Algebra & Number Theory

Volume 18
2024
No. 10

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2024 is US \$525/year for the electronic version, and \$770/year (+\$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 18:10 (2024)
https://doi.org/10.2140/ant.2024.18.1767

A case study of intersections on blowups of the moduli of curves

Sam Molcho and Dhruv Ranganathan

We explain how logarithmic structures select principal components in an intersection of schemes. These manifest in Chow homology and can be understood using strict transforms under logarithmic blowups. Our motivation comes from Gromov–Witten theory. The *toric contact cycles* in the moduli space of curves parameterize curves that admit a map to a fixed toric variety with prescribed contact orders. We show that they are intersections of virtual strict transforms of double ramification cycles in blowups of the moduli space of curves. We supply a calculation scheme for the virtual strict transforms, and deduce that toric contact cycles lie in the tautological ring of the moduli space of curves. This is a higher-dimensional analogue of a result of Faber and Pandharipande. The operational Chow rings of Artin fans play a basic role, and are shown to be isomorphic to rings of piecewise polynomials on associated cone complexes. The ingredients in our analysis are Fulton's blowup formula, Aluffi's formulas for Segre classes of monomial schemes, piecewise polynomials, and degeneration methods. A model calculation in toric intersection theory is treated without logarithmic methods and may be read independently.

1. Introduction

Logarithmic geometry plays a central role in the construction of compact moduli spaces in algebraic geometry. In recent years, researchers have applied these techniques to study a range of questions in enumerative geometry and the moduli space of curves, and our work is especially motivated by [29; 30; 32; 40; 41; 45; 49; 55]. These results bring into sharp focus a basic phenomenon: logarithmic moduli spaces arise naturally as an infinite collection of schemes assembled in an inverse system. The system is analogous to the system of equivariant compactifications of a fixed torus. It occurs frequently that the inverse system does not contain a minimal model, and even when it does, expected intersection theory

MSC2020: 14A21, 14H10.
Keywords: logarithmic geometry, double ramification cycle, Gromov–Witten theory.

statements may fail on the minimal model but hold on finer members of the system. Working with the whole inverse system reveals additional structure.[1] Instances include the following:

(i) The *logarithmic Picard group* is not representable by a scheme with a logarithmic structure; its geometry is fully captured by an inverse system of toroidal compactifications of semiabelian schemes. In what appears to be an analogous situation, no minimal model of the *logarithmic Hilbert scheme of curves* is known [41; 45].

(ii) Toroidal compactifications of the *moduli space of abelian varieties* appear in an inverse system without a minimal representative; the situation is analogous to that of the logarithmic Picard group [36].

(iii) The naive *degeneration formula* in logarithmic Gromov–Witten theory fails to hold for the minimal model of the space of stable maps, but holds on sufficiently fine models [55]. The situation in logarithmic Donaldson–Thomas theory is expected to be similar [41].

(iv) The naive *product formulas* in logarithmic Abel–Jacobi theory and Gromov–Witten theory for logarithmic and rubber geometries fail to hold on the space of stable maps, but hold on sufficiently fine models [29; 32; 49].

We provide a framework to understand these phenomena. We illustrate, by means of a detailed case study on the moduli space of curves and the problem (iv) above, the intersection theory of cycles on such inverse systems. We deduce that certain natural cohomology classes, the *toric contact cycles*, lie in the tautological ring of the moduli space of curves; see Theorem A.

We first outline the basic geometry and then state the main results.

1.1. ***The idealized problem.*** We contemplate a smooth, noncompact moduli space U of nondegenerate objects, and a system of logarithmic compactifications $U \subset Y_\alpha$, indexed by a combinatorial datum α, usually a choice of polyhedral structure on a topological space. In the ideal scenario, each compactification contains U as the complement of a normal crossings divisor. Any two compactifications Y_α, Y_β in the system are dominated by a third Y_γ, via a birational morphism: an iteration of blowups and root stack constructions on strata in the complement of U.

We examine intersections of cycles that occur on spaces in the inverse system. Intersection theory requires us to either pick a compactification arbitrarily or work with the system of all compactifications simultaneously; our paper concerns the comparison between the two. In the latter case, the Chow homology and cohomology groups may be defined formally as limits and colimits of the corresponding groups of the system. A pair (Y_α, V_α) of a compactification of U and a cycle V_α on it determines a compatible system of classes in the limit Chow groups, by pulling back along blowups. By cardinality considerations, comparatively few classes in limit Chow homology arise in this way.

[1] We are by no means the first to notice this, as Mumford puts it: "the nonuniqueness (of compactification) gives one freedom to seek for the most elegant solutions in any particular case."; see [47].

Nevertheless, in logarithmic moduli problems, birational invariance statements imply that many classes of interest do arise by pullback of a cycle along *some* element [2; 41; 55]. If this holds, we are left to understand the intersection product of cycles V_α and V_β on different models. There are two possibilities for this intersection product, and we illustrate the geometry in the simplest situation.

1.2. ***The toric specialization.*** Let U be a torus, and V_1, V_2 two subvarieties intersecting transversely. On any smooth toric compactification $U \subset Y$, we may form the closures

$$\overline{V}_1^Y, \overline{V}_2^Y \subset Y$$

to obtain

$$V_1 \cdot_Y V_2 := \overline{V}_1^Y \cdot \overline{V}_2^Y \quad \text{in } \mathsf{CH}_\star(Y).$$

Compact models are related by blowups and blowdowns along smooth centers. Given a blowup $\pi : Y' \to Y$, the classes $V_1 \cdot_Y V_2$ and $V_1 \cdot_{Y'} V_2$ are typically *not* related either by pushforward or pullback. The closures in different compactifications are instead related by *strict transforms*.

By an elementary but crucial argument, if the pair of cycles is fixed, there is a *stable* answer: if Y is any *sufficiently fine* toric compactification of U, the intersection product is a well-defined cycle class, independent of the choice of compactification. In other words, if Y is replaced by a blowup along a smooth center, the class is pulled back. The question of whether a given compactification is fine enough is a tropical question, answered by a theorem of Tevelev [58]: the tropicalizations of V_1 and V_2 must each be unions of cones in the fan of Y.

We view the stable class $[V_1 \cdot_Y V_2]$ as a prototype for the *logarithmic intersection product* and view it as an element in the direct limit of Chow cohomology groups of all toric compactifications of U.

Question. *How does the class of the logarithmic intersection product differ from the ordinary intersection product on an arbitrarily chosen model Y that is not sufficiently fine*?

We study this problem in a more general context, replacing toric varieties and cycles with logarithmically flat schemes, and examine intersections with logarithmic local complete intersection morphisms. Our analysis involves four tools: (i) the resolution package for toroidal schemes and morphisms [1; 3; 7; 42], (ii) Fulton's blowup formula [24], (iii) Aluffi's formula for Segre classes of monomial subschemes [10], and (iv) the ring of piecewise polynomial functions on a logarithmic scheme [16; 52]. The ingredients are combined by using Artin fans [2; 5; 6].

1.3. ***The moduli space of curves.*** After the formalism has been set up, we illustrate the geometry by a detailed case study on the moduli space of curves. Fix a toric variety X of dimension r and consider smooth n-pointed curves of genus g with a map to X, where each marked point has a fixed contact order with each toric divisor. All contact points with the boundary will be marked and two maps are considered equivalent if they are translates under the torus action. Logarithmic geometry gives rise to a compactification of this cycle, known as the *toric contact cycle*.

Theorem A. *The toric contact cycles are contained in the tautological ring of the moduli space of curves.*

The class arises naturally on the moduli space of logarithmic stable maps to X; see Remark 4.6.4. When X is $\mathbb{P}^1$ the statement was proved by Faber and Pandharipande [23], and their result is an ingredient in ours. The result might be viewed as saying that there is hope in finding a *formula* for the class. Elementary examples in genus 1 and 2 can be computed using our methods, however a complete formula demands further study.

Consider the inverse system of blowups of $\overline{\mathcal{M}}_{g,n}$ along boundary strata. The space of stable maps to an unparameterized $\mathbb{P}^1$ — relative to 0 and ∞ — lifts to a system of classes on the blowups, together called the logarithmic double ramification cycle, compatible under pushforward. The toric contact cycles are not products of double ramification cycles on $\overline{\mathcal{M}}_{g,n}$, but are products on sufficiently refined blowups. On such a blowup

$$\overline{\mathcal{M}}'_{g,n} \to \overline{\mathcal{M}}_{g,n}$$

the lifted cycles can be understood as *virtual strict transforms* of the double ramification cycles. We control virtual strict transforms and calculate them algorithmically in terms of the standard $\overline{\mathcal{M}}_{g,n}$ package — tautological classes, strata of the double ramification cycle, and normal bundles to the exceptional strata. Fulton computes ordinary strict transforms via Segre classes, and our virtual strict transforms are calculated via virtual Segre classes. The latter are virtual pullbacks of Segre classes defined in a universal situation. This type of analysis is expected to occur frequently in nature; see Section 1.6.

1.4. ***Piecewise polynomials.*** When working with cycles on the inverse system of toric compactifications of a fixed torus, the intersection theory can be understood in combinatorial terms, and this is presented in Section 2. The generalization to logarithmic schemes follows this path.

The first observation is that logarithmic schemes come equipped with a natural collection of cohomology classes. An *Artin cone* is the stack quotient of an affine toric variety by its dense torus and an *Artin fan* is a logarithmic algebraic stack that is logarithmically étale over a point and admits a strict cover by Artin cones. Under mild assumptions, every logarithmic scheme Y admits a strict morphism to an Artin fan [6, Section 3]. If A is an Artin fan there is an associated *tropicalization*, a stack over the category of cone complexes, denoted Σ and constructed in [18].

Theorem B. *The operational Chow cohomology group of* A *with rational coefficients is canonically isomorphic to the ring of rational piecewise polynomial functions on* Σ.

If A is a global quotient of a toric variety by its dense torus, this identification was established by Payne [52]. If, in addition, the stack A is smooth, the result is due to Brion [16].[2] The result equips every logarithmic scheme or stack with tautological cohomology classes from combinatorics.

1.5. ***Logarithmic pullbacks.*** In Section 3.6 we introduce a general notion of pullback for a logarithmically flat scheme V mapping to Y, along logarithmic local complete intersection morphism $X \to Y$ between

[2] In fact, the results of Brion and Payne hold integrally. The results of Brion and Payne are about equivariant Chow cohomology groups of toric varieties; these are, by design, the Chow groups of the quotients stacks described.

logarithmically flat schemes. We refer the reader to the main text for complete formal statements. For the moment, we note two special cases in which there is a well-behaved pullback:

(i) If $X \to Y$ is a logarithmic blowup along a regularly embedded center, the pullback is the class of the strict transform.

(ii) If $X \to Y$ is the diagonal morphism of a logarithmically smooth scheme, the pullback gives rise to an intersection product that is calculated by passing to a blowup of Y and intersecting strict transforms.

The most important part of this framework is calculating the class of the strict transform of a cycle under a blowup. With a supply of cohomology classes from piecewise polynomials, we calculate strict transforms in logarithmic geometry. Let Y be a proper and logarithmically smooth scheme over a point, let $Y' \to Y$ be a proper, birational, logarithmically smooth morphism. Let

$$V \to Y$$

be a morphism from a logarithmically flat scheme. Consider a blowup $\widetilde{Y} \to Y$ at a regularly embedded stratum X in Y with exceptional $\widetilde{X}$. Let $[V]^{\mathsf{log}}$ be the logarithmic pullback of V and $[V]^{\mathsf{sch}}$ denote the ordinary pullback along the blowup. The situation is summarized by the diagram

$$\begin{tikzcd} V^{\mathsf{log}} \arrow[r] & V^{\mathsf{sch}} \arrow[r] \arrow[d] & V \arrow[d] \\ & \widetilde{Y} \arrow[r] & Y \end{tikzcd}$$

Theorem C. *The difference between the classes* $[V]^{\mathsf{sch}}$ *and* $[V]^{\mathsf{log}}$ *can be calculated using the standard operations of intersection theory and*

(i) *the Chow homology classes associated to strata of* V *decorated by piecewise polynomial functions*,

(ii) *the piecewise polynomial on* $\widetilde{Y}$ *corresponding to the excess normal bundle of* $\widetilde{X}$.

1.6. ***Related problems.*** Toric contact cycles illustrate a method to understand a simple underlying geometry: natural intersection products on open moduli problems only extend correctly to sufficiently fine logarithmic compactifications, and they are related to naive extensions by blowup formulas. We list a few instances:

(i) Given logarithmically smooth X and Y, one can consider virtual classes of spaces of maps to X and Y, and to the product $X \times Y$. After push forward to an appropriate blowup of the moduli space of curves, the three classes are related by a product formula. They are related to the corresponding classes on $\overline{\mathcal{M}}_{g,n}$ by virtual strict transforms; see [53].

(ii) Given a smooth projective variety X and transverse, smooth divisors D and E, one considers virtual classes of logarithmic maps to (X, D), (X, E), and $(X, D+E)$. In genus 0 and in the presence of positivity conditions, the classes are again related by strict transform operations [48; 49].

(iii) The degeneration formula relates the virtual class of the space of stable maps on the general fiber of a simple normal crossings degeneration to the classes of stable maps to components on the central fiber. The formulation is again in terms of virtual strict transforms of products of spaces of stable maps with appropriate diagonal classes [55, Theorem B].

1.7. ***Past and parallel developments.*** Independent of logarithmic geometry, two papers have served as inspiration. The direct limit of operational Chow rings of toric blowups of a toric variety was completely described by Fulton and Sturmfels [25]. Aluffi's work on modification systems also played a significant role in our formulation of this problem [8]. Intersection theory within logarithmic geometry has attracted significant recent attention. Herr constructs a logarithmic Gysin morphism via the logarithmic normal cone, and building on this, Barrott describes a logarithmic bivariant formalism [14; 29]; see also [15]. Their work has influenced our study, but our goals are orthogonal to theirs. Our methods explain precisely how logarithmic intersection differs from standard intersection on a model, so that existing calculations can be bootstrapped to new ones.

We are motivated by geometric phenomena that have recently come to light, particularly by work in Abel–Jacobi theory and Gromov–Witten theory [32; 55]. The ideas here are closely related to and informed by the strategy followed in [49]. A related direction is work by Pandharipande, Schmitt, and the first author on the top Chern class of the Hodge bundle on blowups of $\overline{\mathcal{M}}_{g,n}$ [46]. Piecewise polynomials appear here, but the strict transform plays no role in this work, and the *formula* for the top Chern class of the Hodge bundle becomes the object of study.

Remark 1.7.1. The proposal of computing logarithmic intersections via the ring of piecewise polynomials and blowup formulas was presented by the second author at the ETH Algebraic Geometry and Moduli seminar in 2020 [54]. In the intervening time, piecewise polynomials have appeared in nearby contexts [31; 46]. The ring appears to be a natural and interesting invariant of a logarithmic scheme.

Remark 1.7.2. During the preparation of this paper, we were informed of work of Holmes and Schwarz, who prove that the toric contact cycles are tautological based on considerations in logarithmic Abel–Jacobi theory [31]. Their approach is based on new invariance properties of the logarithmic double ramification cycle. Ours is a blunt tool, and does not uncover or rely on new geometry of the cycle, but we hope this may give it broad applicability, as noted above.

Remark 1.7.3 (recent progress). A preprint of the present paper first appeared in July 2021. In the intervening years, additional progress has been made in this area. By using methods from the theory of compactified Jacobians, Holmes, Pandharipande, Pixton, Schmitt, and the first author have made additional progress on toric contact cycles [33]. Precisely, they give a formula for the virtual strict transform above, i.e., the "logarithmic double ramification cycle" in terms of certain basic classes, including the "piecewise polynomial functions" studied here. A different generalization of Theorem A, involving Brill–Noether classes, and also using compactified Jacobian techniques, is established in [43]. However, the methods in the present paper are applicable more broadly. For example, the blueprint here should also show that

Gromov–Witten cycles of products of logarithmically smooth curves always lie in the tautological ring, bootstrapping [23; 34]. In a different direction, logarithmic Gromov–Witten classes of toric varieties with their canonical logarithmic structure, including insertions, are shown to lie in the tautological ring in [56], and intersection numbers against the logarithmic double ramification cycle are studied in [19].

Conventions and terminology. We work over an algebraically closed field of characteristic zero. All schemes and stacks will be of finite type unless otherwise stated. Logarithmic structures will always be fine, and sometimes saturated; in the latter case this will be explicitly stated. A logarithmic scheme X is *tropically smooth* if the characteristic monoid at any point of X is free. Terminology such as "logarithmically flat logarithmic scheme" is shortened to "logarithmically flat scheme", etc. A *logarithmic modification* $\mathcal{X}' \to \mathcal{X}$ of logarithmically smooth stacks is a proper and birational logarithmically étale map, which include logarithmic blowups and generalized root constructions; a logarithmic modification of a general X is the pullback of such under a strict morphism $X \to \mathcal{X}$. A logarithmically smooth morphism is *weakly semistable* if it is flat with reduced fibers and *strongly semistable* or *semistable* if it is weakly semistable with smooth source and target.

2. Intersections on toric varieties

We motivate the core ideas — logarithmic flatness, the ring of piecewise polynomials, the blowup formula, and Aluffi's Segre formula — in the toric setting. The route here is not the most efficient analysis of the toric problem, but is parallel to the moduli calculation that appears later.

Let Y be a complete toric variety and $V \hookrightarrow Y$ be an irreducible subvariety that meets the dense torus of Y. Let V° be the intersection with the torus. Suppose $f : \widetilde{Y} \to Y$ is a blowup of Y at a regularly embedded stratum X. We describe the class of the closure of $f^{-1}(V^\circ)$, or equivalently, the strict transform of V. The strict transform will be denoted $V^\dagger$. The upshot will be that if the class of V in Y is known, the strict transform can be calculated if certain "decorated strata classes" of V are known. The latter are pushforwards to X of Chow operators applied to the fundamental classes of strata of V.

2.1. ***Setup.*** A reformulation will be convenient. Fix a toric proper birational morphism $Y' \to Y$, with dense torus T, and choose a subvariety V on Y' that is transverse in the sense that

$$\text{the multiplication map } V \times T \to Y' \text{ is flat.}$$

Equivalently, the map $V \to [Y'/T]$ is flat. In later sections, this will become the condition that V is logarithmically flat. As a consequence, if Y' is blown up, the strict and total transforms of V under that blowup coincide.

Remark 2.1.1 (platification). Given a subvariety of a torus, its closure in any sufficiently fine toric compactification is transverse in the above sense by a theorem of Tevelev [58].

The setup is summarized by the basic blowup diagram

$$\begin{tikzcd} V^{\dagger} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & \widetilde{V} \arrow[r] \arrow[d, "\nu"'] \arrow[dr, phantom, "\square"] & V \arrow[d] \\ \widetilde{Y}^{\dagger'} \arrow[r] & \widetilde{Y}' \arrow[r, "\varphi"] \arrow[d, "\nu"'] \arrow[dr, phantom, "\square"] & Y' \arrow[d, "\pi"] \\ & \widetilde{Y} \arrow[r, "f"'] & Y \end{tikzcd} \qquad \text{with } \widetilde{Y} = \mathsf{Bl}_X Y.$$

The map f is the blowup of the regularly embedded X. The space $\widetilde{Y}^{\dagger'}$ is the strict transform of Y', obtained by pulling back ideal of X and blowing it up. The pullback ideal is monomial, so the blowup map

$$\widetilde{Y}^{\dagger'} \to Y'$$

is equivariant. It is given by a subdivision of the fan of Y'. The remainder of the diagram is defined by pullback.

Problem 2.1.2. Calculate the difference in the Chow group of $\widetilde{Y}^{\dagger'}$ between the pushforward of $V^{\dagger}$ and the pushforward of the class $f^{!}[V]$ to $\widetilde{Y}^{\dagger'}$.

Since V is flat over $[Y'/T]$ this is equivalent to the one stated at the beginning of the section.

2.2. *Fulton.* We use Fulton's refined blowup formula [24, Example 6.7.1],[3] and first apply this to the middle row of the diagram above. We remind the reader that *refined* in this context means that the pullback along $\widetilde{Y} \to Y$ is defined for an arbitrary variety equipped with a map to Y, rather than a cycle on Y, and produces a class supported on the fiber product.

Fulton's formula involves two ingredients — a total Segre class and a total Chern class. The Segre class term is $s(X', Y')$, where X' is the pullback of the center X along the toric morphism π. The Chern class term is that of the excess normal bundle $\mathbb{E}$ on the exceptional divisor $\widetilde{X} \to X$. The excess normal bundle on a blowup is the quotient of the pullback of the normal bundle $N_{X/Y}$ by the normal bundle of the exceptional divisor $\widetilde{X}$. We introduce notation for the exceptionals:

$$g : \widetilde{X} \to X, \quad g' : \widetilde{X}' \to X', \quad j : \widetilde{X} \hookrightarrow \widetilde{Y}, \quad j' : \widetilde{X}' \hookrightarrow \widetilde{Y}'.$$

Theorem 2.2.1 (Fulton). *There is an equality of classes in the Chow group of $\widetilde{Y}'$ given by*

$$f^{!}[Y'] - [\widetilde{Y}^{\dagger'}] = j'_{\star}\{c(\mathbb{E}) \cap g'^{\star} s(X', Y')\}_{\mathsf{exp}},$$

where the right-hand side takes the expected dimensional piece of the intersection product.

The analogue for V certainly holds, and we come to it momentarily.

[3] There is a minor typographical error in the statement of this formula in Fulton's text. The stated formula is correct, but the formula is an equality in the Chow group of $\widetilde{Y}'$ rather than $\widetilde{X}'$ as stated in the line after the displayed equation. Our notation is consistent with Fulton's.

2.3. ***Chow.*** We review intersection theory on complete toric varieties. Let Y be a toric variety with fan Σ and dense torus T. Given a cone σ in Σ, there is a distinguished set of linear functions defined on σ, identified with the character lattice of the dense torus. Correspondingly, there is a well-defined notion of *polynomial function* from σ to $\mathbb{R}$.

Definition 2.3.1. A *piecewise polynomial function on* Σ is a continuous function

$$f : |\Sigma| \to \mathbb{R}$$

whose restriction to any cones of Σ is polynomial. Let $\mathsf{PP}(\Sigma)$ be the ring of such functions.

Results of Brion for Σ smooth, and Payne in general, give the ring geometric meaning [16; 52].

Theorem 2.3.2 (Brion/Payne). *The equivariant Chow cohomology of a toric variety Y is naturally isomorphic to the ring of piecewise polynomial functions on Σ.*

The theorem holds for arbitrary toric Y, but the nonequivariant Chow cohomology is more delicate. An answer is known in the complete case. Let $\Sigma^{(k)}$ be the cones in Σ of codimension k.

Definition 2.3.3. An integer-valued function c on $\Sigma^{(k)}$ is *balanced* if the relation

$$\sum_{\sigma \in \Sigma^{(k)} : \sigma \supset \tau} \langle u, n_{\sigma,\tau} \rangle \cdot c(\sigma) = 0$$

is satisfied for all cones τ of dimension $k+1$, and the vector $n_{\sigma,\tau}$ is the generator of the lattice of σ relative to that of τ. A balanced function of this form is a *Minkowski weight* of codimension k.

The Minkowski weights of a fixed codimension form a group. The direct sum of the groups admits a ring structure, described in [25]. Fulton and Sturmfels prove the following theorem.

Theorem 2.3.4 (Fulton–Sturmfels). *The operational Chow cohomology ring of Y is naturally isomorphic to the ring* $\mathsf{MW}(\Sigma)$ *of Minkowski weights on* Σ.

The two theorems are connected. Katz and Payne describe the ring map

$$\mathsf{PP}(\Sigma) \to \mathsf{MW}(\Sigma)$$

using localization in Chow cohomology in purely combinatorial terms [37].

2.4. ***Chern.*** For concreteness, we assume that Y is *smooth*. Recall the blowup diagram

$$\begin{array}{ccc} \widetilde{X} & \longrightarrow & \widetilde{Y} \\ \downarrow & \square & \downarrow \\ X & \longrightarrow & Y \end{array}$$

Let σ be the cone corresponding to X and let $\widetilde{\sigma}$ be the additional ray in the fan of $\widetilde{Y}$. Enumerate the rays of σ and consider the piecewise linear function

$$\ell_i : \Sigma \to \mathbb{R}$$

that has slope 1 along the i-th ray of σ and slope zero on all other rays. The function

$$\tilde{\ell} := \min_i \{\ell_i\}$$

is piecewise linear on the stellar subdivision $\widetilde{\Sigma}$ of Σ along σ.

Lemma 2.4.1. *The total Chern class of the excess normal bundle of $\widetilde{X}$ is the image in the Chow cohomology of $\widetilde{X}$ of the following piecewise polynomial class on $\widetilde{Y}$:*

$$c(\mathbb{E}) = \frac{\prod_i (1+\ell_i)}{1+\tilde{\ell}}.$$

Precisely, the denominator is expanded as a formal power series, and its image in the nonequivariant Chow cohomology of $\widetilde{Y}$, and the Gysin pullback to $\widetilde{X}$ is equal to the required Chern class.

Proof. The excess normal bundle is the quotient of the normal bundle of the center, pulled back to $\widetilde{X}$, by the natural relative tautological bundle on $\widetilde{X}$, when viewed as a projective bundle. Since the center of the blowup is the torus invariant subvariety dual to σ, and the ambient toric variety is smooth, the subvariety is a complete intersection. The normal bundle is therefore given the numerator (after interpreting each ℓ_i as the Chern class associated to the divisor of the i-th ray. Similarly, by hyperplane bundle is the line bundle on $\widetilde{Y}$ of the exceptional $\widetilde{X}$, restricted to $\widetilde{X}$. By multiplicativity of the total Chern class and functoriality of the identification with piecewise polynomials, the lemma follows. □

2.5. *Segre.* We have $Y' \to Y$ a proper birational map and $V \hookrightarrow Y'$ a subvariety that is flat over $[Y'/T]$. Our goals require an understanding of the two classes

$$s(X', Y') \quad \text{and} \quad s(X' \cap V, V).$$

The subscheme X' is the vanishing locus of a monomial ideal sheaf on Y' since $Y' \to Y$ is a monomial map. Similarly, the intersection $X' \cap V$ is monomial on V with respect to the Cartier divisors on V given by the intersections with V of the toric divisors on Y'.

Fix a pure scheme S equipped with n Cartier divisors $D_1, \dots, D_n$ with regular crossings. They are said to have *regular crossings* if the local defining equations of these divisors form regular sequences. A monomial subscheme $Z \subset S$ is determined by a finite collection of lattice points $q_1, \dots, q_r$ in $\mathbb{Z}^n_{\geq 0}$. Each q_i determines a reducible hypersurface supported on the D_i, and Z is their intersection. A connected intersection of divisors contained in Z will is called a *stratum* in Z.

A combinatorial manifestation of the normal cone is the *Newton region* of Z. It is the complement in $\mathbb{R}^n_{\geq 0}$ of the convex hull of the union of the positive orthants centered at $q_1, \dots, q_r$. The Segre classes are computed by a beautiful formula of Aluffi [10]. We require the following consequence.

Theorem 2.5.1 (Aluffi). *There exists a universal formula for the Segre class of a subscheme $Z \subset S$ that is monomial with respect to n Cartier divisors with regular crossings, depending only on the Newton region of the subscheme. If $L_1, \dots, L_n$ denote the line bundles associated to these Cartier divisors, the Segre class is equal to a sum of terms, with each term equal to a polynomial in the Chern classes of $L_1, \dots, L_n$ applied to strata in Z.*

A word on the meaning of "universal" here: for a monomial subscheme on a regular crossings pair, the Segre class is calculated by the *same* formal expression — a sum of strata that are contained in the subscheme, decorated by Chern class operators, depending only on the Newton region.

2.6. ***Synthesis.*** We still have the running assumption that Y is smooth. We can put the pieces together now. The difference between the strict and total transforms of V can be computed from the pushforwards of strata of V decorated by polynomials in the Chern operators.

Fulton's formula asserts that the difference between the strict transform and the Gysin pullback along a blowup is a class that is supported on the exceptional locus. This part of the diagram is reproduced below, with X' denoting the scheme-theoretic preimage of the blowup center in Y':

$$\begin{tikzcd} \widetilde{X}' \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & X' \arrow[d] \\ \widetilde{X} \arrow[r] & X \end{tikzcd}$$

The horizontal maps are projective bundles. The Chern class of the excess bundle $\mathbb{E}$ is supported on $\widetilde{X}$. The term $s(X', Y')$ is handled by Aluffi's formula. Recall that $V \hookrightarrow Y'$ is the cycle of interest.

Lemma 2.6.1. *The Segre class $s(V \cap X', V)$, after pushforward to the Chow group of X', is computed as*

$$s(V \cap X', V) = s(X', Y')(V),$$

where the right-hand side is interpreted as follows. The Segre class $s(X', Y')$ according to Aluffi's universal formula is a sum of terms — each term is a stratum of X' times a polynomial in Chern classes of line bundles on Y'. Define $s(X', Y')(V)$ as the same formal sum, where the strata terms are replaced by their intersection with V, and the Chern classes are pulled back to these strata.

Proof. By our transversality hypothesis for $V \hookrightarrow Y'$, the pullback of the divisors on Y' exhibit V as a regular crossings pair. The Segre class $s(V \cap X', V)$ can be computed by Aluffi's formula, so the lemma follows by combining the transversality of V and Y' with naturality of Chern classes. □

The inclusion $V \cap X' \to X'$ may not admit an obvious Gysin pullback. A more sophisticated view is that since the Segre formula is universal, it holds for monomial substacks of $[Y/T]$; since Y and V are flat this stack, the pullback preserves the Segre classes. We soon adopt this view, but when the definitions are unwound, this does not say anything more than the stated procedure.

We examine the blowup over the center:

$$\begin{tikzcd} \widetilde{V \cap X'} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] \arrow[dd, bend right=60, "q"'] & V \cap X' \arrow[d] \arrow[dd, bend left=60, "p"] \\ \widetilde{X}' \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & X' \arrow[d] \\ \widetilde{X} \arrow[r, "h"] & X \end{tikzcd}$$

The composite vertical maps are proper. The vector bundle $\mathbb{E}$ is pulled back along q. By the projection formula for $c(\mathbb{E})$ and compatibility of pullback and pushforward applied to the Segre class, we obtain the following equality of class in the Chow group of $\widetilde{X}$:

$$q_\star(c(\mathbb{E}) \cap g'^\star s(X' \cap V, V)) = c(\mathbb{E}) \cap h^\star p_\star s(V \cap X', V).$$

We explain the utility of the formula. The spaces in question come equipped with a collection of Cartier divisors, i.e., the boundary divisors. A *standard expression* on $V \cap X'$ is a polynomial in Chern classes of the boundary divisors of V applied to the class of a stratum of V.

Corollary 2.6.2. *The difference of classes*

$$f^\star[V] - [V^\dagger]$$

in the Chow group of $\widetilde{Y}$ is the pushforward of an element in the Chow group of $\widetilde{X}$. The element can be calculated by evaluating a standard expression on $V \cap X'$, pushing forward along p, pulling back along h, and applying the total Chern class of the excess normal bundle.

The expression above is "universal" in a similar sense of Aluffi's formula itself. In this toric setting, the content is that the difference between strict and total transforms can be understood completely in terms of a universal formula, whose input is the pushforward to X of classes of strata of V decorated by Chern classes, also known as *normally decorated strata classes* [46].

2.7. ***Calculus.*** The calculation scheme we have laid out is essentially elementary. The rest of this paper is devoted to implementing in the context of logarithmic schemes, moduli spaces, and virtual classes. We close out the toric discussion by explaining how using elementary toric intersection theory, the ingredients in the universal expression above can be pleasantly written in terms of Minkowski weights and piecewise polynomials. We stress again that Y is smooth.

2.7.1. *The cycle and its strata.* The cycle V in Y' intersects the boundary strata of Y' properly and defines a cohomology class $[V]$ in $\mathsf{CH}^\star(Y')$. It determines a Minkowski weight on $\Sigma_{Y'}$.

Similarly, the strata of V lie in the boundary strata of Y'. If $W \subset V$ is a closed stratum, it is contained in a closed stratum $Z' \subset Y'$. The stratum Z' is a toric variety and has an associated fan. The subvariety W defines a Chow cohomology class in $\mathsf{CH}^\star(Z')$. These Chow cohomology classes are expressed as Minkowski weights on the fan of Z'.

2.7.2. *Normal decoration.* Let $W \subset V$ be a stratum. The intersection procedure requires us to decorate this stratum with polynomials in Chern roots of the normal bundle of W before pushforward.

Assume $W \subset Z'$. The normal bundle splits, and in practice Chern roots can be given in three ways. The first is as a combination of divisorial strata of Z'. In this case, each of these strata is itself a toric variety and the Minkowski weight associated to W described above gives a Minkowski weight on this smaller stratum. Repeating this process for all the divisors, we obtain a Minkowski weight corresponding to a decorated stratum of V.

The second way that the normal roots are given is by a piecewise linear functions. In this case, we convert the piecewise linear function η into a integer combination of boundary divisors. The coefficient of a divisor corresponding to a ray ρ is the slope of η along this ray. We then repeat the procedure above.

The final way in which the normal roots are provided is as a product of codimension-1 Minkowski weights. In this case, the decorated stratum of V is obtained by the fan displacement rule, intersecting the cycle defined by the stratum of V with these codimension-1 weights [25, Section 4].

2.7.3. *Pushforward.* Given an equivariant proper toric morphism $Z' \to Z$ and a Minkowski weight on Z', its pushforward to Z is calculated via the projection formula. The Minkowski weight records the degree of the operator applied to all boundary strata. In order to compute the pushforward, it is therefore sufficient to calculate, with multiplicity, an expression in boundary strata for the pullback to Z' of all boundary strata of Z. Once this is done, the Minkowski weight may be evaluated on these expressions, and this determines the pushforward.

2.7.4. *Pullback.* The morphism $\widetilde{X} \to X$ is a projective bundle and this is visible at the fan level. An explicit description of the fan of a projective bundle may be found in [22, Chapter VII] . The pullback itself is computed by Fulton and Sturmfels' formula [25, Proposition 3.7].

2.8. ***A simple example.*** Let us do the first nontrivial example of a strict transform calculation to illustrate the nature of the calculation. We hope this will allow an interested reader to work through the different pieces of the formula.

Let $Y = \mathbb{P}^3$ and X be one of its four torus invariant points. Let $V \hookrightarrow Y$ be a line passing through X. We maintain the notation from the discussion above. Let us calculate the class of $V^\dagger$ in the scheme above.

According to the scheme, in order to study the class $f^\star[V] - [V^\dagger]$, we need to calculate the excess bundle $\mathbb{E}$. The bundle $\mathbb{E}$ is a quotient of the pullback of a bundle on X, to $\widetilde{X}$. Therefore the first piece is just the normal bundle of X. Of course, in this case X is a point, so the classes ℓ_i are all 0. However, this is a low-dimensional accident.

The point X is an intersection of three torus invariant hyperplanes. Each is associated to a ray, and let the associated piecewise linear functions be ℓ_1, ℓ_2 and ℓ_3. Since X is a complete intersection, the total Chern class of the normal bundle is given by

$$\prod_{i=1}^{3}(1+\ell_i),$$

turned into a Chow cohomology class on Y and then restricted to X. The second piece of the excess normal bundle is the class of $\widetilde{X}$ on $\widetilde{Y}$ *restricted* to $\widetilde{X}$. Let $\tilde{\ell}$ denote the piecewise linear function associated to $\widetilde{X}$ on $\widetilde{Y}$. We can expand it as a power series $1 - \ell + \ell^2 - \cdots$ and then restrict to $\widetilde{X}$. We are interested in the restriction of the linear term in this expansion, by dimension considerations.

The final term is the Segre class of $X \cap V$ in V. Typically this is where we would plug in Aluffi's formula, but in this case the term is just 1. In general, these Segre calculations are essentially independent from the Chern class calculations, so we refer the reader to the discussion in [9, Examples 1.1–1.4].

Putting the pieces together and extracting the degree-2 term, we see that the difference we were trying to calculate is $-\ell$, restricted to $\widetilde{X}$, and then pushed forward to $\widetilde{Y}$. We find that there difference is exactly the class of a line in the exceptional divisor, as expected from elementary calculations.

3. Intersections on logarithmic schemes

In the remainder of the paper, we work with Chow groups and their operational Chow rings with rational coefficients, and assume that all logarithmic stacks have locally connected strata, to guarantee the existence of Artin fans. We write $X \times_Y^{\mathsf{log}} Z$ for the fiber product in the category of fine and saturated logarithmic schemes.

We adapt the toric techniques to the context of logarithmic schemes. Our motivating example, the double ramification cycle, requires an additional ingredient — the virtual class — see Section 4.

3.1. ***The plan.*** We examine the part of intersection theory that behaves well with respect to logarithmic blowups. First, we examine intersections on logarithmically smooth schemes, and use this to model the general formalism. The relevance of Fulton's blowup formula will be apparent.

In order to build operations in intersection theory based on logarithmic fiber products, Fulton's formula demands the following: a notion of logarithmic cycle whose logarithmic pullback is calculated by strict transforms, a supply of Chow cohomology operators on a logarithmic scheme, an understanding of the normal cones of monomial ideals, and a class of morphisms along which we can pullback. The cycles are provided by maps between logarithmically flat schemes, the cohomology operators by the ring of piecewise polynomials on the tropicalization, the normal cones by Aluffi's formula for Segre classes of monomial subschemes, and the morphisms by logarithmic local complete intersections.

3.2. ***Basic model: logarithmically nonsingular schemes.*** Let Y be a logarithmically smooth scheme. A *logarithmic cycle* on Y is a pair (V, α) where V is a logarithmically flat scheme and

$$\alpha : V \to Y$$

is a proper monomorphism of logarithmic schemes. Equivalently, there exists a logarithmic modification $Y' \to Y$ such that the fine logarithmic pullback $V' \to Y'$ is a *strict* closed embedding. The notion was considered in an early version of Barrott's paper [14]. When the morphism α is clear from the context, we drop it from the notation.

Definition 3.2.1. Let (V, α) and (W, β) be logarithmic cycles on Y; their *logarithmic intersection* is defined as follows. Let $\pi : Y' \to Y$ be a logarithmic modification of Y with Y' smooth such that, denoting the fine and saturated pullbacks of V and W by primes, the morphisms

$$\alpha' : V' \to Y', \quad \beta' : W' \to Y'$$

are strict. Then define

$$V \cdot_{\mathsf{log}} W := \pi_\star(V' \cdot W'),$$

where the product on the right-hand side is the standard intersection product on the smooth Y'.

Proposition 3.2.2. *Let (V, α) and (W, β) be logarithmic cycles on Y. The class of the logarithmic intersection $V \cdot_{\log} W$ is independent of all choices.*

Proof. By toroidal weak factorization [3], it suffices to take a blowup $Y' \to Y$ such that

$$\alpha' : V' \to Y', \quad \beta' : W' \to Y'$$

are flat and compare it to a further logarithmic blowup $Y'' \to Y$ along a smooth center. We do this by reduction to the diagonal. The blowup gives rise to a local complete intersection morphism

$$Y'' \times Y'' \to Y' \times Y'$$

and the diagonal class on the left pushes forward to the diagonal on the right. The intersection product can be defined in two ways. First, we intersect $V'' \times W''$ with the diagonal in $Y'' \times Y''$ and push it forward. Second, we intersect $V' \times W'$ with the diagonal in $Y' \times Y'$. The cycle $V' \times W'$ in $Y' \times Y'$ is strict and therefore meets the strata of $X' \times X'$ in the expected dimension. Therefore the fine and saturated logarithmic pullback $V'' \times W''$ has the same cycle class as the total transform. Therefore the two definitions above agree by the projection formula. □

Remark 3.2.3 (logarithmic flatness). Logarithmic flatness over the ground appears frequently in our setup. There are two practical reasons. First, if Y is logarithmically flat, the locus where the logarithmic structure is nontrivial has positive codimension. Second, if Y is logarithmically flat, a logarithmic blowup can be understood as a birational modification of Y. It may be possible to generalize this, but we have made no attempt to do so. In the final section of the paper, we work with a cycle, namely the double ramification cycle, that is not logarithmically flat, but we factorize out the failure of logarithmic flatness into a strict map that is handled separately, and reduce to the logarithmically flat situation. The reason this is possible is because the double ramification is still, "virtually logarithmically flat". The same is true for all moduli spaces of logarithmic stable maps, and the methods are easily adaptable to that setting.

Remark 3.2.4 (strict transforms). Fix logarithmic cycle $\alpha : V \to Y$. Logarithmic flatness implies that the Chow homology $\alpha_\star[V]$ is determined by the image of the interior of V, where the logarithmic structure is trivial. If $Y' \to Y$ is a projective logarithmic modification, the fine logarithmic pullback V' is, by definition, obtained by pulling back the monomial ideal whose blowup defines $Y' \to Y$ to V and blowing it up there. Since V is logarithmically flat, this has the same Chow homology class in Y' as the closure of its interior. It coincides with the class of the strict transform.

3.3. ***Chow operators: piecewise polynomials.*** Let Y be a logarithmic scheme. There are two combinatorial constructions associated to Y: its cone complex[4] Σ_Y and its Artin fan A_Y. The cone complex is a classical construction from toroidal geometry originating in [38] and developed further by many others [4; 18; 59; 60]. The cone complex has a realization within logarithmic algebraic stacks via Artin fans

[4]In standard terminology, a general logarithmic scheme only has a cone stack or generalized cone complex rather than a cone complex, where self-gluings along faces and quotients by automorphisms may be allowed. We will use the terminology "cone complex" with the allowance of the generalized situation.

[2; 51]; see also the survey [5]. The output of the equivalence is that the cone complex and the Artin fan have presentations as the same colimit over all points y in Y of simpler pieces:

$$\Sigma_Y = \varinjlim \sigma_y \quad \text{and} \quad \mathsf{A}_Y = \varinjlim \mathsf{A}_y,$$

where σ_y are polyhedral cones, and $\mathsf{A}_y = \mathrm{Spec}(k[P_y])/\mathrm{Spec}(k[P_y^{\mathrm{gp}}])$ for P_y the dual monoid of σ_y. The first colimit may be taken in generalized cone complexes [4] or, 2-categorically, in the category of stacks over cone complexes [18]. The latter colimit is in logarithmic algebraic stacks. Both colimits are taken over points of Y, with arrows given by generization maps. Define

$$\mathrm{PP}(\Sigma_Y) := \varprojlim \mathrm{PP}(\sigma_y).$$

We pause to flag a potential point of confusion. A cone complex Σ_Y can be presented as a colimit of cones in multiple ways; for example, a cone can be presented as a trivial colimit, or as a colimit of all of its faces. Therefore, a priori, there is a question concerning the independence of presentation of Σ_Y in the definition, and while there is essentially a best presentation [4, Section 2.6], the issue will be dodged — the independence will follow from the main result of this section. Given a presentation, we have maps

$$\mathsf{CH}^\star_{\mathrm{op}}(\mathsf{A}_Y) \to \varprojlim \mathsf{CH}^\star_{\mathrm{op}}(\mathsf{A}_y) \xrightarrow{\sim} \varprojlim \mathrm{PP}(\sigma_y) = \mathrm{PP}(\Sigma_Y).$$

The middle isomorphism is due to Payne [52, Theorem 1]. When A_Y is a global quotient of a toric variety by the dense torus, Payne's result equates the outer two rings. We have the following.

Theorem 3.3.1. *Let Y be a logarithmic stack with Artin fan A_Y. There is a natural isomorphism*

$$\mathsf{CH}^\star_{\mathrm{op}}(\mathsf{A}_Y) = \mathrm{PP}(\Sigma_Y),$$

and therefore a functorial map

$$\mathrm{PP}(\Sigma_Y) \to \mathsf{CH}_{\mathrm{op}}(Y).$$

3.3.1. *Proof of Theorem 3.3.1.* We begin with a result of Bae and Park that the operational Chow rings of Artin stacks (of finite type, that admit stratifications by global quotients) with rational coefficients satisfy Kimura's descent with respect to blowups.[5] A blowup $\mathcal{Y}' \to \mathcal{Y}$ induces two exact sequences:

$$\mathsf{CH}_\star(\mathcal{Y}' \times_{\mathcal{Y}} \mathcal{Y}') \to \mathsf{CH}_\star(\mathcal{Y}') \to \mathsf{CH}_\star(\mathcal{Y}) \to 0, \quad \mathsf{CH}^\star_{\mathrm{op}}(\mathcal{Y}' \times_{\mathcal{Y}} \mathcal{Y}') \leftarrow \mathsf{CH}^\star_{\mathrm{op}}(\mathcal{Y}') \leftarrow \mathsf{CH}^\star_{\mathrm{op}}(\mathcal{Y}) \leftarrow 0.$$

Consequently, an analogue of [39, Theorem 3.2] holds, by cosmetic changes to the proof:

Lemma 3.3.2. *Let $p : \mathcal{Y}' \to \mathcal{Y}$ be a blowup of (finite type, that admit stratification by global quotients) Artin stacks. Let S_i be the irreducible components of the locus in $\mathcal{Y}'$ over which p is not an isomorphism, and $T_i = p^{-1}(S_i)$ their preimages. Then a class $\alpha \in \mathsf{CH}^\star_{\mathrm{op}}(\mathcal{Y}')$ is in the image of $\mathsf{CH}^\star_{\mathrm{op}}(\mathcal{Y})$ if and only if its restriction to each T_i is in the image of the Chow ring of S_i.*

[5] The details of the argument of Bae and Park will appear in forthcoming work [12]. We take the opportunity to thank them for sharing their expertise on these matters.

Proof of Lemma 3.3.2. Let now Y be a logarithmic stack, and A_Y its Artin fan, as in the theorem statement. There is a blowup $\mathsf{A}_{Y'} \to \mathsf{A}_X$ with $\mathsf{A}_{Y'}$ smooth. For smooth Artin fans, we have

$$\mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_{Y'}) = \mathsf{PP}(\Sigma_{Y'})$$

by [46, Theorem 14]. Therefore, Payne's theorem and proof, on the equivariant cohomology of toric varieties, generalizes. We provide the details.

Factor $\mathsf{A}_{Y'} \to \mathsf{A}_Y$ as a sequence of blowups along smooth centers. By induction on the number of blowups necessary to reach from a nonsmooth A_X to a smooth A_Y, we may assume that $\mathsf{A}_{Y'} \to \mathsf{A}_Y$ is a single blowup along a smooth stratum A_S. Let A_T be it's preimage, but we note that despite the notation, it is not an Artin fan but a gerbe over one. From the Kimura descent sequence, an operator $\alpha \in \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_{Y'})$ comes from A_Y if its restriction to A_T comes from A_S. Let η be the open stratum of A_S, corresponding to a cone $\sigma \in \Sigma_Y$. The stratum may have self-intersections or automorphisms. The stabilizer group of this point is $\mathbb{G}_m^n \rtimes G$. Then $\mathbb{G}_m^n \rtimes G$ injects into the automorphism groups of every point of A_S, and the rigidification $\mathsf{A}_S/\mathbb{G}_m^n \rtimes G$ is an Artin fan. By construction, the cone complex of A_S is

$$\mathsf{Star}(\mathsf{A}_S) := \varinjlim_{\sigma \prec \tau \in \Sigma_Y} \tau$$

On the other hand, let $\overline{\tau}$ denote the image of τ in the quotient of the lattice generated by τ by the lattice generated by σ, i.e., the reduced star of τ, defined as

$$\overline{\mathsf{Star}(\mathsf{A}_S)} := \varinjlim_{\sigma \prec \tau \in \Sigma_Y} \overline{\tau}.$$

The cone complex $\overline{\mathsf{Star}(\mathsf{A}_S)}$ then corresponds to the cone complex of the rigidification $\mathsf{A}_S/\mathbb{G}_m^n \rtimes G$, and, by induction on the maximal rank of stabilizer groups, we have an isomorphism

$$\mathsf{PP}(\overline{\mathsf{Star}(\mathsf{A}_S)}) \cong \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_S/\mathbb{G}_m^n \rtimes G).$$

We are working with $\mathbb{Q}$ coefficients, so the finite group G does not contribute either to piecewise polynomials or Chow rings, and since A_S is a gerbe over $\mathsf{A}_S/\mathbb{G}_m^n \rtimes G$, we have

$$\mathsf{PP}(\mathsf{Star}(\mathsf{A}_S)) = \mathsf{PP}(\overline{\mathsf{Star}(\mathsf{A}_S)}) \otimes \mathbb{Q}[\mathbb{Z}^n] \cong \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_S/\mathbb{G}_m^n \rtimes G) \otimes \mathbb{Q}[\mathbb{Z}^n] \cong \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_S).$$

Looking at the map of short exact sequences

$$\begin{tikzcd}
0 \arrow[r] & \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_Y) \arrow[r] \arrow[d] & \mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_{Y'}) \arrow[d] \\
0 \arrow[r] & \mathsf{PP}(\Sigma_Y) \arrow[r] & \mathsf{PP}(\Sigma_{Y'})
\end{tikzcd}$$

and applying the inductive hypothesis,

$$\mathsf{CH}^{\star}_{\mathsf{op}}(\mathsf{A}_{Y'}) = \mathsf{PP}(\Sigma_{Y'}),$$

and the map $\mathsf{CH}^\star_{\mathsf{op}}(\mathsf{A}_Y) \to \mathsf{PP}(\Sigma_Y)$ is injective. On the other hand, let $p \in \mathsf{PP}(\Sigma_Y)$ be a piecewise polynomial. Then p pulls back to a piecewise polynomial on $\Sigma_{Y'}$, i.e., a Chow class on $\mathsf{A}_{Y'}$. By Kimura's sequence, this comes from a class on A_Y if and only if its restriction to A_T comes from A_S. But the pullback of p to A_T comes from a class on the cone complex of A_T, which is a subdivision of $\mathsf{Star}(\mathsf{A}_Y)$, and we've seen that piecewise polynomials on the latter are the same as $\mathsf{CH}^\star_{\mathsf{op}}(\mathsf{A}_S)$. Therefore, the pullback of p to A_T comes from a class on A_S, and thus p is the image of this class in $\mathsf{PP}(\Sigma_X)$. Thus the map $\mathsf{CH}^\star_{\mathsf{op}}(\mathsf{A}_Y) \to \mathsf{PP}(\Sigma_Y)$ is also surjective, and hence an isomorphism. Since we have a canonical map, given by the logarithmic structure, $Y \to \mathsf{A}_Y$, the pullback gives rise to the claimed homomorphism to the operational Chow ring. □

Remark 3.3.3. The piecewise polynomial rings of smooth Artin fans can be exotic. For example, the Chow ring of the stack of expanded smooth pairs is an Artin fan, and its Chow ring has been computed by Oesinghaus, and identified with the ring of quasisymmetric functions [50], which is of significant combinatorial interest. The above result applies to the Chow ring of the stack of expansions of simple normal crossings pairs, constructed in work of Maulik and the second author [41]. The resulting ring is combinatorial but not well understood.

Remark 3.3.4 (K-theory and cobordism). Anderson and Payne [11] and González and Karu [26] describe the equivariant operational K-theory and algebraic cobordism rings of toric varieties as the rings of piecewise exponential functions and piecewise graded power series functions over the Lazard ring. The method of proof is to establish an analogue of the descent sequence. A logarithmic scheme whose Artin fan is a global quotient (which can always be achieved after a blowup) is endowed with tautologically defined classes in these theories. It is plausible that their theorems will generalize to describe the operational theories on general Artin fans. Some logarithmic aspects of K-theory have been explored by Chou, Herr, and Lee [21].

Remark 3.3.5. Holmes and Schwarz define piecewise polynomial functions on logarithmic schemes without using Artin fans [31, Section 3]. The approach is based on the identification between piecewise linear functions on Σ_Y and sections of the characteristic abelian sheaf $\overline{M}_Y^{\mathsf{gp}}$ of Y.

3.4. ***Aluffi's formula and strict transforms.*** Let Y be a logarithmic scheme with structure morphism $\epsilon : M_Y \to \mathcal{O}_Y$. A *monomial subscheme* of Y is a subscheme of X isomorphic to the vanishing of an ideal $\epsilon(I) \subset \mathcal{O}_Y$, for $I \subset M_Y$ an ideal in the sense of monoid theory. Ideals $I \subset M_X$ are in bijection with ideals of $\overline{I} \subset \overline{M}_Y$. An ideal $I \subset M_Y$ is equivalent to the choice of a substack of the Artin fan A_Y. If Y is logarithmically flat, the map $Y \to \mathsf{A}_Y$ is faithfully flat, the ideal $I \subset \overline{M}_Y$ is determined by $\epsilon(I) \subset \mathcal{O}_Y$. We therefore have the following:

Lemma 3.4.1. *Let Y be a logarithmically flat scheme, and $\alpha_Y : Y \to \mathsf{A}_Y$ the map to its Artin fan. The monomial subschemes of Y are precisely the subschemes of the form $\alpha_Y^{-1}(T)$ for T a substack of A_Y.*

Suppose Y is logarithmically flat logarithmic scheme (or stack), and $S = \alpha_Y^{-1}(T)$ a monomial subscheme of Y. The Segre class is preserved under flat base change [24, Proposition 4.2] and we have

$$s(S, Y) = \alpha_X^{\star} s(T, \mathsf{A}_Y)$$

Monomial substacks $T \subset \mathsf{A}_Y$ can be both nonequidimensional and nonreduced. The operational Chow ring of T is typically far from the Chow homology groups. A supply of *homology* classes can be extracted using piecewise polynomials.

Definition 3.4.2. A homology class in $\mathsf{CH}_\star(T)$ is said to *come from piecewise polynomials* if it is a linear combination of classes of the form $\pi_\star(\gamma \cap [R])$ where π is a blowup of A_Y, the class γ is a piecewise polynomial on this blowup, and R is a pure-dimensional substack of the blowup that maps to T. A homology class in $\mathsf{CH}_\star(\alpha_Y^{-1}(T))$ is said to *come from piecewise polynomials* if it is if it the flat pullback of a homology class on T that comes from piecewise polynomials.

We have the following:

Proposition 3.4.3. *The Segre class of a monomial subscheme of a logarithmically flat and tropically smooth scheme comes from piecewise polynomials.*

This lemma is existential and is an immediate consequence of a constructive theorem, namely Aluffi's formula for the Segre class of a monomial subscheme presented in [10]. Aluffi states his formula in terms of *regular crossings* pairs. Let Y be an integral scheme and $D_1, \ldots, D_s$ be divisors on X. They have *regular crossings* if at every point p in the intersection $D_{i_1} \cap \cdots \cap D_{i_j}$, the local equations for the divisors meeting p form regular sequences in the local ring of X at p.

Lemma 3.4.4. *A pair (Y, D) is regular crossings if and only if the tautological induced map*

$$Y \to [\mathbb{A}^s/\mathbb{G}_m^s]$$

is flat.

Proof. The toric Artin stack $[\mathbb{A}^s/\mathbb{G}_m^s]$ with its toric boundary has regular crossings, because it has normal crossings. Regular sequences are preserved by flat pullback, so (Y, D) is regular crossings. Conversely, given a regular crossings pair, since the codomain is smooth, we apply the criterion in [57, Lemma 07DY] to conclude flatness. □

Proof of Proposition 3.4.3 via construction. There is a sequence of logarithmic blowups $p : Y' \to Y$ along smooth centers with $\mathsf{A}_{Y'}$ tropically smooth, monodromy free, and admits a strict morphism to $[\mathbb{A}^s/\mathbb{G}_m^s]$. The statement of Proposition 3.4.3 is known in this case. As Y is logarithmically flat and blowups are logarithmically étale, the space Y' is also logarithmically flat and $Y' \to Y$ is birational. The birational invariance of Segre classes [24, Proposition 4.2] gives

$$s(S, Y) = p_\star s(S', Y')$$

for $S' = p^{-1}(S)$. The result follows. □

We return to logarithmic intersection products. Fix Y logarithmically smooth and logarithmic cycles $V_1, V_2 \to Y$. In order to study the intersection product, we calculate the strict transform of a the cycle under a blowup $Y' \to Y$. It is sufficient to treat the case of a sequence of blowups of Y along smooth centers, and in turn, a single blowup

$$\widetilde{Y} \to Y.$$

Theorem C controls the logarithmic intersection product. It is restated and proved below.

Theorem 3.4.5. *Let $V \to Y$ be a logarithmic cycle and let $\widetilde{Y} \to Y$ be a blowup at a smooth center X. Let $\widetilde{X}$ be the exceptional divisor. The difference between the class of the pullback $[V]^{\mathsf{sch}}$ and the class of the strict transform $[V]^{\mathsf{log}}$ is given by a universal formula involving the standard operations of intersection theory and the following*:

(i) *the Chow homology classes associated to strata of V decorated by piecewise polynomial functions*,

(ii) *the piecewise polynomial on $\widetilde{Y}$ corresponding to the excess normal bundle of $\widetilde{X}$.*

Proof. The proof is parallel to the toric case. Let $\mathbb{E}$ denote the excess normal bundle of the exceptional divisor $\widetilde{X}$ over the center X of the blowup. The difference between the strict and total transforms is determined by Fulton's formula, and is the pushforward of a class on $\widetilde{X}$. The Segre class of $X \cap V$ in V comes from piecewise polynomials. This accounts for the first term. Since the maps $Y' \to \mathsf{A}_{Y'}$ and $Y \to \mathsf{A}_Y$ are smooth, the class $c(\mathbb{E})$ is pulled back from $\mathsf{A}_{Y'}$, and therefore comes from piecewise polynomials as well. $\square$

3.5. ***Interlude: Gysin pullback.*** We have discussed an intersection product on nonsingular logarithmic schemes based on strict transforms. We turn to logarithmic local complete intersections.

We require generalities on Gysin pullbacks induced by maps on Artin fans. The results appear already in [14; 29]. We rework the formalism to be more transparently connected to our calculations. Fix a morphism $f : X \to Y$ between fine logarithmic schemes. By the *substitute for functoriality* in [2, Section 2.5] this map induces the commutative diagram

$$\begin{tikzcd} X \arrow[r, "\alpha_X"] \arrow[d, "f"'] & \mathsf{A}_{X/Y} \arrow[d, "\mathsf{A}_f"] \\ Y \arrow[r, "\alpha_Y"'] & \mathsf{A}_Y \end{tikzcd}$$

with strict horizontal arrows. The diagram is typically *not* Cartesian so we factorize it:

$$\begin{tikzcd} X \arrow[r, "h"] & \mathsf{A}_{X/Y} \times_{\mathsf{A}_Y} Y \arrow[r, "\alpha'_Y"] \arrow[d, "\mathsf{A}'_f"'] \arrow[dr, phantom, "\square"] & \mathsf{A}_{X/Y} \arrow[d, "\mathsf{A}_f"] \\ & Y \arrow[r, "\alpha_Y"'] & \mathsf{A}_Y \end{tikzcd}$$

with h strict. If we further assume that the X and Y are tropically smooth, the stacks $\mathsf{A}_{X/Y}$ and A_Y are smooth, so the map A_f is lci. Therefore, we have a pullback

$$\mathsf{A}_f^! : \mathsf{CH}_\star(Y) \to \mathsf{CH}_\star(Y \times_{\mathsf{A}_Y} \mathsf{A}_X).$$

Assuming that the map h is lci, we then define

$$h^! \mathsf{A}^!_f : \mathsf{CH}_\star(Y) \to \mathsf{CH}_\star(X).$$

We spell out the situations of interest where h is lci below. The second case will be our case of interest, but other situations may be useful to others:

Case I: If Y is logarithmically flat over the base field, the map α_Y is flat. Therefore the map A'_f, being the flat pullback of an lci map is lci. Therefore, when h is lci, the map f is lci, and the pullback is the ordinary lci pullback $f^!$.

Case II: If X, Y are logarithmically smooth, then, being tropically smooth implies that the underlying schemes of X, Y are smooth and the maps f, h are automatically lci. The pullback is $f^!$, as in case I, without further assumptions. This is our main case of interest.

Case III: The map $f : X \to Y$ is logarithmically flat. Then the map h, though not necessarily lci, is flat, and thus a pullback $h^\star \mathsf{A}^!_f$ still exists. It yields the same pullback as $h^! \mathsf{A}^!_f$ if A_f is also lci.

Case IV: The map f is weakly semistable in the sense of [1; 42]. Then A_f is flat, and thus so is A'_f. Thus, in this case the pullback reduces to $h^! \mathsf{A}'^\star_f$, without assumptions about X, Y being tropically smooth. If f is weakly semistable, the relevant diagram above reduces to the simpler diagram

$$\begin{tikzcd} X \arrow[r, "\alpha_X"] \arrow[d, "f"] & \mathsf{A}_X \arrow[d, "\mathsf{A}_f"] \\ Y \arrow[r] & \mathsf{A}_Y \end{tikzcd}$$

3.6. *Intersection via strict transforms: logarithmic complete intersections.* Olsson [51] has explained that there is no theory of the logarithmic cotangent complex $L^{\log}_{X/Y}$ of a morphism of logarithmic schemes that is perfectly parallel to the classical theory. It is not possible to construct a logarithmic cotangent complex $L^{\log}_{X/Y}$ for any morphism of logarithmic schemes, which at once reduces to the ordinary cotangent complex of a strict map, is represented by the logarithmic cotangent bundle for a logarithmically smooth map, is stable under logarithmically flat maps, and has a distinguished triangle associated to a composition. Thus, there is no "best" possible theory of logarithmic lci maps: there is no class of maps, closed under composition and logarithmically flat pullback, in which every logarithmically smooth morphism is a logarithmically local complete intersection, and the strict logarithmic lci maps are precisely the ordinary lci maps. We work with Olsson's version. However, Olsson proves that for logarithmically flat maps, his version coincides with Gabber's and since we have this assumption throughout, the reader may choose either formalism.

We content ourselves with schemes which are *logarithmically flat* over the base field. The assumption is not sharp, but simplifies the theory and covers the examples of Section 1.6 and more.

Definition 3.6.1. A morphism $f : X \to Y$ of fine log schemes is a *logarithmic local complete intersection morphism* or *log lci* for short, if it factors as $X \to P \to Y$ with $i : X \to P$ a strict regular embedding and $g : P \to Y$ logarithmically smooth.

Lemma 3.6.2. *The log cotangent complex of a log lci morphism $f : X \to Y$ is perfect.*

Proof. Factor $f : X \to Y$ as $i : X \to P$, $g : P \to Y$. As g is logarithmically flat, there is a distinguished triangle of log cotangent complexes (that is, even in Olsson's formulation of the cotangent complex)

$$i^\star L^{\mathsf{log}}_{P/Y} \to L^{\mathsf{log}}_{X/Y} \to L^{\mathsf{log}}_{X/P}.$$

Perfect complexes are closed under taking cones, so we conclude. □

Lemma 3.6.3. *Let $f : X \to Y$ be a log lci map between logarithmically flat schemes. Then, for any logarithmically étale map $a : Y' \to Y$, the base change $f' : X' = X \times^{\mathsf{log}}_Y Y' \to Y'$ is log lci.*

Proof. Factor f as $X \to P \to Y$ into a logarithmically smooth map followed by a regular closed immersion. Since logarithmically smooth morphisms pull back under arbitrary maps, it suffices to prove the statement for $f : X \to Y$ a strict regular immersion. Then, the log fiber product X' coincides with the ordinary fiber product. Since $X \to Y$ is a regular immersion, its normal cone $C_{X/Y}$ is a vector bundle. As X has been assumed to be logarithmically flat, the map $X' \to X$ is birational. Thus, the inclusion of the subscheme $C_{X'/Y'}$ to $a^\star C_{X/Y}$ is closed and birational, thus an isomorphism. Therefore $C_{X'/Y'}$ is a strict regular immersion, and thus $X' \to Y'$ is log lci. □

We come to the main goal. Starting with a log lci map $f : X \to Y$ between logarithmically flat schemes, we construct a *refined logarithmic pullback*, associating to each map $\varphi : V \to Y$ from a logarithmically flat scheme V to Y a class in $\mathsf{CH}_\star(X \times^{\mathsf{log}}_Y V)$. We first discuss the ordinary pullbacks for f.

From the previous subsection, there is a pullback $\mathsf{CH}_\star(Y) \to \mathsf{CH}_\star(P)$. Composing with the pullback $i^! : \mathsf{CH}_\star(P) \to \mathsf{CH}_\star(X)$, we get a pullback $\mathsf{CH}_\star(Y) \to \mathsf{CH}_\star(X)$. We do not know when this pullback is independent of the choice of factorization $f = g \circ i$. However, if we impose additional constraints, the pullback is independent of those choices.

Lemma 3.6.4. *Suppose $f : X \to Y$ is a log lci morphism. Assume that either*

(i) *X, Y are logarithmically flat and tropically smooth,*

(ii) *Y is logarithmically smooth and tropically smooth, i.e., smooth, or*

(iii) *f can be factored as $i : X \to P$, $g : P \to Y$ with g logarithmically smooth and weakly semistable.*

Then the pullback is independent of the choices made.

Proof. We argue that when the hypotheses are met, the morphism $g : P \to Y$ is more well behaved than an arbitrary logarithmically smooth map[6] $P \to Y$; if (i) is satisfied, factor f as $i : X \to P'$, $g' : P \to Y$ with g' log smooth. Removing all strata of P' which are disjoint from the image of X produces an open $P \subset P'$ with a factorization $i : X \to P$, $g : P \to Y$ with P tropically smooth as well, and thus $f = g \circ i$ is then an ordinary local complete intersection by cases I and III of the analysis of the previous subsection; if (ii) is satisfied, it is a map to a smooth space and possesses a pullback associated to the diagonal of this space; if (iii) is satisfied the map g is flat, by case IV of the previous subsection. Furthermore, the pullback

[6] The reader is warned that logarithmically smooth maps can have nonequidimensional and nonreduced fibers.

$\mathsf{CH}_\star(Y) \to \mathsf{CH}_\star(P)$ defined via the relative Artin fan is simply the Gysin pullback or flat pullback $g^!$.[7] In each of these cases, by [24, Example 18.3.17], f defines a bivariant class $f^!$ in $\mathsf{CH}^\star_{\mathsf{op}}(X \to Y)$, and we have $f^! = i^! \circ g^!$. Thus, as long as one of the hypotheses are satisfied, the pullback depends only on f and not the particular choice of factorization meeting these hypotheses. □

We may unambiguously denote the pullback in those cases by $f^!$. We want to emphasize here that we *have not* defined a Gysin pullback for the underlying map of schemes of a log lci morphism that in any way generalizes the operations in [24]. We have only defined a pullback when $f : X \to Y$ can be factored as $X \to P \to Y$ of a very particular form, precisely because the constructions of [24] then apply. An arbitrary log lci map does not necessarily have a factorization $X \to P \to Y$ satisfying the additional assumptions of Lemma 3.6.4. In order to achieve such a factorization, the map f may have to be modified.

Definition 3.6.5. Let $f : X \to Y$ be a map of logarithmic schemes. A *logarithmic modification* of f is a map $f' : X' \to Y'$ where Y' is a logarithmic modification of Y and X' a logarithmic modification of X, and f' factors through the fine and saturated logarithmic base change of f along $Y' \to Y$.

Suppose now that $f : X \to Y$ is log lci, X, Y logarithmically flat. We factor f as $i : X \to P$, $g : P \to Y$ with i a strict regular embedding, g logarithmically smooth. We can perform strong semistable reduction, and there are two ways to proceed. We may apply difficult but complete semistable reduction in [7], or use nonrepresentable morphisms. In the latter case, we first perform universal *weak* semistable reduction [42], pass to barycentric subdivisions, then root, to obtain a nonrepresentable semistable morphism of Deligne–Mumford stacks.

In any event, we obtain a modification $g' : P' \to Y'$ with g' logarithmically smooth, weakly semistable, and P', Y' tropically smooth. Pull back X via $P' \to P$ to get a strict regular embedding

$$X' = X \times_P P' = X \times_P^{\mathsf{log}} P' \to P',$$

by Lemma 3.6.3. For a log lci map $f : X \to Y$, we can find a modification $f' : X' \to Y'$ which is log lci and semistable, i.e., log lci, weakly semistable, and with X', Y' tropically smooth. Furthermore, Y' can be chosen to refine any particular log modification $\tilde{Y} \to Y$: simply, the pullback $\tilde{f} : \tilde{Y} \times_Y^{\mathsf{log}} X \to \tilde{Y}$ remains log lci, and performing this construction to $\tilde{f}$ produces such a Y'. In total, starting from

(i) a log lci map $f : X \to Y$ between logarithmically flat log schemes,

(ii) a log map $\varphi : V \to Y$ from a logarithmically flat scheme V,

we find a modification $f' : X' \to Y'$ of f with f' log lci and semistable, and for which the map $\varphi' : V' = V \times_Y^{\mathsf{log}} Y' \to Y'$ is weakly semistable.

We use these observations to construct the logarithmic Gysin map.

[7] We have used the notation $g^!$ here for both flat and local complete intersection pullback.

Definition 3.6.6 (logarithmic intersection). Let $f : X \to Y$ be a logarithmic local complete intersection morphism, and $\varphi : V \to Y$ be a logarithmic morphism. Assume X, Y, V are logarithmically flat over the base field. Let

$$\begin{array}{ccc} X' & \xrightarrow{f'} & Y' \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} \\ X & \xrightarrow{f} & Y \end{array}$$

be a log modification of f with X', Y' tropically smooth, and f', and $\varphi' : V' \to Y'$ weakly semistable, where $V' = V \times_Y^{\mathsf{log}} Y'$ is the fiber product in the category of fine logarithmic schemes. Define

$$f^!_{\mathsf{log}}(V) = (p)_\star (f')^!(V') \in \mathsf{CH}_\star(V \times_Y^{\mathsf{log}} X)$$

We prove that the cycle is well defined.

Theorem 3.6.7. *The cycle that is obtained by logarithmic refined intersection is independent of choices.*

Proof. We show that if $f : X \to Y$ is log lci and semistable, $\varphi : V \to Y$ weakly semistable, and

$$\begin{array}{ccc} X'' & \xrightarrow{f''} & Y'' \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle q} \\ X & \xrightarrow{f} & Y \end{array}$$

a second diagram, as in the paragraph before the theorem, with f'', φ'' having the same properties, then $(f)^!(V) = (p)_\star (f'')^!(V')$. Since φ is weakly semistable, $V'' = V \times_Y Y''$ is Cartesian in the category of schemes as well, and since V is logarithmically flat, we can replace V with the closure of its interior without changing it class. The map $V'' \to V$ is birational on every irreducible component. As X, Y, X'', Y'' are tropically smooth, the maps p, q are lci, and we have $q^!(V) = V''$, as $q^!(V)$ is a dim V''-dimensional class supported on V''. On the other hand, we have $f''^! q^!(V) = p^! f^!(V)$. Therefore, $p_\star f''^!(V') = p_\star p^! f^!(V)$. As $p : X'' \to X$ is proper and birational, we have $p_\star p^! = \mathsf{id}$, and so the result follows. □

We end the section with a series on remarks on practicalities of this definition.

Remark 3.6.8 (complexity of the operation). The most significant combinatorial complexity comes from finding an appropriate logarithmic modification. This is a problem in polyhedral geometry. Once done, the obstruction theory on f' can be broken into a regular embedding and a weakly semistable morphism, which can be understood separately.

Remark 3.6.9 (complexity of semistabilization). In practice, semistabilizing the map $X \to Y$ can be difficult. Even when the simpler nonrepresentable strong semistabilization is used, the barycentric subdivision that is used in the construction adds enormous combinatorial complexity to the problem. When $X \to Y$ is logarithmically a monomorphism, and so $\Sigma_X \to \Sigma_Y$ is injective on cones, semistabilizing $X \to Y$ is equivalent to weakly semistabilizing it, which is a much simpler process.

Remark 3.6.10 (key example). The most important logarithmic Gysin map for us is when the map $X \to Y$ is a logarithmic modification, with X, Y smooth and logarithmically smooth. Spelling out the construction, given $f : X \to Y$, form the diagram

$$\begin{tikzcd} X \arrow[r] \arrow[d] & X \arrow[d] \\ X \arrow[r] & Y \end{tikzcd}$$

In this case, $f^!_{\mathsf{log}}(V) = X \times_Y^{\mathsf{log}} V$ is the class of the strict transform of V. The construction depends on the choice of map $\varphi : V \to Y$, and does not pass to rational equivalence. The construction is expected to pass to any reasonable notion of strict rational equivalence, which should include tropical homotopies, but we have not pursued this.

3.7. ***Calculation strategy.*** The following situation is common, and paradigmatic.[8] Consider

$$\begin{tikzcd} & \mathcal{Z} \arrow[d] \\ X \arrow[r] & Y \end{tikzcd}$$

where $\mathcal{Z}$ and X are logarithmic modifications of a logarithmically and tropically smooth Y. The three spaces are birational, but as we exhibit, there is a subtle *refined* class on the fiber product.

We form the fiber product either to the diagram as given, or after weakly semistabilizing the maps as above. Up to further birational modifications, the latter coincides with the fine and saturated logarithmic fiber product $\mathcal{Z} \times_Y^{\mathsf{log}} X$. This latter fiber product birational to X, but it may only be one component in the schematic pullback $\mathcal{Z} \times_Y X$; the ordinary pullback possesses *additional components* of expected or even excess dimension. By Fulton's theory, these excess components support a class in expected dimension, namely the dimension of any of X, Y, or $\mathcal{Z}$.

The refined pullback of the fundamental class $[\mathcal{Z}]$ is a class on $\mathcal{Z} \times_Y X$. It *differs* from the logarithmic pullback, which the class of $\mathcal{Z} \times_Y^{\mathsf{log}} X$. These spurious classes are supported on the excess components. The following examples illustrate these phenomena.

Example 3.7.1. Let Y be $\mathbb{A}^2$ and X be the blowup at the origin with exceptional divisor E. Let Z be the open subset of X obtained by removing the two torus fixed points in E; it is abstractly isomorphic to $\mathbb{A}^1 \times \mathbb{G}_m$. Everything is endowed with its toric logarithmic structure. The logarithmic fiber product is a toric variety, and in particular irreducible. The schematic fiber product contains a copy of $\mathbb{P}^1 \times \mathbb{G}_m$ that maps to the origin in Y. Since everything has the expected dimension, the spurious component contributes positively to the intersection product [24, Chapter 7].

Components can have excess dimension: if Y is $\mathbb{A}^r$, with X its blowup at the origin, and $\mathcal{Z}$ the inclusion of $\mathbb{A}^1 \times \mathbb{G}_m^{r-1}$, then the schematic pullback has a component of dimension $2r - 2$.

[8]This subsection was written in response to a suggestion of David Holmes to emphasize the importance of the refined version of Fulton's Gysin pullback formalism. We take the opportunity to thank him for these comments.

The passage to the "real world" situation in the next section requires replacing $\mathcal{Z} \to Y$ with a proper morphism. In favorable situations, we are routinely led to consider diagrams of the form

$$\begin{tikzcd}
X \times_Y^{\mathsf{log}} Z \arrow[r] \arrow[d, "q"] & X \times_Y Z \arrow[r] \arrow[d, "r"] & Z \arrow[d, "p"] \\
X \times_Y^{\mathsf{log}} \mathcal{Z} \arrow[r] & X \times_Y \mathcal{Z} \arrow[r] \arrow[d] & \mathcal{Z} \arrow[d] \\
& X \arrow[r, "f"] & Y
\end{tikzcd}$$

with the upper squares Cartesian, p, q and r, carrying compatible obstruction theories. We are interested in the pushforward to X of the difference

$$r^! f^!(\mathcal{Z}) - q^! f^!_{\mathsf{log}}(\mathcal{Z}),$$

which amounts to applying the virtual intersection formalism to piecewise polynomials on the excess components of $X \times_Y \mathcal{Z}$. While the treatment of the class $f^!(\mathcal{Z}) - f^!_{\mathsf{log}}(\mathcal{Z})$ is totally uniform and always produces a piecewise polynomial class, the interaction of this with the virtual class is more delicate, and depends on the obstruction theory. In the case of the double ramification cycle, this interaction is dealt with by means of the gluing formula.

4. Toric contact cycles

The techniques will now be used study *toric contact cycles*, and the remainder of this paper sets up and executes a calculation scheme for these cycles.

4.1. ***The plan.*** The toric contact cycles, recalled below, parameterize curves that admit a prescribed map to a fixed toric variety. In $\mathcal{M}_{g,n}$ these are intersection products of double ramification cycles. The first steps in this section describe how this product relation extends to sufficiently refined blowups of $\overline{\mathcal{M}}_{g,n}$; see Section 4.6. On any blowup, there is a lifted double ramification cycle that pushes down to the ordinary double ramification cycle on $\overline{\mathcal{M}}_{g,n}$, explained in Section 4.7. The products of the lifted classes recover toric contact cycles. The lifted cycle is identified with a virtual strict transform of the double ramification cycle.[9] We explain the idea behind this concept.

There is a moduli space $\mathsf{DR}_g(\mathsf{V})$ that is equipped with a morphism

$$\mathsf{DR}_g(\mathsf{V}) \to \overline{\mathcal{M}}_{g,n},$$

[9]The construction of a lifted class is not original to this paper. It was first considered by Holmes, Pixton, and Schmitt [32]. Our perspective builds on its Gromov–Witten theory incarnation [55; 53]. The contribution here is its understanding as a virtual strict transform.

which carries a virtual class in homological degree $2g-3+n$. Its pushforward to $\overline{\mathcal{M}}_{g,n}$ is well understood [23; 35]. The morphism may be factored as

$$\mathsf{DR}_g(\mathsf{V}) \to t\mathsf{DR}_g(\mathsf{V}) \to \overline{\mathcal{M}}_{g,n},$$

where the second morphism is an "open" logarithmic modification, i.e., an open in a logarithmic modification, and the first arrow is strict and virtually smooth of expected codimension g. The first arrow contains the virtual structure but it is strict, and so has no tropical complexity. The second arrow is far from strict, but is birational and has no virtual complexity.

We blowup $\overline{\mathcal{M}}_{g,n}$ and apply Fulton's *refined* blowup formula [24, Example 6.7.1] to the fundamental class of $t\mathsf{DR}_g(\mathsf{V})$. As in the examples of Section 3.7, we obtain a class measuring the difference between the strict and total transform of $[t\mathsf{DR}_g(\mathsf{V})]$. Virtual pullback along the first arrow in the factorization produces an expression for the virtual pullback of the strict transform, that is, the virtual strict transform. This is the heart of the analysis, and takes place starting in Section 4.9.

4.2. ***Nonsingular and compact type curves.*** We begin in earnest. Let $\mathsf{V}_1, \ldots, \mathsf{V}_r$ each be vectors in $\mathbb{Z}^n$ whose components have vanishing sum, referred to as *contact vectors*. Denote the ordered collection as $\underline{\mathsf{V}}$. Each V_i determines contact (or ramification) data for a map from n-pointed curves to $\mathbb{P}^1$: the j-th entry of V_i is the zero order of the map at the j-th marking, where a negative zero order is interpreted as a pole order. The list $\underline{\mathsf{V}}$ determines contact data for a map from an n-pointed curve to $(\mathbb{P}^1)^r$.

Let $\mathsf{TC}_g^\circ(\underline{\mathsf{V}})$ be stack of maps from smooth n-pointed curves of genus g to the toric pair

$$(C, p_1, \ldots, p_n) \to (\mathbb{P}^1)^r,$$

with ramification specified by $\underline{\mathsf{V}}$ and considered up to the torus action on the codomain.

Proposition 4.2.1. *The map* $\mathsf{TC}_g^\circ(\underline{\mathsf{V}}) \to \mathcal{M}_{g,n}$ *is a closed immersion.*

Proof. On a fixed pointed curve $(C, p_1, \ldots, p_n)$, if such a map exists it is unique up to the rank-r torus scaling accounted for above. Indeed, it is equivalent to the condition that each of the Cartier divisors $\mathsf{V}_i \cdot \underline{p}$ is trivial in the Picard group of C. If the r line bundles are trivial, each has a unique section up to scaling, giving the morphism to $\mathbb{P}^1$ on each factor. □

The construction therefore determines a cycle (and not merely a cycle class) in $\mathcal{M}_{g,n}$ whose expected codimension is rg. There is a basic intersection relation

$$\mathsf{TC}_g^\circ(\underline{\mathsf{V}}) = \bigcap_{i=1}^{r} \mathsf{DR}_g^\circ(\mathsf{V}_i).$$

This relation is visible already at the level of *cycles*, rather than merely cycle classes. The cycle class is 0 in Chow homology, because the tautological ring is 0 in codimension g. The product relation extends to the moduli space of curves of compact type, and passing to Chow homology, gives a basic product relation. The extension to $\overline{\mathcal{M}}_{g,n}$ is much more subtle, and is essentially our main topic in this part of the paper. We require tropical techniques to go further.

4.3. ***A refresher on tropical moduli.*** Tropical geometry provides a route to a compactification of this locus. The discussion here is contained in [40] but we rework it into a setup that is more tailored to the double ramification cycle.

The moduli space $\overline{\mathcal{M}}_{g,n}$ of stable genus g curves with n markings carries a logarithmic structure from the divisor of singular curves. It has an Artin fan $\mathsf{A}_{g,n}$ and a strict morphism

$$\overline{\mathcal{M}}_{g,n} \to \mathsf{A}_{g,n}.$$

The Artin fan is the colimit over points x in $\overline{\mathcal{M}}_{g,n}$ of stacks $\mathsf{A}_x = \operatorname{Spec}(k[P_x])/\operatorname{Spec}(k[P_x^{\mathsf{gp}}])$ where P_x is the characteristic monoid at x. The colimit is taken in logarithmic algebraic stacks.

It is convenient to understand the stack $\mathsf{A}_{g,n}$ in more combinatorial terms. A notion of *stack over the category of cone complexes* has been introduced in [18] leading to a categorical treatment of tropical moduli problems, and the moduli space of curves.

To each n-pointed curve $(C, p_1, \ldots, p_n)$ one may associate a *weighted dual graph*

$$\mathsf{G} = (V, E, L, h),$$

where

(i) the vertex set V is the set of components of C;

(ii) the edge set E is the set of nodes of C, where an edge $e \in E$ is incident to vertices v_1, v_2 if the corresponding node lies on both corresponding components;

(iii) the ordered set of legs of L labeled $\{1, \ldots, n\}$ with a marked leg incident to the vertex whose associated component contains the marking;

(iv) the function $h : V \to \mathbb{Z}_{\geq 0}$ is the genus function, where $h(v)$ is the geometric genus of the component corresponding to v.

