Pacific Journal of Mathematics

Volume 296 No. 2 October 2018

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

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

The subscription price for 2018 is US $475/year for the electronic version, and $640/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.257

CONGRUENCE SUBGROUPS AND SUPER-MODULAR CATEGORIES

PARSA BONDERSON, ERIC C. ROWELL,
ZHENGHAN WANG AND QING ZHANG

A super-modular category is a unitary premodular category with Müger center equivalent to the symmetric unitary category of super-vector spaces. Super-modular categories are important alternatives to modular categories as any unitary premodular category is the equivariantization of a either a modular or super-modular category. Physically, super-modular categories describe universal properties of quasiparticles in fermionic topological phases of matter. In general one does not have a representation of the modular group $\mathrm{SL}(2, \mathbb{Z})$ associated to a super-modular category, but it is possible to obtain a representation of the (index 3) θ-subgroup: $\Gamma_\theta < \mathrm{SL}(2, \mathbb{Z})$. We study the image of this representation and conjecture a super-modular analogue of the Ng–Schauenburg congruence subgroup theorem for modular categories, namely that the kernel of the Γ_θ representation is a congruence subgroup. We prove this conjecture for any super-modular category that is a subcategory of modular category of twice its dimension, i.e., admitting a minimal modular extension. Conjecturally, every super-modular category admits (precisely 16) minimal modular extensions and our conjecture would be a consequence.

1. Introduction

A key part of the data for a modular category $\mathcal{C}$ is the S and T matrices encoding the nondegeneracy of the braiding and the twist coefficients, respectively. We will denote by $\tilde{S}$ the unnormalized matrix obtained as the invariants of the Hopf link so that $\tilde{S}_{0,0} = 1$, while $S = \tilde{S}/D$ will denote the (unitary) normalized S-matrix where $D^2 = \dim(\mathcal{C})$ is the categorical dimension and $D > 0$. Later, we will use the same conventions for any premodular category (for which S may not be invertible). The diagonal matrix $T := \theta_i \delta_{i,j}$ has finite order (Vafa's theorem, see [Bakalov and

E. Rowell and Q. Zhang were partially supported by NSF grants DMS-1410144 and DMS-1664359, and Z. Wang by NSF grant DMS-1411212. The authors thank M. Cheng, M. Papanikolas and Z. Šunić for valuable discussions.
MSC2010: 18D10.
Keywords: spin mapping class group, super-modular category, fermionic modular category.

Kirillov 2001]) for any premodular category. For a modular category the S and T matrices satisfy (see, e.g., [Bakalov and Kirillov 2001, Theorem 3.1.7]):

(1) $S^2 = C$ where $C_{i,j} = \delta_{i,j^*}$ (so $S^4 = C^2 = I$).

(2) $(ST)^3 = \frac{D_+}{D} S^2$ where $D_+ = \sum_i \tilde{S}_{0,i}^2 \theta_i$.

(3) $TC = CT$.

These imply that from any modular category $\mathcal{C}$ of rank r (i.e., with r isomorphism classes of simple objects) one obtains a projective unitary representation of the modular group $\rho : SL(2, \mathbb{Z}) \to \mathrm{PSU}(r)$ defined on generators by $\mathfrak{s} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \to S$ and $\mathfrak{t} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \to T$ composed with the canonical projection $\pi_r : \mathrm{U}(r) \to \mathrm{PSU}(r)$. By rescaling the S and T matrices, ρ may be lifted to a linear representation of $SL(2, \mathbb{Z})$, but these lifts are not unique. This representation has topological significance: one identifies the modular group with the mapping class group $\mathrm{Mod}(\Sigma_{1,0})$ of the torus ($\mathfrak{t}$ and $\mathfrak{s}\mathfrak{t}^{-1}\mathfrak{s}^{-1}$ correspond to Dehn twists about the meridian and parallel) and this projective representation is the action of the mapping class group on the Hilbert space associated to the torus by the modular functor obtained from $\mathcal{C}$.

A subgroup $H < \mathrm{SL}(2, \mathbb{Z})$ is called a *congruence subgroup* if H contains a principal congruence subgroup $\Gamma(n) := \{A \in \mathrm{SL}(2, \mathbb{Z}) : A \equiv I \pmod{n}\}$ for some $n \geq 1$. Since $\Gamma(n)$ is the kernel of the reduction modulo n map $\mathrm{SL}(2, \mathbb{Z}) \to \mathrm{SL}(2, \mathbb{Z}/n\mathbb{Z})$, any congruence subgroup has finite index. The *level* of a congruence subgroup H is the minimal n so that $\Gamma(n) < H$. More generally, for $G < \mathrm{SL}(2, \mathbb{Z})$ we say $H < G$ is a congruence subgroup if $G \cap \Gamma(n) < H$ with the level of H defined similarly.

The connection between topology and number theory found through the representation above is deepened by the following congruence subgroup theorem:

Theorem 1.1 [Ng and Schauenburg 2010]. *Let $\mathcal{C}$ be a modular category of rank r with T-matrix of order N. Then the projective representation $\rho : \mathrm{SL}(2, \mathbb{Z}) \to \mathrm{PSU}(r)$ has* $\ker(\rho)$ *a congruence subgroup of level N.*

In particular the image of ρ factors over $\mathrm{SL}(2, \mathbb{Z}/N\mathbb{Z})$ and hence is a finite group. This fact has many important consequences: for example, it is related to rank-finiteness [Bruillard et al. 2016a] and can be used in classification problems [Bruillard et al. 2016b].

A *super-modular* category is a unitary ribbon fusion category whose Müger center is equivalent, as a unitary symmetric ribbon fusion category, to the category sVec of super-vector spaces (equipped with its unique structure as a unitary spherical symmetric fusion category). Super-modular categories (or slight variations) have been studied from several perspectives; see [Bonderson 2007; Davydov et al. 2013a; Bruillard et al. 2017; Lan et al. 2016] for a few examples. An algebraic motivation for studying these categories is the following: any unitary braided fusion category is the equivariantization [Drinfeld et al. 2010] of either a modular or super-modular category (see [Sawin 2002, Theorem 2]). Physically, super-modular categories provide a framework for studying fermionic topological phases of matter [Bruillard et al. 2017]. Topological motivations include the study of spin 3-manifold invariants [Sawin 2002; Blanchet 2005; Blanchet and Masbaum 1996] and (3+1)-TQFTs [Walker and Wang 2012].

Remark. We restrict to unitary categories both for mathematical convenience and for their physical significance. On the other hand, there is a nonunitary version sVec^- of sVec: the underlying (non-Tannakian) symmetric fusion category is the same, but with the other possible spherical structure, which leads to negative dimensions. We could define super-modular categories more generally as premodular categories $\mathcal{B}$ with Müger center equivalent to either of sVec or sVec^-. However, we do not know of any examples $\mathcal{B}$ with $\mathcal{B}' \cong \mathrm{sVec}^-$ that are not simply of the form $\mathcal{C} \boxtimes \mathrm{sVec}^-$ for some modular category $\mathcal{C}$ (A. Bruguières asked Rowell and Wang for such an example in 2016).

One interesting feature of super-modular categories $\mathcal{B}$ is that their S and T matrices have tensor decompositions [Bonderson et al. 2013, Appendix; Bruillard et al. 2017, Theorem III.5]):

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} \otimes \hat{S}, \quad T = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \otimes \hat{T}, \tag{1-1}$$

where $\hat{S}$ is unitary and $\hat{T}$ is a diagonal (unitary) matrix, depending on $r/2-1$ sign choices. Two naive questions motivated by the above are whether $\hat{S}$ and a choice of $\hat{T}$ provide a (projective) representation of $\mathrm{SL}(2,\mathbb{Z})$, and whether the group generated by $\hat{S}$ and a choice of $\hat{T}$ is finite. Of course if $\mathcal{B} = \mathrm{sVec} \boxtimes \mathcal{D}$ for some modular category $\mathcal{D}$ (*split super-modular*) then the answer to both is "yes". More generally, as Example 2.1 below illustrates, the answer to both questions is "no".

The physical and topological applications of super-modular categories motivate a more refined question as follows. The consideration of fermions on a torus [Alvarez-Gaumé et al. 1986] leads to the study of spin structures on the torus $\Sigma_{1,0}$: there are three even spin structures (A, A), (A, P), (P, A) and one odd spin structure (P, P), where A, P denote antiperiodic and periodic boundary conditions. The full mapping class group $\mathrm{Mod}(\Sigma_{1,0}) = \mathrm{SL}(2,\mathbb{Z})$ permutes the even spin structures: $\mathfrak{s}$ interchanges (P, A) and (A, P), and preserves (A, A), whereas $\mathfrak{t}$ interchanges (A, A) and (P, A) and preserves (A, P). Note that both $\mathfrak{s}$ and $\mathfrak{t}^2$ preserve (A, A), so that the index 3 subgroup $\Gamma_\theta := \langle \mathfrak{s}, \mathfrak{t}^2\rangle < \mathrm{SL}(2,\mathbb{Z})$ is the spin mapping class group of the torus equipped with spin structure (A, A). The spin mapping class group of the torus with spin structure (A, P) or (P, A) is similarly generated by $\mathfrak{s}^2$ and $\mathfrak{t}$, which is projectively isomorphic to $\mathbb{Z}$. On the other hand, Γ_θ is projectively the free product of $\mathbb{Z}/2\mathbb{Z}$ with $\mathbb{Z}$ [Rademacher 1929]. Now the matrix $\hat{T}^2$ is unambiguously defined for any super-modular category $\mathcal{B}$, and in [Bruillard et al. 2017, Theorem II.7] it is shown that $\mathfrak{s} \to \hat{S}$ and $\mathfrak{t}^2 \to \hat{T}^2$ define a projective representation $\hat{\rho}$ of Γ_θ. We propose the following:

Conjecture 1.2. *Let $\mathcal{B}$ be a super-modular category of rank $2k$ and $\hat{S}$ and $\hat{T}^2$ the corresponding matrices as in* (1-1). *Then the kernel of the projective representation $\hat{\rho} : \Gamma_\theta \to \mathrm{PSU}(k)$ given by $\hat{\rho}(\mathfrak{s}) = \pi_k(\hat{S})$ and $\hat{\rho}(\mathfrak{t}^2) = \pi_k(\hat{T}^2)$ is a congruence subgroup.*

In particular if this conjecture holds then $\hat{\rho}(\Gamma_\theta)$ is finite. We do not know what to expect the level of $\ker \hat{\rho}$ to be (in terms of, say, the order of $\hat{T}^2$), but we provide some examples below.

An important outstanding conjecture [Davydov et al. 2013b, Question 5.15; Bruillard et al. 2017, Conjecture III.9; Müger 2003, Conjecture 5.2] is that every super-modular category $\mathcal{B}$ has a *minimal modular extension*, that is, $\mathcal{B}$ can be embedded in a modular category $\mathcal{C}$ of dimension $\dim(\mathcal{C}) = 2\dim(\mathcal{B})$. One may characterize such $\mathcal{C}$: they are called *spin modular categories* [Beliakova et al. 2017]; see Section 3A below. Our main result proves Conjecture 1.2 for super-modular categories admitting minimal modular extensions.

2. Preliminaries

2A. ***Super-modular categories.*** Though one may always define an S-matrix for any ribbon fusion category $\mathcal{B}$, it may be degenerate. This failure of modularity is encoded in the subcategory of transparent objects called the *Müger center* $\mathcal{B}'$. Here an object X is called *transparent* if all the double braidings with X are trivial:

$$c_{Y,X}c_{X,Y} = \mathrm{Id}_{X\otimes Y}.$$

By Proposition 1.1 of [Bruguières 2000], the simple objects in $\mathcal{B}'$ are those X with $\tilde{S}_{X,Y} = d_X d_Y$ for all simple Y, where $d_Y = \dim(Y) = \tilde{S}_{\mathbf{1},Y}$ is the categorical dimension of the object Y. The Müger center is obviously *symmetric*, that is, $c_{Y,X}c_{X,Y} = \mathrm{Id}_{X\otimes Y}$ for all $X, Y \in \mathcal{B}'$. Symmetric fusion categories have been classified by Deligne [1990], in terms of representations of supergroups. In the case that $\mathcal{B}' \cong \mathrm{Rep}(G)$ (i.e., is Tannakian), the modularization (de-equivariantization) procedure of Bruguières [2000] and Müger [2004] yields a modular category $\mathcal{B}_G$ of dimension $\dim(\mathcal{B})/|G|$. Otherwise, by taking a maximal Tannakian subcategory $\mathrm{Rep}(G) \subset \mathcal{B}'$, the de-equivariantization $\mathcal{B}_G$ has Müger center $(\mathcal{B}_G)' \cong \mathrm{sVec}$, the symmetric fusion category of super-vector spaces. Generally, a braided fusion category $\mathcal{B}$ with $\mathcal{B}' \cong \mathrm{sVec}$ as symmetric fusion categories is called *slightly degenerate* [Drinfeld et al. 2010].

The symmetric fusion category sVec has a unique spherical structure compatible with unitarity and has S- and T-matrices: $S_{\mathrm{sVec}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $T_{\mathrm{sVec}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

From this point on we will assume that all our categories are unitary, so that sVec is a unitary symmetric fusion category. A unitary slightly degenerate ribbon category will be called *super-modular*. In other terminology, we say $\mathcal{B}$ is super-modular if its Müger center is generated by a *fermion*, that is, an object ψ with $\psi^{\otimes 2} \cong \mathbf{1}$ and $\theta_\psi = -1$.

Equation (1-1) shows that the S and T matrices of any super-modular category can be expressed as (Kronecker) tensor products: $S = S_{\mathrm{sVec}} \otimes \hat{S}$ and $T = T_{\mathrm{sVec}} \otimes \hat{T}$ with $\hat{S}$ uniquely determined and $\hat{T}$ determined by some sign choices. The projective group generated by $\hat{S}$ and $\hat{T}$ may be infinite for all choices of $\hat{T}$ as the following example illustrates:

Example 2.1. Consider the modular category $\mathrm{SU}(2)_6$. The label set is

$$I = \{0, 1, 2, 3, 4, 5, 6\}.$$

The subcategory $\mathrm{PSU}(2)_6$ is generated by four simple objects with even labels: $X_0 = \mathbf{1}$, X_2, X_4, X_6. We have

$$\hat{S} = \frac{1}{\sqrt{4+2\sqrt{2}}} \begin{pmatrix} 1 & 1+\sqrt{2} \\ 1+\sqrt{2} & -1 \end{pmatrix} \quad \text{and} \quad \hat{T} = \begin{pmatrix} 1 & 0 \\ 0 & \pm i \end{pmatrix}.$$

For either choice of $\hat{T}$ the eigenvalues of $\hat{S}\hat{T}$ are not roots of unity: one checks that they satisfy the irreducible polynomial $x^{16} - x^{12} + \frac{1}{4}x^8 - x^4 + 1$, which has nonabelian Galois group and is not monic over $\mathbb{Z}$.

2B. *The θ-subgroup of* SL(2, ℤ). The index 3 subgroup $\Gamma_\theta < \mathrm{SL}(2,\mathbb{Z})$ generated by $\mathfrak{s}$ and $\mathfrak{t}^2$ has a uniform description (see, e.g., [Köhler 1988]):

$$\Gamma_\theta = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : ac \equiv bd \equiv 0 \pmod 2 \right\}.$$

The notation Γ_θ comes from the fact that Jacobi's θ series $\theta(z) := \sum_{n=-\infty}^{\infty} e^{n^2 \pi \mathrm{i} z}$ is a modular form of weight $\frac{1}{2}$ on Γ_θ. Moreover, Γ_θ is isomorphic to $\Gamma_0(2)$, the Hecke congruence subgroup of level 2 defined as those matrices in $\mathrm{SL}(2,\mathbb{Z})$ that are upper triangular modulo 2, and $\Gamma(2)$ is a subgroup of both $\Gamma_0(2)$ and Γ_θ. In particular, $\Gamma_0(2)$ and Γ_θ are distinct, yet isomorphic, congruence subgroups of level 2. An explicit isomorphism $\vartheta : \Gamma_\theta \to \Gamma_0(2)$ is given by $\vartheta(\mathfrak{g}) = M\mathfrak{g}M^{-1}$ where $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$. This can be verified directly, via:

$$M \begin{pmatrix} a & b \\ c & d \end{pmatrix} M^{-1} = \begin{pmatrix} a+c & \frac{d+b-a-c}{2} \\ 2c & d-c \end{pmatrix}.$$

Observe that $\vartheta(\Gamma(n)) = \Gamma(n)$ for any n, and for n even $\Gamma(n) \lhd \Gamma_\theta$. In particular, we see that $\Gamma_\theta/\Gamma(n) < \mathrm{SL}(2,\mathbb{Z})/\Gamma(n)$ is isomorphic to an index 3 subgroup of $\mathrm{SL}(2,\mathbb{Z}/n\mathbb{Z})$ that is not normal. Suppose $\varphi : \Gamma_\theta \to H$ has a kernel which is a congruence subgroup, i.e., $\Gamma(n) < \ker(\varphi)$. The congruence level of $\ker(\varphi)$, i.e., the minimal n with $\Gamma(n) < \ker(\varphi)$, is the minimal n so that $\Gamma_\theta/\Gamma(n) \twoheadrightarrow \varphi(\Gamma_\theta)$. The following provides a characterization of such quotients:

Lemma 2.2. *Suppose that $n = 2^k q$ with $k \geq 1$ and q odd. Denote by P_k a 2-Sylow subgroup of* $\mathrm{SL}(2, \mathbb{Z}/2^k\mathbb{Z})$. *Then,*

$$\Gamma_\theta/\Gamma(n) \cong P_k \times \mathrm{SL}(2, \mathbb{Z}/q\mathbb{Z}).$$

Proof. By the Chinese remainder theorem, nonnormal index 3 subgroups of

$$\mathrm{SL}(2, \mathbb{Z}/n\mathbb{Z}) \cong \prod_{p|n} \mathrm{SL}(2, \mathbb{Z}/p^{\ell_p}\mathbb{Z})$$

correspond to nonnormal index 3 subgroups of $\mathrm{SL}(2, \mathbb{Z}/p^{\ell_p}\mathbb{Z})$ where $n = \prod_{p|n} p^{\ell_p}$ is the prime factorization of n. Any 2-Sylow subgroup of $\mathrm{SL}(2, \mathbb{Z}/2^k\mathbb{Z})$ has index 3 and is not normal (since reduction modulo 2 gives a surjection to $\mathrm{SL}(2, \mathbb{Z}/2\mathbb{Z}) \cong \mathfrak{S}_3$) so it is enough to show that this fails for $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$ with $p > 2$.

In general, if $H < G$ is a nonnormal subgroup of index 3 then the (transitive) left action of G on the coset space G/H provides a homomorphism to the symmetric group on three letters, i.e., $\phi : G \to \mathfrak{S}_3$. If $\phi(G) = \mathfrak{A}_3$ (the alternating group on three letters) then we would have $\ker(\phi) = H \lhd G$. Thus $\phi(G) = \mathfrak{S}_3$, so that any such group G must have an irreducible 2-dimensional representation with character values $2, -1, 0$.

By [Nobs 1976; Eholzer 1995], we see that for $p > 2$, the groups $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$ only have 2-dimensional irreducible representations for $p = 3, 5$, and each of these representations factor over the reduction modulo p map $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z}) \twoheadrightarrow \mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$. By inspection neither $\mathrm{SL}(2, \mathbb{Z}/3\mathbb{Z})$ nor $\mathrm{SL}(2, \mathbb{Z}/5\mathbb{Z})$ have $\mathfrak{S}_3$ as quotients. □

3. Main results

In this section we prove Conjecture 1.2 for any super-modular category that admits a minimal (spin) modular extension.

3A. ***Spin modular categories.*** A *spin modular category* $\mathcal{C}$ is a modular category with a (chosen) fermion. Let $\mathcal{C}$ be a spin modular category, with fermion ψ, (unnormalized) S-matrix $\tilde{S}$ and T-matrix T. Proposition II.3 of [Bruillard et al. 2017] provides a number of useful symmetries of $\tilde{S}$ and T:

(1) $\tilde{S}_{\psi,\alpha} = \epsilon_\alpha d_\alpha$, where $\epsilon_\alpha = \pm 1$ and $\epsilon_\psi = 1$.

(2) $\theta_{\psi\alpha} = -\epsilon_\alpha \theta_\alpha$.

(3) $\tilde{S}_{\psi\alpha,\beta} = \epsilon_\beta \tilde{S}_{\alpha,\beta}$.

We have a canonical $\mathbb{Z}/2\mathbb{Z}$-grading $\mathcal{C}_0 \oplus \mathcal{C}_1$ with simple objects $X \in \mathcal{C}_0$ if $\epsilon_X = 1$

and $X \in \mathcal{C}_1$ when $\epsilon_X = -1$. The trivial component $\mathcal{C}_0$ is a super-modular category, since $\mathcal{C}_0' = \langle \psi \rangle \cong \mathrm{sVec}$.

Since $\theta_X = -\epsilon_X \theta_{\psi X}$ it is clear that $\psi X \ncong X$ for $X \in \mathcal{C}_0$. However, objects in $\mathcal{C}_1$ may be fixed by $- \otimes \psi$ or not. This provides another canonical decomposition $\mathcal{C}_1 = \mathcal{C}_v \oplus \mathcal{C}_\sigma$ as abelian categories, where a simple object $X \in \mathcal{C}_v \subset \mathcal{C}_1$ if $X\psi \ncong X$ and $X \in \mathcal{C}_\sigma \subset \mathcal{C}_1$ if $X\psi \cong X$. Finally, using the action of $- \otimes \psi$ we make a (noncanonical) decomposition of $\mathcal{C}_0 = \breve{\mathcal{C}}_0 \oplus \psi\breve{\mathcal{C}}_0$ and $\mathcal{C}_v = \breve{\mathcal{C}}_v \oplus \psi\breve{\mathcal{C}}_v$ so that when $X \in \breve{\mathcal{C}}_0$ we have $X\psi \in \psi\breve{\mathcal{C}}_0$ and similarly for $\mathcal{C}_v$. Notice that for $X \in \mathcal{C}_0$ we have $X^* \ncong \psi \otimes X$ since $\theta_X = \theta_{X^*}$, so that we may ensure X and X^* are both in $\breve{\mathcal{C}}_0$ or both in $\psi\breve{\mathcal{C}}_0$. On the other hand, for $Y \in \mathcal{C}_v$ it is possible that $X^* \cong \psi \otimes X$ — for example, this occurs for $\mathrm{SO}(2)_1$.

We choose an ordered basis

$$\Pi = \Pi_0 \sqcup \psi\Pi_0 \sqcup \Pi_v \sqcup \psi\Pi_v \sqcup \Pi_\sigma$$

for the Grothendieck ring of $\mathcal{C}$ that is compatible with the above partition $\mathcal{C} = \breve{\mathcal{C}}_0 \oplus \psi\breve{\mathcal{C}}_0 \oplus \breve{\mathcal{C}}_v \oplus \psi\breve{\mathcal{C}}_v \oplus \mathcal{C}_\sigma$. Using [Bruillard et al. 2017, Proposition II.3] we have the block matrix decomposition for the S and T matrices:

$$S = \begin{pmatrix} \frac{1}{2}\hat{S} & \frac{1}{2}\hat{S} & A & A & X \\ \frac{1}{2}\hat{S} & \frac{1}{2}\hat{S} & -A & -A & -X \\ A^T & -A^T & B & -B & 0 \\ A^T & -A^T & -B & B & 0 \\ X^T & -X^T & 0 & 0 & 0 \end{pmatrix} \quad T = \begin{pmatrix} \hat{T} & 0 & 0 & 0 & 0 \\ 0 & -\hat{T} & 0 & 0 & 0 \\ 0 & 0 & \hat{T}_v & 0 & 0 \\ 0 & 0 & 0 & \hat{T}_v & 0 \\ 0 & 0 & 0 & 0 & T_\sigma \end{pmatrix}.$$

Here B and $\hat{S}$ are symmetric matrices, and each of $\hat{T}$, $\hat{T}_v$ and T_σ are diagonal matrices.

Now consider the following ordered partitioned basis:

(1) $\Pi_0^+ := \{X_i + \psi X_i : X_i \in \Pi_0\}$,

(2) $\Pi_0^- := \{X_i - \psi X_i : X_i \in \Pi_0\}$,

(3) $\Pi_v^+ := \{Y_i + \psi Y_i : Y_i \in \Pi_v\}$,

(4) $\Pi_\sigma := \{Z_i \in \Pi_\sigma\}$ and

(5) $\Pi_v^- := \{Y_i - \psi Y_i : Y_i \in \Pi_v\}$.

With respect to this partitioned basis, the S and T matrices have the block form:

$$S' = \begin{pmatrix} \hat{S} & 0 & 0 & 0 & 0 \\ 0 & 0 & 2A & X & 0 \\ 0 & 2A^T & 0 & 0 & 0 \\ 0 & 2X^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2B \end{pmatrix} \quad T' = \begin{pmatrix} 0 & \hat{T} & 0 & 0 & 0 \\ \hat{T} & 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{T}_v & 0 & 0 \\ 0 & 0 & 0 & T_\sigma & 0 \\ 0 & 0 & 0 & 0 & \hat{T}_v \end{pmatrix}.$$

From this choice of basis one sees that the representation ρ restricted to $\Gamma_\theta = \langle \mathfrak{s}, \mathfrak{t}^2 \rangle$ has three invariant (projective) subspaces, spanned by Π_0^+, $\Pi_0^- \cup \Pi_v^+ \cup \Pi_\sigma$ and Π_v^-, respectively. In particular, we have a surjection $\rho(\Gamma_\theta) \twoheadrightarrow \hat{\rho}(\Gamma_\theta)$, mapping the image of S in $\mathrm{PSU}(|\Pi|)$ to the image of $\hat{S}$ in $\mathrm{PSU}(|\Pi_0^+|)$. We can now prove:

Theorem 3.1. *Suppose that $\mathcal{B}$ is a super-modular category with minimal modular extension $\mathcal{C}$ so that $\mathcal{B} = \mathcal{C}_0$. Assume further that the T-matrix of $\mathcal{C}$ has order N. Then $\hat{\rho} : \Gamma_\theta \to \mathrm{PSU}(k)$ has* $\ker(\hat{\rho})$ *which is a congruence subgroup of level at most N.*

Proof. Let S and T be the S-matrix and T-matrix of $\mathcal{C}$. Consider the projective representation ρ of $\mathrm{SL}(2, \mathbb{Z})$ defined by $\rho(\mathfrak{s}) = S$ and $\rho(\mathfrak{t}) = T$. By Theorem 1.1, $\ker(\rho)$ is a congruence subgroup of level N, i.e., $\Gamma(N) < \ker(\rho)$. Now the restriction of $\rho_{|\Gamma_\theta}$ to Γ_θ has

$$\ker(\rho_{|\Gamma_\theta}) = \ker(\rho) \cap \Gamma_\theta \supset \Gamma(N) \cap \Gamma_\theta.$$

However, since $\mathcal{C}$ contains a fermion N is even, so $\Gamma(N) < \Gamma(2) < \Gamma_\theta$ hence $\Gamma(N) \cap \Gamma_\theta = \Gamma(N)$. It follows that $\Gamma(N) < \ker(\rho_{|\Gamma_\theta})$. The discussion above now implies $\Gamma(N) < \ker(\rho_{|\Gamma_\theta}) < \ker(\hat{\rho})$ as we have a surjection $\rho(\Gamma_\theta) \twoheadrightarrow \hat{\rho}(\Gamma_\theta)$. Thus, we have shown that $\ker(\hat{\rho})$ is a congruence subgroup of level at most N, and in particular, $\hat{\rho}$ has finite image. $\square$

3B. *Further questions.* The charge conjugation matrix C in the basis above has the form $C'_{i,j} = \pm\delta_{i,j^*}$. Since we have arranged that $X_i \in \Pi_0$ implies $X_i^* \in \Pi_0$, $C'_{i,j} = -1$ can only occur for $i = j \in \Pi_v^-$: if $(W - \psi W)^* = -(W - \psi W)$ for some simple object W, then $W^* = \psi W$. We see that this can only happen if $W \in \mathcal{C}_v$ by comparing twists. Under this change of basis, we have

$$(S')^2 = \dim(\mathcal{C}) C' \quad \text{and} \quad (S'T')^3 = \frac{D_+}{D}(S')^2.$$

It would be interesting to explore the extra relations among the various submatrices of S' and T'.

The 16 spin modular categories of dimension 4 are of the form $\mathrm{SO}(n)_1$ (where $\mathrm{SO}(n)_1 \cong \mathrm{SO}(m)_1$ if and only if $n \cong m \pmod{16}$). For n odd $\mathrm{SO}(n)_1$ has rank 3 whereas for n even $\mathrm{SO}(n)_1$ has rank 4. For example, the Ising modular category corresponds to $n = 1$ and $\mathrm{SO}(2)_1$ has fusion rules like the group $\mathbb{Z}_4$. For any modular category $\mathcal{D}$ and $1 \leq n \leq 16$ the spin modular category $\mathrm{SO}(n)_1 \boxtimes \mathcal{D}$ with fermion $(\psi, \mathbf{1})$ has either $\mathcal{C}_\sigma = \varnothing$ or $\mathcal{C}_v = \varnothing$. An interesting problem is to classify spin modular categories with either $\mathcal{C}_\sigma = \varnothing$ or $\mathcal{C}_v = \varnothing$, particularly those with no $\boxtimes$-factorization.

4. A case study

Our result gives an upper bound on the level of $\ker(\hat{\rho})$ for super-modular categories $\mathcal{B}$ with minimal modular extensions $\mathcal{C}$: the level of $\ker(\hat{\rho})$ is at most the order of the T-matrix of $\mathcal{C}$. The actual level can be lower: for a trivial example we consider the super-modular category sVec. In this case $\hat{S} = \hat{T}^2 = I$ so the level $\ker(\hat{\rho})$ is 1, yet the order of the T matrix for its (sixteen) minimal modular extensions can be $2, 4, 8$ or 16. More generally for any split super-modular category $\mathcal{B} = \mathcal{D} \boxtimes \mathrm{sVec} \subset \mathcal{D} \boxtimes \mathrm{SO}(n)_1 = \mathcal{C}$ (with fermion $(\mathbf{1}, \psi)$) the ratio of the levels of the kernels of the $\mathrm{SL}(2, \mathbb{Z})$ (for $\mathcal{C}$) and Γ_θ (for $\mathcal{B}$, i.e., $\mathcal{D}$) representations can be 2^k for $0 \le k \le 4$.

To gain further insight we consider a family of nonsplit super-modular categories obtained from the spin modular category (see [Bruillard et al. 2017, Lemma III.7]) $\mathrm{SU}(2)_{4m+2}$. This has modular data:

$$\tilde{S}_{i,j} := \frac{\sin \dfrac{(i+1)(j+1)\pi}{4m+4}}{\sin \dfrac{\pi}{4m+4}}, \qquad T_{j,j} := e^{\pi \,\mathrm{i}(j^2+2j)/(8m+8)},$$

where $0 \le i, j \le 4m+2$. Since T has order $16(m+1)$, Theorem 1.1 implies that the image of the projective representation $\rho : \mathrm{SL}(2, \mathbb{Z}) \to \mathrm{PSU}(4m+3)$ defined via the normalized S-matrix S and T factors over $\mathrm{SL}(2, \mathbb{Z}/N\mathbb{Z})$ where $N = 16(m+1)$.

The super-modular subcategory $\mathrm{PSU}(2)_{4m+2}$ has simple objects labeled by even i, j. The factorization (1-1) yields

$$\hat{S}_{i,j} = \frac{\sin \dfrac{(2i+1)(2j+1)\pi}{4m+4}}{\Xi \sin \dfrac{\pi}{4m+4}}, \qquad \hat{T}_{j,j} = e^{\pi \,\mathrm{i}(j^2+j)/(2m+2)} \tag{4-1}$$

for $0 \le i, j \le m$, where

$$\Xi = \sqrt{\frac{m+1}{2}} \Big/ \sin \frac{\pi}{4m+4}.$$

In [Bruillard et al. 2017] all 16 minimal modular extensions of $\mathrm{PSU}(2)_{4m+2}$ are explicitly constructed and each has T-matrix of order $16(m+1)$ so that the kernel of the corresponding projective $\mathrm{SL}(2, \mathbb{Z})$ representation is a congruence subgroup of level $16(m+1)$. Our computations suggest the following conjecture, with cases verified using Magma indicated in parentheses. A sample of the results of these computations are found in Table 1. The notation $\langle n, k\rangle$ indicates the kth group of order n in the GAP library of small groups [Besche et al. 2002]. In the last column, we sometimes give a slightly different description than is indicated in part (f) below. We include the groups $\hat{\rho}(\Gamma_\theta)$, $A_m' := [A_m, A_m]$ and $\overline{A}_m := A_m/Z(A_m)$. As $\hat{\rho}$ is not necessarily irreducible, we have $\hat{\rho}(\Gamma_\theta) \twoheadrightarrow \overline{A}_m$. The congruence level of $\ker \hat{\rho}$ is computed using Lemma 2.2.

m	$\lvert\overline{A}_m\rvert$	$\overline{A}_m$	A'_m	$\hat{\rho}(\Gamma_\theta)$
1	2^4	D_{16}	$\mathbb{Z}_8$	$D_{16} = A'_1 \rtimes \mathbb{Z}_2$
2	12	$\mathrm{PSL}(2, \mathbb{Z}_3)$	$\boldsymbol{Q}_8$	$\mathrm{SL}(2, \mathbb{Z}_3) \rtimes \mathbb{Z}_2$
3	2^7	$\langle 128, 71\rangle$	$\langle 64, 184\rangle$	$\langle 128, 71\rangle$
4	60	$\mathrm{PSL}(2, \mathbb{Z}_5)$	$\mathrm{SL}(2, \mathbb{Z}_5)$	$A'_4 \rtimes \mathbb{Z}_2$
5	$2^4 \cdot 12$	$D_{16} \times \mathrm{PSL}(2, \mathbb{Z}_3)$	$\mathbb{Z}_8 \times \boldsymbol{Q}_8$	$(\mathbb{Z}_8 \times \mathrm{SL}(2, \mathbb{Z}_3)) \rtimes \mathbb{Z}_2$
6	168	$\mathrm{PSL}(2, \mathbb{Z}_7)$	$\mathrm{SL}(2, \mathbb{Z}_7)$	$A'_6 \rtimes \mathbb{Z}_2$
7	2^{10}	$\overline{A}_7$	$\lvert\cdot\rvert = 2^9$	$\overline{A}_7$
8	324	$\mathrm{PSL}(2, \mathbb{Z}_9)$	$(\mathbb{Z}_3)^3 \rtimes \boldsymbol{Q}_8$	$(A'_8 \rtimes \mathbb{Z}_3) \rtimes \mathbb{Z}_2$
9	$2^4 \cdot 60$	$D_{16} \times \mathrm{PSL}(2, \mathbb{Z}_5)$	$\mathbb{Z}_8 \times \mathrm{SL}(2, \mathbb{Z}_5)$	$A'_9 \rtimes \mathbb{Z}_2$
10	660	$\mathrm{PSL}(2, \mathbb{Z}_{11})$	$\mathrm{SL}(2, \mathbb{Z}_{11})$	$A'_{10} \rtimes \mathbb{Z}_2$
11	$2^7 \cdot 12$	$\langle 128, 71\rangle \times \mathrm{PSL}(2, \mathbb{Z}_3)$	$\langle 64, 184\rangle \times \boldsymbol{Q}_8$	$\mathrm{SL}(2, \mathbb{Z}_3) \rtimes \langle 128, 71\rangle$
12	1092	$\mathrm{PSL}(2, \mathbb{Z}_{13})$	$\mathrm{SL}(2, \mathbb{Z}_{13})$	$\mathrm{SL}(2, \mathbb{Z}_{13}) \rtimes \mathbb{Z}_2$
13	$2^4 \cdot 168$	$D_{16} \times \mathrm{PSL}(2, \mathbb{Z}_7)$	$\mathbb{Z}_8 \times \mathrm{SL}(2, \mathbb{Z}_7)$	$A'_{13} \rtimes \mathbb{Z}_2$
14	720	$\mathrm{PSL}(2, \mathbb{Z}_{15})$	$\boldsymbol{Q}_8 \times \mathrm{SL}(2, \mathbb{Z}_5)$	$\mathrm{SL}(2, \mathbb{Z}_{15}) \rtimes \mathbb{Z}_2$

Table 1. A sample of $\mathrm{PSU}(2)_{4k+2}$ results.

Conjecture 4.1. *Let A_m be the subgroup of $\mathrm{SU}(k)$ generated by $\hat{S}$ and $\hat{T}^2$ associated with $\mathrm{PSU}(2)_{4m+2}$, the quotient $\overline{A}_m := A_m/Z(A_m)$ and the commutator subgroup $A'_m := [A_m, A_m]$. Then:*

(a) *When $m+1 = q$ is odd, $\overline{A}_m = \overline{A}_{q-1} \cong \mathrm{PSL}(2, \mathbb{Z}/q\mathbb{Z})$ (verified for $2 \le m \le 18$).*

(b) *When $m+1 = 2^n$ we have $\lvert\overline{A}_m\rvert = \lvert\overline{A}_{2^n-1}\rvert = 2^{3n+1}$ (verified for $1 \le n \le 5$).*

(c1) *If we write $m+1 = 2^n q$ where q is odd, then $\overline{A}_m \cong \overline{A}_{2^n-1} \times \overline{A}_{q-1}$ (verified for $1 \le m \le 14$).*

(c2) *If we write $m+1 = 2^n q$ where q is odd, $\lvert\overline{A}_m\rvert = 2^{3n+1} q^3 \prod_{p\mid q}(p^2-1)/2p^2$ (primes p) (verified for $1 \le m \le 21$).*

(d) *For $5 \le m+1 = p$ prime, $A'_{p-1} \cong \mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$ (verified for $4 \le m \le 12$).*

(e) *If we write $m+1 = 2^n q$ where q is odd, then $A'_m \cong A'_{2^n-1} \times A'_{q-1}$ (verified for $1 \le m \le 14$).*

(f) *For $m+1 \not\equiv 0 \pmod 4$, we have $A'_m \lhd \hat{\rho}(\Gamma_\theta)$ and $\hat{\rho}(\Gamma_\theta)$ is an iterated semidirect product of A'_m with cyclic group actions (verified for $1 \le m \le 14$). In general, $\ker(\hat{\rho})$ is a congruence subgroup of level $4(m+1)$ (verified for $1 \le m \le 12$).*

Appendix: Magma code

For our computational experiments we used the symbolic algebra software Magma [Bosma et al. 1997]. In this appendix we give some basic pseudo-code and some sample Magma code to illustrate how we found the image of $\hat{\rho}(\Gamma_\theta)$ in our case study, so that the interested reader can do similar explorations. Given an integer m, the $(m+1)\times(m+1)$ $\hat{S}$ and $\hat{T}^2$ matrices obtained from $\mathrm{PSU}(2)_{4m+2}$ are given in (4-1). In order to use the Magma software we express the entries of $\hat{S}$ and $\hat{T}^2$ in the cyclotomic field $\boldsymbol{Q}(\omega)$, where ω is an $(8m{+}8)$-th root of unity. For this we must write

$$\sin\frac{(2i+1)(2j+1)\pi}{4m+4} \quad\text{and}\quad \sqrt{2(m+1)}$$

in terms of ω, for which we use the result of generalized form of quadratic Gauss sums [Berndt and Evans 1981].

Here is the pseudocode to find $\hat{\rho}(\Gamma_\theta)$ for $\mathrm{PSU}(2)_{4m+2}$:

Algorithm: **projective image**
input: an integer m
output: $\hat{\rho}(\Gamma_\theta)$ for $\mathrm{PSU}(2)_{4m+2}$

set K to the cyclotomic field $\boldsymbol{Q}(\omega)$, where ω is an $(8m{+}8)$-th root of unity.
set $M = 2(m+1)$.
initialize S and $T2$ to be $(m+1)\times(m+1)$ zero matrices over K.

Step 1: *calculate auxiliary factor* α.
if $M \equiv 0 \pmod 4$
  set $\alpha = \sum_{n=0}^{M-1} \omega^{4n^2}/(1+\omega^M)$
else
  set $\alpha = \big((\omega^{m+1} - \omega^{-(m+1)}) \sum_{n=0}^{m} \omega^{8n^2}\big)/\omega^{2m+2}$
  if $m+1 \equiv 3 \pmod 4$
    set $\alpha = \alpha/\omega^M$
set $\alpha = 2/\alpha$

Step 2: *define the entries of S and $T2$.*
for $1 \le i, j \le m+1$
  set $S_{i,j} = \alpha\big(\omega^{(2i-1)(2j-1)} - \omega^{-(2i-1)(2j-1)}\big)/(2\omega^M)$
for $1 \le j \le m+1$
  set $T2_{j,j} = \omega^{(2(j-1))^2+4(j-1)}$

Step 3: *find the projective image.*
set A to the matrix group generated by S and $T2$
set ZK to the group of scalar matrices over K
return $A/(ZK \cap A)$, the projective image of A.

The following code can be used in Magma [Bosma et al. 1997] to find the $\hat{\rho}(\Gamma_\theta)$ in this case, and slight modifications will give the other headings of Table 1:

```
m:=1;
K<w>:=CyclotomicField(8*m+8);
GL:=GeneralLinearGroup(m+1,K);
M:=2*(m+1);
alpha:=0;
if M mod 4 eq 0 then
    for n:=0 to M-1 do
        alpha:=alpha + w^(4*(n^2));
    end for;
    alpha:=alpha/(w^M+1);
else
    for n:=0 to m do
        alpha:= alpha + w^(8*(n^2));
    end for;
    if (m+1) mod 4 eq 3 then
        alpha:=alpha/(w^M);
    end if;
    alpha:=((w^(m + 1) - w^(-(m + 1)))/(w^(2*m + 2)))*alpha;
end if;
alpha:=2/alpha;
S:=ZeroMatrix(K,m+1,m+1);
for i:=1 to m+1 do
    for j:=1 to m+1 do
        S[i,j]:=(w^((2*i-1)*(2*j-1))-w^(-(2*i-1)*(2*j-1)))/(2*(w^M));
        S[i,j]:=S[i,j]*alpha;
    end for;
end for;
T2:=ZeroMatrix(K,m+1,m+1);
for j:=1 to m+1 do
    T2[j,j]:=w^((2*(j-1))^2+4*(j-1));
end for;
A:=MatrixGroup<m+1,K|S,T2>;
ZK:=MatrixGroup<m+1,K|w*IdentityMatrix(K,m+1)>;
F:=(A/(A meet ZK));
```

References

[Alvarez-Gaumé et al. 1986] L. Alvarez-Gaumé, G. Moore, and C. Vafa, "Theta functions, modular invariance, and strings", *Comm. Math. Phys.* **106**:1 (1986), 1–40. MR Zbl

[Bakalov and Kirillov 2001] B. Bakalov and A. Kirillov, Jr., *Lectures on tensor categories and modular functors*, University Lecture Series **21**, American Mathematical Society, Providence, RI, 2001. MR Zbl

[Beliakova et al. 2017] A. Beliakova, C. Blanchet, and E. Contreras, "Spin modular categories", *Quantum Topol.* **8**:3 (2017), 459–504. MR Zbl

[Berndt and Evans 1981] B. C. Berndt and R. J. Evans, "The determination of Gauss sums", *Bull. Amer. Math. Soc.* (*N.S.*) **5**:2 (1981), 107–129. MR Zbl

[Besche et al. 2002] H. U. Besche, B. Eick, and E. O'Brien, "GAP: The SmallGroups Library", software package, 2002, available at https://www.gap-system.org/Packages/sgl.html.

[Blanchet 2005] C. Blanchet, "A spin decomposition of the Verlinde formulas for type A modular categories", *Comm. Math. Phys.* **257**:1 (2005), 1–28. MR Zbl

[Blanchet and Masbaum 1996] C. Blanchet and G. Masbaum, "Topological quantum field theories for surfaces with spin structure", *Duke Math. J.* **82**:2 (1996), 229–267. MR Zbl

[Bonderson 2007] P. H. Bonderson, *Non-abelian anyons and interferometry*, Ph.D. thesis, California Institute of Technology, 2007, available at https://tinyurl.com/bondanyon.

[Bonderson et al. 2013] P. Bonderson, C. Nayak, and X.-L. Qi, "A time-reversal invariant topological phase at the surface of a 3D topological insulator", *J. Stat. Mech. Theory Exp.* **2013**:9 (2013), art. id. P09016. MR

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Bruguières 2000] A. Bruguières, "Catégories prémodulaires, modularisations et invariants des variétés de dimension 3", *Math. Ann.* **316**:2 (2000), 215–236. MR Zbl

[Bruillard et al. 2016a] P. Bruillard, S.-H. Ng, E. C. Rowell, and Z. Wang, "Rank-finiteness for modular categories", *J. Amer. Math. Soc.* **29**:3 (2016), 857–881. MR Zbl

[Bruillard et al. 2016b] P. Bruillard, S.-H. Ng, E. C. Rowell, and Z. Wang, "On classification of modular categories by rank", *Int. Math. Res. Not.* **2016**:24 (2016), 7546–7588. MR Zbl

[Bruillard et al. 2017] P. Bruillard, C. Galindo, T. Hagge, S.-H. Ng, J. Y. Plavnik, E. C. Rowell, and Z. Wang, "Fermionic modular categories and the 16-fold way", *J. Math. Phys.* **58**:4 (2017), art. id. 041704. MR Zbl

[Davydov et al. 2013a] A. Davydov, M. Müger, D. Nikshych, and V. Ostrik, "The Witt group of non-degenerate braided fusion categories", *J. Reine Angew. Math.* **677** (2013), 135–177. MR Zbl

[Davydov et al. 2013b] A. Davydov, D. Nikshych, and V. Ostrik, "On the structure of the Witt group of braided fusion categories", *Selecta Math.* (*N.S.*) **19**:1 (2013), 237–269. MR Zbl

[Deligne 1990] P. Deligne, "Catégories tannakiennes", pp. 111–195 in *The Grothendieck Festschrift, II*, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, Boston, 1990. MR Zbl

[Drinfeld et al. 2010] V. Drinfeld, S. Gelaki, D. Nikshych, and V. Ostrik, "On braided fusion categories, I", *Selecta Math.* (*N.S.*) **16**:1 (2010), 1–119. MR Zbl

[Eholzer 1995] W. Eholzer, "On the classification of modular fusion algebras", *Comm. Math. Phys.* **172**:3 (1995), 623–659. MR Zbl

[Köhler 1988] G. Köhler, "Theta series on the theta group", *Abh. Math. Sem. Univ. Hamburg* **58** (1988), 15–45. MR Zbl

[Lan et al. 2016] T. Lan, L. Kong, and X.-G. Wen, "Theory of $(2+1)$-dimensional fermionic topological orders and fermionic/bosonic topological orders with symmetries", *Phys. Rev. B* **94**:15 (2016), art. id. 155113.

[Müger 2003] M. Müger, "On the structure of modular categories", *Proc. London Math. Soc.* (3) **87**:2 (2003), 291–308. MR Zbl

[Müger 2004] M. Müger, "Galois extensions of braided tensor categories and braided crossed G-categories", *J. Algebra* **277**:1 (2004), 256–281. MR Zbl

[Ng and Schauenburg 2010] S.-H. Ng and P. Schauenburg, "Congruence subgroups and generalized Frobenius–Schur indicators", *Comm. Math. Phys.* **300**:1 (2010), 1–46. MR Zbl

[Nobs 1976] A. Nobs, "Die irreduziblen Darstellungen der Gruppen $\mathrm{SL}_2(\mathbb{Z}_p)$, insbesondere $\mathrm{SL}_2(\mathbb{Z}_2)$, I", *Comment. Math. Helv.* **51**:4 (1976), 465–489. MR Zbl

[Rademacher 1929] H. Rademacher, "Über die Erzeugenden von Kongruenzuntergruppen der Modulgruppe", *Abh. Math. Sem. Univ. Hamburg* **7**:1 (1929), 134–148. MR Zbl

[Sawin 2002] S. F. Sawin, "Invariants of spin three-manifolds from Chern–Simons theory and finite-dimensional Hopf algebras", *Adv. Math.* **165**:1 (2002), 35–70. MR Zbl

[Walker and Wang 2012] K. Walker and Z. Wang, "$(3+1)$-TQFTs and topological insulators", *Front. Phys.* **7**:2 (2012), 150–159.

Received April 15, 2017.

PARSA BONDERSON
MICROSOFT RESEARCH-STATION Q
UNIVERSITY OF CALIFORNIA,
SANTA BARBARA, CA
UNITED STATES
parsab@microsoft.com

ERIC C. ROWELL
DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TX
UNITED STATES
rowell@math.tamu.edu

ZHENGHAN WANG
MICROSOFT RESEARCH-STATION Q
UNIVERSITY OF CALIFORNIA,
SANTA BARBARA, CA
UNITED STATES
zhenghwa@microsoft.com

QING ZHANG
DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TX
UNITED STATES
zhangqing@math.tamu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.271

ON THE CHOW RING OF THE STACK OF TRUNCATED BARSOTTI–TATE GROUPS

DENNIS BROKEMPER

We determine the Chow ring of the stack of truncated displays and more generally the Chow ring of the stack of G-zips. We also investigate the pullback morphism of the truncated display functor. From this we can determine the Chow ring of the stack of truncated Barsotti–Tate groups over a field of characteristic p up to p-torsion.

Introduction

Edidin and Graham [1998] developed an equivariant intersection theory for actions of linear algebraic groups G on algebraic spaces X. For such G-spaces they defined G-equivariant Chow groups $A_*^G(X)$ generalizing Totaro's definition [1999] of the G-equivariant Chow ring of a point. They are an invariant of the corresponding quotient stack $[X/G]$, i.e., they are independent of the choice of presentation. Hence they can be used to define the integral Chow group of a quotient stack. If X is smooth these groups carry a ring structure making them into commutative graded rings. Edidin and Graham used their theory to compute the Chow ring of the stacks $\mathscr{M}_{1,1}$ and $\bar{\mathscr{M}}_{1,1}$ of elliptic curves. In an appendix to that paper, Vistoli computed the Chow ring of $\mathscr{M}_2$. Edidin and Fulghesu [2009] computed the integral Chow ring of the stack of hyperelliptic curves of even genus. In this article we investigate the Chow ring of the stack of truncated Barsotti–Tate groups over a field of characteristic $p > 0$.

Let us denote the stack of level-n Barsotti-Tate groups by BT_n. A level-n BT group has a height and a dimension, which are locally constant functions on the base. If $\mathrm{BT}_n^{h,d}$ denotes the stack of level-n BT groups of constant height h and

Part of the author's doctoral work at the University of Paderborn, Germany.
MSC2010: 14C15.
Keywords: equivariant Chow groups, Barsotti–Tate groups, displays, G-zips.

dimension d we obtain a decomposition $\mathrm{BT}_n = \coprod_{0\le d\le h} \mathrm{BT}_n^{h,d}$ into open and closed substacks. For example, if A is an abelian scheme of relative dimension g then its p^n-torsion subscheme $A[p^n]$ is a level-n BT group of height $2g$ and dimension g.

Although $\mathrm{BT}_n^{h,d}$ has a natural presentation $[X/\operatorname{GL}_{p^{nh}}]$ as a quotient stack with quasiaffine and smooth X (see [Wedhorn 2001]), it seems unlikely that this presentation can be used directly to compute the Chow ring. Instead we compare the stack of truncated Barsotti–Tate groups to a stack whose Chow ring is easier to compute, but still closely related to the Chow ring of BT_n.

Our choice for this stack is the stack $\mathcal{D}\mathrm{isp}_n$ of truncated displays introduced in [Lau 2013]. Displays were first introduced in [Zink 2002] to provide a Dieudonné theory that is valid not only over perfect fields but more generally over $\mathbb{F}_p$-algebras or p-adic rings. While displays are given by an invertible matrix with entries in the ring of Witt vectors $W(R)$, if a basis of the underlying modules is fixed, a truncated display is given by an invertible matrix over the truncated Witt ring $W_n(R)$.

Using crystalline Dieudonné theory one can associate to every p-divisible group a display. This induces a morphism $\phi : \mathrm{BT} \to \mathcal{D}\mathrm{isp}$ from the stack of Barsotti–Tate groups to the stack of displays, which in turn induces a morphism,

$$\phi_n : \mathrm{BT}_n \to \mathcal{D}\mathrm{isp}_n,$$

compatible with the truncations on both sides. By [Lau 2013], this morphism is a smooth morphism of smooth algebraic stacks over k and an equivalence on geometric points.

Theorem A. *The pullback $\phi_n^* : A^*(\mathcal{D}isp_n) \to A^*(\mathrm{BT}_n)$ is injective and an isomorphism after inverting p.*

Let us sketch the proof. Consider a field L and a morphism $\operatorname{Spec} L \to \mathrm{BT}_n$. After base change to a finite field extension of p-power degree the fiber $\phi_n^{-1}(\operatorname{Spec} L)$ is equal to the classifying space of an infinitesimal group scheme necessarily of p-power degree. It follows that the pullback map of Bloch's higher Chow groups $A_*(\operatorname{Spec} L, m) \to A_*(\phi_n^{-1}(\operatorname{Spec} L), m)$ becomes an isomorphism after inverting p. Using the long localization exact sequence the theorem follows from a limit argument and noetherian induction similar to that in [Quillen 1973, Proposition 4.1]. The injectivity assertion follows since $A^*(\mathcal{D}\mathrm{isp}_n)$ is p-torsion free.

Thus to compute the Chow ring of BT_n at least up to p-torsion it suffices to compute the Chow ring of $\mathcal{D}\mathrm{isp}_n$, which is much easier due to the simpler presentation as a quotient stack. More precisely, if $\mathcal{D}\mathrm{isp}_n^{h,d}$ denotes the open and closed substack in $\mathcal{D}\mathrm{isp}_n$ of truncated displays with constant dimension d and height h we have

$$\mathcal{D}\mathrm{isp}_n^{h,d} = [\operatorname{GL}_h(W_n(\,\cdot\,))/G_n^{h,d}],$$

where W_n refers to the ring of truncated Witt vectors and $G_n^{h,d}$ is an extension of

$\mathrm{GL}_d \times \mathrm{GL}_{h-d}$ by a unipotent group. The following result reduces the calculation of $A^*(\mathcal{D}\mathrm{isp}_n)$ to the case $n = 1$.

Theorem B. *The pullback* $\tau_n^* : A^*(\mathcal{D}isp_1) \to A^*(\mathcal{D}isp_n)$ *of the truncation map* $\tau_n : \mathcal{D}isp_n \to \mathcal{D}isp_1$ *is an isomorphism.*

This is proved using the factorization

$$[\mathrm{GL}_h(W_n(\cdot))/G_n^{h,d}] \to [GL_h/G_n^{h,d}] \to [GL_h/G_1^{h,d}]$$

of τ_n and the fact that the first map is an affine bundle and that $G_n^{h,d}$ is an extension of $G_1^{h,d}$ by a unipotent group.

In a similar way one shows that the Chow ring of $\mathcal{D}\mathrm{isp}_1^{h,d}$ coincides with that of the quotient stack

$$[\mathrm{GL}_h\,/(\mathrm{GL}_d \times \mathrm{GL}_{h-d})],$$

where the action is given by conjugation with the Frobenius. This situation is a special case of Proposition 2.3.2.

Theorem C. *The following equation holds*:

$$\begin{aligned} A^*(\mathcal{D}isp_1^{h,d}) &= A^*_{\mathrm{GL}_d \times \mathrm{GL}_{h-d}}(\mathrm{GL}_h) \\ &= \mathbb{Z}[t_1, \dots, t_h]^{S_d \times S_{h-d}}/((p-1)c_1, \dots, (p^h-1)c_h), \end{aligned}$$

where $c_1, \dots, c_h$ *are the elementary symmetric polynomials in the variables* $t_1, \dots, t_h$.

Moreover, $t_1, \dots, t_d$ and $t_{d+1}, \dots, t_h$ are the Chern roots of the vector bundle $\mathcal{L}ie$ and ${}^t\mathcal{L}ie^\vee$, respectively, over $\mathcal{D}\mathrm{isp}_1^{h,d}$. Here $\mathcal{L}ie$ is a vector bundle of rank d assigning to a display its Lie algebra, and ${}^t\mathcal{L}ie^\vee$ is of rank $h-d$ assigning to a display the dual Lie algebra of its dual display.

It follows that the $\mathbb{Q}$-vectorspace $A^*(\mathcal{D}\mathrm{isp}_1^{h,d})_{\mathbb{Q}}$ is finite-dimensional of dimension $\binom{h}{d}$, which also equals the number of isomorphism classes of truncated displays of level 1 with height h and dimension d over an algebraically closed field. We show that a basis is given by the cycles of the closures of the respective EO strata. We prove this fact in greater generality for the stack of G-zips [Pink et al. 2011] in Section 2.4. In that section we will also compute the Chow ring of the stack of G-zips for a connected algebraic zip datum. As in the case of displays the computation can be reduced to the situation of Proposition 2.3.2. In fact, truncated displays of level 1 are a special case of G-zips.

To state our result we recall that an algebraic zip datum $\mathcal{Z}$ is a 4-tuple (G, P, Q, φ) consisting of a split reductive group G, parabolic subgroups P and Q, and an isogeny $\varphi : P/R_u(P) \to Q/R_u(Q)$. To $\mathcal{Z}$ one associates the stack of G-zips $[G/E_{\mathcal{Z}}]$, where $E_{\mathcal{Z}}$ is the group $\{(p, q) \in P \times Q \mid \varphi(\pi_P(p)) = \pi_Q(q)\}$ acting on G by the rule $((p, q), g) \mapsto pgq^{-1}$. We also recall that an algebraic group is called *special* if every principal G-bundle is locally trivial for the Zariski topology.

Theorem D. *Let $\mathcal{Z} = (G, P, Q, \varphi)$ be an algebraic zip datum, where G is connected. Let $W_G = W(G, T)$ be the Weyl group of G and $W_L = W(L, T)$ be the Weyl group of a Levi component L of P with respect to a split maximal torus $T \subset L$ of G. Let $g_0 \in G(k)$ be such that $\varphi(T) = {}^{g_0}T$ and let $\tilde{\varphi} : T \to T$ denote the composition of φ followed by conjugation with g_0^{-1}. Then $\tilde{\varphi}$ induces an action on $S = \mathrm{Sym}(\widehat{T})$, which we will also denote by $\tilde{\varphi}$. We then have*

$$A^*([G/E_{\mathcal{Z}}])_{\mathbb{Q}} = S_{\mathbb{Q}}^{W_L}/(f - \tilde{\varphi} f \mid f \in S_+^{W_G})_{\mathbb{Q}}.$$

If G is special we have

$$A^*([G/E_{\mathcal{Z}}]) = S^{W_L}/(f - \tilde{\varphi} f \mid f \in S_+^{W_G}).$$

(*Note that the action of $\tilde{\varphi}$ on S^{W_G} is independent of the choice of g_0 since two choices differ by an element of $N_G(T)$.*)

Gathering the above results we gain the following information on the Chow ring of the stack of truncated Barsotti–Tate groups.

Theorem E. (i) *We have*

$$A^*(\mathrm{BT}_n^{h,d})_p = \mathbb{Z}[p^{-1}][t_1, \dots, t_h]^{S_d \times S_{h-d}}/((p-1)c_1, \dots, (p^h - 1)c_h),$$

where c_i denotes the i-th elementary symmetric polynomial in the variables $t_1, \dots, t_h$, and $t_1, \dots, t_d$ and $t_{d+1}, \dots, t_h$ are the Chern roots of $\mathcal{L}ie$ and ${}^t\mathcal{L}ie^\vee$, respectively.

(ii) *We have* $\dim_{\mathbb{Q}} A^*(\mathrm{BT}_n^{h,d})_{\mathbb{Q}} = \binom{h}{d}$ *and a basis is given by the cycles of the closures of the EO strata.*

(iii) $$(\mathrm{Pic}\,\mathrm{BT}_n^{h,d})_p = \begin{cases} \mathbb{Z}[p^{-1}]/(p-1) & \text{if } d = 0, h, \\ \mathbb{Z}[p^{-1}] \times \mathbb{Z}[p^{-1}]/(p-1) & \text{otherwise}, \end{cases}$$

where the generators for the free and torsion parts are, respectively, $\det(\mathcal{L}ie)$ *and* $\det(\mathcal{L}ie \otimes {}^t\mathcal{L}ie^\vee)$.

It would be interesting to know if the Chow ring of BT_n has p-torsion, and more specifically if the Picard group of BT_n has p-torsion. However, since ϕ_n^* is injective and the Chow ring of $\mathcal{D}\mathrm{isp}_n$ is p-torsion free, p-torsion in the Chow ring of BT_n cannot be constructed using displays.

Terminology and notation. Every scheme is assumed to be of finite type and separated over the base field k. In Section 2, we assume k to be of characteristic $p > 0$. Algebraic groups are affine smooth group schemes over k. We call an algebraic group G unipotent if G admits a filtration $G = G_0 \supset G_1 \supset \cdots \supset G_e = \{1\}$ by subgroups such that G_i is normal in G_{i-1} with quotient isomorphic to $\mathbb{G}_a$. The character group of an algebraic group G will be denoted by $\widehat{G}$. If X is a scheme, $A^*(X)$ will always denote the operational Chow ring of X [Fulton 1998, Chapter 17]. $A_*(X)$

and $\mathrm{CH}^*(X)$ will be the Chow group of X graded by dimension and codimension, respectively. If X is an algebraic space over k with a left action of an algebraic group G we will refer to X as a G-space. We write $[X/G]$ for the corresponding quotient stack. If G acts freely on X, i.e., the stabilizer of every point is trivial, then $[X/G]$ is an algebraic space. In this case we will write X/G instead of $[X/G]$ and call $X \to X/G$ the principal bundle quotient of X with structure group G.

1. Equivariant intersection theory

1.1. ***Equivariant Chow groups.*** Consider an algebraic group G over k. By [Edidin and Graham 1998, Lemma 9], we can find a representation V of G, and an open subset U in V such that the complement of U has arbitrary high codimension, and such that the principal bundle quotient U/G exists in the category of schemes. If X is an algebraic space on which G acts then G acts diagonally on $X \times U$ and we will denote the principal bundle quotient $(X \times U)/G$ by X_G.

Convention 1.1.1. We call a pair (V, U) consisting of a G-representation V and an open subset U a good pair for G if G acts freely on U, i.e., the stabilizer of every point is trivial. Sometimes we will call the quotient $X_G = (X \times U)/G$ a mixed space for the G-space X. If (V, U) is a good pair for G with $\mathrm{codim}(U^c, V) > i$ we will also call $(X \times U)/G$ an approximation of $[X/G]$ up to codimension i.

If X has dimension n the i-th equivariant Chow group $A_i^G(X)$ is defined in the following way. Choose a good pair (V, U) for G such that the complement of U has codimension greater than $n - i$. Then one defines

$$A_i^G(X) = A_{i+l-g}(X_G),$$

where l denotes the dimension of V and g is the dimension of G. The definition is independent of the choice of the pair (V, U) as long as $\mathrm{codim}(U^c, V) > n - i$ holds [Edidin and Graham 1998, Definition-Proposition 1].

The equivariant Chow groups have the same functorial properties as ordinary Chow groups [Edidin and Graham 1998, Section 2]. In particular, we have an operational equivariant Chow ring $A_G^*(X)$ [Edidin and Graham 1998, Section 2.6], i.e., an element $c \in A_G^i(X)$ consists of operations $c(Y \to X) : A_*^G(Y) \to A_{*-i}^G(Y)$ for each G-equivariant map $Y \to X$ that are compatible with flat pullback, proper pushforward and Gysin homomorphisms.

We will denote by $\mathrm{CH}_G^*(X)$ the G-equivariant Chow group of X graded by codimension. If X is a pure dimensional G-scheme and (V, U) a good pair for G with $\mathrm{codim}(U^c, V) > i$ then

$$\mathrm{CH}_G^j(X) = \mathrm{CH}^j((X \times U)/G)$$

for all $j \le i$. This motivates the term "approximation of $[X/G]$ up to codimension i" in Convention 1.1.1.

If X is smooth then $\mathrm{CH}^*_G(X)$ carries a ring structure which makes it into a commutative graded ring with unit element. Moreover, there is a natural isomorphism $A^*_G(X) \cong \mathrm{CH}^*_G(X)$ of graded rings [Edidin and Graham 1998, Proposition 4].

By [Edidin and Graham 1998, Proposition 16], the equivariant Chow groups do not depend on the presentation as a quotient, meaning if X is a G-space and Y is an H-space such that $[X/G] \cong [Y/H]$, then $A^G_{i+g}(X) = A^H_{i+h}(Y)$, where $g = \dim G$ and $h = \dim H$. Hence one can define the Chow group of a quotient stack $[X/G]$ to be

$$A_i([X/G]) = A^G_{i+g}(X)$$

with $g = \dim G$. By [Edidin and Graham 1998, Proposition 19], one has

$$A^*([X/G]) \cong A_*([X/G]),$$

whenever X is smooth.

1.2. ***Higher equivariant Chow groups.*** The reason we shall need higher Chow groups is that they extend the localization exact sequence to the left. Higher Chow groups were introduced by Bloch [1986]. For a scheme X, higher Chow groups $A_i(X, m)$ are defined as the homology of the complex $z_i(X, *)$, where $z_i(X, m)$ is the group of cycles of dimension $m + i$ in $X \times \Delta^m$ meeting all faces properly. For $m = 0$ one gets back the usual Chow group $A_*(X)$, and $A_i(X, m)$ may be nontrivial for $-m \le i \le \dim X$. The definition of these higher Chow groups also works for algebraic spaces.

In order to define G-equivariant versions $A^G_*(X, m)$ of higher Chow groups we need the homotopy property for the mixed spaces X_G, i.e., the pullback map

$$A_*(X_G, m) \to A_*(\mathcal{E}, m)$$

for a vector bundle $\mathcal{E}$ over X_G is an isomorphism. This is true for any scheme if $\mathcal{E}$ is trivial by [Bloch 1986, Theorem 2.1]. To prove the assertion for arbitrary vector bundles one needs the localization exact sequence of higher Chow groups proved by Bloch in the case of quasiprojective schemes: if X is an equidimensional, quasiprojective scheme over k and $Y \subset X$ is a closed subscheme with complement $U = X - Y$, then there is a long exact sequence of higher Chow groups

$$\cdots \to A_*(Y, m) \to A_*(X, m) \to A_*(U, m) \to A_*(Y, m-1) \to \\ \cdots \to A_*(Y) \to A_*(X) \to A_*(U) \to 0.$$

For a proof see [Edidin and Graham 1998, Lemma 4] and [Bloch 1986, Theorem 3.1].

Remark 1.2.1. Levine extended Bloch's proof of the existence of the long localization exact sequence to all separated schemes of finite type over k [Levine 2001, Theorem 1.7]. Hence for the equivariant higher Chow groups to be well defined it suffices that we can choose the mixed spaces to be separated schemes over k. However, in all applications we have in mind the conditions of Lemma 1.2.2 will be satisfied.

Lemma 1.2.2. *Let G be an algebraic group and X a normal, quasiprojective G-scheme. Then for any $i > 0$ there is a representation V of G and an invariant open subset $U \subset V$ whose complement has codimension greater than i such that G acts freely on U and the principal bundle quotient $(X \times U)/G$ is a quasiprojective scheme. In other words, the quotient stack $[X/G]$ can be approximated by quasiprojective schemes.*

Proof. Embed G into GL_n for some n. Then there is a representation V of GL_n and an open subset $U \subset V$, whose complement has codimension greater than i such that U/GL_n is a Grassmannian (see [Edidin and Graham 1998, Lemma 9]). Since GL_n is special the GL_n/G-bundle $\pi : U/G \to U/\mathrm{GL}_n$ is locally trivial for the Zariski topology, and we will first show that π is quasiprojective.

Since GL_n/G is quasiprojective and normal there is an ample GL_n-linearizable line bundle $L \to \mathrm{GL}_n/G$ [Thomason 1988, Section 5.7]. Then

$$(U \times L)/\mathrm{GL}_n \to (U \times (\mathrm{GL}_n/G))/\mathrm{GL}_n = U/G$$

is a line bundle relatively ample for π. This shows that π is quasiprojective. The same holds then for U/G. Again by [Thomason 1988, Section 5.7], there is an ample G-linearizable line bundle on X. The pullback to $X \times U$ is then relatively ample for the projection $X \times U \to U$. Applying [Mumford et al. 1994, Proposition 7.1] to this situation yields the claim. □

Definition 1.2.3. (i) A pair (V, U) will be called an admissible pair for a G-scheme X if (V, U) is a good pair for G and if the mixed space X_G is quasiprojective and (locally) equidimensional over k. X will be called an admissible G-scheme if for any i there is an admissible pair (V, U) for X with $\mathrm{codim}(U^c, V) > i$.

(ii) If X is an admissible G-scheme we define its higher equivariant Chow groups to be

$$A_i^G(X, m) = A_{i+l-g}(X_G, m),$$

where $g = \dim G$ and X_G is formed from an l-dimensional admissible pair (V, U) such that $\mathrm{codim}(U^c, V) > \dim X + m - i$.

(iii) We will say that a stack $\mathscr{X}$ admits an admissible presentation if there exists an admissible G-scheme X such that $\mathscr{X} = [X/G]$.

(iv) Let $\mathscr{X}$ be a quotient stack that admits a presentation $\mathscr{X} = [X/G]$ by an admissible G-scheme X. We define the higher equivariant Chow groups of $\mathscr{X}$ as

$$A_*(\mathscr{X}, m) = A_{*+g}^G(X, m),$$

where $g = \dim G$.

Remark 1.2.4. The proof that (ii) and (iv) of Definition 1.2.3 are independent of the choice of the admissible pair (V, U) and the presentation $[X/G]$, respectively, is the same as for ordinary equivariant Chow groups (see Definition-Proposition 1 and Proposition 16, respectively, in [Edidin and Graham 1998]) by using the homotopy property for the mixed spaces.

Remark 1.2.5. We will frequently encounter the situation of a morphism $T \to X$ of G-schemes such that T is open in a G-equivariant vector bundle over X. We remark that, if X is an admissible G-scheme, so is T. This follows since a vector bundle over a quasiprojective scheme is again quasiprojective.

Lemma 1.2.6. *Let $f : \mathscr{X} \to \mathscr{Y}$ be a flat map of quotient stacks of relative dimension r. Then there is a flat pullback map $f^* : A_*(\mathscr{Y}) \to A_{*+r}(\mathscr{X})$ between the Chow groups. If $\mathscr{X}$ and $\mathscr{Y}$ admit admissible presentations the same assertion holds for the higher Chow groups.*

Furthermore, if $\mathscr{X}$ and $\mathscr{Y}$ are smooth then under the identification $A_*(\mathscr{X}) = A^*(\mathscr{X})$, the above morphism is just the natural pullback map between the operational Chow rings.

Proof. Consider presentations $\mathscr{X} = [X/G]$ and $\mathscr{Y} = [Y/H]$. By definition $A_i(\mathscr{X}) = A^G_{i+g}(X)$ with $g = \dim G$ and similarly for $A_i(\mathscr{Y})$. Choose a good pair (V_1, U_1) for G and a good pair (V_2, U_2) for H. Let $l_i = \dim V_i$. As usual we will write X_G and Y_H for the mixed spaces $(X \times U_1)/G$ and $(Y \times U_2)/H$, respectively. Consider the following fiber square:

$$\begin{tikzcd} Z' \arrow[r] \arrow[d] & \mathscr{Z} \arrow[r] \arrow[d] & Y_H \arrow[d] \\ X_G \arrow[r] & \mathscr{X} \arrow[r] & \mathscr{Y} \end{tikzcd}$$

Then Z' is a bundle over X_G and $\mathscr{Z}$ with fibers U_2 and U_1, respectively, and $Z' \to Y_H$ is a flat map of algebraic spaces of relative dimension $l_1 + r$. Hence

$$A_{i+l_1+l_2+r}(Z') = A_{i+l_1+r}(X_G) = A_{i+r}(\mathscr{X})$$

and we define f^* to be the ordinary pullback of the flat map $Z' \to Y_H$. The exact same construction works for the higher equivariant Chow groups if $\mathscr{X}$ and $\mathscr{Y}$ admit admissible presentations.

For the last part we recall that the isomorphism $A^i(\mathscr{X}) \cong A^G_{\dim X - i}(X)$ maps $c \in A^i(\mathscr{X})$ to $c(X_G \to \mathscr{X}) \cap [X_G] \in A^G_{\dim X-i}(X)$. Thus we need to check the equality

$$f^*(d(Y_H \to \mathscr{Y}) \cap [Y_H]) = d(X_G \to \mathscr{X} \to \mathscr{Y}) \cap [X_G]$$

for $d \in A^i(\mathscr{Y})$. This follows from the compatibility of d with flat pullbacks. $\square$

1.3. *Auxiliary results.*

Lemma 1.3.1. *Let $X \to Y$ be a flat morphism of schemes and $Y' \to Y$ be a finite, flat and surjective map of degree d. Let $X' \to Y'$ be the base change of $X \to Y$ along $Y' \to Y$. Assume the pullback $A_*(Y', m) \to A_*(X', m)$ becomes an isomorphism after inverting some integer d'. Then the pullback $A_*(Y, m) \to A_*(X, m)$ is an isomorphism after inverting dd'.*

Proof. The injectivity of the pullback $A_*(Y, m)_{dd'} \to A_*(X, m)_{dd'}$ follows from the exact diagram:

$$\begin{tikzcd}
0 \arrow[r] & A_*(Y, m)_{dd'} \arrow[r] \arrow[d] & A_*(Y', m)_{dd'} \arrow[d, "\cong"'] \\
0 \arrow[r] & A_*(X, m)_{dd'} \arrow[r] & A_*(X', m)_{dd'}
\end{tikzcd}$$

and the surjectivity from the exact diagram

$$\begin{tikzcd}
A_*(Y', m)_{dd'} \arrow[r] \arrow[d, "\cong"'] & A_*(Y, m)_{dd'} \arrow[r] \arrow[d] & 0 \\
A_*(X', m)_{dd'} \arrow[r] & A_*(X, m)_{dd'} \arrow[r] & 0
\end{tikzcd}$$

where the horizontal maps in the first diagram are induced by pullback and in the second diagram by pushforward. The commutativity of the second diagram is shown by [Fulton 1998, Proposition 1.7]. □

Lemma 1.3.2. *Let $T \to X$ be a morphism of quasiprojective schemes over k. We assume that X is equidimensional and that $T \to X$ is flat of relative dimension a. Let $d, i \in \mathbb{Z}$ and for $x \in X$ let $h(x)$ denote the dimension of the closure of $\{x\}$ in X. If the pullback $A_{i-h(x)}(\operatorname{Spec} k(x), m)_d \to A_{i-h(x)+a}(T_x, m)_d$ is an isomorphism for every $x \in X$ and for any m, then $A_i(X, m)_d \to A_{i+a}(T, m)_d$ is an isomorphism.*

Proof. We follow Quillen's proof of the analogous result in higher K-theory [1973, Proposition 4.1]. First we may assume that X is irreducible for if $X = W_1 \cup \cdots \cup W_r$ is a decomposition into irreducible components we may consider the long localization exact sequence of the pair $(W_1, X - W_1)$. By induction we are thus reduced to the irreducible case. Since the Chow groups only depend on the reduced structure, we may also assume that X is reduced. Let K denote the function field of X. We have

$$A_{i-n}(\operatorname{Spec} K, m) = \varinjlim_U A_i(U, m),$$
$$A_{i-n+a}(T_K, m) = \varinjlim_U A_{i+a}(T_U, m),$$

where the limit goes over all nonempty open subsets of X and n denotes the dimension of X. In fact, it suffices to go over all nonempty open subsets with equidimensional complement, since for all nonempty open U in X there exists a nonempty open subset U' contained in U with equidimensional complement. We obtain a commutative diagram

$$\begin{tikzcd}
A_{i-n}(\operatorname{Spec} K, m+1) \arrow[r] \arrow[d] & A_{i-n+a}(T_K, m+1) \arrow[d] \\
\varinjlim_Y A_i(Y, m) \arrow[r] \arrow[d] & \varinjlim_Y A_{i+a}(T_Y, m) \arrow[d] \\
A_i(X, m) \arrow[r] \arrow[d] & A_{i+a}(T, m) \arrow[d] \\
A_{i-n}(\operatorname{Spec} K, m) \arrow[r] \arrow[d] & A_{i-n+a}(T_K, m) \arrow[d] \\
\varinjlim_Y A_i(Y, m-1) \arrow[r] & \varinjlim_Y A_{i+a}(T_Y, m-1)
\end{tikzcd}$$

with exact rows, where the limit goes over all proper closed equidimensional subsets of X. After inverting d the first and fourth vertical maps become isomorphisms and we conclude by noetherian induction. □

Corollary 1.3.3. *Let $T \to X$ be a flat morphism of quasiprojective schemes over k with fibers being affine spaces of some dimension n. Then the pullback $A_*(X, m) \to A_{*+n}(T, m)$ is an isomorphism.*

Proof. This is an immediate consequence of Lemma 1.3.2. □

Remark 1.3.4. The assertion of the above corollary in the case $m = 0$ also holds without the quasiprojective assumption. One can use the same proof but using Gillet's higher Chow groups. For his higher Chow groups a long localization exact sequence exists for arbitrary schemes. For details see Chapter 8 in [Gillet 1981].

Lemma 1.3.5. *Let K be a unipotent subgroup of an algebraic group G such that the quotient G/K is finite of degree d. Then the pullback $A^*_G(m) \to A^*_{\{0\}}(m)$ is an isomorphism after inverting d.*

Proof. Let (V, U) be an admissible pair for G. Then $U/K \to U/G$ is a G/K-bundle locally trivial for the flat topology. By assumption on G/K the morphism $U/K \to U/G$ is therefore finite, flat and surjective of degree d. It follows that the pullback $A_*(U/G, m) \to A_*(U/K, m) \cong A_*(U, m)$ is injective after inverting d. Also for sufficiently high degree we know that $A_*(\operatorname{Spec} k, m) \to A_*(U, m)$ is surjective. Since we can assume the codimension of U^c in V to be arbitrarily high, we obtain the surjectivity of $A^*_G(m) \to A^*_{\{0\}}(m)$. □

Lemma 1.3.6. *Let K/k be a Galois extension with Galois group G and let X be a scheme over k. Then pulling back along $X_K \to X$ induces an isomorphism $A_*(X, m)_{\mathbb{Q}} \cong A_*(X_K, m)^G_{\mathbb{Q}}$. If K/k is a finite Galois extension of degree d it suffices to invert d.*

Proof. We first assume that K/k is finite of degree d. Then on the level of cycles we have an injection $z_*(X, \cdot)_d \hookrightarrow z_*(X_K, \cdot)^G_d$ since $X_K \to X$ is finite and flat of degree d. We claim that this map is also surjective. Let $W \subset X_K \times_K \Delta^r_K$ be a subvariety meeting all faces properly. Let $S \subset G$ be the isotropy group of W. It suffices to see that $\sum_{g\in G/S}[gW]$ lies in $z_*(X, \cdot)_d$. For this consider the closed subscheme $V = \cup_{g\in G/S}\, gW$ (equipped with the reduced structure). Then V is a G-invariant equidimensional subscheme of $X_K \times_K \Delta^r_K$ that meets all faces properly. Thus it has a model $\widetilde{V}$ over k also meeting all faces properly. Finally all components gW have the same multiplicity 1 in the cycle $[V]$ and therefore $\sum_{g\in G/S}[gW] = [\widetilde{V}_K]$. To complete the proof in the finite case it suffices now to note that taking G-invariants is an exact functor on the category of $\mathbb{Z}\big[\frac{1}{d}\big]$-modules with G-action, hence $H_i(z_*(X_K, \cdot)^G_d) = H_i(z_*(X_K, \cdot))^G_d$. The general case follows from the finite case and the fact that $A_*(X_K, m)^G = \varinjlim_{L/k} A_*(X_L, m)^{G(L/k)}$, where the limit goes over all finite Galois subextensions L/k of K. □

1.4. ***A pullback lemma.*** Throughout we consider the situation of an exact sequence

$$0 \longrightarrow A \longrightarrow G \longrightarrow H \longrightarrow 0$$

of algebraic groups and an admissible H-scheme X such that the induced G-action on X makes X also into an admissible G-scheme. These conditions are always satisfied if X is quasiprojective and normal by Lemma 1.2.2. We are then interested in properties of the pullback homomorphism (Lemma 1.2.6),

$$A_*([X/H], m) \to A_*([X/G], m).$$

Proposition 1.4.1. *Let*

$$0 \longrightarrow A \longrightarrow G \longrightarrow H \longrightarrow 0$$

be an exact sequence of algebraic groups and X an admissible H-scheme such that the induced G-action makes X also into an admissible G-scheme. We also assume H to be special.

Let $d \in \mathbb{Z}$ be such that $A^*_{A_L}(m) \to A^*_{\{0\}}(m)$ becomes an isomorphism after inverting d for every field extension L of k and every m. Then the pullback $A_*([X/H], m) \to A_*([X/G], m)$ becomes an isomorphism after inverting d.

Proof. First note that the natural map $[X/G] \to [X/H]$ is flat of relative dimension $-a$ with $a = \dim A$. We can choose for any $i \in \mathbb{Z}$ an admissible pair (V, U) for the H-action such that $A_{j+l}([(X \times U)/G], m) = A_j([X/G], m)$ and

$A_{j+l}((X\times U)/H, m) = A_j([X/H], m)$ for all $j > i$. Here l denotes the dimension of V. Note that $X \times U$ is again an admissible G-scheme (see Remark 1.2.5). Replacing X by $X \times U$ we may thus assume that $[X/H]$ is a quasiprojective scheme.

Now, let $(X \times U)/G$ be a quasiprojective mixed space for G. Let $\overline{U}$ be the quotient U/A. Then we can identify $(X \times U)/G$ with the quotient $(X \times \overline{U})/H$ and under this identification the map $(X \times U)/G \to X/H$ corresponds to the $\overline{U}$-bundle $(X \times \overline{U})/H \to X/H$. It is Zariski locally trivial since H is special. We are left to show that the pullback of this map is an isomorphism after inverting d. This will follow from Lemma 1.3.2 once we have seen that the pullback $A_{j-h(x)}(\operatorname{Spec} k(x), m)_d \to A_{j-h(x)+l-a}(\overline{U}_{k(x)}, m)_d$ is an isomorphism for every $x \in X/H$. Here $h(x)$ is the dimension of the closure of $\{x\}$ in X/H. Let us write $L = k(x)$. Assuming the codimension of U^c in V to be sufficiently large we obtain by assumption

$$A_{j-h(x)}(\operatorname{Spec} L, m)_d = A_{j-h(x)+l}(U_L, m)_d = A_{j-h(x)+l-a}(\overline{U}_L, m)_d.$$

For this recall $A_{j+l-a}(\overline{U}_L, m) = A_j^{A_L}(m)$ and $A_{j+l}(U_L, m) = A_j^{\{0\}}(m)$. This proves the claim. $\square$

The above proposition applies to the following cases.

Corollary 1.4.2. *In the situation of Proposition 1.4.1 the following assertions hold.*

(i) *If A is unipotent then $A_*([X/H], m) \to A_*([X/G], m)$ is an isomorphism.*

(ii) *If A is finite of degree d then $A_*([X/H], m) \to A_*([X/G], m)$ becomes an isomorphism after inverting d.*

Proof. The first part follows from Corollary 1.3.3 and the second part follows from Lemma 1.3.5 applied to the case $K = \{0\}$. $\square$

The assumption on H to be special is crucial for the proof of Proposition 1.4.1, since we need to know that the fibers of the $\overline{U}$-bundle $(X \times \overline{U})/H \to X/H$ appearing in the proof are given by $\overline{U}$ in order to apply Lemma 1.3.2. However, we have the following version when H is finite.

Proposition 1.4.3. *Let*

$$0 \longrightarrow A \longrightarrow G \longrightarrow H \longrightarrow 0$$

be an exact sequence of algebraic groups and X an admissible H-scheme such that the induced G-action makes X also into an admissible G-scheme. We assume that H is finite of degree d.

Let $d' \in \mathbb{Z}$ be such that $A^*_{A_L}(m) \to A^*_{\{0\}}(m)$ becomes an isomorphism after inverting d' for every field extension L of k and any m. Then the pullback $A_*([X/H], m) \to A_*([X/G], m)$ becomes an isomorphism after inverting dd'.

Proof. We argue the same way as in Proposition 1.4.1 and then have to see that the pullback of $(X \times \overline{U})/H \to X/H$ becomes an isomorphism after inverting dd'. As mentioned earlier we cannot apply Lemma 1.3.2 since the above $\overline{U}$-bundle is not locally trivial for the Zariski topology. Instead it becomes trivial after the finite, flat and surjective base change $X \to X/H$ of degree d, i.e., there is a cartesian diagram

$$\begin{tikzcd} X \times \overline{U} \arrow[r] \arrow[d] & X \arrow[d] \\ (X \times \overline{U})/H \arrow[r] & X/H \end{tikzcd}$$

The claim thus follows from Lemma 1.3.1. □

Corollary 1.4.4. *In the situation of Proposition 1.4.3 the following assertions hold.*

(i) *If A is unipotent then $A_*([X/H], m)_d \to A_*([X/G], m)_d$ is an isomorphism.*

(ii) *If A is finite of degree d' then $A_*([X/H], m)_{dd'} \to A_*([X/G], m)_{dd'}$ is an isomorphism.*

In the next proposition we show that the assertion of Proposition 1.4.1 is valid over $\mathbb{Q}$ for arbitrary H.

Proposition 1.4.5. *Let*

$$0 \longrightarrow A \longrightarrow G \longrightarrow H \longrightarrow 0$$

be an exact sequence of algebraic groups and X an admissible H-scheme such that the induced G-action makes X also into an admissible G-scheme.

