Algebra & Number Theory

Volume 12
2018
No. 10

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2018 is US $340/year for the electronic version, and $535/year (+$55, 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. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2237

Higher weight on GL(3), II: The cusp forms

Jack Buttcane

The purpose of this paper is to collect, extend, and make explicit the results of Gel'fand, Graev and Piatetski-Shapiro and Miyazaki for the GL(3) cusp forms which are nontrivial on $SO(3,\mathbb{R})$. We give new descriptions of the spaces of cusp forms of minimal K-type and from the Fourier–Whittaker expansions of such forms give a complete and completely explicit spectral expansion for $L^2(SL(3,\mathbb{Z})\backslash PSL(3,\mathbb{R}))$, accounting for multiplicities, in the style of Duke, Friedlander and Iwaniec's paper. We do this at a level of uniformity suitable for Poincaré series which are not necessarily K-finite. We directly compute the Jacquet integral for the Whittaker functions at the minimal K-type, improving Miyazaki's computation. These results will form the basis of the nonspherical spectral Kuznetsov formulas and the arithmetic/geometric Kuznetsov formulas on GL(3). The primary tool will be the study of the differential operators coming from the Lie algebra on vector-valued cusp forms.

1. Introduction

In the previous paper [Buttcane 2018], we worked out the continuous and residual spectra in the Langlands decomposition in the case of $L^2(SL(3,\mathbb{Z})\backslash PSL(3,\mathbb{R}))$. We turn now specifically to the cuspidal spectral decomposition. The initial decomposition into eigenspaces of the Casimir operators is originally due, in much greater generality, to Gel'fand, Graev and Piatetski-Shapiro [Gel'fand et al. 1969], who described

During the time of this research, the author was supported by NSF grant DMS-1601919.
MSC2010: primary 11F72; secondary 11F30.
Keywords: Maass forms, automorphic forms, GL(3), SO(3), weight, raising and lowering operators.

the decomposition in terms of integral operators (in the style of [Selberg 1956]) operating on representation spaces in the L^2-space. In modern terminology, this is the study of irreducible unitary representations and $(\mathfrak{g}, K)$ modules, and in the particular case of $\mathrm{PSL}(3, \mathbb{R})$, the former was initiated by Vahutinskiĭ [1968] and the latter by Howe [2000] and Miyazaki [2008].

From the representation-theoretic description, this paper will analyze the structure of the cuspidal part of the principal series representations as they decompose over the entries of the Wigner $\mathcal{D}$-matrices, the intertwining operators, and the Jacquet integral as an operator from the principal series representation to its Whittaker model. From the analytic perspective, we are decomposing the eigenspaces of the Casimir operators on the L^2-space via raising and lowering operators (in the style of [Maass 1952]) obtained from the Lie algebra; we show the equivalence of several descriptions of the minimal K-types, and compute Mellin–Barnes integrals for the Whittaker functions attached to those K-types. We will largely avoid the representation-theoretic description outside of the Introduction, Section 3, and Appendix A; partial descriptions in that language maybe found in [Miller and Schmid 2011, Appendix A] and [Buttcane and Miller 2017].

The goal of this paper is to describe, as completely, explicitly and uniformly as possible, the spectral expansion of the cuspidal part of $L^2(\mathrm{SL}(3, \mathbb{Z})\backslash\mathrm{PSL}(3, \mathbb{R}))$, and provide all of the associated information necessary for a number theorist to apply analysis on this space as well as at the minimal K-types. We classify the spectral parameters of the cusp forms at the minimal K-types, and then, taking as input the spectral parameters and Fourier–Whittaker coefficients of such minimal cusp forms, generate the rest of the spectral expansion. The classification of the minimal K-types is given in Theorem 3, and the spectral expansion is Theorem 6.

The uniformity of Theorem 6 is necessary if one needs to expand, say, a Poincaré series which is not K-finite. Since certain cusp forms (and Eisenstein series) miss a number of K-types (i.e., generalized principal series forms), one would typically expect such an expansion to require evaluating the Wigner coefficients of the K-part of the Poincaré series, a task which is at best difficult for any large class of test functions. However, in Theorem 6, the sum is always taken over *all* K-types, relying on the fact that the Jacquet–Whittaker function for any given cusp form is simply zero on the types missed by that form. It is not too hard to see that the Archimedean weight function for a generalized principal series form in such an expansion, viewed as a sum over minimal-weight forms, is just the analytic continuation of the weight function for a full principal series form, and this can be evaluated without ever mentioning Wigner $\mathcal{D}$-matrices at all. We anticipate using such expansions to study smooth sums of exponential sums on $\mathrm{GL}(3)$.

On the other hand, the majority of the study of analytic number theory on automorphic forms takes place at the minimal K-types. With the longer history of automorphic forms on $\mathrm{GL}(2)$, it is perhaps easy to lose track of the fundamental importance of the connection between the three realizations of the Whittaker functions:

(1) The Whittaker functions of holomorphic and Maass cusp forms and Eisenstein series, which arise through their Fourier coefficients. These are the main number-theoretic objects which occur in L-functions (see [Goldfeld 2006, Section 6.5]), various period integral formulas [Jacquet et al. 1983; Watson 2002], etc.

(2) The classical Whittaker functions, which arise as the solutions to differential equations or through generalized hypergeometric series. These are the main analytic objects, and show up in a multitude of integral formulas (see the index entry for "Whittaker functions" in [Gradshteyn and Ryzhik 2015]), treatises on asymptotics [Olver 1975], etc.

(3) The Jacquet–Whittaker functions, which arise via an integral operator between representation spaces. These are the main algebraic objects, which have numerous functional identities (for GL(3), these are equations (3.4)–(3.8) of [Buttcane 2018], to name a few), and are the focus of much study in representation theory ([Shalika 1974] is fundamental there).

The bulk of this paper is concerned with making this connection very precise at the minimal K-types, from which the spectral expansion follows fairly easily via a double induction argument in Section 9.

We give expressions for the Whittaker functions at the minimal K-types in Theorem 5; these are due, in greatly different language, to Miyazaki [2010], who obtained them by studying their differential equations and then applying a certain induction argument, but we take this a step farther by evaluating the base case directly from the Jacquet integral, thus fixing the relationship between elements of the principal series representation and the Whittaker model, as well as the cusp forms, via the Fourier expansion. This is significant for the theory of special functions on GL(3) because the Mellin–Barnes integral is simpler analytically, but the Jacquet integral has a certain algebraic structure (i.e., the many identities of [Buttcane 2018, Section 3.1]) that we fundamentally rely on in constructions such as the Kuznetsov formulas (see [Buttcane 2016; 2017b; 2017c], and note the extensive use of the properties of the Jacquet–Whittaker functions there). Conversely, we must then rely on Miyazaki's solution of the differential equations (or representation theory) for the uniqueness property.

Such precise knowledge on the minimal-weight forms is, to the author, of primary use in developing spectral Kuznetsov/Petersson trace formulas and thereby Weyl laws and the many derived applications for studying such forms. These will appear in future papers on the topic, e.g., [Buttcane 2017b; 2017c].

A primary tenet of this paper, in combination with the previous part [Buttcane 2018], is to be self-contained. We generally succeed here except for three external pieces which are collected into Theorem 1: First, we do not prove the initial spectral decomposition; this is a standard proof using the techniques of Selberg, studying integral operators first carried out by [Gel'fand et al. 1969], and can be found in the first few pages of [Harish-Chandra 1968], or [Langlands 1976], or in a number of other books, e.g., [Osborne and Warner 1981]; see also [Iwaniec 1995, Appendix A]. Second, we quote the trivial bound for the principal series representations; this is a bound on Langlands parameters for the conjecturally nonexistent complementary series and can be found in [Vahutinskiĭ 1968]. Lastly, we do not prove multiplicity one for Whittaker functions; from the analytic perspective, we are avoiding the solution to a lengthy problem in partial differential equations, but it follows in the representation-theoretic language from [Shalika 1974]. The details of these connections may be found in the discussion following Theorem 1.

The proofs below will generally assume that the quotient is by $\Gamma = \mathrm{SL}(3, \mathbb{Z})$, but of course only the Fourier expansion will see the particular discrete group in use.

2. Some notation and background from Part I

We recall the notation of Part I of this series of papers. Throughout the current paper, section, equation and theorem numbers beginning with an I reference [Buttcane 2018], so for example Theorem I.1.1 references [Buttcane 2018, Theorem 1.1] and (I.2.3) references [Buttcane 2018, equation (2.3)].

Let $G = \mathrm{PSL}(3,\mathbb{R}) = \mathrm{GL}(3,\mathbb{R})/\mathbb{R}^\times$ and $\Gamma = \mathrm{SL}(3,\mathbb{Z})$. The Iwasawa decomposition of G is $G = U(\mathbb{R})Y^+K$ using the groups $K = \mathrm{SO}(3,\mathbb{R})$,

$$U(R) = \left\{ \begin{pmatrix} 1 & x_2 & x_3 \\ & 1 & x_1 \\ & & 1 \end{pmatrix} \,\middle|\, x_i \in R \right\}, \qquad R \in \{\mathbb{R}, \mathbb{Q}, \mathbb{Z}\},$$

$$Y^+ = \{\mathrm{diag}\{y_1y_2, y_1, 1\} \mid y_1, y_2 > 0\}.$$

The measure on the space $U(\mathbb{R})$ is simply $dx := dx_1\,dx_2\,dx_3$, and the measure on Y^+ is

$$dy := \frac{dy_1\,dy_2}{(y_1y_2)^3},$$

so that the measure on G is $dg := dx\,dy\,dk$, where dk is the Haar probability measure on K (see Section I.2.2.1). We generally identify elements of quotient spaces with their coset representatives, and in particular, we view $U(\mathbb{R})$, Y^+, K and Γ as subsets of G. (Note that these subspaces of $\mathrm{GL}(3,\mathbb{R})$ do inject into the quotient G, which is not generally the case for equivalent subspaces of $\mathrm{GL}(2n,\mathbb{R})$, since $-I \in \mathrm{SL}(2n,\mathbb{R})$.)

Characters of $U(\mathbb{R})$ are given by

$$\psi_m(x) = \psi_{m_1,m_2}(x) = e(m_1x_1 + m_2x_2), \qquad e(t) = e^{2\pi i t},$$

where $m \in \mathbb{R}^2$. Characters of Y^+ are given by the power function on 3×3 diagonal matrices, defined by

$$p_\mu(\mathrm{diag}\{a_1,a_2,a_3\}) = |a_1|^{\mu_1}|a_2|^{\mu_2}|a_3|^{\mu_3},$$

where $\mu \in \mathbb{C}^3$. We assume $\mu_1+\mu_2+\mu_3 = 0$ so this is defined modulo $\mathbb{R}^\times$, renormalize by $\rho = (1,0,-1)$, and extend by the Iwasawa decomposition

$$p_{\rho+\mu}(rxyk) = y_1^{1-\mu_3}y_2^{1+\mu_1}, \qquad r \in \mathbb{R}^\times,\ x \in U(\mathbb{R}),\ y \in Y^+,\ k \in K.$$

The Weyl group W of G contains the six matrices

$$I = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad w_2 = -\begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}, \quad w_3 = -\begin{pmatrix} 1 & & \\ & & 1 \\ & 1 & \end{pmatrix},$$

$$w_4 = \begin{pmatrix} & 1 & \\ & & 1 \\ 1 & & \end{pmatrix}, \quad w_5 = \begin{pmatrix} & & 1 \\ 1 & & \\ & 1 & \end{pmatrix}, \quad w_l = -\begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}.$$

The group of diagonal, orthogonal matrices $V \subset G$ contains the four matrices $v_{\varepsilon_1,\varepsilon_2} = \mathrm{diag}\{\varepsilon_1, \varepsilon_1\varepsilon_2, \varepsilon_2\}$, $\varepsilon \in \{\pm 1\}^2$, which we abbreviate $V = \{v_{++}, v_{+-}, v_{-+}, v_{--}\}$. The Weyl group induces an action on the coordinates of μ by $p_{\mu^w}(a) := p_\mu(waw^{-1})$, and we denote the coordinates of the permuted parameters by $\mu_i^w := (\mu^w)_i$, $i = 1, 2, 3$.

Part I made explicit the continuous and residual parts of the Langlands spectral expansion, and it remains to do this for $L^2_{\text{cusp}}(\Gamma\backslash G)$, the space of square-integrable functions $f : \Gamma\backslash G \to \mathbb{C}$ satisfying the cuspidality condition described below Theorem I.1.1, or equivalently, whose degenerate Fourier coefficients are all zero:

$$\int_{U(\mathbb{Z})\backslash U(\mathbb{R})} f(ug)\psi_n(u)\,du = 0 \quad \text{whenever } n_1 n_2 = 0,\ n \in \mathbb{Z}^2. \tag{1}$$

We will describe such functions in terms of their Fourier expansion (see Section I.3.6) and their decomposition over the Wigner $\mathcal{D}$-matrices.

If we describe elements $k = k(\alpha, \beta, \gamma) \in K$ in terms of the Z-Y-Z Euler angles

$$k(\alpha,\beta,\gamma) := k(\alpha,0,0)\,w_3\,k(-\beta,0,0)\,w_3\,k(\gamma,0,0), \quad k(\theta,0,0) := \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2}$$

then the Wigner $\mathcal{D}$-matrix $\mathcal{D}^d$ is the $(2d+1)$-dimensional representation of K primarily characterized by

$$\mathcal{D}^d(k(\theta,0,0)) = \mathcal{R}^d(e^{i\theta}), \qquad \mathcal{R}^d(s) := \operatorname{diag}\{s^d, \ldots, s^{-d}\}, \quad s \in \mathbb{C}. \tag{3}$$

The entries of the matrix-valued function $\mathcal{D}^d$ are indexed from the center:

$$\mathcal{D}^d = \begin{pmatrix} \mathcal{D}^d_{-d,-d} & \cdots & \mathcal{D}^d_{-d,d} \\ \vdots & \ddots & \vdots \\ \mathcal{D}^d_{d,-d} & \cdots & \mathcal{D}^d_{d,d} \end{pmatrix},$$

so in particular $\mathcal{D}^d_{m',m}(k(\theta,0,0)) = e^{-im'\theta}$ when $m' = m$ and zero otherwise, as the indices m' and m run through the integers $-d, \ldots, d$. Similarly, we index the rows $\mathcal{D}^d_{m'} = (\mathcal{D}^d_{m',-d}, \ldots, \mathcal{D}^d_{m',d})$ and columns $\mathcal{D}^d_{\cdot,m} = (\mathcal{D}^d_{-d,m}, \ldots, \mathcal{D}^d_{d,m})^T$ from the central entry as well. The entries, rows, and columns of the derived matrix- and vector-valued functions (e.g., the Whittaker function (7)) will be indexed similarly. The Wigner $\mathcal{D}$-matrices exhaust the equivalence classes of unitary, irreducible representations of the compact group K; hence they give a basis of $L^2(K)$, as in Section I.2.2.1, by the Peter–Weyl theorem.

The entries of the matrix $\mathcal{D}^d(k(0,\beta,0)) = \mathcal{D}^d(w_3)\mathcal{D}^d(k(-\beta,0,0))\mathcal{D}^d(w_3)$ are known as the Wigner d-polynomials. For the most part, we will avoid the Wigner d-polynomials by treating $\mathcal{D}^d(w_3)$ as a black box, that is, as some generic orthogonal matrix. The notable exceptions are in Section 4.8 and Proposition 15, and we frequently use the facts (see Section I.2.2.2)

$$\mathcal{D}^d(v_{\varepsilon,+1}) = \operatorname{diag}\{\varepsilon^d, \ldots, \varepsilon^{-d}\}, \quad \mathcal{D}^d_{m',m}(v_{\varepsilon,-1}) = (-1)^d \varepsilon^{m'} \delta_{m'=-m}. \tag{4}$$

The notation δ_P here is 1 if the predicate P is true, and 0 if it is false.

A complete list of the characters of V is given by $\chi_{\varepsilon_1,\varepsilon_2}$, $\varepsilon \in \{\pm 1\}^2$, which act on the generators by

$$\chi_{\varepsilon_1,\varepsilon_2}(v_{-+}) = \varepsilon_1, \quad \chi_{\varepsilon_1,\varepsilon_2}(v_{+-}) = \varepsilon_2. \tag{5}$$

These give rise to the projection operators

$$\Sigma_\chi^d = \frac{1}{4}\sum_{v\in V}\chi(v)\mathcal{D}^d(v), \tag{6}$$

which are written out explicitly in Section I.2.2.2 using the description (4). We sometimes use the abbreviation $\Sigma^d_{\varepsilon_1,\varepsilon_2} = \Sigma^d_{\chi_{\varepsilon_1,\varepsilon_2}}$.

Throughout the paper, we take the term "smooth", in reference to some function, to mean infinitely differentiable on the domain. The letters x, y, k, g, v and w will generally refer to elements of $U(\mathbb{R})$, Y^+, K, G, V, and W, respectively. The letters χ and ψ will generally refer to characters of V and $U(\mathbb{R})$, respectively, and μ will always refer to an element of $\mathbb{C}^3$ satisfying $\mu_1+\mu_2+\mu_3=0$. Vectors or matrices not directly associated with the Wigner $\mathcal{D}$-matrices, e.g., elements $n\in\mathbb{Z}^2$, are indexed in the traditional manner from the left-most entry or the top-left entry, respectively, e.g., $n=(n_1,n_2)$; when it becomes necessary to make it explicit, we will refer to this as "standard indexing".

There is a technical point in relation to differentiability on quotient spaces: we denote by $C^\infty(\Gamma\backslash G)$ the space of infinitely differentiable functions on $\Gamma\backslash G$, and one may define the differentiability on either of the smooth manifolds $\Gamma\backslash G$ or G itself (i.e., $C^\infty(\Gamma\backslash G)$ as the space of infinitely differentiable functions of G which are left-invariant by Γ) or conceivably by restricting to just the left-translation-invariant differential operators, i.e., those coming from the Lie algebra of G (see Section 4.1). In the present case, it turns out the space of functions does not depend on these choices, and this is a discussion best left to a text on smooth manifolds; see [Lee 2013, Theorem 9.16].

The majority of the paper will be concerned with the matrix-valued Jacquet–Whittaker function at each K-type $\mathcal{D}^d$:

$$W^d(g,\mu,\psi) := \int_{U(\mathbb{R})} I^d(w_l u g,\mu)\overline{\psi(u)}\,du, \quad I^d(xyk,\mu) := p_{\rho+\mu}(y)\mathcal{D}^d(k), \tag{7}$$

whose functional equations in μ (Proposition I.3.3),

$$W^d(g,\mu,\psi_{1,1}) = T^d(w,\mu)W^d(g,\mu^w,\psi_{1,1}), \quad w\in W, \tag{8}$$

are generated by the matrices

$$T^d(w_2,\mu) := \pi^{\mu_1-\mu_2}\Gamma^d_{\mathcal{W}}(\mu_2-\mu_1,+1), \tag{9}$$

$$T^d(w_3,\mu) := \pi^{\mu_2-\mu_3}\mathcal{D}^d(v_{--}w_l)\Gamma^d_{\mathcal{W}}(\mu_3-\mu_2,+1)\mathcal{D}^d(w_l v_{--}), \tag{10}$$

and $\Gamma^d_{\mathcal{W}}(u,\varepsilon)$ is a diagonal matrix coming from the functional equation of the classical Whittaker function (I.2.20): if $\mathcal{W}^d(y,u)$ is the diagonal matrix-valued function with entries (see Section I.2.3.1)

$$\begin{aligned}\mathcal{W}^d_{m,m}(y,u) &= \int_{-\infty}^{\infty}(1+x^2)^{-\frac12(1+u)}\left(\frac{1+ix}{\sqrt{1+x^2}}\right)^{-m} e(-yx)\,dx \\ &= \begin{cases} \dfrac{(\pi|y|)^{\frac12(1+u)}}{|y|\Gamma\left(\frac12(1-\varepsilon m+u)\right)} W_{-\frac12\varepsilon m,\frac12 u}(4\pi|y|) & \text{if } y\neq 0, \\ \dfrac{2^{1-u}\pi\,\Gamma(u)}{\Gamma\left(\frac12(1+u+m)\right)\Gamma\left(\frac12(1+u-m)\right)} & \text{if } y=0,\end{cases}\end{aligned} \tag{11}$$

(where $W_{\alpha,\beta}(y)$ is the classical Whittaker function), then for $y \neq 0$, we have the functional equations

$$\mathcal{W}^d(y,-u) = (\pi|y|)^{-u}\Gamma^d_{\mathcal{W}}(u,\operatorname{sgn}(y))\mathcal{W}^d(y,u), \quad \Gamma^d_{\mathcal{W},m,m}(u,\varepsilon) = \frac{\Gamma\big(\frac{1}{2}(1-\varepsilon m+u)\big)}{\Gamma\big(\frac{1}{2}(1-\varepsilon m-u)\big)}. \tag{12}$$

The function $W^d(g,\mu,\psi)$, as an integral of a Wigner $\mathcal{D}$-matrix, is again matrix-valued, and we index its rows $W^d_{m'}$, columns $W^d_{\cdot,m}$, and entries $W^d_{m',m}$ from the central entry, i.e., by the same convention as the Wigner $\mathcal{D}$-matrices. The matrices $T^d(w,\mu)$ are indexed by the central entry as well, and satisfy the composition

$$T^d(ww',\mu) = T^d(w,\mu)T^d(w',\mu^w). \tag{13}$$

3. The main results

Let $\mathcal{A}$ be the space of smooth, scalar-valued cusp forms, equipped with the usual L^2 inner product, that is, the functions $f \in L^2(\Gamma\backslash G)$ which are both infinitely differentiable and satisfy the cuspidality condition (1). We further impose on elements $f \in \mathcal{A}$ the technical requirement that

$$\text{there exists } r > 0 \text{ such that for all } X \in \mathfrak{g}_{\mathbb{C}} \text{ we have } \sup_{g\in G}|(Xf)(g)|\,\|g\|^{-r} < \infty. \tag{14}$$

Here $\mathfrak{g}_{\mathbb{C}}$ is the complexified Lie algebra of G (see Section 4.1), and $\|g\|^2 = \sum_{j,k}|g_{j,k}|^2$ is the Euclidean norm resulting from the natural inclusion $G \subset \mathbb{R}^9$.

Now let $\mathcal{A}^d$, $d \geq 0$, be the space of smooth, vector-valued cusp forms with K-type $\mathcal{D}^d$, that is, the functions $f : \Gamma\backslash G \to \mathbb{C}^{2d+1}$ taking values in the complex $(2d+1)$-dimensional row vectors, satisfying $f(gk) = f(g)\mathcal{D}^d(k)$ and having finite norm under the natural inner product

$$\langle f_1, f_2\rangle = \int_{\Gamma\backslash G} f_1(g)\overline{f_2(g)}^T\,dg = \int_{\Gamma\backslash G/K} f_1(z)\overline{f_2(z)}^T\,dz. \tag{15}$$

We again impose on elements of $\mathcal{A}^d$ the moderate growth condition (14).

In Section 4.3, we describe the decomposition of $\mathcal{A}$ into the spaces $\mathcal{A}^d$ by the projection operators on K. As discussed in the Introduction, we will use three external results about these cusp forms.

Theorem 1. (1) *The space of scalar-valued cusp forms decomposes into a direct sum of simultaneous eigenspaces $\mathcal{A}_\mu$ of the Casimir differential operators Δ_1 and Δ_2 (see Section 4.2). We parametrize the eigenspaces by the eigenvalues of the corresponding power function*

$$\Delta_i p_{\rho+\mu} = \lambda_i(\mu)p_{\rho+\mu}, \quad \lambda_1(\mu) = 1-\tfrac{1}{2}(\mu_1^2+\mu_2^2+\mu_3^2), \quad \lambda_2(\mu) = \mu_1\mu_2\mu_3;$$

i.e., functions $f \in \mathcal{A}_\mu$ satisfy $\Delta_i f = \lambda_i(\mu) f$. Again, the eigenspaces $\mathcal{A}_\mu$ further decompose into $\mathcal{A}^d_\mu$. The space $\mathcal{A}^d_\mu$ is finite-dimensional.

(2) *If $\mathcal{A}^d_\mu \neq \{0\}$ with $d \in \{0,1\}$ and μ of the form $(x+it, -x+it, -2it)$, $x,t \in \mathbb{R}$, then $|x| < \frac{1}{2}$.*

(3) *The n-th Fourier coefficient of $\phi \in \mathcal{A}_\mu$ lies in the image of the Jacquet integrals. That is, for any $\phi \in \mathcal{A}_\mu$ which lies in an irreducible component of the right-regular representation of G, there exists some $f \in C^\infty(K)$ and some coefficients $c_n \in \mathbb{C}$, $n \in \mathbb{Z}^2$, so that for any $n \in \mathbb{Z}^2$,*

$$\int_{U(\mathbb{Z})\backslash U(\mathbb{R})} \phi(ug)\overline{\psi_n(u)}\,du = c_n \int_{U(\mathbb{R})} I_{\mu,f}(w_l ug)\overline{\psi_n(u)}\,du,$$

with $I_{\mu,f}(xyk) = p_{\rho+\mu}(y)f(k)$. The integral on the right converges absolutely on $\mathrm{Re}(\mu_1) > \mathrm{Re}(\mu_2) > \mathrm{Re}(\mu_3)$, and must be interpreted by analytic continuation when one or more $\mathrm{Re}(\mu_i) = \mathrm{Re}(\mu_j)$, $i \neq j$.

Part (1) is Theorem 3 and Lemma 18 in [Harish-Chandra 1968]. Part (2) is the unitary dual estimate noted in [Miller and Schmid 2011, Appendix A.1]; of course, stronger estimates are known for $d = 0$, e.g., [Kim 2003, Appendix 2], and the $d = 1$ case follows by Rankin–Selberg theory (as [Jacquet and Shalika 1981] does for the unramified case), and this computation is given in [Buttcane 2017a]. Part (3) is a well-known reformulation of Shalika's multiplicity-one theorem [1974, Theorem 3.1]; it is the combination of several deep results of representation theory, and requires a representation-theoretic proof; we give the details in Appendix A. In particular, the meaning of the Jacquet integral on the right-hand side in the case that some $\mathrm{Re}(\mu_i) = \mathrm{Re}(\mu_j)$, $i \neq j$, is made explicit there. The coefficients c_n are called the Fourier–Whittaker coefficients, though in practice they are usually scaled by a factor $|n_1 n_2|$, as in (I.3.31), and the Whittaker function used is instead the completed Whittaker function at the minimal K-type, as in Theorem 5, below.

The assumption that ϕ lies in an irreducible component of the right-regular representation of G is not restrictive, and this, as well as the meaning of the statement, is explained in Appendix A.

3.1. *The minimal-weight forms.* In Section 5.2, we will show the Lie algebra of G gives rise to five differential-vector operators $Y^a : \mathcal{A}^d \to \mathcal{A}^{d+a}$, $a = -2, -1, 0, 1, 2$ (see (82) and (83)) which are left-translation-invariant. We describe the cusp forms by their behavior under the action of these operators. In Section 10.1, we will show:

Proposition 2. *Define the differential operator*

$$\Lambda_x = 27\Delta_2^2 + 4(\Delta_1 + x^2 - 1)(\Delta_1 + 4x^2 - 1)^2, \quad x \in \mathbb{R}. \tag{16}$$

Then for $\phi \in \mathcal{A}^d$, the following are equivalent:

(1) *ϕ is orthogonal to $Y^1\mathcal{A}^{d-1}$, $Y^2\mathcal{A}^{d-2}$ and $Y^0Y^2\mathcal{A}^{d-2}$.*

(2) *ϕ is a zero of Y^{-1}, Y^{-2} and an eigenfunction of Y^0.*

(3) *$d = 0$, $d = 1$ or ϕ is a zero of the operator $\Lambda_{\frac{1}{2}(d-1)}$.*

We say that such a ϕ is at its minimal K-type, or that ϕ is a minimal-weight form, and we define $\mathcal{A}^{d*} \subset \mathcal{A}^d$ to be the subspace of such forms. Further define $\mathcal{A}^{d*}_\mu$ to be the subspace of simultaneous eigenfunctions of Δ_1 and Δ_2 with eigenvalues matching $p_{\rho+\mu}$.

The value of a description as in part (3) is that it involves only the particular K-type we are interested in, without reference to $\mathcal{A}^{d-1}$ or $\mathcal{A}^{d-2}$. We initially take part (2) of the proposition as our working definition of minimal-weight forms, and the equivalence of the two remaining conditions comes at the very end of the paper in Section 10.1; in particular, after we have Theorem 3, below.

The appearance of Y^0 in the above proposition is strictly to accommodate an exceptional case that occurs at $d = 2$ and begins to trouble us in Section 8 (see the discussion following Proposition 14).

For vectors in $\mathbb{C}^{2d+1}$ associated in some manner with the Wigner $\mathcal{D}$-matrix (or the Jacquet–Whittaker function), we again index from the central entry. For $|j| \le d$, let $\mathfrak{v}^d_j$ be the $(2d+1)$-dimensional row vector with entries

$$\mathfrak{v}^d_{j,m'} = \delta_{m'=j}, \quad |m'| \le d, \tag{17}$$

and set

$$\mathfrak{u}^{d,\pm}_j = \tfrac{1}{2}(\mathfrak{v}^d_j \pm (-1)^d \mathfrak{v}^d_{-j}). \tag{18}$$

In Section 8.3, we prove:

Theorem 3. *Suppose $\phi \in \mathcal{A}^{d*}_\mu$. Then its n-th Fourier coefficient, $n \in \mathbb{Z}^2$, is a multiple of $f W^d(\cdot, \mu, \psi_n)$, where $f \in \mathbb{C}^{2d+1}$ and μ can be taken as one of the following*:

(1) *for $d = 0$, we use $f = 1$ and $\mathrm{Re}(\mu) = 0$ or $\mu = (x+it, -x+it, -2it)$, $|x| < \frac{1}{2}$,*

(2) *for $d = 1$, we use $f = \mathfrak{u}^{1,-}_0$ and $\mathrm{Re}(\mu) = 0$ or $\mu = (x+it, -x+it, -2it)$, $|x| < \frac{1}{2}$,*

(3) *for $d \ge 2$, we use $f = \mathfrak{u}^{d,+}_d$ and $\mu = \left(\frac{1}{2}(d-1)+it, -\frac{1}{2}(d-1)+it, -2it\right)$,*

with x, t real.

The (scalar) multiple of the theorem is again the Fourier–Whittaker coefficient. Notice that the components of $\phi = (\phi_{-d}, \dots, \phi_d) \in \mathcal{A}^{d*}_\mu$ are scalar-valued cusp forms $\phi_j \in \mathcal{A}_\mu$; then for ϕ_j, the element of $C^\infty(K)$ whose existence is assured by Theorem 1(3), call it $\tilde{f}$, is precisely $\tilde{f}(k) = f\mathcal{D}^d_{\cdot,j}(k)$, where f is the vector given in the above theorem.

We note that the symmetric square of a holomorphic modular form of even weight k will occur at minimal weight $d = 2k-1$ with μ as in Theorem 3 part (3) at $t = 0$; cf. [Buttcane and Miller 2017, Section 6.4].

The strong Selberg eigenvalue conjecture states that cuspidal representations occurring in the spectral expansion should be tempered; in the context of Theorem 3, this is precisely the statement that only the case $\mathrm{Re}(\mu) = 0$ occurs in parts (1) and (2). The argument of Section 10.1 also gives an equivalent statement of this conjecture:

Conjecture 4. *The null space of Λ_x in the cusp forms $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$ is trivial for $x \in \left(0, \frac{1}{2}\right)$.*

In fact, we expect Λ_x, $x \in \left(0, \frac{1}{2}\right)$, to be a positive operator on the cusp forms.

A by-product of the analysis of Section 8.3 is the explicit determination of all vectors $f \in \mathbb{C}^{2d+1}$ such that $f W^d(\cdot, \mu, \psi_n)$ is identically zero, when μ is in the standard form (107). In studying the action of the Lie algebra, i.e., the action of the Y^a operators, on the Whittaker functions of cusp forms, Theorem 1(3) allows us to pass from Whittaker functions defined as Fourier coefficients of cusp forms to Jacquet integrals of power functions (that is, elements of the principal series representation). In this way, and having studied the operation of the Y^a operators on the power functions, the determination of precisely which power functions are killed by the Jacquet integral becomes the main obstruction to studying the action of the Y^a operators on the Whittaker functions. Once we have a solid grasp on the action of the Lie algebra on the Whittaker functions, the return to studying cusp forms is more or less immediate from the Fourier expansion.

In Theorem 3, the vectors $f = \mathrm{u}^{d,\varepsilon}_{2j+\delta}$, with $\delta \in \{0, 1\}$ and $\varepsilon = \pm 1$, are the rows of the matrices Σ^d_χ as in (6) where $\chi = \chi_{(-1)^\delta,\varepsilon}$ as in (5). Specifically, for $\Phi \in \mathcal{A}^{d*}$, the character χ is

$$\begin{cases} \chi_{++} & \text{if } d = 0, \\ \chi_{+-} & \text{if } d = 1, \\ \chi_{(-1)^d,+} & \text{if } d \geq 2. \end{cases} \tag{19}$$

We have essentially forced a choice of the character χ in Theorem 3 to make the vector f as nice as possible, and this is done by applying the functional equations of the Jacquet–Whittaker function as in Proposition 15 (which permutes both the character and the coordinates of μ, see (I.3.27)); other allowable choices are χ_{--} or χ_{-+} when $d = 1$ and $\chi_{(-1)^d,-}$ or $\chi_{\pm,(-1)^d}$ for $d \geq 2$.

In Section 7, we compute the Mellin–Barnes integrals for the Whittaker functions of GL(3) automorphic forms at their minimal K-types. These have been computed by Miyazaki [2010] for $d \geq 2$, Manabe, Ishii and Oda [Manabe et al. 2004] for $d = 1$ and Bump [1984] for $d = 0$. In case $d \geq 1$, the bases used are somewhat different than ours, and for $d \geq 2$, the results here are somewhat stronger than Miyazaki's because we have solidified the connection between the Jacquet–Whittaker function and the Mellin–Barnes integral. (Theorem 5.9 of [Miyazaki 2010] contains the phrase "there is an element [of the Whittaker model]"; we have very precisely answered the question of which vector in the principal series representation gives rise to that element.)

For $\alpha \in \{0, 1\}^3$, $\beta, \eta \in \mathbb{Z}^3$, $\ell \in \mathbb{Z}^2$ and $s \in \mathbb{C}^2$ define

$$\Lambda_\alpha(\mu) = \pi^{-\frac{3}{2}+\mu_3-\mu_1}\Gamma\left(\tfrac{1}{2}(1+\alpha_1+\mu_1-\mu_2)\right)\Gamma\left(\tfrac{1}{2}(1+\alpha_2+\mu_1-\mu_3)\right)\Gamma\left(\tfrac{1}{2}(1+\alpha_3+\mu_2-\mu_3)\right), \tag{20}$$

$$\widetilde{G}(d, \beta, \eta, s, \mu) = \frac{\prod_{i=1}^3 \Gamma\left(\frac{1}{2}(\beta_i + s_1 - \mu_i)\right)\Gamma\left(\frac{1}{2}(\eta_i + s_2 + \mu_i)\right)}{\Gamma\left(\frac{1}{2}(s_1 + s_2 + \sum_i (\beta_i + \eta_i) - 2d)\right)}, \tag{21}$$

$$\widetilde{G}^0(\ell, s, \mu) = \widetilde{G}(0, 0, 0, s, \mu), \quad \widetilde{G}^1(\ell, s, \mu) = \widetilde{G}(1, (\ell_1, \ell_1, 1-\ell_1), (\ell_2, \ell_2, 1-\ell_2), s, \mu), \tag{22}$$

and for $d \geq 2$,

$$\Lambda^*(\mu) = (-1)^d \pi^{-\frac{3}{2}+\mu_3-\mu_1}\Gamma(d)\Gamma\left(\tfrac{1}{2}(1+\mu_1-\mu_3)\right)\Gamma\left(\tfrac{1}{2}(2+\mu_1-\mu_3)\right), \tag{23}$$

$$\widetilde{G}^d(\ell, s, \mu) = \widetilde{G}(d, (d, 0, \ell_1), (0, d, \ell_2), s, \mu). \tag{24}$$

Now for $|m'| \le d$, write $m' = \varepsilon m$ with $\varepsilon = \pm 1$ and $0 \le m \le d$, set

$$G^d_{m'}(s,\mu) = \sqrt{\binom{2d}{d+m}} \sum_{\ell=0}^{m} \varepsilon^\ell \binom{m}{\ell} \widetilde{G}^d((d-m,\ell), s, \mu), \tag{25}$$

and take $G^d(s,\mu)$ to be the vector with coordinates $G^d_{m'}(s,\mu)$, $m' = -d, \ldots, d$.

We define the completed minimal-weight Whittaker function at each weight d as

$$W^{d*}(y,\mu) = \frac{1}{4\pi^2} \int_{\mathrm{Re}(s)=\mathfrak{s}} (\pi y_1)^{1-s_1} (\pi y_2)^{1-s_2} G^d(s,\mu) \frac{ds}{(2\pi i)^2} \tag{26}$$

for any $\mathfrak{s} \in (\mathbb{R}^+)^2$.

Theorem 5. *The Whittaker functions at the minimal K-types are*

$$\begin{aligned}
\Lambda_{(0,0,0)}(\mu) W^0(y,\mu,\psi_{1,1}) &= W^{0*}(y,\mu),\\
\sqrt{2}\Lambda_{(0,1,1)}(\mu) \mathfrak{u}^{1,-}_0 W^1(y,\mu,\psi_{1,1}) &= W^{1*}(y,\mu),\\
-2\Lambda_{(1,0,1)}(\mu) \mathfrak{u}^{1,-}_1 W^1(y,\mu,\psi_{1,1}) &= W^{1*}(y,\mu^{w_4}),\\
2\Lambda_{(1,1,0)}(\mu) \mathfrak{u}^{1,+}_1 W^1(y,\mu,\psi_{1,1}) &= W^{1*}(y,\mu^{w_5}),
\end{aligned}$$

and for $\mu = \left(\tfrac{1}{2}(d-1)+it, -\tfrac{1}{2}(d-1)+it, -2it\right)$ with $d \ge 2$,

$$\Lambda^*(\mu) W^d_{-d}(y,\mu,\psi_{1,1}) = W^{d*}(y,\mu), \qquad W^d_{d,m} = 0.$$

As mentioned in the Introduction, this is the main vehicle for the analytic study of GL(3) automorphic forms. It has applications in the functional equations of GL(3) L-functions (see, e.g., Section 6.5, especially Lemma 6.5.21, of [Goldfeld 2006] or [Hirano et al. 2012]) and their Rankin–Selberg convolutions (see, e.g., [Stade 2002; Hirano et al. 2016; Buttcane 2017b; 2017c]), and is fundamental to the creation of Kuznetsov-type trace formulas (see [Buttcane 2016; 2017b; 2017c]), among other applications.

3.2. ***The full spectral expansion.*** If $\Phi \in \mathcal{A}^{d_0*}_\mu$, $d_0 \ge 0$, has Fourier coefficients

$$\rho_\Phi(n) f W^{d_0}(\,\cdot\,,\mu,\psi_n)$$

with the parameters μ and f given by Theorem 3 and χ is given by (19), then for all d (not just $d \ge d_0$), we may construct two matrix-valued forms

$$\Phi^d(g) = \sum_{\gamma \in (U(\mathbb{Z})V)\backslash \mathrm{SL}(2,\mathbb{Z})} \sum_{v \in V} \sum_{n \in \mathbb{N}^2} \rho_\Phi(n) \Sigma^d_\chi W^d(\gamma v g, \mu, \psi_n), \tag{27}$$

and $\widetilde{\Phi}^d(g)$ defined similarly but using the spectral parameters $-\bar{\mu}$ instead. Note that this is not the dual form. We will insist on a slightly unusual normalization (166) of Φ.

In Section 9, we look at the spaces of vector-valued forms generated by applying the Y^a operators to the minimal-weight cusp forms. More precisely, we operate at the level of the coefficient vectors. That is, if v lies in an appropriate subspace of $\mathbb{C}^{2d+1}$ (the row space of Σ^d_χ), then the entries of $p_{\rho+\mu}(y) v \mathcal{D}^d(k)$

lie in the principal series representation and the entries of $vW^d(g,\mu,\psi_n)$ lie in the Whittaker model, and we will show the entries of $v\Phi^d(g)$ (or more precisely $v\widetilde{\Phi}^d(g)$) are cusp forms. Since the Y^a operators function identically on all three objects by left-translation invariance, we are free to study their behavior on objects of the first type.

Now let $\mathcal{S}^d_{3,\mu}$ be a basis of $\mathcal{A}^{d*}_\mu$ for each $d \geq 0$, and take $\mathcal{S}^d_3$ to be the union of $\mathcal{S}^d_{3,\mu}$ for all μ and $\mathcal{S}_3 = \bigcup_d \mathcal{S}^d_3$. As described above, for each $\Phi \in \mathcal{S}_3$ and every d, we may construct the two matrix-valued forms Φ^d and $\widetilde{\Phi}^d$. It is the case that the function $\widetilde{\Phi}^d$ will be identically zero when d is less than the minimal weight of Φ, and in Section 10, we pull everything together into the following very uniform expansion of cusp forms:

Theorem 6. *For $f \in \mathcal{A}$, we have*

$$f(g) = \sum_{\Phi\in\mathcal{S}_3} \sum_{d=0}^{\infty} (2d+1)\operatorname{Tr}\left(\Phi^d(g)\int_{\Gamma\backslash G} f(g')\overline{\widetilde{\Phi}^d(g')}^T\,dg'\right).$$

The equivalent statement for a vector-valued form $f \in \mathcal{A}^d$ is

$$f(g) = \sum_{\Phi\in\mathcal{S}_3} \int_{\Gamma\backslash G} f(g')\overline{\widetilde{\Phi}^d(g')}^T\,dg'\,\Phi^d(g). \tag{28}$$

This completes the spectral expansion.

Each vector-valued cusp form of K-type $\mathcal{D}^d$ corresponds to $2d+1$ scalar-valued forms, and the multiplicities of a given $\Phi \in \mathcal{S}^{d_0*}_3$ in the vector-valued forms are essentially the rank of the matrices Σ^d_χ, with the middle $2d_0-1$ rows removed if $d_0 \geq 2$. By the types listed in Theorem 3, these multiplicities are

(1) $\frac{1}{2}d+1$ if d is even and $\frac{1}{2}(d-1)$ if d is odd,

(2) $\left\lfloor \frac{1}{2}(d+1)\right\rfloor$,

(3) $\left\lfloor \frac{1}{2}(d-d_0)\right\rfloor+1$ if $d \geq d_0$.

4. Background

4.1. ***The Lie algebras.*** The complexified Lie algebra of $G = \mathrm{PSL}(3,\mathbb{R})$ is $\mathfrak{g}_\mathbb{C} = \mathfrak{sl}(3)\otimes_\mathbb{R}\mathbb{C}$, the space of complex matrices of trace zero, and the complexified Lie algebra of $K = \mathrm{SO}(3,\mathbb{R})$ is $\mathfrak{k}_\mathbb{C}$, the space of antisymmetric matrices. These matrices act as differential operators on smooth functions of G by

$$(Xf)(g) := \frac{d}{dt} f(g\exp tX)\Big|_{t=0}, \qquad (X+iY)f = Xf+iYf, \tag{29}$$

when the entries of X and Y are real. We will not differentiate notationally between a matrix in the Lie algebra and the associated differential operator. Note that the commutator of the differential operators associated to two such matrices is then given by the differential operator associated to the commutator of the matrices; that is,

$$[X,Y] := X\circ Y - Y\circ X = XY-YX.$$

We take as our basis of $\mathfrak{k}_{\mathbb{C}}$ the three matrices

$$K_{\pm 1} = \begin{pmatrix} 0 & 0 & \mp 1 \\ 0 & 0 & -i \\ \pm 1 & i & 0 \end{pmatrix}, \quad K_0 = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{30}$$

In terms of the Z-Y-Z coordinates of Section I.2.2, the associated differential operators acting on a function of $k(\alpha, \beta, \gamma)$, are

$$K_{\pm 1} = e^{\mp i \gamma}(i \cot \beta \, \partial_\gamma \mp \partial_\beta - i \csc \beta \, \partial_\alpha), \quad K_0 = -\sqrt{2} \partial_\gamma. \tag{31}$$

On an entry of the Wigner-$\mathcal{D}$ matrix, the $K_{\pm 1}$ operators increase or decrease the column [Biedenharn and Louck 1981, Section 3.8],

$$K_{\pm 1} \mathcal{D}^d_{m',m} = \sqrt{d(d+1) - m(m \pm 1)} \mathcal{D}^d_{m',m \pm 1}, \quad |m'|, |m| \le d, \tag{32}$$

and K_0 does not,

$$K_0 \mathcal{D}^d_{m',m} = \sqrt{2} i m \mathcal{D}^d_{m',m}, \quad |m'|, |m| \le d. \tag{33}$$

The Laplacian on K is given by

$$2\Delta_K = K_1 \circ K_{-1} + K_{-1} \circ K_1 - K_0 \circ K_0, \tag{34}$$

and the Wigner-$\mathcal{D}$ matrix satisfies

$$\Delta_K \mathcal{D}^d = d(d+1) \mathcal{D}^d. \tag{35}$$

We extend to a basis for $\mathfrak{g}_{\mathbb{C}}$ by adding the five matrices

$$X_{\pm 2} = \begin{pmatrix} 1 & \pm i & 0 \\ \pm i & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_{\pm 1} = - \begin{pmatrix} 0 & 0 & \pm 1 \\ 0 & 0 & i \\ \pm 1 & i & 0 \end{pmatrix}, \quad X_0 = -\frac{\sqrt{6}}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{36}$$

Though the details of their construction are not used in this paper, these matrices are deliberately constructed so that

$$k(\alpha, 0, 0) X_j k(-\alpha, 0, 0) = e^{-ij\alpha} X_j, \quad |j| \le 2, \tag{37}$$

and they are orthogonal to $\mathfrak{k}_{\mathbb{C}}$ and have identical norm under the Killing form (up to a constant, the natural inner-product resulting from the inclusion $\mathfrak{g}_{\mathbb{C}} \subset \mathbb{C}^9$):

$$\sum_{i,j} [X_{k_1}]_{i,j} \overline{[X_{k_2}]_{i,j}} = 4\delta_{k_1 = k_2}, \quad \sum_{i,j} [X_{k_1}]_{i,j} \overline{[K_{k_2}]_{i,j}} = 0, \quad \sum_{i,j} [K_{k_1}]_{i,j} \overline{[K_{k_2}]_{i,j}} = 4\delta_{k_1 = k_2}, \tag{38}$$

where $[X_k]_{i,j}$ means the entry at index i, j of the matrix X_k as in (36), and similarly for $[K_k]_{i,j}$. Such a choice makes the description (74) relatively nice.

For use in Sections 5.1 and 5.2, we introduce the right K-invariant operators (which act on the left): For smooth functions of K, define

$$(K_j^{\mathrm{Left}} f)(k) = \frac{d}{dt} f(\exp(t K_j) k) \Big|_{t=0}.$$

Then we have [Biedenharn and Louck 1981, Section 3.8]

$$K_0^{\mathrm{Left}} = -\sqrt{2}\partial_\alpha, \quad K_{\pm 1}^{\mathrm{Left}} = e^{\pm i\alpha}(i\csc\beta\,\partial_\gamma - i\cot\beta\,\partial_\alpha \mp \partial_\beta). \tag{39}$$

These perform the symmetric operation to the K_j on entries of the Wigner-$\mathcal{D}$ matrix:

$$K_{\pm 1}^{\mathrm{Left}}\mathcal{D}^d_{m',m} = \sqrt{d(d+1)-m'(m'\mp 1)}\,\mathcal{D}^d_{m'\mp 1,m}, \quad K_0^{\mathrm{Left}}\mathcal{D}^d_{m',m} = \sqrt{2}im'\mathcal{D}^d_{m',m}, \quad |m'|, |m| \le d. \tag{40}$$

And the K Laplacian may also be written as

$$2\Delta_K = K_1^{\mathrm{Left}}\circ K_{-1}^{\mathrm{Left}} + K_{-1}^{\mathrm{Left}}\circ K_1^{\mathrm{Left}} - K_0^{\mathrm{Left}}\circ K_0^{\mathrm{Left}}. \tag{41}$$

We extend these operators to smooth functions on G by the expressions (39).

4.2. ***The Casimir operators.*** The (normalized) Casimir operators are defined by

$$\Delta_1 = -\frac{1}{2}\sum_{i,j} E_{i,j}\circ E_{j,i}, \quad \Delta_2 = \frac{1}{3}\sum_{i,j,k} E_{i,j}\circ E_{j,k}\circ E_{k,i} + \Delta_1, \tag{42}$$

where $E_{i,j}$, $i, j \in \{1,2,3\}$, is the element of the Lie algebra of $\mathrm{GL}(3,\mathbb{R})$ whose matrix has a 1 at position i, j and zeros elsewhere, using the standard indexing. In the Lie algebra, these are sometimes called the Capelli elements, and we require their expressions as differential operators. For operators of this complexity, this is typically done on a computer, but in Appendix B, we give sufficient details that the computation could (but likely shouldn't) be carried out by hand.

On functions of G/K these become [Goldfeld 2006, equations (6.1.1)]

$$\Delta_1^\circ = -y_1^2\partial_{y_1}^2 - y_2^2\partial_{y_2}^2 + y_1y_2\partial_{y_1}\partial_{y_2} - y_1^2(x_2^2+y_2^2)\partial_{x_3}^2 - y_1^2\partial_{x_1}^2 - y_2^2\partial_{x_2}^2 - 2y_1^2x_2\partial_{x_1}\partial_{x_3}, \tag{43}$$

$$\begin{aligned}\Delta_2^\circ = {}& -y_1^2y_2\partial_{y_1}^2\partial_{y_2} + y_1y_2^2\partial_{y_1}\partial_{y_2}^2 - y_1^3y_2\partial_{x_3}^2\partial_{y_1} + y_1y_2^2\partial_{x_2}^2\partial_{y_1} - 2y_1^2y_2x_2\partial_{x_1}\partial_{x_3}\partial_{y_2}\\ & + (-x_2^2+y_2^2)y_1^2y_2\partial_{x_3}^2\partial_{y_2} - y_1^2y_2\partial_{x_1}^2\partial_{y_2} + 2y_1^2y_2^2\partial_{x_1}\partial_{x_2}\partial_{x_3} + 2y_1^2y_2^2x_2\partial_{x_2}\partial_{x_3}^2\\ & + y_1^2\partial_{y_1}^2 - y_2^2\partial_{y_2}^2 + 2y_1^2x_2\partial_{x_1}\partial_{x_3} + (x_2^2+y_2^2)y_1^2\partial_{x_3}^2 + y_1^2\partial_{x_1}^2 - y_2^2\partial_{x_2}^2.\end{aligned} \tag{44}$$

We will require expressions for the full operators, i.e., those acting on functions of G.

We define the operators

$$\begin{aligned} Z_{\pm 2} &= (2y_2\partial_{y_2} - y_1\partial_{y_1}) \pm 2iy_2\partial_{x_2},\\ Z_{\pm 1} &= -2iy_1(\partial_{x_1} + x_2\partial_{x_3}) \mp 2y_1y_2\partial_{x_3},\\ Z_0 &= -\sqrt{6}y_1\partial_{y_1};\end{aligned} \tag{45}$$

then the full Casimir operators on G are given by the following lemma.

Lemma 7. *Acting on functions of the Iwasawa coordinates $x_1, x_2, x_3, y_1, y_2, \alpha, \beta, \gamma$, the Casimir operators have the form*

$$8\Delta_1 = 8\Delta_1^\circ - 2K_1^{\mathrm{Left}}\circ Z_{-1} - \sqrt{2}iK_0^{\mathrm{Left}}\circ(Z_2 - Z_{-2}) - 2K_{-1}^{\mathrm{Left}}\circ Z_1, \tag{46}$$

$$96\Delta_2 = 96\Delta_2^\circ + 2K_1^{\mathrm{Left}}\circ\mathcal{T}_1 + \sqrt{2}iK_0^{\mathrm{Left}}\circ\mathcal{T}_0 + 2K_{-1}^{\mathrm{Left}}\circ\mathcal{T}_{-1} + 6\sqrt{2}i(K_1^{\mathrm{Left}} + K_{-1}^{\mathrm{Left}})\circ K_0\circ(Z_{-1} - Z_1), \tag{47}$$

where

$$\mathcal{T}_{\pm1} = \sqrt{6} Z_{\mp1} \circ Z_0 + 6(1 - Z_{\mp2}) \circ Z_{\pm1},$$
$$\mathcal{T}_0 = 3(Z_1 \circ Z_1 - Z_{-1} \circ Z_{-1}) + 2(\sqrt{6} Z_0 + 6) \circ (Z_{-2} - Z_2).$$

Given a computer algebra package (and a sufficiently fast computer), one can compute these expressions by computing the differential operator for each $E_{i,j}$ acting on

$$f = f(x_1, x_2, x_3, y_1, y_2, \alpha, \beta, \gamma)$$

from the Iwasawa decomposition (as in Section I.2.4) for

$$xyk \exp(tE_{i,j}) \approx xyk(I + tE_{i,j}) \quad (t \text{ small})$$

(one can simplify this process using (76) or (168)), then forming $\Delta_i f$ directly from (42) and having the computer algebra package collect terms according to the derivatives of f. We give a somewhat more explicit proof in Appendix B.

4.3. ***The projection operators.*** Let $\widetilde{C}(G, \mathcal{D}^d)$ be the space of smooth functions on G that lie in the span of the entries of $\mathcal{D}^d$. That is $f \in \widetilde{C}(G, \mathcal{D}^d)$ if and only if

$$f(xyk) = \sum_{|m'|,|m| \le d} f_{m',m}(xy) \mathcal{D}^d_{m',m}(k)$$

for some collection of functions $f_{m',m} : G \to \mathbb{C}$. For $f \in \widetilde{C}(G, \mathcal{D}^d)$, taking the expansion (I.2.10) of $f_g(k) := f(gk)$ into Wigner $\mathcal{D}$-functions and evaluating at the identity $k = I$ gives

$$f(g) = (2d+1) \sum_{|m| \le d} \int_K f(gk') \overline{\mathcal{D}^d_{m,m}(k')}\, dk' = \sum_{|m| \le d} (\widetilde{\mathcal{P}}^d_m f)(g),$$

where

$$(\widetilde{\mathcal{P}}^d_m f)(g) = (2d+1) \int_K f(gk) \overline{\mathcal{D}^d_{m,m}(k)}\, dk \tag{48}$$

is the projection onto the span of the m-th column of $\mathcal{D}^d$.

The projection operators may also be written as

$$\widetilde{\mathcal{P}}^d_m f(xyk) = (2d+1) \sum_{|m'| \le d} \int_K f(xyk') \overline{\mathcal{D}^d_{m',m}(k')}\, dk'\, \mathcal{D}^d_{m',m}(k).$$

Also let

$$\widetilde{\mathcal{P}}^d = \sum_{|m| \le d} \widetilde{\mathcal{P}}^d_m.$$

Note that the projection operators (48) commute with the differential operators Δ_1 and Δ_2 by translation invariance, but the projection onto the span of a single Wigner function $\mathcal{D}^d_{m',m}$ might not.

Now if $f \in \widetilde{C}(G, \mathcal{D}^d)$, we may write f in terms of row-vector-valued functions which transform by $\mathcal{D}^d$,

$$f(xyk) = \sum_{|\ell| \le d} f_{\ell,\ell}(xyk), \quad f_\ell(xyk) := (2d+1) \int_K f(xykk') \mathcal{D}^d_\ell(k'^{-1})\, dk' = f_\ell(xy) \mathcal{D}^d(k)$$

(here $f_{\ell,\ell}$ is the entry at index ℓ of the vector-valued function f_ℓ), so we may move to considering vector-valued functions; say $C(G,\mathcal{D}^d)$ is the space of smooth vector-valued functions $f(xyk) = f(xy)\mathcal{D}^d(k)$. Define the vector projection operators $\mathcal{P}^d_m : \widetilde{C}(G,\mathcal{D}^d) \to C(G,\mathcal{D}^d)$ by

$$(\mathcal{P}^d_m f)(g) = (2d+1)\int_K f(gk')\mathcal{D}^d_m(k'^{-1})\,dk'.$$

Note that we have an isomorphism $\widetilde{C}(G,\mathcal{D}^d) \cong C(G,\mathcal{D}^d)^{2d+1}$ as each scalar-valued function gives rise to one vector-valued function for each row of $\mathcal{D}^d$. These functions are inherently linear combinations of the rows of $\mathcal{D}^d(k)$,

$$(f_{-d}(xyk),\ldots,f_d(xyk)) = \sum_{|m'|\le d} f_{m'}(xy)\mathcal{D}^d_{m'}(k),$$

and the entries may be written as

$$f_m(xyk) = \sum_{|m'|\le d} f_{m'}(xy)\mathcal{D}^d_{m',m}(k).$$

The raising and lowering operators on K then give

$$K_{\pm 1} f_m(xyk) = \sqrt{d(d+1)-m(m\pm 1)}\, f_{m\pm 1}(xyk), \tag{49}$$

using $f_m = 0$ when $|m| > d$. The operators $\widetilde{\mathcal{P}}^d_\ell$ and $K_{\pm 1}$ have simple commutation relations,

$$\widetilde{\mathcal{P}}^d_\ell K_{\pm 1} = K_{\pm 1}\widetilde{\mathcal{P}}^d_{\ell\mp 1}, \tag{50}$$

and this gives

$$\mathcal{P}^d_\ell K_{\pm 1} = \sqrt{d(d+1)-\ell(\ell\pm 1)}\,\mathcal{P}^d_{\ell\mp 1}. \tag{51}$$

4.4. ***Clebsch–Gordan coefficients.*** It is a fact from the theory of the Wigner-$\mathcal{D}$ matrices that any product $\mathcal{D}^a_{j,\ell}\mathcal{D}^d_{m',m}$ lies in the span of $\{\mathcal{D}^{d+b}_{m'+j,m+\ell} \mid |b| \le a\}$. More precisely, we have [Biedenharn and Louck 1981, equation (3.189)],

$$\mathcal{D}^k_{j,i}\mathcal{D}^d_{m',m} = \sum_{|a|\le k} \mathfrak{A}^{d,k,a}_{m',j}\mathfrak{A}^{d,k,a}_{m,i}\mathcal{D}^{d+a}_{m'+j,m+i}, \tag{52}$$

using the Clebsch–Gordan coefficients

$$\mathfrak{A}^{d,k,a}_{m,i} = \langle k\, i\, d\, m \mid (d+a)\,(i+m)\rangle,$$

(we prefer the shorthand $\mathfrak{A}^{d,k,a}_{m,i}$ over the standard notation $\langle\cdots\mid\cdots\rangle$ in the interest of compactness) or the Wigner three-j symbols via

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} = \frac{(-1)^{j_1-j_2+m_3}}{\sqrt{2j_3+1}}\langle j_1\, m_1\, j_2\, m_2 \mid j_3\, m_3\rangle. \tag{53}$$

The coefficients are defined to be zero unless

$$|m| \le d, \quad |i| \le k, \quad |m+i| \le d+a, \quad |d-k| \le d+a \le d+k. \tag{54}$$

One particular symmetry relation is [DLMF, equation 34.3.10]

$$\mathfrak{A}^{d,k,a}_{m,i} = (-1)^{k-a}\mathfrak{A}^{d,k,a}_{-m,-i}. \tag{55}$$

Because we only use $k = 2$ and $k = 1$, these may be found in [Abramowitz and Stegun 1964, Tables 27.9.2, 27.9.4], and we list the relevant values here: set

$$\mathfrak{C}^{d,k,a}_{m,i} = \frac{\sqrt{(2d+2a+1)\,(2d+a-k)!}}{\sqrt{(k-i)!\,(k+i)!\,(2d+a+k+1)!}} \times \begin{cases} \dfrac{(-1)^i\sqrt{(d-m)!\,(d+m)!}}{\sqrt{(d+a-i-m)!\,(d+a+i+m)!}} & \text{if } a \le 0, \\[2ex] \dfrac{\sqrt{(d+a-i-m)!\,(d+a+i+m)!}}{\sqrt{(d-m)!\,(d+m)!}} & \text{otherwise;} \end{cases}$$

then

$$\mathfrak{A}^{d,2,-2}_{m,i} = 2\sqrt{6}\mathfrak{C}^{d,2,-2}_{m,i}, \qquad \mathfrak{A}^{d,2,2}_{m,i} = 2\sqrt{6}\mathfrak{C}^{d,2,2}_{m,i}, \tag{56}$$

$$\mathfrak{A}^{d,2,-1}_{m,i} = 2\sqrt{6}(i(d+1)+2m)\mathfrak{C}^{d,2,-1}_{m,i}, \quad \mathfrak{A}^{d,2,1}_{m,i} = 2\sqrt{6}(di-2m)\mathfrak{C}^{d,2,1}_{m,i}, \tag{57}$$

$$\mathfrak{A}^{d,2,0}_{m,i} = 2(2d^2(i^2-1)+d(5i^2+6im-2)+3(i+m)(i+2m))\mathfrak{C}^{d,2,0}_{m,i}, \tag{58}$$

$$\mathfrak{A}^{d,1,-1}_{m,i} = -\sqrt{2}\mathfrak{C}^{d,1,-1}_{m,i}, \quad \mathfrak{A}^{d,1,0}_{m,i} = -2(i(d+1)+m)\mathfrak{C}^{d,1,1}_{m,i}, \quad \mathfrak{A}^{d,1,1}_{m,i} = \sqrt{2}\mathfrak{C}^{d,1,1}_{m,i}. \tag{59}$$

4.5. *The commutation relations.* We record the commutation relations amongst the differential operators. This is most easily expressed in terms of the Clebsch–Gordan coefficients

$$[X_j, X_k] = 2\sqrt{5}i^{1-j-k}\mathfrak{A}^{2,2,-1}_{j,k}K_{j+k}, \quad [X_j, K_k] = \sqrt{12}i^{1+k}\mathfrak{A}^{2,1,0}_{j,k}X_{j+k}, \tag{60}$$

$$[K_j, K_k] = 2i\mathfrak{A}^{1,1,0}_{j,k}K_{j+k}, \tag{61}$$

and the Clebsch–Gordan coefficients, as matrices $\mathfrak{A}^{d,k,a}$, are

$$\mathfrak{A}^{2,2,-1} = \frac{1}{10}\begin{pmatrix} 0 & 0 & 0 & 2\sqrt{5} & 2\sqrt{10} \\ 0 & 0 & -\sqrt{30} & -\sqrt{10} & 2\sqrt{5} \\ 0 & \sqrt{30} & 0 & -\sqrt{30} & 0 \\ -2\sqrt{5} & \sqrt{10} & \sqrt{30} & 0 & 0 \\ -2\sqrt{10} & -2\sqrt{5} & 0 & 0 & 0 \end{pmatrix}, \quad \mathfrak{A}^{2,1,0} = \frac{1}{6}\begin{pmatrix} 0 & 2\sqrt{6} & 2\sqrt{3} \\ -2\sqrt{3} & \sqrt{6} & 3\sqrt{2} \\ -3\sqrt{2} & 0 & 3\sqrt{2} \\ -3\sqrt{2} & -\sqrt{6} & 2\sqrt{3} \\ -2\sqrt{3} & -2\sqrt{6} & 0 \end{pmatrix},$$

$$\mathfrak{A}^{1,1,0} = \frac{1}{2}\begin{pmatrix} 0 & \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 0 & \sqrt{2} \\ -\sqrt{2} & -\sqrt{2} & 0 \end{pmatrix}.$$

4.6. *Barnes integrals.* We will make use of Barnes' second lemma [1910, Section 6]: for $a, b, c, d, e \in \mathbb{C}$,

$$\int_{-i\infty}^{i\infty} \frac{\Gamma(a+s)\Gamma(b+s)\Gamma(c+s)\Gamma(d-s)\Gamma(e-s)}{\Gamma(f+s)}\frac{ds}{2\pi i} = \frac{\Gamma(a+e)\Gamma(b+e)\Gamma(c+e)\Gamma(a+d)\Gamma(b+d)\Gamma(c+d)}{\Gamma(f-a)\Gamma(f-b)\Gamma(f-c)}, \tag{62}$$

where $f = a+b+c+d+e$.

4.7. ***The Weyl group and the V group.*** In Section 8, we will make extensive use of the Weyl and V group elements and their images under the Wigner $\mathcal{D}$-matrices directly, and in preparation, we give a few identities that will smooth out the analysis there. First, we note that the Weyl group is generated by the transpositions w_2 and w_3:

$$w_4 = w_3 w_2, \quad w_5 = w_2 w_3, \quad w_l = w_2 w_3 w_2 = w_3 w_2 w_3. \tag{63}$$

Next, we investigate the commutation relations between W and V:

$$\begin{gathered} w_2 v_{\varepsilon_1,\varepsilon_2} w_2 = v_{\varepsilon_1\varepsilon_2,\varepsilon_2}, \quad w_3 v_{\varepsilon_1,\varepsilon_2} w_3 = v_{\varepsilon_1,\varepsilon_1\varepsilon_2}, \quad w_4 v_{\varepsilon_1,\varepsilon_2} w_5 = v_{\varepsilon_1\varepsilon_2,\varepsilon_1}, \\ w_5 v_{\varepsilon_1,\varepsilon_2} w_4 = v_{\varepsilon_2,\varepsilon_1\varepsilon_2}, \quad w_l v_{\varepsilon_1,\varepsilon_2} w_l = v_{\varepsilon_2,\varepsilon_1}. \end{gathered} \tag{64}$$

Now we need to see how the matrices $\mathcal{D}^d(w_2)$ and $\mathcal{D}^d(v)$, $v \in V$, interact with the rows of the Σ_χ^d matrices, which are the vectors $\mathrm{u}_m^{d,\pm}$ as in (18). From (4), we can see that

$$\mathrm{u}_m^{d,+} \mathcal{D}^d(v_{\varepsilon_1,\varepsilon_2}) = \varepsilon_1^j \mathrm{u}_m^{d,+}, \quad \mathrm{u}_m^{d,-} \mathcal{D}^d(v_{\varepsilon_1,\varepsilon_2}) = \varepsilon_1^j \varepsilon_2 \mathrm{u}_m^{d,-}, \tag{65}$$

or equivalently,

$$\mathrm{u}_m^{d,\varepsilon} \mathcal{D}^d(v_{-+}) = (-1)^j \mathrm{u}_m^{d,\varepsilon}, \quad \mathrm{u}_m^{d,\varepsilon} \mathcal{D}^d(v_{+-}) = \varepsilon \mathrm{u}_m^{d,\varepsilon}. \tag{66}$$

From Sections I.2.2.2 and I.2.2.3, we can see that

$$\mathcal{D}^d(w_2) = \mathcal{D}^d(v_{+-}) \mathcal{R}^d(i) = \mathcal{R}^d(i) \mathcal{D}^d(v_{--}), \tag{67}$$

and it follows that if $\varepsilon = (-1)^m$, then

$$\mathrm{u}_m^{d,\pm} \mathcal{R}^d(i) = i^{-m} \mathrm{u}_m^{d,\pm\varepsilon}, \quad \mathrm{u}_m^{d,\pm} \mathcal{D}^d(w_2) = \mathrm{u}_m^{d,\pm} \mathcal{R}^d(i) \mathcal{D}^d(v_{--}) = \pm i^{-m} \mathrm{u}_m^{d,\pm\varepsilon}. \tag{68}$$

4.8. ***Wigner*** d***-polynomials and Jacobi P-polynomials.*** The Wigner d-polynomials

$$\mathrm{d}_{m',m}^d(\cos\beta) = \mathcal{D}_{m',m}^d(k(0,\beta,0)) \tag{69}$$

may be given in terms of Jacobi polynomials [Biedenharn and Louck 1981, equation (3.72)]

$$\mathrm{d}_{m',m}^d(x) = 2^{-m} \sqrt{\frac{(d+m)!\,(d-m)!}{(d+m')!\,(d-m')!}} (1-x)^{\frac{1}{2}(m-m')} (1+x)^{\frac{1}{2}(m+m')} P_{d-m}^{(m-m',m+m')}(x), \tag{70}$$

and they satisfy the symmetries [Biedenharn and Louck 1981, equations (3.80)–(3.82)]

$$\mathrm{d}_{m',m}^d(x) = (-1)^{m'+m} \mathrm{d}_{-m',-m}^d(x) = (-1)^{m'+m} \mathrm{d}_{m,m'}^d(x). \tag{71}$$

We will need the first two Jacobi polynomials; they are [DLMF, equation 18.5.7]

$$P_0^{(\alpha,\beta)}(x) = 1, \quad P_1^{(\alpha,\beta)}(x) = \tfrac{1}{2}(\alpha - \beta + x(2+\alpha+\beta)). \tag{72}$$

5. The action of the X_j operators

5.1. ***The X_j as differential operators.*** Starting from (45), we define the differential operators

$$\widetilde{Z}_{\pm 2} = Z_{\pm 2} \mp i\tfrac{\sqrt{2}}{2} K_0^{\mathrm{Left}}, \quad \widetilde{Z}_{\pm 1} = Z_{\pm 1} - K_{\pm 1}^{\mathrm{Left}}, \quad \widetilde{Z}_0 = Z_0. \tag{73}$$

Lemma 8. *As differential operators in the Iwasawa U-Y-K, i.e., $g = xyk$, coordinates, the X_j operators have the form*

$$X_j = \sum_{\ell=-2}^{2} \mathcal{D}^2_{\ell,j}(k) \widetilde{Z}_\ell. \tag{74}$$

We have expressed, as we may, all of the trigonometric polynomials in α, β, γ in terms of the Wigner $\mathcal{D}$-matrices in preparation for application of the Clebsch–Gordan multiplication rules; see Section 4.4.

Proof. We have the relation (37), and it can be computed directly using

$$\mathcal{D}^2(w_3) = \frac{1}{4}\begin{pmatrix} 1 & 2i & -\sqrt{6} & -2i & 1 \\ -2i & 2 & 0 & 2 & 2i \\ -\sqrt{6} & 0 & -2 & 0 & -\sqrt{6} \\ 2i & 2 & 0 & 2 & -2i \\ 1 & -2i & -\sqrt{6} & 2i & 1 \end{pmatrix}$$

that

$$w_3 X_j w_3 = \sum_{|\ell| \le 2} \mathcal{D}^2_{\ell,j}(w_3) X_\ell,$$

so from (2) and (3) it follows that for $k \in K$,

$$k X_j k^{-1} = \sum_{|\ell| \le 2} \mathcal{D}^2_{\ell,j}(k) X_\ell. \tag{75}$$

This applies to the simple trick

$$(X_j f)(xyk) = \frac{d}{dt} f(xy \exp(t k X_j k^{-1}) k)\Big|_{t=0} \tag{76}$$

for a smooth function f on G. We switch to the Iwasawa-friendly basis

$$N_1 = E_{2,3}, \quad N_2 = E_{1,2}, \quad N_3 = E_{1,3}, \quad 3A_1 = E_{1,1} + E_{2,2} - 2E_{3,3}, \quad 3A_2 = 2E_{1,1} - E_{2,2} - E_{3,3}.$$

The conversion is

$$X_{\pm 2} = \pm 2i N_2 - A_1 + \sqrt{2} A_2 \mp i K_0, \quad X_{\pm 1} = -2i N_1 \mp 2N_3 - K_{\pm 1}, \quad X_0 = -\sqrt{6} A_1, \tag{77}$$

and we can compute directly from the definition (29) that

$$N_1 = y_1 \partial_{x_1} + y_1 x_2 \partial_{x_3}, \quad N_2 = y_2 \partial_{x_2}, \quad N_3 = y_1 y_2 \partial_{x_3}, \quad A_1 = y_1 \partial_{y_1}, \quad A_2 = y_2 \partial_{y_2} \tag{78}$$

as differential operators on functions of $U(\mathbb{R})Y^+$ alone. For example, if we write $f(xy) = f(x_1, x_2, x_3, y_1, y_2)$ for a smooth function on $U(\mathbb{R})Y^+$, the action of N_1 is computed by

$$(N_1 f)(xy) := \frac{d}{dt} f(xy \exp(t N_1))\Big|_{t=0} = \frac{d}{dt} f(x_1 + t y_1, x_2, x_3 + t y_1 x_2, y_1, y_2)\Big|_{t=0}.$$

The lemma follows from applying (75), (77) and (78) in (76). Note that the N_j and A_j operators are still acting on the right of the $U(\mathbb{R})Y^+$ part of the Iwasawa decomposition, and this produces the Z_j operators, but the K_j operators are acting now *on the left* of the K part. $\square$

5.2. ***The action of X_j on vector-valued functions.*** We wish to describe the effect of each operator X_j on functions $f \in C(G, \mathcal{D}^d)$. We may restrict our attention to $\widetilde{\mathcal{P}}^{d+a}_\ell X_j f$ with $|a| \le 2$, $|\ell| \le d+a$, by the explicit form of the X_j in (74) and the fact that a product $\mathcal{D}^a_{i,j}\mathcal{D}^d_{m',m}$ lies in the span of $\{\mathcal{D}^{d+b}_{m'+i,m+j} \mid |b| \le a\}$, as in (52). Further, the components of f may be obtained by applying $K_{\pm 1}$ repeatedly to the central entry f_0 using (49), and we have already described the commutation relations of these operators with $\widetilde{\mathcal{P}}^{d+a}$ and X_j in (51) and Section 4.5, so we need only consider $\widetilde{\mathcal{P}}^{d+a}_\ell X_j f_0$. Lastly, by (74) and (52) only the j-th column of $\mathcal{D}^{d+a}$ is involved in the operator $X_j f_0$, so it is sufficient to consider $\widetilde{\mathcal{P}}^{d+a}_j X_j f_0$ for $|a| \le 2$. This gives, in principle, 25 operators to consider, but up to applying $K_{\pm 1}$ to the result, we have only the following five operators: Define the vector

$$\mathfrak{B}^d = (\mathfrak{B}^d_{-2}, \dots, \mathfrak{B}^d_2) := \tfrac{\sqrt{6}}{3}(-2(d+1), -(d+3), -3, (d-2), 2d),$$

and for $|a|, |j| \le 2$, $|m'| \le d$, set

$$\widetilde{Y}^{d,a}_{m',j} = \mathfrak{A}^{d,2,a}_{m'-j,j} \times \begin{cases} Z_{-2} - m' - 2 & \text{if } j = -2, \\ Z_{-1} & \text{if } j = -1, \\ Z_0 - \mathfrak{B}^d_a & \text{if } j = 0, \\ Z_1 & \text{if } j = 1, \\ Z_2 + m' - 2 & \text{if } j = 2. \end{cases} \tag{79}$$

Then we define the operator on $\widetilde{Y}^{d,a}$ on smooth, scalar-valued functions $f \in C^\infty(G)$ by

$$(\widetilde{Y}^{d,a} f)(xyk) = \sum_{m'=-d}^{d} \mathcal{D}^{d+a}_{m',0}(k) \sum_{j=-2}^{2} \widetilde{Y}^{d,a}_{m',j} f_{m'-j}(xy), \tag{80}$$

where

$$\sum_{m'=-d}^{d} \mathcal{D}^d_{m',0}(k) f_{m'}(xy) = (\widetilde{\mathcal{P}}^d_0 f)(xyk). \tag{81}$$

Here, and throughout, we define $f_{m'} = 0$ for $|m'| > d$.

Proposition 9. *Suppose $f \in \widetilde{C}(G, \mathcal{D}^d)$. Then we have*

$$\begin{aligned}
\widetilde{\mathcal{P}}^{d+a}_0 X_0 f &= \mathfrak{A}^{d,2,a}_{0,0} \widetilde{Y}^{d,a} f, \\
\widetilde{\mathcal{P}}^{d+a}_{\pm 1} X_{\pm 1} f &= \frac{1}{\sqrt{(d+a)(d+a+1)}} \mathfrak{A}^{d,2,a}_{0,\pm 1} K_{\pm 1} \widetilde{Y}^{d,a} f, \\
\widetilde{\mathcal{P}}^{d+a}_{\pm 2} X_{\pm 2} f &= \frac{1}{\sqrt{(d+a)(d+a+1)}\sqrt{(d+a)(d+a+1)-2}} \mathfrak{A}^{d,2,a}_{0,\pm 2} K_{\pm 1} K_{\pm 1} \widetilde{Y}^{d,a} f.
\end{aligned}$$

Then we can realize $\widetilde{Y}^{d,a}$ as a composition of left-invariant projection and differential operators:

Corollary 10. *The operator* $\widetilde{Y}^{d,a} : C^\infty(G) \to C^\infty(G)$, *defined by* (79) *and* (80), *is given by*

$$\widetilde{Y}^{d,a} = \begin{cases} \dfrac{1}{\mathfrak{A}^{d,2,a}_{0,0}} \widetilde{\mathcal{P}}^{d+a}_0 X_0 \widetilde{\mathcal{P}}^d_0 & \text{if } a = -2, 0, 2, \\ \dfrac{1}{\mathfrak{A}^{d,2,a}_{0,1}\sqrt{(d+a)(d+a+1)}} K_{-1} \widetilde{\mathcal{P}}^{d+a}_1 X_1 \widetilde{\mathcal{P}}^d_0 & \text{if } a = \pm 1. \end{cases}$$

In particular, $\widetilde{Y}^{d,a}$ *is left-translation-invariant.*

Note that $\mathfrak{A}^{d,2,\pm1}_{0,0} = 0$ by (57), so we cannot use the top expression for $\widetilde{Y}^{d,a}$ when $a = \pm 1$.

For each d, we may extend this to an operator $Y^a : C^\infty(G, \mathbb{C}^{2d+1}) \to C^\infty(G, \mathbb{C}^{2d+2a+1})$ on smooth, vector-valued functions

$$C^\infty(G, \mathbb{C}^{2d+1}) := \{ f : G \to \mathbb{C}^{2d+1} \mid f \text{ smooth} \}$$

by applying $\widetilde{Y}^{d,a}$ to the central entry and projecting back to a vector-valued function:

$$Y^a f = \mathcal{P}^{d+a}_0 \widetilde{Y}^{d,a} f_0, \qquad f = (f_{-d}, \dots, f_d) \in C^\infty(G, \mathbb{C}^{2d+1}).$$

The components of $Y^a f \in C^\infty(G, \mathbb{C}^{2d+2a+1})$ are then given by

$$(Y^a f)_{m'}(xy) = \sum_{j=-2}^{2} \widetilde{Y}^{d,a}_{m',j} f_{m'-j}(xy), \tag{82}$$

and again we may realize this as a left-invariant operator by

$$Y^a f = \begin{cases} \dfrac{1}{\mathfrak{A}^{d,2,a}_{0,0}} \mathcal{P}^{d+a}_0 X_0 f_0 & \text{if } a = -2, 0, 2, \\ \dfrac{1}{\mathfrak{A}^{d,2,a}_{0,1}} \mathcal{P}^{d+a}_1 X_1 f_0 & \text{if } a = \pm 1. \end{cases} \tag{83}$$

The dimension d in the domain of the operator Y^a is to be understood from context; alternately, via the natural extension, one may think of Y^a as an operator on the union of vector-valued functions of all (odd) dimensions

$$Y^a : \bigcup_d C^\infty(G, \mathbb{C}^{2d+1}) \to \bigcup_d C^\infty(G, \mathbb{C}^{2d+1}), \tag{84}$$

which satisfies $Y^a C^\infty(G, \mathbb{C}^{2d+1}) \subset C^\infty(G, \mathbb{C}^{2d+2a+1})$ for any $d \geq 0$, $|a| \leq 2$, $d + a \geq 0$. In the course of the current paper, we have no need to place any algebraic structure on the space $\bigcup_d C^\infty(G, \mathbb{C}^{2d+1})$.

Proof of Proposition 9. Consider $\widetilde{\mathcal{P}}^{d+a}_j X_j f$, and assume for convenience $\mathfrak{A}^{d,2,a}_{0,j} \neq 0$, since otherwise the result is trivial. The function of the operators

$$\frac{1}{\sqrt{(d+a)(d+a+1)}} K_{\pm1} \quad \text{and} \quad \frac{1}{\sqrt{(d+a)(d+a+1)}\sqrt{(d+a)(d+a+1)-2}} K_{\pm1} K_{\pm1}$$

is precisely to shift the column of the Wigner $\mathcal{D}$-matrix on which $\widetilde{Y}^{d,a} f$ is supported from zero (by (80)) to ±1 and ±2, respectively.

We apply (40) and (52) to (74), giving

$$\frac{1}{\mathfrak{A}^{d,2,a}_{0,j}}\widetilde{\mathcal{P}}^{d+a}_j X_j f = \sum_{m'=-d}^{d}\sum_{\ell=-2}^{2}\mathfrak{A}^{d,2,a}_{m',\ell}\mathcal{D}^{d+a}_{m'+\ell,j}(k)Z_\ell f_{m'}(xy) + \sum_{m'=-d}^{d} f_{m'}(xy)\sum_{\pm}\pm m'\mathfrak{A}^{d,2,a}_{m',\pm 2}\mathcal{D}^{d+a}_{m'\pm 2,j}(k)$$
$$- \sum_{m'=-d}^{d} f_{m'}(xy)\sum_{\pm}\mathfrak{A}^{d,2,a}_{m'\mp 1,\pm 1}\sqrt{d(d+1)-m'(m'\mp 1)}\mathcal{D}^{d+a}_{m',j}(k).$$

Now we send $m'\mapsto m'-\ell$ in the first two sums (using $\ell=\pm 2$ in the second sum), and the result follows from

$$\sum_{\pm}\mathfrak{A}^{d,2,a}_{m'\mp 1,\pm 1}\sqrt{d(d+1)-m'(m'\mp 1)} = \mathfrak{A}^{d,2,a}_{m',0}\mathfrak{B}^d_a,$$

which can be verified directly from (56)-(59) or by the properties of the Wigner $3j$-symbol [DLMF, Section 34.3]. □

5.3. ***The action of Y^a on the power function.*** For each d, define $Y^a_\mu : \mathbb{C}^{2d+1}\to\mathbb{C}^{2d+2a+1}$ by

$$Y^a_\mu f = p_{-\rho-\mu}(y)Y^a p_{\rho+\mu}(y)f. \tag{85}$$

As with the definition (84), either the dimension of the domain of Y^a_μ is to be understood from context, or we may use the natural extension

$$Y^a_\mu : \bigcup_d \mathbb{C}^{2d+1}\to\bigcup_d \mathbb{C}^{2d+1}, \tag{86}$$

which satisfies $Y^a_\mu\mathbb{C}^{2d+1}\subset\mathbb{C}^{2d+2a+1}$, and again, we have no need to place an algebraic structure on $\bigcup_d \mathbb{C}^{2d+1}$.

The Z_j operators collectively act on the power function by eigenvalues as

$$(Z_{-2}p_{\rho+\mu},\ldots,Z_2 p_{\rho+\mu}) = (\mu_1-\mu_2+1, 0, \sqrt{6}(\mu_3-1), 0, \mu_1-\mu_2+1)p_{\rho+\mu},$$

and using $\mathfrak{v}^d_j$ and $\mathfrak{u}^{d,\pm}_j$ as in (17) and (18), we have

$$Y^a\mathfrak{v}^d_j p_{\rho+\mu} = \big(\mathfrak{v}^{d+a}_{j-2}\mathfrak{A}^{d,2,a}_{j,-2}(\mu_1-\mu_2+1-j) + \mathfrak{v}^{d+a}_j\mathfrak{A}^{d,2,a}_{j,0}(\sqrt{6}(\mu_3-1)-\mathfrak{B}^d_a)$$
$$+\mathfrak{v}^{d+a}_{j+2}\mathfrak{A}^{d,2,a}_{j,2}(\mu_1-\mu_2+1+j)\big)p_{\rho+\mu}, \tag{87}$$

and by the symmetry (55),

$$Y^a_\mu\mathfrak{u}^{d,\pm}_j = \mathfrak{A}^{d,2,a}_{j,-2}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d+a,\pm}_{j-2} + \mathfrak{A}^{d,2,a}_{j,0}(\sqrt{6}(\mu_3-1)-\mathfrak{B}^d_a)\mathfrak{u}^{d+a,\pm}_j$$
$$+\mathfrak{A}^{d,2,a}_{j,2}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d+a,\pm}_{j+2}. \tag{88}$$

The reason for the $(-1)^d$ in the definition of $\mathfrak{u}^{d,\pm}_j$ is to maintain consistency across parities of a in the above equation.

Because this is the key identity, we write it out explicitly for each a. We assume $|j| \le d$. At $a=0$, $Y^0_\mu \mathfrak{u}^{0,\pm}_j = 0$, and when $d>0$,

$$\begin{aligned} 0 &= \sqrt{6(d+2-j)(d+1-j)(d+j)(d-1+j)}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d,\pm}_{j-2} \\ &\quad -2\big(\sqrt{6}(d(d+1)-3j^2)\mu_3+\sqrt{d(d+1)(2d-1)(2d+3)}Y^0_\mu\big)\mathfrak{u}^{d,\pm}_j \\ &\quad +\sqrt{6(d+2+j)(d+1+j)(d-j)(d-1-j)}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d,\pm}_{j+2}. \end{aligned} \tag{89}$$

At $a=1$,

$$\begin{aligned} &\sqrt{2d(d+1)(d+2)(2d+1)}Y^1_\mu \mathfrak{u}^{d,\pm}_j \\ &\quad = -\sqrt{(d+1-j)(d+2-j)(d+3-j)(d+j)}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d+1,\pm}_{j-2} \\ &\qquad -2j\sqrt{(d+1-j)(d+1+j)}(3\mu_3-d-1)\mathfrak{u}^{d+1,\pm}_j \\ &\qquad +\sqrt{(d-j)(d+1+j)(d+2+j)(d+3+j)}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d+1,\pm}_{j+2}. \end{aligned} \tag{90}$$

At $a=-1$,

$$\begin{aligned} &\sqrt{2d(d-1)(d+1)(2d+1)}Y^{-1}_\mu \mathfrak{u}^{d,\pm}_j \\ &\quad = -\sqrt{(d+1-j)(d-2+j)(d-1+j)(d+j)}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d-1,\pm}_{j-2} \\ &\qquad +2j\sqrt{(d-j)(d+j)}(3\mu_3+d)\mathfrak{u}^{d-1,\pm}_j \\ &\qquad +\sqrt{(d-2-j)(d-1-j)(d-j)(d+1+j)}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d-1,\pm}_{j+2}. \end{aligned} \tag{91}$$

At $a=2$,

$$\begin{aligned} &2\sqrt{(d+1)(d+2)(2d+1)(2d+3)}Y^2_\mu \mathfrak{u}^{d,\pm}_j \\ &\quad = \sqrt{(d+1-j)(d+2-j)(d+3-j)(d+4-j)}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d+2,\pm}_{j-2} \\ &\qquad +2\sqrt{(d+1-j)(d+2-j)(d+1+j)(d+2+j)}(3\mu_3-2d-3)\mathfrak{u}^{d+2,\pm}_j \\ &\qquad +\sqrt{(d+1+j)(d+2+j)(d+3+j)(d+4+j)}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d+2,\pm}_{j+2}. \end{aligned} \tag{92}$$

At $a=-2$,

$$\begin{aligned} &2\sqrt{d(d-1)(2d-1)(2d+1)}Y^{-2}_\mu \mathfrak{u}^{d,\pm}_j \\ &\quad = \sqrt{(d-3+j)(d-2+j)(d-1+j)(d+j)}(\mu_1-\mu_2+1-j)\mathfrak{u}^{d-2,\pm}_{j-2} \\ &\qquad +2\sqrt{(d-1-j)(d-j)(d-1+j)(d+j)}(3\mu_3+2d-1)\mathfrak{u}^{d-2,\pm}_j \\ &\qquad +\sqrt{(d-3-j)(d-2-j)(d-1-j)(d-j)}(\mu_1-\mu_2+1+j)\mathfrak{u}^{d-2,\pm}_{j+2}. \end{aligned} \tag{93}$$

6. Interactions with the Lie algebra

From the discussion of the previous section and Sections 4.1 and 4.3, in studying the action of the Lie algebra $\mathfrak{g}_\mathbb{C}$ on $\mathcal{A}$, it is sufficient to study the action of the Y^a operators on $\mathcal{A}^d$. Further, by the Fourier expansion (I.3.30), it is sufficient to study the action of Y^a_μ on $\mathbb{C}^{2d+1}$, provided we know which vectors $f \in \mathbb{C}^{2d+1}$ are sent to zero by the Jacquet–Whittaker function, which we will carefully study in Section 8.2.

For the moment, we need to study the interaction of the Lie algebra, via the Y^a_μ operators, with several common operations.

6.1. ***The adjoint of Y^a.*** First, we compute the adjoint of the Y^a operators with respect to the inner product on $\mathcal{A}^d$. We use the composition of operators given in (83).

Suppose $f \in \widetilde{C}(\Gamma\backslash G, \mathcal{D}^d)$ and $h \in C(\Gamma\backslash G, \mathcal{D}^d)$ are square-integrable, then

$$\begin{aligned}\int_{\Gamma\backslash G} (\mathcal{P}^d_m f)(g)\, \overline{h(g)^T}\, dg &= (2d+1) \int_{\Gamma\backslash G} f(g) \int_K \mathcal{D}^d_m(k^{-1}) \mathcal{D}^d(k) dk\, \overline{h(g)^T}\, dg \\ &= (2d+1) \int_{\Gamma\backslash G} f(g)\, \overline{h_m(g)}\, dg,\end{aligned}$$

since the integral over K in the middle is just the m-th row of the identity matrix (indexing from the center).

If instead $f, h \in \mathcal{A}$, with $\mathcal{A}$ as in Section 3, then

$$\int_{\Gamma\backslash G} (Xf)(g)\, \overline{h(g)}\, dg = \int_{\Gamma\backslash G} f(g)\, \overline{(-\overline{X}h)(g)}\, dg, \tag{94}$$

directly from the definitions (29) and the Haar measure dg.

Now suppose $f \in \mathcal{A}^d$ and $h \in \mathcal{A}^{d+a}$; then if a is even,

$$\begin{aligned}\int_{\Gamma\backslash G} (Y^a f)(g)\, \overline{h(g)^T}\, dg &= \frac{(2d+2a+1)}{\mathfrak{A}^{d,2,a}_{0,0}} \int_{\Gamma\backslash G} (X_0 f_0)(g)\, \overline{h_0(g)}\, dg \\ &= \frac{(2d+2a+1)}{\mathfrak{A}^{d,2,a}_{0,0}} \int_{\Gamma\backslash G} f_0(g)\, \overline{(-\overline{X}_0 h_0)(g)}\, dg \\ &= \frac{(2d+2a+1)}{(2d+1)\mathfrak{A}^{d,2,a}_{0,0}} \int_{\Gamma\backslash G} f(g) \overline{(\mathcal{P}^d_0(-\overline{X}_0) h_0)(g)^T}\, dg,\end{aligned}$$

so the adjoint of Y^a is

$$\widehat{Y}^a h = -\frac{(2d+2a+1)}{(2d+1)\mathfrak{A}^{d,2,a}_{0,0}} \mathcal{P}^d_0 X_0 h_0 = -\frac{(2d+2a+1)\mathfrak{A}^{d+a,2,-a}_{0,0}}{(2d+1)\mathfrak{A}^{d,2,a}_{0,0}} Y^{-a} h,$$

since the Clebsch–Gordan coefficients and X_0 are real. Similarly, if a is odd, the adjoint is

$$\widehat{Y}^a h = \frac{(2d+2a+1)}{(2d+1)\mathfrak{A}^{d,2,a}_{0,1}} \mathcal{P}^d_0 (-\overline{X}_1) h_1,$$

but we would prefer to use h_0, since this is where we have done all of our computations so far, so we use

$$\widehat{Y}^a h = \frac{(2d+2a+1)}{(2d+1)\sqrt{(d+a)(d+a+1)}\mathfrak{A}^{d,2,a}_{0,1}} \mathcal{P}^d_0 X_{-1} K_1 h_0,$$

and

$$\mathcal{P}_0^d X_{-1}K_1h_0 = \mathcal{P}_0^d K_1 X_{-1}h_0 - \sqrt{6}\mathcal{P}_0^d X_0 h_0 = \sqrt{d(d+1)}\mathcal{P}_{-1}^d X_{-1}h_0 - 0$$
$$= \sqrt{d(d+1)}\mathfrak{A}_{0,-1}^{d+a,2,-a} Y^{-a},$$

(recall the remark below Corollary 10) giving

$$\widehat{Y}^a = \frac{(2d+2a+1)\sqrt{d(d+1)}\mathfrak{A}_{0,-1}^{d+a,2,-a}}{(2d+1)\sqrt{(d+a)(d+a+1)}\mathfrak{A}_{0,1}^{d,2,a}} Y^{-a}.$$

In general, after writing out the Clebsch–Gordan coefficients, we have:

Proposition 11. *With respect to the inner products (15) on $\mathcal{A}^d$ and $\mathcal{A}^{d+a}$, the operator Y^a has adjoint*

$$\widehat{Y}^a = -(-1)^a \sqrt{\frac{2d+2a+1}{2d+1}} Y^{-a}. \tag{95}$$

By analogy with (85), we define the adjoint action on power functions $\widehat{Y}_\mu^a : \mathbb{C}^{2d+2a+1} \to \mathbb{C}^{2d+1}$ by

$$\widehat{Y}_\mu^a f := p_{-\rho-\mu}(y)\widehat{Y}^a p_{\rho+\mu}(y) f = -(-1)^a \sqrt{\frac{2d+2a+1}{2d+1}} Y_\mu^{-a} f. \tag{96}$$

6.2. *The intertwining operators.* We require an understanding of the interaction between the Lie algebra, i.e., the Y^a operators, and the intertwining operators, a.k.a. the matrices $T^d(w,\mu)$ occurring in the functional equations of the Whittaker function, as in (8).

First, a lemma:

Lemma 12. *Suppose none of the differences $\mu_i - \mu_j$, $i \neq j$, are in $\mathbb{Z}$; then for a vector $f \in \mathbb{C}^{2d+1}$, $d \geq 0$, and any $U(\mathbb{R})$-character ψ, the vector-valued function $f W^d(g,\mu,\psi)$ is identically zero as a function of g exactly when $f = 0$.*

Proof. For a nondegenerate character ψ_y, $y_1 y_2 \neq 0$, and either $\mathrm{Re}(\mu_1) > \mathrm{Re}(\mu_2) > \mathrm{Re}(\mu_3)$ or $\mathrm{Re}(\mu) = 0$, this follows from (I.3.28) and (I.3.29) and the fact that $W^d(g,\mu,\psi_{0,0})$ and $T^d(w,\mu)$ are given by products of $\mathcal{D}^d(v_{--})$, $\mathcal{D}^d(w_l)$, $\Gamma_{\mathcal{W}}^d(u,+1)$, and $\mathcal{W}^d(0,u)$, which are invertible matrices when $u \notin \mathbb{Z}$ (see (12) and (I.2.18)). This extends to the cases where only one $\mathrm{Re}(\mu_i) = \mathrm{Re}(\mu_j)$, $i \neq j$, and degenerate characters ψ_{0,y_2}, $\psi_{y_1,0}$, $\psi_{0,0}$ by the same argument as in Section I.3.5. □

Suppose none of $\mu_i - \mu_j \in \mathbb{Z}$, $i \neq j$; then for $f \in \mathbb{C}^{2d+1}$ and any $w \in W$,

$$\begin{aligned}(Y_\mu^a(fT^d(w^{-1},\mu^w)))W^{d+a}(g,\mu,\psi_{1,1}) &= Y^a(fT^d(w^{-1},\mu^w))W^d(g,\mu,\psi_{1,1})\\ &= Y^a(fW^d(g,\mu^w,\psi_{1,1}))\\ &= (Y_{\mu^w}^a f)W^{d+a}(g,\mu^w,\psi_{1,1})\\ &= ((Y_{\mu^w}^a f)T^{d+a}(w^{-1},\mu^w))W^{d+a}(g,\mu,\psi_{1,1}),\end{aligned}$$

and the lemma implies

$$Y_\mu^a(fT^d(w^{-1},\mu^w)) = (Y_{\mu^w}^a f)T^{d+a}(w^{-1},\mu^w). \tag{97}$$

This continues to an equality of meromorphic functions in μ.

6.3. *Duality.* As in (I.5.4), we define the involution $\iota : G \to G$ by $g^\iota = w_l (g^{-1})^T w_l$. To every cusp form $\phi \in \mathcal{A}^d$, we may associate a dual form $\check{\phi}(g) = \phi(g^\iota w_l) \in \mathcal{A}^d$, which is frequently used to show isomorphisms between subspaces of cusp forms. We will require an understanding of this duality, and in particular, its interaction with the Lie algebra, at the level of Whittaker functions.

Comparing (I.3.22) to (I.3.24), we have the relation

$$\mathcal{D}^d(v_{--} w_l) W^d(I, \mu, \psi_y) \mathcal{D}^d(w_l v_{--}) = W^d(I, -\mu^{w_l}, \psi_{y^\iota}),$$

and by (I.3.7), we have

$$W^d(y, \mu, \psi_{1,1}) = \mathcal{D}^d(v_{--} w_l) W^d(y^\iota, -\mu^{w_l}, \psi_{1,1}) \mathcal{D}^d(v_{--} w_l).$$

Thus in general, we have

$$W^d(g, \mu, \psi_{1,1}) = \mathcal{D}^d(v_{--} w_l) W^d(v_{--} g^\iota w_l, -\mu^{w_l}, \psi_{1,1}). \tag{98}$$

For $f \in C^\infty(G)$, set $\check{f}(g) = f(g^\iota w_l)$; then the action of the K_j operators on the dual form $\check{f}$ is given by

$$K_j \check{f}(g) = \frac{d}{dt} f(g^\iota w_l \exp(-t K_j{}^T)) \Big|_{t=0} = \frac{d}{dt} f(g^\iota w_l \exp t K_j) \Big|_{t=0} = \widetilde{K_j} f(g^\iota).$$

Similarly $X_j \check{f}(g) = -\widetilde{X_j} f(g^\iota)$. Therefore, for $f \in \mathbb{C}^d$,

$$\begin{aligned}
(Y^a_\mu(f \mathcal{D}^d(v_{--} w_l))) W^{d+a}(g, \mu, \psi_{1,1}) &= Y^a f W^d(v_{--} g^\iota w_l, -\mu^{w_l}, \psi_{1,1}) \\
&= -(Y^a_{-\mu^{w_l}} f) W^{d+a}(v_{--} g^\iota w_l, -\mu^{w_l}, \psi_{1,1}) \\
&= -(Y^a_{-\mu^{w_l}} f) \mathcal{D}^d(v_{--} w_l) W^{d+a}(g, \mu, \psi_{1,1}),
\end{aligned}$$

and we conclude as in the previous subsection that

$$Y^a_\mu(f \mathcal{D}^d(v_{--} w_l)) = -(Y^a_{-\mu^{w_l}} f) \mathcal{D}^d(v_{--} w_l). \tag{99}$$

7. The minimal-weight Whittaker functions

In this section, we will prove Theorem 5 and analyze some additional, degenerate cases of the Jacquet–Whittaker integral for the purpose of proving they cannot occur among the cusp forms. First, we give some general results on Whittaker functions that were not needed for the previous paper.

7.1. *The differential equations satisfied by the Whittaker functions.* We wish to develop raising and lowering operators on the entries of the vector-valued Whittaker functions. Suppose $\Phi : G \to \mathbb{C}^{2d+1}$ is defined over the Iwasawa decomposition by

$$\Phi(xyk) = \phi(xy) \mathcal{D}^d(k), \quad \phi(xy) = (\phi_{-d}(xy), \dots, \phi_d(xy));$$

then the components of $\Phi = (\Phi_{-d}, \dots, \Phi_d)$ are given by

$$\Phi_m = \sum_{m'=-d}^{d} \phi_{m'} \mathcal{D}^d_{m',m}, \quad \Phi = \sum_{m'=-d}^{d} \phi_{m'} \mathcal{D}^d_{m'}$$

and requiring $\Delta_i \Phi = \lambda_i \Phi,\ i = 1, 2$, is equivalent to

$$\begin{aligned}0 &= (4\Delta_1^\circ + m'(Z_2 - Z_{-2}) - 4\lambda_1)\phi_{m'} \\ &\quad - \sqrt{d(d+1) - m'(m'+1)}\,Z_{-1}\phi_{m'+1} \\ &\quad - \sqrt{d(d+1) - m'(m'-1)}\,Z_1\phi_{m'-1}, \\ 0 &= (48\Delta_2^\circ - m'\mathcal{T}_0 - 48\lambda_2)\phi_{m'} \\ &\quad + \sqrt{d(d+1) - m'(m'+1)}\,(\mathcal{T}_1 - 6(m'+1)(Z_{-1} - Z_1))\phi_{m'+1} \\ &\quad + \sqrt{d(d+1) - m'(m'-1)}\,(\mathcal{T}_{-1} - 6(m'-1)(Z_{-1} - Z_1))\phi_{m'-1},\end{aligned}$$

by (46) and (47).

Suppose also that $\phi(xy) = \psi_{1,1}(x)\phi(y)$; then

$$\begin{aligned}0 &= (\Delta_1^{\mathrm{Whitt}} - 2\pi m' y_2 - \lambda_1)\phi_{m'}(y) \\ &\quad - \pi\sqrt{d(d+1) - m'(m'+1)}\,y_1\phi_{m'+1}(y) \\ &\quad - \pi\sqrt{d(d+1) - m'(m'-1)}\,y_1\phi_{m'-1}(y),\end{aligned} \tag{100}$$

$$\begin{aligned}0 &= (\Delta_2^{\mathrm{Whitt}} + 2\pi m' y_2(y_1\partial_{y_1} - 1) - \lambda_2)\phi_{m'}(y) \\ &\quad + \pi\sqrt{d(d+1) - m'(m'+1)}\,y_1(1 - y_2\partial_{y_2} - 2\pi y_2)\phi_{m'+1}(y) \\ &\quad + \pi\sqrt{d(d+1) - m'(m'-1)}\,y_1(1 - y_2\partial_{y_2} + 2\pi y_2)\phi_{m'-1}(y),\end{aligned} \tag{101}$$

where

$$\begin{aligned}\Delta_1^{\mathrm{Whitt}} &= -y_1^2\partial_{y_1}^2 - y_2^2\partial_{y_2}^2 + y_1y_2\partial_{y_1}\partial_{y_2} + 4\pi^2y_1^2 + 4\pi^2y_2^2, \\ \Delta_2^{\mathrm{Whitt}} &= -y_1^2y_2\partial_{y_1}^2\partial_{y_2} + y_1y_2^2\partial_{y_1}\partial_{y_2}^2 - 4\pi^2y_1y_2^2\partial_{y_1} + 4\pi^2y_1^2y_2\partial_{y_2} + y_1^2\partial_{y_1}^2 - y_2^2\partial_{y_2}^2 - 4\pi^2y_1^2 + 4\pi^2y_2^2.\end{aligned}$$

Some rearranging gives

$$S_{m'}^{\pm}\phi_{m'}(y) = \pm\sqrt{d(d+1) - m'(m'\pm 1)}\,\phi_{m'\pm 1}(y), \tag{102}$$

$$\begin{aligned}S_{m'}^{\pm} := \frac{1}{4\pi^2y_1y_2}\big((1 - y_2\partial_{y_2} \pm 2\pi y_2)(\Delta_1^{\mathrm{Whitt}} - \lambda_1) \\ + (\Delta_2^{\mathrm{Whitt}} - \lambda_2) + 2\pi m' y_2(y_1\partial_{y_1} + y_2\partial_{y_2} - 1 \mp 2\pi y_2)\big).\end{aligned} \tag{103}$$

One particular consequence is that the full vector-valued Whittaker function may be generated from any given entry, so any vector-valued Whittaker function is identically zero exactly when any entry is identically zero. Precisely, we have:

Lemma 13. *For any vector-valued function $\Phi(xyk) = \psi_{1,1}(x)\phi(y)\mathcal{D}^d(k)$, with $\phi(y) = (\phi_{-d}(y), \ldots, \phi_d(y))$, which satisfies $\Delta_i\Phi = \lambda_i\Phi,\ i = 1, 2$, and any $-d \le m_1 < m_2 \le d$, we have*

$$\phi_{m_2} = C_1 S_{m_2-1}^{+} \cdots S_{m_1+1}^{+} S_{m_1}^{+}\phi_{m_1}, \qquad \phi_{m_1} = C_2 S_{m_1+1}^{-} \cdots S_{m_2-1}^{-} S_{m_2}^{-}\phi_{m_2},$$

where

$$\frac{1}{C_1} = \prod_{j=m_1}^{m_2-1}\sqrt{d(d+1) - j(j+1)}, \qquad \frac{1}{C_2} = (-1)^{m_2-m_1}\prod_{j=m_1+1}^{m_2}\sqrt{d(d+1) - j(j-1)}.$$

7.2. ***The central entry.*** We turn now to the proof of Theorem 5, which is the evaluation of the Jacquet integral at the minimal K-types. As mentioned in the Introduction, the proof occurs in two steps; in this subsection, we directly evaluate the Jacquet integral at the central entry, and in the next, we show the Mellin–Barnes integrals of Theorem 5 satisfy (102) and (100) to complete the theorem. For the central entry, we also consider the evaluation of $W^d_{-d,0}(g,\mu,\psi_{1,1})$, where μ is of the form $(d-1,0,1-d)$; we will see this cannot occur as the Whittaker function of a cusp form, but we still require knowledge of precisely when the function is identically zero.

Suppose $\mu_1-\mu_2=d-1$, $d\geq 2$, and temporarily suppose $\mathrm{Re}(\mu_2-\mu_3)$ is large; we will reach the cases of Theorem 5 by analytic continuation, below. The cases $d=0,1$ of Theorem 5 may be handled similarly. Starting from (I.3.22), we send $u_3\mapsto u_3\sqrt{1+u_2^2}$ to obtain

$$W^d_{-d,0}(I,\mu,\psi_y)=(-1)^d\int_{\mathbb{R}^2}(1+u_3^2)^{\frac12(-1+\mu_3-\mu_1)}(1+u_2^2)^{\frac12(-1+\mu_3-\mu_2)}\mathcal{W}_{-d}\left(y_1\frac{\sqrt{1+u_3^2}}{\sqrt{1+u_2^2}},\mu_1-\mu_2\right)\\ \times\mathrm{d}^d_{-d,0}\left(\frac{-u_3}{\sqrt{1+u_3^2}}\right)e\left(-y_1\frac{u_2u_3}{\sqrt{1+u_2^2}}\right)e(-y_2u_2)\,du_2\,du_3. \quad (104)$$

Then (11) and [DLMF, equation 13.18.2] imply

$$\mathcal{W}_{-d}(y,d-1)=\frac{(2\pi)^d y^{d-1}}{(d-1)!}\exp(-2\pi y). \quad (105)$$

We also have, by (70) and (72),

$$\mathrm{d}^d_{-d,0}(x)=\mathrm{d}^d_{0,d}(x)=\frac{\sqrt{(2d)!}}{d!\,2^d}(1-x^2)^{\frac12 d}.$$

Plugging in the two previous displays, applying (I.3.17) (or the definition of the gamma function) to each of the three exponentials, and evaluating the u integrals with (I.2.27) gives

$$W^d_{-d,0}(I,\mu,\psi_y)\\ =(-1)^d\frac{\sqrt{(2d)!}\pi^d y_1^{d-1}}{d!\,(d-1)!}\int_{\mathrm{Re}(s)=\epsilon}(2\pi y_1)^{-s_1-s_3}(2\pi y_2)^{-s_2}\Gamma(s_1)\Gamma(s_3)\Gamma(s_2)\cos\left(\tfrac12\pi s_2\right)\cos\left(\tfrac12\pi s_3\right)\\ \times B\left(\tfrac12(1-s_3-s_2),\tfrac12(\mu_1-\mu_2-s_1+s_2)\right)\\ \times B\left(\tfrac12(1-s_3),\tfrac12(d+\mu_3-\mu_2+s_1+s_3)\right)\frac{ds}{(2\pi i)^3},$$

where

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

is the Euler beta function. We may now suppose $\mu=\left(\frac12(d-1)+it,-\frac12(d-1)+it,-2it\right)$ or $\mu=(d-1,0,1-d)$, as the convergence is clear; i.e., the above integral converges to an entire function of μ by the exponential decay coming from Stirling's formula, so we have the anticipated analytic continuation.

Send $s_1 \mapsto s_1 - s_3$, and apply the duplication and reflection properties of the gamma function so that

$$W^d_{-d,0}(I,\mu,\psi_y) \\ = \frac{(-1)^d\sqrt{(2d)!}\,\pi^{\frac{3}{2}}}{8\,d!\,(d-1)!}\int_{\mathrm{Re}(s)=\epsilon}(\pi y_1)^{d-1-s_1}(\pi y_2)^{-s_2}\frac{\Gamma\big(\tfrac12 s_2\big)\Gamma\big(\tfrac12(1+\mu_1-\mu_2+s_1)\big)}{\Gamma\big(\tfrac12(1-s_2)\big)\Gamma\big(\tfrac12(1+\mu_1-\mu_2-s_1)\big)} \\ \times\frac{\Gamma\big(\tfrac12(1-s_3-s_2)\big)\Gamma\big(\tfrac12(\mu_1-\mu_2-s_1+s_3+s_2)\big)\Gamma\big(\tfrac12 s_3\big)\Gamma\big(\tfrac12(s_1-s_3)\big)\Gamma\big(\tfrac12(1+s_1-s_3)\big)}{\Gamma\big(\tfrac12(d+1+\mu_3-\mu_2+s_1-s_3)\big)}\frac{ds}{(2\pi i)^3}$$

using

$$1+\mu_1-\mu_3 = d+\mu_2-\mu_3 = \begin{cases} \tfrac12(d+1)+3it & \text{if } \mu=\big(\tfrac12(d-1)+it,\,-\tfrac12(d-1)+it,\,-2it\big),\\ 2d-1 & \text{if } \mu=(d-1,0,1-d).\end{cases}$$

We apply (62) on $-\tfrac12 s_3$, substitute $(s_1,s_2)\mapsto(s_1+d-1-\mu_1,\,s_2+\mu_3)$, and use (I.3.6) to obtain

$$W^d_{-d,0}(y,\mu,\psi_{1,1}) = \frac{(-1)^d}{4\pi^2}\int_{\mathrm{Re}(s)=\mathfrak{s}}(\pi y_1)^{1-s_1}(\pi y_2)^{1-s_2}\frac{G^d_0(s,\mu)}{\Lambda^*(\mu)}\frac{ds}{(2\pi i)^2}, \tag{106}$$

with G^d as in (24), and using the contour

$$\mathfrak{s} = \begin{cases} (\epsilon,\epsilon) & \text{if } \mu=\big(\tfrac12(d-1)+it,\,-\tfrac12(d-1)+it,\,-2it\big),\\ (\epsilon,\,d-1+\epsilon) & \text{if } \mu=(d-1,0,1-d).\end{cases}$$

The opposite case $W^d_{d,0}(y,\mu,\psi_{1,1})=0$ follows from the pole of the gamma function of (11) in (104).

7.3. ***The remaining components of the Whittaker function.*** We now complete the proof of Theorem 5. We continue to assume $d\ge 2$ and specify $\mu=\big(\tfrac12(d-1)+it,\,-\tfrac12(d-1)+it,\,-2it\big)$. We've computed the base case $m'=0$ directly, and the cases $m'=\pm1$ can be verified from the $S^{\pm}_{m'}$ operator, so we have to compute $|m'|\ge 2$. It is sufficient to verify the terms $1\le|m'|\le d-1$ satisfy (100), and in the following proof we assume $|m'|\ge 2$ for convenience. Let $m'=\varepsilon m$; then we need to verify

$$0=\int_{\mathrm{Re}(s)=\mathfrak{s}}\Bigg((\Delta_1^{\mathrm{Whitt}}-2\pi m' y_2-\lambda_1(\mu))(\pi y_1)^{1-s_1}(\pi y_2)^{1-s_2}\sum_{\ell=0}^{m}\varepsilon^\ell\binom{m}{\ell}\widetilde{G}^d((d-m,\ell),s,\mu) \\ -(d-m)(\pi y_1)^{1-s_1}(\pi y_2)^{1-s_2}\sum_{\ell=0}^{m+1}\varepsilon^\ell\binom{m+1}{\ell}\widetilde{G}^d((d-m-1,\ell),s,\mu) \\ -(d+m)(\pi y_1)^{1-s_1}(\pi y_2)^{1-s_2}\sum_{\ell=0}^{m-1}\varepsilon^\ell\binom{m-1}{\ell}\widetilde{G}^d((d-m+1,\ell),s,\mu)\Bigg)\frac{ds}{(2\pi i)^2}.$$

We apply the differential operator and shift s on the various terms so we may deal with the Mellin transform directly. We factor out

$$\Gamma\big(\tfrac12(d-1-\mu_1+s_1)\big)\Gamma\big(\tfrac12(d-\mu_1+s_1)\big)\Gamma\big(\tfrac12(\mu_1+s_2)\big)\Gamma\big(\tfrac12(1+\mu_1+s_2)\big),$$

which replaces the terms $\widetilde{G}^d(\cdots)$ with a polynomial times a beta function of the form

$$B\big(\tfrac{1}{2}(d-m+a-\mu_3+s_1),\,\tfrac{1}{2}(\ell+b+\mu_3+s_2)\big),\qquad a=0,2,\ b=0,1,2.$$

Now apply

$$B(x+1,y)=B(x,y)-B(x,y+1)$$

to normalize $a\mapsto 0$, and substitute $\ell\mapsto\ell-b$ to align the beta functions. After some algebra and removing a factor (s_1-s_2-2it), we need to show

$$\begin{aligned}0=&(s_2+\mu_3)B\big(\tfrac{1}{2}(d-m-\mu_3+s_1),\,\tfrac{1}{2}(\mu_3+s_2)\big)+\varepsilon m(s_2+\mu_3+1)B\big(\tfrac{1}{2}(d-m-\mu_3+s_1),\,\tfrac{1}{2}(\mu_3+s_2+1)\big)\\&+\sum_{\ell=2}^{m}\varepsilon^\ell\bigg(\binom{m}{\ell}(s_2+\mu_3+\ell)-\binom{m}{\ell-2}(d-m+s_1+s_2+\ell-2)\bigg)B\big(\tfrac{1}{2}(d-m-\mu_3+s_1),\,\tfrac{1}{2}(\mu_3+s_2+\ell)\big)\\&\quad-\varepsilon^{m+1}m(d-1+s_1+s_2)B\big(\tfrac{1}{2}(d-m-\mu_3+s_1),\,\tfrac{1}{2}(\mu_3+s_2+m+1)\big)\\&\quad\quad-\varepsilon^{m}(d+s_1+s_2)B\big(\tfrac{1}{2}(d-m-\mu_3+s_1),\,\tfrac{1}{2}(\mu_3+s_2+m+2)\big)\end{aligned}$$

The even and odd terms of this sum then separately telescope to zero.

7.4. ***The bad Whittaker functions.*** Suppose $\mu=(d-1,0,1-d)$. It is a well-known fact (see the corollary to [Harish-Chandra 1968, Lemma 15]) that cusp forms, and hence their Fourier coefficients are bounded (as functions of G). Then (106) shows the Whittaker functions $W^d_{-d}(\cdot,\mu,\psi_{1,1})$ are nonzero and have bad asymptotics; i.e., they are not suitable for cusp forms since the central entry tends to infinity as $y_2\to 0$ (shift the s contours to the left past the pole at $(s_1,s_2)=(0,d-1)$). We may instead rule out any cusp form having such a Whittaker function by noticing it would necessarily have a real eigenvalue of the skew-symmetric operator Y^0, and this is what we do in Section 8.3.

8. Going down

We determine the spectral parameters and Whittaker functions of all forms which are killed by both of Y^{-1}, Y^{-2}. Equivalently, we determine vectors in $\mathbb{C}^{2d+1}$ which are killed by both of Y^{-1}_μ, Y^{-2}_μ. The unitaricity conditions on the spectral parameters of Maass cusp forms, meaning the symmetries of the differential operators Δ_1 and Δ_2, imply that, for such forms, $-\bar{\mu}$ is a permutation of μ, and further, because of the absence of poles in the required functional equations of the Whittaker function, we may assume $\mathrm{Re}(\mu_1)\geq\mathrm{Re}(\mu_2)\geq\mathrm{Re}(\mu_3)$, so

$$\mu=(it_1,it_2,-i(t_1+t_2))\quad\text{or}\quad\mu=(x+it,-2it,-x+it),\tag{107}$$

where $t_1-t_2,\,2t_1+t_2,\,t_1+2t_2\neq 0,\ x\geq 0$, and possibly $t=0$.

8.1. ***For the power function.*** Define

$$g_1=-\sqrt{6}(1+\mu_3)\mathrm{u}^{2,+}_2+(\mu_1-\mu_2-1)\mathrm{u}^{2,+}_0,\tag{108}$$

$$g^{d,\delta,\varepsilon}_2=\sum_{0\leq 2j+\delta\leq d}g^d_{2,\delta,2j+\delta}\mathrm{u}^{d,\varepsilon}_{2j+\delta},\tag{109}$$

where the coefficients are given by $g^d_{2,0,0} = \frac{1}{2}$ and

$$g^d_{2,k,2j+k} = (-1)^j \prod_{i=0}^{j-1} \frac{\sqrt{(d-1-2i-k)(d-2i-k)}(3\mu_3+2d-1+2i+k)}{\sqrt{(d+1+2i+k)(d+2+2i+k)}(\mu_1-\mu_2-1-2i-k)} \tag{110}$$

for $0 < 2j+k \le d$. We also define $g_3^{d,\delta,\varepsilon}$ using the same coefficients as $g_2^{d,\delta,\varepsilon}$; i.e.,

$$g^d_{3,\delta,2j+\delta} := g^d_{2,\delta,2j+\delta}, \quad 0 \le 2j+\delta < d, \tag{111}$$

but take

$$g^d_{3,\delta,d} := \frac{g^d_{2,\delta,d-2}}{\sqrt{d(2d-1)}} \tag{112}$$

to be the coefficient of $\mathfrak{u}_d^{d,\varepsilon}$ instead. Lastly, for d odd, define

$$g_4^{d,\varepsilon} = \sum_{\kappa \le 2j+\kappa \le d} g^d_{2,\kappa,2j+\kappa} \mathfrak{u}^{d,\varepsilon}_{2j+\kappa}, \tag{113}$$

where $\kappa = \frac{1}{2}(d+1)$.

Proposition 14. *If $f \in \mathbb{C}^{2d+1}$ is such that*

$$Y_\mu^{-1} f = 0 \quad \textit{and} \quad Y_\mu^{-2} f = 0, \tag{114}$$

subject to (107), then at least one of the following is true:

(1) $f = 0$.

(2) $d \in \{0, 1\}$.

(3) *$d = 2$ and f is a multiple of g_1.*

(4) *$\mu = \left(\frac{1}{2}(d-1)+it, -2it, -\frac{1}{2}(d-1)+it\right)$, with either d even or $t \ne 0$, and f is a linear combination of $g_2^{d,0,\varepsilon}$ and $g_2^{d,1,-\varepsilon}$, with $\varepsilon = (-1)^d$.*

(5) *$\mu = \left(\frac{1}{2}(d-1), 0, -\frac{1}{2}(d-1)\right)$, d odd, and f is a linear combination of $g_2^{d,\delta,\varepsilon}$, $g_4^{d,+}$ and $g_4^{d,-}$, with $\delta \equiv \kappa+1 \pmod 2$, $\varepsilon = (-1)^\kappa$, $\kappa = \frac{1}{2}(d+1)$.*

(6) *$\mu = (d-1, 0, 1-d)$, and f is a linear combination of $g_3^{d,1,\varepsilon}$, $g_3^{d,0,\varepsilon}$, $\mathfrak{u}_d^{d,+}$ and $\mathfrak{u}_d^{d,-}$, with $\varepsilon = (-1)^d$.*

We call such an f *power-function minimal* for μ at weight d.

We note that g_1 in case (3) may be a false positive in the sense that it is killed by the lowering operators, but not necessarily an eigenfunction of Y^0_μ, while the vectors of the other cases are all also eigenfunctions of Y^0_μ. (We will not show this directly, but it can be deduced from Proposition 15 and (97).) That is, if none of differences $\mu_i - \mu_j$ are 1, then

$$\sqrt{35} Y_\mu^{-2} Y_\mu^0 g_1 = -8\sqrt{2}(\mu_1-\mu_2-1)(\mu_1-\mu_3-1)(\mu_2-\mu_3-1)\mathfrak{u}_0^{0,+} \ne 0, \tag{115}$$

so the minimal weight of the corresponding principle series representation is $d = 0$ instead of $d = 2$, and of course, the instances where some $\mu_i - \mu_j = 1$ reduce to cases (4) or (6).

Proof of Proposition 14. The result is trivial for $d \in \{0, 1\}$, and $d = 2$ is simple to compute (note that $Y_\mu^{-1}\mathfrak{u}_0^{2,\pm}$, $Y_\mu^{-1}\mathfrak{u}_2^{2,\pm}$ and $Y_\mu^{-2}\mathfrak{u}_1^{2,\pm}$ all vanish), so we may assume $d \geq 3$.

For clarity, we write $f^d := f$. We may split f^d according to characters of V,

$$f^d = \sum_{|m'|\leq d} f^d_{m'}\mathfrak{v}^d_{m'} = \sum_{\pm,\delta\in\{0,1\}} f^{d,\delta,\pm},$$

$$f^{d,\delta,\pm} = \sum_{2j+\delta\leq d} (f^d_{2j+\delta} \pm (-1)^d f^d_{-2j-\delta})\mathfrak{u}^{d,\pm}_{2j+\delta} \times \begin{cases} \frac{1}{2} & \text{if } 2j+\delta = 0, \\ 1 & \text{otherwise,} \end{cases}$$

and f^d is a zero of both Y_μ^{-1} and Y_μ^{-2} exactly when all of the $f^{d,\delta,\pm}$ are, so we now fix d, μ and a choice of $\delta \in \{0, 1\}$ and parity $\varepsilon = \pm 1$ and assume f^d is of the form

$$f^d = \sum_{0\leq 2j+\delta\leq d} f^d_{2j+\delta}\mathfrak{u}^{d,\varepsilon}_{2j+\delta},$$

after relabeling the subscripts. Let

$$f^{d-1} = \sqrt{2d(d-1)(d+1)(2d+1)}\,Y_\mu^{-1} f^d \quad \text{and} \quad f^{d-2} = 2\sqrt{d(d-1)(2d-1)(2d+1)}\,Y_\mu^{-2} f^d.$$

From (91) and (93), their coefficients are given by

$$\begin{aligned} f^{d-1}_j = &-\sqrt{(d-1-j)(d+j)(d+1+j)(d+2+j)}(\mu_1-\mu_2-1-j) f^d_{j+2} \\ &+2j\sqrt{(d-j)(d+j)}(3\mu_3+d) f^d_j \\ &\quad + c_j\sqrt{(d-j)(d+1-j)(d+2-j)(d-1+j)}(\mu_1-\mu_2-1+j) f^d_{j-2}, \end{aligned} \tag{116}$$

$$\begin{aligned} f^{d-2}_j = &\sqrt{(d-1+j)(d+j)(d+1+j)(d+2+j)}(\mu_1-\mu_2-1-j) f^d_{j+2} \\ &+2\sqrt{(d-1-j)(d-j)(d-1+j)(d+j)}(3\mu_3+2d-1) f^d_j \\ &\quad + c_j\sqrt{(d-1-j)(d-j)(d+1-j)(d+2-j)}(\mu_1-\mu_2-1+j) f^d_{j-2} \end{aligned} \tag{117}$$

for $j \geq 2$, where $c_j = 2$ if $j = 2$ and 1 otherwise. When $\varepsilon = -(-1)^d$, we set $f^d_0 = 0$. Some care must be taken when dealing with the coefficients f^{d-1}_j, f^{d-2}_j for $j = 0, 1$, as the previous expressions do not necessarily apply:

$$\begin{aligned} f^{d-1}_1 = &-\sqrt{(d-2)(d+1)(d+2)(d+3)}(\mu_1-\mu_2-2) f^d_3 \\ &+\sqrt{(d-1)(d+1)}(2(3\mu_3+d)+\varepsilon(-1)^d d(\mu_1-\mu_2)) f^d_1, \end{aligned} \tag{118}$$

$$\begin{aligned} f^{d-2}_1 = &\sqrt{d(d+1)(d+2)(d+3)}(\mu_1-\mu_2-2) f^d_3 \\ &+\sqrt{d(d-2)(d-1)(d+1)}(2(3\mu_3+2d-1)+\varepsilon(-1)^d(\mu_1-\mu_2)) f^d_1, \end{aligned} \tag{119}$$

$$f^{d-1}_0 = -\sqrt{d(d-1)(d+1)(d+2)}(\mu_1-\mu_2-1) f^d_2\, \delta_{\varepsilon=-(-1)^d}, \tag{120}$$

$$f^{d-2}_0 = \sqrt{d(d-1)}\big(\sqrt{(d+1)(d+2)}(\mu_1-\mu_2-1) f^d_2 + 2\sqrt{d(d-1)}(3\mu_3+2d-1) f^d_0\big)\delta_{\varepsilon=+(-1)^d}. \tag{121}$$

Two useful linear combinations are

$$s_{1,j} := \frac{1}{2(d-1)\sqrt{d+j}}(\sqrt{d-1-j}\, f_j^{d-1} - \sqrt{d-1+j}\, f_j^{d-2})$$
$$= -\sqrt{(d+1+j)(d+2+j)}(\mu_1-\mu_2-1-j)\, f_{j+2}^d - \sqrt{(d-1-j)(d-j)}(3\mu_3+2d-1+j)\, f_j^d, \quad (122)$$
$$s_{2,j} := \frac{1}{2(d-1)\sqrt{d-j}}(\sqrt{d-1+j}\, f_j^{d-1} + \sqrt{d-1-j}\, f_j^{d-2})$$
$$= \sqrt{(d-1+j)(d+j)}(3\mu_3+2d-1-j)\, f_j^d + \sqrt{(d+1-j)(d+2-j)}(\mu_1-\mu_2-1+j)\, f_{j-2}^d \quad (123)$$

for $3 \le j \le d-2$. The given expression (122) for $s_{1,j}$ continues to hold for $j = 1, 2$, and for $s_{2,j}$, $j = 2, 1$, we have the expressions

$$s_{2,2} = \sqrt{(d+1)(d+2)}(3\mu_3+2d-3)\, f_2^d + 2\sqrt{d(d-1)}(\mu_1-\mu_2+1)\, f_0^d, \quad (124)$$
$$s_{2,1} = -2 f_1^d \sqrt{d(d+1)} \times \begin{cases} (\mu_2-\mu_3+1-d) & \text{if } \varepsilon = +(-1)^d, \\ (\mu_1-\mu_3+1-d) & \text{if } \varepsilon = -(-1)^d, \end{cases} \quad (125)$$

but for $j = 0$, it is simplest to use f_0^{d-1} and f_0^{d-2} directly. Note that $s_{1,j} = s_{2,j} = 0$ is fully equivalent to $f_j^{d-1} = f_j^{d-2} = 0$ for each $1 \le j \le d-2$.

For $1 \le j \le d-4$, we further cancel to obtain

$$s_{3,j} := (3\mu_3+2d-3-j)s_{1,j} + (\mu_1-\mu_2-1-j)s_{2,j+2}$$
$$= -4\sqrt{(d-j)(d-1-j)}(\mu_1-\mu_3+1-d)(\mu_2-\mu_3+1-d)\, f_j^d, \quad (126)$$

and, for $3 \le j \le d-2$, we have

$$s_{4,j} := (3\mu_3+2d-3-j)s_{1,j-2} + (\mu_1-\mu_2-1-j)s_{2,j}$$
$$= 4\sqrt{(d+j)(d-1+j)}(\mu_1-\mu_3+1-d)(\mu_2-\mu_3+1-d)\, f_j^d. \quad (127)$$

Then

$$\left\{\begin{array}{l} f^{d-1} = 0 \\ f^{d-2} = 0 \end{array}\right\} \iff \begin{cases} f_{d-1}^{d-1} = f_0^{d-1} = f_0^{d-2} = 0, \\ s_{1,j} = 0,\ j = 1, \dots, d-2, \\ s_{2,1} = s_{2,2} = 0, \\ s_{3,j-2} = s_{4,j} = 0,\ j = 3, \dots, d-2, \end{cases} \quad (128)$$

because for each $3 \le j \le d-2$ the coefficient of $s_{2,j}$ in one of $s_{3,j-2}$ or $s_{4,j}$ is nonzero.

The proof now proceeds by cases on μ.

Case I: $\mu_1-\mu_3+1 \ne d$, $\mu_2-\mu_3+1 \ne d$. These assumptions (recall (107)) imply $\mu_1-\mu_2+1 \ne d$, as well. Then $s_{3,j} = 0$ is equivalent to $f_j^d = 0$ for all $1 \le j \le d-4$, and similarly $s_{4,j} = 0$ is equivalent to $f_j^d = 0$ on $3 \le j \le d-2$. (Note that for $d \ge 6$, $\{1 \le j \le d-4\} \cup \{3 \le j \le d-2\} = \{1 \le j \le d-2\}$, but the same does not hold for $3 \le d \le 5$.)

Now suppose $f^{d-1} = 0$ and $f^{d-2} = 0$ with $d \ge 6$. Then (126), (127) and (125) imply $f_j^d = 0$, $1 \le j \le d-2$. In the case $\varepsilon = +(-1)^d$ and $\delta = 0$, (124) implies $f_0^d = 0$ since $\mu_1-\mu_2 \ne -1$.

If $d \equiv \delta \pmod 2$, then $s_{1,d-2} = 0$ implies $f_d^d = 0$ and we have $f^d = 0$. If $d \not\equiv \delta \pmod 2$, then $f_{d-3}^{d-2} = f_{d-1}^{d-1} = 0$ implies $f_{d-1}^d = 0$ and we have $f^d = 0$.

For $3 \le d \le 5$, first suppose $\delta = 1$, and note (125) implies $f_1^d = 0$. Then $s_{1,1} = 0$ implies $f_3^d = 0$ unless $\mu_1 - \mu_2 = 2$, in which case $\mu = (2, 0, -2)$ (recall (107)), so $d = 4$ and $s_{2,3} = 0$ again implies $f_3^d = 0$. Lastly, if $d = 5$, then $s_{1,3} = 0$ implies $f_5^d = 0$.

Now suppose $3 \le d \le 5$ with $\delta = 0$; then, as before, (124) implies $f_0^d = 0$, and $f_0^{d-1} = f_0^{d-2} = 0$ implies $f_2^d = 0$ unless $\mu_1 - \mu_2 = 1$, in which case (124) works since $\mu = (1, 0, -1)$ implies $d \ne 3$. When $d \in \{4, 5\}$, $s_{1,2} = 0$ implies $f_4^d = 0$ unless $\mu_1 - \mu_2 = 3$, in which case $f_4^{d-1} = 0$ works since $d = 5$ and $\mu_3 = -3$.

Case II: $\mu = \left(\frac{1}{2}(d-1) + it, -2it, -\frac{1}{2}(d-1) + it\right)$ with $t \ne 0$ or d even. Under the current assumptions, we have all $s_{3,j} = 0$, $s_{4,j} = 0$, and we use (128). When $\delta = 0$, $s_{2,1} = 0$ becomes trivial, and similarly, when $\delta = 1$, $s_{2,2} = f_0^{d-1} = f_0^{d-2} = 0$ becomes trivial.

Since $\mu_1 - \mu_2 - 1 \notin \mathbb{Z}$, $s_{1,j} = 0$ is equivalent to

$$f_{j+2}^d = -\frac{\sqrt{(d-1-j)(d-j)}(3\mu_3 + 2d - 1 + j)}{\sqrt{(d+1+j)(d+2+j)}(\mu_1 - \mu_2 - 1 - j)} f_j^d. \tag{129}$$

If $\delta = 0$ and $\varepsilon = +(-1)^d$, then $s_{2,2} = 0$ is redundant over $f_0^{d-2} = 0$. If $\delta = 0$ and $\varepsilon = -(-1)^d$, then $f_0^{d-1} = 0$ implies $f_2^d = 0$ and hence $f^d = 0$. If $\delta = 1$ and $\varepsilon = +(-1)^d$, then $s_{2,1} = 0$ implies $f_1^d = 0$ and hence $f^d = 0$. If $\delta = 1$ and $\varepsilon = -(-1)^d$, then $s_{2,1} = 0$ is trivial.

This completes the analysis of conclusion (4) of the proposition. The need for $g_{2,0,0}^{d,\delta,\varepsilon} = \frac{1}{2}$ comes from the spare 2 in f_0^{d-2}, as compared to $s_{1,j}$, $j \ge 1$.

Case III: $\mu = (d-1, 0, 1-d)$. Since $\mu_1 - \mu_2 - 1 = -(3\mu_3 + 2d - 1) = d - 2$, this is identical to Case II, except that $s_{1,d-2} = 0$ is trivial, and now when $\delta = 1$, $s_{2,1} = 0$ implies $f_1^d = 0$ unless $\varepsilon = +(-1)^d$. The coefficient $g_{2,1,d}^d$ is undefined, but we may freely choose the coefficient of $\mathrm{u}_d^{d,\varepsilon}$ since $Y_\mu^{-1}\mathrm{u}_d^{d,\varepsilon} = 0$ and $Y_\mu^{-2}\mathrm{u}_d^{d,\varepsilon} = 0$. Our particular choice (112) makes $g_3^{d,\delta,\varepsilon}$ an eigenfunction of Y_μ^0, as we will see later. This completes the analysis of conclusion (6) of the proposition.

Case IV: $\mu = \left(\frac{1}{2}(d-1), 0, -\frac{1}{2}(d-1)\right)$, d odd. Since $\mu_1 - \mu_2 - 1 = \kappa - 2$, the case $\delta \not\equiv \kappa \pmod 2$ is identical to Case II. In the case $\delta \equiv \kappa \pmod 2$, $s_{1,\kappa-2} = 0$ implies $f_{\kappa-2}^d = 0$, and recursively $s_{1,j} = 0$ implies $f_j^d = 0$ for $1 \le j < \kappa$ and $s_{2,2} = 0$ implies $f_0^d = 0$. Then (129) applies for $j \ge \kappa$, and this completes the analysis of conclusion (5) of the proposition. □

8.2. ***For Whittaker functions.*** We wish to prove a version of Proposition 14 for Whittaker functions, but this is somewhat complicated by the possibility that the Whittaker function itself may be zero, and by the unpleasant shape of the vectors in Proposition 14, so we start by a careful examination of the intertwining operators.

Proposition 15. *Suppose $d \ge 2$, and let C be a nonzero constant, depending on d and μ, whose value may differ between occurrences:*

(1) *If* $\mu = \left(\frac{1}{2}(d-1)+it, -2it, -\frac{1}{2}(d-1)+it\right)$, *with either* d *even or* $t \neq 0$, *then*

$$g_2^{d,0,\varepsilon} = C\mathfrak{u}_d^{d,\varepsilon} T^d(w_3, \mu^{w_3}), \quad g_2^{d,1,-\varepsilon} = C\mathfrak{u}_d^{d,-\varepsilon} T^d(w_3, \mu^{w_3}),$$

with $\varepsilon = (-1)^d$.

(2) *If* $\mu = \left(\frac{1}{2}(d-1), 0, -\frac{1}{2}(d-1)\right)$ *with* $d \equiv 1 \pmod 4$, *then*

$$g_2^{d,0,-} = C\mathfrak{u}_d^{d,-} T^d(w_3, \mu^{w_3}), \quad g_4^{d,+} = C\mathfrak{u}_d^{d,+} T^d(w_3, \mu^{w_3}), \quad g_4^{d,-} = C\mathfrak{u}_d^{d,+} T^d(w_4, \mu^{w_5}).$$

(3) *If* $\mu = \left(\frac{1}{2}(d-1), 0, -\frac{1}{2}(d-1)\right)$ *with* $d \equiv 3 \pmod 4$, *then*

$$g_2^{d,1,+} = C\mathfrak{u}_d^{d,+} T^d(w_3, \mu^{w_3}), \quad g_4^{d,-} = C\mathfrak{u}_d^{d,-} T^d(w_3, \mu^{w_3}), \quad g_4^{d,+} = C\mathfrak{u}_{d-1}^{d,+} T^d(w_4, \mu^{w_5}).$$

(4) *If* $\mu = (d-1, 0, 1-d)$, *then*

$$g_3^{d,1,\varepsilon} = C\mathfrak{u}_{d-1}^{d,\varepsilon} T^d(w_3, \mu^{w_3}), \quad g_3^{d,0,\varepsilon} = C\mathfrak{u}_d^{d,\varepsilon} T^d(w_3, \mu^{w_3}),$$

with $\varepsilon = (-1)^d$.

Note. We have suppressed the constants for purely aesthetic reasons, their values may be extracted from the computations below, but they are irrelevant in the context of the current paper. To be precise, the proof proceeds by comparing ratios of successive coefficients, and its use in Theorem 3 is such that the nonzero constant is simply included in the (scalar) Fourier–Whittaker coefficients.

We note that some care must be taken in the use of the intertwining operators when two of the complex parameters simultaneously encounter a singularity. This occurs, e.g., in case (3) of the proposition when considering $T^d(w_4, \mu^{w_5})$ since both $\mu_1 - \mu_3$ and $\mu_1 - \mu_2$ are integral, and we consider this case in particular as an example at the end of the proof.

Proof of Proposition 15. We wish to compute

$$\mathfrak{u}_d^{d,\pm} T^d(w_3, \mu^{w_3}), \quad \mathfrak{u}_{d-1}^{d,\pm} T^d(w_3, \mu^{w_3}), \quad \mathfrak{u}_d^{d,+} T^d(w_4, \mu^{w_5}), \quad \mathfrak{u}_{d-1}^{d,-} T^d(w_4, \mu^{w_5})$$

for all d and μ. We give the general procedure first, then illustrate with an example below.