The *genus* of a weighted dual graph is the sum of the values of the genus function at all vertices, plus the first Betti number of the geometric realization of G. A *tropical curve* is a pair (G, ℓ) where G is a weighted dual graph as above and

$$\ell : E \to \mathbb{R}_{>0}$$

is a length function. The function ℓ enhances the topological realization of the graph to a metric space. We enhance further this by attaching, for each leg, a copy of $\mathbb{R}_{\geq 0}$ to this metric where 0 is identified with the vertex which supports the leg.

If a weighted dual graph G is fixed, the interior of the cone $\sigma_{\mathsf{G}} := \mathbb{R}^E_{\geq 0}$ is the moduli space of all tropical curves equipped with an identification of the underlying weighted dual graph with G. We refer to it as the *moduli cone of* G. A *graph contraction* of a weighted dual graph is a sequence of edge contractions $\pi : \mathsf{G} \to \mathsf{G}'$ of the underlying graph with its canonically defined genus function given by assigning a vertex v' the genus of $\pi^{-1}(v')$. An edge contraction $\mathsf{G} \to \mathsf{G}'$ induces a morphism

$$\sigma_{\mathsf{G}'} \hookrightarrow \sigma_{\mathsf{G}}$$

identifying the former with a face of the latter, obtained by setting the length function on coordinates corresponding to contracted edges to be zero.

The moduli space of tropical curves may now be constructed. There are multiple approaches: via generalized cone complexes [4], via cone stacks, and via combinatorial cone stacks [18]. We opt for the third, following [18, Section 3.4] which is expanded on in [62, Section 3.2].

Let $I_{g,n}$ be the category whose objects are weighted dual graphs of stable curves of genus g with n-marked points. The arrows are graph contractions that are compatible with all labels. Let $\mathbf{RPC}^f$ be the category of sharp rational polyhedral cones with face morphisms. The functor

$$\sigma : I_{g,n} \to \mathbf{RPC}^f$$

defines a category fibered in groupoids. This is explained in [62, Proposition 3.2] together with the discussion following Definition 3.1 in the same paper. The diagram of cones associated to the functor above precisely describes the moduli stack $\mathcal{M}_{g,n}^{\mathrm{trop}}$.

The 2-categories of cone stacks and Artin fans are equivalent [18, Theorem 3]. The cone stack $\mathcal{M}_{g,n}^{\mathrm{trop}}$ defined above defines an Artin fan which we denote $a^\star\mathcal{M}_{g,n}^{\mathrm{trop}}$. There is a morphism

$$\mathrm{trop}_{g,n} : \overline{\mathcal{M}}_{g,n} \to a^\star\mathcal{M}_{g,n}^{\mathrm{trop}}$$

that is logarithmically smooth, strict, and surjective.

4.4. ***A subdivision of the moduli space.*** The Artin fan can be used to produced an open set in birational models of $\overline{\mathcal{M}}_{g,n}$. Fix a length n integer vector V_i whose components sum to 0. Fix a weighted dual graph G and let σ_{G} denote its moduli cone. As described in the paragraphs above, if a point p in $\sigma_{\mathsf{G}}^\circ$ is fixed, we obtain a metric graph structure enhancing the realization of G as a CW complex. We denote it by $\sqsubset_p$.

Definition 4.4.1. A *balanced function on* $\sqsubset_p$ *with slopes* V_i is a continuous function $\sqsubset \to \mathbb{R}$ such that

(i) the function restricted to any edge of $\sqsubset_p$ is linear with integer slope,

(ii) the sum of the outgoing slopes at every vertex of $\sqsubset_p$ is zero, and

(iii) the slope along the leg labeled $j \in \{1, \ldots, n\}$ is the index j entry in the contact vector V_i.

Balancing ensures that such a function is unique up to additive constant if it exists, but it may not.

Remark 4.4.2. We follow the treatment of the double ramification cycle via logarithmic geometry and curves carrying piecewise linear functions on their tropicalizations found in [40]. In this paper, the balancing condition is not insisted upon. We insist on it here because it makes the cone complexes that we consider somewhat more concrete. Our interest here is only in the "usual" double ramification cycle rather than its pluricanonical variants, which allows us to fully classify the necessary piecewise linear functions from the outset.

Lemma 4.4.3. *The closure in* σ_{G} *of the set of points* p *in* $\sigma_{\mathsf{G}}^\circ$ *such that the tropical curve* $\sqsubset_p$ *admits a balanced function with slopes given by the contact vector* V_i *is a subfan of* σ_{G}*, that is, a union of cones in a subdivision.*

Proof. This is the main result in [62, Section 4]. □

For each cone σ_{G}, the lemma above gives rise to a union of cones contained inside $\sigma_{\mathsf{G}}(\mathsf{V}_i)$. These cones are typically not full dimensional, and in particular, the subfan is not complete. One can see this by elementary geometry. First note that the coordinates in σ_{G} are the edge lengths, which we can change freely. But, if we are given a map from $\sqsubset_p$ to $\mathbb{R}$, where the underlying graph is G, there are constraints on moving the edge lengths while keeping the map to $\mathbb{R}$. Indeed, each cycle imposes one condition on the edge lengths. Therefore, one expects, and it is often the case in practice, that the subfan in the lemma has codimension equal to the number of cycles.

The intersection of this conical subfan with a face of σ_{G} corresponding to a graph contraction $\mathsf{G} \to \mathsf{G}'$ is the set of tropical curves with underlying weighted graph G' that admit a balanced function with slopes V_i. It follows from the work of Ulirsch and Zakharov [62, Theorem B] that the resulting cones glue to form a substack and is equipped with a map to the moduli stack of tropical curves

$$\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#} \to \mathcal{M}^{\mathrm{trop}}_{g,n}.$$

The domain is a moduli space for tropical curves with a balanced function with slopes V_i. We use # to indicate that we will momentarily change the conical and integral structure on this substack.

Remark 4.4.4. Ulirsch and Zakharov construct and study the space above, as the *space of principal divisors*, as a generalized cone complex. The generalized cone complex comes with a presentation as a colimit of a diagram of cones and face morphisms parameterized by an index category; the functor defines a category fibered in groupoids and gives rise to a cone stack [62, Section 3.2]. The relation between generalized cone complexes and cone stacks is handled in [61, Section 1.3].

We refine the cone structure on $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ to guarantee nonsingularity of its Artin fan. In above cone structure, it may occur that, in the interior of a cone σ in $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ where the weighted graph is constant, the difference in the positions of two vertices may change from positive to negative depending on the chosen point in σ. We replace the existing polyhedral structure on $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ to a finer one, avoiding this phenomenon.

By [40, Section 5.5], there is a subdivision of $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$, possibly including passage to a sublattice,[10] where in the interior of each cone, the images of the vertices of the universal curve in $\mathbb{R}$ are totally ordered. Define $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ to be this polyhedral subdivision. We explain it in more combinatorial terms, and record it for ease of access.

Terminology 4.4.5. Let p be a point in $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$. The point p determines a piecewise linear function on $\sqsubset$ up to constant with slopes given by V_i. Choose any such piecewise linear function

$$F : \sqsubset_p \to \mathbb{R}.$$

[10] It was pointed out to the authors by Holmes that there is a minor error in the description of the rubber moduli space of maps in [40], concerning the integral structure on the base of this tropical moduli space, or equivalently, a certain root construction along the logarithmic boundary; we are informed that the clarifications will appear in a revised version of [13]. The integral structure is not critical for us, so long as it ensures that the universal curve is reduced.

(i) We refer to the underlying weighted graph G of $\sqsubset_p$ as the *stable source graph at p*.

(ii) Let $\mathcal{R}$ be the subdivision of $\mathbb{R}$ obtained by placing a vertex at the image of each vertex of $\sqsubset_p$. Let T be the graph underlying $\mathcal{R}$. This graph is independent of the choice of F. We refer to the graph T as the *target graph at p*.

(iii) Let $\sqsubset'_p$ be obtained by subdividing $\sqsubset_p$ at the preimages of all vertices of $\mathcal{R}$. Let Γ be its underlying weighted dual graph, labeling all new vertices as genus 0. We refer to Γ as the *semistable source graph at p*.

(iv) The *rubber combinatorial type at p* is the data of the weighted dual graph Γ, the target graph T, the graph morphism between them, and the slope decoration for each edge and leg of Γ.

If the rubber type is fixed, the moduli of maps with this type is a simplicial cone.

Proposition 4.4.6. *Fix V_i as above and let G be a weighted dual graph. Let σ_{G} denote the tropical moduli cone. The closure of the set of points of σ_{G} that lie in $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ and have a fixed rubber combinatorial type forms a simplicial cone.*

Proof. Fix a rubber combinatorial type and examine the graph map $\Gamma \to T$. Let $C \subset E$ be the subset of edges of Γ that have slope 0 and let N be the set of edges in the target graph T. The set of points in $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ with this rubber combinatorial type is parameterized by arbitrary positive edge lengths corresponding to the contracted edges in Γ, and arbitrary positive edge lengths corresponding to the edges in the target graph T. Every noncontracted edge of Γ maps with positive slope onto one of the edges of T and therefore its length is determined as well. □

The space $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ comes equipped with a universal stable tropical curve $\sqsubset \to \mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$, obtained by pulling back the universal tropical curve from $\mathcal{M}^{\mathrm{trop}}_{g,n}$. It is also equipped with a universal semistable tropical curve, coming from Terminology 4.4.5(iii). Let $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ be the cone structure refining $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}_{\#}$ such that two points lie in the interior of a cone if and only if their rubber combinatorial types coincide. The proposition above shows that the cones are simplicial. We choose an integral structure on $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ such that (i) the cones all in fact smooth and (ii) the universal semistable source tropical curve is weakly semistable over the tropical moduli space.

Remark 4.4.7. The subdivision studied by Marcus and Wise is referred to as the *rubber* space for its connections to Gromov–Witten theory of nonrigid targets [27]. The subdivision was studied earlier in the genus 0 context by Cavalieri, Markwig, and the second author [17, Section 3.1] and the pictures in loc. cit. may help a reader better visualize the cones in the subdivision. A similar construction appears in [49, Section 2.4].

With this conical and integral structure fixed, we pass to Artin fans and define

$$t\mathsf{DR}_g(\mathsf{V}) := a^{\star}\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}} \times_{a^{\star}\mathcal{M}^{\mathrm{trop}}_{g,n}} \overline{\mathcal{M}}_{g,n}.$$

The arrow from this fiber product to $\overline{\mathcal{M}}_{g,n}$ is a logarithmically étale and birational morphism. There are two flat families of universal curves over $t\mathsf{DR}_g(\mathsf{V})$: (i) a universal *stable curve* pulled back from $\overline{\mathcal{M}}_{g,n}$; (ii) a universal *semistable curve* obtained from the universal stable curve by inserting 2-pointed $\mathbb{P}^1$-components. The space is stratified by rubber combinatorial types. On the locally closed strata, the universal semistable curve is obtained from Terminology 4.4.5(iv). The details may be found in [40, Section 5.2].

We proceed analogously for the toric contact cycles. Fix a vector $\underline{\mathsf{V}}$ of n-tuples and construct a subcomplex

$$\mathsf{TC}_g(\underline{\mathsf{V}})^{\mathrm{trop}} \hookrightarrow \mathcal{M}_{g,n}^{\mathrm{trop}}$$

as the fiber product[11] of the all tropical spaces $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ over the inputs in $\underline{\mathsf{V}}$, over the space $\mathcal{M}_{g,n}^{\mathrm{trop}}$. In identical fashion, the subcomplex above determines a Deligne–Mumford stack

$$t\mathsf{TC}_g(\underline{\mathsf{V}}) \to \overline{\mathcal{M}}_{g,n},$$

where the map to curves is a logarithmically étale and birational morphism.

4.5. ***Toric contacts.*** Each stack $t\mathsf{DR}_g(\mathsf{V}_i)$ is equipped with a universal stable curve. The birational spaces induced by the subdivisions above admit Abel–Jacobi morphisms, extending

$$(C, p_1, \dots, p_n) \mapsto \mathcal{O}_C(V_i \cdot \underline{p}).$$

The construction here is due to Marcus and Wise [40], but see also [28; 30]. The universal curve is equipped with a well-defined piecewise linear function up to constant functions, and the toric dictionary gives rise to a line bundle. This gives rise to an Abel–Jacobi section associated to V_i

$$t\mathsf{DR}_g(\underline{\mathsf{V}}_i) \to \mathsf{Pic}_{g,n},$$

where the codomain is the pullback of the universal Picard group from the moduli stack of curves. Since it is a section of a smooth fibration, this Abel–Jacobi section is a regular embedding, and can be equipped with a perfect obstruction theory, and so comes equipped with a Gysin pullback [40, Section 4.6]. Analogously, the factorwise morphism associated to $\underline{\mathsf{V}}$ is

$$t\mathsf{TC}_g(\underline{\mathsf{V}}) \to \mathsf{Pic}_{g,n} \times_{t\mathsf{TC}_g(\underline{\mathsf{V}})} \cdots \times_{t\mathsf{TC}_g(\underline{\mathsf{V}})} \mathsf{Pic}_{g,n}.$$

Definition 4.5.1. The *toric contact space* $\mathsf{TC}_g(\underline{\mathsf{V}})$ is the pullback of the factorwise zero section in $t\mathsf{TC}_g(\underline{\mathsf{V}})$ under the morphism above. The virtual class, defined by Gysin pullback of the zero section is the *toric contact class*. Its pushforward to $\overline{\mathcal{M}}_{g,n}$ is the *toric contact cycle*. For a single factor, we replace TC_g with DR_g and refer to it as *double ramification* space, class, or cycle.

A word of justification is required concerning the pushforward operation in the definition. The space $\mathsf{TC}_g(\underline{\mathsf{V}})$ is defined by pulling back a closed subscheme of a nonproper space — the universal Picard — to a nonproper space. It is nevertheless proper; see [40, Section 4] for the rank-1 case. Since the diagonal is closed, the general case follows from this.

[11] There is no difficulty in defining the fiber product. The map from $\mathsf{DR}_g^{\mathrm{trop}}(\mathsf{V}_i)$ to $\mathcal{M}_{g,n}^{\mathrm{trop}}$ is representable by cone complexes, so we work on a cover by cone complexes, where it is defined in the standard fashion in [42, Section 2.2].

By [40, Section 5] the double ramification space defined above coincides up to saturation of the logarithmic structure with the moduli space of rubber stable maps to a nonrigid $\mathbb{P}^1$ with its toric logarithmic structure. We use the Gromov–Witten perspective heavily in the analysis below; the standard reference is [27, Section 2.4]. A modern treatment via logarithmic geometry is given in [40, Section 5], and a careful explanation is recorded in [13, Section 6].

For the higher-rank case, we note that $\underline{\mathsf{V}}$ consists of r entries, the space $\mathsf{TC}_g(\mathsf{V})$ can be considered a compactification of the space of maps to $(\mathbb{P}^1)^r$ up to the $\mathbb{G}_m^r$ rubber.[12] By birational invariance, the target $(\mathbb{P}^1)^r$ can be replaced with any toric compactification of the same torus [2]. A more precise relationship can be deduced from what follows; see Remark 4.6.4.

4.6. ***Product rule.*** For ease of bookkeeping, we assume $\underline{\mathsf{V}}$ has two entries V_1 and V_2. We have the basic product diagram

$$\begin{tikzcd}
\mathsf{TC}_g(\underline{\mathsf{V}}) \arrow[r, "\upsilon"] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{P}(\underline{\mathsf{V}}) \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V}_1) \times \mathsf{DR}_g(\mathsf{V}_2) \arrow[d] \\
t\mathsf{TC}_g(\underline{\mathsf{V}}) \arrow[r] & t\mathsf{P}(\underline{\mathsf{V}}) \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & t\mathsf{DR}_g(\mathsf{V}_1) \times t\mathsf{DR}_g(\mathsf{V}_2) \arrow[d, "\psi"] \\
 & \overline{\mathcal{M}}_{g,n} \arrow[r, "\Delta"'] & \overline{\mathcal{M}}_{g,n} \times \overline{\mathcal{M}}_{g,n}
\end{tikzcd}$$

There are two deficiencies in this diagram. First, the spaces on the fundamental class of $t\mathsf{TC}_g(\underline{\mathsf{V}})$ does not pushforward to that of $t\mathsf{P}(\underline{\mathsf{V}})$; the map υ is correspondingly not virtually birational. Second, while the diagonal is equipped with a normal bundle, there is no guarantee that the arrow above it is. The issues are resolved simultaneously by weakly semistabilizing ψ.

Fx a complete subdivision, possibly including passage to finite index sublattices,

$$\mathcal{M}^{\mathrm{trop}}_{g,\underline{\mathsf{V}}} \to \mathcal{M}^{\mathrm{trop}}_{g,n}$$

with the following two properties:

(i) The cone stack $\mathcal{M}^{\mathrm{trop}}_{g,\underline{\mathsf{V}}}$ is smooth.

(ii) For each vector V_i the pullback of $\mathsf{DR}_g(\mathsf{V}_i)^{\mathrm{trop}}$ mapping $\mathcal{M}^{\mathrm{trop}}_{g,\underline{\mathsf{V}}}$ has the property that cones of the source surject on to cones, and the lattices of the source cones are saturated in the image.

We form the fine and saturated logarithmic base change to obtain

$$\begin{tikzcd}
t\mathsf{DR}_g(\mathsf{V}_i)^{\dagger} \arrow[r] \arrow[d] & \mathcal{M}^{\mathrm{trop}}_{g,\underline{\mathsf{V}}} \arrow[d] \\
t\mathsf{DR}_g(\mathsf{V}_i) \arrow[r] & \mathcal{M}^{\mathrm{trop}}_{g,n}
\end{tikzcd}$$

[12] These spaces have not appeared in the literature, but appear implicitly in the boundary structure of the spaces constructed in [55] and will appear in forthcoming work of Carocci and Nabijou.

The space $t\mathsf{DR}_g(\mathsf{V}_i)^\dagger$ inherits an Abel–Jacobi section to its universal Picard variety, and Gysin pullback of the zero section gives rise to a space $\mathsf{DR}_g(\mathsf{V}_i)^\dagger$ equipped with a virtual class. There is a new diagram

$$\begin{tikzcd}
\mathsf{TC}_g(\underline{\mathsf{V}})^\dagger \arrow[r, "\upsilon"] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V}_1)^\dagger \times \mathsf{DR}_g(\mathsf{V}_2)^\dagger \arrow[d] \\
t\mathsf{TC}_g(\underline{\mathsf{V}})^\dagger \arrow[r, "\delta"'] \arrow[d] & t\mathsf{DR}_g(\mathsf{V}_1)^\dagger \times t\mathsf{DR}_g(\mathsf{V}_2)^\dagger \arrow[d, "\psi"] \\
\overline{\mathcal{M}}_{g,\underline{\mathsf{V}}} \arrow[r, "\Delta_{\underline{\mathsf{V}}}"'] & \overline{\mathcal{M}}_{g,\underline{\mathsf{V}}} \times \overline{\mathcal{M}}_{g,\underline{\mathsf{V}}}
\end{tikzcd}$$

where the lower square is Cartesian in the category of fine and saturated logarithmic stacks; the upper square is Cartesian in all categories.

Lemma 4.6.1. *The morphism ψ is flat of relative dimension* 0 *with reduced fibers.*

Proof. The map ψ is a logarithmically étale with smooth target. The source is a locally toric, therefore Cohen–Macaulay. Flatness will follow from equidimensionality of the fibers. Equidimensionality and reducedness of the fibers follow from the toroidal criteria [1, Section 4]; see also [42]. □

Lemma 4.6.2. *The squares in the diagram are Cartesian diagrams of algebraic stacks and of fine and saturated logarithmic stacks.*

Proof. The squares are defined to be Cartesian in fine and saturated logarithmic stacks. The morphism ψ is integral and saturated by construction, so the two pullbacks coincide. □

Theorem 4.6.3. *The toric contact cycle in $\overline{\mathcal{M}}_{g,\underline{\mathsf{V}}}$ is the product of the strict transforms of the double ramification cycles, i.e., there is an equality of classes in the Chow groups of $\mathsf{TC}_g(\underline{\mathsf{V}})^\dagger$:*

$$[\mathsf{TC}_g(\underline{\mathsf{V}})^\dagger]^{\mathrm{vir}} = \Delta^!_{\underline{\mathsf{V}}}[\mathsf{DR}_g(\mathsf{V}_1)^\dagger \times \mathsf{DR}_g(\mathsf{V}_2)^\dagger]^{\mathrm{vir}}.$$

Proof. We examine the diagram above. The morphism $\Delta_{\underline{\mathsf{V}}}$ is a regular embedding since $\overline{\mathcal{M}}_{g,\underline{\mathsf{V}}}$ is smooth. By the lemma, the map ψ is flat and therefore the middle row map δ is also a regular embedding. The Gysin morphisms for δ and $\Delta_{\underline{\mathsf{V}}}$ coincide [24, Theorem 6.2(c)]. We will show

$$[\mathsf{TC}_g(\underline{\mathsf{V}})^\dagger]^{\mathrm{vir}} = \delta^![\mathsf{DR}_g(\mathsf{V}_1)^\dagger \times \mathsf{DR}_g(\mathsf{V}_2)^\dagger]^{\mathrm{vir}}.$$

The obstruction theory on the morphism

$$\mathsf{DR}_g(\mathsf{V}_1)^\dagger \times \mathsf{DR}_g(\mathsf{V}_2)^\dagger \to t\mathsf{DR}_g(\mathsf{V}_1)^\dagger \times t\mathsf{DR}_g(\mathsf{V}_2)^\dagger$$

is obtained, on each factor, by pulling back the obstruction theory of the appropriate Abel–Jacobi section; the latter is a regular embedding since it is a section of a smooth fibration and therefore has a normal bundle. Similarly, the Picard group of the universal curve over $t\mathsf{TC}_g(\underline{\mathsf{V}})$ is the pullback along δ of the Picard group from either factor, and Abel–Jacobi sections on $t\mathsf{TC}_g(\underline{\mathsf{V}})$ obtained by composition. It follows that the obstruction theory on the vertical arrows in the top square are the same. The theorem follows by compatibility of Gysin pullback. ☐

Remark 4.6.4 (via Gromov–Witten theory). There is different route to the toric contact cycle, via Gromov–Witten theory. Fix the vector of r contact vectors $\underline{\mathsf{V}}$. Let X be any proper toric variety equipped with its toric logarithmic structure. The data $\underline{\mathsf{V}}$ determines a logarithmic stable maps space $\mathsf{K}_{g,\underline{\mathsf{V}}}(X)$. We make a simplifying assumption that the first marked point has contact order 0 with all divisors, that is, the first entry in each contact vector is 0. There is a morphism

$$\mathsf{ev}_1 : \mathsf{K}_{g,\underline{\mathsf{V}}}(X) \to X.$$

The target can be taken to be a product of projective lines. The Gromov–Witten cycle

$$\mathsf{ev}_1{}^\star([\mathsf{pt}]) \cap [\mathsf{K}_{g,\underline{\mathsf{V}}}(X)]^{\mathsf{vir}}$$

can be pushed forward to $\overline{\mathcal{M}}_{g,n}$. This cycle completes $\mathsf{TC}_g^\circ(\underline{\mathsf{V}})$ on $\mathcal{M}_{g,n}$: curves that admit a map to X with contact orders $\underline{\mathsf{V}}$, up to the torus action. When X is $\mathbb{P}^1$, this completed class coincides with the double ramification cycle. Since $\mathsf{TC}_g(\underline{\mathsf{V}})$ satisfies a product rule and the Gromov–Witten class also satisfies a product rule [29; 53] the higher-rank classes also coincide. The virtual class is pulled back under forgetting points with trivial contact order, and the pullback along $\overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ is injective on Chow groups, so there is no loss of information in the assumption.

4.7. ***The lightning class.*** Fix a single contact vector V. Consider any *complete* smooth subdivision

$$\mathcal{M}_{g,\mathsf{V}}^{\mathrm{trop}} \to \mathcal{M}_{g,n}^{\mathrm{trop}}$$

with the property that the pullback of $\mathsf{DR}_g^{\mathrm{trop}}(\mathsf{V})$ is weakly semistable over $\mathcal{M}_{g,n}^{\mathrm{trop}}$. A subdivision satisfying this property is stable under further subdivision, so this property holds for any sufficiently fine subdivision. The constructions above determine a class

$$[\mathsf{DR}_g^{\lightning}(\mathsf{V})] \quad \text{in } \mathsf{CH}^g(\overline{\mathcal{M}}_{g,\mathsf{V}}; \mathbb{Q}),$$

by pushing forward the double ramification class. We verify the independence of choices.

Proposition 4.7.1. *The double ramification class constructed above is stable under pullback*, i.e., *for any further logarithmic blowup*

$$\overline{\mathcal{M}}'_{g,\mathsf{V}} \to \overline{\mathcal{M}}_{g,\mathsf{V}}$$

with smooth source the two double ramification cycles are related by pullback along the local complete intersection morphism above.

Proof. The blowup in question is given by a subdivision of the tropical moduli stack $\mathcal{M}_{g,n}^{\mathrm{trop}}$. By supposition, the pullback of $\mathsf{DR}_g(\mathsf{V})^{\mathrm{trop}}$ to $\mathcal{M}_{g,\mathsf{V}}^{\mathrm{trop}}$ is a union of cones. These cones determine an open substack $\mathcal{U}_g(\mathsf{V})$; exactly as in the previous section, the substack has an Abel–Jacobi section, and the pullback of the zero section is the proper stack which we denote $\mathsf{Z}_g(\mathsf{V})$ to avoid a clash of notation. Denoting the analogous constructions on $\overline{\mathcal{M}}'_{g,V}$ with primes, we have two fiber squares in the category of algebraic stacks:

$$\begin{tikzcd}
\mathsf{Z}_g(\mathsf{V})' \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{Z}_g(\mathsf{V}) \arrow[d] \\
U_g(\mathsf{V})' \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & U_g(\mathsf{V}) \arrow[d] \\
\overline{\mathcal{M}}'_{g,\mathsf{V}} \arrow[r] & \overline{\mathcal{M}}_{g,V}
\end{tikzcd}$$

The vertical arrows in the lower square are flat and therefore it suffices to check that the double ramification cycle on $\mathsf{Z}_g(\mathsf{V})$ pulls back to that on $\mathsf{Z}_g(\mathsf{V})'$. Arguing as in the proof of Theorem 4.6.3, the obstruction theories for the two vertical arrows in the upper square coincide, and the proposition follows by compatibility of Gysin pullbacks. □

We obtain this basic consequence: the classes $[\mathsf{DR}_g^{\not{\,}}(\mathsf{V})]$ form a compatible system of classes on all blowups of $\overline{\mathcal{M}}_{g,n}$ and we find

$$[\mathsf{DR}_g^{\not{\,}}(\mathsf{V})] \quad \text{in} \quad \varprojlim \mathsf{CH}_\star(\overline{\mathcal{M}}'_{g,n}),$$

where the limit is taken over all smooth $\overline{\mathcal{M}}'_{g,n}$ obtained from $\overline{\mathcal{M}}_{g,n}$ by blowups. In fact, the class comes from one in the colimit of Chow cohomology groups under pullback. The determination of $[\mathsf{DR}_g^{\not{\,}}(\mathsf{V})]$ classes in any blowup satisfying the properties above determines the toric contact cycles.

4.8. ***Projective bundle geometry.*** We study $[\mathsf{DR}_g^{\not{\,}}(\mathsf{V})]$ by an inductive argument, using the self-similar structure of the boundary of the space $\mathsf{DR}_g(\mathsf{V})$ that is guaranteed by the gluing formula in relative Gromov–Witten theory: the boundary is built from smaller double ramification problems. The smaller cycles can be pushed forward to smaller moduli spaces of curves, but the induction will require a more refined pushforward — to certain projective bundles over moduli of curves.

The following is a rephrasing of results on the boundary structure of spaces of stable maps to rubber [23; 27]. Logarithmic strata of $t\mathsf{DR}_g(\mathsf{V})$ are specified by rubber combinatorial types

$$\Theta = [\Gamma \to T]$$

as in Section 4.4. We remind the reader that Γ is the semistable source graph, that is, a weighted dual graph of a semistable curve, with edges and legs decorated by integers slopes and T is a graph underlying a subdivision of $\mathbb{R}$.

Fix such a type Θ and inspect a vertex u of Γ. Each flag of an edge emanating from u is decorated by a slope, and we collect these slopes in a integer vector V_u of length equal to the number of flags exiting u. The flags exiting u are not labeled, so we pass to a finite cover to label them. The group Aut_Θ of deck transformations of the cover is precisely the group of automorphisms of the graph Γ preserving labels and commuting with the map to T. We denote the space $t\mathsf{DR}_{g(u)}(\mathsf{V}_u)$ by $t\mathsf{DR}(u)$. There is a morphism

$$\mathsf{cut} : t\mathsf{DR}_{[\Gamma\to T]} \to \Big[\prod_{u\in V(\Gamma)} t\mathsf{DR}(u)/\mathsf{Aut}_\Theta\Big]$$

obtained by cutting the target graph, as in the degeneration/gluing formulae [20; 27; 44].

Lemma 4.8.1. *The morphism* cut *is a torus torsor.*

The source and target of cut are both determined by purely tropical data, and experts in toric geometry will note that it is possible to give an essentially combinatorial proof of this lemma. We record a more geometric proof.

Proof. We will pass to Aut_Θ covers first, thereby replacing the moduli stratum $t\mathsf{DR}_{[\Gamma\to T]}$ with a cover $t\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$. The cover parameterizes curves equipped with a rubber combinatorial type $[\Gamma' \to T']$ *together* with the data of a contraction of $\Gamma' \to T'$ to a fixed cover $[\Gamma \to T]$.

We describe the product space first. A point in this product yields a nodal curve associated to each vertex u in the graph Γ. The marked points on this curve correspond to the flags of edges in Γ leaving u. If u is adjacent to u' via an edge e, we glue the curves associated to u and u' at the points corresponding to the flags. We have passed to the cover, so the flags may be unambiguously identified. We repeat this for all edges of T. By construction, the curve associated to each factor admits a unique rubber combinatorial type. The types are compatible, in the sense that if there are flags emanating from u and u' that together form an edge in Γ, the slopes on these flags are equal.[13] There is a universal *glued* curve

$$\mathcal{C}^{\mathsf{glued}} \to \prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$$

At each point of the moduli space, there is a well-defined rubber combinatorial type, together with a contraction to the fixed cover $[\Gamma \to T]$. The glued curve is *semistable*, and it may have 2-pointed $\mathbb{P}^1$ components, or "trivial bubbles".

Fix an edge e of T and let $e_1, \dots, e_r$ be the edges of Γ that map to e in T, and each edge e_i consists of two flags $f_{e_i,1}$ and $f_{e_i,2}$. Suppose the vertices bearing these flags are u_1 and u_2. Each flag determines a line bundle on the moduli space $t\mathsf{DR}_g(u_i)$, namely the cotangent line bundle $\Psi_{e_i,j}$ of the corresponding universal semistable curve at this marked point, where j is either 1 or 2. On the product moduli space, we pullback and continue to denote these line bundles $\Psi_{e_i,j}$. The line bundle $\Psi_{e_i,1}^{-1} \otimes \Psi_{e_i,2}^{-1}$ is denoted $\mathcal{O}(e_i)$.

We examine the difference between this product space and the covered stratum $t\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$. This latter space is obtained by a modification induced by the tropical inclusion

$$\mathsf{DR}_g(\mathsf{V})^{\mathrm{trop}} \hookrightarrow \mathcal{M}_{g,n}^{\mathrm{trop}}.$$

Fix an edge e of the target graph T, and let $e_1, \dots, e_r$ be the edges of Γ mapping to e. Let w_i be the slope along e_i. Given a point of the product space above, there is a canonically associated glued curve

[13] Longtime readers of Gromov–Witten theory will notice that this is essentially the predeformability condition.

with a rubber combinatorial type. However, in order to produce a point of $t\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$, we must also constrain the logarithmic structure at the nodes. In order to see this, note that the weighted lengths of these corresponding edges are always equal in (the Θ-cone of) $\mathsf{DR}_g(\mathsf{V})^{\mathrm{trop}}$. These edge lengths are piecewise linear functions on the tropical moduli space, or equivalently, elements of the characteristic sheaf of the logarithmic structure. Unwinding the manner in which $\mathsf{DR}_g(\mathsf{V})^{\mathrm{trop}}$ determines a birational modification of $\overline{\mathcal{M}}_{g,n}$, to produce an element of $t\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$ we must choose an isomorphism between the line bundles $\mathcal{O}(e_i)^{\otimes w_i}$ for different i whenever the elements of the characteristic are equated.

We repeat the analysis for all edges of T; this exhibits the morphism

$$\mathsf{cut} : t\widetilde{\mathsf{DR}}_{[\Gamma\to T]} \to \prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$$

as the total space of direct sum of torsors. Specifically, for each edge e of T with edges $e_1, \dots, e_r$ of Γ mapping to it, we form the total space of the torsors

$$\mathcal{O}(w_1\cdot e_1 - w_r\cdot e_r)\oplus\cdots\oplus\mathcal{O}(w_{r-1}e_{r-1} - w_r\cdot e_r).$$

Taking another direct sum over all edges e of T, we obtain a space isomorphic to $t\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$, and this proves the lemma. $\square$

There is a map equipped with a perfect obstruction theory

$$\mathsf{DR}_g(\mathsf{V}) \to t\mathsf{DR}_g(\mathsf{V}),$$

and since the inclusion $t\mathsf{DR}_{[\Gamma\to T]} \hookrightarrow t\mathsf{DR}_g(\mathsf{V})$ is a regular embedding, the pullback of the virtual class gives a refined cycle of codimension g on $t\mathsf{DR}_{[\Gamma\to T]}$ supported on a compact subset. We denote it $\mathsf{DR}^{\mathsf{vir}}_{[\Gamma\to T]}$. The induction will require pushing forward this refined class to compactifications of $t\mathsf{DR}_{[\Gamma\to T]}$. In order to induct, we relate this to a refined class on $\prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$ and compactifications of it.

The torus bundle $t\mathsf{DR}_{[\Gamma\to T]}$ contains a compact subset and class on, namely the virtual class $\mathsf{DR}_{[\Gamma\to T]}$ of the stratum. Our main argument will be to study this class on pushforwards to compactifications of $t\mathsf{DR}_{[\Gamma\to T]}$. The study will be inductive, using the fact that $t\mathsf{DR}_{[\Gamma\to T]}$ is a bundle over a product of simpler spaces. This requires a compactification of the map cut above.

Construction 4.8.2 (a compactification of cut). The proof of the lemma shows that the morphism

$$\mathsf{cut} : t\widetilde{\mathsf{DR}}_{[\Gamma\to T]} \to \prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$$

is the total space of a direct sum of $\mathbb{G}_m$-torsors. Each torsor in the direct sum was obtained as a tensor product of cotangent line bundles at the marked points. Fix a marked point $\star$ and let $\mathbb{L}_\star$ denote the cotangent line bundle on the universal *stable curve*. On the locus of nonbroken curves in $t\mathsf{DR}(u)$ the line bundle $\mathbb{L}_\star$ coincides with the cotangent line bundle $\Psi_\star$. By the standard comparison of cotangent lines, the difference $\mathbb{L}_\star\otimes\Psi_\star^{-1}$ is the divisor associated to a linear combination of boundary divisors on $t\mathsf{DR}(u)$.

Specifically, it is the sum of boundary divisors where the marked point $\star$ is supported on an unstable component of the universal semistable curve.

Choose compactifications of each factor in $\prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$ to form $\prod_{u\in V(\Gamma)} \overline{\mathcal{M}}(u)$. Each $t\mathsf{DR}(u)$ is birational to (a finite group quotient of) an appropriate moduli space of curves, with genus and number of marked points determined by u. We require the compactification to be a proper toroidal modification of the space of stable curves corresponding to these data.

For the marked point $\star$, we choose a Cartier divisor on the compactification that restricts to $\mathbb{L}_\star \otimes \Psi_\star^{-1}$. Since each line bundle extends, the torus torsor also extends to the compactification $\prod_{u\in V(\Gamma)} \overline{\mathcal{M}}(u)$. Since the source of cut is a contained in a direct sum of line bundles that extend to the compactification in the base direction, we compactify the fiber directions of the map via the projective completion of these line bundles to obtain

$$\mathsf{kut} : \widetilde{\mathcal{M}}_{[\Gamma\to T]} \to \prod_{u\in V(\Gamma)} \overline{\mathcal{M}}(u).$$

By construction, this coincides with cut after restricting to $\prod_{u\in V(\Gamma)} t\mathsf{DR}(u)$ in the base directions and the tori in each fiber direction. □

Write $[\mathsf{DR}_{[\Gamma\to T]}]^{\mathsf{push}}$ pushforward to $\widetilde{\mathcal{M}}_{[\Gamma\to T]}$ of the virtual class of the stratum corresponding to $[\Gamma \to T]$. We write $\left[\prod_{u\in V(\Gamma)} \mathsf{DR}(u)\right]^{\mathsf{push}}$ for the factorwise double ramification cycle determined by each vertex. The following proposition shows that the split class can be used to recover the class of the stratum.

Proposition 4.8.3. *Let H denote the operational class of the fiberwise hyperplane in the bundle $\widetilde{\mathcal{M}}_{[\Gamma\to T]}$. There is an equality of classes on $\widetilde{\mathcal{M}}_{[\Gamma\to T]}$ given by*

$$[\mathsf{DR}_{[\Gamma\to T]}]^{\mathsf{push}} = \lambda \cdot H^k \cap \mathsf{kut}^\star \Bigg[\prod_{u\in V(\Gamma)} \mathsf{DR}(u)\Bigg]^{\mathsf{push}},$$

where λ is a nonzero rational number and k is the fiber dimension of kut.

In the following proof, we use results from Chen's treatment of the degeneration formula for logarithmic expanded degenerations [20, Section 7]. Our setting is slightly different, in that we describe the boundary of the space of maps to a rubber target, rather than a stratum in a degeneration of the moduli space. The precise results we need apply without change in our setting.

Proof. By the gluing formula in relative Gromov–Witten theory, after passing to a finite cover to kill the automorphisms, the arrow

$$\widetilde{\mathsf{DR}}_{[\Gamma\to T]} \to \prod_{u\in V(\Gamma)} \overline{\mathcal{M}}(u),$$

obtained by restricting kut, is finite and étale over its image and its image is precisely the external product $\prod_{u\in V(\Gamma)} \mathsf{DR}(u)$. This result follows immediately from [20, Lemma 7.9.1]. We examine the projective bundle obtained by pulling back the bundle from Construction 4.8.2 to this product $\prod_{u\in V(\Gamma)} \mathsf{DR}(u)$. The space $\widetilde{\mathsf{DR}}_{[\Gamma\to T]}$ is naturally contained in this bundle.

We claim the bundle becomes trivial after replacing the base with a finite cover which we denote $\prod^{\sim}_{u\in V(\Gamma)} \mathsf{DR}(u)$. Indeed, on restricting to this locus, the morphism admits an orbifold section by [20, Equation 7.9.2], and therefore the torsor is trivial on this locus; see also [27, Remark 3.4]. The natural obstruction theory on $\prod^{\sim}_{u\in V(\Gamma)} \mathsf{DR}(u)$ pulls back to the obstruction theory on the section given by $\mathsf{DR}_{[\Gamma\to T]}$; this follows by the same argument presented in [20, Section 4.10]. Therefore, on this trivial projective bundle, the virtual class of the section is equal to the pullback of the virtual class on $\prod^{\sim}_{u\in V(\Gamma)} \mathsf{DR}(u)$ times a power of the fiberwise hyperplane class. The result follows from the projection formula. □

4.9. ***Computing virtual strict transforms.*** The class $[\mathsf{DR}_g(\mathsf{V})]$ in the Chow cohomology of $\overline{\mathcal{M}}_{g,n}$ is known to be tautological, either by the result of Faber and Pandharipande [23] or via the explicit description in terms of the standard basis of the tautological ring [35]. We compute $[\mathsf{DR}_g^{\lightning}(\mathsf{V})]$ on a sufficient blowup of the moduli space of curves and show that it lies in the tautological ring.

The class $[\mathsf{DR}_g^{\lightning}(\mathsf{V})]$ can be regarded as a *virtual strict transform* of the refined cycle $[\mathsf{DR}_g(\mathsf{V})]$ in $\overline{\mathcal{M}}_{g,n}$ in the following sense. We begin with a morphism

$$t\mathsf{DR}_g(\mathsf{V}) \to \overline{\mathcal{M}}_{g,n}$$

and examine its pullback under any sufficient projective birational modification $\overline{\mathcal{M}}_{g,\mathsf{V}}$, which we assume to be smooth. This modification is given by the blowup of a monomial ideal sheaf, in the sense discussed previously. The pullback of $t\mathsf{DR}_g(\mathsf{V})$ is the *total transform* of the morphism.

There is a well-defined *strict transform* which is obtained by pulling back the ideal sheaf above to $t\mathsf{DR}_g(\mathsf{V})$ and blowing it up there. But the ideal sheaf is monomial and therefore pulled back from the Artin fan of $\overline{\mathcal{M}}_{g,n}$, so this strict transform is obtained by first subdividing $\overline{\mathcal{M}}_{g,n}$, then pulling back the subdivision to $\mathsf{DR}_g(\mathsf{V})^{\mathrm{trop}}$, and performing the associated birational modification on $t\mathsf{DR}_g(\mathsf{V})$. Since the blowup is sufficiently fine, the morphism from this modification of $t\mathsf{DR}_g(\mathsf{V})$ to $\overline{\mathcal{M}}_{g,\mathsf{V}}$ is the inclusion of an open substack.

There are now three spaces in play: the space $t\mathsf{DR}_g(\mathsf{V})$, its strict transform, and its total transform. The first two are pure of dimension $3g-g+n$, while the latter is not pure, but is equipped with an excess class in homological degree $3g-3+n$ by Fulton's *refined* Gysin pullback applied to the local complete intersection morphism

$$\overline{\mathcal{M}}_{g,\mathsf{V}} \to \overline{\mathcal{M}}_{g,n};$$

see [24, Section 6.6]. The vector V determines Abel–Jacobi sections on all three spaces, and we can perform refined intersection for each one with the 0 section of the universal Picard group. We obtain the *virtual strict transform* and *virtual total transform*. By definition, after pushing forward to $\overline{\mathcal{M}}_{g,\mathsf{V}}$, the virtual total transform is the class of the Chow cohomological pullback under the blowup and is known [23; 35]. We compute the virtual strict transform.

The situation is complicated by the fact that a blowup along a monomial ideal is singular and difficult to control in intersection theory, and we instead argue inductively along a factorization.

Factorization. As a first step, we require a sequence of subdivisions and root constructions

$$\mathcal{M}_{g,n}^{\mathrm{trop}}\langle m\rangle \to \mathcal{M}_{g,n}^{\mathrm{trop}}\langle m-1\rangle \to \cdots \to \mathcal{M}_{g,n}^{\mathrm{trop}}\langle 0\rangle = \mathcal{M}_{g,n}^{\mathrm{trop}}$$

such that (i) each arrow is a blowup at a stack theoretically smooth center and (ii) the pullback of $\mathsf{DR}_g^{\mathrm{trop}}(\mathsf{V}) \to \mathcal{M}_{g,n}^{\mathrm{trop}}$ to $\mathcal{M}_{g,n}^{\mathrm{trop}}\langle m\rangle$ is weakly semistable.

Lemma 4.9.1. *A sequence of blowups with the properties above exists.*

Proof. Sequentially blowup smooth strata of $\overline{\mathcal{M}}_{g,n}$ until the irreducible components of the boundary are each smooth and nonempty intersections of these irreducible strata are connected; this can be achieved by two barycentric subdivisions; see, for instance, [6, Section 4.6]. Once this is done, we use the fact that sequences of stellar subdivisions of a fan are cofinal in the system of all subdivisions. The result follows. □

As our goal is to describe the difference between the virtual strict and virtual total transforms, we begin by describing this difference on the base of the obstruction theories.

4.10. *The first blowup: subdivision.* Consider the diagram

$$\begin{tikzcd} t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle \arrow[r, hook] & t\mathsf{DR}_g(\mathsf{V})^{\mathsf{tot}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & t\mathsf{DR}_g(\mathsf{V}) \arrow[d, "\kappa"] \\ & \overline{\mathcal{M}}_{g,n}\langle 1\rangle \arrow[r, "\nu"] & \overline{\mathcal{M}}_{g,n} \end{tikzcd}$$

where the middle square is Cartesian and defines the total transform, while $t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle$ is the strict transform. Recall that the space $t\mathsf{DR}_g(\mathsf{V})$ is an open in a birational modification of $\overline{\mathcal{M}}_{g,n}$ and contains the double ramification space as a closed and proper substack (we do not yet involve the space $\mathsf{DR}_g(\mathsf{V})$ itself). In particular, aside from $t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle$, the remaining four spaces in the diagram are birational to each other. As explained in Section 3.7 there is still very nontrivial refined intersection theory on the total transform!

The morphism ν is a blowup morphism and its source and target are smooth Deligne–Mumford stacks. It is therefore a local complete intersection morphism. There is a refined Gysin pullback $\nu^!$ which we apply to obtain

$$\nu^![t\mathsf{DR}_g(\mathsf{V})] - [t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle] = [\underline{\mathsf{correction}}].$$

The term $[\underline{\mathsf{correction}}]$ is supported on the exceptional divisor. It is described by Fulton's blowup formula in its refined version [24, Example 6.7.1]. Let $Z\langle 1\rangle$ be the center of the blowup morphism ν. Denote the exceptional divisor by $\mathbb{P}\langle 1\rangle$. It is the projectivized normal bundle of the center, and can examine the bundle map

$$\mu : \mathbb{P}\langle 1\rangle \to Z\langle 1\rangle.$$

In Fulton's formula, the $[\underline{\mathsf{correction}}]$ term above is pushed forward from $\mathbb{P}\langle 1\rangle \times_{\overline{\mathcal{M}}_{g,n}\langle 1\rangle} \kappa^{-1}Z\langle 1\rangle$ to the total transform. The class on $\mathbb{P}\langle 1\rangle$ is the expected dimensional, that is homological degree $3g-3+n$. Precisely, it is the degree $3g-3+n$ piece in the product of the following two terms:

- the Chern class of the excess normal bundle — the quotient of the normal bundle of $Z\langle 1\rangle$ by the fiberwise hyperplane in $\mathbb{P}\langle 1\rangle$,
- the Segre class of the scheme-theoretic preimage $\kappa^{-1}(Z\langle 1\rangle)$ in $t\mathsf{DR}_g(\mathsf{V})$.

Let us interpret the product. The support of the Segre term is, by definition, contained in the fiber product $Z\langle 1\rangle \times_{\overline{\mathcal{M}}_{g,n}} t\mathsf{DR}_g(\mathsf{V})$. In a moment, we will explain how to handle this Segre class term using Aluffi's formula. For now, we blackbox the class. This fiber product maps to the center $Z\langle 1\rangle$ of the blowup, and therefore we can pullback the projective bundle μ to this fiber product. Now apply flat pullback to this Segre class term, along the bundle, and then apply Chern classes of the excess bundle. The result of a process is the correction term. As yet, there are still no virtual class considerations.

We come to the Segre class term and Aluffi's formula. A stratum in any logarithmic scheme is cut out by a monomial ideal sheaf. This class of ideals is stable under pullback by logarithmic morphism. The center of the blowup $Z\langle 1\rangle$ is such a stratum. It follows that the Segre class term $s(\kappa^{-1}Z\langle 1\rangle, t\mathsf{DR}_g(\mathsf{V}))$ is given by a monomial ideal. We can replace $t\mathsf{DR}_g(\mathsf{V})$ with a logarithmic modification without changing the problem, and therefore we can assume that $t\mathsf{DR}_g(\mathsf{V})$ is a simple normal crossings pair, i.e., tropically smooth.

Since the Segre class on a logarithmically and tropically smooth space comes from piecewise polynomials, we have the following.

Corollary 4.10.1. *The correction term comes from piecewise polynomials. Explicitly, it is a sum of classes that are obtained by the following three steps*:

(i) *Choose a closed stratum $W \subset t\mathsf{DR}_g(\mathsf{V})$ mapping to the center $Z\langle 1\rangle$ of the blowup.*

(ii) *Pullback the projective bundle μ to W and obtain a projective bundle $\mathbb{P}_W \to W$.*

(iii) *Apply to the fundamental class of the $\mathbb{P}_W$ a polynomial in the Chern roots of the normal bundle $N_{Z\langle 1\rangle/\overline{\mathcal{M}}_{g,n}}$, the Chern class of the relative hyperplane bundle $\mathbb{P}_W \to W$, and the Chern roots of the normal bundle $N_{W/t\mathsf{DR}_g(\mathsf{V})}$.*

4.11. ***The first blowup: virtual structure.*** The preceding arguments account for the purely tropical part of the argument. We now wish to calculate the pushforward of the virtual strict transform of $\mathsf{DR}_g(\mathsf{V})$ to the blowup $\overline{\mathcal{M}}_{g,n}\langle 1\rangle$. To achieve this, we apply virtual pullback to correction and push it down to $\overline{\mathcal{M}}_{g,n}\langle 1\rangle$. We begin by adding a row to our previous diagram

$$\begin{tikzcd}
\mathsf{DR}_g(\mathsf{V})\langle 1\rangle \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V})^{\mathsf{tot}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V}) \arrow[d] \\
t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle \arrow[r, hook] & t\mathsf{DR}_g(\mathsf{V})^{\mathsf{tot}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & t\mathsf{DR}_g(\mathsf{V}) \arrow[d, "\kappa"] \\
 & \overline{\mathcal{M}}_{g,n}\langle 1\rangle \arrow[r, "\nu"] & \overline{\mathcal{M}}_{g,n}
\end{tikzcd}$$

In the middle row, the setup of the previous section gives us

$$\nu^![t\mathsf{DR}_g(\mathsf{V})] - [t\mathsf{DR}_g(\mathsf{V})\langle 1\rangle] = [\underline{\mathsf{correction}}].$$

The equality is as classes on the total space of the pullback, namely $t\mathsf{DR}_g(\mathsf{V})^{\mathsf{tot}}$. The vertical arrows are all equipped with obstruction theories, and we pull [$\underline{\mathsf{correction}}$] up along the vertical arrows to obtain a *virtual* correction, and push it down to $\overline{\mathcal{M}}_{g,n}\langle 1\rangle$. By compatibility of Gysin maps [24, Proposition 6.6] on the top right square, the class of $\mathsf{DR}_g(\mathsf{V})\langle 1\rangle$ can be computed by pulling back the known formula on $\overline{\mathcal{M}}_{g,n}$ and subtracting the virtual correction term.

The term [$\underline{\mathsf{correction}}$] is supported over the blowup center. Furthermore, by Aluffi's formula, it is a sum of Chern class operators applied to strata of $t\mathsf{DR}_g(\mathsf{V})$ that map to the blowup center $Z\langle 1\rangle$. Let us now choose a single such summand. We may therefore fix a stratum W — it amounts to choosing a cone in the tropical moduli space of Section 4.4. Restricting attention here, we have

$$\begin{tikzcd}
\mathsf{DR}_W^{\mathsf{exc}}(\mathsf{V}) \arrow[r] \arrow[d, "\epsilon"] \arrow[dr, phantom, "\square"] & \mathsf{DR}_W(\mathsf{V}) \arrow[d, "\varphi"] \\
\mathbb{P}_W\langle 1\rangle \arrow[r] \arrow[d, "\tau"] \arrow[dr, phantom, "\square"] & W \arrow[d, "\kappa"] \\
\mathbb{P}\langle 1\rangle \arrow[r, "\varpi"] & Z\langle 1\rangle
\end{tikzcd}$$

The space $\mathsf{DR}_W(\mathsf{V})$ is simply the preimage of W in $\mathsf{DR}_g(\mathsf{V})$. The term $\mathsf{DR}_W^{\mathsf{exc}}(\mathsf{V})$ is defined by pullback. The key property is that it is a projective space bundle over the stratum $\mathsf{DR}_W(\mathsf{V})$.

Before proceeding, we observe that there are natural classes defined in different parts of this diagram. The center $Z\langle 1\rangle$ carries the Chern classes of its normal bundle. The exceptional divisor $\mathbb{P}\langle 1\rangle$ carries the fiberwise hyperplane class. The stratum W carries cohomology class that comes from piecewise polynomials, according to Aluffi's formula.

Proposition 4.11.1. *The pushforward to $\overline{\mathcal{M}}_{g,n}\langle 1\rangle$ of the difference of classes*

$$\nu^![\mathsf{DR}_g(\mathsf{V})] - [\mathsf{DR}_g(\mathsf{V})\langle 1\rangle]$$

is tautological. More precisely, it is the pushforward of a sum of terms that are supported on the exceptional divisor of ν, with each given by a product of tautological classes with powers of the normal bundle of the exceptional divisor.

Proof. We apply Fulton's blowup formula to $t\mathsf{DR}_g(\mathsf{V})$ as stated in [24, Example 6.7.1] to obtain the term [$\underline{\mathsf{correction}}$]. We perform virtual pullback and pushforward of this class to control it. By Aluffi's Segre class formula, we see that the classes of interest in Fulton's formula are sums over strata W in $t\mathsf{DR}_g(\mathsf{V})$ of classes that are obtained in the following steps:

(i) Choose any piecewise polynomial p on $t\mathsf{DR}_g(\mathsf{V})$ and pull it back to W.

(ii) Choose a polynomial c in the Chern classes of the normal bundle to $Z\langle 1\rangle$.

(iii) Perform the smooth pullback along ϖ of the Chow homology class $c \cup p \cap [W]$ to obtain a class on $\mathbb{P}_W\langle 1\rangle$.

(iv) Apply any polynomial in the Chern class of the hyperplane bundle on $\mathbb{P}_W\langle 1\rangle$.

(v) Perform virtual pullback along ϵ.

(vi) Perform proper pushforward along the composite morphism $\tau \circ \epsilon$.

We can absorb the terms p and c into a single cohomological term which we write as q. The hyperplane bundle is pulled back from $\mathbb{P}\langle 1\rangle$, and by applying the projection formula and commutativity of smooth and Gysin pullback, we can obtain the same class instead by *first* calculating the $\varphi \circ \kappa$ pushforward of the class $\varphi^!(q \cap [W])$, pulling back to $\mathbb{P}\langle 1\rangle$, applying operators coming from the hyperplane bundle and pushing forward to $\overline{\mathcal{M}}_{g,n}\langle 1\rangle$. We will prove that $\varphi^!(q \cap [W])$ pushes forward to a tautological class.

The virtual stratum $\mathsf{DR}_W(\mathsf{V})$ may be written as a product of smaller strata by the gluing formula for strata in Gromov–Witten theory of rubber targets; see, for instance, [44, Section 5.3] or [27, Section 3]. The class q is piecewise polynomial and therefore restricts to a tautological class on each of the relative stable maps spaces. We can now use the result of Faber and Pandharipande, that the pushforwards of tautological Gromov–Witten cycles, including those with disconnected domain curves, to the moduli space of curves are tautological [23, Theorem 2]. We conclude that the virtual strict transform under one blowup is tautological. □

4.12. *The inductive procedure.* The calculations performed for the first blowup guarantee that $\mathsf{DR}_g(\mathsf{V})\langle 1\rangle$ is equal to the pullback of a tautological class on the moduli space of curves, corrected by classes on strata that come from piecewise polynomials. Practically, these are pushforwards from strata of expressions in the Chern roots of the normal bundles to those strata.

We now climb up the tower of blowups of $\overline{\mathcal{M}}_{g,n}$. At each stage, we blowup a smooth center, and we calculate the strict transform exactly as in the previous section. Subtracting off the correction terms at each stage, eventually, we are left with the class $\mathsf{DR}_g^{\lightning}(\mathsf{V})$.

Notation. The nature of the inductive procedure means that this final section is notation heavy. We will use $\overline{\mathcal{M}}_{g,n}\langle k\rangle$ to denote the k-th blowup of the moduli space of curves, and $t\mathsf{DR}_g(\mathsf{V})\langle k\rangle$ to be the *strict* transform of $t\mathsf{DR}_g(\mathsf{V})$ under this blowup. The superscript $(-)^{\mathsf{tot}}$ is used for the *total* transform and $(-)^{\mathsf{exc}}$ is used for objects living over the exceptional divisor of the blowup.

The inductive setup. Let us now describe the inductive step after k blowups have been performed. The setup is

$$\begin{tikzcd}
\mathsf{DR}_g(\mathsf{V})\langle k+1\rangle \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V})\langle k\rangle^{\mathsf{tot}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & \mathsf{DR}_g(\mathsf{V})\langle k\rangle \arrow[d] \\
t\mathsf{DR}_g(\mathsf{V})\langle k+1\rangle \arrow[r, hook] & t\mathsf{DR}_g(\mathsf{V})\langle k\rangle^{\mathsf{tot}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\square"] & t\mathsf{DR}_g(\mathsf{V})\langle k\rangle \arrow[d, "\kappa"] \\
 & \overline{\mathcal{M}}_{g,n}\langle k+1\rangle \arrow[r, "\nu"] & \overline{\mathcal{M}}_{g,n}\langle k\rangle
\end{tikzcd}$$

We know from the earlier steps in the recursion that the class $\mathsf{DR}_g(\mathsf{V})\langle k\rangle$ pushes forward in $\overline{\mathcal{M}}_{g,n}\langle k\rangle$ to a class that differs from the pullback of a tautological class on the moduli space of curves by classes that come from piecewise polynomials and cotangent classes. This class is pulled back along the next blowup, which is a local complete intersection morphism. The pullback is of the virtual total transform, and we are interested in the virtual strict transform.

The virtual strict transform, as before, is controlled by Aluffi's Segre class formula. The correction term in the middle row is a sum of classes involving the hyperplane bundle on the exceptional divisor and classes on strata that come from piecewise polynomials.[14] By treating the terms one at a time, we may focus on a fixed stratum W of $t\mathsf{DR}_g(\mathsf{V})\langle k\rangle$.

The terms over a fixed stratum. Restrict the diagram above to a stratum W of $t\mathsf{DR}_g(\mathsf{V})\langle k\rangle$. We have, over the exceptional divisor, the diagram

$$\begin{tikzcd}
\mathsf{DR}_W^{\mathsf{exc}}(\mathsf{V})\langle k\rangle \arrow[r] \arrow[d, "\epsilon"] \arrow[dr, phantom, "\square"] & \mathsf{DR}_W(\mathsf{V})\langle k\rangle \arrow[d, "\varphi"] \\
\mathbb{P}_W\langle k\rangle \arrow[r, "g"] \arrow[d, "\tau"] \arrow[dr, phantom, "\square"] & W \arrow[d, "\kappa"] \\
\mathbb{P}\langle k\rangle \arrow[r, "\varpi"] & Z\langle k\rangle
\end{tikzcd}$$

The horizontal arrows are all projective bundles, and in particular are flat. We recall the definition of correction term in the blowup: (i) place a piecewise polynomial on W, (ii) apply Gysin pullback along φ, (iii) apply refined pullback along ϖ, and (iv) apply the Chern class in the excess normal bundle.

By compatibility, we can instead apply $\varphi^!$ to an operator on W and push it all the way to $Z\langle k\rangle$, before pulling back to the bundle and only then the excess operator. Therefore we examine the class $\varphi^!(p\cap[W])$ pushed forward to $Z\langle k\rangle$ where p is a piecewise polynomial. There is a sequence of maps

$$W \hookrightarrow t\mathsf{DR}_g(\mathsf{V})\langle k\rangle \to t\mathsf{DR}_g(\mathsf{V}).$$

The first map is an inclusion of a stratum. The composite morphism is a partially compactified torus bundle over a stratum in the second codomain. Denote this latter stratum by $W^\flat$ and write the composite morphism above as

$$W \to W^\flat.$$

By hypothesis, the stratum $W^\flat$ is irreducible, so its fundamental class pulls back to the fundamental class of W, using the diagonal pullback of [24, Section 8.1]. The stratum W is contained in $t\mathsf{DR}_g(\mathsf{V})$ and is therefore equipped with a refined virtual class by virtual pullback along the map $r : \mathsf{DR}_g(\mathsf{V}) \to t\mathsf{DR}_g(\mathsf{V})$. The correction is calculated by refined virtual pullback of this virtual class along $W \to W^\flat$, decorating by piecewise polynomials, and pushing forward to $\mathbb{P}\langle k\rangle$. The diagonal pullback commutes with the virtual pullback along $r : \mathsf{DR}_g(\mathsf{V}) \to t\mathsf{DR}_g(\mathsf{V})$, so we focus our attention on the class $r^![W^\flat]$.

[14] We repurpose the notation from the previous section.

Decomposing the stratum. The key fact that we will now use is that strata of the double ramification cycle can be expressed as finite quotients of products of smaller double ramification cycle. This allows us to run the induction; let us now provide the details.

The class $r^![W^\flat]$ is a virtual stratum in a space of relative stable maps, to which we may apply the gluing formula. By the discussion in Section 4.4 stratum $W^\flat$ corresponds uniquely to a stratum in the moduli space of maps to expansions of the rubber $\mathbb{P}^1$ geometry, relative to 0 and ∞. A stratum is indexed by a rubber combinatorial type $[\Gamma \to T]$ where as before, T is a line graph and Γ is a weighted dual graph. The combinatorial type fixes the ramification orders along all edges and half edges.

By cutting all the edges of the graph T above, we obtain a collection of tropical maps $[\Gamma_i \to T_i]$ where T_i is a 1-vertex graph with 2 legs, with decorations of all edge flags by slopes. The graph Γ_i determines a moduli space $\mathcal{M}_{\Gamma_i}$ of curves whose genus is the genus label at Γ and whose number of markings is equal to the number of flags at this vertex. A graph cover $\Theta = [\Gamma_i \to T_i]$ of this type determines a double ramification cycle problem.

We obtain a product decomposition, parallel to Section 4.8,

$$W^\flat \to W^\flat_{\Gamma_1} \times \cdots \times W^\flat_{\Gamma_r}/\mathsf{Aut}(\Theta),$$

exhibiting the stratum as a torus bundle, as we explained in that section. These come with numerical specifications $\mathsf{U}_1, \ldots, \mathsf{U}_r$ of contact orders for a smaller double ramification problem. We choose any compactification of the map with smooth source and target space to obtain

$$\overline{W}^\flat \to \overline{W}^\flat_{\Gamma_1} \times \cdots \times \overline{W}^\flat_{\Gamma_r}/\mathsf{Aut}(\Theta).$$

The external product on Chow groups gives rise to a *split DR cycle* that is picked out individually on the factors

$$\mathsf{DR}_g(\mathsf{U}_1) \times \cdots \times \mathsf{DR}_g(\mathsf{U}_r) \in \mathsf{CH}_\star(\overline{W}^\flat; \mathbb{Q}).$$

On the other hand, $\overline{W}^\flat$ also carries a double ramification cycle: the virtual pullback on the stratum $W^\flat$ under the Abel–Jacobi morphism produces a refined class which we pushforward to the compact $\overline{W}^\flat$.

We come to the final result. Fix a sequence of blowups giving rise to a morphism

$$\nu_m : \overline{\mathcal{M}}_{g,n}\langle m\rangle \to \overline{\mathcal{M}}_{g,n}.$$

Theorem 4.12.1. *The class* $[\mathsf{DR}_g(\mathsf{V})\langle m\rangle]$ *in the Chow ring of* $\overline{\mathcal{M}}_{g,n}\langle m\rangle$ *of the virtual strict transform lies in the subring generated by the pullbacks of tautological classes from the moduli space of curves and Chern roots of the normal bundles to the strata of* $\overline{\mathcal{M}}_{g,n}\langle m\rangle$*. In particular, the class* $[\mathsf{DR}_g^{\lightning}(\mathsf{V})]$ *is tautological.*

Proof. First note that in genus 0 the lightning class is uncorrected for all contact data. In all genus, the class is uncorrected when V is the 0 vector or when it is $(1, -1)$. Consider a proper substratum of the double ramification cycle. Each proper stratum of $t\mathsf{DR}_g(\mathsf{V})$ gives rise to smaller double ramification cycle problems — either the genus or the total number of markings decreases — and we may inductively assume that all the associated lightning classes are known to be tautological.

We induct on the number of blowups. We have completed the base case, and come to the inductive step. By the discussion above, the correction class is the pushforward of a sum of classes supported on the exceptional divisor of the blowup. Each of these classes corresponds to a stratum and we call it W. As before, we have

$$\begin{tikzcd}
\mathsf{DR}^{\mathsf{exc}}_W(\mathsf{V})\langle k+1\rangle \arrow[r] \arrow[d, "\epsilon"] \arrow[dr, phantom, "\square"] & \mathsf{DR}_W(\mathsf{V})\langle k+1\rangle \arrow[d, "\varphi"] \\
\mathbb{P}_W\langle k+1\rangle \arrow[r, "g"] \arrow[d, "\tau"] \arrow[dr, phantom, "\square"] & W \arrow[d, "\kappa"] \\
\mathbb{P}\langle k+1\rangle \arrow[r, "\varpi"] & Z\langle k+1\rangle
\end{tikzcd}$$

and it will suffice for us to show that if we start with a class of the form $p \cap [W]$, apply the virtual pullback along φ, followed by the pushforward to $Z\langle k+1\rangle$, the result is a tautological class. As noted already, the refined class $\varphi^![W]$ is pulled back from the corresponding refined class on $W^\flat \subset t\mathsf{DR}_g(\mathsf{V})$. We compactify, as above,

$$\overline{W} \to \overline{W}^\flat$$

and observe that any sufficiently fine compactification of $\overline{W}$ will map to $Z\langle k+1\rangle$. It suffices then to show that the following class on $\overline{W}$ is inductively tautological: We do this as follows: (i) compute the refined virtual class on $W^\flat$ obtained via the virtual pullback $r : \mathsf{DR}_g(\mathsf{V}) \to t\mathsf{DR}_g(\mathsf{V})$; (ii) pushforward to a sufficiently fine compactification $\overline{W}^\flat$; (iii) pullback along $\overline{W} \to \overline{W}^\flat$; (iv) apply a piecewise polynomial operator p.

The operator p can be arranged to be the restriction of an operator on the compactification, which we continue to denote as p. It will therefore suffice to show that the virtual class on $\overline{W}^\flat$ is tautological. By the discussion in Section 4.8, and the weak factorization theorem, we may connect the compactification $\overline{W}^\flat$ by a sequence of blowups and blowdowns along smooth centers, to the compactification studied in Construction 4.8.2. We recall that this construction outputs a projective bundle,

$$\mathsf{kut} : \widetilde{\mathcal{M}}_{[\Gamma \to T]} \to \prod_{u \in V(\Gamma)} \overline{\mathcal{M}}(u).$$

The source carries a double ramification class, as it contains $W^\flat$, and we can pushforward the refined virtual class here. As this class is stable under pullbacks along blowups, it suffices to show that the double ramification class on $\widetilde{\mathcal{M}}_{[\Gamma \to T]}$ is tautological. We apply Proposition 4.8.3, which reduces the question to the smaller double ramification problems associated to the vertices of Γ. These are known to be tautological inductively, and we conclude the result. □

Corollary 4.12.2. *The toric contact cycles lie in the tautological ring of the moduli space of curves.*

Proof. We find a sequence of blowups $\overline{\mathcal{M}}_{g,n}\langle m\rangle \to \overline{\mathcal{M}}_{g,n}$ such that all double ramification cycles have been transversalized. By the projective bundle formula and Fulton's key formula [24, Theorem 3.3 and

Proposition 6.7], under each blowup morphism a polynomial in the Chern roots of the normal bundle to a stratum pushes forward to a class spanned by tautological classes from the moduli space of curves and piecewise polynomial operators applied to the strata. These push forward to tautological classes on the moduli space of curves, and we conclude the main result. □

Acknowledgements

The idea originates from discussions with Davesh Maulik concerning the degeneration formula in Gromov–Witten theory. Ranganathan is especially grateful to Navid Nabijou for a collaboration in which this type of study was first implemented and lessons were learned [49] and to Tom Graber and Rahul Pandharipande for many useful conversations about localization and the toric contact cycles in the summer of 2019. Discussions with Paolo Aluffi, Younghan Bae, Luca Battistella, Renzo Cavalieri, David Holmes, Johannes Schmitt, and Jonathan Wise have been important and inspiring. We are grateful to Dan Abramovich and Sam Payne for countless conversations that have impacted our understanding of the tropical and logarithmic geometry. We thank the referee for comments and corrections.

The work here was presented at the ETH Algebraic Geometry and Moduli seminar, the working seminar on bChow intersections via subdivisions, and at the Stanford Algebraic Geometry Seminar. We thank the organizers and participants for enthusiasm and for numerous discussions.

Molcho was supported by ERC-2017-AdG-786580-MACI. This project has received funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation program (grant agreement no. 786580).

References

[1] D. Abramovich and K. Karu, "Weak semistable reduction in characteristic 0", *Invent. Math.* **139**:2 (2000), 241–273. MR Zbl

[2] D. Abramovich and J. Wise, "Birational invariance in logarithmic Gromov–Witten theory", *Compos. Math.* **154**:3 (2018), 595–620. MR Zbl

[3] D. Abramovich, K. Karu, K. Matsuki, and J. Włodarczyk, "Torification and factorization of birational maps", *J. Amer. Math. Soc.* **15**:3 (2002), 531–572. MR Zbl

[4] D. Abramovich, L. Caporaso, and S. Payne, "The tropicalization of the moduli space of curves", *Ann. Sci. École Norm. Sup.* (4) **48**:4 (2015), 765–809. MR Zbl

[5] D. Abramovich, Q. Chen, S. Marcus, M. Ulirsch, and J. Wise, "Skeletons and fans of logarithmic structures", pp. 287–336 in *Nonarchimedean and tropical geometry* (Saint John, USVI, 2013/Puerto Rico, 2015), edited by M. Baker and S. Payne, Springer, 2016. MR Zbl

[6] D. Abramovich, Q. Chen, S. Marcus, and J. Wise, "Boundedness of the space of stable logarithmic maps", *J. Eur. Math. Soc.* **19**:9 (2017), 2783–2809. MR Zbl

[7] K. Adiprasito, G. Liu, and M. Temkin, "Semistable reduction in characteristic 0", *Sém. Lothar. Combin.* **82B** (2020), art. id. 25. MR Zbl

[8] P. Aluffi, "Modification systems and integration in their Chow groups", *Selecta Math.* (*N.S.*) **11**:2 (2005), 155–202. MR Zbl

[9] P. Aluffi, "Segre classes of monomial schemes", *Electron. Res. Announc. Math. Sci.* **20** (2013), 55–70. MR Zbl

[10] P. Aluffi, "Segre classes as integrals over polytopes", *J. Eur. Math. Soc.* **18**:12 (2016), 2849–2863. MR Zbl

[11] D. Anderson and S. Payne, "Operational K-theory", *Doc. Math.* **20** (2015), 357–399. MR Zbl

[12] Y. Bae and H. Park, "A comparison theorem for cycle theories for algebraic stacks", in preparation.

[13] Y. Bae, D. Holmes, R. Pandharipande, J. Schmitt, and R. Schwarz, "Pixton's formula and Abel–Jacobi theory on the Picard stack", *Acta Math.* **230**:2 (2023), 205–319. MR Zbl

[14] L. J. Barrott, "Logarithmic Chow theory", preprint, 2018. arXiv 1810.03746

[15] C. Böhning, H.-C. G. von Bothmer, and M. van Garrel, "Prelog Chow rings and degenerations", *Rend. Circ. Mat. Palermo* (2) **72**:5 (2023), 2861–2894. MR Zbl

[16] M. Brion, "Piecewise polynomial functions, convex polytopes and enumerative geometry", pp. 25–44 in *Parameter spaces* (Warsaw, 1994), edited by P. Pragacz, Banach Center Publ. **36**, Polish Acad. Sci. Inst. Math., Warsaw, 1996. MR Zbl

[17] R. Cavalieri, H. Markwig, and D. Ranganathan, "Tropical compactification and the Gromov–Witten theory of $\mathbb{P}^1$", *Selecta Math.* (*N.S.*) **23**:2 (2017), 1027–1060. MR Zbl

[18] R. Cavalieri, M. Chan, M. Ulirsch, and J. Wise, "A moduli stack of tropical curves", *Forum Math. Sigma* **8** (2020), art. id. e23. MR Zbl

[19] R. Cavalieri, H. Markwig, and D. Ranganathan, "Pluricanonical cycles and tropical covers", 2022. To appear in *Trans. Amer. Math. Soc.* arXiv 2206.14034

[20] Q. Chen, "The degeneration formula for logarithmic expanded degenerations", *J. Algebraic Geom.* **23**:2 (2014), 341–392. MR Zbl

[21] Y.-C. Chou, L. Herr, and Y.-P. Lee, "The log product formula in quantum K-theory", *Math. Proc. Cambridge Philos. Soc.* **175**:2 (2023), 225–252. MR Zbl

[22] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Grad. Stud. Math. **124**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[23] C. Faber and R. Pandharipande, "Relative maps and tautological classes", *J. Eur. Math. Soc.* **7**:1 (2005), 13–49. MR Zbl

[24] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Math. (3) **2**, Springer, 1998. MR Zbl

[25] W. Fulton and B. Sturmfels, "Intersection theory on toric varieties", *Topology* **36**:2 (1997), 335–353. MR Zbl

[26] J. L. González and K. Karu, "Bivariant algebraic cobordism", *Algebra Number Theory* **9**:6 (2015), 1293–1336. MR Zbl

[27] T. Graber and R. Vakil, "Relative virtual localization and vanishing of tautological classes on moduli spaces of curves", *Duke Math. J.* **130**:1 (2005), 1–37. MR Zbl

[28] J. Guéré, "A generalization of the double ramification cycle via log-geometry", preprint, 2016. arXiv 1603.09213

[29] L. Herr, "The log product formula", *Algebra Number Theory* **17**:7 (2023), 1281–1323. MR Zbl

[30] D. Holmes, "Extending the double ramification cycle by resolving the Abel–Jacobi map", *J. Inst. Math. Jussieu* **20**:1 (2021), 331–359. MR Zbl

[31] D. Holmes and R. Schwarz, "Logarithmic intersections of double ramification cycles", *Algebr. Geom.* **9**:5 (2022), 574–605. MR Zbl

[32] D. Holmes, A. Pixton, and J. Schmitt, "Multiplicativity of the double ramification cycle", *Doc. Math.* **24** (2019), 545–562. MR Zbl

[33] D. Holmes, S. Molcho, R. Pandharipande, A. Pixton, and J. Schmitt, "Logarithmic double ramification cycles", preprint, 2022. arXiv 2207.06778

[34] F. Janda, "Gromov–Witten theory of target curves and the tautological ring", *Michigan Math. J.* **66**:4 (2017), 683–698. MR Zbl

[35] F. Janda, R. Pandharipande, A. Pixton, and D. Zvonkine, "Double ramification cycles on the moduli spaces of curves", *Publ. Math. Inst. Hautes Études Sci.* **125** (2017), 221–266. MR Zbl

[36] T. Kajiwara, K. Kato, and C. Nakayama, "Logarithmic abelian varieties, VII: Moduli", *Yokohama Math. J.* **67** (2021), 9–48. MR Zbl

[37] E. Katz and S. Payne, "Piecewise polynomials, Minkowski weights, and localization on toric varieties", *Algebra Number Theory* **2**:2 (2008), 135–155. MR Zbl

[38] G. Kempf, F. F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings, I*, Lecture Notes in Math. **339**, Springer, 1973. MR Zbl

[39] S.-i. Kimura, “Fractional intersection and bivariant theory”, *Comm. Algebra* **20**:1 (1992), 285–302. MR Zbl

[40] S. Marcus and J. Wise, “Logarithmic compactification of the Abel–Jacobi section”, *Proc. Lond. Math. Soc.* (3) **121**:5 (2020), 1207–1250. MR Zbl

[41] D. Maulik and D. Ranganathan, “Logarithmic Donaldson–Thomas theory”, *Forum Math. Pi* **12** (2024), art. id. e9. MR Zbl

[42] S. Molcho, “Universal stacky semistable reduction”, *Israel J. Math.* **242**:1 (2021), 55–82. MR Zbl

[43] S. Molcho, “Pullbacks of Brill–Noether classes under Abel–Jacobi sections”, preprint, 2022. arXiv 2212.14368

[44] S. Molcho and E. Routis, “Localization for logarithmic stable maps”, *Trans. Amer. Math. Soc. Ser. B* **6** (2019), 80–113. MR Zbl

[45] S. Molcho and J. Wise, “The logarithmic Picard group and its tropicalization”, *Compos. Math.* **158**:7 (2022), 1477–1562. MR Zbl

[46] S. Molcho, R. Pandharipande, and J. Schmitt, “The Hodge bundle, the universal 0-section, and the log Chow ring of the moduli space of curves”, *Compos. Math.* **159**:2 (2023), 306–354. MR Zbl

[47] D. Mumford, “An analytic construction of degenerating abelian varieties over complete rings”, *Compos. Math.* **24** (1972), 239–272. MR Zbl

[48] N. Nabijou, *Recursion formulae in Gromov–Witten theory and quasimap theory*, Ph.D. thesis, Imperial College London, 2018, available at https://webspace.maths.qmul.ac.uk/n.nabijou/NabijouThesis.pdf.

[49] N. Nabijou and D. Ranganathan, “Gromov–Witten theory with maximal contacts”, *Forum Math. Sigma* **10** (2022), art. id. e5. MR Zbl

[50] J. Oesinghaus, “Quasisymmetric functions and the Chow ring of the stack of expanded pairs”, *Res. Math. Sci.* **6**:1 (2019), art. id. 5. MR Zbl

[51] M. C. Olsson, “Logarithmic geometry and algebraic stacks”, *Ann. Sci. École Norm. Sup.* (4) **36**:5 (2003), 747–791. MR Zbl

[52] S. Payne, “Equivariant Chow cohomology of toric varieties”, *Math. Res. Lett.* **13**:1 (2006), 29–41. MR Zbl

[53] D. Ranganathan, “A note on the cycle of curves in a product of pairs”, preprint, 2019. arXiv 1910.00239

[54] D. Ranganathan, “Gromov–Witten theory and logarithmic intersection theory”, lecture notes, ETH Zürich, 2020, available at https://www.dpmms.cam.ac.uk/∼dr508/ETHApril2020.pdf.

[55] D. Ranganathan, “Logarithmic Gromov–Witten theory with expansions”, *Algebr. Geom.* **9**:6 (2022), 714–761. MR Zbl

[56] D. Ranganathan and A. U. Kumaran, “Logarithmic Gromov–Witten theory and double ramification cycles”, *J. Reine Angew. Math.* **809** (2024), 1–40. MR Zbl

[57] “The Stacks project”, electronic reference, 2005–, available at http://stacks.math.columbia.edu.

[58] J. Tevelev, “Compactifications of subvarieties of tori”, *Amer. J. Math.* **129**:4 (2007), 1087–1104. MR Zbl

[59] A. Thuillier, “Géométrie toroïdale et géométrie analytique non archimédienne: application au type d’homotopie de certains schémas formels”, *Manuscripta Math.* **123**:4 (2007), 381–451. MR Zbl

[60] M. Ulirsch, “Functorial tropicalization of logarithmic schemes: the case of constant coefficients”, *Proc. Lond. Math. Soc.* (3) **114**:6 (2017), 1081–1113. MR Zbl

[61] M. Ulirsch, “A non-Archimedean analogue of Teichmüller space and its tropicalization”, *Selecta Math.* (*N.S.*) **27**:3 (2021), art. id. 39. MR Zbl

[62] M. Ulirsch and D. Zakharov, “Tropical double ramification loci”, preprint, 2019. arXiv 1910.01499

Communicated by Martin Olsson
Received 2021-07-14 Revised 2023-06-02 Accepted 2023-10-31

samouil.molcho@math.ethz.ch — *Department of Mathematics, ETH Zürich, Zürich, Switzerland*

dr508@cam.ac.uk — *Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, United Kingdom*

ALGEBRA AND NUMBER THEORY 18:10 (2024)
https://doi.org/10.2140/ant.2024.18.1817

Spectral moment formulae for GL(3) x GL(2) L-functions I: The cuspidal case

Chung-Hang Kwan

Spectral moment formulae of various shapes have proven very successful in studying the statistics of central L-values. We establish, in a completely explicit fashion, such formulae for the family of $\mathrm{GL}(3) \times \mathrm{GL}(2)$ Rankin–Selberg L-functions using the period integral method. Our argument does not rely on either the Kuznetsov or Voronoi formulae. We also prove the essential analytic properties and derive explicit formulae for the integral transform of our moment formulae. We hope that our method will provide deeper insights into moments of L-functions for higher-rank groups.

1. Introduction

1A. ***Background.*** The study of L-values at the central point $s = \frac{1}{2}$ has taken center stage in many branches of number theory over the past decades due to their profound arithmetic significance. A variety of perspectives have enriched our understanding of the nature of central L-values. In particular, a statistical perspective can offer valuable insights. Fundamental questions in this direction include the determination of (non)vanishing and sizes of these L-values. An effective way to approach problems of this sort is via *moments of L-functions*. Techniques from analytic number theory have proven fruitful in estimating the sizes of moments of all kinds. Moreover, spectacular results can be obtained when moment estimates join forces with arithmetic geometry and automorphic representations.

This line of investigation is nicely exemplified by the landmark result of [Conrey and Iwaniec 2000]. Let χ be a real primitive Dirichlet character $\pmod q$ with q odd and square-free. The main object of [loc. cit.] is the cubic moment of $\mathrm{GL}(2)$ automorphic L-functions of the congruence subgroup $\Gamma_0(q)$ twisted by χ. An *upper bound* of Lindelöf strength in the q-aspect was established therein. When combining this upper bound with the celebrated formula of [Waldspurger 1981], the famous Burgess $\frac{3}{16}$-bound for Dirichlet L-functions was improved for the first time since the 1960's. In fact, Conrey and Iwaniec [2000] proved the bound

$$L\left(\tfrac{1}{2}, \chi\right) \ll_\epsilon q^{1/6+\epsilon}. \tag{1-1}$$

This research is supported by the EPSRC grant: EP/W009838/1.
MSC2020: primary 11F55; secondary 11F72.
Keywords: automorphic forms, automorphic L-functions, Maass forms, moments of L-functions, Rankin–Selberg L-functions, period integrals, Whittaker functions, hypergeometric functions, Poincare series.