Assume $A^*_{A_L}(m)_{\mathbb{Q}} \to A^*_{\{0\}}(m)_{\mathbb{Q}}$ is an isomorphism for every field extension L of k and any m. Then the pullback $A_*([X/H], m)_{\mathbb{Q}} \to A_*([X/G], m)_{\mathbb{Q}}$ is an isomorphism.

Proof. Using the notation of the proof of Proposition 1.4.1 we need to see that the pullback of the $\overline{U}$-bundle $T := (X \times \overline{U})/H \to X/H$ is an isomorphism over $\mathbb{Q}$. It suffices to see that $A_*(\operatorname{Spec} k(x), m)_{\mathbb{Q}} \to A_*(T_x, m)_{\mathbb{Q}}$ is an isomorphism for $x \in X/H$. The above $\overline{U}$-bundle may not be trivial for the Zariski topology, but we still have $T_{\bar{x}} = \overline{U}_{\bar{x}}$ and thus $A_*(\operatorname{Spec} k(x)^{\mathrm{sep}}, m)_{\mathbb{Q}} \to A_*(T_{\bar{x}}, m)_{\mathbb{Q}}$ is an isomorphism by assumption. The claim then follows from Lemma 1.3.6 and the fact that the Galois action is compatible with pullback. □

Corollary 1.4.6. *In the situation of Proposition 1.4.5 the following assertions hold.*

(i) *If A is unipotent then $A_*([X/H], m)_{\mathbb{Q}} \to A_*([X/G], m)_{\mathbb{Q}}$ is an isomorphism.*

(ii) *If A is finite then $A_*([X/H], m)_{\mathbb{Q}} \to A_*([X/G], m)_{\mathbb{Q}}$ is an isomorphism.*

Lemma 1.4.7. *Let G be a split extension*

$$0 \longrightarrow K \longrightarrow G \longrightarrow H \longrightarrow 0$$

of an algebraic group H by a unipotent group K. Choose a splitting $H \hookrightarrow G$ and let X be a normal, quasiprojective G-scheme. Then the pullback map

$$A_*^G(X, m)_{\mathbb{Q}} \to A_*^H(X, m)_{\mathbb{Q}}$$

is an isomorphism. If G is special, this above map is an isomorphism over $\mathbb{Z}$.

Proof. Let (V, U) be an admissible pair for the G-action on X. It follows from the proof of Lemma 1.2.2 that (V, U) is then also admissible for the induced H-action. The morphism $(X \times U)/H \to (X \times U)/G$ is a G/H-bundle. If G is special this bundle is locally trivial for the Zariski topology. Hence the lemma follows from Corollary 1.3.3 in the special case and Lemmas 1.3.2 and 1.3.6 in the general case. $\square$

1.5. ***The restriction map.*** We want to describe properties of the restriction map $\mathrm{res}_T^G : A_*^G(X) \to A_*^T(X)$, where T is a split torus in G. This map is defined via flat pullback of the natural map $X_T \to X_G$ between the mixed spaces. Note that more generally one has a restriction map $\mathrm{res}_H^G : A_*^G(X) \to A_*^H(X)$ for every subgroup H of G. We will need the following result.

Theorem 1.5.1. *Let G be a connected reductive group with split maximal torus T and Weyl group $W = W(G, T)$. Let X be a G-scheme.*

(i) *W acts on $A_*^T(X)$. Furthermore, the restriction morphism $A_*^G(X) \to A_*^T(X)$ induces a map $r : A_*^G(X) \to A_*^T(X)^W$.*

(ii) *Assume X is smooth. Then r is an isomorphism after tensoring with $\mathbb{Q}$.*

(iii) *Assume X is smooth and that G is special. Then r is injective. Moreover, r is an isomorphism if $A_T^*(X)$ is $\mathbb{Z}$-torsion free (e.g., if $X = \mathrm{Spec}\, k$).*

Part (iii) is basically proved in [Edidin and Graham 1997], where the case $X = \mathrm{Spec}\, k$ is considered. However, there seems to be no complete proof of part (ii) in the literature. We therefore give a proof.

In the following $A^*(X; \mathbb{Q})$ will denote the operational Chow ring of X consisting of characteristic classes with values in rational Chow groups, i.e., an element $c \in A^*(X; \mathbb{Q})$ assigns to each $T \to X$ a morphism

$$c(T \to X) : A_*(T)_{\mathbb{Q}} \to A_*(T)_{\mathbb{Q}}$$

satisfying the usual compatibility conditions [Fulton 1998, Section 17.1]. A proper map $\pi : \widetilde{X} \to X$ is called an envelope if for each irreducible subspace $V \subset X$ there exists an irreducible subspace $\widetilde{V} \subset \widetilde{X}$ such that π maps $\widetilde{V}$ birationally onto V.

Remark 1.5.2. There is a natural map $A^*(X)_{\mathbb{Q}} \to A^*(X;\mathbb{Q})$ and this map is an isomorphism if X is smooth. This follows from

$$\begin{array}{ccc} A^*(X)_{\mathbb{Q}} & \xrightarrow[\cong]{\cap[X]} & A_*(X)_{\mathbb{Q}} \\ \downarrow & & \| \\ A^*(X;\mathbb{Q}) & \xrightarrow[\cong]{\cap[X]} & A_*(X)_{\mathbb{Q}} \end{array}$$

We recall the following easy lemma.

Lemma 1.5.3. (i) *Let $\pi : \widetilde{X} \to X$ be a proper surjective map. Then*

$$\pi_* : A_*(\widetilde{X})_{\mathbb{Q}} \to A_*(X)_{\mathbb{Q}}$$

is surjective and $\pi^* : A^*(X;\mathbb{Q}) \to A^*(\widetilde{X};\mathbb{Q})$ is injective.

(ii) *Let $\pi : \widetilde{X} \to X$ be a birational envelope. Then $\pi_* : A_*(\widetilde{X}) \to A_*(X)$ is surjective and $\pi^* : A^*(X) \to A^*(\widetilde{X})$ is injective.*

Proof. The first part of (i) is [Kimura 1992, Proposition 1.3]. The first part of (ii) follows immediately from the definition of an envelope. The second part of (i) and (ii) are formal consequences of their first parts. □

In order to prove Theorem 1.5.1 we consider the following situation: Let G be a connected reductive group with split maximal torus T and Weyl group $W = W(G,T)$. Let M be smooth and $E \to M$ be a principal G-bundle. Consider a Borel subgroup $B \supset T$. Now W acts on $A^*(E/B)$ in the following way. We identify $W = N_G(T)/T$ and choose $w \in N_G(T)$. Then w induces an automorphism $w : E/T \to E/T$ and hence an automorphism $w^* : A^*(E/T) \to A^*(E/T)$. This defines an action of W on $A^*(E/T) = A^*(E/B)$. The following lemma is also mentioned (without proof) in [Vistoli 1989, Section 2.5].

Lemma 1.5.4. *Pullback induces an isomorphism $A^*(M)_{\mathbb{Q}} \cong A^*(E/B)^W_{\mathbb{Q}}$.*

Proof. Let $w \in N_G(T)$. Since w lies in G the diagram

$$\begin{tikzcd} E/T \arrow[r] \arrow[d, "w"'] & E/G = M \\ E/T \arrow[ur] & \end{tikzcd}$$

commutes and this implies that the image of the pullback $A^*(M) \to A^*(E/B)$ lies in $A^*(E/B)^W$. We are left to show that

$$A^*(M)_{\mathbb{Q}} \to A^*(E/B)^W_{\mathbb{Q}}$$

is an isomorphism. Let us first show that $A_*(M)_{\mathbb{Q}} \to A_*(E/B)^W_{\mathbb{Q}}$ is surjective. For this the smoothness assumption on M is not needed. We recall that every G-torsor is locally isotrivial by [Raynaud 1970, XIV, Lemma 1.4]. This means that there exists a covering of M by open subsets U with the property that for each U there is a finite, étale and surjective map $U' \to U$ such that $E_{U'} = E \times_M U' \to U'$ becomes a trivial G-torsor. Let V denote the complement of such a U in M and consider the commutative diagram

$$\begin{tikzcd}
A_*(V)_Q \arrow[r] \arrow[d] & A_*(M)_{\mathbb{Q}} \arrow[r] \arrow[d] & A_*(U)_{\mathbb{Q}} \arrow[r] \arrow[d] & 0 \\
A_*(E_V/B)^W_{\mathbb{Q}} \arrow[r] & A_*(E/B)^W_{\mathbb{Q}} \arrow[r] & A_*(E_U/B)^W_{\mathbb{Q}} \arrow[r] & 0
\end{tikzcd}$$

with exact rows. An easy diagram chase shows that if the first and last vertical map are surjective so is $A^*(M)_{\mathbb{Q}} \to A^*(E/B)^W_{\mathbb{Q}}$. Using noetherian induction we are thus reduced to the case that there exists a proper surjective map $M' \to M$ such that $E_{M'} \to M'$ is trivial. Since the diagram

$$\begin{tikzcd}
A_*(M')_{\mathbb{Q}} \arrow[r] \arrow[d] & A_*(E_{M'}/B)^W_{\mathbb{Q}} \arrow[d] \\
A_*(M)_{\mathbb{Q}} \arrow[r] & A_*(E/B)^W_{\mathbb{Q}}
\end{tikzcd}$$

commutes [Fulton 1998, Proposition 1.7], and since $A_*(E_{M'}/B)^W_{\mathbb{Q}} \to A_*(E/B)^W_{\mathbb{Q}}$ is surjective by part (i) of the previous lemma we are further reduced to the case of a trivial G-torsor $E = G \times M \to M$. Now G/B has a decomposition into affine cells and therefore we obtain in the case of a trivial G-torsor $A_*(E/B)_{\mathbb{Q}} = A_*(G/B)_{\mathbb{Q}} \otimes A_*(M)_{\mathbb{Q}}$ by [Totaro 2014, Section 3]. From [Demazure 1973, Section 8] we get $A_*(G/B)_{\mathbb{Q}} = S_{\mathbb{Q}}/(S^W_+)$, where $S = \mathrm{Sym}(\widehat{T})$ and S^W_+ denotes the submodule generated by homogeneous W-invariant elements of positive degree. Since $(S_{\mathbb{Q}}/(S^W_+))^W = \mathbb{Q}$ we obtain $A_*(E/B)^W_{\mathbb{Q}} = A_*(M)$ as wanted.

By the previous lemma we know that $A^*(M;\mathbb{Q}) \to A^*(E/B;\mathbb{Q})$ is injective but since M (and therefore E) is smooth we obtain the injectivity of $A^*(M)_{\mathbb{Q}} \to A^*(E/B)_{\mathbb{Q}}$. □

Proof of Theorem 1.5.1. The assertions (i) and (ii) are immediate consequences of Lemma 1.5.4. Under the assumption that $A^*_T(X)$ is $\mathbb{Z}$-torsion free the surjectivity of r follows from part (ii) by using the argumentation of the proof of Lemma 5 in [Edidin and Graham 1997]. □

2. The Chow ring of the stack of level-n Barsotti–Tate groups

2.1. ***The stack of truncated displays.*** Let R be an $\mathbb{F}_p$-algebra. We denote by $W_n(R)$ the ring of truncated Witt vectors of length n. Let $I_{n,R} \subset W_n(R)$ be the image of the Verschiebung $W_{n-1}(R) \to W_n(R)$ and $J_{n,R} \subset W_n(R)$ be the kernel of the projection $W_n(R) \to W_{n-1}(R)$. The Frobenius on R induces a ring homomorphism $\sigma : W_n(R) \to W_n(R)$ and the inverse of the Verschiebung induces a bijective σ-linear map $\sigma_1 : I_{n+1,R} \to W_n(R)$. Note that $pR = 0$ implies $I_{n,R} J_{n,R} = 0$, hence we may view $I_{n+1,R}$ as a $W_n(R)$-module. We call a σ-linear map $f : M \to N$ between $W_n(R)$-modules a σ-linear isomorphism, if its linearization $f^\sharp : W_n(R) \otimes_{\sigma, W_n(R)} M \to M$ is an isomorphism of $W_n(R)$-modules.

Truncated displays were introduced in [Lau 2013]. Let us recall the necessary notation. We are only going to need the following description of truncated displays.

Definition 2.1.1. A truncated display of level n over an $\mathbb{F}_p$-algebra R is a triple (L, T, Ψ) consisting of projective $W_n(R)$-modules L and T of finite rank and a σ-linear automorphism $\Psi : L \oplus T \to L \oplus T$.

A morphism between truncated displays is defined as follows. First we can use Ψ to define σ-linear maps

$$F : L \oplus T \to L \oplus T, \quad l + t \mapsto p\Psi(l) + \Psi(t),$$

$$F_1 : L \oplus (T \otimes_{W_n(R)} I_{n+1,R}) \to L \oplus T, \quad l + (t \otimes \omega) \mapsto \Psi(l) + \sigma_1(\omega)\Psi(t).$$

Then a morphism between two truncated displays (L, T, Ψ) and (L', T', Ψ') of level n is given by a matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where $A \in \mathrm{Hom}(L, L')$, $B \in \mathrm{Hom}(T, L')$, $C \in \mathrm{Hom}(L, T' \otimes_{W_n(R)} I_{n+1,R})$ and $D \in \mathrm{Hom}(T, T')$ such that

$$\begin{tikzcd} L \oplus T \arrow[r, "F"] \arrow[d] & L \oplus T \arrow[d] \\ L' \oplus T' \arrow[r, "F'"] & L' \oplus T' \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} L \oplus (T \otimes_{W_n(R)} I_{n+1,R}) \arrow[r, "F_1"] \arrow[d] & L \oplus T \arrow[d] \\ L' \oplus (T' \otimes_{W_n(R)} I_{n+1,R}) \arrow[r, "F_1'"] & L' \oplus T' \end{tikzcd}$$

commute.

The height of a truncated display is defined as the rank of $L \oplus T$ and the dimension as the rank of T. Both are locally constant functions on $\mathrm{Spec}\, R$. Let $\mathcal{D}\mathrm{isp}_n \to \mathrm{Spec}\, \mathbb{F}_p$ denote the stack of truncated displays of level n. That is, for R an $\mathbb{F}_p$-algebra, $\mathcal{D}\mathrm{isp}_n(\mathrm{Spec}\, R)$ is the groupoid of truncated displays of level n. It is proved in [Lau 2013, Proposition 3.15] that $\mathcal{D}\mathrm{isp}_n$ is a smooth Artin algebraic stack of dimension zero over $\mathbb{F}_p$ with affine diagonal.

For $h \in \mathbb{N}$ and $0 \le d \le h$ we denote by $\mathcal{D}\mathrm{isp}_n^{h,d}$ the open and closed substack of truncated displays of level n with constant height h and constant dimension d. Then

$$\mathcal{D}\mathrm{isp}_n = \coprod_{h,d} \mathcal{D}\mathrm{isp}_n^{h,d}.$$

A presentation of $\mathcal{D}isp_n^{h,d}$. We will adopt the notation of the proof of Proposition 3.15 in [Lau 2013]. Let $X_n^{h,d}$ be the functor on affine $\mathbb{F}_p$-schemes with

$X_n^{h,d}(R) = \mathrm{GL}_h(W_n(R))$. This is an affine open subscheme of $\mathbb{A}^{nh^2}$. Furthermore, let $G_n^{h,d}$ be the functor such that $G_n^{h,d}(R)$ is the group of invertible matrices $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with $A \in \mathrm{GL}_{h-d}(W_n(R))$, $B \in \mathrm{Hom}(W_n(R)^d, W_n(R)^{h-d})$, $C \in \mathrm{Hom}(W_n(R)^{h-d}, I_{n+1,R}^d)$ and $T \in \mathrm{GL}_d(W_n(R))$. Then $G_n^{h,d}$ is a connected algebraic group of dimension nh^2.

Remark 2.1.2. Since $I_{2,R}$ is in bijection to R via σ_1 we may view $G_1^{h,d}(R)$ as the group of invertible matrices with entries in R with respect to the multiplication given by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' & AB' + BD' \\ C\sigma(A') + \sigma(D)C' & DD' \end{pmatrix},$$

where in the four blocks we have the usual matrix multiplication.

Let $\pi_n^{h,d} : X_n^{h,d} \to \mathcal{D}\mathrm{isp}_{n,d}$ be the functor that assigns to an invertible matrix $\Psi \in \mathrm{GL}_h(W_n(R))$ the truncated display $(W_n(R)^{h-d}, W_n(R)^d, \Psi)$, where we view Ψ as a σ-linear map $W_n(R)^h \to W_n(R)^h$ via $x \mapsto \Psi \cdot \sigma x$. Now if we let $G_n^{h,d}$ act on $X_n^{h,d}$ via

$$G \cdot \Psi = G \Psi \sigma_1(G)^{-1}$$

where

$$\sigma_1(G) = \begin{pmatrix} \sigma(A) & p\sigma(B) \\ \sigma_1(C) & \sigma(D) \end{pmatrix},$$

then every $G \in G_n^{h,d}$ defines an isomorphism $\pi_n^{h,d}(\Psi) \to \pi_n^{h,d}(G \cdot \Psi)$ of truncated displays. On the other hand, if G defines an isomorphism $\pi_n^{h,d}(\Psi) \to \pi_n^{h,d}(\Psi')$ then necessarily $\Psi' = G\Psi\sigma_1(G)^{-1}$. We thus obtain the following theorem.

Theorem 2.1.3. *The functor $\pi_n^{h,d}$ induces an isomorphism of stacks*

$$[X_n^{h,d} / G_n^{h,d}] \cong \mathcal{D}isp_n^{h,d}.$$

There are the following two obvious vector bundles on $\mathcal{D}\mathrm{isp}_n^{h,d}$.

Definition 2.1.4. Let $\operatorname{Spec} R \to \mathcal{D}\mathrm{isp}_n^{h,d}$ be a map corresponding to a truncated display $\mathcal{P} = (L, T, \Psi)$.

(i) We denote by $\mathcal{L}ie$ the vector bundle of rank d over $\mathcal{D}\mathrm{isp}_n^{h,d}$ that assigns to $\operatorname{Spec} R \to \mathcal{D}\mathrm{isp}_n^{h,d}$ the vector bundle $\mathrm{Lie}(\mathcal{P}) = T/I_{n,R}T$ of rank d over R.

(ii) We denote by ${}^t\mathcal{L}ie^\vee$ the vector bundle of rank $h-d$ that assigns to $\operatorname{Spec} R \to \mathcal{D}\mathrm{isp}_n^{h,d}$ the vector bundle $L/I_{n,R}L$ of rank $h-d$ over R.

Remark 2.1.5. The notation ${}^t\mathcal{L}ie^\vee$ in the above definition stems from the fact that the dual of $L/I_{n,R}L$ gives the Lie algebra of the dual display $\mathcal{P}^t$. For the definition of the dual display see [Zink 2002, Definition 19].

The truncated display functor. As mentioned in the introduction the strategy for computing the Chow ring of the stack of truncated Barsotti–Tate groups is to relate it to the stack of truncated displays. This happens via the truncated display functor

$$\phi_n : \mathrm{BT}_n \to \mathcal{D}\mathrm{isp}_n$$

constructed in [Lau 2013]. Let us briefly sketch the construction.

Let G be a p-divisible group over an $\mathbb{F}_p$-algebra R. The ring of Witt vectors $W(R)$ is p-adically complete and the ideal I_R in $W(R)$ carries natural divided powers compatible with the canonical divided powers of p. Let $\mathbb{D}(G)$ denote the covariant Dieudonné crystal of G. We can evaluate $\mathbb{D}(G)$ at $W(R) \to R$ and set $P = \mathbb{D}(G)_{W(R)\to R}$ and $Q = \mathrm{Ker}(P \to \mathrm{Lie}(G))$. Furthermore, let $F^\sharp : P^\sigma \to P$ and $V^\sharp : P \to P^\sigma$ be the maps induced by the Frobenius and Verschiebung of G. One can show that there are σ-linear maps $F : P \to P$ and $\dot{F} : Q \to P$ compatible with base change in R such that $(P, Q, F, \dot{F})$ is a display which induces the maps $F^\sharp$ and $V^\sharp$. See [Lau 2013, Proposition 2.4] for the precise statement. This construction yields a 1-morphism

$$\phi : \mathrm{BT} \to \mathcal{D}\mathrm{isp}$$

from the stack of Barsotti–Tate groups to the stack of displays. It is clear from the construction that the Lie algebra of G is equal to the Lie algebra of $\phi(G)$ defined by P/Q.

Moreover, one can prove that for all n there are maps $\phi_n : \mathrm{BT}_n \to \mathcal{D}\mathrm{isp}_n$ compatible with the truncation maps on both sides such that ϕ is the projective limit of the system $(\phi_n)_{n\geq 1}$. The central result in [Lau 2013] is that ϕ_n is a smooth morphism of smooth algebraic stacks over $\mathbb{F}_p$ which is an equivalence on geometric points.

2.2. *Group theoretic properties of $G_n^{h,d}$.* We denote by $K_{(n,m)}^{h,d}$ the kernel of the projection $G_n^{h,d} \to G_m^{h,d}$ for $m < n$ and by $\widetilde{K}_n^{h,d}$ the kernel of the projection $G_n^{h,d} \to \mathrm{GL}_{h-d} \times \mathrm{GL}_d$. Note that $G_n^{h,0} = \mathrm{GL}_h(W_n(\,\cdot\,))$. We recall the following well known facts about the Witt ring. For an $\mathbb{F}_p$-algebra R we denote by $[\,\cdot\,] : R \to W_n(R)$ the map $r \mapsto (r, 0, \dots, 0)$ and ${}^V(\,\cdot\,) : W(R) \to W(R)$ is the Verschiebung.

Lemma 2.2.1. *Let R be an $\mathbb{F}_p$-algebra and $x, y \in R$. Then $[x+y]-[x]-[y]$ lies in ${}^VW(R)$. Furthermore, ${}^{V^r}W(R) \cdot {}^{V^s}W(R) \subset {}^{V^{r+s}}W(R)$.*

Proof. The first part follows immediately from the fact that ${}^VW(R)$ is the kernel of the ring homomorphism $\mathbb{W}_0 : W(R) \to R$ and the fact $\mathbb{W}_0([x]) = x$ for all $x \in R$.

For the second part we may assume $r \geq s$. We then write

$${}^{V^r}x\,{}^{V^s}y = {}^{V^r}(x\,{}^{F^rV^s}y) = p^s \cdot {}^{V^r}(x\,{}^{F^{r-s}}y).$$

Since $pR = 0$ we have $p(x_0, x_1, \dots) = (0, x_0^p, x_1^p, \dots)$ in $W(R)$ and the lemma follows. ☐

Lemma 2.2.2. (i) *$K^{h,d}_{(n,m)}$ is unipotent.*

(ii) *$\widetilde{K}^{h,d}_n$ is unipotent.*

Proof. (i) First note that $K^{h,0}_{(n,n-1)} = \ker(\mathrm{GL}_h(W_n(\,\cdot\,)) \to \mathrm{GL}_h(W_{n-1}(\,\cdot\,)))$ is unipotent. To see this we consider the Verschiebung ${}^V(\,\cdot\,)$ as a map $W_n(R) \to W_n(R)$. Then by the above lemma the map

$$\mathbb{G}_a^{h^2} \to K^{h,0}_{(n,n-1)}, \quad A \mapsto I_h + {}^{V^{n-1}}[A]$$

is an isomorphism of algebraic groups.

Next we show that $K^{h,d}_{(n,n-1)}$ is unipotent. This is the group of matrices $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with $A \in K^{h-d,0}_{(n,n-1)}$, $B \in J_n^{(h-d)\times d}$, $C \in J_{n+1}^{d\times(h-d)}$ and $D \in K^{d,0}_{(n,n-1)}$. The multiplication in this group is given by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' & AB' + BD' \\ CA' + DC' & DD' \end{pmatrix}.$$

Starting with the normal subgroup

$$\begin{pmatrix} I_{h-d} & J_n^{(h-d)\times d} \\ J_{n+1}^{d\times(h-d)} & I_d \end{pmatrix},$$

which is isomorphic to $\mathbb{G}_a^{2d(h-d)}$, and then using the fact that $K^{h-d,0}_{(n,n-1)}$ and $K^{d,0}_{(n,n-1)}$ are isomorphic to $\mathbb{G}_a^{(h-d)^2}$ and $\mathbb{G}_a^{d^2}$, respectively, one obtains a filtration of $K^{h,d}_{(n,n-1)}$ by normal subgroups, whose successive quotients are isomorphic to a product of copies of $\mathbb{G}_a$. Now we have an exact sequence

$$0 \longrightarrow K^{h,d}_{(n,n-1)} \longrightarrow K^{h,d}_{(n,m)} \longrightarrow K^{h,d}_{(n-1,m)} \longrightarrow 0$$

and by induction we may assume that $K^{h,d}_{(n-1,m)}$ is unipotent. It follows that $K^{h,d}_{(n,m)}$ is unipotent.

(ii) For $n = 1$ the assertion is obvious in view of Remark 2.1.2. For $n > 1$ we use the exact sequence

$$0 \longrightarrow K^{h,d}_{(n,n-1)} \longrightarrow \widetilde{K}^{h,d}_n \longrightarrow \widetilde{K}^{h,d}_{n-1} \longrightarrow 0.$$

By induction and part (i) it follows that $\widetilde{K}^{h,d}_n$ is unipotent. □

Corollary 2.2.3. (i) *$G^{h,d}_n$ is special.*

(ii) *$\widetilde{K}^{h,d}_n$ is the unipotent radical of $G^{h,d}_n$.*

(iii) *The projection $X^{h,d}_n \to X^{h,d}_1$ is a trivial $K^{h,0}_{(n,1)}$-torsor.*

Proof. We have the exact sequence

$$0 \longrightarrow \widetilde{K}^{h,d}_n \longrightarrow G^{h,d}_n \longrightarrow \mathrm{GL}_{h-d} \times \mathrm{GL}_d \longrightarrow 0.$$

Now $\widetilde{K}^{h,d}_n$ is unipotent, and thus special. Since $\mathrm{GL}_{h-d} \times \mathrm{GL}_d$ is also special part (i) follows.

Clearly the projection $X_n^{h,d} \to X_1^{h,d}$ is a $K_{(n,1)}^{h,0}$-torsor by definition of $K_{(n,1)}^{h,0}$. It is trivial since $K_{(n,1)}^{h,0}$ is unipotent and $X_1^{h,d}$ is affine. $\square$

2.3. *The Chow ring of $\mathcal{D}isp_n$.* We start with the following result which reduces the calculation of $A^*(\mathcal{D}\mathrm{isp}_n)$ to the case $n = 1$.

Theorem 2.3.1. *The pullback*

$$\tau_n^* : A^*(\mathcal{D}isp_1^{h,d}) \to A^*(\mathcal{D}isp_n^{h,d})$$

of the truncation $\tau_n : \mathcal{D}isp_n^{h,d} \to \mathcal{D}isp_1^{h,d}$ is an isomorphism.

Proof. Under the presentation $\mathcal{D}\mathrm{isp}_n^{h,d} = [X_n^{h,d}/G_n^{h,d}]$ the truncation τ_n is induced by the natural projections $X_n^{h,d} \to X_1^{h,d}$ and $G_n^{h,d} \to G_1^{h,d}$. Thus τ_n factors as

$$[X_n^{h,d}/G_n^{h,d}] \to [X_1^{h,d}/G_n^{h,d}] \to [X_1^{h,d}/G_1^{h,d}].$$

By Lemma 2.2.2 and Corollary 1.4.2, the pullback of the second map is an isomorphism. To show that the pullback of the first map is also an isomorphism let us abbreviate $X = X_1^{h,d}$ and $G = G_n^{h,d}$. By part (iii) of Corollary 2.2.3 we know that $X_n^{h,d} = X \times K$ with $K = K_{(n,1)}^{h,0}$, and the projection $X \times K \to X$ is G-equivariant. Moreover, K is an affine space by Lemma 2.2.2. After replacing $[X/G]$ by an appropriate mixed space (see Convention 1.1.1), i.e., replacing X by $X \times U$ where (V, U) is an admissible pair with high codimension, we may assume that $[X/G]$ is a quasiprojective scheme. We claim that $(X \times K)/G \to X/G$ is a Zariski locally trivial affine bundle. Since G is special by part (i) of Corollary 2.2.3 the principal G-bundle $X \to X/G$ is locally trivial for the Zariski topology and after replacing X/G by an appropriate open subset we may assume $X = G \times (X/G)$. We then have an isomorphism $(G \times (X/G) \times K)/G \cong (X/G) \times K$ given by the assignment $(g, x, k) \mapsto (x, k')$, where k' is defined by $g^{-1}(g, x, k) = (1, x, k')$. This proves the claim and hence the pullback of the first map is also an isomorphism by Corollary 1.3.3. $\square$

The main ingredient of the computation of $A^*\mathcal{D}\mathrm{isp}_1^{h,d}$ is the following proposition.

Proposition 2.3.2. *Let G be a connected split reductive group over a field k with split maximal torus T. Consider an isogeny $\varphi : L \to M$, where L and M are Levi components of parabolic subgroups P and Q of G. Assume $T \subset L$ and let $g_0 \in G(k)$ such that $\varphi(T) = {}^{g_0}T$. Let $\tilde{\varphi} : T \to T$ denote the isogeny φ followed by conjugation with g_0^{-1}. We write $S = \mathrm{Sym}(\widehat{T}) = A_T^*$ and $S_+ = A_T^{\geq 1}$. We have a natural action of $\tilde{\varphi}$ on S, that we will also denote by $\tilde{\varphi}$.*

Consider the action of L on G by φ-conjugation. If $W_G = W(G, T)$ and $W_L = W(L, T)$ denote the respective Weyl groups we have

$$A_L^*(G)_{\mathbb{Q}} = S_{\mathbb{Q}}^{W_L}/(f - \tilde{\varphi} f \mid f \in S_+^{W_G})_{\mathbb{Q}}.$$

If G is special we have

$$A_L^*(G) = S^{W_L}/(f - \tilde{\varphi} f \mid f \in S_+^{W_G}).$$

(*Note that the action of $\tilde{\varphi}$ on S^{W_G} is independent of the choice of g_0 since two choices differ by an element of $N_G(T)$.*)

Proof. The case of special G is proven in [Brokemper 2016, Proposition 1.1]. Let I denote the ideal $(f - \varphi f \mid f \in S_+^{W_G})_{\mathbb{Q}}$ in $S_{\mathbb{Q}}^{W_G}$. It remains to show $A_L^*(G)_{\mathbb{Q}} = S_{\mathbb{Q}}^{W_L}/I S_{\mathbb{Q}}^{W_L}$ in the nonspecial case. Using the same argumentation as in the special case we arrive at

$$A_T^*(G)_{\mathbb{Q}} = S_{\mathbb{Q}}/I S_{\mathbb{Q}}.$$

Now by Theorem 1.5.1 we know $A_L^*(G)_{\mathbb{Q}} = A_T^*(G)_{\mathbb{Q}}^{W_L}$. Since $S_{\mathbb{Q}}^{W_L} \hookrightarrow S_{\mathbb{Q}}$ is finite free [Demazure 1973, Theorem 2(d)], it is also faithfully flat. Hence we obtain $S_{\mathbb{Q}}^{W_L} \cap I S_{\mathbb{Q}} = I S_{\mathbb{Q}}^{W_L}$ and the assertion follows. □

In the following we will write c_i for the i-th elementary symmetric polynomial in the variables $t_1, \dots, t_h$ and $c_i^{(j,k)}$ will denote the i-th elementary symmetric polynomial in the variables $t_j, \dots, t_k$, where $1 \le j < k \le h$ and $1 \le i \le k - j + 1$. We then have $\mathbb{Z}[t_1, \dots, t_n]^{S_{h-d} \times S_d} = \mathbb{Z}[c_1^{(1,h-d)}, \dots, c_{h-d}^{(1,h-d)}, c_1^{(h-d+1,h)}, \dots, c_d^{(h-d+1,h)}]$.

Theorem 2.3.3. $A^*(\mathcal{D}isp_1^{h,d}) = A^*_{\mathrm{GL}_{h-d} \times \mathrm{GL}_d}(\mathrm{GL}_h)$

$$= \mathbb{Z}[t_1, \dots, t_n]^{S_{h-d} \times S_d}/((p-1)c_1, \dots, (p^h - 1)c_h),$$

where the $c_i^{(1,h-d)}$ and $c_i^{(h-d+1,h)}$ are the Chern classes of ${}^t\mathcal{L}ie^\vee$ and $\mathcal{L}ie$, respectively.

Proof. We have that $G_1^{h,d}$ is a split extension of the group $\mathrm{GL}_{h-d} \times \mathrm{GL}_d$ by the unipotent group

$$\left\{ \begin{pmatrix} I_{h-d} & * \\ * & I_d \end{pmatrix} \right\},$$

where $*$ denotes an arbitrary matrix (see Remark 2.1.2). The splitting is given by the canonical inclusion $\mathrm{GL}_{h-d} \times \mathrm{GL}_d \hookrightarrow G_1^{h,d}$. Hence by Lemma 1.4.7 we know

$$A^*(\mathcal{D}\mathrm{isp}_1^{h,d}) = A^*_{\mathrm{GL}_{h-d} \times \mathrm{GL}_d}(\mathrm{GL}_h),$$

where the action of $\mathrm{GL}_{h-d} \times \mathrm{GL}_d$ on GL_h is given by σ-conjugation. Since $\mathrm{GL}_{h-d} \times \mathrm{GL}_d$ is special with Weyl group $S_{h-d} \times S_d$ we obtain from Proposition 2.3.2

$$A^*_{\mathrm{GL}_{h-d} \times \mathrm{GL}_d}(\mathrm{GL}_h) = \mathbb{Z}[t_1, \dots, t_n]^{S_{h-d} \times S_d}/((p-1)c_1, \dots, (p^h - 1)c_h).$$

For this, note that the Frobenius σ acts on $S = \mathbb{Z}[t_1, \dots, t_n]$ via $\sigma t_i = p t_i$ and that c_i is a homogenous polynomial in $t_1, \dots, t_n$ of degree i.

The assertion that the $c_i^{(1,h-d)}$ and $c_i^{(h-d+1,h)}$ are the Chern classes of $\mathcal{L}ie$ and ${}^t\mathcal{L}ie^\vee$, respectively, follows from the following simple fact. Let us write $\mathcal{E}_d$

(resp. $\mathcal{E}_{h-d}$) for the vector bundle over $[*/\operatorname{GL}_d]$ (resp. $[*/\operatorname{GL}_{h-d}]$) that corresponds to the canonical representation of GL_d (resp. GL_{h-d}). Then $\mathcal{L}ie$ is the pullback of $\mathcal{E}_d$ under the natural map

$$\mathcal{D}\mathrm{isp}_1^{h,d} = [GL_h/G_1^{h,d}] \longrightarrow [*/(\operatorname{GL}_d \times GL_{h-d})] \longrightarrow [*/GL_d]$$

and similarly for ${}^t\mathcal{L}ie^\vee$. □

Corollary 2.3.4. $\operatorname{Pic}(\mathcal{D}isp_1^{h,d}) = \begin{cases} \mathbb{Z}/(p-1)\mathbb{Z} & \text{if } d = 0, h, \\ \mathbb{Z} \times \mathbb{Z}/(p-1)\mathbb{Z} & \text{otherwise.} \end{cases}$

A generator for the free part is $\det(\mathcal{L}ie)$ *and a generator for the torsion part is* $\det(\mathcal{L}ie \otimes {}^t\mathcal{L}ie^\vee)$.

Proof. Note $\operatorname{Pic} \mathcal{D}\mathrm{isp}_n^{h,d} = A^1\mathcal{D}\mathrm{isp}_n^{h,d}$ by [Edidin and Graham 1998, Corollary 1]. □

Remark 2.3.5. There is also a more direct approach to compute the above Picard groups. By using a theorem of Rosenlicht, namely that for irreducible varieties X and Y the natural map

$$\mathcal{O}(X)^* \times \mathcal{O}(Y)^* \to \mathcal{O}(X \times Y)^*$$

is surjective, it is not difficult to establish the exact sequence

$$\mathcal{O}(X)^*/k^* \longrightarrow \widehat{G} \longrightarrow \operatorname{Pic}^G(X) \longrightarrow \operatorname{Pic}(X)$$

for G connected and X an irreducible G-scheme. The first map assigns to a nonvanishing regular function on X its eigenvalue. In our case we have $G = \operatorname{GL}_{h-d} \times \operatorname{GL}_d$ and $X = \operatorname{GL}_h$. Then $\mathcal{O}(\operatorname{GL}_h)^*/k^* = \mathbb{Z}$ with generator given by the determinant and eigenvalue given by the character $(p-1)(\det_{\operatorname{GL}_{h-d}} + \det_{\operatorname{GL}_d}) \in \widehat{G}$. Since $\operatorname{Pic}(\operatorname{GL}_h) = 0$ we again obtain the result of the above Corollary.

Remark 2.3.6. The fact that $(\det \mathcal{L}ie \otimes \det {}^t\mathcal{L}ie^\vee)^{p-1}$ is trivial can also be seen directly as follows: $(\det \mathcal{L}ie \otimes \det {}^t\mathcal{L}ie^\vee)^{p-1}$ being trivial means that $\det \mathcal{L}ie \otimes \det {}^t\mathcal{L}ie^\vee$ is fixed under the pullback of the Frobenius map $\operatorname{Frob} : \mathcal{D}\mathrm{isp}_1^{2,1} \to \mathcal{D}\mathrm{isp}_1^{2,1}$ assigning to a display $\mathcal{P}$ over an $\mathbb{F}_p$-algebra R the display $\mathcal{P}^\sigma$ obtained by base change via the Frobenius $\sigma : R \to R$. But by definition of a truncated display we have an isomorphism

$$\Psi : L \oplus T \cong L^\sigma \oplus T^\sigma$$

of R-modules. Taking the determinant of Ψ yields the desired isomorphism

$$\det L \otimes \det T \cong \det L^\sigma \otimes \det T^\sigma.$$

Remark 2.3.7. Let us put this result into context by relating it to the corresponding result for elliptic curves. Let $\mathcal{M}_{1,1} \to \operatorname{Spec} k$ denote the moduli stack of elliptic curves. A morphism $\operatorname{Spec} R \to \mathcal{M}_{1,1}$ corresponds to a pair $(C \to \operatorname{Spec} R, \sigma)$ where

$C \to \operatorname{Spec} R$ is a smooth projective curve of genus 1 and $\sigma : \operatorname{Spec} R \to C$ is a smooth section. We now have the diagram

$$\begin{tikzcd}
\mathcal{M}_{1,1} \arrow[r] & \mathrm{BT}^{h=2,d=1} \arrow[r, "\phi"] \arrow[d] & \mathcal{D}\mathrm{isp}^{h=2,d=1} \arrow[d] \\
& \mathrm{BT}^{h=2,d=1}_{n=1} \arrow[r, "\phi_1"'] & \mathcal{D}\mathrm{isp}^{h=2,d=1}_{n=1}
\end{tikzcd}$$

where $\mathcal{M}_{1,1} \to \mathrm{BT}^{h=2,d=1}$ sends an elliptic curve C to its associated Barsotti–Tate group $C[p^\infty]$. Let us consider the pullback map $A^*(\mathcal{D}\mathrm{isp}^{2,1}_1) \to A^*(\mathcal{M}_{1,1})$. In characteristic p different from 2 and 3, Edidin and Graham computed $A^*(\mathcal{M}_{1,1}) = \mathbb{Z}[t]/(12t)$, where t is given by the first Chern class of the Hodge bundle on $\mathcal{M}_{1,1}$ [Edidin and Graham 1998, Proposition 21].

By construction of the truncated display functor the pullback of $\mathcal{L}ie$ to $\mathcal{M}_{1,1}$ is the dual of the Hodge bundle on $\mathcal{M}_{1,1}$. Since the dual of an elliptic curve is the elliptic curve itself, it follows from Remark 2.1.5 that the pullback of ${}^t\mathcal{L}ie^\vee$ is given by the Hodge bundle. Hence $A^*(\mathcal{D}\mathrm{isp}^{2,1}_1) \to A^*(\mathcal{M}_{1,1})$ is the map

$$\mathbb{Z}[t_1, t_2]/((p-1)c_1, (p^2-1)c_2) \to \mathbb{Z}[t]/(12t)$$

that sends t_1 to $-t$ and t_2 to t. Note that $p^2 - 1$ is divisible by 12 if and only if $p \geq 5$. In particular, there can be no such map for $p = 2, 3$, and we deduce that the description $A^*(\mathcal{M}_{1,1}) = \mathbb{Z}[t]/(12t)$ does not hold in characteristic 2 and 3.

2.4. ***The Chow ring of the stack of G-zips.*** Let us first consider the case of F-zips introduced in [Moonen and Wedhorn 2004]. We denote by F-zip the stack of F-zips over a field k of characteristic $p > 0$. For S a k-scheme F-zip(S) is the groupoid of F-zips over S. If $\tau : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ is a function with finite support we denote by F-zip$^\tau$ the open and closed substack of F-zips of type τ. Note that

$$F\text{-zip} = \coprod_\tau F\text{-zip}^\tau.$$

The stacks F-zip$^\tau$ are smooth Artin algebraic stacks over k which follows for example from the following representation as a quotient stack. Let X_τ denote the k-scheme whose S-valued points are given by

$$X_\tau(S) = \{\underline{M} = (M, C^\bullet, D_\bullet, \varphi_\bullet) \mid \underline{M} \text{ is an } F\text{-zip of type } \tau,\ M = \mathcal{O}_S^h\}.$$

This is a smooth scheme of dimension h^2. Here $h = \sum_{i \in \mathbb{Z}} \tau(i)$ is also called the height of $\underline{M}$. The group GL_h acts on X_τ by

$$G \cdot \underline{M} = (\mathcal{O}_S^h, G(C^\bullet), G(D_\bullet), G\varphi_\bullet(G^{-1})^\sigma).$$

It is easy to see that two F-zips over S of the above form are isomorphic if and only if they lie in the same $\mathrm{GL}_h(S)$-orbit. Thus

$$F\text{-zip}^{\tau} = [X_{\tau}/\,\mathrm{GL}_h].$$

An F-zip $\underline{M}$ of type τ with support in $\{0, 1\}$ over an $\mathbb{F}_p$-algebra R is just a tuple

$$\underline{M} = (M, C, D, \varphi_0, \varphi_1),$$

where M is a projective R-module with submodules C and D, which are direct summands of M and isomorphisms

$$\varphi_0 : C^{\sigma} \to M/D, \quad \varphi_1 : (M/C)^{\sigma} \to D.$$

Lemma 2.4.1. *Let R be an $\mathbb{F}_p$-algebra. Then we have an equivalence of categories*

$$\mathcal{D}isp_1(R) \to \coprod_{\tau,\mathrm{Supp}(\tau)\in\{0,1\}} F\text{-zip}^{\tau}(R)$$

given by

$$(L, T, \Psi) \mapsto (L \oplus T, T, \Psi^{\sigma}(L^{\sigma}), \Psi^{\sigma}|_{T^{\sigma}}, \Psi^{\sigma}|_{L^{\sigma}}).$$

The above assignment commutes with pulling back. In particular, we get an isomorphism of stacks

$$F\text{-zip}^{\tau} \cong \mathcal{D}isp_1^{\tau(0)+\tau(1),\tau(1)}$$

for every type τ with support lying in $\{0, 1\}$.

Proof. An inverse functor is given by the assignment

$$(M, C, D, \varphi_0, \varphi_1) \mapsto (C, M/C, \varphi_0 \oplus \varphi_1). \qquad \square$$

More generally, there is the stack of G-zips introduced in [Pink et al. 2011]. Here G refers to an arbitrary reductive group. It is defined as follows. Let $\mathcal{Z}$ be an algebraic zip datum, i.e., a 4-tuple (G, P, Q, φ) consisting of a split reductive group G, parabolic subgroups P and Q and an isogeny $\varphi : P/R_u(P) \to Q/R_u(Q)$. To $\mathcal{Z}$ one associates the group

$$E_{\mathcal{Z}} = \{(p, q) \in P \times Q \mid \varphi(\pi_P(p)) = \pi_Q(q)\}.$$

Now $E_{\mathcal{Z}}$ acts on G by the rule

$$((p, q), g) \mapsto pgq^{-1}$$

and the quotient stack $[G/E_{\mathcal{Z}}]$ is called the stack of G-zips. If G is connected $\mathcal{Z}$ is called a connected zip datum [Pink et al. 2011, Definition 3.1].

Let us recall how the stack of F-zips is just a special case of this construction. For this let $\tau : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ be a function with finite support, say $i_1 \leq \cdots \leq i_r$. If we denote $n_k = \tau(i_k)$, then $(n_1, \ldots, n_r)$ defines a partition of $h = \sum_k n_k$. We denote the standard parabolic of type $(n_1, \ldots, n_r)$ in GL_h by P_{τ}.

Lemma 2.4.2. *Let $\tau : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ be a function with finite support and let $\mathcal{Z} = (\mathrm{GL}_h, P_\tau, P_\tau^-, \sigma)$ be the algebraic zip datum with P_τ^- the opposite parabolic of P_τ and σ the Frobenius isogeny. Then there is an isomorphism of stacks*

$$[\mathrm{GL}_h / E_{\mathcal{Z}}] \xrightarrow{\sim} F\text{-zip}^\tau.$$

Proof. Let S be an k-scheme. We denote by $C_\tau^\bullet$ the descending filtration

$$C_\tau^\bullet = \mathcal{O}_S^h \supset \mathcal{O}_S^{n_1+\cdots+n_{r-1}} \supset \cdots \supset$$

in $\mathcal{O}_S^h$ given by the standard flag of type $(n_1, \ldots, n_r)$ and by $D_\bullet^{\tau^-}$ the ascending filtration

$$D_\bullet^{\tau^-} = 0 \subset \mathcal{O}_S^{n_r} \subset \cdots \subset \mathcal{O}_S^{n_r+\cdots+n_2} \subset \mathcal{O}_S^h$$

given by the flag of type opposite to $(n_1, \ldots, n_r)$. To $g \in \mathrm{GL}_h(S)$ we assign the F-zip

$$\underline{M}_g = (\mathcal{O}_S^h, C_\tau^\bullet, g(D_\bullet^{\tau^-}), \varphi_\bullet),$$

where φ is given by the restriction of g to the successive quotients of $C_\tau^\bullet$. Note that we can consider g as a σ-linear map.

If (p, q) is an element of $E_{\mathcal{Z}}$ we get an isomorphism $M_g \to M_{pgq^{-1}}$ of F-zips induced by p. The fact that p commutes with the φ_i is exactly the condition $\sigma(\pi(p)) = \pi(q)$. On the other hand if an isomorphism $p : M_g \to M_{g'}$ of F-zips is given, we see that $g'^{-1}pg$ preserves the flag of type opposite to $(n_1, \ldots, n_r)$. Thus $q = g'^{-1}pg \in P_\tau^-$ and again the compatibility of p with the φ_i implies the condition $\sigma(\pi(p)) = \pi(q)$. □

We can also use Proposition 2.3.2 to say something about the Chow ring of the stack of G-zips for an arbitrary connected algebraic zip datum.

Definition 2.4.3. We call an algebraic zip datum $\mathcal{Z} = (G, P, Q, \varphi)$ special, if G is special.

Theorem 2.4.4. *Let $\mathcal{Z} = (G, P, Q, \varphi)$ be a connected algebraic zip datum. Let $W_G = W(G, T)$ be the Weyl group of G and $W_L = W(L, T)$ be the Weyl group of a Levi component L of P with respect to a split maximal torus $T \subset L$ of G. Let $g_0 \in G(k)$ be such that $\varphi(T) = {}^{g_0}T$ and let $\tilde{\varphi} : T \to T$ denote the composition of φ followed by conjugation with g_0^{-1}. Then $\tilde{\varphi}$ induces an action on $S = \mathrm{Sym}(\widehat{T})$ that we will also denote by $\tilde{\varphi}$. We then have*

$$A^*([G/E_{\mathcal{Z}}])_{\mathbb{Q}} = S_{\mathbb{Q}}^{W_L} / (f - \tilde{\varphi} f \mid f \in S_+^{W_G})_{\mathbb{Q}}.$$

If $\mathcal{Z}$ is special we have

$$A^*([G/E_{\mathcal{Z}}]) = S^{W_L} / (f - \tilde{\varphi} f \mid f \in S_+^{W_G}).$$

(*Note that the action of $\tilde{\varphi}$ on S^{W_G} is independent of the choice of g_0 since two choices differ by an element of $N_G(T)$.*)

Proof. By definition of the group $E_{\mathcal{Z}}$ we have a split exact sequence

$$0 \longrightarrow R_u(P) \times R_u(Q) \longrightarrow E_{\mathcal{Z}} \longrightarrow L \longrightarrow 0,$$

where the splitting is given by $L \hookrightarrow E_{\mathcal{Z}},\ l \mapsto (l, \varphi(l))$. From Lemma 1.4.7 we deduce

$$A^*([G/E_{\mathcal{Z}}])_{\mathbb{Q}} = A^*_L(G)_{\mathbb{Q}},$$

where the action of L on G is given by φ-conjugation. If G is special the above equality holds over $\mathbb{Z}$. We conclude by Proposition 2.3.2. □

Example 2.4.5. We consider the case $\mathcal{Z} = (\mathrm{Sp}(2n),\ P,\ P^-,\ \sigma)$, where σ denotes the q-th power Frobenius. Recall that $\mathrm{Sp}(2n)$ is special and the Weyl group of $\mathrm{Sp}(2n)$ is the wreath product $S_n \wr (\mathbb{Z}/2\mathbb{Z}) = S_n \ltimes (\mathbb{Z}/2\mathbb{Z})^n$. It acts on $\mathrm{Sym}(\widehat{T}) = \mathbb{Z}[t_1, \dots, t_n]$ in the following way. S_n acts by permuting the variables $t_1, \dots, t_n$ and after identifying $\mathbb{Z}/2\mathbb{Z} = \{\pm 1\}$ an element $(\varepsilon_1, \dots, \varepsilon_n) \in \mathbb{Z}/2\mathbb{Z}^n$ acts by $(\varepsilon_1, \dots, \varepsilon_n) \cdot t_i = \varepsilon_i t_i$.

If P is a Borel we obtain from the above theorem that

$$A^*([Sp(2n)/E_{\mathcal{Z}}])) = \mathbb{Z}[t_1, \dots, t_n]/((q^2-1)c_1(\underline{t}^2), \dots, (q^{2n}-1)c_n(\underline{t}^2)).$$

If P is the maximal parabolic subgroup fixing a maximal isotropic subspace then $L = \mathrm{GL}_n$ and $W_L = S_n$ and therefore

$$A^*([Sp(2n)/E_{\mathcal{Z}}]) = \mathbb{Z}[c_1, \dots, c_n]/((q^2-1)c_1(\underline{t}^2), \dots, (q^{2n}-1)c_n(\underline{t}^2)).$$

It turns out that a $\mathbb{Q}$-basis of the Chow ring of the stack of G-zips is given by the closures of the orbits of the action of $E_{\mathcal{Z}}$ on G. To prove this let us introduce the naive Chow group of a quotient stack.

Definition 2.4.6. Let G be an algebraic group and let X be a G-scheme. Let $Z_*([X/G])$ be the free abelian group generated by the set of G-invariant closed subvarieties of X graded by dimension. Let $W_i([X/G])$ be the group $\bigoplus_Y k(Y)^G$, where the sum goes over all G-invariant closed subvarieties of X of dimension $i+1$. There is the usual divisor map $\mathrm{div} : W_i([X/G]) \to Z_i([X/G])$ and we define the i-th naive Chow group of $[X/G]$ to be

$$A^o_i[X/G] = Z_i([X/G])/\operatorname{div}(W_i([X/G])).$$

Remark 2.4.7. There is more generally a definition of naive Chow groups for arbitrary algebraic stacks ([Kresch 1999, Definition 2.1.4]) which in the case of a quotient stack agrees with the one given above. Thus the above definition is independent of the presentation as a quotient stack.

Remark 2.4.8. There is a natural map $A^o_*[X/G] \to A_*[X/G]$. When X is Deligne–Mumford, i.e., the stabilizer of every point is finite and geometrically reduced, the induced map $A^o_*[X/G]_{\mathbb{Q}} \to A_*[X/G]_{\mathbb{Q}}$ is an isomorphism of groups and an isomorphism of rings if $[X/G]$ is smooth [Kresch 1999, Theorem 2.1.12(ii)].

The stack of G-zips is not Deligne–Mumford. However, we still have the following proposition.

Proposition 2.4.9. *Let G be a connected algebraic group and X be an admissible G-scheme* (see Definition 1.2.3) *with finitely many orbits such that the stabilizer of every point is an extension of a finite group by a unipotent group. Then $A^o_*[X/G]_{\mathbb{Q}} \to A_*[X/G]_{\mathbb{Q}}$ is an isomorphism.*

Proof. We prove this by induction on the number of orbits. Let U denote the open G-orbit and W its complement. We have a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & A^o_*[W/G]_{\mathbb{Q}} \arrow[r] \arrow[d] & A^o_*[X/G]_{\mathbb{Q}} \arrow[r] \arrow[d] & A^o_*[U/G]_{\mathbb{Q}} \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & A_*[W/G]_{\mathbb{Q}} \arrow[r] & A_*[X/G]_{\mathbb{Q}} \arrow[r] & A_*[U/G]_{\mathbb{Q}} \arrow[r] & 0 \end{tikzcd}$$

and we claim that the rows of this diagram are exact. Since there are only finitely many orbits every G-invariant subvariety Y of X is the closure of a G-orbit. Since Y admits a dense G-invariant subset every G-invariant rational function on Y is constant. It follows that $A^o_*[X/G] = \bigoplus_Z \mathbb{Z}[\overline{Z}]$ where the sum goes over all G-orbits Z of X. From this we obtain the exactness of the top row. For the exactness of the lower row we need to see that the pullback map $A_*([X/G], 1)_{\mathbb{Q}} \to A_*([U/G], 1)_{\mathbb{Q}}$ is surjective. But $[U/G]$ is isomorphic to the classifying space of the stabilizer group scheme of U. By assumption and Corollary 1.4.4 we get that $A_*([U/G], m)_{\mathbb{Q}} \to A_*(B\{0\}, m)_{\mathbb{Q}}$ is an isomorphism. Equivalently the pullback of the structure morphism $[U/G] \to \operatorname{Spec} k$ is an isomorphism for the higher Chow groups with rational coefficients and hence the claim follows.

Now the right vertical arrow is an isomorphism since both groups are isomorphic to $\mathbb{Q}$. By induction we may assume that the first vertical arrow is also an isomorphism. □

Recall that an algebraic zip datum $\mathcal{Z}$ is called orbitally finite if G has finitely many $E_{\mathcal{Z}}$-orbits [Pink et al. 2011, Definition 7.2].

Theorem 2.4.10. *Let $\mathcal{Z}$ be an orbitally finite connected algebraic zip datum and $[G/E_{\mathcal{Z}}]$ be the corresponding stack of G*-zips*. Then the following assertions hold.*

(i) *$A^o_*[G/E_{\mathcal{Z}}]_{\mathbb{Q}} \to A_*[G/E_{\mathcal{Z}}]_{\mathbb{Q}}$ is an isomorphism.*

(ii) *$A^o_*[G/E_{\mathcal{Z}}] = \bigoplus_Z \mathbb{Z}[\overline{Z}]$ where the sum goes over all orbits Z.*

In particular, *the dimension of $A_*[G/E_{\mathcal{Z}}]_{\mathbb{Q}}$ as a $\mathbb{Q}$-vector space is equal to the number of orbits.*

Proof. The assumption of the previous proposition on the stabilizer group schemes hold by [Pink et al. 2011, Theorem 8.1]. □

Corollary 2.4.11. *Let $\mathcal{Z} = (G, P, Q, \varphi)$ be a connected algebraic zip datum and T be a split maximal torus of G in a Levi component L of P. If $\mathcal{Z}$ is orbitally finite the $\mathbb{Q}$-vectorspace $A^*([G/E_{\mathcal{Z}}])_{\mathbb{Q}}$ is finite dimensional of dimension $|W_G/W_L|$, where as usual $W_G = W(G, T)$ is the Weyl group of G and $W_L = W(L, T)$ is the Weyl group of L.*

Proof. By the above theorem $\dim_{\mathbb{Q}} A^*([G/E_{\mathcal{Z}}])_{\mathbb{Q}}$ equals the number of $E_{\mathcal{Z}}$-orbits in G. This number equals $|W_G/W_L|$ by [Pink et al. 2011, Theorem 7.5]. □

In the case of F-zips the above results read as follows.

Corollary 2.4.12. *Let $\tau : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ be a function with finite support $i_1 \leq \cdots \leq i_r$ and $n_k = \tau(i_k)$. Let $h = \sum_i n_i$ be its height. Then:*

(i) $$A^* F\text{-zip}^\tau = \mathbb{Z}[t_1, \ldots, t_h]^{S_{n_1} \times \cdots \times S_{n_r}} / ((p-1)c_1, \ldots, (p^h - 1)c_h),$$

with c_i the i-th elementary symmetric polynomial in the variables $t_1, \ldots, t_h$.

(ii) $$\operatorname{Pic}(F\text{-zip}^\tau) = \mathbb{Z}^{r-1} \times \mathbb{Z}/(p-1)\mathbb{Z}.$$

(iii) $$\dim_{\mathbb{Q}} A^*(F\text{-zip}^\tau)_{\mathbb{Q}} = \frac{h!}{n_1! \times \cdots \times n_r!}.$$

2.5. *The Chow ring of* $\mathbf{BT}_n$. The goal of this section is to prove the following result.

Theorem 2.5.1. *The pullback $\phi_n^* : A^*(\mathcal{D}isp_n) \to A^*(\mathrm{BT}_n)$ is injective and an isomorphism after inverting p.*

We know that $\mathcal{D}\mathrm{isp}_n = \coprod_{d \leq h} \mathcal{D}\mathrm{isp}_n^{h,d}$ is a decomposition into open and closed substacks. The same holds for BT_n and the morphism ϕ_n maps $\mathrm{BT}_n^{h,d}$ to $\mathcal{D}\mathrm{isp}_n^{h,d}$. It suffices to prove the theorem for the restriction of ϕ_n to $\mathrm{BT}_n^{h,d}$. The following proposition is the crucial point in the proof of Theorem 2.5.1.

Proposition 2.5.2. *Let L be a field extension of k and* $\operatorname{Spec} L \to \mathcal{D}isp_n$ *be a morphism. Then there is a finite field extension L' of L of p-power degree and an infinitesimal commutative group scheme A over L' such that the fiber ϕ_n^{-1}*$(\operatorname{Spec} L')$ *is the classifying space of A.*

Proof. The diagonal $\Delta : \mathrm{BT}_n \to \mathrm{BT}_n \times_{\mathcal{D}\mathrm{isp}_n} \mathrm{BT}_n$ is flat and surjective by [Lau 2013, Theorem 4.7]. This means that two Barsotti–Tate groups of level n having the same associated display become isomorphic when pulled back to a suitable fppf-covering. It follows that the fiber $(\mathrm{BT}_n)_L$ of a display P over some field L is a gerbe over L. If L is perfect there is a truncated Barsotti–Tate group G over L with $\phi_n(G) = P$, i.e., $(\mathrm{BT}_n)_L$ is a neutral gerbe. In this case $(\mathrm{BT}_n)_L = B\underline{\mathrm{Aut}}^o(G)$ where $\underline{\mathrm{Aut}}^o(G) = \operatorname{Ker}(\underline{\mathrm{Aut}} G \to \underline{\mathrm{Aut}} P)$ is commutative and infinitesimal again by [Lau 2013, Theorem 4.7]. If L is not perfect we may consider the perfect hull $L^{p^{-\infty}}$ in an algebraic closure of L. Then $L \subset L^{p^{-\infty}}$ is purely inseparable and

$(\mathrm{BT}_n)_L(L^{p^{-\infty}})$ is nonempty. Since $(\mathrm{BT}_n)_L(L^{p^{-\infty}}) = \varinjlim_{L'}(\mathrm{BT}_n)_L(L')$, where the limit goes over all finite subextensions $L \subset L' \subset L^{p^{-\infty}}$, we find some L' such that $(\mathrm{BT}_n)_{L'}$ has a section corresponding to a truncated Barsotti–Tate group G over L'. Thus $A = \underline{\mathrm{Aut}}^o(G)$ and L' have the desired properties. □

Remark 2.5.3. Over the open and closed substack of BT_n consisting of level-n BT-groups of constant dimension d and codimension c the degree of $\mathrm{Aut}^o(G^{\mathrm{univ}})$ is p^{ncd}. See Remark 4.8 in [Lau 2013].

Note that $\mathcal{D}\mathrm{isp}_n^{h,d}$ and $\mathrm{BT}_n^{h,d}$ both admit admissible presentations in the sense of Definition 1.2.3. In the case of $\mathcal{D}\mathrm{isp}_n^{h,d}$ this follows from Theorem 2.1.3 and Lemma 1.2.2. To obtain the assertion for $\mathrm{BT}_n^{h,d}$ we use [Wedhorn 2001, Proposition 1.8] which yields a presentation $\mathrm{BT}_n^h = [Y_n^h / \mathrm{GL}_{p^{nh}}]$ with Y_n^h quasiaffine and of finite type over k. Now BT_n^h is smooth over $\operatorname{Spec} k$ [Lau 2013]. Hence Y_n^h is also smooth and in particular normal and equidimensional.

We now consider the flat pullback map

$$\phi_n^* : A_*(\mathcal{D}\mathrm{isp}_n^{h,d}, m) \to A_*(\mathrm{BT}_n^{h,d}, m)$$

from Lemma 1.2.6.

Proposition 2.5.4. $\phi_n^* : A_*(\mathcal{D}isp_n^{h,d}, m) \to A_*(\mathrm{BT}_n^{h,d}, m)$ *is an isomorphism after inverting* p.

Proof. Let us write $\mathscr{X} = \mathrm{BT}_n^{h,d}$ and $\mathscr{Y} = \mathcal{D}\mathrm{isp}_n^{h,d}$. We fix some $i_o \in \mathbb{Z}$ and show that $\phi_n : A_{i_o}(\mathcal{D}\mathrm{isp}_n^{h,d}, m)_p \to A_{i_o}(\mathrm{BT}_n^{h,d}, m)_p$ is an isomorphism.

Consider an approximation of $\mathscr{Y}$ (see Convention 1.1.1) by a quasiprojective scheme $Y \to \mathscr{Y}$ so that $A_{i_o}(\mathscr{Y}, m) = A_{i_o}(Y, m)$ and similarly an approximation $X \to \mathscr{X}$ of $\mathscr{X}$. Let r denote the relative dimension of $X \to \mathscr{X}$. Let Z be the fiber product $X \times_{\mathscr{Y}} Y$. The morphism $Z \to Y$ is then smooth of relative dimension r and we need to see that the pullback $A_{i_o}(Y, m)_p \to A_{i_o+r}(Z, m)_p$ is an isomorphism. Note that Z is again quasiprojective since it is open in a vector bundle over the quasiprojective scheme X (see Remark 1.2.5). We have the cartesian diagram

$$\begin{tikzcd}
Z_y \arrow[r] \arrow[d] & \mathscr{X}_{k(y)} \arrow[r] \arrow[d] & \operatorname{Spec} k(y) \arrow[d] \\
Z \arrow[r] \arrow[d] & \mathscr{X}_Y \arrow[r] \arrow[d] & Y \arrow[d] \\
X \arrow[r] & \mathscr{X} \arrow[r] & \mathscr{Y}
\end{tikzcd}$$

By Lemma 1.3.2 it suffices to see that $A_i(\operatorname{Spec} k(y), m)_p \to A_{i+r}(Z_y, m)_p$ with $i = i_o - \dim \overline{\{y\}}$ is an isomorphism. According to the previous proposition there is a finite field extension K of $k(y)$ of p-power degree such that $\mathscr{X}_K = BA$ holds for an infinitesimal group scheme A over K.

Since Z_K is open in a vector bundle over $\mathscr{X}_K$ of rank r we have $Z_K = U/A$, where U is open in a representation V of A. Note that V is of dimension r. Hence by choosing $\operatorname{codim} X^c$ to be big enough, we may assume $A_i(\operatorname{Spec} K, m) \to A_{i+r}(U, m)$ is an isomorphism. Since A is of p-power degree it follows that the map $A_i(\operatorname{Spec} K, m)_p \to A_{i+r}(Z_K, m)_p$ is an isomorphism. Now since the field extension $K \supset k(y)$ is of p-power degree it follows from Lemma 1.3.1 that

$$A_i(\operatorname{Spec} k(y), m)_p \to A_{i+r}(Z_y, m)_p$$

is also an isomorphism. We are done. □

Proof of Theorem 2.5.1. Since BT_n and $\mathcal{D}\mathrm{isp}_n$ are smooth the pullback

$$(\phi_n)_p^* : A^*(\mathcal{D}\mathrm{isp}_n)_p \to A^*(\mathrm{BT}_n)_p$$

is an isomorphism by Lemma 1.2.6 and the proposition above. We already know $A^*(\mathcal{D}\mathrm{isp}_n)$ is p-torsion free by Theorems 2.3.1 and 2.3.3. Thus ϕ_n^* is injective. □

Gathering the results of Section 2, we obtain the following theorem:

Theorem 2.5.5. (i) *We have*

$$A^*(\mathrm{BT}_n^{h,d})_p = \mathbb{Z}[p^{-1}][t_1, \dots, t_h]^{S_d \times S_{h-d}} / ((p-1)c_1, \dots, (p^h - 1)c_h),$$

where c_i denotes the i-th elementary symmetric polynomial in the variables $t_1, \dots, t_h$, and $t_1, \dots, t_d$ and $t_{d+1}, \dots, t_h$ are the Chern roots of $\mathcal{L}ie$ and ${}^t\mathcal{L}ie^\vee$, respectively.

(ii) *We have* $\dim_{\mathbb{Q}} A^*(\mathrm{BT}_n^{h,d})_{\mathbb{Q}} = \binom{h}{d}$ *and a basis is given by the cycles of the closures of the EO strata.*

(iii) $$(\operatorname{Pic} \mathrm{BT}_n^{h,d})_p = \begin{cases} \mathbb{Z}[p^{-1}]/(p-1) & \text{if } d = 0, h, \\ \mathbb{Z}[p^{-1}] \times \mathbb{Z}[p^{-1}]/(p-1) & \text{otherwise}, \end{cases}$$

where the generator for the free part is $\det(\mathcal{L}ie)$ *and for the torsion part is* $\det(\mathcal{L}ie \otimes {}^t\mathcal{L}ie^\vee)$.

Proof. By Theorem 2.5.1 we know $A^*(\mathcal{D}\mathrm{isp}_n^{h,d})_p \cong A^*(\mathrm{BT}_n^{h,d})_p$. Further, we have

$$A^*(\mathcal{D}\mathrm{isp}_n^{h,d}) \cong A^*(\mathcal{D}\mathrm{isp}_1^{h,d})$$

by Theorem 2.3.1 and $A^*(\mathcal{D}\mathrm{isp}_1^{h,d})$ was computed in Theorem 2.3.3. This proves part (i). By Lemmas 2.4.1 and 2.4.2 we know that $\mathcal{D}\mathrm{isp}_1^{h,d}$ is isomorphic to the stack $[\mathrm{GL}_h / E_{\mathcal{Z}}]$ corresponding to the Frobenius zip datum $\mathcal{Z} = (\mathrm{GL}_h, P, P^-, \sigma)$, where P is the standard parabolic of type (d, h), P^- is the opposite parabolic and σ is the Frobenius isogeny. Now the dimension of $A^*(\mathcal{D}\mathrm{isp}_1^{h,d})_{\mathbb{Q}}$ as a $\mathbb{Q}$-vectorspace follows from Corollary 2.4.12 and a basis is given by Theorem 2.4.10. This proves (ii). Finally (iii) follows from (i) together with the fact that $\operatorname{Pic} \mathrm{BT}_n^{h,d} = A^1(\mathrm{BT}_n^{h,d})$. □

Acknowledgements

This article is part of the author's PhD thesis. He wishes to thank his advisor Eike Lau for guidance and support, and for many helpful comments on earlier versions of this article. He is also grateful to Jean-Stefan Koskivirta for many valuable discussions on G-zips and mathematics in general.

References

[Bloch 1986] S. Bloch, "Algebraic cycles and higher K-theory", *Adv. in Math.* **61**:3 (1986), 267–304. MR Zbl

[Brokemper 2016] D. Brokemper, "On the Chow ring of the classifying space of some Chevalley groups", preprint, 2016. arXiv

[Demazure 1973] M. Demazure, "Invariants symétriques entiers des groupes de Weyl et torsion", *Invent. Math.* **21** (1973), 287–301. MR Zbl

[Edidin and Fulghesu 2009] D. Edidin and D. Fulghesu, "The integral Chow ring of the stack of hyperelliptic curves of even genus", *Math. Res. Lett.* **16**:1 (2009), 27–40. MR Zbl

[Edidin and Graham 1997] D. Edidin and W. Graham, "Characteristic classes in the Chow ring", *J. Algebraic Geom.* **6**:3 (1997), 431–443. MR Zbl

[Edidin and Graham 1998] D. Edidin and W. Graham, "Equivariant intersection theory", *Invent. Math.* **131**:3 (1998), 595–634. MR Zbl

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

[Gillet 1981] H. Gillet, "Riemann–Roch theorems for higher algebraic K-theory", *Adv. in Math.* **40**:3 (1981), 203–289. MR Zbl

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

[Kresch 1999] A. Kresch, "Cycle groups for Artin stacks", *Invent. Math.* **138**:3 (1999), 495–536. MR Zbl

[Lau 2013] E. Lau, "Smoothness of the truncated display functor", *J. Amer. Math. Soc.* **26**:1 (2013), 129–165. MR Zbl

[Levine 2001] M. Levine, "Techniques of localization in the theory of algebraic cycles", *J. Algebraic Geom.* **10**:2 (2001), 299–363. MR Zbl

[Moonen and Wedhorn 2004] B. Moonen and T. Wedhorn, "Discrete invariants of varieties in positive characteristic", *Int. Math. Res. Not.* **2004**:72 (2004), 3855–3903. MR Zbl

[Mumford et al. 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Math. (2) **34**, Springer, 1994. MR Zbl

[Pink et al. 2011] R. Pink, T. Wedhorn, and P. Ziegler, "Algebraic zip data", *Doc. Math.* **16** (2011), 253–300. MR Zbl

[Quillen 1973] D. Quillen, "Higher algebraic K-theory. I", pp. 85–147 in *Algebraic K-theory, I: Higher K-theories*, edited by H. Bass, Lecture Notes in Math. **341**, Springer, 1973. MR Zbl

[Raynaud 1970] M. Raynaud, *Faisceaux amples sur les schémas en groupes et les espaces homogènes*, Lecture Notes in Mathematics **119**, Springer, 1970. MR Zbl

[Thomason 1988] R. W. Thomason, "Equivariant algebraic vs. topological K-homology Atiyah–Segal-style", *Duke Math. J.* **56**:3 (1988), 589–636. MR Zbl

[Totaro 1999] B. Totaro, “The Chow ring of a classifying space”, pp. 249–281 in *Algebraic K-theory*, edited by W. Raskind and C. Weibel, Proc. Sympos. Pure Math. **67**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Totaro 2014] B. Totaro, “Chow groups, Chow cohomology, and linear varieties”, *Forum Math. Sigma* **2** (2014). MR Zbl

[Vistoli 1989] A. Vistoli, “Characteristic classes of principal bundles in algebraic intersection theory”, *Duke Math. J.* **58**:2 (1989), 299–315. MR Zbl

[Wedhorn 2001] T. Wedhorn, “The dimension of Oort strata of Shimura varieties of PEL-type”, pp. 441–471 in *Moduli of abelian varieties* (Texel Island, 1999), edited by C. Faber et al., Progr. Math. **195**, Birkhäuser, Basel, 2001. MR Zbl

[Zink 2002] T. Zink, “The display of a formal p-divisible group”, pp. 127–248 in *Cohomologies p-adiques et applications arithmétiques, I*, edited by P. Berthelot et al., Astérisque **278**, Société Mathématique de France, 2002. MR Zbl

Received February 2, 2018. Revised April 16, 2018.

DENNIS BROKEMPER
UNIVERSITY OF PADERBORN
PADERBORN
GERMANY
dennis.brokemper@gmx.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.305

HYPERBOLIC MANIFOLDS CONTAINING HIGH TOPOLOGICAL INDEX SURFACES

MARION CAMPISI AND MATT RATHBUN

If a graph is in bridge position in a 3-manifold so that the graph complement is irreducible and boundary-irreducible, we generalize a result of Bachman and Schleimer to prove that the complexity of a surface properly embedded in the complement of the graph bounds the graph distance of the bridge surface. We use this result to construct, for any natural number n, a hyperbolic manifold containing a surface of topological index n.

1. Introduction

It has become increasingly common and useful to measure distances in complexes associated to surfaces between certain important subcomplexes associated with the surface embedded in a 3-manifold. These techniques provide a means to indicate the inherent complexity of links in a manifold, decomposing surfaces, or the manifold itself. Bachman [2010] defined the topological index of a surface as a topological analogue of the index of an unstable minimal surface. When the distance is small, the notion of topological index refines this distance, by looking at the *homotopy type* of a certain subcomplex.

In the same way that incompressible surfaces share important properties with strongly irreducible surfaces (distance > 2) despite being compressible, the topological index provides a degree of measurement of how similar irreducible, but weakly reducible (distance $= 1$) surfaces are to incompressible surfaces. Bachman [2012a; 2012b; 2012c] has shown that surfaces with a well-defined topological index in a 3-manifold can be put into a sort of normal form with respect to a triangulation of the manifold, generalizing the ideas of normal form introduced by Kneser [1929] and almost normal form introduced by Rubinstein [1995], and mirroring results about geometrically minimal surfaces due to Colding and Minicozzi [2004a; 2004b; 2004c; 2004d; 2015].

Lee [2015] has shown that an irreducible manifold containing an incompressible surface contains topologically minimal surfaces of arbitrarily high genus, but has

MSC2010: primary 55P15, 57M20, 57M27; secondary 57M10, 57M15.
Keywords: topological index, topologically minimal, hyperbolic, bridge position, distance, bridge distance, graph.

only shown that the topological index of such surfaces is at least two. Bachman and Johnson [2010] showed that surfaces of arbitrarily high index exist. These surfaces are the lifts of Heegaard surfaces in an n-fold cover of a manifold obtained by gluing together boundary components of the complement of a link in S^3. A byproduct of their construction is that the resulting manifolds are toroidal.

This leaves open the question of whether the much more ubiquitous class of hyperbolic manifolds can also contain high topological index surfaces. Here we construct certain hyperbolic manifolds containing such surfaces. We generalize the construction in [Bachman and Johnson 2010] by gluing along the boundary components of the complement of a graph in S^3 to show:

Theorem 1.1. *There is a closed* 3*-manifold* M^1, *with an index* 1 *Heegaard surface* S, *such that for each* n, *the lift of* S *to some* n*-fold cover* M^n *of* M^1 *has topological index* n. *Moreover*, M^n *is hyperbolic for all* n.

In order to guarantee the hyperbolicity of M^n, we must rule out the existence of high Euler characteristic surfaces in the graph complement. To that end, we define the *graph distance*, $d_{\mathcal{G}}$, of graphs in S^3, an analogue of bridge distance of links. In the spirit of Hartshorn [2002] and Bachman and Schleimer [2005], we show that the complexity of an essential surface is bounded below by the graph bridge distance:

Theorem 1.2. *Let* Γ *be a graph in a closed*, *orientable* 3*-manifold*, M, *which is in bridge position with respect to a Heegaard surface*, B, *so that* $M \setminus n(\Gamma)$ *is irreducible and boundary-irreducible. Let* S *be a properly embedded*, *orientable*, *incompressible*, *boundary-incompressible*, *non-boundary-parallel surface in* $M \setminus n(\Gamma)$. *Then* $d_{\mathcal{G}}(B, \Gamma)$ *is bounded above by* $2(2g(S) + |\partial S| - 1)$.

In Section 2 we lay out the definitions of the various complexes and distances we will use, and prove Theorem 1.2. In Section 3, we prove Theorem 1.1.

2. Definitions

Given a link $\mathcal{L} \subset S^3$, a *bridge sphere for* $\mathcal{L}$ is a sphere, B, embedded in S^3, intersecting the link $\mathcal{L}$ transversely, and dividing S^3 into two 3-balls, V and W, so that there exist disks D_V and D_W properly embedded in V and W, respectively, so that $\mathcal{L} \cap V \subset D_V$ and $\mathcal{L} \cap W \subset D_W$ are each a collection of arcs. If there are b arcs, the link is said to be b-bridge with respect to B.

Goda [1997] introduced the notion of a bridge sphere for a spatial θ-graph, and this was extended by Ozawa [2012]. A *bridge sphere for a (spatial) graph* Γ is a sphere, B, embedded in S^3, intersecting Γ transversely in the interior of edges, and dividing S^3 into two 3-balls, V and W, so that there exist disks D_V and D_W properly embedded in V and W, respectively, so that $\Gamma \cap V \subset D_V$ and $\Gamma \cap W \subset D_W$ are each a collection of trees and/or arcs.

If B is a bridge sphere for a link $\mathcal{L}$, then a *bridge disk* is a disk properly embedded in one of the components of $\overline{(S^3 \smallsetminus n(\mathcal{L})) \smallsetminus B}$, whose boundary consists of exactly two arcs, meeting at their endpoints, with one arc essential in $B \smallsetminus n(\mathcal{L})$, and the other essential in $\partial n(\mathcal{L}) \smallsetminus B$. We refer to the arc in the boundary of the disk that is contained in B as a *bridge arc*. Similarly, if B is a bridge sphere for a graph Γ, then a *graph-bridge disk* is a disk properly embedded in one of the components of $\overline{(S^3 \smallsetminus n(\Gamma)) \smallsetminus B}$, whose boundary consists of exactly two arcs, meeting at their endpoints, with one arc essential in $B \smallsetminus n(\Gamma)$, and the other essential in $\partial n(\Gamma) \smallsetminus B$. We refer to the arc in the boundary of the disk that is contained in B as a *graph-bridge arc*.

Definition 2.1. The *curve complex* for a surface B with (possibly empty) boundary is the complex with vertices corresponding to the isotopy classes of essential simple closed curves in B, so that a collection of vertices defines a simplex if representatives of the corresponding isotopy classes can be chosen to be pairwise disjoint. We will denote the curve complex for a surface B by $\mathcal{C}(B)$.

Definition 2.2. The *arc and curve complex* for a surface B' with boundary is the complex with vertices corresponding to the (free) isotopy classes of essential simple closed curves and properly embedded arcs in B'. A collection of vertices defines a simplex if representatives of the corresponding isotopy classes can be chosen to be pairwise disjoint. We will denote the arc and curve complex for a surface B' by $\mathcal{AC}(B')$.

If B is a surface embedded in a manifold, and a 1-dimensional complex intersects B transversely, we will refer to the surface obtained by removing a neighborhood of the 1-complex by B'. We will often refer to $\mathcal{C}(B')$ simply by $\mathcal{C}(B)$, and $\mathcal{AC}(B')$ simply by $\mathcal{AC}(B)$.

Definition 2.3. Let B be a surface with at least two distinct, essential curves. Given two collections X and Y of vertices in the complex $\mathcal{C}(B)$ (resp., $\mathcal{AC}(B)$), the distance between X and Y, denoted $d_{\mathcal{C}(B)}(X, Y)$ (resp., $d_{\mathcal{AC}(B)}(X, Y)$), is the minimal number of edges in any path in $\mathcal{C}(B)$ (resp., $\mathcal{AC}(B)$) from a vertex in X to a vertex in Y. When the surface is understood, we often just write $d_{\mathcal{C}}$ (resp., $d_{\mathcal{AC}}$).

We will be working with four subtly different but closely related subcomplexes, and some associated notions of distance.

Definition 2.4. Let B be a properly embedded surface separating a manifold M into two components, V and W. Define the *disk set of* V (resp., W), denoted $\mathcal{D}_V \subset \mathcal{C}(B)$, (resp., $\mathcal{D}_W \subset \mathcal{C}(B)$), as the set of all vertices corresponding to essential simple closed curves in B that bound embedded disks in V (resp., W). Define the *disk set of* B, denoted $\mathcal{D}_B$, as the set of all vertices corresponding to essential simple closed curves in B that bound embedded disks in M.

Definition 2.5. Let B be a bridge sphere for a link $\mathcal{L}$, bounding 3-balls V and W, with at least 6 marked points corresponding to the transverse intersections of $\mathcal{L}$ with B. The *distance of the bridge surface*, denoted $d_{\mathcal{C}}(B, \mathcal{L})$, is $d_{\mathcal{C}(B')}(\mathcal{D}_V, \mathcal{D}_W)$, the distance in the curve complex of B' between $\mathcal{D}_V$ and $\mathcal{D}_W$.

The fundamental building block in our construction will be the exterior of a graph that is highly complex as viewed from the arc and curve complex. The existence of such a block will follow from a result of Blair, Tomova, and Yoshizawa, using "warped pants decompositions" and Dehn twists to construct gluing maps resulting in high bridge distance link complements. It is a special case of [Blair et al. 2013, Corollary 5.3 and the proof of Theorem 4.9].

Theorem 2.6 [Blair et al. 2013]. *Given nonnegative integers b_1, b_2 and d, with $b_1 + b_2 \geq 3$, there exists a 2-component link $\mathcal{L}$ in S^3, and a bridge sphere B for $\mathcal{L}$ so that $\mathcal{L}$ is $(b_1{+}b_2)$-bridge with respect to B, the components of $\mathcal{L}$ are b_1- and b_2-bridge with respect to B, and $d_{\mathcal{C}}(B, \mathcal{L}) \geq d$.*

Definition 2.7. Let B be a bridge sphere for a link $\mathcal{L}$, bounding 3-balls V and W. Define the *bridge disk set* of V (resp., W), denoted $\mathcal{BD}_V \subset \mathcal{AC}(B)$ (resp., $\mathcal{BD}_W$), as the set of all vertices either corresponding to essential simple closed curves in B' that bound embedded disks in $V \setminus \mathcal{L}$ (resp., $W \setminus \mathcal{L}$), or corresponding to bridge arcs in B' contained in the boundaries of bridge disks in V (resp., W).

Definition 2.8. Let B be a bridge sphere for a link $\mathcal{L}$, bounding 3-balls V and W. The *bridge distance of the bridge surface B*, which we denote by $d_{\mathcal{BD}}(B, \mathcal{L})$, is $d_{\mathcal{AC}(B')}(\mathcal{BD}_V, \mathcal{BD}_W)$, the distance in the arc and curve complex of B' between $\mathcal{BD}_V$ and $\mathcal{BD}_W$.

Lemma 2.9 [Blair et al. 2017, Lemma 2]. *If B is a bridge surface which is not a sphere with four or fewer punctures, then $d_{\mathcal{BD}}(B, \mathcal{L}) \leq d_{\mathcal{C}}(B, \mathcal{L}) \leq 2d_{\mathcal{BD}}(B, \mathcal{L})$.*

Definition 2.10. Let B be a bridge sphere for graph Γ, bounding 3-balls V and W. The *graph disk set* of V (resp., W) denoted $\mathcal{GD}_V \subset \mathcal{AC}(B)$ (resp., $\mathcal{GD}_W \subset \mathcal{AC}(B)$), is the set of all vertices either corresponding to essential simple closed curves in $B \setminus n(\Gamma)$ that bound embedded disks in $V \setminus n(\Gamma)$ (resp., $W \setminus n(\Gamma)$), or corresponding to graph-bridge arcs in $B \setminus n(\Gamma)$ contained in the boundaries of graph-bridge disks in V (resp., W).

Definition 2.11. Let B be a bridge sphere for graph Γ. The *graph distance of the bridge surface*, denoted $d_{\mathcal{G}}(B, \Gamma)$ is $d_{\mathcal{AC}(B')}(\mathcal{GD}_V, \mathcal{GD}_W)$, the distance in the arc and curve complex of $B' = B \setminus n(\Gamma)$ between $\mathcal{GD}_V$ and $\mathcal{GD}_W$.

Lemma 2.12. *Let $\mathcal{L}$ be a link in bridge position with respect to a bridge sphere B, bounding 3-balls V and W, and let $\Gamma_{\mathcal{L}}$ be a graph in bridge position with respect to B formed by adding edges to $\mathcal{L}$ in V that are simultaneously parallel into B in the complement of $\mathcal{L}$, and so that $\Gamma_{\mathcal{L}} \cap V$ has at least two components.*

If $D \subset (V \setminus n(\Gamma_{\mathcal{L}}))$ is a graph-bridge disk for $\Gamma_{\mathcal{L}}$, then there is a bridge disk D' for $\mathcal{L}$ in $(V \setminus n(\mathcal{L}))$ which is disjoint from D.

Proof. Let $\Gamma_1, \ldots, \Gamma_\ell$ be the connected components $\Gamma_{\mathcal{L}} \cap V$, and let Γ_i be the component of $\Gamma_{\mathcal{L}} \cap V$ to which D is incident.

Over all bridge disks $E \subset V$ for $\mathcal{L}$ disjoint from Γ_i, choose one which minimizes $|D \cap E|$. Suppose the intersection is nonempty. Any loops of intersection can be removed because $(V \setminus n(\Gamma))$ is a handlebody and therefore irreducible. Any points of intersection between ∂D and ∂E are contained in $\partial D \cap B$ and $\partial E \cap B$. Choose an arc γ of $|D \cap E|$. The arc γ cuts D into two disks D_{γ_1} and D_{γ_2}. For one of $j = 1$ or 2, $\partial D_{\gamma_j} \cap \partial D$ is contained in B. Call that disk D_γ. Consider an arc α of $|D \cap E|$ outermost in D_γ. If the interior of D_γ is disjoint from E then take α to be γ. The arc α cuts off a disk D_α from D_γ and cuts E into two disks E_1 and E_2, only one of whose (say E_2) boundary is incident to $\mathcal{L}$. The disk $E_2 \cup D_\alpha = E'$ is a bridge disk for $\mathcal{L}$ and intersects D fewer times than E, contradicting the minimality of $|D \cap E|$. □

The above implies that the distance in the arc and curve complex of $B \setminus n(\Gamma)$ between $\mathcal{GD}_V$ and $\mathcal{BD}_V$ is less than or equal to 1.

