Algebra & Number Theory

Volume 17
2023
No. 6

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2023 is US $485/year for the electronic version, and $705/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 17:6 (2023)
https://doi.org/10.2140/ant.2023.17.1127

On Héthelyi–Külshammer's conjecture for principal blocks

Nguyen Ngoc Hung and A. A. Schaeffer Fry

We prove that the number of irreducible ordinary characters in the principal p-block of a finite group G of order divisible by p is always at least $2\sqrt{p-1}$. This confirms a conjecture of Héthelyi and Külshammer (2000) for principal blocks and provides an affirmative answer to Brauer's problem 21 (1963) for principal blocks of bounded defect. Our proof relies on recent works of Maróti (2016) and Malle and Maróti (2016) on bounding the conjugacy class number and the number of p'-degree irreducible characters of finite groups, earlier works of Broué, Malle and Michel (1993) and Cabanes and Enguehard (2004) on the distribution of characters into unipotent blocks and e-Harish-Chandra series of finite reductive groups, and known cases of the Alperin–McKay conjecture.

1. Introduction

Bounding the number $k(G)$ of conjugacy classes of a finite group G in terms of a certain invariant associated to G is a fundamental problem in group representation theory. An equally important problem in modular representation theory is to bound the number $k(B)$ of ordinary irreducible characters in a block B of G. It is not surprising that these two problems are closely related to each other. For instance, the p-solvable case of the Brauer's celebrated $k(B)$-conjecture [Brauer 1963, Problem 20], which asserts that $k(B)$ is bounded *above* by the order of a defect group for B, was known to be equivalent to the coprime $k(GV)$-problem (by [Nagao 1962]), which in turn was eventually solved in [Gluck et al. 2004]; see also [Schmid 2007]. While there have been a number of results on upper bounds for $k(B)$ [Brauer and Feit 1959; Robinson 2004; Sambale 2017; Malle 2018], not much has been done on lower bounds.

Let p be a prime dividing the order of G. A result of Brauer [1942] on characters and blocks of groups G of order divisible by p but not by p^2 implies that $k(G) \geq 2\sqrt{p-1}$ for those groups, and the bound was later conjectured to be true for all finite groups. After several partial results [Héthelyi and Külshammer 2000; 2003; Malle 2006; Keller 2009; Héthelyi et al. 2011], the conjecture was finally proved by Maróti [2016]. In the proof of the conjecture for solvable groups, Héthelyi and Külshammer [2000] speculated that "*perhaps it is even true that $k(B) \geq 2\sqrt{p-1}$ for every p-block B of positive defect, where $k(B)$*

The first author thanks Gunter Malle for several stimulating discussions on the relation between the relative Weyl groups of e-split Levi subgroups and maximal tori, and both authors thank him for his comments on an earlier version of the manuscript. The second author was partially supported by the National Science Foundation under Grants No. DMS-1801156 and DMS-2100912. Finally, we thank the anonymous referee for his/her comments and careful reading of the manuscript.
MSC2020: 20C15, 20C20, 20C33, 20D06.
Keywords: finite groups, principal blocks, characters, Héthelyi–Külshammer conjecture, Alperin–McKay conjecture.

denotes the number of irreducible ordinary characters in B". Of course, they were aware of blocks of defect zero, which have a unique irreducible ordinary character (whose degree has the same p-part as the order of the group) and a unique irreducible Brauer character as well; see [Navarro 1998, Theorem 3.18].

The main aim of this paper is to confirm Héthelyi and Külshammer's conjecture for principal blocks. Throughout, we use $B_0(G)$ to denote the principal p-block of G.

Theorem 1.1. *Let G be a finite group and p a prime such that $p \mid |G|$. Then $k(B_0(G)) \geq 2\sqrt{p-1}$.*

Problem 21 in Brauer's famous list [1963] asks whether there exists a function $f(q)$ on prime powers q such that $f(q) \to \infty$ for $q \to \infty$ and that $k(B) \geq f(p^{d(B)})$ for every p-block B of defect $d(B) > 0$. Our Theorem 1.1 provides an affirmative answer to this question for principal blocks of *bounded* defect. See [Külshammer 1990] for more discussion on this problem.

One may expect to improve the bound in Theorem 1.1 when the power of p in $|G|$ is large. Kovács and Leedham-Green [1986] constructed, for each odd prime p, a p-group P of order p^p with $k(P) = (p^3 - p^2 + p + 1)/2$. Therefore, the bound $k(B_0(G)) \geq 2\sqrt{p-1}$ cannot be replaced by $k(B_0(G)) \geq p^3$, for example, even when any given large power of p divided the group order.

Building upon the ideas of Maróti [2016] and the subsequent paper by Malle and Maróti [2016] on bounding the number of p'-degree irreducible characters in a finite group, we observe that Héthelyi and Külshammer's conjecture for principal blocks essentially reduces to bounding the number of irreducible ordinary characters in principal blocks of almost simple groups, as well as bounding the number of orbits of irreducible characters in principal blocks of simple groups under the action of their automorphism groups.

Theorem 1.2. *Let S be a finite nonabelian simple group and p a prime such that $p \mid |S|$. Let G be an almost simple group with socle S such that $p \nmid |G/S|$. Then*:

(i) $k(B_0(G)) \geq 2\sqrt{p-1}$. *Moreover*, $k(B_0(G)) > 2\sqrt{p-1}$ *if S does not have cyclic Sylow p-subgroups.*

(ii) *Assume further that $p \geq 11$ and S does not have cyclic Sylow p-subgroups. Then the number of* $\mathrm{Aut}(S)$*-orbits on* $\mathrm{Irr}(B_0(S))$ *is at least* $2(p-1)^{1/4}$.

As we will explain in the next section, Theorem 1.1 is a consequence of [Maróti 2016] and the well-known Alperin–McKay conjecture, which asserts that the number of irreducible characters of height 0 in a block B of a finite group G coincides with the number of irreducible characters of height 0 in the Brauer correspondent of B of the normalizer of a defect subgroup for B in G. We take advantage of the recent advances on the conjecture in the proof of our results, particularly the fact that Späth's inductive Alperin–McKay conditions hold for all p-blocks with cyclic defect groups [Späth 2013; Koshitani and Späth 2016]. This explains why simple groups with cyclic Sylow p-subgroups are excluded in Theorem 1.2(ii). Additionally, we take advantage of recent results on the possible structure of defect groups of principal blocks with few ordinary characters [Koshitani and Sakurai 2021; Rizo et al. 2021], and this explains why the smaller values of p are excluded in Theorem 1.2(ii).

Theorem 1.2 turns out to be straightforward for alternating groups or groups of Lie type in characteristic p, but highly nontrivial for groups of Lie type in characteristic not equal to p. We make use of Cabanes and Enguehard's results [1994; 2004] on the distribution of characters into unipotent blocks and Broué, Malle and Michel's results [Broué et al. 1993] on the compatibility between the distributions of unipotent characters into unipotent blocks and e-Harish-Chandra series to obtain a general bound for the number of $\mathrm{Aut}(S)$-orbits of characters in $\mathrm{Irr}(B_0(S))$ in terms of certain data associated to S, for S a simple group of Lie type, see Theorem 5.4. We hope this result will be useful in other purposes.

The next result classifies groups for which $k(B_0(G))$ is minimal in the sense of Theorem 1.1.

Theorem 1.3. *Let G be a finite group and p a prime. Let P be a Sylow p-subgroup of G. Then $k(B_0(G))=2\sqrt{p-1}$ if and only if $\sqrt{p-1}\in\mathbb{N}$ and $\boldsymbol{N}_G(P)/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P))$ is isomorphic to the Frobenius group $C_p\rtimes C_{\sqrt{p-1}}$.*

We remark that, in the situation of Theorem 1.3, the number of p'-degree irreducible characters in $B_0(G)$ is also equal to $2\sqrt{p-1}$. In general, if a p-block B of a finite group has an abelian defect group, then every ordinary irreducible character of B has height zero. This is the "if direction" of Brauer's height-zero conjecture, which is now known to be true, thanks to the work of Kessar and Malle [2013]. Theorem 1.1 therefore implies that if $P\in\mathrm{Syl}_p(G)$ is abelian and nontrivial then $k_0(B_0(G))\geq 2\sqrt{p-1}$, where $k_0(B)$ denotes the number of height zero ordinary irreducible characters of a block B.

Theorems 1.1 and 1.3 are useful in the study of principal blocks with few height zero ordinary irreducible characters. In fact, using them, we are able to show in [Hung et al. 2023] that $k_0(B_0(G))=3$ if and only if $P\cong C_3$, and that $k_0(B_0(G))=4$ if and only if $|P/P'|=4$ or $P\cong C_5$ and $\boldsymbol{N}_G(P)/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P))$ is isomorphic to the dihedral group D_{10}. These results have been known only in the case $p\leq 3$; see [Navarro et al. 2018, Theorems A and C].

The paper is organized as follows. In Section 2, we recall some known results on the Alperin–McKay conjecture and prove that our results follow when all the nonabelian composition factors of G have cyclic Sylow p-subgroups. We also prove Theorem 1.2 for the sporadic simple groups and groups of Lie type defined in characteristic p in Section 2. The alternating groups are treated in Section 3. Section 4 takes care of the case when the Sylow p-subgroups of S are nonabelian. Sections 6, 7, and 8 are devoted to proving Theorem 1.2 for simple groups of Lie type defined in characteristics different from p. To do so, in Section 5, we prove a bound for the number of $\mathrm{Aut}(S)$-orbits of characters in $\mathrm{Irr}(B_0(S))$. Finally, we finish the proofs of Theorems 1.1 and 1.3 in Section 9.

2. Some first observations

In this section we make some observations toward the proofs of the main results.

2A. ***The Alperin–McKay conjecture.*** The well-known Alperin–McKay (AM) conjecture predicts that the number of irreducible characters of height zero in a block B of a finite group G coincides with the number of irreducible characters of height zero in the Brauer correspondent of B of the normalizer of a defect subgroup of B in G. For the principal blocks, the conjecture is equivalent to

$$k_{p'}(B_0(G)) = k_{p'}(B_0(\boldsymbol{N}_G(P))),$$

where P is a Sylow p-subgroup of G and $k_{p'}(B_0(G))$ denotes the number of p'-degree irreducible ordinary characters in $B_0(G)$.

On the other hand, if $p \mid |G|$, we have

$$\begin{aligned} k_{p'}(B_0(\boldsymbol{N}_G(P))) &\geq k_{p'}(B_0(\boldsymbol{N}_G(P)/P')) \\ &= k(B_0(\boldsymbol{N}_G(P)/P')) \\ &= k(B_0((\boldsymbol{N}_G(P)/P')/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P)/P'))) \\ &= k((\boldsymbol{N}_G(P)/P')/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P)/P')) \\ &\geq 2\sqrt{p-1}, \end{aligned}$$

where the first inequality follows from [Navarro 1998, page 137], the first equality follows from the fact that every irreducible ordinary character of $\boldsymbol{N}_G(P)/P'$ has p'-degree, the last two equalities follow from [loc. cit., Theorem 9.9] and Fong's theorem (see [loc. cit., Theorem 10.20]), and the last inequality follows from [Maróti 2016]. Therefore, if the AM conjecture holds for G and p, then the number of p'-degree irreducible ordinary characters in $B_0(G)$ is bounded below by $2\sqrt{p-1}$.

From this, we see that Theorems 1.1 and 1.2(i) hold if the AM conjecture holds for (G, p). We now prove that the same is true for Theorem 1.3. Note that the "if" implication of this theorem is clear. Assume that the AM conjecture holds for $B_0(G)$ and $k(B_0(G)) = 2\sqrt{p-1}$ for some prime p such that $\sqrt{p-1} \in \mathbb{N}$. Then, as seen above, we have

$$2\sqrt{p-1} = k(B_0(G)) \geq k((\boldsymbol{N}_G(P)/P')/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P)/P')) \geq 2\sqrt{p-1},$$

implying

$$k((\boldsymbol{N}_G(P)/P')/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P)/P')) = 2\sqrt{p-1},$$

and thus $(\boldsymbol{N}_G(P)/P')/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P)/P')$ is isomorphic to the Frobenius group $C_p \rtimes C_{\sqrt{p-1}}$, by [Maróti 2016, Theorem 1]. In particular, $P/P' \cong C_p$, implying that $P \cong C_p$, and hence it follows that $\boldsymbol{N}_G(P)/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P))$ is isomorphic to the Frobenius group $C_p \rtimes C_{\sqrt{p-1}}$, as wanted.

The AM conjecture is known to be true when G has a cyclic Sylow p-subgroup by Dade's theory [1966]. In fact, by [Späth 2013; Koshitani and Späth 2016], the so-called *inductive Alperin–McKay conditions* are satisfied for all blocks with cyclic defect groups. Therefore, we have:

Lemma 2.1 (Koshitani–Späth). *Let p be a prime. Assume that all the composition factors of a finite group G have cyclic Sylow p-subgroups. Then the Alperin–McKay conjecture holds for G and p, and thus Theorems 1.1, 1.2(i), and 1.3 hold for G and p.*

Note that the linear groups $\mathrm{PSL}_2(q)$, the Suzuki groups ${}^2B_2(2^{2f+1})$ and the Ree groups ${}^2G_2(3^{2f+1})$ all have cyclic Sylow p-subgroups for odd p different from the defining characteristic of the group. So

Theorem 1.2 automatically follows from Lemma 2.1 for these groups in characteristic not equal to p, when p is odd.

2B. ***Small blocks.*** Blocks with a small number of ordinary characters have been studied significantly in the literature. In particular, the possible structure of defect groups of principal blocks with at most 5 ordinary irreducible characters are now known; see [Brandt 1982; Belonogov 1990; Koshitani and Sakurai 2021; Rizo et al. 2021]. (B. Sambale informed us that he and S. Koshitani think that Belonogov's work [1990] for the case $k(B_0) = 3$ is not conclusive. However, this case has been recently reproved in [Koshitani and Sakurai 2021, Section 3].) Using these results, we can easily confirm our results for $p \leq 7$. For instance, to prove Theorems 1.1 and 1.2 for $p = 7$ it is enough to assume that $k(B_0(G)) \leq 4$, but by going through the list of possible defect groups of $B_0(G)$, we then have $\mathrm{Syl}_p(G) \in \{1, C_2, C_3, C_2 \times C_2, C_4, C_5\}$, which cannot happen. To prove Theorem 1.3 for $p < 7$ we note that if $p = 5$ and $k(B_0(G)) = 4$ then $P = C_5$; and if $p = 2$ and $k(B_0(G)) = 2$ then $P = C_2$, in both of which cases P is cyclic, and thus the result of Section 2A applies.

Therefore we will assume from now on that $p \geq 11$, unless stated otherwise.

2C. ***Sporadic groups and the Tits group.*** We remark that Theorem 1.2 can be confirmed directly using [Conway et al. 1985; Jansen et al. 1995] or [GAP 2020] for sporadic simple groups and the Tits group. Therefore, we are left with the alternating groups and groups of Lie type, which will be treated in the subsequent sections.

2D. ***Groups of Lie type in characteristic p.*** Let S be a simple group of Lie type defined over the field of $q = p^f$ elements, where p is a prime and f a positive integer. According to results of Dagger and Humphreys on defect groups of finite reductive groups in defining characteristic; see [Cabanes 2018, Proposition 1.18 and Theorem 3.3] for instance, S has only two p-blocks. The only nonprincipal block is a defect-zero block containing only the Steinberg character of S. Therefore,

$$k(B_0(S)) = k(S) - 1.$$

Let $\boldsymbol{G}$ be a simple algebraic group of simply connected type and let F be a Steinberg endomorphism on $\boldsymbol{G}$ such that $S = X/\boldsymbol{Z}(X)$, where $X = \boldsymbol{G}^F$. Assume that the rank of $\boldsymbol{G}$ is r. By a result of Steinberg (see [Fulman and Guralnick 2012, Theorem 3.1]), X has at least q^r semisimple conjugacy classes, and thus $k(X) > q^r$. It follows that

$$k(B_0(S)) > \frac{q^r}{|\boldsymbol{Z}(X)|} - 1,$$

which yields $k(B_0(S)) \geq \lfloor q^r/|\boldsymbol{Z}(X)| \rfloor$. Using the values of $|\boldsymbol{Z}(X)|$ and $|\mathrm{Out}(S)|$ available in [Conway et al. 1985, page xvi], it is straightforward to check that $\lfloor q^r/|\boldsymbol{Z}(X)| \rfloor \geq 2\sqrt{p-1}|\mathrm{Out}(S)|$, proving Theorem 1.2 for the relevant S and p.

3. Alternating groups

In this section we prove Theorem 1.2 for the alternating groups. The background on block theory of symmetric and alternating groups can be found in [Olsson 1993] for instance.

The ordinary irreducible characters of S_n are naturally labeled by partitions of n. Two characters are in the same p-block if and only if their corresponding partitions have the same p-cores, which are obtained from the partitions by successive removals of rim p-hooks until no p-hook is left. Therefore, p-blocks of S_n are in one-to-one correspondence with p-cores of partitions of n.

Let B be a p-block of S_n. The number $k(B)$ of ordinary irreducible characters in B turns out to depend only on p and the so-called *weight* of B, which is defined to be $w(B) := (n-|\mu|)/p$, where μ is the p-core corresponding to B under the aforementioned correspondence. In fact,

$$k(B) = k(p, w(B)) := \Sigma_{(w_0, w_1, \dots, w_{p-1})} \pi(w_0)\pi(w_1)\cdots\pi(w_{p-1}),$$

where $(w_0, w_1, \dots, w_{p-1})$ runs through all p-tuples of nonnegative integers such that $w(B) = \Sigma_{i=0}^{p-1} w_i$ and $\pi(x)$ is the number of partitions of x; see [Olsson 1993, Proposition 11.4]. Note that $k(p, w(B))$ is precisely the number of p-tuples of partitions of $w(B)$.

For the principal block $B_0(\mathsf{S}_n)$ of S_n, we have $w(B_0(\mathsf{S}_n)) = \lfloor n/p \rfloor$, which is at least 1 by the assumption $p \mid |S|$. It follows that

$$k(B_0(\mathsf{S}_n)) \geq k(p, 1) = p \geq 2\sqrt{p-1}.$$

Moreover, according to [Olsson 1992, Proposition 2.8], when p is odd and $\widetilde{B}$ is a block of A_n covered by B, then B and $\widetilde{B}$ have the same number of irreducible ordinary characters (and indeed the same number of irreducible Brauer characters as well). In particular, when p is odd, we have $k(B_0(\mathsf{A}_n)) = k(B_0(\mathsf{S}_n)) \geq 2\sqrt{p-1}$, which proves Theorem 1.2(i) for the alternating groups.

For part (ii) of Theorem 1.2, recall that $p \geq 11$, and thus $n \geq 11$ and $\mathrm{Aut}(S) = \mathsf{S}_n$. The number of S_n-orbits on $\mathrm{Irr}(B_0(\mathsf{A}_n))$ is at least $1 + (k(B_0(\mathsf{A}_n)) - 1)/2$, which in turn is at least

$$1 + \frac{p-1}{2} = \frac{p+1}{2} > 2(p-1)^{1/4},$$

and this proves Theorem 1.2(ii) for the alternating groups.

4. Groups of Lie type: the nonabelian Sylow case

In this section, we let $\boldsymbol{G}$ be a simple algebraic group of adjoint type and F a Steinberg endomorphism on $\boldsymbol{G}$ such that $S \cong [\mathbb{G}, \mathbb{G}]$ where $\mathbb{G} := \boldsymbol{G}^F$. Let ℓ be a prime different from p and assume $q = \ell^f$ is the absolute value of all eigenvalues of F on the character group of an F-stable maximal torus of G. Recall that we are assuming $p \geq 11$.

In this section we prove Theorem 1.2 for those S of Lie type in characteristic different from p such that the Sylow p-subgroups of $\mathbb{G}$ are nonabelian.

In that case, there are then more than one $d \in \mathbb{N}$ such that $p \mid \Phi_d(q)$ with Φ_d dividing the order polynomial of $(\boldsymbol{G}, F)$. Here, as usual, Φ_d denotes the d-th cyclotomic polynomial. In fact, if there a unique such d, then a Sylow p-subgroup of $\mathbb{G}$ is contained in a Sylow d-torus of $\mathbb{G}$, and hence is abelian; see [Malle and Testerman 2011, Theorem 25.14].

Let $e_p(q)$ denote the multiplicative order of q modulo p. Note that, by [Malle and Testerman 2011, Lemma 25.13], $p \mid \Phi_d(q)$ if and only if $d = e_p(q)p^i$ for some $i \geq 0$. Therefore, as there is more than one $d \in \mathbb{N}$ such that $p \mid \Phi_d(q)$, we must have $p \mid d$ for some $d \in \mathbb{N}$ such that Φ_d divides the order polynomial of $(\boldsymbol{G}, F)$. The fact that $p \geq 11$ then rules out the cases when $\boldsymbol{G}$ is of exceptional type and thus we are left with only the classical types. That is, $\mathbb{G} = \mathrm{PGL}_n(q)$, $\mathrm{PGU}_n(q)$, $\mathrm{SO}_{2n+1}(q)$, $\mathrm{PCSp}_{2n}(q)$, or $\mathrm{P}(\mathrm{CO}_{2n}^{\pm}(q))^0$.

For $\mathbb{G} = \mathrm{PGL}_n(q)$ or $\mathrm{PGU}_n(q)$, we define e to be the smallest positive integer such that $p \mid (q^e - (\epsilon)^e)$ ($\epsilon = 1$ for linear groups and $\epsilon = -1$ for unitary groups), so that $e = e_p(q)$ when $\mathbb{G} = \mathrm{PGL}_n(q)$ or $\mathbb{G} = \mathrm{PGU}_n(q)$ and $4 \mid e_p(q)$, $e = e_p(q)/2$ when $\mathbb{G} = \mathrm{PGU}_n(q)$ and $2 \mid e_p(q)$ but $4 \nmid e_p(q)$, and $e = 2e_p(q)$ when $\mathbb{G} = \mathrm{PGU}_n(q)$ and $2 \nmid e_p(q)$. For $\mathbb{G} = \mathrm{SO}_{2n+1}(q)$, $\mathrm{PCSp}_{2n}(q)$, or $\mathrm{P}(\mathrm{CO}_{2n}^{\pm}(q))^0$, we define e to be the smallest positive integer such that $p \mid (q^e \pm 1)$, so that $e = e_p(q)$ when $e_p(q)$ is odd and $e = e_p(q)/2$ when $e_p(q)$ is even.

Let $n = we + m$ where w and m are integers with $0 \leq m < e$. We claim that $p \leq w$. To see this, first assume that $\mathbb{G} = \mathrm{PGL}_n(q)$. Then, as mentioned above, $ep \leq n$, which implies that $ep < (w+1)e$, and thus $p \leq w$. Next, assume that $\mathbb{G} = \mathrm{SO}_{2n+1}(q)$, $\mathrm{PCSp}_{2n}(q)$, or $\mathrm{P}(\mathrm{CO}_{2n}^{\pm}(q))^0$. If $e = e_p(q)$ is odd, then since $p \mid (q^e - 1)$ and $\gcd(q^e - 1, q^i + 1) \leq 2$ for every $i \in \mathbb{N}$, we have $p \mid (q^j - 1)$ for some $e < j \leq n$, and it follows that $ep \leq n$, implying $p \leq w$. On the other hand, if $2e = e_p(q)$ is even then

$$2ep = e_p(q)p \leq 2n < 2(w+1)e,$$

which also implies that $p \leq w$. Finally, assume $\mathbb{G} = \mathrm{PGU}_n(q)$. The case $4 \mid e_p(q)$ is argued as in the case $S = \mathrm{PGL}_n(q)$; the case $2 \mid e_p(q)$ but $4 \nmid e_p(q)$ is argued as in the case $S = \mathrm{SO}_{2n+1}(q)$ and $2 \mid e_p(q)$. For the last case $2 \nmid e_p(q)$, we have $ep/2 = e_p(q)p$, and in order for $\Phi_{e_p(q)p}$ to divide the generic order of $|\mathrm{PGU}_n(q)|$, $e_p(q)p \leq n/2$, and hence it follows that $ep \leq n$, which also implies that $p \leq w$. The claim is fully proved.

Since p is good for $\boldsymbol{G}$, by [Broué et al. 1993, Theorem 3.2] and [Cabanes and Enguehard 1994, main theorem], the number of unipotent characters of $\mathbb{G}$ in the principal block $B_0(\mathbb{G})$ is equal to $k(W_e)$ — the number of irreducible complex characters of the relative Weyl group W_e of a Sylow $e_p(q)$-torus of $\mathbb{G}$. This W_e is the wreath product $C_e \wr \mathsf{S}_w$ when $\boldsymbol{G}$ is of type A and is a subgroup of index 1 or 2 of $C_{2e} \wr \mathsf{S}_w$ when $\boldsymbol{G}$ is of type B, C, or D; see [Broué et al. 1993, Section 3A]. In any case, W_e has a quotient S_w, so we have that the number of unipotent characters in $\mathrm{Irr}(B_0(\mathbb{G}))$ is at least $k(\mathsf{S}_w) = \pi(w)$, which in turns is at least $\pi(p)$ as $p \leq w$. Since every unipotent character of $\mathbb{G}$ restricts irreducibly to S and $B_0(\mathbb{G})$ covers a unique block of S, it follows that the number of unipotent characters in $\mathrm{Irr}(B_0(S))$ is at least $\pi(p)$.

By a result of Lusztig (see [Malle 2008, Theorem 2.5]), every unipotent character of a simple group of Lie type lies in a $\mathrm{Aut}(S)$-orbit of length at most 3. (In fact, every $\mathrm{Aut}(S)$-orbit on unipotent characters of S has length 1 or 2, except when $S = P\Omega_8^+(q)$ whose graph automorphism of order 3 produces two orbits of length 3.) Therefore, together with the conclusion of the previous paragraph, we deduce that the number of $\mathrm{Aut}(S)$-orbits on $\mathrm{Irr}(B_0(S))$ is at least $\pi(p)/3$. This bound is greater than $2\sqrt{p-1}$ when $p \geq 11$, as required.

5. A general bound for the number of Aut(S)-orbits on Irr($B_0(S)$)

The aim of this section is to obtain a general bound for the number of $\mathrm{Aut}(S)$-orbits on irreducible ordinary characters in the principal block of S, for S a simple group of Lie type.

5A. *Semisimple characters.* Before continuing with our proof of Theorem 1.2 for groups of Lie type, we recall some background on certain characters known as *semisimple characters* and the fact that they fall into the principal block in a certain situation. Background on character theory of finite reductive groups can be found in [Carter 1985; Cabanes and Enguehard 2004; Digne and Michel 1991]. Let $\boldsymbol{G}$ be a connected reductive group defined over $\mathbb{F}_q$ and F an associated Frobenius endomorphism on $\boldsymbol{G}$. Let $\boldsymbol{G}^*$ be an algebraic group with a Frobenius endomorphism which, for simplicity, we denote by the same F, such that $(\boldsymbol{G}, F)$ is in duality to $(\boldsymbol{G}^*, F)$.

Let t be a semisimple element of $(\boldsymbol{G}^*)^F$. The rational Lusztig series $\mathcal{E}(\boldsymbol{G}^F, (t))$ associated to the $(\boldsymbol{G}^*)^F$-conjugacy class (t) of t is defined to be the set of irreducible characters of $\boldsymbol{G}^F$ occurring in some Deligne–Lusztig character $R_{\boldsymbol{T}}^{\boldsymbol{G}}\theta$, where $\boldsymbol{T}$ is an F-stable maximal torus of $\boldsymbol{G}$ and $\theta \in \mathrm{Irr}(\boldsymbol{T}^F)$ such that $(\boldsymbol{T}, \theta)$ corresponds in duality to a pair $(\boldsymbol{T}^*, s)$ with $s \in \boldsymbol{T}^* \cap (t)$. Here we recall from [Digne and Michel 1991, Proposition 13.13] that there is a one-to-one duality correspondence between $\boldsymbol{G}^F$-conjugacy classes of pairs $(\boldsymbol{T}, \theta)$, where $\boldsymbol{T}$ is an F-stable maximal torus of $\boldsymbol{G}$ and $\theta \in \mathrm{Irr}(\boldsymbol{T}^F)$, and the $(\boldsymbol{G}^*)^F$-conjugacy classes of pairs $(\boldsymbol{T}^*, s)$, where $\boldsymbol{T}^*$ is dual to $\boldsymbol{T}$ and $s \in (\boldsymbol{T}^*)^F$.

We continue to let t be a semisimple element of $(\boldsymbol{G}^*)^F$ and assume furthermore that $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ is a Levi subgroup of $\boldsymbol{G}^*$. Let $\boldsymbol{G}(t)$ be an F-stable Levi subgroup of $\boldsymbol{G}$ in duality with $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ and $\boldsymbol{P}$ be a parabolic subgroup of $\boldsymbol{G}$ for which $\boldsymbol{G}(t)$ is the Levi complement. The twisted induction $R^{\boldsymbol{G}}_{\boldsymbol{G}(t)\subseteq \boldsymbol{P}}$ and the multiplication by $\hat{t}$, a certain linear character of $\mathrm{Irr}(\boldsymbol{G}(t)^F)$ naturally defined by t (see [Cabanes and Enguehard 2004, (8.19)]), then induce a bijection between the Lusztig series $\mathcal{E}(\boldsymbol{G}(t)^F, 1)$ and $\mathcal{E}(\boldsymbol{G}^F, (t))$; see [Cabanes and Enguehard 2004, Proposition 8.26 and Theorem 8.27]. In fact, for each $\lambda \in \mathcal{E}(\boldsymbol{G}(t)^F, 1)$, one has

$$\varepsilon_{\boldsymbol{G}}\varepsilon_{\boldsymbol{G}(t)} R^{\boldsymbol{G}}_{\boldsymbol{G}(t)\subseteq P}(\hat{t}\lambda) \in \mathcal{E}(\boldsymbol{G}^F, (t)),$$

where $\varepsilon_{\boldsymbol{G}} := (-1)^{\sigma(\boldsymbol{G})}$ with $\sigma(\boldsymbol{G})$ the $\mathbb{F}_q$-rank of $\boldsymbol{G}$. Taking λ to be trivial, we have the character

$$\chi_{(t)} := \varepsilon_{\boldsymbol{G}}\varepsilon_{\boldsymbol{G}(t)} R^{\boldsymbol{G}}_{\boldsymbol{G}(t)\subseteq P}(\hat{t}\mathbf{1}_{\boldsymbol{G}(t)^F}) \in \mathcal{E}(\boldsymbol{G}^F, (t)),$$

which is often referred to as a semisimple character of $\boldsymbol{G}^F$, of degree

$$\chi_{(t)}(1) = |(\boldsymbol{G}^*)^F : \boldsymbol{C}_{\boldsymbol{G}^{*F}}(t)|_{\ell'},$$

where ℓ is the defining characteristic of $\boldsymbol{G}$; see [Digne and Michel 1991, Theorem 13.23].

By [Cabanes and Enguehard 2004, Theorem 9.12], every element of $B_0(\boldsymbol{G}^F)$ lies in a Lusztig series $\mathcal{E}(\boldsymbol{G}^F, (t))$ where t is a p-element of $\boldsymbol{G}^{*F}$. Hence one might ask which such t indeed produce semisimple characters that contribute to the principal block. We will see in the following theorem that in a certain nice situation which is indeed enough for our purpose, the centralizer $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ is a Levi subgroup of $\boldsymbol{G}^*$, and thus the semisimple character $\chi_{(t)}$ associated to (t) is well-defined and belongs to $B_0(\boldsymbol{G}^F)$.

Recall that a prime p is good for $\boldsymbol{G}$ if it does not divide the coefficients of the highest root of the root system associated to $\boldsymbol{G}$. The following result, mainly due to Hiss [1990, Corollary 3.4] and Cabanes and Enguehard [2004, Theorem 21.13], will be very useful in later sections.

Theorem 5.1. *Let $(\boldsymbol{G}, F)$ be a connected reductive group defined over $\mathbb{F}_q$. Let p be a prime not dividing q. Let t be a p-element of $\boldsymbol{G}^{*F}$. If $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ is connected and p is good for $\boldsymbol{G}$, then the semisimple character $\chi_{(t)} \in \mathrm{Irr}(\boldsymbol{G}^F)$ belongs to the principal p-block of $\boldsymbol{G}^F$. Also, if $\boldsymbol{Z}(\boldsymbol{G})$ is connected, then $\chi_{(t)}$ belongs to the principal block of $\boldsymbol{G}^F$.*

We note that Theorem 5.1 can also be deduced from [Cabanes and Enguehard 1994, main theorem], but with more restricted conditions on p.

Building on Theorem 5.1, we observe that the principal block of S contains *many* irreducible semisimple characters. By controlling the length of $\mathrm{Aut}(S)$-orbits on these characters, we are able to bound below the number of $\mathrm{Aut}(S)$-orbits on $\mathrm{Irr}(B_0(S))$. The bound turns out to be enough to prove Theorem 1.2, at least in the case when the Sylow p-subgroups of the group of inner and diagonal automorphisms of S are abelian but non-cyclic, which is precisely the case we need after Sections 2A and 4.

5B. ***Specific setup for our purpose.*** For the rest of this section, we will work with the following setup: $\boldsymbol{G}$ is a simple algebraic group of adjoint type defined over $\mathbb{F}_q$ and F a Steinberg endomorphism on $\boldsymbol{G}$ such that $S = [\mathbb{G}, \mathbb{G}]$ with $\mathbb{G} = \boldsymbol{G}^F$. Let $(\boldsymbol{G}^*, F^*)$ be the dual pair of $(\boldsymbol{G}, F)$ and for simplicity we will use the same notation F for F^*, and thus $\boldsymbol{G}^*$ is a simple algebraic group of simply connected type and $S = \mathbb{G}^*/\boldsymbol{Z}(\mathbb{G}^*)$, where $\mathbb{G}^* := (\boldsymbol{G}^*)^F$.

Theorem 5.1 has the following consequence.

Lemma 5.2. *Assume the above notation. Let p be a prime not dividing q. For every p-element t of $\mathbb{G}^*$, the semisimple character $\chi_{(t)} \in \mathcal{E}(\mathbb{G}, (t))$ belongs to the principal block of $\mathbb{G}$.*

Proof. Since $\boldsymbol{G}^*$ has connected center, the lemma follows from Theorem 5.1; see also [Bessenrodt et al. 2007, Lemma 3.1]. □

5C. ***Orbits of semisimple characters.*** Knowing that the semisimple characters $\chi_{(t)} \in \mathrm{Irr}(\mathbb{G})$ associated to $\mathbb{G}^*$-conjugacy classes of p-elements all belong to $B_0(\mathbb{G})$, we now wish to control the number of orbits of the action of the automorphism group $\mathrm{Aut}(S)$ on these characters. By a result of Bonnafé [Navarro et al. 2008, Section 2] (see also [Taylor 2018, Section 7]), this action turns out to be well-behaved.

Let $\alpha \in \mathrm{Aut}(\mathbb{G})$, which in our situation will be a product of a field automorphism and a graph automorphism. It is easy to see that α then can be extended to a bijective morphism $\bar{\alpha} : \boldsymbol{G} \to \boldsymbol{G}$ such that $\bar{\alpha}$ commutes with F. This $\bar{\alpha}$ induces a bijective morphism $\bar{\alpha}^* : \boldsymbol{G}^* \to \boldsymbol{G}^*$ which commutes with the dual of F. The restriction of $\bar{\alpha}^*$ to $\mathbb{G}^*$, which we denote by α^*, is now an automorphism of $\mathbb{G}^*$. Recall that $\alpha \in \mathrm{Aut}(\mathbb{G})$ induces a natural action on $\mathrm{Irr}(\mathbb{G})$ by $\chi^\alpha := \chi \circ \alpha^{-1}$. By [Navarro et al. 2008, Section 2], α maps the Lusztig series $\mathcal{E}(\mathbb{G}, (t))$ of $\mathbb{G}$ associated to (t) to the series $\mathcal{E}(\mathbb{G}, (\alpha^*(t)))$ associated to $(\alpha^*(t))$. Consequently,

$$\chi_{(t)}{}^{\alpha} = \chi_{(\alpha^*(t))}, \tag{5.1}$$

which means that an automorphism of $\mathbb{G}$ maps the semisimple character associated to a conjugacy class (t) (of $\mathbb{G}^*$) to the semisimple character associated to $(\alpha^*(t))$. Here we note that if $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ is connected, then $\boldsymbol{C}_{\boldsymbol{G}^*}(\alpha^*(t))$ is also connected.

Due to Section 4 and Section 2A, we may assume that the Sylow p-subgroups of $\mathbb{G}$ are abelian. Assume for a moment that $\mathbb{G}$ is not of type ${}^2B_2$, ${}^2G_2$, or ${}^2F_4$. Then there is a unique positive integer e such that $p \mid \Phi_e(q)$ and Φ_e divides the generic order of $\mathbb{G}$. (Recall that Φ_e denotes the e-th cyclotomic polynomial.) This e then must be the multiplicative order of q modulo p, which means that $p \mid (q^e - 1)$ but $p \nmid (q^i - 1)$ for every $0 < i < e$. In the case where $\mathbb{G}$ is of type ${}^2B_2$, ${}^2G_2$, or ${}^2F_4$, what we just discussed still holds with slight modification on e, Φ_e, and $\boldsymbol{S}_e$; see [Malle 2007, Section 8] for more details.

Let $\Phi_e(q) = p^a m$ where $\gcd(p, m) = 1$ and $\Phi_e^{k_e}$ the precise power of Φ_e dividing the generic order of $\mathbb{G}$. We will use k for k_e for convenience if e is not specified. A Sylow e-torus of $\mathbb{G}^*$ has order $\Phi_e(q)^k$ and contains a Sylow p-subgroup of $\mathbb{G}^*$. Sylow p-subgroups of $\mathbb{G}^*$ (and $\mathbb{G}$) are then isomorphic to

$$\underbrace{C_{p^a} \times C_{p^a} \times \cdots \times C_{p^a}}_{k \text{ times}};$$

see [Malle and Testerman 2011, Theorem 25.14].

Assume that ℓ is the defining characteristic of S.

Lemma 5.3. *Assume the above notation. Let α be a field automorphism of $\mathbb{G}$. Each α-orbit on semisimple characters $\chi_{(t)} \in \mathrm{Irr}(\mathbb{G})$ associated to conjugacy classes of p-elements ($p \neq \ell$) of $\mathbb{G}^*$ has length at most $p^a - p^{a-1}$.*

Proof. Let α^* be an automorphism of $\mathbb{G}^*$ constructed from α by the process described above. For simplicity we use α for α^*. By (5.1), it is enough to show that each α-orbit on $\mathbb{G}^*$-conjugacy classes of (semisimple) p-elements of $\mathbb{G}^*$ has length at most $p^a - p^{a-1}$.

Let $t \in \mathbb{G}^*$ be a p-element. Note that each element in $\mathbb{G}^*$ conjugate to t under $\boldsymbol{G}^*$ is automatically conjugate to t under $\mathbb{G}^*$, by [Digne and Michel 1991, (3.25)] and the fact that $\boldsymbol{C}_{\boldsymbol{G}^*}(t)$ is connected. Let t be conjugate to $h_{\alpha_1}(\lambda_1) \cdots h_{\alpha_n}(\lambda_n)$, where the h_{α_i} are the coroots corresponding to a set of fundamental roots with respect to a maximal torus $\boldsymbol{T}^*$ of $\boldsymbol{G}^*$ and n is the rank of $\boldsymbol{G}^*$. Since $\boldsymbol{G}^*$ is simply connected, note that $(t_1, \dots, t_n) \mapsto h_{\alpha_1}(t_1) \cdots h_{\alpha_n}(t_n)$ is an isomorphism from $(\bar{\mathbb{F}}_q^\times)^n$ to $\boldsymbol{T}^*$; see [Gorenstein et al. 1994, Theorem 1.12.5].

Now, if $\lambda = \lambda_i$ for some $1 \leq i \leq n$, then $\lambda^{p^a} = 1$, since $|t| \mid p^a$. Recall that $\ell \neq p$, and thus $\ell^{p^a - p^{a-1}} \equiv 1 \pmod{p^a}$ by Euler's totient theorem. It follows that $\lambda^{\ell^{p^a - p^{a-1}}} = \lambda$, which yields that the α-orbit on (t) is contained in $\{(t), (\alpha(t)), \dots, (\alpha^{p^a - p^{a-1} - 1}(t))\}$, as desired. $\square$

5D. ***A bound for the number of*** $\mathbf{Aut}(S)$***-orbits on*** $\mathbf{Irr}(B_0(S))$**.** Let S_e be a Sylow e-torus of G^* and let $P \subseteq S_e$ be Sylow p-subgroup of $\mathbb{G}^*$. Such P exists by [Malle and Testerman 2011, Theorem 25.14]. Let $W(L_e)$ denote the relative Weyl group of the centralizer $L_e := C_{G^*}(S_e)$ of S_e in G^*. Here we note that L_e is a minimal e-split Levi subgroup of G^* and $W(L_e) \cong N_{\mathbb{G}^*}(S_e)/C_{\mathbb{G}^*}(S_e)$. By [Malle 2007, Proposition 5.11], $N_{\mathbb{G}^*}(S_e)$ controls $\mathbb{G}^*$-fusion in $C_{\mathbb{G}^*}(S_e)$, and since $P \subseteq S_e$, the number of conjugacy classes of (nontrivial) p-elements of $\mathbb{G}^*$ is at least

$$\frac{|P|-1}{|W(L_e)|} = \frac{p^{ak}-1}{|W(L_e)|}.$$

Note that $\chi_{(t)}$ belongs to the Lusztig series $\mathcal{E}(\mathbb{G}, (t))$ defined by the conjugacy class (t) and the Lusztig series are disjoint, and so two semisimple characters $\chi_{(t)}$ and $\chi_{(t_1)}$ are equal if and only if t and t_1 are conjugate in $\mathbb{G}^*$. Also, since $\mathbb{G}^*$ has abelian Sylow p-subgroup, p is a good prime for G, by [Malle 2014, Lemma 2.1]. Therefore, using Lemma 5.2, we deduce that

$$|\mathrm{Irr}_{ss}(B_0(\mathbb{G}))| \geq \frac{p^{ak}-1}{|W(L_e)|}, \tag{5.2}$$

where $\mathrm{Irr}_{ss}(B_0(\mathbb{G}))$ denotes the set of (nontrivial) semisimple characters (associated to p-elements of $\mathbb{G}^*$) in $B_0(\mathbb{G})$. Let $n(X, Y)$ denote the number of X-orbits on a set Y. Using Lemma 5.3, we then have

$$n(\mathrm{Aut}(S), \mathrm{Irr}_{ss}(B_0(\mathbb{G}))) \geq \frac{p^{ak}-1}{g(p^a-p^{a-1})|W(L_e)|} \geq \frac{p^k-1}{g(p-1)|W(L_e)|},$$

where g is the order of the group of graph automorphisms of S. Let $d := |\mathbb{G}/S|$ — the order of the group of diagonal automorphisms of S and viewing the irreducible constituents of the restrictions of semisimple characters of $\mathbb{G}$ to S as semisimple characters of S, we now have

$$n(\mathrm{Aut}(S), \mathrm{Irr}_{ss}(B_0(S))) \geq \frac{p^k-1}{dg(p-1)|W(L_e)|}. \tag{5.3}$$

We note that values of d, f, and g for various families of simple groups are known; see [Conway et al. 1985, page xvi] for instance.

We now turn to unipotent characters in the principal block $B_0(S)$. Broué, Malle and Michel [1993, Theorem 3.2] partitioned the set $\mathcal{E}(\mathbb{G}^*, 1)$ of unipotent characters of $\mathbb{G}^*$ into e-Harish-Chandra series associated to e-cuspidal pairs of G^*, and furthermore obtained one-to-one correspondences between e-Harish-Chandra series and the irreducible characters of the relative Weyl groups of the e-cuspidal pairs defining these series. Broué, Malle and Michel [1993, Theorem 5.24] then show that, when the Sylow p-subgroups of $\mathbb{G}^*$ is abelian, the partition of unipotent characters of $\mathbb{G}^*$ by e-Harish-Chandra series is compatible with the partition of unipotent characters by unipotent blocks; see [Cabanes and Enguehard 2004, Theorem 21.7] for a more general result. These results imply that the number of unipotent characters in $B_0(S)$ (and $B_0(\mathbb{G}^*)$ as well) is the same as the number $k(W(L_e))$ of conjugacy classes of the relative Weyl group $W(L_e)$ with $L_e := C_{G^*}(S_e)$, where S_e is a Sylow e-torus of G^*, as mentioned above.

By the aforementioned result of Lusztig (see [Malle 2008, Theorem 2.5] and also [Malle 2007, Theorem 3.9] for the corrected version), every unipotent character of a simple group of Lie type lies in an $\mathrm{Aut}(S)$-orbit of length at most 3. In fact, every unipotent character of S is $\mathrm{Aut}(S)$-invariant, except in the following cases:

(1) $S = P\Omega_{2n}^{+}(q)$ (n even), the graph automorphism of order 2 has $o_2(S)$ orbits of length 2, where $o_2(S)$ is the number of degenerate symbols of defect 0 and rank n parametrizing unipotent characters of S; see [Carter 1985, page 471].

(2) $S = P\Omega_8^{+}(q)$, the graph automorphism of order 3 has $o_3(S) = 2$ orbits of length 3, each of which contains one pair of characters parametrized by one degenerate symbol of defect 0 and rank 2 in (1).

(3) $S = \mathrm{Sp}_4(2^f)$, the graph automorphism of order 2 has $o_2(S) = 1$ orbit of length 2.

(4) $S = G_2(3^f)$, the graph automorphism of order 2 has $o_2(S) = 1$ orbit of length 2 on unipotent characters.

(5) $S = F_4(2^f)$, the graph automorphism of order 2 has $o_2(S) = 8$ orbits of length 2 on unipotent characters.

Combining this with the bound (5.3), we obtain:

Theorem 5.4. *Let S be a simple group of Lie type (including Suzuki and Ree groups). Let p be a prime different from the defining characteristic of S. Assume that Sylow p-subgroups of the group of inner and diagonal automorphisms of S are abelian. Let k, d, f, g, and $\boldsymbol{L}_e$ be as above and let $n(S)$ denote the number of* $\mathrm{Aut}(S)$*-orbits on irreducible ordinary characters in $B_0(S)$. Then*

$$n(S) \geq k(W(\boldsymbol{L}_e)) + \frac{p^k - 1}{dg(p-1)|W(\boldsymbol{L}_e)|},$$

except possibly the above cases (1), (3), (4), *and* (5) *in which the bound is lower by the number $o_2(S)$ of orbits of length* 2 *on unipotent characters and case* (2) *in which the bound is lower by* 4.

We remark that when the Sylow p-subgroups of the group of inner and diagonal automorphisms of S are furthermore noncyclic, then $k \geq 2$, and, away from those exceptions, we have a rougher bound

$$n(S) \geq k(W(\boldsymbol{L}_e)) + \frac{p+1}{dg|W(\boldsymbol{L}_e)|}, \tag{5.4}$$

but this turns out to be sufficient for our purpose in most cases.

6. Linear and unitary Groups

In this section, we let $S = \mathrm{PSL}_n^{\epsilon}(q)$, where $p \nmid q$ and $\epsilon \in \{\pm 1\}$. Here $\mathrm{PSL}_n^{\epsilon}(q) := \mathrm{PSL}_n(q)$ in the case $\epsilon = 1$ and $\mathrm{PSU}_n(q)$ in the case $\epsilon = -1$, and analogous for $\mathrm{SL}_n^{\epsilon}(q)$, $\mathrm{GL}_n^{\epsilon}(q)$, and $\mathrm{PGL}_n^{\epsilon}(q)$. We further let $\bar{q} := q$ if $\epsilon = 1$ and $\bar{q} := q^2$ if $\epsilon = -1$. Note that with our notation, $\mathrm{SL}_n^{\epsilon}(q)$ and $\mathrm{GL}_n^{\epsilon}(q)$ are naturally subgroups of $\mathrm{SL}_n(\bar{q})$ and $\mathrm{GL}_n(\bar{q})$, respectively.

We are now ready to prove Theorem 1.2 in the case of linear and unitary groups. Since $k(B_0(A))$ is bounded below by the number of A-orbits on $\mathrm{Irr}(B_0(S))$ for any $S \le A \le \mathrm{Aut}(S)$, our strategy in most cases will be to prove that there are more than $2\sqrt{p-1}$ orbits under $\mathrm{Aut}(S)$ in $\mathrm{Irr}(B_0(S))$, thus proving parts (i) and (ii) of Theorem 1.2 simultaneously.

Proposition 6.1. *Let $S = \mathrm{PSL}_n^\epsilon(q)$ and let $p \nmid q$ be a prime. Then Theorem 1.2 holds for any almost simple group A with socle S and $p \nmid |A/S|$.*

Proof. With the results of the previous sections, we may assume $n \ge 3$ and $p \ge 11$.

Write $S = \mathrm{PSL}_n^\epsilon(q)$, $\mathbb{G} = \mathrm{PGL}_n^\epsilon(q)$, $\widetilde{G} = \mathrm{GL}_n^\epsilon(q)$, and $G = \mathrm{SL}_n^\epsilon(q)$. Then we have $G = [\widetilde{G}, \widetilde{G}]$, $S = G/\boldsymbol{Z}(G)$, and $\mathbb{G} = \widetilde{G}/\boldsymbol{Z}(\widetilde{G})$. From Section 4, we may assume that Sylow p-subgroups of $\mathbb{G}$ are abelian, which implies that there is a unique e such that $p \mid \Phi_e(q)$ and Φ_e divides the generic order polynomial of $\mathbb{G}$. Here e must be $e_p(q)$, the multiplicative order of q modulo p. Note that this also forces $p \nmid n$ by again appealing to [Malle and Testerman 2011, Lemma 25.13].

We will further define $\bar{e} := e_p(\bar{q})$ and e' as follows:

$$e' := \begin{cases} \bar{e} & \text{if } \epsilon = 1 \text{ or if } \epsilon = -1 \text{ and } p \mid q^{\bar{e}} - (-1)^{\bar{e}}, \\ 2\bar{e} & \text{if } \epsilon = -1 \text{ and } p \mid q^{\bar{e}} + (-1)^{\bar{e}}. \end{cases}$$

To prove Theorem 1.2, our aim is to show that when a Sylow p-subgroup of S is not cyclic, then the number of $\mathrm{Aut}(S)$-orbits on $\mathrm{Irr}(B_0(S))$ is larger than $2\sqrt{p-1}$.

Note that since $p \nmid \gcd(n, q-\epsilon) = |\boldsymbol{Z}(G)|$, the irreducible characters in the principal block of S are the same as those of G, under inflation; see [Navarro 1998, Theorem 9.9]. Similarly, if $e' > 1$, then $p \nmid (q-\epsilon) = |\boldsymbol{Z}(\widetilde{G})|$ and an analogous statement holds for $\mathbb{G}$ and $\widetilde{G}$. Hence, we begin by studying $B_0(\widetilde{G})$, which will be sufficient for our purposes in the case $e' > 1$.

Let $n = we' + m$ with $0 \le m < e'$. Set $p^a := (\bar{q}^{\bar{e}} - 1)_p \ge p$. The case $p \le w$ was treated in Section 4, so we assume that $p > w$. Note that by [Michler and Olsson 1983, Theorem 1.9], $B_0(\widetilde{G})$ and $B_0(\mathrm{GL}_{we'}^\epsilon(q))$ have the same number of ordinary irreducible characters, so we may assume that $n = we'$. (Note that the action of $\mathrm{Aut}(S)$ is analogous as well.)