Note that

$$T^d(w_4, \mu^{w_5}) = T^d(w_3, \mu^{w_5}) T^d(w_2, \mu^{w_2}),$$

and $T^d(w_2, \mu^{w_2})$ is diagonal and satisfies a symmetry relation under $(m, m) \mapsto (-m, -m)$, so it suffices to compute

$$\mathfrak{u}_d^{d,\pm} \mathcal{D}^d(w_l) \Gamma_{\mathcal{W}}^d(u, +1) \mathcal{D}^d(w_l), \quad \mathfrak{u}_{d-1}^{d,\pm} \mathcal{D}^d(w_l) \Gamma_{\mathcal{W}}^d(u, +1) \mathcal{D}^d(w_l).$$

There is a coincidence of form among the intertwining operators $T^d(w, \mu)$, the constant terms of the Eisenstein series and the Whittaker functions at a degenerate character: Superficially, this is because they are all generated by compositions (compare (13) and (I.4.12)) of analogous diagonal matrices (compare (I.2.18), (9), (I.2.20), the definition of M^d in Section I.4.4.1 and (I.3.11)). More fundamentally, composition of the constant terms of the minimal parabolic Eisenstein series, which are the Whittaker functions $\Sigma^d_{\chi_{++}} W^d(I, w, \mu, \psi_{00})$ (see (I.3.3)), gives the functional equations of the Eisenstein series, and hence also the functional equations of its nondegenerate Fourier coefficients, which are the Whittaker functions $\Sigma^d_{\chi_{++}} W^d(g, \mu, \psi_{1,1})$. This fact implies the relationship precisely, at least in the case $\chi = \chi_{++}$. We use this commonality here in the form

$$\Gamma^d_{\mathcal{W}}(u, +1) = \frac{i\,\Gamma(1+u)}{2^{1+u}\pi}\big(\exp\big(\tfrac{1}{2}i\pi u\big)\mathcal{D}^d(v_{++}) - \exp\big(-\tfrac{1}{2}i\pi u\big)\mathcal{D}^d(v_{-+})\big)\mathcal{R}^d(-i)\mathcal{W}^d(0, -u). \tag{130}$$

In fact, from (4), (3) and (11), the (diagonal) entry at row m of the right-hand side is

$$2i^m \frac{\Gamma(1+u)\Gamma(-u)}{\Gamma\big(\frac{1}{2}(1-u+m)\big)\Gamma\big(\frac{1}{2}(1-u-m)\big)} \times \begin{cases} -\sin\big(\frac{1}{2}\pi u\big) & \text{for even } m, \\ i\cos\big(\frac{1}{2}\pi u\big) & \text{for odd } m, \end{cases}$$

and the reflection formula for the gamma function shows the quotient formed by dividing this by $\Gamma^d_{\mathcal{W},m,m}(u, +1)$ is 1.

From (67) and (4), we see that conjugating a diagonal matrix by $\mathcal{D}^d(w_2)$ simply reverses the order of the diagonal entries. So using (130), (11) and the results of Section 4.7 (or by the reflection formula, see (I.2.20)), we can see

$$\mathcal{D}^d(w_2)\Gamma^d_{\mathcal{W}}(u, +1)\mathcal{D}^d(w_2) = \mathcal{D}^d(v_{-+})\Gamma^d_{\mathcal{W}}(u, +1).$$

Applying this back in (130), we have

$$\begin{aligned}\mathcal{D}^d(v_{--}w_l)\Gamma^d_{\mathcal{W}}(u, +1)\mathcal{D}^d(w_l v_{--}) \\ = \frac{i\,\Gamma(1+u)}{2^{1+u}\pi}\mathcal{R}^d(i)\big(\exp\big(\tfrac{1}{2}i\pi u\big)\mathcal{D}^d(v_{+-}) - \exp\big(-\tfrac{1}{2}i\pi u\big)\mathcal{D}^d(v_{-+})\big)F^d(u)\mathcal{R}^d(i),\end{aligned} \tag{131}$$

where

$$F^d(u) := \mathcal{D}^d(w_3)\mathcal{R}^d(-i)\mathcal{W}^d(0, -u)\mathcal{D}^d(w_3). \tag{132}$$

So we need to compute the first and last two rows of the matrix-valued function

$$F^d(u) = \int_{-\infty}^{\infty} (1+x^2)^{\frac{1}{2}(-1+u)}\mathcal{D}^d\Big(\tilde{k}\Big(1, \frac{x+i}{\sqrt{1+x^2}}, 1\Big)\Big)\,dx$$

(using $\tilde{k}(e^{i\alpha}, e^{i\beta}, e^{i\gamma}) = k(\alpha, \beta, \gamma)$ as in (I.2.5)), which has components

$$F^d_{m',m}(u) = \int_{-1}^{1} (1-x^2)^{-1-\frac{1}{2}u}\,\mathrm{d}^d_{m',m}(x)\,dx,$$

recalling (2), (3), (11) and (69). For the moment, we assume $-\operatorname{Re}(u) > 1$ so this and the subsequent integrals converge.

We use [Gradshteyn and Ryzhik 2015, equation 3.196.3] in the form

$$\int_{-1}^{1} (1-x)^{a-1}(1+x)^{b-1}\,dx = 2^{a+b-1}B(a, b), \quad \operatorname{Re}(a), \operatorname{Re}(b) > 0, \tag{133}$$

where $B(a,b)$ is again the Euler beta function. Then from (70) and (72), we have

$$F^d_{m',d}(u)=2^{-d}\pi^{\frac12}\sqrt{\frac{(2d)!}{(d+m)!\,(d-m)!}}\frac{\Gamma\big(\frac12(d-m'-u)\big)\Gamma\big(\frac12(d+m'-u)\big)}{\Gamma\big(\frac12(d-u)\big)\Gamma\big(\frac12(d+1-u)\big)}. \tag{134}$$

When $m=d-1$, (72) becomes $P_1^{(d-1-m',d-1+m')}(x)=dx-m'$, so

$$\begin{aligned}\frac{F^d_{m',d-1}(u)}{2^{1-d}\sqrt{\frac{(2d-1)!}{(d+m')!\,(d-m')!}}}&=d\int_{-1}^{1}(1-x)^{-1+\frac12(d-1-m'-u)}(1+x)^{-1+\frac12(d+1+m'-u)}\,dx\\&\quad-d\int_{-1}^{1}(1-x)^{-1+\frac12(d-1-m'-u)}(1+x)^{-1+\frac12(d-1+m'-u)}\,dx\\&\quad-m'\int_{-1}^{1}(1-x)^{-1+\frac12(d-1-m'-u)}(1+x)^{-1+\frac12(d-1+m'-u)}\,dx\\&=-m'\pi^{\frac12}\frac{\Gamma\big(\frac12(d-1-m'-u)\big)\Gamma\big(\frac12(d-1+m'-u)\big)\Gamma\big(\frac12(1-u)\big)}{\Gamma\big(\frac12(d-u)\big)\Gamma\big(\frac12(d+1-u)\big)\Gamma\big(-\frac12(u+1)\big)},\end{aligned}\tag{135}$$

using the usual recurrence relation of the gamma function.

We may deduce from the symmetries (71) with (134) and (135) that

$$\begin{aligned}F^d_{\pm d,m}(u)&=(\mp1)^{d+m}F^d_{\pm m,d}(u)=(\mp1)^{d+m}F^d_{|m|,d}(u),\\F^d_{\pm(d-1),m}(u)&=(\mp1)^{d-1+m}F^d_{\pm m,d-1}(u)=-\operatorname{sgn}(m)(\mp1)^{d+m}F^d_{|m|,d-1}(u),\end{aligned}$$

and it follows that for $\delta\in\{0,1\}$

$$\mathfrak{u}^{d,(-1)^\delta}_d F^d(u)=\tfrac12F^d_d(u)\pm(-1)^d\tfrac12F^d_{-d}(u)=2\sum_{0\le m\equiv\delta}c_mF^d_{m,d}(u)(-1)^{d+m}\mathfrak{u}^{d,(-1)^d}_m, \tag{136}$$

$$\mathfrak{u}^{d,(-1)^\delta}_{d-1} F^d(u)=2\sum_{0\le m\equiv 1-\delta}c_mF^d_{m,d-1}(u)(-1)^{d-1+m}\mathfrak{u}^{d,(-1)^{d-1}}_m, \tag{137}$$

where $c_0=\frac12,\ c_m=1,\ m\neq0$.

Now if we define $\delta,\eta\in\{0,1\}$ for a choice of the parity $\pm$ by $\pm1=(-1)^{d+\delta}=(-1)^\eta$, then (131), (68) and (136) imply

$$\begin{aligned}&\mathfrak{u}^{d,\pm}_d\mathcal{D}^d(v_{--}w_l)\Gamma^d_{\mathcal{W}}(u,+1)\mathcal{D}^d(w_lv_{--})\\&\quad=\pm(-1)^d\frac{i^{1-d}\Gamma(1+u)}{2^{1+u}\pi}\big(\exp\big(\tfrac12i\pi u\big)\mp\exp\big(-\tfrac12i\pi u\big)\big)\mathfrak{u}^{d,\pm(-1)^d}_dF^d(u)\mathcal{D}^d(w_2)\\&\quad=\frac{i^{d+\eta}}{2^{d-1}}\frac{\Gamma\big(\frac12(1+\eta+u)\big)}{\Gamma\big(\frac12(\eta-u)\big)}\sum_{0\le m\equiv\delta}c_m\mathfrak{u}^{d,\pm}_mi^{-m}\sqrt{\frac{(2d)!}{(d+m)!\,(d-m)!}}\frac{\Gamma\big(\frac12(d-m-u)\big)\Gamma\big(\frac12(d+m-u)\big)}{\Gamma\big(\frac12(d-u)\big)\Gamma\big(\frac12(d+1-u)\big)}.\end{aligned}\tag{138}$$

Here, we have used the reflection and duplication properties of the gamma function to simplify the leading coefficient, but again, this value is not actually relevant to the rest of the paper. Similarly

$$\begin{aligned}
&\mathfrak{u}_{d-1}^{d,\pm}\mathcal{D}^d(v_{--}w_l)\Gamma_{\mathcal{W}}^d(u,+1)\mathcal{D}^d(w_l v_{--})\\
&\quad=\frac{i^{d+1+\eta}}{2^{d-2}}\frac{\Gamma\left(\frac{1}{2}(1+\eta+u)\right)\Gamma\left(\frac{1}{2}(1-u)\right)}{\Gamma\left(\frac{1}{2}(\eta-u)\right)\Gamma\left(-\frac{1}{2}(u+1)\right)}\sum_{0\le m\equiv 1-\delta} c_m\mathfrak{u}_m^{d,\pm} i^{-m} m\\
&\qquad\times\sqrt{\frac{(2d-1)!}{(d+m)!\,(d-m)!}}\frac{\Gamma\left(\frac{1}{2}(d-1-m-u)\right)\Gamma\left(\frac{1}{2}(d-1+m-u)\right)}{\Gamma\left(\frac{1}{2}(d-u)\right)\Gamma\left(\frac{1}{2}(d+1-u)\right)},
\end{aligned}\tag{139}$$

and using

$$\mathfrak{u}_m^{d,\pm}\Gamma_{\mathcal{W}}^d(u,+1)=\frac{\Gamma\left(\frac{1}{2}(1-m+u)\right)}{\Gamma\left(\frac{1}{2}(1-m-u)\right)}\mathfrak{u}_m^{d,\pm(-1)^m},\quad 0\le m\le d$$

(which follows from the reflection formula, see (I.2.20)), we have

$$\begin{aligned}
&\mathfrak{u}_{d}^{d,\pm}\mathcal{D}^d(v_{--}w_l)\Gamma_{\mathcal{W}}^d(u_1,+1)\mathcal{D}^d(w_l v_{--})\Gamma_{\mathcal{W}}^d(u_2,+1)\\
&\quad=\frac{i^{d+\eta}}{2^{d-1}}\frac{\Gamma\left(\frac{1}{2}(1+\eta+u_1)\right)}{\Gamma\left(\frac{1}{2}(\eta-u_1)\right)}\sum_{0\le m\equiv \delta} c_m\mathfrak{u}_m^{d,\varepsilon} i^{-m}\\
&\qquad\times\sqrt{\frac{(2d)!}{(d+m)!\,(d-m)!}}\frac{\Gamma\left(\frac{1}{2}(d-m-u_1)\right)\Gamma\left(\frac{1}{2}(d+m-u_1)\right)}{\Gamma\left(\frac{1}{2}(d-u_1)\right)\Gamma\left(\frac{1}{2}(d+1-u_1)\right)}\frac{\Gamma\left(\frac{1}{2}(1-m+u_2)\right)}{\Gamma\left(\frac{1}{2}(1-m-u_2)\right)},
\end{aligned}\tag{140}$$

$$\begin{aligned}
&\mathfrak{u}_{d-1}^{d,\pm}\mathcal{D}^d(v_{--}w_l)\Gamma_{\mathcal{W}}^d(u_1,+1)\mathcal{D}^d(w_l v_{--})\Gamma_{\mathcal{W}}^d(u_2,+1)\\
&\quad=\frac{i^{d+1+\eta}}{2^{d-2}}\frac{\Gamma\left(\frac{1}{2}(1+\eta+u_1)\right)\Gamma\left(\frac{1}{2}(1-u_1)\right)}{\Gamma\left(\frac{1}{2}(\eta-u_1)\right)\Gamma\left(-\frac{1}{2}(u_1+1)\right)}\sum_{0\le m\equiv 1-\delta} c_m\mathfrak{u}_m^{d,-\varepsilon} i^{-m} m\\
&\qquad\times\sqrt{\frac{(2d-1)!}{(d+m)!\,(d-m)!}}\frac{\Gamma\left(\frac{1}{2}(d-1-m-u_1)\right)\Gamma\left(\frac{1}{2}(d-1+m-u_1)\right)}{\Gamma\left(\frac{1}{2}(d-u_1)\right)\Gamma\left(\frac{1}{2}(d+1-u_1)\right)}\frac{\Gamma\left(\frac{1}{2}(1-m+u_2)\right)}{\Gamma\left(\frac{1}{2}(1-m-u_2)\right)},
\end{aligned}\tag{141}$$

where $\varepsilon=(-1)^d$.

One can check that the parities $\pm$ and δ match between the preceding formulas and the claims of the theorem. Note that when $u=\mu_2-\mu_3=\frac{1}{2}(d-1)$ is an integer with $u\equiv\eta\equiv d-\delta \pmod 2$, then the coefficients of $\mathfrak{u}_m^{d,\pm}$ are zero unless $d-m-u\le 0$ in the first formula or $d-1-m-u\le 0$ in the second by the poles of the gamma functions; here we are not directly evaluating the gamma functions at a value of u, but taking the value of the whole meromorphic function at u. One may compute the ratio of the coefficient of $\mathfrak{u}_{m+2}^{d,\cdot}$ to the coefficient of $\mathfrak{u}_m^{d,\cdot}$ in (138)–(141), giving

$$-\sqrt{\frac{(d-m)(d-1-m)}{(d+2+m)(d+1+m)}}\frac{(d-u+m)}{(d-2-u-m)},\tag{142}$$

$$-\sqrt{\frac{(d-m)(d-1-m)}{(d+2+m)(d+1+m)}}\frac{(m+2)(d-1-u+m)}{m(d-3-u-m)},\tag{143}$$

$$-\sqrt{\frac{(d-m)(d-1-m)}{(d+2+m)(d+1+m)}}\frac{(d-u_1+m)(-1-m-u_2)}{(d-2-u_1-m)(-1-m+u_2)}, \tag{144}$$

$$-\sqrt{\frac{(d-m)(d-1-m)}{(d+2+m)(d+1+m)}}\frac{(m+2)(d-1-u_1+m)(-1-m-u_2)}{m(d-3-u_1-m)(-1-m+u_2)}, \tag{145}$$

respectively, with the understanding that the ratio is to be multiplied by 2 when $m=0$ (to accommodate $c_2/c_0=2$). The proposition then follows by applying the explicit form of μ in each case with $u=\mu_2-\mu_3$, $u_1=\mu_1-\mu_3$, $u_2=\mu_1-\mu_2$.

As an example, consider case (3) of the proposition where $\mu=\left(\frac{1}{2}(d-1),0,-\frac{1}{2}(d-1)\right)$ with $d\equiv 3 \pmod 4$. Using the table of μ^w given in Section I.2.1 and the definition (9), (10) and (13) of $T^d(w,\mu)$, we have

$$\begin{aligned} \mathfrak{u}^{d,+}_{d-1}T^d(w_4,\mu^{w_5}) &= \mathfrak{u}^{d,+}_{d-1}T^d(w_3,\mu^{w_5})T^d(w_2,\mu^{w_2}) \\ &= \pi^{-\frac{3}{2}(d-1)}\mathfrak{u}^{d,+}_{d-1}\mathcal{D}^d(v_{--}w_l)\Gamma^d_{\mathcal{W}}(d-1,+1)\mathcal{D}^d(w_l v_{--})\Gamma^d_{\mathcal{W}}\left(\tfrac{1}{2}(d-1),+1\right). \end{aligned}$$

The coefficients of $\mathfrak{u}^{d,+}_m$ in $g^{d,+}_4$ (recall (113)) are supported on $\frac{1}{2}(d+1)\le m\equiv\frac{1}{2}(d+1)\equiv 0 \pmod 2$, and we compare this to (141) with $u_1=d-1$, $u_2=\frac{1}{2}(d-1)$ (or rather, the analytic continuation to this point), $\delta=1$, $\eta=0$ and $\varepsilon=-1$. As mentioned in the previous paragraph, (141) has a removable singularity at this (u_1,u_2)-point, and the summand is zero unless $m\ge\frac{1}{2}(d+1)$, so the support of the coefficients matches that of $g^{d,+}_4$. The ratio of successive coefficients (145) reduces to

$$\sqrt{\frac{(d-m)(d-1-m)}{(d+2+m)(d+1+m)}}\frac{\left(m+\frac{1}{2}(d+1)\right)}{\left(m-\frac{1}{2}(d-3)\right)};$$

this matches the ratio $g^d_{2,\kappa,2j+\kappa+2}/g^d_{2,\kappa,2j+\kappa}$ with $m=2j+\kappa$. So we conclude $g^{d,+}_4$ and $\mathfrak{u}^{d,+}_{d-1}T^d(w_4,\mu^{w_5})$ are the same, up to a nonzero constant, and since $g^d_{2,\kappa,0}=1$, we conclude the value of the constant is the coefficient of $\mathfrak{u}^{d,+}_m$ in $\mathfrak{u}^{d,+}_{d-1}T^d(w_4,\mu^{w_5})$ at $m=\kappa$, which is

$$C=\pi^{-\frac{3}{2}(d-1)}2^{3-2d}d!\sqrt{\frac{(2d-1)!\,(\kappa-1)!}{(3\kappa-1)!}}.$$

Here we have written the coefficient $C=C(0,0)$ of $\mathfrak{u}^{d,+}_\kappa$ in (141) by taking $u_1=d-1+a_1$ and $u_2=\frac{1}{2}(d-1)+a_2$ and applying the reflection formula for the gamma functions (keeping in mind $d\equiv 3 \pmod 4$) so that

$$C(a_1,a_2)=\frac{2^{1-2d-2a_1}(d+1)}{\pi^{\frac{3}{2}(d-1)+a_1+a_2}}\sqrt{\frac{(2d-1)!}{(\kappa-1)!\,(3\kappa-1)!}}\frac{\Gamma\left(\frac{1}{2}(\kappa-a_1)\right)\Gamma\left(\frac{1}{2}(d+1+a_2)\right)\Gamma\left(d+1+a_1\right)}{\Gamma(1-a_1)\Gamma\left(\frac{1}{2}(2-b)\right)\Gamma\left(\frac{1}{2}(\kappa+2+2a)\right)},$$

and this holds in a neighborhood of $(a_1,a_2)=(0,0)$, as desired. Of course, $\mathfrak{u}^{d,+}_{d-1}T^d(w_4,\mu^{w_5})/C(a_1,a_2)$ may be defined in terms of the ratios (145) (including the terms with $m<\kappa$), and in this way our expression for $\mathfrak{u}^{d,+}_{d-1}T^d(w_4,\mu^{w_5})$ is holomorphic in a neighborhood of $(a_1,a_2)=(0,0)$. □

We must clarify precisely when the Whittaker function is identically zero, and we begin with the minimal K-types, as described in Proposition 14. The following proposition follows from Theorem 5 and the results of Section 7.2:

Proposition 16. *Suppose f is power-function minimal for μ at weight d:*

(1) *If $d=0$, then $fW^d(\cdot,\mu,\psi_{1,1})$ is identically zero if and only if $f=0$.*

(2) *If $d=1$ and all μ_i are distinct, then $fW^d(\cdot,\mu,\psi_{1,1})$ is identically zero if and only if $f=0$.*

(3) *If $\mu=\left(\frac{1}{2}(d-1)+it,-2it,-\frac{1}{2}(d-1)+it\right)$ with $d\geq 1$, and either d even or $t\neq 0$, then f is a linear combination of the vectors $\mathfrak{v}^d_{\pm d}T^d(w_3,\mu^{w_3})$, and $fW^d(\cdot,\mu,\psi_{1,1})$ is identically zero if and only if f is a multiple of $\mathfrak{v}^d_d T^d(w_3,\mu^{w_3})$.*

(4) *If $\mu=\left(\frac{1}{2}(d-1),0,-\frac{1}{2}(d-1)\right)$ with d odd, then f is a linear combination of the vectors $g_4^{d,+}$, $g_4^{d,-}$ and $(g_4^{d,+}+g_4^{d,-})\mathcal{D}^d(v_{--}w_l)$, and $fW^d(\cdot,\mu,\psi_{1,1})$ is identically zero if and only if f is a linear combination of $g_4^{d,+}+g_4^{d,-}$ and $(g_4^{d,+}+g_4^{d,-})\mathcal{D}^d(v_{--}w_l)$.*

(5) *If $\mu=(d-1,0,1-d)$ and $d\geq 1$, then f a linear combination of the four vectors $\mathfrak{v}^d_{\pm d}$ and $\mathfrak{v}^d_{\pm d}\mathcal{D}^d(v_{--}w_l)$, and $fW^d(\cdot,\mu,\psi_{1,1})$ is identically zero if and only if f is a linear combination of $\mathfrak{v}^d_d$ and $\mathfrak{v}^d_d\mathcal{D}^d(v_{--}w_l)$.*

Note that in the case $d=1$, $g_4^{\pm}=\mathfrak{u}_1^{1,\pm}$, so there is no inconsistency between (4) and (5) in that case.

Proof of Proposition 16. To see case (2) we note that the eigenvalues under Y^0 are

$$Y^0_\mu\mathfrak{u}_0^{1,-}=-2\sqrt{\tfrac{3}{5}}\mu_3\mathfrak{u}_0^{1,-},\quad Y^0_\mu\mathfrak{u}_1^{1,-}=-2\sqrt{\tfrac{3}{5}}\mu_2\mathfrak{u}_1^{1,-},\quad Y^0_\mu\mathfrak{u}_1^{1,+}=-2\sqrt{\tfrac{3}{5}}\mu_1\mathfrak{u}_1^{1,+},\tag{146}$$

and we may see directly from Theorem 5 that each of the associated Whittaker functions is nonzero.

In case (3), we note that

$$\mathfrak{v}^d_{-d}T^d(w_3,\mu^{w_3})W^d(g,\mu,\psi_{1,1})\neq 0,\quad \mathfrak{v}^d_d T^d(w_3,\mu^{w_3})W^d(g,\mu,\psi_{1,1})=0,$$

and by Proposition 15, those two vectors are a basis of the required space.

In case (4), we note that

$$\mathfrak{v}^d_{-d}T^d(w_3,\mu^{w_3})W^d(\cdot,\mu,\psi_{1,1})\neq 0,$$
$$(g_4^{d,+}+g_4^{d,-})W^d(g,\mu,\psi_{1,1})=0.$$

The second equality follows because the vector $g_4^{d,+}+g_4^{d,-}$ is supported on $\mathfrak{v}^d_m$ with

$$m\equiv\kappa\equiv 1+\mu_1-\mu_2 \pmod 2,\qquad m\geq\kappa=1+\mu_1+\mu_2,$$

and again the gamma function in (11) has a pole there (recall the discussion at the end of Section 7.2). Then by duality, i.e., (98) and (99),

$$(g_4^{d,+}+g_4^{d,-})\mathcal{D}^d(v_{--}w_l)W^d(g,\mu,\psi_{1,1})=(g_4^{d,+}+g_4^{d,-})W^d(v_{--}g^\iota w_l,\mu,\psi_{1,1})=0,$$

and $(g_4^{d,+}+g_4^{d,-})\mathcal{D}^d(v_{--}w_l)$ is still power-function minimal. Again, these are three linearly independent vectors (they have different parities, i.e., live in the row space of Σ_χ^d for different characters χ) in a three-dimensional space.

In case (5), we have

$$\mathfrak{v}_{-d}^d W^d(\cdot,\mu,\psi_{1,1})\neq 0,\quad \mathfrak{v}_d^d W^d(\cdot,\mu,\psi_{1,1})=0,$$

and by duality

$$\mathfrak{v}_{-d}^d\mathcal{D}^d(v_{--}w_l)W^d(\cdot,\mu,\psi_{1,1})\neq 0,\quad \mathfrak{v}_d^d\mathcal{D}^d(v_{--}w_l)W^d(\cdot,\mu,\psi_{1,1})=0.$$

It is easy to see that the vectors $\mathfrak{v}_{-d}^d$ and $\mathfrak{v}_{-d}^d\mathcal{D}^d(v_{--}w_l)$ are linearly independent; however, we require something a bit stronger: we need to know that any nonzero linear combination

$$(a_1\mathfrak{v}_{-d}^d+a_2\mathfrak{v}_{-d}^d\mathcal{D}^d(v_{--}w_l))W^d(\cdot,\mu,\psi_{1,1}) \tag{147}$$

with $a_1a_2\neq 0$ yields a function that is not identically zero. Notice that

$$\sqrt{(d+1)(2d+3)}Y_\mu^0\mathfrak{v}_{-d}^d=-(d-1)\sqrt{6d(2d-1)}\mathfrak{v}_{-d}^d, \tag{148}$$

and by duality $\mathfrak{v}_{-d}^d\mathcal{D}^d(v_{--}w_l)$ is also an eigenfunction of Y_μ^0, but its eigenvalue has the opposite sign, and this is sufficient for our purposes. (Consider applying the operators $\sqrt{(d+1)(2d+3)}Y^0\pm(d-1)\sqrt{6d(2d-1)}$ to (147).) □

Proposition 17. *If $f\in\mathbb{C}^{2d+1}$ is such that*

$$fW^d(g,\mu,\psi_{1,1})\neq 0,\quad Y^{-1}fW^d(g,\mu,\psi_{1,1})=0,\quad Y^{-2}fW^d(g,\mu,\psi_{1,1})=0, \tag{149}$$

subject to (107), then f is power-function minimal.

Proof. The proposition is trivial by Lemma 12 when none of the differences $\mu_i-\mu_j$, $i\neq j$, are integers, and the case $\mu=\left(\frac{1}{2}(\kappa-1)+it,-2it,-\frac{1}{2}(\kappa-1)+it\right)$ for some $\kappa\geq 1$ with κ even or $t\neq 0$, is also relatively simple, after applying the w_3 functional equation of the Whittaker functions (as we may). So suppose $\mu=(\kappa-1,0,1-\kappa)$ for some $\kappa\geq 1$.

For any d, define the spaces

$$\begin{aligned} V_\delta^d&=\operatorname{span}\{\mathfrak{v}_j^d\mid j\equiv\delta\pmod 2\},\\ V_{\delta,\kappa}^d&=\operatorname{span}\{\mathfrak{v}_j^d\mid |j|\geq\kappa,\ j\equiv\delta\pmod 2\},\\ V_{\delta,\kappa,\pm}^d&=\operatorname{span}\{\mathfrak{v}_j^d\mid \pm j\geq\kappa,\ j\equiv\delta\pmod 2\}.\end{aligned}$$

It follows immediately from the argument of Proposition 14 that for f which is not already power-function minimal,

$$f\in V_\kappa^d,\quad Y_\mu^{-1}f\in V_{\kappa,\kappa}^{d-1},\quad Y_\mu^{-2}f\in V_{\kappa,\kappa}^{d-2}\implies f\in V_{\kappa,\kappa}^d, \tag{150}$$

and noticing that the Y_μ^a operators act on $\mathfrak{v}_j^d$ precisely the same as on $\mathfrak{u}_j^{d,\pm}$ for $j\geq 3$, we also have

$$f\in V_\kappa^d,\quad Y_\mu^{-1}f\in V_{\kappa,\kappa,\pm}^{d-1},\quad Y_\mu^{-2}f\in V_{\kappa,\kappa,\pm}^{d-2}\implies f\in V_{\kappa,\kappa,\pm}^d. \tag{151}$$

We know that $\mathfrak{v}_\kappa^\kappa W^\kappa(\cdot, \mu, \psi_{1,1}) = 0$, and the argument of Section 9 applied to $\mathfrak{v}_\kappa^\kappa$ (in place of $\mathfrak{u}_\kappa^{\kappa,\pm}$) implies that $vW^d(\cdot, \mu, \psi_{1,1}) = 0$ for all $v \in V_{\kappa,\kappa,+}^d$ for all d. Similarly, since $V_{\kappa,\kappa,-}^d$ is closed under the Y_μ^a operators (because $\mu_1 - \mu_2 + 1 - \kappa = 0$ in (89)–(93)), and

$$\mathfrak{v}_{-\kappa}^\kappa W^\kappa(\cdot, \mu, \psi_{1,1}) \neq 0, \quad (g_4^{2\kappa-1,+} - g_4^{2\kappa-1,-}) W^{2\kappa-1}(\cdot, \mu, \psi_{1,1}) \neq 0,$$

Proposition 14 implies $vW^d(\cdot, \mu, \psi_{1,1}) \neq 0$ for all $v \in V_{\kappa,\kappa}^d$, $v \notin V_{\kappa,\kappa,+}^d$ for all d. Even stronger, by (150), Proposition 14, and Proposition 16 at $d = 0, 1$, we know $vW^d(\cdot, \mu, \psi_{1,1}) \neq 0$ for all $v \in V_\kappa^d$, $v \notin V_{\kappa,\kappa,+}^d$ for all d.

Suppose

$$Y^{-1} f W^d(g, \mu, \psi_{1,1}) = 0, \quad Y^{-2} f W^d(g, \mu, \psi_{1,1}) = 0,$$

and f is not power-function minimal. Set $\varepsilon = (-1)^\kappa$ and $\delta \in \{0, 1\}$, $\delta \equiv \kappa \pmod 2$. Since the zero function cannot descend to a nonzero Whittaker function, f must descend, via Y_μ^{-1} and Y_μ^{-2} to a nonzero linear combination of either

$$g_4^{2\kappa-1,+} + g_4^{2\kappa-1,-} \quad \text{and} \quad (g_4^{2\kappa-1,+} + g_4^{2\kappa-1,-})\mathcal{D}^d(v_{--}w_l),$$

or $\mathfrak{v}_\kappa^\kappa$ and $\mathfrak{v}_\kappa^\kappa \mathcal{D}^d(v_{--}w_l)$.

Set

$$h = \sum_{\substack{j \geq \kappa \\ j \equiv \kappa \ (\mathrm{mod}\ 2)}} \tfrac{1}{2}(f_j + (-1)^{d-\kappa} f_{-j})\mathfrak{v}_j^d \in V_{\kappa,\kappa,+}^d;$$

then $hW^d(g, \mu, \psi_{1,1})$ is zero. Subtracting h removes the projection onto $\chi_{\varepsilon,\varepsilon}$, i.e., $(f - h)\Sigma_{\varepsilon,\varepsilon}^d = 0$, and so by the above arguments, $f - h$ must descend to a multiple of either

$$(g_4^{2\kappa-1,+} + g_4^{2\kappa-1,-})\mathcal{D}^d(v_{--}w_l) \quad \text{or} \quad \mathfrak{v}_\kappa^\kappa \mathcal{D}^d(v_{--}w_l).$$

Now we switch to the dual Whittaker function (note $-\mu^{w_l} = \mu$): Set $\tilde{f} = (f - h)\mathcal{D}^d(v_{--}w_l)$; then $\tilde{f}$ must descend to a multiple of either $g_4^{2\kappa-1,+} + g_4^{2\kappa-1,-} \in V_{\kappa,\kappa,+}^{2\kappa-1}$ or $\mathfrak{v}_\kappa^\kappa \in V_{\kappa,\kappa,+}^\kappa$. As before, we set

$$\tilde{h} = \sum_{\substack{j \geq \kappa \\ j \equiv \kappa \ (\mathrm{mod}\ 2)}} \tfrac{1}{2}(\tilde{f}_j - (-1)^{d-\kappa} \tilde{f}_{-j})\mathfrak{v}_j^d \in V_{\kappa,\kappa,+}^d.$$

Now consider the vector $v = f - h - \tilde{h}\mathcal{D}^d(v_{--}w_l)$ which has the projections $v\Sigma_{\varepsilon,\varepsilon}^d = v\Sigma_{-\varepsilon,\varepsilon}^d = 0$. The projection of v onto V_κ^d is not contained in $V_{\kappa,\kappa,+}^d$ unless it is zero, and the same is true for the projection of $v\mathcal{D}^d(v_{--}w_l)$. We still have

$$Y^{-1} vW^d(g, \mu, \psi_{1,1}) = 0, \quad Y^{-2} vW^d(g, \mu, \psi_{1,1}) = 0,$$

and applying Proposition 14 to v, we conclude $v = 0$ because any nonzero minimal descendant could only meet the conclusions (4)–(6), and we have constructed v so this is impossible. Therefore,

$$f W^d(g, \mu, \psi_{1,1}) = (h + \tilde{h}\mathcal{D}^d(v_{--}w_l))W^d(g, \mu, \psi_{1,1}) = 0. \qquad \square$$

Using the bases for the spaces of power-function minimal vectors given in Proposition 16, together with Propositions 14 and 15, and Theorem 5 gives:

Corollary 18. *Suppose $d \geq 2$, μ and $f \in \mathbb{C}^{2d+1}$ are such that (149) holds and further that $f W^d(g, \mu, \psi_{1,1})$ is an eigenfunction of Y^0 if $d = 2$. Then $f W^d(g, \mu, \psi_{1,1}) = C f' W^d(g, \mu', \psi_{1,1})$, for some nonzero constant C, where*

(1) $\mu' = \left(\frac{1}{2}(d-1)+it, -\frac{1}{2}(d-1)+it, -2it\right)$ *with* $f' = \mathfrak{u}_d^{d,+}$, *or*

(2) $\mu' = (d-1, 0, 1-d)$ *with f' some linear combination of* $\mathfrak{v}_{-d}^d$ *and* $\mathfrak{v}_{-d}^d \mathcal{D}^d(v_{--} w_l)$.

This follows, for example, when $\mu = \left(\frac{1}{2}(d-1), 0, -\frac{1}{2}(d-1)\right)$ with $d \equiv 3 \pmod 4$ by writing

$$f = a_1 g_2^{d,0,-} + a_2 g_4^{d,+} + a_3 g_4^{d,-} = (a_1 C_1 \mathfrak{u}_d^{d,-} + (a_2 - a_3) C_2 \mathfrak{u}_d^{d,+}) T^d(w_3, \mu^{w_3}) + a_3 (g_4^{d,+} + g_4^{d,-}),$$

and using (8) and

$$\mathfrak{v}_d^d W^d(g, \mu^{w_3}, \psi_{1,1}) = (g_4^{d,+} + g_4^{d,-}) W^d(g, \mu, \psi_{1,1}) = 0.$$

In the case $d = 2$, the additional assumption about the behavior under Y^0 rules out the false positive posed by g_1, as in the discussion following Proposition 14. As mentioned in Section 7.4, the second case will not occur as the Whittaker function of a cusp form.

8.3. ***For cusp forms.*** We may now prove Theorem 3. As mentioned in the discussion preceding the theorem, we take condition (2) of Proposition 2 as our working definition of minimal-weight forms. Suppose ϕ has minimal weight d and spectral parameters μ.

The n-th Fourier coefficient of ϕ is of the form $f W^d(\cdot, \mu, \psi_n)$, $f \in \mathbb{C}^{2d+1}$, by Theorem 1(3) (and the discussion of Section 4.3), and Proposition 17 gives the allowed values of d, f and μ. If $d = 0$, there is nothing to do, and if $d \geq 2$, we apply Corollary 18 to arrive at the parameter set described in Theorem 3. There is a minor caveat that this corollary is given in terms of the character $\psi_{1,1}$ and not ψ_n, but this is readily fixed: if $n \in \mathbb{Z}^2$, $n_1 n_2 \neq 0$, we may use the isomorphism $n \in (\mathbb{R}^\times)^2 \cong V Y^+$ to write $n = v\tilde{n}$ with $v \in V$, and $\tilde{n} \in Y^+$, then using (I.3.7) and (I.3.9), we have

$$W^d(g, \mu, \psi_n) = p_{-\rho-\mu^{w_l}}(\tilde{n}) \mathcal{D}^d(w_l v w_l) W^d(ng, \mu, \psi_{1,1}),$$

and (65) shows the extra matrix $\mathcal{D}^d(w_l v w_l)$ at worst alters the sign of the coefficient.

A small bit more needs to be said about the cases $\mu = (d-1, 0, 1-d)$ and $d = 1$. When $\mu = (d-1, 0, 1-d)$, we have already pointed out in Section 7.4 that the asymptotics of $W_{-d}^d(g, \mu, \psi_n)$ are not compatible with the boundedness of cusp forms. Another way to see these don't occur is to note that such a cusp form would have a real eigenvalue for the skew-symmetric operator Y^0 (recall (148) and the following discussion).

Now suppose $d = 1$ and μ is arbitrary. The eigenvalues under Y^0 are given in (146); if the components of μ are distinct, these choices of f give three linearly independent Whittaker functions. In the case $\mu = (x+it, -2it, -x+it)$, $x \neq 0$, only $\mathfrak{u}_1^{1,-}$ may give the Whittaker function of a cusp form since the others are again eigenfunctions of a skew-symmetric operator with eigenvalues that are not purely imaginary. Further, the functional equations of the Whittaker function yield

$$\mathrm{u}_1^{1,-}T(w_3,\mu^{w_3})=C\mathrm{u}_0^{1,-},$$
$$\mathrm{u}_1^{1,+}T(w_5,\mu^{w_4})=C\mathrm{u}_0^{1,-}.$$

So it suffices to take $\mathrm{u}_0^{1,-}W^1(y,\mu,\psi_{1,1})$ to be the $d=1$ Whittaker function with either $\mu=(it_1,it_2,-i(t_1+t_2))$ or $\mu=(x+it,-x+it,-2it)$.

9. Going up

Throughout this section, we fix a triple of spectral parameters μ and a character $\chi=\chi_{(-1)^\delta,\varepsilon}$ with its two parities $\delta\in\{0,1\}$ and $\varepsilon=\pm1$, as in (5). For any $0\le\kappa\in\mathbb{Z}$, let

$$\mathcal{V}^d_{\kappa,\chi}=\operatorname{span}_{j\ge\kappa}\mathrm{u}^{d,\varepsilon}_{2j+\delta}\subset\mathbb{C}^{2d+1}.$$

As in the previous section, we may safely assume $-\bar{\mu}$ is a permutation of μ, i.e., that μ is a permutation of either

$$(it_1,it_2,-i(t_1+t_2))\quad\text{or}\quad(x+it,-x+it,-2it),\qquad t_i,x,t\in\mathbb{R},\ t_1>t_2>-t_1-t_2,\ x\ge0.$$

We choose the trivial permutation. In the case $\mu_1-\mu_2+1\in(2\mathbb{Z}+\delta)$, we define $\kappa=\mu_1-\mu_2+1$, and otherwise we set $\kappa=0$. For convenience, we define $d_0=\kappa$ when $\kappa>0$ and

$$d_0=\begin{cases}1 & \text{if } \delta=1 \text{ or } \varepsilon=-1,\\ 0 & \text{otherwise,}\end{cases}$$

when $\kappa=0$. Lastly and for this section only, we apply the shorthand $\mathrm{u}^d_j=\mathrm{u}^{d,\varepsilon}_j$ and $\mathcal{V}^d=\mathcal{V}^d_{\kappa,\chi}$.

For each dimension/weight d, we have a "minimal-weight vector"

$$\mathrm{u}^d_{\min}=\mathrm{u}^d_{j_{\min}},\qquad j_{\min}=\begin{cases}2 & \text{if } \kappa=\delta=0,\ d\not\equiv d_0 \pmod 2,\\ 1 & \text{if } \delta=1 \text{ and } \kappa=0,\\ \kappa & \text{otherwise.}\end{cases}$$

We wish to show for $d\ge d_0$ that $\mathcal{V}^d$ can be generated by applying the Y^a_μ operators to $\mathrm{u}^{d_0}_{\min}$. We accomplish this through induction by showing that suitable combinations of operators applied to $\mathrm{u}^d_{\min}$ will give $\mathrm{u}^{d+1}_{\min}$ or $\mathrm{u}^{d+2}_{\min}$; we call this the $d\to d+1$ or $d\to d+2$ step. We can then fill out the remainder of $\mathcal{V}^{d+1}$ or $\mathcal{V}^{d+2}$ by repeatedly applying (89). Precisely, we show:

Proposition 19.
$$\mathbb{C}[Y^0_\mu,Y^1_\mu,Y^2_\mu]\mathrm{u}^{d_0}_{\min}=\bigcup_{d\ge d_0}\mathcal{V}^d,$$

where $\mathbb{C}[Y^0_\mu,Y^1_\mu,Y^2_\mu]$ is the complex algebra generated by Y^0_μ,Y^1_μ,Y^2_μ under composition, and the Y^a_μ operators are viewed in the sense of (86).

All of the raising operators described in the induction argument below follow from either (90) and (92) directly or from the following two linear combinations: set

$$\begin{aligned} R^{d,1}_{\mu,j} &= \sqrt{d(d+2)^2(2d+1)^2(2d+5)}\,Y^0_\mu Y^1_\mu \\ &\quad + 2\sqrt{6d(d+1)(d+2)(2d+1)}(d-1-2j-2\mu_3)Y^1_\mu \\ &\quad - \sqrt{d^2(d+2)(2d-1)(2d+1)(2d+3)}\,Y^1_\mu Y^0_\mu, \end{aligned} \tag{152}$$

$$\begin{aligned} R^{d,2}_{\mu,j} &= \sqrt{(d+1)(d+2)^2(d+3)(2d+1)(2d+7)}\,Y^0_\mu Y^2_\mu \\ &\quad + 2\sqrt{6(d+1)(d+2)(2d+1)(2d+3)}(2d+1-2j-2\mu_3)Y^2_\mu \\ &\quad - \sqrt{d(d+1)^2(d+2)(2d-1)(2d+1)}\,Y^2_\mu Y^0_\mu; \end{aligned} \tag{153}$$

then

$$\begin{aligned} R^{d,1}_{\mu,j}\mathrm{u}^d_j &= 8j\sqrt{3(d+1-j)(d+2-j)(d+3-j)(d+j)}(\mu_1-\mu_2+1-j)\mathrm{u}^{d+1}_{j-2} \\ &\quad - 8j\sqrt{3(d+1-j)(d+1+j)}\big((d-j-2)(j+1+3\mu_3) \\ &\qquad - 2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1)+4(j+1)\big)\mathrm{u}^{d+1}_j, \end{aligned} \tag{154}$$

$$\begin{aligned} R^{d,2}_{\mu,j}\mathrm{u}^d_j &= -4j\sqrt{6(d+1-j)(d+2-j)(d+3-j)(d+4-j)}(\mu_1-\mu_2+1-j)\mathrm{u}^{d+2}_{j-2} \\ &\quad - 4\sqrt{6(d+1-j)(d+2-j)(d+1+j)(d+2+j)}\big((2d-j)(d+2-3\mu_3) \\ &\qquad + 2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1)-j(d+1-j)\big)\mathrm{u}^{d+2}_j. \end{aligned} \tag{155}$$

In both operators, it is not too hard to see that the constant multiplying u^{d+a}_j is nonzero, but we will show a stronger statement about the norms of the new vectors. These become true raising operators because in the cases we use them one of the following is true:

(1) $j=0$.

(2) $\mathrm{u}^{d+a}_{j-2}=0$.

(3) $\mu_1-\mu_2+1-j=0$.

(4) $j=1$ (so $\mathrm{u}^{d+a}_{j-2}=\pm\mathrm{u}^{d+a}_j$).

Note that this is where the argument would fail if we attempted to use the highest-weight vector.

Proposition 20. *Suppose we have a sequence of sesquilinear forms $\langle\cdot,\cdot\rangle$ on $\mathcal{V}^d\times\mathcal{V}^d_{0,\chi}$, $d\ge d_0$, that satisfy*

$$\langle su,tv\rangle=s\bar{t}\langle u,v\rangle,\ s,t\in\mathbb{C},\quad \langle u,v\rangle=\langle u,\mathrm{proj}_{\mathcal{V}^d}v\rangle,\quad \langle Y^a_\mu u,v\rangle=\langle u,\widehat{Y}^a_{-\bar\mu}v\rangle, \tag{156}$$

and that, for $d=d_0$, we have

$$\langle \mathrm{u}^d_{2i+\delta},\mathrm{u}^d_{2i'+\delta}\rangle=\mathrm{u}^d_{2i+\delta}(\mathrm{u}^d_{2i'+\delta})^T,\quad 2i+\delta,2i'+\delta\ge\kappa. \tag{157}$$

Then (157) continues to hold for all $d\ge d_0$.

Remarks. (1) By construction $\dim \mathcal{V}^{d_0} = 1$, so the assumption that (157) holds for $d = d_0$ can be reduced to just

$$\langle \mathrm{u}^{d_0}_{\min}, \mathrm{u}^{d_0}_{\min} \rangle = \begin{cases} 1 & \text{if } \kappa = \delta = 0,\ \varepsilon = +1, \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

(2) The projection assumption $\langle u, v\rangle = \langle u, \mathrm{proj}_{\mathcal{V}^d} v\rangle$ is necessary since the $Y^a_{-\bar{\mu}}$ operators do not respect the spaces $\mathcal{V}^d$, even though the Y^a_μ do. In practice, this assumption is met since the irksome rows of the incomplete Whittaker function $W^d(g, -\bar{\mu}, \psi)$ on the right-hand side of our inner product on Maass forms will be zero.

(3) The actual sequence of sesquilinear forms we will use is given by the left-hand side of (167); that is,

$$\langle v, v'\rangle = \int_{\Gamma\backslash G} (v\Phi^d(g))\overline{(v'\widetilde{\Phi}^d(g))^T}\, dg, \qquad v \in \mathcal{V}^d,\ v' \in \mathcal{V}^d_{0,\chi}, \tag{158}$$

where Φ^d and $\widetilde{\Phi}^d$ are constructed from the Fourier expansion of a single minimal-weight form as in (27) and the comment that follows.

Proof of Propositions 19 and 20. With respect to the sesquilinear forms of Proposition 20, the adjoints of the operators $R^{d,a}_{\mu,j}$ act on the particular vectors u^{d+a}_j as

$$\begin{aligned}\overline{\widehat{R}^{d,1}_{\mu,j}\mathrm{u}^{d+1}_j} &= -8\sqrt{3(d+2-j)(d-1+j)(d+j)(d+1+j)}(\mu_1-\mu_2-1+j)\mathrm{u}^d_{j-2} \\ &\quad - 8j\sqrt{3(d+1-j)(d+1+j)} \\ &\qquad \times \big((d-j-2)(j+1+3\mu_3) - 2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1) + 4(j+1)\big)\mathrm{u}^d_j \\ &\quad + 8(j+1)\sqrt{3(d-1-j)(d-j)(d+1-j)(d+2+j)}(\mu_1-\mu_2-1-j)\mathrm{u}^d_{j+2},\end{aligned} \tag{159}$$

$$\begin{aligned}\overline{\widehat{R}^{d,2}_{\mu,j}\mathrm{u}^{d+2}_j} &= -4\sqrt{6(d-1+j)(d+j)(d+1+j)(d+2+j)}(\mu_1-\mu_2-1+j)\mathrm{u}^d_{j-2} \\ &\quad - 4\sqrt{6(d+1-j)(d+2-j)(d+1+j)(d+2+j)} \\ &\qquad \times \big((2d-j)(d+2-3\mu_3) + 2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1) - j(d+1-j)\big)\mathrm{u}^d_j \\ &\quad - 4(1+j)\sqrt{6(d-1-j)(d-j)(d+1-j)(d+2-j)}(\mu_1-\mu_2-1-j)\mathrm{u}^d_{j+2}\end{aligned} \tag{160}$$

Note the coefficient on u^d_j matches those on u^{d+a}_j in (154) and (155). It follows that if $\langle \mathrm{u}^d_j, \mathrm{u}^d_{j+2}\rangle = 0$ (which will be the induction assumption), then

$$\begin{aligned} &j\sqrt{(d+1-j)(d+1+j)}\big((d-j-2)(j+1+3\mu_3) - 2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1) \\ &\qquad\qquad + 4(j+1)\big)(\langle \mathrm{u}^{d+1}_j, \mathrm{u}^{d+1}_j\rangle - \langle \mathrm{u}^d_j, \mathrm{u}^d_j\rangle), \\ &\quad = \sqrt{(d+2-j)(d+j)(d-1+j)(d+1+j)}(\mu_1-\mu_2-1+j)\langle \mathrm{u}^d_j, \mathrm{u}^d_{j-2}\rangle \\ &\qquad + j\sqrt{(d+2-j)(d+j)(d+1-j)(d+3-j)}(\mu_1-\mu_2+1-j)\langle \mathrm{u}^{d+1}_{j-2}, \mathrm{u}^{d+1}_j\rangle, \end{aligned} \tag{161}$$

$$\sqrt{(d+1-j)(d+2-j)(d+1+j)(d+2+j)}\big((2d-j)(d+2-3\mu_3) \\ +2(\mu_1-\mu_3+1)(\mu_2-\mu_3+1)-j(d+1-j)\big)(\langle \mathrm{u}_j^{d+2}, \mathrm{u}_j^{d+2}\rangle - \langle \mathrm{u}_j^d, \mathrm{u}_j^d\rangle) \\ = \sqrt{(d-1+j)(d+j)(d+1+j)(d+2+j)}(\mu_1-\mu_2-1+j)\langle \mathrm{u}_j^d, \mathrm{u}_{j-2}^d\rangle \\ -j\sqrt{(d+1-j)(d+2-j)(d+3-j)(d+4-j)}(\mu_1-\mu_2+1-j)\langle \mathrm{u}_{j-2}^{d+2}, \mathrm{u}_j^{d+2}\rangle. \quad (162)$$

We wish to show for $d \geq d_0$ that $\mathcal{V}^d$ is in the image of the raising operators, i.e.,

$$\mathcal{V}^d = \mathbb{C}[Y_\mu^0]\begin{cases} (\mathbb{C}Y_\mu^1\mathcal{V}^{d-1} + \mathbb{C}Y_\mu^2\mathcal{V}^{d-2}) & \text{if } d \geq d_0+2, \\ Y_\mu^1\mathcal{V}^{d-1} & \text{if } d = d_0+1, \\ \mathrm{u}_{\min}^d & \text{if } d = d_0, \end{cases}$$

and that (157) continues to hold in the higher weight. We prove this in two steps: First, we show that $\mathrm{u}_{\min}^d$ is in the image of the raising operators and (157) holds for $\mathrm{u}_{2i+\delta}^d = \mathrm{u}_{2i'+\delta}^d = \mathrm{u}_{\min}^d$ for all $d \geq d_0$ (the base case of the double induction). Second, we extend this to all of $\mathcal{V}^d$ for all $d \geq d_0$ (the induction step of the double induction).

The first step, itself an induction argument on d, proceeds by cases. Note that the base case of the induction is the statement that $\mathrm{u}_{\min}^{d_0}$ itself is in the image of $\mathrm{u}_{\min}^{d_0}$ under the raising operators, i.e., elements of $\mathbb{C}[Y_\mu^0, Y_\mu^1, Y_\mu^2]$, but this is obvious. The cases are:

case	conditions	step
Ia	$\kappa > 1$	$d_0 \to d_0+1$
Ib	$\kappa > 1$	$d \to d+2$
IIa	$\kappa = \delta = 0,\ \varepsilon = +1$	$0 \to 2$
IIb	$\kappa = \delta = 0,\ \varepsilon = +1$	$2 \to 3$
IIc	$\kappa = \delta = 0,\ \varepsilon = +1$	$d \to d+2$
IIIa	$\max\{\kappa, \delta\} = 1,\ \varepsilon = (-1)^d$	$d \to d+1$
IIIb	$\max\{\kappa, \delta\} = 1,\ \varepsilon = -(-1)^d$	$d \to d+1$

For the cases incrementing d by 1, we use the raising operator (154) and apply (161) for the orthonormality. For the cases incrementing d by 2, we use the raising operator (155) and apply (162) for the orthonormality. The reasons for the separation of cases are: firstly, when $j_{\min} = 1$, we have $\mathrm{u}_{j_{\min}-2}^d = \varepsilon(-1)^d \mathrm{u}_{j_{\min}}^d$, so the form of the raising operator changes (slightly) with the parity of d; secondly, if $j_{\min} = 2$ for some d, then $j_{\min} = 0$ for $d+1$, and the raising operator cannot lift from $j = 2$ at d to $j = 0$ at $d+1$ (but lifting from $j = 2$ at d to $j = 2$ at $d+2$ is fine); lastly $\dim \mathcal{V}^1 = 0$ when $\kappa = \delta = 0, \varepsilon = +1$.

We now prove Case IIIa, the others are similar: At $j = 1$, (154) becomes

$$R_{\mu,1}^{d,1}\mathrm{u}_1^d = 16\sqrt{3d(d+2)}(\mu_1-\mu_3+d)(\mu_2-\mu_3-1)\mathrm{u}_1^{d+1}.$$

For the base case of the orthonormality condition (157), we apply (161) with $j = 1$ and $\mathrm{u}_{-1}^{d+a} = (-1)^a \mathrm{u}_1^{d+a}$ (by assumption) and put everything on the same side, giving

$$2\sqrt{d(d+2)}(\mu_1-\mu_3+d)(\mu_2-\mu_3-1)(\langle \mathrm{u}_1^{d+1}, \mathrm{u}_1^{d+1}\rangle - \langle \mathrm{u}_1^d, \mathrm{u}_1^d\rangle) = 0.$$

In both equations, we know $\sqrt{d(d+2)}(\mu_1-\mu_3+d)(\mu_2-\mu_3-1)\neq 0$, and this completes the induction step in this case.

We now proceed to the second step. Again, the base case that $\mathcal{V}^{d_0}=\mathbb{C}\mathfrak{u}_{\min}^{d_0}$ is in the image of $\mathfrak{u}_{\min}^{d_0}$ under the raising operators and (157) holds at $d=d_0$ is obvious. As mentioned above, $\mathcal{V}^d$ can be generated from $\mathfrak{u}_{\min}^d$ by repeatedly applying (89), though there is a little extra work going from $j=j_{\min}$ to $j+2$ in the case where $j_{\min}=0$ since then $\mathfrak{u}_{-2}^d=\mathfrak{u}_2^d$ (note that $\mathfrak{u}_{-2}^d=-\mathfrak{u}_2^d$ cannot occur for $\delta=j_{\min}=0$). Even in that case, it is easy to see that the coefficient of $\mathfrak{u}_{j+2}^d$ in (89) is nonzero, and this is enough to conclude that $\mathcal{V}^d=\mathbb{C}[Y_\mu^0]\mathfrak{u}_{\min}^d$ and hence is in the image of $\mathfrak{u}_{\min}^{d_0}$ under the raising operators (by the conclusion of the first step).

Now we show (157) holds for $d>d_0$, assuming it holds for all $d_0\leq d'<d$ and that it holds at $2i+\delta=2i'+\delta=j_{\min}$. By symmetry, it suffices to assume $i\geq i'$, since for $i'\geq i+1$ all of the relevant adjoint operators will respect the spaces $\mathcal{V}^{d+a}$ (which was the only asymmetry in the hypotheses on the sesquilinear forms).

For convenience, let $j=2i+\delta$ and $j'=2i'+\delta$. We now proceed by induction on j. The base cases are $j\in\{0,1,\kappa\}$, which necessarily imply $j=j'=j_{\min}$, and this case is implied by the induction assumption above. Note that $j=\kappa+1$ implies $\kappa=0$ so that again $j=1$.

Assume $j\geq\max\{3,\kappa+2\}$. First assume $j>j'$; then by (90), we have

$$\begin{aligned}
&\sqrt{2d(d+1)(d+2)(2d+1)}\langle Y_\mu^1\mathfrak{u}_{j-2}^d,\mathfrak{u}_{j'}^{d+1}\rangle\\
&\qquad=\sqrt{(d+2-j)(d-1+j)(d+j)(d+1+j)}(\mu_1-\mu_2-1+j)\langle\mathfrak{u}_j^{d+1},\mathfrak{u}_{j'}^{d+1}\rangle\\
&\qquad\quad-\sqrt{(d+3-j)(d+4-j)(d+5-j)(d+j-2)}(\mu_1-\mu_2+3-j)\tfrac12\delta_{j-4=j'}\\
&\qquad\quad-2(j-2)\sqrt{(d+3-j)(d-1+j)}(3\mu_3-d-1)\tfrac12\delta_{j-2=j'}.
\end{aligned}\tag{163}$$

From the definition (156) and using (96) with our induction assumption and (91), this may also be written

$$\begin{aligned}
&\sqrt{2d(d+1)(d+2)(2d+3)}\langle\mathfrak{u}_{j-2}^d,Y_{-\bar\mu}^{-1}\mathfrak{u}_{j'}^{d+1}\rangle\\
&\quad=-\sqrt{(d+2-j')(d-1+j')(d+j')(d+1+j')}(-\mu_1+\mu_2+1-j')\tfrac12\delta_{j-2=j'-2}\\
&\qquad+2j'\sqrt{(d+1-j')(d+1+j')}(-3\mu_3+d+1)\tfrac12\delta_{j-2=j'}\\
&\qquad+\sqrt{(d-1-j')(d-j')(d+1-j')(d+2+j')}(-\mu_1+\mu_2+1+j')\tfrac12\delta_{j-2=j'+2},
\end{aligned}\tag{164}$$

from which it follows

$$\langle\mathfrak{u}_j^{d+1},\mathfrak{u}_{j'}^{d+1}\rangle=\tfrac12\delta_{j=j'}.$$

Now having shown the case $j=j'+2$ (and by symmetry $j=j'-2$), the proof applies verbatim at $j=j'$.

The remaining case is $j=2$, $j'\in\{0,2\}$, with $\kappa=0$, $\varepsilon=(-1)^d$ and $d\geq d_0+2$, and this has a proof identical to the previous using (92). ☐

10. The structure of the cusp forms

10.1. ***The orthogonal and harmonic descriptions of minimal K-types.*** We now finish the proof of Proposition 2.

First, we show condition (1) implies condition (2): If ϕ is orthogonal to the raises of all lower-weight forms, then some word L in Y^{-1} and Y^{-2}, i.e., $L = Y^{a_1} Y^{a_2} \cdots Y^{a_k}$ for some $k \geq 0$ and $a_i \in \{-1, -2\}$, makes $0 \neq L\phi =: \phi^{d_0}$ minimal in the sense that $Y^{-1}\phi^{d_0} = 0$ and $Y^{-2}\phi^{d_0} = 0$. But if $L \neq 1$, then

$$0 = \langle \widehat{L}\phi^{d_0}, \phi\rangle = \langle \phi^{d_0}, L\phi\rangle = \langle \phi^{d_0}, \phi^{d_0}\rangle \neq 0,$$

a contradiction, so ϕ must already satisfy $Y^{-1}\phi = 0$ and $Y^{-2}\phi = 0$.

As mentioned in the discussions following Corollary 18 and Proposition 14, this is already sufficient to say ϕ is an eigenfunction of Y^0, except possibly when $d = 2$. If $d = 2$ and ϕ is not an eigenfunction of Y^0, i.e., when the Fourier coefficients of ϕ are multiples of $g_1 W^2(\cdot, \mu, \psi_{1,1})$ with all $\mu_i - \mu_j \neq 1$, then $Y^{-2}Y^0\phi$ is not zero (recall (115)), and the above argument works taking L to be $Y^{-2}Y^0$. So ϕ is minimal in the sense of condition (2).

Condition (1) follows from condition (2) by considering minimal-weight ancestors: We use the decomposition $\mathcal{A}^d = \bigoplus_\mu \mathcal{A}^d_\mu$ implied by Theorem 1(1). Suppose $\phi \in \mathcal{A}^d_\mu$ is minimal in the sense of condition (2). Now if $\phi' \in \mathcal{A}^{d-a}_{\mu'}$ for some $a \in \{1, 2\}$, then by applying Y^{-1} and Y^{-2} (and possibly Y^0) to ϕ', we must arrive at some cusp form which satisfies condition (2), but by Theorem 3 (and the fact that $Y^1 : \mathcal{A}^0 \to \mathcal{A}^1$ is the zero operator), we know $\mu \neq \mu'$, and hence ϕ is orthogonal to $Y^a\phi'$ as they belong to different eigenspaces of the Casimir operators.

It remains to prove the equivalence of conditions (2) and (3). Since Λ_X necessarily commutes with Δ_1 and Δ_2, we may assume the cusp forms in question are eigenfunctions of the latter two operators. We know that such forms have spectral parameters of the form $(it_1, it_2, -i(t_1+t_2))$ or $(x+it, -x+it, -2it)$, up to permutation. In the first case, the Λ_X eigenvalue is computed to be

$$((t_1 - t_2)^2 + 4X^2)((2t_1 + t_2)^2 + 4X^2)((t_1 + 2t_2)^2 + 4X^2),$$

and, for $X \neq 0$, this is not zero. In the second case, the eigenvalue is

$$4(X - x)(X + x)(9t^2 + (2X - x)^2)(9t^2 + (2X + x)^2),$$

and this is zero exactly when $x = \pm X$ or $(x, t) = (\pm 2X, 0)$. For $d \geq 2$, since we have shown the cusp forms having minimal K-type $\mathcal{D}^d$ are exactly those with spectral parameters of the form

$$\left(\tfrac{1}{2}(d-1) + it, -\tfrac{1}{2}(d-1) + it, -2it\right)$$

and there are no cusp forms of K-type $\mathcal{D}^d$ with spectral parameters of the form $(d-1, 1-d, 0)$, this gives the claim.

10.2. ***The cuspidal spectral expansion.*** We have some final calculations to complete the proof of Theorem 6. Consider $\Phi \in \mathcal{S}_3$ and Φ^d, $\widetilde{\Phi}^d$ as in Section 3 and $\mathcal{V}^d_{\kappa,\chi}$ as in Section 9. Then $\Phi \in \mathcal{A}^{d_0 *}_\mu$ for some μ and d_0, and we determine χ and κ from d_0 by (19) and

$$\kappa = \begin{cases} d_0 & \text{if } d_0 \geq 2, \\ 0 & \text{otherwise.} \end{cases} \tag{165}$$

First, we point out that in the proof of Proposition 17, we showed $v\widetilde{\Phi}^d(g) = (\mathrm{proj}_{\mathcal{V}^d_{\kappa,\chi}} v)\widetilde{\Phi}^d(g)$ for $v \in \mathcal{V}^d_{0,\chi}$, since all of the Whittaker functions of $v\widetilde{\Phi}^d(g)$ are zero when also $v \in (\mathcal{V}^d_{\kappa,\chi})^\perp$; in particular, the projection assumption of (156) is met for the inner product (158). Moreover, in Proposition 19, we showed that all vectors $v\widetilde{\Phi}^d(g)$ with $v \in \mathcal{V}^d_{\kappa,\chi}$ are obtainable by applying suitable combinations of the Y^a operators to $\Phi = \mathrm{u}^{d_0,\varepsilon}_{\min}\Phi^{d_0}$ itself, and hence they are all true cusp forms by the left-translation invariance of the Y^a operators. So the rows of $\Phi^d(g)$ which do not correspond to cusp forms also do not contribute to $\mathrm{Tr}(\Phi^d(g)\overline{\widetilde{\Phi}^d(g')}^T)$.

One might wonder about the need for $\widetilde{\Phi}$ when a similar situation does not happen for the maximal parabolic Eisenstein series. The simple answer is that for the Eisenstein series, our choice of normalization constants effectively completes the Whittaker function under the $\mu \mapsto \mu^{w_2}$ functional equation. (We cannot formulate a matrix-valued Whittaker function which is complete under all of the functional equations, because the w_3 functional equation, in particular, is not a diagonal matrix, while the w_2 functional equation is diagonal with distinct entries, so they cannot be simultaneously diagonalized.)

We normalize the Fourier–Whittaker coefficients as follows: The space $\mathcal{V}^{d_0}_{\kappa,\chi}$ has dimension 1, and hence the matrix $\widetilde{\Phi}^{d_0}(g)$ has exactly one distinct nonzero row (which may occur twice, due to the symmetry of Σ_χ), call it $\widetilde{\Phi}$. If we insist that

$$\int_{\Gamma\backslash G} \Phi(g)\overline{\widetilde{\Phi}(g)}^T\, dg = \begin{cases} 1 & \text{if } \kappa = \delta = 0, \\ \frac{1}{2} & \text{otherwise,} \end{cases} \tag{166}$$

then Proposition 20 implies that the rows of Φ^d and $\widetilde{\Phi}^d$ have the desired orthonormality. That is,

$$\int_{\Gamma\backslash G} (v\Phi^d(g))\overline{(v'\widetilde{\Phi}^d(g))}^T\, dg = v\,\overline{v'}^T \tag{167}$$

for $v, v' \in \mathcal{V}^d_{\kappa,\chi}$. Note that the rows of $\Phi^d(g)$ which apparently have norm $\frac{1}{\sqrt{2}}$ actually occur twice in $\Phi^d(g)$, so there is no discrepancy. Since $-\bar{\mu}$ is either μ or μ^{w_2} and the w_2 functional equation of the Whittaker function acts by a nonzero scalar on the minimal Whittaker function, it is always possible to arrange (166).

The final step is to convert back to scalar-valued forms. We notice that

$$\begin{aligned}(2d+1)\int_{\Gamma\backslash G} \Phi^d_{i,j}(g)\overline{\widetilde{\Phi}^d_{k,\ell}(g)}\, dg &= (2d+1)\sum_{m,n}\int_{\Gamma\backslash G/K} \Phi^d_{i,m}(z)\overline{\widetilde{\Phi}^d_{n,k}(z)}\int_K \mathcal{D}^d_{m,j}(k)\overline{\mathcal{D}^d_{n,\ell}(k)}\, dk\, dz \\ &= \delta_{j=\ell}\int_{\Gamma\backslash G} \Phi^d_i(g)\overline{\widetilde{\Phi}^d_k(g)}^T\, dg,\end{aligned}$$

and this produces the factor $2d+1$ in Theorem 6.

Appendix A: Shalika's multiplicity-one theorem

We now show that Theorem 1(3) follows from Shalika's local multiplicity-one theorem [1974, Theorem 3.1]; for this we follow the notation of [Knapp 1986]. Assume that $n_1 n_2 \neq 0$, since otherwise $c_n = 0$

works. We note that ϕ generates a unitary, admissible representation (finite-dimensionality of the K-types is given by part (1) of Theorem 1) of G via right translation

$$\mathcal{R}_\phi := \operatorname{span}\{\phi(\cdot\, g) \mid g \in G\} \subset L^2(\Gamma\backslash G).$$

We have assumed that $\mathcal{R}_\phi$ is irreducible; in general, $\mathcal{A}^d_\mu$ is a finite (by part (1) of the theorem) span of elements of irreducible representations, so that the c_n and f of the theorem will be replaced by linear combinations of some finite set $f_1, \dots, f_k$.

This representation is infinitesimally equivalent[1] to a subrepresentation [Knapp 1986, Theorem 8.37] of a principal series representation

$$\mathcal{P}_{\mu,\chi} = \left\{ f : G \to \mathbb{C} \;\middle|\; f(xyvk) = p_{\rho+\mu}(y)\chi(v) f(k),\ \int_K |f(k)|^2\, dk < \infty \right\}$$

for some χ (a character of the diagonal, orthogonal matrices V); denote the isomorphism by $\mathcal{L} : \mathcal{R}_\phi \to \mathcal{P}_{\mu,\chi}$.

Denote the subspace of smooth functions in $\mathcal{R}_\phi$ and $\mathcal{P}_{\mu,\chi}$ as $\mathcal{R}^\infty_\phi$ and $\mathcal{P}^\infty_{\mu,\chi}$ respectively. We give these spaces the Fréchet topology generated by the seminorms $\|f\|^2_{R,X} := \int_{\Gamma\backslash G} |(Xf)(g)|^2\, dg$, $f \in \mathcal{R}^\infty_\phi$, and $\|f\|^2_{P,X} := \int_K |(Xf)(k)|^2\, dk$, $f \in \mathcal{P}^\infty_{\mu,\chi}$, for all $X \in \mathfrak{g}_\mathbb{C}$. The property of infinitesimal equivalence means that the isomorphism $\mathcal{L}$ preserves the $(\mathfrak{g}, K)$-module structure of the admissible representations; in particular, the action of the Lie algebra $\mathfrak{g}_\mathbb{C}$, and hence the generated Fréchet topology is preserved. Thus $\mathcal{L}$ restricts to $\mathcal{L}^\infty : \mathcal{R}^\infty_\phi \to \mathcal{P}^\infty_{\mu,\chi}$.

We have the Whittaker model

$$\mathcal{W}_{n,\mu} = \{ f \in C^\infty(G) \mid f(xg) = \psi_n(x) f(g),\ \Delta_i f = \lambda_i(\mu) f,\ i = 1, 2\},$$

which is once again given the Fréchet topology generated by the action of the Lie algebra. The operator $\mathcal{F}_n : \mathcal{R}^\infty_\phi \to \mathcal{W}_{n,\mu}$, which takes a cusp form to its n-th Fourier coefficient,

$$(\mathcal{F}_n f)(g) := \int_{U(\mathbb{Z})\backslash U(\mathbb{R})} f(ug)\overline{\psi_n(u)}\, du, \qquad f \in \mathcal{R}^\infty_\phi,$$

has image in the Whittaker model, as does the Jacquet integral,

$$(\mathcal{J}_n f)(g) := \int_{U(\mathbb{R})} f(w_l u g)\overline{\psi_n(u)}\, du, \qquad f \in \mathcal{P}^\infty_{\mu,\chi},$$

viewed as an operator $\mathcal{J}_n : \mathcal{P}^\infty_{\mu,\chi} \to \mathcal{W}_{n,\mu}$. Both operators commute with the action of the Lie algebra; hence they are continuous with respect to the Fréchet topology.

The convergence and analytic continuation of the Jacquet integral was originally studied in [Jacquet 1967], but the necessary extension to $\operatorname{Re}(\mu_i) = \operatorname{Re}(\mu_j)$, $i \neq j$, can instead be deduced from Propositions I.3.1 and I.3.3. Indeed, Proposition I.3.1 gives the analytic continuation, and the functional equations; Proposition I.3.3, plus the usual Phragmén–Lindelöf argument, shows the entries of $W^d(g, \mu, \psi_n)$ are

[1] That is, isomorphic by a bounded, unitary intertwining operator that commutes with the action of the Lie algebra $\mathfrak{g}_\mathbb{C}$; see [Knapp 1986, Corollary 9.2].

polynomially bounded in d. Then on $\mathrm{Re}(\mu_i)=\mathrm{Re}(\mu_j)$, we may define $\mathcal{J}_n f$ by the expansion (I.2.10) of f into Wigner $\mathcal{D}$-matrices,

$$(\mathcal{J}_n f)(g) := \sum_{d\geq 0}(2d+1)\,\mathrm{Tr}\biggl(\int_K f(k)\overline{\mathcal{D}^d(k)^T}\,dk\,W^d(g,\mu,\psi_n)\biggr),$$

since the entries of $\int_K f(k)\overline{\mathcal{D}^d(k)^T}\,dk$ have superpolynomial decay in d. This is the natural, continuous extension which gives $(\mathcal{J}_n f)(g)=W^d_{m',m}(g,\mu,\psi_n)$ for $f(xyk)=p_{\rho+\mu}(y)\mathcal{D}^d_{m',m}(k)$.

As an aside, we note that the extension of Jacquet integral to $\mathrm{Re}(\mu_i)=\mathrm{Re}(\mu_j)$ may instead be accomplished by interpreting the integral in the Riemannian sense, $\int_{U(\mathbb{R})}=\lim_{R\to\infty}\int_{[-R,R]^2\times\mathbb{R}}$, by integration by parts; this is done very explicitly for $f=1$ in the analysis of X_3 in [Buttcane 2013, Section 4.3], but easily extends to nontrivial f.

Shalika's local multiplicity-one theorem [1974, Theorem 3.1] states the operators $\mathcal{F}_n$ and $\mathcal{J}_n\circ\mathcal{L}^\infty$ are identical, up to a constant. In particular, $\mathcal{F}_n(\phi)=c_n\mathcal{J}_n(\mathcal{L}^\infty(\phi))$ for some constant $c_n\in\mathbb{C}$, and we take the f in the theorem to be the restriction to K of $\mathcal{L}^\infty(\phi)$.

In terms of Shalika's notation, we notice that Shalika's $\mathcal{D}(\Pi)$ is a dense subspace of the smooth vectors $\mathcal{R}^\infty_\phi$, the continuity condition on elements of Shalika's $\mathcal{D}'(\Pi)$ is with respect to the Fréchet topology described above, and the images of both $\mathcal{F}_n$ and $\mathcal{J}_n\circ\mathcal{L}^\infty$ trivially lie in Shalika's $\mathcal{D}'_{\psi_n}(\Pi)\subset\mathcal{W}_{n,\mu}$.

Appendix B: Computing the Casimir operators

We now prove Lemma 7. The first step is to write the $E_{i,j}$ basis in terms of the X_j, K_j basis: As operators on smooth functions of G (that is, we drop the identity matrix which acts as the zero operator),

$$4E_{i,j}=\sum_{|k|\leq 2}\overline{[X_k]}_{i,j}X_k+\sum_{|k|\leq 1}\overline{[K_{-k}]}_{i,j}K_k, \tag{168}$$

where, as in (38), $[X_k]_{i,j}$ means the entry at index i,j of the matrix X_k and similarly for $[K_k]_{i,j}$.

By the symmetries of the matrices X_k and K_k, the coefficients satisfy

$$\begin{aligned}\overline{[X_k]}_{i,j}&=(-1)^k[X_{-k}]_{i,j}, &\quad \overline{[K_k]}_{i,j}&=(-1)^k[K_{-k}]_{i,j},\\ [X_k]_{j,i}&=[X_k]_{i,j}, & [K_k]_{j,i}&=-[K_k]_{i,j}.\end{aligned}$$

Inserting these expressions into (42) gives

$$-32\Delta_1=\sum_{i,j}\biggl(\sum_{|k_1|\leq 2}(-1)^{k_1}[X_{-k_1}]_{i,j}X_{k_1}+\sum_{|k_1|\leq 1}(-1)^{k_1}[K_{-k_1}]_{i,j}K_{k_1}\biggr)\circ\biggl(\sum_{|k_2|\leq 2}\overline{[X_{k_2}]}_{i,j}X_{k_2}-\sum_{|k_2|\leq 1}\overline{[K_{k_2}]}_{i,j}K_{k_2}\biggr).$$

Now we interchange the sums and use the orthonormality (38), so we have

$$-8\Delta_1=\sum_{|j|\leq 2}(-1)^jX_j\circ X_{-j}+2\Delta_K. \tag{169}$$

We reduce to the $\widetilde{Z}_j$ operators by (74), so that

$$-8\Delta_1 - 2\Delta_K = \sum_{|\ell_1|,|\ell_2|\le 2} \left(\sum_{|j|\le 2} (-1)^j \mathcal{D}^2_{\ell_1,j}(k) \mathcal{D}^2_{\ell_2,-j}(k) \right) \widetilde{Z}_{\ell_1} \circ \widetilde{Z}_{\ell_2} + \sum_{|\ell_1|,|\ell_2|\le 2} \left(\sum_{|j|\le 2} (-1)^j \mathcal{D}^2_{\ell_1,j}(k) \widetilde{Z}_{\ell_1} \mathcal{D}^2_{\ell_2,-j}(k) \right) \widetilde{Z}_{\ell_2}. \tag{170}$$

The symmetry (71) implies $(-1)^{\ell+j}\mathcal{D}^2_{\ell,-j}(k) = \overline{\mathcal{D}^2_{-\ell,j}(k)}$ and so the orthogonality of the rows of $\mathcal{D}^2(k)$ gives

$$\sum_{|j|\le 2} (-1)^j \mathcal{D}^2_{\ell_1,j}(k) \mathcal{D}^2_{\ell_2,-j}(k) = (-1)^{\ell_2} \delta_{\ell_1=-\ell_2}. \tag{171}$$

Also, we may compute from (73) and (40) that

$$\widetilde{Z}_{\pm 2} \mathcal{D}^d_{\ell,j}(k) = \pm \ell \mathcal{D}^d_{\ell,j}(k), \tag{172}$$

$$\widetilde{Z}_{\pm 1} \mathcal{D}^d_{\ell,j}(k) = -\sqrt{d(d+1) - \ell(\ell \mp 1)}\, \mathcal{D}^d_{\ell \mp 1,j}(k), \tag{173}$$

$$\widetilde{Z}_0 \mathcal{D}^d_{\ell,j}(k) = 0. \tag{174}$$

Applying these two facts to (170), we have

$$-8\Delta_1 - 2\Delta_K = \sum_{|\ell|\le 2} (-1)^\ell \widetilde{Z}_\ell \circ \widetilde{Z}_{-\ell} - 2\widetilde{Z}_{-2} + 2\sqrt{6}\widetilde{Z}_0 - 2\widetilde{Z}_2.$$

From (73), (41) and the commutativity of all K^{Left}_j and Z_j, we have

$$-8\Delta_1 = \sum_{\ell=-2}^{2} (-1)^\ell Z_\ell \circ Z_{-\ell} - 2Z_2 + 2\sqrt{6} Z_0 - 2Z_{-2} + \sqrt{2} i K^{\text{Left}}_0 \circ (Z_2 - Z_{-2}) + 2K^{\text{Left}}_1 \circ Z_{-1} + 2K^{\text{Left}}_{-1} \circ Z_1. \tag{175}$$

Since the operators K^{Left}_j are zero on spherical functions, we see that

$$-8\Delta_1^\circ = \sum_{\ell=-2}^{2} (-1)^\ell Z_\ell \circ Z_{-\ell} - 2Z_2 - 2Z_{-2} + 2\sqrt{6} Z_0, \tag{176}$$

and (46) follows.

The degree-3 operator is somewhat more complicated, so we increase the formalism a little: We collect the X_ℓ and K_ℓ operators and the $\widetilde{Z}_\ell$ and K^{Left}_ℓ operators by defining

$$\mathcal{X}^d_\ell = \begin{cases} X_\ell & \text{if } d = 2, \\ K_\ell & \text{if } d = 1, \end{cases} \qquad \mathcal{Z}^d_\ell = \begin{cases} \widetilde{Z}_\ell & \text{if } d = 2, \\ K^{\text{Left}}_\ell & \text{if } d = 1. \end{cases}$$

In the same manner as we derived (75), we have

$$k K_j k^{-1} = \sum_{|\ell|\le 1} i^{j-\ell} \mathcal{D}^1_{\ell,j}(k) K_j, \qquad k \in K, \tag{177}$$

which follows from

$$k(\alpha,0,0)K_j k(-\alpha,0,0) = e^{-ij\alpha}K_j,$$
$$w_3 K_j w_3^{-1} = \sum_{|\ell|\le 1} i^{j-\ell}\mathcal{D}^1_{\ell,w_3}(k)K_j, \tag{178}$$

and this may be checked using

$$\mathcal{D}^1(w_3) = \frac{1}{2}\begin{pmatrix} -1 & -i\sqrt{2} & 1 \\ i\sqrt{2} & 0 & i\sqrt{2} \\ 1 & -i\sqrt{2} & -1 \end{pmatrix}.$$

Then the trick (76) implies

$$K_j = \sum_{|\ell|\le 1} i^{j-\ell}\mathcal{D}^1_{\ell,j}(k)K_j^{\mathrm{Left}}, \tag{179}$$

as differential operators in the Iwasawa coordinates. Collectively, we may now write

$$\mathcal{X}_j^d = \sum_{|\ell|\le d} i^{d^2(j-\ell)}\mathcal{Z}_\ell^d. \tag{180}$$

(The factor d^2 here should not be confused with the Wigner d-polynomial.)

Now the orthonormality relations (38) are replaced with properties of Clebsch–Gordan coefficients. The identities are best expressed in terms of the Wigner three-j symbols (recall (53)), so we define

$$\mathfrak{D}^{d_1,d_2,d_3}_{\ell_1,\ell_2,\ell_3} := \delta_{\ell_1+\ell_2+\ell_3=0}\, i^{d_1\ell_1+d_2\ell_2+d_3\ell_3}\begin{pmatrix} d_1 & d_2 & d_3 \\ \ell_1 & \ell_2 & \ell_3\end{pmatrix} \times \begin{cases} -\sqrt{105} & \text{if } d_1+d_2+d_3=6, \\ -3\sqrt{15}i & \text{if } d_1+d_2+d_3=5, \\ 3\sqrt{5} & \text{if } d_1+d_2+d_3=4, \\ -3\sqrt{3}i & \text{if } d_1+d_2+d_3=3. \end{cases}$$

Then

$$\mathcal{U} := 48\sum_{i,j,k} E_{i,j}\circ E_{j,k}\circ E_{j,i} = \sum_{d_1,d_2,d_3\in\{1,2\}}\ \sum_{\substack{|j_1|\le d_1,|j_2|\le d_2,|j_3|\le d_3 \\ j_1+j_2+j_3=0}} \mathfrak{D}^{d_1,d_2,d_3}_{j_1,j_2,j_3}\mathcal{X}^{d_1}_{j_1}\circ\mathcal{X}^{d_2}_{j_2}\circ\mathcal{X}^{d_3}_{j_3}, \tag{181}$$

which follows from

$$\frac{3}{4}\sum_{i,j,k}\overline{[\mathcal{X}^{d_1}_{\ell_1}]}_{i,j}\overline{[\mathcal{X}^{d_2}_{\ell_2}]}_{j,k}\overline{[\mathcal{X}^{d_3}_{\ell_3}]}_{k,i} = \frac{3}{4}\overline{\mathrm{Tr}(\mathcal{X}^{d_2}_{\ell_2}\mathcal{X}^{d_2}_{\ell_2}\mathcal{X}^{d_3}_{\ell_3})} = \mathfrak{D}^{d_1,d_2,d_3}_{\ell_1,\ell_2,\ell_3},$$

and this may be checked directly.

To (181), we apply (180), and carefully interchange summations. For $d, j, \ell\in\mathbb{Z}^3$, let

$$\mathfrak{E}^d_{j,\ell} = i^{d_1^2(j_1-\ell_1)+d_2^2(j_2-\ell_2)+d_3^2(j_3-\ell_3)}\mathfrak{D}^{d_1,d_2,d_3}_{j_1,j_2,j_3}.$$

Then the key identity is

$$\sum_{\substack{|j_1|\le d_1,|j_2|\le d_2,|j_3|\le d_3 \\ j_1+j_2+j_3=0}} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)\mathcal{D}^{d_2}_{\ell_2,j_2}(k)\mathcal{D}^{d_3}_{\ell_3,j_3}(k) = \mathfrak{D}^{d_1,d_2,d_3}_{\ell_1,\ell_2,\ell_3}. \tag{182}$$

This can be seen by the definition (52) and orthogonality [DLMF, equation 34.3.16] of the Clebsch–Gordan coefficients as follows: Suppose, for convenience, that $d_1 = d_2 = d_3 = 2$; then

$$\sum_{j_1+j_2+j_3=0} \begin{pmatrix} 2 & 2 & 2 \\ j_1 & j_2 & j_3 \end{pmatrix} \mathcal{D}^2_{\ell_1,j_1}(k)\mathcal{D}^2_{\ell_2,j_2}(k)\mathcal{D}^2_{\ell_3,j_3}(k)$$
$$=(-1)^{\ell_3}\sum_{d_4=0}^{4} \begin{pmatrix} 2 & 2 & d_4 \\ \ell_1 & \ell_2 & -\ell_1-\ell_2 \end{pmatrix} \sum_{|j_3|\le 2} (-1)^{j_3}\mathcal{D}^{d_4}_{\ell_1+\ell_2,-j_3}(k)\mathcal{D}^2_{\ell_3,j_3}(k) \sum_{j_1+j_2=-j_3} (2d_4+1) \begin{pmatrix} 2 & 2 & 2 \\ j_1 & j_2 & j_3 \end{pmatrix}\begin{pmatrix} 2 & 2 & d_4 \\ j_1 & j_2 & j_3 \end{pmatrix},$$

and the inner sum over $j_1 + j_2 = -j_3$ is $\delta_{d_4=2}$ by orthogonality. The identity follows by applying (171) on the j_3 sum.

We have $\mathcal{U} = \mathcal{U}_3 + \mathcal{U}_2 + \mathcal{U}_1$, where

$$\mathcal{U}_3 = \sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)\mathcal{D}^{d_2}_{\ell_2,j_2}(k)\mathcal{D}^{d_3}_{\ell_3,j_3}(k)\Bigg) \mathcal{Z}^{d_1}_{\ell_1}\circ\mathcal{Z}^{d_2}_{\ell_2}\circ\mathcal{Z}^{d_3}_{\ell_3},$$

$$\begin{aligned}\mathcal{U}_2 = &\sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)(\mathcal{Z}^{d_1}_{\ell_1}\mathcal{D}^{d_2}_{\ell_2,j_2}(k))\mathcal{D}^{d_3}_{\ell_3,j_3}(k)\Bigg) \mathcal{Z}^{d_2}_{\ell_2}\circ\mathcal{Z}^{d_3}_{\ell_3}\\
&+\sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)\mathcal{D}^{d_2}_{\ell_2,j_2}(k)(\mathcal{Z}^{d_1}_{\ell_1}\mathcal{D}^{d_3}_{\ell_3,j_3}(k))\Bigg) \mathcal{Z}^{d_2}_{\ell_2}\circ\mathcal{Z}^{d_3}_{\ell_3}\\
&+\sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)\mathcal{D}^{d_2}_{\ell_2,j_2}(k)(\mathcal{Z}^{d_2}_{\ell_2}\mathcal{D}^{d_3}_{\ell_3,j_3}(k))\Bigg) \mathcal{Z}^{d_1}_{\ell_1}\circ\mathcal{Z}^{d_3}_{\ell_3},\end{aligned}$$

$$\begin{aligned}\mathcal{U}_1 = &\sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)\mathcal{D}^{d_2}_{\ell_2,j_2}(k)(\mathcal{Z}^{d_1}_{\ell_1}\circ\mathcal{Z}^{d_2}_{\ell_2}\mathcal{D}^{d_3}_{\ell_3,j_3}(k))\Bigg) \mathcal{Z}^{d_3}_{\ell_3}\\
&+\sum_{\substack{d_1,d_2,d_3\\ \ell_1,\ell_2,\ell_3}} \Bigg(\sum_{j_1+j_2+j_3=0} \mathfrak{E}^d_{j,\ell}\mathcal{D}^{d_1}_{\ell_1,j_1}(k)(\mathcal{Z}^{d_1}_{\ell_1}\mathcal{D}^{d_2}_{\ell_2,j_2}(k))(\mathcal{Z}^{d_2}_{\ell_2}\mathcal{D}^{d_3}_{\ell_3,j_3}(k))\Bigg) \mathcal{Z}^{d_3}_{\ell_3}.\end{aligned}$$

Though it is far more involved than the degree-2 operator, from (172)-(174) and (182), we can compute

$$\mathcal{U}^3 = \sum_{\substack{d_1,d_2,d_3\\ \ell_1+\ell_2+\ell_3=0}} \mathfrak{D}^{d_1,d_2,d_3}_{\ell_1,\ell_2,\ell_3}\mathcal{Z}^{d_1}_{\ell_1}\circ\mathcal{Z}^{d_2}_{\ell_2}\circ\mathcal{Z}^{d_3}_{\ell_3}, \tag{183}$$

$$\mathcal{U}^2 = \sum_{\substack{d_2,d_3\\ \ell_2,\ell_3}} \mathfrak{F}^{d_2,d_3}_{\ell_2,\ell_3}\mathcal{Z}^{d_2}_{\ell_2}\circ\mathcal{Z}^{d_3}_{\ell_3}, \tag{184}$$

$$\mathcal{U}^1 = 48\sqrt{6}\widetilde{Z}_0 = 48\sqrt{6}Z_0, \tag{185}$$

where the coefficients matrices $\mathfrak{F}^{d_2,d_3}$ are given by

$$\mathfrak{F}^{1,1} = \begin{pmatrix} 9 & 0 & 15 \\ 0 & 12 & 0 \\ 15 & 0 & 9 \end{pmatrix}, \quad \mathfrak{F}^{1,2} = \begin{pmatrix} 0 & -3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -3 & 0 \end{pmatrix}, \tag{186}$$

$$\mathfrak{F}^{1,2} = \begin{pmatrix} 0 & 6i\sqrt{2} & 0 \\ 3 & 0 & -15 \\ 0 & 0 & 0 \\ -15 & 0 & 3 \\ 0 & -6i\sqrt{2} & 0 \end{pmatrix}, \quad \mathfrak{F}^{2,2} = \begin{pmatrix} 0 & 0 & 2\sqrt{6} & 0 & 0 \\ 0 & -9 & 0 & -27 & \\ 4\sqrt{6} & 0 & 36 & 0 & 4\sqrt{6} \\ 0 & -27 & 0 & 9 & 0 \\ 0 & 0 & 2\sqrt{6} & 0 & 0 \end{pmatrix}, \tag{187}$$

indexing from the center. We have now completely removed the Wigner $\mathcal{D}$-matrices from the above expressions, and the rest is purely computational.

As before, we have

$$144\Delta_2^\circ = \sum_{\ell_1+\ell_2+\ell_3=0} \mathfrak{D}^{2,2,2}_{\ell_1,\ell_2,\ell_3} Z_{\ell_1} \circ Z_{\ell_2} \circ Z_{\ell_3} + \sum_{\ell_2,\ell_3} \mathfrak{F}^{2,2}_{\ell_2,\ell_3} Z_{\ell_2} \circ Z_{\ell_3} + 48\sqrt{6} Z_0 + 144\Delta_1^\circ,$$

and (47) follows by using

$$[K_0^{\mathrm{Left}}, K_{\pm1}^{\mathrm{Left}}] = \mp\sqrt{2} i K_{\pm1}^{\mathrm{Left}}, \quad [K_1^{\mathrm{Left}}, K_{-1}^{\mathrm{Left}}] = \sqrt{2} i K_0^{\mathrm{Left}}, \tag{188}$$

and

$$[Z_{\pm2}, Z_0] = 0, \qquad [Z_{\pm1}, Z_0] = \sqrt{6} Z_{\pm1}, \qquad [Z_{-1}, Z_1] = 0 \tag{189}$$

$$[Z_{\pm2}, Z_{\pm1}] = -Z_{\pm1}, \quad [Z_{\pm2}, Z_{\mp1}] = Z_{\mp1} - 2Z_{\pm1}, \quad [Z_{-2}, Z_2] = 2Z_2 - 2Z_{-2}. \tag{190}$$

Acknowledgements

The author would like to thank Stephen D. Miller for his insight into the nonspherical Maass forms; in particular, Prof. Miller and the author independently and near-simultaneously arrived at forms of (74). The author would also like to thank James Cogdell and Joseph Hundley for their comments and helpful discussions along the way, and the referee for a careful reading of the paper.

References

[Abramowitz and Stegun 1964] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards Applied Mathematics Series **55**, U.S. Government Printing Office, Washington, D.C., 1964. MR Zbl

[Barnes 1910] E. W. Barnes, "A transformation of generalised hypergeometric series", *Quart. J. Pure and Applied Math.* **41** (1910), 136–140. Zbl

[Biedenharn and Louck 1981] L. C. Biedenharn and J. D. Louck, *Angular momentum in quantum physics: theory and application*, Encyclopedia of Mathematics and its Applications **8**, Addison-Wesley, Reading, MA, 1981. MR Zbl

[Bump 1984] D. Bump, *Automorphic forms on* $\mathrm{GL}(3, \mathbb{R})$, Lecture Notes in Mathematics **1083**, Springer, 1984. 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 2017a] J. Buttcane, "The arithmetic Kuznetsov formula on GL(3), I: The Whittaker case", preprint, 2017. To appear in *Rev. Mat. Iberoam.* arXiv

[Buttcane 2017b] J. Buttcane, "Kuznetsov, Petersson and Weyl on GL(3), I: The principal series forms", preprint, 2017. To appear in *Amer. J. Math.* arXiv

[Buttcane 2017c] J. Buttcane, "Kuznetsov, Petersson and Weyl on GL(3), II: The generalized principal series forms", preprint, 2017. arXiv

[Buttcane 2018] J. Buttcane, "Higher weight on GL(3), I: The Eisenstein series", *Forum Math.* **30**:3 (2018), 681–722. MR Zbl

[Buttcane and Miller 2017] J. Buttcane and S. D. Miller, "Weights, raising and lowering operators, and K-types for automorphic forms on $SL(3, \mathbb{R})$", preprint, 2017. arXiv

[DLMF] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (editors), *NIST digital library of mathematical functions*, release 1.0.14.

[Gel'fand et al. 1969] I. M. Gel'fand, M. I. Graev, and I. I. Pyatetskii-Shapiro, *Generalized functions*, Representation theory and automorphic functions **6**, W. B. Saunders, Philadelphia, PA, 1969. MR Zbl

[Goldfeld 2006] D. Goldfeld, *Automorphic forms and L-functions for the group* $GL(n, \mathbb{R})$, Cambridge Studies in Advanced Mathematics **99**, Cambridge University Press, Cambridge, 2006. MR Zbl

[Gradshteyn and Ryzhik 2015] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 8th ed., Elsevier/Academic Press, Amsterdam, 2015. MR Zbl

[Harish-Chandra 1968] Harish-Chandra, *Automorphic forms on semisimple Lie groups*, Lecture Notes in Mathematics **62**, Springer, 1968. MR Zbl

[Hirano et al. 2012] M. Hirano, T. Ishii, and T. Miyazaki, "The Archimedean Whittaker functions on GL(3)", pp. 77–109 in *Geometry and analysis of automorphic forms of several variables*, edited by Y. Hamahata et al., Ser. Number Theory Appl. **7**, World Scientific, Hackensack, NJ, 2012. MR Zbl

[Hirano et al. 2016] M. Hirano, T. Ishii, and T. Miyazaki, "The archimedean zeta integrals for $GL(3) \times GL(2)$", *Proc. Japan Acad. Ser. A Math. Sci.* **92**:2 (2016), 27–32. MR Zbl

[Howe 2000] R. Howe, "K-type structure in the principal series of GL_3, I", pp. 77–98 in *Analysis on homogeneous spaces and representation theory of Lie groups* (Okayama–Kyoto, 1997), edited by T. Kobayashi et al., Adv. Stud. Pure Math. **26**, Math. Soc. Japan, Tokyo, 2000. MR Zbl

[Iwaniec 1995] H. Iwaniec, *Introduction to the spectral theory of automorphic forms*, Revista Matemática Iberoamericana, Madrid, 1995. MR Zbl

[Jacquet 1967] H. Jacquet, "Fonctions de Whittaker associées aux groupes de Chevalley", *Bull. Soc. Math. France* **95** (1967), 243–309. MR Zbl

[Jacquet and Shalika 1981] H. Jacquet and J. A. Shalika, "On Euler products and the classification of automorphic forms, II", *Amer. J. Math.* **103**:4 (1981), 777–815. MR Zbl

[Jacquet et al. 1983] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, "Rankin–Selberg convolutions", *Amer. J. Math.* **105**:2 (1983), 367–464. MR Zbl

[Kim 2003] H. H. Kim, "Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2", *J. Amer. Math. Soc.* **16**:1 (2003), 139–183. MR Zbl

[Knapp 1986] A. W. Knapp, *Representation theory of semisimple groups: an overview based on examples*, Princeton Mathematical Series **36**, Princeton University Press, 1986. Zbl

[Langlands 1976] R. P. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics **544**, Springer, 1976. MR Zbl

[Lee 2013] J. M. Lee, *Introduction to smooth manifolds*, 2nd ed., Graduate Texts in Mathematics **218**, Springer, 2013. MR Zbl

[Maass 1952] H. Maass, "Die Differentialgleichungen in der Theorie der elliptischen Modulfunktionen", *Math. Ann.* **125** (1952), 235–263. MR Zbl

[Manabe et al. 2004] H. Manabe, T. Ishii, and T. Oda, "Principal series Whittaker functions on $SL(3, R)$", *Japan. J. Math.* (*N.S.*) **30**:1 (2004), 183–226. MR Zbl

[Miller and Schmid 2011] S. D. Miller and W. Schmid, "Adelization of automorphic distributions and mirabolic Eisenstein series", pp. 289–334 in *Representation theory and mathematical physics*, edited by J. Adams et al., Contemp. Math. **557**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Miyazaki 2008] T. Miyazaki, "The structures of standard $(\mathfrak{g}, K)$-modules of $SL(3, \mathbb{R})$", *Glas. Mat. Ser. III* **43(63)**:2 (2008), 337–362. MR Zbl

[Miyazaki 2010] T. Miyazaki, "Whittaker functions for generalized principal series representations of $SL(3, \mathbb{R})$", pp. 103–110 in *New viewpoints of representation theory and noncommutative harmonic analysis*, edited by M. Itoh and H. Ochiai, RIMS Kôkyûroku Bessatsu **B20**, RIMS, Kyoto, 2010. MR Zbl

[Olver 1975] F. W. J. Olver, "Unsolved problems in the asymptotic estimation of special functions", pp. 99–142 in *Theory and application of special functions* (Madison, WI, 1975), Math. Res. Center, Univ. Wisconsin Publ. **35**, Academic Press, New York, 1975. MR Zbl

[Osborne and Warner 1981] M. S. Osborne and G. Warner, *The theory of Eisenstein systems*, Pure and Applied Mathematics **99**, Academic, New York, 1981. MR Zbl

[Selberg 1956] A. Selberg, "Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series", *J. Indian Math. Soc.* (*N.S.*) **20** (1956), 47–87. MR Zbl

[Shalika 1974] J. A. Shalika, "The multiplicity one theorem for GL_n", *Ann. of Math.* (2) **100** (1974), 171–193. MR Zbl

[Stade 2002] E. Stade, "Archimedean L-factors on $GL(n) \times GL(n)$ and generalized Barnes integrals", *Israel J. Math.* **127** (2002), 201–219. MR Zbl

[Vahutinskiĭ 1968] I. J. Vahutinskiĭ, "Unitary irreducible representations of the group $GL(3, \mathbb{R})$ of real matrices of the third order", *Mat. Sb.* (*N.S.*) **75** (1968), 303–320. In Russian; translated in *Math. USSR-Sbornik* **4**:2 (1968) 273–291. MR Zbl

[Watson 2002] T. C. Watson, *Rankin triple products and quantum chaos*, Ph.D. thesis, Princeton University, 2002, available at https://search.proquest.com/docview/252081445. MR

Communicated by Philippe Michel
Received 2017-04-09 Revised 2018-04-15 Accepted 2018-08-19

buttcane@buffalo.edu Mathematics Department, University at Buffalo, Buffalo, NY, United States

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2295

Stark systems over Gorenstein local rings

Ryotaro Sakamoto

In this paper, we define a Stark system over a complete Gorenstein local ring with a finite residue field. Under some standard assumptions, we show that the module of Stark systems is free of rank 1 and that these systems control all the higher Fitting ideals of the Pontryagin dual of the dual Selmer group. This is a generalization of the theory, developed by B. Mazur and K. Rubin, on Stark (or Kolyvagin) systems over principal ideal local rings. Applying our result to a certain Selmer structure over the cyclotomic Iwasawa algebra, we propose a new method for controlling Selmer groups using Euler systems.

1. Introduction

Euler systems were introduced by V. A. Kolyvagin in order to study Selmer groups. An Euler system is a collection of cohomology classes with certain norm relations. He constructed a collection of cohomology classes which is called Kolyvagin's derivative classes from an Euler system. He used these classes which come from the Euler system of Heegner points to bound the order of the Selmer group of certain elliptic curves over imaginary quadratic fields. Note that the Euler system of Heegner points does not fit in the formalism of this paper; ring class fields of an imaginary quadratic field are used to define this Euler system, while we consider an Euler system with respect to ray class fields in this paper. By using the Euler system of the cyclotomic units, K. Rubin gave another proof of the classical Iwasawa main conjecture over $\mathbb{Q}$, which had been proved by B. Mazur and A. Wiles. After that, many mathematicians, especially K. Kato, B. Perrin-Riou, and K. Rubin, have studied Selmer groups using Euler systems.

The images of Kolyvagin's derivative classes in the local cohomology groups satisfy certain relations which come from the norm relations of Euler systems. B. Mazur and K. Rubin noticed that, after minor modification, these classes satisfy certain additional local conditions. In [Mazur and Rubin 2004], they defined a Kolyvagin system to be a system of cohomology classes satisfying these good interrelations.

To explain their work in [Mazur and Rubin 2004; 2016], we introduce some notation. Let K be a number field and G_K denote the absolute Galois group of K. Let R be a complete noetherian local ring with a finite residue field of odd characteristic and T be a free R-module with an R-linear continuous G_K-action which is unramified outside a finite set of places of K. Let $\mathcal{F}$ be a Selmer structure on T; see Definition 3.1 or [Mazur and Rubin 2004, Definition 2.1.1]. Suppose that T satisfies Hypothesis 3.12 and $\mathcal{F}$ is cartesian (see Definition 3.8). Let $\chi(\mathcal{F}) \in \mathbb{Z}$ be the core rank of $\mathcal{F}$ (see Definition 3.19).

MSC2010: primary 11R23; secondary 11F80, 11S25.
Keywords: Stark systems, Euler systems, Selmer groups, Iwasawa theory.

Suppose that R is a principal ideal local ring and $\chi(\mathcal{F}) = 1$. In [Mazur and Rubin 2004], the authors proved that the module of Kolyvagin systems for $\mathcal{F}$ is free of rank 1 and that the R-module structure of the dual Selmer group associated with $\mathcal{F}$ is determined by these systems. Furthermore, if R is the ring of integers in a finite extension of the field $\mathbb{Q}_p$, then the authors constructed a canonical morphism from the module of Euler systems to the module of Kolyvagin systems for the canonical Selmer structure (see Example 3.4).

We still assume that R is a principal ideal local ring. Suppose that $\chi(\mathcal{F}) > 0$. T. Sano [2014] and B. Mazur and K. Rubin [2016] defined Stark systems, independently (Sano called them unit systems). Mazur and Rubin [2016] removed the assumption $\chi(\mathcal{F}) = 1$ by using Stark systems instead of Kolyvagin systems. Namely, they proved that the R-module structure of the dual Selmer group associated with $\mathcal{F}$ is determined by Stark systems. Furthermore, they showed that the module of Stark systems is free of rank 1 and that there is an isomorphism from the module of Stark systems to the module of Kolyvagin systems if $\chi(\mathcal{F}) = 1$.

To attempt Iwasawa main conjectures using an Euler system, we need a Stark system over a complete regular local ring. In [Mazur and Rubin 2016], the elementary divisor theorem is used to define the module of Stark systems. Hence Stark systems are not defined in that paper when R is not a principal ideal ring. We therefore have two important problems. The first problem is to define a Stark system over an arbitrary complete noetherian local ring. The second problem is to control Selmer groups using Stark systems. We obtain the following results in relation to these problems.

Stark systems over Gorenstein local rings. Suppose that R is a zero-dimensional Gorenstein local ring. To define a Stark system over R, we use a natural generalization of Rubin's lattice, called the exterior bidual, instead of the exterior power. In [Burns et al. 2016], D. Burns, M. Kurihara, and T. Sano used the exterior bidual to state some conjectures about Rubin–Stark elements.

By using the exterior bidual, we generalize some linear-algebraic lemmas in [Mazur and Rubin 2016] to zero-dimensional Gorenstein local rings (see Section 2). Thus, in the same way as that paper, we are able to define a Stark system over R. Furthermore, under the assumptions that T satisfies Hypothesis 3.12 and that $\mathcal{F}$ is cartesian (which are precisely the same kind of the assumptions in [Mazur and Rubin 2004; 2016]), we prove that the module of Stark systems over R is free of rank 1 (Theorem 4.7).

When R is an arbitrary complete Gorenstein local ring, we can construct an inverse system of the modules of Stark systems over some zero-dimensional Gorenstein quotients of R. We define the module of Stark systems over R to be its inverse limit and also prove that this module is free of rank 1 (Theorem 5.4).

Remark 1.1. When R is the cyclotomic Iwasawa algebra, $\mathcal{F}$ is the canonical Selmer structure, and $\chi(\mathcal{F}) = 1$, K. Büyükboduk [2011] proved that the module of Kolyvagin systems for $\mathcal{F}$ is free of rank 1. More generally, when R is a certain Gorenstein local ring, $\mathcal{F}$ is a certain Selmer structure, and $\chi(\mathcal{F}) = 1$, he showed that the module of Kolyvagin systems for $\mathcal{F}$ is free of rank 1 in [Büyükboduk 2016]. In this case, we are able to construct an isomorphism from the module of Kolyvagin systems to the module of Stark systems. Hence Theorem 5.4 is a generalization of these results.

Controlling Selmer groups using Stark systems. Under the same assumptions as above, we prove that all the higher Fitting ideals of the Pontryagin dual of the dual Selmer group associated with $\mathcal{F}$ are determined by Stark systems for $\mathcal{F}$ when R is a complete Gorenstein local ring (Theorem 5.6).

Remark 1.2. If R is a zero-dimensional principal ideal ring and M is a finitely generated R-module, then there is a noncanonical isomorphism $M \simeq \mathrm{Hom}(M, \mathbb{Q}_p/\mathbb{Z}_p)$ as R-modules, which implies

$$\mathrm{Fitt}^i_R(M) = \mathrm{Fitt}^i_R(\mathrm{Hom}(M, \mathbb{Q}_p/\mathbb{Z}_p))$$

for any nonnegative integer i. Thus Theorem 5.6 is a generalization of [Mazur and Rubin 2016, Theorem 8.6] (see Proposition 4.17 and Remark 5.7).

Applying our result to a certain Selmer structure over the cyclotomic Iwasawa algebra Λ, we prove that all the higher Fitting ideals of the Pontryagin dual of the Λ-adic dual Selmer group are determined by Λ-adic Stark systems (Theorem 6.10). In particular, we show that Λ-adic Stark systems determine not only the characteristic ideal of this Λ-module but also its pseudoisomorphism class. In [Mazur and Rubin 2004], the authors showed that a primitive Λ-adic Kolyvagin system controls the characteristic ideal of the Pontryagin dual of the Λ-adic dual Selmer group when the core rank is 1. We will show that Theorem 6.10 is a generalization of this result (Proposition 6.11). However, our proof of Theorem 6.10 is different from the proof in [Mazur and Rubin 2004]. Furthermore, there is a canonical map from the module of Euler systems to the module of Λ-adic Stark systems when the core rank is 1. Hence we propose a new method for controlling Selmer groups using Euler systems.

Remark 1.3. After the author had almost all the results in this paper, T. Sano told the author that D. Burns and he also were studying Kolyvagin and Stark systems over zero-dimensional Gorenstein rings using an exterior bidual; see [Burns and Sano 2016].

Notation. Let p be an odd prime number and K a number field. Let K_v denote the completion of K at a place v. For a field L, let $\bar{L}$ denote a fixed separable closure of L and $G_L := \mathrm{Gal}(\bar{L}/L)$ the absolute Galois group of L.

Throughout this paper, R denotes a complete noetherian local ring with a finite residue field $\Bbbk$ of characteristic p. Let $\mathfrak{m}_R$ be the maximal ideal of R. In addition, T denotes a free R-module of finite rank with an R-linear continuous G_K-action which is unramified outside a finite set of places of K. For any nonnegative integer i, let $H^i(K, T)$ be the i-th continuous cohomology group of G_K with coefficient T.

If A is a commutative ring and I is an ideal of A, then define

$$M^* := \mathrm{Hom}_A(M, A)$$

and $M[I] := \{x \in M \mid Ix = 0\}$ for any A-module M. When $I = (x_1, \dots, x_n)$, we write $M[x_1, \dots, x_n]$ instead of $M[I]$. Let $\bigwedge^r_A M$ denote the r-th exterior power of M in the category of A-modules. We denote by $\ell_A(M) \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ the A-length of M.

For any topological $\mathbb{Z}_p$-module M, we define the Pontryagin dual $M^\vee$ of M to be

$$M^\vee := \mathrm{Hom}_{\mathrm{cont}}(M, \mathbb{Q}_p/\mathbb{Z}_p).$$

2. Preliminaries

Let A be a commutative ring, M an A-module, and r a nonnegative integer. We define the r-th exterior bidual $\bigcap_A^r M$ of M by

$$\bigcap_A^r M := \Big(\bigwedge_A^r M^*\Big)^* = \mathrm{Hom}_A\Big(\bigwedge_A^r \mathrm{Hom}_A(M, A), A\Big).$$

The pairing

$$\bigwedge_A^r M \times \bigwedge_A^r M^* \to A, \quad (m_1 \wedge \cdots \wedge m_r, f_1 \wedge \cdots \wedge f_r) \mapsto \det(f_i(m_j)),$$

induces an A-homomorphism

$$\xi_M^r : \bigwedge_A^r M \to \bigcap_A^r M.$$

Note that the map ξ_M^r is an isomorphism if M is a finitely generated projective A-module. However ξ_M^r is, in general, neither injective nor surjective.

Lemma 2.1. *Let $F \xrightarrow{h} M \xrightarrow{g} N \to 0$ be an exact sequence of A-modules and r a nonnegative integer. If F is a free A-module of rank $s \le r$, then there is a unique A-homomorphism*

$$\phi : \det(F) \otimes_A \bigwedge_A^{r-s} N \to \bigwedge_A^r M$$

such that

$$\phi(a \otimes \wedge^{r-s} g(b)) = (\wedge^s h(a)) \wedge b$$

for any $a \in \det(F)$ and $b \in \bigwedge_A^{r-s} M$.

Proof. Since g is surjective and $\bigwedge_A^{s+1} F = 0$, the map ϕ is well-defined and characterized by the relation $\phi(a \otimes \wedge^{r-s} g(b)) = (\wedge^s h(a)) \wedge b$. □

Lemma 2.2. *Let r be a nonnegative integer. Suppose that we have the following commutative diagram of A-modules*:

$$\begin{tikzcd} F \arrow[r, "h"] \arrow[d, "\alpha"] & M \arrow[r, "g"] \arrow[d, "\beta"] & N \arrow[r] \arrow[d, "\gamma"] & 0 \\ F' \arrow[r, "h'"] & M' \arrow[r, "g'"] & N' \arrow[r] & 0 \end{tikzcd}$$

where the horizontal rows are exact and both F and F' are free A-modules of rank $s \le r$. Then the following diagram is commutative:

$$\begin{tikzcd} \det(F) \otimes_A \bigwedge_A^{r-s} N \arrow[r, "\phi"] \arrow[d, "\wedge^s\alpha \otimes \wedge^{r-s}\gamma"] & \bigwedge_A^r M \arrow[d, "\wedge^r \beta"] \\ \det(F') \otimes_A \bigwedge_A^{r-s} N' \arrow[r, "\phi'"] & \bigwedge_A^r M' \end{tikzcd}$$

where ϕ and ϕ' are the maps defined in Lemma 2.1.

Proof. Let $a \in \det(F)$ and $b \in \bigwedge_A^{r-s} M$. By Lemma 2.1, we have

$$\begin{aligned}(\phi' \circ (\wedge^s \alpha \otimes \wedge^{r-s} \gamma))(a \otimes \wedge^{r-s} g(b)) &= \phi'\big(\wedge^s \alpha(a) \otimes \wedge^{r-s} (g' \circ \beta)(b)\big) \\ &= (\wedge^s (h' \circ \alpha)(a)) \wedge (\wedge^{r-s} \beta(b)) \\ &= \wedge^r \beta((\wedge^s h(a)) \wedge b) \\ &= (\wedge^r \beta \circ \phi)(a \otimes \wedge^{r-s} g(b)). \end{aligned}$$
□

Suppose that A is a zero-dimensional Gorenstein local ring. Note that all free A-modules are injective by the definition of A. In particular, the functor $(-)^* := \mathrm{Hom}_A(-, A)$ is exact. Furthermore, by Matlis duality, we have $\ell_A(M) = \ell_A(M^*)$ and the canonical map $\xi_M^1 : M \to \bigcap_A^1 M$ is an isomorphism for any finitely generated A-module M.

For $i \in \{1, 2\}$, let M_i be an A-module and F_i a free A-module of rank s_i. Suppose that the following diagram is cartesian:

$$\begin{tikzcd} M_1 \arrow[r, hook] \arrow[d] & M_2 \arrow[d] \\ F_1 \arrow[r, hook] & F_2 \end{tikzcd}$$

where the horizontal maps are injective. Note that F_2/F_1 is a free A-module of rank $s_2 - s_1$ since F_1 is an injective A-module. Applying Lemma 2.1 to the exact sequence

$$(F_2/F_1)^* \to M_2^* \to M_1^* \to 0,$$

we obtain an A-homomorphism

$$\widetilde{\Phi} : \det((F_2/F_1)^*) \otimes_A \bigcap\nolimits_A^r M_2 \to \bigcap\nolimits_A^{r-s_2+s_1} M_1$$

for any nonnegative integer $r \geq s_2 - s_1$. Therefore we get the following map, which plays an important role in this paper.

Definition 2.3. For any cartesian diagram as above and a nonnegative integer $r \geq s_2 - s_1$, we define an A-homomorphism

$$\begin{aligned}\Phi : \det(F_2^*) \otimes_A \bigcap\nolimits_A^r M_2 &\xrightarrow{\sim} \det((F_2/F_1)^*) \otimes_A \det(F_1^*) \otimes_A \bigcap\nolimits_A^r M_2 \\ &\longrightarrow \det(F_1^*) \otimes_A \bigcap\nolimits_A^{r-s_2+s_1} M_1,\end{aligned}$$

where the first map is induced by the isomorphism

$$\det((F_2/F_1)^*) \otimes_A \det(F_1^*) \xrightarrow{\sim} \det(F_2^*), \quad a \otimes b \longrightarrow a \wedge \tilde{b},$$

where $\tilde{b}$ is a lift of b in $\bigwedge_A^{s_1} F_2^*$ and the second map is induced by $\widetilde{\Phi}$.

The map Φ is a generalization of the map defined in [Mazur and Rubin 2016, Proposition A.1]. Furthermore, we have the following proposition which is a generalization of [loc. cit., Proposition A.2] to zero-dimensional Gorenstein local rings.

Proposition 2.4. *Let A be a zero-dimensional Gorenstein local ring. Suppose that we have the following commutative diagram of A-modules*:

$$\begin{tikzcd} M_1 \arrow[r, hook] \arrow[d] & M_2 \arrow[r, hook] \arrow[d] & M_3 \arrow[d] \\ F_1 \arrow[r, hook] & F_2 \arrow[r, hook] & F_3 \end{tikzcd}$$

where F_1, F_2, and F_3 are free of finite rank and the two squares are cartesian. Let $s_i = \mathrm{rank}_A(F_i)$ and $r \geq s_3 - s_1$. If $i, j \in \{1, 2, 3\}$ and $j < i$, we denote by

$$\Phi_{ij} : \det(F_i^*) \otimes_A \bigcap\nolimits_A^{r-s_3+s_i} M_i \to \det(F_j^*) \otimes_A \bigcap\nolimits_A^{r-s_3+s_j} M_j$$

the map given by Definition 2.3. Then we have

$$\Phi_{31} = \Phi_{21} \circ \Phi_{32}.$$

Proof. We may assume that $F_1 = 0$. Hence we have $s_1 = 0$. For $i \in \{1, 2\}$, applying Lemma 2.1 to the exact sequence $(F_{i+1}/F_i)^* \xrightarrow{h_i} M_{i+1}^* \xrightarrow{g_i} M_i^* \to 0$, we have a map

$$\phi_i : \det((F_{i+1}/F_i)^*) \otimes_A \bigwedge\nolimits_A^{r-s_3+s_i} M_i^* \to \bigwedge\nolimits_A^{r-s_3+s_{i+1}} M_{i+1}^*.$$

Let

$$\phi_3 : \det(F_3^*) \otimes_A \bigwedge\nolimits_A^{r-s_3} M_1^* \to \bigwedge\nolimits_A^{r} M_3^*$$

be the map defined by the exact sequence $F_3^* \xrightarrow{h_3} M_3^* \xrightarrow{g_3} M_1^* \to 0$ using Lemma 2.1. Put $q : F_3^* \to F_2^*$. Let $x \in \det((F_3/F_2)^*)$, $y \in \bigwedge_A^{s_2} F_3^*$, and $z \in \bigwedge_A^{r-s_3} M_3^*$. Then we compute

$$\begin{aligned}(\phi_2 \circ (\mathrm{id} \otimes \phi_1))(x \otimes \wedge^{s_2} q(y) \otimes \wedge^{r-s_3} g_3(z)) &= \phi_2\bigl(x \otimes ((\wedge^{s_2}(h_1 \circ q)(y)) \wedge (\wedge^{r-s_3} g_2(z)))\bigr) \\ &= (\wedge^{s_3-s_2} h_2(x)) \wedge (\wedge^{s_2} h_3(y)) \wedge z \\ &= (\wedge^{s_3} h_3(x \wedge y)) \wedge z \\ &= \phi_3\bigl((x \wedge y) \otimes \wedge^{r-s_3} g_3(z)\bigr).\end{aligned}$$

Since the image of $x \otimes \wedge^{s_2} q(y)$ under the isomorphism

$$\det((F_3/F_2)^*) \otimes_A \det(F_2^*) \xrightarrow{\sim} \det(F_3^*)$$

in Definition 2.3 is $x \wedge y$, by taking the dual, we get the desired equality. □

3. Selmer structures

In this section, we review the results of [Mazur and Rubin 2004; 2016]. Recall that p is an odd prime number, K is a number field, $(R, \mathfrak{m}_R)$ is a complete noetherian local ring with finite residue field $\Bbbk := R/\mathfrak{m}_R$ of characteristic p, and T is a free R-module of finite rank with an R-linear continuous G_K-action which is unramified outside a finite set of places of K.

Throughout this paper, we assume that the field $\overline{K}$ is contained in the complex number field $\mathbb{C}$ and that the fixed separable closure $\overline{K}_{\mathfrak{q}}$ of $K_{\mathfrak{q}}$ contains $\overline{K}$ for each prime $\mathfrak{q}$ of K. Let $K(\mathfrak{q})$ denote the p-part of the ray class field of K modulo $\mathfrak{q}$ and $K(\mathfrak{q})_{\mathfrak{q}}$ the closure of $K(\mathfrak{q})$ in $\overline{K}_{\mathfrak{q}}$. Let $\mathcal{D}_{\mathfrak{q}} := \mathrm{Gal}(\overline{K}_{\mathfrak{q}}/K_{\mathfrak{q}})$ be the decomposition group at $\mathfrak{q}$ in G_K. Set $H^1(K_{\mathfrak{q}}, T) = H^1(\mathcal{D}_{\mathfrak{q}}, T)$. Let $\mathcal{I}_{\mathfrak{q}} \subseteq \mathcal{D}_{\mathfrak{q}}$ be the inertia group at $\mathfrak{q}$ and $\mathrm{Fr}_{\mathfrak{q}}$ the Frobenius element of $\mathcal{D}_{\mathfrak{q}}/\mathcal{I}_{\mathfrak{q}}$. We write $\mathrm{loc}_{\mathfrak{q}}$ for the localization map $H^1(K, T) \to H^1(K_{\mathfrak{q}}, T)$. Define

$$H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, T) := \ker(H^1(K_{\mathfrak{q}}, T) \to H^1(\mathcal{I}_{\mathfrak{q}}, T)),$$
$$H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T) := \ker(H^1(K_{\mathfrak{q}}, T) \to H^1(K(\mathfrak{q})_{\mathfrak{q}}, T))$$

for each prime $\mathfrak{q}$ of K. Furthermore, we set $H^1_f(K_{\mathfrak{q}}, T) = H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, T)$ if T is unramified at a prime $\mathfrak{q}$ of K.

Definition 3.1. A Selmer structure $\mathcal{F}$ on T is a collection of the following data:

- a finite set $\Sigma(\mathcal{F})$ of places of K, including all the infinite places, all the primes above p, and all the primes where T is ramified,
- a choice of R-submodule $H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T) \subseteq H^1(K_{\mathfrak{q}}, T)$ for each $\mathfrak{q} \in \Sigma(\mathcal{F})$.

Put $H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T) := H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, T)$ for each prime $\mathfrak{q} \notin \Sigma(\mathcal{F})$. We call $H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T)$ the local condition of $\mathcal{F}$ at a prime $\mathfrak{q}$ of K.

Since p is an odd prime, we have $H^1(K_v, T) = 0$ for any infinite place v of K. Thus we ignore the local condition at any infinite place of K.

Remark 3.2. Let $\mathcal{F}$ be a Selmer structure on T and $R \to S$ a surjective ring homomorphism. Then $\mathcal{F}$ induces a Selmer structure on the S-module $T \otimes_R S$, which we will denote by $\mathcal{F}_S$. If there is no risk of confusion, then we write $\mathcal{F}$ instead of $\mathcal{F}_S$.

Definition 3.3. Let $\mathcal{F}$ be a Selmer structure on T. Set

$$H^1_{/\mathcal{F}}(K_{\mathfrak{q}}, T) := H^1(K_{\mathfrak{q}}, T)/H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T)$$

for any prime $\mathfrak{q}$ of K. We define the Selmer group $H^1_{\mathcal{F}}(K, T) \subseteq H^1(K, T)$ associated with $\mathcal{F}$ to be the kernel of the direct sum of localization maps

$$H^1_{\mathcal{F}}(K, T) := \ker\left(H^1(K, T) \to \bigoplus_{\mathfrak{q}} H^1_{/\mathcal{F}}(K_{\mathfrak{q}}, T)\right),$$

where $\mathfrak{q}$ runs through all the primes of K.

Example 3.4. Suppose that R is p-torsion-free. Then we define a canonical Selmer structure $\mathcal{F}_{\mathrm{can}}$ on T by the following data:

- $\Sigma(\mathcal{F}_{\mathrm{can}}) := \{\mathfrak{q} \mid T \text{ is ramified at } \mathfrak{q}\} \cup \{\mathfrak{p} \mid p\} \cup \{v \mid \infty\}$.
- We define a local condition at a prime $\mathfrak{q} \nmid p$ to be
$$H^1_{\mathcal{F}_{\mathrm{can}}}(K_{\mathfrak{q}}, T) := H^1_f(K_{\mathfrak{q}}, T) := \ker(H^1(K_{\mathfrak{q}}, T) \to H^1(\mathcal{I}_{\mathfrak{q}}, T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)).$$
- We define a local condition at a prime $\mathfrak{p} \mid p$ to be $H^1_{\mathcal{F}_{\mathrm{can}}}(K_{\mathfrak{p}}, T) := H^1(K_{\mathfrak{p}}, T)$.

Note that $H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, T) = H^1_f(K_{\mathfrak{q}}, T)$ for any prime $\mathfrak{q}$ where T is unramified since the map $H^1(\mathcal{I}_{\mathfrak{q}}, T) = \mathrm{Hom}(\mathcal{I}_{\mathfrak{q}}, T) \to \mathrm{Hom}(\mathcal{I}_{\mathfrak{q}}, T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p) = H^1(\mathcal{I}_{\mathfrak{q}}, T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$ is injective.

Definition 3.5. Let $\mathcal{F}$ be a Selmer structure on T and let $\mathfrak{a}$, $\mathfrak{b}$, and $\mathfrak{c}$ be pairwise relatively prime integral ideals of K. Define a Selmer structure $\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})$ on T by the following data:

- $\Sigma(\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})) := \Sigma(\mathcal{F}) \cup \{\mathfrak{q} \mid \mathfrak{abc}\}$.
- Define $H^1_{\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T) := \begin{cases} H^1(K_{\mathfrak{q}}, T) & \text{if } \mathfrak{q} \mid \mathfrak{a}, \\ 0 & \text{if } \mathfrak{q} \mid \mathfrak{b}, \\ H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T) & \text{if } \mathfrak{q} \mid \mathfrak{c}, \\ H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T) & \text{otherwise.} \end{cases}$

Definition 3.6. For any positive integer n, let μ_{p^n} denote the group of all p^n-th roots of unity. The Cartier dual of T is defined by

$$T^\vee(1) := T^\vee \otimes_{\mathbb{Z}_p} \varprojlim_n \mu_{p^n} = \mathrm{Hom}_{\mathrm{cont}}(T, \mathbb{Q}_p/\mathbb{Z}_p) \otimes_{\mathbb{Z}_p} \varprojlim_n \mu_{p^n}.$$

Then we have the local Tate paring

$$\langle \cdot, \cdot \rangle_{\mathfrak{q}} : H^1(K_{\mathfrak{q}}, T) \times H^1(K_{\mathfrak{q}}, T^\vee(1)) \to \mathbb{Q}_p/\mathbb{Z}_p$$

for each prime $\mathfrak{q}$ of K. Let $\mathcal{F}$ be a Selmer structure on T. Put

$$H^1_{\mathcal{F}^*}(K_{\mathfrak{q}}, T^\vee(1)) := \{x \in H^1(K_{\mathfrak{q}}, T^\vee(1)) \mid \langle y, x \rangle_{\mathfrak{q}} = 0 \text{ for any } y \in H^1_{\mathcal{F}}(K_{\mathfrak{q}}, T)\}.$$

In this manner, the Selmer structure $\mathcal{F}$ on T gives rise to a Selmer structure $\mathcal{F}^*$ on $T^\vee(1)$. We define the dual Selmer group $H^1_{\mathcal{F}^*}(K, T^\vee(1))$ associated with $\mathcal{F}$ to be the kernel of the sum of localization maps

$$H^1_{\mathcal{F}^*}(K, T^\vee(1)) := \ker\Bigl(H^1(K, T^\vee(1)) \to \bigoplus_{\mathfrak{q}} H^1_{/\mathcal{F}^*}(K_{\mathfrak{q}}, T^\vee(1))\Bigr),$$

where $\mathfrak{q}$ runs through all the primes of K.

Theorem 3.7 (Poitou–Tate global duality). *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be Selmer structures on T. If $H^1_{\mathcal{F}_1}(K_{\mathfrak{q}}, T) \subseteq H^1_{\mathcal{F}_2}(K_{\mathfrak{q}}, T)$ for any prime $\mathfrak{q}$ of K, then we have the exact sequence*

$$\begin{aligned} 0 \to H^1_{\mathcal{F}_1}(K, T) \to H^1_{\mathcal{F}_2}(K, T) \to \bigoplus_{\mathfrak{q}} H^1_{\mathcal{F}_2}(K_{\mathfrak{q}}, T)/H^1_{\mathcal{F}_1}(K_{\mathfrak{q}}, T) \\ \to H^1_{\mathcal{F}_1^*}(K, T^\vee(1))^\vee \to H^1_{\mathcal{F}_2^*}(K, T^\vee(1))^\vee \to 0 \end{aligned}$$

where $\mathfrak{q}$ runs through all the primes of K which satisfy $H^1_{\mathcal{F}_1}(K_{\mathfrak{q}}, T) \neq H^1_{\mathcal{F}_2}(K_{\mathfrak{q}}, T)$.

Proof. This is [Mazur and Rubin 2004, Theorem 2.3.4]. □

Definition 3.8 (cartesian condition). Suppose that R is a zero-dimensional Gorenstein local ring. We fix an injective map $\Bbbk \to R$. It induces an injective map $T/\mathfrak{m}_R T \to T$. We say that a Selmer structure $\mathcal{F}$ on T is cartesian if the map

$$H^1_{/\mathcal{F}_{\Bbbk}}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}}(K_{\mathfrak{q}}, T)$$

induced by the map $T/\mathfrak{m}_R T \to T$ is injective for any prime $\mathfrak{q} \in \Sigma(\mathcal{F})$.

Remark 3.9. When R is a principal artinian local ring, the definitions of a cartesian Selmer structure in this paper and [Mazur and Rubin 2004] are equivalent. Furthermore, under the setting of [Büyükboduk 2011], we see that the conditions (C2) and (C3) in Definition 2.5 of that paper are equivalent to the cartesian condition in the sense of this paper.

Remark 3.10. Suppose that R is a zero-dimensional Gorenstein local ring. By the definition of Gorenstein ring, we have $\dim_{\Bbbk} \mathrm{Hom}_R(\Bbbk, R) = 1$. Hence the definition of cartesian condition is independent of the choice of the injective map $\Bbbk \to R$.

Let R and S be zero-dimensional Gorenstein local rings and $\pi : R \to S$ a surjective ring homomorphism. Put $I = \ker(\pi)$. Since R is an injective R-module, $R[I] \simeq \mathrm{Hom}_R(S, R)$ is an injective S-module. Hence we conclude that there is an isomorphism $S \xrightarrow{\sim} R[I]$ as R-modules since S is an injective S-module. In particular, there is an injective R-module homomorphism $S \to R$.

Lemma 3.11. *Let R and S be zero-dimensional Gorenstein local rings and $\pi : R \to S$ a surjective ring homomorphism. If a Selmer structure $\mathcal{F}$ on T is cartesian, then so is $\mathcal{F}_S$.*

Proof. Let $S \to R$ and $\Bbbk \to S$ be injective R-module homomorphisms and $\mathfrak{q} \in \Sigma(\mathcal{F})$. Then these maps induce maps $H^1_{/\mathcal{F}_S}(K_\mathfrak{q}, T \otimes_R S) \to H^1_{/\mathcal{F}}(K_\mathfrak{q}, T)$ and $H^1_{/\mathcal{F}_\Bbbk}(K_\mathfrak{q}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}_S}(K_\mathfrak{q}, T \otimes_R S)$. By the definition of cartesian condition and Remark 3.10, the composition of maps

$$H^1_{/\mathcal{F}_\Bbbk}(K_\mathfrak{q}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}_S}(K_\mathfrak{q}, T \otimes_R S) \to H^1_{/\mathcal{F}}(K_\mathfrak{q}, T)$$

is injective, and thus so is $H^1_{/\mathcal{F}_\Bbbk}(K_\mathfrak{q}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}_S}(K_\mathfrak{q}, T \otimes_R S)$. □

Let $\mathcal{H}$ be the Hilbert class field of K and $\mathcal{O}_K$ the ring of integers of K. Set $\mu_{p^\infty} := \bigcup_{n \geq 0} \mu_{p^n}$ and $\mathcal{H}_\infty := \mathcal{H}(\mu_{p^\infty}, (\mathcal{O}_K^\times)^{p^{-\infty}})$. Here $(\mathcal{O}_K^\times)^{p^{-n}} := \{x \in \overline{K} \mid x^{p^n} \in \mathcal{O}_K^\times\}$ and $(\mathcal{O}_K^\times)^{p^{-\infty}} := \bigcup_{n \geq 0} (\mathcal{O}_K^\times)^{p^{-n}}$. In order to use the results of [Mazur and Rubin 2004; 2016], we will usually assume the following additional conditions.

Hypothesis 3.12. (H.1) $(T/\mathfrak{m}_R T)^{G_K} = (T^\vee(1)[\mathfrak{m}_R])^{G_K} = 0$ and $T/\mathfrak{m}_R T$ is an irreducible $\Bbbk[G_K]$-module.

(H.2) There is a $\tau \in \mathrm{Gal}(\overline{K}/\mathcal{H}_\infty)$ such that $T/(\tau - 1)T \simeq R$ as R-modules.

(H.3) $H^1(\mathcal{H}_\infty(T)/K, T/\mathfrak{m}_R T) = H^1(\mathcal{H}_\infty(T)/K, T^\vee(1)[\mathfrak{m}_R]) = 0$, where $\mathcal{H}_\infty(T)$ is the fixed field of the kernel of the map $\mathrm{Gal}(\overline{K}/\mathcal{H}_\infty) \to \mathrm{Aut}(T)$.

Lemma 3.13. *Suppose that R is a zero-dimensional Gorenstein local ring and $(T/\mathfrak{m}_R T)^{G_K} = 0$. Let $\mathcal{F}$ be a cartesian Selmer structure on T. Then an injective map $\Bbbk \to R$ induces isomorphisms*

(1) $H^1(K, T/\mathfrak{m}_R T) \xrightarrow{\sim} H^1(K, T)[\mathfrak{m}_R]$,

(2) $H^1_\mathcal{F}(K, T/\mathfrak{m}_R T) \xrightarrow{\sim} H^1_\mathcal{F}(K, T)[\mathfrak{m}_R]$.

Proof. Since R is a zero-dimensional Gorenstein local ring, an injective map $\Bbbk \to R$ induces an isomorphism $T/\mathfrak{m}_R T \xrightarrow{\sim} T[\mathfrak{m}_R]$ as $R[G_K]$-modules. Hence the proof of assertion (1) is the same as the first part of the proof of [Mazur and Rubin 2004, Lemma 3.5.3]. Furthermore, the proof of assertion (2) is the same as the last part of the proof of [loc. cit., Lemma 3.5.3] since $\mathcal{F}$ is a cartesian Selmer structure. □

Lemma 3.14. *Let $\mathcal{F}$ be a Selmer structure on T and I an ideal of R. Suppose that R is an artinian local ring and $(T^\vee(1)[\mathfrak{m}_R])^{G_K}=0$. Then the inclusion map $T^\vee(1)[I]\to T^\vee(1)$ induces isomorphisms*

(1) $H^1(K,T^\vee(1)[I]) \xrightarrow{\sim} H^1(K,T^\vee(1))[I]$,

(2) $H^1_{\mathcal{F}^*}(K,T^\vee(1)[I]) \xrightarrow{\sim} H^1_{\mathcal{F}^*}(K,T^\vee(1))[I]$.

Proof. This is [Mazur and Rubin 2004, Lemma 3.5.3]. □

Definition 3.15. Let $\mathcal{F}$ be a Selmer structure on T and n a positive integer. Define a set $\mathcal{P}(\mathcal{F})$ of primes of K by

$$\mathcal{P}(\mathcal{F}) := \{\mathfrak{q} \mid \mathfrak{q}\notin\Sigma(\mathcal{F})\}.$$

Let τ be as in (H.2) and q the order of the residue field $\Bbbk$ of R. Set $\mathcal{H}_n := \mathcal{H}(\mu_{q^n}, (\mathcal{O}_K^\times)^{q^{-n}})$. Here $(\mathcal{O}_K^\times)^{q^{-n}} := \{x\in\overline{K}\mid x^{q^n}\in\mathcal{O}_K^\times\}$. Note that $\mathfrak{q}\in\mathcal{P}(\mathcal{F})$ is unramified in $\mathcal{H}_n(T/\mathfrak{m}_R^nT)/K$. We define a set $\mathcal{P}_n(\mathcal{F})$ of primes of K as

$$\mathcal{P}_n(\mathcal{F}) := \{\mathfrak{q}\in\mathcal{P}(\mathcal{F}) \mid \mathrm{Fr}_\mathfrak{q} \text{ is conjugate to } \tau \text{ in } \mathrm{Gal}(\mathcal{H}_n(T/\mathfrak{m}_R^nT)/K)\}.$$

Remark 3.16. If $\pi: R\to S$ is a surjective ring homomorphism, then we have $\mathcal{P}_n(\mathcal{F})\subseteq\mathcal{P}_n(\mathcal{F}_S)$ for any positive integer n.

Let $\mathfrak{q}$ be a prime of K where T is unramified. We define the singular quotient at $\mathfrak{q}$ as

$$H^1_{/f}(K_\mathfrak{q},T) := H^1(K_\mathfrak{q},T)/H^1_f(K_\mathfrak{q},T).$$

We denote by $\mathrm{loc}_\mathfrak{q}^{/f}: H^1(K,T)\to H^1_{/f}(K_\mathfrak{q},T)$ the composition of the localization map at $\mathfrak{q}$ and a surjective map $H^1(K_\mathfrak{q},T)\to H^1_{/f}(K_\mathfrak{q},T)$.

Let $\mathcal{N}(\mathcal{P})$ denote the set of square-free products of primes in a set $\mathcal{P}$ of primes of K. By considering the empty product, the trivial ideal 1 is contained in $\mathcal{N}(\mathcal{P})$.

Lemma 3.17. *Suppose that R is an artinian ring and T satisfies (H.2). If $\mathfrak{q}\in\mathcal{P}_{\ell_R(R)}(\mathcal{F})$, then both $H^1_f(K_\mathfrak{q},T)$ and $H^1_{/f}(K_\mathfrak{q},T)$ are free R-modules of rank 1 and the composition of maps*

$$H^1_{\mathrm{tr}}(K_\mathfrak{q},T)\to H^1(K_\mathfrak{q},T)\xrightarrow{\mathrm{loc}_\mathfrak{q}^{/f}} H^1_{/f}(K_\mathfrak{q},T)$$

is an isomorphism. Furthermore, we have canonical isomorphisms

(1) $H^1_f(K_\mathfrak{q},T)\otimes_R R/I \xrightarrow{\sim} H^1_f(K_\mathfrak{q},T/IT)$,

(2) $H^1_{/f}(K_\mathfrak{q},T)\otimes_R R/I \xrightarrow{\sim} H^1_{/f}(K_\mathfrak{q},T/IT)$

for any ideal I of R. In particular, the natural map $H^1(K_\mathfrak{q},T)\otimes_R R/I\to H^1(K_\mathfrak{q},T/IT)$ is an isomorphism.

Proof. This lemma follows from [Mazur and Rubin 2004, Lemmas 1.2.3 and 1.2.4]. □

Corollary 3.18. *Suppose that R is a zero-dimensional Gorenstein local ring and T satisfies (H.2). Let $\mathfrak{a},\mathfrak{b},\mathfrak{c}\in\mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$ with $\mathfrak{abc}\in\mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$. If a Selmer structure $\mathcal{F}$ on T is cartesian, then so is $\mathcal{F}_\mathfrak{b}^\mathfrak{a}(\mathfrak{c})$.*

Proof. We fix an injective map $i:\Bbbk \to R$. Since $\mathcal{F}$ is cartesian, we only need to show that the map

$$H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T)$$

induced by the map $i:\Bbbk \to R$ is injective for any prime $\mathfrak{q} \mid \mathfrak{abc}$ of K. By the definition of the Selmer structure $\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})$ and Lemma 3.17, the map $H^1(K_{\mathfrak{q}}, T) \to H^1(K_{\mathfrak{q}}, T/\mathfrak{m}_R T)$ induces an isomorphism $H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T) \otimes_R \Bbbk \xrightarrow{\sim} H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T)$. Since $H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T)$ is a free R-module by Lemma 3.17 and the composition of the maps

$$H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T) \otimes_R \Bbbk \xrightarrow{\sim} H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T) \longrightarrow H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T)$$

is $\mathrm{id}_{H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T)} \otimes i$, the map $H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T) \to H^1_{/\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})}(K_{\mathfrak{q}}, T)$ is injective. □

Definition 3.19. We define the core rank $\chi(\mathcal{F})$ of a Selmer structure $\mathcal{F}$ on T by

$$\chi(\mathcal{F}) := \dim_{\Bbbk} H^1_{\mathcal{F}}(K, T/\mathfrak{m}_R T) - \dim_{\Bbbk} H^1_{\mathcal{F}^*}(K, T^\vee(1)[\mathfrak{m}_R]).$$

Example 3.20. Let R be the ring of integers of a finite extension of the field $\mathbb{Q}_p$. Suppose that $(T/\mathfrak{m}_R T)^{G_K} = (T^\vee(1)[\mathfrak{m}_R])^{G_K} = 0$. Then by the same proof as [Mazur and Rubin 2004, Theorem 5.2.15], we have

$$\chi(\mathcal{F}_{\mathrm{can}}) = \sum_{v \mid \infty} \mathrm{rank}_R(H^0(K_v, T^\vee(1))^\vee) + \sum_{\mathfrak{p} \mid p} \mathrm{rank}_R(H^2(K_{\mathfrak{q}}, T)),$$

where $\mathcal{F}_{\mathrm{can}}$ is the canonical Selmer structure defined in Example 3.4.