Understanding the effects of a sequence of intricate arithmetic and analytic transformations constitutes a significant part of moment calculations as seen in [Conrey and Iwaniec 2000]. Surprisingly, such a sequence of [loc. cit.] ends up in a single elegant *identity* showcasing a duality between the cubic average over a basis of GL(2) automorphic forms (Maass or holomorphic) and the fourth moment of GL(1) L-functions. This remarkable phenomenon was uncovered relatively recently in [Petrow 2015]. His work consists of new elaborate analysis (see also [Young 2017]) building upon the foundation of [Conrey and Iwaniec 2000]. Further contributions to this topic include those in [Frolenkov 2020] and the earlier works [Ivić 2001; 2002], which studied other aspects of the problem. In its basic form, the identity roughly takes the shape

$$\sum_{f:\mathrm{GL}(2)\text{-Maass/Holomorphic}} L\left(\tfrac{1}{2}, f\right)^3 = \int_{-\infty}^{\infty} \left|\zeta\left(\tfrac{1}{2}+it\right)\right|^4 dt + (***), \tag{1-2}$$

where the weight functions for the moments are suppressed and $(***)$ represents certain polar contributions.

Besides its structural elegance, the identity (1-2) comes with immediate applications. It leads to sharp moment estimates as a consequence of exact evaluation. As an extra benefit, it streamlines the analysis in the traditional, approximate approach. In [Petrow 2015], this identity was referred to as a *"Motohashi-type identity"*. Previously, Motohashi [1993; 1997] discovered a similar identity but with the test function chosen on the fourth moment side, i.e., in the reverse direction of [Conrey and Iwaniec 2000; Petrow 2015; Young 2017; Ivić 2001; 2002]. It greatly enhances our understanding of the fourth moment of the ζ-function. The recent works [Young 2011; Blomer et al. 2020; Topacogullari 2021; Kaneko 2022] have extended Motohashi's work to Dirichlet L-functions.

Conrey and Iwaniec [2000, Introduction] further envisioned the possibilities and challenges of extending their method to a setting involving a GL(3) automorphic form. This is natural because the cubic moment of GL(2) L-functions can be regarded as the first moment of GL(3) × GL(2) Rankin–Selberg L-functions, averaged over a basis of GL(2) automorphic forms, where the GL(3) automorphic form is a minimal parabolic Eisenstein series. It is anticipated that advances in harmonic analysis of GL(3) could provide new perspectives towards the Conrey–Iwaniec method. Furthermore, the GL(3) set-up introduces an important new example: the first moment for the GL(3) × GL(2) family involving a GL(3) cusp form, which necessitates the use of genuine GL(3) techniques.

In the decade following [Conrey and Iwaniec 2000], two key breakthroughs made this extension possible for GL(3). Firstly, the GL(3) *Voronoi formula* was developed in [Miller and Schmid 2006] (see also [Goldfeld and Li 2006; Ichino and Templier 2013]), making it usable for a variety of analytic applications. Notably, the Hecke combinatorics of GL(3) associated to twisting and ramifications are considerably more involved than the classical GL(2) counterpart. Secondly, the GL(3) Voronoi formula was successfully applied in [Li 2011] together with new techniques to obtain strong *upper bounds* for the first moment of GL(3) × GL(2) Rankin–Selberg L-functions in the GL(2) spectral aspect. As a corollary, she obtained the first instance of subconvexity for GL(3) automorphic L-functions.

1B. *Main results.* The purpose of this article is to further the investigation of $\mathrm{GL}(3) \times \mathrm{GL}(2)$ moments of L-functions. However, we will depart from the standard approaches in the existing literature. We are interested in understanding the *intrinsic mechanisms* and examining the *essential ingredients* that lead directly to the complete structure of these moments, including both main terms and off-diagonals. Addressing these aspects carefully is crucial for enabling generalizations to higher-rank groups. We find that the formalism of *period integrals* for GL(3) is particularly effective in achieving these objectives.

We are ready to state the main result of this article, which is the Motohashi type moment identity behind the work of Li [2011].

Theorem 1.1. *Let*:

- Φ *be a fixed, Hecke-normalized Maass cusp form of* $\mathrm{SL}_3(\mathbb{Z})$ *with the Langlands parameters* $(\alpha_1, \alpha_2, \alpha_3) \in (i\mathbb{R})^3$, *and* $\widetilde{\Phi}$ *be the dual form of* Φ.
- $(\phi_j)_{j=1}^{\infty}$ *be an orthogonal basis of* ***even***, *Hecke-normalized Maass cusp forms of* $\mathrm{SL}_2(\mathbb{Z})$ *which satisfy*
$$\Delta\phi_j = \left(\tfrac{1}{4} - \mu_j^2\right)\phi_j.$$
- $L(s, \phi_j \otimes \Phi)$ *and* $L(s, \Phi)$ *be the Rankin–Selberg L-function of the pair* (ϕ_j, Φ) *and the standard L-function of* Φ *respectively*, *where* Λ *denotes the corresponding complete L-functions.*
- $\mathcal{C}_\eta$ $(\eta > 40)$ *be the class of holomorphic functions* H *defined on the vertical strip* $|\mathrm{Re}\,\mu| < 2\eta$ *such that* $H(\mu) = H(-\mu)$ *and has rapid decay*
$$H(\mu) \ll e^{-2\pi|\mu|} \quad (|\mathrm{Re}\,\mu| < 2\eta).$$
- *For* $H \in \mathcal{C}_\eta$, $(\mathcal{F}_\Phi H)(s_0, s)$ *is the integral transform defined in* (7-6) *and it only depends on the Langlands parameters of* Φ.

Then on the domain $\frac{1}{4} + \frac{1}{200} < \sigma < \frac{3}{4}$, *we have the following moment identity*:

$$\sum_{j=1}^{\infty} H(\mu_j)\frac{\Lambda(s, \phi_j \otimes \widetilde{\Phi})}{\langle \phi_j, \phi_j\rangle} + \int_{(0)} H(\mu)\frac{\Lambda(s+\mu, \widetilde{\Phi})\Lambda(1-s+\mu, \Phi)}{|\Lambda(1+2\mu)|^2}\frac{d\mu}{4\pi i}$$
$$= \frac{\pi^{-3s}}{2}L(2s, \Phi)\int_{(0)} \frac{H(\mu)}{|\Gamma(\mu)|^2}\cdot\prod_{i=1}^{3}\Gamma\left(\frac{s+\mu-\alpha_i}{2}\right)\Gamma\left(\frac{s-\mu-\alpha_i}{2}\right)\frac{d\mu}{2\pi i}$$
$$+ \tfrac{1}{2}L(2s-1, \Phi)(\mathcal{F}_\Phi H)(2s-1, s) \tag{1-3}$$
$$+ \tfrac{1}{2}\int_{(1/2)} \zeta(2s-s_0)L(s_0, \Phi)(\mathcal{F}_\Phi H)(s_0, s)\frac{ds_0}{2\pi i}.$$

The function $s \mapsto (\mathcal{F}_\Phi H)(2s-1, s)$ *can be computed explicitly, see* Theorem 1.2 *below.*

The temperedness assumption $(\alpha_1, \alpha_2, \alpha_3) \in (i\mathbb{R})^3$ for our fixed Maass cusp form Φ is very mild — it merely serves as a simplification of our exposition (when applying Stirling's formula in Section 8) and can be removed with a little more effort. In fact, all Maass cusp forms of $\mathrm{SL}_3(\mathbb{Z})$ are conjectured to be tempered and it was proved in [Miller 2001] that the nontempered forms constitute a density zero set.

We have made no attempt to enlarge the class of test functions for Theorem 1.1 since this is not the focus of this article (but is certainly doable by more refined analysis). The regularity assumptions of $\mathcal{C}_\eta$ essentially follow from those of the Kontorovich–Lebedev inversion (see Section 5B). As in [Goldfeld and Kontorovich 2013; Goldfeld et al. 2021; 2022; Buttcane 2020], the class $\mathcal{C}_\eta$ already includes good test functions that are useful in a number of applications and allows us to deduce a version of Theorem 1.1 for incomplete L-functions (see Remark 5.27).

Also, we have obtained the analytic properties and several explicit expressions for the integral transform $(\mathcal{F}_\Phi H)(s_0, s)$. They are written in terms of Mellin–Barnes integrals or hypergeometric functions as in [Motohashi 1993; 1997]. We do not record the full formulae here but refer the readers to Section 10 for the detailed discussions. However, we record an interesting identity of special functions as follow:

Theorem 1.2 (Theorem 10.2). *For* $\frac{1}{2} + \frac{1}{100} < \sigma < 1$, *we have*

$$(\mathcal{F}_\Phi H)(2s-1, s) = \pi^{\frac{1}{2}-s} \prod_{i=1}^{3} \frac{\Gamma\left(s - \frac{1}{2} + \frac{\alpha_i}{2}\right)}{\Gamma\left(1 - s - \frac{\alpha_i}{2}\right)} \cdot \int_{(0)} \frac{H(\mu)}{|\Gamma(\mu)|^2} \cdot \prod_{i=1}^{3} \prod_{\pm} \Gamma\left(\frac{1 - s + \alpha_i \pm \mu}{2}\right) \frac{d\mu}{2\pi i}. \tag{1-4}$$

There are actually two additional identities of Barnes type that account for the origins and the combinatorics of six (out of eight) of the off-diagonal main terms for the cubic moment of GL(2) L-functions. The results align nicely with the predictions of the "*moment conjecture*" (or "*recipe*") of [Conrey et al. 2005]. We refer the interested readers to our papers [Kwan 2023; 2024].

1C. ***Follow-up works.*** The current work aims to illustrate the key ideas and address the main analytic issues of our period integral approach. It is the simplest to illustrate all these using the cuspidal case for Φ. However, this is by no means the end of the scope of our method. In our upcoming works [Kwan 2023; 2024], we demonstrate the versatility of our method by:

(1) Providing a new proof of the cubic moment identity (1-2) (actually for the more general "*shifted moment*") with a number of technical advantages, as well as a new unified way of extracting the full set of main terms. There are considerable recent interests in understanding the deep works of [Motohashi 1993; 1997] and [Conrey and Iwaniec 2000] from different perspectives, e.g., [Nelson 2019; Wu 2022; Balkanova et al. 2021].

(2) Establishing a Motohashi's formula of GL(3) in the nonarchimedean aspect which dualizes GL(2) twists of Hecke eigenvalues into GL(1) twists by Dirichlet characters. This offers insights into the celebrated works [Young 2011; Blomer et al. 2020] on the fourth moment of Dirichlet L-functions. In their works, this kind of change of structures was the result of a long sequence of spectral/harmonic transformations and it was surprising (and useful) to observe such a nice phenomenon.

2. Outline

In Section 3, we discuss the technical features of the method used in this article and draw comparisons with the current literature. In Section 4, we include a sketch of our arguments to demonstrate the essential ideas of our method and sidestep the technical points. In Section 5, we collect the essential notions and results for later parts of the article.

The proof of Theorem 1.1 is divided into four sections. In Section 6, we prove the key identity of this article (see Corollary 6.2). In Section 7, we develop such an identity into moments of L-functions on the region of absolute convergence. In particular, the intrinsic structure of the problem allows one to easily see the shape of the dual moment (see Proposition 7.2). In Section 8, we obtain the region of holomorphy and growth of the archimedean transform. In Section 9, a step-by-step analytic continuation argument is performed based on the analytic information obtained in Section 8.

In Section 10, we prove Theorem 1.2. and provide several explicit formulae of the integral transforms.

3. Technical features of our method

3A. ***Period reciprocity.*** Our work adds a new instance to the recent banner *"period reciprocity"* which seeks to uncover the underlying structures of moments of L-functions through the lenses of period integrals. The general philosophy of this method is to evaluate a period integral in two distinct manners. Under favorable circumstances, the intrinsic structures of period integrals would lead to interesting, nontrivial moment identities, say connecting two different-looking families of L-functions.

In our case, the generalized Motohashi-type phenomenon of Theorem 1.1 at $s = \frac{1}{2}$ will be shown to be an intrinsic property of a given Maass cusp form Φ of $\mathrm{SL}_3(\mathbb{Z})$ via the following trivial identity

$$\int_0^1 \left[\int_0^\infty \Phi\left(\begin{pmatrix} y_0 & & \\ & y_0 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix} \right) d^\times y_0 \right] e(-u)\, du = \int_0^\infty \left[\int_0^1 \Phi\left(\begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 & & \\ & y_0 & \\ & & 1 \end{pmatrix} \right) e(-u)\, du \right] d^\times y_0. \tag{3-1}$$

Roughly speaking, Theorem 1.1 follows from (1) spectrally expanding the innermost integral on the left in terms of a basis of GL(2) automorphic forms, and (2) computing the innermost integral on the right in terms of the GL(3) Fourier–Whittaker period. A sketch of this will be provided in Section 4. In practice, it turns out to be convenient to work with a more general set-up

$$\int_{\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})} P(g; h)\Phi \begin{pmatrix} g & \\ & 1 \end{pmatrix} |\det g|^{s-1/2}\, dg \tag{3-2}$$

so as to bypass certain technical difficulties, where $P(*; h)$ is a Poincaré series of $\mathrm{SL}_2(\mathbb{Z})$.

The current examples for period reciprocity occur rather sporadically and there is currently no systematic method for constructing new examples. Also, techniques differ greatly in each known instance; see [Michel and Venkatesh 2006; 2010; Nelson 2019; Blomer 2012a; Nunes 2023; Jana and Nunes 2021; Zacharias 2021; 2019]. This stands in stark contrast to the more traditional "Kuznetsov–Voronoi" framework (see Section 3B). However, period reciprocity seems to address some of the technical complications more softly than the Kuznetsov–Voronoi approach. We shall elaborate more in the upcoming subsections.

Regarding the "classical" Motohashi phenomenon (1-2), a strategy was very recently proposed in [Michel and Venkatesh 2006; 2010] that developed into a fully rigorous method by Nelson [2019] through the use of regularized period integrals, incorporating new insights from automorphic representations. This article provides an alternative approach, which not only includes (1-2) but also generalizes several related instances of this phenomenon. We address the structural and analytic aspects of the formulae rather differently using unipotent integration for GL(3) and method of analytic continuation. (We begin by considering (3-2) for $\operatorname{Re} s \gg 1$.) For further discussions, see Section 4.

We would also like to mention the works [2022; 2021] in which an interesting framework in terms of tempered distributions and relative trace formula of Godement–Jacquet type was developed to address the phenomenon (1-2).

3B. ***Comparisons with the Conrey–Iwaniec–Li method.*** The celebrated works of Conrey and Iwaniec [2000] and Li [2009; 2011] are known for their successful analysis based on the Kuznetsov trace formulae and summation formulae of Poisson/Voronoi type. Their accomplishments include a delicate treatment of the arithmetic of exponential sums as well as the stationary phase analysis.

The Kuznetsov trace formula (or more generally the relative trace formula) has been a cornerstone in the analytic theory of L-functions over the past few decades. In the context of Theorem 1.1, which involves summing over a basis of even Maass forms for $\mathrm{SL}_2(\mathbb{Z})$ (or equivalently, Maass forms for $\mathrm{PGL}_2(\mathbb{Z}) \backslash \mathrm{PGL}_2(\mathbb{R})$), it is an equality of the shape

$$\sum_j H(\mu_j)\frac{\lambda_j(n)\overline{\lambda_j(m)}}{L(1, \mathrm{Ad}^2 \phi_j)} + (\mathrm{cts}) = \delta_{m=n}\int_{\mathbb{R}} H(\mu)d_{\mathrm{spec}}\mu + \sum_{\pm}\sum_c \frac{S(\pm m, n; c)}{c}\mathcal{J}^{\pm}\left(\frac{4\pi\sqrt{mn}}{c}\right) \tag{3-3}$$

between the spectral bilinear form of Hecke eigenvalues and the geometric expansion, which consists of Kloosterman sums $S(m, n; c)$ and oscillatory integrals $\mathcal{J}^+$ and $\mathcal{J}^-$ involving the J-Bessel and K-Bessel function in their kernels respectively. These two pieces have to be treated separately.

As noticed in [Conrey and Iwaniec 2000; Li 2009; 2011; Blomer 2012b] and a number of subsequent works, the J-Bessel piece is particularly interesting due to its striking technical features. These features are crucial for achieving significant cancellations in geometric sums and integrals, a property that appears to be distinctive to higher-rank settings. (In view of this, readers may wish to compare with the analysis in [Liu and Ye 2002] in the GL(2) settings.) More concretely, Li [2011] was able to apply the GL(3) Voronoi formula *twice*, which were surprisingly noninvoluntary, because of a subtle cancellation taking place between the *arithmetic phase* coming from Voronoi and the *analytic phase* coming from the J-Bessel transform.

The treatment of the J-Bessel piece in the Kuznetsov–Voronoi approach is crucial for analyzing more general moments of L-functions, including those involving nonselfdual L-functions or noncentral L-values, as demonstrated in Theorem 1.1.

In our period integral approach, the Kuznetsov formula, the Voronoi formula, and the approximate functional equation, which belong to the standard toolbox in analytic number theory, are completely avoided. This is motivated by several conceptual reasons, which we will now explain:

- Firstly, since the $\mathrm{GL}(3) \times \mathrm{GL}(2)$ L-functions on the spectral side are interpreted as period integrals, we never need to open up those L-functions into Dirichlet series. As a result, averaging over the Hecke eigenvalues of our basis of $\mathrm{GL}(2)$ Maass forms using the Kuznetsov formula is unnecessary.
- Secondly, the dual arithmetic object in our moment identity (1-3) contains the standard L-function of $\mathrm{GL}(3)$. The standard L-function is constructed solely from the $\mathrm{GL}(3)$ Hecke eigenvalues, whereas the $\mathrm{GL}(3)$ Voronoi formula involves *general* Fourier coefficients of $\mathrm{GL}(3)$ due to arithmetic twisting. It is thus reasonable to expect a proof of (1-3) that does not rely on the $\mathrm{GL}(3)$ Voronoi formula of [Miller and Schmid 2006] nor the full Fourier expansions of [Jacquet et al. 1979a; 1979b]. The set-up (3-1) already suggests that our method meets such an expectation, but see Proposition 6.1 for full details.
- Thirdly, we do not encounter any intermediate exponential sums (e.g., Kloosterman/Ramanujan sums), slow-decaying/very oscillatory special functions, nor shifted convolution sums which are necessary in [Ivić 2001; 2002; Frolenkov 2020] for (1-2). Also, we handle the archimedean component of (1-3) in a unified manner, rather than handling the J- and K-Bessel pieces separately as done in [Conrey and Iwaniec 2000; Li 2009; 2011]. We directly work with the $\mathrm{GL}(3)$ Whittaker function associated with the automorphic form Φ.
- Fourthly, we take advantage of the equivariance of the Whittaker functions under unipotent translations which helps to simplify many formulae.

Our period integral approach offers several technical advantages and is fundamentally distinct from the Kuznetsov–Voronoi approach. Indeed, our approach is *local* and the key result Proposition 6.1 can be easily phrased in terms of adeles (see (4-7)), whereas the Kuznetsov–Voronoi approach is *global* and *nonadelic*. In this article, we focus on the level 1 case and the spectral aspect as a proof of concept and thus we use the classical language of real groups. In our upcoming work, we wish to extend our method in various nonarchimedean aspects.

3C. *Prospects for higher-rank.* Once we reach $\mathrm{GL}(3)$, the geometric expansion for the Kuznetsov formula becomes significantly more intricate and presents a number of obstacles in generalizing the Kuznetsov-based approaches to moments of L-functions of higher-rank:

Remark 3.1 (oscillatory integrals). In $\mathrm{GL}(2)$, a couple of coincidences allow us to identify the oscillatory integrals with some well-studied special functions; see [Motohashi 1997; Iwaniec 2002]. However, such phenomena do not occur in $\mathrm{GL}(3)$, where unexpected analytic difficulties arise; see [Buttcane 2013; 2016]. The complicated formulae for the oscillatory integrals make the Kuznetsov trace formula for $\mathrm{GL}(3)$ challenging to apply; see [Blomer and Buttcane 2020].

Remark 3.2 (Kloosterman sums). The GL(3) Kloosterman sums, e.g.,

$$S(m_1, m_2, n_1, n_2; D_1, D_2) := \sideset{}{^\dagger}\sum_{\substack{B_1\ (D_1),\ B_2\ (D_2)\\ C_1\ (D_1),\ C_2\ (D_2)}} e\left(\frac{m_1 B_1 + n_1(Y_1 D_2 - Z_1 B_2)}{D_1}\right) e\left(\frac{m_2 B_2 + n_2(Y_2 D_1 - Z_2 B_1)}{D_2}\right), \tag{3-4}$$

are clearly much harder to work with than the usual one, where the definitions of Y_i, Z_i's along with a couple of congruence and coprimality conditions are suppressed. There are two other Kloosterman sums for GL(3); see [Buttcane 2013] for details.

As discussed in Section 3B, further transformations of the exponential sums from the Kuznetsov formulae encode important arithmetic information about the moments of L-functions. In [Blomer and Buttcane 2020] it was demonstrated that this approach for (3-4) after applying a four-fold Poisson summation. However, beyond this specific instance, the general applicability of such transformations to (3-4) remains unclear. On the other hand, applications of Voronoi formulae for GL(3) (see [Conrey and Iwaniec 2000; Li 2009; 2011; Blomer 2012b; Blomer and Khan 2019a; 2019b]) and for GL(4) (see [Blomer et al. 2019; Chandee and Li 2020]) are currently limited to the usual Kloosterman sums of GL(2), with complications arising quickly beyond this familiar context.

Conceptually speaking, the challenges associated with Remarks 3.1–3.2 stem from the *Bruhat decomposition*, which is fundamental to the framework of relative trace formulae in general. However, ideas from period reciprocity offers a way to bypass the Bruhat decomposition and the related geometric sums and integrals, which is a welcoming feature.

Regarding Remark 3.1, the advantages of our method are visible even in the context of Theorem 1.1. Even though we work with the group GL(3) on the dual side, the oscillatory factor in our approach (see (6-8)) is actually simpler than the ones encountered in the "Kuznetsov–Voronoi" approaches (see [Li 2011]). It is more structured in two key ways: (1) It arises naturally from the definition of the archimedean Whittaker function. (2) It serves as an important constituent of the exact Motohashi structure, the exact structures of the main terms predicted by [Conrey et al. 2005], as well as for the analytic continuation past $\operatorname{Re} s = \frac{1}{2}$. Furthermore, our approach is devoid of integrals over noncompact subsets of the unipotent subgroups (or the complements) which are known to result in intricate dual calculations and exponential phases in case of GL(3) Voronoi formula (see Section 4 of [Ichino and Templier 2013]) and Kuznetsov formulae (see Chapter 11 of [Goldfeld 2015]).

It is worth pointing out the crucial archimedean ingredient in our proof generalizes to GL(n) through *Stade's formula* (see [Stade 2001]), which allows us to rewrite the archimedean part completely in terms of integrals Γ-functions. This representation is sufficient for our purposes and possesses remarkable recursive structures beneficial for further analytic manipulations, as detailed in Section 10. Another notable recent application of Stade's formula can be found in [Goldfeld et al. 2021; 2022]. We anticipate that our method will provide insights into the structures of archimedean transforms, pave the way for generalizing to moments of higher-rank L-functions and overcome the technical challenges posed by the "Kuznetsov–Voronoi" method. We shall return to this subject in our upcoming works, together with treatment of the nonarchimedean places.

4. Informal sketch and discussion

To assist the readers, we first outline the main ideas of this article, before diving into any of the analytic subtleties of our actual argument. In fact, this represents the most intrinsic picture of our method and facilitates comparisons with the strategy of [Michel and Venkatesh 2006]. The style of this section will be largely informal — we shall suppress the constant multiples (say those 2's and π's), assume convergence, and set aside the treatment of main terms.

According to [Michel and Venkatesh 2006], the classical Motohashi formula can be understood as an intrinsic property of the GL(2) Eisenstein series (denoted by E^* below) via the ("regularized") geodesic period

$$\int_0^\infty |E^*(iy)|^2 \, d^\times y,$$

which can be evaluated in two ways according to $|E^*|^2$ and $E^* \cdot \overline{E^*}$ respectively:

(1) (GL(2) spectral expansion)

$$\sum_{\phi:\mathrm{GL}(2)} \langle |E^*|^2, \phi\rangle \int_0^\infty \phi(iy)\, d^\times y = \sum_{\phi:\mathrm{GL}(2)} \Lambda\bigl(\tfrac{1}{2},\phi\bigr)^2 \cdot \Lambda\bigl(\tfrac{1}{2},\phi\bigr) + (\cdots). \tag{4-1}$$

(2) (GL(1) × GL(1) expansion, or the Mellin–Plancherel formula)

$$\int_{(1/2)} |\widetilde{E^*}(s)|^2 \frac{ds}{2\pi i} = \int_{\mathbb{R}} \bigl|\Lambda\bigl(\tfrac{1}{2}+it\bigr)^2\bigr|^2 \frac{dt}{2\pi}. \tag{4-2}$$

This seemingly simple sketch turns out to require rather sophisticated regularizations but was skillfully executed very recently in [Nelson 2019].

We now turn to our sketch of the (generalized) Motohashi phenomenon as described in Theorem 1.1. Let Φ be a Maass cusp form of $\mathrm{SL}_3(\mathbb{Z})$. As mentioned in the introduction, our starting point is the trivial identity

$$\int_0^1 \biggl[\int_0^\infty \Phi\begin{pmatrix} y_0 \begin{pmatrix} 1 & u \\ & 1\end{pmatrix} & \\ & 1\end{pmatrix} d^\times y_0\biggr] e(-u)\, du = \int_0^\infty \biggl[\int_0^1 \Phi\begin{pmatrix} \begin{pmatrix} 1 & u \\ & 1\end{pmatrix} y_0 & \\ & 1\end{pmatrix} e(-u)\, du\biggr] d^\times y_0. \tag{4-3}$$

For symmetry, observe that the right side of (4-3) can be written as

$$\int_0^\infty \left[\int_0^1 \widetilde{\Phi}\left[\begin{pmatrix} 1 & & \\ & 1 & u \\ & & 1\end{pmatrix}\begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1\end{pmatrix}\right] e(-u)\, du\right] d^\times y_0 \tag{4-4}$$

with $\widetilde{\Phi}(g) := \Phi({}^t g^{-1})$ being the dual form of Φ.

Remark 4.1. Indeed, the center-invariance of Φ implies that

$$\text{(4-3)} = \int_0^\infty \int_0^1 \Phi\left[\begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix}\begin{pmatrix} 1 & & \\ & 1 & \\ & & y_0 \end{pmatrix}\right] e(-u)\, du\, d^\times y_0.$$

Let $w_\ell := \begin{pmatrix} & & -1 \\ & 1 & \\ 1 & & \end{pmatrix}$. The observation

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & y_0 \end{pmatrix} = w_\ell^{-1}\begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix} w_\ell \quad \text{and} \quad \begin{pmatrix} 1 & & \\ & 1 & \\ & -u & 1 \end{pmatrix} = w_\ell \begin{pmatrix} 1 & u & \\ & 1 & \\ & & 1 \end{pmatrix} w_\ell^{-1}$$

together with the left and right invariance of Φ by w_ℓ further rewrite (4-3) as

$$\int_0^\infty \int_0^1 \Phi\left[\begin{pmatrix} 1 & & \\ & 1 & \\ & -u & 1 \end{pmatrix}\begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right] e(-u)\, du\, d^\times y_0$$
$$= \int_0^\infty \int_0^1 \widetilde{\Phi}\left[\begin{pmatrix} 1 & & \\ & 1 & u \\ & & 1 \end{pmatrix}\begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right] e(-u)\, du\, d^\times y_0.$$

As an overview of our strategy:

(1) *Similar to* Michel–Venkatesh's strategy, the integral over $(0, \infty)$ (or the center $Z^+_{\mathrm{GL}_2}(\mathbb{R})$) yields Rankin–Selberg L-functions on the spectral side and a t-integral on the dual side.

(2) *Different from* Michel–Venkatesh's strategy, our approach introduces an extra integral over $[0, 1]$ (or the quotient $U_2(\mathbb{Z}) \backslash U_2(\mathbb{R})$ of the unipotent subgroup U_2 of $\mathrm{GL}(2)$). This integral results in Whittaker functions as weight functions on the spectral side, and leads to a product of two distinct L-functions on the dual side.

(3) The Mellin–Plancherel of (4-2) is *replaced by* two Fourier expansions over $\mathbb{Z} \backslash \mathbb{R}$ below.

In fact, the *unipotent* nature of our period method is crucial in realizing the spectral duality for the fourth moment of Dirichlet L-functions (see [Kwan 2024]), as well as in ensuring the abundance of admissible test functions on the spectral side, but these features will not be displayed in this section.

4A. ***The* GL(2) *(spectral) side.*** This side is relatively straight-forward and gives the desired $\mathrm{GL}(3) \times \mathrm{GL}(2)$ moment. Regard Φ as a function of $L^2(\Gamma_2 \backslash \mathfrak{h}^2)$ via

$$(\mathrm{Proj}_2^3\, \Phi)(g) := \int_0^\infty \Phi\begin{pmatrix} y_0 g & \\ & 1 \end{pmatrix} d^\times y_0 \quad (g \in \mathfrak{h}^2),$$

which in turn can be expanded spectrally as

$$(\mathrm{Proj}_2^3\, \Phi)(g) = \sum_j \frac{\langle \mathrm{Proj}_2^3\, \Phi, \phi_j \rangle}{\|\phi_j\|^2} \phi_j(g) + \frac{\langle \mathrm{Proj}_2^3\, \Phi, 1 \rangle}{\|1\|^2} \cdot 1 + (\mathrm{cont}).$$

The spectral coefficients $\langle \mathrm{Proj}_2^3\,\Phi, \phi_j \rangle$ are precisely the GL(3) × GL(2) Rankin–Selberg L-functions. Hence,

$$\text{LHS of (4-3)} = \int_0^1 (\mathrm{Proj}_2^3\,\Phi) \begin{pmatrix} 1 & u \\ & 1 \end{pmatrix} e(-u)\,du = \sum_j W_{\mu_j}(1) \cdot \frac{\Lambda\left(\frac{1}{2}, \phi_j \otimes \Phi\right)}{\|\phi_j\|^2} + (\text{cont}), \tag{4-5}$$

where $\mu \mapsto W_\mu(1)$ is a weight function.

4B. ***The* GL(1) *(dual) side.*** In view of Point (3) above, we evaluate the innermost integral of (4-4) in terms of the *Fourier–Whittaker periods* for $\widetilde{\Phi}$, denoted by $(\widehat{\widetilde{\Phi}})_{(\cdot,\cdot)}$ (see Definition 5.12). From Proposition 6.1, (4-4) is given by

$$\begin{aligned}\int_0^\infty \int_0^1 \int_0^1 \widetilde{\Phi}\left[\begin{pmatrix} 1 & & u_{1,3} \\ & 1 & u_{2,3} \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right] e(-u_{2,3})\,du_{1,3}\,du_{2,3}d^\times y_0 \\ + \sum_{a_0 \in \mathbb{Z}-\{0\}} \sum_{a_1 \in \mathbb{Z}-\{0\}} \int_0^\infty (\widehat{\widetilde{\Phi}})_{(1,a_1)} \left[\begin{pmatrix} 1 & & \\ a_0 & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right] d^\times y_0.\end{aligned} \tag{4-6}$$

The first line of (4-6) corresponds to the diagonal term and is precisely the integral representation of the standard L-function of $\widetilde{\Phi}$. It is equal to $L(1, \widetilde{\Phi}) \cdot Z_\infty(1, \widetilde{\Phi})$, where $Z_\infty(\,\cdot\,, \widetilde{\Phi})$ is the GL(3) local zeta integral at ∞. The second line of (4-6) is the off-diagonal contribution, denoted by OD_Φ below, and is expressed in terms of the Fourier coefficients of $\widetilde{\Phi}$:

$$OD_\Phi = \sum_{a_0 \in \mathbb{Z}-\{0\}} \sum_{a_1 \in \mathbb{Z}-\{0\}} \frac{\mathcal{B}_{\widetilde{\Phi}}(1, a_1)}{|a_1|} \int_0^\infty (\widehat{\widetilde{\Phi}})_{(1,1)} \left[\begin{pmatrix} a_1/a_0 & & \\ 1 & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\right] d^\times y_0. \tag{4-7}$$

It can be further explicated as

$$OD_\Phi = \sum_{a_0 \in \mathbb{Z}-\{0\}} \sum_{a_1 \in \mathbb{Z}-\{0\}} \frac{\mathcal{B}_{\widetilde{\Phi}}(1, a_1)}{|a_1|} \cdot \int_0^\infty W_{\alpha(\Phi)}\left(\left|\frac{a_1}{a_0}\right| \frac{y_0}{1+y_0^2}, 1\right) \cdot e\left(\frac{a_1}{a_0} \frac{y_0^2}{1+y_0^2}\right) d^\times y_0 \tag{4-8}$$

using the GL(3) Whittaker function $W_{\alpha(\Phi)}$, where the oscillatory factor $e(\cdots)$ originates from the unipotent translation of Whittaker function.

Roughly speaking, (4-8) suggests some forms of (multiplicative) convolutions between the GL(3) and GL(1) data at both the archimedean and the nonarchimedean places:

(1) (Archimedean) We apply the Mellin inversion formula for $W_{\alpha(\Phi)}$, a standard result in GL(3) theory, together with the local functional equation for GL(1) in the form

$$e(x) + e(-x) = \int_{-i\infty}^{i\infty} \frac{\Gamma_{\mathbb{R}}(u)}{\Gamma_{\mathbb{R}}(1-u)} |x|^{-u} \frac{du}{2\pi i} \quad (x \neq 0). \tag{4-9}$$

(2) (Nonarchimedean) Observe the following identity of the double Dirichlet series:

$$\sum_{a_0\neq 0}\sum_{a_1\neq 0}\frac{\mathcal{B}_\Phi(a_1,1)}{|a_1|}\left|\frac{a_1}{a_0}\right|^{1-s_0-u}=L(s_0+u,\widetilde{\Phi})\zeta(1-s_0-u). \tag{4-10}$$

We thus arrive at

$$OD_\Phi=\int_{(1/2)}\zeta(1-s_0)L(s_0,\widetilde{\Phi})\cdot(\cdots)\frac{ds_0}{2\pi i}, \tag{4-11}$$

where "$(\cdots)$" stands for a certain integral transform that can be described purely in terms of Γ-functions.

Remark 4.2. (1) In (3-2), the test function h of the Poincaré series $P(*;h)$ will be transformed into the Kontorovich–Lebedev transform $h^{\#}$ on the GL(2) side (see Proposition 5.25) and into the Mellin transform $\widetilde{h}$ on the GL(1) side (see (7-6)). This is consistent with the sketch above.

(2) Readers may wish to compare the integral transforms obtained in the sketch with the one described in Section 1.3 of [Balkanova et al. 2021].

Remark 4.3. The choices of unipotent subgroups have been important in the constructions of various L-series for the group GL(3):

- $\left\{\begin{pmatrix}1 & & * \\ & 1 & * \\ & & 1\end{pmatrix}\right\}$ or $\left\{\begin{pmatrix}1 & * & * \\ & 1 & \\ & & 1\end{pmatrix}\right\}$ for the standard L-function.
- $\left\{\begin{pmatrix}1 & & * \\ & 1 & \\ & & 1\end{pmatrix}\right\}$ for Bump's double Dirichlet series [Bump 1984].
- $\left\{\begin{pmatrix}1 & & \\ & 1 & * \\ & & 1\end{pmatrix}\right\}$ or $\left\{\begin{pmatrix}1 & * & \\ & 1 & \\ & & 1\end{pmatrix}\right\}$ for the Motohashi phenomenon of this article.

5. Preliminary

The analytic theory of automorphic forms for the group GL(3) has undergone considerable development in the past decade. Readers should beware that the recent articles in the field (e.g., [Buttcane 2013; 2016; 2020; Goldfeld et al. 2021]) have adopted a different set of conventions and normalizations from those in the standard text [Goldfeld 2015]. (Nevertheless, [Goldfeld 2015] remains a useful reference as it thoroughly documents many standard results and their proofs.)

In this article, we follow the more recent conventions (closest to [Buttcane 2020]), which is better aligned with the theory of automorphic representation. We will summarize the essential notions and results below, with extra attention on the archimedean calculations involving Whittaker functions, as they play a key role in our analysis.

5A. ***Notations and conventions.*** Throughout this article, we use the following notations: $\Gamma_{\mathbb{R}}(s) := \pi^{-s/2}\Gamma(s/2)$ $(s \in \mathbb{C})$; $e(x) := e^{2\pi i x}$ $(x \in \mathbb{R})$; $\Gamma_n := \mathrm{SL}_n(\mathbb{Z})$ $(n \geq 2)$. Without otherwise specified, our test function H lies in the class $\mathcal{C}_\eta$ and $H = h^{\#}$. We will often use the same symbol to denote a function (in s) and its analytic continuation.

We will frequently encounter contour integrals of the shape

$$\int_{-i\infty}^{i\infty} \cdots \int_{-i\infty}^{i\infty} (\cdots) \frac{ds_1}{2\pi i} \cdots \frac{ds_k}{2\pi i}$$

where the contours involved should follow Barnes' convention: they pass to the right of all of the poles of the gamma functions in the form $\Gamma(s_i + a)$ and to the left of all of the poles of the gamma functions in the form $\Gamma(a - s_i)$.

We also adopt the following set of conventions:

(1) All Maass cusp forms will be simultaneous eigenfunctions of the Hecke operators and will be either even or odd. Also, their first Fourier coefficients are equal to 1. In this case, the forms are said to be *Hecke-normalized*. Note that there are no odd form for $\mathrm{SL}_3(\mathbb{Z})$; see Proposition 9.2.5 of [Goldfeld 2015].

(2) Our fixed Maass cusp form Φ of $\mathrm{SL}_3(\mathbb{Z})$ is assumed to be *tempered at* ∞, i.e., its Langlands parameters are purely imaginary.

(3) Denote by θ the best progress towards the Ramanujan conjecture for the Maass cusp forms of $\mathrm{SL}_3(\mathbb{Z})$. We have $\theta \leq \frac{1}{2} - \frac{1}{10}$; see Theorem 12.5.1 of [Goldfeld 2015].

5B. ***(Spherical) Whittaker functions and transforms.*** In the rest of this article, all Whittaker functions will refer to the spherical ones. The Whittaker function of $\mathrm{GL}_2(\mathbb{R})$ is more familiar and is given by

$$W_\mu(y) := 2\sqrt{y}K_\mu(2\pi y) \tag{5-1}$$

for $\mu \in \mathbb{C}$ and $y > 0$. Under this normalization, the following holds:

Proposition 5.1. *For* $\mathrm{Re}\big(w + \frac{1}{2} \pm \mu\big) > 0$, *we have*

$$\int_0^\infty W_\mu(y) y^w \, d^\times y = \frac{\pi^{-w-1/2}}{2} \Gamma\left(\frac{w + \frac{1}{2} + \mu}{2}\right)\Gamma\left(\frac{w + \frac{1}{2} - \mu}{2}\right). \tag{5-2}$$

Proof. Standard, see (2.5.2) of [Motohashi 1997] for instance. □

For the group $\mathrm{GL}_3(\mathbb{R})$, we first introduce the function

$$I_\alpha(y_0, y_1) = I_\alpha\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix} := y_0^{1-\alpha_3} y_1^{1+\alpha_1}$$

for $y_0, y_1 > 0$ and $\alpha \in \mathfrak{a}_{\mathbb{C}}^{(3)} := \{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{C}^3 : \alpha_1 + \alpha_2 + \alpha_3 = 0\}$. Then the Whittaker function for $\mathrm{GL}_3(\mathbb{R})$, denoted by

$$W_\alpha(y_0, y_1) = W_\alpha \begin{pmatrix} y_0 y_1 & & \\ & y_0 & \\ & & 1 \end{pmatrix},$$

is defined in terms of *Jacquet's integral*

$$\prod_{1 \le j < k \le 3} \Gamma_{\mathbb{R}}(1+\alpha_j-\alpha_k) \times \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} I_\alpha \left[\begin{pmatrix} & & 1 \\ & -1 & \\ 1 & & \end{pmatrix} \begin{pmatrix} 1 & u_{1,2} & u_{1,3} \\ & 1 & u_{2,3} \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 y_1 & & \\ & y_0 & \\ & & 1 \end{pmatrix} \right] e(-u_{1,2}-u_{2,3})\, du_{1,2}\, du_{1,3}\, du_{2,3} \tag{5-3}$$

for $y_0, y_1 > 0$ and $\alpha \in \mathfrak{a}_{\mathbb{C}}^{(3)}$; see Chapter 5.5 of [Goldfeld 2015] for details.

Remark 5.2. Notice the differences in the normalizations of I_α here compared to that in equation (5.1.1) of [Goldfeld 2015]. Also, the Whittaker functions discussed here are the *complete* Whittaker functions as defined in [loc. cit.].

Moreover, the Whittaker function of $\mathrm{GL}_3(\mathbb{R})$ admits the following useful Mellin–Barnes representation commonly known as the *Vinogradov–Takhtadzhyan formula*:

Proposition 5.3. *Assume $\alpha \in \mathfrak{a}_{\mathbb{C}}^{(3)}$ is tempered, i.e.,* $\operatorname{Re} \alpha_i = 0$ $(i = 1, 2, 3)$. *Then for any $\sigma_0, \sigma_1 > 0$,*

$$W_{-\alpha}(y_0, y_1) = \frac{1}{4} \int_{(\sigma_0)} \int_{(\sigma_1)} G_\alpha(s_0, s_1) y_0^{1-s_0} y_1^{1-s_1} \frac{ds_0}{2\pi i} \frac{ds_1}{2\pi i}, \qquad y_0, y_1 > 0, \tag{5-4}$$

where

$$G_\alpha(s_0, s_1) := \frac{\prod_{i=1}^{3} \Gamma_{\mathbb{R}}(s_0 + \alpha_i) \Gamma_{\mathbb{R}}(s_1 - \alpha_i)}{\Gamma_{\mathbb{R}}(s_0 + s_1)}. \tag{5-5}$$

Proof. This can be verified $\left(\text{up to the constant } \frac{1}{4}\right)$ by a brute force yet elementary calculation, i.e., checking the right side of (5-4) satisfies the differential equations of GL(3); see pages 38–39 of [Bump 1984]. For a cleaner proof starting from (5-3); see Chapter X of [loc. cit.]. □

Remark 5.4. Notice the sign convention of the α_i in formula (5-4) — it is consistent with [Buttcane 2020] but is opposite to that of (6.1.4)–(6.1.5) in [Goldfeld 2015].

Corollary 5.5. *For any $-\infty < A_0, A_1 < 1$, we have*

$$|W_{-\alpha}(y_0, y_1)| \ll y_0^{A_0} y_1^{A_1}, \qquad y_0, y_1 > 0, \tag{5-6}$$

where the implicit constant depends only on α, A_0, A_1.

Proof. Follows directly from Proposition 5.3 by contour shifting. □

We will need the explicit evaluation of the $\mathrm{GL}_3(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$ Rankin–Selberg integral. It is a consequence of the *second Barnes lemma* stated as follows.

Lemma 5.6. *For $a, b, c, d, e, f \in \mathbb{C}$ with $f = a+b+c+d+e$, we have*

$$\int_{-i\infty}^{i\infty} \frac{\Gamma(w+a)\Gamma(w+b)\Gamma(w+c)\Gamma(d-w)\Gamma(e-w)}{\Gamma(w+f)} \frac{dw}{2\pi i} = \frac{\Gamma(d+a)\Gamma(d+b)\Gamma(d+c)\Gamma(e+a)\Gamma(e+b)\Gamma(e+c)}{\Gamma(f-a)\Gamma(f-b)\Gamma(f-c)}. \tag{5-7}$$

The contours of integration must adhere to Barnes' convention; *see Section 5A for details.*

Proof. See [Bailey 1935]. □

Proposition 5.7. *Let W_μ and $W_{-\alpha}$ be the Whittaker functions of* $\mathrm{GL}_2(\mathbb{R})$ *and* $\mathrm{GL}_3(\mathbb{R})$ *respectively. For* $\operatorname{Re} s \gg 0$, *we have*

$$\begin{aligned}\mathcal{Z}_\infty(s; W_\mu, W_{-\alpha}) :&= \int_0^\infty \int_0^\infty W_\mu(y_1) W_{-\alpha}(y_0, y_1)(y_0^2 y_1)^{s-1/2} \frac{dy_0 dy_1}{y_0 y_1^2} \\ &= \tfrac{1}{4} \prod_{\pm} \prod_{k=1}^{3} \Gamma_{\mathbb{R}}(s \pm \mu - \alpha_k).\end{aligned} \tag{5-8}$$

Proof. See [Bump 1988]. □

The following pair of integral transforms plays an important role in the archimedean aspect of this article.

Definition 5.8. Let $h : (0, \infty) \to \mathbb{C}$ and $H : i\mathbb{R} \to \mathbb{C}$ be measurable functions with $H(\mu) = H(-\mu)$. Let $W_\mu(y) := 2\sqrt{y} K_\mu(2\pi y)$. Then the Kontorovich–Lebedev transform of h is defined by

$$h^{\#}(\mu) := \int_0^\infty h(y) W_\mu(y) \frac{dy}{y^2}, \tag{5-9}$$

whereas its inverse transform is defined by

$$H^{\flat}(y) = \frac{1}{4\pi i} \int_{(0)} H(\mu) W_\mu(y) \frac{d\mu}{|\Gamma(\mu)|^2}, \tag{5-10}$$

provided the integrals converge absolutely. Note: the normalization constant $1/4\pi i$ in (5-10) is consistent with that in [Motohashi 1997; Iwaniec 2002].

Definition 5.9. Let $\mathcal{C}_\eta$ be the class of holomorphic functions H on the vertical strip $|\operatorname{Re} \mu| < 2\eta$ such that

(1) $H(\mu) = H(-\mu)$,

(2) H has rapid decay in the sense that

$$H(\mu) \ll e^{-2\pi|\mu|} \quad (|\operatorname{Re} \mu| < 2\eta). \tag{5-11}$$

In this article, we take $\eta > 40$ without otherwise specifying.

By contour-shifting and Stirling's formula, we have:

Proposition 5.10. *For any $H \in \mathcal{C}_\eta$, the integral (5-10) defining $H^\flat$ converges absolutely. Moreover, we have*

$$H^\flat(y) \ll \min\{y, y^{-1}\}^\eta \quad (y > 0). \tag{5-12}$$

Proof. See Lemma 2.10 of [Motohashi 1997]. □

Proposition 5.11. *Under the same assumptions of Proposition 5.10, we have*

$$(h^\#)^\flat(g) = h(g) \quad \textit{and} \quad (H^\flat)^\#(\mu) = H(\mu). \tag{5-13}$$

Proof. See Lemma 2.10 of [Motohashi 1997]. It is a consequence of the Rankin–Selberg calculation for $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$. □

5C. *Automorphic forms of* GL(2) *and* GL(3).

Let

$$\mathfrak{h}^2 := \left\{ \begin{pmatrix} 1 & u \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & 1 \end{pmatrix} : u \in \mathbb{R},\, y > 0 \right\}$$

with its invariant measure given by $y^{-2}\, du\, dy$. Let $\Delta := -y^2(\partial_x^2 + \partial_y^2)$. An automorphic form $\phi : \mathfrak{h}^2 \to \mathbb{C}$ of $\Gamma_2 = \mathrm{SL}_2(\mathbb{Z})$ satisfies $\Delta\phi = \left(\frac{1}{4} - \mu^2\right)\phi$ for some $\mu = \mu(\phi) \in \mathbb{C}$. It is often handy to identify μ with the pair $(\mu, -\mu) \in \mathfrak{a}_{\mathbb{C}}^{(2)}$.

For $a \in \mathbb{Z} - \{0\}$, the a-th Fourier coefficient of ϕ, denoted by $\mathcal{B}_\phi(a)$, is defined by

$$(\widehat{\phi})_a(y) := \int_0^1 \phi\left[\begin{pmatrix} 1 & u \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & 1 \end{pmatrix} \right] e(-au)\, du = \frac{\mathcal{B}_\phi(a)}{\sqrt{|a|}} \cdot W_{\mu(\phi)}(|a|y). \tag{5-14}$$

In the case of the Eisenstein series of Γ_2, i.e.,

$$\phi = E(z; \mu) := \tfrac{1}{2} \sum_{\gamma \in U_2(\mathbb{Z}) \backslash \Gamma_2} I_\mu(\operatorname{Im} \gamma z) \quad (z \in \mathfrak{h}^2), \tag{5-15}$$

where $I_\mu(y) := y^{\mu+1/2}$, it is well-known that $\Delta E(*; \mu) = \left(\frac{1}{4} - \mu^2\right) E(*; \mu)$ and the Fourier coefficients $\mathcal{B}(a; \mu)$ of $E(*; \mu)$ is given by

$$\mathcal{B}(a; \mu) = \frac{|a|^\mu \sigma_{-2\mu}(|a|)}{\Lambda(1 + 2\mu)}, \tag{5-16}$$

where

$$\Lambda(s) := \pi^{-s/2} \Gamma(s/2) \zeta(s) \quad \text{and} \quad \sigma_{-2\mu}(|a|) := \sum_{d|a} d^{-2\mu}.$$

The series (5-15) converges absolutely for $\operatorname{Re} \mu > \frac{1}{2}$ and it admits a meromorphic continuation to $\mathbb{C}$.

Next, let

$$\mathfrak{h}^3 := \left\{ \begin{pmatrix} 1 & u_{1,2} & u_{1,3} \\ & 1 & u_{2,3} \\ & & 1 \end{pmatrix} \begin{pmatrix} y_0 y_1 & & \\ & y_0 & \\ & & 1 \end{pmatrix} : u_{i,j} \in \mathbb{R},\, y_k > 0 \right\}.$$

Let $\Phi : \mathfrak{h}^3 \to \mathbb{C}$ be a Maass cusp form of Γ_3 as defined in Definition 5.1.3 of [Goldfeld 2015]. In particular, there exists $\alpha = \alpha(\Phi) \in \mathfrak{a}_{\mathbb{C}}^{(3)}$ such that for any $D \in Z(U\mathfrak{gl}_3(\mathbb{C}))$ (the center of the universal enveloping algebra of the Lie algebra $\mathfrak{gl}_3(\mathbb{C})$), we have

$$D\Phi = \lambda_D \Phi \quad \text{and} \quad DI_\alpha = \lambda_D I_\alpha$$

for some $\lambda_D \in \mathbb{C}$. The triple $\alpha(\Phi)$ is said to be the *Langlands parameters* of Φ.

Definition 5.12. Let $m = (m_1, m_2) \in (\mathbb{Z} - \{0\})^2$ and $\Phi : \mathfrak{h}^3 \to \mathbb{C}$ be a Maass cusp form of $\mathrm{SL}_3(\mathbb{Z})$. For any $y_0, y_1 > 0$, the integral defined by

$$(\widehat{\Phi})_{(m_1,m_2)}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}$$
$$:= \int_0^1\int_0^1\int_0^1 \Phi\left[\begin{pmatrix} 1 & u_{1,2} & u_{1,3} \\ & 1 & u_{2,3} \\ & & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-m_1u_{2,3} - m_2u_{1,2})\, du_{1,2}\, du_{1,3}\, du_{2,3}. \tag{5-17}$$

is said to be the (m_1, m_2)-th *Fourier–Whittaker period* of Φ. Moreover, the (m_1, m_2)-th *Fourier coefficient* of Φ is the complex number $\mathcal{B}_\Phi(m_1, m_2)$ for which

$$(\widehat{\Phi})_{(m_1,m_2)}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix} = \frac{\mathcal{B}_\Phi(m_1,m_2)}{|m_1m_2|}\, W_{\alpha(\Phi)}^{\mathrm{sgn}(m_2)}\begin{pmatrix} (|m_1|y_0)(|m_2|y_1) & & \\ & |m_1|y_0 & \\ & & 1\end{pmatrix} \tag{5-18}$$

holds for any $y_0, y_1 > 0$.

Remark 5.13. (1) The multiplicity-one theorem of Shalika (see Theorem 6.1.6 of [Goldfeld 2015]) guarantees the well-definedness of the Fourier coefficients for Φ.

(2) If Φ is Hecke-normalized (see Section 5A.(1)), then $\mathcal{B}_\Phi(1, n)$ can be shown to be a Hecke eigenvalue of Φ; see Section 6.4 of [Goldfeld 2015].

5D. *Automorphic L-functions.* The Maass cusp forms Φ and ϕ below are Hecke-normalized and their Langlands parameters are denoted by $\alpha \in \mathfrak{a}_{\mathbb{C}}^{(3)}$ and $\mu \in \mathfrak{a}_{\mathbb{C}}^{(2)}$ respectively. Let $\widetilde{\Phi}(g) := \Phi({}^t g^{-1})$ be the dual form of Φ. It is not hard to show that the Langlands parameters of $\widetilde{\Phi}$ are given by $-\alpha$.

Definition 5.14. Suppose Φ and ϕ are Maass cusp forms of Γ_3 and Γ_2 respectively. For $\mathrm{Re}\, s \gg 1$, the Rankin–Selberg L-function of Φ and ϕ is defined by

$$L(s, \phi \otimes \Phi) := \sum_{m_1=1}^{\infty}\sum_{m_2=1}^{\infty} \frac{\mathcal{B}_\phi(m_2)\mathcal{B}_\Phi(m_1,m_2)}{(m_1^2m_2)^s}. \tag{5-19}$$

Although we do not make use of the Dirichlet series for $L(s, \phi \otimes \Phi)$ in this article, it is frequently used in the literature, especially in the "Kuznetsov–Voronoi" method. We take this opportunity to indicate our normalization in terms of Dirichlet series to facilitate conversion and comparison, and to correct some minor inaccuracies in Section 12.2 of [Goldfeld 2015].

Proposition 5.15. *Suppose Φ and ϕ are Maass cusp forms of Γ_3 and Γ_2 respectively. In addition, assume that ϕ is even. Then for any* $\operatorname{Re} s \gg 1$, *we have*

$$\int_{\Gamma_2\backslash \mathrm{GL}_2(\mathbb{R})} \phi(g)\widetilde{\Phi}\begin{pmatrix} g & \\ & 1 \end{pmatrix} |\det g|^{s-1/2}\, dg = \tfrac{1}{2}\cdot \Lambda(s, \phi\otimes\widetilde{\Phi}), \tag{5-20}$$

where

$$\Lambda(s,\phi\otimes\widetilde{\Phi}) := L_\infty(s,\phi\otimes\widetilde{\Phi})\cdot L(s,\phi\otimes\widetilde{\Phi}) \tag{5-21}$$

and

$$L_\infty(s,\phi\otimes\widetilde{\Phi}) := \prod_{k=1}^{3} \Gamma_{\mathbb{R}}(s\pm\mu-\alpha_k). \tag{5-22}$$

Proof. The assumption on the parity of ϕ is missing in [Goldfeld 2015]. Also, the pairing should be taken over the quotient $\Gamma_2\setminus \mathrm{GL}_2(\mathbb{R})$ instead of $\Gamma_2\setminus\mathfrak{h}^2$ in [loc. cit.].

As a brief sketch, we replace $\widetilde{\Phi}\begin{pmatrix} g & \\ & 1\end{pmatrix}$ by its Fourier–Whittaker expansion (see Theorem 5.3.2 of [Goldfeld 2015]) on the left side of (5-20) and unfold. Then one may extract the Dirichlet series in (5-19) by using (5-14) and (5-17). The integral of Whittaker functions can be computed by Proposition 5.7. □

In the rest of this article, we will often make use of the shorthands $(\mathbb{P}_2^3\Phi)(g) := \Phi\begin{pmatrix} g & \\ & 1\end{pmatrix}$ and the pairing

$$(\phi, (\mathbb{P}_2^3\Phi)\cdot|\det *|^{\bar{s}-1/2})_{\Gamma_2\backslash \mathrm{GL}_2(\mathbb{R})}$$

for the integral on the left side of (5-20). By the rapid decay of Φ at ∞, this integral converges absolutely for any $s\in\mathbb{C}$ and uniformly on any compact subset of $\mathbb{C}$. Thus, the L-function $L(s,\phi\otimes\widetilde{\Phi})$ admits an entire continuation.

Remark 5.16. (1) When ϕ is even, the involution $g\mapsto {}^tg^{-1}$ gives the functional equation

$$\Lambda(s,\phi\otimes\widetilde{\Phi}) = \Lambda(1-s,\phi\otimes\Phi).$$

(2) When ϕ is odd, the right side of (5-20) is identical to 0 and hence *does not* provide an integral representation for $\Lambda(s,\phi\otimes\widetilde{\Phi})$. One must alter Proposition 5.15 accordingly in this case, say using the raising/lowering operators, or proceed adelically with an appropriate choice of test vector at ∞. However, we shall not go into these as our spectral average is taken over even Maass forms of Γ_2 only.

(3) As discussed in Section 3B, the roles of parities and root numbers are rather intricate in the study of moments of L-functions, especially regarding the archimedean integral transforms.

Definition 5.17. Let $\Phi:\mathfrak{h}^3\to\mathbb{C}$ be a Maass cusp form of Γ_3. For $\operatorname{Re} s\gg 1$, the standard L-function of Φ is defined by

$$L(s,\Phi) := \sum_{n=1}^{\infty} \frac{\mathcal{B}_\Phi(1,n)}{n^s}. \tag{5-23}$$

In the rest of this article, we will not make use of the integral representation of $L(s,\Phi)$, i.e., the first line of (4-6) with $\widetilde{\Phi}$ replaced by Φ. It suffices to note that $L(s,\Phi)$ admits an entire continuation and satisfies the following functional equation:

Proposition 5.18. *Let $\Phi:\mathfrak{h}^3\to\mathbb{C}$ be a Maass cusp form of Γ_3. For any $s\in\mathbb{C}$, we have*

$$\Lambda(s,\Phi)=\Lambda(1-s,\widetilde{\Phi}), \tag{5-24}$$

where

$$\Lambda(s,\Phi):=L_\infty(s,\Phi)\cdot L(s,\Phi) \tag{5-25}$$

and

$$L_\infty(s,\Phi):=\prod_{k=1}^{3}\Gamma_{\mathbb{R}}(s+\alpha_k). \tag{5-26}$$

Proof. See Chapter 6.5 of [Goldfeld 2015] or [Jacquet et al. 1979a; 1979b]. □

Furthermore, since ϕ and Φ are assumed to be Hecke-normalized, the standard L-functions $L(s,\phi)$ and $L(s,\Phi)$ admit Euler products of the form

$$L(s,\phi)=\prod_p\prod_{j=1}^{2}(1-\beta_{\phi,j}(p)p^{-s})^{-1},\quad L(s,\Phi)=\prod_p\prod_{k=1}^{3}(1-\alpha_{\Phi,k}(p)p^{-s})^{-1} \tag{5-27}$$

for $\operatorname{Re}s\gg 1$. Then one can show that

$$L(s,\phi\otimes\Phi)=\prod_p\prod_{j=1}^{2}\prod_{k=1}^{3}(1-\beta_{\phi,j}(p)\alpha_{\Phi,k}(p)p^{-s})^{-1} \tag{5-28}$$

by Cauchy's identity, see the argument of Proposition 7.4.12 of [Goldfeld 2015].

Proposition 5.19. *For* $\operatorname{Re}(s\pm\mu)\gg 1$, *we have*

$$\big(E(*;\mu),(\mathbb{P}_2^3\Phi)\cdot|\det *|^{\bar{s}-1/2}\big)_{\Gamma_2\backslash\mathrm{GL}_2(\mathbb{R})}=\tfrac{1}{2}\frac{\Lambda(s+\mu,\widetilde{\Phi})\Lambda(s-\mu,\widetilde{\Phi})}{\Lambda(1+2\mu)}. \tag{5-29}$$

Proof. Parallel to Proposition 5.15. Meanwhile, we make use of (5-16). □

Remark 5.20. By analytic continuation, (5-20) and (5-29) hold for $s\in\mathbb{C}$ and away from the poles of $E(*;\mu)$. In fact, the rapid decay of Φ at ∞ guarantees the pairings converge absolutely.

5E. *Calculation on the spectral side.* As noted before, our approach diverges from the "Kuznetsov–Voronoi" method from the outset. We express the moment of GL(3) × GL(2) L-functions via the period integral in Proposition 5.15 using a Poincaré series.

Definition 5.21. Let $a\ge 1$ be an integer and $h\in C^\infty(0,\infty)$. The Poincaré series of Γ_2 is defined as

$$P^a(z;h):=\sum_{\gamma\in U_2(\mathbb{Z})\backslash\Gamma_2}h(a\operatorname{Im}\gamma z)e(a\operatorname{Re}\gamma z)\quad (z\in\mathfrak{h}^2) \tag{5-30}$$

provided the series converges absolutely.

It is not hard to see that if the bounds

$$h(y) \ll y^{1+\epsilon} \quad (\text{as } y \to 0) \quad \text{and} \quad h(y) \ll y^{1/2-\epsilon} \quad (\text{as } y \to \infty) \tag{5-31}$$

are satisfied, then the Poincaré series $P^a(z; h)$ converges absolutely and represents an L^2-function. In this article, we take $h := H^\flat$ with $H \in \mathcal{C}_\eta$ and $\eta > 40$. By Proposition 5.10, the conditions in (5-31) are clearly met. We will often use the shorthand $P^a := P^a(*; h)$. Also, we denote the Petersson inner product on $\Gamma_2 \backslash \mathfrak{h}^2$ by $\langle \cdot, \cdot \rangle$, defined as

$$\langle \phi_1, \phi_2 \rangle := \int_{\Gamma_2 \backslash \mathfrak{h}^2} \phi_1(g) \cdot \overline{\phi_2(g)}\, dg$$

with dg being the invariant measure on $\mathfrak{h}^2$.

Lemma 5.22. *Let ϕ be a Maass cusp form of Γ_2, $\Delta\phi = \left(\frac{1}{4} - \mu^2\right)\phi$, and $\mathcal{B}_\phi(a)$ be the a-th Fourier coefficient of ϕ. Then*

$$\langle P^a, \phi \rangle = |a|^{1/2} \cdot \overline{\mathcal{B}_\phi(a)} \cdot h^\#(\bar{\mu}).$$

Proof. Replace P^a in $\langle P^a, \phi \rangle$ by its definition and unfold, we easily find that

$$\langle P^a, \phi \rangle = \int_0^\infty h(ay) \cdot \overline{(\widehat{\phi})_a(y)}\, \frac{dy}{y^2}.$$

The result follows at once upon plugging-in (5-14) and making the change of variable $y \to |a|^{-1} y$. □

Similarly, the following holds away from the poles of $E(*; \mu)$:

Lemma 5.23. *We have*

$$\langle P^a, E(*; \mu) \rangle = |a|^{1/2} \cdot \frac{|a|^{\bar{\mu}} \sigma_{-2\bar{\mu}}(|a|)}{\zeta^*(1 + 2\bar{\mu})} \cdot h^\#(\bar{\mu}). \tag{5-32}$$

Proposition 5.24 (spectral expansion). *Suppose $f \in L^2(\Gamma_2 \backslash \mathfrak{h}^2)$ and $\langle f, 1 \rangle = 0$. Then*

$$f(z) = \sum_{j=1}^\infty \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle} \cdot \phi_j(z) + \int_{(0)} \langle f, E(*; \mu) \rangle \cdot E(z; \mu) \frac{d\mu}{4\pi i} \quad (z \in \mathfrak{h}^2) \tag{5-33}$$

where $(\phi_j)_{j \ge 1}$ is any orthogonal basis of Maass cusp forms for Γ_2.

Proof. See Theorem 3.16.1 of [Goldfeld 2015]. □

Proposition 5.25. *Let Φ be a Maass cusp form of Γ_3 and P^a be a Poincaré series of Γ_2. Then*

$$2|a|^{-1/2} (P^a, (\mathbb{P}_2^3 \Phi) |\det *|^{\bar{s}-1/2})_{\Gamma_2 \backslash \mathrm{GL}_2(\mathbb{R})}$$
$$= \sum_{j=1}^{\infty}{}' h^\#(\overline{\mu_j}) \frac{\overline{\mathcal{B}_j(a)} \Lambda(s, \phi_j \otimes \widetilde{\Phi})}{\langle \phi_j, \phi_j \rangle} + \int_{(0)} h^\#(\mu) \frac{\sigma_{-2\mu}(|a|) |a|^{-\mu} \Lambda(s+\mu, \widetilde{\Phi}) \Lambda(1-s+\mu, \Phi)}{|\Lambda(1+2\mu)|^2} \frac{d\mu}{4\pi i} \tag{5-34}$$

for any $s \in \mathbb{C}$, where the sum is restricted to an orthogonal basis (ϕ_j) of even Hecke-normalized Maass cusp forms for Γ_2 with $\Delta\phi_j = \left(\frac{1}{4} - \mu_j^2\right)\phi_j$ and $\mathcal{B}_j(a) := \mathcal{B}_{\phi_j}(a)$.

Proof. Substitute the spectral expansion of P^a as in (5-33) into the pairing $(P^a, (\mathbb{P}_2^3\Phi)\cdot|\det *|^{\bar{s}-1/2})_{\Gamma_2\backslash \mathrm{GL}_2(\mathbb{R})}$. The inner products involved have been computed in Lemmas 5.22–5.23 and Definitions 5.15–5.19. □

Remark 5.26. Good control over spectral aspects and integral transforms, along with flexibility in choosing test functions on the spectral side, are crucial in applications. Also, this helps eliminate extraneous polar contributions (e.g., those not predicted by [Conrey et al. 2005]) for Eisenstein cases. These explain the preference of Kuznetsov-based methods over period-based methods (see the discussions in [Blomer 2012a; Nunes 2023; Zacharias 2019; 2021]).

While our method is period-based, it accommodates a broad class of test functions similar to the Kuznetsov approaches, thanks to the transforms in Definition 5.8. These transforms, generalized to $\mathrm{GL}(n)$ as in [Goldfeld and Kontorovich 2012], have significantly contributed to the development of higher-rank Kuznetsov formulae; see [Goldfeld and Kontorovich 2013; Goldfeld et al. 2021; 2022; Buttcane 2020].

Our method effectively combines the strengths of both Kuznetsov and period approaches, balancing precision in the archimedean aspect with structural insights into the nonarchimedean aspect.

Remark 5.27. Within our class $\mathcal{C}_\eta$ of test functions, a good choice of test function is given by

$$H(\mu) := \left(e^{((\mu+iM)/R)^2}+e^{((\mu-iM)/R)^2}\right)\cdot \frac{\Gamma(2\eta+\mu)\Gamma(2\eta-\mu)}{\prod_{i=1}^{3}\Gamma\left(\frac{1/2+\mu-\alpha_i}{2}\right)\Gamma\left(\frac{1/2-\mu-\alpha_i}{2}\right)}, \tag{5-35}$$

where $\eta > 40$, $M \gg 1$, and $R = M^\gamma$ $(0<\gamma\le 1)$. In (5-35),

- the factor $e^{((\mu+iM)/R)^2}+e^{((\mu-iM)/R)^2}$ serves as a smooth cut-off for $|\mu_j| \in [M-R, M+R]$ and gives the needed decay in Proposition 5.10;
- the factors $\prod_{i=1}^{3}\Gamma\left(\frac{1}{2}\left(\frac{1}{2}+\mu-\alpha_i\right)\right)\Gamma\left(\frac{1}{2}\left(\frac{1}{2}-\mu-\alpha_i\right)\right)$ cancel out the archimedean factors of $\Lambda\left(\frac{1}{2}, \phi_j\otimes\widetilde{\Phi}\right)$ on the spectral expansion (5-34) and in the diagonal contribution (6-9);
- the factors $\Gamma(2\eta+\mu)\Gamma(2\eta-\mu)$ balance off the exponential growth from $d\mu/|\Gamma(\mu)|^2$, $\|\phi_j\|^{-2}$ and $|\Lambda(1+2i\mu)|^{-2}$. Also, a large enough region of holomorphy of (5-35) is maintained so that $h(y) := H^\flat(y)$ has sufficient decay at 0 and ∞.

Remark 5.28. One might consider using an automorphic kernel instead of a Poincaré series for Theorem 1.1. While this offers more structural flexibility, the analysis of the spherical transforms becomes quite complicated; see [Zagier 1981; Buttcane 2013]. The Poincaré series approach appears better suited to the analytic number theory of higher-rank groups.

6. Basic identity for dual moment

6A. *Unipotent integration.* We are ready to work on the dual side of our moment formula. To simplify our argument, we will only consider $P = P^a(*; h)$ with $a = 1$ in the following. Suppose $\operatorname{Re} s > 1+\frac{\theta}{2}$, where θ is defined in Section 5A. We start by substituting the definition of P into the pairing in (5-34).

We find upon unfolding

$$(P, \mathbb{P}_2^3\Phi \cdot |\det *|^{\bar{s}-1/2})_{\Gamma_2\backslash \mathrm{GL}_2(\mathbb{R})}$$
$$= \int_0^\infty \int_0^\infty h(y_1)\cdot (y_0^2 y_1)^{s-1/2} \cdot \int_0^1 \widetilde{\Phi}\left[\begin{pmatrix} 1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(u_{1,2})\, du_{1,2} \frac{dy_0 dy_1}{y_0 y_1^2}. \tag{6-1}$$

The main task of this section is to compute the inner, "incomplete" unipotent integral in (6-1) in terms of the Fourier–Whittaker periods of Φ (see Definition 5.12), which are relevant in constructing various L-functions associated with Φ, as discussed in Section 5D.

While this can be achieved using the *full* Fourier expansion of [Jacquet et al. 1979a; 1979b] (see [Goldfeld 2015, Theorem 5.3.2]) and simplifying, we opt for a self-contained and conceptual treatment, which follows from two one-dimensional Fourier expansions and the automorphy of Φ. Essentially, this is where "summation formulae" come into play in our method, presented in an elementary, clean, and global manner.

Proposition 6.1. *For any automorphic function Φ of Γ_3, we have, for any $y_0, y_1 > 0$,*

$$\int_0^1 \Phi\left[\begin{pmatrix} 1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u_{1,2})\, du_{1,2}$$
$$= \sum_{a_0,a_1=-\infty}^{\infty} (\widehat{\Phi})_{(a_1,1)}\left[\begin{pmatrix} 1 & & \\ & 1 & \\ & -a_0 & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right]. \tag{6-2}$$

Proof. Firstly, we Fourier-expand along the abelian subgroup $\left\{\begin{pmatrix} 1 & & * \\ & 1 & \\ & & 1\end{pmatrix}\right\}$:

$$\int_0^1 \Phi\left[\begin{pmatrix} 1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u_{1,2})\, du_{1,2}$$
$$= \sum_{a_0=-\infty}^{\infty} \int_{\mathbb{Z}^2\backslash\mathbb{R}^2} \Phi\left[\begin{pmatrix} 1 & u_{1,2} & u_{1,3} \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix} y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u_{1,2} - a_0\cdot u_{1,3})\, du_{1,2}\, du_{1,3}. \tag{6-3}$$

Secondly, for each $a_0 \in \mathbb{Z}$, consider a unimodular change of variables of the form $(u_{1,2}, u_{1,3}) = (u_{1,2}', u_{1,3}')\cdot \left(\begin{smallmatrix} 1 & \\ -a_0 & 1\end{smallmatrix}\right)$. One can readily observe that

$$\begin{pmatrix} 1 & u_{1,2} & u_{1,3} \\ & 1 & \\ & & 1\end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & a_0 & 1\end{pmatrix}\begin{pmatrix} 1 & u_{1,2}' & u_{1,3}' \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix} 1 & & \\ & 1 & \\ & -a_0 & 1\end{pmatrix}.$$

Together with the automorphy of Φ with respect to Γ_3, we have

$$\int_0^1 \Phi\left[\begin{pmatrix}1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-a_2\cdot u_{1,2})\, du_{1,2}$$
$$= \sum_{a_0=-\infty}^{\infty} \int_{\mathbb{Z}^2\backslash\mathbb{R}^2} \Phi\left[\begin{pmatrix}1 & u'_{1,2} & u'_{1,3} \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix}1 & & \\ & 1 & \\ & -a_0 & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u'_{1,2})\, du'_{1,2}\, du'_{1,3}. \tag{6-4}$$

The result follows from a third and final Fourier expansion along the abelian subgroup $\left\{\begin{pmatrix}1 & & \\ & 1 & * \\ & & 1\end{pmatrix}\right\}$:

$$\int_0^1 \Phi\left[\begin{pmatrix}1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u_{1,2})\, du_{1,2}$$
$$= \sum_{a_0,a_1=-\infty}^{\infty} \int_0^1\int_0^1\int_0^1 \Phi\left[\begin{pmatrix}1 & u_{1,2} & u_{1,3} \\ & 1 & u_{2,3} \\ & & 1\end{pmatrix}\begin{pmatrix}1 & & \\ & 1 & \\ & -a_0 & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right]$$
$$\cdot e(-u_{1,2}-a_1\cdot u_{2,3})\, du_{1,2}\, du_{1,3}\, du_{2,3}.$$

□

We then explicate Proposition 6.1 when Φ is a Maass cusp form of Γ_3. This constitutes the *basic identity* of the present article. Theorem 1.1 is a natural consequence of this identity and the diagonal/off-diagonal structures on the dual side become apparent (see Proposition 7.2).

Corollary 6.2. *Suppose Φ is a Maass cusp form of Γ_3. Then*

$$\int_0^1 \Phi\left[\begin{pmatrix}1 & u_{1,2} & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix}\right] e(-u_{1,2})\, du_{1,2}$$
$$= \sum_{a_1\neq 0} \frac{\mathcal{B}_\Phi(a_1,1)}{|a_1|}\cdot W_{\alpha(\Phi)}(|a_1|y_0, y_1) + \sum_{a_0\neq 0}\sum_{a_1\neq 0} \frac{\mathcal{B}_\Phi(a_1,1)}{|a_1|}\cdot W_{\alpha(\Phi)}\left(\frac{|a_1|y_0}{1+(a_0y_0)^2}, y_1\sqrt{1+(a_0y_0)^2}\right)$$
$$\cdot e\left(-\frac{a_0a_1y_0^2}{1+(a_0y_0)^2}\right). \tag{6-5}$$

Proof. By cuspidality, $(\widehat{\Phi})_{(0,1)} \equiv 0$. The result follows from a straight-forward linear algebra calculation

$$\begin{pmatrix}1 & & \\ & 1 & \\ & -a_0 & 1\end{pmatrix}\begin{pmatrix}y_0y_1 & & \\ & y_0 & \\ & & 1\end{pmatrix} \equiv \begin{pmatrix}1 & & \\ & 1 & -\frac{a_0y_0^2}{1+(a_0y_0)^2} \\ & & 1\end{pmatrix}\begin{pmatrix}\frac{y_0}{1+(a_0y_0)^2}\cdot y_1\sqrt{1+(a_0y_0)^2} & & \\ & \frac{y_0}{1+(a_0y_0)^2} & \\ & & 1\end{pmatrix} \tag{6-6}$$

under the right quotient by $O_3(\mathbb{R})\cdot\mathbb{R}^{\times}$. One can verify Equation (6-6) using the formula stated in Section 2.4 of [Buttcane 2018] or the mathematica command *IwasawaForm*[] in the GL(n)*pack* (*gln.m*). The user manual and the package can both be downloaded from Kevin A. Broughan's website: see https://www.math.waikato.ac.nz/ kab/glnpack.html. ◻

6B. *Initial simplification and absolute convergence.* We temporarily restrict ourselves to the vertical strip $1+\frac{\theta}{2}<\sigma:=\operatorname{Re}s<4$. As we will see, this guarantees absolute convergence of sums and integrals.

Suppose $H\in\mathcal{C}_\eta$ with $\eta>40$ (see Proposition 5.10). Then the bound (5-12) for $h:=H^{\flat}$ implies its Mellin transform $\widetilde{h}(w):=\int_0^\infty h(y)y^w\,d^\times y$ is holomorphic on the strip $|\operatorname{Re}w|<\eta$. Substituting (6-5) into (6-1), and apply the changes of variables $y_0\to|a_1|^{-1}y_0$, $y_0\to|a_0|^{-1}y_0$ to the first, second piece of the resultant, we have

$$\begin{aligned}(P,\mathbb{P}_2^3\Phi\cdot|\det *|^{\bar{s}-1/2})_{\Gamma_2\backslash \mathrm{GL}_2(\mathbb{R})}&\\ =2\cdot L(2s,\Phi)\cdot\int_0^\infty\int_0^\infty h(y_1)\cdot(y_0^2y_1)^{s-1/2}\cdot W_{-\alpha(\Phi)}(y_0,y_1)\frac{dy_0dy_1}{y_0y_1^2}+OD_\Phi(s),\end{aligned}\tag{6-7}$$

where $OD_\Phi(s)$ is defined below.

Definition 6.3. Define $OD_\Phi(s)$ as

$$\begin{aligned}OD_\Phi(s):=\sum_{a_0\neq0}\sum_{a_1\neq0}\frac{\mathcal{B}_\Phi(1,a_1)}{|a_0|^{2s-1}|a_1|}\cdot\int_0^\infty\int_0^\infty h(y_1)\cdot(y_0^2y_1)^{s-1/2}\cdot e\left(\frac{a_1}{a_0}\cdot\frac{y_0^2}{1+y_0^2}\right)&\\ \cdot W_{-\alpha(\Phi)}\left(\left|\frac{a_1}{a_0}\right|\cdot\frac{y_0}{1+y_0^2},\,y_1\sqrt{1+y_0^2}\right)\frac{dy_0dy_1}{y_0y_1^2}.&\end{aligned}\tag{6-8}$$

Proposition 6.4. *When $H\in\mathcal{C}_\eta$ and $4>\sigma>\frac{1+\theta}{2}$, we have*

$$\begin{aligned}\int_0^\infty\int_0^\infty h(y_1)\cdot(y_0^2y_1)^{s-1/2}\cdot W_{-\alpha(\Phi)}(y_0,y_1)\frac{dy_0dy_1}{y_0y_1^2}&\\ =\frac{\pi^{-3s}}{8}\cdot\int_{(0)}\frac{H(\mu)}{|\Gamma(\mu)|^2}\cdot\prod_{i=1}^3\Gamma\left(\frac{s+\mu-\alpha_i}{2}\right)\Gamma\left(\frac{s-\mu-\alpha_i}{2}\right)\frac{d\mu}{2\pi i}.&\end{aligned}\tag{6-9}$$

Proof. From Proposition 5.11, we have

$$\begin{aligned}\int_0^\infty\int_0^\infty h(y_1)\cdot(y_0^2y_1)^{s-1/2}\cdot W_{-\alpha(\Phi)}(y_0,y_1)\frac{dy_0dy_1}{y_0y_1^2}&\\ =\frac{1}{2}\cdot\int_{(0)}\frac{H(\mu)}{|\Gamma(\mu)|^2}\cdot\int_0^\infty\int_0^\infty W_\mu(y_1)W_{-\alpha(\Phi)}(y_0,y_1)(y_0^2y_1)^{s-1/2}\frac{dy_0dy_1}{y_0y_1^2}\frac{d\mu}{2\pi i}.&\end{aligned}$$

The y_0, y_1-integrals can be evaluated by Proposition 5.7 and (6-9) follows. Moreover, the right side of (6-9) is holomorphic on $\sigma>0$. ◻

Proposition 6.5. *The off-diagonal $OD_\Phi(s)$ converges absolutely when $4>\sigma>1+\frac{\theta}{2}$ and $H\in\mathcal{C}_\eta$ $(\eta>40)$.*

Proof. Upon inserting absolute values, breaking up the y_0-integral into $\int_0^1+\int_1^\infty$, and applying the bounds (5-6) and $|\mathcal{B}_\Phi(1,a_1)|\ll|a_1|^\theta$, observe that

$$OD_\Phi(s)$$
$$\ll\sum_{a_0=1}^\infty\sum_{a_1=1}^\infty\frac{1}{a_0^{2\sigma-1}a_1^{1-\theta}}\left(\int_{y_0=1}^\infty+\int_{y_0=0}^1\right)\int_{y_1=0}^\infty|h(y_1)|(y_0^2y_1)^{\sigma-1/2}\left(\frac{a_1a_0^{-1}y_0}{1+y_0^2}\right)^{A_0}\left(y_1\sqrt{1+y_0^2}\right)^{A_1}\frac{dy_0dy_1}{y_0y_1^2},$$

where the implicit constant depends only on Φ, A_0, A_1 with $-\infty<A_0,A_1<1$. We are allowed to choose different A_0, A_1 in different ranges of the y_0, y_1-integrals.

The convergence of both of the series is guaranteed if

$$A_0<-\theta\quad\text{and}\quad\sigma>1-\frac{A_0}{2}.\tag{6-10}$$

We now show that if (6-10) and

$$A_1<A_0-2\sigma+1\tag{6-11}$$

both hold, then the y_0-integrals converge. Indeed, observe that $2\sigma+A_0-2>-1$ (by (6-10)), and

$$\int_{y_0=0}^1y_0^{2\sigma+A_0-2}(1+y_0^2)^{A_1/2-A_0}\,dy_0\asymp_{A_0,A_1}\int_{y_0=0}^1y_0^{2\sigma+A_0-2}\,dy_0.$$

So, the last integral converges. Also, (6-10) and (6-11) imply $A_1<\min\{1,2A_0\}$ and thus,

$$\int_{y_0=1}^\infty y_0^{2\sigma+A_0-2}(1+y_0^2)^{A_1/2-A_0}\,dy_0\le\int_{y_0=1}^\infty y_0^{2\sigma+A_1-A_0-2}\,dy_0.$$

The last integral converges because of (6-11).

For the y_1-integral, the integrals

$$\int_{y_1=1}^\infty|h(y_1)|y_1^{\sigma+A_1-5/2}\,dy_1\quad\text{and}\quad\int_{y_1=0}^1|h(y_1)|y_1^{\sigma+A_1-5/2}\,dy_1$$

converge whenever $H\in\mathcal{C}_\eta$ (we then have (5-12)) and

$$\eta>\left|\sigma+A_1-\tfrac{3}{2}\right|.\tag{6-12}$$

Let $\delta:=\sigma-1-(\theta/2)(>0)$. In view of (6-10) and (6-11), we may take $A_0:=-\theta-\delta$ and $A_1:=-2\theta-1-4\delta$. Also, (6-12) trivially holds as $\eta>40$ and $\sigma<4$. The result follows. ☐

Remark 6.6. Readers may notice the similarity between (3-2) and the inner product construction of the Kuznetsov formula. Indeed, $\mathbb{P}_2^3\Phi$ is an infinite sum of Poincaré series for $SL_2(\mathbb{Z})$ due to its Fourier expansion, though we never adopt this perspective in this article. This serves as a $GL(3)\times GL(2)$ analog to the Kuznetsov formula. However, there are key differences. Our moment identity equates two unfoldings, rather than comparing spectral and geometric expansions.