Let $\mathcal{F}(p, a)$ denote the set of monic polynomials over $\mathbb{F}_{\bar{q}}$ in the set $\mathscr{F}$ defined in [Fong and Srinivasan 1982] whose roots have p-power order in $\bar{\mathbb{F}}_q^\times$ at most p^a. Note that $|\mathcal{F}(p, a)| = 1 + (p^a - 1)/e'$; see [Michler and Olsson 1983, page 211].

The conjugacy classes $(t) := t^{\widetilde{G}}$ of p-elements in $\widetilde{G}$ are parametrized by p-weight vectors of w, which are functions $\boldsymbol{w} := \boldsymbol{w}_{(t)} : \mathcal{F}(p, a) \to \mathbb{Z}_{\ge 0}$ such that $w = \sum_{g \in \mathcal{F}(p,a)} \boldsymbol{w}(g)$. The characteristic polynomial of elements in (t) is

$$(x-1)^{e'\boldsymbol{w}(x-1)} \prod_{x-1 \ne g \in \mathcal{F}(p,a)} g^{\boldsymbol{w}(g)},$$

and the centralizer of t is

$$\boldsymbol{C}_{\widetilde{G}}(t) = \mathrm{GL}^\epsilon_{e'\boldsymbol{w}(x-1)}(q) \times \prod_{x-1 \ne g \in \mathcal{F}(p,a)} \mathrm{GL}^\eta_{\boldsymbol{w}(g)}(q^{e'})$$

where $\eta = \epsilon$ unless $\epsilon = -1$ and $e' = 2\bar{e}$, in which case $\eta = 1$.

Each character in the Lusztig series $\mathcal{E}(\widetilde{G}, t)$ is labeled by $\chi_{t,\psi}$ where ψ is a unipotent character of $\boldsymbol{C}_{\widetilde{G}}(t)$. So $\psi = \prod_{g\in\mathcal{F}(p,a)} \psi_g$ where ψ_g is a unipotent character of $\mathrm{GL}^{\eta}_{\boldsymbol{w}(g)}(q^{e'})$ if $g \neq x-1$ and of $\mathrm{GL}_{e'\boldsymbol{w}(x-1)}(q)$ if $g = x-1$. Note that there is a canonical correspondence between unipotent characters of $\mathrm{GL}^{\pm}_x(q)$ and partitions of x, so we may view ψ_g as a partition of $\boldsymbol{w}(g)$ when $g \neq x-1$ and of $e'\boldsymbol{w}(x-1)$ when $g = x-1$. Further, by [Fong and Srinivasan 1982, Theorem (7A)], the characters of $B_0(\widetilde{G})$ are exactly those $\chi_{t,\psi}$ satisfying t is a p-element and the partition ψ_{x-1} has trivial e'-core.

By [Olsson 1984, Proposition 6], since $w < p$, we have

$$k(B_0(\widetilde{G})) = k\left(e' + \frac{p^a-1}{e'}, w\right),$$

where $k(x, y)$ is as defined in Section 3 above. This number is at least

$$e' + \frac{p^a-1}{e'} \geq 2\sqrt{p^a-1} \geq 2\sqrt{p-1}. \tag{6.1}$$

But, recall that we wish to show that there are at least $2\sqrt{p-1}$ orbits on $\mathrm{Irr}(B_0(S))$ under $\mathrm{Aut}(S)$.

Now, by taking $t = 1$, the number of unipotent characters in $B_0(\widetilde{G})$ is precisely $k(e', w)$. Note that $k(e', w) \geq k(e', 1) = e'$, and that further $k(e', w) \geq 2e'$ if $w \geq 2$ with strict inequality for $(e', w) \neq (1, 2)$, and each unipotent character is $\mathrm{Aut}(S)$-invariant. So we have at least e' $\mathrm{Aut}(S)$-orbits of unipotent characters in $B_0(\mathbb{G})$, and hence of $B_0(S)$, since restriction yields a bijection between unipotent characters of S and $\mathbb{G}$.

Let $\widetilde{\boldsymbol{G}} := \mathrm{GL}_n(\overline{\mathbb{F}}_q)$ so that $\widetilde{G} = \widetilde{\boldsymbol{G}}^F$. Since $\boldsymbol{Z}(\widetilde{\boldsymbol{G}})$ is connected, [Cabanes and Späth 2013, Theorem 3.1] yields that the "Jordan decomposition" $\psi_{t,\psi} \leftrightarrow (t, \psi)$ can be chosen to be $\mathrm{Aut}(S)$-equivariant. Since ψ is a unipotent character of a product of groups of the form $\mathrm{GL}^{\pm}_x(q^d)$, which are invariant under automorphisms as discussed above, it follows that the orbit of $\chi_{t,\psi}$ is completely determined by the action of $\mathrm{Aut}(S)$ on the class (t).

Now, recall that the $\widetilde{G}$-class of t is completely determined by its eigenvalues. Let $|t| = p^c$ and note that $c \leq a$. By viewing t as an element $1 \times \prod_{x-1\neq g\in\mathcal{F}(p,a)} \zeta_g$ of

$$\boldsymbol{Z}(\boldsymbol{C}_{\widetilde{G}}(t)) \cong C_{q-\epsilon} \times \prod_{x-1\neq g\in\mathcal{F}(p,a)} C_{q^{e'}-\eta}$$

we see that for $\alpha \in \mathrm{Aut}(S)$, the eigenvalues of t^α are those of t raised to some power ηq_0^e for some $\eta \in \{\pm 1\}$ and some q_0 such that $\bar{q}$ is a power of q_0. This implies that the $\mathrm{Aut}(S)$-orbit of (t) has size at most $(p^c-1)/e' \leq (p^a-1)/e'$.

Now, a Sylow p-subgroup P of $\widetilde{G}$ is of the form $C^w_{p^a} \leq (\mathbb{F}^{\times}_{\bar{q}^{\bar{e}}})^w$. Then if $w = 1$, P is cyclic, and hence we may assume that $w \geq 2$. In this case, we have at least $\frac{1}{2}((p^a-1)/e')^2$ choices for $(t) \neq (1)$, and hence at least $\frac{1}{2}((p^a-1)/e')^2$ nonunipotent characters in $B_0(\widetilde{G})$ by taking ψ_{x-1} to be trivial. This gives at least $(p^a-1)/2e'$ distinct orbits of nonunipotent characters, and hence more than $2\sqrt{p-1}$ orbits of characters in $B_0(\mathbb{G})$ under $\mathrm{Aut}(S)$ when $e' > 1$, by (6.1) with $2e'$ rather than e'. This completes the proof of Theorem 1.2 for S in the case $e' > 1$ by the discussion at the beginning of the section.

Finally suppose $e' = 1$, so $w = n \geq 3$ and we may continue to assume $p > w$. Consider the elements t of $\widetilde{G}$ whose eigenvalues are of the form $\{\zeta, \xi, (\zeta\xi)^{-1}, 1, \dots, 1\}$ with ζ and ξ p-elements of $C_{q-\epsilon} \leq \mathbb{F}_{\bar{q}}^{\times}$. Note that each member of $\mathcal{E}(\widetilde{G}, t)$ lies in the principal block of $\widetilde{G}$, using [Cabanes and Enguehard 2004, Theorem 9.12] and that every unipotent character lies in $B_0(\widetilde{G})$ since $e' = 1$. Further, t lies in $G = [\widetilde{G}, \widetilde{G}]$ and $|\boldsymbol{C}_{\boldsymbol{G}^*}(t)/\boldsymbol{C}^{\circ}_{\boldsymbol{G}^*}(t)| = 1$ since this number must divide both the order of t and $|\boldsymbol{Z}(G)|$, contradicting $p > n$. Hence each character in such a $\mathcal{E}(\widetilde{G}, t)$ is irreducible on restriction to G, yielding at least $(p^a - 1)^2/2$ nonunipotent members of $B_0(G)$. Since the $\mathrm{Aut}(S)$-orbits of such characters are again of size at most $p^a - 1$, this yields at least $2 + (p^a - 1)/2$ distinct orbits, which is larger than $2\sqrt{p-1}$. This completes the proof of Theorem 1.2 in the case that $S = \mathrm{PSL}_n^{\epsilon}(q)$. □

7. Symplectic and orthogonal Groups

In this section, we consider the simple groups coming from orthogonal and symplectic groups. That is, simple groups of Lie type B_n, C_n, D_n, and ${}^2D_n$. We let $\epsilon \in \{\pm\}$, and let $\mathrm{P}\Omega^{\epsilon}_{2n}(q)$ denote the simple group of Lie type $D_n(q)$ for $\epsilon = +$ and of type ${}^2D_n(q)$ for $\epsilon = -$.

Proposition 7.1. *Let q be a power of a prime different from p and let $S = \mathrm{PSp}_{2n}(q)$ with $n \geq 2$, $\mathrm{P}\Omega_{2n+1}(q)$ with $n \geq 3$, or $\mathrm{P}\Omega^{\epsilon}_{2n}(q)$ with $n \geq 4$. Then Theorem 1.2 holds for any almost simple group A with socle S and $p \nmid |A/S|$.*

Proof. With the results of the previous sections, we may again assume that $p \geq 11$ and that a Sylow p-subgroup of S is abelian, but not cyclic.

Let H be the corresponding symplectic or special orthogonal group $\mathrm{Sp}_{2n}(q)$, $\mathrm{SO}_{2n+1}(q)$, or $\mathrm{SO}^{\epsilon}_{2n}(q)$ and let $(\boldsymbol{H}, F)$ be the corresponding simple algebraic group and Frobenius endomorphism so that $H = \boldsymbol{H}^F$. Let $G = \boldsymbol{G}^F$ be the corresponding group of simply connected type, so that $G = H$ in the symplectic case or G is the appropriate spin group in the orthogonal cases. Further, let $(\boldsymbol{H}^*, F)$ and $(\boldsymbol{G}^*, F)$ be dual to $(\boldsymbol{H}, F)$ and $(\boldsymbol{G}, F)$, respectively, and $H^* = \boldsymbol{H}^{*F}$ and $G^* = \boldsymbol{G}^{*F}$.

Define $\overline{H}$ to be the group $\mathrm{GO}^{\epsilon}_{2n}(q)$ in the case $S = \mathrm{P}\Omega^{\epsilon}_{2n}(q)$, and $\overline{H} := H$ otherwise. We also let Ω be the unique subgroup of index 2 in H for the orthogonal cases when q is odd, and let $\Omega = H$ otherwise, so that $\Omega/\boldsymbol{Z}(\Omega) = S = G/\boldsymbol{Z}(G)$ and $\Omega \lhd \overline{H}$. Note that since $p \neq 2$, $B_0(S)$ can be identified with $B_0(\Omega)$ or with $B_0(G)$, by [Navarro 1998, Theorem 9.9].

Now, let $e := e_p(q)/\gcd(e_p(q), 2)$ and write $n = we + m$ with $0 \leq m < e$. From Section 4, we may again assume $w < p$. To obtain our result, we will rely on the case of linear groups and use some of the ideas of the arguments used in [Malle 2018, Propositions 5.4 and 5.5], which provides an analogue in this situation to the results of Michler and Olsson discussed above. Namely, [Malle 2018, Propositions 5.4 and 5.5] tells us

$$k(B_0(\overline{H})) = k\left(2e + \frac{p^a - 1}{2e}, w\right),$$

where $p^a = (q^{2e} - 1)_p$. Note that again, this number is at least $2\sqrt{p-1}$ (with strict inequality when $w \geq 2$), but that we wish to show the inequality for $k(B_0(A))$. In most cases, we will again show that the number of $\mathrm{Aut}(S)$-orbits of characters in $B_0(S)$ is at least $2\sqrt{p-1}$.

If $w = 1$, a Sylow p-subgroup of Ω, G, H, or $\overline{H}$ (recall $p \geq 11$) is cyclic, so we may assume by Lemma 2.1 that $w \geq 2$. Note that the unipotent characters of H are irreducible on restriction to Ω. Assume first that $S \neq D_4(q)$ nor $\mathrm{Sp}_4(2^f)$. By [Malle 2018, Discussion before Propositions 5.4 and 5.5], the number of unipotent characters in $B_0(\overline{H})$ is $k(2e, w)$. If further $S \neq \mathrm{P}\Omega^{\epsilon}_{2n}(q)$, then since all unipotent characters are $\mathrm{Aut}(S)$-invariant, this yields $k(2e, w)$ $\mathrm{Aut}(S)$-orbits of unipotent characters in $B_0(H)$, and hence $B_0(S)$. Note that $k(2e, w) > 4e$ since $w \geq 2$. If $S = \mathrm{P}\Omega^{\epsilon}_{2n}(q)$, then note that $n \geq 4$ forces $e \geq 2$ if $w = 2$. Now, in this case, the proof of [Malle 2018, Lemma 5.6 and Corollary 5.7] yields that the number of $\overline{H}$-orbits of unipotent characters in $B_0(H)$ is at least $k(2e, w)/2$, and that this number is $(k(2e, w) + k(e, w/2))/2$ if w is even. Now, if $w \geq 3$, we have $k(2e, w)/2 > 4e$. If $w = 2$ and $e \geq 2$, we have

$$\frac{k(2e, 2) + k(e, 1)}{2} = e^2 + 2e \geq 4e.$$

Hence in all cases, the number of $\mathrm{Aut}(S)$-orbits of unipotent characters in $B_0(S)$ is at least $4e$, and is strictly greater unless $e = 2 = w$ in the case $S = \mathrm{P}\Omega^{\epsilon}_{2n}(q)$.

The characters in $B_0(H)$ and $B_0(G)$ lie in Lusztig series indexed by p-elements t of H^*, respectively G^*; by [Cabanes and Enguehard 2004, Theorem 9.12]. Note that centralizers of odd-order elements of $\boldsymbol{H}^*$ and of $\boldsymbol{G}^*$ are always connected (see, e.g., [Malle and Testerman 2011, Exercise (20.16)]) and that every odd p is good for $\boldsymbol{H}$ and $\boldsymbol{G}$, so that $\chi_{(t)}$ lies in $B_0(H)$, respectively $B_0(G)$, for every p-element t of H^*, respectively G^*, by Theorem 5.1. Further, note that the action on $\chi_{(t)}$ under a graph-field automorphism of H is determined by the action of a corresponding graph-field automorphism on (t), by [Navarro et al. 2008, Corollary 2.8]; see also (5.1) above.

Now let $\boldsymbol{G} \hookrightarrow \widetilde{\boldsymbol{G}}$ be a regular embedding as in [Cabanes and Enguehard 2004, 15.1] and let $\widetilde{G} := \widetilde{\boldsymbol{G}}^F$. Then the action of $\widetilde{G}$ on G induces all diagonal automorphisms of S. Now, since $C_{\boldsymbol{G}^*}(t)$ is connected for any p-element $t \in G^*$, we have every character in $\mathcal{E}(G, (t))$ extends to a character in $\widetilde{G}$. (Indeed, since $\widetilde{G}/G$ is abelian and restrictions from $\widetilde{G}$ to G are multiplicity-free, the number of characters lying below a given $\widetilde{\chi} \in \mathrm{Irr}(\widetilde{G})$ is the number of $\beta \in \mathrm{Irr}(\widetilde{G}/G)$ such that $\widetilde{\chi}\beta = \widetilde{\chi}$, as noted in [Rizo et al. 2021, Lemma 1.4]. Hence [Bonnafé 2005, Corollary 2.8] and [Schaeffer Fry and Taylor 2023, Proposition 2.6] yields the claim.) Therefore, each member of $B_0(S)$ is invariant under diagonal automorphisms.

First consider the case $H = \mathrm{SO}_{2n+1}(q)$ or $\mathrm{Sp}_{2n}(q)$, so $H^* = \mathrm{Sp}_{2n}(q)$ or $\mathrm{SO}_{2n+1}(q)$, respectively. Note that $\mathrm{Aut}(S)/S$ in this case is generated by field automorphisms, which also act on H, along with a diagonal or graph automorphism of order at most 2.

If $H = \mathrm{SO}_{2n+1}(q)$, then $\mathrm{GL}_n(q)$ may be embedded into $H^* = \mathrm{Sp}_{2n}(q)$ in a natural way (namely, block diagonally as the set of matrices of the form (A, A^{-T}) for $A \in \mathrm{GL}_n(q)$), and the conjugacy class of t is again determined by its eigenvalues. Arguing as in the case of $\mathrm{SL}_n(q)$ above and noting that every eigenvalue of t must have the same multiplicity as its inverse, we then have at least $(p^a - 1)/4e$ distinct orbits of nonunipotent characters in $B_0(H)$ under the field automorphisms, and hence at least $(p^a - 1)/4e$ orbits in $B_0(S)$ under $\mathrm{Aut}(S)$. This gives more than $4e + (p^a - 1)/4e$ orbits in $\mathrm{Irr}(B_0(S))$ under $\mathrm{Aut}(S)$, which proves Theorem 1.2 in this case using (6.1).

If $H = \mathrm{Sp}_{2n}(q)$, by [Geck and Hiss 1991, Theorem 4.2], there is a bijection between classes of p-elements of H and H^*, and we note that field automorphisms act analogously on the p-elements of H and H^*. Then the above again yields the result in this case as long as $S \neq \mathrm{Sp}_4(2^f)$.

If $S = \mathrm{Sp}_4(2^f)$, then we must have $e = 1$ and $w = 2$. Here [Malle 2008, Theorem 2.5] tells us that there is a pair of unipotent characters permuted by the exceptional graph automorphism, leaving $k(2, 2) - 1 = 4$ orbits of unipotent characters in $B_0(S)$ under $\mathrm{Aut}(S)$. In this case, arguing as before and considering the action of the graph automorphism gives at least $4 + (p^a - 1)/8$ orbits in $B_0(S)$ under $\mathrm{Aut}(S)$, which is at least $2(p-1)^{1/4}$. Hence part (ii) of Theorem 1.2 holds. So let $S \leq A \leq \mathrm{Aut}(S)$, and we wish to show that $B_0(A)$ contains more than $2\sqrt{p-1}$ characters. Note that in this case, $\mathrm{Aut}(S)/S$ is cyclic. Let $X := SC_A(P)$ for $P \in \mathrm{Syl}_p(S)$. Then A/X is cyclic, say of size b, and $B_0(A)$ is the unique block covering $B_0(X)$ by [Navarro 1998, (9.19) and (9.20)]. Note that since at least 3 of the unipotent characters of S are A-invariant, we have at least $3b$ characters in $B_0(A)$ lying above unipotent characters. Further, since the automorphisms corresponding to those in X stabilize p-classes in G^*, the arguments above give at least $\frac{1}{2} \cdot ((p^a - 1)/2)^2$ members of $B_0(X)$ lying above semisimple characters of S, and hence there are at least $(p^a - 1)^2/8b$ members of $B_0(A)$ lying above semisimple characters of S. Note then that the size of $B_0(A)$ is at least $3b + (p^a - 1)^2/8b$, which is larger than $2\sqrt{p-1}$, completing the proof in this case.

Now, suppose we are in the case that $\overline{H} = \mathrm{GO}^{\epsilon}_{2n}(q)$. Note that the action of $\overline{H}/H$ induces a graph automorphism of order 2 in the case $\epsilon = 1$, and that $\mathrm{Aut}(S)/S$ is generated by a group of diagonal automorphisms of size at most 4, along with graph and field automorphisms. Further, note that the action of H on Ω induces a diagonal automorphism of order 2 on S. We may embed $\overline{H}$ in $\mathrm{SO}_{2n+1}(q)$, and by [Malle 2018, proof of Proposition 5.5], the classes of p-elements t with Lusztig series contributing to $B_0(\overline{H})$ are parametrized exactly as in the case of $\mathrm{SO}_{2n+1}(q)$ above.

Assume that $(n, \epsilon) \neq (4, 1)$. By again considering semisimple characters $\chi_{(t)}$ of H for p-elements $t \in H^*$, we may conclude that the number of orbits of nonunipotent characters in $B_0(S)$ under $\mathrm{Aut}(S)$ is at least $(p^a - 1)/(4e)$. This yields at least $4e + (p^a - 1)/(4e)$ orbits in $\mathrm{Irr}(B_0(S))$ under $\mathrm{Aut}(S)$, with strict inequality unless $e = 2$. Hence we have the number of $\mathrm{Aut}(S)$ orbits in $B_0(S)$ is strictly larger than $4e + (p^a - 1)/(4e)$, completing Theorem 1.2 again in this case using (6.1), unless possibly if $e = 2$. But in the latter situation, we have $8 + (p^a - 1)/8 > 2\sqrt{p^a - 1}$ unless $8 = p^a - 1$, contradicting $p \geq 11$ and we are again done.

Finally, suppose $S = D_4(q) = \mathrm{P}\Omega^+_8(q)$ so $\overline{H} = \mathrm{GO}^+_8(q)$. In this case, the graph automorphisms generate a group of size 6, and a triality graph automorphism of order 3 permutes two triples of unipotent characters; see [Malle 2008, Theorem 2.5]. Since $w \geq 2$, we have $(e, w) \in \{(1, 4), (2, 2)\}$. The arguments above give at least $\big(\big(k(2e, w) + k(e, w/2)\big)/2\big) - 4 + (p^a - 1)/12e$ distinct $\mathrm{Aut}(S)$-orbits in $\mathrm{Irr}(B_0(S))$. Since $k(2, 4) = 20$, $k(1, 2) = 2 = k(2, 1)$, and $k(4, 2) = 14$, we have

$$\frac{k(2e, w) + k\big(e, \frac{w}{2}\big)}{2} - 4 + \frac{p^a - 1}{12e} > 2(p-1)^{1/4},$$

so Theorem 1.2(ii) is proved in this case.

Now, let $S \le A \le \mathrm{Aut}(S)$, let Γ be the subgroup of $\mathrm{Aut}(S)$ generated by inner, diagonal, and graph automorphisms, and let $X := (\Gamma \cap A)C_A(P)$. Then A/X is cyclic, and by [Navarro 1998, (9.19) and (9.20)], $B_0(A)$ is the unique block covering $B_0(X)$. Let $b := |A/X|$. Now, the arguments above give at least $\frac{1}{3} \cdot \frac{1}{2} \cdot ((p-1)/4)^2$ members of $B_0(X)$ lying above semisimple characters of S, since members of $C_A(P)$ correspond to automorphisms stabilizing classes of p-elements of G^*, and hence there are at least $(p-1)^2/(96b)$ members of $B_0(A)$ lying above semisimple characters of S. Further, there are at least 10 characters in $B_0(X)$ lying above unipotent characters in $B_0(S)$. Since unipotent characters extend to their inertia groups and are invariant under field automorphisms (see [Malle 2008, Theorems 2.4 and 2.5]), this gives at least $4b$ elements of $B_0(A)$ lying above unipotent characters of S. Together, this gives $k(B_0(A)) \ge 4b + (p-1)^2/(96b) > 2\sqrt{p-1}$ since $p \ge 11$, proving part (i) of Theorem 1.2. □

8. Groups of exceptional types

In this section we prove Theorem 1.2 for S being of exceptional type. This is achieved by considering each type case by case, with the help of Theorem 5.4.

We keep all the notation in Section 5. In particular, the underlying field of S has order $q = \ell^f$. By Section 2A, we may assume that $\ell \ne p \ge 11$. This assumption on p guarantees that Sylow p-subgroups of $\mathbb{G}$ are abelian. Recall also that e is the multiplicative order of q modulo p (when S is not of Suzuki or Ree type), $p^a = \Phi_e(q)_p$, and $\Phi_e^k = \Phi_e^{k_e}$ is the precise power of Φ_e dividing the generic order of $\mathbb{G}$. By Section 2A, we may assume that the Sylow p-subgroups of S are not cyclic, and thus $k_e \ge 2$. Also, $\boldsymbol{S}_e$ is a Sylow e-torus of a simple algebraic group $\boldsymbol{G}^*$ of simply connected type associated with a Steinberg endomorphism F such that $S = \mathbb{G}^*/\boldsymbol{Z}(\mathbb{G}^*)$ and $\mathbb{G}^* := \boldsymbol{G}^{*F}$, and $\boldsymbol{L}_e := \boldsymbol{C}_{\mathbb{G}^*}(\boldsymbol{S}_e)$ is a minimal e-split Levi subgroup of $\boldsymbol{G}^*$. Note that $\boldsymbol{L}_e$ is then a maximal torus of $\boldsymbol{G}^*$ (in other words, e is regular for $\boldsymbol{G}^*$), except the single case of type E_7 and $e = 4$. The relative Weyl groups $W(\boldsymbol{L}_e)$ are always finite complex reflection groups, and we will follow the notation for these groups in [Benard 1976]. Relative Weyl groups for various $\boldsymbol{L}_e$ are available in [Broué et al. 1993, Tables 1 and 3]. The structure of $\mathrm{Out}(S)$ is available in [Gorenstein et al. 1994, Theorem 2.5.12]. We will use these data freely without further notice.

It turns out that Theorem 5.4 is sufficient to prove Theorem 1.2 whenever $k_e \ge 3$. In fact, even when $k_e = 2$, Theorem 5.4 is also sufficient for Theorem 1.2(ii). We have to work harder, though, to achieve Theorem 1.2(i) in the case $k_e = 2$ for some types.

Proposition 8.1. *Theorem 1.2 holds for simple groups of exceptional types.*

Proof. (1) $\underline{S = G_2(q) \text{ and } S = F_4(q)}$: First we consider $S = G_2(q)$ (so $S = \mathbb{G}$) with $q > 2$. Then $e \in \{1, 2\}$ and $k_1 = k_2 = 2$. Also, the Sylow e-tori are maximal tori, and their relative Weyl groups are the dihedral group D_{12}. The bound (5.4) implies that $n(S) \ge 5 + (p+1)/12$ for $q = 3^f$ with odd f, and $n(S) \ge 6 + (p+1)/12$ otherwise. In any case it follows that $n(S) > 2(p-1)^{1/4}$, proving Theorem 1.2(ii) for $G_2(q)$.

Note that $\mathrm{Aut}(S)$ is a cyclic extension of S. First assume that $q \ne 3^f$ or G does not contain the graph automorphism of S. In particular, every unipotent character of S is extendible to G. Let $H := \langle S, \boldsymbol{C}_G(P)\rangle$, where P is a Sylow p-subgroup of G (and S as well by the assumption $p \nmid |G/S|$). Since $P\boldsymbol{C}_G(P)$

is contained in H, $B_0(H)$ is covered by a unique block of G, which is $B_0(G)$. It follows that, each unipotent character in $B_0(S)$ extends to an irreducible character in $B_0(H)$, which in turn lies under $|G/H|$ irreducible characters in $B_0(G)$. Therefore, the number of irreducible characters in $B_0(G)$ lying over unipotent characters of S is at least $k(D_{12})|G/H| = 6|G/H|$. When $q = 3^f$ and G does contain the nontrivial graph automorphism, similar arguments yield that the number of irreducible characters in $B_0(G)$ lying over unipotent characters of S is at least $5|G/H|$.

On the other hand, each G-orbit on semisimple characters (associated to p-elements) of S now has length at most $|G/H|$ by (5.1) and the fact that $H = \langle S, \boldsymbol{C}_G(P)\rangle$ fixes every conjugacy class of p-elements of S. Therefore, the bound (5.2) yields

$$n(G, \mathrm{Irr}_{ss}(B_0(S))) \geq \frac{p^2-1}{12|G/H|}.$$

This and the conclusion of the last paragraph imply that

$$k(B_0(G)) \geq 5|G/H| + \frac{p^2-1}{12|G/H|} \geq 2\sqrt{\frac{5(p^2-1)}{12}},$$

which in turn implies the desired bound $k(B_0(G)) > 2\sqrt{p-1}$ for all $p \geq 11$.

For $S = F_4(q)$, we have $e \in \{1, 2\}$ for which $k_e = 4$, or $e \in \{3, 4, 6\}$ for which $k_e = 2$. Therefore all the Sylow e-tori are maximal tori, and their relative Weyl groups are $G_{28} = \mathrm{GO}_4^+(3)$ for $e = 1, 2$; $G_5 = \mathrm{SL}_2(3) \times C_3$ for $e = 3, 6$; and $G_8 = C_4.\mathrm{S}_4$ for $e = 4$. Now we just follow along similar arguments as above to prove the theorem for this type.

(2) $S = {}^2F_4(q)$ with $q = 2^{2n+1} \geq 8$ and $S = {}^3D_4(q)$: These two types are treated in a fairly similar way as for G_2. Note that $\mathrm{Out}(S)$ here is always cyclic. First let $S = {}^2F_4(q)$. Then $e \in \{1, 2, 4^+, 4^-\}$ and $k_e = 2$ for all e. All the Sylow e-tori are maximal. The relative Weyl groups of these tori are D_{16}, $G_{12} = \mathrm{GL}_2(3)$, $G_8 = C_4.\mathrm{S}_4$ and G_8 for $e = 1, 2, 4^+$, and 4^-, respectively. One can now easily check the inequality $n(S) \geq 2(p-1)^{1/4}$, using (5.4). The bound $k(B_0(G)) > 2\sqrt{p-1}$ is proved similarly as in type G_2.

Now let $S = {}^3D_4(q)$. Then $e \in \{1, 2, 3, 6\}$ and $k_e = 2$ for all e. For $e \in \{3, 6\}$, a Sylow e-torus is maximal with the relative Weyl group $G_4 = \mathrm{SL}_2(3)$. For $e = 1$ or 2, Sylow e-tori of S are not maximal anymore but are contained in maximal tori of orders $\Phi_1^2(q)\Phi_3(q)$ and $\Phi_2^2(q)\Phi_6(q)$, respectively. The relative Weyl groups of these tori are both isomorphic to D_{12}. Now the routine estimates are applied to achieve the required bounds.

(3) $S = E_6(q)$ and $S = {}^2E_6(q)$: These two types are approached similarly and we will provide details only for E_6. Then $e = 1$ for which $k_e = 6$, or $e = 2$ for which $k_e = 4$, or $e = 3$ for which $k_e = 3$, or $e \in \{4, 6\}$ for which $k_e = 2$.

Assume $e = 1$. Then $\boldsymbol{S}_1$ is a maximal torus and its Weyl group is $G_{35} = \mathrm{SO}_5(3)$. Theorem 5.4 then implies that

$$n(S) \geq k(\mathrm{SO}_5(3)) + \frac{p^6-1}{6(p-1)|\mathrm{SO}_5(3)|} = 25 + \frac{p^6-1}{311040(p-1)} > 2\sqrt{p-1},$$

proving both parts of Theorem 1.2 in this case. The case $e \in \{2, 3\}$ is similar. We note that S_3 is a maximal torus with the relative Weyl group $G_{25} = 3^{1+2}.\mathrm{SL}_2(3)$, and a maximal torus containing a Sylow 2-torus has relative Weyl group G_{28}.

Assume $e = 4$. Then a maximal torus containing a Sylow 4-torus of $E_6(q)_{sc}$ has order $\Phi_4^2(q)\Phi_1^2(q)$ and its relative Weyl group is $G_8 = C_4.\mathrm{S}_4$, whose order is 96 and class number is 16. Now the bound (5.4) yields $n(S) > 2(p-1)^{1/4}$, proving part (ii) of the theorem.

We need to do more to obtain part (i) in this case. In fact, when $2\sqrt{p-1} \le 16$, which means that $p \le 65$, we have $n(S) > 16 \ge 2\sqrt{p-1}$, which proves part (i) as well. So let us assume that $p > 65$.

Note that $\mathrm{Out}(S)$ is a semidirect product $C_{(3,q-1)} \rtimes (C_f \times C_2)$, which may not be abelian but every unipotent character of S is still fully extendible to $\mathrm{Aut}(S)$ by [Malle 2008, Theorems 2.4 and 2.5]. As before, let $\mathbb{G}$ be the extension of S by diagonal automorphisms. Similar to the proof for type G_2, let $H := \langle G \cap \mathbb{G}, \boldsymbol{C}_G(P)\rangle$, where P is a Sylow p-subgroup of S. Each unipotent character in $B_0(S)$ then lies under at least $|\mathrm{Irr}(G/H)| = |G/H|$ irreducible characters in $B_0(G)$. (Here we note that G/H is abelian.) Thus, the number of irreducible characters in $B_0(G)$ lying over unipotent characters of S is at least $16|G/H|$.

As in Section 5D, here we have

$$|\mathrm{Irr}_{ss}(B_0(\mathbb{G}))| \ge \frac{p^2-1}{|W(\boldsymbol{L}_4)|} = \frac{p^2-1}{96}.$$

Let $\mathrm{Irr}_{ss}(B_0(S))$ be the set of restrictions of characters in $\mathrm{Irr}_{ss}(B_0(\mathbb{G}))$ to S. These restrictions are irreducible as the semisimple elements of $\mathbb{G}^*$ associated to these semisimple characters are p-elements whose orders are coprime to $|\boldsymbol{Z}(\mathbb{G}^*)| = \gcd(3, q-1)$. Moreover, if the restrictions of $\chi_{(t)}$ and $\chi_{(t_1)}$ to S are the same, then $(t) = (t_1 z)$ for some $z \in \boldsymbol{Z}(\mathbb{G}^*)$ (see [Tiep 2015, Proposition 5.1]), which happens only when z is trivial since t and t_1 are p-elements. It follows that

$$|\mathrm{Irr}_{ss}(B_0(S))| = |\mathrm{Irr}_{ss}(B_0(\mathbb{G}))| \ge \frac{p^2-1}{96}.$$

Note that $\mathbb{G} \cap G = \mathbb{G}$ or S and each G-orbit of relevant semisimple characters in $B_0(\mathbb{G} \cap G)$, and hence in $B_0(S)$, has length at most $|G/H|$. It follows that the number of irreducible characters in $B_0(G)$ lying over semisimple characters in $B_0(S)$ is at least $(p^2-1)/(96|G/H|)$. Together with the bound of $16|G/H|$ for the number of irreducible characters in $B_0(G)$ lying over unipotent characters of S, we deduce that

$$k(B_0(G)) \ge 16|G/H| + \frac{p^2-1}{96|G/H|} \ge 2\sqrt{\frac{16(p^2-1)}{96}},$$

and thus, when $p > 65$, the desired bound $k(B_0(G)) > 2\sqrt{p-1}$ follows.

The last case $e = 6$ can be argued in a similar way, with notice that a maximal torus containing a Sylow 6-torus of $E_6(q)_{sc}$ has order $\Phi_6^2(q)\Phi_3(q)$ and its relative Weyl group is $G_5 = \mathrm{SL}_2(3) \times C_3$, whose order is 72 and class number is 21.

(4) $\underline{S = E_7(q)}$: Then $e \in \{1, 2\}$ for which $k_e = 7$, or $e \in \{3, 6\}$ for which $k_e = 3$, or $e = 4$ for which $k_e = 2$. When $k_e > 2$, the bound (5.4) again is sufficient to achieve the desired bound $n(S) > 2\sqrt{p-1}$. In fact,

even for the case $k_e = 2$, we have $n(S) \geq 2(p-1)^{1/4}$. So it remains to prove Theorem 1.2(i) for $e = 4$, in which case e is not regular and the relative Weyl group of the minimal e-split Levi subgroup $\boldsymbol{L}_e = \boldsymbol{S}_e.A_1^3$ is G_8. The estimates are now similar to those in the case $e = 4$ of the type E_6.

(5) $\underline{S = E_8(q)}$: Then $e \in \{1, 2\}$ for which $k_e = 8$, or $e \in \{3, 4, 6\}$ for which $k_e = 4$, or $e \in \{5, 8, 10, 12\}$ for which $k_e = 2$. The standard approach as above works for all e with $k_e > 2$.

Assume that $e \in \{5, 10\}$. Then a Sylow e-torus of S is maximal and its relative Weyl group is $G_{16} \cong \mathrm{SL}_2(5) \times C_5$. A similar proof to the case of type G_2 yields $k(B_0(G)) \geq 2\sqrt{45(p^2-1)/600}$, which is certainly greater than $2\sqrt{p-1}$ for $p \geq 13$. On the other hand, we always have $k(B_0(G)) \geq 45 > 2\sqrt{p-1}$ for smaller p, and thus the desired bound holds for all p. Finally, the case $e \in \{8, 12\}$ is entirely similar, with notice that the relative Weyl groups of Sylow e-tori are $G_9 = C_8.\mathrm{S}_4$ and $G_{10} = C_{12}.\mathrm{S}_4$ for $e = 8$ and 12, respectively. □

Theorem 1.2 is now completely proved.

9. Proof of Theorems 1.1 and 1.3

We are now ready to prove the main results.

Proof of Theorems 1.1 and 1.3. First we remark that the "if" implication of Theorem 1.3 is clear, and moreover, we are done if the Sylow p-subgroups of G are cyclic, thanks to Section 2A.

Let (G, p) be a counterexample to either Theorem 1.1 or the "only if" implication of Theorem 1.3 with $|G|$ minimal. In particular, Sylow p-subgroups of G are not cyclic and $k(B_0(G)) \leq 2\sqrt{p-1}$. Let N be a minimal normal subgroup of G. Note that $N = G$ if G turns out to be simple.

Assume first that $p \mid |G/N|$. Then, since $\mathrm{Irr}(B_0(G/N)) \subseteq \mathrm{Irr}(B_0(G))$ and by the minimality of $|G|$, we have

$$2\sqrt{p-1} \geq k(B_0(G)) \geq k(B_0(G/N)) \geq 2\sqrt{p-1},$$

and thus

$$k(B_0(G)) = k(B_0(G/N)) = 2\sqrt{p-1}.$$

The minimality of G again then implies that G/N is isomorphic to the Frobenius group $C_p \rtimes C_{\sqrt{p-1}}$. It follows that $p \mid |N|$, and thus there exists a nontrivial irreducible character $\theta \in \mathrm{Irr}(B_0(N))$. As $B_0(G)$ covers $B_0(N)$, there is some $\chi \in \mathrm{Irr}(B_0(G))$ lying over θ, implying that $k(B_0(G)) > k(B_0(G/N))$, a contradiction.

So we must have $p \nmid |G/N|$, and it follows that $p \mid |N|$. This in fact also yields that N is the unique minimal normal subgroup of G. Assume first that N is abelian. We then have that G is p-solvable, and hence Fong's theorem (see [Navarro 1998, Theorem 10.20]) implies that

$$k(B_0(G)) = k(B_0(G/\boldsymbol{O}_{p'}(G))) = k(G/\boldsymbol{O}_{p'}(G)),$$

which is greater than $2\sqrt{p-1}$ by the main result of [Maróti 2016].

We now may assume that $N \cong S_1 \times S_2 \times \cdots \times S_k$, a direct product of $k \in \mathbb{N}$ copies of a nonabelian simple group S. If S has cyclic Sylow p-subgroups, then G is not a counterexample for Theorem 1.1 by Lemma 2.1. Furthermore,

$$k(B_0(G)) \geq k(\boldsymbol{N}_G(P)/\boldsymbol{O}_{p'}(\boldsymbol{N}_G(P))) > 2\sqrt{p-1}$$

by the analysis in Section 2A, and thus G is not a counterexample for Theorem 1.3 either.

So the Sylow p-subgroups of S are not cyclic. Let n be the number of $\boldsymbol{N}_G(S_1)/N$-orbits on $\mathrm{Irr}(B_0(S_1))$. By Theorem 1.2(ii), we have $n \geq 2(p-1)^{1/4}$. Therefore, if $k \geq 2$, the number of G-orbits on $\mathrm{Irr}(B_0(N)) = \prod_{i=1}^{k} \mathrm{Irr}(B_0(S_i))$ is at least $n(n+1)/2 \geq 2(p-1)^{1/4}(2(p-1)^{1/4}+1)/2 > 2\sqrt{p-1}$, and it follows that $k(B_0(G)) > 2\sqrt{p-1}$, a contradiction. Hence, $N = S$ and G is then an almost simple group with socle S. Furthermore, $p \nmid |G/S|$. But such a group G cannot be a counterexample by Theorem 1.2(i). The proof is complete. □

In regard to Theorem 1.1, we remark that Kovács and Leedham-Green constructed, for any odd prime p, a family of p-groups P of order p^p with $k(P) = (p^3 - p^2 + p + 1)/2$; see [Pyber 1992]. Therefore the bound $k(B_0(G)) \geq 2\sqrt{p-1}$ cannot be replaced by $k(B_0(G)) \geq p^3$, even if one assumes $|G|$ to be divisible by a certain fixed power of p.

With Theorem 1.1 in mind, it follows that for any p-block B for a finite group such that $k(B) = k(B_0(H))$ for some finite group H of order divisible by p, we have $k(B) \geq 2\sqrt{p-1}$. In particular, we may record the following:

Corollary 9.1. *Let G be one of the classical groups $\mathrm{GL}_n(q)$, $\mathrm{GU}_n(q)$, $\mathrm{Sp}_{2n}(q)$, $\mathrm{SO}_{2n+1}(q)$, or $\mathrm{GO}_{2n}^{\pm}(q)$. Let p be a prime dividing $|G|$ and not dividing q. Then for any p-block B of G with positive defect, we have $k(B) \geq 2\sqrt{p-1}$.*

Proof. If $p = 2$, then the statement is clear, so we assume p is odd. First, if $G = \mathrm{GL}_n(q)$ or $\mathrm{GU}_n(q)$, the statement follows immediately from Theorem 1.1 and [Michler and Olsson 1983, Theorem (1.9)], which states that B has the same number of irreducible characters as the principal block of a product of lower-rank general linear and unitary groups of order also divisible by p.

Now suppose that G is $\mathrm{Sp}_{2n}(q)$, $\mathrm{SO}_{2n+1}(q)$, or $\mathrm{GO}_{2n}^{\pm}(q)$. If B is a unipotent block, then by [Malle 2018, Proposition 5.4 and 5.5], B has the same number of irreducible characters as a block of an appropriate general linear group of order also divisible by p. (In the case $\mathrm{GO}_{2n}^{\pm}(q)$, we define a unipotent block to be one lying above a unipotent block of $\mathrm{SO}_{2n}^{\pm}(q)$.) Hence the statement holds if B is a unipotent block.

Now, the block B determines a class of semisimple p'-elements (s) of the dual group G^* (see [Cabanes and Enguehard 2004, Theorem 9.12]) such that B contains some member of $\mathcal{E}(G, (s))$. By [Enguehard 2008, Théorème 1.6], there exists a group $G(s)$ dual to $\boldsymbol{C}_{G^*}(s)$ such that $k(B) = k(b)$ for some unipotent block b of $G(s)$. Now, in the cases under consideration, $\boldsymbol{C}_{G^*}(s)$ and $G(s)$ are direct products of lower-rank classical groups of the types being considered here, completing the proof. □

References

[Belonogov 1990] V. A. Belonogov, "Finite groups with a small principal p-block", pp. 8–30 in *Group-theoretic investigations*, edited by A. I. Starostin, Akad. Nauk SSSR Ural. Otdel., Sverdlovsk, 1990. In Russian. MR Zbl

[Benard 1976] M. Benard, "Schur indices and splitting fields of the unitary reflection groups", *J. Algebra* **38**:2 (1976), 318–342. MR Zbl

[Bessenrodt et al. 2007] C. Bessenrodt, G. Navarro, J. B. Olsson, and P. H. Tiep, "On the Navarro–Willems conjecture for blocks of finite groups", *J. Pure Appl. Algebra* **208**:2 (2007), 481–484. MR Zbl

[Bonnafé 2005] C. Bonnafé, "Quasi-isolated elements in reductive groups", *Comm. Algebra* **33**:7 (2005), 2315–2337. MR

[Brandt 1982] J. r. Brandt, "A lower bound for the number of irreducible characters in a block", *J. Algebra* **74**:2 (1982), 509–515. MR

[Brauer 1942] R. Brauer, "On groups whose order contains a prime number to the first power, I", *Amer. J. Math.* **64** (1942), 401–420. MR Zbl

[Brauer 1963] R. Brauer, "Representations of finite groups", pp. 133–175 in *Lectures on Modern Mathematics*, vol. I, edited by T. L. Saaty, Wiley, New York, 1963. MR Zbl

[Brauer and Feit 1959] R. Brauer and W. Feit, "On the number of irreducible characters of finite groups in a given block", *Proc. Nat. Acad. Sci. U.S.A.* **45** (1959), 361–365. MR Zbl

[Broué et al. 1993] M. Broué, G. Malle, and J. Michel, "Generic blocks of finite reductive groups", pp. 7–92 in *Représentations unipotentes génériques et blocs des groupes réductifs finis*, Astérisque **212**, Société Mathématique de France, Paris, 1993. MR Zbl

[Cabanes 2018] M. Cabanes, "Local methods for blocks of finite simple groups", pp. 179–265 in *Local representation theory and simple groups*, edited by R. Kessar et al., Eur. Math. Soc., Zürich, 2018. MR Zbl

[Cabanes and Enguehard 1994] M. Cabanes and M. Enguehard, "On unipotent blocks and their ordinary characters", *Invent. Math.* **117**:1 (1994), 149–164. MR Zbl

[Cabanes and Enguehard 2004] M. Cabanes and M. Enguehard, *Representation theory of finite reductive groups*, New Mathematical Monographs **1**, Cambridge University Press, 2004. MR Zbl

[Cabanes and Späth 2013] M. Cabanes and B. Späth, "Equivariance and extendibility in finite reductive groups with connected center", *Math. Z.* **275**:3-4 (2013), 689–713. MR Zbl

[Carter 1985] R. W. Carter, *Finite groups of Lie type*: *conjugacy classes and complex characters*, Wiley, New York, 1985. 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*, Oxford University Press, 1985. MR

[Dade 1966] E. C. Dade, "Blocks with cyclic defect groups", *Ann. of Math.* (2) **84** (1966), 20–48. MR Zbl

[Digne and Michel 1991] F. Digne and J. Michel, *Representations of finite groups of Lie type*, London Mathematical Society Student Texts **21**, Cambridge University Press, 1991. MR Zbl

[Enguehard 2008] M. Enguehard, "Vers une décomposition de Jordan des blocs des groupes réductifs finis", *J. Algebra* **319**:3 (2008), 1035–1115. MR Zbl

[Fong and Srinivasan 1982] P. Fong and B. Srinivasan, "The blocks of finite general linear and unitary groups", *Invent. Math.* **69**:1 (1982), 109–153. MR Zbl

[Fulman and Guralnick 2012] J. Fulman and R. Guralnick, "Bounds on the number and sizes of conjugacy classes in finite Chevalley groups with applications to derangements", *Trans. Amer. Math. Soc.* **364**:6 (2012), 3023–3070. MR Zbl

[GAP 2020] The GAP Group, "GAP: groups, algorithms, and programming", 2020, available at http://www.gap-system.org. Version 4.11.0.

[Geck and Hiss 1991] M. Geck and G. Hiss, "Basic sets of Brauer characters of finite groups of Lie type", *J. Reine Angew. Math.* **418** (1991), 173–188. MR Zbl

[Gluck et al. 2004] D. Gluck, K. Magaard, U. Riese, and P. Schmid, "The solution of the $k(GV)$-problem", *J. Algebra* **279**:2 (2004), 694–719. MR

[Gorenstein et al. 1994] D. Gorenstein, R. Lyons, and R. Solomon, *The classification of the finite simple groups*, Mathematical Surveys and Monographs **40**, American Mathematical Society, Providence, RI, 1994. MR Zbl

[Héthelyi and Külshammer 2000] L. Héthelyi and B. Külshammer, "On the number of conjugacy classes of a finite solvable group", *Bull. London Math. Soc.* **32**:6 (2000), 668–672. MR Zbl

[Hethelyi and Külshammer 2003] L. Hethelyi and B. Külshammer, "On the number of conjugacy classes of a finite solvable group, II", *J. Algebra* **270**:2 (2003), 660–669. MR Zbl

[Héthelyi et al. 2011] L. Héthelyi, E. Horváth, T. M. Keller, and A. Maróti, "Groups with few conjugacy classes", *Proc. Edinb. Math. Soc.* (2) **54**:2 (2011), 423–430. MR Zbl

[Hiss 1990] G. Hiss, "Regular and semisimple blocks of finite reductive groups", *J. London Math. Soc.* (2) **41**:1 (1990), 63–68. MR Zbl

[Hung et al. 2023] N. N. Hung, A. A. Schaeffer Fry, and C. Vallejo, "Height zero characters in principal blocks", *J. Algebra* **622** (2023), 197–219. MR Zbl

[Jansen et al. 1995] C. Jansen, K. Lux, R. Parker, and R. Wilson, *An atlas of Brauer characters*, London Mathematical Society Monographs. New Series **11**, The Clarendon Press Oxford University Press, New York, 1995. MR Zbl

[Keller 2009] T. M. Keller, "Lower bounds for the number of conjugacy classes of finite groups", *Math. Proc. Cambridge Philos. Soc.* **147**:3 (2009), 567–577. MR Zbl

[Kessar and Malle 2013] R. Kessar and G. Malle, "Quasi-isolated blocks and Brauer's height zero conjecture", *Ann. of Math.* (2) **178**:1 (2013), 321–384. MR Zbl

[Koshitani and Sakurai 2021] S. Koshitani and T. Sakurai, "The principal p-blocks with four irreducible characters", *Bull. Lond. Math. Soc.* **53**:4 (2021), 1124–1138. MR Zbl

[Koshitani and Späth 2016] S. Koshitani and B. Späth, "The inductive Alperin–McKay and blockwise Alperin weight conditions for blocks with cyclic defect groups and odd primes", *J. Group Theory* **19**:5 (2016), 777–813. MR Zbl

[Kovács and Leedham-Green 1986] L. G. Kovács and C. R. Leedham-Green, "Some normally monomial p-groups of maximal class and large derived length", *Quart. J. Math. Oxford Ser.* (2) **37**:145 (1986), 49–54. MR Zbl

[Külshammer 1990] B. Külshammer, "Landau's theorem for p-blocks of p-solvable groups", *J. Reine Angew. Math.* **404** (1990), 189–191. MR Zbl

[Malle 2006] G. Malle, "Fast-einfache Gruppen mit langen Bahnen in absolut irreduzibler Operation", *J. Algebra* **300**:2 (2006), 655–672. MR Zbl

[Malle 2007] G. Malle, "Height 0 characters of finite groups of Lie type", *Represent. Theory* **11** (2007), 192–220. MR Zbl

[Malle 2008] G. Malle, "Extensions of unipotent characters and the inductive McKay condition", *J. Algebra* **320**:7 (2008), 2963–2980. MR Zbl

[Malle 2014] G. Malle, "On the inductive Alperin–McKay and Alperin weight conjecture for groups with abelian Sylow subgroups", *J. Algebra* **397** (2014), 190–208. MR Zbl

[Malle 2018] G. Malle, "On a minimal counterexample to Brauer's $k(B)$-conjecture", *Israel J. Math.* **228**:2 (2018), 527–556. MR Zbl

[Malle and Maróti 2016] G. Malle and A. Maróti, "On the number of p'-degree characters in a finite group", *Int. Math. Res. Not.* **2016**:20 (2016), 6118–6132. MR Zbl

[Malle and Testerman 2011] G. Malle and D. Testerman, *Linear algebraic groups and finite groups of Lie type*, Cambridge Studies in Advanced Mathematics **133**, Cambridge University Press, 2011. MR Zbl

[Maróti 2016] A. Maróti, "A lower bound for the number of conjugacy classes of a finite group", *Adv. Math.* **290** (2016), 1062–1078. MR Zbl

[Michler and Olsson 1983] G. O. Michler and J. r. B. Olsson, "Character correspondences in finite general linear, unitary and symmetric groups", *Math. Z.* **184**:2 (1983), 203–233. MR Zbl

[Nagao 1962] H. Nagao, "On a conjecture of Brauer for p-solvable groups", *J. Math. Osaka City Univ.* **13** (1962), 35–38. MR Zbl

[Navarro 1998] G. Navarro, *Characters and blocks of finite groups*, London Mathematical Society Lecture Note Series **250**, Cambridge University Press, 1998. MR Zbl

[Navarro et al. 2008] G. Navarro, P. H. Tiep, and A. Turull, "Brauer characters with cyclotomic field of values", *J. Pure Appl. Algebra* **212**:3 (2008), 628–635. MR

[Navarro et al. 2018] G. Navarro, B. Sambale, and P. H. Tiep, "Characters and Sylow 2-subgroups of maximal class revisited", *J. Pure Appl. Algebra* **222**:11 (2018), 3721–3732. MR Zbl

[Olsson 1984] J. r. B. Olsson, "On the number of characters in blocks of finite general linear, unitary and symmetric groups", *Math. Z.* **186**:1 (1984), 41–47. MR Zbl

[Olsson 1992] J. r. B. Olsson, "The number of modular characters in certain p-blocks", *Proc. London Math. Soc.* (3) **65**:2 (1992), 245–264. MR Zbl

[Olsson 1993] J. B. Olsson, *Combinatorics and representations of finite groups*, Vorlesungen Fachbereich Math. Univ. Essen **20**, Univ. Essen, 1993. MR Zbl

[Pyber 1992] L. Pyber, "Finite groups have many conjugacy classes", *J. London Math. Soc.* (2) **46**:2 (1992), 239–249. MR Zbl

[Rizo et al. 2021] N. Rizo, A. A. Schaeffer Fry, and C. Vallejo, "Principal blocks with 5 irreducible characters", *J. Algebra* **585** (2021), 316–337. MR Zbl

[Robinson 2004] G. R. Robinson, "On Brauer's $k(B)$-problem for blocks of p-solvable groups with non-Abelian defect groups", *J. Algebra* **280**:2 (2004), 738–742. MR Zbl

[Sambale 2017] B. Sambale, "On blocks with abelian defect groups of small rank", *Results Math.* **71**:1-2 (2017), 411–422. MR Zbl

[Schaeffer Fry and Taylor 2023] A. A. Schaeffer Fry and J. Taylor, "Galois automorphisms and classical groups", *Transform. Groups* **28**:1 (2023), 439–486. MR Zbl

[Schmid 2007] P. Schmid, *The solution of the $k(GV)$ problem*, ICP Advanced Texts in Mathematics **4**, Imperial College Press, London, 2007. MR Zbl

[Späth 2013] B. Späth, "A reduction theorem for the Alperin–McKay conjecture", *J. Reine Angew. Math.* **680** (2013), 153–189. MR Zbl

[Taylor 2018] J. Taylor, "Action of automorphisms on irreducible characters of symplectic groups", *J. Algebra* **505** (2018), 211–246. MR Zbl

[Tiep 2015] P. H. Tiep, "Real ordinary characters and real Brauer characters", *Trans. Amer. Math. Soc.* **367**:2 (2015), 1273–1312. MR Zbl

Communicated by Michael J. Larsen
Received 2021-04-22 Revised 2022-01-16 Accepted 2022-07-06

hungnguyen@uakron.edu *Department of Mathematics, Buchtel College of Arts and Sciences, The University of Akron, Akron, OH, United States*

aschaef6@msudenver.edu *Department of Mathematics and Statistics, Metropolitan State University of Denver, Denver, CO, United States*

ALGEBRA AND NUMBER THEORY 17:6 (2023)
https://doi.org/10.2140/ant.2023.17.1153

Shintani–Barnes cocycles and values of the zeta functions of algebraic number fields

Hohto Bekki

We construct a new Eisenstein cocycle, called the Shintani–Barnes cocycle, which specializes in a uniform way to the values of the zeta functions of general number fields at positive integers. Our basic strategy is to generalize the construction of the Eisenstein cocycle presented in the work of Vlasenko and Zagier by using some recent techniques developed by Bannai, Hagihara, Yamada, and Yamamoto in their study of the polylogarithm for totally real fields. We also closely follow the work of Charollois, Dasgupta, and Greenberg. In fact, one of the key ingredients which enables us to deal with general number fields is the introduction of a new technique, called the "exponential perturbation", which is a slight modification of the Q-perturbation studied in their work.