Corollary 3.21. *Suppose that R is an artinian local ring and T satisfies (H.2). Let $\mathcal{F}$ be a Selmer structure on T and $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} \in \mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$ with $\mathfrak{abc} \in \mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$. Then we have*

$$\chi(\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})) = \chi(\mathcal{F}) + \nu(\mathfrak{a}) - \nu(\mathfrak{b}),$$

where $\nu(\mathfrak{n})$ denotes the number of prime factors of $\mathfrak{n} \in \mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$.

Proof. Applying Theorem 3.7 with $\mathcal{F}_1 = \mathcal{F}$ and $\mathcal{F}_2 = \mathcal{F}^{\mathfrak{ac}}$, we have an exact sequence

$$0 \to H^1_{\mathcal{F}}(K, T/\mathfrak{m}_R T) \to H^1_{\mathcal{F}^{\mathfrak{ac}}}(K, T/\mathfrak{m}_R T) \to \bigoplus_{\mathfrak{q} \mid \mathfrak{ac}} H^1_{/f}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T)$$
$$\to H^1_{\mathcal{F}^*}(K, (T/\mathfrak{m}_R T)^\vee(1))^\vee \to H^1_{(\mathcal{F}^{\mathfrak{ac}})^*}(K, (T/\mathfrak{m}_R T)^\vee(1))^\vee \to 0.$$

Since $(T/\mathfrak{m}_R T)^\vee = T^\vee[\mathfrak{m}_R]$, by the definition of core rank and Lemma 3.17, we have

$$\chi(\mathcal{F}^{\mathfrak{ac}}) = \chi(\mathcal{F}) + \sum_{\mathfrak{q} \mid \mathfrak{ac}} \dim_{\Bbbk} H^1_{/f}(K_{\mathfrak{q}}, T/\mathfrak{m}_R T) = \chi(\mathcal{F}) + \nu(\mathfrak{a}) + \nu(\mathfrak{c}).$$

Again using Lemma 3.17 and Theorem 3.7 with $\mathcal{F}_1 = \mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})$ and $\mathcal{F}_2 = \mathcal{F}^{\mathfrak{ac}}$, we see that $\chi(\mathcal{F}^{\mathfrak{a}}_{\mathfrak{b}}(\mathfrak{c})) = \chi(\mathcal{F}^{\mathfrak{ac}}) - \nu(\mathfrak{b}) - \nu(\mathfrak{c}) = \chi(\mathcal{F}) + \nu(\mathfrak{a}) - \nu(\mathfrak{b})$. □

Proposition 3.22. *Suppose that R is an artinian local ring and T satisfies Hypothesis 3.12. Let $c \in H^1(K, T)$ and $d \in H^1(K, T^\vee(1))$ be nonzero elements and n a positive integer. Then there is a set $\mathcal{Q} \subseteq \mathcal{P}_n(\mathcal{F})$ of positive density such that $\mathrm{loc}_{\mathfrak{q}}(c) \neq 0$ and $\mathrm{loc}_{\mathfrak{q}}(d) \neq 0$ for every $\mathfrak{q} \in \mathcal{Q}$.*

Proof. The proof of this proposition is the same as that of [Mazur and Rubin 2004, Proposition 3.6.1]. □

Remark 3.23. Proposition 3.22 is a weaker version of [Mazur and Rubin 2004, Proposition 3.6.1]. Thus we do not need the assumption (H.4) introduced in [Mazur and Rubin 2004; 2016] to prove Proposition 3.22.

Lemma 3.24. *Suppose that R is an artinian local ring and T satisfies Hypothesis 3.12. Let $c \in H^1(K, T)$ with $Rc \simeq R$ as R-modules and n a positive integer. Then there are infinitely many primes $\mathfrak{q} \in \mathcal{P}_n(\mathcal{F})$ such that $R \cdot \mathrm{loc}_{\mathfrak{q}}(c) = H^1_f(K_{\mathfrak{q}}, T)$.*

Proof. We may assume that $n > \ell_R(R)$. Note that $\mathrm{loc}_{\mathfrak{q}}(c) \in H^1_f(K_{\mathfrak{q}}, T)$ for all but finitely many primes $\mathfrak{q}$ of K and that $H^1_f(K_{\mathfrak{q}}, T)$ is free of rank 1 for any prime $\mathfrak{q} \in \mathcal{P}_n(\mathcal{F})$ by Lemma 3.17.

Since R is an artinian local ring, there is an element $m \in R$ such that $\mathfrak{m}_R = R[m]$. Put $c' := mc$. Since $Rc \simeq R$ as R-modules, the element c' is nonzero. Hence by Proposition 3.22, there are infinitely many primes $\mathfrak{q} \in \mathcal{P}_n(\mathcal{F})$ such that $\mathrm{loc}_{\mathfrak{q}}(c') \neq 0$ and $\mathrm{loc}_{\mathfrak{q}}(c) \in H^1_f(K_{\mathfrak{q}}, T)$. Since $\mathfrak{m}_R = R[m]$ and $H^1_f(K_{\mathfrak{q}}, T) \simeq R$ as R-modules, we have $R \cdot \mathrm{loc}_{\mathfrak{q}}(c) = H^1_f(K_{\mathfrak{q}}, T)$. □

Lemma 3.25. *Suppose that R is an artinian local ring and T satisfies Hypothesis 3.12. Let M be a free R-submodule of $H^1(K, T)$ of rank $s \geq 0$ and $R \to A$ a ring homomorphism. Then the composition of maps*

$$M \otimes_R A \to H^1(K, T) \otimes_R A \to H^1(K, T \otimes_R A)$$

is a split injection. Here $T \otimes_R A$ is equipped with the discrete topology.

Proof. Set $n = l_R(R)$. By Lemma 3.24, there are primes $\mathfrak{q}_1, \ldots, \mathfrak{q}_s \in \mathcal{P}_n(\mathcal{F})$ such that the sum of localization maps $\bigoplus_{i=1}^s \mathrm{loc}_{\mathfrak{q}_i} : H^1(K, T) \to \bigoplus_{i=1}^s H^1(K_{\mathfrak{q}_i}, T)$ induces an isomorphism $M \xrightarrow{\sim} \bigoplus_{i=1}^s H^1_f(K_{\mathfrak{q}}, T)$. For any $1 \leq i \leq s$, T is unramified at $\mathfrak{q}_i$, which implies

$$\begin{aligned} H^1_f(K_{\mathfrak{q}_i}, T) \otimes_R A &\xrightarrow{\sim} (T/(\mathrm{Fr}_{\mathfrak{q}_i} - 1)T) \otimes_R A \\ &\xrightarrow{\sim} (T \otimes_R A)/((\mathrm{Fr}_{\mathfrak{q}_i} - 1)(T \otimes_R A)) \\ &\xleftarrow{\sim} H^1_f(K_{\mathfrak{q}_i}, T \otimes_R A). \end{aligned}$$

Therefore the natural map $M \otimes_R A \to \bigoplus_{i=1}^s H^1_f(K_{\mathfrak{q}}, T \otimes_R A)$ is an isomorphism. □

4. Stark systems over zero-dimensional Gorenstein local rings

We will now define a Stark system over a zero-dimensional Gorenstein local ring. Under Hypothesis 3.12, we will prove that the module of Stark systems associated with a cartesian Selmer structure is free of rank 1 and control all the higher Fitting ideals of the Pontryagin dual of dual Selmer groups using Stark systems (Theorems 4.7 and 4.10). Furthermore, we will show that these results generalize [Mazur and Rubin 2004; 2016].

Throughout this section, we assume that R is a zero-dimensional Gorenstein local ring and that $\mathcal{F}$ is a Selmer structure on T. Suppose that T satisfies Hypothesis 3.12.

For simplicity, we write $\mathcal{P} = \mathcal{P}_{\ell_R(R)}(\mathcal{F})$ and $\mathcal{N} = \mathcal{N}(\mathcal{P}_{\ell_R(R)}(\mathcal{F}))$. Let $\nu(\mathfrak{n})$ denote the number of prime factors of $\mathfrak{n} \in \mathcal{N}$.

Stark systems over zero-dimensional Gorenstein local rings. Let r be a nonnegative integer and $\mathfrak{n} \in \mathcal{N}$. Define $W_{\mathfrak{n}} := \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} H^1_{/f}(K_{\mathfrak{q}}, T)^* = \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} \mathrm{Hom}_R(H^1_{/f}(K_{\mathfrak{q}}, T), R)$ and

$$X_{\mathfrak{n}}(T, \mathcal{F}) := \det(W_{\mathfrak{n}}) \otimes_R \bigcap\nolimits_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T).$$

If there is no risk of confusion, $X_{\mathfrak{n}}(T, \mathcal{F})$ is abbreviated to $X_{\mathfrak{n}}$. For an ideal $\mathfrak{m} \mid \mathfrak{n}$, we have the following cartesian diagram:

$$\begin{tikzcd}
H^1_{\mathcal{F}^{\mathfrak{m}}}(K, T) \arrow[r, hook] \arrow[d] & H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \arrow[d] \\
\bigoplus_{\mathfrak{q} \mid \mathfrak{m}} H^1_{/f}(K_{\mathfrak{q}}, T) \arrow[r, hook] & \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} H^1_{/f}(K_{\mathfrak{q}}, T)
\end{tikzcd}$$

Hence by Definition 2.3, we obtain a map

$$\Phi_{\mathfrak{n},\mathfrak{m}} : X_{\mathfrak{n}} \to X_{\mathfrak{m}}.$$

Lemma 4.1. *We have* $\Phi_{\mathfrak{n}_1,\mathfrak{n}_3} = \Phi_{\mathfrak{n}_2,\mathfrak{n}_3} \circ \Phi_{\mathfrak{n}_1,\mathfrak{n}_2}$ *for any* $\mathfrak{n}_1 \in \mathcal{N}$, $\mathfrak{n}_2 \mid \mathfrak{n}_1$, *and* $\mathfrak{n}_3 \mid \mathfrak{n}_2$.

Proof. This lemma follows from Proposition 2.4. □

Definition 4.2. Let $\mathcal{Q}$ be a subset of $\mathcal{P}$. We define the module $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ of Stark systems of rank r for $(T, \mathcal{F}, \mathcal{Q})$ by

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) := \varprojlim_{\mathfrak{n} \in \mathcal{N}(\mathcal{Q})} X_{\mathfrak{n}},$$

where the inverse limit is taken with respect to the maps $\Phi_{\mathfrak{n},\mathfrak{m}}$. We call an element of $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ a Stark system of rank r associated with the tuple $(T, \mathcal{F}, \mathcal{Q})$.

Definition 4.3. We say that an ideal $\mathfrak{n} \in \mathcal{N}$ is a core vertex (for $\mathcal{F}$) if

$$H^1_{(\mathcal{F}^{\mathfrak{n}})^*}(K, T^\vee(1)) = H^1_{\mathcal{F}^*_{\mathfrak{n}}}(K, T^\vee(1)) = 0.$$

For any ideal $\mathfrak{n} \in \mathcal{N}$ we have $H^1_{\mathcal{F}^*_{\mathfrak{n}}}(K, T^\vee(1)) \subseteq H^1_{\mathcal{F}^*(\mathfrak{n})}(K, T^\vee(1))$. Hence a core vertex in the sense of [Mazur and Rubin 2004; 2016] is a core vertex in the sense of this paper.

Proposition 4.4. *The set* $\mathcal{N}(\mathcal{P}_n(\mathcal{F}))$ *has a core vertex for any positive integer* n.

Proof. Let n be a positive integer. If $H^1_{\mathcal{F}^*}(K, T^\vee(1)) \neq 0$, by Proposition 3.22, there is a prime $\mathfrak{q} \in \mathcal{P}_n(\mathcal{F})$ such that $\mathrm{loc}_{\mathfrak{q}}(H^1_{\mathcal{F}^*}(K, T^\vee(1))) \neq 0$. Then we have

$$\ell_R(H^1_{\mathcal{F}^*_{\mathfrak{q}}}(K, T^\vee(1))) < \ell_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))) < \infty.$$

By repeating this argument, we get an ideal $\mathfrak{n} \in \mathcal{N}(\mathcal{P}_n(\mathcal{F}))$ such that $H^1_{\mathcal{F}^*_{\mathfrak{n}}}(K, T^\vee(1)) = 0$. □

Remark 4.5. (1) If $\mathfrak{q} \in \mathcal{P}$ and $\mathfrak{n} \in \mathcal{N}(\mathcal{P} \setminus \{\mathfrak{q}\})$ is a core vertex, then $\mathfrak{n}\mathfrak{q} \in \mathcal{N}$ is also a core vertex.

(2) Let $\mathfrak{n} \in \mathcal{N}$ be a core vertex, S a zero-dimensional Gorenstein local ring, and $\pi : R \to S$ a surjective ring homomorphism. Then $\mathfrak{n}$ is also a core vertex for $\mathcal{F}_S$ by Lemma 3.14.

Lemma 4.6. *Suppose that the Selmer structure $\mathcal{F}$ on T is cartesian. If $\mathfrak{n} \in \mathcal{N}$ is a core vertex, then the R-module $H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ is free of rank $\chi(\mathcal{F}) + \nu(\mathfrak{n})$.*

Proof. Let $\mathfrak{n} \in \mathcal{N}$ be a core vertex. We take an ideal $\mathfrak{a} \in \mathcal{N}$ such that $(\mathfrak{a}, \mathfrak{n}) = 1$ and $\nu(\mathfrak{a}) = \chi(\mathcal{F})$. By Corollary 3.21, we have $\chi(\mathcal{F}_{\mathfrak{a}}) = \chi(\mathcal{F}) - \nu(\mathfrak{a}) = 0$. Then by [Mazur and Rubin 2016, Theorem 11.6], there is an ideal $\mathfrak{m} \in \mathcal{N}$ such that $(\mathfrak{a}\mathfrak{n}, \mathfrak{m}) = 1$ and $H^1_{\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})^*}(K, T^\vee(1)[\mathfrak{m}_R]) = 0$. Then we have

$$\dim_{\Bbbk} H^1_{\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})}(K, T/\mathfrak{m}_R T) = \chi(\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})) = \chi(\mathcal{F}_{\mathfrak{a}}) = 0,$$

where the first equality follows from $H^1_{\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})^*}(K, T^\vee(1)[\mathfrak{m}_R]) = 0$ and the second equality follows from Corollary 3.21. Since $\mathcal{F}$ is cartesian, it follows from Lemmas 3.13, 3.14, and Corollary 3.18 that

$$H^1_{\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})}(K, T) = H^1_{\mathcal{F}_{\mathfrak{a}}(\mathfrak{m})^*}(K, T^\vee(1)) = 0.$$

Applying Theorem 3.7 with $\mathcal{F}_1 = \mathcal{F}_{\mathfrak{a}}(\mathfrak{m})$ and $\mathcal{F}_2 = \mathcal{F}^{\mathfrak{m}\mathfrak{n}}$, we see that the canonical map

$$H^1_{\mathcal{F}^{\mathfrak{m}\mathfrak{n}}}(K, T) \to \bigoplus_{\mathfrak{q} \mid \mathfrak{a}} H^1_f(K_{\mathfrak{q}}, T) \oplus \bigoplus_{\mathfrak{q} \mid \mathfrak{m}} H^1(K_{\mathfrak{q}}, T)/H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T) \oplus \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} H^1_{/f}(K_{\mathfrak{q}}, T)$$

is an isomorphism. Hence by Lemma 3.17, $H^1_{\mathcal{F}^{\mathfrak{m}\mathfrak{n}}}(K, T)$ is a free R-module of rank $\chi(\mathcal{F}) + \nu(\mathfrak{m}) + \nu(\mathfrak{n})$. Since $\mathfrak{n}$ is a core vertex, by Theorem 3.7, we have an exact sequence

$$0 \to H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \to H^1_{\mathcal{F}^{\mathfrak{m}\mathfrak{n}}}(K, T) \to \bigoplus_{\mathfrak{q} \mid \mathfrak{m}} H^1_{/f}(K_{\mathfrak{q}}, T) \to 0.$$

Again by Lemma 3.17, $\bigoplus_{\mathfrak{q} \mid \mathfrak{m}} H^1_{/f}(K, T)$ is a free R-module of rank $\nu(\mathfrak{m})$, which implies $H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ is free of rank $\chi(\mathcal{F}) + \nu(\mathfrak{n})$. □

Theorem 4.7. *Suppose that $\mathcal{F}$ is cartesian and $r = \chi(\mathcal{F}) \geq 0$. Let $\mathcal{Q}$ be a subset of $\mathcal{P}$. If the set $\mathcal{N}(\mathcal{Q})$ has a core vertex, then the projection map*

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \to X_{\mathfrak{n}}(T, \mathcal{F})$$

is an isomorphism for any core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$. In particular, the module $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ is a free R-module of rank 1.

Proof. Let $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ be a core vertex. We only need to show that the map $\Phi_{\mathfrak{n}\mathfrak{q},\mathfrak{n}}$ is an isomorphism for any prime $\mathfrak{q} \in \mathcal{Q}$ with $\mathfrak{q} \nmid \mathfrak{n}$. Note that $\mathfrak{n}\mathfrak{q}$ is also a core vertex. By Theorem 3.7 and Lemma 4.6, we have the exact sequence of free R-modules

$$0 \to H^1_{/f}(K_{\mathfrak{q}}, T)^* \to H^1_{\mathcal{F}^{\mathfrak{n}\mathfrak{q}}}(K, T)^* \to H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)^* \to 0.$$

Since $\mathrm{rank}_R(H^1_{\mathcal{F}^{\mathfrak{n}\mathfrak{q}}}(K, T)) = r + \nu(\mathfrak{n}\mathfrak{q})$ and $\mathrm{rank}_R(H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)) = r + \nu(\mathfrak{n})$, the map

$$H^1_{/f}(K_{\mathfrak{q}}, T)^* \otimes_R {\textstyle\bigwedge_R^{r+\nu(\mathfrak{n})}} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)^* \to {\textstyle\bigwedge_R^{r+\nu(\mathfrak{n}\mathfrak{q})}} H^1_{\mathcal{F}^{\mathfrak{n}\mathfrak{q}}}(K, T)^*$$

defined by Lemma 2.1 is an isomorphism, and thus so is $\Phi_{\mathfrak{n}\mathfrak{q},\mathfrak{n}}$. □

Lemma 4.8. *Let s and t be nonnegative integers and*

$$0 \to N \to R^{s+t} \to R^s \to M \to 0$$

an exact sequence of R-modules. If $\phi \in \mathrm{Hom}_R\big(\bigwedge_R^t N^*, R\big)$ is the image of a basis of $\det((R^s)^*) \otimes_R \bigcap_R^{s+t} R^{s+t}$ under the map

$$\det((R^s)^*) \otimes_R \bigcap\nolimits_R^{s+t} R^{s+t} \to \bigcap\nolimits_R^{t} N = \mathrm{Hom}_R\Big(\bigwedge\nolimits_R^t N^*, R\Big)$$

defined in Definition 2.3, then we have $\mathrm{im}(\phi) = \mathrm{Fitt}_R^0(M)$.

Proof. Let $r_1, \ldots, r_{s+t}$ be the standard basis of R^{s+t} and $r'_1, \ldots, r'_s$ be the standard basis of R^s. We denote by $r_1^*, \ldots, r_{s+t}^*$ the dual basis of $r_1, \ldots, r_{s+t}$ and by $r_1'^*, \ldots, r_s'^*$ the dual basis of $r'_1, \ldots, r'_s$. Let $A = (a_{ij})$ be the $s \times (s+t)$-matrix defined by the map $R^{s+t} \to R^s$.

Applying Lemma 2.1 to the exact sequence $(R^s)^* \to (R^{s+t})^* \to N^* \to 0$, we have a map

$$\det((R^s)^*) \otimes_R \bigwedge\nolimits_R^t N^* \longrightarrow \bigwedge\nolimits_R^{s+t} (R^{s+t})^* \xrightarrow{\sim} R$$

and its image is equal to $\mathrm{im}(\phi)$. Put

$$\Psi : \det((R^s)^*) \otimes_R \bigwedge\nolimits_R^t (R^{s+t})^* \to \det((R^s)^*) \otimes_R \bigwedge\nolimits_R^t N^* \to R.$$

Since a natural map $(R^{s+t})^* \to N^*$ is surjective, we have $\mathrm{im}(\phi) = \mathrm{im}(\Psi)$. If $\{i_1 < \cdots < i_t\} \subseteq \{1, \ldots, s+t\}$, then we compute

$$\Psi((r_1'^* \wedge \cdots \wedge r_s'^*) \otimes (r_{i_1}^* \wedge \cdots \wedge r_{i_t}^*)) = \pm \det((a_{kj_\nu})_{k,\nu}),$$

where $\{j_1 < \cdots < j_s\}$ is the complement of $\{i_1 < \cdots < i_t\}$ in $\{1, \ldots, s+t\}$. By the definition of Fitting ideals, we conclude that $\mathrm{im}(\phi) = \mathrm{im}(\Psi) = \mathrm{Fitt}_R^0(M)$. □

Definition 4.9. Let $\mathcal{Q}$ be a subset of $\mathcal{P}$ and $\epsilon \in \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$. Fix an isomorphism $H^1_{/f}(K_{\mathfrak{q}}, T) \simeq R$ for each prime $\mathfrak{q} \in \mathcal{Q}$. Using these isomorphisms, we have an isomorphism

$$i_{\mathfrak{n}} : X_{\mathfrak{n}} \xrightarrow{\sim} \bigcap\nolimits_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) = \mathrm{Hom}_R\Big(\bigwedge\nolimits_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)^*, R\Big)$$

for each ideal $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$. Then we define an ideal $I_i(\epsilon) \subseteq R$ by

$$I_i(\epsilon) := \sum_{\mathfrak{n} \in \mathcal{N}(\mathcal{Q}), \nu(\mathfrak{n}) = i} \mathrm{im}(i_{\mathfrak{n}}(\epsilon_{\mathfrak{n}}))$$

for any nonnegative integer i. It is easy to see that the ideal $I_i(\epsilon)$ is independent of the choice of the isomorphisms $H^1_{/f}(K_{\mathfrak{q}}, T) \simeq R$.

Theorem 4.10. *Suppose that $\mathcal{F}$ is cartesian and $r = \chi(\mathcal{F}) \geq 0$. Let $\mathcal{Q}$ be an infinite subset of $\mathcal{P}$ such that $\mathcal{N}(\mathcal{Q})$ has a core vertex. If $\epsilon = \{\epsilon_{\mathfrak{n}}\}_{\mathfrak{n} \in \mathcal{N}(\mathcal{Q})}$ is a basis of $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$, then we have*

$$I_i(\epsilon) = \mathrm{Fitt}_R^i(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee)$$

for any nonnegative integer i.

Proof. Let i be a nonnegative integer. Since $\mathcal{N}(\mathcal{Q})$ has a core vertex and $\mathcal{Q}$ is an infinite set, we can take a core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ with $\nu(\mathfrak{n}) \geq i$. Fix an isomorphism $H^1_{/f}(K_{\mathfrak{q}}, T) \simeq R$ for each prime $\mathfrak{q} \in \mathcal{Q}$. These isomorphisms induce an isomorphism

$$i_{\mathfrak{m}} : X_{\mathfrak{m}} \xrightarrow{\sim} \bigcap\nolimits_R^{r+\nu(\mathfrak{m})} H^1_{\mathcal{F}^{\mathfrak{m}}}(K, T)$$

for each ideal $\mathfrak{m} \in \mathcal{N}(\mathcal{Q})$. Since $\mathfrak{n}$ is a core vertex, by Theorem 3.7, we have an exact sequence

$$0 \to H^1_{\mathcal{F}^{\mathfrak{m}}}(K, T) \to H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \to \bigoplus_{\mathfrak{q} \mid \mathfrak{m}^{-1}\mathfrak{n}} H^1_{/f}(K_{\mathfrak{q}}, T) \to H^1_{\mathcal{F}^*_{\mathfrak{m}}}(K, T^\vee(1))^\vee \to 0$$

for an ideal $\mathfrak{m} \mid \mathfrak{n}$. We denote by $e(\mathfrak{m})$ this exact sequence. Note that $H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ is free of rank $r + \nu(\mathfrak{n})$ by Lemma 4.6 and that $H^1_{/f}(K_{\mathfrak{q}}, T)$ is free of rank 1 for any prime $\mathfrak{q} \in \mathcal{Q}$ by Lemma 3.17. Since $\mathfrak{n}$ is a core vertex, by Theorem 4.7, the element $\epsilon_{\mathfrak{n}}$ is a basis of $X_{\mathfrak{n}}$. Thus for any ideal $\mathfrak{m} \mid \mathfrak{n}$, the exact sequence $e(\mathfrak{m})$ and the element $i_{\mathfrak{m}}(\epsilon_{\mathfrak{m}})$ satisfy the assumptions in Lemma 4.8. Hence we have

$$\mathrm{im}(i_{\mathfrak{m}}(\epsilon_{\mathfrak{m}})) = \mathrm{Fitt}^0_R(H^1_{\mathcal{F}^*_{\mathfrak{m}}}(K, T^\vee(1))^\vee).$$

By comparing the exact sequences $e(1)$ and $e(\mathfrak{m})$ for any $\mathfrak{m} \mid \mathfrak{n}$ with $\nu(\mathfrak{m}) = i$, we see that

$$\mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee) = \sum_{\mathfrak{m} \mid \mathfrak{n}, \nu(\mathfrak{m}) = i} \mathrm{Fitt}^0_R(H^1_{\mathcal{F}^*_{\mathfrak{m}}}(K, T^\vee(1))^\vee).$$

Thus we have

$$\mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee) = \sum_{\mathfrak{m} \mid \mathfrak{n}, \nu(\mathfrak{m}) = i} \mathrm{im}(i_{\mathfrak{m}}(\epsilon_{\mathfrak{m}})) \subseteq I_i(\epsilon)$$

for any core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ with $\nu(\mathfrak{n}) \geq i$. Let $\mathfrak{m} \in \mathcal{Q}$ with $\nu(\mathfrak{m}) = i$. Then there is a core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ with $\mathfrak{m} \mid \mathfrak{n}$. Thus we have

$$\mathrm{im}(i_{\mathfrak{m}}(\epsilon_{\mathfrak{m}})) \subseteq \sum_{\mathfrak{m} \mid \mathfrak{n}, \nu(\mathfrak{n}) = i} \mathrm{im}(i_{\mathfrak{m}}(\epsilon_{\mathfrak{m}})) = \mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee).$$

Hence we get the desired equality. □

The following proposition will be used in Section 6.

Proposition 4.11. *Suppose that $\mathcal{F}$ is cartesian and $r = \chi(\mathcal{F}) \geq 0$. Let $\mathcal{Q}$ be a subset of $\mathcal{P}$ such that $\mathcal{N}(\mathcal{Q})$ has a core vertex. Let $\epsilon = \{\epsilon_{\mathfrak{n}}\}_{\mathfrak{n} \in \mathcal{N}(\mathcal{Q})}$ be a basis of $\mathbf{SS}_r(T, \mathcal{F}, \mathcal{Q})$. For $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$, let*

$$\xi^{r+\nu(\mathfrak{n})}_{\mathfrak{n}} : \bigwedge\nolimits_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \to \bigcap\nolimits_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$$

denote the canonical map.

If there is a free R-submodule M of $H^1_{\mathcal{F}}(K, T)$ with a basis $c_1, \ldots, c_r$, then there is a unique element $\theta(\epsilon) \in R$ such that $\theta(\epsilon)R = \mathrm{Fitt}^0_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee)$ and $\epsilon_1 = \xi^r_1(\theta(\epsilon)c_1 \wedge \cdots \wedge c_r)$. In particular, we have $\epsilon_1 = 0$ if there is an injective map $R^{r+1} \to H^1_{\mathcal{F}}(K, T)$.

Proof. Let $c := c_1 \wedge \cdots \wedge c_r$. Note that M is an injective R-module since R is a zero-dimensional Gorenstein local ring. Since the map $M \to H^1_{\mathcal{F}}(K, T)$ is a split injection, the composition of maps

$$\bigwedge\nolimits^r_R M \xrightarrow{\sim} \bigcap\nolimits^r_R M \longrightarrow \bigcap\nolimits^r_R H^1_{\mathcal{F}}(K, T)$$

is injective. Thus uniqueness follows from $\epsilon_1 = \xi^r_1(\theta(\epsilon)c)$.

Since $H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ is a free R-module of rank $r+\nu(\mathfrak{n})$ by Lemma 4.6, we can take elements $d_1, \ldots, d_{\nu(\mathfrak{n})} \in H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ such that $c_1, \ldots, c_r, d_1, \ldots, d_{\nu(\mathfrak{n})}$ is a basis of $H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$. Let $N = \sum_{i=1}^{\nu(\mathfrak{n})} Rd_i$ and $d := d_1 \wedge \cdots \wedge d_{\nu(\mathfrak{n})}$. We also take a basis $w^* \in \det(W_{\mathfrak{n}})$ such that $w^* \otimes \xi^{r+\nu(\mathfrak{n})}_{\mathfrak{n}}(d \wedge c) = \epsilon_{\mathfrak{n}}$. Put $W^*_{\mathfrak{n}} := \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} H^1_{/f}(K_{\mathfrak{q}}, T)$ and $\varphi : N \hookrightarrow H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \to W^*_{\mathfrak{n}}$. We denote by $w \in \det(W^*_{\mathfrak{n}})$ the dual basis of w^*.

Since $\mathfrak{n}$ is a core vertex and $c_1, \ldots, c_r \in H^1_{\mathcal{F}}(K, T)$, by Theorem 3.7, we have a canonical isomorphism $\operatorname{coker}(\varphi) \xrightarrow{\sim} H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee$. We define an element $\theta(\epsilon) \in R$ by the relation $\wedge^{\nu(\mathfrak{n})}\varphi(d) = \theta(\epsilon)w$. Then by the definition of Fitting ideals, we have

$$\theta(\epsilon)R = \mathrm{Fitt}^0_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee).$$

Applying Lemma 2.2 to the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & W_{\mathfrak{n}} \arrow[r] \arrow[d, "\varphi^*"] & H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)^* \arrow[r] \arrow[d, "\simeq"] & H^1_{\mathcal{F}}(K, T)^* \arrow[d] & \\
0 \arrow[r] & N^* \arrow[r] & N^* \oplus M^* \arrow[r] & M^* \arrow[r] & 0
\end{tikzcd}$$

we get the commutative diagram

$$\begin{tikzcd}
\bigcap^{r+\nu(\mathfrak{n})}_R (N \oplus M) \arrow[r, "\simeq"] \arrow[d, "\mathrm{id}"] & \det(N) \otimes_R \bigcap^r_R M \arrow[d] \\
\bigcap^{r+\nu(\mathfrak{n})}_R H_{\mathcal{F}^{\mathfrak{n}}}(K, T) \arrow[r] & \det(W^*_{\mathfrak{n}}) \otimes_R \bigcap^r_R H^1_{\mathcal{F}}(K, T)
\end{tikzcd}$$

Since the right vertical map sends $d \otimes \xi^r_M(c)$ to $w \otimes \xi^r_1(\theta(\epsilon)c)$, we have

$$\epsilon_1 = \Phi_{\mathfrak{n},1}(\epsilon_{\mathfrak{n}}) = \Phi_{\mathfrak{n},1}(w^* \otimes \xi^r_1(d \wedge c)) = \xi^r_1(\theta(\epsilon)c). \qquad \square$$

Let $\mathcal{Q}$ be a subset of $\mathcal{P}$ such that $\mathcal{N}(\mathcal{Q})$ has a core vertex and let S be a zero-dimensional Gorenstein local ring. Suppose that $\mathcal{F}$ is cartesian and that there is a surjective ring homomorphism $\pi : R \to S$. Take a core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ for $\mathcal{F}$. By Lemma 3.11, Remark 4.5(2), and Lemma 4.6, we obtain a map

$$\begin{aligned}
\bigcap\nolimits^{r+\nu(\mathfrak{n})}_R H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \xleftarrow{\sim} {} & \bigwedge\nolimits^{r+\nu(\mathfrak{n})}_R H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \\
& \longrightarrow \bigwedge\nolimits^{r+\nu(\mathfrak{n})}_R H^1_{\mathcal{F}^{\mathfrak{n}}_S}(K, T \otimes_R S) \\
& \xrightarrow{\sim} \bigcap\nolimits^{r+\nu(\mathfrak{n})}_R H^1_{\mathcal{F}^{\mathfrak{n}}_S}(K, T \otimes_R S).
\end{aligned}$$

Hence if $r = \chi(\mathcal{F}) \geq 0$, we get a map

$$\phi_\pi : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \xrightarrow{\sim} X_{\mathfrak{n}}(T, \mathcal{F}) \longrightarrow X_{\mathfrak{n}}(T \otimes_R S, \mathcal{F}_S) \xleftarrow{\sim} \boldsymbol{SS}_r(T \otimes_R S, \mathcal{F}, \mathcal{Q})$$

by Theorem 4.7. It is easy to see that the map ϕ_π is independent of the choice of the core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$.

Proposition 4.12. *Suppose that $\mathcal{F}$ is cartesian and $r = \chi(\mathcal{F}) \geq 0$. Let $\mathcal{Q}$ be a subset of $\mathcal{P}$ such that $\mathcal{N}(\mathcal{Q})$ has a core vertex, let S be a zero-dimensional Gorenstein local ring, and let $\pi : R \to S$ be a surjective ring homomorphism:*

(1) *The map ϕ_π induces an isomorphism*

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \otimes_R S \xrightarrow{\sim} \boldsymbol{SS}_r(T \otimes_R S, \mathcal{F}_S, \mathcal{Q}).$$

(2) *We have $I_i(\epsilon)S = I_i(\phi_\pi(\epsilon))$ for any $\epsilon \in \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ and nonnegative integer i.*

Proof. Let $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ be a core vertex. Note that $\mathcal{F}_S$ is cartesian by Lemma 3.11. Hence by Lemma 3.25 and Lemma 4.6, the map $X_{\mathfrak{n}}(T, \mathcal{F}) \to X_{\mathfrak{n}}(S, \mathcal{F}_S)$ induces an isomorphism $X_{\mathfrak{n}}(T, \mathcal{F}) \otimes_R S \xrightarrow{\sim} X_{\mathfrak{n}}(S, \mathcal{F}_S)$. Hence the map ϕ_π induces an isomorphism $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \otimes_R S \xrightarrow{\sim} \boldsymbol{SS}_r(S, \mathcal{F}_S, \mathcal{Q})$ by Theorem 4.7.

We will prove assertion (2). Let $\mathfrak{m} \in \mathcal{N}(\mathcal{Q})$. Since $\mathcal{N}(\mathcal{Q})$ has a core vertex, there is a core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ with $\mathfrak{m} \mid \mathfrak{n}$. We fix an isomorphism $H^1_{/f}(K_{\mathfrak{q}}, T) \simeq R$ for each prime $\mathfrak{q} \mid \mathfrak{n}$. Let $? \in \{\mathfrak{m}, \mathfrak{n}\}$. Using these isomorphisms, we regard the element $\epsilon_?$ as an element of $\bigcap_R^{r+\nu(?)} H^1_{\mathcal{F}^?}(K, T)$ and $\phi_\pi(\epsilon)_?$ as an element of $\bigcap_S^{r+\nu(?)} H^1_{\mathcal{F}^?_S}(K, T \otimes_R S)$. We put $H_? := H^1_{\mathcal{F}^?}(K, T)^*$ and $H'_? := H^1_{\mathcal{F}^?_S}(K, T \otimes_R S)^*$. Applying Lemma 2.1 to the exact sequence $W_{\mathfrak{m}^{-1}\mathfrak{n}} \to H_{\mathfrak{n}} \to H_{\mathfrak{m}} \to 0$, we get a map

$$\psi : \bigwedge\nolimits_R^{r+\nu(\mathfrak{m})} H_{\mathfrak{m}} \xrightarrow{\sim} \det(W_{\mathfrak{m}^{-1}\mathfrak{n}}) \otimes_R \bigwedge\nolimits_R^{r+\nu(\mathfrak{m})} H_{\mathfrak{m}} \longrightarrow \bigwedge\nolimits_R^{r+\nu(\mathfrak{n})} H_{\mathfrak{n}},$$

where the first isomorphism is induced by the fixed isomorphisms $H^1_{/f}(K_{\mathfrak{q}}, T) \simeq R$. In the same way, we also get a map $\psi' : \bigwedge_S^{r+\nu(\mathfrak{m})} H'_{\mathfrak{m}} \to \bigwedge_S^{r+\nu(\mathfrak{n})} H'_{\mathfrak{n}}$. Then we have the following commutative diagram:

$$\begin{tikzcd}
\bigwedge_R^{r+\nu(\mathfrak{m})} H_{\mathfrak{n}} \arrow[r, "j"] \arrow[d, "s"] & \bigwedge_R^{r+\nu(\mathfrak{m})} H_{\mathfrak{m}} \arrow[r, "\psi"] & \bigwedge_R^{r+\nu(\mathfrak{n})} H_{\mathfrak{n}} \arrow[r, "\epsilon_{\mathfrak{n}}"] \arrow[d, "t"] & R \arrow[d, "\pi"] \\
\bigwedge_S^{r+\nu(\mathfrak{m})} H'_{\mathfrak{n}} \arrow[r, "j'"] & \bigwedge_S^{r+\nu(\mathfrak{m})} H'_{\mathfrak{m}} \arrow[r, "\psi'"] & \bigwedge_S^{r+\nu(\mathfrak{n})} H'_{\mathfrak{n}} \arrow[r, "\phi_\pi(\epsilon)_{\mathfrak{n}}"] & S
\end{tikzcd}$$

where the maps s and t are induced by a surjective homomorphism

$$H_{\mathfrak{n}} \longrightarrow H_{\mathfrak{n}} \otimes_R S \xrightarrow{\sim} (H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T) \otimes_R S)^* \xleftarrow{\sim} H'_{\mathfrak{n}},$$

the map j is induced by the surjective map $H_{\mathfrak{n}} \to H_{\mathfrak{m}}$, and j' is induced by the surjective map $H'_{\mathfrak{n}} \to H'_{\mathfrak{m}}$. By the definition of Stark system, we have $\phi_\pi(\epsilon)_{\mathfrak{m}} = \phi_\pi(\epsilon)_{\mathfrak{n}} \circ \psi'$ and $\epsilon_{\mathfrak{m}} = \epsilon_{\mathfrak{n}} \circ \psi$. Since the maps s, j, and j' are surjective, we have

$$\operatorname{im}(\phi_\pi(\epsilon)_{\mathfrak{m}}) = \operatorname{im}(\phi_\pi(\epsilon)_{\mathfrak{m}} \circ j' \circ s) = \operatorname{im}(\pi \circ \epsilon_{\mathfrak{m}} \circ j) = \operatorname{im}(\epsilon_{\mathfrak{m}})S. \qquad \square$$

Stark systems over principal Artinian local rings. Suppose that R is a principal artinian local ring, $\mathcal{F}$ is cartesian, and $r = \chi(\mathcal{F}) \geq 0$. We will show that there is a canonical isomorphism from the module of Stark systems defined in [Mazur and Rubin 2016] to the module of Stark systems defined in the previous subsection.

Let $\mathcal{Q}$ be an infinite subset of $\mathcal{P}_{\ell_R(R)}(\mathcal{F})$ such that $\mathcal{N}(\mathcal{Q})$ has a core vertex. We put

$$W'_{\mathfrak{n}} := \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T)^* = \bigoplus_{\mathfrak{q} \mid \mathfrak{n}} \mathrm{Hom}_R(H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T), R)$$

for each ideal $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$. By Lemma 3.17, the composition of maps

$$H^1_{\mathrm{tr}}(K_{\mathfrak{q}}, T) \to H^1(K, T) \xrightarrow{\mathrm{loc}^{/f}_{\mathfrak{q}}} H^1_{/f}(K_{\mathfrak{q}}, T)$$

is an isomorphism for any prime $\mathfrak{q} \in \mathcal{Q}$. For any ideal $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$, we denote by $j_{\mathfrak{n}} : W'_{\mathfrak{n}} \xrightarrow{\sim} W_{\mathfrak{n}}$ the inverse of a natural isomorphism $W_{\mathfrak{n}} \xrightarrow{\sim} W'_{\mathfrak{n}}$.

Let $Y_{\mathfrak{n}} := \det(W'_{\mathfrak{n}}) \otimes_R \bigwedge_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^{\mathfrak{n}}}(K, T)$ and let $\Psi_{\mathfrak{n},\mathfrak{m}} : Y_{\mathfrak{n}} \to Y_{\mathfrak{m}}$ be the map defined in [Mazur and Rubin 2016, Definition 6.3]. Then the module $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})'$ of Stark systems defined in that paper is the inverse limit

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})' := \varprojlim_{\mathfrak{n} \in \mathcal{N}(\mathcal{Q})} Y_{\mathfrak{n}}$$

with respect to the maps $\Psi_{\mathfrak{n},\mathfrak{m}}$.

Lemma 4.13. *Let $0 \to N \xrightarrow{g} M \xrightarrow{h} F$ be an exact sequence of finitely generated R-modules, F a free R-module of rank 1, and s a positive integer. Then the following diagram commutes*:

$$\begin{array}{ccc} \bigwedge_R^s M & \xrightarrow{\hat{h}} & F \otimes_R \bigwedge_R^{s-1} N \\ \big\downarrow{\scriptstyle \xi^s_M} & & \big\downarrow{\scriptstyle \mathrm{id}_F \otimes \xi^{s-1}_N} \\ \bigcap_R^s M & \xrightarrow{\tilde{h}} & F \otimes_R \bigcap_R^{s-1} N \end{array}$$

where $\hat{h}$ is the map defined in [Mazur and Rubin 2016, Proposition A.1] and $\tilde{h}$ is the dual of the map defined by the exact sequence $F^* \xrightarrow{h^*} M^* \xrightarrow{g^*} N^* \to 0$ using Lemma 2.1.

Proof. We may assume that $F = R$. Since R is a principal artinian local ring, we can write $M = Rm \oplus N_0$ and $N = Im \oplus N_0$ for some element m of M and ideal I of R. Then $\bigwedge_R^s M$ is generated by the set $E = \{m_1 \wedge \cdots \wedge m_s \mid m_2, \ldots, m_s \in N_0, m_1 \in \{m\} \cup N_0\}$. Let $m = m_1 \wedge \cdots \wedge m_s \in E$, $f = f_2 \wedge \cdots \wedge f_s \in \bigwedge_R^{s-1} M^*$, and $x = \wedge^{s-1} g^*(f)$. Note that $m_2 \wedge \cdots \wedge m_s \in \bigwedge_R^{s-1} N$. Put $f_1 = h$. Then we have

$$\begin{aligned} \tilde{h}(\xi^s_M(m))(x) &= \det(f_i(m_j)) \\ &= h(m_1) \det((f_i(m_j))_{2 \leq i \leq s, 2 \leq j \leq s}) \\ &= h(m_1) \xi^{s-1}_N(m_2 \wedge \cdots \wedge m_s)(x) \\ &= \xi^{s-1}_N(\hat{h}(m))(x), \end{aligned}$$

where the second equality follows from $f_s(m_i) = h(m_i) = 0$ for $2 \leq i \leq s$ and the last equality follows from $\hat{h}(m) = h(m_1)m_2 \wedge \cdots \wedge m_s$. $\square$

Let $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ and $\xi_\mathfrak{n} : \bigwedge_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^\mathfrak{n}}(K, T) \to \bigcap_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^\mathfrak{n}}(K, T)$ be the canonical map. Define a map $C_\mathfrak{n} : Y_\mathfrak{n} \to X_\mathfrak{n}$ by $C_\mathfrak{n} = j_\mathfrak{n} \otimes \xi_\mathfrak{n}$.

Proposition 4.14. *The maps $C_\mathfrak{n} : Y_\mathfrak{n} \to X_\mathfrak{n}$ induce an isomorphism*

$$C : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})' \xrightarrow{\sim} \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}).$$

Proof. By Lemma 4.13 and the definition of the transition maps $\Psi_{\mathfrak{n},\mathfrak{m}}$ and $\Phi_{\mathfrak{n},\mathfrak{m}}$, the maps $C_\mathfrak{n}$ induce a map $C : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})' \to \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$. By Lemma 4.6 or [Mazur and Rubin 2016, Proposition 3.3], the map $C_\mathfrak{n} : Y_\mathfrak{n} \to X_\mathfrak{n}$ is an isomorphism for any core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$. Thus by Theorem 4.7 and [loc. cit., Theorem 6.7], the map C is an isomorphism. $\square$

Lemma 4.15. *Let A be a zero-dimensional Gorenstein local ring, M a finitely generated A-module, and s a positive integer. Let $\xi_M^s : \bigwedge_A^s M \to \bigcap_A^s M = \mathrm{Hom}_R\big(\bigwedge_A^s M^*, A\big)$ denote the canonical map and $x \in \bigwedge_A^s M$:*

(1) *Let J be an ideal of A. If $x \in J \bigwedge_A^s M$, then $\mathrm{im}(\xi_M^r(x)) \subseteq J$.*

(2) *Suppose that there is a free A-submodule $M' \subseteq M$ of rank s such that $x \in \mathrm{im}\big(\bigwedge_A^s M' \to \bigwedge_A^s M\big)$. Then $x \in \mathrm{im}(\xi_M^s(x)) \bigwedge_A^s M$.*

Proof. Since $\mathrm{im}(\xi_M^s(ay)) = \mathrm{im}(a \cdot \xi_M^s(y)) = a \cdot \mathrm{im}(\xi_M^s(y))$ for any $y \in \bigwedge_A^s M$ and $a \in A$, assertion (1) holds.

We will show assertion (2). Since A is a zero-dimensional Gorenstein local ring, we can write $M = M' \oplus N$ for some A-submodule $N \subseteq M$. Let $y_1, \ldots, y_s$ be a basis of M'. We take the element $y_i^* \in M^*$ such that $y_i^*(y_i) = 1$ and $y_i^*(y_j) = 0$ if $i \neq j$. We write $x = ay_1 \wedge \cdots \wedge y_s$ for some $a \in A$. Then we have $\xi_M^s(x)(y_1^* \wedge \cdots \wedge y_s^*) = a$. $\square$

Lemma 4.16. *Let $0 \to N \to M \xrightarrow{h} F$ be an exact sequence of finitely generated R-modules, F a free R-module of rank 1, and s a positive integer. Let $\hat{h} : \bigwedge_R^s M \to F \otimes_R \bigwedge_R^{s-1} N$ denote the map defined in [Mazur and Rubin 2016, Proposition A.1] and $x \in \bigwedge_R^s M$.*

Suppose that there is a free R-submodule $M' \subseteq M$ of rank s such that $x \in \mathrm{im}\big(\bigwedge_R^s M' \to \bigwedge_R^s M\big)$. Then there is a free R-submodule $N' \subseteq M' \cap N$ of rank $s-1$ such that

$$\hat{h}(y) \in F \otimes_R \mathrm{im}\left(\bigwedge\nolimits_R^{s-1} N' \to \bigwedge\nolimits_R^{s-1} N\right) \subseteq F \otimes_R \bigwedge\nolimits_R^{s-1} N.$$

Proof. Since R is a principal artinian local ring, there is a basis $y_1, \ldots, y_s$ of M' such that $y_2, \ldots, y_s \in N$. Put $N' := \sum_{i=2}^s Ry_i$. Then N' is a free R-submodule of $M' \cap N$ of rank $s-1$ and $\hat{h}(y_1 \wedge \cdots \wedge y_s) = h(y_1) \otimes y_2 \wedge \cdots \wedge y_s \in F \otimes \mathrm{im}\big(\bigwedge_R^{s-1} N' \to \bigwedge_R^{s-1} N\big)$. $\square$

The following proposition says that Theorem 4.10 is a generalization of [Mazur and Rubin 2016, Theorem 8.5].

Proposition 4.17. *Let ϵ' be a basis of $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})'$ and $k = \ell_R(R)$. Then we have*

$$\partial\varphi_{\epsilon'}(t) = \max\{n \in \{0, 1, \dots, k-1, \infty\} \mid I_t(C(\epsilon')) \subseteq \mathfrak{m}_R^n\}.$$

Here $\partial\varphi_{\epsilon'} : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ is the map defined in [Mazur and Rubin 2016, Definition 8.1].

Proof. We fix an isomorphism $H^1_{/f}(K_\mathfrak{q}, T) \simeq R$ for each prime $\mathfrak{q} \in \mathcal{Q}$. Then we get an isomorphism $i_\mathfrak{n} : X_\mathfrak{n} \xrightarrow{\sim} \bigcap_R^{r+\nu(\mathfrak{n})} H^1_{\mathcal{F}^\mathfrak{n}}(K, T)$ for each ideal $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$. Let c be a nonnegative integer and $\mathfrak{m} \in \mathcal{N}(\mathcal{Q})$. Then we only need to show that the following assertions are equivalent:

(i) $\epsilon'_\mathfrak{m} \in \mathfrak{m}_R^c Y_\mathfrak{m}$.

(ii) $\mathrm{im}(i_\mathfrak{m}(C(\epsilon')_\mathfrak{m})) = \mathrm{im}(i_\mathfrak{m}(C_\mathfrak{m}(\epsilon'_\mathfrak{m}))) \subseteq \mathfrak{m}_R^c$.

By Lemma 4.15(1), (i) implies (ii). We will show that (ii) implies (i). Since we assume that $\mathcal{N}(\mathcal{Q})$ has a core vertex, there is a core vertex $\mathfrak{n} \in \mathcal{N}(\mathcal{Q})$ with $\mathfrak{m} \mid \mathfrak{n}$. Since $H^1_{\mathcal{F}^\mathfrak{n}}(K, T)$ is a free R-module of rank $r + \nu(\mathfrak{n})$, by Lemma 4.16, there is a free R-submodule $M \subseteq H^1_{\mathcal{F}^\mathfrak{m}}(K, T)$ of rank $r + \nu(\mathfrak{m})$ such that $\epsilon'_\mathfrak{m} \in \det(W'_\mathfrak{m}) \otimes_R \mathrm{im}\big(\bigwedge_R^{r+\nu(\mathfrak{m})} M \to \bigwedge_R^{r+\nu(\mathfrak{m})} H^1_{\mathcal{F}^\mathfrak{m}}(K, T)\big)$. Thus by Lemma 4.15(2), we see that (ii) implies (i). □

5. Stark systems over Gorenstein local rings

In this section, we assume that R is a complete Gorenstein local ring and that T satisfies Hypothesis 3.12.

Since R is a complete Gorenstein local ring, there is an increasing sequence $J_1 \subseteq J_2 \subseteq J_3 \subseteq \cdots$ of ideals of R such that R/J_n is a zero-dimensional Gorenstein local ring and a natural map $R \to \varprojlim_n R/J_n$ is an isomorphism. In fact, since R is Cohen–Macaulay, there is a regular sequence $x_1, \dots, x_d \in \mathfrak{m}_R$ such that $\sqrt{(x_1, \dots, x_d)} = \mathfrak{m}_R$. Put $J_n = (x_1^n, \dots, x_d^n)$ for each positive integer n. Since $x_1^n, \dots, x_d^n$ is a regular sequence and $\sqrt{J_n} = \mathfrak{m}_R$, R/J_n is a zero-dimensional Gorenstein local ring. Furthermore the map $R \to \varprojlim_n R/J_n$ is an isomorphism since R is a complete noetherian local ring. Throughout this section, we fix such a sequence $\{J_n\}_{n\geq 1}$ of ideals of R.

Let n be a positive integer and $\mathcal{F}_n$ a Selmer structure on T/J_nT such that

$$H^1_{\mathcal{F}_n}(K_\mathfrak{q}, T/J_nT) = H^1_{(\mathcal{F}_{n+1})_{R/J_n}}(K_\mathfrak{q}, T/J_nT)$$

for any prime $\mathfrak{q}$ of K. Then we have $r := \chi(\mathcal{F}_n) = \chi(\mathcal{F}_{n+1})$ for any positive integer n. In order to use the results of the previous section, we assume that $r \geq 0$ and that the Selmer structure $\mathcal{F}_n$ is cartesian for all positive integers n. Put $\mathcal{F} := \{\mathcal{F}_n\}_{n\geq 1}$.

Remark 5.1. Let S be a zero-dimensional Gorenstein local ring and $\pi : R \to S$ a surjective ring homomorphism. Since the kernel of π is an open ideal, there is a positive integer n such that $I_n \subseteq \ker(\pi)$. Hence the collection of the Selmer structures $\mathcal{F}$ defines a Selmer structure $\mathcal{F}_S$ on $T \otimes_R S$. Note that $\chi(\mathcal{F}_S) = r$ and $\mathcal{F}_S$ is cartesian by Remark 3.10.

Remark 5.2. Let S be a complete Gorenstein local ring and $\pi : R \to S$ a surjective ring homomorphism. Take an increasing sequence $\{J_{S,n}\}_{n\geq 1}$ of ideals of S such that $S/J_{S,n}$ is a zero-dimensional Gorenstein local ring and the map $S \to \varprojlim_n S/J_{S,n}$ is an isomorphism. By Remark 5.1, we get a Selmer structure $\mathcal{F}_{S,n}$ on $T \otimes_R S/J_{S,n}$ for each positive integer n. Then we have $\chi(\mathcal{F}_{S,n}) = r$, the Selmer structure $\mathcal{F}_{S,n}$ is cartesian, and

$$H^1_{\mathcal{F}_{S,n}}(K_\mathfrak{q}, T \otimes_R S/J_{S,n}S) = H^1_{(\mathcal{F}_{S,n+1})_{R/J_n}}(K_\mathfrak{q}, T \otimes_R S/J_{S,n}S)$$

for any prime $\mathfrak{q}$ of K and positive integer n. We denote by $\mathcal{F}_S$ the collection $\{\mathcal{F}_{S,n}\}_{n\geq 1}$ of the Selmer structures.

Example 5.3. The following example is closely related to the results in [Büyükboduk 2014, Appendix A; Büyükboduk and Lei 2015, Appendix A]. Let $\mathcal{O}$ be the ring of integers of a finite extension of the field $\mathbb{Q}_p$ and $\mathcal{T}$ a free $\mathcal{O}$-module of finite rank with an $\mathcal{O}$-linear continuous G_K-action which is unramified outside a finite set of places of K. Let d be a positive integer and L a free $R := \mathcal{O}[[X_1, \dots, X_d]]$-module of rank 1 with an R-linear continuous G_K-action which is unramified outside primes above p. Put $T := \mathcal{T} \otimes_\mathcal{O} L$ and $J_n := (p^n, X_1^n, \dots, X_d^n)$. Suppose that the following conditions hold:

- The module $\mathcal{T}$ satisfies Hypothesis 3.12.
- The module $L \otimes_R R/(X_1, \dots, X_d)$ has trivial G_K-action.
- The module $(\mathcal{T} \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)^{\mathcal{I}_\mathfrak{q}}$ is divisible for every prime $\mathfrak{q} \nmid p$ of K.
- $H^2(K_\mathfrak{p}, \mathcal{T}) = 0$ for each prime $\mathfrak{p} \mid p$ of K.

For a positive integer n, we define a Selmer structure $\mathcal{F}_n$ on T/J_nT by the following data:

- $\Sigma(\mathcal{F}_n) := \{v \mid p\infty\} \cup \{\mathfrak{q} \mid \mathcal{T} \text{ is ramified at } \mathfrak{q}\}$.
- $H^1_{\mathcal{F}_n}(K_\mathfrak{q}, T/J_nT) := H^1_{\mathrm{ur}}(K_\mathfrak{q}, T/J_nT)$ for each prime $\mathfrak{q} \nmid p$ of K.
- $H^1_{\mathcal{F}_n}(K_\mathfrak{p}, T/J_nT) := H^1(K_\mathfrak{p}, T/J_nT)$ for each prime $\mathfrak{p} \mid p$ of K.

Since $(\mathcal{T} \otimes_{\mathbb{Z}_p} \mathbb{Q}_p/\mathbb{Z}_p)^{\mathcal{I}_\mathfrak{q}}$ is divisible and $L^{\mathcal{I}_\mathfrak{q}} = L$ for every prime $\mathfrak{q} \nmid p$ of K, a canonical map $(T/J_{n+1}T)^{\mathcal{I}_\mathfrak{q}} \to (T/J_nT)^{\mathcal{I}_\mathfrak{q}}$ is surjective. Note that the cohomological dimension of $\mathcal{D}_\mathfrak{q}/\mathcal{I}_\mathfrak{q} \simeq \widehat{\mathbb{Z}}$ is 1. Therefore a canonical map $H^1_{\mathrm{ur}}(K_\mathfrak{q}, T/J_{n+1}T) \to H^1_{\mathrm{ur}}(K_\mathfrak{q}, T/J_nT)$ is surjective. By the inflation-restriction exact sequence, we have an isomorphism

$$H^1_{/\mathcal{F}_n}(K_\mathfrak{q}, T/J_nT) \xrightarrow{\sim} H^1(\mathcal{I}_\mathfrak{q}, T/J_nT)^{\mathrm{Fr}_\mathfrak{q}=1}.$$

Hence we see that the map

$$H^1_{/\mathcal{F}_n}(K_\mathfrak{q}, T/\mathfrak{m}_RT) \to H^1_{/\mathcal{F}_n}(K_\mathfrak{q}, T/J_nT)$$

induced by an injection $\Bbbk \to R$ is injective. Let $\mathfrak{p} \mid p$ be a prime of K. Since the cohomological dimension of $\mathcal{D}_\mathfrak{p}$ is 2, we have

$$H^2(K_\mathfrak{p}, T) \otimes_R R/(X_1, \dots, X_d) \xrightarrow{\sim} H^2(K_\mathfrak{p}, \mathcal{T}).$$

Because $H^2(K_{\mathfrak{p}}, \mathcal{T}) = 0$, we have $H^2(K_{\mathfrak{p}}, T) = 0$ by Nakayama's lemma, and a canonical map $H^1(K_{\mathfrak{p}}, T/J_{n+1}T) \to H^1(K_{\mathfrak{p}}, T/J_nT)$ is surjective for any positive integer n. Furthermore, by [Mazur and Rubin 2016, Theorem 5.4], we have

$$\chi(\mathcal{F}_n) = \sum_{v \mid \infty} \mathrm{rank}_{\mathcal{O}}(H^0(K_v, \mathcal{T}^\vee(1))^\vee) \geq 0,$$

where v runs through all the infinite places of K. Hence the collection of the Selmer structures $\{\mathcal{F}_n\}_{n\geq 1}$ satisfies all the conditions in the paragraph before Remark 5.1.

Stark systems over Gorenstein local rings. We fix a decreasing sequence $\mathcal{Q}_1 \supseteq \mathcal{Q}_2 \supseteq \mathcal{Q}_3 \supseteq \cdots$ such that $\mathcal{Q}_n$ is an infinite subset of $\mathcal{P}_{\ell_R(R/J_n)}(\mathcal{F}_n)$ and $\mathcal{N}(\mathcal{Q})$ has a core vertex for $\mathcal{F}_n$. For example, we take $\mathcal{Q}_n = \mathcal{P}_{\ell_R(R/J_n)}(\mathcal{F}_n) \setminus (\Sigma(\mathcal{F}_1) \cup \cdots \cup \Sigma(\mathcal{F}_n))$ for each positive integer n. We denote by $\mathcal{Q}$ the collection of the sequences $\{\mathcal{Q}_n\}_{n\geq 1}$.

Let n be a positive integer. Since $\mathcal{N}(\mathcal{Q}_{n+1})$ has a core vertex for $\mathcal{F}_{n+1}$, the map $R/J_{n+1} \to R/J_n$ induces a map $\boldsymbol{SS}_r(T/J_{n+1}T, \mathcal{F}_{n+1}, \mathcal{Q}_{n+1}) \to \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_{n+1})$ as in the paragraph before Proposition 4.12. Since, by Theorem 4.7, a restriction map $\boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n) \to \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_{n+1})$ is an isomorphism, we get a map

$$\phi_{n+1,n} : \boldsymbol{SS}_r(T/J_{n+1}T, \mathcal{F}_{n+1}, \mathcal{Q}_{n+1}) \to \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n).$$

We define the module $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ of Stark systems of rank r for $\mathcal{F}$ by

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) := \varprojlim_{n\geq 1} \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n),$$

where the inverse limit is taken with respect to the maps $\phi_{n+1,n}$.

Let S be a complete Gorenstein local ring and $\pi : R \to S$ a surjective ring homomorphism. Take an increasing sequence $J_{S,1} \subseteq J_{S,2} \subseteq J_{S,3} \subseteq \cdots$ of ideals of S such that $S/J_{S,n}$ is a zero-dimensional Gorenstein local ring and the map $S \to \varprojlim_n S/J_{S,n}$ is an isomorphism. Let $\mathcal{F}_S$ be the collection of the Selmer structures defined in Remark 5.2. For a positive integer n, we fix a positive integer n_S such that $n_S \geq n$ and such that the map $\pi : R \to S$ induces a map $\pi_n : R/J_{n_S} \to S/J_{S,n}$. Then we have the map

$$\phi_{\pi_n} : \boldsymbol{SS}_r(T/J_{n_S}T, \mathcal{F}_{n_S}, \mathcal{Q}_{n_S}) \to \boldsymbol{SS}_r(T \otimes_R S/J_{S,n}, \mathcal{F}_{S,n}, \mathcal{Q}_{n_S}).$$

Hence we get a canonical map

$$\phi_\pi : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \to \boldsymbol{SS}_r(T \otimes_R S, \mathcal{F}_S, \{\mathcal{Q}_{n_S}\}_{n\geq 1}).$$

For simplicity, we also denote by $\mathcal{Q}$ the subsequence $\{\mathcal{Q}_{n_S}\}_{n\geq 1}$ of $\mathcal{Q} = \{\mathcal{Q}_n\}_{n\geq 1}$.

Theorem 5.4. (1) *The R-module $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ is free of rank 1.*

(2) *Let S be a complete Gorenstein local ring and $\pi : R \to S$ a surjective ring homomorphism. Then the map ϕ_π induces an isomorphism*

$$\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \otimes_R S \xrightarrow{\sim} \boldsymbol{SS}_r(T \otimes_R S, \mathcal{F}_S, \mathcal{Q}).$$

Proof. For any positive integers $n_1 < n_2$, by Proposition 4.12(1), a canonical map

$$\boldsymbol{SS}_r(T/J_{n_2}T, \mathcal{F}_{n_2}, \mathcal{Q}_{n_2}) \otimes_{R/J_{n_2}} R/J_{n_1} \to \boldsymbol{SS}_r(T/J_{n_1}T, \mathcal{F}_{n_1}, \mathcal{Q}_{n_1})$$

is an isomorphism. Furthermore, for any positive integer n, the module $\boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n)$ is a free R/J_n-module of rank 1 by Theorem 4.7. Hence by definition, $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$ is free of rank 1, so (1) holds. Note that assertion (1) and Proposition 4.12(1) imply $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \otimes_R R/J_n \xrightarrow{\sim} \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n)$. Again by Proposition 4.12(1), the map ϕ_{π_n} induces an isomorphism

$$\boldsymbol{SS}_r(T/J_{n_S}T, \mathcal{F}_{n_S}, \mathcal{Q}_{n_S}) \otimes_{R/J_{n_S}} S/J_{S,n} \xrightarrow{\sim} \boldsymbol{SS}_r(T \otimes_R S/J_{S,n}, \mathcal{F}_{S,n}, \mathcal{Q}_{n_S}).$$

Assertion (2) follows from these isomorphisms. □

Let i be a nonnegative integer and $\epsilon \in \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$. For a positive integer n, let

$$\phi_n : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}) \to \boldsymbol{SS}_r(T/J_nT, \mathcal{F}_n, \mathcal{Q}_n)$$

denote the projection map. By Proposition 4.12(2), we can define an ideal $I_i(\epsilon)$ of R by

$$I_i(\epsilon) := \varprojlim_{n\geq 1} I_i(\phi_n(\epsilon)) = \varprojlim_{n\geq 1}\Bigg(\sum_{\mathfrak{n}\in\mathcal{N}(\mathcal{Q}_n),\, \nu(\mathfrak{n})=i} \mathrm{im}(\phi_n(\epsilon)_{\mathfrak{n}})\Bigg),$$

where the inverse limit is taken with respect to the natural maps $R/J_{n+1} \to R/J_n$.

Definition 5.5. We define the dual Selmer group $H^1_{\mathcal{F}^*}(K, T^\vee(1))$ associated with the collection of the Selmer structures $\mathcal{F}$ by

$$H^1_{\mathcal{F}^*}(K, T^\vee(1)) := \varinjlim_{n\geq 1} H^1_{\mathcal{F}_n^*}(K, (T/J_nT)^\vee(1)),$$

where the injective limit is taken with respect to the maps induced by the Cartier duals of the maps $T/J_{n+1}T \to T/J_nT$.

The following theorem is the main result of this paper.

Theorem 5.6. *Let i be a nonnegative integer*:

(1) *If ϵ is a basis of $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$, then we have*

$$I_i(\epsilon) = \mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee).$$

(2) *Let S be a complete Gorenstein local ring and $\pi : R \to S$ a surjective ring homomorphism. Then we have $I_i(\epsilon)S = I_i(\phi_\pi(\epsilon))$ for any $\epsilon \in \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})$.*

Proof. By Lemma 3.14, the canonical map

$$H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee \otimes_R R/J_n \to H^1_{\mathcal{F}_n^*}(K, (T/J_nT)^\vee(1))^\vee$$

is an isomorphism for any positive integer n. Thus we have

$$\mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee)R/J_n = \mathrm{Fitt}^i_{R/J_n}(H^1_{\mathcal{F}_n^*}(K, (T/J_nT)^\vee(1))^\vee) = I_i(\phi_n(\epsilon)) = I_i(\epsilon)R/J_n,$$

where the second equality follows from Theorem 4.10, and the third follows from Proposition 4.12(2). Since R is a complete noetherian local ring, all the ideals of R are closed. Hence we have $I_i(\epsilon) = \mathrm{Fitt}^i_R(H^1_{\mathcal{F}^*}(K, T^\vee(1))^\vee)$. The second assertion follows from Proposition 4.12(2). $\square$

Remark 5.7. Suppose that R is a discrete valuation ring. Let $\boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})'$ be the module of Stark systems defined in [Mazur and Rubin 2016, Definition 7.1] and n a positive integer. Let

$$C_n : \boldsymbol{SS}_r(T/\mathfrak{m}_R^n T, \mathcal{F}_n, \mathcal{Q}_n)' \to \boldsymbol{SS}_r(T/\mathfrak{m}_R^n T, \mathcal{F}_n, \mathcal{Q}_n)$$

be the isomorphism defined in Proposition 4.14. By the definition of the map C_n, we have the following commutative diagram:

$$\begin{tikzcd} \boldsymbol{SS}_r(T/\mathfrak{m}_R^{n+1}T, \mathcal{F}_{n+1}, \mathcal{Q}_{n+1})' \arrow[r, "C_{n+1}"] \arrow[d] & \boldsymbol{SS}_r(T/\mathfrak{m}_R^{n+1}T, \mathcal{F}_n, \mathcal{Q}_{n+1}) \arrow[d] \\ \boldsymbol{SS}_r(T/\mathfrak{m}_R^n T, \mathcal{F}_n, \mathcal{Q}_n)' \arrow[r, "C_n"] & \boldsymbol{SS}_r(T/\mathfrak{m}_R^n T, \mathcal{F}_n, \mathcal{Q}_n) \end{tikzcd}$$

Thus by [loc. cit., Proposition 7.3], the maps C_n induce an isomorphism

$$C : \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})' \xrightarrow{\sim} \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q}).$$

Furthermore, if $\epsilon' \in \boldsymbol{SS}_r(T, \mathcal{F}, \mathcal{Q})'$ is nonzero, we see that

$$\partial\varphi_{\epsilon'}(t) = \max\{n \in \mathbb{Z}_{\geq 0} \cup \{\infty\} \mid I_t(C(\epsilon')) \subseteq \mathfrak{m}_R^n\}$$

for any nonnegative integer t by Proposition 4.17. This shows that Theorem 5.6 implies [loc. cit., Theorem 8.7].

6. Controlling Selmer groups using Λ-adic Stark systems

Let $\mathcal{O}$ be the ring of integers of a finite extension of the field $\mathbb{Q}_p$ and $\mathcal{T}$ a free $\mathcal{O}$-module of finite rank with an $\mathcal{O}$-linear continuous G_K-action which is unramified outside a finite set of places of K. We write K_∞ for the cyclotomic $\mathbb{Z}_p$-extension of K. Fix a topological generator γ of $\mathrm{Gal}(K_\infty/K)$. Let R be the ring of formal power series $\mathcal{O}[[X]]$. By using the element γ, we get an isomorphism $\mathcal{O}[[\mathrm{Gal}(K_\infty/K)]] \xrightarrow{\sim} R$, $\gamma \mapsto 1 + X$. It induces an R-module structure on $\mathcal{O}[[\mathrm{Gal}(K_\infty/K)]]$.

Following the notation in [Mazur and Rubin 2004, Section 5.3], we write Λ instead of R and put $\mathbb{T} := \mathcal{T} \otimes_\Lambda \mathcal{O}[[\mathrm{Gal}(K_\infty/K)]]$. We assume the following conditions:

- $(\mathcal{T}/\varpi\mathcal{T})^{G_K} = (\mathcal{T}^\vee(1)[\varpi])^{G_K} = 0$ and $\mathcal{T}/\varpi\mathcal{T}$ is an absolutely irreducible $\Bbbk[G_K]$-module, where ϖ is a uniformizer of $\mathcal{O}$.
- The module $\mathcal{T}$ satisfies conditions (H.2) and (H.3).
- The module $(\mathcal{T} \otimes \mathbb{Q}_p/\mathbb{Z}_p)^{\mathcal{I}_\mathfrak{q}}$ is divisible for every prime $\mathfrak{q} \nmid p$ of K.
- $H^2(K_\mathfrak{p}, \mathcal{T}) = 0$ for each prime $\mathfrak{p} \mid p$ of K.

Let n be a positive integer and $J_n := (p^n, X^n)$. Let $\mathcal{F}_n$ be the Selmer structure on $\mathbb{T}/J_n\mathbb{T}$ defined in Example 5.3. Note that $\chi(\mathcal{F}_n) = \sum_{v \mid \infty} \mathrm{rank}_{\mathcal{O}}(H^0(K_v, \mathcal{T}^\vee)^\vee)$, where v runs through all the infinite places of K. Set $\mathcal{F} = \{\mathcal{F}_n\}_{n\geq 1}$ and $r = \chi(\mathcal{F}_n)$. We can define the module of Stark systems by

$$\boldsymbol{SS}_r(\mathbb{T}) := \boldsymbol{SS}_r(\mathbb{T}, \mathcal{F}, \{\mathcal{P}_{n^2}(\mathcal{F}_n)\}_{n\geq 1}) = \varprojlim_{n\geq 1} \boldsymbol{SS}_r(\mathbb{T}/J_n\mathbb{T}, \mathcal{F}_n, \mathcal{P}_{n^2}(\mathcal{F}_n)).$$

Note that $\boldsymbol{SS}_r(\mathbb{T})$ is a free Λ-module of rank 1 by Theorem 5.4. We call an element of $\boldsymbol{SS}_r(\mathbb{T})$ a Λ-adic Stark system. We say that a Λ-adic Stark system is primitive if it is a basis of $\boldsymbol{SS}_r(\mathbb{T})$.

Remark 6.1. Assume that $r = 1$. Let $\boldsymbol{KS}(\mathbb{T})$ be the module of Λ-adic Kolyvagin systems defined in [Mazur and Rubin 2004, Chapter 5; Büyükboduk 2011]. Note that the module of Λ-adic Stark systems is not defined in [Mazur and Rubin 2016] even if $r = 1$. We can construct a canonical isomorphism $\boldsymbol{SS}_1(\mathbb{T}) \xrightarrow{\sim} \boldsymbol{KS}(\mathbb{T})$; see [Büyükboduk 2011, Section 3.1.2]. Thus we have a natural map from the module of Euler systems to the module of Λ-adic Stark systems $\boldsymbol{SS}_1(\mathbb{T})$ by [Mazur and Rubin 2004, Theorem 5.3.3].

Example 6.2. Suppose that K is a totally real number field. Let $\chi : G_K \to \mathcal{O}^\times$ be an even character of finite prime-to-p order and let $\mathfrak{f}_\chi$ be the conductor of χ. Put $\mathcal{T} := \mathcal{O}(1) \otimes_{\mathcal{O}} \chi^{-1}$. Assume the following hypotheses:

- $(p, \mathfrak{f}_\chi) = 1$.
- K is unramified at all primes above p.
- $\chi(\mathrm{Frob}_{\mathfrak{q}}) \neq 1$ for any prime $\mathfrak{p}$ of K above p.

Then $\mathcal{T}$ satisfies all the conditions in this section and we have $r = [K : \mathbb{Q}]$ by [Mazur and Rubin 2016, Corollary 5.6]. Thus if $K = \mathbb{Q}$, then we get a Λ-adic Stark system from the Euler systems of cyclotomic units. By using the analytic class number formula, we see that this Stark system is primitive. More generally, Büyükboduk constructed an $\mathbb{L}$-restricted Kolyvagin system from conjectural Rubin–Stark units in [Büyükboduk 2009]. By [Büyükboduk 2011, Section 3.1.2], this Kolyvagin system gives rise to an element of $SS_1(\mathbb{T}, \mathcal{F}_{\mathbb{L}_\infty}, \{\mathcal{P}_{n^2}(\mathcal{F}_n)\}_{n\geq 1})$. Here $\mathcal{F}_{\mathbb{L}_\infty}$ is the $\mathbb{L}_\infty$-modified Selmer structure defined in [loc. cit., Definition 4.8]. Moreover, we see that there is a canonical isomorphism $SS_r(\mathbb{T}) \xrightarrow{\sim} SS_1(\mathbb{T}, \mathcal{F}_{\mathbb{L}_\infty}, \{\mathcal{P}_{n^2}(\mathcal{F}_n)\}_{n\geq 1})$. Hence we get a Λ-adic Stark system $\epsilon^{\text{R-S}} = \{\epsilon_{\mathfrak{n}}^{\text{R-S}}\}_{\mathfrak{n}}$ such that the leading term $\epsilon_1^{\text{R-S}}$ is an inverse limit of Rubin–Stark elements.

Furthermore, when K is a CM-field, Büyükboduk [2014; 2018] constructed an $\mathbb{L}$-restricted Kolyvagin system from conjectural Rubin–Stark units. In this case, we also get a primitive Λ-adic Stark system such that the leading term is an inverse limit of Rubin–Stark elements.

Example 6.3. Suppose that $\mathcal{O} = \mathbb{Z}_p$. Let A be an abelian variety of dimension d defined over K and let $\mathcal{T}$ be the Tate module $T_p(A) := \varprojlim A[p^n]$. Assume the following hypotheses:

- A has good reduction at all the primes of K above p.
- The image of G_K in $\mathrm{Aut}(A[p]) \simeq \mathrm{GL}_{2d}(\mathbb{F}_p)$ contains $\mathrm{Sp}_{2d}(\mathbb{F}_p)$.

- $p \nmid 6\prod_{\mathfrak{q} \nmid p} c_{\mathfrak{q}}$, where $c_{\mathfrak{q}} \in \mathbb{Z}_{>0}$ is the Tamagawa factor at $\mathfrak{q} \nmid p$.
- $\widehat{A}(K_{\mathfrak{q}})[p] = 0$ for any prime $\mathfrak{p} \mid p$, where $\widehat{A}$ is the dual abelian variety of A.

Then $\mathcal{T} = T_p(A)$ satisfies all the conditions in this section and we have $r = d[K : \mathbb{Q}]$ by [Mazur and Rubin 2016, Proposition 5.7]. If $K = \mathbb{Q}$ and $d = 1$, then Kato's Euler system gives rise to a Λ-adic Stark system; see [Mazur and Rubin 2004, Theorem 6.2.4 and Proposition 6.2.6; Büyükboduk 2011, Proposition 4.2]. Under the assumptions of and by [Büyükboduk 2011, Proposition 4.2], and by Theorem 5.6, we see that this Λ-adic Stark system determines all the higher Fitting ideals of the Pontryagin dual of the Λ-adic Selmer group.

Under slightly different hypotheses, there is a nontrivial conjectural example in the case of $r > 1$. Assume that $\mathcal{T}$ is the twist of $T_p(A)$ by certain character of G_K and A has supersingular reduction at p and a complex multiplication. Under the Perrin-Riou–Stark conjecture, see [Büyükboduk and Lei 2015, Conjecture 4.14], which is closely related to Rubin–Stark conjecture, Büyükboduk and Lei [2015] constructed an $\mathbb{L}$-restricted Kolyvagin system. For the same reason as Example 6.2, we see that this Kolyvagin system gives rise to a Λ-adic Stark system.

By using [Rubin 2000, Proposition B.3.4], we have $H^1(K_{\mathfrak{q}}, \mathbb{T}) = H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, \mathbb{T})$ for each prime $\mathfrak{q} \nmid p$ of K. Hence we can define a Selmer structure $\mathcal{F}_\Lambda$ on $\mathbb{T}$ by the following data:

- $\Sigma(\mathcal{F}_\Lambda) = \Sigma := \{v \mid p\infty\} \cup \{\mathfrak{q} \mid \mathcal{T} \text{ is ramified at } \mathfrak{q}\}$.
- $H^1_{\mathcal{F}_\Lambda}(K_{\mathfrak{q}}, \mathbb{T}) = H^1(K_{\mathfrak{q}}, \mathbb{T})$ for each prime $\mathfrak{q}$ of K.

Remark 6.4. Since $H^1(K_{\mathfrak{q}}, \mathbb{T}) = H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, \mathbb{T})$ for each prime $\mathfrak{q} \nmid p$ of K, we have

$$H^1_{\mathcal{F}_\Lambda}(K, \mathbb{T}) = H^1(K, \mathbb{T}) = H^1(K_\Sigma/K, \mathbb{T}).$$

Here K_Σ is the maximal extension of K unramified outside Σ and put

$$H^1(K_\Sigma/K, M) := H^1(\mathrm{Gal}(K_\Sigma/K), M)$$

for any continuous $\mathrm{Gal}(K_\Sigma/K)$-module M. Furthermore, the maps $\mathbb{T} \to \mathbb{T}/J_n\mathbb{T}$ induce isomorphisms

$$H^1(K, \mathbb{T}) \xrightarrow{\sim} \varprojlim_{n \geq 1} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$$

and $H^1_{\mathcal{F}^*}(K, \mathbb{T}^\vee(1)) = H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))$. Note that $H^1(K_\Sigma/K, \mathbb{T})$, $H^2(K_\Sigma/K, \mathbb{T})$, and $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ are finitely generated Λ-modules; see [Perrin-Riou 1992, Section 3].

Remark 6.5. Since $\mathcal{T}$ satisfies condition (H.1), the $\mathcal{O}$-module $H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda/(X)$ is free and the map $H^1(K, \mathbb{T}) \xrightarrow{\times X} H^1(K, \mathbb{T})$ is injective. Thus $H^1(K, \mathbb{T})$ is a free Λ-module.

Proposition 6.6. *Let $\mathfrak{q}$ be a height-1 prime of Λ with $\mathfrak{q} \neq p\Lambda$ and let $S_{\mathfrak{q}}$ denote the integral closure of $\Lambda/\mathfrak{q}$ in its field of fractions. If $H^2(K_\Sigma/K, \mathbb{T})[\mathfrak{q}]$ is finite, then we have*

$$\mathrm{rank}_\Lambda(H^1(K, \mathbb{T})) = r + \mathrm{rank}_{S_{\mathfrak{q}}}(H^1_{\mathcal{F}^*_{\mathrm{can}}}(K, (\mathbb{T} \otimes_\Lambda S_{\mathfrak{q}})^\vee(1))^\vee),$$

where $\mathcal{F}_{\mathrm{can}}$ is the canonical Selmer structure defined in Example 3.4.

Proof. Note that $\mathbb{T} \otimes_\Lambda S_\mathfrak{q}$ satisfies Hypothesis 3.12. Since $H^2(K_\Sigma/K, \mathbb{T})[\mathfrak{q}]$ is finite, by the same proof as [Mazur and Rubin 2004, Proposition 5.3.14], one can show that the map

$$H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda/\mathfrak{q} \to H^1_{\mathcal{F}_{\mathrm{can}}}(K, \mathbb{T} \otimes_\Lambda S_\mathfrak{q})$$

induced by the map $\mathbb{T} \to \mathbb{T} \otimes_\Lambda S_\mathfrak{q}$ is injective and its cokernel is finite. Since $H^1(K, \mathbb{T})$ is a free Λ-module by Remark 6.5, we have

$$\mathrm{rank}_\Lambda(H^1(K, \mathbb{T})) = \mathrm{rank}_{S_\mathfrak{q}}(H^1_{\mathcal{F}_{\mathrm{can}}}(K, \mathbb{T} \otimes_\Lambda S_\mathfrak{q})).$$

Since $H^2(K_\mathfrak{p}, \mathcal{T}) = 0$ for any prime $\mathfrak{p} \mid p$ of K, the core rank of $\mathcal{F}_{\mathrm{can}}$ on $S_\mathfrak{q}$ is r by Example 3.20. Thus by the same proof as [Mazur and Rubin 2004, Corollary 5.2.6], we have

$$\mathrm{rank}_{S_\mathfrak{q}}(H^1_{\mathcal{F}_{\mathrm{can}}}(K, \mathbb{T} \otimes_\Lambda S_\mathfrak{q})) = r + \mathrm{rank}_{S_\mathfrak{q}}(H^1_{\mathcal{F}^*_{\mathrm{can}}}(K, (\mathbb{T} \otimes_\Lambda S_\mathfrak{q})^\vee(1))^\vee). \qquad \square$$

Proposition 6.7. *The Λ-module $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is torsion if and only if $H^1(K, \mathbb{T})$ is a free Λ-module of rank r.*

Proof. Since $H^2(K_\Sigma/K, \mathbb{T})$ is a finitely generated Λ-module, there is a height-1 prime $\mathfrak{q} \neq p\Lambda$ of Λ such that $H^2(K_\Sigma/K, \mathbb{T})[\mathfrak{q}]$ is finite. Then by the same proof as [Mazur and Rubin 2004, Proposition 5.3.14], we see that the kernel and the cokernel of the map

$$H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee \otimes_\Lambda \Lambda/\mathfrak{q} \to H^1_{\mathcal{F}^*_{\mathrm{can}}}(K, (\mathcal{T} \otimes_\mathcal{O} S_\mathfrak{q})^\vee(1))^\vee$$

induced by the map $(\mathbb{T} \otimes_\Lambda S_\mathfrak{q})^\vee(1) \to \mathbb{T}^\vee(1)$ are finite. Thus if $H^1(K, \mathbb{T})$ is free of rank r, then $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee \otimes_\Lambda \Lambda/\mathfrak{q}$ is finite by Proposition 6.6. Hence $H^1_{\mathcal{F}_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module. If $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module, then there is a height-1 prime $\mathfrak{q} \neq p\Lambda$ of Λ such that both $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee \otimes_\Lambda \Lambda/\mathfrak{q}$ and $H^2(K_\Sigma/K, \mathbb{T})[\mathfrak{q}]$ are finite. Thus by Remark 6.5 and Proposition 6.6, $H^1(K, \mathbb{T})$ is a free Λ-module of rank r. $\square$

Remark 6.8. The assertion that $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module is a form of the weak Leopoldt conjecture; see [Perrin-Riou 2000, Section 1.3]. When $T = \mathbb{Z}_p(1)$, this assertion is closely related to boundedness of $\{\delta(K_n)\}_{n \in \mathbb{Z}_{>0}}$; see [Neukirch et al. 2008, Theorem 10.3.22]. Here K_n is the n-th layer of K_∞/K and $\delta(K_n)$ is the Leopoldt defect of K_n.

Lemma 6.9. *Let n be a positive integer. Then we have an exact sequence*

$$0 \to H^2(K_\Sigma/K, \mathbb{T})[J_n] \to H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda/J_n \to H^1(K_\Sigma/K, \mathbb{T}/J_n\mathbb{T}).$$

Proof. Note that, by using the exact sequence $0 \to \mathbb{T}/X^n\mathbb{T} \xrightarrow{\times p^n} \mathbb{T}/X^n\mathbb{T} \to \mathbb{T}/J_n\mathbb{T} \to 0$, we have an exact sequence

$$0 \to H^1(K_\Sigma/K, \mathbb{T}/X^n\mathbb{T}) \xrightarrow{\times p^n} H^1(K_\Sigma/K, \mathbb{T}/X^n\mathbb{T}) \to H^1(K_\Sigma/K, \mathbb{T}/J_n\mathbb{T})$$

since we assume $(\mathbb{T}/(\varpi, X)\mathbb{T})^{G_K} = (\mathcal{T}/\varpi\mathcal{T})^{G_K} = 0$, where ϖ is a uniformizer of $\mathcal{O}$. Furthermore, the exact sequence $0 \to \mathbb{T} \xrightarrow{\times X^n} \mathbb{T} \to \mathbb{T}/X^n\mathbb{T} \to 0$ induces an exact sequence

$$0 \to H^1(K_\Sigma/K, \mathbb{T}) \otimes_\Lambda \Lambda/X^n\Lambda \to H^1(K_\Sigma/K, \mathbb{T}/X^n\mathbb{T}) \to H^2(K_\Sigma/K, \mathbb{T})[X^n] \to 0.$$

Applying $- \otimes_{\Lambda/(X^n)} \Lambda/J_n$ to the above exact sequence, we get the exact sequence

$$0 \to H^2(K_\Sigma/K, \mathbb{T})[J_n] \to H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda/J_n \to H^1(K_\Sigma/K, \mathbb{T}/X^n\mathbb{T}) \otimes_{\Lambda/(X^n)} \Lambda/J_n. \qquad \square$$

The following theorem is a generalization of [Mazur and Rubin 2004, Theorems 5.3.6 and 5.3.10]. However, the proof of this theorem is different from the proof in that paper.

Theorem 6.10. *If ϵ is a primitive Λ-adic Stark system and i is a nonnegative integer, then we have*

$$I_i(\epsilon) = \mathrm{Fitt}^i_\Lambda(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee).$$

In particular, $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module if and only if there is a Λ-adic Stark system $\eta = \{\eta^{(n)}\}_{n\geq 1}$ such that $\eta_1^{(n)} \neq 0$ for some positive integer n.

Proof. The first assertion follows from $H^1_{\mathcal{F}^*}(K, \mathbb{T}^\vee(1)) = H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))$ and Theorem 5.6. Since

$$\mathrm{Ann}_\Lambda(M)^m \subseteq \mathrm{Fitt}^0_\Lambda(M) \subseteq \mathrm{Ann}_\Lambda(M)$$

for any Λ-module M which is generated by m elements, $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module if and only if $I_0(\epsilon) \neq 0$. Furthermore, by Proposition 4.12(2), there is a Λ-adic Stark system $\eta = \{\eta^{(n)}\}_{n\geq 1}$ such that $\eta_1^{(n)} \neq 0$ for some positive integer n if and only if $I_0(\epsilon) \neq 0$. $\square$

We will show that Theorem 6.10 implies [Mazur and Rubin 2004, Theorems 5.3.6 and 5.3.10]. Assume $r = 1$. Let n be a positive integer. By Matlis duality, the natural map

$$H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T}) \to \bigcap\nolimits^1_{\Lambda/J_n} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$$

is an isomorphism. Hence we have a map

$$\Theta^{(n)} : \boldsymbol{SS}_1(\mathbb{T}/J_n\mathbb{T}, \mathcal{F}_n, \mathcal{P}_{n^2}(\mathcal{F}_n)) \longrightarrow \bigcap\nolimits^1_{\Lambda/J_n} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T}) \xleftarrow{\sim} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T}),$$

where the first map is a projection map. It is easy to see that the maps $\Theta^{(n)}$ induce a map $\Theta : \boldsymbol{SS}_1(\mathbb{T}) \to H^1(K, \mathbb{T})$.

Proposition 6.11. *Let $H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}}$ denote the maximal finite Λ-submodule of $H^2(K_\Sigma/K, \mathbb{T})$. Suppose that $r = 1$. If $\mathrm{im}(\Theta) \neq 0$, then $H^1(K, \mathbb{T})$ is free of rank 1 and $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module. Furthermore, we have*

$$\mathrm{Fitt}^0_\Lambda(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee) = \mathrm{char}_\Lambda(H^1(K, \mathbb{T})/\Lambda\Theta(\epsilon))\,\mathrm{Ann}_\Lambda(H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}})$$

for any primitive Λ-adic Stark system ϵ. In particular,

$$\mathrm{char}_\Lambda(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee) = \mathrm{char}_\Lambda(H^1(K, \mathbb{T})/\Lambda\Theta(\epsilon)).$$

Proof. Since the map $H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T}) \to \bigcap^1_{\Lambda/J_n} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$ is an isomorphism for any positive integer n and $\mathrm{im}(\Theta) \neq 0$, the Λ-module $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion by Theorem 6.10. Thus $H^1(K, \mathbb{T})$ is a free Λ-module of rank 1 by Proposition 6.7. Let $\epsilon = \{\epsilon^{(n)}\}_{n\geq 1}$ be a primitive Λ-adic Stark system. Put $\mathrm{Ind}(\epsilon) := \mathrm{char}_\Lambda(H^1(K, \mathbb{T})/\Lambda\Theta(\epsilon))$. Then we have

$$\mathrm{Ind}(\epsilon) = \{f(\Theta(\epsilon)) \mid f \in \mathrm{Hom}_\Lambda(H^1(K, \mathbb{T}), \Lambda)\}.$$

Let n be a positive integer and $\Lambda_n := \Lambda/J_n$. We will compute the ideal $I_0(\epsilon^{(n)}) \subseteq \Lambda_n$. Let M_n be the image of the map $H^1(K, \mathbb{T}) \to H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$. Fix an isomorphism $H^1(K, \mathbb{T}) \simeq \Lambda$. Let Z_n be the image of the injective map $H^2(K_\Sigma/K, \mathbb{T})[J_n] \to H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda_n \simeq \Lambda_n$. By Lemma 6.9, the fixed isomorphism $H^1(K, \mathbb{T}) \simeq \Lambda$ induces an isomorphism $\Lambda_n/Z_n \xrightarrow{\sim} M_n$. Let $\Theta(\epsilon)_n = \Theta(\epsilon) \bmod J_n \in H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$. Since the map $H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})^* \to M_n^*$ is surjective and $\Theta(\epsilon)_n \in M_n$, we have

$$I_0(\epsilon^n) = \{f(\Theta(\epsilon)_n) \mid f \in M_n^*\} = \mathrm{Ind}(\epsilon)\Lambda_n[Z_n].$$

By the definition of the ideal Z_n, we have

$$\ker(\Lambda \to \Lambda_n/\Lambda_n[Z_n]) = \mathrm{Ann}_\Lambda(H^2(K_\Sigma/K, \mathbb{T})[J_n]).$$

Since the Λ-module $H^2(K_\Sigma/K, \mathbb{T})$ is of finite type, we have

$$\ker(\Lambda \to \Lambda_n/\Lambda_n[Z_n]) = \mathrm{Ann}_\Lambda(H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}})$$

for all sufficiently large integers n. Thus we conclude that

$$I_0(\epsilon) = \varprojlim_{n\geq 1} \mathrm{Ind}(\epsilon)\Lambda_n[Z_n] = \mathrm{Ind}(\epsilon)\,\mathrm{Ann}_\Lambda(H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}}).$$

Hence this proposition follows from Theorem 6.10. □

Proposition 6.12. *For a positive integer n, let*

$$\xi_n : \bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda/J_n \to \bigcap\nolimits^r_{\Lambda/J_n} H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$$

denote a natural map. If $H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}} = 0$, there is a unique Λ-module homomorphism

$$\Theta : \boldsymbol{SS}_r(\mathbb{T}) \to \bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T})$$

such that $\xi_n(\Theta(\epsilon) \bmod J_n) = \epsilon_1^{(n)}$ for any positive integer n and Λ-adic Stark system $\epsilon = \{\epsilon^{(n)}\}_{n\geq 1}$. Furthermore, if $\mathrm{im}(\Theta) \neq 0$, then $H^1(K, \mathbb{T})$ is free of rank r, $H^1_{\mathcal{F}^*}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module, and

$$\mathrm{Fitt}^0_\Lambda(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee) = \mathrm{char}_\Lambda\left(\bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T}) \Big/ \Lambda\Theta(\epsilon)\right)$$

for any primitive Λ-adic Stark system ϵ.

Proof. Let n be a positive integer and $s = \mathrm{rank}_\Lambda(H^1(K, \mathbb{T}))$. Note that $s \geq r$ by Proposition 6.6. Let $\Lambda_n := \Lambda/J_n$. By Lemma 6.9 and $H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}} = 0$, the map $H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda_n \to H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$ is a split injection. Thus $H^1_{\mathcal{F}_n}(K, \mathbb{T}/J_n\mathbb{T})$ has a free Λ_n-submodule of rank s and the map ξ_n is injective.

Let $\epsilon = \{\epsilon^{(n)}\}_{n\geq 1}$ be a Λ-adic Stark system. By Proposition 4.11, there is a unique element $\Theta(\epsilon)_n \in \bigwedge^r_\Lambda H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda_n$ such that $\xi_n(\Theta(\epsilon)_n) = \epsilon_1^{(n)}$. Since the set $\{\Theta(\epsilon)_n\}_{n\geq 1}$ becomes an inverse system, we have the desired map $\Theta(\epsilon) := \varprojlim_{n\geq 1} \Theta(\epsilon)_n$.

Let ϵ be a primitive Λ-adic Stark system. If $\mathrm{im}(\Theta) \neq 0$, we have $s = r$ by Proposition 4.11. Hence $H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee$ is a torsion Λ-module by Proposition 6.7. Furthermore, we have

$$\begin{aligned}\mathrm{Fitt}^0_\Lambda\Big(\bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T}) \Big/ \Lambda\Theta(\epsilon)\Big)\Lambda_n &= \mathrm{Fitt}^0_{\Lambda_n}\Big(\bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T}) \otimes_\Lambda \Lambda_n \Big/ \Lambda_n\Theta(\epsilon)_n\Big)\\ &= \mathrm{Fitt}^0_{\Lambda_n}(H^1_{\mathcal{F}^*_n}(K, (\mathbb{T}/J_n\mathbb{T})^\vee(1))^\vee)\\ &= \mathrm{Fitt}^0_\Lambda(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee)\Lambda_n,\end{aligned}$$

where the second equality follows from Proposition 4.11 and the third equality follows from Lemma 3.14. Since Λ is a complete noetherian local ring, this completes the proof. □

Remark 6.13. We use the same notation as in Example 6.2. If we assume the Rubin–Stark conjecture, then we have a primitive Λ-adic stark system $\epsilon^{\text{R-S}}$; see [Büyükboduk 2011, Section 4.3; 2009] and Example 6.2. Since $H^2(K_\Sigma/K, \mathbb{T})_{\mathrm{fin}} = 0$, we have

$$\mathrm{char}(H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee) = \mathrm{char}\Big(\bigwedge\nolimits^r_\Lambda H^1(K, \mathbb{T}) \Big/ \Lambda\Theta(\epsilon^{\text{R-S}})\Big)$$

by Proposition 6.12. Let $H^1(K_p, \mathbb{T}) = \bigoplus_{\mathfrak{p}\mid p} H^1(K_\mathfrak{p}, \mathbb{T})$ and $\mathrm{loc}_p : H^1(K, \mathbb{T}) \to H^1(K_p, \mathbb{T})$ denote the localization map at p. Suppose that the map loc_p is injective. Then by Theorem 3.7, we obtain an exact sequence

$$0 \to H^1(K, \mathbb{T}) \to H^1(K_p, \mathbb{T}) \to H^1_{\mathcal{F}^*_{\mathrm{str}}}(K, \mathbb{T}^\vee(1))^\vee \to H^1_{\mathcal{F}^*_\Lambda}(K, \mathbb{T}^\vee(1))^\vee \to 0.$$

Here the Selmer structure $\mathcal{F}_{\mathrm{str}}$ is defined in [Büyükboduk 2011, Proposition 4.11]. Then we have $\mathrm{rank}_\Lambda(H^1(K, \mathbb{T})) = \mathrm{rank}_\Lambda(H^1(K_p, \mathbb{T})) = r$ by Proposition 6.12 and [loc. cit., Proposition 4.7(i)]. Thus we get

$$\mathrm{char}(H^1_{\mathcal{F}^*_{\mathrm{str}}}(K, \mathbb{T}^\vee(1))^\vee) = \mathrm{char}\Big(\bigwedge\nolimits^r_\Lambda H^1(K_p, \mathbb{T}) \Big/ \Lambda \cdot \mathrm{loc}_p(\Theta(\epsilon^{\text{R-S}}))\Big).$$

Therefore we see that Proposition 6.12 implies [loc. cit., Theorem 4.15].

Acknowledgement

The author would like to express his gratitude to his supervisor Professor Takeshi Tsuji for much helpful advice. The author is also grateful to an anonymous referee for careful reading and many helpful comments. The author is supported by the FMSP program at the University of Tokyo.

References

[Burns and Sano 2016] D. Burns and T. Sano, "On the theory of higher rank Euler, Kolyvagin and Stark systems", preprint, 2016. arXiv

[Burns et al. 2016] D. Burns, M. Kurihara, and T. Sano, "On Stark elements of arbitrary weight and their p-adic families", preprint, 2016. arXiv

[Büyükboduk 2009] K. Büyükboduk, "Stark units and the main conjectures for totally real fields", *Compos. Math.* **145**:5 (2009), 1163–1195. MR Zbl

[Büyükboduk 2011] K. Büyükboduk, "Λ-adic Kolyvagin systems", *Int. Math. Res. Not.* **2011**:14 (2011), 3141–3206. MR Zbl

[Büyükboduk 2014] K. Büyükboduk, "Main conjectures for CM fields and a Yager-type theorem for Rubin–Stark elements", *Int. Math. Res. Not.* **2014**:21 (2014), 5832–5873. MR Zbl

[Büyükboduk 2016] K. Büyükboduk, "Deformations of Kolyvagin systems", *Ann. Math. Qué.* **40**:2 (2016), 251–302. MR Zbl

[Büyükboduk 2018] K. Büyükboduk, "On the Iwasawa theory of CM fields for supersingular primes", *Trans. Amer. Math. Soc.* **370**:2 (2018), 927–966. MR Zbl

[Büyükboduk and Lei 2015] K. Büyükboduk and A. Lei, "Coleman-adapted Rubin–Stark Kolyvagin systems and supersingular Iwasawa theory of CM abelian varieties", *Proc. Lond. Math. Soc.* (3) **111**:6 (2015), 1338–1378. MR Zbl

[Mazur and Rubin 2004] B. Mazur and K. Rubin, *Kolyvagin systems*, Mem. Amer. Math. Soc. **799**, 2004. MR Zbl

[Mazur and Rubin 2016] B. Mazur and K. Rubin, "Controlling Selmer groups in the higher core rank case", *J. Théor. Nombres Bordeaux* **28**:1 (2016), 145–183. MR Zbl

[Neukirch et al. 2008] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, 2nd ed., Grundlehren der Math. Wissenschaften **323**, Springer, 2008. MR Zbl

[Perrin-Riou 1992] B. Perrin-Riou, "Théorie d'Iwasawa et hauteurs p-adiques", *Invent. Math.* **109**:1 (1992), 137–185. MR Zbl

[Perrin-Riou 2000] B. Perrin-Riou, *p-adic L-functions and p-adic representations*, SMF/AMS Texts and Monographs **3**, Amer. Math. Soc., Providence, RI, 2000. MR Zbl

[Rubin 2000] K. Rubin, *Euler systems*, Ann. Math. Studies **147**, Princeton Univ. Press, 2000. MR Zbl

[Sano 2014] T. Sano, "A generalization of Darmon's conjecture for Euler systems for general p-adic representations", *J. Number Theory* **144** (2014), 281–324. MR Zbl

Communicated by Christopher Skinner
Received 2017-04-19 Revised 2018-02-26 Accepted 2018-08-23

sakamoto@ms.u-tokyo.ac.jp *Graduate School of Mathematical Sciences, The University of Tokyo, Tokyo, Japan*

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2327

Jordan blocks of cuspidal representations of symplectic groups

Corinne Blondel, Guy Henniart and Shaun Stevens

Let G be a symplectic group over a nonarchimedean local field of characteristic zero and odd residual characteristic. Given an irreducible cuspidal representation of G, we determine its Langlands parameter (equivalently, its *Jordan blocks* in the language of Mœglin) in terms of the local data from which the representation is explicitly constructed, up to a possible unramified twist in each block of the parameter. We deduce a ramification theorem for G, giving a bijection between the set of *endoparameters* for G and the set of restrictions to wild inertia of discrete Langlands parameters for G, compatible with the local Langlands correspondence. The main tool consists in analyzing the Hecke algebra of a good cover, in the sense of Bushnell–Kutzko, for parabolic induction from a cuspidal representation of $G \times \mathrm{GL}_n$, seen as a maximal Levi subgroup of a bigger symplectic group, in order to determine reducibility points; a criterion of Mœglin then relates this to Langlands parameters.

Introduction

0.1. Let F be a locally compact nonarchimedean local field of odd residual characteristic and denote by W_F the Weil group of F. Let G be the symplectic group preserving a nondegenerate alternating form on a $2N$-dimensional F-vector space. The local Langlands conjectures for G (now a theorem of [Arthur 2013] when F has characteristic zero) stipulate that to an irreducible (smooth, complex) representation π of G is attached a Langlands parameter, and the representations with a given parameter form a finite set of isomorphism classes, called an L-packet for G.

MSC2010: primary 22E50; secondary 11F70.
Keywords: local Langlands correspondence, symplectic group, p-adic group, Jordan block, endoparameter, types and covers.

Since the symplectic group is split, Langlands parameters for G are simply continuous homomorphisms ϕ from $W_F \times \mathrm{SL}_2(\mathbb{C})$ into the dual group $\widehat{G} = \mathrm{SO}_{2N+1}(\mathbb{C})$, taken up to conjugation, such that the $(2N+1)$-dimensional representation $\iota \circ \phi$ of $W_F \times \mathrm{SL}_2(\mathbb{C})$, obtained from the inclusion ι of $\widehat{G}$ into $\mathrm{GL}_{2N+1}(\mathbb{C})$, is semisimple. If π is a discrete series representation of G, then its parameter ϕ is discrete, that is, the image of ϕ is not contained in a proper parabolic subgroup of $\widehat{G}$; equivalently, $\iota\circ\phi$ is the direct sum of inequivalent irreducible orthogonal representations of $W_F \times \mathrm{SL}_2(\mathbb{C})$, and has determinant 1. In that case giving $\iota \circ \phi$ up to equivalence is the same as giving ϕ up to conjugation in $\widehat{G}$.

On the other hand, we have an explicit description of the cuspidal representations of G via the theory of types [Stevens 2008], in the spirit of the classification of the irreducible representations of $\mathrm{GL}_n(F)$ of [Bushnell and Kutzko 1993]. It is our goal in this paper to describe as much as possible of the Langlands parameter of a cuspidal representation of G from its explicit construction. We will denote by $\mathrm{Cusp}(G)$ the set of equivalence classes of cuspidal representations of G, and by $\Phi^{\mathrm{cusp}}(G)$ the subset of discrete Langlands parameters consisting of those parameters with a cuspidal representation in the corresponding L-packet (see Section 0.5 below for a more detailed description).

0.2. At the technical and arithmetic heart of the construction of cuspidal representations of G and $\mathrm{GL}_n(F)$ is the theory of *endoclasses* of simple characters — families of very special characters of compact open subgroups. An irreducible cuspidal representation of $\mathrm{GL}_n(F)$ contains, up to conjugacy, a unique such simple character and thus determines an endoclass. By considering the endoclasses in its cuspidal support, an arbitrary irreducible representation of $\mathrm{GL}_n(F)$ then determines a formal sum of endoclasses (with multiplicities), which we call an *endoparameter of degree n* (see Section 2.7). We write $\mathcal{E}\mathcal{E}_n(F)$ for the set of endoparameters of degree n.

Similarly, an irreducible cuspidal representation of G is constructed from a *semi*simple character, and thus also comes from an endoparameter, the weighted formal sum of the endoclasses of its simple components; moreover, the semisimple character is self-dual so that every endoclass appearing must also be self-dual. Thus the construction of an irreducible cuspidal representation of G gives rise to a self-dual endoparameter of degree $2N$. We write $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F)$ for the set of these self-dual endoparameters.

0.3. The notions of endoclass and endoparameter admit an instructive interpretation via the local Langlands correspondence. Denote by $\mathscr{P}_F$ the wild ramification subgroup of the Weil group W_F. Then the (first) ramification theorem [Bushnell and Henniart 2003, 8.2, Theorem] says that there is a unique bijection between the set of endoclasses over F and the set of W_F-orbits of irreducible complex representations of $\mathscr{P}_F$, which is compatible with the local Langlands correspondence for general linear groups. This then induces a bijection, again compatible with the Langlands correspondence, between the set of endoparameters of degree n and the set of equivalence classes of n-dimensional complex representations of $\mathscr{P}_F$ which are invariant under conjugation by W_F (see 7.3 Theorem for a precise statement). We call these representations of $\mathscr{P}_F$ *wild parameters*.

Our first main result (or, rather, the last in the scheme of proof) is an analogous ramification theorem for the symplectic group G. First we see that the bijection above restricts to a bijection between self-dual endoclasses and self-dual W_F-orbits of irreducible complex representations of $\mathscr{P}_F$. (Note that we really mean that the *orbit* is self-dual: the only self-dual irreducible complex representation of $\mathscr{P}_F$ is the trivial representation, since p is odd.) We say that a $(2N+1)$-dimensional wild parameter is *discrete self-dual* if it is a sum of self-dual W_F-orbits of irreducible complex representations of $\mathscr{P}_F$, and write $\Psi^{\mathrm{sd}}_{2N+1}(F)$ for the set of such wild parameters. These are precisely the restrictions to wild inertia of discrete Langlands parameters. We prove the following ramification theorem for G (see the end of the Introduction for remarks on the characteristic).

7.6 Theorem. *Suppose F is of characteristic zero. There is a unique bijection $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F) \to \Psi^{\mathrm{sd}}_{2N+1}(F)$ which is compatible with the Langlands correspondence for cuspidal representations of G:*

$$\begin{array}{ccc} \mathrm{Cusp}(G) & \twoheadrightarrow & \Phi^{\mathrm{cusp}}(G) \\ \downarrow & & \downarrow \\ \mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F) & \xrightarrow{\sim} & \Psi^{\mathrm{sd}}_{2N+1}(F) \end{array}$$

The bijection here is not just that in the case of general linear groups (indeed, the degree has changed): one must first take the square of every endoclass in the support of the endoparameter, then map across using the bijection for general linear groups, and finally add the trivial representation of $\mathscr{P}_F$.

0.4. The ramification theorem for G is in fact a consequence of rather more precise results, proved on the automorphic side of the Langlands correspondence. To explain the connection, we recall in more detail the structure of discrete Langlands parameters, and the results of Mœglin.

There is, up to isomorphism, exactly one irreducible m-dimensional representation St_m of $\mathrm{SL}_2(\mathbb{C})$ for each $m \geq 1$. Thus an irreducible representation of $W_F \times \mathrm{SL}_2(\mathbb{C})$ is a tensor product $\sigma \otimes \mathrm{St}_m$, where σ is an irreducible representation of W_F; moreover it is orthogonal if and only if either σ is self-dual symplectic and m is even, or σ is self-dual orthogonal and m is odd. By the Langlands correspondence for GL_n [Laumon et al. 1993; Harris and Taylor 2001; Henniart 2000], such a σ is the Langlands parameter of a (single) cuspidal representation ρ of $\mathrm{GL}_n(F)$, where $n = \dim\sigma$. Saying that σ is self-dual is saying that ρ is self-dual (i.e., isomorphic to its contragredient), and σ is then symplectic (resp. orthogonal) if the Langlands–Shahidi L-function $L(s, \Lambda^2, \rho)$ (resp. $L(s, \mathrm{Sym}^2, \rho)$) has a pole at $s = 0$ [Henniart 2010], in which case we say that ρ is of symplectic (resp. orthogonal) type.

Thus, a discrete parameter ϕ for G can be given by a set of (distinct) pairs (ρ_i, m_i), where ρ_i is an isomorphism class of irreducible cuspidal representations of $\mathrm{GL}_{n_i}(F)$, with n_i and m_i positive integers, and

- $\sum_i n_i m_i = 2N+1$,
- each ρ_i is self-dual, of symplectic type if m_i is even and of orthogonal type if m_i is odd,
- if ω_i is the central character of ρ_i then $\prod_i \omega_i^{m_i} = 1$.

0.5. If π is an irreducible cuspidal representation of G and ϕ is its parameter, [Mœglin 2014] gives a criterion to determine the set attached to ϕ as above, i.e., the pairs (ρ_i, m_i) that she calls the “Jordan blocks” of π; we write $\mathrm{Jord}(\pi)$ for this set of pairs. Let us explain her results.

For any positive integer n, the group $\mathrm{GL}_n(F) \times G$ appears naturally as a standard maximal Levi subgroup of $\mathrm{Sp}_{2(N+n)}(F)$. If ρ is a cuspidal representation of $\mathrm{GL}_n(F)$, we can form the parabolically induced representation $\rho\nu^s \rtimes \pi$ (we use normalized induction and induce via the standard parabolic), where s is here a *real* parameter and ν is the character $g \mapsto |\det g|_F$ of $\mathrm{GL}_n(F)$. If no unramified twist of ρ is self-dual then $\rho\nu^s \rtimes \pi$ is always irreducible. On the other hand, if ρ is self-dual, there is a unique $s_\pi(\rho) \geq 0$ such that $\rho\nu^s \rtimes \pi$ is reducible if and only if $s = \pm s_\pi(\rho)$.

We define the *reducibility set* $\mathrm{Red}(\pi)$ to be the set of isomorphism classes of cuspidal representations ρ of some $\mathrm{GL}_n(F)$, with $n \geq 1$, for which $2s_\pi(\rho) - 1$ is a positive integer. Indeed, it is known that $2s_\pi(\rho)$ is an integer [Mœglin and Tadić 2002], so the condition for ρ to lie in $\mathrm{Red}(\pi)$ is that $s_\pi(\rho)$ is neither 0 nor $\frac{1}{2}$. The *Jordan set* $\mathrm{Jord}(\pi)$ is then the set of pairs (ρ, m), where $\rho \in \mathrm{Red}(\pi)$ and $2s_\pi(\rho) - 1 = m + 2k$ for some integer $k \geq 0$.

From its construction, $\mathrm{Jord}(\pi)$ is "without holes" in the sense that if it contains (ρ, m) then it also contains $(\rho, m-2)$ whenever $m - 2 > 0$. However there may be discrete series noncuspidal representations of G with the same parameter as π; this happens as soon as $\mathrm{Jord}(\pi)$ contains a pair (ρ, m) with $m > 1$. For the number of cuspidal representations of G with a given parameter (without holes), see [Mœglin 2011] (recalled in Section 7.4 below).

0.6. The results of Mœglin described in the previous subsection now say that, in order to determine the Langlands parameter of an irreducible cuspidal representation π of G, we need only compute the reducibility points $s_\pi(\rho)$ for ρ an irreducible self-dual representation of some $\mathrm{GL}_n(F)$. Moreover, we need only find *enough* reducibility points $s_\pi(\rho) \geq 1$ to fill the parameter.

To compute these reducibility points, we use Bushnell and Kutzko's theory [1998] of types and covers. The representation π takes the form $\text{c-Ind}_{J_\pi}^G \lambda_\pi$, for some irreducible representation λ_π of a compact open subgroup J_π; this pair (J_π, λ_π) is a *type* for π. Similarly, we have a Bushnell–Kutzko-type $(\tilde{J}_\rho, \tilde{\lambda}_\rho)$ for ρ. Moreover, from [Miyauchi and Stevens 2014] we have a *cover* (J, λ) in $\mathrm{Sp}_{2(N+n)}(F)$ of $(\tilde{J}_\rho \times J_\pi, \tilde{\lambda}_\rho \otimes \lambda_\pi)$.

The reducibility of the parabolically induced representation $\rho\nu^s \rtimes \pi$ for *complex* s is translated, via category equivalence, to the reducibility of induction from modules over the spherical Hecke algebra $\mathcal{H}(\mathrm{GL}_n(F) \times G, \tilde{\lambda}_\rho \otimes \lambda_\pi)$ to $\mathcal{H}(\mathrm{Sp}_{2(N+n)}(F), \lambda)$. The former algebra is isomorphic to $\mathbb{C}[Z^{\pm 1}]$, while the latter is a Hecke algebra on an infinite dihedral group, with two generators each satisfying a quadratic relation of the form $(T+1)(T - q^r)$, with $r \geq 0$ an integer and q the cardinality of the residue field of F. The results of [Blondel 2012] then translate the values of the parameters r for the two generators into the *real parts* of those $s \in \mathbb{C}$ for which $\rho\nu^s \rtimes \pi$ is reducible.

In the inertial class $[\rho] = \{\rho\nu^s \mid s \in \mathbb{C}\}$, there are precisely two inequivalent self-dual representations, and we write ρ' for the other one. Thus the method described above allows one to compute the set $\{s_\pi(\rho), s_\pi(\rho')\}$ but not to distinguish between the two values if they are distinct. Thus our method computes the *inertial Jordan set* $\mathrm{IJord}(\pi)$, which is the *multiset* of pairs $([\rho], m)$ such that $(\rho, m) \in \mathrm{Jord}(\pi)$.

0.7. According to the previous subsection, computing $\mathrm{IJord}(\pi)$ explicitly comes down to computing the parameters in the quadratic relations for the spherical Hecke algebra of the cover. We do this in two steps.

First, we consider the special case when the semisimple character θ_π in π, from which the type (J_π, λ_π) is built, is in fact *simple*. In this case, it determines a self-dual endoclass Θ and we consider only those irreducible cuspidal representations of some $\mathrm{GL}_n(F)$ which have endoclass Θ^2. We prove that just these representations already give us enough to fill the Jordan set (see 2.5 Theorem) and describe an algorithm to determine $\mathrm{IJord}(\pi)$ (see Section 5.10).

Here the computation of the parameters can be done using results of Lusztig on finite reductive groups: if Θ is the trivial endoclass Θ_0, so that we are in depth zero, this was done already in [Lust and Stevens 2016]; otherwise, the groups in question are the reductive quotients of maximal parahoric subgroups in a unitary group (ramified or unramified). There is also an added subtlety which does not arise in the depth-zero case: two signature characters of certain permutations (coming from a comparison of so-called *beta-extensions*) cause an extra twist which must be taken care of in the algorithm and counting.

In the second step, we consider an arbitrary irreducible cuspidal representation π and reduce to the first case. More precisely, the semisimple character θ_π determines by restriction its simple components θ_i for $0 \le i \le l$, and hence endoclasses Θ_i. From the construction of the type (J_π, λ_π), we define types (J_i, λ_i) in symplectic groups $\mathrm{Sp}_{2N_i}(F)$, with $\sum_{i=0}^{l} N_i = N$, which induce to irreducible cuspidal representations π_i containing a simple character of endoclass Θ_i. (See Section 2.6 for details.)

The reduction is obtained by showing that elements of $\mathrm{IJord}(\pi)$ with endoclass Θ_i can be obtained from those of $\mathrm{IJord}(\pi_i)$ by a simple twisting process, by a character of order 1 or 2 (see 2.6 Theorem). This character arises as the comparison of pairs of signature characters as in the first case for π and for π_i; the point that is both crucial and subtle is that, although we need to make two comparisons, they turn out to be equal. Now the first case, together with a dimension count, ensures that we have filled the expected size of $\mathrm{IJord}(\pi)$. If F is of characteristic zero then, by the results of Mœglin, this is indeed the entire inertial Jordan set (see 2.6 Corollary).

0.8. From our explicit description of the set $\mathrm{IJord}(\pi)$, we know the endoclass of every self-dual irreducible cuspidal representation of some $\mathrm{GL}_n(F)$ which appears in $\mathrm{Jord}(\pi)$. From this we deduce the following result, which gives the compatibility of taking endoparameters with the endoscopic transfer from G to $\mathrm{GL}_{2N+1}(F)$ and from which, via the results of Arthur, we deduce compatibility with the local Langlands correspondence. In the following, the map ι_{2N} sends a (self-dual) endoparameter $\sum m_\Theta \Theta$ of degree $2N$ to the endoparameter $\sum m_\Theta \Theta^2 + \Theta_0$ of degree $2N+1$, where Θ_0 denotes the trivial endoclass.

2.8 Theorem. *Suppose F has characteristic* 0*. Then the following diagram commutes*:

$$\begin{tikzcd} \mathrm{Cusp}(G) \arrow[r, "\text{transfer}"] \arrow[d] & \mathrm{Irr}(\mathrm{GL}_{2N+1}(F)) \arrow[d] \\ \mathcal{EE}^{\mathrm{sd}}_{2N}(F) \arrow[r, hook, "\iota_{2N}"'] & \mathcal{EE}_{2N+1}(F) \end{tikzcd}$$

It is very tempting to think that this result could be an instance of a general theory of endoparameters for arbitrary reductive groups, which would be in bijection with suitably defined wild parameters and would be compatible with (twisted) endoscopy.

0.9. Let π be an irreducible cuspidal representation of G. Having given an explicit description of $\mathrm{IJord}(\pi)$, we can ask whether we can then determine $\mathrm{Jord}(\pi)$ precisely; that is, given $([\rho], m) \in \mathrm{IJord}(\pi)$, can we tell whether it is (ρ, m) or (ρ', m) in $\mathrm{Jord}(\pi)$, where ρ' is the self-dual unramified twist of ρ which is inequivalent to ρ. In certain cases the answer is yes: often the representations ρ, ρ' have opposite parities (that is, one is symplectic and the other orthogonal) and then we know that we must have the representation of symplectic type if m is even, and the one of orthogonal type if m is odd. In the exceptional case where ρ, ρ' have the same parity, we can only recover $\mathrm{Jord}(\pi)$ if it happens that *both* appear (that is, $([\rho], m)$ appears in $\mathrm{IJord}(\pi)$ with multiplicity 2); otherwise, we are left with an ambiguity. (See 4.4 Remark for more on this.)

In Section 6, we explore this exceptional case on the Galois side — that is, we look at the self-dual irreducible representations of W_F which have the same parity as their self-dual unramified twist. It turns out that they have quite a special structure and that one can determine their parity (see 6.6 Proposition). This also translates to a criterion for determining the parity of a self-dual cuspidal representation ρ (such that ρ and its self-dual unramified twist ρ' have the same parity), in terms of the type it contains (see Section 6.8).

It is also possible, at least in certain cases, to be more precise in the analysis of the category equivalences and reducibility, in order to elucidate the ambiguity and recover $\mathrm{Jord}(\pi)$ completely. We hope to come back to this in the case of $\mathrm{Sp}_4(F)$ in a sequel to this paper.

0.10. As one of the referees has pointed out, given a *generic* cuspidal representation π of G, it follows from the results of Arthur and Mœglin that, for every (ρ, m) appearing in $\mathrm{Jord}(\pi)$, we have $m = 1$; we say that the Jordan set (or the corresponding L-packet) is *regular* in this case.

In general, determining the genericity of a cuspidal representation of G from the data used in its construction is difficult, as the example of $\mathrm{Sp}_4(F)$ shows; see [Blondel and Stevens 2009]. However, the principal difficulties occur when trying to determine which cuspidal representations in a regular L-packet are generic, rather than in proving that the cuspidals in a nonregular L-packet are nongeneric. Moreover, the case of depth-zero representations is much simpler. Since the appearance of a pair (ρ, m) in $\mathrm{Jord}(\pi)$ with $m > 1$ arises from the "depth-zero data" used in the construction of π, it may be possible to use the techniques of [Blondel and Stevens 2009] to prove that any cuspidal in a nonregular L-packet is nongeneric. We leave this as an interesting question to return to later.

A remark on characteristic. The bulk of our work is on the representation theory of symplectic groups; for this, while we require that the residual characteristic be odd, we have no further conditions on the characteristic — that is, we do *not* require F to be of characteristic zero. In particular, our description of the inertial Jordan set in 2.5 Theorem and 2.6 Theorem does not require characteristic zero. It is only when interpreting these results in terms of the Langlands correspondence (or the endoscopic transfer map) where, until these results have been proved with F of positive characteristic, we require characteristic zero.

Structure of the paper. In Section 1, we recall the basic structure of types for cuspidal representations, in particular semisimple characters and beta-extensions, including the choice of a base point for beta-extensions. Section 2 contains the statements of the main results on (inertial) Jordan sets, remaining entirely on the automorphic side, while the following three sections are devoted to their proofs: in Section 3, we recall the theory of covers and the results of [Blondel 2012; Miyauchi and Stevens 2014] on their Hecke algebras and reducibility of parabolic induction; in Section 4 we prove the reduction to the simple case which is at the heart of our method; and in Section 5 we prove the result in the simple case. The exploration of self-dual irreducible representations of W_F is given in Section 6 and finally, in Section 7, we interpret our results via the local Langlands correspondence.

Notation

Throughout the paper, F will be a locally compact nonarchimedean local field, with ring of integers $\mathfrak{o}_F$, maximal ideal $\mathfrak{p}_F$, and residue field $k_F = \mathfrak{o}_F/\mathfrak{p}_F$ of cardinality $q = q_F$ and *odd* characteristic p; similar notation will be used for extensions of F. The absolute value $|\cdot|_F$ on F is normalized to have image $q^{\mathbb{Z}}$ and we write ν for the character $g \mapsto |\det g|_F$ of $\mathrm{GL}_n(F)$.

All representations we consider here will be smooth and complex. By a *cuspidal* representation of the group of rational points of a connected reductive group over F, we mean a representation which is smooth, irreducible and cuspidal (i.e., killed by all proper Jacquet functors).

1. Cuspidal types and primary beta-extensions

In this section we fix notation following mostly [Stevens 2008]. We recall, in the first sections, the main features of the construction of cuspidal representations of symplectic groups achieved in that paper, to which we refer for relevant definitions. We do not give references for the by now classical definitions and constructions previously made for general linear groups by Bushnell and Kutzko. One of the key steps in the construction is the existence of a so-called *beta-extension*. We will have to compare such beta-extensions across different groups but, unfortunately, they are not uniquely defined. Here, following [Bushnell and Henniart 2005], we explain one way of picking out a particular beta-extension (which we call *p-primary*, see 1.8 Definition) in each case, giving a base point to make comparisons.

1.1. We recall the notation for skew semisimple strata and related objects. Let V be a finite-dimensional symplectic space over F of dimension $2N$. We denote by h the symplectic form on V, by $x \mapsto \bar{x}$ the corresponding adjoint (anti-)involution on $\mathrm{End}_F(V)$ and by σ the corresponding involution on $\mathrm{GL}_F(V)$. We put $G = \mathrm{Sp}_F(V) \simeq \mathrm{Sp}_{2N}(F)$, where $\mathrm{Sp}_F(V)$ is the isometry group of h, which is the group of fixed points of σ in $\mathrm{GL}_F(V)$.

Let $[\Lambda, n, 0, \beta]$ be a skew semisimple stratum in $\mathrm{End}_F(V)$ [Stevens 2008, Definitions 2.4 and 2.5]. In particular Λ is a self-dual $\mathfrak{o}_F$-lattice sequence and $\beta = -\bar{\beta}$ belongs to the Lie algebra $\mathfrak{sp}_F(V)$. We write B for the commuting algebra of β in $\mathrm{End}_F(V)$.

Remark. Following [Stevens 2008] we always normalize self-dual lattice sequences such that their period over any relevant field is even and their duality invariant d is 1. With this convention, for any self-dual lattice sequence Λ and any multiple s of the period e of Λ, there is a unique self-dual lattice sequence of period s having the form $t \mapsto \Lambda((t+a)/(s/e))$. There is thus a well-defined way of summing two self-dual lattice sequences, by first transforming both so that they have the same period; see [Bushnell and Kutzko 1999]. When performing such transformations, the valuation n of β relative to the lattice sequence Λ undergoes changes that are of no importance to us, since the associated groups H^1, J^1, J and characters (see Sections 1.2 and 1.4 below) are left unchanged; we will thus ignore this parameter and write the stratum in the form $[\Lambda, -, 0, \beta]$.

The characteristic spaces of β determine a canonical orthogonal splitting $V = \perp_{i=0}^{l} V^i$ for the stratum $[\Lambda, -, 0, \beta]$ such that, letting $\Lambda^i = \Lambda \cap V^i$ (that is, $\Lambda^i(t) = \Lambda(t) \cap V^i$ for any $t \in \mathbb{Z}$) and $\beta^i = \beta_{|V^i}$, the strata $[\Lambda^i, -, 0, \beta^i]$, $0 \leq i \leq l$, are skew simple strata which are "sufficiently distant" in the sense of [Stevens 2008, Definition 2.4]. We put $E = F[\beta] = \bigoplus_{i=1}^{l} E^i$, where $E^i = F[\beta^i]$, and write $\mathfrak{o}_E^i$ for the ring of integers of E^i. We recall that Λ is an $\mathfrak{o}_E$*-lattice sequence*, by which we mean that each Λ^i is an $\mathfrak{o}_E^i$-lattice sequence in V^i.

Convention. In this paper we also take the convention that, for any skew semisimple stratum $[\Lambda, n, 0, \beta]$ with splitting $V = \perp_{i=0}^{l} V^i$, we have $\beta^0 = 0$. When 0 is not an eigenvalue of β, this can be achieved by taking V^0 to be the zero-dimensional space over F; since, in that case, $\dim_F V^0 = 0$, it does not affect any of the following constructions. The reason for this convention will become apparent later.

1.2. From the datum $[\Lambda, -, 0, \beta]$ are built open compact subrings

- $\tilde{\mathfrak{H}}^1(\beta, \Lambda) \subseteq \tilde{\mathfrak{J}}^1(\beta, \Lambda)$ of $\mathrm{End}_F(V)$,
- $\mathfrak{H}^1(\beta, \Lambda) \subseteq \mathfrak{J}^1(\beta, \Lambda)$ of $\mathfrak{sp}_F(V)$, the fixed points of the former ones under the adjoint involution on $\mathrm{End}_F(V)$,

and open compact subgroups

- $\tilde{H}^1(\beta, \Lambda) \subseteq \tilde{J}^1(\beta, \Lambda) \subset \tilde{J}(\beta, \Lambda)$ of $\mathrm{GL}_F(V)$,
- $H^1(\beta, \Lambda) \subseteq J^1(\beta, \Lambda) \subset J(\beta, \Lambda)$ of G, the subgroups of fixed points of the former ones under the adjoint involution on $\mathrm{GL}_F(V)$.

We will frequently write $H^1_\Lambda = H^1(\beta, \Lambda)$ and so on.

1.3. We introduce more notation relative to Λ. For $n \in \mathbb{Z}$ we write

$$\mathfrak{a}_n(\Lambda) = \{x \in \mathrm{End}_F(V) \mid \forall t \in \mathbb{Z},\ x\Lambda(t) \subseteq \Lambda(t+n)\}, \quad \mathfrak{b}_n(\Lambda) = \mathfrak{a}_n(\Lambda) \cap B.$$

In particular $\mathfrak{a}_0(\Lambda)$ is a hereditary $\mathfrak{o}_F$-order in $\mathrm{End}_F(V)$ with Jacobson radical $\mathfrak{a}_1(\Lambda)$. Let $\tilde{P}(\Lambda) = \mathfrak{a}_0(\Lambda)^\times$ and $\tilde{P}_1(\Lambda) = 1 + \mathfrak{a}_1(\Lambda)$. Then $P_1(\Lambda) = \tilde{P}_1(\Lambda) \cap G$ is the pro-p-radical of $P(\Lambda) = \tilde{P}(\Lambda) \cap G$. The quotient groups

$$\tilde{\mathcal{G}}(\Lambda) = \tilde{P}(\Lambda)/\tilde{P}_1(\Lambda) \quad \text{and} \quad \mathcal{G}(\Lambda) = P(\Lambda)/P_1(\Lambda)$$

are (the groups of rational points of) finite reductive groups over k_F. The latter may be disconnected so we let $\mathcal{G}^0(\Lambda)$ be (the group of rational points of) its neutral component and call $P^0(\Lambda)$ the inverse image of $\mathcal{G}^0(\Lambda)$ in $P(\Lambda)$; this is a parahoric subgroup of G.

Actually we will mainly work with $\mathfrak{b}_0(\Lambda) = \mathfrak{a}_0(\Lambda) \cap B$ and with $P(\Lambda_{\mathfrak{o}_E}) := P(\Lambda) \cap B$, with $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) = P(\Lambda_{\mathfrak{o}_E})/P_1(\Lambda_{\mathfrak{o}_E})$ and its neutral component $\mathcal{G}^0(\Lambda_{\mathfrak{o}_E})$, and with the parahoric subgroup $P^0(\Lambda_{\mathfrak{o}_E})$ of $G_E := B \cap G$, inverse image of $\mathcal{G}^0(\Lambda_{\mathfrak{o}_E})$ in $P(\Lambda_{\mathfrak{o}_E})$. Indeed we have

$$J(\beta, \Lambda) = P(\Lambda_{\mathfrak{o}_E}) J^1(\beta, \Lambda) \quad \text{and} \quad J(\beta, \Lambda)/J^1(\beta, \Lambda) \simeq P(\Lambda_{\mathfrak{o}_E})/P_1(\Lambda_{\mathfrak{o}_E}) = \mathcal{G}(\Lambda_{\mathfrak{o}_E}).$$

Moreover, we have natural isomorphisms

$$\mathcal{G}(\Lambda_{\mathfrak{o}_E}) \simeq \prod_{i=0}^{l} \mathcal{G}(\Lambda^i_{\mathfrak{o}^i_E}) \quad \text{and} \quad \mathcal{G}^0(\Lambda_{\mathfrak{o}_E}) \simeq \prod_{i=0}^{l} \mathcal{G}^0(\Lambda^i_{\mathfrak{o}^i_E}).$$

Note that, writing E^i_{o} for the field of fixed points of E^i under the adjoint involution $x \mapsto \bar{x}$ and k^i_{o} for its residue field, the groups on the right-hand side here are reductive groups over k^i_{o}. We also have similar decompositions and isomorphisms for the group $\tilde{J}(\beta, \Lambda)$.

1.4. On the group $\widetilde{H}^1(\beta, \Lambda)$ lives a family of one-dimensional representations endowed with very strong properties, called *semisimple characters* [Stevens 2008, §3.1], that restricts to a family of *skew semisimple characters* on $H^1(\beta, \Lambda)$. In particular, a skew semisimple character of $H^1(\beta, \Lambda)$, say θ, restricts to a skew *simple* character θ_i of $H^1(\beta, \Lambda) \cap \mathrm{Sp}_F(V^i) = H^1(\beta^i, \Lambda^i)$ for $0 \le i \le l$. Among the properties of these families, the "transfer property" is especially important. It asserts that if $[\Lambda', -, 0, \beta]$ is another skew semisimple stratum in $\mathrm{End}_F(V)$, then there is a canonical bijection between the sets of skew semisimple characters on $H^1(\beta, \Lambda)$ and $H^1(\beta, \Lambda')$ [loc. cit., Proposition 3.2]. The image of θ under this bijection is called the *transfer* of θ.

To any semisimple character $\tilde{\theta}$ of $\widetilde{H}^1(\beta, \Lambda)$ is associated the unique (up to equivalence) irreducible representation $\tilde{\eta}$ of $\tilde{J}^1(\beta, \Lambda)$ that contains $\tilde{\theta}$ upon restriction; actually $\tilde{\eta}$ restricts to a multiple of $\tilde{\theta}$ on $\widetilde{H}^1(\beta, \Lambda)$. Now $\widetilde{H}^1_\Lambda$ and $\tilde{J}^1_\Lambda$ are pro-p-groups with p odd, on which the adjoint involution σ acts. The Glauberman correspondence hence relates their representations to those of the fixed point subgroups H^1_Λ and J^1_Λ. Indeed if $\tilde{\theta}$ is fixed under the involution σ so is $\tilde{\eta}$ and its image η under the Glauberman correspondence is the unique (up to equivalence) irreducible representation of $J^1(\beta, \Lambda)$ that contains θ; it actually restricts to a multiple of θ on $H^1(\beta, \Lambda)$.

1.5. In turn the representation $\tilde{\eta}$ has special extensions to $\tilde{J}(\beta, \Lambda)$ called *beta-extensions* and denoted by $\tilde{\kappa}$. These beta-extensions in $\mathrm{GL}_F(V)$ are characterized by the fact that they are intertwined by $B^\times$ [Bushnell and Kutzko 1993, (5.2.1)].

Remark. In the literature, these extensions are usually called β-extensions. However, the simple stratum $[\Lambda, -, 0, \beta]$ giving rise to a particular simple character θ is not unique, while the notion of beta-extension turns out to be independent of the choice of β. It is thus convenient to write *beta-extension*, especially since we also have strata indexed by i so we would otherwise need to talk about β_i-extensions etc.

The definition of beta-extensions in classical groups is more delicate [Stevens 2008, §4]. A skew semisimple stratum as above is called *maximal* if $\mathfrak{b}_0(\Lambda)$ is a maximal self-dual $\mathfrak{o}_E$-order in B. If $[\Lambda, -, 0, \beta]$

is a maximal skew semisimple stratum, a beta-extension of η is an extension κ of η to $J(\beta, \Lambda)$ such that the restriction of κ to any pro-p-Sylow subgroup is intertwined by G_E [Stevens 2008, Corollary 3.11, Theorem 4.1]. In the general case, the notion of beta-extension is a relative one. Given a maximal skew semisimple stratum $[\mathfrak{M}, -, 0, \beta]$ in $\mathrm{End}_F(V)$ such that $\mathfrak{b}_0(\mathfrak{M}) \supset \mathfrak{b}_0(\Lambda)$, given the transfer $\theta_{\mathfrak{M}}$ of θ to $H^1_{\mathfrak{M}}$ and the representation $\eta_{\mathfrak{M}}$ of $J^1_{\mathfrak{M}}$ determined by $\theta_{\mathfrak{M}}$, there is a canonical way to associate to a beta-extension $\kappa_{\mathfrak{M}}$ of $\eta_{\mathfrak{M}}$, an extension κ of η, called *the beta-extension of η to J_Λ relative to $\mathfrak{M}$, compatible with $\kappa_{\mathfrak{M}}$* [Stevens 2008, Lemma 4.3, Definition 4.5]. (We can also call κ a beta-extension of θ.)

Note that the groups $\tilde{J}(\beta, \Lambda)$ and $J(\beta, \Lambda)$ are not pro-p-groups: the notation κ here should not call to mind a Glauberman-like connection with the former $\tilde{\kappa}$.

1.6. Let $J = J(\beta, \Lambda)$, for a skew semisimple stratum $[\Lambda, -, 0, \beta]$ as above, let θ be a skew semisimple character of $H^1(\beta, \Lambda)$ and let λ be an irreducible representation of J of the form $\lambda = \kappa \otimes \tau$, with κ some beta-extension of θ, and τ the inflation of a cuspidal representation of $J/J^1 \simeq \mathcal{G}(\Lambda_{\mathfrak{o}_E})$. Under the additional assumptions that the group G_E has compact centre and that $P^0(\Lambda_{\mathfrak{o}_E})$ is a maximal parahoric subgroup of G_E, the pair (J, λ) is called a *cuspidal type* for G. Recall from [Stevens 2008] (see also [Miyauchi and Stevens 2014] for complements):

Theorem [Stevens 2008, Corollary 6.19, Theorem 7.14]. *A cuspidal type in G induces to a cuspidal representation of G and any cuspidal representation of G is thus obtained.*

1.7. There is of course a similar result for the group $\mathrm{GL}_F(V)$. Here we let $\tilde{J} = \tilde{J}(\beta, \Lambda)$ for a *simple* stratum $[\Lambda, -, 0, \beta]$ (so that $E = F[\beta]$ is a field) and let $\tilde{\lambda}$ be an irreducible representation of $\tilde{J}$ of the form $\tilde{\lambda} = \tilde{\kappa} \otimes \tilde{\tau}$, with $\tilde{\kappa}$ some beta-extension of $\tilde{\theta}$, and $\tilde{\tau}$ the inflation of a cuspidal representation of $\tilde{J}/\tilde{J}^1$. Under the additional assumptions that $\widetilde{P}(\Lambda) \cap B$ is a maximal parahoric subgroup of $B^\times$, the pair $(\tilde{J}, \tilde{\lambda})$ is called a *maximal simple type* for $\mathrm{GL}_F(V)$.

Theorem [Bushnell and Kutzko 1993, Definition 5.5.10, Theorems 6.2.4 and 8.4.1]. *A maximal simple type in* $\mathrm{GL}_F(V)$ *extends to an irreducible representation of its normalizer, which then induces to a cuspidal representation of* $\mathrm{GL}_F(V)$; *any cuspidal representation ρ of* $\mathrm{GL}_F(V)$ *is thus obtained and the maximal simple type yielding ρ is unique up to conjugacy in* $\mathrm{GL}_F(V)$.

Remark. This theorem includes depth-zero representations, by formally considering the null stratum $[\Lambda, -, 0, 0]$ to be simple.

1.8. In order to compare representations across different groups, we need a way to compare beta-extensions. (The transfer of semisimple characters already allows a comparison.) Two beta-extensions only differ by a character (of a specific shape); however we will need to choose beta-extensions in a unique way as in [Bushnell and Henniart 2005, §2.3, Lemma 1], which amounts to the GL-case in the following lemma.