The second difference is technical. In the Kuznetsov formula, the oscillatory factors can be eliminated to obtain a "primitive" trace formula, see [Goldfeld and Kontorovich 2013; Zhou 2014; Goldfeld et al.

2021]. However, this does not work here — we have yet to analytically continue into the critical strip in Proposition 6.5. Here, the oscillatory factor in $OD_\Phi(s)$ is crucial, arising naturally from the abstract characterization of Whittaker functions.

7. Structure of the off-diagonal

Fix $\epsilon := \frac{1}{100}$ (say), $0 < \phi < \frac{\pi}{2}$, and consider the domain $1 + \frac{\theta}{2} + \epsilon < \sigma < 4$ in this section to maintain absolute convergence. We will stick with this choice of ϵ for the rest of this article and the number ϕ here should not pose any confusion with the basis of cusp forms (ϕ_j) of Γ_2. We define a perturbed version of $OD_\Phi(s)$ as follows:

$$OD_\Phi(s;\phi) := \sum_{a_0\neq 0}\sum_{a_1\neq 0} \frac{\mathcal{B}_\Phi(1,a_1)}{|a_0|^{2s-1}|a_1|}\int_0^\infty\int_0^\infty h(y_1)\cdot(y_0^2y_1)^{s-1/2}W_{-\alpha(\Phi)}\left(\left|\frac{a_1}{a_0}\right|\frac{y_0}{1+y_0^2}, y_1\sqrt{1+y_0^2}\right) \cdot e\left(\frac{a_1}{a_0}\frac{y_0^2}{1+y_0^2};\phi\right)\frac{dy_0dy_1}{y_0y_1^2}, \tag{7-1}$$

where

$$e(x;\phi) := \int_{(\epsilon)} |2\pi x|^{-u}e^{iu\phi\,\mathrm{sgn}(x)}\Gamma(u)\frac{du}{2\pi i} \quad (x\in\mathbb{R}-\{0\}). \tag{7-2}$$

In Proposition 7.3, we will show that

$$\lim_{\phi\to\pi/2} OD_\Phi(s;\phi) = OD_\Phi(s) \tag{7-3}$$

on a smaller region of absolute convergence.

Remark 7.1. The goals of this section is to obtain an expression of $OD_\Phi(s;\phi)$ that

- reveals the structure of the dual moment;
- can be analytically continued into the critical strip;
- and will allow us to pass to the limit $\phi\to\pi/2$ (in the critical strip).

Given these considerations, it is natural to work on the dual side of the Mellin transforms, which also allows for the separation of variables. The main result of this section is as follows:

Proposition 7.2 (dual moment). *Let $H\in\mathcal{C}_\eta$ $(\eta > 40)$ and $\phi\in(0,\pi/2)$. On the vertical strip*

$$1 + \frac{\theta}{2} + \epsilon < \sigma < 4, \tag{7-4}$$

we have

$$OD_\Phi(s;\phi) = \frac{1}{4}\int_{(1+\theta+2\epsilon)} \zeta(2s-s_0)L(s_0,\Phi)\cdot\sum_{\delta=\pm}(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi)\frac{ds_0}{2\pi i}, \tag{7-5}$$

where the transform of H is given by

$$(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi) := \int_{(15)}\int_{(\epsilon)} \widetilde{h}\left(s-s_1-\tfrac{1}{2}\right)\cdot\mathcal{G}_\Phi^{(\delta)}(s_1,u;s_0,s;\phi)\frac{du}{2\pi i}\frac{ds_1}{2\pi i}, \tag{7-6}$$

with $h := H^\flat$, $G_\Phi := G_{\alpha(\Phi)}$ *as defined in* (5-5), *and*

$$\mathcal{G}_\Phi^{(\delta)}(s_1, u; s_0, s; \phi) := G_\Phi(s_0-u, s_1)\cdot(2\pi)^{-u}e^{i\delta\phi u}\Gamma(u)\cdot\frac{\Gamma\big(\frac{u+1-2s+s_1-s_0}{2}\big)\Gamma\big(\frac{2s-s_0-u}{2}\big)}{\Gamma\big(\frac{1+s_1}{2}-s_0\big)}. \tag{7-7}$$

Proof. Plug-in the expression for $W_{-\alpha(\Phi)}$ from Proposition 5.3 into $OD_\Phi(s;\phi)$ with

$$\sigma_1 := 15 \quad \text{and} \quad 1+\theta<\sigma_0<2\sigma-1-\epsilon. \tag{7-8}$$

Insert absolute values to the resulting expression, the sums and integrals are bounded by

$$\sum_{\delta:=\operatorname{sgn}(a_0a_1)=\pm}\bigg(\sum_{a_0\neq 0}\frac{1}{|a_0|^{2\sigma-\sigma_0-\epsilon}}\bigg)\bigg(\sum_{a_1\neq 0}\frac{|\mathcal{B}_\Phi(1,a_1)|}{|a_1|^{\sigma_0+\epsilon}}\bigg)\bigg(\int_{(\sigma_0)}\int_{(\sigma_1)}|G_\Phi(s_0,s_1)|\,|ds_0||ds_1|\bigg)$$
$$\cdot\bigg(\int_{(\epsilon)}|e^{i\delta\phi u}\Gamma(u)|\,|du|\bigg)\bigg(\int_0^\infty y_0^{-\sigma_0-2\epsilon+2\sigma}(1+y_0^2)^{\sigma_0+\epsilon-(1+\sigma_1)/2}\,d^\times y_0\bigg)$$
$$\cdot\bigg(\int_0^\infty|h(y_1)|\cdot y_1^{\sigma-\sigma_1-1/2}\,d^\times y_1\bigg). \tag{7-9}$$

Observe that:

- By Stirling's formula, the s_0, s_1, u-integrals converge as long as

$$\sigma_0,\sigma_1,\epsilon>0, \qquad \phi\in(0,\pi/2). \tag{7-10}$$

- The y_0-integral converges as long as

$$\sigma_0+2\epsilon<2\sigma<\sigma_1-\sigma_0+1. \tag{7-11}$$

- By the bound $|\mathcal{B}_\Phi(1,a_1)|\ll|a_1|^\theta$, the a_0-sum and the a_1-sum converge as long as

$$2\sigma-1>\sigma_0+\epsilon>1+\theta. \tag{7-12}$$

Under (7-8), items (7-10), (7-11), (7-12) hold. Moreover, the y_1-integral converges by (5-12) and $H\in\mathcal{C}_\eta$ ($\eta>40$). Now, upon rearranging sums and integrals, and noticing that $\mathcal{B}_\Phi(1,a_1)=\mathcal{B}_\Phi(1,-a_1)$, we have

$$OD_\Phi(s;\phi)=2\sum_{\delta=\pm}\int_{(\sigma_0)}\int_{(\sigma_1)}\int_{(\epsilon)}\frac{G_\Phi(s_0,s_1)}{4}\cdot(2\pi)^{-u}e^{i\delta\phi u}\Gamma(u)$$
$$\cdot\bigg(\int_0^\infty h(y_1)y_1^{s-s_1-1/2}d^\times y_1\bigg)\bigg(\int_0^\infty y_0^{-s_0-2u+2s}(1+y_0^2)^{s_0+u-(1+s_1)/2}\,d^\times y_0\bigg)$$
$$\cdot\bigg(\sum_{a_0=1}^\infty\sum_{a_1=1}^\infty\frac{\mathcal{B}_\Phi(1,a_1)}{a_0^{2s-1}a_1}\Big(\frac{a_1}{a_0}\Big)^{1-s_0-u}\bigg)\frac{ds_0}{2\pi i}\frac{ds_1}{2\pi i}\frac{du}{2\pi i}. \tag{7-13}$$

Recall the integral identity

$$\int_0^\infty y_0^v(1+y_0^2)^A\,d^\times y_0=\tfrac{1}{2}\frac{\Gamma(-A-v/2)\Gamma(v/2)}{\Gamma(-A)} \tag{7-14}$$

for $0 < \operatorname{Re} v < -2\operatorname{Re} A$. It follows that

$$OD_\Phi(s;\phi) = 2\sum_{\delta=\pm}\int_{(\sigma_0)}\int_{(\sigma_1)}\int_{(\epsilon)} \zeta(2s-s_0-u)L(s_0+u;\,\Phi)\cdot\widetilde{h}\big(s-s_1-\tfrac{1}{2}\big) \cdot\frac{G_\Phi(s_0,s_1)}{4}\cdot(2\pi)^{-u}e^{i\delta\phi u}\Gamma(u)\cdot\frac{1}{2}\frac{\Gamma\big(s-\frac{s_0}{2}-u\big)\Gamma\big(\frac{1+s_1-s_0}{2}-s\big)}{\Gamma\big(\frac{1+s_1}{2}-s_0-u\big)}\frac{ds_0}{2\pi i}\frac{ds_1}{2\pi i}\frac{du}{2\pi i}. \tag{7-15}$$

We pick the contour $(\sigma_0) := (1+\theta+\epsilon)$, thus imposing (7-4). To isolate the nonarchimedean part of $OD_\Phi(s;\phi)$, we change variables to $s_0' = s_0+u$. Substituting the expression for $G_\Phi(s_0'-u, s_1)$ (see (5-5)), we obtain (7-5)–(7-7). The absolute convergence proven earlier also ensures the holomorphy of the integral transform $(\mathcal{F}_\Phi^{(\delta)}h)(s_0', s;\phi)$ on the domain

$$\sigma < 4 \quad\text{and}\quad 1+\theta+\epsilon < \sigma_0' < 2\sigma-1. \tag{7-16}$$

This completes the proof. □

Proposition 7.3. *For $4 > \sigma > (3+\theta)/2$ and $H \in \mathcal{C}_\eta$, we have*

$$\lim_{\phi\to\pi/2} OD_\Phi(s;\phi) = OD_\Phi(s). \tag{7-17}$$

Proof. Let $\epsilon := \frac{1}{100}$, $\sigma_1 := 15$, and pick any σ_0 satisfying

$$\tfrac{3}{2}+\theta+\epsilon < \sigma_0 < 2\sigma-1-\epsilon. \tag{7-18}$$

Denote by $\mathcal{C}_\epsilon$ the indented path consisting of the line segments:

$$-\tfrac{1}{2}-\epsilon-i\infty \to -\tfrac{1}{2}-\epsilon-i \to \epsilon-i \to \epsilon+i \to -\tfrac{1}{2}-\epsilon+i \to -\tfrac{1}{2}-\epsilon+i\infty.$$

Replace $e(x;\phi)$ in (7-13) by the expression:

$$e(x;\phi) = \int_{\mathcal{C}_\epsilon} |2\pi x|^{-u} e^{iu\phi\,\operatorname{sgn}(x)}\Gamma(u)\frac{du}{2\pi i}. \tag{7-19}$$

Note that $|e^{iu\phi\,\operatorname{sgn}(x)}\Gamma(u)| \ll_\epsilon (1+|\operatorname{Im} u|)^{-1-\epsilon}$ for $u \in \mathcal{C}_\epsilon$ and $\phi \in (0,\pi/2]$. Insert absolute values in (7-13). The resulting sums and integrals converge absolutely when $\phi \in (0,\pi/2]$ and (7-18) holds, which can be seen by the same argument following (7-9). Apply dominated convergence and shift the contour of the u-integral to $-\infty$, the residual series obtained is exactly $e((a_1/a_0)(y_0^2/(1+y_0^2)))$. This completes the proof. □

Now, $OD_\Phi(s;\phi)$ is expressed as Mellin-Barnes integrals. The Γ-factors from Proposition 5.3 and (7-2) alone are not sufficient for our goals (see Remark 7.1 and (7-10), (7-11), (7-12)). The three extra Γ-factors brought by the y_0-integral, which mix all integration variables, will play an important role in Section 8-9.

8. Analytic properties of the archimedean transform

In (7-5), the factors $\zeta(2s-s_0)$ and $L(s_0,\Phi)$ are known to admit holomorphic continuation and have polynomial growth in vertical strips, except on the line $2s-s_0=1$. We also examine the archimedean part of (7-5), i.e., the integral transform

$$(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi) := \int_{(15)}\int_{(\epsilon)} \widetilde{h}\big(s-s_1-\tfrac{1}{2}\big)\cdot \mathcal{G}_\Phi^{(\delta)}(s_1,u;s_0,s;\phi)\frac{du}{2\pi i}\frac{ds_1}{2\pi i}, \tag{8-1}$$

where $h := H^\flat$ and $\mathcal{G}_\Phi^{(\delta)}(\cdots)$ is defined in (7-7). In Section 7, we have shown that when $\phi\in(0,\pi/2)$, the function $(s_0,s)\mapsto(\mathcal{F}_\Phi^{(\delta)}h)(s_0,s;\phi)$ is holomorphic on the domain (7-16), i.e.,

$$\sigma<4 \quad\text{and}\quad 1+\theta+\epsilon<\sigma_0<2\sigma-1.$$

In this section, we establish a larger region of holomorphy for $(s_0,s)\mapsto(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi)$ that holds for $\phi\in(0,\pi/2]$. We write

$$s=\sigma+it,\qquad s_0=\sigma_0+it_0,\qquad s_1=\sigma_1+it_1,\qquad\text{and}\qquad u=\epsilon+iv,$$

with $\epsilon := \frac{1}{100}$. It is sufficient to consider s inside the rectangular box $\epsilon<\sigma<4$ and $|t|\le T$, for any given $T\ge 1000$. Moreover, $\alpha_k := i\gamma_k\in i\mathbb{R}$ $(k=1,2,3)$ by our assumptions on Φ. The main result of this section can be stated as follows:

Proposition 8.1. *Suppose $H\in\mathcal{C}_\eta$:*

(1) *For any $\phi\in(0,\pi/2]$, the transform $(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi)$ is holomorphic on the domain*

$$\sigma_0>\epsilon,\qquad \sigma<4,\qquad\textit{and}\qquad 2\sigma-\sigma_0-\epsilon>0. \tag{8-2}$$

(2) *Whenever $(\sigma_0,\sigma)\in$ (8-2), $|t|<T$, and $\phi\in(0,\pi/2)$, the transform $(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi)$ has exponential decay as $|t_0|\to\infty$. Note: The explicit estimate is stated in the proof below and the implicit constant depends only on T and Φ.*

Remark 8.2. The domain (8-2) is chosen in a way that the function $(s_0,s)\mapsto\mathcal{G}_\Phi^{(\delta)}(s_1,u;s_0,s;\phi)$ is holomorphic on (8-2) when $\operatorname{Re}s_1=\sigma_1\ge 15$ and $\operatorname{Re}u=\epsilon$. Moreover, if we have $15\le\sigma_1\le\eta-\frac{1}{2}$ and (8-2), then $s-s_1-\frac{1}{2}$ lies inside the region of holomorphy of $\widetilde{h}$.

Remark 8.3. As we shall see in Proposition 9.2, the region of holomorphy (8-2) is essentially optimal in terms of σ_0.

Proof. The proof is based on a careful application of the Stirling estimate

$$|\Gamma(a+ib)|\asymp_a(1+|b|)^{a-1/2}e^{-(\pi/2)|b|}\qquad(a\neq 0,-1,-2,\dots,\ b\in\mathbb{R}) \tag{8-3}$$

to the kernel function $\mathcal{G}_{\Phi}^{(\delta)}(s_1, u; s_0, s; \phi)$. The following set of conditions will be repeated throughout the proof:

$$\begin{gathered} 0 < \phi \leq \pi/2, \\ \sigma_0 > \epsilon, \quad \sigma < 4, \quad 2\sigma - \sigma_0 - \epsilon > 0, \\ \operatorname{Re} s_1 = \sigma_1 \geq 15, \quad \operatorname{Re} u = \epsilon. \end{gathered} \tag{8-4}$$

Assuming (8-4), we apply (8-3) to the kernel function (7-7). It follows that

$$\begin{aligned} &|\mathcal{G}_{\Phi}^{(\delta)}(s_1, u; s_0, s; \phi)| \\ &\quad \asymp (1+|v|)^{\epsilon-1/2} e^{-((\pi/2)-\phi)|v|} \cdot \prod_{k=1}^{3} (1+|t_1-\gamma_k|)^{(\sigma_1-1)/2} e^{-(\pi/4)|t_1-\gamma_k|} \\ &\qquad \cdot \prod_{k=1}^{3} (1+|t_0-v+\gamma_k|)^{(\sigma_0-\epsilon-1)/2} e^{-(\pi/4)|t_0-v+\gamma_k|} \cdot (1+|2t-t_0-v|)^{(2\sigma-1-\sigma_0-\epsilon)/2} e^{-(\pi/4)|2t-t_0-v|} \\ &\qquad \cdot (1+|v-2t+t_1-t_0|)^{(\epsilon-2\sigma+\sigma_1-\sigma_0)/2} e^{-(\pi/4)|v-2t+t_1-t_0|} \\ &\qquad \cdot (1+|t_1-2t_0|)^{-(\sigma_1/2-\sigma_0)} e^{\frac{\pi}{4}|t_1-2t_0|} \cdot (1+|t_0+t_1-v|)^{-(\sigma_0+\sigma_1-\epsilon-1)/2} e^{(\pi/4)|t_0+t_1-v|}, \end{aligned} \tag{8-5}$$

where the implicit constant depends at most on σ_1. Note that the domain (8-2) for (σ, σ_0) is bounded and thus the estimate is uniform in $\sigma, \sigma_0, \epsilon$. This will be assumed for all estimates in the rest of this section.

Let $\mathcal{P}_s^{\Phi}(t_0, t_1, v)$ be the "polynomial part" of (8-5) and the "exponential phase" of (8-5) be

$$\mathcal{E}_s^{\Phi}(t_0, t_1, v) := \sum_{k=1}^{3} \{|t_1-\gamma_k| + |t_0-v+\gamma_k|\} + |2t-t_0-v| + |v-2t+t_1-t_0| - |t_1-2t_0| - |t_0+t_1-v|.$$

We first examine $\mathcal{E}_s^{\Phi}(t_0, t_1, v)$, which determines the effective support of $(\mathcal{F}_{\Phi}^{(\delta)} H)(s_0, s; \phi)$. By the triangle inequality and the fact $\gamma_1 + \gamma_2 + \gamma_3 = 0$, we have

$$|\mathcal{G}_{\Phi}^{(\delta)}(s_1, u; s_0, s; \phi)| \ll_{\sigma_1} e^{\pi T} \cdot \mathcal{P}_s^{\Phi}(t_0, t_1, v) \cdot \exp\!\left(-\tfrac{\pi}{4} \mathcal{E}(t_0, t_1, v)\right) \cdot e^{-(\pi/2-\phi)|v|} \tag{8-6}$$

with

$$\mathcal{E}(t_0, t_1, v) := 3|t_1| + 3|t_0 - v| - |t_1 - 2t_0| + |v + t_1 - t_0| + |t_0 + v| - |t_0 + t_1 - v|, \tag{8-7}$$

whenever we have (8-4) and $|t| \leq T$.

Claim 8.4. *For any $t_0, t_1, v \in \mathbb{R}$, we have $\mathcal{E}(t_0, t_1, v) \geq 0$. Equality holds if and only if*

$$t_1 = 0 \quad \textit{and} \quad t_0 - v = 0. \tag{8-8}$$

Proof. Adding up the inequalities $|t_1| + |t_0 - v| \geq |t_0 + t_1 - v|$ and $|v + t_1 - t_0| + |t_0 + v| \geq |t_1 - 2t_0|$, we have

$$\mathcal{E}(t_0, t_1, v) \geq 2(|t_1| + |t_0 - v|) \geq 0. \tag{8-9}$$

The equality case is apparent. ∎

Claim 8.5. *When (8-4) and $|t| \le T$ hold, the integral*

$$\iint_{\substack{(\operatorname{Re} s_1, \operatorname{Re} u)=(\sigma_1,\epsilon),\\ (t_1,v):\text{(8-11) holds}}} \widetilde{h}\big(s - s_1 - \tfrac{1}{2}\big) \cdot \mathcal{G}_\Phi^{(\delta)}(s_1, u; s_0, s; \phi) \frac{du}{2\pi i} \frac{ds_1}{2\pi i} \tag{8-10}$$

has exponential decay as $|t_0| \to \infty$, where

$$|t_1| > \log^2(3 + |t_0|) \quad \text{or} \quad |v - t_0| > \log^2(3 + |t_0|). \tag{8-11}$$

Proof. In the case of (8-11), we have

$$\mathcal{E}(t_0, t_1, v) > \log^2(3 + |t_0|) + |t_1| + |t_0 - v| \tag{8-12}$$

from (8-9). The polynomial part $\mathcal{P}_s^\Phi(t_0, t_1, v)$ can be crudely bounded by

$$\mathcal{P}_s^\Phi(t_0, t_1, v) \ll_{\Phi,\sigma_1,T} [(1 + |t_1|)(1 + |v - t_0|)(1 + |t_0|)]^{A(\sigma_1)}, \tag{8-13}$$

where $A(\sigma_1) > 0$ is some constant.

Putting (8-12), (8-13), and the bound $e^{-(\pi/2-\phi)|v|} \le 1$ ($\phi \in (0, \pi/2]$) into (8-6), we obtain

$$|\mathcal{G}_\Phi^{(\delta)}(s_1, u; s_0, s; \phi)| \\ \ll_{\Phi,\sigma_1,T} (1 + |t_0|)^{A(\sigma_1)} e^{-(\pi/4)\log^2(3+|t_0|)} \cdot [(1 + |t_1|)(1 + |v - t_0|)]^{A(\sigma_1)} e^{-(\pi/4)[|t_1|+|t_0-v|]} \tag{8-14}$$

whenever (8-11), (8-4), and $|t| \le T$ hold. The boundedness of $\widetilde{h}$ on vertical strips implies that (8-10) is

$$\ll_{\sigma_1,\Phi,T} (1 + |t_0|)^{A(\sigma_1)} e^{-(\pi/4)\log^2(3+|t_0|)}. \tag{8-15}$$

This proves Claim 8.5. □

Now, let $\phi \in (0, \pi/2]$ and consider $(\mathcal{F}_\Phi^{(\delta)} H)(s_0, s; \phi)$ as a function on the bounded domain

$$(\sigma_0, \sigma) \in \text{(8-2)}, \quad |t|, |t_0| \le T. \tag{8-16}$$

When $|t_1| > \log^2(3 + T)$ or $|v| > T + \log^2(3 + T)$, observe that (8-11) is satisfied and from (8-14),

$$|\mathcal{G}_\Phi^{(\delta)}(s_1, u; s_0, s; \phi)| \ll_{\Phi,T} [(1 + |t_1|)(1 + |v|)]^{A(15)} \cdot e^{-(\pi/4)[|t_1|+|v|]}. \tag{8-17}$$

The last function is clearly jointly integrable with respect to t_1, v, and by Remark 8.2, $(\mathcal{F}_\Phi^{(\delta)} H)(s_0, s; \phi)$ is a holomorphic function on (8-16). Since the choice of T is arbitrary, we arrive at the first conclusion of Proposition 8.1.

In the remaining part of this section, we prove the second assertion of Proposition 8.1. We estimate the contribution from

$$|t_1| \le \log^2(3 + |t_0|) \quad \text{and} \quad |v - t_0| \le \log^2(3 + |t_0|), \tag{8-18}$$

where the complementary part has been treated in Claim 8.5.

It suffices to restrict to the effective support (8-8). The polynomial part can be essentially computed by substituting $t_1 := 0$ and $v := t_0$. More precisely, when (8-18) and $|t_0| \gg_T 1$ hold, there are only two possible scenarios for the factors $1+|(\cdots)|$ in (8-5): either $1+|(\cdots)| \asymp |t_0|$, or $\log^{-C}(3+|t_0|) \ll 1+|(\cdots)| \ll \log^{C}(3+|t_0|)$ for some absolute constant $C>0$.

In the case of (8-18), we apply the bounds $e^{-(\pi/4)\mathcal{E}(t_0,t_1,v)} \le 1$ and $e^{-(\pi/2-\phi)|v|} \le e^{-(1/2)(\pi/2-\phi)|t_0|}$ for $|t_0| \gg 1$ to (8-6). As a result, if we also have (8-4), $|t|<T$, and $|t_0|>8T$, then

$$|\mathcal{G}_{\Phi}^{(\delta)}(s_1, u; s_0, s; \phi)| \ll_{\sigma_1,\Phi,T} |t_0|^{7-\sigma_1/2} e^{-(1/2)(\pi/2-\phi)|t_0|} \log^{B(\sigma_1)}|t_0| \tag{8-19}$$

and

$$\iint_{\substack{(\operatorname{Re} s_1, \operatorname{Re} u)=(\sigma_1,\epsilon),\\ (t_1,v):\text{(8-18) holds}}} \widetilde{h}\bigl(s-s_1-\tfrac{1}{2}\bigr)\cdot \mathcal{G}_{\Phi}^{(\delta)}(s_1,u;s_0,s;\phi)\frac{du}{2\pi i}\frac{ds_1}{2\pi i}$$
$$\ll_{\sigma_1,\Phi,T} |t_0|^{7-\sigma_1/2} e^{-(1/2)(\pi/2-\phi)|t_0|} \log^{4+B(\sigma_1)}|t_0|, \tag{8-20}$$

where $B(\sigma_1)>0$ is some constant. If $\phi<\pi/2$, then there is exponential decay in (8-20) as $|t_0|\to\infty$. Therefore, the second conclusion of the proposition follows from (8-20) and (8-15) (putting $\sigma_1=15$). $\square$

9. Analytic continuation of the off-diagonal (proof of Theorem 1.1)

Recall that

$$OD_{\Phi}(s;\phi) = \tfrac{1}{4}\int_{(1+\theta+2\epsilon)} \zeta(2s-s_0)L(s_0,\Phi)\cdot \sum_{\delta=\pm}(\mathcal{F}_{\Phi}^{(\delta)}H)(s_0,s;\phi)\frac{ds_0}{2\pi i} \tag{9-1}$$

for $1+\frac{\theta}{2}+\epsilon<\sigma<4$ and $\phi\in(0,\pi/2)$, see Proposition 7.2.

9A. ***Step 1:*** We first obtain a holomorphic continuation of $OD_{\Phi}(s;\phi)$ up to $\operatorname{Re} s>\frac{1}{2}+\epsilon$ by shifting the s_0-integral to the left.

Fix any $\phi\in(0,\pi/2)$ and $T\ge 1000$. We first restrict ourselves to

$$1+\frac{\theta}{2}+2\epsilon<\sigma<4, \quad |t|<T. \tag{9-2}$$

Clearly, the pole $s_0=2s-1$ of $\zeta(2s-s_0)$ is on the right of the contour $\operatorname{Re} s_0 = 1+\theta+2\epsilon$ of the integral (7-5).

Let $T_0\gg 1$. The rectangle with vertices $2\epsilon\pm iT_0$ and $(1+\theta+2\epsilon)\pm iT_0$ in the s_0-plane lies inside the region of holomorphy (8-2) of $(\mathcal{F}_{\Phi}^{(\delta)}H)(s_0,s;\phi)$. The contribution from the horizontal segments $[2\epsilon\pm iT_0, (1+\theta+2\epsilon)\pm iT_0]$ tends to 0 as $T_0\to\infty$ by the exponential decay of $(\mathcal{F}_{\Phi}^{(\delta)}H)(s_0,s;\phi)$ (see Proposition 8.1), which surely counteracts the polynomial growth from $L(s_0,\Phi)$ and $\zeta(2s-s_0)$. As a result, we may shift the line of integration to $\operatorname{Re} s_0=2\epsilon$ and no pole is crossed. Hence,

$$OD_{\Phi}(s;\phi) = \tfrac{1}{4}\int_{(2\epsilon)} \zeta(2s-s_0)L(s_0,\Phi)\cdot \sum_{\delta=\pm}(\mathcal{F}_{\Phi}^{(\delta)}H)(s_0,s;\phi)\frac{ds_0}{2\pi i} \tag{9-3}$$

on (9-2). The right side of (9-3) is holomorphic on

$$\tfrac{1}{2}+\epsilon<\sigma<4, \quad |t|<T \tag{9-4}$$

and serves as an analytic continuation of $OD_\Phi(s;\phi)$ to (9-4) by using Proposition 8.1. Note that $\sigma > \frac{1}{2}+\epsilon$ implies the holomorphy of $\zeta(2s-s_0)$.

9B. ***Step 2: Crossing the polar line (shifting the s_0-integral again).*** Consider a subdomain of (9-4):

$$\tfrac{1}{2}+\epsilon < \sigma < \tfrac{3}{4}, \quad |t| < T. \tag{9-5}$$

Different from step 1, the pole $s_0 = 2s-1$ is now inside the rectangle with vertices $2\epsilon \pm iT_0$ and $\frac{1}{2} \pm iT_0$ provided $T_0 > 4T$. Such a rectangle lies in the region of holomorphy (8-2) of $(\mathcal{F}_\Phi^{(\delta)}H)(s_0, s; \phi)$. When $\phi < \pi/2$, the exponential decay of $(\mathcal{F}_\Phi^{(\delta)}H)(s_0, s; \phi)$ once again allows us to shift the line of integration from $\operatorname{Re} s_0 = 2\epsilon$ to $\operatorname{Re} s_0 = \frac{1}{2}$, crossing the pole of $\zeta(2s-s_0)$ which has residue -1. In other words,

$$OD_\Phi(s;\phi) = \tfrac{1}{4}L(2s-1,\Phi)\sum_{\delta=\pm}(\mathcal{F}_\Phi^{(\delta)}H)(2s-1,s;\phi) \\ + \tfrac{1}{4}\int_{(1/2)} \zeta(2s-s_0)L(s_0,\Phi)\cdot\sum_{\delta=\pm}(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\phi)\frac{ds_0}{2\pi i}. \tag{9-6}$$

On the line $\operatorname{Re} s_0 = \frac{1}{2}$, observe that $s \mapsto (\mathcal{F}_\Phi^{(\delta)}H)(s_0, s; \phi)$ is holomorphic on $\sigma > \frac{1}{4}+\frac{\epsilon}{2}$ by (8-2); whereas $s \mapsto \zeta(2s-s_0)$ is holomorphic on $\sigma < \frac{3}{4}$ as $2\sigma - \sigma_0 < 1$. As a result, the function $s \mapsto \int_{(1/2)}(\cdots)\frac{ds_0}{2\pi i}$ in (9-6) is holomorphic on the vertical strip

$$\tfrac{1}{4}+\tfrac{\epsilon}{2} < \sigma < \tfrac{3}{4}, \tag{9-7}$$

which is sufficient for our purpose.

Proposition 8.1 only asserts that the function $s \mapsto (\mathcal{F}_\Phi^{(\delta)}H)(2s-1, s; \phi)$ is holomorphic on $\frac{1}{2}+\epsilon < \sigma < 4$. However, it actually admits a continuation to the domain $\epsilon < \sigma < 4$ as we will see in Proposition 9.2.

9C. ***Step 3: Putting back $\phi \to \pi/2$—shifting the s_1-integral and refining Steps 1–2.*** By using estimate (8-14) and dominated convergence,

$$\lim_{\phi \to \pi/2}(\mathcal{F}_\Phi^{(\delta)}H)(2s-1,s;\phi) = (\mathcal{F}_\Phi^{(\delta)}H)(2s-1,s;\pi/2) \tag{9-8}$$

for $\frac{1}{2}+\epsilon < \sigma < 4$ and $|t| < T$. For the continuous part of (9-6), we need to extend Proposition 8.1 to pass to the limit $\phi \to \pi/2$. Using the Γ-factors from Proposition 5.3 and the analytic properties of $\widetilde{h}$, we shift the line of integration for the s_1-integral to achieve the necessary polynomial decay.

Proposition 9.1. *Let $H \in \mathcal{C}_\eta$. There exists a constant $B = B_\eta$ such that whenever $(\sigma_0, \sigma) \in$ (8-2), $|t| < T$, and $|t_0| \gg_T 1$, we have the estimate*

$$|(\mathcal{F}_\Phi^{(\delta)}H)(s_0,s;\pi/2)| \ll |t_0|^{8-\eta/2}\log^B|t_0|, \tag{9-9}$$

where the implicit constant depends only on η, T, Φ.

Proof. On domain (8-2), observe that the vertical strip $\operatorname{Re} s_1 \in \left[15, \eta - \frac{1}{2}\right]$ contains no pole of the function $s_1 \mapsto \mathcal{G}_\Phi^{(\delta)}(s_1, u; s_0, s; \phi)$, and it lies within the region of holomorphy of $\widetilde{h}$ (see Remark 8.2). The estimate

(8-14) allows us to shift the line of integration from $\mathrm{Re}\, s_1 = 15$ to $\mathrm{Re}\, s_1 = \eta - \frac{1}{2}$ in (7-6). Notice that the estimates done in Proposition 8.1 works for $\phi = \pi/2$ too. In particular, from (8-20) and (8-15), the bound (9-9) follows by taking $\sigma_1 := \eta - \frac{1}{2}$ therein (after the contour shift). This completes the proof. □

Suppose $(3+\theta)/2 < \sigma < 4$. By Proposition 7.3, (7-5) and (9-3), we have

$$\begin{aligned} OD_\Phi(s) &= \lim_{\phi \to \pi/2} OD_\Phi(s;\phi) \\ &= \lim_{\phi \to \pi/2} \tfrac{1}{4} \int_{(1+\theta+2\epsilon)} \zeta(2s-s_0) L(s_0,\Phi) \cdot \sum_{\delta=\pm} (\mathcal{F}_\Phi^{(\delta)} H)(s_0,s;\phi) \frac{ds_0}{2\pi i} \\ &= \lim_{\phi \to \pi/2} \tfrac{1}{4} \int_{(2\epsilon)} \zeta(2s-s_0) L(s_0,\Phi) \cdot \sum_{\delta=\pm} (\mathcal{F}_\Phi^{(\delta)} H)(s_0,s;\phi) \frac{ds_0}{2\pi i}. \end{aligned} \tag{9-10}$$

Proposition 9.1 ensures enough polynomial decay and hence the absolute convergence of (9-11) at $\phi = \pi/2$:

$$OD_\Phi(s) = \tfrac{1}{4} \int_{(2\epsilon)} \zeta(2s-s_0) L(s_0,\Phi) \cdot \sum_{\delta=\pm} (\mathcal{F}_\Phi^{(\delta)} H)(s_0,s;\pi/2) \frac{ds_0}{2\pi i}. \tag{9-11}$$

Now, (9-11) serves as an analytic continuation of $OD_\Phi(s)$ to the domain $\frac{1}{2} + \epsilon < \sigma < 4$.

On the smaller domain $\frac{1}{2} + \epsilon < \sigma < \frac{3}{4}$, the expressions (9-10) and (9-6) are equal. Then

$$\begin{aligned} OD_\Phi(s) &= \text{(9-10)} \\ &= \tfrac{1}{4} L(2s-1,\Phi) \sum_{\delta=\pm} (\mathcal{F}_\Phi^{(\delta)} H)(2s-1,s;\pi/2) \\ &\qquad + \tfrac{1}{4} \int_{(1/2)} \zeta(2s-s_0) L(s_0,\Phi) \cdot \sum_{\delta=\pm} (\mathcal{F}_\Phi^{(\delta)} H)(s_0,s;\pi/2) \frac{ds_0}{2\pi i} \end{aligned} \tag{9-12}$$

by dominated convergence and Proposition 8.1. The last integral is holomorphic on $\frac{1}{4} + \frac{\epsilon}{2} < \sigma < \frac{3}{4}$.

In the following, we write $(\mathcal{F}_\Phi H)(s_0,s) := (\mathcal{F}_\Phi^{+} H)(s_0,s;\pi/2) + (\mathcal{F}_\Phi^{-} H)(s_0,s;\pi/2)$. Duplication and reflection formulae of Γ-functions in the form

$$2^{-u}\Gamma(u) = \frac{1}{2\sqrt{\pi}} \cdot \Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{u+1}{2}\right) \quad \text{and} \quad \Gamma\left(\frac{1+u}{2}\right)\Gamma\left(\frac{1-u}{2}\right) = \pi \sec \frac{\pi u}{2},$$

lead to

$$\begin{aligned} (\mathcal{F}_\Phi H)(s_0,s) = \sqrt{\pi} \int_{(\eta-1/2)} \widetilde{h}\big(s-s_1-\tfrac{1}{2}\big) \pi^{-s_1} \frac{\prod_{i=1}^{3} \Gamma\big(\frac{s_1-\alpha_i}{2}\big)}{\Gamma\big(\frac{1+s_1}{2} - s_0\big)} \\ \cdot \int_{(\epsilon)} \frac{\Gamma\big(\frac{u}{2}\big)\Gamma\big(\frac{s_1-(s_0-u)}{2} + \frac{1}{2} - s\big) \cdot \prod_{i=1}^{3} \Gamma\big(\frac{(s_0-u)+\alpha_i}{2}\big)\Gamma\big(s - \frac{s_0+u}{2}\big)}{\Gamma\big(\frac{1-u}{2}\big)\Gamma\big(\frac{(s_0-u)+s_1}{2}\big)} \frac{du}{2\pi i}\frac{ds_1}{2\pi i}. \end{aligned} \tag{9-13}$$

In Section 10, we will work with this expression further.

9D. ***Step 4: Continuation of the residual term — shifting the u-integral.***

Proposition 9.2. *Let $H \in \mathcal{C}_\eta$. The function $s \mapsto (\mathcal{F}_\Phi H)(2s-1, s)$ can be holomorphically continued to the vertical strip $\epsilon < \sigma < 4$ except at the three simple poles: $s = (1-\alpha_i)/2$ $(i = 1, 2, 3)$, where $(\alpha_1, \alpha_2, \alpha_3)$ are the Langlands parameters of the Maass cusp form Φ.*

Proof. We will prove a stronger result in Theorem 10.2. However, a simpler argument suffices for the time being. Suppose $\frac{1}{2} + \epsilon < \sigma < 4$ and $s_0 = 2s - 1$. In (9-13), we shift the line of integration from $\operatorname{Re} u = \epsilon$ to $\operatorname{Re} u = -1.9$:

$$(\mathcal{F}_\Phi H)(2s-1, s) = 2\sqrt{\pi} \prod_{i=1}^{3} \Gamma\big(s - \tfrac{1}{2} + \tfrac{\alpha_i}{2}\big) \int_{(\eta-1/2)} \widetilde{h}\big(s - s_1 - \tfrac{1}{2}\big) \frac{\pi^{-s_1} \prod_{i=1}^{3} \Gamma\big(\frac{s_1 - \alpha_i}{2}\big) \Gamma\big(\frac{s_1}{2} + 1 - 2s\big)}{\Gamma\big(\frac{1+s_1}{2} + 1 - 2s\big) \Gamma\big(s - \frac{1}{2} + \frac{s_1}{2}\big)} \frac{ds_1}{2\pi i}$$
$$+ \sqrt{\pi} \int_{(\eta-1/2)} \int_{(-1.9)} (\text{same as the integrand of (9-13)}) \frac{du}{2\pi i} \frac{ds_1}{2\pi i}.$$

By Stirling's formula and the same argument following (8-17), the integrals above represent holomorphic functions on $\epsilon < \sigma < 4$. □

9E. ***Step 5: Conclusion.*** Apply Proposition 9.2 to (9-12) and observe that the poles of

$$s \mapsto (\mathcal{F}_\Phi H)(2s-1, s)$$

are exactly the trivial zeros of the arithmetic factor $L(2s-1, \Phi)$ in (9-6). We conclude that the product of functions

$$s \mapsto L(2s-1, \Phi) \cdot (\mathcal{F}_\Phi H)(2s-1, s)$$

is holomorphic on $\epsilon < \sigma < 4$ and thus (9-12) provides a holomorphic continuation of $OD_\Phi(s)$ to the vertical strip $\frac{1}{4} + \frac{\epsilon}{2} < \sigma < \frac{3}{4}$. By the rapid decay of Φ at ∞, the pairing $s \mapsto (P, \mathbb{P}_2^3 \Phi \cdot |\det *|^{\bar{s}-1/2})_{\Gamma_2 \backslash \mathrm{GL}_2(\mathbb{R})}$ represents an entire function. Putting Proposition 5.25, (6-9) and (9-12) together, we arrive at Theorem 1.1.

Remark 9.3. Analytic continuation for moments of degree 6 automorphic L-functions (i.e., our case) is significantly more complicated than those of degree 4, as seen in the second moment formula for GL(2) by Iwaniec–Sarnak and Motohashi. This complication is partly due to the off-diagonal main terms when Φ is an Eisenstein series (see [Kwan 2023]), which are absent in degree 4 cases (see [Conrey et al. 2005, page 35]).

The key distinction lies in the off-diagonal arithmetic. For Iwaniec–Sarnak and Motohashi, the arithmetic is captured by the shifted Dirichlet series of divisor functions, with holomorphy for the dual side depending on the absolute convergence of this series. In our case, the absolute convergence provided by Proposition 7.2 is insufficient. We must carefully move the contour to ensure the L-functions in the off-diagonal evaluate on $\operatorname{Re} s_0 = \frac{1}{2}$ when $s = \frac{1}{2}$.

10. Explication of the off-diagonal — main terms and integral transform

The power of spectral summation formulae (including Theorem 1.1) is encoded in the archimedean transforms. It is important to express these transforms explicitly, often in terms of *special functions*. While the special functions for GL(2) exhibit numerous symmetries and identities, this is less true for higher-rank groups, leaving much to explore.

Nevertheless, there has been success in higher-rank cases. For example, Stade [2001; 2002] computed the Mellin transforms and certain Rankin–Selberg integrals of Whittaker functions for $GL_n(\mathbb{R})$; Goldfeld et al. [Goldfeld and Kontorovich 2013; Goldfeld et al. 2021; 2022] obtained (harmonic-weighted) spherical Weyl laws of $GL_3(\mathbb{R})$, $GL_4(\mathbb{R})$ and $GL_n(\mathbb{R})$ with strong power-saving error terms; Buttcane [2013; 2016] developed the Kuznetsov formulae for GL(3). These works heavily rely on *Mellin-Barnes integrals*, suggesting this approach effectively handles the archimedean aspects of higher-rank problems.

In this final section, we continue such investigation and record several formulae for the archimedean transform $(\mathcal{F}_\Phi H)(s_0, s)$.

Lemma 10.1. *Suppose $H \in \mathcal{C}_\eta$ and $h := H^\flat$. On the vertical strip $-\frac{1}{2} < \operatorname{Re} w < \eta$, we have*

$$\widetilde{h}(w) := \int_0^\infty h(y) y^w \, d^\times y = \frac{\pi^{-w-1/2}}{4} \int_{(0)} H(\mu) \cdot \frac{\Gamma\left(\frac{w+1/2+\mu}{2}\right)\Gamma\left(\frac{w+\frac{1}{2}-\mu}{2}\right)}{|\Gamma(\mu)|^2} \frac{d\mu}{2\pi i}, \tag{10-1}$$

Proof. Since $H \in \mathcal{C}_\eta$, both sides of (10-1) converge absolutely on the strip $-\frac{1}{2} < \operatorname{Re} w < \eta$ by Stirling's formula and Proposition 5.11. Substituting the definition of h as in (5-10) into $\widetilde{h}(w)$, the result follows from (5-2). □

10A. ***The off-diagonal main term in Theorem 1.1.*** In this subsection, we show that the off-diagonal main term of Theorem 1.1 (i.e., $L(2s-1, \Phi) \cdot (\mathcal{F}_\Phi H)(2s-1, s)/2$) aligns with the prediction of [Conrey et al. 2005]. This follows immediately from proving a Mellin–Barnes integral identity, after which the matching follows from the functional equation (5-24).

The proof is more involved than that of Proposition 5.7, as the u-integral (see Section 7) adds intricacies. However, the introduction of new Γ-factors reveals symmetries in the u-integral, leading to several cancellations and reductions.

Theorem 10.2. *Suppose $\frac{1}{2} + \epsilon < \sigma < 1$. Then*

$$(\mathcal{F}_\Phi H)(2s-1, s) = \pi^{1/2-s} \cdot \prod_{i=1}^{3} \frac{\Gamma\left(s - \frac{1}{2} + \frac{\alpha_i}{2}\right)}{\Gamma\left(1 - s - \frac{\alpha_i}{2}\right)} \cdot \int_{(0)} \frac{H(\mu)}{|\Gamma(\mu)|^2} \cdot \prod_{i=1}^{3} \prod_{\pm} \Gamma\left(\frac{1 - s + \alpha_i \pm \mu}{2}\right) \frac{d\mu}{2\pi i}. \tag{10-2}$$

Proof. Suppose $\frac{1}{2}+\epsilon<\sigma<4$. When $s_0=2s-1$, observe that the factor $\Gamma((1-u)/2)$ in the denominator of (9-13) cancels with the factor $\Gamma(s-(s_0+u)/2)$ in the numerator of (9-13). This gives

$$(\mathcal{F}_\Phi H)(2s-1,s)=\sqrt{\pi}\int_{(\eta-1/2)}\widetilde{h}\big(s-s_1-\tfrac{1}{2}\big)\frac{\pi^{-s_1}\prod_{i=1}^{3}\Gamma\big(\frac{s_1-\alpha_i}{2}\big)}{\Gamma\big(\frac{1+s_1}{2}+1-2s\big)} \cdot\int_{(\epsilon)}\frac{\Gamma\big(\frac{u}{2}\big)\Gamma\big(\frac{u+s_1}{2}+1-2s\big)\cdot\prod_{i=1}^{3}\Gamma\big(s-\frac{1}{2}+\frac{\alpha_i-u}{2}\big)}{\Gamma\big(s-\frac{1}{2}+\frac{s_1-u}{2}\big)}\frac{du}{2\pi i}\frac{ds_1}{2\pi i}. \tag{10-3}$$

We make the change of variable $u\to-2u$ and take

$$(a,b,c;d,e)=\big(s-\tfrac{1}{2}+\tfrac{\alpha_1}{2},\,s-\tfrac{1}{2}+\tfrac{\alpha_2}{2},\,s-\tfrac{1}{2}+\tfrac{\alpha_3}{2};\,0,\,\tfrac{s_1}{2}+1-2s\big)$$

in (5-7). Notice that

$$(a+b+c)+d+e=3\big(s-\tfrac{1}{2}\big)+\tfrac{s_1}{2}+1-2s=s-\tfrac{1}{2}+\tfrac{s_1}{2}(:=f)$$

because of $\alpha_1+\alpha_2+\alpha_3=0$. We find the u-integral is equal to

$$2\cdot\prod_{i=1}^{3}\frac{\Gamma\big(s-\frac{1}{2}+\frac{\alpha_i}{2}\big)\Gamma\big(\frac{1}{2}-s+\frac{s_1+\alpha_i}{2}\big)}{\Gamma\big(\frac{s_1-\alpha_i}{2}\big)}. \tag{10-4}$$

Notice that the three Γ-factors in denominator of the last expression cancel with the three in the numerator of the first line of (10-3). Hence, we have

$$(\mathcal{F}_\Phi H)(2s-1,s) =2\sqrt{\pi}\cdot\prod_{i=1}^{3}\Gamma\big(s-\tfrac{1}{2}+\tfrac{\alpha_i}{2}\big)\int_{(\eta-1/2)}\widetilde{h}\big(s-s_1-\tfrac{1}{2}\big)\frac{\pi^{-s_1}\prod_{i=1}^{3}\Gamma\big(\frac{1}{2}-s+\frac{s_1+\alpha_i}{2}\big)}{\Gamma\big(\frac{1+s_1}{2}+1-2s\big)}\frac{ds_1}{2\pi i}. \tag{10-5}$$

We must now further restrict to $\frac{1}{2}+\epsilon<\sigma<1$. We shift the line of integration to the left from $\operatorname{Re}s_1=\eta-\frac{1}{2}$ to $\operatorname{Re}s_1=\sigma_1$ satisfying

$$2\sigma-1<\sigma_1<\sigma.$$

It is easy to see no pole is crossed and we may now apply Lemma 10.1:

$$(\mathcal{F}_\Phi H)(2s-1,s)=\frac{\pi^{1/2-s}}{2}\cdot\prod_{i=1}^{3}\Gamma\big(s-\tfrac{1}{2}+\tfrac{\alpha_i}{2}\big)\cdot\int_{(0)}\frac{H(\mu)}{|\Gamma(\mu)|^2} \cdot\int_{(\sigma_1)}\frac{\prod_{i=1}^{3}\Gamma\big(\frac{1}{2}-s+\frac{s_1+\alpha_i}{2}\big)\cdot\Gamma\big(\frac{s-s_1+\mu}{2}\big)\Gamma\big(\frac{s-s_1-\mu}{2}\big)}{\Gamma\big(\frac{1+s_1}{2}+1-2s\big)}\frac{ds_1}{2\pi i}\frac{d\mu}{2\pi i}. \tag{10-6}$$

For the s_1-integral, applying the change of variable $s_1\to 2s_1$ and (5-7) the second time but with

$$(a,b,c;d,e)=\big(\tfrac{1}{2}-s+\tfrac{\alpha_1}{2},\,\tfrac{1}{2}-s+\tfrac{\alpha_2}{2},\,\tfrac{1}{2}-s+\tfrac{\alpha_3}{2};\,\tfrac{s+\mu}{2},\,\tfrac{s-\mu}{2}\big). \tag{10-7}$$

Observe that

$$(a+b+c)+(d+e)=3\left(\tfrac{1}{2}-s\right)+s:=\tfrac{3}{2}-2s(:=f).$$

The s_1-integral is thus equal to

$$\prod_{i=1}^{3}\frac{\prod_{\pm}\Gamma\left(\frac{1-s+\alpha_i\pm\mu}{2}\right)}{\Gamma\left(1-s-\frac{\alpha_i}{2}\right)}$$

and the result follows. □

10B. *Integral transform.* Based on the experience of Stade [2001; 2002], we *do not* expect the Mellin–Barnes integrals of $(\mathcal{F}_\Phi H)(s_0,s)$ (see (10-12) below) to be completely reducible as in Theorem 10.2 if (s_0,s) is in a *general position*. However, reductions can occur if the integrals take certain special forms, most clearly seen when expressed as *hypergeometric functions*.

We define

$$\begin{aligned}
{}_4\widehat{F}_3&\left(\begin{matrix}A_1 & & A_2 & & A_3 & & A_4\\ & B_1 & & B_2 & & B_3 & \end{matrix}\,\middle|\,z\right)\\
&:=\frac{\Gamma(A_1)\Gamma(A_2)\Gamma(A_3)\Gamma(A_4)}{\Gamma(B_1)\Gamma(B_2)\Gamma(B_3)}\cdot{}_4F_3\left(\begin{matrix}A_1 & & A_2 & & A_3 & & A_4\\ & B_1 & & B_2 & & B_3 & \end{matrix}\,\middle|\,z\right)\\
&:=\sum_{n=0}^{\infty}\frac{\Gamma(A_1+n)\Gamma(A_2+n)\Gamma(A_3+n)\Gamma(A_4+n)}{\Gamma(B_1+n)\Gamma(B_2+n)\Gamma(B_3+n)}\frac{z^n}{n!}.
\end{aligned}\tag{10-8}$$

The series converges absolutely when $|z|<1$ and $A_1,A_2,A_3,A_4\notin\mathbb{Z}_{\le 0}$; and on $|z|=1$ if

$$\operatorname{Re}(B_1+B_2+B_3-A_1-A_2-A_3-A_4)>0.$$

In fact, our hypergeometric functions are of *Saalschütz* type, i.e., $B_1+B_2+B_3-A_1-A_2-A_3-A_4=1$. Only such special type of hypergeometric functions at $z=1$ possess many functional relations and integral representations; see [Mishev 2012].

Proposition 10.3. *Suppose $H\in\mathcal{C}_\eta$ and $h:=H^\flat$. On the region $\sigma_0>\epsilon$, $\sigma<4$, and $2\sigma-\sigma_0-\epsilon>0$, we have $(\mathcal{F}_\Phi H)(s_0,s)$ equal to $2\pi^{3/2}$ times*

$$\begin{aligned}
&\int_{(\eta-1/2)}\widetilde{h}\left(s-s_1-\tfrac{1}{2}\right)\cdot\frac{\prod_{i=1}^{3}\Gamma\left(\frac{s_1-\alpha_i}{2}\right)}{\Gamma\left(\frac{1+s_1}{2}-s_0\right)}\cdot\pi^{-s_1}\sec\tfrac{\pi}{2}(2s+s_0-s_1)\\
&\qquad {}_4\widehat{F}_3\left(\begin{matrix}s-\frac{s_0}{2} & & \frac{s_0+\alpha_1}{2} & & \frac{s_0+\alpha_2}{2} & & \frac{s_0+\alpha_3}{2}\\ & \frac{1}{2} & & \frac{s_0+s_1}{2} & & s+\frac{1}{2}+\frac{s_0-s_1}{2} & \end{matrix}\,\middle|\,1\right)\frac{ds_1}{2\pi i}\\
&-\int_{(\eta-1/2)}\widetilde{h}\left(s-s_1-\tfrac{1}{2}\right)\cdot\frac{\prod_{i=1}^{3}\Gamma\left(\frac{s_1-\alpha_i}{2}\right)}{\Gamma\left(\frac{1+s_1}{2}-s_0\right)}\cdot\pi^{-s_1}\sec\tfrac{\pi}{2}(2s+s_0-s_1)\\
&\qquad {}_4\widehat{F}_3\left(\begin{matrix}\frac{1}{2}-s_0+\frac{s_1}{2} & & \frac{1}{2}-s+\frac{s_1+\alpha_1}{2} & & \frac{1}{2}-s+\frac{s_1+\alpha_2}{2} & & \frac{1}{2}-s+\frac{s_1+\alpha_3}{2}\\ & \frac{1}{2}-s+s_1 & & 1-s-\frac{s_0-s_1}{2} & & \frac{3}{2}-s-\frac{s_0-s_1}{2} & \end{matrix}\,\middle|\,1\right)\frac{ds_1}{2\pi i}.
\end{aligned}\tag{10-9}$$

Proof. By Stirling's formula, we can shift the line of integration of the u-integral in (9-13) to $-\infty$. The residual series obtained can then be identified in terms of hypergeometric series as asserted in the present proposition. This can also be verified by *InverseMellinTransform*[] command in mathematica. More systematically, one rewrites the u-integral in the form of a Meijer's G-function. The conversion between Meijer's G-functions and generalized hypergeometric functions is known as *Slater's theorem*; see Chapter 8 of [Prudnikov et al. 1990]. □

Recently, the articles [Balkanova et al. 2020; Balkanova et al. 2021] have brought in powerful asymptotic analysis of hypergeometric functions to study moments, yielding sharp spectral estimates. Our class of admissible test functions in Theorem 1.1 is broad enough for such prospects, see Remark 5.27.

Next, we establish the existence of a kernel function for the integral transform $(\mathcal{F}_\Phi H)(s_0, s)$ when integrating against a chosen test function $H(\mu)$ on the spectral side. This formula serves as a step toward a more practical result for $(\mathcal{F}_\Phi H)(s_0, s)$. While the proof requires care, it is relatively manageable for our case. However, this is not always true; for example, in the spectral Kuznetsov formulae for GL(2) and GL(3), kernel existence can be more challenging, as noted by [Buttcane 2016; Motohashi 1997].

Proposition 10.4. *Suppose $H \in \mathcal{C}_\eta$. On the domain*

$$\sigma_0 > \epsilon := \tfrac{1}{100}, \quad \sigma < 4, \quad 2\sigma - \sigma_0 - \epsilon > 0, \quad \sigma_0 + 2\sigma - 1 - \epsilon > 0, \quad 1 + \epsilon - \sigma_0 - \sigma > 0, \tag{10-10}$$

we have

$$(\mathcal{F}_\Phi H)(s_0, s) = \frac{\pi^{1/2-s}}{4} \int_{(0)} \frac{H(\mu)}{|\Gamma(\mu)|^2} \cdot \mathcal{K}(s_0, s; \alpha, \mu) \frac{d\mu}{2\pi i}, \tag{10-11}$$

where the kernel function $\mathcal{K}(s_0, s; \alpha, \mu)$ is given explicitly by the double Barnes integrals

$$\begin{aligned}\mathcal{K}(s_0, s; \alpha, \mu) := \int_{-i\infty}^{i\infty} \int_{-i\infty}^{i\infty} & \frac{\Gamma\left(\frac{s-s_1+\mu}{2}\right)\Gamma\left(\frac{s-s_1-\mu}{2}\right)\prod_{i=1}^{3}\Gamma\left(\frac{s_1-\alpha_i}{2}\right)}{\Gamma\left(\frac{1+s_1}{2}-s_0\right)} \\ & \cdot \frac{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{s_1-s_0+u}{2}+\frac{1}{2}-s\right)\prod_{i=1}^{3}\Gamma\left(\frac{s_0-u+\alpha_i}{2}\right)\Gamma\left(s-\frac{s_0+u}{2}\right)}{\Gamma\left(\frac{1-u}{2}\right)\Gamma\left(\frac{s_0-u+s_1}{2}\right)} \frac{du}{2\pi i}\frac{ds_1}{2\pi i},\end{aligned} \tag{10-12}$$

and the contours follow the Barnes convention.

Remark 10.5. (1) The domain (10-10) is certainly nonempty as it includes our point of interest

$$(\sigma_0, \sigma) = \left(\tfrac{1}{2}, \tfrac{1}{2}\right).$$

(2) The contours of (10-12) may be taken explicitly as the vertical lines $\operatorname{Re} u = \epsilon$ and $\operatorname{Re} s_1 = \sigma_1$ with

$$\sigma_0 + 2\sigma - 1 - \epsilon < \sigma_1 < \sigma. \tag{10-13}$$

Proof. Suppose

$$\sigma_0 > \epsilon, \quad \sigma < 4, \quad \text{and} \quad 2\sigma - \sigma_0 - \epsilon > 0 \tag{10-14}$$

as in Proposition 8.1. Recall the expression (9-13) for $(\mathcal{F}_\Phi H)(s_0, s)$. This time, we shift the line of integration of the s_1-integral to $\operatorname{Re} s_1 = \sigma_1$ satisfying

$$\sigma_1 < \sigma \tag{10-15}$$

and no pole is crossed during this shift as long as

$$\sigma_1 > 0 \quad \text{and} \quad \sigma_1 > \sigma_0 + 2\sigma - 1 - \epsilon. \tag{10-16}$$

Now, assume (10-10). The restrictions (10-14), (10-15), (10-16) hold and such a line of integration for the s_1-integral exists. Upon shifting the line of integration to such a position, substituting (10-1) into (9-13) and the result follows. □

The second step is to apply a very useful rearrangement of the Γ-factors in the $(n-1)$-fold Mellin transform of the $GL(n)$ spherical Whittaker function as discovered in [Ishii and Stade 2007]. We shall only need the case of $n = 3$ which we describe as follows. Recall

$$G_\alpha(s_1, s_2) := \pi^{-s_1-s_2} \cdot \frac{\prod_{i=1}^{3} \Gamma\left(\frac{s_1+\alpha_i}{2}\right)\Gamma\left(\frac{s_2-\alpha_i}{2}\right)}{\Gamma\left(\frac{s_1+s_2}{2}\right)} \tag{10-17}$$

from Proposition 5.3. The first Barnes lemma, i.e.,

$$\int_{-i\infty}^{i\infty} \Gamma(w+\alpha)\Gamma(w+\mu)\Gamma(\gamma-w)\Gamma(\delta-w)\frac{dw}{2\pi i} = \frac{\Gamma(\alpha+\gamma)\Gamma(\alpha+\delta)\Gamma(\mu+\gamma)\Gamma(\gamma+\delta)}{\Gamma(\alpha+\mu+\gamma+\delta)}, \tag{10-18}$$

can be applied *in reverse* such that (10-17) can be rewritten as

$$\begin{aligned} G_\alpha(s_1, s_2) = \pi^{-s_1-s_2} \cdot \Gamma\left(\frac{s_1+\alpha_1}{2}\right)\Gamma\left(\frac{s_2-\alpha_1}{2}\right) \\ \cdot \int_{-i\infty}^{i\infty} \Gamma\left(z+\frac{s_1}{2}-\frac{\alpha_1}{4}\right)\Gamma\left(z+\frac{s_2}{2}+\frac{\alpha_1}{4}\right)\Gamma\left(\frac{\alpha_2}{2}+\frac{\alpha_1}{4}-z\right)\Gamma\left(\frac{\alpha_3}{2}+\frac{\alpha_1}{4}-z\right)\frac{dz}{2\pi i}, \end{aligned} \tag{10-19}$$

see Section 2 of [Ishii and Stade 2007]. Although (10-19) is less symmetric than (10-17), it more clearly displays the recursive structure of the $GL(3)$ Whittaker function in terms of the K-Bessel function.

Theorem 10.6. *Suppose* $\operatorname{Re} s_0 = \operatorname{Re} s = \frac{1}{2}$ *and* $\operatorname{Re} \alpha_i = \operatorname{Re} \mu = 0$. *Then* $\mathcal{K}(s_0, s; \alpha, \mu)$ *is equal to*

$$\begin{aligned} 4 \cdot \gamma\left(-\frac{s_0+\alpha_1}{2}\right) \cdot \prod_{\pm} \Gamma\left(\frac{s \pm \mu - \alpha_1}{2}\right) \cdot \int_{-i\infty}^{i\infty}\int_{-i\infty}^{i\infty} \Gamma\left(s+t\right)\Gamma\left(\frac{1-\alpha_1}{2}+t\right) \cdot \Gamma\left(\frac{\alpha_2}{2}+\frac{\alpha_1}{4}-z\right) \\ \cdot \Gamma\left(\frac{\alpha_3}{2}+\frac{\alpha_1}{4}-z\right) \cdot \prod_{\pm} \Gamma\left(\frac{-s \pm \mu}{2}+\frac{\alpha_1}{4}+z-t\right) \cdot \frac{\gamma(t+s_0/2)\gamma(\alpha_1/4-z-s_0/2)}{\gamma(\alpha_1/4+t-z)} \frac{dz}{2\pi i}\frac{dt}{2\pi i}, \end{aligned} \tag{10-20}$$

where the contours may be explicitly taken as the vertical lines $\mathrm{Re}\, t = a$ *and* $\mathrm{Re}\, z = b$ *satisfying*

$$-\tfrac{1}{2} < a < -\tfrac{1}{4}, \quad -\tfrac{1}{4} < b < 0, \quad \textit{and} \quad b - a > \tfrac{1}{4} \tag{10-21}$$

and

$$\gamma(x) := \frac{\Gamma(-x)}{\Gamma(1/2 + x)}. \tag{10-22}$$

Remark 10.7. (1) The assumptions in Theorem 10.6 cover the most interesting cases of Theorem 1.1, particularly on the critical line and for tempered forms, though they are not strictly necessary. These were chosen for a clean description of the contours (10-21).

(2) Furthermore, if either of the following holds:

(a) The cusp form Φ is fixed, allowing implicit constants to depend on $\alpha(\Phi)$.
(b) $\Phi = E^{(3)}_{\min}(*; \alpha)$, where the "shifts" α_i are small as in [Conrey et al. 2005] (i.e., $\ll 1/\log R$, per Remark 5.27).

Then by continuity, it suffices to assume $\alpha_1 = \alpha_2 = \alpha_3 = 0$. With $s = \frac{1}{2}$, this leads to a simpler formula for (10-20):

$$4 \cdot \gamma\left(-\frac{s_0}{2}\right) \prod_{\pm} \Gamma\left(\frac{\frac{1}{2} \pm \mu}{2}\right) \int_{-i\infty}^{i\infty} \int_{-i\infty}^{i\infty} \Gamma\left(\tfrac{1}{2} + t\right)^2 \Gamma(-z)^2 \prod_{\pm} \Gamma\left(\frac{-\frac{1}{2} \pm \mu}{2} + z - t\right) \cdot \frac{\gamma\left(t + \frac{s_0}{2}\right)\gamma\left(-z - \frac{s_0}{2}\right)}{\gamma(t - z)} \frac{dz}{2\pi i} \frac{dt}{2\pi i}.$$

(3) For analytic applications involving Whittaker functions for GL(n), the formula from [Ishii and Stade 2007] has proven more effective than the ones obtained previously. For example:

(a) Buttcane [2020] used the formula (10-19) to significantly simplify the archimedean Rankin–Selberg calculation for GL(3), earlier done in [Stade 1993].
(b) In [Goldfeld et al. 2022], it was crucial for deriving strong bounds for Whittaker functions and their inverse transforms, and the Weyl law.

(This was pointed out to the author by Prof. Eric Stade and Prof. Dorian Goldfeld. The author would like to thank them for their comments here.)

(4) Finally, Stirling's formula shows that the integrand in the Mellin–Barnes representation (10-20) decays exponentially as $|\mathrm{Im}\, z|, |\mathrm{Im}\, t| \to \infty$, independent of $|\mathrm{Im}\, s_0|$. This advantage is not shared by the integrand in (8-1).

Proof of Theorem 10.6. Substitute (10-19) into (10-12) rearrange the integrals, we find that

$$\begin{aligned}&\mathcal{K}(s_0, s; \alpha, \mu)\\ &:= \int_{-i\infty}^{i\infty} \frac{\Gamma\left(\frac{s-s_1+\mu}{2}\right)\Gamma\left(\frac{s-s_1-\mu}{2}\right)\Gamma\left(\frac{s_1-\alpha_1}{2}\right)}{\Gamma\left(\frac{1+s_1}{2}-s_0\right)}\\ &\quad\cdot\int_{-i\infty}^{i\infty} \Gamma\left(\frac{\alpha_2}{2}+\frac{\alpha_1}{4}-z\right)\Gamma\left(\frac{\alpha_3}{2}+\frac{\alpha_1}{4}-z\right)\Gamma\left(z+\frac{s_1}{2}+\frac{\alpha_1}{4}\right)\\ &\quad\cdot\int_{-i\infty}^{i\infty} \frac{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{s_1-(s_0-u)}{2}+\frac{1}{2}-s\right)\Gamma\left(s-\frac{s_0+u}{2}\right)\Gamma\left(\frac{s_0-u+\alpha_1}{2}\right)\Gamma\left(z+\frac{s_0-u}{2}-\frac{\alpha_1}{4}\right)}{\Gamma\left(\frac{1-u}{2}\right)} \frac{du}{2\pi i}\frac{dz}{2\pi i}\frac{ds_1}{2\pi i}.\end{aligned}\tag{10-23}$$

The innermost u-integral, originally of ${}_4F_3(1)$-type (Saalschütz), reduces to a ${}_3F_2(1)$-type (non-Saalschütz), allowing further transformations. We apply the following Barnes integral identity for ${}_3F_2(1)$-type (see [Bailey 1935]):

$$\begin{aligned}&\int_{-i\infty}^{i\infty} \frac{\Gamma(a+u)\Gamma(b+u)\Gamma(c+u)\Gamma(f-u)\Gamma(-u)}{\Gamma(e+u)}\frac{du}{2\pi i}\\ &\quad= \frac{\Gamma(b)\Gamma(c)\Gamma(f+a)}{\Gamma(f+a+b+c-e)\Gamma(e-b)\Gamma(e-c)}\\ &\qquad\qquad\cdot\int_{-i\infty}^{i\infty} \frac{\Gamma(a+t)\Gamma(e-c+t)\Gamma(e-b+t)\Gamma(f+b+c-e-t)\Gamma(-t)}{\Gamma(e+t)}\frac{dt}{2\pi i}.\end{aligned}\tag{10-24}$$

Make a change of variable $u \to -2u$ and take

$$a = s - \tfrac{1}{2}s_0, \quad b = \tfrac{1}{2}(s_0+\alpha_1), \quad c = z + \tfrac{1}{2}s_0 - \tfrac{1}{4}\alpha_1, \quad f = \tfrac{1}{2}(s_1 - s_0) + \tfrac{1}{2} - s, \quad e = \tfrac{1}{2}\tag{10-25}$$

in (10-24), the u-integral of (10-23) can be written as

$$\begin{aligned}&2\cdot\frac{\Gamma\left(\frac{s_0+\alpha_1}{2}\right)\Gamma\left(z+\frac{s_0}{2}-\frac{\alpha_1}{4}\right)\Gamma\left(\frac{1+s_1}{2}-s_0\right)}{\Gamma\left(\frac{s_1}{2}+z+\frac{\alpha_1}{4}\right)\Gamma\left(\frac{1-s_0-\alpha_1}{2}\right)\Gamma\left(\frac{1}{2}-z-\frac{s_0}{2}+\frac{\alpha_1}{4}\right)}\\ &\quad\cdot\int_{-i\infty}^{i\infty} \frac{\Gamma\left(t+s-\frac{s_0}{2}\right)\Gamma\left(t+\frac{1}{2}-z-\frac{s_0}{2}+\frac{\alpha_1}{4}\right)\Gamma\left(t+\frac{1}{2}-\frac{s_0+\alpha_1}{2}\right)\Gamma\left(\frac{s_0+s_1}{2}+z-s+\frac{\alpha_1}{4}-t\right)\Gamma(-t)}{\Gamma\left(\frac{1}{2}+t\right)}\frac{dt}{2\pi i}.\end{aligned}\tag{10-26}$$

Putting this back into (10-23). Observe that two pairs of Γ-factors involving s_1 will be canceled and we can then execute the s_1-integral. More precisely,

$$\begin{aligned}&\tfrac{1}{2}\cdot\mathcal{K}(s_0, s; \alpha, \mu)\\ &= \frac{\Gamma\left(\frac{s_0+\alpha_1}{2}\right)}{\Gamma\left(\frac{1-s_0-\alpha_1}{2}\right)}\cdot\int_{-i\infty}^{i\infty}\frac{dt}{2\pi i}\frac{\Gamma\left(t+s-\frac{s_0}{2}\right)\Gamma\left(t+\frac{1}{2}-\frac{s_0+\alpha_1}{2}\right)\Gamma(-t)}{\Gamma\left(\frac{1}{2}+t\right)}\\ &\quad\cdot\int_{-i\infty}^{i\infty}\frac{dz}{2\pi i}\frac{\Gamma\left(\frac{\alpha_2}{2}+\frac{\alpha_1}{4}-z\right)\Gamma\left(\frac{\alpha_3}{2}+\frac{\alpha_1}{4}-z\right)\Gamma\left(z+\frac{s_0}{2}-\frac{\alpha_1}{4}\right)}{\Gamma\left(\frac{1}{2}-z-\frac{s_0}{2}+\frac{\alpha_1}{4}\right)}\Gamma\left(t+\frac{1}{2}-z-\frac{s_0}{2}+\frac{\alpha_1}{4}\right)\\ &\quad\cdot\int_{-i\infty}^{i\infty}\frac{ds_1}{2\pi i}\Gamma\left(\frac{s_0+s_1}{2}+z-s+\frac{\alpha_1}{4}-t\right)\Gamma\left(\frac{s-s_1+\mu}{2}\right)\Gamma\left(\frac{s-s_1-\mu}{2}\right)\Gamma\left(\frac{s_1-\alpha_1}{2}\right).\end{aligned}\tag{10-27}$$

Applying (10-18) once again, we obtain

$$\begin{aligned}\tfrac{1}{4}\cdot\mathcal{K}(s_0,s;\alpha,\mu) &= \frac{\Gamma\big(\frac{s_0+\alpha_1}{2}\big)}{\Gamma\big(\frac{1-s_0-\alpha_1}{2}\big)}\Gamma\Big(\frac{s+\mu-\alpha_1}{2}\Big)\Gamma\Big(\frac{s-\mu-\alpha_1}{2}\Big)\\ &\quad\cdot\int_{-i\infty}^{i\infty}\frac{dt}{2\pi i}\Gamma\Big(s+t-\frac{s_0}{2}\Big)\Gamma\Big(\frac{1-\alpha_1}{2}+t-\frac{s_0}{2}\Big)\frac{\Gamma(-t)}{\Gamma\big(\frac{1}{2}+t\big)}\\ &\quad\cdot\int_{-i\infty}^{i\infty}\frac{dz}{2\pi i}\Gamma\Big(\frac{\alpha_2}{2}+\frac{\alpha_1}{4}-z\Big)\Gamma\Big(\frac{\alpha_3}{2}+\frac{\alpha_1}{4}-z\Big)\frac{\Gamma\big(z+\frac{s_0}{2}-\frac{\alpha_1}{4}\big)}{\Gamma\big(\frac{1}{2}-z-\frac{s_0}{2}+\frac{\alpha_1}{4}\big)}\\ &\quad\cdot\Gamma\Big(\frac{-s+\mu}{2}+\frac{\alpha_1}{4}+\frac{s_0}{2}+z-t\Big)\Gamma\Big(\frac{-s-\mu}{2}+\frac{\alpha_1}{4}+\frac{s_0}{2}+z-t\Big)\\ &\quad\cdot\frac{\Gamma\big(\frac{1}{2}+\frac{\alpha_1}{4}-\frac{s_0}{2}-z+t\big)}{\Gamma\big(-\frac{\alpha_1}{4}+\frac{s_0}{2}+z-t\big)}.\end{aligned}\tag{10-28}$$

A final cleaning can be performed via the change of variables $t\to t+\frac{s_0}{2}$. This leads to (10-20) and completes the proof. □

11. Notes

Remark 11.1 (Note added in December 2021). The first version of our preprint appeared on Arxiv in December 2021. Peter Humphries has kindly informed the author that the moment of Theorem 1.1 arises naturally from the context of the L^4-norm problem of GL(2) Maass forms and can also be investigated under another set of "Kuznetsov–Voronoi" method (see [Blomer and Khan 2019a; Blomer and Khan 2019b; Blomer et al. 2019]) that is distinct from [Li 2009; 2011]. This is his on-going work with Rizwanur Khan.

Remark 11.2 (Note added in October 2022/April 2023). The preprint of Humphries–Khan has now appeared; see [Humphries and Khan 2022]. The spectral moments considered in [loc. cit.] and the present paper are distinct in a number of ways. In one case, our spectral moments coincide when both $\Phi=\widetilde{\Phi}$ and $s=\frac{1}{2}$ hold true, but otherwise extra twistings by root numbers are present in the one considered by [loc. cit.]. This would then lead to different conclusions in view of the moment conjecture of [Conrey et al. 2005] (see the discussions in Section 3B). In the other case, our spectral moments differ by a full holomorphic spectrum and thus give rise to distinct conclusions in applications toward nonvanishing (say). All these result in different ways of making choices of test functions, as well as different shapes of the dual sides. The self-duality assumption was used in [Humphries and Khan 2022] to annihilate two of the terms in their proof, but no such treatment is necessary for our method.

There is also the recent preprint [Biró 2022] which studies another instance of reciprocity closely related to ours, but with the decomposition "$4=2\times 2$" on the dual side instead. His integral construction consists of a product of an automorphic kernel with a copy of θ-function and Maass cusp form of $\mathrm{SL}_2(\mathbb{Z})$ attached to each variable. The integration is taken over both variables and over the quotient $\Gamma_0(4)\backslash\mathfrak{h}^2$; see equation (3.15) therein.

Acknowledgements

This paper is an extension of the author's thesis [Kwan 2021; 2022]. It is a great pleasure to thank Jack Buttcane, Peter Humphries, Eric Stade, and my Ph.D. advisor Dorian Goldfeld for helpful and interesting discussions, as well as the reviewers for a careful reading of the article and their valuable comments. Part of the work was completed during the author's stay at the American Institute of Mathematics (AIM) in Summer 2021. I would like to thank AIM for the generous hospitality.

References

[Bailey 1935] W. N. Bailey, *Generalized hypergeometric series*, Cambridge Tracts in Math. and Math. Phys. **32**, Cambridge Univ. Press, 1935. Zbl

[Balkanova et al. 2020] O. Balkanova, G. Bhowmik, D. Frolenkov, and N. Raulf, "Mixed moment of GL(2) and GL(3) L-functions", *Proc. Lond. Math. Soc.* (3) **121**:2 (2020), 177–219. MR Zbl

[Balkanova et al. 2021] O. Balkanova, D. Frolenkov, and H. Wu, "On Weyl's subconvex bound for cube-free Hecke characters: totally real case", preprint, 2021. arXiv 2108.12283

[Biró 2022] A. Biró, "Triple product integrals and Rankin–Selberg L-functions", preprint, 2022. arXiv 2209.08913

[Blomer 2012a] V. Blomer, "Period integrals and Rankin–Selberg L-functions on GL(n)", *Geom. Funct. Anal.* **22**:3 (2012), 608–620. MR Zbl

[Blomer 2012b] V. Blomer, "Subconvexity for twisted L-functions on GL(3)", *Amer. J. Math.* **134**:5 (2012), 1385–1421. MR Zbl

[Blomer and Buttcane 2020] V. Blomer and J. Buttcane, "On the subconvexity problem for L-functions on GL(3)", *Ann. Sci. École Norm. Sup.* (4) **53**:6 (2020), 1441–1500. MR Zbl

[Blomer and Khan 2019a] V. Blomer and R. Khan, "Twisted moments of L-functions and spectral reciprocity", *Duke Math. J.* **168**:6 (2019), 1109–1177. MR Zbl

[Blomer and Khan 2019b] V. Blomer and R. Khan, "Uniform subconvexity and symmetry breaking reciprocity", *J. Funct. Anal.* **276**:7 (2019), 2315–2358. MR Zbl

[Blomer et al. 2019] V. Blomer, X. Li, and S. D. Miller, "A spectral reciprocity formula and non-vanishing for L-functions on $\mathrm{GL}(4)\times\mathrm{GL}(2)$", *J. Number Theory* **205** (2019), 1–43. MR Zbl

[Blomer et al. 2020] V. Blomer, P. Humphries, R. Khan, and M. B. Milinovich, "Motohashi's fourth moment identity for non-Archimedean test functions and applications", *Compos. Math.* **156**:5 (2020), 1004–1038. MR Zbl

[Bump 1984] D. Bump, *Automorphic forms on* $\mathrm{GL}(3,\mathbb{R})$, Lecture Notes in Math. **1083**, Springer, 1984. MR Zbl

[Bump 1988] D. Bump, "Barnes' second lemma and its application to Rankin–Selberg convolutions", *Amer. J. Math.* **110**:1 (1988), 179–185. MR Zbl

[Buttcane 2013] J. Buttcane, "On sums of $\mathrm{SL}(3,\mathbb{Z})$ Kloosterman sums", *Ramanujan J.* **32**:3 (2013), 371–419. MR Zbl

[Buttcane 2016] J. Buttcane, "The spectral Kuznetsov formula on SL(3)", *Trans. Amer. Math. Soc.* **368**:9 (2016), 6683–6714. MR Zbl

[Buttcane 2018] J. Buttcane, "Higher weight on GL(3), I: The Eisenstein series", *Forum Math.* **30**:3 (2018), 681–722. MR Zbl

[Buttcane 2020] J. Buttcane, "Kuznetsov, Petersson and Weyl on GL(3), I: The principal series forms", *Amer. J. Math.* **142**:2 (2020), 595–626. MR Zbl

[Chandee and Li 2020] V. Chandee and X. Li, "The second moment of $\mathrm{GL}(4)\times\mathrm{GL}(2)$ L-functions at special points", *Adv. Math.* **365** (2020), art. id. 107060. MR Zbl

[Conrey and Iwaniec 2000] J. B. Conrey and H. Iwaniec, "The cubic moment of central values of automorphic L-functions", *Ann. of Math.* (2) **151**:3 (2000), 1175–1216. MR Zbl

[Conrey et al. 2005] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, "Integral moments of L-functions", *Proc. Lond. Math. Soc.* (3) **91**:1 (2005), 33–104. MR Zbl

[Frolenkov 2020] D. Frolenkov, "The cubic moment of automorphic L-functions in the weight aspect", *J. Number Theory* **207** (2020), 247–281. MR Zbl

[Goldfeld 2015] D. Goldfeld, *Automorphic forms and L-functions for the group* $\mathrm{GL}(n, \mathbb{R})$, Cambridge Stud. Adv. Math. **99**, Cambridge Univ. Press, 2015. MR Zbl

[Goldfeld and Kontorovich 2012] D. Goldfeld and A. Kontorovich, "On the determination of the Plancherel measure for Lebedev–Whittaker transforms on $\mathrm{GL}(n)$", *Acta Arith.* **155**:1 (2012), 15–26. MR Zbl

[Goldfeld and Kontorovich 2013] D. Goldfeld and A. Kontorovich, "On the GL(3) Kuznetsov formula with applications to symmetry types of families of L-functions", pp. 263–310 in *Automorphic representations and L-functions* (Mumbai, 2012), edited by D. Prasad et al., Tata Inst. Fundam. Res. Stud. Math. **22**, Tata Inst. Fund. Res., Mumbai, 2013. MR Zbl

[Goldfeld and Li 2006] D. Goldfeld and X. Li, "Voronoi formulas on $\mathrm{GL}(n)$", *Int. Math. Res. Not.* **2006** (2006), art. id. 86295. MR Zbl

[Goldfeld et al. 2021] D. Goldfeld, E. Stade, and M. Woodbury, "An orthogonality relation for $\mathrm{GL}(4, \mathbb{R})$", *Forum Math. Sigma* **9** (2021), art. id. e47. MR Zbl

[Goldfeld et al. 2022] D. Goldfeld, E. Stade, and M. Woodbury, "An asymptotic orthogonality relation for $\mathrm{GL}(n, \mathbb{R})$", preprint, 2022. arXiv 2212.14534

[Humphries and Khan 2022] P. Humphries and R. Khan, "L^p-norm bounds for automorphic forms via spectral reciprocity", preprint, 2022. arXiv 2208.05613

[Ichino and Templier 2013] A. Ichino and N. Templier, "On the Voronoĭ formula for $\mathrm{GL}(n)$", *Amer. J. Math.* **135**:1 (2013), 65–101. MR Zbl

[Ishii and Stade 2007] T. Ishii and E. Stade, "New formulas for Whittaker functions on $\mathrm{GL}(n, \mathbb{R})$", *J. Funct. Anal.* **244**:1 (2007), 289–314. MR Zbl

[Ivić 2001] A. Ivić, "On sums of Hecke series in short intervals", *J. Théor. Nombres Bordeaux* **13**:2 (2001), 453–468. MR Zbl

[Ivić 2002] A. Ivić, "On the moments of Hecke series at central points", *Funct. Approx. Comment. Math.* **30** (2002), 49–82. MR Zbl

[Iwaniec 2002] H. Iwaniec, *Spectral methods of automorphic forms*, 2nd ed., Grad. Stud. Math. **53**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Jacquet et al. 1979a] H. Jacquet, I. I. Piatetski-Shapiro, and J. Shalika, "Automorphic forms on GL(3), I", *Ann. of Math.* (2) **109**:1 (1979), 169–212. MR Zbl

[Jacquet et al. 1979b] H. Jacquet, I. I. Piatetski-Shapiro, and J. Shalika, "Automorphic forms on GL(3), II", *Ann. of Math.* (2) **109**:2 (1979), 213–258. MR Zbl

[Jana and Nunes 2021] S. Jana and R. Nunes, "Spectral reciprocity for $\mathrm{GL}(n)$ and simultaneous non-vanishing of central L-values", preprint, 2021. arXiv 2111.02297

[Kaneko 2022] I. Kaneko, "Motohashi's formula for the fourth moment of individual Dirichlet L-functions and applications", *Forum Math. Sigma* **10** (2022), art. id. e41. MR Zbl

[Kwan 2021] C.-H. Kwan, "Spectral moment formulae for GL(3) × GL(2) L-functions", preprint, 2021. arXiv 2112.08568v1

[Kwan 2022] C.-H. Kwan, *Spectral moments of Rankin–Selberg L-functions*, Ph.D. thesis, Columbia University, 2022, available at https://www.proquest.com/docview/2656912304.