1. Introduction

It is classically known that the Hecke integral formula [1917] expresses the zeta function of a number field of degree g as an integral of the Eisenstein series over a certain torus orbit on the locally symmetric space for $\mathrm{SL}_g(\mathbb{Z})$.

In some special cases, typically in the case where the number field is totally real, it is known that such an integral formula has a cohomological interpretation, and this often enables us to access the algebraic properties of the special values of the zeta function. More precisely, one can construct a certain $(g-1)$-cocycle on $\mathrm{SL}_g(\mathbb{Z})$ which can be thought as an algebraic counterpart of the Eisenstein series, and a $(g-1)$-cycle on $\mathrm{SL}_g(\mathbb{Z})$ which can be thought as an algebraic counterpart of the torus orbit, so that their pairing gives the value of the zeta function of a given totally real number field. Such a

MSC2020: primary 11R42; secondary 11F75, 55N91.
Keywords: Eisenstein cocycle, Shintani cocycle, special values of L-functions.

cocycle is often called the Eisenstein cocycle. Actually, many different kinds of Eisenstein cocycles have been constructed and studied by Harder [1987], Sczech [1993], Nori [1995], Solomon [1998], Hill [2007], Vlasenko and Zagier [2013], Charollois, Dasgupta, and Greenberg [Charollois et al. 2015], Beilinson, Kings, and Levin [Beilinson et al. 2018], Bergeron, Charollois, and Garcia [Bergeron et al. 2020], Flórez, Karabulut, and Wong [Flórez et al. 2019], Lim and Park [2019], Bannai, Hagihara, Yamada and Yamamoto [Bannai et al. 2023], and Sharifi and Venkatesh [2020], and various applications have been obtained. However, the number fields previously treated are basically limited to totally real fields or totally imaginary fields. The aim of this paper is to propose a new formulation in which we can treat all number fields in a uniform way.

1.1. ***Shintani cocycles.*** Among these many kinds of construction of the Eisenstein cocycle, a method we use in this paper is called Shintani's method, and the Eisenstein cocycles constructed by Shintani's method are often called the Shintani cocycles;[1] see [Solomon 1998; Hill 2007; Charollois et al. 2015; Lim and Park 2019; Bannai et al. 2023]. Roughly speaking, a Shintani cocycle is constructed as a family of objects (e.g., functions, formal power series, distributions, etc.) indexed by rational cones in $\mathbb{R}^g$. Therefore, what we do in this paper is basically the following:

(1) Define a certain object "ψ_C" for each rational cone $C \subset \mathbb{R}^g$.

(2) Prove that the family $(\psi_C)_C$ satisfies the "cocycle relation".

(3) Prove that the cohomology class defined by $(\psi_C)_C$ specializes to the special values of the zeta function of a given number field.

Let $g, k \geq 1$ be integers. In this paper, we say that a matrix $Q \in \mathrm{GL}_g(\mathbb{Q})$ is *irreducible* if its characteristic polynomial is irreducible over $\mathbb{Q}$. In Section 6, for a rational open cone

$$C_I = \sum_{i=1}^{g} \mathbb{R}_{>0}\alpha_i \subset \mathbb{R}^g$$

generated by $I = (\alpha_1, \ldots, \alpha_g) \in (\mathbb{Q}^g - \{0\})^g$, and an irreducible matrix $Q \in \mathrm{GL}_g(\mathbb{Q})$, we consider a holomorphic function

$$\psi^{Q}_{kg,I}(y) := \mathrm{sgn}(I) \sum_{x \in C_I^Q \cap \mathbb{Z}^g - \{0\}} \frac{1}{\langle x, y\rangle^{g+kg}}$$

on

$$\left\{ y \in \mathbb{C}^g \mid \text{there exists } \lambda \in \mathbb{C}^\times \text{ such that for all } i \in \{1, \ldots, g\},\ \mathrm{Re}(\langle \alpha_i, \lambda y\rangle) > 0 \right\} \subset \mathbb{C}^g - \{0\},$$

where

- $\mathrm{sgn}(I) = \mathrm{sgn}(\det(\alpha_1, \ldots, \alpha_g)) \in \{0, \pm 1\}$,
- the bracket $\langle x, y\rangle = {}^t x y$ denotes the dot product,
- C_I^Q is the "exponential Q-perturbation" of the cone C_I (Section 5.1).

[1] The terminology seems to depend on the authors. We adopt this convention in this paper.

Then we prove that the collection $(\psi_{kg,I}^{Q})_{I,Q}$ defines a class

$$[\Psi_{kg}] \in H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_{kg}^{\Xi})$$

of the equivariant cohomology of a certain $\mathrm{SL}_g(\mathbb{Z})$-equivariant sheaf $\mathscr{F}_{kg}^{\Xi}$ on $Y^\circ := \mathbb{C}^g - \boldsymbol{i}\mathbb{R}^g$; see Section 3 and Theorem 6.2.5. We call our Shintani cocycle the Shintani–Barnes cocycle because the function $\psi_{kg,I}^{Q}(y)$ is essentially the Barnes zeta function.

Then for a number field $F/\mathbb{Q}$ of degree g, a fractional ideal $\mathfrak{a} \subset F$, and a continuous map $\chi : F_{\mathbb{R}}^{\times} \to \mathbb{Z}$, we construct a specialization map

$$H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_{kg}^{\Xi}) \to H_{\mathrm{sing}}^{g-1}(F_{\mathbb{R}}^{\times}/\mathcal{O}_{F,+}^{\times}, \mathbb{C}) \to \mathbb{C},$$

using a certain integral operator; see (8-11). The image of the Shintani–Barnes cocycle $[\Psi_{kg}]$ under this specialization map can be computed using the classical Hurwitz formula (Proposition 7.1.3, Example 7.2.4) and a version of the Shintani cone decomposition (Proposition 8.2.1). As a result, we prove that the class $[\Psi_{kg}]$ maps to the value of the partial zeta function,

$$\pm \frac{\sqrt{D_{\mathcal{O}_F}}\, N\mathfrak{a}(k!)^g}{(g+gk-1)!} \zeta_{\mathcal{O}_F}(\boldsymbol{\varepsilon}^{k+1}\chi, \mathfrak{a}^{-1}, k+1),$$

under the specialization map, where $\boldsymbol{\varepsilon} : F_{\mathbb{R}}^{\times} \to \{\pm 1\}$ is the sign character; see Theorem 8.3.2.

The idea of using the Barnes zeta functions is based on the work of Vlasenko and Zagier [2013] dealing with the values of the zeta functions of real quadratic fields at positive integers, and the idea of constructing the Shintani cocycle as a Čech cocycle of an equivariant sheaf is based on the work of Bannai, Hagihara, Yamada, and Yamamoto [Bannai et al. 2023], in which the higher-dimensional polylogarithm associated to a totally real field is studied. Moreover, the concept of the exponential Q-perturbation C_I^Q of a cone C_I is a slight modification of the Q-perturbation studied by Charollois, Dasgupta, Greenberg [Charollois et al. 2015] and Yamamoto [2010]. We use irreducible matrices $Q \in \mathrm{GL}_g(\mathbb{Q})$ instead of the “irrational vectors” used in [Charollois et al. 2015]. These three ideas are the main ingredients in this paper which enable us to deal with general number fields.

1.2. ***Structure of the paper.*** Sections 2–5 are devoted to preparing some tools that are necessary for the definition of the Shintani–Barnes cocycle. More precisely, in Section 2 we review some elementary facts about irreducible matrices of $\mathrm{GL}_g(\mathbb{Q})$ and their relationship to number fields. In Section 3 we introduce the sheaves $\mathscr{F}_d$ and $\mathscr{F}_d^{\Xi}$ on $Y^\circ = \mathbb{C}^g - \boldsymbol{i}\mathbb{R}^g$, and examine the basic properties of these sheaves. Then in Section 4 we compute the equivariant cohomology groups of these sheaves using the equivariant Čech complex. In Section 5 we introduce the notion of the exponential perturbation, and prove the cocycle relation satisfied by rational cones. Based on these preparations, in Section 6 we give the definition of the Shintani–Barnes cocycle.

The remaining sections (Sections 7 and 8) are devoted to showing that we can obtain the special values of the zeta functions as a specialization of the Shintani–Barnes cocycle. In Section 7 we first introduce a certain integral operator, and construct the first half of the specialization map. In Section 8 we finish the construction of the specialization map using a version of the Shintani cone decomposition, and finally prove the main result, Theorem 8.3.2.

2. Preliminaries

Conventions.

- Throughout the paper we fix an integer $g \geq 1$.
- For a ring R, a vector $x \in R^g$ is always regarded as a column vector, and the matrix algebra $M_g(R)$ acts on R^g by the matrix multiplication from the left.
- For $x_1, \ldots, x_g \in R^g$, we often regard $(x_1, \ldots, x_g)$ as a $g \times g$-matrix whose columns are $x_1, \ldots, x_g$.
- For $\gamma \in M_g(R)$, its transpose is denoted by ${}^t\gamma \in M_g(R)$.
- The bracket
$$\langle\ ,\ \rangle : R^g \times R^g \to R, \quad (x, y) \mapsto \langle x, y\rangle = {}^t x y$$
denotes the standard scalar product (the dot product, not a Hermitian product even if $R = \mathbb{C}$).
- If A and B are sets, then $A - B$ denotes the relative complement of B in A.
- Let $\{S_\lambda\}_{\lambda\in\Lambda}$ be a family of sets. For $s \in \prod_{\lambda\in\Lambda} S_\lambda$, the λ-component of s is often denoted by $s_\lambda \in S_\lambda$.

2.1. *Irreducible matrices.* In this subsection we review some basic facts about irreducible matrices of $\mathrm{GL}_g(\mathbb{Q})$. We say that a matrix $Q \in \mathrm{GL}_g(\mathbb{Q})$ is *irreducible over* $\mathbb{Q}$ if the characteristic polynomial of Q is an irreducible polynomial over $\mathbb{Q}$. We often drop "over $\mathbb{Q}$" if it is obvious from the context. Let

$$\Xi := \{Q \in \mathrm{GL}_g(\mathbb{Q}) \mid Q \text{ is irreducible over } \mathbb{Q}\}$$

denote the set of irreducible matrices of $\mathrm{GL}_g(\mathbb{Q})$. The group $\mathrm{GL}_g(\mathbb{Q})$ acts on Ξ by the conjugate action. For $Q \in \Xi$ and $\gamma \in \mathrm{GL}_g(\mathbb{Q})$, let

$$[\gamma](Q) := \gamma Q \gamma^{-1} \in \Xi$$

denote this conjugate action.

Now, for $Q \in \Xi$, let

$$\Gamma_Q := \mathrm{Stab}_{\mathrm{SL}_g(\mathbb{Z})}(Q) = \left\{\gamma \in \mathrm{SL}_g(\mathbb{Z}) \mid [\gamma](Q) = \gamma Q \gamma^{-1} = Q\right\}$$

denote the subgroup of $\mathrm{SL}_g(\mathbb{Z})$ stabilizing Q. Moreover, let

$$F_Q := \mathbb{Q}[Q] \subset M_g(\mathbb{Q}) \quad \text{and} \quad \mathcal{O}_Q := F_Q \cap M_g(\mathbb{Z}) \subset F_Q$$

denote the subalgebras of $M_g(\mathbb{Q})$ generated by Q over $\mathbb{Q}$ and its "$M_g(\mathbb{Z})$-part" respectively.

Lemma 2.1.1. *Let $Q \in \Xi$, and let $f_Q(X) \in \mathbb{Q}[X]$ be the characteristic polynomial of Q.*

(1) *Q has g distinct eigenvalues in $\mathbb{C}$, and hence Q is diagonalizable in* $\mathrm{GL}_g(\mathbb{C})$.

(2) *There are no nonzero proper Q-stable $\mathbb{Q}$-subspaces of $\mathbb{Q}^g$.*

(3) *For any nonzero vector* $x \in \mathbb{Q}^g - \{0\}$, *the map*

$$F_Q \xrightarrow{\sim} \mathbb{Q}^g, \quad \gamma \mapsto \gamma x$$

is an isomorphism of $\mathbb{Q}$*-vector spaces.*

(4) *The* $\mathbb{Q}$*-algebra* F_Q *is a field of degree* g *over* $\mathbb{Q}$, *and we have*

$$N_{F_Q/\mathbb{Q}}(\gamma) = \det \gamma$$

for $\gamma \in F_Q$, *where* $N_{F_Q/\mathbb{Q}}$ *is the norm of the field extension* $F_Q/\mathbb{Q}$.

(5) *We have*

$$F_Q = \{\gamma \in M_g(\mathbb{Q}) \mid \gamma Q = Q\gamma\}.$$

(6) *We have*

$$\Gamma_Q = \{\gamma \in \mathcal{O}_Q \mid N_{F_Q/\mathbb{Q}}(\gamma) = 1\} \subset \mathcal{O}_Q^\times,$$

i.e., Γ_Q *is the norm-one unit group of* $\mathcal{O}_Q$.

(7) *The action of* Γ_Q *on* $\mathbb{Q}^g - \{0\}$ *is free*, *i.e.*, *for any* $x \in \mathbb{Q}^g - \{0\}$ *and* $\gamma \in \Gamma_Q$, *we have* $\gamma x = x$ *if and only if* $\gamma = 1$.

Proof. (1) This follows from the fact that $f_Q(X)$ is an irreducible polynomial over $\mathbb{Q}$.

(2) This also follows from the irreducibility of $f_Q(X)$. Indeed, if $V \subset \mathbb{Q}^g$ is a Q-stable $\mathbb{Q}$-subspace, then the characteristic polynomial of $Q|_V$ divides $f_Q(X)$.

(3) and (4) First, since $x \neq 0$, the image of the map

$$F_Q \to \mathbb{Q}^g, \quad \gamma \mapsto \gamma x$$

is a nonzero Q-stable $\mathbb{Q}$-subspace. Hence, by (2), this map is surjective. Now, again since $f_Q(X)$ is an irreducible polynomial over $\mathbb{Q}$, we see that $F_Q \simeq \mathbb{Q}[X]/(f_Q(X))$ is a field of degree g over $\mathbb{Q}$. Therefore, by comparing the dimension, we find that the above map is an isomorphism. The identity $N_{F_Q/\mathbb{Q}}(\gamma) = \det \gamma$ is nothing but the definition of the norm.

(5) Let F'_Q denote the right-hand side. The inclusion $F_Q \subset F'_Q$ is obvious. We compare the dimension. First we have

$$F'_Q \otimes_{\mathbb{Q}} \mathbb{C} \subset F''_Q := \{\gamma \in M_g(\mathbb{C}) \mid \gamma Q = Q\gamma\}.$$

Then, by (1), the right-hand side F''_Q is simultaneously diagonalizable in $M_g(\mathbb{C})$. Therefore, F''_Q is isomorphic to the space of diagonal matrices. Thus we find

$$\dim_{\mathbb{Q}} F'_Q = \dim_{\mathbb{C}} F'_Q \otimes_{\mathbb{Q}} \mathbb{C} \leq \dim_{\mathbb{C}} F''_Q = g = \dim_{\mathbb{Q}} F_Q,$$

and hence we obtain $F_Q = F'_Q$.

(6) This follows directly from (4) and (5).

(7) By (6), we see that $\Gamma_Q \subset F_Q^\times$, and by (3) and (4), we see that $F_Q^\times$ acts freely on $\mathbb{Q}^g - \{0\}$. $\square$

2.2. ***Review on number fields.*** In this subsection we take a closer look at the relationship between irreducible matrices and number fields.

Let $F/\mathbb{Q}$ be a number field of degree g, and let

$$\tau_1, \dots, \tau_g : F \hookrightarrow \mathbb{C}$$

be the field embeddings of F into $\mathbb{C}$, i.e., $\{\tau_1, \dots, \tau_g\} = \mathrm{Hom}_{\mathrm{field}}(F, \mathbb{C})$.[2] Let $\mathcal{O} \subset F$ be an order in F, i.e., $\mathcal{O} \subset F$ is a subring which is a finitely generated $\mathbb{Z}$-module and generates F over $\mathbb{Q}$. Let $\mathfrak{a} \subset F$ be a proper fractional $\mathcal{O}$-ideal, i.e., $\mathfrak{a} \subset F$ is a finitely generated $\mathcal{O}$-submodule such that

$$\{\alpha \in F \mid \alpha\mathfrak{a} \subset \mathfrak{a}\} = \mathcal{O}. \tag{2-1}$$

Let $w_1, \dots, w_g \in \mathfrak{a}$ be a basis of $\mathfrak{a}$ over $\mathbb{Z}$, and put

$$w := {}^t(w_1, \dots, w_g) \in F^g \quad \text{and} \quad w^{(i)} := \tau_i(w) = {}^t(\tau_i(w_1), \dots, \tau_i(w_g)) \in \mathbb{C}^g$$

for $i = 1, \dots, g$. We define the norm polynomial $N_w(x) = N_w(x_1, \dots, x_g) \in \mathbb{Q}[x_1, \dots, x_g]$ with respect to this basis by

$$N_w(x) := \prod_{i=1}^{g} \langle x, w^{(i)} \rangle \in \mathbb{Q}[x_1, \dots, x_g],$$

where $x = (x_1, \dots, x_g)$. The situation can be summarized in the following diagram:

$$\begin{tikzcd}
x \arrow[d, mapsto] & \in & \mathbb{Z}^g \arrow[d, "\wr"] & \subset & \mathbb{Q}^g \arrow[d, "\wr"] \arrow[dr, "N_w"] & \\
\langle x, w \rangle & \in & \mathfrak{a} & \subset & F \arrow[r, "N_{F/\mathbb{Q}}"'] & \mathbb{Q}
\end{tikzcd}$$

Moreover, let

$$\rho_w : F \to M_g(\mathbb{Q})$$

be the regular representation of F with respect to the basis $w_1, \dots, w_g$, i.e., for $\alpha \in F$ and $x \in \mathbb{Q}^g$, we have

$$\langle \rho_w(\alpha)x, w \rangle = \alpha \langle x, w \rangle = \langle x, \alpha w \rangle \in F. \tag{2-2}$$

Dual objects. Let $w_1^*, \dots, w_g^* \in F$ be the dual basis of $w_1, \dots, w_g$ with respect to the field trace $\mathrm{Tr}_{F/\mathbb{Q}}$, i.e.,

$$\mathrm{Tr}_{F/\mathbb{Q}}(w_i w_j^*) = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

Then it is easy to see that $w_1^*, \dots, w_g^*$ form a $\mathbb{Z}$-basis of a proper fractional $\mathcal{O}$-ideal

$$\mathfrak{a}^* := \{\alpha \in F \mid \mathrm{Tr}_{F/\mathbb{Q}}(\alpha\mathfrak{a}) \subset \mathbb{Z}\}.$$

[2] At this stage we don't make a distinction between real embeddings and complex embeddings. Later, in Section 8.2, we will make such a distinction for convenience.

We define

$$w^* := {}^t(w_1^*, \dots, w_g^*) \in F^g,$$
$$w^{*(i)} := \tau_i(w^*) = {}^t(\tau_i(w_1^*), \dots, \tau_i(w_g^*)) \in \mathbb{C}^g,$$
$$N_{w^*}(x) := \prod_{i=1}^{g} \langle x, w^{*(i)} \rangle \in \mathbb{Q}[x_1, \dots, x_g],$$

and

$$\rho_{w^*} : F \to M_g(\mathbb{Q})$$

in the same way as above, starting from the dual basis $w_1^*, \dots, w_g^*$.

Lemma 2.2.1. *Let $\theta \in F^\times$ be an element such that $F = \mathbb{Q}(\theta)$. Put $Q = \rho_w(\theta) \in \mathrm{GL}_g(\mathbb{Q})$.*

(1) *We have $Q \in \Xi$. Conversely, any element of Ξ can be obtained in this way.*

(2) *The regular representation $\rho_w : F \to M_g(\mathbb{Q})$ induces isomorphisms*

$$\begin{array}{ccc} F & \xrightarrow[\sim]{\rho_w} & F_Q \\ \cup & & \cup \\ \mathcal{O} & \xrightarrow[\sim]{} & \mathcal{O}_Q \\ \cup & & \cup \\ \mathcal{O}^1 & \xrightarrow[\sim]{} & \Gamma_Q \end{array}$$

where $\mathcal{O}^1 := \{u \in \mathcal{O}^\times \mid N_{F/\mathbb{Q}}(u) = 1\}$ is the norm-one unit group of $\mathcal{O}$.

(3) *$w^{*(1)}, \dots, w^{*(g)} \in \mathbb{C}^g$ are the dual basis of $w^{(1)}, \dots, w^{(g)} \in \mathbb{C}^g$ with respect to the scalar product $\langle\ ,\ \rangle$, i.e., we have*

$$\langle w^{*(i)}, w^{(j)} \rangle = \delta_{ij}.$$

(4) *For $\alpha \in F$, we have*

$$\rho_{w^*}(\alpha) = {}^t\rho_w(\alpha).$$

(5) *Let $\alpha \in F$. Then $w^{(i)}$ is an eigenvector of ${}^t\rho_w(\alpha)$ with eigenvalue $\tau_i(\alpha)$.*

(6) *Let $\alpha \in F$. Then $w^{*(i)}$ is an eigenvector of $\rho_w(\alpha)$ with eigenvalue $\tau_i(\alpha)$.*

(7) *For $\gamma \in \Gamma_Q$, we have*

$$N_w(\gamma x) = N_w(x) \quad \text{and} \quad N_{w^*}({}^t\gamma x) = N_{w^*}(x).$$

Proof. (1) Since θ generates F, the characteristic polynomial of $Q = \rho_w(\theta)$ is irreducible, and hence $Q \in \Xi$. The latter half of the statement follows from Lemma 2.1.1(3), (4). Indeed, for $Q \in \Xi$, fix a nonzero vector $x \in \mathbb{Q}^g$ and take a basis $w_1, \dots, w_g \in F_Q$ corresponding to the standard basis of $\mathbb{Q}^g$ via the isomorphism

$$F_Q \xrightarrow{\sim} \mathbb{Q}^g, \quad \gamma \mapsto \gamma x.$$

Let $\mathfrak{a} \subset F_Q$ be the subset corresponding to $\mathbb{Z}^g \subset \mathbb{Q}^g$ under this isomorphism. Then we easily see that $\mathfrak{a}$ is a proper $\mathcal{O}_Q$-ideal and that ρ_w is the natural inclusion $F_Q \hookrightarrow M_g(\mathbb{Q})$. Hence we find that $Q = \rho_w(Q)$.

(2) The first isomorphism $F \xrightarrow{\sim} F_Q$ is obvious. The second isomorphism follows from (2-1), and the third follows from Lemma 2.1.1(6).

(3) Put

$$W := (w^{(1)}, \dots, w^{(g)}) = (\tau_j(w_i))_{ij} \in M_g(\mathbb{C}) \quad \text{and} \quad W^* := (w^{*(1)}, \dots, w^{*(g)}) = (\tau_j(w_i^*))_{ij} \in M_g(\mathbb{C}).$$

Then, by definition, we have

$$W^tW^* = (\mathrm{Tr}_{F/\mathbb{Q}}(w_i w_j^*))_{ij} = 1 \in M_g(\mathbb{C}), \tag{2-3}$$

and hence

$$(\langle w^{*(i)}, w^{(j)}\rangle)_{ij} = {}^tW^*W = 1.$$

(4)–(6) First, by (2-2), we have

$$\langle x, \alpha w\rangle = \langle \rho_w(\alpha)x, w\rangle = \langle x, {}^t\rho_w(\alpha)w\rangle \in F$$

for all $x \in \mathbb{Q}^g$. Therefore, we find that $\alpha w = {}^t\rho_w(\alpha)w \in F^g$. By applying τ_i, we obtain (5). In particular,

$$W\mathrm{diag}(\tau_1(\alpha), \dots, \tau_g(\alpha)) = {}^t\rho_w(\alpha)W, \tag{2-4}$$

where $\mathrm{diag}(\tau_1(\alpha), \dots, \tau_g(\alpha)) \in M_g(\mathbb{C})$ is the diagonal matrix with diagonal entries $\tau_1(\alpha), \dots, \tau_g(\alpha)$. Similarly, we have

$$W^*\mathrm{diag}(\tau_1(\alpha), \dots, \tau_g(\alpha)) = {}^t\rho_{w^*}(\alpha)W^*. \tag{2-5}$$

On the other hand, by using (2-3) and (2-4) we also find that

$$\mathrm{diag}(\tau_1(\alpha), \dots, \tau_g(\alpha)){}^tW^* = {}^tW^*{}^t\rho_w(\alpha),$$

and hence, by taking the transpose, we have

$$W^*\mathrm{diag}(\tau_1(\alpha), \dots, \tau_g(\alpha)) = \rho_w(\alpha)W^*. \tag{2-6}$$

By comparing (2-5) and (2-6), we obtain (4) and (6).

(7) This follows from (2), (5), and (6). Indeed, take $u \in \mathcal{O}^1$ such that $\rho_w(u) = \gamma$. Then we have

$$N_w(\gamma x) = \prod_{i=1}^{g}\langle \gamma x, w^{(i)}\rangle = \prod_{i=1}^{g}\langle x, {}^t\rho_w(u)w^{(i)}\rangle = N_{F/\mathbb{Q}}(u)N_w(x) = N_w(x).$$

The statement for $N_{w^*}(x)$ can be proved similarly. $\square$

3. The space Y° and the sheaves $\mathscr{F}_d$ and $\mathscr{F}_d^\Xi$

3.1. *Definitions.* Let $\mathbb{P}^{g-1}(\mathbb{C}) = (\mathbb{C}^g - \{0\})/\mathbb{C}^\times$ be the complex projective $(g-1)$−space, and let

$$\pi_\mathbb{C} : \mathbb{C}^g - \{0\} \to \mathbb{P}^{g-1}(\mathbb{C})$$

be the natural projection. We define an open subset Y° of $\mathbb{C}^g - \{0\}$ by

$$Y^\circ := \mathbb{C}^g - \boldsymbol{i}\mathbb{R}^g \subset \mathbb{C}^g - \{0\},$$

where $\boldsymbol{i} \in \mathbb{C}$ is the imaginary unit. The group $\mathrm{GL}_g(\mathbb{Q})$ acts on $\mathbb{C}^g - \{0\}$, Y°, and $\mathbb{P}^{g-1}(\mathbb{C})$ by the matrix action from the left. For an integer $d \geq 0$, we define a sheaf $\mathscr{F}_d$ on Y° as

$$\mathscr{F}_d := \pi_\mathbb{C}^{-1} \Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)|_{Y^\circ},$$

where $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$ is the $(-d)$-th Serre twist of the sheaf $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}$ of holomorphic $(g-1)$-forms on $\mathbb{P}^{g-1}(\mathbb{C})$, and $\pi_\mathbb{C}^{-1}$ is the inverse image functor of sheaves. Furthermore, we define

$$\mathscr{F}_d^\Xi := \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{F}_d) \simeq \prod_{Q \in \Xi} \mathscr{F}_d,$$

where $\underline{\mathbb{Z}[\Xi]}$ is the constant sheaf associated to the free abelian group $\mathbb{Z}[\Xi]$ generated by the set Ξ of irreducible matrices of $\mathrm{GL}_g(\mathbb{Q})$, and $\underline{\mathrm{Hom}}$ is the sheaf Hom. For $Q \in \Xi$, let

$$\mathrm{ev}_Q : \mathscr{F}_d^\Xi \to \mathscr{F}_d \tag{3-1}$$

denote the evaluation map at Q. See Remark 3.1.1 below.

Remark 3.1.1. (1) More generally, for a sheaf $\mathscr{F}$ (of abelian groups) on Y°, we define

$$\mathscr{F}^\Xi := \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{F}).$$

Note that for an open subset $U \subset Y^\circ$, we have

$$\Gamma(U, \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{F})) = \mathrm{Hom}(\underline{\mathbb{Z}[\Xi]}|_U, \mathscr{F}|_U) = \mathrm{Hom}(\mathbb{Z}[\Xi], \Gamma(U, \mathscr{F})) = \mathrm{Map}(\Xi, \Gamma(U, \mathscr{F})).$$

Then the evaluation map $\mathrm{ev}_Q : \mathscr{F}^\Xi \to \mathscr{F}$ is given by

$$\mathrm{ev}_Q : \Gamma(U, \mathscr{F}^\Xi) = \mathrm{Map}(\Xi, \Gamma(U, \mathscr{F})) \to \Gamma(U, \mathscr{F}), \quad \phi \mapsto \phi(Q).$$

(2) By (1) we also see that $\mathscr{F}^\Xi \simeq \prod_{Q\in\Xi} \mathscr{F}$.

(3) The sheaf $\mathscr{F}_d^\Xi$ is an analogue of the group $\mathcal{N}$ considered in [Charollois et al. 2015].

Remark 3.1.2. The sections of the sheaf $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$ on an open subset $U \subset \mathbb{P}^{g-1}(\mathbb{C})$ can be described as follows. First, let ω be a holomorphic $(g-1)$-form on $\mathbb{C}^g - \{0\}$ defined by

$$\omega(y_1, \dots, y_g) := \sum_{i=1}^{g} (-1)^{i-1} y_i \, dy_1 \wedge \cdots \wedge \check{dy_i} \wedge \cdots \wedge dy_g$$

for $y = {}^t(y_1, \dots, y_g) \in \mathbb{C}^g - \{0\}$, where $\check{dy_i}$ means that dy_i is omitted.

Then we have

$$\begin{aligned}&\Gamma(U, \Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d))\\ &\quad\simeq \{f\omega \mid f \text{ holomorphic function on } \pi_{\mathbb{C}}^{-1}(U) \text{ such that } f(\lambda y) = \lambda^{-g-d} f(y) \text{ for all } \lambda \in \mathbb{C}^{\times}\}. \qquad (3\text{-}2)\end{aligned}$$

In this paper we use this as a definition of the sheaf $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$.

The sheaf $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$ has a natural $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure via the pullback of differential forms. Since $\pi_{\mathbb{C}}$ is a $\mathrm{GL}_g(\mathbb{Q})$-equivariant map, this induces $\mathrm{GL}_g(\mathbb{Q})$-equivariant structures on $\mathscr{F}_d$ and $\mathscr{F}_d^{\Xi}$. We describe these $\mathrm{GL}_g(\mathbb{Q})$-equivariant structures more explicitly in Section 3.3.

3.2. *A vanishing result.* Here our aim is to compute the cohomology group $H^q(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d))$ for convex open subsets $U \subset \mathbb{C}^g - \{0\}$. Actually, we will show that

$$H^q(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)) = 0$$

for $q \geq 1$, and also give an explicit description of $H^0(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d))$.

Let

$$\mathbb{D} := \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0\}$$

be the right half-plane. We start with the following elementary lemma.

Lemma 3.2.1. *Let X be a paracompact manifold, and let $\mathrm{pr}_1 : X \times \mathbb{D} \to X$ be the first projection. Let $U \subset X \times \mathbb{D}$ be an open subset such that for any $x \in X$, the set*

$$\{z \in \mathbb{D} \mid (x, z) \in U\}$$

is a nonempty convex subset of $\mathbb{D}$. Then there exists a continuous section $s : X \to U$ of $\mathrm{pr}_1|_U : U \to X$ such that $s \circ \mathrm{pr}_1$ is homotopic to the identity map id_U over X, i.e., there exists a continuous map

$$h : [0, 1] \times U \to U$$

such that $h(0, u) = s \circ \mathrm{pr}_1(u)$, $h(1, u) = u$, and $\mathrm{pr}_1 \circ h(t, u) = \mathrm{pr}_1(u)$ for $t \in [0, 1]$ and $u \in U$.

Proof. In order to construct a section, it suffices to construct a continuous map

$$f : X \to \mathbb{D}$$

such that $(x, f(x)) \in U$ for all $x \in X$. First, by assumption, for each $x \in X$ we can take $z_x \in \mathbb{D}$ such that $(x, z_x) \in U$. Then there exist an open neighborhood $U_x \subset X$ of x and an open neighborhood $V_x \subset \mathbb{D}$ of z_x such that $U_x \times V_x \subset U$. Since $X = \bigcup_{x \in X} U_x$ and X is paracompact, there exists a subset $\Lambda \subset X$ such that $\{U_\lambda\}_{\lambda \in \Lambda}$ is a locally finite open covering of X. Note that for $x \in U_\lambda$, we have

$$(x, z_\lambda) \in U_\lambda \times V_\lambda \subset U.$$

By using the paracompactness once again, there exists a partition of unity with respect to the open covering $\{U_\lambda\}_{\lambda \in \Lambda}$, i.e., a collection $\{\phi_\lambda\}_{\lambda \in \Lambda}$ of continuous maps

$$\phi_\lambda : X \to [0, 1]$$

such that $\operatorname{supp}(\phi_\lambda) \subset U_\lambda$ and $\sum_{\lambda\in\Lambda} \phi_\lambda(x) = 1$ for all $x \in X$. Put

$$f := \sum_{\lambda\in\Lambda} z_\lambda \phi_\lambda : X \to \mathbb{D}.$$

Then, by the convexity assumption, we see that

$$(x, f(x)) = \Bigl(x, \sum_{\lambda\in\Lambda} z_\lambda \phi_\lambda(x)\Bigr) \in U$$

for all $x \in X$. Thus we obtain a section

$$s : X \to U, \quad x \mapsto (x, f(x)).$$

Again by the convexity assumption, we see that $s \circ \mathrm{pr}_1$ is homotopic to the identity map id_U over X. Indeed,

$$h : [0, 1] \times U \to U, \quad (t, (x, z)) \mapsto (x, tz + (1 - t) f(x))$$

gives a homotopy between $s \circ \mathrm{pr}_1$ and id_U over X. $\square$

Lemma 3.2.2. *Let $U \subset \mathbb{C}^g - \{0\}$ be a convex open subset.*

(1) *There exists $x \in \mathbb{C}^g - \{0\}$ such that $U \subset V_x := \bigl\{y \in \mathbb{C}^g - \{0\} \mid \mathrm{Re}(\langle x, y\rangle) > 0\bigr\}$.*

(2) *The projection $\pi_{\mathbb{C}}|_U : U \to \pi_{\mathbb{C}}(U)$ has a continuous section $s : \pi_{\mathbb{C}}(U) \to U$ such that $s \circ \pi_{\mathbb{C}}|_U$ is homotopic to the identity map id_U over $\pi_{\mathbb{C}}(U)$.*

(3) *The image $\pi_{\mathbb{C}}(U)$ is a Stein manifold.*

Proof. (1) By the so-called hyperplane separation theorem [Rudin 1991, Theorem 3.4(a)] applied to U and $\{0\}$, there exist $x \in \mathbb{C}^g - \{0\}$ and $\mu \in \mathbb{R}$ such that

$$0 = \mathrm{Re}(\langle x, 0\rangle) \le \mu < \mathrm{Re}(\langle x, y\rangle)$$

for all $y \in U$, and hence $U \subset V_x = \bigl\{y \in \mathbb{C}^g - \{0\} \mid \mathrm{Re}(\langle x, y\rangle) > 0\bigr\}$.

(2) We first construct a section $s_x : \pi_{\mathbb{C}}(V_x) \to V_x$ of $\pi_{\mathbb{C}}|_{V_x}$ as follows. Set

$$V_x^1 := \bigl\{y \in \mathbb{C}^g - \{0\} \mid \langle x, y\rangle = 1\bigr\} \subset V_x.$$

Then we easily see that $\pi_{\mathbb{C}}|_{V_x^1} : V_x^1 \xrightarrow{\sim} \pi_{\mathbb{C}}(V_x)$ is a biholomorphism. Thus we define

$$s_x := (\pi_{\mathbb{C}}|_{V_x^1})^{-1} : \pi_{\mathbb{C}}(V_x) \xrightarrow{\sim} V_x^1 \subset V_x$$

to be the inverse map of $\pi_{\mathbb{C}}|_{V_x^1}$, which is clearly a section of $\pi_{\mathbb{C}}|_{V_x}$. Then we have a trivialization φ of $\pi_{\mathbb{C}}|_{V_x}$

$$\begin{tikzcd} \pi_{\mathbb{C}}(V_x) \times \mathbb{D} \arrow[rr, "\varphi", "\sim"'] \arrow[dr, "\mathrm{pr}_1"'] & & V_x \arrow[dl, "\pi_{\mathbb{C}}|_{V_x}"] \\ & \pi_{\mathbb{C}}(V_x) & \end{tikzcd}$$

defined by $\varphi(z, \lambda) := \lambda s_x(z)$ for $(z, \lambda) \in \pi_{\mathbb{C}}(V_x) \times \mathbb{D}$.

Therefore, it suffices to construct a continuous section s' of

$$p := \mathrm{pr}_1|_{\varphi^{-1}(U)} : \varphi^{-1}(U) \xrightarrow{\mathrm{pr}_1} \pi_{\mathbb{C}}(U)$$

such that $s' \circ p$ is homotopic to $\mathrm{id}_{\varphi^{-1}(U)}$ over $\pi_{\mathbb{C}}(U)$. By Lemma 3.2.1, it suffices to show the following:

Claim. *For any $z \in \pi_{\mathbb{C}}(U)$, the set*

$$\mathbb{D}_z := \{\lambda \in \mathbb{D} \mid (z, \lambda) \in \varphi^{-1}(U)\}$$

is a nonempty convex subset of $\mathbb{D}$.

Proof of claim. Let $z \in \pi_{\mathbb{C}}(U)$. The set $\mathbb{D}_z$ is obviously nonempty. Suppose that $\lambda, \lambda' \in \mathbb{D}_z$, i.e., $\lambda s_x(z), \lambda' s_x(z) \in U$. Then for $t \in [0, 1]$, we have $(t\lambda + (1-t)\lambda')s_x(z) \in U$ because U is convex, and hence $t\lambda + (1-t)\lambda' \in \mathbb{D}_z$. □

(3) From the above argument, we see that $\pi_{\mathbb{C}}(U)$ is an open subset of

$$\pi_{\mathbb{C}}(V_x) \simeq V_x^1 \simeq \mathbb{C}^{g-1}.$$

Since every pseudoconvex open subset of $\mathbb{C}^{g-1}$ is a Stein manifold (see [Hörmander 1973, Theorem 4.2.8, Example after Definition 5.1.3]), it suffices to see that $\pi_{\mathbb{C}}(U)$ is pseudoconvex. This follows, for example, from [Hörmander 1994, Proposition 4.6.3, Theorem 4.6.8]. (Use [Hörmander 1994, Theorem 4.6.8] for $X = U$, $z_0 = 0$, and $L(y) = \langle x, y\rangle$. Note that a convex set U is obviously $\mathbb{C}$ convex.) □

Proposition 3.2.3. *Let $U \subset \mathbb{C}^g - \{0\}$ be a convex open subset.*

(1) *The natural map*

$$H^q(\pi_{\mathbb{C}}(U), \Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)) \xrightarrow{\sim} H^q(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d))$$

is an isomorphism for all $q \geq 0$.

(2) *Under this identification, we have*

$$\Gamma(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d))$$
$$= \{f\omega \mid f \text{ holomorphic function on } \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U)) \text{ such that } f(\lambda y) = \lambda^{-g-d} f(y) \text{ for all } \lambda \in \mathbb{C}^\times\}.$$

(3) *For all $q \geq 1$, we have*

$$H^q(U, \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)) = 0.$$

Proof. (1) This follows from Lemma 3.2.2(2) and [Kashiwara and Schapira 1990, Corollary 2.7.7(ii)].

(2) This follows directly from (1) and the description of $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$; see Remark 3.1.2.

(3) By Lemma 3.2.2(3), we know $\pi_{\mathbb{C}}(U)$ is a Stein manifold. Moreover, $\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)$ is a coherent sheaf on $\mathbb{P}^{g-1}(\mathbb{C})$. So (3) follows from (1) and Cartan's Theorem B; see [Hörmander 1973, Theorem 7.4.3]. □

3.3. $\mathrm{GL}_g(\mathbb{Q})$*-equivariant structures.* In this subsection we explicitly describe the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structures on $\mathscr{F}_d$ and $\mathscr{F}_d^{\Xi}$.

In this paper, for a subgroup $G \subset \mathrm{GL}_g(\mathbb{Q})$ and a sheaf $\mathscr{F}$ (of abelian groups) on Y°, we define a G-equivariant structure on $\mathscr{F}$ to be a collection $\{[\gamma]\}_{\gamma \in G}$ of isomorphisms

$$[\gamma] : \mathscr{F} \xrightarrow{\sim} ({}^t\gamma)_* \mathscr{F}$$

subject to the conditions

(i) $[1] = \mathrm{id}_{\mathscr{F}}$,

(ii) $[\gamma_1\gamma_2] = ({}^t\gamma_2)_*[\gamma_1] \circ [\gamma_2]$ for all $\gamma_1, \gamma_2 \in G$.

Here, ${}^t\gamma$ is the transpose matrix of γ, and $({}^t\gamma)_*\mathscr{F}$ (resp. $({}^t\gamma_2)_*[\gamma_1]$) is the direct image of $\mathscr{F}$ (resp. $[\gamma_1]$) with respect to the map ${}^t\gamma : Y^\circ \to Y^\circ$ (resp. ${}^t\gamma_2 : Y^\circ \to Y^\circ$).[3]

The $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d$ can be defined as follows. First, by Proposition 3.2.3(2),

$$\Gamma(U, \mathscr{F}_d) = \{ f\omega \mid f \text{ holomorphic function on } \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U)) \text{ such that } f(\lambda y) = \lambda^{-g-d} f(y) \text{ for all } \lambda \in \mathbb{C}^\times \}$$

for a convex open subset $U \subset Y^\circ$, where

$$\omega(y_1, \dots, y_g) := \sum_{i=1}^{g} (-1)^{i-1} y_i \, dy_1 \wedge \cdots \wedge \check{dy_i} \wedge \cdots \wedge dy_g.$$

Lemma 3.3.1. *For $\gamma \in \mathrm{GL}_g(\mathbb{Q})$, we have*

$$\omega(\gamma y) = \det(\gamma)\omega(y).$$

Proof. It suffices to prove the identity for elementary matrices γ. This case can be checked easily. $\square$

Definition 3.3.2. For $\gamma \in \mathrm{GL}_g(\mathbb{Q})$ and a convex open subset $U \subset Y^\circ$, let $[\gamma]_U$ denote the pullback map

$$\begin{aligned} [\gamma]_U : \Gamma(U, \mathscr{F}_d) &\xrightarrow{\sim} \Gamma(U, ({}^t\gamma)_*\mathscr{F}_d) = \Gamma(({}^t\gamma)^{-1}U, \mathscr{F}_d), \\ f(y)\omega(y) &\longmapsto f({}^t\gamma y)\omega({}^t\gamma y) = \det(\gamma) f({}^t\gamma y)\omega(y). \end{aligned}$$

Here $f({}^t\gamma y)$ is regarded as a holomorphic function of $y \in ({}^t\gamma)^{-1}\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U)) = \pi_{\mathbb{C}}^{-1}\big(\pi_{\mathbb{C}}(({}^t\gamma)^{-1}U)\big)$. We may drop the subscript U and write as $[\gamma] = [\gamma]_U$ if there is no confusion.

Lemma 3.3.3. (1) *Let $V, U \subset Y^\circ$ be convex open subsets such that $V \subset U$, and let $s \in \Gamma(U, \mathscr{F}_d)$ be a section. Then we have*

$$[\gamma]_U(s)|_V = [\gamma]_V(s|_V)$$

in $\Gamma(V, ({}^t\gamma)_*\mathscr{F}_d)$.

(2) *The collection $\{[\gamma]_U \mid U \subset Y^\circ$ convex open$\}$ defines an isomorphism of sheaves*

$$[\gamma] : \mathscr{F}_d \xrightarrow{\sim} ({}^t\gamma)_*\mathscr{F}_d.$$

(3) *The collection $\{[\gamma]\}_{\gamma \in \mathrm{GL}_g(\mathbb{Q})}$ defines a $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d$.*

Proof. (1) is clear, and (2) follows from (1) since convex open subsets form a basis of open subsets of Y°. We prove (3).

[3] We consider the action of ${}^t\gamma$ on Y° instead of γ since it is more convenient later when we use the identity $\langle \gamma x, y \rangle = \langle x, {}^t\gamma y \rangle$.

Condition (i) of the definition is obvious.

Let $U \subset Y^\circ$ be a convex open subset, and let $s(y) = f(y)\omega(y) \in \Gamma(U, \mathscr{F}_d)$ be a section. Then for $\gamma_1, \gamma_2 \in \mathrm{GL}_g(\mathbb{Q})$, we have

$$({}^t\gamma_2)_*[\gamma_1] \circ [\gamma_2](s(y)) = [\gamma_1]_{{}^t\gamma_2^{-1}U} \circ [\gamma_2]_U(s(y)) = [\gamma_1]_{{}^t\gamma_2^{-1}U}(s({}^t\gamma_2 y)) = s({}^t\gamma_2{}^t\gamma_1 y) = [\gamma_1\gamma_2](s(y)).$$

Since convex open subsets form a basis of open subsets of Y°, this shows condition (ii). □

This describes the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d$. Next we describe the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d^\Xi$. First, note that the conjugate action

$$[\gamma] : \mathbb{Z}[\Xi] \xrightarrow{\sim} \mathbb{Z}[\Xi], \quad Q \mapsto [\gamma](Q) = \gamma Q \gamma^{-1}$$

of $\mathrm{GL}_g(\mathbb{Q})$ on $\mathbb{Z}[\Xi]$ naturally induces a $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on the associated constant sheaf $\underline{\mathbb{Z}[\Xi]}$. Therefore, for a $\mathrm{GL}_g(\mathbb{Q})$-equivariant sheaf $\mathscr{F}$, the sheaf

$$\mathscr{F}^\Xi = \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{F})$$

has a natural $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure induced from those of $\underline{\mathbb{Z}[\Xi]}$ and $\mathscr{F}$. In particular, we obtain a $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d^\Xi$.

More concretely, for an open subset $U \subset Y^\circ$ and a section

$$\phi \in \Gamma(U, \mathscr{F}^\Xi) = \mathrm{Map}(\Xi, \Gamma(U, \mathscr{F}))$$

(see Remark 3.1.1) the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}^\Xi$ can be computed as

$$[\gamma](\phi)(Q) = [\gamma]\big(\phi\big([\gamma^{-1}](Q)\big)\big) = [\gamma](\phi(\gamma^{-1}Q\gamma))$$

for $\gamma \in \mathrm{GL}_g(\mathbb{Q})$ and $Q \in \Xi$. In particular, we see that for $Q \in \Xi$, the evaluation map

$$\mathrm{ev}_Q : \mathscr{F}^\Xi \to \mathscr{F}$$

(see Remark 3.1.1) is a Γ_Q-equivariant map, where $\Gamma_Q = \mathrm{Stab}_{\mathrm{SL}_g(\mathbb{Z})}(Q) \subset \mathrm{SL}_g(\mathbb{Z})$ is the stabilizer of $Q \in \Xi$ in $\mathrm{SL}_g(\mathbb{Z})$.

4. Equivariant cohomology

Recall that $\Gamma_Q = \mathrm{Stab}_{\mathrm{SL}_g(\mathbb{Z})}(Q) \subset \mathrm{SL}_g(\mathbb{Z})$ denotes the stabilizer of $Q \in \Xi$ in $\mathrm{SL}_g(\mathbb{Z})$. In this section we compute the equivariant cohomology groups

$$H^q(Y^\circ, \Gamma_Q, \mathscr{F}_d) \quad \text{and} \quad H^q(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^\Xi)$$

using the equivariant Čech complexes; see Corollary 4.3.4. We closely follow the argument in [Bannai et al. 2023].

