G(\mathbb{A}_F)$. Then π is tempered.

This is in agreement with the widely expected validity of the conjecture for $GL_n(\mathbb{A}_F)$ as every cuspidal representation of $GL_n(\mathbb{A}_F)$ is globally generic [83].

One of the consequences of Theorem 7.2 is the following result proved in [17].

Theorem 8.2 ([17]). *a) Assume that the Ramanujan conjecture is valid for $GL_m(\mathbb{A}_F)$ for all $m \le N$; then it is valid for the globally generic cuspidal spectrum of $SO_{2n}(\mathbb{A}_F)$ and $SO_{2n+1}(\mathbb{A}_F)$ with $N = 2n$, as well as that of $Sp_{2n}(\mathbb{A}_F)$ with $N = 2n + 1$.*

b) Write $\pi = \otimes_v \pi_v$, where π is a globally generic cuspidal representation of either $SO_m(\mathbb{A}_F)$ or $Sp_{2n}(\mathbb{A}_F)$. Write

$$\pi_v \simeq \mathrm{Ind}(\tau_{1,v}|\det(\cdot)|^{b_{1,v}} \otimes \ldots \otimes \tau_{t,v}|\det(\cdot)|^{b_{t,v}} \otimes \tau_{0,v}),$$

where $b_{1,v} > \ldots > b_{t,v}$ with $\tau_{i,v}$ irreducible tempered representations of appropriate GL-groups and $\tau_{0,v}$ is a generic tempered representation of a similar classical group but of lower rank (Muić). Then

$$-\left(\frac{1}{2} - \frac{1}{N^2+1}\right) \le b_{i,v} \le \frac{1}{2} - \frac{1}{N^2+1}, \; 1 \le i \le t.$$

Here N is the minimal rank of $GL_N(\mathbb{C})$ in which ${}^LG$ is embedded (cf. (7.1)).

We point out that in part b) of Theorem 8.2 no assumptions concerning the validity of the Ramanujan conjecture for $GL_n(\mathbb{A}_F)$ is necessary as this time one uses certain estimates of Luo–Rudnick–Sarnak in the Ramanujan conjecture for $GL_N(\mathbb{A}_F)$ proved in [58].

We conclude this section by referring to [5] for results concerning globally generic cuspidal spectrum of $GSp_4(\mathbb{A}_F)$. In fact, besides the new estimates proved in [5] towards the conjecture, we should mention that, as it is proved in [5], they are in fact also weakly Ramanujan, i.e., that they come as close as possible to be tempered at all their unramified places. This is done by using functoriality (Remark 7.3) to reduce the question to forms on $GL_4(\mathbb{A}_F)$ for which the weak Ramanujan property is proved in [43].

9 The method

The cases of functoriality conjecture proved in [4, 5, 16, 17, 43, 45, 48] are all established by applying converse theorems of Cogdell and Piatetski-Shapiro [14, 15] to analytic properties of certain automorphic L-functions obtained from the Langlands–Shahidi method (cf. [10, 25, 41, 55, 56, 70, 72, 73, 74, 75, 79]).

We will discuss this through the case of functoriality of $\mathrm{Sym}^3\pi$ for a cuspidal representation π of $GL_2(\mathbb{A}_F)$. This is proved by establishing another transfer of automorphic forms. More precisely, consider the homomorphism

$$\rho\colon GL_2(\mathbb{C}) \times GL_3(\mathbb{C}) \to GL_6(\mathbb{C}) \qquad (9.1)$$
$$(g_1, g_2) \mapsto g_1 \otimes g_2.$$

It is proved in [48] that ρ is functorial, again in the strongest possible form as we now explain.

Let $\pi_1 = \otimes_v \pi_{1v}$ and $\pi_2 = \otimes_v \pi_{2v}$ be cuspidal representations of $GL_2(\mathbb{A}_F)$ and $GL_3(\mathbb{A}_F)$, respectively. By the local Langlands conjecture proved in these cases by Kutzko [50], Henniart [30] and Langlands ([57] for archimedean places and in full generality), there exist homomorphisms

$$\phi_{iv}\colon W'_{F_v} \to GL_{i+1}(\mathbb{C}) \quad (i = 1, 2), \qquad (9.2)$$

parametrizing π_{iv}, $i = 1, 2$, preserving root numbers and L-functions. Then

$$\phi_{1v} \otimes \phi_{2v}\colon W'_{F_v} \to GL_6(\mathbb{C}) \qquad (9.3)$$

is a 6-dimensional representation of W'_{F_v} for each v. By recent results of Harris–Taylor [28] and Henniart [30] (and Langlands [57]) $\phi_{1v} \otimes \phi_{2v}$ defines an irreducible admissible representation of $GL_6(F_v)$ which we denote by $\pi_{1v} \boxtimes \pi_{2v}$. Set

$$\pi_1 \boxtimes \pi_2 = \otimes_v (\pi_{1v} \boxtimes \pi_{2v}). \qquad (9.4)$$

Note that for all the unramified places $\pi_{1v} \boxtimes \pi_{2v}$ is the unramified representation determined by $\rho(t_{1v} \otimes t_{2v})$, where t_{iv} is the semisimple conjugacy class parametrizing π_{iv}, $i = 1, 2$. The following theorem is one of the main results of Kim–Shahidi proved in [48].

Theorem 9.1 ([48]). *The irreducible admissible representation $\pi_1 \boxtimes \pi_2$ of $GL_6(\mathbb{A}_F)$ is automorphic, i.e., it appears in $L^2(GL_6(F)\backslash GL_6(\mathbb{A}_F), \omega_{\pi_1}^3 \omega_{\pi_2}^2)$, where ω_{π_i} is the central character of π_i, $i = 1, 2$.*

Corollary 9.2 ([48]). *Let π be a cuspidal representation of $GL_2(\mathbb{A}_F)$. Then $\mathrm{Sym}^3\pi$ is an automorphic representation of $GL_4(\mathbb{A}_F)$.*

Proof. One needs to use the identity

$$\pi \boxtimes \mathrm{Sym}^2\pi = \mathrm{Sym}^3\pi \boxplus (\pi \otimes \omega_\pi), \qquad (9.5)$$

together with the classification theorem of Jacquet and Shalika [39] and a simple argument using L-functions to conclude the automorphy of $\mathrm{Sym}^3\pi$.

We shall now give a sketch of the proof of Theorem 9.1.

Let S be a non-empty finite set of finite places such that for every $v \notin S$, π_{1v} and π_{2v} are unramified. Let η be a *grössencharacter* of F which is highly ramified at least at one place $v \in S$. Given a positive integer n, let $\tau_n(S)$ be the set of cuspidal representations of $GL_n(\mathbb{A}_F)$, which are unramified at all $v \in S$. Set

$$\tau_n(\eta, S) = \tau_n(S) \otimes (\eta \cdot \det). \tag{9.6}$$

Let $\sigma = \otimes_v \sigma_v \in \tau_n(\eta, S)$ for $n = 1, 2, 3$ or 4. For each v, let

$$\Sigma_v : W'_{F_v} \to GL_n(\mathbb{C}) \tag{9.7}$$

be the n-dimensional representation of W'_{F_v} parametrizing σ_v ([28, 30, 57]). Set

$$L(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v) = L(s, \phi_{1v} \otimes \phi_{2v} \otimes \Sigma_v), \tag{9.8}$$

where the L-function on the right is that of Artin [53].

Given a non-trivial additive character $\psi = \otimes_v \psi_v$ of $F\backslash\mathbb{A}_F$, we let

$$\varepsilon(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v, \psi_v) = \varepsilon(s, \phi_{1v} \otimes \phi_{2v} \otimes \Sigma_v, \psi_v) \tag{9.9}$$

be the corresponding Artin root number.

We recall that by the results in [28, 30, 57], both $L(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v)$ and $\varepsilon(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v, \psi_v)$ are factors defined for the Rankin products of $GL_6(F_v) \times GL_n(F_v)$ by both the Rankin–Selberg method ([12, 36, 38, 39]) and that of Langlands–Shahidi ([10, 55, 60, 71, 72, 73]).

Finally set

$$L(s, \pi_1 \boxtimes \pi_2) \times \sigma) = \prod_v L(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v), \tag{9.10}$$

and

$$\varepsilon(s, (\pi_1 \boxtimes \pi_2) \times \sigma) = \prod_v \varepsilon(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v, \psi_v). \tag{9.11}$$

An appropriate version of the converse theorems of Cogdell and Piatetski-Shapiro [11, 12, 13, 14, 15] in this case can be formulated as:

Theorem 9.3 ([15]). *Assume that for all $\sigma \in \tau_n(\eta, S)$ with $n = 1, 2, 3$ or 4*

a) $L(s, (\pi_1 \boxtimes \pi_2) \times \sigma)$ is entire,

b) is bounded in vertical strips of finite width and satisfies

c) $L(s, (\pi_1 \boxtimes \pi_2) \times \sigma) = \varepsilon(s, (\pi_1 \boxtimes \pi_2) \times \sigma) L(1 - s, (\tilde{\pi}_1 \boxtimes \tilde{\pi}_2) \times \tilde{\sigma})$.

Then there exists an automorphic representation $\Pi' = \otimes_v \Pi'_v$ of $GL_6(\mathbb{A}_F)$ such that $\Pi'_v = \pi_{1v} \boxtimes \pi_{2v}$ for all $v \notin S$.

Therefore to apply Theorem 9.3 one needs to verify the validity of 9.3.a, 9.3.b and 9.3.c. A priori, this seems quite out of reach since (9.10) and (9.11) are produced by multiplying together an infinite set of local factors, but a closer look proves otherwise. In fact, these properties can all be verified using Langlands–Shahidi method ([10, 25, 41, 55, 56, 70, 72, 73, 74, 75]) which we shall now explain in this case.

Starting with (π_1, π_2, σ), one builds a cuspidal representation $\Pi = \otimes_v \Pi_v$ of $M(\mathbb{A}_F)$, the adelic points of a maximal Levi subgroup of a connected reductive group G over F, in such a way that

$$L(s, (\pi_{1v} \boxtimes \pi_{2v}) \times \sigma_v) = L(s, \Pi_v) \tag{9.12}$$

for all $v \notin S$, where S is assumed large enough so that for all $v \notin S$, $\pi_{iv}, \sigma_v, \eta_v$ and ψ_v are all unramified; and

$$L_S(s, \Pi) = \prod_{v \notin S} L(s, \Pi_v) \tag{9.13}$$

appears in the constant term of the Eisenstein series built on $(G(\mathbb{A}_F), M(\mathbb{A}_F))$ by means of Π (induced from Π). Here is the table of possibilities in which M_D stands for the derived group of M and "sc" denotes the simply connected form of the group (cf. [48]):

n	G	M_D
1	SL_5	$SL_2 \times SL_3$
2	D_5^{sc}	$SL_2 \times SL_3 \times SL_2$
3	E_6^{sc}	$SL_2 \times SL_3 \times SL_3$
4	E_7^{sc}	$SL_2 \times SL_3 \times SL_4$

The machinary developed in [70, 71, 72, 73, 74, 75] defines appropriate local factors from our method and proves the functional equation 9.3.c. The local L-function $L(s, \Pi_v)$ is then what one usually calls a triple product L-function denoted by $L(s, \pi_{1v} \times \pi_{2v} \times \sigma_v)$. These factors are defined by methods of harmonic analysis through standard intertwining operators and Whittaker functionals and are very much of analytic nature. One bi-product of our results is the equalities

$$L(s, \pi_{1v} \times \pi_{2v} \times \sigma_v) = L(s, \phi_{1v} \otimes \phi_{2v} \otimes \sigma_v) \tag{9.14}$$

and

$$\varepsilon(s, \pi_{1v} \times \pi_{2v} \times \sigma_v, \psi_v) = \varepsilon(s, \phi_{1v} \otimes \phi_{2v} \otimes \sigma_v, \psi_v) \tag{9.15}$$

which are quite subtle since the Artin factors on the right-hand side are defined completely by arithmetic means.

The validity of statement 9.3.a is due to the fact that twisting by a highly ramified η_v kills all the symmetries in Π_v, and thus Π, under the action of Weyl group which leads to an Eisenstein series with no poles ([27, 56, 59]). As this is equivalent to regularity of the constant term of the Eisenstein series, it reflects itself in the L-functions appearing there and in particular $L(s, \Pi)$. This is Kim's observation (cf. [41], also see Proposition 2.1 of [48]).

The boundedness in vertical strips is a special case of a general result of Gelbart and Shahidi [25]. It exploits analytic properties of both the constant and the non-constant terms of the Eisenstein series [27, 56, 59, 61].

In all of these, there is an induction, proved in [74, 75], since in most interesting cases there are at least two L-functions appearing in the constant term. (In the cases of E_6^{sc} and E_7^{sc} there are 3 and 4 L-functions, respectively). One then inductively removes the L-functions other than the main L-function $L(s, \Pi)$, using another triple $(G^\vee(\mathbb{A}_F), M^\vee(\mathbb{A}_F), \Pi^\vee)$ for which a given other L-function of the setting $(G(\mathbb{A}_F), M(\mathbb{A}_F), \Pi)$ appears as the main one ([55, 74, 75]).

Finally, to show that

$$\Pi'_v = \pi_{1v} \boxtimes \pi_{2v} \tag{9.16}$$

everywhere, a lot more work is required. One needs to use base change, both normal and non-normal cubic [2, 37] as well as a result of Bushnell and Henniart [8] on local Galois representations and types, which was prepared as an appendix to supplement [48], to prove (9.16) everywhere. This is what is usually called the strong lift (or transfer) as opposed to the weak lift (or transfer) proved by Theorem 9.2 as a first step.

We refer to [12, 13, 80] for the sketches of the proofs of different parts of Theorem 7.1. The proof of functoriality of $\mathrm{Sym}^4\pi$, due to Kim [43], is also sketched in [31, 40, 77]. It requires establishing another functoriality, namely to show that

$$\Lambda^2 \colon GL_4(\mathbb{C}) \to GL_6(\mathbb{C})$$

is functorial.

Functoriality of $\mathrm{Sym}^4\pi$ then follows from

$$\Lambda^2(\mathrm{Sym}^3\pi) = (\mathrm{Sym}^4\pi \otimes \omega_\pi) \boxplus \omega_\pi^3 .$$

Incidentally, identifying $G\mathrm{Spin}_6(\mathbb{C})$ as a covering of $GL_4(\mathbb{C})$, the (weak) functoriality of Λ^2 becomes a special case of the results proved in [4].

References

1. Arthur, J.: The principle of functoriality, *Bull. Amer. Math. Soc.* (N.S.) **40**, no. 1, 39–53 (2002); Mathematical Challenges of the 21st century (Los Angeles, CA, 2000).
2. Arthur, J., Clozel, L.: Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula, *Ann. of Math. Studies*, **120**, Princeton Univ. Press, Princeton, NJ (1989).
3. Asgari, M.: Local L-functions for split spinor groups, *Canad. J. Math.*, **54**, 673–693 (2002).
4. Asgari, M., Shahidi, F.: Generic transfer for general spin groups, *Duke Math. J.*, **132**, 137–190 (2006).

5. Asgari, M., Shahidi, F.: Generic transfer form $GSp(4)$ to $GL(4)$, *Comp. Math.*, **142**, 541–550 (2006).
6. Booker, A.: Poles of Artin L-functions and the strong Artin conjecture, *Ann. of Math.*, **158**, 1089–1098 (2003).
7. Borel, A.: Automorphic L-functions, *Proc. Sympos. Pure Math.*, **33**, Part 2, 27–61 (1979).
8. Bushnell, C.J., Henniart, G.: On certain dyadic representations (Appendix to [46]), *Ann. of Math.*, **155**, 883–893 (2002).
9. Buzzard, K., Dickinson, M., Shepherd-Barron, N., Taylor, R.: On icosahedral Artin representations, *Duke Math. J.*, **109**, 283–318 (2001).
10. Casselman, W., Shalika, J.A.: The unramified principal series of p-adic groups II, The Whittaker function, *Comp. Math.*, **41**, 207–231 (1980).
11. Cogdell, J.W.: L-functions and Converse Theorems for $GL(n)$, IAS/Park City Lecture Notes, Park City, Utah (2002).
12. Cogdell, J.W.: Lectures on L-functions, Converse Theorems, and Functoriality for $GL(n)$. In: *Lectures on Automorphic L-Functions*, J.W. Cogdell, H.H. Kim and M.R. Murty, Fields Institute Monographs, Vol. 20, AMS, Providence, RI, pp. 1–96 (2004).
13. Cogdell, J.W., and Piatetski-Shapiro, I.I.: Converse theorems, functoriality, and applications to number theory, *Proceedings of the International Congress of Mathematicians*, Vol. II (Beijing, 2002), Higher Ed. Press, Beijing, 2002, 119–128.
14. Cogdell, J.W., Piatetski-Shapiro, I.I.: Converse theorems for GL_n, *Publ. Math. Inst. Hautes Études Sci.*, **79**, 157–214 (1994).
15. Cogdell, J.W., Piatetski-Shapiro, I.I.: Converse theorems for GL_n II, *J. Reine Angew. Math.*, **507**, 165–188 (1999).
16. Cogdell, J.W., Kim, H.H., Piatetski-Shapiro, I.I., Shahidi, F.: On lifting from classical groups to GL_N, *Publ. Math. Inst. Hautes Études Sci.*, **93**, 5–30 (2001).
17. Cogdell, J.W., Kim, H.H., Piatetski-Shapiro, I.I., Shahidi, F.: Functoriality for the classical groups, *Publ. Math. Inst. Hautes Études Sci.*, **99**, 163–233 (2004).
18. Deligne, P.: La conjecture de Weil. I, *Publ. Math. Inst. Hautes Études Sci.*, **43**, 273–307 (1974).
19. Deligne, P., Serre, J-P.: Formes modulaires de poids 1, *Ann. Sci.Éc. Norm. Supér.*, **7**, no. 4, 507–530 (1974).
20. Drinfeld, V.G.: Langlands' conjecture for $GL(2)$ over functional fields, *Proceedings of ICM*, Helsinki, 565–574 (1978).
21. Gelbart, S.: Automorphic Forms on Adèle Groups, Princeton University Press, Princeton, NJ, 1975, Annals of Mathematics Studies, no. 83.
22. Gelbart, S.: Three lectures on the modularity of $\overline{\rho}_{E,3}$ and the Langlands reciprocity conjecture, *Modular Forms and Fermat's Last Theorem*, Springer–Verlag, New York (1997), pp. 155–207.
23. Gelbart, S., Jacquet, H.: A relation between automorphic representations of $GL(2)$ and $GL(3)$, *Ann. Sci. École Norm. Sup.*, **11**, no. 4, 471–542 (1978).
24. Gelbart, S., Miller, S.: Riemann's zeta function and beyond, *Bulletin of AMS*, **41**, no. 1, 59–112 (2003).
25. Gelbart, S., Shahidi, F.: Boundedness of automorphic L-functions in vertical strips, *J. Amer. Math. Soc.*, **14**, no. 1, 79–107 (2001).
26. Good, A.: On various means involving the Fourier coefficients of cusp forms, *Math. Z.*, **183**, 95–129 (1983).
27. Harish-Chandra: *Automorphic forms on semisimple Lie groups*, Notes by J.G.M. Mars. Lecture Notes in Mathematics, no. 62, Springer–Verlag, Berlin (1968).

28. Harris, M., Taylor, R.: *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, Vol. 151, Princeton University Press, Princeton, NJ (2001).
29. Henniart, G.: La conjecture de Langlands locale pour $GL(3)$, *Mémoires de la Soc. Math. de France*, **11/12**, 1–186 (1984).
30. Henniart, G.: Une preuve simple des conjectures de Langlands pour $GL(n)$ sur un corps p-adique, *Invent. Math.*, **139**, 439–455 (2000).
31. Henniart, G.: Progrés récents en fonctorialité de Langlands. Séminaire Bourbaki, Vol. 2000/2001, *Astérisque*, No. 282, Exp. 890, 301–322 (2002).
32. Howe, R., Piatetski-Shapiro, I.I.: A counter-example to the "generalized Ramanujan conjecture" for (quasi)-split groups, *Proc. Sympos. Pure Math.*, **33**, Part 1, 315–322 (1979).
33. Iwaniec, H.: *Spectral Methods of Automorphic Forms*. 2nd ed., Graduate Studies in Mathematics, Vol. 53, American Mathematical Society, Providence, RI (2002).
34. Iwaniec, H., Kowalski, E.: *Analytic Number Theory*, Colloquium Publications, Vol. 53, American Mathematical Society, Providence, RI (2004).
35. Jacquet, H., Langlands, R.P.: *Automorphic Forms on GL_2*, Lect. Notes Math., **114**, Springer (1970).
36. Jacquet, H., Piatetski-Shapiro, I.I., Shalika, J.A.: Rankin–Selberg convolutions, *Amer. J. Math.*, **105**, no. 2, 367–464 (1983).
37. Jacquet, H., Piatetski-Shapiro, I.I., Shalika, J.A.: Relèvement cubique non-normal, *C.R. Acad. Sci. Paris Sér. I Math.*, **292**, 567–571 (1981).
38. Jacquet, H., Shalika, J.A.: On Euler products and the classification of automorphic representations, I. *Amer. J. Math.*, **103**, no. 3, 499–558 (1981).
39. Jacquet, H., Shalika, J.A.: On Euler products and the classification of automorphic representations, II. *Amer. J. Math.*, **103**, 777–815 (1981).
40. Kim, H.H.: *Automorphic L-functions*. In: Lectures on Automorphic L-functions, J.W. Cogdell, H.H. Kim and M.R. Murty, Fields Institute Monographs, Vol. 20, AMS, Providence, RI, 2004, pp. 97–201.
41. Kim, H.H.: Langlands–Shahidi method and poles of automorphic L-functions: application to exterior square L-functions, *Canad. J. Math.*, **51**, no. 4, 835–849 (1999).
42. Kim, H.H.: An example of non-normal quintic automorphic induction and modularity of symmetric powers of cusp forms of icosahedral type, *Invent. Math.*, **156**, 495–502 (2004).
43. Kim, H.H.: Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. With appendix 1 by D. Ramakrishnan and appendix 2 by H.H. Kim and P. Sarnak, *J. Amer. Math. Soc.*, **6**, no. 1, 139–183 (2002).
44. Kim, H.H.: On local L-functions and noramlized intertwining operators, *Canad. J. Math.*, **57**, 535–597 (2005).
45. Kim, H.H., Krishnamurthy, M.: Base change lift from unitary groups to GL_N, *IMRP*, 2005; no. 1, 1–52 (2005).
46. Kim, H.H., Sarnak, P.: Refined estimates towards the Ramanujan and Selberg conjectures. Appendix 2 to [41], *J. Amer. Math. Soc.*, **16**, no. 1, 175–181 (2002).
47. Kim, H.H., Shahidi, F.: Cuspidality of symmetric powers with applications, *Duke Math. J.*, **112**, no. 1, 177–197 (2002).
48. Kim, H.H., Shahidi, F.: Functorial products for $GL_2 \times GL_3$ and the symmetric cube for GL_2. With an appendix by C.J. Bushnell and G. Henniart, *Ann. of Math.*, **155**, no. 3, 837–893 (2002).
49. Kurokawa, N.: Examples of eigenvalues of Hecke operators on Siegel cusp forms of degree two, *Invent. Math.*, **49**, 149–165 (1978).
50. Kutzko, P.: The Langlands conjecture for GL_2 of a local field, *Ann. of Math.*, **112**, 381–412 (1980).