Corollary 2.13. *Let $\mathcal{L}$ and $\Gamma_{\mathcal{L}}$ be as above. Then $d_{\mathcal{BD}}(B, \mathcal{L}) \leq 1 + d_{\mathcal{G}}(B, \Gamma_{\mathcal{L}})$.*

Proof. Since $W \setminus n(\Gamma)$ contains no graph-bridge disks, $\mathcal{GD}_W = \mathcal{BD}_W$. Suppose that the distance in $\mathcal{AC}(B')$ between $\mathcal{GD}_W = \mathcal{BD}_W$ and $\mathcal{GD}_V$ is realized by a path between vertices $X \in \mathcal{GD}_W$ and $Y \in \mathcal{GD}_V$. Then, by Lemma 2.12, there is a vertex Z of $\mathcal{BD}_V$ so that the distance between Y and Z is at most 1, and therefore $d_{\mathcal{AC}(B')}(\mathcal{BD}_W, \mathcal{BD}_V) \leq d_{\mathcal{AC}(B')}(\mathcal{GD}_W, \mathcal{GD}_V) + 1$. □

Hartshorn [2002] proved that an essential closed surface in a 3-manifold creates an upper bound on the possible distances of Heegaard splittings of that manifold in terms of the genus of the essential surface.

Theorem 2.14 [Hartshorn 2002, Theorem 1.2]. *Let M be a Haken 3-manifold containing an incompressible surface of genus g. Then any Heegaard splitting of M has distance at most $2g$.*

This idea has been generalized in numerous ways, including in [Bachman and Schleimer 2005] where it is shown that the distance of a bridge Heegaard surface in a knot complement is bounded by twice the genus plus the number of boundary components of an essential properly embedded surface.

Theorem 2.15 [Bachman and Schleimer 2005, Theorem 5.1]. *Let K be a knot in a closed, orientable 3-manifold M which is in bridge position with respect to a Heegaard surface B. Let S be a properly embedded, orientable, essential surface in $M \setminus n(K)$. Then the distance of K with respect to B is bounded above by twice the genus of S plus $|\partial S|$.*

We will need a yet more general version, since we will be concerned with surfaces properly embedded in *graph* complements.

The essence of both results is that the distance of a bridge or Heegaard surface is bounded above in terms of the *complexity* of an essential properly embedded surface. We will generalize this result to link and graph complements, with the additional benefit of avoiding many of the technical details of [Bachman and Schleimer 2005] necessary to treat the boundary components. Unfortunately, our bound will be worse than that obtained by Bachman and Schleimer, though it will be sufficient for many applications of this type of bound (see, e.g., [Mossessian 2016; Du and Qiu 2016; Ohshika and Sakuma 2016; Bachman 2013; Namazi 2007]). We note also that our proof requires a minimal starting position similar to that used by Hartshorn, an assumption Bachman and Schleimer's method was able to avoid.

We now prove Theorem 1.2.

Theorem 1.2. *Let Γ be a graph in a closed, orientable 3-manifold, M, which is in bridge position with respect to a Heegaard surface, B, so that $M \setminus n(\Gamma)$ is irreducible and boundary-irreducible. Let S be a properly embedded, orientable, incompressible, boundary-incompressible, non-boundary-parallel surface in $M \setminus n(\Gamma)$. Then $d_{\mathcal{G}}(B, \Gamma)$ is bounded above by $2(2g(S) + |\partial S| - 1)$.*

Proof of Theorem 1.2. In the case that S is closed, we note that the proofs of Theorems 2.14 and 2.15 both apply to closed surfaces in manifolds with boundary as long as the manifold is irreducible. In the case that $\partial S \neq \varnothing$ we will double $M \setminus n(\Gamma)$ along $\partial n(\Gamma)$ to obtain a closed surface and show that the surface can be made to fulfill all the hypotheses necessary to use the machinery in the proof of Theorem 2.14 to obtain the bound on distance.

First, isotope S to intersect B minimally, among all isotopy representatives of S. Let V and W be the handlebodies on either side of B. Double $M \setminus n(\Gamma)$ along $\partial n(\Gamma)$, and call the resulting manifold $\widehat{M}$. Let the doubles of S, B, V, and W be $\widehat{S}$, $\widehat{B}$, $\widehat{V}$, and $\widehat{W}$, respectively, and let G be $\partial n(\Gamma)$ in $\widehat{M}$, with respective copies M_i, S_i, B_i, V_i, and W_i, for $i = 1, 2$.

Note that $\widehat{B}$ is a Heegaard surface for $\widehat{M}$. (The proof of this is very similar to the proof of Proposition 3.2 below.) Also, note that since S is incompressible and ∂-incompressible in $M \setminus n(\Gamma)$, $\widehat{S}$ is an incompressible closed surface in $\widehat{M}$, for otherwise an outermost arc of intersection between a compressing disk and G would show S to have been ∂-compressible in $M \setminus n(\Gamma)$. Since $\partial n(\Gamma)$ was incompressible in $M \setminus n(\Gamma)$, G is incompressible in $\widehat{M}$.

Claim 1. *Each of $\widehat{S} \cap \widehat{V}$ and $\widehat{S} \cap \widehat{W}$ are incompressible.*

Proof. If, say, $\widehat{S} \cap \widehat{V}$ had a compressing disk D, then since $\widehat{S}$ is incompressible in $\widehat{M}$, there would have to be a disk D' in $\widehat{S}$ with $\partial D' = \partial D$, and $D' \cap \widehat{B} \neq \varnothing$. We may choose D to be a compressing disk which intersects G minimally. Further, since G is incompressible, we may choose D to intersect G only in arcs, if at all. But $\widehat{M}$ is irreducible, so $D \cup D'$ bounds a ball and we may isotope $\widehat{S}$ across this ball from D' to D, lowering the number of intersections between $\widehat{S}$ and $\widehat{B}$.

If $D' \cap G = \varnothing$, then this can be viewed as an isotopy of S in $M \setminus n(\Gamma)$ which reduces the number of intersections between S and B, a contradiction.

If $D' \cap G \neq \varnothing$ we still arrive at a contradiction. Consider a loop, ℓ, of intersection in $(D \cup D') \cap G$, innermost in $D \cup D'$. Since $D \cap G$ only contains arcs, ℓ consists of two arcs, α and α' in D and D' respectively. Thus ℓ bounds a disk D_ℓ in G, α cuts off a subdisk D_α of D and α' cuts off a subdisk $D_{\alpha'}$ of D', both of which are in either M_1 or M_2, say M_1. Now we have an isotopy of S_1 from $D_\alpha \cup D_{\alpha'}$ to D_ℓ.

Independent of whether $D_{\alpha'}$ intersected B, we could have chosen D to have fewer intersections with G, contradicting our choice of D to minimize intersections. □

Claim 2. *Every intersection of $\widehat{S}$ with $\widehat{B}$ is essential in $\widehat{B}$.*

Proof. Curves of intersection in $\widehat{S} \cap \widehat{B}$ which are inessential in both surfaces would either give rise to a reduction in $|S \cap B|$ or could have come from the doubling of arcs in $S \cap B$ which would give rise to a reduction in $|S \cap B|$ in a fashion similar to the previous claim. □

Claim 3. *There are no ∂-parallel annular components of $\widehat{S} \cap \widehat{W}$ or $\widehat{S} \cap \widehat{V}$.*

Proof. Any such component disjoint from G would have been eliminated when $|S \cap B|$ was minimized. The intersection of any such component intersecting G with M_1 would be a ∂-parallel disk which also would have been eliminated when $|S \cap B|$ was minimized. □

Now we have satisfied all the hypotheses to obtain the sequence of isotopic copies of $\widehat{S}$ described in Lemmas 4.4 and 4.5 of [Hartshorn 2002]. Depending on whether either of $\widehat{S} \cap \widehat{V}$ or $\widehat{S} \cap \widehat{W}$ contains disk components or not, we apply either Lemma 4.4 or 4.5, respectively, of [Hartshorn 2002] to obtain a sequence of boundary compressions of $\widehat{S}$ in $\widehat{V}$ or $\widehat{W}$, which gives rise to a path in $\mathcal{C}(\widehat{S})$. A priori, this path would not restrict to a path in $\mathcal{AC}(S)$, but the following claim shows that we can choose the compressions to be symmetric across G, and so each compression will correspond to an edge in $\mathcal{AC}(S)$.

Claim 4. *If there exists an elementary ∂-compression of $\widehat{S}$ in $\widehat{V}$ (resp., $\widehat{W}$), then there exists an elementary compression of $\widehat{S}$ in $\widehat{V}$ (resp., $\widehat{W}$) which is symmetric across G in the sense that either*

(1) *the ∂-compressing disk D_1 is disjoint from G in M_1, and there is a corresponding ∂-compressing disk D_2 in M_2, or*

(2) *the ∂-compression is along a disk that is symmetric across G.*

Proof. Let D be an elementary ∂-compression disk for, say, $\widehat{S} \cap \widehat{V}$ chosen to minimize $|D \cap G|$. We may restrict attention to such disks with $|D \cap G| > 0$.

First, we observe that $D \cap G$ cannot contain any loops of intersection, for a loop of $D \cap G$ innermost in D bounds a subdisk of D which would either give rise to a compression for G or would provide a means of isotoping D so as to lower $|D \cap G|$. Thus, $D \cap G$ consists only of arcs. These arcs are either

- vertical arcs, with one endpoint on each of $\widehat{S}$ and $\widehat{B}$,
- $\widehat{S}$-arcs, with both endpoints on $\widehat{S}$, or
- $\widehat{B}$-arcs, with both endpoints on $\widehat{B}$.

Consider an $\widehat{S}$-arc of $D \cap G$, outermost in D, cutting off subdisk D' from D, with boundary consisting of σ in $\widehat{S}$ and γ in G. Without loss of generality, assume $D' \subset M_1$. If σ is essential in $\widehat{S} \cap M_1$, then D' is a boundary-compression disk for S in M, which is impossible. If σ is inessential in $\widehat{S} \cap M_1$, then it must cobound a disk E in $\widehat{S} \cap M_1$ together with an arc $\sigma' \subseteq \partial(\widehat{S} \cap M_1)$. The curve $\gamma \cup \sigma'$ cannot be essential in G, else $D' \cup E$ would be a compressing disk for G. Thus, $\gamma \cup \sigma'$ bounds a disk, $F \subseteq G$. Now $F \cup D' \cup E$ is a sphere bounding a ball in M_1, so $D \cup E$ is isotopic to F, and replacing D' with F results in an elementary boundary-compressing disk for $\widehat{S} \cap V$ with fewer intersections with G than D. Thus we may assume that $D \cap G$ contains no $\widehat{S}$-arcs.

Now consider a subdisk D' of D which is cut off by all the arcs of $D \cap G$ and whose boundary consists of no more than one vertical arc. Without loss of generality, assume $D' \subseteq M_1$. Suppose $\partial D'$ has $\widehat{B}$-arcs, $\beta_1, \beta_2, \ldots, \beta_k$. Then all the β_i are disjoint arcs on G. If any of them are inessential in $G \cap \widehat{V}$ then they bound disks $B_i \subseteq G \cap V_1$. If any of the β_i are essential in $G \cap \widehat{V}$, then they bound disks $B_i \subseteq V_1$ that are bridge disks for $n(\Gamma)$ in V_1. In either case, $D' \cup \left(\bigcup_{i=1}^{k} B_i\right)$ results in a boundary-compressing disk for $S \cap \widehat{V}$ with fewer intersections with G than D. This boundary-compressing disk is still elementary as the arc in $\widehat{S}$ remains unchanged. Thus, we may assume that $D \cap G$ consists solely of vertical arcs.

Let γ be an arc of $D \cap G$ outermost in D, cutting off a subdisk D_1 from D. Without loss of generality, $D_1 \subseteq M_1$. The boundary of D_1 consists of three arcs; $\gamma \subseteq G$, $\sigma_1 \subseteq S_1$ and $\beta_1 \subseteq B_1$. By symmetry, there exists disk $D_2 \subseteq M_2$ in M_2, so that $D_1 \cup D_2$ is a disk in $\widehat{V}$ with boundary consisting of arcs $\sigma = \sigma_1 \cup \sigma_2 \subseteq \widehat{S}$ and $\beta = \beta_1 \cup \beta_2 \subseteq \widehat{B}$, intersecting G in exactly one arc, γ. Finally, we must show that σ is a "strongly essential" arc in $\widehat{S} \cap \widehat{V}$.

If σ is not strongly essential then it is either the meridian of a boundary-parallel annulus of $\widehat{S} \cap \widehat{V}$, which is not possible since σ_1 was a subarc of the original elementary compression disk D, or σ is inessential in $\widehat{S} \cap \widehat{V}$. If σ is inessential then it would cobound a disk E in $\widehat{S}$ together with an arc $\sigma' \subseteq \widehat{S} \cap \widehat{B}$. This disk provides an isotopy in $\widehat{S}$ of σ_1 to σ_2.

If the disk $D' = D \setminus D_1$ only intersects D_2 in γ then $D' \cup D_2$ is a compressing disk for $\widehat{S} \cap \widehat{V}$ with fewer arcs of intersection with G, as the disk can be isotoped away from γ. This disk is still an elementary compressing disk because σ_1 is isotopic to σ_2, and so contradicts our original choice of D.

Thus, σ is strongly essential in $\widehat{S} \cap \widehat{V}$, and $D_1 \cup D_2$ is a new compressing disk for $\widehat{S} \cap \widehat{V}$ that is symmetric across G. □

We may, thus, proceed exactly as in Theorem 2.14. Each elementary boundary compression of $\widehat{S}$ towards either of $\widehat{V}$ or $\widehat{W}$ can be performed in a symmetric way, demonstrating a path from $\mathcal{D}_{\widehat{V}}$ to $\mathcal{D}_{\widehat{W}}$ in $\mathcal{C}(\widehat{S})$ of length no greater than twice the genus of $\widehat{S}$, which is $2(g(S) + |\partial S| - 1)$.

Each time a boundary compression for $\widehat{S}$ corresponds to a pair of curves $\hat{c}_i$ and $\hat{c}_{i+1}$ in S_1 that contribute an edge in a path in $\mathcal{C}(\widehat{S})$ from $\mathcal{D}_{\widehat{V}}$ to $\mathcal{D}_{\widehat{W}}$, there is immediately a pair of curves $\hat{c}_{i+2}$ and $\hat{c}_{i+3}$ in S_2 also contributing an edge in a path from $\mathcal{D}_V$ to $\mathcal{D}_W$, and this pair of paths corresponds to a single pair of curves c_i and c_{i+1} in S contributing a single edge in $\mathcal{AC}(S)$. Each time a boundary compression for $\widehat{S}$ corresponds to a pair of curves intersecting G that contributes an edge in a path in $\mathcal{C}(\widehat{S})$ from $\mathcal{D}_{\widehat{V}}$ to $\mathcal{D}_{\widehat{W}}$, the restriction of these curves to S_1 is a pair of arcs contributing an edge in $\mathcal{AC}(S)$.

Further, since the boundary compressions (and elimination of boundary-parallel annuli) are all being performed symmetrically, the resulting disks $D_{\widehat{V}} \in \mathcal{D}_{\widehat{V}}$ from $\widehat{S} \cap \widehat{V}$ and $D_{\widehat{W}} \in \mathcal{D}_{\widehat{W}}$ from $\widehat{S} \cap \widehat{W}$ are symmetric. That is, either $D_{\widehat{V}}$ (resp., $D_{\widehat{W}}$) is disjoint from G, so that we may assume that it sits in V_1 (resp., W_1), or it is symmetric across G so that $D_{\widehat{V}} \cap M_1$ (resp., $D_{\widehat{W}} \cap M_1$) is a graph bridge disk for Γ in M. In either case, this demonstrates a path in $\mathcal{AC}(S)$ from $\mathcal{DG}_V$ to $\mathcal{DG}_W$ of length no greater than $2(g(S) + |\partial S| - 1)$. □

3. Theorem 1.1

Bachman [2010] defined the topological index of a surface. In contrast to the distances between subcomplexes each corresponding to some disks discussed in Section 2, he exploits the homotopy type of the complex of all disks.

Definition 3.1. The surface B is said to be *topologically minimal* if either $\mathcal{D}_B$ is empty, or if there exists an $n \in \mathbb{N}$ so that $\pi_n(\mathcal{D}_B) \neq 0$. If a surface B is topologically minimal, then the *topological index* is defined to be the smallest $n \in \mathbb{N}$ so that $\pi_{n-1}(\mathcal{D}_B) \neq 0$, or 0 if $\mathcal{D}_B$ is empty.

Bachman and Johnson [2010] showed that surfaces of arbitrarily high index exist, but their manifolds all contain essential tori. We prove an analogue of this.

Theorem 1.1. *There is a closed* 3*-manifold M^1, with an index* 1 *Heegaard surface S, such that for each n, the lift of S to some n-fold cover M^n of M^1 has topological index n. Moreover, M^n is hyperbolic for all n.*

3A. ***The construction.*** Let n be a positive integer. We will construct a hyperbolic manifold containing a Heegaard surface of topological index n.

Using the machinery in Theorem 2.6, let $\mathcal{L}$ be a link in S^3 with two components, L and K, that are each 2-bridge with respect to a bridge sphere B of distance at least $24n+7$. Let V and W be the two 3-balls bounded by B. Since $\mathcal{L}$ is in bridge position, there exist disks D_V and D_W properly embedded in V and W, respectively, with $(\mathcal{L} \cap V) \subset D_V$, and $(\mathcal{L} \cap W) \subset D_W$. By modifying D_V if necessary, we can find two arcs τ_L and τ_K in the interior of V such that

(1) $\tau_L \cup \tau_K \subset D_V$,

(2) $\tau_L \cap \tau_K = \varnothing$,

(3) $\tau_L \cap \mathcal{L} = \partial\tau_L \subset L$ and $\tau_K \cap \mathcal{L} = \partial\tau_K \subset K$,

(4) the endpoints of τ_K are on different components of $K \cap V$, and the endpoints of τ_L are on different components of $L \cap V$.

Let $L' = L \cup \tau_L$, let $G_L = \partial n(L')$, let $K' = K \cup \tau_K$, let $G_K = \partial n(K')$, and let $\Gamma = \mathcal{L} \cup \tau_L \cup \tau_K = L' \cup K'$. Observe that Γ is a graph in bridge position with respect to B. Let $M' = \overline{S^3 \setminus n(\Gamma)}$, let $V' = \overline{V \setminus n(\Gamma)}$, and let $W' = \overline{W \setminus n(\Gamma)} = \overline{W \setminus n(\mathcal{L})}$, and $B' = B \setminus n(\Gamma) = B \setminus n(\mathcal{L})$.

For each $i = 1, 2, \dots, n$, let M'_i be homeomorphic to M', along with homeomorphic copies $\mathcal{L}_i$ of $\mathcal{L}$, $(G_L)_i$ of G_L, $(G_K)_i$ of G_K, and B'_i of B'.

Then, for each $i = 1, 2, \dots, (n-1)$, identify $(G_K)_i$ with $(G_L)_{i+1}$ and identify $(G_K)_n$ with $(G_L)_1$, all via the same homeomorphism. Call the resulting closed 3-manifold M^n. Observe that the union of the B'_i is a closed surface that we will call B^n. We will show that B^n is a Heegaard surface for M^n, that B^n has high topological index, and that M^n is hyperbolic.

Proposition 3.2. *For each n, $B^n \subset M^n$ is a genus $3n+1$ Heegaard surface.*

Proof. That the genus of B^n is $3n+1$ can be verified by an Euler characteristic count. It suffices, then, to verify that the complement of B^n is two handlebodies, V^n and W^n.

Since Γ was in bridge position with respect to B, there are disks D_V and D_W properly embedded in V and W, respectively, so that $\Gamma \cap V \subset D_V$ and $\Gamma \cap W \subset D_W$. Then D_V and D_W cut along Γ is a collection of subdisks.

The result of cutting $V \setminus n(\Gamma)$ along all these subdisks of D_V is a pair of 3-balls, each with two subdisks, D_1^+ and D_2^+, of $n(\Gamma)$ contained in the boundary. Each identification of $(G_K)_i$ with $(G_L)_{i+1}$ (indices mod n) glues pairs of these subdisks along arcs, resulting in disks in V^n, and further cutting along $(n-1)$ copies of each of D_1^+ and D_2^+ results in a collection of 3-balls, showing that V^n is a handlebody.

Similarly, the result of cutting $W \setminus n(\Gamma)$ along all of the subdisks of D_W is a pair of 3-balls, each with four subdisks of $n(\Gamma)$ contained in the boundary, D_1^-, D_2^-, D_3^-, and D_4^-. Each identification of $(G_K)_i$ with $(G_L)_{i+1}$ (indices mod n) glues pairs of these subdisks along arcs, resulting in disks in W^n, and further cutting along $(n-1)$ copies of each of D_1^-, D_2^-, D_3^-, and D_4^- results in a collection of 3-balls, showing that W^n is a handlebody. □

3B. ***Bounding from above.***

Proposition 3.3. *The surface B^n has topological index at most n.*

Proof. Our proof will follow almost exactly the proof of Proposition 5 from [Bachman and Johnson 2010]. In each copy M_i' of the manifold M', we have the surface B_i', a copy of B', dividing the manifold into V_i' and W_i', copies of V' and W'. Observe that in each V_i', there is exactly one essential disk, D_i^+ with boundary contained in B_i', just as in [Bachman and Johnson 2010]. However, in each W_i', there are several essential disks with boundary contained in B_i'. We will call this collection of disks $\mathscr{D}_i^-$. From each $\mathscr{D}_i^-$, choose a single representative D_i^-.

Define the subcomplex, P, of $\mathcal{D}_M$ spanned by the vertices corresponding to $\bigcup_i \{D_i^+, D_i^-\}$, which is homeomorphic to an $(n-1)$-sphere. Then, define a map $F : \mathcal{D}_M \to P$ by the identity on P, and by sending a vertex corresponding to a disk $D \notin \bigcup_i \{D_i^+, D_i^-\}$ to the vertex corresponding to D_j^+ or D_j^-, where either $D \in \mathscr{D}_j^-$, or j is the smallest index for which an essential outermost subdisk of $D \setminus \left(\bigcup_i G_i\right)$ is contained in V_j' or W_j', respectively.

Just as in [Bachman and Johnson 2010], we claim that this map F is a simplicial map that fixes each vertex of P. To see this, consider any two disks D_1 and D_2 connected by an edge in $\mathcal{D}_M$ (so that the disks are realized disjointly in M). Observe that by our construction of M' and Corollary 2.13, any disk contained in V_j' must intersect any disk contained in W_j' (whether either disk is a bridge disk, a graph-bridge disk, or the boundary is contained in B_j'). So, if $D_i^{\pm} = F(D_1) \neq F(D_2) = D_j^{\pm}$, then $i \neq j$, and $F(D_1)$ is joined to $F(D_2)$ in P. Thus, F is a retraction onto the $(n-1)$-sphere, P, showing that $\pi_{n-1}(\mathcal{D}_M)$ is nontrivial, so the topological index of B^n is at most n. □

Corollary 3.4. *The topological index of B^n is well defined, and B^n is topologically minimal.*

3C. ***Bounding from below.*** We make use of an important theorem in the development of the topological index by Bachman:

Theorem 3.5 [Bachman 2010, Theorem 3.7]. *Let G be a properly embedded, incompressible surface in an irreducible* 3*-manifold M. Let B be a properly embedded surface in M with topological index n. Then B may be isotoped so that*

(1) *B meets G in p saddles, for some $p \leq n$, and*

(2) *the sum of the topological indices of the components of $B \setminus n(G)$, plus p, is at most n.*

Proposition 3.6. *The surface B^n has topological index no smaller than n.*

Proof. Suppose B^n had topological index $\iota < n$. Let G be the union of all the genus two surfaces $G_i^n := (G_K)_i = (G_L)_{i+1}$ (indices mod n) in the manifold M^n. By Theorem 3.5, B^n can be isotoped to a surface, B_+^n, so that B_+^n meets G in σ saddles, the sum of the topological indices of the components of $B_+^n \setminus n(G)$ is k, and $k+\sigma \leq \iota$. Observe that $\chi(B_+^n \setminus n(G)) = -6n + \sigma$. We may isotope any annular components of $B_+^n \setminus n(G)$ that are boundary-parallel into $\partial n(G)$ completely into $n(G)$. Note that this will have no effect on the Euler characteristic of $B_+^n \setminus n(G)$, nor any effect on the topological index, since such a component will have topological index 0.

Any component, Q, of $B_+^n \cap n(G)$ is contained in $n(G_i^n)$ for some i. Any such Q is a punctured sphere with, say, d boundary components, has $d-2$ saddles, and we will show that at most $d-2$ of its boundary components can bound disks of $B_+^n \setminus n(G)$ that are boundary-parallel into $\partial n(G)$ in $M_i \setminus n(G)$ or $M_{i+1} \setminus n(G)$.

As B_+^n is connected and not a sphere, all the boundary curves of Q cannot bound disks. Suppose, then, that $d-1$ of the curves bound disks that are boundary-parallel into $\partial n(G)$ in $M_i \setminus n(G)$ or $M_{i+1} \setminus n(G)$, and let c be the remaining boundary component of Q. As the other curves all bound disks that can be isotoped into $n(G_i^n)$, and G_i^n is incompressible in M^n, c must bound a disk in $\partial n(G_i^n)$. By pushing this disk slightly into M_i or M_{i+1}, we have a compressing disk for a component of $B_+^n \setminus n(G)$ that is disjoint from all other compressing disks for that component. Thus, the disk complex for that component is contractible, contrary to the fact that it is topologically minimal. Thus, at most $d-2$ of the boundary components of Q can bound disks that are boundary-parallel into $\partial n(G)$ in $M_i \setminus n(G)$ or $M_{i+1} \setminus n(G)$.

Therefore, the total number of disk components of $B_+^n \setminus n(G)$ that are boundary-parallel in $M^n \setminus n(G)$ is $\beta \leq \sigma$. So we may further isotope all β such boundary-parallel disks into $n(G)$, and call the resulting surface B_0^n. Still, then, each component of $B_0^n \setminus n(G)$ is topologically minimal, the topological index will be unchanged as each boundary-parallel disk has topological index 0, $B_0^n \setminus n(G)$ has no boundary-parallel disks or annuli, and

$$\chi(B_0^n \setminus n(G)) = \chi(B_+^n \setminus n(G)) - \beta \geq \chi(B_+^n \setminus n(G)) - \sigma = -6n.$$

First, suppose that there is some component of $B_0^n \setminus n(G)$ with Euler characteristic less than $-6n$. In this case, because the Euler characteristic of $B_0^n \setminus n(G)$ is greater than or equal to $-6n$, there must be a component of $B_0^n \setminus n(G)$ with positive Euler characteristic. But there are no disks, as we have eliminated boundary-parallel disks and an essential disk would be a compression of G in M^n, and it cannot be a sphere, so this is impossible.

Thus, we may suppose that the Euler characteristic of each component of $B_0^n \setminus n(G)$ is bounded below by $-6n$. Observe that each component of G is an incompressible surface, so B^n cannot be made disjoint from any component of G, and so $(B_0^n \setminus n(G)) \cap M_i$ is nonempty for all i. As the sum of the topological indices of the components of $B_0^n \setminus n(G)$ is $k < n$, there must be at least one index j so that every component of $(B_0^n \setminus n(G)) \cap M_j$ has topological index 0. Thus, there is some component of $(B_0^n \setminus n(G)) \cap M_j$, and all such components are incompressible and have Euler characteristic bounded below by $-6n$. If necessary, maximally boundary compress $(B_0^n \setminus n(G)) \cap M_j$, and isotope any resulting boundary-parallel components into $n(G)$. As B_0^n cannot be isotoped away from any copy of G_i^n, there must be some component remaining that is incompressible, boundary-incompressible, and not boundary-parallel. Since boundary compressions only increase Euler characteristic, the resulting component has Euler characteristic bounded below by $-6n$. Call this component B''.

By Lemma 2.9 and Corollary 2.13, in M_j with B_j a copy of B', we have

$$d_{\mathcal{C}}(B_j, \mathcal{L}) \leq 2d_{\mathcal{BD}}(B_j, \mathcal{L}) \leq 2(1 + d_{\mathcal{G}}(B_j, \Gamma)).$$

By Theorem 1.2, $d_{\mathcal{G}}(B_j, \Gamma) \leq 2(2g(B'') + |\partial B''| - 1)$. By our choice of $\mathcal{L}$ and the fact that $\chi(S) = 2 - 2g(S) - |\partial S|$, we have

$$24n + 7 \leq d_{\mathcal{C}}(B_j, \mathcal{L}) \leq 2 + 2d_{\mathcal{G}}(B_j, \Gamma) \leq 8g(B'') + 4|\partial B''| - 2 = -4\chi(B'') + 6.$$

On the other hand we have just shown that $-6n \leq \chi(B'')$, a contradiction. Thus, the topological index of B^n cannot be less than n. □

3D. *Hyperbolicity.* We have now shown that M^n contains a surface of topological index n. To prove Theorem 1.1 it remains to show that M^n is hyperbolic.

Proposition 3.7. *For all n, M^n is hyperbolic.*

Proof. Consider an essential surface S in M^n with Euler characteristic bounded below by 0, chosen to intersect G minimally. If $S \cap G = \varnothing$, we arrive at a contradiction to Theorem 1.2 as S would lie in one of the copies of M'. If $S \cap G \neq \varnothing$, the incompressibility and boundary-incompressibility of G guarantees that the curves of $S \cap G$ are essential in S. Thus $S \cap M_i'$ is a collection of one or more planar surfaces for some i. This again contradicts Theorem 1.2. Thus, in particular, M^n is prime and atoroidal for all n. Then, as G is an incompressible surface in M^n, we conclude that M^n is hyperbolic. □

Now the proof of Theorem 1.1 follows.

Proof of Theorem 1.1. Let M^n and B^n be as in Section 3A. We note that M^n is an n-fold cover of M^1. By Proposition 3.2, B^n is a genus $3n + 1$ Heegaard surface. By Propositions 3.3 and 3.6, B^n has topological index n, and by Proposition 3.7, M^n is hyperbolic. □

References

[Bachman 2010] D. Bachman, "Topological index theory for surfaces in 3-manifolds", *Geom. Topol.* **14**:1 (2010), 585–609. MR Zbl

[Bachman 2012a] D. Bachman, "Normalizing topologically minimal surfaces I: Global to local index", preprint, 2012. arXiv

[Bachman 2012b] D. Bachman, "Normalizing topologically minimal surfaces II: Disks", preprint, 2012. arXiv

[Bachman 2012c] D. Bachman, "Normalizing topologically minimal surfaces III: Bounded combinatorics", preprint, 2012. arXiv

[Bachman 2013] D. Bachman, "Stabilizing and destabilizing Heegaard splittings of sufficiently complicated 3-manifolds", *Math. Ann.* **355**:2 (2013), 697–728. MR Zbl

[Bachman and Johnson 2010] D. Bachman and J. Johnson, "On the existence of high index topologically minimal surfaces", *Math. Res. Lett.* **17**:3 (2010), 389–394. MR Zbl

[Bachman and Schleimer 2005] D. Bachman and S. Schleimer, "Distance and bridge position", *Pacific J. Math.* **219**:2 (2005), 221–235. MR Zbl

[Blair et al. 2013] R. Blair, M. Tomova, and M. Yoshizawa, "High distance bridge surfaces", *Algebr. Geom. Topol.* **13**:5 (2013), 2925–2946. MR Zbl

[Blair et al. 2017] R. Blair, M. Campisi, J. Johnson, S. A. Taylor, and M. Tomova, "Exceptional and cosmetic surgeries on knots", *Math. Ann.* **367**:1-2 (2017), 581–622. MR Zbl

[Colding and Minicozzi 2004a] T. H. Colding and W. P. Minicozzi, II, "The space of embedded minimal surfaces of fixed genus in a 3-manifold, I: Estimates off the axis for disks", *Ann. of Math.* (2) **160**:1 (2004), 27–68. MR Zbl

[Colding and Minicozzi 2004b] T. H. Colding and W. P. Minicozzi, II, "The space of embedded minimal surfaces of fixed genus in a 3-manifold, II: Multi-valued graphs in disks", *Ann. of Math.* (2) **160**:1 (2004), 69–92. MR Zbl

[Colding and Minicozzi 2004c] T. H. Colding and W. P. Minicozzi, II, "The space of embedded minimal surfaces of fixed genus in a 3-manifold, III: Planar domains", *Ann. of Math.* (2) **160**:2 (2004), 523–572. MR Zbl

[Colding and Minicozzi 2004d] T. H. Colding and W. P. Minicozzi, II, "The space of embedded minimal surfaces of fixed genus in a 3-manifold, IV: Locally simply connected", *Ann. of Math.* (2) **160**:2 (2004), 573–615. MR Zbl

[Colding and Minicozzi 2015] T. H. Colding and W. P. Minicozzi, II, "The space of embedded minimal surfaces of fixed genus in a 3-manifold V; fixed genus", *Ann. of Math.* (2) **181**:1 (2015), 1–153. MR Zbl

[Du and Qiu 2016] K. Du and R. Qiu, "A note on the uniqueness of unstabilized Heegaard splittings of amalgamated 3-manifolds", *Topology Appl.* **204** (2016), 135–148. MR Zbl

[Goda 1997] H. Goda, "Bridge index for theta curves in the 3-sphere", *Topology Appl.* **79**:3 (1997), 177–196. MR Zbl

[Hartshorn 2002] K. Hartshorn, "Heegaard splittings of Haken manifolds have bounded distance", *Pacific J. Math.* **204**:1 (2002), 61–75. MR Zbl

[Kneser 1929] H. Kneser, "Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten", *Jahresbericht der Deutschen Mathematiker-Vereinigung* **38** (1929), 248–259. Zbl

[Lee 2015] J. H. Lee, "On topologically minimal surfaces of high genus", *Proc. Amer. Math. Soc.* **143**:6 (2015), 2725–2730. MR Zbl

[Mossessian 2016] G. Mossessian, "Stabilizing Heegaard splittings of high-distance knots", *Algebr. Geom. Topol.* **16**:6 (2016), 3419–3443. MR Zbl

[Namazi 2007] H. Namazi, "Big Heegaard distance implies finite mapping class group", *Topology Appl.* **154**:16 (2007), 2939–2949. MR Zbl

[Ohshika and Sakuma 2016] K. Ohshika and M. Sakuma, "Subgroups of mapping class groups related to Heegaard splittings and bridge decompositions", *Geom. Dedicata* **180** (2016), 117–134. MR Zbl

[Ozawa 2012] M. Ozawa, "Bridge position and the representativity of spatial graphs", *Topology Appl.* **159**:4 (2012), 936–947. MR Zbl

[Rubinstein 1995] J. H. Rubinstein, "An algorithm to recognize the 3-sphere", pp. 601–611 in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), vol. 1, edited by S. D. Chatterji, Birkhäuser, Basel, 1995. MR Zbl

Received June 16, 2017. Revised February 22, 2018.

MARION CAMPISI
MATHEMATICS AND STATISTICS DEPARTMENT
SAN JOSÉ STATE UNIVERSITY
SAN JOSÉ, CA
UNITED STATES
marion.campisi@sjsu.edu

MATT RATHBUN
DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY, FULLERTON
FULLERTON, CA
UNITED STATES
mrathbun@fullerton.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.321

LENGTH SPECTRA OF SUB-RIEMANNIAN METRICS ON COMPACT LIE GROUPS

ANDRÁS DOMOKOS, MATTHEW KRAUEL, VINCENT PIGNO,
COREY SHANBROM AND MICHAEL VANVALKENBURGH

Length spectra for Riemannian metrics have been well studied, while sub-Riemannian length spectra remain largely unexplored. Here we give the length spectrum for a canonical sub-Riemannian structure attached to any compact Lie group by restricting its Killing form to the sum of the root spaces. Surprisingly, the shortest loops are the same in both the Riemannian and sub-Riemannian cases. We provide specific calculations for SU(2) and SU(3).

1. Introduction

While much is known about the existence and geometric properties of closed geodesics on Riemannian manifolds in general [Klingenberg 1978], and Lie groups in particular, we cannot say the same thing about their connection with the algebraic structure of Lie groups. Moreover, the sub-Riemannian setting has been mostly neglected.

In the case of simple, simply connected, compact Lie groups, Helgason [2001, Proposition 11.9] obtained the length of the shortest Riemannian geodesic loop in terms of the length of the highest root. We expand upon Helgason's work using more algebraic methods, obtaining the sub-Riemannian and Riemannian geodesic loop length spectra. The sub-Riemannian structure consists of the horizontal distribution defined by the orthogonal complement of a Cartan subalgebra and the restriction of the bi-invariant metric defined by the Killing form. To our knowledge, nothing was previously known about the connection between root systems and lengths of sub-Riemannian geodesic loops.

In Section 2 we provide the background for the root space decomposition of semisimple, compact Lie algebras and prove Theorem 2.7, which shows that all sub-Riemannian geodesics are normal. In Section 3 we work in a simple, simply connected, compact Lie group. We find connections between the algebraic information encoded in the root system of the Lie algebra and properties of Riemannian and sub-Riemannian geodesic loops. In Theorems 3.3 and 3.6 we describe the entire

MSC2010: primary 53C17, 53C22; secondary 22E30, 51N30.
Keywords: sub-Riemannian geometry, geodesics, root systems, compact Lie groups.

length spectra of the Riemannian and certain sub-Riemannian geodesic loops. In Theorem 3.9 we find properties that help describe the remaining sub-Riemannian geodesic loops. In Theorem 3.7, we compute the lengths of the shortest Riemannian and sub-Riemannian loops, which unexpectedly turn out to be equal. Further, in Corollary 3.8 we derive a purely algebraic formula for the length of the highest root. In Sections 4 and 5 we provide relevant examples in SU(2) and SU(3).

Note that the terms *length spectrum* and *geodesic* have varying definitions in the literature. By length spectrum, we mean the set of lengths of all primitive geodesic loops. A sub-Riemannian geodesic is defined as in [Montgomery 2002] as a locally length minimizing curve. While in general such curves may not satisfy the geodesic equations, in our setting we show that the two notions coincide (see Theorem 2.7).

2. General results

In this section, we assume that $\mathbb{G}$ is a semisimple, connected, compact matrix Lie group. This assumption is suited to present and prove some general results about sub-Riemannian geodesics, and we will use the more restrictive simple and simply connected assumptions in the following sections, where we prove results about sub-Riemannian geodesic loops. Our notation and definitions will be geared toward the presentation of the sub-Riemannian geometry, rather than the algebraic theory of Lie groups.

The Lie algebra of $\mathbb{G}$ can be defined in terms of the matrix exponential:

$$\mathcal{G} = \{X \in \mathcal{M}_n \,:\, e^{tX} \in \mathbb{G} \text{ for all } t \in \mathbb{R}\},$$

where $\mathcal{M}_n$ is the linear space of $n \times n$ real or complex matrices in which $\mathbb{G}$ is included. Then $\mathcal{G}$ is a real Lie algebra endowed with the commutator operator

$$[X, Y] = XY - YX.$$

A Lie algebra is called simple if it is noncommutative and does not have any nontrivial ideals, and it is called semisimple if it is the direct sum of simple Lie algebras. A Lie group is simple or semisimple if its Lie algebra has the corresponding property.

The adjoint representation of $\mathbb{G}$ is the group homomorphism

$$\mathrm{Ad} : \mathbb{G} \to \mathrm{Aut}(\mathcal{G}), \quad \mathrm{Ad}(g)(X) = gXg^{-1},$$

while its differential at the identity is the adjoint representation of its Lie algebra

$$\mathrm{ad} : \mathcal{G} \to \mathrm{End}(\mathcal{G}), \quad \mathrm{ad}\, X(Y) = [X, Y].$$

Note that, among semisimple Lie algebras, Ad is an irreducible representation of $\mathbb{G}$ if and only if $\mathbb{G}$ is simple.

The Killing form,

$$K(X, Y) = \mathrm{trace}(\mathrm{ad}\, X \cdot \mathrm{ad}\, Y),$$

is negative definite and nondegenerate on the Lie algebra of a semisimple, compact Lie group, and hence we can define an inner product on $\mathcal{G}$ as

$$\langle X, Y\rangle = -\rho\, K(X, Y), \tag{2-1}$$

where $\rho > 0$ is a constant which can be adjusted according to our normalization preferences. The inner product (2-1) generates a bi-invariant metric on $\mathbb{G}$. The Killing form is Ad-invariant, so $\mathrm{Ad}(g)$ is a unitary linear transformation of $\mathcal{G}$ for all $g \in \mathbb{G}$ and $\mathrm{ad}\, X$ is skew-symmetric for all $X \in \mathcal{G}$.

Let $\mathbb{T}$ be a maximal torus in $\mathbb{G}$ and $\mathcal{T}$ be its Lie algebra. In this case, $\mathcal{T}$ is a maximal commutative subalgebra of $\mathcal{G}$ called the Cartan subalgebra. Its dimension is called the rank of $\mathcal{G}$, and also the rank of $\mathbb{G}$. Consider an orthonormal basis $\mathcal{B}_\mathcal{T} = \{T_1, \dots, T_r\}$ of $\mathcal{T}$, which will be fixed throughout the paper.

We extend the inner product (2-1) on $\mathcal{G}$ bilinearly to the complexified Lie algebra $\mathcal{G}_\mathbb{C} = \mathcal{G} \oplus i\mathcal{G}$. The mappings $\mathrm{ad}\, T : \mathcal{G}_\mathbb{C} \to \mathcal{G}_\mathbb{C}$, $T \in \mathcal{T}$, commute and are skew-symmetric, so they share eigenspaces and have purely imaginary eigenvalues.

Once we fix the orthonormal basis $\mathcal{B}_\mathcal{T}$, we can identify $\mathcal{T}^*$ with $\mathcal{T}$ and define the roots as elements of the Cartan subalgebra, as in [Domokos 2015].

Definition 2.1. We define $R \in \mathcal{T}$ to be a root if $R \neq 0$ and the root space $\mathcal{G}_R \neq \{0\}$, where

$$\mathcal{G}_R = \{Z \in \mathcal{G}_\mathbb{C} : \mathrm{ad}\, T(Z) = i\, \langle R, T\rangle\, Z \ \text{ for all } \ T \in \mathcal{T}\}.$$

Additionally, we use the notation $\mathcal{G}_0 = \mathcal{T}_\mathbb{C} = \mathcal{T} \oplus i\mathcal{T}$.

Let $\mathcal{R}$ be the set of all roots, which will be partially ordered by the relation $R_1 > R_2$ if the first nonzero coordinate of $R_1 - R_2$ relative to the ordered basis $\mathcal{B}_\mathcal{T}$ is positive. We call a root positive if its first nonzero coordinate is positive and let $\mathcal{R}^+$ denote the set of all positive roots. For the most important properties of $\mathcal{G}_R$ we quote [Duistermaat and Kolk 2000; Knapp 1986]:

(i) $\dim_\mathbb{C} \mathcal{G}_R = 1$.

(ii) If $R \in \mathcal{R}$ then $-R \in \mathcal{R}$.

(iii) $\mathcal{G}_{-R} = \{X - iY : X + iY \in \mathcal{G}_R\}$.

(iv) $\langle \mathcal{G}_{R_1}, \mathcal{G}_{R_2}\rangle = 0$ if $R_1, R_2 \in \mathcal{R} \cup \{0\}$, $R_1 \neq \pm R_2$.

(v) $[\mathcal{G}_{R_1}, \mathcal{G}_{R_2}] \begin{cases} = \mathcal{G}_{R_1+R_2} & \text{if } R_1 + R_2 \in \mathcal{R} \\ = \{0\} & \text{if } R_1 + R_2 \notin \mathcal{R} \text{ and } R_1 + R_2 \neq 0 \\ \subset i\mathcal{T} & \text{if } \ R_1 + R_2 = 0. \end{cases}$

(vi) If $Z_R \in \mathcal{G}_R$ and $Z_{-R} \in \mathcal{G}_{-R}$ then $[Z_R, Z_{-R}] = i\langle Z_R, Z_{-R}\rangle R$.

The above properties of $\mathcal{G}_R$ and the real root space decomposition

$$\mathcal{G} = \mathcal{T} \oplus \mathcal{H},$$

where

$$\mathcal{H} = \mathcal{T}^{\perp} = \bigoplus_{R \in \mathcal{R}^+} (\mathcal{G}_R \oplus \mathcal{G}_{-R}) \cap \mathcal{G}, \tag{2-2}$$

allow us to choose an orthonormal basis of $\mathcal{H}$,

$$\mathcal{B}_{\mathcal{H}} = \{X_1, \dots, X_k, Y_1, \dots, Y_k\}, \tag{2-3}$$

with the following properties:

(i) For all $1 \le j \le k$ there exists $R_j \in \mathcal{R}^+$ such that $\{X_j, Y_j\} \subset (\mathcal{G}_{R_j} \oplus \mathcal{G}_{-R_j}) \cap \mathcal{G}$.

(ii) $E_{\pm j} = X_j \pm i Y_j \in \mathcal{G}_{\pm R_j}$.

(iii) $\langle E_j, E_{-j} \rangle = 2$.

(iv) $[X_j, Y_j] = -R_j$.

Notice that $\{(g, \mathcal{H}_g) : g \in \mathbb{G}\}$, where $\mathcal{H}_g = g\mathcal{H}$, forms a sub-bundle of the tangent bundle of $\mathbb{G}$, which we call the horizontal sub-bundle. The property $\mathcal{T} \subset [\mathcal{H}, \mathcal{H}]$ shows that this horizontal sub-bundle is bracket-generating, hence its choice defines a sub-Riemannian metric on $\mathbb{G}$ in the following way (see [Montgomery 2002]).

We call an absolutely continuous curve $\gamma : [a, b] \to \mathbb{G}$ horizontal if $\gamma'(t) \in \mathcal{H}_{\gamma(t)}$ for every $t \in [a, b]$ where $\gamma'(t)$ exists. The length of a horizontal curve is defined as

$$\text{Length}(\gamma) = \int_a^b \|\gamma'(t)\| \, dt. \tag{2-4}$$

The bracket-generating property implies that any two points can be connected by horizontal curves and therefore we can define a sub-Riemannian (also called Carnot–Carathéodory) distance as

$$d(x, y) = \inf\{\text{Length}(\gamma) : \gamma \text{ is a horizontal curve connecting } x \text{ and } y\}.$$

We say that a horizontal curve γ is a sub-Riemannian geodesic if locally it is a length minimizer. We call a sub-Riemannian geodesic $\gamma : [0, 1] \to \mathbb{G}$ a sub-Riemannian geodesic loop if $\gamma(0) = \gamma(1) = I$ and $\gamma(t) \neq I$ for all $t \in (0, 1)$. Here, I denotes the identity matrix.

If we do not restrict the curve γ to be horizontal, then similar definitions lead to Riemannian geodesics and geodesic loops. With the choice of the bi-invariant inner product (2-1), the Riemannian geodesics through the identity and in the direction of an arbitrary $X \in \mathcal{G}$ have the form

$$\gamma(t) = e^{tX};$$

see [Arvanitoyeorgos 2003, Chapter 3].

Remark 2.2. With our assumptions on $\mathbb{G}$ and $\mathcal{H}$, all sub-Riemannian geodesics are smooth [Montgomery 1994, Theorem 3]. Moreover, as the inner product on $\mathcal{H}$ is the restriction of the inner product (2-1) defined on $\mathcal{G}$, a sub-Riemannian geodesic is also a smooth curve of equal Riemannian length.

Sub-Riemannian geodesics can be characterized in various ways. We follow the description from [Montgomery 1994; 1995; 2002], but also see [Agrachev and Sarychev 1999; Boscain et al. 2002]. If a sub-Riemannian geodesic is a projection to $\mathbb{G}$ of a solution to Hamilton's equations for the sub-Riemannian Hamiltonian, then we call it normal, otherwise we call it abnormal. If a sub-Riemannian geodesic is a critical point of the endpoint map, then we call it singular, otherwise we call it regular [Montgomery 1994]. The following implications hold.

Proposition 2.3 [Montgomery 2002, Theorem 5.8]. *All regular sub-Riemannian geodesics are normal and, therefore, all abnormal geodesics are singular.*

If the horizontal distribution is fat, which means that for all $X \in \mathcal{H}$

$$\mathcal{H} + [X, \mathcal{H}] = \mathcal{G},$$

then all sub-Riemannian geodesics are normal [Montgomery 1995, Proposition 4]. For example, the horizontal distribution is fat in the case of SU(2), but not in the case of SU(3).

Regarding the form of the normal geodesics we have the following result, which is [Montgomery 2002, Theorem 11.8] adapted to our setting. See also [Boscain et al. 2002].

Proposition 2.4. *Consider a semisimple, connected, compact Lie group $\mathbb{G}$ endowed with horizontal distribution defined by the orthogonal complement $\mathcal{H}$ of a Cartan subalgebra $\mathcal{T}$, and inner product* (2-1). *Then the normal sub-Riemannian geodesics through the identity are of the form*

$$\gamma(t) = e^{tX} \cdot e^{-tX^{\perp}}, \tag{2-5}$$

where X is any element of $\mathcal{G}$ and $X^{\perp}$ is the orthogonal projection of X onto $\mathcal{T}$.

Definition 2.5. If $X \in \mathcal{H}$, then we call $\gamma_X(t) = e^{tX}$ a horizontal Riemannian geodesic.

These are precisely the Riemannian geodesics which are also sub-Riemannian. As we will see, they can be regular or singular.

If $R \in \mathcal{R}^+$, then let us use the notation $\mathcal{H}_R = (\mathcal{G}_R \oplus \mathcal{G}_{-R}) \cap \mathcal{G}$. With this notation we can rewrite (2-2) as

$$\mathcal{H} = \mathcal{T}^{\perp} = \bigoplus_{R \in \mathcal{R}^+} \mathcal{H}_R. \tag{2-6}$$

From the relations

$$[\mathcal{H}_R, \mathcal{H}_R] = \operatorname{span}\{R\} \quad \text{and} \quad [R, \mathcal{H}_R] = \mathcal{H}_R,$$

we conclude that

$$\mathrm{su}(2)_R = \mathcal{H}_R \oplus \operatorname{span}\{R\} \tag{2-7}$$

is a subalgebra of $\mathcal{G}$, isomorphic to su(2).

For each $T \in \mathcal{T}$ let

$$\mathcal{R}^T = \{R \in \mathcal{R}^+ : \langle R, T\rangle = 0\}$$

and

$$\mathcal{G}^T = \bigoplus_{R \in \mathcal{R}^T} \mathrm{su}(2)_R. \tag{2-8}$$

If $\mathcal{R}^T \neq \varnothing$, then $\mathcal{G}^T$ is a nontrivial Lie subalgebra of $\mathcal{G}$ and therefore we can find a closed, connected subgroup $\mathbb{G}^T$ of $\mathbb{G}$, which has $\mathcal{G}^T$ as its Lie algebra. Note that $\mathbb{G}^T$ carries a sub-Riemannian geometry, for which the horizontal distribution is

$$\mathcal{H}^T = \bigoplus_{R \in \mathcal{R}^T} \mathcal{H}_R. \tag{2-9}$$

Therefore, horizontal curves in $\mathbb{G}^T$ are also horizontal in $\mathbb{G}$ and if a normal sub-Riemannian geodesic of $\mathbb{G}$ lies in $\mathbb{G}^T$, then it is a normal sub-Riemannian geodesic of $\mathbb{G}^T$ too.

A characteristic subgroup for a singular sub-Riemannian geodesic γ is a closed connected subgroup within which γ is regular.

Proposition 2.6 [Montgomery 1994, Theorem 2]. *Every singular sub-Riemannian geodesic of $\mathbb{G}$ lies in some characteristic subgroup $\mathbb{G}^T$ with dimension strictly less than the dimension of $\mathbb{G}$.*

Propositions 2.3 and 2.6 allow us to give a simple algebraic proof of the following result, which is also proved using control theoretic methods, including generalized Maslov index theory, in [Boscain et al. 2002].

Theorem 2.7. *Consider a semisimple, connected, compact Lie group $\mathbb{G}$ endowed with the horizontal distribution defined by the orthogonal complement $\mathcal{H}$ of a Cartan subalgebra $\mathcal{T}$, and inner product* (2-1). *Then we have the following results.*

(i) *All sub-Riemannian geodesics are normal.*

(ii) *All sub-Riemannian geodesics through the identity have the form*

$$\gamma(t) = e^{tX} \cdot e^{-tX^\perp},$$

where $X \in \mathcal{G}$ and $X^\perp$ is the orthogonal projection of X onto $\mathcal{T}$.

Proof. Let us assume that γ is an abnormal sub-Riemannian geodesic of $\mathbb{G}$. Then, by Proposition 2.3, γ is singular and by Proposition 2.6, there exists $T \in \mathcal{T}$ such that γ lies in a characteristic subgroup $\mathbb{G}^T$. But, as γ is regular in $\mathbb{G}^T$, by Proposition 2.3 it is also normal in $\mathbb{G}^T$. Hence, γ must have the form (2-5) in $\mathbb{G}^T$, which, by (2-6)–(2-9), gives a normal sub-Riemannian geodesic of $\mathbb{G}$.

Once all sub-Riemannian geodesics are normal, part (ii) is a direct consequence of Proposition 2.4. □

3. Lengths of sub-Riemannian geodesic loops

In this section we assume that $\mathbb{G}$ is a simple, simply connected, compact matrix Lie group.

For each root $R \in \mathcal{R}$ and $n \in \mathbb{Z}$ we define the hyperplane in $\mathcal{T}$:

$$P(R, 2\pi n) = \{T \in \mathcal{T} : \langle R, T\rangle = 2\pi n\}.$$

The reflections in $\mathcal{T}$ across the hyperplanes $P(R, 0)$ will be denoted by r_R. Note that

$$r_R(T) = T - \frac{2\langle R, T\rangle}{\|R\|^2} R.$$

The Weyl group of $\mathbb{G}$ can be defined as the group W generated by the reflections $\{r_R : R \in \mathcal{R}\}$.

The set

$$\mathcal{T} \setminus \bigcup_{R\in\mathcal{R}} P(R, 0)$$

is a union of disjoint, open cones, called Weyl chambers. The Weyl group acts transitively on the Weyl chambers. We define the positive Weyl chamber by

$$C = \{T \in \mathcal{T} : \langle R, T\rangle > 0, \ \text{ for all } R \in \mathcal{R}^+\},$$

and let $\overline{C}$ denote its closure.

Let us choose the simple roots $\mathcal{R}_s = \{R_1, \dots, R_m\}$. In the case of a simple Lie algebra, the root system is irreducible and the length of the roots can take at most 2 values, which implies that the entries of the Cartan matrix,

$$N(R_j, R_k) = \frac{2\langle R_j, R_k\rangle}{\|R_k\|^2},$$

can take only the following values:

$$N(R_j, R_k) = \begin{cases} 2 & \text{if } j = k, \\ 0, -1, -2, -3 & \text{if } j \neq k, \end{cases}$$

where at most one of -2 or -3 can appear in the matrix. For each $R \in \mathcal{R}$ we denote by

$$P_R = \frac{2\pi R}{\|R\|^2} \tag{3-1}$$

the orthogonal projection of the origin onto the hyperplane $P(R, 2\pi)$. It is known from [Helgason 2001, Chapter 7, Lemma 7.6] that

$$e^{2P_R} = I, \text{ for all } R \in \mathcal{R}. \tag{3-2}$$

The unit lattice in $\mathcal{T}$ is defined by

$$\mathcal{L}_{\mathcal{T}} = \{T \in \mathcal{T} : e^T = I\},$$

and let us also set

$$\mathcal{Z}_{\mathcal{T}} = \{n_1 2P_{R_1} + \cdots + n_m 2P_{R_m} : n_1, \ldots, n_m \in \mathbb{Z}\}.$$

By the commutativity of $\mathcal{T}$, it is evident that $\mathcal{Z}_{\mathcal{T}} \subset \mathcal{L}_{\mathcal{T}}$. By [Simon 1996, Theorem IX.1.6] we know that $\mathcal{L}_{\mathcal{T}}/\mathcal{Z}_{\mathcal{T}} \cong \pi_1(\mathbb{G})$, where $\pi_1(\mathbb{G})$ is the fundamental group of $\mathbb{G}$. Since $\mathbb{G}$ is simply connected, it follows that

$$\mathcal{L}_{\mathcal{T}} = \mathcal{Z}_{\mathcal{T}}. \tag{3-3}$$

From [Simon 1996, Theorem IX.1.4], it is also known that

$$\mathcal{L}_{\mathcal{T}} \subseteq \{T \in \mathcal{T} : \langle R, T\rangle \in 2\pi\mathbb{Z} \text{ for all } R \in \mathcal{R}\}, \tag{3-4}$$

and the two sets in (3-4) are equal only if the center of $\mathbb{G}$ equals $\{I\}$.

Definition 3.1. We call the numbers $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ relatively prime if at least one of the numbers is nonzero and the greatest common factor of the nonzero numbers is 1. In particular, if we have only one nonzero number, then it must be 1.

Remark 3.2. By (3-3), if the numbers $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime, then the line segment joining the origin to $n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}$ intersects $\mathcal{L}_{\mathcal{T}}$ only at the endpoints.

Theorem 3.3. *Let $\mathbb{G}$ be a simple, simply connected, compact Lie group endowed with the bi-invariant inner product* (2-1).

(a) *If the numbers $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime and*

$$T = n_1 2P_{R_1} + \cdots + n_m 2P_{R_m},$$

then $\gamma_T(t) = e^{tT}$, $0 \le t \le 1$, is a Riemannian geodesic loop with length

$$\|n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}\|.$$

(b) *All Riemannian geodesic loops in $\mathbb{G}$ have lengths*

$$\|n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}\|,$$

where $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime.

Proof. (a) If the rank of $\mathbb{G}$ is 1, then $\mathbb{T} = \mathrm{U}(1)$ and any geodesic loop in $\mathbb{T}$ has length 2π. Now suppose the rank of $\mathbb{G}$ is greater than or equal to 2. Let $T = n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}$, where $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime and $\gamma_T(t) = e^{tT}$. By the commutativity of the elements of $\mathcal{T}$ we know that $\gamma_T(1) = I$. If, for some $0 < t < 1$, we have $\gamma_T(t) = I$, then $t(n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}) \in \mathcal{L}_{\mathcal{T}}$ and, by Remark 3.2, this contradicts the fact that $n_1, \ldots, n_m$ are relatively prime. Hence, the length of one loop described by γ_T is

$$\text{Length}(\gamma_T) = \int_0^1 \|T\|\,dt = \|n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}\|. \tag{3-5}$$

(b) Let $X \in \mathcal{G}$ and $\gamma_X(t) = e^{tX}$. Assume that $\gamma_X(1) = I$ and $\gamma_X(t) \neq I$ if $0 < t < 1$. Since $\mathrm{Ad}(\mathbb{G})(X) \cap \mathcal{T}$ is nonempty and finite, and the Weyl group acts transitively on the Weyl chambers, and each element of the Weyl group can be written as $\mathrm{Ad}(g)$ for some $g \in \mathbb{G}$, it follows that there exists $g \in \mathbb{G}$ such that $T_X = \mathrm{Ad}(g)X \in \overline{C}$. Hence, $e^{T_X} = ge^X g^{-1} = I$ and therefore $T_X \in \mathcal{L}_{\mathcal{T}}$. By (3-3) we have that $T_X = n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}$, where $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime. Using the fact that $\|T_X\| = \|X\|$ we find that

$$\text{Length}(\gamma_X) = \int_0^1 \|T_X\|\,dt = \|n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}\|. \qquad \square$$

Remark 3.4. Moreover, for any $0 \neq T = n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}$ we have that $\mathrm{Ad}(\mathbb{G})(T) \cap (\mathcal{G} \setminus \mathcal{T}) \neq \varnothing$, so there exists $X \notin \mathcal{T}$ in the same conjugacy class with T. Hence we have a Riemannian geodesic loop outside of $\mathbb{T}$, corresponding to X, which has length equal to $\|T\|$ in (3-5).

We need the following lemma to generalize Theorem 3.3 to the case of horizontal Riemannian geodesic loops (see Definition 2.5).

Lemma 3.5. *For any $T \in \mathcal{T}$ we have $\mathrm{Ad}(\mathbb{G})(T) \cap \mathcal{H} \neq \varnothing$.*

Proof. By [D'Andrea and Maffei 2016, Lemma 2.2], given $\mathcal{T}$, we can construct another Cartan subalgebra $\mathcal{T}'$ which is orthogonal to $\mathcal{T}$. Hence, $\mathcal{T}' \subset \mathcal{H}$ and, as any two Cartan subalgebras are conjugate, there exists some $g \in \mathbb{G}$ such that $\mathrm{Ad}(g)\mathcal{T} = \mathcal{T}'$. Hence, we conclude that for any $T \in \mathcal{L}_{\mathcal{T}}$ we have that $\mathrm{Ad}(\mathbb{G})(T) \cap \mathcal{H} \neq \varnothing$. $\square$

Theorem 3.6. *Consider a simple, simply connected, compact Lie group $\mathbb{G}$ endowed with horizontal distribution defined by the orthogonal complement $\mathcal{H}$ of a Cartan subalgebra $\mathcal{T}$, and inner product* (2-1)*. Then the horizontal Riemannian geodesic loops have lengths*

$$\|n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}\|,$$

where $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime.

Proof. Let $X \in \mathcal{H}$ and $\gamma_X(t) = e^{tX}$. If $\gamma_X(1) = I$ and $\gamma_X(t) \neq I$ for all $0 < t < 1$, then we can follow the proof of Theorem 3.3(b), to conclude that there exist $n_1, \dots, n_m \in \mathbb{N} \cup \{0\}$ relatively prime such that

$$\mathrm{Length}(\gamma_X) = \|n_1 P_{R_1} + \cdots + n_m P_{R_m}\|.$$

By Lemma 3.5, the entire length spectrum of $\|n_1 P_{R_1} + \cdots + n_m P_{R_m}\|$, where $n_1, \dots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime, is covered, and this finishes the proof. □

One might expect the shortest sub-Riemannian geodesic loops to be longer than their Riemannian counterparts. Surprisingly, the following result, which generalizes the Riemannian case of [Helgason 2001, Chapter 7, Proposition 11.9], proves otherwise.

Theorem 3.7. *The shortest sub-Riemannian geodesic loops are also the shortest Riemannian geodesic loops. Their common length is* $4\pi/\|R^*\|$, *where* R^* *is the highest root.*

Proof. We first consider the Riemannian case. Without loss of generality we can assume that the rank of $\mathcal{G}$ is greater than 1. Let $\gamma^*(t) = e^{t\,2P_{R^*}}$. By (3-2) we know that $\gamma^*(1) = I$. Moreover, there exists $R \in \mathcal{R}^+$ such that

$$N(R, R^*) = 2\frac{\langle R, R^*\rangle}{\|R^*\|^2} = 1.$$

Therefore, for any $0 < t < 1$ we have

$$\langle R, t\,2P_{R^*}\rangle = 2\pi t,$$

which, by (3-4), implies that $\gamma^*(t) \neq I$ if $0 < t < 1$. Hence, the length of one loop described by γ^* is

$$\mathrm{Length}(\gamma^*) = \int_0^1 \|2P_{R^*}\|\,dt = \frac{4\pi}{\|R^*\|}.$$

Let $T = n_1 2P_{R_1} + \cdots + n_m 2P_{R_m}$, where $n_1, \dots, n_m \in \mathbb{N} \cup \{0\}$ are relatively prime and let $\gamma_T(t) = e^{tT}$. Assume that $\gamma_T(t) \neq I$ if $0 < t < 1$ and

$$\mathrm{Length}(\gamma_T) \leq \mathrm{Length}(\gamma^*).$$

Hence,

$$\|T\| \leq \frac{4\pi}{\|R^*\|} = \|2P_{R^*}\|.$$

As in the proof of Theorem 3.3, by the fact that the Weyl group acts transitively on the Weyl chambers, there exist $g \in \mathbb{G}$ and $T_1 \in \overline{C}$ such that $T_1 = \mathrm{Ad}(g)T$. Therefore, $e^{T_1} = I$ and hence $\langle R^*, T_1\rangle = 2\pi n$ for some $n \in \mathbb{N}$. By [Helgason 2001, Chapter 7, Theorem 6.1],

$$P(R^*, 2\pi) \cap \overline{C} \cap \mathcal{L}_{\mathcal{T}} = \varnothing,$$

which implies that $n \neq 1$. On the other hand, $\|T_1\| = \|T\| \leq \|2P_{R^*}\|$, which is the shortest distance from the origin to $P(R^*, 4\pi)$. Therefore, $n = 2$ and this implies that $T_1 = 2P_{R^*}$. In conclusion, we have $\mathrm{Length}(\gamma_T) = 4\pi/\|R^*\|$, which establishes the length of the shortest Riemannian geodesic loops. Note that this slight generalization of [Helgason 2001, Chapter 7, Proposition 11.9] is proved differently here.

We now consider the sub-Riemannian case. Theorem 3.6 implies that the shortest horizontal Riemannian geodesic loops have length equal to $4\pi/\|R^*\|$, which equals the length of the shortest Riemannian geodesic loops by the argument above. By Remark 2.2, every sub-Riemannian geodesic is a smooth Riemannian curve of equal length, so we conclude that $4\pi/\|R^*\|$ is the shortest length for any sub-Riemannian geodesic loop. $\square$

Theorem 3.7 implies the following result concerning the highest root.

Corollary 3.8. *We have*

$$\|R^*\| = \max \frac{4\pi}{\|n_1 2P_{R_1} + \cdots + n_m P_{R_m}\|},$$

where $n_1, \ldots, n_m \in \mathbb{N} \cup \{0\}$ *are relatively prime.*

Regarding the sub-Riemannian geodesic loops which are not necessarily horizontal Riemannian, we have the following result.

Theorem 3.9. *Let* $X = H + X^\perp$ *be such that* $H \in \mathcal{H}$ *and* $X^\perp \in \mathcal{T}$. *Consider* $\gamma(t) = e^{tX} \cdot e^{-tX^\perp}$ *and assume that* $\gamma(t) \neq I$ *if* $0 < t < 1$ *and* $\gamma(1) = I$. *Then*:

(a) *The length of* γ *satisfies*

$$\mathrm{Length}(\gamma) = \|H\| \geq \frac{4\pi}{\|R^*\|},$$

and there is an $X = H + X^\perp$ *for which* $4\pi/\|R^*\|$ *is attained.*

(b) *We have*

$$H = e^{X^\perp} H e^{-X^\perp}. \tag{3-6}$$

(c) *If*

$$\mathrm{Ad}(\mathbb{G})(X^\perp) \cap \mathcal{T} = \{S_1, \ldots, S_l\},$$

then for all $1 \leq j \leq l$ *there exist* $L_j \in \mathcal{L}_{\mathcal{T}}$ *such that*

$$\mathrm{Ad}(\mathbb{G})(X) \cap \mathcal{T} = \{S_1 + L_1, \ldots, S_l + L_l\}.$$

Proof. (a) Note that $\gamma(1) = I$ implies that $e^X = e^{X^\perp}$. Then,

$$\gamma'(t) = e^{tX} \cdot X \cdot e^{-tX^\perp} - e^{tX} \cdot X^\perp \cdot e^{-tX^\perp} = e^{tX}(X - X^\perp)e^{-tX^\perp}, \tag{3-7}$$

and

$$\|\gamma'(t)\|_{\gamma(t)} = \|\gamma(t)^{-1} \cdot \gamma'(t)\|_I = \|e^{tX^\perp} \cdot (X - X^\perp) \cdot e^{-tX^\perp}\| = \|X - X^\perp\|.$$

Hence, the length of γ is $\|X - X^{\perp}\| = \|H\|$. The fact that $\|H\|$ is at least $4\pi/\|R^*\|$ is an immediate consequence of Theorem 3.6.

Consider the simply connected Lie subgroup $\mathrm{SU}(2)_{R^*}$ of $\mathbb{G}$ which has its Lie algebra equal to $\mathrm{su}(2)_{R^*}$, and denote by X^* and Y^* those elements of (2-3) which, together with R^*, generate $\mathrm{su}(2)_{R^*}$. The relations

$$[R^*, X^*] = -\|R^*\|^2 Y^* \quad \text{and} \quad [R^*, Y^*] = \|R^*\|^2 X^*$$

show that the only positive root, and hence the highest root in $\mathrm{su}(2)_{R^*}$, is R^*. In a similar way to the proof of (4-2), we can obtain a sub-Riemannian geodesic loop in $\mathrm{SU}(2)_{R^*}$ whose length is $4\pi/\|R^*\|$.

(b) We claim that $\gamma'(0) = \gamma'(1)$. This information can be found in [Helgason 2001, page 148, Exercise 3] and its proof is based on the fact that for all $t \in \mathbb{R}$,

$$\begin{aligned}\gamma(t+1) &= e^{(t+1)X}\, e^{-(t+1)X^{\perp}} = e^{tX}\, e^{X}\, e^{-tX^{\perp}}\, e^{-X^{\perp}} \\ &= e^{tX}\, e^{X}\, e^{-X^{\perp}}\, e^{-tX^{\perp}} = e^{tX}\, e^{-tX^{\perp}} = \gamma(t).\end{aligned}$$

By (3-7), it follows that

$$X - X^{\perp} = e^{X}\,(X - X^{\perp})\, e^{-X^{\perp}},$$

which clearly implies (3-6).

(c) By the properties of the adjoint representation there exist $S_1, \dots, S_l \in \mathcal{T}$ and $S'_1, \dots, S'_l \in \mathcal{T}$, where l is the number of Weyl chambers, such that

$$\mathrm{Ad}(\mathbb{G})(X^{\perp}) \cap \mathcal{T} = \{S_1, \dots, S_l\} \tag{3-8}$$

and

$$\mathrm{Ad}(\mathbb{G})(X) \cap \mathcal{T} = \{S'_1, \dots, S'_l\}. \tag{3-9}$$

Note that in (3-8) and (3-9) some of the S_j and S'_j might be repeated if they belong to one of the hyperplanes $P(R, 0)$ for $R \in \mathcal{R}$.

Therefore,

$$\mathrm{Ad}(\mathbb{G})(e^{X^{\perp}}) \cap \mathbb{T} = \{e^{S_1}, \dots, e^{S_l}\}, \tag{3-10}$$

and

$$\mathrm{Ad}(\mathbb{G})(e^{X}) \cap \mathbb{T} = \{e^{S'_1}, \dots, e^{S'_l}\}. \tag{3-11}$$

The fact that $e^{X^{\perp}} = e^{X}$ implies that the sets in (3-10) and (3-11) must coincide. Therefore, by rearranging the elements if necessary, we can suppose that for all $1 \le j \le l$, we have $e^{S_j} = e^{S'_j}$, which immediately implies the existence of $L_j \in \mathcal{L}_{\mathcal{T}}$ such that $S'_j = S_j + L_j$. □

Since $X^{\perp} \in \mathcal{T}$, we can see that one of $S_1, \dots, S_l$ in (3-8) must be $X^{\perp}$. Therefore, we have the following corollary.

Corollary 3.10. *Under the assumptions of Theorem 3.9, there exist $g \in \mathbb{G}$ and $L \in \mathcal{L}_{\mathcal{T}}$ such that $X = \operatorname{Ad}(g)(X^\perp + L)$.*

4. The case of SU(2)

The special unitary group of 2×2 complex matrices is

$$\mathrm{SU}(2) = \left\{ g = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

Its Lie algebra is the three dimensional real Lie algebra

$$\mathrm{su}(2) = \left\{ X = \begin{pmatrix} ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & -ix_1 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\}.$$

The Killing form of su(2) is

$$K(X, Y) = 4 \operatorname{trace}(XY),$$

while the inner product (2-1) is defined as

$$\langle X, Y \rangle = -\tfrac{1}{2} \operatorname{trace}(XY).$$

The Cartan subalgebra $\mathcal{T}$ is spanned by the unit vector

$$T_1 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

and the orthonormal basis of $\mathcal{H}$ is formed by

$$X_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ and } Y_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

The exponential map $\exp : \mathrm{su}(2) \to \mathrm{SU}(2)$ has the following simple form:

$$\exp(X) = e^X = \cos(\|X\|)\, I + \frac{\sin(\|X\|)}{\|X\|}\, X.$$

Consider $X = aX_1 + bY_1 + cT_1$. Then, for $0 \le t \le 1$,

$$e^{tX} = \cos(t\sqrt{a^2 + b^2 + c^2})\, I + \frac{\sin(t\sqrt{a^2 + b^2 + c^2})}{\sqrt{a^2 + b^2 + c^2}}\, X,$$

and

$$e^{tX^\perp} = \cos(tc)\, I + \sin(tc)\, T_1.$$

If we have $\sqrt{a^2 + b^2 + c^2} = 2\pi$, the Riemannian geodesic e^{tX} closes the first time at $t = 1$, which shows that the Riemannian length spectrum equals $\{2\pi\}$. For the sub-Riemannian geodesic $\gamma(t) = e^{tX} e^{-tX^\perp}$, the condition $e^X = e^{X^\perp}$ implies that

$$\sqrt{a^2 + b^2 + c^2} = n\pi \quad \text{and} \quad |c| = m\pi, \tag{4-1}$$

where $m, n \in \mathbb{N} \cup \{0\}$, $m \leq n$, are both even or both odd. To ensure $\gamma(t) \neq I$ for all $0 < t < 1$, we require that $m, n \in \mathbb{N} \cup \{0\}$ are both odd and relatively prime or both even and $\frac{m}{2}, \frac{n}{2}$ are relatively prime with one of them odd and the other even.

Notice that, as the only positive root of SU(2) is $R_1 = 2T_1$, we have $\|R^*\| = \|2T_1\| = 2$, and therefore

$$\|H\| = \|X - X^{\perp}\| = \pi\sqrt{n^2 - m^2} = \frac{2\pi\sqrt{n^2 - m^2}}{\|R^*\|}. \tag{4-2}$$

The same result can be obtained by Theorem 3.9. In SU(2) the unit lattice is $\mathcal{L}_{\mathcal{T}} = \{2k\pi T_1 : k \in \mathbb{Z}\}$. The formula (3-6) implies that

$$c = m\pi \in \pi\mathbb{Z}.$$

The matrices S_1' and S_2' from (3-9) are diagonal with entries consisting of the eigenvalues of X. Thus, Theorem 3.9 implies that there exists some $k \in \mathbb{N}$ such that

$$\sqrt{a^2 + b^2 + m^2\pi^2} - m\pi = 2k\pi,$$

and this implies (4-1).

We have therefore presented two algebraic proofs of the following proposition, which is a special case of Theorems 3.3, 3.6, and 3.7, and which extends the results from [Chang et al. 2011; Klapheck and VanValkenburgh 2019].

Proposition 4.1. *In* SU(2) *the following properties hold.*

(a) *The Riemannian geodesic loops have length equal to* 2π.

(b) *The horizontal Riemannian geodesic loops have length equal to* 2π.

(c) *The shortest sub-Riemannian geodesic loops have length equal to* 2π.

(d) *The sub-Riemannian geodesic loops have lengths equal to* $\pi\sqrt{n^2 - m^2}$, *where* $m, n \in \mathbb{N} \cup \{0\}$ *are odd and relatively prime or even and* $\frac{m}{2}, \frac{n}{2}$ *are relatively prime with one of them odd and the other even.*

Remark 4.2. As an introduction to the next section, let us show that we can use Viète's formulas to get the result of Proposition 4.1(d). Indeed, the characteristic polynomial of X is

$$P(\lambda) = \lambda^2 + (a^2 + b^2 + c^2),$$

and by Theorem 3.9 and the first Viète formula, the eigenvalues of X must be of the form $\lambda_1 = -ci - 2k\pi i$ and $\lambda_2 = ci + 2k\pi i$, where $k \in \mathbb{N}$. The second Viète formula gives

$$\lambda_1\lambda_2 = (c + 2k\pi)^2 = a^2 + b^2 + c^2,$$

which leads to (4-1).

Remark 4.3. For comparison with the case of SU(3) in the next section, note that in SU(2) the sub-Riemannian geodesic loops have the form

$$\gamma(t) = e^{t\left(a_1 X_1 + b_1 Y_1 + \frac{m}{2} R_1\right)} e^{-t\frac{m}{2} R_1}, \tag{4-3}$$

where a, b, c, m satisfy (4-1).

5. The case of SU(3)

Consider the special unitary group of 3×3 complex matrices

$$\mathrm{SU}(3) = \{g \in \mathrm{GL}(3, \mathbb{C}) : g \cdot g^* = I,\ \det g = 1\},$$

and its Lie algebra

$$\mathrm{su}(3) = \{X \in \mathrm{gl}(3, \mathbb{C}) : X + X^* = 0,\ \mathrm{trace}\, X = 0\}.$$

The inner product is defined by

$$\langle X, Y \rangle = -\tfrac{1}{2}\,\mathrm{trace}(XY).$$

We consider the maximal torus

$$\mathbb{T} = \left\{ \begin{pmatrix} e^{ia_1} & 0 & 0 \\ 0 & e^{ia_2} & 0 \\ 0 & 0 & e^{ia_3} \end{pmatrix} : a_1,\ a_2,\ a_3 \in \mathbb{R},\ a_1 + a_2 + a_3 = 0 \right\}$$

and its Lie algebra

$$\mathcal{T} = \left\{ \begin{pmatrix} ia_1 & 0 & 0 \\ 0 & ia_2 & 0 \\ 0 & 0 & ia_3 \end{pmatrix} : a_1,\ a_2,\ a_3 \in \mathbb{R},\ a_1 + a_2 + a_3 = 0 \right\},$$

which is our choice for the Cartan subalgebra. The following are the Gell-Mann matrices, which form an orthonormal basis of su(3) and satisfy the relations in (2-2), (2-3), and (i)–(iv) on page 324:

$$T_1 = \begin{pmatrix} -i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \frac{-i}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{-i}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{2i}{\sqrt{3}} \end{pmatrix}, \quad X_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y_1 = \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad Y_3 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}.$$

The positive roots are

$$R_1 = \begin{pmatrix} -2i & 0 & 0 \\ 0 & 2i & 0 \\ 0 & 0 & 0 \end{pmatrix} = 2T_1,$$

$$R_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2i & 0 \\ 0 & 0 & -2i \end{pmatrix} = T_1 - \sqrt{3}T_2,$$

$$R_3 = \begin{pmatrix} -2i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2i \end{pmatrix} = T_1 + \sqrt{3}T_2.$$

The highest root is $R^* = R_1$, while the two simple roots are R_2 and R_3. The unit lattice is

$$\mathcal{L}_{\mathcal{T}} = \{n\pi R_2 + m\pi R_3 : n,\ m \in \mathbb{Z}\},$$

and observe that

$$e^{\pi R_1} = e^{\pi R_2} = e^{\pi R_3} = I. \tag{5-1}$$

For $k = 1, 2, 3$, the projections of the origin onto the hyperplanes $P(R_k, 2\pi)$ are

$$P_{R_k} = \frac{\pi}{2} R_k,$$

and, indeed, (5-1) is equivalent to

$$e^{2P_{R_k}} = I, \quad k = 1, 2, 3.$$

Observing that

$$\|n\pi R_2 + m\pi R_3\| = 2\pi\sqrt{n^2 - nm + m^2},$$

we conclude that, in SU(3), Theorems 3.3, 3.6, and 3.7 have the following special form.

Proposition 5.1. *In* SU(3) *the following properties hold.*

(a) *The Riemannian geodesic loops have lengths equal to*

$$2\pi\sqrt{n^2 - nm + m^2},$$

where $n, m \in \mathbb{N} \cup \{0\}$ *are relatively prime.*

(b) *The horizontal Riemannian geodesic loops have lengths equal to*

$$2\pi\sqrt{n^2 - nm + m^2},$$

where $n, m \in \mathbb{N} \cup \{0\}$ *are relatively prime.*

(c) *The shortest sub-Riemannian geodesic loops have length equal to* 2π.

To obtain information about the full sub-Riemannian length spectrum in SU(3), consider

$$H = a_1X_1 + b_1Y_1 + a_2X_2 + b_2Y_2 + a_3X_3 + b_3Y_3,$$

$$X^{\perp} = \frac{c_1}{2}R_3 + \frac{c_2}{2}R_2 = \begin{pmatrix} -c_1 i & 0 & 0 \\ 0 & c_2 i & 0 \\ 0 & 0 & (c_1-c_2)i \end{pmatrix},$$

$$X = H + X^{\perp}, \text{ and } \gamma(t) = e^{tX}e^{-tX^{\perp}}.$$

The characteristic polynomial of X is

$$P(\lambda) = -\lambda^3 - p\lambda + qi,$$

where

$$p = \sum_{j=1}^{3}(a_j^2 + b_j^2) + c_1^2 + c_2^2 - c_1c_2,$$

and

$$q = c_2(a_3^2 + b_3^2 - a_1^2 - b_1^2 + c_1^2) - c_1(a_2^2 + b_2^2 - a_1^2 - b_1^2 + c_2^2) \\ +2(a_1a_2b_3 + a_1b_2a_3 - b_1a_2a_3 + b_1b_2b_3).$$

Note that $p = \|H\|^2 + \|X^{\perp}\|^2$. Formula (3-6) gives

$$\begin{aligned} c_1 + c_2 &\in 2\pi\mathbb{Z} \quad \text{if } a_1 + b_1 i \neq 0, \\ c_1 - 2c_2 &\in 2\pi\mathbb{Z} \quad \text{if } a_2 + b_2 i \neq 0, \\ 2c_1 - c_2 &\in 2\pi\mathbb{Z} \quad \text{if } a_3 + b_3 i \neq 0. \end{aligned} \tag{5-2}$$

To see the connection with the case of SU(2), let us start with the following simple cases.

Case 1: Consider $c_2 = -c_1$, $a_2 = b_2 = a_3 = b_3 = 0$. This corresponds to the case of SU(2) from the previous section and these geodesics are singular in SU(3). Therefore the sub-Riemannian geodesics have the form (4-3) and the lengths $\pi\sqrt{n^2 - m^2}$ from Proposition 4.1(d).

Case 2: Consider $c_2 = 0$, $a_2 = b_2 = a_3 = b_3 = 0$. These geodesics are not contained in any copy of SU(2) and are regular in SU(3). Here, $c_1 = 2m\pi$, $m \in \mathbb{Z}$, and the eigenvalues of X are

$$-2m\pi i \quad \text{and} \quad \Big(m\pi \pm \sqrt{a_1^2 + b_1^2 + m^2\pi^2}\Big)i.$$

By Theorem 3.9(c) we have that

$$|c_1| = 2m\pi \quad \text{and} \quad a_1^2 + b_1^2 + m^2\pi^2 = n^2\pi^2, \tag{5-3}$$

where $m, n \in \mathbb{N} \cup \{0\}$, $m \le n$, are both odd or even. Therefore, the sub-Riemannian geodesic loop corresponding to (5-3) is

$$\gamma(t) = e^{t(a_1 X_1 + b_1 Y_1 + m R_3)}\, e^{-tmR_3},$$

and its length is $\pi\sqrt{n^2 - m^2}$.

Case 3: If at least two of $a_j + b_j i$, $j = 1, 2, 3$, are not zero, then

$$c_1 = \frac{4n + 2m}{3}\pi, \quad c_2 = \frac{2n - 2m}{3}\pi,$$

where $n, m \in \mathbb{Z}$. Theorem 3.9(c) and the first Viète formula for the characteristic polynomial imply that the eigenvalues of X must have the form

$$\begin{aligned} \lambda_1 &= (-c_1 - 2k\pi)i, \\ \lambda_2 &= (c_2 + 2l\pi)i, \\ \lambda_3 &= (c_1 - c_2 + 2(k-l)\pi)i. \end{aligned} \tag{5-4}$$

The second Viète formula gives

$$\|H\| = \sqrt{c_1(4k - 2l)\pi + c_2(4l - 2k)\pi + 4(k^2 + l^2 - kl)\pi^2}. \tag{5-5}$$

From the third Viète formula we find

$$\begin{aligned} 4c_1c_2(k-l)\pi + 4c_1 l(2k-l)\pi^2 + 4c_2 k(k-2l)\pi^2 + 2c_1^2 l\pi - 2c_2^2 k\pi + 8kl(k-l)\pi^3 \\ = (a_3^2 + b_3^2 - a_1^2 - b_1^2)c_2 - (a_2^2 + b_2^2 - a_1^2 - b_1^2)c_1 \\ + 2(a_1a_2b_3 + a_1b_2a_3 - b_1a_2a_3 + b_1b_2b_3). \end{aligned}$$

The complexity of this formula hides its true geometric meaning; however, in the case when $q = 0$, it reduces to $0 = 0$, and we have the following eigenvalues for X:

$$0 \quad \text{and} \quad \pm\sqrt{\|H\|^2 + \|X^\perp\|^2}\, i.$$

Without loss of generality we can assume that $\lambda_1 = 0$. Then $c_1 = 2k\pi$, which implies that $m = 3k - 2n$ and $c_2 = 2(n - k)$. From (5-5) it follows that

$$\|H\| = 2\pi\sqrt{(2k - l)^2 - nk + 2nl},$$

which in the case of $k = l$ reduces to

$$\|H\| = 2\pi\sqrt{k^2 + nk} = 2\pi\sqrt{\left(k + \tfrac{n}{2}\right)^2 - \tfrac{n^2}{4}}.$$

This shows that, as expected, formula (5-5) includes the sub-Riemannian geodesic loop length spectrum of SU(2).