Lemma. (i) *Let $[\Lambda, n, 0, \beta]$ be a simple stratum in* $\mathrm{End}_F(V)$, *let $\tilde{\theta}$ be a simple character of $\widetilde{H}^1(\beta, \Lambda)$, and let $\tilde{\eta}$ be the irreducible representation of $\tilde{J}^1(\beta, \Lambda)$ containing $\tilde{\theta}$. There exists one and only one beta-extension $\tilde{\kappa}$ of $\tilde{\eta}$ to $\tilde{J}(\beta, \Lambda)$ whose determinant has order a power of p.*

(ii) *With notation as in (i), assume the stratum and the simple character are skew so that the involution* σ *on* $\mathrm{GL}_F(V)$ *stabilizes* $\widetilde{H}^1(\beta, \Lambda)$, $\tilde{J}^1(\beta, \Lambda)$, $\tilde{J}(\beta, \Lambda)$ *and* $\tilde{\theta}$. *The beta-extension* $\tilde{\underline{\kappa}}$ *in (i) satisfies* $\tilde{\underline{\kappa}} \simeq \tilde{\underline{\kappa}} \circ \sigma$.

(iii) *Let* $[\Lambda, n, 0, \beta]$ *be a maximal skew semisimple stratum in* $\mathrm{End}_F(V)$, *let* θ *be a skew semisimple character of* $H^1(\beta, \Lambda)$, *and let* η *be the irreducible representation of* $J^1(\beta, \Lambda)$ *containing* θ. *There exists one and only one beta-extension of* η *to* $J(\beta, \Lambda)$ *whose determinant has order a power of* p.

Proof. (i) The reference is [Bushnell and Kutzko 1993, Theorem 5.2.2], which we imitate below to conclude the proof of (iii).

(ii) Self-duality with respect to σ follows from uniqueness. Indeed $\tilde{\eta} \circ \sigma$ is equivalent to η so there is an intertwining operator T such that $\tilde{\eta}(x) = T(\tilde{\eta} \circ \sigma(x))T^{-1}$ for $x \in \tilde{J}^1(\beta, \Lambda)$. Since σ stabilizes $\mathrm{GL}_{F[\beta]}(V)$, the representation $T(\tilde{\underline{\kappa}} \circ \sigma(x))T^{-1}$ for $x \in \tilde{J}(\beta, \Lambda)$ is a beta-extension of $\tilde{\eta}$ by [Bushnell and Kutzko 1993, Definition 5.2.1]; its determinant is a power of p, so it is equal to $\tilde{\underline{\kappa}}$.

(iii) Let κ be a beta-extension of η and let $\Phi = \det(\kappa_{|P(\Lambda_{\mathfrak{o}_E})})$. The main point is to prove that the character Φ of $P(\Lambda_{\mathfrak{o}_E})$ factors through the determinant $\det_E$. By this we mean, as usual, that $\Phi_{|P(\Lambda_{\mathfrak{o}^i_E})}$ factors through $\det_{E^i}$ for $0 \leq i \leq l$; the remainder of the proof uses this convention.

Since θ is equal to $\chi \circ \det_E$ on $P^1(\Lambda_{\mathfrak{o}_E})$ for some character χ of $1 + \mathfrak{p}_E$, we have that $\kappa_{|P^1(\Lambda_{\mathfrak{o}_E})}$ is the sum of $\dim \eta$ copies of $\chi \circ \det_E$. Now χ extends to a character $\tilde{\chi}$ of $\mathfrak{o}_E^\times$ and $\Phi' = (\tilde{\chi} \circ \det_E)^{-\dim \eta} \Phi$ is then a character of $P(\Lambda_{\mathfrak{o}_E})/P^1(\Lambda_{\mathfrak{o}_E})$. From [Stevens 2008, Lemma 3.10, Corollary 3.11 and Theorem 4.1], the character Φ' is trivial on all p-Sylow subgroups of $P(\Lambda_{\mathfrak{o}_E})/P^1(\Lambda_{\mathfrak{o}_E})$ and so factors as $\Phi' = \psi \circ \det_E$, where ψ is a character of $\mathfrak{o}_E^\times$ trivial on $1 + \mathfrak{p}_E$ (and depends on the choice of extension $\tilde{\chi}$).

Let us write $\mathfrak{o}_E^\times = \mu'_E(1 + \mathfrak{p}_E)$, where μ'_E is the group of roots of unity in $E^\times$ of order prime to p, and, in the above, let us choose $\tilde{\chi}$ trivial on μ'_E so that the order of $\tilde{\chi}$ is a power of p. The corresponding character ψ has order prime to p, so prime to $\dim \eta$, and there is a character α of $\mathfrak{o}_E^\times$ (trivial on $1 + \mathfrak{p}_E$) such that $\psi = \alpha^{\dim \eta}$.

The representation $\underline{\kappa} = (\alpha \circ \det_E)^{-1}\kappa$ satisfies the required condition. It is unique since any other beta-extension has the form $(\psi \circ \det_E)\underline{\kappa}$, with ψ as above, and if ψ is nontrivial then no p^i-th power of ψ can be trivial. $\square$

Definition. With the notations of (i) above, we denote by $\tilde{\underline{\kappa}}$ the unique beta-extension of $\tilde{\eta}$ whose determinant has order a power of p. We call $\tilde{\underline{\kappa}}$ the *p-primary beta-extension* of $\tilde{\eta}$.

With the notations of (iii) above, we denote by $\underline{\kappa}$ the unique beta-extension of η whose determinant has order a power of p. We call $\underline{\kappa}$ the *p-primary beta-extension* of η.

We remark that, while the p-primary beta-extensions give a useful way of picking a base point amongst the beta-extensions, sufficient for our needs here, it is not clear whether this is the best choice of base point.

2. Inertial Jordan blocks

In this section, we state the main results on Jordan blocks and the consequences for the endoscopic transfer map. We continue with the notation from the previous section.

2.1. Let π be a cuspidal representation of $G \simeq \mathrm{Sp}_{2N}(F)$. We recall the *reducibility set* $\mathrm{Red}(\pi)$ and the *Jordan set* $\mathrm{Jord}(\pi)$ from the Introduction. For any positive integer n, the group $\mathrm{GL}_n(F) \times G$ appears naturally as a standard maximal Levi subgroup of $\mathrm{Sp}_{2(N+n)}(F)$. If ρ is a cuspidal representation of $\mathrm{GL}_n(F)$ we can form the normalized parabolically induced representation $\rho\nu^s \rtimes \pi$ (we use normalized induction and induce via the standard parabolic), where s is here a *real* parameter and ν is the character $g \mapsto |\det g|_F$ of $\mathrm{GL}_n(F)$. If no unramified twist of ρ is self-dual (i.e., isomorphic to its contragredient) then $\rho\nu^s \rtimes \pi$ is always irreducible. On the other hand, if ρ is self-dual, there is a unique $s_\pi(\rho) \geq 0$ such that $\rho\nu^s \rtimes \pi$ is reducible if and only if $s = \pm s_\pi(\rho)$.

Definition. Let π be a cuspidal representation of G:

- The *reducibility set* $\mathrm{Red}(\pi)$ is the set of isomorphism classes of self-dual cuspidal representations ρ of some $\mathrm{GL}_n(F)$, with $n \geq 1$, for which $s_\pi(\rho) \geq 1$.
- The *Jordan set* $\mathrm{Jord}(\pi)$ is the set of pairs (ρ, m), where $\rho \in \mathrm{Red}(\pi)$ and m is a positive integer such that $2s_\pi(\rho) - 1 - m$ is a nonnegative even integer.

Note that, if $\rho \in \mathrm{Red}(\pi)$ then $2s_\pi(\rho) - 1$ is a positive integer by [Mœglin and Tadić 2002], so that there is a positive integer m such that $(\rho, m) \in \mathrm{Jord}(\pi)$.

2.2. For ρ an irreducible representation of some $\mathrm{GL}_n(F)$, we write $n = \deg \rho$. Recall that the *inertial class* $[\rho]$ of a cuspidal representation ρ of $\mathrm{GL}_n(F)$ is the equivalence class of ρ under the equivalence relation defined by twisting by an unramified character (that is, twisting by $\omega \circ \det$ where ω is a character of $F^\times$ trivial on $\mathfrak{o}_F^\times$). If ρ is self-dual then the inertial class $[\rho]$ contains precisely two self-dual representations: if $t(\rho)$ denotes the number of unramified characters χ of $\mathrm{GL}_n(F)$ such that $\rho \otimes \chi \simeq \rho$, and if χ' is an unramified character of order $2t(\rho)$, then $\rho' = \rho \otimes \chi'$ is the other self-dual representation in $[\rho]$.

Definition. Let π be a cuspidal representation of G. The *inertial Jordan set* of π is the *multiset* $\mathrm{IJord}(\pi)$ consisting of all pairs $([\rho], m)$ with $(\rho, m) \in \mathrm{Jord}(\pi)$.

Note that, if $([\rho], m) \in \mathrm{IJord}(\pi)$, with ρ a self-dual cuspidal representation of $\mathrm{GL}_n(F)$, then either $(\rho, m) \in \mathrm{Jord}(\pi)$ or $(\rho', m) \in \mathrm{Jord}(\pi)$, where ρ' as above is the second self-dual representation in the inertial class $[\rho]$. As discussed in the Introduction, if one of ρ, ρ' is of symplectic type and the other of orthogonal type, then which occurs in $\mathrm{Jord}(\pi)$ is determined by the parity of m. On the other hand, if ρ, ρ' are both of the same parity then the inertial Jordan set $\mathrm{IJord}(\pi)$ does *not* distinguish them; of course, if $([\rho], m)$ occurs with multiplicity 2 in $\mathrm{IJord}(\pi)$, then both (ρ, m) and (ρ', m) occur in $\mathrm{Jord}(\pi)$ and there is no ambiguity; see 4.4 Remark for more on this.

2.3. In order to refine further the (inertial) Jordan set, we need to use the notion of the *endoclass* of a simple character, as defined in [Bushnell and Henniart 1996]. To any cuspidal representation ρ of $\mathrm{GL}_n(F)$ is attached in [Bushnell and Henniart 2003, §1.4] an endoclass of simple characters, denoted by $\Theta(\rho)$, as follows. As recalled in 1.7 Theorem, there is a maximal simple type $(\tilde{J}, \tilde{\lambda})$ in $\mathrm{GL}_n(F)$ which occurs in ρ, and ρ determines the $\mathrm{GL}_n(F)$-conjugacy class of $(\tilde{J}, \tilde{\lambda})$. This maximal simple type is built from a simple character $\tilde{\theta}$ and we define $\Theta(\rho)$ to be the endoclass of $\tilde{\theta}$. (In fact, this is also the endoclass of *any* simple character contained in ρ.) Note that we are allowing here the case of depth-zero representations (where ρ contains the trivial character of $\widetilde{P}^1(\Lambda)$ for some lattice sequence Λ), in which case $\Theta(\rho) = \Theta_0^F$ is the trivial endoclass over F.

Definition. Let π be a cuspidal representation of G and let Θ be an endoclass of simple characters over F. The *inertial Jordan set of π relative to* Θ is the *multiset* $\mathrm{IJord}(\pi, \Theta)$ consisting of all pairs $([\rho], m)$ with $(\rho, m) \in \mathrm{Jord}(\pi)$ and $\Theta(\rho) = \Theta$.

2.4. We will also need to twist inertial Jordan blocks as follows. With notation as in the previous subsection, the $\mathrm{GL}_n(F)$-conjugacy class of $(\tilde{J}, \tilde{\lambda})$ depends only on the inertial class $[\rho]$; it also determines $[\rho]$ by [Bushnell and Kutzko 1998, (5.5)]. The quotient group $\tilde{J}/\tilde{J}^1$ is a linear group over a finite field, say $\mathrm{GL}(m_{[\rho]}, k_{[\rho]})$. We define the twist of the inertial class $[\rho]$ by a character χ of $k_{[\rho]}^\times$ to be the inertial class $[\rho]_\chi$ determined by the maximal simple type $(\tilde{J}, \tilde{\lambda} \otimes \chi \circ \det)$ — that is, in the decomposition $\tilde{\lambda} = \tilde{\kappa} \otimes \tilde{\tau}$ with $\tilde{\kappa}$ a beta-extension, we replace the cuspidal representation $\tilde{\tau}$ by $\tilde{\tau} \otimes \chi \circ \det$.

Let Θ be an endoclass of simple characters. By [Bushnell and Henniart 1996, Proposition 8.11], it determines a finite extension k_Θ of k_F such that, for any cuspidal representation ρ of some $\mathrm{GL}_n(F)$ satisfying $\Theta(\rho) = \Theta$, if $(\tilde{J}, \tilde{\lambda})$ is a maximal simple type in ρ then the quotient group $\tilde{J}/\tilde{J}^1$ is a linear group over k_Θ (that is, $k_{[\rho]} = k_\Theta$ in the notation above). It is thus meaningful to give the following definition:

Definition. Let π be a cuspidal representation of G, let Θ be an endoclass of simple characters, and let χ be a character of $k_\Theta^\times$. The *χ-twisted inertial Jordan set of π relative to* Θ is the multiset $\mathrm{IJord}(\pi, \Theta)_\chi$ consisting of all pairs $([\rho]_\chi, m)$ with $(\rho, m) \in \mathrm{Jord}(\pi)$ and $\Theta(\rho) = \Theta$.

The relevant case for us will be the case where χ is quadratic (that is, of order dividing 2).

Remark. Since p is odd, we have a squaring map $\Theta \mapsto \Theta^2$ on endoclasses: if θ is a simple character with endoclass Θ, associated to a simple stratum $[\Lambda, -, 0, \beta]$, then the character θ^2 is a simple character for the stratum $[\Lambda, -, 0, 2\beta]$ and Θ^2 is the endoclass corresponding to θ^2. This is well-defined and moreover gives a bijection on the set of endoclasses (again, since p is odd). We note also that the fields k_Θ and k_{Θ^2} coincide.

2.5. We begin the computation of the inertial Jordan set with a special case, to which we will reduce in the next subsection. We call a cuspidal representation of G *simple* if it contains a simple character; that is, it contains a semisimple character θ of $H^1(\beta, \Lambda)$ associated to a skew semisimple stratum $[\Lambda, -, 0, \beta]$ such that $E = F[\beta]$ is a field. We allow the degenerate case $\beta = 0$, in which case π is of depth zero (and every depth-zero representation is simple with $\beta = 0$); we also allow, in the case $\beta = 0$, the degenerate case that G is the trivial group, so that the trivial representation of the trivial group is regarded as being simple of depth zero.

Remark. Our use of the word *simple* here is consistent with, but not the same as, the use in [Gross and Reeder 2010] where, for symplectic groups, it means of minimal positive depth $1/(2N)$. More precisely, all cuspidal representations of depth $1/(2N)$ are simple in our sense, but the converse is false.

The following theorem tells us that, in the case of simple cuspidals, the Jordan set is filled by representations with the expected endoclass.

Theorem. *Let π be a simple cuspidal representation of G and let $\tilde{\theta}$ be a self-dual simple character whose restriction to G is contained in π. Let Θ be the endoclass of the simple character $\tilde{\theta}$. Then*

$$\sum_{([\rho],m)\in \mathrm{IJord}(\pi,\Theta^2)} m \deg \rho = \begin{cases} 2N+1 & \text{if } \Theta \text{ is } \Theta_0^F, \text{ the trivial endoclass,} \\ 2N & \text{otherwise.} \end{cases} \tag{2-1}$$

Note that we have $\Theta = \Theta_0^F$ if and only if π is of depth zero (which includes the degenerate case $N = 0$ where G is the trivial group). In this case, the theorem is a special case of the main result of [Lust and Stevens 2016].

Remark. Since the dual group of G is $\mathrm{SO}_{2N+1}(\mathbb{C})$, the reader may be surprised to see the sum in (2-1) being $2N$ rather than $2N+1$ in most cases. The reason is as follows. The Jordan set of π always contains a pair $(\chi, 1)$, with χ a quadratic character; since χ is tame, it has trivial endoclass and so contributes to the sum in (2-1) if and only if $\Theta = \Theta_0^F$. The point then of Theorem 2.5 is that, apart from this quadratic character, every other cuspidal representation appearing in the Jordan set of π has endoclass Θ^2.

We will prove Theorem 2.5 in Section 5 by computing the real parts of the complex reducibility points of parabolically induced representations of the form $\rho\nu^s \rtimes \pi$, with ρ a self-dual cuspidal representation of some general linear group with endoclass Θ^2, using the theory of types and covers to reduce the calculation to computations of Lusztig for finite reductive groups. We note also that the proof not only gives the equality above but also gives an algorithm to compute the multiset $\mathrm{IJord}(\pi, \Theta^2)$ (see Section 5.10 for more detail).

2.6. Now let π be an arbitrary cuspidal representation of G. Recall from 1.6 Theorem that π can be constructed by induction, starting with a maximal skew semisimple stratum $[\Lambda, -, 0, \beta]$ and a skew semisimple character θ of $H^1(\beta, \Lambda)$, which decomposes into a family of skew simple characters θ_i of $H^1(\beta^i, \Lambda^i)$ for $i \in \{0, \dots, l\}$. Let $\underline{\kappa}$ be the p-primary beta-extension of θ to J_Λ and, similarly, let $\underline{\kappa}_i$ be the p-primary beta-extension of θ_i to J_{Λ^i} (in $\mathrm{Sp}_F(V^i)$) for $0 \le i \le l$.

Let τ be the cuspidal representation of $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) = P(\Lambda_{\mathfrak{o}_E})/P^1(\Lambda_{\mathfrak{o}_E})$ such that π is induced from $\lambda = \underline{\kappa} \otimes \tau$. Then we can uniquely decompose τ as $\tau = \bigotimes_{i=0}^l \tau_i$, with τ_i an irreducible (cuspidal) representation of $\mathcal{G}(\Lambda_{\mathfrak{o}_E^i})$. We may then define, for each i, the cuspidal representation π_i of $\mathrm{Sp}_F(V^i)$ by

$$\pi_i = \text{c-Ind}_{J_{\Lambda^i}}^{\mathrm{Sp}_F(V^i)} \underline{\kappa}_i \otimes \tau_i.$$

Note that this representation is simple, in the sense of the previous subsection.

Remark. Recall that we are using the notation of Section 1.1, in particular 1.1 Convention so that we are assuming $\beta^0 = 0$. If the space V^0 is trivial then the representation π_0 is the trivial representation of the trivial group.

We can now state the crucial reduction theorem, which allows us to determine the inertial Jordan set of π from those of the simple cuspidals π_i.

Theorem. *With notation as above, for $0 \le i \le l$, let $\tilde{\theta}_i$ be the unique self-dual simple character of $\widetilde{H}^1(\beta^i, \Lambda^i)$ restricting to θ_i on $H^1(\beta^i, \Lambda^i)$. Let Θ_i be the endoclass of the simple character $\tilde{\theta}_i$ and let k_{Θ_i} be the corresponding extension of k_F. Then there is a character χ_i of $k_{\Theta_i}^\times$ of order at most 2 such that we have an equality of multisets*

$$\mathrm{IJord}(\pi, \Theta_i^2) = \mathrm{IJord}(\pi_i, \Theta_i^2)_{\chi_i}.$$

The character χ_i appearing here is in some sense explicit, coming from certain permutation characters (see 4.4 Theorem, 4.3 Proposition and 3.10 Proposition for more details). The proof of the theorem will be given in Section 4, following preparation in Section 3 (which is also needed for the proof of 2.5 Theorem). Again, the principle is to use the theory of types and covers to compare the real parts of the complex reducibility points of $\rho\nu^s \rtimes \pi$ with those of $\rho_i\nu^s \rtimes \pi_i$, for ρ a self-dual cuspidal representation of some general linear group with endoclass Θ_i^2 and ρ_i self-dual in the inertial class $[\rho]_{\chi_i}$.

For now, we put together the two previous theorems to get:

Corollary. *Suppose F is of characteristic zero. With the notation of the theorem, we have*

$$\mathrm{IJord}(\pi) = \bigsqcup_{i=0}^{l} \mathrm{IJord}(\pi_i, \Theta_i^2)_{\chi_i}.$$

Since the proof of 2.5 Theorem gives us an algorithm to compute the multisets $\mathrm{IJord}(\pi_i, \Theta_i^2)$, we can then use this also to compute $\mathrm{IJord}(\pi)$ for any cuspidal representation π.

Proof. The theorem says that $\mathrm{IJord}(\pi)$ contains the right-hand side. On the other hand, by [Mœglin 2014, Theorem 3.2.1] the multiset $\mathrm{IJord}(\pi)$ is finite and we have

$$\sum_{([\rho],m)\in\mathrm{IJord}(\pi)} m \deg\rho = \sum_{(\rho,m)\in\mathrm{Jord}(\pi)} m \deg\rho = 2N+1.$$

However, writing $\dim_F V^i = 2N_i$, we get from 2.5 Theorem that

$$\sum_{i=0}^{l} \sum_{([\rho],m)\in\mathrm{IJord}(\pi_i,\Theta_i^2)} m \deg\rho = (2N_0+1) + \sum_{i=1}^{l} 2N_i = 2N+1.$$

Thus we have equality, as required. □

We remark that the proof does not require the full strength of [Mœglin 2014, Theorem 3.2.1]; indeed, it only uses the inequality

$$\sum_{(\rho,m)\in\mathrm{Jord}(\pi)} m \deg\rho \le 2N+1,$$

which was proved previously in [Mœglin 2003, §4, Corollaire]. Thus it does not in fact depend on Arthur's endoscopic classification of discrete series representations of G. One could also prove it (without the restriction on the characteristic of F) by checking that $\mathrm{IJord}(\pi, \Theta)$ is empty for any self-dual endoclass $\Theta \neq \Theta_i^2$; indeed, the methods of Section 4 together with results from [Kurinczuk et al. 2016] would allow this.

2.7. In this and the following subsection, we interpret our results in terms of the endoscopic transfer map from cuspidal representations of G to $\mathrm{GL}_{2N+1}(F)$.

For Θ an endoclass over F, we recall that the *degree* $\deg \Theta$ of Θ is the degree of an extension $F[\beta]/F$ for which there are a simple stratum $[\Lambda, -, 0, \beta]$ with a simple character of endoclass Θ. Although the stratum and the field extension are not uniquely determined by Θ, this degree is; see [Bushnell and Henniart 1996, Proposition 8.11].

Let N' be a positive integer and write $\mathcal{E}(F)$ for the set of endoclasses of simple characters over F. An *endoparameter of degree N' over F* is a formal sum

$$\sum_{\Theta \in \mathcal{E}(F)} m_\Theta \Theta, \quad m_\Theta \in \mathbb{Z}_{\geq 0},$$

such that

$$\sum_{\Theta \in \mathcal{E}(F)} m_\Theta \deg \Theta = N'.$$

In particular, such a formal sum has finite support $\{\Theta \in \mathcal{E}(F) \mid m_\Theta \neq 0\}$. (In [Sécherre and Stevens 2016], these formal sums are called *semisimple endoclasses*; the nomenclature *endoparameter* comes from [Kurinczuk et al. 2016].) We write $\mathcal{EE}_{N'}(F)$ for the set of endoparameters of degree N' over F. We then have, for each positive integer N', a well-defined map

$$\mathsf{e}_{N'} : \mathrm{Irr}(\mathrm{GL}_{N'}(F)) \to \mathcal{EE}_{N'}(F)$$

given by mapping a cuspidal representation ρ to $(N'/\deg \Theta(\rho))\Theta(\rho)$, and mapping an arbitrary representation to the sum of the endoparameters of its cuspidal support.

2.8. We call an endoclass Θ over F *self-dual* if there is a self-dual simple character $\tilde{\theta}$ with endoclass Θ. We write $\mathcal{E}^{\mathrm{sd}}(F)$ for the set of self-dual endoclasses over F. An endoparameter of degree N' over F is called *self-dual* if its support is contained in $\mathcal{E}^{\mathrm{sd}}(F)$, and we write $\mathcal{EE}^{\mathrm{sd}}_{N'}(F)$ for the set of self-dual endoparameters of degree N' over F.

Since p is odd, the only self-dual endoclass over F of odd degree is the trivial endoclass Θ_0^F, which has degree 1. Indeed, if $\tilde{\theta}$ is a self-dual simple character which is not the trivial character, then [Stevens 2001, Theorem 6.3] implies that $\tilde{\theta}$ is associated to a skew simple stratum, whose associated field extension E/F is therefore of even degree. This implies, in particular, that there is a canonical bijection

$$\mathcal{EE}^{\mathrm{sd}}_{2N}(F) \to \mathcal{EE}^{\mathrm{sd}}_{2N+1}(F), \quad \sum_{\Theta \in \mathcal{E}^{\mathrm{sd}}} m_\Theta \Theta \mapsto \sum_{\Theta \in \mathcal{E}^{\mathrm{sd}}} m_\Theta \Theta + \Theta_0^F.$$

For any N', there is also the natural squaring map

$$\mathcal{EE}_{N'}(F) \to \mathcal{EE}_{N'}(F), \quad \sum_{\Theta \in \mathcal{E}} m_\Theta \Theta \mapsto \sum_{\Theta \in \mathcal{E}} m_\Theta \Theta^2,$$

which is a bijection since p is odd. Combining these, we get a natural inclusion map

$$\iota_{2N} : \mathcal{EE}^{\mathrm{sd}}_{2N}(F) \hookrightarrow \mathcal{EE}_{2N+1}(F), \qquad \sum_{\Theta\in\mathcal{E}^{\mathrm{sd}}} m_\Theta \Theta \mapsto \sum_{\Theta\in\mathcal{E}^{\mathrm{sd}}} m_\Theta \Theta^2 + \Theta_0^F.$$

Given a maximal skew semisimple stratum $[\Lambda, n, 0, \beta]$ and a skew semisimple character θ of $H^1(\beta, \Lambda)$, which decomposes into a family of skew simple characters θ_i of $H^1(\beta^i, \Lambda^i)$ for $i \in \{0, \dots, l\}$, we define the self-dual endoparameter of θ to be

$$\sum_{i=0}^{l} \frac{\dim_F V^i}{\deg \Theta_i} \Theta_i,$$

where Θ_i is the endoclass of the unique self-dual simple character $\tilde\theta_i$ which restricts to θ_i. This is a self-dual endoparameter of degree $2N$.

We write $\mathrm{Cusp}(G)$ for the set of equivalence classes of cuspidal representations of G. From 2.6 Theorem and 2.5 Theorem, we derive the following result.

Theorem. *Suppose that F is of characteristic zero. Let π be a cuspidal representation of G and let θ be a skew semisimple character contained in π. Then the self-dual endoparameter of θ depends only on π. Moreover, the diagram*

$$\begin{tikzcd}
\mathrm{Cusp}(G) \arrow[r, "\text{transfer}"] \arrow[d, "\mathsf{e}_G"'] & \mathrm{Irr}(\mathrm{GL}_{2N+1}(F)) \arrow[d, "\mathsf{e}_{2N+1}"] \\
\mathcal{EE}^{\mathrm{sd}}_{2N}(F) \arrow[r, hook, "\iota_{2N}"'] & \mathcal{EE}_{2N+1}(F)
\end{tikzcd}$$

commutes, where $\mathsf{e}_G(\pi)$ denotes the endoparameter of any skew semisimple character contained in π.

We remark that the fact that the map e_G is well-defined is also proved, in much greater generality and without the assumption that F has characteristic zero, in [Kurinczuk et al. 2016]; the proof here is quite different and long predates the one in that paper. We also remark that we will see later (7.6 Theorem) that the map e_G is in fact surjective.

Proof. Let π be a cuspidal representation of G and let θ be a skew semisimple character contained in π, with all the notation from above. In particular, we have a family of skew simple characters θ_i for $0 \le i \le l$, and, for each i, the unique self-dual simple character $\tilde\theta_i$ restricting to θ_i and the self-dual endoclass Θ_i of $\tilde\theta_i$.

For $(\rho, m) \in \mathrm{Jord}(\pi)$, we write Θ_ρ for the endoclass of any simple character in ρ. Then 2.6 Corollary implies that $\Theta_\rho = \Theta_i^2$ for some $0 \le i \le l$; moreover, together with 2.5 Theorem it implies

$$\sum_{(\rho,m)\in\mathrm{Jord}(\pi)} \frac{m \deg \rho}{\deg \Theta_\rho} \Theta_\rho = \sum_{i=0}^{l} \frac{\dim_F V^i}{\deg \Theta_i} \Theta_i^2 + \Theta_0^F. \tag{2-2}$$

In particular, the right-hand side here is Θ_0^F plus the square of the endoparameter of θ; since the squaring map is a bijection, this endoparameter is therefore independent of the choice of θ in π since the left-hand side is.

Now, according to [Mœglin 2014, Theorem 3.2.1], the Jordan set exactly determines the endoscopic transfer of π to $\mathrm{GL}_{2N+1}(F)$; more precisely, the transfer of π is

$$\prod_{(\rho,m)\in\mathrm{Jord}(\pi)} \mathrm{St}(\rho, m),$$

where $\mathrm{St}(\rho, m)$ denotes the unique irreducible quotient of the normalized parabolically induced representation

$$\rho\nu^{(1-m)/2} \times \rho\nu^{(3-m)/2} \times \cdots \times \rho\nu^{(m-1)/2}$$

of $\mathrm{GL}_{m \deg \rho}(F)$. The endoparameter of the transfer of π is thus

$$\sum_{(\rho,m)\in\mathrm{Jord}(\pi)} \frac{m \deg \rho}{\deg \Theta_\rho} \Theta_\rho,$$

where Θ_ρ is the endoclass of (any simple character in) ρ. In particular, this lies in $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N+1}(F)$ and (2-2) now implies that the diagram commutes. $\square$

3. Types, covers and reducibility

In the following subsections we recall the main results about covers and their Hecke algebras, from [Bushnell and Kutzko 1998] in the general situation and from [Miyauchi and Stevens 2014] in the particular situation of interest to us: induction from a maximal parabolic subgroup of a symplectic group. One of the key features in [Miyauchi and Stevens 2014] is the presence of quadratic characters arising from the comparison of beta-extensions. Using the notion of p-primary beta-extension, together with results from [Blondel 2012], we describe these characters as signatures of permutations and recall the implications of the structure of the Hecke algebra (including its parameters) for the reducibility of parabolic induction from that paper.

3.1. We briefly recall the general notion of a *type* as defined in [Bushnell and Kutzko 1998]. Let for a moment G be the group of F-points of an arbitrary connected reductive group defined over F, let L be a Levi subgroup of G and let σ be a cuspidal representation of L. The pair (L, ρ) determines, through G-conjugacy and twisting by unramified characters of L, an inertial class $\mathfrak{s} = [L, \rho]_G$ in G. This class $\mathfrak{s}$ indexes the Bernstein block $\mathcal{R}^{\mathfrak{s}}(G)$ (in the category $\mathcal{R}(G)$ of smooth representations of G) which is the direct factor of $\mathcal{R}(G)$ consisting of representations all of whose irreducible subquotients are subquotients of a representation parabolically induced from an element of $\mathfrak{s}$.

Let (J, λ) be a pair made of an open compact subgroup J of G and an irreducible smooth representation λ of J, acting on the finite-dimensional space V_λ. The Hecke algebra of the pair (J, λ) is the intertwining algebra of the representation $\text{c-Ind}_J^G \lambda$, traditionally viewed as

$$\mathcal{H}(G, \lambda) = \{ f : G \to \mathrm{End}(V_\lambda) \mid f \text{ compactly supported and } \forall g \in G,\ \forall j, k \in J,\ f(jgk) = \lambda(j) f(g) \lambda(k) \}.$$

The pair (J, λ) is an $\mathfrak{s}$-type if the irreducible objects of $\mathcal{R}^{\mathfrak{s}}(G)$ are exactly the irreducible representations of G that contain λ upon restriction to J. In this case there is an equivalence of categories

$$\mathcal{M}_\lambda : \mathcal{R}^{\mathfrak{s}}(G) \to \text{Mod-}\mathcal{H}(G, \lambda), \quad \mathcal{M}_\lambda(\pi) = \mathrm{Hom}_J(\lambda, \pi).$$

3.2. There is a counterpart of parabolic induction for types: the notion of G-cover, also defined in [Bushnell and Kutzko 1998]. Let M be a Levi subgroup of G, let J_M be a compact open subgroup of M and let λ_M be a smooth irreducible representation of J_M. A *G-cover* of the pair (J_M, λ_M) is an analogous pair (J, λ) in G satisfying the following conditions, for any parabolic subgroup P of G of Levi M, where we write N for the unipotent radical of P, and P^- for the parabolic subgroup opposite to P with respect to M, with unipotent radical N^-:

(i) J has an Iwahori decomposition with respect to $(M; P)$; i.e.,

$$J = (J \cap N^-)(J \cap M)(J \cap N), \text{ and } J \cap M = J_M.$$

(ii) λ restricts to λ_M on J_M and to a multiple of the trivial representation on $J \cap N^-$ and $J \cap N$.

(iii) The Hecke algebra $\mathcal{H}(G, \lambda)$ contains an invertible element supported on the double coset of a strongly positive element of the centre of M [Bushnell and Kutzko 1998, §7].

If the pair (J_M, λ_M) is an $\mathfrak{s}_M$-type in M for an inertial class $\mathfrak{s}_M = [L, \sigma]_M$ (so that L is a Levi subgroup of M) and if (J, λ) is a G-cover of (J_M, λ_M), then the pair (J, λ) is an $\mathfrak{s}_G$-type in G for the inertial class $s_G = [L, \sigma]_G$ [Bushnell and Kutzko 1998, §8]. Furthermore, the third condition above provides us with an injective morphism of algebras $t : \mathcal{H}(M, \lambda_M) \hookrightarrow \mathcal{H}(G, \lambda)$ that induces on modules a morphism t_* yielding a commutative diagram:

$$\begin{tikzcd}
\mathcal{R}^{\mathfrak{s}_G}(G) \arrow[r, "\mathcal{M}_\lambda"] & \text{Mod-}\mathcal{H}(G, \lambda) \\
\mathcal{R}^{\mathfrak{s}_M}(M) \arrow[u, "\mathrm{Ind}_P^G"] \arrow[r, "\mathcal{M}_{\lambda_M}"] & \text{Mod-}\mathcal{H}(M, \lambda_M) \arrow[u, "t_*"']
\end{tikzcd}$$

The reducibility of parabolically induced representations from P to G, on the left side, can thus be studied in terms of Hecke algebra modules, on the right side.

3.3. This is the tool we use in this paper, where the types will be cuspidal types as in Sections 1.6 and 1.7, simple types and semisimple types. As for the relevant Levi and parabolic subgroups, they will come in most cases as follows — and now we come back to the symplectic group $G = \mathrm{Sp}_F(V)$ and the setting of Section 1.1. Thus we have a skew semisimple stratum $[\Lambda, -, 0, \beta]$ with associated orthogonal decomposition $V = \perp_{i=0}^{l} V^i$, as well as all the other notation from Section 1.

Let $V = \bigoplus_{j=-m}^{m} W^j$ be a *self-dual* decomposition of V (i.e., for which the orthogonal space of W^j is $\bigoplus_{k \neq -j} W^k$) such that:

(a) $W^j = \bigoplus_{i=0}^{l} W^j \cap V^i$ and $W^j \cap V^i$ is an E_i-subspace of V^i.

(b) $\Lambda(t) = \bigoplus_{j=-m}^{m} \Lambda(t) \cap W^{(j)}$ for all $t \in \mathbb{Z}$.

(c) For any $r \in \mathbb{Z}$ and i with $0 \le i \le l$, there is at most one j, with $-m \le j \le m$, such that $\Lambda(r) \cap V^i \cap W^{(j)} \supsetneq \Lambda(r+1) \cap V^i \cap W^{(j)}$.

(d) For $j \neq 0$ there exists $0 \leq i \leq l$ such that $W^j \subset V^i$, and $\widetilde{P}((\Lambda \cap W^j)_{\mathfrak{o}_E^i})$ is a maximal parahoric subgroup of $\mathrm{GL}_{E_i}(W^j)$.

(e) $P^\circ((\Lambda \cap W^0)_{\mathfrak{o}_E})$ is a maximal parahoric subgroup of $G \cap \prod_{i=0}^{l} \mathrm{GL}_{E_i}(W^0 \cap V^i)$, which is a group with compact centre.

Such a decomposition is called *exactly subordinate* to the stratum $[\Lambda, -, 0, \beta]$ (compare to [Stevens 2008, Definition 6.5]).

Let then $V = \bigoplus_{j=-m}^{m} W^j$ be a self-dual decomposition of V exactly subordinate to the stratum $[\Lambda, -, 0, \beta]$, let M be the Levi subgroup of G stabilizing this decomposition and let P be a parabolic subgroup of G with Levi component M. Then the pairs $(H^1(\beta, \Lambda), \theta)$, $(J^1(\beta, \Lambda), \eta)$ and $(J(\beta, \Lambda), \kappa)$ all satisfy conditions (i) and (ii) of Section 3.2 above. In fact, for the first two pairs we need only conditions (a)–(b) and a self-dual decomposition satisfying these will be called *subordinate* to the stratum; for the final pair we need only (a)–(c).

3.4. In the next few subsections, we subsume the results of [Stevens 2008], in a form easier to refer to taken from [Miyauchi and Stevens 2014] and in the case that we will focus on, that is, parabolic induction of self-dual cuspidal representations of a maximal Levi subgroup in a symplectic group. We thus continue with the notation of Section 1 and fix a cuspidal representation π of $G = \mathrm{Sp}_F(V)$. We also fix a finite-dimensional vector space W over F and a self-dual cuspidal representation ρ of $\mathrm{GL}_F(W)$. We consider the symplectic space $X = V \perp (W \oplus W^*)$ over F, with form

$$h_X(v_1 + w_1 + w_1^*, v_2 + w_2 + w_2^*) = h(v_1, v_2) + \langle w_1, w_2^* \rangle - \langle w_2, w_1^* \rangle,$$

where h is the symplectic form on V and $\langle \cdot, \cdot \rangle$ is the pairing $W \times W^* \to F$. We put M equal to $\mathrm{GL}_F(W) \times G$, a maximal Levi subgroup of $\mathrm{Sp}_F(X)$. According to [Dat 2009, Proposition 8.4] and [Miyauchi and Stevens 2014, §4.1], one can find a type $(\tilde{J}_W \times J_V, \tilde{\lambda}_W \otimes \lambda_V)$ in M for the cuspidal representation $\rho \otimes \pi$ of M and a G-cover of this M-type as follows.

3.5. There exist a skew semisimple stratum $[\Lambda, -, 0, \beta]$ in $\mathrm{End}_F(X)$ and a skew semisimple character θ of $H^1(\beta, \Lambda)$ with the following properties:

- The decomposition $X = V \perp (W \oplus W^*)$ is exactly subordinate to the stratum $[\Lambda, -, 0, \beta]$. In particular, letting

$$\Lambda \cap V = \Lambda_V, \quad \beta_{|V} = \beta_V, \quad \Lambda \cap W = \Lambda_W \quad \text{and} \quad \beta_{|W} = \beta_W,$$

the stratum $[\Lambda_V, -, 0, \beta_V]$ in $\mathrm{End}_F(V)$ is skew semisimple maximal and the stratum $[\Lambda_W, -, 0, \beta_W]$ in $\mathrm{End}_F(W)$ is simple maximal. Moreover, the self-duality of ρ is reflected in the fact that the restriction of β to $W \oplus W^*$ generates a field (equivalently, the restricted stratum $[\Lambda \cap (W \oplus W^*), -, 0, \beta_{|W \oplus W^*}]$ is skew simple). We also have

$$H^1(\beta, \Lambda) \cap M \simeq \widetilde{H}^1(\beta_W, \Lambda_W) \times H^1(\beta_V, \Lambda_V),$$

where the isomorphism is given by restriction, and similarly for $J^1(\beta, \Lambda)$ and for $J(\beta, \Lambda)$. We will abbreviate $H^1(\beta, \Lambda) = H^1_\Lambda$, $H^1(\beta_V, \Lambda_V) = H^1_V$ and $\widetilde{H}^1(\beta_W, \Lambda_W) = \widetilde{H}^1_W$, and similarly for J^1 and J.

- Let $\tilde{\vartheta}_W$ be the restriction of θ to $\widetilde{H}^1_W$; this is a self-dual simple character. There are the p-primary beta-extension $\tilde{\underline{\kappa}}_W$ of $\tilde{\vartheta}_W$ and a self-dual cuspidal representation $\tilde{\tau}_W$ of $\tilde{J}_W/\tilde{J}^1_W$ such that ρ is induced by an extension of $\tilde{\lambda}_W = \tilde{\underline{\kappa}}_W \otimes \tilde{\tau}_W$ to the normalizer of $\tilde{J}_W$.
- Let θ_V be the restriction of θ to H^1_V; this is a skew semisimple character. There are the p-primary beta-extension $\underline{\kappa}_V$ of θ_V and a cuspidal representation τ_V of J_V/J^1_V such that π is induced by $\lambda_V = \underline{\kappa}_V \otimes \tau_V$.

3.6. Let P be the parabolic subgroup of $\mathrm{Sp}_F(X)$ which is the stabilizer of the subspace W (so stabilizes the flag $W \subset W \perp V \subset X$), let U be the unipotent radical of P and let P^- be the parabolic subgroup opposite to P with respect to M (the stabilizer of W^*). Also set $J_P = H^1_\Lambda(J_\Lambda \cap P)$ and $J^1_P = H^1_\Lambda(J^1_\Lambda \cap P)$.

For any extension κ of η to J_Λ we denote by κ_P the natural representation of J_P in the space of $(J_\Lambda \cap U)$-fixed vectors under κ. In particular, there is a beta-extension κ_Λ of θ such that $\kappa_{\Lambda,P\,|J\cap M} = \tilde{\underline{\kappa}}_W \otimes \underline{\kappa}_V$. We can view $\tau = \tilde{\tau}_W \otimes \tau_V$ as a cuspidal representation of $J_P/J^1_P \simeq \tilde{J}_W/\tilde{J}^1_W \times J_V/J^1_V$. Then, letting $\lambda_P = \kappa_{\Lambda,P} \otimes \tau$, we have:

Theorem [Miyauchi and Stevens 2014, §4.1]. *(J_P, λ_P) is an $\mathrm{Sp}_F(X)$-cover of $(\tilde{J}_W \times J_V, \tilde{\lambda}_W \otimes \lambda_V)$.*

3.7. Furthermore precise information about the Hecke algebra of this cover is given in [loc. cit.]:

Theorem [Miyauchi and Stevens 2014, Theorem B]. *The Hecke algebra $\mathscr{H}(\mathrm{Sp}_F(X), \lambda_P)$ is a Hecke algebra on a dihedral group: it is generated by T_0 and T_1, each invertible and supported on a single double coset, with relations*

$$(T_i - q^{r_i})\,(T_i + 1) = 0, \quad i = 0, 1,\ r_0, r_1 \in \mathbb{Z}.$$

3.8. In fact, the parameters come from rank-2 Hecke algebras of finite reductive groups as follows [Stevens 2008, (7.3) and §7.2.2]. There are two self-dual $\mathfrak{o}_E$-lattice sequences $\mathfrak{M}_0$ and $\mathfrak{M}_1$ in X such that $[\mathfrak{M}_t, -, 0, \beta]$, for $t = 0, 1$, are semisimple strata and

- the hereditary orders $\mathfrak{b}_0(\mathfrak{M}_0)$ and $\mathfrak{b}_0(\mathfrak{M}_1)$ are maximal self-dual $\mathfrak{o}_E$-orders containing $\mathfrak{b}_0(\Lambda)$;
- the decomposition $X = V \perp (W \oplus W^*)$ is subordinate to the strata $[\mathfrak{M}_t, -, 0, \beta]$ for $t = 0, 1$;
- we have $P(\Lambda_{\mathfrak{o}_E}) = (P(\mathfrak{M}_{1,\mathfrak{o}_E}) \cap P^-)P^1(\mathfrak{M}_{1,\mathfrak{o}_E}) = (P(\mathfrak{M}_{0,\mathfrak{o}_E}) \cap P)P^1(\mathfrak{M}_{0,\mathfrak{o}_E})$.

The representation $\tau = \tilde{\tau}_W \otimes \tau_V$ is a cuspidal representation of the Levi subgroup $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) = P(\Lambda_{\mathfrak{o}_E})/P^1(\Lambda_{\mathfrak{o}_E})$ of $\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}) = P(\mathfrak{M}_{t,\mathfrak{o}_E})/P^1(\mathfrak{M}_{t,\mathfrak{o}_E})$, for $t = 0, 1$, that can be inflated to the parabolic subgroup $P(\Lambda_{\mathfrak{o}_E})/P^1(\mathfrak{M}_{t,\mathfrak{o}_E})$, then induced to the full group $\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E})$. A specific use of the notion of beta-extension relative to $\mathfrak{M}_{t,\mathfrak{o}_E}$ leads to self-dual characters χ_t of $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$, for $t = 0, 1$, giving rise to injective homomorphisms of algebras

$$\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \chi_t \otimes \tau) \hookrightarrow \mathscr{H}(G, \lambda_P) \quad (t = 0, 1). \tag{3-1}$$

We will elaborate on this in Section 3.12 below.

3.9. In order to make use of this, we need some control on the characters χ_t and it is here that we really need to use the notion of p-primary beta-extension. We continue with the notation of the previous subsections but, for the moment, drop the subscript t on $\mathfrak{M}_t$. We will assume that $P(\Lambda_{\mathfrak{o}_E}) = (P(\mathfrak{M}_{\mathfrak{o}_E}) \cap P)P^1(\mathfrak{M}_{\mathfrak{o}_E})$ so that, in the notation above, we are doing the case $t = 0$; the case $t = 1$ is obtained by exchanging the parabolic P with its opposite P^-. Denote by $\theta_{\mathfrak{M}}$ the transfer of $\theta_\Lambda = \theta$ to $H^1_{\mathfrak{M}} = H^1(\beta, \mathfrak{M})$, and denote by η_Λ, $\eta_{\mathfrak{M}}$ the unique irreducible representations of J^1_Λ, $J^1_{\mathfrak{M}}$ which contain θ_Λ, $\theta_{\mathfrak{M}}$ respectively. Similarly, we have the representations $\tilde{\eta}_W$, η_V of $\tilde{J}^1_W$, J^1_V which contain $\tilde{\vartheta}_W$, θ_V respectively.

For a moment, let (J, J^1, η) be either $(J_{\mathfrak{M}}, J^1_{\mathfrak{M}}, \eta_{\mathfrak{M}})$ or $(J_\Lambda, J^1_\Lambda, \eta_\Lambda)$, and let κ be any extension of η to J. We define $\boldsymbol{r}_P(\kappa)$, the *Jacquet restriction* of κ, as the natural representation of $J \cap M$ on the space of $J^1 \cap U$-invariants of κ [Blondel 2012, Corollaire 1.12, Lemme 1.18]; that is, $\boldsymbol{r}_P(\kappa)$ is the restriction to $J \cap M$ of κ_P, in the notation of Section 3.6.

3.10. In order to compute the character χ from (3-1), we need to compare the following two representations of J_Λ:

- the beta-extension $\underline{\kappa}_{\Lambda,\mathfrak{M}}$ of η_Λ to J_Λ which is compatible with the p-primary beta-extension $\underline{\kappa}_{\mathfrak{M}}$ of $\eta_{\mathfrak{M}}$ to $J_{\mathfrak{M}}$ (in the sense of [Stevens 2008, Definition 4.5]);
- the extension $\kappa_\Lambda = \kappa^P_\Lambda$ of η_Λ to J_Λ characterized by the property
$$\boldsymbol{r}_P(\kappa^P_\Lambda) \simeq \tilde{\underline{\kappa}}_W \otimes \underline{\kappa}_V,$$
where $\tilde{\underline{\kappa}}_W$ and $\underline{\kappa}_V$ are the p-primary beta-extensions of $\tilde{\eta}_W$ and η_V respectively, as above. (See [Blondel 2012, Lemme 1.16].)

We apply Jacquet restriction to $\underline{\kappa}_{\Lambda,\mathfrak{M}}$. The groups $J_{\mathfrak{M}} \cap M$ and $J_\Lambda \cap M$ are both equal to $\tilde{J}_W \times J_V$ and the representations $\boldsymbol{r}_P(\underline{\kappa}_{\mathfrak{M}})$ and $\boldsymbol{r}_P(\underline{\kappa}_{\Lambda,\mathfrak{M}})$ both extend $\tilde{\eta}_W \otimes \eta_V$. From [Blondel 2012, Proposition 1.20], the beta-extension $\underline{\kappa}_{\Lambda,\mathfrak{M}}$ is characterized by

$$\boldsymbol{r}_P(\underline{\kappa}_{\Lambda,\mathfrak{M}}) = \boldsymbol{r}_P(\underline{\kappa}_{\mathfrak{M}}). \tag{3-2}$$

Proposition. *For $m \in J_{\mathfrak{M}} \cap M$, define $\epsilon_{\mathfrak{M}}(m)$ as the signature of the permutation*
$$\mathrm{Ad}\, m : u \mapsto m^{-1}um, \quad u \in J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-.$$
The p-primary beta-extension $\underline{\kappa}_{\mathfrak{M}}$ of $\eta_{\mathfrak{M}}$ to $J_{\mathfrak{M}}$ satisfies
$$\boldsymbol{r}_P(\underline{\kappa}_{\mathfrak{M}}) \simeq \epsilon_{\mathfrak{M}}(\tilde{\underline{\kappa}}_W \otimes \underline{\kappa}_V). \tag{3-3}$$

This proposition and (3-2) immediately imply:

Corollary. *The extensions $\underline{\kappa}_{\Lambda,\mathfrak{M}}$ and κ^P_Λ of η_Λ to J_Λ are related by*
$$\boldsymbol{r}_P(\kappa^P_\Lambda) = \epsilon_{\mathfrak{M}} \boldsymbol{r}_P(\underline{\kappa}_{\Lambda,\mathfrak{M}}).$$

Proof of Proposition. Let ϕ be an arbitrary extension of $\eta_{\mathfrak{M}}$ to $J_{\mathfrak{M}}$. By [Blondel 2012, Lemme 1.10] the restriction of ϕ to $(J_{\mathfrak{M}} \cap P)J^1_{\mathfrak{M}}$ is induced from the natural representation ϕ_P of $(J_{\mathfrak{M}} \cap P)H^1_{\mathfrak{M}}$ on the space of $J^1_{\mathfrak{M}} \cap U$-invariants of ϕ. Hence we can realize $\phi_{|J_{\mathfrak{M}} \cap M}$ as the action by right translation on functions taking values in the space of ϕ_P. We also have the representation $\eta_{\mathfrak{M},P}$ of $(J^1_{\mathfrak{M}} \cap P)H^1_{\mathfrak{M}}$ on the space of $J^1_{\mathfrak{M}} \cap U$-invariants of $\eta_{\mathfrak{M}}$.

Let $\widetilde{\mathcal{S}}$ be the space of $\tilde{\eta}_W$ and let $\mathcal{S}$ be the space of η_V, so that $\widetilde{\mathcal{S}} \otimes \mathcal{S}$ is the space of $\eta_{\mathfrak{M},P}$. The representation ϕ_P itself extends $\eta_{\mathfrak{M},P}$, so our representation $\phi_{|J_{\mathfrak{M}} \cap M}$ acts by right translation on the space of functions

$$f : (J_{\mathfrak{M}} \cap P)J^1_{\mathfrak{M}} \to \widetilde{\mathcal{S}} \otimes \mathcal{S}$$

satisfying, for all $x \in (J_{\mathfrak{M}} \cap P)H^1_{\mathfrak{M}}$ and all $g \in (J_{\mathfrak{M}} \cap P)J^1_{\mathfrak{M}}$,

$$f(xg) = \phi_P(x) f(g).$$

Using Iwahori decompositions as in [Blondel 2012, §1.3] we identify this space with the space $\mathcal{T}$ of functions on $J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-$ with values in $\widetilde{\mathcal{S}} \otimes \mathcal{S}$. The action of $m \in J_{\mathfrak{M}} \cap M$ on $f \in \mathcal{T}$ is now given by

$$\phi(m) f(u) = f(um) = f(m.m^{-1}um) = \boldsymbol{r}_P(\phi)(m) f(m^{-1}um)$$

for $u \in J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-$.

Let $\mathcal{T}_0$ be the space of complex functions on $J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-$ and $\mathcal{E}$ the permutation representation of $J_{\mathfrak{M}} \cap M$ on $\mathcal{T}_0$,

$$\mathcal{E}(m) f(u) = f(m^{-1}um) \quad \text{for } f \in \mathcal{T}_0,\ m \in J_{\mathfrak{M}} \cap M,\ u \in J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-.$$

We can further identify $\mathcal{T}$ with $\mathcal{T}_0 \otimes (\widetilde{\mathcal{S}} \otimes \mathcal{S})$ to obtain $\phi_{|J_{\mathfrak{M}} \cap M} \simeq \mathcal{E} \otimes \boldsymbol{r}_P(\phi)$.

All of this applies to $\underline{\kappa}_{\mathfrak{M}}$, so

$$\underline{\kappa}_{\mathfrak{M}|J_{\mathfrak{M}} \cap M} \simeq \mathcal{E} \otimes \boldsymbol{r}_P(\underline{\kappa}_{\mathfrak{M}}).$$

The determinant of this representation has order a power of p, a property that is unchanged by taking p^k-th powers. Recall that the determinant of some $x \otimes y$ acting on $X \otimes Y$ is $(\det x)^{\dim Y} (\det y)^{\dim X}$. The two spaces here, $\mathcal{T}_0$ and $\widetilde{\mathcal{S}} \otimes \mathcal{S}$, have dimension a power of p, which is odd, and the determinant of $\mathcal{E}$ acting on $\mathcal{T}_0$ is $\epsilon_{\mathfrak{M}}$.

We now write $\boldsymbol{r}_P(\underline{\kappa}_{\mathfrak{M}}) = \tilde{\kappa}_W \otimes \kappa_V$, where $\tilde{\kappa}_W$ is a beta-extension of $\tilde{\eta}_W$ and κ_V is a beta-extension of η_V; see (3-2) and [Stevens 2008, Proposition 6.3]. It is enough to prove

$$\epsilon_{\mathfrak{M}}\ \det(\tilde{\kappa}_W \otimes \kappa_V) \text{ has order a power of } p.$$

Writing $\epsilon_{\mathfrak{M}}$ for the restrictions of $\epsilon_{\mathfrak{M}}$ to $\tilde{J}_W$ and to J_V, this condition transforms into

$$\det(\epsilon_{\mathfrak{M}} \otimes \tilde{\kappa}_W) \text{ and } \det(\epsilon_{\mathfrak{M}} \otimes \kappa_V) \text{ have order a power of } p.$$

The character $\epsilon_{\mathfrak{M}}$ is trivial on pro-p-subgroups so $\epsilon_{\mathfrak{M}} \otimes \tilde{\kappa}_W$ and $\epsilon_{\mathfrak{M}} \otimes \kappa_V$ are beta-extensions of $\tilde{\eta}_W$ and η_V respectively [Stevens 2008, Theorem 4.1]. This last condition actually means that they are the p-primary beta-extensions of $\tilde{\eta}_W$ and η_V respectively, and (3-3) follows. □

3.11. Before returning to the implications on reducibility, we examine the character $\epsilon_{\mathfrak{M}}$ a little further. We begin with a general lemma.

Lemma. *Let Z be a finite-dimensional vector space over a finite field $\mathbb{F}_q$ with odd cardinality q and let $g \in \mathrm{GL}_{\mathbb{F}_q}(Z)$. The signature of the permutation g of Z is equal to $(\det_{\mathbb{F}_q} g)^{(q-1)/2}$.*

Proof. As a character of $\mathrm{GL}_{\mathbb{F}_q}(Z)$, the signature is trivial on the derived subgroup, which is $\mathrm{SL}_{\mathbb{F}_q}(Z)$, as $q > 2$, and hence factors through a character χ of the determinant over $\mathbb{F}_q$. We know χ^2 is trivial and it remains to show that χ is not identically trivial on $\mathrm{GL}_{\mathbb{F}_q}(Z)$.

Consider the permutation of $\mathbb{F}_q$ given by multiplication by an element ζ of $\mathbb{F}_q^\times$ of order 2^t, with $(q-1)/2^t$ an odd integer. This permutation fixes 0 and has $(q-1)/2^t$ cycles of length 2^t in $\mathbb{F}_q^\times$, and so has odd signature. Then the element $g = \mathrm{diag}(\zeta, 1, \dots, 1)$ has odd signature so χ is nontrivial. □

Proposition. *For $m \in J_{\mathfrak{M}} \cap M$, the permutation $\mathrm{Ad}\, m$ of $J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-$ is an $\mathbb{F}_p$-linear transformation of this $\mathbb{F}_p$-vector space and*

$$\epsilon_{\mathfrak{M}}(m) = [\det_{\mathbb{F}_p} \mathrm{Ad}\, m]^{(p-1)/2}.$$

Moreover, the permutation $u \mapsto m^{-1}um$ of the space $J^1_{\mathfrak{M}} \cap U / H^1_{\mathfrak{M}} \cap U$ also has signature $\epsilon_{\mathfrak{M}}(m)$.

Proof. The first part follows from the previous lemma. Since the decomposition $X = V \perp (W \oplus W^*)$ is subordinate to $[\mathfrak{M}, -, 0, \beta]$, the pairing

$$\langle x, y \rangle = \theta_{\mathfrak{M}}([x, y]) \quad \text{for } x \in J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-, \ \ y \in J^1_{\mathfrak{M}} \cap U / H^1_{\mathfrak{M}} \cap U,$$

identifies each of those $\mathbb{F}_p$-vector spaces to the dual of the other [Stevens 2008, Lemma 5.6] in such a way that, for $m \in J_{\mathfrak{M}} \cap M$, the transpose of the map $x \mapsto m^{-1}xm$, for $x \in J^1_{\mathfrak{M}} \cap U^- / H^1_{\mathfrak{M}} \cap U^-$, is $y \mapsto mym^{-1}$, for $y \in J^1_{\mathfrak{M}} \cap U / H^1_{\mathfrak{M}} \cap U$. The result follows. □

3.12. We return to the notation of Sections 3.4–3.8 and now put together the Hecke algebra homomorphisms (3-1) with 3.10 Proposition. Let $t = 0$ or 1. We recall from [Stevens 2008, (7.3)] (rephrased in the present framework in [Blondel 2012, Proposition 3.6]) that if $\kappa = \text{c-Ind}_{J_P}^{J_\Lambda} \kappa_P$ is a beta-extension of $\eta_\Lambda = \text{c-Ind}_{J_P^1}^{J_\Lambda^1} \eta_P$ *relative to* $\mathfrak{M}_t$, then there is an injective morphism of algebras

$$\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \tilde{\tau}_W \otimes \tau_V) \hookrightarrow \mathscr{H}(\mathrm{Sp}_F(X), \kappa_P \otimes (\tilde{\tau}_W \otimes \tau_V))$$

that preserves support. We want to express this with the fixed representation $\lambda_P = \kappa_{\Lambda,P}$ on the right, where $\kappa_{\Lambda,P|J\cap M} = \tilde{\underline{\kappa}}_W \otimes \underline{\kappa}_V$, as in Section 3.6. We thus plug in 3.10 Proposition above and get:

Theorem. *Let $t = 0$ or 1. There is an injective morphism of algebras*

$$j_t : \mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_V)) \hookrightarrow \mathscr{H}(\mathrm{Sp}_F(X), \lambda_P)$$

that preserves support; i.e., $\mathrm{Supp}(j_t(\phi)) = J_P \,\mathrm{Supp}(\phi) J_P$ *for all* $\phi \in \mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_V))$.

3.13. We now focus on the finite-dimensional algebra $\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_V))$, a Hecke algebra on the finite reductive group $\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E})$ relative to a cuspidal representation of the parabolic subgroup $P(\Lambda_{\mathfrak{o}_E})/P^1(\mathfrak{M}_{t,\mathfrak{o}_E})$.

Let $X = \perp_{j=0}^{l} X^j$ be the splitting associated to the skew semisimple stratum $[\Lambda, -, 0, \beta]$. Since the stratum $[\Lambda_W, -, 0, \beta_W]$ is simple, there is a unique index i such that $W \subseteq X^i$, and then $W^* \subseteq X^i$ also. *This index i will be fixed until the end of the section.*

Writing $V^j = V \cap X^j$, the skew semisimple stratum $[\Lambda_V, -, 0, \beta_V]$ then has splitting consisting of the nonzero spaces in $V = \perp_{j=0}^{l} V^j$; the only spaces which may be zero here are V^0 (since we have the convention that $\beta^0 = 0$) and V^i (which is zero if and only if $X^i = W \oplus W^*$).

The ambient finite group $\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E})$ is a product over j, for $0 \le j \le l$, of analogous groups relative to X^j, but in all of them except X^i the parabolic subgroup considered is the full group:

$$P(\mathfrak{M}_{t,\mathfrak{o}_E})/P^1(\mathfrak{M}_{t,\mathfrak{o}_E}) \simeq P(\mathfrak{M}^i_{t,\mathfrak{o}^i_E})/P^1(\mathfrak{M}^i_{t,\mathfrak{o}^i_E}) \times \prod_{j\neq i} P(\Lambda^j_{\mathfrak{o}^j_E})/P^1(\Lambda^j_{\mathfrak{o}^j_E}),$$

$$P(\Lambda_{\mathfrak{o}_E})/P^1(\mathfrak{M}_{t,\mathfrak{o}_E}) \simeq P(\Lambda^i_{\mathfrak{o}^i_E})/P^1(\mathfrak{M}^i_{t,\mathfrak{o}^i_E}) \times \prod_{j\neq i} P(\Lambda^j_{\mathfrak{o}^j_E})/P^1(\Lambda^j_{\mathfrak{o}^j_E}).$$

The representation $\tilde{\tau}_W \otimes \tau_V$ decomposes accordingly using $\tau_V = \bigotimes_{j=0}^{l} \tau_j$ and we finally get an isomorphism of algebras:

$$\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_V)) \simeq \mathscr{H}(\mathcal{G}(\mathfrak{M}^i_{t,\mathfrak{o}^i_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_i)), \tag{3-4}$$

where $\tilde{\tau}_W \otimes \tau_i$ is a cuspidal representation of $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}^i_E}) \times \mathcal{G}(\Lambda^i_{V,\mathfrak{o}^i_E})$, identified with a maximal Levi subgroup of each finite reductive group $\mathcal{G}(\mathfrak{M}^i_{t,\mathfrak{o}^i_E})$ for $t = 0, 1$.

It follows from [Lusztig 1984], as recalled in Section 5, that this algebra is two-dimensional, because $\tilde{\tau}_W$ and $\epsilon_{\mathfrak{M}_t}$ are self-dual. It has basis given by the identity element and an element $\mathcal{T}_t$ supported on the double coset of a certain Weyl group element, called s_i if $t = 0$, or s_i^{ϖ} if $t = 1$, in [Stevens 2008, §7.2.2]; this only defines $\mathcal{T}_t$ up to a nonzero scalar, which will not matter to us at first. Lusztig gives an algorithm permitting the actual computation of the quadratic relation satisfied by $\mathcal{T}_t$. This relation always has the following shape, for some nonzero complex number ω_t:

$$(\mathcal{T}_t - q^{r_t}\,\omega_t)(\mathcal{T}_t + \omega_t) = 0, \quad \text{where } r_t = r_t(\epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_i)) \ge 0. \tag{3-5}$$

We emphasize the dependency in the inducing cuspidal representation $\epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau_i)$.

3.14. Finally, we can restate [Blondel 2012, Proposition 3.12], describing the real parts of the reducibility points we wish to compute, in our notation. Recall that, for ρ a cuspidal representation of $\mathrm{GL}_F(W)$ as above, we write $t(\rho)$ for the number of unramified characters χ of $\mathrm{GL}_F(W)$ such that $\rho \otimes \chi \simeq \rho$. Recall also that, if ρ is self-dual, then there are precisely two representations ρ, ρ' in the inertial class of ρ which are self-dual.

Let π be a cuspidal representation of G. Recall that there is a real number $s_\pi(\rho) \ge 0$ such that, for *real* s, the normalized induced representation $\nu^s\rho \times \pi$ of $\mathrm{Sp}_F(X)$ is reducible if and only if $s = \pm s_\pi(\rho)$, and similarly we have $s_\pi(\rho')$. Then, for *complex* s, if $\nu^s\rho\times\pi$ is reducible then the real part of s must be $\pm s_\pi(\rho)$ or $\pm s_\pi(\rho')$; we say that these are the real parts of the reducibility points of $\nu^s\rho\times\pi$.

Proposition [Blondel 2012, Proposition 3.12]. *Let π be a cuspidal representation of G, let ρ be an irreducible self-dual cuspidal representation of $\mathrm{GL}_F(W)$, and take all the notation of the previous subsections. Then the real parts of the reducibility points of the normalized induced representation $\nu^s\rho\times\pi$ are the elements of the set*

$$\left\{\pm\frac{r_0+r_1}{2t(\rho)},\pm\frac{r_0-r_1}{2t(\rho)}\right\},\quad \textit{where } r_0=r_0(\epsilon_{\mathfrak{M}_0}(\tilde\tau_W\otimes\tau_i)),\ r_1=r_1(\epsilon_{\mathfrak{M}_1}(\tilde\tau_W\otimes\tau_i)). \tag{3-6}$$

Note that, by [Bushnell and Kutzko 1993, Lemma 6.2.5], the unramified twist number $t(\rho)$ can also be computed from the formula $t(\rho)=\dim_F W/e(F[\beta_W]/F)$.

3.15. We can also apply the discussion of the previous subsections in the space $X^i=W\oplus V^i\oplus W^*$. From the splitting of our strata, we have the lattice sequence $\Lambda_V^i=\Lambda\cap V^i$ and the simple stratum $[\Lambda_V^i,-,0,\beta_V^i]$ in V^i. We write $J_i=J(\beta_V^i,\Lambda_V^i)$, and similarly for J_i^1 and H_i^1, and let $\underline{\kappa}_i$ be the p-primary beta-extension of the simple character $\theta_{|H_i^1}$. Then τ_i is a representation of the reductive quotient J_i/J_i^1 and, putting $G_i=\mathrm{Sp}_F(V^i)$ we can define the cuspidal representation $\pi_i=\text{c-Ind}_{J_{\Lambda^i}}^{G_i}\underline{\kappa}_i\otimes\tau_i$ of G_i. (Note that, if $V^i=\{0\}$, then π_i is the trivial representation of the trivial group.)

Applying the discussion above to the representation π_i and the space X^i, we find that the real parts of the reducibility points of the normalized induced representation $\nu^s\rho\times\pi_i$ of $\mathrm{Sp}_F(X^i)$ are the elements of the set

$$\left\{\pm\frac{r_0'+r_1'}{2t(\rho)},\pm\frac{r_0'-r_1'}{2t(\rho)}\right\},\quad \text{where } r_0'=r_0(\epsilon_{\mathfrak{M}_0^i}(\tilde\tau_W\otimes\tau_i)),\ r_1'=r_1(\epsilon_{\mathfrak{M}_1^i}(\tilde\tau_W\otimes\tau_i)). \tag{3-7}$$

The comparison between (3-6) and (3-7) will be crucial.

3.16. We end this section with the simplest example of the computation of the parameters r_0,r_1 in (3-6), for positive-depth representations. Continuing in the notation above, we assume that $i>0$ and that β^i is *maximal* in the following sense: we have $[F[\beta^i]:F]=\dim_F V^i$, so that (the image of) $F[\beta^i]$ is a maximal extension of F in $\mathrm{End}_F(V^i)$. In particular, this implies that $V^i\neq\{0\}$. We assume moreover that $\dim_F W=\dim_F V^i$, the smallest example of the situation above. (It will turn out that this is in fact the only situation of interest, in this context.)

Let E_0^i be the fixed field of the adjoint involution acting on $E^i=F[\beta^i]$. The centralizer of β^i in $\mathrm{Sp}_F(X^i)$ is thus isomorphic to the unitary group $\mathrm{U}(2,1)(E^i/E_0^i)$. In the latter group, there are two conjugacy classes of maximal compact subgroups, the reductive quotients of which are, for some a and b with $\{a,b\}=\{0,1\}$ depending on the initial lattice sequence Λ^i,

- $\mathcal{G}(\mathfrak{M}_{a,\mathfrak{o}_E})\simeq U(2,1)(k_{E^i}/k_{E_0^i})$ and $\mathcal{G}(\mathfrak{M}_{b,\mathfrak{o}_E})\simeq U(1,1)\times U(1)(k_{E^i}/k_{E_0^i})$ if E^i/E_0^i is unramified;
- $\mathcal{G}(\mathfrak{M}_{a,\mathfrak{o}_E})\simeq \mathrm{SL}(2,k_{E^i})\times\{\pm1\}$ and $\mathcal{G}(\mathfrak{M}_{b,\mathfrak{o}_E})\simeq \mathrm{O}(2,1)(k_{E^i})$ if E^i/E_0^i is ramified.

We set $f = f(E_0^i/F)$. From the calculations in Section 5 the possible values for r_a, r_b and the sets of real parts of reducibility points in (3-6) are

- if E^i/E_0^i is unramified: $r_a = 3f$ or f, and $r_b = f$; real parts $\{\pm 1, \pm\frac{1}{2}\}$ or $\{\pm\frac{1}{2}, 0\}$;
- if E^i/E_0^i is ramified: $r_a = f$ or 0, and $r_b = f$; real parts $\{\pm 1, 0\}$ or $\{\pm\frac{1}{2}, \pm\frac{1}{2}\}$.

In both cases the value of $r_b = r_b(\epsilon_{\mathfrak{M}_b}\tilde{\tau}_W \otimes \tau_i)$ is independent of the representation. We choose $\tilde{\tau}_W$ such that $r_a(\epsilon_{\mathfrak{M}_a}(\tilde{\tau}_W \otimes \tau_i)) = 3f$ if E^i/E_0^i is unramified, or $r_a(\epsilon_{\mathfrak{M}_a}(\tilde{\tau}_W \otimes \tau_i)) = f$ if E^i/E_0^i is ramified. This choice, which we denote by $\tilde{\underline{\tau}}_W$, is unique and provides us with a reducibility with real part 1.

We conclude that there exists one and only one self-dual cuspidal representation ρ of $\mathrm{GL}_F(W)$ containing the simple character $\tilde{\vartheta}_W$ such that the parabolically induced representation $\nu^1\rho\otimes\pi$ is reducible. The representation ρ contains the type $(\tilde{J}_W, \tilde{\underline{\kappa}}_W \otimes \tilde{\underline{\tau}}_W)$. However, as discussed previously, this does not give us a full description of the self-dual representation ρ: we know its inertial class but this still leaves two possibilities. This situation is explored more fully in Section 6.

3.17. Applying the previous subsection again to the representation π_i of $G_i = \mathrm{Sp}_F(V^i)$ and comparing (3-6) and (3-7), we remark that the relevant choice of $\tilde{\rho}_W$ for the situation in X, with the cuspidal representation π of $G = \mathrm{Sp}_F(V)$, differs from the analogous choice relative to the situation in X^i, with the cuspidal representation π_i of G_i, by a simple twist by the character $\epsilon_{\mathfrak{M}_a}\epsilon_{\mathfrak{M}_a^i}$. Indeed, in our example, the value of r_b is independent of the representation. In the next section we will study the general case, when r_a and r_b may both depend on the representation.

4. Reduction to the simple case

In this section, we make the reduction to the simple case, proving 2.6 Theorem. As intimated at the end of the last chapter, the key point to prove is that the character $\epsilon_{\mathfrak{M}_t}\epsilon_{\mathfrak{M}_t^i}$ is independent of t (see 4.3 Proposition). Note that the character $\epsilon_{\mathfrak{M}_t}\epsilon_{\mathfrak{M}_t^i}$ is the character χ_i appearing in the statement of 2.6 Theorem. While we have a description of it as a permutation character and, through careful analysis of this permutation, give a recipe by which one could compute it, we do not here compute it precisely; we only check that it is independent of t.

There is one further subtlety which should be remarked upon. In Section 3, we began with a pair of cuspidal representations (ρ, π) and built from them a cover of a type, without starting from types for ρ and π. In this section, we begin just with a cuspidal representation π of G and a cuspidal type λ for it, and use this to define certain cuspidal representations ρ of general linear groups, and maximal simple types $\tilde{\lambda}$ for them. The cover obtained in Section 3 is then indeed a cover of $\tilde{\lambda}\otimes\lambda$ but this is only clear because the (semi-)simple characters in λ and $\tilde{\lambda}$ are suitably related. Thus we take great care to set up the notation in this section.

4.1. We first review the notation that we need. This is the notation as in Section 2.6 so that it differs slightly from the notation of the previous section. In particular, objects in the symplectic space V do not have the subscript V; instead, the corresponding objects in X (which we have yet to define) will have the subscript X.

Throughout this and the following subsections, we fix a cuspidal representation π of $G = \mathrm{Sp}_F(V)$. We have the following data.

In the symplectic space V:

- A maximal skew semisimple stratum $[\Lambda, -, 0, \beta]$ in $\mathrm{End}_F(V)$ and a skew semisimple character θ of $H^1 = H^1(\beta, \Lambda)$ such that θ occurs in π.
- The irreducible representation η of $J^1 = J^1(\beta, \Lambda)$ containing θ and the p-primary beta-extension $\underline{\kappa}$ of η to $J = J(\beta, \Lambda)$.
- A cuspidal representation τ of $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) = P(\Lambda_{\mathfrak{o}_E})/P^1(\Lambda_{\mathfrak{o}_E})$ such that π is induced from $\lambda = \underline{\kappa} \otimes \tau$.

The stratum $[\Lambda, -, 0, \beta]$ can be written (uniquely) as an orthogonal direct sum of skew simple strata $[\Lambda^j, -, 0, \beta^j]$ in $\mathrm{End}_F(V^j)$, for $j = 0, \dots, l$, with the convention that $\beta^0 = 0$. The data above then give us the following data in the spaces V^j:

- Skew simple characters θ_j of $H^1_j = H^1(\beta^j, \Lambda^j)$, which are the restriction of θ.
- The irreducible representation η_j of $J^1_j = J^1(\beta^j, \Lambda^j)$ containing θ_j and the p-primary beta-extension $\underline{\kappa}_j$ of η_j to $J_j = J(\beta^j, \Lambda^j)$.
- The cuspidal representations τ_j of $\mathcal{G}(\Lambda^j_{\mathfrak{o}^j_E})$ such that, via the isomorphism $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) \simeq \prod_{j=0}^l \mathcal{G}(\Lambda^j_{\mathfrak{o}^j_E})$, we have $\tau = \bigotimes_{j=0}^l \tau_j$.
- The representation $\lambda_j = \underline{\kappa}_j \otimes \tau_j$ of J_j.

Note that, writing $G_j = \mathrm{Sp}_F(V^j)$, the representation $\pi_j = \text{c-Ind}_{J_j}^{G_j} \lambda_j$ is a cuspidal representation. A priori, it is not determined uniquely by the representation π, but it is determined by our choice of data $([\Lambda, -, 0, \beta], \theta)$ such that π contains θ.

We now fix $i \in \{0, \dots, l\}$ and choose an F-vector space W whose dimension is divisible by the degree $[E^i : F]$. We then have the following data.

In the vector space W:

- A maximal simple stratum $[\Lambda_W, -, 0, \beta_W]$ in $\mathrm{End}_F(W)$, together with a field isomorphism $E^i = F[\beta^i] \to F[\beta_W] = E_W$ fixing F and taking β^i to β_W.
- The simple character $\tilde{\vartheta}_W$ of $\widetilde{H}^1_W = \widetilde{H}^1(\beta_W, \Lambda_W)$ which is the transfer of the square $(\tilde{\theta}_i)^2$ of the unique self-dual simple character of $\widetilde{H}^1(\beta^i, \Lambda^i)$ restricting to θ_i.
- The p-primary beta-extension $\underline{\tilde{\kappa}}_W$ of $\tilde{\vartheta}_W$ to $\tilde{J}_W = dJ(\beta_W, \Lambda_W)$ and an irreducible self-dual cuspidal representation $\tilde{\tau}_W$ of $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_W})$, inflated to $\tilde{J}_W$, where we have written $\mathfrak{o}_W$ for the ring of integers of E_W.
- A self-dual cuspidal representation ρ of $\mathrm{GL}_F(W)$ containing $\tilde{\lambda}_W = \underline{\tilde{\kappa}}_W \otimes \tilde{\tau}_W$.

These data also induce data in the dual space W^* as follows. By duplicating if necessary, we assume that Λ_W has period divisible by 4 and that $\Lambda_W(-1) \neq \Lambda_W(0)$. (The reason for doing this is to ensure that the self-dual lattice sequence we will obtain conforms to our standard normalization — see 1.1 Remark.) Writing $\langle\cdot,\cdot\rangle$ for the pairing $W\times W^*\to F$, we define Λ_W^* by

$$\Lambda_W^*(r)=\{w^*\in W^*\mid\langle\Lambda_W(1-r),w^*\rangle\subseteq\mathfrak{p}_F\}\quad\text{for } r\in\mathbb{Z};$$

then the lattice sequence $\Lambda_W\oplus\Lambda_W^*$ is self-dual with respect to the natural symplectic structure on $W\oplus W^*$. We also define β_W^* in $\mathrm{End}_F(W^*)$ by

$$\langle w,\beta_W^*(w^*)\rangle=-\langle\beta_W(w),w^*\rangle\quad\text{for all } w\in W,\ w^*\in W^*.$$

Note that, by the fact that $[\Lambda^i,-,0,\beta^i]$ is skew, there is a unique isomorphism $E^i\to F[\beta_W^*]$ which takes β^i to β_W^*.

We now use these data to define corresponding data in the larger spaces on which we will have covers (as in Section 3). We define the symplectic space $X^i=(W\oplus W^*)\perp V^i$, for which we have the following.

In the symplectic space X^i:

- The maximal Levi subgroup $M_i\simeq\mathrm{GL}_F(W)\times\mathrm{Sp}_F(V^i)$ of $\mathrm{Sp}_F(X^i)$ which stabilizes the decomposition $X^i=(W\oplus W^*)\perp V^i$, and the maximal parabolic subgroup $P_i=M_iU_i$ which stabilizes the subspace W (and so stabilizes the flag $W\subseteq W\perp V^i\subset X^i$).
- The skew simple stratum $[\Lambda_X^i,-,0,\beta_X^i]$ in $\mathrm{End}_F(X^i)$, where $\Lambda_X^i=(\Lambda_W\oplus\Lambda_W^*)\perp\Lambda$ and β_X^i is the unique skew simple element which stabilizes the decomposition $X^i=(W\oplus W^*)\perp V^i$ and acts as β^i on V^i and as β_W on W; it then acts as β_W^* on W^*. We identify E^i with $F[\beta_X^i]$ via the isomorphism which takes β^i to β_X^i.
- Two further skew simple strata in $\mathrm{End}_F(X^i)$,
$$[\mathfrak{M}_0^i,-,0,\beta_X^i],\quad[\mathfrak{M}_1^i,-,0,\beta_X^i],$$
such that $\mathfrak{b}_0(\mathfrak{M}_t^i)$, for $t=0,1$, are the two maximal self-dual $\mathfrak{o}_E^i$-orders in the commuting algebra of β_X^i which contain $\mathfrak{b}_0(\Lambda_X^i)$.
- The unique skew simple character θ_X^i of $H_{X^i}^1=H^1(\beta_X^i,\Lambda_X^i)$ that restricts to θ_i on H_i^1 and to $\tilde{\vartheta}_W$ on $\widetilde{H}_W^1$; this is the transfer to Λ_X^i of the skew simple character θ_i.
- For $t=0,1$, the skew simple character $\theta_{\mathfrak{M}_t^i}$ of $H_{\mathfrak{M}_t^i}^1$ that is transferred from θ_X^i; the corresponding irreducible representation $\eta_{\mathfrak{M}_t^i}$ of $J_{\mathfrak{M}_t^i}^1$; and the p-primary beta-extension $\underline{\kappa}_{\mathfrak{M}_t^i}$ of $\eta_{\mathfrak{M}_t^i}$ to $J_{\mathfrak{M}_t^i}$.
- An $\mathrm{Sp}_F(X^i)$-cover (J_P^i,λ_P^i) of the pair $(\tilde{J}_W\times J_i,\tilde{\lambda}_W\otimes\lambda_i)$ in M_i.

Finally, we define the symplectic space $X=(W\oplus W^*)\perp V=X^i\perp V^{\vee i}$, where $V^{\vee i}=\perp_{j\neq i}V^j$, for which we have the following.

In the symplectic space X:

- The maximal Levi subgroup $M\simeq\mathrm{GL}_F(W)\times\mathrm{Sp}_F(V)$ of $\mathrm{Sp}_F(X)$ which stabilizes the decomposition $X=(W\oplus W^*)\perp V$, and the maximal parabolic subgroup $P=MU$ which stabilizes the subspace W.

- The skew semisimple stratum $[\Lambda_X, -, 0, \beta_X]$, where $\Lambda_X = \Lambda^i_X \perp \Lambda^{\vee i}$, with $\Lambda^{\vee i} = \bigperp_{j\neq i} \Lambda^j$, and β the unique skew semisimple element which stabilizes the decomposition $X = (W \oplus W^*) \perp V$ and acts as β on V and β_W on W (or, equivalently, acts as β^i_X on X^i and as $\beta^{\vee i} = \bigoplus_{j\neq i} \beta^j$ on $V^{\vee i}$, from which it is clear that the resulting stratum is indeed semisimple). We identify E with $F[\beta_X]$ via the isomorphism which takes β^i to β^i_X and β^j to itself, for $j \neq i$.
- Two further skew semisimple strata

$$[\mathfrak{M}_0, -, 0, \beta_X], \quad [\mathfrak{M}_1, -, 0, \beta_X],$$

where $\mathfrak{M}_t = \mathfrak{M}^i_t \perp \Lambda^{\vee i}$ for $t = 0, 1$; then $\mathfrak{b}_0(\mathfrak{M}_t)$ are the two maximal self-dual $\mathfrak{o}_E$-orders in the commuting algebra of β_X which contain $\mathfrak{b}_0(\Lambda_X)$.
- The unique skew semisimple character θ_X of $H^1_X = H^1(\beta_X, \Lambda_X)$ which restricts to θ on H^1 and to $\tilde{\vartheta}_W$ on $\widetilde{H}^1_W$; it is the transfer to H^1_X of the skew semisimple character θ, and restricts to θ^i_X on $H^1_{X^i}$.
- For $t = 0, 1$, the skew semisimple character $\theta_{\mathfrak{M}_t}$ of $H^1_{\mathfrak{M}_t}$ that is transferred from θ_X; the corresponding irreducible representation $\eta_{\mathfrak{M}_t}$ of $J^1_{\mathfrak{M}_t}$; and the p-primary beta-extension $\underline{\kappa}_{\mathfrak{M}_t}$ of $\eta_{\mathfrak{M}_t}$ to $J_{\mathfrak{M}_t}$.
- An $\mathrm{Sp}_F(X)$-cover (J_P, λ_P) of the pair $(\tilde{J}_W \times J, \tilde{\lambda}_W \otimes \lambda)$ in M.

4.2. We use the setup in the previous subsection and come back to the comparison of real parts of reducibility points, as in Section 3.17. The comparison of beta-extensions yields, as in 3.10 Proposition, characters $\epsilon_{\mathfrak{M}^i_t}$ and $\epsilon_{\mathfrak{M}_t}$ for $t = 0, 1$.

We fix $t = 0, 1$ and temporarily drop the subscript t. By definition $\epsilon_{\mathfrak{M}}(m)$, for $m \in J_{\mathfrak{M}} \cap M$, is the signature of the permutation $\mathrm{Ad}\, m : u \mapsto m^{-1}um$ of the quotient $J^1_{\mathfrak{M}} \cap U / H^1_{\mathfrak{M}} \cap U$, isomorphic to the $\mathbb{F}_p$-vector space $\mathfrak{J}^1_{\mathfrak{M}} \cap \mathbb{U} / \mathfrak{H}^1_{\mathfrak{M}} \cap \mathbb{U}$, where $\mathbb{U}$ is the Lie algebra of U (see 3.10 Proposition and 3.11 Proposition). The same holds with $\epsilon_{\mathfrak{M}^i}(m)$ for $m \in J_{\mathfrak{M}^i} \cap M^i$: it is the signature of the same permutation on $\mathfrak{J}^1_{\mathfrak{M}^i} \cap \mathbb{U}_i / \mathfrak{H}^1_{\mathfrak{M}^i} \cap \mathbb{U}_i$. On the other hand $\mathbb{U}$ is isomorphic to $\mathbb{U}_i \oplus \mathrm{Hom}_F(V^{\vee i}, W)$ in an M_i-equivariant way, and the action of $(m, y) \in \mathrm{GL}_F(W) \times \mathrm{Sp}_F(V^i)$ on $\phi \in \mathrm{Hom}_F(V^{\vee i}, W)$ is given by $\phi \mapsto m\phi$. The associated decompositions of the lattices $\mathfrak{J}^1$ and $\mathfrak{H}^1$ (as in [Bushnell and Kutzko 1993, Proposition 7.1.12]) lead to:

Lemma. *Let $(m, y) \in \widetilde{P}(\Lambda_{W,\mathfrak{o}_W}) \times P(\Lambda^i_{\mathfrak{o}^i_E})$. Then $(\epsilon_{\mathfrak{M}^i}\epsilon_{\mathfrak{M}})((m, y))$ is the signature of the permutation $\phi \mapsto m\phi$ of*

$$\mathfrak{X} := \mathfrak{J}^1_{\mathfrak{M}} \cap \mathrm{Hom}_F(V^{\vee i}, W)/\mathfrak{H}^1_{\mathfrak{M}} \cap \mathrm{Hom}_F(V^{\vee i}, W).$$

Now the quotient group $\widetilde{P}(\Lambda_{W,\mathfrak{o}_W})/\widetilde{P}^1(\Lambda_{W,\mathfrak{o}_W})$ is a general linear group $\mathrm{GL}_{m_W}(k_W)$ over the finite extension $k_W = k_{E^i}$ of k_F; this extension depends only on the endoclass of the simple character $\tilde{\vartheta}_W$. The lemma actually asserts that the character $\epsilon_{\mathfrak{M}}\epsilon_{\mathfrak{M}^i}$ is trivial on $P(\Lambda^i_{\mathfrak{o}^i_E})$ and factors through the signature of the natural left action of $\mathrm{GL}_{m_W}(k_W)$ on $\mathfrak{X}$.

4.3. Retrieving the subscripts t, our main tool is the following comparison of characters:

Proposition. *With notation as above, we have*

$$\epsilon_{\mathfrak{M}_0^i}\epsilon_{\mathfrak{M}_0} = \epsilon_{\mathfrak{M}_1^i}\epsilon_{\mathfrak{M}_1}.$$

This character, as a character of $\mathrm{GL}_{m_W}(k_W)$, *can be written as* $\chi_i \circ \det_{k_W}$, *where* χ_i *is a quadratic character of* $k_W^\times$ *which is independent of the choice of the space* W.

The independence on the space W (for a fixed choice of i) is particularly important. We postpone the proof of the proposition for now and, taking it for granted, deduce 2.6 Theorem.

4.4. Recall that we have written $\pi = \text{c-Ind}_J^G \lambda$ and, for $j = 0, \dots, l$, we have the cuspidal representation $\pi_j = \text{c-Ind}_{J_j}^{G_j} \lambda_j$ of $G_j = \mathrm{Sp}_F(V^j)$. We have θ_j, the simple character of H_j^1 contained in λ_j, and we write $\tilde{\theta}_j$ for the self-dual simple character of $\widetilde{H}_j^1$ which restricts to θ_j. Let Θ_j be the endoclass of the simple character $(\tilde{\theta}_j)^2$, which is a simple character for the stratum $[\Lambda^j, -, 0, 2\beta^j]$, and k_{Θ_j} for the corresponding extension of k_F.

Recall that, for an endoclass Θ and a character χ of the multiplicative group of the corresponding finite field k_Θ, we have

- $\mathrm{Jord}(\pi)$, the Jordan set of π (see Section 2.1);
- $\mathrm{IJord}(\pi, \Theta)$, the inertial Jordan set of π relative to Θ, which is the multiset of pairs $([\rho], m)$ for $(\rho, m) \in \mathrm{Jord}(\pi)$ such that ρ has endoclass Θ;
- $\mathrm{IJord}(\pi, \Theta)_\chi$, the χ-twisted inertial Jordan set of π relative to Θ, which is the multiset of pairs $([\rho]_\chi, m)$ with $(\rho, m) \in \mathrm{Jord}(\pi, \Theta)$.

Recall here that, if ρ contains a maximal simple type $(\tilde{J}, \tilde{\lambda})$, then $[\rho]_\chi$ denotes the inertial class of cuspidal representations containing $(\tilde{J}, \tilde{\lambda} \otimes \chi \circ \det)$ (see Section 2.4). Also, when χ is the trivial character we just write $\mathrm{IJord}(\pi, \Theta)$.

We restate 2.6 Theorem in a refined form:

Theorem. *Fix i with $0 \le i \le l$, and let χ_i be the character of $k_{\Theta_i}^\times$ such that $\chi_i \circ \det_{k_{\Theta_i}}$ is the twisting character in 4.3 Proposition. We have an equality of multisets*

$$\mathrm{IJord}(\pi, \Theta_i) = \mathrm{IJord}(\pi_i, \Theta_i)_{\chi_i}.$$

Proof. This is now just a matter of putting together the previous results. Let ρ be a cuspidal representation with endoclass Θ_i and use the notation of Section 4.1 so that ρ is a representation of $\mathrm{GL}_F(W)$ containing the maximal simple type $\tilde{\lambda}_W = \underline{\tilde{\kappa}}_W \otimes \tilde{\tau}_W$. The values of m, if any, for which $([\rho], m) \in \mathrm{IJord}(\pi_i, \Theta_i)$ can then be computed from (3-7): more precisely, they are

$$\left| \frac{r_0(\epsilon_{\mathfrak{M}_0^i}(\tilde{\tau}_W \otimes \tau_i)) \pm r_1(\epsilon_{\mathfrak{M}_1^i}(\tilde{\tau}_W \otimes \tau_i))}{t(\rho)} \right| - 1, \tag{4-1}$$

whenever these integers are strictly positive, together with positive integers less than this and of the same parity.

Now we consider the inertial class $[\rho]_{\chi_i}$. The cuspidal representations in this class contain the maximal simple type $\tilde{\lambda}_W \otimes \chi_i \circ \det = \underline{\tilde{\kappa}}_W \otimes (\tilde{\tau}_W \otimes \chi_i \circ \det)$. Then, using (3-6), we see that the values of m for which $([\rho]_{\chi_i}, m) \in \mathrm{IJord}(\pi, \Theta_i)$ are

$$\left| \frac{r_0(\epsilon_{\mathfrak{M}_0}((\tilde{\tau}_W \otimes \chi_i \circ \det) \otimes \tau_i)) \pm r_1(\epsilon_{\mathfrak{M}_1}((\tilde{\tau}_W \otimes \chi_i \circ \det) \otimes \tau_i))}{t(\rho)} \right| - 1,$$

whenever these integers are strictly positive, together with positive integers less than this and of the same parity. But 4.3 Proposition says that these are precisely the same integers as those in (4-1) (recall that all the characters here are quadratic), and the result follows. □

Remark. As we have seen in the proof, the pairs $([\rho], m)$ which appear in $\mathrm{IJord}(\pi)$ are determined by the values of $r_t = r_t(\epsilon_{\mathfrak{M}^i}(\tilde{\tau}_W \otimes \tau_i))$. Denote by ρ' the other self-dual cuspidal in the inertial class $[\rho]$. If ρ and ρ' are of opposite parity, say ρ is of symplectic type and ρ' is of orthogonal type, then we also recover this part of the full Jordan set $\mathrm{Jord}(\pi)$: if m is even then it is (ρ, m) which appears in $\mathrm{Jord}(\pi)$, while if m is odd then it is (ρ', m).

Suppose now that ρ, ρ' are of the same parity and $([\rho], m)$ appears in $\mathrm{IJord}(\pi)$. Then ρ and ρ' both appear *with the same multiplicities* in $\mathrm{Jord}(\pi)$ if and only if $r_0 r_1 = 0$. Thus in this case we also recover this part of the full Jordan set. Both ρ and ρ' appear with *some* multiplicity in $\mathrm{Jord}(\pi)$ if and only if $|r_0 - r_1| > t(\rho)$; when ρ, ρ' are both of orthogonal type, this condition simplifies to $r_0 \neq r_1$, since the reducibility points must be integers in this case.

The situations in which ρ, ρ' have the same parity are examined more closely from the Galois point of view in Section 6.

It remains now to prove 4.3 Proposition, which will take up the remainder of this section.

4.5. In this and the next few subsections, we define and study an auxiliary lattice sequence which will be needed for the calculations. Let Λ_W and Λ_Y be $\mathfrak{o}_F$-lattice sequences in finite-dimensional F-vector spaces W and Y respectively, with the same $\mathfrak{o}_F$-period e. We define an $\mathfrak{o}_F$-lattice sequence $\mathcal{C} = \mathcal{C}(\Lambda_Y, \Lambda_W)$ in the vector space $C = \mathrm{Hom}_F(Y, W)$ by

$$\mathcal{C}(t) = \{g \in C \mid g\Lambda_Y(i) \subseteq \Lambda_W(i+t) \text{ for all } i \in \mathbb{Z}\} \quad \text{for } t \in \mathbb{Z}.$$

We call the *jumps* of Λ_Y those integers i such that $\Lambda_Y(i) \neq \Lambda_Y(i+1)$ (and similarly for any lattice sequence). The set of jumps of Λ_Y is also the image of $Y \setminus \{0\}$ by the valuation map attached to Λ_Y, given by $\mathrm{val}_Y(y) = \max\{k \in \mathbb{Z} \mid y \in \Lambda_Y(k)\}$, for $y \in Y \setminus \{0\}$.

We make the following assumptions:

(i) The set of jumps of Λ_W is equal to $a_W + s_W\mathbb{Z}$ and the set of jumps of Λ_Y is equal to $a_Y + s_Y\mathbb{Z}$.

(ii) The orders $\mathfrak{a}(\Lambda_W)$ and $\mathfrak{a}(\Lambda_Y)$ are principal orders, in other words nonzero quotients $\Lambda_W(i)/\Lambda_W(i+1)$ are all isomorphic, and the same for Λ_Y. In particular there are an element $\Pi_W \in \mathfrak{a}(\Lambda_W)$ such that $\Pi_W(\Lambda_W(i)) = \Lambda_W(i+1)$ whenever i is a jump of Λ_W, and an element $\Pi_Y \in \mathfrak{a}(\Lambda_Y)$ such that $\Pi_Y(\Lambda_Y(i)) = \Lambda_Y(i+1)$ whenever i is a jump of Λ_Y [Bushnell and Kutzko 1993, §5.5].

Lemma. *The set of jumps of $\mathcal{C}$ is equal to $(a_W - a_Y) + \gcd(s_W, s_Y)\mathbb{Z}$. Moreover, the quotient spaces $\mathcal{C}(i)/\mathcal{C}(i+1)$ that are nonzero are all isomorphic as k_F-vector spaces, and their common dimension is*

$$c = c(\Lambda_Y, \Lambda_W) = \frac{\gcd(s_W, s_Y)}{e} \dim_F W \dim_F Y.$$

Proof. Proving that the set of jumps is contained in the given $\mathbb{Z}$-coset is straightforward using only (i). Now we use (ii) and remark that Π_W and Π_Y satisfy $\Pi_W(\Lambda_W(i)) = \Lambda_W(i+s_W)$ and $\Pi_Y(\Lambda_Y(i)) = \Lambda_Y(i+s_Y)$ for any integer i. For any $\phi \in C$ we check that

$$\mathrm{val}_{\mathcal{C}}(\Pi_W \phi) = \mathrm{val}_{\mathcal{C}}(\phi) + s_W, \quad \mathrm{val}_{\mathcal{C}}(\phi \Pi_Y) = \mathrm{val}_{\mathcal{C}}(\phi) + s_Y.$$

Thus left multiplication by Π_W is an isomorphism of $\mathfrak{o}_F$-modules from $\mathcal{C}(t)$ onto $\mathcal{C}(t+s_W)$ and right multiplication by Π_Y is an isomorphism of $\mathfrak{o}_F$-modules from $\mathcal{C}(t)$ onto $\mathcal{C}(t+s_Y)$, and the isomorphy follows.

To compute the dimension we use the generalized index notation $[A:B]$ for two lattices A and B in a same finite-dimensional vector space: $[A:B]$ is just the ordinary quotient of $[A:X]$ and $[B:X]$ for any lattice X contained in A and B.

The common $\mathfrak{o}_F$-period e is a multiple of s_W and s_Y, say $e = r_W s_W = r_Y s_Y$. Write $s = \gcd(s_W, s_Y)$ and pick integers n, m such that $s = n s_W + m s_Y$. We have, for any integer k,

$$\begin{aligned}
[\mathcal{C}(k) : \mathcal{C}(k+s)] &= [\mathcal{C}(k) : \mathcal{C}(k+ns_W)]\, [\mathcal{C}(k+ns_W) : \mathcal{C}(k+ns_W+ms_Y)] \\
&= [\mathcal{C}(k) : \mathcal{C}(k+s_W)]^n\, [\mathcal{C}(k) : \mathcal{C}(k+s_Y)]^m \\
&= [\mathcal{C}(k) : \varpi_F \mathcal{C}(k)]^{n/r_W}\, [\mathcal{C}(k) : \varpi_F \mathcal{C}(k)]^{m/r_Y} \\
&= [\mathcal{C}(k) : \varpi_F \mathcal{C}(k)]^{s/e},
\end{aligned}$$

and the result follows. □

4.6. We will need to determine the effect on $\mathcal{C}(\Lambda_Y, \Lambda_W)$ of a shift in indices on Λ_W. We further assume the following.

Notation. (i) The space W is an E_W-vector space for some finite extension E_W of F, with ramification index e_W and residue field k_W of cardinality q_W.

(ii) We fix two $\mathfrak{o}_W$-lattice sequences $\Lambda_{W,0}$ and $\Lambda_{W,1}$ in W with the same underlying lattice chain of period 1 over E_W (so that $s_{W,0} = s_{W,1} = e/e_W$) and with jumps at $a_{W,0} = 0$ and $a_{W,1} = e/(2e_W)$ respectively.

We write $s_Y = e/r_Y$ and put $\mathcal{C}_t = \mathcal{C}(\Lambda_Y, \Lambda_{W,t})$ for $t = 0, 1$. The sets of jumps of $\mathcal{C}_0, \mathcal{C}_1$ are respectively

$$-a_Y + \gcd\left(\frac{e}{r_Y}, \frac{e}{e_W}\right)\mathbb{Z} \quad \text{and} \quad \frac{e}{2e_W} - a_Y + \gcd\left(\frac{e}{r_Y}, \frac{e}{e_W}\right)\mathbb{Z};$$

they are the same when $e/(2e_W)$ divides $\gcd(e/r_Y, e/e_W)$. We get the following, where val_2 is the 2-adic valuation of an integer.

Lemma. *$\mathcal{C}_0$ and $\mathcal{C}_1$ have the same jumps if and only if* $\mathrm{val}_2(e_W) < \mathrm{val}_2(r_Y)$. *Otherwise the jumps of $\mathcal{C}_0$ and $\mathcal{C}_1$ are shifted by* $\frac{1}{2}\gcd(e/r_Y, e/e_W)$.

4.7. We now observe that the group $\mathrm{GL}_{m_W}(k_W)$ acts on the quotients $\mathcal{C}_t(i)/\mathcal{C}_t(i+1)$ by left multiplication, where $m_W = \dim_{E_W} W$. These actions commute with the left action of $E_W^\times$ and with the right action of Π_Y so, on the nonzero quotients, they are all equivalent and the corresponding permutations of the nonzero sets $\mathcal{C}_t(i)/\mathcal{C}_t(i+1)$ all have the same signature.

In the same fashion the nonzero quotients $\mathcal{C}_t(i)/\mathcal{C}_t(i+1)$ are isomorphic left modules over $\mathfrak{a}_0(\Lambda_{W,t,\mathfrak{o}_W})/\mathfrak{a}_1(\Lambda_{W,t,\mathfrak{o}_W}) \simeq M_{m_W}(k_W)$. The latter is a simple algebra; hence those modules have composition series with d simple quotients all isomorphic to the natural module $k_W^{m_W}$. The determinant of the action of $g \in \mathrm{GL}_{m_W}(k_W)$ on any such module is thus $(\det_{k_W} g)^d$ and the signature of the corresponding permutation is $((\det_{k_W} g)^{(q_W-1)/2})^d$ by 3.11 Lemma. The associated character of $\mathrm{GL}_{m_W}(k_W)$ is then trivial if and only if d is even. Now 4.5 Lemma gives us

$$d = \frac{1}{m_W[k_W : k_F]}\,\frac{\gcd(e/r_Y, e/e_W)}{e}\,\dim_F W \dim_F Y.$$

Since $\dim_F W = e_W[k_W : k_F]m_W$, we conclude:

Lemma. *The signature of the natural left action of* $\mathrm{GL}_{m_W}(k_W)$ *on the nontrivial quotients* $\mathcal{C}_t(i)/\mathcal{C}_t(i+1)$ *is the trivial character if and only if*

$$d = \frac{e_W}{\mathrm{lcm}(r_Y, e_W)}\dim_F Y$$

*is **even**; otherwise it is the unique character of* $\mathrm{GL}_{m_W}(k_W)$ *of order* 2. *In particular,*

- *this signature only depends on e_W, not on W itself*;
- *when $\mathcal{C}_0$ and $\mathcal{C}_1$ do not have the same jumps, we have* $d \equiv \dim_F Y \pmod 2$.

4.8. We return to the notation of Sections 4.1–4.2 but, for now, we drop the subscript t so that $\mathfrak{M}$ denotes either of the orders $\mathfrak{M}_0$ or $\mathfrak{M}_1$. We first detail the structure of the $\mathfrak{b}_0(\mathfrak{M})$-bimodule $\mathfrak{J}^1_{\mathfrak{M}} \cap \mathbb{U}/\mathfrak{H}^1_{\mathfrak{M}} \cap \mathbb{U}$, isomorphic to $J^1_{\mathfrak{M}} \cap U/H^1_{\mathfrak{M}} \cap U$ by the Cayley map, or equivalently by $Y \mapsto 1+Y$. (Recall that $\mathbb{U}$ denotes the Lie algebra of U.) We use the inductive definition of the orders $\mathfrak{J}_{\mathfrak{M}}$ and $\mathfrak{H}_{\mathfrak{M}}$ given in [Stevens 2005, §3.2].

We have, for some $u \ge 1$, a sequence $(\gamma_0 = \beta, \gamma_1, \dots, \gamma_u = 0)$ and a strictly increasing sequence of integers $0 < r_0 < \dots < r_{u-1} = n := v_{\mathfrak{M}}(\beta)$ such that, for $0 \le v \le u-1$, the stratum $[\mathfrak{M}, -, r_v - 1, \gamma_v]$ is semisimple and the stratum $[\mathfrak{M}, -, r_v, \gamma_v]$ is equivalent to $[\mathfrak{M}, -, r_v, \gamma_{v+1}]$. Using the inductive definition and writing $[Z]$ for the image in the Grothendieck group of a $\mathfrak{b}_0(\mathfrak{M})$-bimodule Z, we find that

$$[\mathfrak{J}^1_{\mathfrak{M}}/\mathfrak{H}^1_{\mathfrak{M}}] = [\mathfrak{a}_{\mathfrak{M}}^{n/2}/\mathfrak{a}_{\mathfrak{M}}^{(n/2)+}] - [\mathfrak{b}^{r_0/2}_{\gamma_0,\mathfrak{M}}/\mathfrak{b}^{(r_0/2)+}_{\gamma_0,\mathfrak{M}}] + \sum_{v=1}^{u-1}\big([\mathfrak{b}^{r_{v-1}/2}_{\gamma_v,\mathfrak{M}}/\mathfrak{b}^{(r_{v-1}/2)+}_{\gamma_v,\mathfrak{M}}] - [\mathfrak{b}^{r_v/2}_{\gamma_v,\mathfrak{M}}/\mathfrak{b}^{(r_v/2)+}_{\gamma_v,\mathfrak{M}}]\big), \tag{4-2}$$

where $\mathfrak{b}^{r/2}_{\gamma,\mathfrak{M}}$ is shorthand for the intersection of $\mathfrak{a}_{r/2}(\mathfrak{M})$ with the centraliser of γ.

From [Stevens 2005, Proposition 3.4], we may choose the elements γ_v so that the decomposition $X = V \perp (W \oplus W^*)$ is subordinate to all strata considered above; in particular we can take intersections with $\mathbb{U}$ in every term in the above equality. Then the value $\epsilon_{\mathfrak{M}}(m)$ of the quadratic character $\epsilon_{\mathfrak{M}}$ can be calculated as the product of the signatures of the permutation $\operatorname{Ad} m$ on each resulting quotient.

4.9. We now begin the proof of 4.3 Proposition. Recall that, by 4.2 Lemma, the character $\epsilon_{\mathfrak{M}^i}\epsilon_{\mathfrak{M}}$ is given by the signature of the permutation $\phi \mapsto m\phi$ on

$$\mathfrak{X} := \mathfrak{J}^1_{\mathfrak{M}} \cap \operatorname{Hom}_F(V^{\vee i}, W)/\mathfrak{H}^1_{\mathfrak{M}} \cap \operatorname{Hom}_F(V^{\vee i}, W)$$

for $m \in \widetilde{P}(\Lambda_{W,\mathfrak{o}_W})$.

The space $\operatorname{Hom}_F(V^{\vee i}, W)$ decomposes as a direct sum $\bigoplus_{j\neq i} \operatorname{Hom}_F(V^j, W)$. Moreover, each V^j in turn decomposes as a direct sum $V^j = \perp_{s=1}^{s_j} Y^{j,s}$ of subspaces $Y^{j,s}$ for which the assumptions of Section 4.5–4.7 are satisfied, and such that the resulting decomposition of V is subordinate to $[\Lambda, -, 0, \beta]$. Precisely:

- If β_j is nonzero, we take a direct sum of lines over E_j that splits the lattice sequence Λ^j as in [Bushnell and Kutzko 1999, §5.3, Lemma].
- If $\beta_j = 0$, the reductive quotient of the maximal parahoric subgroup $P(\Lambda^j)$ is isomorphic to the direct product of at most two symplectic groups over k_F, whence a decomposition of V^j as an orthogonal sum of at most two symplectic spaces satisfying the conditions required.

The action of $\mathrm{GL}_{m_W}(k_W)$ on $\mathfrak{X}$ then decomposes as a direct sum over j, s of actions on

$$\mathfrak{X}^{j,s} = \mathfrak{J}^1_{\mathfrak{M}} \cap \boldsymbol{Y}^{j,s}/\mathfrak{H}^1_{\mathfrak{M}} \cap \boldsymbol{Y}^{j,s},$$

where $\boldsymbol{Y}^{j,s} = \operatorname{Hom}_F(Y^{j,s}, W)$.

Using [Stevens 2005, Proposition 3.4] and [Bushnell and Kutzko 1999, §5.3, Corollary], we may choose the elements γ_v for (4-2) so that the decomposition $X = \perp_{j,s} Y^{j,s} \perp (W \oplus W^*)$ is subordinate to all strata considered. The action of $\mathrm{GL}_{m_W}(k_W)$ then decomposes further along (4-2) into pieces that fit the hypotheses of 4.7 Lemma, namely pieces of the forms

$$\begin{aligned}
Q_1 &= [\mathfrak{a}_{\mathfrak{M}}^{n/2} \cap \boldsymbol{Y}^{j,s}/\mathfrak{a}_{\mathfrak{M}}^{(n/2)+} \cap \boldsymbol{Y}^{j,s}],\\
Q_2 &= [\mathfrak{b}_{\gamma_0,\mathfrak{M}}^{r_0/2} \cap \boldsymbol{Y}^{j,s}/\mathfrak{b}_{\gamma_0,\mathfrak{M}}^{(r_0/2)+} \cap \boldsymbol{Y}^{j,s}],\\
Q_3 &= [\mathfrak{b}_{\gamma_v,\mathfrak{M}}^{r_{v-1}/2} \cap \boldsymbol{Y}^{j,s}/\mathfrak{b}_{\gamma_v,\mathfrak{M}}^{(r_{v-1}/2)+} \cap \boldsymbol{Y}^{j,s}] - [\mathfrak{b}_{\gamma_v,\mathfrak{M}}^{r_v/2} \cap \boldsymbol{Y}^{j,s}/\mathfrak{b}_{\gamma_v,\mathfrak{M}}^{(r_v/2)+} \cap \boldsymbol{Y}^{j,s}].
\end{aligned}$$

4.10. At last we come to the point, which is not actually to compute the character $\epsilon_{\mathfrak{M}^i}\epsilon_{\mathfrak{M}}$, but rather to prove that this character does not depend on the maximal self-dual order $\mathfrak{M}$. In our setting there are exactly two choices for $\mathfrak{M}$ with a given period e and duality invariant $d = 1$. Indeed, the lattice chain underlying the self-dual lattice sequence $\Lambda_X \cap (W \oplus W^*)$ is the disjoint union of two self-dual lattice chains, one containing a self-dual lattice and its multiples, the other containing a non-self-dual lattice (whose dual is $\mathfrak{p}_W$ times it) and its multiples. Let $\mathfrak{M}_0$ and $\mathfrak{M}_1$ be the two possible choices and write $\Lambda_{W,0}=\mathfrak{M}_0\cap W$ and $\Lambda_{W,1}=\mathfrak{M}_1\cap W$. According to [Stevens 2008, Lemma 6.7], the sets of jumps of $\Lambda_{W,0}$ and $\Lambda_{W,1}$ are $(e/e_W)\mathbb{Z}$ and $(e/(2e_W))+(e/e_W)\mathbb{Z}$ respectively, and all results in Sections 4.5–4.7 apply.

We can thus compare $\epsilon_{\mathfrak{M}_0^i}\epsilon_{\mathfrak{M}_0}$ and $\epsilon_{\mathfrak{M}_1^i}\epsilon_{\mathfrak{M}_1}$ term by term.

Term Q_1: We apply Sections 4.5–4.7, replacing Y by $Y^{j,s}$, Λ_Y by $\Lambda\cap Y^{j,s}$ and using $\Lambda_{W,t}=\mathfrak{M}_t\cap W$ as above for $t=0,1$. We remark that $\dim_F Y^{j,s}$ is always even. Hence, by 4.7 Lemma, the signature on Q_1 is trivial unless $\mathcal{C}_0$ and $\mathcal{C}_1$ have the same jumps, so give the same signature.

Term Q_2: We actually have $\gamma_0=\beta$; hence this term is zero if the centralizer of β does not intersect $\mathrm{Hom}_F(V^j,W)$. This condition holds under the assumptions of 4.3 Proposition because $j\neq i$.

Term Q_3: Since $\mathfrak{M}_0$ and $\mathfrak{M}_1$ have the same intersection with V, we may and do choose the same sequence (γ_v,r_v) for both. We may also scale all our lattice sequences to make the period big enough so that all numbers $r_v/2$ are integers. Now Q_3 is zero unless the centralizer of γ_v intersects $\mathrm{Hom}_F(V^j,W)$, which we now assume. We then apply Sections 4.5–4.7 over $F[\gamma_v]$.

If the lattice sequences $\mathcal{C}_0$ and $\mathcal{C}_1$ have the same jumps we have the equality we want. Otherwise, they are shifted by half a period (4.6 Lemma) and the integer d given by 4.7 Lemma is equal to $\dim_{F[\gamma_v]}Y^{j,s}$. If d is even we are also done. Otherwise we have $\beta_j\neq 0$ and $s_\Lambda=e/e_j$, and the period of $\mathcal{C}_0$ and $\mathcal{C}_1$ is $e/\mathrm{lcm}(e_W,e_j)$.

Since Q_3 is the difference of two terms $[\mathcal{C}_t(a)]-[\mathcal{C}_t(b)]$ in the lattice sequence $\mathcal{C}_t$, for $t=0,1$, over $F[\gamma_v]$, the values of Q_3 for $t=0$ and $t=1$ will be the same on condition that the difference $a-b$ is a multiple of half the period. This is what we will now prove.

In the notation of (4-2), we let $h\geq 1$ be the smallest integer such that the centralizer of γ_h intersects $\mathrm{Hom}_F(V^j,W)$, so that we only need to consider terms with $v\geq h$. If $h=u$ there is nothing to do. Otherwise, we need to examine the values of r_h and r_{h-1} more closely, in terms of the *normalized critical exponents* $k_0^F[\gamma_v]$; see [Stevens 2005, pp. 129,141–142]. We use Lemma 3.7(ii) of that work for $r_{h-1}=-k_0(\gamma_{h-1},\mathfrak{M})$ (the unnormalized critical exponent relative to $\mathfrak{M}$) and case (i) for $r_h=-k_0(\gamma_h,\mathfrak{M})$ to get

$$r_{h-1}=v_{F[\gamma_h]}(c)\frac{e(\mathfrak{M}|\mathfrak{o}_F)}{e(F[\gamma_h]/F)},\qquad r_h=-k_0^F(\gamma_h)\frac{e(\mathfrak{M}|\mathfrak{o}_F)}{e(F[\gamma_h]/F)}$$

for some element c in $F[\gamma_h]$, so that

$$\frac{r_{h-1}}{2}-\frac{r_h}{2}=\frac{e(\mathfrak{M}|\mathfrak{o}_{F[\gamma_h]})}{2}(v_{F[\gamma_h]}(c)+k_0^F(\gamma_h)).$$

This is indeed an integer multiple of the half-period of jumps

$$\frac{e(\mathfrak{M}|\mathfrak{o}_{F[\gamma_h]})}{2}\left(\frac{\mathrm{lcm}(e_W,e_j)}{e(F[\gamma_h]/F)}\right)^{-1}$$

since the last term is the inverse of an integer.

For $v > h$ we use [Stevens 2005, Lemma 3.7(i)] and get

$$\frac{r_{v-1}}{2} - \frac{r_v}{2} = \frac{e(\mathfrak{M}|\mathfrak{o}_{F[\gamma_v]})}{2}\left(-k_0^F(\gamma_{v-1})\frac{e(F[\gamma_v]/F)}{e(F[\gamma_{v-1}]/F)} + k_0^F(\gamma_v)\right).$$

This is a multiple of the half-period of jumps if and only if

$$\frac{-k_0^F(\gamma_{v-1})e(F[\gamma_v]/F) + k_0^F(\gamma_v)e(F[\gamma_{v-1}]/F)}{e(F[\gamma_{v-1}]/F)}\,\frac{\operatorname{lcm}(e_W, e_j)}{e(F[\gamma_v]/F)}$$
$$= \left(-k_0^F(\gamma_{v-1}) + k_0^F(\gamma_v)\frac{e(F[\gamma_{v-1}]/F)}{e(F[\gamma_v]/F)}\right)\frac{\operatorname{lcm}(e_W, e_j)}{e(F[\gamma_{v-1}]/F)}$$

is an integer, which is the case because $e(F[\gamma_v]/F)$ divides $e(F[\gamma_{v-1}]/F)$; see [Bushnell and Kutzko 1993, 2.4.1].

Putting this together, we obtain the character $\epsilon_{\mathfrak{M}^i}\epsilon_{\mathfrak{M}}$ as a product of signatures, each of them only depending on e_W by 4.7 Lemma; hence our character only depends on e_W, not on W. Furthermore, the extension E_W is isomorphic to $F[\beta_i]$; hence e_W is equal to $e(F[\beta_i]/F)$, independent of the choice of W. This completes the proof of 4.3 Proposition, and hence that of 4.4 Theorem.

5. The simple case

In this section we prove 2.5 Theorem. Recalling that, by 3.12 Theorem, the parameters of the Hecke algebra of our cover are those in the Hecke algebra of a finite reductive group, we are required to analyze these Hecke algebras. Fortunately, these are described by [Lusztig 1984] and have been computed in our cases in [Lust and Stevens 2016]. One subtlety is that the twisting characters $\epsilon_{\mathfrak{M}_t}$ give rise to involutions which we have not computed explicitly and so remain unknown. Fortunately, the numerics are such that an exact description of these involutions is not needed.

5.1. Let π be a simple cuspidal representation of G in the sense of Section 2.5. Since the case of depth-zero representations is already dealt with in [Lust and Stevens 2016], we assume moreover that π has positive depth. Thus π contains a skew *simple* character θ of $H^1 = H^1(\beta, \Lambda)$, for some maximal skew simple stratum $[\Lambda, -, 0, \beta]$, with $\beta \neq 0$, and $E = F[\beta]$ is a field. We write Θ for the endoclass of the unique self-dual simple character $\tilde{\theta}$ which restricts to θ. We retain all the notation of Section 4.1 and so interpret simplicity as meaning that $l = 1$ and drop the index 1 for notation. We will be considering the space $X = X^1$, while varying the self-dual cuspidal representation ρ of $\mathrm{GL}_F(W)$ (and the space W). Note that we have $E_W \simeq E$ so we will identify them.

For a self-dual cuspidal representation ρ of some $\mathrm{GL}_F(W)$, recall that we write $\deg\rho = \dim_F W$ and $s_\pi(\rho)$ for the unique nonnegative real number such that the normalized induced representation $\nu^s\rho\times\pi$ is reducible. Then the description of the Jordan set in Section 2.1 shows that, in order to prove 2.5 Theorem, the equality we must prove is

$$\sum_\rho \lfloor s_\pi(\rho)^2\rfloor \deg\rho = 2N, \tag{5-1}$$

where the sum runs over all self-dual cuspidal representations ρ with endoclass $\Theta(\rho) = \Theta^2$.

5.2. Recall that we have $\pi = \text{c-Ind}_J^G \lambda$, with $\lambda = \underline{\kappa} \otimes \tau$ and that ρ contains the maximal simple type $\tilde{\lambda}_W = \underline{\tilde{\kappa}}_W \otimes \tilde{\tau}_W$ and has unramified twist number $t(\rho) = (\dim_F W)/e(\Theta)$, where we have written $e(\Theta) = e(\Theta^2) = e(E/F)$ since it depends only on the endoclass. Moreover, by 3.14 Proposition, we have that the real parts of the reducibility points of the normalized induced representation $\nu^s \rho \times \pi$ are the elements of the set

$$\left\{ \pm \frac{r_0 + r_1}{2t(\rho)}, \pm \frac{r_0 - r_1}{2t(\rho)} \right\},$$

where, for $t = 0, 1$, the integers $r_t = r_t(\epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau))$ come from the quadratic relations in the finite Hecke algebra $\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau))$ as in (3-5).

Remark. It will be crucial to note that the character $\epsilon_{\mathfrak{M}_t}$ depends only on the dimension $\deg \rho = \dim_F W$, and not on the representation ρ itself.

The contribution to the sum (5-1) of the inertial class $[\rho]$ (that is, writing $\rho' = \nu^{\pi i/t(\rho)\log(q)} \rho$ for the other self-dual representation in the inertial class, the combined contributions of ρ and ρ') is

$$\left\lfloor \left(\frac{r_0 + r_1}{2t(\rho)} \right)^2 \right\rfloor + \left\lfloor \left(\frac{r_0 - r_1}{2t(\rho)} \right)^2 \right\rfloor.$$

From results of Lusztig (see [Lust and Stevens 2016, §8] and also Section 5.6 below), the numbers $r_t/t(\rho)$ are either both integers or both half-integers so that this simplifies to

$$\left\lfloor \frac{r_0^2 + r_1^2}{2t(\rho)^2} \right\rfloor. \tag{5-2}$$

5.3. In order to prove (5-1) we will need to recall Lusztig's parametrization of cuspidal representations of classical groups, and the computation of the parameter r_t in the Hecke algebra $\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau))$. We follow the description in [Lust and Stevens 2016, §2, 3, 6 and, especially, 7], to which we refer for details and references for the assertions made here.

In almost all cases, we have

$$\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau)) \simeq \mathscr{H}(\mathcal{G}^\circ(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau^\circ)),$$

where τ° is an irreducible component of the restriction $\tau_{|\mathcal{G}^\circ(\Lambda_{\mathfrak{o}_E})}$, and it is here that we will perform our calculations. In the exceptional cases we have $r_t = 0$ and it will turn out that this matches the formula one would obtain by following the recipe for computing the parameters in the connected component $\mathcal{G}^\circ(\mathfrak{M}_{t,\mathfrak{o}_E})$. Thus we will assume first that the calculation is to be done in $\mathcal{G}^\circ(\mathfrak{M}_{t,\mathfrak{o}_E})$ and then, in Section 5.7, we will treat the exceptional cases.

5.4. Since $P(\Lambda_{\mathfrak{o}_E})$ is the normalizer of a maximal parahoric subgroup of the centraliser G_E, we have the decomposition

$$\mathcal{G}^\circ(\Lambda_{\mathfrak{o}_E}) = \mathcal{G}^{(0)}(\Lambda_{\mathfrak{o}_E}) \times \mathcal{G}^{(1)}(\Lambda_{\mathfrak{o}_E}),$$

which is a product of two connected classical groups over k_E° (the residue field of the fixed points $E_\circ$ in E under the involution on A). We have a similar decomposition of $\mathcal{G}^\circ(\mathfrak{M}_{t,\mathfrak{o}_E})$ with, moreover,

$$\mathcal{G}^{(1)}(\mathfrak{M}_{0,\mathfrak{o}_E}) = \mathcal{G}^{(1)}(\Lambda_{\mathfrak{o}_E}) \quad \text{and} \quad \mathcal{G}^{(0)}(\mathfrak{M}_{1,\mathfrak{o}_E}) = \mathcal{G}^{(0)}(\Lambda_{\mathfrak{o}_E}),$$

and the Levi subgroup

$$\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \times \mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E}) \subseteq \mathcal{G}^{(t)}(\mathfrak{M}_{t,\mathfrak{o}_E}).$$

We choose an irreducible component τ° of the restriction $\tau_{|\mathcal{G}^\circ(\Lambda_{\mathfrak{o}_E})}$ and write it as $\tau^{(0)} \otimes \tau^{(1)}$. Writing the character $\epsilon_{\mathfrak{M}_t}$ as $\epsilon_{\mathfrak{M}_t}^W \otimes \epsilon_{\mathfrak{M}_t}^{(0)} \otimes \epsilon_{\mathfrak{M}_t}^{(1)}$, we have isomorphisms of Hecke algebras

$$\mathscr{H}(\mathcal{G}^\circ(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau^\circ)) \simeq \mathscr{H}(\mathcal{G}^{(t)}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}^W \tilde{\tau}_W \otimes \epsilon_{\mathfrak{M}_t}^{(t)} \tau^{(t)}),$$

and it is in this Hecke algebra that we compute the parameter r_t.

5.5. We now fix $t = 0$ or 1 and so drop the sub/superscript t from our notations for now. Thus we have

- a connected classical group $\mathcal{G}^\circ(\mathfrak{M}_{\mathfrak{o}_E})$ over k_E°, with Levi subgroup $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \times \mathcal{G}(\Lambda_{\mathfrak{o}_E})$ and $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \simeq \mathrm{GL}_m(k_E)$, where $m = \dim_E W$;
- a self-dual cuspidal representation $\tilde{\tau}_W \otimes \tau$ of $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \times \mathcal{G}(\Lambda_{\mathfrak{o}_E})$;
- a character $\epsilon_{\mathfrak{M}}^W \otimes \epsilon_{\mathfrak{M}}$ of $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \times \mathcal{G}(\Lambda_{\mathfrak{o}_E})$ of order at most 2, which depends on $m = \dim_E W$ but not on $\tilde{\tau}_W$.

By Green's parametrization (and after fixing an isomorphism $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E}) \simeq \mathrm{GL}_m(k_E)$), the cuspidal representation $\tilde{\tau}_W$ corresponds to an irreducible monic polynomial $Q \in k_E[X]$ of degree m. Moreover, this polynomial is k_E/k_E°-self-dual, that is,

$$Q(X) = (\overline{Q(0)})^{-1} X^{\deg Q}\, \overline{Q}(1/X),$$

where $x \mapsto \bar{x}$ is the automorphism of k_E with fixed field k_E°, extended to $k_E[X]$ coefficientwise; see [Lust and Stevens 2016, §7.1]. Since a cuspidal representation $\tilde{\tau}_W$ of $\widetilde{\mathcal{G}}(\Lambda_{W,\mathfrak{o}_E})$ is self-dual if and only if $\tilde{\tau}_W \epsilon_{\mathfrak{M}}^W$ is cuspidal self-dual, twisting by $\epsilon_{\mathfrak{M}}^W$ induces an involution on the set of irreducible k_E/k_E°-self-dual monic polynomials of degree m. We denote this involution by $\sigma_{m,W}$; it is either trivial, or given by $Q(X) \mapsto (-1)^{\deg Q} Q(-X)$.

Similarly, by Lusztig's parametrization, the cuspidal representation τ lies in a *rational Lusztig series* $\mathcal{E}(\boldsymbol{s})$ corresponding to (the rational conjugacy class of) a semisimple element $\boldsymbol{s}$ of the dual group of $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$. Since its series contains a cuspidal representation, this semisimple element $\boldsymbol{s}$ has characteristic polynomial of a particular form, namely

$$P_{\boldsymbol{s}}(X) = \prod_P P(X)^{a_P},$$

where the product is over all irreducible k_E/k_E°-self-dual monic polynomials and the integers a_P satisfy certain combinatorial constraints (see [Lust and Stevens 2016, (7.2) and §7.7]); more precisely, we have:

- $\sum_P a_P \deg P$ is the dimension of the space $\mathcal{V}$ on which the dual group of $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$ naturally acts.
- If either $k_E \neq k_E^\circ$ or $P(X) \neq (X \pm 1)$, then $a_P = \frac{1}{2}(b_P^2 + b_P)$, for some nonnegative integers b_P.
- If $P(X) = X \pm 1$ then, writing $a_+ := a_{(X-1)}$ and $a_- := a_{(X+1)}$, there are integers $b_+, b_- \geq 0$ such that
  (i) if $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$ is odd special orthogonal then $a_+ = 2(b_+^2 + b_+)$ and $a_- = 2(b_-^2 + b_-)$,
  (ii) if $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$ is symplectic then $a_+ = 2(b_+^2 + b_+) + 1$ and $a_- = 2b_-^2$,
  (iii) if $\mathcal{G}(\Lambda_{\mathfrak{o}_E})$ is even special orthogonal then $a_+ = 2b_+^2$ and $a_- = 2b_-^2$,

  and, in case (iii), the (± 1)-eigenspace in $\mathcal{V}$ is an even-dimensional orthogonal space of type $(-1)^{b_\pm}$, and the same in case (ii) for the (-1)-eigenspace only.

As above, twisting by the character $\epsilon_{\mathfrak{M}}$ will induce a degree-preserving involution on the set of irreducible k_E/k_E°-self-dual monic polynomials. If the character $\epsilon_{\mathfrak{M}}$ is trivial then this involution is trivial. If the character $\epsilon_{\mathfrak{M}}$ is nontrivial quadratic then, by [Cabanes and Enguehard 2004, Proposition 8.26], twisting by $\epsilon_{\mathfrak{M}}$ induces a bijection between rational Lusztig series

$$\mathcal{E}(s) \xrightarrow{\sim} \mathcal{E}(-s),$$

and the involution is given by $P(X) \mapsto (-1)^{\deg P} P(-X)$. In either case, we denote by $\sigma_{m,G}$ the involution induced by twisting by $\epsilon_{\mathfrak{M}}$. (Note that this is a degree-preserving involution on the set of *all* irreducible k_E/k_E°-self-dual monic polynomials; the subscript m is included to indicate that the involution depends on m.) The characteristic polynomial corresponding to the cuspidal representation $\tau\epsilon_{\mathfrak{M}}$ is then

$$\prod_P P(X)^{a_{\sigma_{m,G}(P)}}.$$

Putting together our two involutions, we get an involution on the set of irreducible k_E/k_E°-self-dual monic polynomials of degree m given by

$$\sigma_m = \sigma_{m,G} \circ \sigma_{m,W}.$$

5.6. Recall that the Hecke algebra $\mathscr{H}(\mathcal{G}(\mathfrak{M}_{\mathfrak{o}_E}), \epsilon_{\mathfrak{M}}^W \tilde{\tau}_W \otimes \epsilon_{\mathfrak{M}}\tau)$ is generated by an element $\mathcal{T}$ satisfying a quadratic relation

$$(\mathcal{T} - q^r\omega)(\mathcal{T} + \omega) = 0,$$

where q is the cardinality of the residue field of k_F. The work of Lusztig, as presented in [Lust and Stevens 2016, §7], allows one to write down explicitly the parameter r in terms of the characteristic polynomials of the previous subsection, as follows.

Let $Q(X)$ be the irreducible k_E/k_E°-self-dual monic polynomial of degree m corresponding to $\tilde{\tau}_W$, and let $P_s(X) = \prod_P P(X)^{a_P}$ be the monic polynomial corresponding to τ, where the a_P are as described in the previous subsection. Writing f for the degree of the extension k_E/k_F, one gets the following values:

- If $k_E = k_E^\circ$ and $\sigma_1(Q) = X - 1$ then

$$\frac{r}{f} = \begin{cases} 2b_+ & \text{if } \mathcal{G} \text{ is even special orthogonal,} \\ 2b_+ + 1 & \text{otherwise.} \end{cases}$$

- If $k_E = k_E^{\circ}$ and $\sigma_1(Q) = X+1$ then

$$\frac{r}{f} = \begin{cases} 2b_- + 1 & \text{if } \mathcal{G} \text{ is odd special orthogonal,} \\ 2b_- & \text{otherwise.} \end{cases}$$

- If $k_E \neq k_E^{\circ}$ or m is even then

$$\frac{r}{f} = (2b_{\sigma_m(Q)} + 1)\frac{m}{2}.$$

Note that, since $t(\rho) = mf$, the number $r/t(\rho)$ is a half-integer, as asserted above. Moreover, $r/t(\rho)$ is an integer precisely when E/E_{o} is ramified and E/F is a maximal extension (i.e., of degree $\dim_F W$); in particular, this depends only on the polynomial Q (that is, on $\tilde{\tau}_W$, so on the representation ρ) and not on either the representation τ or on the involution σ_1.

5.7. In this subsection, we treat the exceptional cases, where we do *not* have an isomorphism

$$\mathscr{H}(\mathcal{G}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau)) \simeq \mathscr{H}(\mathcal{G}^{\circ}(\mathfrak{M}_{t,\mathfrak{o}_E}), \epsilon_{\mathfrak{M}_t}(\tilde{\tau}_W \otimes \tau^{\circ})). \tag{5-3}$$

According to the description in [Miyauchi and Stevens 2014, §6.3], this occurs precisely when

- E/E_{o} is ramified;
- $\dim_E W = 1$, so that $\tilde{\tau}_W$ is a character of order at most 2;
- and either $\mathcal{G}(\Lambda_{\mathfrak{o}_E}) = \mathcal{G}^{\circ}(\Lambda_{\mathfrak{o}_E})$ or $\epsilon_{\mathfrak{M}_t}\tau_{|\mathcal{G}^{\circ}(\Lambda_{\mathfrak{o}_E})}$ is reducible.

We remark that $\epsilon_{\mathfrak{M}_t}\tau_{|\mathcal{G}^{\circ}(\Lambda_{\mathfrak{o}_E})}$ is reducible if and only if $\tau_{|\mathcal{G}^{\circ}(\Lambda_{\mathfrak{o}_E})}$ is reducible.

In these cases, writing $\mathcal{G}^{\circ}(\mathfrak{M}_{t,\mathfrak{o}_E}) = \mathcal{G}^{(0)}(\Lambda_{\mathfrak{o}_E}) \times \mathcal{G}^{(1)}(\Lambda_{\mathfrak{o}_E})$, there is one value of t for which $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ is an even special orthogonal group (for the other it is a symplectic group) and it is precisely for this value of t that we do not have an isomorphism (5-3) and we get $r_t = 1$.

As above, we write $\tau^{(0)} \otimes \tau^{(1)}$ for an irreducible component of $\tau_{|\mathcal{G}^{\circ}(\Lambda_{\mathfrak{o}_E})}$, and write $\epsilon_{\mathfrak{M}_t}$ as $\epsilon^W_{\mathfrak{M}_t} \otimes \epsilon^{(0)}_{\mathfrak{M}_t} \otimes \epsilon^{(1)}_{\mathfrak{M}_t}$. Writing $P_s(X) = \prod P(X)^{a_P}$ for the polynomial corresponding to the cuspidal representation $\tau^{(t)}$, the fact that it does *not* extend to the full even orthogonal group implies, by [Lust and Stevens 2016, Proposition 7.9], that ± 1 are *not* roots of P_s, that is, $a_+ = a_- = 0$.

Since $\tilde{\tau}_W$ is a character of order at most 2, the corresponding polynomial is $Q(X) = X \pm 1$. In particular, since we have $b_+ = b_- = 0$, the formulae of Section 5.6 are still valid, since they too give $r_t = 0$. Thus those formulae are valid in every case.

5.8. Finally, using the formulae of Section 5.6, we return to computing the contribution (5-2), so we retrieve the sub/superscripts t. We have

- an irreducible k_E/k_E°-self-dual monic polynomial $Q(X)$, corresponding to the cuspidal representation $\tilde{\tau}_W$;
- for $t = 0, 1$, a polynomial $\prod_P P(X)^{a_P^{(t)}}$ corresponding to the cuspidal representation $\tau^{(t)}$;
- for $t = 0, 1$, an involution $\sigma_m^{(t)}$ on the set of irreducible k_E/k_E°-self-dual monic polynomials of degree m.

Suppose first that either $k_E \neq k_E^\circ$ or m is even; then we get

$$\left\lfloor \frac{r_0^2+r_1^2}{2t(\rho)^2} \right\rfloor = \left\lfloor \frac{(2b^{(0)}_{\sigma_m^{(0)}(Q)}+1)^2+(2b^{(1)}_{\sigma_m^{(1)}(Q)}+1)^2}{8} \right\rfloor$$
$$= \left\lfloor \tfrac{1}{2}b^{(0)}_{\sigma_m^{(0)}(Q)}(b^{(0)}_{\sigma_m^{(0)}(Q)}+1)+\tfrac{1}{2}b^{(1)}_{\sigma_m^{(1)}(Q)}(b^{(1)}_{\sigma_m^{(1)}(Q)}+1)+\tfrac{1}{2}\right\rfloor = a^{(0)}_{\sigma_m^{(0)}(Q)}+a^{(1)}_{\sigma_m^{(1)}(Q)}.$$

If $k_E = k_E^\circ$ then one of the groups $\mathcal{G}^{(t)}(\mathfrak{M}_{t,\mathfrak{o}_E})$ is symplectic, while the other is orthogonal. Here we can treat each case, each polynomial $X \pm 1$, and each possibility for the involutions $\sigma_1^{(t)}$, separately. Up to permuting $\{0, 1\}$ we are in one of the following two cases:

If $\mathcal{G}^{(0)}(\Lambda_{\mathfrak{o}_E})$ is odd special orthogonal and $\mathcal{G}^{(1)}(\Lambda_{\mathfrak{o}_E})$ is symplectic, then the contribution of $\sigma_1^{(1)}(X-1)$ is

$$\left\lfloor \frac{(2b^{(0)}_\zeta+1)^2+(2b^{(1)}_+ +1)^2}{2} \right\rfloor = 2b^{(0)}_\zeta(b^{(0)}_\zeta+1)+2b^{(1)}_+(b^{(1)}_+ +1)+1 = a^{(0)}_\zeta + a^{(1)}_+,$$

where ζ is the sign defined by $\sigma_1^{(0)}\sigma_1^{(1)}(X-1) = X-\zeta$; and the contribution of $\sigma_1^{(1)}(X+1)$ is

$$\left\lfloor \frac{(2b^{(0)}_{-\zeta}+1)^2+(2b^{(1)}_-)^2}{2} \right\rfloor = 2b^{(0)}_{-\zeta}(b^{(0)}_{-\zeta}+1)+2(b^{(1)}_-)^2 = a^{(0)}_{-\zeta}+a^{(1)}_-.$$

In particular, the sum of the contributions of $X \pm 1$ is

$$a^{(0)}_+ + a^{(0)}_- + a^{(1)}_+ + a^{(1)}_-.$$

If $\mathcal{G}^{(0)}(\Lambda_{\mathfrak{o}_E})$ is even special orthogonal and $\mathcal{G}^{(1)}(\Lambda_{\mathfrak{o}_E})$ is symplectic, then the contribution of $\sigma_1^{(1)}(X-1)$ is

$$\left\lfloor \frac{(2b^{(0)}_\zeta)^2+(2b^{(1)}_+ +1)^2}{2} \right\rfloor = 2(b^{(0)}_\zeta)^2+2b^{(1)}_+(b^{(1)}_+ +1) = a^{(0)}_\zeta + a^{(1)}_+ - 1,$$

where ζ is again the sign defined by $\sigma_1^{(0)}\sigma_1^{(1)}(X-1) = X-\zeta$, and the contribution of $\sigma_1^{(1)}(X+1)$ is

$$\left\lfloor \frac{(2b^{(0)}_{-\zeta})^2+(2b^{(1)}_-)^2}{2} \right\rfloor = 2(b^{(0)}_{-\zeta})^2+2(b^{(1)}_-)^2 = a^{(0)}_{-\zeta}+a^{(1)}_-.$$

In this second case, the sum of the contributions of $X \pm 1$ is

$$a^{(0)}_+ + a^{(0)}_- + a^{(1)}_+ + a^{(1)}_- - 1;$$

the term -1 reflects the fact that the sum of the dimensions of the spaces on which the dual groups of $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ act naturally is 1 more than the sum of the dimensions of the spaces on which the groups $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ act naturally. Note also that this latter sum of dimensions is precisely $\dim_E V$, where we recall that V is the symplectic space on which our group G acts.

5.9. Having computed all the contributions to the sum (5-1) in the previous subsection, we can now sum them over all possible Q, noting that, if the cuspidal representation ρ corresponds to the polynomial Q, then $\deg\rho = [E:F]\deg Q$. If $k_E \neq k_E^\circ$, this is straightforward and we obtain

$$\sum_\rho \lfloor s_\pi(\rho)^2\rfloor \deg\rho = [E:F]\sum_m\Bigl(m\sum_{\deg Q=m}(a^{(0)}_{\sigma_m^{(0)}(Q)} + a^{(1)}_{\sigma_m^{(1)}(Q)})\Bigr)$$
$$= [E:F]\Bigl(\sum_P a_P^{(0)}\deg P + \sum_P a_P^{(1)}\deg P\Bigr) = [E:F]\dim_E V = 2N,$$

as required. Here the penultimate equality occurs because each group $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ is a unitary group (whose dual group is then a unitary group acting naturally on a space of the same dimension), and the sum of the dimensions of the spaces on which they act is $\dim_E V$.