51. Lafforgue, L.: Chtoucas de Drinfeld et correspondance de Langlands, *Invent. Math.*, **147**, no. 1, 1–241 (2002).
52. Langlands, R.P.: Base Change for $GL(2)$, *Ann. of Math. Studies*, vol. 96, Princeton University Press, Princeton, NJ (1980).
53. Langlands, R.P.: *On Artin's L-functions*, Rice Univ. Studies, **56**, Houston, TX (1970).
54. Langlands, R.P.: *Problems in the theory of automorphic forms*. In Lecture Notes in Math. 170, 18–86, Springer–Verlag, Berlin–Heidelberg–New York (1970).
55. Langlands, R.P.: *Euler Products*, Yale Mathematical Monographs, 1. Yale University Press, New Haven, CT (1971).
56. Langlands, R.P.: *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics, Vol. 544, Springer–Verlag, Berlin (1976).
57. Langlands, R.P.: On the classification of irreducible representations of real algebraic groups. *Representation Theory and Harmonic Analysis on Semisimple Lie Groups*, Math. Surveys Monogr., Vol. 31, pp. 101–170, Amer. Math. Soc., Providence, RI (1989).
58. Luo, W., Rudnick, Z., Sarnak, P.: On the generalized Ramanujan conjecture for $GL(n)$. Automorphic forms, automorphic representations, and arithmetic, *Proc. Sympos. Pure Math.*, **66**, 301–310 (1999).
59. Moeglin, C., Waldspurger, J.-L.: *Spectral decomposition and Eisenstein series*, Cambridge Tracts in Mathematics, vol. 113, Cambridge University Press, Cambridge (1995).
60. Moeglin, C., Waldspurger, J.-L.: Le spectre résiduel de $GL(n)$, *Ann. Sci. Ećole Norm. Sup.*, **22**, 605–674 (1989).
61. Müller, W.: The trace class conjecture in the theory of automorphic forms, *Ann. of Math.*, **130**, 473–529 (1989).
62. Piatetski-Shapiro, I.I.: Multiplicity one theorems, *Proc. Sympos. Pure Math.*, **33**, Part 1, 209–212 (1979).
63. Ramakrishnan, D.: Modularity of the Rankin–Selberg L-series, and multiplicity one for $SL(2)$, *Ann. of Math.*, **152**, 45–111 (2000).
64. Ramakrishnan, D.: Modularity of solvable Artin representations of $GO(4)$-type, *IMRN*, No. 1, 1–54 (2002).
65. Sarnak, P.: Spectra of hyperbolic surfaces, *Bull. Amer. Math. Soc.* (N.S.), **40**, 441–478 (2003).
66. Sarnak, P.: Notes on the generalized Ramanujan conjectures. Fields Institute Lectures, June (2003).
67. Selberg, A.: On the estimation of Fourier coefficients of modular forms, *Proc. Sympos. Pure Math.*, **8**, 1–15 (1965).
68. Serre, J-P.: *Abelian l-adic Representations and Elliptic Curves*, W.A. Benjamin, New York (1968).
69. Serre, J-P.: Appendix to [84], in *Elliptic Curves and Related Topics*, CRM Proc. Lecture Notes, Vol. 4, AMS, Providence, RI, 175–180 (1994).
70. Shahidi, F.: Functional equations satisfied by certain L-functions, *Comp. Math.*, **37**, 171–207 (1978).
71. Shahidi, F.: Fourier transforms of intertwining operators and Plancherel measures for $GL(n)$, *Amer. J. Math.*, **106**, 67-111 (1984).
72. Shahidi, F.: On certain L-functions, *Amer. J. Math.*, **103**, no. 2, 297–355 (1981).
73. Shahidi, F.: Local coefficients as Artin factors for real groups, *Duke Math. J.*, **52**, no. 4, 973–1007 (1985).
74. Shahidi, F.: On the Ramanujan conjecture and finiteness of poles for certain L-functions, *Ann. of Math.*, **127**, no. 3, 547–584 (1988).
75. Shahidi, F.: A proof of Langlands' conjecture on Plancherel measures; complementary series for p-adic groups, *Ann. of Math.*, **132**, no. 2, 273–330 (1990).

76. Shahidi, F.: Symmetric power L-functions for $GL(2)$. In: *Elliptic curves and related topics*, with an appendix by J-P. Serre, CRM Proc. Lecture Notes, vol. 4, Amer. Math. Soc., Providence, RI, 159–182 (1994).
77. Shahidi, F.: Automorphic L-functions and functoriality, *Proceedings of the International Congress of Mathematicians*, Vol. II (Beijing, 2002), Higher Ed. Press, Beijing, 655–666 (2002).
78. Shahidi, F.: Langlands–Shahidi Method. IAS/Park City Lecture Notes, Park City, Utah (2002).
79. Shahidi, F.: Local coefficients as Mellin transforms of Bessel functions; Towards a general stability, *IMRN*, **2002**, no. 39, 2075–2119 (2002).
80. Shahidi, F.: On the Ramanujan conjecture for quasisplit groups, *Asian J. Math.*, **8**, 813–836 (2004).
81. Shahidi, F.: Functoriality and small eigenvalues of Laplacian on Riemann surfaces. In: Eigenvalues of Laplacians and other geometric operators, A. Grigor'yan and S-T. Yau (ed) *Surveys in Differential Geometry*, Vol. IX, International Press, Somerville, MA, pp. 385–400 (2004).
82. Shahidi, F.: Infinite dimensional groups and automorphic L-functions, *Pure and Applied Math.* Quarterly 1, no. 3, Special Issue: In Memory of Armand Borel, Part 2 of 3, 683–699 (2005).
83. Shalika, J.: The multiplicity one theorem for $GL(n)$, *Ann. of Math.*, **100**, 171–193 (1974).
84. Shimura, G.: *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press (1994).
85. Tate, J.: Number theoretic background, *Proc. Sympos. Pure Math.*, **33**, Part 2, 3–26 (1979).
86. Tunnell, J.: Artin's conjecture for representations of octahedral type, *Bull. Amer. Math. Soc.*, **5**, 173–175 (1981).

6

Discriminant of Certain $K3$ Surfaces

Ken-Ichi Yoshikawa

Graduate School of Mathematical Sciences
University of Tokyo
3-8-1 Komaba, Tokyo 153-8914, Japan
yosikawa@ms.u-tokyo.ac.jp

Dedicated to Professor Kyoji Saito on his sixtieth birthday

Summary. In this article we study the discriminant of those $K3$ surfaces with involution which were introduced and investigated by Matsumoto, Sasaki, and Yoshida. We extend several classical results on the discriminant of elliptic curves to the discriminant of Matsumoto–Sasaki–Yoshida's $K3$ surfaces.

Key words: $K3$ surface, automorphic form, moduli space, discriminant, analytic torsion

Subject Classifications: 58J52, 14J15, 14J28, 11F55, 32N15

1 Introduction – Discriminant of elliptic curves

Let $M(n, 2n; \mathbf{C})$ be the vector space of all complex $n \times 2n$-matrices and consider the following subset

$$M^o(n, 2n) = \left\{(\mathbf{a}_1, \dots, \mathbf{a}_{2n}) \in M(n, 2n; \mathbf{C});\ \mathbf{a}_{i_1} \wedge \cdots \wedge \mathbf{a}_{i_n} \neq \mathbf{0},\ \forall\, i_1 < \cdots < i_n\right\}.$$

On $M(n, 2n; \mathbf{C})$, the group $GL_n(\mathbf{C}) \times (\mathbf{C}^*)^{2n}$ acts by

$$(g, \lambda_1, \dots, \lambda_{2n}) \cdot A = g \cdot A \cdot \mathrm{diag}(\lambda_1, \dots, \lambda_{2n}), \qquad A \in M(n, 2n; \mathbf{C})$$

where $\mathrm{diag}(\lambda_1, \dots, \lambda_{2n})$ denotes the diagonal matrix $(\lambda_i\, \delta_{ij})$. The configuration space of $2n$ points (or $2n$ hyperplanes) in the general position of $\mathbf{P}^{n-1}$ is defined as

$$X^o(n, 2n) = GL_n(\mathbf{C}) \backslash M^o(n, 2n)/(\mathbf{C}^*)^{2n}.$$

*The author is partially supported by the Grants-in-Aid for Scientific Research for young scientists (B) 16740030, JSPS.

Let us consider the case $n = 2$. On $M^o(2,4)$, we have a family of curves $\pi : \mathcal{E} \to M^o(2,4)$ with fiber

$$\pi^{-1}(A) = E_A = \{((x_1 : x_2), y) \in \mathcal{O}_{\mathbf{P}^1}(2);\ y^2 = \prod_{i=1}^{4}(a_{i1}\,x_1 + a_{i2}\,x_2)\}, \tag{1.1}$$

where $(x_1 : x_2)$ denotes the homogeneous coordinates of $\mathbf{P}^1$. The natural projection $\mathrm{pr}_1 : E_A \to \mathbf{P}^1$ is a double covering with 4 branch points, so that E_A is an elliptic curve. It is classical that $X^o(2,4)$ is a moduli space of elliptic curves with level 2 structure.

We define the discriminant of $A \in M^o(2,4)$ by

$$\Delta_{(2,4)}(A) = \prod_{\{i,j\}\cup\{k,l\}=\{1,2,3,4\},\ i<j,\ k<l} \det(\mathbf{a}_i, \mathbf{a}_j)\ \det(\mathbf{a}_k, \mathbf{a}_l). \tag{1.2}$$

Set $dx = x_2\,dx_1 - x_1\,dx_2 = x_2^2 d(x_1/x_2)$, and define the norm of $\Delta_{(2,4)}(A)$ by

$$\|\Delta_{(2,4)}(A)\|^2 = \left(\frac{\mathrm{i}}{2\pi}\int_{E_A} \frac{dx}{y} \wedge \overline{\left(\frac{dx}{y}\right)}\right)^6 |\Delta_{(2,4)}(A)|^2. \tag{1.3}$$

Since $\|\Delta_{(2,4)}\|^2$ is invariant under the action of $GL_2(\mathbf{C}) \times (\mathbf{C}^*)^4$, it descends to a function on $X^o(2,4)$. There is an analytic expression of $\|\Delta_{(2,4)}\|$.

Let $\det^* \square_A$ be the regularized determinant of the Laplacian of E_A with respect to the normalized flat Kähler metric of volume 1. Since the isomorphism class of E_A is constant along the $GL_2(\mathbf{C}) \times (\mathbf{C}^*)^4$-orbit, $\det^*\square_A$ is constant on each $GL_2(\mathbf{C}) \times (\mathbf{C}^*)^4$-orbit. For all $A \in M^o(2,4)$, by [11] we get

$$\det{}^* \square_A = \|\Delta_{(2,4)}(A)\|^{-1/3}. \tag{1.4}$$

In fact, Eq. (1.4) follows from the classical Kronecker limit formula, which can be seen as follows. For $z \in \mathbf{C}$ and $\tau \in \mathbf{H} := \{\tau \in \mathbf{C};\ \operatorname{Im}\tau > 0\}$, let

$$\wp(z,\tau) = \frac{1}{z^2} + \sum_{(m,n)\in\mathbf{Z}^2\setminus\{(0,0)\}} \left\{\frac{1}{(z + m\tau + n)^2} - \frac{1}{(m\tau + n)^2}\right\}$$

be the Weierstrass $\wp$-function and set

$$A(\tau) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & -\wp\left(\frac{1}{2},\tau\right) & -\wp\left(\frac{\tau}{2},\tau\right) & -\wp\left(\frac{1+\tau}{2},\tau\right)\end{pmatrix} \in M^o(2,4).$$

By setting $x_1 = u\,x_2$ and $y = v\,x_2^2/2$ in (1.1), $E_{A(\tau)}$ is isomorphic to the cubic curve in $\mathbf{P}^2$ defined by the inhomogeneous equation in the variables u, v:

$$\begin{aligned} v^2 &= 4\left\{u - \wp\left(\frac{1}{2},\tau\right)\right\}\left\{u - \wp\left(\frac{\tau}{2},\tau\right)\right\}\left\{u - \wp\left(\frac{1+\tau}{2},\tau\right)\right\} \\ &= 4u^3 - g_2(\tau)\,u - g_3(\tau) \end{aligned}$$

with $g_2(\tau) = 60\, E_4(\tau)$ and $g_3(\tau) = 140\, E_6(\tau)$, where $E_k(\tau)$ denotes the Eisenstein series of weight k (cf. [28, p. 11]). Hence the complex torus $\mathbf{C}/\mathbf{Z} + \tau\,\mathbf{Z}$, $\tau \in \mathbf{H}$, is isomorphic to $E_{A(\tau)}$ via the map

$$f : \mathbf{C}/\mathbf{Z} + \tau\,\mathbf{Z} \ni z \to ((\wp(z) : 1), \wp'(z)\, x_2^2/2) \in \mathcal{O}_{\mathbf{P}^2}(2). \tag{1.5}$$

Since $dx/y = 4\, f^*(dz)$ by (1.5) and since $\Delta_{(2,4)}(A(\tau))^2 = g_2(\tau)^3 - 27\, g_3(\tau)^2$, Eq. (1.4) is deduced from the Kronecker limit formula:

$$\det{}^* \square_{A(\tau)} = C_1\, \|\Delta(\tau)\|^{-1/6}, \tag{1.6}$$

where $C_1 \neq 0$ is an absolute constant, $\Delta(\tau) = (2\pi)^{-12}(g_2(\tau)^3 - 27\, g_3(\tau)^2)$ is the Jacobi Δ-function and $\|\Delta(\tau)\|^2 = (\mathrm{Im}\tau)^{12}|\Delta(\tau)|^2$ is its Petersson norm. Recall that one has the following expressions of the Jacobi Δ-function:

$$\Delta(\tau) = \left(\prod_{\text{even}} \theta_{ab}(\tau)\right)^8 = q \prod_{n=1}^{\infty}(1 - q^n)^{24}, \qquad q = e^{2\pi i \tau}, \tag{1.7}$$

where $\theta_{ab}(\tau)$ denotes the theta constants.

In [39], we extended Eq. (1.6) to $K3$ surfaces with involution. Let us explain our results briefly. Let X be a $K3$ surface and let $\iota : X \to X$ be a holomorphic involution acting non-trivially on holomorphic 2-forms on X. Let $H^2_+(X, \mathbf{Z})$ be the invariant part of the ι-action on $H^2(X, \mathbf{Z})$. The free $\mathbf{Z}$-module $H^2(X, \mathbf{Z})$ of rank 22 endowed with the cup product is an even unimodular lattice of signature (3, 19) isometric to the $K3$ lattice $\mathbb{L}_{K3}$. By Nikulin, the topological type of ι is determined by $H^2_+(X, \mathbf{Z})$, which is a primitive 2-elementary hyperbolic sublattice of $H^2(X, \mathbf{Z})$. Let $M \subset \mathbb{L}_{K3}$ be a primitive 2-elementary hyperbolic sublattice with rank $r(M)$. The pair (X, ι) is called a 2-elementary $K3$ surface of type M if $H^2_+(X, \mathbf{Z}) \cong M$. The period of a 2-elementary $K3$ surface of type M lies in $\Omega^o_M = \Omega_M \setminus \mathcal{D}_M$, where Ω_M is isomorphic to a symmetric bounded domain of type IV of dimension $20 - r(M)$ and $\mathcal{D}_M$ is a divisor of Ω_M called the discriminant locus. The moduli space of 2-elementary $K3$ surfaces of type M is isomorphic to the quotient Ω^o_M / Γ_M, where Γ_M is an arithmetic subgroup of $\mathrm{Aut}(\Omega_M)$. We assume that $r(M) \leq 17$.

For a 2-elementary $K3$ surface (X, ι) of type M, we constructed an invariant $\tau_M(X, \iota)$ by using the equivariant analytic torsion of (X, ι) (cf. [5]). We regard τ_M as a function on the moduli space Ω^o_M / Γ_M. The main result of [39] is that τ_M is expressed as the norm of the "automorphic form" Φ_M on Ω_M characterizing the discriminant locus $\mathcal{D}_M$. Here Φ_M is an automorphic form with values in some Γ_M-equivariant coherent sheaf λ_M on Ω_M. If the fixed point set of ι consists of only rational curves, then $\lambda_M \cong \mathcal{O}_{\Omega_M}$ and hence Φ_M is an automorphic form in the classical sense. By Nikulin [33], there exist only seven isometry classes of lattices M with $r(M) \leq 17$ such that the fixed point set of a 2-elementary $K3$ surface of type M consists of only rational curves. Let $\mathbb{S}_k$, $1 \leq k \leq 7$, be those seven lattices, where $r(\mathbb{S}_k) = 10 + k$. In Section 6, we shall express $\Phi_{\mathbb{S}_k}$ as a Borcherds product [8]. Thus the infinite product expansion (1.7) shall be extended to 2-elementary $K3$ surfaces of type $\mathbb{S}_k$.

The case $k = 6$ is of particular interest. In [30], [31], [36], 2-elementary $K3$ surfaces of type $\mathbb{S}_6$ with level 2 structure were studied by Matsumoto, Sasaki, Yoshida; they proved that the moduli space of 2-elementary $K3$ surfaces of type $\mathbb{S}_6$ with level 2 structure is isomorphic to $X^o(3,6)$. In Section 7.3, we shall extend the definitions (1.2), (1.3) to 3×6-matrices and get a function $\|\Delta_{(3,6)}\|$ on the configuration space $X^o(3,6)$. By Freitag, there exist theta functions $\{\Theta\binom{a}{b}\}$ on the period domain $\Omega_{\mathbb{S}_6}$, ten of which are called even. We define the Matsumoto–Sasaki–Yoshida form $\Delta_{\mathrm{MSY}}(W)$ as the product of all even Freitag theta functions: $\Delta_{\mathrm{MSY}}(W) := \prod_{\mathrm{even}} \Theta\binom{a}{b}$. Let $\|\Delta_{\mathrm{MSY}}\|$ denote the Petersson norm of Δ_{MSY}, which descends to a function on $X^o(3,6)$. The main result here is the following identity, which can be regarded as an analogue of Eqs. (1.4), (1.6), (1.7) in dimension 2:

Theorem 1.1. *The following identity of functions on $X^o(3,6)$ holds:*

$$\tau_{\mathbb{S}_6} = C_2 \, \|\Delta_{(3,6)}\|^{-1/4} = C_3 \, \|\Delta_{\mathrm{MSY}}\|^{-1/2}, \tag{1.8}$$

where C_2, C_3 are non-zero absolute constants.

This article is organized as follows. In Section 2, we recall $K3$ surfaces with involution and their moduli spaces. In Section 3, we recall automorphic forms on the moduli space. In Section 4, we recall the invariant τ_M. In Section 5, we recall Borcherds products. In Section 6, we give an expression of $\tau_{\mathbb{S}_k}$ as the Petersson norm of an interesting Borcherds product, whose proof shall be given in the forthcoming paper [41]. In Section 7, we prove Eq. (1.8). In Section 8, we prove that the discriminant of smooth quartic hypersurfaces of $\mathbf{P}^3$ is expressed as the norm of an interesting Borcherds product.

2 $K3$ surfaces with involution and their moduli spaces

In this section we recall the definition of $K3$ surfaces with involution. We refer to [39] for more details about $K3$ surfaces with involution.

Let X be a compact, connected, smooth complex surface with canonical line bundle K_X. Then X is called a $K3$ surface if

$$H^1(X, \mathcal{O}_X) = 0, \qquad K_X \cong \mathcal{O}_X.$$

Every $K3$ surface is Kähler [2, Chap. 8 Th. 14.5]. By the second condition, there exists a nowhere vanishing holomorphic 2-form η_X on X. Notice that η_X is uniquely determined up to a non-zero constant. The cohomology group $H^2(X, \mathbf{Z})$ is a free $\mathbf{Z}$-module endowed with the cup product pairing. There exists an isometry of lattices:

$$\alpha\colon H^2(X, \mathbf{Z}) \cong \mathbb{L}_{K3} := \mathbb{U} \oplus \mathbb{U} \oplus \mathbb{U} \oplus \mathbb{E}_8 \oplus \mathbb{E}_8.$$

Here $\mathbb{U} = \binom{0\,1}{1\,0}$ and $\mathbb{E}_8$ is the *negative-definite* Cartan matrix of type E_8 under the identification of a lattice with its Gram matrix. The isometry α as above is called a *marking*, and the pair (X, α) is called a *marked* $K3$ *surface*. For a marked $K3$ surface (X, α), the point

$$\pi(X, \alpha) := [\alpha(\eta_X)] \in \mathbf{P}(\mathbb{L}_{K3\,\mathbf{C}}), \qquad \eta_X \in H^0(X, K_X) \setminus \{0\}$$

is called the period of X, where $L_{\mathbf{K}} := L \otimes \mathbf{K}$ for a lattice L and a field $\mathbf{K}$.