[Kwan 2023] C.-H. Kwan, "Spectral moment formulae for GL(3) × GL(2) L-functions, II: The Eisenstein case", preprint, 2023. arXiv 2310.09419

[Kwan 2024] C.-H. Kwan, "Spectral moment formulae for GL(3) × GL(2) L-functions, III: Twisted case", *Math. Ann.* (2024).

[Li 2009] X. Li, "The central value of the Rankin–Selberg L-functions", *Geom. Funct. Anal.* **18**:5 (2009), 1660–1695. MR Zbl

[Li 2011] X. Li, "Bounds for GL(3) × GL(2) L-functions and GL(3) L-functions", *Ann. of Math.* (2) **173**:1 (2011), 301–336. MR Zbl

[Liu and Ye 2002] J. Liu and Y. Ye, "Subconvexity for Rankin–Selberg L-functions of Maass forms", *Geom. Funct. Anal.* **12**:6 (2002), 1296–1323. MR Zbl

[Michel and Venkatesh 2006] P. Michel and A. Venkatesh, "Equidistribution, L-functions and ergodic theory: on some problems of Yu. Linnik", pp. 421–457 in *International Congress of Mathematicians*, *II* (Madrid, 2006), edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Michel and Venkatesh 2010] P. Michel and A. Venkatesh, "The subconvexity problem for GL_2", *Publ. Math. Inst. Hautes Études Sci.* **111** (2010), 171–271. MR Zbl

[Miller 2001] S. D. Miller, "On the existence and temperedness of cusp forms for $\mathrm{SL}_3(\mathbb{Z})$", *J. Reine Angew. Math.* **533** (2001), 127–169. MR Zbl

[Miller and Schmid 2006] S. D. Miller and W. Schmid, "Automorphic distributions, L-functions, and Voronoi summation for GL(3)", *Ann. of Math.* (2) **164**:2 (2006), 423–488. MR Zbl

[Mishev 2012] I. D. Mishev, "Coxeter group actions on Saalschützian ${}_4F_3(1)$ series and very-well-poised ${}_7F_6(1)$ series", *J. Math. Anal. Appl.* **385**:2 (2012), 1119–1133. MR Zbl

[Motohashi 1993] Y. Motohashi, "An explicit formula for the fourth power mean of the Riemann zeta-function", *Acta Math.* **170**:2 (1993), 181–220. MR Zbl

[Motohashi 1997] Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Math. **127**, Cambridge Univ. Press, 1997. MR Zbl

[Nelson 2019] P. D. Nelson, "Eisenstein series and the cubic moment for PGL(2)", preprint, 2019. arXiv 1911.06310

[Nunes 2023] R. M. Nunes, "Spectral reciprocity via integral representations", *Algebra Number Theory* **17**:8 (2023), 1381–1409. MR Zbl

[Petrow 2015] I. N. Petrow, "A twisted Motohashi formula and Weyl-subconvexity for L-functions of weight two cusp forms", *Math. Ann.* **363**:1-2 (2015), 175–216. MR Zbl

[Prudnikov et al. 1990] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and series*, *III*: *More special functions*, Gordon & Breach, New York, 1990. MR Zbl

[Stade 1993] E. Stade, "Hypergeometric series and Euler factors at infinity for L-functions on $\mathrm{GL}(3,\mathbb{R})\times\mathrm{GL}(3,\mathbb{R})$", *Amer. J. Math.* **115**:2 (1993), 371–387. MR Zbl

[Stade 2001] E. Stade, "Mellin transforms of $\mathrm{GL}(n,\mathbb{R})$ Whittaker functions", *Amer. J. Math.* **123**:1 (2001), 121–161. MR Zbl

[Stade 2002] E. Stade, "Archimedean L-factors on $\mathrm{GL}(n)\times\mathrm{GL}(n)$ and generalized Barnes integrals", *Israel J. Math.* **127** (2002), 201–219. MR Zbl

[Topacogullari 2021] B. Topacogullari, "The fourth moment of individual Dirichlet L-functions on the critical line", *Math. Z.* **298**:1-2 (2021), 577–624. MR Zbl

[Waldspurger 1981] J.-L. Waldspurger, "Sur les coefficients de Fourier des formes modulaires de poids demi-entier", *J. Math. Pures Appl.* (9) **60**:4 (1981), 375–484. MR Zbl

[Wu 2022] H. Wu, "On Motohashi's formula", *Trans. Amer. Math. Soc.* **375**:11 (2022), 8033–8081. MR Zbl

[Young 2011] M. P. Young, "The fourth moment of Dirichlet L-functions", *Ann. of Math.* (2) **173**:1 (2011), 1–50. MR Zbl

[Young 2017] M. P. Young, "Weyl-type hybrid subconvexity bounds for twisted L-functions and Heegner points on shrinking sets", *J. Eur. Math. Soc.* **19**:5 (2017), 1545–1576. MR Zbl

[Zacharias 2019] R. Zacharias, "Periods and reciprocity, II", preprint, 2019. arXiv 1912.01512

[Zacharias 2021] R. Zacharias, "Periods and reciprocity, I", *Int. Math. Res. Not.* **2021**:3 (2021), 2191–2209. MR Zbl

[Zagier 1981] D. Zagier, "Eisenstein series and the Selberg trace formula, I", pp. 303–355 in *Automorphic forms*, *representation theory and arithmetic* (Bombay, 1979), Tata Inst. Fundam. Res. Stud. Math. **10**, Springer, 1981. MR Zbl

[Zhou 2014] F. Zhou, "Weighted Sato–Tate vertical distribution of the Satake parameter of Maass forms on $\mathrm{PGL}(N)$", *Ramanujan J.* **35**:3 (2014), 405–425. MR Zbl

Communicated by Philippe Michel
Received 2021-12-29 Revised 2023-08-29 Accepted 2023-10-12

c.kwan@ucl.ac.uk *Department of Mathematics, University College London, London, United Kingdom*

ALGEBRA AND NUMBER THEORY 18:10 (2024)
https://doi.org/10.2140/ant.2024.18.1863

The wavefront sets of unipotent supercuspidal representations

Dan Ciubotaru, Lucas Mason-Brown and Emile Okada

We prove that the double (or canonical unramified) wavefront set of an irreducible depth-0 supercuspidal representation of a reductive p-adic group is a singleton provided $p > 3(h-1)$, where h is the Coxeter number. We deduce that the geometric wavefront set is also a singleton in this case, proving a conjecture of Mœglin and Waldspurger. When the group is inner to split and the representation belongs to Lusztig's category of unipotent representations, we give an explicit formula for the double and geometric wavefront sets. As a consequence, we show that the nilpotent part of the Deligne–Langlands–Lusztig parameter of a unipotent supercuspidal representation is precisely the image of its geometric wavefront set under Spaltenstein's duality map.

1. Introduction

Let $\boldsymbol{G}$ be a connected reductive algebraic group defined over a p-adic field k with residue field $\mathbb{F}_q$ and let $\boldsymbol{G}(\mathsf{k})$ be the group of k-rational points. Let (π, X) be an irreducible smooth representation of $\boldsymbol{G}(\mathsf{k})$ with distribution character Θ_X. Let $\mathcal{N}_o^*(\mathsf{k})$ denote the set of nilpotent orbits in $\mathfrak{g}^*(\mathsf{k})$ — the linear dual of the Lie algebra $\mathfrak{g}(\mathsf{k})$ of $\boldsymbol{G}(\mathsf{k})$ — and for each $\mathbb{O}^* \in \mathcal{N}_o^*(\mathsf{k})$ let $\hat{\mu}_{\mathbb{O}^*}$ denote the Fourier transform of the associated orbital integral. In [14], Harish-Chandra proved that there are unique complex numbers $c_{\mathbb{O}^*}(X) \in \mathbb{C}$ such that

$$\Theta_X(\exp(\xi)) = \sum_{\mathbb{O}^* \in \mathcal{N}_o^*(\mathsf{k})} c_{\mathbb{O}^*}(X)\hat{\mu}_{\mathbb{O}^*}(\xi) \tag{1.0.1}$$

MSC2020: 20C33, 22E50.
Keywords: wavefront set, supercuspidal representations, Langlands correspondence, unipotent representations, local character expansion.

for $\xi \in \mathfrak{g}(\mathsf{k})$ a regular element in a small neighborhood of 0. The formula (1.0.1) is called the *local character expansion* of X.

One of the most important invariants which can be extracted from the local character expansion is the *wavefront set* of X. This is the set of nilpotent coadjoint orbits which appear with nonzero multiplicity in the local character expansion

$$\mathrm{WF}(X) = \max\{\mathbb{O}^* \subset \mathfrak{g}^*(\mathsf{k}) \mid c_{\mathbb{O}^*}(X) \neq 0\}.$$

It is common in the literature to consider a slightly coarser invariant called the *geometric wavefront set*. This is the set of maximal nilpotent coadjoint orbits over an *algebraic closure* $\bar{\mathsf{k}}$ of k which meet some orbit in $\mathrm{WF}(X)$. This set is denoted by ${}^{\bar{\mathsf{k}}}\mathrm{WF}(X)$. A longstanding conjecture of Moeglin and Waldspurger [23, Page 429] asserts that ${}^{\bar{\mathsf{k}}}\mathrm{WF}(X)$ is a singleton for all X. Our first main result is that this conjecture is true when X is a depth-0 supercuspidal representation of any reductive group subject only to the assumption that the residue characteristic of k is sufficiently large. In fact, we prove something stronger. Namely we give formulas for two finer invariants in terms of the unrefined minimal K-type of X. The first of these finer invariants is the *unramified wavefront set* ${}^{K}\mathrm{WF}(X)$. This is the set of maximal nilpotent coadjoint orbits over the maximal unramified extension K of k in $\bar{\mathsf{k}}$ which meet some orbit in $\mathrm{WF}(X)$. The second of these finer invariants is the *double wavefront set* $\mathrm{DWF}(X)$ (also known as the *canonical unramified wavefront set* in earlier work of the authors). This invariant is a subset of $\mathcal{N}_o^* \times \mathcal{N}_o^{*\vee}$ where $\mathcal{N}_o^*$ denotes the set of coadjoint nilpotent orbits of the complex reductive group G with the same absolute root datum as $\boldsymbol{G}$, while $\mathcal{N}_o^{*\vee}$ denotes the analogous set for the Langlands dual group $G^\vee$ of $\boldsymbol{G}$. This set comes equipped with a partial order, defined by the formula

$$(\mathbb{O}_1^*, \mathbb{O}_1^{*\vee}) \leq (\mathbb{O}_2^*, \mathbb{O}_2^{*\vee}) \quad \text{if } \mathbb{O}_1^* \subseteq \overline{\mathbb{O}_2^*},\ \mathbb{O}_2^{*\vee} \subseteq \overline{\mathbb{O}_1^{*\vee}}$$

and receives a natural map from $\mathcal{N}_o^*(\mathsf{k})$

$$i_{\mathsf{k}} : \mathcal{N}_o^*(\mathsf{k}) \to \mathcal{N}_o^* \times \mathcal{N}_o^{*\vee};$$

see Section 2.5 below. The double wavefront set is the set

$$\mathrm{DWF}(X) = \max\{i_{\mathsf{k}}(\mathbb{O}^*) \mid c_{\mathbb{O}^*}(X) \neq 0\} \subseteq \mathcal{N}_o^* \times \mathcal{N}_o^{*\vee}.$$

As this is a relatively new invariant we will say a few words here to motivate its study.

Much like the geometric wavefront set we expect that the double wavefront set of any irreducible depth-0 representation (π, X) is a singleton when the residue characteristic of k is sufficiently large. Under such assumptions the double wavefront set takes on a particularly simple form

$$\mathrm{DWF}(X) = ({}^{\bar{\mathsf{k}}}\mathrm{WF}(X), \mathbb{O}^{*\vee}), \tag{1.0.2}$$

where we view the geometric wavefront set as an element of $\mathcal{N}_o^*$ under a suitable identification of $\bar{\mathsf{k}}$-coadjoint orbits with $\mathcal{N}_o^*$. In this setting, it is natural to ask for a description of the nilpotent orbit $\mathbb{O}^{*\vee}$. The results in the second half of this paper and in [9] strongly suggest that when X is unipotent in the sense of [22], $\mathbb{O}^{*\vee}$ should be the open orbit in the singular support of the perverse sheaf corresponding (in the sense of [11]) to (π, X) under the Langlands correspondence. A theorem to this effect would impose a strict relation between the singular support of the character distribution of (π, X) at the identity and the singular support of its parameterizing sheaf. The relation isn't precise (it may not always be the case that $\mathbb{O}^{*\vee}$ determines ${}^{\bar{\mathsf{k}}}\mathrm{WF}(X)$ as demonstrated in [9, Example 1.4.2]) but the image of i_{k} in $\mathcal{N}_o \times \mathcal{N}_o^\vee$ is small in some sense, and so the possible pairs $(\mathbb{O}, \mathbb{O}^\vee)$ which can arise as double wavefront sets is still very limited.

In the second half of this paper, we specialize our attention to inner to split reductive groups and supercuspidal representations with L-parameters which are trivial on the inertia subgroup of the Weil group. Such representations belong to Lusztig's category of *unipotent representations*, defined in [22, Section 0.3] (see also Definition 4.4.1 below). Irreducible unipotent representations of $\boldsymbol{G}(\mathsf{k})$ are parameterized by so-called Deligne–Langlands–Lusztig (or enhanced Langlands) parameters. The Deligne–Langlands–Lusztig parameter of an irreducible unipotent $\boldsymbol{G}(\mathsf{k})$-representation X is a $G^\vee$-orbit of triples (τ, n, ρ), where τ is a semisimple element in the Langlands dual group $G^\vee$ of $\boldsymbol{G}$, n is a nilpotent element of the Lie algebra $\mathfrak{g}^\vee$ such that $\mathrm{Ad}(\tau)n = qn$, and ρ is an irreducible representation of a certain finite group A^1_φ; see Section 4.5. When $\boldsymbol{G}$ is adjoint, A^1_φ is the component group of the centralizer of τ and n in $G^\vee$.

It is natural to ask how the local character expansion of X is related to its Deligne–Langlands–Lusztig parameter (τ, n, ρ). At one extreme, we have the coefficient $c_0(X)$ of the zero orbit in the local character expansion. It is known that when X is tempered, $c_0(X) \neq 0$ if and only if X is square integrable, and in this case, $c_0(X)$ equals, up to a sign, the ratio between the formal degrees of X and the Steinberg representation. An interpretation of the formal degree in terms of the Langlands parameter was conjectured first by Reeder [32] for unipotent representations and then vastly generalized by Hiraga, Ichino, and Ikeda [15] for all discrete series representations of a semisimple group over a local field. For unipotent representations, this interpretation was verified in the case of split exceptional groups by Reeder [32] and in the remaining cases by Opdam [28]; see also [13].

At the other extreme, we have the various wavefront sets. Our second main result relates the unramified, geometric, and double wavefront sets to the nilpotent element n in the Deligne–Langland–Lusztig parameter (s, n, τ) of X. More precisely, let $\mathbb{O}^\vee_X$ denote the nilpotent $G^\vee$-orbit of n. We prove in Theorem 5.0.2 that

$$d_S({}^K\mathrm{WF}(X)) = \mathbb{O}^\vee_X, \quad \mathrm{DWF}(X) = d_A(\mathbb{O}^\vee_X, 1) = (d(\mathbb{O}^\vee_X), \mathbb{O}^\vee_X), \quad {}^{\bar{\mathsf{k}}}\mathrm{WF}(X) = d(\mathbb{O}^\vee_X). \tag{1.0.3}$$

Here we have identified coadjoint nilpotent orbits with adjoint nilpotent orbits using the Killing form, and d, d_S and d_A are the duality maps defined by Spaltenstein, Sommers and Achar, respectively; see Section 2.1. A notable consequence of this result is that the unramified and double wavefront sets (and in fact also the geometric wavefront set since $\mathbb{O}^\vee_X$ is always special for the unipotent supercuspidal representations) determine the nilpotent part of the Deligne–Langlands–Lusztig parameter. We emphasize that the simplicity of the formulas (1.0.3) is due to the fact that X is supercuspidal and therefore equal to its Aubert–Zelevinsky dual [2]. In general, one expects that the wavefront set of X is related by duality to the nilpotent parameter associated to the *AZ dual* of X; see [10]. This expression for the wavefront set is reminiscent of Lusztig's formula for the Kawanaka wavefront set of an irreducible unipotent representation of a finite reductive group [21, Theorem 11.2]. In fact, the finite reductive group results from loc. cit. play an important role in the construction and analysis of test functions in the local character expansion [5; 27]. For a positive-depth analogue of the depth-0 Barbasch–Moy methods, see [8].

1.1. ***Structure of paper.*** In Section 2, we collect some preliminaries on nilpotent orbits, wavefront sets, and depth-0 representations. In Section 3, we prove our main result on depth-0 supercuspidal representations (this result is proved without restrictions on the group). In Section 4, we collect some additional preliminaries on nilpotent orbits, wavefront sets, and Bruhat–Tits theory, necessary for our study of unipotent representations. In Section 5, we state our main result on unipotent supercuspidal representations (this result requires that $\boldsymbol{G}$ is inner to split). The proof of this result, which is completed in Section 8, requires some explicit tabulation of unipotent cuspidal representations of finite groups of Lie type (appearing in Section 6), and unipotent supercuspidal representations of simple adjoint groups (appearing in Section 7).

2. Preliminaries, I

Let k be a nonarchimedean local field of characteristic 0 with residue field $\mathbb{F}_q$ of sufficiently large characteristic, ring of integers $\mathfrak{o} \subset \mathsf{k}$, and valuation $\mathsf{val}_{\mathsf{k}}$. Fix an algebraic closure $\bar{\mathsf{k}}$ of k with Galois group Γ, and let $K \subset \bar{\mathsf{k}}$ be the maximal unramified extension of k in $\bar{\mathsf{k}}$. Let E be a minimal Galois extension of K in $\bar{\mathsf{k}}$ such that $\boldsymbol{G}$ is E-split. Let $\mathfrak{O}$ be the ring of integers of K. Let Frob be the geometric Frobenius element of $\mathrm{Gal}(K/\mathsf{k})$, the topological generator which induces the inverse of the automorphism $x \to x^q$ of $\mathbb{F}_q$. Let $\boldsymbol{G}$ be a connected reductive algebraic group defined over k and write $\boldsymbol{G}(\mathsf{k})$ for the group of k-rational points of $\boldsymbol{G}$. Let G be the complex reductive group defined over $\mathbb{C}$ with the same absolute root datum as $\boldsymbol{G}$, and let $G^\vee$ denote the Langlands dual group. Let $\mathcal{C}(\boldsymbol{G}(\mathsf{k}))$ be the category of smooth complex $\boldsymbol{G}(\mathsf{k})$-representations and let $\Pi(\boldsymbol{G}(\mathsf{k})) \subset \mathcal{C}(\boldsymbol{G}(\mathsf{k}))$ be the set of irreducible objects. Let $R(\boldsymbol{G}(\mathsf{k}))$ denote the Grothendieck group of $\mathcal{C}(\boldsymbol{G}(\mathsf{k}))$. Let $\mathcal{B}(\boldsymbol{G}, \mathsf{k})$ denote the extended Bruhat–Tits building for $\boldsymbol{G}$ over k and $\mathcal{A}(\boldsymbol{T}, \mathsf{k})$ the apartment of $\mathcal{B}(\boldsymbol{G}, \mathsf{k})$ corresponding to a maximal k-split torus $\boldsymbol{T}$ of $\boldsymbol{G}$. For a subset $c \subseteq \mathcal{B}(\boldsymbol{G}, \mathsf{k})$ write $c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})$ to indicate that it is a face of the building. For a face $c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})$ there is a group $\boldsymbol{P}_c^\dagger$ defined over $\mathfrak{o}$ such that $\boldsymbol{P}_c^\dagger(\mathfrak{o})$ identifies with the stabilizer of c in $\boldsymbol{G}(\mathsf{k})$. There is an exact sequence

$$1 \to \boldsymbol{U}_c(\mathfrak{o}) \to \boldsymbol{P}_c^\dagger(\mathfrak{o}) \to \boldsymbol{L}_c^\dagger(\mathbb{F}_q) \to 1, \tag{2.0.1}$$

where $\boldsymbol{U}_c(\mathfrak{o})$ is the pro-unipotent radical of $\boldsymbol{P}_c^\dagger(\mathfrak{o})$ and $\boldsymbol{L}_c^\dagger$ is the reductive quotient of the special fiber of $\boldsymbol{P}_c^\dagger$. Let $\boldsymbol{L}_c$ denote the identity component of $\boldsymbol{L}_c^\dagger$, and let $\boldsymbol{P}_c$ be the subgroup of $\boldsymbol{P}_c^\dagger$ defined over $\mathfrak{o}$ such that $\boldsymbol{P}_c(\mathfrak{o})$ is the inverse image of $\boldsymbol{L}_c(\mathbb{F}_q)$ in $\boldsymbol{P}_c^\dagger(\mathfrak{o})$. The groups $\boldsymbol{P}_c(\mathfrak{o})$ are the so-called parahoric subgroups. When c is a chamber, the group $\boldsymbol{P}_c(\mathfrak{o})$ is called an Iwahori subgroup.

2.1. ***Nilpotent orbits.*** For rest of the paper we identify coadjoint nilpotent orbits with adjoint nilpotent orbits using the Killing form. Let $\mathcal{N}$ be the functor which takes a field F to the set of nilpotent elements in $\mathfrak{g}(F)$, and let $\mathcal{N}_o$ be the functor which takes F to the set of adjoint $\boldsymbol{G}(F)$-orbits on $\mathcal{N}(F)$. When F is k or K, we view $\mathcal{N}_o(F)$ as a partially ordered set with respect to the closure ordering in the topology induced by the topology on F. When F is algebraically closed, we view $\mathcal{N}_o(F)$ as a partially ordered set with respect to the closure ordering in the Zariski topology. For brevity we will write $\mathcal{N}(F'/F)$ (resp. $\mathcal{N}_o(F'/F)$) for $\mathcal{N}(F \to F')$ (resp. $\mathcal{N}_o(F \to F')$) where $F \to F'$ is a morphism of fields. For $(F, F') = (\mathsf{k}, K)$ (resp. $(\mathsf{k}, \bar{\mathsf{k}})$, $(K, \bar{\mathsf{k}})$), the map $\mathcal{N}_o(F'/F)$ is strictly increasing (resp. strictly increasing, nondecreasing). When we wish to emphasize the group we are working with we include it as a superscript, e.g., $\mathcal{N}_o^{\boldsymbol{G}}$. Define

$$\mathcal{I}_o = \{(c, \mathbb{O}) \mid c \subseteq_f \mathcal{B}(\boldsymbol{G}),\ \mathbb{O} \in \mathcal{N}_o^{\boldsymbol{L}_c}(\bar{\mathbb{F}}_q)\}. \tag{2.1.1}$$

There is a partial order on $\mathcal{I}_o$, defined by

$$(c_1, \mathbb{O}_1) \le (c_2, \mathbb{O}_2) \iff c_1 = c_2 \text{ and } \mathbb{O}_1 \le \mathbb{O}_2.$$

By [27, Section 1.1.2] there is a strictly increasing surjective map

$$\mathcal{L} : (\mathcal{I}_o, \le) \to (\mathcal{N}_o(K), \le). \tag{2.1.2}$$

For a face $c \subseteq_f \mathcal{B}(\boldsymbol{G}, F)$ let $\mathcal{L}_c : \mathcal{N}_o^{\boldsymbol{L}_c}(\bar{F}) \to \mathcal{N}_o^{\boldsymbol{G}}(\mathsf{k})$ denote the map $\mathbb{O} \mapsto \mathcal{L}(c, \mathbb{O})$.

Finally recall the following classical result on nilpotent orbits.

Lemma 2.1.1 [29, Theorem 1.5; 30, Corollary 3.5]. *Let $\boldsymbol{G}$ a connected reductive group defined over a field F with good characteristic for $\boldsymbol{G}$. For any algebraically closed field extension F' of F there is canonical isomorphism of partially ordered sets $\Lambda^{F'} : \mathcal{N}_o^{\boldsymbol{G}}(F') \xrightarrow{\sim} \mathcal{N}_o$.*

2.2. *Dualities on nilpotent orbits.* Recall the groups $G, G^\vee$ from Section 2 and the functors $\mathcal{N}, \mathcal{N}_0$ from Section 2.1. We will extend the functors $\mathcal{N}, \mathcal{N}_o$ to include the field $\mathbb{C}$: Define $\mathcal{N}(\mathbb{C})$ to be the nilpotent elements of the Lie algebra of G, and $\mathcal{N}_o(\mathbb{C})$ to be the G-orbits on $\mathcal{N}$. Since it won't ever cause confusion we will often omit the $\mathbb{C}$ and simply write $\mathcal{N}, \mathcal{N}_o$ instead of $\mathcal{N}(\mathbb{C}), \mathcal{N}_o(\mathbb{C})$. Define $\mathcal{N}_{o,c}$ (resp. $\mathcal{N}_{o,\bar{c}}$) to be the set of all pairs $(\mathbb{O}, C)$ such that $\mathbb{O} \in \mathcal{N}_o$ and C is a conjugacy class in the fundamental group $A(\mathbb{O})$ of $\mathbb{O}$ (resp. Lusztig's canonical quotient $\bar{A}(\mathbb{O})$ of $A(\mathbb{O})$; see [33, Section 5]). There is a natural map

$$\mathfrak{Q} : \mathcal{N}_{o,c} \to \mathcal{N}_{o,\bar{c}}, \quad (\mathbb{O}, C) \mapsto (\mathbb{O}, \bar{C}), \tag{2.2.1}$$

where $\bar{C}$ is the image of C in $\bar{A}(\mathbb{O})$ under the natural homomorphism $A(\mathbb{O}) \twoheadrightarrow \bar{A}(\mathbb{O})$. There are also projection maps $\mathrm{pr}_1 : \mathcal{N}_{o,c} \to \mathcal{N}_o$, $\mathrm{pr}_1 : \mathcal{N}_{o,\bar{c}} \to \mathcal{N}_o$. We will typically write $\mathcal{N}^\vee, \mathcal{N}_o^\vee, \mathcal{N}_{o,c}^\vee$, and $\mathcal{N}_{o,\bar{c}}^\vee$ for the sets $\mathcal{N}, \mathcal{N}_o, \mathcal{N}_{o,c}$, and $\mathcal{N}_{o,\bar{c}}$ associated to the Langlands dual group $G^\vee$.

Let

$$d : \mathcal{N}_0 \to \mathcal{N}_0^\vee, \quad d : \mathcal{N}_0^\vee \to \mathcal{N}_0 \tag{2.2.2}$$

be the duality maps defined by Spaltenstein [34, Proposition 10.3] (see also Lusztig [20, §13.3] and Barbasch and Vogan [6, Appendix A]). Let

$$d_S : \mathcal{N}_{o,c} \twoheadrightarrow \mathcal{N}_o^\vee, \quad d_S : \mathcal{N}_{o,c}^\vee \twoheadrightarrow \mathcal{N}_o \tag{2.2.3}$$

be the duality maps defined by Sommers [33, Section 6] and

$$d_A : \mathcal{N}_{o,\bar{c}} \to \mathcal{N}^\vee_{o,\bar{c}}, \quad d_A : \mathcal{N}^\vee_{o,\bar{c}} \to \mathcal{N}_{o,\bar{c}} \tag{2.2.4}$$

be the duality maps defined by Achar [1, Section 1]. These duality maps are compatible in the following sense. For $\mathbb{O} \in \mathcal{N}_o$,

$$d_S(\mathbb{O}, 1) = d(\mathbb{O})$$

and, for $\xi \in \mathcal{N}_{o,c}$,

$$d_A(\mathfrak{Q}(\xi)) = (d_S(\xi), \bar{C}')$$

for some $\bar{C}'$. In particular

$$d_A(\mathbb{O}, 1) = (d(\mathbb{O}), \bar{C}) \tag{2.2.5}$$

for some $\bar{C}$.

There is a natural map

$$\iota : \mathcal{N}_{o,c} \to \mathcal{N}_o \times \mathcal{N}^\vee_o, \quad \xi \mapsto (\mathrm{pr}_1(\xi), d_S(\xi)).$$

The set $\mathcal{N}_o \times \mathcal{N}^\vee_o$ is equipped with a natural partial order

$$(\mathbb{O}_1, \mathbb{O}^\vee_1) \le (\mathbb{O}_2, \mathbb{O}^\vee_2) \iff \mathbb{O}_1 \le \mathbb{O}_2,\ \mathbb{O}^\vee_2 \le \mathbb{O}^\vee_1.$$

This partial order pulls back to a preorder on $\mathcal{N}_{o,c}$ via ι which coincides with the preorder defined in [1, Introduction]. For $\xi, \xi' \in \mathcal{N}_{o,c}$ define $\xi \sim_A \xi'$ if $\iota(\xi) = \iota(\xi')$. Write $[\xi]$ for the equivalence class of $\xi \in \mathcal{N}_{o,c}$. By [1, Theorem 1], the $\sim_A$-equivalence classes in $\mathcal{N}_{o,c}$ coincide precisely with the fibers of the projection map $\mathfrak{Q} : \mathcal{N}_{o,c} \to \mathcal{N}_{o,\bar{c}}$ and so ι descends to an injection

$$\bar{\iota} : \mathcal{N}_{o,\bar{c}} \to \mathcal{N}_o \times \mathcal{N}^\vee_o.$$

In particular $\le_A$ descends to a partial order on $\mathcal{N}_{o,\bar{c}}$ which we also call $\le_A$. The maps d, d_S, d_A are all order reversing with respect to the relevant pre/partial orders.

Achar duality admits a particularly simple interpretation with this setup. We can view $\mathcal{N}_{o,\bar{c}}$ as a subset of $\mathcal{N}_o \times \mathcal{N}^\vee_o$ via $\bar{\iota}$ and we can similarly view $\mathcal{N}^\vee_{o,\bar{c}}$ as a subset of $\mathcal{N}^\vee_o \times \mathcal{N}_o$. Define

$$D : \mathcal{N}_o \times \mathcal{N}^\vee_o \to \mathcal{N}^\vee_o \times \mathcal{N}_o, \quad (\mathbb{O}, \mathbb{O}^\vee) \mapsto (\mathbb{O}^\vee, \mathbb{O}).$$

Then for the so-called special elements $\bar{\xi} \in \mathcal{N}_{o,\bar{c}}$,

$$\bar{\iota}^\vee \circ d_A(\bar{\xi}) = D \circ \bar{\iota}(\bar{\xi}).$$

In particular the elements of the form $(\mathbb{O}, 1) \in \mathcal{N}_{o,\bar{c}}$ are all special and

$$\bar{\iota}(\mathbb{O}, 1) = (\mathbb{O}, d_{\mathrm{BV}}(\mathbb{O})), \quad \bar{\iota}^\vee(d_A(\mathbb{O}, 1)) = (d_{\mathrm{BV}}(\mathbb{O}), \mathbb{O}). \tag{2.2.6}$$

2.3. ***Nilpotent orbits over E.*** Recall from Section 2 that E denotes a minimal Galois extension of K such that $\boldsymbol{G}$ is E-split. In [27, Section 2] the third author establishes a number of results about the structure of $\mathcal{N}_o^{\boldsymbol{G}}(E)$ which we now briefly summarize.

Let $\boldsymbol{T}$ be a maximal E-split torus of $\boldsymbol{G}$ and let x_0 be a special point in $\mathcal{A}(\boldsymbol{T}, E)$. In [27, Section 2.1.5], we construct a bijection

$$\theta_{x_0,\boldsymbol{T}} : \mathcal{N}_o^{\boldsymbol{G}}(E) \xrightarrow{\sim} \mathcal{N}_{o,c}.$$

Theorem 2.3.1 [27, Theorem 2.20, Theorem 2.27, Proposition 2.29]. *The bijection*

$$\theta_{x_0,\boldsymbol{T}} : \mathcal{N}_o^{\boldsymbol{G}}(E) \xrightarrow{\sim} \mathcal{N}_{o,c}$$

is natural in $\boldsymbol{T}$, *equivariant in* x_0, *and makes the following diagram commute*:

$$\begin{tikzcd} \mathcal{N}_o^{\boldsymbol{G}}(E) \arrow[r, "\theta_{x_0,\boldsymbol{T}}"] \arrow[d, "\mathcal{N}_o(\bar{\mathsf{k}}/E)"'] & \mathcal{N}_{o,c} \arrow[d, "\mathrm{pr}_1"] \\ \mathcal{N}_o(\bar{\mathsf{k}}) \arrow[r, "\Lambda^{\bar{\mathsf{k}}}"] & \mathcal{N}_o \end{tikzcd} \tag{2.3.1}$$

The composition

$$d_{S,\boldsymbol{T}} := d_S \circ \theta_{x_0,\boldsymbol{T}}$$

is independent of the choice of x_0 and natural in $\boldsymbol{T}$ [27, Proposition 2.32] and so we get a map

$$i : \mathcal{N}_o^{\boldsymbol{G}}(E) \to \mathcal{N}_o \times \mathcal{N}_o^{\vee}, \quad \mathbb{O} \mapsto (\Lambda^{\bar{\mathsf{k}}} \circ \mathcal{N}_o(\bar{\mathsf{k}}/E)(\mathbb{O}), d_{S,\boldsymbol{T}}(\mathbb{O})),$$

which only depends naturally on the torus $\boldsymbol{T}$. Let $\le_A$ denote the preorder obtained by pulling back $\le$ along i. So in particular for $\mathbb{O}_1, \mathbb{O}_2 \in \mathcal{N}_o(E)$ we have $\mathbb{O}_1 \le_A \mathbb{O}_2$ if

$$\mathcal{N}_o(\bar{\mathsf{k}}/E)(\mathbb{O}_1) \le \mathcal{N}_o(\bar{\mathsf{k}}/E)(\mathbb{O}_2) \quad \text{and} \quad d_{S,\boldsymbol{T}}(\mathbb{O}_1) \ge d_{S,\boldsymbol{T}_2}(\mathbb{O}_2)$$

and let $\sim_A$ denote the equivalence classes of this preorder. By naturality of $d_{S,\boldsymbol{T}}$, this preorder is independent of the choice of $\boldsymbol{T}$ and the map

$$\theta_{x_0,\boldsymbol{T}} : (\mathcal{N}_o(E), \le_A) \to (\mathcal{N}_{o,c}, \le_A)$$

is an isomorphism of preorders.

2.4. ***Wavefront sets.*** Let (π, X) be an admissible smooth representation of $\boldsymbol{G}(\mathsf{k})$ and let Θ_X be the character of X. Recall that for each nilpotent orbit $\mathbb{O} \in \mathcal{N}_o(\mathsf{k})$ there is an associated distribution $\mu_{\mathbb{O}}$ on $C_c^\infty(\mathfrak{g}^\omega(\mathsf{k}))$ called the *nilpotent orbital integral* of $\mathbb{O}$ [31]. Write $\hat{\mu}_{\mathbb{O}}$ for the Fourier transform of this distribution. Generalizing a result of Howe [16], Harish-Chandra [14] showed that there are complex numbers $c_{\mathbb{O}}(X) \in \mathbb{C}$ such that

$$\Theta_X(\exp(\xi)) = \sum_{\mathbb{O}} c_{\mathbb{O}}(X)\hat{\mu}_{\mathbb{O}}(\xi) \tag{2.4.1}$$

for $\xi \in \mathfrak{g}^{\omega}(\mathsf{k})$ a regular element in a small neighborhood of 0. The formula (2.4.1) is called the *local character expansion* of π. The (*p-adic*) *wavefront set* of X is

$$\mathrm{WF}(X) := \max\{\mathbb{O} \mid c_{\mathbb{O}}(X) \neq 0\} \subseteq \mathcal{N}_o(\mathsf{k}).$$

The *unramified wavefront set* of X is

$$^{K}\mathrm{WF}(X) := \max\{\mathcal{N}_o(K/\mathsf{k})(\mathbb{O}) \mid c_{\mathbb{O}}(X) \neq 0\} \subseteq \mathcal{N}_o(K).$$

The *geometric wavefront set* of X is

$$^{\bar{\mathsf{k}}}\mathrm{WF}(X) := \max\{\mathcal{N}_o(\bar{\mathsf{k}}/\mathsf{k})(\mathbb{O}) \mid c_{\mathbb{O}}(X) \neq 0\} \subseteq \mathcal{N}_o(\bar{\mathsf{k}}).$$

Using the map $\Lambda^{\bar{\mathsf{k}}}$ from Lemma 2.1.1 we will interchangeably view the geometric wavefront set as living in $\mathcal{N}_o(\bar{\mathsf{k}})$ and $\mathcal{N}_o$. The way we will compute the unramified wavefront set is using the tools developed in [27] based on ideas of Barbasch and Moy [5].

Definition 2.4.1. For every face $c \subseteq_f \mathcal{B}(\boldsymbol{G})$, the space of invariants $X^{\boldsymbol{U}_c(\mathfrak{o})}$ is a (finite-dimensional) $\boldsymbol{L}_c(\mathbb{F}_q)$-representation. Let $\mathrm{WF}(X^{\boldsymbol{U}_c(\mathfrak{o})}) \subseteq \mathcal{N}_o^{\boldsymbol{L}_c}(\bar{\mathbb{F}}_q)$ denote the *Kawanaka wavefront set* [18] of $X^{\boldsymbol{U}_c(\mathfrak{o})}$, and define *the local unramified wavefront set of X at c* to be

$$^{K}\mathrm{WF}_c(X) := \{\mathcal{L}_c(\mathbb{O}) \mid \mathbb{O} \in \mathrm{WF}(X^{\boldsymbol{U}_c(\mathfrak{o})})\} \subseteq \mathcal{N}_o(K). \tag{2.4.2}$$

Theorem 2.4.2 [27, Theorem 0.1]. *Let (π, X) be a depth-0 representation of $\boldsymbol{G}(\mathsf{k})$ and $\mathcal{C}$ be a collection of faces of $\mathcal{B}(\boldsymbol{G}, \mathsf{k})$ such that every nilpotent orbit lies in the image of $\mathcal{L}_c$ for some $c \in \mathcal{C}$. Then*

$$^{K}\mathrm{WF}(X) := \max\{^{K}\mathrm{WF}_c(X) \mid c \in \mathcal{C}\} \subseteq \mathcal{N}_o(K). \tag{2.4.3}$$

2.5. *The double wavefront set.* In [27, Section 2.2.3] the third author introduced a refinement of the geometric wavefront set called the *double wavefront set*.

Remark 2.5.1. (1) The double wavefront set is called the *canonical unramified wavefront set* in [27], but this name is not very descriptive and double wavefront set seems more appropriate due to its cosmetic similarities to the double affine Hecke algebra.

(2) The double wavefront set is only defined for unramified groups in [27], but the definition given here makes sense for any reductive group defined over a nonarchimedean local field with good residue characteristic. The proof of Theorem 2.5.2 in this generality is essentially the same as for unramified groups.

The double wavefront set is defined as follows. Recall the map i and partial order $\leq$ on $\mathcal{N}_o \times \mathcal{N}_o^{\vee}$ from Section 2.3, and the coefficients $c_{\mathbb{O}}(X)$ of the local character expansion from Section 2.4. Let

$$i_{\mathsf{k}} : \mathcal{N}_o(\mathsf{k}) \to \mathcal{N}_o \times \mathcal{N}_o^{\vee}$$

denote the composition $i \circ \mathcal{N}_o(E/\mathsf{k})$. The double wavefront set of a smooth admissible representation (π, X) of $\boldsymbol{G}(\mathsf{k})$ is defined to be

$$\mathrm{DWF}(X) := \max\{i_{\mathsf{k}}(\mathbb{O}) : \mathbb{O} \in \mathcal{N}_o(\mathsf{k}),\ c_{\mathbb{O}}(X) \neq 0\} \subseteq \mathcal{N}_o \times \mathcal{N}_o^\vee.$$

Let

$$\mathrm{pr}_{\mathcal{N}} : \mathcal{N}_o \times \mathcal{N}_o^\vee \to \mathcal{N}_o, \quad \mathrm{pr}_{\mathcal{N}^\vee} : \mathcal{N}_o \times \mathcal{N}_o^\vee \to \mathcal{N}_o^\vee$$

denote the projection maps. It is clear from the definition that the geometric wavefront set is simply the maximal orbit lying in $\mathrm{pr}_{\mathcal{N}} \circ \mathrm{DWF}(X)$. In particular when the double wavefront set is a singleton, then so is the geometric wavefront set, and the double wavefront set must be of the form

$$\mathrm{DWF}(X) = ({}^{\bar{\mathsf{k}}}\mathrm{WF}(X), \mathbb{O}^\vee) \tag{2.5.1}$$

for some orbit $\mathbb{O}^\vee \in \mathcal{N}_o^\vee$.

Let i_K be the composition $i \circ \mathcal{N}_o(E/K)$. Much like the unramified wavefront set, the double wavefront set can be computed building-locally.

Theorem 2.5.2 [27, Lemma 2.36]. *Let (π, X) be a depth-0 representation of $\boldsymbol{G}(\mathsf{k})$ and $\mathcal{C}$ be a collection of faces of $\mathcal{B}(\boldsymbol{G}, \mathsf{k})$ such that every nilpotent orbit lies in the image of $\mathcal{L}_c$ for some $c \in \mathcal{C}$. Then*

$$\mathrm{DWF}(X) := \max\{i_K({}^K\mathrm{WF}_c(X)) \mid c \in \mathcal{C}\} \subseteq \mathcal{N}_o \times \mathcal{N}_o^\vee. \tag{2.5.2}$$

2.6. ***Depth-0 representations.*** Let (π, X) be a smooth irreducible representation of $\boldsymbol{G}(\mathsf{k})$. We say that X is depth-0 if $X^{\boldsymbol{U}_c(\mathfrak{o})} \neq 0$ for some face $c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})$. Write $\Pi^0(\boldsymbol{G}(\mathsf{k}))$ for the subset of $\Pi(\boldsymbol{G}(\mathsf{k}))$ consisting of depth-0 representations. Let

$$S(\boldsymbol{G}) := \{(c, \sigma) : c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k}),\ \sigma \text{ a cuspidal representation of } \boldsymbol{L}_c(\mathbb{F}_q)\}$$

and for $(c_1, \sigma_1), (c_2, \sigma_2) \in S(\boldsymbol{G})$ write $(c_1, \sigma_1) \sim (c_2, \sigma_2)$ if they are associate in the sense of [26, Section 5]. By [25, Theorem 5.2] there is a well-defined map

$$\mathrm{type} : \Pi^0(\boldsymbol{G}(\mathsf{k})) \to S(\boldsymbol{G})/\sim,$$

which attaches to X a well-defined and unique association class $(c, \sigma) \in S(\boldsymbol{G})$ such that σ appears as a subrepresentation of $X^{\boldsymbol{U}_c(\mathfrak{o})}$. This association class is called the unrefined minimal K-type of X, and we write $\mathrm{type}(X)$ for brevity (for depth-0 representations unrefined minimal K-types are types). For $(c_1, \sigma_1), (c_2, \sigma_2) \in S(\boldsymbol{G})$, if $(c_1, \sigma_1) \sim (c_2, \sigma_2)$ then $(c_1, \mathrm{WF}(\sigma_1)) \sim_K (c_2, \mathrm{WF}(\sigma_2))$ where WF denotes the Kawanaka wavefront set and $\sim_K$ is the equivalence relation defined in [27, Section 1.1.1]. In particular, since $\mathcal{L}$ is constant on $\sim_K$-classes we have that

$$\mathcal{L}(c_1, \mathrm{WF}(\sigma_1)) = \mathcal{L}(c_2, \mathrm{WF}(\sigma_2)). \tag{2.6.1}$$

For $X \in \Pi^0(\boldsymbol{G}(\mathsf{k}))$ define ${}^K\mathbb{O}_{\mathrm{type}}(X)$ to be the well-defined orbit $\mathcal{L}_c(\mathrm{WF}(\sigma))$ where $(c, \sigma) = \mathrm{type}(X)$.

3. Wavefront sets of depth-0 supercuspidal representations

Lemma 3.0.1 [24, Proposition 1.4, Proposition 2.1, Corollary 3.5]. *Let $(\pi, X) \in \Pi^0(\mathbf{G}(\mathsf{k}))$ be a depth-0 representation and let* $\mathrm{type}(X) = [(c, \sigma)]$. *Then (π, X) is supercuspidal if and only if c is a minimal facet of the building. When X is supercuspidal it is of the form* $\mathrm{ind}_{\mathbf{P}_c^\dagger(\mathfrak{o})}^{\mathbf{G}(\mathsf{k})}(\sigma^\dagger)$ *where $\sigma^\dagger$ is some irreducible representation of $\mathbf{P}_c^\dagger(\mathfrak{o})$ which contains the inflation of an irreducible Deligne–Lusztig cuspidal unipotent representation σ of $\mathbf{L}_c(\mathbb{F}_q)$ upon restriction to $\mathbf{P}_c(\mathfrak{o})$.*

Lemma 3.0.2. *For $X \in \Pi^0(\mathbf{G}(\mathsf{k}))$ we have that ${}^K\mathbb{O}_{\mathrm{type}}(X) \le \mathbb{O}$ for some $\mathbb{O} \in {}^K\mathrm{WF}(X)$. In particular, when ${}^K\mathrm{WF}(X)$ is a singleton we have*

$$ {}^K\mathbb{O}_{\mathrm{type}}(X) \le {}^K\mathrm{WF}(X). $$

Proof. Let $\mathrm{type}(X) = [(c, \sigma)]$. Then σ is a subrepresentation of $X^{\mathbf{U}_c(\mathfrak{o})}$ and so $\mathrm{WF}(\sigma) \le \mathbb{O}$ for some $\mathbb{O} \in \mathrm{WF}(X^{\mathbf{U}_c(\mathfrak{o})})$. Thus

$$ \mathcal{L}_c(\mathrm{WF}(\sigma)) \le \mathcal{L}_c(\mathbb{O}) \in {}^K\mathrm{WF}_c(X). $$

The result then follows from the fact that

$$ {}^K\mathrm{WF}(X) = \max\{{}^K\mathrm{WF}_c(X) : c \subseteq_f \mathcal{B}(\mathbf{G}, \mathsf{k})\}. \qquad \square $$

For an Iwahori-spherical representation X this inequality says nothing because $\mathrm{type}(X) = [(c_0, \mathrm{triv})]$ where c_0 is a chamber of the building and so ${}^K\mathbb{O}_{\mathrm{type}}(X)$ is the zero orbit. For supercuspidal representations however, the inequality is in fact an equality. We now proceed to prove this.

Lemma 3.0.3. *Let X be a depth-0 supercuspidal representation. Let c' be a face of $\mathcal{B}(\mathbf{G}, \mathsf{k})$ with $X^{\mathbf{U}_{c'}(\mathfrak{o})} \ne 0$ and suppose that τ is an irreducible constituent of $X^{\mathbf{U}_{c'}(\mathfrak{o})}$. Then τ is a cuspidal representation of $\mathbf{L}_{c'}(\mathbb{F}_q)$ and in particular* $[(c', \tau)] = \mathrm{type}(X)$.

Proof. Let $\mathrm{type}(X) = [(c, \sigma)]$ and c', τ be as in the statement of the lemma. Let $[(M, \tau')]$ be the cuspidal data for τ (i.e., a conjugacy class of Levi of $\mathbf{L}_{c'}(\mathbb{F}_q)$ and cuspidal representation of said Levi). In particular, if M is included into any parabolic P so that P has Levi decomposition $P = MU$, then τ' is a subrepresentation of τ^U. Now, all of the parabolics of $\mathbf{L}_{c'}(\mathfrak{o})$ are conjugate to a parabolic of the form $\mathbf{P}_{c''}(\mathfrak{o})/\mathbf{U}_{c'}(\mathfrak{o})$ where $c' \subseteq_f \overline{c''}$. Thus (conjugating M appropriately) we can find a c'' such that M is a Levi factor of $\mathbf{P}_{c''}(\mathfrak{o})/\mathbf{U}_{c'}(\mathfrak{o})$ and so $\mathbf{L}_{c''}(\mathbb{F}_q) \simeq M$. We thus have that

$$ \tau' \subseteq \tau^U \subseteq (X^{\mathbf{U}_{c'}(\mathfrak{o})})^{\mathbf{U}_{c''}(\mathfrak{o})/\mathbf{U}_{c'}(\mathfrak{o})} = X^{\mathbf{U}_{c''}(\mathfrak{o})}. $$

In particular (c'', τ') is an unrefined minimal K-type for X. Thus by [26, Theorem 5.2], we have that $(c'', \tau') \sim (c, \sigma)$. In particular c'' is also a minimal face and so $c'' = c'$ and $\tau = \tau'$. Thus τ is a cuspidal representation of $\mathbf{L}_{c'}(\mathbb{F}_q)$ and $[(c, \sigma)] = [(c', \tau)]$. $\square$

Theorem 3.0.4. *Let (π, X) be a depth-0 supercuspidal representation of $\mathbf{G}(\mathsf{k})$. Then*

$$ {}^K\mathrm{WF}(X) = {}^K\mathbb{O}_{\mathrm{type}}(X), \quad \mathrm{DWF}(X) = i_K({}^K\mathbb{O}_{\mathrm{type}}(X)). $$

In particular the unramified, double, and geometric wavefront sets are all singletons.

Proof. By Theorems 2.4.2 and 2.5.2

$$^{K}\mathrm{WF}(X) = \max\{^{K}\mathrm{WF}_c(X) : c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})\}$$
$$\mathrm{DWF}(X) = \max\{i_K(^{K}\mathrm{WF}_c(X)) : c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})\}.$$

Suppose $c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})$ is such that $X^{\boldsymbol{U}_c(\mathfrak{o})} \neq 0$. Then for any τ an irreducible constituent of $X^{\boldsymbol{U}_c(\mathfrak{o})}$ we have by Lemma 3.0.3 that $[(c, \tau)] = \mathrm{type}(X)$. Thus we must have that $^{K}\mathrm{WF}_c(X) = {}^{K}\mathbb{O}_{\mathrm{type}}(X)$. If $c \subseteq_f \mathcal{B}(\boldsymbol{G}, \mathsf{k})$ is such that $X^{\boldsymbol{U}_c(\mathfrak{o})} = 0$ then $^{K}\mathrm{WF}_c(X, \mathbb{C})$ is the zero orbit. Thus

$$^{K}\mathrm{WF}(X) = {}^{K}\mathbb{O}_{\mathrm{type}}(X), \quad \mathrm{DWF}(X) = i_K(^{K}\mathbb{O}_{\mathrm{type}}(X)). \qquad \square$$

4. Preliminaries, II

In Section 5 we will use Theorem 3.0.4 to deduce a formula for the wavefront set of a unipotent supercuspidal representation in terms of its Langlands parameters, in the special case when the group is inner to split. For this, we will need some additional notation and preliminaries.

Let $\boldsymbol{G}$ denote a split group defined over $\mathbb{Z}$ and let $\boldsymbol{T} \subset \boldsymbol{G}$ be a maximal torus. For any field F, we write $\boldsymbol{G}(F)$, $\boldsymbol{T}(F)$, etc. for the groups of F-rational points. The $\mathbb{C}$-points are denoted by G, T, etc. Let $\boldsymbol{G}_{\mathrm{ad}} = \boldsymbol{G}/Z(\boldsymbol{G})$ denote the adjoint group of $\boldsymbol{G}$. Write $X^*(\boldsymbol{T}, \bar{\mathsf{k}})$ (resp. $X_*(\boldsymbol{T}, \bar{\mathsf{k}})$) for the lattice of algebraic characters (resp. cocharacters) of $\boldsymbol{T}(\bar{\mathsf{k}})$, and write $\Phi(\boldsymbol{T}, \bar{\mathsf{k}})$ (resp. $\Phi^\vee(\boldsymbol{T}, \bar{\mathsf{k}})$) for the set of roots (resp. coroots). Let

$$\mathcal{R} = (X^*(\boldsymbol{T}, \bar{\mathsf{k}}),\ \Phi(\boldsymbol{T}, \bar{\mathsf{k}}),\ X_*(\boldsymbol{T}, \bar{\mathsf{k}}),\ \Phi^\vee(\boldsymbol{T}, \bar{\mathsf{k}}),\ \langle\,\cdot\,,\cdot\,\rangle)$$

be the root datum corresponding to $\boldsymbol{G}$, and let W be the associated (finite) Weyl group. Let $\boldsymbol{G}^\vee$ be the Langlands dual group of $\boldsymbol{G}$, that is, the connected reductive algebraic group defined over $\mathbb{Z}$ corresponding to the dual root datum

$$\mathcal{R}^\vee = (X_*(\boldsymbol{T}, \bar{\mathsf{k}}),\ \Phi^\vee(\boldsymbol{T}, \bar{\mathsf{k}}),\ X^*(\boldsymbol{T}, \bar{\mathsf{k}}),\ \Phi(\boldsymbol{T}, \bar{\mathsf{k}}),\ \langle\,\cdot\,,\cdot\,\rangle).$$

Set $\Omega = X_*(\boldsymbol{T}, \bar{\mathsf{k}})/\mathbb{Z}\Phi^\vee(\boldsymbol{T}, \bar{\mathsf{k}})$. The center $Z(\boldsymbol{G}^\vee)$ can be naturally identified with the irreducible characters $\mathsf{Irr}\Omega$, and dually, $\Omega \cong X^*(Z(\boldsymbol{G}^\vee))$. For $\omega \in \Omega$, let ζ_ω denote the corresponding irreducible character of $Z(\boldsymbol{G}^\vee)$.

For details regarding the parameterization of inner twists of a group $\boldsymbol{G}(\mathsf{k})$, see [3, §1.3; 13, §1; 17, §2; 19; 36, §2]. We record here only that the set of equivalence classes of inner twists of the split form of $\boldsymbol{G}$ are parameterized by the Galois cohomology group

$$H^1(\Gamma, \boldsymbol{G}_{\mathrm{ad}}) \cong H^1(F, \boldsymbol{G}_{\mathrm{ad}}(K)) \cong \Omega_{\mathrm{ad}} \cong \mathrm{Irr}\, Z(\boldsymbol{G}^\vee_{\mathsf{sc}}),$$

where $\boldsymbol{G}^\vee_{\mathsf{sc}}$ is the Langlands dual group of $\boldsymbol{G}_{\mathrm{ad}}$, i.e., the simply connected cover of $\boldsymbol{G}^\vee$, and F denotes the action of Frob on $\boldsymbol{G}(K)$. We identify Ω_{ad} with the fundamental group of $\boldsymbol{G}_{\mathrm{ad}}$. The isomorphism above is determined as follows: For a cohomology class h in $H^1(F, \boldsymbol{G}_{\mathrm{ad}}(K))$, let z be a representative cocycle. Let $u \in \boldsymbol{G}_{\mathrm{ad}}(K)$ be the image of F under z, and let ω denote the image of u in Ω_{ad}. Set $F_\omega = \mathrm{Ad}(u) \circ F$.

The corresponding rational structure of $\boldsymbol{G}$ is given by F_ω. Let $\boldsymbol{G}^\omega$ be the connected reductive group defined over k such that $\boldsymbol{G}(K)^{F_\omega} = \boldsymbol{G}^\omega(\mathsf{k})$. Note that $\boldsymbol{G}^1 = \boldsymbol{G}$ (where we view $\boldsymbol{G}$ as an algebraic group over k for this equality). The minimal Galois extension of K over which $\boldsymbol{G}^\omega$ splits is K itself. So we may take $E = K$ for the rest of the paper.

If H is a complex reductive group and x is an element of H or $\mathfrak{h}$, we write $H(x)$ for the centralizer of x in H, and $A_H(x)$ for the group of connected components of $H(x)$. If S is a subset of H or $\mathfrak{h}$ (or indeed, of $H \cup \mathfrak{h}$), we can similarly define $H(S)$ and $A_H(S)$. We will sometimes write $A(x)$, $A(S)$ when the group H is implicit. The subgroups of H of the form $H(x)$ where x is a semisimple element of H are called *pseudo-Levi* subgroups of H.

4.1. ***The Bruhat–Tits building.*** We will recall some standard facts about the Bruhat–Tits building (all of which can be found in [35]).

Fix a $\omega \in \Omega$ and let $\boldsymbol{G}^\omega$ be the inner twist of $\boldsymbol{G}$ corresponding to ω as defined in the previous section. We write $\boldsymbol{T}$ for the split torus scheme over $\mathfrak{o}$ with generic fiber $\boldsymbol{T}$. This scheme $\boldsymbol{T}$ defined over $\mathfrak{o}$ is a subgroup of $\boldsymbol{P}_c$ for any $c \subseteq_f \mathcal{B}(\boldsymbol{G}^\omega, \mathsf{k})$ and the special fiber of $\boldsymbol{T}$, denoted $\bar{\boldsymbol{T}}$, is a maximal torus of $\boldsymbol{L}_c$.

For an apartment $\mathcal{A}$ of $\mathcal{B}(\boldsymbol{G}, K)$ and $\Omega \subseteq \mathcal{A}$ we write $\mathcal{A}(\Omega, \mathcal{A})$ for the smallest affine subspace of $\mathcal{A}$ containing Ω. The inner twist $\boldsymbol{G}^\omega$ of $\boldsymbol{G}$ gives rise to an action of the Galois group $\mathrm{Gal}(K/k)$ on $\mathcal{B}(\boldsymbol{G}, K)$ and we can (and will) identify $\mathcal{B}(\boldsymbol{G}^\omega, \mathsf{k})$ with the fixed points of this action. Write $\Phi(\boldsymbol{T}, K)$ (resp. $\Psi(\boldsymbol{T}, K)$) for the set of roots of $\boldsymbol{G}(K)$ (resp. affine roots) of $\boldsymbol{G}^\omega(K) = \boldsymbol{G}(K)$ relative to $\boldsymbol{T}$. For $\psi \in \Psi(\boldsymbol{T}, \mathsf{k})$ write $\dot{\psi} \in \Phi(\boldsymbol{T}, \mathsf{k})$ for the gradient of ψ, and $W = W(\boldsymbol{T}, \mathsf{k})$ for the Weyl group of $\boldsymbol{G}(\mathsf{k})$ with respect to $\boldsymbol{T}(\mathsf{k})$.

For $c \subseteq_f \mathcal{B}(\boldsymbol{G}, K)$, the stabilizer of c in $\boldsymbol{G}(K)$ identifies with $\boldsymbol{P}_c^\dagger(\mathfrak{O})$. It has pro-unipotent radical $\boldsymbol{U}_c(\mathfrak{O})$ and $\boldsymbol{P}_c^\dagger(\mathfrak{O})/\boldsymbol{U}_c(\mathfrak{O}) = \boldsymbol{L}_c^\dagger(\bar{\mathbb{F}}_q)$. For c a face lying in $\mathcal{B}(\boldsymbol{G}^\omega, \mathsf{k}) \subseteq \mathcal{B}(\boldsymbol{G}, K)$, F_ω stabilizes $\boldsymbol{P}_c(\mathfrak{O})$ and induces a Frobenius on $\boldsymbol{L}_c(\bar{\mathbb{F}}_q)$. The group $\boldsymbol{L}_c(\mathbb{F}_q)$ consists of the fixed points of this Frobenius.

For this paper it will be convenient to fix a maximal k-split torus $\boldsymbol{T}_0$ of $\boldsymbol{G}^\omega$ lying in $\boldsymbol{T}$ of $\boldsymbol{G}^\omega$. We have that $\mathcal{A}(\boldsymbol{T}_0, \mathsf{k}) = \mathcal{A}(\boldsymbol{T}, K)^{\mathrm{Gal}(K/k)}$. We will also fix a $\mathrm{Gal}(K/\mathsf{k})$-stable chamber c_0 of $\mathcal{A}(\boldsymbol{T}, K)$ and a special point $x_0 \in c_0$. Let $\widetilde{W} = W \ltimes X_*(\boldsymbol{T}, K)$ be the (extended) affine Weyl group. The choice of special point x_0 of $\mathcal{B}(\boldsymbol{G}, K)$ fixes an inclusion $\Phi(\boldsymbol{T}, K) \to \Psi(\boldsymbol{T}, K)$ and an isomorphism between $\widetilde{W}$ and $N_{\boldsymbol{G}(K)}(\boldsymbol{T}(K))/\boldsymbol{T}(\mathfrak{O}^\times)$. Write

$$\widetilde{W} \to W, \quad w \mapsto \dot{w}, \tag{4.1.1}$$

for the natural projection map. For a face $c \subseteq_f \mathcal{A}$ let W_c be the subgroup of $\widetilde{W}$ generated by reflections in the hyperplanes through c. The special fiber of $\boldsymbol{T}$ (as a scheme over $\mathfrak{O}$) which we denote by $\bar{\boldsymbol{T}}$, is a split maximal torus of $\boldsymbol{L}_c(\bar{\mathbb{F}}_q)$. Write $\Phi_c(\bar{\boldsymbol{T}}, \bar{\mathbb{F}}_q)$ for the root system of $\boldsymbol{L}_c$ with respect to $\bar{\boldsymbol{T}}$. Then $\Phi_c(\bar{\boldsymbol{T}}, \bar{\mathbb{F}}_q)$ naturally identifies with the set of $\psi \in \Psi(\boldsymbol{T}, K)$ that vanish on c, and the Weyl group of $\bar{\boldsymbol{T}}$ in $\boldsymbol{L}_c$ is isomorphic to W_c.

Recall that a choice of x_0 fixes an embedding $\Phi(\boldsymbol{T}, K) \to \Psi(\boldsymbol{T}, L)$. If we fix a set of simple roots $\Delta \subset \Phi(\boldsymbol{T}, K)$, this embedding determines a set of extended simple roots $\tilde{\Delta} \subseteq \Psi(\boldsymbol{T}, K)$. When $\Phi(\boldsymbol{T}, K)$ is irreducible, $\tilde{\Delta}$ is just the set $\Delta \cup \{1 - \alpha_0\}$ where α_0 is the highest root of $\Phi(\boldsymbol{T}, K)$ with respect to Δ.

When $\Phi(\boldsymbol{T}, K)$ is reducible, say $\Phi(\boldsymbol{T}, K) = \cup_i \Phi_i$ where each Φ_i is irreducible, then $\tilde{\Delta} = \cup_i \tilde{\Delta}_i$ where $\Delta_i = \Phi_i \cap \Delta$. Fix Δ so that the chamber cut out by $\tilde{\Delta}$ is c_0. Let

$$\boldsymbol{P}(\tilde{\Delta}) := \{J \subsetneq \tilde{\Delta} : J \cap \tilde{\Delta}_i \subsetneq \tilde{\Delta}_i, \forall i\}.$$

Each $J \in \boldsymbol{P}(\tilde{\Delta})$ cuts out a face of c_0 which we denote by $c(J)$. In particular $c(\Delta) = x_0$. Since $\Omega \simeq \widetilde{W}/W \ltimes \mathbb{Z}\Phi(\boldsymbol{T}, K)$, and $W \ltimes \mathbb{Z}\Phi(\boldsymbol{T}, K)$ acts simply transitively on the chambers of $\mathcal{A}(\boldsymbol{T}, K)$, the action of $\widetilde{W}$ on $\mathcal{A}(\boldsymbol{T}, K)$ induces an action of Ω on the faces of c_0 and hence on $\tilde{\Delta}$ (and $\boldsymbol{P}(\tilde{\Delta})$). For $\omega \in \Omega$ let σ_ω denote the corresponding permutation of $\tilde{\Delta}$. Let

$$\boldsymbol{P}^\omega(\tilde{\Delta}) := \{J \in \boldsymbol{P}(\tilde{\Delta}) \mid \sigma_\omega(J) = J\}$$

and let c_0^ω be the chamber of $\mathcal{B}(\boldsymbol{G}^\omega, \mathsf{k})$ lying in c_0. The set $\boldsymbol{P}^\omega(\tilde{\Delta})$ is an indexing set for the faces of c_0^ω. For $J \in \boldsymbol{P}^\omega(\tilde{\Delta})$ write $c^\omega(J)$ for the face of c_0^ω corresponding to J. The face $c^\omega(J)$ lies in $c(J)$. For $J, J' \in \boldsymbol{P}^\omega(\tilde{\Delta})$ (resp. $\boldsymbol{P}(\tilde{\Delta})$) we have $J \subseteq J'$ if and only if $\overline{c^\omega(J)} \supseteq c^\omega(J')$ (resp. $\overline{c(J)} \supseteq c(J')$).

4.2. ***Lifting nilpotent orbits.*** Recall the group G from Section 2 and the definition of psuedo-Levi subgroups from Section 4. Fix a maximal torus T of G. Call a pseudo-Levi subgroup L of G *standard* if it contains T and write Z_L for its center. Let $\mathcal{A} = \mathcal{A}(\boldsymbol{T}, K)$.

Lemma 4.2.1 [27, Section 2.14, Corollary 2.19]. *There is a W-equivariant map*

$$\mathfrak{L}_{x_0} : \{\textit{faces of } \mathcal{A}\} \to \{(L, tZ_L^\circ) \mid L \textit{ a standard pseudo-Levi},\ Z_G^\circ(tZ_L^\circ) = L\}, \tag{4.2.1}$$

where c_1, c_2 lie in the same fiber if and only if

$$\mathcal{A}(c_1, \mathcal{A}) + X_*(\boldsymbol{T}, K) = \mathcal{A}(c_2, \mathcal{A}) + X_*(\boldsymbol{T}, K).$$

If $\mathfrak{L}_{x_0}(c) = (L, tZ_L^\circ)$ then L is the complex reductive group with the same root datum as $\boldsymbol{L}_c(\bar{F}_q)$ and thus there is an isomorphism $\Lambda_c^{\bar{F}_q} : \mathcal{N}_o^{\boldsymbol{L}_c}(\bar{\mathbb{F}}_q) \xrightarrow{\sim} \mathcal{N}_o^L$.

Recall from the end of Section 4.1 the definitions of Δ, $\tilde{\Delta}$, c_0, $\boldsymbol{P}(\tilde{\Delta})$ and $c(J)$. The definitions of c_0 and $c(J)$ depend on a choice of x_0 and Δ. Let L_J denote the pseudo-Levi subgroup of G generated by T and the root groups corresponding to $\dot{\alpha}$ for $\alpha \in J$. Then $\mathrm{pr}_1 \circ \mathfrak{L}_{x_0}(c(J)) = L_J$ (by [27, Lemma 2.21], this group should not depend on x_0). Define

$$\begin{aligned} \mathcal{I}_{x_0, \tilde{\Delta}} &= \{(J, \mathbb{O}) \mid J \in \boldsymbol{P}(\tilde{\Delta}),\ \mathbb{O} \in \mathcal{N}_o^{\boldsymbol{L}_{c(J)}}(\bar{\mathbb{F}}_q)\}, \\ \mathcal{K}_{\tilde{\Delta}} &= \{(J, \mathbb{O}) \mid J \in \boldsymbol{P}(\tilde{\Delta}),\ \mathbb{O} \in \mathcal{N}_o^{L_J}(\mathbb{C})\}. \end{aligned} \tag{4.2.2}$$

The map

$$\iota_{x_0} : \mathcal{I}_{x_0, \tilde{\Delta}} \to \mathcal{K}_{\tilde{\Delta}}, \quad (J, \mathbb{O}) \mapsto (J, \Lambda_{c(J)}^{\bar{F}_q}(\mathbb{O})),$$

is an isomorphism. Let

$$\mathbb{L} : \mathcal{K}_{\tilde{\Delta}} \to \mathcal{N}_{c,o} \tag{4.2.3}$$

be the map that sends $(J, \mathbb{O})$ to $(Gx, tZ_G^\circ(x))$ where $x \in \mathbb{O}$ and $(L, tZ_L^\circ) = \mathfrak{L}(c(J))$ where $\mathfrak{L}$ is the map from [27, Corollary 2.19]. By [10, Theorem 2.1.7] the diagram

$$\begin{tikzcd} \mathcal{I}_{x_0,\tilde{\Delta}} \arrow[r, "\sim", "\iota_{x_0}"'] \arrow[d, two heads, "\mathcal{L}"] & \mathcal{K}_{\tilde{\Delta}} \arrow[d, two heads, "\mathbb{L}"] \\ \mathcal{N}_o(K) \arrow[r, "\sim", "\theta_{x_0,\boldsymbol{T}}"'] & \mathcal{N}_{o,c} \end{tikzcd} \tag{4.2.4}$$

commutes. Define

$$\bar{\mathbb{L}} = \mathfrak{Q} \circ \mathbb{L}.$$

This map can be computed using Achar's algorithms in [1, Section 3.4].

4.3. ***Isogenies.*** Let $f : \boldsymbol{H}' \to \boldsymbol{H}$ be an isogeny of connected reductive groups defined over k. Let $f_k : \boldsymbol{H}'(\mathsf{k}) \to \boldsymbol{H}(\mathsf{k})$ denote the corresponding homomorphism of k-points. We note that $\mathcal{N}_{o,\bar{c}}^{\boldsymbol{H}} \simeq \mathcal{N}_{o,\bar{c}}^{\boldsymbol{H}'}$ and so we can compare the double wavefront sets of representations of the two groups.

Lemma 4.3.1. *Let X be an irreducible smooth depth-0 representation of $\boldsymbol{H}(\mathsf{k})$ and write X' for the representation of $\boldsymbol{H}'(\mathsf{k})$ obtained by pulling back along f_k. Then X' decomposes as a finite sum of irreducible smooth representations $X' = \bigoplus_i X_i'$ and* $\mathrm{DWF}(X) = \mathrm{DWF}(X_i')$ *for all i.*

Proof. This is an easy consequence of Clifford theory since the image of G under the isogeny is normal and of finite index. □

4.4. ***Unipotent supercuspidal representations.*** Recall Lusztig's notion of a unipotent representation of a finite group of Lie type [20, Section 6.5]. Let

$$S_{\text{unip}}(\boldsymbol{G}) := \{(c, \sigma) \in S(\boldsymbol{G}) : \sigma \text{ is } \textit{unipotent}\}.$$

Definition 4.4.1. Let X be a depth-0 irreducible $\boldsymbol{G}(\mathsf{k})$-representation. We say that X has *unipotent cuspidal support* if $\mathrm{type}(X) \in S_{\text{unip}}(\boldsymbol{G})$. Write $\Pi^{\mathsf{Lus}}(\boldsymbol{G}(\mathsf{k}))$ for the subset of $\Pi^0(\boldsymbol{G}(\mathsf{k}))$ consisting of all such representations. We call a supercuspidal representation 'unipotent' if it is depth-0 and has unipotent cuspidal support.

4.5. ***Langlands classification of unipotent supercuspidal representations.*** Let W_k be the Weil group of k with inertia subgroup I_k and set $W_\mathsf{k}' = W_\mathsf{k} \times \mathrm{SL}(2, \mathbb{C})$. We will think of a *Langlands parameter* for $\boldsymbol{G}$ as a continuous morphism $\varphi : W_\mathsf{k}' \to G^\vee$ such that $\varphi(w)$ is semisimple for each $w \in W_\mathsf{k}$ and the restriction of φ to $\mathrm{SL}(2, \mathbb{C})$ is algebraic. A Langlands parameter φ is called *unramified* if $\varphi(I_\mathsf{k}) = \{1\}$. Let $G^\vee(\varphi)$ denote the centralizer of $\varphi(W_\mathsf{k}')$ in $G^\vee$. Define

$$Z^1_{G^\vee_{\mathsf{sc}}}(\varphi) = \text{ preimage of } G^\vee(\varphi)/Z(G^\vee) \text{ under the projection } G^\vee_{\mathsf{sc}} \to G^\vee_{\mathrm{ad}},$$

and let A^1_φ denote the component group of $Z^1_{G^\vee_{\mathsf{sc}}}(\varphi)$. An *enhanced Langlands parameter* is a pair (φ, ρ), where $\rho \in \mathrm{Irr}(A^1_\varphi)$. A parameter (φ, ρ) is called $\boldsymbol{G}^\omega$*-relevant* (recall that $\boldsymbol{G}^\omega$ is an inner twist of the split form, $\omega \in \Omega_{\mathrm{ad}}$) if ρ acts on $Z(G^\vee_{\mathsf{sc}})$ by a multiple of the character ζ_ω.

Define the elements

$$s_\varphi = \varphi(\mathrm{Frob}, 1), \quad u_\varphi = \varphi\left(1, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right).$$

Following [4], consider the possibly disconnected reductive group

$$\mathcal{G}_\varphi = Z^1_{G^\vee_{\mathrm{sc}}}(\varphi(W_{\mathsf{k}})),$$

which is defined analogously to $Z^1_{G^\vee_{\mathrm{sc}}}(\varphi)$. Then $u_\varphi \in \mathcal{G}_\varphi^\circ$ and by [4, (92)]

$$A^1_\varphi \cong \mathcal{G}_\varphi(u_\varphi)/\mathcal{G}_\varphi(u_\varphi)^\circ.$$

An enhanced Langlands parameter (φ, ρ) is called *discrete* if $G^\vee(\varphi)$ does not contain a nontrivial torus (this notion is independent of ρ). A discrete parameter is called *cuspidal* if (u_φ, ρ) is a cuspidal pair. This means that every ρ° which occurs in the restriction of ρ to $A_{\mathcal{G}_\varphi^\circ}(u_\varphi)$ defines a $\mathcal{G}_\varphi^\circ$-equivariant local system on the $\mathcal{G}_\varphi^\circ$-conjugacy class of u_φ which is cuspidal in the sense of Lusztig.

A Langlands correspondence for unipotent supercuspidal representations has been obtained by [24] when $\boldsymbol{G}$ is simple and adjoint; see also [22]. For arbitrary reductive K-split groups, this correspondence is available by [12; 13]. Let $\mathrm{Irr}(\boldsymbol{G}^\omega(\mathsf{k}))_{\mathrm{cusp,unip}}$ denote the set of equivalence classes of irreducible unipotent supercuspidal $\boldsymbol{G}^\omega(\mathsf{k})$-representations. Let $\Phi(G^\vee)^\omega_{\mathrm{cusp,nr}}$ denote the set of $G^\vee$-equivalence classes of unramified cuspidal enhanced Langlands parameters (φ, ρ) which are $\boldsymbol{G}^\omega$-relevant.

Theorem 4.5.1. *For every $\omega \in \Omega_{\mathrm{ad}}$, there is a bijection*

$$\Phi(G^\vee)^\omega_{\mathrm{cusp,nr}} \longleftrightarrow \mathrm{Irr}(\boldsymbol{G}^\omega(\mathsf{k}))_{\mathrm{cusp,unip}}.$$

This bijection satisfies several natural desiderata (*including formal degrees*, *equivariance with respect to tensoring by weakly unramified characters*); *see* [13, Theorem 2].

For X a unipotent supercuspidal representation of $\boldsymbol{G}^\omega(\mathsf{k})$ let φ denote the corresponding Langlands parameter. We will write $\mathbb{O}^\vee_X \in \mathcal{N}^\vee_o$ for the $G^\vee$-orbit of

$$n_\varphi = d\varphi\left(0, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right).$$

Lemma 4.5.2. *Let $X, X' \in \mathrm{Irr}(\boldsymbol{G}^\omega(\mathsf{k}))_{\mathrm{cusp,unip}}$. If* $\mathrm{type}(X) = \mathrm{type}(X')$ *then* $\mathbb{O}^\vee_X = \mathbb{O}^\vee_{X'}$.

Proof. This follows by inspecting the explicit classification in [13]. □

For $[(c, \sigma)] \in S_{\mathrm{unip}}(\boldsymbol{G})$ with c a minimal face we write $\mathbb{O}^\vee(c, \sigma)$ for the common nilpotent parameter of all $X \in \mathrm{Irr}(\boldsymbol{G}^\omega(\mathsf{k}))_{\mathrm{cusp,unip}}$ with $\mathrm{type}(X) = [(c, \sigma)]$.

We will recall the explicit classification in Section 7.

5. Wavefront sets of unipotent supercuspidal representations

Proposition 5.0.1. *Suppose $\boldsymbol{G}$ is simple and adjoint and let $[(c,\sigma)] \in S_{\mathrm{unip}}(\boldsymbol{G}^\omega)$ be such that c is a minimal face. Then*

$$\bar{\mathbb{L}}(c, \mathrm{WF}(\sigma)) = d_A(\mathbb{O}^\vee(c,\sigma), 1).$$

This proposition will be proved in Section 8.

Theorem 5.0.2. *Let $\boldsymbol{G}$ be a split reductive group defined over k. Let $\omega \in \Omega$ and let $\boldsymbol{G}^\omega$ denote the corresponding inner twist of $\boldsymbol{G}$. Let (π, X) be an irreducible supercuspidal $\boldsymbol{G}^\omega(\mathsf{k})$-representation with unipotent cuspidal support. Then ${}^K\mathrm{WF}(X)$, $\mathrm{DWF}(X)$, ${}^{\bar{\mathsf{k}}}\mathrm{WF}(X)$ are singletons, and*

$$d_S({}^K\mathrm{WF}(X)) = \mathbb{O}_X^\vee, \quad \mathrm{DWF}(X) = (d(\mathbb{O}_X^\vee), \mathbb{O}_X^\vee), \quad {}^{\bar{\mathsf{k}}}\mathrm{WF}(X) = d(\mathbb{O}_X^\vee).$$

Proof. By Theorem 3.0.4, the second component of $\mathrm{DWF}(X)$ is exactly $d_S({}^K\mathrm{WF}(X))$. By (2.5.1) the first component of $\mathrm{DWF}(X)$ is ${}^{\bar{\mathsf{k}}}\mathrm{WF}(X)$. Thus it suffices to prove the second equality only. Suppose first that $\boldsymbol{G}$ is simple and adjoint.

Let X be a unipotent supercuspidal representation of $\boldsymbol{G}^\omega(\mathsf{k})$. By Theorem 3.0.4 we have that

$$\mathrm{DWF}(X) = i_K({}^K\mathbb{O}_{\mathrm{type}}(X)).$$

Write $\mathrm{type}(X) = [(c,\sigma)]$. By definition

$${}^K\mathbb{O}_{\mathrm{type}}(X) = \mathbb{L}(c, \mathrm{WF}(\sigma)).$$

By Proposition 5.0.1 we have that

$$\bar{\mathbb{L}}(c, \mathrm{WF}(\sigma)) = d_A(\mathbb{O}^\vee(c,\sigma), 1).$$

Since $\mathbb{O}_X^\vee = \mathbb{O}^\vee(c,\sigma)$ we get that

$$\mathrm{DWF}(X) = d_A(\mathbb{O}_X^\vee, 1)$$

as required.

Applying Lemma 4.3.1 we get that the theorem holds for all simply connected simple groups. Since wavefront sets behave as expected with respect to products, the theorem holds for all simply connected reductive groups. Finally, applying Lemma 4.3.1 again we get that the theorem holds for all split reductive groups $\boldsymbol{G}$. □

6. Unipotent cuspidal representations of finite reductive groups

Let $\boldsymbol{G}$ be a reductive group defined over $\mathbb{F}_q$. We list the unipotent cuspidal representations of $\boldsymbol{G}(\mathbb{F}_q)$ and their Kawanaka wavefront sets. Since the classification of unipotent representations is independent of the isogeny, we may assume without loss of generality that the group G is simple and adjoint. For the explicit results about the parameterization of unipotent representations of finite groups of Lie type, we refer to [20, §4, §8.1] and [7, §13.8, §13.9]. The relevant results for the Kawanaka wavefront sets and unipotent support are in [21, §10, §11].

6.1. *Classical groups.*

6.1.1. $A_{n-1}(q)$. The group $\boldsymbol{G} = \mathrm{PGL}(n)$ does not have unipotent cuspidal representations.

6.1.2. ${}^2A_n(q^2)$. The group $\boldsymbol{G} = \mathrm{PU}(n+1)$ has unipotent representations if and only if $n = \frac{r(r+1)}{2} - 1$, for some integer $r \geq 2$. The unipotent ${}^2A_n(q^2)$-representations are in one-to-one correspondence with partitions of $n+1$, and so are the geometric nilpotent orbits of $\boldsymbol{G}$. When $n = \frac{r(r+1)}{2} - 1$, the cuspidal unipotent representation σ is unique and it is parameterized by the partition

$$(1, 2, 3, \ldots, r).$$

Its Kawanaka wavefront set is $\mathrm{WF}(\sigma) = (1, 2, 3, \ldots, r)$.