Here, for a subgroup $G \subset \mathrm{GL}_g(\mathbb{Q})$, the equivariant cohomology

$$H^q(Y^\circ, G, -) : \mathbf{Sh}(Y^\circ, G) \to \mathbf{Ab}$$

is defined to be the right derived functor of the G-invariant global section functor

$$\Gamma(Y^\circ, G, -) : \mathbf{Sh}(Y^\circ, G) \to \mathbf{Ab}, \quad \mathscr{F} \mapsto \Gamma(Y^\circ, \mathscr{F})^G,$$

where $\mathbf{Sh}(Y^\circ, G)$ is the category of G-equivariant sheaves on Y°, $\mathbf{Ab}$ is the category of abelian groups, and $\Gamma(Y^\circ, \mathscr{F})^G$ is the G-invariant part of the global section $\Gamma(Y^\circ, \mathscr{F})$.

4.1. ***Open covering.*** In this subsection we introduce a certain $\mathrm{GL}_g(\mathbb{Q})$-stable open covering of Y°. For $\alpha \in \mathbb{C}^g - \{0\}$, we define an open subset $V_\alpha \subset \mathbb{C}^g - \{0\}$ by

$$V_\alpha := \{y \in \mathbb{C}^g \mid \mathrm{Re}(\langle \alpha, y \rangle) > 0\} \subset \mathbb{C}^g - \{0\}.$$

Clearly, $V_\alpha \subset \mathbb{C}^g - \{0\}$ is a convex open subset. Let

$$X_{\mathbb{Q}} := \mathbb{Q}^g - \{0\}$$

denote the set of all nonzero rational vectors on which $\mathrm{GL}_g(\mathbb{Q})$ acts by the matrix multiplication from the left. Then we easily see that

$$Y^\circ = \bigcup_{\alpha \in X_{\mathbb{Q}}} V_\alpha.$$

Let $\mathscr{X}_{\mathbb{Q}} := \{V_\alpha\}_{\alpha \in X_{\mathbb{Q}}}$ denote this open covering of Y°. For $r \geq 0$ and $I = (\alpha_1, \dots, \alpha_r) \in (X_{\mathbb{Q}})^r$, set

$$V_I := \bigcap_{i=1}^{r} V_{\alpha_i} = \left\{ y \in Y^\circ \mid \mathrm{Re}(\langle \alpha_i, y \rangle) > 0 \text{ for all } i \right\} \subset Y^\circ.$$

In the case $r = 0$, we set $(X_{\mathbb{Q}})^0 = \{\varnothing\}$ and $V_\varnothing = Y^\circ$ by convention. Let

$$j_I : V_I \hookrightarrow Y^\circ$$

denote the inclusion map.

First, we show that $\mathscr{X}_{\mathbb{Q}} = \{V_\alpha\}_{\alpha \in X_{\mathbb{Q}}}$ is a $\mathrm{GL}_g(\mathbb{Q})$-stable open covering. Note that the group $\mathrm{GL}_g(\mathbb{Q})$ acts diagonally on $(X_{\mathbb{Q}})^r$. For $I = (\alpha_1, \dots, \alpha_r) \in (X_{\mathbb{Q}})^r$ and $\gamma \in \mathrm{GL}_g(\mathbb{Q})$, let

$$\gamma I = (\gamma\alpha_1, \dots, \gamma\alpha_r) \in (X_{\mathbb{Q}})^r$$

denote this diagonal action of γ on I.

Lemma 4.1.1. *For $r \geq 0$, $I = (\alpha_1, \dots, \alpha_r) \in (X_{\mathbb{Q}})^r$, and $\gamma \in \mathrm{GL}_g(\mathbb{Q})$, we have*

$$V_{\gamma I} = {}^t\gamma^{-1} V_I.$$

In other words, we have the following commutative diagram:

$$\begin{tikzcd} V_I \arrow[r, hook, "j_I"] \arrow[d, "{}^t\gamma^{-1}"', "\wr"] & Y^\circ \arrow[d, "{}^t\gamma^{-1}"] \\ V_{\gamma I} \arrow[r, hook, "j_{\gamma I}"] & Y^\circ \end{tikzcd}$$

Proof. For $y \in Y^\circ$, we have $y \in V_{\gamma I}$ if and only if

$$0 < \mathrm{Re}(\langle \gamma\alpha_i, y\rangle) = \mathrm{Re}(\langle \alpha_i, {}^t\gamma y\rangle)$$

for all $i \in \{1, \ldots, r\}$. This proves the lemma. □

4.2. ***The equivariant Čech complex.*** Let $\mathscr{F}$ be a $\mathrm{GL}_g(\mathbb{Q})$-equivariant sheaf on Y°. We consider the $\mathrm{GL}_g(\mathbb{Q})$-equivariant "sheaf Čech complex"

$$\mathscr{C}^\bullet(\mathcal{X}_\mathbb{Q}, \mathscr{F}) : \mathscr{C}^0(\mathcal{X}_\mathbb{Q}, \mathscr{F}) \xrightarrow{\mathrm{d}^0} \mathscr{C}^1(\mathcal{X}_\mathbb{Q}, \mathscr{F}) \xrightarrow{\mathrm{d}^1} \mathscr{C}^2(\mathcal{X}_\mathbb{Q}, \mathscr{F}) \xrightarrow{\mathrm{d}^2} \cdots$$

defined as follows. For $q \geq 0$, put

$$\mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F}) := \prod_{I \in (X_\mathbb{Q})^{q+1}} j_{I*} j_I^{-1} \mathscr{F},$$

where j_{I*} (resp. j_I^{-1}) is the direct image (resp. inverse image) functor induced by the inclusion map $j_I : V_I \hookrightarrow Y^\circ$. By Lemma 4.1.1, the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure

$$[\gamma] : \mathscr{F} \xrightarrow{\sim} ({}^t\gamma)_* \mathscr{F}$$

of $\mathscr{F}$ induces isomorphisms

$$[\gamma] : j_{I*} j_I^{-1} \mathscr{F} \xrightarrow{\sim} j_{I*} j_I^{-1} ({}^t\gamma)_* \mathscr{F} \simeq ({}^t\gamma)_* j_{\gamma I*} j_{\gamma I}^{-1} \mathscr{F} \quad \text{and} \quad [\gamma] : \mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F}) \xrightarrow{\sim} ({}^t\gamma)_* \mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F}).$$

We easily see that this defines a $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F})$. More concretely, for an open subset $U \subset Y^\circ$ and a section

$$s = (s_I)_{I \in (X_\mathbb{Q})^{q+1}} \in \Gamma(U, \mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F})) = \prod_{I \in (X_\mathbb{Q})^{q+1}} \Gamma(U \cap V_I, \mathscr{F}),$$

we have

$$([\gamma](s))_I = [\gamma](s_{\gamma^{-1} I}), \tag{4-1}$$

where $([\gamma](s))_I$ is the I-th component of $[\gamma](s)$.

The differential map

$$\mathrm{d}^q : \mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F}) \to \mathscr{C}^{q+1}(\mathcal{X}_\mathbb{Q}, \mathscr{F})$$

is given by

$$(\mathrm{d}^q(s))_{(\alpha_0, \ldots, \alpha_{q+1})} = \sum_{i=0}^{q+1} (-1)^i s_{(\alpha_0, \ldots, \check{\alpha}_i, \ldots, \alpha_{q+1})}|_{U \cap V_{(\alpha_0, \ldots, \alpha_{q+1})}}$$

for an open subset $U \subset Y^\circ$ and a section $s = (s_I)_{I \in (X_\mathbb{Q})^{q+1}} \in \Gamma(U, \mathscr{C}^q(\mathcal{X}_\mathbb{Q}, \mathscr{F}))$. Here $\check{\alpha}_i$ means that α_i is omitted. Moreover, there is a map

$$\mathrm{d}^{-1} : \mathscr{F} \to \mathscr{C}^0(\mathcal{X}_\mathbb{Q}, \mathscr{F}) = \prod_{\alpha \in X_\mathbb{Q}} j_{\alpha*} j_\alpha^{-1} \mathscr{F}$$

induced by the natural maps $\mathscr{F} \to j_{\alpha*} j_\alpha^{-1} \mathscr{F}$.

Then we have the following.

Lemma 4.2.1. (1) *For $q \geq -1$, the differential map* d^q *is a* $\mathrm{GL}_g(\mathbb{Q})$*-equivariant map, i.e.,* $[\gamma] \circ \mathrm{d}^q = \mathrm{d}^q \circ [\gamma]$ *for* $\gamma \in \mathrm{GL}_g(\mathbb{Q})$.

(2) *For any* $\alpha_0 \in X_{\mathbb{Q}}$*, the sequence*

$$0 \longrightarrow \mathscr{F}|_{V_{\alpha_0}} \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})|_{V_{\alpha_0}} \xrightarrow{\mathrm{d}^0} \mathscr{C}^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})|_{V_{\alpha_0}} \longrightarrow \cdots$$

is homotopic to zero. In particular, the sequence

$$0 \longrightarrow \mathscr{F} \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \xrightarrow{\mathrm{d}^0} \mathscr{C}^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \longrightarrow \cdots$$

is an exact sequence of $\mathrm{GL}_g(\mathbb{Q})$*-equivariant sheaves since* $Y^\circ = \bigcup_{\alpha_0 \in X_{\mathbb{Q}}} V_{\alpha_0}$.

Proof. (1) Let $U \subset Y^\circ$ be an open subset, and let

$$s = (s_I)_{I \in (X_{\mathbb{Q}})^{q+1}} \in \Gamma(U, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})) = \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \Gamma(U \cap V_I, \mathscr{F})$$

be a section. Let $J = (\alpha_0, \ldots, \alpha_{q+1}) \in (X_{\mathbb{Q}})^{q+2}$, and put $J^{(i)} := (\alpha_0, \ldots, \check{\alpha}_i, \ldots, \alpha_{q+1}) \in (X_{\mathbb{Q}})^{q+1}$ for $i = 0, \ldots, q+1$. Then we have

$$\begin{aligned}\big(\mathrm{d}^q([\gamma](s))\big)_J &= \sum_{i=0}^{q+1} (-1)^i [\gamma](s_{\gamma^{-1}J^{(i)}})|_{{}^t\gamma^{-1}U \cap V_J} = \sum_{i=0}^{q+1} (-1)^i [\gamma](s_{\gamma^{-1}J^{(i)}}|_{U \cap V_{\gamma^{-1}J}}) \\ &= [\gamma]\bigg(\sum_{i=0}^{q+1} (-1)^i s_{\gamma^{-1}J^{(i)}}|_{U \cap V_{\gamma^{-1}J}}\bigg) = \big([\gamma](\mathrm{d}^q(s))\big)_J.\end{aligned}$$

(2) See [Godement 1973, Théorème 5.2.1] or [Stacks 2005–, Lemma 02FU]. Although they prove only the exactness of the sequence, we can prove the statement in this lemma using essentially the same argument. See also [Kashiwara and Schapira 1990, Lemma 2.8.2, Remark 2.8.3]. □

By applying the additive functor

$$\underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, -) : \mathbf{Sh}(Y^\circ, \mathrm{GL}_g(\mathbb{Q})) \to \mathbf{Sh}(Y^\circ, \mathrm{GL}_g(\mathbb{Q})), \quad \mathscr{G} \mapsto \mathscr{G}^\Xi := \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{G}),$$

we obtain the following.

Corollary 4.2.2. *The sequence*

$$0 \longrightarrow \mathscr{F}^\Xi \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \xrightarrow{\mathrm{d}^0} \mathscr{C}^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \longrightarrow \cdots$$

is an exact sequence of $\mathrm{GL}_g(\mathbb{Q})$*-equivariant sheaves.*

Proof. Since the homotopy is preserved by the additive functor, by Lemma 4.2.1(2), we see that for any $\alpha_0 \in X_{\mathbb{Q}}$, the sequence

$$0 \longrightarrow \mathscr{F}^\Xi|_{V_{\alpha_0}} \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi|_{V_{\alpha_0}} \xrightarrow{\mathrm{d}^0} \mathscr{C}^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi|_{V_{\alpha_0}} \longrightarrow \cdots$$

is homotopic to zero, and hence exact. □

Now, by taking the global section, set

$$C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) := \Gamma(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})) = \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \Gamma(V_I, \mathscr{F}).$$

Then we obtain a complex

$$C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) : C^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \xrightarrow{\mathrm{d}^0} C^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \xrightarrow{\mathrm{d}^1} C^2(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \xrightarrow{\mathrm{d}^2} \cdots$$

of $\mathrm{GL}_g(\mathbb{Q})$-modules. Note that this is the usual Čech complex associated to the open covering $\mathcal{X}_{\mathbb{Q}}$. Furthermore, set

$$C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi := \Gamma(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi) = \mathrm{Map}(\Xi, C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})).$$

Then we obtain another complex

$$C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi : C^0(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \xrightarrow{\mathrm{d}^0} C^1(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \xrightarrow{\mathrm{d}^1} C^2(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \xrightarrow{\mathrm{d}^2} \cdots$$

of $\mathrm{GL}_g(\mathbb{Q})$-modules. For $Q \in \Xi$, the evaluation map

$$\mathrm{ev}_Q : C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F})^\Xi \to C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}) \tag{4-2}$$

is a Γ_Q-equivariant morphism of complexes.

4.3. ***Acyclicity.*** Our aim here is to prove the acyclicity of the sheaves $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ and $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi$; see Proposition 4.3.3. Then we can compute the equivariant cohomology groups $H^q(Y^\circ, \Gamma_Q, \mathscr{F}_d)$ and $H^q(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^\Xi)$ using the Čech complexes $C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ and $C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi$; see Corollary 4.3.4.

Lemma 4.3.1. *Let $r \geq 1$ and $I = (\alpha_1, \ldots, \alpha_r) \in (X_{\mathbb{Q}})^r$.*

(1) *For all $q \geq 1$, we have*

$$H^q(V_I, \mathscr{F}_d) = 0.$$

(2) *For all $q \geq 1$, we have*

$$R^q j_{I*}(j_I^{-1}\mathscr{F}_d) = 0,$$

where $R^q j_{I*}$ (resp. j_I^{-1}) is the higher direct image (resp. inverse image) functor induced by the inclusion map $j_I : V_I \hookrightarrow Y^\circ$.

(3) *For any open subset $U \subset Y^\circ$ and $q \geq 0$, we have an isomorphism*

$$H^q(U, j_{I*}j_I^{-1}\mathscr{F}_d) \xrightarrow{\sim} H^q(U \cap V_I, \mathscr{F}_d).$$

Proof. (1) This follows directly from Proposition 3.2.3(3) since V_I is convex.

(2) Let $x \in Y^\circ$. Since convex open subsets form a basis of open subsets of Y°, we have

$$(R^q j_{I*}(j_I^{-1}\mathscr{F}_d))_x = \varinjlim_{x \in U \text{ convex}} H^q(U \cap V_I, j_I^{-1}\mathscr{F}_d) = \varinjlim_{x \in U \text{ convex}} H^q(U \cap V_I, \mathscr{F}_d) = 0.$$

Here the last vanishing follows from Proposition 3.2.3(3). This proves (2).

(3) This follows from (2) and the Leray spectral sequence. $\square$

Proposition 4.3.2. *For $q \geq 0$, the sheaves $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ and $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}$ are $\Gamma(Y^\circ, -)$-acyclic, i.e.,*

$$H^p(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) = 0 \quad \textit{and} \quad H^p(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}) = 0 \qquad \textit{for } p \geq 1.$$

Proof. We imitate the argument in [Bannai et al. 2023, Proposition 3.4, Lemma 3.5]. For $I \in (X_{\mathbb{Q}})^{q+1}$, put $\mathscr{F}_I := j_{I*} j_I^{-1} \mathscr{F}_d$, and let

$$0 \to \mathscr{F}_I \to \mathscr{I}_I^\bullet$$

be an injective resolution of $\mathscr{F}_I$. First we show that

$$0 \to \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d) = \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{F}_I \to \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet, \tag{4-3}$$

$$0 \to \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi} = \Bigl(\prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{F}_I\Bigr)^{\Xi} \to \Bigl(\prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet\Bigr)^{\Xi} \tag{4-4}$$

are both injective resolutions of $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ and $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}$ respectively. It is clear that

$$\prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^p \quad \text{and} \quad \Bigl(\prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^p\Bigr)^{\Xi} \simeq \prod_{Q \in \Xi} \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^p$$

are injective sheaves because they are products of injective sheaves; see Remark 3.1.1(2). We must show the exactness of (4-3) and (4-4). Let $U \subset Y^\circ$ be any convex open subset. By Lemma 4.3.1(3) and Proposition 3.2.3(3), we have

$$H^p(U, \mathscr{F}_I) \xrightarrow{\sim} H^p(U \cap V_I, \mathscr{F}_d) = 0$$

for $p \geq 1$. Therefore, we find that

$$0 \to \mathscr{F}_I(U) \to \mathscr{I}_I^\bullet(U)$$

is exact because $H^p(U, \mathscr{F}_I)$ is the cohomology of this complex. Hence,

$$0 \to \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{F}_I(U) \to \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet(U) \quad \text{and} \quad 0 \to \prod_{Q \in \Xi} \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{F}_I(U) \to \prod_{Q \in \Xi} \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet(U)$$

are also exact. Since convex open subsets of Y° form a basis of open subsets, we obtain the exactness of (4-3) and (4-4).

Then for $p \geq 1$, we have

$$\begin{aligned} H^p(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) &\simeq H^p\Bigl(\Gamma\Bigl(Y^\circ, \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet\Bigr)\Bigr) \\ &\simeq \prod_{I \in (X_{\mathbb{Q}})^{q+1}} H^p(\Gamma(Y^\circ, \mathscr{I}_I^\bullet)) \\ &\simeq \prod_{I \in (X_{\mathbb{Q}})^{q+1}} H^p(Y^\circ, \mathscr{F}_I) \simeq \prod_{I \in (X_{\mathbb{Q}})^{q+1}} H^p(V_I, \mathscr{F}_d) = 0, \end{aligned}$$

and similarly,

$$\begin{aligned} H^p(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi) &\simeq H^p\Big(\Gamma\Big(Y^\circ, \Big(\prod_{I\in(X_{\mathbb{Q}})^{q+1}} \mathscr{I}_I^\bullet\Big)^\Xi\Big)\Big) \\ &\simeq \prod_{Q\in\Xi} \prod_{I\in(X_{\mathbb{Q}})^{q+1}} H^p(\Gamma(Y^\circ, \mathscr{I}_I^\bullet)) \simeq \prod_{Q\in\Xi} \prod_{I\in(X_{\mathbb{Q}})^{q+1}} H^p(V_I, \mathscr{F}_d) = 0. \qquad \square \end{aligned}$$

Proposition 4.3.3. (1) *Let $Q \in \Xi$. For $q \geq 0$, the sheaf $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ is $\Gamma(Y^\circ, \Gamma_Q, -)$-acyclic, i.e.,*

$$H^p(Y^\circ, \Gamma_Q, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) = 0$$

for $p \geq 1$. In particular, the complex

$$0 \to \mathscr{F}_d \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$$

gives a $\Gamma(Y^\circ, \Gamma_Q, -)$-acyclic resolution of $\mathscr{F}_d$.

(2) *For $q \geq 0$, the sheaf $\mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi$ is $\Gamma(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), -)$-acyclic, i.e., we have*

$$H^p(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi) = 0$$

for $p \geq 1$. In particular, the complex

$$0 \to \mathscr{F}_d^\Xi \xrightarrow{\mathrm{d}^{-1}} \mathscr{C}^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^\Xi$$

gives a $\Gamma(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), -)$-acyclic resolution of $\mathscr{F}_d^\Xi$.

Proof. (1) First note that the functor $\Gamma(Y^\circ, \Gamma_Q, -)$ is a composition of two left exact functors $\Gamma(Y^\circ, -)$ and $(-)^{\Gamma_Q}$. Moreover, $\Gamma(Y^\circ, -)$ sends injective objects to injective objects. Therefore, we have a spectral sequence

$$E_2^{ab} = H^a\big(\Gamma_Q, H^b(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d))\big) \Rightarrow H^{a+b}(Y^\circ, \Gamma_Q, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)),$$

where $H^a(\Gamma_Q, -)$ is the usual group cohomology of Γ_Q. Now, by Proposition 4.3.2, we already have

$$H^b(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) = 0 \quad \text{for all } b \geq 1.$$

Therefore, it suffices to show

$$H^a\big(\Gamma_Q, \Gamma(Y^\circ, \mathscr{C}^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d))\big) = H^a(\Gamma_Q, C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) = 0 \quad \text{for all } a \geq 1.$$

Actually, we will prove that $C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ is a coinduced Γ_Q-module. First, recall that

$$C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d) = \prod_{I\in(X_{\mathbb{Q}})^{q+1}} \Gamma(V_I, \mathscr{F}_d),$$

and that Γ_Q acts freely on $(X_{\mathbb{Q}})^{q+1}$ by Lemma 2.1.1(7). Let $A \subset (X_{\mathbb{Q}})^{q+1}$ be a system of representatives of $\Gamma_Q \backslash (X_{\mathbb{Q}})^{q+1}$, and set

$$M := \prod_{I\in A} \Gamma(V_I, \mathscr{F}_d).$$

Then recall that the $\mathrm{GL}_g(\mathbb{Q})$-equivariant structure on $\mathscr{F}_d$ gives an isomorphism

$$[\gamma] : \Gamma(V_I, \mathscr{F}_d) \xrightarrow{\sim} \Gamma(({}^t\gamma)^{-1} V_I, \mathscr{F}_d) = \Gamma(V_{\gamma I}, \mathscr{F}_d) \tag{4-5}$$

for each $I \in (X_{\mathbb{Q}})^{q+1}$ and $\gamma \in \mathrm{GL}_g(\mathbb{Q})$; see Lemma 4.1.1. Therefore, for each $\gamma \in \Gamma_Q$, we have an isomorphism

$$M = \prod_{I \in A} \Gamma(V_I, \mathscr{F}_d) \xrightarrow{\sim} \prod_{I \in A} \Gamma(V_{\gamma I}, \mathscr{F}_d), \quad (s_I)_{I \in A} \mapsto ([\gamma](s_I))_{I \in A},$$

and hence we obtain an isomorphism

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[\Gamma_Q], M) = \prod_{\gamma \in \Gamma_Q} M \xrightarrow{\sim} \prod_{\gamma \in \Gamma_Q} \prod_{I \in A} \Gamma(V_{\gamma I}, \mathscr{F}_d) = C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d).$$

Since this is clearly a Γ_Q-equivariant isomorphism, we see $C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)$ is a coinduced Γ_Q-module.

(2) This can be proved similarly. First, by the spectral sequence

$$E_2^{ab} = H^a\big(\mathrm{SL}_g(\mathbb{Z}), H^b(Y^\circ, \mathscr{C}^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi})\big) \Rightarrow H^{a+b}\big(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}\big)$$

and Proposition 4.3.2, it suffices to show

$$H^a\big(\mathrm{SL}_g(\mathbb{Z}), C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}\big) = 0 \quad \text{for all } a \geq 1.$$

Again, we will prove that

$$C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi} \simeq \prod_{Q \in \Xi} \prod_{I \in (X_{\mathbb{Q}})^{q+1}} \Gamma(V_I, \mathscr{F}_d)$$

is a coinduced $\mathrm{SL}_g(\mathbb{Z})$-module. Note that the action of $\mathrm{SL}_g(\mathbb{Z})$ on $\Xi \times (X_{\mathbb{Q}})^{q+1}$ is free. Indeed, if

$$\gamma(Q, I) = ([\gamma](Q), \gamma I) = (Q, I),$$

then it follows that $\gamma \in \Gamma_Q$, and hence $\gamma = 1$, since the action of Γ_Q on $(X_{\mathbb{Q}})^{q+1}$ is free. Let $A' \subset \Xi \times (X_{\mathbb{Q}})^{q+1}$ be a system of representatives of $\mathrm{SL}_g(\mathbb{Z}) \backslash (\Xi \times (X_{\mathbb{Q}})^{q+1})$, and set

$$M' := \prod_{(Q,I) \in A'} \Gamma(V_I, \mathscr{F}_d).$$

Then again by using (4-5), we obtain an isomorphism

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[\mathrm{SL}_g(\mathbb{Z})], M') = \prod_{\gamma \in \mathrm{SL}_g(\mathbb{Z})} M' \xrightarrow{\sim} \prod_{\gamma \in \mathrm{SL}_g(\mathbb{Z})} \prod_{(Q,I) \in A'} \Gamma(V_{\gamma I}, \mathscr{F}_d) \simeq C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}$$

of $\mathrm{SL}_g(\mathbb{Z})$-modules. Thus we find that $C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}$ is a coinduced $\mathrm{SL}_g(\mathbb{Z})$-module. □

Corollary 4.3.4. (1) *Let $Q \in \Xi$. For $q \geq 0$, we have*

$$H^q(Y^\circ, \Gamma_Q, \mathscr{F}_d) \simeq H^q\big(\Gamma(Y^\circ, \Gamma_Q, \mathscr{C}^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d))\big) = H^q(C^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Gamma_Q}),$$

where the second and third H^q are the cohomology of complexes.

(2) *For $q \geq 0$, we have*

$$H^q(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^{\Xi}) \simeq H^q\big(\Gamma(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi})\big) = H^q\big(\mathrm{Map}_{\mathrm{SL}_g(\mathbb{Z})}(\Xi, C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d))\big),$$

where $\mathrm{Map}_{\mathrm{SL}_g(\mathbb{Z})}(-,-)$ *is the set of* $\mathrm{SL}_g(\mathbb{Z})$*-equivariant maps.*

(3) *For $Q \in \Xi$, we have the commutative diagram*

$$\begin{array}{ccc} H^q(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^{\Xi}) & \xrightarrow{\mathrm{ev}_Q} & H^q(Y^\circ, \Gamma_Q, \mathscr{F}_d) \\ \Big\downarrow{\scriptstyle\wr} & & \Big\downarrow{\scriptstyle\wr} \\ H^q\big(\mathrm{Map}_{\mathrm{SL}_g(\mathbb{Z})}(\Xi, C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d))\big) & \xrightarrow{\mathrm{ev}_Q} & H^q(C^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Gamma_Q}) \end{array}$$

where the two ev_Q *are the evaluation maps induced by (3-1) and (4-2).*

We end this section with one more corollary, concerning an operation which shifts the index $d \geq 0$ of $\mathscr{F}_d$.

Corollary 4.3.5. *Let $P(y_1, \dots, y_g) \in \mathbb{C}[y_1, \dots, y_g]$ be a homogeneous polynomial of degree $d' \leq d$ such that*

$$P({}^t\gamma y) = P(y) \quad \text{for all } \gamma \in \Gamma_Q.$$

Then the multiplication by P,

$$P : C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d) \to C^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_{d-d'}), \quad (s_I(y))_{I \in (X_{\mathbb{Q}})^{q+1}} \mapsto (P(y)s_I(y))_{I \in (X_{\mathbb{Q}})^{q+1}},$$

gives a Γ_Q-equivariant map of complexes, and hence induces a map

$$P : H^q(Y^\circ, \Gamma_Q, \mathscr{F}_d) \to H^q(Y^\circ, \Gamma_Q, \mathscr{F}_{d-d'}).$$

Example 4.3.6. A typical example of such a Γ_Q-invariant homogeneous polynomial P is the norm polynomial N_{w^*} defined in Section 2.2; see Lemma 2.2.1. More generally, let $k \geq 1$ be an integer. Under the notation in Lemma 2.2.1, the k-th power $N_{w^*}^k$ of the norm polynomial N_{w^*} is a Γ_Q-invariant homogeneous polynomial of degree kg. In particular, we have a map

$$N_{w^*}^k : H^q(Y^\circ, \Gamma_Q, \mathscr{F}_{kg}) \to H^q(Y^\circ, \Gamma_Q, \mathscr{F}_0).$$

5. Cones and the exponential perturbation

In this section we introduce the notion of exponential perturbation, which is a modification of the so-called upper closure or Q-perturbation (Colmez perturbation) used in [Yamamoto 2010; Bannai et al. 2023; Charollois et al. 2015]. This is one of the key ingredients enabling us to deal with general number fields.

For $r \geq 0$, $I = (\alpha_1, \dots, \alpha_r) \in (\mathbb{R}^g - \{0\})^r$, let

$$C_I := \sum_{i=1}^r \mathbb{R}_{>0}\alpha_i \subset \mathbb{R}^g$$

denote the *open cone generated by* $\alpha_1, \dots, \alpha_r$. In the case $r = 0$ and $I = \varnothing$, we set $C_\varnothing := \{0\}$.

Remark 5.0.1. We follow the convention to call C_I an "open" cone although it is not necessarily an open subset of $\mathbb{R}^g$. Note that, however, C_I is open in $\mathrm{Span}_{\mathbb{R}}\{\alpha_1, \dots, \alpha_r\}$, where $\mathrm{Span}_{\mathbb{R}}\{\alpha_1, \dots, \alpha_r\} \subset \mathbb{R}^g$ is the $\mathbb{R}$-subspace spanned by $\alpha_1, \dots, \alpha_r$; see Lemma 5.2.4.

Recall that $X_{\mathbb{Q}} := \mathbb{Q}^g - \{0\}$ denotes the set of nonzero vectors of $\mathbb{Q}^g$. In this paper we fix the terminology concerning cones as follows.

Definition 5.0.2. (1) An open cone C_I is called *rational* if we can take $I \in (X_{\mathbb{Q}})^r$.

(2) An open cone C_I is called *simplicial* if $\alpha_1, \dots, \alpha_r$ are linearly independent over $\mathbb{R}$.

(3) We refer to a subset of $\mathbb{R}^g$ which can be written as a disjoint union of a finite number of rational simplicial open cones as a *rational constructible cone*.

5.1. ***The exponential perturbation.*** Recall that

$$\Xi = \{Q \in \mathrm{GL}_g(\mathbb{Q}) \mid Q \text{ is irreducible over } \mathbb{Q}\}$$

denotes the set of irreducible matrices of $\mathrm{GL}_g(\mathbb{Q})$; see Section 2.1.

Definition 5.1.1. For $Q \in \Xi$ and a subset $A \subset \mathbb{R}^g$, we define the *exponential Q-perturbation* A^Q of A as

$$A^Q := \{x \in \mathbb{R}^g \mid \text{there exists } \delta > 0 \text{ such that for all } \varepsilon \in (0, \delta),\ \exp(\varepsilon Q)x \in A\},$$

where $\exp(\varepsilon Q) \in \mathrm{GL}_g(\mathbb{R})$ is the matrix exponential of $\varepsilon Q \in \mathrm{GL}_g(\mathbb{R})$.

Remark 5.1.2. This exponential Q-perturbation is defined by considering the perturbation of $x \in \mathbb{R}^g$ by the matrix action of $\exp(\varepsilon Q)$, and we call this process the *exponential perturbation*. The original Q-perturbation used in [Charollois et al. 2015] is the perturbation of x by the vectors $Q \in \mathbb{R}^g$ whose components are linearly independent over $\mathbb{Q}$.

Lemma 5.1.3. *Let $Q \in \Xi$.*

(1) *Let $A, B \subset \mathbb{R}^g$ be subsets such that $A \subset B$. Then we have*

$$A^Q \subset B^Q.$$

(2) *Let $A_1, \dots, A_m \subset \mathbb{R}^g$ be subsets. Then we have*

$$(A_1 \cap \cdots \cap A_m)^Q = A_1^Q \cap \cdots \cap A_m^Q.$$

In particular, if $A_1 \cap \cdots \cap A_m = \varnothing$, then $A_1^Q \cap \cdots \cap A_m^Q = \varnothing$.

Proof. (1) is obvious. We prove (2). The inclusion $\subset$ is clear. We prove $\supset$. Let $x \in A_1^Q \cap \cdots \cap A_m^Q$. Then, by definition, there exist $\delta_1, \dots, \delta_m > 0$ such that

$$\exp((0, \delta_i)Q)x \subset A_i$$

for $i = 1, \dots, m$. Put $\delta := \min\{\delta_1, \dots, \delta_m\} > 0$. Then we have

$$\exp((0, \delta)Q)x \subset A_1 \cap \cdots \cap A_m,$$

and hence $x \in (A_1 \cap \cdots \cap A_m)^Q$. □

In the following, we study the exponential Q-perturbation C_I^Q of rational open cones C_I, which play an important role in the construction of our Shintani cocycle.

Lemma 5.1.4. *For $r \geq 0$, $I = (\alpha_1, \ldots, \alpha_r) \in (X_{\mathbb{Q}})^r$, $Q \in \Xi$, and $\gamma \in \mathrm{GL}_g(\mathbb{Q})$, we have*

$$\gamma(C^Q_{\gamma^{-1}I}) = C_I^{[\gamma](Q)},$$

where $[\gamma](Q) = \gamma Q \gamma^{-1} \in \Xi$.

Proof. Indeed, for $x \in \mathbb{R}^g$ and $\varepsilon > 0$, we see that

$$\exp(\varepsilon[\gamma](Q))x \in C_I \iff \exp(\varepsilon\gamma Q\gamma^{-1})x \in C_I \iff \exp(\varepsilon Q)\gamma^{-1}x \in \gamma^{-1}(C_I) = C_{\gamma^{-1}I}.$$

This proves the lemma. □

5.2. *Rationality.* The aim of this subsection is to prove the following proposition:

Proposition 5.2.1. *Let $0 \leq r \leq g$, $I = (\alpha_1, \ldots, \alpha_r) \in (X_{\mathbb{Q}})^r$, and $Q \in \Xi$.*

(1) *Suppose* $\dim_{\mathbb{Q}} \mathrm{Span}_{\mathbb{Q}}\{\alpha_1, \ldots, \alpha_r\} \leq g - 1$. *Then*

$$C_I^Q = \begin{cases} \{0\} & \text{if } 0 \in C_I, \\ \varnothing & \text{if } 0 \notin C_I. \end{cases}$$

(2) *The exponential Q-perturbation C_I^Q of the rational open cone C_I generated by I is a rational constructible cone, i.e., a disjoint union of a finite number of rational simplicial open cones.*

To prove this proposition, we first prepare several lemmas. In the following, for $\alpha \in \mathbb{R}^g - \{0\}$, we put

$$U_{\alpha,\pm} := \{x \in \mathbb{R}^g \mid \pm\langle x, \alpha\rangle > 0\} \quad \text{and} \quad H_\alpha := \{x \in \mathbb{R}^g \mid \langle x, \alpha\rangle = 0\}.$$

We start with recalling the following fact.

Lemma 5.2.2 [Shintani 1976, Section 1.2; Hida 1993, pp. 68–69, Lemma 1]. (1) *Let $W \subset \mathbb{Q}^g$ be a $\mathbb{Q}$-subspace, and let $l_1, \ldots, l_m \in \mathbb{Q}^g - \{0\}$. Then the subset*

$$X = \{x \in W \otimes_{\mathbb{Q}} \mathbb{R} \subset \mathbb{R}^g \mid \langle x, l_i\rangle > 0 \text{ for } i = 1, \ldots, m\} \subset \mathbb{R}^g$$

is a rational constructible cone.

(2) *Let $C, C' \subset \mathbb{R}^g$ be rational constructible cones. Then $C \cup C'$, $C \cap C'$, and $C - C'$ are rational constructible cones.*

Proof. See [Shintani 1976, Lemma 2, Corollary to Lemma 2] and [Hida 1993, pp. 68–69, Lemma 1]. Although, in [Hida 1993], it is assumed that the total space is of the form $F \otimes_{\mathbb{Q}} \mathbb{R}$ for a number field F and that W is a subspace generated by elements in F, the proof there does not use this special assumption. □

The following is the key lemma of this section.

Lemma 5.2.3. *Let $Q \in \Xi$ and $\alpha \in \mathbb{Q}^g - \{0\}$. For $k \geq 0$, put*

$$H_{\pm}^{(k)} := \left\{x \in \mathbb{R}^g \mid \pm\langle x, {}^tQ^k\alpha\rangle > 0 \text{ and } \langle x, {}^tQ^j\alpha\rangle = 0 \text{ for } 0 \leq j \leq k-1\right\}.$$

Note that $H_{\pm}^{(0)} = U_{\alpha,\pm}$ by definition.

(1) *There exists $k_0 \geq 0$ such that $H_{\pm}^{(k)} = \varnothing$ for all $k \geq k_0 + 1$. Moreover, we have*

$$\mathbb{R}^g - \{0\} = \bigsqcup_{k=0}^{k_0} (H_+^{(k)} \sqcup H_-^{(k)}),$$

where $\bigsqcup$ and $\sqcup$ denote the disjoint union.

(2) *For all $k \geq 0$, the sets $H_+^{(k)}$ and $H_-^{(k)}$ are rational constructible cones.*

(3) *For all $k \geq 0$, we have $H_+^{(k)} \subset (H_+^{(0)})^Q = (U_{\alpha,+})^Q$ and $H_-^{(k)} \subset (H_-^{(0)})^Q = (U_{\alpha,-})^Q$.*

(4) *We have $H_\alpha^Q = \{0\}$ and*

$$\mathbb{R}^g - \{0\} = (U_{\alpha,+})^Q \sqcup (U_{\alpha,-})^Q.$$

In particular, $\mathbb{R}^g = H_\alpha^Q \sqcup (U_{\alpha,+})^Q \sqcup (U_{\alpha,-})^Q$.

(5) *We have*

$$(U_{\alpha,+})^Q = \bigsqcup_{k=0}^{k_0} H_+^{(k)} \quad \text{and} \quad (U_{\alpha,-})^Q = \bigsqcup_{k=0}^{k_0} H_-^{(k)}.$$

In particular, $(U_{\alpha,+})^Q$ and $(U_{\alpha,-})^Q$ are rational constructible cones.

Proof. (1) and (2) For $k \geq 0$, put

$$H^{(k)} := \{x \in \mathbb{R}^g \mid \langle x, {}^tQ^j\alpha\rangle = 0 \text{ for } 0 \leq j \leq k-1\}.$$

Then we have a descending chain

$$\mathbb{R}^g = H^{(0)} \supset H^{(1)} \supset H^{(2)} \supset \cdots$$

of $\mathbb{R}$-vector spaces. Note that the subspaces $H^{(k)}$ are all defined over $\mathbb{Q}$ since we have ${}^tQ^j\alpha \in \mathbb{Q}^g - \{0\}$ for $j \geq 0$. Since $\mathbb{R}^g$ is a finite-dimensional vector space, there exists $k_0 \geq 0$ such that $H^{(k)} = H^{(k_0+1)}$ for all $k \geq k_0 + 1$.

Claim. $$H^{(k_0+1)} = 0.$$

Proof of claim. Indeed, let $x \in H^{(k_0+1)} = H^{(k_0+2)}$. Then we have

$$\langle Qx, {}^tQ^j\alpha\rangle = \langle x, {}^tQ^{j+1}\alpha\rangle = 0 \quad \text{for } 0 \leq j \leq k_0,$$

and hence $Qx \in H^{(k_0+1)}$. Therefore, $H^{(k_0+1)}$ is a Q-stable subspace of $\mathbb{R}^g$ defined over $\mathbb{Q}$. Moreover, since $\alpha \neq 0$, we have

$$H^{(k_0+1)} \subset H^{(1)} \subsetneq \mathbb{R}^g.$$

Therefore, we obtain $H^{(k_0+1)} = 0$ by Lemma 2.1.1(2). □

Now (1) follows from the fact

$$H^{(k)}-H^{(k+1)}=H_+^{(k)}\sqcup H_-^{(k)}\quad \text{for all } k\geq 0,$$

and (2) follows from Lemma 5.2.2(1).

(3) Let $x\in H_+^{(k)}$. Then we have

$$\langle \exp(\varepsilon Q)x,\alpha\rangle=\sum_{m\geq k}\frac{\langle x,{}^tQ^m\alpha\rangle}{m!}\varepsilon^m.$$

Now since $\langle x,{}^tQ^k\alpha\rangle>0$, there exists $\delta>0$ such that

$$\langle \exp(\varepsilon Q)x,\alpha\rangle=\sum_{m\geq k}\frac{\langle x,{}^tQ^m\alpha\rangle}{m!}\varepsilon^m>0$$

for all $\varepsilon\in(0,\delta)$. Hence $x\in(H_+^{(0)})^Q$. The inclusion $H_-^{(k)}\subset(H_-^{(0)})^Q$ can be proved similarly.

(4) First, by Lemma 5.1.3(2), we see $(U_{\alpha,+})^Q\cap(U_{\alpha,-})^Q=\varnothing$, and $H_\alpha^Q\cap(U_{\alpha,\pm})^Q=\varnothing$. On the other hand, we obviously have $0\in H_\alpha^Q$, and hence $0\notin(U_{\alpha,\pm})^Q$. Therefore, by (1) and (3), we obtain

$$\mathbb{R}^g-\{0\}\subset(U_{\alpha,+})^Q\sqcup(U_{\alpha,-})^Q\subset\mathbb{R}^g-\{0\}.$$

Thus we find $\mathbb{R}^g-\{0\}=(U_{\alpha,+})^Q\sqcup(U_{\alpha,-})^Q$ and $H_\alpha^Q=\{0\}$.

(5) The first part follows from (1), (3), and (4). Then the latter part follows from (2). □

Lemma 5.2.4. *Let $I=(\alpha_1,\dots,\alpha_r)\in(X_{\mathbb{Q}})^r$ such that $\alpha_1,\dots,\alpha_r\in\mathbb{Q}^g-\{0\}$ are linearly independent. Note that we automatically have $r\leq g$.*

(1) *There exist $\alpha'_1,\dots,\alpha'_r,\beta'_1,\dots,\beta'_{g-r}\in\mathbb{Q}^g-\{0\}$ such that*

$$C_I=\left(\bigcap_{i=1}^{r}U_{\alpha'_i,+}\right)\cap\left(\bigcap_{i=1}^{g-r}H_{\beta'_i}\right).$$

(2) *Let $Q\in\Xi$. Then we have*

$$\mathbb{R}^g=C_I^Q\sqcup(\mathbb{R}^g-C_I)^Q.$$

Proof. (1) Put $W:=\mathrm{Span}_{\mathbb{Q}}\{\alpha_1,\dots,\alpha_r\}\subset\mathbb{Q}^g$, and let $W^\perp\subset\mathbb{Q}^g$ be its orthogonal complement with respect to the scalar product $\langle\ ,\ \rangle$. Let $\alpha'_1,\dots,\alpha'_r\in W$ be the dual basis of $\alpha_1,\dots,\alpha_r$ in W with respect to $\langle\ ,\ \rangle$, i.e.,

$$\langle\alpha_i,\alpha'_j\rangle=\begin{cases}1 & (i=j),\\ 0 & (i\neq j),\end{cases}$$

and let $\beta'_1,\dots,\beta'_{g-r}\in W^\perp$ be a basis of $W^\perp$ over $\mathbb{Q}$. Then $\alpha'_1,\dots,\alpha'_r,\beta'_1,\dots,\beta'_{g-r}$ satisfy the desired property. Indeed, let $\beta_1,\dots,\beta_{g-r}\in W^\perp$ be the dual basis of $\beta'_1,\dots,\beta'_{g-r}$ in $W^\perp$, and let $x\in\mathbb{R}^g$. Since $\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_{g-r}$ form a basis of $\mathbb{R}^g$, we have

$$x=\sum_{i=1}^{r}c_i\alpha_i+\sum_{j=1}^{g-r}d_j\beta_j$$

for some $c_i, d_j \in \mathbb{R}$. Then we have $x \in C_I$ if and only if

$$\langle x, \alpha'_i \rangle = c_i > 0 \quad \text{and} \quad \langle x, \beta'_j \rangle = d_j = 0 \qquad \text{for all } i, j.$$

This proves (1).

(2) Using (1), we take $\alpha'_1, \dots, \alpha'_r, \beta'_1, \dots, \beta'_{g-r} \in \mathbb{Q}^g - \{0\}$ such that

$$C_I = \Bigl(\bigcap_{i=1}^{r} U_{\alpha'_i,+}\Bigr) \cap \Bigl(\bigcap_{i=1}^{g-r} H_{\beta'_i}\Bigr). \tag{5-1}$$

We then have

$$\mathbb{R}^g - C_I = \bigcup_{i=1}^{r} (U_{\alpha'_i,-} \cup H_{\alpha'_i}) \cup \bigcup_{i=1}^{g-r} (U_{\beta'_i,+} \cup U_{\beta'_i,-}).$$

By taking the exponential Q-perturbation and using Lemma 5.1.3(1), we obtain

$$\bigcup_{i=1}^{r} \bigl((U_{\alpha'_i,-})^Q \cup H^Q_{\alpha'_i}\bigr) \cup \bigcup_{i=1}^{g-r} \bigl((U_{\beta'_i,+})^Q \cup (U_{\beta'_i,-})^Q\bigr) \subset (\mathbb{R}^g - C_I)^Q. \tag{5-2}$$

On the other hand, by (5-1) and Lemmas 5.1.3(2) and 5.2.3(4), we obtain

$$\begin{aligned} \mathbb{R}^g - C^Q_I &= \mathbb{R}^g - \Bigl(\Bigl(\bigcap_{i=1}^{r} (U_{\alpha'_i,+})^Q\Bigr) \cap \Bigl(\bigcap_{i=1}^{g-r} H^Q_{\beta'_i}\Bigr)\Bigr) \\ &= \bigcup_{i=1}^{r} \bigl((U_{\alpha'_i,-})^Q \cup H^Q_{\alpha'_i}\bigr) \cup \bigcup_{i=1}^{g-r} \bigl((U_{\beta'_i,+})^Q \cup (U_{\beta'_i,-})^Q\bigr). \end{aligned} \tag{5-3}$$

Thus, by combining (5-2) and (5-3), we find that $\mathbb{R}^g - C^Q_I \subset (\mathbb{R}^g - C_I)^Q$, and hence $\mathbb{R}^g = C^Q_I \cup (\mathbb{R}^g - C_I)^Q$. Finally, since we have $C^Q_I \cap (\mathbb{R}^g - C_I)^Q = \varnothing$ by Lemma 5.1.3(2), we obtain $\mathbb{R}^g = C^Q_I \sqcup (\mathbb{R}^g - C_I)^Q$. □

Proof of Proposition 5.2.1. (1) Since $\mathrm{Span}_{\mathbb{Q}}\{\alpha_1, \dots, \alpha_r\} \subsetneq \mathbb{Q}^g$, there exists $\beta \in \mathbb{Q}^g - \{0\}$ such that

$$C_I \subset \mathrm{Span}_{\mathbb{R}}\{\alpha_1, \dots, \alpha_r\} \subset H_\beta.$$

Therefore, by Lemmas 5.1.3(1) and 5.2.3(4), we have either $C^Q_I = \varnothing$ or $C^Q_I = \{0\}$. Then it is clear that $C^Q_I = \{0\}$ if and only if $0 \in C_I$. This proves (1).

(2) Since $\varnothing$ and $\{0\}$ are obviously rational constructible cones, we may assume $\alpha_1, \dots, \alpha_r$ generates $\mathbb{R}^g$. In particular, we have $r = g$ and C_I is a rational simplicial open cone. By Lemma 5.2.4(1), there exist $\alpha'_1, \dots, \alpha'_g \in \mathbb{Q}^g - \{0\}$ such that

$$C_I = \bigcap_{i=1}^{g} U_{\alpha'_i,+}.$$

Then, by Lemma 5.1.3(2), we have

$$C^Q_I = \bigcap_{i=1}^{g} (U_{\alpha'_i,+})^Q.$$

Now, we already know that $(U_{\alpha'_i,+})^Q$ is a rational constructible cone by Lemma 5.2.3(5), and hence C_I^Q is also a rational constructible cone by Lemma 5.2.2(2). □

5.3. *Cocycle relation.*

Definition 5.3.1. (1) For a subset $A \subset \mathbb{R}^g$, let

$$\mathbf{1}_A : \mathbb{R}^g \to \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{if } x \notin A, \\ 1 & \text{if } x \in A \end{cases}$$

denote the characteristic function of A.

(2) For $I = (\alpha_1, \dots, \alpha_g) \in (\mathbb{R}^g - \{0\})^g$, we set

$$\operatorname{sgn}(I) := \operatorname{sgn}\det(\alpha_1, \dots, \alpha_g) \in \{-1, 0, 1\},$$

where $(\alpha_1, \dots, \alpha_g)$ is regarded as an element in $M_g(\mathbb{R})$. We assume $\operatorname{sgn} 0 := 0$.

(3) Let $r \geq 1$ and $I = (\alpha_1, \dots, \alpha_r) \in (\mathbb{R}^g - \{0\})^r$. We say that $x \in \mathbb{R}^g$ is *in general position relative to I* if x is not contained in any proper $\mathbb{R}$-subspace of $\mathbb{R}^g$ generated by a subset of $\{\alpha_1, \dots, \alpha_r\}$.

Remark 5.3.2. The condition "in general position relative to I" is slightly more strict than the condition "generic with respect to $\{\alpha_1, \dots, \alpha_r\}$" in the sense of Yamamoto [2010, p. 471]. Actually, this difference is not important at all, but we adopt this definition since it is more useful in this paper.

Lemma 5.3.3. *Let $r \geq 1$, $I = (\alpha_1, \dots, \alpha_r) \in (X_{\mathbb{Q}})^r$, $x \in \mathbb{R}^g - \{0\}$, and $Q \in \Xi$. Then there exists $\delta > 0$ such that* $\exp(\varepsilon Q)x$ *is in general position relative to I for all $\varepsilon \in (0, \delta)$.*

Proof. Let $W_1, \dots, W_m \subsetneq \mathbb{R}^g$ be all the proper $\mathbb{R}$-subspaces which can be generated by some subset of $\{\alpha_1, \dots, \alpha_r\}$. In particular, $y \in \mathbb{R}^g$ is in general position relative to I if and only if $y \notin \bigcup_{j=1}^m W_j$.

Take $\beta_1, \dots, \beta_m \in \mathbb{Q}^g - \{0\}$ such that $W_j \subset H_{\beta_j}$ for $j = 1, \dots, m$. (See Section 5.2 for the definition of H_{β_j}.) Then, by Lemma 5.2.3(4), for each j, there exists $\delta_j > 0$ such that

$$\exp((0, \delta_j)Q)x \subset U_{\beta_j,+} \cup U_{\beta_j,-} = \mathbb{R}^g - H_{\beta_j}.$$

Put $\delta := \min\{\delta_1, \dots, \delta_m\} > 0$. Then for all $\varepsilon \in (0, \delta)$, we have

$$\exp(\varepsilon Q)x \notin \bigcup_{j=1}^m H_{\beta_j} \supset \bigcup_{j=1}^m W_j,$$

and hence $\exp(\varepsilon Q)x$ is in general position relative to I. □

The following is the main proposition of this subsection.

Proposition 5.3.4. *Let $J = (\alpha_0, \dots, \alpha_g) \in (X_{\mathbb{Q}})^{g+1}$ and $Q \in \Xi$. Assume that there exists $y \in \mathbb{R}^g - \{0\}$ such that for all $i = 0, \dots, g$ we have $\langle \alpha_i, y \rangle > 0$. Then we have*

$$\sum_{i=0}^g (-1)^i \operatorname{sgn}(J^{(i)}) \mathbf{1}_{C^Q_{J^{(i)}}}(x) = 0$$

for $x \in \mathbb{R}^g - \{0\}$, where $J^{(i)} = (\alpha_0, \dots, \check{\alpha}_i, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$.