For a lattice L with bilinear form $\langle \cdot, \cdot \rangle$, we denote by $L(k)$ the lattice with bilinear form $k\langle \cdot, \cdot \rangle$. The set of roots of L is defined by $\Delta_L := \{d \in L;\ \langle d, d \rangle = -2\}$. The isometry group of L is denoted by $O(L)$. Let $L^\vee = \mathrm{Hom}_{\mathbf{Z}}(L, \mathbf{Z})$ be the dual lattice of L, which is naturally embedded into $L_{\mathbf{Q}}$. The finite abelian group $A_L := L^\vee / L$ is called the *discriminant group* of L. For a primitive sublattice $L \subset \mathbb{L}_{K3}$, $L^\perp$ denotes the orthogonal complement of L in $\mathbb{L}_{K3}$.

Definition 2.1. *For a primitive hyperbolic sublattice* $S \subset \mathbb{L}_{K3}$*, define*

$$\Omega_S = \Omega_{S^\perp} := \{[x] \in \mathbf{P}(S^\perp_{\mathbf{C}});\ \langle x, x \rangle = 0,\ \langle x, \bar{x} \rangle > 0\}.$$

We set $r(S) := \mathrm{rank}_{\mathbf{Z}} S$. Then $\dim \Omega_S = 20 - r(S)$. There are two connected components of Ω_S, each of which is biholomorphic to a symmetric bounded domain of type IV of dimension $20 - r(S)$ (cf. [2, Chap. 8, Lemma 20.1]).

Definition 2.2. *An even lattice* S *is said to be* 2*-elementary if there is an integer* $l \geq 0$ *with* $A_S \cong (\mathbf{Z}/2\mathbf{Z})^l$*. For a* 2*-elementary lattice* S*, set* $l(S) := \dim_{\mathbf{F}_2} A_S$.

Let $M \subset \mathbb{L}_{K3}$ be a primitive 2-elementary hyperbolic sublattice. Let I_M be the involution on $M \oplus M^\perp$ defined by

$$I_M(x, y) = (x, -y).$$

Then I_M extends uniquely to an involution on $\mathbb{L}_{K3}$. For $l \in M^\perp_{\mathbf{R}}$, we set

$$\mathcal{H}_l := \{[x] \in \Omega_M;\ \langle x, l \rangle = 0\}.$$

Then $\mathcal{H}_l \neq \emptyset$ if and only if $\langle l, l \rangle < 0$. We define

$$\mathcal{D}_M := \bigcup_{d \in \Delta_{M^\perp}} \mathcal{H}_d, \qquad \Omega^o_M := \Omega_M \setminus \mathcal{D}_M.$$

We regard $\mathcal{D}_M$ as a reduced divisor of Ω_M.

Definition 2.3. *A* $K3$ *surface* X *equipped with a holomorphic involution* $\iota\colon X \to X$ *is called a* 2*-elementary* $K3$ *surface if*

$$\iota^*|_{H^0(X, K_X)} = -1.$$

The pair (X, ι) *is called a* 2*-elementary* $K3$ *surface of type* M *if there exists a marking* α *of* X *with* $\iota^* = \alpha^{-1} \circ I_M \circ \alpha$.

Let (X, ι) be a 2-elementary $K3$ surface of type M and let α be a marking with $\iota^* = \alpha^{-1} \circ I_M \circ \alpha$. Let $\eta_X \in H^0(X, K_X) \setminus \{0\}$. Then $\pi(X, \alpha) \in \Omega^o_M$. The $O(M^\perp)$-orbit of $\pi(X, \alpha)$ is independent of the choice of a marking α with $\iota^* = \alpha^{-1} \circ I_M \circ \alpha$. The Griffiths period of (X, ι), which is denoted by $\varpi_M(X, \iota)$, is defined as the $O(M^\perp)$-orbit

$$\varpi_M(X, \iota) := O(M^\perp) \cdot \pi(X, \alpha) \in \Omega^o_M / O(M^\perp).$$

Theorem 2.4. *The coarse moduli space of* 2*-elementary* $K3$ *surfaces of type* M *is isomorphic to the analytic space* $\Omega^o_M / O(M^\perp)$.

Proof. See [39, Th. 1.8]. □

We set

$$\mathcal{M}_M := \Omega_M / O(M^\perp), \qquad \mathcal{M}^o_M := \Omega^o_M / O(M^\perp).$$

Let $\Omega^\pm_M$ be the connected components of Ω_M and set

$$O^+(M^\perp) := \{g \in O(M^\perp);\ g(\Omega^\pm_M) = \Omega^\pm_M\}.$$

Then $O^+(M^\perp) \subset O(M^\perp)$ is a subgroup of index 2 with $\mathcal{M}_M = \Omega^+_M / O(M^\perp)^+$ and $\mathcal{M}^o_M = (\Omega^+_M \setminus \mathcal{D}_M)/O^+(M^\perp)$. We consider Ω^+_M as the period domain for 2-elementary $K3$ surfaces of type M. By Baily–Borel–Satake, both of $\mathcal{M}_M$ and $\mathcal{M}^o_M$ are quasi-projective algebraic varieties.

The topological type of the set of fixed points of (X, ι) was determined by Nikulin. We need the following partial result. See [33] for the general cases.

Lemma 2.5. *Let* (X, ι) *be a* 2*-elementary* $K3$ *surface of type* M *and let*

$$X^\iota := \{x \in X;\ \iota(x) = x\}.$$

If $r(M) + l(M) = 22$*, then* X^ι *is the disjoint union of* $(r(M) - 10)$*-smooth rational curves.*

By the adjunction formula, a smooth irreducible curve of a $K3$ surface is rational if and only if its self-intersection number is equal to -2.

3 Automorphic forms on the moduli space

Throughout this section we assume that $M \subset \mathbb{L}_{K3}$ *is a primitive* 2*-elementary hyperbolic sublattice*. We recall the definition of automorphic forms on the period domain Ω^+_M and give its differential geometric characterization.

Let us fix a vector $\mathbf{l}_M \in M^\perp_{\mathbf{R}}$ with $\langle \mathbf{l}_M, \mathbf{l}_M \rangle \geq 0$. We set

$$j_M(\gamma, [z]) := \frac{\langle \gamma \cdot z, \mathbf{l}_M \rangle}{\langle z, \mathbf{l}_M \rangle} \qquad [z] \in \Omega^+_M, \quad \gamma \in O^+(M^\perp).$$

Since $\mathcal{H}_{\mathbf{l}_M} = \emptyset$, $j_M(\gamma, \cdot)$ is a nowhere vanishing holomorphic function on Ω^+_M.

Definition 3.1. *Let $\Gamma \subset O^+(M^\perp)$ be a cofinite subgroup. A holomorphic function $f \in \mathcal{O}(\Omega_M^+)$ is called an automorphic form for Γ of weight p if*

$$f(\gamma \cdot [z]) = \chi(\gamma)\, j_M(\gamma, [z])^p f([z]), \qquad [z] \in \Omega_M^+, \quad \gamma \in \Gamma,$$

where $\chi : \Gamma \to \mathbf{C}^*$ is a character.

Let $K_M([z])$ be the Bergman kernel function of Ω_M^+:

$$K_M([z]) := \frac{\langle z, \bar{z} \rangle}{|\langle z, \mathbf{l}_M \rangle|^2}.$$

For an automorphic form of weight p, the Petersson norm of f is the function on Ω_M^+ defined as

$$\|f([z])\|^2 := K_M([z])^p\, |f([z])|^2.$$

If $r(M) \leq 17$ and if $\Gamma \subset O^+(M^\perp)$ is a cofinite subgroup, then $\|f\|^2$ is a Γ-invariant C^∞ function on Ω_M^+, because the group $\Gamma/[\Gamma, \Gamma]$ is finite and abelian in this case.

Let ω_M be the Kähler form of the Bergman metric on Ω_M^+:

$$\omega_M := -dd^c \log K_M, \tag{3.1}$$

where $d^c = (\partial - \bar{\partial})/4\pi\mathrm{i}$ and hence $dd^c = \bar{\partial}\partial/2\pi\mathrm{i}$ for complex manifolds. For a divisor D on Ω_M^+, let δ_D be the Dirac δ-current on Ω_M^+ with support D.

Theorem 3.2. *Let $p \in \mathbf{N}$ and let D be a divisor on Ω_M^+. Let $\Gamma \subset O^+(M^\perp)$ be a cofinite subgroup. Let φ be a non-negative, Γ-invariant C^∞ function on $\Omega_M \setminus D$ satisfying $\log \varphi \in L^1_{\mathrm{loc}}(\Omega_M)$ and the equation of currents on Ω_M^+:*

$$dd^c \log \varphi = \delta_D - p\, \omega_M. \tag{3.2}$$

If $r(M) \leq 17$, then there exists an automorphic form F for Γ of weight p with zero divisor D such that $\varphi = \|F\|^2$.

Proof. Set $\psi = \varphi\, K_M^{-p}$. Then $\log \psi \in L^1_{\mathrm{loc}}(\Omega_M^+)$. We get the following equation of currents on Ω_M^+ by (3.1), (3.2):

$$dd^c \log \psi = \delta_D,$$

so that $\partial \log \psi$ is a meromorphic 1-form on Ω_M^+ with at most logarithmic poles along D. Fix a point $[\eta_0] \in \Omega_M^+ \setminus D$, and set

$$F([\eta]) := \exp\left(\int_{[\eta_0]}^{[\eta]} \partial \log \psi \right), \qquad [\eta] \in \Omega_M^+.$$

Since the residues of $\partial \log \psi$ are integers, we get $F \in \mathcal{O}(\Omega_M^+)$ and

$$d \log F = \partial \log \psi, \qquad \mathrm{div}(F) = D.$$

Let $\gamma \in \Gamma$. By the identity $K_M(\gamma[\eta]) = |j_M(\gamma, [\eta])|^{-2} K_M([\eta])$ and by the Γ-invariance of φ, we get $\psi(\gamma[\eta]) = |j_M(\gamma, [\eta])|^{2p}\, \psi([\eta])$, which yields that

$$\gamma^* \partial \log \psi = \partial \log \psi + p \cdot d \log j_M(\gamma, \cdot).$$

Namely, $d \log(\gamma^* F / j_M(\gamma, \cdot)^p F) = 0$ and hence

$$\chi(\gamma) := F(\gamma[\eta])\, j_M(\gamma, [\eta])^{-p} F([\eta])^{-1}$$

is a non-zero constant on Ω_S^+. Then χ is a character of Γ because for every $\gamma, \gamma' \in \Gamma$,

$$\begin{aligned}\chi(\gamma\gamma') &= \frac{F(\gamma\gamma'[\eta])}{j_M(\gamma\gamma', [\eta])^p F([\eta])} \\ &= \frac{F(\gamma\gamma'[\eta])}{j_M(\gamma, \gamma'[\eta])^p F(\gamma'[\eta])} \times \frac{j_M(\gamma, \gamma'[\eta])^p F(\gamma'[\eta])}{j_M(\gamma\gamma', [\eta])^p F([\eta])} = \chi(\gamma)\chi(\gamma').\end{aligned}$$

Hence F is an automorphic form on Ω_M^+ for Γ of weight p with character χ such that $\mathrm{div}(F) = D$. Since $r(M) \leq 17$, $\|F\|$ is Γ-invariant. By the Poincaré–Lelong formula, the following equation of currents on Ω_M^+ holds:

$$dd^c \log \|F\|^2 = \delta_D - p\,\omega_M. \tag{3.3}$$

By comparing (3.1) and (3.3), $\log(\varphi/\|F\|^2)$ is a Γ-invariant pluriharmonic function on Ω_M^+, so that $\log(\varphi/\|F\|^2)$ descends to a pluriharmonic function on $\mathcal{M}_M$. Since $\overline{\mathcal{M}_M}$, the Baily–Borel–Satake compactification of $\mathcal{M}_M$ is a normal projective variety with $\mathrm{codim}(\overline{\mathcal{M}_M} \setminus \mathcal{M}_M) \geq 2$ when $r(M) \leq 17$, $\log(\varphi/\|F\|^2)$ extends to a pluriharmonic function on $\overline{\mathcal{M}_M}$ by Grauert–Remmert. Since $\overline{\mathcal{M}_M}$ is compact, $\log(\varphi/\|F\|^2)$ is constant by the maximum principle for pluriharmonic functions. This proves the existence of a positive constant C with $\varphi = C\,\|F\|^2$. □

4 Equivariant analytic torsion and 2-elementary $K3$ surfaces

4.1 Equivariant analytic torsion

Let (X, κ) be a compact Kähler manifold. Let G be a finite group acting holomorphically on X and preserving κ. Let $\square_q = (\bar\partial + \bar\partial^*)^2$ be the $\bar\partial$-Laplacian acting on C^∞ $(0, q)$-forms on X. Let $\sigma(\square_q)$ be the spectrum of $\square_q$. For $\lambda \in \sigma(\square_q)$, let $E_q(\lambda)$ be the eigenspace of $\square_q$ with respect to the eigenvalue λ. Since G preserves κ, $E_q(\lambda)$ is a finite-dimensional unitary representation of G. For $g \in G$ and $s \in \mathbf{C}$, set

$$\zeta_q(g)(s) := \sum_{\lambda \in \sigma(\square_q) \setminus \{0\}} \mathrm{Tr}\,(g|_{E_q(\lambda)})\, \lambda^{-s}.$$

Then $\zeta_q(g)(s)$ converges absolutely when $\mathrm{Re}\, s > \dim X$ admits a meromorphic continuation to the complex plane $\mathbf{C}$, and is holomorphic at $s = 0$.

Definition 4.1. *The equivariant analytic torsion of* (X,κ) *is the class function on* G *defined by*

$$\tau_G(X,\kappa)(g) := \exp[-\sum_{q\geq 0}(-1)^q q\,\zeta_q'(g)(0)], \qquad g\in G.$$

When $g=1$, $\tau_G(X,\kappa)(1)$ *is denoted by* $\tau(X,\kappa)$ *and is called the analytic torsion of* (X,κ).

4.2 An invariant for 2-elementary $K3$ surfaces

Let (X,ι) be a 2-elementary $K3$ surface of type M. Let $\mathbf{Z}_2$ be the subgroup of $\mathrm{Aut}(X)$ generated by ι. Let κ be a $\mathbf{Z}_2$-invariant Kähler form on X. Set $\mathrm{vol}(X,\kappa) := (2\pi)^{-2}\int_X \kappa^2/2!$. Let η_X be a nowhere vanishing holomorphic 2-form on X. The L^2-norm of η_X is defined as $\|\eta_X\|^2_{L^2} := (2\pi)^{-2}\int_X \eta_X\wedge\bar{\eta}_X$.

Let $X^\iota = \sum_i C_i$ be the decomposition of the fixed point set of ι into the connected components. Set $\mathrm{vol}(C_i,\kappa|_{C_i}) := (2\pi)^{-1}\int_{C_i}\kappa|_{C_i}$. Let $c_1(C_i,\kappa|_{C_i})$ be the Chern form of $(TC_i,\kappa|_{C_i})$.

Definition 4.2. *Define*

$$\tau_M(X,\iota) : = \mathrm{vol}(X,\kappa)^{\frac{14-r(M)}{4}}\tau_{\mathbf{Z}_2}(X,\kappa)(\iota)\prod_i \mathrm{vol}(C_i,\kappa|_{C_i})\tau(C_i,\kappa|_{C_i})$$
$$\times\exp\left[\frac{1}{8}\int_{C_i}\log\left(\frac{\eta_X\wedge\bar{\eta}_X}{\kappa^2/2!}\cdot\frac{\mathrm{vol}(X,\kappa)}{\|\eta_X\|^2_{L^2}}\right)\Bigg|_{C_i} c_1(C_i,\kappa|_{C_i})\right].$$

Obviously, $\tau_M(X,\iota)$ is independent of the choice of η_X. It is worth remarking that if κ is *Ricci-flat*, then

$$\tau_M(X,\iota) = \mathrm{vol}(X,\kappa)^{\frac{14-r(M)}{4}}\tau_{\mathbf{Z}_2}(X,\kappa)(\iota)\prod_i \mathrm{vol}(C_i,\kappa|_{C_i})\tau(C_i,\kappa|_{C_i}).$$

Theorem 4.3. *Let* $M\subset\mathbb{L}_{K3}$ *be a primitive* 2*-elementary hyperbolic sublattice satisfying* $11\leq r(S)\leq 17$ *and* $r(M)+l(M)=22$. *Then there exists an automorphic form* Φ_M *on* Ω^+_M *for* $O^+(M^\perp)$ *of weight* $(r(S)-6)$ *with zero divisor* $\mathcal{D}_S$ *such that for every* 2*-elementary* $K3$ *surface* (X,ι) *of type* M *and for every* $\mathbf{Z}_2$*-invariant Kähler form* κ *on* X,

$$\tau_M(X,\iota) = \|\Phi_M(\varpi_M(X,\iota))\|^{-\frac{1}{2}}.$$

Proof. Theorem 4.3 follows from the following two claims:

• The number $\tau_M(X,\iota)$ is independent of the choice of a $\mathbf{Z}_2$-invariant Kähler form, and it gives rise to a function τ_M on $\mathcal{M}^o_M$.

• Regarded as a Γ_M-invariant function on Ω^o_M, $\log\tau_M$ lies in $L^1_{\mathrm{loc}}(\Omega_M)$ and satisfies the following equation of currents on Ω_M:

$$dd^c \log \tau_M = \frac{r(M)-6}{4}\,\omega_M - \frac{1}{4}\delta_{\mathcal{D}_M}. \tag{4.1}$$

The first claim follows immediately from the curvature formula for equivariant Quillen metrics [7], [29]. To prove the second claim, it suffices to determine the singularity of τ_M near the divisor $\mathcal{D}_M$. Let $\gamma : \Delta \to \Omega_M$ be a holomorphic curve intersecting $\mathcal{D}^o$ transversally at $t = 0$. Then Eq. (4.1) is deduced from the following estimate:

$$\log \tau_M(\gamma(t)) = -\frac{1}{4}\log|t|^2 + O(\log(-\log|t|)), \qquad t \to 0. \tag{4.2}$$

Under a certain technical assumption on the curve γ, Eq. (4.2) follows from the embedding formula of Bismut [5] for equivariant Quillen metrics. See [39] for more details about the proof. □

In the cases $M = \mathbb{U}(2) \oplus \mathbb{E}_8(2)$ and $M = \mathbb{U} \oplus \mathbb{E}_8(2)$, an explicit formula for Φ_M was given in [39, Sect. 8]; in the first case, Φ_M is given by the Borcherds Φ-function of dimension 10; in the second case, Φ_M is given by the restriction of the Borcherds Φ-function of dimension 26 to Ω_M.

5 The Borcherds products

In this section we recall Borcherds products. For simplicity, we restrict our explanation to those lattices that splits into two hyperbolic lattices.

Let $Mp_2(\mathbf{Z})$ be the metaplectic group (cf. [8], [9]):

$$Mp_2(\mathbf{Z}) := \left\{ \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \sqrt{c\tau + d} \right);\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}),\ \sqrt{c\tau+d} \in \mathcal{O}(\mathbf{H}) \right\},$$

which is generated by the following two elements

$$S := \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \sqrt{\tau} \right), \qquad T := \left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, 1 \right).$$

Let K be an even hyperbolic lattice with signature $(1, b^- - 1)$. Let $N \in \mathbf{N}$. Let $\mathbf{f}, \mathbf{f}'$ be a basis of $\mathbb{U}(N)$ such that

$$\mathbf{f}\cdot\mathbf{f} = \mathbf{f}'\cdot\mathbf{f}' = 0, \qquad \mathbf{f}\cdot\mathbf{f}' = N.$$

Set

$$L := \mathbb{U}(N) \oplus K.$$

The signature of L is $(2, b^-)$. Let $l \in \mathbf{N}$ be the *level* of L; i.e., l is the smallest natural number such that $l\langle\gamma,\gamma\rangle/2 \in \mathbf{Z}$ and $l\langle\gamma,\delta\rangle \in \mathbf{Z}$ for all $\gamma,\delta \in A_L$.

Let $\mathbf{C}[A_L]$ be the group ring of the discriminant group A_L. Let $\{\mathbf{e}_\gamma\}_{\gamma \in A_L}$ be the standard basis of $\mathbf{C}[A_L]$. Let $\rho_L\colon Mp_2(\mathbf{Z}) \to \mathrm{GL}(\mathbf{C}[A_L])$ be the *Weil representation*. Namely, for the generators S and T of $Mp_2(\mathbf{Z})$, we define

$$\rho_L(T)\,\mathbf{e}_\gamma = e^{\pi i \gamma^2}\mathbf{e}_\gamma, \qquad \rho_L(S)\,\mathbf{e}_\gamma = \frac{i^{\frac{b^- - 2}{2}}}{\sqrt{|A_L|}} \sum_{\delta \in A_L} e^{-2\pi i \gamma\cdot\delta}\mathbf{e}_\delta. \tag{5.1}$$

Here the bilinear form on L is denoted by $x \cdot y = \langle x, y\rangle$ for simplicity. Then ρ_L extends to a group homomorphism from $Mp_2(\mathbf{Z})$ to $\mathrm{GL}(\mathbf{C}[A_L])$.

Definition 5.1. *A $\mathbf{C}[A_L]$-valued holomorphic function $F(\tau)$ on the complex upper-half plane $\mathbf{H}$ is a modular form of weight $1-\frac{b^-}{2}$ of type ρ_L if the following conditions are satisfied:*
(1) *For all $\left(\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right), \sqrt{c\tau+d}\right) \in Mp_2(\mathbf{Z})$ and $\tau \in \mathbf{H}$,*

$$F\left(\frac{a\tau+b}{c\tau+d}\right) = \sqrt{c\tau+d}^{\,2-b^-}\rho_L\left(\left(\begin{pmatrix} a & b\\ c & d\end{pmatrix}, \sqrt{c\tau+d}\right)\right)\cdot F(\tau).$$

(2) *$F(\tau)$ is meromorphic at $+i\infty$ and admits the integral Fourier expansion:*

$$F(\tau) = \sum_{\gamma\in A_L} \mathbf{e}_\gamma \sum_{k \in \frac{1}{l}\mathbf{Z}} c_\gamma(k)\, e^{2\pi i k\tau},$$

where $c_\gamma(k) \in \mathbf{Z}$ for all $k \in \frac{1}{l}\mathbf{Z}$ and $c_\gamma(k) = 0$ for $k \ll 0$.