If $k_E = k_E^\circ$ then we need to be a little more careful with the polynomials $X \pm 1$ (that is, the k_E/k_E°-self-dual monic polynomials of degree 1), as described at the end of the previous subsection. If one of the $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ is *even* special orthogonal (and the other symplectic) then we get that $\sum_\rho \lfloor s_\pi(\rho)^2\rfloor \deg\rho$ is

$$[E:F]\Bigl(\sum_{m\geq 2}\Bigl(m\sum_{\deg Q=m}(a^{(0)}_{\sigma_m^{(0)}(Q)} + a^{(1)}_{\sigma_m^{(1)}(Q)})\Bigr) + a_+^{(0)} + a_-^{(0)} + a_+^{(1)} + a_-^{(1)} - 1\Bigr)$$
$$= [E:F]\Bigl(\sum_P a_P^{(0)}\deg P + \sum_P a_P^{(1)}\deg P - 1\Bigr) = [E:F]\dim_E V = 2N,$$

where the penultimate equality uses the fact that the dual of a symplectic group acts naturally on a space of one dimension greater, while the dual of an even special orthogonal group acts naturally on a space of the same dimension.

On the other hand, if one of the $\mathcal{G}^{(t)}(\Lambda_{\mathfrak{o}_E})$ is *odd* special orthogonal (and the other symplectic) then we get the same sum except without the term -1, and the penultimate equality uses the fact that the dual of an odd special orthogonal group acts naturally on a space of one dimension smaller, while the dual of a symplectic group acts naturally on a space of one dimension greater.

This completes the proof of (5-1), and hence that of 2.5 Theorem.

5.10. The results in this section not only prove 2.5 Theorem but also give an algorithm to compute the inertial Jordan set of a positive-depth simple cuspidal representation of G. (The case of depth zero is treated already in [Lust and Stevens 2016].) Moreover, 2.6 Corollary then gives the inertial Jordan set for any cuspidal representation of G.

Indeed, suppose π is a simple cuspidal representation of G, induced from a cuspidal type $\lambda = \underline{\kappa}\otimes\tau$. With the usual notation, let τ° be any irreducible component of the restriction of τ to the maximal parahoric subgroup $P^\circ(\Lambda_{\mathfrak{o}_E})$. Then τ° is the inflation of a representation $\tau^{(0)}\otimes\tau^{(1)}$, with each $\tau^{(t)}$ a cuspidal representation of a finite reductive group over k_E°. These each appear in some rational Lusztig series and we consider the set $\mathcal{Q}^{(t)}$ of monic irreducible polynomials dividing the characteristic polynomial (over k_E) of the corresponding semisimple conjugacy class for $t = 0, 1$, all of which are k_E/k_E°-self-dual.

For each $m \in \deg \mathcal{Q}^{(t)}$, we compute the signature character $\epsilon_{\mathfrak{M}_t}$, and thus deduce the involution $\sigma_m^{(t)}$ as in Section 5.5. We set

$$\mathcal{Q} = \{\sigma_m^{(t)}(Q) \mid Q \in \mathcal{Q}^{(t)},\ \deg Q = m,\ t = 0, 1\}.$$

Now let Θ be the endoclass of the self-dual simple character lifting any skew simple character in π and let $Q \in \mathcal{Q}$. We put $n = \deg Q \deg \Theta$ and let $\tilde{\theta}$ be the unique (up to conjugacy) *m-simple* character in $\mathrm{GL}_n(F)$ with endoclass Θ^2 (in the language of [Bushnell and Henniart 2014], for example). Let $\underline{\tilde{\kappa}}$ be the p-primary extension of $\tilde{\theta}$, a representation of a group $\tilde{J}$. The group $\tilde{J}/\tilde{J}^1$ is then a finite general linear group of rank $\deg Q$ over k_E, and we let $\tilde{\tau}_Q$ be the unique cuspidal representation in the Lusztig series corresponding to a semisimple conjugacy class with characteristic polynomial Q. Write $[\rho_Q]$ for the inertial class of cuspidal representations of $\mathrm{GL}_n(F)$ containing $\underline{\tilde{\kappa}} \otimes \tilde{\tau}_Q$.

The inertial classes in $\{[\rho_Q] \mid Q \in \mathcal{Q}\}$ are precisely the inertial classes which will appear in $\mathrm{IJord}(\pi)$. In order to compute the multiplicities with which $[\rho_Q]$ appears, we follow the recipe of Section 5.6 to compute the corresponding Hecke algebra parameters r_0 and r_1, and hence the real parts of the reducibility points $|r_0 \pm r_1|/(\deg Q[k_E : k_F])$ and the multiplicities from Mœglin's criterion. In the case that $k_E \neq k_E^\circ$ or $m = \deg Q > 1$, this is straightforward, with the real parts of the reducibility points given by

$$\frac{b^{(0)}_{\sigma_m^{(0)}(Q)} + b^{(1)}_{\sigma_m^{(1)}(Q)} + 1}{2} \quad \text{and} \quad \frac{b^{(0)}_{\sigma_m^{(0)}(Q)} - b^{(1)}_{\sigma_m^{(1)}(Q)}}{2},$$

where $a_P^{(t)} = \frac{1}{2} b_P^{(t)}(b_P^{(t)} + 1)$ is the power to which P divides the characteristic polynomial corresponding to $\tau^{(t)}$. By the construction of $\mathcal{Q}$, the first of these is certainly greater than $\frac{1}{2}$. In the case $k_E = k_E^\circ$ and $\deg Q = 1$ (so that Q is $X \pm 1$) there is no such simple universal formula, and instead one must proceed in a case-by-case analysis as in Section 5.8. We leave this as an exercise to the reader; a similar calculation is done in [Lust and Stevens 2016, §8].

6. Galois parameters

In this section we study self-duality in terms of Galois parameters with a view, in particular, to understanding the ambiguities in our results in terms of the local Langlands correspondence.

6.1. We denote by $\bar{F}$ a fixed separable closure of F and by W_F the absolute Weil group of F (with similar notation for intermediate fields). We would like to explore a self-dual irreducible representation σ of W_F, with a view to determining its parity (that is, whether it is symplectic or orthogonal); in particular, we would like to know when the self-dual irreducible representation σ' which is an unramified twist of (and not isomorphic to) σ has the same parity as σ, since it is in this case that we have ambiguity. For now, we do not require p to be odd.

Let χ be an unramified character of W_F. Then $(\chi\sigma)^\vee$ is isomorphic to $\chi^{-1}\sigma^\vee$, so, σ being self-dual, $\chi\sigma$ is self-dual if and only if $\chi^2\sigma \simeq \sigma$.

We let $t(\sigma)$ be the number of unramified characters η of W_F such that $\eta\sigma \simeq \sigma$ — such characters form a cyclic group. We deduce that the only unramified character twist σ' of σ which is self-dual but not isomorphic to σ is obtained as $\chi\sigma$, where χ is an unramified character of order $2t(\sigma)$. (If $r = \mathrm{val}_2(t(\sigma))$, any unramified character of order 2^{r+1} would do equally well.)

Let E be the unramified extension of F in $\overline{F}$ of degree $t(\sigma)$. Then σ is induced from a representation τ of W_E; the restriction of σ to W_E is the direct sum of the conjugates of τ under $\mathrm{Gal}(E/F)$, which are pairwise inequivalent. As σ is self-dual, $\tau^\vee$ is one of those conjugates.

Assume first that τ is self-dual — which, we remark, is necessarily true if $t(\sigma)$ is odd. Since $t(\tau) = 1$, the unramified twist τ' of τ which is self-dual but not isomorphic to τ has the form $\chi\tau$, where χ is the order-2 unramified character of W_E, and it has the same parity as τ. Since induction for self-dual representations preserves the parity, we deduce that σ and σ' share the same parity too.

Assume then that τ is not self-dual. Then $\tau^\vee$ is necessarily isomorphic to τ^γ, where γ is the order-2 element of $\mathrm{Gal}(E/F)$. Let $\widetilde{E} = E^\gamma$, so that $E/\widetilde{E}$ is quadratic, and let T be the (irreducible) representation of $W_{\widetilde{E}}$ induced from τ. As $\tau^\vee \simeq \tau^\gamma$, we see that T is self-dual. Its restriction to W_E is $\tau \oplus \tau^\vee$, with τ not isomorphic to $\tau^\vee$, so the W_E-invariant bilinear forms on the space of T form a space of dimension 2, with a line of alternating forms and a line of symmetric ones. Each of these lines is invariant under $\mathrm{Gal}(E/F)$, one offering the trivial representation, the other the order-2 character ω of $\mathrm{Gal}(E/F)$. The self-dual unramified twist T' of T which is not isomorphic to T is $T' = \eta T$ where η is unramified of order 4, so that $T' \otimes T' \simeq \omega T \otimes T$. From the previous analysis, we deduce that if T is symplectic then T' is orthogonal and conversely, T and T' have different parities. By induction again we see that σ and σ' have different parities.

6.2. Let us look at some special cases. Assume first that σ is tame. Then $t(\sigma) = \dim\sigma$. Introducing E and τ as in Section 6.1, we have that τ is a character, regular under the action of $\mathrm{Gal}(E/F)$. If τ were self-dual it would have order 1 or 2, but any character of $E^\times$ of order 1 or 2 factors through $N_{E/F}$, and hence can be regular under the action of $\mathrm{Gal}(E/F)$ only if $t(\sigma) = \dim\sigma = 1$, so $E = F$. Thus, apart from quadratic characters of W_F, *tame* self-dual irreducible representations σ of W_F have even dimension, and we can apply the discussion of Section 6.1 to them, concluding that σ and σ' have different parities.

6.3. We now assume that σ is *not tame*, but we concentrate on our case of interest; that is, we assume from now on that *p is odd.* We want in that case to spot when σ and σ' have the same parity, and then try to say whether they are orthogonal or symplectic.

Let us first analyze σ. Its restriction to the wild ramification subgroup $\mathscr{P}_F$ of W_F is nontrivial, since σ is not tame. Let γ be an irreducible component of this restriction — so that γ is not the trivial character of $\mathscr{P}_F$ — and $S = S_\gamma$ its stabilizer in W_F. Then by Clifford theory σ is induced from the representation of S on the isotypical component $\mathcal{V}(\gamma)$ of γ in the space $\mathcal{V}$ of σ.

Now by assumption σ is self-dual, and so is its restriction to $\mathscr{P}_F$. But $\mathscr{P}_F$ is a pro-p-group and p is odd, so no nontrivial irreducible representation of $\mathscr{P}_F$ is self-dual, and we see that $\gamma^\vee$ is not isomorphic to γ. Thus there is g in $W_F \setminus S$ with ${}^g\gamma$ isomorphic to $\gamma^\vee$; the coset gS is the same for all possible choices of g, and g^2 belongs to S, so $\widetilde{S} = S \cup gS$ is a subgroup of W_F containing S as an index-2 subgroup.

To get σ, we can first induce $\mathcal{V}(\gamma)$ from S to $\widetilde{S}$, and then from $\widetilde{S}$ to W_F. We shall prove now that $\mathrm{Ind}_S^{\widetilde{S}}\,\mathcal{V}(\gamma)$ is self-dual; its parity is then inherited by σ. This reduces the problem to understanding the parity of $\mathrm{Ind}_S^{\widetilde{S}}\,\mathcal{V}(\gamma)$.

6.4. To prove that $\mathrm{Ind}_S^{\widetilde{S}}\,\mathcal{V}(\gamma)$ is self-dual, we take an abstract viewpoint:

Proposition. *Let $\mathscr{G}$ be a group with a subgroup $\mathscr{H}$ of index* 2, *and let $g \in \mathscr{G} \setminus \mathscr{H}$. Let $(\rho, \mathcal{V})$ be an irreducible representation of $\mathscr{H}$. Assume that ρ is not self-dual, but that $\rho^\vee$ is equivalent to ${}^g\rho$. Then $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,\rho$ is irreducible and self-dual. If $\dim \mathcal{V}$ is odd, then $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,\rho$ is symplectic if and only if its determinant is trivial.*

Proof. Since ${}^g\rho$ is not isomorphic to ρ, the induced representation $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\rho$ is irreducible, and it is self-dual because $(\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\rho)^\vee$ is isomorphic to $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\rho^\vee$ and hence to $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,{}^g\rho$, which is isomorphic to $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\rho$. If $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\rho$ is symplectic, then clearly its determinant is trivial. To prove the converse statement when $\dim \mathcal{V}$ is odd, we need to analyze the situation carefully.

Since $\rho^\vee$ is equivalent to ${}^g\rho$, there is a nondegenerate bilinear form $\Phi : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ such that

$$\Phi(hv, ghg^{-1}v') = \Phi(v, v') \quad \text{for all } h \in \mathscr{H},\ v, v' \in \mathcal{V}.$$

It is unique up to scalar. We claim that the form Ψ, defined by $\Psi(v, v') = \Phi(v', g^2 v)$ for v, v' in $\mathcal{V}$, is proportional to Φ. Indeed, for $v, v' \in \mathcal{V}$ and $h \in \mathscr{H}$, we find

$$\Psi(hv, ghg^{-1}v') = \Phi(ghg^{-1}v', g^2hv) = \Phi(v', g^2v) = \Psi(v, v').$$

Writing $\Psi = \lambda\Phi$ with $\lambda \in \mathbb{C}^\times$, we compute

$$\Phi(v', v) = \Phi(g^2v', g^2v) = \Psi(v, g^2v') = \lambda\Phi(v, g^2v') = \lambda\Psi(v', v) = \lambda^2\Phi(v', v)$$

for v, v' in $\mathcal{V}$ so that $\lambda^2 = 1$. We shall see that the parity of $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,\rho$ is governed by the scalar λ.

On the space $\mathcal{V}\oplus\mathcal{V}$ equipped with the representation $\rho \oplus {}^g\rho$, there is an $\mathscr{H}$-invariant symplectic form f, unique up to scalar, which we can take to be

$$f : ((v_1, v_2), (w_1, w_2)) \mapsto \Phi(v_1, w_2) - \Phi(w_1, v_2).$$

The space of $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,\rho$ can be taken as $\mathcal{V}\oplus\mathcal{V}$, where $\mathscr{H}$ acts as $\rho \oplus {}^g\rho$ and g acts via

$$g(v_1, v_2) = (v_2, g^2v_1).$$

Since $\Phi(v_2, g^2v_1) = \Psi(v_1, v_2) = \lambda\Phi(v_1, v_2)$, we get that g acts on f by multiplication by $-\lambda$, so $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}}\,\rho$ is symplectic if and only if $\lambda = -1$.

Let us choose a basis $(e_1, \dots, e_d)$ of $\mathcal{V}$, where $d = \dim\rho$. Then $\Phi(v_1, v_2) = ({}^t\boldsymbol{x}_1)H\boldsymbol{x}_2$ for $v_1, v_2 \in \mathcal{V}$ with coordinates given by $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^d$ respectively, and H the $d \times d$ Gram matrix of Φ in the basis. If M_{g^2} is the matrix of $\rho(g^2)$ we get $\Phi(v_2, g^2v_1) = ({}^t\boldsymbol{x}_2)HM_{g^2}\boldsymbol{x}_1$, from which we deduce that $HM_{g^2} = \lambda({}^tH)$, which implies that $\det\rho(g^2) = \lambda^d$.

Now $\det(\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}} \rho)$ is an order-2 character of $\mathscr{G}$ which is trivial on $\mathscr{H}$; in fact it is given by

$$(\det \rho \circ \mathrm{Ver})\, \omega^{\dim \rho},$$

where $\mathrm{Ver} : \mathscr{G} \mapsto \mathscr{G}^{\mathrm{ab}} \mapsto \mathscr{H}^{\mathrm{ab}}$ is the transfer and ω is the nontrivial character of $\mathscr{G}$ trivial on $\mathscr{H}$. In this special case where $\mathscr{H}$ has index 2 in $\mathscr{G}$, the transfer map Ver is trivial on $\mathscr{H}$ and sends g to g^2, so $\det(\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}} \rho(g)) = (-\lambda)^d$.

When d is odd, we find that $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}} \rho$ is symplectic if and only if its determinant is trivial, as desired. □

Remark. When d is even, $\mathrm{Ind}_{\mathscr{H}}^{\mathscr{G}} \rho$ always has trivial determinant, regardless of its parity. Determining the parity amounts to computing the scalar λ.

6.5. We revert to the context of Sections 6.1–6.3. We want to spot the cases where σ and σ' (in the notation of Section 6.1) have the same parity, and in those cases possibly apply 6.4 Proposition to determine that parity. For that we have to analyze the situation further.

It is known (see [Bushnell and Henniart 2014, 1.3, Proposition]) that γ extends to a representation Γ of $S = S_\gamma$, and we can even impose that $\det \Gamma$ have order a power of p; then Γ is unique up to twist by an unramified character of S, of order a power of p. Since ${}^g\gamma$ is equivalent to $\gamma^\vee$, we see that ${}^g\Gamma$ is equivalent to $\chi\Gamma^\vee$, where χ is an unramified character of S of order a power of p. Such a χ has a unique square root η with order a power of p and replacing Γ with $\eta^{-1}\Gamma$, we may — and do — assume that ${}^g\Gamma \simeq \Gamma^\vee$. This now specifies Γ completely.

As a representation of S, the space $\mathcal{V}(\gamma)$ is a tensor product $\Gamma \otimes \delta$, where δ is an irreducible representation of S trivial on $\mathscr{P}_F$, well-defined up to isomorphism. Since ${}^g\mathcal{V}(\gamma) \simeq \mathcal{V}(\gamma)^\vee$ as representations of S, we get that ${}^g\delta \simeq \delta^\vee$.

Let K be the fixed field of S, and $\widetilde{K}$ that of $\widetilde{S}$; thus the extension $K/\widetilde{K}$ is quadratic and, in particular, tame. Writing $d = \dim \delta$, the representation δ is induced from a character α of the unramified degree-d extension K_d of K in $\overline{F}$, with α tamely ramified and regular under the action of $\mathrm{Gal}(K_d/K)$; this character α is determined up to the action of $\mathrm{Gal}(K_d/K)$.

In those terms, we try to see when σ and σ' have the same parity; that is, writing $\sigma = \mathrm{Ind}\,\tau$ as in Section 6.1, where τ is a representation of W_E with E/F unramified of degree $t(\sigma)$, we want to know if τ is self-dual. Note that $t(\mathcal{V}(\gamma)) = d$, so $t(\sigma) = df(K/F)$, where $f(K/F)$ is the inertia degree of K/F. The extension K_d/E is totally tamely ramified, and we can take τ to be $\mathrm{Ind}(\Gamma \otimes \alpha)$, where the induction is from W_{K_d} to W_E (and we first restrict Γ from S to W_{K_d}).

6.6. The following result describes when σ, σ' have the same parity.

Proposition. *Let σ be a self-dual irreducible representation of W_F. Assume σ is not tame, and adopt the above notation. Then the following are equivalent*:

(i) *σ and σ' have the same parity.*

(ii) *$K/\widetilde{K}$ is ramified and $d = 1$.*

When these conditions are satisfied, σ and σ' are symplectic if and only if the character α is ramified.

Remark. When $d=1$, we see that α is a tame character of $K^\times$ which satisfies ${}^g\alpha=\alpha^{-1}$. If $K/\widetilde{K}$ is ramified, g acts trivially on the residue field of K, and $\alpha_{|U_K}$ has order 1 or 2. In that case, let ϖ be a uniformizer of K with $\varpi^2\in\widetilde{K}$; then the condition ${}^g\alpha=\alpha^{-1}$ translates into $\alpha(-\varpi^2)=1$: either α is unramified of order 1 or 2 or $\alpha_{|\widetilde{K}^\times}$ is the quadratic character $\omega_{K/\widetilde{K}}$ defining K.

Proof. To prove the proposition, we need to see when $\tau=\mathrm{Ind}_{W_{K_d}}^{W_E}(\Gamma\otimes\alpha)$ is self-dual. The restriction of $\Gamma\otimes\alpha$ to $\mathscr{P}_F$ is γ, so τ can be self-dual only if there is h in W_E such that ${}^h\gamma\simeq{}^g\gamma$ — that is $h\in gS$, or equivalently $W_E\cap S\neq W_E\cap\widetilde{S}$. Recalling that E is the maximal unramified extension of F in K_d, we see that the fixed field of $W_E\cap S$ is K_d; if $K/\widetilde{K}$ were unramified, the fixed field of $W_E\cap\widetilde{S}$ would also be K_d, so that τ could not be self-dual.

Thus, if τ is self-dual then $K/\widetilde{K}$ is ramified and we take g in $W_E\cap\widetilde{S}$. Reasoning as in Section 6.3 and using 6.4 Proposition, we see that τ is self-dual if and only if $\Gamma\otimes\alpha$ induces to a self-dual representation of $W_{\widetilde{K}_d}$, where $\widetilde{K}_d$ is the fixed field of g in K_d (so that $K_d/\widetilde{K}_d$ is quadratic ramified); in particular we then have ${}^g(\Gamma\otimes\alpha)\simeq(\Gamma\otimes\alpha)^\vee$. Since ${}^g\Gamma\simeq\Gamma^\vee$ by construction, this implies ${}^g\alpha=\alpha^{-1}$ and since $K_d/\widetilde{K}_d$ is ramified, g acts trivially on the residue field of K_d so $\alpha_{|U_{K_d}}$ has order 1 or 2 and regularity with respect to $\mathrm{Gal}(K/K_d)$ implies $d=1$. Thus if τ is self-dual then $d=1$, which proves (i) $\Rightarrow$ (ii).

Conversely if (ii) is satisfied then τ is self-dual if and only if ${}^g\alpha=\alpha^{-1}$ by the above analysis, which gives (ii) $\Rightarrow$ (i).

Assume finally that conditions (i) and (ii) are satisfied. Using again 6.4 Proposition, we have to check whether the determinant of $\mathrm{Ind}_{W_K}^{W_{\widetilde{K}}}(\Gamma\otimes\alpha)$ — which by self-duality has order 1 or 2 — is trivial. Seeing that determinant as a character of $\widetilde{K}^\times$ (via class field theory), it is equal to

$$\nu=\det(\Gamma\otimes\alpha)_{|\widetilde{K}^\times}(\omega_{K/\widetilde{K}})^{\dim\gamma}.$$

But $\det\Gamma$ has order a power of p and p is odd, and α has order at most 4 (cf. the remark above) so we find $\nu=(\alpha_{|\widetilde{K}^\times}\omega_{K/\widetilde{K}})^{\dim\gamma}$. If α is unramified then $\nu=\omega_{K/\widetilde{K}}$ (since $\dim\gamma$ is odd) is nontrivial; if α is ramified then $\alpha_{|\widetilde{K}^\times}=\omega_{K/\widetilde{K}}$ by the remark and ν is trivial. The final claim of the proposition now follows from 6.4 Proposition. $\square$

6.7. Now we interpret the conditions of 6.4 Proposition in terms of the cuspidal representation ρ of $\mathrm{GL}_n(F)$, with $n=\dim\sigma$, which corresponds to σ under the Langlands correspondence. To describe this representation ρ we will use the machinery of the construction of cuspidal representations as in Section 1.

Assume σ is not tame, i.e., ρ is not of depth zero. Then ρ contains a simple character $\tilde{\theta}$, belonging to a set of simple characters built using an element $\beta\in\mathrm{GL}_n(F)$ which generates a field $F[\beta]$. We have $n=d[F[\beta]:F]$, so that d is determined by ρ. Moreover the extension K/F which appears above in the discussion on the construction of σ is isomorphic to the maximal tame subextension L/F of $F[\beta]/F$; see [Bushnell and Henniart 2014, tame parameter theorem].

When ρ — equivalently σ — is self-dual, we can choose β such that the self duality comes from an automorphism $x\mapsto\bar{x}$ of $F[\beta]$, sending β to $-\beta$, and $\tilde{\theta}$ to $\tilde{\theta}^{-1}$; see [Blondel 2004, Theorem 1]. That automorphism induces an order-2 automorphism of L; let $\widetilde{L}$ be its fixed field.

Proposition. *The extensions $K/\widetilde{K}$ and $L/\widetilde{L}$ are isomorphic.*

Thus condition (ii) in 6.6 Proposition can be translated in terms of ρ. See below (Section 6.8) for a translation of the last assertion of [Blondel 2004].

Proof. The proof relies on the compatibility of tame lifting of simple characters with the induction process for Weil group representations [Bushnell and Henniart 1996; 2005; 2014]. Choose an isomorphism ι of K/F onto L/F.

The representation σ_K of W_K on $\mathcal{V}(\sigma)$ corresponds to a (cuspidal) representation ρ_L of $\mathrm{GL}_m(L)$, where $m = d[F[\beta] : L]$; the simple character $\tilde{\theta}_L$ appearing in ρ_L is an L/F-lift of $\tilde{\theta}$ and L/F is the maximal tame extension such that $\tilde{\theta}$ has a lift to $\mathrm{GL}_{[F[\beta]:L]}(L)$.

If K' is intermediate between F and K, and $L' = \iota(K')$, then $\sigma_{K'} = \mathrm{Ind}_{W_K}^{W_{K'}} \mathcal{V}(\sigma)$ corresponds to a (cuspidal) representation $\rho_{L'}$ of $\mathrm{GL}_{m'}(L')$, with $m' = d[F[\beta] : L']$, and the simple character $\tilde{\theta}_{L'}$ appearing in $\rho_{L'}$ is an L'/F-lift of $\tilde{\theta}$ and lifts to $\tilde{\theta}_L$ in L/L'. But $\widetilde{K}$ is the maximal intermediate subfield K' such that $\sigma_{K'}$ is self-dual. Because the Langlands correspondence is compatible with taking contragredients, the field $\iota(\widetilde{K})$ is the maximal field L' intermediate between F and L such that $\tilde{\theta}_{L'}$ is self-dual (i.e., conjugate to $\tilde{\theta}_{L'}^{-1}$ in $\mathrm{GL}_{[F[\beta]:L']}(L')$). Thus $\iota(\widetilde{K}) = \widetilde{L}$. □

6.8. Now assume that $d = 1$ and $L/\widetilde{L}$ (or equivalently $K/\widetilde{K}$) is ramified. We want to express the condition that α is ramified in 6.6 Proposition in terms of ρ. For that we have to review a little bit the construction of ρ from $\tilde{\theta}$ from Section 1, whose notation we use.

We also continue with the notation for ρ introduced in the previous subsection. Recall that, since $d = 1$, we have $n = [F[\beta] : F]$. The simple character $\tilde{\theta}$ is a character of $\widetilde{H}^1$ and we have the open subgroups $\tilde{J}^1$, $\tilde{J}$ of $\mathrm{GL}_n(F)$. We write $\tilde{\eta}$ for the unique irreducible representation of $\tilde{J}^1$ containing $\tilde{\theta}$. Then $\tilde{J} = U_{F[\beta]}\tilde{J}^1$ and, by the types theorem [Bushnell and Henniart 2014, §7.6, Theorem], there is a unique beta-extension $\tilde{\kappa}$ such that $\mathrm{tr}\,\tilde{\kappa}$ is constant on the roots of unity of $F[\beta]$ of order prime to p which are regular for the action of $\mathrm{Gal}(K_{nr}/F)$, where K_{nr} is the maximal unramified extension of F in K. Moreover, the same result gives that ρ contains the representation $\tilde{\kappa} \otimes \omega\alpha$ of J, where α is seen as a character of $J/J^1 \simeq U_{F[\beta]}/U^1_{F[\beta]} \simeq U_K/U^1_K$ and ω is the order-2 character of $U_{F[\beta]}$.

Thus we conclude that ρ is symplectic when ρ contains $\tilde{\kappa}$, and is orthogonal when ρ contains $\omega\tilde{\kappa}$.

6.9. We have discussed at length above the ambiguity between σ and σ' inherent to our method — of course when σ and σ' have different parities it is the orthogonal one that features.

Let us now briefly mention a few favourable circumstances when our methods do allow us to determine completely the parameter of a cuspidal representation π of $\mathrm{Sp}_{2N}(F)$.

Since the parameter ϕ of π is orthogonal of dimension $2N + 1$, one irreducible component must have odd dimension. But in our case where p is odd, the only irreducible orthogonal representations of W_F with odd dimension are the four quadratic characters of W_F. Thus at least one of them, say ω, has to occur in the parameter, and if the Jordan block it belongs to is (ω, m) then m has to be congruent to $1 \pmod 4$, to yield an odd-dimensional contribution to ϕ; the contribution to the determinant is then ω. We then see that if we know all other components, then we can decide between ω and ω' by taking into account the condition $\det\phi = 1$. To know all the other components σ, it is necessary that for each of them, σ and σ' have different parities. We conclude that it will be rather rare that we determine ϕ without ambiguity.

Let us give just a few examples in low dimension. See [Lust and Stevens 2016] for a discussion of depth-zero cases.

$N = 1$, $\mathrm{SL}_2(F)$: The parameter is either $\rho\oplus\omega$ with ρ irreducible orthogonal of dimension 2 and $\omega = \det\rho$, or $\omega_1 \oplus \omega_2 \oplus \omega_3$ where the ω_i's are the nontrivial quadratic characters of W_F. In terms of homomorphisms $W_F \to \mathrm{SO}_3(\mathbb{C}) \simeq \mathrm{PGL}_2(\mathbb{C})$, the second case corresponds to a triply imprimitive representation of W_F, and the first case to a simply imprimitive one [Bushnell and Henniart 2006]. In the first case, our methods allow us to determine ρ only if it is induced from the quadratic unramified extension of F (i.e., in fact, when ω is unramified of order 2).

$N = 2$, $\mathrm{Sp}_4(F)$: There has to be a quadratic character ω of W_F occurring with Jordan block $(\omega, 1)$ only. If another quadratic character η occurs, the Jordan block can be $(\eta, 1)$ or $(\eta, 3)$. In the latter case $\phi = \omega \oplus \eta \oplus \eta \otimes \mathrm{St}_3$ and the determinant condition implies that ω is trivial and consequently that η is not trivial. If our computation shows that both 1 and the nontrivial quadratic unramified character ω_{nr} occur, then the parameter is necessarily $\phi = 1 \oplus \omega_{nr} \oplus \omega_{nr} \otimes \mathrm{St}_3$; if, on the contrary, our method gives that a ramified quadratic character η occurs, then we cannot distinguish between η and $\eta' = \eta\omega_{nr}$.

Let us look at the case where two distinct characters ω, η occur with Jordan blocks $(\omega, 1)$ and $(\eta, 1)$ only. Then a third character, ν say, must also occur and $\phi = \omega \oplus \eta \oplus \nu \oplus \rho$, where ρ is irreducible orthogonal of dimension 2. The determinant of ρ is the quadratic character $\omega_{E/F}$ defining the extension from which ρ is induced so that the determinant condition on ϕ is $\omega\eta\nu\omega_{E/F} = 1$.

When E/F is unramified, there is no ambiguity in ρ in our computation, and the parameter is

$$\phi = 1 \oplus \mu \oplus \mu' \oplus \rho,$$

where μ, μ' are the two ramified quadratic characters of W_F.

When E/F is ramified, the parameter could be

$$\phi = 1 \oplus \omega_{nr} \oplus \omega_{nr}\omega_{E/F} \oplus \rho \quad \text{or} \quad \phi = 1 \oplus \omega_{nr} \oplus \omega_{nr}\omega_{E/F} \oplus \rho'$$

and we cannot resolve the ambiguity between ρ and ρ'.

Finally if there is only one quadratic character ω of W_F occurring in ϕ, we can compute ω, and thus determine ϕ completely, only if the other components (necessarily even-dimensional) offer no ambiguity.

We hope to come back to the case of $\mathrm{Sp}_4(F)$ in a sequel to this paper, where a refinement of our methods will allow a more complete determination of ϕ.

7. Langlands correspondence and ramification

In this final section we interpret our results on the endoscopic transfer map in terms of the Langlands correspondence for G. In particular, we prove a ramification theorem for the symplectic group G, giving a bijection between self-dual endoclasses and self-dual orbits of irreducible representations of the wild inertia group $\mathscr{P}_F$ which is simultaneously compatible (in a suitable sense) with the Langlands correspondence for symplectic groups over F in all dimensions.

7.1. We first recall the ramification theorem for general linear groups, from [Bushnell and Henniart 2003, 8.2, Theorem]; see also [Bushnell and Henniart 2014, 6.3, Theorem]. Recall that $\mathcal{E}(F)$ denotes the set of endoclasses over F. We write $W_F\backslash\operatorname{Irr}(\mathscr{P}_F)$ for the set of W_F-orbits of irreducible representations of $\mathscr{P}_F$. By abuse of notation, we will identify such an orbit with the direct sum of the inequivalent irreducible representations in the orbit; thus, for γ an irreducible representation of $\mathscr{P}_F$ with stabilizer S, we identify its W_F-orbit $[\gamma]$ with $\bigoplus_{W_F/S} {}^g\gamma$. In particular, we can then talk of the dimension of an orbit.

Given an irreducible representation of W_F, by Mackey theory its restriction to $\mathscr{P}_F$ is a multiple of a single W_F-orbit of irreducible representations, so we get a natural map $\operatorname{Irr}(W_F)\to W_F\backslash\operatorname{Irr}(\mathscr{P}_F)$, which is surjective.

Theorem. *There is a unique bijection* $\mathcal{E}(F)\to W_F\backslash\operatorname{Irr}(\mathscr{P}_F)$, $\Theta\mapsto[\gamma(\Theta)]$, *which is compatible with the local Langlands correspondence*:

$$\begin{array}{ccc}
\bigcup_{n\ge1}\operatorname{Cusp}(\mathrm{GL}_n(F)) & \xrightarrow{\sim} & \operatorname{Irr}(W_F)\\
\downarrow & & \downarrow\\
\mathcal{E}(F) & \xrightarrow{\quad\sim\quad} & W_F\backslash\operatorname{Irr}(\mathscr{P}_F)
\end{array}$$

Moreover we have $\deg\Theta=\dim[\gamma(\Theta)]$.

7.2. Now we consider how this bijection behaves with respect to duality. Recall that we write $\mathcal{E}^{\mathrm{sd}}(F)$ for the set of self-dual endoclasses, that is, those endoclasses Θ for which there is a self-dual simple character $\tilde\theta$ with endoclass Θ. If the endoclass is nontrivial then $\tilde\theta$ is associated to a skew simple stratum $[\Lambda,-,0,\beta]$ and the associated field $E=F[\beta]$ has degree n over F and is equipped with a Galois involution with fixed field $E_\circ$. If Θ is the trivial endoclass then we have $E=E_\circ=F$.

It will be useful to have the following result, which guarantees the existence of self-dual cuspidal representations of general linear groups with given (self-dual) endoclass.

Lemma. *Let Θ be a self-dual endoclass and $E/E_\circ$ as above. Let m be an integer which is*

(i) *odd if $E/E_\circ$ is unramified quadratic*,

(ii) 1 *or even if $E/E_\circ$ is ramified quadratic*,

(iii) *even if $E=F$*,

and put $n=m\deg\Theta$. *Then there are* (*at least*) *two inequivalent orthogonal self-dual cuspidal representations of* $\mathrm{GL}_n(F)$ *with endoclass* Θ, *and two inequivalent symplectic self-dual cuspidal representations of* $\mathrm{GL}_n(F)$ *with endoclass* Θ.

Note that, in the case that $E = F$ (so Θ is trivial) and $m = 1$, there are four inequivalent self-dual (cuspidal) representations of $\mathrm{GL}_1(F)$ with endoclass Θ but all four are orthogonal; they are the four quadratic characters.

Proof. Suppose first that Θ is nontrivial. Let $\tilde\theta$ be a self-dual simple character with endoclass Θ, as above, with associated skew simple stratum $[\Lambda, -, 0, \beta]$ and $E = F[\beta]$. Then any transfer (in the sense of simple characters) of $\tilde\theta$ is also self-dual, by [Stevens 2005, Corollary 2.13].

Let m be an integer as in the hypotheses of the lemma and let f be a nondegenerate skew-hermitian form on an m-dimensional E-vector space V such that the associated unitary group (a group over $E_\circ$) is quasisplit. We write $\mathfrak{o}_E^\circ$ for the ring of integers of $E_\circ$ and $\mathfrak{p}_E^\circ$ for its unique maximal ideal, with k_E° the residue field. Fix $\lambda_\circ$ an F-linear form on $E_\circ$ such that $\{e \in E_\circ \mid \lambda_\circ(e\mathfrak{o}_E^\circ) \subseteq \mathfrak{p}_F\} = \mathfrak{p}_E^\circ$, and consider the form $h = \lambda_\circ \circ \mathrm{tr}_{E/E_\circ} \circ f$ on V. Thinking of V as an n-dimensional F-vector space, this is a nondegenerate alternating form. We take the transfer $\tilde\theta_m$ of $\tilde\theta$ to the unique (up to conjugacy) self-dual $\mathfrak{o}_E$-lattice chain Λ_m on V such that $\Lambda_m(0) \neq \Lambda_m(1)$. Thus $\tilde\theta$ is a self-dual simple character of endoclass Θ.

Denote by $\tilde\kappa_m$ the unique p-primary extension of $\tilde\theta_m$, and denote by $\tilde J_m$ the group on which it lives; then, by uniqueness, $\tilde\kappa_m$ is self-dual (that is, invariant under the involution σ defining the symplectic group $\mathrm{Sp}_F(V)$). Now $\tilde\kappa_m$ extends to a representation $\widetilde{\mathcal{K}}_m$ of $E^\times \tilde J_m$ with determinant a power of p and any two such extensions differ by an unramified character of order a power of p. In particular, $\widetilde{\mathcal{K}}_m \circ \sigma$ is another such extension and so has the form $\widetilde{\mathcal{K}}_m \otimes \chi$ for χ unramified of order a power of p. Since p is odd, χ has a unique square root χ' of order a power of p, and then we can replace $\widetilde{\mathcal{K}}_m$ by $\widetilde{\mathcal{K}}_m \otimes \chi'$, which is self-dual.

Now we consider the quotient

$$\tilde J_m / \tilde J_m^1 \simeq \widetilde P(\Lambda_{m,\mathfrak{o}_E}) / \widetilde P^1(\Lambda_{m,\mathfrak{o}_E}) \simeq \mathrm{GL}_m(k_E).$$

The involution σ also acts here, with fixed points a unitary group if $E/E_\circ$ is unramified and a symplectic group if $E/E_\circ$ is ramified (in the latter case, it is symplectic rather than orthogonal because $\Lambda_m(0) \neq \Lambda_m(1)$); the action of σ is conjugate to the map transpose-inverse-$\mathrm{Gal}(k_E/k_E^\circ)$-conjugate. The conditions on m are then precisely those required for the existence of a $\mathrm{Gal}(k_E/k_E^\circ)$-self-dual cuspidal representation $\tilde\tau$ of $\mathrm{GL}_m(k_E)$ (that is, such that the Galois conjugate of $\tilde\tau$ is equivalent to $\tilde\tau^\vee$) — see [Adler 1997, Theorem 7.1] in the case $k_E = k_E^\circ$ and [Kariyama 2008, Corollary 5.8] in the case $k_E \neq k_E^\circ$.

Let ω be a quadratic character of E, necessarily tame since p is odd. We also write ω for the character of $k_E^\times$ induced by restricting ω; then the representation $\tilde\tau\omega$ is also $\mathrm{Gal}(k_E/k_E^\circ)$-self-dual. We inflate $\tilde\tau\omega$ to $\tilde J_m$ and extend to a representation $\widetilde{\mathcal{T}}_\omega$ of $E^\times \tilde J_m$ by setting $\widetilde{\mathcal{T}}_\omega(\varpi_E) = \omega(\varpi_E)\mathrm{Id}_{\tilde\tau\omega}$, for ϖ_E a fixed uniformizer of E such that $\overline{\varpi}_E = (-1)^{e(E/E_\circ)}\varpi_E$, where $x \mapsto \bar x$ denotes the generator of $\mathrm{Gal}(E/E_\circ)$. This representation $\widetilde{\mathcal{T}}_\omega$ is then self-dual, that is, equivalent to $\widetilde{\mathcal{T}}_\omega \circ \sigma$.

Finally, the representation $\rho_\omega = \text{c-Ind}_{E^\times \tilde J}^{\mathrm{GL}_n(F)} \widetilde{\mathcal{K}}_m \otimes \widetilde{\mathcal{T}}_\omega$ is then irreducible and cuspidal, and equivalent to $\rho_\omega \circ \sigma$. Since the involution σ is a conjugate of the involution transpose-inverse, by [Gelfand and Kajdan 1975, Theorem 2], the representation $\rho_\omega \circ \sigma$ is equivalent to $\rho_\omega^\vee$.

Thus we have constructed four self-dual cuspidal representations ρ_ω of $\mathrm{GL}_m(F)$ with endoclass Θ, and it remains only to see that two are orthogonal and two symplectic. Note that ρ_ω and $\rho_{\omega'}$ are unramified twists of each other if and only if $\omega^{-1}\omega'$ is the nontrivial unramified quadratic character ω_{nr}. If either $m>1$ or $E/E_\circ$ is unramified then, by 6.6 Proposition and 6.7 Proposition, the representations ρ_ω and its self-dual unramified twist $\rho_{\omega\omega_{nr}}$ have opposite parities so we are done. (Note that, writing L for the maximal tame subextension of E/F and $L_\circ$ for that of $E_\circ/F$, we have that $L/L_\circ$ is ramified if and only if $E/E_\circ$ is ramified, since p is odd.)

On the other hand, if $m=1$ and $E/E_\circ$ is ramified then we are in the situation of Section 6.8, and the argument there explains that one pair ρ_ω, $\rho_{\omega\omega_{nr}}$ consists of two orthogonal representations, while the other pair consists of two symplectic representations, as required.

We are left with the case that Θ is the trivial endoclass and m is even. The existence of self-dual cuspidal depth-zero representations is [Adler 1997, Theorem 7.1] and the argument that there are (at least) two orthogonal and two symplectic is formally exactly as in the previous case, with $\widetilde{\underline{\mathcal{K}}}$ the trivial representation. □

We say that an orbit $[\gamma]$ in $W_F\backslash\operatorname{Irr}(\mathscr{P}_F)$ is self-dual if it is self-dual when considered as a representation of $\mathscr{P}_F$; that is, if there is $g\in W_F$ such that $\gamma^\vee\simeq{}^g\gamma$. We write $(W_F\backslash\operatorname{Irr}(\mathscr{P}_F))^{\mathrm{sd}}$ for the set of self-dual orbits. Then we have:

Proposition. *The bijection of 7.1 Theorem restricts to a bijection,*

$$\mathcal{E}^{\mathrm{sd}}(F)\to(W_F\backslash\operatorname{Irr}(\mathscr{P}_F))^{\mathrm{sd}}. \tag{7-1}$$

Proof. Let γ be an irreducible representation of $\mathscr{P}_F$ and put $n=\dim[\gamma]$. Suppose that $[\gamma]$ is a self-dual orbit and let $g\in W_F$ be such that ${}^g\gamma\simeq\gamma^\vee$. Then, as in Section 6.5, there is a unique irreducible representation Γ of the stabilizer S of γ such that $\det\Gamma$ has order a power of p and ${}^g\Gamma\simeq\Gamma^\vee$. Then the representation $\operatorname{Ind}_S^{W_F}\Gamma$ is irreducible self-dual so the corresponding cuspidal representation ρ of $\mathrm{GL}_n(F)$ is also self-dual. By [Blondel 2004, 2.2, Corollary] (see also [Goldberg et al. 2007, p. 10]), ρ contains a simple character with self-dual transfer to $\mathrm{GL}_{2n}(F)$, so the endoclass $\Theta(\rho)$, which corresponds to $[\gamma]$ by 7.1 Theorem, is self-dual.

Conversely, let Θ be a self-dual endoclass and put $n=\deg\Theta$. By the lemma, there is a self-dual cuspidal representation ρ of $\mathrm{GL}_n(F)$ with endoclass Θ. Then the corresponding irreducible representation of W_F is self-dual so the orbit in its restriction to $\mathscr{P}_F$ is also self-dual, as required. □

7.3. We now introduce the notion of *wild parameter*.

Definition. A *wild parameter* (*over* F) is a finite-dimensional semisimple complex representation $\mathcal{V}$ of $\mathscr{P}_F$ such that ${}^g\mathcal{V}\simeq\mathcal{V}$ for all $g\in W_F$. We write $\Psi(F)$ for the set of equivalence classes of wild parameters over F, and $\Psi_n(F)$ for the set of equivalence classes of n-dimensional wild parameters over F.

Equivalently, we can think of an element of $\Psi_n(F)$ as the $\mathrm{GL}_n(\mathbb{C})$-conjugacy class of a homomorphism $\psi:\mathscr{P}_F\to\mathrm{GL}_n(\mathbb{C})$ for which there exists $A\in\mathrm{GL}_n(\mathbb{C})$ such that $\psi\circ\operatorname{Ad}g=\operatorname{Ad}A\circ\psi$ for all $g\in W_F$.

Thus a finite-dimensional semisimple complex representation $\mathcal{V}$ of $\mathscr{P}_F$ is a wild parameter if and only if, when we decompose it into its isotypic components $\mathcal{V}=\bigoplus_{\gamma\in\operatorname{Irr}(\mathscr{P}_F)}\mathcal{V}(\gamma)$, we have

$$\dim\mathcal{V}(\gamma)=\dim\mathcal{V}({}^g\gamma)\quad\text{for all } g\in W_F.$$

Therefore a wild parameter is equivalent to

$$\bigoplus_{W_F \backslash \operatorname{Irr}(\mathscr{P}_F)} m_{[\gamma]}[\gamma],$$

where we are thinking of the orbit $[\gamma]$ as the sum over the W_F-conjugates of $\gamma \in \operatorname{Irr}(\mathscr{P}_F)$, and $m_{[\gamma]} \in \mathbb{Z}_{\geq 0}$.

Equivalently, the n-dimensional wild parameters are precisely the restrictions to $\mathscr{P}_F$ of the Langlands parameters for $\mathrm{GL}_n(F)$; that is, writing $\Phi_n(F)$ for the set of admissible homomorphisms $\phi : W_F \times \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ up to conjugacy, and $\Phi(F) = \bigcup_{n\geq 1} \Phi_n(F)$, the natural map

$$\Phi(F) \to \Psi(F)$$

induced by $\phi \mapsto \phi_{|\mathscr{P}_F}$ is surjective. Indeed, by taking direct sums one need only check that, for any $\gamma \in \mathscr{P}_F$, there is a Langlands parameter ϕ whose restriction to $\mathscr{P}_F$ is isomorphic to $[\gamma]$. This, however, follows from the discussion in Section 6.5: γ extends to a representation Γ of its stabilizer S_γ by [Bushnell and Henniart 2014, 1.3, Proposition], and then $\operatorname{Ind}_{S_\gamma}^{W_F} \Gamma$ is the required Langlands parameter (with trivial $\mathrm{SL}_2(\mathbb{C})$ action).

Recall from Section 2.7 that an endoparameter of degree n over F is a formal sum

$$\sum_{\Theta\in\mathcal{E}} m_\Theta \Theta, \quad m_\Theta \in \mathbb{Z}_{\geq 0}, \qquad \text{such that} \quad \sum_{\Theta\in\mathcal{E}} m_\Theta \deg \Theta = n.$$

We write $\mathcal{EE}_n(F)$ for the set of endoparameters of degree n over F. Then the ramification theorem for GL_n (7.1 Theorem) together with the compatibility of the Langlands correspondence with parabolic induction immediately give:

Theorem. *The bijection of 7.1 Theorem induces, for each n, a bijection $\mathcal{EE}_n(F) \to \Psi_n(F)$ which is compatible with the Langlands correspondence*:

$$\begin{array}{ccc} \operatorname{Irr}(\mathrm{GL}_n(F)) & \xrightarrow{\sim} & \Phi_n(F) \\ \downarrow & & \downarrow \\ \mathcal{EE}_n(F) & \xrightarrow{\sim} & \Psi_n(F) \end{array}$$

7.4. Now we turn to the case of the symplectic group G and recall Arthur's local Langlands correspondence in this case.

We denote by $\Phi(G)$ the set of Langlands parameters for G, that is, the set of conjugacy classes of homomorphisms $\phi : W_F \times \mathrm{SL}_2(\mathbb{C}) \to \mathrm{SO}_{2N+1}(\mathbb{C})$ such that the representation obtained by composing with the natural inclusion map $\iota : \mathrm{SO}_{2N+1}(\mathbb{C}) \hookrightarrow \mathrm{GL}_{2N+1}(\mathbb{C})$ is semisimple.

We denote by $\Phi^{\mathsf{disc}}(G)$ the set of *discrete* Langlands parameters, that is, those whose image is not contained in a proper parabolic subgroup of $\mathrm{SO}_{2N+1}(\mathbb{C})$; equivalently, $\iota \circ \phi$ is a direct sum of inequivalent irreducible orthogonal representations of $W_F \times \mathrm{SL}_2(\mathbb{C})$ and has determinant 1. Thus, given ϕ a discrete

Langlands parameter, the representation $\iota \circ \phi$ decomposes as a multiplicity-free direct sum

$$\bigoplus_{i\in I} \sigma_i \otimes \mathrm{St}_{m_i}, \tag{7-2}$$

where St_m denotes the unique m-dimensional irreducible algebraic representation of $\mathrm{SL}_2(\mathbb{C})$ for $m \geq 1$, and the σ_i are irreducible self-dual representations of W_F, such that

- $\sum_{i\in I} m_i \dim \sigma_i = 2N+1$,
- σ_i is symplectic if m_i is even and orthogonal if m_i is odd,
- $\prod_{i\in I} \det(\sigma_i)^{m_i} = 1$.

We say that a discrete Langlands parameter ϕ is *cuspidal* if, whenever $\sigma \otimes \mathrm{St}_m$ is a subrepresentation of $\iota\circ\phi$ and $m > 2$, the representation $\sigma\otimes\mathrm{St}_{m-2}$ is also a subrepresentation of $\iota\circ\phi$. We denote by $\Phi^{\mathrm{cusp}}(G)$ the set of cuspidal Langlands parameters.

As usual, for ϕ a discrete Langlands parameter, we denote by $\mathscr{S}_\phi$ the group of connected components of the centralizer in $\mathrm{SO}_{2N+1}(\mathbb{C})$ of the image of ϕ. This is a finite product of copies of the cyclic group of order 2; if $\iota\circ\phi$ decomposes as in (7-2), then $\mathscr{S}_\phi$ has order $2^{\#I-1}$.

Theorem [Arthur 2013, Theorems 1.5.1 and 2.2.1; Mœglin 2011, Theorem 1.5.1]. *Suppose that F is of characteristic zero. There is a natural surjective map from the set of discrete series representations of G to $\Phi^{\mathrm{disc}}(G)$ with finite fibres, characterized by an equality of stable distributions via transfer to $\mathrm{GL}_{2N+1}(\mathbb{C})$. Moreover,*

- *the fibre of $\phi \in \Phi^{\mathrm{disc}}(G)$ is in bijection with the set of characters of $\mathscr{S}_\phi$;*
- *the fibre Π_ϕ of $\phi\in\Phi^{\mathrm{disc}}(G)$ contains a cuspidal representation of G if and only if ϕ is cuspidal, in which case $\Pi_\phi \cap \mathrm{Cusp}(G)$ is in bijection with the set of **alternating** characters of $\mathscr{S}_\phi$.*

We do not recall the definition of *alternating* character (see [Mœglin 2011, §1.5]) but only recall that if, for ϕ a cuspidal Langlands parameter as in (7-2), we set $I_0 = \{\sigma_i \mid \sigma_i \text{ is orthogonal}\}$, then there are $2^{\#I_0-1}$ alternating characters of $\mathscr{S}_\phi$. (Note that I_0 is nonempty, since one of the σ_i must be a quadratic character.) In particular, the L-packet of a cuspidal Langlands parameter ϕ consists only of cuspidal representations if and only if $m_i = 1$ for all $i \in I$, in the description (7-2); that is, each self-dual irreducible representation of W_F which appears in ϕ is orthogonal and appears with multiplicity 1. In this case, we say that ϕ is *regular*.

7.5. We say that a wild parameter $\mathcal{V}$ is *self-dual* if it is self-dual as a representation of $\mathscr{P}_F$, in which case $\det(\mathcal{V})$ is trivial (since p is odd).

Given a self-dual wild parameter, we would like to see that there is a unique choice of orthogonal structure on it. This is indeed a special case of the following result on the existence and uniqueness of orthogonal structures on self-dual representations of groups of odd order.

Proposition. *Let $\mathscr{G}$ be a finite group of odd order and let $\mathcal{V}$ be a finite-dimensional complex representation of $\mathscr{G}$. If $\mathcal{V}$ is self-dual, then $\mathcal{V}$ is orthogonal: there is on $\mathcal{V}$ a $\mathscr{G}$-invariant nondegenerate symmetric bilinear form; moreover such a form is unique up to the action of* $\mathrm{Aut}_{\mathscr{G}}(\mathcal{V})$.

In other words, a self-dual representation of $\mathscr{G}$ is underlying a unique (up to isomorphism) orthogonal representation.

Proof. As $\mathscr{G}$ has odd order, the only self-dual irreducible representation of $\mathscr{G}$ is the trivial representation $1_{\mathscr{G}}$. For an irreducible representation γ of $\mathscr{G}$, let $\mathcal{V}(\gamma)$ be the γ-isotypic component of $\mathcal{V}$, and put $\mathcal{V}_\gamma = \mathrm{Hom}_{\mathscr{G}}(\gamma, \mathcal{V})$, so that $\mathcal{V}(\gamma)$ decomposes canonically as $\gamma \otimes \mathcal{V}_\gamma$. Then $\mathcal{V}$ is self-dual if and only if $\mathcal{V}_\gamma$ and $\mathcal{V}_{\gamma^\vee}$ have the same dimension for all γ.

Assume $\mathcal{V}$ is self-dual. For any $\mathscr{G}$-invariant nondegenerate symmetric bilinear form on $\mathcal{V}$, we can write $\mathcal{V}$ as the orthogonal direct sum of its subspaces $\mathcal{V}(1_{\mathscr{G}})$ and $\mathcal{V}(\gamma) \oplus \mathcal{V}(\gamma^\vee)$ for γ running through a set of representatives of the nontrivial irreducible representations up to contragredient. On $\mathcal{V}(1_{\mathscr{G}})$, where $\mathscr{G}$ acts trivially, there is a nondegenerate symmetric bilinear form, unique up to the action of $\mathrm{Aut}(\mathcal{V}(1_{\mathscr{G}}))$.

Therefore, for existence and uniqueness, it is enough to consider the case where $\mathcal{V} = \mathcal{V}(\gamma) \oplus \mathcal{V}(\gamma^\vee)$ for some nontrivial γ. Then the dual of $\mathcal{V}(\gamma)$ is $\gamma^\vee \otimes (\mathcal{V}_\gamma)^*$, whereas the dual of $\mathcal{V}(\gamma^\vee)$ is $\gamma \otimes (\mathcal{V}_{\gamma^\vee})^*$. An isomorphism $j : \mathcal{V} \to \mathcal{V}^\vee$ (that is, a self-duality on $\mathcal{V}$) is the direct sum of $\mathrm{Id}_\gamma \otimes i$ and $\mathrm{Id}_{\gamma^\vee} \otimes i'$, where i is an isomorphism of $\mathcal{V}_\gamma$ onto $(\mathcal{V}_{\gamma^\vee})^*$, and i' an isomorphism of $\mathcal{V}_{\gamma^\vee}$ onto $(\mathcal{V}_\gamma)^*$. The self-duality is orthogonal if and only if i and i' are transpose to each other.

Obviously there exists then an orthogonal structure on $\mathcal{V}$, and moreover all such structures are given by the choice of i (with i' its transpose). Since $\mathrm{Aut}_{\mathscr{G}}(\mathcal{V})$, which is the product $\mathrm{Aut}(\mathcal{V}_\gamma) \times \mathrm{Aut}(\mathcal{V}_{\gamma^\vee})$, acts transitively on the set of i, we have uniqueness too. □

7.6. Now let $\mathcal{V}$ be an n-dimensional self-dual wild parameter over F. By 7.5 Proposition, $\mathcal{V}$ then carries a $\mathscr{P}_F$-invariant nondegenerate symmetric bilinear form, unique up to the action of $\mathrm{Aut}_{\mathscr{P}_F}(\mathcal{V})$. Thus we can regard $\mathcal{V}$ as a homomorphism $\psi : \mathscr{P}_F \to \mathrm{SO}(\mathcal{V}) \simeq \mathrm{SO}_n(\mathbb{C})$.

For $\gamma \in \mathrm{Irr}(\mathscr{P}_F)$, we write $\mathcal{V}[\gamma]$ for the component of $\mathcal{V}$ corresponding to the orbit of γ under W_F; that is,

$$\mathcal{V}[\gamma] = \sum_{g \in W_F} \mathcal{V}({}^g\gamma).$$

We consider the stabilizer in $\mathrm{SO}(\mathcal{V})$ of the self-dual decomposition

$$\mathcal{V} = \bigoplus_{W_F \backslash \mathrm{Irr}(\mathscr{P}_F)} \mathcal{V}[\gamma]$$

and say that $\mathcal{V}$ is *discrete* if this stabilizer is contained in no proper Levi subgroup of $\mathrm{SO}(\mathcal{V})$. Equivalently, the self-dual parameter $\mathcal{V}$ is discrete if and only if every orbit $[\gamma]$ in the support of $\mathcal{V}$ (that is, such that $\mathcal{V}[\gamma]$ is nonzero) is self-dual.

We write $\Psi_n^{\mathrm{sd}}(F)$ for the set of equivalence classes of discrete self-dual n-dimensional wild parameters over F. Note that the restriction to $\mathscr{P}_F$ of any discrete Langlands parameter for G is a discrete self-dual wild parameter of dimension $2N+1$, which explains the nomenclature.

Recall also that we have the set $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_n(F)$ of self-dual endoparameters of degree n over F, which consists of those endoparameters of degree n with support in the set $\mathcal{E}^{\mathrm{sd}}(F)$ of self-dual endoclasses. Then we have the following ramification theorem for G.

Theorem. *The bijection (7-1) induces, for each $N \geq 1$, a bijection $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F) \to \Psi^{\mathrm{sd}}_{2N+1}(F)$ which, when F is of characteristic zero, is compatible with the Langlands correspondence for cuspidal representations of G*:

$$\begin{tikzcd} \mathrm{Cusp}(G) \arrow[r, two heads] \arrow[d, two heads] & \Phi^{\mathrm{cusp}}(G) \arrow[d, two heads] \\ \mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F) \arrow[r, "\sim"] & \Psi^{\mathrm{sd}}_{2N+1}(F) \end{tikzcd}$$

The induced bijection $\mathcal{E}\mathcal{E}^{\mathrm{sd}}_{2N}(F) \to \Psi^{\mathrm{sd}}_{2N+1}(F)$ is not as obvious as in the case of general linear groups. If we denote the bijection (7-1) by $\Theta \mapsto [\gamma(\Theta)]$ then the induced map is

$$\sum_{\Theta} m_\Theta \Theta \mapsto 1_{\mathscr{P}_F} \oplus \bigoplus_{\Theta} m_\Theta [\gamma(\Theta^2)]. \tag{7-3}$$

We remark also that this theorem asserts that the restriction map $\Phi^{\mathrm{cusp}}(G) \to \Psi^{\mathrm{sd}}_{2N+1}(F)$ is surjective, so that every discrete self-dual wild parameter of dimension $2N+1$ occurs as the restriction of not only some discrete Langlands parameter for G but of some cuspidal parameter. In fact, we show that it occurs as the restriction of a *regular* parameter (i.e., one whose L-packet consists only of cuspidal representations).

Proof. Since the only irreducible self-dual representation of $\mathscr{P}_F$ is the trivial representation (so the only odd-dimensional self-dual class $[\gamma]$ is that of the trivial representation), while the squaring map on endoclasses is a bijection (since p is odd), it is clear that (7-3) defines a bijection. Its compatibility with the Langlands correspondence is now just a reinterpretation of 2.8 Theorem, using 7.3 Theorem.

It remains to prove that the vertical maps are surjective. We prove that the map on the right is surjective, and then surjectivity on the left follows. So let $\mathcal{V} = \bigoplus m_{[\gamma]}[\gamma]$ be a $(2N+1)$-dimensional self-dual wild parameter (where the sum is over the W_F orbits in $\mathrm{Irr}(\mathscr{P}_F)$ as usual). We will define a regular Langlands parameter $\sigma = \bigoplus \sigma[\gamma]$ for G such that $\sigma[\gamma]$ restricts to $\mathcal{V}[\gamma] = m_{[\gamma]}[\gamma]$.

Let $\gamma \in \mathrm{Irr}(\mathscr{P}_F)$ be a nontrivial representation. If $m_{[\gamma]} = 0$ then we put $\sigma[\gamma] = \{0\}$ so assume $m_{[\gamma]} > 0$, in which case the orbit $[\gamma]$ is self-dual. Let Θ be corresponding (self-dual) endoclass and let $E/E_\circ$ be the quadratic extension associated to a skew simple stratum which has a simple character with endoclass Θ. We pick nonnegative integers m_1, m_2 with $m_1 + m_2 = m_{[\gamma]}$ such that

(i) m_1, m_2 are odd or 0 if $E/E_\circ$ is unramified;

(ii) m_1, m_2 are even or 1 if $E/E_\circ$ is ramified.

For $i = 1, 2$ we put $n_i = m_i \deg \Theta$. Then, by 7.2 Lemma, there exist inequivalent orthogonal self-dual cuspidal representations ρ_1, ρ_2 of $\mathrm{GL}_{n_1}(F), \mathrm{GL}_{n_2}(F)$ respectively, both with endoclass Θ. Let σ_1, σ_2 denote the corresponding Langlands parameters, which are orthogonal and put $\sigma[\gamma] = \sigma_1 \oplus \sigma_2$; then the restriction of $\sigma[\gamma]$ to $\mathscr{P}_F$ is $\mathcal{V}[\gamma]$, as required, by 7.1 Theorem.

Finally, put $m = m_{[1_{\mathscr{P}_F}]} - 1$, which is even. By 7.2 Lemma, there is an orthogonal self-dual depth-zero cuspidal representation of $\mathrm{GL}_m(F)$, and let δ be the corresponding representation of W_F. We put $\sigma[1_{\mathscr{P}_F}] = \delta \oplus \omega$, where $\omega = \det(\delta)\prod_{[\gamma]\neq[1_{\mathscr{P}_F}]} \det \sigma[\gamma]$.

Then $\sigma = \bigoplus_{[\gamma]} \sigma[\gamma]$ is a regular cuspidal Langlands parameter for G which restricts to $\mathcal{V}$. □

7.7. In the proof of 7.6 Theorem we saw that, for any self-dual wild parameter $\mathcal{V}$ of odd dimension, there is a regular Langlands parameter for G which restricts to $\mathcal{V}$. As well as this, one can (in general) cook up other examples of Langlands parameters which restrict to $\mathcal{V}$ and are highly irregular. Since we find it amusing, we include here a description of how to find a highly irregular Langlands parameter which restricts to $\mathcal{V}$.

We begin with the following observation, which is just the translation of 7.2 Lemma (with $m = 1$) to Galois representations. Suppose $\gamma \in \mathrm{Irr}(\mathscr{P}_F)$ is nontrivial with self-dual W_F-orbit. Then there are four self-dual representations of W_F whose restriction to $\mathscr{P}_F$ is $[\gamma]$, two of which are orthogonal and two of which are symplectic. We write $\sigma_{\gamma,1}, \sigma_{\gamma,2}$ for the two orthogonal ones, and $\sigma_{\gamma,3}, \sigma_{\gamma,4}$ for the two symplectic ones.

Now we decompose $\mathcal{V} = \bigoplus m_{[\gamma]}[\gamma]$ as above. As before, we will define a Langlands parameter $\sigma = \bigoplus \sigma[\gamma]$, with $\sigma[\gamma]|_{\mathscr{P}_F} = m_{[\gamma]}[\gamma]$. We will obtain a parameter which is not regular whenever either $m_{[1_{\mathscr{P}_F}]} > 3$ or $m_{[\gamma]} > 1$ for some nontrivial self-dual $[\gamma]$.

By Lagrange's four-square theorem, we can find nonnegative integers such that

$$4m_{[\gamma]} + 2 = a_1^2 + a_2^2 + a_3^2 + a_4^2.$$

Moreover, two of the a_i are even and the other two odd. We label them so that a_1, a_2 are even and a_3, a_4 are odd and, when $m_{[\gamma]} = 2$, we take the solution with $a_1 = a_2 = 0$. Then we set

$$\sigma[\gamma] = \bigoplus_{i=1}^{4} \sigma_{\gamma,i} \otimes (\mathrm{St}_{a_i-1} \oplus \mathrm{St}_{a_i-3} \oplus \cdots),$$

where we understand that we ignore the terms on the right where $a_i \leq 1$.

Finally, write $\omega_1 = \prod_{[\gamma]\neq[1_{\mathscr{P}_F}]} \det \sigma[\gamma]$, which is a quadratic character, and let $\omega_2, \omega_3, \omega_4$ denote the other three quadratic characters. Again, there are nonnegative integers such that

$$m_{[1_{\mathscr{P}_F}]} = a_1^2 + a_2^2 + a_3^2 + a_4^2.$$

Since $m_{[1_{\mathscr{P}_F}]}$ is odd, there is exactly one a_i which has opposite parity to the other three, and we choose our numbering so that this is a_1. Then we take

$$\sigma[1_{\mathscr{P}_F}] = \bigoplus_{i=1}^{4} \omega_i \otimes (\mathrm{St}_{2a_i-1} \oplus \mathrm{St}_{2a_i-3} \oplus \cdots \oplus \mathrm{St}_1),$$

where, again, we ignore the terms for which $a_i = 0$.

Acknowledgements

The authors would like to thank the Schrödinger Institute in Vienna, where this work was initiated back in winter 2009. Blondel and Henniart would like to thank the University of East Anglia for hosting them during several visits in the course of the present work. Henniart would like to thank the Université de Paris-Sud and the Institut Universitaire de France. Stevens would like to thank the Université Paris 7 for inviting and hosting him in spring 2012. Stevens was supported by EPSRC grants EP/G001480/1 and EP/H00534X/1, and by the Heilbronn Institute for Mathematical Research for the end of the writing-up period. We also thank the referees for their careful reading of the paper.

References

[Adler 1997] J. D. Adler, "Self-contragredient supercuspidal representations of GL_n", *Proc. Amer. Math. Soc.* **125**:8 (1997), 2471–2479. MR Zbl

[Arthur 2013] J. Arthur, *The endoscopic classification of representations*, Amer. Math. Soc. Colloq. Publ. **61**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Blondel 2004] C. Blondel, "$\mathrm{Sp}(2N)$-covers for self-contragredient supercuspidal representations of $\mathrm{GL}(N)$", *Ann. Sci. École Norm. Sup.* (4) **37**:4 (2004), 533–558. MR Zbl

[Blondel 2012] C. Blondel, "Représentation de Weil et β-extensions", *Ann. Inst. Fourier* (*Grenoble*) **62**:4 (2012), 1319–1366. MR Zbl

[Blondel and Stevens 2009] C. Blondel and S. Stevens, "Genericity of supercuspidal representations of p-adic Sp_4", *Compos. Math.* **145**:1 (2009), 213–246. MR Zbl

[Bushnell and Henniart 1996] C. J. Bushnell and G. Henniart, "Local tame lifting for $\mathrm{GL}(N)$, I: Simple characters", *Inst. Hautes Études Sci. Publ. Math.* **83** (1996), 105–233. MR Zbl

[Bushnell and Henniart 2003] C. J. Bushnell and G. Henniart, "Local tame lifting for $\mathrm{GL}(n)$, IV: Simple characters and base change", *Proc. London Math. Soc.* (3) **87**:2 (2003), 337–362. MR Zbl

[Bushnell and Henniart 2005] C. J. Bushnell and G. Henniart, "The essentially tame local Langlands correspondence, I", *J. Amer. Math. Soc.* **18**:3 (2005), 685–710. MR Zbl

[Bushnell and Henniart 2006] C. J. Bushnell and G. Henniart, *The local Langlands conjecture for* $\mathrm{GL}(2)$, Grundlehren der Math. Wissenschaften **335**, Springer, 2006. MR Zbl

[Bushnell and Henniart 2014] C. J. Bushnell and G. Henniart, *To an effective local Langlands correspondence*, Mem. Amer. Math. Soc. **1087**, 2014. MR Zbl

[Bushnell and Kutzko 1993] C. J. Bushnell and P. C. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Ann. of Math. Studies **129**, Princeton Univ. Press, 1993. MR Zbl

[Bushnell and Kutzko 1998] C. J. Bushnell and P. C. Kutzko, "Smooth representations of reductive p-adic groups: structure theory via types", *Proc. London Math. Soc.* (3) **77**:3 (1998), 582–634. MR Zbl

[Bushnell and Kutzko 1999] C. J. Bushnell and P. C. Kutzko, "Semisimple types in GL_n", *Compos. Math.* **119**:1 (1999), 53–97. MR Zbl

[Cabanes and Enguehard 2004] M. Cabanes and M. Enguehard, *Representation theory of finite reductive groups*, New Math. Monographs **1**, Cambridge Univ. Press, 2004. MR Zbl

[Dat 2009] J.-F. Dat, "Finitude pour les représentations lisses de groupes p-adiques", *J. Inst. Math. Jussieu* **8**:2 (2009), 261–333. MR Zbl

[Gelfand and Kajdan 1975] I. M. Gel'fand and D. A. Kajdan, "Representations of the group $\mathrm{GL}(n, K)$ where K is a local field", pp. 95–118 in *Lie groups and their representations* (Budapest, 1971), edited by I. M. Gel'fand, Halsted, New York, 1975. MR Zbl

[Goldberg et al. 2007] D. Goldberg, P. Kutzko, and S. Stevens, "Covers for self-dual supercuspidal representations of the Siegel Levi subgroup of classical p-adic groups", *Int. Math. Res. Not.* **2007**:22 (2007), art. id. rnm085. MR Zbl

[Gross and Reeder 2010] B. H. Gross and M. Reeder, "Arithmetic invariants of discrete Langlands parameters", *Duke Math. J.* **154**:3 (2010), 431–508. MR Zbl

[Harris and Taylor 2001] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, Ann. of Math. Studies **151**, Princeton Univ. Press, 2001. MR Zbl

[Henniart 2000] G. Henniart, "Une preuve simple des conjectures de Langlands pour $\mathrm{GL}(n)$ sur un corps p-adique", *Invent. Math.* **139**:2 (2000), 439–455. MR Zbl

[Henniart 2010] G. Henniart, "Correspondance de Langlands et fonctions L des carrés extérieur et symétrique", *Int. Math. Res. Not.* **2010**:4 (2010), 633–673. MR Zbl

[Kariyama 2008] K. Kariyama, "On types for unramified p-adic unitary groups", *Canad. J. Math.* **60**:5 (2008), 1067–1107. MR Zbl

[Kurinczuk et al. 2016] R. Kurinczuk, D. Skodlerack, and S. Stevens, "Endo-classes for p-adic classical groups", preprint, 2016. arXiv

[Laumon et al. 1993] G. Laumon, M. Rapoport, and U. Stuhler, "$\mathscr{D}$-elliptic sheaves and the Langlands correspondence", *Invent. Math.* **113**:2 (1993), 217–338. MR Zbl

[Lust and Stevens 2016] J. Lust and S. Stevens, "On depth zero L-packets for classical groups", preprint, 2016. arXiv

[Lusztig 1984] G. Lusztig, *Characters of reductive groups over a finite field*, Ann. of Math. Studies **107**, Princeton Univ. Press, 1984. MR Zbl

[Miyauchi and Stevens 2014] M. Miyauchi and S. Stevens, "Semisimple types for p-adic classical groups", *Math. Ann.* **358**:1-2 (2014), 257–288. MR Zbl

[Mœglin 2003] C. Mœglin, "Points de réductibilité pour les induites de cuspidales", *J. Algebra* **268**:1 (2003), 81–117. MR Zbl

[Mœglin 2011] C. Mœglin, "Multiplicité 1 dans les paquets d'Arthur aux places p-adiques", pp. 333–374 in *On certain L-functions* (West Lafayette, IN, 2007), edited by J. Arthur et al., Clay Math. Proc. **13**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Mœglin 2014] C. Mœglin, "Paquets stables des séries discrètes accessibles par endoscopie tordue: leur paramètre de Langlands", pp. 295–336 in *Automorphic forms and related geometry: assessing the legacy of I. I. Piatetski-Shapiro* (New Haven, CT, 2012), edited by J. W. Cogdell et al., Contemp. Math. **614**, Amer. Math. Soc., Providence, RI, 2014. MR Zbl

[Mœglin and Tadić 2002] C. Mœglin and M. Tadić, "Construction of discrete series for classical p-adic groups", *J. Amer. Math. Soc.* **15**:3 (2002), 715–786. MR Zbl

[Sécherre and Stevens 2016] V. Sécherre and S. Stevens, "Towards an explicit local Jacquet–Langlands correspondence beyond the cuspidal case", preprint, 2016. arXiv

[Stevens 2001] S. Stevens, "Intertwining and supercuspidal types for p-adic classical groups", *Proc. London Math. Soc.* (3) **83**:1 (2001), 120–140. MR Zbl

[Stevens 2005] S. Stevens, "Semisimple characters for p-adic classical groups", *Duke Math. J.* **127**:1 (2005), 123–173. MR Zbl

[Stevens 2008] S. Stevens, "The supercuspidal representations of p-adic classical groups", *Invent. Math.* **172**:2 (2008), 289–352. MR Zbl

Communicated by Michael Rapoport
Received 2017-05-30 Revised 2018-06-02 Accepted 2018-07-20

corinne.blondel@imj-prg.fr *CNRS-IMJ-PRG, Université Paris Diderot, Paris, France*

guy.henniart@math.u-psud.fr *Université de Paris-Sud, Laboratoire de Mathématiques d'Orsay, Orsay, France*

shaun.stevens@uea.ac.uk *School of Mathematics, University of East Anglia, Norwich, United Kingdom*

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2387

Realizing 2-groups as Galois groups following Shafarevich and Serre

Peter Schmid

Let G be a finite p-group for some prime p, say of order p^n. For odd p the inverse problem of Galois theory for G has been solved through the (classical) work of Scholz and Reichardt, and Serre has shown that their method leads to fields of realization where at most n rational primes are (tamely) ramified. The approach by Shafarevich, for arbitrary p, has turned out to be quite delicate in the case $p = 2$. In this paper we treat this exceptional case in the spirit of Serre's result, bounding the number of ramified primes at least by an integral polynomial in the rank of G, the polynomial depending on the 2-class of G.

1. Introduction

Let p be prime and G a finite p-group. By [Scholz 1937; Reichardt 1937] there is a Galois extension $K\,|\,\mathbb{Q}$ with group G provided p is odd. The general case, allowing $p = 2$, has been treated by Shafarevich [1954] in a different and somehow more complicated way. Actually the $p = 2$ case led to controversial discussions some years ago, because Shafarevich used in his proof "something on free groups (and their p-filtration) which is false for $p = 2$" (Serre in a letter of May 10, 1988). Shafarevich [1989] corrected this by suggesting to use a refined filtration. The proof of Shafarevich's theorem (for solvable groups) given in [Neukirch et al. 2000] is based on this filtration; it employs deep results and techniques in cohomology of number fields.

The Scholz–Reichardt method has been explained further by Serre. In a letter of September 6, 1988 he wrote: "I have now looked into Reichardt's 1937 paper in Crelle, and it is quite nice. The proof gives a rather surprising result, namely: if G has order p^n, $p \neq 2$, then G can be realized as $\mathrm{Gal}(K\,|\,\mathbb{Q})$ where K is ramified at most n primes. However, $p \neq 2$ seems indeed essential." This was elaborated in [Serre 1992, Chapter 2]. A slight improvement was given in [Plans 2004]; see also [Geyer and Jarden 1998].

The field K in Serre's letter refers to so-called *Scholz fields* (Section 2). Only tame ramification happens in these fields, so that the inertia groups are all cyclic. This implies that the cardinality of the set $\mathrm{Ram}(K)$ of rational primes ramified in K must be at least equal to the rank $d(G)$ of the p-group $G = \mathrm{Gal}(K\,|\,\mathbb{Q})$, its minimum number of generators (in view of Burnside's basis theorem and the Hermite–Minkowski theorem). Indeed at least $d(G)$ primes must ramify in the socle $\mathfrak{S}(K)$ of K, the fixed field of the Frattini subgroup $\Phi(G)$ of G (where $G/\Phi(G)$ is an $\mathbb{F}_p$-vector space of dimension $d(G)$).

MSC2010: primary 11R32; secondary 20D15.
Keywords: Galois 2-groups, Scholz fields, tame ramification, Shafarevich, Serre.

The *p-class* (sometimes also called Frattini class) of the p-group G is the least positive integer c such that G has a central series of length c with all factors being elementary.

Theorem. *Let G be a nontrivial finite 2-group with rank d and 2-class c. There exist infinitely many Scholz fields K with pairwise coprime (absolute) discriminants realizing G as Galois group over $\mathbb{Q}$ and satisfying* $\mathrm{Ram}(K) \subseteq 1 + 2^c\mathbb{Z}$ *and* $|\mathrm{Ram}(K)| \leq f_c(d)$ *for some integral polynomial f_c of degree* $(c+3)!/24$.

The upper bound on the cardinality of $\mathrm{Ram}(K)$ is rather weak (compared with the odd case). This is primarily due to the (inductive) shrinking process needed (see below). The polynomial $f_c \in \mathbb{Z}[X]$ will be defined recursively ($f_1 = X,\ f_2 = 2X^5 + X^2, \ldots$).

For any number field K_0 there is a field K as above having discriminant coprime to that of K_0. Then the compositum K_0K (in $\mathbb{C}$) is Galois over K_0 and admits G. Reichardt [1937] proved a corresponding result in the odd case.

The proof of the theorem utilizes ideas from Scholz, Reichardt and Serre, as well as from Shafarevich. Up to isomorphism there is a unique p-group $G_d^c(p)$ of minimal order with rank d and p-class c which has every (finite) p-group of rank d and p-class c as epimorphic image. We will have to consider, like Shafarevich [1954], this so-called *disposition p-group* (for $p = 2$). In order to eliminate certain (*Scholz*) *obstructions* we also use a *shrinking process*, a technique also developed in [Shafarevich 1954]. However, avoiding Shafarevich's "graded functions on canonical homomorphisms", this will be based on the Chevalley–Warning theorem, in a manner as proposed in [Meshulam and Sonn 1999; Neukirch et al. 2000, Proposition (9.5.4)]. It is possible to derive upper bounds on $|\mathrm{Ram}(K)|$ following the lines of proof given by Shafarevich; e.g., see [Rabayev 2013].

The "minimal ramification problem" for a p-group G is the question whether G can be realized as the group of a tamely ramified Galois extension of $\mathbb{Q}$ in which exactly $d(G)$ primes are ramified. Kisilevsky, Neftin and Sonn [Kisilevsky et al. 2010] answered this question to the affirmative in the case where G is semiabelian. At present a general answer seems to be out of reach; no counterexample is known so far.

2. Scholz fields

In this section G is a finite p-group for some prime p. As usual $G_{\mathbb{Q}}$ denotes the absolute Galois group of the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ contained in $\mathbb{C}$. Number fields are understood to be subfields of $\overline{\mathbb{Q}}$. For any rational prime q we fix one of the $G_{\mathbb{Q}}$-conjugate prime ideals $\mathfrak{Q}$ above q of the ring of algebraic integers in $\overline{\mathbb{Q}}$, and let $I_q \subset D_q$ denote the inertia and decomposition groups of $\mathfrak{Q}$ ($D_q/I_q \cong \mathrm{Gal}(\overline{\mathbb{F}}_q \,|\, \mathbb{F}_q) \cong \widehat{\mathbb{Z}}$).

Definition 1. Let N be a positive integer with $p^N \geq \exp(G)$, where $\exp(G)$ denotes the exponent of G. Suppose we have a (continuous) epimorphism $\varphi : G_{\mathbb{Q}} \twoheadrightarrow G$. The fixed field $K = \overline{\mathbb{Q}}^{\mathrm{Ker}(\varphi)}$ of $\mathrm{Ker}(\varphi)$ (having Galois group G over the rationals) is a *Scholz field* with respect to N provided:

(S1) Each $q \in \mathrm{Ram}(K)$ belongs to $1 + p^N\mathbb{Z}$.

(S2) Each $q \in \mathrm{Ram}(K)$ is *busy* in K (in German: "fleissig"); that is, $\varphi(I_q) = \varphi(D_q)$.

We say that K is a Scholz field (per se) if it is Scholz with respect to N for some N with $p^N \geq \exp(G)$. Normal subfields of Scholz fields obviously are Scholz fields. We also say that K is a *strong* Scholz field (with respect to N) if in addition the socle satisfies $\mathfrak{S}(K) = P_1 \cdots P_d$, where $d = d(G)$ and the sets $\mathrm{Ram}(P_i)$ for the (cyclic) fields P_i are pairwise disjoint and of the same cardinality.

By (S1) ramification in a Scholz field K is always tame, and by (S2) the residue class degrees of the primes of K ramified over $\mathbb{Q}$ are 1. Our definition of a Scholz field is in accordance with that given in [Scholz 1937; Reichardt 1937; Serre 1992] (for odd p), but differs from that in [Shafarevich 1954]. In the $p = 2$ case from (S1), with $N \geq 3$, it follows that 2 splits completely in $\mathfrak{S}(K)$ and that this is a (totally) real field, which just says that $\mathfrak{S}(K)$ is a Scholz field in the sense of Shafarevich.

Proposition 2.1. *Let $Z \rightarrowtail H \overset{\pi}{\twoheadrightarrow} G$ be a nonsplit central extension of the p-group G where $Z = Z_p$ is cyclic of order p. Assume that $K = \overline{\mathbb{Q}}^{\mathrm{Ker}(\varphi)}$ is a Scholz field with respect to N where $p^N \geq \exp(H)$. Then the embedding problem $(G_{\mathbb{Q}}, \varphi, \pi)$ has a proper solution $E = \overline{\mathbb{Q}}^{\mathrm{Ker}(\psi)}$, with $\psi : G_{\mathbb{Q}} \twoheadrightarrow H$ lifting φ, such that* $\mathrm{Ram}(E) = \mathrm{Ram}(K)$.

Since Z is contained in the Frattini subgroup of H, every solution of the embedding problem $(G_{\mathbb{Q}}, \varphi, \pi)$ is proper. Let $\rho \in H^2(G, Z)$ be the cohomology class of the extension. Recall that $(G_{\mathbb{Q}}, \varphi, \pi)$ has a solution if and only if the map $\varphi^* : H^2(G, Z) \to H^2(G_{\mathbb{Q}}, Z)$ induced by φ vanishes at ρ [Neukirch et al. 2000, Proposition (9.4.2)]. The existence of a solution then follows by using standard global-local techniques, as described in [Serre 1992, Lemma 2.1.5]. Actually Serre treats only the case where p is odd; see also [Scholz 1937; Reichardt 1937]. Let $p = 2$ (and $Z = Z_2$). The map $\tau \mapsto \tau^2$ is an epimorphism of $\overline{\mathbb{Q}}^*$ onto itself with kernel $\{\pm 1\} \cong Z$. Using that $H^1(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^*) = 0$ (Hilbert's Theorem 90) we get that $H^2(G_{\mathbb{Q}}, Z)$ is isomorphic to $\mathrm{Br}_2(\mathbb{Q})$, the 2-torsion of the Brauer group $\mathrm{Br}(\mathbb{Q}) = H^2(G_{\mathbb{Q}}, \overline{\mathbb{Q}}^*)$ of $\mathbb{Q}$. Restriction to the decomposition groups $D_q \cong G_{\mathbb{Q}_q}$ gives rise to a map

$$\mathrm{Br}_2(\mathbb{Q}) \to \bigoplus_q \mathrm{Br}_2(\mathbb{Q}_q),$$

and this is injective when q varies over all (finite) rational primes (ignoring the infinite place ∞). This follows from the celebrated Brauer–Hasse–Noether theorem, which tells us that an element of $\mathrm{Br}(\mathbb{Q})$ is trivial provided it is locally trivial everywhere, except possibly at one place (*Hasse reciprocity*; see [Weil 1967, Chapter XIII, Theorem 2]). Now the arguments given in [Serre 1992] apply as in the odd case. (For odd p the archimedean places can be ignored, and by Hasse reciprocity one could allow that p is ramifying [Reichardt 1937].)

Having found a solution of the embedding problem $(G_{\mathbb{Q}}, \varphi, \pi)$ from [Serre 1992, Proposition 2.1.7] it follows that there is a solution E with $\mathrm{Ram}(E) = \mathrm{Ram}(K)$; see also [Scholz 1937, Section 5].

Usually the field $E = \overline{\mathbb{Q}}^{\mathrm{Ker}(\psi)}$ will not be a Scholz field, because condition (S2) may fail. It is the unique solution of $(G_{\mathbb{Q}}, \varphi, \pi)$ with $\mathrm{Ram}(E) = \mathrm{Ram}(K)$ only when $|\mathrm{Ram}(\mathfrak{S}(K))| = d(G)$. In fact, by the Kronecker–Weber theorem $\mathfrak{S}(K)$ is a subfield of a cyclotomic field. Let $q \in \mathrm{Ram}(\mathfrak{S}(K))$, and let P_q be the (unique) subfield of the q-th cyclotomic field $\mathbb{Q}(\zeta_q)$ of absolute degree p, which exists by (S1). Then

$\mathrm{Ram}(P_q) = \{q\}$. There is an epimorphism $\chi_q : G_{\mathbb{Q}} \twoheadrightarrow Z$ with $\overline{\mathbb{Q}}^{\mathrm{Ker}(\chi_q)} = P_q$, and $E_q = \overline{\mathbb{Q}}^{\mathrm{Ker}(\psi\chi_q)}$ is a solution of $(G_{\mathbb{Q}}, \varphi, \pi)$ with $\mathrm{Ram}(E_q) = \mathrm{Ram}(E) = \mathrm{Ram}(K)$. We have $E_q = E$ only when $P_q \subseteq \mathfrak{S}(K)$. Hence uniqueness happens only when $\mathfrak{S}(K) = \prod_{q \in \mathrm{Ram}(\mathfrak{S}(K))} P_q$.

Lemma 2.2. *For any positive integers N, d there exist infinitely many pairwise disjoint d-sets of primes $\{q_1, \dots, q_d\}$ such that each q_i is in $1 + p^N\mathbb{Z}$ and is a p^N-th power in $\mathbb{F}_{q_j}^* = (\mathbb{Z}/q_j\mathbb{Z})^*$ whenever $j \neq i$.*

Let q_1 be one of the infinitely many (Chebotarev) primes which split completely in the p^N-th cyclotomic field $K_1 = \mathbb{Q}(\zeta_{p^N})$. Let q_2 split completely in $K_2 = K_1(\zeta_{q_1}, \sqrt[p^N]{q_1})$, $\dots$, and let finally q_d split completely in $K_d = K_{d-1}(\zeta_{q_{d-1}}, \sqrt[p^N]{q_{d-1}})$. Each q_i is in $1 + p^N\mathbb{Z}$ as it splits completely in $\mathbb{Q}(\zeta_{p^N})$. In

$$K_{i+1} = \mathbb{Q}(\zeta_{p^N}; \zeta_{q_1}, \dots, \zeta_{q_i}; \sqrt[p^N]{q_1}, \dots, \sqrt[p^N]{q_i})$$

the prime q_{i+1} is completely split, whereas $q_1, \dots, q_i$ are ramified ($1 \le i < d$). For $1 \le j \le i$ we have $q_{i+1} \in 1 + q_j\mathbb{Z}$ since it is totally split in $\mathbb{Q}(\zeta_{q_j})$; in this case q_{i+1} obviously is a p^N-th power in $\mathbb{F}_{q_j}^*$. Since q_{i+1} splits completely in $\mathbb{Q}(\zeta_{p^N}, \sqrt[p^N]{q_j})$ for $j \le i$, the congruence $x^{p^N} \equiv q_j \pmod{q_{i+1}}$ is solvable in $\mathbb{Z}$ (Kummer's theorem).

Having found this d-set $\{q_1, \dots, q_d\}$ of primes, let q_{d+1} be a prime splitting totally in $K_{d+1} = K_d(\zeta_{q_d}, \sqrt[p^N]{q_d})$, and proceed in this manner.

Lemma 2.3. *Given positive integers N, d, let $\{q_1, \dots, q_d\}$ be a d-set of primes as constructed in the preceding lemma. Let also n_i be integers with $1 \le n_1 \le n_2 \le \cdots \le n_d \le N$, and let G be abelian of type $(p^{n_1}, \dots, p^{n_d})$. For $i = 1, \dots, d$ let P_i be the (unique) subfield of $\mathbb{Q}(\zeta_{q_i})$ of absolute degree p^{n_i} (which exists). Then $K = P_1 \cdots P_d$ is a Scholz field with respect to N realizing G as Galois group over $\mathbb{Q}$, with* $\mathrm{Ram}(K) = \{q_1, \dots, q_d\}$.

By construction and the decomposition law in cyclotomic fields, for each $i = 1, \dots, d$ the prime q_i is in $1 + p^N\mathbb{Z}$, is totally ramified in P_i and is completely split in all P_j, $j \neq i$.

Remark. Let $S = \{q_1, \dots, q_d\}$ be as constructed in Lemma 2.2, and let $G_S(p)$ be the absolute Galois group of the maximal p-extension of $\mathbb{Q}$ unramified outside $S \cup \{\infty\}$. By [Fröhlich 1983, Theorem 4.11] $G_S(p)$ maps onto every p-group of rank d, exponent p^N and nilpotency class 2. This solves the minimal ramification problem for p-groups of nilpotency class at most 2 (varying d and N). However, it is easily seen that such groups are semiabelian (so that [Kisilevsky et al. 2010] applies).

By recursive definition a finite group G is *semiabelian* if either G is abelian or $G = AH$ for some normal abelian subgroup A of G and some *proper* semiabelian subgroup H. So G is an epimorphic image of a split group extension with abelian kernel. In an analogous manner finite solvable groups might be called *seminilpotent* (see Proposition 2.2.4 and Claim 2.2.5 in [Serre 1992], and the elegant proof of this claim in the case of abelian kernels).

The bound $|\mathrm{Ram}(K)| = d(G)$ can be diminished if one allows also wild ramification. Examples for $p = 2$ are the (semiabelian) dihedral, semidihedral and modular 2-groups, with $\mathrm{Ram}(K) = \{2\}$, whereas the (generalized) quaternion 2-groups require a further (odd) ramifying prime; e.g., see [Schmid 2014].

3. Disposition 2-groups

For subgroups X, Y of a group G we let $[X, Y]$ be the subgroup of G generated by the commutators $[x, y] = x^{-1}y^{-1}xy = x^{-1}x^{y}$ $(x \in X,\ y \in Y)$. We define recursively $[x_1, \dots, x_{n+1}] = [[x_1, \dots, x_n], x_{n+1}]$, and $\gamma_1(G) = G,\ \gamma_{n+1}(G) = [\gamma_n(G), G]$ describing the lower central series of G. As usual we write $G' = \gamma_2(G) = [G, G]$. We also denote by $Z(G)$ the centre of G, and we write $Z'(G) = Z(G) \cap G'$.

The lower (central Frattini) 2-series of the group G is defined inductively by $\lambda_1(G) = G$ and $\lambda_{n+1}(G) = [\lambda_n(G), G]\lambda_n(G)^2$. If $G \neq 1$ is a finite 2-group then $\Phi(G) = \lambda_2(G)$ is the Frattini subgroup of G, and G has 2-class c if $\lambda_{c+1}(G) = 1$ but $\lambda_c(G) \neq 1$. Letting F_d be "the" free group of finite rank $d \geq 1$, for any integer $c \geq 1$ the quotient

$$G_d^c = G_d^c(2) = F_d/\lambda_{c+1}(F_d)$$

is a finite 2-group of rank d and 2-class c, which will be called a "disposition group" (with respect to the prime $p = 2$; of course we could replace F_d by the free pro-2-group of rank d). Every (finite) 2-group G of rank $\leq d$ and 2-class $\leq c$ is an epimorphic image of G_d^c. In fact, by the universal property of free groups, and by Burnside's basis theorem, any epimorphism $F_d/\lambda_2(F_d) \twoheadrightarrow G/\lambda_2(G)$ lifts to an epimorphism $\pi : F_d \twoheadrightarrow G$, and $\lambda_{c+1}(G) = 1$ implies that $\lambda_{c+1}(F_d) \subseteq \mathrm{Ker}(\pi)$.

The disposition p-groups have been studied in the literature quite intensively (see for instance [Shafarevich 1989; Neukirch et al. 2000; Schmid 2017]). We summarize the basic facts (for the somewhat exceptional case $p = 2$).

Proposition 3.1. *Let $G = G_d^c$ for $d \geq 2$ and $c \geq 2$, and let*

$$\ell_d^\kappa = \frac{1}{k}\sum_{k|\kappa} \mu(k) d^{\kappa/k}$$

for $\kappa = 1, \dots, c$ (where $\mu(k)$ denotes the Möbius function). The group G has rank d, exponent 2^c and nilpotency class c, with centre $Z(G) = \lambda_c(G)$. So both $V = G/\Phi(G)$ and $Z(G)$ are $\mathbb{F}_2$-vector spaces (often written additively):

(a) *The assignment $x\Phi(G) \mapsto x^{2^{c-1}}$ for $x \in G$ is a well-defined injection of V into $Z(G)$. Fix a basis $\{x_i\Phi(G)\}_{i=1}^d$ of V $(x_i \in G)$, and let $z_i = x_i^{2^{c-1}}$ and $L_d^1 = \langle z_1, \dots, z_d\rangle$. Then $Z(G) = L_d^1 \oplus Z'(G)$, and $x_i\Phi(G) \mapsto z_i$ defines a linear isomorphism $\psi_d^1 : V \xrightarrow{\sim} L_d^1$.*

(b) *For $\kappa \in \{2, \dots, c\}$ the $2^{c-\kappa}$-th power map on $\gamma_\kappa(G)$ is a homomorphism with kernel $\gamma_{\kappa+1}(G)\gamma_\kappa(G)^2$ and image $L_d^\kappa = \gamma_\kappa(G)^{2^{c-\kappa}}$ in $Z'(G)$, and we have the (natural) direct decomposition*

$$Z'(G) = L_d^2 \oplus \cdots \oplus L_d^c.$$

The "Lie module" L_d^κ has the $\mathbb{F}_2$-dimension ℓ_d^κ, and the assignments $\bar{x}_1 \otimes \cdots \otimes \bar{x}_\kappa \mapsto [x_1, \dots, x_\kappa]^{2^{c-\kappa}}$, for $x_i \in G$ and $\bar{x}_i = x_i\Phi(G)$, define an epimorphism $\psi_d^\kappa : V^{\otimes\kappa} \twoheadrightarrow L_d^\kappa$.

Proposition 3.1 is contained in the (Main) Theorem of [Schmid 2017] (where one can also find an explanation of the notion "Lie module"). Actually we shall only use the $\mathbb{F}_2$-vector space decomposition

$Z(G_d^c) = \bigoplus_{\kappa=1}^{c} L_d^{\kappa}$, together with the epimorphisms ψ_d^{κ} described above. We emphasize that ψ_d^1 depends on the choice of a basis for $G_d^c/\Phi(G_d^c)$ (in the $p=2$ case).

Lemma 3.2. *Let $G_\delta^c = G_\delta^c$ and $G_d^c = G_d^c$ be disposition 2-groups with $\delta > d \geq 2$ and $c \geq 2$, and let $\alpha : G_\delta^c/\Phi(G_\delta^c) \twoheadrightarrow G_d^c/\Phi(G_d^c)$ be an epimorphism. Then all lifts of α to G_δ^c (which exist) restrict to the same epimorphism $\alpha_z : Z(G_\delta^c) \twoheadrightarrow Z(G_d^c)$, and α_z maps $Z'(G_\delta^c)$ onto $Z'(G_d^c)$ respecting the direct decompositions into Lie modules.*

This lemma follows from [Schmid 2017, Proposition 3]. If α sends basis vectors to basis vectors or zero, and L_δ^1, L_d^1 are computed with regard to these bases (see above), then α_z maps L_δ^1 onto L_d^1. The following lemma is Proposition 4 in [Schmid 2017].

Lemma 3.3. *Let $H = G_d^c$ with $d \geq 2$, $c \geq 2$, and let $G = H/Z(H)$ ($\cong G_d^{c-1}$). There is a natural (transgression) isomorphism* $\mathrm{Hom}(Z(H), \mathbb{F}_2) \xrightarrow{\sim} H^2(G, \mathbb{F}_2)$. *Choose a basis $\{\rho_\tau\}$ of $H^2(G, \mathbb{F}_2)$, and let H_τ for each τ be an extension of G by $Z_2 \cong \mathbb{F}_2$ with cohomology class ρ_τ. Then the fibre product of the H_τ amalgamating G is isomorphic to H.*

4. The Scholz obstructions

Let $d \geq 2$, $c \geq 2$, and let $G = G_d^{c-1}$. Let N be an integer with $N \geq c$, and suppose we have a strong Scholz field K with respect to N realizing G as Galois group over $\mathbb{Q}$. Let $\{\rho_\tau\}$ be a basis of $H^2(G, \mathbb{F}_2)$. For any τ let H_τ be a (central) extension of G by $Z_2 \cong \mathbb{F}_2$ with cohomology class ρ_τ, and let E_τ be a solution of the corresponding nonsplit embedding problem with $\mathrm{Ram}(E_\tau) = \mathrm{Ram}(K)$ (see Proposition 2.1). The compositum $E = \prod_\tau E_\tau$ is a normal number field containing K with $\mathrm{Ram}(E) = \mathrm{Ram}(K)$, and $H = \mathrm{Gal}(E\,|\,\mathbb{Q})$ is the fibre product of the H_τ amalgamating G. Hence $H \cong G_d^c$ by Lemma 3.3.

For proof-technical reasons we assume in what follows merely that the E_τ are chosen such that if there is $q \in \mathrm{Ram}(E)\backslash\mathrm{Ram}(K)$, then $q \in 1+2^N\mathbb{Z}$ and q splits completely in $\mathfrak{S}(K)$. We also will choose the basis $\{\rho_\tau\}$ of $H^2(G, Z_2)$ suitably, without altering the field E (see below). Let $t = \dim H^2(G, Z_2) = \dim Z(H)$ (Lemma 3.3).

By Proposition 3.1 we have $Z(H) = \lambda_c(H) \subseteq \Phi(H)$ and $H/Z(H) \cong G$. Hence by assumption $\mathfrak{S}(E) = \mathfrak{S}(K) = P_1 \cdots P_d$, where the $\mathrm{Ram}(P_i)$ are pairwise disjoint and of the same cardinality. Let b_i be the discriminant of P_i. By (S1) 2 is unramified in $P_i = \mathbb{Q}(\sqrt{b_i})$ and hence

$$b_i = \prod_{q \in \mathrm{Ram}(P_i)} q \in 1 + 2^N\mathbb{Z}.$$

Given a prime q we simply write $I_q \subseteq D_q$ for the inertia and decomposition groups in H of some fixed prime $\mathfrak{Q}$ of E above q (determined up to H-conjugacy). The images of these groups in $H/\Phi(H) = \mathrm{Gal}(\mathfrak{S}(E)\,|\,\mathbb{Q})$ and their intersections with $Z(H) = \mathrm{Gal}(E\,|\,K)$ are independent of the choice of $\mathfrak{Q}$. Recall that I_q is cyclic (by tame ramification).

Lemma 4.1. *Let $I_q = \langle x_i \rangle$ be the inertia group in H of some $q \in \mathrm{Ram}(P_i)$ ($1 \leq i \leq d$). Then $\bar{x}_i = x_i\Phi(H)$ and $z_i = x_i^{2^{c-1}}$ are independent of the choice of the prime q in $\mathrm{Ram}(P_i)$, and $\{x_i\}_{i=1}^d$ is a minimal system* *of generators for H. For any $q \in \mathrm{Ram}(P_i)$ we have $I_q \cap Z(H) = \langle z_i \rangle$, and the primes of K above q are unramified in $E^{\langle z_i \rangle}$.*

This is immediate from the structure of $\mathfrak{S}(E) = E^{\Phi(H)}$ and from Proposition 3.1. We also get that $L_d^1 = \langle z_1, \dots, z_d \rangle$ is an $\mathbb{F}_2$-subspace of $Z(H)$ of dimension $d = \ell_d^1$ complementary to $Z'(H) = H' \cap Z(H)$ in $Z(H)$. Hence letting $E^{\perp} = \bigcap_{i=1}^{d} E^{\langle z_i \rangle}$ be the fixed field of this L_d^1 we have $E = E^{\perp} \cdot E^{Z'(H)}$ and $E^{\perp} \cap E^{Z'(H)} = K$. Let also

$$E(i) = \bigcap_{j \neq i} E^{\langle z_j \rangle} \cap E^{Z'(H)}$$

for each $i = 1, \dots, d$. Now choose a basis $\{\chi_\tau\}$ of $\mathrm{Hom}(Z(H), \mathbb{F}_2)$ such that $E_\tau = E^{\mathrm{Ker}(\chi_\tau)}$ either is contained in $E^{\perp}$ or is equal to $E(i)$ for some i, and let $\{\rho_\tau\}$ correspond to $\{\chi_\tau\}$ under the transgression isomorphism $\mathrm{Hom}(Z(H), \mathbb{F}_2) \xrightarrow{\sim} H^2(G, \mathbb{F}_2)$ (Lemma 3.3).

For each τ write $E_\tau = K(\sqrt{\mu_\tau})$ for some $\mu_\tau \in K^*$ (determined mod $(K^*)^2$). Then every group extension of G by Z_2 with cohomology class ρ_τ is realized as $K(\sqrt{m\mu_\tau})$ for some (square-free) integer $m \neq 0$, because it is obtained by Baer addition of H_τ with the split extension of G by Z_2 realized as $K(\sqrt{m}) = \mathbb{Q}(\sqrt{m}) \cdot K$ (with $\mathbb{Q}(\sqrt{m}) \nsubseteq K$; alternately, multiply $\psi_\tau : G_{\mathbb{Q}} \twoheadrightarrow H_\tau$ with $\chi_m : G_{\mathbb{Q}} \twoheadrightarrow Z_2$ having $\overline{\mathbb{Q}}^{\mathrm{Ker}(\chi_m)} = \mathbb{Q}(\sqrt{m})$).

Let $q \in \mathrm{Ram}(P_i)$ for some i, and let $I_q \subseteq D_q$ be as above. For any prime $\mathfrak{q}$ of K above q, determined up to G-conjugacy, the Frobenius

$$\bar{\phi}_q = \left(\frac{E^{\langle z_i \rangle} \,|\, K}{\mathfrak{q}} \right)$$

(Artin symbol) is an element of $\mathrm{Gal}(E^{\langle z_i \rangle} \,|\, K)$ and independent of the choice of $\mathfrak{q}$ above q. Since q is busy in the Scholz field K, both I_q and D_q have the same image (of order 2^{c-1}) in $H/Z(H) \cong G$. Hence $D_q = I_q(D_q \cap Z(H))$ and

$$D_q \cap Z(H) = \langle z_i \rangle \times \langle \sigma_q \rangle$$

for some element σ_q (of order 2 or 1) mapping onto $\bar{\phi}_q$, which in turn maps onto the generator of D_q/I_q. If $\sigma_q \neq 0$ (additive notation), we may replace σ_q by $z_i + \sigma_q$. In spite of this ambiguity we call σ_q "the" *Scholz obstruction* for E associated to q. This will be no problem since we only shall consider the restrictions of σ_q to fields $E_\tau \subseteq E^{\langle z_i \rangle}$.

Proposition 4.2. *Let $\sigma_i = \sum_{q \in \mathrm{Ram}(P_i)} \sigma_q$, and assume that $\sigma_i = 0$ (or that $\sigma_i \in \langle z_i \rangle$) for all $i = 1, \dots, d$. Then there exist infinitely many pairwise disjoint t-sets $\{p_1, \dots, p_t\}$ of rational primes such that $\widehat{E} = \prod_{\tau=1}^{t} K(\sqrt{p_\tau e \mu_\tau})$ is a strong Scholz field with respect to N admitting G_d^c as Galois group over $\mathbb{Q}$ and having $\mathrm{Ram}(\widehat{E}) = \mathrm{Ram}(K) \cup \{p_1, \dots, p_t\}$.*

Proof. We argue by induction. Suppose that either $K_0 = K$ or that $K_0 = \prod_{\tau'} K(\sqrt{p_{\tau'} e \mu_{\tau'}})$ is a strong Scholz field with respect to N (with corresponding $\mathrm{Ram}(K_0)$) for certain τ' and primes $p_{\tau'}$, but that there is still some τ different from all these τ'. We prove that there are infinitely many primes $p_\tau \in 1 + 2^N\mathbb{Z}$

which split completely in K_0 such that $\hat{E}_0 = K_0(\sqrt{p_\tau e \mu_\tau})$ is a strong Scholz field with respect to N with $\mathrm{Ram}(\hat{E}_0) = \mathrm{Ram}(K_0) \cup \{p_\tau\}$.

Let $G_0 = \mathrm{Gal}(K_0|\mathbb{Q})$. By Lemma 3.3 this represents a central Frattini extension of G and is an epimorphic image over G of $H \cong G_d^c$. In particular $\mathfrak{S}(K_0) = \mathfrak{S}(K) = \mathfrak{S}(E)$. By construction the image ρ_0 of ρ_τ under the inflation map $\inf : H^2(G, \mathbb{F}_2) \to H^2(G_0, \mathbb{F}_2)$ is nontrivial. Let $E_0 = K_0 E_\tau = K_0(\sqrt{\mu_\tau})$ and $H_0 = \mathrm{Gal}(E_0|\mathbb{Q})$. For $x \in G_0$ choose an inverse image $\tilde{x} \in H_0$, and observe that $E_0 = K_0(\sqrt{\mu_\tau}^{\tilde{x}})$ and $(\sqrt{\mu_\tau}^{\tilde{x}})^2 = (\sqrt{\mu_\tau}^2)^{\tilde{x}} = \mu_\tau^x$. Hence

$$\mu_\tau^x = \beta_x^2 \mu_\tau$$

for some $\beta_x \in K_0^*$. Let $\mathfrak{q}_0 | q$ be primes of $K_0|\mathbb{Q}$. For the $\mathfrak{q}_0$-adic valuation we have $v_{\mathfrak{q}_0}(\mu_\tau) = v_{\mathfrak{q}_0{}^x}(\mu_\tau^x) = 2 \cdot v_{\mathfrak{q}_0{}^x}(\beta_x) + v_{\mathfrak{q}_0{}^x}(\mu_\tau)$. This shows that the fractional ideal (μ_τ) of K_0 generated by μ_τ splits into a square of a fractional ideal $\mathfrak{b}$ and a G_0-invariant square-free (integral) ideal of K_0, the latter being decomposed into products of G_0-conjugates of primes of K_0 ramified over $\mathbb{Q}$ and those which are not. Hence may write uniquely

$$(\mu_\tau) = \mathfrak{b}^2 \cdot \mathfrak{D} \cdot (e),$$

where $\mathfrak{D}$ is a G_0-invariant ideal of K_0 composed of pairwise distinct prime ideals of K_0 ramified over the rationals, and where e is a square-free positive integer relatively prime to the discriminant of K_0.

Let again $\mathfrak{q}_0 | q$ be primes of $K_0|\mathbb{Q}$. By [Hecke 1981, Theorem 120], $\mathfrak{q}_0$ is ramified in $E_0 = K_0(\sqrt{\mu_\tau})$ if and only if $v_{\mathfrak{q}_0}(\mu_\tau)$ is odd, except possibly when $\mathfrak{q}_0$ is dyadic (lying above 2). But 2 is not ramified in the Scholz field K_0 by (S1), and $\mathrm{Ram}(E_\tau) \subseteq 1 + 2^N \mathbb{Z}$ by assumption. So this exception does not happen. Hence $\mathfrak{q}_0$ is ramified in E_0 if and only if either $\mathfrak{q}_0$ appears in $\mathfrak{D}$ or q is a divisor of e. Each rational prime dividing e is unramified in K_0 but ramified in $E_0 = K_0 E_\tau$ and hence in E_τ. The prime divisors of e (if any) therefore are in $\mathrm{Ram}(E) \setminus \mathrm{Ram}(K)$, thus belong to $1 + 2^N \mathbb{Z}$ and split completely in $\mathfrak{S}(K)$ (by our convention).

Let R be the subset of $\mathrm{Ram}(K)$ consisting of those rational primes q for which the primes $\mathfrak{q}$ of K above q do not ramify in E_τ. We claim that $R = R(\rho_\tau)$ is an invariant of the cohomology class ρ_τ. By Hecke $v_\mathfrak{q}(\mu)$ is even, and knowing that $q \neq 2$ one just has to show that $v_\mathfrak{q}(m\mu) = v_\mathfrak{q}(m) + v_\mathfrak{q}(\mu)$ also is even for any integer $m \neq 0$. But this is clear since $v_\mathfrak{q}(m) = e(\mathfrak{q}|q) \cdot v_q(m)$ and the ramification index $e(\mathfrak{q}|q) = |I_\mathfrak{q}|$ is a proper power of 2. Similarly, the set R_0 of rational primes ramified in K_0 but not in $E_0 = K_0 E_\tau$ is an invariant of $\rho_0 = \inf(\rho_\tau)$.

Now let $q \in \mathrm{Ram}(\mathfrak{S}(K_0)) = \mathrm{Ram}(\mathfrak{S}(K))$, say $q \in \mathrm{Ram}(P_i)$. We assert that $q \in R$ if and only if $q \in R_0$. Let $\mathfrak{q}_0 | \mathfrak{q}$ be primes of $K_0|K$ above q. We have $q \in R$ if and only if $\mathfrak{q}$ is unramified in E_τ ($E_\tau \subseteq E^{\langle z_i \rangle}$), and then (obviously) $\mathfrak{q}_0$ is unramified in $E_0 = K_0 E_\tau$ and hence $q \in R_0$. Conversely, suppose that $\mathfrak{q}_0$ is unramified in E_0 ($q \in R_0$). Assume that $q \notin R$; that is, $\mathfrak{q} = \mathfrak{q}_0 \cap K$ is ramified in E_τ. Then $E_\tau = E(i)$ by construction. Moreover then $\mathfrak{q}$ is unramified in $E_{\tau'} = K(\sqrt{\mu_{\tau'}})$ for all $\tau' \neq \tau$ since then $E_{\tau'} \subseteq E^{\langle z_i \rangle}$. As this is a property of the cohomology class ρ, we know $\mathfrak{q}$ is unramified in the fields $K(\sqrt{p_{\tau'} e \mu_{\tau'}})$ generating K_0. Consequently $\mathfrak{q}$ is unramified in K_0, whence splits completely in the Scholz field K_0. It follows that $\mathfrak{q}_0$ must be ramified in E_0. This is the desired contradiction. (The "converse" statement is not true in general; a corresponding argument is missing in [Shafarevich 1954, p. 121].)

Let $\mathfrak{q}_0 \,|\, q$ be primes of $K_0 \,|\, \mathbb{Q}$ with $q \in R_0$. Then the Frobenius

$$\phi_q = \left(\frac{E_0 \,|\, K_0}{\mathfrak{q}_0} \right)$$

is defined, which is a central element in $H_0 = \mathrm{Gal}(E_0 \,|\, \mathbb{Q})$ and so depends only on q. Independent of the choice of the square root $\sqrt{\mu_\tau}$ we have

$$(\sqrt{\mu_\tau})^{\phi_q} = \left(\frac{\mu_\tau}{\mathfrak{q}_0} \right) \sqrt{\mu_\tau},$$

where

$$\left(\frac{\mu_\tau}{\mathfrak{q}_0} \right) = \pm 1$$

is the Legendre symbol (quadratic residue symbol). Since

$$\left(\frac{\mu_\tau}{\mathfrak{q}_0} \right) = \left(\frac{\mu_\tau^x}{\mathfrak{q}_0{}^x} \right) = \left(\frac{\beta_x^2 \mu_\tau}{\mathfrak{q}_0{}^x} \right) = \left(\frac{\mu_\tau}{\mathfrak{q}_0{}^x} \right)$$

for each $x \in G_0$, and since q is the absolute norm of $\mathfrak{q}_0$ by (S2), it is appropriate to write this symbol as $\left[\frac{\mu_\tau}{q} \right]$ (like in [Shafarevich 1954]). As usual the Legendre symbol is extended multiplicatively to products of nondyadic primes in the denominator (Jacobi symbol), yielding also certain extensions of the Shafarevich symbol. For an integer $m \neq 0$ the symbol $\left[\frac{m\mu_\tau}{q} \right]$ is defined since R_0 is an invariant of ρ_0, and if m is not divisible by q then

$$\left[\frac{m\mu_\tau}{q} \right] = \left(\frac{m}{q} \right) \left[\frac{\mu_\tau}{q} \right].$$

In this case $\mathfrak{q}_0 \,|\, q$ are unramified in $K_0(\sqrt{m}) \,|\, \mathbb{Q}(\sqrt{m})$ and

$$\left(\frac{K_0(\sqrt{m}) \,|\, K_0}{\mathfrak{q}_0} \right) \quad \text{restricts to} \quad \left(\frac{\mathbb{Q}(\sqrt{m}) \,|\, \mathbb{Q}}{(q)} \right)$$

since q is busy in K_0 by (S2). Thus $\left(\frac{m}{q} \right) = \left(\frac{m}{\mathfrak{q}_0} \right)$, and the result follows since evidently

$$\left(\frac{m}{\mathfrak{q}_0} \right) \left(\frac{\mu_\tau}{\mathfrak{q}_0} \right) = \left(\frac{m\mu_\tau}{\mathfrak{q}_0} \right).$$

Let again $q \in \mathrm{Ram}(P_i)$ for some i, and let $q \in R_0$. Using that q is busy in the Scholz field K_0, the restrictions to E_τ of σ_q and of the Frobenius ϕ_q introduced above agree. Therefore

$$(\sqrt{\mu_\tau})^{\sigma_q} = (\sqrt{\mu_\tau})^{\phi_q} = \left[\frac{\mu_\tau}{q} \right] \sqrt{\mu_\tau}.$$

We know that $q \in R \cap \mathrm{Ram}(P_i)$, and this implies that all primes in $\mathrm{Ram}(P_i)$ belong to R and hence to R_0 (see Lemma 4.1). Therefore

$$\left[\frac{\mu_\tau}{b_i} \right] = \prod_{q \in \mathrm{Ram}(P_i)} \left[\frac{\mu_\tau}{q} \right]$$

is defined, and

$$(\sqrt{\mu_\tau})^{\sigma_i} = \left[\frac{\mu_\tau}{b_i}\right]\sqrt{\mu_\tau}.$$

Thus

$$\left[\frac{\mu_\tau}{b_i}\right] = 1$$

by the hypothesis of the proposition.

Recall that e is coprime to the discriminant of K_0 and so not divisible by any prime in R_0. By the Chinese remainder theorem there is an *odd* integer m such that

$$\left(\frac{m}{q}\right) = \left[\frac{e\mu_\tau}{q}\right] = \left(\frac{e}{q}\right)\left[\frac{\mu_\tau}{q}\right]$$

for each $q \in R_0$. Then m is prime to every $q \in R_0$ and

$$\left[\frac{me\mu_\tau}{q}\right] = \left(\frac{m}{q}\right)\left[\frac{e\mu_\tau}{q}\right] = 1.$$

Since $R_0 \subseteq \mathrm{Ram}(K_0) \subseteq 1 + 2^N\mathbb{Z}$ (with $N \geq c \geq 2$), replacing m by $-m$ if necessary, we may assume that $m > 0$. We assert that

$$\left(\frac{b_i}{m}\right) = 1$$

whenever $\mathrm{Ram}(P_i) \subseteq R_0$. By quadratic reciprocity

$$\left(\frac{b_i}{m}\right) = \left(\frac{m}{b_i}\right),$$

because b_i and m are relatively prime positive odd integers and $b_i \in 1 + 2^N\mathbb{Z}$. We have

$$\left(\frac{e}{b_i}\right) = \left(\frac{b_i}{e}\right) = 1$$

since the primes dividing e are in $1 + 2^N\mathbb{Z}$ and split completely in $P_i = \mathbb{Q}(\sqrt{b_i})$. Consequently

$$1 = \prod_{q \in \mathrm{Ram}(P_i)} \left[\frac{me\mu_\tau}{q}\right] = \left[\frac{me\mu_\tau}{b_i}\right] = \left(\frac{m}{b_i}\right)\left(\frac{e}{b_i}\right)\left[\frac{\mu_\tau}{b_i}\right] = \left(\frac{m}{b_i}\right),$$

as required.

Let $T = \prod_{q \in R_0} \mathbb{Q}(\sqrt{q})$. Using that $\mathfrak{S}(T) = T$, $\mathrm{Ram}(\mathbb{Q}(\zeta_{2^N})) = \{2\}$ and $2 \notin R_0 = \mathrm{Ram}(T)$ we get

$$T \cap K_0(\zeta_{2^N}) = \prod_i{}' \mathbb{Q}(\sqrt{b_i}),$$

where the product is taken over those indices i where $\mathrm{Ram}(P_i) \subseteq R_0$. (This is always true when $E_\tau \subseteq E^\perp$, and if $E_\tau = E(i)$ for some i then $\mathrm{Ram}(P_j) \subseteq R_0$ for all $j \neq i$ and $\mathrm{Ram}(P_i) \cap R_0 = \varnothing$.) By construction the Artin automorphism

$$\phi = \left(\frac{T\,|\,\mathbb{Q}}{(m)}\right)$$

is defined and is trivial on $\prod_i' \mathbb{Q}(\sqrt{b_i})$. Hence ϕ can be extended to an automorphism $\hat{\phi}$ of $T \cdot K_0(\zeta_{2^N})$ which is trivial on $K_0(\zeta_{2^N})$. By Chebotarev's density theorem there are infinitely many rational primes p_τ which are unramified in $T \cdot K_0(\zeta_{2^N})$ and for which some prime above p_τ has $\hat{\phi}$ as Frobenius automorphism. Then p_τ splits completely in $K_0(\zeta_{2^N})$; that is, p_τ splits completely in K_0 and belongs to $1 + 2^N\mathbb{Z}$. Moreover

$$\phi = \left(\frac{T \,|\, \mathbb{Q}}{(p_\tau)}\right).$$

Thus

$$\left(\frac{q}{p_\tau}\right) = \left(\frac{q}{m}\right)$$

for all $q \in R_0$, in view of the action of ϕ on $\mathbb{Q}(\sqrt{q})$ (and consistency of the Artin symbol). But

$$\left(\frac{q}{p_\tau}\right) = \left(\frac{p_\tau}{q}\right) \quad \text{and} \quad \left(\frac{q}{m}\right) = \left(\frac{m}{q}\right)$$

by quadratic reciprocity. Consequently

$$\left[\frac{p_\tau e\mu_\tau}{q}\right] = \left(\frac{p_\tau}{q}\right)\left[\frac{e\mu_\tau}{q}\right] = \left(\frac{m}{q}\right)\left[\frac{e\mu_\tau}{q}\right] = \left[\frac{me\mu_\tau}{q}\right] = 1.$$

Let $\widehat{E}_0 = K_0(\sqrt{p_\tau e\mu_\tau})$. By the above, every prime in R_0 is busy in $\widehat{E}_0$. From

$$(p_\tau e\mu_\tau) = (e\mathfrak{b})^2 \cdot \mathfrak{D} \cdot (p_\tau)$$

we infer that $\mathrm{Ram}(\widehat{E}_0) = \mathrm{Ram}(K_0) \cup \{p_\tau\}$ (Hecke; 2 does not ramify in $\widehat{E}_0$ as it does not ramify in $E_0 = K_0 E_\tau$ or in $\mathbb{Q}(\sqrt{p_\tau e})$). Consequently $\widehat{E}_0$ is a (strong) Scholz field with respect to N. The proposition follows by induction and by appealing to Lemma 3.3. □

5. The shrinking process

We are going to construct Scholz fields fulfilling the assumptions made in Proposition 4.2. As above we consider disposition 2-groups G_d^c. Arguing by induction on the 2-class c (varying d) this will prove the theorem. For $c = 1$ (or $d = 1$) Lemmas 2.2 and 2.3 apply, in which case we may define the polynomial $f_c = X$. So let $N \ge c \ge 2$ be integers. We assume that for every $d \ge 2$ there are infinitely many strong Scholz fields K_d^{c-1} with respect to N with pairwise coprime discriminants admitting G_d^{c-1} as Galois group over the rationals, all these fields having the property that $|\mathrm{Ram}(K_d^{c-1})| \le f_{c-1}(d)$ for some (unique) polynomial $f_{c-1} \in \mathbb{Z}[X]$ with $\deg f_{c-1} = (c+2)!/24$.

Fixing $d \ge 2$ we let $\delta = r \cdot d$ where

$$r = 2d^2 \sum_{\kappa=1}^{c} \kappa \cdot \ell_d^\kappa$$

(see Proposition 3.1 for notation). We know that $r = r(d)$ is an integral polynomial in d of degree $c + 2$. By our inductive hypothesis there is a strong Scholz field $K_\delta = K_\delta^{c-1}$ with respect to N admitting G_δ^{c-1}

as Galois group over the rationals. Indeed there are infinitely many such fields with pairwise coprime discriminants. By Proposition 2.1 and Lemma 3.3 we can embed K_δ into a normal number field E_δ with group G_δ^c having $\mathrm{Ram}(E_\delta)=\mathrm{Ram}(K_\delta)$. In particular $K_\delta = E_\delta^{Z(G_\delta^c)}$ and

$$\mathfrak{S}(E_\delta)=\mathfrak{S}(K_\delta)=\prod_{j=1}^{r}\prod_{i=1}^{d} P_{ij},$$

where the $\mathrm{Ram}(P_{ij})$ have the same cardinality and are pairwise disjoint. Adapted to this decomposition there is a minimal system $\{x_{ij}\}_{i,j}$ of generators of G_δ^c such that the image $\bar{x}_{ij}$ in $W=G_\delta^c/\Phi(G_\delta^c)$ of x_{ij} generates the image in W of the inertia group I_q in G_δ^c for any $q\in\mathrm{Ram}(P_{ij})$, and $z_{ij}=x_{ij}^{2^{c-1}}$ has order 2 and generates $I_q\cap Z(G_\delta^c)$ (see Lemma 4.1).

For every $q\in\mathrm{Ram}(\mathfrak{S}(E_\delta))$ we choose a Scholz obstruction σ_q for E_δ (determined by q up to adding z_{ij} if $\sigma_q\neq 0$ and $q\in\mathrm{Ram}(P_{ij})$). Define

$$\sigma_{ij}=\sum_{q\in\mathrm{Ram}(P_{ij})}\sigma_q$$

for each pair i,j. Let L_δ^1 be the subspace of $Z(G_\delta^c)$ generated by all the z_{ij}, and let $\psi_\delta^1: W \xrightarrow{\sim} L_\delta^1$ be the linear map given by $\bar{x}_{ij}\mapsto z_{ij}$ for all i,j. By Proposition 3.1 we have the decomposition $Z(G_\delta^c)=\bigoplus_{\kappa=1}^{c} L_\delta^\kappa$ into $\mathbb{F}_2$-vector spaces. We also introduce a "target" disposition 2-group G_d^c of rank d and class c with generators $x_1,\dots,x_d$, yielding the basis $\bar{x}_i=x_i\Phi(G_d^c)$ of $V=G_d^c/\Phi(G_d^c)$, and let L_d^1 be the subspace of $Z(G_d^c)$ generated by the $z_i=x_i^{2^{c-1}}$ $(1\le i\le d)$. Then we have again the vector space decomposition $Z(G_d^c)=\bigoplus_{\kappa=1}^{c} L_d^\kappa$. Let $\psi_d^1: V\xrightarrow{\sim} L_d^1$ be the linear isomorphism given by $\bar{x}_i\mapsto z_i$ for each i, and define the epimorphisms $\psi_\delta^\kappa: W^{\otimes\kappa}\twoheadrightarrow L_\delta^\kappa$ and $\psi_d^\kappa: V^{\otimes\kappa}\twoheadrightarrow L_d^\kappa$ for $2\le\kappa\le c$ as in Proposition 3.1.

Now let $\alpha=(a_j)$ be any nontrivial r-tuple in $\mathbb{F}_2^{(r)}$. We shall also write $\alpha: W\twoheadrightarrow V$ for the (surjective) linear map given by $\alpha(\bar{x}_{ij})=a_j\bar{x}_i$ for all pairs i,j (additive notation). By Lemma 3.2 every lift of α to G_δ^c gives rise to the same epimorphism $\alpha_z: Z(G_\delta^c)\twoheadrightarrow Z(G_d^c)$, and α_z respects the corresponding vector space decompositions. From Proposition 3.1 it follows that $\alpha_z\circ\psi_\delta^\kappa=\psi_d^\kappa\circ\alpha^{\otimes\kappa}$ for each $\kappa=1,\dots,c$ (where $\alpha^{\otimes\kappa}: W^{\otimes\kappa}\twoheadrightarrow V^{\otimes\kappa}$ is the κ-th tensor power of α). In particular $\alpha_z(z_{ij})=a_jz_i$ for all i,j (additive notation).

Though irrelevant for our purposes, but following [Shafarevich 1954], we consider the "canonical" epimorphism $\pi(\alpha): G_\delta^c\twoheadrightarrow G_d^c$ given by mapping x_{ij} onto x_i for all i if $a_j=1$ and to 1 if $a_j=0$. This is a distinguished lift of α to G_δ^c. (Writing $G_\delta^c=F_\delta/\lambda_{c+1}(F_\delta)$ and letting $\{t_{ij}\}$ be a basis of the free group, there is an automorphism of G_δ^c sending x_{ij} to $t_{ij}\lambda_{c+1}(F_\delta)$ for all i,j. Then $\pi(\alpha)$ is given via the assignments $t_{ij}\mapsto x_i$ if $a_j=1$ and $t_{ij}\mapsto 1$ otherwise.) Let

$$E(\alpha)=E_\delta^{\mathrm{Ker}(\pi(\alpha))}\quad\text{and}\quad K(\alpha)=E(\alpha)\cap K_\delta.$$

Obviously $K(\alpha)$ is a Scholz field with respect to N; condition (S2) might fail for the field $E(\alpha)$.

It is convenient to identify $\mathrm{Gal}(E(\alpha)\,|\,\mathbb{Q})$ with G_d^c through the isomorphism induced by $\pi(\alpha)$. Then every element of G_δ^c is sent by $\pi(\alpha)$ to its restriction on $E(\alpha)$. In particular $K(\alpha) = E(\alpha)^{Z(G_d^c)}$ since $K_\delta = E_\delta^{Z(G_\delta^c)}$ and $\pi(\alpha)$ (resp. α_z) maps $Z(G_\delta^c)$ onto $Z(G_d^c)$. It follows that $\mathrm{Gal}(K(\alpha)\,|\,\mathbb{Q}) \cong G_d^{c-1}$ and that $\mathfrak{S}(K(\alpha)) = \mathfrak{S}(E(\alpha))$.

If there is a prime $q \in \mathrm{Ram}(E(\alpha)) \setminus \mathrm{Ram}(K(\alpha))$, then $q \in \mathrm{Ram}(K_\delta) \subseteq 1 + 2^N\mathbb{Z}$ and q is busy in the Scholz field K_δ. It follows that q splits completely in $K(\alpha)$ (being busy and unramified). In particular q splits completely in $\mathfrak{S}(K(\alpha))$.

We have $\mathfrak{S}(E(\alpha)) = E(\alpha) \cap \mathfrak{S}(E_\delta)$ since $\pi(\alpha)(\Phi(G_\delta^c)) = \Phi(G_d^c)$. For each $i = 1, \dots, d$ we let

$$P_i(\alpha) = E(\alpha) \cap \prod_{j=1}^{r} P_{ij}.$$

If I_q is the inertia group in G_δ^c for some $q \in \mathrm{Ram}(P_{ij})$, then $\pi(\alpha)(I_q)$ maps onto $\langle \bar{x}_i \rangle$ if $a_j = 1$ (which exists) and $\pi(a)(I_q) \subseteq \Phi(G_d^c)$ otherwise. So $P_i(\alpha)$ is the (cyclic) subfield of $\mathfrak{S}(E(\alpha))$ fixed (centralized) by all $\bar{x}_{i'}$ for $i' \neq i$ (but not by $\bar{x}_i$), and $\mathrm{Ram}(P_i(\alpha)) = \biguplus_j' \mathrm{Ram}(P_{ij})$, where j varies over the indices in $\{1, \dots, r\}$ for which $a_j = 1$. Hence we have

$$\mathfrak{S}(K(\alpha)) = \mathfrak{S}(E(\alpha)) = P_1(\alpha) \cdots P_d(\alpha),$$

and we infer that $K(\alpha)$ is *strongly* Scholz.

Let $q \in \mathrm{Ram}(P_i(\alpha))$ for some i. Then $q \in \mathrm{Ram}(P_{ij})$ for a unique j, and $\alpha_z(z_{ij}) = a_j z_i = z_i$. Every prime $\mathfrak{q}$ of K_δ above q is unramified in $E_\delta^{\langle z_{ij}\rangle}$, and $\mathfrak{q}_\alpha = \mathfrak{q} \cap K(\alpha)$ is unramified in $E(\alpha)^{\langle z_i \rangle} = E(\alpha) \cap E_\delta^{\langle z_{ij}\rangle}$. The restriction of

$$\left(\frac{E_\delta^{\langle z_{ij}\rangle}\,|\,K_\delta}{\mathfrak{q}}\right)$$

to $E(\alpha)^{\langle z_i \rangle}$ agrees with

$$\left(\frac{E(\alpha)^{\langle z_i \rangle}\,|\,K(\alpha)}{\mathfrak{q}_a}\right)$$

as q is busy in K_δ. Hence $\alpha_z(\sigma_q)$ may be identified with "the" Scholz obstruction for $E(\alpha)$ associated to q. We have

$$\sum_{q \in \mathrm{Ram}(P_i(\alpha))} \alpha_z(\sigma_q) = \sum_j{}' \alpha_z(\sigma_{ij}),$$

where the sum is taken over all j for which $a_j = 1$.

Consider $Z(G_\delta^c) \otimes L_\delta^1 = \bigoplus_{\kappa=1}^{c} (L_\delta^\kappa \otimes L_\delta^1)$. Let α_z^κ denote the restriction to L_δ^κ of α_z. The map $\alpha_z \otimes \alpha_z^1 : Z(G_\delta^c) \otimes L_\delta^1 \to Z(G_d^c) \otimes L_d^1$ respects the corresponding decompositions, and the diagram

$$\begin{tikzcd}
V^{\otimes\kappa} \otimes V \arrow[d, "\psi_d^\kappa \otimes \psi_d^1"'] & W^{\otimes\kappa} \otimes W \arrow[l, "\alpha^{\otimes\kappa} \otimes \alpha"'] \arrow[d, "\psi_\delta^\kappa \otimes \psi_\delta^1"] \\
L_d^\kappa \otimes L_d^1 & L_\delta^\kappa \otimes L_\delta^1 \arrow[l, "\alpha_z^\kappa \otimes \alpha_z^1"]
\end{tikzcd}$$

commutes for each κ. All maps in this square are surjections. On the obvious bases of $W^{\otimes\kappa+1}$ and $V^{\otimes\kappa+1}$ we have

$$\alpha^{\otimes\kappa+1}(\bar{x}_{i_1,j_1}\otimes\cdots\otimes\bar{x}_{i_{\kappa+1},j_{\kappa+1}})=(\bar{x}_{i_1}\otimes\cdots\otimes\bar{x}_{i_{\kappa+1}})a_{j_1}\cdots a_{j_{\kappa+1}}.$$

Let $z\in Z(G_\delta^c)\otimes L_\delta^1$ with κ-component $z^\kappa\in L_\delta^\kappa\otimes L_\delta^1$, and let $\widehat{z^\kappa}\in W^{\otimes\kappa}\otimes W$ be an inverse image of z^κ with regard to $\psi_\delta^\kappa\otimes\psi_\delta^1$. Given a nontrivial linear form $\chi\in\mathrm{Hom}(L_d^\kappa\otimes L_d^1,\mathbb{F}_2)$, the element $\chi\circ(\psi_d^\kappa\otimes\psi_d^1)\circ\alpha^{\otimes\kappa+1}(\widehat{z^\kappa})$ may be interpreted as evaluation at (a_j) of some homogeneous polynomial of degree $\kappa+1$ in r variables over $\mathbb{F}_2$ determined by $\widehat{z^\kappa}$ and χ. But $(\psi_d^\kappa\otimes\psi_d^1)\circ\alpha^{\otimes\kappa+1}(\widehat{z^\kappa})=(\alpha_z^\kappa\otimes\alpha_z^1)(z^\kappa)$ and so this evaluation only relies on z^κ (and on χ). Hence we may state that $\chi\circ(\alpha_z^\kappa\otimes\alpha_z^1)(z^\kappa)=0$ if (a_j) is a (nontrivial) zero of a certain homogeneous polynomial of degree $\kappa+1$ in r variables over $\mathbb{F}_2$. Varying χ over a basis for $\mathrm{Hom}(L_d^\kappa\otimes L_d^1,\mathbb{F}_2)$ we obtain that $(\alpha_z^\kappa\otimes\alpha_z^1)(z^\kappa)=0$ if (a_j) is a common zero of $\ell_d^\kappa\cdot d$ such polynomials, and we get $(\alpha_z\otimes\alpha_z^1)(z)=0$ if (a_j) is a common zero of $d\sum_{\kappa=1}^c\ell_d^\kappa$ such homogeneous polynomials in r variables over $\mathbb{F}_2$ of respective degrees $\kappa+1=2,\dots,c+1$.

Now consider for each $i=1,\dots,d$ the element $z(i)=\sum_{j=1}^r\sigma_{ij}\otimes z_{ij}$ of $Z(G_\delta^c)\otimes L_\delta^1$. Since by definition $r>d^2\sum_{\kappa=1}^c(\kappa+1)\cdot\ell_d^\kappa$, the Chevalley–Warning theorem guarantees that we may choose $\alpha=(a_j)$ nontrivial in $\mathbb{F}_2^{(r)}$ such that $(\alpha_z\otimes\alpha_z^1)(z(i))=0$ for all i. We have

$$(\alpha_z\otimes\alpha_z^1)(z(i))=\sum_{j=1}^r\alpha_z(\sigma_{ij})\otimes\alpha_z(z_{ij})=\sum_{j=1}^r\alpha_z(\sigma_{ij})\otimes a_jz_i=\Bigg(\sum_{q\in\mathrm{Ram}(P_i(\alpha))}\alpha_z(\sigma_q)\Bigg)\otimes z_i.$$

Hence $\sum_{q\in\mathrm{Ram}(P_i(\alpha))}\alpha_z(\sigma_q)=0$ for all $i=1,\dots,d$, so that Proposition 4.2 applies. Consequently there is a strong Scholz field E with respect to N containing $K(\alpha)=E(\alpha)\cap K_\delta$ and admitting G_d^c as Galois group over the rationals. We also get

$$\mathrm{Ram}(E)=\mathrm{Ram}(K(\alpha))\cup\{p_1,\dots,p_t\},$$

where $t=\dim Z(G_d^c)=\sum_{\kappa=1}^c\ell_d^\kappa$. Here the t-set $\{p_1,\dots,p_t\}$ of rational primes may be chosen in infinitely many pairwise disjoint ways.

By induction $|\mathrm{Ram}(E_\delta)|=|\mathrm{Ram}(K_\delta)|\le f_{c-1}(\delta)$. Define f_c such that $f_c(d)=f_{c-1}(\delta)+\sum_{\kappa=1}^c\kappa\cdot\ell_d^\kappa$. Then $|\mathrm{Ram}(E)|\le f_c(d)$. Since $\delta=rd$ is an integral polynomial in d of degree $c+3$, this f_c is an integral polynomial of degree $(c+3)\deg f_{c-1}=(c+3)!/24$. This completes the proof of the theorem.

Acknowledgements

I owe special thanks to J.-P. Serre for some revealing discussions on the inverse problem of Galois theory for nilpotent (and for solvable) groups. I am also indebted to the referee for some valuable comments.

References

[Fröhlich 1983] A. Fröhlich, *Central extensions, Galois groups, and ideal class groups of number fields*, Contemp. Math. **24**, Amer. Math. Soc., Providence, RI, 1983. MR Zbl

[Geyer and Jarden 1998] W.-D. Geyer and M. Jarden, "Bounded realization of l-groups over global fields: the method of Scholz and Reichardt", *Nagoya Math. J.* **150** (1998), 13–62. MR Zbl

[Hecke 1981] E. Hecke, *Lectures on the theory of algebraic numbers*, Graduate Texts in Math. **77**, Springer, 1981. MR Zbl

[Kisilevsky et al. 2010] H. Kisilevsky, D. Neftin, and J. Sonn, "On the minimal ramification problem for semiabelian groups", *Algebra Number Theory* **4**:8 (2010), 1077–1090. MR Zbl

[Meshulam and Sonn 1999] R. Meshulam and J. Sonn, "A quantitative version of a lemma of Shafarevich–Ishkhanov", *Comm. Algebra* **27**:3 (1999), 1255–1259. MR Zbl

[Neukirch et al. 2000] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, Grundlehren der Math. Wissenschaften **323**, Springer, 2000. MR Zbl

[Plans 2004] B. Plans, "On the minimal number of ramified primes in some solvable extensions of $\mathbb{Q}$", *Pacific J. Math.* **215**:2 (2004), 381–391. MR Zbl

[Rabayev 2013] D. Rabayev, *Upper bound on the minimal number of ramified primes for solvable groups over the rationals*, Ph.D. thesis, Technion-Israel Institute of Technology, 2013.