6.1.3. $B_n(q)$, $C_n(q)$. Suppose $\boldsymbol{G}$ is $\mathrm{SO}(2n+1)$ or $\mathrm{PSp}(2n)$ over $\mathbb{F}_q$. The group $\boldsymbol{G}(\mathbb{F}_q)$ has a unipotent cuspidal representation (and in this case the cuspidal representation is unique) if and only if $n = r^2 + r$ for a positive integer r. The unipotent representations of $\boldsymbol{G}(\mathbb{F}_q)$ are parameterized by symbols

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_a \\ & \mu_1 & \cdots & \mu_b \end{pmatrix},$$

$0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_a$, $0 \leq \mu_1 < \mu_2 < \cdots < \mu_b$, $a - b$ odd and positive, and λ_1, μ_1 are not both zero, such that

$$n = \sum \lambda_i + \sum \mu_j - \left(\frac{a+b-1}{2}\right)^2.$$

Let $d = a - b$ be the defect of the symbol. Two unipotent representations belong to the same family if their symbols have the same entries with the same multiplicities. For the unipotent cuspidal representation σ, the corresponding symbol has defect $d = 2r + 1$ and it is

$$\begin{pmatrix} 0 & 1 & 2 & \cdots & 2r \\ & & - & & \end{pmatrix}.$$

The geometric nilpotent orbits of $\mathrm{SO}(2n+1)$ (resp. $\mathrm{PSp}(2n)$) are parameterized by partitions of $2n+1$ (resp. $2n$), where the even (resp. odd) parts occur with even multiplicity. The Kawanaka wavefront set of the unipotent cuspidal representation σ is

$$\mathrm{WF}(\sigma) = \begin{cases} (1, 1, 3, 3, \ldots, 2r-1, 2r-1, 2r+1) & \text{if } G = \mathrm{SO}(2n+1), \\ (2, 2, 4, 4, \ldots, 2r, 2r) & \text{if } G = \mathrm{PSp}(2n). \end{cases} \tag{6.1.1}$$

6.1.4. $D_n(q)$. Suppose $\boldsymbol{G}$ is the split orthogonal group $\mathrm{PSO}(2n)$ over $\mathbb{F}_q$. There exists a unipotent cuspidal representation (and in this case it is unique) if and only if $n = r^2$ for a positive even integer r. The type D_n-symbols are

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_a \\ & \mu_1 & \cdots & \mu_b \end{pmatrix},$$

$0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_a$, $0 \le \mu_1 < \mu_2 < \cdots < \mu_b$, $a-b$ is divisible by 4, and λ_1, μ_1 are not both zero, such that

$$n = \sum \lambda_i + \sum \mu_j - \frac{(a+b)(a+b-2)}{4}.$$

One symbol and the symbol if the row swapped are regarded the same. The irreducible unipotent $\boldsymbol{G}(\mathbb{F}_q)$-representations are in one-to-one correspondence with the type D_n-symbols, except if the symbol has identical rows; in that case there are two nonisomorphic irreducible unipotent representations attached to it. The defect $d = a-b$ is even.

For the unipotent cuspidal representation σ, the corresponding symbol has defect $d = 2r$ and it is

$$\begin{pmatrix} 0 & 1 & 2 & \cdots & 2r-1 \\ & & & - & \end{pmatrix}.$$

The geometric nilpotent orbits of $\mathrm{PSO}(2n)$ are parameterized by partitions of $2n$ with the even parts occurring with even multiplicity. The Kawanaka wavefront set of the unipotent cuspidal representation is

$$\mathrm{WF}(\sigma) = (1, 1, 3, 3, \ldots, 2r-1, 2r-1). \tag{6.1.2}$$

6.1.5. ${}^2D_n(q^2)$. The group ${}^2D_n(q^2)$ admits unipotent cuspidal representations if and only if $n = r^2$, for some odd positive integer r, and in this case the unipotent cuspidal representation is unique. The type ${}^2D_n$-symbols are

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_a \\ & \mu_1 & \cdots & \mu_b \end{pmatrix},$$

$0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_a$, $0 \le \mu_1 < \mu_2 < \cdots < \mu_b$, $a-b \equiv 2 \pmod 4$, and λ_1, μ_1 are not both zero, such that

$$n = \sum \lambda_i + \sum \mu_j - \frac{(a+b)(a+b-2)}{4}.$$

One symbol and the symbol if the row swapped are regarded the same. The irreducible unipotent ${}^2D_n(q^2)$-representations are in one-to-one correspondence with the type ${}^2D_n$ symbols.

For the unipotent cuspidal representation σ, the corresponding symbol and Kawanaka wavefront set are the same as in the split case $D_n(q)$ (except r is now odd).

6.1.6. ${}^3D_4(q^3)$. The group ${}^3D_4(q^3)$ has eight unipotent representations: six are in the principal series, in one-to-one correspondence with the irreducible representations of the Weyl group of type G_2, and two unipotent cuspidal representations, denoted ${}^3D_4[1]$ and ${}^3D_4[-1]$.

The geometric nilpotent orbits of ${}^3D_4(q^3)$ are parameterized by partitions of 8 with even parts occurring with even multiplicity. The unipotent cuspidal representations have Kawanaka wavefront set

$$\mathrm{WF}({}^3D_4[1]) = \mathrm{WF}({}^3D_4[1]) = (1, 1, 3, 3). \tag{6.1.3}$$

6.2. ***Exceptional groups.*** Suppose $\boldsymbol{G}(\mathbb{F}_q)$ is $\mathbb{F}_q$-split. In Table 1, we list all unipotent cuspidal $\boldsymbol{G}(\mathbb{F}_q)$-representations. The irreducible unipotent representations of $\boldsymbol{G}(\mathbb{F}_q)$ are partitioned into families, each

$\boldsymbol{G}(\mathbb{F}_q)$	cuspidal σ	$\mathrm{WF}(\sigma)$	$\mathbb{O}^\vee_\sigma$	$\bar{A}(\mathbb{O}^\vee)$	(x,τ)
G_2	$G_2[1]$ $G_2[-1]$ $G_2[\theta^l]$, $l=1,2$	$G_2(a_1)$	$G_2(a_1)$	S_3	$(1,\epsilon)$ (g_2,ϵ) (g_3,θ^l)
F_4	$F_4^{II}[1]$ $F_4[-1]$ $F_4^I[1]$ $F_4[\theta^l]$, $l=1,2$ $F_4[\pm i]$	$F_4(a_3)$	$F_4(a_3)$	S_4	$(1,\lambda^3)$ (g_2,ϵ) (g_2',ϵ) (g_3,θ^l) $(g_4,\pm i)$
E_6	$E_6[\theta^l]$, $l=1,2$	$D_4(a_1)$	$D_4(a_1)$	S_3	(g_3,θ^l)
E_7	$E_7[\zeta]$ $E_7[\zeta]$	A_4+A_1	A_4+A_1	$\mathbb{Z}/2$	$(g_2,1)$ (g_2,ϵ)
E_8	$E_8^{II}[1]$ $E_8[-1]$ $E_8^I[1]$ $E_8[\theta^l]$, $l=1,2$ $E_8[-\theta^l]$, $l=1,2$ $E_8[\pm i]$ $E_8[\zeta^j]$, $1\le j\le 4$	$E_8(a_7)$	$E_8(a_7)$	S_5	$(1,\lambda^4)$ $(g_2,-\epsilon)$ (g_2',ϵ) $(g_3,\epsilon\theta^l)$ $(g_6,-\theta^l)$ $(g_4,\pm i)$ (g_5,ζ^j)

Table 1. Unipotent cuspidal representations of exceptional groups $\boldsymbol{G}(\mathbb{F}_q)$.

family being in one-to-one correspondence with the set

$$M(\Gamma)=\Gamma\text{-orbits in }\{(x,\tau)\mid x\in\Gamma,\ \tau\in\widehat{Z_\Gamma(x)}\},$$

for a finite group Γ. Each group Γ is uniquely attached to a special nilpotent orbit $\mathbb{O}^\vee$ in the dual Lie algebra, such that $\Gamma=\bar{A}(\mathbb{O}^\vee)$, where $\bar{A}(\mathbb{O}^\vee)$ is Lusztig's canonical quotient.

In Table 1, for each unipotent cuspidal representation σ, we will record the corresponding Kawanaka wavefront set, the nilpotent orbit $\mathbb{O}^\vee$ corresponding to σ and its canonical quotient $\bar{A}(\mathbb{O}^\vee)$, the pair $(x,\tau)\in M(\bar{A}(\mathbb{O}^\vee))$ that parameterizes σ. The geometric nilpotent orbits are given in the Bala–Carter notation.

Finally, for the twisted group ${}^2E_6(q^2)$, there are three unipotent cuspidal representations, denoted ${}^2E_6[1]$, ${}^2E_6[\theta]$, ${}^2E_6[\theta^2]$. All three of them have Kawanaka wavefront set $D_4(a_1)$ in E_6.

7. Langlands parameters for unipotent supercuspidal representations

Recall from Section 4.5 the notation $\mathbb{O}^\vee_X$ and $\mathbb{O}^\vee(c,\sigma)$. We record $\mathbb{O}^\vee_X$ for each unipotent supercuspidal representations (π,X) of an inner to split simple adjoint algebraic group. To only list the information we need we make the following observations. By Lemma 4.5.2, $\mathbb{O}^\vee_X$ only depends on $\mathrm{type}(X)=[(c,\sigma)]$ and by conjugating appropriately we may assume that c is of the form $c^\omega(J)$ for some $J\in\boldsymbol{P}^\omega(\tilde{\Delta})$. Since $c^\omega(J)$

is a minimal face if and only if J is maximal in $\boldsymbol{P}^{\omega}(\tilde{\Delta})$, the possible values for type(X) can be indexed by pairs (J,σ) where J is maximal in $\boldsymbol{P}^{\omega}(\tilde{\Delta})$ and σ is a unipotent cuspidal representation of $\boldsymbol{L}_{c^{\omega}(J)}(\mathbb{F}_q)$. So in this section we list the possible ω along with the set of possible pairs (J,σ) (up to $\sim$) and their corresponding $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$. We use the conventions of [22, Section 6.10] to specify the set $J\subseteq\tilde{\Delta}$. When $\boldsymbol{G}$ is of classical type, the group $\boldsymbol{L}_{c^{\omega}(J)}(\mathbb{F}_q)$ is also of classical type and so if it admits a unipotent cuspidal representation, then it has *exactly one* unipotent cuspidal representation. Thus for the classical types we will only record the J and $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$. The explicit parameters can be found in [12, §4.7; 13; 22; 32].

7.1. *Classical groups.*

7.1.1. PGL(n). If $\boldsymbol{G}=\mathrm{PGL}(n)$, then $G^{\vee}=\mathrm{SL}(n,\mathbb{C})$ and $Z(G^{\vee})=\mathbb{Z}/n\mathbb{Z}$. Hence $\Omega=\mathrm{Irr}(Z(G^{\vee})$ can be identified with C_n. For $\omega\in\Omega$, the inner form $\boldsymbol{G}^{\omega}$ admits unipotent supercuspidal representations if and only if ω has order n and $J=\varnothing$. In this case $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$ is the principal nilpotent orbit.

7.1.2. SO$(2n+1)$. If $\boldsymbol{G}=\mathrm{SO}(2n+1)$, $G^{\vee}=\mathrm{Sp}(2n,\mathbb{C})$ and $Z(G^{\vee})=\mathbb{Z}/2\mathbb{Z}$. The inner forms are parameterized by $\widehat{Z(G^{\vee})}\cong C_2=\{1,-1\}$.

(1) If $\omega=1$, then J is of the form $D_{\ell}\times B_t$, where $\ell+t=n$, $\ell=a^2$, $t=b(b+1)$, a, b nonnegative integers, a even. Let

$$\delta=\begin{cases}b-a & \text{if } b\geq a,\\ a-b-1 & \text{if } a>b,\end{cases}\tag{7.1.1}$$

and $\Sigma=a+b$. The nilpotent orbit $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$ is parameterized by the partition

$$\lambda=(2,4,\dots,2\delta)\cup(2,4,\dots,2\Sigma).\tag{7.1.2}$$

(2) If $\omega=-1$, then J is of the form $D_{\ell}\times B_t$, where $\ell+t=n$, $\ell=a^2$, $t=b(b+1)$, a, b nonnegative integers, where a is now odd. The nilpotent orbit $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$ is defined analogously to the $\omega=1$ case.

7.1.3. PSp$(2n)$. If $\boldsymbol{G}=\mathrm{PSp}(2n)$, then $G^{\vee}=\mathrm{Spin}(2n+1,\mathbb{C})$, and $Z(G^{\vee})=\mathbb{Z}/2\mathbb{Z}$. The inner forms are parameterized by $\widehat{Z(G^{\vee})}\cong C_2=\{1,-1\}$.

(1) If $\omega=1$, then J is of the form $C_{\ell}\times C_t$, where $\ell+t=n$, $\ell=a(a+1)$, $t=b(b+1)$, a, b nonnegative integers and $a\geq b$. Let $\delta=a-b$ and $\Sigma=a+b$. The nilpotent orbit $\mathbb{O}^{\vee}(c^{\omega}(J),\sigma)$ is parameterized by the partition

$$\lambda=(1,3,\dots,2\delta-1)\cup(1,3,\dots,2\Sigma+1),\tag{7.1.3}$$

where $\cup$ means union of partitions.

(2) If $\omega=-1$, then J is of the form $J=C_{\ell}\,{}^2A_t\,C_{\ell}$, where $2\ell+t=n-1$ and $t=\frac{a(a+1)}{2}-1$, $\ell=b(b+1)$, a, b are nonnegative integers. If $a=0,1$, we interpret J as being $J=C_{\ell}\times C_{\ell}$. Let a' be such that $a=2a'$ if a is even and $a=2a'+1$ if a is odd. Let $\Sigma=b+a'$ and

$$\delta=\begin{cases}b-a' & \text{if } 2b\geq a,\\ a'-b & \text{if } 2b<a.\end{cases}$$

The nilpotent orbit $\mathbb{O}^\vee(c^\omega(J),\sigma)$ is parameterized by the partition

$$\lambda=\begin{cases}(1,5,\dots,4\Sigma+1)\cup(3,7,\dots,4\delta-1) & \text{if } a \text{ is even and } 2b\geq a,\\(1,5,\dots,4\Sigma+1)\cup(1,5,\dots,4\delta-3) & \text{if } a \text{ is even and } 2b< a,\\(3,7,\dots,4\Sigma+3)\cup(1,5,\dots,4\delta-3) & \text{if } a \text{ is odd and } 2b\geq a,\\(3,7,\dots,4\Sigma+3)\cup(3,7,\dots,4\delta-1) & \text{if } a \text{ is odd and } 2b< a.\end{cases}$$

7.1.4. $\mathrm{PSO}(2n)$. If $\boldsymbol{G}=\mathrm{PSO}(2n)$, then $G^\vee=\mathrm{Spin}(2n,\mathbb{C})$, and

$$Z(G^\vee)=\begin{cases}(\mathbb{Z}/2\mathbb{Z})^2 & \text{if } n \text{ is even},\\ \mathbb{Z}/4\mathbb{Z} & \text{if } n \text{ is odd}.\end{cases}$$

Let τ be the standard diagram automorphism of type D_n. Let $\{1,-1\}$ be the kernel of the isogeny $\mathrm{Spin}(2n,\mathbb{C})\to\mathrm{SO}(2n,\mathbb{C})$. Write the four characters of $Z(G^\vee)$ as $\Omega=\{1,\eta,\rho,\eta\rho\}$, where $\tau(\eta)=\eta$ and $\eta(-1)=1$.

(1) If $\omega=1$, then J is of the form $D_\ell\times D_t$, where $\ell+t=n$, $\ell=a^2$, $t=b^2$, a,b even nonnegative integers, $a\geq b$. Let $\delta=a-b$, $\Sigma=a+b$. The nilpotent orbit $\mathbb{O}^\vee(c^\omega(J),\sigma)$ is parameterized by the partition

$$\lambda=(1,3,\dots,2\delta-1)\cup(1,3,\dots,2\Sigma-1). \tag{7.1.4}$$

(2) If $\omega=\eta$, then J is of the form ${}^2D_\ell\times{}^2D_t$, where $\ell+t=n$, $\ell=a^2$, $t=b^2$, and a,b are odd positive integers, $a\geq b$. The nilpotent orbit $\mathbb{O}^\vee(c^\omega(J),\sigma)$ is defined analogously to the $\omega=1$ case.

(3) If $\omega=\rho,\eta\rho$, then J can take one of the following two forms:

(i) J is of the form ${}^2A_t$, where $t=n-1$ is even, $t=\frac{a(a+1)}{2}-1$, a is a nonnegative integer. This means that $a\equiv 0,3 \pmod 4$. There are four ways to embed J into the affine Dynkin diagram $\widetilde{D}_n$, two of them are ρ-stable, and the other two $\eta\rho$-stable. In all cases the nilpotent orbit $\mathbb{O}^\vee(c^\omega(J),\sigma)$ is parameterized by the partition

$$\lambda=\begin{cases}(3,3,7,7,\dots,2a-1,2a-1) & \text{if } a\equiv 0 \pmod 4,\\(1,1,5,5,\dots,2a-1,2a-1) & \text{if } a\equiv 3 \pmod 4.\end{cases} \tag{7.1.5}$$

(ii) J is of the form $D_\ell\,{}^2A_tD_\ell$, where $2\ell+t=n-1$, $t=\frac{a(a+1)}{2}-1$ and $\ell=b^2$, a,b are nonnegative integers. Let a' be such that $a=2a'$ if a is even and $a=2a'+1$ if a is odd. Let $\Sigma=b+a'$ and

$$\delta=\begin{cases}b-a' & \text{if } 2b>a,\\ a'-b & \text{if } 2b\leq a.\end{cases}$$

The nilpotent orbit $\mathbb{O}^\vee(c^\omega(J),\sigma)$ is parameterized by the partition

$$\lambda=\begin{cases}(3,7,\dots,4\Sigma-1)\cup(1,5,\dots,4\delta-3) & \text{if } a \text{ is even and } 2b> a,\\(3,7,\dots,4\Sigma-1)\cup(3,7,\dots,4\delta-1) & \text{if } a \text{ is even and } 2b\leq a,\\(1,5,\dots,4\Sigma+1)\cup(3,7,\dots,4\delta-5) & \text{if } a \text{ is odd and } 2b> a,\\(1,5,\dots,4\Sigma+1)\cup(1,5,\dots,4\delta+1) & \text{if } a \text{ is odd and } 2b\leq a.\end{cases}$$

7.2. ***Exceptional groups.***

7.2.1. G_2. If $\boldsymbol{G} = G_2$, then $G^\vee = G_2(\mathbb{C})$, and $Z(G^\vee) = \{1\}$. If $\omega = 1$ then J is of the form G_2 and there are four choices for σ as enumerated in Table 1. In all cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = G_2(a_1).$$

7.2.2. F_4. If $\boldsymbol{G} = F_4$, then $G^\vee = F_4(\mathbb{C})$, and $Z(G^\vee) = \{1\}$. If $\omega = 1$ then J is of the form F_4 and there are seven choices for σ as enumerated in Table 1. In all cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = F_4(a_3).$$

7.2.3. E_6. If $\boldsymbol{G} = E_6$, then $G^\vee = E_6(\mathbb{C})$, and $Z(G^\vee) = \{1, \zeta, \zeta^2\}$.

(1) If $\omega = 1$ then J is of the form E_6 and there are two choices for σ as enumerated in Table 1. In both cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = D_4(a_1).$$

(2) If $\omega \in \{\zeta, \zeta^2\}$ then J is of the form ${}^3D_4$ and $\sigma = D_4[1]$ or $D_4[-1]$. In both cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = E_6(a_3).$$

7.2.4. E_7. If $\boldsymbol{G} = E_7$, then $G^\vee = E_7(\mathbb{C})$, and $Z(G^\vee) = \{1, -1\}$.

(1) If $\omega = 1$ then J is of the form E_7 and there are two choices for σ as enumerated in Table 1. In both cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = A_4 + A_1.$$

(2) If $\omega = -1$, then J is of the form ${}^2E_6$. There are three cuspidal unipotent representations afforded by J: ${}^2E_6[1]$, ${}^2E_6[\theta]$, ${}^2E_6[\theta^2]$. In all cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = E_7(a_5).$$

7.2.5. E_8. If $\boldsymbol{G} = E_8$, then $G^\vee = E_8(\mathbb{C})$, and $Z(G^\vee) = \{1\}$. If $\omega = 1$ then J is of the form E_8 and there are thirteen choices for σ as enumerated in Table 1. In all cases

$$\mathbb{O}^\vee(c^\omega(J), \sigma) = E_8(a_7).$$

8. Proof of Proposition 5.0.1

8.1. ***Classical groups.*** In each case we show that

$$\bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) = d_A(\lambda, 1),$$

where λ is the partition parameterizing $\mathbb{O}^\vee(c^\omega(J), \sigma)$. We will use the machinery of [1, Section 3.4] to prove this equality.

8.1.1. PGL(n). Let $\omega \in \Omega \simeq C_n$ be of order n. Let $J = \varnothing$. Then $\sigma = \mathrm{triv}$, $\mathrm{WF}(\sigma) = \{0\}$, and $\mathbb{O}^\vee(c^\omega(\varnothing), \sigma)$ is the principal orbit $\mathbb{O}^\vee_{\mathrm{prin}}$. We need to show that

$$\bar{\mathbb{L}}(\varnothing, \{0\}) = d_A(\mathbb{O}^\vee_{\mathrm{prin}}, 1).$$

But both sides are equal to $(\{0\}, 1)$ and so we have equality.

8.1.2. SO$(2n+1)$. Consider the cases $\omega = 1, -1$ simultaneously. Fix integers a, b as in Section 7.1.2 to fix J and hence σ. By Section 6

$$\mathrm{WF}(\sigma) = (1, 1, 3, 3, \dots, 2a-1, 2a-1) \times (1, 1, 3, 3, \dots, 2b-1, 2b-1, 2b+1).$$

Let δ, Σ, λ be as in Section 7.1.2. We have that

$$\begin{aligned} \lambda^t &= (\delta, \delta, \delta-1, \delta-1, \dots, 1, 1) \vee (\Sigma, \Sigma, \Sigma-1, \Sigma-1, \dots, 1, 1) \\ &= \begin{cases} (2b, 2b, 2b-2, 2b-2, \dots, 2a, 2a, 2a-1, 2a-1, \dots, 1, 1) & \text{if } b \geq a, \\ (2a-1, 2a-1, 2a-3, 2a-3, \dots, 2b+1, 2b+1, 2b, 2b, \dots, 1, 1) & \text{if } a > b, \end{cases} \end{aligned}$$

so $\pi(\lambda) = \varnothing$. We also have

$$d(\lambda) = (2b+1, 2b-1, 2b-1, \dots, 1, 1) \cup (2a-1, 2a-1, \dots, 1, 1).$$

Since $(1, 1, 3, 3, \dots, 2a-1, 2a-1)$ only has parts with even multiplicity,

$$\bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) = {}^{\langle(1,1,3,3,\dots,2a-1,2a-1)\rangle}d(\lambda) = {}^{\langle\varnothing\rangle}d(\lambda) = {}^{\langle\pi(\lambda)\rangle}d(\lambda) = d_A(\lambda, 1),$$

where $\pi(\lambda)$ is the subpartition of λ^t defined by Achar in [1, Equation 8].

8.1.3. PSp$(2n)$.

(1) Let $\omega = 1$. Fix integers a, b as in Section 7.1.3(1) to fix J and hence σ. By Section 6

$$\mathrm{WF}(\sigma) = (2, 2, 4, 4, \dots, 2a, 2a) \times (2, 2, 4, 4, \dots, 2b, 2b).$$

Let δ, Σ, λ be as in Section 7.1.3(1). We have that

$$\begin{aligned} \lambda^t &= (\delta, \delta-1, \delta-1, \dots, 1, 1) \vee (\Sigma+1, \Sigma, \Sigma, \dots, 1, 1) \\ &= (2a+1, 2a-1, 2a-1, \dots, 2b+1, 2b+1, 2b, 2b, \dots, 1, 1) \end{aligned}$$

so $\pi(\lambda) = \varnothing$. We also have

$$d(\lambda) = (2a, 2a, \dots, 2b+2, 2b+2, 2b, 2b, 2b, 2b, \dots, 2, 2, 2, 2).$$

Since $(2, 2, 4, 4, \dots, 2a, 2a)$ only has parts with even multiplicity,

$$\bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) = {}^{\langle(2,2,4,4,\dots,2a,2a)\rangle}d(\lambda) = {}^{\langle\varnothing\rangle}d(\lambda) = {}^{\langle\pi(\lambda)\rangle}d(\lambda) = d_A(\lambda, 1).$$

(2) Let $\omega = -1$. Fix integers a, b as in Section 7.1.3(2) to fix J and hence σ. Then

$$\mathrm{WF}(\sigma) = (2, 2, 4, 4, \ldots, 2b, 2b) \times (1, 2, \ldots, a) \times (2, 2, 4, 4, \ldots, 2b, 2b).$$

Let δ, Σ, λ be as in Section 7.1.3(2). We have that

$$\lambda^t = \begin{cases} (\Sigma+1, \Sigma^4, \ldots, 1^4) \vee (\delta^3, (\delta-1)^4, \ldots, 1^4) & \text{if } a \text{ is even and } 2b \geq a, \\ (\Sigma+1, \Sigma^4, \ldots, 1^4) \vee (\delta, (\delta-1)^4, \ldots, 1^4) & \text{if } a \text{ is even and } 2b < a, \\ ((\Sigma+1)^3, \Sigma^4, \ldots, 1^4) \vee (\delta, (\delta-1)^4, \ldots, 1^4) & \text{if } a \text{ is odd and } 2b \geq a, \\ ((\Sigma+1)^3, \Sigma^4, \ldots, 1^4) \vee (\delta^3, (\delta-1)^4, \ldots, 1^4) & \text{if } a \text{ is odd and } 2b < a, \end{cases} \tag{8.1.1}$$

$$= \begin{cases} (2b+1, (2b)^2, \ldots, (a+1)^2, a^4, \ldots, 1^4) & \text{if } a \text{ is even and } 2b \geq a, \\ (a+1, (a-1)^4, \ldots, (2b+1)^4, (2b)^4, \ldots, 1^4) & \text{if } a \text{ is even and } 2b < a, \\ (2b+1, (2b)^2, \ldots, (a+1)^2, a^4, \ldots, 1^4) & \text{if } a \text{ is odd and } 2b \geq a, \\ (a^3, (a-2)^4, \ldots, (2b+1)^4, (2b)^4, \ldots, 1^4) & \text{if } a \text{ is odd and } 2b < a. \end{cases} \tag{8.1.2}$$

Thus $\pi(\lambda) = \varnothing$ since all even parts of λ^t have even multiplicity. Moreover

$$d(\lambda) = (2, 2, 4, 4, \ldots, 2b, 2b) \cup (1, 1, 2, 2, \ldots, a, a) \cup (2, 2, 4, 4, \ldots, 2b, 2b)$$

in all cases. Thus

$$\begin{aligned} \bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) &= \bar{\mathbb{L}}(\tilde{J}, (2, 2, \ldots, 2b, 2b) \times (1, 1, \ldots, a, a) \cup (2, 2, \ldots, 2b, 2b)) \\ &= {}^{\langle (2,2,4,4,\ldots,2b,2b) \rangle} d(\lambda) = {}^{\langle \varnothing \rangle} d(\lambda) = {}^{\langle \pi(\lambda) \rangle} d(\lambda) = d_A(\lambda, 1), \end{aligned}$$

where $\tilde{J} = C_l \times C_{t+1+l}$.

8.1.4. $\mathrm{PSO}(2n)$.

(1) Let $\omega \in \{1, \eta\}$. Fix integers a, b as in Section 7.1.4(1) and (2) to fix J and hence σ. By Section 6

$$\mathrm{WF}(\sigma) = (1, 1, 3, 3, \ldots, 2a-1, 2a-1) \times (1, 1, 3, 3, \ldots, 2b-1, 2b-1).$$

Let δ, Σ, λ be as in Section 7.1.4(1). We have that

$$\begin{aligned} \lambda^t &= (\delta, \delta-1, \delta-1, \ldots, 1, 1) \vee (\Sigma, \Sigma-1, \Sigma-1, \ldots, 1, 1) \\ &= (2a, 2a-2, 2a-2, \ldots, 2b, 2b, 2b-1, 2b-1, \ldots, 1, 1) \end{aligned}$$

so $\pi(\lambda) = \varnothing$ since all odd parts have even multiplicity. We also have

$$d(\lambda) = (2a-1, 2a-1, \ldots, 2b+1, 2b+1, 2b-1, 2b-1, 2b-1, 2b-1, \ldots, 1, 1, 1, 1).$$

Since $(1, 1, 3, 3, \ldots, 2a-1, 2a-1)$ only has parts with even multiplicity,

$$\bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) = {}^{\langle (1,1,3,3,\ldots,2a-1,2a-1) \rangle} d(\lambda) = {}^{\langle \varnothing \rangle} d(\lambda) = {}^{\pi(\lambda)} d(\lambda) = d_A(\lambda, 1).$$

(2) Let $\omega \in \{\rho, \eta\rho\}$. We will treat the cases (i) and (ii) simultaneously. Fix integers a, b as in Section 7.1.4(3)(ii) to fix J and hence σ (we treat (i) as the case with $b = 0$). By Section 6

$$\mathrm{WF}(\sigma) = (1, 1, 3, 3, \ldots, 2b-1, 2b-1) \times (1, 2, \ldots, a) \times (1, 1, 3, 3, \ldots, 2b-1, 2b-1).$$

Let δ, Σ, λ be as in Section 7.1.4(3)(ii). We have that

$$\lambda^t = \begin{cases} (\Sigma^3, (\Sigma-1)^4, \dots, 1^4) \vee (\delta, (\delta-1)^4, \dots, 1^4) & \text{if } a \text{ is even and } 2b \le a, \\ (\Sigma^3, (\Sigma-1)^4, \dots, 1^4) \vee (\delta^3, (\delta-1)^4, \dots, 1^4) & \text{if } a \text{ is even and } 2b \le a, \\ (\Sigma+1, \Sigma^4, \dots, 1^4) \vee (\delta, (\delta-1)^4, \dots, 1^4) & \text{if } a \text{ is odd and } 2b < a, \\ ((\Sigma+1)^3, \Sigma^4, \dots, 1^4) \vee (\delta^3, (\delta-1)^4, \dots, 1^4) & \text{if } a \text{ is odd and } 2b < a, \end{cases} \tag{8.1.3}$$

$$= \begin{cases} (2b+1, (2b)^2, \dots, (a+1)^2, a^4, \dots, 1^4) & \text{if } a \text{ is even and } 2b \ge a, \\ (a+1, (a-1)^4, \dots, (2b+1)^4, (2b)^4, \dots, 1^4) & \text{if } a \text{ is even and } 2b < a, \\ (2b+1, (2b)^2, \dots, (a+1)^2, a^4, \dots, 1^4) & \text{if } a \text{ is odd and } 2b \ge a, \\ (a^3, (a-2)^4, \dots, (2b+1)^4, (2b)^4, \dots, 1^4) & \text{if } a \text{ is odd and } 2b < a. \end{cases} \tag{8.1.4}$$

Thus $\pi(\lambda) = \varnothing$ since all even parts of λ^t have even multiplicity. Moreover

$$d(\lambda) = (2, 2, 4, 4, \dots, 2b, 2b) \cup (1, 1, 2, 2, \dots, a, a) \cup (2, 2, 4, 4, \dots, 2b, 2b)$$

in all cases. Thus

$$\begin{aligned} \bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) &= \bar{\mathbb{L}}(\tilde{J}, (1, 1, \dots, 2b-1, 2b-1) \times (1, 1, \dots, a, a) \cup (1, 1, \dots, 2b-1, 2b-1)) \\ &= {}^{\langle (2,2,4,4,\dots,2b,2b) \rangle} d(\lambda) = {}^{\langle \varnothing \rangle} d(\lambda) = {}^{\langle \pi(\lambda) \rangle} d(\lambda) = d_A(\lambda, 1), \end{aligned}$$

where $\tilde{J} = D_l \times D_{t+1+l}$.

8.2. *Exceptional groups.*

8.2.1. *Split forms.* Suppose that $\boldsymbol{G}$ is split, of exceptional type, and that $\omega = 1$. As can be seen in Section 7.2, J is always equal to Δ. Thus,

$$\bar{\mathbb{L}}(J, \mathrm{WF}(\sigma)) = (\mathrm{WF}(\sigma), 1).$$

On the other hand, the nilpotent orbit $\mathbb{O}^\vee := \mathbb{O}^\vee(c^\omega(J), \sigma)$ is always special. Thus,

$$d_A(\mathbb{O}^\vee, 1) = (d(\mathbb{O}^\vee), 1)$$

by the general properties of d_A; see [1, Section 3]. So for Proposition 5.0.1 it suffices to show that

$$\mathrm{WF}(\sigma) = d(\mathbb{O}^\vee)$$

for all σ. This follows by comparing the orbits in Table 1 and in Section 7.2.

8.2.2. *Nonsplit forms of E_6.* Suppose $\boldsymbol{G}$ is of type E_6 and $\omega \in \{\zeta, \zeta^2\}$. Then J is of the form ${}^3D_4$, and $\mathrm{WF}(\sigma) = (1, 1, 3, 3)$ for both $\sigma = D_4[1]$ and $\sigma = D_4[-1]$. The orbit $(1, 1, 3, 3)$ is the orbit A_2 in Bala–Carter notation. Thus we need to show that

$$\bar{\mathbb{L}}(J, (1, 1, 3, 3)) = d_A(E_6(a_3), 1).$$

We note that $E_6(a_3)$ is special and $d(E_6(a_3)) = A_2$ so we must show that

$$\bar{\mathbb{L}}(J, A_2) = (A_2, 1).$$

Since $J \subseteq \Delta$ this follows from [27, Proposition 2.30].

8.2.3. *Nonsplit forms of E_7.* Suppose $\boldsymbol{G}$ is of type E_7 and $\omega = -1$. Then J is of the form ${}^2E_6$, and $\mathrm{WF}(\sigma) = D_4(a_1)$ for all possible σ. Thus we need to show that

$$\bar{\mathbb{L}}(J, D_4(a_1)) = d_A(E_7(a_5), 1).$$

We note that $E_7(a_5)$ is special and $d(E_7(a_5)) = D_4(a_1)$ so we must show that

$$\bar{\mathbb{L}}(J, D_4(a_1)) = (D_4(a_1), 1).$$

Since $J \subseteq \Delta$ this follows from [27, Proposition 2.30].

References

[1] P. N. Achar, "An order-reversing duality map for conjugacy classes in Lusztig's canonical quotient", *Transform. Groups* **8**:2 (2003), 107–145. MR Zbl

[2] A.-M. Aubert, "Dualité dans le groupe de Grothendieck de la catégorie des représentations lisses de longueur finie d'un groupe réductif p-adique", *Trans. Amer. Math. Soc.* **347**:6 (1995), 2179–2189. MR Zbl

[3] A.-M. Aubert, P. Baum, R. Plymen, and M. Solleveld, "Conjectures about p-adic groups and their noncommutative geometry", pp. 15–51 in *Around Langlands correspondences*, edited by F. Brumley et al., Contemp. Math. **691**, Amer. Math. Soc., Providence, RI, 2017. MR Zbl

[4] A.-M. Aubert, A. Moussaoui, and M. Solleveld, "Generalizations of the Springer correspondence and cuspidal Langlands parameters", *Manuscripta Math.* **157**:1-2 (2018), 121–192. MR Zbl

[5] D. Barbasch and A. Moy, "Local character expansions", *Ann. Sci. École Norm. Sup.* (4) **30**:5 (1997), 553–567. MR Zbl

[6] D. Barbasch and D. A. Vogan, Jr., "Unipotent representations of complex semisimple groups", *Ann. of Math.* (2) **121**:1 (1985), 41–110. MR Zbl

[7] R. W. Carter, *Finite groups of Lie type*: *conjugacy classes and complex characters*, Wiley, Chichester, 1993. MR Zbl

[8] D. Ciubotaru and E. Okada, "Local character expansions via positive depth Barbasch–Moy theory", preprint, 2023. arXiv 2307.06780

[9] D. Ciubotaru, L. Mason-Brown, and E. Okada, "Wavefront sets of unipotent representations of reductive p-adic groups, II", preprint, 2023. arXiv 2303.10713

[10] D. Ciubotaru, L. Mason-Brown, and E. Okada, "Some unipotent Arthur packets for reductive p-adic groups", *Int. Math. Res. Not.* **2024**:9 (2024), 7502–7525. MR Zbl

[11] S. Evens and I. Mirković, "Fourier transform and the Iwahori–Matsumoto involution", *Duke Math. J.* **86**:3 (1997), 435–464. MR Zbl

[12] Y. Feng and E. Opdam, "On a uniqueness property of supercuspidal unipotent representations", *Adv. Math.* **375** (2020), art. id. 107406. MR Zbl

[13] Y. Feng, E. Opdam, and M. Solleveld, "On formal degrees of unipotent representations", *J. Inst. Math. Jussieu* **21**:6 (2022), 1947–1999. MR Zbl

[14] Harish-Chandra, *Admissible invariant distributions on reductive p-adic groups*, University Lecture Series **16**, American Mathematical Society, Providence, RI, 1999. MR Zbl

[15] K. Hiraga, A. Ichino, and T. Ikeda, "Formal degrees and adjoint γ-factors", *J. Amer. Math. Soc.* **21**:1 (2008), 283–304. MR Zbl

[16] R. Howe, "The Fourier transform and germs of characters (case of Gl_n over a p-adic field)", *Math. Ann.* **208** (1974), 305–322. MR Zbl

[17] T. Kaletha, "The local Langlands conjectures for non-quasi-split groups", pp. 217–257 in *Families of automorphic forms and the trace formula*, edited by W. Müller et al., Springer, 2016. MR Zbl

[18] N. Kawanaka, "Shintani lifting and Gel'fand–Graev representations", pp. 147–163 in *The Arcata Conference on Representations of Finite Groups* (Arcata, CA, 1986), edited by P. Fong, Proc. Sympos. Pure Math. **47, Part 1**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[19] R. E. Kottwitz, "Stable trace formula: cuspidal tempered terms", *Duke Math. J.* **51**:3 (1984), 611–650. MR Zbl

[20] G. Lusztig, *Characters of reductive groups over a finite field*, Annals of Mathematics Studies **107**, Princeton University Press, 1984. MR Zbl

[21] G. Lusztig, "A unipotent support for irreducible representations", *Adv. Math.* **94**:2 (1992), 139–179. MR Zbl

[22] G. Lusztig, "Classification of unipotent representations of simple p-adic groups", *Internat. Math. Res. Notices* 11 (1995), 517–589. MR Zbl

[23] C. Mœglin and J.-L. Waldspurger, "Modèles de Whittaker dégénérés pour des groupes p-adiques", *Math. Z.* **196**:3 (1987), 427–452. MR Zbl

[24] L. Morris, "Tamely ramified supercuspidal representations", *Ann. Sci. École Norm. Sup.* (4) **29**:5 (1996), 639–667. MR Zbl

[25] A. Moy and G. Prasad, "Unrefined minimal K-types for p-adic groups", *Invent. Math.* **116**:1-3 (1994), 393–408. MR Zbl

[26] A. Moy and G. Prasad, "Jacquet functors and unrefined minimal K-types", *Comment. Math. Helv.* **71**:1 (1996), 98–121. MR Zbl

[27] E. Okada, "The wavefront set over a maximal unramified field extension", preprint, 2021. arXiv 2107.10591

[28] E. Opdam, "Spectral transfer morphisms for unipotent affine Hecke algebras", *Selecta Math. (N.S.)* **22**:4 (2016), 2143–2207. MR Zbl

[29] K. Pommerening, "Über die unipotenten Klassen reduktiver Gruppen", *J. Algebra* **49**:2 (1977), 525–536. MR Zbl

[30] K. Pommerening, "Über die unipotenten Klassen reduktiver Gruppen, II", *J. Algebra* **65**:2 (1980), 373–398. MR Zbl

[31] R. Ranga Rao, "Orbital integrals in reductive groups", *Ann. of Math.* (2) **96** (1972), 505–510. MR Zbl

[32] M. Reeder, "Formal degrees and L-packets of unipotent discrete series representations of exceptional p-adic groups", *J. Reine Angew. Math.* **520** (2000), 37–93. MR Zbl

[33] E. Sommers, "Lusztig's canonical quotient and generalized duality", *J. Algebra* **243**:2 (2001), 790–812. MR Zbl

[34] N. Spaltenstein, *Classes unipotentes et sous-groupes de Borel*, Lecture Notes in Mathematics **946**, Springer, 1982. MR Zbl

[35] J. Tits, "Reductive groups over local fields", pp. 29–69 in *Automorphic forms, representations and L-functions*, *I* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR Zbl

[36] D. A. Vogan, Jr., "The local Langlands conjecture", pp. 305–379 in *Representation theory of groups and algebras*, edited by J. Adams et al., Contemp. Math. **145**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

Communicated by Akshay Venkatesh
Received 2022-07-11 Revised 2023-07-27 Accepted 2023-10-24

dan.ciubotaru@maths.ox.ac.uk *Mathematical Institute, University of Oxford, Oxford, United Kingdom*

lucas.mason-brown@maths.ox.ac.uk *Mathematical Institute, University of Oxford, Oxford, United Kingdom*

okada@nus.edu.sg *Department of Mathematics, National University of Singapore, Singapore*

ALGEBRA AND NUMBER THEORY 18:10 (2024)
https://doi.org/10.2140/ant.2024.18.1891

A geometric classification of the holomorphic vertex operator algebras of central charge 24

Sven Möller and Nils R. Scheithauer

We associate with a generalised deep hole of the Leech lattice vertex operator algebra a generalised hole diagram. We show that this Dynkin diagram determines the generalised deep hole up to conjugacy and that there are exactly 70 such diagrams. In an earlier work we proved a bijection between the generalised deep holes and the strongly rational, holomorphic vertex operator algebras of central charge 24 with nontrivial weight-1 space. Hence, we obtain a new, geometric classification of these vertex operator algebras, generalising the classification of the Niemeier lattices by their hole diagrams.

1. Introduction

In 1968 Niemeier classified the positive-definite, even, unimodular lattices of rank 24 [54]. He showed that up to isomorphism there are exactly 24 such lattices and that the isomorphism class of each lattice is uniquely determined by its root system. The Leech lattice Λ is the unique Niemeier lattice without roots. There are at least five proofs of this classification result. Niemeier applied Kneser's neighbourhood method. Venkov found a proof based on harmonic theta series [57]. It can also be derived from Conway, Parker and Sloane's classification of the deep holes of the Leech lattice [3; 12] and from the Smith–Minkowski–Siegel mass formula [9; 11]. Finally, it also follows from the classification of certain automorphic representations of O_{24} [7].

We describe the third proof in more detail. Borcherds [3] showed that the Leech lattice Λ is the unique Niemeier lattice without roots (see also [8]) and that the orbits of deep holes of Λ, i.e., points in $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$ that have maximal distance to Λ, are in natural bijection with the other Niemeier lattices. These results

MSC2020: primary 17B69; secondary 17B22.
Keywords: conformal field theory, vertex operator algebra, Leech lattice, Schellekens' list, deep hole, generalised deep hole, central charge 24.

are proved without explicitly classifying the deep holes or the Niemeier lattices. Conway, Parker and Sloane [12] associate with a deep hole in Λ a hole diagram, a certain affine Dynkin diagram whose vertices are the closest lattice points, and classify the possible diagrams by geometric methods. They find 23 diagrams and show that a deep hole is fixed up to equivalence by its hole diagram. This implies that there are exactly 23 Niemeier lattices with roots. We generalise this approach to strongly rational, holomorphic vertex operator algebras of central charge 24.

Vertex operator algebras and their representations axiomatise 2-dimensional conformal field theories [4; 29]. They have found various applications in mathematics and mathematical physics, e.g., in geometry, group theory and the theory of automorphic forms. The theory of these algebras is in certain aspects similar to the theory of even lattices over the integers.

The weight-1 subspace V_1 of a strongly rational, holomorphic vertex operator algebra V of central charge 24 is a reductive Lie algebra. In 1993 Schellekens [55] (see also [25]) showed that there are at most 71 possibilities for this Lie algebra using the theory of Jacobi forms. He conjectured that all potential Lie algebras are realised and that the V_1-structure fixes the vertex operator algebra up to isomorphism. By the work of many authors over the past three decades the following result is now proved (see, e.g., [37; 40]).

Theorem. *Up to isomorphism there are exactly* 70 *strongly rational, holomorphic vertex operator algebras V of central charge* 24 *with* $V_1 \neq \{0\}$. *Such a vertex operator algebra is uniquely determined by its V_1-structure.*

The proof is based on a case-by-case analysis and uses a variety of methods.

The 24 vertex operator algebras V_N associated with the Niemeier lattices N are examples of vertex operator algebras on Schellekens' list.

We give a uniform, geometric proof of the theorem based on the results in [50], which generalises the classification of the Niemeier lattices by enumeration of the corresponding deep holes of the Leech lattice Λ [3; 12].

One method to construct vertex operator algebras is the cyclic orbifold construction [25]. Let V be a strongly rational, holomorphic vertex operator algebra and g an automorphism of V of finite order n. Then the fixed-point subalgebra V^g is a strongly rational vertex operator algebra with n^2 nonisomorphic irreducible modules, which can be realised as the eigenspaces of g acting on the unique irreducible twisted modules $V(g^i)$ of V. If the twisted modules $V(g^i)$ have conformal weights in $(1/n)\mathbb{Z}_{>0}$ for $i \neq 0 \bmod n$, then the direct sum $V^{\mathrm{orb}(g)} := \bigoplus_{i\in\mathbb{Z}_n} V(g^i)^g$ is again a strongly rational, holomorphic vertex operator algebra of the same central charge as V. There is also an inverse orbifold construction, that is, an automorphism h of $V^{\mathrm{orb}(g)}$ such that $(V^{\mathrm{orb}(g)})^{\mathrm{orb}(h)} = V$.

Suppose that V has central charge 24 and that $n > 1$. Pairing the character of V^g with a certain vector-valued Eisenstein series of weight 2 we obtain [50]:

Theorem (dimension formula). *The dimension of the weight-*1 *subspace of $V^{\mathrm{orb}(g)}$ is given by*

$$\dim(V_1^{\mathrm{orb}(g)}) = 24 + \sum_{d\mid n} c_n(d)\dim(V_1^{g^d}) - R(g),$$

where the $c_n(d) \in \mathbb{Q}$ *are defined by* $\sum_{d|n} c_n(d)(t,d) = n/t$ *for all* $t \mid n$ *and the remainder term* $R(g)$ *is nonnegative. In particular,*

$$\dim(V_1^{\mathrm{orb}(g)}) \leq 24 + \sum_{d|n} c_n(d) \dim(V_1^{g^d}).$$

The remainder term $R(g)$ is described explicitly. It depends on the dimensions of the weight spaces of the irreducible V^g-modules of weight less than 1.

The upper bound in the theorem motivates the following definition. The automorphism g is called a *generalised deep hole* of V if

(1) the upper bound in the dimension formula is attained, i.e., $\dim(V_1^{\mathrm{orb}(g)}) = 24 + \sum_{d|n} c_n(d) \dim(V_1^{g^d})$,

(2) any Cartan subalgebra of V_1^g is also a Cartan subalgebra of $V_1^{\mathrm{orb}(g)}$, i.e., $\mathrm{rk}(V_1^{\mathrm{orb}(g)}) = \mathrm{rk}(V_1^g)$.

We also call the identity automorphism a generalised deep hole.

Let V_Λ be the vertex operator algebra of the Leech lattice Λ. Recall that algebraic conjugacy means conjugacy of cyclic subgroups. An averaged version of Kac's very strange formula implies [50] (cf. [27]):

Theorem (holey correspondence). *The orbifold construction* $g \mapsto V_\Lambda^{\mathrm{orb}(g)}$ *defines a bijection between the algebraic conjugacy classes of generalised deep holes* g *in* $\mathrm{Aut}(V_\Lambda)$ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$ *and the isomorphism classes of strongly rational, holomorphic vertex operator algebras* V *of central charge* 24 *with* $V_1 \neq \{0\}$.

Let $g \in \mathrm{Aut}(V_\Lambda)$ be a generalised deep hole. Then $\mathfrak{h} := (V_\Lambda^g)_1$ is a Cartan subalgebra of $(V_\Lambda^{\mathrm{orb}(g)})_1$. It acts on $(V_\Lambda(g))_1$. The corresponding weights form a Dynkin diagram, which we denote by $\Phi(g)$. Then the *generalised hole diagram* of g is defined as the pair $(\varphi(g), \Phi(g))$, where $\varphi(g)$ denotes the cycle shape of the image of g under the natural projection $\mathrm{Aut}(V_\Lambda) \to O(\Lambda)$. For example, if $V_\Lambda^{\mathrm{orb}(g)}$ is isomorphic to the vertex operator algebra V_N of the Niemeier lattice with Dynkin diagram N, then the generalised hole diagram of g is $(1^{24}, \tilde{N})$ where $\tilde{N}$ is the affine Dynkin diagram corresponding to N, that is, the hole diagram of the Niemeier lattice inside the Leech lattice Λ.

Our main result is the following (see Theorem 5.25):

Theorem (classification of generalised deep holes). *There are exactly* 70 *conjugacy classes of generalised deep holes* g *in* $\mathrm{Aut}(V_\Lambda)$ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$*. The class of a generalised deep hole is uniquely determined by its generalised hole diagram.*

We outline the proof. The holey correspondence together with the lowest-order trace identity (see (S) in Section 3.1) imply that there are at most 82 possible generalised hole diagrams. These are described in Tables 1 and 2. Then, using geometric arguments similar to those by Conway, Parker and Sloane [12] we reduce this number to 70. On the other hand, in [50] we explicitly list 70 generalised deep holes with distinct diagrams. It follows that these are exactly the generalised deep holes g of V_Λ with $\mathrm{rk}((V_\Lambda^g)_1) > 0$.

We observe (see Theorem 5.27):

Theorem (projection to Co_0). *Under the natural projection* $\mathrm{Aut}(V_\Lambda) \to O(\Lambda)$ *the* 70 *conjugacy classes of generalised deep holes* g *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$ *map to the* 11 *conjugacy classes in* $O(\Lambda) \cong \mathrm{Co}_0$ *with cycle shapes* 1^{24}, $1^8 2^8$, $1^6 3^6$, 2^{12}, $1^4 2^2 4^4$, $1^4 5^4$, $1^2 2^2 3^2 6^2$, $1^3 7^3$, $1^2 2^1 4^1 8^2$, $2^3 6^3$ *and* $2^2 10^2$.

This recovers the decomposition of the Schellekens vertex operator algebras into 12 families described by Höhn in [32] (cf. [34] in the fermionic case). The precise connection is explored in [33, Section 4.2].

A consequence of the classification of generalised deep holes is:

Theorem (classification of vertex operator algebras). *Up to isomorphism there are exactly* 70 *strongly rational, holomorphic vertex operator algebras* V *of central charge* 24 *with* $V_1 \neq \{0\}$. *Such a vertex operator algebra is uniquely determined by its* V_1*-structure.*

In contrast to the previous proof, our argument is uniform and, except for the lowest-order trace identity, independent of Schellekens' results.

Outline. In Section 2 we describe the orbifold construction, lattice vertex operator algebras and the automorphisms of the Leech lattice vertex operator algebra.

In Section 3 we recall some results on strongly rational, holomorphic vertex operator algebras of central charge 24, in particular the bijection with the generalised deep holes of the Leech lattice vertex operator algebra.

In Section 4 we associate a generalised hole diagram with a generalised deep hole of the Leech lattice vertex operator algebra.

In Section 5 we finally use the generalised hole diagrams to classify the generalised deep holes of the Leech lattice vertex operator algebra.

2. Vertex operator algebras and their automorphisms

We review the cyclic orbifold construction and describe the automorphisms of the Leech lattice vertex operator algebra V_Λ.

A vertex operator algebra V is called *strongly rational* if it is rational (as defined, e.g., in [21]), C_2-cofinite (or lisse), self-contragredient (or self-dual) and of CFT-type. This already implies that V is simple.

A simple vertex operator algebra V is said to be *holomorphic* if V itself is the only irreducible V-module. The central charge of a strongly rational, holomorphic vertex operator algebra V is necessarily a nonnegative multiple of 8.

Examples of strongly rational vertex operator algebras are those associated with positive-definite, even lattices. If the lattice is unimodular, then the associated vertex operator algebra is holomorphic.

2.1. ***Orbifold construction.*** The cyclic orbifold construction [25; 48] is an important tool that can be used to construct new vertex operator algebras from known ones.

Let V be a strongly rational, holomorphic vertex operator algebra and $G = \langle g \rangle \cong \mathbb{Z}_n$ a finite, cyclic group of automorphisms of V of order n.

By [22] there is an up to isomorphism unique irreducible g^i-twisted V-module $V(g^i)$ for each $i \in \mathbb{Z}_n$. The uniqueness of $V(g^i)$ implies that there is a representation $\phi_i : G \to \mathrm{Aut}_{\mathbb{C}}(V(g^i))$ of G on the vector space $V(g^i)$ such that

$$\phi_i(g)Y_{V(g^i)}(v,x)\phi_i(g)^{-1} = Y_{V(g^i)}(gv,x)$$

for all $v \in V$, $i \in \mathbb{Z}_n$. This representation is unique up to an n-th root of unity. Denote the eigenspace of $\phi_i(g)$ in $V(g^i)$ corresponding to the eigenvalue $\mathrm{e}^{2\pi \mathrm{i} j/n}$ by $W^{(i,j)}$. On $V(g^0) = V$ we choose $\phi_0(g) = g$.

By [6; 16; 45; 46] the fixed-point vertex operator subalgebra $V^g = W^{(0,0)}$ is again strongly rational. It has exactly n^2 irreducible modules, namely the $W^{(i,j)}$, $i, j \in \mathbb{Z}_n$ [23; 47]. One can further show that the conformal weight $\rho(V(g))$ of $V(g)$ is in $(1/n^2)\mathbb{Z}$, and we define the type $t \in \mathbb{Z}_n$ of g by $t = n^2\rho(V(g)) \bmod n$.

Assume for simplicity that g has type 0, i.e., that $\rho(V(g)) \in (1/n)\mathbb{Z}$. Then it is possible to choose the representations ϕ_i such that the conformal weights satisfy

$$\rho(W^{(i,j)}) = \frac{ij}{n} \mod 1$$

and V^g has fusion rules

$$W^{(i,j)} \boxtimes W^{(k,l)} \cong W^{(i+k,j+l)}$$

for all $i, j, k, l \in \mathbb{Z}_n$, that is, the fusion ring of V^g is the group ring $\mathbb{C}[\mathbb{Z}_n \times \mathbb{Z}_n]$ [25]. In particular, all irreducible V^g-modules are simple currents.

In essence, the results in [25] show that for cyclic $G \cong \mathbb{Z}_n$ and strongly rational, holomorphic V the module category of V^G is the *twisted group double* $\mathcal{D}_\omega(G)$ where the 3-cocycle $[\omega] \in H^3(G, \mathbb{C}^\times) \cong \mathbb{Z}_n$ is determined by the type $t \in \mathbb{Z}_n$. This proves a special case of a conjecture in [13; 14] stated for arbitrary finite G. The general case is proved in [24].

In general, a simple vertex operator algebra V is said to satisfy the *positivity condition* if the conformal weight $\rho(W)$ is greater than 0 for any irreducible V-module $W \not\cong V$ and $\rho(V) = 0$.

Now, if V^g satisfies the positivity condition (it is shown in [49] that this condition is almost automatically satisfied if V is strongly rational) and g has type 0, then the direct sum of V^g-modules

$$V^{\mathrm{orb}(g)} := \bigoplus_{i \in \mathbb{Z}_n} W^{(i,0)}$$

admits the structure of a strongly rational, holomorphic vertex operator algebra of the same central charge as V and is called *orbifold construction* associated with V and g [25]. Note that $\bigoplus_{j \in \mathbb{Z}_n} W^{(0,j)}$ gives back the original vertex operator algebra V.

We briefly describe the *inverse* (or *reverse*) orbifold construction [25; 39]. Suppose that the strongly rational, holomorphic vertex operator algebra $V^{\mathrm{orb}(g)}$ is obtained by an orbifold construction as described above. Then via $\zeta v := \mathrm{e}^{2\pi \mathrm{i} j/n} v$ for $v \in W^{(j,0)}$ we define an automorphism ζ of $V^{\mathrm{orb}(g)}$ of order n and type 0, and the unique irreducible ζ^j-twisted $V^{\mathrm{orb}(g)}$-module is given by $V^{\mathrm{orb}(g)}(\zeta^j) = \bigoplus_{i \in \mathbb{Z}_n} W^{(i,j)}$,

$j \in \mathbb{Z}_n$. Then

$$(V^{\mathrm{orb}(g)})^{\mathrm{orb}(\zeta)} = \bigoplus_{j \in \mathbb{Z}_n} W^{(0,j)} = V,$$

i.e., orbifolding with ζ is inverse to orbifolding with g.

2.2. ***Automorphisms of the Leech lattice vertex operator algebra.*** We describe lattice vertex operator algebras [4; 29], the automorphism group of the Leech lattice vertex operator algebra V_Λ and in particular its conjugacy classes, which were determined in [50].

For a positive-definite, even lattice L with bilinear form $\langle\cdot,\cdot\rangle\colon L \times L \to \mathbb{Z}$ the associated vertex operator algebra is given by

$$V_L = M(1) \otimes \mathbb{C}_\varepsilon[L],$$

where $M(1)$ is the Heisenberg vertex operator algebra of rank $\mathrm{rk}(L)$ associated with $\mathfrak{h}_L = L \otimes_{\mathbb{Z}} \mathbb{C}$ and $\mathbb{C}_\varepsilon[L]$ the twisted group algebra, that is, the algebra with basis $\{\mathfrak{e}_\alpha \mid \alpha \in L\}$ and products $\mathfrak{e}_\alpha \mathfrak{e}_\beta = \varepsilon(\alpha,\beta)\mathfrak{e}_{\alpha+\beta}$ where $\varepsilon\colon L \times L \to \{\pm 1\}$ is a 2-cocycle satisfying $\varepsilon(\alpha,\beta)/\varepsilon(\beta,\alpha) = (-1)^{\langle\alpha,\beta\rangle}$.

Let $O(L)$ denote the orthogonal group (or automorphism group) of the lattice L. For $\nu \in O(L)$ and a function $\eta\colon L \to \{\pm 1\}$ the map $\phi_\eta(\nu)$ acting on $\mathbb{C}_\varepsilon[L]$ as $\phi_\eta(\nu)(\mathfrak{e}_\alpha) = \eta(\alpha)\mathfrak{e}_{\nu\alpha}$ for $\alpha \in L$ and as ν on $M(1)$ defines an automorphism of V_L if and only if

$$\frac{\eta(\alpha)\eta(\beta)}{\eta(\alpha+\beta)} = \frac{\varepsilon(\alpha,\beta)}{\varepsilon(\nu\alpha,\nu\beta)}$$

for all $\alpha, \beta \in L$. In this case, $\phi_\eta(\nu)$ is called a *lift* of ν and all such automorphisms form the subgroup $O(\hat{L})$ of $\mathrm{Aut}(V_L)$. There is a short exact sequence

$$1 \to \mathrm{Hom}(L,\{\pm 1\}) \to O(\hat{L}) \to O(L) \to 1$$

with the surjection $O(\hat{L}) \to O(L)$ given by $\phi_\eta(\nu) \mapsto \nu$. The image of $\mathrm{Hom}(L,\{\pm 1\})$ in $O(\hat{L})$ is exactly the lifts of $\mathrm{id} \in O(L)$.

If the restriction of η to the fixed-point lattice L^ν is trivial, we call $\phi_\eta(\nu)$ a *standard lift* of ν. It is always possible to choose η in this way [42]. It was proved in [25] that all standard lifts of a given $\nu \in O(L)$ are conjugate in $\mathrm{Aut}(V_L)$.

For any vertex operator algebra V, $K := \langle\{\mathrm{e}^{v_0} \mid v \in V_1\}\rangle$ defines a normal subgroup of $\mathrm{Aut}(V)$ called the *inner automorphism group* of V. By [20] the automorphism group of V_L is of the form

$$\mathrm{Aut}(V_L) = O(\hat{L}) \cdot K,$$

$\mathrm{Hom}(L,\{\pm 1\})$ is a subgroup of $K \cap O(\hat{L})$ and $\mathrm{Aut}(V_L)/K$ isomorphic to a quotient of $O(L)$.

In the following, we specialise to the Leech lattice Λ, the up to isomorphism unique unimodular, positive-definite, even lattice of rank 24 without roots, i.e., without vectors of norm 2. The automorphism group $O(\Lambda)$ is Conway's group Co_0. Since $(V_\Lambda)_1 = \{h(-1) \otimes \mathfrak{e}_0 \mid h \in \mathfrak{h}_\Lambda\} \cong \mathfrak{h}_\Lambda$ with $\mathfrak{h}_\Lambda = \Lambda \otimes_{\mathbb{Z}} \mathbb{C}$, the inner automorphism group is given by

$$K = \{\mathrm{e}^{h_0} \mid h \in \mathfrak{h}_\Lambda\}$$

and is abelian. Because $K \cap O(\hat{\Lambda}) = \operatorname{Hom}(\Lambda, \{\pm 1\})$ in the special case of the Leech lattice, there is a short exact sequence

$$1 \to K \to \operatorname{Aut}(V_\Lambda) \to O(\Lambda) \to 1.$$

Hence, every automorphism of V_Λ is of the form

$$\phi_\eta(\nu)\sigma_h$$

with a lift $\phi_\eta(\nu)$ of some $\nu \in O(\Lambda)$ and with $\sigma_h = \mathrm{e}^{2\pi i h_0}$ for some $h \in \mathfrak{h}_\Lambda$. The surjection $\operatorname{Aut}(V_\Lambda) \to O(\Lambda)$ in the short exact sequence is given by $\phi_\eta(\nu)\sigma_h \mapsto \nu$. It suffices to take a standard lift $\phi_\eta(\nu)$ of ν because any two lifts of ν only differ by a homomorphism $\Lambda \to \{\pm 1\}$, which can be absorbed into σ_h. Moreover, since $\sigma_h = \mathrm{id}$ if and only if $h \in \Lambda' = \Lambda$, it is enough to take $h \in \mathfrak{h}_\Lambda/\Lambda$.

We describe the conjugacy classes of $\operatorname{Aut}(V_\Lambda)$. For $\nu \in O(\Lambda)$ let

$$\pi_\nu = \frac{1}{|\nu|}\sum_{i=0}^{|\nu|-1} \nu^i$$

denote the projection from $\mathfrak{h}_\Lambda$ onto the elements of $\mathfrak{h}_\Lambda$ fixed by ν. The automorphism $\phi_\eta(\nu)\sigma_h$ is conjugate to $\phi_\eta(\nu)\sigma_{\pi_\nu(h)}$ for any $h \in \mathfrak{h}_\Lambda$, and $\phi_\eta(\nu)$ and $\sigma_{\pi_\nu(h)}$ commute.

In [50] all automorphisms in $\operatorname{Aut}(V_\Lambda)$ were classified up to conjugacy. A similar result for arbitrary lattice vertex operator algebras was proved in [33].

Proposition 2.1 [50]. *Let $Q := \{(\nu, h) \mid \nu \in N,\ h \in H_\nu\}$ where*

(1) *N is a set of representatives for the conjugacy classes in $O(\Lambda)$,*

(2) *H_ν is a set of representatives for the orbits of the action of the centraliser $C_{O(\Lambda)}(\nu)$ on $\pi_\nu(\mathfrak{h}_\Lambda)/\pi_\nu(\Lambda)$.*

Fix a section $\nu \mapsto \phi_\eta(\nu)$. Then the map $(\nu, h) \mapsto \phi_\eta(\nu)\sigma_h$ is a bijection from the set Q to the conjugacy classes of $\operatorname{Aut}(V_\Lambda)$.

Since $h \in \pi_\nu(\mathfrak{h}_\Lambda)$, $\phi_\eta(\nu)$ and σ_h commute. The automorphism $\phi_\eta(\nu)\sigma_h$ in $\operatorname{Aut}(V_\Lambda)$ has finite order if and only if h is in $\pi_\nu(\Lambda \otimes_\mathbb{Z} \mathbb{Q})$.

We also describe the conjugacy classes in $\operatorname{Aut}(V_\Lambda)$ of a given finite order n. First note that a standard lift $\phi_\eta(\nu)$ of ν has order $m = |\nu|$ if m is odd or if m is even and $\langle \alpha, \nu^{m/2}\alpha\rangle \in 2\mathbb{Z}$ for all $\alpha \in \Lambda$, and order $2m$ otherwise. In the latter case we say that ν exhibits *order doubling*. Then $\phi_\eta(\nu)^m \mathfrak{e}_\alpha = (-1)^{m\langle \pi_\nu(\alpha), \pi_\nu(\alpha)\rangle}\mathfrak{e}_\alpha = (-1)^{\langle \alpha, \nu^{m/2}\alpha\rangle}\mathfrak{e}_\alpha$ for all $\alpha \in \Lambda$. Note that the map sending α to $m\langle \pi_\nu(\alpha), \pi_\nu(\alpha)\rangle = \langle \alpha, \nu^{m/2}\alpha\rangle \bmod 2$ defines a homomorphism $\Lambda \to \mathbb{Z}_2$.

Let $\phi_\eta(\nu)$ be a standard lift. If ν exhibits order doubling, then there exists a vector $s_\nu \in (1/2m)\Lambda^\nu$ defining an inner automorphism $\sigma_{s_\nu} = \mathrm{e}^{2\pi \mathrm{i}(s_\nu)_0}$ of order $2m$ such that $\phi_\eta(\nu)\sigma_{s_\nu}$ has order m. If ν does not exhibit order doubling, we set $s_\nu = 0$. Then the order of an automorphism $\phi_\eta(\nu)\sigma_{s_\nu + f}$ for $f \in \Lambda \otimes_\mathbb{Z} \mathbb{Q}$ is given by $\operatorname{lcm}(m,k)$ where k is the smallest positive integer such that kf is in Λ or equivalently in the fixed-point lattice Λ^ν.

For convenience, we define the s_ν-shifted action of $C_{O(\Lambda)}(\nu)$ on $\pi_\nu(\mathfrak{h}_\Lambda)$ by

$$\tau.f = \tau f + (\tau - \mathrm{id})s_\nu$$

for all $\tau \in C_{O(\Lambda)}(\nu)$ and $f \in \pi_\nu(\mathfrak{h}_\Lambda)$. Then:

Proposition 2.2 [27]. *Fix a section $\nu \mapsto \phi_\eta(\nu)$ mapping only to standard lifts. A complete system of representatives for the conjugacy classes of automorphisms in* $\operatorname{Aut}(V_\Lambda)$ *of order n is given by the $\phi_\eta(\nu)\sigma_{s_\nu+f}$ where*

(1) *ν is from the representatives in $N \subseteq O(\Lambda)$ of order m dividing n,*

(2) *f is from the orbit representatives of the s_ν-shifted action of $C_{O(\Lambda)}(\nu)$ on $(\Lambda^\nu/n)/\pi_\nu(\Lambda)$*

such that $\operatorname{lcm}(m, |\sigma_f|) = n$.

We conclude this section by recalling some results on the twisted modules of lattice vertex operator algebras. For a standard lift $\phi_\eta(\nu)$ the irreducible $\phi_\eta(\nu)$-twisted modules of a lattice vertex operator algebra V_L are described in [1; 15]. Together with the results in [44] this allows us to describe the irreducible g-twisted V_L-modules for all finite-order automorphisms $g \in \operatorname{Aut}(V_L)$.

For simplicity, let L be unimodular. Then V_L is holomorphic and there is a unique irreducible g-twisted V_L-module $V_L(g)$ for each $g \in \operatorname{Aut}(V_L)$ of finite order. Let $g = \phi_\eta(\nu)\sigma_h$ for some standard lift $\phi_\eta(\nu)$ and $\sigma_h = \mathrm{e}^{2\pi i h_0}$ for some $h \in \pi_\nu(L \otimes_{\mathbb{Z}} \mathbb{Q})$. Then

$$V_L(g) = M(1)[\nu] \otimes \mathbb{C}[-h + \pi_\nu(L)] \otimes \mathbb{C}^{d(\nu)}$$

with the twisted Heisenberg module $M(1)[\nu]$, grading by the lattice coset $-h + \pi_\nu(L)$ and defect $d(\nu) \in \mathbb{Z}_{>0}$. (The minus sign in $-h+\pi_\nu(\Lambda)$ has to do with the sign convention in the definition of twisted modules. Here, we follow the convention in, e.g., [22] as opposed to some older texts.)

Assume that ν has order m and cycle shape $\prod_{t|m} t^{b_t}$ with $b_t \in \mathbb{Z}$, that is, the extension of ν to $\mathfrak{h}_L$ has characteristic polynomial $\prod_{t|m}(x^t-1)^{b_t}$. Then the conformal weight of $V_L(g)$ is given by

$$\rho(V_L(g)) = \frac{1}{24}\sum_{t|m} b_t \left(t - \frac{1}{t}\right) + \min_{\alpha \in -h+\pi_\nu(L)} \frac{\langle\alpha,\alpha\rangle}{2} \geq 0,$$

where $\rho_\nu = \frac{1}{24}\sum_{t|m} b_t\big(t - \frac{1}{t}\big)$ is called the *vacuum anomaly* of $V_L(g)$ [15]. Note that ρ_ν is positive for $\nu \neq \mathrm{id}$. The second term is half of the norm of a shortest vector in the lattice coset $-h + \pi_\nu(L)$.

3. Holomorphic vertex operator algebras of central charge 24

We recall the notion of the affine structure of a strongly rational, holomorphic vertex operator algebra of central charge 24 and describe the bijection between these vertex operator algebras and the generalised deep holes of the Leech lattice vertex operator algebra [50].

3.1. ***Affine structure.*** Let $V=\bigoplus_{n=0}^{\infty} V_n$ be a vertex operator algebra of CFT-type. Then the zero modes

$$[a,b]:=a_0b$$

for $a,b\in V_1$ endow the weight-1 space V_1 with the structure of a finite-dimensional Lie algebra. Moreover, the zero modes a_0 for $a\in V_1$ equip each V-module with an action of this Lie algebra.

If $g\in\operatorname{Aut}(V)$ is an automorphism of the vertex operator algebra V, fixing the vacuum vector $\mathbf{1}\in V_0$ and the Virasoro vector $\omega\in V_2$ by definition, then the restriction of g to V_1 is a Lie algebra automorphism, possibly of smaller order.

If V is also self-contragredient, then there exists a nondegenerate, invariant bilinear form $\langle\,\cdot\,,\cdot\,\rangle$ on V, which is unique up to a nonzero scalar and symmetric [30; 43]. We normalise this bilinear form such that $\langle\mathbf{1},\mathbf{1}\rangle=-1$. Then $a_1b=b_1a=\langle a,b\rangle\mathbf{1}$ for all for $a,b\in V_1$.

Let $\mathfrak{g}$ be a simple, finite-dimensional Lie algebra with the nondegenerate, invariant bilinear form $(\,\cdot\,,\cdot\,)$ normalised such that $(\alpha,\alpha)=2$ for all long roots α. The affine Kac–Moody algebra $\hat{\mathfrak{g}}$ associated with $\mathfrak{g}$ is the Lie algebra $\hat{\mathfrak{g}}:=\mathfrak{g}\otimes\mathbb{C}[t,t^{-1}]\oplus\mathbb{C}K$ with central element K and Lie bracket

$$[a\otimes t^m,b\otimes t^n]:=[a,b]\otimes t^{m+n}+m(a,b)\delta_{m+n,0}K$$

for $a,b\in\mathfrak{g}$, $m,n\in\mathbb{Z}$.

A $\hat{\mathfrak{g}}$-module is said to have level $k\in\mathbb{C}$ if K acts as $k\,\mathrm{id}$. Let $\lambda\in P_+$ be a dominant integral weight and $k\in\mathbb{C}$. Then we denote by $L_{\hat{\mathfrak{g}}}(k,\lambda)$ the irreducible quotient of the $\hat{\mathfrak{g}}$-module of level k induced from the irreducible highest-weight $\mathfrak{g}$-module $L_{\mathfrak{g}}(\lambda)$ (see, for example, [35]).

For a positive integer $k\in\mathbb{Z}_{>0}$, $L_{\hat{\mathfrak{g}}}(k,0)$ admits the structure of a rational vertex operator algebra, called the simple affine vertex operator algebra of level k, whose irreducible modules are given by the modules $L_{\hat{\mathfrak{g}}}(k,\lambda)$ for $\lambda\in P_+^k$, the subset of the dominant integral weights P_+ of level at most k [28].

If V is a self-contragredient vertex operator algebra of CFT-type, the commutator formula implies that the modes satisfy

$$[a_m,b_n]=(a_0b)_{m+n}+m(a_1b)_{m+n-1}=[a,b]_{m+n}+m\langle a,b\rangle\delta_{m+n,0}\,\mathrm{id}_V$$

for all $a,b\in V_1$, $m,n\in\mathbb{Z}$. Comparing this with the definition above we see that for a simple Lie subalgebra $\mathfrak{g}$ of V_1 the map $a\otimes t^n\mapsto a_n$ for $a\in\mathfrak{g}$ and $n\in\mathbb{Z}$ defines a representation of $\hat{\mathfrak{g}}$ on V of some level $k_{\mathfrak{g}}\in\mathbb{C}$ with $\langle\,\cdot\,,\cdot\,\rangle|_{\mathfrak{g}}=k_{\mathfrak{g}}(\,\cdot\,,\cdot\,)$.

Suppose that V is strongly rational. Then it is shown in [18] that the Lie algebra V_1 is reductive, that is, a direct sum of a semisimple and an abelian Lie algebra. Moreover, it is stated in [19] that for a simple Lie subalgebra $\mathfrak{g}$ of V_1 the restriction of $\langle\,\cdot\,,\cdot\,\rangle$ to $\mathfrak{g}$ is nondegenerate, the level $k_{\mathfrak{g}}$ is a positive integer, the vertex operator subalgebra of V generated by $\mathfrak{g}$ is isomorphic to $L_{\hat{\mathfrak{g}}}(k_{\mathfrak{g}},0)$ and V is an integrable $\hat{\mathfrak{g}}$-module.

Assume in addition that V is holomorphic and of central charge 24. Then the Lie algebra V_1 is zero, abelian of dimension 24 or semisimple of rank at most 24 [17]. If the Lie algebra V_1 is semisimple, then it decomposes into a direct sum

$$V_1 \cong \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$$

of simple ideals $\mathfrak{g}_i$ and the vertex operator subalgebra $\langle V_1 \rangle$ of V generated by V_1 is isomorphic to the tensor product of simple affine vertex operator algebras

$$\langle V_1 \rangle \cong L_{\hat{\mathfrak{g}}_1}(k_1, 0) \otimes \cdots \otimes L_{\hat{\mathfrak{g}}_r}(k_r, 0)$$

with levels $k_i := k_{\mathfrak{g}_i} \in \mathbb{Z}_{>0}$ and has the same Virasoro vector as V. The tensor-product decomposition of the vertex operator algebra $\langle V_1 \rangle$ is called the *affine structure* of V and denoted by $\mathfrak{g}_{1,k_1} \cdots \mathfrak{g}_{r,k_r}$.

Since $\langle V_1 \rangle \cong L_{\hat{\mathfrak{g}}_1}(k_1, 0) \otimes \cdots \otimes L_{\hat{\mathfrak{g}}_r}(k_r, 0)$ is rational, V decomposes into the direct sum of finitely many irreducible $\langle V_1 \rangle$-modules

$$V \cong \bigoplus_{\lambda} m_\lambda L_{\hat{\mathfrak{g}}_1}(k_1, \lambda_1) \otimes \cdots \otimes L_{\hat{\mathfrak{g}}_r}(k_r, \lambda_r),$$

where $m_\lambda \in \mathbb{Z}_{\geq 0}$ and the sum ranges over finitely many $\lambda = (\lambda_1, \ldots, \lambda_r)$ with dominant integral weights $\lambda_i \in P_+^{k_i}(\mathfrak{g}_i)$, that is, of level at most k_i.

Let $h_i^\vee$ denote the dual Coxeter number of $\mathfrak{g}_i$. The fact that the character of V is a Jacobi form of lattice index implies the trace identity

$$\frac{h_i^\vee}{k_i} = \frac{\dim(V_1) - 24}{24} \tag{S}$$

for all $i = 1, \ldots, r$ (see [17; 25; 55]). As a consequence, the Lie algebra V_1 uniquely determines the affine structure, i.e., the levels k_i. The equation has exactly 221 solutions (see Table 3 in [27]).

In [55] Schellekens also derived so-called higher-order trace identities (cf. [25], Theorem 6.1), which allowed him to reduce the above 221 affine structures down to 69 by solving large integer linear programming problems on the computer. Together with the zero Lie algebra and the 24-dimensional abelian Lie algebra this gives Schellekens' list of 71 Lie algebras (see Table 2) that occur as the weight-1 space of a strongly rational, holomorphic vertex operator algebra of central charge 24 [55].

We shall however not make use of Schellekens' classification result but give an independent proof based on the classification of certain geometric structures in the Leech lattice Λ.

3.2. *Generalised deep holes.* One of the main results of [50] is a dimension formula for the weight-1 space of the cyclic orbifold construction $V^{\mathrm{orb}(g)}$.

Theorem 3.1 (dimension formula, [50, Theorem 5.3 and Corollary 5.7]). *Let V be a strongly rational, holomorphic vertex operator algebra of central charge* 24 *and g an automorphism of V of finite order $n > 1$ and type* 0 *such that V^g satisfies the positivity condition. Then the dimension of the weight-*1 *subspace of $V^{\mathrm{orb}(g)}$ is*

$$\dim(V_1^{\mathrm{orb}(g)}) = 24 + \sum_{d|n} c_n(d) \dim(V_1^{g^d}) - R(g),$$

where the $c_n(d) \in \mathbb{Q}$ are defined by $\sum_{d|n} c_n(d)(t,d) = n/t$ for all $t \mid n$ and the remainder term $R(g)$ is nonnegative. In particular,

$$\dim(V_1^{\mathrm{orb}(g)}) \leq 24 + \sum_{d|n} c_n(d) \dim(V_1^{g^d}).$$

This dimension formula is obtained by pairing the vector-valued character of the fixed-point vertex operator subalgebra V^g with a vector-valued Eisenstein series of weight 2, and it generalises earlier results in [26; 39; 48; 51] under the assumption that the modular curve $\Gamma_0(n)\backslash\mathbb{H}^*$ has genus zero.

We point out that the upper bound in the dimension formula depends only on the action of g on the weight-1 Lie algebra V_1.

An automorphism g such that $\dim(V_1^{\mathrm{orb}(g)})$ attains the above upper bound is called *extremal*. We also call the identity automorphism extremal.

The upper bound in the dimension formula motivates the following definition.

Definition 3.2 (generalised deep hole, [50]). Let V be a strongly rational, holomorphic vertex operator algebra of central charge 24 and $g \in \mathrm{Aut}(V)$ of finite order $n > 1$. Suppose g has type 0 and V^g satisfies the positivity condition. Then g is called a *generalised deep hole* of V if

(1) g is extremal, i.e., $\dim(V_1^{\mathrm{orb}(g)}) = 24 + \sum_{d|n} c_n(d) \dim(V_1^{g^d})$,

(2) $\mathrm{rk}(V_1^{\mathrm{orb}(g)}) = \mathrm{rk}(V_1^g)$.

In other words, we demand the dimension of the Lie algebra $V_1^{\mathrm{orb}(g)}$ to be maximal with respect to the upper bound from the dimension formula and the rank to be minimal with respect to the obvious lower bound $\mathrm{rk}(V_1^g)$.

By convention, we call the identity automorphism a generalised deep hole.

Recall that the Lie algebras V_1^g and $V_1^{\mathrm{orb}(g)}$ are reductive. By Lemma 8.1 in [35] the centraliser in $V_1^{\mathrm{orb}(g)}$ of any choice of Cartan subalgebra of V_1^g is a Cartan subalgebra of $V_1^{\mathrm{orb}(g)}$. Condition (2) is hence equivalent to demanding that any Cartan subalgebra of V_1^g also be a Cartan subalgebra of $V_1^{\mathrm{orb}(g)}$. It can be replaced by the equivalent condition that the inverse-orbifold automorphism restricts to an inner automorphism on $V_1^{\mathrm{orb}(g)}$.

If $V \cong V_\Lambda$, the vertex operator algebra associated with the Leech lattice Λ, then the rank condition is equivalent to demanding that $(V_\Lambda^g)_1$, which as a subalgebra of $(V_\Lambda)_1$ is abelian, be a Cartan subalgebra of $(V_\Lambda^{\mathrm{orb}(g)})_1$.

The second main result of [50] is a natural bijection between the generalised deep holes of the Leech lattice vertex operator algebra V_Λ and the strongly rational, holomorphic vertex operator algebras of central charge 24 with nonvanishing weight-1 space.

Theorem 3.3 (holey correspondence, [50]). *The cyclic orbifold construction $g \mapsto V_\Lambda^{\mathrm{orb}(g)}$ defines a bijective correspondence between the algebraic conjugacy classes of generalised deep holes $g \in \mathrm{Aut}(V_\Lambda)$ with $\mathrm{rk}((V_\Lambda^g)_1) > 0$ and the isomorphism classes of strongly rational, holomorphic vertex operator algebras V of central charge 24 with $V_1 \neq \{0\}$.*

The proof combines the dimension formula with an averaged version of Kac's very strange formula [35]. It does not use any classification result for either side of the correspondence.

This theorem generalises the natural bijection between the deep holes of the Leech lattice Λ and the Niemeier lattices with roots [3], which is mediated by the holey construction [10].

Recall that the weight-1 Lie algebra V_1 of a strongly rational, holomorphic vertex operator algebra V of central charge 24 is either abelian or semisimple. If g is a generalised deep hole of V_Λ with nonzero $(V_\Lambda^g)_1$ and $V \cong V_\Lambda^{\operatorname{orb}(g)}$, then V_1 is abelian if and only if $\dim(V_1) = 24$ if and only if $V \cong V_\Lambda$ if and only if $g = \mathrm{id}$.

The inverse orbifold construction corresponding to a generalised deep hole g of the Leech lattice vertex operator algebra V_Λ takes a very simple form [27]. Assume that $V = V_\Lambda^{\operatorname{orb}(g)}$ is a strongly rational, holomorphic vertex operator algebra V of central charge 24 with $V_1 = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ semisimple. Then the inverse-orbifold automorphism of g (which must be of type 0 and extremal [50]) is given by the inner automorphism

$$\sigma_u = \mathrm{e}^{2\pi \mathrm{i} u_0} \quad \text{with } u := \sum_{i=1}^{r} \frac{\rho_i}{h_i^\vee},$$

where $h_i^\vee$ is the dual Coxeter number and ρ_i the Weyl vector of $\mathfrak{g}_i$. The order of σ_u on each simple ideal $\mathfrak{g}_i$ is $l_i h_i^\vee$ where $l_i \in \{1, 2, 3\}$ is the lacing number of $\mathfrak{g}_i$. Hence, the order on V_1 is $\operatorname{lcm}(\{l_i h_i^\vee\}_{i=1}^r)$, which can be shown to equal the order n of σ_u on the whole vertex operator algebra V. Of course, this equals the order of the corresponding generalised deep hole $g \in \operatorname{Aut}(V_\Lambda)$.