By [8, p.512 Th. 5.3], $F(\tau)$ induces an elliptic modular form $F_K(\tau)$ of the same weight $1-\frac{b^-}{2}$ of type ρ_K.

As before, define

$$\Omega_L := \{[x] \in \mathbf{P}(L_{\mathbf{C}});\ \langle x, x\rangle = 0,\ \langle x, \bar{x}\rangle > 0\}.$$

For $\lambda \in L_{\mathbf{R}}$ with $\langle\lambda,\lambda\rangle < 0$, we define $\mathcal{H}_\lambda$ as before in Section 2. Let

$$C_K = \{v \in K_{\mathbf{R}};\ \langle v, v\rangle > 0\}$$

be the *light cone* of K. Then the tube domain $K_{\mathbf{R}} + i\,C_K$ is identified with Ω_L by the map

$$K_{\mathbf{R}} + i\,C_K \ni z \to \left[\mathbf{f} - \frac{\langle z, z\rangle}{2}\mathbf{f}' + z\right] \in \mathbf{P}(L_{\mathbf{C}}). \tag{5.2}$$

Since K is hyperbolic, C_K consists of two connected components. Let C_K^+ be one of them. Let Ω_L^+ be the component of Ω_L corresponding to $K_{\mathbf{R}} + i\,C_K^+$ via the isomorphism (5.2).

By [8, p.517], $F_K(\tau)$ induces a chamber structure of the cone C_K^+. Each chamber of C_K^+ is called a *Weyl chamber*. Let W be a Weyl chamber of C_K^+. The dual cone of W is defined by

$$W^\vee = \{v \in K_{\mathbf{R}};\ \langle v, w\rangle > 0,\ \forall\, w \in W\}.$$

Theorem 5.2. *(Borcherds) There exists an automorphic form $\Psi_L(z, F)$ on Ω_L^+ with the following properties:*
(1) *$\Psi_L(z, F)$ is an automorphic form of weight $c_0(0)/2$ for a cofinite subgroup of $O^+(L)$.*
(2) *The divisor of $\Psi_L(\cdot, F)$ is given by*

$$\mathrm{div}(\Psi_L(\cdot, F)) = \sum_{\lambda \in L^\vee,\, \lambda^2<0} c_\lambda\left(\frac{\lambda^2}{2}\right)\mathcal{H}_\lambda.$$

(3) *There exists a vector $\rho = \rho(K, F_K(\tau), W) \in K_{\mathbf{Q}}$ determined by K, $F_K(\tau)$ and W such that $\Psi_L(z, F)$ admits the following infinite product expansion for $z \in K_{\mathbf{R}} + iW$ with $\langle z, z\rangle \gg 0$:*

$$\Psi_L(z, F) = e^{2\pi i\rho\cdot z} \prod_{\lambda \in K^\vee \cap W^\vee} \prod_{n \in \mathbf{Z}/N\mathbf{Z}} \left(1 - e^{2\pi i(\lambda\cdot z + \frac{n}{N})}\right)^{c_{\lambda + \frac{n}{N}\mathfrak{f}'}(\lambda^2/2)}.$$

The automorphic form $\Psi_L(z, F)$ is called the Borcherds product associated with L and $F(\tau)$. The vector ρ is called the Weyl vector of $\Psi_L(z, F)$.

Proof. See [8] and [12].

6 Borcherds products for odd unimodular lattices

Define the symmetric unimodular matrix $\mathbb{I}_{1,m}$ of rank $m+1$ and with signature $(1, m)$ by

$$\mathbb{I}_{1,m} := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

We identify $\mathbb{I}_{1,m}$ with the corresponding unimodular hyperbolic lattice. Define 2-elementary lattices $\mathbb{S}_k$, $\mathbb{T}_k$ $(1 \le k \le 9)$ by

$$\mathbb{T}_k := \mathbb{U}(2) \oplus \mathbb{I}_{1,9-k}(2), \qquad \mathbb{S}_k := \mathbb{T}_k^\perp.$$

Then $\mathbb{S}_k$ verifies the conditions in Theorem 4.3:

$$11 \le r(\mathbb{S}_k) = 22 - r(\mathbb{T}_k) \le 17, \qquad r(\mathbb{S}_k) + l(\mathbb{S}_k) = 22 - r(\mathbb{T}_k) + l(\mathbb{T}_k) = 22. \tag{6.1}$$

By Nikulin [33, Th. 4.2.2, P.1434 table 1], $\mathbb{S}_k$ are the only 2-elementary hyperbolic lattices satisfying (6.1), up to an automorphism of $\mathbb{L}_{K3}$. In this section we give an explicit expression of the automorphic form $\Phi_{\mathbb{S}_k}$ in Theorem 4.3 as a Borcherds product.

We define the Weyl vector of $\mathbb{I}_{1,9-k}(2)$ by

$$\rho_k := \frac{1}{2}(3, -1, \dots, -1) \in \mathbb{I}_{1,9-k}(2)^\vee.$$

We set

$$V := S^{-1}T^2S = \left(\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \sqrt{-2\tau + 1} \right) \in Mp_2(\mathbf{Z})$$

and we define $\mathbf{e}_0, \mathbf{e}_1, \mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{C}[A_{\mathbb{T}_k}]$ by

$$\mathbf{e}_0 := \mathbf{e}_{(0,0,0)}, \qquad \mathbf{e}_1 := \mathbf{e}_{(0,0,\rho_k)}, \qquad \mathbf{v}_i := \sum_{\delta \in A_{\mathbb{T}_k},\, 2\langle \delta, \delta \rangle \equiv i \bmod 4} \mathbf{e}_\delta,$$

where vectors in $\mathbb{T}_k$ are denoted by (m, n, λ), $m, n \in \mathbf{Z}$, $\lambda \in \mathbb{I}_{1,9-k}(2)$.

Set $q = e^{2\pi i \tau}$. For $\tau \in \mathbf{H}$, let $\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty}(1 - q^n)$ be the Dedekind η-function and let

$$\theta_2(\tau) = \sum_{m \in \mathbb{Z}} q^{(m+\frac{1}{2})^2/2}, \qquad \theta_3(\tau) = \sum_{m \in \mathbb{Z}} q^{m^2/2}, \qquad \theta_4(\tau) = \sum_{m \in \mathbb{Z}} (-1)^m q^{m^2/2}$$

be Jacobi theta functions. Notice that we use the notation $q = e^{2\pi i \tau}$ while $q = e^{\pi i \tau}$ in [13, Chap. 4]. For $\delta \in \{0, 1/2\}$, let $\theta_{\mathbb{A}_1 + \delta/2}(\tau)$ be the theta function of the A_1-lattice

$$\theta_{\mathbb{A}_1}(\tau) := \theta_3(2\tau), \qquad \theta_{\mathbb{A}_1 + 1/2}(\tau) := \theta_2(2\tau).$$

Define holomorphic functions $f_k^{(0)}(\tau)$, $f_k^{(1)}(\tau)$ and the series $\{c_k^{(0)}(l)\}_{l \in \mathbf{Z}}$, $\{c_k^{(1)}(l)\}_{l \in \mathbf{Z}+1/4}$ by

$$f_k^{(0)}(\tau) := \frac{\eta(2\tau)^8\, \theta_{\mathbb{A}_1}(\tau)^k}{\eta(\tau)^8 \eta(4\tau)^8} = \sum_{l \in \mathbf{Z}} c_k^{(0)}(l)\, q^l = q^{-1} + 8 + 2k + O(q),$$

$$f_k^{(1)}(\tau) := -16 \frac{\eta(4\tau)^8\, \theta_{\mathbb{A}_1 + 1/2}(\tau)^k}{\eta(2\tau)^{16}} = \sum_{l \in 1/4 + \mathbf{Z}} 2c_k^{(1)}(l)\, q^l.$$

We define holomorphic functions $g_k^{(i)}(\tau)$, $i \in \mathbf{Z}/4\mathbf{Z}$ by

$$g_k^{(i)}(\tau) = \sum_{l \equiv i \bmod 4} c_k^{(0)}(l)\, q^{l/4}.$$

By definition,

$$\sum_{i \in \mathbf{Z}/4\mathbf{Z}} g_k^{(i)}(\tau) = \frac{\eta(\tau/2)^8\, \theta_{\mathbb{A}_1}(\tau/4)^k}{\eta(\tau)^8 \eta(\tau/4)^8} = f_k^{(0)}(\tau/4).$$

Define a $\mathbf{C}[A_{\mathbb{T}_k}]$-valued holomorphic function on $\mathbf{H}$ by

$$F_k(\tau) := f_k^{(0)}(\tau)\, \mathbf{e}_0 + f_k^{(1)}(\tau)\, \mathbf{e}_1 + \sum_{i \in \mathbf{Z}/4\mathbf{Z}} g_k^{(i)}(\tau)\, \mathbf{v}_i.$$

Theorem 6.1. *For* $1 \leq k \leq 9$*, the following hold:*
(1) $F_k(\tau)$ *is a modular form for* $Mp_2(\mathbf{Z})$ *of type* $\rho_{\mathbb{T}_k}$ *and of weight* $(k-8)/2$*;*
(2) *the Weyl vector of* $\Psi_{\mathbb{T}_k}(z, F_k)$ *is given by* $2\rho_k$*;*
(3) *there exists a generalized Kac–Moody superalgebra with denominator function* $\Phi_{\mathbb{S}_k}$*;*
(4) *if* $k < 8$*, then there exists a constant* $C_k \neq 0$ *such that*

$$\Phi_{\mathbb{S}_k}(z)^2 = C_k\, \Psi_{\mathbb{T}_k}(z, F_k).$$

The modular form $F_k(\tau)$ for $Mp_2(\mathbf{Z})$ is induced from the modular form $f_k^{(0)}(\tau)$ for $\Gamma_0(4)$. The modular form $f_k^{(0)}(\tau)$ is *reflective* for $\mathbb{T}_k$ in the sense of Borcherds [9, Sect. 11, pp. 350–351].

Remark 6.2. Theorem 6.1 (2) is closely related to an example of Borcherds [8, Example 15.3]. Theorem 6.1 (3), (4) seem to be closely related to a problem of Borcherds [8, Problem 16.2] and conjectures of Harvey–Moore [20, Sect. 7 Conjecture] and Gritsenko–Nikulin [17]. See [40, Sect. 7] for more explanations. The automorphic form $\Phi_{\mathbb{S}_7}$ was already found by Gritsenko–Nikulin [19].

We shall give a detailed proof of Theorem 6.1 in the forthcoming paper [41]. In fact, the norm of $\Phi_{\mathbb{S}_k}$ is regarded as an invariant of certain Calabi–Yau threefolds, which was introduced by Bershadsky–Cecotti–Ooguri–Vafa [4] and by Fang–Lu–Yoshikawa [15] using analytic torsion.

7 *K*3 surfaces of Matsumoto–Sasaki–Yoshida

In Sections 4 and 6, we extended Eqs. (1.6) and (1.7) to 2-elementary $K3$ surfaces of type $\mathbb{S}_k$. In this section we consider an analogue of Eq. (1.4) in dimension 2. We focus on 2-elementary $K3$ surfaces of type $\mathbb{S}_6$. Those $K3$ surfaces were studied in detail by Matsumoto–Sasaki–Yoshida [30], [31], [36].

7.1 The construction of Matsumoto–Sasaki–Yoshida

Recall that

$$M^o(3,6) := \{A = (\mathbf{a}_1, \dots, \mathbf{a}_6) \in M(3,6);\ \mathbf{a}_i \wedge \mathbf{a}_j \wedge \mathbf{a}_k \neq 0 \text{ for } i < j < k\}.$$

For $A \in M^o(3,6)$, we define

$$S_A := \{((x_1 : x_2 : x_3), y) \in \mathcal{O}_{\mathbf{P}^2}(3);\ y^2 = \prod_{i=1}^{6}(a_{1i}\, x_1 + a_{2i}\, x_2 + a_{3i}\, x_3)\}.$$

The natural projection $p = \mathrm{pr}_2 \colon S_A \to \mathbf{P}^2$ is a double covering with branch divisor

$$L_{A,1} \cup \cdots \cup L_{A,6},$$

where

$$L_{A,i} := \{(x_1 : x_2 : x_3) \in \mathbf{P}^2;\ a_{1i}\, x_1 + a_{2i}\, x_2 + a_{3i}\, x_3 = 0\} \cong \mathbf{P}^1.$$

Set $E_{A,ij} := L_{A,i} \cap L_{A,j}$. Corresponding to the 15 points $\{E_{A,ij}\ i \neq j\} \subset \mathbf{P}^2$, S_A has 15 ordinary double points. By [31], [36, Sect. 9.1], the minimal resolution of S_A, i.e., the blow-up of these 15 singular points, is a $K3$ surface. In fact, the following 2-form η_A on X_A is nowhere vanishing:

$$\eta_A := \frac{dx}{y} = \frac{dx}{\prod_{i=1}^{6}(a_{1i}\, x_1 + a_{2i}\, x_2 + a_{3i}\, x_3)^{1/2}}, \tag{7.1}$$

where

$$dx = x_1\, dx_2 \wedge dx_3 - x_2\, dx_1 \wedge dx_3 + x_3\, dx_1 \wedge dx_2.$$

Let $\theta_A\colon S_A \to S_A$ be the involution defined as the non-trivial covering transformation of the double covering $p\colon S_A \to \mathbf{P}^2$.

Definition 7.1. *Let X_A be the minimal resolution of S_A, and let $\iota_A\colon X_A \to X_A$ be the involution on X_A induced by θ_A.*
(1) *The pair (X_A, ι_A) is called a Matsumoto–Sasaki–Yoshida (MSY) $K3$ surface associated with A.*
(2) *Let $L_A := (L_{A,1}, \ldots, L_{A,6})$ be the* ordered *set of lines of* $\mathbf{P}^2$ *associated with A. The triple (X_A, ι_A, L_A) is called a MSY-$K3$ surface with level 2 structure associated with A.*

Two MSY-$K3$ surfaces with level 2 structure (X_A, ι_A, L_A) and (X_B, ι_B, L_B) are isomorphic if there exists an isomorphism $\varphi\colon X_A \to X_B$ such that

$$\varphi \circ \iota_A = \iota_B \circ \varphi, \qquad \varphi(L_A) = L_B.$$

Let $\widetilde{E}_{A,ij} \subset X_A$ be the proper transform of $E_{A,ij}$ by the blow-up of $X_A \to S_A$. Let $H_A \subset \mathbf{P}^2$ be a line which does not pass any points $E_{A,ij}$, and let $\widetilde{H}_A \subset X_A$ be the proper transform of $p^{-1}(H_A)$ by the blow-up of $X_A \to S_A$. Let $\widetilde{L}_{A,i}$ be the proper transform of $p^{-1}(L_{A,i})$ by the blow-up of $X_A \to S_A$. By [31, Prop. 2.1.5], there exists a system of generators E_{ij} $(1 \leq i < j \leq 6)$, H, L_i $(1 \leq i \leq 6)$ of $\mathbb{S}_6$ such that for every MSY-$K3$ surfaces with level 2 structure (X_A, ι_A, L_A), there exists a marking α with

$$\alpha^{-1}(E_{ij}) = c_1([\widetilde{E}_{A,ij}]), \qquad \alpha^{-1}(H) = c_1([\widetilde{H}_A]), \qquad \alpha^{-1}(L_i) = c_1([\widetilde{L}_{A,i}]).$$

Here $[D]$ denotes the line bundle on X_A associated with the divisor D. The triple (X_A, ι_A, L_A) defines a $\mathbb{S}_6$-polarized $K3$ surface in the sense of Dolgachev [14]. A marking of (X_A, ι_A) satisfying these conditions is called a marking of MSY-$K3$ surfaces with level 2 structure (X_A, ι_A, L_A).

Define

$$O(\mathbb{T}_6)(2) := \ker\{O(\mathbb{T}_6) \to O(A_{\mathbb{T}_6})\}.$$

If α, β are markings of (X_A, ι_A, L_A), then

$$\beta \circ \alpha^{-1}|_{\mathbb{S}_6} = \mathrm{id}_{\mathbb{S}_6}, \qquad \beta \circ \alpha^{-1}|_{\mathbb{T}_6} \in O(\mathbb{T}_6)(2).$$

Since $\beta \circ \alpha^{-1}|_{\mathbb{T}_6} \in O(\mathbb{T}_6)(2)$, the $O(\mathbb{T}_6)(2)$-orbit of the period $\pi(X_A, \iota_A, \alpha)$ is independent of the choice of a marking of MSY-$K3$ surface with level 2 structure. The $O(\mathbb{T}_6)(2)$-orbit

$$O(\mathbb{T}_6)(2) \cdot \pi(X_A, \iota_A, \alpha) \in \Omega^o_{\mathbb{S}_6}/O(\mathbb{T}_6)(2)$$

is called the *Griffiths period* of a MSY-$K3$ surface (X_A, ι_A, L_A).

Lemma 7.2. *A MSY-$K3$ surface is a 2-elementary $K3$ surface of type $\mathbb{S}_6$.*

Proof. Let (X_A, ι_A) be a MSY-$K3$ surface. Since X_A/ι_A is the blow-up of $\mathbf{P}^2$ at the 15 points $\{E_{A,ij}\}_{i<j}$ and is a rational surface, ι_A acts non-trivially on $H^0(X_A, K_{X_A})$. The type of (X_A, ι_A) is $\mathbb{S}_6$ by [31, Prop. 2.1.5]. □

We have a family of $K3$ surfaces with involution $\pi : (\mathcal{X}, \iota) \to M^o(3,6)$ such that $\pi^{-1}(A) = (X_A, \iota_A)$. On $M^o(3,6)$ the group $GL_3(\mathbf{C}) \times (\mathbf{C}^*)^6$, acts by

$$(g, \lambda_1, \ldots, \lambda_6) \cdot A := g \cdot A \cdot \mathrm{diag}(\lambda_1, \ldots, \lambda_6).$$

Definition 7.3. *Define the configuration space of six lines in general position on $\mathbf{P}^2$ by*

$$X^o(3,6) := GL_3(\mathbf{C})\backslash M^o(3,6)/(\mathbf{C}^*)^6.$$

The configuration space $X^o(3,6)$ is a Zariski open subset of $\mathbf{C}^4$. In fact, every element of $X^o(3,6)$ has a unique representative of the form (cf. [36, Chap. 7 Sect. 2]):

$$\begin{pmatrix} a_1 & a_2 & 1 & 1 & 0 & 0 \\ a_3 & a_4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \qquad a_1, \ldots, a_4 \in \mathbf{C}.$$

Hence there exists an embedding $j\colon X^o(3,6) \hookrightarrow M^o(3,6)$ with

$$j(X^o(3,6)) = \left\{ \begin{pmatrix} a_1 & a_2 & 1 & 1 & 0 & 0 \\ a_3 & a_4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \in M^o(3,6);\ a_1, a_2, a_3, a_4 \in \mathbf{C} \right\}. \tag{7.2}$$

By the expression (7.2), there exist 15 hyperplanes $H_1, \ldots, H_{15} \subset \mathbf{C}^4$ and a hyperquadric $Q \in \mathbf{C}^4$ such that $X^o(3,6) = \mathbf{C}^4 \setminus H_1 \cup \cdots \cup H_{15} \cup Q$.

The permutation group on 6 letters $\mathfrak{S}_6$ acts on $M(3,6;\mathbf{C})$ by

$$\sigma \cdot (\mathbf{a}_1, \ldots, \mathbf{a}_6) := (\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(6)}), \qquad (\mathbf{a}_1, \ldots, \mathbf{a}_6) \in M(3,6;\mathbf{C}), \quad \sigma \in \mathfrak{S}_6.$$

Following [36, Chap. 7 Sect. 3], we define an automorphism of $M^o(3,6)$ by

$$T(U, V) := (\det(U)^t U^{-1}, \det(V)^t V^{-1}), \qquad U, V \in GL_3(\mathbf{C}).$$

Notice that the (i, j)-entry of $\det(U)^t U^{-1}$ is the (i, j)-minor of U for $U \in GL_3(\mathbf{C})$. For all $A \in M^o(3,6)$, $g \in GL_3(\mathbf{C})$, $\lambda_1, \dots, \lambda_6 \in \mathbf{C}^*$, one has

$$\sigma(gA) = g\sigma(A), \qquad \sigma(A \cdot \operatorname{diag}(\lambda_1, \cdots, \lambda_6)) = \sigma(A) \cdot \operatorname{diag}(\lambda_{\sigma(1)}, \cdots, \lambda_{\sigma(6)})$$

$$T(gA) = {}^t g^{-1} T(A), \quad T(A \cdot \operatorname{diag}(\lambda_1, \cdots, \lambda_6)) = T(A) \cdot \operatorname{diag}(\mu_{\sigma(1)}, \cdots, \mu_{\sigma(6)})$$

where $\mu_i = \lambda_1\lambda_2\lambda_3/\lambda_i$ for $i = 1,2,3$ and $\mu_j = \lambda_4\lambda_5\lambda_6/\lambda_j$ for $j = 4,5,6$. Hence the actions of $\mathfrak{S}_6$ and T on $M^o(3,6)$ descend to those on $X^o(3,6)$.

Let $\langle T \rangle \cong \mathbf{Z}_2$ be the subgroup of $\operatorname{Aut}(X^o(3,6))$ generated by T. Let G be the finite automorphism group of $X^o(3,6)$ generated by $\mathfrak{S}_6$ and T. Since $\mathfrak{S}_6$ commutes with $\langle T \rangle$ by [36, Chap. 7 Prop. 3.3], one has $G \cong \mathfrak{S}_6 \times \mathbf{Z}_2$.