Proof. Take such $y \in \mathbb{R}^g - \{0\}$. We will reduce the problem to the "generic case". First, we claim that for each $i = 0, \ldots, g$, there exists $\delta_i > 0$ such that

$$\exp((0, \delta_i)Q)x \subset C_{J^{(i)}} \quad \text{or} \quad \exp((0, \delta_i)Q)x \subset \mathbb{R}^g - C_{J^{(i)}}.$$

Indeed, if $\alpha_0, \ldots, \check{\alpha}_i, \ldots, \alpha_g$ (α_i is omitted) are linearly independent, then this follows directly from Lemma 5.2.4(2). On the other hand, if $\alpha_0, \ldots, \check{\alpha}_i, \ldots, \alpha_g$ are linearly dependent, then we have $\mathrm{Span}_{\mathbb{Q}}\{\alpha_0, \ldots, \check{\alpha}_i, \ldots, \alpha_g\} \subsetneq \mathbb{Q}^g$, and hence there exists $\alpha \in \mathbb{Q}^g - \{0\}$ such that $C_{J^{(i)}} \subset H_\alpha$. Therefore, by Lemma 5.2.3(4) along with Lemma 5.1.3, we find

$$\mathbb{R}^g - \{0\} = (U_{\alpha,+})^Q \sqcup (U_{\alpha,-})^Q \subset (\mathbb{R}^g - C_{J^{(i)}})^Q,$$

and we can take such $\delta_i > 0$.

Consequently, for $i = 0, \ldots, g$, we obtain

$$\mathbf{1}_{C^Q_{J^{(i)}}}(x) = \mathbf{1}_{C_{J^{(i)}}}(\exp(\varepsilon Q)x) \quad \text{for all } \varepsilon \in (0, \delta_i).$$

On the other hand, by Lemma 5.3.3, there exists $\delta > 0$ such that for all $\varepsilon \in (0, \delta)$, $\exp(\varepsilon Q)x$ is in general position relative to J. Set $\varepsilon_0 := \frac{1}{2}\min\{\delta_0, \ldots, \delta_g, \delta\}$, and put $x' := \exp(\varepsilon_0 Q)x$. Then

- $\mathbf{1}_{C^Q_{J^{(i)}}}(x) = \mathbf{1}_{C_{J^{(i)}}}(x')$ for $i = 0, \ldots, g$,
- x' is in general position relative to J.

Therefore, it suffices to prove

$$\sum_{i=0}^{g} (-1)^i \operatorname{sgn}(J^{(i)}) \mathbf{1}_{C_{J^{(i)}}}(x') = 0 \tag{5-4}$$

for any x' which is in general position relative to J. First, if $\langle x', y \rangle \le 0$, then we have

$$\mathbf{1}_{C_{J^{(i)}}}(x') = 0 \quad \text{for all } i \in \{0, \ldots, g\}$$

because $\langle \alpha_i, y \rangle > 0$ for all $i = 0, \ldots, g$. Therefore, we may assume $\langle x', y \rangle > 0$. In this case, the identity (5-4) follows from [Yamamoto 2010, Proposition 6.2].

Indeed, let $\gamma \in \mathrm{GL}_g(\mathbb{R})$ such that ${}^t\gamma e_g = y$, where $e_g = {}^t(0, \ldots, 0, 1) \in \mathbb{R}^g$. Then

- $\gamma x', \gamma\alpha_0, \ldots, \gamma\alpha_g \in \mathcal{H} := \{v \in \mathbb{R}^g \mid \langle v, e_g \rangle > 0\}$,
- $\gamma x'$ is in general position relative to γJ,
- $\operatorname{sgn}(\gamma J^{(i)}) = \operatorname{sgn}(\det(\gamma)) \operatorname{sgn}(J^{(i)})$,
- $\mathbf{1}_{C_{J^{(i)}}}(x') = \mathbf{1}_{C_{\gamma J^{(i)}}}(\gamma x')$,

and hence we can use [Yamamoto 2010, Proposition 6.2]. This completes the proof. □

Remark 5.3.5. It is also possible to prove the last part using [Charollois et al. 2015, Theorem 2.1].

6. Construction of the Shintani–Barnes cocycle

Recall that for $d \geq 0$, we have sheaves

$$\mathscr{F}_d = \pi_{\mathbb{C}}^{-1}\Omega^{g-1}_{\mathbb{P}^{g-1}(\mathbb{C})}(-d)|_{Y^\circ} \quad \text{and} \quad \mathscr{F}_d^{\Xi} = \underline{\mathrm{Hom}}(\underline{\mathbb{Z}[\Xi]}, \mathscr{F}_d) \simeq \prod_{Q\in\Xi} \mathscr{F}_d$$

on $Y^\circ = \mathbb{C}^g - \boldsymbol{i}\mathbb{R}^g$. In this section we construct a certain cohomology class in $H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^{\Xi})$ using the Čech complex $\mathscr{C}^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}$.

6.1. ***Barnes zeta function associated to C_I^Q.*** Recall that for $I = (\alpha_1, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$, the open subset $V_I \subset Y^\circ$ is defined as

$$V_I = \{y \in Y^\circ \mid \mathrm{Re}(\langle \alpha_i, y\rangle) > 0 \text{ for } i = 1, \dots, g\},$$

and we have

$$\Gamma(V_I, \mathscr{F}_d) = \{f\omega \mid f \text{ holomorphic function on } \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I)) \text{ such that } f(\lambda y) = \lambda^{-g-d} f(y) \text{ for all } \lambda \in \mathbb{C}^\times\}$$

by Proposition 3.2.3. Note that $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I)) \subset \mathbb{C}^g - \{0\}$ is an open subset of the following form:

$$\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I)) = \{y \in \mathbb{C}^g \mid \text{there exists } \lambda \in \mathbb{C}^\times \text{ such that } \lambda y \in V_I\} \subset \mathbb{C}^g - \{0\}.$$

Definition 6.1.1. For $d \geq 1$, $I = (\alpha_1, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$, $Q \in \Xi$, and $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$, set

$$\psi_{d,I}^{Q}(y) := \mathrm{sgn}(I) \sum_{x \in C_I^Q \cap \mathbb{Z}^g - \{0\}} \frac{1}{\langle x, y\rangle^{g+d}}, \tag{6-1}$$

where $\mathrm{sgn}(I) = \mathrm{sgn}\det(\alpha_1, \dots, \alpha_g) \in \{-1, 0, 1\}$; see Section 5.3.

Proposition 6.1.2. *The infinite series (6-1) converges absolutely and locally uniformly for $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$. In particular, $\psi_{d,I}^Q$ is a holomorphic function on $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$. Moreover, we have*

$$\psi_{d,I}^Q(\lambda y) = \lambda^{-g-d}\psi_{d,I}^Q(y)$$

for all $\lambda \in \mathbb{C}^\times$ and $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$.

Proof. If $\mathrm{sgn}(I) = 0$, then by Proposition 5.2.1(1), we see that $C_I^Q \cap \mathbb{Z}^g - \{0\} = \varnothing$, and hence the sum is zero. (In particular, the series converges.) Therefore, we may assume that $\alpha_1, \dots, \alpha_g$ form a basis of $\mathbb{Q}^g$. Furthermore, since $\mathrm{sgn}(I)$ and C_I^Q do not change if we replace α_i by its multiple by positive integers, we may assume that $\alpha_1, \dots, \alpha_g \in \mathbb{Z}^g - \{0\}$.

Let $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$ and take $\lambda \in \mathbb{C}^\times$ such that $\lambda y \in V_I$. Then take a relatively compact open neighborhood $U \subset V_I$ of λy, i.e., U is an open neighborhood of λy such that its closure $\overline{U}$ is compact and $\overline{U} \subset V_I$. Since $y \in \lambda^{-1}\overline{U} \subset \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$, it suffices to show that (6-1) converges absolutely and uniformly on $\lambda^{-1}\overline{U}$.

First, note that by the definition of C_I^Q, we have

$$C_I^Q \subset \overline{C}_I = \sum_{i=1}^{g} \mathbb{R}_{\geq 0}\alpha_i,$$

where $\bar{C}_I$ is the closed cone generated by I. Put

$$R_I := \sum_{i=1}^{g} [0, 1)\alpha_i.$$

Then we see

- $C_I^Q \cap \mathbb{Z}^g \subset \bar{C}_I \cap \mathbb{Z}^g = \{x + \sum_{i=1}^g n_i\alpha_i \mid x \in R_I \cap \mathbb{Z}^g,\ n_i \in \mathbb{Z}_{\geq 0}\}$,
- $R_I \cap \mathbb{Z}^g$ is a finite set,
- $\{\mathrm{Re}(\langle x, y'\rangle) \mid x \in R_I \cap \mathbb{Z}^g - \{0\},\ y' \in \bar{U}\}$ is a compact subset of $\mathbb{R}_{>0}$.

Therefore, set

$$b := \min\{\mathrm{Re}(\langle x, y'\rangle) \mid x \in R_I \cap \mathbb{Z}^g - \{0\},\ y' \in \bar{U}\} > 0.$$

Moreover, for $i = 1, \dots, g$, set

$$a_i := \min\{\mathrm{Re}(\langle \alpha_i, y'\rangle) \mid y' \in \bar{U}\} > 0.$$

Then for $y'' = \lambda^{-1}y' \in \lambda^{-1}\bar{U}$, where $y' \in \bar{U}$, we have

$$\begin{aligned}
&\sum_{x \in C_I^Q \cap \mathbb{Z}^g - \{0\}} \left| \frac{1}{\langle x, y''\rangle^{g+d}} \right| \\
&\leq |\lambda|^{g+d} \sum_{x \in \bar{C}_I \cap \mathbb{Z}^g - \{0\}} \frac{1}{|\langle x, y'\rangle|^{g+d}} \\
&\leq |\lambda|^{g+d} \sum_{x \in \bar{C}_I \cap \mathbb{Z}^g - \{0\}} \frac{1}{\big(\mathrm{Re}(\langle x, y'\rangle)\big)^{g+d}} \\
&\leq |\lambda|^{g+d} \sum_{\substack{x' \in R_I \cap \mathbb{Z}^g,\ (n_1, \dots, n_g) \in (\mathbb{Z}_{\geq 0})^g, \\ x' + \sum_{i=1}^g n_i\alpha_i \neq 0}} \frac{1}{\big(\mathrm{Re}(\langle x', y'\rangle) + \sum_{i=1}^g n_i \,\mathrm{Re}(\langle \alpha_i, y'\rangle)\big)^{g+d}} \\
&\leq |\lambda|^{g+d} \sum_{(n_1, \dots, n_g) \in (\mathbb{Z}_{\geq 0})^g - \{0\}} \frac{1}{\big(\sum_{i=1}^g n_i a_i\big)^{g+d}} + |\lambda|^{g+d} \#(R_I \cap \mathbb{Z}^g - \{0\}) \sum_{(n_1, \dots, n_g) \in (\mathbb{Z}_{\geq 0})^g} \frac{1}{\big(b + \sum_{i=1}^g n_i a_i\big)^{g+d}},
\end{aligned}$$

where $\#(R_I \cap \mathbb{Z}^g - \{0\})$ is the order of the finite set $R_I \cap \mathbb{Z}^g - \{0\}$. It is now clear that the last two series converge for $d \geq 1$. The last statement in the proposition follows directly from the definition. □

Remark 6.1.3. Since C_I^Q is a rational constructible cone (see Proposition 5.2.1), we see that $\psi_{d,I}^Q$ can be written as a sum of a finite number of the Barnes zeta functions; see [Barnes 1904; Yamamoto 2010]. Conceptually, we may also view $\psi_{d,I}^Q$ as a decomposed piece of the "Eisenstein series"

$$\psi_d(y) = \sum_{x \in \mathbb{Z}^g - \{0\}} \frac{1}{\langle x, y\rangle^{g+d}},$$

which coincides with the classical holomorphic Eisenstein series of weight $2 + d$ if $g = 2$, $d \geq 2$ is even, and $y = (1, z)$ with $\mathrm{Im}(z) > 0$, but does not converge if $g \geq 3$. Therefore, the following construction of the Shintani–Barnes cocycle can be seen as a cohomological realization of this (generally) nonconvergent Eisenstein series.

Corollary 6.1.4. *Let $d \geq 1$. For $I = (\alpha_1, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$ and $Q \in \Xi$, we have*

$$\psi^Q_{d,I}\omega \in \Gamma(V_I, \mathscr{F}_d), \quad \textit{where } \omega(y) = \sum_{i=1}^{g} (-1)^{i-1} y_i \, dy_1 \wedge \cdots \wedge \check{dy_i} \wedge \cdots \wedge dy_g.$$

Proof. This follows directly from Propositions 3.2.3(2) and 6.1.2. □

6.2. *The Shintani–Barnes cocycle.*

Definition 6.2.1. For $d \geq 1$, we define a map $\Psi_d : \Xi \to C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)$ by

$$\Psi_d(Q) := (\psi^Q_{d,I}\omega)_{I \in (X_{\mathbb{Q}})^g} \in C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d) = \prod_{I \in (X_{\mathbb{Q}})^g} \Gamma(V_I, \mathscr{F}_d) \qquad \text{for } Q \in \Xi.$$

We aim to show that Ψ_d defines a class in $H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^{\Xi})$ via Corollary 4.3.4.

Proposition 6.2.2. *The map Ψ_d is a $\mathrm{SL}_g(\mathbb{Z})$-equivariant map, i.e., we have*

$$\Psi_d\big([\gamma](Q)\big) = [\gamma](\Psi_d(Q))$$

for $Q \in \Xi$ and $\gamma \in \mathrm{SL}_g(\mathbb{Z})$. In other words, we have

$$\Psi_d \in \mathrm{Map}_{\mathrm{SL}_g(\mathbb{Z})}(\Xi, C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)) = \Gamma\big(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}\big).$$

Proof. Let $I = (\alpha_1, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$. We need to show

$$\Psi_d([\gamma](Q))_I = \big([\gamma](\Psi_d(Q))\big)_I \in \Gamma(V_I, \mathscr{F}_d),$$

where $\Psi_d([\gamma](Q))_I$ $\big($resp. $\big([\gamma](\Psi_d(Q))\big)_I\big)$ is the I-th component of $\Psi_d\big([\gamma](Q)\big)$ $\big($resp. $[\gamma](\Psi_d(Q))\big)$ as always. Indeed, we have

$$\begin{aligned}
\big([\gamma](\Psi_d(Q))\big)_I(y) &= \big([\gamma](\psi^Q_{d,\gamma^{-1}I}\omega)\big)(y) \\
&= \psi^Q_{d,\gamma^{-1}I}({}^t\gamma y)\omega({}^t\gamma y) \\
&= \mathrm{sgn}(\gamma^{-1}I) \sum_{x \in C^Q_{\gamma^{-1}I} \cap \mathbb{Z}^g - \{0\}} \frac{\omega({}^t\gamma y)}{\langle x, {}^t\gamma y\rangle^{g+d}} \\
&= \mathrm{sgn}(\det(\gamma^{-1}))\, \mathrm{sgn}(I) \det({}^t\gamma) \sum_{x \in C^Q_{\gamma^{-1}I} \cap \mathbb{Z}^g - \{0\}} \frac{\omega(y)}{\langle \gamma x, y\rangle^{g+d}} \\
&= \mathrm{sgn}(I) \sum_{x \in \gamma(C^Q_{\gamma^{-1}I}) \cap \mathbb{Z}^g - \{0\}} \frac{\omega(y)}{\langle x, y\rangle^{g+d}} \\
&= \mathrm{sgn}(I) \sum_{x \in C^{[\gamma](Q)}_I \cap \mathbb{Z}^g - \{0\}} \frac{\omega(y)}{\langle x, y\rangle^{g+d}} = \Psi_d([\gamma](Q))_I(y)
\end{aligned}$$

for $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$. Here, the first and second equalities follow from the definition of $[\gamma]$ (see (4-1) and Definition 3.3.2), the fourth equality follows from Lemma 3.3.1, and the sixth equality follows from Lemma 5.1.4. □

Corollary 6.2.3. *For $Q \in \Xi$, we have*

$$\Psi_d(Q) \in C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Gamma_Q} = \Gamma(Y^\circ, \Gamma_Q, \mathscr{C}^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)).$$

Proof. Because Γ_Q is the stabilizer of Q in $\mathrm{SL}_g(\mathbb{Z})$ and Ψ_d is a $\mathrm{SL}_g(\mathbb{Z})$-equivariant map, it follows that $\Psi_d(Q)$ is a Γ_Q-invariant element. □

Proposition 6.2.4. (1) *Let $Q \in \Xi$. We have*

$$\mathrm{d}^{g-1}(\Psi_d(Q)) = 0$$

under the differential map

$$\mathrm{d}^{g-1} : C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d) \to C^g(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d).$$

(2) *We have*

$$\mathrm{d}^{g-1}(\Psi_d) = 0$$

under the differential map

$$\mathrm{d}^{g-1} : \Gamma\big(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}\big) \to \Gamma\big(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{C}^{g}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Xi}\big).$$

In the following, we refer to Ψ_d as the Shintani–Barnes cocycle.

Proof. (1) Let $J = (\alpha_0, \dots, \alpha_g) \in (X_{\mathbb{Q}})^{g+1}$. For $i = 0, \dots, g$, put $J^{(i)} = (\alpha_0, \dots, \check{\alpha}_i, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$. We need to show

$$\big(\mathrm{d}^{g-1}(\Psi_d(Q))\big)_J = \sum_{i=0}^{g} (-1)^i \Psi_d(Q)_{J^{(i)}}|_{V_J} = 0. \tag{6-2}$$

First if $V_J = \varnothing$, then (6-2) is obvious because $\Gamma(\varnothing, \mathscr{F}_d) = 0$. Assume $V_J \neq \varnothing$, and take $y' \in V_J$. Then we have $\langle \alpha_i, \mathrm{Re}(y')\rangle = \mathrm{Re}(\langle \alpha_i, y'\rangle) > 0$ for all $i = 0, \dots, g$, and hence the assumption in Proposition 5.3.4 is satisfied. Therefore, by Proposition 5.3.4, we find

$$\begin{aligned}
\sum_{i=0}^{g} (-1)^i \Psi_d(Q)_{J^{(i)}}|_{V_J}(y) &= \sum_{i=0}^{g} (-1)^i \,\mathrm{sgn}(J^{(i)}) \sum_{x \in C^Q_{J^{(i)}} \cap \mathbb{Z}^g - \{0\}} \frac{1}{\langle x, y\rangle^{g+d}} \omega(y) \\
&= \sum_{x \in \mathbb{Z}^g - \{0\}} \left(\sum_{i=0}^{g} (-1)^i \,\mathrm{sgn}(J^{(i)}) \mathbf{1}_{C^Q_{J^{(i)}}}(x) \right) \frac{\omega(y)}{\langle x, y\rangle^{g+d}} \\
&= 0
\end{aligned}$$

for $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_J))$. This proves (1).

(2) This follows from (1). □

We obtain the following.

Theorem 6.2.5. *For $d \geq 1$, the Shintani–Barnes cocycle Ψ_d defines a class*

$$[\Psi_d] \in H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_d^{\Xi}).$$

Moreover, for $Q \in \Xi$, the element $\Psi_d(Q) \in C^{g-1}(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_d)^{\Gamma_Q}$ defines a class

$$[\Psi_d(Q)] \in H^{g-1}(Y^\circ, \Gamma_Q, \mathscr{F}_d),$$

and we have

$$\mathrm{ev}_Q([\Psi_d]) = [\Psi_d(Q)].$$

Proof. This follows from Corollary 4.3.4, Proposition 6.2.2, Corollary 6.2.3, and Proposition 6.2.4. □

7. Integration

The goal of the remaining sections is to construct a specialization map (8-11), and prove that the Shintani–Barnes cocycle class $[\Psi_d]$ specializes to the special value of the zeta functions of number fields; see Theorem 8.3.2.

Let $Q \in \Xi$ be fixed throughout this section. In this section we define an integral map

$$\int_Q : H^q(Y^\circ, \Gamma_Q, \mathscr{F}_0) \to H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C}),$$

where $H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C})$ is a certain auxiliary cohomology group defined later; see Section 7.2. This group $H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C})$ will be studied more closely in Section 8 using a topological method.

7.1. ***Integration and the Hurwitz formula.*** For $q \geq 0$, let

$$\Delta^q := \left\{(t_1, \ldots, t_{q+1}) \in \mathbb{R}^{q+1} \,\middle|\, \sum_{i=1}^{q+1} t_i = 1,\ t_i \geq 0\right\}$$

denote the standard q-simplex. Note that we can also embed Δ^q into $\mathbb{R}^q$ by

$$\Delta^q \hookrightarrow \mathbb{R}^q, \quad (t_1, \ldots, t_{q+1}) \mapsto (t_2, \ldots, t_{q+1}),$$

and we equip Δ^q with an orientation induced from the standard orientation of $\mathbb{R}^q$. Moreover, for $\xi_1, \ldots, \xi_{q+1} \in \mathbb{C}^g - \{0\}$, let

$$\sigma_{(\xi_1,\ldots,\xi_{q+1})} : \Delta^q \to \mathbb{C}^g, \quad (t_1, \ldots, t_{q+1}) \mapsto \sum_{i=1}^{q+1} t_i \xi_i$$

denote the affine q-simplex with vertices $\xi_1, \ldots, \xi_{q+1}$, and let

$$|\sigma_{(\xi_1,\ldots,\xi_{q+1})}| := \sigma_{(\xi_1,\ldots,\xi_{q+1})}(\Delta^q) \subset \mathbb{C}^g$$

denote the image of $\sigma_{(\xi_1,\ldots,\xi_{q+1})}$.

Now, let $U \subset \mathbb{C}^g - \{0\}$ be a convex open subset and let $\xi_1, \ldots, \xi_g \in U$ be a basis of $\mathbb{C}^g$. Then for a homogeneous holomorphic function f on $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U))$ of degree $-g$, (i.e., $f(\lambda y) = \lambda^{-g} f(y)$ for all $\lambda \in \mathbb{C}^\times$), we consider the integral

$$\int_{\sigma_{(\xi_1,\ldots,\xi_g)}} f\omega := \int_{\Delta^{g-1}} (\sigma_{(\xi_1,\ldots,\xi_g)})^*(f\omega), \tag{7-1}$$

where

$$\omega(y) = \sum_{i=1}^{g} (-1)^{i-1} y_i \, dy_1 \wedge \cdots \wedge \check{dy_i} \wedge \cdots \wedge dy_g.$$

Here note that $f\omega$ is a holomorphic $(g-1)$-form on $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U)) \supset U$, and we have $|\sigma_{(\xi_1,\ldots,\xi_g)}| \subset U$ since U is convex.

Remark 7.1.1. Via the identification (3-2), the above $f\omega$ corresponds to a holomorphic $(g-1)$-form on $\pi_{\mathbb{C}}(U) \subset \mathbb{P}^{g-1}(\mathbb{C})$. More precisely, there is a holomorphic $(g-1)$-form η on $\pi_{\mathbb{C}}(U) \subset \mathbb{P}^{g-1}(\mathbb{C})$ such that

$$(\pi_{\mathbb{C}})^* \eta = f\omega.$$

Then we see that the integral (7-1) is actually an integral on $\mathbb{P}^{g-1}(\mathbb{C})$:

$$\int_{\sigma_{(\xi_1,\ldots,\xi_g)}} f\omega = \int_{\pi_{\mathbb{C}} \circ \sigma_{(\xi_1,\ldots,\xi_g)}} \eta.$$

Lemma 7.1.2. *Let $U \subset \mathbb{C}^g - \{0\}$ be a convex open subset, and let $\xi_1, \ldots, \xi_g \in \mathbb{C}^g - \{0\}$ be a basis of $\mathbb{C}^g$ such that*

$$\xi_1, \ldots, \xi_g \in U.$$

Furthermore, let $\lambda_1, \ldots, \lambda_g \in \mathbb{C}^\times$ be any complex numbers such that

$$\lambda_1 \xi_1, \ldots, \lambda_g \xi_g \in U.$$

Then for a homogeneous holomorphic function f on $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(U))$ of degree $-g$, we have

$$\int_{\sigma_{(\xi_1,\ldots,\xi_g)}} f\omega = \int_{\sigma_{(\lambda_1\xi_1,\ldots,\lambda_g\xi_g)}} f\omega.$$

Proof. Let

$$h : [0,1] \times \Delta^{g-1} \to U, \quad (u,t) \mapsto u\sigma_{(\xi_1,\ldots,\xi_g)}(t) + (1-u)\sigma_{(\lambda_1\xi_1,\ldots,\lambda_g\xi_g)}(t) = \sum_{i=1}^{g} (u + (1-u)\lambda_i) t_i \xi_i \tag{7-2}$$

be a homotopy between $\sigma_{(\xi_1,\ldots,\xi_g)}$ and $\sigma_{(\lambda_1\xi_1,\ldots,\lambda_g\xi_g)}$. Note that we have $h(u,t) \in U$ because U is convex. We regard h as a singular g-chain in a usual way using the standard decomposition of the prism $[0,1] \times \Delta^{g-1}$; see [Hatcher 2002, Section 2.1, Proof of 2.10]. Then we have

$$\partial h = \sigma_{(\xi_1,\ldots,\xi_g)} - \sigma_{(\lambda_1\xi_1,\ldots,\lambda_g\xi_g)} + h',$$

where

$$h' : [0,1] \times \partial\Delta^{g-1} \to U, \quad (u,t) \mapsto h(u,t),$$

which is also regarded as a singular $(g-1)$-chain. Let $\xi_1^*, \dots, \xi_g^* \in \mathbb{C}^g$ be the dual basis of $\xi_1, \dots, \xi_g$, and let

$$Z := \bigcup_{i=1}^{g} \{y \in \mathbb{C}^g \mid \langle \xi_i^*, y\rangle = 0\}$$

be the union of hyperplanes defined by $\xi_1^*, \dots, \xi_g^*$. Then, by (7-2), we easily see

$$h'([0,1] \times \partial\Delta^{g-1}) \subset Z.$$

Now, by Remark 7.1.1, there exists a holomorphic $(g-1)$-form η on $\pi_{\mathbb{C}}(U)$ such that

$$(\pi_{\mathbb{C}})^*\eta = f\omega.$$

In particular, we have

$$d(f\omega) = (\pi_{\mathbb{C}})^*(d\eta) = 0,$$

where d is the usual derivative of differential forms. Moreover, we also have

$$\int_{h'} f\omega = \int_{\pi_{\mathbb{C}} \circ h'} \eta = 0$$

because $\pi_{\mathbb{C}} \circ h'$ is contained in a divisor $\pi_{\mathbb{C}}(Z-\{0\}) \subset \mathbb{P}^{g-1}(\mathbb{C})$. Therefore, we obtain

$$0 = \int_h d(f\omega) = \int_{\partial h} f\omega = \int_{\sigma(\xi_1,\dots,\xi_g)} f\omega - \int_{\sigma(\lambda_1\xi_1,\dots,\lambda_g\xi_g)} f\omega + \int_{h'} f\omega = \int_{\sigma(\xi_1,\dots,\xi_g)} f\omega - \int_{\sigma(\lambda_1\xi_1,\dots,\lambda_g\xi_g)} f\omega.$$

This completes the proof. □

An important example of such an integral is the following Hurwitz formula (see [Hurwitz 1922; Sczech 1993]), which is also known as the Feynman parametrization.

Proposition 7.1.3 [Hurwitz 1922]. *Let $x \in \mathbb{C}^g - \{0\}$, and let $\xi_1, \dots, \xi_g \in \mathbb{C}^g - \{0\}$ be a basis of $\mathbb{C}^g$ such that*

$$\xi_1, \dots, \xi_g \in V_x = \{y \in \mathbb{C}^g - \{0\} \mid \mathrm{Re}(\langle x, y\rangle) > 0\}.$$

(1) *We have*

$$\int_{\sigma(\xi_1,\dots,\xi_g)} \frac{\omega(y)}{\langle x, y\rangle^g} = \frac{1}{(g-1)!} \frac{\det(\xi_1, \dots, \xi_g)}{\langle x, \xi_1\rangle \cdots \langle x, \xi_g\rangle}. \qquad (*)$$

(2) *Let $\xi_1^*, \dots, \xi_g^* \in \mathbb{C}^g$ be the dual basis of $\xi_1, \dots, \xi_g$, and let $\underline{k} = (k_1, \dots, k_g) \in (\mathbb{Z}_{\geq 0})^g$. Then*

$$\int_{\sigma(\xi_1,\dots,\xi_g)} \langle \xi_1^*, y\rangle^{k_1} \cdots \langle \xi_g^*, y\rangle^{k_g} \frac{\omega(y)}{\langle x, y\rangle^{g+|\underline{k}|}} = \frac{\underline{k}!}{(g+|\underline{k}|-1)!} \frac{\det(\xi_1, \dots, \xi_g)}{\langle x, \xi_1\rangle^{k_1+1} \cdots \langle x, \xi_g\rangle^{k_g+1}},$$

where $|\underline{k}| := k_1 + \cdots + k_g$ and $\underline{k}! := k_1! \cdots k_g!$.

Proof. (1) Let $W := (\xi_1, \dots, \xi_g) \in \mathrm{GL}_g(\mathbb{C})$ be the matrix whose columns are $\xi_1, \dots, \xi_g$ so that the $(g-1)$-simplex $\sigma_{(\xi_1,\dots,\xi_g)}$ is represented by the linear transformation W, i.e., we have $\sigma_{(\xi_1,\dots,\xi_g)}(t_1, \dots, t_g) = W^t(t_1, \dots, t_g)$ for $(t_1, \dots, t_g) \in \Delta^{g-1} \subset \mathbb{R}^g$. Then

$$\int_{\sigma(\xi_1,\dots,\xi_g)} \frac{\omega(y)}{\langle x, y\rangle^g} = \int_{\Delta^{g-1}} \frac{\omega(Wy)}{\langle x, Wy\rangle^g} = \det W \int_{\Delta^{g-1}} \frac{\omega(y)}{\langle {}^tWx, y\rangle^g}.$$

For $i = 1, \dots, g$, put

$$a_i := \langle x, \xi_i \rangle \neq 0,$$

and let $e_1, \dots, e_g \in \mathbb{C}^g$ be the standard basis, i.e., $e_i = {}^t(0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$. Then we find

$$\begin{aligned}
\det W \int_{\Delta^{g-1}} \frac{\omega(y)}{\langle {}^tWx, y\rangle^g} &= \frac{\det W}{a_1 \cdots a_g} \int_{\Delta^{g-1}} \frac{\omega((a_1 y_1, \dots, a_g y_g))}{(a_1 y_1 + \cdots + a_g y_g)^g} \\
&= \frac{\det W}{a_1 \cdots a_g} \int_{\sigma(a_1 e_1, \dots, a_g e_g)} \frac{\omega(y)}{(y_1 + \cdots + y_g)^g} \\
&= \frac{\det W}{a_1 \cdots a_g} \int_{\sigma(e_1, \dots, e_g)} \frac{\omega(y)}{(y_1 + \cdots + y_g)^g} \\
&= \frac{\det W}{a_1 \cdots a_g} \int_{\sigma(e_1, \dots, e_g)} \omega(y) \\
&= \frac{1}{(g-1)!} \frac{\det W}{a_1 \cdots a_g}.
\end{aligned}$$

Here, the third equality follows from Lemma 7.1.2, and the last equality follows from an elementary computation. This proves (1).

(2) First note that for fixed $\xi_1, \dots, \xi_g$, the formula $(*)$ can be seen as an equality of holomorphic functions in the x-variable. Thus, for $1 \le i \le g$, we consider a linear differential operator

$$D_i := \left\langle \xi_i^*, \frac{\partial}{\partial x} \right\rangle = \xi_{i1}^* \frac{\partial}{\partial x_1} + \cdots + \xi_{ig}^* \frac{\partial}{\partial x_g},$$

where ξ_{ij}^* is the j-th component of ξ_i^*. Then we can compute the action of D_i on the both sides of $(*)$ using the formula

$$D_i \frac{1}{\langle x, y\rangle^n} = -n \langle \xi_i^*, y\rangle \frac{1}{\langle x, y\rangle^{n+1}},$$

where $y \in \mathbb{C}^g$, $\langle x, y\rangle \neq 0$, and $n \ge 1$. Now (2) follows from (1) by applying to $(*)$ the operator

$$D_1^{k_1} \cdots D_g^{k_g}. \qquad \square$$

Remark 7.1.4. The right-hand side of the Hurwitz formula (Proposition 7.1.3) is exactly the building block of Sczech's Eisenstein cocycle [Sczech 1993].

7.2. *The integral map $\int_Q$.* Let $Q \in \Xi$, and let $\theta^{(1)}, \dots, \theta^{(g)} \in \mathbb{C}$ be the distinct eigenvalues of ${}^tQ$. Note that by Lemma 2.1.1(1), ${}^tQ$ has g distinct eigenvalues.

We will introduce an auxiliary cohomology group $H_Q^q(Y^\circ, \Gamma_Q, \mathbb{C})$ and define the integral map $\int_Q$.

Definition 7.2.1. Let $q \ge 0$. We say that $I \in (X_{\mathbb{Q}})^{q+1}$ is *Q-admissible* if we can take a system of eigenvectors $\xi_1, \dots, \xi_g$ of ${}^tQ$ in V_I, i.e., if

$$\text{there exists } \xi_1, \dots, \xi_g \in V_I \text{ such that } {}^tQ\xi_i = \theta^{(i)}\xi_i \text{ for } i = 1, \dots, g.$$

We define $(X_{\mathbb{Q}})_Q^{q+1}$ to be the set of all Q-admissible elements of $(X_{\mathbb{Q}})^{q+1}$.

Recall that

$\Gamma(V_I, \mathscr{F}_0) = \{f\omega \mid f$ holomorphic function on $\pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$ such that $f(\lambda y) = \lambda^{-g} f(y)$ for all $\lambda \in \mathbb{C}^\times\}$.

Definition 7.2.2. For $q \geq 0$ and a Q-admissible $I \in (X_{\mathbb{Q}})_Q^{q+1}$, we define a map

$$\int_{Q,I} : \Gamma(V_I, \mathscr{F}_0) \to \mathbb{C}, \quad s \mapsto \int_{Q,I} s \tag{7-3}$$

as follows. Take $\xi_1, \dots, \xi_g \in V_I$ such that ${}^t Q \xi_i = \theta^{(i)} \xi_i$ for $i = 1, \dots, g$, and define

$$\int_{Q,I} f\omega := \int_{\sigma(\xi_1, \dots, \xi_g)} f\omega$$

for $f\omega \in \Gamma(V_I, \mathscr{F}_0)$. Note that by Lemma 7.1.2, the map $\int_{Q,I}$ is independent of the choice of the eigenvectors $\xi_1, \dots, \xi_g$.

Remark 7.2.3. Strictly speaking, the map $\int_Q$ is depending on the (fixed) choice of the order of the eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ up to sign.

Example 7.2.4. Let the notation be the same as in Section 2.2. Furthermore, let $\theta \in F^\times$ and $Q = \rho_w(\theta) \in \Xi$ be as in Lemma 2.2.1, and let $I \in (X_{\mathbb{Q}})_Q^g$.

(1) For $k \geq 0$ and $x \in C_I^Q - \{0\}$, we have

$$N_{w^*}(y)^k \frac{\omega(y)}{\langle x, y\rangle^{g+kg}} \in \Gamma(V_I, \mathscr{F}_0),$$

and

$$\int_{Q,I} N_{w^*}(y)^k \frac{\omega(y)}{\langle x, y\rangle^{g+kg}} = \frac{(k!)^g}{(g+kg-1)!} \frac{\det(w^{(1)}, \dots, w^{(g)})}{N_w(x)^{k+1}}.$$

(2) For $k \geq 1$, we have

$$N_{w^*}(y)^k \psi_{kg,I}^Q(y)\omega(y) \in \Gamma(V_I, \mathscr{F}_0),$$

and

$$\int_{Q,I} N_{w^*}(y)^k \psi_{kg,I}^Q(y)\omega(y) = \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+kg-1)!} \operatorname{sgn}(I) \sum_{x \in C_I^Q \cap \mathbb{Z}^g - \{0\}} \frac{1}{N_w(x)^{k+1}}.$$

Proof. (1) First, since $x \in C_I^Q - \{0\}$, we easily see $\operatorname{Re}(\langle x, y\rangle) > 0$ for all $y \in V_I$, i.e., $V_I \subset V_x$. In particular, $\langle x, y\rangle \neq 0$ for all $y \in \pi_{\mathbb{C}}^{-1}(\pi_{\mathbb{C}}(V_I))$, and hence we obtain the first assertion. Now, by Lemma 2.2.1(5), we know $w^{(1)}, \dots, w^{(g)} \in \mathbb{C}^g$ are the eigenvectors of ${}^t Q$ with eigenvalues $\theta^{(1)} := \tau_1(\theta), \dots, \theta^{(g)} := \tau_g(\theta) \in \mathbb{C}$, respectively. Take $\mu_1, \dots, \mu_g \in \mathbb{C}^\times$ so that $\xi_1 := \mu_1 w^{(1)}, \dots, \xi_g := \mu_g w^{(g)} \in V_I$. This is possible since I is Q-admissible. Then, by Lemma 2.2.1(3), we see that $\xi_1^* := \mu_1^{-1} w^{*(1)}, \dots, \xi_g^* := \mu_g^{-1} w^{*(g)}$ form the dual basis of $\xi_1, \dots, \xi_g$. Thus, by Proposition 7.1.3, we find

$$\begin{aligned}\int_{Q,I} N_{w^*}(y)^k \frac{\omega(y)}{\langle x, y\rangle^{g+kg}} &= \int_{\sigma(\xi_1,\dots,\xi_g)} \prod_{i=1}^{g} \langle \mu_i \xi_i^*, y\rangle^k \frac{\omega(y)}{\langle x, y\rangle^{g+kg}} \\ &= (\mu_1 \cdots \mu_g)^k \frac{(k!)^g}{(g+kg-1)!} \frac{\det(\xi_1,\dots,\xi_g)}{\prod_{i=1}^{g}\langle x, \xi_i\rangle^{k+1}} \\ &= (\mu_1 \cdots \mu_g)^k \frac{(k!)^g}{(g+kg-1)!} \frac{\det(\mu_1 w^{(1)},\dots,\mu_g w^{(g)})}{\prod_{i=1}^{g}\langle x, \mu_i w^{(i)}\rangle^{k+1}} \\ &= \frac{(k!)^g}{(g+kg-1)!} \frac{\det(w^{(1)},\dots,w^{(g)})}{N_w(x)^{k+1}}.\end{aligned}$$

(2) The first assertion follows from Proposition 6.1.2. The integral formula follows from (1) by taking the sum over $x \in C_I^Q \cap \mathbb{Z}^g - \{0\}$. □

Next, we extend the map (7-3) to the cohomology group.

Lemma 7.2.5. *Let $I = (\alpha_0, \dots, \alpha_q) \in (X_{\mathbb{Q}})^{q+1}$.*

(1) *If $q \geq 1$ and I is Q-admissible, then so is $I^{(i)} = (\alpha_0, \dots, \check{\alpha}_i, \dots, \alpha_q)$ for $i = 0, \dots, q$.*

(2) *Let $\gamma \in \Gamma_Q$. If I is Q-admissible, then so is γI, i.e., $(X_{\mathbb{Q}})_Q^{q+1}$ is a Γ_Q-stable subset of $(X_{\mathbb{Q}})^{q+1}$.*

Proof. (1) This follows from the fact $V_I = V_{I^{(i)}} \cap V_{\alpha_i} \subset V_{I^{(i)}}$.

(2) Take $\xi_1, \dots, \xi_g \in V_I$ such that ${}^tQ\xi_i = \theta^{(i)}\xi_i$ for $i = 1, \dots, g$. Then since ${}^tQ{}^t\gamma = {}^t\gamma{}^tQ$, we see that ${}^t\gamma^{-1}\xi_1, \dots, {}^t\gamma^{-1}\xi_g$ are again eigenvectors of ${}^tQ$ with eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ respectively. On the other hand, by Lemma 4.1.1, we have

$${}^t\gamma^{-1}\xi_i \in {}^t\gamma^{-1}V_I = V_{\gamma I}$$

for $i = 1, \dots, g$. Thus we find that ${}^t\gamma^{-1}\xi_1, \dots, {}^t\gamma^{-1}\xi_g$ are a system of eigenvectors of ${}^tQ$ in $V_{\gamma I}$. □

For a Γ_Q-equivariant sheaf $\mathscr{F}$ on Y°, set

$$C_Q^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}) := \prod_{I \in (X_{\mathbb{Q}})_Q^{q+1}} \Gamma(V_I, \mathscr{F}) \quad \text{and} \quad {}_QC^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}) := \prod_{\substack{I \in (X_{\mathbb{Q}})^{q+1}, \\ I \notin (X_{\mathbb{Q}})_Q^{q+1}}} \Gamma(V_I, \mathscr{F}).$$

Then we have a natural short exact sequence

$$0 \to {}_QC^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}) \to C^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}) \xrightarrow{p_Q} C_Q^q(\mathscr{X}_{\mathbb{Q}}, \mathscr{F}) \to 0, \tag{7-4}$$

where p_Q is the natural projection. By Lemma 7.2.5, we easily see that ${}_QC^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})$ becomes a Γ_Q-equivariant subcomplex of $C^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})$, and hence $C_Q^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})$ has a natural structure of Γ_Q-equivariant complex induced from that of $C^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})$. For a subgroup $\Gamma \subset \Gamma_Q$, we define

$$H_Q^q(Y^\circ, \Gamma, \mathscr{F}) := H^q(C_Q^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})^\Gamma)$$

to be the q-th cohomology group of the complex $C_Q^\bullet(\mathscr{X}_{\mathbb{Q}}, \mathscr{F})^\Gamma$.

Now, by taking the product of (7-3) over $I \in (X_{\mathbb{Q}})_Q^{q+1}$, we define

$$\int_Q : C_Q^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_0) \to C_Q^q(\mathcal{X}_{\mathbb{Q}}, \mathbb{C}), \quad (s_I)_{I \in (X_{\mathbb{Q}})_Q^{q+1}} \mapsto \left(\int_{Q,I} s_I \right)_{I \in (X_{\mathbb{Q}})_Q^{q+1}}.$$

Here $\mathbb{C}$ is regarded as a constant sheaf associated to $\mathbb{C}$ with the trivial Γ_Q-equivariant structure.

Proposition 7.2.6. *The map*

$$\int_Q : C_Q^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_0) \to C_Q^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathbb{C})$$

is a morphism of Γ_Q*-equivariant complexes, and hence induces a map*

$$\int_Q : H_Q^q(Y^\circ, \Gamma_Q, \mathscr{F}_0) \to H_Q^q(Y^\circ, \Gamma_Q, \mathbb{C})$$

for $q \geq 0$.

Proof. First we must show $\int_Q \circ \mathrm{d}^q = \mathrm{d}^q \circ \int_Q$ for $q \geq 0$. Let $J = (\alpha_0, \dots, \alpha_{q+1}) \in (X_{\mathbb{Q}})_Q^{q+2}$, and let $\xi_1, \dots, \xi_g \in V_J$ be a system of eigenvectors of ${}^tQ$ with eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ respectively. Then for $s = (s_I)_{I \in (X_{\mathbb{Q}})_Q^{q+1}} \in C_Q^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_0)$, we have

$$\begin{aligned} \left(\int_Q d^q(s) \right)_J = \int_{Q,J} (d^q(s))_J &= \int_{\sigma(\xi_1,\dots,\xi_g)} \sum_{i=0}^{q+1} (-1)^i s_{J^{(i)}}|_{V_J} \\ &= \sum_{i=0}^{q+1} (-1)^i \int_{\sigma(\xi_1,\dots,\xi_g)} s_{J^{(i)}} = \sum_{i=0}^{q+1} (-1)^i \left(\int_Q s \right)_{J^{(i)}} = \left(d^q \left(\int_Q s \right) \right)_J, \end{aligned}$$

where $J^{(i)} = (\alpha_0, \dots, \check{\alpha}_i, \dots, \alpha_{q+1})$.

Next we must show $\int_Q \circ [\gamma] = [\gamma] \circ \int_Q$ for $\gamma \in \Gamma_Q$. Let $J = (\alpha_1, \dots, \alpha_{q+1}) \in (X_{\mathbb{Q}})_Q^{q+1}$, and let again $\xi_1, \dots, \xi_g \in V_J$ be a system of eigenvectors of ${}^tQ$ with eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ respectively. Then as in the proof of Lemma 7.2.5, we see that ${}^t\gamma\xi_1, \dots, {}^t\gamma\xi_g$ are eigenvectors of ${}^tQ$ in $V_{\gamma^{-1}J}$ with eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ respectively. Therefore, for $s = (s_I)_{I \in (X_{\mathbb{Q}})_Q^{q+1}} \in C_Q^q(\mathcal{X}_{\mathbb{Q}}, \mathscr{F}_0)$, we have

$$\begin{aligned} \left(\int_Q [\gamma](s) \right)_J = \int_{Q,J} ([\gamma](s))_J &= \int_{\sigma(\xi_1,\dots,\xi_g)} s_{\gamma^{-1}J}({}^t\gamma y) \\ &= \int_{\sigma({}^t\gamma\xi_1,\dots,{}^t\gamma\xi_g)} s_{\gamma^{-1}J}(y) \\ &= \int_{Q,\gamma^{-1}J} s_{\gamma^{-1}J} = \left(\int_Q s \right)_{\gamma^{-1}J} = \left([\gamma] \left(\int_Q s \right) \right)_J. \end{aligned}$$

This completes the proof. □

Let $\int_Q$ also denote the composition

$$\int_Q : H^q(Y^\circ, \Gamma_Q, \mathscr{F}_0) \xrightarrow{p_Q} H_Q^q(Y^\circ, \Gamma_Q, \mathscr{F}_0) \xrightarrow{\int_Q} H_Q^q(Y^\circ, \Gamma_Q, \mathbb{C}), \tag{7-5}$$

where p_Q is the natural map induced from the projection p_Q in (7-4). See also Corollary 4.3.4.

8. Specialization to the zeta values

In this section we compute the group $H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C})$ explicitly, and show that we can get the values of the zeta function as a specialization of the Shintani–Barnes cocycle $[\Psi_d]$.

First we return to the setting in Section 2.2. Let

- $F/\mathbb{Q}$ be a number field of degree g,
- $\tau_1, \dots, \tau_g : F \hookrightarrow \mathbb{C}$ be the field embeddings of F into $\mathbb{C}$,
- $\mathcal{O} \subset F$ be an order,
- $\mathfrak{a} \subset F$ be a proper fractional $\mathcal{O}$-ideal,
- $w_1, \dots, w_g \in \mathfrak{a}$ be a basis of $\mathfrak{a}$ over $\mathbb{Z}$,
- $w := {}^t(w_1, \dots, w_g) \in F^g$, and $w^{(i)} := \tau_i(w) = {}^t(\tau_i(w_1), \dots, \tau_i(w_g)) \in \mathbb{C}^g$,
- $\rho_w : F \to M_g(\mathbb{Q})$ be the regular representation with respect to

$$w : \mathbb{Q}^g \xrightarrow{\sim} F, \quad x \mapsto \langle x, w\rangle,$$

- $N_w(x_1, \dots, x_g) \in \mathbb{Q}[x_1, \dots, x_g]$ be the norm polynomial with respect to w,
- $w_1^*, \dots, w_g^* \in F$ be the dual basis of $w_1, \dots, w_g$ with respect to the trace $\mathrm{Tr}_{F/\mathbb{Q}}$,
- $w^*, w^{*(i)}, N_{w^*}, \rho_{w^*}$ be the dual objects obtained from $w_1^*, \dots, w_g^*$.

Take $\theta \in F^\times$ such that $F = \mathbb{Q}(\theta)$ and put $Q := \rho_w(\theta) \in \Xi$. Also, set $\theta^{(1)} := \tau_1(\theta), \dots, \theta^{(g)} := \tau_g(\theta) \in \mathbb{C}^\times$ to be the eigenvalues of ${}^tQ$. We fix this notation.

8.1. ***Computation of $H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C})$.*** Define

$$T_w := \{x \in \mathbb{R}^g \mid N_w(x) \neq 0\} \subset \mathbb{R}^g - \{0\}$$

to be the set of real vectors whose norm is nonzero. By Lemma 2.2.1(7), it is clear that T_w is a Γ_Q-stable subset of $\mathbb{R}^g - \{0\}$. Note that under the isomorphism

$$w : \mathbb{R}^g \xrightarrow{\sim} F_{\mathbb{R}} := F \otimes_{\mathbb{Q}} \mathbb{R}, \quad x \mapsto \langle x, w\rangle,$$

T_w corresponds to $F_{\mathbb{R}}^\times = \{\alpha \in F_{\mathbb{R}} \mid N_{F/\mathbb{Q}}(\alpha) \neq 0\}$, i.e.,

$$w : T_w \xrightarrow{\sim} F_{\mathbb{R}}^\times. \tag{8-1}$$

The aim of this subsection is to obtain an isomorphism

$$H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C}) \xleftarrow{\sim} H^q(T_w/\Gamma_Q, \mathbb{C}) \simeq H^q(F_{\mathbb{R}}^\times/\mathcal{O}^1, \mathbb{C}), \tag{8-2}$$

where the last two cohomology groups are the usual singular cohomology groups.

As in Section 7, for $I = (\alpha_1, \dots, \alpha_{q+1}) \in (X_{\mathbb{Q}})^{q+1}$, let

$$\sigma_I : \Delta^q \to \mathbb{R}^g, \quad t = (t_1, \dots, t_{q+1}) \mapsto \sum_{i=1}^{q+1} \alpha_i t_i$$

denote the affine q-simplex with vertices $\alpha_1, \dots, \alpha_{q+1}$, and let $|\sigma_I| := \sigma_I(\Delta^q) \subset \mathbb{R}^g$ denote the image of σ_I. The following lemma enables us to compute the group $H^q_Q(Y^\circ, \Gamma_Q, \mathbb{C})$ using these simplices.

Lemma 8.1.1. *Let $q \geq 0$ and $I = (\alpha_1, \dots, \alpha_{q+1}) \in (X_{\mathbb{Q}})^{q+1}$. The following conditions are equivalent*:

(i) *I is Q-admissible.*

(ii) $|\sigma_I| \subset T_w$.

To prove this lemma, recall the following fact:

Lemma 8.1.2. *Let $A \subset \mathbb{C}$ be a convex compact subset. The following conditions are equivalent*:

(i) $0 \notin A$.

(ii) *There exists $\lambda \in \mathbb{C}^\times$ such that* $\mathrm{Re}(\lambda A) \subset \mathbb{R}_{>0}$.

Proof. This follows from [Rudin 1991, Theorem 3.4(b)]. □

Proof of Lemma 8.1.1. First, by Lemma 2.2.1(5), we know that $w^{(1)}, \dots, w^{(g)}$ are the eigenvectors of ${}^tQ$ with eigenvalues $\theta^{(1)}, \dots, \theta^{(g)}$ respectively. Therefore,

I is Q-admissible

$\iff$ for all $j \in \{1, \dots, g\}$ there exists $\lambda_j \in \mathbb{C}^\times$ such that $\lambda_j w^{(j)} \in V_I$

$\iff$ for all $j \in \{1, \dots, g\}$ there exists $\lambda_j \in \mathbb{C}^\times$ such that for all $i \in \{1, \dots, q+1\}$, $\mathrm{Re}(\langle \alpha_i, \lambda_j w^{(j)} \rangle) > 0$

$\iff$ for all $j \in \{1, \dots, g\}$ there exists $\lambda_j \in \mathbb{C}^\times$ such that $\mathrm{Re}(\lambda_j \langle |\sigma_I|, w^{(j)} \rangle) \subset \mathbb{R}_{>0}$

$\overset{*}{\iff}$ $0 \notin \langle |\sigma_I|, w^{(j)} \rangle$ for all $j \in \{1, \dots, g\}$

$\iff$ $N_w(x) \neq 0$ for all $x \in |\sigma_I|$

$\iff$ $|\sigma_I| \subset T_w$.