4. Generalised hole diagrams

We associate generalised hole diagrams with automorphisms of the Leech lattice vertex operator algebra V_Λ. They will be the main datum we use to classify the generalised deep holes in $\operatorname{Aut}(V_\Lambda)$.

Let V_Λ be the Leech lattice vertex operator algebra and $g \in \operatorname{Aut}(V_\Lambda)$ of order $n > 1$ and type 0 such that V_Λ^g satisfies the positivity condition. Let ν be the projection of g to $O(\Lambda)$. Consider the orbifold construction $V_\Lambda^{\operatorname{orb}(g)} = \bigoplus_{i \in \mathbb{Z}_n} W_\Lambda^{(i,0)}$ and assume that $\operatorname{rk}((V_\Lambda^{\operatorname{orb}(g)})_1) = \operatorname{rk}((V_\Lambda^g)_1) > 0$. Then $\mathfrak{g} := (V_\Lambda^{\operatorname{orb}(g)})_1$ is a semisimple or abelian Lie algebra and $\mathfrak{h} = (V_\Lambda^g)_1 = \{h(-1) \otimes \mathfrak{e}_0 \mid h \in \pi_\nu(\mathfrak{h}_\Lambda)\}$ is a Cartan subalgebra of $\mathfrak{g}$.

The nondegenerate, invariant bilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ on $V_\Lambda^{\operatorname{orb}(g)}$, normalised such that $\langle \mathbf{1}, \mathbf{1} \rangle = -1$, restricts to a nondegenerate, invariant bilinear form on $\mathfrak{g}$. The Cartan subalgebra $\mathfrak{h}$ with the form $\langle\,\cdot\,,\,\cdot\,\rangle$ is naturally isometric to the subspace $\pi_\nu(\mathfrak{h}_\Lambda)$ of $\mathfrak{h}_\Lambda = \mathbb{C} \otimes_{\mathbb{Z}} \Lambda$. We may also identify $\mathfrak{h}$ with $\mathfrak{h}^*$ via $\langle\,\cdot\,,\,\cdot\,\rangle$. We write the Cartan decomposition corresponding to $\mathfrak{h}$ as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$$

with root system $\Phi \subseteq \mathfrak{h}^*$, which is empty if and only if $\mathfrak{g}$ is abelian. The inverse orbifold automorphism ζ of g restricts to an inner automorphism of $\mathfrak{g}$, and $\mathfrak{g}$ decomposes into eigenspaces

$$\mathfrak{g} = \mathfrak{g}_{(0)} \oplus \mathfrak{g}_{(1)} \oplus \cdots \oplus \mathfrak{g}_{(n-1)}$$

where $\mathfrak{g}_{(i)} = \mathfrak{g} \cap W_\Lambda^{(i,0)} = (W_\Lambda^{(i,0)})_1$ and $\mathfrak{g}_{(0)} = (V_\Lambda^g)_1 = \mathfrak{h}$. Since the action of ζ commutes with the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ and the spaces $\mathfrak{g}_\alpha$ are 1-dimensional, each $\mathfrak{g}_\alpha$ lies in exactly one $\mathfrak{g}_{(i)}$. Hence, the root system Φ is a disjoint union

$$\Phi = \Phi_{(1)} \cup \cdots \cup \Phi_{(n-1)}$$

with $\Phi_{(i)} = \{\alpha \in \Phi \mid \mathfrak{g}_\alpha \subseteq \mathfrak{g}_{(i)}\}$. We define

$$\Pi(g) := \Phi_{(1)}.$$

Since $(1, n) = 1$, the weight-1 subspace of the irreducible g-twisted V-module $V_\Lambda(g)$ is $(W_\Lambda^{(1,0)})_1$. Hence, $\Pi(g) \subseteq \mathfrak{h}^*$ can also be defined as the set of weights of the adjoint action of $\mathfrak{h}$ on $V_\Lambda(g)_1$.

Proposition 4.1. *If the root system Φ of $\mathfrak{g}$ is nonempty, then the inner products $2\langle\alpha_i, \alpha_j\rangle/\langle\alpha_i, \alpha_i\rangle$ for $\alpha_i, \alpha_j \in \Pi(g)$ form a generalised Cartan matrix with Dynkin diagram $\Phi(g)$ given by a subdiagram of the extended affine Dynkin diagram associated with the (finite) Dynkin diagram of Φ.*

Proof. This is Proposition 8.6c) in [35] together with the fact that $\mathfrak{g}_{(0)} = \mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. $\square$

Let $\varphi(g)$ be the cycle shape of the image ν of g under the natural projection $\mathrm{Aut}(V_\Lambda) \to O(\Lambda)$. We define the *generalised hole diagram* of g as the pair

$$(\varphi(g), \Phi(g)).$$

The generalised hole diagram only depends on the algebraic conjugacy class of g in $\mathrm{Aut}(V_\Lambda)$. For $g = \mathrm{id}$ we set $\Phi(g) = \emptyset$.

In the following, we study the weights $\Pi(g) \subseteq \mathfrak{h}^*$ and the corresponding Dynkin diagram $\Phi(g)$ in more detail. These results will allow us to classify these diagrams in Section 5 in the case where g is a generalised deep hole.

By Proposition 2.1, up to conjugacy, $g = \phi_\eta(\nu)\sigma_h$ for some standard lift $\phi_\eta(\nu)$ of $\nu \in O(\Lambda) \cong \mathrm{Co}_0$ and some $h \in \pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q})$. Define $m = |\nu|$, which divides $n = |g|$, and let $\prod_{t|m} t^{b_t}$ be the cycle shape of ν. Recall that the unique irreducible g-twisted V_Λ-module is of the form

$$V_\Lambda(g) = M(1)[\nu] \otimes \mathbb{C}[-h + \pi_\nu(\Lambda)] \otimes \mathbb{C}^{d(\nu)}$$

with the twisted Heisenberg algebra $M(1)[\nu]$ and defect $d(\nu) \in \mathbb{Z}_{>0}$. $V_\Lambda(g)$ is spanned by the vectors

$$v = h_1(-n_1) \cdots h_r(-n_r) \otimes \mathfrak{e}_\alpha \otimes t,$$

where the h_i are in certain eigenspaces of $\mathfrak{h}_\Lambda$, $n_i \in (1/m)\mathbb{Z}_{>0}$, $\alpha \in -h + \pi_\nu(\Lambda)$ and $t \in \mathbb{C}^{d(\nu)}$. Such a vector has L_0-weight

$$\mathrm{wt}(v) = \rho_\nu + n_1 + \cdots + n_r + \frac{\langle\alpha, \alpha\rangle}{2}$$

with vacuum anomaly $\rho_\nu = \frac{1}{24}\sum_{t|m} b_t(t - 1/t)$ and is acted on by the Cartan subalgebra $\mathfrak{h} = (V_\Lambda^g)_1 \cong \pi_\nu(\mathfrak{h}_\Lambda)$ of $\mathfrak{g} = (V_\Lambda^{\mathrm{orb}(g)})_1$ as

$$h_0 v = \langle h, \alpha\rangle v \quad \text{for } h \in \pi_\nu(\mathfrak{h}_\Lambda).$$

Proposition 4.2. *The weights of the action of* $\mathfrak{h}$ *on* $V_\Lambda(g)_1$ *are given by*

$$\Pi(g) = \left\{\alpha \in -h + \pi_\nu(\Lambda) \;\middle|\; \frac{\langle\alpha,\alpha\rangle}{2} = 1 - \rho_\nu\right\}$$

if $d(\nu) = 1$ *and*

$$\Pi(g) = \varnothing$$

if $d(\nu) > 1$.

Proof. First, note that $\rho_\nu \geq 1 - 1/m$ for all $\nu \in O(\Lambda)$. Moreover, $V_\Lambda(g)_1$ cannot contain any vector $v = \cdots \otimes \mathfrak{e}_0 \otimes t$ as this would lie in the centraliser of $\mathfrak{h} = (V_\Lambda^g)_1$, contradicting the assumption that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Therefore, a vector $v \in V_\Lambda(g)_1$ must be of the form $v = 1 \otimes \mathfrak{e}_\alpha \otimes t$ for some (nonzero) $\alpha \in -h + \pi_\nu(\Lambda)$ and $t \in \mathbb{C}^{d(\nu)}$, i.e., there can be no contribution to the weight from the twisted Heisenberg algebra except for the vacuum anomaly. Hence,

$$V_\Lambda(g)_1 = \left\{1 \otimes \mathfrak{e}_\alpha \otimes t \;\middle|\; \alpha \in -h + \pi_\nu(\Lambda) \text{ such that } \frac{\langle\alpha,\alpha\rangle}{2} = 1 - \rho_\nu,\ t \in \mathbb{C}^{d(\nu)}\right\}.$$

Since the action of the Cartan subalgebra $\mathfrak{h}$ is independent of t and all weight spaces are 1-dimensional, either $d(\nu) = 1$ or $V_\Lambda(g)_1 = \{0\}$. In the first case

$$\Pi(g) = \left\{\alpha \in -h + \pi_\nu(\Lambda) \;\middle|\; \frac{\langle\alpha,\alpha\rangle}{2} = 1 - \rho_\nu\right\},$$

while $\Pi(g) = \varnothing$ if $d(\nu) > 1$. □

Even if $d(\nu) = 1$, it is possible for $\Pi(g)$ to be empty, for instance if the shortest vectors in $-h + \pi_\nu(\Lambda)$ have norm greater than $2(1 - \rho_\nu)$. This is in particular the case if $\rho_\nu > 1$.

Proposition 4.3. *The Dynkin diagram* $\Phi(g)$ *of* $\Pi(g)$ *can also be obtained as follows. Each vector* $\alpha_i \in \Pi(g) \subseteq \mathfrak{h}^*$ *defines a node of* $\Phi(g)$. *The nodes* α_i *and* α_j *for* $i \neq j$ *are joined by*

(1) *no edge if* $\langle\alpha_i - \alpha_j, \alpha_i - \alpha_j\rangle/2 = 2(1 - \rho_\nu)$,

(2) *a single edge if* $\langle\alpha_i - \alpha_j, \alpha_i - \alpha_j\rangle/2 = 3(1 - \rho_\nu)$,

(3) *an undirected double edge if* $\langle\alpha_i - \alpha_j, \alpha_i - \alpha_j\rangle/2 = 4(1 - \rho_\nu)$,

corresponding to angles of $2\pi/4$, $2\pi/3$ *and* $2\pi/2$, *respectively, between* α_i *and* α_j.

We define the *shifted weights*

$$\tilde{\Pi}(g) := \Pi(g) + h = \left\{\beta \in \pi_\nu(\Lambda) \;\middle|\; \frac{\langle\beta - h, \beta - h\rangle}{2} = 1 - \rho_\nu\right\} \subseteq \pi_\nu(\Lambda).$$

We can associate a Dynkin diagram with $\tilde{\Pi}(g)$ in the same way as with $\Pi(g)$, using Proposition 4.3. Since the translation by h does not affect the distances between the weights, both diagrams coincide. Geometrically, $\tilde{\Pi}(g)$ is given by the elements in $\pi_\nu(\Lambda)$ lying on the sphere in $\pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})$ with centre h and radius $\sqrt{2(1 - \rho_\nu)}$.

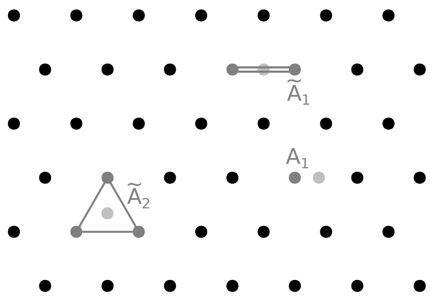

Figure 1. Dynkin diagrams in the lattice A_2.

Examples of Dynkin diagrams inside the lattice A_2 (with different radii) are shown in Figure 1. The centres are shown light grey, the diagrams middle gray.

In a connected extended affine Dynkin diagram with simple roots $\alpha_0, \ldots, \alpha_l$ there is a linear relation between the α_i. More precisely, there are positive integers a_i such that $\sum_{i=0}^{l} a_i\alpha_i = 0$. If chosen coprime, the a_i are unique and sometimes called *Kac labels* (see, e.g., Table Aff 1 in Section 4.8 of [35]).

Proposition 4.4. *If $\Phi(g)$ contains a connected component of affine type, then the centre h of $\tilde{\Pi}(g)$ can be reconstructed from the weights in $\tilde{\Pi}(g)$.*

Proof. Denote the shifted weights of the connected affine component by $\beta_0, \ldots, \beta_l$. Write $\beta_i = \alpha_i + h$. Then

$$h = \sum_{i=0}^{l} a_i\beta_i \Big/ \sum_{i=0}^{l} a_i. \qquad \square$$

We now additionally assume that the automorphism $g = \phi_\eta(\nu)\sigma_h \in \operatorname{Aut}(V_\Lambda)$ is extremal, i.e., that g is a generalised deep hole. Then $\rho(V_\Lambda(g)) \geq 1$, so that

$$\min_{\beta\in\pi_\nu(\Lambda)} \frac{\langle \beta - h, \beta - h\rangle}{2} \geq 1 - \rho_\nu$$

(see Proposition 5.9 in [50]). Hence, if the hole diagram $\Phi(g)$ is nonempty, then the points in $\tilde{\Pi}(g)$ are exactly the *closest vectors* to h in $\pi_\nu(\Lambda)$.

As a side remark (cf. [38]), we note that h is in general not a *deep hole* or even just a *hole* of the lattice $\pi_\nu(\Lambda)$. Indeed, for most $\nu \in O(\Lambda)$ the covering radius of $\pi_\nu(\Lambda)$ is greater than $\sqrt{2(1-\rho_\nu)}$ so that h cannot be a deep hole of $\pi_\nu(\Lambda)$. In fact, usually the number of points in $\tilde{\Pi}(g)$ is less than $\operatorname{rk}(\pi_\nu(\Lambda)) + 1$, which means that h cannot be a hole. On the other hand, if $\nu \in O(\Lambda)$ is such that the covering radius of $\pi_\nu(\Lambda)$ is less than $\sqrt{2(1-\rho_\nu)}$, then there can be no extremal automorphism in $\operatorname{Aut}(V_\Lambda)$ projecting to ν.

We now exploit the fact that the inverse-orbifold automorphism of such a generalised deep hole g is known [27]. Since we assumed that g has order $n > 1$, $\mathfrak{g} = (V_\Lambda^{\mathrm{orb}(g)})_1$ must be semisimple, with decomposition $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ into simple ideals. Recall that the inverse-orbifold automorphism is given by $\sigma_u = e^{2\pi i u_0} \in \operatorname{Aut}(V_\Lambda^{\mathrm{orb}(g)})$ with $u = \sum_{i=1}^{r} \rho_i / h_i^\vee$ where $h_i^\vee$ is the dual Coxeter number and ρ_i the Weyl vector of $\mathfrak{g}_i$ (see Section 3). The restriction of σ_u to $\mathfrak{g}$ only depends on the Lie algebra structure of $\mathfrak{g}$, which means that the Dynkin diagram $\Phi(g)$ can be easily read off from the isomorphism type of $\mathfrak{g}$:

Proposition 4.5. *Let g be a generalised deep hole of V_Λ of order $n > 1$ with $\operatorname{rk}((V_\Lambda^g)_1) > 0$. Then $(V_\Lambda^{\operatorname{orb}(g)})_1 = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ is semisimple and the Dynkin diagram $\Phi(g)$ is of type*

$$\bigcup_{\substack{i=1\\ l_i h_i^\vee = n}}^{r} \begin{cases} \tilde{A}_l & \text{if } \mathfrak{g}_i \text{ has type } A_l,\ l \geq 1,\\ A_1 & \text{if } \mathfrak{g}_i \text{ has type } B_l,\ l \geq 2,\\ A_{l-1} & \text{if } \mathfrak{g}_i \text{ has type } C_l,\ l \geq 3,\\ \tilde{D}_l & \text{if } \mathfrak{g}_i \text{ has type } D_l,\ l \geq 4,\\ \tilde{E}_l & \text{if } \mathfrak{g}_i \text{ has type } E_l,\ l \in \{6,7,8\},\\ A_2 & \text{if } \mathfrak{g}_i \text{ has type } F_4,\\ A_1 & \text{if } \mathfrak{g}_i \text{ has type } G_2, \end{cases}$$

where $l_i \in \{1, 2, 3\}$ is the lacing number of the simple ideal $\mathfrak{g}_i$.

The order of σ_u on each simple ideal $\mathfrak{g}_i$ is $l_i h_i^\vee$ so that the order of σ_u on $(V_\Lambda^{\operatorname{orb}(g)})_1$ is $\operatorname{lcm}(\{l_i h_i^\vee\}_{i=1}^r)$, which can be shown to equal the order n of σ_u on the whole vertex operator algebra $V_\Lambda^{\operatorname{orb}(g)}$. The proposition states in particular that only those simple ideals contribute to the Dynkin diagram $\Phi(g)$, on which σ_u assumes its order.

Proof. Recall that the inverse orbifold automorphism acts on $(W_\Lambda^{(1,0)})_1 = V_\Lambda(g)_1$ as multiplication by $\mathrm{e}^{2\pi\mathrm{i}/n}$. Hence, the simple ideal $\mathfrak{g}_i$ can only contribute to $(V_\Lambda(g))_1$ if the order of σ_u restricted to $\mathfrak{g}_i$, which is $l_i h_i^\vee$, equals n. On a simple ideal where this is the case, the eigenspace for the eigenvalue $\mathrm{e}^{2\pi\mathrm{i}/n}$ is now determined following Proposition 8.6c) in [35]. For this one uses the type (in the language of [35]) of σ_u restricted to $\mathfrak{g}_i$, which is described in the proof of Proposition 5.1 in [27]. □

The special case of the proposition for types A, D and E was already discussed in [41] (see Lemma 2.6).

From what we have seen so far, the Dynkin diagram $\Phi(g)$ of a generalised deep hole could in principle be empty. The following is immediate:

Corollary 4.6. *Let g be a generalised deep hole of V_Λ of order $n > 1$ with $\operatorname{rk}((V_\Lambda^g)_1) > 0$. Then the following are equivalent*:

(1) *The Dynkin diagram $\Phi(g)$ is nonempty.*

(2) *The set of shifted weights $\tilde{\Pi}(g)$ is nonempty.*

(3) $\rho(V_\Lambda(g)) = 1$.

(4) $l_i h_i^\vee = \operatorname{lcm}(\{l_j h_j^\vee\}_{j=1}^r)$ *for some $i \in \{1, \ldots, r\}$.*

(5) $|\sigma_u| = \big|\sigma_u|_{\mathfrak{g}_i}\big|$ *for some $i \in \{1, \ldots, r\}$.*

We now discuss the special case of g being an inner automorphism. In this case, we exactly recover the classical hole diagrams in [12]:

Proposition 4.7. *Let g be a generalised deep hole of V_Λ of order $n > 1$ with $\operatorname{rk}((V_\Lambda^g)_1) > 0$. Assume that g is inner. Then $g = \sigma_h$ for some deep hole $h \in \Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ corresponding to the Niemeier lattice N. Let $\tilde{N}$ be the extended affine Dynkin diagram corresponding to N, which is the hole diagram of h. Then $V_\Lambda^{\operatorname{orb}(g)} \cong V_N$ and g has the generalised hole diagram $(1^{24}, \tilde{N})$.*

Proof. Since g is inner, $g = \sigma_h$ for some $h \in \Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$. The extremality of g implies that $\rho(V(g)) \geq 1$. But the covering radius of the Leech lattice Λ is $\sqrt{2}$, so that

$$\rho(V(g)) = \min_{\beta \in \Lambda} \frac{\langle \beta - h, \beta - h \rangle}{2} = 1,$$

that is, h is a deep hole of Λ. The remaining claims follow from Proposition 4.3 and the results in [12]. $\square$

5. Classification of generalised deep holes

We classify the generalised deep holes of the Leech lattice vertex operator algebra by enumerating the corresponding generalised hole diagrams. As a consequence we obtain a new, geometric classification of the strongly rational, holomorphic vertex operator algebras of central charge 24 with nontrivial weight-1 space, which is independent of Schellekens' results.

The possible generalised hole diagrams are strongly restricted by the following result (see Lemma 6.1 in [27]):

Proposition 5.1. *Let V be a strongly rational, holomorphic vertex operator algebra of central charge* 24 *with V_1 semisimple. Let $\mathfrak{g}_{1,k_1} \cdots \mathfrak{g}_{r,k_r}$ denote the affine structure* (*with dual Coxeter numbers $h_i^\vee$ and lacing numbers l_i*). *Then*

(1) $h_i^\vee / k_i = (\dim(V_1) - 24)/24$

for all $i = 1, \ldots, r$, and there exists a $\nu \in O(\Lambda)$ such that

(2) $\mathrm{rk}(\Lambda^\nu) = \mathrm{rk}(V_1)$,

(3) $|\nu| \,\big|\, \mathrm{lcm}(\{l_i h_i^\vee\}_{i=1}^r)$,

(4) $1/(1 - \rho_\nu) = \mathrm{lcm}(\{l_i k_i\}_{i=1}^r)$.

The automorphism ν is exactly the projection $\mathrm{Aut}(V_\Lambda) \to O(\Lambda)$ of the generalised deep hole g corresponding to V by Theorem 3.3. Recall that ρ_ν denotes the vacuum anomaly of ν and only depends on the cycle shape of ν.

The first equation is Schellekens' lowest-order trace identity (S). The other conditions follow from the bijection in Theorem 3.3.

It is straightforward to list all solutions, i.e., pairs of affine structures and automorphisms of the Leech lattice Λ, to the equations in Proposition 5.1 (see Proposition 6.2 in [27]):

Proposition 5.2. *There are exactly* 82 *pairs of affine structures and conjugacy classes in $O(\Lambda)$ satisfying the four equations in Proposition 5.1. These are the* 69 *cases described in Table 2 plus the* 13 *spurious cases listed in Table 1.*

There is no affine structure that appears in more than one pair. By Proposition 4.5, the affine structure fixes the generalised hole diagram of the corresponding generalised deep hole g. However, there could still be multiple nonconjugate generalised deep holes for a given pair (or generalised hole diagram).

We observe that, except for $g = \mathrm{id}$, the Dynkin diagram $\Phi(g)$ of a generalised deep hole is never empty.

$\nu \in O(\Lambda)$	$\lvert\phi_\eta(\nu)\rvert$	ρ_ν	n	affine structure	$\Phi(g)$	norms
6^4	12	$\frac{35}{36}$	6	$D_{4,36}$	$\tilde{D}_4$	$\frac{2}{18}, \frac{2}{12}$
4^6	8	$\frac{15}{16}$	4	$A_{3,16}^2$	$\tilde{A}_3^2$	$\frac{2}{8}, \frac{6}{16}$
			8	$C_{3,8}A_{3,8}$	A_2	$\frac{6}{16}$
3^8	3	$\frac{8}{9}$	3	$A_{2,9}^4$	$\tilde{A}_2^4$	$\frac{4}{9}, \frac{2}{3}$
			6	$D_{4,9}A_{1,3}^4$	$\tilde{D}_4$	$\frac{4}{9}, \frac{2}{3}$
			12	$G_{2,3}^4$	A_1^4	$\frac{4}{9}$
$2^4 4^4$	4	$\frac{7}{8}$	4	$A_{3,8}^2 A_{1,4}^2$	$\tilde{A}_3^2$	$\frac{2}{4}, \frac{6}{8}$
			8	$C_{3,4}^2 A_{1,2}^2$	A_2^2	$\frac{2}{4}, \frac{6}{8}$
$1^2 2^2 3^2 6^2$	6	$\frac{5}{6}$	6	$D_{4,6}B_{2,3}^2$	$\tilde{D}_4 A_1^2$	$\frac{2}{3}, \frac{2}{2}$
2^{12}	4	$\frac{3}{4}$	4	$A_{3,4}^2 A_{1,2}^6$	$\tilde{A}_3^2$	$\frac{2}{2}, \frac{6}{4}$
			8	$D_{5,4}A_{3,2}A_{1,1}^4$	$\tilde{D}_5$	$\frac{2}{2}, \frac{6}{4}$
			8	$C_{3,2}^3 A_{1,1}^3$	A_2^3	$\frac{2}{2}, \frac{6}{4}$
			8	$C_{3,2}^2 A_{3,2}^2$	A_2^3	$\frac{2}{2}, \frac{6}{4}$

Table 1. 13 spurious cases in Proposition 5.2.

Lemma 5.3. *There are no generalised deep holes in* $\operatorname{Aut}(V_\Lambda)$ *corresponding to the eight spurious cases in Table 1 with cycle shapes* 6^4, 4^6, 3^8 *and* $2^4 4^4$.

Proof. We write the potential generalised deep hole as $g = \phi_\eta(\nu)\sigma_h$ where $\phi_\eta(\nu)$ is a standard lift of $\nu \in O(\Lambda)$. Note that $\langle\beta,\beta\rangle/2 \in (1/\lvert\phi_\eta(\nu)\rvert)\mathbb{Z}$ for all $\beta \in \pi_\nu(\Lambda)$. The hole diagrams $\Phi(g)$ are determined by Proposition 4.5 and listed in Table 1. Based on Proposition 4.3 we can also read off the norms of the differences of the elements in $\tilde{\Pi}(g) \subseteq \pi_\nu(\Lambda)$. Hence, none of the eight cases in the assertion can occur as in each case not all the computed norms are in $(2/\lvert\phi_\eta(\nu)\rvert)\mathbb{Z}$. □

As a consequence, we are left with 5 spurious cases, namely those entries in Table 1 with cycle shapes $1^2 2^2 3^2 6^2$ and 2^{12}.

5.1. ***Affine case.*** Suppose that g is a generalised deep hole projecting to $\nu \in O(\Lambda)$ and that the corresponding set of shifted weights $\tilde{\Pi}(g) \subseteq \pi_\nu(\Lambda)$ contains a connected affine component $\tilde{X}_l$. Our strategy will be to search for the Dynkin diagram $\tilde{X}_l$ inside $\pi_\nu(\Lambda)$ as lattice points lying on a sphere around some point $h \in \pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q})$ of radius $\sqrt{2(1-\rho_\nu)}$ with edges defined as in Proposition 4.3 (see also Figure 1). We enumerate the occurrences of $\tilde{X}_l$ in $\pi_\nu(\Lambda)$, more precisely the finitely many orbits under the action of $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$. This can be done by moving one vertex of $\tilde{X}_l$ to the origin and then performing a short vector search in `Magma` [5] (code available at arxiv.org). In principle, this could also be done by hand, as is demonstrated in [12] in the case $\nu = \mathrm{id}$. Note that $C_{O(\Lambda)}(\nu)$ in general only induces a subgroup

number	number	$(V_\Lambda^{\mathrm{orb}(g)})_1$	dimension	n	$\rho(V_\Lambda(g^m))$	$\Phi(g)$
			rank 24, cycle shape 1^{24}			
70	A1	$D_{24,1}$	1128	46	$1, \frac{22}{23}, 1, 0$	$\tilde{D}_{24}$
69	A2	$D_{16,1}E_{8,1}$	744	30	$1, \frac{14}{15}, 1, 1, 1, 1, 1, 0$	$\tilde{D}_{16}\tilde{E}_8$
68	A3	$E_{8,1}^3$	744	30	$1, \frac{14}{15}, \frac{9}{10}, \frac{5}{6}, 1, 1, 1, 0$	$\tilde{E}_8^3$
67	A4	$A_{24,1}$	624	25	$1, 1, 0$	$\tilde{A}_{24}$
66	A5	$D_{12,1}^2$	552	22	$1, \frac{10}{11}, 1, 0$	$\tilde{D}_{12}^2$
65	A6	$A_{17,1}E_{7,1}$	456	18	$1, 1, 1, 1, 1, 0$	$\tilde{A}_{17}\tilde{E}_7$
64	A7	$D_{10,1}E_{7,1}^2$	456	18	$1, \frac{8}{9}, 1, 1, 1, 0$	$\tilde{D}_{10}\tilde{E}_7^2$
63	A8	$A_{15,1}D_{9,1}$	408	16	$1, 1, 1, 1, 0$	$\tilde{A}_{15}\tilde{D}_9$
61	A9	$D_{8,1}^3$	360	14	$1, \frac{6}{7}, 1, 0$	$\tilde{D}_8^3$
60	A10	$A_{12,1}^2$	336	13	$1, 0$	$\tilde{A}_{12}^2$
59	A11	$A_{11,1}D_{7,1}E_{6,1}$	312	12	$1, 1, 1, 1, 1, 0$	$\tilde{A}_{11}\tilde{D}_7\tilde{E}_6$
58	A12	$E_{6,1}^4$	312	12	$1, 1, \frac{3}{4}, 1, 1, 0$	$\tilde{E}_6^4$
55	A13	$A_{9,1}^2D_{6,1}$	264	10	$1, 1, 1, 0$	$\tilde{A}_9^2\tilde{D}_6$
54	A14	$D_{6,1}^4$	264	10	$1, \frac{4}{5}, 1, 0$	$\tilde{D}_6^4$
51	A15	$A_{8,1}^3$	240	9	$1, 1, 0$	$\tilde{A}_8^3$
49	A16	$A_{7,1}^2D_{5,1}^2$	216	8	$1, 1, 1, 0$	$\tilde{A}_7^2\tilde{D}_5^2$
46	A17	$A_{6,1}^4$	192	7	$1, 0$	$\tilde{A}_6^4$
43	A18	$A_{5,1}^4D_{4,1}$	168	6	$1, 1, 1, 0$	$\tilde{A}_5^4\tilde{D}_4$
42	A19	$D_{4,1}^6$	168	6	$1, \frac{2}{3}, 1, 0$	$\tilde{D}_4^6$
37	A20	$A_{4,1}^6$	144	5	$1, 0$	$\tilde{A}_4^6$
30	A21	$A_{3,1}^8$	120	4	$1, 1, 0$	$\tilde{A}_3^8$
24	A22	$A_{2,1}^{12}$	96	3	$1, 0$	$\tilde{A}_2^{12}$
15	A23	$A_{1,1}^{24}$	72	2	$1, 0$	$\tilde{A}_1^{24}$
1	A24	$\mathbb{C}^{24}$	24	1	0	$\varnothing$

Table 2. The 70 generalised deep holes of V_Λ whose corresponding orbifold constructions realise all nonzero Lie algebras on Schellekens' list (continued below).

of $O(\pi_\nu(\Lambda))$, but in view of Proposition 2.1 it is important to consider the orbits under $C_{O(\Lambda)}(\nu)$ rather than under the full orthogonal group $O(\pi_\nu(\Lambda))$.

Then, since $\tilde{X}_l$ is of affine type, its centre h is uniquely determined by the concrete realisation of $\tilde{X}_l$ inside $\pi_\nu(\Lambda)$ (see Proposition 4.4). For each orbit, this immediately yields the complete hole diagram $\tilde{X}_l \ldots$ defined by h, which is some Dynkin diagram containing $\tilde{X}_l$ as a connected component.

Finally, by Proposition 2.1, each generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ defining a hole diagram containing $\tilde{X}_l$ in $\pi_\nu(\Lambda)$ must be conjugate to $g = \phi_\eta(\nu)\sigma_h$ where $\phi_\eta(\nu)$ is a standard lift of ν and h is one of the centres in the finite list of orbits.

Now, we go through the potential generalised deep holes in Tables 1 and 2 containing a connected affine component (54 plus three spurious cases) and show that the entries of Table 1 cannot be realised

number	number	$(V_\Lambda^{\mathrm{orb}(g)})_1$	dimension	n	$\rho(V_\Lambda(g^m))$	$\Phi(g)$
		rank 16, cycle shape $1^8 2^8$				
62	B1	$B_{8,1}E_{8,2}$	384	30	$1, \frac{14}{15}, 1, 1, 1, 1, 1, 0$	$A_1\tilde{E}_8$
56	B2	$B_{6,1}C_{10,1}$	288	22	$1, \frac{10}{11}, 1, 0$	A_1A_9
52	B3	$C_{8,1}F_{4,1}^2$	240	18	$1, \frac{8}{9}, 1, 1, 1, 0$	$A_2^2A_7$
53	B4	$B_{5,1}E_{7,2}F_{4,1}$	240	18	$1, \frac{8}{9}, 1, 1, 1, 0$	$A_1A_2\tilde{E}_7$
50	B5	$A_{7,1}D_{9,2}$	216	16	$1, 1, 1, 1, 0$	$\tilde{D}_9$
47	B6	$B_{4,1}^2D_{8,2}$	192	14	$1, \frac{6}{7}, 1, 0$	$A_1^2\tilde{D}_8$
48	B7	$B_{4,1}C_{6,1}^2$	192	14	$1, \frac{6}{7}, 1, 0$	$A_1A_5^2$
44	B8	$A_{5,1}C_{5,1}E_{6,2}$	168	12	$1, 1, 1, 1, 1, 0$	$A_4\tilde{E}_6$
40	B9	$A_{4,1}A_{9,2}B_{3,1}$	144	10	$1, 1, 1, 0$	$A_1\tilde{A}_9$
39	B10	$B_{3,1}^2C_{4,1}D_{6,2}$	144	10	$1, \frac{4}{5}, 1, 0$	$A_1^2A_3\tilde{D}_6$
38	B11	$C_{4,1}^4$	144	10	$1, \frac{4}{5}, 1, 0$	A_3^4
33	B12	$A_{3,1}A_{7,2}C_{3,1}^2$	120	8	$1, 1, 1, 0$	$A_2^2\tilde{A}_7$
31	B13	$A_{3,1}^2D_{5,2}^2$	120	8	$1, 1, 1, 0$	$\tilde{D}_5^2$
26	B14	$A_{2,1}^2A_{5,2}^2B_{2,1}$	96	6	$1, 1, 1, 0$	$A_1\tilde{A}_5^2$
25	B15	$B_{2,1}^4D_{4,2}^2$	96	6	$1, \frac{2}{3}, 1, 0$	$A_1^4\tilde{D}_4^2$
16	B16	$A_{1,1}^4A_{3,2}^4$	72	4	$1, 1, 0$	$\tilde{A}_3^4$
5	B17	$A_{1,2}^{16}$	48	2	$1, 0$	$\tilde{A}_1^{16}$
		rank 12, cycle shape $1^6 3^6$				
45	C1	$A_{5,1}E_{7,3}$	168	18	$1, 1, 1, 1, 1, 0$	$\tilde{E}_7$
34	C2	$A_{3,1}D_{7,3}G_{2,1}$	120	12	$1, 1, 1, 1, 1, 0$	$A_1\tilde{D}_7$
32	C3	$E_{6,3}G_{2,1}^3$	120	12	$1, 1, \frac{3}{4}, 1, 1, 0$	$A_1^3\tilde{E}_6$
27	C4	$A_{2,1}^2A_{8,3}$	96	9	$1, 1, 0$	$\tilde{A}_8$
17	C5	$A_{1,1}^3A_{5,3}D_{4,3}$	72	6	$1, 1, 1, 0$	$\tilde{A}_5\tilde{D}_4$
6	C6	$A_{2,3}^6$	48	3	$1, 0$	$\tilde{A}_2^6$
		rank 12, cycle shape 2^{12} (order doubling)				
57	D1a	$B_{12,2}$	300	46	$1, \frac{22}{23}, 1, 0$	A_1
41	D1b	$B_{6,2}^2$	156	22	$1, \frac{10}{11}, 1, 0$	A_1^2
29	D1c	$B_{4,2}^3$	108	14	$1, \frac{6}{7}, 1, 0$	A_1^3
23	D1d	$B_{3,2}^4$	84	10	$1, \frac{4}{5}, 1, 0$	A_1^4
12	D1e	$B_{2,2}^6$	60	6	$1, \frac{2}{3}, 1, 0$	A_1^6
2	D1f	$A_{1,4}^{12}$	36	2	$1, 0$	$\tilde{A}_1^{12}$
36	D2a	$A_{8,2}F_{4,2}$	132	18	$1, 1, 1, 1, 1, 0$	A_2
22	D2b	$A_{4,2}^2C_{4,2}$	84	10	$1, 1, 1, 0$	A_3
13	D2c	$A_{2,2}^4D_{4,4}$	60	6	$1, 1, 1, 0$	$\tilde{D}_4$

Table 2. (continued).

number	number	$(V_\Lambda^{\mathrm{orb}(g)})_1$	dimension	n	$\rho(V_\Lambda(g^m))$	$\Phi(g)$
rank 10, cycle shape $1^4 2^2 4^4$						
35	E1	$A_{3,1}C_{7,2}$	120	16	$1,1,1,1,0$	A_6
28	E2	$A_{2,1}B_{2,1}E_{6,4}$	96	12	$1,1,1,1,1,0$	$\tilde{E}_6$
18	E3	$A_{1,1}^3 A_{7,4}$	72	8	$1,1,1,0$	$\tilde{A}_7$
19	E4	$A_{1,1}^2 C_{3,2} D_{5,4}$	72	8	$1,1,1,0$	$A_2\tilde{D}_5$
7	E5	$A_{1,2}A_{3,4}^3$	48	4	$1,1,0$	$\tilde{A}_3^3$
rank 8, cycle shape $1^4 5^4$						
20	F1	$A_{1,1}^2 D_{6,5}$	72	10	$1,1,1,0$	$\tilde{D}_6$
9	F2	$A_{4,5}^2$	48	5	$1,0$	$\tilde{A}_4^2$
rank 8, cycle shape $1^2 2^2 3^2 6^2$						
21	G1	$A_{1,1}C_{5,3}G_{2,2}$	72	12	$1,1,1,1,1,0$	A_1A_4
8	G2	$A_{1,2}A_{5,6}B_{2,3}$	48	6	$1,1,1,0$	$A_1\tilde{A}_5$
rank 6, cycle shape $1^3 7^3$						
11	H1	$A_{6,7}$	48	7	$1,0$	$\tilde{A}_6$
rank 6, cycle shape $1^2 2^1 4^1 8^2$						
10	I1	$A_{1,2}D_{5,8}$	48	8	$1,1,1,0$	$\tilde{D}_5$
rank 6, cycle shape $2^3 6^3$ (order doubling)						
14	J1a	$A_{2,2}F_{4,6}$	60	18	$1,1,1,1,1,0$	A_2
3	J1b	$A_{2,6}D_{4,12}$	36	6	$1,1,1,0$	$\tilde{D}_4$
rank 4, cycle shape $2^2 10^2$ (order doubling)						
4	K1	$C_{4,10}$	36	10	$1,1,1,0$	A_3

Table 2. (continued).

by generalised deep holes while the candidates of Table 2 by at most one algebraic conjugacy class in $\mathrm{Aut}(V_\Lambda)$.

In analogy to [12], we introduce the notation

$$\tilde{X}_l \Longrightarrow \tilde{X}_l \ldots$$

to mean that there is a unique (unless otherwise noted) orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ of the connected affine diagram $\tilde{X}_l$ in $\pi_\nu(\Lambda)$ (as lattice points sitting on a sphere of radius $\sqrt{2(1-\rho_\nu)}$ around the centre of $\tilde{X}_l$) and that it defines the complete diagram $\tilde{X}_l \ldots$ (*all* the lattice points sitting on said sphere). If there are several orbits, each defining a different diagram $\tilde{X}_l \ldots$, we shall separate these by *or*. If $\tilde{X}_l$ does not appear at all in $\pi_\nu(\Lambda)$, we write $\tilde{X}_l \Longrightarrow \varnothing$.

The first case was already covered in [12]:

Lemma 5.4 [12]. *Let $\nu \in O(\Lambda)$ be of cycle shape 1^{24}. Then in $\pi_\nu(\Lambda) = \Lambda$:*

$\tilde{A}_1 \Longrightarrow \tilde{A}_1^{24}$, $\tilde{A}_2 \Longrightarrow \tilde{A}_2^{12}$, $\tilde{A}_3 \Longrightarrow \tilde{A}_3^{8}$, $\tilde{A}_4 \Longrightarrow \tilde{A}_4^{6}$,
$\tilde{A}_5 \Longrightarrow \tilde{A}_5^{4}\tilde{D}_4$, $\tilde{A}_6 \Longrightarrow \tilde{A}_6^{4}$, $\tilde{A}_7 \Longrightarrow \tilde{A}_7^{2}\tilde{D}_5^{2}$, $\tilde{A}_8 \Longrightarrow \tilde{A}_8^{3}$,
$\tilde{A}_9 \Longrightarrow \tilde{A}_9^{2}\tilde{D}_6$, $\tilde{A}_{11} \Longrightarrow \tilde{A}_{11}\tilde{D}_7\tilde{E}_6$, $\tilde{A}_{12} \Longrightarrow \tilde{A}_{12}^{2}$, $\tilde{A}_{15} \Longrightarrow \tilde{A}_{15}\tilde{D}_9$,
$\tilde{A}_{17} \Longrightarrow \tilde{A}_{17}\tilde{E}_7$, $\tilde{A}_{24} \Longrightarrow \tilde{A}_{24}$, $\tilde{D}_4 \Longrightarrow \tilde{D}_4^{6}$ *or* $\tilde{A}_5^{4}\tilde{D}_4$, $\tilde{D}_5 \Longrightarrow \tilde{A}_7^{2}\tilde{D}_5^{2}$,
$\tilde{D}_6 \Longrightarrow \tilde{D}_6^{4}$ *or* $\tilde{A}_9^{2}\tilde{D}_6$, $\tilde{D}_7 \Longrightarrow \tilde{A}_{11}\tilde{D}_7\tilde{E}_6$, $\tilde{D}_8 \Longrightarrow \tilde{D}_8^{3}$, $\tilde{D}_9 \Longrightarrow \tilde{A}_{15}\tilde{D}_9$,
$\tilde{D}_{10} \Longrightarrow \tilde{D}_{10}\tilde{E}_7^{2}$, $\tilde{D}_{12} \Longrightarrow \tilde{D}_{12}^{2}$, $\tilde{D}_{16} \Longrightarrow \tilde{D}_{16}\tilde{E}_8$, $\tilde{D}_{24} \Longrightarrow \tilde{D}_{24}$,
$\tilde{E}_6 \Longrightarrow \tilde{E}_6^{4}$ *or* $\tilde{A}_{11}\tilde{D}_7\tilde{E}_6$, $\tilde{E}_7 \Longrightarrow \tilde{D}_{10}\tilde{E}_7^{2}$ *or* $\tilde{A}_{17}\tilde{E}_7$, $\tilde{E}_8 \Longrightarrow \tilde{E}_8^{3}$ *or* $\tilde{D}_{16}\tilde{E}_8$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *for each of the* 23 *nonempty hole diagrams listed in Table 2.*

Lemma 5.5. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^8 2^8$. Then in $\pi_\nu(\Lambda)$:*

$\tilde{A}_1 \Longrightarrow \tilde{A}_1^{8}$ *or* $\tilde{A}_1^{16}$, $\tilde{A}_3 \Longrightarrow A_1^4\tilde{A}_3^{2}$ *or* $\tilde{A}_3^{4}$,
$\tilde{A}_5 \Longrightarrow A_1^2 A_3\tilde{A}_5$ *or* $A_1\tilde{A}_5^{2}$, $\tilde{A}_7 \Longrightarrow A_1^2\tilde{A}_7$ (*at most* 2 *orbits*) *or* $A_2^2\tilde{A}_7$,
$\tilde{A}_9 \Longrightarrow A_1\tilde{A}_9$, $\tilde{D}_4 \Longrightarrow A_1^8\tilde{D}_4$ *or* $\tilde{D}_4^2$ *or* $A_1^4\tilde{D}_4^2$,
$\tilde{D}_5 \Longrightarrow D_4\tilde{D}_5$ *or* $\tilde{D}_5^2$, $\tilde{D}_6 \Longrightarrow A_1^4\tilde{D}_6$ *or* $A_1^2 A_3\tilde{D}_6$,
$\tilde{D}_8 \Longrightarrow \tilde{D}_8$ (*at most* 2 *orbits*) *or* $A_1^2\tilde{D}_8$, $\tilde{D}_9 \Longrightarrow \tilde{D}_9$,
$\tilde{E}_6 \Longrightarrow A_3\tilde{E}_6$ *or* $A_4\tilde{E}_6$, $\tilde{E}_7 \Longrightarrow A_1^2\tilde{E}_7$ *or* $A_1A_2\tilde{E}_7$,
$\tilde{E}_8 \Longrightarrow \tilde{E}_8$ *or* $A_1\tilde{E}_8$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *for each of the hole diagrams* $A_1\tilde{E}_8$, $A_1A_2\tilde{E}_7$, $\tilde{D}_9$, $A_1^2\tilde{D}_8$, $A_4\tilde{E}_6$, $A_1\tilde{A}_9$, $A_1^2A_3\tilde{D}_6$, $A_2^2\tilde{A}_7$, $\tilde{D}_5^2$, $A_1\tilde{A}_5^2$, $A_1^4\tilde{D}_4^2$, $\tilde{A}_3^4$ *and* $\tilde{A}_1^{16}$.

We can explicitly check, for instance, that the automorphism $g = \phi_\eta(\nu)\sigma_h$ defined by the diagram $\tilde{A}_1^{16}$ and its centre h is a generalised deep hole, while for the diagram $\tilde{A}_1^{8}$ this is not the case.

Lemma 5.6. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^6 3^6$. Then in $\pi_\nu(\Lambda)$:*

$\tilde{A}_2 \Longrightarrow \tilde{A}_2^{3}$ *or* $\tilde{A}_2^{6}$, $\tilde{A}_5 \Longrightarrow A_2\tilde{A}_5$ (*at most* 2 *orbits*) *or* $\tilde{A}_5\tilde{D}_4$, $\tilde{A}_8 \Longrightarrow \tilde{A}_8$, $\tilde{D}_4 \Longrightarrow A_2^2\tilde{D}_4$ *or* $\tilde{A}_5\tilde{D}_4$,
$\tilde{D}_7 \Longrightarrow A_1\tilde{D}_7$, $\tilde{E}_6 \Longrightarrow \tilde{E}_6$ (*at most* 2 *orbits*) *or* $A_1^3\tilde{E}_6$, $\tilde{E}_7 \Longrightarrow \tilde{E}_7$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *for each of the hole diagrams* $\tilde{E}_7$, $A_1\tilde{D}_7$, $A_1^3\tilde{E}_6$, $\tilde{A}_8$, $\tilde{A}_5\tilde{D}_4$ *and* $\tilde{A}_2^6$.

For the cycle shape 2^{12} we first remove two more spurious cases.

Lemma 5.7. *Let $\nu \in O(\Lambda)$ be of cycle shape 2^{12}. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_3 \Longrightarrow \tilde{A}_3, \quad \tilde{D}_5 \Longrightarrow \varnothing.$$

In particular, there is no generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *projecting to* ν *with hole diagram* $\tilde{A}_3^2$ *or* $\tilde{D}_5$.

Lemma 5.8. *Let $\nu \in O(\Lambda)$ be of cycle shape 2^{12}. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_1 \Longrightarrow \tilde{A}_1^4 \text{ or } \tilde{A}_1^{12}, \quad \tilde{D}_4 \Longrightarrow \tilde{D}_4 \text{ (2 orbits)}.$$

One of the two orbits of type $\tilde{D}_4$ has a centre defining an automorphism of order 12, *the other one an automorphism of order* 6.

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν for each of the hole diagrams $\tilde{A}_1^{12}$ and $\tilde{D}_4$.*

Lemma 5.9. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^4 2^2 4^4$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_3 \Longrightarrow \tilde{A}_3^2 \text{ or } A_1^2\tilde{A}_3 \text{ (at most 2 orbits) or } A_1^4\tilde{A}_3 \text{ or } \tilde{A}_3^3, \quad \tilde{A}_7 \Longrightarrow \tilde{A}_7, \quad \tilde{D}_5 \Longrightarrow \tilde{D}_5 \text{ or } A_2\tilde{D}_5, \quad \tilde{E}_6 \Longrightarrow \tilde{E}_6.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν for each of the hole diagrams $\tilde{E}_6$, $\tilde{A}_7$, $A_2\tilde{D}_5$ and $\tilde{A}_3^3$.*

Lemma 5.10. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^4 5^4$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_4 \Longrightarrow \tilde{A}_4 \text{ (at most 3 orbits) or } \tilde{A}_4^2, \quad \tilde{D}_6 \Longrightarrow \tilde{D}_6.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν for each of the hole diagrams $\tilde{D}_6$ and $\tilde{A}_4^2$.*

We remove the spurious case for the cycle shape $1^2 2^2 3^2 6^2$.

Lemma 5.11. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^2 2^2 3^2 6^2$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{D}_4 \Longrightarrow \varnothing$$

In particular, there is no generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *projecting to ν with hole diagram $A_1^2\tilde{D}_4$.*

Lemma 5.12. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^2 2^2 3^2 6^2$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_5 \Longrightarrow A_1\tilde{A}_5.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν with hole diagram $A_1\tilde{A}_5$.*

Lemma 5.13. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^3 7^3$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{A}_6 \Longrightarrow \tilde{A}_6.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν with hole diagram $\tilde{A}_6$.*

Lemma 5.14. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^2 2^1 4^1 8^2$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{D}_5 \Longrightarrow \tilde{D}_5.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν with hole diagram $\tilde{D}_5$.*

Lemma 5.15. *Let $\nu \in O(\Lambda)$ be of cycle shape $2^3 6^3$. Then in $\pi_\nu(\Lambda)$:*

$$\tilde{D}_4 \Longrightarrow \tilde{D}_4.$$

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν with hole diagram $\tilde{D}_4$.*

5.2. ***Nonaffine case.*** We now consider the more difficult case of potential generalised deep holes g with hole diagrams that do not contain any affine component. These are 15 plus two spurious cases (see Table 2 and Table 1).

First, we enumerate the orbits under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ of the diagram realised in $\pi_\nu(\Lambda)$ as lattice points with relative distances defined as in Proposition 4.3. This is the same computation as in the affine case, with the exception that we are now directly searching for the complete diagram. Again, this is a relatively cheap computation using the short vector search in `Magma` [5] (code made available at arxiv.org).

The points of the hole diagram must lie on a sphere of radius $\sqrt{2(1-\rho_\nu)}$ around some centre $h \in \pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q})$. However, in contrast to the affine case, this centre is not uniquely determined by the diagram. (In the most extreme case of the diagram A_1, h could be any point at distance $\sqrt{2(1-\rho_\nu)}$ from the single vertex defining A_1.) The second and generally computationally more expensive part is to determine all the possible $h \in \pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q})$ that could be the centre of the diagram.

We employ two methods to facilitate this search. First, from Proposition 5.2 we know the order n of the generalised deep hole (see Tables 1 and 2). Then Proposition 2.2 implies that h must lie in $s_\nu + \Lambda^\nu/n$ (where s_ν is nonzero if and only if ν exhibits order doubling). Second, as h must have distance $\sqrt{2(1-\rho_\nu)}$ to all the vertices in the hole diagram, it must in particular lie on the hyperplanes of points equidistant to all pairs of vertices. This reduces the dimension of the eventual close vector search to find h, which is again performed in `Magma` [5].

As a result, for each orbit of the original diagram search we obtain a finite list of possible centres h. We then only keep those h

(1) whose corresponding automorphism $g = \phi_\eta(\nu)\sigma_h$ with standard lift $\phi_\eta(\nu)$ has order n (i.e., g must satisfy $\operatorname{lcm}(|\nu|, |\sigma_{h-s_\nu}|) = n$),

(2) such that $\tilde{\Pi}(g) = \{\beta \in \pi_\nu(\Lambda) \mid \langle \beta - h, \beta - h\rangle/2 = 1 - \rho_\nu\}$ has exactly the diagram we are searching for (a priori we only know that it contains this diagram as a subdiagram),

(3) that actually correspond to a generalised deep hole $g = \phi_\eta(\nu)\sigma_h$ (in particular, g must be extremal).

Again, we sort the results by cycle shape and treat the case 2^{12} last because it is the most complicated one.

Lemma 5.16. *Let $\nu \in O(\Lambda)$ be of cycle shape $1^8 2^8$. Then in $\pi_\nu(\Lambda)$:*

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ of the diagram $A_1 A_9$. There are 16 *possible centres $h \in \Lambda^\nu/22$ of this diagram, but only one of them satisfies (1) to (3).*

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* $A_2^2 A_7$. *There are six possible centres* $h \in \Lambda^\nu/18$ *of this diagram, but only two of them satisfy (1) to (3). They are both in the same orbit under* $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* $A_1 A_5^2$. *There are seven possible centres* $h \in \Lambda^\nu/14$ *of this diagram, but only one of them satisfies (1) to (3).*

There are exactly two orbits under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* A_3^4. *For the first orbit there is one possible centre* $h \in \Lambda^\nu/10$ *and it satisfies (1) to (3). For the second orbit there are six possible centres* $h \in \Lambda^\nu/10$, *but none of them satisfy (1) to (3).*

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *for each of the hole diagrams* $A_1 A_9$, $A_2^2 A_7$, $A_1 A_5^2$ *and* A_3^4.

Lemma 5.17. *Let* $\nu \in O(\Lambda)$ *be of cycle shape* $1^4 2^2 4^4$. *Then in* $\pi_\nu(\Lambda)$:

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* A_6. *There are nine possible centres* $h \in \Lambda^\nu/16$ *of this diagram, but only one of them satisfies (1) to (3).*

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *with hole diagram* A_6.

Lemma 5.18. *Let* $\nu \in O(\Lambda)$ *be of cycle shape* $1^2 2^2 3^2 6^2$. *Then in* $\pi_\nu(\Lambda)$:

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* $A_1 A_4$. *There are seven possible centres* $h \in \Lambda^\nu/12$ *of this diagram, but only two of them satisfy (1) to (3). They are both in the same orbit under* $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *with hole diagram* $A_1 A_4$.

Lemma 5.19. *Let* $\nu \in O(\Lambda)$ *be of cycle shape* $2^3 6^3$. *Then in* $\pi_\nu(\Lambda)$:

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* A_2. *There are* 98 *possible centres* $h \in s_\nu + \Lambda^\nu/18$ *of this diagram, but only* 36 *of them satisfy (1) to (3). They are all in the same orbit under* $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *with hole diagram* A_2.

Lemma 5.20. *Let* $\nu \in O(\Lambda)$ *be of cycle shape* $2^2 10^2$. *Then in* $\pi_\nu(\Lambda)$:

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ *of the diagram* A_3. *There are two possible centres* $h \in s_\nu + \Lambda^\nu/10$ *of this diagram and both of them satisfy (1) to (3). They are both in the same orbit under* $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.

In particular, there is at most one generalised deep hole in $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to* ν *with hole diagram* A_3.

Now we consider the case 2^{12}. We start by excluding the last two spurious cases.

Lemma 5.21. *Let $\nu \in O(\Lambda)$ be of cycle shape 2^{12}. There is no hole diagram in $\pi_\nu(\Lambda)$ containing A_2^3. In particular, there is no generalised deep hole in* $\operatorname{Aut}(V_\Lambda)$ *projecting to ν with hole diagram A_2^3.*

Lemma 5.22. *Let $\nu \in O(\Lambda)$ be of cycle shape 2^{12}. Then in $\pi_\nu(\Lambda)$:*

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ of the diagram A_2. There are 9,132,200 *possible centres $h \in s_\nu + \Lambda^\nu/18$ of this diagram, but only* 31,680 *of them satisfy (1) to (3). They are all in the same orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.*

There is exactly one orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$ of the diagram A_3. There are 432 *possible centres $h \in \Lambda^\nu/10$ of this diagram, but only* 72 *of them satisfy (1) to (3). They are all in the same orbit under $C_{O(\Lambda)}(\nu) \ltimes \pi_\nu(\Lambda)$.*

In particular, there is at most one generalised deep hole in $\operatorname{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν for each of the hole diagrams A_2 and A_3.*

We now discuss the most difficult cases, the generalised deep holes with cycle shape 2^{12} and hole diagrams A_1, A_1^2, A_1^3, A_1^4 and A_1^6. The hardest case is the potential generalised deep hole of order 46 with hole diagram A_1. Here, the hole diagram merely implies that the vector h is the only and closest point at distance $1-\rho_\nu = \frac{1}{4}$ from the single vertex $\beta_1 \in \pi_\nu(\Lambda)$ defining A_1.

Fortunately, we can exploit that the fixed-point lattice Λ^ν has a very symmetric embedding into Euclidean space. Indeed, let D_{12}^+ denote the positive-definite, integral lattice

$$D_{12}^+ := \left\{ (x_1, \ldots, x_{12}) \in \mathbb{R}^{12} \;\middle|\; \text{all } x_i \in \mathbb{Z} \text{ or all } x_i \in \mathbb{Z} + \tfrac{1}{2} \text{ and } \sum_{i=1}^{12} x_i \in 2\mathbb{Z} \right\}$$

embedded into Euclidean space $\mathbb{R}^{12}$ with the standard scalar product. It is the unique indecomposable, positive-definite, integral, unimodular lattice of rank 12. Let $K := \sqrt{2}D_{12}^+$ denote the lattice with lattice vectors scaled by $\sqrt{2}$. Then K is even and

$$\Lambda^\nu \cong K.$$

We note that

$$\pi_\nu(\Lambda) = (\Lambda^\nu)' = \frac{\Lambda^\nu}{2}.$$

The first equality holds since Λ is unimodular, but the second equality (which is not just an isomorphism but a proper equality) is a special property of ν.

The automorphism group of D_{12}^+ (and of K) is generated by permutations and even sign changes, that is,

$$O(K) = S_{12} \ltimes 2^{11}.$$

The kernel of the map $C_{O(\Lambda)}(\nu) \to O(\Lambda^\nu) \cong O(K)$ has order 2 and is generated by ν. The image has index 5040 and is of the form $P \ltimes 2^{11}$ where P is some permutation group of index 5040 in S_{12}.

In the following, we want to show that there is a unique $h \in \pi_\nu(\Lambda \otimes_\mathbb{Z} \mathbb{Q})/\pi_\nu(\Lambda) \cong (K \otimes_\mathbb{Z} \mathbb{Q})/(K/2)$ with a certain list of properties up to the action of $C_{O(\Lambda)}(\nu)$. The number of elements we have to consider is too big to be amenable to a brute-force approach. We therefore split the computation into three parts, first considering only properties invariant under the much bigger group $S_{12} \ltimes 2^{12}$ (where we allow all sign changes) and computing the orbits satisfying these, then under the group $O(K) = S_{12} \ltimes 2^{11}$ and finally under the group $C_{O(\Lambda)}(\nu)$, that is, $P \ltimes 2^{11}$. In each step the number of orbits we consider remains manageable.

Lemma 5.23. *Let $\nu \in O(\Lambda)$ be of cycle shape 2^{12}. Then there is at most one generalised deep hole in* $\mathrm{Aut}(V_\Lambda)$ *up to conjugacy projecting to ν for each of the hole diagrams A_1, A_1^2, A_1^3, A_1^4 and A_1^6.*

Proof. We only describe the hardest case of the diagram A_1. The other cases are treated analogously.

A generalised deep hole with hole diagram A_1 has order 46 and is conjugate to $g = \phi_\eta(\nu)\sigma_h$ for some standard lift $\phi_\eta(\nu)$ and $h \in s_\nu + \Lambda^\nu/46 \subseteq \Lambda^\nu/92$. By applying a translation in $\pi_\nu(\Lambda)$, we may assume that the only vertex of the hole diagram A_1 in $\pi_\nu(\Lambda)$ is the origin. Then h has the properties

(1) $h \in \Lambda^\nu/92$,

(2) $\langle h, h\rangle/2 = \frac{1}{4}$,

(3) $\langle h-\beta, h-\beta\rangle/2 \geq \frac{1}{4}$ for all $\beta \in \pi_\nu(\Lambda)$,

(4) $\langle h-\beta, h-\beta\rangle/2 = \frac{1}{4}$ and $\beta \in \pi_\nu(\Lambda)$ if and only if $\beta = 0$.

We consider $\tilde{h} := 92h$. The above conditions are equivalent to

(1) $\tilde{h} \in \Lambda^\nu$,

(2) $\langle \tilde{h}, \tilde{h}\rangle/2 = 46^2$,

(3) $\langle \tilde{h}, \beta\rangle \leq 46\langle \beta, \beta\rangle/2$ for all $\beta \in \Lambda^\nu$,

(4) $\langle \tilde{h}, \beta\rangle = 46\langle \beta, \beta\rangle/2$ and $\beta \in \Lambda^\nu$ if and only if $\beta = 0$.

We identify Λ^ν with K and write $\tilde{h} = \sqrt{2}(h_1, \ldots, h_{12})$. Then the first condition is equivalent to either all $h_i \in \mathbb{Z}$ or all $h_i \in \mathbb{Z} + \frac{1}{2}$, and moreover $\sum_{i=1}^{12} h_i \in 2\mathbb{Z}$. We actually know that $\tilde{h} \in 92s_\nu + 2\Lambda^\nu$ for some $s_\nu \in \Lambda^\nu/4$ (such that $\phi_\eta(\nu)\sigma_{s_\nu}$ has order 2). By choosing an s_ν we see that either all $h_i \in 2\mathbb{Z}$ or all $h_i \in 2\mathbb{Z}+1$. In total, the above conditions imply

(1′) all $h_i \in 2\mathbb{Z}$ or all $h_i \in 2\mathbb{Z}+1$,

(2′) $\sum_{i=1}^{12} h_i^2 = 46^2$,

(3′) $|h_i| + |h_j| < 46$ for $i \neq j$.

We determine the orbits of the solutions of these three conditions up to the action of $S_{12} \ltimes 2^{12}$, i.e., we ignore signs and permutations. This is a simple combinatorial problem with 10, 301 solutions.

We then consider the corresponding orbits under $O(K) = S_{12} \ltimes 2^{11}$, i.e., each orbit represented by a sequence $(h_1, \ldots, h_{12})$ not containing a 0 splits up into two orbits by introducing a sign at, e.g., the first entry. The fact that g is extremal implies that the twisted modules $V(g)$, $V(g^5)$, $V(g^9)$, $V(g^{13})$, $V(g^{17})$ and $V(g^{21})$ each have conformal weight at least 1. Since $\phi_\eta(\nu)^4 = \mathrm{id}$, it follows that $g^{4k+1} = \phi_\eta(\nu)\sigma_{(4k+1)h}$

so that these conditions translate to

$$\min_{\beta\in 46\Lambda^\nu} \frac{\langle (4k+1)\tilde{h}-\beta, (4k+1)\tilde{h}-\beta\rangle}{2} = 46^2 \tag{4'}$$

for $k = 0, 1, \ldots, 11$. In fact, for $k = 0$ we require equality and that there is exactly one closest vector, namely the one forming the diagram A_1. These conditions are invariant under $O(K) = S_{12} \ltimes 2^{11}$. The result is that there is exactly one orbit under $O(K)$ satisfying conditions (1′) to (4′), namely (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 24).

Finally, we split up this orbit into the orbits under the action of the centraliser $C_{O(\Lambda)}(\nu)$, that is, under $P \ltimes 2^{11}$. In this case, since all the h_i are distinct, these orbits are in natural bijection with the 5040 cosets of P in S_{12}, which can be computed using `GAP` [31]. For these orbits we then explicitly check if they can be generalised deep holes of order 46, that is, in particular, extremal. In the end, this leaves us with just one orbit $h \in \pi_\nu(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q})/\pi_\nu(\Lambda)$ under the action of $C_{O(\Lambda)}(\nu)$, which concludes the proof. □

We remark that the generalised deep holes (of order n) for the diagrams A_1, A_1^2, A_1^3, A_1^4 and A_1^6 correspond to the vectors $h = \sqrt{2}(h_1, \ldots, h_{12})/(2n)$ in $K/(2n) \subseteq \mathbb{R}^{12}$ specified by the following h_i:

A_1	0	2	4	6	8	10	12	14	16	18	20	24
A_1^2	0	0	2	2	4	4	6	6	8	8	10	12
A_1^3	0	0	0	2	2	2	4	4	4	6	6	8
A_1^4	0	0	0	0	2	2	2	2	4	4	4	6
A_1^6	0	0	0	0	0	0	2	2	2	2	2	4
$\tilde{A}_1^{12}$	0	0	0	0	0	0	0	0	0	0	0	2

Here, we ignore signs and the order of the entries, which in any case depend on the concrete choice of the isomorphism $\Lambda^\nu \cong K$.

5.3. *Classification results.* We summarise the above results:

Proposition 5.24. *There are at most* 70 *conjugacy classes of generalised deep holes g in* $\mathrm{Aut}(V_\Lambda)$ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$. *They are described in Table 2.*

In [50] we list 70 generalised deep holes g in $\mathrm{Aut}(V_\Lambda)$ with $\mathrm{rk}((V_\Lambda^g)_1) > 0$. Using Proposition 4.5 we can easily determine their generalised hole diagrams, which are all distinct. This implies the main result:

Theorem 5.25 (classification of generalised deep holes). *There are exactly* 70 *conjugacy classes of generalised deep holes g in* $\mathrm{Aut}(V_\Lambda)$ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$. *The conjugacy class of g is uniquely fixed by its generalised hole diagram.*

An automorphism g of order n is called *rational* if g is conjugate to g^i for all $i \in \mathbb{Z}_n$ with $(i, n) = 1$ (see, e.g., Chapter 7 in [56]). Equivalently, the conjugacy class and the algebraic conjugacy class (i.e., the conjugacy class of the cyclic subgroup) of g coincide. The following observation is immediate:

Corollary 5.26. *The generalised deep holes g in* $\mathrm{Aut}(V_\Lambda)$ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$ *are rational*, *that is*, *conjugacy is equivalent to algebraic conjugacy.*

We also recover the decomposition of the Schellekens vertex operator algebras into 12 families described by Höhn in [32] (see also [52; 53]):

Theorem 5.27 (projection to Co_0). *Under the natural projection* $\mathrm{Aut}(V_\Lambda) \to O(\Lambda)$ *the generalised deep holes* g *of* V_Λ *with* $\mathrm{rk}((V_\Lambda^g)_1) > 0$ *map to the* 11 *algebraic conjugacy classes in* $O(\Lambda) \cong \mathrm{Co}_0$ *with cycle shapes* 1^{24}, $1^8 2^8$, $1^6 3^6$, 2^{12}, $1^4 2^2 4^4$, $1^4 5^4$, $1^2 2^2 3^2 6^2$, $1^3 7^3$, $1^2 2^1 4^1 8^2$, $2^3 6^3$ *and* $2^2 10^2$.

A consequence of the above classification of generalised deep holes and the holey correspondence in [50] is a new, geometric proof of the following result:

Theorem 5.28 (classification of vertex operator algebras). *Up to isomorphism there are exactly* 70 *strongly rational, holomorphic vertex operator algebras* V *of central charge* 24 *with* $V_1 \neq \{0\}$. *Such a vertex operator algebra is uniquely determined by its* V_1*-structure.*

We have thus obtained a geometric proof of this classification that is analogous to the classification of the Niemeier lattices by enumeration of the corresponding deep holes of the Leech lattice Λ [3; 12]. In fact, it includes it as a special case (see Proposition 4.7).

We mention that [38] give an interpretation of the generalised deep holes of V_Λ in terms of actual deep holes of Λ after rescaling.

We remark that Höhn's approach to the classification problem in [32] (and [36]) based on coset constructions can also be used to give a uniform proof of the above classification result [2; 32].

Acknowledgements

The authors thank Tomoyuki Arakawa and Gerald Höhn for valuable discussions, and the referees for helpful comments. Sven Möller was supported by a JSPS *Postdoctoral Fellowship for Research in Japan* and by JSPS Grant-in-Aid KAKENHI 20F40018. Nils Scheithauer acknowledges support by the LOEWE research unit *Uniformized Structures in Arithmetic and Geometry* and by the DFG through the CRC *Geometry and Arithmetic of Uniformized Structures*, project number 444845124.

References

[1] B. Bakalov and V. G. Kac, "Twisted modules over lattice vertex algebras", pp. 3–26 in *Lie theory and its applications in physics*, *V* (Varna, Bulgaria, 2003), edited by H.-D. Doebner and V. K. Dobrev, World Sci., River Edge, NJ, 2004. MR Zbl

[2] K. Betsumiya, C. H. Lam, and H. Shimakura, "Automorphism groups and uniqueness of holomorphic vertex operator algebras of central charge 24", *Comm. Math. Phys.* **399**:3 (2023), 1773–1810. MR Zbl

[3] R. E. Borcherds, "The Leech lattice", *Proc. Roy. Soc. Lond. Ser. A* **398**:1815 (1985), 365–376. MR Zbl

[4] R. E. Borcherds, "Vertex algebras, Kac–Moody algebras, and the Monster", *Proc. Nat. Acad. Sci. U.S.A.* **83**:10 (1986), 3068–3071. MR Zbl

[5] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[6] S. Carnahan and M. Miyamoto, "Regularity of fixed-point vertex operator subalgebras", preprint, 2016. arXiv 1603.05645

[7] G. Chenevier and J. Lannes, *Automorphic forms and even unimodular lattices: Kneser neighbors of Niemeier lattices*, Ergebnisse der Math. (3) **69**, Springer, 2019. MR Zbl

[8] J. H. Conway, "A characterisation of Leech's lattice", *Invent. Math.* **7** (1969), 137–142. MR Zbl

[9] J. H. Conway and N. J. A. Sloane, "On the enumeration of lattices of determinant one", *J. Number Theory* **15**:1 (1982), 83–94. MR Zbl

[10] J. H. Conway and N. J. A. Sloane, "Twenty-three constructions for the Leech lattice", *Proc. Roy. Soc. Lond. Ser. A* **381**:1781 (1982), 275–283. MR Zbl

[11] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Grundl. Math. Wissen. **290**, Springer, 1988. MR Zbl

[12] J. H. Conway, R. A. Parker, and N. J. A. Sloane, "The covering radius of the Leech lattice", *Proc. Roy. Soc. Lond. Ser. A* **380**:1779 (1982), 261–290. MR Zbl

[13] R. Dijkgraaf, C. Vafa, E. Verlinde, and H. Verlinde, "The operator algebra of orbifold models", *Comm. Math. Phys.* **123**:3 (1989), 485–526. MR Zbl

[14] R. Dijkgraaf, V. Pasquier, and P. Roche, "Quasi Hopf algebras, group cohomology and orbifold models", *Nuclear Phys. B Proc. Suppl.* **18B** (1990), 60–72. MR Zbl

[15] C. Dong and J. Lepowsky, "The algebraic structure of relative twisted vertex operators", *J. Pure Appl. Algebra* **110**:3 (1996), 259–295. MR Zbl

[16] C. Dong and G. Mason, "On quantum Galois theory", *Duke Math. J.* **86**:2 (1997), 305–321. MR Zbl

[17] C. Dong and G. Mason, "Holomorphic vertex operator algebras of small central charge", *Pacific J. Math.* **213**:2 (2004), 253–266. MR Zbl

[18] C. Dong and G. Mason, "Rational vertex operator algebras and the effective central charge", *Int. Math. Res. Not.* **2004**:56 (2004), 2989–3008. MR Zbl

[19] C. Dong and G. Mason, "Integrability of C_2-cofinite vertex operator algebras", *Int. Math. Res. Not.* **2006** (2006), art. id. 80468. MR Zbl

[20] C. Dong and K. Nagatomo, "Automorphism groups and twisted modules for lattice vertex operator algebras", pp. 117–133 in *Recent developments in quantum affine algebras and related topics* (Raleigh, NC, 1998), edited by N. Jing and K. C. Misra, Contemp. Math. **248**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[21] C. Dong, H. Li, and G. Mason, "Regularity of rational vertex operator algebras", *Adv. Math.* **132**:1 (1997), 148–166. MR Zbl

[22] C. Dong, H. Li, and G. Mason, "Modular-invariance of trace functions in orbifold theory and generalized moonshine", *Comm. Math. Phys.* **214**:1 (2000), 1–56. MR Zbl

[23] C. Dong, L. Ren, and F. Xu, "On orbifold theory", *Adv. Math.* **321** (2017), 1–30. MR Zbl

[24] C. Dong, S.-H. Ng, and L. Ren, "Orbifolds and minimal modular extensions", preprint, 2021. arXiv 2108.05225

[25] J. van Ekeren, S. Möller, and N. R. Scheithauer, "Construction and classification of holomorphic vertex operator algebras", *J. Reine Angew. Math.* **759** (2020), 61–99. MR Zbl

[26] J. van Ekeren, S. Möller, and N. R. Scheithauer, "Dimension formulae in genus zero and uniqueness of vertex operator algebras", *Int. Math. Res. Not.* **2020**:7 (2020), 2145–2204. MR Zbl

[27] J. van Ekeren, C. H. Lam, S. Möller, and H. Shimakura, "Schellekens' list and the very strange formula", *Adv. Math.* **380** (2021), art. id. 107567. MR Zbl

[28] I. B. Frenkel and Y. Zhu, "Vertex operator algebras associated to representations of affine and Virasoro algebras", *Duke Math. J.* **66**:1 (1992), 123–168. MR Zbl

[29] I. Frenkel, J. Lepowsky, and A. Meurman, *Vertex operator algebras and the Monster*, Pure Appl. Math. **134**, Academic Press, Boston, MA, 1988. MR Zbl

[30] I. B. Frenkel, Y.-Z. Huang, and J. Lepowsky, *On axiomatic approaches to vertex operator algebras and modules*, Mem. Amer. Math. Soc. **494**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

[31] The GAP Group, "GAP: groups, algorithms, and programming", 2019, available at http://www.gap-system.org. Version 4.10.2.

[32] G. Höhn, "On the genus of the moonshine module", preprint, 2017. arXiv 1708.05990

[33] G. Höhn and S. Möller, "Systematic orbifold constructions of Schellekens' vertex operator algebras from Niemeier lattices", *J. Lond. Math. Soc.* (2) **106**:4 (2022), 3162–3207. MR Zbl

[34] G. Höhn and S. Möller, "Classification of self-dual vertex operator superalgebras of central charge at most 24", preprint, 2023. arXiv 2303.17190

[35] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge Univ. Press, 1990. MR Zbl

[36] C. H. Lam, "Cyclic orbifolds of lattice vertex operator algebras having group-like fusions", *Lett. Math. Phys.* **110**:5 (2019), 1081–1112. MR Zbl

[37] C. H. Lam and X. Lin, "A holomorphic vertex operator algebra of central charge 24 with the weight one Lie algebra $F_{4,6}A_{2,2}$", *J. Pure Appl. Algebra* **224**:3 (2020), 1241–1279. MR Zbl

[38] C. H. Lam and M. Miyamoto, "A lattice theoretical interpretation of generalized deep holes of the Leech lattice vertex operator algebra", *Forum Math. Sigma* **11** (2023), art. id. e86. MR Zbl

[39] C. H. Lam and H. Shimakura, "Reverse orbifold construction and uniqueness of holomorphic vertex operator algebras", *Trans. Amer. Math. Soc.* **372**:10 (2019), 7001–7024. MR Zbl

[40] C. H. Lam and H. Shimakura, "Inertia groups and uniqueness of holomorphic vertex operator algebras", *Transform. Groups* **25**:4 (2020), 1223–1268. MR Zbl

[41] C. H. Lam and H. Shimakura, "On orbifold constructions associated with the Leech lattice vertex operator algebra", *Math. Proc. Cambridge Philos. Soc.* **168**:2 (2020), 261–285. MR Zbl

[42] J. Lepowsky, "Calculus of twisted vertex operators", *Proc. Nat. Acad. Sci. U.S.A.* **82**:24 (1985), 8295–8299. MR Zbl

[43] H. S. Li, "Symmetric invariant bilinear forms on vertex operator algebras", *J. Pure Appl. Algebra* **96**:3 (1994), 279–297. MR Zbl

[44] H.-S. Li, "Local systems of twisted vertex operators, vertex operator superalgebras and twisted modules", pp. 203–236 in *Moonshine*, *the Monster*, *and related topics* (South Hadley, MA, 1994), edited by C. Dong and G. Mason, Contemp. Math. **193**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[45] R. McRae, "On rationality for C_2-cofinite vertex operator algebras", preprint, 2021. arXiv 2108.01898

[46] M. Miyamoto, "C_2-cofiniteness of cyclic-orbifold models", *Comm. Math. Phys.* **335**:3 (2015), 1279–1286. MR Zbl

[47] M. Miyamoto and K. Tanabe, "Uniform product of $A_{g,n}(V)$ for an orbifold model V and G-twisted Zhu algebra", *J. Algebra* **274**:1 (2004), 80–96. MR Zbl

[48] S. Möller, *A cyclic orbifold theory for holomorphic vertex operator algebras and applications*, Ph.D. thesis, Technische Universität Darmstadt, 2016. arXiv 1611.09843

[49] S. Möller, "Orbifold vertex operator algebras and the positivity condition", pp. 163–171 in *Research on algebraic combinatorics and representation theory of finite groups and vertex operator algebras*, edited by T. Abe, RIMS Kôkyûroku **2086**, Res. Inst. Math. Sci., Kyoto, 2018. Zbl

[50] S. Möller and N. R. Scheithauer, "Dimension formulae and generalised deep holes of the Leech lattice vertex operator algebra", *Ann. of Math.* (2) **197**:1 (2023), 221–288. MR Zbl

[51] P. S. Montague, "Orbifold constructions and the classification of self-dual $c = 24$ conformal field theories", *Nuclear Phys. B* **428**:1-2 (1994), 233–258. MR Zbl

[52] Y. Moriwaki, "Genus of vertex algebras and mass formula", *Math. Z.* **299**:3-4 (2021), 1473–1505. MR Zbl

[53] Y. Moriwaki, "Two-dimensional conformal field theory, full vertex algebra and current-current deformation", *Adv. Math.* **427** (2023), art. id. 109125. MR Zbl

[54] H.-V. Niemeier, "Definite quadratische Formen der Dimension 24 und Diskriminante 1", *J. Number Theory* **5** (1973), 142–178. MR Zbl

[55] A. N. Schellekens, "Meromorphic $c = 24$ conformal field theories", *Comm. Math. Phys.* **153**:1 (1993), 159–185. MR Zbl

[56] J.-P. Serre, *Topics in Galois theory*, Res. Notes in Math. **1**, Jones & Bartlett, Boston, MA, 1992. MR Zbl

[57] B. B. Venkov, "On the classification of integral even unimodular 24-dimensional quadratic forms", *Trudy Mat. Inst. Steklov.* **148** (1978), 65–76. In Russian; translated in *Proc. Steklov Inst. Math.* **148** (1980), 63–74. MR Zbl

Communicated by Edward Frenkel
Received 2022-11-10 Revised 2023-07-20 Accepted 2023-10-12

math@moeller-sven.de *Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany*

scheithauer@mathematik.tu-darmstadt.de *Fachbereich Mathematik, Technische Universität Darmstadt, Darmstadt, Germany*

ALGEBRA AND NUMBER THEORY 18:10 (2024)
https://doi.org/10.2140/ant.2024.18.1923

A short resolution of the diagonal for smooth projective toric varieties of Picard rank 2

Michael K. Brown and Mahrud Sayrafi

Given a smooth projective toric variety X of Picard rank 2, we resolve the diagonal sheaf on $X \times X$ by a linear complex of length $\dim X$ consisting of finite direct sums of line bundles. As applications, we prove a new case of a conjecture of Berkesch, Erman and Smith that predicts a version of Hilbert's syzygy theorem for virtual resolutions, and we obtain a Horrocks-type splitting criterion for vector bundles over smooth projective toric varieties of Picard rank 2, extending a result of Eisenbud, Erman and Schreyer. We also apply our results to give a new proof, in the case of smooth projective toric varieties of Picard rank 2, of a conjecture of Orlov concerning the Rouquier dimension of derived categories.

1. Introduction

Beilinson's resolution [1978] of the diagonal over a projective space is a powerful tool in algebraic geometry. For instance, this resolution may be used to show that the bounded derived category $\mathrm{D^b}(\mathbb{P}^n)$ is generated by the line bundles $\mathcal{O}, \mathcal{O}(1), \dots, \mathcal{O}(n)$. Additionally, taking a Fourier–Mukai transform with kernel given by Beilinson's resolution yields a representation of any object in $\mathrm{D^b}(\mathbb{P}^n)$ as a complex of vector bundles, called a *Beilinson monad*, which has been used to great effect in computational algebraic geometry, e.g., [Eisenbud and Schreyer 2003; 2009].

We aim to construct a Beilinson-type resolution of the diagonal over a smooth projective toric variety X of Picard rank 2. More specifically, with a view toward proving a new case of a conjecture of Berkesch, Erman and Smith (Conjecture 1.2 below), we construct such a resolution of length $\dim X$ — the shortest possible length — whose terms are finite direct sums of line bundles. While the existence of a full strong exceptional collection of line bundles [Costa and Miró-Roig 2004; Borisov and Hua 2009] implies that X admits a resolution of the diagonal via a tilting bundle construction [King 1997, Proposition 3.1], it follows from a result of Ballard and Favero [2012, Proposition 3.33] that this resolution may have length greater than $\dim X$. Our main result is as follows:

Theorem 1.1. *Let X be the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_s))$ over $\mathbb{P}^r$, where $1 \le r, s$ and $0 \le a_1 \le \cdots \le a_s$. Denote by $\mathbb{F}_{a_s}$ the Hirzebruch surface of type a_s, and equip $\mathrm{Pic}(\mathbb{F}_{a_s}) \cong \mathbb{Z}^2$ with the basis described in Convention 3.1 below. There is a complex R of finitely generated graded free modules over the Cox ring of $X \times X$ such that*:

MSC2020: primary 13D02; secondary 14F06, 14F08.
Keywords: Horrocks splitting criterion, resolution of the diagonal, toric variety, virtual resolution.