Note that $q = 0$ is satisfied if $c_2 = 0$, $a_1^2 + b_1^2 = a_2^2 + b_2^2$ and $a_3 = b_3 = 0$. As a numerical example we can give the sub-Riemannian geodesic loop of length 8π described by

$$\gamma(t) = e^{\pi(5X_1+\sqrt{7}Y_1+5X_2+\sqrt{7}Y_2+3R_3)t}e^{-3\pi R_3 t}.$$

In conclusion, we have the following result.

Proposition 5.2. *In* $\mathrm{SU}(3)$ *the sub-Riemannian geodesic loops have lengths equal to*

$$2\pi\sqrt{\left(\frac{(2n+m)(2k-l)}{3}+\frac{(n-m)(2l-k)}{3}\right)+(k^2+l^2-kl)},$$

where $m, n, k, l \in \mathbb{Z}$.

Acknowledgements

The authors are grateful to Richard Montgomery for helpful discussions and advice on exposition. We also thank the referees for valuable suggestions.

References

[Agrachev and Sarychev 1999] A. A. Agrachev and A. V. Sarychev, "Sub-Riemannian metrics: minimality of abnormal geodesics versus subanalyticity", *ESAIM Control Optim. Calc. Var.* **4** (1999), 377–403. MR Zbl

[Arvanitoyeorgos 2003] A. Arvanitoyeorgos, *An introduction to Lie groups and the geometry of homogeneous spaces*, Student Mathematical Library **22**, American Mathematical Society, Providence, RI, 2003. MR Zbl

[Boscain et al. 2002] U. Boscain, T. Chambrion, and J.-P. Gauthier, "On the $K + P$ problem for a three-level quantum system: optimality implies resonance", *J. Dynam. Control Systems* **8**:4 (2002), 547–572. MR Zbl

[Chang et al. 2011] D.-C. Chang, I. Markina, and A. Vasil'ev, "Hopf fibration: geodesics and distances", *J. Geom. Phys.* **61**:6 (2011), 986–1000. MR Zbl

[D'Andrea and Maffei 2016] A. D'Andrea and A. Maffei, "Commutators of small elements in compact semisimple groups and Lie algebras", *J. Lie Theory* **26**:3 (2016), 683–690. MR

[Domokos 2015] A. Domokos, "Subelliptic Peter–Weyl and Plancherel theorems on compact, connected, semisimple Lie groups", *Nonlinear Anal.* **126** (2015), 131–142. MR Zbl

[Duistermaat and Kolk 2000] J. J. Duistermaat and J. A. C. Kolk, *Lie groups*, Springer, 2000. MR Zbl

[Helgason 2001] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Graduate Studies in Mathematics **34**, American Mathematical Society, Providence, RI, 2001. MR Zbl

[Klapheck and VanValkenburgh 2019] D. Klapheck and M. VanValkenburgh, "The length spectrum of the sub-Riemannian three-sphere", *Involve* **12**:1 (2019), 45–61.

[Klingenberg 1978] W. Klingenberg, *Lectures on closed geodesics*, Grundlehren der Math. Wissenschaften **230**, Springer, 1978. MR Zbl

[Knapp 1986] A. W. Knapp, *Representation theory of semisimple groups*, Princeton Mathematical Series **36**, Princeton University Press, 1986. MR Zbl

[Montgomery 1994] R. Montgomery, "Singular extremals on Lie groups", *Math. Control Signals Systems* **7**:3 (1994), 217–234. MR Zbl

[Montgomery 1995] R. Montgomery, "A survey of singular curves in sub-Riemannian geometry", *J. Dynam. Control Systems* **1**:1 (1995), 49–90. MR Zbl

[Montgomery 2002] R. Montgomery, *A tour of sub-Riemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs **91**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Simon 1996] B. Simon, *Representations of finite and compact groups*, Graduate Studies in Mathematics **10**, American Mathematical Society, Providence, RI, 1996. MR Zbl

Received April 7, 2017. Revised December 11, 2017.

ANDRÁS DOMOKOS
domokos@csus.edu

MATTHEW KRAUEL
krauel@csus.edu

VINCENT PIGNO
vincent.pigno@csus.edu

COREY SHANBROM
corey.shanbrom@csus.edu

MICHAEL VANVALKENBURGH
mjv@csus.edu
.

(*all authors*)

DEPARTMENT OF MATHEMATICS AND STATISTICS
CALIFORNIA STATE UNIVERSITY, SACRAMENTO
SACRAMENTO, CA
UNITED STATES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.341

THE ACTION OF THE HECKE OPERATORS ON THE COMPONENT GROUPS OF MODULAR JACOBIAN VARIETIES

TAEKYUNG KIM AND HWAJONG YOO

For a prime number $q \geq 5$ and a positive integer N prime to q, Ribet proved the action of the Hecke algebra on the component group of the Jacobian variety of the modular curve of level Nq at q is "Eisenstein", which means the Hecke operator T_ℓ acts by $\ell + 1$ when ℓ is a prime number not dividing the level. We completely compute the action of the Hecke algebra on this component group by a careful study of supersingular points with extra automorphisms.

1. Introduction

Let $q \geq 5$ be a prime number, and let N be a positive integer. Let $X_0(Nq)$ denote the modular curve over $\mathbb{Q}$ and $J_0(Nq)$ its Jacobian variety. For any integer n, there is the Hecke operator T_n acting on $J_0(Nq)$. Let $\Phi_q(Nq)$ denote the component group of the special fiber $\mathcal{J}$ of the Néron model of $J_0(Nq)$ at q. According to the theorems of Ribet [1988; 1990] (when q does not divide N) and Edixhoven [1991] (in general), the action of the Hecke algebra on $\Phi_q(Nq)$ is "Eisenstein." Here by "Eisenstein" we mean the Hecke operator T_ℓ acts on $\Phi_q(Nq)$ by $\ell + 1$ when a prime number ℓ does not divide Nq.[1] In this article, we compute the action of the Hecke operators T_ℓ on the component group $\Phi_q(Nq)$ when ℓ divides Nq and q does not divide N.

Here is an exotic example[2] which leads us to this study: Let $N = \prod_{i=1}^{\nu} p_i$ be the product of distinct prime numbers with $\nu \geq 1$, and let $q \equiv 2$ or $5 \pmod 9$ be an odd prime number. Assume that $p_i \equiv 4$ or $7 \pmod 9$ for all $1 \leq i \leq \nu$. Let $\mathbb{T}(Nq)$ and $\mathbb{T}(N)$ denote the $\mathbb{Z}$-subalgebras of $\mathrm{End}(J_0(Nq))$ and $\mathrm{End}(J_0(N))$, respectively, generated by all the Hecke operators T_n for $n \geq 1$. Let

$$\mathfrak{m} := \big(3,\, T_{p_i} - 1,\, T_q + 1,\, T_\ell - \ell - 1 :\ \text{for all } 1 \leq i \leq \nu, \\ \text{and for primes } \ell \nmid Nq\big) \subset \mathbb{T}(Nq)$$

MSC2010: 11G05, 11G18, 14G35.
Keywords: Hecke operators, Hecke action, component group, modular Jacobian varieties.

[1] On the other hand, Ribet and Edixhoven did not proceed to compute the action of the Hecke operator T_p on $\Phi_q(Nq)$ for a prime divisor p of the level Nq because their results were enough for their applications.

[2] This phenomenon cannot occur when the residual characteristic is greater than 3.

and

$$\mathfrak{n} := \big(3,\, T_{p_i} - 1,\, T_\ell - \ell - 1 : \text{ for all } 1 \le i \le \nu, \text{ and for primes } \ell \nmid N\big) \subset \mathbb{T}(N)$$

be Eisenstein ideals. By [Yoo 2016, Theorem 1.4], $\mathfrak{m}$ is maximal. Furthermore, $\mathfrak{n}$ is maximal if and only if $\nu \ge 2$.

The dimension of $J_0(N)[\mathfrak{n}]$ is equal to ν if $\mathfrak{n}$ is maximal, i.e., $\nu \ge 2$. (Here $J_0(N)[\mathfrak{n}] := \{x \in J_0(N)(\overline{\mathbb{Q}}) : Tx = 0 \text{ for all } T \in \mathfrak{n}\}$.) It is an extension of $\mu_3^{\oplus \nu-1}$ by $\mathbb{Z}/3\mathbb{Z}$, and it does not contain a submodule isomorphic to μ_3. On the other hand, the dimension of $J_0(Nq)[\mathfrak{m}]$ is either 2ν or $2\nu+1$. Furthermore $J_0(Nq)[\mathfrak{m}]$ contains a submodule $\mathcal{N}$ isomorphic to $J_0(N)[\mathfrak{n}]$, and it also contains $\mu_3^{\oplus\nu}$ (which is contributed from the Shimura subgroup). As $\mathcal{N}$ is unramified at q, by [Serre and Tate 1968], $\mathcal{N}$ maps injectively into $\mathcal{J}[\mathfrak{m}]$ and it turns out that its image is isomorphic to $\mathcal{J}^0[\mathfrak{m}]$, where $\mathcal{J}^0$ is the identity component of $\mathcal{J}$. (Note that $\Phi_q(Nq)$ is the quotient of $\mathcal{J}$ by $\mathcal{J}^0$.) Since $\mu_3^{\oplus\nu}$ is also unramified at q, it maps into $\mathcal{J}[\mathfrak{m}]$ and therefore its image maps injectively to $\Phi_q(Nq)[\mathfrak{m}]$. (This statement is also true when $\nu = 1$.) The structure of the component group $\Phi_q(Nq)$ is known by the work of Mazur and Rapoport [1977]:[3]

$$\Phi_q(Nq) = \Phi \oplus (\mathbb{Z}/3\mathbb{Z})^{2^\nu - 1},$$

where Φ is cyclic and generated by the image of the cuspidal divisor $(0) - (\infty)$. The action of the Hecke operators on Φ is well known (e.g., [Yoo 2014, Appendix A1]), and so $\Phi[\mathfrak{m}] = 0$. Therefore $(\mathbb{Z}/3\mathbb{Z})^{2^\nu-1}[\mathfrak{m}] \neq 0$ and its dimension is at least ν. Indeed it is equal to $2^{\nu-1}$, which can easily be computed by the theorems below.

Now, we introduce our results.

Theorem 1.1. *For a prime divisor p of N, the Hecke operator T_p acts on $\Phi_q(Nq)$ by p.*

The key idea of the proof is that the two degeneracy maps coincide on the component group (see [Ribet 1988; Edixhoven 1991, §4.2, Lemme 2]).

Now, the missing action is that of the Hecke operator T_q on $\Phi_q(Nq)$. Note that T_q acts on $\Phi_q(Nq)$ by an involution because the action of the Hecke algebra on $\Phi_q(Nq)$ is "q-new." To describe its action more precisely, we define some notation: for $N = \prod_{p \mid N} p^{n_p}$ being the prime factorization of N (i.e., $n_p > 0$), let $\nu := \#\{p : p \neq 2, 3\}$ and let

$$u := \begin{cases} 0 & \text{if } q \equiv 1 \pmod 4 \text{ or } 4 \mid N \text{ or if there exists } p \equiv -1 \pmod 4, \\ 1 & \text{otherwise,} \end{cases}$$

$$v := \begin{cases} 0 & \text{if } q \equiv 1 \pmod 3 \text{ or } 9 \mid N \text{ or if there exists } p \equiv -1 \pmod 3, \\ 1 & \text{otherwise.} \end{cases}$$

[3] There are some minor errors in the paper, which are corrected by Edixhoven [1991, §4.4.1]

Suppose that $(u, v) = (0, 0)$ or $\nu = 0$. Then $\Phi_q(Nq) = \Phi$ and T_q acts on Φ by 1, where Φ is the cyclic subgroup generated by the image of the cuspidal divisor $(0) - (\infty)$ (Proposition 4.1). If $\nu \geq 1$, $\Phi_q(Nq)$ becomes isomorphic to

$$\Phi' \oplus \boldsymbol{A} \oplus \boldsymbol{B},$$

where $\boldsymbol{A} \simeq (\mathbb{Z}/2\mathbb{Z})^{\oplus u(2^\nu - 2)}$, $\boldsymbol{B} \simeq (\mathbb{Z}/3\mathbb{Z})^{\oplus v(2^\nu - 1)}$ and Φ' is a cyclic group containing Φ and $\Phi'/\Phi \simeq (\mathbb{Z}/2^u\mathbb{Z})$.[4]

Theorem 1.2. *Assume that $(u, v) \neq (0, 0)$ and $\nu \geq 1$.*

(1) *Suppose that $v = 1$. Then there are distinct subgroups $B_i \simeq \mathbb{Z}/3\mathbb{Z}$ of $\boldsymbol{B}$ so that $\boldsymbol{B} = \bigoplus B_i$. For any $1 \leq i \leq (2^\nu - 1)$, T_q acts on B_i by $(-1)^i$.*

(2) *Suppose that $u = 1$. Then there are distinct subgroups $A_i \simeq \mathbb{Z}/2\mathbb{Z}$ of $\boldsymbol{A}$ so that $\boldsymbol{A} = \oplus A_i$. For any $1 \leq k \leq (2^{\nu-1} - 2)$, T_q acts on $A_{2k-1} \oplus A_{2k}$ by the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$.[5] In other words, if $A_{2k-1} = \langle \boldsymbol{u}_{2k-1} \rangle$ and $A_{2k} = \langle \boldsymbol{u}_{2k} \rangle$, then*

$$T_q(\boldsymbol{u}_{2k-1}) = \boldsymbol{u}_{2k-1} + \boldsymbol{u}_{2k} \quad \textit{and} \quad T_q(\boldsymbol{u}_{2k}) = \boldsymbol{u}_{2k}.$$

For a complete description of the action of T_q on each subgroup, see Section 4.

2. Supersingular points of $X_0(N)$

From now on, we always assume that $q \geq 5$ is a prime number and N is a positive integer which is prime to q. Let p denote a prime divisor of N. Let $\boldsymbol{F}$ be an algebraically closed field of characteristic q.

Let $\Sigma(N)$ denote the set of supersingular points of $X_0(N)(\boldsymbol{F})$. Since we assume that $q \geq 5$, the group of automorphisms of supersingular points is cyclic of order 2, 4 or 6. Let

$$\Sigma_n(N) := \{s \in \Sigma(N) : \#\mathrm{Aut}(s) = n\} \quad \text{and} \quad s_n(N) := \#\Sigma_n(N).$$

Note that $s_4(N) = u \cdot 2^\nu$ and $s_6(N) = v \cdot 2^\nu$ (see [Edixhoven 1991, §4.2, Lemme 1]), where u, v and ν are as in Section 1. Moreover $s_2(N)$ can be computed using Eichler's mass formula [Katz and Mazur 1985, Theorem 12.4.5, Corollary 12.4.6]:

$$\frac{s_2(N)}{2} + \frac{s_4(N)}{4} + \frac{s_6(N)}{6} = \frac{(q-1)Q}{24}, \tag{2-1}$$

where $Q := N \prod_{p \mid N} (1 + p^{-1})$ is the degree of the degeneracy map $X_0(N) \to X_0(1)$.

[4] The structure of $\Phi_q(Nq)$ is already known by [Mazur and Rapoport 1977] when N is square-free and prime to 6, and by [Edixhoven 1991, §4.4.1] in general.

[5] This reminds us of the result by Mazur [1977]: when N is a prime number, the kernel of the Eisenstein prime of $J_0(N)$ containing a prime number ℓ is completely reducible when ℓ is odd, and is indecomposable when $\ell = 2$.

In the remainder of this section, we study $\Sigma_4(N)$ and $\Sigma_6(N)$ in detail. (See also [Ribet 1988, §2; 1995, §4; Edixhoven 1991, §4.2].) In the section below, we always assume that $\nu \geq 1$, i.e., there is a prime divisor $p \geq 5$ of N. (If $\nu = 0$ then $s_{2e}(N) \leq 1$ for $e = 2$ or 3, and the description is very simple.)

Let $\mathcal{E}$ be a supersingular elliptic curve with $\mathrm{Aut}(\mathcal{E}) = \langle \sigma \rangle$, and let C be a cyclic subgroup of $\mathcal{E}$ of order N. Assume that $q \equiv -1 \pmod 4$ (resp. $q \equiv -1 \pmod 3$) if $\sigma = \sigma_4$ (resp. $\sigma = \sigma_6$), where σ_k is a primitive k-th root of unity.

Proposition 2.1. *Let $N = p^n$ for some $n \geq 1$ with $p \geq 5$. Suppose* $\mathrm{Aut}(\mathcal{E}, C) = \langle \sigma \rangle$. *Then, there exists another cyclic subgroup D of order N such that $\mathcal{E}[N] \simeq C \oplus D$. Moreover,* $\mathrm{Aut}(\mathcal{E}, D) = \langle \sigma \rangle$ *and $(\mathcal{E}, C)$ is not isomorphic to $(\mathcal{E}, D)$.*

Proof. Here, we closely follow the argument in the proof of Proposition 1 in [Ribet 1988, §2].

Let R be the subring $\mathbb{Z}[\sigma]$ of $\mathrm{End}(\mathcal{E}, C)$. Since $\mathrm{Aut}(\mathcal{E}, C) = \langle \sigma \rangle$, $p \equiv 1 \pmod 4$ (resp. $p \equiv 1 \pmod 3$) if $\sigma = \sigma_4$ (resp. $\sigma = \sigma_6$). Therefore p splits completely in R. Note that $R = \mathbb{Z}[\sigma]$ is a principal ideal domain and therefore

$$R/pR \simeq R/\gamma R \oplus R/\delta R \simeq \delta R/pR \oplus \gamma R/pR$$

with $p = \gamma\delta$. Moreover,

$$R/NR = R/p^n R \simeq R/\gamma^n R \oplus R/\delta^n R \simeq \delta^n R/NR \oplus \gamma^n R/NR.$$

Note that $\mathcal{E}[N]$ is a free module of rank 1 over R/NR by the action of R on $\mathcal{E}$. We may identify C with the quotient I/NR for some ideal I of R containing N if we fix an R-isomorphism between $\mathcal{E}[N]$ and R/NR. Thus, $I = \delta^n R$ or $\gamma^n R$. Suppose that $I = \delta^n R$. Then, by the fixed isomorphism, $C = \mathcal{E}[\gamma^n]$. Let $D := \mathcal{E}[\delta^n]$ so that its corresponding ideal is $\gamma^n R$. Then, $\mathcal{E}[N] \simeq C \oplus D$. Moreover since $\gamma^n R$ is also an ideal of R, D is also stable under the action of σ. In other words, $\mathrm{Aut}(\mathcal{E}, D) = \langle \sigma \rangle$. Also, $(\mathcal{E}, C)$ cannot be isomorphic to $(\mathcal{E}, D)$ since $\mathrm{Aut}(\mathcal{E}) = \langle \sigma \rangle$ and $\sigma(C) = C$. □

From now on, we use the same notation as in the proof of Proposition 2.1.

Definition 2.2. By the above formulas, for every $n \geq 1$ and $p \equiv 1 \pmod 4$ (resp. $p \equiv 1 \pmod 3$), there are precisely two cyclic subgroups C, D of $\mathcal{E}$ of order p^n such that $\mathrm{Aut}(\mathcal{E}, C) = \mathrm{Aut}(\mathcal{E}, D) = \langle \sigma \rangle$ (and $\mathcal{E}[p^n] \simeq C \oplus D$) if $\sigma = \sigma_4$ (resp. if $\sigma = \sigma_6$). Thus, for each $n \geq 1$ we define $\mathcal{C}_{p^n}$ and $\mathcal{D}_{p^n}$ by

$$\mathcal{C}_{p^n} := \mathcal{E}[\gamma^n] \quad \text{and} \quad \mathcal{D}_{p^n} := \mathcal{E}[\delta^n].$$

Proposition 2.3. *For each $n \geq 1$, $\mathcal{C}_{p^{n+1}}[p^n] = \mathcal{C}_{p^n}$ and $\mathcal{D}_{p^{n+1}}[p^n] = \mathcal{D}_{p^n}$.*

Proof. By the fixed R-isomorphism ι between $\mathcal{E}[p^{n+1}]$ and $R/p^{n+1}R$, we identify $\mathcal{C}_{p^{n+1}}$ with $I/p^{n+1}R$, where $I = \delta^{n+1}R$. As I is an ideal of R, $\gamma I = p(\delta^n R) \subset I$

and $I/\gamma I \simeq R/\gamma R \simeq \mathbb{Z}/p\mathbb{Z}$. Therefore

$$\mathcal{C}_{p^{n+1}}[p^n] \xrightarrow{\iota} (I/p^{n+1}R)[p^n] = \gamma I/p^{n+1}R \xrightarrow[\times 1/p]{\sim} (\delta^n R)/p^n R,$$

which corresponds to $\mathcal{C}_{p^n}$. Similarly, we prove that $\mathcal{D}_{p^{n+1}}[p^n] = \mathcal{D}_{p^n}$, and the proposition follows. $\square$

Let $N = Mp^n$ with $(6M, p) = 1$ and $n \geq 1$. Let L be a cyclic subgroup of $\mathcal{E}$ of order M.

Proposition 2.4. *Suppose that* $\mathrm{Aut}(\mathcal{E}, \mathcal{C}_{p^{n+1}}, L) = \langle\sigma\rangle$. *Then, there is an isomorphism between* $(\mathcal{E}/\mathcal{C}_p, \mathcal{C}_{p^{n+1}}/\mathcal{C}_p, (L \oplus \mathcal{C}_p)/\mathcal{C}_p))$ *and* $(\mathcal{E}, \mathcal{C}_{p^n}, L)$.

Proof. We mostly follow the idea of the proof of Proposition 2 in [Ribet 1988, §2].

The endomorphism γ sends $\mathcal{E}[\gamma^{n+1}] = \mathcal{C}_{p^{n+1}}$ to $\mathcal{E}[\gamma^n] = \mathcal{C}_{p^n}$, and L to itself (because $L \cap \mathcal{E}[p] = 0$). Now we denote by $\bar{\gamma}$ the map $\mathcal{E}/\mathcal{C}_p \to \mathcal{E}$ induced by γ. Note that $\bar{\gamma}$ is an isomorphism because $\mathcal{C}_p$ is $\mathcal{E}[\gamma]$, the kernel of γ. By the above consideration, this isomorphism $\bar{\gamma}$ sends $(\mathcal{C}_{p^{n+1}}/\mathcal{C}_p, (L \oplus \mathcal{C}_p)/\mathcal{C}_p)$ to $(\mathcal{C}_p, L)$ because $\mathcal{C}_{p^{n+1}}/\mathcal{C}_p$ and $(L \oplus \mathcal{C}_p)/\mathcal{C}_p$, respectively, are the images of $\mathcal{C}_{p^{n+1}}$ and L by the quotient map $\mathcal{E} \to \mathcal{E}/\mathcal{C}[p]$. Therefore $\bar{\gamma}$ gives rise to the desired isomorphism between triples. $\square$

Corollary 2.5. *The map* $(\mathcal{E}, C, L) \to (\mathcal{E}, C[p^n], L)$ *induces a bijection between* $\Sigma_{2e}(Np)$ *and* $\Sigma_{2e}(N)$, *where* $\sigma = \sigma_{2e}$. *Moreover if* $(\mathcal{E}, C, L) \in \Sigma_{2e}(Np)$, *we have*

$$(\mathcal{E}, C[p^n], L) \simeq (\mathcal{E}/C[p], C/C[p], (L \oplus C[p])/C[p]).$$

The corollary tells us that two degeneracy maps α_p and β_p in Section 3 coincide on $\Sigma_{2e}(Np)$, which is a generalization of [Edixhoven 1991, §4.2, Lemme 2].

Proposition 2.6. *Suppose that* $\mathrm{Aut}(\mathcal{E}, \mathcal{C}_{p^n}, L) = \langle\sigma\rangle$. *Then*, $\mathrm{Frob}(\mathcal{E}) = \mathcal{E}$ *and* $\mathrm{Frob}(\mathcal{C}_{p^n}) = \mathcal{D}_{p^n}$, *where* Frob *is the Frobenius morphism in characteristic* q. *Furthermore*, $\mathrm{Frob}^2(\mathcal{E}, \mathcal{C}_{p^n}, L) = (\mathcal{E}, \mathcal{C}_{p^n}, L)$.

Proof. Since $\mathcal{E}$ is isomorphic to the reduction of the elliptic curve with j-invariant 1728 (resp. 0) if $\sigma = \sigma_4$ (resp. $\sigma = \sigma_6$), the Frobenius morphism is an endomorphism of $\mathcal{E}$ (see [Silverman 2009, Chapter V, Examples 4.4 and 4.5]). Moreover, the Frobenius morphism and σ generate $\mathrm{End}(\mathcal{E})$, which is a quaternion algebra. (Note that the degree of the Frobenius morphism is q.) Since $\mathrm{End}(\mathcal{E})$ is a quaternion algebra, we have

$$\sigma \circ \mathrm{Frob} = \mathrm{Frob} \circ \bar{\sigma} = \mathrm{Frob} \circ \sigma^{-1},$$

where $\bar{\sigma}$ denotes the complex conjugation in $R = \mathbb{Z}[\sigma]$. Analogously, we have

$$\gamma \circ \mathrm{Frob} = \mathrm{Frob} \circ \bar{\gamma} = \mathrm{Frob} \circ \delta.$$

Since $\sigma(\mathrm{Frob}(\mathcal{C}_{p^n})) = \mathrm{Frob}(\sigma^{-1}(\mathcal{C}_{p^n})) = \mathrm{Frob}(\mathcal{C}_{p^n})$, $\mathrm{Frob}(\mathcal{C}_{p^n})$ is also stable under the action of σ. Moreover $\mathcal{C}_{p^n}$ does not intersect with the kernel of Frob.

Thus, $\mathrm{Frob}(\mathcal{C}_{p^n})$ is either $\mathcal{C}_{p^n}$ or $\mathcal{D}_{p^n}$. As an endomorphism of $\mathcal{E}$, γ sends $\mathcal{C}_{p^n}$ (resp. $\mathcal{D}_{p^n}$) to $\mathcal{C}_{p^{n-1}}$ (resp. $\mathcal{D}_{p^n}$). Similarly, δ maps $\mathcal{C}_{p^n}$ (resp. $\mathcal{D}_{p^n}$) to $\mathcal{C}_{p^n}$ (resp. $\mathcal{D}_{p^{n-1}}$). Therefore if $\mathrm{Frob}(\mathcal{C}_{p^n}) = \mathcal{C}_{p^n}$, then

$$\gamma \circ \mathrm{Frob}(\mathcal{C}_{p^n}) = \gamma(\mathcal{C}_{p^n}) = \mathcal{C}_{p^{n-1}} \quad \text{and} \quad \mathrm{Frob} \circ \delta(\mathcal{C}_{p^n}) = \mathrm{Frob}(\mathcal{C}_{p^n}) = \mathcal{C}_{p^n},$$

which is a contradiction. Thus, we get $\mathrm{Frob}(\mathcal{C}_{p^n}) = \mathcal{D}_{p^n}$.

Since every supersingular point can be defined over $\mathbb{F}_{q^2}$, the quadratic extension of $\mathbb{F}_q$, Frob^2 acts trivially on $\Sigma(N)$ (see [Ribet 1990, Remark 3.5.b]), which proves the last claim. $\square$

Remark 2.7. By taking $H = (\mathbb{Z}/N\mathbb{Z})^*$ in Lemma 1 of [Ribet 1995], we can obtain a similar result if we show that the Atkin–Lehner style involution in [Ribet 1995, §4] is equal to the Frobenius morphism.

3. The action of T_p on the component group

Before discussing the action of the Hecke operators on the component group, we study it on the group of divisors supported on supersingular points, which we denote by $\mathrm{Div}(\Sigma(N))$.

Let $N = Mp^n$ with $(M, p) = 1$ and $n \geq 1$, and assume that $(N, q) = 1$. Let $\alpha_p,\ \beta_p : X_0(Npq) \rightrightarrows X_0(Nq)$ denote two degeneracy maps of degree p, defined by

$$\alpha_p(E, C, L) := (E, C[p^n], L)$$

and

$$\beta_p(E, C, L) := (E/C[p], C/C[p], (L + C[p])/C[p]),$$

where C (resp. L) denotes a cyclic subgroup of order p^{n+1} (resp. Mq) in an elliptic curve E (see [Mazur and Ribet 1991, §13]). Let T_p and ξ_p be two Hecke correspondences defined by the following diagram:

$$\begin{tikzcd} & X_0(Npq) \arrow[dl, "\alpha_p"'] \arrow[dr, "\beta_p"] & \\ X_0(Nq) \arrow[rr, dashed, shift right, "T_p"'] & & X_0(Nq) \arrow[ll, dashed, shift right, "\xi_p"'] \end{tikzcd}$$

By pullback, the Hecke correspondence T_p (resp. ξ_p) induces the Hecke operator $T_p := \beta_{p,*} \circ \alpha_p^*$ (resp. $\xi_p := \alpha_{p,*} \circ \beta_p^*$) on $J_0(Nq)$.

The same description of the Hecke operator T_p on $\mathrm{Div}(\Sigma(N))$ as above works. In other words, we have two degeneracy maps[6] $\alpha_p, \beta_p : \Sigma(Np) \rightrightarrows \Sigma(N)$ of degree p, defined by

$$\alpha_p(E, C, L) := (E, C[p^n], L)$$

[6]Every elliptic curve isogenous to a supersingular one is also supersingular

and

$$\beta_p(E, C, L) := (E/C[p], C/C[p], (L + C[p])/C[p]),$$

where C (resp. L) denotes a cyclic subgroup of order p^{n+1} (resp. M) in a supersingular elliptic curve E over $\boldsymbol{F}$. These maps induce the maps

$$\mathrm{Div}(\Sigma(N)) \underset{\beta_p^*}{\overset{\alpha_p^*}{\rightrightarrows}} \mathrm{Div}(\Sigma(Np)) \underset{\beta_{p,*}}{\overset{\alpha_{p,*}}{\rightrightarrows}} \mathrm{Div}(\Sigma(N))$$

on their divisor groups, and the Hecke operator T_p (resp. ξ_p) can be defined by $\beta_{p,*} \circ \alpha_p^*$ (resp. $\alpha_{p,*} \circ \beta_p^*$). (For the details when $n = 0$, see [Ribet 1990, §3; 1991, pp. 18–22; Edixhoven 1991, §4.1; Emerton 2002, §7]. By the same method, we get the above description without further difficulties.)

Now, let $\Phi_q(Nq)$ denote the component group of the special fiber $\mathcal{J}$ of the Néron model of $J_0(Nq)$ at q. To compute the action of T_p on it, we closely follow the method of Ribet (see [Ribet 1988; 1990, §2, §3; Edixhoven 1991, §1]). Since N is not divisible by q, the identity component $\mathcal{J}^0$ of $\mathcal{J}$ is a semiabelian variety by Deligne and Rapoport [1973] and Raynaud [1970]. Moreover, $\mathcal{J}^0$ is an extension of $J_0(N)_{\boldsymbol{F}} \times J_0(N)_{\boldsymbol{F}}$ by $\mathcal{T}$, the torus of $\mathcal{J}^0$. Let $\mathcal{X}$ be the character group of the torus $\mathcal{T}$. By Grothendieck, there is a (Hecke-equivariant) monodromy exact sequence [SGA 7_{I} 1972] (see also [Ribet 1990, §2, §3; Raynaud 1991; Illusie 2015, §4]),

$$0 \longrightarrow \mathcal{X} \xrightarrow{\iota} \mathrm{Hom}(\mathcal{X}^t, \mathbb{Z}) \longrightarrow \Phi_q(Nq) \longrightarrow 0.$$

Here $\mathcal{X}^t$ denotes the character group corresponding to the dual abelian variety of $J_0(Nq)$, which is equal to $J_0(Nq)$. Namely, $\mathcal{X}^t = \mathcal{X}$ as sets, but the action of the Hecke operator T_ℓ on $\mathcal{X}^t$ is equal to the action of its dual ξ_ℓ on $\mathcal{X}$ (see [Ribet 1988; 1990, §3; Emerton 2002, §7]). Note that $\mathcal{X}$ is the group of degree 0 elements in $\mathbb{Z}^{\Sigma(N)}$. For $s, t \in \Sigma(N)$, let $e(s) := \frac{1}{2}\#\mathrm{Aut}(s)$ and

$$\phi_s(t) := \begin{cases} e(s) & \text{if } s = t, \\ 0 & \text{otherwise,} \end{cases}$$

and extends via linearity, i.e., $\phi_s\left(\sum a_i t_i\right) = \sum a_i \phi_s(t_i)$. Then, $\iota(s - t) = \phi_s - \phi_t$. Note also that $\mathrm{Hom}(\mathbb{Z}^{\Sigma(N)}, \mathbb{Z})$ is generated by $\psi_s := 1/e(s)\phi_s$, and $\mathrm{Hom}(\mathcal{X}^t, \mathbb{Z})$ is its quotient by the relation

$$\sum_{s \in \Sigma(N)} \psi_s = \sum_{s \in \Sigma(N)} \frac{1}{e(s)} \phi_s = 0.$$

(This is the minimal relation to make $\sum a_w \psi_w$ vanish for all the divisors of the form $s - t$, which are the generators of $\mathcal{X}$.) For more details, see [Ribet 1990, §2, §3, Raynaud 1991].

In conclusion, the component group $\Phi_q(Nq)$ is isomorphic to

$$\mathrm{Hom}(\mathbb{Z}^{\Sigma(N)}, \mathbb{Z})/R,$$

where R is the set of relations

$$R = \left\{ e(s)\psi_s = e(t)\psi_t \quad \text{for any } s, t \in \Sigma(N), \sum_{t \in \Sigma(N)} \psi_t = 0 \right\}. \tag{3-1}$$

Let Ψ_s denote the image of ψ_s by the natural projection $\mathrm{Hom}(\mathbb{Z}^{\Sigma(N)}, \mathbb{Z}) \to \Phi_q(Nq)$. The Hecke operator T_p acts on $\mathrm{Hom}(\mathbb{Z}^{\Sigma(N)}, \mathbb{Z})$ via the action of ξ_p on $\mathrm{Div}(\Sigma(N))$, i.e.,

$$T_p(\psi_s)(t) := \psi_s(\xi_p(t)) = \psi_s(\alpha_{p,*} \circ \beta_p^*(t)).$$

For $s \in \Sigma(N)$, we temporarily denote $\alpha_p^*(s) = \sum_{i=1}^p A^i(s)$ and $\beta_p^*(s) = \sum_{i=1}^p B^i(s)$ (allowing repetition). We note that if $e(s) = 1$ then there is no repetition, i.e., $A^i(s) \not\simeq A^j(s)$ and $B^i(s) \not\simeq B^j(s)$ if $i \neq j$. If $e(s) = e > 1$, then after renumbering the index properly we have

$$e(A^i(s)) = 1 \quad \text{for } 1 \le i \le p-1 \text{ and } e(A^p(s)) = e.$$

Moreover, we have

$$A^{e(k-1)+1}(s) \simeq \cdots \simeq A^{ek}(s) \quad \text{for } 1 \le k \le \frac{p-1}{e},$$

and

$$A^i(s) \not\simeq A^j(s) \quad \text{if } \left[\frac{i-1}{e}\right] \neq \left[\frac{j-1}{e}\right],$$

where $[x]$ denotes the largest integer less than or equal to x. This can be seen as follows: Let $\sigma = \sigma_{2e}$, and let s represent a pair $(\mathcal{E}, C)$, where C is a cyclic subgroup of E of order N. Since $e(s) = e$, $\sigma(C) = C$. Suppose that $s' \in \Sigma(Np)$ with $\alpha_{p,*}(s') = s$. Then s' represents a pair $(\mathcal{E}, D)$ with $D[N] = C$. If $\sigma(D) = D$, then $\mathrm{Aut}([(\mathcal{E}, D)]) = \langle\sigma\rangle$ and $(\mathcal{E}, D) \not\simeq (\mathcal{E}, D')$ if $D \neq D'$. (Note that there is a unique such D.) On the other hand, if $\sigma(D) \neq D$ then

$$(\mathcal{E}, D) \simeq (\mathcal{E}, \sigma(D)) \simeq \cdots \simeq (\mathcal{E}, \sigma^{e-1}(D)) \simeq (\mathcal{E}, \sigma^e(D)) = (\mathcal{E}, D)$$

and $\mathrm{Aut}([(\mathcal{E}, D)]) = \{\pm 1\}$. Thus, we can rearrange $A^i(s)$ as above. (Note that this can only be possible when $p \equiv 1 \pmod{2e}$, which is true because $e(s) = e$.)

Now, we claim that $\phi_s(\alpha_{p,*}(t)) = \phi_t(\alpha_p^*(s))$. Indeed, $\phi_s(\alpha_{p,*}(t))$ is nonzero if and only if $t \in \{A^1(s), \dots, A^p(s)\}$. So, it suffices to show this equality when $t \in \{A^1(s), \dots, A^p(s)\}$. If $e(s) = 1$, then there is no repetition and the claim follows clearly (both are 1). Now, let $e(s) = e > 1$. If $e(t) = 1$, then $t = A^i(s)$ for some $1 \le i \le p-1$. Since the number of repetitions of $t = A^i(s)$ in $\{A^1(s), \dots, A^p(s)\}$ is e, the above equality holds. If $e(t)=e$, then $t=A^p(s)$ and $\phi_s(\alpha_{p,*}(t))=e=\phi_t(\alpha_p^*(s))$, as claimed. Analogously, we have

$$\phi_t(\beta_{p,*}(s))=\phi_s(\beta_p^*(t)).$$

More generally, we get

$$\begin{aligned}\phi_s(\alpha_{p,*}\circ\beta_p^*(t)) &= \sum_{i=1}^{p}\phi_s(\alpha_{p,*}(B^i(t))) = \sum_{i=1}^{p}\sum_{j=1}^{p}\phi_{B^i(t)}(A^j(s))\\ &= \sum_{j=1}^{p}\sum_{i=1}^{p}\phi_{A^j(s)}(B^i(t)) = \sum_{j=1}^{p}\phi_{A^j(s)}(\beta_p^*(t))\\ &= \sum_{j=1}^{p}\phi_t(\beta_{p,*}(A^j(s))) = \phi_t(\beta_{p,*}\circ\alpha_p^*(s)) = \phi_t(T_p(s)).\end{aligned}$$

If we set $T_p(s)=\sum s_j$, then $\phi_t(T_p(s))=\sum\phi_{s_i}(t)=\sum e(s_i)\psi_{s_i}(t)$ and hence for any $t\in\Sigma(N)$,

$$e(s)T_p(\psi_s)(t)=\phi_s(\alpha_{p,*}\circ\beta_p^*(t))=\phi_t(T_p(s))=e(s_i)\psi_{s_i}(t).$$

In other words, we get

$$T_p(\Psi_s)=\frac{1}{e(s)}\sum e(s_i)\Psi_{s_i}. \tag{3-2}$$

We can also define the action of T_p on the component group via functorialities. Namely, let

$$\Phi_q(Nq)\ \underset{\beta_p^*}{\overset{\alpha_p^*}{\rightrightarrows}}\ \Phi_q(Npq)\ \underset{\beta_{p,*}}{\overset{\alpha_{p,*}}{\rightrightarrows}}\ \Phi_q(Nq)$$

denote the maps functorially induced from the degeneracy maps.[7] Then, as before, $T_p:=\beta_{p,*}\circ\alpha_p^*$. Note that since the degrees of α_p and β_p are p, we have $\alpha_{p,*}\circ\alpha_p^*=\beta_{p,*}\circ\beta_p^*=p$.

Lemma 3.1. *The operator $\alpha_{p,*}$ is equal to $\beta_{p,*}$ on $\Phi_q(Npq)$.*

Proof. For $s\in\Sigma_{2e}(Npq)$ with $e=2$ or 3, $\alpha_p(s)=\beta_p(s)$ by Corollary 2.5, and hence $\alpha_{p,*}(\Psi_s)=\beta_{p,*}(\Psi_s)$. For $s\in\Sigma_2(Npq)$, let $\alpha_p(s)=t$ and $\beta_p(s)=w$. Then, $\alpha_{p,*}(\Psi_s)=e(t)\Psi_t=e(w)\Psi_w=\beta_{p,*}(\Psi_s)$. In other words, for any $s\in\Sigma(Npq)$, $\alpha_{p,*}(\Psi_s)=\beta_{p,*}(\Psi_s)$. Since Ψ_s's generate $\Phi_q(Npq)$, the result follows. □

In fact, Theorem 1.1 is an easy corollary of the above lemma.

[7] If $\alpha_p^*(s)=\sum t_j$ then $\alpha_p^*(\Psi_s)=\sum\Psi_{t_j}$ and if $\alpha_p(t)=s$ then $\alpha_{p,*}(\Psi_t)=e(s)/e(t)\Psi_s$; and similarly for β_p^* and $\beta_{p,*}$.

Proof of Theorem 1.1. Since $\alpha_{p,*} = \beta_{p,*}$ on $\Phi_q(Npq)$, we have

$$T_p(\Psi_s) = \beta_{p,*} \circ \alpha_p^*(\Psi_s) = \alpha_{p,*} \circ \alpha_p^*(\Psi_s) = p\Psi_s,$$

which implies the result. □

4. The action of T_q on the component group

In this section, we provide a complete description of the action of T_q on the component group $\Phi_q(Nq)$. See Propositions 4.2, 4.3 and 4.4, which imply Theorem 1.2.

Note that the Hecke operator T_q acts on $\Sigma(N)$ by the Frobenius morphism [Ribet 1990, Proposition 3.8], and the same is true for ξ_q. Since the Frobenius morphism is an involution on $\Sigma(N)$ (see Proposition 2.6), we have

$$T_q(\psi_s)(t) = \psi_s(\xi_q(t)) = \psi_s(\mathrm{Frob}(t)) = \psi_{\mathrm{Frob}(s)}(t) \quad \text{for any} \quad t \in \Sigma(N), \tag{4-1}$$

which implies that $T_q(\psi_s) = \psi_{\mathrm{Frob}(s)}$.

From now on, if there is no confusion we remove (N) from the notation for simplicity. Let $n := \frac{1}{12}(q-1)Q$ (which is not necessarily an integer), and let Φ denote the cyclic subgroup of $\Phi_q(Nq)$ generated by $\Psi_{\mathfrak{s}}$ for a fixed $\mathfrak{s} \in \Sigma_2$. (Note that this Φ is the same as that of Mazur and Rapoport [1977], namely, Φ is equal to the cyclic subgroup generated by the image of the cuspidal divisor $(0) - (\infty)$.)

***Case* 1: $(u, v) = (0, 0)$ *or* $v = 0$.** Let $e = 1$ if $(u, v) = (0, 0)$ and $e = 2u + 3v$ if $(u, v) \neq (0, 0)$ and $v = 0$. If $(u, v) = (0, 0)$, $s_2 = n$ and $s_4 = s_6 = 0$. If $(u, v) \neq (0, 0)$ and $v = 0$, then $s_{2e} = 1$ and $s_2 = \frac{1}{e}(en - 1)$. (Note that s_2 is an integer but n is not.)

Proposition 4.1. *The component group $\Phi_q(Nq)$ is equal to Φ, which is cyclic of order en. The Hecke operator T_q acts on it by* 1.

Proof. First, we assume that $(u, v) = (0, 0)$. Then $\Psi_s = \Psi_{\mathfrak{s}}$ for any $s \in \Sigma = \Sigma_2$. Therefore $\Phi_q(Nq) = \Phi$ and $n\Psi_{\mathfrak{s}} = \sum_{s \in \Sigma} \Psi_s = 0$. Moreover, $T_q(\Psi_{\mathfrak{s}}) = \Psi_{s'} = \Psi_{\mathfrak{s}}$, where $s' = \mathrm{Frob}(\mathfrak{s})$.

Now, we assume that $(u, v) \neq (0, 0)$ and $v = 0$. In this case, either $N = 2q$ (with $(u, v) = (1, 0)$ and $e = 2$) or $N = 3q$ (with $(u, v) = (0, 1)$ and $e = 3$). In each case, let $z \in \Sigma_{2e}$. Then

$$\sum_{s \in \Sigma_2} \Psi_s + \Psi_z = s_2\Psi_{\mathfrak{s}} + \Psi_z = 0 \quad \text{and} \quad \Psi_{\mathfrak{s}} = e\Psi_z.$$

Therefore the component group is generated by Ψ_z, and its order is $(es_2 + 1) = en$. Since $en = es_2 + 1$ is prime to e, this group is also generated by $\Psi_{\mathfrak{s}} = e\Psi_z$. (In fact, $\Psi_z = -s_2\Psi_{\mathfrak{s}}$.) Moreover we have $T_q(\Psi_{\mathfrak{s}}) = \Psi_{\mathfrak{s}}$ as above. □

***Case* 2: $(u, v) = (0, 1)$ *and* $v \geq 1$.** In this case, $s_4 = 0$, $s_6 = 2^\nu$, and $s_2 = \frac{1}{3}(3n - 2^\nu)$. Let $\Sigma_6 := \{t_1, t_2, \ldots, t_{2^\nu}\}$. Here we assume that $\mathrm{Frob}(t_{2k-1}) = t_{2k}$ for $1 \leq k \leq 2^{\nu-1}$.[8] Let $t := t_{2^\nu-1}$ and $t' := t_{2^\nu}$.

Proposition 4.2. *The component group $\Phi_q(Nq)$ decomposes as follows*:

$$\Phi_q(Nq) = \bigoplus_{i=0}^{2^\nu-1} B_i =: B_0 \oplus \boldsymbol{B},$$

where $B_0 = \Phi$ is cyclic of order $3n$, and for $1 \leq i \leq 2^\nu - 1$, B_i is cyclic of order 3. For $1 \leq k \leq 2^{\nu-1}$, B_{2k-1} and B_{2k} are generated by

$$\boldsymbol{v}_{2k-1} := \Psi_{t_{2k-1}} - \Psi_{t_{2k}}$$

and

$$\boldsymbol{v}_{2k} := \Psi_{t_{2k-1}} + \Psi_{t_{2k}} - \Psi_t - \Psi_{t'},$$

respectively. The Hecke operator T_q acts on B_i by $(-1)^i$.

Proof. Note that $\Psi_s = 3\Psi_{t_i} = 3\Psi_{t_j}$ for all i, j and $\sum_{i=1}^{2^\nu} \Psi_{t_i} + s_2\Psi_s = 0$. Therefore $\Phi_q(Nq)$ is generated by Ψ_{t_i} for $1 \leq i \leq 2^\nu - 1$. The order of each group $\langle\Psi_{t_i}\rangle$ is $9n$ because

$$9n\Psi_{t_i} = 3s_2(3\Psi_{t_i}) + \sum_{i=1}^{2^\nu} 3\Psi_{t_i} = 3\Bigg(\sum_{s\in\Sigma_2} \Psi_s + \sum_{i=1}^{2^\nu} \Psi_{t_i}\Bigg) = 0,$$

and $9n$ is the smallest positive integer to make this happen. Moreover $\langle\Psi_{t_i}\rangle \cap \langle\Psi_{t_j}\rangle$ is of order $3n$ for any $i \neq j$. Since $3n = 3s_2 + 2^\nu$ is prime to 3, we can decompose the component group into

$$\langle 3\Psi_t\rangle \oplus \langle(3s_2 + 2^\nu)\Psi_t\rangle \bigoplus_{i=1}^{2^\nu-2} \langle\Psi_{t_i} - \Psi_t\rangle. \tag{4-2}$$

Since $\Psi_s = 3\Psi_{t_i} = 3\Psi_t = 3\Psi_{t'}$ for any i and

$$\sum_{i=1}^{2^\nu} \Psi_{t_i} = -3s_2\Psi_t,$$

we have

$$\begin{aligned}
\Psi_{2k-1} - \Psi_t &= 2\boldsymbol{v}_{2k-1} + 2\boldsymbol{v}_{2k} + \boldsymbol{v}_{2^\nu-1},\\
\Psi_{2k} - \Psi_t &= \boldsymbol{v}_{2k-1} + 2\boldsymbol{v}_{2k} + \boldsymbol{v}_{2^\nu-1},\\
(3s_2 + 2^\nu)\Psi_t &= \sum_{i=1}^{2^\nu} (\Psi_t - \Psi_{t_i}) = -\sum_{k=1}^{2^{\nu-1}} \boldsymbol{v}_{2k} - (-1)^\nu \boldsymbol{v}_{2^\nu-1}.
\end{aligned}$$

Therefore the decomposition in the proposition is isomorphic to (4-2). The action of T_q on each B_i is obvious from its construction. □

[8] By Proposition 2.6, we know that Frob is an involution of Σ_6 without fixed points.

***Case* 3: $(u, v) = (1, 0)$ *and* $\nu \geq 1$.** Note that $s_4 = 2^\nu$, $s_6 = 0$, and $s_2 = n - 2^{\nu-1}$. Let $\Sigma_4 = \{w_1, w_2, \dots, w_{2^\nu}\}$. As before, we assume that $\mathrm{Frob}(w_{2k-1}) = w_{2k}$ for $1 \leq k \leq 2^{\nu-1}$.[9] Let $w := w_{2^\nu-1}$ and $w' := w_{2^\nu}$.

Proposition 4.3. *The component group $\Phi_q(Nq)$ decomposes as*

$$\Phi_q(Nq) = \bigoplus_{i=0}^{2^\nu-2} A_i = A_0 \oplus \boldsymbol{A},$$

where A_0 is cyclic of order $4n$ generated by Ψ_w, and for $1 \leq i \leq 2^\nu - 2$, A_i is cyclic of order 2. For $1 \leq k \leq 2^{\nu-1} - 2$, A_{2k-1} and A_{2k} are generated by

$$\boldsymbol{u}_{2k-1} := \Psi_{w_{2k-1}} - \Psi_w \quad \text{and} \quad \boldsymbol{u}_{2k} := \Psi_{w_{2k-1}} + \Psi_{w_{2k}} - \Psi_w - \Psi_{w'}, \quad \text{respectively.}$$

And $A_{2^\nu-3}$ and $A_{2^\nu-2}$ are generated by

$$\boldsymbol{u}_{2^\nu-3} := \Psi_{w_{2^\nu-3}} - \Psi_w \quad \text{and} \quad \boldsymbol{u}_{2^\nu-2} := \Psi_{w_{2^\nu-3}} - \Psi_{w_{2^\nu-2}}, \quad \text{respectively.}$$

Moreover, the action of the Hecke operator T_q on each group is as follows:

$$T_q(\Psi_w) = (1+2n)\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i},$$

$$T_q(\boldsymbol{u}_{2k-1}) = \boldsymbol{u}_{2k-1} + \boldsymbol{u}_{2k} \quad \text{and} \quad T_q(\boldsymbol{u}_{2k}) = \boldsymbol{u}_{2k} \quad \text{for } 1 \leq k \leq 2^{\nu-1} - 2,$$

$$T_q(\boldsymbol{u}_{2^\nu-3}) = 2n\Psi_w + \boldsymbol{u}_{2^\nu-3} + \sum_{i=1}^{2^{\nu-1}-2} \boldsymbol{u}_{2i} \quad \text{and} \quad T_q(\boldsymbol{u}_{2^\nu-2}) = \boldsymbol{u}_{2^\nu-2}.$$

Proof. The argument in Proposition 4.2 applies *mutatis mutandis*. For instance, when $\nu \geq 2$ an isomorphism between $A_0 \bigoplus_{i=1}^{2^\nu-2} \langle \Psi_{w_i} - \Psi_w \rangle$ and $A_0 \oplus \boldsymbol{A}$ can be given as follows: for $1 \leq k \leq 2^{\nu-1} - 2$,

$$\Psi_{w_{2k}} - \Psi_w = \boldsymbol{u}_{2k} + \boldsymbol{u}_{2k-1} + (\Psi_{w'} - \Psi_w),$$

$$\Psi_w - \Psi_{w'} = 2n\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i},$$

$$\Psi_{w_{2^\nu-2}} - \Psi_w = \boldsymbol{u}_{2^\nu-3} + \boldsymbol{u}_{2^\nu-2}.$$

The action of the Hecke operator T_q on each A_i is clear except

$$T_q(\Psi_w) = \Psi_{w'} = \Psi_w - (\Psi_w - \Psi_{w'}) = (1+2n)\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i},$$

$$\begin{aligned} T_q(\boldsymbol{u}_{2^\nu-3}) = \Psi_{w_{2^\nu-2}} - \Psi_{w'} &= \boldsymbol{u}_{2^\nu-3} + \boldsymbol{u}_{2^\nu-2} + (\Psi_w - \Psi_{w'}) \\ &= 2n\Psi_w + \boldsymbol{u}_{2^\nu-3} + \sum_{i=1}^{2^{\nu-1}-2} \boldsymbol{u}_{2i}. \quad \square \end{aligned}$$

[9]By Proposition 2.6, we know that Frob is an involution of Σ_4 without fixed points.

***Case* 4: $(\boldsymbol{u}, \boldsymbol{v}) = (1, 1)$ *and* $\nu \geq 1$.** Note that $s_4 = s_6 = 2^\nu$ and $s_2 = \frac{1}{6}(6n - 5 \cdot 2^\nu)$. Let $\Sigma_4 = \{w_1, \ldots, w_{2^\nu}\}$ and $\Sigma_6 := \{t_1, \ldots, t_{2^\nu}\}$. As before, we assume that $\mathrm{Frob}(w_{2k-1}) = w_{2k}$ and $\mathrm{Frob}(t_{2k-1}) = t_{2k}$ for $1 \leq k \leq 2^{\nu-1}$. Let $w := w_{2^\nu - 1}$ and $w' := w_{2^\nu}$. Also, let $t := t_{2^\nu - 1}$ and $t' := t_{2^\nu}$.

Proposition 4.4. *The component group $\Phi_q(Nq)$ decomposes as*

$$\Phi_q(Nq) = A_0 \oplus \boldsymbol{A} \oplus \boldsymbol{B},$$

where A_0 is cyclic of order $12n$ generated by Ψ_w. The structures of $\boldsymbol{A}$ and $\boldsymbol{B}$ are the same as those in Propositions 4.2 and 4.3. The actions of T_q on $\boldsymbol{A}$ and $\boldsymbol{B}$ are the same as before except on $A_{2^\nu - 3}$ (when $\nu \geq 2$), where T_q acts by

$$T_q(\boldsymbol{u}_{2^\nu - 3}) = 6n\Psi_w + \boldsymbol{u}_{2^\nu - 3} + \sum_{i=1}^{2^{\nu-1}-2} \boldsymbol{u}_{2i}.$$

Moreover, the action of T_q on A_0 is analogous to the previous case:

$$T_q(\Psi_w) = (1 + 6n)\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i}.$$

Proof. Note that from (3-1), we have

$$s_2\Psi_s + \Psi_{w_1} + \cdots + \Psi_{w'} + \Psi_{t_1} + \cdots + \Psi_{t'} = 0.$$

Multiplying by 3, we have

$$\Psi_{w_1} + \cdots + \Psi_{w'} = -(3s_2 + 2 \cdot 2^\nu)\Psi_s = -(6s_2 + 4 \cdot 2^\nu)\Psi_w. \tag{4-3}$$

Also, multiplying by 4, we have

$$\Psi_{t_1} + \cdots + \Psi_{t'} = -(4s_2 + 3 \cdot 2^\nu)\Psi_s = -(12s_2 + 9 \cdot 2^\nu)\Psi_t. \tag{4-4}$$

Therefore $\Psi_{w_1}, \ldots, \Psi_w, \Psi_{t_1}, \ldots, \Psi_t$ can generate the whole group. By a similar computation, the order of $\langle \Psi_{w_i} \rangle$ is $12n$ and the order of $\langle \Psi_{t_i} \rangle$ is $18n$. All of them contain Φ as a subgroup, which is of order $6n$. Here we note that $\langle \Psi_t \rangle = \langle 3\Psi_t \rangle \oplus \langle 6n\Psi_t \rangle$ because $6n = 6s_2 + 5 \cdot 2^\nu$ is prime to 3. Therefore we can decompose $\Phi_q(Nq)$ into

$$\langle \Psi_w \rangle \bigoplus_{i=1}^{2^\nu - 2} \langle \Psi_{w_i} - \Psi_w \rangle \bigoplus_{i=1}^{2^\nu - 2} \langle \Psi_{t_i} - \Psi_t \rangle \bigoplus \langle 6n\Psi_t \rangle. \tag{4-5}$$

As in Propositions 4.2 and 4.3, we can find an isomorphism between (4-5) and $A_0 \oplus \boldsymbol{A} \oplus \boldsymbol{B}$, which proves the first part. From (4-3) (and the previous discussions) we have

$$\Psi_w - \Psi_{w'} = (6s_2 + 5 \cdot 2^\nu)\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i} = 6n\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i}.$$

The action of T_q on each component is also obvious except

$$T_q(\Psi_w) = \Psi_{w'} = \Psi_w - (\Psi_w - \Psi_{w'}) = (1+6n)\Psi_w + \sum_{i=1}^{2^{\nu-1}-1} \boldsymbol{u}_{2i},$$

$$\begin{aligned} T_q(\boldsymbol{u}_{2^\nu-3}) = \Psi_{w_{2^\nu-2}} - \Psi_{w'} = \boldsymbol{u}_{2^\nu-3} + \boldsymbol{u}_{2^\nu-2} + (\Psi_w - \Psi_{w'}) \\ = 6n\Psi_w + \boldsymbol{u}_{2^\nu-3} + \sum_{i=1}^{2^{\nu-1}-2} \boldsymbol{u}_{2i}. \quad \square \end{aligned}$$

Acknowledgements

Yoo would like to thank Kenneth Ribet for a number of very helpful discussions about Eisenstein ideals and component groups. The referee deserves special thanks for a thorough reading of the manuscript and for many useful comments and suggestions. This work was supported by IBS-R003-D1.

References

[Deligne and Rapoport 1973] P. Deligne and M. Rapoport, "Les schémas de modules de courbes elliptiques", pp. 143–316 in *Modular functions of one variable, II* (Antwerp, 1972), edited by P. Deligne and W. Kuyk, Lecture Notes in Math. **349**, Springer, 1973. MR Zbl

[Edixhoven 1991] B. Edixhoven, "L'action de l'algèbre de Hecke sur les groupes de composantes des jacobiennes des courbes modulaires est 'Eisenstein'", pp. 159–170 in *Courbes modulaires et courbes de Shimura* (Orsay, 1987/1988), Astérisque **196-197**, Société Mathématique de France, Paris, 1991. MR Zbl

[Emerton 2002] M. Emerton, "Supersingular elliptic curves, theta series and weight two modular forms", *J. Amer. Math. Soc.* **15**:3 (2002), 671–714. MR Zbl

[Illusie 2015] L. Illusie, "Grothendieck and vanishing cycles", preprint, 2015, Available at https://tinyurl.com/lucillus.

[Katz and Mazur 1985] N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Math. Studies **108**, Princeton Univ. Press, 1985. MR Zbl

[Mazur 1977] B. Mazur, "Modular curves and the Eisenstein ideal", *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 33–186. MR Zbl

[Mazur and Rapoport 1977] B. Mazur and M. Rapoport, "Behavior of the Néron model of the Jacobian of $X_0(N)$ at bad primes", *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 173–186. Appendix to [Mazur 1977].

[Mazur and Ribet 1991] B. Mazur and K. A. Ribet, "Two-dimensional representations in the arithmetic of modular curves", pp. 215–255 in *Courbes modulaires et courbes de Shimura* (Orsay, 1987/1988), Astérisque **196-197**, Société Mathématique de France, Paris, 1991. MR Zbl

[Raynaud 1970] M. Raynaud, "Spécialisation du foncteur de Picard", *Inst. Hautes Études Sci. Publ. Math.* **38** (1970), 27–76. MR Zbl

[Raynaud 1991] M. Raynaud, "Jacobienne des courbes modulaires et opérateurs de Hecke", pp. 9–25 in *Courbes modulaires et courbes de Shimura* (Orsay, 1987/1988), Astérisque **196-197**, Société Mathématique de France, Paris, 1991. MR Zbl

[Ribet 1988] K. A. Ribet, "On the component groups and the Shimura subgroup of $J_0(N)$", exp. 6, 10 pp. in *Séminaire de Théorie des Nombres* (Talence, France, 1987/1988), Univ. Bordeaux I, Talence, 1988. MR Zbl

[Ribet 1990] K. A. Ribet, "On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms", *Invent. Math.* **100**:2 (1990), 431–476. MR Zbl

[Ribet 1995] K. A. Ribet, "Irreducible Galois representations arising from component groups of Jacobians", pp. 131–147 in *Elliptic curves, modular forms, & Fermat's last theorem* (Hong Kong, 1993), edited by J. Coates, Ser. Number Theory **1**, Int. Press, Cambridge, 1995. MR Zbl

[Serre and Tate 1968] J.-P. Serre and J. Tate, "Good reduction of abelian varieties", *Ann. of Math.* (2) **88** (1968), 492–517. MR Zbl

[SGA 7_{I} 1972] A. Grothendieck, "Modèles de Néron et monodromie", exp. IX, 313–523 in *Groupes de monodromie en géométrie algébrique, I* (Séminaire de Géométrie Algébrique du Bois Marie 1967-1969), edited by A. Grothendieck, Lecture Notes in Math. **288**, Springer, 1972. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Math. **106**, Springer, 2009. MR Zbl

[Yoo 2014] H. Yoo, "Non-optimal levels of a reducible mod ℓ modular representation", 2014. To appear in *Trans. Amer. Math. Soc.* arXiv

[Yoo 2016] H. Yoo, "On Eisenstein ideals and the cuspidal group of $J_0(N)$", *Israel J. Math.* **214**:1 (2016), 359–377. MR Zbl

Received October 23, 2017. Revised February 5, 2018.

TAEKYUNG KIM
CENTER FOR GEOMETRY AND PHYSICS
INSTITUTE FOR BASIC SCIENCE
POHANG
SOUTH KOREA
taekyung.kim.maths@gmail.com

HWAJONG YOO
CENTER FOR GEOMETRY AND PHYSICS
INSTITUTE FOR BASIC SCIENCE
POHANG
SOUTH KOREA
hwajong@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.357

LIOUVILLE THEOREMS, VOLUME GROWTH, AND VOLUME COMPARISON FOR RICCI SHRINKERS

LI MA

We study volume growth, a Liouville theorem for f-harmonic functions, and a volume comparison property of unit balls in complete noncompact gradient Ricci shrinkers and gradient steady Ricci solitons. We also study integral properties of f-harmonic functions and harmonic functions on complete manifolds, such as the Ricci–Einstein solitons.

1. Introduction

In the study of Ricci flow on a compact Riemannian manifold, because of its complicated nonlinearity, one meets singularities of the flow in finite time. After blowing up, one expects to get self-similar Ricci shrinkers or steady Ricci solitons [Hamilton 1995]. In the case of a type-I singularity in dimension three, one gets nontrivial gradient Ricci shrinkers via the use of the Ivey–Hamilton pinching estimate, and the classification of this type of self-similar Ricci shrinker was done by G. Perelman [2000]. In the case of a type-I singularity of dimension four, A. Naber [2010] showed that one gets a gradient Ricci shrinker, and nontrivial properties of this Ricci shrinker have been studied by others. These solitons can be considered as special examples of weighted Riemannian manifolds or metric measure spaces [Bakry and Émery 1985; Wei and Wylie 2009; Lott and Villani 2009; Chen 2009; Chow et al. 2006; Lichnerowicz 1970; Lott 2003; Ma 2013; Sturm 2006a; 2006b; Yang 2009; Munteanu and Wang 2011; 2014; 2015]. Because of the importance of four-dimensional Ricci shrinkers, many people study various properties of them; see for example [Ma 2013; Cao and Zhou 2010; Cao 2007; 2010; Carrillo and Ni 2009; Munteanu and Wang 2015; Haslhofer and Müller 2011]. In this paper, we examine three questions about Ricci shrinking and steady solitons, in particular for Ricci shrinkers which are the complete Riemannian manifold (M, g) such that $\mathrm{Ric}_f := \mathrm{Ric} + D^2 f = \lambda g$ on M, where Ric is the Ricci curvature of (M, g), $f : M \to R$ is a smooth function in M, and $\lambda > 0$ is a constant. One question under

This research is partially supported by the National Natural Science Foundation of China, grant nos. 11771124 and 11271111.
MSC2010: 53C21, 53C44, 53C25.
Keywords: Ricci shrinkers, f-harmonic functions, Liouville theorem, volume comparison.

consideration in this paper is about the volume comparison of unit balls on Ricci shrinkers. The other two are about f-harmonic functions and harmonic functions with finite energy. The volume comparison of unit balls is an important step to understanding the volume growth of geodesic balls in gradient Ricci shrinkers. The Liouville theorems for f-harmonic and harmonic functions with finite energy are important to understanding the connectivity at infinity about gradient Ricci solitons; see [Munteanu and Wang 2015].

We have the following new results. The first is the volume comparison of unit balls at any point $x \in M$ with a unit ball at a fixed point $p \in M$, which deals with the injectivity radius decay from the point p to the point x and is important to understanding the topology of the underlying manifold at infinity.

Theorem 1. *On the complete noncompact gradient Ricci shrinker (M, g, f) with Ricci curvature bounded by $-(n-1)k^2$ for some constant $k \geq 0$, we have*

$$\mathrm{Vol}(B_x(1)) \geq \exp(-\sqrt{c_0(n-1)R})\,\mathrm{Vol}(B_p(1)),$$

where c_0 is a uniform constant which does not depend on x and $R = d(p, x) > R_0$ for some uniform constant $R_0 > 1$.

The above result is motivated by the work of Ovidiu Munteanu and Jiaping Wang [2014] and the result is not sharp, as pointed out by the referee. In fact, if the Ricci curvature is bounded from below, the logarithmic Sobolev inequality of [Carrillo and Ni 2009] implies a uniform lower bound of the volume of unit balls, $\mathrm{Vol}(B_x(1)) \geq \frac{1}{c}\mathrm{Vol}(B_p(1))$. However, from our argument, our result is true on a complete smooth measure space of dimension n with $\mathrm{Ric}_f \geq \frac{1}{2}g$ and $|\nabla f|^2 \leq f$ and with Ricci curvature bounded below by $-(n-1)k^2$ for some constant $k \geq 0$ on M. Notice that with the extra bound on Ricci curvature, our result improves the result of Lemma 5.2 in [Munteanu and Wang 2014]. We remark that the above result is still true for steady Ricci solitons. See Theorem 5 in Section 2.

We want to understand the topology and geometry from properties of (super- or sub-) harmonic functions and f-harmonic functions on Ricci solitons. By a well-known argument, we know that there is no nontrivial positive f-harmonic function on a gradient Ricci shrinker. In fact there is no nonconstant positive f-superharmonic function u ($\Delta_f u := \Delta u - \nabla f.\nabla u \leq 0$) on the complete Riemannian manifold (M, g) with $\mathrm{Ric}_f \geq \frac{1}{2}g$ on M. The process of proving this is below. Recall that the weighted volume of (M, g) is finite [Morgan 2005]; i.e., $V_f(M) := \int_M e^{-f}\,dv_g < \infty$. Assume $u > 0$ is such a positive f-superharmonic function on M. Let $v = \log u$. Then we have

$$\Delta_f v = \frac{\Delta_f u}{u} - |\nabla v|^2 \leq -|\nabla v|^2.$$

Thus, for any cut-off function $\phi \geq 0$ on the ball $B_p(2R) \subset M$ with $\phi = 1$ in $B_p(R)$, we have

$$\int_M |\nabla v|^2 \phi^2 e^{-f}\, dv_g \leq 2 \int_M \phi \nabla\phi . \nabla v e^{-f}\, dv_g.$$

By the Cauchy–Schwarz inequality we get

$$\int_M |\nabla v|^2 \phi^2 e^{-f}\, dv_g \leq 4 \int_M |\nabla\phi|^2 e^{-f}\, dv_g \leq \frac{16}{R^2} V_f(M) \to 0$$

as $R \to \infty$. Hence, $|\nabla v|^2 = 0$ in M, which implies that u is a constant on M.

Although this result is known to experts, we cannot find it in the literature. We formulate it as below.

Proposition 2. *Let (M, g, f) be a complete Riemannian manifold (M, g) with potential function f satisfying $\mathrm{Ric}_f \geq \frac{1}{2} g$ on M. Then there is no nonconstant positive f-superharmonic function on (M, g).*

We are now trying to find another kind of Liouville theorem for an f-harmonic function with weighted finite energy on a gradient Ricci shrinker. We show that as a direct consequence of a Bochner-type formula, see [Ma and Du 2010], we have the following Liouville-type theorem.

Theorem 3. *Let (M, g) be a complete noncompact Riemannian manifold such that $\mathrm{Ric}_f \geq h(x)g$ for some potential functions $f(x)$ and some nontrivial nonnegative function $h(x)$ in M. There is no nontrivial f-harmonic function u defined in (M, g) with weighted finite energy*; *i.e.*,

$$\int_M |\nabla u|^2 e^{-f}\, dv_g < \infty.$$

The proof of this result is given in Section 3. We remark that when one studies self-similar solutions to the Ricci–Einstein-type flow

$$\partial_t g = -2\,\mathrm{Rc}(g) + \gamma R g,$$

where γ is a physical constant and $R = R(g)$ is the scalar curvature of the manifold (M, g), one gets the soliton equation of it in the form

$$\mathrm{Ric}_f = \gamma R g + \lambda g,$$

which gives what may be called Ricci–Einstein solitons. Theorem 3 can be used for these solitons.

Proposition 4. *Fix any $p \in M$. Assume that the complete noncompact Riemannian manifold (M, g) satisfies $\mathrm{Ric}_f \geq h(x)g$ for some potential functions $f(x)$ and some smooth function $h(x)$ with*

$$|\nabla f(x)| \leq \alpha d(x, p) + b$$

for some uniform constants $\alpha \geq 0$ and $b > 0$ in M. Then for any harmonic function u with finite energy, i.e.,

$$\int_M |\nabla u|^2 < \infty,$$

we have the integral inequality

$$\int_M |\nabla^2 u|^2 + \int_M h(x)|\nabla u|^2 \leq \frac{1}{2}\int_M \Delta f |\nabla u|^2.$$

As a consequence, when $\Delta f(x) \leq 2h(x)$ on M, we have $D^2 u = 0$ in M; i.e., ∇u is a parallel vector field on M.

We now give a few remarks related to this result.

(1) Naber [2010] proved that for the weighted smooth metric space (M, g, f) satisfying $\mathrm{Ric}_f \leq \frac{1}{2}g$ and $|\mathrm{Ric}| \leq C$, there exists $\alpha > 0$ such that if $\Delta_f u := \Delta u - \nabla f.\nabla u = 0$ on M with $|u(x)| \leq A \exp(\alpha d(x, p)^2)$ for some $A > 0$ and $p \in M$, then u is a constant.

(2) Munteanu and Sesum [2013] proved that for gradient shrinking Kähler–Ricci solitons, if the harmonic function u has finite energy, i.e., $\int_M |\nabla u|^2 < \infty$, then u is a constant. As a consequence of this result, they showed that such a manifold has at most one nonparabolic end; see [Munteanu and Sesum 2013] for the definition of nonparabolic end. In the earlier work [Munteanu and Wang 2011], they proved that on a weighted smooth metric space (M, g, f) satisfying $\mathrm{Ric}_f \geq 0$ and f is a bounded function, any sublinear growth f-harmonic function on M must be a constant.

(3) Some consequences of Proposition 4 are given in Section 3.

Here is the plan of the paper. We study the volume comparison of unit balls of gradient Ricci shrinkers and gradient steady Ricci solitons in Section 2. We prove Proposition 2 and Theorem 3 in Section 3. In the last section, we consider integral properties for harmonic functions on Ricci solitons and prove Proposition 4.

2. Volume comparison of unit balls

We now prove Theorems 1 and 5. We use an idea similar to the proof of Theorem 2.3 in [Munteanu and Wang 2014].

We first give a proof of an improved volume comparison of unit balls on the weighted Riemannian manifold of shrinking type.

Proof of Theorem 1. Again, we take any point $x \in M$ and express the volume form in the geodesic polar coordinates centered at x as

$$dV|_{\exp_x(r\xi)} = J(x, r, \xi)\, dr\, d\xi$$

for $r > 0$ and $\xi \in S_x M$ a unit tangent vector at x. We let $R = d(p, x)$ and omit the dependence of the geometric quantities on ξ. Let $R = d(p, x) > 2$. Let $\gamma(s)$ be the minimizing geodesic starting from x connecting $\gamma(0) = x$ to any point $\gamma(T) \in B_p(1)$. By the triangle inequality we know that $T \in [R-1, R+1]$. It is well known [Li 1993] that along the minimizing geodesic curve γ, we have

$$m'(r) + \frac{1}{n-1} m^2(r) + \mathrm{Ric}(\partial_r, \partial_r) \leq 0,$$

where $r > 0$ and $m = m(r) = \frac{d}{dr}(\log J)(r)$. Using the Ricci soliton equation $\mathrm{Ric}_f = \frac{1}{2} g$ we immediately obtain

$$m'(r) + \frac{1}{n-1} m^2(r) \leq -\frac{1}{2} + f''(r).$$

Integrating this relation, we get for $r > 1$,

$$m(r) + \frac{1}{n-1} \int_1^r m^2(t)\, dt \leq -\frac{r-1}{2} + f'(r) - f'(1) + m(1).$$

Recall the well-known fact (see (2.3) and (2.8) in [Cao and Zhou 2010] in different notation) that for any $T > \tau > 0$, we have

$$-\tfrac{1}{2}(T-\tau) - c \leq f'(\tau) \leq -\tfrac{1}{2}(T-\tau) + c \tag{1}$$

and $|f'(T)| \leq c$ since $\gamma(T) \in B_p(1)$. Here and everywhere in the proofs, c denotes a constant depending only on the dimension n and $f(p)$. By this we know

$$-\frac{r-1}{2} + f'(r) - f'(1) \leq c_0$$

for some uniform constant $c_0 > 0$.

Using the Ricci curvature lower bound, a standard argument shows, see (1.1.8) in [Schoen and Yau 1994], that there is a uniform constant $c_1 > 0$ such that $m(s) \leq c_1$ for $s \geq \frac{1}{2}$. Then we have for another uniform constant $c_0 > 0$ and for any $r > 1$,

$$m(r) + \frac{1}{n-1} \int_1^r m^2(t)\, dt \leq c_0.$$

By the Cauchy–Schwarz inequality we obtain

$$m(r) + \frac{1}{(n-1)r} \left(\int_1^r m(t)\, dt \right)^2 < c \tag{2}$$

for any $c > c_0$ and $r > 1$.

Claim. For any $r > 1$,

$$\int_1^r m(t)\, dt \leq \sqrt{c(n-1)r}. \tag{3}$$

In fact, let

$$v(t) = \sqrt{c(n-1)t} - \int_1^t m(r)\,dr.$$

Then

$$v'(t) = \frac{\sqrt{c(n-1)}}{2\sqrt{t}} - m(t).$$

Clearly $v(1) > 0$ by the fact that $c > c_0$. Suppose that v is negative somewhere for $t > 1$. Let $d > 1$ be the first zero point of v; i.e., $v(d) = 0$. Then by the choice of d, we have $v'(d) \le 0$. That is,

$$\int_1^d m(t)\,dt = \sqrt{c(n-1)d}$$

and

$$m(d) \ge \frac{\sqrt{c(n-1)}}{2\sqrt{d}}.$$

By direct computation we know

$$m(d) + \frac{1}{(n-1)d}\left(\int_1^d m(t)\,dta\right)^2 \ge \frac{\sqrt{c(n-1)}}{2\sqrt{d}} + c,$$

which is a contradiction with (2) at $r = d$.

The relation (3) implies

$$\log J(x, r, \xi)/J(x, 1, \xi) \le \sqrt{c(n-1)r}$$

for any $r > 1$ and we have at $r = R$,

$$J(x, 1, \xi) \ge \exp(-\sqrt{c(n-1)R})J(x, R, \xi).$$

Integrating over the unit tangent vectors ξ we get

$$\mathrm{Area}(\partial B_x(1)) \ge \exp(-\sqrt{c(n-1)R})\,\mathrm{Vol}(B_p(1)),$$

where $R = d(p, x) > 2$. Similarly we have

$$\mathrm{Area}(\partial B_x(s)) \ge \exp(-\sqrt{c(n-1)R})\,\mathrm{Vol}(B_p(1))$$

for any $s \in \left[\frac{1}{2}, 1\right]$. Hence, we have

$$\mathrm{Vol}(B_x(1)) \ge \exp(-\sqrt{c(n-1)R})\,\mathrm{Vol}(B_p(1)). \qquad \square$$

A similar argument gives us the result below for gradient steady Ricci solitons.

Theorem 5. *On the complete noncompact Riemannian manifold* (M, g, f) *with Ricci curvature bounded below by* $-(n-1)k^2$ *for some constant* $k \ge 0$ *and* $\mathrm{Rc}_f \ge 0$ *on* M *and* $|\nabla f| \le C$ *on* M *for some uniform constant* $C > 0$, *we have*

$$\mathrm{Vol}(B_x(1)) \ge \exp(-\sqrt{c_0(n-1)R})\,\mathrm{Vol}(B_p(1)),$$

where c_0 is a uniform constant which does not depends on p and $R = d(p, x) > R_0$ for some uniform constant $R_0 > 1$.

The proof of Theorem 5 is almost the same as the previous proof. So we only present the necessary modification and use the same notation as above. Again we have along the minimizing geodesic curve γ,

$$m'(r) + \frac{1}{n-1} m^2(r) + \mathrm{Ric}(\partial_r, \partial_r) \le 0.$$

Using the relation $\mathrm{Ric}_f \ge 0$ we immediately obtain

$$m'(r) + \frac{1}{n-1} m^2(r) \le f''(r).$$

Integrating this relation we get for $t > 1$ and $s \in \left[\frac{1}{2}, 1\right]$,

$$m(t) + \frac{1}{n-1} \int_s^t m^2(r)\, dr \le f'(t) - f'(1) + m(s).$$

This means that for any $t > 1$ and $s = 1$,

$$m(t) + \frac{1}{n-1} \int_1^t m^2(r)\, dr \le C_0$$

for some uniform constant C_0. The remaining part of the proof is the same as the proof of Theorem 1 and we omit the details.

We now consider the volume growth of geodesic balls in manifolds with density and we show that for $(M, g, e^{-f}\, dv)$ a complete smooth metric measure space of dimension n with $\mathrm{Ric}_f \ge \frac{1}{2}$, $|\nabla f|^2 \le f$, and also with both Ricci curvature and Δf bounded from above, the volume growth of geodesic balls is in polynomial order (which may be smaller than n).

Proposition 6. *Let $(M, g, e^{-f} dv)$ be a complete smooth metric measure space of dimension n. Assume that* $\mathrm{Ric}_f \ge \frac{1}{2}$, $|\nabla f|^2 \le f$. *Assume further that $\Delta f \le K$ and* $\mathrm{Ric} \le K_1$ *for some constants $K > 0$ and $K_1 > 0$. Then for some $p \in M$, the volume growth of geodesic balls $B_p(r)$ is of polynomial order*; i.e., *there is a uniform constant $C > 0$ such that*

$$V(r) \le C r^{2K}.$$

Proof. Recall that under the conditions $\mathrm{Ric}_f \ge \frac{1}{2}$ and $|\nabla f|^2 \le f$, there are a point $p \in M$, two constants $r_0 > 0$ and a depending only on n and $f(p)$ such that

$$\left(\tfrac{1}{2} d(x, p) - a\right)^2 \le f(x) \le \left(\tfrac{1}{2} d(x, p) + a\right)^2. \tag{4}$$

This is from Proposition 4.2 in the interesting paper [Munteanu and Wang 2014]. By this we know that $|\nabla f(x)| \le \frac{1}{2} d(x, p) + a$. We may assume that $d(x, p) > 2$. Consider any minimizing normal geodesic $\gamma(s)$, $0 \le s \le r := d(x, p)$, starting from the point $\gamma(0) = p$ to the point $\gamma(r) = x$. Let $X = \dot\gamma(s)$. By the second variation formula of arc length we know

$$\int_0^r \phi^2 \operatorname{Ric}(X, X)\, ds \le (n-1)\int_0^r |\dot\phi(s)|^2\, ds$$

for any $\phi \in C_0^{1-}([0, r])$. Let $\phi(s) = s$ on $[0, 1]$, $\phi(s) = r - s$ on $[r-1, r]$, and $\phi(s) = 1$ on $[1, r-1]$. Then we have

$$\int_0^r \operatorname{Ric}(X, X)\, ds = \int_0^r \phi^2 \operatorname{Ric}(X, X)\, ds + \int_0^r (1-\phi^2) \operatorname{Ric}(X, X)\, ds.$$

We derive via the use of $\operatorname{Ric} \le K_1$, and by an argument similar to the proof before (2.8) in [Cao and Zhou 2010], that

$$\int_0^r \operatorname{Ric}(X, X)\, ds \le 2(n-1) + 2K_1.$$

Since

$$\nabla_X \dot f = \nabla^2 f(X, X) \ge \tfrac12 - \operatorname{Ric}(X, X),$$

integrating it from 0 to r, we get

$$\dot f(r) = \tfrac12 r - \int_0^r \operatorname{Ric}(X, X)\, ds \ge \tfrac12 r - c \tag{5}$$

for some constant c depending only on K_1, n and $f(p)$. Hence,

$$|\nabla f|(x) \ge \dot f(r) \ge \tfrac12 d(x, p) - c.$$

Define

$$\rho(x) = 2\sqrt{f(x)}.$$

Then,

$$|\nabla\rho| = \frac{|\nabla f|}{\sqrt f} \le 1.$$

Let, for $r > 0$ large,

$$D(r) = \{x \in M; \rho(x) \le r\}, \quad V(r) = \operatorname{Vol}(D(r)).$$

As in [Cao and Zhou 2010], by using the coarea formula we have

$$V(r) = \int_0^r ds \int_{\partial D(r)} \frac{1}{|\nabla\rho|}\, dA,$$

$$V'(r) = \int_{\partial D(r)} \frac{1}{|\nabla\rho|}\, dA = \frac{r}{2}\int_{\partial D(r)} \frac{1}{|\nabla f|}\, dA.$$

By the assumption $\Delta f \le K$ and the divergence theorem we have

$$2K V(r) \ge 2\int_{D(r)} \Delta f = 2\int_{\partial D(r)} |\nabla f|\, dA.$$

By (5) we know that on $\partial D(r)$, there is a uniform constant $C > 2$ such that for $r \geq 2C$,

$$|\nabla f|^2 \geq f - C.$$

Then we have

$$2\int_{\partial D(r)} |\nabla f|\, dA \geq 2\int_{\partial D(r)} \frac{f - C}{|\nabla f|}\, dA.$$

The right side of above inequality is greater than or equal to

$$(r-2)V'(r).$$

Hence we have

$$2K V(r) \geq (r-2)V'(r),$$

which then implies

$$V(r) \leq V(2C) r^{2K}$$

for $r > 2C$. □

We remark that the above argument is motivated by the proof of the volume growth estimate in [Cao and Zhou 2010]. Our result is different from the deep result Theorem 1.4 in [Munteanu and Wang 2014] in the case when the constant $2K$ is smaller than n.

3. Harmonic and f-harmonic functions on Ricci–Einstein solitons

We now prove Theorem 3. We wish that we could use the Caccioppoli argument (see the proof of Proposition 8.1 in [Naber 2010], with the use of Lemma 2.2 replaced by Proposition 4.2 in [Munteanu and Wang 2014]) to conclude that with some decay assumption such as finite energy, an f-harmonic function u is a constant function on M. However, we have a simpler proof of this result below.

Proof of Theorem 3. Recall the Bochner formula for the f-harmonic function $u : M \to R$,

$$\tfrac{1}{2}\Delta_f |\nabla u|^2 = |\nabla^2 u|^2 + \mathrm{Rc}_f(\nabla u, \nabla u).$$

By our assumption that $\mathrm{Ric}_f \geq h(x)g$, we know

$$\tfrac{1}{2}\Delta_f |\nabla u|^2 \geq |\nabla^2 u|^2 + h(x)|\nabla u|^2.$$

Let ϕ be the standard cut-off function on $B_p(2r)$ and let $dm = \exp(-f)\, dv_g$. Then we have

$$\int_M (|\nabla^2 u|^2 + h(x)|\nabla u|^2)\phi\, dm \leq \int_M \bigl(\tfrac{1}{2}\Delta_f \phi\bigr)|\nabla u|^2\, dm.$$

The right side goes to zero as $r \to \infty$. Hence we have

$$\int_M (|\nabla^2 u|^2 + h(x)|\nabla u|^2)\, dm = 0,$$

which implies that u is a constant. □

Let (M, g) be a complete noncompact Riemannian manifold of dimension n. Fix $p \in M$. In this section we always assume that (M, g) satisfies $\mathrm{Ric}_f \geq h(x)g$ for some function $h(x)$ and $|\nabla f| \leq \alpha d(x, p) + b$. Then we have $R + \Delta f \geq nh(x)$ on M. We study the L^2 estimate for the Hessian matrix for harmonic functions with finite energy.

Proof of Proposition 4. Let $u : M \to R$ be a harmonic function on (M, g, f) with finite energy

$$\int_M |\nabla u|^2 < \infty.$$

Recall the Bochner formula for the harmonic function $u : M \to R$,

$$\tfrac{1}{2}\Delta|\nabla u|^2 = |\nabla^2 u|^2 + \mathrm{Rc}(\nabla u, \nabla u).$$

Using the assumption $\mathrm{Ric}_f \geq h(x)g$ we have

$$\tfrac{1}{2}\Delta|\nabla u|^2 \geq |\nabla^2 u|^2 + h(x)|\nabla u|^2 - \nabla^2 f(\nabla u, \nabla u). \tag{6}$$

Recall the Hessian matrix $\nabla^2 f = (f_{ij})$ in local coordinates (x_i) in M.