[Reichardt 1937] H. Reichardt, "Konstruktion von Zahlkörpern mit gegebener Galoisgruppe von Primzahlpotenzordnung", *J. Reine Angew. Math.* **177** (1937), 1–5. MR Zbl

[Schmid 2014] P. Schmid, "On 2-extensions of the rationals with restricted ramification", *Acta Arith.* **163**:2 (2014), 111–125. MR Zbl

[Schmid 2017] P. Schmid, "Disposition p-groups", *Arch. Math.* (*Basel*) **108**:2 (2017), 113–121. MR Zbl

[Scholz 1937] A. Scholz, "Konstruktion algebraischer Zahlkörper mit beliebiger Gruppe von Primzahlpotenzordnung, I", *Math. Z.* **42**:1 (1937), 161–188. MR Zbl

[Serre 1992] J.-P. Serre, *Topics in Galois theory*, Research Notes in Math. **1**, Jones and Bartlett, Boston, 1992. MR Zbl

[Shafarevich 1954] I. R. Shafarevich, "On the construction of fields with a given Galois group of order l^{α}", *Izv. Akad. Nauk SSSR Ser. Mat.* **18**:3 (1954), 261–296. In Russian; translated in *Amer. Math. Soc. Trans.* (2) 4 (1956), 107–142. Zbl

[Shafarevich 1989] I. R. Shafarevich, "Factors of a descending central series", *Mat. Zametki* **45**:3 (1989), 114–117. In Russian; translated in *Math. Notes* **45**:3 (1989), 262–264. MR Zbl

[Weil 1967] A. Weil, *Basic number theory*, Grundlehren der Math. Wissenschaften **144**, Springer, 1967. MR Zbl

Communicated by Pham Huu Tiep
Received 2017-07-26 Revised 2018-07-21 Accepted 2018-08-26

peter.schmid@uni-tuebingen.de *Mathematisches Institut, Universität Tübingen, Tübingen, Germany*

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2403

Heights of hypersurfaces in toric varieties

Roberto Gualdi

For a cycle of codimension 1 in a toric variety, its degree with respect to a nef toric divisor can be understood in terms of the mixed volume of the polytopes associated to the divisor and to the cycle. We prove here that an analogous combinatorial formula holds in the arithmetic setting: the global height of a 1-codimensional cycle with respect to a toric divisor equipped with a semipositive toric metric can be expressed in terms of mixed integrals of the v-adic roof functions associated to the metric and the Legendre–Fenchel dual of the v-adic Ronkin function of the Laurent polynomial of the cycle.

Introduction

The arithmetic intersection theory of toric varieties with respect to toric line bundles equipped with their canonical metric was first studied by Maillot [2000]. Later, the systematic extension of the toric dictionary to Arakelov geometry was carried out by Burgos Gil, Philippon and Sombra [Burgos Gil et al. 2014]. It turns out from their study that suitably metrized toric line bundles can be expressed in terms of families of concave functions on convex polytopes and that the height of the toric variety with respect to this choice is related to the integral of such functions. Their theory allows one to treat a large spectrum of height functions, namely the ones arising from toric line bundles equipped with toric metrics; this includes the canonical heights studied by Maillot and the Fubini–Study height. On the other hand, the techniques developed in [Burgos Gil et al. 2014] only apply to the computation of the height of toric subvarieties and do not solve, for instance, the problem of determining the height of a general cycle of codimension 1. This question was answered in a very special case by [Maillot 2000], where a relation between the canonical height of a hypersurface in a smooth projective toric variety and the Mahler measure of the corresponding

MSC2010: primary 14M25; secondary 11G50, 14G40, 52A39.
Keywords: toric variety, height of a variety, Ronkin function, Legendre–Fenchel duality, mixed integral.

polynomial was given. Other computations have been performed by Cassaigne and Maillot [2000] for the Fubini–Study height of hypersurfaces in projective spaces. Extending the techniques of [Burgos Gil et al. 2014], we give here a combinatorial formula for the height of a 1-codimensional cycle in a toric variety for a much more general choice of metrics.

For the sake of simplicity, we restrict for the moment to the case of an ambient proper toric variety X_Σ of dimension n over $\mathbb{Q}$, leaving the treatment of the case of an arbitrary base adelic field to the body of the paper. Let $\mathfrak{M}$ stand for the set of places of $\mathbb{Q}$. As usual in toric geometry, we denote by M the lattice of characters of the torus of X_Σ, by N its dual lattice, and by $M_\mathbb{R}$ and $N_\mathbb{R}$ the corresponding real vector spaces. We are interested in a combinatorial expression for the height of a cycle of codimension 1 in X_Σ with respect to a suitable choice of a metrized (Cartier) divisor. By the linearity of height functions, we can restrict to the case of an irreducible hypersurface Y. Moreover, since irreducible hypersurfaces in X_Σ not intersecting its dense open torus have to coincide with 1-codimensional toric orbits, whose heights have already been calculated in [Burgos Gil et al. 2014, Proposition 5.1.11], we can assume that the generic point of Y lies in the dense open orbit of X_Σ. Under this assumption, Y is described by an irreducible Laurent polynomial f with rational coefficients. Its Newton polytope $\mathrm{NP}(f)$ is a nonempty subset of $M_\mathbb{R}$ capturing enough information for the intersection-theoretical properties of Y. For instance, Proposition 5.2 implies that the degree of Y with respect to a toric divisor D on X_Σ generated by its global sections is given by

$$\deg_D(Y) = \mathrm{MV}_M(\Delta, \dots, \Delta, \mathrm{NP}(f)),$$

where Δ is the polytope in $M_\mathbb{R}$ associated to D and MV_M denotes the mixed volume of convex bodies in $M_\mathbb{R}$ with respect to a suitably normalized Haar measure.

The height of a cycle in X is the arithmetic counterpart of its degree with respect to a divisor D. Its definition requires as an extra datum the choice of an adelic semipositive metric on D; see Section 3 for a precise definition. To have a combinatorial description of heights in toric varieties, it is necessary to ask D to be a toric divisor (with associated polytope Δ) and the metric on it to be "toric invariant" in some sense. In such a situation, Burgos Gil, Philippon and Sombra have shown that a combinatorial description is possible, translating the additional information of the metric into an extra dimension on the convex geometrical side: an adelic semipositive toric metric on D is associated to a family $(\vartheta_v)_{v\in\mathfrak{M}}$ of continuous concave functions on Δ, called the *roof functions* of the metric, such that $\vartheta_v = 0$ for all but finitely many v. We show how the height of Y with respect to the adelic semipositive toric metrized divisor $\overline{D}$ can be expressed using such an extra-dimensional representation, in a spirit analogous to the formula for its degree mentioned above. The key idea consists in associating to the polynomial f defining Y, for every place v of $\mathbb{Q}$, a suitable function which we call the *v-adic Ronkin function* of f and denote by $\rho_{f,v}$. It is a concave function on $N_\mathbb{R}$ whose value at u can be interpreted as an average of $-\log|f|$ on the fiber of the tropicalization map over u. When v is archimedean, it is the Ronkin function studied by Passare and Rullgård among others, while for nonarchimedean places it coincides with the v-adic tropicalization of the polynomial f. Its Legendre–Fenchel dual $\rho_{f,v}^\vee$ is a concave function on $M_\mathbb{R}$ which is supported on the

Newton polytope of f. Recall now that Philippon and Sombra [2008a] introduced a polarized version of the integration of a concave function with bounded support, called the *mixed integral*. For the choice of a suitably normalized Haar measure on the vector space $M_{\mathbb{R}}$, it is a multilinear symmetric real-valued function MI_M taking as entries $n+1$ concave functions supported on convex bodies in $M_{\mathbb{R}}$.

Theorem 1. *The height of Y with respect to $\overline{D}$ is given by*

$$h_{\overline{D}}(Y) = \sum_{v \in \mathfrak{M}} \mathrm{MI}_M(\vartheta_v, \dots, \vartheta_v, \rho_{f,v}^{\vee}).$$

Despite the complexity of the computation of the archimedean Ronkin function, the formula in the previous theorem clarifies the relation between the defining polynomial of an irreducible hypersurface and its height with respect to an adelic semipositive toric metrized divisor. It is easy to specialize it to the case of the canonical metric on D, where it reduces to the equality proved in [Maillot 2000], or of the Fubini–Study metric in the projective setting. We hope that a better understanding of the properties of mixed integrals and archimedean Ronkin functions could be used to deduce both lower and upper bounds for the height of Y. More importantly, our result asserts that the collection of the v-adic Ronkin functions of a hypersurface contains enough information to determine its height; we wonder whether other arithmetical properties of Y might be read in terms of such functions.

To show the stated result, we prove more precise formulas for the local height and the toric local height of a 1-codimensional cycle. We also show some new properties of mixed integrals and we propose a more uniform definition and study of v-adic Ronkin functions which is independent of whether the place v is archimedean or not. The obtained formulas for the height extend to the case of admissible adelic toric metrized divisors as alternated sums of mixed integrals, as in [Burgos Gil et al. 2014, Remark 5.1.10].

For an arbitrary adelic base field K, we remark that one needs to prove that the global height of Y with respect to an adelic semipositive toric metrized divisor is a finite sum and hence well-defined. This is automatic if K is a global field, because of [loc. cit., Proposition 1.5.14 and Theorem 4.9.3]. We show it here for an arbitrary adelic field K with product formula, in which case the formulas for the height stay true. In the more general setting of an adelic field K not satisfying the product formula, it is easy to verify that the same equality for the global height holds up to the sum by the *defect* of K; see [loc. cit., Definition 1.5.9]. Finally, the recent work [Gubler and Hertel 2017] suggests that similar statements might hold for a base M-field.

We now briefly summarize the content of each section.

In Section 1, we recall the tools from convex geometry which are needed throughout the paper: Legendre–Fenchel duality of concave functions, real Monge–Ampère measures and mixed integrals. In particular, we reinterpret the recursive formula for mixed integrals proved by Philippon and Sombra in terms of mixed real Monge–Ampère measures. We then make use of it to deduce two elementary, though useful, properties of such operators. Finally, we describe what happens when one of the functions appearing in the mixed integral is the indicator function of a line segment.

Section 2 deals with the key object of our work: v-adic Ronkin functions of Laurent polynomials, which are introduced and described after recalling the needed preliminaries in tropical and nonarchimedean geometry. In this context, the discussion of a notion of minimal boundaries allows one to treat the archimedean and nonarchimedean cases homogeneously.

In Section 3 we briefly recall the general adelic Arakelov framework and focus then on the results obtained by Burgos Gil, Philippon and Sombra in the toric setting. To keep the treatment of archimedean and nonarchimedean places on equal footing, we rephrase their description of the Chambert-Loir measure of semipositive toric metrized divisors in terms of minimal boundaries of tropical fibers.

As a needed step for the main proof, we combinatorially describe the Weil divisor of the rational function defined by a Laurent polynomial on a toric variety. This result can be of independent interest and has thus been set aside in Section 4.

Section 5 is dedicated to the proofs of our main results Theorems 5.9 and 5.12, which are formulas for the local height and the toric local height of cycles of codimension 1 in toric varieties. We then make use of them to prove the integrability statement and a formula for their global heights, with respect to the choice of adelic semipositive toric metrized divisors.

For binomial hypersurfaces, such a formula is compatible with the one deduced from [Burgos Gil et al. 2014]. This is shown in Section 6, where we also apply our results to some other particular cases. We provide convex geometrical formulas for the canonical height of 1-codimensional cycles, obtaining the quoted result by Maillot, and for the Fubini–Study height of a projective hypersurface. We also propose a new height function, the ρ-height, for which we give a compact formula. We do not know any application of such a height, which could be anyway worth studying.

Terminology and notation. A *variety* X is assumed to be a reduced and irreducible separated scheme of finite type over a field. By an *irreducible hypersurface* in it we mean a closed integral subscheme of codimension 1 in X. A *divisor* on X is a Cartier divisor, unless otherwise stated. Toric varieties are assumed to be normal; whenever the choice of the base field K is clear from the context, the notation X_Σ will refer to the toric variety over K associated to the fan Σ.

The term *measure* on a topological space stands for a signed Borel measure on it; in particular, measures admit a well-defined push-forward via continuous mappings. A measure which only takes nonnegative real values on Borel subsets is called a *positive measure*.

1. Preliminaries in convex geometry

This section is devoted to recalling notions from convex geometry that will be useful in the sequel. We follow the conventions and notation of [Burgos Gil et al. 2014, Chapter 2], referring to [Rockafellar 1970, §12] for a more complete treatment of the subject. We refer to these two sources for the proofs of the statements we make here.

For the whole section, let N be a lattice of rank n and $M := \mathrm{Hom}(N, \mathbb{Z})$ its dual lattice. Denote by $N_\mathbb{R} = N \otimes_\mathbb{Z} \mathbb{R}$ and by $M_\mathbb{R} = M \otimes_\mathbb{Z} \mathbb{R}$ the corresponding n-dimensional real vector spaces.

By a *polyhedron* in $N_{\mathbb{R}}$ we mean a convex subset of $N_{\mathbb{R}}$ obtained as the intersection of finitely many closed half-spaces $\{u \in N_{\mathbb{R}} : \langle x, u\rangle + c \geq 0\}$, with $x \in M_{\mathbb{R}}$ and $c \in \mathbb{R}$. If all the slopes x can be chosen in M, the polyhedron is said to be *rational*. A *polytope* is a bounded polyhedron. A polytope in $N_{\mathbb{R}}$ whose vertices all lie in N is called a *lattice polytope*; it is in particular a rational polytope. A compact convex subset of $N_{\mathbb{R}}$ is called a *convex body*.

1A. ***Legendre–Fenchel duality.*** A function $f : N_{\mathbb{R}} \to \mathbb{R} \cup \{-\infty\}$ is said to be *concave* if it is not identically $-\infty$ and, for every $u_1, u_2 \in N_{\mathbb{R}}$ and for every $t \in [0, 1]$, one has the inequality

$$f(tu_1 + (1-t)u_2) \geq tf(u_1) + (1-t)f(u_2).$$

The *effective domain* of a concave function f is the set on which the function takes values different from $-\infty$ and it is denoted by $\mathrm{dom}(f)$: it is a convex subset of $N_{\mathbb{R}}$. A concave function is said to be *closed* if it is upper semicontinuous. Every concave function with closed effective domain, on which it is continuous, is closed. The *recession function* of a closed concave function f is the concave conical function which takes on $u \in N_{\mathbb{R}}$ the value

$$\mathrm{rec}(f)(u) := \lim_{\lambda\to\infty} \frac{f(v_0 + \lambda u)}{\lambda} \in \mathbb{R} \cup \{-\infty\}$$

for any $v_0 \in \mathrm{dom}(f)$; see [Rockafellar 1970, Theorem 8.5]. Finally, a concave function f with effective domain a polyhedron in $N_{\mathbb{R}}$ is *piecewise affine* if

$$f(u) = \min_{\alpha\in S}(\langle \alpha, u\rangle + c_\alpha)$$

for every $u \in \mathrm{dom}(f)$, with S a finite subset of $M_{\mathbb{R}}$ and $c_\alpha \in \mathbb{R}$ for every $\alpha \in S$.

To each concave function f on $N_{\mathbb{R}}$, one can associate its *Legendre–Fenchel dual*, which is the closed concave function $f^\vee$ on $M_{\mathbb{R}}$ defined as

$$f^\vee(x) := \inf_{u\in N_{\mathbb{R}}} (\langle x, u\rangle - f(u))$$

for every $x \in M_{\mathbb{R}}$; see [Burgos Gil et al. 2014, §2.2]. If f is closed, $(f^\vee)^\vee = f$. The effective domain of $f^\vee$ is a convex subset of $M_{\mathbb{R}}$, which one calls the *stability set of f* and denotes by $\mathrm{stab}(f)$.

The following example is classical and will play a role later on.

Example 1.1. Any nonempty convex body B in $M_{\mathbb{R}}$ induces a concave function Ψ_B on $N_{\mathbb{R}}$, called the *support function* of B and defined as

$$\Psi_B(u) := \min_{x\in B}\langle x, u\rangle$$

for every $u \in N_{\mathbb{R}}$. Its Legendre–Fenchel dual is the *indicator function ι_B of B*, which is the function taking the value 0 on B and $-\infty$ elsewhere. Hence, $\mathrm{dom}(\Psi_B) = N_{\mathbb{R}}$ and $\mathrm{stab}(\Psi_B) = B$. Notice that, whenever B is a polytope, Ψ_B is a conic piecewise affine concave function.

We also recall that there exist a number of operations that one can define on concave functions, in addition to the usual pointwise sum and scalar multiplication. Among these, the *sup-convolution* of two concave functions f and g on $N_{\mathbb{R}}$ with nondisjoint stability sets is defined as

$$(f \boxplus g)(v) := \sup_{u_1+u_2=v} (f(u_1)+g(u_2))$$

and the *right scalar multiplication* of f by $\lambda \in \mathbb{R}_{\geq 0}$ as

$$(f\lambda)(u) := \lambda f(u/\lambda).$$

Also, the *translate* of a concave function f on $N_{\mathbb{R}}$ by a point $u_0 \in N_{\mathbb{R}}$ is set to be

$$(\tau_{u_0} f)(u) := f(u-u_0).$$

These operations are dual, via Legendre–Fenchel duality, to the usual pointwise addition, scalar multiplication and sum by a linear function, respectively; see [Burgos Gil et al. 2014, Propositions 2.3.1 and 2.3.3].

1B. *Real Monge–Ampère measures.* For any closed concave function f on $N_{\mathbb{R}}$ with $\operatorname{dom}(f) = N_{\mathbb{R}}$ and for any Haar measure μ on $M_{\mathbb{R}}$, one can define a corresponding *real Monge–Ampère measure* $\mathcal{M}_\mu(f)$, as in [Burgos Gil et al. 2014, §2.7]. It is a measure on $N_{\mathbb{R}}$, of total mass $\mu(\operatorname{stab}(f))$, being supported on finitely many points if f is piecewise affine. The Monge–Ampère operator, associating to each closed concave function f with $\operatorname{dom}(f) = N_{\mathbb{R}}$ the corresponding measure $\mathcal{M}_\mu(f)$ on $N_{\mathbb{R}}$ is homogeneous of degree n with respect to pointwise scalar multiplication. It was shown in [Passare and Rullgård 2004] that such an operator admits a polarization: for $f_1, \dots, f_n$ closed concave functions with effective domain $N_{\mathbb{R}}$, their *mixed real Monge–Ampère measure* is defined as the measure

$$\mathrm{M}\mathcal{M}_\mu(f_1, \dots, f_n) := \sum_{k=1}^{n} (-1)^{n-k} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mathcal{M}_\mu(f_{i_1} + \cdots + f_{i_k}). \tag{1-1}$$

Notice that this definition differs from [Passare and Rullgård 2004, formula (14)] and [Burgos Gil et al. 2014, Definition 2.7.12] by a multiplicative constant. It follows from [Passare and Rullgård 2004, §5] that the measure in (1-1) is in fact a positive measure. The so-obtained mixed Monge–Ampère operator is, by definition, symmetric and multilinear in its entries (with respect to pointwise sum) and it satisfies

$$\mathrm{M}\mathcal{M}_\mu(f, \dots, f) = n!\, \mathcal{M}_\mu(f)$$

for every closed concave function f. In particular, if $f_1, \dots, f_n$ are closed concave functions with convex bodies as stability sets, their mixed real Monge–Ampère measure is a finite measure on $N_{\mathbb{R}}$ of total mass $\mathrm{MV}_\mu(\operatorname{stab}(f_1), \dots, \operatorname{stab}(f_n))$. Here, MV_μ denotes the mixed volume of convex bodies with respect to the measure μ, normalized in such a way that $\mathrm{MV}_\mu(Q, \dots, Q) = n!\, \mu(Q)$ for every convex body Q in $M_{\mathbb{R}}$.

Remark 1.2. The integral structure on $M_{\mathbb{R}}$ coming from the subjacent lattice M gives a distinguished measure vol_M on $M_{\mathbb{R}}$, which is the unique Haar measure for which the volume of a fundamental domain of M is 1 (this does not depend on the choice of a basis of M). To lighten the notation, the corresponding (mixed) real Monge–Ampère operator will be denoted by $\mathrm{M}\mathcal{M}_M$.

1C. ***Mixed integrals.*** For any Haar measure μ on $M_{\mathbb{R}}$, the application mapping a compactly supported concave function g on $M_{\mathbb{R}}$ to its Lebesgue integral with respect to μ is homogeneous of degree $n+1$ with respect to right scalar multiplication. As a consequence of [Philippon and Sombra 2008a, Proposition 4.5], there exists a polarized operator: for $g_0, \dots, g_n$ concave functions on $M_{\mathbb{R}}$ with respective domains the convex bodies $Q_0, \dots, Q_n$, their *mixed integral* with respect to μ is defined as

$$\mathrm{MI}_{\mu}(g_0, \dots, g_n) := \sum_{k=0}^{n} (-1)^{n-k} \sum_{0 \leq i_0 < \dots < i_k \leq n} \int_{Q_{i_0}+\dots+Q_{i_k}} (g_{i_0} \boxplus \dots \boxplus g_{i_k})\, d\mu.$$

This notion was introduced and studied in [Philippon and Sombra 2008a; 2008b]. The mixed integral operator is symmetric and multilinear in its entries (with respect to sup-convolution) and it satisfies

$$\mathrm{MI}_{\mu}(g, \dots, g) = (n+1)! \int_{Q} g\, d\mu$$

for every concave function g on a convex body Q. As for the mixed Monge–Ampère operator, we will denote by MI_M the mixed integral computed with respect to the Haar measure vol_M on $M_{\mathbb{R}}$.

The following proposition describes the behavior of mixed integrals with respect to translation of the entries.

Proposition 1.3. *Let g_i be a concave function defined on a convex body Q_i in $M_{\mathbb{R}}$ for $i = 0, \dots, n$. Let also $x_0 \in M_{\mathbb{R}}$. Then,*

$$\mathrm{MI}_{\mu}(\tau_{x_0} g_0, g_1, \dots, g_n) = \mathrm{MI}_{\mu}(g_0, \dots, g_n).$$

Proof. For any subset $\{i_1, \dots, i_r\} \subseteq \{1, \dots, n\}$ one has, directly by definition, that

$$((\tau_{x_0} g_0) \boxplus g_{i_1} \boxplus \dots \boxplus g_{i_r})(x + x_0) = (g_0 \boxplus g_{i_1} \boxplus \dots \boxplus g_{i_r})(x)$$

for every $x \in M_{\mathbb{R}}$. The change of variables formula implies then that the integrals appearing in the definitions of $\mathrm{MI}_{\mu}(\tau_{x_0} g_0, g_1, \dots, g_n)$ and of $\mathrm{MI}_{\mu}(g_0, \dots, g_n)$ are pairwise equal, from which the statement follows trivially. □

We now focus on a recursive formula for mixed integrals. For any closed convex subset C in $M_{\mathbb{R}}$, and for every $u \in N_{\mathbb{R}}$, one can consider the subset

$$C^u := \left\{ x \in C : \langle x, u \rangle = \inf_{y \in C} \langle y, u \rangle \right\}$$

of C. For $u \neq 0$, this subset is contained in an affine subspace of $M_{\mathbb{R}}$ of codimension 1 and parallel to $u^{\perp} := \{x \in M_{\mathbb{R}} : \langle x, u \rangle = 0\}$. It is immediate from the definition that for every pair of convex bodies C_1 and C_2 in $M_{\mathbb{R}}$, and for every $u \in N_{\mathbb{R}}$, one has $C_1^u + C_2^u = (C_1 + C_2)^u$, where the plus sign denotes the Minkowski sum of convex sets. When $u \in N_{\mathbb{Q}} \setminus \{0\}$, the intersection $M(u) := M \cap u^{\perp}$ is a lattice of rank $n-1$ spanning the linear space $u^{\perp}$ and hence it induces a normalized Haar measure $\mathrm{vol}_{M(u)}$ on $u^{\perp}$, as in Remark 1.2.

Remark 1.4. Let $B_1, \dots, B_n$ be n convex bodies in $M_\mathbb{R}$ and let $u \in N_\mathbb{R}$. The invariance under translation of $\mathrm{vol}_{M(u)}$ allows one to consistently consider the mixed volume of $B_1^u, \dots, B_n^u$ as convex bodies in $u^\perp$.

Similarly, let $g_0, \dots, g_n$ be concave functions defined on convex bodies $B_0, \dots, B_n$ in $M_\mathbb{R}$, respectively. For every $u \in N_\mathbb{Q} \setminus \{0\}$, Proposition 1.3 allows one to consider the mixed integral with respect to $\mathrm{vol}_{M(u)}$ of $g_0|_{B_0^u}, \dots, g_n|_{B_n^u}$, as functions defined on convex subsets of $u^\perp$.

For a concave function g with effective domain a polytope Q in $M_\mathbb{R}$, its *hypograph* is the closed convex set

$$\Gamma(g) := \{(x,t) : x \in Q, t \leq g(x)\} \subseteq M_\mathbb{R} \times \mathbb{R}.$$

With the pairing between $N_\mathbb{R} \times \mathbb{R}$ and $M_\mathbb{R} \times \mathbb{R}$ given by $\langle (x,t),(u,\lambda)\rangle := \langle x,u\rangle + t\lambda$ for every $(u,\lambda) \in N_\mathbb{R} \times \mathbb{R}$ and $(x,t) \in M_\mathbb{R} \times \mathbb{R}$, one can consider the subset $\Gamma(g)^{(u,\lambda)}$ of $\Gamma(g)$ for every $(u,\lambda) \in N_\mathbb{R} \times \mathbb{R}$. It is empty when $\lambda > 0$.

The mixed real Monge–Ampère measure of piecewise affine concave functions can be made explicit in terms of the hypographs of their Legendre–Fenchel duals. We denote by δ_v the Dirac measure supported on v.

Proposition 1.5. *For $i = 1, \dots, n$, let g_i be a piecewise affine concave function with effective domain a polytope $Q_i \subset M_\mathbb{R}$, and Γ_i the hypograph of g_i. Denote by $\pi : M_\mathbb{R} \times \mathbb{R} \to M_\mathbb{R}$ the projection onto the first factor. For any choice of a Haar measure μ on $M_\mathbb{R}$ one has*

$$\mathrm{M}\mathcal{M}_\mu(g_1^\vee, \dots, g_n^\vee) = \sum_{v \in N_\mathbb{R}} \mathrm{MV}_\mu(\pi(\Gamma_1^{(v,-1)}), \dots, \pi(\Gamma_n^{(v,-1)}))\, \delta_v,$$

and the sum is finite.

Proof. For a piecewise affine concave function g with bounded domain in $M_\mathbb{R}$, its Legendre–Fenchel dual is a piecewise affine concave function with domain $N_\mathbb{R}$ and [Burgos Gil et al. 2014, Proposition 2.7.4] affirms that

$$\mathcal{M}_\mu(g^\vee) = \sum_{v \in N_\mathbb{R}} \mu(v^*)\delta_v,$$

with $v^* = \{x \in M_\mathbb{R} : g^\vee(v) = \langle x, v\rangle - g(x)\}$. From the definition of the Legendre–Fenchel duality, one has hence that

$$v^* = \Big\{x \in M_\mathbb{R} : \langle x, v\rangle - g(x) = \min_{y \in M_\mathbb{R}} (\langle y, v\rangle - g(y))\Big\} = \{x \in M_\mathbb{R} : (x, g(x)) \in \Gamma(g)^{(v,-1)}\},$$

and so

$$\mathcal{M}_\mu(g^\vee) = \sum_{v \in N_\mathbb{R}} \mu(\pi(\Gamma(g)^{(v,-1)}))\delta_v.$$

The sum is moreover supported on finitely many $v \in N_\mathbb{R}$, corresponding to the directions of the finitely many exposed faces of $\Gamma(g)$.

By [Attouch and Wets 1989, §2], the relation

$$\Gamma(g_i \boxplus g_j) = \Gamma(g_i) + \Gamma(g_j)$$

on the hypographs of g_i and g_j holds for any $i, j \in \{1, \dots, n\}$. As a consequence, for every subset $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$, [Burgos Gil et al. 2014, Proposition 2.3.1] and the linearity of π yield

$$\begin{aligned}\mathcal{M}_\mu(g_{i_1}^\vee + \cdots + g_{i_k}^\vee) &= \mathcal{M}_\mu((g_{i_1} \boxplus \cdots \boxplus g_{i_k})^\vee) \\ &= \sum_{v \in N_\mathbb{R}} \mu(\pi(\Gamma_{i_1}^{(v,-1)} + \cdots + \Gamma_{i_k}^{(v,-1)}))\delta_v \\ &= \sum_{v \in N_\mathbb{R}} \mu(\pi(\Gamma_{i_1}^{(v,-1)}) + \cdots + \pi(\Gamma_{i_k}^{(v,-1)}))\delta_v,\end{aligned}$$

and the sum is finite. The statement follows then from the definition of the mixed real Monge–Ampère measure, rearranging the terms. □

We can now prove a recursive formula relating the notions of the mixed real Monge–Ampère measure and the mixed integral of concave functions, via Legendre–Fenchel duality. A vector $u \in N$ is said to be *primitive* if it is nonzero and there is no other element $u' \in N$ such that $ku' = u$ for some positive integer k.

Theorem 1.6. *For $i = 0, \dots, n$, let g_i be a continuous concave function on a rational polytope Q_i in $M_\mathbb{R}$. Then*

$$\mathrm{MI}_M(g_0, \dots, g_n) = -\sum_{\substack{u \in N \\ \text{primitive}}} \Psi_{Q_0}(u)\, \mathrm{MI}_{M(u)}(g_1|_{Q_1^u}, \dots, g_n|_{Q_n^u}) - \int_{N_\mathbb{R}} g_0^\vee \, d\, \mathrm{M}\mathcal{M}_M(g_1^\vee, \dots, g_n^\vee),$$

the first sum being finite.

In particular, if g_i is a piecewise affine concave function on Q_i with hypograph Γ_i for any $i = 0, \dots, n$, denoting by $\pi : M_\mathbb{R} \times \mathbb{R} \to M_\mathbb{R}$ the projection onto the first factor, one has

$$\begin{aligned}&\mathrm{MI}_M(g_0, \dots, g_n) \\ &\quad = -\sum_{\substack{u \in N \\ \text{primitive}}} \Psi_{Q_0}(u)\, \mathrm{MI}_{M(u)}(g_1|_{Q_1^u}, \dots, g_n|_{Q_n^u}) - \sum_{v \in N_\mathbb{R}} g_0^\vee(v)\, \mathrm{MV}_M(\pi(\Gamma_1^{(v,-1)}), \dots, \pi(\Gamma_n^{(v,-1)})).\end{aligned}$$

Proof. By [Burgos Gil et al. 2014, Proposition 2.5.23(1)], any continuous concave function on a polytope can be approximated, with respect to uniform convergence, by a sequence of piecewise affine concave functions on the polytope itself. On the other hand, the Legendre–Fenchel duality and the real Monge–Ampère operator are continuous with respect to uniform limits of concave functions; see [Burgos Gil et al. 2014, Proposition 2.2.3] and [Rauch and Taylor 1977, §3], respectively. It is not difficult to show that the same holds for mixed integrals. Thanks to Proposition 1.5, it is hence enough to prove the formula in the particular case of $g_0, \dots, g_n$ being piecewise affine concave functions.

Let hence g_i be a concave piecewise affine function on the rational polytope Q_i in $M_\mathbb{R}$, and Γ_i its hypograph, for $i = 0, \dots, n$. The choice of a basis of N (and of the dual basis of M) endows $N_\mathbb{R}$ and $M_\mathbb{R}$ with a euclidean structure, allowing one to consider the sets

$$\mathbb{S}^{n-1} := \{w \in N_\mathbb{R} : \|w\| = 1\} \subseteq N_\mathbb{R}$$

and

$$\mathbb{S}^n_- := \{(v,t) \in N_{\mathbb{R}} \times \mathbb{R} : \|(v,t)\| = 1, t < 0\} \subseteq N_{\mathbb{R}} \times \mathbb{R}.$$

After a change of sign due to the use of different notation, [Philippon and Sombra 2008b, Proposition 8.5] affirms that

$$\mathrm{MI}_M(g_0,\dots,g_n) = -\sum_{w\in\mathbb{S}^{n-1}} \Psi_{Q_0}(w)\,\mathrm{MI}_{n-1}(g_1|_{Q_1^w},\dots,g_n|_{Q_n^w}) - \sum_{r\in\mathbb{S}^n_-} \Psi_{\Gamma_0}(r)\ \mathrm{MV}_n(\Gamma_1^r,\dots,\Gamma_n^r), \quad (1\text{-}2)$$

where, on the right-hand side, one refers to the mixed integral with respect to the measure obtained restricting vol_M to $w^\perp$ and to the mixed volume with respect to the restriction of $\mathrm{vol}_{M\oplus\mathbb{Z}}$ to $r^\perp$.

Concerning the first sum on the right-hand side of (1-2), if a term in the sum is different from zero, then there exists a subset $I \subset \{1,\dots,n\}$ such that the Minkowski sum of Q_i^w, with $i \in I$, is of dimension $n-1$; in particular, if one sets $Q := Q_1 + \dots + Q_n$, then $Q^w = Q_1^w + \dots + Q_n^w$ needs to be of dimension $n-1$. As a consequence, one can restrict the sum to the set of vectors $w \in \mathbb{S}^{n-1}$ for which Q^w is an $(n-1)$-dimensional face of Q. This set is included in the set of vectors of unitary length which are perpendicular to an $(n-1)$-dimensional face of Q; hence it is finite since Q is a polytope. Moreover, since Q is rational, the ray spanned by such a vector w contains a unique primitive vector $u \in N$. The linearity of Ψ_{Q_0} yields hence the equality

$$\sum_{w\in\mathbb{S}^{n-1}} \Psi_{Q_0}(w)\,\mathrm{MI}_{n-1}(g_1|_{Q_1^w},\dots,g_n|_{Q_n^w}) = \sum_{\substack{u\in N\\ \text{primitive}}} \frac{\Psi_{Q_0}(u)}{\|u\|}\,\mathrm{MI}_{n-1}(g_1|_{Q_1^u},\dots,g_n|_{Q_n^u}).$$

The fact that the restriction of vol_M to $u^\perp$ is equal to the measure $\mathrm{vol}_{M(u)}$ multiplied by $\|u\|$, see [Burgos Gil et al. 2014, proof of Corollary 2.7.10], allows one to conclude that the first sum in (1-2) coincides with the desired one.

Regarding the second sum in (1-2), there exists an obvious bijection between $\mathbb{S}^n_-$ and $N_{\mathbb{R}}$ given by associating to each $r \in \mathbb{S}^n_-$ the only vector $v \in N_{\mathbb{R}}$ such that $(v,-1)$ lies on the line spanned by r. Hence,

$$\sum_{r\in\mathbb{S}^n_-} \Psi_{\Gamma_0}(r)\ \mathrm{MV}_n(\Gamma_1^r,\dots,\Gamma_n^r) = \sum_{v\in N_{\mathbb{R}}} \frac{\Psi_{\Gamma_0}(v,-1)}{\|(v,-1)\|}\ \mathrm{MV}_n(\Gamma_1^{(v,-1)},\dots,\Gamma_n^{(v,-1)}).$$

Directly by the definition of Legendre–Fenchel duality, one has that $\Psi_{\Gamma_0}(v,-1) = g_0^\vee(v)$. The statement follows then from the fact that for every Borel set E in $(v,-1)^\perp$, the measure of E with respect to the restriction of $\mathrm{vol}_{M\oplus\mathbb{Z}}$ to $(v,-1)^\perp$ equals $\|(v,-1)\|\cdot\mathrm{vol}_M(\pi(E))$, again by [Burgos Gil et al. 2014, proof of Corollary 2.7.10]. □

Remark 1.7. For a rational polytope P of full dimension n in $M_{\mathbb{R}}$, every facet F of P, that is a face of dimension $n-1$, admits a distinguished orthogonal vector: it is the unique primitive vector $v_F \in N$ which satisfies $P^{v_F} = F$. Under the additional assumption that the Minkowski sum $Q := Q_1 + \dots + Q_n$ is of dimension n in $M_{\mathbb{R}}$, the formula in Theorem 1.6 can be written as

$$\mathrm{MI}_M(g_0,\dots,g_n) = -\sum_F \Psi_{Q_0}(v_F)\,\mathrm{MI}_{M(v_F)}(g_1|_{Q_1^{v_F}},\dots,g_n|_{Q_n^{v_F}}) - \int_{N_{\mathbb{R}}} g_0^\vee\, d\,\mathrm{M}\mathcal{M}_M(g_1^\vee,\dots,g_n^\vee),$$

the first sum being over the finite set of facets of the polytope Q. Indeed, in such a situation the application $F \mapsto v_F$ realizes a bijection between the set of facets of Q and the set of primitive vectors $u \in N$ for which Q^u is an $(n-1)$-dimensional face of Q, which are the only vectors for which the term of the sum in the statement of the theorem does not vanish.

Remark 1.8. The statement of Theorem 1.6 can be reformulated in terms of Legendre–Fenchel duality. For $i = 0, \dots, n$, let f_i be a concave function on $N_{\mathbb{R}}$ with stability set a rational polytope Q_i in $M_{\mathbb{R}}$. Under the assumption that $Q_1 + \cdots + Q_n$ is of dimension n in $M_{\mathbb{R}}$, Remark 1.7 yields

$$\mathrm{MI}_M(f_0^{\vee}, \dots, f_n^{\vee}) = -\sum_F \Psi_{Q_0}(v_F)\,\mathrm{MI}_{M(v_F)}(f_1^{\vee}|_{Q_1^{v_F}}, \dots, f_n^{\vee}|_{Q_n^{v_F}}) - \int_{N_{\mathbb{R}}} f_0 \, d\,\mathrm{M}\mathcal{M}_M(f_1, \dots, f_n). \tag{1-3}$$

Indeed, it is sufficient to readily apply the previous theorem to the functions $f_0^{\vee}, \dots, f_n^{\vee}$, which are continuous on their domain and satisfy the equality $(f_i^{\vee})^{\vee} = f_i$ for each $i = 0, \dots, n$ by concavity and closedness. It is easy to verify that the choice $f_0 = \cdots = f_n = f$ in (1-3) yields the formula in [Burgos Gil et al. 2014, Corollary 2.7.10].

We present now two applications of the recursive formula proved above. The first one concerns the computation of the mixed integral when all except one entry are indicator functions in the sense of Example 1.1.

Corollary 1.9. *Let $Q_1, \dots, Q_n$ be rational polytopes in $M_{\mathbb{R}}$ and f a concave function on $N_{\mathbb{R}}$ with stability set a rational polytope. Then*

$$\mathrm{MI}_M(\iota_{Q_1}, \dots, \iota_{Q_n}, f^{\vee}) = -\,\mathrm{MV}_M(Q_1, \dots, Q_n) \cdot f(0).$$

Proof. By symmetry, one can develop the recursive formula in Remark 1.8 with respect to $f^{\vee}$ to obtain

$$\mathrm{MI}_M(\iota_{Q_1}, \dots, \iota_{Q_n}, f^{\vee}) = -\int_{N_{\mathbb{R}}} f \, d\,\mathrm{M}\mathcal{M}_M(\iota_{Q_1}^{\vee}, \dots, \iota_{Q_n}^{\vee}),$$

the indicator functions $\iota_{Q_1}, \dots, \iota_{Q_n}$ being zero where defined. The duality in Example 1.1 and the fact that

$$\mathrm{M}\mathcal{M}_M(\Psi_{Q_1}, \dots, \Psi_{Q_n}) = \mathrm{MV}_M(Q_1, \dots, Q_n)\delta_0$$

because of Proposition 1.5 conclude the proof. □

The second application explains how the mixed integral behaves with respect to pointwise sum by a constant in one entry.

Corollary 1.10. *Let g_i be a concave function defined on a rational polytope $Q_i \subseteq M_{\mathbb{R}}$ for $i = 0, \dots, n$ and $c \in \mathbb{R}$. Then*

$$\mathrm{MI}_M(g_0, \dots, g_{n-1}, g_n + c) = \mathrm{MI}_M(g_0, \dots, g_n) + c \cdot \mathrm{MV}_M(Q_0, \dots, Q_{n-1}).$$

Proof. Denoting by $c\delta_0$ the concave function which has value c at 0 and $-\infty$ otherwise, it follows from the definitions that $g_n + c = g_n \boxplus c\delta_0$. The multilinearity of mixed integrals implies then that

$$\mathrm{MI}_M(g_0, \dots, g_{n-1}, g_n + c) = \mathrm{MI}_M(g_0, \dots, g_n) + \mathrm{MI}_M(g_0, \dots, g_{n-1}, c\delta_0).$$

Using the fact that $(c\delta_0)^\vee = -c$, the recursive formula in Theorem 1.6, developed with respect to $c\delta_0$, yields

$$\mathrm{MI}_M(g_0,\dots,g_{n-1},c\delta_0)=\int_{N_\mathbb{R}} c\, d\,\mathrm{M}\mathcal{M}_M(g_0^\vee,\dots,g_{n-1}^\vee)=c\cdot \mathrm{M}\mathcal{M}_M(g_0^\vee,\dots,g_{n-1}^\vee)(N_\mathbb{R}).$$

The statement follows then from the fact that the total volume of the mixed Monge–Ampère measure of $g_0^\vee,\dots,g_{n-1}^\vee$ is equal to $\mathrm{MV}_M(\mathrm{dom}(g_0),\dots,\mathrm{dom}(g_{n-1}))$ by [Passare and Rullgård 2004, Proposition 3(iv)]. □

We conclude the section by proving a formula expressing the mixed integral of an $(n+1)$-tuple of concave functions on $M_\mathbb{R}$ where one of them is the indicator function of a line segment.

Let m be a primitive vector of M and consider the quotient $P := M/\mathbb{Z}m$. Since m is primitive, P is a lattice of rank $n-1$. By abuse of notation, let π denote both the projection from M to P and the induced linear map from $M_\mathbb{R}$ to $P_\mathbb{R}$. For each closed concave function g defined on a compact subset B of $M_\mathbb{R}$, let

$$\pi_*g:\pi(B)\to\mathbb{R},\qquad x\mapsto \max_{y\in\pi^{-1}(x)} g(y), \tag{1-4}$$

be the *direct image* of g by π. It is a well-defined closed concave function with domain a bounded subset of $P_\mathbb{R}$; see [Rockafellar 1970, Theorems 5.7 and 9.2]. Finally, for $x_1, x_2\in M_\mathbb{R}$, denote by $\overline{x_1x_2}$ the line segment in $M_\mathbb{R}$ with extremal points x_1 and x_2. The following lemma is a generalization of [Ewald 1996, Exercise 3, p. 128] and seems to be well known to experts, though we could not find an adequate reference in the literature for its proof.

Lemma 1.11. *In the above hypotheses and notation and for $n\ge 2$, let $Q_1,\dots,Q_{n-1}$ be polytopes in $M_\mathbb{R}$. Then,*

$$\mathrm{MV}_M(\overline{0m},Q_1,\dots,Q_{n-1})=\mathrm{MV}_P(\pi(Q_1),\dots,\pi(Q_{n-1})).$$

Proof. Since the vector m is primitive, it can be extended to a basis of the lattice M; see for instance [Lekkerkerker 1969, Theorem 5, p. 21]. We suppose fixed throughout the proof such a basis $(m_1,\dots,m_{n-1},m)$ of M and the induced isomorphism $M_\mathbb{R}\simeq\mathbb{R}^n$; under this identification, the normalized volume vol_M corresponds to the Lebesgue measure vol_n on $\mathbb{R}^n$. Since $(\pi(m_1),\dots,\pi(m_{n-1}))$ is a basis of P, such a lattice is isomorphic to the span of $m_1,\dots,m_{n-1}$ in M and hence it is identified with the linear subspace $\mathbb{R}^{n-1}\times\{0\}$ of $\mathbb{R}^n$. Moreover, vol_P corresponds to the $(n-1)$-dimensional Lebesgue measure vol_{n-1} on $\mathbb{R}^{n-1}\times\{0\}$ and the map π to the vertical projection.

The claim reduces then to the particular case of a family of polytopes $Q_1,\dots,Q_{n-1}$ in $\mathbb{R}^n$, $m=(0,\dots,0,1)$ and π the vertical projection. Denoting by S the vertical segment of unitary length and rearranging the terms in the definition of the mixed volume given for instance in [Burgos Gil et al. 2014, Definition 2.7.14] one obtains

$$\mathrm{MV}_n(S,Q_1,\dots,Q_{n-1})=\sum_{k=1}^{n-1}(-1)^{n-1-k}\sum_{1\le i_1<\dots<i_k\le n-1}(\mathrm{vol}_n(S+Q_{i_1}+\dots+Q_{i_k})-\mathrm{vol}_n(Q_{i_1}+\dots+Q_{i_k}))$$

since the n-dimensional volume of a line segment vanishes for $n \geq 2$. To prove the claim it is hence enough to show that for each polytope Q in $\mathbb{R}^n$ the equality

$$\mathrm{vol}_n(S+Q) - \mathrm{vol}_n(Q) = \mathrm{vol}_{n-1}(\pi(Q))$$

holds. But $Q \subset S+Q$ and the difference of their volumes coincides with the integral over $\pi(Q) = \pi(S+Q)$ of the difference between the concave functions parametrizing the roof of the polytope $S+Q$ and Q respectively. Such a difference being constantly equal to 1 on $\pi(Q)$, the claim follows from the definition of the Lebesgue integral, concluding the proof. □

Proposition 1.12. *In the above hypotheses and notation, let g_i be a continuous concave function defined on a polytope Q_i in $M_\mathbb{R}$, for $i = 1, \dots, n$. Then,*

$$\mathrm{MI}_M(\iota_{\overline{0m}}, g_1, \dots, g_n) = \mathrm{MI}_P(\pi_* g_1, \dots, \pi_* g_n).$$

Proof. For $n = 1$, the claim follows from Corollary 1.9. Assume hence $n \geq 2$. Choose for each $i = 1, \dots, n$ a nonpositive real number γ_i such that $\gamma_i \leq \min_{x \in Q_i} g_i(x)$ and consider the convex body

$$Q_{g_i,\gamma_i} := \{(x,t) \in Q_i \times \mathbb{R} : \gamma_i \leq t \leq g_i(x)\}$$

in $M_\mathbb{R} \times \mathbb{R}$. The formula in [Philippon and Sombra 2008a, Proposition IV.5(d)] implies that

$$\begin{aligned}\mathrm{MI}_M(\iota_{\overline{0m}}, g_1, \dots, g_n) = \mathrm{MV}_{M\oplus\mathbb{Z}}(\overline{(0,0)(m,0)}, Q_{g_1,\gamma_1}, \dots, Q_{g_n,\gamma_n})& \\ + \sum_{i=1}^{n} \gamma_i \,\mathrm{MV}_M(\overline{0m}, Q_1, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n)&.\end{aligned}$$

By the hypotheses on m, $(m,0)$ is a nonzero primitive vector of the lattice $M \oplus \mathbb{Z}$. The map $\pi' := \pi \times \mathrm{id}_\mathbb{Z} : M \oplus \mathbb{Z} \to P \oplus \mathbb{Z}$ is a surjective group homomorphism, giving $(M \oplus \mathbb{Z})/\mathbb{Z}(m,0) \simeq P \oplus \mathbb{Z}$. By Lemma 1.11,

$$\begin{aligned}\mathrm{MI}_M(\iota_{\overline{0m}}, g_1, \dots, g_n) = \mathrm{MV}_{P\oplus\mathbb{Z}}(\pi'(Q_{g_1,\gamma_1}), \dots, \pi'(Q_{g_n,\gamma_n}))& \\ + \sum_{i=1}^{n} \gamma_i \,\mathrm{MV}_P(\pi(Q_1), \dots, \pi(Q_{i-1}), \pi(Q_{i+1}), \dots, \pi(Q_n))&.\end{aligned}$$

The statement follows hence from the equality in [Philippon and Sombra 2008a, Proposition IV.5(d)] applied to the concave functions $\pi_* g_1, \dots, \pi_* g_n$, the direct image of g_i by π being a concave function defined on $\pi(Q_i)$ and satisfying

$$\begin{aligned}\pi'(Q_{g_i,\gamma_i}) &= \{(\pi(y), t) \in \pi(Q_i) \times \mathbb{R} : \gamma_i \leq t \leq g_i(y)\} \\ &= \{(x,t) \in \pi(Q_i) \times \mathbb{R} : \gamma_i \leq t \leq (\pi_* g_i)(x)\}\end{aligned}$$

for every $i = 1, \dots, n$. □

2. Ronkin functions

Fix for the whole section an algebraically closed complete field $(K, |\cdot|)$, that is, a pair of an algebraically closed field and an absolute value with respect to which the field is complete. Depending on the nature of the absolute value, $(K, |\cdot|)$ is said to be archimedean or nonarchimedean. As a consequence of the Gelfand–Mazur theorem, the only archimedean algebraically closed complete field is $\mathbb{C}$ endowed with a power of the usual absolute value. When the choice of the absolute value is clear from the context, an algebraically closed complete field will be simply denoted by K.

2A. *Tropicalization.* Given an affine variety $X = \operatorname{Spec} A$ over an algebraically closed complete field K, let $\iota : K \to A$ be the corresponding ring homomorphism. Berkovich's construction allows one to define a locally ringed space $(X^{\mathrm{an}}, \mathcal{O}_{X^{\mathrm{an}}})$, called the *analytification* of X; as a set,

$$X^{\mathrm{an}} := \{\|\cdot\|_x \text{ multiplicative seminorm on } A : \|\iota(k)\|_x = |k| \text{ for all } k \in K\}.$$

We refer to [Berkovich 1990, §1.5] (or also to [Burgos Gil et al. 2014, §1.2] for a more concise treatment) for the definition of the topology and the structure sheaf on X^{an}. By a gluing argument, one can then extend the definition to any variety over K.

Remark 2.1. When $K = \mathbb{C}$ endowed with the usual absolute value, the Gelfand–Mazur theorem implies that the locally ringed space produced by Berkovich's construction agrees with the standard complex analytification of a variety over $\operatorname{Spec} \mathbb{C}$.

Let $\mathbb{T}$ denote a split torus over K of dimension n and M its character lattice, in such a way that $\mathbb{T} = \operatorname{Spec} K[M]$. A basis of the K-algebra $K[M]$ will be denoted, as in [Fulton 1993, beginning of §1.3], by $(\chi^m)_{m\in M}$ and the elements of $K[M]$ will be called *Laurent polynomials* over K. Let N be the dual lattice of M and denote by $N_{\mathbb{R}} = N \otimes \mathbb{R}$ the associated real vector space.

Definition 2.2. The *tropicalization map* $\operatorname{trop} : \mathbb{T}^{\mathrm{an}} \to N_{\mathbb{R}} = \operatorname{Hom}(M, \mathbb{R})$ is the application defined by

$$(\operatorname{trop}(\|\cdot\|_x))(m) := -\log \|\chi^m\|_x$$

for every $\|\cdot\|_x \in \mathbb{T}^{\mathrm{an}}$.

The tropicalization map turns out to be a continuous application with respect to the Berkovich topology of $\mathbb{T}^{\mathrm{an}}$ and the Euclidean topology of $N_{\mathbb{R}}$. Moreover, in the archimedean case, the choice of a basis of M allows one to write such map in the more familiar form

$$\operatorname{trop}((z_1, \dots, z_n)) = (-\log|z_1|, \dots, -\log|z_n|).$$

Remark 2.3. The tropicalization map coincides with the valuation map val used by Burgos Gil, Philippon and Sombra [Burgos Gil et al. 2014, equation 4.1.2].

When the absolute value on K is nonarchimedean, one can construct a suitable section of trop. For each $u \in N_{\mathbb{R}}$, consider the map which associates to a Laurent polynomial $f = \sum c_m \chi^m$ the real value

$$\|f\|_{\kappa(u)} := \max_m |c_m| e^{-\langle m,u\rangle}.$$

One can verify that $\|\cdot\|_{\kappa(u)}$ is a multiplicative seminorm on $K[M]$ extending the absolute value on K, and hence it corresponds to a point $\kappa(u) \in \mathbb{T}^{\mathrm{an}}$, which one calls the *Gauss point* over u. The application

$\kappa : N_{\mathbb{R}} \to \mathbb{T}^{\mathrm{an}}$ defined by

$$\kappa : u \to \|\cdot\|_{\kappa(u)}$$

is proved to be a continuous section of trop; see for instance [Burgos Gil et al. 2014, Proposition-Definition 4.2.12], noting that κ coincides with $\theta_0 \circ \boldsymbol{e}$ in the cited reference.

It is easily seen that for any closed subvariety X of $\mathbb{T}$, there is a natural inclusion of sets $X^{\mathrm{an}} \subseteq \mathbb{T}^{\mathrm{an}}$, making the following definition meaningful.

Definition 2.4. Let f be a nonzero Laurent polynomial in $K[M]$ and $V(f)$ the associated closed subvariety of $\mathbb{T}$. The subset

$$\mathcal{A}_f := \mathrm{trop}(V(f)^{\mathrm{an}}) \subseteq N_{\mathbb{R}}$$

is called the *amoeba* of f.

The so-defined set has been widely studied in the literature. In the archimedean case, it coincides (up to a change of sign) with the notion of amoeba studied by Gelfand, Kapranov and Zelevinsky [1994] and by Passare and Rullgård [2004]. In the nonarchimedean case, the amoeba of a nonzero Laurent polynomial f coincides with the corner locus of the associated tropical polynomial f^{trop}; see [Einsiedler et al. 2006, Theorem 2.1.1].

2B. ***Minimal boundaries.*** To keep the treatment of the archimedean and nonarchimedean settings uniform, the following notion is crucial.

Definition 2.5. Let K be an algebraically closed complete field and A a K-algebra. For any set $\mathcal{S}$ of multiplicative seminorms on A extending the absolute value of K, a *boundary* of $\mathcal{S}$ is a subset $\mathcal{B} \subseteq \mathcal{S}$ such that

$$\max_{x \in \mathcal{S}} \|a\|_x = \max_{x \in \mathcal{B}} \|a\|_x$$

for every $a \in A$.

In other words, any boundary of $\mathcal{S}$ contains all of the information about the maximal values that the seminorms in $\mathcal{S}$ can attain. As a trivial example, $\mathcal{S}$ is a boundary of $\mathcal{S}$.

We are especially interested in the boundary of the fibers of the tropicalization map. Keeping the notation of the previous subsection, for a point $u \in N_{\mathbb{R}}$, it is immediate to show that

$$\mathrm{trop}^{-1}(u) = \{\|\cdot\|_x \in \mathbb{T}^{\mathrm{an}} : \|\chi^m\|_x = e^{-\langle m,u\rangle} \text{ for all } m \in M\}.$$

In the complex case, the existence of a unique minimal boundary for an algebra of functions on a compact space has been proved by Shilov. The following result, which equally holds in the nonarchimedean case, is well-known by experts.

Proposition 2.6. *The set* $\mathrm{trop}^{-1}(u)$ *has a unique minimal boundary. In the archimedean case, it coincides with the whole* $\mathrm{trop}^{-1}(u)$, *while in the nonarchimedean case it consists of the Gauss point over* u.

Proof. In the archimedean case, after having chosen a basis of M, one can treat $\mathbb{T}^{\mathrm{an}}$ as $(\mathbb{C}^*)^n$, by identifying a point $z = (z_1, \ldots, z_n)$ of $(\mathbb{C}^*)^n$ with the seminorm $\|\cdot\|_z$ defined as

$$\|f\|_z = |f(z)|$$

for every $f \in K[T_1, \ldots, T_n]$. Assuming that the coordinates of u in the basis dual to the chosen one are $(u_1, \ldots, u_n)$, the set $\mathrm{trop}^{-1}(u)$ corresponds to the subset of $(\mathbb{C}^*)^n$ consisting of the points $(z_1, \ldots, z_n)$ satisfying $-\log|z_i| = u_i$ for every $i = 1, \ldots, n$. A point $\tilde{z}$ of $\mathrm{trop}^{-1}(u)$ is then of the form

$$\tilde{z} := (e^{-u_1+i\theta_1}, \ldots, e^{-u_n+i\theta_n}),$$

with $\theta_i \in [0, 2\pi)$ for every $i = 1, \ldots, n$. It is easy to check that such a point is the only one in $\mathrm{trop}^{-1}(u)$ for which

$$\|(T_1 + e^{-u_1+i\theta_1}) \cdots (T_n + e^{-u_n+i\theta_n})\|_z$$

is maximal. As a consequence, $\tilde{z}$ must belong to any boundary of $\mathrm{trop}^{-1}(u)$. This proves that the only boundary of $\mathrm{trop}^{-1}(u)$ is the set itself.

In the nonarchimedean case, for every $\|\cdot\|_x \in \mathrm{trop}^{-1}(u)$ and for every $f = \sum c_m \chi^m \in K[M]$ one has (by definition of nonarchimedean absolute value) that

$$\|f\|_x \leq \max_m |c_m| \|\chi^m\|_x = \max_m |c_m| e^{-\langle m,u\rangle}.$$

This trivially implies that the Gauss norm $\|f\|_{\kappa(u)} = \max_m |c_m| e^{-\langle m,u\rangle}$ is a boundary for $\mathrm{trop}^{-1}(u)$. □

For $u \in N_{\mathbb{R}}$, one denotes the unique minimal boundary of $\mathrm{trop}^{-1}(u)$ described in the previous proposition by $\mathcal{B}(u)$. If not otherwise mentioned, such a set is considered to be endowed with the topology induced from $\mathbb{T}^{\mathrm{an}}$. If one sets

$$\mathcal{B}_K := \begin{cases} (\mathbb{S}^1)^n & \text{if } K \text{ is archimedean,} \\ \{1\} & \text{otherwise,} \end{cases}$$

the boundary $\mathcal{B}(u)$ is homeomorphic to the compact group $\mathcal{B}_K$ for every $u \in N_{\mathbb{R}}$. This allows one to define the measure

$$\sigma_u := \mathrm{Haar}_{\mathcal{B}(u)}$$

on $\mathrm{trop}^{-1}(u)$, which is the Haar measure on the compact group $\mathcal{B}(u)$ normalized to have total mass 1; it is a finite measure on $\mathrm{trop}^{-1}(u)$, supported on $\mathcal{B}(u)$ and distributing homogeneously on this set. In the nonarchimedean case it coincides with the Dirac delta at the Gauss point over u.

There exists an embedding $\iota : N_{\mathbb{R}} \times \mathcal{B}_K \to \mathbb{T}^{\mathrm{an}}$ fitting in the commutative diagram

$$\begin{tikzcd} N_{\mathbb{R}} \times \mathcal{B}_K \arrow[r, hook, "\iota"] \arrow[d] & \mathbb{T}^{\mathrm{an}} \arrow[dl, "\mathrm{trop}"] \\ N_{\mathbb{R}} & \end{tikzcd} \tag{2-1}$$

with the vertical arrow being the projection onto the first factor. In the archimedean case, it is determined by the choice of a homeomorphism $(\mathbb{C}^*)^n \simeq N_{\mathbb{R}} \times (\mathbb{S}^1)^n$, while in the nonarchimedean case it coincides with the map $(u, 1) \mapsto \kappa(u)$. The image of ι is homeomorphic to $N_{\mathbb{R}} \times \mathcal{B}_K$ and it is a deformation retract of the analytic torus, coinciding with it if K is archimedean.

2C. *Ronkin functions.* The terminology and notation introduced in the previous subsection allow one to define Ronkin functions in the archimedean and nonarchimedean cases simultaneously.

Definition 2.7. Let f be a nonzero Laurent polynomial over K. The *Ronkin function* of f is the map $\rho_f : N_{\mathbb{R}} \to \mathbb{R}$ defined as

$$\rho_f(u) := \int_{\mathrm{trop}^{-1}(u)} -\log \|f\|_x \, d\sigma_u(x)$$

for every $u \in N_{\mathbb{R}}$.

The integral in the previous definition is finite. Indeed, logarithmic singularities are integrable in the archimedean case, and the Gauss norm of a nonzero Laurent polynomial is positive in the nonarchimedean case.

Remark 2.8. In the archimedean case, $\rho_f(u) = -N_f(-u)$, where N_f is the classical Ronkin function associated to a complex Laurent polynomial; refer for instance to [Passare and Rullgård 2004]. In the nonarchimedean case, it is easily checked that $\rho_f = f^{\mathrm{trop}}$, where the tropicalization of a Laurent polynomial $f = \sum c_m \chi^m$ is defined by

$$f^{\mathrm{trop}}(u) = \min_m(\langle m, u\rangle - \log |c_m|)$$

for every $u \in N_{\mathbb{R}}$.

The following property of the Ronkin function follows immediately from its definition.

Proposition 2.9. *For every pair of nonzero Laurent polynomials f and g, one has $\rho_{f \cdot g} = \rho_f + \rho_g$. For every $\lambda \in K$, moreover, $\rho_{\lambda \cdot f} = -\log |\lambda| + \rho_f$.*

Recall that for any Laurent polynomial $f = \sum_m c_m \chi^m$, we denote by $\mathrm{NP}(f)$ the *Newton polytope of f*, that is, the convex hull in $M_{\mathbb{R}}$ of the set $\{m \in M : c_m \neq 0\}$. Its *support function*, as defined in Example 1.1, is denoted by $\Psi_{\mathrm{NP}(f)}$.

Proposition 2.10. *Let f be a nonzero Laurent polynomial. Then*:

(1) *ρ_f is a continuous concave function on $N_{\mathbb{R}}$ (in particular it is closed) and it is affine on each connected component of the complement of the amoeba of f.*

(2) *$|\rho_f - \Psi_{\mathrm{NP}(f)}|$ is bounded on $N_{\mathbb{R}}$.*

(3) *The stability set of ρ_f coincides with* $\mathrm{NP}(f)$ *and* $\mathrm{rec}(\rho_f) = \Psi_{\mathrm{NP}(f)}$.

Proof. The statements in (1) are trivial in the nonarchimedean case. In the archimedean case, the concavity of ρ_f and its affinity outside the amoeba are shown in [Passare and Rullgård 2004, Theorem 1]. As a consequence of concavity, ρ_f is continuous on $N_{\mathbb{R}}$, and hence closed.

To prove (2), suppose that $f = \sum_m c_m \chi^m$ and let $\gamma_K(f)$ be the number of nonzero coefficients of f if K is archimedean, and 1 otherwise. For any $x \in \operatorname{trop}^{-1}(u)$, the inequality $\|f\|_x \leq \gamma_K(f) \cdot \max_m(|c_m| \|\chi^m\|_x)$ implies

$$-\log \|f\|_x \geq \min_m(\langle m, u\rangle - \log|c_m|) - \log \gamma_K(f) \geq \Psi_{\mathrm{NP}(f)}(u) - \max_m \log|c_m| - \log \gamma_K(f)$$

and hence

$$\rho_f(u) \geq \Psi_{\mathrm{NP}(f)}(u) - \log(\gamma_K(f) \max_m |c_m|)$$

for every $u \in N_{\mathbb{R}}$. For a reverse inequality, denote by $\mathcal{V}(f)$ the set of vertices of $\mathrm{NP}(f)$. Then, for every $m \in \mathcal{V}(f)$ the Ronkin function of f coincides with $\langle m, u\rangle - \log|c_m|$ in a nonempty open subset of $N_{\mathbb{R}}$ (this follows from [Passare and Rullgård 2004, Proposition 2] in the archimedean setting and from [Einsiedler et al. 2006, Corollary 2.1.2] in the nonarchimedean one). By the concavity of ρ_f, one deduces hence that

$$\rho_f(u) \leq \Psi_{\mathrm{NP}(f)}(u) - \log \min_{m \in \mathcal{V}(f)} |c_m|$$

for every $u \in N_{\mathbb{R}}$, concluding the proof of (2).

The statements in (3) follow directly from (2). Indeed, since $|\rho_f - \Psi_{\mathrm{NP}(f)}|$ is bounded, $\operatorname{stab}(\rho_f) = \operatorname{stab}(\Psi_{\mathrm{NP}(f)}) = \mathrm{NP}(f)$; the last equality follows then from [Rockafellar 1970, Theorem 13.3]. □

The calculation of the Ronkin function of a nonzero Laurent polynomial is typically very difficult. Anyway, an explicit expression for it is available in the following two simple situations.

Example 2.11. It follows from the definition that the Ronkin function of the monomial χ^m coincides with the linear function m on $N_{\mathbb{R}}$ for every $m \in M$.

Example 2.12. For any $m, m' \in M$ with $m \neq m'$, the Ronkin function of the binomial $f = \chi^m - \chi^{m'}$ coincides with the support function of the segment $\overline{mm'}$, that is,

$$\rho_f(u) = \min(\langle m, u\rangle, \langle m', u\rangle)$$

for every $u \in N_{\mathbb{R}}$. To prove this, note first that one can restrict to the case of the binomial $f = \chi^m - 1$ with $m \neq 0$ because of Proposition 2.9 and Example 2.11 and by factoring with a monomial. In the nonarchimedean case the statement follows immediately from Remark 2.8. In the archimedean one, the choice of a basis for M allows one to write

$$\rho_f(u) = -\frac{1}{(2\pi)^n} \int_{\theta_1, \dots, \theta_n \in [0, 2\pi]} \log\left|e^{-m_1 u_1 + i m_1 \theta_1} \cdot \dots \cdot e^{-m_n u_n + i m_n \theta_n} - 1\right| d\theta_1 \cdots d\theta_n,$$

with $m_1, \dots, m_n$ being the coordinates of m in such a basis and $u_1, \dots, u_n$ the coordinates of u in the dual one. Assuming that $m_1 > 0$, which is always possible since $m \neq 0$, Jensen's formula yields, for every $\theta_2, \dots, \theta_n$,

$$\int_{\theta_1 \in [0, 2\pi]} \log\left|e^{-m_1 u_1 + i m_1 \theta_1} \cdot \dots \cdot e^{-m_n u_n + i m_n \theta_n} - 1\right| d\theta_1 = -2\pi \sum_{j=1}^{k} \log \frac{|\alpha_j|}{e^{-m_1 u_1}} \tag{2-2}$$

with $\alpha_1, \dots, \alpha_k$ being the zeros of the univariate polynomial

$$(e^{-m_2u_2+im_2\theta_2} \cdot \dots \cdot e^{-m_nu_n+im_n\theta_n})T - 1$$

lying inside the closed disk of radius $e^{-m_1u_1}$, repeated according to multiplicity. The only complex zero of the above polynomial has modulus $e^{m_2u_2+\dots+m_nu_n}$; the integral in (2-2) is then zero if $m_1u_1+\dots+m_nu_n > 0$; otherwise it equals $-2\pi(m_1u_1+\dots+m_nu_n)$. It follows that

$$\rho_f(u) = \min(m_1u_1+\dots+m_nu_n, 0),$$

and hence the claim.

3. Heights of toric varieties

A well-suited framework to develop Arakelov geometry is provided by the study of varieties over adelic fields. In this setting, local and global heights of cycles of arbitrary dimension can be defined following [Zhang 1995; Gubler 1998; Chambert-Loir 2006]. A more general approach involving M-fields has been suggested in [Gubler 1997]. Even if the theory is often phrased in terms of line bundles, we adopt here the equivalent point of view of divisors, which turns out to be more convenient in the toric case; see for instance [Burgos Gil et al. 2015].

3A. *Adelic fields.* By a *place* on a field K, we mean an equivalence class of absolute values on K, which could be either archimedean or nonarchimedean. Whenever $\mathfrak{M}$ is a collection of places on K, the subset of archimedean places in $\mathfrak{M}$ is denoted by $\mathfrak{M}_\infty$.

Definition 3.1. Let K be a field. A family of places $\mathfrak{M}$ on K is said to be *adelic* if it satisfies the following properties:

(1) For every $v \in \mathfrak{M}\setminus\mathfrak{M}_\infty$, one (and hence all) absolute value in the class of v is associated to a nontrivial discrete valuation.

(2) For each $\alpha \in K^*$, the set of places v for which $|\alpha|_v \neq 1$ for any $|\cdot|_v \in v$ is finite.

It is clear that the two conditions of the previous definition do not depend on the choice of the representative of the class v.

Definition 3.2. An *adelic field* is a field K together with an adelic family of places $\mathfrak{M}$ on K and a choice of an absolute value $|\cdot|_v$ and of a real positive number n_v for each place $v \in \mathfrak{M}$. An adelic field $(K, (|\cdot|_v, n_v)_{v\in\mathfrak{M}})$ is said to satisfy the *product formula* if for every $\alpha \in K^*$

$$\sum_{v\in\mathfrak{M}} n_v \log|\alpha|_v = 0.$$

Whenever there is no ambiguity on its adelic structure, an adelic field will be simply denoted by K.

The following property is an easy, though fundamental, consequence of the definition.

Lemma 3.3. *Any adelic field K only admits finitely many archimedean places.*

Proof. Since an absolute value on K is nonarchimedean if and only if it is bounded on the image of $\mathbb{Z}$ in K, a field with positive characteristic has no archimedean absolute values. Suppose hence that K has characteristic zero. In this case it contains a copy of $\mathbb{Q}$ and any archimedean absolute value $|\cdot|_v$ on K restricts to an archimedean absolute value on $\mathbb{Q}$. By Ostrowski's theorem, one has $|2|_v > 1$. The second axiom in Definition 3.1 then allows one to conclude the claim. □

For an adelic field K and a finite field extension F of K, there exists a canonical way of endowing F with the structure of an adelic field; see [Gubler 1997, Remark 2.5] and [Martínez and Sombra 2018, §3] for the detailed construction. With this induced adelic structure, F satisfies the product formula whenever K does.

Example 3.4. The archetypal example of an adelic field satisfying the product formula is given by the field $\mathbb{Q}$, together with the collection of all its nontrivial places, the standard normalized absolute value for each of them and weights equal to 1.

More generally, any *global field*, that is, a number field or the function field of a smooth projective curve over a field k with the structure described in [Burgos Gil et al. 2014, Example 1.5.4], is an adelic field satisfying the product formula.

3B. ***Local and global heights.*** Let K be an adelic field satisfying the product formula and X a proper variety of dimension n over K. For every place $v \in \mathfrak{M}$, denote by K_v the completion of K with respect to $|\cdot|_v$ and by $\mathbb{C}_v$ the completion of an algebraic closure of K_v with respect to the unique extension of the absolute value. It is a well-known fact that $\mathbb{C}_v$ is algebraically closed; moreover, $\mathbb{C}_v$ comes with an absolute value that one denotes, with abuse of notation, by $|\cdot|_v$. The pair $(\mathbb{C}_v, |\cdot|_v)$ is hence an algebraically closed complete field as in Section 2.

The base change $X_{\mathbb{C}_v}$ is a scheme of finite type over $\operatorname{Spec} \mathbb{C}_v$ to which one associates its Berkovich analytification $(X_v^{\mathrm{an}}, \mathscr{O}_{X_v^{\mathrm{an}}})$, whose underlying topological space is compact because of the properness of X. To stress its dependence on the choice of the place v, X_v^{an} is called the *v-adic analytification* of X. Similarly, one can consider the base change X_{K_v} of X over $\operatorname{Spec} K_v$ and consider its Berkovich analytification $X_{K_v}^{\mathrm{an}}$. The two spaces are related by the isomorphism

$$X_{K_v}^{\mathrm{an}} \simeq X_v^{\mathrm{an}} / \operatorname{Gal}(\overline{K}_v^{\mathrm{sep}}/K_v),$$

as shown in [Berkovich 1990, Proposition 1.3.5]. Moreover, there exists a surjective morphism of locally ringed space $\pi_v : X_v^{\mathrm{an}} \to X_{\mathbb{C}_v}$.

Remark 3.5. By Ostrowski's theorem and the Gelfand–Mazur theorem, if v is an archimedean absolute value on K, then $\mathbb{C}_v$ is isometric to the field $\mathbb{C}$ endowed with a power of the usual absolute value. In this case, the Berkovich space $(X_v^{\mathrm{an}}, \mathscr{O}_{X_v^{\mathrm{an}}})$ is isomorphic to the usual complex analytification of $X_{\mathbb{C}}$.

For any line bundle L on X, its v-adic analytification is the analytic line bundle

$$L_v^{\mathrm{an}} := \pi_v^* L_{\mathbb{C}_v}$$

on X_v^{an}. Continuous metrics on L_v^{an} are defined as in [Chambert-Loir 2011, §1.1.1], independently of the nature of the place v. Relevant classes of metrics on L_v^{an} are *smooth* metrics in the archimedean case and *algebraic* (or, equivalently, formal, see [Gubler and Künnemann 2017, Proposition 8.13]) metrics when v is nonarchimedean; see for example [Chambert-Loir 2011, §1] and [Gubler and Künnemann 2017, 8.8 and 8.12] for the precise definitions. A divisor D on X together with a continuous $\mathrm{Gal}(\overline{K}_v^{\mathrm{sep}}/K_v)$-invariant metric $\|\cdot\|_v$ on the analytic line bundle $\mathscr{O}(D)_v^{\mathrm{an}}$ is called a *v-adic metrized divisor* and it is denoted by $\overline{D}_v$ or also by $(\mathscr{O}(D), \|\cdot\|_v)$. Sums and pull-backs of v-adic metrized divisors can be defined as in [Chambert-Loir 2011, §1.2].

A v-adic metrized divisor $\overline{D}_v$ is said to be *semipositive* if the corresponding metric can be approximated by semipositive smooth (when v is archimedean) or algebraic (when v is nonarchimedean) semipositive metrics in the sense of [Burgos Gil et al. 2014, §1.4]. For any d-dimensional subvariety Y of X and for any d-tuple of v-adic semipositive metrized divisors $\overline{D}_{0,v}, \dots, \overline{D}_{d-1,v}$ on X, there exists a positive measure

$$c_1(\overline{D}_{0,v}) \wedge \cdots \wedge c_1(\overline{D}_{d-1,v}) \wedge \delta_Y \tag{3-1}$$

on X_v^{an}, which was first introduced in [Chambert-Loir 2006, Définition 2.4 and Proposition 2.7(b)] in the nonarchimedean setting and extended in [Gubler 2007, §3.8] under weaker assumptions. The suggestive notation for the measure in (3-1) is compatible with the wedge product of first Chern forms in the smooth archimedean case, while it is justified by the recent advances in the theory of forms and currents on Berkovich spaces otherwise, as shown in [Chambert-Loir and Ducros 2012, §6.9] and [Gubler and Künnemann 2017, Theorem 10.5].

Recall also that for a d-dimensional cycle Z in X and a family $(D_0, s_0), \dots, (D_d, s_d)$ of divisors on X with rational sections of the associated line bundles, one says that $s_0, \dots, s_d$ *meet Z properly* if for every $J \subseteq \{0, \dots, d\}$, each irreducible component of $|Z| \cap \bigcap_{i \in J} |\mathrm{div}(s_i)|$ has dimension $d - \#J$, where $|\cdot|$ denotes here the support of a cycle.

Definition 3.6. Let Z be a d-dimensional cycle in X and $(\overline{D}_{0,v}, s_0), \dots, (\overline{D}_{d,v}, s_d)$ a collection of v-adic semipositive metrized divisors on X with rational sections of the corresponding line bundles, with $s_0, \dots, s_d$ meeting Z properly. The *v-adic local height of Z in X with respect to* $(\overline{D}_{i,v}, s_i)$ for $i = 0, \dots, d$ is defined, linearly in its irreducible components, by the recursive formula

$$\begin{aligned} &h_{\overline{D}_{0,v},\dots,\overline{D}_{d,v}}(Z; s_0, \dots, s_d) \\ &\qquad := h_{\overline{D}_{0,v},\dots,\overline{D}_{d-1,v}}(Z \cdot \mathrm{div}(s_d); s_0, \dots, s_{d-1}) - \int_{X_v^{\mathrm{an}}} \log \|s_d\|_{d,v}\, c_1(\overline{D}_{0,v}) \wedge \cdots \wedge c_1(\overline{D}_{d-1,v}) \wedge \delta_Z, \end{aligned}$$

where $\|\cdot\|_{d,v}$ denotes the metric of $\overline{D}_{d,v}$ and one sets the height of the zero cycle to be zero.

The integrals appearing in the previous definition are well-defined, as shown in [Chambert-Loir and Thuillier 2009, Théorème 4.1] in both the archimedean and nonarchimedean settings, and in [Gubler and Hertel 2017, Theorem 1.4.3] for the case of nonarchimedean valuations which are not necessarily discrete. The v-adic local height function is moreover symmetric and multilinear with respect to sums of metrized divisors with rational sections of the associated line bundles; see [Gubler 2003, Proposition 3.4 and Remark 9.3].

The adelic structure on the field K allows one to define a *semipositive metrized divisor* $\overline{D}$ on X by the choice, for every place $v \in \mathfrak{M}$, of a continuous semipositive metric on $\mathscr{O}(D)_v^{\mathrm{an}}$. This global definition induces a notion of a v-adic local height function at each place of K. Some care has to be taken when defining global heights as sums of such v-adic local heights, since they do not need to be well-defined in general.

Definition 3.7. A d-dimensional irreducible subvariety Y of X is said to be *integrable* with respect to the choice of $d+1$ semipositive metrized divisors $\overline{D}_0, \dots, \overline{D}_d$ if there exists a birational proper map $\varphi : Y' \to Y$, with Y' projective, and sections s_i of $\varphi^*\mathscr{O}(D_i)$ for each $i = 0, \dots, d$, meeting Y' properly, such that the v-adic local height

$$h_{\varphi^*\overline{D}_{0,v},\dots,\varphi^*\overline{D}_{d,v}}(Y'; s_0, \dots, s_d)$$

is zero for all but finitely many places $v \in \mathfrak{M}$. A d-dimensional cycle is said to be *integrable* if each of its irreducible components is. If Y is an integrable d-dimensional irreducible subvariety, the *global height of Y in X with respect to $\overline{D}_0, \dots, \overline{D}_d$* is defined as

$$h_{\overline{D}_0,\dots,\overline{D}_d}(Y) := \sum_{v\in\mathfrak{M}} n_v \, h_{\varphi^*\overline{D}_{0,v},\dots,\varphi^*\overline{D}_{d,v}}(Y'; s_0, \dots, s_d).$$

The *global height* of integrable cycles is defined by linearity.

The previous definition does not depend on the choice of the projective resolution Y' of Y nor of the sections $s_0, \dots, s_d$, as a consequence of [Gubler 2003, Proposition 3.6 and Remark 9.3], [Burgos Gil et al. 2014, Theorem 1.4.17(3)] and the product formula on K. As its local counterparts, the global height is symmetric and multilinear with respect to sums of metrized divisors. Moreover, it is well-behaved under proper transformations, in the sense of the next proposition.

Proposition 3.8. *Let $\varphi : X' \to X$ be a dominant morphism of proper varieties over K, $\overline{D}_0, \dots, \overline{D}_d$ semipositive metrized divisors over X and Z' a d-dimensional cycle in X'. The cycle $\varphi_* Z'$ is integrable with respect to $\overline{D}_0, \dots, \overline{D}_d$ if and only if Z' is integrable with respect to $\varphi^*\overline{D}_0, \dots, \varphi^*\overline{D}_d$ and in this case*

$$h_{\overline{D}_0,\dots,\overline{D}_d}(\varphi_* Z') = h_{\varphi^*\overline{D}_0,\dots,\varphi^*\overline{D}_d}(Z').$$

Proof. The statement about integrability is [Burgos Gil et al. 2014, Proposition 1.5.8(2)], while the equality of the global heights follows from the same property on local heights, as proved in [Gubler 2003, Proposition 3.6 and Remark 9.3] in the more general context of pseudodivisors. □

3C. *Heights on toric varieties.* In this section the basic constructions and results of the arithmetic geometry of toric varieties are recalled, following the treatment of [Burgos Gil et al. 2014]. Let hence X_Σ be a proper toric variety of dimension n over an adelic field K, with torus $\mathbb{T}$ and dense open orbit X_0. Denote by N and M the character and cocharacter groups of $\mathbb{T}$ and by $N_\mathbb{R}$ and $M_\mathbb{R}$ the associated (reciprocally dual) real vector spaces. The toric variety X_Σ is associated with a complete fan Σ in $N_\mathbb{R}$, whose collection of k-dimensional cones is written $\Sigma^{(k)}$. For any $\tau \in \Sigma^{(1)}$, denote by v_τ its minimal nonzero integral vector and by $V(\tau)$ its associated orbit closure, as in [Fulton 1993, §3.1].

Divisors on X_Σ which are invariant under the torus action are called *toric divisors* and admit a nice combinatorial description, as follows. By [Kempf et al. 1973, §I.2, Theorem 9], there exists a bijection between the set of toric Cartier divisors on X_Σ and the set of *virtual support functions* on Σ, that is, piecewise linear real-valued functions on the support of Σ, with integral slope on each cone of Σ. The toric Cartier divisor constructed from the virtual support function Ψ is denoted by D_Ψ and it defines a distinguished rational section s_Ψ of $\mathscr{O}(D_\Psi)$ satisfying $\operatorname{div}(s_\Psi) = D_\Psi$. The corresponding Weil divisor is given by

$$[D_\Psi] = \sum_{\tau \in \Sigma^{(1)}} -\Psi(v_\tau) V(\tau). \tag{3-2}$$

In particular, the rational section s_Ψ is regular and nowhere-vanishing on X_0. A toric divisor D_Ψ also determines a polyhedron

$$\Delta_\Psi := \{x \in M_\mathbb{R} : x - \Psi \geq 0\}$$

in $M_\mathbb{R}$, which is bounded because of the properness of X_Σ, see [Fulton 1993, Proposition on p. 67], and coincides with the stability set of Ψ if Ψ is concave. Many algebro-geometric properties of a toric divisor are read from its associated virtual support function: for instance, D_Ψ is generated by global sections if and only if Ψ is concave, and it is ample if and only if Ψ is concave and restricts to different linear functions on different maximal cones of Σ.

Regarding the arithmetic part of the toric dictionary, let D_Ψ be a toric divisor on X_Σ, with associated virtual support function Ψ. The continuous metrics on $\mathscr{O}(D_\Psi)$ admitting a combinatorial description are the ones which are invariant under the action of a certain compact torus; see [Burgos Gil et al. 2014, §4.2] for more details about this notion. In concrete terms, a continuous metric $\|\cdot\|_v$ on $\mathscr{O}(D_\Psi)_v^{\mathrm{an}}$ is called a *v-adic toric metric* if the map

$$(X_0)_v^{\mathrm{an}} \to \mathbb{R}, \qquad p \mapsto \|s_\Psi(p)\|_v,$$

is constant along the fibers of the v-adic tropicalization map $\operatorname{trop}_v : (X_0)_v^{\mathrm{an}} \to N_\mathbb{R}$ introduced in Definition 2.2. A toric divisor D together with a v-adic toric metric on $\mathscr{O}(D)$ is called a *v-adic toric metrized divisor*. To a v-adic toric metrized divisor $\overline{D}_v$ one can associate the real-valued map $\psi_{\overline{D}_v}$ on $N_\mathbb{R}$ satisfying the equality

$$\psi_{\overline{D}_v} \circ \operatorname{trop}_v = \log \|s_D\|_v \tag{3-3}$$

on the analytic torus $(X_0)_v^{\mathrm{an}}$, where s_D is the distinguished rational section of $\mathscr{O}(D)$.

The map $\psi_{\overline{D}_v}$, which will be referred to as the *metric function* of $\overline{D}_v$, was introduced by Burgos Gil, Philippon and Sombra in their study of Arakelov geometry of toric varieties to encode many arithmetic properties of $\overline{D}_v$; see [Burgos Gil et al. 2014, Chapter 4]. For instance, it is smooth in the archimedean case if the metric is smooth, while in the nonarchimedean setting it is rational piecewise affine if the metric is algebraic; see [Burgos Gil et al. 2014, Theorem 4.5.10(1)] and [Gubler and Hertel 2017, Proposition 2.5.5]. Also, the semipositivity of $\overline{D}_v$ is translated into the concavity of its corresponding metric function.

Theorem 3.9. *Let D be the toric divisor associated to the virtual support function Ψ. The assignment $\|\cdot\|_v \mapsto \psi_{\overline{D}_v}$ is a bijection between the space of v-adic semipositive toric metrics on $\mathscr{O}(D)_v^{\mathrm{an}}$ and the space of concave functions ψ on $N_{\mathbb{R}}$ such that $|\psi - \Psi|$ is bounded.*

Proof. This is [Burgos Gil et al. 2014, Theorem 4.8.1(1)]. The extension to the general nonarchimedean case is [Gubler and Hertel 2017, Theorem 2.5.8]. □

If $\overline{D}_v$ is a v-adic semipositive toric metrized divisor, the Legendre–Fenchel dual of the metric function of $\overline{D}_v$ is called the *roof function* of $\overline{D}_v$ and denoted by $\vartheta_{\overline{D}_v}$: it is a concave function on $M_{\mathbb{R}}$ with effective domain the polytope Δ_Ψ.

A toric divisor admits a v-adic semipositive metric if and only if it is generated by global sections, as proved in [Burgos Gil et al. 2014, Corollary 4.8.5]. For such divisors, moreover, there exists a distinguished choice of a v-adic semipositive metric.

Definition 3.10. Let D be a toric divisor generated by global sections, and Ψ its associated virtual support function. The *v-adic canonical metric* on D is the semipositive toric metric on $\mathscr{O}(D)_v^{\mathrm{an}}$ corresponding to Ψ in the bijection of Theorem 3.9.

In the nonarchimedean case, the canonical metric on D coincides with the algebraic metric induced by the canonical model of X_Σ and D; see [Burgos Gil et al. 2014, Example 4.5.4].

For semipositive v-adic toric metrized divisors, the measure in (3-1) can be expressed in terms of the associated metric functions. Indeed, recall from Section 2B that there exists an embedding

$$\iota_v : N_{\mathbb{R}} \times \mathcal{B}_{\mathbb{C}_v} \to X_{0,v}^{\mathrm{an}}$$

which fits into the commutative diagram (2-1), and denote by $\mathrm{Haar}_{\mathcal{B}_{\mathbb{C}_v}}$ the Haar measure on $\mathcal{B}_{\mathbb{C}_v}$ normalized to have total mass 1.

Theorem 3.11. *For $i = 0, \dots, n-1$, let $\overline{D}_{i,v}$ be a semipositive v-adic toric metrized divisor on X_Σ, Ψ_i the virtual support function associated to D_i and $\psi_{i,v}$ the metric function of $\overline{D}_{i,v}$. Then, the positive measure*

$$c_1(\overline{D}_{0,v}) \wedge \cdots \wedge c_1(\overline{D}_{n-1,v}) \wedge \delta_{X_\Sigma}$$

is zero outside $X_{0,v}^{\mathrm{an}}$ and

$$c_1(\overline{D}_{0,v}) \wedge \cdots \wedge c_1(\overline{D}_{n-1,v}) \wedge \delta_{X_\Sigma}|_{X_{0,v}^{\mathrm{an}}} = (\iota_v)_*(\mathrm{M}\mathcal{M}_M(\psi_{0,v}, \dots, \psi_{n-1,v}) \times \mathrm{Haar}_{\mathcal{B}_{\mathbb{C}_v}}).$$

In particular,

$$(\mathrm{trop}_v)_*(c_1(\overline{D}_{0,v}) \wedge \cdots \wedge c_1(\overline{D}_{n-1,v}) \wedge \delta_{X_\Sigma}|_{X_{0,v}^{\mathrm{an}}}) = \mathrm{M}\mathcal{M}_M(\psi_{0,v}, \dots, \psi_{n-1,v})$$

as measures on $N_{\mathbb{R}}$.

Proof. The first statement follows from [Burgos Gil et al. 2014, Theorem 1.4.10(1)] and [Gubler and Hertel 2017, Corollary 1.4.5]. The expression for the measure in the archimedean and the discrete nonarchimedean case is obtained from [Burgos Gil et al. 2014, Theorem 4.8.11] and multilinearity; the general nonarchimedean case is deduced from [Gubler and Hertel 2017, Theorem 2.5.10]. The last assertion is an easy consequence of the commutativity of the diagram (2-1). □

Moving to the global case, a (*semipositive*) *toric metric* on a toric divisor D is a choice, for each place $v \in \mathfrak{M}$, of a (semipositive) v-adic toric metric on the line bundle $\mathscr{O}(D)$. The toric divisor D together with a (semipositive) toric metric is called a (*semipositive*) *toric metrized divisor* and it is denoted by $\overline{D}$. From the point of view of convex geometry, the semipositive toric metrized divisor $\overline{D}$ is completely described by the collection $(\psi_v)_{v\in\mathfrak{M}}$ of its metric functions or, equivalently, by the collection $(\vartheta_v)_{v\in\mathfrak{M}}$ of its roof functions.

A notion of well-behaving toric metrics was defined in [Burgos Gil et al. 2014, Definition 4.9.1].

Definition 3.12. A toric metric $(\|\cdot\|_v)_{v\in\mathfrak{M}}$ on a toric divisor is said to be *adelic* if for all but finitely many $v \in \mathfrak{M}$ the v-adic toric metric $\|\cdot\|_v$ is the canonical one, in the sense of Definition 3.10.

In convex terms, a toric metric on the toric divisor D associated to the virtual support function Ψ is adelic if and only if the family $(\psi_v)_v$ of its metric functions satisfies $\psi_v = \Psi$ for all but finitely many $v \in \mathfrak{M}$.

It follows from [Burgos Gil et al. 2014, Theorem 5.2.4] that any toric subvariety of X_Σ is integrable with respect to the choice of adelic semipositive toric metrized divisors. In particular, one can compute the global height of the n-dimensional cycle X_Σ with respect to such choices.

Theorem 3.13. *Let $\overline{D}_0, \dots, \overline{D}_n$ be toric divisors over X_Σ, equipped with adelic semipositive toric metrics. Then*

$$h_{\overline{D}_0,\dots,\overline{D}_n}(X_\Sigma) = \sum_{v\in\mathfrak{M}} n_v \operatorname{MI}_M(\vartheta_{0,v}, \dots, \vartheta_{n,v}),$$

where $\vartheta_{i,v}$ is the roof function of $\overline{D}_{i,v}$ for every $i = 0, \dots, n$ and $v \in \mathfrak{M}$.

Proof. This is [Burgos Gil et al. 2014, Theorem 5.2.5]. □

4. Divisors of rational functions

The present section focuses on the combinatorial description of the Weil divisor on a toric variety of the rational function coming from a Laurent polynomial. This result will be used in the proof of the main theorems in the next section.

To fix notation, let X_Σ be a proper smooth toric variety of dimension n over a field K, M the character lattice of its torus $\mathbb{T}$ and $N_\mathbb{R}$ the associated dual vector space. The toric variety X_Σ has a dense open orbit X_0 isomorphic to $\mathbb{T}$ and hence the function field of X_Σ coincides with $K(M)$. In particular, any Laurent polynomial $f = \sum c_m \chi^m$ gives rise to a rational function on X_Σ, which is regular and coincides with f on X_0. To avoid confusion, one will denote by $\tilde{f}$ such a rational function.

Recall from [Cox et al. 2011, Theorem 3.1.19(a)] that the toric variety X_Σ is smooth if and only if each cone of Σ is a smooth cone, that is, it is generated by a part of a basis of the lattice N. For each cone τ in Σ of dimension 1, denote by v_τ its minimal nonzero integral vector, which generates $\tau \cap N$ as a monoid. If σ is a smooth cone of dimension n in $N_\mathbb{R}$, the collection $(v_\tau)_\tau$, with τ ranging in the set of 1-dimensional faces of σ, is a basis of N and hence gives a dual basis $(v_\tau^\vee)_\tau$ of the lattice M.

Lemma 4.1. *Let σ be a strongly convex polyhedral rational cone in $N_\mathbb{R}$. For every face τ of σ of dimension 1, the orbit closure $V(\tau)$ in the affine toric variety X_σ is the subvariety corresponding to the prime ideal*

$$\mathfrak{p} = (\chi^m : m \in \sigma^\vee \cap M, m \notin \tau^\perp)$$

of $\mathscr{O}(X_\sigma) = K[\sigma^\vee \cap M]$. Moreover, if σ is smooth and of maximal dimension in $N_\mathbb{R}$, then $\mathfrak{p}$ is principal and generated by $\chi^{v_\tau^\vee}$.

Proof. Recall for example from [Fulton 1993, §3.1] that the orbit closure $V(\tau)$ is the toric variety $\operatorname{Spec} K[\sigma^\vee \cap \tau^\perp \cap M]$ and can be embedded in $X_\sigma = \operatorname{Spec} K[\sigma^\vee \cap M]$ via the surjection of rings sending χ^m to itself if $m \in \tau^\perp$, and to 0 otherwise. Then $V(\tau)$ is seen as the subvariety of X_σ corresponding to the kernel of such homomorphism, that is,

$$\mathfrak{p} = \bigoplus_{\substack{m \in \sigma^\vee \cap M \\ m \notin \tau^\perp}} K\chi^m = (\chi^m : m \in \sigma^\vee \cap M, m \notin \tau^\perp),$$

proving the first statement.

Suppose now that σ is smooth and of dimension n in $N_\mathbb{R}$; denote by $v_1, v_2, \ldots, v_n$ the basis of N given by the minimal integral vectors of the rays of σ, with the assumption that $v_1 = v_\tau$. By definition,

$$\sigma = \mathbb{R}_{\geq 0} v_1 + \cdots + \mathbb{R}_{\geq 0} v_n.$$

As a result, denoting by $(v_i^\vee)_{i=1,\ldots,n}$ the basis of M dual to $(v_i)_{i=1,\ldots,n}$, one has that

$$\langle v_i^\vee, u \rangle = \lambda_i \geq 0$$

for every $i \in \{1, \ldots, n\}$ and for every $u = \sum_i \lambda_i v_i \in \sigma$. In particular, $v_\tau^\vee \in \sigma^\vee$. It is easy to check that $v_\tau^\vee$ is integrally valued on each element of N and hence it belongs to M. It follows then from $\langle v_\tau^\vee, v_\tau \rangle = 1$ that

$$(\chi^{v_\tau^\vee}) \subseteq \mathfrak{p}.$$

For the reverse inclusion, consider $m \in \sigma^\vee \cap M$ with $m \notin \tau^\perp$. By assumption, $\langle m, v_\tau \rangle \in \mathbb{Z}$ and $\langle m, v_\tau \rangle \geq 0$; moreover, since $m \notin \tau^\perp$, one has $\langle m, v_\tau \rangle \geq 1$. For each $u = \sum_i \lambda_i v_i \in \sigma$ one has

$$\langle m - v_\tau^\vee, u \rangle = \lambda_1 \langle m - v_\tau^\vee, v_\tau \rangle + \sum_{i \geq 2} \lambda_i \langle m, v_i \rangle \geq \lambda_1 (\langle m, v_\tau \rangle - 1) \geq 0.$$

As a result, $m - v_\tau^\vee \in \sigma^\vee \cap M$ and hence $\chi^m = \chi^{v_\tau^\vee} \cdot \chi^{m - v_\tau^\vee} \in (\chi^{v_\tau^\vee})$, completing the proof. $\square$

Remark 4.2. The last statement of the previous lemma is not true for a general strongly convex polyhedral rational cone σ of maximal dimension in N. For example, if σ has more than n faces of dimension 1, the divisor class group of X_σ, which is generated by the classes of the orbit closures associated to the rays, turns out to be nontrivial, as a consequence of [Fulton 1993, Proposition on p. 63].

For a nonzero Laurent polynomial $f \in K[M]$, the subset $V(f)$ of zeros of f in X_0 is a closed subscheme of the dense open orbit. Its closure in X_Σ is a closed subscheme of X_Σ, denoted by $\overline{V(f)}$. Taking into account multiplicities, one can consider the associated Weil divisor $[\overline{V(f)}]$. It is the zero cycle when f is a monomial.

Theorem 4.3. *Let f be a nonzero Laurent polynomial and $\tilde{f}$ the rational function on X_Σ arising from f. Then,*

$$\operatorname{div}(\tilde{f}) = [\overline{V(f)}] + \sum_{\tau \in \Sigma^{(1)}} \Psi_{\mathrm{NP}(f)}(v_\tau) V(\tau),$$

where $\Psi_{\mathrm{NP}(f)}$ denotes the support function of the Newton polytope of f. In particular, $[\overline{V(f)}]$ is rationally equivalent to the cycle $-\sum_{\tau \in \Sigma^{(1)}} \Psi_{\mathrm{NP}(f)}(v_\tau) V(\tau)$ on X_Σ.

Proof. By [Fulton 1993, formula on p. 55], the irreducible components of $X_\Sigma \setminus X_0$ are exactly the orbit closures $V(\tau)$, with τ ranging in the set of 1-dimensional cones of Σ. Since moreover the restriction of $\tilde{f}$ to X_0 is the regular function f, it follows from the classical theory of divisors that

$$\operatorname{div}(\tilde{f}) = [\overline{V(f)}] + \sum_{\tau} \nu_\tau(\tilde{f}) V(\tau),$$

where $\nu_\tau(\tilde{f}) \in \mathbb{Z}$ is the order of vanishing of $\tilde{f}$ along $V(\tau)$. The statement of the theorem then follows from the fact that, for every $\tau \in \Sigma^{(1)}$, such an order equals $\Psi_{\mathrm{NP}(f)}(v_\tau)$.

This claim can be proved locally; fix a ray $\tau \in \Sigma^{(1)}$ and let σ be any maximal-dimensional cone of Σ containing τ. The fan being complete and consisting of smooth cones, such a σ exists and the minimal integral vectors $v_1, \dots, v_n$ of its rays are a basis of N. Assume moreover that $v_1 = v_\tau$ and, for simplicity, denote by $R := K[\sigma^\vee \cap M]$ the ring of regular functions over X_σ. The order of vanishing of $\tilde{f}$ along $V(\tau)$ is computed as the valuation of $\tilde{f}$ determined by the valuation ring $R_{\mathfrak{p}}$, the localization of R at the prime ideal $\mathfrak{p}$ corresponding to the subvariety $V(\tau)$ in X_σ. By Lemma 4.1, since the cone σ is smooth and maximal-dimensional, one has that $\mathfrak{p} = (\chi^{v_\tau^\vee})$. The maximal ideal $\mathfrak{p}R_{\mathfrak{p}}$ of $R_{\mathfrak{p}}$ is hence the principal ideal generated by $\chi^{v_\tau^\vee}$.

Suppose first that $\tilde{f} = \sum_m c_m \chi^m$ lies in R, that is, every m appearing in $\tilde{f}$ belongs to $\sigma^\vee \cap M$. By definition of the valuation in $R_{\mathfrak{p}}$,

$$\begin{aligned} \nu_\tau(\tilde{f}) &= \max\{l \in \mathbb{N} : \tilde{f} \in (\mathfrak{p}R_{\mathfrak{p}})^l\} = \max\,\{l \in \mathbb{N} : \tilde{f} \in (\chi^{l v_\tau^\vee})\} \\ &= \max\,\{l \in \mathbb{N} : \chi^{m - l v_\tau^\vee} \in R_{\mathfrak{p}} \text{ for all } m \text{ with } c_m \neq 0\}. \end{aligned}$$

The condition $\chi^{m - l v_\tau^\vee} \in R_{\mathfrak{p}}$ is equivalent to the fact that $\langle m, v_\tau \rangle \geq l$. Indeed, if the first is true, then

$$\langle m, v_\tau \rangle - l = \langle m - l v_\tau^\vee, v_\tau \rangle \geq 0.$$

Conversely, for each $u = \sum_i \lambda_i v_i \in \sigma$ one has

$$\langle m - l v_\tau^\vee, u\rangle = \lambda_1(\langle m, v_\tau\rangle - l) + \sum_{i\geq 2} \lambda_i \langle m, v_i\rangle \geq 0,$$

and so $m - l v_\tau^\vee \in \sigma^\vee \cap M$. As a consequence,

$$\begin{aligned} \nu_\tau(\tilde{f}) &= \max\{l \in \mathbb{N} : \langle m, v_\tau\rangle \geq l \text{ for all } m \text{ with } c_m \neq 0\} \\ &= \min\{\langle m, v_\tau\rangle : m \text{ with } c_m \neq 0\} = \Psi_{\mathrm{NP}(f)}(v_\tau). \end{aligned}$$

For a general $\tilde{f} = \sum_m c_m \chi^m$, the fact that $\sigma^\vee$ has dimension n in $M_\mathbb{R}$ (σ is indeed strongly convex) ensures that there exists a big enough vector $m_0 \in \sigma^\vee \cap M$ for which $m + m_0 \in \sigma^\vee \cap M$ for each m such that $c_m \neq 0$. Hence

$$\tilde{f} = \frac{\sum_m c_m \chi^{m+m_0}}{\chi^{m_0}},$$

with both the numerator and the denominator belonging to R. Applying the result for such elements one deduces

$$\nu_\tau(\tilde{f}) = \nu_\tau\left(\sum_m c_m \chi^{m+m_0}\right) - \nu_\tau(\chi^{m_0}) = \Psi_{\mathrm{NP}(f)+m_0}(v_\tau) - \langle m_0, v_\tau\rangle = \Psi_{\mathrm{NP}(f)}(v_\tau),$$

concluding the proof. □

5. Local and global heights of hypersurfaces

Fix for the whole section an adelic field $(K, (|\cdot|_v, n_v)_{v\in\mathfrak{M}})$ satisfying the product formula. Let X_Σ be a proper toric variety over K, of dimension n, with torus $\mathbb{T} = \operatorname{Spec} K[M]$ and dense open orbit X_0. For an effective cycle Z on X_Σ of pure codimension 1, whose prime components intersect X_0, we present a series of results concerning its integrability and its local and global height with respect to suitable choices of metrized divisors on X_Σ.

5A. ***Degrees.*** With the notation and assumptions given above, the effective cycle Z can be written as

$$Z = \sum_{i=1}^r \ell_i Y_i$$

for prime divisors $Y_1, \ldots, Y_r$ intersecting X_0. For every $i = 1, \ldots, r$, the closed irreducible subvariety of X_0 obtained as the intersection between Y_i and X_0 is associated to a prime ideal of height 1 in $K[M]$, which is principal since $K[M]$ is a unique factorization domain; denote by f_i an irreducible Laurent polynomial generating such an ideal. The Laurent polynomial $f = f_1^{\ell_1} \cdot \cdots \cdot f_r^{\ell_r}$ is called a *defining polynomial* for the cycle Z and is uniquely defined up to multiplication by an invertible element of $K[M]$, which means by a monomial. Moreover,

$$[\overline{V(f)}] = Z; \tag{5-1}$$

that is, the cycle associated to the closure of the subscheme $V(f)$ in X_Σ agrees with Z. Let

$$\Psi_f := \Psi_{\mathrm{NP}(f)}$$

be the support function, in the sense of Example 1.1, of the Newton polytope $\mathrm{NP}(f)$ of f; it is a piecewise linear function on $N_{\mathbb{R}}$. It is not necessarily a virtual support function on the fan Σ, but it can always be made such after a suitable refinement of the fan.

Lemma 5.1. *For any proper toric variety X_Σ there exist a smooth projective toric variety $X_{\Sigma'}$ with fan Σ' in $N_{\mathbb{R}}$ and a proper toric morphism $\pi : X_{\Sigma'} \to X_\Sigma$ satisfying*:

(1) *π restricts to the identity on the dense open orbit of $X_{\Sigma'}$ and X_Σ.*

(2) *Ψ_f is a virtual support function on Σ'.*

Proof. One can always refine the complete fan Σ to a fan Σ' in such a way that Ψ_f is a virtual support function on Σ'. After possibly refining again, one can suppose that Σ' is the fan of a projective toric variety (because of the toric Chow lemma, see [Cox et al. 2011, Theorem 6.1.18]) and that each of its cones is generated by a part of a basis of N; see [Fulton 1993, §2.6]. The associated toric variety $X_{\Sigma'}$ is smooth, projective and it satisfies (2). Finally, since Σ' is a refinement of Σ, the toric morphism π given by [Cox et al. 2011, Theorem 3.3.4] is proper and restricts to the identity on the dense open orbit of $X_{\Sigma'}$. □

The previous lemma, together with the fact that intersection-theoretical properties are stable under birational transformations, allows one to compute the degree of a cycle of codimension 1 in a toric variety with respect to a family of toric divisors generated by global sections.

Proposition 5.2. *Let $D_{\Psi_1}, \dots, D_{\Psi_{n-1}}$ be toric divisors on X_Σ generated by global sections and Z an effective cycle on X_Σ of pure codimension 1 and prime components intersecting X_0, with defining polynomial f. Then*

$$\deg_{D_{\Psi_1},\dots,D_{\Psi_{n-1}}}(Z) = \mathrm{MV}_M(\Delta_{\Psi_1}, \dots, \Delta_{\Psi_{n-1}}, \mathrm{NP}(f)),$$

where MV_M denotes the mixed volume function associated to the measure vol_M (see Remark 1.2) and Δ_{Ψ_i} the polytope associated to the toric divisor D_{Ψ_i} for each $i = 1, \dots, n-1$.

Proof. Consider the smooth projective toric variety $X_{\Sigma'}$ and the proper toric morphism $\pi : X_{\Sigma'} \to X_\Sigma$ given by Lemma 5.1. Since the support function Ψ_f is a virtual support function on Σ', one can consider the corresponding toric divisor D_f on $X_{\Sigma'}$ and the associated distinguished rational section s_f of $\mathscr{O}(D_f)$. The product $\tilde{f} s_f$ is a rational section of $\mathscr{O}(D_f)$, with associated Weil divisor satisfying

$$\pi_* \operatorname{div}(\tilde{f} s_f) = \pi_*(\operatorname{div}(\tilde{f}) + \operatorname{div}(s_f)) = Z$$

by Theorem 4.3, (3-2), (5-1) and the definition of π. The projection formula in [Fulton 1998, Proposition 2.3(c)] and Bézout's theorem yield

$$\deg_{D_{\Psi_1},\dots,D_{\Psi_{n-1}}}(Z) = \deg_{\pi^* D_{\Psi_1},\dots,\pi^* D_{\Psi_{n-1}},D_f}(X_{\Sigma'}).$$

Since the function Ψ_f is concave, D_f is generated by global sections. Moreover, the virtual support function associated to the toric divisor $\pi^* D_{\Psi_i}$ on $X_{\Sigma'}$ agrees with Ψ_i for every $i = 1, \dots, n-1$. The combinatorial description in [Oda 1988, Proposition 2.10] of the degree of a toric variety with respect to toric divisor generated by global sections then concludes the proof. $\square$

Remark 5.3. By [Fulton 1993, formula on page 55], the irreducible components of $X_\Sigma \setminus X_0$ are the orbit closures $V(\tau)$, with τ ranging in the set of 1-dimensional cones of Σ. It follows that if Z is a prime divisor of X_Σ not intersecting X_0, it coincides with $V(\tau)$ for some $\tau \in \Sigma^{(1)}$. In such a case, the degree of Z with respect to a collection $D_{\Psi_1}, \dots, D_{\Psi_{n-1}}$ of toric divisors on X_Σ generated by global sections is given by

$$\deg_{D_{\Psi_1},\dots,D_{\Psi_{n-1}}}(V(\tau)) = \mathrm{MV}_{M(v_\tau)}(\Delta_{\Psi_1}^{v_\tau}, \dots, \Delta_{\Psi_{n-1}}^{v_\tau}),$$

where v_τ is the minimal nonzero integral vector of τ; see [Burgos Gil et al. 2014, formulae (3.4.1) and (3.4.4)].

Remark 5.4. The reduction to the case of a smooth projective toric variety employed in the proof of Proposition 5.2 equally works when computing the local height of the cycle Z with respect to a family of v-adic semipositive toric metrized divisors $\overline{D}_{0,v}, \dots, \overline{D}_{n-1,v}$. Indeed, let f be a defining polynomial for Z, and $X_{\Sigma'}$ and π be as in the statement of Lemma 5.1. For every family of rational sections $s_0, \dots, s_{n-1}$ of $\mathscr{O}(D_0), \dots, \mathscr{O}(D_{n-1})$ respectively for which the following local heights are well-defined, the local arithmetic projection formula in [Burgos Gil et al. 2014, Theorem 1.4.17(2)] asserts that

$$h_{\overline{D}_{0,v},\dots,\overline{D}_{n-1,v}}(Z; s_0, \dots, s_{n-1}) = h_{\pi^*\overline{D}_{0,v},\dots,\pi^*\overline{D}_{n-1,v}}(Z'; \pi^* s_0, \dots, \pi^* s_{n-1}),$$

where Z' is the cycle in $X_{\Sigma'}$ associated to the subscheme obtained as the closure of $V(f)$ and has hence f as a defining polynomial. Because of [Burgos Gil et al. 2014, Proposition 4.3.19], the pull-back of $\overline{D}_{i,v}$ via π is a v-adic semipositive toric metrized divisor on $X_{\Sigma'}$ whose metric function coincides with the one of $\overline{D}_{i,v}$ for every $i = 0, \dots, n-1$ and $v \in \mathfrak{M}$. It follows that any combinatorial formula for the local height of Z' in $X_{\Sigma'}$ with respect to $\pi^*\overline{D}_{0,v}, \dots, \pi^*\overline{D}_{n-1,v}$ only involving the defining polynomial of Z and the metric functions of the metrized divisors equally holds for the local height of Z in X_Σ with respect to $\overline{D}_{0,v}, \dots, \overline{D}_{n-1,v}$. Similarly, the reduction step can be adopted when dealing with the integrability and the global height of Z, because of Proposition 3.8.

5B. *Local heights.* Let f be a defining polynomial for the cycle Z and, as in the previous subsection, denote by Ψ_f the support function of its Newton polytope. Under the assumption that Ψ_f is a virtual support function on the fan of X_Σ, it determines a toric divisor D_f on X_Σ together with a distinguished rational section s_f of $\mathscr{O}(D_f)$, as in Section 3C.

Definition 5.5. In the notation above, and for a place $v \in \mathfrak{M}$, the *v-adic Ronkin metric* on D_f is the v-adic semipositive toric metric on $\mathscr{O}(D_f)_v^{\mathrm{an}}$ corresponding to the v-adic Ronkin function $\rho_{f,v}$ via Theorem 3.9.

The previous definition makes sense since, for every $v \in \mathfrak{M}$, the v-adic Ronkin function of f is concave on $N_\mathbb{R}$ and has bounded difference from Ψ_f because of Proposition 2.10. If not otherwise specified, $\overline{D}_{f,v}$

will denote the divisor D_f equipped with the v-adic Ronkin metric $\|\cdot\|_{f,v}$ defined above. By definition,

$$\log\|s_f\|_{f,v} = \rho_{f,v}\circ \mathrm{trop}_v \tag{5-2}$$

on $X_{0,v}^{\mathrm{an}}$. To lighten the notation, we will drop the subscript v whenever the choice of the place is clear from the context.

Proposition 5.6. *Let f and g be two nonzero Laurent polynomials and assume that Ψ_f and Ψ_g are virtual support functions on the fan of X_Σ. Then,*

$$\overline{D}_f + \overline{D}_g = \overline{D}_{f\cdot g}.$$

Proof. The equality $\mathrm{NP}(f\cdot g) = \mathrm{NP}(f)+\mathrm{NP}(g)$ implies that $\Psi_{f\cdot g} = \Psi_f + \Psi_g$. In particular, $\Psi_{f\cdot g}$ is a virtual support function on the fan Σ and then defines a toric divisor $D_{f\cdot g}$ on X_Σ which satisfies $D_{f\cdot g} = D_f + D_g$ because of [Fulton 1993, §3.4]. The statement follows now from [Burgos Gil et al. 2014, Proposition 4.3.14(1)] and Proposition 2.9. □

The key property of the Ronkin metric is given in the following proposition.

Proposition 5.7. *Let X_Σ be a smooth projective toric variety and Z an effective cycle on X_Σ of pure codimension 1 and prime components intersecting X_0. Let f be a defining polynomial for Z and $\tilde{f}$ the associated rational function on X_Σ. Assume moreover that Ψ_f is a virtual support function on the fan Σ. For a fixed place v of K, let $\overline{D}_0,\dots,\overline{D}_{n-1}$ be toric divisors on X_Σ, equipped with v-adic semipositive toric metrics. Then*

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z; s_0,\dots,s_{n-1}) = h_{\overline{D}_0,\dots,\overline{D}_{n-1},\overline{D}_f}(X_\Sigma; s_0,\dots,s_{n-1},\tilde{f}s_f) \tag{5-3}$$

for every choice of rational sections $s_0,\dots,s_{n-1}$ of $\mathscr{O}(D_0),\dots,\mathscr{O}(D_{n-1})$ respectively with $\mathrm{div}(s_0),\dots,$ $\mathrm{div}(s_{n-1})$, *Z intersecting properly.*

Proof. The product $\tilde{f}s_f$ is a rational section of $\mathscr{O}(D_f)$ on X_Σ with associated Weil divisor

$$\mathrm{div}(\tilde{f}s_f) = \mathrm{div}(\tilde{f}) + \mathrm{div}(s_f) = Z$$

by Theorem 4.3, (3-2) and (5-1). Hence, the sections $s_0,\dots,s_{n-1},\tilde{f}s_f$ meet X_Σ properly and the right-hand side term in (5-3) is well-defined.

Definition 3.6 states that

$$\begin{aligned} h_{\overline{D}_0,\dots,\overline{D}_{n-1}}&(Z; s_0,\dots,s_{n-1}) \\ &= h_{\overline{D}_0,\dots,\overline{D}_{n-1},\overline{D}_f}(X_\Sigma; s_0,\dots,s_{n-1},\tilde{f}s_f) + \int_{X_\Sigma^{\mathrm{an}}} \log\|\tilde{f}s_f\|_f\, c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}), \end{aligned}$$

and thus the proposition follows from the vanishing of the integral on the right-hand side. Indeed, thanks to Theorem 3.11, such an integral is supported on the analytification of the dense open orbit of X_Σ, where the rational function $\tilde{f}$ coincides with the regular function f. Together with the definition of the Ronkin metric in (5-2), this yields

$$\int_{X_\Sigma^{\mathrm{an}}} \log \|\tilde{f} s_f\|_f \; c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}) \\ = \int_{X_0^{\mathrm{an}}} \log|f| \; c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}) + \int_{X_0^{\mathrm{an}}} (\rho_{f,v}\circ \mathrm{trop}_v)\; c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}).$$

For every $i=0,\dots,n-1$, denote by ψ_i the metric function of $\overline{D}_i$. The tropicalization map being continuous, the change of variables formula and Theorem 3.11 imply on the one hand that

$$\int_{X_0^{\mathrm{an}}} (\rho_{f,v}\circ \mathrm{trop}_v)\; c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}) = \int_{N_\mathbb{R}} \rho_{f,v}\, d\,\mathrm{M}\mathcal{M}_M(\psi_0,\dots,\psi_{n-1}).$$

On the other hand, Theorem 3.11, together with the change of variables formula and Fubini's theorem, gives

$$\int_{X_0^{\mathrm{an}}} \log|f| \; c_1(\overline{D}_0)\wedge\cdots\wedge c_1(\overline{D}_{n-1}) = \int_{N_\mathbb{R}} \left(\int_{\mathcal{B}_{\mathbb{C}_v}} (\log|f|\circ\iota_v)\, d\,\mathrm{Haar}_{\mathcal{B}_{\mathbb{C}_v}} \right) d\,\mathrm{M}\mathcal{M}_M(\psi_0,\dots,\psi_{n-1}).$$

The definition of the maps ι_v and $\rho_{f,v}$ ensures that the inner integral coincides with the opposite of the v-adic Ronkin function of f, concluding the proof. □

5C. *Toric local heights.* Recall from Definition 3.10 that any toric divisor generated by global sections admits a distinguished v-adic semipositive toric metric, the canonical metric. This allows one to define a local height with respect to toric divisors that is independent of the choice of the sections. Such a notion was introduced in [Burgos Gil et al. 2014, §5.1] as a key step in the proof of the formula for the global height of a toric variety.

Definition 5.8. For a place v of K, let $\overline{D}_0,\dots,\overline{D}_d$ be toric divisors on X_Σ, endowed with v-adic semipositive toric metrics. Denote by $\overline{D}_0^{\mathrm{can}},\dots,\overline{D}_d^{\mathrm{can}}$ the same divisors equipped with their v-adic canonical metric. Let Y be an irreducible d-dimensional subvariety of X_Σ and $\varphi: Y'\to Y$ a birational morphism, with Y' projective. The *v-adic toric local height* of Y with respect to $\overline{D}_0,\dots,\overline{D}_d$ is defined as

$$h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_d}(Y) := h_{\varphi^*\overline{D}_0,\dots,\varphi^*\overline{D}_d}(Y';s_0,\dots,s_d) - h_{\varphi^*\overline{D}_0^{\mathrm{can}},\dots,\varphi^*\overline{D}_d^{\mathrm{can}}}(Y';s_0,\dots,s_d),$$

where s_i is a rational section of $\varphi^*\mathscr{O}(D_i)$ for every $i=0,\dots,d$ and $s_0,\dots,s_d$ meet Y' properly. The definition extends by linearity to any cycle of dimension d.

The toric local height of a cycle neither depends on the choice of the sections $s_0,\dots,s_d$, nor on the birational model Y' of Y because of [Burgos Gil et al. 2014, Theorem 1.4.17(2) and (3)]. Moreover, the definition is nonempty: Chow's lemma provides Y with a projective birational model, while the moving lemma ensures the existence of rational sections meeting Y' properly.

We prove here a formula for the toric local height of an effective cycle Z on X_Σ of pure codimension 1 and prime components intersecting X_0.

Theorem 5.9. *Let X_Σ be a proper toric variety, Z an effective cycle on X_Σ of pure codimension* 1 *and prime components intersecting X_0. For a place v of K, let $\overline{D}_0, \dots, \overline{D}_{n-1}$ be toric divisors on X_Σ, equipped with v-adic semipositive toric metrics. Then*

$$h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = \mathrm{MI}_M(\vartheta_0, \dots, \vartheta_{n-1}, \rho^\vee_{f,v}) + \deg_{D_0,\dots,D_{n-1}}(X_\Sigma) \cdot \rho_{f,v}(0),$$

where f is a defining polynomial for Z and ϑ_i is the roof function of $\overline{D}_i$ for $i = 0, \dots, n-1$.

Proof. Because of Remark 5.4, one can assume that X_Σ is a smooth projective toric variety on whose fan Ψ_f is a virtual support function. Thanks to the moving lemma, one can choose rational sections $s_0, \dots, s_{n-1}$ of $\mathscr{O}(D_0), \dots, \mathscr{O}(D_{n-1})$ respectively such that $\operatorname{div}(s_0), \dots, \operatorname{div}(s_{n-1})$, Z intersect properly. Proposition 5.7 implies then that

$$h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = h_{\overline{D}_0,\dots,\overline{D}_{n-1},\overline{D}_f}(X_\Sigma; s_0, \dots, s_{n-1}, \tilde{f} s_f) - h_{\overline{D}_0^{\mathrm{can}},\dots,\overline{D}_{n-1}^{\mathrm{can}},\overline{D}_f}(X_\Sigma; s_0, \dots, s_{n-1}, \tilde{f} s_f).$$

By adding and subtracting the quantity

$$h_{\overline{D}_0^{\mathrm{can}},\dots,\overline{D}_{n-1}^{\mathrm{can}},\overline{D}_f^{\mathrm{can}}}(X_\Sigma; s_0, \dots, s_{n-1}, \tilde{f} s_f)$$

on the right-hand side, one obtains that

$$h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_{n-1},\overline{D}_f}(X_\Sigma) - h^{\mathrm{tor}}_{\overline{D}_0^{\mathrm{can}},\dots,\overline{D}_{n-1}^{\mathrm{can}},\overline{D}_f}(X_\Sigma).$$

Denote by Ψ_i the virtual support function on Σ associated to the toric divisor D_i for every $i = 0, \dots, n-1$. Thanks to [Burgos Gil et al. 2014, Corollary 5.1.9], the previous equality yields

$$h^{\mathrm{tor}}_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = \mathrm{MI}_M(\vartheta_0, \dots, \vartheta_{n-1}, \rho^\vee_{f,v}) - \mathrm{MI}_M(\Psi_0^\vee, \dots, \Psi_{n-1}^\vee, \rho^\vee_{f,v}).$$

Since they admit by hypothesis a semipositive toric metric, the toric divisors $D_0, \dots, D_{n-1}$ are generated by global sections. For every $i = 0, \dots, n-1$, the function Ψ_i is hence concave and conic and so it is the support function of the polytope $\Delta_i := \operatorname{stab}(\Psi_i) \subseteq M_\mathbb{R}$. The statement of the theorem follows from a combination of Example 1.1, Corollary 1.9 and the combinatorial expression for the degree of a toric variety with respect to toric divisors generated by their global sections; see for example [Oda 1988, Proposition 2.10]. □

5D. *Global heights.* To state the result concerning the global case, recall that for a defining polynomial f for Z, the support function of its Newton polytope is denoted by Ψ_f; whenever it is linear on each cone of a complete fan Σ, it defines a toric divisor D_f on the toric variety X_Σ, together with a distinguished rational section s_f of $\mathscr{O}(D_f)$.

Definition 5.10. In the above assumptions, the *Ronkin metric* on D_f is the choice, for every place $v \in \mathfrak{M}$, of the v-adic Ronkin metric on D_f defined in Definition 5.5.

Unless otherwise stated, $\overline{D}_f$ will denote the toric divisor D_f equipped with its Ronkin metric. By definition, it is a semipositive toric metrized divisor.

Lemma 5.11. *The Ronkin metric on D_f is adelic.*

Proof. For a nonarchimedean place $v \in \mathfrak{M}$, the function $\rho_{f,v}$ coincides with the tropicalization of the Laurent polynomial f, as claimed in Remark 2.8. The fact that f has finitely many nonzero coefficients and the second axiom in Definition 3.1 imply that $\rho_{f,v} = \Psi_f$ for all but finitely many nonarchimedean places. The statement follows then from Lemma 3.3. □

The definition of such a toric metrized divisor and the study of the local height of Z in Section 5B allow one to answer the question of the integrability of Z and to give a formula for its global height, implying Theorem 1 in the Introduction.

Theorem 5.12. *Let X_Σ be a proper toric variety and Z an effective cycle on X_Σ of pure codimension 1 and prime components intersecting X_0. Let $\overline{D}_0, \dots, \overline{D}_{n-1}$ be toric divisors on X_Σ, equipped with adelic semipositive toric metrics. Then, Z is integrable with respect to $\overline{D}_0, \dots, \overline{D}_{n-1}$ and its global height is given by*

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = \sum_{v \in \mathfrak{M}} n_v \, \mathrm{MI}_M(\vartheta_{0,v}, \dots, \vartheta_{n-1,v}, \rho_{f,v}^\vee),$$

where f is a defining polynomial for Z and $\vartheta_{i,v}$ is the roof function of $\overline{D}_{i,v}$ for every $i = 0, \dots, n-1$ and $v \in \mathfrak{M}$.

Proof. Because of Remark 5.4, one can assume that X_Σ is a smooth projective toric variety on whose fan Ψ_f is a virtual support function. Let hence $s_0, \dots, s_{n-1}$ be rational sections of $\mathscr{O}(D_0), \dots, \mathscr{O}(D_{n-1})$ respectively such that $\mathrm{div}(s_0), \dots, \mathrm{div}(s_{n-1})$, Z intersect properly. Because of Proposition 5.7, the v-adic local height of Z with respect to the above choice of sections is given by

$$h_{\overline{D}_{0,v},\dots,\overline{D}_{n-1,v}}(Z; s_0, \dots, s_{n-1}) = h_{\overline{D}_{0,v},\dots,\overline{D}_{n-1,v},\overline{D}_{f,v}}(X_\Sigma; s_0, \dots, s_{n-1}, \tilde{f} s_f). \tag{5-4}$$

Because of Lemma 5.11, each member of the family $\overline{D}_0, \dots, \overline{D}_{n-1}, \overline{D}_f$ is an adelic semipositive toric metrized divisor on X_Σ. As a consequence of the first assertion in [Burgos Gil et al. 2014, Proposition 5.2.4], X_Σ is integrable with respect to such a choice of metrized divisors and hence [loc. cit., Proposition 1.5.8(1)] allows one to conclude that

$$h_{\overline{D}_{0,v},\dots,\overline{D}_{n-1,v},\overline{D}_{f,v}}(X_\Sigma; s_0, \dots, s_{n-1}, \tilde{f} s_f) = 0$$

for all but finitely many places $v \in \mathfrak{M}$. Comparing with (5-4), one deduces that Z is integrable with respect to $\overline{D}_0, \dots, \overline{D}_{n-1}$.

From the same equality, the global height of Z is seen to satisfy

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = h_{\overline{D}_0,\dots,\overline{D}_{n-1},\overline{D}_f}(X_\Sigma).$$

The formula for the global height of Z follows then from Theorem 3.13. □

Remark 5.13. As in Remark 5.3, if Z is an irreducible hypersurface on X_Σ not intersecting X_0 it coincides with $V(\tau)$ for a 1-dimensional cone τ of the fan Σ. In such a case, Z is integrable with respect to a family of adelic semipositive toric metrized divisors $\overline{D}_0, \dots, \overline{D}_{n-1}$ on X_Σ and its global height is given by

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(V(\tau)) = \sum_{v\in\mathfrak{M}} n_v \operatorname{MI}_{M(v_\tau)}(\vartheta_{0,v}|_{\Delta_0^{v_\tau}}, \dots, \vartheta_{n-1,v}|_{\Delta_{n-1}^{v_\tau}});$$

see [Burgos Gil et al. 2014, Propositions 5.1.11 and 5.2.4]. In the previous formula, Δ_i is the polytope associated to the divisor D_i and $\vartheta_{i,v}$ is the roof function of $\overline{D}_{i,v}$ for every $i = 0, \dots, n-1$ and $v \in \mathfrak{M}$, while v_τ is the minimal nonzero integral vector of τ.

Remark 5.14. Local and global heights of cycles are symmetric and multilinear with respect to sums of semipositive metrized divisors, provided that all terms are defined. The formulas obtained for 1-codimensional cycles in toric varieties are consistent with these properties, the sum of semipositive toric metrized divisors corresponding to the sup-convolution of the associated roof functions, as shown in [Burgos Gil et al. 2014, Proposition 4.3.14(1)].

6. Examples

For a fixed adelic field $(K, (|\cdot|_v, n_v)_{v\in\mathfrak{M}})$ satisfying the product formula and a proper toric variety X_Σ over K, of dimension n, with torus $\mathbb{T} = \operatorname{Spec} K[M]$ and dense open orbit X_0, we apply in this section the formula in Theorem 5.12 to four particular cases. In the first one, we focus on specific hypersurfaces of X_Σ, while in the following three we make relevant choices of the metrized divisors.

6A. ***Binomial hypersurfaces.*** For a primitive vector m in M one can consider the Laurent binomial $f = \chi^m - 1$; it is irreducible in $K[M]$, as can be verified by considering its Newton polytope. Hence the closure Z in X_Σ of the subvariety $V(f)$ of the torus $\operatorname{Spec} K[M]$ is an irreducible hypersurface of X_Σ with defining polynomial f.

Let $\overline{D}_0, \dots, \overline{D}_{n-1}$ be toric divisors on X_Σ, equipped with adelic semipositive toric metrics, with $\vartheta_{i,v}$ the roof function of $\overline{D}_{i,v}$ for every $i = 0, \dots, n-1$ and $v \in \mathfrak{M}$. By Example 2.12, $\rho_{f,v}$ coincides for every $v \in \mathfrak{M}$ with the support function of the segment $\overline{0m}$ in $M_\mathbb{R}$. The formula in Theorem 5.12 implies then that Z is integrable with respect to $\overline{D}_0, \dots, \overline{D}_{n-1}$ and that its global height is given by

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = \sum_{v\in\mathfrak{M}} n_v \operatorname{MI}_M(\vartheta_{0,v}, \dots, \vartheta_{n-1,v}, \iota_{\overline{0m}}),$$

because of Example 1.1. Considering, as at the end of Section 1C, the quotient lattice $P := M/\mathbb{Z}m$ and the associated projection $\pi : M \to P$, Proposition 1.12 allows one to deduce

$$h_{\overline{D}_0,\dots,\overline{D}_{n-1}}(Z) = \sum_{v\in\mathfrak{M}} n_v \operatorname{MI}_P(\pi_*\vartheta_{0,v}, \dots, \pi_*\vartheta_{n-1,v}), \tag{6-1}$$

with $\pi_*\vartheta_{i,v}$ denoting the direct image of $\vartheta_{i,v}$ by π for every $i = 0, \dots, n-1$ and $v \in \mathfrak{M}$; see (1-4).

Remark 6.1. Let Q be the dual lattice of $P = M/\mathbb{Z}m$. The projection $\pi : M \to P$ induces an injective dual map $Q \to N$, with image $m^\perp \cap N$. By identifying Q with such an image, which is a saturated sublattice of N, one can consider the restriction of the fan Σ to $Q_{\mathbb{R}}$; its corresponding toric variety X_{Σ_Q} is proper and has torus Spec $K[P]$. It also comes with a toric morphism $\varphi : X_{\Sigma_Q} \to X_\Sigma$, whose restriction to the dense open orbit coincides with the closed immersion of split tori Spec $K[P] \to$ Spec $K[M]$ given by the surjection $\pi : M \to P$; see [Burgos Gil et al. 2014, pp. 81–83]. Finally, the push-forward of the cycle X_{Σ_Q} by φ is the cycle Z associated to the hypersurface defined by $\chi^m - 1$. Indeed, the image of φ coincides by properness with the closure in X_Σ of the image of Spec $K[P] \to$ Spec $K[M]$, which is an irreducible $(n-1)$-dimensional subscheme of Spec $K[M]$ contained in $V(\chi^m - 1)$ as $\chi^{\pi(m)} - 1 = 0$.

Hence, equality (6-1) can also be obtained from Theorem 3.13, the arithmetic projection formula and the fact that $\pi_* \vartheta_{i,v}$ is the roof function of the pull-back of $\bar{D}_{i,v}$ via π for every $i = 0, \ldots, n-1$ and $v \in \mathfrak{M}$, because of [loc. cit., Propositions 4.3.19 and 2.3.8(3)].

6B. ***The canonical height.*** A toric divisor D on X_Σ generated by global sections admits by Definition 3.10 a distinguished semipositive toric metric at any place. The metrized divisor obtained by the choice of such a family of v-adic canonical metrics is denoted by $\bar{D}^{\mathrm{can}}$; it is an adelic semipositive toric metrized divisor.

For a cycle Z of dimension d in X_Σ, the *canonical global height* of Z with respect to a family $D_0, \ldots, D_d$ of toric divisors on X_Σ generated by global sections is defined to be its global height with respect to $\bar{D}_0^{\mathrm{can}}, \ldots, \bar{D}_d^{\mathrm{can}}$ and it is also denoted by $h^{\mathrm{can}}_{D_0,\ldots,D_d}(Z)$. The machinery developed in the previous sections allows one to express the canonical global height of an effective cycle on X_Σ of pure codimension 1 via convex geometry.

Proposition 6.2. *Let Z be an effective cycle on X_Σ of pure codimension* 1 *and prime components intersecting X_0 and $D_0, \ldots, D_{n-1}$ a family of toric divisors on X_Σ generated by global sections. The canonical global height of Z with respect to $D_0, \ldots, D_{n-1}$ is given by*

$$h^{\mathrm{can}}_{D_0,\ldots,D_{n-1}}(Z) = -\deg_{D_0,\ldots,D_{n-1}}(X_\Sigma) \cdot \sum_{v \in \mathfrak{M}} n_v \rho_{f,v}(0)$$

for any choice of a defining polynomial f for Z.

Proof. Denoting by Ψ_i, for any $i = 0, \ldots, n-1$, the function associated to D_i, the property of being globally generated implies that Ψ_i is the support function of the lattice polytope $\Delta_i := \mathrm{stab}(\Psi_i) \subseteq M_{\mathbb{R}}$. The roof function of $\bar{D}^{\mathrm{can}}_{i,v}$ is hence ι_{Δ_i} for every $i = 0, \ldots, n-1$ and for every $v \in \mathfrak{M}$, because of Example 1.1. It follows from Theorem 5.12 and Corollary 1.9 that

$$h^{\mathrm{can}}_{D_0,\ldots,D_{n-1}}(Z) = -\mathrm{MV}_M(\Delta_0, \ldots, \Delta_{n-1}) \cdot \sum_{v \in \mathfrak{M}} n_v \rho_{f,v}(0),$$

with f any defining polynomial for Z. To conclude, recall that the degree of X_Σ with respect to $D_0, \ldots, D_{n-1}$ is given by the mixed volume of the associated polytopes, as proved in [Oda 1988, Proposition 2.10]. □

The case of the base field $\mathbb{Q}$ with the adelic structure described in Example 3.4 is particularly interesting for arithmetic purposes. For a Laurent polynomial f in n variables and complex coefficients, one defines its (*logarithmic*) *Mahler measure* to be

$$m(f) := \frac{1}{(2\pi)^n} \int_{\theta_1,\dots,\theta_n\in[0,2\pi]} \log|f(e^{i\theta_1},\dots,e^{i\theta_n})|\,d\theta_1\cdots d\theta_n.$$

Such a quantity is notoriously difficult to compute and is sometimes related to special values of L-functions; see [Smyth 1981; Deninger 1997; Boyd 1998; Lalín 2008].

Maillot [2000, Proposition 7.2.1] expressed the canonical height of a hypersurface in a toric variety over $\mathbb{Q}$ in terms of the Mahler measure of the associated section. While its proof relies on the study of the arithmetic Chow ring of the ambient toric variety, we here deduce his result from Proposition 6.2.

Corollary 6.3 (Maillot). *In the hypotheses and notation of Proposition 6.2, assume moreover that the base adelic field is $\mathbb{Q}$ with its usual adelic structure. Let f be a defining polynomial for Z having as coefficients integers with greatest common divisor* 1. *Then,*

$$h^{\mathrm{can}}_{D_0,\dots,D_{n-1}}(Z) = \deg_{D_0,\dots,D_{n-1}}(X_\Sigma)\cdot m(f).$$

Proof. Let f be a defining polynomial for Z satisfying the assumptions. Because of Remark 2.8, for every nonarchimedean place v of $\mathbb{Q}$

$$\rho_{f,v}(0) = f^{\mathrm{trop}}(0) = 0.$$

At the unique archimedean place v of $\mathbb{Q}$ one has by definition that $\rho_{f,v}(0) = -m(f)$. The statement follows then directly from Proposition 6.2. □

6C. ***The ρ-height.*** The strategy adopted in the previous section to prove the main results of the paper suggests the introduction of a distinguished height function. Let Z be an effective cycle on X_Σ of pure codimension 1 and prime components intersecting X_0 and assume that the support function of the Newton polytope of a defining polynomial for Z is a virtual support function on the fan Σ. By Lemma 5.1, this is always the case up to a birational toric transformation. In this setting, the choice of a defining polynomial f for Z determines a toric divisor D_f on X_Σ and a distinguished toric metric on it, the Ronkin metric, as introduced in Definition 5.10. The so-obtained metrized divisor, which is denoted by $\overline{D}_f$, is an adelic semipositive toric metrized divisor by Lemma 5.11.

Definition 6.4. In the above hypotheses and notation, the ρ-*height* of Z, denoted by $h_\rho(Z)$, is defined as its global height with respect to $\overline{D}_f,\dots,\overline{D}_f$ for a choice of a defining polynomial f for Z.

As shown below, the ρ-height of Z is independent of the choice of the defining polynomial f. Even if it is not clear whether such a height has a significant geometrical interpretation or arithmetical application, it is straightforward to give a combinatorial formula for it.

Proposition 6.5. *In the above hypotheses and notation, the ρ-height of Z is given by*

$$h_\rho(Z) = (n+1)!\sum_{v\in\mathfrak{M}} n_v \int_{\mathrm{NP}(f)} \rho^\vee_{f,v}\,d\,\mathrm{vol}_M,$$

where f is a defining polynomial for Z and NP(f) *is its Newton polytope.*

Proof. The statement follows trivially from Theorem 5.12, Proposition 2.10(3) and the properties of mixed integrals. □

Remark 6.6. The equality in Proposition 6.5 shows that the ρ-height of Z does not depend on the choice of a defining polynomial for it. Indeed, if f' is another such polynomial, it must satisfy $f' = c \cdot \chi^m \cdot f$ for some nonzero monomial $c \cdot \chi^m \in K[M]$. For every $v \in \mathfrak{M}$, one has then that

$$\rho_{f',v} = -\log |c|_v + m + \rho_{f,v}$$

by Proposition 2.9 and Example 2.11. The stated independence follows hence from the relation

$$\rho^{\vee}_{f',v} = \tau_m \rho^{\vee}_{f,v} + \log |c|_v$$

obtained using [Burgos Gil et al. 2014, Proposition 2.3.3] and from the product formula on K.

It is significant to stress that the formula in Proposition 6.5, though compact, is difficult to evaluate because of the complexity of the archimedean Ronkin function.

6D. *The Fubini–Study height.* As a last example, consider the ambient toric variety X_Σ to be the n-dimensional projective space over K. Denote by D_∞ the toric divisor on $\mathbb{P}^n_K$ whose associated Weil divisor is the hyperplane at infinity; the corresponding sheaf is the universal line bundle $\mathscr{O}(1)$ on $\mathbb{P}^n_K$. If not otherwise specified, the notation $\overline{D}_\infty$ will refer to D_∞ equipped with the Fubini–Study metric at archimedean places, see [Burgos Gil et al. 2014, Example 1.1.2], and the canonical one at nonarchimedean places, in the sense of Definition 3.10. It turns out that $\overline{D}_\infty$ is an adelic semipositive toric metrized divisor. Thanks to Theorem 5.12, any effective cycle Z on $\mathbb{P}^n_K$ of pure codimension 1 is then integrable with respect to $\overline{D}_\infty, \dots, \overline{D}_\infty$ and the corresponding global height

$$h_{\mathrm{FS}}(Z) := h_{\overline{D}_\infty, \dots, \overline{D}_\infty}(Z)$$

is called the *Fubini–Study height* of Z.

Remark 6.7. The Fubini–Study height defined here coincides with the one introduced in [Faltings 1991] and studied in [Philippon 1995]. Examples of the computation of such height for projective hypersurfaces can be found in [Cassaigne and Maillot 2000].

Specializing Theorem 5.12, one can write the Fubini–Study height of a projective hypersurface in terms of convex geometry. To do so, denote by $\mathfrak{M}_\infty$ the collection of archimedean places of K, which is a finite set by Lemma 3.3. After fixing an isomorphism $M \simeq \mathbb{Z}^n$, consider the standard simplex

$$\Delta^n := \{(x_1, \dots, x_n) : x_1 + \dots + x_n \leq 1, x_i \geq 0 \text{ for all } i = 1, \dots, n\}$$

in $M_\mathbb{R} \simeq \mathbb{R}^n$ and, agreeing that $x_0 := 1 - \sum_{i=1}^n x_i$, set the function $\vartheta_{\mathrm{FS}} : \Delta^n \to \mathbb{R}$ to be

$$\vartheta_{\mathrm{FS}}(x) := -\frac{1}{2} \sum_{i=0}^{n} x_i \log x_i,$$

which is defined on the boundary of Δ^n by continuity.

Proposition 6.8. *Let Z be an effective cycle on $\mathbb{P}^n_K$ of pure codimension* 1 *and prime components intersecting X_0. The Fubini–Study height of Z is given by*

$$h_{\mathrm{FS}}(Z) = \sum_{v \in \mathfrak{M}_\infty} n_v \, \mathrm{MI}_M(\vartheta_{\mathrm{FS}}, \dots, \vartheta_{\mathrm{FS}}, \rho^\vee_{f,v}) - \sum_{v \in \mathfrak{M} \setminus \mathfrak{M}_\infty} n_v \rho_{f,v}(0),$$

where f is a defining polynomial for Z.