Note that the fourth equivalence $\overset{*}{\iff}$ follows from Lemma 8.1.2 since $\langle |\sigma_I|, w^{(i)} \rangle \subset \mathbb{C}$ is a convex compact subset. This proves the lemma. □

For $q \geq 0$, let $\Sigma_q := \{\sigma : \Delta^q \to T_w \text{ continuous}\}$ denote the set of singular q-simplices in T_w, and let

$$S_q := \mathbb{Z}[\Sigma_q]$$

denote the group of singular q-chains of T_w. For $j = 1, \dots, q+1$, let

$$\delta^q_j : \Delta^{q-1} \to \Delta^q, \quad (t_1, \dots, t_q) \mapsto (t_1, \dots, t_{j-1}, 0, t_j, \dots, t_q)$$

denote the j-th face map. Then we have a boundary map $\partial : S_q \to S_{q-1}$ which maps $\sigma \in \Sigma_q$ to

$$\partial\sigma = \sum_{j=1}^{q+1} (-1)^{j-1} \sigma \circ \delta_j^q \in S_{q-1}.$$

The action of Γ_Q on T_w naturally induces an action of Γ_Q on S_q, and we have a Γ_Q-equivariant singular chain complex $S_\bullet$. Moreover, let

$$K_q := \mathbb{Z}[(X_{\mathbb{Q}})_Q^{q+1}]$$

denote the free abelian group generated by $(X_{\mathbb{Q}})_Q^{q+1}$. By Lemma 7.2.5(2), we have a natural action of Γ_Q on K_q. Then, by Lemma 8.1.1, we have a natural injective homomorphism

$$K_q \hookrightarrow S_q, \quad I \mapsto \sigma_I,$$

which is clearly a Γ_Q-equivariant map. In the following, we identify K_q with a Γ_Q-submodule

$$\mathbb{Z}\big[\sigma_I \mid I \in (X_{\mathbb{Q}})_Q^{q+1}\big] = \mathbb{Z}\big[\sigma_I \mid I \in (X_{\mathbb{Q}})^{q+1}, |\sigma_I| \subset T_w\big] \subset S_q$$

of S_q via this injective map. Then, by Lemma 7.2.5(1), we see that the boundary map ∂ maps K_q to K_{q-1}, and hence $K_\bullet \subset S_\bullet$ becomes a Γ_Q-equivariant subcomplex of $S_\bullet$.

Note that we have a natural isomorphism

$$K_{\mathbb{C}}^\bullet := \mathrm{Hom}_{\mathbb{Z}}(K_\bullet, \mathbb{C}) \simeq \prod_{I \in (X_{\mathbb{Q}})_Q^{\bullet+1}} \mathbb{C} = C_Q^\bullet(\mathcal{X}_{\mathbb{Q}}, \mathbb{C})$$

of Γ_Q-equivariant complexes, and hence

$$H_Q^q(Y^\circ, \Gamma_Q, \mathbb{C}) \simeq H^q((K_{\mathbb{C}}^\bullet)^{\Gamma_Q}).$$

Therefore, in order to obtain (8-2), we compare $K_\bullet$ and $S_\bullet$.

Proposition 8.1.3. (1) *Let $\Gamma \subset \Gamma_Q$ be a subgroup. For $q \geq 0$, the quotient group S_q/K_q is an induced Γ-module.*

(2) *The inclusion map*

$$K_\bullet \hookrightarrow S_\bullet$$

is a quasi-isomorphism. In other words, the quotient complex $S_\bullet/K_\bullet$ is exact.

Proof. (1) This is clear since we have

$$S_q/K_q \simeq \mathbb{Z}[\sigma \in \Sigma_q \mid \sigma \notin K_q]$$

and $\Gamma \subset \Gamma_Q$ acts freely on the basis $\{\sigma \in \Sigma_q \mid \sigma \notin K_q\}$.

(2) This kind of fact may be well known to experts, but here we give a proof for the sake of completeness of the paper. First take any finite open covering

$$T_w = \bigcup_{k=1}^{N} U_k$$

of T_w such that U_k is a convex open subset of T_w for all k. The existence of such a covering can be easily seen from the identification $w : T_w \xrightarrow{\sim} F_{\mathbb{R}}^{\times}$.

We will prove that the quotient complex $S_\bullet/K_\bullet$ is exact. Let $q \geq 0$ and let $a \in S_q$ such that $\partial a \in K_{q-1}$. We need to show the following:

Aim. *There exist $\eta \in S_{q+1}$ and $b \in K_q$ such that $a = \partial\eta + b$.*

Suppose $a \in S_q$ is of the form

$$a = \sum_{i=1}^{r} c_i \sigma_i,$$

where σ_i are distinct singular q-simplices in T_w, and $c_i \in \mathbb{Z}$. By using the barycentric subdivision if necessary, without loss of generality we may assume

$$\text{for each } i \in \{1, \dots, r\} \text{ there exists } \kappa_i \in \{1, \dots, N\} \text{ such that } \sigma_i(\Delta^q) \subset U_{\kappa_i}. \tag{8-3}$$

Indeed, let

$$S : S_n \to S_n \quad \text{and} \quad T : S_n \to S_{n+1}$$

be the subdivision operator and the chain homotopy between S and id_{S_n} defined as in [Hatcher 2002, Section 2.1, Proof of Proposition 2.21]. Then taking into account the fact that the barycenter of any $\sigma_I \in K_n$ ($I \in (X_{\mathbb{Q}})_Q^{n+1}$) belongs to $\mathbb{Q}^g \cap |\sigma_I|$, we easily see that S (resp. T) maps K_n to K_n (resp. K_{n+1}). Hence we have

$$\partial S(a) = S(\partial a) \in K_{q-1} \quad \text{and} \quad a - S(a) = \partial T(a) + T(\partial a) \in \partial S_{q+1} + K_q.$$

Therefore, we can replace a with its (iterated) barycentric subdivision $S^m(a)$ (m sufficiently large) until we have (8-3).

We fix such κ_i for each $i = 1, \dots, r$.

Step 1: In order to "approximate" σ_i by the elements in K_q, we first approximate their vertices "simultaneously". For $i = 1, \dots, r$ and $j = 1, \dots, q+1$, let $v_{ij} \in U_{\kappa_i} \subset T_w$ denote the j-th vertex of σ_i, i.e.,

$$v_{ij} = \sigma_i(0, \dots, 0, \overset{j}{1}, 0, \dots, 0) \in T_w.$$

Then for $i = 1, \dots, r$ and $j = 1, \dots, q+1$, take $v'_{ij} \in U_{\kappa_i} \cap \mathbb{Q}^g$ satisfying the following conditions:

(V1) If $v_{ij} \in \mathbb{Q}^g$, then $v'_{ij} = v_{ij}$.

(V2) If $v_{ij} = v_{mn}$ for some $i, m \in \{1, \dots, r\}$ and $j, n \in \{1, \dots, q+1\}$, then $v'_{ij} = v'_{mn}$. (In other words, if the j-th vertex of σ_i and the n-th vertex of σ_m are the same, then v'_{ij} and v'_{mn} are the same as well.)

This is possible because $\mathbb{Q}^g$ is dense in $\mathbb{R}^g$. Then set

$$I_i := (v'_{i1}, \dots, v'_{i,q+1}) \in (X_{\mathbb{Q}})^{q+1} \quad \text{for } i = 1, \dots, r \qquad \text{and} \qquad a' := \sum_{i=1}^{r} c_i \sigma_{I_i}.$$

Since U_{κ_i} is convex, we have $\sigma_{I_i} \subset U_{\kappa_i} \subset T_w$, and hence $\sigma_{I_i} \in K_q$. Therefore, we see that $a' \in K_q$.

Now, recall that for $j = 1, \dots, q+1$,

$$\delta^q_j : \Delta^{q-1} \to \Delta^q, \quad (t_1, \dots, t_q) \mapsto (t_1, \dots, t_{j-1}, 0, t_j, \dots, t_q)$$

denotes the j-th face map. Then, by the conditions (V1) and (V2), we have the following:

(F1) If $\sigma_i \circ \delta^q_j \in K_{q-1}$, then $\sigma_{I_i} \circ \delta^q_j = \sigma_i \circ \delta^q_j$.

(F2) If $\sigma_i \circ \delta^q_j = \sigma_m \circ \delta^q_n$ for $i, m \in \{1, \dots, r\}$ and $j, n \in \{1, \dots, q+1\}$, then $\sigma_{I_i} \circ \delta^q_j = \sigma_{I_m} \circ \delta^q_n$. (In other words, if the j-th face of σ_i and the n-th face of σ_m are the same, then the j-th face of σ_{I_i} and the n-th face of σ_{I_m} are the same as well.)

Step 2: Next we consider the homotopy between a and a'. For $i = 1, \dots, r$, let

$$h_i : [0, 1] \times \Delta^q \to T_w, \quad (u, t) \mapsto u\sigma_i(t) + (1-u)\sigma_{I_i}(t)$$

be a homotopy between σ_i and σ_{I_i}. Note that since U_{κ_i} is convex, we have

$$h_i([0, 1] \times \Delta^q) \subset U_{\kappa_i}.$$

The homotopy h_i defines a $(q+1)$-chain $\eta_i \in S_{q+1}$in a usual way using the standard decomposition of the prism $[0, 1] \times \Delta^q$. More precisely, for $j = 1, \dots, q+1$, put

$$\epsilon^q_j : \Delta^{q+1} \to [0, 1] \times \Delta^q, \quad (t_1, \dots, t_{q+2}) \mapsto \Bigg(\sum_{m \geq j+1} t_m, (t_1, \dots, t_{j-1}, t_j + t_{j+1}, t_{j+2}, \dots, t_{q+2})\Bigg).$$

Using these maps, the $(q+1)$-chain $\eta_i \in S_{q+1}$ is defined as

$$\eta_i := \sum_{j=1}^{q+1} (-1)^{j-1} h_i \circ \epsilon^q_j.$$

Set $\eta := \sum_{i=1}^r c_i \eta_i \in S_{q+1}$.

Step 3: Now we examine the assumption $\partial a \in K_{q-1}$. First, we have

$$\partial a = \sum_{i=1}^r \sum_{j=1}^{q+1} (-1)^{j-1} c_i \sigma_i \circ \delta^q_j.$$

For each singular $(q-1)$-simplex $\sigma \in \Sigma_{q-1}$, set

$$C_\sigma := \sum_{\substack{i=1,\dots,r,\\ j=1,\dots,q+1,\\ \sigma_i \circ \delta^q_j = \sigma}} (-1)^{j-1} c_i \in \mathbb{Z}.$$

In the case where the index set of the sum is empty, we set $C_\sigma = 0$ by convention. Then we can rewrite ∂a as

$$\partial a = \sum_{\sigma \in \Sigma_{q-1}} C_\sigma \sigma.$$

Then, by the assumption $\partial a \in K_{q-1}$, we find that $C_\sigma = 0$ for all $\sigma \notin K_{q-1}$ since the set Σ_{q-1} of singular $(q-1)$-simplices is a basis of S_{q-1}.

Step 4: Next we compute the boundary of the homotopy $\eta \in S_{q+1}$. By an elementary computation we see

$$\partial \eta_i = \sigma_i - \sigma_{I_i} - \sum_{j=1}^{q+1} \sum_{m=1}^{q} (-1)^{j+m} h_{ij} \circ \epsilon_m^{q-1},$$

where

$$h_{ij} : [0, 1] \times \Delta^{q-1} \to T_w, \quad (u, t) \mapsto u\sigma_i \circ \delta_j^q(t) + (1-u)\sigma_{I_i} \circ \delta_j^q(t)$$

is a homotopy between $\sigma_i \circ \delta_j^q$ and $\sigma_{I_i} \circ \delta_j^q$; see [Hatcher 2002, Section 2.1, Proof of 2.10].

Now, by the properties (F1) and (F2), we see the following:

(H1) If $\sigma_i \circ \delta_j^q \in K_{q-1}$, then $h_{ij}(u, t) = \sigma_i \circ \delta_j^q(t)$ for $(u, t) \in [0, 1] \times \Delta^{q-1}$.

(H2) If $\sigma_i \circ \delta_j^q = \sigma_m \circ \delta_n^q$ for $i, m \in \{1, \dots, r\}$ and $j, n \in \{1, \dots, q+1\}$, then $h_{ij} = h_{mn}$.

Then for each singular $(q-1)$-simplex $\sigma \in \Sigma_{q-1}$, we define a map

$$h_\sigma : [0, 1] \times \Delta^{q-1} \to T_w$$

as follows: If σ is of the form $\sigma = \sigma_i \circ \delta_j^q$ for some $i \in \{1, \dots, r\}$ and $j \in \{1, \dots, q+1\}$, we set $h_\sigma := h_{ij}$. This is well defined by the property (H2). If σ is not of the form $\sigma_i \circ \delta_j^q$, then simply set $h_\sigma(u, t) := \sigma(t)$ for $(u, t) \in [0, 1] \times \Delta^{q-1}$.

Then we find

$$\begin{aligned}
\partial \eta &= a - a' - \sum_{i=1}^{r} \sum_{j=1}^{q+1} \sum_{m=1}^{q} (-1)^{j+m} c_i h_{ij} \circ \epsilon_m^{q-1} \\
&= a - a' - \sum_{i=1}^{r} \sum_{j=1}^{q+1} \sum_{m=1}^{q} (-1)^{j+m} c_i h_{\sigma_i \circ \delta_j^q} \circ \epsilon_m^{q-1} \\
&= a - a' - \sum_{\sigma \in \Sigma_{q-1}} \sum_{m=1}^{q} (-1)^{m-1} h_\sigma \circ \epsilon_m^{q-1} \sum_{\substack{i=1,\dots,r, \\ j=1,\dots,q+1, \\ \sigma_i \circ \delta_j^q = \sigma}} (-1)^{j-1} c_i \\
&= a - a' - \sum_{\sigma \in \Sigma_{q-1}} C_\sigma \sum_{m=1}^{q} (-1)^{m-1} h_\sigma \circ \epsilon_m^{q-1} \\
&= a - a' - \sum_{\sigma \in \Sigma_{q-1} \cap K_{q-1}} C_\sigma \sum_{m=1}^{q} (-1)^{m-1} h_\sigma \circ \epsilon_m^{q-1}.
\end{aligned}$$

Note that the last equality holds since we have $C_\sigma = 0$ for $\sigma \notin K_{q-1}$. Moreover, by the property (H1), we easily see that if $\sigma = \sigma_i \circ \delta_j^q \in K_{q-1}$, then $h_\sigma \circ \epsilon_m^{q-1} \in K_q$ for all $m = 1, \dots, q$. Therefore, by setting

$$b := a' + \sum_{\sigma \in \Sigma_{q-1} \cap K_{q-1}} C_\sigma \sum_{m=1}^{q} (-1)^{m-1} h_\sigma \circ \epsilon_m^{q-1} \in K_q,$$

we obtain the desired identity $a = \partial \eta + b$. □

Let $S^\bullet_{\mathbb{C}} := \mathrm{Hom}_{\mathbb{Z}}(S_\bullet, \mathbb{C})$ denote the singular cochain complex of T_w with coefficients in $\mathbb{C}$.

Corollary 8.1.4. *Let $\Gamma \subset \Gamma_Q$ be a subgroup.*

(1) *The map $K_\bullet \hookrightarrow S_\bullet$ induces a quasi-isomorphism*

$$(K_\bullet)_\Gamma \to (S_\bullet)_\Gamma,$$

where $(-)_\Gamma$ denotes the Γ-coinvariant part. In particular, we obtain an isomorphism

$$H_q((K_\bullet)_\Gamma) \xrightarrow{\sim} H_q(T_w/\Gamma, \mathbb{Z}).$$

(2) *The map $K_\bullet \hookrightarrow S_\bullet$ induces a quasi-isomorphism*

$$(S^\bullet_{\mathbb{C}})^\Gamma \to (K^\bullet_{\mathbb{C}})^\Gamma.$$

In particular, we obtain an isomorphism

$$H^q(T_w/\Gamma, \mathbb{C}) \xrightarrow{\sim} H^q((K^\bullet_{\mathbb{C}})^\Gamma) \simeq H^q_Q(Y^\circ, \Gamma, \mathbb{C}).$$

Proof. First note that since the action of Γ_Q on T_w is free and properly discontinuous, the singular homology $H_q(T_w/\Gamma, \mathbb{Z})$ (resp. singular cohomology $H^q(T_w/\Gamma, \mathbb{C})$) can be computed by the equivariant singular homology (resp. equivariant singular cohomology), i.e., we have

$$H_q(T_w/\Gamma, \mathbb{Z}) \simeq H_q((S_\bullet)_\Gamma) \quad \text{and} \quad H^q(T_w/\Gamma, \mathbb{C}) \simeq H^q((S^\bullet_{\mathbb{C}})^\Gamma).$$

See [Cartan and Eilenberg 1956, Chapter XVI, Section 9].

(1) We consider the tautological exact sequence

$$0 \to K_q \to S_q \to S_q/K_q \to 0. \tag{8-4}$$

By Proposition 8.1.3(1), we obtain a short exact sequence

$$0 = H_1(\Gamma, S_q/K_q) \to (K_q)_\Gamma \to (S_q)_\Gamma \to (S_q/K_q)_\Gamma \to 0,$$

where $H_1(\Gamma, -)$ is the first group homology of Γ. This induces a long exact sequence

$$\cdots \to H_{q+1}((S_\bullet/K_\bullet)_\Gamma) \to H_q((K_\bullet)_\Gamma) \to H_q((S_\bullet)_\Gamma) \to H_q((S_\bullet/K_\bullet)_\Gamma) \to \cdots.$$

Therefore, it remains to show

$$H_q((S_\bullet/K_\bullet)_\Gamma) = 0$$

for $q \geq 0$. Indeed, by Proposition 8.1.3, we see that

$$\cdots \to S_2/K_2 \to S_1/K_1 \to S_0/K_0 \to 0 \tag{8-5}$$

is an exact sequence of induced Γ-modules. Therefore, (8-5) can be seen as a $(-)_\Gamma$-acyclic resolution of 0. Thus we see

$$H_q((S_\bullet/K_\bullet)_\Gamma) = H_q(\Gamma, 0) = 0$$

for all $q \geq 0$.

(2) This can be proved similarly. By applying $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{C})$ to (8-4), we obtain a short exact sequence

$$0 \to (S_q/K_q)^\vee_{\mathbb{C}} \to S^q_{\mathbb{C}} \to K^q_{\mathbb{C}} \to 0,$$

where $(S_q/K_q)^\vee_{\mathbb{C}} := \mathrm{Hom}_{\mathbb{Z}}(S_q/K_q, \mathbb{C})$. Then, by Proposition 8.1.3(1), we see that $(S_q/K_q)^\vee_{\mathbb{C}}$ is a coinduced Γ-module, and hence we obtain another short exact sequence

$$0 \to ((S_q/K_q)^\vee_{\mathbb{C}})^\Gamma \to (S^q_{\mathbb{C}})^\Gamma \to (K^q_{\mathbb{C}})^\Gamma \to H^1(\Gamma, (S_q/K_q)^\vee_{\mathbb{C}}) = 0.$$

Furthermore, this exact sequence induces a long exact sequence

$$\cdots \to H^q\big(((S_\bullet/K_\bullet)^\vee_{\mathbb{C}})^\Gamma\big) \to H^q((S^\bullet_{\mathbb{C}})^\Gamma) \to H^q((K^\bullet_{\mathbb{C}})^\Gamma) \to H^{q+1}\big(((S_\bullet/K_\bullet)^\vee_{\mathbb{C}})^\Gamma\big) \to \cdots.$$

Therefore, it remains to show that

$$H^q\big(((S_\bullet/K_\bullet)^\vee_{\mathbb{C}})^\Gamma\big) = 0$$

for $q \geq 0$. Indeed, by applying $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{C})$ to (8-5), we see that

$$0 \to (S_0/K_0)^\vee_{\mathbb{C}} \to (S_1/K_1)^\vee_{\mathbb{C}} \to (S_2/K_2)^\vee_{\mathbb{C}} \to \cdots$$

is a $(-)^\Gamma$-acyclic resolution of 0, and hence

$$H^q\big(((S_\bullet/K_\bullet)^\vee_{\mathbb{C}})^\Gamma\big) \simeq H^q(\Gamma, 0) = 0$$

for all $q \geq 0$. □

As a result, for a subgroup $\Gamma \subset \Gamma_Q$ and a homology class $\mathfrak{z} \in H_{g-1}(T_w/\Gamma, \mathbb{Z})$, we can define an evaluation map

$$\langle \mathfrak{z}, \ \rangle : H^{g-1}_Q(Y^\circ, \Gamma_Q, \mathbb{C}) \simeq H^{g-1}(T_w/\Gamma_Q, \mathbb{C}) \to H^{g-1}(T_w/\Gamma, \mathbb{C}) \xrightarrow{\langle \mathfrak{z}, \ \rangle} \mathbb{C} \tag{8-6}$$

by taking the pairing with $\mathfrak{z}$.

8.2. *Shintani decomposition.* Using Corollary 8.1.4, here we construct a cone decomposition of a homology class $\mathfrak{z} \in H_{g-1}(T_w/\Gamma, \mathbb{Z})$. See Proposition 8.2.1 and Remark 8.2.2. We need such a cone decomposition in order to compute the specialization of the Shintani–Barnes cocycle.

Recall that $\tau_1, \dots, \tau_g$ are the field embeddings of F into $\mathbb{C}$. Clearly, τ_i extends to

$$\tau_i : F_{\mathbb{R}} = F \otimes \mathbb{R} \to \mathbb{C}.$$

Let F_{τ_i} denote the completion of F with respect to the embedding τ_i. In the following, we assume for simplicity that $\tau_1, \dots, \tau_{r_1}$ are the real embeddings, i.e., $F_{\tau_i} = \mathbb{R}$ for $i = 1, \dots, r_1$, and $\tau_{r_1+1}, \dots, \tau_g$ are the nonreal embeddings, i.e., $F_{\tau_i} = \mathbb{C}$ for $i = r_1 + 1, \dots, g$.

For $\mu = (\mu_1, \dots, \mu_{r_1}) \in \{\pm 1\}^{r_1}$ $(:= \{-1, 1\}^{r_1})$, set

$$F^\times_{\mathbb{R},\mu} := \{x \in F^\times_{\mathbb{R}} \mid \mu_i \tau_i(x) > 0 \ \text{ for } \ i = 1, \dots, r_1\}.$$

Clearly, $\{F^\times_{\mathbb{R},\mu} \mid \mu \in \{\pm 1\}^{r_1}\}$ are the connected components of $F^\times_{\mathbb{R}}$, and we have $F^\times_{\mathbb{R}} = \bigsqcup_{\mu\in\{\pm1\}^{r_1}} F^\times_{\mathbb{R},\mu}$. Then let $T_{w,\mu} \subset T_w$ be the connected component of T_w corresponding to $F^\times_{\mathbb{R},\mu}$ via the identification (8-1):

$$w : T_w \xrightarrow{\sim} F^\times_{\mathbb{R}}.$$

If $\mu = (1, 1, \dots, 1)$, then $F^\times_{\mathbb{R},\mu}$ is the totally positive component of $F^\times_{\mathbb{R}}$, and simply denoted by $F^\times_{\mathbb{R},+}$. Furthermore, let

$$F^\times_+ := F^\times \cap F^\times_{\mathbb{R},+} = \{x \in F^\times \mid \tau_i(x) > 0 \text{ for } i = 1, \dots, r_1\},$$
$$\mathcal{O}^\times_+ := \mathcal{O}^\times \cap F^\times_{\mathbb{R},+} = \{u \in \mathcal{O}^\times \mid \tau_i(u) > 0 \text{ for } i = 1, \dots, r_1\}$$

denote the totally positive parts of $F^\times$ and $\mathcal{O}^\times$ respectively, and let $\Gamma^+_Q \subset \Gamma_Q$ be the image of $\mathcal{O}^\times_+$ under the isomorphism

$$\rho_w : \mathcal{O}^1 \xrightarrow{\sim} \Gamma_Q$$

(see Section 2.2).

By Dirichlet's unit theorem, we know that

$$T_w/\mathbb{R}_{>0}\Gamma^+_Q \simeq F^\times_{\mathbb{R}}/\mathbb{R}_{>0}\mathcal{O}^\times_+$$

is compact, and its connected components

$$T_{w,\mu}/\mathbb{R}_{>0}\Gamma^+_Q \simeq F^\times_{\mathbb{R},\mu}/\mathbb{R}_{>0}\mathcal{O}^\times_+ \quad \text{for } \mu \in \{\pm 1\}^{r_1}$$

are homeomorphic to $(g-1)$-dimensional topological tori. Therefore, we have

$$H_{g-1}(T_w/\Gamma^+_Q, \mathbb{Z}) \simeq H_{g-1}(T_w/\mathbb{R}_{>0}\Gamma^+_Q, \mathbb{Z}) \simeq \mathbb{Z}^{\{\pm1\}^{r_1}}. \tag{8-7}$$

Here the first isomorphism is a canonical isomorphism induced from the projection

$$T_w/\Gamma^+_Q \to T_w/\mathbb{R}_{>0}\Gamma^+_Q,$$

which is clearly a homotopy equivalence. In order to fix the second isomorphism of (8-7), we equip $T_w/\mathbb{R}_{>0}\Gamma^+_Q$ with an orientation as follows.

Orientation. Set

$$\boldsymbol{T}_\mu := T_{w,\mu}/\mathbb{R}_{>0}\Gamma^+_Q \subset \boldsymbol{T} := T_w/\mathbb{R}_{>0}\Gamma^+_Q$$

for simplicity. Recall that an orientation of a $(g-1)$-dimensional manifold X is defined as a system $(\nu_x)_{x\in X}$ of generators $\nu_x \in H_{g-1}(X, X-\{x\}, \mathbb{Z}) \simeq \mathbb{Z}$ with a certain compatibility; see [Hatcher 2002, Section 3.3]. Note that giving a generator ν_x of $H_{g-1}(X, X-\{x\}, \mathbb{Z}) \simeq \mathbb{Z}$ is equivalent to giving an isomorphism

$$o_x : H_{g-1}(X, X-\{x\}, \mathbb{Z}) \xrightarrow{\sim} \mathbb{Z}, \qquad \nu_x \mapsto 1.$$

We first fix an orientation of the $(g-1)$-sphere $\boldsymbol{S}^{g-1} = (\mathbb{R}^g - \{0\})/\mathbb{R}_{>0}$ as follows. Let $x \in \mathbb{R}^g - \{0\}$ and let $\bar{x} \in \boldsymbol{S}^{g-1}$ be its image. Moreover, let $I = (\alpha_1, \dots, \alpha_g) \in (X_{\mathbb{Q}})^g$ such that $0 \notin |\sigma_I|$ and $x \notin \partial C_I$, where ∂C_I is the boundary of the cone C_I. Then we see

$$\bar{\sigma}_I : \Delta^{g-1} \xrightarrow{\sigma_I} \mathbb{R}^g - \{0\} \to \boldsymbol{S}^{g-1}$$

defines a class $[\bar{\sigma}_I] \in H_{g-1}(\boldsymbol{S}^{g-1}, \boldsymbol{S}^{g-1} - \{\bar{x}\}, \mathbb{Z})$. We fix the isomorphism $o_{\bar{x}}$ so that we have

$$o_{\bar{x}}([\bar{\sigma}_I]) = \operatorname{sgn}(I)\mathbf{1}_{C_I}(x)$$

for all such I, where $\operatorname{sgn}(I) = \operatorname{sgn}(\det I) \in \{0, \pm 1\}$. This defines an orientation of $\boldsymbol{S}^{g-1}$. Then this orientation of $\boldsymbol{S}^{g-1}$ induces orientations of $T_w/\mathbb{R}_{>0} \subset \boldsymbol{S}^{g-1}$ and $\boldsymbol{T} = T_w/\mathbb{R}_{>0}\Gamma_Q^+$ because the action of Γ_Q^+ on $T_w/\mathbb{R}_{>0}$ is free, properly discontinuous, and orientation-preserving. More explicitly, for $x \in T_w$ and its image $\boldsymbol{x} \in \boldsymbol{T}$, the local orientation isomorphism

$$o_{\boldsymbol{x}} : H_{g-1}(\boldsymbol{T}, \boldsymbol{T} - \{\boldsymbol{x}\}, \mathbb{Z}) \xrightarrow{\sim} \mathbb{Z}$$

can be computed as follows. let $I = (\alpha_1, \ldots, \alpha_g) \in (X_{\mathbb{Q}})_Q^g$ such that $\gamma x \notin \partial C_I$ for all $\gamma \in \Gamma_Q^+$. Then

$$\boldsymbol{\sigma}_I : \Delta^{g-1} \xrightarrow{\sigma_I} T_w \to \boldsymbol{T}$$

defines a class $[\boldsymbol{\sigma}_I] \in H_{g-1}(\boldsymbol{T}, \boldsymbol{T} - \{\boldsymbol{x}\}, \mathbb{Z})$, and we have

$$o_{\boldsymbol{x}}([\boldsymbol{\sigma}_I]) = \operatorname{sgn}(I) \sum_{\gamma \in \Gamma_Q^+} \mathbf{1}_{C_I}(\gamma x). \tag{8-8}$$

Now, since $\{\boldsymbol{T}_\mu \mid \mu \in \{\pm 1\}^{r_1}\}$ are the connected components of $\boldsymbol{T}$, this orientation defines isomorphisms

$$o_\mu : H_{g-1}(\boldsymbol{T}_\mu, \mathbb{Z}) \xrightarrow{\sim} \mathbb{Z}, \quad \mu \in \{\pm 1\}^{r_1},$$
$$o = \bigoplus_\mu o_\mu : H_{g-1}(\boldsymbol{T}, \mathbb{Z}) \simeq \bigoplus_{\mu \in \{\pm 1\}^{r_1}} H_{g-1}(\boldsymbol{T}_\mu, \mathbb{Z}) \xrightarrow{\sim} \bigoplus_{\mu \in \{\pm 1\}^{r_1}} \mathbb{Z}$$

such that for all $\boldsymbol{x} \in \boldsymbol{T}_\mu$, the following diagram is commutative:

$$\begin{tikzcd}
H_{g-1}(\boldsymbol{T}_\mu, \mathbb{Z}) \arrow[r, "o_\mu", "\sim"'] \arrow[d, "\mathrm{loc}_{\boldsymbol{x}}"'] & \mathbb{Z} \arrow[d, equal] \\
H_{g-1}(\boldsymbol{T}, \boldsymbol{T} - \{\boldsymbol{x}\}, \mathbb{Z}) \arrow[r, "o_{\boldsymbol{x}}", "\sim"'] & \mathbb{Z}
\end{tikzcd} \tag{8-9}$$

Here the left vertical arrow is the natural localization map; see [Hatcher 2002, Theorem 3.26, Lemma 3.27].

For $\chi = (\chi_\mu)_\mu \in \bigoplus_{\mu \in \{\pm 1\}^{r_1}} \mathbb{Z}$, let

$$\mathfrak{z}_\chi \in H_{g-1}(\boldsymbol{T}, \mathbb{Z}) \simeq H_{g-1}(T_w/\,\Gamma_Q^+, \mathbb{Z})$$

denote the class such that $o(\mathfrak{z}_\chi) = \chi$. Note that if $\mathfrak{z}_\mu$ denotes the fundamental class of $\boldsymbol{T}_\mu$, then $\mathfrak{z}_\chi$ can be written as $\mathfrak{z}_\chi = \sum_\mu \chi_\mu \mathfrak{z}_\mu$.

Proposition 8.2.1. *Let $\chi = (\chi_\mu)_\mu \in \bigoplus_{\mu \in \{\pm 1\}^{r_1}} \mathbb{Z}$.*

(1) *There exists*

$$\Phi = \sum_{i=1}^r c_i \sigma_{I_i} \in K_{g-1} = \mathbb{Z}[\sigma_I \mid I \in (X_{\mathbb{Q}})_Q^g] \subset S_{g-1}$$

which represents the homology class $\mathfrak{z}_\chi \in H_{g-1}(T_w/\,\Gamma_Q^+, \mathbb{Z})$, where $I_1, \ldots, I_r \in (X_{\mathbb{Q}})_Q^g$, and $c_i \in \mathbb{Z}$.

(2) *Then for* $x \in \mathbb{R}^g - \{0\}$, *we have*

$$\sum_{\gamma \in \Gamma_Q^+} \sum_{i=1}^{r} c_i \operatorname{sgn}(I_i) \mathbf{1}_{C_{I_i}^Q}(\gamma x) = \chi(x) \mathbf{1}_{T_w}(x),$$

where χ *is regarded as a locally constant function* $\chi : T_w \to \mathbb{Z}$ *which has value* χ_μ *on* $T_{w,\mu}$, *i.e.,* $\chi(x) = \chi_\mu$ *for* $x \in T_{w,\mu}$.

Proof. (1) This is a direct consequence of Corollary 8.1.4(1).

(2) First note that we have

$$\mathbf{1}_{C_{I_i}^Q}(\gamma x) = \mathbf{1}_{\gamma^{-1} C_{I_i}^Q}(x) = \mathbf{1}_{C_{\gamma^{-1} I_i}^Q}(x)$$

for $\gamma \in \Gamma_Q^+$. Now, since the action of Γ_Q^+ on $T_w/\mathbb{R}_{>0}$ is properly discontinuous, the collection $\{\gamma^{-1} C_{I_i}\}_{i,\gamma}$ of subsets of T_w is locally finite. Therefore, as in the proof of Proposition 5.3.4, by using Lemma 5.3.3, we can find $\delta > 0$ such that

$$\exp(\varepsilon Q) x \notin \partial C_{\gamma^{-1} I_i}$$

for all $\varepsilon \in (0, 2\delta)$, $i = 1, \dots, r$, and $\gamma \in \Gamma_Q^+$. Set

$$x' := \exp(\delta Q) x.$$

Then we have

$$\mathbf{1}_{C_{I_i}^Q}(\gamma x) = \mathbf{1}_{C_{\gamma^{-1} I_i}^Q}(x) = \mathbf{1}_{C_{\gamma^{-1} I_i}}(x') = \mathbf{1}_{C_{I_i}}(\gamma x').$$

Moreover, by using Lemma 2.2.1(5), we see that $\exp(\delta Q)$ preserves the connected components $T_{w,\mu}$ of T_w, and hence we have

$$\chi(x) \mathbf{1}_{T_w}(x) = \chi(x') \mathbf{1}_{T_w}(x').$$

Therefore, it suffices to show

$$\sum_{\gamma \in \Gamma_Q^+} \sum_{i=1}^{r} c_i \operatorname{sgn}(I_i) \mathbf{1}_{C_{I_i}}(\gamma x') = \chi(x') \mathbf{1}_{T_w}(x'). \tag{8-10}$$

First, by Lemma 8.1.1, all of the terms in (8-10) are 0 if $x' \notin T_w$. Therefore, we assume $x' \in T_{w,\mu}$ for some $\mu \in \{\pm 1\}^{r_1}$. Set

$$\boldsymbol{T}_\mu := T_{w,\mu}/\mathbb{R}_{>0}\Gamma_Q^+ \quad \text{and} \quad \boldsymbol{x}' := \mathbb{R}_{>0}\Gamma_Q^+ x' \in \boldsymbol{T}_\mu.$$

Then, by (8-8), we see that the image of Φ under the localization map

$$o_{\boldsymbol{x}'} \circ \mathrm{loc}_{\boldsymbol{x}'} : H_{g-1}(T_{w,\mu}/\Gamma_Q^+, \mathbb{Z}) \simeq H_{g-1}(\boldsymbol{T}_\mu, \mathbb{Z}) \xrightarrow{\mathrm{loc}_{\boldsymbol{x}'}} H_{g-1}(\boldsymbol{T}, \boldsymbol{T} - \{\boldsymbol{x}'\}, \mathbb{Z}) \xrightarrow[\sim]{o_{\boldsymbol{x}'}} \mathbb{Z}$$

is equal to

$$\sum_{\gamma \in \Gamma_Q^+} \sum_{i=1}^{r} c_i \operatorname{sgn}(I_i) \mathbf{1}_{C_{I_i}}(\gamma x').$$

On the other hand, by (8-9), $o_{\boldsymbol{x}'} \circ \mathrm{loc}_{\boldsymbol{x}'}(\Phi) = o_\mu(\mathfrak{z}_\chi) = \chi_\mu$ because Φ represents $\mathfrak{z}_\chi$. This completes the proof. $\square$

Remark 8.2.2. In the case where $\mathfrak{z}_\chi = \mathfrak{z}_\mu$ is the fundamental class of a connected component $\boldsymbol{T}_\mu$, Proposition 8.2.1 says that

$$\sum_{i=1}^{r} c_i \operatorname{sgn}(I_i) \mathbf{1}_{C^Q_{I_i}}$$

gives a signed fundamental domain for $T_{w,\mu}/\Gamma_Q^+$ in the sense of Charollois, Dasgupta, and Greenberg [Charollois et al. 2015, Definition 2.4], which is a "weighted version" of the Shintani cone decomposition; see also [Diaz y Diaz and Friedman 2014; Espinoza and Friedman 2020].

Remark 8.2.3. Let the notation χ, $\mathfrak{z}_\chi$, and $\Phi = \sum_{i=1}^{r} c_i \sigma_{I_i}$ be the same as in Proposition 8.2.1. We can compute the evaluation map

$$\langle \mathfrak{z}_\chi, \ \rangle : H_Q^{g-1}(Y^\circ, \Gamma_Q, \mathbb{C}) \simeq H^{g-1}(T_w/\Gamma_Q, \mathbb{C}) \to H^{g-1}(T_w/\Gamma_Q^+, \mathbb{C}) \xrightarrow{\langle \mathfrak{z}_\chi, \ \rangle} \mathbb{C}$$

(see (8-6)) explicitly as follows. Let

$$s = (s_I)_{I \in (X_\mathbb{Q})^g_Q} \in C_Q^{g-1}(\mathcal{X}_\mathbb{Q}, \mathbb{C}) = \prod_{I \in (X_\mathbb{Q})^g_Q} \mathbb{C}$$

be a Γ_Q-invariant cocycle and let $[s] \in H_Q^{g-1}(Y^\circ, \Gamma_Q, \mathbb{C})$ be the class represented by s. Then we have

$$\langle \mathfrak{z}_\chi, [s] \rangle = \sum_{i=1}^{r} c_i s_{I_i}.$$

8.3. ***Values of the zeta functions.*** Recall that F is a number field of degree g, $\mathcal{O}$ is an order in F, and $\mathfrak{a} \subset F$ is a proper fractional $\mathcal{O}$-ideal.

Definition 8.3.1. (1) For a continuous map

$$\chi : F_\mathbb{R}^\times = (F \otimes_\mathbb{Q} \mathbb{R})^\times \to \mathbb{Z},$$

let

$$\zeta_\mathcal{O}(\chi, \mathfrak{a}^{-1}, s) := \sum_{x \in (\mathfrak{a} - \{0\})/\mathcal{O}_+^\times} \frac{\chi(x)}{|N_{F/\mathbb{Q}}(x)|^s}, \qquad \operatorname{Re}(s) > 1$$

denote the partial zeta function associated to χ and a proper fractional $\mathcal{O}$-ideal $\mathfrak{a}^{-1}$. Here, note that χ is constant on each connected component of $F_\mathbb{R}^\times$, and thus invariant under the action of $\mathcal{O}_+^\times$.

(2) Let

$$\boldsymbol{\varepsilon} : F_\mathbb{R}^\times \to \{\pm 1\}, \qquad x \mapsto \frac{N_{F/\mathbb{Q}}(x)}{|N_{F/\mathbb{Q}}(x)|}$$

denote the sign character.

Now, let $k \geq 1$, and let $\chi \in \bigoplus_{\mu \in \{\pm 1\}^{r_1}} \mathbb{Z}$. Note that χ can be regarded as a continuous map $\chi : F_\mathbb{R}^\times \to \mathbb{Z}$ via

$$\chi : F_\mathbb{R}^\times \to F_\mathbb{R}^\times / F_{\mathbb{R},+}^\times \simeq \{\pm 1\}^{r_1} \xrightarrow{\chi} \mathbb{Z}.$$

So far, we have defined the following series of maps between cohomology groups:

$$
\begin{tikzcd}
H^{g-1}(Y^\circ, \mathrm{SL}_g(\mathbb{Z}), \mathscr{F}_{kg}^{\Xi}) \arrow[d, "\mathrm{ev}_Q"] & \ni & {[\Psi_{kg}]} \arrow[dddd, mapsto] \\
H^{g-1}(Y^\circ, \Gamma_Q, \mathscr{F}_{kg}) \arrow[d, "N_{w^*}^k"] & & \\
H^{g-1}(Y^\circ, \Gamma_Q, \mathscr{F}_{0}) \arrow[d, "\int_Q"] & & \\
H_Q^{g-1}(Y^\circ, \Gamma_Q, \mathbb{C}) \arrow[d, "\langle \mathfrak{z}_\chi ,\ \rangle"] & & \\
\mathbb{C} & \ni & \left\langle \mathfrak{z}_\chi, \int_Q N_{w^*}^k \mathrm{ev}_Q([\Psi_{kg}])\right\rangle
\end{tikzcd}
\tag{8-11}
$$

See Corollary 4.3.4, Example 4.3.6, (7-5), and Remark 8.2.3 for the definitions of these maps.

Theorem 8.3.2. *We have*

$$
\left\langle \mathfrak{z}_\chi, \int_Q N_{w^*}^k \mathrm{ev}_Q([\Psi_{kg}])\right\rangle = \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \zeta_{\mathcal{O}}(\boldsymbol{\varepsilon}^{k+1}\chi, \mathfrak{a}^{-1}, k+1),
$$

where $\boldsymbol{\varepsilon}^{k+1}\chi(x) = \boldsymbol{\varepsilon}(x)^{k+1}\chi(x)$.

Proof. By Hurwitz' formula (Example 7.2.4), we see that the class

$$
\int_Q N_{w^*}^k \mathrm{ev}_Q([\Psi_{kg}]) \in H_Q^{g-1}(Y^\circ, \Gamma_Q, \mathbb{C})
$$

is represented by

$$
\begin{aligned}
&\left(\int_{Q,I} N_{w^*}(y)^k \psi_{kg,I}^{Q}(y) \omega(y) \right)_{I \in (X_{\mathbb{Q}})_Q^g} \\
&\qquad = \left(\frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \operatorname{sgn}(I) \sum_{x \in C_I^Q \cap \mathbb{Z}^g - \{0\}} \frac{1}{N_w(x)^{k+1}} \right)_{I \in (X_{\mathbb{Q}})_Q^g}.
\end{aligned}
$$

On the other hand, by Proposition 8.2.1(1), we can take a representative

$$
\Phi = \sum_{i=1}^{r} c_i \sigma_{I_i} \in K_{g-1} = \mathbb{Z}[\sigma_I \mid I \in (X_{\mathbb{Q}})_Q^g] \subset S_{g-1}
$$

of $\mathfrak{z}_\chi \in H_{g-1}(T_w/\Gamma_Q^+, \mathbb{Z})$. Then, by using Remark 8.2.3 and Proposition 8.2.1(2), we find

$$\begin{aligned}\left\langle \mathfrak{z}_\chi, \int_Q N_{w^*}^k \mathrm{ev}_Q([\Psi_{kg}]) \right\rangle &= \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \sum_{i=1}^r c_i \operatorname{sgn}(I_i) \sum_{x \in C_{I_i}^Q \cap \mathbb{Z}^g - \{0\}} \frac{1}{N_w(x)^{k+1}} \\ &= \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \sum_{x \in \mathbb{Z}^g - \{0\}} \sum_{i=1}^r c_i \operatorname{sgn}(I_i) \mathbf{1}_{C_{I_i}^Q}(x) \frac{1}{N_w(x)^{k+1}} \\ &= \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \sum_{x \in (\mathbb{Z}^g - \{0\})/\Gamma_Q^+} \sum_{\gamma \in \Gamma_Q^+} \sum_{i=1}^r c_i \operatorname{sgn}(I_i) \mathbf{1}_{C_{I_i}^Q}(\gamma x) \frac{1}{N_w(x)^{k+1}} \\ &= \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \sum_{x \in (\mathbb{Z}^g - \{0\})/\Gamma_Q^+} \frac{\chi(x)}{N_w(x)^{k+1}} \\ &= \frac{(k!)^g \det(w^{(1)}, \dots, w^{(g)})}{(g+gk-1)!} \sum_{x \in (\mathfrak{a} - \{0\})/\mathcal{O}_+^\times} \frac{\boldsymbol{\varepsilon}(x)^{k+1} \chi(x)}{|N_{F/\mathbb{Q}}(x)|^{k+1}}. \qquad \square\end{aligned}$$

Remark 8.3.3. It is easy to see that

$$\det(w^{(1)}, \dots, w^{(g)})^2 = D_{\mathcal{O}} N\mathfrak{a}^2,$$

where $D_{\mathcal{O}}$ is the discriminant of the order $\mathcal{O}$. Moreover, we also know that $\operatorname{sgn}(D_{\mathcal{O}}) = (-1)^{r_2}$, where r_2 is the number of complex places of F. Therefore, by permuting the order of the embeddings $\tau_1, \dots, \tau_g$ if necessary, we have

$$\det(w^{(1)}, \dots, w^{(g)}) = \boldsymbol{i}^{r_2} \sqrt{|D_{\mathcal{O}}|} N\mathfrak{a},$$

where $\boldsymbol{i} \in \mathbb{C}$ is the imaginary unit. Hence (under a suitable ordering of $\tau_1, \dots, \tau_g$), Theorem 8.3.2 can be also written as

$$\left\langle \mathfrak{z}_\chi, \int_Q N_{w^*}^k \mathrm{ev}_Q([\Psi_{kg}]) \right\rangle = \boldsymbol{i}^{r_2} \frac{\sqrt{|D_{\mathcal{O}}|} N\mathfrak{a} (k!)^g}{(g+gk-1)!} \zeta_{\mathcal{O}}(\boldsymbol{\varepsilon}^{k+1}\chi, \mathfrak{a}^{-1}, k+1).$$

Acknowledgements. I would like to express my gratitude to Takeshi Tsuji for the constant encouragement and for valuable advice during my research. I am also grateful to Kenichi Bannai, Kei Hagihara, Kazuki Yamada, and Shuji Yamamoto for answering my questions about their work and for their helpful comments. Thanks are also due to Ryotaro Sakamoto for helpful discussions. Some of the ideas in this paper were obtained during my stay at the Max Planck Institute for Mathematics in 2018. I would like to express my appreciation to Don Zagier and Günter Harder for the valuable discussions and for their great hospitality during the stay. Finally, I would also like to thank the referees for reading my manuscript carefully and giving many helpful comments. This work is supported by JSPS Overseas Challenge Program for Young Researchers Grant Number 201780267, JSPS KAKENHI Grant Number JP18J12744, JSPS KAKENHI Grant Number JP18H05233, and JSPS KAKENHI Grant Number JP20J01008.

References

[Bannai et al. 2023] K. Bannai, K. Hagihara, K. Yamada, and S. Yamamoto, "Canonical equivariant cohomology classes generating zeta values of totally real fields", *Trans. Amer. Math. Soc. Ser. B* **10** (2023), 613–635. MR Zbl

[Barnes 1904] E. W. Barnes, "On the theory of the multiple Gamma function", *Cambr. Trans.* **19** (1904), 374–425. Zbl

[Beilinson et al. 2018] A. Beilinson, G. Kings, and A. Levin, "Topological polylogarithms and p-adic interpolation of L-values of totally real fields", *Math. Ann.* **371**:3-4 (2018), 1449–1495. MR Zbl

[Bergeron et al. 2020] N. Bergeron, P. Charollois, and L. E. Garcia, "Transgressions of the Euler class and Eisenstein cohomology of $\mathrm{GL}_N(\mathbb{Z})$", *Jpn. J. Math.* **15**:2 (2020), 311–379. MR Zbl

[Cartan and Eilenberg 1956] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton Univ. Press, 1956. MR Zbl

[Charollois et al. 2015] P. Charollois, S. Dasgupta, and M. Greenberg, "Integral Eisenstein cocycles on GL_n, II: Shintani's method", *Comment. Math. Helv.* **90**:2 (2015), 435–477. MR Zbl

[Diaz y Diaz and Friedman 2014] F. Diaz y Diaz and E. Friedman, "Signed fundamental domains for totally real number fields", *Proc. Lond. Math. Soc.* (3) **108**:4 (2014), 965–988. MR Zbl

[Espinoza and Friedman 2020] M. Espinoza and E. Friedman, "Twisters and signed fundamental domains for number fields", *Ann. Inst. Fourier* (*Grenoble*) **70**:2 (2020), 479–521. MR Zbl

[Flórez et al. 2019] J. Flórez, C. Karabulut, and T. A. Wong, "Eisenstein cocycles over imaginary quadratic fields and special values of L-functions", *J. Number Theory* **204** (2019), 497–531. MR Zbl

[Godement 1973] R. Godement, *Topologie algébrique et théorie des faisceaux*, Publications de l'Institut de Mathématique de l'Université de Strasbourg **13**, Hermann, Paris, 1973. MR Zbl

[Harder 1987] G. Harder, "Eisenstein cohomology of arithmetic groups: the case GL_2", *Invent. Math.* **89**:1 (1987), 37–118. MR Zbl

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge Univ. Press, 2002. MR Zbl

[Hecke 1917] E. Hecke, "Über die Kroneckersche Grenzformel für reelle quadratische Körper und die Klassenzahl relativ-Abelscher Körper", *Verh. Naturf. Ges. Basel* **28** (1917), 363–372. Zbl

[Hida 1993] H. Hida, *Elementary theory of L-functions and Eisenstein series*, London Mathematical Society Student Texts **26**, Cambridge Univ. Press, 1993. MR Zbl

[Hill 2007] R. Hill, "Shintani cocycles on GL_n", *Bull. Lond. Math. Soc.* **39**:6 (2007), 993–1004. MR Zbl

[Hörmander 1973] L. Hörmander, *An introduction to complex analysis in several variables*, North-Holland Mathematical Library **7**, North-Holland, 1973. MR Zbl

[Hörmander 1994] L. Hörmander, *Notions of convexity*, Progress in Mathematics **127**, Birkhäuser, Boston, 1994. MR Zbl

[Hurwitz 1922] A. Hurwitz, "Über die Anzahl der Klassen positiver ternärer quadratischer Formen von gegebener Determinante", *Math. Ann.* **88**:1-2 (1922), 26–52. MR Zbl

[Kashiwara and Schapira 1990] M. Kashiwara and P. Schapira, *Sheaves on manifolds*, Grundl. Math. Wissen. **292**, Springer, 1990. MR Zbl

[Lim and Park 2019] S. H. Lim and J. Park, "The Milnor K-theory and the Shintani cocycle", 2019. arXiv 1909.03450

[Nori 1995] M. V. Nori, "Some Eisenstein cohomology classes for the integral unimodular group", pp. 690–696 in *Proceedings of the International Congress of Mathematicians*, *I* (Zürich, 1994), edited by S. D. Chatterji, Birkhäuser, Basel, 1995. MR Zbl

[Rudin 1991] W. Rudin, *Functional analysis*, 2nd ed., McGraw-Hill, New York, 1991. MR Zbl

[Sczech 1993] R. Sczech, "Eisenstein group cocycles for GL_n and values of L-functions", *Invent. Math.* **113**:3 (1993), 581–616. MR Zbl

[Sharifi and Venkatesh 2020] R. Sharifi and A. Venkatesh, "Eisenstein cocycles in motivic cohomology", 2020. arXiv 2011.07241

[Shintani 1976] T. Shintani, "On evaluation of zeta functions of totally real algebraic number fields at non-positive integers", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **23**:2 (1976), 393–417. MR Zbl

[Solomon 1998] D. Solomon, "Algebraic properties of Shintani's generating functions: Dedekind sums and cocycles on $\mathrm{PGL}_2(\mathbb{Q})$", *Compositio Math.* **112**:3 (1998), 333–362. MR Zbl

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

[Vlasenko and Zagier 2013] M. Vlasenko and D. Zagier, "Higher Kronecker 'limit' formulas for real quadratic fields", *J. Reine Angew. Math.* **679** (2013), 23–64. MR Zbl

[Yamamoto 2010] S. Yamamoto, "On Shintani's ray class invariant for totally real number fields", *Math. Ann.* **346**:2 (2010), 449–476. MR Zbl

Communicated by Samit Dasgupta
Received 2021-09-18 Revised 2022-04-29 Accepted 2022-07-06

bekki@mpim-bonn.mpg.de *Max Planck Institute for Mathematics, Bonn, Germany*

msp

ALGEBRA AND NUMBER THEORY 17:6 (2023)
https://doi.org/10.2140/ant.2023.17.1209

On the commuting probability of p-elements in a finite group

Timothy C. Burness, Robert Guralnick, Alexander Moretó and Gabriel Navarro

Let G be a finite group, let p be a prime and let $\mathrm{Pr}_p(G)$ be the probability that two random p-elements of G commute. In this paper we prove that $\mathrm{Pr}_p(G) > (p^2+p-1)/p^3$ if and only if G has a normal and abelian Sylow p-subgroup, which generalizes previous results on the widely studied commuting probability of a finite group. This bound is best possible in the sense that for each prime p there are groups with $\mathrm{Pr}_p(G) = (p^2+p-1)/p^3$ and we classify all such groups. Our proof is based on bounding the proportion of p-elements in G that commute with a fixed p-element in $G \setminus \boldsymbol{O}_p(G)$, which in turn relies on recent work of the first two authors on fixed point ratios for finite primitive permutation groups.