Let $\phi = \phi_r$ be the cut-off function on $B_{2r}(p)$. We write by $o(1)$ the quantities such that $o(1) \to 0$ as $r \to \infty$. Then, we have

$$\int_M (|\nabla^2 u|^2 + h(x)|\nabla u|^2)\phi^2 \leq \int_M \big(\tfrac{1}{2}\Delta|\nabla u|^2 + \nabla^2 f(\nabla u, \nabla u)\big)\phi^2. \tag{7}$$

By direct computation we have, for $\epsilon > 0$ small,

$$\int_M \big(\tfrac{1}{2}\Delta|\nabla u|^2\big)\phi^2 = -2\int_M \phi D^2 u(\nabla u, \nabla\phi) \leq \epsilon \int_M |\nabla^2 u|^2\phi^2 + o(1).$$

Using $|\nabla f| \leq \alpha d(x, p) + b$ and integrating by parts, we obtain

$$\int_M f_{ij} u_i u_j \phi^2 = -\int f_i u_{ij} u_j \phi^2 + o(1).$$

Furthermore, we have

$$\int_M f_{ij} u_i u_j \phi^2 = \frac{1}{2}\int_M \Delta f |\nabla u|^2 \phi^2 + o(1).$$

Hence by (6) we have

$$\int_M \big((1-\epsilon)|\nabla^2 u|^2 + h(x)|\nabla u|^2\big)\phi^2 - \frac{1}{2}\int_M \Delta f|\nabla u|^2\phi^2 \leq o(1).$$

That is,

$$\int_M (1-\epsilon)|\nabla^2 u|^2\phi^2 + \int_M h(x)|\nabla u|^2\phi^2 \le \frac{1}{2}\int_M \Delta f|\nabla u|^2\phi^2 + o(1).$$

Sending $r \to \infty$ and letting $\epsilon \to 0$, we obtain

$$\int_M |\nabla^2 u|^2 + \int_M h(x)|\nabla u|^2 \le \frac{1}{2}\int_M \Delta f|\nabla u|^2. \tag{8}$$

Note that when $\Delta f \le 2h(x)$ on M, by (8) we have $D^2 u = 0$ in M; i.e., ∇u is a parallel vector field on M. □

We remark that when $\mathrm{Ric}_f = \lambda g$ (and $R + \Delta f = n\lambda$) with λ being a constant, by the proof of Proposition 4 we have

$$\int_M |\nabla^2 u|^2 + \frac{1}{2}\int_M R|\nabla u|^2 \le \int_M \frac{(n-2)\lambda}{2}|\nabla u|^2.$$

Note that the assumption about the potential function f in Proposition 4 is true on the steady soliton (M, g); see [Hamilton 1995]. We now give an application of this integral inequality (8). A special case of the Liouville-type theorem below, due to Munteanu and Sesum [2013, Theorem 4.1], can be derived from the integral estimate (8).

Proposition 7. *Let $n = 2$. Assume that the complete noncompact surface (M, g, f) satisfies* $\mathrm{Ric}_f = h(x)g$ *on M with $R \ge 0$, and $|\nabla f| \le b$ for some $b > 0$ in M. Then there is no nontrivial harmonic function on (M, g) with finite energy on (M, g).*

Proof. Note that $R + \Delta f = nh(x) = 2h(x)$ in M. Then $\Delta f = 2h(x) - R$ in M. By (8) we have

$$\int_M |\nabla^2 u|^2 + \frac{1}{2}\int_M R|\nabla u|^2 \le 0.$$

If $R = 0$, then (M, g) is flat and the result follows from Theorem 4.1 in [Munteanu and Sesum 2013].

We may assume $R > 0$ in M. We argue by contradiction. Assume that there is a nontrivial harmonic function with finite energy on (M, g). By (8) we know

$$\int_M |\nabla^2 u|^2 + \frac{1}{2}\int_M R|\nabla u|^2 = 0.$$

Hence ∇u is a parallel vector field on M and $R = 0$, a contradiction with $R > 0$. □

We remark that we can give a new proof of Theorem 4.1 in [Munteanu and Sesum 2013], which is on a gradient steady Ricci soliton. It says that there is no nontrivial harmonic function with finite energy on the steady Ricci soliton (M, g). The proof is below. We may assume that (M, g, f) is a nontrivial steady Ricci soliton. Recall that it is well known that either $R > 0$ or $R = 0$ on M. By (8), we have $R = 0$, and then

$$\Delta_f R = -2|\mathrm{Ric}|^2,$$

and we know that $\mathrm{Ric} = 0$ on M. By [Schoen and Yau 1994], we know that there is no nontrivial harmonic function with finite energy.

Acknowledgements

The author thanks the anonymous referees very much for useful suggestions. This work was done when the author was visiting the University of Minnesota in September 2016; he would like to thank Prof. Jiaping Wang for helpful discussions and the hospitality of the School of Mathematics.

References

[Bakry and Émery 1985] D. Bakry and M. Émery, "Diffusions hypercontractives", pp. 177–206 in *Séminaire de Probabilités XIX* (Strasbourg, 1983/84), edited by J. Azéma and M. Yor, Lecture Notes in Math. **1123**, Springer, 1985. MR Zbl

[Cao 2007] X. Cao, "Compact gradient shrinking Ricci solitons with positive curvature operator", *J. Geom. Anal.* **17**:3 (2007), 425–433. MR Zbl

[Cao 2010] H.-D. Cao, "Recent progress on Ricci solitons", pp. 1–38 in *Recent advances in geometric analysis*, edited by Y.-I. Lee et al., Adv. Lect. Math. (ALM) **11**, International, Somerville, MA, 2010. MR Zbl

[Cao and Zhou 2010] H.-D. Cao and D. Zhou, "On complete gradient shrinking Ricci solitons", *J. Differential Geom.* **85**:2 (2010), 175–185. MR Zbl

[Carrillo and Ni 2009] J. A. Carrillo and L. Ni, "Sharp logarithmic Sobolev inequalities on gradient solitons and applications", *Comm. Anal. Geom.* **17**:4 (2009), 721–753. MR Zbl

[Chen 2009] B.-L. Chen, "Strong uniqueness of the Ricci flow", *J. Differential Geom.* **82**:2 (2009), 363–382. MR Zbl

[Chow et al. 2006] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Graduate Studies in Mathematics **77**, American Mathematical Society, Providence, RI, 2006. MR Zbl

[Hamilton 1995] R. S. Hamilton, "The formation of singularities in the Ricci flow", pp. 7–136 in *Surveys in differential geometry, II* (Cambridge, MA, 1993), edited by S.-T. Yau, International, Cambridge, MA, 1995. MR Zbl

[Haslhofer and Müller 2011] R. Haslhofer and R. Müller, "A compactness theorem for complete Ricci shrinkers", *Geom. Funct. Anal.* **21**:5 (2011), 1091–1116. MR Zbl

[Li 1993] P. Li, *Lecture notes on geometric analysis*, Lecture Notes Series **6**, Seoul National University, Seoul, 1993. MR Zbl

[Lichnerowicz 1970] A. Lichnerowicz, "Variétés riemanniennes à tenseur C non négatif", *C. R. Acad. Sci. Paris Sér. A-B* **271** (1970), A650–A653. MR Zbl

[Lott 2003] J. Lott, "Some geometric properties of the Bakry–Émery–Ricci tensor", *Comment. Math. Helv.* **78**:4 (2003), 865–883. MR Zbl

[Lott and Villani 2009] J. Lott and C. Villani, "Ricci curvature for metric-measure spaces via optimal transport", *Ann. of Math.* (2) **169**:3 (2009), 903–991. MR Zbl

[Ma 2013] L. Ma, "Remarks on compact shrinking Ricci solitons of dimension four", *C. R. Math. Acad. Sci. Paris* **351**:21-22 (2013), 817–823. MR Zbl

[Ma and Du 2010] L. Ma and S.-H. Du, "Extension of Reilly formula with applications to eigenvalue estimates for drifting Laplacians", *C. R. Math. Acad. Sci. Paris* **348**:21-22 (2010), 1203–1206. MR Zbl

[Morgan 2005] F. Morgan, "Manifolds with density", *Notices Amer. Math. Soc.* **52**:8 (2005), 853–858. MR Zbl

[Munteanu and Sesum 2013] O. Munteanu and N. Sesum, "On gradient Ricci solitons", *J. Geom. Anal.* **23**:2 (2013), 539–561. MR Zbl

[Munteanu and Wang 2011] O. Munteanu and J. Wang, "Smooth metric measure spaces with non-negative curvature", *Comm. Anal. Geom.* **19**:3 (2011), 451–486. MR Zbl

[Munteanu and Wang 2014] O. Munteanu and J. Wang, "Geometry of manifolds with densities", *Adv. Math.* **259** (2014), 269–305. MR Zbl

[Munteanu and Wang 2015] O. Munteanu and J. Wang, "Geometry of shrinking Ricci solitons", *Compos. Math.* **151**:12 (2015), 2273–2300. MR Zbl

[Naber 2010] A. Naber, "Noncompact shrinking four solitons with nonnegative curvature", *J. Reine Angew. Math.* **645** (2010), 125–153. MR Zbl

[Perelman 2000] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2000. Zbl arXiv

[Schoen and Yau 1994] R. Schoen and S.-T. Yau, *Lectures on differential geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology **1**, International, Cambridge, MA, 1994. MR Zbl

[Sturm 2006a] K.-T. Sturm, "On the geometry of metric measure spaces, I", *Acta Math.* **196**:1 (2006), 65–131. MR Zbl

[Sturm 2006b] K.-T. Sturm, "On the geometry of metric measure spaces, II", *Acta Math.* **196**:1 (2006), 133–177. MR Zbl

[Wei and Wylie 2009] G. Wei and W. Wylie, "Comparison geometry for the Bakry–Emery Ricci tensor", *J. Differential Geom.* **83**:2 (2009), 377–405. MR Zbl

[Yang 2009] N. Yang, "A note on nonnegative Bakry–Émery Ricci curvature", *Arch. Math.* (*Basel*) **93**:5 (2009), 491–496. MR Zbl

Received January 3, 2017. Revised December 5, 2017.

LI MA
SCHOOL OF MATHEMATICS AND PHYSICS
UNIVERSITY OF SCIENCE AND TECHNOLOGY BEIJING
BEIJING
CHINA

and

DEPARTMENT OF MATHEMATICS
HENAN NORMAL UNIVERSITY
XINXIANG
CHINA

lma@tsinghua.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.371

PRESENTATIONS OF GENERALISATIONS OF THOMPSON'S GROUP V

CONCHITA MARTÍNEZ-PÉREZ, FRANCESCO MATUCCI AND BRITA NUCINKIS

We consider generalisations of Thompson's group V, denoted by $V_r(\Sigma)$, which also include the groups of Higman, Stein and Brin. We showed earlier (*Forum Math.* 28:5 (2016), 909–921) that under some mild conditions these groups and centralisers of their finite subgroups are of type F_∞. Under more general conditions we show that the groups $V_r(\Sigma)$ are finitely generated and, under the mild conditions mentioned above for which they are of type F_∞ and hence finitely presented, we give a recipe to find explicit presentations. For the centralisers of finite subgroups we find a suitable infinite presentation and then show how to apply a general procedure to shorten this presentation. In the appendix, we give a proof of this general shortening procedure.

1. Introduction

The original Thompson groups $F \leq T \leq V$ are groups of homeomorphisms of the unit interval, the circle and the Cantor set respectively. In this note we consider generalisations of these groups, which are described as groups of automorphisms of certain Cantor algebras. These groups include Higman's [1974], Stein's [1992] and Brin's [2004] generalisations of V.

The groups F, T and V have attracted the attention of group theorists for several reasons, one of them being that there are nice presentations and ways to represent elements available, making it possible to prove interesting results about metrics, geodesics and decision problems. However, the situation changes when one moves to some of their generalisations. There are presentations available for Higman's groups $V_{n,r}$ [1974], Stein's generalisations [Brin and Squier 2001; Stein 1992]

This work was partially funded by LMS Scheme 4 grant 41209. The first named author was supported by Gobierno de Aragón, European Regional Development Funds and MTM2015-67781-P (MINECO/FEDER). The second author is a member of the Gruppo Nazionale per le Strutture Algebriche, Geometriche e le loro Applicazioni (GNSAGA) of the Istituto Nazionale di Alta Matematica (INdAM) and gratefully acknowledges the support of the Fondation Mathématique Jacques Hadamard (ANR-10-CAMP-0151-02 - FMJH - Investissement d'Avenir) and the Fundação para a Ciência e a Tecnologia (FCT projects UID/Multi/04621/2013 (CEMAT-Ciências) and PEst-OE/MAT/UI0143/2014).
MSC2010: 20J05.
Keywords: generalized Thompson groups, finite presentations.

and Brin's higher dimensional Thompson groups sV [Hennig and Matucci 2012], but not for more complicated generalisations such as the groups $V_r(\Sigma)$ we are considering here. These were defined in [Kochloukova et al. 2013; Martínez-Pérez and Nucinkis 2013] and were denoted $G_r(\Sigma)$. It is worth pointing out that elements in $V_r(\Sigma)$ admit a tree-pair representation similar to that of the original groups F, T and V. The authors show in [Martínez-Pérez et al. 2016] that, under some mild hypotheses, being *valid* and *bounded*, $V_r(\Sigma)$ is the full automorphism group of a Cantor algebra. In the same paper it is shown that under some further minor restrictions, being *complete*, these groups are of type F_∞ and that this also implies that centralisers of finite subgroups are of type F_∞. We introduce all necessary background in Section 2. The structure of centralisers in $V_r(\Sigma)$ is studied in detail in [Martínez-Pérez et al. 2016; Martínez-Pérez and Nucinkis 2013].

One of the objectives of the present paper is to introduce a common framework providing recipes; first to find explicit finite generating sets for the groups $V_r(\Sigma)$ in the case when the underlying Cantor algebra $U_r(\Sigma)$ is valid and bounded, and second, to find explicit presentations under the additional assumption that $U_r(\Sigma)$ is complete. To do that, we construct a model for the classifying space for free actions $\mathrm{E}G$ for $G = V_r(\Sigma)$, and use this model to obtain presentations of these groups. As far as we are aware, this construction is new even for the group V, and hence could be of independent interest.

In Section 7 we also give an explicit finite presentation for centralisers of finite subgroups for those $V_r(\Sigma)$ that are finitely presented. To do so we use the so-called *Burnside procedure* as used by Guralnick, Kantor, Kassabov and Lubotzky [Guralnick et al. 2011].

In the Appendix we shall give an outline and proof of the Burnside procedure as used in [Guralnick et al. 2011]. This procedure is well known, but we are not aware of any proofs elsewhere. The point is to look for a simple presentation for G that is somehow symmetric and elementary. Initially it may have infinitely many generators and relations; the Burnside procedure offers a way to cut it down to a more manageable, and sometimes finite, presentation.

2. Background on generalised Thompson groups

In this section we introduce only those properties of valid bounded Cantor algebras used to make this paper self-contained. For detailed definitions and notation the reader is referred to [Martínez-Pérez and Nucinkis 2013, Section 2], and for proofs of statements cited here, to [Kochloukova et al. 2013; Martínez-Pérez et al. 2016; Martínez-Pérez and Nucinkis 2013].

Let $S = \{1, \dots, s\}$ be a finite set of colours and associate to each $i \in S$ an integer $n_i > 1$, called the arity of the colour i. For every $i \in S$ consider the following right operations on a set U:

(i) One n_i-ary operation $\lambda_i : U^{n_i} \to U$, and

(ii) n_i unary operations $\alpha_i^1, \dots, \alpha_i^{n_i}$, where α_i^j is a map $U \to U$.

We also consider, for each $i \in S$ and $v \in U$, the map

$$\alpha_i : U \to U^{n_i}$$

given by $v\alpha_i := (v\alpha_i^1, v\alpha_i^2, \dots, v\alpha_i^{n_i})$. The maps α_i are called *descending* operations, or expansions, and the maps λ_i are called *ascending* operations, or contractions.

Fix a finite set X_r of cardinality $|X_r| = r$. One can define the free object on the set X_r with respect to the previous operations which we denote U. To define our generalisations of Thompson's group V, we will be interested in the free object constructed under the extra requirement that a certain set of laws Σ described below must be satisfied. We denote this last free object by $U_r(\Sigma)$ and call it the (free) Cantor algebra on X_r satisfying Σ.

Definition 2.1 [Martínez-Pérez and Nucinkis 2013, Section 2]. Fix a finite set X_r of cardinality $|X_r| = r$ and consider the free object U on X_r with respect to operations (i) and (ii) above. Then $\Sigma = \Sigma_1 \cup \Sigma_2$ with Σ_1 and Σ_2 the following set of laws:

(i) Σ_1 is given by

$$u\alpha_i\lambda_i = u, \qquad (u_1, \dots, u_{n_i})\lambda_i\alpha_i = (u_1, \dots, u_{n_i}),$$

for every $u \in U$, $i \in S$, and n_i-tuple, $(u_1, \dots, u_{n_i}) \in U^{n_i}$.

(ii) Σ_2 is given by

$$\Sigma_2 = \bigcup_{1 \le i < i' \le s} \Sigma_2^{i,i'},$$

where each $\Sigma_2^{i,i'}$ is either empty or consists of the following laws: consider first i and fix a map $f : \{1, \dots, n_i\} \to \{1, \dots, s\}$. For each $1 \le j \le n_i$, we see $\alpha_i^j \alpha_{f(j)}$ as a set of sequences of length 2 of descending operations and let

$$\Lambda_i = \cup_{j=1}^{n_i} \alpha_i^j \alpha_{f(j)}.$$

Do the same for i' (with a corresponding map f') to get $\Lambda_{i'}$. We need to assume that f and f' are chosen so that $|\Lambda_i| = |\Lambda_{i'}|$ and fix a bijection $\phi : \Lambda_i \to \Lambda_{i'}$. Then $\Sigma_2^{i,i'}$ is

$$u\nu = u\phi(\nu), \quad \nu \in \Lambda_i, u \in U.$$

Let $U_r(\Sigma)$ be the algebra obtained from U by quotienting out the relations in Σ. We say that $U_r(\Sigma)$ is *valid* if for any set $Y \in U$, we have $|Y| = |\overline{Y}|$, where $\overline{Y}$ is the image of Y in $U_r(\Sigma)$. In particular this implies that $U_r(\Sigma)$ is a free object on X in the class of those algebras with the descending and ascending operations (i) and (ii) above which satisfy the identities Σ.

From now on we work with the free object $U_r(\Sigma)$ only. Let $B \subset U_r(\Sigma)$, $b \in B$, and let i be a colour of arity n_i. The set

$$(B \smallsetminus \{b\}) \cup \{b\alpha_i^1, \dots, b\alpha_i^{n_i}\}$$

is called a *simple expansion* of B. Analogously, if $b_1, \dots, b_{n_i} \subseteq B$ are pairwise distinct,

$$(B \smallsetminus \{b_1, \dots, b_{n_i}\}) \cup \{(b_1, \dots, b_{n_i})\lambda_i\}$$

is a *simple contraction* of B. A finite chain of simple expansions is an expansion and a finite chain of simple contractions is a contraction. A subset $A \subseteq U_r(\Sigma)$ is called *admissible* if it can be obtained from the set X_r by finitely many expansions or contractions. If a subset A_1 is obtained from a subset A by an expansion (simple or not), then we write $A \leq A_1$.

Remark 2.2. Recall that $U_r(\Sigma)$ is said to be *bounded* (see [Martínez-Pérez and Nucinkis 2013, Definition 2.7]) if for all admissible subsets Y and Z such that there is some admissible $A \leq Y, Z$, there is a unique least upper bound of Y and Z. By a unique least upper bound we mean an admissible subset T such that $Y \leq T$ and $Z \leq T$, and whenever there is an admissible set S also satisfying $Y \leq S$ and $Z \leq S$, then $T \leq S$.

By [Kochloukova et al. 2013, Lemma 2.5], any admissible set is a basis of $U_r(\Sigma)$. Conversely, by [Martínez-Pérez et al. 2016, Theorem 2.5], if Σ is valid and bounded, any basis of $U_r(\Sigma)$ is also an admissible set. Furthermore, for every admissible subset of cardinality m, we have that

$$m \equiv r \bmod d \qquad \text{for } d := \gcd\{n_i - 1 \mid i = 1, \dots, s\}.$$

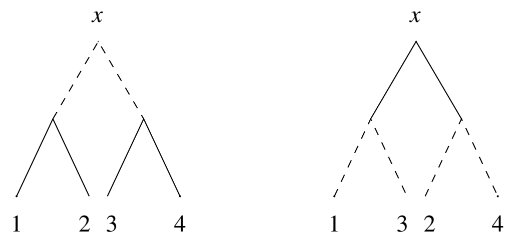
In particular, any basis with m elements can be transformed into one of r elements. Hence $U_r(\Sigma) = U_m(\Sigma)$ and we may assume that $r \leq d$.

Definition 2.3. [Martínez-Pérez and Nucinkis 2013, Definition 2.12] Let $U_r(\Sigma)$ be a valid Cantor algebra. We denote the group of all Cantor algebra automorphisms of $U_r(\Sigma)$ by $V_r(\Sigma)$. In particular, these automorphisms are induced by a map $V \to W$, where V and W are admissible subsets of $U_r(\Sigma)$ of the same cardinality. In our notation automorphisms act on the left.

For example, when $s = 1$ we have $\Sigma_2 = \varnothing$ and we retrieve the original Higman–Thompson groups $G_{r,n}$ (here, $n = n_1$) [Higman 1974]. For

$$s = 2, \quad r = 1 \quad \text{and} \quad n_1 = n_2 = 2,$$

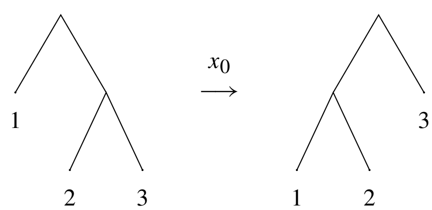
the Brin–Thompson groups are now given by the set Σ_2 that can be visualised as follows:



Here dashed and solid lines represent expansions of different colours. For more examples the reader is referred to [Martínez-Pérez et al. 2016; Martínez-Pérez and Nucinkis 2013].

Remark 2.4. If $U_r(\Sigma)$ is valid and bounded every element of $V_r(\Sigma)$ can be given by a bijection $V \to W$, where V and W are descendants of the fixed basis X_r.

For $r = 1$, this means that we can visualise elements of $V_1(\Sigma)$ by tree-pair diagrams of rooted trees, where the root represents the basis $X_1 = \{x\}$. So, for example, when $s = 1$ and $n_1 = 2$, $V_1(\Sigma)$ is equal to V (the original Thompson group), and the well-known generator $x_0 \in F \subset V$ is visualised as



Definition 2.5 [Martínez-Pérez et al. 2016, Definition 3.2]. Let $B \leq A$ be admissible subsets of $U_r(\Sigma)$. We say that the expansion $B \leq A$ is *elementary* if there are no repeated colours in the paths from elements in B to their descendants in A. We denote an elementary expansion by $B \preceq A$. We say that the expansion is *very elementary* if all paths have length at most 1.

Denote by $\mathcal{P}_r$ the poset of admissible subsets in $U_r(\Sigma)$, and by $|\mathcal{P}_r|$ its geometric realisation. (It was shown in [Martínez-Pérez and Nucinkis 2013] that $|\mathcal{P}_r|$ is a model for $\underline{E}G$, the classifying space for proper actions). We now describe the Stein complex $\mathcal{S}_r(\Sigma)$ [Stein 1992], which is a subcomplex of $|\mathcal{P}_r|$. The vertices in $\mathcal{S}_r(\Sigma)$ are given by the admissible subsets of $U_r(\Sigma)$. The k-simplices are given by chains of expansions $Y_0 \leq \cdots \leq Y_k$, where $Y_0 \preceq Y_k$ is an elementary expansion.

From now on we will denote $V_r(\Sigma)$ by G. In the next section we will use $\mathcal{S}_r(\Sigma)$ to construct a model for EG. Recall that by [Martínez-Pérez et al. 2016, Lemma 3.6 and Remark 3.7], $\mathcal{S}_r(\Sigma)$ is contractible and has finite stabilisers.

3. A model for EG

In this section we construct a model for the space EG when G is the automorphism group of a valid and bounded Cantor algebra $U_r(\Sigma)$ as before. We shall use this model to get, initially infinite, presentations for our groups, which we will then reduce to obtain a finite generating set, and later a finite presentation under some extra hypothesis on $U_r(\Sigma)$.

3A. ***Some technical observations.*** To begin with, we collect few technical observations that we will use later on. As seen before, the elements g in our group G can be expressed via a bijection between a pair of admissible subsets (or bases) (B, B') both of the same cardinality. Observe that the pair above is not enough to determine g and that we have to specify the explicit bijection. A way to overcome this problem is to work with ordered bases, in the sense that instead of a basis B viewed as a set, we will be considering an ordered tuple A with underlying set B. We say $u(A) = B$ (u for *underlying*). A pair of ordered tuples (A, A'), with both A and A' of the same cardinality, uniquely determines the element of G mapping the elements of A to the elements of A' in the prescribed order; conversely, any group element is expressible in this way. Of course, just as for the representation of the pair of bases, there is not a unique pair (A, A') determining a given $g \in G$: we may apply descending or ascending operations to A and A' in a consistent way to get a new pair of ordered tuples representing the same group element. Moreover, we may also permute the elements of both tuples in a consistent way and still get the same g. This means that when we represent elements of G as pairs (A, A') we should be talking about equivalence classes of pairs under the obvious equivalence relation that identifies pairs yielding the same element. However, to make the notation lighter we will talk about pairs and denote them as above. The following definition will be useful later on: given tuples A_1, A_2 with bases as underlying sets we put

$$A_1 \precsim A_2 \iff u(A_1) \leq u(A_2) \text{ is an elementary expansion.}$$

Equivalently, $A_1 \precsim A_2$ if $u(A_1) \leq u(A_2)$ in the Stein poset. Observe that this is not a partial order, as it is not antisymmetric: we could have $A_1 \precsim A_2$ and $A_2 \precsim A_1$ but $A_1 \neq A_2$. When $A_1 \precsim A_2$, abusing the terminology slightly, we will say that A_2 is obtained from A_1 by *descending operations*. Essentially this means that we are considering the permutation of the elements of a tuple as a new type of descending operation. Of course this could equally be viewed as an "ascending" operation, but it turns out to be convenient to view it as descending. If we want to record the precise operations that yield A_2 when applied to A_1 we will write

$$A_1 \overset{\varepsilon}{\precsim} A_2$$

and will also set $A_2 = A_1\varepsilon$. Observe that ε can be seen as a precise recipe to get A_2, and ε encodes exactly which elements are modified, permuted and so on.

3B. ***The model for EG.*** Let Z be the complex constructed as follows: The points of Z are the ordered tuples A with underlying set a basis $u(A)$ in the Stein complex $\mathcal{S}_r(\Sigma)$. For each chain

$$A_0 \precsim \cdots \precsim A_k$$

we attach an (oriented) k-simplex at the vertices $A_0, \dots, A_k$. Observe that there might be repeated vertices, so this is not a simplicial complex but rather has the structure of a Δ-complex; see [Hatcher 2002, Section 2.1]. The group G acts on the set of bases, and using that action one can define a G-action on Z in the obvious way. Note that this action is free. In particular this implies that two different 1-simplices starting in A_0, say $A_0 \precsim A_1$ and $A_0 \precsim A_1'$ cannot be in the same G-orbit. Hence they yield different 1-simplices in the quotient complex Z/G. Conversely, if $\overline{A}_0 \xrightarrow{\bar{\varepsilon}} \overline{A}_1$ is an edge in Z/G, then once we have fixed a lift A_0 of $\overline{A}_0$ to Z, $\bar{\varepsilon}$ lifts to a unique 1-simplex of Z. Therefore there is some well-defined set of descending operations giving a tuple A_1' which is uniquely determined so that $A_0 \overset{\varepsilon}{\precsim} A_1'$ is the lift of $\bar{\varepsilon}$. Moreover, the tuple A_1' is uniquely determined. Note that we have the extra restriction coming from the Stein poset: we can only apply descending operations of the same colour once to any element of A_0.

Applying the same argument implies that this also holds for any lift of a path in Z/G to Z.

We now show that Z is contractible by using the contractibility of $\mathcal{S}_r(\Sigma)$, [Stein 1992]. There is a G-map

$$u : Z \to \mathcal{S}_r(\Sigma)$$

associating the underlying basis to an ordered tuple.

Fix a basis $B \in \mathcal{S}_r(\Sigma)$ of cardinality k. Then $u^{-1}(B)$ is the full subcomplex of Z with 0-simplices given by the tuples with underlying set B, i.e., given by all possible permutations of the elements in B. Let H the stabiliser of B in G. Then H is isomorphic to the symmetric group of degree k and acts freely on the 0-simplices of $u^{-1}(B)$. In fact we may choose a bijection between the 0-simplices of $u^{-1}(B)$ and the elements of H and the definition of the complex structure of Z means that any $(k+1)$-element subset of 0-simplices spans a k-simplex.

For example if $H = S_2$ is the symmetric group on two letters with elements 1 and x, then the 1-simplices are $\{1, 1\}$, $\{1, x\}$, $\{x, 1\}$ and $\{x, x\}$, and the 2-simplices are $\{1, 1, 1\}$, $\{1, 1, x\}$ etc.

In other words, $u^{-1}(B)$ is easily seen to be the usual complex associated to the bar resolution of the finite group H; see for example [Hatcher 2002, Example 1B.7]. In particular this shows that $u^{-1}(B)$ is contractible.

Using [Quillen 1973, Theorem A], we can now show that u is a homotopy equivalence. To see this, let J_Z be the category with objects the simplices of Z and morphisms given by the face relations. Note that since Z is not a simplicial complex, this is not a poset. Let J_S be the poset of simplices in $\mathcal{S}_r(\Sigma)$. The map u induces a functor

$$J_u : J_Z \to J_S,$$

and the geometric realisations of nerves of the categories J_Z and J_S are the barycentric subdivisions of Z and $\mathcal{S}_r(\Sigma)$, respectively. Once we show that for any $\sigma : B_0 < B_1 < \cdots < B_t$ in J_S,

$$J_u/\sigma := \{\tau \in J_Z \mid J_u(\tau) \text{ is a subsimplex of } \sigma\}$$

is a contractible subcategory of J_Z, we can use Quillen's Theorem A to deduce that J_u is a homotopy equivalence. The category J_u/σ is just the category with objects the simplices in the join

$$u^{-1}(B_0) \star \cdots \star u^{-1}(B_t)$$

and morphisms given by face relations. As $u^{-1}(B_0) \star \cdots \star u^{-1}(B_t)$ is contractible, this category is also contractible. Hence J_u is a homotopy equivalence and thus u is, too. Since $\mathcal{S}_r(\Sigma)$ is contractible we deduce that Z is contractible as required.

4. An infinite presentation

In this section we use the model for EG that we have just constructed to obtain a presentation for our group. As the model is of infinite type, our presentation will initially be infinite. But in the case when the Cantor algebra is valid and bounded it is possible to reduce the generating system to a finite one, as we will see in the next section.

We obtain our presentation using the following well known result that we recall here for the reader's convenience.

Theorem 4.1 [Geoghegan 2008, Theorem 3.1.16 and Corollary 3.1.17]. *Let G be a group and Z a simply connected CW-complex with a free G-action such that Z/G is oriented and path connected. Let $\mathcal{T}$ be a maximal tree in Z/G. Let*:

- *W be the set of* (*oriented*) 1*-cells of Z/G.*
- *R be the set of words in the alphabet $W \cup W^{-1}$ obtained as follows*: *for each* (*oriented*) 2*-cell e^2_γ in Z/G, let $\tau(e^2_\gamma)$ be a word representing the boundary δe^2_γ and set*
$$R = \{\tau(e^2_\gamma) \mid e^2_\gamma \text{ is an oriented 2-cell of } Z/G\}.$$
- *$S \subset W$ be the set of* (*oriented*) 1*-cells of $\mathcal{T}$* (*seen as one letter words in W*).

Then

$$\langle W \mid R \cup S\rangle$$

is a presentation of the group $G \cong \pi_1(Z/G)$. If, *moreover*, *Z/G has a finite* 2*-skeleton*, *then this is a finite presentation.*

4A. ***The isomorphism $G \cong \pi_1(Z/G)$.*** We now give an explicit isomorphism between G and the fundamental group of Z/G, where we return to our previous notation so $G = V_r(\Sigma)$ and Z is the same complex as in Section 3B. The standard way to show this isomorphism is to fix some point $x_0 \in Z$ and map the element $g \in G$ to the path in Z/G obtained by taking the quotient of a path from x_0 to gx_0 in Z. As x_0 and gx_0 have the same cardinality, what we get is a loop path in Z/G. We shall take as x_0 a tuple with underlying set our preferred basis of r elements X_r. To ease notation, we denote this tuple by X_r as well. As the G-action on Z preserves the cardinality of each tuple, the 0-simplices of Z/G correspond to the possible cardinalities of tuples (or of bases). By Remark 2.2, we recall that the possible cardinalities of the bases are exactly the integers congruent to r modulo d where $n_1, \dots, n_s$ are the arities and

$$d = \gcd(n_1-1, \dots, n_s-1).$$

So the 0-simplices of Z/G can be labelled as

$$\{\overline{X}_i \mid i \equiv r \bmod d\},$$

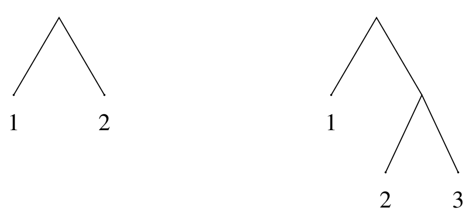
where the subscript is the cardinality of the associated bases. Now, choose a maximal tree $\mathcal{T}$ in Z/G. The vertices of $\mathcal{T}$ are all the 0-simplices above and there is a unique path in $\mathcal{T}$ from $\overline{X}_r$ to every other $\overline{X}_i$. This path determines uniquely a precise tuple X_i that is a lift of $\overline{X}_i$ (observe that X_i depends on the choice of $\mathcal{T}$).

Let $\overline{X}_i \xrightarrow{\bar{\varepsilon}} \overline{X}_j$ be an edge (thus $i \le j$). By the comments above there is a uniquely determined lift $X_i \xrightarrow{\varepsilon} X'_j$ of $\bar{\varepsilon}$; here X'_j is a new tuple which is in the same orbit as X_j. Therefore there is a uniquely determined $g \in G$ such that $X'_j = gX_j$, and this is precisely the element in G corresponding to the generator

$$\bar{\varepsilon} \in \pi_1(Z/G).$$

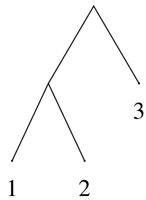
We have $g = (X_j, X'_j)$ and $X'_j = X_i\varepsilon$.

Example 4.2. Let G be the original Thompson group V. In particular, $r = 1$, $s = 1$, $n_1 = 2$. We can represent bases of $U_1(\Sigma)$ by finite rooted binary trees, and hence can choose X_1 to be a single point, and X_2 and X_3 to be the bases represented thus:



X_2 $\qquad\qquad$ X_3

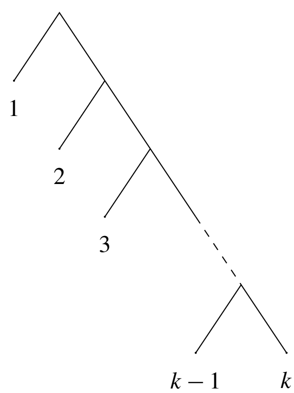
Suppose we take ε to be the expansion of X_2 on the left-hand leaf. This gives us X_3' as



and the corresponding element of V is x_0 as described after Remark 2.4.

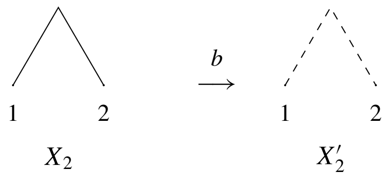
4B. ***The maximal tree $\mathcal{T}$.*** To be able to write down an explicit presentation, the choice for $\mathcal{T}$ becomes important. This relies heavily on the choice of representative for $\overline{X_i}$ above. This amounts to choosing a particular set of bases X_i in $U_r(\Sigma)$, where $i \equiv r \bmod d$.

Example 4.3. For $G = sV$ we have $r = 1$ and for each $k \in \mathbb{N}$ there is a basis X_k. Again, these can be represented by finite rooted binary trees. Now fix a colour $i \in S$ and choose the X_k as follows: we begin with X_1 our fixed one-element basis represented by a single point. Now X_2 is the basis obtained by applying the descending operation of colour i to X_1. We successively chose X_k as obtained from X_{k-1} by applying the descending operation of colour i to the last element of X_{k-1} and labelling the elements in successive order. The representation for X_k by a binary tree then looks as follows:



Notice that for $V_1(\Sigma)$ we can always choose the X_k to be represented by a rightmost tree as above, provided that all colours have the same arity. Now the construction shows that the maximal tree $\mathcal{T}$ in Z/G is a rooted infinite line.

For example, the baker's map $b \in 2V$ can easily be described using the bases chosen above. Let X_1 be a single point and X_2 be as in Example 4.2; note that we expanded with colour 1. Now we consider X_2' the basis obtained from X_1 by expanding once with colour 2 (represented by a dashed line). Hence this gives rise to the element $b \in 2V$.



For the general case with mixed arities we will not be able to find such a straightforward set of representatives X_i as before. We will show that we can, however, find a maximal tree $\mathcal{T}$ in Z/G, whose vertices are all but a finite number obtained by a step-by-step process beginning with our fixed basis X_r and then expanding the last element of a basis previously constructed.

Example 4.4. Let $V_r(\Sigma)$ be the group given by $r = 1,\ s = 2,\ n_1 = 5$ and $n_2 = 7$. Then $d = 2$ and our chosen set of bases is of the form

$$\{X_i \mid i \equiv 1 \bmod 2\}.$$

By simply expanding X_1 by the colours 1 and 2 respectively, we obtain X_5 and X_7. To obtain X_3 we could contract the last 5 elements of X_7 by colour 1, but there is no way to obtain X_3 from X_1 by simply expanding.

Remark 4.5. We now describe the construction of our preferred maximal tree $\mathcal{T}$ in Z/G, where $G = V_r(\Sigma)$ is the automorphism group of a valid and bounded Cantor algebra. We begin by showing that we can obtain all but finitely many of the bases

$$\{X_i \mid i \equiv r \bmod d\}$$

from X_r applying descending operations only. In other words

$$\left\{ r + \sum_{i=1}^{s} k_i(n_i - 1) \mid 0 \leq k_1, \ldots, k_s \right\} \cup P = \{r + kd \mid 0 \leq k\},$$

where P is a finite set of integers. To see this, observe first that the problem can be reduced to the case when $r = 0$ and $d = 1$. Now choose integers $k_1, \ldots, k_s$ such that

$$1 = \sum_{i=1}^{s} k_i(n_i - 1)$$

and use them to produce integers $m_1, \ldots, m_s$ with $0 \leq m_2, \ldots, m_s$ such that

$$1 = \sum_{i=1}^{s} m_i(n_i - 1).$$

Hence,

$$1 \equiv \sum_{i=2}^{s} m_i(n_i - 1) \bmod m_1.$$

Multiplying this expression by the integers $2, \dots, m_1$ we get positive numbers $a_1, \dots, a_{m_1}$ which are a complete set of representatives of the residues modulo m_1 and such that they all belong to $\sum_{i=1}^{s} \mathbb{N}(n_i - 1)$. Now, let m be any integer with $m \geq \max\{a_i \mid 1 \leq i \leq m_1\}$. Then for some such i, we have $m \equiv a_i$ modulo m_1 and therefore $m - a_i = lm_1$ for some $l \geq 0$. From this we deduce that m also belongs to $\sum_{i=1}^{s} \mathbb{N}(n_i - 1)$.

It is now easy to find a (nonmaximal) directed tree in Z/G having $\overline{X}_r$ as a root and such that the cardinalities of the vertices are precisely the set $r + \sum_{i=1}^{s} \mathbb{N}(n_i - 1)$. Here, a root is the only vertex of the tree from which all other vertices can be reached by paths respecting the directions of the edges. Moreover, we can do it in such a way that the descending operations are always applied to the last element of each tuple. There are only finitely many points of our space Z/G not in this tree. Choose one of them and consider a directed path from that point to some point of the tree. Adding this directed path we get a new tree which no longer has a single root in the above sense. If there are still points left, repeat the process. Eventually, we get a directed tree with the desired properties and with only finitely many roots.

4C. ***The presentation.*** Now we apply Theorem 4.1 to produce an explicit presentation. We do get an abstract group presentation but we can also write it down as a presentation in terms of elements given by pairs of ordered bases using the explicit isomorphism in Section 4A, which allows one to recognise the group elements in a much more familiar way. Recall that we have fixed a set of tuples

$$\{X_i \mid i \equiv r \bmod d\}$$

which are lifts of the nodes of our tree $\mathcal{T}$. Moreover there is a tree in Z that is a lift of $\mathcal{T}$.

By [Geoghegan 2008, Theorem 3.1.16], $\pi_1(Z/G)$ is generated by the edges in Z/G, i.e., by the 1-simplices $\overline{X}_i \xrightarrow{\bar{\varepsilon}} \overline{X}_j$ in Z/G. As we have seen before, these correspond to elements $g \in G$ which are given by pairs $(X_j, X_i\varepsilon)$ where ε is a set of descending operations.

There are two sets of relators:

(i) Relators of the form $\bar{\varepsilon} = 1$ whenever $\bar{\varepsilon}$ is an edge in the tree $\mathcal{T}$. This means that there are tuples X_i and X_j in $\mathcal{T}$ such that X_j is obtained from X_i performing the operations ε. The group element that corresponds to ε is (X_j, X_j).

(ii) Relators obtained from the boundaries of the 2-cells in Z/G. The 2-cells of Z/G come from 2-cells in Z and these are of the form $A_0 \precsim A_1 \precsim A_2$. Let ε_1

be the set of operations needed to obtain A_1 from A_0, ε_2 the set of operations needed to obtain A_2 from A_1 and ε the composition of ε_1 and ε_2. Passing down to the quotient Z/G we get a 2-cell with boundary labelled $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$ and $\bar{\varepsilon}$. So we have the relator

$$\bar{\varepsilon} = \bar{\varepsilon}_1 \bar{\varepsilon}_2.$$

All this means that this second set of relators consists of the "composition of paths". We want to write this down in terms of pairs of ordered bases. Let i be the cardinality of A_0 and j_1 the cardinality of A_1. The edge $\bar{\varepsilon}_1$ represents the element $g_1 = (X_{j_1}, X_i\varepsilon_1) \in G$. We may apply the descending operations ε_2 to this pair and then we observe that also $g_1 = (X_{j_1}\varepsilon_2, X_i\varepsilon_1\varepsilon_2)$. Note here that this follows from the definition of tree pair representation, and we do not need to impose any conditions on the presentation that we are building. Let j_2 be the cardinality of A_2, then $\bar{\varepsilon}_2$ represents the element $g_2 = (X_{j_2}, X_{j_1}\varepsilon_2)$ and $\bar{\varepsilon}$ represents $g = (X_{j_2}, X_i\varepsilon)$. So we get the relator $g = g_1 g_2$. In the particular case when $X_{j_1}\varepsilon_2$ belongs to the lift of our tree $\mathcal{T}$, or equivalently when $\bar{\varepsilon}_2$ belongs to $\mathcal{T}$, there is also a relator $g_2 = 1$ and we deduce $g = g_1$. This can also be seen using tree pairs: as X_{j_2} belongs to the prefixed set of nodes and has the same cardinality as $X_{j_1}\varepsilon_2$, we must have $X_{j_1}\varepsilon_2 = X_{j_2}$.

We may summarise as follows:

$$G = \langle W \mid R \rangle,$$

where

$$W = \{(X_j, X_i\varepsilon) \mid \varepsilon \text{ is a sequence of descending operations and } X_i \neq X_j\},$$

$$R = \{g = g_2 g_1 \mid g = (X_{j_2}, X_i\varepsilon),\ g_1 = (X_{j_1}, X_i\varepsilon_1),\ g_2 = (X_{j_2}, X_{j_1}\varepsilon_2),\ \varepsilon = \varepsilon_1\varepsilon_2\}.$$

Alternatively, we may delete those pairs $(X_j, X_i\varepsilon)$ where $\bar{\varepsilon}$ lies in the tree $\mathcal{T}$ from our list of generators.

4D. ***Reducing the generating set.*** A quick look to the generating set we have just obtained shows that it is far too big. Reducing it can be a complicated task but there is a reduction that seems natural: our generators come from edges in Z/G and these edges come from descending operations, so one expects that edges coming from "very elementary operations" should be enough. This is in fact the case but to make it more precise we need now some additional technicalities. Let us fix what should be called "very elementary" in our context. An edge $A_1 \overset{\varepsilon}{\precsim} A_2$ in Z is *very elementary* if it consists of a single operation, i.e., if it is either a permutation or it is a single descending operation (in this case $u(A_1) < u(A_2)$ is very elementary) but we do not allow composition of both. The case when $u(A_1) < u(A_2)$ will be termed *strict* and for these type of operations we will assume that if the original tuple is $(x_1, \dots, x_i)$ and we apply the descending operation α at the k-th element then the resulting tuple is

$$(x_1, \dots, x_{k-1}, x_k\alpha^1, \dots, x_k\alpha^{n_\alpha}, x_{k+1}, \dots, x_i).$$

Any ε can be written as a composition of very elementary operations. Of course it may happen that different sequences of operations give the same result when applied to the same tuple. This happens in the following four ways, which we shall refer to as *moves*:

(i) Disjoint type: we may apply two very elementary strict descending operations acting on distinct elements of a tuple and we get the same result regardless of the order of application of these two operations.

(ii) Σ type: we have different chains of elementary strict descending operations such that, up to a permutation, they give the same result when applied to any element of any tuple and which come from the defining relations for the algebra encoded in Σ.

(iii) Permutation-descending: we may first permute the elements of a tuple and then apply a very elementary strict descending operation or do it the other way around in a consistent manner and get the same result.

(iv) Permutation: the composition of two permutations is still a permutation.

Lemma 4.6. *Let A_1, A_2 be tuples. If two different chains of very elementary descending operations yield A_2 when applied to A_1, then one can be obtained from the other by repeated application of moves of the four types above.*

Proof. By making moves of types (iii) and (iv) only we may assume that our two chains are of the form

$$\varepsilon_1\varepsilon_2 \cdots \varepsilon_t\sigma,$$
$$\varepsilon_1'\varepsilon_2' \cdots \varepsilon_{t'}'\sigma',$$

where all $\varepsilon_i,\ \varepsilon_i'$ are very elementary and strict and $\sigma,\ \sigma'$ are permutations. Consider first what happens when we look at the underlying sets $u(A_1)$ and $u(A_2)$. The fact that both series of operations give the same set when applied to $u(A_1)$, implies that, for each particular element, we are either performing the same operation or the same operation up to applying some of the relators encoded in Σ. This means that $\varepsilon_1\varepsilon_2 \cdots \varepsilon_t$ can be transformed to $\varepsilon_1'\varepsilon_2' \cdots \varepsilon_{t'}'$ by making moves of types (i) or (ii) without taking the order of the elements into account. The fact that the relations in Σ involve certain permutations implies that what we really get is that via some extra moves of types (iii) and (iv), $\varepsilon_1\varepsilon_2 \cdots \varepsilon_t$ is transformed to $\varepsilon_1'\varepsilon_2' \cdots \varepsilon_{t'}'\tau$ for a certain permutation τ. So at this point our two sequences are

$$\varepsilon_1'\varepsilon_2' \cdots \varepsilon_{t'}'\tau\sigma, \qquad \varepsilon_1'\varepsilon_2' \cdots \varepsilon_{t'}'\sigma'.$$

The fact that both sequences yield A_2 when applied to A_1 implies that $B\tau\sigma = B\sigma'$ for $B = A_1\varepsilon'_1\varepsilon'_2\cdots\varepsilon'_{t'}$, which is a move of type (iv). $\square$

We next use Tietze transformations to change the presentation above. Essentially, what we need to do is the following: whenever there is a relator $g = g_1g_2$ we delete g from our set of generators. The effect of this transformation on the generating set is that we no longer have elements g coming from edges which are not very elementary. Moreover we will have only two kinds of generators: strict generators coming from strict very elementary edges, and finite order generators coming from permutations. We denote these sets by

$$W_s = \{(X_j, X_i\varepsilon) \mid \varepsilon \text{ is a very elementary strict expansion, } j = |X_i\varepsilon|\}$$

and call these *very elementary strict generators*. We also consider the elements of the set

$$W_p = \{(X_i, X_i\sigma) \mid \sigma \text{ is a permutation}\},$$

and call them *permutations*. From now on we will use the term *strict generators* for elements in W_s instead of the more precise very elementary strict generator.

The effect of this transformation on the set of relators is as follows: we no longer have to consider relators coming from edges in the tree. Whenever there are two sequences of very elementary operations that give the same A_2 when applied to some A_1, we have a new relator. Lemma 4.6 implies that these relators can be obtained from relators of the following types:

(i) R_D contains relators of the form $g_1g_2 = g'_2g'_1$ with g_1, g_2, g'_1, g'_2 strict generators coming from moves of disjoint type.

(ii) R_Σ contains relators between strict generators possibly followed by a permutation coming from moves of Σ type.

(iii) R_{PD} contains relators of the form $g\sigma = \sigma g$ with g a strict generator and σ a permutation coming from moves of type (iii).

(iv) R_P contains relators of the form $\sigma = \sigma_1\sigma_2$ with $\sigma,\ \sigma_1$ and σ_2 permutations coming from moves of type (iv).

Thus G admits the following (infinite) presentation:

$$\langle W_s \cup W_p \mid R_D \cup R_\Sigma \cup R_{PD} \cup R_P\rangle. \tag{1}$$

4E. ***Being more explicit.*** Let us consider an arbitrary strict generator $(X_j, X_i\varepsilon)$ associated to the strict edge $\bar{\varepsilon}$. It is completely determined by a triple (i, k, t) meaning that $\bar{\varepsilon}$ is obtained by applying the descending operation of colour t to the k-th element of an orbit representative of the set of tuples of order i. We will use the triple to denote the generator. Now we are going to write down explicitly what relators of disjoint type look like with this new notation. Recall that these relators come from very elementary strict descending and disjoint operations $\varepsilon_1, \varepsilon_2$ on one hand, and $\varepsilon_2', \varepsilon_1'$ on the other. They are such that

$$\varepsilon_1\varepsilon_2 = \varepsilon_2'\varepsilon_1',$$

where ε_1 and ε_1' are operations of the same colour, say t, whereas ε_2 and ε_2' are of colour s. Moreover ε_1 acts at the k_1-th and ε_2' acts at the k_2-th elements of X_i. We may assume that $k_1 < k_2$. Observe that this means that if we apply a descending operation to the k_2-th element first then the k_1-th element remains the same, but if we do it the other way around, i.e., apply a descending operation of colour t to the k_1-th element first, then the former k_2-th element becomes the (k_2+n_t-1)-th. Therefore the triples associated to each of $\bar\varepsilon_1, \bar\varepsilon_2, \bar\varepsilon_2', \bar\varepsilon_1'$ are

$$\begin{aligned}
&\bar\varepsilon_1 : (i, k_1, t) = (X_{i+n_t-1}, X_i\varepsilon_1),\\
&\bar\varepsilon_2 : (i+n_t-1, k_2+n_t-1, s) = (X_{i+n_t-1+n_s-1}, X_{i+n_t-1}\varepsilon_2),\\
&\bar\varepsilon_2' : (i, k_2, s) = (X_{i+n_s-1}, X_i\varepsilon_2'),\\
&\bar\varepsilon_1' : (i+n_s-1, k_1, t) = (X_{i+n_s-1+n_t-1}, X_{i+n_s-1}\varepsilon_1),
\end{aligned}$$

and our relator is

$$(i, k_1, t)(i+n_t-1, k_2+n_t-1, s) = (i, k_2, s)(i+n_s-1, k_1, t). \tag{2}$$

Analogously, it is possible to represent a generator $(X_i, \sigma(X_i))$ of "permutation type" using the pair (i, σ). Now, relators of type R_{PD} come from the fact that applying first a permutation and then a very elementary strict operation to a tuple, yields the same as doing it the other way around for a suitable permutation. More explicitly, assume that we start with the tuple X_i. Let ε be the operation associated to the triple, say, (i, k, t) and consider a permutation σ represented by the pair (i, σ). Slightly abusing notation view σ as a permutation of the numbers $\{1, \dots, i\}$. Starting with X_i and performing first the permutation σ and then applying the strict descending operation associated to $\bar\varepsilon' = (i, \sigma(k), t)$, yields the tuple $X_i\sigma\varepsilon'$ whose underlying set is the same as that of the tuple $X_i\varepsilon$. Therefore there is some permutation σ' such that the tuples $X_i\sigma\varepsilon'$ and $X_i\varepsilon\sigma'$ coincide. And this implies that we have a relator $\bar\sigma \cdot \bar\varepsilon' = \bar\varepsilon \cdot \bar\sigma'$ or

$$(i, \sigma)(i, \sigma(k), t) = (i, k, t)(i+n_t-1, \sigma'). \tag{3}$$

5. A finite generating set

In this section, we show that the generating system $W_s \cup W_p$ can be reduced to a finite one. We begin with W_s. We will use the following two particular cases of relators of disjoint type.

Case 1: Let (i, k, t) be a triple such that $i - k > n_l - 1$ for any colour l where we include the case $l = t$. Assume moreover that the terminal point of the associated edge in Z/G, i.e., $\overline{X}_{i+n_t-1}$, is not a root of the tree $\mathcal{T}$. Recall that this edge consists of applying a descending operation of colour t, which increases the cardinality in $n_t - 1$. Then there is some edge of $\mathcal{T}$ ending in $\overline{X}_{i+n_t-1}$. Let s be the colour of this last edge which is represented as a triple by $(i + n_t - n_s, i + n_t - n_s, s)$ (recall that we constructed the tree $\mathcal{T}$ in such a way that the last element of each tuple is always being expanded). Now, as $i - k > n_s - 1$ we deduce $k < i - n_s + 1$. Thus there is a relator of disjoint type such as in (2) but with $i - n_s + 1$ instead of i, k instead of k_1 and $i - n_s + 1$ instead of k_2. This relator is

$$(i - n_s{+}1, k, t)(i + n_t - n_s, i + n_t - n_s, s) = (i - n_s{+}1, i - n_s{+}1, s)(i, k, t).$$

Since there is also a relator

$$(i + n_t - n_s, i + n_t - n_s, s) = 1,$$

because it belongs to $\mathcal{T}$, we deduce

$$(i, k, t) = (i - n_s{+}1, i - n_s{+}1, s)^{-1}(i - n_s{+}1, k, t). \tag{4}$$

This means that (i, k, t) can be expressed in terms of triples with a smaller value of i.

Case 2: Let (i, k, t) be a triple such that $i \geq k \geq n_t + 1$. Then $k - n_t + 1 > 1$ and $i - n_t + 1 \geq 2$. This means that there is a relator of disjoint type such as in (2) but with $i - n_t + 1$ instead of i, 1 instead of k_1 and $k - n_t + 1$ instead of k_2. This relator is

$$(i - n_t{+}1, 1, t)(i, k, t) = (i - n_t{+}1, k - n_t{+}1, t)(i, 1, t).$$

From this we deduce

$$(i, k, t) = (i - n_t{+}1, 1, t)^{-1}(i - n_t{+}1, k - n_t{+}1, t)(i, 1, t), \tag{5}$$

meaning that (i, k, t) can be expressed in terms of triples with either a smaller value of i or with $k = 1$.

Observe now that arguing by induction on $i + k$, equations (4) and (5) imply that any element in W_s lies in the finite subgroup generated by the finite subset

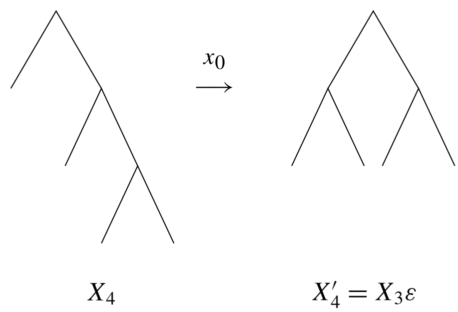
$$\{g \in W_s \mid \text{ the associated triple fails to fulfil both the conditions in Case 1 and in Case 2}\}.$$

Example 5.1. Let us consider the group V, i.e., where we have one colour t and $n_t = 2$. For now let us only concentrate on the strict generators W_s. Note that an element (i, i, t) is the identity. Looking at the representation by tree-pair diagrams, and the choice of X_i in Example 4.3, we see that we expand the rightmost leaf of the rightmost tree X_i, hence we obtain X_{i+1} and the group element is represented by (X_{i+1}, X_{i+1}), which is the identity. Now consider elements (i, k, t), where

$k < i-1$. Again, using the rightmost-tree, we see that after deleting unnecessary carets on the right, we get

$$(i, k, t) = (k+1, k, t),$$

which is exactly the relator (4). For example, consider $(3, 1, t)$. Then the corresponding tree-pair diagram is



In particular, after deleting the rightmost caret in each tree, this is exactly the element x_0, see the picture after Remark 2.4.

Writing

$$x_{i-2} = (i, i-1, t),$$

we recover the well-known infinite generating set $\{x_k \mid k \geq 0\}$ for $F < V$. Furthermore, this enables us to simplify the relator (2) above. We have

$$(i, k_1, t)(i+1, k_2+1, t) = (i, k_2, t)(i+1, k_1, t).$$

Using that $(i, k, t) = (k+1, k, t)$ for $k < i-1$, we get the well-known relator

$$x_k^{-1} x_l x_k = x_{l+1}$$

for any k and l. Moreover, observe that strict generators and disjoint relators give us the well-known infinite presentation of Thompson's group F; see [Cannon et al. 1996].

Now we want to reduce W_p in a similar way. The most natural way to do that is using relators of type R_{PD}, i.e., those mixing permutations and strict generators. To be able to argue by induction as before, we need to show that if i is big enough, any element of the form $(X_i, \sigma(X_i))$, where σ is a permutation, can be expressed in terms of permutations with a smaller i and possibly strict generators. As the group of permutations of the tuple X_i is generated by transpositions, we may assume that σ itself is a transposition. Now, assume that $i \geq 3n_t$ for t a colour with smallest possible arity n_t. As σ only moves two elements, we may find n_t consecutive elements in X_i which are untouched by σ. Let k be such that the k-th element in X_i is the first one of those n_t consecutive elements, and consider the strict generator associated to the triple $(i - n_t+1, k, t)$. Let σ' be the transposition of X_{i-n_t+1}

that moves precisely the elements that are also moved by σ. Then the associated relator (3) with $i - n_t + 1$ instead of i, and σ and σ' interchanged is

$$(i - n_t{+}1, \sigma')(i - n_t{+}1, k, t) = (i - n_t{+}1, k, t)(i, \sigma).$$

Thus

$$(i, \sigma) = (i - n_t{+}1, k, t)^{-1}(i - n_t{+}1, \sigma')(i - n_t{+}1, k, t) \tag{6}$$

as we wanted to show.

This discussion can be summarised as follows:

Theorem 5.2. *Assume that $U_r(\Sigma)$ is valid and bounded. Then $V_r(\Sigma)$ is generated by the finite set consisting of elements of the following three types*:

(1) *Strict generators associated to triples (i, k, t) with $i \leq n_t + 1$ and $i - k \leq n_s$ for any colour s.*

(2) *Strict generators associated to triples (i, k, t) such that $\overline{X}_{i+n_t-1}$ is a root of the tree $\mathcal{T}$.*

(3) *Permutations associated to pairs (i, σ) such that $i < 3n_t$ for some colour t.*

Example 5.3. Consider $G = V$. In Example 5.1, we have already recovered the infinite presentation for $F < V$. In the tree of Example 4.3, the triples have a single root X_1 so we do not have to consider generators as in item (2) of Theorem 5.2. As before, let $i \geq 2$ and denote by x_{i-2} the group element associated to the triple $(i, i{-}1, t)$. Then from Theorem 5.2 one deduces the well-known fact that the elements x_i, $i \geq 1$, together with the permutations generate the group and that x_0 and x_1 plus permutations are enough.

Remark 5.4. Similar generating systems can be obtained without using the space Z by proceeding as Burillo and Cleary did for the Brin–Thompson groups sV [Burillo and Cleary 2010, Theorem 2.1]. Instead of our first step (Section 4A), fix a set of tuples, one for each possible cardinality, which are to be the "source tree" of our tree pairs, and as "target tree" we allow anything that is obtained from one of these tuples by descending operations and permutations only. If $g \in G$ is an arbitrary element, it follows from the fact that any two bases have a common descendant that $g = (Y_1, Y_2)$ where Y_1 and Y_2 are obtained in that way. Then, choose X_i in our previously fixed set of tuples (what used to be the set of nodes in $\mathcal{T}$) of the same cardinality as Y_1 and Y_2, and observe that $g = g_2 g_1^{-1}$ with $g_1 = (X_i, Y_1)$ and $g_2 = (X_i, Y_2)$. These are precisely the type of elements we wanted to verify to be the generators of the group.

The choice of that fixed set of tuples can be the same as in Section 4B, but now we no longer need to construct the actual tree $\mathcal{T}$, we only need the nodes. For example, we can proceed as follows: as done before, fix a tuple X_r with r elements and choose integers $m_1, \dots, m_s$ with

$$d = \sum_{t=1}^{s} m_t(n_t - 1).$$

There is a sequence of operations (first descending, then ascending) that we can perform on the last element of X_r to get a new tuple with exactly $r + d$ elements that we denote $X_r\tau$. We may repeat the process to get a new tuple $X_r\tau^2$ and so on. We set $X_r\tau^0 := X_r$, let $X_{i+rd} := X_r\tau^i$ for $i \geq 0$ and take the obtained family as our prefixed set of "sources".

As seen above, our first set of generators is then

$$\{(X_j, X_i\varepsilon) \mid \varepsilon \text{ is a sequence of descending operations}\}.$$

Using Section 4D this can be further reduced to

$$\{(X_j, X_i\varepsilon) \mid \varepsilon \text{ is a single strict descending operation or permutation}\}.$$

Again, there is no serious need of the space Z to see that this reduction is possible. One can just check that composition of these elements corresponds to composition of the associated descending operations, in a way similar to that of [Burillo and Cleary 2010]. The same happens with the reduction performed in Section 4E: basically, we used Z only to have some identities available that allowed us to eliminate some elements from our generating system, but all those identities can be easily checked by hand and one gets the same finite set in the end.

6. Finite presentations

In this section, we still assume that $U_r(\Sigma)$ is valid and bounded and we add the extra hypothesis that it is also complete to exhibit a procedure that gives a finite presentation. To do that, we just replace Z by a truncated version Z^n and we use the results of Section 4 to obtain an explicit finite presentation.

Definition 6.1. Using the notation of Definition 2.1, suppose that for all $i \neq i'$, $i, i' \in S$ we have that $\Sigma_2^{i,i'} \neq \varnothing$ and that $f(j) = i'$ for all $j = 1, \dots, n_i$ and $f'(j') = i$ for all $j' = 1, \dots, n_{i'}$. Then we say that $U_r(\Sigma)$ is *complete*.

Considering the Morse function $t(A) = |A|$ in $\mathcal{S}_r(\Sigma)$ we can filter the complex with respect to t, and define the truncated Stein complex

$$\mathcal{S}_r(\Sigma)^n := \text{full subcomplex supported on } \{A \in \mathcal{S}_r(\Sigma) \mid t(A) \leq n\}.$$

In particular this is just the simplicial complex $\mathcal{S}_r(\Sigma)^n$ obtained by considering bases of cardinality bounded by n only. Note that in [Martínez-Pérez et al. 2016, Theorem 3.1] this complex was used to show that under the conditions above $V_r(\Sigma)$

is of type F_∞. The purpose of this section is to give a recipe for constructing explicit presentations.

Obviously, we can do the same with the complex Z and consider its truncated version Z^n where the tuples have at most n elements. The map u restricts to these truncated versions and the same argument as in Section 3B shows that there is a homotopy equivalence

$$u : Z^n \to |\mathcal{S}_r(\Sigma)^n|.$$

By [Fluch et al. 2013, Corollary 3.9] for the special case of sV and [Martínez-Pérez et al. 2016, Section 3] for the general case, assuming that $U_r(\Sigma)$ is valid, bounded and complete, there is some positive integer n_0 depending on Σ, such that for any $n \geq n_0$ and any basis $B \in \mathcal{S}_r(\Sigma)$ with cardinality $|B| = n+1$ the descending link of B in the Stein complex $\mathcal{S}_r(\Sigma)$ is simply connected. Using Morse theory ([Bestvina and Brady 1997, Corollary 2.6]) we deduce that for $n \geq n_0$ the inclusion

$$\mathcal{S}_r(\Sigma)^n \subseteq \mathcal{S}_r(\Sigma)^{n+1}$$

induces an isomorphism in π_1 and π_0. As the space $\mathcal{S}_r(\Sigma)$ is contractible we have

$$1 = \pi_1(\mathcal{S}_r(\Sigma)) = \lim \pi_1(\mathcal{S}_r(\Sigma)^n),$$
$$1 = \pi_0(\mathcal{S}_r(\Sigma)) = \lim \pi_0(\mathcal{S}_r(\Sigma)^n),$$

and $1 = \pi_1(\mathcal{S}_r(\Sigma)^n) = \pi_0(\mathcal{S}_r(\Sigma)^n)$ for $n \geq n_0$. From this we deduce that $\mathcal{S}_r(\Sigma)^n$ is path connected and simply connected for $n \geq n_0$. This, together with the fact that u is a homotopy equivalence, implies that the same holds true for Z^n. Finally, observe that Z^n being path connected implies that the same is true for Z^n/G. Therefore we can use Z^n instead of Z in Theorem 4.1 and as Z^n/G is finite we get a finite presentation. Hence we have the following theorem.

Theorem 6.2. *Let $U_r(\Sigma)$ be a valid, bounded and complete Cantor algebra, and let $n \geq 1$ be such that Z^n is simply connected. Then there is a finite presentation of $V_r(\Sigma)$ involving only strict generators (i, k, t) with $i + n_t - 1 \leq n$, permutations (i, σ) with $i \leq n$, and relators involving these generators only, and which is obtained by truncating the presentation*

$$\langle W_s \cup W_p \mid R_D \cup R_\Sigma \cup R_{PD} \cup R_P \rangle$$

given in (1).

The main difference with the reduction process of Section 5 is that we are now also reducing the set of relators. Moreover, the "truncated" set of generators in the finite presentation obtained this way can be further reduced using the same arguments as in Section 5.

Example 6.3. For $G = V$, in [Fluch et al. 2013, Corollary 3.9] there is an explicit condition on n that implies that Z^n is simply connected: we need

$$1 \le \left\lfloor \frac{n-1}{3} \right\rfloor - 1;$$

thus we can take $n = 7$. This means that the set of strict generators in Example 5.1 can be reduced to $x_0, \dots, x_4$ and the relators of disjoint type can be reduced to

$$x_k^{-1} x_l x_k = x_{l+1},$$

where $(k, l, l+1)$ is one of the following tuples: $(0, 1, 2)$, $(0, 2, 3)$, $(0, 3, 4)$, $(1, 2, 3)$, $(1, 3, 4)$, $(2, 3, 4)$. At this point it is not difficult to write down a finite presentation of V. Note also that in Example 5.3 we had already reduced to two strict generators x_0 and x_1.

Recently, Bleak and Quick [2017] found a short finite presentation for V with two generators and nine relations using different methods.

Using our methods we get a finite presentation of Thompson's group F, and by using Tietze moves this presentation can be transformed to the well-known

$$\langle x_0, x_1 \mid x_0^{-3} x_1 x_0^3 = x_1^{-1} x_0^{-2} x_1 x_0^2 x_1,\, x_0^{-2} x_1 x_0^2 = x_1^{-1} x_0^{-1} x_1 x_0 x_1 \rangle$$

Example 6.4. For $G = sV$ we can also use [Fluch et al. 2013, Corollary 3.9] to compute the value of n making Z^n simply connected: we need

$$1 \le \left\lfloor \frac{n-1}{2^s} \right\rfloor - 1,$$

thus we can take $n = 1 + 2^{s+1}$. Recall that when choosing the maximal tree in Z/G we chose expansion by one colour only (see Example 4.3). Let that colour be denoted by 1. For the same reason as in Example 5.1 we now have that elements of the form $(i, i, 1)$ are the identity, and that for any colour t and any $k < i - 1$, we have $(i, k, t) = (k+1, k, t)$.

This now gives an infinite W_s, which for $G = 2V$ can be listed as

$$(i+1, i, 1), \quad (i+1, i, 2), \quad \text{and} \quad (k, k, 2),$$

which corresponds to the infinite order generators A_{i-1}, B_{i-1} and C_i of Brin's infinite generating set of $2V$; see [Brin 2004] or [Burillo and Cleary 2010]. Now by Theorem 5.2(1), this can be reduced to a finite generating set with seven strict generators; those where $i \le 2$ and $k \le 3$, as well as a finite number of permutation generators. Using Theorem 6.2 without any further reductions, we get a finite presentation where $i \le 7$ and $k \le 8$.

7. Finite presentation for centralisers of finite subgroups

The proof of [Martínez-Pérez et al. 2016, Theorem 4.9] can be used to show that whenever the group $V_k(\Sigma)$ is finitely presented for any k, then so is $C_{V_r(\Sigma)}(Q)$ for any finite $Q \leq V_r(\Sigma)$, but the proof there does not yield an explicit finite presentation. In this section we are going to construct a finite presentation of $C_{V_r(\Sigma)}(Q)$. To do that, we proceed as follows. Note first that, by [Martínez-Pérez et al. 2016, Theorem 4.2], the group $C_{V_r(\Sigma)}(Q)$ is a direct product of groups of the form

$$\varinjlim(U_{r'}(\Sigma), L) \rtimes V_{r'}(\Sigma).$$

We now summarise the notation developed in [Martínez-Pérez et al. 2016]. The semidirect product above is associated to a fixed transitive permutation representation $\varphi : Q \to S_m$ of the finite group Q, where S_m is the symmetric group of degree m, the orbit length. Then L is the centraliser of the image $\varphi(Q)$ in S_m and thus is a finite group. The number r' depends on φ (see [Martínez-Pérez et al. 2016, Theorem 4.2]), but in order to simplify notation we will just set $r' = r$. The set of bases in $U_r(\Sigma)$ together with the expansion maps can be viewed as a directed graph and we let $(U_r(\Sigma), L)$ be the following diagram of groups associated to this graph: To each basis A we associate $\mathrm{Maps}(A, L)$, the group with elements the maps from A to L where the group operation is induced by multiplication in L. Each simple expansion $A \leq B$ corresponds to the diagonal map $\delta : \mathrm{Maps}(A, L) \to \mathrm{Maps}(B, L)$ with $\delta(f)(a\alpha_i^j) = f(a)$, where $a \in A$ is the expanded element. Then we consider the direct limit $\varinjlim(U_r(\Sigma), L)$ whose elements are determined by some basis A and a map $A \to L$. Observe that we may always assume that the basis A satisfies $X_r \leq A$.

We begin by studying presentations for $\varinjlim(U_r(\Sigma), L)$. We will obtain an infinite presentation (see Lemma 7.1 below) and then we will use the semidirect product action of $V_r(\Sigma)$ on this presentation together with the so-called Burnside procedure described in the Appendix to get a (finite) presentation of the group $\varinjlim(U_r(\Sigma), L) \rtimes V_r(\Sigma)$.

We begin by constructing a generating system for the group $\varinjlim(U_r(\Sigma), L)$. Take $x \in L$ and A a basis with $X_r \leq A$. Take some subset $A_1 \subseteq A$ and let $\chi_{A_1,x} \in \varinjlim(U_r(\Sigma), L)$ be the element that maps every $a \in A_1$ to x and every $a \in A \smallsetminus A_1$ to the identity $1 \in L$. It is easy to see that the set of all the elements of this form generates our group, but observe that there might be a uniqueness issue because if we had another basis C with $A \leq C$ and C_1 were the subset of those elements in C coming from elements in A_1, then $\chi_{A_1,x}$ would equal $\chi_{C_1,x}$. To avoid this problem we set

$$\omega(A_1) := \{b \text{ is a descendant of elements in } X_r \mid aw = bw' \text{ for some } a \in A_1 \text{ and descending words } w, w'\}$$

(this was denoted $A_1(\mathcal{L})$ in [Martínez-Pérez et al. 2016]) and

$$\Omega := \{\omega(A_1) \mid A_1 \text{ is a subset of some basis } A \geq X_r\}.$$

At first sight, this set Ω seems different from the set Ω defined in [Martínez-Pérez et al. 2016], which was defined for arbitrary finite subsets of the set of all descendants of elements in X_r, but Lemma 4.5(i) in that paper shows that since we are assuming that our Cantor algebra is valid and bounded they are in fact equal.

We set $\chi_{\omega,x} := \chi_{A_1,x}$, where $\omega = \omega(A_1)$. Observe that the proof of [Martínez-Pérez et al. 2016, Lemma 4.5(i)] also implies that $\omega(A_1) = \omega(C_1)$, provided that $A \leq C$ and C_1 is the subset of those elements in C coming from elements in A_1 (or, in other words, $C_1 = C \cap \omega(A_1)$). As a consequence one easily sees that for any B_1 subset of a basis B with $X_r \leq B$,

$$\chi_{A_1,x} = \chi_{B_1,x} \iff \omega(A_1) = \omega(B_1),$$

implying that $\chi_{\omega,x}$ is well defined.

We will need a bit more of the notation from [Martínez-Pérez et al. 2016]. Let $\omega \in \Omega$ and $A_1 \subseteq A \geq X_r$ with $\omega = \omega(A_1)$. We set

$$\|\omega\| = \begin{cases} t & \text{if } |A_1| \equiv t \bmod d \text{ with } 0 < t \leq d, \\ 0 & \text{if } \omega = \varnothing. \end{cases}$$

This does not depend on A_1; see [Martínez-Pérez et al. 2016, Lemma 4.5(v)]. Now, let $\omega_1, \omega_2 \in \Omega$ and $A_1, A_2 \subseteq A \geq X_r$ with $\omega_i = \omega(A_i)$ for $i = 1, 2$. Observe that the fact that our Cantor algebra is bounded means that we can always find such A_1 and A_2. If $A_1 \cap A_2 = \varnothing$, we write $\omega_1 \wedge \omega_2 = \varnothing$. Again, this is well defined, by [Martínez-Pérez et al. 2016, Lemma 4.5(vi)].

Lemma 7.1. *The following is a presentation of* $\varinjlim(U_r(\Sigma), L)$:

$$\langle (\chi_{\omega,x})_{\omega \in \Omega \smallsetminus \varnothing, x \in L} \mid \mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3 \rangle,$$

where

$$\mathcal{R}_1 = \{\chi_{\omega,xy}^{-1} \chi_{\omega,x} \chi_{\omega,y} \mid \omega \in \Omega, x, y \in L\},$$
$$\mathcal{R}_2 = \{[\chi_{\omega,x}, \chi_{\omega',y}] \mid \omega, \omega' \in \Omega, \omega \wedge \omega' = \varnothing\},$$
$$\mathcal{R}_3 = \{\chi_{\omega,x}^{-1} \chi_{\omega_1,x} \chi_{\omega_2,x} \mid \omega, \omega_1, \omega_2 \in \Omega, \omega = \omega_1 \dot\cup \omega_2\},$$

where $\omega_1 \dot\cup \omega_2$ *denotes the disjoint union. Moreover* $V_r(\Sigma)$ *acts by permutations with finitely many orbits on this presentation.*

Proof. As observed above, any $\chi \in \varinjlim(U_r(\Sigma), L)$ is a product of elements of the form $\chi_{\omega,x}$ for a suitable $\omega \in \Omega$ and $x \in L$. Let F denote the free group on the set $\{\widetilde{\chi}_{\omega,x} \mid \omega \in \Omega \smallsetminus \varnothing, x \in L\}$. There is an epimorphism

$$F \overset{\tau}{\twoheadrightarrow} \varinjlim(U_r(\Sigma), L)$$

with $\tau(\widetilde{\chi}_{\omega,x}) = \chi_{\omega,x}$. Let G be the abstract group defined in the statement of the result for the generators $\widetilde{\chi}_{\omega,x}$. It is immediate to verify that the epimorphism τ

defined above induces an epimorphism from G to $\varinjlim(U_r(\Sigma), L)$ which we still call τ. This follows since all relations inside G are easily verified to hold for the images $\tau(\widetilde{\chi}_{\omega,x})$. Assume that we have a word $\widetilde{w} = w(\widetilde{\chi}_{\omega_1,x_1}, \dots, \widetilde{\chi}_{\omega_k,x_k})$, for some $\omega_1, \dots, \omega_k \in \Omega$ and $x_1, \dots, x_k \in L$. Assume further that

$$1 = \tau(\widetilde{w}) = \tau(w(\widetilde{\chi}_{\omega_1,x_1}, \dots, \widetilde{\chi}_{\omega_k,x_k})) = w(\tau(\widetilde{\chi}_{\omega_1,x_1}), \dots, \tau(\widetilde{\chi}_{\omega_k,x_k})).$$

Let $X_r \leq A$ be a basis with subsets $A_i \subseteq A$ such that $\omega_i = A_i(\mathcal{L})$ for $i = 1, \dots, k$. We now refine the set $\{A_1, \dots, A_k\}$ to a set $\{A'_1, \dots, A'_{k'}\}$ of subsets of A such that for all $i, j \leq k'$ either $A'_i \cap A'_j = \varnothing$ or $A'_i = A'_j$. By suitably applying the relations in $\mathcal{R}_3$ to both the original word $w(\widetilde{\chi}_{\omega_1,x_1}, \dots, \widetilde{\chi}_{\omega_k,x_k})$ and its image

$$w := \tau(\widetilde{w}) = w(\chi_{\omega_1,x_1}, \dots, \chi_{\omega_k,x_k}),$$

we may rewrite each occurrence of χ_{ω_i,x_i} and $\widetilde{\chi}_{\omega_i,x_i}$ in terms of suitable new elements $\tau(\widetilde{\chi}_{\omega'_j,y_j})$ and $\chi_{\omega'_j,y_j}$ for $1 \leq j \leq k'$, so that either $\omega'_j \wedge \omega'_i = \varnothing$ or $\omega'_j = \omega'_i$.

Reordering them so that $\omega_1, \dots, \omega_u$ for $1 \leq u \leq k'$ are pairwise distinct and applying the relations in $\mathcal{R}_2$ and $\mathcal{R}_1$ to group together the suitable products of the y_j's we obtain new words

$$\widetilde{w} \sim \widetilde{w}' = \widetilde{\chi}_{\omega'_1,z_1} \cdots \widetilde{\chi}_{\omega'_u,z_u}, \qquad w \sim w' = \chi_{\omega'_1,z_1} \cdots \chi_{\omega'_u,z_u},$$

where the ω'_i's are pairwise disjoint.

If $w' \sim 1$, we must have $z_i = 1$ for any $1 \leq i \leq u$, by applying the word w' to an $a \in A_i$ such that $A_i(\mathcal{L}) = \omega'_i$. From $\mathcal{R}_1$ it is immediate to see that $\widetilde{\chi}_{\omega,1} = 1$ for any $\omega \in \Omega$ so we also have $\widetilde{w} \sim \widetilde{w}' \sim 1$ and G gives a presentation of $\varinjlim(U_r(\Sigma), L)$.

By [Martínez-Pérez et al. 2016, Lemma 4.7], the group $V_r(\Sigma)$ acts by permutations on Ω. Moreover, for any $g \in V_r(\Sigma)$, if $\omega, \omega' \in \Omega$ are such that $\omega \wedge \omega' = \varnothing$, then $g\omega \wedge g\omega' = \varnothing$ and if $\omega = \omega_1 \cup \omega_2$ for $\omega_1, \omega_2 \in \Omega$, then $g\omega = g\omega_1 \cup g\omega_2$. Therefore $V_r(\Sigma)$ acts by permutations on this presentation. To prove the last statement, it suffices to check the following:

Claim 1. The set of generators is $V_r(\Sigma)$-finite.

Claim 2. Each of the sets of relations $\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3$ is $V_r(\Sigma)$-finite.

As the group L is finite, both claims follow from slight variations of the proof of [Martínez-Pérez et al. 2016, Lemma 4.7]. For example, for Claim 2 for $\mathcal{R}_2$, it suffices to check that whenever we have $\omega, \omega', \widehat{\omega}, \widehat{\omega}' \in \Omega$ with

$$\omega \wedge \omega' = \varnothing, \quad \widehat{\omega} \wedge \widehat{\omega}' = \varnothing, \quad \|\omega\| = \|\widehat{\omega}\| \quad \text{and} \quad \|\omega'\| = \|\widehat{\omega}'\|,$$

then there is some $g \in V_r(\Sigma)$ such that for any $x \in L$, we have $\chi_{\widehat{\omega},x} = \chi^g_{\omega,x}$ and $\chi_{\widehat{\omega}',x} = \chi^g_{\omega',x}$. To get a suitable g, choose bases $X_r \leq A, \widehat{A}$ so that for $B, B' \subseteq A$

and $\widehat{B}, \widehat{B}' \subseteq \widehat{A}$, we have

$$\omega = \omega(B), \quad \omega' = \omega(B'), \quad \widehat{\omega} = \omega(\widehat{B}), \quad \widehat{\omega}' = \omega(\widehat{B}'),$$
$$|A| = |\widehat{A}|, \quad |B| = |\widehat{B}|, \quad |B'| = |\widehat{B}'|.$$

The assumptions imply that $B \cap B' = \varnothing = \widehat{B} \cap \widehat{B}'$. So we may choose a $g \in V_r(\Sigma)$ with $gA = \widehat{A}$, $gB = \widehat{B}$ and $gB' = \widehat{B}'$.

In a completely analogous way one proves that for $\omega, \omega_1, \omega_2, \widehat{\omega}, \widehat{\omega}_1, \widehat{\omega}_2 \in \Omega$ with

$$\omega = \omega_1 \cup \omega_2, \quad \widehat{\omega} = \widehat{\omega}_1 \cup \widehat{\omega}_2, \quad \|\omega\| = \|\widehat{\omega}\|, \quad \|\omega_1\| = \|\widehat{\omega}_1\|, \quad \|\omega_2\| = \|\widehat{\omega}_2\|,$$

there is some $g \in V_r(\Sigma)$ such that for any $x \in L$,

$$\chi_{\widehat{\omega},x} = \chi^g_{\omega,x}, \quad \chi_{\widehat{\omega}_1,x} = \chi^g_{\omega_1,x} \quad \text{and} \quad \chi_{\widehat{\omega}_2,x} = \chi^g_{\omega_2,x}. \qquad \square$$

Proposition 7.2. *Assume that the group $V_r(\Sigma)$ is finitely presented. Let $Q \le V_r(\Sigma)$ be a finite subgroup. Given a finite presentation of $V_r(\Sigma)$, Lemma 7.1 together with Theorem A.3 yield an explicit finite presentation of $C_{V_r(\Sigma)}(Q)$.*

Proof. By [Martínez-Pérez et al. 2016, Theorem 4.2], it suffices to construct an explicit finite presentation of a group of the form

$$H = \varinjlim(U_r(\Sigma), L) \rtimes V_r(\Sigma)$$

when L is an arbitrary finite group. Let $V_r(\Sigma) = \langle Z \mid T \rangle$ be a finite presentation of $V_r(\Sigma)$ and let

$$\varinjlim(U_r(\Sigma), L) = \langle Y \mid R \rangle$$

be the presentation constructed in Lemma 7.1. We need to verify the hypotheses of Theorem A.3. In Lemma 7.1 we have already checked that the group $V_r(\Sigma)$ acts by permutations in this presentation and that there are only finitely many orbits under that action. We may therefore choose $Y_0 \subseteq Y$ and $R_0 \subseteq R$ to be finite sets of representatives of these orbits.

The argument in Section A1 thus implies that the group H has the presentation

$$\langle Y_0, Z \mid \widehat{R}_0, T, [\mathrm{Stab}_{V_r(\Sigma)}(y), y], y \in Y_0 \rangle.$$

We can give explicit descriptions of possible choices for the sets Y_0, R_0. Set $X_r = \{x_1, \dots, x_r\}$ and let $\omega_i = \omega(\{x_1, \dots, x_i\})$ for $i = 1, \dots, r$. Then:

$$Y_0 = \{\chi_{\omega_i,z} \mid 1 \le i \le r, z \in L\}.$$

To describe R_0, we are going to split it into three pairwise disjoint subsets $R_0 = R_0^1 \cup R_0^2 \cup R_0^3$, according to the three subsets of relations $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ of Lemma 7.1. The simplest one is R_0^1:

$$R_0^1 = \{\chi^{-1}_{\omega_i,zy}\chi_{\omega_i,z}\chi_{\omega_i,y} \mid 1 \le i \le r, z \in L\}.$$

For R_0^2, R_0^3 it is more convenient to fix a basis $X_r \leq A$ with $|A| \geq 2r$. Then we may choose

$$R_0^2 = \{[\chi_{\omega,z}, \chi_{\omega',z}] \mid z \in L, \omega = \omega(A_1), \omega' = \omega(A_1'), A_1, A_1' \subseteq A, A_1 \cap A_1' = \varnothing\},$$
$$R_0^3 = \{\chi_{\omega,z}^{-1}\chi_{\omega_1,z}\chi_{\omega_2,z} \mid z \in L, \omega_1 = \omega(A_1), \omega_2 = \omega(A_2), \omega = \omega_1 \dot\cup \omega_2, A_1, A_2 \subseteq A\}.$$

Observe that these choices of R_0^2 and R_0^3 yield redundant presentations.