Set

$$\mathcal{M}^o_{\mathbb{S}_6}(2) := \mathbf{\Omega}^o_{\mathbb{S}_6}/O(\mathbb{T}_6)(2).$$

Theorem 7.4 (Matsumoto–Sasaki–Yoshida). *The period map for the family of MSY-$K3$ surfaces with level 2 structure $\pi : (\mathcal{X}, \iota) \to M^o(3,6)$ with fiber $\pi^{-1}(A) = (X_A, \iota_A, L_A)$, induces an isomorphisms of analytic spaces*

$$X^o(3,6) \cong \mathcal{M}^o_{\mathbb{S}_6}(2), \qquad X^o(3,6)/G \cong \mathcal{M}^o_{\mathbb{S}_6}.$$

In particular, $X^o(3,6)/G$ (resp. $X^o(3,6)$) is a coarse moduli space of MSY-$K3$ *surfaces (resp. with level 2 structure).*

Proof. By [31, Prop. 2.10.1] [36, Sect. 9.5], the period map for the family $\pi : (\mathcal{X}, \iota) \to M^o(3,6)$ induces an isomorphism of analytic spaces $\varphi : X^o(3,6) \cong \mathcal{M}^o_{\mathbb{S}_6}(2)$ such that the following diagram is commutative:

$$\begin{array}{ccc} M^o(3,6) & \xrightarrow{\text{id}} & M^o(3,6) \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle \phi} \\ X^o(3,6) & \xrightarrow{\varphi} & \mathcal{M}^o_{\mathbb{S}_6}(2), \end{array} \qquad (7.3)$$

where $q : M^o(3,6) \to X^o(3,6)$ is the natural projection and $\phi : M^o(3,6) \to \mathcal{M}^o_{\mathbb{S}_6}(2)$ is the period map for the family $\pi : (\mathcal{X}, \iota) \to M^o(3,6)$. This proves the first isomorphism. Since the isomorphism φ induces an isomorphism of groups $G \cong O(\mathbb{T}_6)/O(\mathbb{T}_6)(2)$ by [36, Prop. 9.4], we have $X^o(3,6)/G \cong \mathbf{\Omega}^o_{\mathbb{S}_6}/O(\mathbb{T}_6) = \mathcal{M}^o_{\mathbb{S}_6}$. This proves the second assertion. See [30], [31, Prop. 2.10.1, p.22, l.7], [36] for more details. □

7.2 The Freitag theta functions

Let $M(2,\mathbf{C})$ denote the vector space of 2×2 complex matrices. Then $\mathbf{\Omega}_{\mathbb{S}_6}$ is biholomorphic to a tube domain $\mathbf{H}_2 \subset M(2,\mathbf{C})$ defined by

$$\mathbf{H}_2 := \Lambda + \mathrm{i}\, C_\Lambda = \left\{ W \in M(2,\mathbf{C});\ \frac{W - W^*}{2\mathrm{i}} > 0 \right\}, \qquad W^* := {}^t\overline{W}.$$

The isomorphism between $\Omega_{\mathbb{S}_6}$ and $\mathbf{H}_2$ is given as follows. Let Λ be the real vector space of 2×2 Hermitian matrices:

$$\Lambda := \left\{ \begin{pmatrix} u & w \\ \bar{w} & v \end{pmatrix} \in M(2, \mathbf{C});\ u, v \in \mathbf{R},\ w \in \mathbf{C} \right\}.$$

Let $C_\Lambda := \{H \in \Lambda :\ H > 0\}$ be the light cone of Λ, where $H > 0$ if and only if H is positive-definite. Let $\{\mathbf{h}, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ be the basis of Λ defined as

$$\mathbf{h} = \begin{pmatrix} 1 & \frac{-1}{1+\mathrm{i}} \\ \frac{-1}{1-\mathrm{i}} & 1 \end{pmatrix}, \quad \mathbf{d}_1 = \begin{pmatrix} 0 & \frac{-1}{1+\mathrm{i}} \\ \frac{-1}{1-\mathrm{i}} & 1 \end{pmatrix}, \quad \mathbf{d}_2 = \begin{pmatrix} 1 & \frac{-1}{1+\mathrm{i}} \\ \frac{-1}{1-\mathrm{i}} & 0 \end{pmatrix}, \quad \mathbf{d}_3 = \begin{pmatrix} 0 & \frac{\mathrm{i}}{1+\mathrm{i}} \\ \frac{-\mathrm{i}}{1-\mathrm{i}} & 0 \end{pmatrix}$$

We consider the following coordinates $\mathbf{y} = (y_0, y_1, y_2, y_3)$ on $\mathbf{H}_2 = \Lambda + \mathrm{i}\, C_\Lambda$:

$$\begin{aligned} \mathbf{y} &= y_0\, \mathbf{h} + y_1\, \mathbf{d}_1 + y_2\, \mathbf{d}_2 + y_3\, \mathbf{d}_3 \\ &= \begin{pmatrix} y_0 + y_2 & (-y_0 - y_1 - y_2 + \mathrm{i} y_3)/(1+\mathrm{i}) \\ (-y_0 - y_1 - y_2 - i y_3)/(1-\mathrm{i}) & y_0 + y_1 \end{pmatrix} \in \mathbf{H}_2. \end{aligned}$$

The period domain $\Omega_{\mathbb{S}_6}$ is isomorphic to the tube domain $\mathbf{H}_2$ by the map:

$$\mu \colon \mathbf{H}_2 \ni \mathbf{y} \to (1 : -\det(\mathbf{y}) : y_0 : y_1 : y_2 : y_3) \in \Omega_{\mathbb{S}_6}. \tag{7.4}$$

Definition 7.5. (1) *For $a, b \in \{0, \frac{1+\mathrm{i}}{2}\}^2$ and $W \in \mathbf{H}_2$, define*

$$\Theta\begin{pmatrix} a \\ b \end{pmatrix}(W) = \sum_{m \in \mathbf{Z}[\mathrm{i}]^2} \exp \pi \mathrm{i} \left\{ \left(m + \frac{a}{1+\mathrm{i}}\right)^* W \left(m + \frac{a}{1+\mathrm{i}}\right) + 2\mathrm{Re} \left(\frac{b}{1+\mathrm{i}}\right)^* m \right\}.$$

The Freitag theta function $\Theta\binom{a}{b}(W)$ is said to be even if $a^* b \in \mathbf{Z}$.
(2) *Define the Matsumoto–Sasaki–Yoshida form Δ_{MSY} by*

$$\Delta_{\mathrm{MSY}}(W) := \prod_{\binom{a}{b}\ \text{even}} \Theta\begin{pmatrix} a \\ b \end{pmatrix}(W).$$

Let $\mathcal{P}$ be the set of all partitions $\binom{ijk}{lmn}$ of the set $\{1, \ldots, 6\}$, where

$$\begin{pmatrix} ijk \\ lmn \end{pmatrix} := \{i, j, k\} \cup \{l, m, n\} = \{1, \ldots, 6\}, \qquad i < j < k, \quad l < m < n.$$

There exists a one to one correspondence between $\mathcal{P}$ and the set of even Freitag theta functions. Since $\#\mathcal{P} = 10$, there exists ten even Freitag theta functions. The Freitag theta function corresponding to the partition $\binom{ijk}{lmn}$ is denoted by $\Theta\binom{ijk}{lmn}(W)$. Hence

$$\Delta_{\mathrm{MSY}}(W) = \prod_{\binom{ijk}{lmn} \in \mathcal{P}} \Theta\begin{pmatrix} ijk \\ lmn \end{pmatrix}(W).$$

See [30, Sect. 2.3], [36, Sect. 9.12.5] for the explicit correspondence between the even characteristics $\{\binom{a}{b}\}$ and the partitions $\{\binom{ijk}{lmn}\}$.

Proposition 7.6. *Under the identification* $\mu\colon \mathbf{H}_2 \cong \Omega^+_{\mathbb{S}_6}$, *the Matsumoto–Sasaki–Yoshida form* $\Delta_{\mathrm{MSY}}(W)$ *is an automorphic form on* $\mathbf{H}_2$ *for* $O(\mathbb{T}_6)^+$ *of weight* 10 *with*

$$\mathrm{div}(\Delta_{\mathrm{MSY}}) = \mathcal{D}_{\mathbb{S}_6} = \sum_{\delta\in\Delta_{\mathbb{T}_6}} \mathcal{H}_\delta.$$

Proof. See [30, Lemma 2.3.1 and Prop. 3.1.1]. □

7.3 The discriminant of MSY $K3$ surfaces

We introduce an analogue of the function $\Delta_{(2,4)}$ in the case $M^o(3,6)$.

Definition 7.7. (1) *For* $A = (\mathbf{a}_1, \dots, \mathbf{a}_6) \in M^o(3,6)$ *and a partition* $\binom{ijk}{lmn} \in \mathcal{P}$, *define*

$$D\binom{ijk}{lmn}(A) := \det(\mathbf{a}_i, \mathbf{a}_j, \mathbf{a}_k)\, \det(\mathbf{a}_l, \mathbf{a}_m, \mathbf{a}_n).$$

(2) *Define a holomorphic function* $\Delta_{(3,6)}$ *on* $M^o(3,6)$ *by*

$$\Delta_{(3,6)}(A) := \prod_{\binom{ijk}{lmn}\in\mathcal{P}} D\binom{ijk}{lmn}(A) = \prod_{\binom{ijk}{lmn}\in\mathcal{P}} \det(\mathbf{a}_i, \mathbf{a}_j, \mathbf{a}_k)\cdot\det(\mathbf{a}_l, \mathbf{a}_m, \mathbf{a}_n).$$

(3) *Define a real-valued function* $\|\Delta_{(3,6)}\|$ *on* $M^o(3,6)$ *by*

$$\|\Delta_{(3,6)}(A)\| := \left(\frac{1}{(2\pi)^2}\int_{X_A} \eta_A \wedge \overline{\eta_A}\right)^{10} |\Delta_{(3,6)}(A)|.$$

Lemma 7.8. (1) $M^o(3,6) = M(3,6;\mathbf{C}) \setminus \mathrm{div}(\Delta_{(3,6)})$.
(2) $\|\Delta_{(3,6)}\|$ *is* $GL_3(\mathbf{C}) \times (\mathbf{C}^*)^6$*-invariant.*

Proof. **(1)** The first assertion follows from the definition of $M^o(3,6)$.
(2) Let $A = (a_{ij}) \in M^o(3,6)$ and $g \in GL_3(\mathbf{C})$. We write $gA = (a^{(g)}_{ij})$. We identify g with the corresponding projective transformation. Then the projective transformation $\mathbf{P}^2 \ni [x] \to [{}^t g^{-1} x] \in \mathbf{P}^2$ lifts to an isomorphism $f_g\colon X_A \to X_{gA}$ such that

$$f_g^*(\eta_{gA}) = f_g^*\left(\frac{dx}{\prod_{i=1}^6 (a^{(g)}_{1i}\, x_1 + a^{(g)}_{2i}\, x_2 + a^{(g)}_{3i}\, x_3)^{1/2}}\right)$$

$$= \frac{d({}^t g^{-1} x)}{\prod_{i=1}^6 (a_{1i}\, x_1 + a_{2i}\, x_2 + a_{3i}\, x_3)^{1/2}} = \det(g)^{-1}\, \eta_A.$$

This, together with $\Delta_{(3,6)}(gA) = \det(g)^{20}\, \Delta_{(3,6)}(A)$, implies the $GL_3(\mathbf{C})$-invariance of $\|\Delta_{(3,6)}\|$. Let us see the $(\mathbf{C}^*)^6$-invariance of $\|\Delta_{(3,6)}\|$. Identify $\lambda = (\lambda_i)_{i=1}^6 \in (\mathbf{C}^*)^6$ with the invertible diagonal matrix $\lambda = (\delta_{ij}\lambda_i)_{1\le i,j\le 6} \in GL_6(\mathbf{C})$. Since $\eta_{A\lambda} = (\det\lambda)^{-1/2}\eta_A$ and $\Delta_{(3,6)}(A\lambda) = \det(\lambda)^{10}\, \Delta_{(3,6)}(A)$, we get the $(\mathbf{C}^*)^6$-invariance of $\|\Delta_{(3,6)}\|$. □

By Lemma 7.8, $\|\Delta_{(3,6)}\|$ descends to a function on $X^o(3,6)$. We identify $\|\Delta_{(3,6)}\|$ with the corresponding function on $X^o(3,6)$.

Theorem 7.9. (1) *There exist non-zero constants C_1, C_2 such that the following identity holds under the identification* (7.4)*:*

$$\Phi_{\mathbb{S}_6} = C_1\, \Delta_{\mathrm{MSY}} = C_2\, \Psi_{\mathbb{T}_6}(\cdot, F_6)^{1/2}.$$

(2) *There exists an absolute constant $C_3 \neq 0$ such that for all $A \in M^o(3,6)$,*

$$\tau_{\mathbb{S}_6}(X_A, \iota_A) = C_3\, \|\Delta_{(3,6)}(A)\|^{-1/4}.$$

By Theorems 6.1 and 7.9 (1), we get an infinite product expansion of the Igusa cusp form, i.e., the restriction of Δ_{MSY} to the Siegel upper-half space $\mathfrak{S}_2 = \{W \in \mathbf{H}_2;\ {}^tW = W\}$. The infinite product expansion of the Igusa cusp form was first obtained by Gritsenko–Nikulin [18].

For the proof of Theorem 7.9, we recall the results of Matsumoto–Saasaki–Yoshida in more details.

7.4 A compactification of $X^o(3,6)$

For $1 \leq i < j < k \leq 6$, we define

$$M_{ijk}(3,6) := \left\{ A \in M(3,6;\mathbf{C});\ \begin{matrix} \mathbf{a}_i \wedge \mathbf{a}_j \wedge \mathbf{a}_k = \mathbf{0} & \text{if } (k,l,m) = (i,j,k) \\ \mathbf{a}_l \wedge \mathbf{a}_m \wedge \mathbf{a}_n \neq \mathbf{0} & \text{if } (k,l,m) \neq (i,j,k) \end{matrix} \right\},$$

$$X_{ijk}(3,6) := GL_3(\mathbf{C})\backslash M_{ijk}(3,6)/(\mathbf{C}^*)^6,$$

and we set

$$M^*(3,6) := M^o(3,6) \cup \amalg_{i<j<k} M_{ijk}(3,6)$$
$$X^*(3,6) := X^o(3,6) \cup \amalg_{i<j<k} X_{ijk}(3,6).$$

Notice that if $i < j < k, l < m < n$ and $(i,j,k) \neq (l,m,n)$, then

$$M_{ijk}(3,6) \cap M_{lmn}(3,6) = \emptyset, \qquad X_{ijk}(3,6) \cap X_{lmn}(3,6) = \emptyset.$$

The subset $M^*(3,6)$ is open in $M(3,6;\mathbf{C})$.

For $A \in \amalg_{i<j<k} M_{ijk}(3,6)$, we define S_A and $L_{A,i}$, $i = 1,\ldots,6$ as in Section 7.1. Then Sing S_A consists of only rational double points, i.e., 12 ordinary double points and one A_3-singularity. For $A \in M^*(3,6)$, we define η_A as in (7.1). Since η_A is nowhere vanishing on the regular part of S_A, the minimal resolution of S_A, denoted again by X_A, is a $K3$ surface. We have a flat family of surfaces $\pi : \mathcal{S} \to M^*(3,6)$ with fiber $\pi^{-1}(A) = S_A$.

With respect to the trivial $GL_3(\mathbf{C}) \times (\mathbf{C}^*)^6$-action on $\mathbf{P}^{29}$, there exists by [31], [36] a $GL_3(\mathbf{C}) \times (\mathbf{C}^*)^6$-equivariant holomorphic map $F\colon M^*(3,6) \to \mathbf{P}^{29}$ that induces an injection $f\colon X^*(3,6) \hookrightarrow \mathbf{P}^{29}$. We consider the topology on $X^*(3,6)$

induced from the one on $f(X^*(3,6))$ via f; we identify $X^*(3,6)$ with $f(X^*(3,6))$ as a topological space. Let $\overline{X}(3,6)$ be the closure of $f(X^*(3,6))$ in $\mathbf{P}^{29}$ and let $\overline{X}_{ijk}(3,6)$ be the closure of $f(X_{ijk}(3,6))$ in $\mathbf{P}^{29}$.

Set $\mathcal{M}_{\mathbb{S}_6}(2) := \Omega^+_{\mathbb{S}_6}/O^+(\mathbb{T}_6)(2)$. Since $O^+(\mathbb{T}_6)(2)$ is generated by reflections by [31, Prop. 2.5.2], $\mathcal{M}_{\mathbb{S}_6}(2)$ is smooth.

Theorem 7.10. (1) $\overline{X}(3,6)$ *is a projective variety of dimension* 4. *The isomorphism* φ *in* (7.3) *extends to an isomorphism* $\overline{\varphi}$ *between* $\overline{X}(3,6)$ *and the Baily–Borel–Satake compactification of* $\mathcal{M}_{\mathbb{S}_6}(2)$.
(2) $X^*(3,6) \subset \overline{X}(3,6)_{\mathrm{reg}} := \overline{X}(3,6) \setminus \mathrm{Sing}\,\overline{X}(3,6)$.
(3) $X^*(3,6)$ *is a Zariski open subset of* $\overline{X}(3,6)$ *with* $\dim \overline{X}(3,6) \setminus X^*(3,6) \le 2$.
(4) $\overline{X}_{ijk}(3,6) \cap X^*(3,6)$ *is a smooth hypersurface of* $X^*(3,6)$.

Proof. See [31, Th. A6.2] for the first part of (1) and [30, Th. 3.2.4, Cor. 4.4.2] for the second part of (1). Since X_A is a $K3$ surface with at most rational double points for $A \in \amalg_{i<j<k} M_{ijk}(3,6)$, $\overline{X}_{ijk}(3,6)$ is identified with a divisor of $\mathcal{M}_{\mathbb{S}_6}(2)$ via $\overline{\varphi}$. Hence $X^*(3,6)$ is regarded as a subset of $\mathcal{M}_{\mathbb{S}_6}(2)$ via $\overline{\varphi}$. Since $\mathcal{M}_{\mathbb{S}_6}(2)$ is smooth, $X^*(3,6)$ consists of smooth points of $\overline{X}(3,6)$. This proves (2). See also [36, p.244] for the proof of (2). See [31, Prop. A5.3, Cor. A5.4, Th. A6.2] for the proof of (3). Consider the following subset of $M^*(3,6)$:

$$\mathcal{U} := \left\{ \begin{pmatrix} 1 & 0 & a & 0 & 1 & c \\ 0 & 1 & b & 0 & 1 & d \\ 0 & 0 & z & 1 & 1 & 1 \end{pmatrix} \in M^*(3,6);\ a,b,c,d,z \in \mathbf{C} \right\}.$$

Let $U_{123} \subset X^*(3,6)$ be the image of $\mathcal{U}$ by the natural projection $M^*(3,6) \to X^*(3,6)$. By [31, Lemmas A6.8 and A6.9 and their proofs], U_{123} is an open subset of $X^*(3,6)$ containing $X^o(3,6) \cup X_{123}(3,6)$. Since U_{123} is isomorphic to an open subset of $\mathbf{P}^2 \times \mathbf{C}^2$ and since $X_{123}(3,6) \cap U_{123}$ is defined by the equation $z = 0$, $X_{123}(3,6)$ is a smooth hypersurface of $X^*(3,6)$. This proves (4). By [36, p.244], $X_{ijk}(3,6)$ is identified with a certain smooth hypersurface of $\mathcal{M}_{\mathbb{S}_6}(2)$, which also proves (4). □

See [36, Chap. 7 Sect. 5] for the interpretation of the boundary locus $\overline{X}(3,6) \setminus X^o(3,6)$ in terms of degenerate matrices in $M(3,6;\mathbf{C})$.

Define a function K on $M^*(3,6)$ by

$$K(A) := \int_{X_A} \eta_A \wedge \bar{\eta}_A, \qquad A \in M^*(3,6).$$

Lemma 7.11. *K is a nowhere vanishing continuous function on $M^*(3,6)$.*

Proof. Let $\mathfrak{U} \cong \Delta^{18}$ be a small neighborhood of A in $M^*(3,6)$ such that $\mathfrak{U} \cap \amalg_{i<j<k} M_{ijk}(3,6) \cong \Delta^{17}$. By [25, Th. 4.28], there exists a finite holomorphic map $h\colon \mathfrak{V} \to \mathfrak{U}$ with branch divisor $\mathfrak{U} \cap \amalg_{i<j<k} M_{ijk}(3,6)$ such that the family $\mathrm{pr}_2\colon \mathcal{S} \times_{\mathfrak{U}} \mathfrak{V} \to \mathfrak{V}$ induced from $\pi : \mathcal{S} \to M^*(3,6)$ by $h\colon \mathfrak{V} \to \mathfrak{U}$ admits a simultaneous resolution. Namely, there exist a complex manifold $\mathcal{X}$, holomorphic maps $p\colon \mathcal{X} \to \mathcal{S} \times_{\mathfrak{U}} \mathfrak{V}$ and $\widetilde{\pi} : \mathcal{X} \to \mathfrak{V}$ such that the diagram

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{p} & \mathcal{S} \times_{\mathfrak{U}} \mathfrak{V} \\ {\scriptstyle \tilde{\pi}}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_2} \\ \mathfrak{V} & \xrightarrow{\mathrm{id}} & \mathfrak{V} \end{array}$$

is commutative and such that $p\colon X_B := \tilde{\pi}^{-1}(B) \to S_B$ is the minimal resolution for all $B \in \mathfrak{V}$. Hence $\{p^*\eta_B\}_{B\in\mathfrak{U}}$ is a nowhere vanishing relative holomorphic 2-form on $\mathcal{X}$. Since every fiber of $\tilde{\pi}$ is smooth and since $h^*K(B) = \int_{X_B} p^*\eta_B \wedge p^*\bar{\eta}_B$ for all $B \in \mathfrak{V}$, h^*K is a continuous function on $\mathfrak{V}$. Since $p^*\eta_B \neq 0$ for all $B \in \mathfrak{V}$, h^*K is nowhere vanishing on $\mathfrak{B}$. This proves the assertion on $\mathfrak{U}$. Since $A \in \amalg_{i<j<k} M_{ijk}(3,6)$ is an arbitrary point, K is a nowhere vanishing continuous function on $M^*(3,6)$. □

7.5 An intermediate modular variety

Let $z = (z_{\binom{ijk}{lmn}})_{\binom{ijk}{lmn}\in\mathcal{P}}$ be the homogeneous coordinates of $\mathbf{P}^9$ and define

$$Z := \{z \in \mathbf{P}^9;\ \mathrm{Plk}_{ij}(z) = 0 \text{ for all } i < j\},$$

where $\mathrm{Plk}_{ij}(z) := z_{\binom{ijk}{lmn}} - z_{\binom{ijl}{mnk}} + z_{\binom{ijm}{nkl}} - z_{\binom{ijn}{klm}}$ are the Plucker relations. Then $Z \subset \mathbf{P}^9$ is a linear subspace of dimension 4.