1. Introduction

The *commuting probability* of a finite group G is the probability that two random elements of G commute, namely

$$\mathrm{Pr}(G) = \frac{|\{(x, y) \in G \times G : xy = yx\}|}{|G|^2}.$$

A celebrated, but elementary, result of Gustafson [1973] asserts that $\mathrm{Pr}(G) > \frac{5}{8}$ if and only if G is abelian, which is best possible since $\mathrm{Pr}(D_8) = \frac{5}{8}$. This concept has been widely studied in recent years and some natural analogues for infinite groups have also been investigated; see, for instance, [Antolín et al. 2017; Eberhard 2015; Guralnick and Robinson 2006; Lescot 1995; Neumann 1989; Tointon 2020]. In addition, the commuting variety of elements in Lie algebras and algebraic groups has been a subject of great interest for several decades. This was originally introduced by Motzkin and Taussky [1955] and further studied by Richardson [1979], Ginzburg [2000], Premet [2003] and others.

In this paper, we pursue a local version of Gustafson's theorem, which turns out to be significantly more challenging.

Burness thanks the Department of Mathematics at the University of Padua for their generous hospitality during a research visit in autumn 2021. Guralnick was partially supported by the NSF grant DMS-1901595 and a Simons Foundation Fellowship 609771. Moretó and Navarro are supported by Ministerio de Ciencia e Innovación (Grant PID2019-103854GB-I00 funded by MCIN/AEI/10.13039/501100011033). Moretó is also supported by Generalitat Valenciana AICO/2020/298. We thank J. Martínez and P.H. Tiep for useful conversations on this paper. We thank the referee for their helpful comments.
MSC2020: 20D20.
Keywords: finite groups, commuting probability, p-elements.

Definition. Let G be a finite group, let p be a prime and let G_p be the set of p-elements in G (that is, the set of elements in G of order p^m for some $m \geqslant 0$). Then

$$\mathrm{Pr}_p(G) = \frac{|\{(x, y) \in G_p \times G_p : xy = yx\}|}{|G_p|^2}$$

is the probability that two random p-elements of G commute. Note that $\mathrm{Pr}_p(G) = 1$ if and only if G has a normal and abelian Sylow p-subgroup.

Local versions of the commuting probability have also been studied in the context of algebraic groups and Lie algebras. In particular, Premet [2003] identified the irreducible components of the commuting variety of nilpotent elements of a reductive Lie algebra defined over an algebraically closed field of good characteristic (and similarly, as an immediate consequence, for unipotent elements in the corresponding reductive algebraic groups). The set of commuting r-tuples of elements of order p (or commuting nilpotent elements of nilpotence degree p in a p-restricted Lie algebra) has also been studied for its connection to problems in representation theory; see [Carlson et al. 2016]. For finite groups, a generating function is presented in [Fulman and Guralnick 2018] for counting the number of commuting pairs of p-elements in some finite classical groups in good characteristic.

In this paper we consider arbitrary finite groups. Given a prime number p, set

$$f(p) = \frac{p^2 + p - 1}{p^3}.$$

Our first main result is the following.

Theorem A. *Let G be a finite group and let p be a prime. Then $\mathrm{Pr}_p(G) > f(p)$ if and only if G has a normal and abelian Sylow p-subgroup.*

In particular, if G is a nonabelian finite simple group and $|G|$ is divisible by p, then $\mathrm{Pr}_p(G) \leqslant f(p)$. We can say more in this situation.

Theorem B. *Let G be a nonabelian finite simple group and let p be a prime divisor of $|G|$. Then $\mathrm{Pr}_p(G) = f(p)$ if and only if $p \geqslant 5$ and G is isomorphic to $\mathrm{PSL}_2(p)$.*

In fact, we can classify all the finite groups G with $G = \boldsymbol{O}^{p'}(G)$ and $\mathrm{Pr}_p(G) = f(p)$, where $\boldsymbol{O}^{p'}(G)$ is the subgroup of G generated by G_p. See Theorem 5.2 for a precise statement. In particular, we observe that there is no nonsolvable group with $\mathrm{Pr}_2(G) = \frac{5}{8}$ and no nonsolvable group $G = \boldsymbol{O}^{3'}(G)$ with $\mathrm{Pr}_3(G) = \frac{11}{27}$. In addition, if G is given as in Theorem B with p a fixed prime, then $\mathrm{Pr}_p(G)$ tends to 0 as $|G|$ tends to infinity; we refer the reader to the end of Section 5 for further details.

Our next result, which may be of independent interest, is a key ingredient in the proof of Theorem A. Recall that $\boldsymbol{O}_p(G)$ denotes the largest normal p-subgroup of G.

Theorem C. *Let G be a finite group and let p be a prime. Then*

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} \leqslant \frac{1}{p}$$

for every p-element $x \in G \setminus \boldsymbol{O}_p(G)$.

This can be extended as follows.

Theorem D. *Let G be a finite group and let p be a prime. If $x \in G$ is a p-element and*

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} > \frac{1}{p},$$

then $x \in \boldsymbol{Z}(\boldsymbol{O}_p(G))$.

Remark 1. It is easy to see that the converse of Theorem D is false. For example, if $G = D_{8(2m+1)}$, then $|\boldsymbol{C}_G(x)_2|/|G_2| = 1/(2m+2)$ if $x \in \boldsymbol{Z}(\boldsymbol{O}_2(G))$ has order 4. On the other hand, in Examples 3.16 (p odd) and 3.17 ($p = 2$) we present a family of examples (G, p, x), where $x \in \boldsymbol{Z}(\boldsymbol{O}_p(G))$ is nontrivial and $|\boldsymbol{C}_G(x)_p|/|G_p|$ tends to 1 as p tends to infinity.

Remark 2. Let G be a finite group with $\boldsymbol{O}_p(G) = 1$. Then the conclusions in Theorems A and C are still valid if we work with elements of order p instead of all p-elements (with essentially no change in the proofs). And similarly for Theorem 5.2, which includes Theorem B as a special case.

The proofs of our main results depend upon the classification of finite simple groups. However, it is worth noting that our proof of Theorem C does not require the classification if we assume that x normalizes, but does not centralize, some normal p'-subgroup of G. This implies that the classification is not required for Theorem A under the assumption that the generalized Fitting subgroup of G is a p'-group (and so in particular, if G is p-solvable). In order to handle the general case, we use a recent result of the first two authors [Burness and Guralnick 2022] on fixed point ratios of elements of prime order in primitive permutation groups (see Theorem 3.4).

Remark 3. Let us observe that

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} = \frac{\Psi(x)}{\Psi(1)},$$

where Ψ is the permutation character for the action of G on its p-elements by conjugation. In the language of permutation groups, this number coincides with the fixed point ratio of x with respect to this action, which explains why the main theorem of [Burness and Guralnick 2022] will be an important ingredient in the proof of Theorem C.

2. Some preliminary results

For the remainder of this paper, all groups are finite and p is a prime number. We will frequently use the elementary fact that if G is a group and $H, K \leqslant G$ are subgroups, then

$$|H : H \cap K| \leqslant |G : K| \tag{1}$$

with equality if and only if $G = HK$.

Lemma 2.1. *Let G be a finite group and let N be a normal p-subgroup.*

(i) *If $x \in G$ is a p-element, then*

$$|\boldsymbol{C}_G(x)_p|/|G_p| \leqslant |\boldsymbol{C}_{G/N}(Nx)_p|/|(G/N)_p|.$$

(ii) $\mathrm{Pr}_p(G) \leqslant \mathrm{Pr}_p(G/N)$.

Proof. Both parts quickly follow from the fact that $|G_p| = |(G/N)_p||N|$. □

Remark 2.2. In the previous lemma, the assumption that N is a p-subgroup is essential. For example, there is a semidirect product $G = C_{35}{:}D_{12}$ with a normal subgroup N of order 3 such that $G/N = D_{10} \times D_{14}$ and we compute

$$\mathrm{Pr}_2(G) = \tfrac{211}{1296} > \tfrac{11}{72} = \mathrm{Pr}_2(G/N).$$

(Here G is `SmallGroup`(420, 30) in the GAP Small Groups library [GAP 2020].) One can check that this is the smallest finite group with $\mathrm{Pr}_p(G) > \mathrm{Pr}_p(G/N)$ for some prime p.

There is a special case where quotients by normal subgroups of order prime to p do not change the proportions.

Lemma 2.3. *Let G be a finite group and let N be a central p'-subgroup.*

(i) *If $x \in G$ is a p-element, then*

$$|\boldsymbol{C}_G(x)_p|/|G_p| = |\boldsymbol{C}_{G/N}(Nx)_p|/|(G/N)_p|.$$

(ii) *If N is central in G, then* $\mathrm{Pr}_p(G) = \mathrm{Pr}_p(G/N)$.

Proof. Let $x \in G$ be a p-element and suppose that $[x, y] \in N$ for some $y \in G$. Since $[x, N] = 1$ it follows that $[x^p, y] = [x, y]^p$, so $[x, y] = 1$. In addition, if y is a p-element, then y is the only p-element in the coset Ny and so (i) follows. Now (ii) follows from (i), noting that G and G/N both have the same number of p-elements. □

Lemma 2.4. *Let P be a p-group acting on a p'-group K and let L be a P-invariant subgroup of K. If*

$$\frac{|\boldsymbol{C}_K(P) : \boldsymbol{C}_L(P)|}{|K : L|} < 1, \tag{2}$$

then

$$\frac{|\boldsymbol{C}_K(P) : \boldsymbol{C}_L(P)|}{|K : L|} \leqslant \frac{1}{p+1}.$$

Proof. Let $C = \boldsymbol{C}_K(P)$ and note that $K \neq CL$ in view of the inequality in (2). For any prime q, let L_q be a P-invariant Sylow q-subgroup of L, which is contained in a P-invariant Sylow q-subgroup K_q of K; see [Isaacs 2008, Corollary 3.25]. Thus $K_q \cap L = L_q$. By coprime action, $C_q := C \cap K_q$ and $C \cap L_q$ are Sylow q-subgroups of C and $C \cap L$, respectively; see [Isaacs 2008, Lemma 3.32], for example. In view of (2) we have

$$\prod_q \frac{|C_q : C_q \cap L|}{|K_q : L_q|} = \frac{|C : C \cap L|}{|K : L|} < 1$$

and we note that

$$\frac{|C_q : C_q \cap L|}{|K_q : L_q|} = \frac{|C_q : C_q \cap L_q|}{|K_q : L_q|} \leqslant 1$$

for every prime q (see (1)). Therefore,

$$\frac{|C : C \cap L|}{|K : L|} \leqslant \frac{|C_q : C_q \cap L|}{|K_q : L_q|}$$

for every q, so the bound in (2) implies that

$$\frac{|C_q : C_q \cap L|}{|K_q : L_q|} < 1$$

for some q. As a consequence, we are free to assume that K is a q-group.

Arguing by induction on $|K : L|$, we may assume that L is a maximal P-invariant subgroup of K. Then L is normal in K and K/L does not have any proper nontrivial P-invariant subgroups, whence (2) implies that $C = \boldsymbol{C}_K(P) = \boldsymbol{C}_L(P)$. If $|\boldsymbol{C}_K(y) : \boldsymbol{C}_L(y)| = |K : L|$ for every $y \in P$, then P acts trivially on K/L and thus $K = CL$, which is incompatible with (2). Therefore, we may assume that $P = \langle y \rangle$ is cyclic. Then the action of P on K/L is a Frobenius action, which implies that if $x \in K \setminus L$, then $\{L, Lx^z : z \in P\}$ is a set of distinct cosets of L in K. Therefore $|K : L| \geqslant |P| + 1 \geqslant p + 1$, as required. □

Next we record the following well known result.

Lemma 2.5. *Let G be a finite group, let $x, y \in G$ and let $K \leqslant G$ be a subgroup normalized by x and y. If $Kx = Ky$ and $|K|$ is coprime with $o(x)o(y)$, then x and y are K-conjugate.*

Proof. We may assume $G = K\langle x, y \rangle$ and thus K is normal in G. Since $Kx = Ky$, it follows that $K\langle x \rangle = K\langle y \rangle = G$. Now, $K \cap \langle x \rangle = K \cap \langle y \rangle = 1$ and we also note that $o(x) = o(y)$ and $\langle x \rangle$, $\langle y \rangle$ are Hall π-subgroups of G, where π is the set of primes dividing $o(x)$. By the Schur–Zassenhaus theorem, we have $\langle x \rangle^k = \langle y \rangle$ for some $k \in K$ and thus $x^k = y^n$ for some integer n. Now, $Ky = Kx = Kx^k = Ky^n$ and $y^n y^{-1} \in K \cap \langle y \rangle = 1$, so $y^n = y$ and the result follows. □

We shall need one more well known fact about coprime actions, which follows from [Isaacs 2008, Theorem 3.27].

Lemma 2.6. *Let G and A be finite groups with coprime orders and suppose that A acts on G by automorphisms. Set $C = \boldsymbol{C}_G(A)$. Then $G = C[A, G]$ and $[A, [A, G]] = [A, G]$.*

3. Proofs of Theorems C and D

In this section we prove Theorems C and D. We begin by handling a special case of Theorem C, which relies on the following proposition. In part (i), we write $(Ky)_p$ for the set of p-elements in the coset Ky, where p is a fixed prime throughout this section.

Proposition 3.1. *Let G be a finite group and let K be a normal p'-subgroup of G. Let $x \in G$ be an element of order p such that $K = [x, K]$ and let $y \in G$ be a p-element with $[x, y] = 1$.*

(i) *If $[y, K] \neq 1$, then the proportion of elements in $y^K = (Ky)_p$ which commute with x is at most $1/(p+1)$.*

(ii) *If $L = \langle K, x\rangle$, then the proportion of p-elements in the coset Ly which commute with x is at most $1/p$.*

Proof. First consider (i). Since K is a p'-group, Lemma 2.5 implies that y^K is precisely the set of p-elements in the coset Ky. Next observe that $y^K \cap \boldsymbol{C}_G(x) = (yK)_p \cap \boldsymbol{C}_G(x) = (Ay)_p$, where $A = \boldsymbol{C}_K(x)$, and another application of Lemma 2.5 gives $(Ay)_p = y^A$. Therefore, the proportion of elements in y^K which commute with x is equal to

$$\frac{|y^K \cap \boldsymbol{C}_G(x)|}{|y^K|} = \frac{|\boldsymbol{C}_K(x) : \boldsymbol{C}_K(x) \cap \boldsymbol{C}_K(y)|}{|K : \boldsymbol{C}_K(y)|}. \tag{3}$$

If every element in y^K commutes with x, then $[y, K] \leqslant \boldsymbol{C}_K(x)$. But then the three subgroups lemma implies that y centralizes $[x, K] = K$, which is incompatible with the condition $[y, K] \neq 1$ in (i). Therefore, the proportion in (3) is less than 1 and by applying Lemma 2.4 (with $L = \boldsymbol{C}_K(y)$ and $P = \langle x\rangle$) we deduce that it is at most $1/(p+1)$ as required.

We now prove (ii). For $0 \leqslant i < p$, let a_i be the number of p-elements in Kx^iy commuting with x and let $b_i = |(Kx^iy)_p|$, so $\alpha = \sum_i a_i / \sum_i b_i$ is the proportion of p-elements in Ly which commute with x. If $[x^iy, K] \neq 1$ for all i, then (i) implies that $a_i/b_i \leqslant 1/(p+1)$ and we immediately deduce that $\alpha \leqslant 1/(p+1)$. Therefore, we may assume $[y, K] = 1$ (otherwise replace y by x^iy for some i). For $1 \leqslant i < p$ it follows that $[x^iy, K] \neq 1$ (since $[x, K] = K$) and thus $a_i/b_i \leqslant 1/(p+1)$. Since $|y^K| = 1$ we have $a_0 = b_0 = 1$ and we deduce that

$$\alpha \leqslant \frac{1}{p+1} + \frac{p}{(p+1)m},$$

where $m = |(Ly)_p|$. Finally, we note that $b_i \geqslant (p+1)a_i \geqslant p+1$ for $1 \leqslant i < p$ (since $x^iy \in Kx^iy$ is a p-element commuting with x), so $m \geqslant 1 + (p-1)(p+1) = p^2$ and we conclude that $\alpha \leqslant 1/p$. □

We are now ready to prove a special case of Theorem C.

Theorem 3.2. *Let G be a finite group and let $x \in G$ be an element of order p. If there exists a normal p'-subgroup K of G with $[x, K] \neq 1$, then*

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} \leqslant \frac{1}{p}. \tag{4}$$

Proof. By Lemma 2.1, we may assume that $\boldsymbol{O}_p(G) = 1$. We can also assume that $G = K\boldsymbol{C}_G(x)$ and we may replace K by any proper normal subgroup of G contained in K that does not centralize x. In particular, by Lemma 2.6, we can replace K by $[x, K]$ and so we may assume that $K = [x, K]$.

Set $L = \langle K, x\rangle$ and let $y \in G$ be a p-element. It suffices to show that the proportion of p-elements in the coset Ly which commute with x is at most $1/p$. Clearly, if no p-element in Ly commutes with x, then this proportion is 0, so we may assume $[x, y] = 1$. Now apply Proposition 3.1(ii). □

Remark 3.3. Let $\boldsymbol{F}^*(G)$ be the generalized Fitting subgroup of G. If $\boldsymbol{F}^*(G)$ is a p'-group, then $\boldsymbol{O}_p(G) = 1$ and the statement of Theorem 3.2 holds for every nontrivial p-element x because we can replace x by an element of order p in $\langle x\rangle$. Of course, if the upper bound in (4) holds for all elements in G of order p (modulo $\boldsymbol{O}_p(G)$), then the same bound holds for every nontrivial p-element in G.

Recall that if G is a permutation group on a finite set Ω, then the *fixed point ratio* of an element $z \in G$, denoted $\mathrm{fpr}(z, \Omega)$, is the proportion of points in Ω fixed by z. It is easy to see that if G is transitive and H is a point stabilizer, then

$$\mathrm{fpr}(z, \Omega) = \frac{|z^G \cap H|}{|z^G|}.$$

The following is a simplified version of the main theorem of [Burness and Guralnick 2022].

Theorem 3.4. *Let $G \leqslant \mathrm{Sym}(\Omega)$ be a finite primitive permutation group with point stabilizer H. If $z \in G$ has prime order p, then either*

$$\mathrm{fpr}(z, \Omega) \leqslant \frac{1}{p+1},$$

or one of the following holds (up to permutation isomorphism):

(i) *G is almost simple and either*
   (a) *$G = S_n$ or A_n acting on k-element subsets of $\{1, \dots, n\}$ with $1 \leqslant k < n/2$; or*
   (b) *$(G, H, z, \mathrm{fpr}(z, \Omega))$ is known.*

(ii) *G is an affine group, $\boldsymbol{F}^*(G) = \boldsymbol{F}(G) = (C_p)^d$, $z \in \mathrm{GL}_d(p)$ is a transvection and $\mathrm{fpr}(z, \Omega) = 1/p$.*

(iii) *$G \leqslant A \wr S_t$ is a product type group with its product action on $\Omega = \Gamma^t$ and $z \in A^t \cap G$, where $A \leqslant \mathrm{Sym}(\Gamma)$ is one of the almost simple primitive groups in part* (i).

We will also need the following corollary to Theorem 3.4 in the almost simple setting; see [Burness and Guralnick 2022, Corollary 3]. Recall that the *socle* of an almost simple group G is its unique minimal normal subgroup, which coincides with $\boldsymbol{F}^*(G)$.

Corollary 3.5. *Let $G \leqslant \mathrm{Sym}(\Omega)$ be a finite almost simple primitive permutation group with socle J. If $z \in G$ has prime order p, then either*

$$\mathrm{fpr}(z, \Omega) \leqslant \frac{1}{p},$$

or one of the following holds (up to permutation isomorphism):

(i) *$J = A_n$ and Ω is the set of k-element subsets of $\{1, \dots, n\}$ for some $1 \leqslant k < n/2$.*

(ii) *$(J, p) = (\mathrm{PSL}_2(q), q-1)$, $(\mathrm{Sp}_6(2), 3)$, $(\mathrm{PSU}_4(2), 2)$, $(\mathrm{Sp}_n(2), 2)$ or $(\Omega_n^{\epsilon}(2), 2)$.*

We will now use Theorem 3.4 and Corollary 3.5 to handle two more special cases of Theorem C, which will then be applied to obtain the result in full generality. In the following proposition, the *components* of K are the quasisimple groups referred to in the statement.

Proposition 3.6. *Let K be a central product of quasisimple groups with $\boldsymbol{O}_p(K)=1$ and let $x, y \in \mathrm{Aut}(K)$ be nontrivial p-elements such that x does not normalize any component of K. Assume that the simple quotients of the components of K are isomorphic. Then the proportion of elements in y^K which commute with x is at most $1/(p+1)$.*

Proof. We may assume that $[x, y]=1$ and x has order p. Let $K_1, \ldots, K_t$ be the components of K and set $L_i = K_i/\boldsymbol{Z}(K_i) \cong L$. Note that t is a multiple of p since x acts fixed point freely on the set of components. We can now view x and y as commuting automorphisms of the direct product $J := L^t$, with $o(x)=p$ and $o(y)=p^m$ for some $m \geqslant 1$. Set $G=\langle J, x, y\rangle \leqslant \mathrm{Aut}(J)$ and note that J is the unique minimal normal subgroup of G. Now

$$\frac{|y^J \cap \boldsymbol{C}_G(x)|}{|y^J|} = \frac{|x^J \cap \boldsymbol{C}_G(y)|}{|x^J|}$$

and it suffices to show that

$$\frac{|x^J \cap \boldsymbol{C}_G(y)|}{|x^J|} \leqslant \frac{1}{p+1}. \tag{5}$$

Let M be a maximal subgroup of G containing $\boldsymbol{C}_G(y)$ and observe that M does not contain J since $G = J\boldsymbol{C}_G(y)$. This allows us to view G acting primitively on the set of cosets $\Omega = G/M$ and we note that

$$\frac{|x^J \cap \boldsymbol{C}_G(y)|}{|x^J|} \leqslant \frac{|x^G \cap M|}{|x^G|} = \mathrm{fpr}(x, \Omega).$$

Then by applying Theorem 3.4, noting that $x \notin \mathrm{Aut}(L)^t \cap G$ by hypothesis, it follows that $\mathrm{fpr}(x, \Omega) \leqslant 1/(p+1)$ and thus (5) holds. □

Next we seek a version of Proposition 3.6 in the special case where K is quasisimple (see Propositions 3.10 and 3.12). In order to do this, we will need the following elementary result.

Lemma 3.7. *Let G be a finite group, let $x \in G \setminus \boldsymbol{O}_p(G)$ be a p-element and set*

$$D = \{y \in G : \langle y\rangle \text{ is } G\text{-conjugate to } \langle x\rangle\}.$$

Then $|D| \geqslant p^2-1$.

Proof. Without loss of generality, we may assume that $\boldsymbol{O}_p(G)=1$ and x has order p. Consider the natural action of G on the set C of conjugates of $\langle x\rangle$ and note that $|D| = (p-1)|C|$, so it suffices to show that $|C| \geqslant p+1$. Note that x fixes $\langle x\rangle \in C$, so it has at least one fixed point on C. If x acts trivially on C, then x centralizes each of its conjugates and thus, by Baer's theorem, $x \in \boldsymbol{O}_p(G)$, which is a contradiction. Therefore, x acts nontrivially on C and we conclude that $|C| \geqslant p+1$. □

Remark 3.8. Let G, D and p be given as in Lemma 3.7. Then $|D| = p^2-1$ if and only if $|G_p| = p^2$ and the groups with this property are determined in Lemma 5.1.

We also need the following result, which is a corollary of Theorem 3.4.

Lemma 3.9. *Let G be an almost simple group with socle J and assume J is not isomorphic to an alternating group. Let p be a prime divisor of $|J|$ and suppose $x \in G$ has order p. Then there exists an element $y \in J$ of order p such that*

$$\frac{|y^G \cap \boldsymbol{C}_G(x)|}{|y^G|} \leqslant \frac{1}{p+1}.$$

Proof. We may assume $G = \langle J, x \rangle$ and we may embed $\boldsymbol{C}_G(x)$ in a core-free maximal subgroup H of G, so

$$\frac{|y^G \cap \boldsymbol{C}_G(x)|}{|y^G|} \leqslant \frac{|y^G \cap H|}{|y^G|} = \mathrm{fpr}(y, G/H)$$

for every element $y \in J$ of order p. Clearly, the desired conclusion holds if there exists such an element with $\mathrm{fpr}(y, G/H) \leqslant 1/(p+1)$, so we may assume otherwise, in which case (G, H, y) is one of the special cases arising in part (i)(b) of Theorem 3.4. More precisely, [Burness and Guralnick 2022, Theorem 1] implies that either G is a classical group in a subspace action (and the special cases that arise are recorded in [loc. cit., Table 6]), or $G = \mathrm{M}_{22}:2$, $H = \mathrm{PSL}_3(4).2_2$ and $p = 2$. In the latter case one can check that $\mathrm{fpr}(y, G/H) = \frac{3}{11}$ if $y \in J$ is an involution, so we may assume G is a classical group in a subspace action. We now inspect the cases in [loc. cit., Table 6].

If J is a unitary, symplectic or orthogonal group, then it is easy to check that in every case (G, H) there exists an element $y \in J$ of order p such that $\mathrm{fpr}(y, G/H) \leqslant 1/(p+1)$. For example, if $J = \mathrm{PSp}_n(q)$ with $n \geqslant 4$, $H = P_1$ is the stabilizer of a 1-space and $p = q$, then we can take $y = (J_2^2, J_1^{n-4})$, where J_i denotes a standard unipotent Jordan block of size i.

To complete the proof, let us assume $J = \mathrm{PSL}_n(q)$ is a linear group and note that $H = P_1$ is the stabilizer of a 1-space. If $n \geqslant 4$ then once again it is straightforward to see that there is an element $y \in J$ of order p with $\mathrm{fpr}(y, G/H) \leqslant 1/(p+1)$, so we may assume $n \in \{2, 3\}$. Suppose $n = 3$. If $p = q \geqslant 3$ then we can choose $y = (J_3)$, while for $q = 2$ we must take $y = (J_2, J_1)$ and one can use GAP [2020] to verify the desired bound in the statement of the lemma. Similarly, if $p = q - 1 \geqslant 3$ then we can take y to be the image (modulo scalars) of a diagonal matrix $(\omega, \omega^{-1}, I_1)$, where $\omega \in \mathbb{F}_q^{\times}$ has order p. And if $(q, p) = (3, 2)$ then $y = (-I_2, I_1)$ is the only option and the result can be checked using GAP.

Finally, suppose $J = \mathrm{PSL}_2(q)$, so $q \geqslant 7$ since $\mathrm{PSL}_2(4)$ and $\mathrm{PSL}_2(5)$ are both isomorphic to A_5. If $p = q-1$ then $|y^G| = q(q+1)$ and $|\boldsymbol{C}_G(x)| < q$, so the desired bound holds. Now assume $q = p$. Here both x and y are regular unipotent elements and we compute $|y^G| = (p^2-1)/2$ and $|y^G \cap \boldsymbol{C}_G(x)| = (p-1)/2$, which implies that

$$\frac{|y^G \cap \boldsymbol{C}_G(x)|}{|y^G|} = \frac{1}{p+1}.$$

The result follows. □

Proposition 3.10. *Let K be a quasisimple group such that $\boldsymbol{O}_p(K) = 1$ and $K/\boldsymbol{Z}(K)$ is not isomorphic to an alternating group. Let $x \in \mathrm{Aut}(K)$ be a nontrivial p-element.*

(i) *There is a normal subset D of nontrivial p-elements in K such that $|D| \geqslant p^2 - 1$ and the proportion of elements in D which commute with x is at most $1/(p+1)$.*

(ii) *Let $y \in \mathrm{Aut}(K)$ be a nontrivial p-element.*

(a) *The proportion of elements in y^K which commute with x is at most $1/p$, unless $K = \Omega_n^+(2)$, $n \geqslant 8$, $p = 2$ and both x and y are transvections, in which case the proportion is $\frac{1}{2} + \frac{1}{2(2^{n/2}-1)}$.*

(b) *The proportion of elements in $(Ky)_p$ which commute with x is at most $1/p$.*

Proof. We may assume x has order p. Set $J = K/\boldsymbol{Z}(K)$ and view x as an automorphism of J of order p. Set $G = \langle J, x\rangle$. By Lemma 3.9, there exists an element $y \in J$ of order p such that

$$\frac{|y^J \cap \boldsymbol{C}_G(x)|}{|y^J|} \leqslant \frac{|y^G \cap \boldsymbol{C}_G(x)|}{|y^G|} \leqslant \frac{1}{p+1}. \tag{6}$$

If we write y for the corresponding element in K, then by applying Lemma 3.7 we deduce that the normal subset

$$D = \{z \in K : \langle z\rangle \text{ is } K\text{-conjugate to } \langle y\rangle\} = \bigcup_{i=1}^{t} z_i^K$$

contains at least $p^2 - 1$ elements. Moreover, (6) implies that the proportion of elements in z_i^K which commute with x is at most $1/(p+1)$ for $1 \leqslant i \leqslant t$ and thus part (i) follows.

Now let us turn to part (ii). We may assume $[x, y] = 1$ and we may view y as an automorphism of J with $o(y) = p^a$ for some $a \geqslant 1$. Set $G = \langle J, x, y\rangle \leqslant \mathrm{Aut}(J)$ and embed $\boldsymbol{C}_G(y)$ in a core-free maximal subgroup H of G, which allows us to view G as an almost simple primitive permutation group on $\Omega = G/H$.

For now, let us exclude the special cases (J, p) in Corollary 3.5(ii). Then Corollary 3.5 implies that

$$\frac{|y^J \cap \boldsymbol{C}_G(x)|}{|y^J|} = \frac{|x^J \cap \boldsymbol{C}_G(y)|}{|x^J|} \leqslant \frac{|x^G \cap H|}{|x^G|} = \mathrm{fpr}(x, G/H) \leqslant \frac{1}{p} \tag{7}$$

and thus the proportion of elements in y^K which commute with x is at most $1/p$.

Next consider the coset Ky. Write $(Ky)_p = y_1^K \cup \cdots \cup y_r^K$ as a disjoint union of K-classes. If $Ky \neq K$ then each y_i is a nontrivial p-element and so the proportion of elements in y_i^K commuting with x is at most $1/p$ by (7) and the desired result follows. A very similar argument applies when $Ky = K$, but here we have to account for the identity element. To do this, write $K_p = \{1\} \cup D \cup z_1^K \cup \cdots \cup z_s^K$, where D is the normal subset in (i) and each z_i is nontrivial. Set $a_0 = |D \cap \boldsymbol{C}_K(x)| + 1$, $b_0 = |D| + 1$, $a_i = |z_i^K \cap \boldsymbol{C}_K(x)|$ and $b_i = |z_i^K|$ for $i \geqslant 1$. As above, we have $a_i/b_i \leqslant 1/p$ for $i \geqslant 1$, so it suffices to show that $a_0/b_0 \leqslant 1/p$. If we write $D = y_1^K \cup \cdots \cup y_t^K$, then $|y_i^K \cap \boldsymbol{C}_K(x)|/|y_i^K| \leqslant 1/(p+1)$ for each i and thus

$$\frac{a_0}{b_0} \leqslant \frac{1}{p+1} + \frac{p}{m(p+1)},$$

where $m = |D| + 1$. Since $m \geqslant p^2$ we deduce that $a_0/b_0 \leqslant 1/p$ and the result follows.

To complete the proof of (ii), it remains to consider the special cases (J, p) in Corollary 3.5(ii). In each of these cases, G is an almost simple classical group in a subspace action with point stabilizer H and there exists an element $z \in G$ of order p with $\mathrm{fpr}(z, G/H) > 1/p$. The possibilities for (G, H, z) are recorded in [Burness and Guralnick 2022, Table 1]. By inspection, we observe that either

(a) $\boldsymbol{C}_G(z)$ is contained in a maximal subgroup M of G such that $\mathrm{fpr}(z', G/M) \leqslant 1/p$ for all $z' \in G$ of order p; or

(b) $G = \mathrm{O}_n^+(2)$, $n \geqslant 8$, $p = 2$, H is the stabilizer of a nonsingular 1-space and $z = (J_2, J_1^{n-2})$.

So excluding the special case in (b), the previous argument goes through. In particular, the previous argument applies if $y \in K$ (note that in case (b), z is contained in $\mathrm{O}_n^+(2) \setminus J$).

We have now reduced to the case where $G = \mathrm{O}_n^+(2)$, $p = 2$ and both x and y are transvections. Here $y^J = y^G$ and $\boldsymbol{C}_G(x) = H$ is the stabilizer of a nonsingular 1-space, so [Burness and Guralnick 2022, Theorem 1] gives

$$\frac{|y^J \cap \boldsymbol{C}_G(x)|}{|y^J|} = \mathrm{fpr}(y, G/H) = \frac{1}{2} + \frac{1}{2(2^{n/2} - 1)}$$

for the proportion of elements in y^K commuting with x. So this is an exception to the main bound in (ii)(a), but we still claim that the proportion of 2-elements in Ky commuting with x is at most $\frac{1}{2}$.

To see this, write $(Ky)_2 = y^K \cup y_1^K \cup \cdots \cup y_r^K$ as a disjoint union. By [Burness and Guralnick 2022, Theorem 1], the proportion of elements in y_i^K which commute with x is at most $\frac{1}{3}$ for each $1 \leqslant i \leqslant r$. As a consequence, we deduce that the proportion of 2-elements in Ky commuting with x is at most $\frac{1}{2}$ so long as

$$3.2^{n/2-1} = \frac{3|y^K|}{2^{n/2} - 1} \leqslant \sum_{i=1}^{r} |y_i^K|.$$

But this inequality clearly holds since $|z^G| \geqslant 2^{n/2-1}(2^{n/2} - 1)$ for every nontrivial 2-element $z \in G$. $\square$

Remark 3.11. Let us observe that the upper bound in Proposition 3.10(ii)(b) is best possible. For example, let $K = \mathrm{PSL}_2(p)$ and let x and y be inner automorphisms of K of order p. Then $|(Ky)_p| = p^2$ and $|\boldsymbol{C}_K(x)| = p$, so the relevant proportion is exactly $1/p$.

We need a different result to handle alternating and symmetric groups.

Proposition 3.12. *Let $L = S_n$ and $J = A_n$, where $n \geqslant 5$. Let $x \in L$ be an element of prime order p and let $y \in L$ be a transposition.*

(i) *If p is odd, then the proportion of p-elements in J which commute with x is at most $1/p$.*

(ii) *If $p = 2$ and x is not a transposition, then the proportion of 2-elements in J or Jy which commute with x is at most $\frac{1}{2}$.*

(iii) *If $p = 2$ and x is a transposition, then the proportion of 2-elements in L which commute with x is at most $\frac{1}{2}$.*

Proof. First assume p is odd and $|\mathrm{supp}(x)| = m \geqslant 5$ with respect to the natural action of L on $\{1,\dots,n\}$. Note that p divides m and it suffices to work inside $H := A_m \times A_{n-m}$ since H contains $\boldsymbol{C}_J(x)_p$. In particular, we may assume that $m = n$. The result is clear if $m = p$, so assume $m > p$. Here $\boldsymbol{C}_J(x)$ is contained in an imprimitive subgroup $K = A_p \wr A_{m/p}$ and so if we write $J_p = x_0^J \cup \cdots \cup x_t^J$ with $x_0 = 1$, then

$$\frac{|\boldsymbol{C}_J(x)_p|}{|J_p|} \leqslant \frac{|K_p|}{|J_p|} = \frac{1+\sum_{i=1}^t a_i}{1+\sum_{i=1}^t b_i},$$

where $a_i = |x_i^J \cap K|$ and $b_i = |x_i^J|$. Now Theorem 3.4 implies that $a_i/b_i \leqslant 1/(p+1)$ for all i and we deduce that

$$\frac{|\boldsymbol{C}_J(x)_p|}{|J_p|} \leqslant \frac{1}{p+1} + \frac{p}{(p+1)c},$$

where $c = |J_p|$. Since $c \geqslant p^2$ (for example, this follows from Lemma 3.7) we conclude that this proportion is at most $1/p$, as required.

So to complete the proof of (i), it remains to handle the special case where $x = (1,2,3)$ is a 3-cycle. Set $d = |(S_{n-3})_3|$ and note that $|\boldsymbol{C}_J(x)_3| = 3d$. If $a = (1,2,i) \in J$ with $i \geqslant 4$, then for each 3-element $b \in J$ fixing 1, 2 and i we see that $a^{\pm}b \in J \setminus \boldsymbol{C}_J(x)$ is a 3-element. Therefore, $|J_3| \geqslant 2d(n-3)+3d$ and thus

$$\frac{|\boldsymbol{C}_J(x)_3|}{|J_3|} \leqslant \frac{3}{2n-3} \leqslant \frac{1}{3}$$

for $n \geqslant 6$. The case $n = 5$ can be handled directly.

For the remainder, let us assume $p = 2$ and write $|\mathrm{supp}(x)| = 2m$. For $m \geqslant 4$ we can essentially repeat the argument in (i). Write $\boldsymbol{C}_L(x) = (S_2 \wr S_m) \times S_{n-2m}$ and let a_1 and a_2 be the number of even and odd 2-elements in $S_2 \wr S_m$, respectively. Similarly, let $b_1 = |(A_{n-2m})_2|$ and $b_2 = |(S_{n-2m} \setminus A_{n-2m})_2|$. Then

$$|\boldsymbol{C}_J(x)_2| = a_1b_1 + a_2b_2, \quad |\boldsymbol{C}_L(x)_2 \cap Jy| = a_1b_2 + a_2b_1.$$

We claim that

$$a_1 \leqslant \tfrac{1}{2}|(A_{2m})_2|, \quad a_2 \leqslant \tfrac{1}{2}|(S_{2m} \setminus A_{2m})_2|. \tag{8}$$

To see this, set $K = A_{2m}$ and $H = (S_2 \wr S_m) \cap K$, so $a_1 = |H_2|$. By [Burness and Guralnick 2022, Theorem 1], we observe that $|z^K \cap H|/|z^K| \leqslant \frac{1}{3}$ for every nontrivial 2-element $z \in K$ and by arguing as in case (i) we deduce that $|H_2|/|K_2| \leqslant \frac{1}{2}$. This justifies the first inequality in (8) and a very similar argument establishes the second. As an immediate consequence, we deduce that

$$|\boldsymbol{C}_J(x)_2| \leqslant \tfrac{1}{2}|(S_{2m} \times S_{n-2m})_2 \cap J|, \quad |\boldsymbol{C}_L(x)_2 \cap Jy| \leqslant \tfrac{1}{2}|(S_{2m} \times S_{n-2m})_2 \cap Jy|$$

and thus the proportion of 2-elements in J and Jy commuting with x is at most $\frac{1}{2}$. In the same way, if $m = 3$ then we can reduce to the case $n = 6$ and here we can check the result directly.

Next assume $m = 2$, say $x = (1,2)(3,4)$. Set $d = |(S_{n-4})_2|$ and note that

$$|\boldsymbol{C}_J(x)_2| = |\boldsymbol{C}_L(x)_2 \cap Jy| = 4d.$$

Fix i, j with $4 < i < j$. Let $Z(i, j)$ (respectively, $W(i, j)$) be the set of elements in L of the form uv, where u is a 4-cycle (respectively, a double transposition different from $(1, 2)(i, j)$) on $\{1, 2, i, j\}$ and v is a 2-element fixing each of these 4 points. Then $|Z(i, j)| = 6d$ and $|W(i, j)| = 2d$, so there are at least $2d$ distinct 2-elements of each parity in $Z(i, j) \cup W(i, j)$, none of which commute with x. Since there are $(n-4)(n-5)/2$ choices for $\{i, j\}$, and the corresponding sets of 2-elements are pairwise disjoint, this implies that the proportion of 2-elements in each coset commuting with x is at most

$$\frac{4}{4+(n-4)(n-5)} \leqslant \frac{1}{2}$$

for $n \geqslant 7$. The cases $n = 5, 6$ can be handled directly.

Finally, let us assume $x = (1, 2)$ is a transposition. Set $d = |(S_{n-2})_2|$ and note that $|\boldsymbol{C}_L(x)_2| = 2d$. For each $j \in \{3, \dots, n\}$, let Z_j denote the set of 2-elements in L which interchange 1 and j. Note that the Z_j are pairwise disjoint sets of size d and no element in Z_j commutes with x. Therefore, $|L_2| \geqslant nd$ and we conclude that the proportion of 2-elements in L centralizing x is at most $2/n$. □

Remark 3.13. One can show that the conclusion in part (ii) of Proposition 3.12 also holds when x is a transposition. But the proof is more involved and we do not require the stronger result.

Remark 3.14. In the proof of Theorem C, we will need to extend Proposition 3.12 to central extensions K of A_n with $\boldsymbol{O}_p(K) = 1$. This follows by Lemma 2.3 unless an element of order p does not centralize $\boldsymbol{Z}(K)$. This only occurs when $p = 2$ and K is a 3-fold cover of A_n. One can check the result directly for these cases.

Finally, we are now ready to complete the proof of Theorem C.

Proof of Theorem C. Let G be a finite group and let $x \in G \setminus \boldsymbol{O}_p(G)$ be a p-element. Let $\boldsymbol{F}(G)$ and $\boldsymbol{F}^*(G)$ denote the Fitting and generalized Fitting subgroups of G, respectively, and note that $x \notin \boldsymbol{F}(G)$. By Lemma 2.1, we may assume that $\boldsymbol{O}_p(G) = 1$. Without loss of generality, we may assume $o(x) = p$.

If x does not centralize $\boldsymbol{O}_{p'}(G)$, then the result follows by Theorem 3.2. Therefore, we may assume $x \in \boldsymbol{C}_G(\boldsymbol{F}(G))$ and thus G is nonsolvable. Since x is not in $\boldsymbol{O}_p(G)$, x acts faithfully on $\boldsymbol{F}^*(G)$ and therefore it must act faithfully on some subgroup K, which is a central product of quasisimple components (each with order divisible by p). We may assume that K is a minimal such subgroup, which implies that G acts transitively on the components of K. Note that $\boldsymbol{O}_p(K) = 1$.

We can further assume that $G = K\boldsymbol{C}_G(x)$ since both G and $K\boldsymbol{C}_G(x)$ contain the same number of p-elements commuting with x. Therefore, $\boldsymbol{C}_G(x)$ acts transitively on the components of K, so either

(a) every orbit of x on the components of K has size p; or

(b) x normalizes each component of K, inducing the same automorphism (up to conjugacy) on each component.

For each $y \in \boldsymbol{C}_G(x)_p$ it suffices to show that the proportion of p-elements in the coset Ky which commute with x is at most $1/p$. Fix such an element y and observe that we may assume that $G = \langle K, x, y\rangle$.

In addition, by repeating the argument above, we can reduce to the case where $\langle x, y\rangle$ acts transitively on the components of K. We now consider cases (a) and (b) in turn.

First assume (a) holds, so x does not normalize any component of K. Let $z \in Ky$ be a nontrivial p-element. Then by Proposition 3.6, the proportion of elements in z^K commuting with x is at most $1/(p+1)$. Therefore, if $Ky \neq K$ then the proportion of p-elements in Ky which commute with x is at most $1/(p+1)$. Similarly, if $Ky = K$ then Lemma 3.7 implies that $|K_p| \geqslant p^2$ and by expressing K_p as a union of K-classes we quickly deduce that $|\boldsymbol{C}_K(x)_p|/|K_p| \leqslant 1/p$ as required.

Finally, let us assume (b) holds, in which case y must act transitively on the components of K. If K has two or more components, then y is nontrivial and we can just interchange x and y in the argument above (noting that any element in Ky still acts transitively on the set of components). This allows us to reduce to the case where K is quasisimple. The result now follows by applying Propositions 3.10(ii)(b) and 3.12, except for the case where $K/\boldsymbol{Z}(K) = A_n$ is an alternating group, $p = 2$ and x acts as a transposition on K. In this case, set $L = \langle K, x\rangle \cong S_n$ and note that it suffices to show that the proportion of 2-elements in the coset Ly which commute with x is at most $\frac{1}{2}$. This follows from Proposition 3.12(iii). □

Theorem D now follows by combining Theorem C with the following result.

Proposition 3.15. *Let G be a finite group and let $x \in \boldsymbol{O}_p(G) \setminus \boldsymbol{Z}(\boldsymbol{O}_p(G))$. Then*

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} \leqslant \frac{1}{p}.$$

Proof. Set $Q = \boldsymbol{O}_p(G)$ and note that we may assume $G = Q\boldsymbol{C}_G(x)$. Let $y \in \boldsymbol{C}_G(x)$ be a p-element and note that $(Qy)_p = Qy$. Then the number of elements in the coset Qy commuting with x is equal to $|\boldsymbol{C}_Q(x)|$, which is at most $|Q|/p$ since $x \notin \boldsymbol{Z}(Q)$. Therefore, the proportion of p-elements in Qy commuting with x is at most $1/p$ and the result follows. □

To close this section, we present a family of examples to show that there exist finite groups G with a p-element x such that

$$\frac{1}{p} < \frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} < 1.$$

Note that Theorem D implies that such an element x must be in $\boldsymbol{Z}(\boldsymbol{O}_p(G))$. In fact, our examples have the property that this ratio tends to 1 as $|G|$ tends to infinity.

We consider the cases p odd and $p = 2$ separately.

Example 3.16. Fix an odd prime p and consider the semidirect product $H = A{:}B$, where $A = (C_p)^3$ is elementary abelian and a generator b for $B = C_p$ acts on A with a single Jordan block. Let $a \in A$ be a generator for A as a module for B. Note that A contains a normal subgroup K of H with $|K| = p^2$. Fix an element $x \in K \setminus H$.

Let r be a prime with $r \equiv 1 \pmod{p}$ and fix a scalar $\mu \in \mathbb{F}_r^\times$ of order p. Let $V = (\mathbb{F}_r)^p$ be a p-dimensional vector space over $\mathbb{F}_r$ and consider the semidirect product $G = V{:}H$, where K acts trivially on V, a acts as $(\mu, \dots, \mu)$ and b acts via $(1, \mu, \dots, \mu^{p-1})$. We now compute

$$|\boldsymbol{C}_G(x)_p| = (p^3 - p^2)r^p + p^2, \quad |G_p| = (p^3 - p^2)r^p + (p^4 - p^3)r^{p-1} + p^2. \tag{9}$$

This follows by counting the p-elements in each coset Vh of V, noting that if h centralizes x, then the entire coset does as well. Here it is also helpful to observe that $|(Vh)_p| = |h^V|$, where $|h^V| = r^p$ if $h \in A \setminus K$ and $|h^V| = r$ for $h \in H \setminus A$.

Finally, let us observe that both expressions in (9) are polynomials in r of degree p, with the same leading coefficient, whence $|\boldsymbol{C}_G(x)_p|/|G_p|$ tends to 1 as r tends to infinity.

Similarly, we can present a family of examples for $p = 2$.

Example 3.17. Let $H = \langle a, b\rangle = D_{16}$, where $o(a) = 8$ and $o(b) = 2$. Fix an odd prime r and let V be a 2-dimensional vector space over $\mathbb{F}_r$. Consider the semidirect product $G = V{:}H$, where a acts as $(-1, -1)$ on V and b acts as $(-1, 1)$. Let $x \in H$ be an element of order 4. Since

$$|\boldsymbol{C}_G(x)_2| = 4r^2 + 4, \quad |G_2| = 4r^2 + 8r + 4,$$

we conclude that the ratio $|\boldsymbol{C}_G(x)_2|/|G_2|$ tends to 1 as r tends to infinity.

4. Proof of Theorem A

In this section we prove Theorem A. We begin with some general observations. As always, G is a finite group and p is a prime. As in Section 1, we set

$$f(p) = \frac{p^2 + p - 1}{p^3}.$$

Lemma 4.1. *If G/N is a p'-group, then $\mathrm{Pr}_p(G) = \mathrm{Pr}_p(\langle G_p\rangle) = \mathrm{Pr}_p(N)$.*

Proof. This is clear because $G_p = \langle G_p\rangle_p = N_p$. □

We need the following elementary generalization of Gustafson's theorem [1973] on the commuting probability $\mathrm{Pr}(G)$. This result explains the presence of the term $f(p)$ in Theorem A and it extends [Lescot 1995, Lemma 1.3].

Lemma 4.2. *Let p be the smallest prime divisor of $|G|$. If $\mathrm{Pr}(G) > f(p)$, then G is abelian.*

Proof. Recall that $\mathrm{Pr}(G) = k(G)/|G|$, where $k(G)$ is the number of conjugacy classes in G; see [Gustafson 1973]. Seeking a contradiction, let us assume G is nonabelian. Let $K_1, \dots, K_r$ be the noncentral conjugacy classes of G and note that $|K_i| \geqslant p$ for every i, so we have

$$|G| = |\boldsymbol{Z}(G)| + \sum_{i=1}^{r} |K_i| \geqslant |\boldsymbol{Z}(G)| + rp$$

and thus $r \leqslant (|G| - |\boldsymbol{Z}(G)|)/p$. Therefore,

$$k(G) = |\boldsymbol{Z}(G)| + r \leqslant \left(\frac{p-1}{p}\right)|\boldsymbol{Z}(G)| + \frac{|G|}{p}$$

and thus

$$\mathrm{Pr}(G) \leqslant \frac{(p-1)/p}{|G:\boldsymbol{Z}(G)|} + \frac{1}{p} \leqslant \frac{(p-1)/p}{p^2} + \frac{1}{p} = f(p),$$

where we have used the fact that $G/\boldsymbol{Z}(G)$ is not cyclic in the last inequality. This is a contradiction. □

We now prove Theorem A. Note that if $\boldsymbol{F}^*(G)$ is a p'-group, then the proof does not require the classification of finite simple groups.

Proof of Theorem A. Let G be a finite group and let P be a Sylow p-subgroup of G. As previously noted, if P is both normal and abelian, then $\mathrm{Pr}_p(G) = 1$.

Now assume $\mathrm{Pr}_p(G) > f(p)$. We need to show that P is a normal abelian subgroup of G. To do this, we first use induction on $|G|$ to show that P is normal.