The previous presentation may not be finite because of all the relations needed to form $[\mathrm{Stab}_{V_r(\Sigma)}(y), y]$ where $y \in Y_0$. Notice that $g \in \mathrm{Stab}_{V_r(\Sigma)}(y)$ if and only if $g(\omega) = \omega$ where $y = \chi_{\omega,z}$ for some $z \in L$. By [Martínez-Pérez et al. 2016, Lemma 4.7] and the assumption on $V_r(\Sigma)$ we deduce that $\mathrm{Stab}_{V_r(\Sigma)}(y)$ is finitely generated by some generators $\mu_1, \dots, \mu_m$.

Consider now the following m relations, which are a subset of the stabiliser relations $[\mathrm{Stab}_{V_r(\Sigma)}(y), y]$:

$$\mu_i \chi_{\omega,z} \mu_i^{-1} = \chi_{\omega,z}, \quad i = 1, \dots, m. \tag{7}$$

If $g \in \mathrm{Stab}_{V_r(\Sigma)}(y)$, then $g = w(\mu_1, \dots, \mu_m)$ and the stabiliser relation $g\chi_{\omega,z}g^{-1} = \chi_{\omega,z}$ is thus obtained by starting from relation (7) for some i and then suitably conjugating this relation to build the word w.

Therefore, by Lemmas A.1 and A.2, the group H has the finite presentation

$$\langle Y_0, Z \mid \widehat{R}_0, T, [\mu_i, y], i = 1, \dots, m, y \in Y_0\rangle,$$

where the elements $\mu_1, \dots, \mu_m$ are expressed as words in the generators Z. □

Appendix: The Burnside procedure

We shall now give an outline of the Burnside procedure used in the proof of Proposition 7.2. As mentioned in the Introduction, we do not claim any originality for this. For example, this procedure has been used, without proof, in [Guralnick et al. 2011]. We are not aware of any place where a proof is presented. Hence we include it here for completeness.

The goal is to find a small finite presentation of a group, in the cases where the following procedure can be applied. The idea is to look for a possibly infinite, but well-behaved, presentation of a group G and a group Q such that the action of Q on the generators and relators of G cuts them down to a very small number. At a later stage, the group Q will be assumed to be a subgroup of G and its action will return a new smaller presentation.

A1. ***Preliminary lemmas.*** The beginning of this procedure is general and we only require each of the groups G and Q to have a presentation, without any assumption on them.

Let $G = \langle Y \mid R\rangle$ and $Q = \langle Z \mid T\rangle$ be groups. Let Q act on Y by permutations. Notice that $R \subseteq F(Y)$, where $F(Y)$ is the free group generated by Y, and observe that Q also acts on $F(Y)$. We assume that $Q(R) = R$. Let Y_0 be a set of representatives for the Q-orbits in Y and R_0 be a set of representatives for the Q-orbits in R. We observe that $R_0 \subseteq F(Y) = \langle\, t(a_0) \mid a_0 \in Y_0, t \in Q\,\rangle$, that is, we may express the elements of R_0 as products of the results of Q acting on elements of Y_0. In the special case that Q is a subgroup of G, we will be able to express elements in R_0 as products of conjugates of elements in Y_0 by elements in Q. Hence each element of R_0, seen as an element in G, can be written in more than one way and we fix an expression of the type $t_1(a_1)\cdots t_k(a_k)$ for such elements. We then define the set $\widehat{R_0} \subseteq \langle t a_0 t^{-1} \mid a_0 \in Y_0, t \in Q\rangle$ to be the set of fixed expressions for the elements of R_0, where we have replaced the action of Q on Y_0 by the conjugation of elements. That is, if $t_1(a_1)\cdots t_k(a_k)$ is a fixed expression in R_0, the corresponding element in $\widehat{R_0}$ is $t_1 a_1 t_1^{-1}\cdots t_k a_k t_k^{-1}$. The set $\widehat{R_0}$ is thus a set of formal expressions which will be used later to express relations in the groups.

Lemma A.1. *Following the notation previously defined, we have*

$$G \rtimes Q \cong \langle Y_0, Z \mid \widehat{R_0}, T, [\mathrm{Stab}_Q(y), y], y \in Y_0\rangle,$$

where the semidirect product is given by the action of Q on G as follows: for all $g_1, g_2 \in G$ and $t_1, t_2 \in Q$, multiplication is given by

$$(g_1, t_1)(g_2, t_2) = (g_1 \cdot t_1(g_2), t_1 t_2).$$

Proof. Let H be the group presented by $\langle Y_0, Z \mid \widehat{R_0}, T, [\mathrm{Stab}_Q(y), y], y \in Y_0\rangle$. Define the group homomorphism $\varphi : F(Y_0 \cup Z) \to G \rtimes Q$ by sending $a_0 \in Y_0$ to $(a_0, 1) \in G \rtimes Q$ and $c \in Z$ to $(1, c) \in G \rtimes Q$. By construction we see that

$$(*) \qquad \varphi(t)\varphi(a_0)\varphi(t)^{-1} = (t(a_0), 1)$$

for any word $t \in Q$.

Claim 1. The map φ induces a homomorphism $H \to G \rtimes Q$, which we still call φ.

Proof. If $d \in T$ is a relation in H, then $d = c_1 \cdots c_k$, for some $c_i \in Z$, and $\varphi(c_1)\cdots\varphi(c_k) = (1, 1)$. Let now $\widehat{b_0} \in \widehat{R_0}$ be a relation in H, then

$$\widehat{b_0} = t_1 a_1 t_1^{-1} \cdots t_k a_k t_k^{-1},$$

for some $a_i \in Y_0$ and $t_i \in Q$. Moreover, by applying $(*)$, we get

$$\prod_{i=1}^{k} \varphi(t_i)\varphi(a_i)\varphi(t_i)^{-1} = \left(\prod_{i=1}^{k} t_i(a_i), 1\right) = (1, 1).$$

Finally let $a_0 \in Y_0,\ t \in \mathrm{Stab}_Q(a_0)$. Thus we have, using $(*)$ again,

$$\varphi(t)\varphi(a_0)\varphi(t)^{-1}\varphi(a_0)^{-1} = (t(a_0), 1)(a_0^{-1}, 1) = (1, 1).$$

Now we just apply von Dyck's theorem. $\square$

Claim 2. The map φ is surjective.

Proof. Any element $(1, t) \in \{1\} \times Q := \{(1, s) \mid s \in Q\}$ can be written as $(1, c_1 \cdots c_k)$ for suitable $c_i \in Z$ and so $\varphi(H)$ contains $\{1\} \times Q$. We observe that any element of $G \times \{1\} := \{(h, 1) \mid h \in G\}$ can be written as $(t_1(a_1) \cdots t_k(a_k), 1)$ for suitable $a_i \in Y_0$ and $t_i \in Q$. By arguing as in Claim 1 we have $(g, 1) = \varphi(\prod_{i=1}^{k} t_i a_i t_i^{-1})$. Thus, $\varphi(H) \geq \langle G \times \{1\}, \{1\} \times Q\rangle = G \rtimes Q$. $\square$

Claim 3. The map φ is injective.

Proof. Any element of Y can be written as $t(a_0)$, for some $a_0 \in Y_0$ and $t \in Q$. Define $\overline{Y}^* = \{ta_0t^{-1} \mid a_0 \in Y_0, t \in Q\}$ to be the set of symbols of Y where we have replaced the action of Q with the conjugation of elements. We notice that, if $t(a_0) = s(a_0)$, then $t^{-1}s \in \mathrm{Stab}_Q(a_0)$ and we thus define an equivalence relation on $\overline{Y}^*$ by writing $ta_0t^{-1} \sim sa_0s^{-1}$ if and only if $t^{-1}s \in \mathrm{Stab}_Q(a_0)$. We define $\overline{Y} := \overline{Y}^*/\sim$ to be the collection of equivalence classes.

If $a \in Y$ and $a = t(a_0)$, for some $a_0 \in Y_0$ and $t \in Q$, we define an element $\bar{a}$ of $\overline{Y}$ by setting $\bar{a} = \{sa_0s^{-1} \mid t^{-1}s \in \mathrm{Stab}_Q(a_0)\}$. With this notation, we observe that Q acts on $\overline{Y}$ through

$$(s, \bar{a}) \to s \cdot \bar{a} := \overline{sta_0t^{-1}s^{-1}},$$

for some $a_0 \in Y_0, t \in Q$ such that $\bar{a} = \overline{ta_0t^{-1}}$. Also, notice that the map $\psi : Y \to \overline{Y}$ sending $a \mapsto \bar{a}$ is a Q-equivariant bijection, that is $\psi(sa) = s\psi(a) = s \cdot \bar{a}$ for all $s \in Q$. Hence the action of Q on Y is equivalent to the action of Q on $\overline{Y}$. For each element $\bar{a} \in \overline{Y}$ we can fix a representative $ta_0t^{-1} \in F(Y_0 \cup Z)$ and we call the set of representatives $\widehat{Y}$. By construction, every element $\widehat{b_0} \in \widehat{R_0}$ can be uniquely written as $\widehat{b_0} = t_1a_1t_1^{-1} \cdots t_ka_kt_k^{-1}$, so we define $\overline{R}_0 \subseteq F(\overline{Y})$ be the set of elements

$$\overline{t_1a_1t_1^{-1}} \cdots \overline{t_ka_kt_k^{-1}}.$$

We then let $\overline{R} \subseteq F(\overline{Y})$ be the set of all elements $\overline{tt_1a_1t_1^{-1}t^{-1}} \cdots \overline{tt_ka_kt_k^{-1}t^{-1}}$, for any $t \in Q$.

With these definitions, it makes sense to say that the normal closure $F(\overline{R})^{F(\overline{Y})}$ inside $F(\overline{Y})$ is isomorphic to $F(R)^{F(Y)}$ inside $F(Y)$. Also notice that

$$F(\overline{Y}) \cong F(\overline{Y}^*/\sim) = \langle \overline{Y}^* \mid R_\sim\rangle,$$

where $R_\sim$ is the set of all relations of the type $ta_0t^{-1} \sim sa_0s^{-1}$ if and only if $t^{-1}s \in \mathrm{Stab}_Q(a_0)$.

Let $w \in H$ be such that $\varphi(w) = (1, 1)$. Let $w = c_1 a_1 c_2 a_2 \cdots a_k c_{k+1}$ for $a_i \in Y_0$ and $c_i \in \langle Z \rangle$ and we rewrite w as

$$w = (c_1 a_1 c_1^{-1})(c_1 c_2 a_2 c_2^{-1} c_1^{-1}) \cdots (c_1 c_2 \cdots c_k a_k c_k^{-1} \cdots c_1^{-1}) c_1 c_2 \cdots c_k c_{k+1}.$$

Define $t_i = \underline{c_1 \cdots c_i}$. Then, up to replacing t_i with another suitable $t_i' \in Q$, we can assume that $t_i a_i t_i^{-1} \in \widehat{Y}$. Hence we can write $w = (t_1 a_1 t_1^{-1} \cdots t_k a_k t_k^{-1}) t_{k+1}$ and, applying φ to the rewriting of w we get $(1, 1) = (t_1(a_1) \cdots t_k(a_k), t_{k+1})$.

Since $t_{k+1} = 1$ inside Q, we can use the relations of Q to rewrite $t_{k+1} = 1$ inside H. Similarly, since $t_1(a_1) \cdots t_k(a_k) = 1$ inside G and since the normal closure $F(\bar{R})^{F(\bar{Y})}$ inside $F(\bar{Y})$ is isomorphic to $F(R)^{F(Y)}$ inside $F(Y)$, we can use the relations of G to rewrite $t_1 a_1 t_1^{-1} \cdots t_k a_k t_k^{-1} = 1$ inside H. Therefore $w = 1$ in H and so φ is injective. □

The map φ is thus a group isomorphism and the proof of Lemma A.1 is complete. □

The following result does not depend on the presentations of the relevant groups and relies only on the definition of semidirect product.

Lemma A.2. *Let G be a group and $Q \le G$. Let $G \rtimes Q$ be the semidirect product constructed using the action of Q on G by conjugation inside G. Then*

$$G \rtimes Q \cong G \times Q.$$

Proof. Let $H := G \rtimes Q$ with product given by $(a, x)(b, y) = (axbx^{-1}, xy)$. It is clear that $\widetilde{Q} = \{(t^{-1}, t) \mid t \in Q\}$ is a subgroup of H and $\widetilde{Q} \cong Q$. Since

$$(a, x) = (ax, 1)(x^{-1}, x),$$

H is generated by $G \times \{1\}$ and $\widetilde{Q}$. It is straightforward to verify that Q is normal and so, since $G \times \{1\}$ is normal as well, we get $G \rtimes Q \cong (G \times \{1\}) \times \widetilde{Q} \cong G \times Q$. □

A2. ***The Burnside procedure.*** We are now ready to explain the Burnside procedure. We make two additional assumptions with respect to those in Section A1. We assume

(i) the presentation $Q = \langle Z \mid T \rangle$ is finite,

(ii) the number of Q-orbits in Y is finite (and possibly very small, in practical applications),

(iii) the number of Q-orbits in R is finite (and also possibly very small),

(iv) the stabilisers $\mathrm{Stab}_Q(y)$ are finitely generated, for $y \in Y_0$.

Let G and Q be as defined in Lemma A.1, $Q \le G$ and let Q act by conjugation on G, then Lemmas A.1 and A.2 imply that

$$G \times Q \cong \langle Y_0, Z \mid \widehat{R}_0, T, [\mathrm{Stab}_Q(y), y] \text{ for } y \in Y_0 \rangle.$$

We rewrite Z in terms of Y_0 and then mod out Q. We also use the finite generation of $\mathrm{Stab}_Q(y)$ to rewrite the stabiliser relations as conjugations. Therefore we obtain the following theorem:

Theorem A.3 (Burnside procedure). *Let G, Q be the groups defined in Lemma A.1. Assume that*

(i) *$Q \le G$ and Q acts by conjugation on G,*

(ii) *$Q = \langle Z \mid T\rangle$ is finitely presented,*

(iii) *the number of Q-orbits in Y is finite,*

(iv) *the number of Q-orbits in R is finite,*

(v) *the stabilisers $\mathrm{Stab}_Q(y)$ are finitely generated, for $y \in Y_0$.*

Then there exists a finite presentation of G of the type

$$G = \Big\langle Y_0, Z \mid R_0, T, cyc^{-1} = y, \text{ for } y \in Y_0, \\ \text{a generator } c \text{ of } \mathrm{Stab}_Q(y), \text{finitely many extra relations}\Big\rangle,$$

where the extra relations are obtained in the following way: there is a relation for every element $c \in Z$ and it has the form

$$c = \text{word in conjugates of elements of } Y_0 \text{ by elements of } Z.$$

A3. ***An application.*** The following example is taken from [Guralnick et al. 2011]. Recall the following presentation for the alternating group

$$\mathrm{Alt}(n+2) = \langle x_1, \dots, x_p \mid (x_i)^3, (x_i x_j)^2, i \neq j\rangle,$$

where x_i can be realised as the 3-cycle $(i\ \ n{+}1\ \ n{+}2)$. Hence

$$\mathrm{Alt}(7) = \langle x_1, x_2, x_3, x_4, x_5 \mid (x_i)^3, (x_i x_j)^2, i \neq j\rangle := G.$$

On the other hand, it can be shown that

$$\mathrm{Alt}(5) = \langle a, b \mid a^5, b^2, (ab)^3\rangle := Q,$$

where a can be realised as $(1\,2\,3\,4\,5)$ and $b = (2\,3)(4\,5)$. Let $z := x_1 = (1\,6\,7)$ and observe that $x_i = z^{a^{i-1}}$, for $i = 1, \dots, 5$. Now we check that

$$Y = \{x_1, \dots, x_5\}, \quad Y_0 = \{z\}, \qquad R = \{(x_i)^3, (x_i x_j)^2, i \neq j\},$$
$$R_0 = \{z^3, (zz^a)^2\}, \qquad Z = \{a, b\}, \quad T = \{a^5, b^2, (ab)^3\}$$

satisfy the conditions of Theorem A.3. Noting that $\{[\mathrm{Stab}_Q(y), y] \text{ for } y \in Y_0\} = \{[z, b], [z, (ba)^a]\}$, we have

$$G \times Q = \langle a, b, z \mid a^5, b^2, (ab)^3, z^3, (zz^a)^2, [z, b], [z, (ba)^a]\rangle.$$

We can write $a = w_1(x_1, \ldots, x_5)$ and $b = w_2(x_1, \ldots, x_5)$, for suitable words $w_1, w_2 \in F(x_1, \ldots, x_5)$ and then Theorem A.3 yields the following finite presentation for Alt(7):

$$\mathrm{Alt}(7) = \big\langle a, b, z \mid R_0, T, [z, b], [z, (ba)^a], \\ a^{-1} w_1(z, z^a, \ldots, z^{a^4}), b^{-1} w_2(z, z^a, \ldots, z^{a^4}) \big\rangle.$$

Acknowledgments

We would like to thank Collin Bleak for helpful discussions related to [Bleak et al. 2013] and Martin Kassabov for introducing to us the Burnside procedure used in [Guralnick et al. 2011].

References

[Bestvina and Brady 1997] M. Bestvina and N. Brady, "Morse theory and finiteness properties of groups", *Invent. Math.* **129**:3 (1997), 445–470. MR Zbl

[Bleak and Quick 2017] C. Bleak and M. Quick, "The infinite simple group V of Richard J. Thompson: presentations by permutations", *Groups Geom. Dyn.* **11**:4 (2017), 1401–1436. MR Zbl

[Bleak et al. 2013] C. Bleak, H. Bowman, A. Gordon Lynch, G. Graham, J. Hughes, F. Matucci, and E. Sapir, "Centralizers in the R. Thompson group V_n", *Groups Geom. Dyn.* **7**:4 (2013), 821–865. MR Zbl

[Brin 2004] M. G. Brin, "Higher dimensional Thompson groups", *Geom. Dedicata* **108** (2004), 163–192. MR Zbl

[Brin and Squier 2001] M. G. Brin and C. C. Squier, "Presentations, conjugacy, roots, and centralizers in groups of piecewise linear homeomorphisms of the real line", *Comm. Algebra* **29**:10 (2001), 4557–4596. MR Zbl

[Burillo and Cleary 2010] J. Burillo and S. Cleary, "Metric properties of higher-dimensional Thompson's groups", *Pacific J. Math.* **248**:1 (2010), 49–62. MR Zbl

[Cannon et al. 1996] J. W. Cannon, W. J. Floyd, and W. R. Parry, "Introductory notes on Richard Thompson's groups", *Enseign. Math.* (2) **42**:3-4 (1996), 215–256. MR Zbl

[Fluch et al. 2013] M. G. Fluch, M. Marschler, S. Witzel, and M. C. B. Zaremsky, "The Brin–Thompson groups sV are of type F_∞", *Pacific J. Math.* **266**:2 (2013), 283–295. MR Zbl

[Geoghegan 2008] R. Geoghegan, *Topological methods in group theory*, Graduate Texts in Math. **243**, Springer, 2008. MR Zbl

[Guralnick et al. 2011] R. M. Guralnick, W. M. Kantor, M. Kassabov, and A. Lubotzky, "Presentations of finite simple groups: a computational approach", *J. Eur. Math. Soc.* **13**:2 (2011), 391–458. MR Zbl

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge Univ. Press, 2002. MR Zbl

[Hennig and Matucci 2012] J. Hennig and F. Matucci, "Presentations for the higher-dimensional Thompson groups nV", *Pacific J. Math.* **257**:1 (2012), 53–74. MR Zbl

[Higman 1974] G. Higman, *Finitely presented infinite simple groups*, Notes on Pure Math. **8**, Australian National Univ., Canberra, 1974. MR

[Kochloukova et al. 2013] D. H. Kochloukova, C. Martínez-Pérez, and B. E. A. Nucinkis, "Cohomological finiteness properties of the Brin–Thompson–Higman groups $2V$ and $3V$", *Proc. Edinb. Math. Soc.* (2) **56**:3 (2013), 777–804. MR Zbl

[Martínez-Pérez and Nucinkis 2013] C. Martínez-Pérez and B. E. A. Nucinkis, "Bredon cohomological finiteness conditions for generalisations of Thompson groups", *Groups Geom. Dyn.* **7**:4 (2013), 931–959. MR Zbl

[Martínez-Pérez et al. 2016] C. Martínez-Pérez, F. Matucci, and B. E. A. Nucinkis, "Cohomological finiteness conditions and centralisers in generalisations of Thompson's group V", *Forum Math.* **28**:5 (2016), 909–921. MR Zbl

[Quillen 1973] D. Quillen, "Higher algebraic K-theory, I", pp. 85–147 in *Algebraic K-theory, I: Higher K-theories* (Seattle, 1972), edited by H. Bass, Lecture Notes in Math. **341**, Springer, 1973. MR Zbl

[Stein 1992] M. Stein, "Groups of piecewise linear homeomorphisms", *Trans. Amer. Math. Soc.* **332**:2 (1992), 477–514. MR Zbl

Received January 17, 2017. Revised December 12, 2017.

CONCHITA MARTÍNEZ-PÉREZ
DEPARTAMENTO DE MATEMÁTICAS
UNIVERSIDAD DE ZARAGOZA
ZARAGOZA
SPAIN
conmar@unizar.es

FRANCESCO MATUCCI
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
UNIVERSIDADE ESTADUAL DE CAMPINAS
CAMPINAS
BRAZIL
francesco@ime.unicamp.br

BRITA NUCINKIS
DEPARTMENT OF MATHEMATICS
ROYAL HOLLOWAY
UNIVERSITY OF LONDON
UNITED KINGDOM
brita.nucinkis@rhul.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.405

LOCALIZATION FUNCTORS AND COSUPPORT IN DERIVED CATEGORIES OF COMMUTATIVE NOETHERIAN RINGS

TSUTOMU NAKAMURA AND YUJI YOSHINO

Let R be a commutative Noetherian ring. We introduce the notion of localization functors λ^W with cosupports in arbitrary subsets W of Spec$R$; it is a common generalization of localizations with respect to multiplicatively closed subsets and left derived functors of ideal-adic completion functors. We prove several results about the localization functors λ^W, including an explicit way to calculate λ^W using the notion of Čech complexes. As an application, we can give a simpler proof of a classical theorem by Gruson and Raynaud, which states that the projective dimension of a flat R-module is at most the Krull dimension of R. As another application, it is possible to give a functorial way to replace complexes of flat R-modules or complexes of finitely generated R-modules by complexes of pure-injective R-modules.

1. Introduction

Throughout this paper, we assume that R is a commutative Noetherian ring. We denote by $\mathcal{D} = D(\mathrm{Mod}\, R)$ the derived category of all complexes of R-modules, by which we mean that $\mathcal{D}$ is the unbounded derived category. For a triangulated subcategory $\mathcal{T}$ of $\mathcal{D}$, its left and right orthogonal subcategories are defined as

MSC2010: 13D09, 13D45, 55P60.
Keywords: colocalizing subcategory, cosupport, local homology.

${}^{\perp}\mathcal{T}=\{X\in\mathcal{D}\mid\mathrm{Hom}_{\mathcal{D}}(X,\mathcal{T})=0\}$ and $\mathcal{T}^{\perp}=\{Y\in\mathcal{D}\mid\mathrm{Hom}_{\mathcal{D}}(\mathcal{T},Y)=0\}$, respectively. Moreover, $\mathcal{T}$ is called localizing if $\mathcal{T}$ is closed under arbitrary direct sums, and colocalizing if it is closed under arbitrary direct products.

Recall that the support of a complex $X\in\mathcal{D}$ is defined as

$$\operatorname{supp} X=\{\mathfrak{p}\in\operatorname{Spec} R\mid X\otimes^{\mathrm{L}}_R\kappa(\mathfrak{p})\neq 0\},$$

where $\kappa(\mathfrak{p})=R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$. We write $\mathcal{L}_W=\{X\in\mathcal{D}\mid\operatorname{supp}X\subseteq W\}$ for a subset W of $\operatorname{Spec} R$. Then $\mathcal{L}_W$ is a localizing subcategory of $\mathcal{D}$. Neeman [1992] proved that any localizing subcategory of $\mathcal{D}$ is obtained in this way. The localization theory of triangulated categories [Krause 2010] yields a couple of adjoint pairs (i_W,γ_W) and (λ_W,j_W) as it is indicated in the following diagram:

$$\mathcal{L}_W \underset{\gamma_W}{\overset{i_W}{\rightleftarrows}} \mathcal{D} \underset{j_W}{\overset{\lambda_W}{\rightleftarrows}} \mathcal{L}_W^{\perp} \tag{1.1}$$

Here, i_W and j_W are the inclusion functors $\mathcal{L}_W\hookrightarrow\mathcal{D}$ and $\mathcal{L}_W^{\perp}\hookrightarrow\mathcal{D}$, respectively. In [Nakamura and Yoshino 2018], we introduced the colocalization functor with support in W as the functor γ_W. If V is a specialization-closed subset of $\operatorname{Spec} R$, then γ_V coincides with the right derived functor $\mathrm{R}\Gamma_V$ of the section functor Γ_V with support in V; it induces the local cohomology functors $H^i_V(-)=H^i(\mathrm{R}\Gamma_V(-))$. In [loc. cit.], we established some methods to compute γ_W for general subsets W of $\operatorname{Spec} R$. Furthermore, the local duality theorem and Grothendieck type vanishing theorem of local cohomology were extended to the case of γ_W.

On the other hand, in this paper, we introduce the notion of localization functors with cosupports in arbitrary subsets W of $\operatorname{Spec} R$. Recall that the cosupport of a complex $X\in\mathcal{D}$ is defined as

$$\operatorname{cosupp} X=\{\mathfrak{p}\in\operatorname{Spec} R\mid\mathrm{RHom}_R(\kappa(\mathfrak{p}),X)\neq 0\}.$$

We write $\mathcal{C}^W=\{X\in\mathcal{D}\mid\operatorname{cosupp}X\subseteq W\}$ for a subset W of $\operatorname{Spec} R$. Then $\mathcal{C}^W$ is a colocalizing subcategory of $\mathcal{D}$. Neeman [2011] proved that any colocalizing subcategory of $\mathcal{D}$ is obtained in this way.[1]

We remark that there are equalities

$${}^{\perp}\mathcal{C}^W=\mathcal{L}_{W^c},\qquad \mathcal{C}^W=\mathcal{L}_{W^c}^{\perp}, \tag{1.2}$$

where $W^c=\operatorname{Spec} R\setminus W$. The second equality follows from [Neeman 1992, Theorem 2.8], which states that $\mathcal{L}_{W^c}$ is equal to the smallest localizing subcategory of $\mathcal{D}$ containing the set $\{\kappa(\mathfrak{p})\mid\mathfrak{p}\in W^c\}$. Then it is seen that the first equality holds, since ${}^{\perp}(\mathcal{L}_{W^c}^{\perp})=\mathcal{L}_{W^c}$ (see [Krause 2010, §4.9]).

[1] This result is not needed in this work.

Now we write $\lambda^W = \lambda_{W^c}$ and $j^W = j_{W^c}$. By (1.1) and (1.2), there is a diagram of adjoint pairs:

$$
{}^{\perp}\mathcal{C}^W = \mathcal{L}_{W^c} \underset{\gamma_{W^c}}{\overset{i_{W^c}}{\rightleftarrows}} \mathcal{D} \underset{j^W}{\overset{\lambda^W}{\rightleftarrows}} \mathcal{C}^W = \mathcal{L}_{W^c}^{\perp}
$$

We call λ^W *the localization functor with cosupport in* W.

For a multiplicatively closed subset S of R, the localization functor λ^{U_S} with cosupport in U_S is nothing but $(-) \otimes_R S^{-1}R$, where $U_S = \{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{p} \cap S = \varnothing\}$. Moreover, for an ideal $\mathfrak{a}$ of R, the localization functor $\lambda^{V(\mathfrak{a})}$ with cosupport in $V(\mathfrak{a})$ is isomorphic to the left derived functor $\mathrm{L}\Lambda^{V(\mathfrak{a})}$ of the $\mathfrak{a}$-adic completion functor $\Lambda^{V(\mathfrak{a})} = \varprojlim(- \otimes_R R/\mathfrak{a}^n)$ defined on $\operatorname{Mod} R$. See Section 2 for details.

In this paper, we establish several results about the localization functor λ^W with cosupport in a general subset W of $\operatorname{Spec} R$.

In Section 3, we prove that λ^W is isomorphic to $\prod_{\mathfrak{p} \in W} \mathrm{L}\Lambda^{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}})$ if there is no inclusion relation between two distinct prime ideals in W. Furthermore, we give a method to compute λ^W for a general subset W. We write $\eta^W : \mathrm{id}_{\mathcal{D}} \to \lambda^W$ $(= j^W \lambda^W)$ for the natural morphism given by the adjointness of (λ^W, j^W). In addition, note that when $W_0 \subseteq W$, there is a morphism $\eta^{W_0}\lambda^W : \lambda^W \to \lambda^{W_0}\lambda^W \cong \lambda^{W_0}$. The following theorem is one of the main results of this paper.

Theorem 1.3 (Theorem 3.15). *Let* W, W_0 *and* W_1 *be subsets of* $\operatorname{Spec} R$ *with* $W = W_0 \cup W_1$. *We denote by* $\overline{W_0}^s$ (*resp.* $\overline{W_1}^g$) *the specialization* (*resp. generalization*) *closure of* W. *Suppose that one of the following conditions holds*:

(1) $W_0 = \overline{W_0}^s \cap W$.

(2) $W_1 = W \cap \overline{W_1}^g$.

Then, *for any* $X \in \mathcal{D}$, *there is a triangle*

$$
\lambda^W X \xrightarrow{f} \lambda^{W_1} X \oplus \lambda^{W_0} X \xrightarrow{g} \lambda^{W_1}\lambda^{W_0} X \longrightarrow \lambda^W X[1],
$$

where

$$
f = \begin{pmatrix} \eta^{W_1}\lambda^W X \\ \eta^{W_0}\lambda^W X \end{pmatrix}, \quad g = \begin{pmatrix} \lambda^{W_1}\eta^{W_0} X & (-1)\cdot \eta^{W_1}\lambda^{W_0} X \end{pmatrix}.
$$

This theorem enables us to compute λ^W by using λ^{W_0} and λ^{W_1} for smaller subsets W_0 and W_1. Furthermore, as long as we consider the derived category $\mathcal{D}$, this theorem and Theorem 3.22 generalize Mayer–Vietoris triangles by Benson, Iyengar and Krause [Benson et al. 2008, Theorem 7.5].

In Section 4, as an application, we give a simpler proof of a classical theorem due to Gruson and Raynaud. The theorem states that the projective dimension of a flat R-module is at most the Krull dimension of R.

Section 5 contains some basic facts about cotorsion flat R-modules.

Section 6 is devoted to studying the cosupport of a complex X consisting of cotorsion flat R-modules. As a consequence, we can calculate $\gamma_{V^c}X$ and $\lambda^V X$ explicitly for a specialization-closed subset V of $\operatorname{Spec} R$.

In Section 7, using Theorem 1.3 above, we give a new way to get λ^W. In fact, provided that $d = \dim R$ is finite, we are able to calculate λ^W by a Čech complex of functors of the form

$$\prod_{0\le i\le d} \bar{\lambda}^{W_i} \longrightarrow \prod_{0\le i<j\le d} \bar{\lambda}^{W_j}\bar{\lambda}^{W_i} \longrightarrow \cdots \longrightarrow \bar{\lambda}^{W_d}\cdots\bar{\lambda}^{W_0},$$

where $W_i = \{\mathfrak{p} \in W \mid \dim R/\mathfrak{p} = i\}$ and $\bar{\lambda}^{W_i} = \prod_{\mathfrak{p}\in W_i} \Lambda^{V(\mathfrak{p})}(-\otimes_R R_{\mathfrak{p}})$ for $0 \le i \le d$. This Čech complex sends a complex X of R-modules to a double complex in a natural way. We shall prove that $\lambda^W X$ is isomorphic to the total complex of the double complex if X consists of flat R-modules.

Section 8 treats commutativity of λ^W with tensor products. Consequently, we show that $\lambda^W Y$ can be computed by using the Čech complex above if Y is a complex of finitely generated R-modules.

In Section 9, as an application, we give a functorial way to construct quasi-isomorphisms from complexes of flat R-modules, or complexes of finitely generated R-modules to complexes of pure-injective R-modules.

2. Localization functors

In this section, we summarize some notions and basic facts used in the later sections.

We write $\operatorname{Mod} R$ for the category of all modules over a commutative Noetherian ring R. For an ideal $\mathfrak{a}$ of R, $\Lambda^{V(\mathfrak{a})}$ denotes the $\mathfrak{a}$-adic completion functor $\varprojlim(-\otimes_R R/\mathfrak{a}^n)$ defined on $\operatorname{Mod} R$. Moreover, we also denote by $M^\wedge_\mathfrak{a}$ the $\mathfrak{a}$-adic completion $\Lambda^{V(\mathfrak{a})}M = \varprojlim M/\mathfrak{a}^n M$ of an R-module M. If the natural map $M \to M^\wedge_\mathfrak{a}$ is an isomorphism, then M is called $\mathfrak{a}$-adically complete. In addition, when R is a local ring with maximal ideal $\mathfrak{m}$, we simply write $\widehat{M}$ for the $\mathfrak{m}$-adic completion of M.

We start with the following proposition.

Proposition 2.1. *Let $\mathfrak{a}$ be an ideal of R. If F is a flat R-module, then so is $F^\wedge_\mathfrak{a}$.*

As stated in [Simon 1990, 2.4], this fact is known. For the reader's convenience, we mention that this proposition follows from the two lemmas below.

Lemma 2.2. *Let $\mathfrak{a}$ be an ideal of R and F be a flat R-module. We consider a short exact sequence of finitely generated R-modules*

$$0 \to L \to M \to N \to 0.$$

Then

$$0 \to (F\otimes_R L)^\wedge_\mathfrak{a} \to (F\otimes_R M)^\wedge_\mathfrak{a} \to (F\otimes_R N)^\wedge_\mathfrak{a} \to 0$$

is exact.

Lemma 2.3. *Let $\mathfrak{a}$ and F be as above. Then we have a natural isomorphism,*

$$(F \otimes_R M)^{\wedge}_{\mathfrak{a}} \cong F^{\wedge}_{\mathfrak{a}} \otimes_R M,$$

for any finitely generated R-module M.

Using the Artin–Rees lemma and [Bourbaki 1961, Chap. I, §2.6, Proposition 6], we can prove Lemma 2.2, from which we obtain Lemma 2.3. Furthermore, Lemmas 2.2 and 2.3 imply that $F^{\wedge}_{\mathfrak{a}} \otimes_R (-)$ is an exact functor from the category of finitely generated R-modules to Mod R. Therefore Proposition 2.1 holds. □

It is also possible to show that $F^{\wedge}_{\mathfrak{a}}$ is flat over $R^{\wedge}_{\mathfrak{a}}$ by the same argument as above.

If R is a local ring with maximal ideal $\mathfrak{m}$, then $\mathfrak{m}$-adically complete flat R-modules are characterized as follows:

Lemma 2.4. *Let $(R, \mathfrak{m}, k)$ be a local ring and F a flat R-module. Set $B = \dim_k F/\mathfrak{m}F$. Then there is an isomorphism*

$$\widehat{F} \cong \widehat{\bigoplus_B R},$$

where $\bigoplus_B R$ is the direct sum of B-copies of R.

This lemma is proved in [Raynaud and Gruson 1971, Part. II, Proposition 2.4.3.1]. See also [Enochs and Jenda 2000, Lemma 6.7.4].

As in the introduction, we denote by $\mathcal{D} = D(\operatorname{Mod} R)$ the derived category of all complexes of R-modules. We write complexes X cohomologically:

$$X = (\cdots \to X^{i-1} \to X^i \to X^{i+1} \to \cdots).$$

For a complex P of R-modules, we say that P is K-projective if $\operatorname{Hom}_R(P, -)$ preserves acyclicity of complexes, where a complex is called acyclic if all its cohomology modules are zero. Similarly, for a complex F of R-modules, we say that F is K-flat if $(-) \otimes_R F$ preserves acyclicity of complexes.

Let $\mathfrak{a}$ be an ideal of R and $X \in \mathcal{D}$. If P is a K-projective resolution of X, then we have $\mathrm{L}\Lambda^{V(\mathfrak{a})} X \cong \Lambda^{V(\mathfrak{a})} P$. Moreover, $\mathrm{L}\Lambda^{V(\mathfrak{a})} X$ is also isomorphic to $\Lambda^{V(\mathfrak{a})} F$ if F is a K-flat resolution of X. Further, it is known that the following proposition holds.

Proposition 2.5. *Let $\mathfrak{a}$ be an ideal of R and X be a complex of flat R-modules. Then $\mathrm{L}\Lambda^{V(\mathfrak{a})} X$ is isomorphic to $\Lambda^{V(\mathfrak{a})} X$.*

Proof. To show this, we note there is an integer $n \geq 0$ such that $H^i(\mathrm{L}\Lambda^{V(\mathfrak{a})} M) = 0$ for all $i > n$ and all R-modules M, see [Greenlees and May 1992, Theorem 1.9] or [Alonso Tarrío et al. 1997, p. 15]. Using this fact, we can show that $\Lambda^{V(\mathfrak{a})}$ preserves acyclicity of complexes of flat R-modules. Then it is straightforward to see that $\mathrm{L}\Lambda^{V(\mathfrak{a})} X$ is isomorphic to $\Lambda^{V(\mathfrak{a})} X$. □

Let W be any subset of Spec R. Recall that γ_W denotes a right adjoint to the inclusion functor $i_W : \mathcal{L}_W \hookrightarrow \mathcal{D}$, and λ^W denotes a left adjoint to the inclusion functor $j^W : \mathcal{C}^W \hookrightarrow \mathcal{D}$. Moreover, γ_W and λ^W are identified with $i_W\gamma_W$ and $j^W\lambda^W$, respectively. We write $\varepsilon_W : \gamma_W \to \mathrm{id}_{\mathcal{D}}$ and $\eta^W : \mathrm{id}_{\mathcal{D}} \to \lambda^W$ for the natural morphisms induced by the adjointness of (i_W, γ_W) and (λ^W, j^W), respectively.

Note that $\lambda^W\eta^W$ (resp. $\gamma_W\varepsilon_W$) is invertible, and the equality $\lambda^W\eta^W = \eta^W\lambda^W$ (resp. $\gamma_W\varepsilon_W = \varepsilon_W\gamma_W$) holds, i.e., λ^W (resp. γ_W) is a localization (resp. colocalization) functor on $\mathcal{D}$. See [Krause 2010] for more details. In this paper, we call λ^W the localization functor with cosupport in W.

Using (1.2), we restate [Nakamura and Yoshino 2018, Lemma 2.1] as follows.

Lemma 2.6. *Let W be a subset of* Spec R. *For any $X \in \mathcal{D}$, there is a triangle of the following form*:

$$\gamma_{W^c}X \xrightarrow{\varepsilon_{W^c}X} X \xrightarrow{\eta^W X} \lambda^W X \longrightarrow \gamma_{W^c}X[1].$$

Furthermore, *if*

$$X' \longrightarrow X \longrightarrow X'' \longrightarrow X'[1]$$

is a triangle with $X' \in {}^{\perp}\mathcal{C}^W = \mathcal{L}_{W^c}$ and $X'' \in \mathcal{C}^W = \mathcal{L}_{W^c}^{\perp}$, then there exist unique isomorphisms $a : \gamma_{W^c}X \to X'$ and $b : \lambda^W X \to X''$ such that the following diagram is commutative:

$$\begin{array}{ccccccc}
\gamma_{W^c}X & \xrightarrow{\varepsilon_{W^c}X} & X & \xrightarrow{\eta^W X} & \lambda^W X & \longrightarrow & \gamma_{W^c}X[1] \\
\downarrow{\scriptstyle a} & & \| & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle a[1]} \\
X' & \longrightarrow & X & \longrightarrow & X'' & \longrightarrow & X'[1]
\end{array}$$

Remark 2.7. (i) Let $X \in \mathcal{D}$ and W be a subset of Spec R. By Lemma 2.6, X belongs to ${}^{\perp}\mathcal{C}^W = \mathcal{L}_{W^c}$ if and only if $\lambda^W X = 0$. This is equivalent to saying that $\lambda^{\{\mathfrak{p}\}}X = 0$ for all $\mathfrak{p} \in W$, since ${}^{\perp}\mathcal{C}^W = \mathcal{L}_{W^c} = \bigcap_{\mathfrak{p}\in W} \mathcal{L}_{\{\mathfrak{p}\}^c} = \bigcap_{\mathfrak{p}\in W} {}^{\perp}\mathcal{C}^{\{\mathfrak{p}\}}$.

(ii) Let W_0 and W be subsets of Spec R with $W_0 \subseteq W$. It follows from the uniqueness of adjoint functors that

$$\lambda^{W_0}\lambda^W \cong \lambda^{W_0} \cong \lambda^W\lambda^{W_0};$$

see also [Nakamura and Yoshino 2018, Remark 3.7(i)].

Now we give a typical example of localization functors. Let S be a multiplicatively closed subset S of R, and set $U_S = \{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{p} \cap S = \varnothing\}$. It is known that the localization functor λ^{U_S} with cosupport in U_S is nothing but $(-) \otimes_R S^{-1}R$. For the reader's convenience, we give a proof of this fact. Let $X \in \mathcal{D}$. It is clear that cosupp $X \otimes_R S^{-1}R \subseteq U_S$, or equivalently, $X \otimes_R S^{-1}R \in \mathcal{C}^{U_S}$. Moreover, embedding the natural morphism $X \to X \otimes_R S^{-1}R$ into a triangle,

$$C \longrightarrow X \longrightarrow X \otimes_R S^{-1}R \longrightarrow C[1],$$

we have $C \otimes_R S^{-1}R = 0$. This yields an inclusion relation $\operatorname{supp} C \subseteq (U_S)^c$. Hence it holds that $C \in \mathcal{L}_{(U_S)^c}$. Since we have shown that $C \in \mathcal{L}_{(U_S)^c}$ and $X \otimes_R S^{-1}R \in \mathcal{C}^{U_S}$, it follows from Lemma 2.6 that $\lambda^{U_S} X \cong X \otimes_R S^{-1}R$. Therefore we obtain the isomorphism

$$\lambda^{U_S} \cong (-) \otimes_R S^{-1}R. \tag{2.8}$$

For $\mathfrak{p} \in \operatorname{Spec} R$, we write $U(\mathfrak{p}) = \{\mathfrak{q} \in \operatorname{Spec} R \mid \mathfrak{q} \subseteq \mathfrak{p}\}$. If $S = R \setminus \mathfrak{p}$, then $U(\mathfrak{p})$ is equal to U_S, so that $\lambda^{U(\mathfrak{p})} \cong (-) \otimes_R R_\mathfrak{p}$ by (2.8). We remark that $\lambda^{U(\mathfrak{p})} = \lambda_{U(\mathfrak{p})^c}$ is written as $L_{Z(\mathfrak{p})}$ in [Benson et al. 2008], where $Z(\mathfrak{p}) = U(\mathfrak{p})^c$.

There is another important example of localization functors. Let $\mathfrak{a}$ be an ideal of R. It was proved by [Greenlees and May 1992] and [Alonso Tarrío et al. 1997] that $\mathrm{L}\Lambda^{V(\mathfrak{a})} : \mathcal{D} \to \mathcal{D}$ is a right adjoint to $\mathrm{R}\Gamma_{V(\mathfrak{a})} : \mathcal{D} \to \mathcal{D}$. In [Nakamura and Yoshino 2018, Proposition 5.1], using the adjointness property of $(\mathrm{R}\Gamma_{V(\mathfrak{a})}, \mathrm{L}\Lambda^{V(\mathfrak{a})})$, we proved that $\lambda^{V(\mathfrak{a})} = \lambda_{V(\mathfrak{a})^c}$ coincides with $\mathrm{L}\Lambda^{V(\mathfrak{a})}$. Hence there is an isomorphism

$$\lambda^{V(\mathfrak{a})} \cong \mathrm{L}\Lambda^{V(\mathfrak{a})}. \tag{2.9}$$

The functor $H_i^{\mathfrak{a}}(-) = H^{-i}(\mathrm{L}\Lambda^{V(\mathfrak{a})}(-))$ is called the i-th local homology functor with respect to $\mathfrak{a}$.

A subset W of $\operatorname{Spec} R$ is said to be specialization-closed (resp. generalization-closed) provided that the following condition holds: if $\mathfrak{p} \in W$ and $\mathfrak{q} \in \operatorname{Spec} R$ with $\mathfrak{p} \subseteq \mathfrak{q}$ (resp. $\mathfrak{p} \supseteq \mathfrak{q}$), then $\mathfrak{q} \in W$.

If V is a specialization-closed subset, then we have

$$\gamma_V \cong \mathrm{R}\Gamma_V; \tag{2.10}$$

see [Lipman 2002, Appendix 3.5].

3. Auxiliary results on localization functors

In this section, we give several results to compute localization functors λ^W with cosupports in arbitrary subsets W of $\operatorname{Spec} R$.

We first give the following lemma.

Lemma 3.1. *Let V be a specialization-closed subset of* Spec R. *Then we have the following equalities*;

$$^{\perp}\mathcal{C}^V = \mathcal{L}_{V^c} = \mathcal{L}_V^{\perp} = \mathcal{C}^{V^c}.$$

Proof. This follows from [Nakamura and Yoshino 2018, Lemma 4.3] and (1.2). □

Let W be a subset of $\operatorname{Spec} R$. We denote by $\overline{W}^s$ the specialization closure of W, which is the smallest specialization-closed subset of $\operatorname{Spec} R$ containing W. Moreover, for a subset W_0 of W, we say that W_0 is specialization-closed in W if $V(\mathfrak{p}) \cap W \subseteq W_0$ for any $\mathfrak{p} \in W_0$ (see [Nakamura and Yoshino 2018, Definition 3.10]). This is equivalent to saying that $\overline{W_0}^s \cap W = W_0$.

Corollary 3.2. *Let $W_0 \subseteq W \subseteq \operatorname{Spec} R$ be sets. Suppose that W_0 is specialization-closed in W. Setting $W_1 = W \setminus W_0$, we have $\mathcal{C}^{W_1} \subseteq {}^{\perp}\mathcal{C}^{W_0}$.*

Proof. Note that $W_1 \subseteq (\overline{W_0}^s)^c$. Further, we have ${}^{\perp}\mathcal{C}^{\overline{W_0}^s} = \mathcal{C}^{(\overline{W_0}^s)^c}$ by Lemma 3.1. Hence it holds that $\mathcal{C}^{W_1} \subseteq \mathcal{C}^{(\overline{W_0}^s)^c} = {}^{\perp}\mathcal{C}^{\overline{W_0}^s} \subseteq {}^{\perp}\mathcal{C}^{W_0}$. □

Remark 3.3. For an ideal $\mathfrak{a}$ of R, $\lambda^{V(\mathfrak{a})}$ is a right adjoint to $\gamma_{V(\mathfrak{a})}$ by (2.9) and (2.10). More generally, it is known that for any specialization-closed subset V, $\lambda^V : \mathcal{D} \to \mathcal{D}$ is a right adjoint to $\gamma_V : \mathcal{D} \to \mathcal{D}$. We now prove this fact, which will be used in the next proposition. Let $X, Y \in \mathcal{D}$, and consider the following triangles:

$$\gamma_V X \longrightarrow X \longrightarrow \lambda^{V^c} X \longrightarrow \gamma_V X[1],$$
$$\gamma_{V^c} Y \longrightarrow Y \longrightarrow \lambda^{V} Y \longrightarrow \gamma_{V^c} Y[1].$$

Since $\lambda^{V^c} X \in \mathcal{C}^{V^c} = {}^{\perp}\mathcal{C}^{V}$ by Lemma 3.1, applying $\operatorname{Hom}_{\mathcal{D}}(-, \lambda^V Y)$ to the first triangle, we have $\operatorname{Hom}_{\mathcal{D}}(\gamma_V X, \lambda^V Y) \cong \operatorname{Hom}_{\mathcal{D}}(X, \lambda^V Y)$. Moreover, Lemma 3.1 implies that $\gamma_{V^c} Y \in \mathcal{L}_{V^c} = \mathcal{L}_V^{\perp}$. Hence, applying $\operatorname{Hom}_{\mathcal{D}}(\gamma_V X, -)$ to the second triangle, we have $\operatorname{Hom}_{\mathcal{D}}(\gamma_V X, Y) \cong \operatorname{Hom}_{\mathcal{D}}(\gamma_V X, \lambda^V Y)$. Thus there is a natural isomorphism $\operatorname{Hom}_{\mathcal{D}}(\gamma_V X, Y) \cong \operatorname{Hom}_{\mathcal{D}}(X, \lambda^V Y)$, so that (γ_V, λ^V) is an adjoint pair. See also [Nakamura and Yoshino 2018, Remark 5.2].

Proposition 3.4. *Let V and U be arbitrary subsets of $\operatorname{Spec} R$. Suppose that one of the following conditions holds*:

(1) *V is specialization-closed.*

(2) *U is generalization-closed.*

Then we have an isomorphism

$$\lambda^V \lambda^U \cong \lambda^{V \cap U}.$$

Proof. Let $X \in \mathcal{D}$ and $Y \in \mathcal{C}^{V \cap U} = \mathcal{C}^V \cap \mathcal{C}^U$. Then there are natural isomorphisms

$$\operatorname{Hom}_{\mathcal{D}}(\lambda^V \lambda^U X, Y) \cong \operatorname{Hom}_{\mathcal{D}}(\lambda^U X, Y) \cong \operatorname{Hom}_{\mathcal{D}}(X, Y).$$

Recall that $\lambda^{V \cap U}$ is a left adjoint to the inclusion functor $\mathcal{C}^{V \cap U} \hookrightarrow \mathcal{D}$. Hence, by the uniqueness of adjoint functors, we only have to verify that $\lambda^V \lambda^U X \in \mathcal{C}^{V \cap U}$. Since $\lambda^V \lambda^U X \in \mathcal{C}^V$, it remains to show that $\lambda^V \lambda^U X \in \mathcal{C}^U$.

Case 1: Let $\mathfrak{p} \in U^c$. Since $\operatorname{supp} \gamma_V \kappa(\mathfrak{p}) \subseteq \{\mathfrak{p}\}$, it follows from (1.2) that $\gamma_V \kappa(\mathfrak{p}) \in \mathcal{L}_{U^c} = {}^{\perp}\mathcal{C}^U$. Thus, by the adjointness of (γ_V, λ^V), we have

$$\operatorname{RHom}_R(\kappa(\mathfrak{p}), \lambda^V \lambda^U X) \cong \operatorname{RHom}_R(\gamma_V \kappa(\mathfrak{p}), \lambda^U X) = 0.$$

This implies that $\operatorname{cosupp} \lambda^V \lambda^U X \subseteq U$, i.e., $\lambda^V \lambda^U X \in \mathcal{C}^U$.

Case 2: Since U^c is specialization-closed, Case 1 yields an isomorphism $\lambda^{U^c}\lambda^V \cong \lambda^{U^c\cap V}$. Furthermore, setting $W = (U^c\cap V)\cup U$, we see that $U^c\cap V$ is specialization-closed in W, and $W\setminus(U^c\cap V) = U$. Hence we have $\lambda^{U^c}(\lambda^V\lambda^U X) \cong \lambda^{U^c\cap V}\lambda^U X = 0$, by Corollary 3.2. It then follows from Lemma 3.1 that $\lambda^V\lambda^U X \in {}^{\perp}\mathcal{C}^{U^c} = \mathcal{C}^U$. $\square$

Remark 3.5. For arbitrary subsets W_0 and W_1 of $\operatorname{Spec} R$, Remark 2.7(ii) and Proposition 3.4 yield the isomorphisms

$$\lambda^{W_0}\lambda^{W_1} \cong \lambda^{W_0}\lambda^{\overline{W_0}^s}\lambda^{W_1} \cong \lambda^{W_0}\lambda^{\overline{W_0}^s\cap W_1},$$

$$\lambda^{W_0}\lambda^{W_1} \cong \lambda^{W_0}\lambda^{\overline{W_1}^g}\lambda^{W_1} \cong \lambda^{W_0\cap\overline{W_1}^g}\lambda^{W_1}.$$

The next result is a corollary of (2.8), (2.9) and Proposition 3.4.

Corollary 3.6. *Let S be a multiplicatively closed subset of R and $\mathfrak{a}$ be an ideal of R. We set $W = V(\mathfrak{a})\cap U_S$. Then we have*

$$\lambda^W \cong \mathrm{L}\Lambda^{V(\mathfrak{a})}(-\otimes_R S^{-1}R).$$

Since $V(\mathfrak{p})\cap U(\mathfrak{p}) = \{\mathfrak{p}\}$ for $\mathfrak{p}\in\operatorname{Spec} R$, as a special case of this corollary, we have the following result.

Corollary 3.7. *Let $\mathfrak{p}$ be a prime ideal of R. Then we have*

$$\lambda^{\{\mathfrak{p}\}} \cong \mathrm{L}\Lambda^{V(\mathfrak{p})}(-\otimes_R R_{\mathfrak{p}}).$$

The next lemma follows from this corollary and Lemma 2.4.

Lemma 3.8. *Let $\mathfrak{p}$ be a prime ideal of R and F be a flat R-module. Then $\lambda^{\{\mathfrak{p}\}}F$ is isomorphic to $\left(\bigoplus_B R_{\mathfrak{p}}\right)^{\wedge}_{\mathfrak{p}}$, where $\bigoplus_B R_{\mathfrak{p}}$ is the direct sum of B-copies of $R_{\mathfrak{p}}$ and $B = \dim_{\kappa(\mathfrak{p})} F\otimes_R \kappa(\mathfrak{p})$.*

Remark 3.9. If W_1 and W_2 are both specialization-closed or both generalization-closed, then Proposition 3.4 implies that $\lambda^{W_1}\lambda^{W_2} \cong \lambda^{W_2}\lambda^{W_1}$. However, in general, λ^{W_1} and λ^{W_2} need not commute. For example, let $\mathfrak{p},\mathfrak{q}\in\operatorname{Spec} R$ with $\mathfrak{p}\subsetneq\mathfrak{q}$. Then $(\lambda^{\{\mathfrak{p}\}}R)\otimes_R\kappa(\mathfrak{q}) = \widehat{R}_{\mathfrak{p}}\otimes_R\kappa(\mathfrak{q}) = 0$ and $(\lambda^{\{\mathfrak{q}\}}R)\otimes_R\kappa(\mathfrak{p}) = \widehat{R}_{\mathfrak{q}}\otimes_R\kappa(\mathfrak{p}) \neq 0$. Then we see from Lemma 3.8 that $\lambda^{\{\mathfrak{q}\}}\lambda^{\{\mathfrak{p}\}}R = 0$ and $\lambda^{\{\mathfrak{p}\}}\lambda^{\{\mathfrak{q}\}}R \neq 0$.

Compare this remark with [Benson et al. 2008, Example 3.5]. See also [Nakamura and Yoshino 2018, Remark 3.7(ii)].

Let $\mathfrak{p}$ be a prime ideal which is not maximal. Then $\lambda^{\{\mathfrak{p}\}}$ is distinct from $\Lambda^{\mathfrak{p}} = \mathrm{L}\Lambda^{V(\mathfrak{p})}\operatorname{RHom}_R(R_{\mathfrak{p}}, -)$, which is introduced in [Benson et al. 2012]. To see this, let $\mathfrak{q}$ be a prime ideal with $\mathfrak{p}\subsetneq\mathfrak{q}$. Then it holds that $\operatorname{cosupp}\widehat{R}_{\mathfrak{q}} = \{\mathfrak{q}\}\subseteq U(\mathfrak{p})^c$. Hence $\widehat{R}_{\mathfrak{q}}$ belongs to $\mathcal{C}^{U(\mathfrak{p})^c}$. Then we have $\operatorname{RHom}_R(R_{\mathfrak{p}},\widehat{R}_{\mathfrak{q}}) = 0$ since $R_{\mathfrak{p}}\in\mathcal{L}_{U(\mathfrak{p})} = {}^{\perp}\mathcal{C}^{U(\mathfrak{p})^c}$ by (1.2). This implies that $\Lambda^{\mathfrak{p}}\widehat{R}_{\mathfrak{q}} = \mathrm{L}\Lambda^{V(\mathfrak{p})}\operatorname{RHom}_R(R_{\mathfrak{p}},\widehat{R}_{\mathfrak{q}}) = 0$, while $\lambda^{\{\mathfrak{p}\}}\widehat{R}_{\mathfrak{q}} \cong \lambda^{\{\mathfrak{p}\}}\lambda^{\{\mathfrak{q}\}}R \neq 0$ by Remark 3.9.

Let $X \in \mathcal{D}$, and write $\varGamma_{\mathfrak{p}} = \mathrm{R}\Gamma_{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}})$ (see [Benson et al. 2008]). Recall that $\mathfrak{p} \in \operatorname{supp} X$ (resp. $\mathfrak{p} \in \operatorname{cosupp} X$) if and only if $\varGamma_{\mathfrak{p}} X \neq 0$ (resp. $\varLambda^{\mathfrak{p}} X \neq 0$); see [Foxby and Iyengar 2003, Theorems 2.1 and 4.1] and [Benson et al. 2012, §4]. In contrast, $\mathfrak{p} \in \operatorname{cosupp} X$ (resp. $\mathfrak{p} \in \operatorname{supp} X$) if and only if $\gamma_{\{\mathfrak{p}\}} X \neq 0$ (resp. $\lambda^{\{\mathfrak{p}\}} X \neq 0$), by Lemma 2.6. Here, $\gamma_{\{\mathfrak{p}\}} \cong \mathrm{R}\Gamma_{V(\mathfrak{p})} \operatorname{RHom}_R(R_{\mathfrak{p}}, -)$ by [Nakamura and Yoshino 2018, Corollary 3.3]. See also [Sather-Wagstaff and Wicklein 2017, Propositions 3.6 and 4.4].

Let W be a subset of $\operatorname{Spec} R$. We denote by $\dim W$ the supremum of lengths of chains of distinct prime ideals in W (see [Nakamura and Yoshino 2018, Definition 3.6]).

Theorem 3.10. *Let W be a subset of* $\operatorname{Spec} R$. *We assume that* $\dim W = 0$. *Then there are isomorphisms*

$$\lambda^W \cong \prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} \cong \prod_{\mathfrak{p} \in W} \mathrm{L}\Lambda^{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}}).$$

Proof. Let $X \in \mathcal{D}$, and consider the natural morphisms $\eta^{\{\mathfrak{p}\}} X : X \to \lambda^{\{\mathfrak{p}\}} X$ for $\mathfrak{p} \in W$. Take the product of the morphisms, and we obtain a morphism $f : X \to \prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} X$. Embed f into a triangle

$$C \longrightarrow X \xrightarrow{f} \prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} X \longrightarrow C[1].$$

Note that $\prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} X \in \mathcal{C}^W$. We have to prove that $C \in {}^{\perp}\mathcal{C}^W$. For this purpose, take any prime ideal $\mathfrak{q} \in W$. Then $\{\mathfrak{q}\}$ is specialization-closed in W, because $\dim W = 0$. Hence we have

$$\prod_{\mathfrak{p} \in W \setminus \{\mathfrak{q}\}} \lambda^{\{\mathfrak{p}\}} X \in \mathcal{C}^{W \setminus \{\mathfrak{q}\}} \subseteq {}^{\perp}\mathcal{C}^{\{\mathfrak{q}\}},$$

by Corollary 3.2. Thus an isomorphism $\lambda^{\{\mathfrak{q}\}}\big(\prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} X\big) \cong \lambda^{\{\mathfrak{q}\}} X$ holds. Then it is seen from the triangle above that $\lambda^{\{\mathfrak{q}\}} C = 0$ for all $\mathfrak{q} \in W$, so that $C \in {}^{\perp}\mathcal{C}^W$; see Remark 2.7(i). Therefore Lemma 2.6 yields $\lambda^W X \cong \prod_{\mathfrak{p} \in W} \lambda^{\{\mathfrak{p}\}} X$. The second isomorphism in the theorem follows from Corollary 3.7. $\square$

Example 3.11. Let W be a subset of $\operatorname{Spec} R$ such that W is an infinite set with $\dim W = 0$. Let $X^{\{\mathfrak{p}\}}$ be a complex with $\operatorname{cosupp} X^{\{\mathfrak{p}\}} = \{\mathfrak{p}\}$ for each $\mathfrak{p} \in W$. We take $\mathfrak{p} \in W$. Since $\dim W = 0$, it holds that $X^{\{\mathfrak{q}\}} \in \mathcal{C}^{V(\mathfrak{p})^c}$ for any $\mathfrak{q} \in W \setminus \{\mathfrak{p}\}$. Furthermore, Lemma 3.1 implies that $\mathcal{C}^{V(\mathfrak{p})^c}$ is equal to ${}^{\perp}\mathcal{C}^{V(\mathfrak{p})}$, which is closed under arbitrary direct sums. Thus it holds that

$$\bigoplus_{\mathfrak{q} \in W \setminus \{\mathfrak{p}\}} X^{\{\mathfrak{q}\}} \in \mathcal{C}^{V(\mathfrak{p})^c} = {}^{\perp}\mathcal{C}^{V(\mathfrak{p})} \subseteq {}^{\perp}\mathcal{C}^{\{\mathfrak{p}\}}.$$

Therefore, setting $Y = \bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}$, we have $\lambda^{\{\mathfrak{p}\}} Y \cong X^{\{\mathfrak{p}\}}$. It then follows from Theorem 3.10 that

$$\lambda^{W} Y \cong \prod_{\mathfrak{p}\in W} \lambda^{\{\mathfrak{p}\}} Y \cong \prod_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}.$$

Under this identification, the natural morphism $Y \to \lambda^{W} Y$ coincides with the canonical morphism $\bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}} \to \prod_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}$.

Remark 3.12. Let W, $X^{\{\mathfrak{p}\}}$ be as in Example 3.11, and suppose that each $X^{\{\mathfrak{p}\}}$ is an R-module. Then $\bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}$ is not in $\mathcal{C}^{W}$, because the natural morphism $\bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}} \to \lambda^{W}\big(\bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}\big)$ is not an isomorphism. Hence the cosupport of $\bigoplus_{\mathfrak{p}\in W} X^{\{\mathfrak{p}\}}$ properly contains W. In particular, setting $X^{\{\mathfrak{p}\}} = \kappa(\mathfrak{p})$, we have $W \subsetneq \operatorname{cosupp} \bigoplus_{\mathfrak{p}\in W} \kappa(\mathfrak{p})$. Similarly, we can prove that $W \subsetneq \operatorname{supp} \prod_{\mathfrak{p}\in W} \kappa(\mathfrak{p})$. Nakamura noticed these facts through discussion with Srikanth Iyengar.

It is possible to give another type of example, by which we also see that a colocalizing subcategory of $\mathcal{D}$ is not necessarily closed under arbitrary direct sums. Suppose that $(R, \mathfrak{m})$ is a complete local ring with $\dim R \geq 1$. Then we have $R \cong \widehat{R} \in \mathcal{C}^{V(\mathfrak{m})}$. However the free module $\bigoplus_{\mathbb{N}} R$ is never $\mathfrak{m}$-adically complete, so that $\bigoplus_{\mathbb{N}} R$ is not isomorphic to $\lambda^{V(\mathfrak{m})}\big(\bigoplus_{\mathbb{N}} R\big)$. Hence $\bigoplus_{\mathbb{N}} R$ is not in $\mathcal{C}^{V(\mathfrak{m})}$.

For a subset W of $\operatorname{Spec} R$, $\overline{W}^{g}$ denotes the generalization closure of W, which is the smallest generalization-closed subset of $\operatorname{Spec} R$ containing W. In addition, for a subset $W_1 \subseteq W$, we say that W_1 is generalization-closed in W if $W \cap U(\mathfrak{p}) \subseteq W_1$ for any $\mathfrak{p} \in W_1$. This is equivalent to saying that $W \cap \overline{W_1}^{g} = W_1$.

We extend Proposition 3.4 to the following corollary, which will be used in Theorem 3.15.

Corollary 3.13. *Let W_0 and W_1 be arbitrary subsets of* $\operatorname{Spec} R$. *Suppose that one of the following conditions hold*:

(1) *W_0 is specialization-closed in $W_0 \cup W_1$.*

(2) *W_1 is generalization-closed in $W_0 \cup W_1$.*

Then we have an isomorphism

$$\lambda^{W_0}\lambda^{W_1} \cong \lambda^{W_0\cap W_1}.$$

Proof. Set $W = W_0 \cup W_1$. By the assumption, we have

$$\overline{W_0}^{s} \cap W = W_0 \ \text{ or } \ W \cap \overline{W_1}^{g} = W_1.$$

Therefore, it holds that

$$\overline{W_0}^{s} \cap W_1 = W_0 \cap W_1 \ \text{ or } \ W_0 \cap \overline{W_1}^{g} = W_0 \cap W_1.$$

Hence this proposition follows from Remark 3.5 and Remark 2.7(ii). $\square$

Remark 3.14. (i) Let W_0 and W be subsets of $\operatorname{Spec} R$ with $W_0 \subseteq W$. Under the isomorphism $\lambda^{W_0}\lambda^{W} \cong \lambda^{W_0}$ by Remark 2.7(ii), there is a morphism $\eta^{W_0}\lambda^{W} : \lambda^{W} \to \lambda^{W_0}$.

(ii) Let W_0 and W_1 be subsets of $\operatorname{Spec} R$. Let $X \in \mathcal{D}$. Since $\eta^{W_1} : \mathrm{id}_{\mathcal{D}} \to \lambda^{W_1}$ is a morphism of functors, there is a commutative diagram of the following form:

$$\begin{tikzcd}
X \arrow[r, "\eta^{W_0}X"] \arrow[d, "\eta^{W_1}X"] & \lambda^{W_0}X \arrow[d, "\eta^{W_1}\lambda^{W_0}X"] \\
\lambda^{W_1}X \arrow[r, "\lambda^{W_1}\eta^{W_0}X"] & \lambda^{W_1}\lambda^{W_0}X
\end{tikzcd}$$

Now we prove the following result, which is the main theorem of this section.

Theorem 3.15. *Let W, W_0 and W_1 be subsets of* $\operatorname{Spec} R$ *with $W = W_0 \cup W_1$. Suppose that one of the following conditions holds*:

(1) *W_0 is specialization-closed in W.*

(2) *W_1 is generalization-closed in W.*

Then, for any $X \in \mathcal{D}$, there is a triangle of the form

$$\lambda^{W}X \xrightarrow{f} \lambda^{W_1}X \oplus \lambda^{W_0}X \xrightarrow{g} \lambda^{W_1}\lambda^{W_0}X \longrightarrow \lambda^{W}X[1],$$

where f and g are morphisms represented by the matrices

$$f = \begin{pmatrix} \eta^{W_1}\lambda^{W}X \\ \eta^{W_0}\lambda^{W}X \end{pmatrix}, \qquad g = \begin{pmatrix} \lambda^{W_1}\eta^{W_0}X & (-1)\cdot\eta^{W_1}\lambda^{W_0}X \end{pmatrix}.$$

Proof. We embed the morphism g into a triangle

$$C \xrightarrow{a} \lambda^{W_1}X \oplus \lambda^{W_0}X \xrightarrow{g} \lambda^{W_1}\lambda^{W_0}X \longrightarrow C[1].$$

Notice that $C \in \mathcal{C}^{W}$ since $\mathcal{C}^{W_0}, \mathcal{C}^{W_1} \subseteq \mathcal{C}^{W}$. By Remark 3.14, it is easily seen that $g \cdot f = 0$. Thus there is a morphism $b : \lambda^{W}X \to C$ making the following diagram commutative:

$$\begin{tikzcd}
\lambda^{W}X \arrow[r, equal] \arrow[d, "b"] & \lambda^{W}X \arrow[r] \arrow[d, "f"] & 0 \arrow[r] \arrow[d] & \lambda^{W}X[1] \arrow[d, "b[1]"] \\
C \arrow[r, "a"] & \lambda^{W_1}X \oplus \lambda^{W_0}X \arrow[r, "g"] & \lambda^{W_1}\lambda^{W_0}X \arrow[r] & C[1]
\end{tikzcd} \tag{3.16}$$

We only have to show that b is an isomorphism. To do this, embedding the morphism b into a triangle

$$Z \longrightarrow \lambda^{W}X \xrightarrow{b} C \longrightarrow Z[1], \tag{3.17}$$

we prove that $Z = 0$. Since $\lambda^{W}X, C \in \mathcal{C}^{W}$, Z belongs to $\mathcal{C}^{W}$. Hence it suffices to show that $Z \in {}^{\perp}\mathcal{C}^{W}$.

First, we prove that $\lambda^{W_1} b$ is an isomorphism. We employ a similar argument to [Benson et al. 2008, Theorem 7.5]. Consider the sequence

$$\lambda^W X \xrightarrow{f} \lambda^{W_1} X \oplus \lambda^{W_0} X \xrightarrow{g} \lambda^{W_1}\lambda^{W_0} X, \tag{3.18}$$

and apply λ^{W_1} to it. Then we obtain a sequence which can be completed to a split triangle. The triangle appears in the first row of the diagram below. Moreover, λ^{W_1} sends the second row of the diagram (3.16) to a split triangle, which appears in the second row of the diagram below:

$$\begin{tikzcd}
\lambda^{W_1} X \arrow[r, "\lambda^{W_1} f"] \arrow[d, "\lambda^{W_1} b"] & \lambda^{W_1} X \oplus \lambda^{W_1}\lambda^{W_0} X \arrow[r, "\lambda^{W_1} g"] \arrow[d, equal] & \lambda^{W_1}\lambda^{W_0} X \arrow[r, "0"] \arrow[d, equal] & \lambda^{W_1} X[1] \arrow[d, "\lambda^{W_1} b[1]"] \\
\lambda^{W_1} C \arrow[r, "\lambda^{W_1} a"] & \lambda^{W_1} X \oplus \lambda^{W_1}\lambda^{W_0} X \arrow[r, "\lambda^{W_1} g"] & \lambda^{W_1}\lambda^{W_0} X \arrow[r, "0"] & \lambda^{W_1} C[1]
\end{tikzcd}$$

Since this diagram is commutative, we conclude that $\lambda^{W_1} b$ is an isomorphism.

Next, we prove that $\lambda^{W_0} b$ is an isomorphism. Thanks to Corollary 3.13, we are able to follow the same process as above. In fact, the corollary implies that $\lambda^{W_0}\lambda^{W_1} \cong \lambda^{W_0 \cap W_1}$. Thus, applying λ^{W_0} to the sequence (3.18), we obtain a sequence which can be completed into a split triangle. Furthermore, λ^{W_0} sends the second row of the diagram (3.16) to a split triangle. Consequently we see that there is a morphism of triangles:

$$\begin{tikzcd}
\lambda^{W_0} X \arrow[r, "\lambda^{W_0} f"] \arrow[d, "\lambda^{W_0} b"] & \lambda^{W_0 \cap W_1} X \oplus \lambda^{W_0} X \arrow[r, "\lambda^{W_0} g"] \arrow[d, equal] & \lambda^{W_0 \cap W_1} X \arrow[r, "0"] \arrow[d, equal] & \lambda^{W_0} X[1] \arrow[d, "\lambda^{W_0} b[1]"] \\
\lambda^{W_0} C \arrow[r, "\lambda^{W_0} a"] & \lambda^{W_0 \cap W_1} X \oplus \lambda^{W_0} X \arrow[r, "\lambda^{W_0} g"] & \lambda^{W_0 \cap W_1} X \arrow[r, "0"] & \lambda^{W_0} C[1]
\end{tikzcd}$$

Therefore $\lambda^{W_0} b$ is an isomorphism.

Since we have shown that $\lambda^{W_0} b$ and $\lambda^{W_1} b$ are isomorphisms, it follows from the triangle (3.17) that $\lambda^{W_0} Z = \lambda^{W_1} Z = 0$. Thus we have $Z \in {}^{\perp}\mathcal{C}^W$ by Remark 2.7(i). $\square$

Remark 3.19. Let f, g and a be as above. Let $h : X \to \lambda^{W_1} X \oplus \lambda^{W_0} X$ be a morphism induced by $\eta^{W_1} X$ and $\eta^{W_0} X$. Then $g \cdot h = 0$ by Remark 3.14(ii). Hence there is a morphism $b' : X \to C$ such that the following diagram is commutative:

$$\begin{tikzcd}
X \arrow[r, equal] \arrow[d, "b'"] & X \arrow[r] \arrow[d, "h"] & 0 \arrow[r] \arrow[d] & X[1] \arrow[d, "b'[1]"] \\
C \arrow[r, "a"] & \lambda^{W_1} X \oplus \lambda^{W_0} X \arrow[r, "g"] & \lambda^{W_1}\lambda^{W_0} X \arrow[r] & C[1]
\end{tikzcd}$$

We can regard any morphism b' making this diagram commutative as the natural morphism $\eta^W X$. In fact, since $\lambda^W h = f$, applying λ^W to this diagram, and setting $\lambda^W b' = b$, we obtain the diagram (3.16). Note that $b \cdot \eta^W X = b'$. Moreover, the above proof implies that $b : \lambda^W X \to C$ is an isomorphism. Thus we can identify b' with $\eta^W X$ under the isomorphism b.

We give some examples of Theorem 3.15.

Example 3.20. (1) Let x be an element of R. Recall that $\lambda^{V(x)} \cong \mathrm{L}\Lambda^{V(x)}$ by (2.9). We put $S = \{1, x, x^2, \dots\}$. Since $V(x)^c = U_S$, it holds that $\lambda^{V(x)^c} = \lambda^{U_S} \cong (-)\otimes_R R_x$ by (2.8). Set $W = \operatorname{Spec} R$, $W_0 = V(x)$ and $W_1 = V(x)^c$. Then the theorem yields the triangle

$$R \longrightarrow R_x \oplus R^{\wedge}_{(x)} \longrightarrow (R^{\wedge}_{(x)})_x \longrightarrow R[1].$$

(2) Suppose that $(R, \mathfrak{m})$ is a local ring with $\mathfrak{p} \in \operatorname{Spec} R$ and having $\dim R/\mathfrak{p} = 1$. Setting $W = V(\mathfrak{p})$, $W_0 = V(\mathfrak{m})$ and $W_1 = \{\mathfrak{p}\}$, we see from the theorem and Corollary 3.7 that there is a short exact sequence,

$$0 \longrightarrow R^{\wedge}_{\mathfrak{p}} \longrightarrow \widehat{R_{\mathfrak{p}}} \oplus \widehat{R} \longrightarrow \widehat{(\widehat{R})_{\mathfrak{p}}} \longrightarrow 0.$$

Actually, this gives a pure-injective resolution of $R^{\wedge}_{\mathfrak{p}}$; see Section 9. Moreover, if R is a 1-dimensional local domain with quotient field Q, then this short exact sequence is of the form

$$0 \longrightarrow R \longrightarrow Q \oplus \widehat{R} \longrightarrow \widehat{R} \otimes_R Q \longrightarrow 0.$$

By similar arguments to Proposition 3.4 and Corollary 3.13, one can prove the following proposition, which is a generalized form of [Nakamura and Yoshino 2018, Proposition 3.1].

Proposition 3.21. *Let W_0 and W_1 be arbitrary subsets of* $\operatorname{Spec} R$. *Suppose that one of the following conditions hold*:

(1) *W_0 is specialization-closed in $W_0 \cup W_1$.*

(2) *W_1 is generalization-closed in $W_0 \cup W_1$.*

Then we have an isomorphism

$$\gamma_{W_0}\gamma_{W_1} \cong \gamma_{W_0 \cap W_1}.$$

As with Theorem 3.15, it is possible to prove the following theorem, in which we implicitly use the fact that $\gamma_{W_0}\gamma_W \cong \gamma_{W_0}$ if $W_0 \subseteq W$ (see [Nakamura and Yoshino 2018, Remark 3.7(i)]).

Theorem 3.22. *Let W, W_0 and W_1 be subsets of* $\operatorname{Spec} R$ *with $W = W_0 \cup W_1$. Suppose that one of the following conditions holds*:

(1) *W_0 is specialization-closed in W.*

(2) *W_1 is generalization-closed in W.*

Then, for any $X \in \mathcal{D}$, there is a triangle of the form

$$\gamma_{W_1}\gamma_{W_0}X \xrightarrow{f} \gamma_{W_1}X \oplus \gamma_{W_0}X \xrightarrow{g} \gamma_W X \longrightarrow \gamma_{W_1}\gamma_{W_0}X[1],$$

where f and g are morphisms represented by the following matrices;

$$f = \begin{pmatrix} \gamma_{W_1}\varepsilon_{W_0}X \\ (-1)\cdot\varepsilon_{W_1}\gamma_{W_0}X \end{pmatrix}, \quad g = \begin{pmatrix} \varepsilon_{W_1}\gamma_W X & \varepsilon_{W_0}\gamma_W X \end{pmatrix}.$$

Remark 3.23. As long as we work on the derived category $\mathcal{D}$, Theorem 3.15 and Theorem 3.22 generalize Mayer–Vietoris triangles in the sense of [Benson et al. 2008, Theorem 7.5], in which γ_V and λ_V are written as Γ_V and L_V, respectively, for a specialization-closed subset V of Spec R.

4. Projective dimension of flat modules

As an application of results in Section 3, we give a simpler proof of a classical theorem due to Gruson and Raynaud.

Theorem 4.1 [Raynaud and Gruson 1971, Part. II, Corollary 3.2.7]. *Let F be a flat R-module. Then the projective dimension of F is at most* $\dim R$.

We start by showing the following lemma.

Lemma 4.2. *Let F be a flat R-module and $\mathfrak{p}$ be a prime ideal of R. Suppose that $X \in \mathcal{C}^{\{\mathfrak{p}\}}$. Then there is an isomorphism*

$$\operatorname{RHom}_R(F, X) \cong \prod_B X,$$

where $B = \dim_{\kappa(\mathfrak{p})} F \otimes_R \kappa(\mathfrak{p})$.

Proof. Since $\lambda^{\{\mathfrak{p}\}} : \mathcal{D} \to \mathcal{C}^{\{\mathfrak{p}\}}$ is a left adjoint to the inclusion functor $\mathcal{C}^{\{\mathfrak{p}\}} \hookrightarrow \mathcal{D}$, we have $\operatorname{RHom}_R(F, X) \cong \operatorname{RHom}_R(\lambda^{\{\mathfrak{p}\}} F, X)$. Moreover, it follows from Lemma 3.8 that $\lambda^{\{\mathfrak{p}\}} F \cong \big(\bigoplus_B R_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}} \cong \lambda^{\{\mathfrak{p}\}}\big(\bigoplus_B R\big)$, where $B = \dim_{\kappa(\mathfrak{p})} F \otimes_R \kappa(\mathfrak{p})$. Therefore we obtain isomorphisms

$$\operatorname{RHom}_R(F, X) \cong \operatorname{RHom}_R\Big(\lambda^{\{\mathfrak{p}\}}\Big(\bigoplus_B R\Big), X\Big) \cong \operatorname{RHom}_R\Big(\bigoplus_B R, X\Big) \cong \prod_B X. \quad \square$$

Let $a, b \in \mathbb{Z} \cup \{\pm\infty\}$ with $a \le b$. We write $\mathcal{D}^{[a,b]}$ for the full subcategory of $\mathcal{D}$ consisting of all complexes X of R-modules such that $H^i(X) = 0$ for $i \notin [a, b]$ (see [Kashiwara and Schapira 2006, Notation 13.1.11]). For a subset W of Spec R, $\max W$ denotes the set of prime ideals $\mathfrak{p} \in W$ which are maximal with respect to inclusion in W.

Proposition 4.3. *Let F be a flat R-module and $X \in \mathcal{D}^{[-\infty,0]}$. Suppose that W is a subset of* Spec R *such that $n = \dim W$ is finite. Then we have* $\operatorname{Ext}^i_R(F, \lambda^W X) = 0$ *for $i > n$.*

Proof. We use induction on n. First, we suppose that $n = 0$. It then holds that
$$\lambda^W X \cong \prod_{\mathfrak{p}\in W} \lambda^{\{\mathfrak{p}\}} X \cong \prod_{\mathfrak{p}\in W} \mathrm{L}\Lambda^{V(\mathfrak{p})} X_{\mathfrak{p}} \in \mathcal{D}^{[-\infty,0]},$$
by Theorem 3.10. Hence, noting that
$$\mathrm{RHom}_R(F, \lambda^W X) \cong \prod_{\mathfrak{p}\in W} \mathrm{RHom}_R(F, \lambda^{\{\mathfrak{p}\}} X),$$
we have $\mathrm{Ext}^i_R(F, \lambda^W X) = 0$ for $i > 0$, by Lemma 4.2.

Next, we suppose $n > 0$. Set $W_0 = \max W$ and $W_1 = W \setminus W_0$. By Theorem 3.15, there is a triangle
$$\lambda^W X \longrightarrow \lambda^{W_1} X \oplus \lambda^{W_0} X \longrightarrow \lambda^{W_1}\lambda^{W_0} X \longrightarrow \lambda^W X[1].$$
Note that $\dim W_0 = 0$ and $\dim W_1 = n - 1$. By the argument above, it holds that $\mathrm{Ext}^i_R(F, \lambda^{W_0} X) = 0$ for $i > 0$. Furthermore, since $X, \lambda^{W_0} X \in \mathcal{D}^{[-\infty,0]}$, we have $\mathrm{Ext}^i_R(F, \lambda^{W_1} X) = \mathrm{Ext}^i_R(F, \lambda^{W_1}\lambda^{W_0} X) = 0$ for $i > n-1$, by the inductive hypothesis. Hence it is seen from the triangle that $\mathrm{Ext}^i_R(F, \lambda^W X) = 0$ for $i > n$. □

Proof of Theorem 4.1. We may assume that $d = \dim R$ is finite. Let M be any R-module. We only have to show that $\mathrm{Ext}^i_R(F, M) = 0$ for $i > d$. Setting $W = \mathrm{Spec}\, R$, we have $\dim W = d$ and $M \cong \lambda^W M$. It then follows from Proposition 4.3 that $\mathrm{Ext}^i_R(F, M) \cong \mathrm{Ext}^i_R(F, \lambda^W M) = 0$ for $i > d$. □

5. Cotorsion flat modules and cosupport

In this section, we summarize some basic facts about cotorsion flat R-modules.

Recall that an R-module M is called cotorsion if $\mathrm{Ext}^1_R(F, M) = 0$ for any flat R-module F. This is equivalent to saying that $\mathrm{Ext}^i_R(F, M) = 0$ for any flat R-module F and any $i > 0$. Clearly, all injective R-modules are cotorsion.

A cotorsion flat R-module means an R-module which is cotorsion and flat. If F is a flat R-module and $\mathfrak{p} \in \mathrm{Spec}\, R$, then Corollary 3.7 implies that $\lambda^{\{\mathfrak{p}\}} F$ is isomorphic to $\widehat{F_{\mathfrak{p}}}$, which is a cotorsion flat R-module by Lemma 4.2 and Proposition 2.1. Moreover, recall that $\widehat{F_{\mathfrak{p}}}$ is isomorphic to the $\mathfrak{p}$-adic completion of a free $R_{\mathfrak{p}}$-module by Lemma 3.8.

We remark that arbitrary direct products of flat R-modules are flat, since R is Noetherian. Hence, if $T_{\mathfrak{p}}$ is the $\mathfrak{p}$-adic completion of a free $R_{\mathfrak{p}}$ module for each $\mathfrak{p} \in \mathrm{Spec}\, R$, then $\prod_{\mathfrak{p}\in\mathrm{Spec}\, R} T_{\mathfrak{p}}$ is a cotorsion flat R-module. Conversely, the following fact holds.

Proposition 5.1 [Enochs 1984]. *Let F be a cotorsion flat R-module. Then there is an isomorphism*
$$F \cong \prod_{\mathfrak{p}\in\mathrm{Spec}\, R} T_{\mathfrak{p}},$$
where $T_{\mathfrak{p}}$ is the $\mathfrak{p}$-adic completion of a free $R_{\mathfrak{p}}$ module.

Proof. See [Enochs 1984, Theorem; Enochs and Jenda 2000, Theorem 5.3.28]. □

Let S be a multiplicatively closed subset of R and $\mathfrak{a}$ be an ideal of R. For a cotorsion flat R-module F, we have $\mathrm{RHom}_R(S^{-1}R, F) \cong \mathrm{Hom}_R(S^{-1}R, F)$ and $\mathrm{L}\Lambda^{V(\mathfrak{a})} F \cong \Lambda^{V(\mathfrak{a})} F$. Moreover, by Proposition 5.1, we may regard F as an R-module of the form $\prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T_{\mathfrak{p}}$. Then it holds that

$$\mathrm{RHom}_R\Bigl(S^{-1}R, \prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T_{\mathfrak{p}}\Bigr) \cong \mathrm{Hom}_R\Bigl(S^{-1}R, \prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T_{\mathfrak{p}}\Bigr) \cong \prod_{\mathfrak{p}\in U_S} T_{\mathfrak{p}}. \tag{5.2}$$

This fact appears implicitly in [Xu 1996, §5.2]. Furthermore we have

$$\mathrm{L}\Lambda^{V(\mathfrak{a})} \prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T_{\mathfrak{p}} \cong \Lambda^{V(\mathfrak{a})} \prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T_{\mathfrak{p}} \cong \prod_{\mathfrak{p}\in V(\mathfrak{a})} T_{\mathfrak{p}}. \tag{5.3}$$

One can show (5.2) and (5.3) by Lemma 3.1 and (2.9). See also Thompson's recent lemma [2017b, Lemma 2.2].

Let F be a cotorsion flat R-module with $\operatorname{cosupp} F \subseteq W$ for a subset W of $\mathrm{Spec}\, R$. Then it follows from Proposition 5.1 that F is isomorphic to an R-module of the form $\prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}$. More precisely, using Lemma 2.4, (5.2) and (5.3), one can show the following corollary, which is essentially proved in [Enochs and Jenda 2000, Lemma 8.5.25].