After Matsumoto, we define

$$\mathrm{Pr}\colon X(3,6) \ni [A] \to (\cdots : D\binom{ijk}{lmn}(A) : \cdots)_{\binom{ijk}{lmn}\in\mathcal{P}} \in \mathbf{P}^9$$

and

$$\Theta\colon \mathbf{H}_2 \ni W \to (\cdots : \Theta\binom{ijk}{lmn}^2(W) : \cdots)_{\binom{ijk}{lmn}\in\mathcal{P}} \in \mathbf{P}^9.$$

Recall that the period map $\phi\colon M^o(3,6) \to \mathcal{M}^o_{\mathbb{S}_6}(2)$ induces the isomorphism $\varphi\colon X^o(3,6) \to \mathcal{M}^o_{\mathbb{S}_6}(2)$ in (7.3). Let $\Gamma_M(1+\mathrm{i}) \subset \mathrm{Aut}(\mathbf{H}_2)$ be the subgroup corresponding to $O^+(\mathbb{T}_6)(2) \subset \mathrm{Aut}(\Omega_{\mathbb{S}_6})$ via the isomorphism $\mu\colon \mathbf{H}_2 \cong \Omega_{\mathbb{S}_6}$. Let $\overline{\mathbf{H}_2/\Gamma_M(1+\mathrm{i})}$ be the Baily–Borel–Satake compactification of $\mathbf{H}_2/\Gamma_M(1+\mathrm{i})$.

Theorem 7.12. (1) *The images of* Pr *and* Θ *are contained in* Z*;*
(2) Pr *extends to a double covering* $\overline{\mathrm{Pr}}\colon \overline{X}(3,6) \to Z$*;*
(3) Θ *induces a double covering* $\overline{\Theta}\colon \overline{\mathbf{H}_2/\Gamma_M(1+\mathrm{i})} \to Z$*;*
(4) *The period map for the family* $\pi : (\mathcal{X}, \iota) \to M^o(3,6)$ *induces an isomorphism* $\psi : \overline{X}(3,6) \to \overline{\mathbf{H}_2/\Gamma_M(1+\mathrm{i})}$ *such that the following diagram is commutative:*

$$\begin{array}{ccc} \overline{X}(3,6) & \xrightarrow{\psi} & \overline{\mathbf{H}_2/\Gamma_M(1+\mathrm{i})} \\ {\scriptstyle \overline{\mathrm{Pr}}}\downarrow & & \downarrow{\scriptstyle \overline{\Theta}} \\ Z & \xrightarrow{\mathrm{id}} & Z \end{array} \tag{7.5}$$

Proof. See [30, Th. 4.4.1, Cor. 4.4.2] and [36, Chap. 7 Prop. 6.2, Chap. 9 Th. 12.7]. □

We regard the monomial $\prod_{\mathcal{P}} z_{\binom{ijk}{lmn}}$ as an element of $H^0(\mathbf{P}^9, \mathcal{O}_{\mathbf{P}^9}(10))$. Let $\|\cdot\|_{\mathcal{O}_{\mathbf{P}^9}(10)}$ be the standard hermitian metric on $\mathcal{O}_{\mathbf{P}^9}(10)$ whose Chern form is proportional to the Fubini–Study form on $\mathbf{P}^9$. Then

$$\frac{\|\Delta_{(3,6)}(A)\|^2}{\overline{\mathrm{Pr}}^* \|\prod_{\mathcal{P}} z_{\binom{ijk}{lmn}}\|^2_{\mathcal{O}_{\mathbf{P}^9}(10)}(A)} = \frac{K(A)^{20}}{\left(\sum_{\mathcal{P}} |D\binom{ijk}{lmn}(A)|^2\right)^{10}}, \qquad A \in M^*(3,6).$$

Since $K(A)^{20}/\left(\sum_{\mathcal{P}} |D\binom{ijk}{lmn}(A)|^2\right)^{10}$ descends to a nowhere vanishing continuous function on $X^*(3,6)$ by Lemmas 7.8 and 7.11, there exists a continuous hermitian metric $\|\cdot\|'$ on $\overline{\mathrm{Pr}}^* \mathcal{O}_{\mathbf{P}^9}(10)$ such that

$$\|\overline{\mathrm{Pr}}^* \prod_{\mathcal{P}} z_{\binom{ijk}{lmn}}\|' = \|\Delta_{(3,6)}\|.$$

Lemma 7.13. *Let* $\gamma : \Delta \to X^*(3,6)$ *be a holomorphic curve that intersects* $\amalg_{i<j<k} X_{ijk}(3,6)$ *transversally at* $\gamma(0)$. *Then as* $t \to 0$,

$$\log \|\Delta_{(3,6)}(\gamma(t))\|^2 = \log |t|^2 + O(1).$$

Proof. Let $\gamma(0) \in X_{ijk}(3,6)$. Let f be a local holomorphic function defining the divisor $X_{ijk}(3,6)$ near $\gamma(0)$. Since $\|\cdot\|'$ is a continuous metric on $\overline{\mathrm{Pr}}^* \mathcal{O}_{\mathbf{P}^9}(10)$ and since $\prod_{\mathcal{P}} z_{\binom{ijk}{lmn}} \in H^0(\overline{X}(3,6), \overline{\mathrm{Pr}}^* \mathcal{O}_{\mathbf{P}^9}(10))$, we get

$$\begin{aligned}\log \|\Delta_{(3,6)}(\gamma(t))\|^2 &= \log \left(\left\| \overline{\mathrm{Pr}}^* \prod_{\mathcal{P}} z_{\binom{ijk}{lmn}}(\gamma(t)) \right\|' \right)^2 \\ &= (\mathrm{mult}_{t=0} \gamma^* f) \log |t|^2 + O(1).\end{aligned}$$

Since γ intersects $X_{ijk}(3,6)$ transversally at $\gamma(0)$, we get $\mathrm{mult}_{t=0}\gamma^* f = 1$. □

Proof of Theorem 7.9. We keep the notation in (7.3) and (7.5).

(1) The identity $\Phi_{\mathbb{S}_6} = C_2 \Psi_{\mathbb{T}_6}(\cdot, F_6)$ follows from Theorem 6.1. We compare the weights and the zeros of $\Phi_{\mathbb{S}_6}$ and Δ_{MSY}. By Theorem 4.3 and Proposition 7.6, both of $\Phi_{\mathbb{S}_6}$ and Δ_{MSY} have the same weight 10 and the same zero divisor $\mathcal{D}_{\mathbb{S}_5}$. From the Köcher principle, the assertion follows.

(2) By Theorems 4.3 and 7.9 (1), it suffices to prove that

$$\|\Delta_{(3,6)}\|^2 = \mathrm{Const.}\, \varphi^* \|\Delta_{\mathrm{MSY}}\|^2. \tag{7.6}$$

Let $\Pi \colon \Omega_{\mathbb{S}_6} \to \Omega_{\mathbb{S}_6}/O^+(\mathbb{T}_6)(2)$ be the natural projection, and set

$$f := \Pi^* (\varphi^{-1})^* \log \|\Delta_{(3,6)}\|^2.$$

We compute the $(1,1)$-current $dd^c f$ on $\Omega_{\mathbb{S}_6}$. By Definition 7.7 (3) and the definition of the Bergman metric, we get on $M^o(3,6)$

$$q^* dd^c \log \|\Delta_{(3,6)}\|^2 = -20\, \phi^* \omega_{\mathbb{S}_6}.$$

Since $X^o(3,6)$ is regarded as a subvariety of $M^o(3,6)$ via the embedding (7.2), we get on $X^o(3,6)$

$$dd^c \log \|\Delta_{(3,6)}\|^2 = -20\,\varphi^*\omega_{\mathbb{S}_6} \tag{7.7}$$

because

$$\text{L.H.S.} = j^*q^*dd^c \log \|\Delta_{(3,6)}\|^2 = -20\, j^*\phi^*\omega_{\mathbb{S}_6} = -20\, j^*q^*\varphi^*\omega_{\mathbb{S}_6} = \text{R.H.S.}$$

Since $\Pi^{-1} \circ \varphi(X^o(3,6)) = \mathbf{\Omega}^o_{\mathbb{S}_6}$, we deduce from (7.7) the following equation on $\mathbf{\Omega}^o_{\mathbb{S}_6}$

$$dd^c f = -20\,\omega_{\mathbb{S}_6}. \tag{7.8}$$

By the commutativity of (7.5), we get the equation of sets on $\mathbf{H}_2/\,\Gamma_M(1+\mathrm{i})$:

$$(\psi^{-1})^*\mathrm{div}\,\Delta_{(3,6)} = \mathrm{div}\,\Delta_{\mathrm{MSY}}. \tag{7.9}$$

Let $\overline{\mu}\colon \mathbf{H}_2/\,\Gamma_M(1+\mathrm{i}) \to \mathcal{M}_{\mathbb{S}_6}(2) = \mathbf{\Omega}^+_{\mathbb{S}_6}/O^+(\mathbb{T}_6)(2)$ be the isomorphism induced from the isomorphism $\mu\colon \mathbf{H}_2 \cong \mathbf{\Omega}_{\mathbb{S}_6}$. Set $\overline{\mathcal{D}}_{\mathbb{S}_6} := \mathcal{D}_{\mathbb{S}_6}/O^+(\mathbb{T}_6)(2)$. Since $\varphi = \overline{\mu}\circ\psi$, we get by Proposition 7.6 and (7.9)

$$(\varphi^{-1})^*\mathrm{div}\,\Delta_{(3,6)} = (\overline{\mu}^{-1})^*(\psi^{-1})^*\mathrm{div}\,\Delta_{(3,6)} = (\overline{\mu}^{-1})^*\mathrm{div}\,\Delta_{\mathrm{MSY}} = \overline{\mathcal{D}}_{\mathbb{S}_6},$$

which yields the equation of sets

$$\Pi^*(\varphi^{-1})^*\mathrm{div}\,\Delta_{(3,6)} = \mathcal{D}_{\mathbb{S}_6}. \tag{7.10}$$

Set $\mathcal{D}^o_{\mathbb{S}_6} := \amalg_{i<j<k}\Pi^*(\varphi^{-1})^*(X_{ijk}(3,6))$. By Theorem 7.10 (3), $\mathcal{D}^o_{\mathbb{S}_6}$ is smooth and it is a dense Zariski open subset of $\mathcal{D}_{\mathbb{S}_6}$. Let $x \in \mathcal{D}^o_{\mathbb{S}_6}$ be an arbitrary point. Let $\gamma : \Delta \to X^*(3,6)$ be a holomorphic curve intersecting $\amalg_{i<j<k}X_{ijk}(3,6)$ transversally at $\gamma(0)$. By [39, (2.3)], there exists a holomorphic curve $c\colon \Delta \to \mathbf{\Omega}_{\mathbb{S}_6}$ intersecting $\mathcal{D}^o_{\mathbb{S}_6}$ transversally at $c(0)$ such that

$$\Pi \circ c(t) = \varphi \circ \gamma(t^2), \qquad t \in \Delta.$$

By Lemma 7.13, we get

$$f(c(t)) = 2\,\log|t|^2 + O(1) \tag{7.11}$$

because

$$\log \|\Delta_{(3,6)}(\varphi^{-1} \circ \Pi \circ c(t))\|^2 = \log \|\Delta_{(3,6)}(\gamma(t^2))\|^2 = 2\log|t|^2 + O(1).$$

Since $c(0)$ is an arbitrary point of $\mathcal{D}^o_{\mathbb{S}_6}$ and since $c(t)$ intersects $\mathcal{D}^o_{\mathbb{S}_6}$ transversally at $c(0)$, we deduce from (7.8), (7.11) the following equation of currents on $\mathbf{\Omega}_{\mathbb{S}_6}$:

$$dd^c f = -20\,\omega_{\mathbb{S}_6} + 2\,\delta_{\mathcal{D}_{\mathbb{S}_6}}. \tag{7.12}$$

Since f is $O^+(\mathbb{T}_6)(2)$-invariant, it follows from Theorem 3.2 and (7.12) the existence of an automorphic form F for $O^+(\mathbb{T}_6)(2)$ of weight 20 with zero divisor $2\mathcal{D}_{\mathbb{S}_6}$ such that $f = \log\|F\|^2$. Comparing the weights and the zeros of F and Δ^2_{MSY}, we get $F = \text{Const.}\ \Delta^2_{\mathrm{MSY}}$. This proves (7.6). □

Question 7.14. Recall that the constant C_3 was defined as the ratio of $\tau_{\mathbb{S}_6}$ and $\|\Delta_{(3,6)}\|^{-1/4}$ in Theorem 7.9. Is it possible to compute $\log C_3$ in $\mathbf{R}/\mathbf{Q}\log 2$ by using the arithmetic Lefschetz–Riemann–Roch theorem [5], [23]? The corresponding question for the family of elliptic curves over the configuration space $\pi : \mathcal{E} \to M^o(2,4;\mathbf{C})$ was considered by Bost [11], who obtained Eq. (1.4) from the arithmetic Riemann–Roch theorem [6], [16].

Question 7.15. Let L be an even lattice of signature $(2, b^-)$. In [8, Th. 14.3], Borcherds constructed a correspondence from modular forms for $Mp_2(\mathbf{Z})$ of type ρ_L with weight $1+m^+-b^-/2$ to automorphic forms on Ω_L for some cofinite subgroup of $O^+(L)$ of weight m^+. We call this correspondence the *Borcherds additive lifting*, while we call the correspondence in Theorem 5.2 the *multiplicative Borcherds product*. Is it true that the even Freitag theta functions $\{\Theta\binom{ijk}{lmn}\}$ are the Borcherds additive lifting of some modular forms for $Mp_2(\mathbf{Z})$ of type $\rho_{\mathbb{T}_6}$? If it is the case, Theorem 7.9 (1) may be expressed as follows:

$$\prod_{\text{finite}} (\text{additive Borcherds lifting}) = (\text{multiplicative Borcherds product})^{\text{integer}}. \tag{7.13}$$

There are some examples of Eq. (7.13) given by Allcock–Freitag [1] and Kondo [26]; Allcock–Freitag gave an example where the multiplicative Borcherds product is the one given by Borcherds [10] characterizing the discriminant locus on the moduli space of cubic surfaces; Kondo gave an example where the multiplicative Borcherds product is the Borcherds Φ-function of dimension 10 characterizing the discriminant locus on the moduli space of Enriques surfaces. It may be worth asking whether the existence of additive Borcherds liftings such that Eq. (7.13) holds for the automorphic forms $\Phi_{\mathbb{S}_k}$ in Theorem 6.1. Are there many examples of Eq. (7.13)?

Question 7.16. In [27], Krieg studied automorphic forms on the period domain $\Omega_{\mathbb{S}_4}$. There exist analogues of the Freitag theta functions on the period domain $\Omega_{\mathbb{S}_4}$. Is it true that the automorphic form $\Phi_{\mathbb{S}_4}$ has an expression in terms of those theta functions similar to the Matsumoto–Sasaki–Yoshida form Δ_{MSY}?

Question 7.17. In [24], the moduli space and the period map for 2-elementary $K3$ surfaces of type $\mathbb{S}_4$ were studied by Koike, Shiga, Takayama, and Tsutsui. They proved that a general 2-elementary $K3$ surface of type $\mathbb{S}_4$ is obtained as the minimal resolution of the following double covering of $\mathbf{P}^1 \times \mathbf{P}^1$:

$$S(x) := \left\{ ((s,t), w) \in \mathcal{O}_{\mathbf{P}^1}(4) \boxtimes \mathcal{O}_{\mathbf{P}^1}(4);\ w^2 = \prod_{k=1}^{4} (x_1^{(k)}\, st + x_2^{(k)}\, s + x_3^{(k)}\, t + x_4^{(k)}) \right\},$$

where s, t denote the inhomogeneous coordinates of the first $\mathbf{P}^1$ and the second $\mathbf{P}^1$, respectively. Following Koike–Shiga–Takayama–Tsutsui, we set

$$x_k := \begin{pmatrix} x_1^{(k)} & x_2^{(k)} \\ x_3^{(k)} & x_4^{(k)} \end{pmatrix} \in M(2,\mathbf{C}), \qquad 1 \le k \le 4$$

and define for $x = (x_1, x_2, x_3, x_4) \in M(2, \mathbf{C})^4$

$$\eta_x := \prod_{k=1}^{4} \frac{ds \wedge dt}{(x_1^{(k)}\, st + x_2^{(k)}\, s + x_3^{(k)}\, t + x_4^{(k)})^{1/2}} \in H^0(S(x), K_{S(x)}) \setminus \{0\}.$$

Then the following function seems to be an analogue of $\Delta_{(3,6)}$ in the case of 2-elementary $K3$ surfaces of type $\mathbb{S}_4$:

$$\| \Delta_{\mathrm{KSTT}}(x) \|^2 := \left| \prod_{k=1}^{4} \det(x_k) \right|^2 \left(\frac{1}{(2\pi)^2} \int_{S(x)} \eta_x \wedge \bar{\eta}_x \right)^4 .$$

Let $\widetilde{S}(x)$ be the minimal resolution of $S(x)$. It may be worth asking if an analogue of Theorem 7.9 (2) holds in this case, i.e., the existences of a rational number ν and a non-zero real number C with

$$\tau_{\mathbb{S}_4}(\widetilde{S}(x)) = C\, \| \Delta_{\mathrm{KSTT}}(x) \|^{\nu}.$$

Question 7.18. Let $A \in M^o(3, 6)$. Assume that there exists a smooth conic Q_A such that all the six lines $L_{1,A}, \ldots, L_{6,A}$ are tangent to Q_A. Then X_A is a *Kummer surface*. Let C_A be the double covering of Q_A with 6 branch points $L_{1,A} \cap Q_A, \cdots, L_{6,A} \cap Q_A$. Then C_A is a curve of genus 2 and X_A is the Kummer surface associated with the Jacobian variety of C_A, i.e., $X_A = \mathrm{Km}(\mathrm{Jac}(C_A))$. Let $\tau(C_A)$ be the analytic torsion of C_A with respect to the metric induced from the flat Kähler metric on $\mathrm{Jac}(C_A)$. By e.g., [38], $\tau(C_A)$ is expressed as the Petersson norm of the Igusa cusp form. Explain the coincidence of $\tau(C_A)$ and $\tau_{\mathbb{S}_6}(X_A, \iota_A)$.

8 Discriminant of quartic surfaces

8.1 Discriminant of quartic hypersurfaces of $\mathbf{P}^3$

Let $(Z_0 : Z_1 : Z_2 : Z_3)$ be the homogeneous coordinates of $\mathbf{P}^3$. Let $H = \mathcal{O}_{\mathbf{P}^3}(1)$ be the hyperplane bundle over $\mathbf{P}^3$. We identify $Z_0, \ldots, Z_3$ as a basis of $H^0(\mathbf{P}^3, H)$. For an index $I = (i_0, i_1, i_2, i_3)$, we set $|I| = i_0 + \cdots + i_3$ and define $Z^I := Z_0^{i_0} Z_1^{i_1} Z_2^{i_2} Z_3^{i_3}$. Then $\{Z^I\}_{|I|=4}$ is a basis of $H^0(\mathbf{P}^3, 4H)$. Let $\{\xi_I\}_{|I|=4}$ be the coordinates of $H^0(\mathbf{P}^3, 4H)$ with respect to the basis $\{Z^I\}_{|I|=4}$. Then $\{\xi_I\}_{|I|=4}$ is regarded as a basis of the dual vector space $H^0(\mathbf{P}^3, 4H)^{\vee}$.

Let $\Phi_{|4H|} \colon \mathbf{P}^3 \ni Z \hookrightarrow (Z^I) \in \mathbf{P}(H^0(\mathbf{P}^3, 4H))^{\vee}$ be the projective embedding associated with the very ample line bundle $4H$. Let $\Phi_{|4H|}(\mathbf{P}^3)^{\vee} \subset \mathbf{P}(H^0(\mathbf{P}^3, 4H))$ be the projective dual variety of $\Phi_{|4H|}(\mathbf{P}^3)$ (cf. [22]). Then $\Phi_{|4H|}(\mathbf{P}^3)^{\vee}$ is a hypersurface of $\mathbf{P}(H^0(\mathbf{P}^3, 4H))$. The discriminant of quartic hypersurfaces of $\mathbf{P}^3$ is the reduced homogeneous polynomial $\Delta_{(\mathbf{P}^3, 4H)}(\xi) \in \mathbf{Z}[\xi]$ such that

$$\Phi_{|4H|}(\mathbf{P}^3)^{\vee} = \mathrm{div}\, \Delta_{(\mathbf{P}^3, 4H)}(\xi). \tag{8.1}$$

The choice of $\Delta_{(\mathbf{P}^3,4H)}(\xi)$ is unique, up to a constant. We fix one polynomial $\Delta_{(\mathbf{P}^3,4H)}(\xi)$ satisfying (8.1).