(1) *R is exact in positive degrees.*

(2) *R is linear, in the sense that there exists a basis of R with respect to which the differentials of R are matrices whose entries are $\Bbbk$-linear combinations of the variables.*

(3) *We have* $\operatorname{rank} R_n = \binom{r+s}{n} \dim_{\Bbbk} H^0(\mathbb{F}_{a_s}, \mathcal{O}(r, s))$. *In particular, R has length* $\dim X = r + s$, *and the equality* $\operatorname{rank} R_n = \operatorname{rank} R_{r+s-n}$ *holds.*

(4) *The sheafification $\mathcal{R}$ of R is a resolution of the diagonal sheaf $\mathcal{O}_\Delta$ on $X \times X$.*

We note that, by a result of Kleinschmidt [1988], every smooth projective toric variety of Picard rank 2 arises as a projective bundle as in the hypothesis of Theorem 1.1. We construct the resolution $\mathcal{R}$ in Theorem 1.1 using a variant of Weyman's "geometric technique" [2003, Section 5] for building free resolutions. In a bit more detail: let x_i and x_i' refer to the variables corresponding to the first and second copy of X, respectively, in the Cox ring S of $X \times X$. A first, naive, idea is that the diagonal sheaf $\mathcal{O}_\Delta$ ought to be defined by the relations $x_i - x_i'$ in S. The problem is that these relations are not homogeneous with respect to the $\mathbb{Z}^4$-grading on S. To fix this, we homogenize the relations $x_i - x_i'$ in the Cox ring of a certain toric fiber bundle E over $X \times X$ with fiber given by $\mathbb{F}_{a_s}$. Our resolution $\mathcal{R}$ is obtained by taking the Koszul complex on these homogenized relations over E, twisting it by a certain line bundle, and pushing it forward to $X \times X$. Choosing the toric fiber bundle E is delicate; not only do the degrees of the variables in the Cox ring of E need to be suitable for homogenizing the relations $x_i - x_i'$, but the terms of the Koszul complex on these homogenized relations must enjoy appropriate cohomological vanishing properties in order to conclude that $\mathcal{R}$ is a resolution of the required form. See Section 3C for details.

The simplest case of Theorem 1.1 is the Hirzebruch surface

$$\mathbb{F}_a = \mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a)),$$

where $r = s = 1$ and $a = a_1$. As detailed in Example 3.9, the construction above yields a resolution of the diagonal for $\mathbb{F}_a$ whose terms $\mathcal{R}_0$, $\mathcal{R}_1$, and $\mathcal{R}_2$ are sums of $a+4$, $2a+8$, and $a+4$ line bundles, respectively; cf. [Buchdahl 1987, Section 1].

As we explain in Section 2, the resolution $\mathcal{R}$ in Theorem 1.1 should be considered as a natural extension of Beilinson's resolution over projective space and similar resolutions due to Buchdahl [1987] for Hirzebruch surfaces, Canonaco and Karp [2008] for weighted projective stacks, and Kapranov [1988] for quadrics and flag varieties. See [Brown and Erman 2021, Section 4] for a related idea, where a resolution of the diagonal — with terms given by infinite direct sums of line bundles — is obtained for any projective toric stack.

We apply Theorem 1.1 to make progress on a conjecture concerning virtual resolutions in commutative algebra, a notion introduced by Berkesch, Erman and Smith [Berkesch et al. 2020]. We recall that a *virtual resolution* of a graded module M over the Cox ring S of a toric variety Y is a complex F of graded free S-modules such that the associated complex of sheaves $\widetilde{F}$ on Y is a locally free resolution of $\widetilde{M}$. The following conjecture predicts a version of Hilbert's syzygy theorem for virtual resolutions:

Conjecture 1.2 [Berkesch et al. 2020, Question 6.5]. *If Y is a smooth toric variety with Cox ring S and irrelevant ideal B, and M is a finitely generated, B-saturated S-module, then M admits a virtual resolution of length at most* $\dim(Y)$.

This conjecture was proven by Berkesch and Erman and Smith [2020] for products of projective spaces (see also [Eisenbud et al. 2015, Corollary 2.14]) and for monomial ideals in Cox rings of smooth toric varieties by Yang [2021]. As a consequence of Theorem 1.1, we prove the following:

Corollary 1.3. *Conjecture 1.2 holds for smooth projective toric varieties of Picard rank* 2.

Theorem 1.1 also yields a new proof, in the case of smooth projective toric varieties of Picard rank 2, of the following conjecture of Orlov:

Conjecture 1.4 [Orlov 2009, Conjecture 10]. *Let Y be a smooth quasiprojective scheme. The Rouquier dimension of the bounded derived category* $\mathrm{D}^{\mathrm{b}}(Y)$ *is equal to* $\dim(Y)$.

We refer the reader to the original paper of Rouquier [2008] for background on his notion of dimension for triangulated categories. Since the resolution of the diagonal $\mathcal{R}$ in Theorem 1.1 has length $\dim X$, and each term $\mathcal{R}_i$ is a sum of box products of vector bundles on X, it is an immediate consequence of [loc. cit., Proposition 7.6] that Theorem 1.1 implies Conjecture 1.4 for smooth projective toric varieties of Picard rank 2. Conjecture 1.4 was first proven in this case by Ballard, Favero and Katzarkov [Ballard et al. 2019, Corollary 5.2.6] using an entirely different approach: they first observe that the conjecture holds for a smooth projective Picard rank 2 toric variety that is weakly Fano, and then they apply descent along admissible subcategories. See the discussion beneath [Bai and Côté 2023, Conjecture 1.1] for a list of known cases of Conjecture 1.4.

We also apply Theorem 1.1 to obtain a splitting criterion for vector bundles on smooth projective toric varieties of Picard rank 2. A famous result of Horrocks [1964] states that if a vector bundle on projective space has no intermediate cohomology, then it splits as a sum of line bundles. This splitting criterion has been generalized in many different directions: for instance, to products of projective spaces [Costa and Miró-Roig 2005; Eisenbud et al. 2015; Schreyer 2022], to Grassmannians and quadrics [Ottaviani 1989], and to rank 2 vector bundles on Hirzebruch surfaces [Fulger and Marchitan 2011; Yasutake 2015], among others. Our splitting criterion for smooth projective toric varieties of Picard rank 2 extends Eisenbud, Erman and Schreyer's for products of projective spaces [Eisenbud et al. 2015, Theorem 7.2].

To state the result, we must fix some notation. Given $(a, b), (c, d) \in \mathbb{Z}^2$, we write $(a, b) \le (c, d)$ if $a \le c$ and $b \le d$. For a sheaf $\mathcal{F}$ on X, let $\gamma(\mathcal{F})$ denote its cohomology table

$$\gamma(\mathcal{F}) = (\dim_{\Bbbk} H^i(X, \mathcal{F}(a, b)))_{i \ge 0, (a,b) \in \mathbb{Z}^2}.$$

Here, as in Theorem 1.1, we identify $\operatorname{Pic} X$ with $\mathbb{Z}^2$ via the choice of basis described in Convention 3.1 below. Our splitting criterion is as follows:

Theorem 1.5. *Let $\mathcal{E}$ be a vector bundle on X. Suppose we have*

$$\gamma(\mathcal{E}) = \sum_{i=1}^{t} \gamma(\mathcal{O}(b_i, c_i)^{m_i}).$$

If $(b_t, c_t) \le (b_{t-1}, c_{t-1}) \le \cdots \le (b_1, c_1)$, then $\mathcal{E} \cong \bigoplus_{i=1}^{t} \mathcal{O}(b_i, c_i)^{m_i}$.

Our proof of Theorem 1.5 uses a Beilinson-type spectral sequence built from the resolution of the diagonal in Theorem 1.1. This approach is similar to the technique used by Fulger and Marchitan [2011] to obtain a splitting criterion for rank 2 vector bundles on Hirzebruch surfaces, which involves a Beilinson-type spectral sequence built from Buchdahl's resolution [1987] of the diagonal for Hirzebruch surfaces. See also Aprodu and Marchitan's triviality criterion [2011, Theorem 2] for vector bundles on Hirzebruch surfaces, whose proof also involves a Beilinson-type spectral sequence.

When $X = \mathbb{P}^r \times \mathbb{P}^s$, Theorem 1.5 recovers (a special case of) [Eisenbud et al. 2015, Theorem 7.2]. We note that the nef cone of X is given by $\operatorname{Nef} X = \{\mathcal{O}(a, b) \in \operatorname{Pic} X : a, b \ge 0\}$, and so $(a, b) \le (c, d)$ if and only if the line bundle $\mathcal{O}(c - a, d - b)$ is nef. Theorem 1.5 therefore adds a new wrinkle that is not present on products of projective spaces: we require the twists (b_i, c_i) to be ordered with respect to the nef cone, rather than the effective cone. This distinction is invisible in [loc. cit., Theorem 7.2], as the nef and effective cones of a product of projective spaces coincide.

Motivated by the applications of Theorem 1.1 described above, we pose the following:

Question 1.6. Can Theorem 1.1 be generalized to any smooth projective toric variety X?

The difficulty in generalizing Theorem 1.1 is in choosing an appropriate toric fiber bundle E over $X \times X$. A positive answer to Question 1.6 would immediately resolve the projective case of Conjecture 1.2 and imply a large swath of new cases of Conjecture 1.4.

Overview. We begin in Section 2 by constructing a resolution of the diagonal over $\mathbb{P}^n$ as the pushforward of a Koszul complex over a certain projective bundle, which illustrates our main approach. We prove Theorem 1.1 and Corollary 1.3 in Section 3, and we prove Theorem 1.5 in Section 4.

2. Warm-up: the case of $\mathbb{P}^n$

Throughout the paper, we work over a base field $\Bbbk$. Let $\mathcal{T}_{\mathbb{P}^n}$ denote the tangent bundle on $\mathbb{P}^n$ and $\mathcal{W}$ the vector bundle $\mathcal{O}_{\mathbb{P}^n}(1) \boxtimes \mathcal{T}_{\mathbb{P}^n}(-1)$ on $\mathbb{P}^n \times \mathbb{P}^n$. There is a canonical section $s \in H^0(\mathbb{P}^n \times \mathbb{P}^n, \mathcal{W})$ whose vanishing cuts out the diagonal in $\mathbb{P}^n \times \mathbb{P}^n$; see [Huybrechts 2006, Section 8.3]. The Koszul complex associated to s yields Beilinson's resolution of the diagonal

$$0 \leftarrow \mathcal{O}_\Delta \leftarrow \mathcal{O}_{\mathbb{P}^n \times \mathbb{P}^n} \leftarrow \Lambda^1 \mathcal{W}^\vee \leftarrow \cdots \leftarrow \Lambda^n \mathcal{W}^\vee \leftarrow 0.$$

In this section, we construct another resolution of the diagonal sheaf on $\mathbb{P}^n \times \mathbb{P}^n$, whose terms are direct sums of line bundles; cf. [Canonaco and Karp 2008, Remark 3.3]. We explain in Remark 2.3(3) a sense in which this resolution resembles Beilinson's. As discussed in the introduction, our approach is similar to Weyman's "geometric technique" [2003, Section 5]. In Section 3, we explain how the approach in this section extends to smooth projective toric varieties of Picard rank 2.

Let E denote the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(-1, 1))$ on $\mathbb{P}^n \times \mathbb{P}^n$ and let $\pi : E \to \mathbb{P}^n \times \mathbb{P}^n$ be the canonical map. The projective bundle E is a toric variety with $\mathbb{Z}^3$-graded Cox ring

$$S_E = \Bbbk[x_0, \dots, x_n, y_0, \dots, y_n, u_0, u_1],$$

where $\deg(x_i) = (1, 0, 0)$, $\deg(y_i) = (0, 1, 0)$, $\deg(u_0) = (1, -1, 1)$, and $\deg(u_1) = (0, 0, 1)$. Set $\alpha_i = u_1 x_i - u_0 y_i$ for all i; the intuition here is that u_0 and u_1 are homogenizing variables for the nonhomogeneous equations $x_i - y_i$. Let $\mathcal{K}$ denote the Koszul complex on $\alpha_0, \dots, \alpha_n$, considered as a complex of sheaves on E, and set $\mathcal{V} = \mathcal{O}(-1, 0, 0)^{n+1}$. Twisting $\mathcal{K}$ by $\mathcal{O}(0, 0, n)$ yields a complex of the form

$$\mathcal{O}(0, 0, n) \leftarrow (\Lambda^1 \mathcal{V})(0, 0, n-1) \leftarrow \cdots \leftarrow \Lambda^n \mathcal{V} \leftarrow (\Lambda^{n+1} \mathcal{V})(0, 0, -1).$$

Using [Hartshorne 1977, Chapter III, Exercise 8.4(a)] and the projection formula, $\mathcal{R} = \pi_* \mathcal{K}(0, 0, n)$ has the form

$$\operatorname{Sym}^n \mathcal{Q} \leftarrow \Lambda^1 \mathcal{P} \otimes \operatorname{Sym}^{n-1} \mathcal{Q} \leftarrow \cdots \leftarrow \Lambda^{n-1} \mathcal{P} \otimes \operatorname{Sym}^1 \mathcal{Q} \leftarrow \Lambda^n \mathcal{P}, \tag{2-1}$$

where $\mathcal{P} = \mathcal{O}(-1, 0)^{n+1}$ and $\mathcal{Q} = \mathcal{O} \oplus \mathcal{O}(-1, 1)$. Notice that applying π_* to the $n+1$-th term $(\Lambda^{n+1} \mathcal{V})(0, 0, -1)$ of $\mathcal{K}(0, 0, n)$ gives 0, hence the complex (2-1) has length n.

Proposition 2.1. *The complex $\mathcal{R}$ is a resolution of the diagonal sheaf on $\mathbb{P}^n \times \mathbb{P}^n$. Moreover, the complex $\mathcal{R}$ is isomorphic to (the sheafification of) the n-th symmetric power of the complex*

$$S(-1, 1) \oplus S \xleftarrow{\begin{pmatrix} -y_0 & -y_1 & \cdots & -y_n \\ x_0 & x_1 & \cdots & x_n \end{pmatrix}} S(-1, 0)^{n+1}, \tag{2-2}$$

concentrated in homological degrees 0 *and* 1, *where S denotes the Cox ring of $\mathbb{P}^n \times \mathbb{P}^n$.*

Proof. One can use a slight variation of the proof of Theorem 1.1 below to show that $\mathcal{R}$ is a resolution of the diagonal. As for the second statement: let K denote the Koszul complex on the regular sequence $\alpha_0, \dots, \alpha_n$, considered as a complex of S_E-modules. Let R be the complex of S-modules given by $K(0, 0, n)_{(*,*,0)}$. Since K is exact in positive homological degrees, R is as well. It follows from the description of $\mathcal{R}$ in (2-1) that R sheafifies to $\mathcal{R}$. Let R' denote the n-th symmetric power of (2-2). We observe that R' has exactly the same terms as R. The complex R' is precisely the generalized Eagon–Northcott complex of type $\mathcal{C}^n$, as defined in [Eisenbud 1995, A2.6], associated to the map (2-2). It therefore follows from [loc. cit., Theorem A2.10(c)] that R' is exact in positive homological degrees. By the uniqueness of minimal free resolutions, we need only check that the cokernels of the first differentials of R and R' are isomorphic, and this can be verified by direct computation. $\square$

We now compute a well-known example using this approach; cf. [King 1997, Example 5.2].

Example 2.2. Suppose $n = 2$. The monomials in the u_i give bases for the symmetric powers of $\mathcal{Q}$, and the exterior monomials in the α_i give bases for the terms of $\mathcal{K}$, which correspond to the exterior powers of $\mathcal{P}$. Hence, we may index the summands of (2-1) by monomials in $u_0, u_1, \alpha_0, \alpha_1, \alpha_2$. With this in mind, the complex (2-1) has terms

$$\underbrace{\mathcal{O}(-2,2)}_{u_0^2} \oplus \underbrace{\mathcal{O}(-1,1)}_{u_0u_1} \oplus \underbrace{\mathcal{O}}_{u_1^2} \xleftarrow{\partial_1} \underbrace{\mathcal{O}(-2,1)^3}_{\alpha_0u_0,\alpha_1u_0,\alpha_2u_0} \oplus \underbrace{\mathcal{O}(-1,0)^3}_{\alpha_0u_1,\alpha_1u_1,\alpha_2u_1} \xleftarrow{\partial_2} \underbrace{\mathcal{O}(-2,0)^3}_{\alpha_0\alpha_1,\alpha_0\alpha_2,\alpha_1\alpha_2}$$

and differentials

$$\partial_1 = \begin{pmatrix} -y_0 & -y_1 & -y_2 & 0 & 0 & 0 \\ x_0 & x_1 & x_2 & -y_0 & -y_1 & -y_2 \\ 0 & 0 & 0 & x_0 & x_1 & x_2 \end{pmatrix} \quad \text{and} \quad \partial_2 = \begin{pmatrix} y_1 & y_2 & 0 \\ -y_0 & 0 & y_2 \\ 0 & -y_0 & -y_1 \\ -x_1 & -x_2 & 0 \\ x_0 & 0 & -x_2 \\ 0 & x_0 & x_1 \end{pmatrix}.$$

Remark 2.3. We conclude this section with the following observations:

(1) We have $\operatorname{rank} \mathcal{R}_i = \operatorname{rank} \mathcal{R}_{n-i}$, just as in Theorem 1.1.

(2) The resolutions in Theorem 1.1 cannot arise as symmetric powers of complexes, in general; this follows immediately from rank considerations.

(3) Let us explain a sense in which our resolution $\mathcal{R}$ is modeled on Beilinson's resolution of the diagonal. Consider the external tensor product of $\mathcal{O}(1)$ with the Euler sequence

$$0 \leftarrow \mathcal{O}(1) \boxtimes \mathcal{T}(-1) \leftarrow \mathcal{O}(1,0)^{n+1} \xleftarrow{(y_0 \cdots y_n)^T} \mathcal{O}(1,-1) \leftarrow 0.$$

Letting $\mathcal{C}$ denote the subcomplex $\mathcal{O}(1,0)^{n+1} \leftarrow \mathcal{O}(1,-1)$ concentrated in degrees 0 and 1, there is a quasiisomorphism $\mathcal{C} \xrightarrow{\simeq} \mathcal{O}(1) \boxtimes \mathcal{T}(-1)$. The morphism $s \colon \mathcal{O} \xrightarrow{(x_0 \cdots x_n)^T} \mathcal{C}$, where $\mathcal{O}$ lies in degree 0, gives a hypercohomology class in $\mathbb{H}^0(\mathbb{P}^n \times \mathbb{P}^n, \mathcal{C})$, which is isomorphic to $H^0(\mathbb{P}^n \times \mathbb{P}^n, \mathcal{O}(1) \boxtimes \mathcal{T}(-1))$. By Proposition 2.1, the n-th symmetric power of the dual of s, i.e., the n-th Koszul complex of the dual of s [Köck 2001, Definition 2.3], is isomorphic to the resolution $\mathcal{R}$. In short: the resolution $\mathcal{R}$ is a Koszul complex on a section of $\mathcal{O}(1) \boxtimes \mathcal{T}(-1)$, just like Beilinson's resolution.

3. Smooth projective toric varieties of Picard rank 2

In this section, we extend the construction in Section 2 and prove the main theorem. Let X denote the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_s))$ over $\mathbb{P}^r$, where $a_1 \leq \cdots \leq a_s$. As discussed in [Cox et al. 2011, Section 7.3], the fan $\Sigma_X \subseteq \mathbb{Z}^{r+s}$ of X has $r+s+2$ ray generators given by the rows of the

$(r+s+2)\times(r+s)$ matrix

$$P=\begin{pmatrix} -1 & -1 & \cdots & -1 & a_1 & a_2 & \cdots & a_s \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & -1 & \cdots & -1 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}=\begin{pmatrix}\rho_0\\ \rho_1\\ \rho_2\\ \vdots\\ \rho_r\\ \sigma_0\\ \sigma_1\\ \sigma_2\\ \vdots\\ \sigma_s\end{pmatrix} \tag{3-1}$$

and maximal cones generated by collections of rays of the form

$$\{\rho_0,\dots,\widehat{\rho_i},\dots,\rho_r,\sigma_0,\dots,\widehat{\sigma_j},\dots,\sigma_s\}.$$

Convention 3.1. Throughout the paper, we equip $\operatorname{Pic} X\cong\operatorname{coker}(P)\cong\mathbb{Z}^2$ with the basis given by the divisors corresponding to ρ_0 and σ_0. With this choice of basis, we may view the Cox ring of X as the $\mathbb{Z}^2$-graded ring $\Bbbk[x_0,\dots,x_r,y_0,\dots,y_s]$ whose variables have degrees given by the columns of the matrix

$$A=\begin{pmatrix} 1 & 1 & \cdots & 1 & 0 & -a_1 & \cdots & -a_s \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \end{pmatrix}.$$

A main reason we use this convention is that it is also used by the function `kleinschmidt` in `Macaulay2`, which produces any smooth projective toric variety of Picard rank 2 as an object of type `NormalToricVariety`.

3A. ***Vanishing of sheaf cohomology.*** We will need a calculation of the cohomology of a line bundle on X:

Proposition 3.2. *Let $\mathcal{E}$ be the vector bundle $\mathcal{O}\oplus\mathcal{O}(a_1)\oplus\cdots\oplus\mathcal{O}(a_s)$ on $\mathbb{P}^r$, where $a_1\le\cdots\le a_s$, so that $X=\mathbb{P}(\mathcal{E})$. Write $m=\sum_{i=1}^s a_i$, and consider a line bundle $\mathcal{O}(k,\ell)$ on X. For each $0\le j\le r+s$, we have*:

$$H^j(X,\mathcal{O}(k,\ell))\cong\begin{cases} H^j(\mathbb{P}^r,\mathcal{O}_{\mathbb{P}^r}(k)\otimes\operatorname{Sym}^\ell(\mathcal{E})), & \ell\ge 0;\\ H^{j-s}(\mathbb{P}^r,\mathcal{O}_{\mathbb{P}^r}(k-m)\otimes\operatorname{Sym}^{-\ell-s-1}(\mathcal{E})^\vee), & \ell\le -s-1;\\ 0, & \textit{otherwise.}\end{cases}$$

Proof. Let $\pi: X\to\mathbb{P}^r$ denote the projective bundle map. It follows from a well-known calculation (see e.g., [Thomason and Trobaugh 1990, 4.5(e)]) and the projection formula that

$$\boldsymbol{R}^i\pi_*(\mathcal{O}(k,\ell))=\begin{cases}\mathcal{O}_{\mathbb{P}^r}(k)\otimes\operatorname{Sym}^\ell(\mathcal{E}), & i=0;\\ \mathcal{O}_{\mathbb{P}^r}(k-m)\otimes\operatorname{Sym}^{-\ell-s-1}(\mathcal{E})^\vee, & i=s;\\ 0, & 0<i<s.\end{cases}$$

The conclusion follows from the observation that the second page of the Grothendieck spectral sequence

$$E_2^{p,q} = H^p(\mathbb{P}^r, \boldsymbol{R}^q\pi_*(\mathcal{O}(k,\ell))) \Rightarrow H^{p+q}(X, \mathcal{O}(k,\ell))$$

collapses to row $q=0$ when $\ell \geq 0$ and to row $q=s$ when $\ell \leq -s-1$. □

The following result is an immediate consequence of Proposition 3.2. It will play a key role in the proofs of Theorems 1.1 and 1.5.

Corollary 3.3. *Let X be the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_s))$ over $\mathbb{P}^r$ as above, where $a_1 \leq \cdots \leq a_s$. Write $m = \sum_{i=1}^s a_i$, and consider a line bundle $\mathcal{O}(k,\ell)$ on X:*

(1) *We have*:
  - (a) $H^i(X, \mathcal{O}(k,\ell)) = 0$ *if* $i \notin \{0, r, s, r+s\}$.
  - (b) $H^0(X, \mathcal{O}(k,\ell)) = 0$ *if and only if* $\ell < 0$ *or* $k + a_s\ell < 0$.
  - (c) *If* $r \neq s$ *then*
    - (i) $H^r(X, \mathcal{O}(k,\ell)) = 0$ *if and only if* $-r-1 < k$ *or* $\ell < 0$, *and*
    - (ii) $H^s(X, \mathcal{O}(k,\ell)) = 0$ *if and only if* $-s-1 < \ell$ *or* $k < m$.
  - (d) *If* $r = s$ *then* $H^r(X, \mathcal{O}(k,\ell)) = 0$ *if and only if both of the following hold*:
    - (i) $-r-1 < k$ *or* $\ell < 0$.
    - (ii) $-s-1 < \ell$ *or* $k < m$.
  - (e) *Lastly*, $H^{r+s}(X, \mathcal{O}(k,\ell)) = 0$ *if and only if either of the following hold*:
    - (i) $-r-1-a_s(\ell+s+1)+m < k$.
    - (ii) $-s-1 < \ell$;

(2) *In particular, the line bundle $\mathcal{O}(k,\ell)$ is acyclic ($H^i(X, \mathcal{O}(k,\ell)) = 0$ for $i > 0$) if and only if one of the following holds*:
  - (a) $-s-1 < \ell < 0$.
  - (b) $-r-1 < k$ *and* $0 \leq \ell$.
  - (c) $-r-1-a_s(\ell+s+1)+m < k < m$ *and* $\ell \leq -s-1$.

Remark 3.4. Conditions (1b) and (1e) are Serre dual to one another. Ditto for the two conditions in (1c), as well as the conditions (i) and (ii) in (1d). These calculations are surely well-known; see, for instance, [Lasoń and Michałek 2011, Proposition 3.9] for a criterion for acyclicity of line bundles on toric varieties. We refer the reader to [Brown and Erman 2021, Example 3.14] for a depiction of the regions of $\mathbb{Z}^2$ where each $H^i(X, \mathcal{O}(k,\ell))$ vanishes for the Hirzebruch surface $X = \mathbb{P}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(3))$.

3B. *Toric fiber bundles.* Let E and Y be smooth projective toric varieties of dimensions d_E and d_Y associated to fans Σ_E and Σ_Y. Let $\bar{\pi} : \mathbb{Z}^{d_E} \to \mathbb{Z}^{d_Y}$ be a $\mathbb{Z}$-linear surjection that is compatible with the fans Σ_E and Σ_Y, in the sense of [Cox et al. 2011, Definition 3.3.1], so that it induces a morphism $\pi : E \to Y$. We denote by F the toric variety associated to the fan $\Sigma_F = \{\sigma \in \Sigma_E : \sigma \subseteq \ker(\bar{\pi})_{\mathbb{R}}\}$, and write $d_F = \dim F$. Let us assume that the fan Σ_E is *split* by the fans Σ_Y and Σ_F, in the sense of [loc. cit., Definition 3.3.18]. In this case, the map $\pi: E \to Y$ is a fibration with fiber F; see [loc. cit., Theorem 3.3.19].

Writing the Cox rings of Y and F as $S_Y = \Bbbk[x_1, \dots, x_{n_1}]$ and $S_F = \Bbbk[u_1, \dots, u_{n_2}]$, the Cox ring of E has the form $S_E = \Bbbk[x_1, \dots, x_{n_1}, u_1, \dots, u_{n_2}]$. We have presentations $P_Y : \mathbb{Z}^{d_Y} \to \mathbb{Z}^{n_1}$ and $P_F : \mathbb{Z}^{d_F} \to \mathbb{Z}^{n_2}$ of Pic Y and Pic F whose rows are given by the ray generators of Σ_Y and Σ_F, respectively. The analogous presentation of Pic E is of the form

$$\begin{pmatrix} P_Y & Q \\ 0 & P_F \end{pmatrix}$$

for some $n_1 \times d_F$ matrix Q. One may use this presentation to equip S_E with a $\mathbb{Z}^e \oplus \mathbb{Z}^f$-grading such that $\deg_{S_E}(x_i) = (\deg_{S_Y}(x_i), 0)$, and $\deg_{S_E}(u_i) = (t_i, \deg_{S_F}(u_i))$ for some $t_i \in \mathbb{Z}^e$.

Lemma 3.5 (cf. [Hartshorne 1977, Chapter III, Exercise 8.4(a)]). *Let $\mathcal{L} = \mathcal{O}_E(b_1, \dots, b_e, c_1, \dots, c_f)$, and let $\mathcal{B}$ be a $\Bbbk$-basis of $H^0(F, \mathcal{O}_F(c_1, \dots, c_f))$ given by monomials in S_F. Given $m \in \mathcal{B}$, denote its degree in S_E by $(d_1^m, \dots, d_e^m, c_1, \dots, c_f)$. We have $\pi_*(\mathcal{L}) \cong \bigoplus_{m \in \mathcal{B}} \mathcal{O}_Y(b_1 - d_1^m, \dots, b_e - d_e^m)$. Moreover, if $H^i(F, \mathcal{O}_F(c_1, \dots, c_f)) = 0$, then $\boldsymbol{R}^i\pi_*(\mathcal{L}) = 0$.*

Proof. Let $g\colon \bigoplus_{m \in \mathcal{B}} \mathcal{O}_Y(b_1 - d_1^m, \dots, b_e - d_e^m) \to \pi_*(\mathcal{L})$ be the morphism given on the component corresponding to $m \in \mathcal{B}$ by multiplication by m. Let U be an affine open subset of Y over which the fiber bundle E is trivializable; abusing notation slightly, we denote by π the map $\pi^{-1}(U) \to U$ induced by π. To prove the first statement, it suffices to show that the restriction $g_U : \bigoplus_{m \in \mathcal{B}_i} \mathcal{O}_U \to \pi_*(\mathcal{L}|_U)$ of g to U is an isomorphism. Without loss of generality, we may assume that $\pi^{-1}(U) = U \times F$ and that $\pi : \pi^{-1}(U) \to U$ is the projection onto U. Letting $\gamma : \pi^{-1}(U) \to F$ denote the projection, we have that $\mathcal{L}|_U = \gamma^*(\mathcal{O}_F(c_1, \dots, c_f))$. Finally, we observe that g_U coincides with the base change isomorphism

$$\bigoplus_{m \in \mathcal{B}} \mathcal{O}_U = \mathcal{O}_U \otimes_{\Bbbk} H^0(F, \mathcal{O}_F(c_1, \dots, c_f)) \xrightarrow{\cong} \pi_*(\gamma^*(\mathcal{O}_F(c_1, \dots, c_f)) = \pi_*(\mathcal{L}|_U).$$

As for the last statement: it suffices to observe that, by base change,

$$\boldsymbol{R}^i\pi_*(\mathcal{L}|_U) \cong \mathcal{O}_U \otimes_{\Bbbk} H^i(F, \mathcal{O}_F(c_1, \dots, c_f)) = 0. \qquad \square$$

3C. *Constructing the resolution of the diagonal.* Let X be as defined at the beginning of this section. We will construct our resolution of the diagonal for X as the pushforward of a certain Koszul complex on a fibration E over $X \times X$ whose fiber is the Hirzebruch surface $\mathbb{F}_{a_s}$. We begin by constructing the fiber bundle $\pi : E \to X \times X$. The ray generators of E are given by the rows of the $(2r + 2s + 8) \times (2r + 2s + 2)$ matrix

$$\left(\begin{array}{cc|cc} P & 0 & v & -w \\ 0 & P & -v & w \\ \hline 0 & 0 & -1 & a_s \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right), \tag{3-2}$$

where P is as in (3-1), and v (resp. w) is the $(r+s+2)\times 1$ matrix with unique nonzero entry given by a 1 in the first (resp. $(r+2)$-th) position. Notice that the rows in the top-left quadrant of this matrix are the ray generators of $X\times X$, and the rows in the bottom-right quadrant are the ray generators of $\mathbb{F}_{a_s}$.

Let $\bar{\pi}:\mathbb{Z}^{2r+2s+2}\to\mathbb{Z}^{2r+2s}$ denote the projection onto the first $2r+2s$ coordinates. We define the cones of E to be those of the form $\gamma+\gamma'$, where γ is a cone corresponding to a cone of $\mathbb{F}_{a_s}$ and is spanned by a subset of the bottom 4 rows of (3-2), and γ' is a cone spanned by a collection of the top $2r+2s+4$ rows of (3-2) such that $\bar{\pi}_{\mathbb{R}}(\gamma')$ is a cone of $X\times X$. By [Cox et al. 2011, Theorem 3.3.19], the map $\bar{\pi}$ induces a fibration $\pi: E\to X$ with fiber $\mathbb{F}_{a_s}$.

In order to describe the Cox ring of E, first recall the matrix A from Convention 3.1 whose columns are the degrees of the variables of the Cox ring of X, and consider the matrices

$$B=\begin{pmatrix}1 & -a_s & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}\quad\text{and}\quad C=\begin{pmatrix}1 & -a_s & 1 & 0\\ 0 & 1 & 0 & 1\end{pmatrix}.$$

Notice that the columns of C are the degrees of the variables in the Cox ring of $\mathbb{F}_{a_s}$. We choose a basis of $\operatorname{Pic} E\cong\mathbb{Z}^6$ so that the degrees of the variables in the Cox ring

$$S_E=\Bbbk[x_0,\dots x_r,y_0,\dots,y_s,x'_0,\dots,x'_r,y'_0,\dots,y'_s,u_0,\dots,u_3]$$

of E are given by the columns of the Gale dual of (3-2), which is the $6\times(2r+2s+8)$ matrix

$$\begin{pmatrix}A & 0 & B\\ 0 & A & -B\\ 0 & 0 & C\end{pmatrix}=\begin{pmatrix}
1 & \cdots & 1 & 0 & -a_1 & \cdots & -a_s & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & -a_s & 0 & 0\\
0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & -a_1 & \cdots & -a_s & -1 & a_s & 0 & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & -1 & 0 & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & -a_s & 1 & 0\\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 1
\end{pmatrix}.$$

Let K be the Koszul complex corresponding to the regular sequence $\alpha_0,\dots,\alpha_r,\beta_0,\dots,\beta_s$ given by the homogeneous binomials

$$\begin{aligned}
\alpha_i &= u_2x_i-u_0x'_i && \text{for } 0\le i\le r \text{ and}\\
\beta_i &= u_3y_i-u_0^{a_s-a_i}u_1u_2^{a_i}y'_i && \text{for } 0\le i\le s\ (a_0:=0)
\end{aligned}$$

in the Cox ring S_E. Observe that $\deg(\alpha_i)=(1,0,0,0,1,0)$ and $\deg(\beta_i)=(-a_i,1,0,0,0,1)$. Here, we are using that the columns of B span the effective cone of X to homogenize the relations $x_i-x'_i$ and $y_i-y'_i$. Denote by $\mathcal{K}$ the complex of sheaves on E corresponding to K. The following proposition shows that $\mathcal{K}$ twisted by $\mathcal{O}_E(0,0,0,0,r,s)$ is π_*-acyclic.

Proposition 3.6. *The higher direct images $\boldsymbol{R}^i\pi_*(\mathcal{K}(0,0,0,0,r,s))$ vanish for $i>0$.*

Proof. It suffices to show that $\boldsymbol{R}^i\pi_*(\mathcal{K}_j(0,0,0,0,r,s))=0$ for $i>0$ and all j. Each term of $\mathcal{K}(0,0,0,0,r,s)$ is a direct sum of line bundles of the form $\mathcal{O}_E(a,b,0,0,k,\ell))$ for some $a,b\in\mathbb{Z}$, $-1\le k\le r$, and $-1\le\ell\le s$. By Lemma 3.5, we need only show that $H^i(\mathbb{F}_{a_s},\mathcal{O}(k,\ell))=0$ for $i>0$ and such k and ℓ, which follows from Corollary 3.3(2)(a-b). □

Let S denote the Cox ring of $X \times X$ and R the complex of graded S-modules given by the subcomplex $K(0,0,0,0,r,s)_{(*,*,*,*,0,0)}$ of the Koszul complex K twisted by $S_E(0,0,0,0,r,s)$. We will show that R satisfies the requirements of Theorem 1.1. Observe that, by Lemma 3.5, one can alternatively construct R by applying the twisted global sections functor:

$$R = \bigoplus_{\mathcal{L} \in \mathrm{Pic}(X \times X)} H^0(X \times X, \mathcal{L} \otimes \pi_* \mathcal{K}(0,0,0,0,r,s)).$$

In particular, writing $\mathcal{R}$ for the complex of sheaves on $X \times X$ corresponding to R, we have $\mathcal{R} \cong \pi_* \mathcal{K}(0,0,0,0,r,s)$. Note that Proposition 3.6 implies that $\pi_* \mathcal{K}(0,0,0,0,r,s)$ is quasiisomorphic to $\boldsymbol{R}\pi_*(\mathcal{K}(0,0,0,0,r,s))$.

Before discussing some examples, we must establish a bit of notation:

Notation 3.7. Let $S_F = \Bbbk[u_0, u_1, u_2, u_3]$ denote the Cox ring of the Hirzebruch surface $\mathbb{F}_{a_s}$, equipped with the $\mathbb{Z}^2$-grading so that the degrees of the variables correspond to the columns of the matrix C above. Given $i, j \in \mathbb{Z}$, let $M_{i,j}$ denote the set of monomials in S_F of degree (i, j). For $m \in M_{i,j}$, let $(d_1^m, d_2^m, d_3^m, d_4^m) \in \mathbb{Z}^4$ denote the first four coordinates of the degree of m as an element of the $\mathbb{Z}^6$-graded ring S_E; notice that $d_3^m = -d_1^m$, and $d_4^m = -d_2^m$.

Example 3.8. Let us compute the first differential in R. Using the notation above, we have

$$R_0 = \bigoplus_{m \in M_{r,s}} S(-d_1^m, -d_2^m, d_1^m, d_2^m) \cdot m \quad \text{and} \quad R_1 = R_1^\alpha \oplus R_1^\beta,$$

where

$$R_1^\alpha = \bigoplus_{i=0}^{r} \bigoplus_{m \in M_{r-1,s}} S(-d_1^m - 1, -d_2^m, d_1^m, d_2^m) \cdot \alpha_i m, \quad R_1^\beta = \bigoplus_{i=0}^{s} \bigoplus_{m \in M_{r,s-1}} S(-d_1^m + a_i, -d_2^m - 1, d_1^m, d_2^m) \cdot \beta_i m.$$

Here, the decorations "$\cdot m$" in our description of R_0 are just for bookkeeping, and similarly for the "$\cdot \alpha_i m$" and "$\cdot \beta_i m$" in R_1. Viewing the differential $\partial_1 \colon R_1 \to R_0$ as a matrix with respect to the above basis, the column corresponding to $\alpha_i m$ has exactly two nonzero entries: an entry of x_i corresponding to the monomial $u_2 m \in M_{r,s}$ and an entry of $-x_i'$ corresponding to $u_0 m \in M_{r,s}$. Similarly, the column corresponding to $\beta_i m$ has exactly two nonzero entries: an entry of y_i corresponding to $u_3 m$ and an entry of $-y_i'$ corresponding to $u_0^{a_s - a_i} u_1 u_2^{a_i} m$. That is, the matrix ∂_1 has the following form:

$$\begin{array}{cc}
\left(\begin{array}{ccccc}
 & 0 & & 0 & \\
 & -x_i' & & 0 & \\
 & 0 & & 0 & \\
 & 0 & & -y_i' & \\
\cdots & 0 & \cdots & 0 & \cdots \\
 & x_i & & 0 & \\
 & 0 & & 0 & \\
 & 0 & & y_i & \\
 & 0 & & 0 &
\end{array}\right) &
\begin{array}{c}
\vdots \\ u_0 m \\ \vdots \\ u_0^{a_s - a_i} u_1 u_2^{a_i} m \\ \vdots \\ u_2 m \\ \vdots \\ u_3 m \\ \vdots
\end{array} \\
\begin{array}{ccc} \alpha_i m & \cdots & \beta_i m \end{array} &
\end{array}$$

Example 3.9. Suppose X is the Hirzebruch surface of type a, i.e., the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a))$ over $\mathbb{P}^1$. We have $r = s = 1$ and $a_1 = a$. The Koszul complex K on $\alpha_0, \alpha_1, \beta_0, \beta_1$, twisted by $(0, 0, 0, 0, 1, 1)$, looks like

$$\begin{aligned}
&\underbrace{S_E(0,0,0,0,1,1)}_{1} \\
&\qquad \leftarrow \underbrace{S_E(-1,0,0,0,0,1)^2}_{\alpha_0,\,\alpha_1} \oplus \underbrace{S_E(0,-1,0,0,1,0)}_{\beta_0} \oplus \underbrace{S_E(a,-1,0,0,1,0)}_{\beta_1} \\
&\qquad \leftarrow \underbrace{S_E(-2,0,0,0,-1,1)}_{\alpha_0\alpha_1} \oplus \underbrace{S_E(-1,-1,0,0,0,0)^2}_{\alpha_0\beta_0,\,\alpha_1\beta_0} \oplus \underbrace{S_E(a-1,-1,0,0,0,0)^2}_{\alpha_0\beta_1,\,\alpha_1\beta_1} \\
&\qquad\qquad \oplus \underbrace{S_E(a,-2,0,0,1,-1)}_{\beta_0\beta_1} \\
&\qquad \leftarrow \underbrace{S_E(a-2,-1,0,0,-1,0)}_{\alpha_0\alpha_1\beta_1} \oplus \underbrace{S_E(-2,-1,0,0,-1,0)}_{\alpha_0\alpha_1\beta_0} \oplus \underbrace{S_E(a-1,-2,0,0,0,-1)^2}_{\alpha_0\beta_0\beta_1,\,\alpha_1\beta_0\beta_1} \\
&\qquad\qquad \leftarrow \underbrace{S_E(a-2,-2,0,0,-1,-1)}_{\alpha_0\alpha_1\beta_0\beta_1}.
\end{aligned}$$

Letting $M_{i,j}$ be as in Notation 3.7 (with $a_s = a$), we have:

$$\begin{aligned}
M_{0,0} &= \{1\}, \\
M_{1,0} &= \{u_0, u_2\}, \\
M_{0,1} &= \{u_3\} \cup \{u_0^k u_1 u_2^\ell : k + \ell = a\}, \\
M_{-1,1} &= \{u_0^k u_1 u_2^\ell : k + \ell = a - 1\}, \\
M_{1,1} &= \{u_0u_3, u_2u_3\} \cup \{u_0^k u_1 u_2^\ell : k + \ell = a + 1\}, \\
M_{i,j} &= \varnothing \quad \text{for } (i, j) \in \{(1,-1), (-1,0), (0,-1), (-1,-1)\}.
\end{aligned}$$

It follows that the complex R has terms as follows:

$$\begin{aligned}
R_0 &= \underbrace{S(-1,-1,1,1)}_{u_0^{a+1}u_1} \oplus \underbrace{S(0,-1,0,1)}_{u_0^a u_1 u_2} \oplus \cdots \oplus \underbrace{S(a,-1,-a,1)}_{u_1u_2^{a+1}} \oplus \underbrace{S(-1,0,1,0)}_{u_0u_3} \oplus \underbrace{S(0,0,0,0)}_{u_2u_3}, \\
R_1 &= \underbrace{S(-1,-1,0,1)^2}_{\alpha_0u_0^au_1,\,\alpha_1u_0^au_1} \oplus \underbrace{S(0,-1,-1,1)^2}_{\alpha_0u_0^{a-1}u_1u_2,\,\alpha_1u_0^{a-1}u_1u_2} \oplus \cdots \oplus \underbrace{S(a-1,-1,-a,1)^2}_{\alpha_0u_1u_2^a,\,\alpha_1u_1u_2^a} \oplus \underbrace{S(-1,0,0,0)^2}_{\alpha_0u_3,\,\alpha_1u_3} \\
&\qquad \oplus \underbrace{S(-1,-1,1,0)}_{\beta_0u_0} \oplus \underbrace{S(a-1,-1,1,0)}_{\beta_1u_0} \oplus \underbrace{S(0,-1,0,0)}_{\beta_0u_2} \oplus \underbrace{S(a,-1,0,0)}_{\beta_1u_2}, \\
R_2 &= \underbrace{S(-1,-1,-1,1)}_{\alpha_0\alpha_1u_0^{a-1}u_1} \oplus \underbrace{S(0,-1,-2,1)}_{\alpha_0\alpha_1u_0^{a-2}u_1u_2} \oplus \dots \\
&\qquad \oplus \underbrace{S(a-2,-1,-a,1)}_{\alpha_0\alpha_1u_1u_2^{a-1}} \oplus \underbrace{S(-1,-1,0,0)^2}_{\alpha_0\beta_0,\,\alpha_1\beta_0} \oplus \underbrace{S(a-1,-1,0,0)^2}_{\alpha_0\beta_1,\,\alpha_1\beta_1}.
\end{aligned}$$

The differentials $\partial_1\colon R_0 \leftarrow R_1$ and $\partial_2\colon R_1 \leftarrow R_2$ are given, respectively, by the matrices

$$\partial_1 = \begin{pmatrix}
-x_0' & -x_1' & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & -y_0' & 0 & 0 & 0 \\
x_0 & x_1 & -x_0' & -x_1' & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -y_0' & 0 \\
0 & 0 & x_0 & x_1 & -x_0' & -x_1' & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & x_0 & x_1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & -x_0' & -x_1' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & x_0 & x_1 & -x_0' & -x_1' & 0 & 0 & 0 & -y_1' & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & x_0 & x_1 & 0 & 0 & 0 & 0 & 0 & -y_1' \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -x_0' & -x_1' & y_0 & y_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & x_0 & x_1 & 0 & 0 & y_0 & y_1
\end{pmatrix}$$

and

$$\partial_2 = \begin{pmatrix}
x_1' & 0 & \cdots & 0 & y_0' & 0 & 0 & 0 \\
-x_0' & 0 & \cdots & 0 & 0 & y_0' & 0 & 0 \\
-x_1 & x_1' & \cdots & 0 & 0 & 0 & 0 & 0 \\
x_0 & -x_0' & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & -x_1 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & x_0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \\
0 & 0 & \cdots & x_1' & 0 & 0 & 0 & 0 \\
0 & 0 & \cdots & -x_0' & 0 & 0 & 0 & 0 \\
0 & 0 & \cdots & -x_1 & 0 & 0 & y_1' & 0 \\
0 & 0 & \cdots & x_0 & 0 & 0 & 0 & y_1' \\
0 & 0 & \cdots & 0 & -y_0 & 0 & -y_1 & 0 \\
0 & 0 & \cdots & 0 & 0 & -y_0 & 0 & -y_1 \\
0 & 0 & \cdots & 0 & -x_0' & -x_1' & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & -x_0' & -x_1' \\
0 & 0 & \cdots & 0 & x_0 & x_1 & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & x_0 & x_1
\end{pmatrix}.$$

As predicted by Theorem 1.1 parts (2) and (3), the differentials in R are linear; and the ranks of R_0, R_1, and R_2 are $a+4$, $2a+8$, and $a+4$, respectively.

3D. ***A Fourier–Mukai transform.*** Let π_1 and π_2 denote the projections of $X \times X$ onto X, and let $\Phi_{\mathcal{R}}$ denote the following Fourier–Mukai transform:

$$\Phi_{\mathcal{R}}\colon\ \mathrm{D^b}(X) \xrightarrow{\pi_1^*} \mathrm{D^b}(X \times X) \xrightarrow{\cdot \otimes \mathcal{R}} \mathrm{D^b}(X \times X) \xrightarrow{\boldsymbol{R}\pi_{2*}} \mathrm{D^b}(X).$$

We will prove that $\mathcal{R}$ is a resolution of the diagonal by showing that $\Phi_{\mathcal{R}}$ is isomorphic to the identity functor, and we will do so by directly exhibiting a natural isomorphism $\Phi_\nu : \Phi_{\mathcal{R}} \to \Phi_{\mathcal{O}_\Delta}$. In fact, we show this by proving that Φ_ν induces a quasiisomorphism on a full exceptional collection. To perform this calculation, we will need an explicit model for the functor $\Phi_{\mathcal{R}}$, which we present in this section. We refer the reader to [Huybrechts 2006, Section 8.3] for further background.

Let $\mathrm{coh}(X)$ denote the category of coherent sheaves on X, and suppose $\mathcal{F}_1, \mathcal{F}_2 \in \mathrm{coh}(X)$, where $\mathcal{F}_1$ is locally free. By the projection formula and base change, we have canonical isomorphisms

$$\boldsymbol{R}\pi_{2*}(\mathcal{F}_1 \boxtimes \mathcal{F}_2) \cong \boldsymbol{R}\pi_{2*}\pi_1^*(\mathcal{F}_1) \otimes_{\mathcal{O}_X} \mathcal{F}_2 \cong \boldsymbol{R}\Gamma(X, \mathcal{F}_1) \otimes_k \mathcal{F}_2$$

in $\mathrm{D^b}(X)$. Given $\mathcal{F} \in \mathrm{coh}(X)$, we can use this to explicitly compute $\Phi_{\mathcal{R}}(\mathcal{F})$ as follows. Given $\mathcal{G} \in \mathrm{coh}(X)$, let $\check{C}_{\mathcal{G}}$ denote the Čech complex of $\mathcal{G}$ associated to the affine open cover of X arising from the maximal cones in its fan. Consider the following bicomplex, where the horizontal maps are induced by the differentials in $\mathcal{R}$, the vertical maps are induced by the Čech differentials, N is the length of $\mathcal{R}$, and "$\mathcal{L}_1 \boxtimes \mathcal{L}_2 \in \mathcal{R}_i$" is shorthand for "$\mathcal{L}_1 \boxtimes \mathcal{L}_2$ is a summand of $\mathcal{R}_i$":

$$0 \leftarrow \bigoplus_{\mathcal{L}_1 \boxtimes \mathcal{L}_2 \in \mathcal{R}_0} \check{C}_{\mathcal{F}\otimes\mathcal{L}_1} \otimes \mathcal{L}_2 \leftarrow \cdots \leftarrow \bigoplus_{\mathcal{L}_1 \boxtimes \mathcal{L}_2 \in \mathcal{R}_N} \check{C}_{\mathcal{F}\otimes\mathcal{L}_1} \otimes \mathcal{L}_2 \leftarrow 0. \tag{3-3}$$

Since the differentials of $\check{C}_{\mathcal{G}}$ have entries in $\Bbbk$, the columns of (3-3) split. Thus, we may apply [Eisenbud et al. 2003, Lemma 3.5] to conclude that the totalization of (3-3) is homotopy equivalent to a complex $\boldsymbol{B}(\mathcal{F})$ concentrated in degrees $k = -N, \dots, N$ with terms

$$\boldsymbol{B}(\mathcal{F})_k = \bigoplus_{i-j=k} \bigoplus_{\mathcal{L}_1 \boxtimes \mathcal{L}_2 \in \mathcal{R}_i} H^j(X, \mathcal{F} \otimes \mathcal{L}_1) \otimes \mathcal{L}_2 \cong \bigoplus_{i-j=k} \boldsymbol{R}^j \pi_{2*}(\pi_1^* \mathcal{F} \otimes \mathcal{R}_i). \tag{3-4}$$

The terms of $\boldsymbol{B}(\mathcal{F})$ arise from the totalization of the vertical homology of (3-3).

Over projective space, the analogue of this Fourier–Mukai transform involving Beilinson's resolution of the diagonal is called the Beilinson monad (see e.g., [Eisenbud et al. 2003]), hence the notation $\boldsymbol{B}(-)$. Note that "the" complex $\boldsymbol{B}(\mathcal{F})$ is only well-defined up to homotopy equivalence, since the differential depends on a choice of splitting of the columns in the bicomplex (3-3). More precisely, for each term $Y_{i,j}$ of (3-3), choose a decomposition $Y_{i,j} = B_{i,j} \oplus H_{i,j} \oplus L_{i,j}$ such that $B_{i,j} \oplus H_{i,j} = Z_{i,j}^{\mathrm{vert}}$, where $Z_{i,j}^{\mathrm{vert}}$ denotes the vertical cycles in $Y_{i,j}$. Notice that there is a canonical isomorphism $H_{i,j} \cong \bigoplus_{\mathcal{L}_1 \boxtimes \mathcal{L}_2 \in \mathcal{R}_i} H^{-j}(\mathcal{F} \otimes \mathcal{L}_1) \otimes \mathcal{L}_2$. Let $\sigma_H : Y_{\bullet,\bullet} \to H_{\bullet,\bullet}$ and $\sigma_B : Y_{\bullet,\bullet} \to B_{\bullet,\bullet}$ denote the projections, let $g : L_{\bullet,\bullet} \xrightarrow{\cong} B_{\bullet,\bullet-1}$ denote the isomorphism induced by the vertical differential, and let $\pi = g^{-1}\sigma_B$. By [loc. cit., Lemma 3.5], the differential on $\boldsymbol{B}(\mathcal{F})$ is given by

$$\partial_{\boldsymbol{B}(\mathcal{F})} = \sum_{i \geq 0} \sigma_H (d_{\mathrm{hor}} \pi)^i d_{\mathrm{hor}},$$

where d_{hor} is the horizontal differential in the bicomplex (3-3).

Remark 3.10. The $i = 0$ term in the formula for $\partial_{\boldsymbol{B}(\mathcal{F})}$ is simply the map induced by the differential on $\mathcal{R}$; it is independent of the choices of splittings of the columns of (3-3). Since this is the only part of the differential on $\boldsymbol{B}(\mathcal{F})$ that we will need to explicitly compute, we will ignore the ambiguity of $\boldsymbol{B}(\mathcal{F})$ up to homotopy equivalence from now on.

3E. *Proof of Theorem 1.1.*

Proof. To prove parts (1) and (2), first recall that R is the direct sum of the degree $(d_1, d_2, d_3, d_4, 0, 0)$ components of $K(0, 0, 0, 0, r, s)$ for all $d_1, \dots, d_4 \in \mathbb{Z}$. Thus, since K is exact in positive homological degrees, R is as well; moreover, the differentials of R are linear.[1] We now check that R has property (3). For all $k, \ell \in \mathbb{Z}$, we have

$$\dim_{\Bbbk} H^0(\mathbb{F}_{a_s}, \mathcal{O}(k, \ell)) = \begin{cases} (k+1)(\ell+1) + \binom{\ell+1}{2} a_s, & \ell \geq 0; \\ 0, & \ell < 0. \end{cases} \tag{3-5}$$

We now compute

$$\begin{aligned}
\operatorname{rank} \mathcal{R}_n &= \sum_{i=0}^{n} \binom{r+1}{i} \binom{s+1}{n-i} \dim_{\Bbbk} H^0(\mathbb{F}_{a_s}, \mathcal{O}(r-i, s-(n-i))) \\
&= \sum_{i=0}^{r} \binom{r+1}{i} \binom{s+1}{n-i} \left((r-i+1)(s-(n-i)+1) + \binom{s-(n-i)+1}{2} a_s \right) \\
&\qquad\qquad + \binom{s+1}{n-(r+1)} \binom{s-(n-(r+1))+1}{2} a_s \\
&= \sum_{i=0}^{r} \binom{r}{i} \binom{s}{n-i} (r+1)(s+1) + \sum_{i=0}^{r} \binom{r+1}{i} \binom{s-1}{n-i} \binom{s+1}{2} a_s + \binom{s-1}{n-(r+1)} \binom{s+1}{2} a_s \\
&= \sum_{i=0}^{r} \binom{r}{i} \binom{s}{n-i} (r+1)(s+1) + \sum_{i=0}^{r+1} \binom{r+1}{i} \binom{s-1}{n-i} \binom{s+1}{2} a_s \\
&= \binom{r+s}{n} \dim_{\Bbbk} H^0(\mathbb{F}_{a_s}, \mathcal{O}(r, s)).
\end{aligned}$$

The first equality follows from the definition of $\mathcal{R}$, the second from (3-5), the third from some straightforward manipulations, the fourth by combining the second and third terms, and the last by Vandermonde's identity and the equality $\dim_{\Bbbk} H^0(\mathbb{F}_{a_s}, \mathcal{O}(r, s)) = (r+1)(s+1) + \binom{s+1}{2} a_s$. This proves (3).

Finally, we check property (4): namely, that the cokernel of the differential $\partial_1 : \mathcal{R}_1 \to \mathcal{R}_0$ is $\mathcal{O}_\Delta$. Just as in the proof of [Canonaco and Karp 2008, Proposition 3.2], we will prove that $\mathcal{R}$ is a resolution of $\mathcal{O}_\Delta$ by showing there is a chain map $\mathcal{R} \to \mathcal{O}_\Delta$ that induces a natural isomorphism on certain Fourier–Mukai transforms. In detail: given any $i, j \in \mathbb{Z}$, there is a natural map $\mathcal{O}(i, j, -i, -j) \to \mathcal{O}_\Delta$ given by multiplication. These maps determine a natural map $\nu_0 : \mathcal{R}_0 \to \mathcal{O}_\Delta$, and it is clear from the description of ∂_1 in Example 3.8 that ν_0 determines a chain map $\nu : \mathcal{R} \to \mathcal{O}_\Delta$. Recall that $\Phi_{\mathcal{R}}$ denotes the Fourier–Mukai transform associated to $\mathcal{R}$. To show that ν is a quasiisomorphism, we need only prove that the induced natural transformation $\Phi_\nu : \Phi_{\mathcal{R}} \to \Phi_{\mathcal{O}_\Delta}$ on Fourier–Mukai transforms is a natural isomorphism; indeed, this immediately implies that $\Phi_{\operatorname{cone}(\nu)}$ is isomorphic to the 0 functor, and so $\operatorname{cone}(\nu) = 0$ by [loc. cit., Lemma 2.1].

The category $\mathrm{D^b}(X)$ is generated by the line bundles $\mathcal{O}(b, c)$ with $0 \leq b \leq r$ and $0 \leq c \leq s$; in fact, these bundles form a full exceptional collection in $\mathrm{D^b}(X)$ [Orlov 1992, Corollary 2.7]. Since $\Phi_{\mathcal{O}_\Delta}$ is the identity functor, we need only show that the map $\Phi_{\mathcal{R}}(\mathcal{O}(b, c)) \to \mathcal{O}(b, c)$ induced by Φ_ν is an isomorphism in $\mathrm{D^b}(X)$.

[1] Free complexes that are linear in the sense of Theorem 1.1(2) are called *strongly linear* in [Brown and Erman 2024].

Say $\mathcal{O}(d_1, d_2, d_3, d_4)$ is a summand of $\mathcal{R}$. We first show that the line bundle $\mathcal{O}(d_1+b, d_2+c)$ on X is acyclic, i.e., $H^i(X, \mathcal{O}(d_1+b, d_2+c)) = 0$ for $i > 0$. Say the summand $\mathcal{O}(d_1, d_2, d_3, d_4)$ of $\mathcal{R}$ corresponds to the monomial $\alpha_{i_1} \cdots \alpha_{i_k} \beta_{j_1} \cdots \beta_{j_\ell} m$, where $k \le r+1$, $\ell \le s+1$, and $m \in M_{r-k, s-\ell}$. It follows that $d_1 = -k - t_1$ and $d_2 = -\ell - t_2$ for some $t_1 \le r-k$ and $t_2 \le s-\ell$. In particular, we have $d_1 + b \ge d_1 \ge -r$, and $d_2 + c \ge d_2 \ge -s$. Thus, $\mathcal{O}(d_1+b, d_2+c)$ satisfies either (a) or (b) in Corollary 3.3(2), and so $\mathcal{O}(d_1+b, d_2+c)$ is acyclic.

Recall from Section 3D that, given any sheaf $\mathcal{F}$ on X, $\Phi_{\mathcal{R}}(\mathcal{F})$ may be modeled explicitly as the complex $\boldsymbol{B}(\mathcal{F})$. The previous paragraph implies that the terms in $\boldsymbol{B}(\mathcal{O}(b, c))$ involving higher cohomology vanish; that is, the nonzero terms of $\boldsymbol{B}(\mathcal{O}(b, c))$ are of the form $H^0(\mathcal{L}_1(b, c)) \otimes \mathcal{L}_2$, where $\mathcal{L}_1 \boxtimes \mathcal{L}_2$ is a summand of $\mathcal{R}$. In particular, $\boldsymbol{B}(\mathcal{O}(b, c))$ is concentrated in nonnegative degrees, the map $\boldsymbol{B}_0(\mathcal{O}(b, c)) \to \mathcal{O}(b, c)$ induced by ν is the natural multiplication map, and the differential on $\boldsymbol{B}(\mathcal{O}(b, c))$ is induced by the differential on $\mathcal{R}$. It follows that $\boldsymbol{B}(\mathcal{O}(b, c))$ is exact in positive degrees, since $\mathcal{R}$ has this property. We now show, by direct computation, that the induced map $H_0(\boldsymbol{B}(\mathcal{O}(b, c))) \to \mathcal{O}(b, c)$ is an isomorphism.

It follows from our explicit descriptions of the terms R_0 and R_1 in Example 3.8 that

$$\boldsymbol{B}(\mathcal{O}(b, c))_0 = \bigoplus_{m \in M_{r,s}} H^0(X, \mathcal{O}(b - d_1^m, c - d_2^m)) \otimes \mathcal{O}(d_1^m, d_2^m) \cdot m, \quad \text{and}$$
$$\boldsymbol{B}(\mathcal{O}(b, c))_1 = \boldsymbol{B}(\mathcal{O}(b, c))_1^\alpha \oplus \boldsymbol{B}(\mathcal{O}(b, c))_1^\beta,$$

where

$$\boldsymbol{B}(\mathcal{O}(b, c))_1^\alpha = \bigoplus_{i=0}^{r} \bigoplus_{m \in M_{r-1,s}} H^0(X, \mathcal{O}(b - d_1^m - 1, c - d_2^m)) \otimes \mathcal{O}(d_1^m, d_2^m) \cdot \alpha_i m,$$
$$\boldsymbol{B}(\mathcal{O}(b, c))_1^\beta = \bigoplus_{i=0}^{s} \bigoplus_{m \in M_{r,s-1}} H^0(X, \mathcal{O}(b - d_1^m + a_i, c - d_2^m - 1)) \otimes \mathcal{O}(d_1^m, d_2^m) \cdot \beta_i m.$$

We represent the first differential on $\boldsymbol{B}(\mathcal{O}(b, c))$ as a matrix with respect to the above decomposition, along with the monomial bases of each cohomology group. The column of this matrix corresponding to $\alpha_i m$ and a monomial z in the Cox ring $S = k[x_0, \ldots, x_r, y_0, \ldots, y_s]$ of X of degree $(b - d_1^m - 1, c - d_2^m)$ has exactly two nonzero entries:

- An entry of 1 for $u_2 m \in M_{r,s}$ and $x_i z \in H^0(X, \mathcal{O}(b - d_1^{u_2 m}, c - d_2^{u_2 m}))$.
- An entry of $-x_i'$ for $u_0 m \in M_{r,s}$ and $z \in H^0(X, \mathcal{O}(b - d_1^{u_0 m}, c - d_2^{u_0 m}))$.

Similarly, the column corresponding to $\beta_i m$ and a monomial $w \in S$ of degree $(b - d_1^m + a_i, c - d_2^m - 1)$ has exactly two nonzero entries:

- An entry of 1 for $u_3 m$ and $y_i w \in H^0(X, \mathcal{O}(b - d_1^{u_3 m}, c - d_2^{u_3 m}))$.
- An entry of $-y_i'$ for $u_0^{a_s - a_i} u_1 u_2^{a_i} m$ and $w \in H^0(X, \mathcal{O}(b - d_1^{u_0^{a_s - a_i} u_1 u_2^{a_i} m}, c - d_2^{u_0^{a_s - a_i} u_1 u_2^{a_i} m}))$.

That is, the first differential on $\boldsymbol{B}(\mathcal{O}(b,c))$ has the following form:

$$
\begin{array}{c c}
\left(\begin{array}{ccccc}
 & 0 & & 0 & \\
 & -x_i' & & 0 & \\
 & 0 & & 0 & \\
 & 0 & & -y_i' & \\
\cdots & 0 & \cdots & 0 & \cdots \\
 & 1 & & 0 & \\
 & 0 & & 0 & \\
 & 0 & & 1 & \\
 & 0 & & 0 &
\end{array}\right)
&
\begin{array}{c}
\vdots \\
z \otimes u_0 m \\
\vdots \\
w \otimes u_0^{a_s - a_i} u_1 u_2^{a_i} m \\
\vdots \\
x_i z \otimes u_2 m \\
\vdots \\
y_i w \otimes u_3 m \\
\vdots
\end{array}
\\
z \otimes \alpha_i m \quad \cdots \quad w \otimes \beta_i m &
\end{array}
$$

Now observe: every column of this matrix contains exactly one "1", and there is exactly one row that does not contain a "1": namely, the row corresponding to the summand $H^0(X, \mathcal{O}) \otimes \mathcal{O}(b,c) \cdot u_0^{b+ca_s} u_1^c u_2^{r-b} u_3^{s-c}$. It follows immediately that the cokernel of this matrix is isomorphic to the summand $H^0(X, \mathcal{O}) \otimes \mathcal{O}(b,c)$, and the multiplication map induced by ν from this summand to $\mathcal{O}(b,c)$ is clearly an isomorphism. □

Remark 3.11. Our construction of the resolution $\mathcal{R}$ realizes it as a subcomplex of the (infinite rank) resolution of the diagonal obtained in [Brown and Erman 2021, Theorem 4.1] and therefore yields a positive answer to [loc. cit., Conjecture 7.2] for smooth projective toric varieties of Picard rank 2.

Corollary 3.12. *Given a coherent sheaf $\mathcal{F}$ on X, we have $\boldsymbol{B}(\mathcal{F}) \cong \mathcal{F}$ in $\mathrm{D}^{\mathrm{b}}(X)$.*

Corollary 3.13. *Consider the ideal $I = (\alpha_0, \ldots, \alpha_r, \beta_0, \ldots, \beta_s) \subseteq S_E$, and let $\mathcal{D}$ denote the sheaf $\widetilde{S_E/I}$ on E. We have an isomorphism $\pi_* \mathcal{D}(0,0,0,0,r,s) \cong \mathcal{O}_\Delta$ of sheaves on $X \times X$.*

Proof. Recall that $\mathcal{K}$ is the sheafification of the Koszul complex on the generators of I, which form a regular sequence. Therefore $\mathcal{K}$ is a locally free resolution of $\mathcal{D}$, and using Proposition 3.6 and Theorem 1.1(4) we have $\pi_*\mathcal{D}(0,0,0,0,r,s) \cong \pi_*\mathcal{K}(0,0,0,0,r,s) \cong \mathcal{R} \cong \mathcal{O}_\Delta$. □

We will now prove Conjecture 1.2 for X as in Theorem 1.1.

Proof of Corollary 1.3. Our proof is nearly the same as that of [Berkesch et al. 2020, Proposition 1.2]. Given a finitely generated graded module M over the Cox ring of X, let $\mathcal{F}$ be the associated sheaf on X. Applying the Fujita Vanishing Theorem, choose $i, j \gg 0$ such that, for all summands $\mathcal{L}_1 \boxtimes \mathcal{L}_2$ of the resolution of the diagonal $\mathcal{R}$ from Theorem 1.1, we have $H^q(X, \mathcal{F}(i,j) \otimes \mathcal{L}_1) = 0$ for $q > 0$. The complex $\boldsymbol{B}(\mathcal{F}(i,j))$ is a resolution of $\mathcal{F}(i,j)$ of length at most $\dim(X)$ consisting of finite sums of line bundles, and twisting back by $(-i,-j)$ gives a resolution of $\mathcal{F}$. Now applying the functor $\mathcal{G} \mapsto \bigoplus_{(k,\ell)\in\mathbb{Z}^2} H^0(X, \mathcal{G}(k,\ell))$ to the complex $\boldsymbol{B}(\mathcal{F}(i,j))(-i,-j)$ gives a virtual resolution of M. □

4. A Horrocks-type splitting criterion

Let X denote the projective bundle $\mathbb{P}(\mathcal{O} \oplus \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_s))$ over $\mathbb{P}^r$, where $a_1 \le \cdots \le a_s$. Given a coherent sheaf $\mathcal{F}$ on X, let $\boldsymbol{B}(\mathcal{F})$ be the complex of sheaves on X defined in Section 3D. Recall from the introduction the notation $\gamma(\mathcal{F})$ for the cohomology table of $\mathcal{F}$. We will need the following technical result.

Lemma 4.1 (cf. [Eisenbud et al. 2015, Lemma 7.3]). *Let $\mathcal{E}$ be a vector bundle on X, and suppose we have $\gamma(\mathcal{E}) = \gamma(\mathcal{O}^m) + \gamma(\mathcal{E}')$ for some vector bundle $\mathcal{E}'$ on X with $\boldsymbol{B}(\mathcal{E}')_1 = 0$. There is an isomorphism $\mathcal{E} \cong \mathcal{O}^m \oplus \mathcal{E}''$ for some vector bundle $\mathcal{E}''$ such that $\gamma(\mathcal{E}'') = \gamma(\mathcal{E}')$.*

Proof. Let $\mathcal{R}$ be the resolution of the diagonal for X constructed in Section 3C. We have a Beilinson-type spectral sequence

$$E_1^{-i,j}(\mathcal{E}) = \boldsymbol{R}^j \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_i) \Rightarrow \boldsymbol{R}^{i-j} \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}) \cong \begin{cases} \mathcal{E}, & i = j; \\ 0, & i \neq j. \end{cases}$$

The first page looks as follows:

$$\begin{tikzcd}[column sep=small]
\vdots & & \vdots & & \vdots & \\
\boldsymbol{R}^2 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_0) & \arrow[l] \boldsymbol{R}^2 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_1) & & \arrow[ll] \boldsymbol{R}^2 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_2) & \arrow[l] \cdots \\
\boldsymbol{R}^1 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_0) & \arrow[l] \boldsymbol{R}^1 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_1) & & \arrow[ll] {\color{red}\boldsymbol{R}^1 \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_2)} & \arrow[l] \cdots \\
\pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_0) & \arrow[l] {\color{red}\pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_1)} & & \arrow[ll] \pi_{2*}(\pi_1^* \mathcal{E} \otimes \mathcal{R}_2) & \arrow[l] \cdots \\
k=0 & k=1 & & k=2 &
\end{tikzcd} \tag{4-1}$$

Notice that $E_1^{-i,j}(\mathcal{E}) \cong \bigoplus H^j(X, \mathcal{E} \otimes \mathcal{L}_1) \otimes \mathcal{L}_2$, where the direct sum ranges over the summands $\mathcal{L}_1 \boxtimes \mathcal{L}_2$ of $\mathcal{R}_i$. It follows that $\boldsymbol{B}(\mathcal{E})_1 = \bigoplus_{i-j=1} E_1^{-i,j}(\mathcal{E})$. Moreover, since the terms of the first page only depend on $\gamma(\mathcal{E})$, we have

$$E_1^{-i,j}(\mathcal{E}) = E_1^{-i,j}(\mathcal{O})^m \oplus E_1^{-i,j}(\mathcal{E}').$$

Observe that $E_1^{0,0}(\mathcal{O}) = \mathcal{O}$, and $E_1^{-i,j}(\mathcal{O}) = 0$ when either $i \neq 0$ or $j \neq 0$. In particular, we have $E_1^{0,0}(\mathcal{E}) = \mathcal{O}^m \oplus E_1^{0,0}(\mathcal{E}')$, and it follows from the hypothesis $\boldsymbol{B}(\mathcal{E}')_1 = 0$ that the terms along the $k = 1$ diagonal in (4-1) (colored in red) vanish. Thus, every differential in the spectral sequence with either source or target given by $E_r^{0,0}(\mathcal{E})$ for some r vanishes. We conclude that $\mathcal{O}^m$ is a summand of $E_\infty^{0,0}(\mathcal{E})$, and hence $\mathcal{E}$ as well. $\square$

Remark 4.2. A similar technique was recently utilized by Bruce, Cranton Heller, and Sayrafi [Bruce et al. 2021] to give a characterization of multigraded Castelnuovo–Mumford regularity on products of projective spaces. An interesting question is whether there is a similar result for smooth projective varieties of Picard rank 2 using the resolution $\mathcal{R}$.

We will now prove our splitting criterion.

Proof of Theorem 1.5. By induction, it suffices to show that $\mathcal{O}(b_1, c_1)^{m_1}$ is a summand of $\mathcal{E}$. Without loss of generality, we may assume $(b_2, c_2) < (b_1, c_1)$. We may also twist $\mathcal{E}$ so that $b_1 = c_1 = 0$, which implies

$(b_i, c_i) < 0$ for all $i > 1$. Suppose $(a, b) < 0$. By Lemma 4.1, it suffices to show that $\boldsymbol{B}(\mathcal{O}(a, b))_1 = 0$. This amounts to showing that, for $0 < n \leq r + s$, we have

$$H^{n-1}(X, \mathcal{O}(a, b) \otimes \mathcal{L}_1) = 0, \quad \text{when } \mathcal{L}_1 \boxtimes \mathcal{L}_2 \text{ is a summand of } \mathcal{R}_n. \tag{4-2}$$

By Corollary 3.3(1)(a), we need only show that (4-2) holds for $n \in \{1, r+1, s+1\}$. We recall that any summand of $\mathcal{R}_n$ corresponds to a monomial of the form

$$\alpha_{i_1} \cdots \alpha_{i_e} \cdot \beta_{j_1} \cdots \beta_{j_f} \cdot m, \tag{4-3}$$

where $e + f = n$, and $m \in M_{r-e,s-f}$ (using Notation 3.7). Writing $m = u_0^{c_0} u_1^{c_1} u_2^{c_2} u_3^{c_3}$, we have that the summand of $\mathcal{R}_n$ corresponding to (4-3) is $\mathcal{O}(-e - d_1, -f - d_2, d_1, d_2)$, where $d_1 = c_0 - a_s c_1$ and $d_2 = c_1$. In particular, we have $0 \leq d_2 \leq s - f$, which immediately implies that $-s \leq -f - d_2 \leq 0$. When $-f - d_2 < 0$, (4-2) holds for $n \in \{1, r+1, s+1\}$ by Corollary 3.3(1)(b - d). Suppose $-f - d_2 = 0$. Since $f, d_2 \geq 0$, we have $f = 0 = d_2$. It follows that $e = n$ and $d_1 = c_0 \geq 0$. Corollary 3.3(1)(b) therefore implies that (4-2) holds when $n = 1$. Corollary 3.3(1)(c) implies that (4-2) holds for $n = s + 1$ when $r \neq s$; we may thus reduce to the case where $n = r + 1$. But this case cannot occur, since there is no $m \in M_{-1,s}$ of the form $u_0^{c_0} u_2^{c_2} u_3^{c_3}$. □

Remark 4.3. If we replace the nef ordering with the effective ordering in the statement of Theorem 1.5, our proof fails. The problem arises in the final step: there exist line bundles $\mathcal{O}(b, c) \in \operatorname{Pic} X$ such that $-(b, c)$ is effective but $\boldsymbol{B}(\mathcal{O}(b, c))_1 \neq 0$. For instance, over a Hirzebruch surface of type a, the divisor $-(a, -1)$ is effective, and $\boldsymbol{B}(\mathcal{O}(a, -1))_1 \neq 0$.

Acknowledgments

We thank Matthew Ballard, Christine Berkesch, Lauren Cranton Heller, David Eisenbud, Daniel Erman, David Favero, and Frank-Olaf Schreyer for valuable conversations. We also thank the referee for many helpful comments that improved this paper. The computer algebra system `Macaulay2` [Grayson and Stillman 1992] provided indispensable assistance throughout this project. The second author was partially supported by the NSF grant DMS-2001101.

References

[Aprodu and Marchitan 2011] M. Aprodu and M. Marchitan, "A note on vector bundles on Hirzebruch surfaces", *C. R. Math. Acad. Sci. Paris* **349**:11-12 (2011), 687–690. MR Zbl

[Bai and Côté 2023] S. Bai and L. Côté, "On the Rouquier dimension of wrapped Fukaya categories and a conjecture of Orlov", *Compos. Math.* **159**:3 (2023), 437–487. MR Zbl

[Ballard and Favero 2012] M. Ballard and D. Favero, "Hochschild dimensions of tilting objects", *Int. Math. Res. Not.* **2012**:11 (2012), 2607–2645. MR Zbl

[Ballard et al. 2019] M. Ballard, D. Favero, and L. Katzarkov, "Variation of geometric invariant theory quotients and derived categories", *J. Reine Angew. Math.* **746** (2019), 235–303. MR Zbl

[Berkesch et al. 2020] C. Berkesch, D. Erman, and G. G. Smith, "Virtual resolutions for a product of projective spaces", *Algebr. Geom.* **7**:4 (2020), 460–481. MR Zbl

[Beĭlinson 1978] A. A. Beĭlinson, "Coherent sheaves on $\mathbb{P}^n$ and problems in linear algebra", *Funktsional. Anal. i Prilozhen.* **12**:3 (1978), 68–69. In Russian; translated in *Functional Anal. Appl.* **12**:3 (1978), 214–216. MR Zbl

[Borisov and Hua 2009] L. Borisov and Z. Hua, "On the conjecture of King for smooth toric Deligne–Mumford stacks", *Adv. Math.* **221**:1 (2009), 277–301. MR Zbl

[Brown and Erman 2021] M. K. Brown and D. Erman, "Tate resolutions on toric varieties", 2021. To appear in *J. Eur. Math. Soc.* arXiv 2108.03345

[Brown and Erman 2024] M. K. Brown and D. Erman, "Linear strands of multigraded free resolutions", *Math. Ann.* (online publication February 2024).

[Bruce et al. 2021] J. Bruce, L. Cranton Heller, and M. Sayrafi, "Characterizing multigraded regularity on products of projective spaces", preprint, 2021. arXiv 2110.10705

[Buchdahl 1987] N. P. Buchdahl, "Stable 2-bundles on Hirzebruch surfaces", *Math. Z.* **194**:1 (1987), 143–152. MR Zbl

[Canonaco and Karp 2008] A. Canonaco and R. L. Karp, "Derived autoequivalences and a weighted Beilinson resolution", *J. Geom. Phys.* **58**:6 (2008), 743–760. MR Zbl

[Costa and Miró-Roig 2004] L. Costa and R. M. Miró-Roig, "Tilting sheaves on toric varieties", *Math. Z.* **248**:4 (2004), 849–865. MR Zbl

[Costa and Miró-Roig 2005] L. Costa and R. M. Miró-Roig, "Cohomological characterization of vector bundles on multiprojective spaces", *J. Algebra* **294**:1 (2005), 73–96. MR Zbl

[Cox et al. 2011] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Grad. Stud. in Math. **124**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Grad. Texts in Math. **150**, Springer, 1995. MR Zbl

[Eisenbud and Schreyer 2003] D. Eisenbud and F.-O. Schreyer, "Resultants and Chow forms via exterior syzygies", *J. Amer. Math. Soc.* **16**:3 (2003), 537–579. MR Zbl

[Eisenbud and Schreyer 2009] D. Eisenbud and F.-O. Schreyer, "Betti numbers of graded modules and cohomology of vector bundles", *J. Amer. Math. Soc.* **22**:3 (2009), 859–888. MR Zbl

[Eisenbud et al. 2003] D. Eisenbud, G. Fløystad, and F.-O. Schreyer, "Sheaf cohomology and free resolutions over exterior algebras", *Trans. Amer. Math. Soc.* **355**:11 (2003), 4397–4426. MR Zbl

[Eisenbud et al. 2015] D. Eisenbud, D. Erman, and F.-O. Schreyer, "Tate resolutions for products of projective spaces", *Acta Math. Vietnam.* **40**:1 (2015), 5–36. MR Zbl

[Fulger and Marchitan 2011] M. Fulger and M. Marchitan, "Some splitting criteria on Hirzebruch surfaces", *Bull. Math. Soc. Sci. Math. Roumanie (N.S.)* **54(102)**:4 (2011), 313–323. MR Zbl

[Grayson and Stillman 1992] D. R. Grayson and M. E. Stillman, "Macaulay2, a software system for research in algebraic geometry", 1992, available at https://macaulay2.com/.

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Grad. Texts in Math. **52**, Springer, 1977. MR Zbl

[Horrocks 1964] G. Horrocks, "Vector bundles on the punctured spectrum of a local ring", *Proc. Lond. Math. Soc.* (3) **14** (1964), 689–713. MR Zbl

[Huybrechts 2006] D. Huybrechts, *Fourier–Mukai transforms in algebraic geometry*, Oxford Univ. Press, 2006. MR Zbl

[Kapranov 1988] M. M. Kapranov, "On the derived categories of coherent sheaves on some homogeneous spaces", *Invent. Math.* **92**:3 (1988), 479–508. MR Zbl

[King 1997] A. King, "Tilting bundles on some rational surfaces", unpublished paper, 1997, available at https://people.bath.ac.uk/masadk/papers/tilt.pdf.

[Kleinschmidt 1988] P. Kleinschmidt, "A classification of toric varieties with few generators", *Aequationes Math.* **35**:2-3 (1988), 254–266. MR Zbl

[Köck 2001] B. Köck, "Computing the homology of Koszul complexes", *Trans. Amer. Math. Soc.* **353**:8 (2001), 3115–3147. MR Zbl

[Lasoń and Michałek 2011] M. Lasoń and M. Michałek, "On the full, strongly exceptional collections on toric varieties with Picard number three", *Collect. Math.* **62**:3 (2011), 275–296. MR Zbl

[Orlov 1992] D. O. Orlov, "Projective bundles, monoidal transformations, and derived categories of coherent sheaves", *Izv. Ross. Akad. Nauk Ser. Mat.* **56**:4 (1992), 852–862. In Russian; translated in *Russian Acad. Sci. Izv. Math.* **41**:1 (1993), 133–141. MR Zbl

[Orlov 2009] D. Orlov, "Remarks on generators and dimensions of triangulated categories", *Mosc. Math. J.* **9**:1 (2009), 153–159. MR Zbl

[Ottaviani 1989] G. Ottaviani, "Some extensions of Horrocks criterion to vector bundles on Grassmannians and quadrics", *Ann. Mat. Pura Appl.* (4) **155** (1989), 317–341. MR Zbl

[Rouquier 2008] R. Rouquier, "Dimensions of triangulated categories", *J. K-Theory* **1**:2 (2008), 193–256. MR Zbl

[Schreyer 2022] F.-O. Schreyer, "Horrocks splitting on Segre–Veronese varieties", *Math. Z.* **300**:2 (2022), 1351–1358. MR Zbl

[Thomason and Trobaugh 1990] R. W. Thomason and T. Trobaugh, "Higher algebraic K-theory of schemes and of derived categories", pp. 247–435 in *The Grothendieck Festschrift*, *III*, edited by P. Cartier et al., Progr. Math. **88**, Birkhäuser, Boston, MA, 1990. MR Zbl

[Weyman 2003] J. Weyman, *Cohomology of vector bundles and syzygies*, Cambridge Tracts in Math. **149**, Cambridge Univ. Press, 2003. MR Zbl

[Yang 2021] J. Yang, "Virtual resolutions of monomial ideals on toric varieties", *Proc. Amer. Math. Soc. Ser. B* **8** (2021), 100–111. MR Zbl

[Yasutake 2015] K. Yasutake, "A cohomological splitting criterion for rank 2 vector bundles on Hirzebruch surfaces", *Tokyo J. Math.* **38**:2 (2015), 327–330. MR Zbl

Communicated by Irena Peeva
Received 2022-11-11 Revised 2023-08-30 Accepted 2023-10-12

mkb0096@auburn.edu *Department of Mathematics and Statistics, Auburn University, Auburn, AL, United States*

mahrud@umn.edu *School of Mathematics, University of Minnesota, Minneapolis, MN, United States*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 18 No. 10 2024