Corollary 5.4. *Let F be a cotorsion flat R-module, and set $W = \operatorname{cosupp} F$. Then we have an isomorphism*

$$F \cong \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}},$$

where $T_{\mathfrak{p}}$ is of the form $\bigl(\bigoplus_{B_{\mathfrak{p}}} R_{\mathfrak{p}}\bigr)^{\wedge}_{\mathfrak{p}}$ with $B_{\mathfrak{p}} = \dim_{\kappa(\mathfrak{p})} \mathrm{Hom}_R(R_{\mathfrak{p}}, F) \otimes_R \kappa(\mathfrak{p})$.

6. Complexes of cotorsion flat modules and cosupport

In this section, we study the cosupport of a complex X consisting of cotorsion flat R-modules. As a consequence, we obtain an explicit way to calculate $\gamma_{V^c} X$ and $\lambda^V X$ for a specialization-closed subset V of $\mathrm{Spec}\, R$.

Notation 6.1. Let W be a subset of $\mathrm{Spec}\, R$. Let X be a complex of cotorsion flat R-modules such that $\operatorname{cosupp} X^i \subseteq W$ for all $i \in \mathbb{Z}$. Under Corollary 5.4, we use a presentation of the form

$$X = \Bigl(\cdots \to \prod_{\mathfrak{p}\in W} T^i_{\mathfrak{p}} \to \prod_{\mathfrak{p}\in W} T^{i+1}_{\mathfrak{p}} \to \cdots\Bigr),$$

where $X^i = \prod_{\mathfrak{p}\in \mathrm{Spec}\, R} T^i_{\mathfrak{p}}$ and $T^i_{\mathfrak{p}}$ is the $\mathfrak{p}$-adic completion of a free $R_{\mathfrak{p}}$-module.

Remark 6.2. Let $X = \big(\cdots \to \prod_{\mathfrak{p}\in \operatorname{Spec} R} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in \operatorname{Spec} R} T_{\mathfrak{p}}^{i+1} \to \cdots\big)$ be a complex of cotorsion flat R-modules. Let V be a specialization-closed subset of Spec R. By Lemma 3.1, we have $\operatorname{Hom}_R\big(\prod_{\mathfrak{p}\in V^c} T_{\mathfrak{p}}^i, \prod_{\mathfrak{p}\in V} T_{\mathfrak{p}}^{i+1}\big) = 0$ for all $i \in \mathbb{Z}$. Therefore $Y = \big(\cdots \to \prod_{\mathfrak{p}\in V^c} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in V^c} T_{\mathfrak{p}}^{i+1} \to \cdots\big)$ is a subcomplex of X, where the differentials in Y are the restrictions of ones in X.

We say that a complex X of R-modules is left (resp. right) bounded if $X^i = 0$ for $i \ll 0$ (resp. $i \gg 0$). When X is left and right bounded, X is called bounded.

Proposition 6.3. *Let W be a subset of* Spec R *and X be a complex of cotorsion flat R-modules such that* cosupp $X^i \subseteq W$ *for all $i \in \mathbb{Z}$. Suppose that one of the following conditions holds*:

(1) *X is left bounded.*

(2) *W is equal to $V(\mathfrak{a})$ for an ideal $\mathfrak{a}$ of R.*

(3) *W is generalization-closed.*

(4) dim W *is finite.*

Then it holds that cosupp $X \subseteq W$, *i.e.*, $X \in \mathcal{C}^W$.

To show this, we use the elementary lemma below. Therein, for a complex X and $n \in \mathbb{Z}$, we define the truncations $\tau_{\le n}X$ and $\tau_{>n}X$ as follows (see [Hartshorne 1966, Chapter I, §7]):

$$\tau_{\le n}X = (\cdots \to X^{n-1} \to X^n \to 0 \to \cdots),$$
$$\tau_{>n}X = (\cdots \to 0 \to X^{n+1} \to X^{n+2} \to \cdots).$$

Lemma 6.4. *Let W be a subset of* Spec R. *We assume that $\tau_{\le n}X \in \mathcal{C}^W$ (resp. $\tau_{>n}X \in \mathcal{L}_W$) for all $n \ge 0$ (resp. $n < 0$). Then we have $X \in \mathcal{C}^W$ (resp. $X \in \mathcal{L}_W$).*

Recall that $\mathcal{C}^W$ (resp. $\mathcal{L}_W$) is closed under arbitrary direct products (resp. sums). Then one can show this lemma by using homotopy limits (resp. colimits), see [Bökstedt and Neeman 1993, Remarks 2.2 and 2.3]. □

Proof of Proposition 6.3. Case 1: We have $\tau_{\le n}X \in \mathcal{C}^W$ for all $n \ge 0$, since $\tau_{\le n}X$ are bounded. Thus Lemma 6.4 implies that $X \in \mathcal{C}^W$.

Case 2: By (2.9), Proposition 2.5 and (5.3), it holds that $\lambda^{V(\mathfrak{a})}X \cong \mathrm{L}\Lambda^{V(\mathfrak{a})}X \cong \Lambda^{V(\mathfrak{a})}X \cong X$. Hence X belongs to $\mathcal{C}^{V(\mathfrak{a})}$.

Case 3: It follows from Case 1 that $\tau_{>n}X \in \mathcal{C}^W$ for all $n < 0$. Moreover, we have $\mathcal{C}^W = \mathcal{L}_W$ by Lemma 3.1. Thus Lemma 6.4 implies that $X \in \mathcal{L}_W = \mathcal{C}^W$.

Case 4: Under Notation 6.1, we write $X^i = \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}^i$ for $i \in \mathbb{Z}$. Set $n = \dim W$, and use induction on n. First, suppose that $n = 0$. It is seen from Remark 6.2 that X is the direct product of complexes of the form $Y^{\{\mathfrak{p}\}} = (\cdots \to T_{\mathfrak{p}}^i \to T_{\mathfrak{p}}^{i+1} \to \cdots)$ for $\mathfrak{p} \in W$.

Furthermore, by Cases 2 and 3, we have $\operatorname{cosupp} Y^{\{\mathfrak{p}\}} \subseteq V(\mathfrak{p}) \cap U(\mathfrak{p}) = \{\mathfrak{p}\}$. Thus it holds that $X \cong \prod_{\mathfrak{p}\in W} Y^{\{\mathfrak{p}\}} \in \mathcal{C}^W$.

Next, suppose that $n > 0$. Set $W_0 = \max W$ and $W_1 = W \setminus W_0$. We write $Y = \left(\cdots \to \prod_{\mathfrak{p}\in W_1} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W_1} T_{\mathfrak{p}}^{i+1} \to \cdots\right)$, which is a subcomplex of X by Remark 6.2. Hence there is a short exact sequence of complexes,

$$0 \longrightarrow Y \longrightarrow X \longrightarrow X/Y \longrightarrow 0,$$

where $X/Y = \left(\cdots \to \prod_{\mathfrak{p}\in W_0} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W_0} T_{\mathfrak{p}}^{i+1} \to \cdots\right)$. Note that $\dim W_0 = 0$ and $\dim W_1 = n-1$. Then we have $\operatorname{cosupp} X/Y \subseteq W_0$, by the argument above. Moreover the inductive hypothesis implies that $\operatorname{cosupp} Y \subseteq W_1$. Hence it holds that $\operatorname{cosupp} X \subseteq W_0 \cup W_1 = W$. $\square$

Under some assumption, it is possible to extend condition (4) in Proposition 6.3 to the case where $\dim W$ is infinite; see Remark 7.15. See also [Thompson 2017a, Theorem 2.7].

Corollary 6.5. *Let X be a complex of cotorsion flat R-modules and W be a specialization-closed subset of* $\operatorname{Spec} R$. *Under Notation 6.1*, *we write*

$$X = \left(\cdots \to \prod_{\mathfrak{p}\in \operatorname{Spec} R} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in \operatorname{Spec} R} T_{\mathfrak{p}}^{i+1} \to \cdots\right).$$

Suppose that one of the conditions in Proposition 6.3 holds. Then it holds that

$$\begin{aligned} \gamma_{W^c} X &\cong \left(\cdots \to \prod_{\mathfrak{p}\in W^c} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W^c} T_{\mathfrak{p}}^{i+1} \to \cdots\right), \\ \lambda^W X &\cong \left(\cdots \to \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}^{i+1} \to \cdots\right). \end{aligned} \tag{6.6}$$

Proof. Since $Y = \left(\cdots \to \prod_{\mathfrak{p}\in W^c} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W^c} T_{\mathfrak{p}}^{i+1} \to \cdots\right)$ is a subcomplex of X by Remark 6.2, there is a triangle in $\mathcal{D}$:

$$Y \longrightarrow X \longrightarrow X/Y \longrightarrow Y[1],$$

where $X/Y = \left(\cdots \to \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p}\in W} T_{\mathfrak{p}}^{i+1} \to \cdots\right)$. By Proposition 6.3, we have $X/Y \in \mathcal{C}^W$. Moreover, since W^c is generalization-closed, it holds that $Y \in \mathcal{C}^{W^c} = {}^{\perp}\mathcal{C}^W$ by Proposition 6.3 and Lemma 3.1. Therefore we conclude that $\gamma_{W^c} X \cong Y$ and $\lambda^W X \cong X/Y$ by Lemma 2.6. $\square$

Let X be a complex of cotorsion flat R-modules and S be a multiplicatively closed subset of R. We assume that X is left bounded, or $\dim R$ is finite. It then follows from the corollary and (5.2) that

$$\gamma_{U_S} X \cong \Big(\cdots \to \prod_{\mathfrak{p} \in U_S} T_{\mathfrak{p}}^i \to \prod_{\mathfrak{p} \in U_S} T_{\mathfrak{p}}^{i+1} \to \cdots \Big) \cong \mathrm{Hom}_R(S^{-1}R, X).$$

We now recall that $\gamma_{U_S} \cong \mathrm{RHom}_R(S^{-1}R, -)$; see [Nakamura and Yoshino 2018, Proposition 3.1]. Hence it holds that $\mathrm{RHom}_R(S^{-1}R, X) \cong \mathrm{Hom}_R(S^{-1}R, X)$. This fact also follows from Lemma 9.1.

7. Localization functors via Čech complexes

In this section, we introduce a new notion of Čech complexes to calculate $\lambda^W X$, where W is a general subset W of $\mathrm{Spec}\, R$ and X is a complex of flat R-modules.

We first set the following notation.

Notation 7.1. Let W be a subset of $\mathrm{Spec}\, R$ with $\dim W = 0$. We define a functor $\bar{\lambda}^W : \mathrm{Mod}\, R \to \mathrm{Mod}\, R$ by

$$\bar{\lambda}^W = \prod_{\mathfrak{p} \in W} \Lambda^{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}}).$$

For a prime ideal $\mathfrak{p}$ in W, we write

$$\bar{\eta}^{\{\mathfrak{p}\}} : \mathrm{id}_{\mathrm{Mod}\, R} \to \bar{\lambda}^{\{\mathfrak{p}\}} = \Lambda^{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}})$$

for the composition of the natural morphisms $\mathrm{id}_{\mathrm{Mod}\, R} \to (-) \otimes_R R_{\mathfrak{p}}$ and $(-) \otimes_R R_{\mathfrak{p}} \to \Lambda^{V(\mathfrak{p})}(- \otimes_R R_{\mathfrak{p}})$. Moreover, $\bar{\eta}^W : \mathrm{id}_{\mathrm{Mod}\, R} \to \bar{\lambda}^W = \prod_{\mathfrak{p} \in W} \bar{\lambda}^{\{\mathfrak{p}\}}$ denotes the product of the morphisms $\bar{\eta}^{\{\mathfrak{p}\}}$ for $\mathfrak{p} \in W$.

Notation 7.2. Let $\{W_i\}_{0 \le i \le n}$ be a family of subsets of $\mathrm{Spec}\, R$, and suppose that $\dim W_i = 0$ for $0 \le i \le n$. For a sequence $(i_m, \dots, i_1, i_0)$ of integers with $0 \le i_0 < i_1 < \cdots < i_m \le n$, we write

$$\bar{\lambda}^{(i_m, \dots, i_1, i_0)} = \bar{\lambda}^{W_{i_m}} \cdots \bar{\lambda}^{W_{i_1}} \bar{\lambda}^{W_{i_0}}.$$

If the sequence is empty, then we use the general convention that $\lambda^{(\,)} = \mathrm{id}_{\mathrm{Mod}\, R}$. For an integer s with $0 \le s \le m$, $\bar{\eta}^{W_{i_s}} : \mathrm{id}_{\mathrm{Mod}\, R} \to \bar{\lambda}^{(i_s)}$ induces a morphism

$$\bar{\lambda}^{(i_m, \dots, i_{s+1})} \bar{\eta}^{W_{i_s}} \bar{\lambda}^{(i_{s-1}, \dots, i_0)} : \bar{\lambda}^{(i_m, \dots, \hat{i}_s, \dots, i_0)} \to \bar{\lambda}^{(i_m, \dots, i_0)},$$

where we mean by $\hat{i}_s$ that i_s is omitted. We set

$$\partial^{m-1} : \prod_{0 \le i_0 < \cdots < i_{m-1} \le n} \bar{\lambda}^{(i_{m-1}, \dots, i_0)} \to \prod_{0 \le i_0 < \cdots < i_m \le n} \bar{\lambda}^{(i_m, \dots, i_0)}$$

to be the product of the morphisms $\bar{\lambda}^{(i_m, \dots, \hat{i}_s, \dots, i_0)} \to \bar{\lambda}^{(i_m, \dots, i_0)}$ multiplied by $(-1)^s$.

Remark 7.3. Let $W_0, W_1 \subseteq \operatorname{Spec} R$ be subsets such that $\dim W_0 = \dim W_1 = 0$. As with Remark 3.14(ii), the following diagram is commutative:

$$\begin{array}{ccc} \mathrm{id}_{\mathrm{Mod}\,R} & \xrightarrow{\bar{\eta}^{W_0}} & \bar{\lambda}^{W_0} \\ \Big\downarrow{\scriptstyle \bar{\eta}^{W_1}} & & \Big\downarrow{\scriptstyle \bar{\eta}^{W_1}\bar{\lambda}^{W_0}} \\ \bar{\lambda}^{W_1} & \xrightarrow{\bar{\lambda}^{W_1}\bar{\eta}^{W_0}} & \bar{\lambda}^{W_1}\bar{\lambda}^{W_0} \end{array}$$

Definition 7.4. Let $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ be a family of subsets of $\operatorname{Spec} R$, and suppose that $\dim W_i = 0$ for $0 \le i \le n$. By Remark 7.3, it is possible to construct a Čech complex of functors of the form

$$\prod_{0 \le i_0 \le n} \bar{\lambda}^{(i_0)} \xrightarrow{\partial^0} \prod_{0 \le i_0 < i_1 \le n} \bar{\lambda}^{(i_1, i_0)} \to \cdots \to \prod_{0 \le i_0 < \cdots < i_{n-1} \le n} \bar{\lambda}^{(i_{n-1}, \ldots, i_0)} \xrightarrow{\partial^{n-1}} \bar{\lambda}^{(n, \ldots, 0)},$$

which we denote by $L^{\mathbb{W}}$ and call it *the Čech complex with respect to* $\mathbb{W}$.

For an R-module M, $L^{\mathbb{W}}M$ denotes the complex of R-modules obtained by $L^{\mathbb{W}}$ in a natural way, where it is concentrated in degrees from 0 to n. We call $L^{\mathbb{W}}M$ *the Čech complex of M with respect to* $\mathbb{W}$. Note that there is a chain map $\ell^{\mathbb{W}}M : M \to L^{\mathbb{W}}M$ induced by the map $M \to \prod_{0 \le i_0 \le n} \bar{\lambda}^{(i_0)}M$ in degree 0, which is the product of $\bar{\eta}^{W_{i_0}}M : M \to \bar{\lambda}^{(i_0)}M$ for $0 \le i_0 \le n$.

More generally, we regard every term of $L^{\mathbb{W}}$ as a functor $C(\mathrm{Mod}\,R) \to C(\mathrm{Mod}\,R)$, where $C(\mathrm{Mod}\,R)$ denotes the category of complexes of R-modules. Then $L^{\mathbb{W}}$ naturally sends a complex X to a double complex, which we denote by $L^{\mathbb{W}}X$. Furthermore, we write $\operatorname{tot} L^{\mathbb{W}}X$ for the total complex of $L^{\mathbb{W}}X$. The family of chain maps $\ell^{\mathbb{W}}X^j : X^j \to L^{\mathbb{W}}X^j$ for $j \in \mathbb{Z}$ induces a morphism $X \to L^{\mathbb{W}}X$ as double complexes, from which we obtain a chain map $\ell^{\mathbb{W}}X : X \to \operatorname{tot} L^{\mathbb{W}}X$.

Remark 7.5. (i) We regard $\operatorname{tot} L^{\mathbb{W}}$ as a functor $C(\mathrm{Mod}\,R) \to C(\mathrm{Mod}\,R)$. Then $\ell^{\mathbb{W}}$ is a morphism $\mathrm{id}_{C(\mathrm{Mod}\,R)} \to \operatorname{tot} L^{\mathbb{W}}$ of functors. Moreover, if M is an R-module, then $\operatorname{tot} L^{\mathbb{W}}M = L^{\mathbb{W}}M$.

(ii) Let $a, b \in \mathbb{Z} \cup \{\pm\infty\}$ with $a \le b$ and X be a complex of R-modules such that $X^i = 0$ for $i \notin [a, b]$. Then it holds that $(\operatorname{tot} L^{\mathbb{W}}X)^i = 0$ for $i \notin [a, b+n]$, where n is the number given to $\mathbb{W} = \{W_i\}_{0 \le i \le n}$.

(iii) Let X be a complex of flat R-modules. Then we see that $\operatorname{tot} L^{\mathbb{W}}X$ consists of cotorsion flat R-modules with cosupports in $\bigcup_{0 \le i \le n} W_i$.

Definition 7.6. Let W be a nonempty subset of $\operatorname{Spec} R$ and $\{W_i\}_{0 \le i \le n}$ be a family of subsets of W. We say that $\{W_i\}_{0 \le i \le n}$ is *a system of slices of W* if the following conditions hold:

(1) $W = \bigcup_{0 \le i \le n} W_i$.

(2) $W_i \cap W_j = \varnothing$ if $i \ne j$.

(3) $\dim W_i = 0$ for $0 \le i \le n$.

(4) W_i is specialization-closed in $\bigcup_{i \le j \le n} W_j$ for each $0 \le i \le n$.

Compare this definition with the filtrations in [Hartshorne 1966, Chapter IV, §3].

If $\dim W$ is finite, then there exists at least one system of slices of W. Conversely, if there is a system of slices of W, then $\dim W$ is finite.

Proposition 7.7. *Let W be a subset of* $\operatorname{Spec} R$ *and $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ be a system of slices of W. Then, for any flat R-module F, there is an isomorphism in $\mathcal{D}$;*

$$\lambda^W F \cong L^{\mathbb{W}} F.$$

Under this isomorphism, $\ell^{\mathbb{W}} F : F \to L^{\mathbb{W}} F$ coincides with $\eta^W F : F \to \lambda^W F$ in $\mathcal{D}$.

Proof. We use induction on n, which is the number given to $\mathbb{W} = \{W_i\}_{0 \le i \le n}$. Suppose that $n = 0$. It then holds that $L^{\mathbb{W}} F = \bar{\lambda}^{W_0} F = \bar{\lambda}^W F$ and $\ell^{\mathbb{W}} F = \bar{\eta}^{W_0} F = \bar{\eta}^W F$. Hence this proposition follows from Theorem 3.10.

Next, suppose that $n > 0$, and write $U = \bigcup_{1 \le i \le n} W_i$. Setting $U_{i-1} = W_i$, we obtain a system of slices $\mathbb{U} = \{U_i\}_{0 \le i \le n-1}$ of U. Consider the following two squares, where the first and second are in $C(\operatorname{Mod} R)$ and $\mathcal{D}$, respectively:

$$\begin{array}{ccc} F & \xrightarrow{\bar{\eta}^{W_0} F} & \bar{\lambda}^{W_0} F \\ \big\downarrow{\scriptstyle \ell^{\mathbb{U}} F} & & \big\downarrow{\scriptstyle \ell^{\mathbb{U}} \bar{\lambda}^{W_0} F} \\ L^{\mathbb{U}} F & \xrightarrow{L^{\mathbb{U}} \bar{\eta}^{W_0} F} & L^{\mathbb{U}} \bar{\lambda}^{W_0} F \end{array} \qquad \begin{array}{ccc} F & \xrightarrow{\eta^{W_0} F} & \lambda^{W_0} F \\ \big\downarrow{\scriptstyle \eta^{U} F} & & \big\downarrow{\scriptstyle \eta^{U} \lambda^{W_0} F} \\ \lambda^{U} F & \xrightarrow{\lambda^{U} \eta^{W_0} F} & \lambda^{U} \lambda^{W_0} F \end{array}$$

By Remarks 7.5(i) and 3.14(ii), both of them are commutative. Moreover, $\lambda^U \eta^{W_0} F$ is the unique morphism which makes the right square commutative, because λ^U is a left adjoint to the inclusion functor $\mathcal{C}^U \hookrightarrow \mathcal{D}$. Then, regarding the left square as being in $\mathcal{D}$, we see from the inductive hypothesis that the left and right squares coincide in $\mathcal{D}$.

Let $\bar{g} : L^{\mathbb{U}} F \oplus \bar{\lambda}^{W_0} F \to L^{\mathbb{U}} \bar{\lambda}^{W_0} F$ and $\bar{h} : F \to L^{\mathbb{U}} F \oplus \bar{\lambda}^{W_0} F$ be chain maps represented by the matrices

$$\bar{g} = \begin{pmatrix} L^{\mathbb{U}} \bar{\eta}^{W_0} F & (-1) \cdot \ell^{\mathbb{U}} \bar{\lambda}^{W_0} X \end{pmatrix}, \quad \bar{h} = \begin{pmatrix} \ell^{\mathbb{U}} F \\ \bar{\eta}^{W_0} F \end{pmatrix}.$$

Notice that the mapping cone of $\bar{g}[-1]$ is nothing but $L^{\mathbb{W}} F$. Then we can obtain the following morphism of triangles, regarded as being in $\mathcal{D}$:

$$\begin{array}{ccccccc} F[-1] & \longrightarrow & F[-1] & \longrightarrow & 0 & \longrightarrow & F \\ \big\downarrow{\scriptstyle \ell^{\mathbb{W}} F[-1]} & & \big\downarrow{\scriptstyle \bar{h}[-1]} & & \big\downarrow & & \big\downarrow{\scriptstyle \ell^{\mathbb{W}} F} \\ L^{\mathbb{W}} F[-1] & \longrightarrow & (L^{\mathbb{U}} F \oplus \bar{\lambda}^{W_0} F)[-1] & \xrightarrow{\bar{g}[-1]} & L^{\mathbb{U}} \bar{\lambda}^{W_0} F[-1] & \longrightarrow & L^{\mathbb{W}} F \end{array} \tag{7.8}$$

Therefore, by Theorem 3.15 and Remark 3.19, there is an isomorphism $\lambda^W F \cong L^{\mathbb{W}} F$ such that $\ell^{\mathbb{W}} F$ coincides with $\eta^W F$ under this isomorphism. $\square$

The following corollary is one of the main results of this paper.

Corollary 7.9. *Let W and $\mathbb{W}=\{W_i\}_{0\le i\le n}$ be as above. Let X be a complex of flat R-modules. Then there is an isomorphism in $\mathcal{D}$;*

$$\lambda^W X \cong \operatorname{tot} L^{\mathbb{W}} X.$$

Under this isomorphism, $\ell^{\mathbb{W}}X : X \to \operatorname{tot} L^{\mathbb{W}} X$ coincides with $\eta^W X : X \to \lambda^W X$ in $\mathcal{D}$.

Proof. We embed $\ell^{\mathbb{W}}X : X \to \operatorname{tot} L^{\mathbb{W}} X$ into a triangle

$$C \longrightarrow X \xrightarrow{\ell^{\mathbb{W}}X} \operatorname{tot} L^{\mathbb{W}} X \longrightarrow C[1].$$

Proposition 6.3 and Remark 7.5(iii) imply that $\operatorname{tot} L^{\mathbb{W}} X \in \mathcal{C}^W$. Thus it suffices to show that $\lambda^{W_i} C = 0$ for each i, by Lemma 2.6 and Remark 2.7(i). For this purpose, we prove that $\lambda^{W_i}\ell^{\mathbb{W}}X$ is an isomorphism in $\mathcal{D}$. This is equivalent to showing that $\bar{\lambda}^{W_i}\ell^{\mathbb{W}}X$ is a quasi-isomorphism, since X and $\operatorname{tot} L^{\mathbb{W}} X$ consist of flat R-modules.

Consider the natural morphism $X \to L^{\mathbb{W}} X$ of double complexes, which is induced by the chain maps $\ell^{\mathbb{W}}X^j : X^j \to L^{\mathbb{W}} X^j$ for $j \in \mathbb{Z}$. To prove that $\bar{\lambda}^{W_i}\ell^{\mathbb{W}}X$ is a quasi-isomorphism, it is enough to show that $\bar{\lambda}^{W_i}\ell^{\mathbb{W}}X^j$ is a quasi-isomorphism for each $j \in \mathbb{Z}$; see [Kashiwara and Schapira 2006, Theorem 12.5.4]. Furthermore, by Proposition 7.7, each $\ell^{\mathbb{W}}X^j$ coincides with $\eta^W X^j : X^j \to \lambda^W X^j$ in $\mathcal{D}$. Since $W_i \subseteq W$, it follows from Remark 2.7(ii) that $\lambda^{W_i}\eta^W X^j$ is an isomorphism in $\mathcal{D}$. This means that $\bar{\lambda}^{W_i}\ell^{\mathbb{W}}X^j$ is a quasi-isomorphism. $\square$

Let W be a subset of $\operatorname{Spec} R$, and suppose that $n = \dim W$ is finite. Then Corollary 7.9 implies $\lambda^W R \in \mathcal{D}^{[0,n]}$. We give an example such that $H^n(\lambda^W R) \neq 0$.

Example 7.10. Let $(R, \mathfrak{m})$ be a local ring of dimension $d \geq 1$. Then we have $\dim V(\mathfrak{m})^c = d-1$. By Lemma 2.6, there is a triangle

$$\gamma_{V(\mathfrak{m})} R \longrightarrow R \longrightarrow \lambda^{V(\mathfrak{m})^c} R \longrightarrow \gamma_{V(\mathfrak{m})} R[1].$$

Since $\mathrm{R}\Gamma_{V(\mathfrak{m})} \cong \gamma_{V(\mathfrak{m})}$ by (2.10), Grothendieck's nonvanishing theorem implies that $H^d(\gamma_{V(\mathfrak{m})} R)$ is nonzero. Then we see from the triangle that $H^{d-1}(\lambda^{V(\mathfrak{m})^c} R) \neq 0$.

We denote by $\mathcal{D}^-$ the full subcategory of $\mathcal{D}$ consisting of complexes X such that $H^i(X) = 0$ for $i \gg 0$. Let W be a subset of $\operatorname{Spec} R$ and $X \in \mathcal{D}^-$. If $\dim W$ is finite, then we have $\lambda^W R \in \mathcal{D}^-$ by Corollary 7.9. However, as shown in the following example, it can happen that $\lambda^W R \notin \mathcal{D}^-$ when $\dim W$ is infinite.

Example 7.11. Assume that $\dim R = +\infty$, and set $W = \max(\operatorname{Spec} R)$. Then it holds that $\dim W = 0$ and $\dim W^c = +\infty$. Since each $\mathfrak{m} \in W$ is maximal, there are isomorphisms

$$\gamma_W \cong \mathrm{R}\Gamma_W \cong \bigoplus_{\mathfrak{m}\in W} \mathrm{R}\Gamma_{V(\mathfrak{m})}\,.$$

Thus we see from Example 7.10 that $\gamma_W R \notin \mathcal{D}^-$. Then, considering the triangle

$$\gamma_W R \longrightarrow R \longrightarrow \lambda^{W^c} R \longrightarrow \gamma_W R[1],$$

we have $\lambda^{W^c} R \notin \mathcal{D}^-$.

Let W be a subset of $\operatorname{Spec} R$ and $X \in \mathcal{C}^W$. Then $\eta^W X : X \to \lambda^W X$ is an isomorphism in $\mathcal{D}$. Thus Remark 7.5(iii) and Corollary 7.9 yield the following result.

Corollary 7.12. *Let W be a subset of* $\operatorname{Spec} R$, *and* $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ *be a system of slices of W. Let X be a complex of flat R-modules with* $\operatorname{cosupp} X \subseteq W$. *Then the chain map* $\ell^{\mathbb{W}} X : X \to \operatorname{tot} L^{\mathbb{W}} X$ *is a quasi-isomorphism, where* $\operatorname{tot} L^{\mathbb{W}} X$ *consists of cotorsion flat R-modules with cosupports in W.*

Remark 7.13. If $d = \dim R$ is finite, then any complex Y is quasi-isomorphic to a K-flat complex consisting of cotorsion flat R-modules. To see this, set

$$W_i = \{\mathfrak{p} \in \operatorname{Spec} R \mid \dim R/\mathfrak{p} = i\}$$

for $0 \le i \le d$. Then $\mathbb{W} = \{W_i\}_{0 \le i \le d}$ is a system of slices of $\operatorname{Spec} R$. We take a K-flat resolution X of Y such that X consists of flat R-modules. Corollary 7.12 implies that $\ell^{\mathbb{W}} X : X \to \operatorname{tot} L^{\mathbb{W}} X$ is a quasi-isomorphism, and $\operatorname{tot} L^{\mathbb{W}} X$ consists of cotorsion flat R-modules. At the same time, the chain maps $\ell^{\mathbb{W}} X^i : X^i \to L^{\mathbb{W}} X^i$ are quasi-isomorphisms for all $i \in \mathbb{Z}$. Then it is not hard to see that the mapping cone of $\ell^{\mathbb{W}} X$ is K-flat. Thus $\operatorname{tot} L^{\mathbb{W}} X$ is K-flat.

By Proposition 6.3 and Corollary 7.12, we have the next result.

Corollary 7.14. *Let W be a subset of* $\operatorname{Spec} R$ *such that* $\dim W$ *is finite. Then a complex $X \in \mathcal{D}$ belongs to $\mathcal{C}^W$ if and only if X is isomorphic to a complex Z of cotorsion flat R-modules such that* $\operatorname{cosupp} Z^i \subseteq W$ *for all $i \in \mathbb{Z}$.*

Remark 7.15. If $\dim W$ is infinite, it is possible to construct a similar family to systems of slices. We first put $W_0 = \max W$. Let $i > 0$ be an ordinal, and suppose that subsets W_j of W are defined for all $j < i$. Then we put $W_i = \max\bigl(W \setminus \bigcup_{j<i} W_j\bigr)$. In this way, we obtain the smallest ordinal $o(W)$ satisfying the following conditions:

(1) $W = \bigcup_{0 \le i < o(W)} W_i$.

(2) $W_i \cap W_j = \varnothing$ if $i \ne j$.

(3) $\dim W_i \le 0$ for $0 \le i < o(W)$.

(4) W_i is specialization-closed in $\bigcup_{i \le j < o(W)} W_j$ for each $0 \le i < o(W)$.

One should remark that the ordinal $o(W)$ can be uncountable in general; see [Gordon and Robson 1973, p. 48, Theorem 9.8]. However, if R is an infinite-dimensional commutative Noetherian ring given by Nagata [1962, Appendix A1, Example 1], then $o(W)$ is at most countable. Moreover, using transfinite induction, it is possible to extend condition (4) in Proposition 6.3 and Corollary 6.5 to the case where $o(W)$ is countable. One can also extend Corollary 7.14 to the case where $o(W)$ is countable.

Using Theorem 3.22 and results in [Nakamura and Yoshino 2018, §3], it is possible to give a similar result to Corollary 7.9, for colocalization functors γ_W and complexes of injective R-modules.

8. Čech complexes and complexes of finitely generated modules

Let W be a subset of $\operatorname{Spec} R$ and $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ be a system of slices of W. In this section, we prove that $\lambda^W Y$ is isomorphic to $\operatorname{tot} L^{\mathbb{W}} Y$ if Y is a complex of finitely generated R-modules.

We denote by $\mathcal{D}_{\mathrm{fg}}$ the full subcategory of $\mathcal{D}$ consisting of all complexes with finitely generated cohomology modules, and set $\mathcal{D}^-_{\mathrm{fg}} = \mathcal{D}^- \cap \mathcal{D}_{\mathrm{fg}}$. We first prove the following proposition.

Proposition 8.1. *Let W be a subset of* $\operatorname{Spec} R$ *such that* $\dim W$ *is finite. Let $X, Y \in \mathcal{D}$. We suppose that one of the following conditions holds*:

(1) *$X \in \mathcal{D}^-$ and $Y \in \mathcal{D}^-_{\mathrm{fg}}$.*

(2) *X is a bounded complex of flat R-modules and $Y \in \mathcal{D}_{\mathrm{fg}}$.*

Then there are natural isomorphisms

$$(\gamma_{W^c} X) \otimes^{\mathrm{L}}_R Y \cong \gamma_{W^c}(X \otimes^{\mathrm{L}}_R Y), \quad (\lambda^W X) \otimes^{\mathrm{L}}_R Y \cong \lambda^W (X \otimes^{\mathrm{L}}_R Y).$$

For $X \in \mathcal{D}$ and $n \in \mathbb{Z}$, we define the cohomological truncations $\sigma_{\le n} X$ and $\sigma_{>n} X$ as follows (see [Hartshorne 1966, Chapter I, §7]):

$$\begin{aligned} \sigma_{\le n} X &= (\cdots \to X^{n-2} \to X^{n-1} \to \operatorname{Ker} d^n_X \to 0 \to \cdots), \\ \sigma_{>n} X &= (\cdots \to 0 \to \operatorname{Im} d^n_X \to X^{n+1} \to X^{n+2} \to \cdots). \end{aligned}$$

Proof of Proposition 8.1. Applying $(-) \otimes^{\mathrm{L}}_R Y$ to the triangle

$$\gamma_{W^c} X \to X \to \lambda^W X \to \gamma_{W^c} X[1],$$

we obtain the triangle

$$(\gamma_{W^c} X) \otimes^{\mathrm{L}}_R Y \longrightarrow X \otimes^{\mathrm{L}}_R Y \longrightarrow (\lambda^W X) \otimes^{\mathrm{L}}_R Y \longrightarrow (\gamma_{W^c} X) \otimes^{\mathrm{L}}_R Y[1].$$

Since $\operatorname{supp} \gamma_{W^c} X \subseteq W^c$, we have $\operatorname{supp}(\gamma_{W^c} X) \otimes^{\mathrm{L}}_R Y \subseteq W^c$, i.e., $(\gamma_{W^c} X) \otimes^{\mathrm{L}}_R Y \in \mathcal{L}_{W^c}$. Hence it remains to show that $(\lambda^W X) \otimes^{\mathrm{L}}_R Y \in \mathcal{C}^W$; see Lemma 2.6.

Case 1: We remark that X is isomorphic to a right bounded complex of flat R-modules. Then it is seen from Corollary 7.9 that $\lambda^W X$ is isomorphic to a right bounded complex Z of cotorsion flat R-modules such that $\operatorname{cosupp} Z^i \subseteq W$ for all $i \in \mathbb{Z}$. Furthermore, Y is isomorphic to a right bounded complex P of finite free R-modules. Hence it follows that $X \otimes^{\mathrm{L}}_R Y \cong Z \otimes_R P$, where the second one consists of cotorsion flat R-modules with cosupports in W. Then we have $X \otimes^{\mathrm{L}}_R Y \cong Z \otimes_R P \in \mathcal{C}^W$ by Proposition 6.3.

Case 2: By Corollary 7.9, $\lambda^W X$ is isomorphic to a bounded complex consisting of cotorsion flat R-modules with cosupports in W. Thus it is enough to prove that $Z \otimes_R Y \in \mathcal{C}^W$ for a cotorsion flat R-module Z with $\operatorname{cosupp} Z \subseteq W$.

We consider the triangle $\sigma_{\le n} Y \to Y \to \sigma_{>n} Y \to \sigma_{\le n} Y[1]$ for an integer n. Applying $Z \otimes_R (-)$ to this triangle, we obtain the following one:

$$Z \otimes_R \sigma_{\le n} Y \longrightarrow Z \otimes_R Y \longrightarrow Z \otimes_R \sigma_{>n} Y \longrightarrow Z \otimes_R \sigma_{\le n} Y[1].$$

Let $\mathfrak{p} \in W^c$. Case 1 implies that $Z \otimes_R \sigma_{\le n} Y \in \mathcal{C}^W$ for any $n \in \mathbb{Z}$, since $\lambda^W Z \cong Z$. Thus, applying $\mathrm{RHom}_R(\kappa(\mathfrak{p}), -)$ to the triangle above, we have

$$\mathrm{RHom}_R(\kappa(\mathfrak{p}), Z \otimes_R Y) \cong \mathrm{RHom}_R(\kappa(\mathfrak{p}), Z \otimes_R \sigma_{>n} Y).$$

Furthermore, taking a projective resolution P of $\kappa(\mathfrak{p})$, we have

$$\mathrm{RHom}_R(\kappa(\mathfrak{p}), Z \otimes_R \sigma_{>n} Y) \cong \mathrm{Hom}_R(P, Z \otimes_R \sigma_{>n} Y).$$

Let j be any integer. To see that $\mathrm{RHom}_R(\kappa(\mathfrak{p}), Z \otimes_R Y) = 0$, it suffices to show that there exists an integer n such that $H^0(\mathrm{Hom}_R(P[j], Z \otimes_R \sigma_{>n} Y)) = 0$. Note that $P^i = 0$ for $i > 0$. Moreover, each element of $H^0(\mathrm{Hom}_R(P[j], Z \otimes_R \sigma_{>n} Y)) \cong \mathrm{Hom}_{\mathcal{D}}(P[j], Z \otimes_R \sigma_{>n} Y)$ is represented by a chain map $P[j] \to Z \otimes_R \sigma_{>n} Y$. Therefore it holds that $H^0(\mathrm{Hom}_R(P[j], Z \otimes_R \sigma_{>n} Y)) = 0$ if $n > -j$. $\square$

Remark 8.2. (i) In the proposition, we can remove the finiteness condition on $\dim W$ if $W = V(\mathfrak{a})$ for an ideal $\mathfrak{a}$. In such cases, we need only use $\mathfrak{a}$-adic completions of free R-modules instead of cotorsion flat R-modules.

(ii) If W is a generalization-closed subset of $\operatorname{Spec} R$, then the isomorphisms in the proposition hold for any $X, Y \in \mathcal{D}$ because γ_{W^c} is isomorphic to $\mathrm{R}\Gamma_{W^c}$.

Let W be a subset of $\operatorname{Spec} R$ and $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ be a system of slices of W. Let $Y \in \mathcal{D}_{\mathrm{fg}}$. By Propositions 8.1 and 7.7, we have

$$\lambda^W Y \cong (\lambda^W R) \otimes^{\mathrm{L}}_R Y \cong (L^{\mathbb{W}} R) \otimes_R Y. \tag{8.3}$$

Let F be a flat R-module and M be a finitely generated R-module. Then we see from Lemma 2.3 that

$$(\bar{\lambda}^{W_i} F) \otimes_R M \cong \bar{\lambda}^{W_i}(F \otimes_R M).$$

This fact ensures that $(\bar{\lambda}^{(i_m,\dots,i_1,i_0)} R) \otimes_R M \cong \bar{\lambda}^{(i_m,\dots,i_1,i_0)} M$. Thus, if Y is a complex of finitely generated R-modules, then there is a natural isomorphism

$$(L^{\mathbb{W}} R) \otimes_R Y \cong \operatorname{tot} L^{\mathbb{W}} Y \tag{8.4}$$

in $C(\mathrm{Mod}\, R)$. By (8.3) and (8.4), we have shown the following proposition.

Proposition 8.5. *Let W be a subset of* Spec R *and* $\mathbb{W} = \{W_i\}_{0\le i\le n}$ *be a system of slices of W. Let Y be a complex of finitely generated R-modules. Then there is an isomorphism in $\mathcal{D}$*;

$$\lambda^W Y \cong \operatorname{tot} L^{\mathbb{W}} Y.$$

Under this identification, $\ell^{\mathbb{W}} Y : Y \to \operatorname{tot} L^{\mathbb{W}} Y$ coincides with $\eta^W Y : Y \to \lambda^W Y$ in $\mathcal{D}$.

We see from (8.4) and the remark below that it is also possible to give a quick proof of this proposition, provided that Y is a right bounded complex of finitely generated R-modules.

Remark 8.6. Let W be a subset of Spec R and $\mathbb{W} = \{W_i\}_{0\le i\le n}$ be a system of slices of W. We denote by $K(\operatorname{Mod} R)$ the homotopy category of complexes of R-modules. Note that $\operatorname{tot} L^{\mathbb{W}}$ induces a triangulated functor $K(\operatorname{Mod} R) \to K(\operatorname{Mod} R)$, which we also write $\operatorname{tot} L^{\mathbb{W}}$. Then it is seen from Corollary 7.9 that $\lambda^W : \mathcal{D} \to \mathcal{D}$ is isomorphic to the left derived functor of $\operatorname{tot} L^{\mathbb{W}} : K(\operatorname{Mod} R) \to K(\operatorname{Mod} R)$.

Let W be a subset of Spec R such that $n = \dim W$ is finite. By Proposition 8.5, if an R-module M is finitely generated, then $\lambda^W M \in \mathcal{D}^{[0,n]}$. On the other hand, since $\lambda^{V(\mathfrak{a})} \cong \mathrm{L}\Lambda^{V(\mathfrak{a})}$ for an ideal $\mathfrak{a}$, it can happen that $H^i(\lambda^W M) \neq 0$ for some $i < 0$ when M is not finitely generated; see [Nakamura and Yoshino 2018, Example 5.3].

Remark 8.7. Let $n \ge 0$ be an integer. Let $\mathfrak{a}_i$ be ideals of R and S_i be multiplicatively closed subsets of R for $0 \le i \le n$. In Notation 7.2 and Definition 7.4, one can replace $\bar{\lambda}^{(i)} = \bar{\lambda}^{W_i}$ by $\Lambda^{V(\mathfrak{a}_i)}(- \otimes_R S_i^{-1} R)$, and construct a kind of Čech complex. For this Čech complex and λ^W with $W = \bigcup_{0\le i\le n}(V(\mathfrak{a}_i) \cap U_{S_i})$, it is possible to show similar results to Corollary 7.9 and Proposition 8.5, provided that one of the following conditions holds:

(1) $V(\mathfrak{a}_i) \cap U_{S_i}$ is specialization-closed in $\bigcup_{i\le j\le n}(V(\mathfrak{a}_j) \cap U_{S_j})$ for each $0 \le i \le n$.

(2) $V(\mathfrak{a}_i) \cap U_{S_i}$ is generalization-closed in $\bigcup_{0\le j\le i}(V(\mathfrak{a}_j) \cap U_{S_j})$ for each $0 \le i \le n$.

9. Čech complexes and complexes of pure-injective modules

In this section, as an application, we give a functorial way to construct a quasi-isomorphism from a complex of flat R-modules or a complex of finitely generated R-modules to a complex of pure-injective R-modules.

We start with the following well known fact.

Lemma 9.1. *Let X be a complex of flat R-modules and Y be a complex of cotorsion R-modules. We assume that one of the following conditions holds*:

(1) *X is right bounded and Y is left bounded.*

(2) *X is bounded and* $\dim R$ *is finite.*

Then we have $\operatorname{RHom}_R(X, Y) \cong \operatorname{Hom}_R(X, Y)$.

One can prove this lemma by [Kashiwara and Schapira 2006, Theorem 12.5.4] and Theorem 4.1.

Next, we recall the notion of pure-injective modules and resolutions. We say that a morphism $f : M \to N$ of R-modules is pure if $f \otimes_R L$ is a monomorphism in $\operatorname{Mod} R$ for any R-module L. Moreover, an R-module P is called pure-injective if $\operatorname{Hom}_R(f, P)$ is an epimorphism in $\operatorname{Mod} R$ for any pure morphism $f : M \to N$ of R-modules. Clearly, all injective R-modules are pure-injective. Furthermore, all pure-injective R-modules are cotorsion; see [Enochs and Jenda 2000, Lemma 5.3.23].

Let M be an R-module. A complex P together with a quasi-isomorphism $M \to P$ is called a pure-injective resolution of M if P consists of pure-injective R-modules and $P^i = 0$ for $i < 0$. It is known that any R-module has a minimal pure-injective resolution, which is constructed by using pure-injective envelopes, see [Enochs 1987] and [Enochs and Jenda 2000, Example 6.6.5, Definition 8.1.4]. Moreover, if F is a flat R-module and P is a pure-injective resolution of M, then we have $\operatorname{RHom}_R(F, M) \cong \operatorname{Hom}_R(F, P)$ by Lemma 9.1.

Now we observe that any cotorsion flat R-module is pure-injective. Consider an R-module of the form $\big(\bigoplus_B R_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}}$ with some index set B and a prime ideal $\mathfrak{p}$, which is a cotorsion flat R-module. Writing $E_R(R/\mathfrak{p})$ for the injective hull of $R/\mathfrak{p}$, we have

$$\Big(\bigoplus_B R_{\mathfrak{p}}\Big)^{\wedge}_{\mathfrak{p}} \cong \operatorname{Hom}_R\Big(E_R(R/\mathfrak{p}), \bigoplus_B E_R(R/\mathfrak{p})\Big);$$

see [Enochs and Jenda 2000, Theorem 3.4.1]. It follows from tensor-hom adjunction that $\operatorname{Hom}_R(M, I)$ is pure-injective for any R-module M and any injective R-module I. Hence $\big(\bigoplus_B R_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}}$ is pure-injective. Thus any cotorsion flat R-module is pure-injective; see Proposition 5.1.

There is another example of pure-injective R-modules. Let M be a finitely generated R-module. Using the Five Lemma, we are able to prove an isomorphism

$$\operatorname{Hom}_R\Big(E_R(R/\mathfrak{p}), \bigoplus_B E_R(R/\mathfrak{p})\Big) \otimes_R M \cong \operatorname{Hom}_R\Big(\operatorname{Hom}_R(M, E_R(R/\mathfrak{p})), \bigoplus_B E_R(R/\mathfrak{p})\Big).$$

Therefore $\big(\bigoplus_B R_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}} \otimes_R M$ is pure-injective; it is also isomorphic to $\big(\bigoplus_B M_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}}$ by Lemma 2.3. Further, Proposition 8.1 implies that $\operatorname{cosupp}\big(\bigoplus_B M_{\mathfrak{p}}\big)^{\wedge}_{\mathfrak{p}} \subseteq \{\mathfrak{p}\}$.

By the above observation, we see that Corollary 7.12, (8.4) and Proposition 8.5 yield the following theorem, which is one of the main results of this paper.

Theorem 9.2. *Let W be a subset of* $\operatorname{Spec} R$ *and $\mathbb{W} = \{W_i\}_{0 \le i \le n}$ be a system of slices of W. Let Z be a complex of flat R-modules or a complex of finitely generated*

R-modules. We assume that $\operatorname{cosupp} Z \subseteq W$. *Then* $\ell^{\mathbb{W}} Z : Z \to \operatorname{tot} L^{\mathbb{W}} Z$ *is a quasi-isomorphism, where* $\operatorname{tot} L^{\mathbb{W}} Z$ *consists of pure-injective R-modules with cosupports in W.*

Remark 9.3. Let N be a flat or finitely generated R-module. Suppose that $d = \dim R$ is finite. Set $W_i = \{\mathfrak{p} \in \operatorname{Spec} R \mid \dim R/\mathfrak{p} = i\}$ and $\mathbb{W} = \{W_i\}_{0 \le i \le d}$. By Theorem 9.2, we obtain a pure-injective resolution $\ell^{\mathbb{W}} N : N \to L^{\mathbb{W}} N$ of N, that is, there is an exact sequence of R-modules of the form

$$0 \to N \to \prod_{0 \le i_0 \le d} \bar{\lambda}^{(i_0)} N \to \prod_{0 \le i_0 < i_1 \le d} \bar{\lambda}^{(i_1, i_0)} N \to \cdots \to \bar{\lambda}^{(d, \dots, 0)} N \to 0.$$

We remark that, in $C(\operatorname{Mod} R)$, $L^{\mathbb{W}} N$ need not be isomorphic to a minimal pure-injective resolution P of N. In fact, when N is a projective or finitely generated R-module, it holds that $P^0 \cong \prod_{\mathfrak{m} \in W_0} \widehat{N_{\mathfrak{m}}} = \bar{\lambda}^{(0)} N$ (see [Warfield 1969, Theorem 3] and [Enochs and Jenda 2000, Remark 6.7.12]), while $(L^{\mathbb{W}} N)^0 = \prod_{0 \le i_0 \le d} \bar{\lambda}^{(i_0)} N$. Furthermore, Enochs [1987, Theorem 2.1] proved that if N is a flat R-module, then P^i is of the form $\prod_{\mathfrak{p} \in W_{\ge i}} T_{\mathfrak{p}}^i$ for $0 \le i \le d$ (see Notation 6.1), where

$$W_{\ge i} = \{\mathfrak{p} \in \operatorname{Spec} R \mid \dim R/\mathfrak{p} \ge i\}.$$

On the other hand, for a flat or finitely generated R-module N, the differential maps in the pure-injective resolution $L^{\mathbb{W}} N$ are concretely described. In addition, our approach based on the localization functor λ^W and the Čech complex $L^{\mathbb{W}}$ provide a natural morphism $\ell^{\mathbb{W}} : \mathrm{id}_{C(\operatorname{Mod} R)} \to \operatorname{tot} L^{\mathbb{W}}$ which induces isomorphisms in $\mathcal{D}$ for all complexes of flat R-modules and complexes of finitely generated R-modules. The reader should also compare Theorem 9.2 with [Thompson 2017b, Theorem 5.2].

We close this paper with the following example of Theorem 9.2.

Example 9.4. Let R be a 2-dimensional local domain with quotient field Q. Let $\mathbb{W} = \{W_i\}_{0 \le i \le 2}$ be as in Remark 9.3. Then $L^{\mathbb{W}} R$ is a pure-injective resolution of R, and $L^{\mathbb{W}} R$ is of the following form:

$$0 \to Q \oplus \Big(\prod_{\mathfrak{p} \in W_1} \widehat{R_{\mathfrak{p}}} \Big) \oplus \widehat{R} \to \Big(\prod_{\mathfrak{p} \in W_1} \widehat{R_{\mathfrak{p}}} \Big)_{(0)} \oplus (\widehat{R})_{(0)} \oplus \prod_{\mathfrak{p} \in W_1} \widehat{(\widehat{R})_{\mathfrak{p}}} \to \Big(\prod_{\mathfrak{p} \in W_1} \widehat{(\widehat{R})_{\mathfrak{p}}} \Big)_{(0)} \to 0$$

Acknowledgements

Nakamura is grateful to Srikanth Iyengar for his helpful comments and suggestions. Yoshino was supported by JSPS Grant-in-Aid for Scientific Research 26287008.

References

[Alonso Tarrío et al. 1997] L. Alonso Tarrío, A. Jeremías López, and J. Lipman, "Local homology and cohomology on schemes", *Ann. Sci. Éc. Norm. Supér* (4) **30**:1 (1997), 1–39. Correction available at https://tinyurl.com/lipmanfix. MR Zbl

[Benson et al. 2008] D. Benson, S. B. Iyengar, and H. Krause, "Local cohomology and support for triangulated categories", *Ann. Sci. Éc. Norm. Supér.* (4) **41**:4 (2008), 573–619. MR Zbl

[Benson et al. 2012] D. J. Benson, S. B. Iyengar, and H. Krause, "Colocalizing subcategories and cosupport", *J. Reine Angew. Math.* **673** (2012), 161–207. MR Zbl

[Bökstedt and Neeman 1993] M. Bökstedt and A. Neeman, "Homotopy limits in triangulated categories", *Compositio Math.* **86**:2 (1993), 209–234. MR Zbl

[Bourbaki 1961] N. Bourbaki, *Algèbre commutative: Chapitre 1: Modules plats; Chapitre 2: Localisation*, Hermann, Paris, 1961. MR Zbl

[Enochs 1984] E. Enochs, "Flat covers and flat cotorsion modules", *Proc. Amer. Math. Soc.* **92**:2 (1984), 179–184. MR Zbl

[Enochs 1987] E. E. Enochs, "Minimal pure injective resolutions of flat modules", *J. Algebra* **105**:2 (1987), 351–364. MR Zbl

[Enochs and Jenda 2000] E. E. Enochs and O. M. G. Jenda, *Relative homological algebra*, de Gruyter Expositions in Math. **30**, de Gruyter, Berlin, 2000. MR Zbl

[Foxby and Iyengar 2003] H.-B. Foxby and S. Iyengar, "Depth and amplitude for unbounded complexes", pp. 119–137 in *Commutative algebra* (Grenoble/Lyon, 2001), edited by L. L. Avramov et al., Contemp. Math. **331**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Gordon and Robson 1973] R. Gordon and J. C. Robson, *Krull dimension*, Mem. Amer. Math. Soc. **133**, Amer. Math. Soc., Providence, RI, 1973. MR Zbl

[Greenlees and May 1992] J. P. C. Greenlees and J. P. May, "Derived functors of I-adic completion and local homology", *J. Algebra* **149**:2 (1992), 438–453. MR Zbl

[Hartshorne 1966] R. Hartshorne, *Residues and duality*, Lecture Notes in Math. **20**, Springer, 1966. MR Zbl

[Kashiwara and Schapira 2006] M. Kashiwara and P. Schapira, *Categories and sheaves*, Grundlehren der Math. Wissenschaften **332**, Springer, 2006. MR Zbl

[Krause 2010] H. Krause, "Localization theory for triangulated categories", pp. 161–235 in *Triangulated categories*, edited by T. Holm et al., London Math. Soc. Lecture Note Ser. **375**, Cambridge Univ. Press, 2010. MR Zbl

[Lipman 2002] J. Lipman, "Lectures on local cohomology and duality", pp. 39–89 in *Local cohomology and its applications* (Guanajuato, 1999), edited by G. Lyubeznik, Lecture Notes in Pure and Appl. Math. **226**, Dekker, New York, 2002. MR Zbl

[Nagata 1962] M. Nagata, *Local rings*, Interscience Tracts in Pure and Appl. Math. **13**, Interscience, New York, 1962. MR Zbl

[Nakamura and Yoshino 2018] T. Nakamura and Y. Yoshino, "A local duality principle in derived categories of commutative Noetherian rings", *J. Pure Appl. Algebra* **222**:9 (2018), 2580–2595.

[Neeman 1992] A. Neeman, "The chromatic tower for $D(R)$", *Topology* **31**:3 (1992), 519–532. MR Zbl

[Neeman 2011] A. Neeman, "Colocalizing subcategories of $D(R)$", *J. Reine Angew. Math.* **653** (2011), 221–243. MR Zbl

[Raynaud and Gruson 1971] M. Raynaud and L. Gruson, “Critères de platitude et de projectivité: techniques de ‘platification’ d’un module”, *Invent. Math.* **13** (1971), 1–89. MR Zbl

[Sather-Wagstaff and Wicklein 2017] S. Sather-Wagstaff and R. Wicklein, “Support and adic finiteness for complexes”, *Comm. Algebra* **45**:6 (2017), 2569–2592. MR Zbl

[Simon 1990] A.-M. Simon, “Some homological properties of complete modules”, *Math. Proc. Cambridge Philos. Soc.* **108**:2 (1990), 231–246. MR Zbl

[Thompson 2017a] P. Thompson, “Cosupport computations for finitely generated modules over commutative Noetherian rings”, preprint, 2017. arXiv

[Thompson 2017b] P. Thompson, “Minimal complexes of cotorsion flat modules”, preprint, 2017. arXiv

[Warfield 1969] R. B. Warfield, Jr., “Purity and algebraic compactness for modules”, *Pacific J. Math.* **28** (1969), 699–719. MR Zbl

[Xu 1996] J. Xu, *Flat covers of modules*, Lecture Notes in Math. **1634**, Springer, 1996. MR Zbl

Received October 25, 2017. Revised February 23, 2018.

TSUTOMU NAKAMURA
GRADUATE SCHOOL OF NATURAL SCIENCE AND TECHNOLOGY
OKAYAMA UNIVERSITY
OKAYAMA
JAPAN
t.nakamura@s.okayama-u.ac.jp

YUJI YOSHINO
GRADUATE SCHOOL OF NATURAL SCIENCE AND TECHNOLOGY
OKAYAMA UNIVERSITY
TSUSHIMA-NAKA
OKAYAMA
JAPAN
yoshino@math.okayama-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.437

ZETA INTEGRALS FOR GSP(4) VIA BESSEL MODELS

RALF SCHMIDT AND LONG TRAN

We give a revised treatment of Piatetski-Shapiro's theory of zeta integrals and L-factors for irreducible, admissible representations of GSp(4, F) via Bessel models. We explicitly calculate the local L-factors in the nonsplit case for all representations. In particular, we introduce the new concept of Jacquet–Waldspurger modules which play a crucial role in our calculations.

1. Introduction

An irreducible, admissible representation of an algebraic reductive group over a local field is called *generic* if it has a Whittaker model. Whittaker models are one of the main tools to define local and global L-functions and ε-factors of representations. The theory was developed by Jacquet and Langlands for GL(2) following ideas of Tate's thesis for GL(1). The general case of GL(n) was developed in a series of works by Jacquet, Piatetski-Shapiro and Shalika. It is well known that any infinite dimensional irreducible, admissible representation of GL(2) is always generic.

Let F be a nonarchimedean local field of characteristic zero. Takloo-Bighash [2000] computed L-functions for all generic representations of the group GSp(4, F). It is similar to the theory of GL(n) in that the approach is based on the existence of Whittaker models and zeta integrals. The method was first introduced by Novodvorsky [1979] in the Corvallis conference. However, it turns out that there are many irreducible, admissible representations of GSp(4, F) which are not generic.

In the 1970s, Novodvorsky and Piatetski-Shapiro introduced the concept of Bessel models. In contrast to Whittaker models, every irreducible, admissible, infinite-dimensional representation of GSp(4, F) admits a Bessel model of some

MSC2010: 11F70.
Keywords: L-factors, GSp(4), Bessel models, Jacquet–Waldspurger modules.

kind; see Theorem 6.1.4 of [Roberts and Schmidt 2016]. Piatetski-Shapiro [1997] defined a new type of zeta integral with respect to Bessel models which led to a parallel method to the GL(2) case of defining local factors. However, some of his results were only sketched, and not many factors were calculated explicitly.

Danişman calculated many Piatetski-Shapiro L-factors explicitly in the case of nonsplit Bessel models. In [Danişman 2014], representations were treated whose Jacquet module with respect to the Siegel parabolic has at most length 2. In [Danişman 2015a], this was extended to length at most 3. Nongeneric supercuspidals were the topic of [Danişman 2015b].

In this work we revisit both Piatetski-Shapiro's original theory and Danişman's explicit calculations. We generalize the theory of [Piatetski-Shapiro 1997] in that we do not restrict ourselves to unitary representations. We also fill in some of the missing proofs, for example in the argument that generic representations do not admit "exceptional poles".

Generalizing Danişman's approach, we give a unified treatment of the asymptotics of Bessel functions in the nonsplit case which works for all representations. The key here is to consider a new type of finite-dimensional module $V_{N,T,\Lambda}$ associated to an irreducible, admissible representation (π, V) of $\mathrm{GSp}(4, F)$. These *Jacquet–Waldspurger modules* control the asymptotics of Bessel functions. Table 2 contains the semisimplifications of all Jacquet–Waldspurger modules, and Table 3 contains their precise algebraic structure as $F^\times$-modules. A key lemma in the nonsplit case is due to Danişman; see Proposition 4.3.3.

Once the asymptotic behavior is known, it is easy to calculate the *regular part* $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$ of the Piatetski-Shapiro L-factor; see Table 5. Our results show that in all generic cases, $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$ coincides with the usual spin Euler factor defined via the local Langlands correspondence, but for nongeneric representations these factors generally disagree. The results of Table 5 also imply that $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$ is independent of the choice of Bessel model.

2. Definitions and notations

Let F be a nonarchimedean local field of characteristic zero. Let $\mathfrak{o}$ be its ring of integers, $\mathfrak{p}$ the maximal ideal of $\mathfrak{o}$, and ϖ a generator of $\mathfrak{p}$. Let q be the cardinality of $\mathfrak{o}/\mathfrak{p}$. We fix a nontrivial character ψ of F. Let v be the normalized valuation on F, and let ν or $|\cdot|$ be the normalized absolute value on F. Hence $\nu(x) = q^{-v(x)}$ for $x \in F^\times$.

Let $\mathrm{GSp}(4, F) := \{g \in \mathrm{GL}(4, F) : {}^t g J g = \lambda J, \text{ for some } \lambda = \lambda(g) \in F^\times\}$ be defined with respect to the symplectic form

$$J = \begin{bmatrix} & 1_2 \\ -1_2 & \end{bmatrix}. \tag{1}$$

Let $P = MN$ be the Levi decomposition of the Siegel parabolic subgroup P, where

$$(2)\qquad P = \mathrm{GSp}(4, F) \cap \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ & & * & * \\ & & * & * \end{bmatrix}, \quad N = \left\{ \begin{bmatrix} 1 & & x & y \\ & 1 & y & z \\ & & 1 & \\ & & & 1 \end{bmatrix} : x, y, z \in F \right\}$$

and $M = \left\{ \begin{bmatrix} xA & \\ & {}^tA^{-1} \end{bmatrix} : A \in \mathrm{GL}(2, F),\ x \in F^\times \right\}$. We let

$$(3)\qquad H := \left\{ \begin{bmatrix} xI_2 & \\ & I_2 \end{bmatrix} : x \in F^\times \right\} \cong F^\times.$$

Let

$$(4)\qquad \beta = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}, \quad a, b, c \in F$$

be a symmetric matrix. Then β determines a character ψ_β of N by

$$(5)\qquad \psi_\beta\left(\begin{bmatrix} 1 & X \\ & 1 \end{bmatrix} \right) = \psi(\mathrm{tr}(\beta X)), \quad X = \begin{bmatrix} x & y \\ y & z \end{bmatrix}.$$

Every character of N is of this form for a uniquely determined β. We say that ψ_β is *nondegenerate* if $\beta \in \mathrm{GL}(2, F)$.

Attached to a nondegenerate ψ_β is a quadratic extension L/F. If $-\det(\beta) \notin F^{\times 2}$, we set $L = F(\sqrt{-\det(\beta)})$; this is the *nonsplit case*. If $-\det(\beta) \in F^{\times 2}$, we set $L = F \oplus F$; this is the *split case*. Let

$$(6)\qquad \begin{aligned} A_\beta &= \{ g \in M_2(F) : {}^tg\beta g = \det(g)\beta \} \\ &= \left\{ \begin{bmatrix} x + yb/2 & yc \\ -ya & x - yb/2 \end{bmatrix} : x, y \in F \right\}. \end{aligned}$$

Then A_β is an F-algebra isomorphic to L via the map

$$(7)\qquad \begin{bmatrix} x + yb/2 & yc \\ -ya & x - yb/2 \end{bmatrix} \longmapsto x + y\Delta,$$

where $\Delta = \sqrt{-\det(\beta)}$ in the nonsplit case, and $\Delta = (-\delta, \delta)$ if $-\det(\beta) = \delta^2$.

Let T be the connected component of the stabilizer of ψ_β in M. It is easy to check that $T \cong A_\beta^\times \cong L^\times$. We always consider T a subgroup of $\mathrm{GSp}(4, F)$ via

$$(8)\qquad T \ni g \longmapsto \begin{bmatrix} g & \\ & \det(g)\,{}^tg^{-1} \end{bmatrix}.$$

Explicitly, T consists of all elements

$$(9)\qquad \begin{bmatrix} x+yb/2 & yc & & \\ -ya & x-yb/2 & & \\ & & x-yb/2 & ya \\ & & -yc & x+yb/2 \end{bmatrix}, \qquad x, y \in F,\ x^2 - y^2\Delta^2 \neq 0.$$

Let $R := TN$ be the *Bessel subgroup* of $\mathrm{GSp}(4, F)$. If Λ is a character of T, then we can define a character $\Lambda \otimes \psi_\beta$ of R by $tn \mapsto \Lambda(t)\psi_\beta(n)$ for $t \in T$ and $n \in N$.

Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. Nonzero elements of $\mathrm{Hom}_R(V, \mathbb{C}_{\Lambda\otimes\psi_\beta})$ are called (Λ, β)-*Bessel functionals*. It is known that if such a Bessel functional ℓ exists, then $\mathrm{Hom}_R(V, \mathbb{C}_{\Lambda\otimes\psi_\beta})$ is one-dimensional. In this case the space of functions

$$(10)\qquad \mathcal{B}(\pi, \Lambda, \beta) := \{B_v : g \mapsto \ell(\pi(v)g) : v \in V\},$$

endowed with the action of $\mathrm{GSp}(4, F)$ given by right translations, is called the (Λ, β)-*Bessel model* of π.

3. Jacquet–Waldspurger modules

In this section we introduce a certain finite-dimensional $F^\times$-module attached to an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. Since it is derived from the usual Jacquet module by applying a Waldspurger functor, we call it a *Jacquet–Waldspurger module*. Its relevance is that it controls the asymptotics of Bessel functions along the subgroup H defined in (3). The main result of this section is Table 2, which lists the semisimplifications of the Jacquet–Waldspurger modules in the nonsplit case for all representations.

3.1. ***Jacquet modules.*** Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$,

$$V(N) = \langle \pi(n)v - v \mid v \in V,\ n \in N\rangle \quad \text{and} \quad V_N = V/V(N)$$

be the usual Jacquet module with respect to the Siegel parabolic subgroup. We identify M with $\mathrm{GL}(2, F) \times \mathrm{GL}(1, F)$ via the map

$$(11)\qquad (A, x) \longmapsto \begin{bmatrix} xA & \\ & \det(A)\,{}^tA^{-1} \end{bmatrix}, \quad A \in \mathrm{GL}(2, F),\ x \in F^\times,$$

so V_N carries an action of M, and thus an action of $\mathrm{GL}(2, F) \times \mathrm{GL}(1, F)$ via this isomorphism. We have tabulated the semisimplifications of these Jacquet modules in Table 1. Note that this table differs from Table A.3 of [Roberts and Schmidt 2007] in three ways:

- Roberts and Schmidt used a different version of GSp(4, F). Switching the last two rows and columns provides an isomorphism.
- The Jacquet modules listed in [Roberts and Schmidt 2007, Table A.3] are normalized, while the Jacquet modules listed in Table 1 are not. The normalized Jacquet module is obtained from the unnormalized one by twisting by $\delta_P^{-1/2}$, where
$$\delta_P\left(\begin{bmatrix} A & \\ & x\,{}^tA^{-1}\end{bmatrix}\right) = |x^{-1}\det(A)|^3.$$
Hence, we replace each component $\tau \otimes \sigma$ in [Roberts and Schmidt 2007, Table A.3] by $(\nu^{3/2}\tau)\otimes(\nu^{-3/2}\sigma)$ in order to obtain the unnormalized Jacquet modules.
- Roberts and Schmidt used the isomorphism
$$(A, x) \longmapsto \begin{bmatrix} A & \\ & x\,{}^tA^{-1}\end{bmatrix}, \quad A \in \mathrm{GL}(2, F),\ x \in F^\times. \tag{12}$$
Calculations show that we have to replace each component $(\nu^{3/2}\tau)\otimes(\nu^{-3/2}\sigma)$ of the unnormalized Jacquet module by $(\sigma\tau)\otimes(\nu^{3/2}\omega_\tau\sigma)$.

3.2. ***Waldspurger functionals for* GL(2).** Recall the algebra $A_\beta \subset M_2(F)$ defined in (6), and its unit group $T \subset \mathrm{GL}(2, F)$. Let Λ be a character of T. Let (τ, V) be a smooth representation of GL(2, F) admitting a central character ω_τ. A Λ-*Waldspurger functional* on τ is a nonzero linear map $\delta : V \to \mathbb{C}$ such that
$$\delta(\tau(t)v) = \Lambda(t)\delta(v) \quad \text{for all } v \in V \text{ and } t \in T.$$
Since T contains the center Z of GL(2, F), a necessary condition for such a δ to exist is that $\Lambda|_{F^\times} = \omega_\tau$. As in the case of Bessel functionals, we call a Waldspurger functional *split* if $-\det(\beta) \in F^{\times 2}$, otherwise *nonsplit*.

The (Λ, β)-Waldspurger functionals are the nonzero elements of the space $\mathrm{Hom}_T(\tau, \mathbb{C}_\Lambda)$. If we put
$$V(T, \Lambda) = \langle \tau(t)v - \Lambda(t)v : v \in V,\ t \in T\rangle \quad \text{and} \quad V_{T,\Lambda} = V/V(T, \Lambda), \tag{13}$$
then $\mathrm{Hom}_T(\tau, \mathbb{C}_\Lambda) \cong \mathrm{Hom}(V_{T,\Lambda}, \mathbb{C})$. Note that if L is a field, so that T/Z is compact, then the space $V(T, \Lambda)$ can also be characterized as
$$V(T, \Lambda) = \left\{ v \in V : \int_{T/Z} \Lambda(t)^{-1}\tau(t)v\, dt = 0 \right\}. \tag{14}$$
The map $V \mapsto V_{T,\Lambda}$ defines a functor, called the *Waldspurger functor*, from the category of smooth representations of GL(2, F) to the category of $F^\times$-modules. This can be seen just as the analogous statement in the case of Jacquet modules. In

particular, if L is a field, then the Waldspurger functor is exact; this follows from (14) with similar arguments as in [Bernstein and Zelevinskii 1976, Proposition 2.35].

		representation	semisimplification
I		$\chi_1\times\chi_2\rtimes\sigma$ (irreducible)	$\sigma(\chi_1\times\chi_2)\otimes\nu^{3/2}\chi_1\chi_2\sigma+\sigma(\chi_2\times\chi_1)\otimes\nu^{3/2}\sigma$ $+\sigma(\chi_1\chi_2\times 1_{F^\times})\otimes\nu^{3/2}\chi_1\sigma+\sigma(\chi_1\chi_2\times 1_{F^\times})\otimes\nu^{3/2}\chi_2\sigma$
II	a	$\chi\mathrm{St}_{\mathrm{GL}(2)}\rtimes\sigma$	$\sigma\chi\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu^{3/2}\chi^2\sigma+\sigma\chi\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu^{3/2}\sigma+(\chi^2\sigma\times\sigma)\otimes\nu^2\chi\sigma$
	b	$\chi 1_{\mathrm{GL}(2)}\rtimes\sigma$	$\sigma\chi 1_{\mathrm{GL}(2)}\otimes\nu^{3/2}\chi^2\sigma+\sigma\chi 1_{\mathrm{GL}(2)}\otimes\nu^{3/2}\sigma+(\chi^2\sigma\times\sigma)\otimes\nu\chi\sigma$
III	a	$\chi\rtimes\sigma\mathrm{St}_{\mathrm{GSp}(2)}$	$\sigma(\chi\nu^{-1/2}\times\nu^{1/2})\otimes\chi\nu^2\sigma+\sigma(\chi\nu^{1/2}\times\nu^{-1/2})\otimes\nu^2\sigma$
	b	$\chi\rtimes\sigma 1_{\mathrm{GSp}(2)}$	$\sigma(\chi\nu^{1/2}\times\nu^{-1/2})\otimes\chi\nu\sigma+\sigma(\chi\nu^{-1/2}\times\nu^{1/2})\otimes\nu\sigma$
IV	a	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu^3\sigma$
	b	$L(\nu^2,\nu^{-1}\sigma\mathrm{St}_{\mathrm{GSp}(2)})$	$\sigma 1_{\mathrm{GL}(2)}\otimes\nu^3\sigma+\sigma(\nu^{3/2}\times\nu^{-3/2})\otimes\nu\sigma$
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-3/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\sigma+\sigma(\nu^{3/2}\times\nu^{-3/2})\otimes\nu^2\sigma$
	d	$\sigma 1_{\mathrm{GSp}(4)}$	$\sigma 1_{\mathrm{GL}(2)}\otimes\sigma$
V	a	$\delta([\xi,\nu\xi],\nu^{-1/2}\sigma)$	$\sigma\xi\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu^2\sigma+\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\xi\nu^2\sigma$
	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$	$\sigma\xi\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu\sigma+\sigma 1_{\mathrm{GL}(2)}\otimes\xi\nu^2\sigma$
	c	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\xi\nu^{-1/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\xi\nu\sigma+\sigma\xi 1_{\mathrm{GL}(2)}\otimes\nu^2\sigma$
	d	$L(\nu\xi,\xi\rtimes\nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}\otimes\xi\nu\sigma+\sigma\xi 1_{\mathrm{GL}(2)}\otimes\nu\sigma$
VI	a	$\tau(S,\nu^{-1/2}\sigma)$	$2\cdot(\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu^2\sigma)+\sigma 1_{\mathrm{GL}(2)}\otimes\nu^2\sigma$
	b	$\tau(T,\nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}\otimes\nu^2\sigma$
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu\sigma$
	d	$L(\nu,1_{F^\times}\rtimes\nu^{-1/2}\sigma)$	$2\cdot(\sigma 1_{\mathrm{GL}(2)}\otimes\nu\sigma)+\sigma\mathrm{St}_{\mathrm{GL}(2)}\otimes\nu\sigma$
VII		$\chi\rtimes\pi$	0
VIII	a	$\tau(S,\pi)$	0
	b	$\tau(T,\pi)$	0
IX	a	$\delta(\nu\xi,\nu^{-1/2}\pi(\mu))$	0
	b	$L(\nu\xi,\nu^{-1/2}\pi(\mu))$	0
X		$\pi\rtimes\sigma$	$\sigma\pi\otimes\nu^{3/2}\omega_\pi\sigma+\sigma\pi\otimes\nu^{3/2}\sigma$
XI	a	$\delta(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$	$\sigma\pi\otimes\nu^2\sigma$
	b	$L(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$	$\sigma\pi\otimes\nu\sigma$
		supercuspidal	0

Table 1. Jacquet modules with respect to P, using the isomorphism (11).

Now assume that (τ, V) is irreducible and admissible. Then it is known by [Tunnell 1983; Saito 1993; Waldspurger 1985, Lemme 8] that the space $\mathrm{Hom}_T(\tau, \mathbb{C}_\Lambda)$ is at most one-dimensional. It follows that

$$\dim V_{T,\Lambda} \le 1. \tag{15}$$

The following facts are known for any character Λ of T such that $\Lambda|_{F^\times} = \omega_\tau$:

- For principal series representations, we have

$$\dim(\mathrm{Hom}_T(\chi_1 \times \chi_2, \mathbb{C}_\Lambda)) = 1 \quad \text{for all } \Lambda; \tag{16}$$

  see [Tunnell 1983, Proposition 1.6 and Theorem 2.3].

- For twists of the Steinberg representation, we have

$$\dim(\mathrm{Hom}_T(\sigma \mathrm{St}_{\mathrm{GL}(2)}, \mathbb{C}_\Lambda)) = \begin{cases} 0 & \text{if } L \text{ is a field and } \Lambda = \sigma \circ \mathrm{N}_{L/F}, \\ 1 & \text{otherwise;} \end{cases} \tag{17}$$

  see [Tunnell 1983, Proposition 1.7 and Theorem 2.4].

- If τ is infinite-dimensional and $L = F \times F$, then

$$\dim(\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda)) = 1 \quad \text{for all } \Lambda; \tag{18}$$

  see Lemme 8 of [Waldspurger 1985].

- For one-dimensional representations, we have

$$\dim(\mathrm{Hom}_T(\sigma 1_{\mathrm{GL}(2)}, \mathbb{C}_\Lambda)) = \begin{cases} 1 & \text{if } \Lambda = \sigma \circ \mathrm{N}_{L/F}, \\ 0 & \text{otherwise;} \end{cases} \tag{19}$$

  this is obvious.

3.3. ***Jacquet–Waldspurger modules.*** Recall the groups N and T defined in (2) and (9), respectively. Let (π, V) be an admissible representation of $\mathrm{GSp}(4, F)$. We now consider

$$\begin{aligned} V(N, T, \Lambda) &= \langle \pi(tn)v - \Lambda(t)v \,:\, v \in V,\ t \in T,\ n \in N \rangle \\ V_{N,T,\Lambda} &= V / V(N, T, \Lambda). \end{aligned} \tag{20}$$

Evidently, there is a surjective map $V_N \to V_{N,T,\Lambda}$ which induces an isomorphism

$$(V_N)_{T,\Lambda} \cong V_{N,T,\Lambda}. \tag{21}$$

Here, on the left we use the notation (13) for the $\mathrm{GL}(2, F)$-module V_N. Note that, in view of (8), we have to embed $\mathrm{GL}(2, F)$ into $\mathrm{GSp}(4, F)$ via the map

$$\mathrm{GL}(2, F) \ni g \longmapsto \begin{bmatrix} g & \\ & \det(g)\, {}^t g^{-1} \end{bmatrix}, \tag{22}$$

and consider V_N a GL(2, F)-module via this embedding. We call $V_{N,T,\Lambda}$ the *Jacquet–Waldspurger module* of π. This module retains an action of $F^\times$, coming from the action of the group $\{\mathrm{diag}(x, x, 1, 1) : x \in F^\times\}$ on V. The map $V \mapsto V_{N,T,\Lambda}$ defines a functor, called the *Jacquet–Waldspurger functor*, from the category of admissible GSp(4, F)-representations to the category of $F^\times$-modules.

Lemma 3.3.1. *Let V, V', V'' be admissible representations of* GSp(4, F).

(i) *If $V = V' \oplus V''$ is a direct sum, then*

$$V_{N,T,\Lambda} = V'_{N,T,\Lambda} \oplus V''_{N,T,\Lambda}. \tag{23}$$

(ii) *The Jacquet–Waldspurger functor is right exact, i.e, if $0 \to V' \to V \to V'' \to 0$ is exact, then*

$$V'_{N,T,\Lambda} \to V_{N,T,\Lambda} \to V''_{N,T,\Lambda} \to 0 \tag{24}$$

is exact. Moreover, if we are in the nonsplit case, then the Jacquet–Waldspurger functor is exact.

Proof. These are general properties of Jacquet-type functors. See Proposition 2.35 of [Bernstein and Zelevinskii 1976]. □

Lemma 3.3.2. *Let (π, V) be an admissible representation of* GSp(4, F) *of finite length. Then the $F^\times$-module $V_{N,T,\Lambda}$ is finite-dimensional. More precisely, if n is the length of the* GL(2, F)*-module V_N, then* $\dim V_{N,T,\Lambda} \le n$.

Proof. The proof is by induction on n. If $n = 1$, then V_N is an irreducible, admissible representation of GL(2, F). In this case the assertion follows from (15).

Assume that $n > 1$. Let V' be a submodule of V_N of length $n - 1$. Then $V'' := V_N/V'$ is irreducible. By (24), we have an exact sequence

$$V'_{T,\Lambda} \xrightarrow{\alpha} V_{N,T,\Lambda} \to V''_{T,\Lambda} \to 0. \tag{25}$$

By induction and (15), it follows that

$$\dim V_{N,T,\Lambda} = \dim \mathrm{im}(\alpha) + \dim V''_{T,\Lambda} \le n - 1 + 1 = n. \tag{26}$$

This concludes the proof. □

Assume that we are in the nonsplit case, i.e., the quadratic extension L is a field. Then the semisimplifications of the $V_{N,T,\Lambda}$ can easily be calculated from V_N using (21). We already noted that in the nonsplit case the Waldspurger functor is exact. Therefore, to calculate the $V_{N,T,\Lambda}$, we can simply take $(\tau \otimes \sigma)_{T,\Lambda}$ for each constituent $\tau \otimes \sigma$ occurring in Table 1. If $\tau_{T,\Lambda}$ is one-dimensional, then $(\tau \otimes \sigma)_{T,\Lambda} = \sigma 1_{F^\times}$ as an $F^\times$-module, and if $\tau_{T,\Lambda} = 0$, then $(\tau \otimes \sigma)_{T,\Lambda} = 0$. We have listed the semisimplifications of the $V_{N,T,\Lambda}$ for all irreducible, admissible representations in Table 2.

		representation	semisimplification of $V_{N,T,\Lambda}$
I		$\chi_1 \times \chi_2 \rtimes \sigma$ (irreducible)	$\nu^{3/2}\chi_1\chi_2\sigma 1_{F^\times} + \nu^{3/2}\sigma 1_{F^\times} + \nu^{3/2}\chi_1\sigma 1_{F^\times} + \nu^{3/2}\chi_2\sigma 1_{F^\times}$
II	a	$\chi \mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	$\nu^{3/2}\chi^2\sigma 1_{F^\times} + \nu^{3/2}\sigma 1_{F^\times} + \nu^2\chi\sigma 1_{F^\times}$
	b	$\chi 1_{\mathrm{GL}(2)} \rtimes \sigma$	$\nu^{3/2}\chi^2\sigma 1_{F^\times} + \nu^{3/2}\sigma 1_{F^\times} + \nu\chi\sigma 1_{F^\times}$
III	a	$\chi \rtimes \sigma \mathrm{St}_{\mathrm{GSp}(2)}$	$\chi\nu^2\sigma 1_{F^\times} + \nu^2\sigma 1_{F^\times}$
	b	$\chi \rtimes \sigma 1_{\mathrm{GSp}(2)}$	—
IV	a	$\sigma \mathrm{St}_{\mathrm{GSp}(4)}$	$\nu^3\sigma 1_{F^\times}$
	b	$L(\nu^2, \nu^{-1}\sigma \mathrm{St}_{\mathrm{GSp}(2)})$	$\nu^3\sigma 1_{F^\times} + \nu\sigma 1_{F^\times}$
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma)$	—
	d	$\sigma 1_{\mathrm{GSp}(4)}$	—
V	a	$\delta([\xi, \nu\xi], \nu^{-1/2}\sigma)$	$\nu^2\sigma 1_{F^\times} + \xi\nu^2\sigma 1_{F^\times}$
	b	$L(\nu^{1/2}\xi \mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	$\nu\sigma 1_{F^\times} + \xi\nu^2\sigma 1_{F^\times}$
	c	$L(\nu^{1/2}\xi \mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\xi\sigma)$	$\xi\nu\sigma 1_{F^\times} + \nu^2\sigma 1_{F^\times}$
	d	$L(\nu\xi, \xi \rtimes \nu^{-1/2}\sigma)$	$\xi\nu\sigma 1_{F^\times} + \nu\sigma 1_{F^\times}$
VI	a	$\tau(S, \nu^{-1/2}\sigma)$	$2 \cdot (\nu^2\sigma 1_{F^\times})$
	b	$\tau(T, \nu^{-1/2}\sigma)$	$\nu^2\sigma 1_{F^\times}$
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	—
	d	$L(\nu, 1_{F^\times} \rtimes \nu^{-1/2}\sigma)$	—
VII		$\chi \rtimes \pi$	0
VIII	a	$\tau(S, \pi)$	0
	b	$\tau(T, \pi)$	0
IX	a	$\delta(\nu\xi, \nu^{-1/2}\pi(\mu))$	0
	b	$L(\nu\xi, \nu^{-1/2}\pi(\mu))$	0
X		$\pi \rtimes \sigma$	$\nu^{3/2}\omega_\pi\sigma 1_{F^\times} + \nu^{3/2}\sigma 1_{F^\times}$
XI	a	$\delta(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\nu^2\sigma 1_{F^\times}$
	b	$L(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\nu\sigma 1_{F^\times}$
		supercuspidal	0

Table 2. The semisimplifications of Jacquet–Waldspurger modules. It is assumed that L is a field, and that the representation of GSp(4, F) admits a (Λ, β)-Bessel functional. A "—" indicates that no such Bessel functional exists.

4. Asymptotic behavior

We begin this section by developing a simple theory of finite-dimensional $F^\times$-modules, which applies to the Jacquet–Waldspurger modules of the previous section. In Section 4.2 we clarify the notion of "asymptotic function". Using our previous results on Jacquet–Waldspurger modules, as well as a result of Danişman in the nonsplit case (Proposition 4.3.3), we can calculate the asymptotic behavior of all Bessel functions of all representations; see Table 4. Simultaneously, we obtain the precise structure as an $F^\times$-module of the Jacquet–Waldspurger modules; see Table 3.

4.1. *Finite-dimensional $F^\times$-modules.* Recall that $F^\times = \langle \varpi \rangle \times \mathfrak{o}^\times$. We consider representations of $F^\times$ on finite-dimensional complex vector spaces. All such representations are assumed to be continuous.

Let n be a positive integer and U be an n-dimensional complex vector space with basis $e_1, \dots, e_n$. We define an action of $F^\times$ on U as follows:

- $\mathfrak{o}^\times$ acts trivially on all of U.
- ϖ acts by sending e_j to $e_j + e_{j-1}$ for all $j \in \{1, \dots, n\}$, where we understand $e_0 = 0$. In other words, the matrix of ϖ with respect to the basis $e_1, \dots, e_n$ is a Jordan block

$$\begin{bmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}. \tag{27}$$

We denote the equivalence class of the $F^\times$-module thus defined by $[n]$. Note that $[n]$ is canonically defined, even though ϖ is not. Clearly, $[n]$ is an indecomposable $F^\times$-module. If σ is a character of $F^\times$, then $\sigma[n] := \sigma \otimes [n]$ is also indecomposable.