By Lemma 2.1, $\mathrm{Pr}_p(G/\boldsymbol{O}_p(G)) > f(p)$. If $\boldsymbol{O}_p(G) \neq 1$, then the inductive hypothesis implies that $G/\boldsymbol{O}_p(G)$ has a normal Sylow p-subgroup, so $\boldsymbol{O}_p(G)$ is a Sylow p-subgroup of G and we are done. Now assume $\boldsymbol{O}_p(G) = 1$. By Theorem C, we have

$$\mathrm{Pr}_p(G) = \frac{1}{|G_p|^2} \sum_{x \in G_p} |\boldsymbol{C}_G(x)_p| \leqslant \frac{1}{|G_p|}\left(1 + \frac{|G_p| - 1}{p}\right). \tag{10}$$

Consider the real-valued function

$$\varphi_p(x) = \frac{1}{x}\left(1 + \frac{x-1}{p}\right) = \frac{1}{x}\left(1 - \frac{1}{p}\right) + \frac{1}{p},$$

which is a decreasing function for $x > 0$. Seeking a contradiction, assume that P is not normal. Then $|G_p| \geqslant p^2$ (this is clear if $|P| \geqslant p^2$, and for $|P| = p$ it follows from the fact that G has at least $p+1$ Sylow p-subgroups by Sylow's theorem). Hence,

$$\varphi_p(|G_p|) \leqslant \varphi_p(p^2) = \frac{1}{p^2}\left(1 - \frac{1}{p}\right) + \frac{1}{p} = f(p)$$

and we conclude that

$$\mathrm{Pr}_p(G) \leqslant \varphi_p(|G_p|) \leqslant f(p),$$

a contradiction. Therefore, P is a normal subgroup of G.

Finally, Lemma 4.1 yields

$$\mathrm{Pr}(\boldsymbol{O}_p(G)) = \mathrm{Pr}_p(\boldsymbol{O}_p(G)) = \mathrm{Pr}_p(G) > f(p)$$

and thus Lemma 4.2 implies that $\boldsymbol{O}_p(G)$ is abelian. □

5. Proof of Theorem B

In this final section we determine the finite groups G with $\mathrm{Pr}_p(G) = f(p)$, which will allow us to prove Theorem B as a special case. We will need the following auxiliary result.

Lemma 5.1. *Let p be a prime and let G be a finite group such that $G = \boldsymbol{O}^{p'}(G)$ and G has a Sylow p-subgroup of order p. If $\mathrm{Pr}_p(G) = f(p)$, then either*

(i) *G is isomorphic to $\mathrm{PSL}_2(p)$ or $\mathrm{SL}_2(p)$; or*

(ii) *$p = 2^r - 1 \geqslant 7$ is a Mersenne prime and $G = (C_2)^r{:}C_p$, where C_p acts as a Singer cycle on $(C_2)^r$.*

Proof. If $x \in G$ is a nontrivial p-element, then $|\boldsymbol{C}_G(x)_p| = p$ and we easily deduce that $\mathrm{Pr}_p(G) = f(p)$ if and only if $|G_p| = p^2$, or equivalently if G has precisely $p+1$ Sylow p-subgroups. Let P be a Sylow p-subgroup and let K be the largest normal subgroup of G normalizing each Sylow p-subgroup of G. Then K is a p'-group and so $[K, P] \leqslant K \cap P = 1$. Therefore, K is centralized by every p-element in G, so the condition $G = \boldsymbol{O}^{p'}(G)$ implies that $K \leqslant \boldsymbol{Z}(G)$ and G/K is a doubly transitive subgroup of S_{p+1}. Moreover, each point stabilizer in this action is the normalizer of a Sylow p-subgroup and thus $|G/K| \leqslant p(p^2-1)$.

If $p = 2$, then $G/K \cong S_3$ and it is easy to check that $G = S_3 \cong \mathrm{PSL}_2(2)$ is the only possibility. Similarly, if $p = 3$ then $G/K \cong A_4$ and $G = A_4 \cong \mathrm{PSL}_2(3)$ or $\mathrm{SL}_2(3) \cong Q_8{:}C_3$ are the only options. For the remainder we may assume that $p \geqslant 5$.

Suppose p is not a Mersenne prime. Then G/K is nonsolvable and by inspecting the list of doubly transitive groups [Cameron 1981, Theorem 5.3] we see that $G/K \cong \mathrm{PSL}_2(p)$. Since $\mathrm{PSL}_2(p)$ is perfect and K is central in G, by considering the Schur multiplier of $\mathrm{PSL}_2(p)$ we deduce that $G = \mathrm{PSL}_2(p)$ or $\mathrm{SL}_2(p)$.

Finally, let us assume $p = 2^r - 1 \geqslant 7$ is a Mersenne prime, so r is an odd prime. If G/K is almost simple, we deduce as above that $G = \mathrm{PSL}_2(p)$ or $\mathrm{SL}_2(p)$. The other possibility is that G/K has a normal elementary abelian 2-subgroup of order $p+1 = 2^r$. Thus $G/K \leqslant \mathrm{AGL}_r(2)$. Here each element in $\mathrm{AGL}_r(2)$ of order p corresponds to a Singer cycle in $\mathrm{GL}_r(2)$ and by considering the overgroups of such elements (noting that G/K is generated by p-elements and that a point stabilizer has order at most $p(p-1)$) we deduce that $G/K = (C_2)^r{:}C_p$. Since G is the normal closure of P, it follows that K is a 2-group. But since r is an odd prime, we deduce that $K = 1$ is the only possibility. □

We can now classify all the finite groups with $\mathrm{Pr}_p(G) = f(p)$, which yields Theorem B as an immediate corollary.

Theorem 5.2. *Let p be a prime and G a finite group with $G = \boldsymbol{O}^{p'}(G)$ and $\mathrm{Pr}_p(G) = f(p)$. Let $Q = \boldsymbol{O}_p(G)$ and let k be a positive integer. Then one of the following holds*:

(i) *G is a p-group with $|G : \boldsymbol{Z}(G)| = p^2$.*

(ii) *$p \geqslant 5$, Q is abelian and $G = \mathrm{SL}_2(p) \times Q$ or $\mathrm{PSL}_2(p) \times Q$.*

(iii) *$p = 2^r - 1 \geqslant 3$ is a Mersenne prime and $G = (C_2)^r{:}C_{p^k} \times A$, where C_{p^k} acts as a Singer cycle of order p on $(C_2)^r$ and $A \leqslant Q$ is abelian.*

(iv) *$p = 3$ and $G = Q_8{:}C_{3^k} \times A$, where $A \leqslant Q$ is abelian.*

(v) *$p = 2$ and $G = C_3{:}C_{2^k} \times A$, where $A \leqslant Q$ is abelian.*

Proof. Set $Q = \boldsymbol{O}_p(G)$. If G/Q is abelian, then G is a p-group and by arguing as in the proof of Lemma 4.2, we see that $\mathrm{Pr}_p(G) = f(p)$ if and only if $|G : \boldsymbol{Z}(G)| = p^2$.

For the remainder, we may assume G/Q is nonabelian and thus $\mathrm{Pr}_p(G) = \mathrm{Pr}_p(G/Q)$ by Lemma 2.1. This implies that if $x, y \in G_p$ commute modulo Q, then $[x, y] = 1$, which in turn implies that $Q \leqslant \boldsymbol{Z}(G)$.

First assume that $Q = 1$. Let P be a Sylow p-subgroup of G and let n_p be the number of Sylow p-subgroups of G. As noted in the proof of Theorem A (see (10)), we have

$$\mathrm{Pr}_p(G) \leqslant \frac{1}{|G_p|}\left(1 + \frac{|G_p| - 1}{p}\right).$$

If $|G_p| > p^2$ then

$$\mathrm{Pr}_p(G) \leqslant \frac{p+1}{p^2+1} < \frac{p^2+p-1}{p^3}$$

so we may assume $|G_p| \leqslant p^2$ and thus P is abelian. If $|P| = p^2$, then P is normal and abelian and so $\mathrm{Pr}_p(G) = 1$, a contradiction. So we can reduce further to the case where $|P| = p$ and $n_p = p + 1$. Now apply Lemma 5.1 to conclude.

Finally, let us assume $Q \neq 1$ and note that G/Q is one of the groups described in Lemma 5.1. First assume G/Q is nonsolvable, so $p \geqslant 5$ and G is a central extension of $\mathrm{PSL}_2(p)$ or $\mathrm{SL}_2(p)$ by Q. Here (ii) holds since every p-central extension of one of these groups is split.

Suppose that p is a Mersenne prime with $p \geqslant 7$. Let T be a Sylow 2-subgroup of G. Then T is elementary abelian and $K := TQ = T \times Q$. Thus, T is normal in G and $G = TP$ with P a Sylow p-subgroup of G. Since P/K has order p, P is abelian and so $P = \langle x \rangle \times A$ where A is central and x induces an automorphism of order p on T, whence (iii) holds. If $p = 3$, the same argument applies except that T is either elementary abelian of order 4 or a quaternion group of order 8, leading to (iv).

If $p = 2$, the same argument applies with $T \cong C_3$ a Sylow 3-subgroup of G. This leads to (v). □

Corollary 5.3. *Let G be a finite group such that* $\mathrm{Pr}_p(G) = f(p)$.

(i) *If $p = 2$, then G is solvable and $\boldsymbol{O}^{2'}(G)$ is metabelian.*

(ii) *If $p = 3$, then $\boldsymbol{O}^{3'}(G)$ is solvable.*

Finally, we turn to the asymptotic behavior of $\mathrm{Pr}_p(G)$ with respect to a fixed prime p and a sequence of simple groups of order divisible by p. Set

$$f_p(G) = \max\{f_p(x) : 1 \neq x \in G_p\},$$

where $f_p(x) = |\boldsymbol{C}_G(x)_p|/|G_p|$. Note that

$$\mathrm{Pr}_p(G) = \frac{1}{|G_p|}\sum_{x \in G_p} f_p(x) \leqslant \frac{1}{|G_p|} + \left(1 - \frac{1}{|G_p|}\right) f_p(G).$$

Proposition 5.4. *Fix a prime p and let $G = A_n$ be the alternating group of degree n. Then* $\mathrm{Pr}_p(G) \to 0$ *as $n \to \infty$.*

Proof. Since $|G_p|$ tends to infinity with n, it suffices to show that $f_p(G)$ tends to 0. Let $y \in S_n$ be a nontrivial p-element. It is a straightforward exercise to check that for n large enough, $|\boldsymbol{C}_{S_n}(y)_p|$ is maximal when y is a p-cycle. Let us also observe that S_n contains an equal number of even and odd 2-elements commuting with a given 2-element $z \in S_n$ (this is because $\boldsymbol{O}_2(\boldsymbol{C}_{S_n}(z))$ contains odd permutations when z is nontrivial). Therefore, if n is large enough we have

$$f_p(G) \leqslant \frac{|\boldsymbol{C}_G(x)_p|}{|G_p|}$$

with $x = (1, \dots, p) \in S_n$ a p-cycle. For each integer $p < j \leqslant n$, let $y_j \in S_n$ be a p-cycle with an orbit $\{1, \dots, p-1, j\}$ and let Z_j be the set of p-elements in $\boldsymbol{C}_{S_n}(y_j)$ that act nontrivially on $\{1, \dots, p-1, j\}$. Note that the Z_j are pairwise disjoint.

If p is odd, then $|Z_j| = (p-1)!|(A_{n-p})_p|$ and we have $|\boldsymbol{C}_G(x)_p| = p|Z_j|/(p-1)! \geqslant 2|Z_j|/3$, whence

$$\frac{|\boldsymbol{C}_G(x)_p|}{|G_p|} \leqslant \frac{2}{3(n-p)}$$

and this upper bound tends to 0 as n tends to infinity. Similarly, if $p = 2$ then

$$|Z_j| = |(S_{n-2})_2| = |\boldsymbol{C}_G(x)_2|$$

and the result follows. □

It is possible to establish an analogous result for simple groups of Lie type, but the details are more complicated and they will be given elsewhere. Here we just sketch some of the main ideas. Fix a prime p. Let G be a simple group of Lie type over $\mathbb{F}_q$ of (untwisted) rank r and assume p divides $|G|$. As before, it suffices to show that $f_p(G) \to 0$ as $|G| \to \infty$.

First suppose that q is increasing. Let $x \in G$ be a nontrivial p-element such that $f_p(x) = f_p(G)$ and note that we may assume x has order p. Let $y \in G$ be a nontrivial p-element and observe that

$$\frac{|y^G \cap \boldsymbol{C}_G(x)|}{|y^G|}$$

is the probability that x commutes with a random conjugate of y. By the main theorem of [Liebeck and Saxl 1991], this ratio goes to 0 as q tends to infinity. Since this is true for every nontrivial conjugacy of p-elements, and since the number of p-elements in G tends to infinity as q increases (recall that we are assuming p divides $|G|$), we conclude that $f_p(G) \to 0$.

Now suppose q is fixed and r is increasing, so we may assume G is a classical group and we note that p divides $|G|$ if $r \geqslant p$. First assume p divides q, so we are considering unipotent elements. By a result of Steinberg (see [Liebeck and Seitz 2012, Lemma 2.16], for example) we have $|G_p| = q^{\dim X - r}$, where X is the ambient simple algebraic group. By inspecting [loc. cit.], it is easy to see that $|\boldsymbol{C}_G(x)_p|$ is maximal when x is a long root element and the result follows easily.

Finally, let us assume that p does not divide q and so x is a semisimple element. This situation is somewhat more complicated, but there are several ways to proceed and much stronger results can be established. For example, [Burness et al. 2020, Theorem 16] implies that if p is odd and $r > 2$ then the probability that two random elements of order p generate G tends to 1 as $|G|$ tends to infinity (in particular, the probability that two such elements commute tends to 0). With some additional work, this can be extended to p-elements, including the case $p = 2$ (of course, a pair of involutions will not generate G, but the probability that they commute still goes to 0 as r increases). This stronger result implies that $\mathrm{Pr}_p(G) \to 0$ as r tends to infinity.

It is interesting to consider some extensions of this problem. For example, suppose G is a finite group such that $\boldsymbol{O}_p(G) = 1$ and $G = \boldsymbol{O}^{p'}(G)$. Do we have $\mathrm{Pr}_p(G) \to 0$ as $|G| \to \infty$?

References

[Antolín et al. 2017] Y. Antolín, A. Martino, and E. Ventura, "Degree of commutativity of infinite groups", *Proc. Amer. Math. Soc.* **145**:2 (2017), 479–485. MR

[Burness and Guralnick 2022] T. C. Burness and R. M. Guralnick, "Fixed point ratios for finite primitive groups and applications", *Adv. Math.* **411**:part A (2022), Paper No. 108778, 90. MR Zbl

[Burness et al. 2020] T. C. Burness, S. Gerhardt, and R. M. Guralnick, "Topological generation of exceptional algebraic groups", *Adv. Math.* **369** (2020), 107177, 50. MR Zbl

[Cameron 1981] P. J. Cameron, "Finite permutation groups and finite simple groups", *Bull. London Math. Soc.* **13**:1 (1981), 1–22. MR Zbl

[Carlson et al. 2016] J. F. Carlson, E. M. Friedlander, and J. Pevtsova, "Vector bundles associated to Lie algebras", *J. Reine Angew. Math.* **716** (2016), 147–178. MR

[Eberhard 2015] S. Eberhard, "Commuting probabilities of finite groups", *Bull. Lond. Math. Soc.* **47**:5 (2015), 796–808. MR Zbl

[Fulman and Guralnick 2018] J. Fulman and R. Guralnick, "Enumeration of commuting pairs in Lie algebras over finite fields", *Ann. Comb.* **22**:2 (2018), 295–316. MR Zbl

[GAP 2020] The GAP Group, "GAP: groups, algorithms, and programming", 2020, available at http://www.gap-system.org. Version 4.11.0.

[Ginzburg 2000] V. Ginzburg, "Principal nilpotent pairs in a semisimple Lie algebra, I", *Invent. Math.* **140**:3 (2000), 511–561. MR Zbl

[Guralnick and Robinson 2006] R. M. Guralnick and G. R. Robinson, "On the commuting probability in finite groups", *J. Algebra* **300**:2 (2006), 509–528. MR Zbl

[Gustafson 1973] W. H. Gustafson, "What is the probability that two group elements commute?", *Amer. Math. Monthly* **80** (1973), 1031–1034. MR Zbl

[Isaacs 2008] I. M. Isaacs, *Finite group theory*, Graduate Studies in Mathematics **92**, American Mathematical Society, Providence, RI, 2008. MR Zbl

[Lescot 1995] P. Lescot, "Isoclinism classes and commutativity degrees of finite groups", *J. Algebra* **177**:3 (1995), 847–869. MR Zbl

[Liebeck and Saxl 1991] M. W. Liebeck and J. Saxl, "Minimal degrees of primitive permutation groups, with an application to monodromy groups of covers of Riemann surfaces", *Proc. London Math. Soc.* (3) **63**:2 (1991), 266–314. MR Zbl

[Liebeck and Seitz 2012] M. W. Liebeck and G. M. Seitz, *Unipotent and nilpotent classes in simple algebraic groups and Lie algebras*, Mathematical Surveys and Monographs **180**, American Mathematical Society, Providence, RI, 2012. MR Zbl

[Motzkin and Taussky 1955] T. S. Motzkin and O. Taussky, "Pairs of matrices with property L, II", *Trans. Amer. Math. Soc.* **80** (1955), 387–401. MR Zbl

[Neumann 1989] P. M. Neumann, "Two combinatorial problems in group theory", *Bull. London Math. Soc.* **21**:5 (1989), 456–458. MR Zbl

[Premet 2003] A. Premet, "Nilpotent commuting varieties of reductive Lie algebras", *Invent. Math.* **154**:3 (2003), 653–683. MR Zbl

[Richardson 1979] R. W. Richardson, "Commuting varieties of semisimple Lie algebras and algebraic groups", *Compositio Math.* **38**:3 (1979), 311–327. MR Zbl

[Tointon 2020] M. C. H. Tointon, "Commuting probabilities of infinite groups", *J. Lond. Math. Soc.* (2) **101**:3 (2020), 1280–1297. MR Zbl

Communicated by Michael J. Larsen
Received 2021-12-16 Revised 2022-04-06 Accepted 2022-07-06

t.burness@bristol.ac.uk *School of Mathematics, University of Bristol, Bristol, United Kingdom*

guralnic@usc.edu *Department of Mathematics, University of Southern California, Los Angeles, CA, United States*

alexander.moreto@uv.es *Departament de Matemàtiques, Universitat de València, Burjassot, València, Spain*

gabriel@uv.es *Departament de Matemàtiques, Universitat de València, Burjassot, València, Spain*

ALGEBRA AND NUMBER THEORY 17:6 (2023)
https://doi.org/10.2140/ant.2023.17.1231

Correction to the article Height bounds and the Siegel property

Martin Orr and Christian Schnell

This is a correction to the paper "Height bounds and the Siegel property" (*Algebra Number Theory* **12**:2 (2018), 455–478). We correct an error in the proof of Theorem 4.1. Theorem 4.1 as stated in the original paper is correct, but the correction affects additional information about the theorem which is important for applications.

There is an error in the proof of [Orr 2018, Theorem 4.1]. The statement of Theorem 4.1 is correct, but [loc. cit., Lemma 4.4] is incorrect under the conditions on $K_{\boldsymbol{G}}$ stated above it.

Subsequent applications [Bakker et al. 2020, Theorem 1.1(2); Daw and Orr 2021, Lemma 2.3] have required greater control of the maximal compact subgroup $K_{\boldsymbol{G}}$ than is given by the statement of [Orr 2018, Theorem 4.1]. As a result of the error in the proof, the choice of $K_{\boldsymbol{G}}$ is more constrained than it appears in [loc. cit.]. We therefore state a version of [loc. cit., Theorem 4.1], extended to correctly describe the constraints on $K_{\boldsymbol{G}}$.

Theorem 1. *Let $\boldsymbol{G}$ and $\boldsymbol{H}$ be reductive $\mathbb{Q}$-algebraic groups, with $\boldsymbol{H} \subset \boldsymbol{G}$. Let $\mathfrak{S}_{\boldsymbol{H}}$ be a Siegel set in $\boldsymbol{H}(\mathbb{R})$ with respect to the Siegel triple $(\boldsymbol{P_H}, \boldsymbol{S_H}, K_{\boldsymbol{H}})$. Let $K_{\boldsymbol{G}} \subset \boldsymbol{G}(\mathbb{R})$ be a maximal compact subgroup such that*

(i) *$K_{\boldsymbol{H}} \subset K_{\boldsymbol{G}}$; and*

(ii) *the Cartan involution of $\boldsymbol{G}$ associated with $K_{\boldsymbol{G}}$ stabilises $\boldsymbol{S_H}$.*

Then there exist subgroups $\boldsymbol{P_G}, \boldsymbol{S_G} \subset \boldsymbol{G}$ forming a Siegel triple $(\boldsymbol{P_G}, \boldsymbol{S_G}, K_{\boldsymbol{G}})$, a Siegel set $\mathfrak{S}_{\boldsymbol{G}} \subset \boldsymbol{G}(\mathbb{R})$ with respect to this Siegel triple, and a finite set $C \subset \boldsymbol{G}(\mathbb{Q})$ such that

$$\mathfrak{S}_{\boldsymbol{H}} \subset C.\mathfrak{S}_{\boldsymbol{G}}.$$

Furthermore, $R_u(\boldsymbol{P_H}) \subset R_u(\boldsymbol{P_G})$ and $\boldsymbol{S_H} = \boldsymbol{S_G} \cap \boldsymbol{H}$.

Remark 2. In the setting of Theorem 1, let Θ be the Cartan involution of $\boldsymbol{G}$ associated with $K_{\boldsymbol{G}}$. We now compare (ii) with:

(ii′) Θ stabilises $\boldsymbol{H}$.

MSC2020: primary 11F06; secondary 11G18.
Keywords: reduction theory, Siegel sets.

If (i) and (ii′) are satisfied, then the restriction $\Theta_{|\boldsymbol{H}}$ is the Cartan involution of $\boldsymbol{H}$ associated with $K_{\boldsymbol{H}}$. Hence, by the definition of Siegel triple, (ii) is satisfied. However, if (i) and (ii) are satisfied, then (ii′) does not necessarily hold. This may be seen in the example $\boldsymbol{G} = \mathrm{SL}_2$, $\boldsymbol{H} = \left\{\left(\begin{smallmatrix} a & b \\ db & a \end{smallmatrix}\right) : a^2 - db^2 = 1\right\}$ where d is a nonsquare positive rational number, $K_{\boldsymbol{G}} = \mathrm{SO}_2(\mathbb{R})$, $\boldsymbol{S_H} = \{1\}$, $K_{\boldsymbol{H}} = \{1\}$.

In this note, we explain how to correct the proof of [Orr 2018, Theorem 4.1] and prove Theorem 1. We also give examples showing that condition (ii) of Theorem 1 cannot be deleted from the statement of the theorem: first an example in which $\boldsymbol{H}$ is a torus, then a more sophisticated example in which $\boldsymbol{H}$ is semisimple. At the end of the note, we correct some unrelated minor errors in [loc. cit.].

A. ***Correction to proof of*** **[Orr 2018, Theorem 4.1].** On [Orr 2018, page 470], item (2) (the choice of $K_{\boldsymbol{G}}$) should be replaced by:

(2) $K_{\boldsymbol{G}}$, a maximal compact subgroup of $\boldsymbol{G}(\mathbb{R})$ containing $K_{\boldsymbol{H}}$, such that the Cartan involution of $\boldsymbol{G}$ associated with $K_{\boldsymbol{G}}$ stabilises $\boldsymbol{S_H}$.

Paragraph 1 of the proof of [loc. cit., Lemma 4.4] is incorrect: neither the original constraint on $K_{\boldsymbol{G}}$, nor the corrected constraint, are sufficient to guarantee that Θ restricts to an involution of $\boldsymbol{H}$ (see Remark 2). With the corrected constraint, that paragraph can be ignored and paragraph 2 of the proof of [loc. cit., Lemma 4.4] is valid. Hence the lemma is true under the corrected constraint on $K_{\boldsymbol{G}}$.

The remainder of the proof of [loc. cit., Theorem 4.1] is valid without any changes related to the choice of $K_{\boldsymbol{G}}$ (but see unrelated minor corrections in Section E of this note). No further conditions are imposed on $K_{\boldsymbol{G}}$, so this proves Theorem 1.

In order to establish [loc. cit., Theorem 4.1], it is necessary to verify the existence of $K_{\boldsymbol{G}}$ satisfying (2) above. To show this, choose a faithful representation $\rho \colon \boldsymbol{G}_{\mathbb{R}} \to \mathrm{GL}(V)$ for some real vector space V. By [Mostow 1955, Theorem 7.3], there exists a positive definite symmetric form ψ on V with respect to which the groups $K_{\boldsymbol{H}} \subset \boldsymbol{H}(\mathbb{R}) \subset \boldsymbol{G}(\mathbb{R}) \subset \mathrm{GL}(V)$ are simultaneously self-adjoint. In other words, if Θ denotes the Cartan involution of $\mathrm{GL}(V)$ associated with ψ, then Θ restricts to Cartan involutions of $\boldsymbol{G}$, $\boldsymbol{H}$ and $K_{\boldsymbol{H}}$.

Letting $K_{\boldsymbol{G}}$ denote the stabiliser of ψ in $\boldsymbol{G}(\mathbb{R})$, we obtain $K_{\boldsymbol{H}} \subset K_{\boldsymbol{G}}$. By Remark 2, Θ stabilises $\boldsymbol{S_H}$.

B. ***Counterexample in which condition (ii) of Theorem 1 is not satisfied: a torus.*** Let $\boldsymbol{G} = \mathrm{SL}_2$ and let $(\boldsymbol{P}_0, \boldsymbol{S}_0, K_{\boldsymbol{G}})$ be the standard Siegel triple for $\boldsymbol{G}$, that is, $\boldsymbol{P}_0$ is the subgroup of upper triangular matrices in $\boldsymbol{G}$, $\boldsymbol{S}_0$ is the subgroup of diagonal matrices in $\boldsymbol{G}$ and $K_{\boldsymbol{G}} = \mathrm{SO}_2(\mathbb{R})$.

Let

$$\boldsymbol{H} = \left\{ \begin{pmatrix} x & x^{-1} - x \\ 0 & x^{-1} \end{pmatrix} \right\} \subset \boldsymbol{G}.$$

This is a $\mathbb{Q}$-split torus so it possesses a unique Siegel triple, namely $\boldsymbol{P_H} = \boldsymbol{S_H} = \boldsymbol{H}$, $K_{\boldsymbol{H}} = \{\pm 1\}$, and a unique Siegel set, $\mathfrak{S}_{\boldsymbol{H}} = \boldsymbol{H}(\mathbb{R})$.

Clearly $K_{\boldsymbol{H}} = \{\pm 1\} \subset K_{\boldsymbol{G}}$. Thus $K_{\boldsymbol{G}}$ satisfies condition (i) of Theorem 1. However by [Orr 2018, Lemma 2.1], $\boldsymbol{S}_0$ is the only $\mathbb{Q}$-split torus in $\boldsymbol{P}_0$ stabilised by the Cartan involution of $\boldsymbol{G}$ associated with $K_{\boldsymbol{G}}$.

Hence this Cartan involution does not stabilise $\boldsymbol{S_H}$. In other words, $K_{\boldsymbol{G}}$ does not satisfy condition (ii) of Theorem 1.

Now we shall show that this $\mathfrak{S}_{\boldsymbol{H}}$ and $K_{\boldsymbol{G}}$ do not satisfy the conclusion of Theorem 1. Suppose for contradiction that there exist subgroups $\boldsymbol{P_G}, \boldsymbol{S_G} \subset \boldsymbol{G}$ forming a Siegel triple $(\boldsymbol{P_G}, \boldsymbol{S_G}, K_{\boldsymbol{G}})$, a Siegel set $\mathfrak{S}_{\boldsymbol{G}} \subset \boldsymbol{G}(\mathbb{R})$ with respect to this Siegel triple, and a finite set $C \subset \boldsymbol{G}(\mathbb{Q})$ such that $\mathfrak{S}_{\boldsymbol{H}} \subset C.\mathfrak{S}_{\boldsymbol{G}}$.

By [Borel and Tits 1965, Théorème 4.13], there exists $g \in \boldsymbol{G}(\mathbb{Q})$ such that $\boldsymbol{P}_0 = g\boldsymbol{P_G}g^{-1}$. Writing $g = pk$ where $p \in \boldsymbol{P}_0(\mathbb{R})$ and $k \in K_{\boldsymbol{G}}$, $(\boldsymbol{P}_0, k\boldsymbol{S_G}k^{-1}, K_{\boldsymbol{G}})$ is a Siegel triple and $g\mathfrak{S}_{\boldsymbol{G}}$ is a Siegel set with respect to $(\boldsymbol{P}_0, k\boldsymbol{S_G}k^{-1}, K_{\boldsymbol{G}})$. Hence we can replace $\boldsymbol{P_G}$ by $\boldsymbol{P}_0$, $\boldsymbol{S_G}$ by $k\boldsymbol{S_G}k^{-1}$, $\mathfrak{S}_{\boldsymbol{G}}$ by $g\mathfrak{S}_{\boldsymbol{G}}$ and C by Cg^{-1}. We can thus assume that $\boldsymbol{P_G} = \boldsymbol{P}_0$. By the uniqueness of the torus in a Siegel triple, this implies that $\boldsymbol{S_G} = \boldsymbol{S}_0$ and $\mathfrak{S}_{\boldsymbol{G}}$ is a standard Siegel set in $\boldsymbol{G}(\mathbb{R})$.

The image of $\mathfrak{S}_{\boldsymbol{H}} = \boldsymbol{S_H}(\mathbb{R})$ in $\boldsymbol{G}(\mathbb{R})/K_0$, identified with the upper half-plane, is the ray

$$R = \{(1-y) + yi : y \in \mathbb{R}_{>0}\}.$$

Write $\mathcal{F}_{\boldsymbol{G}}$ for the image of $\mathfrak{S}_{\boldsymbol{G}}$ in the upper half-plane.

Since $R \subset C\mathcal{F}_{\boldsymbol{G}}$ and C is finite, there exists $\gamma \in C \subset \boldsymbol{G}(\mathbb{Q})$ such that $R \cap \gamma\mathcal{F}_{\boldsymbol{G}}$ contains points z where both $\operatorname{Im} z$, $|\operatorname{Re} z| \to \infty$. But this is impossible because:

(i) If $\gamma \notin \boldsymbol{P}_0(\mathbb{Q})$, then $\gamma\mathcal{F}_{\boldsymbol{G}}$ lies below a horizontal line.

(ii) If $\gamma \in \boldsymbol{P}_0(\mathbb{Q})$, then $\gamma\mathcal{F}_{\boldsymbol{G}}$ lies within a vertical strip of finite width.

C. *Counterexample in which condition (ii) of Theorem 1 is not satisfied: a semisimple subgroup.* Let $\boldsymbol{G} = \mathrm{SL}_3$ and let $(\boldsymbol{P}_0, \boldsymbol{S}_0, K_{\boldsymbol{G}})$ be the standard Siegel triple for $\boldsymbol{G}$. Let

$$\boldsymbol{H}_0 = \mathrm{SO}_3(J) \quad \text{where } J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Let Q_J denote the quadratic form on $\mathbb{R}^3$ represented by J. This form is negative definite on the 1-dimensional subspace $L = \mathbb{R}(1, 0, -1)^t \subset \mathbb{R}^3$ and positive definite on the 2-dimensional subspace $M = \mathbb{R}(0, 1, 0)^t + \mathbb{R}(1, 0, 1)^t$. Let

$$K_{\boldsymbol{H}} = \{h \in \boldsymbol{H}_0(\mathbb{R}) : h(L) = L \text{ and } h(M) = M\}.$$

This is a maximal compact subgroup of $\boldsymbol{H}_0(\mathbb{R})$ and is isomorphic to $\mathrm{O}_2(\mathbb{R})$ via restriction to its action on M.

Let $c \in \mathbb{Q} \setminus \{0, \pm 1\}$. Let $\eta \in \mathrm{GL}_3(\mathbb{Q})$ be the linear map which acts as multiplication by c on L and as the identity on M. Explicitly,

$$\eta = \begin{pmatrix} \frac{1}{2}(1+c) & 0 & \frac{1}{2}(1-c) \\ 0 & 1 & 0 \\ \frac{1}{2}(1-c) & 0 & \frac{1}{2}(1+c) \end{pmatrix}.$$

Let

$$\boldsymbol{H} = \eta\boldsymbol{H}_0\eta^{-1} = \mathrm{SO}_3(\eta J \eta^t).$$

By construction, η centralises $K_{\boldsymbol{H}}$. It follows that $\eta K_{\boldsymbol{H}} \eta^{-1} = K_{\boldsymbol{H}} = K_{\boldsymbol{G}} \cap \boldsymbol{H}(\mathbb{R})$ and $K_{\boldsymbol{H}}$ is a maximal compact subgroup of $\boldsymbol{H}(\mathbb{R})$.

Let Q_0 denote the standard quadratic form on $\mathbb{R}^3$. The spaces L and M are orthogonal with respect to Q_0 and $Q_{0\,|\,M} = Q_{J\,|\,M}$. Hence $K_{\boldsymbol{H}} \subset \mathrm{SO}_3(Q_0) = K_{\boldsymbol{G}}$. Thus condition (i) of Theorem 1 is satisfied.

Let $\boldsymbol{P_H} = \eta(\boldsymbol{P}_0 \cap \boldsymbol{H}_0)\eta^{-1}$ and $\boldsymbol{S_H} = \eta(\boldsymbol{S}_0 \cap \boldsymbol{H}_0)\eta^{-1}$. As in [Borel 1969, 11.16], $\boldsymbol{P}_0 \cap \boldsymbol{H}_0$ is a minimal $\mathbb{Q}$-parabolic subgroup of $\boldsymbol{H}_0$ so $(\boldsymbol{P_H}, \boldsymbol{S_H}, K_{\boldsymbol{H}})$ is a Siegel triple in $\boldsymbol{H}$. Let $\mathfrak{S}_{\boldsymbol{H}} = \Omega_{\boldsymbol{H}} A_{\boldsymbol{H},t} K_{\boldsymbol{H}}$ be a Siegel set in $\boldsymbol{H}(\mathbb{R})$ with respect to this Siegel triple.

We shall show that $\mathfrak{S}_{\boldsymbol{H}}$ and $K_{\boldsymbol{G}}$ do not satisfy the conclusion of Theorem 1. Suppose for contradiction that there exist subgroups $\boldsymbol{P_G}, \boldsymbol{S_G} \subset \boldsymbol{G}$ forming a Siegel triple $(\boldsymbol{P_G}, \boldsymbol{S_G}, K_{\boldsymbol{G}})$, a Siegel set $\mathfrak{S}_{\boldsymbol{G}} \subset \boldsymbol{G}(\mathbb{R})$ with respect to this Siegel triple, and a finite set $C \subset \boldsymbol{G}(\mathbb{Q})$ such that $\mathfrak{S}_{\boldsymbol{H}} \subset C.\mathfrak{S}_{\boldsymbol{G}}$. By the same argument as in Section B, we may assume that $\boldsymbol{P_G} = \boldsymbol{P}_0$ and $\boldsymbol{S_G} = \boldsymbol{S}_0$.

Let $\sigma_s = \mathrm{diag}(s, 1, s^{-1})$ for $s \in \mathbb{R}_{>0}$. Now

$$\{\eta\sigma_s\eta^{-1} : s \geq t\} = A_{\boldsymbol{H},t} \subset \mathfrak{S}_{\boldsymbol{H}} \subset C\mathfrak{S}_{\boldsymbol{G}}.$$

Since C is finite, there exists some $\gamma \in C$ such that $\gamma\mathfrak{S}_{\boldsymbol{G}}$ contains elements of the form $\eta\sigma_s\eta^{-1}$ for arbitrarily large s. Consequently $\eta^{-1}\gamma\mathfrak{S}_{\boldsymbol{G}}\eta$ contains σ_s for arbitrarily large s. Furthermore the standard Siegel set $\mathfrak{S}_{\boldsymbol{G}}$ contains $\{\sigma_s : s \geq t'\}$ for some $t' \in \mathbb{R}_{>0}$.

Let χ_1, χ_2 denote the simple roots of $\boldsymbol{G}$ with respect to $\boldsymbol{S}_0$, using the ordering induced by $\boldsymbol{P}_0$. Then $\chi_1(\sigma_s) = \chi_2(\sigma_s) = s$ so the previous paragraph shows that $\mathfrak{S}_{\boldsymbol{G}} \cap \eta^{-1}\gamma\mathfrak{S}_{\boldsymbol{G}}\eta$ contains elements σ_s with arbitrarily large values for χ_1 and χ_2. Applying Lemma 3 below (with $\Omega_{\boldsymbol{G}} = K_{\boldsymbol{G}} \cup K_{\boldsymbol{G}}\eta^{-1}$), we deduce that $\eta^{-1}\gamma$ is contained in the standard parabolic subgroup ${}_{\mathbb{Q}}\boldsymbol{P}_{0,\varnothing} = \boldsymbol{P}_0$.

Let $\boldsymbol{U}_0 = R_u(\boldsymbol{P}_0)$. Write the Iwasawa decomposition of η^{-1} as

$$\eta^{-1} = \mu\alpha\kappa \quad \text{where } \mu \in \boldsymbol{U}_0(\mathbb{R}), \alpha \in \boldsymbol{S}_0(\mathbb{R}), \kappa \in K_{\boldsymbol{G}}.$$

For arbitrarily large real numbers s, we have

$$\sigma_s\mu\sigma_s^{-1}.\sigma_s\alpha.\kappa = \sigma_s\eta^{-1} \in \mathfrak{S}_{\boldsymbol{G}}\eta^{-1} \cap \eta^{-1}\gamma\mathfrak{S}_{\boldsymbol{G}} \subset \eta^{-1}\gamma\mathfrak{S}_{\boldsymbol{G}}.$$

By the definition of Siegel sets and since $\eta^{-1}\gamma \in \boldsymbol{P}_0(\mathbb{R})$, the $\boldsymbol{U}_0(\mathbb{R})$-component in the Iwasawa decomposition of every element of $\eta^{-1}\gamma\mathfrak{S}_{\boldsymbol{G}}$ is bounded. Thus $\sigma_s\mu\sigma_s^{-1}$ lies in a bounded set for arbitrarily large real numbers s. By direct calculation, this implies that $\mu = 1$. (This is the opposite situation to [Borel 1969, Lemme 12.2], adapted to our conventions about Siegel sets.) Hence $\eta^{-1} = \alpha\kappa \in \boldsymbol{S}_0(\mathbb{R})K_{\boldsymbol{G}}$.

It follows that $\eta^t\eta = (\alpha^{-1})^t(\kappa^{-1})^t\kappa^{-1}\alpha^{-1} = \alpha^{-2}$ is diagonal. But $\eta^t\eta$ is not diagonal, as can be seen either by direct calculation or by noting that η is symmetrical so $\eta^t\eta = \eta^2$ has L as a 1-dimensional eigenspace yet L is not a coordinate axis.

D. ***Siegel sets with noncompact intersection.*** In this section, we prove a generalisation of [Borel 1969, Proposition 12.6], replacing a Siegel set $\mathfrak{S} = \Omega_{\boldsymbol{P}} A_t K$ by a set of the form $\Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}}$ where $\Omega_{\boldsymbol{G}}$ may be any compact subset of $\boldsymbol{G}(\mathbb{R})$. This generalisation was used in Section C.

Let $\boldsymbol{G}$ be a reductive $\mathbb{Q}$-algebraic group. Let $\boldsymbol{P}$ be a minimal parabolic $\mathbb{Q}$-subgroup of $\boldsymbol{G}$ and let $\boldsymbol{U}$ be the unipotent radical of $\boldsymbol{P}$. Let $\boldsymbol{S}$ be a maximal $\mathbb{Q}$-split torus in $\boldsymbol{S}$ and let $\boldsymbol{M}$ be the maximal $\mathbb{Q}$-anisotropic subgroup of $Z_{\boldsymbol{G}}(\boldsymbol{S})$. Let t be a positive real number and let A_t be the subset of $\boldsymbol{S}(\mathbb{R})$ defined in [Orr 2018, Section 2B]. Let $\mathfrak{g}$ and $\mathfrak{u}$ denote the Lie algebras of $\boldsymbol{G}$ and $\boldsymbol{U}$ respectively (over $\mathbb{R}$).

Let Δ be the set of simple roots of $\boldsymbol{G}$ with respect to $\boldsymbol{S}$, using the ordering induced by $\boldsymbol{P}$. For $\theta \subset \Delta$, let Ψ_θ denote the set of roots ϕ such that the expression of ϕ as a linear combination of elements of Δ has a positive coefficient for at least one element of θ.

For each character $\chi \in X^*(\boldsymbol{S})$, there is a unique continuous group homomorphism $\boldsymbol{P}(\mathbb{R}) \to \mathbb{R}_{>0}$, which we denote f_χ, with the properties $f_\chi(s) = |\chi(s)|$ for all $s \in \boldsymbol{S}(\mathbb{R})$ and $f_\chi = 1$ on $\boldsymbol{U}(\mathbb{R})\boldsymbol{M}(\mathbb{R})$. (This is because $\boldsymbol{S}(\mathbb{R}) \cap \boldsymbol{U}(\mathbb{R})\boldsymbol{M}(\mathbb{R})$ is finite, so $|\chi(s)| = 1$ for all $s \in \boldsymbol{S}(\mathbb{R}) \cap \boldsymbol{U}(\mathbb{R})\boldsymbol{M}(\mathbb{R})$, and $\boldsymbol{S}$ normalises $\boldsymbol{U}\boldsymbol{M}$.) Choose a maximal compact subgroup $K \subset \boldsymbol{G}(\mathbb{R})$. Then $f_\chi(\boldsymbol{P}(\mathbb{R}) \cap K)$ is a compact subgroup of $\mathbb{R}_{>0}$, so is trivial. Therefore we can extend f_χ to a continuous function $\boldsymbol{G}(\mathbb{R}) = \boldsymbol{P}(\mathbb{R})K \to \mathbb{R}_{>0}$ by setting $f_\chi(pk) = f_\chi(p)$ for all $p \in \boldsymbol{P}(\mathbb{R})$ and $k \in K$. These functions f_χ are not necessarily "of type $(\boldsymbol{P}, \chi)$" as defined in [Borel 1969, 14.1] because $\chi \in X^*(\boldsymbol{S})$ might not extend to a character of $\boldsymbol{P}$, but the argument in [loc. cit., 14.2(c)] still applies to the functions f_χ.

Lemma 3. *Let $\Omega_{\boldsymbol{P}}$ and $\Omega_{\boldsymbol{G}}$ be compact subsets of $\boldsymbol{P}(\mathbb{R})$ and $\boldsymbol{G}(\mathbb{R})$ respectively. Let $\gamma \in \boldsymbol{G}(\mathbb{R})$. If $\Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}} \cap \gamma \Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}}$ is noncompact, then γ is contained in a proper parabolic $\mathbb{Q}$-algebraic subgroup of $\boldsymbol{G}$ containing $\boldsymbol{P}$. More precisely, let*

$$\theta = \{\chi \in \Delta : f_\chi \text{ is bounded above on } \Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}} \cap \gamma \Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}}\}.$$

Then γ lies in the standard parabolic subgroup ${}_{\mathbb{Q}}\boldsymbol{P}_\theta$ in the notation of [Borel and Tits 1965, 5.12].

Proof. Let

$$\Omega = \Bigl(\bigcup_{a \in A_t} a^{-1} \Omega_{\boldsymbol{P}} a\Bigr) \Omega_{\boldsymbol{G}} \subset \boldsymbol{G}(\mathbb{R}).$$

By [Borel 1969, Lemme 12.2], Ω is compact. From the definitions, $\Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}} \subset A_t \Omega$. Hence, for all $\chi \in \Delta \setminus \theta$, f_χ is unbounded on $A_t\Omega \cap \gamma A_t \Omega$.

Let ${}_{\mathbb{Q}}\boldsymbol{U}_\theta$ denote the unipotent radical of ${}_{\mathbb{Q}}\boldsymbol{P}_\theta$ and let ${}_{\mathbb{Q}}\mathfrak{u}_\theta = \mathrm{Lie}({}_{\mathbb{Q}}\boldsymbol{U}_\theta)$. Let

$$Y = \{v \in \mathfrak{g} : (\mathrm{Ad}\, \xi_n^{-1}) v \to 0 \text{ for some sequence } (\xi_n) \text{ in } A_t\Omega \cap \gamma A_t \Omega\}.$$

Let $\langle Y \rangle$ denote the subspace of $\mathfrak{g}$ generated by Y. We shall show that

$$ {}_{\mathbb{Q}}\mathfrak{u}_\theta \subset \langle Y \rangle \subset (\mathrm{Ad}\, \gamma)\mathfrak{u}. \tag{1}$$

To prove the first inclusion of (1), note that ${}_{\mathbb{Q}}\mathfrak{u}_\theta$ is the direct sum of the root spaces $\mathfrak{u}_\phi$ for $\phi \in \Psi_{\Delta \setminus \theta}$, so it suffices to prove that $\mathfrak{u}_\phi \subset Y$ for each $\phi \in \Psi_{\Delta\setminus\theta}$.

Let $\phi \in \Psi_{\Delta\setminus\theta}$ and write ϕ as a linear combination of simple roots: $\phi = \sum_{\psi \in \Delta} m_\psi \psi$. By the definition of $\Psi_{\Delta\setminus\theta}$, there exists some $\chi \in \Delta \setminus \theta$ such that $m_\chi > 0$.

By the definition of θ, f_χ is unbounded on $\Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}} \cap \gamma \Omega_{\boldsymbol{P}} A_t \Omega_{\boldsymbol{G}} \subset A_t \Omega \cap \gamma A_t \Omega$. Choose a sequence (ξ_n) in $A_t \Omega \cap \gamma A_t \Omega$ such that $f_\chi(\xi_n) \to +\infty$. Write $\xi_n = \alpha_n \kappa_n$ where $\alpha_n \in A_t$ and $\kappa_n \in \Omega$.

The argument of [Borel 1969, 14.2(c)] shows that $f_\chi(\xi_n)/f_\chi(\alpha_n)$ is bounded both above and below independently of n. Hence

$$|\chi(\alpha_n)| = f_\chi(\alpha_n) \to +\infty.$$

Since ϕ is a positive root, $m_\psi \geq 0$ for all $\psi \in \Delta$. Since $\alpha_n \in A_t$ and $m_\chi > 0$, it follows that $\phi(\alpha_n) \to +\infty$.

Hence for every $v \in \mathfrak{u}_\phi$, we have $(\operatorname{Ad} \alpha_n^{-1})v \to 0$. Since Ω is compact, after replacing (ξ_n) by a subsequence, we may assume that κ_n converges, say to $\kappa \in \Omega$. Then $(\operatorname{Ad} \xi_n^{-1})v \to (\operatorname{Ad} \kappa)^{-1} 0 = 0$. Thus $\mathfrak{u}_\phi \subset Y$.

To prove the second inclusion of (1), consider an element $v \in Y$. Let (ξ_n) be a sequence in $A_t \Omega \cap \gamma A_t \Omega$ such that $(\operatorname{Ad} \xi_n^{-1})v \to 0$. Write $\xi_n = \gamma \beta_n \lambda_n$ with $\beta_n \in A_t$, $\lambda_n \in \Omega$. Since Ω is compact, after replacing (ξ_n) by a subsequence, we may assume that λ_n converges, say to $\lambda \in \Omega$. Then

$$(\operatorname{Ad} \beta_n^{-1})(\operatorname{Ad} \gamma^{-1})v = (\operatorname{Ad} \lambda_n)(\operatorname{Ad} \xi_n^{-1})v \to (\operatorname{Ad} \lambda)0 = 0.$$

Hence, when we decompose $(\operatorname{Ad} \gamma^{-1})v$ using the root space decomposition of $\mathfrak{g}$, nonzero components can occur only for those roots ϕ satisfying $|\phi(\beta_n)| \to +\infty$. Since $\beta_n \in A_t$, such roots ϕ must be positive roots. Thus $(\operatorname{Ad} \gamma^{-1})v \in \bigoplus_{\phi \in \Phi^+} \mathfrak{u}_\phi = \mathfrak{u}$.

We have proved both parts of (1). Passing from Lie algebras to groups, we obtain

$${}_{\mathbb{Q}}\boldsymbol{U}_\theta \subset \gamma \boldsymbol{U} \gamma^{-1} \subset \gamma \boldsymbol{P} \gamma^{-1} \subset \gamma \, {}_{\mathbb{Q}}\boldsymbol{P}_\theta \, \gamma^{-1}.$$

By [Borel and Tits 1965, Corollaire 4.5], it follows that ${}_{\mathbb{Q}}\boldsymbol{P}_\theta = \gamma \, {}_{\mathbb{Q}}\boldsymbol{P}_\theta \, \gamma^{-1}$. Since a parabolic subgroup of $\boldsymbol{G}$ is its own normaliser, we conclude that $\gamma \in {}_{\mathbb{Q}}\boldsymbol{P}_\theta(\mathbb{R})$. ☐

E. *Additional minor corrections to* [Orr 2018]. The following are additional corrections to [Orr 2018]:

- (page 461, Section 2D) (F2) should begin "For every $g \in \boldsymbol{G}(\mathbb{Q})$."
- (page 474, proof of Proposition 4.7) On the fifth line from the end, should say "$\chi_{|\boldsymbol{S_H}} \in \Phi_\alpha \cup \{0\}$." instead of "$\chi_{|\boldsymbol{S_H}} \in \Phi_\alpha$".
- (page 474, proof of Lemma 4.10) The first paragraph should say "Let $\boldsymbol{T_G}$ be a maximal $\mathbb{R}$-split torus in $\boldsymbol{G}$ which contains $\boldsymbol{S_G}$ and is stabilised by the Cartan involution of $\boldsymbol{G}$ associated with $K_{\boldsymbol{G}}$." This is necessary to apply [Borel and Tits 1965, Section 14].

Acknowledgements

Orr is grateful to Dave Witte Morris for informing him of the errors in [Orr 2018, page 474] which are corrected in Section E.

We are grateful to the referee for very careful reading of this note and for suggesting alternative arguments for Sections C and D, including a more general statement for Lemma 3 than in the original version.

References

[Bakker et al. 2020] B. Bakker, B. Klingler, and J. Tsimerman, "Tame topology of arithmetic quotients and algebraicity of Hodge loci", *J. Amer. Math. Soc.* **33**:4 (2020), 917–939. MR Zbl

[Borel 1969] A. Borel, *Introduction aux groupes arithmétiques*, Publ. Inst. Math. Univ. Strasbourg **1341**, Hermann, Paris, 1969. MR Zbl

[Borel and Tits 1965] A. Borel and J. Tits, "Groupes réductifs", *Inst. Hautes Études Sci. Publ. Math.* 27 (1965), 55–150. MR Zbl

[Daw and Orr 2021] C. Daw and M. Orr, "Unlikely intersections with $E \times$ CM curves in $\mathcal{A}_2$", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **22**:4 (2021), 1705–1745. MR Zbl

[Mostow 1955] G. D. Mostow, "Self-adjoint groups", *Ann. of Math.* (2) **62** (1955), 44–55. MR Zbl

[Orr 2018] M. Orr, "Height bounds and the Siegel property", *Algebra Number Theory* **12**:2 (2018), 455–478. MR Zbl

Communicated by Jonathan Pila
Received 2021-10-14 Revised 2022-06-22 Accepted 2022-07-06

martin.orr@manchester.ac.uk *Department of Mathematics, University of Manchester, Manchester, United Kingdom*

christian.schnell@stonybrook.edu *Department of Mathematics, Stony Brook University, Stony Brook, NY, 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 17 No. 6 2023