We define

$$F(Z,\xi) := \sum_{|I|=4} \xi_I Z^I \in H^0(\mathbf{P}^3, 4H)^\vee \otimes H^0(\mathbf{P}^3, 4H).$$

Set

$$\mathbb{P} := \mathbf{P}(H^0(\mathbf{P}^3, 4H)),$$

$$X_\xi := \{[Z] \in \mathbf{P}^3;\ F(Z,\xi) = 0\}, \qquad \xi \in \mathbb{P},$$

and

$$\mathcal{X} := \{([Z],\xi) \in \mathbf{P}^3 \times \mathbb{P};\ F(Z,\xi) = 0\}, \qquad \pi := \mathrm{pr}_2.$$

Then $\pi : \mathcal{X} \to \mathbb{P}$ is a universal family of quartic hypersurfaces of $\mathbf{P}^3$ with fiber $\pi^{-1}(\xi) = X_\xi$. Let $\mathfrak{D}$ be the discriminant locus of the family $\pi : \mathcal{X} \to \mathbb{P}$:

$$\mathfrak{D} := \{\xi \in \mathbb{P};\ \mathrm{Sing}\, X_\xi \neq \emptyset\},$$

which is an irreducible divisor of $\mathbb{P}$ such that

$$\mathfrak{D} = \mathrm{div}\,\Delta_{(\mathbf{P}^3,4H)}(\xi) = \Phi_{|4H|}(\mathbf{P}^3)^\vee.$$

By a formula of Katz [22, Cor. 5.6], we have

$$\deg \mathfrak{D} = \deg \Delta_{(\mathbf{P}^3,4H)}(\xi) = (-1)^3 \int_{\mathbf{P}^3} \frac{c(T\mathbf{P}^3)}{(1+4\,c_1(H))^2} = 108, \tag{8.2}$$

where $c(T\mathbf{P}^3) = (1 + c_1(H))^4$ denotes the total Chern class of $T\mathbf{P}^3$.

For $\xi \in \mathbb{P} \setminus \mathfrak{D}$, $(X_\xi, H|_{X_\xi})$ is a polarized $K3$ surface of degree 4, i.e., a $K3$ surface equipped with an ample line bundle of degree 4. For $\xi \in \mathbb{P} \setminus \mathfrak{D}$, set

$$\eta_\xi := \mathrm{Res}_{X_\xi} \left(\frac{\sum_{\sigma\in\mathfrak{S}_4} \mathrm{sgn}\,\sigma\ Z_{\sigma(1)}\, dZ_{\sigma(2)} \wedge dZ_{\sigma(3)} \wedge dZ_{\sigma(4)}}{F(Z,\xi)} \right).$$

Then η_ξ is a non-zero holomorphic 2-form on X_ξ.

Definition 8.1. *The norm of* $\Delta_{(\mathbf{P}^3,4H)}(\xi)$ *is defined by*

$$\|\Delta_{(\mathbf{P}^3,4H)}(\xi)\|^2 := \left(\frac{1}{(2\pi)^2} \int_{X_\xi} \eta_\xi \wedge \bar{\eta}_\xi \right)^{108} |\Delta_{(\mathbf{P}^3,4H)}(\xi)|^2.$$

By (8.2), $\|\Delta_{(\mathbf{P}^3,4H)}(\xi)\|$ is a C^∞ function on $\mathbb{P} \setminus \mathfrak{D}$. In this section, we prove that $\|\Delta_{(\mathbf{P}^3,4H)}(\xi)\|$ is expressed as the norm of a Borcherds product on the period domain for polarized $K3$ surfaces of degree 4 (cf. Theorem 8.11).

8.2 The polarized period for quartic surfaces

Fix a primitive vector $\mathbf{h} \in \mathbb{L}_{K3}$ of norm 4. The choice of $\mathbf{h}$ is unique up to an automorphism of $\mathbb{L}_{K3}$. Set

$$\mathbb{T} := \mathbf{h}^{\perp} \cong \mathbb{U} \oplus \mathbb{U} \oplus \mathbb{E}_8 \oplus \mathbb{E}_8 \oplus \langle -4 \rangle.$$

A marking of (X_ξ, H) is an isometry $\alpha\colon H^2(X_\xi, \mathbf{Z}) \cong \mathbb{L}_{K3}$ with $\alpha(c_1(H)) = \mathbf{h}$. There exists a marking of (X_ξ, H). The triple (X_ξ, H, α) is called a marked polarized $K3$ surface of degree 4. The polarized period of (X_ξ, H, α) is the point of $\Omega_{\mathbb{T}}$ defined by

$$\pi(X_\xi, H, \alpha) := [\alpha(\eta_\xi)].$$

We define

$$\mathcal{M}_4 := \Omega_{\mathbb{T}}/O(\mathbb{T}).$$

The Griffiths period of (X_ξ, H) is then defined as the orbit

$$\varpi(X_\xi, H) := O(\mathbb{T}) \cdot [\alpha(\eta_\xi)] \in \mathcal{M}_4.$$

Let $\varpi^o\colon \mathbb{P} \setminus \mathfrak{D} \to \mathcal{M}_4$ be the period map for the universal family of quartic surfaces $\pi\colon (\mathcal{X}, (\mathrm{pr}_1)^* H)|_{\mathbb{P}\setminus\mathfrak{D}} \to \mathbb{P} \setminus \mathfrak{D}$

$$\varpi^o(\xi) := \varpi(X_\xi, H), \qquad \xi \in \mathbb{P} \setminus \mathfrak{D}.$$

As in Section 2, we define the discriminant locus of $\Omega_{\mathbb{T}}$ by

$$\mathcal{D}_{\mathbb{T}} := \bigcup_{d \in \Delta_{\mathbb{T}}} \mathcal{H}_d$$

and set $\overline{\mathcal{D}}_{\mathbb{T}} := \mathcal{D}_{\mathbb{T}}/O(\mathbb{T}) \subset \mathcal{M}_4$. We regard $\mathcal{D}_{\mathbb{T}}$ as a reduced divisor of $\Omega_{\mathbb{T}}$.

Lemma 8.2. *One has* $\varpi^o(\mathbb{P} \setminus \mathfrak{D}) \subset \mathcal{M}_4 \setminus \overline{\mathcal{D}}_{\mathbb{T}}$.

Proof. Let $\xi \in \mathbb{P} \setminus \mathfrak{D}$ and assume that $\varpi(\xi) \in \overline{\mathcal{D}}_{\mathbb{T}}$. There is a marking α of X_ξ and a root $\delta \in \Delta_{\mathbb{T}}$ such that $\pi(X_\xi, H, \alpha) \in \mathcal{H}_\delta$. By the Riemann–Roch theorem, there exists an effective divisor E of X_ξ with $\alpha(c_1([E])) = \pm\delta$. Since $\langle \mathbf{h}, \delta \rangle = 0$, we get $\deg H|_E = 0$, which contradicts the ampleness of H. Hence $\varpi(\xi) \in \mathcal{M}_4 \setminus \overline{\mathcal{D}}_{\mathbb{T}}$. □

Define

$$\mathfrak{D}^o := \{\xi \in \mathbb{P};\ \mathrm{Sing}\, X_\xi \text{ consists of a unique ordinary double point}\}.$$

Let $\mathfrak{D}^o_{\mathrm{reg}} := \mathfrak{D}^o \setminus \mathrm{Sing}\,\mathfrak{D}^o$ be the regular part of $\mathfrak{D}^o$. Since $\mathfrak{D}^o$ is a dense Zariski open subset of $\mathfrak{D}$ by [22, Prop. 3.2], so is $\mathfrak{D}^o_{\mathrm{reg}}$. By the Borel–Kobayashi–Ochiai extension theorem, the period map ϖ^o extends to a holomorphic map from $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}^o_{\mathrm{reg}}$ to $\overline{\mathcal{M}_4}$, the Bail–Borel–Satake compactification of $\mathcal{M}_4$. This extension of ϖ^o is denoted by ϖ.

Let us make a geometric construction of the Borel–Kobayashi–Ochiai extension ϖ. Let $\mathcal{Z} \subset \mathbf{P}^3 \times \Delta$ be a smooth complex threefold such that $p := \mathrm{pr}_2 \colon \mathcal{Z} \to \Delta$ is a proper, surjective holomorphic function without critical points on $\mathcal{Z} \setminus p^{-1}(0)$. Set $Z_t = p^{-1}(t)$ for $t \in \varDelta$. Then $p \colon \mathcal{Z} \to \Delta$ is called an *ordinary singular family* of quartic surfaces if p has a unique, non-degenerate critical point on Z_0 and if Z_t is a quartic surface for all $t \in \varDelta$.

We define

$$\mathcal{H}^o_\delta := \mathcal{H}_\delta \setminus \bigcup_{d \in \varDelta_{\mathbb{T}} \setminus \{\pm\delta\}} \mathcal{H}_d, \qquad \mathcal{D}^o_{\mathbb{T}} := \sum_{d \in \varDelta_{\mathbb{T}}} \mathcal{H}^o_d$$

and set $\overline{\mathcal{D}}^o_{\mathbb{T}} := \mathcal{D}^o_{\mathbb{T}} / O(\mathbb{T})$.

Lemma 8.3. *Let $p \colon \mathcal{Z} \to \Delta$ be an ordinary singular family of quartic surfaces. Let $\overline{c} : \Delta^* \to \mathcal{M}_4 \setminus \overline{\mathcal{D}}_{\mathbb{T}}$ be the Griffiths period map for $p \colon \mathcal{Z} \to \Delta$. Let $\Pi : \Omega_{\mathbb{T}} \to \mathcal{M}_4$ be the natural projection. Then there exist a holomorphic curve $c \colon \Delta \to \Omega_{\mathbb{T}}$ and a root $\delta \in \varDelta_{\mathbb{T}}$ satisfying*
(1) $\Pi \circ c(t) = \overline{c}(t^2)$ *for all $t \in \Delta$ and $c(0) \in \mathcal{H}^o_\delta$;*
(2) *c intersects $\mathcal{H}^o_\delta$ transversally at $c(0)$.*

Proof. **(1)** Let $\widetilde{\Delta}$ be another disc and let $\mathcal{Z} \times_\Delta \widetilde{\Delta}$ be the induced family over $\widetilde{\Delta}$ by the map $\widetilde{\Delta} \ni t \to t^2 \in \Delta$. By e.g., [25, Th. 4.28], there exists a simultaneous resolution $\pi : \widetilde{\mathcal{Z}} \to \mathcal{Z} \times_\Delta \widetilde{\Delta}$, i.e., a resolution satisfying the commutative diagram

$$\begin{array}{ccc} \widetilde{\mathcal{Z}} & \xrightarrow{\ \pi\ } & \mathcal{Z} \times_\Delta \widetilde{\Delta} \\ {\scriptstyle \widetilde{p}}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_2} \\ \widetilde{\Delta} & \xrightarrow{\ \mathrm{id}\ } & \widetilde{\Delta} \end{array}$$

such that $\pi|_{\widetilde{p}^{-1}(t)} \colon \widetilde{p}^{-1}(t) \to \mathrm{pr}_2^{-1}(t)$ is an isomorphism for $t \neq 0$ and it is the minimal resolution for $t = 0$. In particular, $\widetilde{p}$ is a smooth morphism. Set $\widetilde{\pi} := \mathrm{pr}_1 \circ \pi \colon \widetilde{\mathcal{Z}} \to \mathcal{Z}$. For $t \in \widetilde{\Delta}$, we set $\widetilde{Z}_t := \widetilde{p}^{-1}(t)$ and $\widetilde{\pi}_t := \widetilde{\pi}|_{\widetilde{Z}_t} : \widetilde{Z}_t \to Z_{t^2}$. Then $\widetilde{\pi}_t$ is an isomorphism for $t \in \widetilde{\Delta}^*$ and is the minimal resolution for $t = 0$.

Since the family $\widetilde{p} : \widetilde{\mathcal{Z}} \to \widetilde{\Delta}$ is differentiably trivial, it admits a marking α such that $(\widetilde{Z}_t, \widetilde{\pi}^* H, \alpha_t := \alpha|_{\widetilde{Z}_t})$ is a marked polarized $K3$ surface of degree 4 for $t \in \widetilde{\Delta}^*$. Let $c \colon \widetilde{\Delta} \to \Omega_{\mathbb{T}}$ be the period map for the marked family $(\widetilde{p} : \widetilde{\mathcal{Z}} \to \widetilde{\Delta}, \alpha)$. Since $(\widetilde{Z}_t, \widetilde{\pi}^* H) \cong (Z_{t^2}, H)$ for $t \neq 0$, we have $\Pi \circ c(t) = \overline{c}(t^2)$.

Let $E_0 \subset \widetilde{Z}_0$ be the exceptional curve of $\widetilde{p}_0$. Since Z_0 has a unique ordinary double point, the self-intersection number of E_0 is equal to -2. Set

$$\delta := \alpha_0(c_1([E_0])) \in \varDelta_{\mathbb{L}_{K3}}.$$

Since E_0 is an algebraic cycle, we have $c(0) \in \mathcal{H}_\delta$. By the same argument as in [39, p.70 Claim 2], we get $c(0) \in \mathcal{H}^o_\delta$. This proves (1).
(2) Let $K_{\mathcal{Z}}$ be the canonical line bundle of $\mathcal{Z}$. Since $K_{\mathcal{Z}}$ is trivial by e.g., [39, Lem. 2.3], there exists a nowhere vanishing 3-form ξ on $\mathcal{Z}$. For $t \in \Delta$, set

$$\eta_t := \mathrm{Res}_{Z_t} \frac{\xi}{p(z) - t} \in H^0(Z_t, K_{Z_t}) \setminus \{0\}. \tag{8.3}$$

Then $\widetilde{\eta}_t := \eta_{t^2}$ is regarded as a holomorphic 2-form on $\widetilde{Z}_t$ for $t \neq 0$. There exists a system of coordinates (z_1, z_2, z_3) near the critical point of p with

$$p(z) = z_1^2 + z_2^2 + z_3^2. \tag{8.4}$$

In the local expression (8.4), the vanishing cycle $\alpha_t^{-1}(\delta) \in H^2(\widetilde{Z}_t, \mathbf{Z})$ is realized as the following embedded 2-sphere $E_t \subset \widetilde{Z}_t$ under the identification $\widetilde{Z}_t = Z_{t^2}$:

$$E_t := \left\{ (z_1, z_2, z_3) \in \mathbf{C}^3;\ \left(\frac{z_1}{t}\right)^2 + \left(\frac{z_2}{t}\right)^2 + \left(\frac{z_3}{t}\right)^2 = 1,\ \frac{z_1}{t}, \frac{z_2}{t}, \frac{z_3}{t} \in \mathbf{R} \right\}. \tag{8.5}$$

By (8.3), (8.4), (8.5), there exists a germ $\epsilon(t) \in \mathbf{C}\{t\}$ with

$$\langle \alpha_t(\widetilde{\eta}_t), \delta \rangle = \int_{E_t} \eta_{t^2} = t\,\epsilon(t), \qquad \epsilon(0) \neq 0.$$

Fix $l \in \mathbb{T}_{\mathbf{R}}$ with $\langle l, l \rangle \geq 0$. Since $\langle \cdot, \delta \rangle / \langle \cdot, l \rangle$ is an equation defining $\mathcal{H}_\delta^o$, $c(t)$ intersects $\mathcal{H}_\delta^0$ transversally at $c(0)$. This proves (2). □

Lemma 8.4. *The following hold:*
(1) $\varpi(\mathfrak{D}_{\mathrm{reg}}^o) \subset \overline{\mathcal{D}}_{\mathbb{T}}^o$*;*
(2) $\overline{\mathcal{D}}_{\mathbb{T}}^o \subset \mathcal{M}_4 \setminus \mathrm{Sing}\,\mathcal{M}_4$ *and* $\overline{\mathcal{D}}_{\mathbb{T}}^o \subset \overline{\mathcal{D}}_{\mathbb{T}} \setminus \mathrm{Sing}\,\overline{\mathcal{D}}_{\mathbb{T}}$*;*
(3) $\mathfrak{D}_{\mathrm{reg}}^o$ *is an irreducible divisor of* $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}_{\mathrm{reg}}^o$.

Proof. **(1)** The result follows from Lemma 8.3 (1).
(2) One can prove the result by the same argument as in [39, Prop. 1.9].
(3) The result follows from the irreducibility of the divisor $\mathfrak{D}$ of $\mathbb{P}$. □

Let $L \subset \mathbb{P}$ be a line, i.e., a smooth rational curve of degree 1. Then L is *general* if the induced family $\pi|_L \colon \mathcal{X}|_L \to L$ is a Lefschetz pencil, i.e.,
(i) $\mathcal{X}|_L$ is a smooth threefold;
(ii) all the critical points of the projection $\pi|_L$ are non-degenerate;
(iii) any singular fiber of $\pi|_L$ has only one critical point of $\pi|_L$.

By [22, Cor. 3.2.1], the set of general lines of $\mathbb{P}$ is a dense Zariski open subset of the set of all lines of $\mathbb{P}$.

Lemma 8.5. *Let* $L \subset \mathbb{P}$ *be a general line. Let* $\varpi|_L \colon L \to (\mathcal{M}_4 \setminus \overline{\mathcal{D}}_{\mathbb{T}}) \cup \overline{\mathcal{D}}_{\mathbb{T}}^o$ *be the period map for* $\pi|_L \colon \mathcal{X}|_L \to L$*. Then* $\varpi|_L$ *intersects* $\overline{\mathcal{D}}_{\mathbb{T}}^o$ *transversally at* $\varpi(L \cap \mathfrak{D})$.

Proof. The result follows from Lemma 8.3 (2). □

Let $\omega_{\mathbb{T}}$ be the Kähler form of the Bergman metric on $\Omega_{\mathbb{T}}$:

$$\omega_{\mathbb{T}}([\eta]) = -dd^c \log \frac{\langle \eta, \bar{\eta} \rangle}{|\langle \eta, \mathbf{l} \rangle|^2}, \qquad [\eta] \in \Omega_{\mathbb{T}}, \tag{8.6}$$

where $\mathbf{l} \in \mathbb{T}_{\mathbf{R}}$ is a fixed vector with $\langle \mathbf{l}, \mathbf{l} \rangle \geq 0$. Since $\omega_{\mathbb{T}}$ is invariant under the action of $\mathrm{Aut}(\Omega_{\mathbb{T}})$, it descends to a Kähler form $\omega_{\mathcal{M}_4}$ on $\mathcal{M}_4$ in the sense of orbifolds. By (8.6) and the definition of the period map ϖ, we get the following equation of $(1,1)$-forms on $\mathbb{P} \setminus \mathfrak{D}$:

$$dd^c \log \| \Delta_{(\mathbf{P}^3, 4H)}(\xi) \|^2 = -108 \, (\varpi^o)^* \omega_{\mathcal{M}_4}. \tag{8.7}$$

Lemma 8.6. *The semi-positive* $(1,1)$*-form* $(\varpi^o)^*\omega_{\mathcal{M}_4}$ *on* $\mathbb{P} \setminus \mathfrak{D}$ *has Poincaré growth along* $\mathfrak{D}_{\mathrm{reg}}$. *In particular,* $(\varpi^o)^*\omega_{\mathcal{M}_4}$ *extends trivially to a closed positive* $(1,1)$*-current on* $\mathbb{P}$.

Proof. By the same argument as in [39, Prop. 3.8 and Th. 3.9] using the Schwarz lemma for Bergman metrics on symmetric bounded domains, the semi-positive $(1,1)$-form $(\varpi^o)^*\omega_{\mathcal{M}_4}$ has Poincaré growth along $\mathfrak{D}_{\mathrm{reg}}$. It extends trivially to a closed positive $(1,1)$-current on $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}_{\mathrm{reg}}$ by an extension theorem of Skoda. Since $\mathrm{Sing}\, \mathfrak{D}$ is a subvariety of $\mathbb{P}$ with codimension ≥ 2 and with $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}_{\mathrm{reg}} = \mathbb{P} \setminus \mathrm{Sing}\, \mathfrak{D}$, the result follows from Siu's extension theorem [34, p. 53 Th. 1]. □

The trivial extension of $(\varpi^o)^*\omega_{\mathcal{M}_4}$ from $\mathbb{P} \setminus \mathfrak{D}$ to $\mathbb{P}$ is denoted by $\varpi^*\omega_{\mathcal{M}_4}$.

Lemma 8.7. *The function* $\log \| \Delta_{(\mathbf{P}^3, 4H)}(\xi) \|^2$ *is locally integrable on* $\mathbb{P}$ *and satisfies the following equation of* $(1,1)$*-currents on* $\mathbb{P}$*:*

$$dd^c \log \| \Delta_{(\mathbf{P}^3, 4H)}(\xi) \|^2 = \delta_{\mathfrak{D}} - 108 \, \varpi^* \omega_{\mathcal{M}_4}. \tag{8.8}$$

Proof. By (8.7) and Siu's extension theorem, it suffices to prove the assertion on $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}^o_{\mathrm{reg}}$. By the same argument as in [39, Prop. 3.11], it suffices to prove the following: let $\gamma : \Delta \to \mathbb{P}$ be a holomorphic curve intersecting $\mathfrak{D}^o_{\mathrm{reg}}$ transversally at $\gamma(0)$. Then

$$\log \| \Delta_{(\mathbf{P}^3, 4H)}(\gamma(t)) \|^2 = \log |t|^2 + O(1). \tag{8.9}$$

Since $X_{\gamma(0)}$ has only one ordinary double point as its singular set, the function $\log(\int_{X_{\gamma(t)}} \eta_{\gamma(t)} \wedge \bar{\eta}_{\gamma(t)})$ is bounded as $t \to 0$ by [37, Proof of Theorem 8.1]. By Definition 8.1, we get (8.9). □

8.3 A Borcherds product

Let $\mathbb{D}_7$ be the root lattice of type D_7, which is assumed to be *negative-definite*. Then $\mathbb{D}_7$ is a primitive sublattice of $\mathbb{E}_8$ with $\mathbb{D}_7^{\perp} = \langle -4 \rangle$. Hence $\mathbb{T}$ is regarded as the orthogonal complement of $\mathbb{D}_7$ in $\mathbb{U} \oplus \mathbb{U} \oplus \mathbb{E}_8 \oplus \mathbb{E}_8 \oplus \mathbb{E}_8$:

$$\mathbb{T} = \{(x, y, a, b, c) \in \mathbb{U} \oplus \mathbb{U} \oplus \mathbb{E}_8 \oplus \mathbb{E}_8 \oplus \mathbb{E}_8;\ \langle c, \mathbb{D}_7 \rangle = 0\}.$$

Since $\mathbb{U} \oplus \mathbb{U} \oplus \mathbb{E}_8 \oplus \mathbb{E}_8 \oplus \mathbb{E}_8$ is unimodular, we get

$$A_{\mathbb{T}} = A_{\mathbb{D}_7} = A_{\langle -4 \rangle} = \frac{1}{4}\mathbf{Z}/\mathbf{Z} = \left\{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}\right\}. \tag{8.10}$$

In what follows, $0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}$ often denote the corresponding elements of the discriminant group $A_{\mathbb{T}} = A_{\mathbb{D}_7} = A_{\langle -4\rangle}$.