Proof. The roof functions of the metrized divisor $\overline{D}_\infty$ are given by the function ϑ_{FS} at archimedean places, as remarked in [Burgos Gil et al. 2014, Examples 2.4.3 and 4.3.9(2)] and by the indicator function of Δ^n at nonarchimedean places, by [loc. cit., Example 4.3.9(1)] and Example 1.1. The statement follows then from Theorem 5.12 and Corollary 1.9, together with the fact that $\mathrm{MV}_M(\Delta^n, \dots, \Delta^n) = 1$ because of the conventions introduced in Section 1B and Remark 1.2. □

The nonarchimedean contributions to the Fubini–Study height are easily computable, since for every Laurent polynomial f with set of coefficients $\Gamma(f)$,

$$-\rho_{f,v}(0) = \log \max_{c \in \Gamma(f)} |c|_v$$

if $v \in \mathfrak{M} \setminus \mathfrak{M}_\infty$. From this equality one easily obtains the following special case.

Corollary 6.9. *Assume the base adelic field to be $\mathbb{Q}$ with its usual adelic structure. The Fubini–Study height of an effective cycle Z on $\mathbb{P}^n_{\mathbb{Q}}$ of pure codimension* 1 *and prime components intersecting X_0 is given by*

$$h_{\mathrm{FS}}(Z) = \mathrm{MI}_M(\vartheta_{\mathrm{FS}}, \dots, \vartheta_{\mathrm{FS}}, \rho^\vee_{f,\infty}),$$

where f is a defining polynomial for Z whose coefficients are integers with greatest common divisor 1.

Because of the presence of an archimedean Ronkin function, the formula in Corollary 6.9 appears arduous to evaluate. It would anyway be interesting to use it to study arithmetical properties of projective hypersurfaces or recover similar results to the ones obtained in [Cassaigne and Maillot 2000].

Acknowledgments

The author wishes to thank his Ph.D. supervisors Martín Sombra and Alain Yger for their guidance and many fruitful discussions, as well as César Martínez for the interest shown in a preliminary version of this text. Also, he is grateful to the two anonymous referees for the careful reading and the valuable remarks, which highly contributed to the improvement of the quality of the text. The work was prepared at the Université de Bordeaux and at the Universitat de Barcelona as part of the author's Ph.D. project, and was partially supported by the MINECO research project MTM2015-65361-P and the CNRS project PICS 6381 "Géométrie diophantienne et calcul formel".

References

[Attouch and Wets 1989] H. Attouch and R. J.-B. Wets, "Epigraphical analysis", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**:suppl. (1989), 73–100. MR Zbl

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR Zbl

[Boyd 1998] D. W. Boyd, "Mahler's measure and special values of L-functions", *Experiment. Math.* **7**:1 (1998), 37–82. MR Zbl

[Burgos Gil et al. 2014] J. I. Burgos Gil, P. Philippon, and M. Sombra, *Arithmetic geometry of toric varieties: metrics, measures and heights*, Astérisque **360**, Société Mathématique de France, Paris, 2014. MR Zbl

[Burgos Gil et al. 2015] J. I. Burgos Gil, P. Philippon, J. Rivera-Letelier, and M. Sombra, "The distribution of Galois orbits of points of small height in toric varieties", preprint, 2015. arXiv

[Cassaigne and Maillot 2000] J. Cassaigne and V. Maillot, "Hauteur des hypersurfaces et fonctions zêta d'Igusa", *J. Number Theory* **83**:2 (2000), 226–255. MR Zbl

[Chambert-Loir 2006] A. Chambert-Loir, "Mesures et équidistribution sur les espaces de Berkovich", *J. Reine Angew. Math.* **595** (2006), 215–235. MR Zbl

[Chambert-Loir 2011] A. Chambert-Loir, "Heights and measures on analytic spaces: a survey of recent results, and some remarks", pp. 1–50 in *Motivic integration and its interactions with model theory and non-Archimedean geometry, II*, edited by R. Cluckers et al., London Math. Soc. Lecture Note Ser. **384**, Cambridge Univ. Press, 2011. MR Zbl

[Chambert-Loir and Ducros 2012] A. Chambert-Loir and A. Ducros, "Formes différentielles réelles et courants sur les espaces de Berkovich", preprint, 2012. arXiv

[Chambert-Loir and Thuillier 2009] A. Chambert-Loir and A. Thuillier, "Mesures de Mahler et équidistribution logarithmique", *Ann. Inst. Fourier* (*Grenoble*) **59**:3 (2009), 977–1014. MR Zbl

[Cox et al. 2011] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Mathematics **124**, American Mathematical Society, Providence, RI, 2011. MR Zbl

[Deninger 1997] C. Deninger, "Deligne periods of mixed motives, K-theory and the entropy of certain $\mathbb{Z}^n$-actions", *J. Amer. Math. Soc.* **10**:2 (1997), 259–281. MR Zbl

[Einsiedler et al. 2006] M. Einsiedler, M. Kapranov, and D. Lind, "Non-Archimedean amoebas and tropical varieties", *J. Reine Angew. Math.* **601** (2006), 139–157. MR Zbl

[Ewald 1996] G. Ewald, *Combinatorial convexity and algebraic geometry*, Graduate Texts in Mathematics **168**, Springer, 1996. MR Zbl

[Faltings 1991] G. Faltings, "Diophantine approximation on abelian varieties", *Ann. of Math.* (2) **133**:3 (1991), 549–576. MR Zbl

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton Univ. Press, 1993. MR Zbl

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **2**, Springer, 1998. MR Zbl

[Gelfand et al. 1994] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Birkhäuser, Boston, MA, 1994. MR Zbl

[Gubler 1997] W. Gubler, "Heights of subvarieties over M-fields", pp. 190–227 in *Arithmetic geometry* (Cortona, 1994), edited by F. Catanese, Sympos. Math. **37**, Cambridge Univ. Press, 1997. MR Zbl

[Gubler 1998] W. Gubler, "Local heights of subvarieties over non-Archimedean fields", *J. Reine Angew. Math.* **498** (1998), 61–113. MR Zbl

[Gubler 2003] W. Gubler, "Local and canonical heights of subvarieties", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **2**:4 (2003), 711–760. MR Zbl

[Gubler 2007] W. Gubler, "Tropical varieties for non-Archimedean analytic spaces", *Invent. Math.* **169**:2 (2007), 321–376. MR Zbl

[Gubler and Hertel 2017] W. Gubler and J. Hertel, "Local heights of toric varieties over non-archimedean fields", pp. 5–77 in *Actes de la conférence "Non-Archimedean Analytic Geometry: Theory and Practice"* (Papeete, 2015), Publ. Math. Besançon Algèbre Théorie Nr. **2017/1**, Presses Univ. Franche-Comté, Besançon, 2017. MR Zbl

[Gubler and Künnemann 2017] W. Gubler and K. Künnemann, "A tropical approach to nonarchimedean Arakelov geometry", *Algebra Number Theory* **11**:1 (2017), 77–180. MR Zbl

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

[Lalín 2008] M. N. Lalín, "Mahler measures and computations with regulators", *J. Number Theory* **128**:5 (2008), 1231–1271. MR Zbl

[Lekkerkerker 1969] C. G. Lekkerkerker, *Geometry of numbers*, Bibliotheca Mathematica **8**, Wolters-Noordhoff, Groningen, 1969. MR Zbl

[Maillot 2000] V. Maillot, *Géométrie d'Arakelov des variétés toriques et fibrés en droites intégrables*, Mém. Soc. Math. Fr. (N.S.) **80**, Société Mathématique de France, Paris, 2000. MR Zbl

[Martínez and Sombra 2018] C. Martínez and M. Sombra, "An arithmetic Bernštein–Kušnirenko inequality", *Math. Z.* (online publication September 2018).

[Oda 1988] T. Oda, "Convex bodies and algebraic geometry—toric varieties and applications, I", pp. 89–94 in *Algebraic geometry seminar* (Singapore, 1987), edited by M. Nagata and T. A. Peng, World Scientific, Singapore, 1988. MR Zbl

[Passare and Rullgård 2004] M. Passare and H. Rullgård, "Amoebas, Monge–Ampère measures, and triangulations of the Newton polytope", *Duke Math. J.* **121**:3 (2004), 481–507. MR Zbl

[Philippon 1995] P. Philippon, "Sur des hauteurs alternatives, III", *J. Math. Pures Appl.* (9) **74**:4 (1995), 345–365. MR Zbl

[Philippon and Sombra 2008a] P. Philippon and M. Sombra, "Hauteur normalisée des variétés toriques projectives", *J. Inst. Math. Jussieu* **7**:2 (2008), 327–373. MR Zbl

[Philippon and Sombra 2008b] P. Philippon and M. Sombra, "A refinement of the Bernštein–Kušnirenko estimate", *Adv. Math.* **218**:5 (2008), 1370–1418. MR Zbl

[Rauch and Taylor 1977] J. Rauch and B. A. Taylor, "The Dirichlet problem for the multidimensional Monge–Ampère equation", *Rocky Mountain J. Math.* **7**:2 (1977), 345–364. MR Zbl

[Rockafellar 1970] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series **28**, Princeton Univ. Press, 1970. MR Zbl

[Smyth 1981] C. J. Smyth, "On measures of polynomials in several variables", *Bull. Austral. Math. Soc.* **23**:1 (1981), 49–63. MR Zbl

[Zhang 1995] S. Zhang, "Small points and adelic metrics", *J. Algebraic Geom.* **4**:2 (1995), 281–300. MR Zbl

Communicated by Shou-Wu Zhang
Received 2017-11-14 Revised 2018-07-16 Accepted 2018-09-23

roberto.gualdi@math.u-bordeaux.fr *Institut de Mathématiques de Bordeaux, Université de Bordeaux, Talence, France*

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2445

Degree and the Brauer–Manin obstruction

Brendan Creutz and Bianca Viray
Appendix by Alexei N. Skorobogatov

Let $X \subset \mathbb{P}^n_k$ be a smooth projective variety of degree d over a number field k and suppose that X is a counterexample to the Hasse principle explained by the Brauer–Manin obstruction. We consider the question of whether the obstruction is given by the d-primary subgroup of the Brauer group, which would have both theoretic and algorithmic implications. We prove that this question has a positive answer in the case of torsors under abelian varieties, Kummer surfaces and (conditional on finiteness of Tate–Shafarevich groups) bielliptic surfaces. In the case of Kummer surfaces we show, more specifically, that the obstruction is already given by the 2-primary torsion, and indeed that this holds for higher-dimensional Kummer varieties as well. We construct a conic bundle over an elliptic curve that shows that, in general, the answer is no.

1. Introduction

Let X be a smooth projective and geometrically integral variety over a number field k. Manin observed that any adelic point $(P_v) \in X(\mathbb{A}_k)$ that is approximated by a k-rational point must satisfy relations imposed by elements of $\operatorname{Br} X$, the Brauer group of X [Manin 1971]. Indeed, for any element $\alpha \in \operatorname{Br} X := \mathrm{H}^2_{\mathrm{et}}(X, \mathbb{G}_m)$, the set of adelic points on X that are orthogonal to α, denoted $X(\mathbb{A}_k)^\alpha$, is a closed set containing the k-rational points of X. In particular,

$$X(\mathbb{A}_k)^{\mathrm{Br}} := \bigcap_{\alpha \in \operatorname{Br} X} X(\mathbb{A}_k)^\alpha = \varnothing \implies X(k) = \varnothing.$$

In this paper, we investigate whether it is necessary to consider the full Brauer group or whether one can determine *a priori* a proper subgroup $B \subset \operatorname{Br} X$ that *captures the Brauer–Manin obstruction* to the existence of rational points, in the sense that the following implication holds:

$$X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing \implies X(\mathbb{A}_k)^B := \bigcap_{\alpha \in B} X(\mathbb{A}_k)^\alpha = \varnothing.$$

This is of interest from both theoretical and practical perspectives. On the one hand, identifying the subgroups for which this holds may shed considerable light on the nature of the Brauer–Manin obstruction,

The second author was partially supported by NSA Young Investigator's Award #H98230-15-1-0054, NSF CAREER grant DMS-1553459, and a University of Canterbury Visiting Erskine Fellowship.
MSC2010: primary 14G05; secondary 11G35, 14F22.
Keywords: Brauer–Manin obstruction, degree, period, rational points.

and, on the other hand, knowledge of such subgroups can facilitate computation Brauer–Manin obstructions in practice.

We pose the following motivating question.

Question 1.1. Suppose that $X \hookrightarrow \mathbb{P}^n$ is embedded as a subvariety of degree d in projective space. Does the d-primary subgroup of $\operatorname{Br} X$ capture the Brauer–Manin obstruction to rational points on X?

More intrinsically, let us say that *degrees capture the Brauer–Manin obstruction on* X if the d-primary subgroup of $\operatorname{Br} X$ captures the Brauer–Manin obstruction to rational points on X for all integers d that are the degree of some k-rational globally generated ample line bundle on X. Since any such line bundle determines a degree d morphism to projective space and conversely, it is clear that the answer to Question 1.1 is affirmative when degrees capture the Brauer–Manin obstruction.

1A. ***Summary of results.*** In general, the answer to Question 1.1 can be no (see the discussion in Section 1B). However, there are many interesting classes of varieties for which the answer is yes. We prove that degrees capture the Brauer–Manin obstruction for torsors under abelian varieties, for Kummer surfaces, and, assuming finiteness of Tate–Shafarevich groups of elliptic curves, for bielliptic surfaces. We also deduce (from various results appearing in the literature) that degrees capture the Brauer–Manin obstruction for all geometrically rational minimal surfaces.

Assuming finiteness of Tate–Shafarevich groups, one can deduce the result for torsors under abelian varieties rather easily from a theorem of Manin (see Remark 4.4 and Proposition 4.9). In Section 4 we unconditionally prove the following much stronger result.

Theorem 1.2. *Let X be a k-torsor under an abelian variety, let $B \subset \operatorname{Br} X$ be any subgroup, and let d be any multiple of the period of X. In particular, d could be taken to be the degree of a k-rational globally generated ample line bundle. If $X(\mathbb{A})^B = \varnothing$, then $X(\mathbb{A})^{B[d^\infty]} = \varnothing$, where $B[d^\infty] \subset B$ is the d-primary subgroup of B.*

This not only shows that degrees capture the Brauer–Manin obstruction (apply the theorem with $B = \operatorname{Br} X$), but also that the Brauer classes with order relatively prime to d cannot provide *any* obstructions to the existence of rational points.

Remark 1.3. As one ranges over all torsors of period d under all abelian varieties over number fields, elements of arbitrarily large order in $(\operatorname{Br} V)[d^\infty]$ are required to produce the obstruction (see Proposition 4.10).

We use Theorem 1.2 to deduce that degrees capture the Brauer–Manin obstruction on certain quotients of torsors under abelian varieties. This method is formalized in Theorem 5.1 and applied in Section 5A to prove the following.

Theorem 1.4. *Let X be a bielliptic surface and assume that the Albanese torsor Alb^1_X is not a nontrivial divisible element in the Tate–Shafarevich group, $Ш(k, \mathrm{Alb}^0_X)$. Then degrees capture the Brauer–Manin obstruction to rational points on X.*

Remark 1.5. As shown by Skorobogatov [1999], the Brauer–Manin obstruction is insufficient to explain all failures of the Hasse principle on bielliptic surfaces. Therefore, if X is a bielliptic surface of degree d, then it is possible that there are adelic points orthogonal to the d-primary subgroup of $\operatorname{Br} X$ even though $X(k)=\varnothing$. However, Theorem 1.4 shows that in this case one also has that $X(\mathbb{A}_k)^{\mathrm{Br}} \neq \varnothing$.

Before stating our results on Kummer varieties we fix some notation. We say that X satisfies BM_d if the d-primary subgroup of $\operatorname{Br} X$ captures the Brauer–Manin obstruction, i.e., if the following implication holds:

$$X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing \implies X(\mathbb{A}_k)^{\mathrm{Br}_d} := \bigcap_{\alpha \in (\operatorname{Br} X)[d^\infty]} X(\mathbb{A}_k)^{\alpha} = \varnothing.$$

We say that X satisfies $\mathrm{BM}_d^{\perp}$ if there is no prime-to-d Brauer–Manin obstruction, i.e., if the following implication holds:

$$X(\mathbb{A}_k) \neq \varnothing \implies X(\mathbb{A}_k)^{\mathrm{Br}_{d^\perp}} := \bigcap_{\alpha \in (\operatorname{Br} X)[d^\perp]} X(\mathbb{A}_k)^{\alpha} \neq \varnothing,$$

where $(\operatorname{Br} X)[d^\perp]$ denotes the subgroup of elements of order prime to d. These properties are birational invariants of smooth projective varieties (see Lemma 2.5).

Remark 1.6. Note that BM_1 and $\mathrm{BM}_1^{\perp}$ are equivalent; they hold if and only if

$$X(\mathbb{A}_k) \neq \varnothing \iff X(\mathbb{A}_k)^{\mathrm{Br}} \neq \varnothing.$$

More generally, global reciprocity shows that the same is true of BM_d and $\mathrm{BM}_d^{\perp}$ whenever $(\operatorname{Br} X)[d^\infty]$ consists solely of constant algebras. In general, however, BM_d and $\mathrm{BM}_d^{\perp}$ are logically independent.

The following theorem is proved in Section 5B. We refer to that section for the definition of a Kummer variety.

Theorem 1.7. *Kummer varieties satisfy* BM_2.

Since the Picard lattice of any K3 surface is even, it is a straightforward consequence of this theorem that degrees capture the Brauer–Manin obstruction on Kummer surfaces.

Theorem 1.7 complements the recent result of Skorobogatov and Zarhin [2017, Theorem 3.3] that Kummer varieties satisfy $\mathrm{BM}_2^{\perp}$. As remarked above, this is logically independent from Theorem 1.7 except when $(\operatorname{Br} X)[2^\infty]$ consists solely of constant algebras. In [Skorobogatov and Zarhin 2017, Theorem 4.3] it is shown that $(\operatorname{Br} X)[2^\infty]$ consists solely of constant algebras for Kummer varieties attached to 2-coverings of products of hyperelliptic Jacobians with large Galois action on 2-torsion. This is the case of interest in [Harpaz and Skorobogatov 2016] where it is shown that (conditionally on finiteness of Tate–Shafarevich groups) some Kummer varieties of this kind satisfy the Hasse principle. Skorobogatov and Zarhin have shown that these Kummer varieties have no nontrivial 2-primary Brauer classes [Skorobogatov and Zarhin 2017, §4 and §5]. Given this, both Theorem 1.7 and [Skorobogatov and Zarhin 2017, Theorem 3.3] provide an explanation for the absence of any condition on the Brauer group in [Harpaz and Skorobogatov 2016, Theorems A and B].

After seeing a draft of this paper, Skorobogatov noted that it is possible to use our results (specifically Lemma 4.6) to extend the proof of [Skorobogatov and Zarhin 2017, Theorem 3.3] to obtain the following common generalization.

Theorem 1.8. *Let X be a Kummer variety over a number field k and suppose $B \subset \operatorname{Br} X$ is a subgroup. If $X(\mathbb{A})^B = \varnothing$, then $X(\mathbb{A})^{B[2^\infty]} = \varnothing$.*

This is the analog of Theorem 1.2 for Kummer varieties. The proof is given in the Appendix by Skorobogatov.

1B. *Discussion.* Question 1.1 trivially has a positive answer when $X(\mathbb{A}_k)^{\operatorname{Br}} \neq \varnothing$ and, in particular when $X(k) \neq \varnothing$. At the other extreme, the answer is also yes when either $X(\mathbb{A}_k) = \varnothing$ or $\operatorname{Alb}^1_X(\mathbb{A}_k)^{\operatorname{Br}} = \varnothing$ (see Corollary 4.5). In particular, the answer is yes for varieties satisfying the Hasse principle.

Theorem 1.2 shows that the answer is yes for curves of genus 1. When the first draft of the present paper was made available we were unaware of any example of a higher genus curve X and prime p for which one could show that X does not satisfy BM_p. Motivated by this, we undertook a deeper study of the Brauer–Manin obstruction on curves, joint with Voloch [Creutz et al. 2018]. There we produced a genus 3 curve over $\mathbb{Q}$ with a 0-cycle of degree 1 that is a counterexample to the Hasse principle explained by the 2-torsion subgroup of the Brauer group, but with no odd Brauer–Manin obstruction. This example shows that degrees do not capture the Brauer–Manin obstruction on higher genus curves, since the 0-cycle of degree 1 implies that every sufficiently large integer is the degree of a very ample line bundle.

One reason why degrees cannot capture the Brauer–Manin obstruction in general is that while the set $\{d \in \mathbb{N} : X \text{ satisfies } \mathrm{BM}_d\}$ is a birational invariant, the set of integers that arise as degrees of globally generated ample line bundles on X is not. Exploiting this one can construct examples (even among geometrically rational surfaces) for which degrees do not capture the Brauer–Manin obstruction (See Lemma 2.6 and Example 2.7). At least in the case of surfaces, this discrepancy can be dealt with by considering only minimal surfaces, i.e., surfaces which do not contain a Galois-invariant collection of pairwise disjoint (-1)-curves

Different and more serious issues are encountered in the case of nonrational surfaces of negative Kodaira dimension. In Section 6 we give an example of a minimal conic bundle over an elliptic curve for which degrees do not capture the Brauer–Manin obstruction. This is unsurprising (and somewhat less disappointing) given the known pathologies of the Brauer–Manin obstruction on quadric fibrations [Colliot-Thélène et al. 2016]. In Section 6 we note that degrees do capture the Brauer–Manin obstruction on minimal rational conic bundles and on Severi–Brauer bundles over elliptic curves with finite Tate–Shafarevich group (See Theorem 6.1).

For a minimal del Pezzo surface of degree d, degrees capture the Brauer–Manin obstruction as soon as BM_d holds. This is because the canonical class generates the Picard group, except when the surface is $\mathbb{P}^2$, a quadric or a rational conic bundle [Hassett 2009, Theorem 3.9], in which cases it follows from results mentioned above. Moreover, BM_d holds trivially when the degree, d, is not equal to 2 or 3. Indeed, when $d = 1$ there must be a rational point and when $d > 3$ the exponent of $\operatorname{Br} X / \operatorname{Br}_0 X$ divides d

[Corn 2007, Theorem 4.1].[1] Swinnerton-Dyer [1993, Corolarry 1] has shown that any cubic surface such that $\exp(\operatorname{Br} X/\operatorname{Br}_0 X) \nmid 3$ must satisfy the Hasse principle, implying that the answer is yes for $d = 3$. Whether the analogous property holds for del Pezzo surfaces of degree 2 was considered by Corn [2007, Question 4.5], but remained open until recent work of Nakahara [2017]. He showed that odd torsion Brauer classes on a del Pezzo surface of degree 2 cannot obstruct the Hasse principle.

Taken together, the results mentioned in the previous two paragraphs combine to yield the following.

Theorem 1.9. *Degrees capture the Brauer–Manin obstruction on geometrically rational minimal surfaces over number fields.*

Our results above give an affirmative answer (conditional on the finiteness of certain Tate–Shafarevich groups in the case of bielliptic surfaces) for two of the four classes of surfaces of Kodaira dimension 0, the other two being K3 and Enriques surfaces. Until quite recently, all known examples of Brauer–Manin obstructions to the existence of rational points on K3 or Enriques surfaces have come from the 2-torsion subgroup of the Brauer group, implying that degrees capture since the Néron–Severi lattice is even in both cases. In addition, even among diagonal quartic surfaces over $\mathbb{Q}$ where there were known to be nonconstant elements of odd order, Ieronymou and Skorobogatov [2015; 2017] showed that $\mathrm{BM}_2^{\perp}$ holds. Nakahara [2017] has extended this result to some diagonal quartics over general number fields, and, using results of this paper (specifically Lemma 4.6) has strengthened this to show that BM_2 holds and hence that degrees capture. These diagonal quartics are K3 surfaces that are geometrically Kummer, but not necessarily Kummer over their base field.

In contrast, some time after the first draft of the present paper was made available, Corn and Nakahara [2017] produced an example of a degree 2 K3 surface with a 3-torsion Brauer–Manin obstruction, showing that $\mathrm{BM}_2^{\perp}$ does not hold. This may suggest it is unlikely that degrees capture the Brauer–Manin obstruction for K3 surfaces, though they have not ruled out the possibility of an even transcendental Brauer–Manin obstruction.

Taken together, our results and the discussion above indicate that while the degrees of ample line bundles may in general be too crude to determine it, the set of integers d for which BM_d, or the stronger variant appearing in Theorems 1.2 and 1.8, holds are interesting birational invariants intimately related to the geometry and arithmetic of the variety.

1C. ***The analog for* 0*-cycles of degree* 1.** One further motivation for studying the question of whether degrees capture the Brauer–Manin obstruction to rational points is that the analog for 0-cycles of degree 1 is expected to hold Conjecture (E) on the sufficiency of the Brauer–Manin obstruction for 0-cycles implies that the $\deg(\mathcal{L})$-primary subgroup *does* capture the Brauer–Manin obstruction to 0-cycles of degree 1 for any ample globally generated line bundle $\mathcal{L}$. See Section 3 for more details.

[1] In fact, it is well-known that del Pezzo surfaces of degree at least 5 have trivial Brauer group and satisfy the Hasse principle (see, e.g., [Corn 2007, Thm. 4.1] and [Várilly-Alvarado 2013, Thm. 2.1]).

2. Preliminaries

2A. ***Notation.*** For an integer n, we define $\mathrm{Supp}(n)$ to be the set of prime numbers dividing n. For an abelian group G and integer d we use $G[d]$ to denote the subgroup of elements of order dividing d, $G[d^\infty]$ for the subgroup of elements of order dividing a power of d and $G[d^\perp]$ for the subgroup of elements order prime to d.

Throughout k denotes a number field, Ω_k denotes the set of places of k, and $\mathbb{A}_k$ denotes the adele ring of k. For any $v \in \Omega_k$, k_v denotes the completion of k at v. We use K to denote an arbitrary field of characteristic 0. We fix an algebraic closure $\overline{K}$ and write Γ_K for the absolute Galois group of K, with similar notation for k in place of K. For a commutative algebraic group G over k, we use $Ш(k, G)$ to denote the Tate–Shafarevich group of G, i.e., the group of torsors under G that are everywhere locally trivial.

Let X be a variety over K. We say that X is *nice* if it is smooth, projective, and geometrically integral. We write $\overline{X}$ for the base-change of X to $\overline{K}$. If X is defined over a number field k and $v \in \Omega_k$, we write X_v for the base-change of X to k_v.

When X is integral, we use $\boldsymbol{k}(X)$ to denote the function field of X. The Picard group of X, denoted $\mathrm{Pic}\, X$, is the group of isomorphism classes of K-rational line bundles on X. In the case that X is smooth, given a Weil divisor $D \in \mathrm{Div}\, X$ we denote the corresponding line bundle by $\mathcal{O}_X(D)$. The subgroup $\mathrm{Pic}^0 X \subset \mathrm{Pic}\, X$ consists of those elements that map to the identity component of the Picard scheme. The Néron–Severi group of X, denoted $\mathrm{NS}\, X$, is the quotient $\mathrm{Pic}\, X/\mathrm{Pic}^0 X$. The Brauer group of X, denoted $\mathrm{Br}\, X$, is $\mathrm{H}^2_{\mathrm{et}}(X, \mathbb{G}_m)$ and the algebraic Brauer group of X is $\mathrm{Br}_1 X := \ker(\mathrm{Br}\, X \to \mathrm{Br}\, \overline{X})$. The structure morphism yields a map $\mathrm{Br}\, K \to \mathrm{Br}\, X$, whose image is the subgroup of constant algebras denoted $\mathrm{Br}_0 X$.

2B. ***Polarized varieties, degrees and periods.*** In this paper a *nice polarized variety* over K is a pair $(X, \mathcal{L})$ consisting of a nice K-variety X and a globally generated ample line bundle $\mathcal{L} \in \mathrm{Pic}\, X$. We define the *degree* of a nice polarized variety, denoted by $\deg(X, \mathcal{L})$ or $\deg(\mathcal{L})$, to be $\dim(X)!$ times the leading coefficient of the Hilbert polynomial, $h(n) := \chi(\mathcal{L}^{\otimes n})$.

Lemma 2.1. *Suppose $(X, \mathcal{L})$ is a nice polarized variety of degree d over K. Then there is a K-rational 0-cycle of degree d on X.*

Proof. Since $\mathcal{L}$ is generated by global sections it determines a morphism $\phi_{\mathcal{L}} : X \to \mathbb{P}^N$, for some N. Since $\mathcal{L}$ is ample and globally generated, $\phi_{\mathcal{L}}$ is a finite morphism [Lazarsfeld 2004, Corollary 1.2.15, page 28]. The intersection of $\phi_{\mathcal{L}}(X) \subset \mathbb{P}^N$ with a general linear subvariety of codimension equal to $\dim(X)$ is a 0-cycle a on $\phi_{\mathcal{L}}(X)$ which pulls back to a 0-cycle of degree d on X. $\square$

For a nice variety X over K and $i \in \mathbb{Z}$ we write Alb^i_X for degree i component of the Albanese scheme parametrizing 0-cycles on X up to Albanese equivalence. Then Alb^0_X is an abelian variety and Alb^i_X is a K-torsor under Alb^0_X. The (Albanese) *period* of X, denoted $\mathrm{per}(X)$, is the order of Alb^1_X in the Weil–Châtelet group $\mathrm{H}^1(K, \mathrm{Alb}^0_X)$. Equivalently, the period is the smallest positive integer P such that

Alb_X^P has a K-rational point. Any k-rational 0-cycle of degree d determines a k-rational point on Alb_X^d. Thus Lemma 2.1 has the following corollary.

Corollary 2.2. *If $(X, \mathcal{L})$ is a nice polarized variety of degree d over K, then the period of X divides d.*

2C. *Basic properties of* BM_d *and* $\mathrm{BM}_d^\perp$. The definitions of BM_d and $\mathrm{BM}_d^\perp$ yield the following lemma, which we will use freely throughout the paper.

Lemma 2.3. *Let X be a nice variety over a number field k, let d and e be positive integers such that $d \mid e$, and set $d_0 := \prod_{p \in \mathrm{Supp}(d)} p$. Then*

(1) *X satisfies BM_d or $\mathrm{BM}_d^\perp$ if and only if X satisfies BM_{d_0} or $\mathrm{BM}_{d_0}^\perp$, respectively, and*

(2) *if X satisfies BM_d or $\mathrm{BM}_d^\perp$, then X satisfies BM_e or $\mathrm{BM}_e^\perp$, respectively.*

In particular, if d' is a positive integer with $\mathrm{Supp}(d) \subset \mathrm{Supp}(d')$ and X satisfies BM_d or $\mathrm{BM}_d^\perp$, then X satisfies $\mathrm{BM}_{d'}$ or $\mathrm{BM}_{d'}^\perp$, respectively.

Lemma 2.4. *Let $\pi : Y \to X$ be a morphism of nice varieties over a number field k and let d be a positive integer*:

(1) *If $Y(\mathbb{A}_k) \neq \varnothing$ and Y satisfies $\mathrm{BM}_d^\perp$, then X satisfies $\mathrm{BM}_d^\perp$.*

(2) *If $X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$ and X satisfies BM_d, then Y satisfies BM_d.*

Proof.

(1) Suppose that $X(\mathbb{A}_k) \neq \varnothing$, but $X(\mathbb{A}_k)^{\mathrm{Br}_{d^\perp}} = \varnothing$. Then for any $y \in Y(\mathbb{A}_k)$ there exists $\mathcal{A} \in (\mathrm{Br}\, X)[d^\infty]$ such that $0 \neq (\pi(y), \mathcal{A}) = (y, \pi^*\mathcal{A})$. Hence $Y(\mathbb{A}_k)^{\mathrm{Br}_{d^\perp}} = \varnothing$ and it follows that Y is not $\mathrm{BM}_d^\perp$.

(2) Suppose that X is BM_d and $X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$. Then given $y \in Y(\mathbb{A}_k)$, we may find an $\mathcal{A} \in (\mathrm{Br}\, X)[d^\infty]$ such that $0 \neq (\pi(y), \mathcal{A}) = (y, \pi^*\mathcal{A})$, which shows that $y \notin Y(\mathbb{A}_k)^{\mathrm{Br}_d}$. □

Lemma 2.5. *Let $\sigma : Y \dashrightarrow X$ be a birational map of nice varieties over a number field k and let d be a positive integer. Then*:

(1) *X satisfies BM_d if and only if Y satisfies BM_d.*

(2) *X satisfies $\mathrm{BM}_d^\perp$ if and only if Y satisfies $\mathrm{BM}_d^\perp$.*

(3) *The map σ induces an isomorphism $\sigma^* \colon \mathrm{Br}\, X \xrightarrow{\sim} \mathrm{Br}\, Y$ such that for any $B \subset \mathrm{Br}\, X$, $Y(\mathbb{A}_k)^{\sigma^*(B)} = \varnothing$ if and only if $X(\mathbb{A}_k)^B = \varnothing$.*

Proof. By the Lang–Nishimura Theorem [Lang 1954; Nishimura 1955], $X(\mathbb{A}_k) = \varnothing \Leftrightarrow Y(\mathbb{A}_k) = \varnothing$. With this in mind, the first two statements follow easily from the third. Any birational map between smooth projective varieties over a field of characteristic 0 can be factored into a sequence of blowups and blowdowns with smooth centers [Abramovich et al. 2002]. Hence, it suffices to prove (3) under the assumption that $\sigma : Y \to X$ is a birational morphism obtained by blowing up a smooth center $Z \subset X$. Then $\sigma^* \colon \mathrm{Br}\, X \to \mathrm{Br}\, Y$ is an isomorphism [Grothendieck 1968, Corollaire 7.3].

For any field L/k it is clear that $\sigma(Y(L))$ contains $(X \setminus Z)(L)$. Furthermore, since Z is smooth, the exceptional divisor E_Z is a projective bundle over Z, so $\sigma(E_Z(L)) = Z(L)$. Therefore $\sigma(Y(L)) = X(L)$. It follows that the map $\sigma : Y(\mathbb{A}_k) \to X(\mathbb{A}_k)$ is surjective, and so (3) follows from functoriality of the Brauer–Manin pairing. $\square$

Lemma 2.6. *Let d be a positive integer and let X be a nice k-variety of dimension at least 2 with the following properties*:

(1) $X(\mathbb{A}_k) \neq \varnothing$.

(2) $X(\mathbb{A}_k)^{\mathrm{Br}_d} = \varnothing$.

(3) $\mathrm{Br}\, X / \mathrm{Br}_0\, X$ *has exponent d.*

(4) *Degrees capture the Brauer–Manin obstruction on X.*

(5) *X has a closed point P with $\deg(P)$ relatively prime to d.*

Let $Y := \mathrm{Bl}_P\, X$. Then degrees do not capture the Brauer–Manin obstruction on Y.

Proof. Clearly X does not satisfy $\mathrm{BM}_{d'}$ for any d' prime to d. By birational invariance, the same is true of Y. It thus suffices to exhibit a globally generated ample line bundle on Y of degree prime to d. Let $\mathcal{L} \in \mathrm{Pic}(Y)$ be the pullback of an ample line bundle on X and let $\mathcal{E}$ denote the line bundle corresponding to the exceptional divisor. For some integer n, $\mathcal{L}^{\otimes n} \otimes \mathcal{E}^{-1}$ is ample and has degree prime to d. An appropriate multiple of this is very ample, hence globally generated, and has degree prime to d. $\square$

Example 2.7. An example of a variety satisfying the conditions of the lemma with $d = 2$ and $\deg(P) = 3$ is the del Pezzo surface of degree 2 given by the equation $w^2 = 34(x^4 + y^4 + z^4)$ [Corn 2007, Remark 4.3]. The blowup of X at a suitable degree 3 point gives a rational surface for which degrees do not capture the Brauer–Manin obstruction. In particular, the property “degrees capture the Brauer–Manin obstruction” is not a birational invariant of smooth projective varieties.

3. The analog for 0-cycles

Let X be a nice variety over a number field k. The group of 0-cycles on X is the free abelian group on the closed points of X. Two 0-cycles are directly rationally equivalent if their difference is the divisor of a function $f \in \boldsymbol{k}(C)^\times$, for some nice curve $C \subset X$. The Chow group of 0-cycles is denoted $\mathrm{CH}_0\, X$; it is the group of 0-cycles modulo rational equivalence, which is the equivalence relation generated by direct rational equivalence. For a place $v \in \Omega$ one defines the modified Chow group

$$\mathrm{CH}'_0\, X_v = \begin{cases} \mathrm{CH}_0\, X_v & \text{if } v \text{ is finite,} \\ \mathrm{coker}(N_{\bar{k}_v/k_v} : \mathrm{CH}_0\, \overline{X}_v \to \mathrm{CH}_0\, X_v) & \text{if } v \text{ is infinite.} \end{cases}$$

Since X is proper there is a well defined degree map $\mathrm{CH}_0\, X \to \mathbb{Z}$. We denote by $\mathrm{CH}_0^{(i)}\, X$ the preimage of $i \in \mathbb{Z}$. For an infinite place v, the degree of an element in $\mathrm{CH}'_0\, X_v$ is well defined modulo 2.

There is a Brauer–Manin pairing $\prod_v \mathrm{CH}_0 X_v \times \mathrm{Br}\, X \to \mathbb{Q}/\mathbb{Z}$ which, by global reciprocity, induces a complex,

$$\mathrm{CH}_0 X \to \prod_v \mathrm{CH}'_0 X_v \to \mathrm{Hom}(\mathrm{Br}\, X, \mathbb{Q}/\mathbb{Z}).$$

In particular, if there are no classes of degree 1 in the kernel of the map on the right, then there is no global 0-cycle of degree 1. In this case, we say that there is a Brauer–Manin obstruction to the existence of 0-cycles of degree 1.

Conjecture (E) states that the sequence

$$\varprojlim_n (\mathrm{CH}_0 X)/n \to \varprojlim_n \prod_v (\mathrm{CH}'_0 X_v)/n \to \mathrm{Hom}(\mathrm{Br}\, X, \mathbb{Q}/\mathbb{Z}) \tag{3-1}$$

is exact. This conjecture has its origins in work of Colliot-Thélène and Sansuc [1981], Kato and Saito [1986] and Colliot-Thélène [1995; 1999]. It has been stated in this form by van Hamel [2003] and Wittenberg [2012].

Conjecture (E) implies that there is a global 0-cycle of degree 1 if and only if there is no Brauer–Manin obstruction to such (see, e.g., [Wittenberg 2012, Remark 1.1(iii)]). For a nice polarized variety $(X, \mathcal{L})$, Conjecture (E) also implies that this obstruction is captured by the $\deg(\mathcal{L})$-primary part of the Brauer group. We note that this also follows immediately from [Colliot-Thélène 1999, Conjecture 2.2, with $i = \dim X$ and l a divisor of $\deg(\mathcal{L})$].

Proposition 3.1. *Let $(X, \mathcal{L})$ be a nice polarized variety of degree d over k and assume that Conjecture (E) holds for X. Then there exists a k-rational 0-cycle of degree 1 on X if and only if there exists $(z_v) \in \prod_v \mathrm{CH}_0^{(1)} X_v$ that is orthogonal to $(\mathrm{Br}\, X)[d^\infty]$.*

Proof. Set $\mathbb{Q}_d := \prod \mathbb{Q}_p$ and $\mathbb{Z}_d := \prod \mathbb{Z}_p$, where in both cases the product ranges over the primes dividing d. Exactness of (3-1) implies the exactness of its d-adic part,

$$\varprojlim_m (\mathrm{CH}_0 X)/d^m \to \varprojlim_m \prod_v (\mathrm{CH}'_0 X_v)/d^m \to \mathrm{Hom}((\mathrm{Br}\, X)[d^\infty], \mathbb{Q}_d/\mathbb{Z}_d).$$

If $(z_v) \in \prod_v \mathrm{CH}_0^{(1)} X_v$ is orthogonal to $(\mathrm{Br}\, X)[d^\infty]$, then by exactness we can find, for every $m \geq 1$, some $z_m \in \mathrm{CH}_0 X$ such that for every v, $z_m \equiv z_v \pmod{d^m \mathrm{CH}'_0 X_v}$. In particular, the degree of z_m is prime to d, so there is a global 0-cycle of degree prime to d on X. On the other hand, there is a 0-cycle of degree d on X by Lemma 2.1, so there must also be a 0-cycle of degree 1 on X. □

Unconditionally we can show that the $\deg(\mathcal{L})$-primary part of the Brauer group captures the Brauer–Manin obstruction to the existence of a global 0-cycle of degree 1 when $\mathrm{Br}\, X/\mathrm{Br}_0\, X$ has finite exponent.

Proposition 3.2. *Let $(X, \mathcal{L})$ be a nice polarized variety of degree d. Assume that $\mathrm{Br}\, X/\mathrm{Br}_0\, X$ has finite exponent. Then there is no Brauer–Manin obstruction to the existence of a 0-cycle of degree 1 if and only if there exists $(z_v) \in \prod_v \mathrm{CH}_0^{(1)} X_v$ that is orthogonal to $(\mathrm{Br}\, X)[d^\infty]$.*

Proof. If there is no Brauer–Manin obstruction to the existence of a 0-cycles of degree 1, then by definition there exists a $(z_v) \in \prod_v \mathrm{CH}_0^{(1)} X_v$ that is orthogonal to $\mathrm{Br}\, X$ and hence to $(\mathrm{Br}\, X)[d^\infty]$. Conversely, suppose there exists $(z_v) \in \prod_v \mathrm{CH}_0^{(1)} X_v$ that is orthogonal to $(\mathrm{Br}\, X)[d^\infty]$ and let $m \in \mathbb{Z}$ be the maximal divisor of the exponent of $\mathrm{Br}\, X/\mathrm{Br}_0\, X$ that is coprime to d. Since $\deg z_v = \deg z_w$ for all places v and w, the adelic 0-cycle (z_v) is orthogonal to $\mathrm{Br}_0\, X$. Hence, the bilinearity of the pairing and the definition of m imply that (mz_v) is orthogonal to $\mathrm{Br}\, X$. Furthermore, by Lemma 2.1, there is a k-rational 0-cycle of degree d on X; let $(z) \in \prod_v \mathrm{CH}_0\, X_v$ be its image. By global reciprocity every integral linear combination of (z) and (mz_v) is orthogonal to $\mathrm{Br}\, X$. Since m is relatively prime to d, some integral linear combination of (z) and (mz_v) has degree 1, as desired. □

4. Torsors under abelian varieties

In this section we prove the following theorem.

Theorem 4.1. *Let k be a number field, let Y be a smooth projective variety over k that is birational to a k-torsor V under an abelian variety, and let P be a positive integer that is divisible by the period of V. For any subgroup $B \subset \mathrm{Br}\, Y$ the following implication holds:*

$$Y(\mathbb{A}_k)^B = \varnothing \implies Y(\mathbb{A}_k)^{B[P^\infty]} = \varnothing.$$

Remark 4.2. Theorem 4.1 is strongest when $P = \mathrm{per}(V)$. However, determining $\mathrm{per}(V)$ is likely more difficult than determining if $Y(\mathbb{A}_k)^{B[P^\infty]} \neq \varnothing$, so the theorem will often be used for an integer P that is only known to bound the period.

Proof of Theorem 1.2. This follows immediately from Theorem 4.1 and Lemma 2.1. □

Corollary 4.3. *Every k-torsor V of period P under an abelian variety satisfies* BM_P *and* $\mathrm{BM}_P^\perp$.

Proof. We apply the theorem with $B = \mathrm{Br}\, X$ and $B = (\mathrm{Br}\, X)[P^\perp]$. □

Remark 4.4. When $Ш(k, A)$ is finite (i.e., conjecturally always) one can deduce that torsors of period P under A satisfy BM_P using the well-known result of Manin relating the Brauer–Manin and Cassels–Tate pairings. A generalization of Manin's result by Harari and Szamuely [2008] can be used to prove that torsors under semiabelian varieties satisfy BM_P, conditional on finiteness of Tate–Shafarevich groups (see Proposition 4.9).

Corollary 4.5. *Suppose X is a nice k-variety such that $\mathrm{Alb}^1_X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$. Then degrees capture the Brauer–Manin obstruction to rational points on X.*

Proof. Let P denote the Albanese period of X. By Theorem 4.1, there is a P-primary Brauer–Manin obstruction to the existence of rational points on Alb^1_X. This pulls back to give a P-primary Brauer–Manin obstruction to the existence of rational points on X. We conclude by noting that $P \mid \deg(\mathcal{L})$ for any globally generated ample line bundle $\mathcal{L} \in \mathrm{Pic}\, X$ by Lemma 2.1. □

For a nice variety X over K we define the subgroup $\mathrm{Br}_{1/2}\, X \subset \mathrm{Br}\, X$ as follows. Let S denote the image of the map $\mathrm{H}^1(K, \mathrm{Pic}^0\, \overline{X}) \to \mathrm{H}^1(K, \mathrm{Pic}\, \overline{X})$ induced by the inclusion $\mathrm{Pic}^0\, \overline{X} \subset \mathrm{Pic}\, \overline{X}$ and let $\mathrm{Br}_{1/2}\, X$ denote the preimage of S under the map $\mathrm{Br}_1\, X \to \mathrm{H}^1(K, \mathrm{Pic}\, \overline{X})$ that is given by the Hochschild–Serre spectral sequence.

Lemma 4.6. *Let V be a K-torsor under an abelian variety A over K. Let m and d be relatively prime integers with m relatively prime to the period of V. If $\operatorname{Br}_{1/2} V$ has finite index in $\operatorname{Br} V$, then there exists an étale morphism $\rho\colon V \to V$ such that the induced map,*

$$\rho^*\colon \frac{\operatorname{Br} V}{\operatorname{Br}_0 V} \to \frac{\operatorname{Br} V}{\operatorname{Br}_0 V},$$

annihilates the m-torsion subgroup and is the identity on the d-torsion subgroup. Moreover, one may choose ρ so that it agrees, geometrically, with $[m^r]\colon A \to A$ for some integer r.

Proof. Since V is a torsor under A, there is an isomorphism $\psi\colon \overline{V} \to \bar{A}$ of varieties over $\overline{K}$, such that the torsor structure of V is given by $a \cdot v = \psi^{-1}(a + \psi(v))$, for $a \in A(\overline{K})$ and $v \in V(\overline{K})$. Moreover, ψ induces group isomorphisms $\operatorname{Pic}^0 \bar{A} \simeq \operatorname{Pic}^0 \overline{V}$, $\operatorname{NS} \bar{A} \simeq \operatorname{NS} \overline{V}$, and $\operatorname{Br} \bar{A} \simeq \operatorname{Br} \overline{V}$.

Let P be the period of V and let n be a power of m such that $n \equiv 1 \bmod Pd$. Then $nV = V$ in $\mathrm{H}^1(K, A)$, so by [Skorobogatov 2001, Proposition 3.3.5] V can be made into an n-covering of itself. This means that there is an étale morphism $\pi\colon V \to V$ such that $\psi \circ \pi = [n] \circ \psi$ where $[n]$ denotes multiplication by n on A. We will show that an iterate of π has the desired properties.

Since $[n]$ induces multiplication by n on $\operatorname{Pic}^0 \bar{A}$, the morphism π induces multiplication by n on $\operatorname{Pic}^0 \overline{V}$. Indeed, $\pi^* = \psi^*[n]^*(\psi^{-1})^* = \psi^* n(\psi^{-1})^* = n\psi^*(\psi^{-1})^* = n$, where the penultimate equality follows since ψ^* is a homomorphism. Similarly, π induces multiplication by n^2 on $\operatorname{NS} \overline{V}$, since $[n]$ induces multiplication by n^2 on $\operatorname{NS} \bar{A}$. Thus we have a commutative diagram with exact rows

$$\begin{tikzcd}
\mathrm{H}^1(K, \operatorname{Pic}^0 \overline{V}) \arrow[r, "i"] \arrow[d, "n"] & \mathrm{H}^1(K, \operatorname{Pic} \overline{V}) \arrow[r, "j"] \arrow[d, "\pi^*"] & \mathrm{H}^1(k, \operatorname{NS} \overline{V}) \arrow[d, "n^2"] \\
\mathrm{H}^1(K, \operatorname{Pic}^0 \overline{V}) \arrow[r, "i"] & \mathrm{H}^1(K, \operatorname{Pic} \overline{V}) \arrow[r, "j"] & \mathrm{H}^1(k, \operatorname{NS} \overline{V})
\end{tikzcd}$$

We claim $(\pi^2)^*$ annihilates the n-torsion of $\mathrm{H}^1(K, \operatorname{Pic} \overline{V})$. Since $\operatorname{Br}_1 V/\operatorname{Br}_0 V$ embeds into $\mathrm{H}^1(K, \operatorname{Pic} \overline{V})$, this would imply that $(\pi^2)^*$ annihilates the n-torsion in $\operatorname{Br}_1 V/\operatorname{Br}_0 V$. Let us prove the claim. For any $x \in \mathrm{H}^1(K, \operatorname{Pic} \overline{V})[n]$, we have that $j(\pi^*(x)) = n^2 j(x) = j(n^2 x) = 0$, so $\pi^*(x) = i(y)$ for some $y \in \mathrm{H}^1(K, \operatorname{Pic}^0 \overline{V})$ such that $ny \in \ker(i)$. Then $(\pi^2)^*(x) = \pi^*(i(y)) = i(ny) = 0$, as desired.

Multiplication by n on A induces multiplication by n^2 on $\operatorname{Br} \bar{A}$ (see [Berkovič 1972, middle of page 182]). Thus π^* acts as multiplication by n^2 on $\operatorname{Br} \overline{V}$, and we have a commutative diagram with exact rows,

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{Br}_1 V/\operatorname{Br}_0 V \arrow[r, "i'"] \arrow[d, "\pi^*"] & \operatorname{Br} V/\operatorname{Br}_0 V \arrow[r, "j'"] \arrow[d, "\pi^*"] & \operatorname{Br} V/\operatorname{Br}_1 V \arrow[d, "n^2"] \\
0 \arrow[r] & \operatorname{Br}_1 V/\operatorname{Br}_0 V \arrow[r, "i'"] & \operatorname{Br} V/\operatorname{Br}_0 V \arrow[r, "j'"] & \operatorname{Br} V/\operatorname{Br}_1 V
\end{tikzcd}$$

The n-torsion in $\operatorname{Br}_1 V/\operatorname{Br}_0 V$ is killed by $(\pi^2)^*$ and i' is injective, so a similar diagram chase as above shows that $(\pi^3)^*$ kills the n-torsion, and hence the m-torsion, in $\operatorname{Br} V/\operatorname{Br}_0 V$.

It thus suffices to show that some power of $(\pi^3)^*$ is the identity map on the d-torsion subgroup of $\operatorname{Br} V/\operatorname{Br}_0 V$. By definition, the image of the composition $(\operatorname{Br}_{1/2} V \to \operatorname{Br}_1 V \to \mathrm{H}^1(K, \operatorname{Pic} \overline{V}))$ is contained in $\mathrm{H}^1(K, \operatorname{Pic}^0 \overline{V})$, so π^* acts as multiplication by n on $\operatorname{Br}_{1/2} V/\operatorname{Br}_0 V$. In particular, since $n \equiv 1 \bmod d$, π^* acts as the identity on the d-torsion subgroup of $\operatorname{Br}_{1/2} V/\operatorname{Br}_0 V$. Additionally, since the degree of π^3 is relatively prime to d, the induced map $(\pi^3)^* : (\operatorname{Br} V)[d^\infty] \to (\operatorname{Br} V)[d^\infty]$ is injective (see [Ieronymou et al. 2011, Proposition 1.1]). Together this shows that $(\pi^3)^*$ is injective on $(\operatorname{Br} V/\operatorname{Br}_{1/2} V)[d]$. Since $(\operatorname{Br} V/\operatorname{Br}_{1/2} V)[d]$ is finite, some power $\sigma = \pi^{3s}$ of π^3 acts as the identity on it. Thus, for every $\mathcal{A} \in \operatorname{Br} V$ such that $d\mathcal{A} \in \operatorname{Br}_0 V$ there is some $\mathcal{A}' \in \operatorname{Br}_{1/2} V$ such that

$$\sigma^*(\mathcal{A}) = \mathcal{A} + \mathcal{A}' \quad \text{and} \quad d\mathcal{A}' \in \operatorname{Br}_0 V.$$

Again σ^* is the identity on $(\operatorname{Br}_{1/2} V/\operatorname{Br}_0 V)[d]$, so by induction we get $(\sigma^d)^*(\mathcal{A}) \equiv \mathcal{A} + d\mathcal{A}' \equiv \mathcal{A} \pmod{\operatorname{Br}_0 V}$. Therefore $\rho = \sigma^d$ has the desired properties. □

Lemma 4.7. *Let k be a number field. If V is a k-torsor under an abelian variety, then* $\operatorname{Br}_{1/2} V$ *has finite index in* $\operatorname{Br} V$.

Proof. We have a filtration

$$\operatorname{Br}_{1/2} V \subset \operatorname{Br}_1 V \subset \operatorname{Br} V.$$

The second inclusion has finite index by [Skorobogatov and Zarhin 2008, Theorem 1.1]. Now we consider the first inclusion. By the definition of $\operatorname{Br}_{1/2} V$, the quotient $\operatorname{Br}_1 V/\operatorname{Br}_{1/2} V$ injects into the cokernel of the map $\mathrm{H}^1(k, \operatorname{Pic}^0 \overline{V}) \to \mathrm{H}^1(k, \operatorname{Pic} \overline{V})$. Then by the long exact sequence in cohomology associated to the short exact sequence

$$0 \to \operatorname{Pic}^0 \overline{V} \to \operatorname{Pic} \overline{V} \to \operatorname{NS} \overline{V} \to 0,$$

the cokernel in question injects into $\mathrm{H}^1(k, \operatorname{NS} \overline{V})$. The result now follows since $\operatorname{NS} \overline{V}$ is finitely generated and torsion-free. □

Lemma 4.8. *Let X be a smooth proper variety over a number field k and let $B \subset \operatorname{Br} X$ be a subgroup. If $X(\mathbb{A}_k)^B = \varnothing$, then there is a finite subgroup $\tilde{B} \subset B$ such that $X(\mathbb{A}_k)^{\tilde{B}} = \varnothing$.*

Proof. By hypothesis $X(\mathbb{A}_k)$ is compact. The lemma follows from the observation that

$$X(\mathbb{A}_k)^B = \bigcap_{\mathcal{A} \in B} X(\mathbb{A}_k)^{\mathcal{A}}.$$

is an intersection of closed subsets of $X(\mathbb{A}_k)$. If the intersection of these subsets is empty, then there is some finite collection of subsets whose intersection is empty. Since $\operatorname{Br} X$ is torsion and a finitely generated torsion abelian group is finite, this completes the proof. □

Proof of Theorem 4.1. In light of Lemma 2.5, we may assume that $Y = V$.

If $V(\mathbb{A}_k)^B = \varnothing$, then, by Lemma 4.8, there is a finite subgroup $B' \subset B$ with $V(\mathbb{A}_k)^{B'} = \varnothing$. Since $V(\mathbb{A}_k)^{B[P^\infty]} \subset V(\mathbb{A}_k)^{B'[P^\infty]}$, the desired implication holds if

$$V(\mathbb{A}_k)^{B'} = \varnothing \implies V(\mathbb{A}_k)^{B'[P^\infty]} = \varnothing,$$

for all finite subgroups $B' \subset B$. Thus, it suffices to prove the theorem when B is finite.

Let d be the exponent of $B[P^\infty]$ and let $m := \operatorname{exponent}(B)/d$ so that m and d are relatively prime. Since we are working over a number field, Lemma 4.7 allows us to apply Lemma 4.6. Let $\rho\colon V \to V$ be the morphism given by Lemma 4.6; then the functoriality of the Brauer–Manin pairing and global reciprocity give that

$$V(\mathbb{A}_k)^B \supset \rho(V(\mathbb{A}_k)^{\rho^*(B)}) = \rho(V(\mathbb{A}_k)^{\rho^*(B[P^\infty])}) = \rho(V(\mathbb{A}_k)^{B[P^\infty]}).$$

In particular, if $V(\mathbb{A}_k)^B$ is empty, then so must be $V(\mathbb{A}_k)^{B[P^\infty]}$. □

4A. ***A conditional extension to semiabelian varieties.***

Proposition 4.9. *Let k be a number field and let V be a k-torsor under a semiabelian variety G with abelian quotient A. Assume that $\text{Ш}(k, A)$ is finite. Then V satisfies* BM_d *for any integer d that is a multiple of the period of V.*

Proof. It is known that the obstruction coming from the group

$$\text{Б}(V) := \ker\left(\mathrm{Br}_1 V \to \bigoplus_v \mathrm{Br}_1 V_v / \mathrm{Br}_0 V_v\right)$$

is the only obstruction to rational points on V [Harari and Szamuely 2008, Theorem 1.1]. The extreme cases where G is an abelian variety or a torus are due to Manin [1971, Théorème 6] and Sansuc [1981, Corollaire 8.7], respectively. The proof of this fact is as follows. There is a homomorphism $\iota\colon \text{Ш}(k, G^*) \to \text{Б}(X)$ (where G^* denotes the 1-motive dual to G) and a bilinear pairing of torsion abelian groups $\langle\ ,\ \rangle_{\mathrm{CT}} : \text{Ш}(k, G) \times \text{Ш}(k, G^*) \to \mathbb{Q}/\mathbb{Z}$. This is related to the Brauer–Manin pairing via ι in the sense that for any $\beta \in \text{Ш}(k, G^*)$ and $(P_v) \in V(\mathbb{A}_k)$, one has $\langle [V], \beta\rangle_{\mathrm{CT}} = ((P_v), \iota(\beta))$. The assumption that $\text{Ш}(k, A)$ is finite implies that the pairing $\langle\ ,\ \rangle_{\mathrm{CT}}$ is nondegenerate. In particular, if $V(k) \subset V(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$, then there is some $\mathcal{A} \in \text{Б}$ such that $V(\mathbb{A}_k)^{\mathcal{A}} = \varnothing$. It follows from bilinearity that, if there is such an obstruction, it will already come from the $\operatorname{per}(V)^\infty$-torsion elements in $\text{Б}(X)$. □

4B. ***Unboundedness of the exponent.*** One might ask if we can restrict consideration to even smaller subgroups of $\mathrm{Br}\, V$, for instance if there is an integer d that can be determined *a priori* such that the d-torsion (rather than the d-primary torsion) captures the Brauer–Manin obstruction. The following proposition shows that this is not possible for torsors under abelian varieties, at least if Tate–Shafarevich groups of elliptic curves are finite.

Proposition 4.10. *Suppose that Tate–Shafarevich groups of elliptic curves over number fields are finite. For any integers P and n there exists, over a number field k, a torsor V under an elliptic curve with* $\operatorname{per}(V) = P$, $V(\mathbb{A}_k)^{(\mathrm{Br}\, V)[P^n]} \neq \varnothing$, *and* $V(\mathbb{A}_k)^{(\mathrm{Br}\, V)[P^{n+1}]} = \varnothing$.

The following lemma will be helpful in the proof.

Lemma 4.11. *For any integer N there exists an elliptic curve E over a number field k such that $\text{Ш}(k, E)$ has an element of order N.*

Remark 4.12. It would be very interesting to know if the curve E can be taken to be defined over $\mathbb{Q}$. To the best of our knowledge, this is unknown for P a sufficiently large prime and for P an arbitrary power of any single prime. It follows from the lemma that this does hold for abelian varieties over $\mathbb{Q}$, since restriction of scalars gives an element of order N in $\text{Ш}(\mathbb{Q}, \text{Res}_{k/\mathbb{Q}}(E))$.

Proof. Recall that the index of a variety X over a field is the gcd of the degrees of the closed points on X. By work of Clark and Sharif [2010] we can find a torsor V under an elliptic curve E/k of period N and index N^2. Moreover, the proof of [loc. cit.] shows that we can find such V with $V(k_v) = \varnothing$ for exactly two primes v of k, both of which are finite, prime to N and such that E has good reduction. Let L/k be any degree N extension of k which is totally ramified at both of these primes. By a result of Lang and Tate [1958] we have that $V(L_w) \neq \varnothing$ for w a place lying over either of these totally ramified primes, and hence $V_L \in \text{Ш}(L, E_L)$. On the other hand, the index of V_L can drop at most by a factor of $N = [L : k]$. Since V_L is locally trivial its period and index are equal [Cassels 1962, Theorem 1.3]. Therefore we have $N \leq \text{Index}(V_L) = \text{Period}(V_L) \leq \text{Period}(V) = N$, so V_L has order N in $\text{Ш}(L, E_L)$. □

Proof of Proposition 4.10. By the lemma above, we can find an elliptic curve E such that $\text{Ш}(k, E)$ contains an element of order P^{n+1}. Since $\text{Ш}(k, E)$ is finite, we can find $W \in \text{Ш}(k, E)$ such that $V := P^n W$ is not divisible by P^{n+1} in $\text{Ш}(k, E)$. Then a theorem of Manin [1971] shows that $V(\mathbb{A}_k)^{(\text{Br}\, V)[P^{n+1}]} = \varnothing$. On the other hand, $\pi : W \to V$ is a P^n-covering and, by descent theory, the adelic points in $\pi(W(\mathbb{A}_k))$ are orthogonal to $(\text{Br}_1\, V)[P^n] = (\text{Br}\, V)[P^n]$ (here we have used the fact that V is a curve and applied Tsen's theorem). □

5. Quotients of torsors under abelian varieties

Theorem 5.1. *Let k be a number field. Let Y be a smooth projective variety and let d be a positive integer such that for any subgroup $B \subset \text{Br}\, Y$ we have*

$$Y(\mathbb{A}_k)^B = \varnothing \Rightarrow Y(\mathbb{A}_k)^{B[d^\infty]} = \varnothing.$$

Let $\pi : Y \to X$ be a finite flat cover such that one of the following holds:

(1) *$d = 2$, π is a ramified double cover, and $(\text{Br}\, X/\text{Br}_0\, X)[2^\infty]$ is finite.*

(2) *π is a torsor under an abelian k-group of exponent dividing d.*

Then X satisfies BM_d.

Remark 5.2. In fact, a stronger result holds. There is a subgroup $B \subset (\text{Br}\, X)[d]$ such that if $X(\mathbb{A}_k)^B \neq \varnothing$, then X is also $\text{BM}_d^\perp$. If we are in case (1), then this subgroup B is finite, depends only on the Galois action on the geometric irreducible components of the branch locus of π, and is generically trivial.

The idea is that the subgroup B controls whether there is a twist of Y over X that is everywhere locally soluble. If there is such a twist, then we may apply Lemma 2.4(1) and Theorem 4.1 to conclude that X inherits $\text{BM}_d^\perp$ from the covering.

Remark 5.3. Theorem 4.1 implies that the first hypothesis of Theorem 5.1 is satisfied when Y is birational to a torsor V under an abelian variety and d is an integer that is divisible by every prime in $\operatorname{Supp}(\operatorname{per}(V))$.

For the proof, we need a slight generalization of a result of Skorobogatov and Swinnerton-Dyer [2005, Theorem 3 and Lemma 6].

Proposition 5.4. *Let $\pi : Y \to X$ be a ramified double cover over k and assume that $(\operatorname{Br} X/\operatorname{Br}_0 X)[2^\infty]$ is finite. Then*

$$X(\mathbb{A}_k)^{\operatorname{Br}_2} \neq \varnothing \iff \bigcup_{a\in k^\times/k^{\times 2}} \pi^a\big(Y^a(\mathbb{A}_k)^{(\pi^a)^*(\operatorname{Br} X[2^\infty])}\big) \neq \varnothing,$$

where $\pi^a : Y^a \to X$ denotes quadratic twist of $Y \to X$ by a.

Proof. The backwards direction follows from the functoriality of the Brauer group. Thus we consider the forwards direction. This proof follows ideas from [Skorobogatov and Swinnerton-Dyer 2005, §5]. We repeat the details here for the reader's convenience.

Let $f \in \boldsymbol{k}(X)^\times$ be such that $\boldsymbol{k}(Y) = \boldsymbol{k}(X)(\sqrt{f})$. We define a finite dimensional k-algebra

$$L := \bigoplus_{\substack{D\in X^{(1)}\\ v_D(f)\text{ odd}}} (\bar{k} \cap \boldsymbol{k}(D)).$$

Note that L is independent of the choice of f, since the class of f in $\boldsymbol{k}(X)^\times/\boldsymbol{k}(X)^{\times 2}$ is unique. Let $\alpha_1, \dots, \alpha_n \in (\operatorname{Br} X)[2^\infty]$ be representatives for the finitely many classes in $(\operatorname{Br} X/\operatorname{Br}_0 X)[2^\infty]$. Let S be a finite set of places such that for all $v \notin S$, for all $P_v \in X(k_v)$, and for all $1 \le i \le n$, $\alpha_i(P_v) = 0 \in \operatorname{Br} k_v$. (It is well-known that finding such a finite set S is possible, see, e.g., [Skorobogatov 2001, §5.2].) After possibly enlarging S we may assume that S contains all archimedean places and all places that are ramified in a subfield of L. We may also assume that $Y^a(k_v) \neq \varnothing$ for all $v \notin S$ and all $a \in k^\times/k^{\times 2}$ with $v(a)$ even [Skorobogatov 2001, Proposition 5.3.2].

Let $(P_v) \in X(\mathbb{A}_k)^{\operatorname{Br}_2}$. For $v \in S$, let $Q_v \in X(k_v)$ be such that $a_v := f(Q_v) \in k_v^\times$ and be sufficiently close to P_v so that $\alpha_i(P_v) = \alpha_i(Q_v)$ for all i. For $v \notin S$, set $a_v := 1$. Let $c \in k^\times$ be such that the class of c lies in the kernel of the natural map $k^\times/k^{\times 2} \to L^\times/L^{\times 2}$. Then by [Skorobogatov and Swinnerton-Dyer 2005, Theorem 3], the quaternion algebra $\mathcal{A} = (c, f)_2$ lies in $(\operatorname{Br} X)[2]$. Using the aforementioned properties of S and the definition of P_v and a_v, we then conclude

$$\sum_{v\in\Omega_k} \operatorname{inv}_v((c, a_v)) = \sum_{v\in S} \operatorname{inv}_v((c, a_v)) = \sum_{v\in S} \operatorname{inv}_v(\mathcal{A}(P_v)) = \sum_{v\in\Omega_k} \operatorname{inv}_v(\mathcal{A}(P_v)) = 0.$$

Hence, by [Skorobogatov and Swinnerton-Dyer 2005, Lemma 6(ii)], there exists an $a \in k^\times$ with $a/a_v \in k_v^{\times 2}$ such that $Y^a(\mathbb{A}_k) \neq \varnothing$. Since $a/a_v \in k_v^{\times 2}$ for all $v \in S$ we may further assume that there exists an

$(R_v) \in Y^a(\mathbb{A}_k)$ such that $\pi^a(R_v) = Q_v$ for all $v \in S$. We then have

$$\begin{aligned}\sum_v \operatorname{inv}_v(((\pi^a)^*(\alpha_i))(R_v)) &= \sum_v \operatorname{inv}_v(\alpha_i(\pi^a(R_v))) = \sum_{v\in S} \operatorname{inv}_v(\alpha_i(Q_v))\\ &= \sum_{v\in S} \operatorname{inv}_v(\alpha_i(P_v)) = \sum_v \operatorname{inv}_v(\alpha_i(P_v)) = 0,\end{aligned}$$

so $(R_v) \in Y^a(\mathbb{A}_k)^{(\pi^a)^*(\operatorname{Br} X[2^\infty])}$. $\square$

Proof of Theorem 5.1. If π is ramified, let $G = \mathbb{Z}/2$; otherwise let G be the finite k-group such that π is a torsor under G. For each $\tau \in \mathrm{H}^1(k, G)$, let $\pi^\tau \colon Y^\tau \to X$ denote the twisted cover, and define $B^\tau := (\pi^\tau)^*(\operatorname{Br} X)$.

Assume that $X(\mathbb{A}_k)^{\operatorname{Br}_d} \neq \varnothing$. By [Ieronymou et al. 2011, Proposition 1.1], $(\pi^\tau)^*((\operatorname{Br} X)[d^\infty]) = B^\tau[d^\infty]$. Therefore, Proposition 5.4 and descent theory show in cases (1) and (2) respectively, that there exists a $\tau \in \mathrm{H}^1(k, G)$ such that $Y^\tau(\mathbb{A}_k)^{B^\tau[d^\infty]} \neq \varnothing$. Applying the assumption with $B = B^\tau$ we conclude that $Y^\tau(\mathbb{A}_k)^{B^\tau}$ is nonempty. By the functoriality of the Brauer–Manin pairing, we have that $X(\mathbb{A}_k)^{\operatorname{Br}} \neq \varnothing$. $\square$

5A. *Bielliptic Surfaces.* We say that a nice variety X over K is a *bielliptic surface* if X is a minimal algebraic surface of Kodaira dimension 0 and irregularity 1.

5A1. *Geometry of bielliptic surfaces.* If X is a bielliptic surface over K then it is well-known that $\overline{X}$ is isomorphic to $(A \times B)/G$ where A and B are elliptic curves and G is a finite abelian group acting faithfully on A and B such that A/G is an elliptic curve and $B/G \simeq \mathbb{P}^1$ (see, e.g., [Beauville 1996, Chapters VI and VIII]). Furthermore, by the Bagnera–de Franchis classification [Beauville 1996, List VI.20], the pair of $\gamma := |G|$ and $n := \operatorname{exponent}(G)$ must be one of the following:

$$(\gamma, n) \in \{(2,2), (4,2), (4,4), (8,4), (3,3), (9,3), (6,6)\}.$$

In all cases n is the order of the canonical sheaf in $\operatorname{Pic} X$.

For A, B, and G as above, the universal property of the Albanese variety and [Beauville 1996, Example IX.7(1)] imply that the natural k-morphism $\Psi \colon X \to \mathrm{Alb}^1_X$ geometrically agrees with the projection map $(A \times B)/G \to A/G$.

Lemma 5.5. *Let A and B be elliptic curves over K and let G be a finite group acting faithfully on A and B such that A/G is an elliptic curve and $B/G \simeq \mathbb{P}^1$. Set $X := (A \times B)/G$ and let π denote the projection map $X \to B/G$. If $\mathcal{L} \in \operatorname{Pic} X$ is such that $\mathcal{L} \sim_{\mathrm{alg}} \pi^*(\mathcal{O}_{\mathbb{P}^1}(m))$ for some positive integer m and $\mathrm{H}^0(X, \mathcal{L}) \neq 0$, then $\mathcal{L} \simeq \pi^*(\mathcal{O}_{\mathbb{P}^1}(m))$.*

Proof. Since $\mathcal{L} \sim_{\mathrm{alg}} \pi^*(\mathcal{O}_{\mathbb{P}^1}(m))$ and $\operatorname{Pic}^0_X \simeq A/G \simeq \operatorname{Pic}^0_{A/G}$, there exists a line bundle $\mathcal{L}' \in \operatorname{Pic}^0(A/G)$ such that $\mathcal{L} \simeq \pi^*(\mathcal{O}_{\mathbb{P}^1}(m)) \otimes \pi_2^*(\mathcal{L}')$, where π_2 is the projection $X \to A/G$. By assumption $\mathrm{H}^0(X, \mathcal{L}) \neq 0$, so $\mathrm{H}^0(\mathbb{P}^1, \pi_*\mathcal{L}) \neq 0$, which, by the projection formula, implies that $\mathrm{H}^0(\mathbb{P}^1, \pi_*\pi_2^*\mathcal{L}')$, and hence $\mathrm{H}^0(A/G, \mathcal{L}')$, are nonzero. We complete the proof by observing that $\mathrm{H}^0(A/G, \mathcal{L}') \neq 0$ if and only if $\mathcal{L}' \simeq \mathcal{O}_{A/G}$. $\square$

Proposition 5.6. *Let X be a bielliptic surface over K. Then there exists a genus 0 curve C and a K-morphism $\Phi \colon X \to C$ that is geometrically isomorphic to the projection map $(A \times B)/G \to B/G$.*

Proof. By the geometric classification, there exist smooth elliptic curves A and B over $\overline{K}$ such that $\overline{X} \simeq (A \times B)/G$. By abuse of notation, we will also use A and B to refer to the algebraic equivalence class of a smooth fiber of the projection maps $\overline{X} \to B/G$ and $\overline{X} \to A/G$, respectively.

Since $\overline{X}$ has an ample divisor, the sum of the Galois conjugates of this divisor is a K-rational ample divisor D. Then by [Serrano 1990, Lemma 1.3 and Table 2], $[D] \equiv \alpha A + \beta B \in \operatorname{Num} \overline{X}$ for some positive $\alpha, \beta \in \frac{1}{n}\mathbb{Z}$. Since the natural map $X \to \operatorname{Alb}^1_X$ is K-rational, taking the fiber above a closed point yields a K-rational divisor F representing $mB \in \operatorname{Num} \overline{X}$ for some $m > 0$. By taking a suitable integral combination of D and F, we obtain a K-rational divisor D' that is algebraically equivalent to $m'A$ for some positive integer m'. Hence, $m'A \in (\operatorname{NS} \overline{X})^{\operatorname{Gal}(\overline{K}/K)}$.

Applying Lemma 5.5 to the Galois conjugates of $\mathcal{O}_{\overline{X}}(m'A)$ we see that $\mathcal{O}_{\overline{X}}(m'A) \in (\operatorname{Pic} \overline{X})^{\Gamma_K}$. By the Hochschild–Serre spectral sequence the cokernel of $\operatorname{Pic} X \to (\operatorname{Pic} \overline{X})^{\Gamma_K}$ injects into $\operatorname{Br} K$, which is torsion. Therefore some multiple of A is represented by a K-rational divisor D_0.

Since A is the class of the pull back of $\mathcal{O}_{B/G}(1)$ under the projection map $(A \times B)/G \to B/G)$, it follows that $\phi_{|D_0|}$ is geometrically isomorphic to the projection map $(A \times B)/G \to B/G$ composed with an r-tuple embedding. Thus, the image of $\phi_{|D_0|}$ is a K-rational genus 0 curve C and $\phi_{|D_0|}$ is the desired map. □

5A2. *Proof of Theorem 1.4.*

Lemma 5.7. *Let X be a bielliptic surface of period d over a number field k. Then X satisfies* BM_{nd} *where n denotes the order of the canonical sheaf in* $\operatorname{Pic} X$.

Proof. Since the canonical sheaf is n-torsion and is defined over k, there exists a μ_n-torsor $\pi : V \to X$ that is defined over k; by [Basile and Skorobogatov 2003, Proposition 1] V is a torsor under an abelian surface. Since π has degree n and X has period d, the period of V must satisfy $\operatorname{Supp}(\operatorname{per}(V)) \subset \operatorname{Supp}(nd)$. Indeed, for any $i \in \mathbb{Z}$, the degree n map $V \to X$ induces a map $\operatorname{Alb}^i_X \to \operatorname{Alb}^{ni}_V$. Since $\operatorname{Alb}^d_X(k) \neq \varnothing$ we must have $\operatorname{Alb}^{nd}_V(k) \neq \varnothing$. Then Theorem 4.1 and case (2) of Theorem 5.1 (see Remark 5.3) show that X satisfies BM_{nd}. □

We now state and prove our main result on bielliptic surfaces. Theorem 1.4 follows immediately.

Theorem 5.8. *Let $(X, \mathcal{L})$ be a nice polarized bielliptic surface over a number field k. If the canonical sheaf of X has order* 3 *or* 6, *assume that* Alb^1_X *is not a nontrivial divisible element of* $Ш(k, \operatorname{Alb}^0_X)$. *Then X satisfies* $\mathrm{BM}_{\deg(\mathcal{L})}$

Remark 5.9. If the canonical sheaf of X has order 3 or 6, then it follows from the Bagnera–de Franchis classification that the elliptic curve Alb^0_X has j-invariant 0.

Proof. Let n denote the order of the canonical sheaf of X. We will show that one of the four possibilities always occurs:

(1) X is not locally solvable.

(2) There is a k-rational point on X.

(3) There is a Brauer–Manin obstruction to rational points on Alb^1_X.

(4) Every globally generated ample line bundle $\mathcal{L}$ on X has $\operatorname{Supp}(n) \subset \operatorname{Supp}(\deg(\mathcal{L}))$.

In the first two cases BM_d holds trivially for every d. In the third case, the theorem follows from Corollary 4.5. In the fourth case we have $\operatorname{Supp}(\deg(\mathcal{L})) = \operatorname{Supp}(n\deg(\mathcal{L}))$ and, by Corollary 2.2, $\operatorname{Supp}(n\deg(\mathcal{L})) \supset \operatorname{Supp}(n \operatorname{per}(X))$, so we can apply Lemma 5.7 to conclude that $\mathrm{BM}_{\deg(\mathcal{L})}$ holds.

By the adjunction formula, the Néron–Severi lattice is even, so when $n \in \{2, 4\}$ we are in case (4) above. By the Bagnera–de Franchis classification we may therefore assume (γ, n) is one of $(6, 6)$, $(3, 3)$ or $(9, 3)$. Furthermore we can assume X is locally soluble. Then Alb^1_X represents an element in $Ш(k, \mathrm{Alb}^0_X)$. If this class is nontrivial, then Manin's theorem (see the proof of Proposition 4.9) shows that $\mathrm{Alb}^1_X(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$ and we are in case (3) above.

Since X is locally solvable Proposition 5.6 gives a k-morphism $\Phi\colon X \to \mathbb{P}^1$. Since $\mathbb{P}^1$ and Alb^1_X have k-points, we obtain k-rational fibers $A, B \in \operatorname{Div}(X)$ above k-points of $\mathbb{P}^1$ and Alb^1_X, respectively. All fibers of Ψ are smooth genus one curves geometrically isomorphic to B, while the general fiber of Φ is geometrically isomorphic to A, but there are 3 multiple fibers with (at least) one having multiplicity n [Serrano 1990, Table 2]. Let $A_0 \in \operatorname{Div}(\overline{X})$ denote the reduced component of a multiple fiber with multiplicity n. By [Serrano 1990, Theorem 1.4], the classes of A_0 and $n/\gamma B$ give a $\mathbb{Z}$-basis for $\operatorname{Num}(\overline{X})$ and the intersection pairing is given by $B^2 = A^2 = 0$ and $A \cdot B = \gamma$. The ample subset of $\operatorname{Pic} \overline{X}$ maps to the set of positive integral linear combinations of A_0 and $(n/\gamma)B$ in $\operatorname{Num}(\overline{X})$.

Let e_A and e_B be the smallest positive integers such that the classes of $e_A A_0$ and $e_B((\gamma/n)B)$ in $\operatorname{NS}(\overline{X})$ are represented by k-rational divisors on X. Since $nA_0 = A$ in $\operatorname{Pic} \overline{X}$ we have $e_A \mid n$. Moreover, $e_A = 1$ when $(\gamma, n) = (6, 6)$, because in this case Φ has a unique fiber of multiplicity 6, which must lie over a k-rational point. Similarly, $(\gamma/n)((n/\gamma)B) = B$, so $e_B \mid \gamma/n$. Therefore both e_A and e_B must divide 3. It follows that when $e_A e_B > 1$ every k-rational ample divisor on X has degree divisible by 6 and we are in case (2) above. Therefore we are reduced to considering the case $e_A = e_B = 1$, in which case we will complete the proof by showing that $X(k) \neq \varnothing$.

First we claim that when $e_A = 1$ the class of A_0 in $\operatorname{NS}(\overline{X})$ is represented by an effective k-rational divisor. In the case $n = 6$, we have seen that $A_0 \in \operatorname{Div}(X)$. In the case $n = 3$, let $F_1 = A_0$, F_2, F_3 denote the reduced components of the multiple fibers of Φ. The class of the canonical sheaf on X is represented by

$$2F_1 + 2F_2 + 2F_3 - 2A$$

[Serrano 1996, Theorem 4.1]. Since the canonical sheaf is not trivial, the F_i cannot all be linearly equivalent. The assumption that $e_A = 1$ implies that $F_1 = A_0$ is linearly equivalent to each of its Galois conjugates. The Galois action must permute the F_i, but it cannot act transitively, so some F_i must be fixed by Galois. This F_i is an effective k-rational divisor representing the class of A_0. Relabeling if necessary, we may therefore assume that A_0 is an effective k-rational divisor.

When $\gamma = n$, $(n/\gamma)B = B$ is an effective k-rational divisor intersecting A_0 transversally and $A_0 \cdot B = 1$, so $A_0 \cap B$ consists of a k-rational point. When $\gamma \neq n$ we have $(\gamma, n) = (9, 3)$. In this case, $D = A_0 + \frac{5}{3}B$ is very ample [Serrano 1990, Theorem 2.2]. By Bertini's theorem [Hartshorne 1977, Lemma V.1.2] the complete linear system $|D|$ contains a curve $C \in \operatorname{Div}(X)$ intersecting A_0 transversally. This gives a k-rational 0-cycle of degree 5 on A_0. On the other hand, the restriction of Ψ to A_0 gives a degree 3 étale map $A_0 \to \mathrm{Alb}^1_X$. As $\mathrm{Alb}^1_X(k) \neq \varnothing$, this gives a k-rational 0-cycle of degree 3 on A_0. Then A_0 is a genus one curve with a k-rational 0-cycle of degree 1, so it (and hence X) must have a k-point. $\square$

5B. *Kummer varieties.* Let A be an abelian variety over K. Any K-torsor T under $A[2]$ gives rise to a *2-covering* $\rho\colon V \to A$, where V is the quotient of $A \times_K T$ by the diagonal action of $A[2]$ and ρ is the projection onto the first factor. Then $T = \rho^{-1}(0_A)$ and V has the structure of a K-torsor under A. The class of T maps to the class of V under the map $\mathrm{H}^1(K, A[2]) \to \mathrm{H}^1(K, A)$ induced by the inclusion of group schemes $A[2] \hookrightarrow A$ and, in particular, the period of V divides 2.

Let $\sigma\colon \tilde{V} \to V$ be the blow up of V at $T \subset V$. The involution $[-1]\colon A \to A$ fixes $A[2]$ and induces involutions ι on V and $\tilde{\iota}$ on $\tilde{V}$ whose fixed point sets are T and the exceptional divisor, respectively. The quotient $\tilde{V}/\tilde{\iota}$ is geometrically isomorphic to the Kummer variety of A, so in particular is smooth. We call the quotient $\tilde{V}/\tilde{\iota}$ the *Kummer variety* associated to the 2-covering $\rho\colon V \to A$.

Theorem 5.10. *Let X be a Kummer variety over a number field k. Then X satisfies* BM_2.

Proof. By definition every Kummer variety admits a double cover by a smooth projective variety birational to a torsor of period dividing 2 under an abelian variety. Moreover $(\mathrm{Br}\, X/\mathrm{Br}_0\, X)[2^\infty]$ is finite by [Skorobogatov and Zarhin 2017, Corollary 2.8]. So the theorem follows from Theorem 4.1 and case (1) of Theorem 5.1 (see Remark 5.3). $\square$

6. Severi–Brauer bundles

A *Severi–Brauer bundle* is a nice variety X together with a dominant morphism $\pi : X \to Y$ to a nice variety Y such that the generic fiber is a smooth Severi–Brauer variety.

Theorem 6.1. *Let $\pi : X \to C$ be a Severi–Brauer bundle, with C a curve. Assume either*

(1) *$g(C) = 0$ and π is minimal of relative dimension* 1, *or*

(2) *$g(C) = 1$, π is a smooth morphism, and $\text{Ш}(k, \mathrm{Alb}^0_C)$ is finite.*

Then degrees capture the Brauer–Manin obstruction on X.

Remark 6.2. If $\pi : X \to C$ is a minimal conic bundle over a positive genus curve C with singular fibers, then degrees do not necessarily capture the Brauer–Manin obstruction. We construct a counterexample in Section 6A.

In the case that $g(C) = 0$, it is well-known that $\mathrm{Br}\, X/\mathrm{Br}_0\, X$ is 2-torsion, so X trivially satisfies BM_2. Then the result follows from the following lemma.

Lemma 6.3. *Let $\pi : X \to C$ be a minimal conic bundle over a curve C without a section. Then the degree of any globally generated ample line bundle in* $\mathrm{Pic}\, X$ *is divisible by $4i$, where i is the odd part of the index of C.*

Proof. Since π has no section, the class of the generic fiber in $\mathrm{Br}\, \boldsymbol{k}(C)$ has order 2. Then, since $\pi : X \to C$ is minimal, any line bundle in $\mathcal{L} \in \mathrm{Pic}\, X$ is algebraically equivalent (over $\bar{k}$) to $\mathcal{O}_X(2aS + b\,\mathrm{ind}(C)F)$, where $a, b, \in \mathbb{Z}$, S is a geometric section of π (which exists by Tsen's theorem) and F is the class of a fiber over a $\overline{K}$-point of C. Thus, we have

$$\deg(\mathcal{L}) = \deg(\mathcal{L}_{\overline{K}}) = 4a^2S^2 + 4ab\,\mathrm{ind}(C).$$

Further, since there is a double cover $C' \to C$ such that the conic bundle $\pi' \colon X \times_C C' \to C'$ has a section, S^2 is equal to the degree of the wedge power of a rank 2 vector bundle on $X \times_C C'$ [Beauville 1996, p. 30 and Proposition III.18]. Since any degree n point on C yields a degree n or $2n$ point on C', the index of C and C' can differ only by a factor of 2. Therefore, $S^2 \in \mathrm{ind}(C')\mathbb{Z} \subset i\mathbb{Z}$ and so $\deg(\mathcal{L}) = 4a^2S^2 + 4ab\,\mathrm{ind}(Y) \equiv 0 \pmod{4i}$. □

To prove part (2) of Theorem 6.1 we require the following result about Severi–Brauer bundles with no singular fibers.

Proposition 6.4. *Let $\pi : X \to Y$ be a Severi–Brauer bundle over a nice k-variety Y with π a smooth morphism. Let $\mathcal{A} \in \mathrm{Br}\, Y$ denote the class of the generic fiber of π and let d be the order of $\mathcal{A}$. Suppose that*

(1) $Ш(k, \mathrm{Alb}^0_Y)[d^\infty]$ *is finite,*

(2) $Y(k) \neq \varnothing$,

(3) *the canonical map* $Y(k) \to \mathrm{Alb}^1_Y(k)/d\,\mathrm{Alb}^0_Y(k)$ *is surjective, and*

(4) *for every prime v, the evaluation map* $\mathcal{A}_v \colon \mathrm{CH}_0^{(0)}(Y_v) \to \mathrm{Br}\, k_v$ *factors through* $\mathrm{Alb}^0_Y(k_v)$.

Then $X(\mathbb{A})^{\mathrm{Br}_d} \neq \varnothing \Rightarrow X(k) \neq \varnothing$. In particular, X satisfies BM_d.

Remark 6.5. Hypotheses (3) and (4) of the proposition are satisfied if

- Y is an abelian variety;
- Y is a curve such that $\#Y(k) = \#\,\mathrm{Alb}^0_Y(k)$; or
- Y is a curve such that $Y(k) \neq \varnothing$ and $\mathrm{Alb}^0_Y(k)$ is finite of order prime to d.

This is a slight generalization of a result in [Colliot-Thélène et al. 2016, Proposition 6.5] which proves that the Brauer–Manin obstruction is the only one on X under the assumption that Y is an elliptic curve with $Ш(k, Y)$ finite.

Proof of Theorem 6.1(2). If $C(k) = \varnothing$, then the assumption on $Ш(k, \mathrm{Alb}^0_C)$ implies by Manin's theorem that $C(\mathbb{A}_k)^{\mathrm{Br}} = \varnothing$. Since $C \simeq \mathrm{Alb}^1_X$, we may apply Corollary 4.5 to conclude that degrees capture the Brauer–Manin obstruction on X.

Assume that $C(k) \neq \varnothing$ and let d be the order of the generic fiber of π in $\mathrm{Br}\, C$. Then Proposition 6.4 and Remark 6.5 together show that X satisfies BM_d so it suffices to show that d divides the degree of any globally generated ample line bundle.

Restriction to the generic fiber yields an exact sequence

$$0 \to \mathbb{Z}F \to \mathrm{NS}\, X \to \mathrm{NS}\, X_\eta \to 0,$$

where F denotes the class of a fiber of π. (Here we are using the assumption that π is smooth, so there are no reducible fibers.) Since X_η is a Severi–Brauer variety whose class in the Brauer group has order d, $\mathrm{NS}\, X_\eta$ is free of rank 1 and generated by dH, where H is a generator of $\mathrm{NS}\,\overline{X_\eta} \simeq \mathrm{NS}\,\mathbb{P}^r \simeq \mathbb{Z}$. Thus any divisor class is $\mathrm{NS}\, X$ can be represented as $adH' + bF$, where H' is a Zariski closure of H. Since

$$\begin{aligned}(adH'+bF)^{\dim X} &= (adH')^{\dim X} + (\dim X)(ad)^{\dim X-1} b(H')^{\dim X-1}\cdot F\\ &= (ad)^{\dim X-1}(ad\cdot H'^{\dim X} + (\dim X) b(H'^{\dim X-1}\cdot F)),\end{aligned}$$

and $\dim X > 1$, this shows that the degree of any line bundle is divisible by d. □

Proof of Proposition 6.4. To ease notation set $A = \mathrm{Alb}^0_Y$. Choose $P_0 \in Y(k)$ and use this to define a k-morphism $\iota\colon Y \to A$ sending P to the class of the 0-cycle $P - P_0$. Suppose that $X(\mathbb{A}_k)^{\mathrm{Br}_d} \neq \varnothing$. Fix a point $(Q_v) \in X(\mathbb{A}_k)^{\mathrm{Br}_d}$ and let $(P_v) := (\pi(Q_v))$. Then $(\iota(P_v)) \in A(\mathbb{A}_k)^{\mathrm{Br}_d}$ by functoriality. Since $\iota(P_v)$ is orthogonal to $(\mathrm{Br}\, A)[d^\infty]$ it follows from descent theory (e.g., [Skorobogatov 2001, Theorem 6.1.2]) that, for every n, $\iota(P_v)$ lifts to an adelic point on some d^n-covering of A. (For the definition of an N-covering see [Skorobogatov 2001, Definition 3.3.1].) Because $Ш(k, A)[d^\infty]$ is finite, there is some n such that every d^n-covering of A which lifts to a locally soluble d^{n+1}-covering of A is of the form $x \mapsto d^n x + Q$ for some $Q \in A(k)$. In particular, there is some $Q \in A(k)$ such that $(\iota(P_v) - Q) \in dA(\mathbb{A}_k)$. By assumption (3) there is some $P \in Y(k)$ such that $(\iota(P_v) - \iota(P)) \in dA(\mathbb{A}_k)$. Now by assumption (4) and the fact that $d\mathcal{A} = 0$ it follows that, for every prime v, $\mathcal{A}(P)\otimes k_v = \mathcal{A}(P_v)$. Note that $\mathcal{A}(P_v) = 0 \in \mathrm{Br}\, k_v$ since the fiber X_{P_v} has a k_v-point. Thus, $\mathcal{A}(P)\otimes k_v = 0$ for every v. We conclude that the Severi–Brauer variety X_P must be everywhere locally soluble and thus, by the Albert–Brauer–Hasse–Noether theorem, that $X_P(k) \neq \varnothing$. □

6A. *A counterexample.*

Theorem 6.6. *There exists a conic bundle surface $\pi : X \to E$ over an elliptic curve, such that X has an ample globally generated line bundle of degree* 12 *and X does not satisfy* BM_6. *In particular, degrees do not capture the Brauer–Manin obstruction.*

Proof. Let $E/\mathbb{Q}$ be an elliptic curve with a $\mathbb{Q}$-rational cyclic subgroup Z of order 5 such that

$$\mathrm{Gal}(\mathbb{Q}(Z)/\mathbb{Q}) = \mathbb{Z}/4,$$

with $E(\mathbb{Q}) = \{O\}$ and with $Ш(\mathbb{Q}, E)$ finite, e.g., the curve with LMFDB label 11.a1 [LMFDB Collaboration 2013, Elliptic Curve 11.a2]. Then $Z = O \cup P$, where P is a degree 4 closed point. Let p_1 and p_2 be two primes that are congruent to 1 modulo 8 and that split completely in $\boldsymbol{k}(P)$. Finally choose a prime $q \neq p_i$ that is 1 modulo 8 and that is a nonsquare modulo p_1 and modulo p_2.

We will use this chosen data to construct a conic bundle X over E with the required properties. Let $\mathcal{E}$ be the rank 3 vector bundle $\mathcal{O}_E \oplus \mathcal{O}_E \oplus \mathcal{O}_E(2O)$. Since $4O \sim P$, there exists a section $s_2 \in \mathrm{H}^0(E, \mathcal{O}_E(2O)^{\otimes 2})$ such that $\mathrm{div}(s_2) = P$ and such that $s_2(O) = -p_1 p_2$. Let $s_0 := 1$ and $s_1 := -q$ thought of as sections of $\mathrm{H}^0(E, \mathcal{O}_E^{\otimes 2})$. Then we define $\pi : X \to E$ to be the conic bundle given by the vanishing of the section $s := s_0 + s_1 + s_2 \in \mathrm{H}^0(E, \mathrm{Sym}^2\, \mathcal{E})$ inside $\mathbb{P}(\mathcal{E})$; this conic bundle is smooth by [Poonen 2009, Lemma 3.1].

Since s is invertible on $E - P$, the morphism π is smooth away from P. Because P is a single closed point, it follows that $\pi^*\colon \operatorname{Br} E \to \operatorname{Br} X$ is surjective (see [Colliot-Thélène et al. 2016, Proposition 2.2(i)]). Hence, $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}_6} \neq \varnothing$ if and only if $\pi(X(\mathbb{A}_{\mathbb{Q}})) \cap E(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}_6} \neq \varnothing$.

Since the p_i split completely in $\boldsymbol{k}(P)$, the connected components of $P \otimes k_{p_i}$ are points in $E(\mathbb{Q}_{p_i})$. Consider an adelic point $(P_v) \in E(\mathbb{A}_{\mathbb{Q}})$ with $P_v = O$ for all $v \neq p_1, p_2$ and P_v a connected component of $P \otimes k_v$ for $v = p_1, p_2$. Then (P_v) is 5-torsion and, since 5 is relatively prime to 6, the point (P_v) is 6-divisible in $E(\mathbb{A}_{\mathbb{Q}})$. Since $Ш(\mathbb{Q}, E)$ is finite, descent theory implies that $(P_v) \in E(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}_6}$ (see the proof of Proposition 6.4).

Since p_i splits completely in $\boldsymbol{k}(P)$, each connected component of $X_P \otimes \mathbb{Q}_{p_i}$ is a pair of lines intersecting at a unique point, which must be defined over $\mathbb{Q}_{p_i}$. In particular, $P_{p_i} \in \pi(X(\mathbb{Q}_{p_i}))$ for $i = 1, 2$. Further, by construction, the fiber above O is given by

$$X_O : w_0^2 - q w_1^2 = p_1 p_2 w_2^2.$$

By the choice of p_1, p_2 and q, $X_O(\mathbb{Q}_v) \neq \varnothing$ exactly when $v \neq p_1, p_2$. Hence, $(P_v) \in \pi(X(\mathbb{A}_{\mathbb{Q}}))$ and so $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}_6} \neq \varnothing$.

To prove that X does not satisfy BM_6, it remains to show that $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \varnothing$, which, by the same argument as above, is equivalent to showing that $\pi(X(\mathbb{A}_{\mathbb{Q}})) \cap E(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \varnothing$. Since $Ш(k, E)$ is finite, $E(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \prod_{p<\infty}\{O\} \times C_O$, where $C_O \subset E(\mathbb{R})$ is the connected component of O. However, since $s_2(P) = 0$ and $q \notin \mathbb{F}_{p_i}^{\times 2}$ for $i = 1, 2$, $X_O(\mathbb{Q}_{p_1}) = X_O(\mathbb{Q}_{p_2}) = \varnothing$, so $X(\mathbb{A}_{\mathbb{Q}})^{\mathrm{Br}} = \varnothing$.

Now we will show that X has a globally generated k-rational ample line bundle of degree 12. Consider the morphism $f\colon X \to \mathbb{P}^3 \times \mathbb{P}^1$ that is the composition of the following maps

$$X \hookrightarrow \mathbb{P}(\mathcal{O}_E \oplus \mathcal{O}_E \oplus \mathcal{O}_E(2\infty)) \to \mathbb{P}(\mathcal{O}_E^{\oplus 4}) \simeq \mathbb{P}^3 \times E \xrightarrow{(\mathrm{id}, x)} \mathbb{P}^3 \times \mathbb{P}^1,$$

where the second map (from left to right) is induced by a surjection $\mathcal{O}_E^{\oplus 2} \to \mathcal{O}_E(2O)$ and the morphism g that is the composition of f with the Segre embedding $\mathbb{P}^3 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^7$. Note that f, and therefore g, is 2-to-1 onto its image, so $\mathcal{L} := g^*\mathcal{O}_{\mathbb{P}^7}(1)$ is a globally generated ample line bundle on X.

The image of f is the complete intersection of a $(2, 0)$ hypersurface, which comes from the quadric relation s, and a $(1, 1)$ hypersurface, which comes from the kernel of $\mathcal{O}_E^{\oplus 2} \to \mathcal{O}_E(2O)$. Since a quadric surface in $\mathbb{P}^3$ is geometrically isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, the image of g is geometrically isomorphic to a hyperplane intersected with the image Y of $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^7$. Since the Hilbert polynomial of Y is $(1 + x)^3$, the degree of Y, and hence the degree of image g, is $3! = 6$. Therefore, $\deg(\mathcal{L}) = 12$ and so degrees do *not* capture the Brauer–Manin obstruction to rational points on X. □

Appendix: Proof of Theorem 1.8

by Alexei N. Skorobogatov

The following, statement equivalent to Theorem 1.8, is a common generalization of [Skorobogatov and Zarhin 2017, Theorem 3.3] and Theorem 1.7.

Let $\mathrm{Br}(X)_{\mathrm{odd}}$ be the subgroup of $\mathrm{Br}(X)$ formed by the elements of odd order.

Theorem A.1. *Let A be an abelian variety of dimension > 1 over a number field k. Let X be the Kummer variety attached to a 2-covering of A. If B is a subgroup of* $\mathrm{Br}(X)$ *such that $X(\mathbb{A}_k)^B \neq \varnothing$, then $X(\mathbb{A}_k)^{B+\mathrm{Br}(X)_{\mathrm{odd}}} \neq \varnothing$.*

Proof. By [Skorobogatov and Zarhin 2017, Corollary 2.8] the group $\mathrm{Br}(X)/\mathrm{Br}_0(X)$ is finite. Hence we can assume without loss of generality that B is finite. Replacing B by its 2-primary torsion subgroup we can assume that $B \subset \mathrm{Br}(X)[2^\infty]$. There exists an odd integer n such that $\mathrm{Br}(X)[n]$ and $\mathrm{Br}(X)_{\mathrm{odd}}$ have the same image in $\mathrm{Br}(X)$.

By the definition of a Kummer variety [Skorobogatov and Zarhin 2017, Definition 2.1] there exists a double cover $\pi : Y' \to X$, where $\sigma : Y' \to Y$ is the blowing-up of a 2-covering $f : Y \to A$ of A in $f^{-1}(0)$, such that the branch locus of π is the exceptional divisor of σ. Since n is odd and Y is a torsor for A of period dividing 2, we have a well defined morphism $[n] : Y \to Y$ compatible with multiplication by n on A.

Let $(P_v) \in X(\mathbb{A}_k)$. For each v there is a class $\alpha_v \in \mathrm{H}^1(k_v, \mu_2) = k_v^*/k_v^{*2}$ such that P_v lifts to a k_v-point on the quadratic twist Y'_{α_v}, which is a variety defined over k_v. By weak approximation we can assume that α_v comes from k^*/k^{*2} and hence Y'_{α_v} is actually defined over k. Let $\pi_{\alpha_v} : Y'_{\alpha_v} \to X$ be the natural double cover.

Let $D = \pi^*_{\alpha_v}(B)$ be the image of B in $\mathrm{Br}(Y_{\alpha_v})$. We need the following corollary of Lemma 4.6.

Lemma A.2. *There exists a positive integer s such that $[n^s]^* : \mathrm{Br}(Y_{\alpha_v}) \to \mathrm{Br}(Y_{\alpha_v})$ restricted to D is the identity map on D.*

Proof. By Lemma 4.6 for $d = 2$ there is a positive integer a such that $[n^a]$ induces the identity on the quotient of D by $D \cap \mathrm{Br}_0(Y_{\alpha_v})$. Hence for each $\mathcal{A} \in D$ there is an $\mathcal{A}_0 \in \mathrm{Br}_0(Y_{\alpha_v})$ such that $[n^a]^*\mathcal{A} = \mathcal{A} + \mathcal{A}_0$. Let b be the product of orders of $\mathcal{A}_0$, where $\mathcal{A}$ ranges over all elements of D. Then $s = ab$ satisfies the conclusion of the lemma. □

Now assume that $(P_v) \in X(\mathbb{A}_k)^B$. For each place v of k choose a k_v-point of Y'_{α_v} above P_v and let R_v be its image in Y_{α_v}. In the proof of [Skorobogatov and Zarhin 2017, Theorem 3.3] instead of $M_v = [n]R_v$ put $M_v = [n^s]R_v$, where s is as in Lemma A.2. This definition has all the properties needed for the proof of [loc. cit., Theorem 3.3] with the additional property

$$\mathcal{A}(M_v) = ([n^s]^*\mathcal{A})(R_v) = \mathcal{A}(R_v) \tag{A-1}$$

for each $\mathcal{A} \in \pi^*_{\alpha_v}(B)$. We now proceed exactly as in the proof of [loc. cit., Theorem 3.3]. By a small deformation we can assume that M_v avoids the exceptional divisor of $Y'_{\alpha_v} \to Y_{\alpha_v}$. Then M_v lifts to a unique point $M'_v \in Y'_{\alpha_v}(k_v)$. Let $Q_v = \pi_{\alpha_v}(M'_v) \in X(k_v)$. By (A-1) for each $\mathcal{B} \in B$ we have $\mathcal{B}(Q_v) = \mathcal{B}(P_v)$, hence $(Q_v) \in X(\mathbb{A}_k)^B$. The proof of [loc. cit., Theorem 3.3] shows that $(Q_v) \in X(\mathbb{A}_k)^{\mathrm{Br}(X)_{\mathrm{odd}}}$. Thus $(Q_v) \in X(\mathbb{A}_k)^{B+\mathrm{Br}(X)_{\mathrm{odd}}}$. □

Acknowledgements

This article was sparked by a discussion at the *Rational Points* 2015; we thank Michael Stoll for organizing a stimulating workshop. We also thank Jean-Louis Colliot-Thélène for comments regarding the origins of

Conjecture (E), Evis Ieronymou for pointing out a small error in Section 3 of an earlier draft, Sándor Kovács for discussions regarding the proof of Lemma 5.5, and Alexei Skorobogatov for the appendix and for helpful comments. We also thank the referee for several helpful comments.

References

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

[Basile and Skorobogatov 2003] C. L. Basile and A. Skorobogatov, "On the Hasse principle for bielliptic surfaces", pp. 31–40 in *Number theory and algebraic geometry*, edited by M. Reid and A. Skorobogatov, London Math. Soc. Lecture Note Ser. **303**, Cambridge Univ. Press, 2003. MR Zbl

[Beauville 1996] A. Beauville, *Complex algebraic surfaces*, 2nd ed., London Mathematical Society Student Texts **34**, Cambridge University Press, 1996. MR Zbl

[Berkovič 1972] V. G. Berkovič, "The Brauer group of abelian varieties", *Funkcional Anal. i Priložen.* **6**:3 (1972), 10–15. In Russian; translated in Functional Anal. Appl. **6** (1972), 180–184. MR Zbl

[Cassels 1962] J. W. S. Cassels, "Arithmetic on curves of genus 1, IV: Proof of the Hauptvermutung", *J. Reine Angew. Math.* **211** (1962), 95–112. MR Zbl

[Clark and Sharif 2010] P. L. Clark and S. Sharif, "Period, index and potential, III", *Algebra Number Theory* **4**:2 (2010), 151–174. MR Zbl

[Colliot-Thélène 1995] J.-L. Colliot-Thélène, "L'arithmétique du groupe de Chow des zéro-cycles", *J. Théor. Nombres Bordeaux* **7**:1 (1995), 51–73. MR Zbl

[Colliot-Thélène 1999] J.-L. Colliot-Thélène, "Conjectures de type local-global sur l'image des groupes de Chow dans la cohomologie étale", pp. 1–12 in *Algebraic K-theory* (Seattle, 1997), edited by W. Raskind and C. Weibel, Proc. Sympos. Pure Math. **67**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Colliot-Thélène and Sansuc 1981] J.-L. Colliot-Thélène and J.-J. Sansuc, "On the Chow groups of certain rational surfaces: a sequel to a paper of S. Bloch", *Duke Math. J.* **48**:2 (1981), 421–447. MR Zbl

[Colliot-Thélène et al. 2016] J.-L. Colliot-Thélène, A. Pál, and A. N. Skorobogatov, "Pathologies of the Brauer–Manin obstruction", *Math. Z.* **282**:3-4 (2016), 799–817. MR Zbl

[Corn 2007] P. Corn, "The Brauer–Manin obstruction on del Pezzo surfaces of degree 2", *Proc. Lond. Math. Soc.* (3) **95**:3 (2007), 735–777. MR Zbl

[Corn and Nakahara 2017] P. Corn and M. Nakahara, "Brauer–Manin obstructions on genus-2 K3 surfaces", 2017. arXiv

[Creutz et al. 2018] B. Creutz, B. Viray, and J. F. Voloch, "The d-primary Brauer–Manin obstruction for curves", *Res. Number Theory* **4**:2 (2018), Art. 26, 16. MR

[Grothendieck 1968] A. Grothendieck, "Le groupe de Brauer, III: Exemples et compléments", pp. 88–188 in *Dix exposés sur la cohomologie des schémas*, Adv. Stud. Pure Math. **3**, North-Holland, Amsterdam, 1968. MR Zbl

[van Hamel 2003] J. van Hamel, "The Brauer–Manin obstruction for zero-cycles on Severi-Brauer fibrations over curves", *J. London Math. Soc.* (2) **68**:2 (2003), 317–337. MR Zbl

[Harari and Szamuely 2008] D. Harari and T. Szamuely, "Local-global principles for 1-motives", *Duke Math. J.* **143**:3 (2008), 531–557. MR Zbl

[Harpaz and Skorobogatov 2016] Y. Harpaz and A. N. Skorobogatov, "Hasse principle for Kummer varieties", *Algebra Number Theory* **10**:4 (2016), 813–841. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math. **52**, Springer, 1977. MR Zbl

[Hassett 2009] B. Hassett, "Rational surfaces over nonclosed fields", pp. 155–209 in *Arithmetic geometry*, edited by H. Darmon et al., Clay Math. Proc. **8**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Ieronymou and Skorobogatov 2015] E. Ieronymou and A. N. Skorobogatov, "Odd order Brauer–Manin obstruction on diagonal quartic surfaces", *Adv. Math.* **270** (2015), 181–205. MR Zbl

[Ieronymou and Skorobogatov 2017] E. Ieronymou and A. N. Skorobogatov, "Corrigendum to "Odd order Brauer–Manin obstruction on diagonal quartic surfaces"", *Adv. Math.* **307** (2017), 1372–1377. MR Zbl

[Ieronymou et al. 2011] E. Ieronymou, A. N. Skorobogatov, and Y. G. Zarhin, "On the Brauer group of diagonal quartic surfaces", *J. Lond. Math. Soc.* (2) **83**:3 (2011), 659–672. MR Zbl

[Kato and Saito 1986] K. Kato and S. Saito, "Global class field theory of arithmetic schemes", pp. 255–331 in *Applications of algebraic K-theory to algebraic geometry and number theory, Part I, II* (Boulder, Colorado, 1983), edited by S. J. Bloch et al., Contemp. Math. **55**, Amer. Math. Soc., Providence, RI, 1986. MR Zbl

[Lang 1954] S. Lang, "Some applications of the local uniformization theorem", *Amer. J. Math.* **76** (1954), 362–374. MR Zbl

[Lang and Tate 1958] S. Lang and J. Tate, "Principal homogeneous spaces over abelian varieties", *Amer. J. Math.* **80** (1958), 659–684. MR Zbl

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics] **49**, Springer, 2004. MR Zbl

[LMFDB Collaboration 2013] T. LMFDB Collaboration, "The L-functions and modular forms database", 2013, Available at http://www.lmfdb.org.

[Manin 1971] Y. I. Manin, "Le groupe de Brauer–Grothendieck en géométrie diophantienne", pp. 401–411 in *Actes du Congrès International des Mathématiciens, Tome 1* (Nice, 1970), Gauthier-Villars, Paris, 1971. MR Zbl

[Nakahara 2017] M. Nakahara, "Index of fibrations and Brauer classes that never obstruct the Hasse principle", 2017. arXiv

[Nishimura 1955] H. Nishimura, "Some remarks on rational points", *Mem. Coll. Sci. Univ. Kyoto. Ser. A. Math.* **29** (1955), 189–192. MR Zbl

[Poonen 2009] B. Poonen, "Existence of rational points on smooth projective varieties", *J. Eur. Math. Soc. (JEMS)* **11**:3 (2009), 529–543. MR Zbl

[Sansuc 1981] J.-J. Sansuc, "Groupe de Brauer et arithmétique des groupes algébriques linéaires sur un corps de nombres", *J. Reine Angew. Math.* **327** (1981), 12–80. MR Zbl

[Serrano 1990] F. Serrano, "Divisors of bielliptic surfaces and embeddings in $\mathbf{P}^4$", *Math. Z.* **203**:3 (1990), 527–533. MR Zbl

[Serrano 1996] F. Serrano, "Isotrivial fibred surfaces", *Ann. Mat. Pura Appl.* (4) **171** (1996), 63–81. MR Zbl

[Skorobogatov 1999] A. N. Skorobogatov, "Beyond the Manin obstruction", *Invent. Math.* **135**:2 (1999), 399–424. MR Zbl

[Skorobogatov 2001] A. Skorobogatov, *Torsors and rational points*, Cambridge Tracts in Mathematics **144**, Cambridge University Press, 2001. MR Zbl

[Skorobogatov and Swinnerton-Dyer 2005] A. Skorobogatov and P. Swinnerton-Dyer, "2-descent on elliptic curves and rational points on certain Kummer surfaces", *Adv. Math.* **198**:2 (2005), 448–483. MR Zbl

[Skorobogatov and Zarhin 2008] A. N. Skorobogatov and Y. G. Zarhin, "A finiteness theorem for the Brauer group of abelian varieties and $K3$ surfaces", *J. Algebraic Geom.* **17**:3 (2008), 481–502. MR Zbl

[Skorobogatov and Zarhin 2017] A. N. Skorobogatov and Y. G. Zarhin, "Kummer varieties and their Brauer groups", *Pure Appl. Math. Q.* **13**:2 (2017), 337–368. MR Zbl

[Swinnerton-Dyer 1993] P. Swinnerton-Dyer, "The Brauer group of cubic surfaces", *Math. Proc. Cambridge Philos. Soc.* **113**:3 (1993), 449–460. MR Zbl

[Várilly-Alvarado 2013] A. Várilly-Alvarado, "Arithmetic of del Pezzo surfaces", pp. 293–319 in *Birational geometry, rational curves, and arithmetic*, edited by F. Bogomolov et al., Springer, 2013. MR Zbl

[Wittenberg 2012] O. Wittenberg, "Zéro-cycles sur les fibrations au-dessus d'une courbe de genre quelconque", *Duke Math. J.* **161**:11 (2012), 2113–2166. MR Zbl

Communicated by Jean-Louis Colliot-Thélène
Received 2017-12-19 Revised 2018-07-12 Accepted 2018-08-23

brendan.creutz@canterbury.ac.nz School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand

bviray@math.washington.edu
Department of Mathematics, University of Washington, Seattle, WA, United States

a.skorobogatov@imperial.ac.uk
Department of Mathematics, Imperial College London, South Kensington Campus, United Kingdom

Institute for the Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2471

Bounds for traces of Hecke operators and applications to modular and elliptic curves over a finite field

Ian Petrow

We give an upper bound for the trace of a Hecke operator acting on the space of holomorphic cusp forms with respect to certain congruence subgroups. Such an estimate has applications to the analytic theory of elliptic curves over a finite field, going beyond the Riemann hypothesis over finite fields. As the main tool to prove our bound on traces of Hecke operators, we develop a Petersson formula for newforms for general nebentype characters.

1. Introduction

1A. ***Statement of results.*** Let $S_\kappa(\Gamma,\epsilon)$ be the space of holomorphic cusp forms of weight κ, for a subgroup Γ of a Hecke congruence group, and of nebentype character ϵ. We write $\mathrm{Tr}(T|S_\kappa(\Gamma,\epsilon))$ for the trace of a linear operator T acting on $S_\kappa(\Gamma,\epsilon)$. The aim of this paper is to give estimates for $\mathrm{Tr}(T_m|S_\kappa(\Gamma,\epsilon))$, where T_m is the m-th Hecke operator, as the parameters m,κ,Γ, and ϵ vary simultaneously.

Consider first the case that $\Gamma=\Gamma_0(N)$ and ϵ is any Dirichlet character modulo N. Let $d(m)$ denote the number of divisors of m, let $\sigma(m)$ denote the sum of the divisors of m, and let $\psi(N)=[\Gamma_0(N):\mathrm{SL}_2(\mathbb{Z})]=N\prod_{p|N}\bigl(1+\frac{1}{p}\bigr)$. We assume that $\kappa\geqslant 2$ is an integer throughout the paper. Deligne's theorem tells us that each eigenvalue $\lambda(m)$ of T_m satisfies $|\lambda(m)|\leqslant d(m)m^{(\kappa-1)/2}$. Therefore we have the "trivial" estimate on the trace

$$\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon))\leqslant \dim S_\kappa(\Gamma_0(N),\epsilon)d(m)m^{\frac{\kappa-1}{2}}\leqslant\frac{(\kappa-1)\psi(N)}{12}d(m)m^{\frac{\kappa-1}{2}}. \tag{1-1}$$

For the bound on $\dim S_\kappa(\Gamma_0(N),\epsilon)$, see, e.g., [Ross 1992, Corollary 8]. The power of m in (1-1) is sharp by the Sato–Tate distribution for Hecke eigenvalues. On the other hand, by a careful analysis using the Eichler–Selberg trace formula, Conrey, Duke and Farmer [Conrey et al. 1997] and in more generality Serre [1997, Proposition 4] showed that if $\epsilon(-1)=(-1)^\kappa$ then

$$\begin{aligned}&\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon))\\ &\qquad=\frac{\kappa-1}{12}\epsilon(m^{\frac12})m^{\frac{\kappa}{2}-1}\psi(N)+O\Bigl(\bigl(\sigma(m)\max_{f^2<4m}\psi(f)+d(m)N^{\frac12}\bigr)m^{\frac{\kappa-1}{2}}d(N)\Bigr),\end{aligned} \tag{1-2}$$

This work was supported by Swiss National Science Foundation grant PZ00P2_168164.
MSC2010: primary 11F25; secondary 11F11, 11F72, 11G20, 14G15.
Keywords: traces of Hecke operators, modular curves over a finite field, elliptic curves over a finite field, Petersson formula for newforms, Tsfasman–Vlăduț–Zink theorem.

where $\epsilon(m^{1/2})$ is understood to be 0 unless m is a perfect square. If $\epsilon(-1)\neq(-1)^\kappa$ then $S_\kappa(\Gamma_0(N),\epsilon)=\{0\}$, so the left-hand side vanishes identically. We expect the estimate (1-2) to be sharp if m is fixed and $\kappa+N\to\infty$.

Write $\boldsymbol{c}(\epsilon)$ for the conductor of the Dirichlet character ϵ, and $\boldsymbol{c}^*(\epsilon)=\prod_{p\mid \boldsymbol{c}(\epsilon)} p$ for its square-free part. In this paper we prove:

Theorem 1.1. *Suppose that* $\epsilon(-1)=(-1)^\kappa$, $(N,m)=1$, *and* $m\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)\ll(N^4\kappa^{10/3})^{1-\eta}$ *for some* $\eta>0$. *Then we have*

$$\begin{aligned}\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon))\\ &=\frac{\kappa-1}{12}\epsilon(m^{\frac12})m^{\frac{\kappa}{2}-1}\psi(N)+O_{\eta,\varepsilon}(N^{\frac{10}{11}}m^{\frac{\kappa-1}{2}+\frac{1}{44}}\kappa^{\frac{61}{66}}\boldsymbol{c}(\epsilon)^{\frac{1}{44}}\boldsymbol{c}^*(\epsilon)^{\frac{1}{44}}(Nm\kappa)^{\varepsilon}).\end{aligned}\tag{1-3}$$

We remark that the hypothesis that $m\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)\ll(N^4\kappa^{10/3})^{1-\eta}$ for some $\eta>0$ in Theorem 1.1 is no restriction in practice, since if the hypothesis fails then (1-1) is a superior bound anyway. Indeed, the error term in (1-3) is smaller than that in both (1-1) and (1-2) when

$$N^{\frac{8}{13}}\kappa^{\frac{122}{195}}(N\kappa)^{\varepsilon}\boldsymbol{c}(\epsilon)^{\frac{1}{65}}\boldsymbol{c}^*(\epsilon)^{\frac{1}{65}}\ll m\ll\frac{(N^4\kappa^{\frac{10}{3}})^{1-\eta}}{\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)}.$$

For example, if ϵ is trivial and the weight κ is fixed, then (1-3) is better than (1-1) and (1-2) for

$$N^{\frac{8}{13}+\varepsilon}\ll m\ll N^{4-\varepsilon}.$$

Note that our result requires the hypothesis $(N,m)=1$, whereas the estimates (1-1) and (1-2) do not. We discuss the source of this condition in the sketch of the proof, below.

We are also interested in spaces of modular forms for groups other than $\Gamma_0(N)$. In particular, for positive integers $M\mid N$ let

$$\Gamma(M,N)=\left\{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z}):a,d\equiv 1\ (\mathrm{mod}\ N),\ c\equiv 0\ (\mathrm{mod}\ NM)\right\}.\tag{1-4}$$

These congruence groups interpolate between $\Gamma_1(N)=\Gamma(1,N)$ and $\Gamma(N)\simeq\Gamma(N,N)$. We write $S_\kappa(M,N)$ for the space of modular forms of weight κ for the group $\Gamma(M,N)$ (without nebentype character). Let $\delta(a,b)$ be the indicator function of $a=b$ and $\delta_c(a,b)$ be the indicator function of $a\equiv b\ (\mathrm{mod}\ c)$. Let T_m be the m-th Hecke operator acting on $S_\kappa(M,N)$ and for $(d,N)=1$ let $\langle d\rangle$ be the d-th diamond operator. These operators commute and $T_1=\langle 1\rangle=\mathrm{id}$; for definitions see [Diamond and Shurman 2005, §5.1, 5.2] or [Kaplan and Petrow 2017, §4]. In particular, we have

$$\mathrm{Tr}(\langle d\rangle T_m|S_\kappa(\Gamma(M,N)))=\sum_{\epsilon\ (\mathrm{mod}\ N)}\epsilon(d)\,\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(NM),\epsilon)).\tag{1-5}$$

Applying (1-1) to (1-5) we have

$$\mathrm{Tr}(\langle d\rangle T_m|S_\kappa(\Gamma(M,N)))\leqslant\frac{\kappa-1}{12}\varphi(N)\psi(NM)d(m)m^{\frac{\kappa-1}{2}}.\tag{1-6}$$

Meanwhile, summing (1-2) over characters $\epsilon \pmod N$ such that $\epsilon(-1) = (-1)^\kappa$ we find

$$\operatorname{Tr}(\langle d\rangle T_m | S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24} m^{\frac{\kappa}{2}-1}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1) + (-1)^\kappa \delta_N(m^{\frac{1}{2}}d,-1)\big) + O\Big(\big(\sigma(m)\max_{f^2<4m}\psi(f) + d(m)(MN)^{\frac{1}{2}}\big)m^{\frac{\kappa-1}{2}} d(MN)N\Big). \quad (1\text{-}7)$$

The following result improves on both (1-6) and (1-7) in an intermediate range of parameters.

Theorem 1.2. *Suppose that $M \mid N$, $(N,m) = 1$, and $m \ll (N^6\kappa^{10/3})^{1-\eta}$ for some $\eta > 0$. We have*

$$\operatorname{Tr}(\langle d\rangle T_m | S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24} m^{\frac{\kappa}{2}-1}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1) + (-1)^\kappa \delta_N(m^{\frac{1}{2}}d,-1)\big) + O_{\eta,\varepsilon}\big(MN^{\frac{41}{22}} m^{\frac{\kappa-1}{2}+\frac{1}{44}}\kappa^{\frac{61}{66}}(Nm\kappa)^\varepsilon\big).$$

1B. *Applications to modular and elliptic curves over a finite field.* Hecke operators appear throughout number theory, and estimates for their traces are especially relevant to equidistribution problems. See for example [Serre 1997, §5–§8] and [Murty and Sinha 2009]. We mention here a few consequences in the analytic theory of modular and elliptic curves over a finite field.

Let C be a nonsingular projective curve of genus g over a finite field $\mathbb{F}_q$ with q elements. Then we have (see, e.g., [Milne 2017, Chapter 11])

$$|C(\mathbb{F}_{q^n})| = q^n + 1 - \sum_{i=1}^{2g} \alpha_i^n,$$

where $\{\alpha_i\}$ are the inverse zeros of the zeta function of C

$$Z(C,T) = \frac{(1-\alpha_1 T)\cdots(1-\alpha_{2g}T)}{(1-T)(1-qT)}.$$

The Riemann hypothesis for curves over finite fields asserts that $|\alpha_i| = \sqrt{q}$ for all i. Igusa [1959] showed that there exists a nonsingular projective model for $X_0(N)$ over $\mathbb{Q}$ whose reductions modulo primes p, $p \nmid N$, are also nonsingular (see also the survey [Diamond and Im 1995, §9]), and so the preceding discussion applies to $X_0(N)$ when $p \nmid N$. Since $g \sim \psi(N)/12$ as $N \to \infty$ we have

$$|X_0(N)(\mathbb{F}_q)| = q + 1 + O(\psi(N)q^{1/2}). \quad (1\text{-}8)$$

In particular, $|X_0(N)(\mathbb{F}_q)| \sim q$ as $q \to \infty$ as soon as $q \gg N^{2+\delta}$ for some $\delta > 0$. On the other hand, the Eichler–Shimura correspondence (see, e.g., [Milne 2017, Theorem 11.14]) asserts that

$$Z(X_0(N),T) = \frac{\prod_{f\in H_2(N)}(1-\lambda_f(p)T + pT^2)}{(1-T)(1-pT)},$$

where $H_2(N)$ is a basis for $S_2(\Gamma_0(N))$ consisting of eigenforms of $\{T_p : p \nmid N\}$ and $\lambda_f(p)$ is the T_p eigenvalue of f. We therefore have

$$|X_0(N)(\mathbb{F}_q)| = q + 1 - \operatorname{Tr}(T_q | S_2(\Gamma_0(N))) + p\operatorname{Tr}(T_{q/p^2}|S_2(\Gamma_0(N))),$$

where we set $T_{p^{-1}} = 0$. Applying (1-1), (1-2), and Theorem 1.1 we get:

Corollary 1.3. *Suppose $q = p^v$ is a prime power such that $p \nmid N$. We have*

$$|X_0(N)(\mathbb{F}_q)| = q + (p-1)\frac{\psi(N)}{12}\delta_2(v,0) + O_\varepsilon\big(\min(\psi(N), q^{\frac{1}{44}}N^{\frac{10}{11}}(qN)^\varepsilon, (q^{\frac{3}{2}} + N^{\frac{1}{2}})d(N)q^\varepsilon)q^{\frac{1}{2}}\big).$$

In particular, the main term is larger than the error term as soon as $q \gg N^{\frac{40}{21}+\delta}$ for some fixed $\delta > 0$.

Corollary 1.3 shows that there is significant cancellation between the zeros α_i of $Z(X_0(N), T)$, and in this sense it goes beyond the Riemann hypothesis for $Z(X_0(N), T)$. Assuming square-root cancellation between the zeros, one might conjecture an error term of size $(qN)^{1/2+\varepsilon}$ in Corollary 1.3, which would imply that the main term is larger than the error term whenever $q \gg N^{1+\delta}$ for some δ. If one assumes the generalized Lindelöf hypothesis for adjoint square L-functions, then the method in this paper produces an error term of size $q^{1/8+\varepsilon}N^{1/2+\varepsilon}$ in Corollary 1.3 (see Lemma 6.1). In a much more speculative direction, if under the assumption $(mn, W) = 1$ the upper bound $\ll_{\kappa,\varepsilon} (mnW)^\varepsilon W^{-1/2}$ for the sum appearing in Lemma 4.1 holds (cf. the Linnik–Selberg conjecture), then the error term $(qN)^{1/2+\varepsilon}$ in Corollary 1.3 is admissible.

If q is a square then we can compare the second main term in Corollary 1.3 to the error term coming from (1-2) in the range where q is small compared to N. For example, in the special case that p is a prime and $q = p^2$ we have:

Corollary 1.4. *If $p, N \to \infty$ where p runs through primes $p \nmid N$ then for any fixed $\delta > 0$ we have*

$$|X_0(N)(\mathbb{F}_{p^2})| = \begin{cases} p^2 + O(p\psi(N)) & \text{if } p^2 \gg N^{4-\delta}, \\ p^2 + p\frac{1}{12}\psi(N) + O_\varepsilon(p^{\frac{23}{22}}N^{\frac{10}{11}}(qN)^\varepsilon) & \text{if } N^{\frac{40}{21}-\delta} \ll p^2 \ll N^{4-\delta}, \\ p\frac{1}{12}\psi(N) + O_\varepsilon(p^{\frac{23}{22}}N^{\frac{10}{11}}(qN)^\varepsilon) & \text{if } N^{\frac{8}{13}+\delta} \ll p^2 \ll N^{\frac{40}{21}-\delta}, \\ (p-1)\frac{1}{12}\psi(N) + O_\varepsilon((p^4 + N^{\frac{1}{2}}p)d(N)p^\varepsilon) & \text{if } p^2 \ll N^{\frac{2}{3}-\delta}. \end{cases}$$

The first of these cases is just (1-8), and the last is the Tsfasman–Vlăduț–Zink theorem [Tsfasman et al. 1982], which has important applications to algebraic coding theory; see [Moreno 1991, Chapter 5].

Using Theorem 1.2 we can make more explicit statements about elliptic curves themselves. Let E be an elliptic curve defined over $\mathbb{F}_q$ and let $t_E = q + 1 - \#E(\mathbb{F}_q)$ be the trace of the associated Frobenius endomorphism. Hasse's theorem tells us that $|t_E| \leqslant 2\sqrt{q}$. The set of $\mathbb{F}_q$-isomorphism classes of elliptic curves defined over $\mathbb{F}_q$ is naturally a probability space where the probability of a singleton is given by

$$\boldsymbol{P}_q(\{E\}) = \frac{1}{q|\operatorname{Aut}_{\mathbb{F}_q}(E)|}.$$

We would like to study the expectations as $q \to \infty$ of various random variables associated to t_E or the structure of the group of $\mathbb{F}_q$-rational points of E. To be precise: let A be a finite abelian group with at most two generators, and let Φ_A denote the indicator function of the event that there exists an injective group homomorphism $A \hookrightarrow E(\mathbb{F}_q)$. Let $U_j(x)$ for $j \geqslant 0$ be the Chebyshev polynomials of the second kind. The Chebyshev polynomials form an orthonormal basis for the Hilbert space $L^2\big([-1, 1], \frac{2}{\pi}\sqrt{1-x^2}dx\big)$.

N. Kaplan and the author [Kaplan and Petrow 2017, Theorem 2] gave explicit formulas for the expectations

$$\boldsymbol{E}_q(U_j(t_E/2\sqrt{q})\Phi_A) = \frac{1}{q} \sum_{\substack{E/\mathbb{F}_q \\ A \hookrightarrow E(\mathbb{F}_q)}} \frac{U_j(t_E/2\sqrt{q})}{|\operatorname{Aut}_{\mathbb{F}_q}(E)|}$$

in terms of $\operatorname{Tr}(\langle d\rangle T_m | S_\kappa(\Gamma(M,N)))$ and elementary arithmetic functions of m, M, N, and j.

Theorem 1.2 yields the following refinement of the error term in the main corollary of [Kaplan and Petrow 2017]. Let

$$v(n_1,n_2) = \frac{n_1}{\psi(n_1)\varphi(n_1)n_2^2} \prod_{\ell \mid \frac{n_1}{(q-1,n_1)}} (1+\ell^{-1-2v_\ell(\frac{(q-1,n_1)}{n_2})}).$$

Corollary 1.5. *Let $n_1 = n_1(A)$ and $n_2 = n_2(A)$ be the first and second invariant factors of A (i.e., we have $n_2|n_1$). Suppose that $(|A|,q) = 1$ and $q \equiv 1 \pmod{n_2}$. Then*

$$\boldsymbol{E}_q(U_j(t_E/2\sqrt{q})\Phi_A) = v(n_1,n_2)\big(\delta(j,0) + O_{j,\varepsilon}(\min(n_1, q^{\frac{1}{44}} n_1^{\frac{19}{22}})n_1 n_2 q^{-\frac{1}{2}}(qn_1)^{\varepsilon})\big).$$

If $q \not\equiv 1 \pmod{n_2}$, then $\boldsymbol{E}_q(U_j\Phi_A)$ vanishes identically.

In particular, the traces of the Frobenius t_E for $\{E/\mathbb{F}_q : A \hookrightarrow E(\mathbb{F}_q)\}$ become equidistributed with respect to the Sato–Tate measure as $q \to \infty$ through prime powers $q \equiv 1 \pmod{n_2}$. The equidistribution is uniform in A as soon as $q \gg n_2^2 n_1^{41/11+\delta}$ for any fixed $\delta > 0$.

In [Kaplan and Petrow 2017] Kaplan and the author showed that the equidistribution of t_E for $\{E/\mathbb{F}_q : A \hookrightarrow E(\mathbb{F}_q)\}$ is uniform as soon as $q \gg n_2^2 n_1^{4+\delta}$ by applying (1-6) to bound the trace. In this sense, Corollary 1.5 goes beyond what one can conclude using the Riemann hypothesis of Deligne alone. All of the error terms in the theorems and corollaries found in Section 2 of [loc. cit.] are similarly improved by applying Theorem 1.2 in addition to (1-6).

1C. ***Outline of proof.*** Thanks to (1-5), the structural steps of the proof of Theorem 1.2 reduce to those of Theorem 1.1. The details of the analytic arguments differ however (see Section 5). For these reasons, we only discuss the proof of Theorem 1.1 in this outline.

By Atkin–Lehner theory, to estimate $\operatorname{Tr}(T_m | S_\kappa(\Gamma_0(N),\epsilon))$ it suffices to estimate

$$\sum_{f \in H_\kappa^\star(N,\epsilon)} \lambda_f(m), \tag{1-9}$$

where $H_\kappa^\star(N,\epsilon)$ is set of Hecke-normalized newforms of level N and character ϵ, and $\lambda_f(m)$ is the m-th Hecke eigenvalue of f, normalized so that $|\lambda_f(n)| \leqslant d(n)$. Whereas Serre and Conrey, Duke, and Farmer used the Eichler–Selberg trace formula to access the trace of T_m, we take a different path and use the Petersson trace formula.

Let $\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)$ be an orthonormal basis for $S_\kappa(\Gamma_0(N),\epsilon)$. Let $g \in \mathcal{B}_\kappa(\Gamma_0(N),\epsilon)$ and write its Fourier coefficients as $\{b_g(n)\}_{n\geqslant 1}$. Then the Petersson formula [Iwaniec and Kowalski 2004, Proposition 14.5] says that

$$\frac{\Gamma(\kappa-1)}{(4\pi\sqrt{mn})^{\kappa-1}} \sum_{f\in\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_f(n)\overline{b_f(m)} \\ = \delta(m,n) + 2\pi i^{-\kappa} \sum_{\substack{c>0 \\ c\equiv 0 \ (\mathrm{mod}\ N)}} \frac{S_\epsilon(m,n,c)}{c} J_{\kappa-1}\left(\frac{4\pi\sqrt{mn}}{c}\right), \tag{1-10}$$

where J_α is the J-Bessel function, $S_\epsilon(m,n,c)$ is the twisted Kloosterman sum

$$S_\epsilon(m,n,c) = \sum_{d \ (\mathrm{mod}\ c)}^{*} \epsilon(d) e\left(\frac{dm+\bar{d}n}{c}\right),$$

and the $*$ indicates we run over invertible $d \ (\mathrm{mod}\ c)$.

Our goal is to apply the Petersson formula to (1-9), and so we are faced with two technical difficulties:

(1) Only the newforms in $S_\kappa(\Gamma_0(N),\epsilon)$ have Fourier coefficients proportional to the Hecke eigenvalues appearing in (1-9).

(2) If f is a newform, the constant of proportionality between Fourier coefficients $b_f(n)$ and the Hecke eigenvalues $\lambda_f(n)$ is $\approx \|f\|_{L^2}$, which is not constant across $H_\kappa^\star(N,\epsilon)$.

We overcome (1) in Theorem 3.1 by developing a Petersson formula for newforms for $S_\kappa(\Gamma_0(N),\epsilon)$. There has been much recent interest in such formulas; see for example [Barrett et al. 2017; Nelson 2017; Petrow and Young 2018; Young 2018]. Theorem 3.1 is a generalization of [Barrett et al. 2017, Proposition 4.1] to nontrivial central characters, which itself is a generalization of work of Iwaniec, Luo and Sarnak [Iwaniec et al. 2000], Rouymi [2011] and Ng [2012]. Peter Humphries has also shared a preprint with the author in which he independently obtains Theorem 3.1, and uses it to study low-lying zeros of the L-functions associated to $f \in H_\kappa^\star(N,\epsilon)$. Theorem 3.1 is the only place in the proof where we have used the hypothesis $(N,m)=1$, in an essential way, and so is the source of the relatively prime conditions in Theorems 1.1 and 1.2.

We deal with (2) by appealing to the special value formula

$$L(1,\mathrm{Ad}^2 f) = \frac{\zeta^{(N)}(2)(4\pi)^\kappa}{\Gamma(\kappa)} \frac{\|f\|_{L^2}^2}{\mathrm{Vol}\, X_0(N)},$$

where $L(s,\mathrm{Ad}^2 f)$ is a certain Dirichlet series whose coefficients involve $\lambda_f(n^2)$, and which we discuss in more detail in Section 2. One may then swap the sum over f and this Dirichlet series, and apply our Petersson formula for newforms (Theorem 3.1). Estimating the resulting sums directly using the Weil bound for $S_\epsilon(a,b,c)$ (see Lemma 4.2), one recovers that the trace of T_m is $\ll_m (N\kappa)^{1+\varepsilon}$ (compare with (1-1)).

To save a bit more and obtain Theorem 1.1 we remove the weights $\|f\|_{L^2}^2$ more efficiently using a method due to Kowalski and Michel [1999, Proposition 2]. Their method is based on Hölder's inequality and a large sieve inequality due to Duke and Kowalski [2000, Theorem 4] for subfamilies of automorphic forms on GL_3. There are other notable large sieve inequalities for GL_3 in the literature; see, e.g., [Blomer et al. 2017, Theorem 3] and [Venkatesh 2006, Theorem 1]. However, these two are not useful to us since we need a large sieve inequality which is efficient for the proper subfamily of GL_3 forms cut out by the image of the adjoint square lift from GL_2. The inequality of Duke and Kowalski is superior to the results [Blomer et al. 2017, Theorem 3] and [Venkatesh 2006, Theorem 1] in the case of a thin subfamily and a long summation variable, which is the situation of interest to us.

2. Preliminaries on L-series

If $L(s)$ is a meromorphic function defined in $\mathrm{Re}(s) \gg 1$ by an infinite product over primes p of local factors $L_p(s)$, then for any integer N we write

$$L^{(N)}(s) = \prod_{p \nmid N} L_p(s) \quad \text{and} \quad L_N(s) = \prod_{p \mid N} L_p(s),$$

so that $L(s) = L_N(s)L^{(N)}(s)$ for any $N \in \mathbb{N}$. To deal with the $\|f\|_{L^2}^2$-normalization alluded to in Section 1C, we introduce the "naive" adjoint square L-function. For $f \in H_\kappa^\star(N, \epsilon)$, let

$$L(s, \mathrm{Ad}^2 f) = \frac{\zeta^{(N)}(2s)}{\zeta(s)} \sum_{n \geqslant 1} \frac{|\lambda_f(n)|^2}{n^s} = \prod_p L_p(s, \mathrm{Ad}^2 f),$$

where $\zeta(s)$ is the Riemann zeta function, and where

$$L_p(s, \mathrm{Ad}^2 f) = \begin{cases} \left(1 - \dfrac{1}{p^{2s}}\right)^{-1} \displaystyle\sum_{\alpha \geqslant 0} \frac{\bar{\epsilon}(p^\alpha)\lambda_f(p^{2\alpha})}{p^{\alpha s}} & \text{if } p \nmid N, \\ \left(1 - \dfrac{1}{p^s}\right)\left(1 - \dfrac{|\lambda_f(p)|^2}{p^s}\right)^{-1} & \text{if } p \mid N. \end{cases} \tag{2-1}$$

Warning: the $L(s, \mathrm{Ad}^2 f)$ is *not* the true adjoint square L-function of f as defined by functoriality (see [Iwaniec and Kowalski 2004, p. 133] and the online errata). But if $p \nmid N$, then $L_p(s, \mathrm{Ad}^2 f)$ *does* match the local L-factor at p of the true adjoint square L-function. Our "naive" adjoint square L function $L(s, \mathrm{Ad}^2 f)$ is chosen to be the Dirichlet series for which the following lemma is true.

Lemma 2.1. *The series $L(s, \mathrm{Ad}^2 f)$ defined above is holomorphic for* $\mathrm{Re}(s) > 0$ *and*

$$L(1, \mathrm{Ad}^2 f) = \frac{\zeta^{(N)}(2)(4\pi)^\kappa}{\Gamma(\kappa)} \frac{\langle f, f \rangle_N}{\mathrm{Vol}\, X_0(N)}, \tag{2-2}$$

where

$$\langle f, f \rangle_N = \int_{\Gamma_0(N)\backslash\mathcal{H}} |f(z)|^2 y^\kappa \frac{dx\, dy}{y^2}$$

and

$$\mathrm{Vol}\, X_0(N) = \int_{\Gamma_0(N)\backslash\mathcal{H}} \frac{dx\, dy}{y^2} = \frac{\pi}{3}\psi(N).$$

Proof. For the first statement, let π denote the irreducible admissible cuspidal automorphic representation of GL_2 generated by f, and denote by $L(s, \mathrm{Ad}^2\,\pi)$ the L-function of its adjoint square lift. We have by [Gelbart and Jacquet 1978] that $L(s, \mathrm{Ad}^2\,\pi)$ is an entire function of s. Therefore, the prime-to-N part of the naive L-function $L^{(N)}(s, \mathrm{Ad}^2\, f)$ is holomorphic for $\mathrm{Re}(s) > 0$.

For the second statement, take the standard nonholomorphic Eisenstein series for $\Gamma_0(N)$ at the cusp ∞ given by

$$E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \mathrm{Im}(\gamma z)^s.$$

Then we have by the classical Rankin–Selberg unfolding argument

$$\int_{\Gamma_0(N)\backslash\mathcal{H}} |f(z)|^2 E(z,s) y^\kappa \,\frac{dx\,dy}{y^2} = \frac{\Gamma(s+\kappa-1)}{(4\pi)^{s+\kappa-1}} \sum_{n\geqslant 1} \frac{|\lambda_f(n)|^2}{n^s}.$$

We deduce the lemma by taking residues on both sides and recalling [Iwaniec 1997, Theorem 13.2] that

$$\mathrm{Res}_{s=1}\, E(z,s) = \mathrm{Vol}\, X_0(N)^{-1}. \qquad \square$$

Let $\varrho_f(n)$ be the Dirichlet series coefficients of $L^{(N)}(s, \mathrm{Ad}^2\, f)$. Explicitly,

$$\varrho_f(n) = \begin{cases} \sum_{n=m^2\ell} \bar\epsilon(\ell)\lambda_f(\ell^2) & \text{if } (n,N)=1, \\ 0 & \text{if } (n,N)>1. \end{cases} \tag{2-3}$$

Inverting, we also have

$$\bar\epsilon(n)\lambda_f(n^2) = \sum_{m^2\ell=n} \mu(m)\varrho_f(\ell). \tag{2-4}$$

For future reference, we write the partial sums of $L^{(N)}(1, \mathrm{Ad}^2\, f)$ compactly as

$$\omega_f(x) = \sum_{n\leqslant x} \frac{\varrho_f(n)}{n}. \tag{2-5}$$

By contrast, when $p\,|\,N$ we have that $L_p(s, \mathrm{Ad}^2\, f)$ is constant along $f \in H_\kappa^\star(N,\epsilon)$ by the following lemma.

Lemma 2.2 [Ogg 1969, Theorems 2 and 3]. *Let $p\,|\,N$ be a prime, and ϵ a Dirichlet character mod N. Write*

$$a_{N,\epsilon}(p) = \begin{cases} 1 & \text{if } \epsilon \text{ is } \textbf{not} \text{ a character mod } N/p, \\ \frac{1}{p} & \text{if } \epsilon \text{ is a character mod } N/p \text{ and } p^2 \nmid N, \\ 0 & \text{if } \epsilon \text{ is a character mod } N/p \text{ and } p^2\,|\,N. \end{cases}$$

Then we have $|\lambda_f(p)|^2 = a_{N,\epsilon}(p)$.

3. Structural steps

We study the operator $T'_m = T_m/m^{(\kappa-1)/2}$ on $S_\kappa(\Gamma_0(N),\epsilon)$ so that each eigenvalue $\lambda_f(m)$ of the T'_m operator is normalized by Deligne's theorem to have $|\lambda_f(m)| \leqslant d(m)$. We write $H_\kappa^\star(N,\epsilon)$ for the set of

Hecke-normalized newforms in $S_\kappa(N,\epsilon)$ in the sense of Atkin–Lehner theory [1970]; see also [Li 1975]. Also by Atkin–Lehner theory we have when $(m,N)=1$ that

$$\mathrm{Tr}(T'_m|S_\kappa(\Gamma_0(N),\epsilon)) = \sum_{LM=N} d(L) \sum_{f\in H^\star_\kappa(M,\epsilon)} \lambda_f(m), \tag{3-1}$$

where we consider the interior sum to be empty if ϵ is not a character mod M. Thanks to (1-5), we can reduce the structural steps for traces on $S_\kappa(\Gamma(M,N))$ to the case of $S_\kappa(\Gamma_0(N),\epsilon)$.

Recall the notation from Section 1C and write $c_\kappa = \Gamma(\kappa-1)/(4\pi)^{\kappa-1}$. Let

$$\Delta_{\kappa,N,\epsilon}(m,n) = \frac{c_\kappa}{(\sqrt{mn})^{\kappa-1}} \sum_{f\in\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_f(n)\overline{b_f(m)},$$

so that the Petersson formula (1-10) is

$$\Delta_{\kappa,N,\epsilon}(m,n) = \delta(m,n) + 2\pi i^{-\kappa} \sum_{\substack{c>0\\ c\equiv 0 \ (\mathrm{mod}\ N)}} \frac{S_\epsilon(m,n,c)}{c} J_{\kappa-1}\left(\frac{4\pi\sqrt{mn}}{c}\right). \tag{3-2}$$

The following theorem is our main tool for computing sums over the set of newforms $H^\star_\kappa(N,\epsilon)$.

Theorem 3.1. *If $(mn,N)=1$ then we have*

$$c_\kappa \sum_{f\in H^\star_\kappa(N,\epsilon)} \frac{\overline{\lambda_f(m)}\lambda_f(n)}{\langle f,f\rangle_N} = \sum_{LM=N} \mu(L) R(M,L,\epsilon) \sum_{\substack{\ell | L^\infty\\ (\ell,M)=1}} \frac{\bar\epsilon(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m,n\ell^2),$$

where

$$R(M,L,\epsilon) := \frac{1}{L} \prod_{\substack{p^2|L\\ p\nmid M}} \left(1-\frac{1}{p^2}\right)^{-1} \prod_{p|(M,L)} \left(1-\frac{a_{M,\epsilon}(p)}{p}\right)^{-1},$$

and $a_{M,\epsilon}(p)$ was defined in Lemma 2.2.

Proof. See Section 7. □

Theorem 3.1 does not directly apply to (3-1) because of the normalization by $\langle f,f\rangle_N$.

We present a technique for removing the weights $\langle f,f\rangle_N$, which is a slight generalization of [Kowalski and Michel 1999, §3]. The idea for removing such weights first appeared in [Ram Murty 1995]. Let $\alpha = (\alpha_f)$ be a sequence of complex numbers indexed by

$$f \in \bigcup_{N\geq 1} \bigcup_{\epsilon\ (\mathrm{mod}\ N)} H^\star_\kappa(N,\epsilon).$$

Define the natural averaging operator

$$A[\alpha] = A_{N,\epsilon}[\alpha] = \sum_{f\in H^\star_\kappa(N,\epsilon)} \alpha_f.$$

Let

$$\omega_f = c_\kappa \frac{L_N(1, \mathrm{Ad}^2 f)}{\langle f, f\rangle_N}.$$

Then we define the *harmonic* averaging operator

$$A^h[\alpha] = A^h_{N,\epsilon}[\alpha] = \sum_{f \in H^\star_\kappa(N,\epsilon)} \omega_f \alpha_f.$$

The following proposition is a minor generalization of Proposition 2 of [Kowalski and Michel 1999]. It allows us to pass from natural averages of newforms to harmonic averages of newforms.

Proposition 3.2. *Let $\alpha = (\alpha_f)$ be a sequence of complex numbers indexed by $f \in H^\star_\kappa(N, \epsilon)$ running over all N and all ϵ. Suppose that for all $\varepsilon > 0$*

$$A^h[|\alpha_f|] \ll_\varepsilon (N\kappa)^\varepsilon \tag{3-3}$$

and

$$\max_{f \in H^\star_\kappa(N,\epsilon)} |\omega_f \alpha_f| \ll (N\kappa)^{-\delta+\varepsilon} \tag{3-4}$$

for some absolute $\delta > 0$. For any integer $r \geqslant 1$ write $x = (N\kappa)^{10/r}$. Then we have

$$A[\alpha_f] = \frac{\kappa - 1}{4\pi} \frac{\mathrm{Vol}\, X_0(N)}{\zeta^{(N)}(2)} (A^h[\omega_f(x)\alpha_f] + O_{\varepsilon,r}(x^{-\frac{\delta}{20}+\varepsilon} + (N\kappa)^{-1})).$$

Proof. See Section 6. □

One of the main ingredients in the proof of Proposition 3.2 is a large sieve inequality for the Dirichlet series coefficients of the automorphic adjoint square L-function $L(s, \mathrm{Ad}^2 \pi)$, see Proposition 6.2, which is a quotation of [Duke and Kowalski 2000, Corollary 6]. This inequality is only valid when the length of summation X satisfies $X \gg (N\kappa)^8$, which is far from the expected truth. Nonetheless, as of now it is the best available such inequality in the range of parameters of interest to us. The exponent $-\delta/20$ in Proposition 3.2 is optimized given the exponent 8 above, and any improvement over the result of Duke and Kowalski would lead to a corresponding improvement to the value $20 = 2(8 + 2)$.

We apply Proposition 3.2 with $\alpha_f = \overline{\lambda_f(m)}$ to (3-1) to get

$$\begin{aligned}\overline{\mathrm{Tr}(T'_m | S_\kappa(\Gamma_0(N), \epsilon))} &= \sum_{LM=N} d(L) \frac{\kappa - 1}{4\pi} \frac{\mathrm{Vol}\, X_0(M)}{\zeta^{(M)}(2)} A^h_{M,\epsilon}[\omega_f(x)\overline{\lambda_f(m)}] + O(\kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon)\\ &= \frac{\kappa - 1}{12} \sum_{LM=N} \frac{d(L)\psi(M)}{\zeta^{(M)}(2)} \sum_{\substack{n \leqslant x\\ (n,M)=1}} \frac{1}{n} \sum_{n=k^2\ell} \bar{\epsilon}(\ell) A^h_{M,\epsilon}[\overline{\lambda_f(m)}\lambda_f(\ell^2)]\\ &\qquad + O(\kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon). \end{aligned} \tag{3-5}$$

We are now ready to apply Theorem 3.1. We deduce a version of the newform formula for the harmonic averages $A^h[\overline{\lambda_f(m)}\lambda_f(n)]$ appearing in (3-5).

Lemma 3.3. *Let $\boldsymbol{c}_p(\epsilon)$ denote the exponent of the p-part of $\boldsymbol{c}(\epsilon)$. If $(mn, N) = 1$ then we have*

$$A^h_{N,\epsilon}[\overline{\lambda_f(m)}\lambda_f(n)] = \frac{1}{\psi(N)} \sum_{LM=N} \mu(L) M F(M,\epsilon) \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m, n\ell^2), \tag{3-6}$$

where

$$F(M,\epsilon) = \prod_{\substack{p \| M \\ \boldsymbol{c}_p(\epsilon)=1}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p^\alpha \| M \\ \alpha \geqslant 2 \\ \boldsymbol{c}_p(\epsilon)=\alpha}} \left(1 - \frac{1}{p}\right)^{-1}.$$

In particular, if $\epsilon = \epsilon_0$ is trivial we have

$$A^h_{N,\epsilon_0}[\overline{\lambda_f(m)}\lambda_f(n)] = \frac{1}{\psi(N)} \sum_{LM=N} \mu(L) M \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{1}{\ell} \Delta_{\kappa,M,\epsilon_0}(m, n\ell^2). \tag{3-7}$$

Note that formula (3-7) resembles closely the formula found in [Barrett et al. 2017, Proposition 4.1].

Proof. By the definition of $L_p(1, \mathrm{Ad}^2 f)$ and Theorem 3.1 we have

$$A^h[\overline{\lambda_f(m)}\lambda_f(n)] = \prod_{p \mid N} \left(1 - \frac{1}{p}\right) \left(1 - \frac{a_{N,\epsilon}(p)}{p}\right)^{-1} \sum_{LM=N} \mu(L) R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m, n\ell^2).$$

It suffices to show for any L, M that

$$\frac{\psi(LM)}{M} \prod_{p \mid LM} \left(1 - \frac{1}{p}\right) \left(1 - \frac{a_{LM,\epsilon}(p)}{p}\right)^{-1} R(M, L, \epsilon) = \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) F(M, \epsilon). \tag{3-8}$$

We may also assume that $c(\epsilon) \mid M$, since otherwise $\Delta_{\kappa,M,\epsilon}(m, n\ell^2) = 0$. Both sides of (3-8) are multiplicative, so it suffices to check the case $M = p^\alpha$ and $L = p^\beta$ for an arbitrary prime p. The following cases can be easily verified one-by-one:

- $\alpha \geqslant 2,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = \alpha,$
- $\alpha \geqslant 2,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) < \alpha,$
- $\alpha \geqslant 2,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = \alpha,$
- $\alpha \geqslant 2,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) < \alpha,$
- $\alpha = 1,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = 1,$
- $\alpha = 1,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 1,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = 1,$
- $\alpha = 1,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 0,\ \beta \geqslant 2,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 0,\ \beta = 1,$ and $\boldsymbol{c}_p(\epsilon) = 0.$ □

4. Analysis for $\Gamma_0(N)$

Now we put together (3-5), the newform formula (3-6), and the Petersson formula (3-2). By (3-5) and (3-6) we have

$$\overline{\mathrm{Tr}(T'_m \mid S_\kappa(\Gamma_0(N), \epsilon))} = A + E,$$

where for an integer $r \geqslant 1$ to be chosen later we set $x^r = (N\kappa)^{10}$ and have

$$A = \frac{\kappa-1}{12} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{\substack{k \leqslant x^{1/2} \\ (k,M)=1}} \frac{1}{k^2} \sum_{\substack{\ell \leqslant x/k^2 \\ (\ell,M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \times \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{\bar{\epsilon}(q)}{q} \Delta_{\kappa,W,\epsilon}(m, q^2\ell^2), \quad (4\text{-}1)$$

and E is the error term from (3-5) of size

$$E \ll_{r,\varepsilon} \kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon. \quad (4\text{-}2)$$

Applying (3-2) to A we get that

$$A = D + OD,$$

where D and OD are the contributions from the diagonal term and off-diagonal term of (3-2), respectively. We insert $\delta_{m=q^2\ell^2}\delta_{\boldsymbol{c}(\epsilon)\mid W}$ for $\Delta_{\kappa,W,\epsilon}(m, q^2\ell^2)$ in (4-1) to find

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac{1}{2}})}{m^{\frac{1}{2}}} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{\substack{k \leqslant x^{1/2}/m^{1/4} \\ (k,M)=1}} \frac{1}{k^2} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \delta_{\boldsymbol{c}(\epsilon)\mid W}.$$

Extending the sum over k to infinity we conclude that

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac{1}{2}})}{m^{\frac{1}{2}}} \sum_{LM=N} d(L) \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \delta_{\boldsymbol{c}(\epsilon)\mid W} + O_\varepsilon\left(\frac{\kappa N^{1+\varepsilon}}{x^{\frac{1}{2}} m^{\frac{1}{4}}} |\epsilon(m^{\frac{1}{2}})|\right).$$

By a tedious case check on prime powers we have

$$\psi(N)\delta_{\boldsymbol{c}(\epsilon)\mid N} = \sum_{LM=N} M F(M,\epsilon) \delta_{\boldsymbol{c}(\epsilon)\mid M} \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right).$$

Therefore the result of the diagonal contribution is

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac{1}{2}})}{m^{\frac{1}{2}}} \psi(N) + O\left(\frac{\kappa N^{1+\varepsilon}}{x^{\frac{1}{2}} m^{\frac{1}{4}}} |\epsilon(m^{\frac{1}{2}})|\right), \quad (4\text{-}3)$$

which matches what one finds directly from the identity contribution of the Eichler–Selberg trace formula.

Now we treat the off-diagonal terms. Let

$$B(Y,m,W) = \sum_{\substack{\ell \leqslant Y \\ (\ell,M)=1}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{\bar{\epsilon}(q\ell)}{q\ell} \sum_{c \equiv 0 \ (\mathrm{mod}\ W)} \frac{S_\epsilon(m, q^2\ell^2, c)}{c} J_{\kappa-1}\left(\frac{4\pi q\ell\sqrt{m}}{c}\right). \quad (4\text{-}4)$$

Then we have

$$OD = \frac{\kappa-1}{12} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{k \leqslant x^{1/2} \\ (k,M)=1}} \frac{1}{k^2} B\left(\frac{x}{k^2}, m, W\right).$$

Lemma 4.1. *Let d_3 denote the 3-divisor function. For any $m, n \geqslant 1$ we have*

$$\sum_{c \equiv 0 \pmod W} \frac{S_\epsilon(m,n,c)}{c} J_{\kappa-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \ll \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \prod_{p \mid \boldsymbol{c}(\epsilon)} p^{\frac{1}{4}} \frac{(m,n,W)^{\frac{1}{2}} d_3((m,n)) d(W)}{W \kappa^{\frac{5}{6}}} \left(\frac{mn}{\sqrt{mn}+\kappa W}\right)^{\frac{1}{2}} \log 2mn.$$

Proof. The proof is identical to [Iwaniec et al. 2000, Corollary 2.2] but with the following bound on the Kloosterman sum in lieu of the standard bound without nebentype character.

Lemma 4.2. *For integers $c \in N\mathbb{Z}$ and $a, b \in \mathbb{Z}$ with $c \neq 0$ and $\boldsymbol{c}(\epsilon) \mid N$, we have the estimate*

$$|S_\epsilon(a,b;c)| \leq d(c)\,(a,b,c)^{\frac{1}{2}}\, c^{\frac{1}{2}}\, \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}}.$$

Proof. See [Knightly and Li 2013, Theorem 9.2]. □

Applying Lemma 4.1 and estimating sums by integrals we find

$$B(Y,m,W) \ll \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} \frac{d(W) m^{\frac{1}{4}} Y^{\frac{1}{2}}}{W \kappa^{\frac{5}{6}}} \log mY;$$

hence one estimates that

$$OD \ll_\varepsilon \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} x^{\frac{1}{2}} \kappa^{\frac{1}{6}} m^{\frac{1}{4}} N^{\varepsilon} \log mx.$$

We have $\operatorname{Tr}(T'_m | S_\kappa(\Gamma_0(N,\epsilon))) = \overline{D} + \overline{OD} + \overline{E}$, and so collecting error terms we obtain

$$\operatorname{Tr}(T'_m | S_\kappa(\Gamma_0(N,\epsilon))) = \frac{\kappa-1}{12} \frac{\epsilon(m^{\frac{1}{2}})}{m^{\frac{1}{2}}} \psi(N) + O_\varepsilon\big(\boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} x^{\frac{1}{2}} \kappa^{\frac{1}{6}} m^{\frac{1}{4}} N^{\varepsilon} \log mx + \kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^{\varepsilon}\big). \tag{4-5}$$

We now optimize the value of r. By [Goldfeld et al. 1994; Banks 1997], the exponent $\delta = 1$ is admissible. The error in (4-5) is minimized when

$$x^{\frac{11}{20}} = \frac{N \kappa^{\frac{5}{6}}}{m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}}}.$$

Let us assume that there is some $\eta > 0$ such that

$$m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} \ll (N \kappa^{\frac{5}{6}})^{1-\eta}. \tag{4-6}$$

We choose $r \geqslant 1$ to be the nearest integer to

$$\frac{11}{2}\left(1 + \frac{\log(\kappa^{\frac{1}{6}} m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}})}{\log(N\kappa^{\frac{5}{6}}) - \log(m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}})}\right),$$

which by (4-6) is then bounded above uniformly in terms of $\eta > 0$ only.

5. Analysis for $\Gamma(M,N)$

Recall from (1-5) that

$$\operatorname{Tr}(\langle\bar{d}\rangle T'_m|S_\kappa(\Gamma(M,N))) = \sum_{\epsilon\ (\mathrm{mod}\ N)} \bar{\epsilon}(d)\operatorname{Tr}(T'_m|S_\kappa(\Gamma_0(MN),\epsilon)),$$

and that in Section 4 we decomposed the interior of this as

$$\overline{\operatorname{Tr}(T'_m|S_\kappa(\Gamma_0(MN),\epsilon))} = D + OD + E.$$

Summing the formula (4-3) for D and (4-2) for E trivially over characters $\epsilon\ (\mathrm{mod}\ N)$ we get

$$\begin{aligned}\operatorname{Tr}(\langle\bar{d}\rangle T'_m|S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24}m^{-\frac{1}{2}}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1)+(-1)^\kappa\delta_N(m^{\frac{1}{2}}d,-1)\big)\\ + \overline{OD^*} + O_{\eta,\varepsilon}(\kappa(MN^2)^{1+\varepsilon}x^{-\frac{\delta}{20}+\varepsilon}+N(MN)^\varepsilon),\end{aligned} \tag{5-1}$$

where $x^r=(MN\kappa)^{10}$, r is a parameter to be chosen later, and

$$OD^* = \sum_{\substack{\epsilon\ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa}} \epsilon(d)OD.$$

Let

$$B^*(Y,m,W) = \sum_{\substack{\epsilon\ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa}} \epsilon(d)F(W,\epsilon) \sum_{\substack{(\ell,K)=1\\ \ell\leqslant Y}} \sum_{\substack{q\,|\,Q^\infty\\ (q,W)=1}} \frac{\overline{\epsilon(q\ell)}}{q\ell} \sum_{c\equiv 0\ (\mathrm{mod}\ W)} \frac{S_\epsilon(m,q^2\ell^2,c)}{c} J_{\kappa-1}\left(\frac{4\pi\ell q\sqrt{m}}{c}\right),$$

so that we have

$$OD^* = \frac{\kappa-1}{12}\sum_{LK=MN}\frac{d(L)}{\zeta^{(K)}(2)}\sum_{\substack{k\leqslant x^{1/2}\\ (k,K)=1}}\frac{1}{k^2}\sum_{WQ=K}\mu(Q)W\prod_{p^2\,|\,W}\left(1-\frac{1}{p^2}\right)B^*\left(\frac{x}{k^2},m,W\right). \tag{5-2}$$

We would like to utilize the orthogonality of characters over $\epsilon\ (\mathrm{mod}\ N)$. To implement this, we now refresh the notation. Suppose $W,N\geqslant 1$ are integers such that $W\,|\,N^2$. For $a,b,d,\kappa\in\mathbb{Z}$ and $1\leqslant c\equiv 0\ (\mathrm{mod}\ W)$ define

$$T_W(a,b,c) := \sum_{\substack{\epsilon\ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa\\ c(\epsilon)\,|\,W}} \epsilon(d)\bar{\epsilon}(b)F(W,\epsilon)S_\epsilon(a,b,c).$$

With this notation, we have

$$B^*(Y,m,W) = \sum_{\substack{(\ell,K)=1\\ \ell\leqslant Y}}\sum_{\substack{q\,|\,Q^\infty\\ (q,W)=1}}\frac{1}{q\ell}\sum_{c\equiv 0\ (\mathrm{mod}\ W)}\frac{T_W(m,q^2\ell^2,c)}{c}J_{\kappa-1}\left(\frac{4\pi\ell q\sqrt{m}}{c}\right). \tag{5-3}$$

We can derive a bound on T_W by appealing to the Weil bound for Kloosterman sums.

Lemma 5.1. *Suppose $W, N \geqslant 1$ such that $W \mid N^2$, $a, b, d, \kappa \in \mathbb{Z}$ such that $(b, W) = 1$, $(d, N) = 1$, and $1 \leqslant c \equiv 0 \pmod{W}$. We factor $c = c_1 c_2$ with $c_1 \mid W^\infty$ and $(c_2, W) = 1$. Then*

$$|T_W(a, b, c)| \leqslant \psi(c_1) d(c_2)(a, b, c_2)^{\frac{1}{2}} c_2^{\frac{1}{2}}.$$

Proof. Consider the sum

$$T'_W(a, b, c) := \sum_{\substack{\epsilon \pmod{N} \\ \boldsymbol{c}(\epsilon) \mid W}} \epsilon(d)\bar{\epsilon}(b) F(W, \epsilon) S_\epsilon(a, b, c),$$

which is a minor variation of $T_W(a, b, c)$ omitting the global condition $\epsilon(-1) = (-1)^\kappa$. We first consider the sum T'_W locally, returning to T_W at the end of the proof. Let $\alpha, \beta, \gamma \geqslant 0$ such that $\alpha \leqslant \gamma$, $\alpha \leqslant 2\beta$, $(d, p^\beta) = (b, p^\alpha) = 1$, and consider $T'_{p^\alpha}(a, b, p^\gamma)$. Let

$$I(\alpha, \beta) := \sum_{\epsilon \pmod{p^\beta}} \epsilon(dx)\bar{\epsilon}(b)\delta_{\boldsymbol{c}_p(\epsilon) \leqslant \alpha} \begin{cases} 1 + \frac{1}{p} & \text{if } \boldsymbol{c}_p(\epsilon) = \alpha = 1, \\ \left(1 - \frac{1}{p}\right)^{-1} & \text{if } \boldsymbol{c}_p(\epsilon) = \alpha \geqslant 2, \\ 1 & \text{else.} \end{cases}$$

By opening the Kloosterman sum and exchanging order of summation we have

$$T'_{p^\alpha}(a, b, p^\gamma) = \sideset{}{^*}\sum_{x \pmod{p^\gamma}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) I(\alpha, \beta). \tag{5-4}$$

Next we break into four cases:

(1) $\alpha > \beta$.

(2) $0 = \alpha \leqslant \beta$.

(3) $1 = \alpha \leqslant \beta$.

(4) $2 \leqslant \alpha \leqslant \beta$.

Recall the orthogonality relation

$$\sum_{\epsilon \pmod{n}} \epsilon(a)\bar{\epsilon}(b) = \varphi(n)\delta_n(a, b)$$

and the almost-orthogonality relation (see, e.g., [Heath-Brown 1981, Section 2])

$$\sum_{\boldsymbol{c}(\epsilon) = c} \epsilon(a)\bar{\epsilon}(b) = \sum_{\delta \mid (a-b, c)} \varphi(\delta)\mu\left(\frac{c}{\delta}\right).$$

We apply these to evaluate $I(\alpha, \beta)$ in cases (1)–(4). We find

$$I(\alpha, \beta) = \begin{cases} \varphi(p^\beta)\delta_{p^\beta}(xd, b) & \text{if } \alpha > \beta, \\ 1 & \text{if } 0 = \alpha \leqslant \beta, \\ \varphi(p)\delta_p(xd, b) + \frac{1}{p}\sum_{\delta \mid (p, xd-b)} \varphi(\delta)\mu(p^\gamma/\delta) & \text{if } 1 = \alpha \leqslant \beta, \\ \varphi(p^\alpha)\delta_{p^\alpha}(xd, b) + \frac{1}{p-1}\sum_{\delta \mid (p^\alpha, xd-b)} \varphi(\delta)\mu(p^\gamma/\delta) & \text{if } 1 = \alpha \leqslant \beta. \end{cases} \tag{5-5}$$

Recall that $(d, p^\beta) = 1$, so that $d^{-1} \pmod{p^\beta}$ (or $\pmod{p^\gamma}$ in cases (3) and (4)) exists. Inserting (5-5) to (5-4), we find the following.

Case (1): $\beta < \alpha$.

$$T'_{p^\alpha}(a, b, p^\gamma) = \varphi(p^\beta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p^\beta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Case (2): $\beta \geqslant \alpha = 0$.

$$T'_1(a, b, p^\gamma) = \sideset{}{^*}\sum_{x \pmod{p^\gamma}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) = S(a, b, p^\gamma).$$

Case (3): $\beta \geqslant \alpha = 1$.

$$T'_p(a, b, p^\gamma) = \varphi(p) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) + \frac{1}{p} \sum_{\delta \mid p} \varphi(\delta)\mu(p/\delta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{\delta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Case (4): $\beta \geqslant \alpha \geqslant 2$.

$$T'_{p^\alpha}(a, b, p^\gamma) = \varphi(p^\alpha) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p^\alpha}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) + \frac{1}{p-1} \sum_{\delta \mid p^\alpha} \varphi(\delta)\mu(p^\alpha/\delta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{\delta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Using the Weil bound for Kloosterman sums and trivial bounds, we find for all integers a, b, and nonzero integers $0 \leqslant i \leqslant j$, and $(y, p) = 1$ we have

$$\left| \sideset{}{^*}\sum_{\substack{x \pmod{p^j} \\ x \equiv y \pmod{p^i}}} e\left(\frac{ax + b^2\bar{x}}{p^\gamma}\right) \right| \leqslant \begin{cases} d(p^j)(a, b^2, p^j)^{\frac{1}{2}} \sqrt{p^j} & \text{if } i = 0, \\ p^{j-i} & \text{else.} \end{cases} \tag{5-6}$$

Applying (5-6) to the various cases above, we find that for cases (1), (3), and (4), i.e., when $\alpha > 0$, we have the bound

$$|T'_{p^\alpha}(a, b, p^\gamma)| \leqslant \psi(p^\gamma). \tag{5-7}$$

In case (2), i.e., when $\alpha = 0$, we have

$$|T'_{p^\alpha}(a, b, p^\gamma)| \leqslant d(p^\gamma)(a, b, p^\gamma)^{\frac{1}{2}} p^{\frac{\gamma}{2}}. \tag{5-8}$$

Thus the estimation of $T'_{p^\alpha}(a, b, p^\gamma)$ is finished.

Now we return to the case of $T_W(a, b, c)$. We have

$$T_W(a, b, c) = \tfrac{1}{2} T'_W(a, b, c) + \tfrac{1}{2}(-1)^\kappa T'_W(a, -b, c),$$

so it suffices to establish the bound stated in the lemma for $T'_W(a,b,c)$. We have that $T'_W(a,b,c)$ is twisted multiplicative, i.e., we have a factorization

$$T'_W(a,b,c) = \prod_{\substack{p^\alpha \| W \\ p^\gamma \| c}} T'_{p^\alpha}(a\overline{cp^{-\gamma}}, b\overline{cp^{-\gamma}}, p^\gamma). \tag{5-9}$$

Bounding the left-hand side of (5-9) using (5-7) and (5-8), we conclude the proof of the lemma. $\square$

Applying Lemma 5.1 to (5-3) we get

$$B^*(Y,m,W) \leqslant \sum_{\substack{(\ell,K)=1 \\ \ell \leqslant Y}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{1}{q\ell} \sum_{W \mid c_1 \mid W^\infty} \frac{\psi(c_1)}{c_1} \sum_{(c_2,W)=1} \frac{d(c_2)(m,q^2\ell^2,c_2)^{\frac{1}{2}}}{\sqrt{c_2}} \left| J_{\kappa-1}\left(\frac{4\pi q\ell\sqrt{m}}{c_1c_2}\right)\right|.$$

Again following closely the proof of [Iwaniec et al. 2000, Corollary 2.2] we have

$$\sum_{(c_2,W)=1} \frac{d(c_2)(a,b^2,c_2)^{\frac{1}{2}}}{\sqrt{c_2}} \left| J_{\kappa-1}\left(\frac{4\pi b\sqrt{a}}{c_1c_2}\right)\right| \ll \frac{d_3((a,b^2))}{\kappa^{\frac{5}{6}}\sqrt{c_1}} \left(\frac{b^2a}{b\sqrt{a}+c_1\kappa}\right)^{\frac{1}{2}} \log 2b^2a.$$

We have moreover that

$$\sum_{W \mid c_1 \mid W^\infty} \frac{1}{\sqrt{c_1}} \frac{1}{(b\sqrt{a}+c_1\kappa)^{\frac{1}{2}}} \leqslant \frac{2}{W^{\frac{1}{2}} b^{\frac{1}{2}} a^{\frac{1}{4}}}.$$

These last two estimations lead to

$$\begin{aligned} B^*(Y,m,W) &\ll \frac{m^{\frac{1}{4}}}{\kappa^{\frac{5}{6}}} \frac{\psi(W)}{W^{\frac{3}{2}}} \sum_{\substack{(\ell,K)=1 \\ \ell \leqslant Y}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{d_3((m,q^2\ell^2))}{\sqrt{q\ell}} \log(2q^2\ell^2 m) \\ &\ll \frac{m^{\frac{1}{4}} Y^{\frac{1}{2}} (\log Y)^3 \log 2m}{\kappa^{\frac{5}{6}}} \frac{\psi(W)}{W^{\frac{3}{2}}}. \end{aligned}$$

Inserting this into (5-2) we get

$$OD^* \ll \kappa^{\frac{1}{6}} M N^{\frac{1}{2}+\varepsilon} x^{\frac{1}{2}} (\log x)^3 m^{\frac{1}{4}} \log 2m,$$

and inserting this into (5-1) we conclude that

$$\begin{aligned} &\mathrm{Tr}(\langle \bar{d} \rangle T'_m | S_\kappa(\Gamma(M,N))) \\ &\quad = \frac{\kappa-1}{24} m^{-\frac{1}{2}} \varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1) + (-1)^\kappa \delta_N(m^{\frac{1}{2}}d,-1)\big) \\ &\qquad + O_{\eta,\varepsilon}\big(\kappa^{\frac{1}{6}} M N^{\frac{1}{2}} x^{\frac{1}{2}+\varepsilon} m^{\frac{1}{4}} \log 2m + \kappa(MN^2)^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N(MN)^\varepsilon\big). \end{aligned} \tag{5-10}$$

Now we optimize the value of r. The error term is minimized when

$$x^{\frac{11}{20}} = \frac{N^{\frac{3}{2}}\kappa^{\frac{5}{6}}}{m^{\frac{1}{4}}}.$$

Let us assume that there is some $\eta > 0$ such that

$$m^{\frac{1}{4}} \ll (N^{\frac{3}{2}}\kappa^{\frac{5}{6}})^{1-\eta}.$$

We choose $r \geqslant 1$ to be the nearest integer to

$$\frac{11}{2}\left(\frac{\log MN\kappa}{\log N^{\frac{3}{2}}\kappa^{\frac{5}{6}} - \log m^{\frac{1}{4}}}\right),$$

which is then bounded above uniformly in terms of $\eta > 0$ only.

6. Proof of Proposition 3.2

Proof. We have by Lemma 2.1 that

$$A[\alpha_f] = \frac{\kappa-1}{4\pi}\frac{\operatorname{Vol} X_0(N)}{\zeta^{(N)}(2)} \sum_{f \in H_\kappa^\star(N,\epsilon)} \omega_f \alpha_f L^{(N)}(1, \operatorname{Ad}^2 f) = \frac{\kappa-1}{4\pi}\frac{\operatorname{Vol} X_0(N)}{\zeta^{(N)}(2)} A^h[\alpha_f L^{(N)}(1, \operatorname{Ad}^2 f)].$$

Recall that we have set $\varrho_f(n)$ to be the Dirichlet series coefficients of $L^{(N)}(s, \operatorname{Ad}^2 f)$, along with

$$\omega_f(x) = \sum_{n \leqslant x} \frac{\varrho_f(n)}{n} \quad \text{and} \quad \omega_f(x,y) = \sum_{x < n \leqslant y} \frac{\varrho_f(n)}{n}.$$

Lemma 6.1. *We have*

$$L^{(N)}(1, \operatorname{Ad}^2 f) = \omega_f(x) + \omega_f(x,y) + O_\varepsilon((N\kappa)^{\frac{1}{2}} y^{-\frac{1}{2}+\varepsilon}).$$

Assuming the generalized Lindelöf hypothesis, the $(N\kappa)^{1/2}$ *can be reduced to* $(N\kappa)^\varepsilon$.

Proof (*sketch*). For $c, T, y > 0$, we apply Perron's formula (see, e.g., [Davenport 2000, p. 105]) to calculate $\omega_f(y)$, finding

$$\omega_f(y) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} L^{(N)}(1+s, \operatorname{Ad}^2 f)\frac{y^s}{s}\,ds + O\biggl(y^c \sum_{n \geqslant 1} \frac{\varrho_f(n)}{n^{1+c}} \min(1, T^{-1}|\log y/n|^{-1})\biggr).$$

We shift the contour to $\operatorname{Re}(s) = -2$ to get

$$\omega_f(y) = L^{(N)}(1, \operatorname{Ad}^2 f) + \frac{1}{2\pi i}\biggl(\int_{c-iT}^{-2-iT} + \int_{-2-iT}^{-2+iT} + \int_{-2+iT}^{c+iT}\biggr) L^{(N)}(1+s, \operatorname{Ad}^2 f)\frac{y^s}{s}\,ds \\ + O\biggl(y^c \sum_{n \geqslant 1} \frac{\varrho_f(n)}{n^{1+c}} \min(1, T^{-1}|\log y/n|^{-1})\biggr). \quad (6\text{-}1)$$

By an inspection of the functional equation for $L(s, f \otimes \bar{f})$ found in [Li 1979, Example 1], we have the convexity bound (see, e.g., [Iwaniec and Kowalski 2004, (5.20)])

$$L^{(N)}(s, \operatorname{Ad}^2 f) \ll [(\kappa N)^2(1+|t|)^3]^{\frac{1-\sigma}{2}+\varepsilon}, \quad (6\text{-}2)$$

where $s = \sigma + it$, valid for $\sigma \leqslant 1$. Choosing $c = \varepsilon$, $T = (N\kappa)^{-\frac{1}{2}} y^{\frac{1}{2}+\varepsilon}$, and estimating all of the terms in (6-1) directly, one finds the estimate in the statement of the lemma.

If one assumes the generalized Lindelöf hypothesis in place of (6-2), then we shift the contour to $\mathrm{Re}(s) = -\frac{1}{2}$ instead of -2 and follow the same steps. □

By Lemma 6.1 we have

$$A[\alpha_f] = \frac{\kappa - 1}{4\pi} \frac{\mathrm{Vol}\, X_0(N)}{\zeta^{(N)}(2)} \bigl(A^h[\omega_f(x)\alpha_f] + A^h[\omega_f(x, y)\alpha_f] + O((N\kappa)^{\frac{1}{2}} y^{-\frac{1}{2}+\varepsilon} A^h[|\alpha_f|]) \bigr). \tag{6-3}$$

By the hypothesis (3-3) we have $A^h[|\alpha_f|] \ll_\varepsilon (N\kappa)^\varepsilon$, and so taking $y = (N\kappa)^{3+\varepsilon}$, we find that the O term in (6-3) is $\ll (N\kappa)^{-1}$.

Next we consider the second term and treat it using the following large sieve inequality. This is a slight variation on Corollary 6 of [Duke and Kowalski 2000]; see also [Kowalski and Michel 1999, Proposition 1]. Let $\lambda_f^{(2)}(n)$ be the Dirichlet series coefficients of the automorphic adjoint-square L-function $L(s, \mathrm{Ad}^2 \pi)$, where f is a newform for the representation π. If $(n, N) = 1$ then we have $\lambda_f^{(2)}(n) = \varrho_f(n)$.

Proposition 6.2. *Let $X \geqslant (N\kappa)^8$. We have for all $\varepsilon > 0$ that*

$$\sum_{f \in H_\kappa^\star(N,\epsilon)} \Bigl| \sum_{n \leqslant X} a_n \lambda_f^{(2)}(n) \Bigr|^2 \ll_\varepsilon X^{1+\varepsilon} \sum_{n \leqslant X} |a_n|^2 \tag{6-4}$$

for any finite family $(a_n)_{1 \leqslant n \leqslant X}$ of complex numbers, where the constant depends only on ε.

By following closely [Kowalski and Michel 1999, §3.3] one deduces from Proposition 6.2 the following lemma.

Lemma 6.3. *Let $r \geqslant 1$ be an integer such that $x^r \geqslant (N\kappa)^{10}$. Then for all $\varepsilon > 0$ we have*

$$A[\omega_f(x, y)^{2r}] \ll_{r,\varepsilon} (N\kappa)^\varepsilon,$$

where the implied constant depends only on r and ε.

Proof. It suffices to replace instances of $\lambda_f(n^2)$ in [Kowalski and Michel 1999, Lemma 3] by $\bar{\epsilon}(n)\lambda_f(n^2)$ and to use (2-3) and (2-4) in the place of (15) and (16) of [loc. cit.]. □

We now can give an estimate for the second term of (6-3). We use Hölder's inequality to separate $\omega_f(x, y)$ from α_f, and Lemma 6.3 to handle the the sum involving $\omega_f(x, y)$. Precisely, let s be defined by $(2r)^{-1} + s^{-1} = 1$. Applying Hölder's inequality we find for any integer $r \geqslant 1$ that

$$\begin{aligned} A^h[\omega_f(x, y)\alpha_f] = \sum_{f \in H_\kappa^\star} \omega_f \omega_f(x, y)\alpha_f &\leqslant A[\omega_f(x, y)^{2r}]^{\frac{1}{2r}} \Bigl(\sum_{f \in H_\kappa^\star(N,\epsilon)} (\omega_f |\alpha_f|)^s \Bigr)^{\frac{1}{s}} \\ &\leqslant A^{\frac{1}{2r}} A[\omega_f(x, y)^{2r}]^{\frac{1}{2r}} A^h[|\alpha_f|]^{\frac{1}{s}}, \end{aligned}$$

where

$$A = \max_{f \in H_\kappa^\star(N,\epsilon)} \omega_f |\alpha_f| \ll_\varepsilon (N\kappa)^{-\delta+\varepsilon}$$

by hypothesis (3-4). Suppose now that r is sufficiently large so that $x^r \geqslant (N\kappa)^{10}$. Then Lemma 6.3 applies, and we have

$$A[\omega_f(x,y)^{2r}]^{\frac{1}{2r}} \ll_{r,\varepsilon} (N\kappa)^{\varepsilon}.$$

Lastly, by hypothesis (3-3) we have

$$A^h[|\alpha_f|]^{\frac{1}{s}} \ll_{\varepsilon} (N\kappa)^{\varepsilon}.$$

Putting these estimates together, we find that $A^h[\omega_f(x,y)\alpha_f] \ll_{r,\varepsilon} (N\kappa)^{-\frac{\delta}{2r}+\varepsilon}$, and so derive the bound claimed in Proposition 3.2. □

7. Proof of Theorem 3.1

Proof. The strategy of the proof is to pick an orthogonal basis for $S_\kappa(\Gamma_0(N),\epsilon)$ and compute the Fourier coefficients of basis elements explicitly. For f a modular function of weight κ, we define $f_{|d}(z) = d^{\frac{\kappa}{2}} f(dz)$. Atkin–Lehner theory gives an orthogonal direct sum decomposition

$$S_\kappa(\Gamma_0(N),\epsilon) = \bigoplus_{LM=N} \bigoplus_{f\in H_\kappa^\star(M,\epsilon)} S_\kappa(L; f,\epsilon),$$

where $S_\kappa(L; f,\epsilon) = \operatorname{span}\{f_{|\ell} : \ell \mid L\}$ is called an oldclass. Note that the inner sum is $\{0\}$ unless $\boldsymbol{c}(\epsilon) \mid M$, so we may assume this for the remainder of the proof.

To pick an orthogonal basis for $S_\kappa(\Gamma_0(N),\epsilon)$ it then suffices to pick a orthonormal basis for each oldclass $S_\kappa(L; f,\epsilon)$. We use a basis for the oldclasses first due to [Schulze-Pillot and Yenirce 2018, Theorem 8]. The basis constructed by Schulze-Pillot and Yenirce is the same as the one found by Rouymi [2011] in the case of prime power level and trivial nebentypus and Ng [2012] in the case of arbitrary level and trivial nebentypus; see also [Blomer and Milićević 2015, Chapter 5] and [Humphries 2018, Lemma 3.15]. Each of these preceding works used the Rankin–Selberg method to compute inner products and orthonormalize the oldclasses. Schulze-Pillot and Yenirce however took a different and simpler path, using the trace operator to compute the inner products.

Let $f \in H^\star(M,\epsilon)$. For integers $d \mid g$ one defines a joint multiplicative function $\xi_g(d)$. On prime powers, $\xi_g(d)$ is given for $\nu \geqslant 2$ by

$$\xi_1(1) = 1, \quad \xi_{p^\nu}(p^\nu) = \left(1 - \frac{|\lambda_f(p)|^2}{p(1+\frac{\varepsilon_{0,M}(p)}{p})^2}\right)^{-\frac{1}{2}} \left(1 - \frac{\varepsilon_{0,M}(p)^2}{p^2}\right)^{-\frac{1}{2}},$$

$$\xi_p(p) = \left(1 - \frac{|\lambda_f(p)|^2}{p(1+\frac{\epsilon_{0,M}(p)}{p})^2}\right)^{-\frac{1}{2}}, \quad \xi_{p^\nu}(p^{\nu-1}) = \frac{-\overline{\lambda_f(p)}}{\sqrt{p}} \xi_{p^\nu}(p^\nu),$$

$$\xi_p(1) = \frac{-\overline{\lambda_f(p)}}{\sqrt{p}(1+\epsilon_{0,M}(p)/p)} \xi_p(p), \quad \xi_{p^\nu}(p^{\nu-2}) = \frac{\overline{\epsilon(p)}}{p} \xi_{p^\nu}(p^\nu),$$

and $\xi_{p^a}(p^b) = 0$ in all other cases.

Proposition 7.1 [Schulze-Pillot and Yenirce 2018, Theorem 9]. *Let $M \mid N$ and let $f \in H_\kappa^\star(M,\epsilon)$. The set of functions*

$$\{f^{(g)}(z) = \textstyle\sum_{d \mid g} \xi_g(d) d^{\frac{\kappa}{2}} f(dz) : g \mid L\}$$

is an orthogonal basis for $S_k(L; f, \epsilon)$. In fact, if f is $L^2(\Gamma_0(N)\backslash\mathcal{H})$-normalized, then the above set is in fact orthonormal.

Now that we have an orthonormal basis for $S_\kappa(\Gamma_0(N),\epsilon)$, we follow Barrett, Burkhardt, DeWitt, Dorward, and Miller [Barrett et al. 2017] to derive the Petersson formula for newforms Theorem 3.1.

Let $f \in H_\kappa^\star(M,\epsilon)$ have Fourier coefficients $a_f(n)$ and be normalized so that $a_f(1) = 1$. Of course $f(z)/\|f\|_N$ is $L^2(\Gamma_0(N)\backslash\mathcal{H})$-normalized, so using the basis in Proposition 7.1 we have

$$\begin{aligned}\Delta_{\kappa,N,\epsilon}(m,n) &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{g \in \mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_g(n)\overline{b_g(m)} \\ &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{LM=N} \sum_{f \in H_\kappa^\star(m,\epsilon)} \frac{1}{\langle f, f\rangle_N} \sum_{g \mid L} \overline{a_{f^{(g)}}(m)} a_{f^{(g)}}(n). \end{aligned} \tag{7-1}$$

By the definition of $f^{(g)}$ we have

$$a_{f^{(g)}}(n) = \sum_{d \mid (g,n)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{n}{d}\right),$$

which are now expressible in terms of Hecke eigenvalues $\lambda_f(n)$ normalized so that $|\lambda_f(n)| \leqslant d(n)$. We have then that

$$\begin{aligned}&\Delta_{\kappa,N,\epsilon}(m,n) \\ &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{LM=N} \sum_{f \in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{n}{d}\right)\right) \\ &= c_\kappa \sum_{LM=N} \sum_{f \in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{n}{d}\right)\right) \\ &= c_\kappa \sum_{LM=N} \sum_{f \in H_\kappa^\star(N,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \Xi_g(m,n,f),\end{aligned}$$

where we have set

$$\Xi_g(m,n,f) = \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{n}{d}\right)\right)$$

for $g \mid L \mid N$.

Now suppose that $(d_1, d_2) = 1$ and $d_1 d_2 \mid m$. Then by Hecke multiplicativity we have

$$\lambda_f\left(\frac{m}{d_1}\right)\lambda_f\left(\frac{m}{d_2}\right) = \lambda_f(m)\lambda_f\left(\frac{m}{d_1 d_2}\right),$$

so that for $(g_1, g_2) = 1$ we have

$$\Xi_{g_1}(m,n,f)\Xi_{g_2}(m,n,f) = \overline{\lambda_f(m)}\lambda_f(n)\Xi_{g_1g_2}(m,n,f).$$

Therefore

$$\Delta_{\kappa,N,\epsilon}(m,n) = c_\kappa \sum_{LM=N} \sum_{f\in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} (\overline{\lambda_f(m)}\lambda_f(n))^{1-\omega(L)} \prod_{p^\alpha || L} \Bigg(\sum_{d\,|\,p^\alpha} \Xi_d(m,n,f)\Bigg),$$

where $\omega(n)$ is the number of distinct prime factors of n. Let

$$V_{p^\alpha}(m,n,f) = \sum_{d\,|\,p^\alpha} \Xi_d(m,n,f) = (1*\Xi)_{p^\alpha}(m,n,f),$$

where $*$ denotes Dirichlet convolution. We suppose now that $(m,n,N) = 1$ and calculate.

Lemma 7.2 [Barrett et al. 2017, Appendix A]. *If $(m,n,N) = 1$ then we have*

$$\begin{aligned} V_{p^\alpha}(m,n,f) =& \overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2) \\ &+\delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}(\overline{\xi_p(p)}\xi_p(1)+\overline{\xi_{p^2}(p)}\xi_{p^2}(1)) \\ &+\delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}(\overline{\xi_p(1)}\xi_p(p)+\overline{\xi_{p^2}(1)}\xi_{p^2}(p)) \\ &+\delta_{p^2|m}\overline{\lambda_f(m/p^2)}\lambda_f(n)p\overline{\xi_{p^2}(p^2)}\xi_{p^2}(1)+\delta_{p^2|n}\overline{\lambda_f(m)}\lambda_f(n/p^2)p\overline{\xi_{p^2}(1)}\xi_{p^2}(p^2), \end{aligned}$$

if $\alpha \geqslant 2$ and

$$\begin{aligned} V_{p^\alpha}(m,n,f) = &\overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2) \\ &+\delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}\overline{\xi_p(p)}\xi_p(1) \\ &+\delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}\overline{\xi_p(1)}\xi_p(p), \end{aligned}$$

if $\alpha = 1$.

Proof. We actually have if $\alpha \geqslant 2$ that

$$V_{p^\alpha}(m,n,f) = \Xi_1(m,n,f)+\Xi_p(m,n,f)+\Xi_{p^2}(m,n,f).$$

The other summands vanish because by our assumption $(m,n,N) = 1$, since if $p\,|\,m$ then $p \nmid n$ because $p\,|\,N$. So each p divides either m or n but never both. Then, we have $\xi_{p^\beta}(1) = 0$ for $\beta \geqslant 3$. In fact, even more terms vanish. We have

$$\begin{aligned} V_{p^\alpha}(m,n,f) =& \overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2) \\ &+\delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}(\overline{\xi_p(p)}\xi_p(1)+\overline{\xi_{p^2}(p)}\xi_{p^2}(1)) \\ &+\delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}(\overline{\xi_p(1)}\xi_p(p)+\overline{\xi_{p^2}(1)}\xi_{p^2}(p)) \\ &+\delta_{p^2|m}\overline{\lambda_f(m/p^2)}\lambda_f(n)p\overline{\xi_{p^2}(p^2)}\xi_{p^2}(1)+\delta_{p^2|n}\overline{\lambda_f(m)}\lambda_f(n/p^2)p\overline{\xi_{p^2}(1)}\xi_{p^2}(p^2). \end{aligned}$$

Inserting the formulas for ξ, we complete the proof. The formula for the $\alpha = 1$ case is even simpler as we can drop the p^2 terms. ☐

Recall we write $ML = N$ and $f \in H_\kappa^\star(M,\epsilon)$.

Lemma 7.3. *If* $(m,N) = 1$ *and* $(n,N) = 1$ *then we have*

$$(\overline{\lambda_f(m)}\lambda_f(n))^{1-\omega(L)} \prod_{p^\alpha \| L} V_{p^\alpha}(m,n,f) = \overline{\lambda_f(m)}\lambda_f(n) \prod_{p\|L}(1+|\xi_p(1)|^2) \prod_{p^2|L}(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2).$$

Proof. Note that the conditions $(m,N) = 1$ and $(n,N) = 1$ imply that $p \nmid m$ and $p \nmid n$. So the formula above follows immediately from the formulas in Lemma 7.2. □

One has that

$$\|f\|_N^2 = \frac{\psi(N)}{\psi(M)}\|f\|_M^2$$

since $f \in H_\kappa^\star(M,\epsilon)$. Thus

$$\begin{aligned}\Delta_{\kappa,N,\epsilon}(m,n) \\ &= c_\kappa \sum_{LM=N} \frac{\psi(M)}{\psi(N)} \sum_{f\in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_M^2}\overline{\lambda_f(m)}\lambda_f(n) \prod_{p\|L}(1+|\xi_p(1)|^2) \prod_{p^2|L}(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2).\end{aligned}$$

Next we insert the definitions of the ξ functions. Let

$$r_f(p) = 1 - \frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2},$$

so

$$r_f(p)^{-1} = 1 + \frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2} + \left(\frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2}\right)^2 + \cdots,$$

where $\epsilon_{0,M}$ denotes the trivial character modulo M. Observe that

$$1+|\xi_p(1)|^2 = r_f(p)^{-1}$$

and

$$1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2 = r_f(p)^{-1}\left(1-\frac{\epsilon_{0,M}(p)}{p^2}\right)^{-1}.$$

Then we get

$$\Delta_{\kappa,N,\epsilon}(m,n) = c_\kappa \sum_{LM=N} \frac{\psi(M)}{\psi(N)} \prod_{p^2|L}\left(1-\frac{\epsilon_{0,M}(p)}{p^2}\right)^{-1} \sum_{f\in H_\kappa^\star(M,\epsilon)} \frac{\overline{\lambda_f(m)}\lambda_f(n)}{\|f\|_M^2} \prod_{p|L}\frac{1}{r_f(p)}. \tag{7-2}$$

Next we need a formula for $r_f(p)^{-1}$. Recall from (2-1) that at a prime $p \nmid M$ the local adjoint square L function is given by

$$L_p(1,\mathrm{Ad}^2 f) = \frac{1}{1-p^{-2}} \sum_{\alpha\geqslant 0} \frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha} = \frac{1}{\big(1-\frac{\alpha(p)/\beta(p)}{p}\big)\big(1-\frac{1}{p}\big)\big(1-\frac{\beta(p)/\alpha(p)}{p}\big)}$$

so that

$$\sum_{\alpha\geqslant 0}\frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha}=\frac{1+\frac{1}{p}}{\left(1-\frac{\alpha(p)/\beta(p)}{p}\right)\left(1-\frac{\beta(p)/\alpha(p)}{p}\right)}$$
$$=\frac{1+\frac{1}{p}}{\left(1+\frac{1}{p}\right)^2-\frac{|\lambda_f(p)|^2}{p}}=\frac{1}{\left(1+\frac{1}{p}\right)r_f(p)},$$

where the second equals sign follows from the formulas

$$|\lambda_f(p)|^2=\bar\epsilon(p)\lambda_f(p)^2,\quad \lambda_f(p)=\alpha(p)+\beta(p),\quad \alpha(p)\beta(p)=\epsilon(p),$$

which are valid when $p\nmid M$. We can summarize the above calculation and Lemma 2.2 as

$$r_f(p)^{-1}=\begin{cases}\left(1+\dfrac{1}{p}\right)\displaystyle\sum_{\alpha\geqslant 0}\frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha} & \text{if } p\nmid M,\\ \left(1-\dfrac{a_{M,\epsilon}(p)}{p}\right)^{-1} & \text{if } p\mid M.\end{cases}$$

Let

$$\Delta^\star_{\kappa,N,\epsilon}(m,n)=c_\kappa\sum_{f\in H^\star_\kappa(N,\epsilon)}\frac{\overline{\lambda_f(m)}\lambda_f(n)}{\|f\|_N^2}.$$

Recall the definition of $R(M,L,\epsilon)$ from the statement of Theorem 3.1, which we rearrange to

$$R(M,L,\epsilon)=\frac{\psi(M)}{\psi(ML)}\prod_{\substack{p^2\mid L\\ p\nmid M}}\left(1-\frac{1}{p^2}\right)^{-1}\prod_{\substack{p\mid L\\ p\nmid M}}\left(1+\frac{1}{p}\right)\prod_{p\mid(M,L)}\left(1-\frac{a_{M,\epsilon}(p)}{p}\right)^{-1}.$$

We have then that

$$\Delta_{\kappa,N,\epsilon}(m,n)=\sum_{LM=N}R(M,L,\epsilon)\sum_{\substack{\ell\mid L^\infty\\ (\ell,M)=1}}\frac{\bar\epsilon(\ell)}{\ell}\Delta^\star_{\kappa,M,\epsilon}(m,n\ell^2).\tag{7-3}$$

This is analogous to the first half of [Barrett et al. 2017, Proposition 4.1]. Now we would like to invert this formula, and we prepare for this with two lemmas.

Lemma 7.4. *Let $\alpha,\beta\geqslant 0$ and $0\leqslant\gamma\leqslant\beta$ and $\boldsymbol{c}_p(\epsilon)\leqslant\beta-1$. Then*

$$R(p^\beta,p^\alpha,\epsilon)R(p^\gamma,p^{\beta-\gamma},\epsilon)=R(p^\gamma,p^{\alpha+\beta-\gamma},\epsilon).\tag{7-4}$$

Proof. We check cases.

<u>Case $\alpha\geqslant 0$ and $\beta=\gamma$.</u> Note that $R(p^\gamma,1,\epsilon)=1$ for any $\gamma\geqslant 0$.

<u>Case $\alpha=0$.</u> Note that $R(p^\beta,1,\epsilon)=1$ for any $\beta\geqslant 0$.

<u>Case $\alpha\geqslant 1$, $\beta=1$ and $\gamma=0$.</u> We have by hypothesis $\boldsymbol{c}_p(\epsilon)=0$, so

$$R(p,p^\alpha)R(1,p)=\frac{\psi(p)}{\psi(p^{\alpha+1})}\left(1-\frac{1}{p^2}\right)^{-1}\frac{1}{\psi(p)}\left(1+\frac{1}{p}\right).$$

On the other hand, we also have

$$R(1, p^{\alpha+1}) = \frac{1}{\psi(p^{\alpha+1})}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right).$$

Case $\alpha \geqslant 1,\ \beta \geqslant 2$ and $\gamma = 0$. We have $p \mid (p^\beta, p^\alpha)$ and $a_{p^\beta,\epsilon}(p) = 0$, so $R(p^\beta, p^\alpha, \epsilon) = p^{-\alpha}$ and

$$R(1, p^\beta, \epsilon) = \frac{1}{\psi(p^\beta)}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right),$$

$$R(1, p^{\alpha+\beta}, \epsilon) = \frac{1}{\psi(p^{\alpha+\beta})}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right).$$

Generic case $\alpha \geqslant 1,\ \beta \geqslant 2,\ 1 \leqslant \gamma \leqslant \beta - 1$, and $\boldsymbol{c}_p(\epsilon) \leqslant \beta - 1$. We have

$$R(p^\beta, p^\alpha, \epsilon) = \frac{\psi(p^\beta)}{\psi(p^{\alpha+\beta})},$$

$$R(p^\gamma, p^{\beta-\alpha}, \epsilon) = \frac{\psi(p^\gamma)}{\psi(p^\beta)}\left(1-\frac{a_{p^\gamma,\epsilon}(p)}{p}\right)^{-1},$$

$$R(p^\gamma, p^{\beta+\alpha-\gamma}, \epsilon) = \frac{\psi(p^\gamma)}{\psi(p^{\alpha+\beta})}\left(1-\frac{a_{p^\gamma,\epsilon}(p)}{p}\right)^{-1}.$$

The above cover all the cases in the lemma. □

Lemma 7.5. *Let $N \in \mathbb{N}$, $N = LM$, and $M = WQ$. Then*

$$R(M, L, \epsilon)R(W, Q, \epsilon)\delta_{\boldsymbol{c}(\epsilon)\mid W} = R(W, LQ, \epsilon)\delta_{\boldsymbol{c}(\epsilon)\mid W}.$$

Proof. Both sides of the desired formula are multiplicative. Let $\alpha = v_p(L)$, $\beta = v_p(M)$, and $\gamma = v_p(W)$. It then suffices to check that

$$R(p^\beta, p^\alpha, \epsilon)R(p^\gamma, p^{\beta-\gamma}, \epsilon)\delta_{\gamma \geqslant \boldsymbol{c}_p(\epsilon)} = R(p^\gamma, p^{\alpha+\beta-\gamma}, \epsilon)\delta_{\gamma \geqslant \boldsymbol{c}_p(\epsilon)}. \tag{7-5}$$

If $\boldsymbol{c}_p(\epsilon) \leqslant \beta - 1$ then (7-5) is true by Lemma 7.4. So, suppose not. Then $\beta \leqslant \boldsymbol{c}_p(\epsilon) \leqslant \gamma$, but $W \mid M$ so $\gamma \leqslant \beta$ and so $\beta = \gamma$. In the case $\beta = \gamma$, (7-5) is true because $R(p^\beta, 1, \epsilon) = 1$. □

We are now prepared to invert (7-3) using Lemma 7.5. We calculate

$$\sum_{LM=N} \mu(L)R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell}\Delta_{\kappa,M,\epsilon}(m, n\ell^2)$$

$$= \sum_{LM=N} \mu(L)R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \sum_{QW=M} R(W, Q, \epsilon) \sum_{\substack{q \mid Q^\infty \\ (q, W)=1}} \frac{\bar{\epsilon}(q)}{q}\Delta^\star_{\kappa,W,\epsilon}(m, n\ell^2q^2)$$

$$= \sum_{LM=N} \mu(L) \sum_{QW=M} R(M, L, \epsilon)R(W, Q, \epsilon) \sum_{\substack{b \mid (LQ)^\infty \\ (b, W)=1}} \frac{\bar{\epsilon}(b)}{b}\Delta^\star_{\kappa,W,\epsilon}(m, nb^2)$$

$$
\begin{aligned}
&= \sum_{WX=N} R(W,X,\epsilon) \sum_{\substack{b\mid X^\infty\\(b,W)=1}} \frac{\bar{\epsilon}(b)}{b} \Delta^\star_{\kappa,W,\epsilon}(m,nb^2) \sum_{LQ=X} \mu(L)\\
&= R(N,1,\epsilon)\Delta^\star_{\kappa,N,\epsilon}(m,n)\\
&= \Delta^\star_{\kappa,N,\epsilon}(m,n),
\end{aligned}
$$

where the first equality follows by (7-3), the third by Lemma 7.5, and the fourth by Möbius inversion. ☐

Acknowledgements

I would like to thank Nathan Kaplan for careful reading and pointing out the Tsfasman–Vlăduţ–Zink theorem to me, Corentin Perret-Gentil for some helpful discussions, and the anonymous referee for a thorough and detailed report on the first version of this paper.

References

[Atkin and Lehner 1970] A. O. L. Atkin and J. Lehner, "Hecke operators on $\Gamma_0(m)$", *Math. Ann.* **185** (1970), 134–160. MR Zbl

[Banks 1997] W. D. Banks, "Twisted symmetric-square L-functions and the nonexistence of Siegel zeros on GL(3)", *Duke Math. J.* **87**:2 (1997), 343–353. MR Zbl

[Barrett et al. 2017] O. Barrett, P. Burkhardt, J. DeWitt, R. Dorward, and S. J. Miller, "One-level density for holomorphic cusp forms of arbitrary level", *Res. Number Theory* **3** (2017), art. id. 25. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "The second moment of twisted modular L-functions", *Geom. Funct. Anal.* **25**:2 (2015), 453–516. MR Zbl

[Blomer et al. 2017] V. Blomer, J. Buttcane, and P. Maga, "Applications of the Kuznetsov formula on GL(3) II: the level aspect", *Math. Ann.* **369**:1-2 (2017), 723–759. MR Zbl

[Conrey et al. 1997] J. B. Conrey, W. Duke, and D. W. Farmer, "The distribution of the eigenvalues of Hecke operators", *Acta Arith.* **78**:4 (1997), 405–409. MR Zbl

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Mathematics **74**, Springer, 2000. MR Zbl

[Diamond and Im 1995] F. Diamond and J. Im, "Modular forms and modular curves", pp. 39–133 in *Seminar on Fermat's Last Theorem* (Toronto, 1993–1994), CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Diamond and Shurman 2005] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics **228**, Springer, 2005. MR Zbl

[Duke and Kowalski 2000] W. Duke and E. Kowalski, "A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations", *Invent. Math.* **139**:1 (2000), 1–39. MR Zbl

[Gelbart and Jacquet 1978] S. Gelbart and H. Jacquet, "A relation between automorphic representations of GL(2) and GL(3)", *Ann. Sci. École Norm. Sup.* (4) **11**:4 (1978), 471–542. MR Zbl

[Goldfeld et al. 1994] D. Goldfeld, J. Hoffstein, and D. Lieman, "Appendix: an effective zero-free region", (1994). Appendix to J. Hoffstein and P. Lockhart, "Coefficients of Maass forms and the Siegel zero", *Ann. of Math.* (2) **140**:1 (1994), 161–181. MR Zbl

[Heath-Brown 1981] D. R. Heath-Brown, "The fourth power mean of Dirichlet's L-functions", *Analysis* **1**:1 (1981), 25–32. MR Zbl

[Humphries 2018] P. Humphries, "Density theorems for exceptional eigenvalues for congruence subgroups", *Algebra Number Theory* **12**:7 (2018), 1581–1610. MR Zbl

[Igusa 1959] J. Igusa, "Kroneckerian model of fields of elliptic modular functions", *Amer. J. Math.* **81** (1959), 561–577. MR Zbl

[Iwaniec 1997] H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics **17**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Iwaniec et al. 2000] H. Iwaniec, W. Luo, and P. Sarnak, "Low lying zeros of families of L-functions", *Inst. Hautes Études Sci. Publ. Math.* **91** (2000), 55–131. MR Zbl

[Kaplan and Petrow 2017] N. Kaplan and I. Petrow, "Elliptic curves over a finite field and the trace formula", *Proc. Lond. Math. Soc.* (3) **115**:6 (2017), 1317–1372. MR Zbl

[Knightly and Li 2013] A. Knightly and C. Li, *Kuznetsov's trace formula and the Hecke eigenvalues of Maass forms*, Mem. Amer. Math. Soc. **1055**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Kowalski and Michel 1999] E. Kowalski and P. Michel, "The analytic rank of $J_0(q)$ and zeros of automorphic L-functions", *Duke Math. J.* **100**:3 (1999), 503–542. MR Zbl

[Li 1975] W. C. W. Li, "Newforms and functional equations", *Math. Ann.* **212** (1975), 285–315. MR Zbl

[Li 1979] W. C. W. Li, "L-series of Rankin type and their functional equations", *Math. Ann.* **244**:2 (1979), 135–166. MR Zbl

[Milne 2017] J. S. Milne, "Modular functions and modular forms", course notes, 2017, available at https://www.jmilne.org/math/CourseNotes/mf.html. version 1.31.

[Moreno 1991] C. Moreno, *Algebraic curves over finite fields*, Cambridge Tracts in Mathematics **97**, Cambridge University Press, 1991. MR Zbl

[Murty and Sinha 2009] M. R. Murty and K. Sinha, "Effective equidistribution of eigenvalues of Hecke operators", *J. Number Theory* **129**:3 (2009), 681–714. MR Zbl

[Nelson 2017] P. D. Nelson, "Analytic isolation of newforms of given level", *Arch. Math.* (*Basel*) **108**:6 (2017), 555–568. MR Zbl

[Ng 2012] M. Ng, *The basis for space of cusp forms and Petersson trace formula*, master's thesis, University of Hong Kong, 2012.

[Ogg 1969] A. P. Ogg, "On the eigenvalues of Hecke operators", *Math. Ann.* **179** (1969), 101–108. MR Zbl

[Petrow and Young 2018] I. Petrow and M. P. Young, "A generalized cubic moment and the Petersson formula for newforms", *Math. Ann.* (online publication August 2018).

[Ram Murty 1995] M. Ram Murty, "The analytic rank of $J_0(N)(\mathbb{Q})$", pp. 263–277 in *Number theory* (Halifax, 1994), edited by K. Dilcher, CMS Conf. Proc. **15**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Ross 1992] S. L. Ross, II, "A simplified trace formula for Hecke operators for $\Gamma_0(N)$", *Trans. Amer. Math. Soc.* **331**:1 (1992), 425–447. MR Zbl

[Rouymi 2011] D. Rouymi, "Formules de trace et non-annulation de fonctions L automorphes au niveau $\mathfrak{p}^\nu$", *Acta Arith.* **147**:1 (2011), 1–32. MR Zbl

[Schulze-Pillot and Yenirce 2018] R. Schulze-Pillot and A. Yenirce, "Petersson products of bases of spaces of cusp forms and estimates for Fourier coefficients", *Int. J. Number Theory* **14**:8 (2018), 2277–2290. MR Zbl

[Serre 1997] J.-P. Serre, "Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p", *J. Amer. Math. Soc.* **10**:1 (1997), 75–102. MR Zbl

[Tsfasman et al. 1982] M. A. Tsfasman, S. G. Vlăduţ, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound", *Math. Nachr.* **109** (1982), 21–28. MR Zbl

[Venkatesh 2006] A. Venkatesh, "Large sieve inequalities for GL(n)-forms in the conductor aspect", *Adv. Math.* **200**:2 (2006), 336–356. MR Zbl

[Young 2018] M. P. Young, "Explicit calculations with Eisenstein series", *J. Number Theory* (online publication December 2018).

Communicated by Roger Heath-Brown
Received 2018-05-06 Revised 2018-07-22 Accepted 2018-08-23

ian.petrow@math.ethz.ch Departement Mathematik, ETH Zürich, Zürich, Switzerland

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2499

2-parts of real class sizes

Hung P. Tong-Viet

We investigate the structure of finite groups whose noncentral real class sizes have the same 2-part. In particular, we prove that such groups are solvable and have 2-length one. As a consequence, we show that a finite group is solvable if it has two real class sizes. This confirms a conjecture due to G. Navarro, L. Sanus and P. Tiep.

1. Introduction

Let G be a finite group. An element x of G is *real* if x and x^{-1} are conjugate in G. A conjugacy class x^G of G is real if x^G contains a real element. If x^G is a real class of G, then we call the size $|x^G|$ of x^G a *real class size*. We call $|x^G|$ a *noncentral real class size* if x is real and noncentral in G, i.e., x does not lie in the center $\mathbf{Z}(G)$ of G.

A classical result due to Burnside (see [Dornhoff 1971, Corollary 23.4]) states that a finite group is of odd order if and only if the identity element is the only real element. This result has been generalized by Chillag and Mann [1998], who showed that if a finite group G has only one real class size, or equivalently if every real element lies in $\mathbf{Z}(G)$, then G is isomorphic to a direct product of a 2-group and a group of odd order.

The main purpose of this paper is to prove the following.

Theorem A. *Let G be a finite group. If G has two real class sizes, then G is solvable.*

This confirms a conjecture due to G. Navarro, L. Sanus and P. Tiep. Theorem A is best possible in the sense that there are nonsolvable groups with exactly three real class sizes. In fact, the special linear group $\mathrm{SL}_2(q)$ of degree 2 over a finite field of size q, where $q \geq 7$ is a prime power and is congruent to -1 modulo 4, has three real class sizes, namely 1, $q(q-1)$ and $q(q+1)$ [Dornhoff 1971, Theorem 38.1], but $\mathrm{SL}_2(q)$ is nonsolvable.

The extremal condition as in Theorem A has been studied extensively in the literature for conjugacy classes as well as character degrees of finite groups. Itô [1953] has shown that if a finite group has only two class sizes, then it is nilpotent. The alternating group A_4 has two real class sizes, which are 1 and 3, but it is not nilpotent. So, we cannot replace solvability by nilpotency in Theorem A. Itô [1970] also showed that if a finite group has three class sizes, then it is solvable. Clearly, this cannot happen for real class sizes by our example in the previous paragraph.

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2018 semester.
MSC2010: primary 20E45; secondary 20D10.
Keywords: real conjugacy classes, involutions, 2-parts, quasisimple groups.

Recall that a character χ of a finite group G is *real-valued* if χ takes real values, or equivalently, χ coincides with its complex conjugate. Notice that the reality of conjugacy classes of a group can be read off from the character table of the group. This follows from the fact that an element $x \in G$ is real if and only if $\chi(x)$ is real for all complex irreducible characters χ of G [Dornhoff 1971, Lemma 23.2].

For the corresponding results in real-valued characters, Iwasaki [1980] and Moretó and Navarro [2008] have studied the structure of finite groups with two and three real-valued irreducible characters. They show that all these groups must be solvable and their Sylow 2-subgroups have a very restricted structure. By Brauer's lemma on character tables [Dornhoff 1971, Theorem 23.3], the number of real conjugacy classes and the number of real-valued irreducible characters of a finite group are the same. Thus the aforementioned results also give the structure of finite groups with at most three real conjugacy classes. For degrees of real-valued characters, Navarro, Sanus and Tiep [Navarro et al. 2009, Theorem B] proved that a finite group is solvable if it has at most three real-valued character degrees.

As already noted in [Navarro et al. 2009], any possible proof of Theorem A is complicated. Instead of giving a direct proof of Theorem A, we prove a much stronger result which implies Theorem A. For an integer $n \geq 1$ and a prime p, the p-part of n, denoted by n_p, is the largest power of p dividing n.

Theorem B. *Let G be a finite group. Suppose that all noncentral real class sizes of G have the same 2-part. Then G is solvable.*

In other words, if $|x^G|_2 = 2^a$ for all noncentral real elements $x \in G$, where $a \geq 0$ is a fixed integer, then G is solvable. In fact, we can say more about the structure of these groups.

Theorem C. *Let G be a finite group. Suppose that all noncentral real class sizes of G have the same 2-part. Then G has 2-length one.*

Recall that a group G is said to have 2-length one if there exist normal subgroups $N \leq K \leq G$ such that N and G/K have odd order and K/N is a 2-group. Theorem C confirms a conjecture proposed in [Tong-Viet 2013].

Several variations of Theorem B are simply not true. Indeed, if we weaken the hypothesis of Theorem B by assuming that $|x^G|_2$ is 1 or 2^a for all real elements $x \in G$, then G need not be solvable. For example, let $G = \mathrm{SL}_2(2^f)$ with $f \geq 2$. Then every element of G is real and $|x^G|_2 = 1$ or 2^f for all elements $x \in G$. We cannot restrict the hypothesis to only real elements of odd order as $\mathrm{SL}_2(7)$ has only one conjugacy class of noncentral real elements of odd order. Also, Theorem B does not hold for odd primes, at least for primes p with $p \equiv -1 \pmod 4$. In fact, we can take $G = \mathrm{PSL}_2(27)$ if $p = 3$ and $G = \mathrm{SL}_2(p)$ if $p \geq 7$. We can check that all noncentral real class sizes of G have the same p-part. There are also some examples for primes p with $p \equiv 1 \pmod 4$; for example, we can take $G = \mathrm{PSL}_3(3)$ if $p = 13$. (We are unable to find any example with $p = 5$.)

Theorems B and C can be considered as (weak) real conjugacy classes versions (only for even prime) of the famous Thompson's theorem on character degrees stating that if a prime p divides the degrees of all nonlinear irreducible complex characters of a finite group G, then G has a normal p-complement. The exact analog of Thompson's theorem does not hold for conjugacy classes. However, Casolo, Dolfi and

Jabara [Casolo et al. 2012] proved that for a fixed prime p, if all noncentral class sizes of a finite group G have the same p-part, then G is solvable and has a normal p-complement. Some real-valued characters versions of Thompson's theorem were obtained in [Navarro et al. 2009; Navarro and Tiep 2010].

In order to describe our strategy, we need some terminology and results from abstract group theory which can be found in Chapter 31 of [Aschbacher 2000]. For a finite group X, the *layer* of X, denoted by $\boldsymbol{E}(X)$, is the subgroup of X generated by all quasisimple subnormal subgroups (or *components*) of X. A finite group L is said to be *quasisimple* if L is perfect and $L/\boldsymbol{Z}(L)$ is a nonabelian simple group. The *generalized Fitting subgroup* of X, denoted by $\boldsymbol{F}^*(X)$, is the central product of $\boldsymbol{E}(X)$ and the Fitting subgroup $\boldsymbol{F}(X)$, the largest nilpotent normal subgroup of X. Bender's theorem states that $\boldsymbol{C}_X(\boldsymbol{F}^*(X)) \leq \boldsymbol{F}^*(X)$ (see, for example, [Aschbacher 2000, 31.13]).

Returning to our problem, for a finite group G, we denote by $\mathrm{Re}(G)$ the set of all real elements of G. Let G be a finite group with $|x^G|_2 = 2^a$ for all $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$, where $a \geq 0$ is a fixed integer. An important consequence of this hypothesis which is key to our proofs is that if $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ is a 2-element and $t \in G$ is a 2-element inverting x, then $\boldsymbol{C}_G(\langle x, t\rangle)$ has a normal Sylow 2-subgroup; in particular, if i is a noncentral involution of G, then $\boldsymbol{C}_G(i)$ has a normal Sylow 2-subgroup (Lemma 2.4).

Let $K = \boldsymbol{O}^{2'}(G)$. Then K satisfies the same hypothesis as G does (Lemma 2.5). Let H be the quotient group $K/\boldsymbol{O}_{2'}(K)$. Then $H = \boldsymbol{O}^{2'}(H)$ and $\boldsymbol{O}_{2'}(H) = 1$ which implies that $\boldsymbol{F}(H) = \boldsymbol{O}_2(H)$ and $\boldsymbol{F}^*(H) = \boldsymbol{O}_2(H)\boldsymbol{E}(H)$. We consider two cases according to whether $\boldsymbol{E}(H)$ is trivial or not. When $\boldsymbol{E}(H) = 1$, we have $\boldsymbol{F}^*(H) = \boldsymbol{O}_2(H)$. Using [Aschbacher 2000, 31.16],

$$\boldsymbol{F}^*(\boldsymbol{N}_H(U)) = \boldsymbol{O}_2(\boldsymbol{N}_H(U))$$

for all 2-subgroups $U \leq H$. Using this, we can show that H is a 2-group and Theorem B holds in this case (Theorem 3.3). When $\boldsymbol{E}(H)$ is nontrivial, it must be a quasisimple group (Lemma 4.1) and actually it is isomorphic to $\mathrm{SL}_2(q)$ with $q \geq 5$ an odd prime power (Theorem 4.3). Using some ad hoc arguments, we finally show that this cannot occur, proving Theorem B. For Theorem C, we know that G is solvable by Theorem B. Now by Bender's theorem, we know that $\boldsymbol{F}^*(H) = \boldsymbol{O}_2(H)$ and thus H is a 2-group by Theorem 3.3. It follows that G has 2-length one as wanted.

Finally, we would like to point out some connections to other classical results in the literature. If G is a finite group which satisfies the hypothesis of Theorem B, then the centralizer of every noncentral involution of G has a normal Sylow 2-subgroup. Finite groups in which the centralizers of involutions are 2-closed (a group is 2-closed if it has a normal Sylow 2-subgroup) have been studied in [Suzuki 1965]. These are the so-called (C)-groups. On the other hand, the groups in which all involutions are central were classified by R. Griess [1978]. In view of these results, it would be interesting to investigate the structure of finite groups in which in the centralizers of all noncentral involutions are 2-closed.

The paper is organized as follows. In Section 2, we collect some properties of groups whose noncentral real class sizes have the same 2-part. We prove Theorem C in Section 3 and Theorem B in Section 4. In the last section, we classify all finite quasisimple groups that can appear as composition factors of the groups considered in Section 4.

2. 2-parts of noncentral real class sizes

In this section, we collect some important properties of real classes as well as draw some consequences on the structure of groups satisfying the hypothesis of Theorem B. Recall that $\mathrm{Re}(G)$ is the set of all real elements of a finite group G. We begin with some properties of real elements and real class sizes.

Fix a nontrivial real element $x \in G$ and set $\boldsymbol{C}^*_G(x) = \{g \in G \mid x^g \in \{x, x^{-1}\}\}$. Then $\boldsymbol{C}^*_G(x)$ is a subgroup of G containing $\boldsymbol{C}_G(x)$ as a normal subgroup. If $t \in G$ such that $x^t = x^{-1}$, then $x^{t^2} = x$ so $t^2 \in \boldsymbol{C}_G(x)$. Assume that x is not an involution. We see that $t \in \boldsymbol{C}^*_G(x) \setminus \boldsymbol{C}_G(x)$ and if $h \in \boldsymbol{C}^*_G(x) \setminus \boldsymbol{C}_G(x)$, then $x^h = x^{-1} = x^t$ so that $ht^{-1} \in \boldsymbol{C}_G(x)$, or equivalently $h \in \boldsymbol{C}_G(x)t$. Thus $\boldsymbol{C}_G(x)$ has index 2 in $\boldsymbol{C}^*_G(x)$ and $\boldsymbol{C}^*_G(x) = \boldsymbol{C}_G(x)\langle t\rangle$. Hence $|x^G|$ is even whenever x is a real element with $x^2 \neq 1$. In particular, $|x^G|$ is even if x is a nontrivial real element of odd order. Notice that if $x \in \mathrm{Re}(G) \cap \boldsymbol{Z}(G)$, then $x^2 = 1$.

Lemma 2.1. *Let G be a finite group and let $N \trianglelefteq G$.*

(1) *If $x \in G$ is real, then every power of x is also real.*

(2) *If $x^g = x^{-1}$ for some $x, g \in G$, then $x^t = x^{-1}$ for some 2-element $t \in G$.*

(3) *If $x \in \mathrm{Re}(G)$ and $|x^G|$ is odd, then $x^2 = 1$.*

(4) *If $|G : N|$ is odd, then $\mathrm{Re}(G) = \mathrm{Re}(N)$.*

Proof. Let $x \in G$ be a real element. Then $x^g = x^{-1}$ for some $g \in G$. If k is any integer, then

$$(x^k)^g = (x^g)^k = (x^{-1})^k = (x^k)^{-1},$$

so x^k is real which proves (1). Write $o(g) = 2^a m$ with $(2, m) = 1$ and let $t = g^m$. Then t is a 2-element and $x^t = x^{g^m} = x^g = x^{-1}$ as $g^2 \in \boldsymbol{C}_G(x)$ and m is odd. This proves (2). Part (3) follows from the discussion above. For part (4), suppose that $N \trianglelefteq G$ and $|G/N|$ is odd. Let $x \in \mathrm{Re}(G)$. Then there exists a 2-element $t \in G$ inverting x by (2). As $|G/N|$ is odd and t is a 2-element, $t \in N$ and thus $x^{-2} = t^{-1}(xtx^{-1}) \in N$. Again, as $|G/N|$ is odd, $x \in N$ and so $x \in \mathrm{Re}(N)$. □

The first two claims of the following lemma are well-known. The last claim follows from [Dolfi et al. 2008, Proposition 6.4], whose proof uses the Baer–Suzuki theorem.

Lemma 2.2. *Let G be a finite group and let $N \trianglelefteq G$.*

(1) *If $x \in N$, then $|x^N|$ divides $|x^G|$.*

(2) *If $Nx \in G/N$, then $|(Nx)^{G/N}|$ divides $|x^G|$.*

(3) *G has no nontrivial real element of odd order if and only if G has a normal Sylow 2-subgroup.*

We next study the structure of finite groups in which all real 2-elements are central. The proof of the following result is similar to that of [Isaacs and Navarro 2010, Corollary C].

Lemma 2.3. *Let G be a finite group. Assume that all real 2-elements of G lie in $\boldsymbol{Z}(G)$. Then G has a normal 2-complement.*

Proof. Let P be a 2-subgroup of G. Then $\mathrm{Re}(P) \subseteq \boldsymbol{Z}(G)$ by the hypothesis of the lemma. Let $Q \leq \boldsymbol{N}_G(P)$ be a $2'$-group. Since $\mathrm{Re}(P) \subseteq \boldsymbol{Z}(G)$, Q centralizes all real elements of P. Now [Isaacs and Navarro 2010, Theorem B] implies that Q centralizes P. Hence, $Q \leq \boldsymbol{C}_G(P)$. It follows that $\boldsymbol{N}_G(P)/\boldsymbol{C}_G(P)$ is a 2-group. As P is chosen arbitrarily, the Frobenius normal p-complement theorem [Aschbacher 2000, 39.4] implies that G has a normal 2-complement. □

Let G be a finite group with $|G|_2 = 2^{a+b}$, where $a, b \geq 0$ are fixed integers. Observe that the two conditions $|x^G|_2 = 2^a$ for all elements $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ and $|\boldsymbol{C}_G(x)|_2 = 2^b$ for all $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ are equivalent. For brevity, we say that a finite group G is an $\mathcal{R}(a, b)$-*group* if G satisfies one of the two equivalent conditions above.

Lemma 2.4. *Let G be an $\mathcal{R}(a, b)$-group, let $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ and let $t \in G$ be a 2-element such that $x^t = x^{-1}$.*

(a) *If $o(x) = 2m$ for some integer $m > 1$, then $x^m \in \boldsymbol{Z}(G)$.*

(b) *If x is a 2-element, then $\boldsymbol{C}_G(\langle x, t\rangle)$ is 2-closed. Moreover, if x is an involution of G, then $\boldsymbol{C}_G(x)$ is 2-closed.*

Proof. For (a), suppose that $o(x) = 2m$ with $m > 1$. Notice that x^m is an involution, $t \in \boldsymbol{C}^*_G(x) \setminus \boldsymbol{C}_G(x)$ and $|\boldsymbol{C}_G(x)|_2 = 2^b$. We have $(x^m)^t = (x^m)^{-1} = x^m$, so $t \in \boldsymbol{C}_G(x^m)$ and hence $\boldsymbol{C}_G(x) \leq \boldsymbol{C}^*_G(x) \leq \boldsymbol{C}_G(x^m)$. Since $x^2 \neq 1$, we have $|\boldsymbol{C}^*_G(x)|_2 = 2|\boldsymbol{C}_G(x)|_2 = 2^{b+1}$. Therefore $|\boldsymbol{C}_G(x^m)|_2 \geq |\boldsymbol{C}^*_G(x)|_2 > 2^b$. Since G is an $\mathcal{R}(a, b)$-group, this forces $x^m \in \boldsymbol{Z}(G)$.

For (b), assume that x is a real 2-element inverted by t. Let $J := \boldsymbol{C}_G(\langle x, t\rangle)$. We first claim that J has no nontrivial real element of odd order. By contradiction, suppose that there exist $y, s \in J$ with $y^s = y^{-1}$, where $o(y) > 1$ is odd. Observe that $[x, y] = [x, s] = [t, y] = [t, s] = 1$ and $x^t = x^{-1}$, $y^s = y^{-1}$, so

$$(xy)^{st} = x^{st} y^{st} = x^t y^{ts} = x^t y^s = x^{-1} y^{-1} = y^{-1} x^{-1} = (xy)^{-1}.$$

So $xy \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$. Since $(o(x), o(y)) = 1$, $\boldsymbol{C}_G(xy) = \boldsymbol{C}_G(x) \cap \boldsymbol{C}_G(y) = \boldsymbol{C}_A(y)$, where $A = \boldsymbol{C}_G(x) \geq J$. Let U be a Sylow 2-subgroup of $\boldsymbol{C}_G(xy)$. Then U is also a Sylow 2-subgroup of both $\boldsymbol{C}_G(x)$ and $\boldsymbol{C}_G(y)$ as $|\boldsymbol{C}_G(xy)|_2 = |\boldsymbol{C}_G(x)|_2 = |\boldsymbol{C}_G(y)|_2 = 2^b$ (noting $x, y, xy \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$). Thus $\boldsymbol{C}_A(y)$ contains a Sylow 2-subgroup of A. Therefore $|y^A|$ is odd and hence $y^2 = 1$ by Lemma 2.1(3). (Note that $y \in \mathrm{Re}(J) \subseteq \mathrm{Re}(A)$.) However, as $o(y)$ is odd, $y = 1$, which is a contradiction. Thus J has no nontrivial real element of odd order and so by Lemma 2.2(3) J has a normal Sylow 2-subgroup as required. Finally, if x is an involution, then we can choose $t = 1$ and the result follows. □

In the next lemma, we show that every normal subgroup of odd index of an $\mathcal{R}(a, b)$-group is also an $\mathcal{R}(a, b)$-group.

Lemma 2.5. *Suppose that G is an $\mathcal{R}(a, b)$-group. Let K be a normal subgroup of G of odd index. Then K is also an $\mathcal{R}(a, b)$-group.*

Proof. Let K be a normal subgroup of G of odd index. By Lemma 2.1(4), we have $\mathrm{Re}(G) = \mathrm{Re}(K)$. Let $x \in \mathrm{Re}(K) \setminus \boldsymbol{Z}(K)$ and let $C := \boldsymbol{C}_G(x)$. Let $P \in \mathrm{Syl}_2(C)$ and let $S \in \mathrm{Syl}_2(G)$ such that $P \leq S$. Since

$|G:K|$ is odd, we have $P \le S \le K$. In particular, $S \in \mathrm{Syl}_2(K)$. We have $P \le K \cap C = \boldsymbol{C}_K(x) \le C$. Thus P is also a Sylow 2-subgroup of $\boldsymbol{C}_K(x)$. Therefore $|C|_2 = |\boldsymbol{C}_K(x)|_2$ and hence K is an $\mathcal{R}(a, b)$-group. □

The first part of the following lemma is essentially Lemma 2.2 in [Guralnick et al. 2011] and its proof. For the reader's convenience, we include the proof here.

Lemma 2.6. *Let G be a finite group and let $N \trianglelefteq G$ be a normal subgroup of odd order. Let $xN \in G/N$ be a real element. Write $\overline{G} = G/N$.*

(1) *There exist $y \in G$ and a 2-element $t \in G$ such that $\bar{x} = \bar{y}$ and $y^t = y^{-1}$.*

(2) *If $\bar{x}$ is a 2-element, then y can be chosen to be a 2-element. Moreover, if $\bar{x}$ is an involution in $\overline{G}$, then y can be chosen to be an involution in G.*

(3) *In (2), we have $|\bar{x}^{\overline{G}}|_2 = |y^G|_2$ and $\boldsymbol{C}_{\overline{G}}(\langle \bar{x}, \bar{t}\rangle) = \overline{\boldsymbol{C}_G(\langle y, t\rangle)}$.*

Proof. Let xN be a real element of G/N. Then there exists a 2-element $tN \in G/N$ inverting xN by Lemma 2.1(2). By considering the 2-part of t, we can assume that t is a 2-element. The maps $z \mapsto z^{-1}$ and $z \mapsto z^t$ are commuting permutations of G having 2-power order and thus their product σ has 2-power order. Each of these two permutations maps xN to $x^{-1}N$ and $x^{-1}N$ to xN. So σ defines a permutation on xN. Since σ has 2-power order and $|xN| = |N|$ is odd, σ has a fixed point $y \in xN$. Thus $y = y^\sigma = (y^t)^{-1}$. So $y^t = y^{-1}$ as wanted. This proves (1).

Assume next that the order of xN in G/N is 2^k for some integer $k \ge 0$. By part (1), there exist elements $y, t \in G$ such that $xN = yN$ and $y^t = y^{-1}$. Since $|N|$ is odd, $o(y) = 2^k m$, where m is odd and $y^{2^k} \in N$. Since $(2^k, m) = 1$, there exist integers u, v such that $1 = um + v2^k$ and thus $yN = (y^{um}N)(y^{v2^k}N) = y^{um}N$. Clearly, y^{um} is a 2-element and is inverted by t. Replace y by y^{um}, the first claim of part (2) follows. Moreover, as $|N|$ is odd, if $y^2 \in N$ and y is a 2-element, then $y^2 = 1$. Hence the last claim of (2) follows.

Finally, for part (3), by [Isaacs 2008, Lemma 7.7], $\boldsymbol{C}_{\overline{G}}(\bar{x}) = \boldsymbol{C}_{\overline{G}}(\bar{y}) = \overline{\boldsymbol{C}_G(y)}$ since $(o(y), |N|) = 1$. If U is a Sylow 2-subgroup of $\boldsymbol{C}_G(y)$, then $\overline{U}$ is a Sylow 2-subgroup of $\overline{\boldsymbol{C}_G(y)}$ and $|U| = |\overline{U}|$ so $|\bar{y}^{\overline{G}}|_2 = |y^G|_2$. Finally, we have $\boldsymbol{C}_{\overline{G}}(\langle \bar{x}, \bar{t}\rangle) = \boldsymbol{C}_{\overline{G}}(\langle \bar{y}, \bar{t}\rangle) = \overline{\boldsymbol{C}_G(\langle y, t\rangle)}$, where the second equality follows from [Isaacs 2008, Lemma 7.7] again as $\langle y, t\rangle$ is a 2-group. □

In the next lemma, we determine some properties of the quotient group $G/\boldsymbol{O}_{2'}(G)$.

Lemma 2.7. *Let G be an $\mathcal{R}(a, b)$-group and T a subgroup of G containing $\boldsymbol{O}_{2'}(G)$. Let $\overline{G} = G/\boldsymbol{O}_{2'}(G)$.*

(1) *If $\bar{x} \in \mathrm{Re}(\overline{T}) \setminus \boldsymbol{Z}(\overline{T})$ is a 2-element, then there exists a 2-element $\bar{t} \in \overline{T}$ inverting $\bar{x}$ and $\boldsymbol{C}_{\overline{G}}(\langle \bar{x}, \bar{t}\rangle)$ has a normal Sylow 2-subgroup.*

(2) *If $\bar{x} \in \mathrm{Re}(\overline{T}) \setminus \boldsymbol{Z}(\overline{T})$ is an involution, then $\boldsymbol{C}_{\overline{G}}(\bar{x})$ is 2-closed.*

(3) *If $\bar{z}, \bar{x} \in \mathrm{Re}(\overline{G}) \setminus \boldsymbol{Z}(\overline{G})$ and $\bar{x}$ is a 2-element, then $|\bar{z}^{\overline{G}}|_2 \le |\bar{x}^{\overline{G}}|_2 = 2^a$ and $|\boldsymbol{C}_{\overline{G}}(\bar{x})|_2 = 2^b$.*

Proof. For part (1), let $\bar{x} \in \mathrm{Re}(\overline{T}) \setminus \boldsymbol{Z}(\overline{T})$ be a 2-element. By applying Lemma 2.6 for T, there exist 2-elements $y, t \in T$ such that $\bar{x} = \bar{y}$ and $y^t = y^{-1}$. Since $\bar{y} = \bar{x} \notin \boldsymbol{Z}(\overline{T})$, $y \notin \boldsymbol{Z}(T)$ and so $y \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$. By Lemma 2.4(b), $\boldsymbol{C}_G(\langle y, t\rangle)$ is 2-closed and thus by Lemma 2.6(3), $\boldsymbol{C}_{\overline{G}}(\langle \bar{x}, \bar{t}\rangle) = \boldsymbol{C}_{\overline{G}}(\langle \bar{y}, \bar{t}\rangle) = \overline{\boldsymbol{C}_G(\langle y, t\rangle)}$ is 2-closed.

Part (2) follows from Lemma 2.6(2) and part (1) above. For part (3), let $\bar{z}, \bar{x} \in \mathrm{Re}(\overline{G}) \setminus \boldsymbol{Z}(\overline{G})$, where $\bar{x}$ is a 2-element. By Lemma 2.6(1), there exists $w \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ with $\bar{w} = \bar{z}$. We have that $|\bar{w}^{\overline{G}}|_2 \leq |w^G|_2 = 2^a$ as $|\bar{w}^{\overline{G}}|$ divides $|w^G|$. By Lemma 2.6(3), $|\bar{x}^{\overline{G}}|_2 = |y^G|_2 = 2^a$, which implies that $|\boldsymbol{C}_{\overline{G}}(\bar{x})|_2 = 2^b$. The proof is now complete. □

3. Proof of Theorem C

In this section, we prove Theorem C assuming the solvability from Theorem B. Indeed, Theorem C follows immediately from Theorem 3.3. It is convenient to state the following.

Hypothesis A. Let G be a finite group with $\boldsymbol{O}_{2'}(G) = 1$ and $G = \boldsymbol{O}^{2'}(G)$. Let $a, b \geq 0$ be fixed integers. Assume the following hold.

(1) If $T \leq G$ and $x \in \mathrm{Re}(T) \setminus \boldsymbol{Z}(T)$ is a 2-element, then there exists a 2-element $t \in T$ inverting x and $\boldsymbol{C}_G(\langle x, t\rangle)$ is 2-closed.

(2) If $x \in G \setminus \boldsymbol{Z}(G)$ is an involution, then $\boldsymbol{C}_G(x)$ is 2-closed.

(3) If $x \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ is a 2-element, then $|x^G|_2 = 2^a$ and $|\boldsymbol{C}_G(x)|_2 = 2^b$.

(4) If $z \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$, then $|z^G|_2 \leq 2^a$ and $|\boldsymbol{C}_G(z)|_2 \geq 2^b$.

Using Lemmas 2.5 and 2.7, we show that if a finite group G satisfies the hypothesis of Theorem B, then a certain section of G satisfies the hypothesis above.

Lemma 3.1. *If G is an $\mathcal{R}(a, b)$-group, $K = \boldsymbol{O}^{2'}(G)$ and $H = K/\boldsymbol{O}_{2'}(K)$, then H satisfies Hypothesis A.*

Proof. Let G be an $\mathcal{R}(a, b)$-group and let $K = \boldsymbol{O}^{2'}(G)$. We know from Lemma 2.5 that K is also an $\mathcal{R}(a, b)$-group. Let $H = K/\boldsymbol{O}_{2'}(K)$. Then $\boldsymbol{O}^{2'}(H) = H$, $\boldsymbol{O}_{2'}(H) = 1$ and H satisfies the conclusion of Lemma 2.7. Therefore, H satisfies Hypothesis A. □

The following is an easy consequence of Lemma 2.3.

Lemma 3.2. *Let G be a finite group. Assume that $|\boldsymbol{C}_G(u)|_2 = 2^b$ for all 2-elements $u \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$. If $U \leq G$ is a 2-subgroup of G with $|U| \geq 2^b$, then $\boldsymbol{C}_G(U)$ has a normal 2-complement.*

Proof. Let $U \leq G$ be a 2-group with $|U| \geq 2^b$. Let $C = \boldsymbol{C}_G(U)$. Observe that if all 2-elements in $\mathrm{Re}(C)$ lie in $\boldsymbol{Z}(C)$, then C has a normal 2-complement by Lemma 2.3. Thus, by contradiction, assume that there exists a real 2-element $y \in \mathrm{Re}(C) \setminus \boldsymbol{Z}(C)$. Clearly $y \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ and so $|\boldsymbol{C}_G(y)|_2 = 2^b$. We see that $U \leq \boldsymbol{C}_G(y)$ as $y \in \boldsymbol{C}_G(U)$, so $\langle U, y\rangle$ is a 2-subgroup of $\boldsymbol{C}_G(y)$. Since $|\boldsymbol{C}_G(y)|_2 = 2^b$ and $|U| \geq 2^b$, $U \in \mathrm{Syl}_2(\boldsymbol{C}_G(y))$, which implies that $y \in U$. But this implies that $y \in \boldsymbol{Z}(C)$ since $[C, U] = 1$ and $y \in U$, a contradiction. □

We now prove the key result of this section.

Theorem 3.3. *Let G be a finite group satisfying Hypothesis A. If $\boldsymbol{F}^*(G) = \boldsymbol{O}_2(G)$, then G is a 2-group.*

Proof. Suppose that G satisfies Hypothesis A and $\boldsymbol{F}^*(G) = \boldsymbol{O}_2(G)$ but G is not a 2-group. Notice that $G = \boldsymbol{O}^{2'}(G)$. If G has no nontrivial real element of odd order, then G has normal Sylow 2-subgroup by Lemma 2.2(3). But then since $G = \boldsymbol{O}^{2'}(G)$, it must be a 2-group. Therefore, we may assume that there exists an element $z \in \mathrm{Re}(G) \setminus \boldsymbol{Z}(G)$ of odd order.

Let $V \in \mathrm{Syl}_2(\boldsymbol{C}_G(z))$. By Hypothesis A(4), $|V| \geq 2^b$. Clearly, by Hypothesis A(3), G satisfies the hypothesis of Lemma 3.2 and so $\boldsymbol{C}_G(V)$ has a normal 2-complement, say W. Then $W = \boldsymbol{O}_{2'}(\boldsymbol{C}_G(V))$ and $\boldsymbol{C}_G(V)/W$ is a 2-group. Now, we see that $W \trianglelefteq \boldsymbol{C}_G(V) \trianglelefteq \boldsymbol{N}_G(V)$ and W is characteristic in $\boldsymbol{C}_G(V)$ so $W \leq \boldsymbol{O}_{2'}(\boldsymbol{N}_G(V))$. However, by [Aschbacher 2000, 31.16], as $\boldsymbol{F}^*(G) = \boldsymbol{O}_2(G)$, we have $\boldsymbol{F}^*(\boldsymbol{N}_G(V)) = \boldsymbol{O}_2(\boldsymbol{N}_G(V))$, which forces $\boldsymbol{O}_{2'}(\boldsymbol{N}_G(V)) = 1$ (by Bender's theorem [Aschbacher 2000, 31.13]). Hence $W = 1$, so $\boldsymbol{C}_G(V)$ is a 2-group, which is impossible since $z \in \boldsymbol{C}_G(V)$ is a nontrivial element of odd order. This contradiction shows that G is a 2-group. □

Assuming the solvability from Theorem B, we can now prove Theorem C, which is included in the following.

Theorem 3.4. *Let G be an $\mathcal{R}(a,b)$-group. Then $\boldsymbol{O}^{2'}(G)$ has a normal 2-complement N. So G has 2-length one. Moreover, the 2-group $\boldsymbol{O}^{2'}(G)/N$ has at most two real class sizes.*

Proof. By Theorem B, we assume that G is solvable. Let $K = \boldsymbol{O}^{2'}(G)$ and $H := K/\boldsymbol{O}_{2'}(K)$. Then H satisfies Hypothesis A by Lemma 3.1. As H is solvable and $\boldsymbol{O}_{2'}(H) = 1$, $\boldsymbol{F}^*(H) = \boldsymbol{F}(H) = \boldsymbol{O}_2(H)$. By Theorem 3.3, H is a 2-group and thus K has a normal 2-complement $N = \boldsymbol{O}_{2'}(K)$. So G has 2-length one. Finally, as H is a 2-group which satisfies Hypothesis A, for any $x \in \mathrm{Re}(H)$, we have $|x^H| = 1$ or 2^a. □

4. Proof of Theorem B

We prove Theorem B in this section. We first prove some reduction results.

Lemma 4.1. *Let G be a finite nonsolvable group satisfying Hypothesis A. Let $\overline{G} = G/\boldsymbol{O}_2(G)$. Then $L = \boldsymbol{E}(G)$ is a quasisimple group whose center is a 2-group and $\overline{G}$ is an almost simple group with socle $\bar{L} \cong L/\boldsymbol{Z}(L)$.*

Proof. Since G satisfies Hypothesis A, $G = \boldsymbol{O}^{2'}(G)$ and $\boldsymbol{O}_{2'}(G) = 1$. We have $\boldsymbol{F}^*(G) = \boldsymbol{O}_2(G)\boldsymbol{E}(G)$. As G is nonsolvable, $\boldsymbol{E}(G)$ is nontrivial by Theorem 3.3. Let L be a component of G, that is, L is a perfect quasisimple subnormal subgroup of G. Recall that $\boldsymbol{E}(G)$ is generated by all components of G.

We first claim that $L = \boldsymbol{E}(G)$. Suppose by contradiction that G has another component, say $L_1 \neq L$. Since L is nonsolvable, by Lemma 2.3 there exists a real 2-element $x \in \mathrm{Re}(L) \setminus \boldsymbol{Z}(L)$. By Hypothesis A(1), there exists a 2-element $t \in L$ such that $x^t = x^{-1}$ and $\boldsymbol{C}_G(\langle x, t\rangle)$ has a normal Sylow 2-subgroup, so it is solvable. However, by [Aschbacher 2000, 31.5], $[L_1, L] = 1$, and since $\langle x, t\rangle \leq L$, we have $L_1 \leq \boldsymbol{C}_G(\langle x, t\rangle)$, which is impossible. Thus L is the unique component of G and so $L = \boldsymbol{E}(G)$. As $\boldsymbol{O}_{2'}(G) = 1$, $\boldsymbol{Z}(L)$ is a 2-group.

Let $C = \boldsymbol{C}_G(L)$. We claim that $C = \boldsymbol{O}_2(G)$, which implies that $\overline{G}$ is almost simple with socle $\bar{L}$. In fact, we show that every real 2-element of C lies in $\boldsymbol{Z}(C)$ and so by Lemma 2.3, C has a normal 2-complement which is $\boldsymbol{O}_{2'}(C)$. Since $L \trianglelefteq G$, $C \trianglelefteq G$ and so $\boldsymbol{O}_{2'}(C) \leq \boldsymbol{O}_{2'}(G) = 1$. Hence C is a 2-group and thus $C \leq \boldsymbol{O}_2(G)$. Furthermore, by [Aschbacher 2000, 31.12], $[\boldsymbol{O}_2(G), L] = 1$ so $\boldsymbol{O}_2(G) \leq C$. Therefore, $C = \boldsymbol{O}_2(G)$ as wanted.

To finish the proof, suppose by contradiction that there exists a real 2-element $y \in \mathrm{Re}(C)$ which is not in $\boldsymbol{Z}(C)$. By Hypothesis A(1), there exists a 2-element $t \in C$ inverting y, and $\boldsymbol{C}_G(\langle y, t\rangle)$ has a normal Sylow 2-subgroup. As $[L, C] = 1$ and $\langle y, t\rangle \leq C$, we have $L \leq \boldsymbol{C}_G(\langle y, t\rangle)$, which is impossible. This completes our proof. □

In order to classify all the possible finite quasisimple groups which can appear as $\boldsymbol{E}(G)$ in the previous lemma, we need the following.

Lemma 4.2. *Let G be a finite nonsolvable group satisfying Hypothesis A, and let $L = \boldsymbol{E}(G)$ and $\overline{G} = G/\boldsymbol{O}_2(G)$. If $x, z \in \mathrm{Re}(L) \setminus \boldsymbol{Z}(L)$ and x is a 2-element, then*

$$|z^L|_2 \leq |z^G|_2 \leq 2^a = |x^G|_2 \leq |\overline{G} : \bar{L}|_2 \cdot |x^L|_2 \leq |\mathrm{Out}(\bar{L})|_2 \cdot |x^L|_2. \tag{4-1}$$

Proof. Notice that $L = \boldsymbol{E}(G)$ is a quasisimple group by Lemma 4.1, and by [Aschbacher 2000, 31.12], $[\boldsymbol{O}_2(G), L] = 1$. Let $x, z \in \mathrm{Re}(L) \setminus \boldsymbol{Z}(L)$, where x is a 2-element. By Hypothesis A(3) and (4), $|z^G|_2 \leq 2^a = |x^G|_2$. Since $L \trianglelefteq G$, $|z^L|_2 \leq |z^G|_2 \leq 2^a$. For the remaining inequalities, observe that $\boldsymbol{O}_2(G) \leq \boldsymbol{C}_G(x)$ and

$$|x^G| = |G : L\boldsymbol{C}_G(x)| \cdot |L\boldsymbol{C}_G(x) : \boldsymbol{C}_G(x)|.$$

We have $|L\boldsymbol{C}_G(x) : \boldsymbol{C}_G(x)| = |L : \boldsymbol{C}_L(x)| = |x^L|$ and

$$|G : L\boldsymbol{C}_G(x)| = |G : L\boldsymbol{O}_2(G)| \,/\, |\boldsymbol{C}_G(x) : \boldsymbol{C}_G(x) \cap L\boldsymbol{O}_2(G)|.$$

As $|G : L\boldsymbol{O}_2(G)| = |\overline{G} : \bar{L}|$ divides $|\mathrm{Out}(\bar{L})|$, by taking the 2-parts, we obtain

$$|x^G|_2 \leq |\overline{G} : \bar{L}|_2 \cdot |x^L|_2 \leq |\mathrm{Out}(\bar{L})|_2 \cdot |x^L|_2.$$

The proof is now complete. □

Using the classification of finite simple groups, we can show that L in the previous lemma is isomorphic to a special linear group $\mathrm{SL}_2(q)$, where $q \geq 5$ is an odd prime power. We defer the proof of the following theorem until next section.

Theorem 4.3. *Let L be a quasisimple group with center Z and let $S = L/Z$. Suppose that Z is a 2-group and the following conditions hold:*

(1) *If i is a noncentral involution of L, then $\boldsymbol{C}_L(i)$ is 2-closed.*

(2) *If $x, z \in L$ are noncentral real elements and x is a 2-element, then $|z^L|_2 \leq |\mathrm{Out}(S)|_2 \cdot |x^L|_2$.*

Then $L \cong \mathrm{SL}_2(q)$ with $q \geq 5$ odd.

Let p be a prime. If $n \geq 1$ is an integer, then the p-adic valuation of n, denoted by $\nu_p(n)$, is the highest exponent ν such that p^ν divides n. Hence $n_p = p^{\nu_p(n)}$. Notice that $\nu_p(xy) = \nu_p(x) + \nu_p(y)$ for all integers $x, y \geq 1$.

The following number theoretic result is obvious.

Lemma 4.4. *Let $m, k \geq 1$ be integers, where $m \geq 3$ is odd. Then $\nu_2\big(m^{2^k} - 1\big) \geq k + 2$.*

Proof. We proceed by induction on $k \geq 1$. Suppose that $m \equiv \epsilon \pmod 4$, where $\epsilon = \pm 1$. For the base case, assume $k = 1$. Clearly, $m - \epsilon$ is divisible by 4 while $m + \epsilon$ is divisible by 2. So $\nu_2(m^2 - 1) \geq 3 = k + 2$.

Assume that $\nu_2\big(m^{2^t} - 1\big) \geq t + 2$ for some integer $t \geq 1$. We have

$$m^{2^{t+1}} - 1 = \big(m^{2^t} - 1\big)\big(m^{2^t} + 1\big).$$

Since m is odd, $\nu_2\big(m^{2^t} + 1\big) \geq 1$ and $\nu_2\big(m^{2^t} - 1\big) \geq t + 2$ by the induction hypothesis. Therefore,

$$\nu_2\big(m^{2^{t+1}} - 1\big) = \nu_2\big(m^{2^t} + 1\big) + \nu_2\big(m^{2^t} - 1\big) \geq 1 + (t + 2) = (t + 1) + 2.$$

By induction, the lemma follows. □

We are now ready to prove Theorem B, which we restate here.

Theorem 4.5. *If G is an $\mathcal{R}(a, b)$-group, then G is solvable.*

Proof. Let G be a counterexample to the theorem of minimal order. Then G is nonsolvable and by Lemma 3.1, H satisfies Hypothesis A, where $H = K/\boldsymbol{O}_{2'}(K)$ and $K = \boldsymbol{O}^{2'}(G)$. As G is nonsolvable, H is nonsolvable. Let $L = \boldsymbol{E}(H)$. By Lemmas 4.1, 4.2 and Hypothesis A(2), L is a quasisimple group satisfying the hypothesis of Theorem 4.3, so $L \cong \mathrm{SL}_2(q)$ with $q \geq 5$ an odd prime power. Moreover, $\boldsymbol{O}^{2'}(\overline{H}) = \overline{H}$ and $\overline{H}$ is almost simple with socle $\bar{L}$, where $\overline{H} = H/\boldsymbol{O}_2(H)$.

Write $q = p^f \geq 5$, where $p > 2$ is a prime and $f \geq 1$. It is well-known that

$$\mathrm{Out}(\mathrm{PSL}_2(q)) \cong \langle \delta \rangle \times \langle \varphi \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_f,$$

where δ is a diagonal automorphism of order 2 and φ is a field automorphism of order f of $\mathrm{PSL}_2(p^f)$. Assume $q \equiv \eta \pmod 4$ with $\eta = \pm 1$. We have $k := \nu_2(q - \eta) \geq 2$, $\nu_2(q + \eta) = 1$ and $|L|_2 = (q^2 - 1)_2 = 2^{k+1}$.

Now L has two noncentral real elements z and x (which lie in the cyclic subgroups of L of order $q + \eta$ and $q - \eta$) of order $(q + \eta)/2$ and $2^k > 2$, respectively. It is easy to check that $|x^L|_2 = 2$ and $|z^L|_2 = 2^k$. Moreover, z^L is invariant under the diagonal automorphisms of L. As $\overline{H}/\bar{L}$ is a subgroup of $\mathrm{Out}(\bar{L}) \cong \mathbb{Z}_2 \times \mathbb{Z}_f$ with $\boldsymbol{O}^{2'}(\overline{H}/\bar{L}) = \overline{H}/\bar{L}$, it follows that $\overline{H}/\bar{L}$ is an abelian 2-group of order at most 2^{c+1}, where $c = \nu_2(f)$.

(a) Assume that f is odd. Then $\mathrm{PSL}_2(q) \cong \bar{L} \trianglelefteq \overline{H} \leq \mathrm{PGL}_2(q)$.

If $\overline{H} \cong \mathrm{PSL}_2(q)$, then $H = L\boldsymbol{O}_2(H)$. In this case, we see that $|z^H|_2 = |z^L|_2 = 2^k > 2 = |x^L|_2 = |x^H|_2$, violating (4-1).

Assume that $\overline{H} \cong \mathrm{PGL}_2(q)$. From the character table of $\mathrm{PGL}_2(q)$ (see [Steinberg 1951, Table III]), $\overline{H} \cong \mathrm{PGL}_2(q)$ contains a real element $\bar{y}$ of order p (labeled by A_2) with $|\bar{y}^{\overline{H}}| = q^2 - 1$. Since $o(\bar{y}) = p$ is odd, by [Guralnick et al. 2011, Lemma 2.2] there exists a real element $w \in H$ of order p such that $\bar{w} = \bar{y}$. Now $|w^H|_2 \geq |\bar{y}^{\overline{H}}|_2 = 2^{k+1} > 4 = |\mathrm{Out}(\bar{L})|_2 \cdot |x^L|_2$, violating (4-1).

(b) Assume f is even. We have that $q \equiv 1 \pmod 4$ and since p is odd, by Lemma 4.4 we have $k = \nu_2(q-1) \geq c+2$. Recall that $c = \nu_2(f)$.

From [Dornhoff 1971, Theorem 38.1], $L \cong \mathrm{SL}_2(q)$ has a real element

$$y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

of order p (labeled by c). The image $\bar{y}$ of y in $\bar{L} \cong \mathrm{PSL}_2(q)$ is also a real element of order p and $|y^L| = |\bar{y}^{\bar{L}}| = (q^2-1)/2$. Hence $|y^H|_2 \geq |\bar{y}^{\bar{L}}|_2 = 2^k$ as $|\mathrm{PSL}_2(q)|_2 = 2^k$.

Assume that $|\overline{H}/\bar{L}|_2 \leq 2^c$. We have $|\overline{H} : \bar{L}|_2 \cdot |x^L|_2 \leq 2^{c+1} < 2^{c+2} \leq 2^k \leq |y^H|_2$, contradicting (4-1).

Finally, we assume that $|\overline{H}/\bar{L}|_2 = 2^{c+1}$ and so $\overline{H}/\bar{L} = \langle \delta \rangle \times \langle \varphi^m \rangle$ with $m = f/2^c$. We know that $\bar{y} \in \bar{L}$ is φ-invariant but not δ-invariant. Thus $|\bar{y}^{\overline{H}}|_2 = 2^{k+1}$ and hence $|y^H|_2 \geq |\bar{y}^{\overline{H}}|_2 = 2^{k+1} \geq 2^{c+3}$. Now $|\mathrm{Out}(\bar{L})|_2 \cdot |x^L|_2 = 2^{c+2} < 2^{c+3} \leq |y^H|_2$, which violates (4-1) again. The proof is now complete. □

5. Quasisimple groups

The main purpose of this section is to prove Theorem 4.3. We first prove some easy results which will be needed in our classification.

For a prime p and a group X, the p-rank of X, denoted by $r_p(X)$, is the maximum rank of an elementary abelian p-subgroup of X. Recall that the p-rank of an elementary abelian p-group of order p^k is k. The following easy result should be known.

Lemma 5.1. *Let X be a finite group and let Z be a subgroup of $\boldsymbol{Z}(X)$. Let $i \in X$ be a noncentral involution. Let r be the 2-rank of Z and let $\overline{X} = X/Z$. Let $T \leq X$ be the full inverse image of $\boldsymbol{C}_{\overline{X}}(\bar{i})$. Then $T/\boldsymbol{C}_X(i)$ is an elementary abelian 2-group of order at most 2^r and $|i^X|_2 \leq 2^r \cdot |\bar{i}^{\overline{X}}|_2$.*

Proof. Let $g \in T$. Then $i^g = iz$ for some $z \in Z$. As $i^2 = 1$ and $iz = zi$, we have $z^2 = 1$. Now $\boldsymbol{C}_X(i)^g = \boldsymbol{C}_X(i^g) = \boldsymbol{C}_X(iz) = \boldsymbol{C}_X(i)$, so $\boldsymbol{C}_X(i) \trianglelefteq T$. Moreover, $i^{g^2} = (iz)^g = i^g z = i$, and hence $g^2 \in \boldsymbol{C}_X(i)$. Thus $T/\boldsymbol{C}_X(i)$ is an elementary abelian 2-group. Let $\Omega = \{x \in Z : x^2 = 1\}$. Clearly Ω is an elementary abelian 2-subgroup of X of order 2^r. For each $h \in T$, there exists $z \in \Omega$ such that $i^h = iz$. Moreover, if $i^h = i^k$ for some $h, k \in T$, then $h\boldsymbol{C}_X(i) = k\boldsymbol{C}_X(i)$. Thus there is an injective map from the set of left cosets of $\boldsymbol{C}_X(i)$ in T to Ω and hence $|T : \boldsymbol{C}_X(i)| \leq 2^r$. Now

$$|i^X|_2 = |X : T|_2 \cdot |T : \boldsymbol{C}_X(i)|_2 = |\bar{i}^{\overline{X}}|_2 \cdot |T : \boldsymbol{C}_X(i)| \leq 2^r \cdot |\bar{i}^{\overline{X}}|_2.$$

The proof is complete. □

We next determine all quasisimple groups which have an involution i whose centralizer is solvable. Notice the isomorphisms $\mathrm{PSU}_4(2) \cong \mathrm{PSp}_4(3)$, $\mathrm{PSL}_2(7) \cong \mathrm{PSL}_3(2)$, $A_5 \cong \mathrm{PSL}_2(4) \cong \mathrm{PSL}_2(5)$, $A_6 \cong \mathrm{PSL}_2(9)$ and $A_8 \cong \mathrm{PSL}_4(2)$.

Lemma 5.2. *Let L be a quasisimple group and let $S = L/\boldsymbol{Z}(L)$. Suppose that L has an involution i such that $\boldsymbol{C}_L(i)$ is solvable. Then one of the following holds.*

(i) $L \cong \mathrm{M}_{11}$ *or* $S \cong \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{Fi}_{22}$.

(ii) $L \cong A_n$ $(5 \le n \le 12)$, $2 \cdot A_n$ $(8 \le n \le 12)$ *or* $3 \cdot A_n$ $(6 \le n \le 7)$.

(iii) $S \cong \mathrm{PSL}_2(q), \mathrm{PSL}_3(q), \mathrm{PSU}_3(q), \mathrm{Sp}_4(q), {}^2\mathrm{B}_2(q)$ *with* $q = 2^f$.

(iv) $S \cong \mathrm{PSL}_n(2)$ $(n = 4, 5, 6)$, $\mathrm{PSU}_n(2)$ $(4 \le n \le 9)$, $\mathrm{Sp}_{2n}(2)$ $(3 \le n \le 5)$, $\Omega^{\pm}_{2n}(2)$ $(4 \le n \le 5)$, ${}^3\mathrm{D}_4(2), {}^2\mathrm{F}_4(2)', \mathrm{F}_4(2), {}^2\mathrm{E}_6(2)$.

(v) $S \cong \mathrm{PSL}_3(3), \mathrm{PSL}_4(3), \mathrm{PSU}_3(3), \mathrm{PSU}_4(3), \mathrm{PSp}_4(3), \Omega_7(3), \mathrm{P}\Omega^+_8(3), \mathrm{G}_2(3)$.

(vi) $L \cong \mathrm{PSL}_2(q)$ *with q odd.*

Proof. Let $\bar{L} = L/\boldsymbol{Z}(L)$. Since $\boldsymbol{C}_L(i)$ is solvable, $i \notin \boldsymbol{Z}(L)$ and $\boldsymbol{C}_{\bar{L}}(\bar{i})$ is solvable by Lemma 5.1. Thus S is a nonabelian simple group with a solvable involution centralizer.

(1) Assume that S is a sporadic simple group. The information on the centralizers of involutions of L can be read off from Tables 5.3(a)–(z) in [Gorenstein et al. 1998]. It follows that $L \cong \mathrm{M}_{11}$ or $S \cong \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{Fi}_{22}$.

(2) Assume $S \cong A_n$ with $n \ge 5$. We know that every involution j of $S \cong A_n$ is a product of r disjoint transpositions, where r is even. Let $s = n - 2r$. By [Gorenstein et al. 1998, Proposition 5.2.8] $\boldsymbol{C}_S(j) \cong (H_1 \times H_2)\langle t\rangle$, where $H_2 \cong A_s$ and $H_1 \cong R_1 L_1$ with $L_1 \cong \mathrm{S}_r$ and $R_1 \cong C_2^{r-1}$. Hence if r or s is at least 5, then $\boldsymbol{C}_S(j)$ is nonsolvable. Thus both r and s are at most 4 and hence $n \le 12$. Therefore, if $L = A_n$ with $n \ge 5$, then $5 \le n \le 12$. If $L \cong 2 \cdot A_n$, then $n \le 12$ by our observation above. However, $2 \cdot A_n$ has a noncentral involution only when $n \ge 8$. Thus if $L \cong 2 \cdot A_n$, then $8 \le n \le 12$. For $n = 6, 7$, only $3 \cdot A_n$ has a solvable involution centralizer.

(3) Assume S is a finite simple group of Lie type. The centralizers of involutions of L are determined in [Aschbacher and Seitz 1976] and Tables 4.5.1 and 4.5.2 in [Gorenstein et al. 1998]. From these results, it is easy to get all the possibilities for L. (See also [Liebeck and O'Brien 2007, Lemma 3.4].) □

We use the convention that $\mathrm{PSL}^{\epsilon}_n(q)$ is $\mathrm{PSL}_n(q)$ if $\epsilon = +$ and $\mathrm{PSU}_n(q)$ if $\epsilon = -$. A similar convention applies to $\mathrm{SL}^{\epsilon}_n(q)$. We now prove the main result of this section.

Theorem 5.3. *Let L be a quasisimple group with center Z and let $S = L/Z$. Suppose that Z is a 2-group and the following conditions hold*:

(1) *If i is a noncentral involution of L, then $\boldsymbol{C}_L(i)$ is 2-closed.*

(2) *If $x, z \in \mathrm{Re}(L) \setminus \boldsymbol{Z}(L)$ are noncentral real elements, where x is a 2-element, then*

$$|z^L|_2 \le |\mathrm{Out}(S)|_2 \cdot |x^L|_2.$$

Then $L \cong \mathrm{SL}_2(q)$ with $q \ge 5$ odd.

Proof. We consider the following cases.

Case 1. All involutions of L are central. By the main theorem in [Griess 1978], $L \cong \mathrm{SL}_2(q)$ with $q \geq 5$ odd or $2 \cdot A_7$. If the first case holds, then we are done. So, assume that $L \cong 2 \cdot A_7$. We have $|\mathrm{Out}(S)|_2 = 2$. Using [Conway et al. 1985], we can check that L has two real elements z and x of order 3 and 4, respectively, with $|z^L| = 280$ and $|x^L| = 210$. Clearly $|z^L|_2 = 2^3 > |\mathrm{Out}(S)|_2 \cdot |x^L|_2 = 2^2$, violating condition (2). So this case cannot occur.

Case 2. $\boldsymbol{Z}(L)$ is trivial. Then L is a nonabelian simple group. Let $x \in L$ be a 2-central involution of L, i.e., x is an involution that lies in the center of some Sylow 2-subgroup of L. We have $|x^L|_2 = 1$ so (2) implies $|z^L|_2 \leq |\mathrm{Out}(L)|_2$ for all noncentral real elements $z \in L$.

The centralizer of every noncentral involution of L is 2-closed by the hypothesis. Since L is simple, it follows that the centralizer of every involution of L is 2-closed. Now [Suzuki 1965, Theorem 1] yields that L is isomorphic to one of the following groups: A_6, $\mathrm{PSL}_2(p)$ with p a Fermat or a Mersenne prime, $\mathrm{PSL}_2(2^f)$ with $f \geq 2$, $\mathrm{PSL}_3(q)$, $\mathrm{PSU}_3(q)$ with $q = 2^f$ or ${}^2\mathrm{B}_2(2^{2f+1})$ with $f \geq 1$.

Assume that $L \cong {}^2\mathrm{B}_2(2^{2f+1})$ with $f \geq 1$. It follows from Propositions 3 and 16 in [Suzuki 1962] that L has a real element z of order $2^f - 1$ with $|\boldsymbol{C}_L(z)| = 2^f - 1$ and so $|z^L|_2 = 2^{2(2f+1)} \geq 64$. Since $|\mathrm{Out}(L)| = 2f + 1$ is odd, $|z^L|_2 > |\mathrm{Out}(L)|_2$, and hence this case cannot occur.

Assume $L \cong \mathrm{PSL}_2(p)$ with p a Fermat or a Mersenne prime. We have $|\mathrm{Out}(L)| = 2$ and L possesses a real element z of odd order $(p+\delta)/2$, where $p \equiv \delta \pmod 4$ and $|z^L|_2 = |L|_2 \geq 4$. As $|z^L|_2 > 2 = |\mathrm{Out}(L)|_2$, this case cannot happen.

If $L \cong A_6$, then $|\mathrm{Out}(L)| = 2^2$. However, A_6 has a real element z of order 3 with $|z^L| = 40$ and thus $|z^L|_2 = 8 > |\mathrm{Out}(S)|_2 = 4$. Thus this case cannot happen.

Next, if $L \cong \mathrm{PSL}_2(2^f)$ with $f \geq 2$, then L has a real element z of order $2^f - 1$ with $|z^L| = 2^f(2^f + 1)$. Clearly $|z^L|_2 = 2^f > f \geq |\mathrm{Out}(L)|_2$ as $|\mathrm{Out}(L)| = f$.

Finally, assume $L \cong \mathrm{PSL}_3^\epsilon(2^f)$. We have $f \geq 2$ as $\mathrm{PSL}_3(2) \cong \mathrm{PSL}_2(7)$, where $7 = 2^3 - 1$ is a Mersenne prime and $\mathrm{PSU}_3(2)$ is not simple. In both cases, $|\mathrm{Out}(L)| = 2df$ with $d = (3, 2^f - \epsilon 1)$ so $|\mathrm{Out}(L)|_2 = 2^{1+\nu_2(f)}$. The quasisimple group $X = \mathrm{SL}_3^\epsilon(2^f)$ possesses real elements h of order $2^f + \epsilon 1$ with $|\boldsymbol{C}_X(h)| = 4^f - 1$ and g of order $2^f - \epsilon 1$ with $|\boldsymbol{C}_X(g)| = (2^f - \epsilon 1)^2$ [Guralnick et al. 2011, Lemma 4.4(3)]. Now let $y \in \{g, h\}$ be an element with $o(y)$ relatively prime to $d = (3, 2^f - \epsilon 1)$ and let z be the image of y in $L = \mathrm{PSL}_3^\epsilon(2^f) \cong X/\boldsymbol{Z}(X)$. Then $|\boldsymbol{C}_L(z)| = |\boldsymbol{C}_X(y)/\boldsymbol{Z}(X)|$ (by [Isaacs 2008, Lemma 7.7]) is odd, so $|z^L|_2 = |L|_2 = 2^{3f}$. Since $f \geq 2$, $2^{3f} > 2^{2f} \geq 2^{1+\nu_2(f)}$. Hence these cases cannot occur.

Case 3. $\boldsymbol{Z}(L)$ is nontrivial and L has a noncentral involution. Let j be a noncentral involution of L. By condition (1), $\boldsymbol{C}_L(j)$ is 2-closed and thus it is solvable. Hence L is one of the quasisimple groups in Lemma 5.2. Moreover, as $\boldsymbol{Z}(L)$ is a nontrivial 2-group, the Schur multiplier $\mathrm{M}(S)$ of $S \cong L/\boldsymbol{Z}(L)$ is of even order. It follows that we only need to consider the following cases.

(i) $S \cong \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{Fi}_{22}, A_n$ $(8 \leq n \leq 12)$, $\mathrm{PSL}_3(4)$, ${}^2\mathrm{B}_2(8)$.

(ii) $S \cong \mathrm{PSU}_n(2)$ $(n = 4, 6)$, $\mathrm{Sp}_6(2)$, $\Omega_8^+(2)$, $\mathrm{F}_4(2)$, ${}^2\mathrm{E}_6(2)$.

(iii) $S \cong \mathrm{PSU}_4(3)$, $\mathrm{PSp}_4(3)$, $\Omega_7(3)$, $\mathrm{P}\Omega_8^+(3)$.

We show that these cases cannot occur by showing that condition (2) does not hold. Clearly $\bar{j} \in \bar{L}$ is a noncentral involution and by Lemma 5.1, $|j^L|_2 \leq 2^{r_2(\mathbf{Z}(L))} \cdot |\bar{j}^{\bar{L}}|_2$, where $r_2(\mathbf{Z}(L))$ is the 2-rank of $\mathbf{Z}(L)$. Let $\widetilde{S}$ be the perfect central extension of S such that $\widetilde{S}/\mathbf{Z}(\widetilde{S}) \cong S$ and $|\mathbf{Z}(\widetilde{S})| = |\mathrm{M}(S)|$. From [Conway et al. 1985], we see that $\mathrm{M}(S)$ can be written as a direct product of at most two (possibly trivial) cyclic groups, so $r_2(\mathbf{Z}(L)) \leq r_2(\mathrm{M}(S)) \leq 2$. Let $e_2(S) = \max\{\nu_2(|x^S|) : x \text{ is an involution in } S\}$. Then for each noncentral involution $j \in L$, we have $\nu_2(|j^L|) \leq r_2(\mathrm{M}(S)) + e_2(S)$.

Let $z \in S$ be a nontrivial real element of odd order. There exists $y \in \mathrm{Re}(L) \setminus \mathbf{Z}(L)$ of odd order such that z is the image of y in $L/\mathbf{Z}(L) \cong S$ (see [Navarro et al. 2009, Lemma 3.2] or [Guralnick et al. 2011, Lemma 2.2]) and by applying [Isaacs 2008, Lemma 7.7], $|z^S| = |y^L|$ (noting $(o(y), |\mathbf{Z}(L)|) = 1$). Hence to show that condition (2) does not hold, it suffices to find a nontrivial real element $z \in S$ of odd order such that the following inequality holds:

$$\nu_2(|z^S|) > \nu_2(|\mathrm{Out}(S)|) + r_2(\mathrm{M}(S)) + e_2(S). \tag{5-1}$$

For each simple group S in (i)–(iii) above, we list in Table 1 the invariant $e_2(S)$, the largest 2-part of the sizes of conjugacy classes of involutions of S in the second column, the Atlas class name and the 2-part of the conjugacy class of odd order real element $z \in S$, the order of the outer automorphism group

S	$e_2(S)$	z	$\nu_2(\lvert z^S\rvert)$	$\lvert\mathrm{Out}(S)\rvert$	$\mathrm{M}(S)$
M_{12}	2	$3a$	5	2	$\mathbb{Z}_2$
M_{22}	0	$5a$	7	2	$\mathbb{Z}_{12}$
Fi_{22}	1	$3d$	17	2	$\mathbb{Z}_6$
A_8	1	$5a$	6	2	$\mathbb{Z}_2$
A_9	1	$3b$	4	2	$\mathbb{Z}_2$
A_{10}	1	$3c$	7	2	$\mathbb{Z}_2$
A_{11}	1	$3c$	6	2	$\mathbb{Z}_2$
A_{12}	0	$3d$	7	2	$\mathbb{Z}_2$
$\mathrm{PSL}_3(4)$	0	$3a$	6	12	$\mathbb{Z}_4 \times \mathbb{Z}_{12}$
${}^2\mathrm{B}_2(8)$	0	$5a$	6	3	$\mathbb{Z}_2 \times \mathbb{Z}_2$
$\mathrm{PSU}_4(2)$	1	$5a$	6	2	$\mathbb{Z}_2$
$\mathrm{PSU}_6(2)$	3	$3c$	12	6	$\mathbb{Z}_2 \times \mathbb{Z}_6$
$\mathrm{Sp}_6(2)$	2	$3c$	7	1	$\mathbb{Z}_2$
$\Omega_8^+(2)$	2	$3d$	9	6	$\mathbb{Z}_2 \times \mathbb{Z}_2$
$\mathrm{F}_4(2)$	4	$3c$	18	2	$\mathbb{Z}_2$
${}^2\mathrm{E}_6(2)$	5	$3c$	27	6	$\mathbb{Z}_2 \times \mathbb{Z}_6$
$\mathrm{PSU}_4(3)$	0	$3d$	7	8	$\mathbb{Z}_3 \times \mathbb{Z}_{12}$
$\Omega_7(3)$	0	$3g$	8	2	$\mathbb{Z}_6$
$\mathrm{P}\Omega_8^+(3)$	2	$3m$	11	24	$\mathbb{Z}_2 \times \mathbb{Z}_2$

Table 1. Some small simple groups.

Out(S) and in the last column the Schur multiplier $\mathrm{M}(S)$. The information in this table can be read off from the character table of S using [Conway et al. 1985].

From Table 1, we can check that (5-1) holds, which completes our proof. □

Acknowledgments

The author is grateful to Kay Magaard, Ben Brewster and Dan Rossi for their help during the preparation of this paper. He also thanks the anonymous referee for his or her comments, which helped to improve the clarity of this paper.

References

[Aschbacher 2000] M. Aschbacher, *Finite group theory*, 2nd ed., Cambridge Studies in Advanced Mathematics **10**, Cambridge University Press, 2000. MR Zbl

[Aschbacher and Seitz 1976] M. Aschbacher and G. M. Seitz, "Involutions in Chevalley groups over fields of even order", *Nagoya Math. J.* **63** (1976), 1–91. MR Zbl

[Casolo et al. 2012] C. Casolo, S. Dolfi, and E. Jabara, "Finite groups whose noncentral class sizes have the same p-part for some prime p", *Israel J. Math.* **192**:1 (2012), 197–219. MR Zbl

[Chillag and Mann 1998] D. Chillag and A. Mann, "Nearly odd-order and nearly real finite groups", *Comm. Algebra* **26**:7 (1998), 2041–2064. MR Zbl

[Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups: Maximal subgroups and ordinary characters for simple groups*, Oxford University Press, 1985. MR Zbl

[Dolfi et al. 2008] S. Dolfi, G. Navarro, and P. H. Tiep, "Primes dividing the degrees of the real characters", *Math. Z.* **259**:4 (2008), 755–774. MR Zbl

[Dornhoff 1971] L. Dornhoff, *Group representation theory, Part A: Ordinary representation theory*, Pure and Applied Mathematics **7**, Marcel Dekker, New York, 1971. MR Zbl

[Gorenstein et al. 1998] D. Gorenstein, R. Lyons, and R. Solomon, *The classification of the finite simple groups, Number 3, Part I, Chapter A: Almost simple K-groups*, Mathematical Surveys and Monographs **40**, American Mathematical Society, Providence, RI, 1998. MR Zbl

[Griess 1978] R. L. Griess, Jr., "Finite groups whose involutions lie in the center", *Quart. J. Math. Oxford Ser.* (2) **29**:115 (1978), 241–247. MR Zbl

[Guralnick et al. 2011] R. M. Guralnick, G. Navarro, and P. H. Tiep, "Real class sizes and real character degrees", *Math. Proc. Cambridge Philos. Soc.* **150**:1 (2011), 47–71. MR Zbl

[Isaacs 2008] I. M. Isaacs, *Finite group theory*, Graduate Studies in Mathematics **92**, American Mathematical Society, Providence, RI, 2008. MR Zbl

[Isaacs and Navarro 2010] I. M. Isaacs and G. Navarro, "Normal p-complements and fixed elements", *Arch. Math.* (*Basel*) **95**:3 (2010), 207–211. MR Zbl

[Itô 1953] N. Itô, "On finite groups with given conjugate types, I", *Nagoya Math. J.* **6** (1953), 17–28. MR Zbl

[Itô 1970] N. Itô, "On finite groups with given conjugate types, II", *Osaka J. Math.* **7** (1970), 231–251. MR Zbl

[Iwasaki 1980] S. Iwasaki, "On finite groups with exactly two real conjugate classes", *Arch. Math.* (*Basel*) **33**:6 (1980), 512–517. MR Zbl

[Liebeck and O'Brien 2007] M. W. Liebeck and E. A. O'Brien, "Finding the characteristic of a group of Lie type", *J. Lond. Math. Soc.* (2) **75**:3 (2007), 741–754. MR Zbl

[Moretó and Navarro 2008] A. Moretó and G. Navarro, "Groups with three real valued irreducible characters", *Israel J. Math.* **163** (2008), 85–92. MR Zbl

[Navarro and Tiep 2010] G. Navarro and P. H. Tiep, "Degrees of rational characters of finite groups", *Adv. Math.* **224**:3 (2010), 1121–1142. MR Zbl

[Navarro et al. 2009] G. Navarro, L. Sanus, and P. H. Tiep, "Real characters and degrees", *Israel J. Math.* **171** (2009), 157–173. MR Zbl

[Steinberg 1951] R. Steinberg, "The representations of $\mathrm{GL}(3, q)$, $\mathrm{GL}(4, q)$, $\mathrm{PGL}(3, q)$, and $\mathrm{PGL}(4, q)$", *Canadian J. Math.* **3** (1951), 225–235. MR Zbl

[Suzuki 1962] M. Suzuki, "On a class of doubly transitive groups", *Ann. of Math.* (2) **75** (1962), 105–145. MR Zbl

[Suzuki 1965] M. Suzuki, "Finite groups in which the centralizer of any element of order 2 is 2-closed", *Ann. of Math.* (2) **82** (1965), 191–212. MR Zbl

[Tong-Viet 2013] H. P. Tong-Viet, "Groups with some arithmetic conditions on real class sizes", *Acta Math. Hungar.* **140**:1-2 (2013), 105–116. MR Zbl

Communicated by Pham Huu Tiep
Received 2018-05-12 Revised 2018-08-14 Accepted 2018-08-15

tongviet@math.binghamton.edu Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902-6000, 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 12 No. 10 2018