Let $\mathbf{e}_0, \mathbf{e}_{1/4}, \mathbf{e}_{2/4}, \mathbf{e}_{3/4}$ be the standard basis of $\mathbf{C}[A_{\mathbb{D}_7}]$. Let $\Theta_{\mathbb{D}_7}(\tau)$ be the theta series of the lattice $\mathbb{D}_7$:

$$\Theta_{\mathbb{D}_7}(\tau) := \theta_{\mathbb{D}_7}(\tau)\,\mathbf{e}_0 + \theta_{\mathbb{D}_7+1/4}(\tau)\,\mathbf{e}_{1/4} + \theta_{\mathbb{D}_7+2/4}(\tau)\,\mathbf{e}_{2/4} + \theta_{\mathbb{D}_7+3/4}(\tau)\,\mathbf{e}_{3/4},$$

where

$$\theta_{\mathbb{D}_7+\delta/4}(\tau) := \sum_{l\in\mathbb{D}_7+\delta/4} q^{-\langle l,l\rangle}, \qquad q = e^{2\pi i\tau}.$$

Notice that $\mathbb{D}_7$ is negative-definite.

Lemma 8.8. *$\Theta_{\mathbb{D}_7}(\tau)/\Delta(\tau)$ is a modular form for $Mp_2(\mathbf{Z})$ of type $\rho_{\mathbb{T}}$ of weight $-17/2$.*

Proof. Since $\Delta(\tau)$ is a modular form for $SL_2(\mathbf{Z})$ of weight 12 and since $\rho_{\mathbb{T}} = \rho_{\mathbb{D}_7}$ by (8.10), it suffices to prove that $\Theta_{\mathbb{D}_7}(\tau)$ is a modular form for $Mp_2(\mathbf{Z})$ of weight $7/2$ and of type $\rho_{\mathbb{D}_7}$. This follows from [8, Th. 4.1]. □

Lemma 8.9. *The following identity holds:*

$$\Theta_{\mathbb{D}_7}(\tau)/\Delta(\tau) \equiv (q^{-1} + 108)\,\mathbf{e}_0 + 2^6\,q^{-1/8}\,\mathbf{e}_{1/4} + 14\,q^{-1/2}\,\mathbf{e}_{2/4} + 2^6\,\mathbf{e}_3 \mod q.$$

Proof. Recall that the Jacobi theta functions $\theta_2(\tau)$, $\theta_3(\tau)$, $\theta_4(\tau)$ were defined in Sect. 6. By [13, Chap. 4, p.118, Eqs. (8.7), (8.8), (8.9)], we get

$$\Theta_{\mathbb{D}_7}(\tau) = \frac{\theta_3(\tau)^7 + \theta_4(\tau)^7}{2}\,\mathbf{e}_0 + \frac{\theta_2(\tau)^7}{2}\,\mathbf{e}_{1/4} + \frac{\theta_3(\tau)^7 - \theta_4(\tau)^7}{2}\,\mathbf{e}_{2/4} + \frac{\theta_2(\tau)^7}{2}\,\mathbf{e}_{3/4}. \tag{8.11}$$

By the definitions of the Jacobi theta functions, we get

$$\theta_2(\tau)^7 = 2^7\,q^{7/8} + 7\cdot 2^7\,q^{15/8} + O(q^2),$$

$$\theta_3(\tau)^7 = 1 + 14\,q^{1/2} + 84\,q + O(q^{3/2}),$$

$$\theta_4(\tau)^7 = 1 - 14\,q^{1/2} + 84\,q + O(q^{3/2}),$$

which, together with (8.11), yield that

$$\Theta_{\mathbb{D}_7}(\tau) \equiv (1+84\,q)\,\mathbf{e}_0 + 2^6 q^{7/8}\,\mathbf{e}_{1/4} + 14\,q^{1/2}\,\mathbf{e}_{2/4} + 2^6\,q^{7/8}\,\mathbf{e}_3 \mod q^{3/2}. \tag{8.12}$$

The result follows from (8.12) and the identity $1/\Delta(\tau) = q^{-1} + 24 + O(q)$. □

By Lemma 8.8, we can apply Theorem 5.2 to the lattice $\mathbb{T}$ and the modular form $\Theta_{\mathbb{D}_7}(\tau)/\Delta(\tau)$.

Lemma 8.10. *Let $\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)$ be the Borcherds product associated with $\mathbb{T}$ and $\Theta_{\mathbb{D}_7}(\tau)/\Delta(\tau)$. Then $\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)$ has weight* 54 *and the zero divisor*

$$\operatorname{div}\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta) = \mathcal{D}_{\mathbb{T}} + 2^7 \sum_{d\in\mathbb{T}+1/4,\, d^2=-1/4} \mathcal{H}_d + 14 \sum_{d\in\mathbb{T}+2/4,\, d^2=-1} \mathcal{H}_d.$$

Proof. By Theorem 5.2 (1) and Lemma 8.9, the weight of $\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)$ is given by $c_0(0)/2 = 108/2 = 54$. By Theorem 5.2 (2) and Lemma 8.9, we get

$$\begin{aligned}\operatorname{div}\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta) &= \sum_{d\in\Delta_{\mathbb{T}}} \mathcal{H}_d + 2^6 \sum_{d\in\mathbb{T}+1/4,\, d^2=-1/4} \mathcal{H}_d + 14 \sum_{d\in\mathbb{T}+2/4,\, d^2=-1} \mathcal{H}_d \\ &\quad + 2^6 \sum_{d\in\mathbb{T}+3/4,\, d^2=-1/4} \mathcal{H}_d \\ &= \mathcal{D}_{\mathbb{T}} + 2^7 \sum_{d\in\mathbb{T}+1/4,\, d^2=-1/4} \mathcal{H}_d + 14 \sum_{d\in\mathbb{T}+2/4,\, d^2=-1} \mathcal{H}_d,\end{aligned}$$

where we used $\mathcal{H}_d = \mathcal{H}_{-d}$ to get the second equality. □

Define the effective divisor $\mathcal{D}'$ on $\Omega_{\mathbb{T}}$ by

$$\mathcal{D}' := 2^7 \sum_{d\in\mathbb{T}+1/4,\, d^2=-1/4} \mathcal{H}_d + 14 \sum_{d\in\mathbb{T}+2/4,\, d^2=-1} \mathcal{H}_d$$

and set $\overline{\mathcal{D}}' := \mathcal{D}'/O(\mathbb{T})$. Then $\overline{\mathcal{D}}'$ is an effective divisor of $\mathcal{M}_4$.

The discriminant $\Delta_{(\mathbf{P}^3,4H)}(\xi)$ is expressed as the Borcherds product:

Theorem 8.11. *There exists a non-zero constant C such that the following identity of C^∞ functions on $\mathbb{P}\setminus\mathfrak{D}$ holds:*

$$\|\Delta_{(\mathbf{P}^3,4H)}\|^2 = C\,\varpi^*\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4.$$

Proof. By the Poincaré–Lelong formula and Lemma 8.10, we get the following equation of currents on $\Omega_{\mathbb{T}}$:

$$dd^c \log\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^2 = \delta_{\mathcal{D}_{\mathbb{T}}} + \delta_{\mathcal{D}'} - 54\,\omega_{\mathbb{T}},$$

which descends to the following equation of currents on $\mathcal{M}_4$:

$$dd^c \log\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^2 = \frac{1}{2}\delta_{\overline{\mathcal{D}}_{\mathbb{T}}} + \delta_{\overline{\mathcal{D}}'} - 54\,\omega_{\mathcal{M}_4}. \tag{8.13}$$

In (8.13), the coefficient $1/2$ of $\delta_{\overline{\mathcal{D}}_{\mathbb{T}}}$ is necessary because the natural projection $\Omega_{\mathbb{T}} \to \mathcal{M}_4$ doubly ramifies along $\mathcal{D}_{\mathbb{T}}$ (cf. [39, Prop. 1.9 (4)]).

Since $\varpi^*\overline{\mathcal{D}}^o_{\mathbb{T}} \subset \mathfrak{D}^o_{\mathrm{reg}}$ by Lemma 8.4 (1) and since $\mathfrak{D}^o_{\mathrm{reg}}$ is an *irreducible* divisor of $(\mathbb{P}\setminus\mathfrak{D}) \cup \mathfrak{D}^o_{\mathrm{reg}}$ by Lemma 8.4 (3), there exists an integer $\nu \geq 1$ with

$$\varpi^*\overline{\mathcal{D}}^o_{\mathbb{T}} = \nu\,\mathfrak{D}^o_{\mathrm{reg}}. \tag{8.14}$$

Let $L \subset \mathbb{P}$ be a general line. We compute the intersection number of L and the divisor $\varpi^*\overline{\mathcal{D}}^o_{\mathbb{T}}$. Since the period map $\varpi|_L\colon L \to (\mathcal{M}_4 \setminus \overline{\mathcal{D}}_{\mathbb{T}}) \cup \overline{\mathcal{D}}^o_{\mathbb{T}}$ intersects $\overline{\mathcal{D}}^o_{\mathbb{T}}$ transversally at $\varpi(L\cap\mathfrak{D}^o_{\mathrm{reg}})$ by Lemma 8.5, we get by (8.14)

$$\nu\,\#(L\cap\mathfrak{D}^o_{\mathrm{reg}}) = \#(L\cap\varpi^*\overline{\mathcal{D}}^o_{\mathbb{T}}) = \#(\varpi(L)\cap\overline{\mathcal{D}}^o_{\mathbb{T}}) = \#(L\cap\mathfrak{D}^o_{\mathrm{reg}}),$$

which yields that $\nu = 1$.

Let x be an arbitrary point of $\mathcal{M}_4$. Let $f = 0$ be a local equation near x defining the divisor $\overline{\mathcal{D}}_{\mathbb{T}} + 2\overline{\mathcal{D}}'$. (When $x \notin \overline{\mathcal{D}}_{\mathbb{T}} + 2\overline{\mathcal{D}}'$, we can choose f to be a non-zero constant function.) By (8.13), $\log(\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4/|f|^2)$ is a local potential function for $-108\,\omega_{\mathcal{M}_4}$:

$$dd^c \log(\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4/|f|^2) = -108\,\omega_{\mathcal{M}_4} \tag{8.15}$$

as currents on an open subset of $\mathcal{M}_4$. Let $\xi \in \mathbb{P}$ be a point with $\varpi(\xi) = x$. Since $\log(\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4/|f|^2)$ is locally bounded near x, we deduce from [39, Prop. 3.11] the following equation of currents near ξ:

$$dd^c \varpi^* \log(\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4/|f|^2) = -108\,\varpi^*\omega_{\mathcal{M}_4}. \tag{8.16}$$

Since x is an arbitrary point of $\mathcal{M}_4$ and hence ξ is an arbitrary point of $(\mathbb{P}\setminus\mathfrak{D})\cup\mathfrak{D}^o_{\mathrm{reg}}$, we deduce from (8.16) and $\nu = 1$ the following equation of currents on $(\mathbb{P}\setminus\mathfrak{D})\cup\mathfrak{D}^o_{\mathrm{reg}}$:

$$dd^c \varpi^* \log \|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^2 = \frac{1}{2}\,\delta_{\mathfrak{D}^o_{\mathrm{reg}}} + \delta_{\varpi^*\overline{\mathcal{D}}'} - 54\,\varpi^*\omega_{\mathcal{M}_4}. \tag{8.17}$$

Comparing (8.8) and (8.17), we get the equation of currents on $(\mathbb{P} \setminus \mathfrak{D}) \cup \mathfrak{D}^o_{\mathrm{reg}}$:

$$dd^c \log \frac{\|\Delta_{(\mathbf{P}^3,4H)}\|^2}{\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4} = -2\,\delta_{\varpi^*\overline{\mathcal{D}}'}. \tag{8.18}$$

Set $F := \log(\|\Delta_{(\mathbf{P}^3,4H)}\|^2/\|\Psi_{\mathbb{T}}(\cdot, \Theta_{\mathbb{D}_7}/\Delta)\|^4)$. Since $\mathfrak{D} \setminus \mathfrak{D}^o_{\mathrm{reg}}$ is a subvariety of $\mathbb{P}$ whose codimension is strictly greater than 1, we deduce from Siu's extension theorem [34, p.53 Th. 1] that $F \in L^1(\mathbb{P})$ and that Eq. (8.18) holds on $\mathbb{P}$. Assume that $\varpi^*\overline{\mathcal{D}}' \neq \emptyset$. Let $L \subset \mathbb{P}$ be a general line. By (8.18), $\partial F|_L$ is a logarithmic 1-form on L with $\mathrm{div}(\partial F|_L) = (\varpi^*\overline{\mathcal{D}}') \cap L$. Since $\varpi^*\overline{\mathcal{D}}'$ is an effective divisor and hence so is $(\varpi^*\overline{\mathcal{D}}') \cap L$, the sum of the residues of $\partial F|_L$ does not vanish, which contradicts the residue theorem. Hence $\varpi^*\overline{\mathcal{D}}' = \emptyset$ and F is a constant function on $\mathbb{P}$. This proves the theorem. □

We do not know if $\|\Delta_{(\mathbf{P}^3,4H)}\|$ admits an analytic expression using (equivariant) analytic torsion. After Beauville [3, Sect. 6], Voisin [35], Huybrechts [21, Example 2.7], it is possible to associate to X an irreducible compact holomorphic symplectic 4-fold with anti-symplectic involution as follows.

For a smooth quartic surface $X \subset \mathbf{P}^3$, let $\mathrm{Hilb}^{(2)}(X)$ denote the Hilbert scheme of zero-cycles of degree 2 of X, which is a symplectic resolution of the second symmetric product of X. Since X is a quartic surface, $\mathrm{Hilb}^{(2)}(X)$ has a natural involution defined as follows. Let $P_1 + P_2$, $P_1 \neq P_2$, be a point of $\Sigma^{(2)}X$, the second symmetric product of X. Let L be the line of $\mathbf{P}^3$ connecting P_1 and P_2. Then there exist $P_3, P_4 \in X$ such that $X \cap L = \{P_1, P_2, P_3, P_4\}$. Let Δ be the diagonal locus of $\Sigma^{(2)}X$. We define the involution $\theta\colon \Sigma^{(2)}X \setminus \Delta \to \Sigma^{(2)}X \setminus \Delta$ by $\theta(P_1 + P_2) := P_3 + P_4$. By [3, Sect. 6 Prop. 11], θ extends to an anti-symplectic holomorphic involution on $\mathrm{Hilb}^{(2)}(X)$. As an analogue of Theorem 4.3, it may be worth asking the following:

Question 8.12. Is it possible to express $\|\Delta_{(\mathbf{P}^3,4H)}\|^2$ as a combination of the equivariant analytic torsions of the bundles $\Omega^p_{\mathrm{Hilb}^{(2)}(X)}$, $p \geq 0$?

References

1. Allcock, D., Freitag, E. *Cubic surfaces and Borcherds products*, Comment. Math. Helv. **77** (2002), 270–296
2. Barth, W., Peters, C., Van de Ven, A. *Compact Complex Surfaces*, Springer, Berline (1984)
3. Beauville, A. *Some remarks on Kähler manifolds with* $c_1 = 0$, Classification of algebraic and analytic manifolds, Progress Math. **39** Birkhäuser, Boston, (1983), 1–26
4. Bershadsky, M., Cecotti, S., Ooguri, H., Vafa, C. *Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes*, Commun. Math. Phys. **165** (1994), 311–428
5. Bismut, J.-M. *Equivariant immersions and Quillen metrics*, Jour. Differ. Geom. **41** (1995), 53–157
6. Bismut, J.-M., Lebeau, G. *Complex immersions and Quillen metrics*, Publ. Math. IHES **74** (1991), 1–297
7. Bismut, J.-M., Gillet, H., Soulé, C. *Analytic torsion and holomorphic determinant bundles I, II, III*, Comm. Math. Phys. **115** (1988) 49–78, 79–126, 301–351
8. Borcherds, R.E. *Automorphic forms with singularities on Grassmanians*, Invent. Math. **132** (1998), 491–562
9. ——— *Reflection groups of Lorenzian lattices*, Duke Math. Jour. **104** (2000), 319–366
10. ——— *An automorphic form related to cubic surfaces*, math.AG/0002079 (2000)
11. Bost, J.-B. *Théorie de l'intersection et théorème de Riemann–Roch arithmétiques*, Astérisque **201–203** (1991), 43–88
12. Bruinier, J.H. *Borcherds Products on* $O(2,l)$ *and Chern Classes of Heegner Divisors*, Lecture Notes Math. **1780** Springer (2002)
13. Conway, J.H., Sloan, N.J.A. *Sphere packings, Lattices and Groups*, Springer Berlin (1999)
14. Dolgachev, I.V. *Mirror symmetry for lattice polarized* $K3$ *surfaces*, J. Math. Sci. **81** (1996), 2599–2639
15. Fang, H., Lu, Z., Yoshikawa, K.-I. *Analytic torsion for Calabi–Yau threefolds*, preprint, arXiv: math.DG/0601411 (2006)
16. Gillet, H., Soulé, C. *An arithmetic Riemann–Roch theorem*, Invent. Math. **110** (1992), 473–543
17. Gritsenko, V.A., Nikulin, V.V. $K3$ *surfaces, Lorentzian Kac–Moody algebras and mirror symmetry*, Math. Res. Lett. **3** (1996), 211–229
18. ——— *Siegel automorphic form corrections of some Lorentzian Kac–Moody Lie algebras*, Amer. Jour. Math. **119** (1997), 181–224
19. ——— *Automorphic forms and Lorentzian Kac–Moody algebras, I, II*, Internat. Jour. Math. **9** (1998), 153–199, 201–275
20. Harvey, J., Moore, G. *Exact gravitational threshold correction in the Ferrara-Harvey-Strominger-Vafa model*, Phys. Rev. D **57** (1998), 2329–2336
21. Huybrechts, D. *Compact hyper-Kähler manifolds: basic results*, Invent. Math. **135** (1999), 63–113
22. Katz, N. *Pinceaux de Lefschetz: théorème d'existence*, Groupe de monodromie en géométrie algébrique, Exposé XVII, SGA 7,II Lecture Notes Math. **340** Springer (1973), 213–253

23. Köhler, K., Roessler, D. *A fixed point formula of Lefschetz type in Arakelov geometry I*, Invent. Math. **145** (2001), 333–396
24. Koike, K., Shiga, H., Takayama N., Tsutsui, T. *Study on the family of* $K3$ *surfaces induced from the lattice* $(D_4)^3 \oplus \langle -2 \rangle \oplus \langle 2 \rangle$, Internat. Jour. Math. **12** (2001), 1049–1085
25. Kollár, J., Mori, S. *Birational Geometry of Algebraic Varieties*, Cambridge Univ. Press, Cambridge (1998)
26. Kondo, S. *The moduli space of Enriques surfaces and Borcherds products*, J. Alg. Geom. **11** (2002), 601–627
27. Krieg, A. *Modular Forms on Half-Spaces of Quaternions*, Lecture Notes Math. **1143** Springer (1985)
28. Lang, S. *Elliptic Functions*, Springer, Berlin (1987)
29. Ma, X. *Submersions and equivariant Quillen metrics*, Ann. Inst. Fourier **50** (2000), 1539–1588
30. Matsumoto, K. *Theta functions on the bounded symmetric domain of type* $I_{2,2}$ *and the period map of a* 4*-parameter family of* $K3$ *surfaces*, Math. Ann. **295** (1993), 383–409
31. Matsumoto, K., Sasaki, T., Yoshida, M. *The monodromy of the period map of a* 4*-parameter family of* $K3$ *surfaces and the hypergeometric function of type* (3, 6), Internat. Jour. Math. **3** (1992), 1–164
32. Nikulin, V.V. *Integral Symmetric bilinear forms and some of their applications*, Math. USSR Izv. **14** (1980) 103–167
33. ______ *Factor groups of groups of automorphisms of hyperbolic forms with respect to subgroups generated by* 2*-reflections*, J. Soviet Math. **22** (1983), 1401–1476
34. Siu, Y.-T. *Analyticity of sets associated to Lelong numbers and the extension of closed positive currents*, Invent. Math. **27** (1974), 53–156
35. Voisin, C. *Sur la stabilité des sous-variétés lagrangiennes des variétés symplectiques holomorphes*, Complex projective geometry, London Math. Soc. Lecture Note **179** (1989), 294–303
36. Yoshida, M. *Hypergeometric Functions, My Love*, Aspects of Math. Vieweg, Braunschweig (1997)
37. Yoshikawa, K.-I. *Smoothing of isolated hypersurface singularities and Quillen metrics*, Asian J. Math. **2** (1998), 325–344
38. ______ *Discriminant of theta divisors and Quillen metrics*, Jour. Differ. Geom. **52** (1999), 73–115
39. ______ $K3$ *surfaces with involution, equivariant analytic torsion, and automorphic forms on the moduli space*, Invent. Math. **156** (2004), 53–117
40. ______ *Analytic torsion and automorphic forms on the moduli space*, Sugaku Expositions **17** (2004), 1–21
41. ______ *Borcea-Voisin threefolds, analytic torsion, and automorphic forms*, in preparation