Lemma 4.1.1. *Every finite-dimensional indecomposable $F^\times$-module is of the form $\sigma[n]$ for some character σ of $F^\times$ and positive integer n.*

Proof. Let (φ, U) be an indecomposable $F^\times$-module. We may decompose U over $\mathfrak{o}^\times$, i.e.,

$$U = \bigoplus_{i=1}^{r} U(\sigma_i), \tag{28}$$

where the σ_i are pairwise distinct characters of $F^\times$, and

$$U(\sigma_i) = \{u \in U \,:\, \varphi(x)u = \sigma_i(x)u \text{ for all } x \in \mathfrak{o}^\times\}. \tag{29}$$

Let $f = \varphi(\varpi)$. Since each $U(\sigma_i)$ is f-invariant and U is indecomposable, it follows that $r = 1$, i.e., $U = U(\sigma)$ for some character σ of $\mathfrak{o}^\times$. Indecomposability implies that the Jordan normal form of f consists of only one Jordan block

$$\begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}, \qquad \lambda \in \mathbb{C}^\times, \tag{30}$$

of size n. Extend σ to a character of $F^\times$ by setting $\sigma(\varpi) = \lambda$. Then it is easy to see that $\varphi \cong \sigma[n]$. □

Lemma 4.1.2. *Let U be a finite-dimensional $F^\times$-module. Then*

$$U \cong \bigoplus_{i=1}^{r} \sigma_i[n_i] \tag{31}$$

with characters σ_i of $F^\times$ and positive integers n_i. A decomposition as in (31) *is unique up to permutation of the summands.*

Proof. A decomposition as in (31) exists by Lemma 4.1.1. To prove uniqueness, assume that

$$\bigoplus_{i=1}^{r} \sigma_i[n_i] \cong \bigoplus_{j=1}^{s} \tau_j[m_j]. \tag{32}$$

By considering isotypical components with respect to characters of $\mathfrak{o}^\times$, we may assume that all σ_i and τ_j agree when restricted to $\mathfrak{o}^\times$. After appropriate tensoring we may assume this restriction is trivial. The uniqueness statement then follows from the uniqueness of Jordan normal forms. □

Lemma 4.1.3. *Let σ be a character of $F^\times$, and n a positive integer. Let $m \in \{0, \dots, n\}$.*

(i) *There exists exactly one $F^\times$-invariant submodule U_m of $\sigma[n]$ of dimension m. We have $U_k \subset U_m$ for $k \le m$.*

(ii) *The representation of $F^\times$ on U_m is isomorphic to $\sigma[m]$.*

(iii) *The representation of $F^\times$ on $\sigma[n]/U_m$ is isomorphic to $\sigma[n-m]$.*

Proof. (i) Since the invariant subspaces of $[n]$ and $\sigma[n]$ coincide, we may assume that $\sigma = 1$, so that $\sigma[n] = [n]$. Let $e_1, \dots, e_n$ be a basis of $[n]$ with respect to which ϖ acts via the matrix (27). Let $U_m = \langle e_1, \dots, e_m \rangle$. Then U_m is invariant and isomorphic to $[m]$ as an $F^\times$-module.

Conversely, let $U \subset [n]$ be any nonzero invariant subspace. Then U is also invariant under the endomorphism f with matrix

$$\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}. \tag{33}$$

The effect of f on a column vector u is to shift its entries "up" and fill in a 0 at the bottom. Let m be maximal with the property that there exists a $u \in U$ of the form

$$u = {}^t[u_1, \dots, u_m, 0, \dots, 0] \quad \text{with } u_m \neq 0.$$

The vector $f^{m-1}u$ is a nonzero multiple of e_1, showing that $e_1 \in U$. Considering $f^{m-2}u$, we see that $e_2 \in U$ as well. Continuing, we see that $e_1, \dots, e_m \in U$. The maximality of m implies that $U = U_m$.

(ii) We already saw that the subspace U_m of $[n]$ is isomorphic to $[m]$. Hence the subspace $\sigma \otimes U_m$ of $\sigma[n]$ is isomorphic to $\sigma[m]$.

(iii) Clearly $[n]/U_m$ is isomorphic to $[n-m]$. Hence $\sigma[n]/(\sigma \otimes U_m)$ is isomorphic to $\sigma[n-m]$. □

Let U be a finite-dimensional $F^\times$-module. For a character σ of $F^\times$, let U_σ be the sum of all submodules of U isomorphic to $\sigma[n]$ for some n. We call U_σ the *σ-component* of U. By (31), U is the direct sum of its σ-components. A homomorphism $U \to V$ of finite-dimensional $F^\times$-modules induces a map $U_\sigma \to V_\sigma$ for all σ; this follows from Lemma 4.1.3.

4.2. ***Asymptotic functions.*** Let $\mathcal{L}$ be the vector space of functions $f : F^\times \to \mathbb{C}$ with the following properties:

(i) There exists an open-compact subgroup Γ of $F^\times$ such that $f(u\gamma) = f(u)$ for all $u \in F^\times$ and all $\gamma \in \Gamma$.

(ii) $f(u) = 0$ for $v(u) \ll 0$.

Such f arise if we restrict Bessel functions on GSp(4, F) to the subgroup

$$\{\mathrm{diag}(u, u, 1, 1) : u \in F^\times\} \cong F^\times.$$

Clearly $\mathcal{L}$ contains the Schwartz space $\mathcal{S}(F^\times)$, i.e., the space of locally constant, compactly supported functions $F^\times \to \mathbb{C}$. We may think of the quotient $\mathcal{L}/\mathcal{S}(F^\times)$ as a space of "asymptotic functions", in the sense that the image of some $f \in \mathcal{L}$ in this quotient is determined by the values $f(u)$ for $v(u) \gg 0$.

There is an action $\bar{\pi}$ of $F^\times$ on $\mathcal{L}$ given by translation: $(\bar{\pi}(x)f)(u) = f(ux)$ for $x, u \in F^\times$. This is a smooth action by the properties of the elements of $\mathcal{L}$. The action preserves the subspace $\mathcal{S}(F^\times)$, so that we get an action on the quotient $\mathcal{L}/\mathcal{S}(F^\times)$.

For the proof of the following lemma, we will use the formula

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k P(k)=0 \quad \text{for all } P\in\mathbb{C}[X] \text{ with } \deg(P)<n. \tag{34}$$

This formula follows by differentiating the identity $(1+x)^n=\sum_{k=0}^{n}\binom{n}{k}x^k$ repeatedly and setting $x=-1$.

Lemma 4.2.1. *Let $\beta\in\mathbb{C}^\times$. For a positive integer n, let $\mathcal{F}_n(\beta)$ be the space of functions $f:\mathbb{Z}_{\geq 0}\to\mathbb{C}$ satisfying*

$$\sum_{k=0}^{n}\binom{n}{k}(-\beta)^{n-k} f(m+k)=0 \quad \textit{for all } m\geq 0. \tag{35}$$

Then $\dim\mathcal{F}_n(\beta)=n$, *and a basis of $\mathcal{F}_n(\beta)$ is given by the functions*

$$f_j(m)=m^j\beta^m, \quad m\geq 0, \tag{36}$$

for $j=0,\dots,n-1$.

Proof. It is clear from (35) that any $f\in\mathcal{F}_n(\beta)$ is determined by the values $f(0),\dots,f(n-1)$. Hence $\dim\mathcal{F}_n(\beta)\leq n$, and we only need to show that the functions f_j lie in $\mathcal{F}_n(\beta)$ and are linearly independent. The fact that the functions f_j lie in $\mathcal{F}_n(\beta)$ follows from (34). It is easy to prove that they are linearly independent. □

Proposition 4.2.2. *Let $\mathcal{K}$ be an $F^\times$-invariant subspace of $\mathcal{L}$ which contains $\mathcal{S}(F^\times)$ with finite codimension n. Assume that, as an $F^\times$-module, the quotient $\mathcal{K}/\mathcal{S}(F^\times)$ is isomorphic to $\sigma[n]$ for some character σ of $F^\times$. Then there exist $f_0,\dots,f_{n-1}\in\mathcal{K}$ with the following properties*:

(i) *The images of $f_0,\dots,f_{n-1}$ in $\mathcal{K}/\mathcal{S}(F^\times)$ are a basis of the quotient space.*

(ii) *f_j has asymptotic behavior*

$$f_j(x)=v(x)^j\sigma(x) \quad \textit{for all } x\in F^\times \textit{ with } v(x)\gg 0, \tag{37}$$

for all $j\in\{0,\dots,n-1\}$.

Proof. It suffices to show that every $f\in\mathcal{K}$ has the asymptotic form

$$f(x)=\sum_{k=0}^{n-1} c_k\, v(x)^k\sigma(x) \quad \text{for all } x\in F^\times \text{ with } v(x)\gg 0 \tag{38}$$

for some constants c_k. We have $\sigma[n](u)=\sigma(u)\mathrm{id}$ for $u\in\mathfrak{o}^\times$ on all of $\sigma[n]$. Hence, for a fixed unit $u\in\mathfrak{o}^\times$,

$$\bar\pi(u)f-\sigma(u)f\in\mathcal{S}(F^\times). \tag{39}$$

It follows that there exists a $j_0 \geq 0$ such that

$$f(u\varpi^{m+j_0}) = \sigma(u)f(\varpi^{m+j_0}) \quad \text{for all } m \geq 0. \tag{40}$$

Since $\mathfrak{o}^\times$ is compact and both sides of (40) are locally constant, we may choose j_0 large enough so that (40) holds for all $u \in \mathfrak{o}^\times$.

Every vector in $\sigma[n]$ is annihilated by $(\sigma[n](\varpi) - \lambda \,\mathrm{id})^n$, where we abbreviate $\lambda = \sigma(\varpi)$. Hence

$$(\bar{\pi}(\varpi) - \lambda \,\mathrm{id})^n f \in \mathcal{S}(F^\times) \tag{41}$$

for all $f \in \mathcal{K}$, or

$$\sum_{k=0}^{n} \binom{n}{k} (-\lambda)^{n-k} \bar{\pi}(\varpi^k) f \in \mathcal{S}(F^\times). \tag{42}$$

It follows that there exists a $j_0 \geq 0$ such that

$$\sum_{k=0}^{n} \binom{n}{k} (-\lambda)^{n-k} f(\varpi^{m+k+j_0}) = 0 \quad \text{for all } m \geq 0. \tag{43}$$

We may assume that the same j_0 works for both (40) and (43). Setting $h(m) := f(\varpi^{m+j_0})$, (43) reads

$$\sum_{k=0}^{n} \binom{n}{k} (-\lambda)^{n-k} h(m+k) = 0 \quad \text{for all } m \geq 0. \tag{44}$$

By Lemma 4.2.1, there exist constants $d_0, \dots, d_{n-1}$ such that

$$h(m) = \sum_{k=0}^{n-1} d_k m^k \lambda^m \quad \text{for all } m \geq 0. \tag{45}$$

We can then also find constants $c_0, \dots, c_{n-1}$ such that

$$h(m) = \sum_{k=0}^{n-1} c_k (m+j_0)^k \lambda^{m+j_0} \quad \text{for all } m \geq 0. \tag{46}$$

(To get the c_k's from the d_k's, expand $m^k = ((m+j_0) - j_0)^k$ in (45).) For $x \in F^\times$ with $v(x) \geq j_0$, write $x = u\varpi^j$ with $u \in \mathfrak{o}^\times$ and $j \geq j_0$. Then

$$\begin{aligned} f(x) &= \sigma(u) f(\varpi^j) && \text{by (40)} \\ &= \sigma(u) \sum_{k=0}^{n-1} c_k j^k \lambda^j && \text{by (46)} \\ &= \sum_{k=0}^{n-1} c_k \, v(x)^k \sigma(x). \end{aligned}$$

□

Corollary 4.2.3. *Let U be a finite-dimensional submodule of $\mathcal{L}/\mathcal{S}(F^\times)$. Then each σ-component of U is indecomposable.*

Proof. Let $\mathcal{K}$ be the preimage of U under the projection $\mathcal{L} \to \mathcal{L}/\mathcal{S}(F^\times)$. Assume that there exists a σ for which U_σ is decomposable. Then U_σ contains a direct sum $\sigma[n] \oplus \sigma[n']$ with $n, n' > 0$. By Proposition 4.2.2, there exist two functions $f, f' \in \mathcal{K}$ such that the image of f in

$$U = \mathcal{K}/\mathcal{S}(F^\times)$$

lies in $\sigma[n]$, the image of f' lies in $\sigma[n']$, and such that

$$f(x) = \sigma(x) \text{ and } f'(x) = \sigma(x) \quad \text{for all } x \in F^\times \text{ with } v(x) \gg 0. \tag{47}$$

It follows from (47) that f and f' have the same image in $\mathcal{K}/\mathcal{S}(F^\times)$, which is a contradiction. $\square$

4.3. ***Asymptotic behavior of Bessel functions.*** Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. Assume that V is the (Λ, β)-Bessel model of π with respect to a character Λ of T. We associate with each Bessel function $B \in V$ the function $\varphi_B : F^\times \to \mathbb{C}$ defined by

$$\varphi_B(u) = B(\mathrm{diag}(u, u, 1, 1)).$$

Let $\mathcal{K}$ be the space spanned by all functions φ_B.

Lemma 4.3.1. *$\mathcal{K}$ contains $\mathcal{S}(F^\times)$.*

Proof. This follows by the same arguments as in Lemma 4.1 of [Danişman 2014]. $\square$

An easy argument as in Proposition 4.7.2 of [Bump 1997], or as in Proposition 3.1 of [Danişman 2014], shows that if $B \in V(N)$, then φ_B has compact support. It is also true, and equally easy to see, that

$$B \in V(N, T, \Lambda) \quad \Longrightarrow \quad \varphi_B \text{ has compact support in } F^\times.$$

It follows that the linear map $B \mapsto \varphi_B$ induces a surjection

$$V_{N,T,\Lambda} \to \mathcal{K}/\mathcal{S}(F^\times). \tag{48}$$

Lemma 4.3.2. *Assume that the map* (48) *is an isomorphism. Then every σ-component of $V_{N,T,\Lambda}$ is indecomposable as an $F^\times$-module.*

Proof. The map (48) induces an isomorphism of the respective σ-components. Hence the assertion follows from Corollary 4.2.3. $\square$

Proposition 4.3.3. *Suppose we are in the nonsplit case. Then the map* (48) *is an isomorphism.*

Proof. See Theorem 4.9 of [Danişman 2014]. $\square$

Recall that in Table 2 we determined the semisimplifications of the Jacquet–Waldspurger modules for all irreducible, admissible representations. In the nonsplit case, we can now determine the precise algebraic structure of these modules.

Corollary 4.3.4. *The algebraic structure of the Jacquet–Waldspurger modules* $V_{N,T,\Lambda}$ *for all irreducible, admissible representations of* $\mathrm{GSp}(4, F)$ *is given in Table 3, under the assumption that the representation* (π, V) *admits a nonsplit* (Λ, β)*-Bessel functional.* (A "—" indicates that no such Bessel functional exists.)

Proof. By Proposition 4.3.3 and Lemma 4.3.2, every σ-component of $V_{N,T,\Lambda}$ is indecomposable. This information, together with the semisimplifications from Table 2, gives the precise structure. □

For type I, we have to distinguish various cases, depending on the regularity of the inducing character:

$$V_{N,T,\Lambda} = \begin{cases} \nu^{3/2}\chi_1\chi_2\sigma \oplus \nu^{3/2}\chi_1\sigma \oplus \nu^{3/2}\chi_2\sigma \oplus \nu^{3/2}\sigma & \text{if } \chi_1\chi_2, \chi_1, \chi_2, 1 \\ & \qquad \text{are pairwise different,} \\ \nu^{3/2}\chi^2\sigma \oplus (\nu^{3/2}\chi\sigma)[2] \oplus \nu^{3/2}\sigma & \text{if } \chi := \chi_1 = \chi_2 \neq 1,\ \chi^2 \neq 1, \\ (\nu^{3/2}\chi\sigma)[2] \oplus (\nu^{3/2}\sigma)[2] & \text{if } \chi := \chi_1 = \chi_2 \neq 1,\ \chi^2 = 1, \\ (\nu^{3/2}\chi\sigma)[2] \oplus (\nu^{3/2}\sigma)[2] & \text{if } \{\chi_1, \chi_2\} = \{\chi \neq 1, 1\} \\ (\nu^{3/2}\sigma)[4] & \text{if } \chi_1 = \chi_2 = 1. \end{cases} \tag{49}$$

Corollary 4.3.5. *Table 4 shows the asymptotic behavior of the functions*

$$B(\mathrm{diag}(u, u, 1, 1))$$

for all irreducible, admissible representations (π, V) *of* $\mathrm{GSp}(4, F)$, *where* B *runs through a nonsplit* (Λ, β)*-Bessel model of* π. (A "—" indicates that no such Bessel model exists.)

Proof. By Proposition 4.3.3, the map (48) is an isomorphism. We can thus use Proposition 4.2.2, which translates the algebraic structure of $V_{N,T,\Lambda}$ given in Table 3 into the asymptotic behavior of Bessel functions. □

Remark. This result is to be understood in the sense that all the constants given in Table 4 are necessary, i.e., for any choice of $C_1, C_2, \dots$ there exists a Bessel function B such that $B(\mathrm{diag}(u, u, 1, 1))$ has the asymptotic behavior given by this choice of constants.

		representation		$V_{N,T,\Lambda}$
I		$\chi_1 \times \chi_2 \rtimes \sigma$		see (49)
II	a	$\chi \mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	$\chi^2 \neq 1$	$\nu^2\chi\sigma \oplus \nu^{3/2}\chi^2\sigma \oplus \nu^{3/2}\sigma$
			$\chi^2 = 1$	$\nu^2\chi\sigma \oplus (\nu^{3/2}\sigma)[2]$
	b	$\chi 1_{\mathrm{GL}(2)} \rtimes \sigma$	$\chi^2 \neq 1$	$\nu\chi\sigma \oplus \nu^{3/2}\chi^2\sigma \oplus \nu^{3/2}\sigma$
			$\chi^2 = 1$	$\nu\chi\sigma \oplus (\nu^{3/2}\sigma)[2]$
III	a	$\chi \rtimes \sigma \mathrm{St}_{\mathrm{GSp}(2)}$		$\chi\nu^2\sigma \oplus \nu^2\sigma$
	b	$\chi \rtimes \sigma 1_{\mathrm{GSp}(2)}$		—
IV	a	$\sigma \mathrm{St}_{\mathrm{GSp}(4)}$		$\nu^3\sigma$
	b	$L(\nu^2, \nu^{-1}\sigma \mathrm{St}_{\mathrm{GSp}(2)})$		$\nu^3\sigma \oplus \nu\sigma$
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma)$		—
	d	$\sigma 1_{\mathrm{GSp}(4)}$		—
V	a	$\delta([\xi, \nu\xi], \nu^{-1/2}\sigma)$		$\nu^2\sigma \oplus \xi\nu^2\sigma$
	b	$L(\nu^{1/2}\xi \mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$		$\nu\sigma \oplus \xi\nu^2\sigma$
	c	$L(\nu^{1/2}\xi \mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\xi\sigma)$		$\xi\nu\sigma \oplus \nu^2\sigma$
	d	$L(\nu\xi, \xi \rtimes \nu^{-1/2}\sigma)$		$\xi\nu\sigma \oplus \nu\sigma$
VI	a	$\tau(S, \nu^{-1/2}\sigma)$		$(\nu^2\sigma)[2]$
	b	$\tau(T, \nu^{-1/2}\sigma)$		$\nu^2\sigma$
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$		—
	d	$L(\nu, 1_{F^\times} \rtimes \nu^{-1/2}\sigma)$		—
VII		$\chi \rtimes \pi$		0
VIII	a	$\tau(S, \pi)$		0
	b	$\tau(T, \pi)$		0
IX	a	$\delta(\nu\xi, \nu^{-1/2}\pi(\mu))$		0
	b	$L(\nu\xi, \nu^{-1/2}\pi(\mu))$		0
X		$\pi \rtimes \sigma$	$\omega_\pi \neq 1$	$\nu^{3/2}\omega_\pi\sigma \oplus \nu^{3/2}\sigma$
			$\omega_\pi = 1$	$(\nu^{3/2}\sigma)[2]$
XI	a	$\delta(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$		$\nu^2\sigma$
	b	$L(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$		$\nu\sigma$
		supercuspidal		0

Table 3. Jacquet–Waldspurger modules $V_{N,T,\Lambda}$. It is assumed that L is a field, and that the representation of $\mathrm{GSp}(4, F)$ admits a (Λ, ψ_β)-Bessel functional. A "—" indicates that no nonsplit Bessel functional exists.

		representation		$\lvert u\rvert^{-3/2}\,B(\mathrm{diag}(u,u,1,1))$
I		$\chi_1\times\chi_2\rtimes\sigma$		see (50)
II	a	$\chi\mathrm{St}_{\mathrm{GL}(2)}\rtimes\sigma$	$\chi^2\neq 1$	$C_1(\nu^{1/2}\chi\sigma)(u)+C_2(\chi^2\sigma)(u)+C_3\sigma(u)$
			$\chi^2=1$	$C_1(\nu^{1/2}\chi\sigma)(u)+(C_2+C_3v(u))\sigma(u)$
	b	$\chi 1_{\mathrm{GL}(2)}\rtimes\sigma$	$\chi^2\neq 1$	$C_1(\nu^{-1/2}\chi\sigma)(u)+C_2(\chi^2\sigma)(u)+C_3\sigma(u)$
			$\chi^2=1$	$C_1(\nu^{-1/2}\chi\sigma)(u)+(C_2+C_3v(u))\sigma(u)$
III	a	$\chi\rtimes\sigma\mathrm{St}_{\mathrm{GSp}(2)}$		$C_1(\nu^{1/2}\chi\sigma)(u)+C_2(\nu^{1/2}\sigma)(u)$
	b	$\chi\rtimes\sigma 1_{\mathrm{GSp}(2)}$		—
IV	a	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$		$C(\nu^{3/2}\sigma)(u)$
	b	$L(\nu^2,\nu^{-1}\sigma\mathrm{St}_{\mathrm{GSp}(2)})$		$C_1(\nu^{3/2}\sigma)(u)+C_2(\nu^{-1/2}\sigma)(u)$
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-3/2}\sigma)$		—
	d	$\sigma 1_{\mathrm{GSp}(4)}$		—
V	a	$\delta([\xi,\nu\xi],\nu^{-1/2}\sigma)$		$C_1(\nu^{1/2}\xi\sigma)(u)+C_2(\nu^{1/2}\sigma)(u)$
	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$		$C_1(\nu^{1/2}\xi\sigma)(u)+C_2(\nu^{-1/2}\sigma)(u)$
	c	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\xi\sigma)$		$C_1(\nu^{-1/2}\xi\sigma)(u)+C_2(\nu^{1/2}\sigma)(u)$
	d	$L(\nu\xi,\xi\rtimes\nu^{-1/2}\sigma)$		$C_1(\nu^{-1/2}\xi\sigma)(u)+C_2(\nu^{-1/2}\sigma)(u)$
VI	a	$\tau(S,\nu^{-1/2}\sigma)$		$(C_1+C_2v(u))(\nu^{1/2}\sigma)(u)$
	b	$\tau(T,\nu^{-1/2}\sigma)$		$C(\nu^{1/2}\sigma)(u)$
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$		—
	d	$L(\nu,1_{F^\times}\rtimes\nu^{-1/2}\sigma)$		—
VII		$\chi\rtimes\pi$		0
VIII	a	$\tau(S,\pi)$		0
	b	$\tau(T,\pi)$		0
IX	a	$\delta(\nu\xi,\nu^{-1/2}\pi(\mu))$		0
	b	$L(\nu\xi,\nu^{-1/2}\pi(\mu))$		0
X		$\pi\rtimes\sigma$	$\omega_\pi\neq 1$	$C_1(\omega_\pi\sigma)(u)+C_2\sigma(u)$
			$\omega_\pi=1$	$(C_1+C_2v(u))\sigma(u)$
XI	a	$\delta(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$		$C(\nu^{1/2}\sigma)(u)$
	b	$L(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$		$C(\nu^{-1/2}\sigma)(u)$
		supercuspidal		0

Table 4. Asymptotic behavior of $B(\mathrm{diag}(u,u,1,1))$ in the nonsplit case. A "—" indicates that no nonsplit Bessel functional exists.

Again, for type I we have to distinguish various cases:

(50) $|u|^{-3/2}\, B(\mathrm{diag}(u, u, 1, 1))$

$$= \begin{cases} C_1(\chi_1\chi_2\sigma)(u) & \text{if } \chi_1\chi_2,\ \chi_1,\ \chi_2,\ 1 \\ \qquad +C_2(\chi_1\sigma)(u)+C_3(\chi_2\sigma)(u)+C_4\sigma(u) & \qquad \text{are pairwise different,} \\ C_1(\chi^2\sigma)(u) & \text{if } \chi := \chi_1 = \chi_2 \neq 1,\ \chi^2 \neq 1, \\ \qquad +(C_2+C_3v(u))(\chi\sigma)(u)+C_4\sigma(u) & \\ (C_1+C_2v(u))(\chi\sigma)(u)+(C_3+C_4v(u))\sigma(u) & \text{if } \chi := \chi_1 = \chi_2 \neq 1,\ \chi^2 = 1, \\ (C_1+C_2v(u))(\chi\sigma)(u)+(C_3+C_4v(u))\sigma(u) & \text{if } \{\chi_1, \chi_2\} = \{\chi \neq 1, 1\}, \\ (C_1+C_2v(u)+C_3v^2(u)+C_4v^3(u))\sigma(u) & \text{if } \chi_1 = \chi_2 = 1. \end{cases}$$

Remark 4.3.6. The proof of Proposition 4.3.3 given in [Danişman 2014] is based on the exactness of the Waldspurger functor, which is only true in the nonsplit case. Assume that (π, V) is an irreducible, admissible representation of $\mathrm{GSp}(4, F)$ which admits a *split* Bessel model $\mathcal{B}(\pi, \Lambda, \beta)$. Then we still have the surjection (48), which implies that the space of asymptotic functions $\mathcal{K}/\mathcal{S}(F^\times)$, as an $F^\times$-module, is a quotient of the Jacquet–Waldspurger module $V_{N,T,\Lambda}$. Starting from the $V_{N,T}$ given in Table 1, the $V_{N,T,\Lambda}$ can be calculated in many cases, but some of them pose difficulties, again due to the fact that the Waldspurger functor in the split case is not exact. Thus, complete results in the split case would follow from the solution of the following two problems:

- Calculate the Jacquet–Waldspurger modules $V_{N,T,\Lambda}$ in all cases.
- Control the kernel of the map (48).

The current methods still allow for some preliminary results on the asymptotic behavior of the functions $B(\mathrm{diag}(u, u, 1, 1))$ in the split case. More precisely, it is not difficult to create a table similar to Table 4, but it is unclear if all the constants C_i in such a table are really necessary. What is clear is that every $B(\mathrm{diag}(u, u, 1, 1))$ is of the general form

$$B(\mathrm{diag}(u, u, 1, 1)) = \sum_{i=1}^{n} C_i v(u)^{k_i} \sigma_i(u) \quad \text{for } v(u) \gg 0 \tag{51}$$

with k_i nonnegative integers, σ_i characters of $F^\times$, and $C_i \in \mathbb{C}$.

5. Local zeta integrals and L-factors

Given an irreducible, admissible, unitary representation π of $\mathrm{GSp}(4, F)$ and a character μ of $F^\times$, a certain type of zeta integral was introduced in Section 3 of [Piatetski-Shapiro 1997] and used to define an L-factor $L^{\mathrm{PS}}(s, \pi, \mu)$. These zeta integrals depend on a choice of Bessel model for π, and hence the L-factor may also depend on this choice. In many cases, though, one can prove that $L^{\mathrm{PS}}(s,\pi,\mu)$ is independent of the choice of Bessel data.

In Section 5.1 we introduce a simplified type of zeta integral and use it to define the *regular part* $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ of the Piatetski-Shapiro L-factor. The simplified zeta integrals also depend on the choice of a Bessel model for π. Using the asymptotic behavior given in Table 4, we explicitly calculate $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ in the nonsplit case for all representations. It turns out that $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ is independent of the choice of Bessel model, and coincides with the usual degree-4 (spin) Euler factor if π is generic. For nongeneric representations, however, the two factors do not agree in general.

We then investigate the Piatetski-Shapiro zeta integrals (78). Their definition involves a certain subgroup G of $\mathrm{GSp}(4,F)$, to which we dedicate Section 5.2. The resulting L-factor $L^{\mathrm{PS}}(s,\pi,\mu)$ is either equal to $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$, or has an additional factor $L(s+1/2,\Lambda_\mu)$, where $\Lambda_\mu=\Lambda\cdot(\mu\circ N_{L/F})$ depends on the Bessel data. In Section 5.5 we will identify several cases where $L^{\mathrm{PS}}(s,\pi,\mu)=L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$.

Overall in this section we closely follow [Piatetski-Shapiro 1997]. However, we treat all representations, not only unitary ones. Our notion of exceptional pole is slightly more general than the one given in [Piatetski-Shapiro 1997]. Also, we fill in some of the missing proofs of that paper.

5.1. ***The simplified zeta integrals.*** Let π be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$. Let $\mathcal{B}(\pi,\Lambda,\beta)$ be a (Λ,β)-Bessel model for π. Let μ be a character of $F^\times$. For $B\in\mathcal{B}(\pi,\Lambda,\beta)$ and $s\in\mathbb{C}$, we define the *simplified zeta integrals*

$$\zeta(s,B,\mu)=\int\limits_{F^\times} B\left(\begin{bmatrix} x & \\ & 1\end{bmatrix}\right)\mu(x)|x|^{s-3/2}\,d^\times x. \tag{52}$$

The same integrals appear in Proposition 18 of [Danişman 2015b]. Using the general form (51) of the functions $B\left(\left[\begin{smallmatrix} x & \\ & 1\end{smallmatrix}\right]\right)$, which holds both in the split and the nonsplit case, it is easy to see that $\zeta(s,B,\mu)$ converges to an element of $\mathbb{C}(q^{-s})$ for real part of s large enough. Let $I(\pi,\mu)$ be the $\mathbb{C}$-vector subspace of $\mathbb{C}(q^{-s})$ spanned by all $\zeta(s,B,\mu)$ as B runs through $\mathcal{B}(\pi,\Lambda,\beta)$.

Proposition 5.1.1. *Let π be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$ *admitting a (Λ,β)-Bessel model with β as in* (4)*. Then $I(\pi,\mu)$ is a nonzero $\mathbb{C}[q^{-s},q^s]$ module containing $\mathbb{C}$, and there exists $R(X)\in\mathbb{C}[X]$ such that*

$$R(q^{-s})I(\pi,\mu)\subset\mathbb{C}[q^{-s},q^s],$$

so that $I(\pi,\mu)$ is a fractional ideal of the principal ideal domain $\mathbb{C}[q^{-s},q^s]$ whose quotient field is $\mathbb{C}(q^{-s})$. The fractional ideal $I(\pi,\mu)$ admits a generator of the form $1/Q(q^{-s})$ with $Q(0)=1$, where $Q(X)\in\mathbb{C}[X]$.

Proof. One can argue as in the proof of Proposition 2.6.4 of [Roberts and Schmidt 2007]. One step in the proof is to show that $I(\pi,\mu)$ contains $\mathbb{C}$. This follows from Lemma 4.3.1. $\square$

Using the notation of this proposition, we set

$$L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu) := 1/Q(q^{-s}) \tag{53}$$

and call this the *regular part of the Piatetski-Shapiro L-factor*; see [Piatetski-Shapiro 1997]. As the notation indicates, $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ does not depend on the Bessel data β and Λ. This is implied by the following result.

Theorem 5.1.2. *Table 5 shows the factors* $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ *for all irreducible, admissible representations* (π,V) *of* $\mathrm{GSp}(4,F)$ *in the nonsplit case.* (A "—" indicates that no nonsplit Bessel functional exists.)

Proof. Up to an element of $\mathcal{S}(F^\times)$, the functions $x\mapsto B\big(\big[\begin{smallmatrix}x&\\&1\end{smallmatrix}\big]\big)$, where $B\in\mathcal{B}(\pi,\Lambda,\beta)$, are listed in Table 4. Using the fact that

$$\sum_{m=m_0}^{\infty} m^j z^m = g(z)\frac{1}{(1-z)^{j+1}} \tag{54}$$

with a function $g(z)$ which is holomorphic and nonvanishing at $z=1$, the integrals in (52) are thus easily calculated up to elements of $\mathbb{C}[q^s,q^{-s}]$. $\square$

Also indicated in Table 5 are the generic representations (i.e., those that admit a Whittaker model); supercuspidals may or may not be generic. We see that for all generic representations $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)=L(s,\varphi)$ if $\mu=1_{F^\times}$. Here $L(s,\varphi)$ is the L-factor of the Langlands parameter φ of π, as listed in Table A.8 of [Roberts and Schmidt 2007].

5.2. ***The group G.*** We now recall the setup of [Piatetski-Shapiro 1997]. Let L be the quadratic extension of F as in Section 2. Let $V=L^2$, which we consider as a space of row vectors. We endow V with the skew-symmetric F-linear form

$$\rho(x,y)=\mathrm{Tr}_{L/F}(x_1y_2-x_2y_1),\quad x=(x_1,x_2),\ y=(y_1,y_2). \tag{55}$$

Let

$$\mathrm{GSp}_\rho=\big\{g\in\mathrm{GL}(4,F):\ \rho(xg,yg)=\lambda\rho(x,y),\ \text{for some } \lambda=\lambda(g)\in F^\times,\ \text{for all } x,y\in V\big\}$$

be the symplectic similitude group of the form ρ. Let

$$G=\{g\in\mathrm{GL}(2,L):\ \det(g)\in F^\times\}. \tag{56}$$

The group G acts on V by matrix multiplication from the right. A calculation shows

$$\rho(xg,yg)=\det(g)\rho(x,y) \tag{57}$$

		representation	$L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$	generic
I		$\chi_1\times\chi_2\rtimes\sigma$ (irreducible)	$L(s,\chi_1\chi_2\sigma\mu)L(s,\sigma\mu)L(s,\chi_1\sigma\mu)L(s,\chi_2\sigma\mu)$	•
II	a	$\chi\mathrm{St}_{\mathrm{GL}(2)}\rtimes\sigma$	$L(s,\nu^{1/2}\chi\sigma\mu)L(s,\chi^2\sigma\mu)L(s,\sigma\mu)$	•
	b	$\chi 1_{\mathrm{GL}(2)}\rtimes\sigma$	$L(s,\nu^{-1/2}\chi\sigma\mu)L(s,\chi^2\sigma\mu)L(s,\sigma\mu)$	
III	a	$\chi\rtimes\sigma\mathrm{St}_{\mathrm{GSp}(2)}$	$L(s,\nu^{1/2}\chi\sigma\mu)L(s,\nu^{1/2}\sigma\mu)$	•
	b	$\chi\rtimes\sigma 1_{\mathrm{GSp}(2)}$	—	
IV	a	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$	$L(s,\nu^{3/2}\sigma\mu)$	•
	b	$L(\nu^2,\nu^{-1}\sigma\mathrm{St}_{\mathrm{GSp}(2)})$	$L(s,\nu^{3/2}\sigma\mu)L(s,\nu^{-1/2}\sigma\mu)$	
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-3/2}\sigma)$	—	
	d	$\sigma 1_{\mathrm{GSp}(4)}$	—	
V	a	$\delta([\xi,\nu\xi],\nu^{-1/2}\sigma)$	$L(s,\nu^{1/2}\xi\sigma\mu)L(s,\nu^{1/2}\sigma\mu)$	•
	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$	$L(s,\nu^{1/2}\xi\sigma\mu)L(s,\nu^{-1/2}\sigma\mu)$	
	c	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\xi\sigma)$	$L(s,\nu^{-1/2}\xi\sigma\mu)L(s,\nu^{1/2}\sigma\mu)$	
	d	$L(\nu\xi,\xi\rtimes\nu^{-1/2}\sigma)$	$L(s,\nu^{-1/2}\xi\sigma\mu)L(s,\nu^{-1/2}\sigma\mu)$	
VI	a	$\tau(S,\nu^{-1/2}\sigma)$	$L(s,\nu^{1/2}\sigma\mu)^2$	•
	b	$\tau(T,\nu^{-1/2}\sigma)$	$L(s,\nu^{1/2}\sigma\mu)$	
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)},\nu^{-1/2}\sigma)$	—	
	d	$L(\nu,1_{F^\times}\rtimes\nu^{-1/2}\sigma)$	—	
VII		$\chi\rtimes\pi$	1	•
VIII	a	$\tau(S,\pi)$	1	•
	b	$\tau(T,\pi)$	1	
IX	a	$\delta(\nu\xi,\nu^{-1/2}\pi(\mu))$	1	•
	b	$L(\nu\xi,\nu^{-1/2}\pi(\mu))$	1	
X		$\pi\rtimes\sigma$	$L(s,\omega_\pi\sigma\mu)L(s,\sigma\mu)$	•
XI	a	$\delta(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$	$L(s,\nu^{1/2}\sigma\mu)$	•
	b	$L(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$	$L(s,\nu^{-1/2}\sigma\mu)$	
		supercuspidal	1	◦

Table 5. Regular parts of Piatetski-Shapiro L-factors (nonsplit case).

for $x, y \in V$ and $g \in G$. Hence, $G \subset \mathrm{GSp}_\rho$. Since all four-dimensional symplectic F-spaces are isomorphic to the standard space F^4 with the form (1), the groups GSp_ρ and $\mathrm{GSp}(4, F)$ are isomorphic; here, we think of $\mathrm{GSp}(4, F)$ as acting on the right on the space of row vectors F^4. We wish to find one such isomorphism under which the group G takes on a particularly simple shape inside $\mathrm{GSp}(4, F)$.

For this we assume that the matrix β in (4) is diagonal and nondegenerate, i.e., $b = 0$ and $a, c \neq 0$; after a suitable conjugation, every nondegenerate β can be brought into this form. Consider the following F-basis of V,

$$f_1 = (1, 0), \quad f_2 = (\Delta/c, 0), \quad f_3 = (0, 1/2), \quad f_4 = (0, c/(2\Delta)). \tag{58}$$

Let $e_1, \dots, e_4$ be the standard basis of F^4. Then the map $f_i \mapsto e_i$ establishes an isomorphism $V \cong F^4$ preserving the symplectic form on both spaces (the form ρ on V, and the form J defined in (1) on F^4). The resulting isomorphism $\mathrm{GSp}_\rho \cong \mathrm{GSp}(4, F)$ has the following properties:

$$G \ni \begin{bmatrix} x & \\ & 1 \end{bmatrix} \longmapsto \begin{bmatrix} x &&& \\ & x && \\ && 1 & \\ &&& 1 \end{bmatrix}, \tag{59}$$

$$G \ni \begin{bmatrix} 1 & \\ & x \end{bmatrix} \longmapsto \begin{bmatrix} 1 &&& \\ & 1 && \\ && x & \\ &&& x \end{bmatrix}, \tag{60}$$

$$G \ni \begin{bmatrix} t & \\ & \bar{t} \end{bmatrix} \longmapsto \begin{bmatrix} x & yc && \\ -ya & x && \\ && x & ya \\ && -yc & x \end{bmatrix} \quad \text{for } t = x + y\Delta \in L^\times, \tag{61}$$

$$G \ni \begin{bmatrix} 1 & x + y\Delta \\ & 1 \end{bmatrix} \longmapsto \begin{bmatrix} 1 && 2x & -2ay \\ & 1 & -2ay & -2ac^{-1}x \\ && 1 & \\ &&& 1 \end{bmatrix}. \tag{62}$$

Here, $\bar{t} = x - y\Delta$ is the Galois conjugate of t. Recall from (9) that the matrices on the right hand side of (61) are precisely the elements of T. It is easy to verify that the matrices on the right-hand side of (62) are precisely those elements of N that lie in

$$N_0 = \left\{ \begin{bmatrix} 1 & X \\ & 1 \end{bmatrix} : \operatorname{tr}(\beta X) = 0 \right\} = \left\{ \begin{bmatrix} 1 && x & y \\ & 1 & y & z \\ && 1 & \\ &&& 1 \end{bmatrix} : ax + by + cz = 0 \right\}. \tag{63}$$

In particular, if we consider G a subgroup of $\mathrm{GSp}(4, F)$, then we see that

$$G \cap R = T N_0;$$

see Proposition 2.1 of [Piatetski-Shapiro 1997]. We define the following subgroups of G:

$$A^G = G \cap \begin{bmatrix} * & \\ & * \end{bmatrix} = \left\{ \begin{bmatrix} xt & \\ & \bar{t} \end{bmatrix} \in \mathrm{GL}(2, L) \,:\, x \in F^\times,\ t \in L^\times \right\}, \tag{64}$$

$$N_0 = G \cap \begin{bmatrix} 1 & * \\ & 1 \end{bmatrix} = \left\{ \begin{bmatrix} 1 & b \\ & 1 \end{bmatrix} \in \mathrm{GL}(2, L) \,:\, b \in L \right\}, \tag{65}$$

$$B^G = G \cap \begin{bmatrix} * & * \\ & * \end{bmatrix} = \left\{ \begin{bmatrix} a & b \\ & d \end{bmatrix} \in \mathrm{GL}(2, L) \,:\, ad \in F^\times \right\}, \tag{66}$$

$$K^G = G \cap \mathrm{GL}(2, \mathfrak{o}_L) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2, \mathfrak{o}_L) \,:\, ad - bc \in F^\times \right\}. \tag{67}$$

By our remarks above, when embedded into $\mathrm{GSp}(4, F)$, the group N_0 coincides with the group introduced in (63), so that the notation is consistent. The Iwasawa decomposition for $\mathrm{GL}(2, L)$ implies that $G = B^G K^G$. The modular factor for B^G is $\delta\big(\big[\begin{smallmatrix} a & b \\ & d \end{smallmatrix}\big]\big) = |a/d|_L$, where $|\cdot|_L$ is the normalized absolute value on L. Note that $|t|_L = |N_{L/F}(t)|_F$ for $t \in L^\times$. Let dn be the Haar measure on N_0 that gives $N_0 \cap K^G$ volume 1. Let da be the Haar measure on A^G that gives $A^G \cap K^G$ volume 1. Let dk be the Haar measure on K^G with total volume 1. There is a Haar measure on G given by

$$\int\limits_{N_0} \int\limits_{A^G} \int\limits_{K^G} f(nak)\delta(a)^{-1}\, dk\, da\, dn. \tag{68}$$

The measure (68) gives K^G volume 1. We will also use the integration formula

$$\int\limits_{N_0\backslash G} f(g)\, dg = \int\limits_{B^G} f(wb)\, db = \int\limits_{N_0} \int\limits_{A^G} f(wna)\, da\, dn \tag{69}$$

for a function f on G that is left N_0-invariant (the db in the middle integral is a *right* Haar measure on B^G). Here, $w = \big[\begin{smallmatrix} & 1 \\ -1 & \end{smallmatrix}\big] \in G$, which is embedded into $\mathrm{GSp}(4, F)$ as

$$w \longmapsto \begin{bmatrix} 2 & & & \\ & -2ac^{-1} & & \\ & & \frac{1}{2} & \\ & & & -\frac{1}{2}ca^{-1} \end{bmatrix} \begin{bmatrix} & & 1 & \\ & & & 1 \\ -1 & & & \\ & -1 & & \end{bmatrix}. \tag{70}$$

Principal series representations of G. Let Λ be a character of $L^\times$, let μ be a character of $F^\times$, and $s \in \mathbb{C}$. We denote by $\mathcal{J}(\Lambda, \mu, s)$ the induced representation $\mathrm{ind}_{B^G}^G(\chi)$ (unnormalized induction), where

$$\chi\left(\begin{bmatrix} xt & * \\ & \bar{t} \end{bmatrix}\right) = \mu(x)|x|^{s+1/2}\Lambda(t)^{-1}. \tag{71}$$

It is easy to see that the contragredient of $\mathcal{J}(\Lambda, \mu, s)$ is $\mathcal{J}(\Lambda^{-1}, \mu^{-1}, 1-s)$.

Let $V = L^2$, considered as a space of row vectors. Let $\mathcal{S}(V)$ be the space of Schwartz–Bruhat functions on V, i.e., the space of locally constant functions with compact support. For $g \in G$, $\Phi \in \mathcal{S}(V)$ and a complex number s, we define

$$\begin{aligned} &f^{\Phi}(g, \mu, \Lambda, s) \\ &\qquad := \mu(\det(g))|\det(g)|^{s+1/2} \int\limits_{L^\times} \Phi((0, \bar{t})g)|t\bar{t}|^{s+1/2}\mu(t\bar{t})\Lambda(t)\, d^\times t. \end{aligned} \tag{72}$$

This is the same definition as on page 265 of [Piatetski-Shapiro 1997], except we have $(0, \bar{t})$ instead of $(0, t)$, in order to be compatible with our conventions about Bessel models. Assuming convergence, a calculation shows that $f^{\Phi} \in \mathcal{J}(\Lambda, \mu, s)$.

Let $\mathcal{S}_0(V)$ be the subspace of $\Phi \in \mathcal{S}(V)$ for which $\Phi(0,0) = 0$. If $\Phi \in \mathcal{S}_0(V)$ and $g \in G$, then $\Phi((0, \bar{t})g) = 0$ for t outside a compact set of $L^\times$. It follows that the integral (72) converges absolutely for $\Phi \in \mathcal{S}_0(V)$, for any $s \in \mathbb{C}$.

Lemma 5.2.1. $\mathcal{J}(\Lambda, \mu, s) = \{f^{\Phi}(\,\cdot\,, \mu, \Lambda, s) : \Phi \in \mathcal{S}_0(V)\}$.

Proof. Given $f \in \mathcal{J}(\Lambda, \mu, s)$, we need to find $\Phi \in \mathcal{S}_0(V)$ such that $f^{\Phi} = f$. We define Φ by

$$\Phi(x, y) = \begin{cases} \mu^{-1}(\det(k)) f(k) & \text{if } (x, y) = (0, 1)k \text{ for some } k \in K^G, \\ 0 & \text{if } (x, y) \notin (0, 1)K^G. \end{cases} \tag{73}$$

It is straightforward to verify that Φ is well defined, that $\Phi \in \mathcal{S}_0(V)$, and that f^{Φ} is a multiple of f. □

Lemma 5.2.2. *Let* $\Lambda_\mu = \Lambda \cdot (\mu \circ N_{L/F})$.

(i) *The representation* $\mathcal{J}(\Lambda, \mu, s)$ *contains a one-dimensional* G*-invariant subspace if and only if*

$$\Lambda_\mu(t) = |t|_L^{-s-1/2} \quad \text{for all } t \in L^\times. \tag{74}$$

In this case the function

$$f(g) = \mu(\det(g))|\det(g)|^{s+1/2}, \quad g \in G, \tag{75}$$

spans a one-dimensional G*-invariant subspace of* $\mathrm{ind}_{BG}^{G}(\chi)$.

(ii) *The representation* $\mathcal{J}(\Lambda, \mu, s)$ *contains a one-dimensional* G*-invariant quotient if and only if*

$$\Lambda_\mu(t) = |t|_L^{-s+3/2} \quad \text{for all } t \in L^\times. \tag{76}$$

Proof. Part (i) is an easy exercise. Part (ii) follows from (i), observing that the contragredient of $\mathcal{J}(\Lambda, \mu, s)$ is $\mathcal{J}(\Lambda^{-1}, \mu^{-1}, 1-s)$. □

Note that condition (74) is equivalent to saying that s is a pole of $L(s+1/2, \Lambda_\mu)$. Later we will define the notion of *exceptional pole*; see (92). The exceptional poles will be among the poles of $L(s+1/2, \Lambda_\mu)$. Note that, by (73), the function f in (75) is a multiple of f^Φ, where

$$\Phi(x, y) = \begin{cases} 1 & \text{if } (x, y) = (0, 1)k \text{ for some } k \in K^G, \\ 0 & \text{if } (x, y) \notin (0, 1)K^G. \end{cases} \tag{77}$$

Hence, in the nonsplit case, Φ is the characteristic function of $(\mathfrak{o}_L \oplus \mathfrak{o}_L) \setminus (\mathfrak{p}_L \oplus \mathfrak{p}_L)$.

5.3. *The zeta integrals.* Let Λ be a character of $T \cong L^\times$, and let μ be a character of $F^\times$. Recall the definition of the functions $f^\Phi(g, \mu, \Lambda, s)$ in (72). Let π be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. Let $\mathcal{B}(\pi, \Lambda, \beta)$ be a (Λ, β)-Bessel model for π. For $B \in \mathcal{B}(\pi, \Lambda, \beta)$ and $s \in \mathbb{C}$, let

$$Z(s, B, \Phi, \mu) = \int\limits_{TN_0\backslash G} B(g) f^\Phi(g, \mu, \Lambda, s)\, dg, \tag{78}$$

provided this integral converges. (In [Piatetski-Shapiro 1997] this integral was denoted by $L(W, \Phi, \mu, s)$.) Substituting the definition of $f^\Phi(g, \mu, \Lambda, s)$ and unfolding the integral shows that

$$Z(s, B, \Phi, \mu) = \int\limits_{N_0\backslash G} B(g)\Phi((0, 1)g)\mu(\det(g))|\det(g)|^{s+1/2}\, dg. \tag{79}$$

By (68), we have

$$\begin{aligned} Z(s, B, \Phi, \mu) \\ = \int\limits_{A^G}\int\limits_{K^G} \delta(a)^{-1} B(ak)\Phi((0, 1)ak)\mu(\det(ak))|\det(ak)|^{s+1/2}\, dk\, da. \end{aligned} \tag{80}$$

Recall that $\mathcal{S}_0(V)$ is the space of $\Phi \in \mathcal{S}(V)$ satisfying $\Phi(0, 0) = 0$. Let $\Phi_1 \in \mathcal{S}(V)$ be the characteristic function of $\mathfrak{o}_L \oplus \mathfrak{o}_L$. Then every $\Phi \in \mathcal{S}(V)$ can be written in a unique way as $\Phi = \Phi_0 + c\Phi_1$ with $\Phi_0 \in \mathcal{S}_0(V)$ and $c \in \mathbb{C}$. We will first investigate $Z(s, B, \Phi, \mu)$ for $\Phi \in \mathcal{S}_0(V)$.

Lemma 5.3.1. *Let the notations and hypotheses be as above.*

(i) *For any $B \in \mathcal{B}(\pi, \Lambda, \beta)$ and $\Phi \in \mathcal{S}_0(V)$, the function $Z(s, B, \Phi, \mu)$ converges for real part of s large enough to an element of $\mathbb{C}(q^{-s})$. This element lies in the ideal $I(\pi, \mu)$ generated by all simplified zeta integrals*; *see Proposition 5.1.1.*

(ii) *For any $B \in \mathcal{B}(\pi, \Lambda, \beta)$, there exists $\Phi \in \mathcal{S}_0(V)$ such that $Z(s, B, \Phi, \mu) = \zeta(s, B, \mu)$.*

Hence, the integrals $Z(s, B, \Phi, \mu)$, as B runs through $\mathcal{B}(\pi, \Lambda, \beta)$ and Φ runs through $\mathcal{S}_0(V)$, generate the ideal $I(\pi, \mu)$ already exhibited in Proposition 5.1.1.

Proof. (i) Let $\Phi \in \mathcal{S}_0(V)$. We have

$$\Phi((0,1)ak) = \Phi(\bar{t}k_3, \bar{t}k_4) \quad \text{if } a = \begin{bmatrix} xt & \\ & \bar{t} \end{bmatrix} \in A^G,\ k = \begin{bmatrix} k_1 & k_2 \\ k_3 & k_4 \end{bmatrix} \in K^G. \tag{81}$$

Since one of k_3 or k_4 is a unit and $\Phi(0,0) = 0$, it follows that $\Phi((0,1)ak) = 0$ if t is outside a compact set of $L^\times$. As a consequence, there exists a small subgroup Γ of K^G such that

$$\Phi((0,1)ak\gamma) = \Phi((0,1)ak)$$

for all $a \in A^G$, $k \in K^G$ and $\gamma \in \Gamma$. By making Γ even smaller, we may assume that B and $\mu \circ \det$ are right Γ-invariant. It follows that $Z(s, B, \Phi, \mu)$ as in (80) is a finite sum of integrals of the form

$$I(s, B, \Phi, \mu) = \int_{A^G} \delta(a)^{-1} B(a) \Phi((0,1)a) \mu(\det(a)) |\det(a)|^{s+1/2}\, da, \tag{82}$$

with different B and $\Phi \in \mathcal{S}_0(V)$. Using coordinates on A^G, we have

$$\begin{aligned} &I(s, B, \Phi, \mu) \\ &= \int_{F^\times} \int_{L^\times} |xt\bar{t}^{-1}|_L^{-1} B\left(\begin{bmatrix} xt & \\ & \bar{t} \end{bmatrix}\right) \Phi(0, \bar{t}) \mu(xt\bar{t}) |xt\bar{t}|^{s+1/2}\, d^\times t\, d^\times x \\ &= \int_{F^\times} \int_{L^\times} |x|^{-2} \Lambda(t) B\left(\begin{bmatrix} x & \\ & 1 \end{bmatrix}\right) \Phi(0, \bar{t}) \mu(xt\bar{t}) |xt\bar{t}|^{s+1/2}\, d^\times t\, d^\times x \\ &= \left(\int_{F^\times} B\left(\begin{bmatrix} x & \\ & 1 \end{bmatrix}\right) \mu(x) |x|^{s-3/2}\, d^\times x\right) \left(\int_{L^\times} \Lambda(t) \Phi(0, \bar{t}) \mu(t\bar{t}) |t\bar{t}|^{s+1/2}\, d^\times t\right). \end{aligned} \tag{83}$$

The first integral is precisely $\zeta(s, B, \mu)$; see (52). Since the integration in the second integral is over a compact subset of $L^\times$, this integral is in $\mathbb{C}[q^s, q^{-s}]$. It follows that $I(s, B, \Phi, \mu)$ lies in the ideal $I(\pi, \mu)$.

(ii) By (79) and (69), we have

$$\begin{aligned} Z(s, B, \Phi, \mu) &= \int_{N_0} \int_{A^G} B(wna) \Phi((0,1)wna) \mu(\det(a)) |\det(a)|^{s+1/2}\, da\, dn \\ &= \int_{N_0} \int_{A^G} B(wna) \Phi((-1,0)na) \mu(\det(a)) |\det(a)|^{s+1/2}\, da\, dn \end{aligned}$$

$$= \int\limits_{L}\int\limits_{F^\times}\int\limits_{L^\times} B\left(w\begin{bmatrix}1&n\\&1\end{bmatrix}\begin{bmatrix}xt&\\&\bar t\end{bmatrix}\right)\Phi\left((-1,0)\begin{bmatrix}1&n\\&1\end{bmatrix}\begin{bmatrix}xt&\\&\bar t\end{bmatrix}\right) \times \mu(xt\bar t)|xt\bar t|^{s+1/2}\,d^\times t\,d^\times x\,dn$$

$$= \int\limits_{L}\int\limits_{F^\times}\int\limits_{L^\times} B\left(w\begin{bmatrix}xt&\bar t n\\&\bar t\end{bmatrix}\right)\Phi(-xt,-\bar t n)\mu(xt\bar t)|x|^{s+1/2}|t|_L^{s+1/2}\,d^\times t\,d^\times x\,dn$$

$$= \int\limits_{L}\int\limits_{F^\times}\int\limits_{L^\times} B\left(w\begin{bmatrix}xt&n\\&\bar t\end{bmatrix}\right)\Phi(-xt,-n)\mu(xt\bar t)|x|^{s+1/2}|t|_L^{s-1/2}\,d^\times t\,d^\times x\,dn$$

$$= \int\limits_{L}\int\limits_{F^\times}\int\limits_{L^\times} B\left(w\begin{bmatrix}1&\\&x^{-1}\end{bmatrix}\begin{bmatrix}t&n\\&\bar t\end{bmatrix}\right)\Phi(-t,-n) \times \mu(x)^{-1}\mu(t\bar t)|x|^{3/2-s}|t|_L^{s-1/2}\,d^\times t\,d^\times x\,dn.$$

Now choose Φ such that $\Phi(-t,-n)$ is zero unless t is close to 1 and n is close to 0. If the support of Φ is chosen small enough, then, after appropriate normalization,

$$Z(s,B,\Phi,\mu) = \int\limits_{F^\times} B\left(\begin{bmatrix}x^{-1}&\\&1\end{bmatrix}w\right)\mu(x)^{-1}|x|^{3/2-s}\,d^\times x.$$

This is just $\zeta(s,wB,\mu)$. The assertion follows. □

We see from Lemma 5.3.1 that, instead of (53), we could have chosen to define $L^{\mathrm{PS}}_{\mathrm{reg}}(s,\pi,\mu)$ as the gcd of all $Z(s,B,\Phi,\mu)$, as B runs through $\mathcal{B}(\pi,\Lambda,\beta)$ and Φ runs through $\mathcal{S}_0(V)$. The same observation was made in [Danişman 2015b, Proposition 18(i)].

Next we investigate $Z(s,B,\Phi_1,\mu)$, where we recall Φ_1 is the characteristic function of $\mathfrak{o}_L\oplus\mathfrak{o}_L$. In the split case, a character Λ of $L^\times = F^\times\times F^\times$ is a pair (λ_1,λ_2) of characters of $F^\times$, and by $L(s,\Lambda)$ we mean $L(s,\lambda_1)L(s,\lambda_2)$.

Lemma 5.3.2. *Let* $\Lambda_\mu = \Lambda\cdot(\mu\circ N_{L/F})$.

(i) *Assume that* Λ_μ *is ramified. Then* $Z(s,B,\Phi_1,\mu)=0$.

(ii) *Assume that* Λ_μ *is unramified. Then*

$$Z(s,B,\Phi_1,\mu) = \zeta(s,B_\mu,\mu)L(s+1/2,\Lambda_\mu), \tag{84}$$

where

$$B_\mu(g) := \int\limits_{K^G} B(gk)\mu(\det(k))\,dk, \qquad g\in \mathrm{GSp}(4,F). \tag{85}$$

Proof. Evidently, $\Phi_1((x, y)k) = \Phi_1(x, y)$ for all $(x, y) \in V$ and $k \in K^G$. Therefore, from (80), we get

$$(86)\quad Z(s, B, \Phi_1, \mu) = \int_{A^G}\int_{K^G} \delta(a)^{-1} B(ak)\Phi_1((0,1)a)\mu(\det(ak))|\det(a)|^{s+1/2}\,dk\,da = \int_{A^G} \delta(a)^{-1} B_\mu(a)\Phi_1((0,1)a)\mu(\det(a))|\det(a)|^{s+1/2}\,da.$$

Clearly, B_μ is an element of $\mathcal{B}(\pi, \Lambda, \beta)$ satisfying

$$B_\mu(gk) = \mu^{-1}(\det(k))B_\mu(g)$$

for $k \in K^G$. Using coordinates on A^G, we have

$$(87)\quad Z(s, B, \Phi_1, \mu) = \int_{F^\times}\int_{L^\times} |xt\bar{t}^{-1}|_L^{-1} B_\mu(a)\Phi_1((0,\bar{t}))\mu(xt\bar{t})|xt\bar{t}|^{s+1/2}\,d^\times t\,d^\times x$$
$$= \int_{F^\times}\int_{L^\times} B_\mu\left(\begin{bmatrix} xt & \\ & \bar{t}\end{bmatrix}\right)\Phi_1((0,\bar{t}))\mu(xt\bar{t})|t\bar{t}|^{s+1/2}|x|^{s-3/2}\,d^\times t\,d^\times x$$
$$= \int_{F^\times}\int_{L^\times\cap\mathfrak{o}_L} \Lambda(t)B_\mu\left(\begin{bmatrix} x & \\ & 1\end{bmatrix}\right)\mu(xt\bar{t})|t\bar{t}|^{s+1/2}|x|^{s-3/2}\,d^\times t\,d^\times x$$
$$= \zeta(s, B_\mu, \mu)\int_{L^\times\cap\mathfrak{o}_L} \Lambda(t)\mu(t\bar{t})|t\bar{t}|^{s+1/2}\,d^\times t.$$

It is straightforward to calculate that

$$(88)\quad \int_{L^\times\cap\mathfrak{o}_L} \Lambda(t)\mu(t\bar{t})|t\bar{t}|^{s+1/2}\,d^\times t = \begin{cases} L(s+1/2, \Lambda_\mu) & \text{if } \Lambda_\mu \text{ is unramified,}\\ 0 & \text{if } \Lambda_\mu \text{ is ramified.}\end{cases} \qquad \square$$

We see from Lemma 5.3.1 and Lemma 5.3.2 that $Z(s, B, \Phi, \mu)$ converges for real part of s large enough to an element of $\mathbb{C}(q^{-s})$, for any $B \in \mathcal{B}(\pi, \Lambda, \beta)$ and $\Phi \in \mathcal{S}(V)$. Let $I_{\Lambda,\beta}(\pi, \mu)$ be the $\mathbb{C}$-vector subspace of $\mathbb{C}(q^{-s})$ spanned by all $\zeta(s, B, \mu)$ as B runs through $\mathcal{B}(\pi, \Lambda, \beta)$.

Proposition 5.3.3. *Let π be an irreducible, admissible representation of* GSp(4, F) *admitting a (Λ, β)-Bessel model with β as in* (4). *Then $I_{\Lambda,\beta}(\pi, \mu)$ is a nonzero $\mathbb{C}[q^{-s}, q^s]$ module containing $\mathbb{C}$, and there exists $R(X) \in \mathbb{C}[X]$ such that*

$$R(q^{-s})I_{\Lambda,\beta}(\pi, \mu) \subset \mathbb{C}[q^{-s}, q^s],$$

so that $I_{\Lambda,\beta}(\pi,\mu)$ is a fractional ideal of the principal ideal domain $\mathbb{C}[q^{-s},q^{s}]$ whose quotient field is $\mathbb{C}(q^{-s})$. The fractional ideal $I_{\Lambda,\beta}(\pi,\mu)$ admits a generator of the form $1/Q(q^{-s})$ with $Q(0)=1$, where $Q(X)\in\mathbb{C}[X]$.

Proof. The proof is similar to that of Proposition 5.1.1. It follows easily from (79) that $I_{\Lambda,\beta}(\pi,\mu)$ is a $\mathbb{C}[q^{s},q^{-s}]$-module. It follows from Proposition 5.1.1 and Lemma 5.3.1 that $I_{\Lambda,\beta}(\pi,\mu)$ contains $\mathbb{C}$. □

Using the notation of this proposition, we set

$$L_{\Lambda}^{\mathrm{PS}}(s,\pi,\mu):=1/Q(q^{-s}). \tag{89}$$

This is the Piatetski-Shapiro L-factor, as defined in [Piatetski-Shapiro 1997]. Our notation indicates that these factors may depend on Λ (and β, which we suppress from the notation).

We now distinguish two cases. In the first, assume

$$\frac{Z(s,B,\Phi,\mu)}{L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu)} \text{ is entire for all } B\in\mathcal{B}(\pi,\Lambda,\beta) \text{ and } \Phi\in\mathcal{S}(V). \tag{90}$$

Being entire is equivalent to lying in $\mathbb{C}[q^{s},q^{-s}]$. Hence, in this case the fractional ideal generated by all $Z(s,B,\Phi,\mu)$ is generated by $L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu)$, and we have

$$L_{\Lambda}^{\mathrm{PS}}(s,\pi,\mu)=L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu). \tag{91}$$

In particular, the Piatetski-Shapiro L-factor does not depend on Λ in this case.

For the second case, assume

$$\frac{Z(s,B,\Phi,\mu)}{L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu)} \text{ has a pole for some } B\in\mathcal{B}(\pi,\Lambda,\beta) \text{ and } \Phi\in\mathcal{S}(V). \tag{92}$$

Such poles are called *exceptional poles*. By (84), exceptional poles are precisely the poles of

$$\frac{\zeta(s,B_{\mu},\mu)}{L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu)}L(s+1/2,\Lambda_{\mu}), \tag{93}$$

as B runs through $\mathcal{B}(\pi,\Lambda,\beta)$. Since the fraction in (93) is entire, exceptional poles are found among the poles of $L(s+1/2,\Lambda_{\mu})$. If we write

$$L(s,\Lambda_{\mu})=\frac{1}{(1-\gamma_1q^{-s})(1-\gamma_2q^{-s})}, \tag{94}$$

where one of the complex numbers γ_1,γ_2 may be zero, then

$$L^{\mathrm{PS}}(s,\pi,\mu)=L_{\mathrm{reg}}^{\mathrm{PS}}(s,\pi,\mu)\frac{1}{P(q^{-s-1/2})}, \tag{95}$$

where $P\in\mathbb{C}[X]$ is either $1-\gamma_iX$ or $(1-\gamma_1X)(1-\gamma_2X)$.

Remark. Our definition of exceptional pole is slightly more general than the one given in [Piatetski-Shapiro 1997]. Therein, a complex number s_0 is called an exceptional pole if s_0 is a pole of $L^{\mathrm{PS}}(s, \pi, \mu)$ but not of $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$. It follows easily that an exceptional pole according to Piatetski-Shapiro is also an exceptional pole according to our definition. However, the two notions may not coincide if there is overlap between the poles of $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$ and the poles of $L(s+1/2, \Lambda_\mu)$.

The *regular poles* are the poles of $L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$. According to our definition, an exceptional pole can also be regular, while in [Piatetski-Shapiro 1997] the two notions are exclusive. Our definition is designed in such a way that $L^{\mathrm{PS}}(s, \pi, \mu) \neq L^{\mathrm{PS}}_{\mathrm{reg}}(s, \pi, \mu)$ precisely if there exist exceptional poles.

5.4. *Double coset decompositions.* We first prove the following double coset decomposition for GL(2, F). Let β be as in (4), and let T be the group of all

$$\begin{bmatrix} x+yb/2 & yc \\ -ya & x-yb/2 \end{bmatrix} \in \mathrm{GL}(2, F), \quad x^2 - y^2\Big(\frac{b^2}{4} - ac\Big) \neq 0. \tag{96}$$

Recall that we are in the *split case* if and only if $b^2 - 4ac \in F^{\times 2}$. We can and will make the assumption that

$$a, c \neq 0. \tag{97}$$

In the split case, let $r_1, r_2 \in F^\times$ be the two roots of the equation

$$ar^2 + br + c = 0. \tag{98}$$

Let B_1 be the subgroup of GL(2, F) consisting of all elements of the form $\left[\begin{smallmatrix} 1 & * \\ & * \end{smallmatrix}\right]$, and let B_2 be the subgroup consisting of all elements of the form $\left[\begin{smallmatrix} 1 & \\ * & * \end{smallmatrix}\right]$.

Lemma 5.4.1. (i) *In the nonsplit case*, $\mathrm{GL}(2, F) = TB_1 = TB_2$.

(ii) *In the split case*,

$$\begin{aligned} \mathrm{GL}(2, F) &= TB_1 \sqcup Tg_1sB_1 \sqcup Tg_2sB_1 \\ &= TB_2 \sqcup Tg_1B_2 \sqcup Tg_2B_2, \quad \textit{where } g_i = \begin{bmatrix} 1 & r_i \\ & 1 \end{bmatrix}, \ s = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}. \end{aligned} \tag{99}$$

The set TB_1 (*resp.* TB_2) *is open and dense in* GL(2, F), *and consists of all* $\left[\begin{smallmatrix} a_1 & a_2 \\ a_3 & a_4 \end{smallmatrix}\right] \in \mathrm{GL}(2, F)$ *with* $aa_1^2 + ba_1a_3 + ca_3^2 \neq 0$ (*resp.* $aa_2^2 + ba_2a_4 + ca_4^2 \neq 0$). *For* $i = 1$ *or* 2, *the set* Tg_isB_1 (*resp.* Tg_iB_2) *consists of all* $\left[\begin{smallmatrix} a_1 & a_2 \\ a_3 & a_4 \end{smallmatrix}\right] \in \mathrm{GL}(2, F)$ *with* $a_1 = a_3r_i$ (*resp.* $a_2 = a_4r_i$).

Proof. Calculations show that if $aa_1^2 + ba_1a_3 + ca_3^2 \neq 0$, then the equation

$$\begin{bmatrix} x+yb/2 & yc \\ -ya & x-yb/2 \end{bmatrix} \begin{bmatrix} 1 & z \\ & d \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

can be solved for x, y, z, d. Assume that $aa_1^2+ba_1a_3+ca_3^2=0$. Then $a_1=a_3r_i$ for $i=1$ or $i=2$. Calculations show that the equation

$$\begin{bmatrix} x+yb/2 & yc \\ -ya & x-yb/2 \end{bmatrix} g_i s \begin{bmatrix} 1 & z \\ & d \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

can be solved for x, y, z, d. This proves the statements for B_1, and the proof for B_2 is similar. $\square$

Let P be the (F-points of the) Siegel parabolic subgroup of $\mathrm{GSp}(4, F)$; see (2). Let G be the group defined in (56). We assume that $\beta = \left[\begin{smallmatrix} a & \\ & c \end{smallmatrix}\right]$ with $ac \neq 0$, and embed G into $\mathrm{GSp}(4, F)$ such that (59) to (62) hold. More generally, if

$$g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in G,$$

then a calculation shows that, as an element of $\mathrm{GSp}(4, F)$,

$$g = \begin{bmatrix} \alpha_1 & c\alpha_2 & 2\beta_1 & -2a\beta_2 \\ -a\alpha_2 & \alpha_1 & -2a\beta_2 & -\frac{2a}{c}\beta_1 \\ \frac{1}{2}\gamma_1 & \frac{c}{2}\gamma_2 & \delta_1 & -a\delta_2 \\ \frac{c}{2}\gamma_2 & -\frac{c}{2a}\gamma_1 & c\delta_2 & \delta_1 \end{bmatrix}. \tag{100}$$

Here, $\alpha = \alpha_1 + \Delta\alpha_2$ etc., with Δ as defined after (7). The following result is a more precise version of a remark made in the proof of Theorem 4.3 of [Piatetski-Shapiro 1997].

Lemma 5.4.2. *Assume the above notations and hypotheses. Let*

$$s_2 = \begin{bmatrix} & & 1 & \\ & 1 & & \\ -1 & & & \\ & & & 1 \end{bmatrix}. \tag{101}$$

Then

$$\mathrm{GSp}(4, F) = GP \sqcup Gs_2P. \tag{102}$$

The double coset Gs_2P is open and dense in $\mathrm{GSp}(4, F)$, *and*

$$s_2^{-1}Gs_2 \cap P = \left\{ \begin{bmatrix} A & \\ & \det(A)\,{}^tA^{-1} \end{bmatrix} : A \in \mathrm{GL}(2, F) \right\}. \tag{103}$$

We have $Gs_2P = Gs_2HN$, where H and N are defined in (3) *and* (2), *respectively. Furthermore*,

$$GP = \begin{cases} GB_2N & \textit{in the nonsplit case}, \\ GB_2N \sqcup Gg_1B_2N \sqcup Gg_2B_2N & \textit{in the split case}, \end{cases} \tag{104}$$

where

$$B_2 = \left\{ \begin{bmatrix} 1 &&& \\ x & y && \\ && y & -x \\ &&& 1 \end{bmatrix} : x \in F,\ y \in F^\times \right\}, \quad g_i = \begin{bmatrix} 1 & r_i && \\ & 1 && \\ &&1& \\ && -r_i & 1 \end{bmatrix}, \tag{105}$$

with $r_1, r_2 \in F^\times$ *being the two roots of the equation* $ar^2 + c = 0$.

Proof. Using the description (100) of the elements of G, it is easy to verify (103). As a consequence, $Gs_2P = Gs_2HN$. Equation (104) follows from (99); the disjointness in the split case is easy to verify.

By the Bruhat decomposition,

$$\mathrm{GSp}(4, F) = P \sqcup \begin{bmatrix} 1 && * & \\ & 1 && \\ && 1 & \\ &&& 1 \end{bmatrix} s_2P \sqcup \begin{bmatrix} 1 &&& \\ * & 1 && * \\ && 1 & * \\ &&& 1 \end{bmatrix} s_1s_2P \sqcup \begin{bmatrix} 1 && * & * \\ & 1 & * & * \\ && 1 & \\ &&& 1 \end{bmatrix} s_2s_1s_2P. \tag{106}$$

Calculations show that

$$Gs_2P \cap \begin{bmatrix} 1 && * & * \\ & 1 & * & * \\ && 1 & \\ &&& 1 \end{bmatrix} s_2s_1s_2P = \left\{ \begin{bmatrix} 1 & X \\ & 1 \end{bmatrix} s_2s_1s_2p : p \in P,\ \mathrm{tr}(\beta X) \neq 0 \right\}, \tag{107}$$

$$Gs_2P \cap \begin{bmatrix} 1 &&& \\ * & 1 && * \\ && 1 & * \\ &&& 1 \end{bmatrix} s_1s_2P = \left\{ \begin{bmatrix} 1 &&& \\ x & 1 && z \\ && 1 & -x \\ &&& 1 \end{bmatrix} s_1s_2p : p \in P,\ x^2 \neq -a/c \right\}, \tag{108}$$

$$Gs_2P \cap \begin{bmatrix} 1 && * & \\ & 1 && \\ && 1 & \\ &&& 1 \end{bmatrix} s_2P = \begin{bmatrix} 1 && * & \\ & 1 && \\ && 1 & \\ &&& 1 \end{bmatrix} s_2P, \tag{109}$$

$$Gs_2P \cap P = \varnothing, \tag{110}$$

and

$$GP \cap \begin{bmatrix} 1 && * & * \\ & 1 & * & * \\ && 1 & \\ &&& 1 \end{bmatrix} s_2s_1s_2P = \left\{ \begin{bmatrix} 1 & X \\ & 1 \end{bmatrix} s_2s_1s_2p : p \in P,\ \mathrm{tr}(\beta X) = 0 \right\}, \tag{111}$$

$$GP \cap \begin{bmatrix} 1 & & & \\ * & 1 & & * \\ & & 1 & * \\ & & & 1 \end{bmatrix} s_1 s_2 P = \left\{ \begin{bmatrix} 1 & & & \\ x & 1 & & z \\ & & 1 & -x \\ & & & 1 \end{bmatrix} s_1 s_2 p \,:\, p \in P,\, x^2 = -a/c \right\}, \tag{112}$$

$$GP \cap \begin{bmatrix} 1 & & * & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} s_2 P = \varnothing, \tag{113}$$

$$GP \cap P = P. \tag{114}$$

It follows that $\mathrm{GSp}(4, F) = GP \sqcup Gs_2P$. Since the big Bruhat cell is dense in $\mathrm{GSp}(4, F)$, (107) implies that Gs_2P is also dense in $\mathrm{GSp}(4, F)$. Since $GP = K^G B^G P = K^G P$ is the product of a compact and a closed set, it is closed in $\mathrm{GSp}(4, F)$. □

In the proof of the following lemma we will make use of the fact that a continuous bijection $X \to Y$ between p-adic spaces is a homeomorphism. This is because we can cover X with open-compact subsets, and a continuous bijection from a compact topological space to a Hausdorff space is a homeomorphism.

For a locally compact, totally disconnected space X, we denote by $\mathcal{S}(X)$ the space of locally constant functions $X \to \mathbb{C}$ with compact support. If X is a group, $h \in X$ and $\phi \in \mathcal{S}(X)$, we denote by $R_h\phi$ the element of $\mathcal{S}(X)$ given by $x \mapsto \phi(xh)$, and by $L_h\phi$ the element of $\mathcal{S}(X)$ given by $x \mapsto \phi(h^{-1}x)$.

Let U be the unipotent radical of the Borel subgroup of $\mathrm{GSp}(4, F)$. Then U consists of all matrices of the form

$$\begin{bmatrix} 1 & & * & * \\ * & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{bmatrix}$$

in $\mathrm{GSp}(4, F)$. For $c_1, c_2 \in F$, we define a character ψ_{c_1,c_2} of U by

$$\psi_{c_1,c_2}\left(\begin{bmatrix} 1 & & y & * \\ x & 1 & * & * \\ & & 1 & -x \\ & & & 1 \end{bmatrix} \right) = \psi(c_1 x + c_2 y). \tag{115}$$

The statement of the following result was mentioned in the proof of Theorem 4.3 of [Piatetski-Shapiro 1997].

Lemma 5.4.3. *Let* $D : \mathcal{S}(\mathrm{GSp}(4, F)) \to \mathbb{C}$ *be a distribution on* $\mathrm{GSp}(4, F)$ *with the following properties*:

- *There exist $c_1, c_2 \in F^\times$ such that*

$$D(R_u\phi) = \psi_{c_1,c_2}(u)D(\phi) \quad \textit{for all } u \in U \tag{116}$$

  and all $\phi \in \mathcal{S}(\mathrm{GSp}(4, F))$.

- *There exists a character β of G such that*

$$D(L_h\phi) = \beta(h)D(\phi) \quad \textit{for all } h \in G \tag{117}$$

  and all $\phi \in \mathcal{S}(\mathrm{GSp}(4, F))$.

Then $D = 0$.

Proof. Since $\mathrm{GSp}(4, F) = GP \sqcup Gs_2P$, it suffices to show that a distribution on $\mathcal{S}(Gs_2P)$ with the properties (116) and (117) is zero, and a distribution on $\mathcal{S}(GP)$ with those properties is also zero.

(1) First we prove that a distribution D on Gs_2P with the properties (116) and (117) must be zero. For $x \in F^\times$, let $h_x = \mathrm{diag}(x, x, 1, 1)$. By Lemma 5.4.2, $Gs_2P = Gs_2HN$. In fact, every element of Gs_2P can be written in the form gs_2h_xn with $g \in G$ and uniquely determined $x \in F^\times$ and $n \in N$. Hence Gs_2P is homeomorphic to $G \times H \times N$. We consider the continuous map

$$p : Gs_2P \to F^\times \qquad \text{defined by } \; gs_2h_xn \longmapsto x.$$

The set Gs_2P is invariant under the left action of G and the right action of U. It is easy to see that every fiber $p^{-1}(x)$ is $G \times U$-invariant. By Corollary 2.1 of [Aizenbud et al. 2010], Bernstein's localization principle, it is sufficient to prove that any distribution D on $\mathcal{S}(p^{-1}(x))$ with the properties (116) and (117) vanishes, for all $x \in F^\times$.

We apply Proposition 4.3.2 of [Bump 1997] with

$$G \times N \cong Gs_2h_xN = p^{-1}(x).$$

It shows that there exists a constant $c_1 \in \mathbb{C}$ such that

$$D(\phi) = c_1 \int_G \int_N \beta(g)\, \psi_{c_1,c_2}^{-1}(n)\, \phi(gs_2h_xn)\, dn\, dg$$

for all $\phi \in \mathcal{S}(p^{-1}(x))$. We may choose some $z \in F$ such that

$$\psi_{c_1,c_2}(u_z) \neq 1 \quad \text{for } u_z = \begin{bmatrix} 1 & & & \\ z & 1 & & \\ & & 1 & -z \\ & & & 1 \end{bmatrix}.$$

By (62),

$$n_z := s_2 u_z s_2^{-1} = \begin{bmatrix} 1 & & & -z \\ & 1 & -z & \\ & & 1 & \\ & & & 1 \end{bmatrix} \in N_0 \subset G,$$

so that $D(L_{n_z^{-1}}\phi) = \beta(n_z^{-1})D(\phi) = D(\phi)$ by (117). On the other hand, the substitution $g \mapsto n_z^{-1} g n_z$ shows that

$$\begin{aligned} D(L_{n_z^{-1}}\phi) &= c_1 \int\limits_G \int\limits_N \phi(n_z g s_2 h_x n)\beta(g)\psi_{c_1,c_2}^{-1}(n)\,dn\,dg \\ &= c_1 \int\limits_G \int\limits_N \phi(g n_z s_2 h_x n)\beta(g)\psi_{c_1,c_2}^{-1}(n)\,dn\,dg \\ &= c_1 \int\limits_G \int\limits_N \Phi(g s_2 u_z h_x n)\beta(g)\psi_{c_1,c_2}^{-1}(n)\,dn\,dg \\ &= c_1 \int\limits_G \int\limits_N \Phi(g s_2 h_x n u_z)\beta(g)\psi_{c_1,c_2}^{-1}(n)\,dn\,dg \\ &= \psi_{c_1,c_2}(u_z) c_1 \int\limits_G \int\limits_N \Phi(g s_2 h_x n)\beta(g)\psi_{c_1,c_2}^{-1}(n)\,dn\,dg. \end{aligned}$$

In the last step we used (116). Hence $D(\phi) = \psi_{c_1,c_2}(u_z)D(\phi)$, which implies $D = 0$ on $\mathcal{S}(p^{-1}(x))$.

(2) Next, using the decomposition (104), we prove that a distribution D on GP with the properties (116) and (117) must be zero.

(2.1) We will first show that a distribution D on GB_2N with the properties (116) and (117) must be zero. We define the groups

$$H_1 := \left\{ k_x = \begin{bmatrix} 1 & & & \\ & x & & \\ & & x & \\ & & & 1 \end{bmatrix} : x \in F^\times \right\}, \quad U_1 := \begin{bmatrix} 1 & & & \\ * & 1 & & * \\ & & 1 & * \\ & & & 1 \end{bmatrix} \cap \mathrm{GSp}(4, F). \tag{118}$$

Then, with N_0 as in (63),

$$GB_2N = GUH_1 = GN_0U_1H_1 = GU_1H_1 = GH_1U_1. \tag{119}$$

In fact, it is not difficult to see that any element of GP can be written in the form gk_xu with uniquely determined $g \in G$, $x \in F^\times$ and $u \in U_1$. Hence GB_2N is homeomorphic to $G \times H_1 \times U_1$. We consider the continuous map

$$p : GB_2N \to F^\times \quad \text{defined by } gk_xu \longmapsto x.$$

The set GB_2N is invariant under the left action of G and the right action of U. It is easy to see that every fiber $p^{-1}(x)$ is $G\times U$-invariant. By Bernstein's localization principle, it is enough to show that a distribution D on $p^{-1}(x)$ with the properties (116) and (117) vanishes.

We apply Proposition 4.3.2 of [Bump 1997] to

$$G\times U_1 \cong Gk_xU_1 = p^{-1}(x).$$

It shows that there exists a constant $c_2\in\mathbb{C}$ such that

$$D(\phi)=c_2\int\limits_G\int\limits_{U_1}\beta(g)\,\psi^{-1}_{c_1,c_2}(u_1)\,\phi\left(g\begin{bmatrix}1&&&\\&x&&\\&&x&\\&&&1\end{bmatrix}u_1\right)du_1\,dg \tag{120}$$

for any $\phi\in\mathcal{S}(p^{-1}(x))$. Let $t\in F^\times$ be such that $\psi(c_22tx)\neq 1$,

$$n:=\begin{bmatrix}1&2t&&\\&1&&-2ac^{-1}t\\&&1&\\&&&1\end{bmatrix}\in N_0\subset G\quad\text{and}\quad u:=\begin{bmatrix}1&2tx&&\\&1&&\\&&1&\\&&&1\end{bmatrix}. \tag{121}$$

Hence,

$$\psi_{c_1,c_2}(u)=\psi(c_22tx)\neq 1.$$

Much as above, we calculate

$$\begin{aligned}
&D(L_{n^{-1}}\phi)\\
&=c_2\int\limits_G\int\limits_{U_1}\beta(g)\,\psi^{-1}_{c_1,c_2}(u_1)\,\phi(gnk_xu_1)\,du_1\,dg\\
&=c_2\int\limits_G\int\limits_{U_1}\beta(g)\,\psi^{-1}_{c_1,c_2}(u_1)\,\phi\left(gk_x\begin{bmatrix}1&2tx&&\\&1&&-2ac^{-1}tx^{-1}\\&&1&\\&&&1\end{bmatrix}u_1\right)du_1\,dg\\
&=c_2\int\limits_G\int\limits_F\int\limits_F\beta(g)\,\psi^{-1}(c_1y)\,\phi\left(gk_x\begin{bmatrix}1&2tx&&\\&1&&-2ac^{-1}tx^{-1}\\&&1&\\&&&1\end{bmatrix}\begin{bmatrix}1&&&\\y&1&&z\\&&1&-y\\&&&1\end{bmatrix}\right)dy\,dz\,dg\\
&=c_2\int\limits_G\int\limits_F\int\limits_F\beta(g)\,\psi^{-1}(c_1y)\,\phi\left(g\begin{bmatrix}1&&&-2txy\\&1&-2txy&\\&&1&\\&&&1\end{bmatrix}k_x\begin{bmatrix}1&&&\\y&1&&z\\&&1&-y\\&&&1\end{bmatrix}\begin{bmatrix}1&2tx&&\\&1&&\\&&1&\\&&&1\end{bmatrix}\right)dy\,dz\,dg
\end{aligned}$$

$$= c_2 \int_G \int_F \int_F \beta\left(g \begin{bmatrix} 1 & & & 2txy \\ & 1 & 2txy & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) \psi^{-1}(c_1 y)\, \phi\left(g k_x \begin{bmatrix} 1 & & & \\ y & 1 & & z \\ & & 1 & -y \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 2tx & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) dy\, dz\, dg$$

$$= c_2 \int_G \int_{U_1} \beta(g)\, \psi^{-1}(c_1 y)\, \phi\left(g k_x u_1 \begin{bmatrix} 1 & 2tx & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) du_1\, dg$$

$$= D(R_u \phi).$$

Hence, by (116) and (117),

$$\begin{aligned} D(\phi) = D(L_{n^{-1}}\phi) &= D(R_u \phi) \\ &= \psi(c_2 2tx) D(\phi). \end{aligned}$$

It follows that $D(\phi) = 0$.

(2.2) Now assume we are in the split case. Let $i \in \{1, 2\}$. We will show that a distribution D on $Gg_i B_2 N$ with the properties (116) and (117) must be zero. Calculations in coordinates verify that

$$g_i^{-1} G g_i \cap B_2 = \left\{ \begin{bmatrix} 1 & & & \\ \frac{y-1}{2r_i} & y & & \\ & & y & \frac{1-y}{2r_i} \\ & & & 1 \end{bmatrix} : y \in F^\times \right\}. \tag{122}$$

It follows that

$$G g_i B_2 N = G g_i H_1 N \sqcup G g_i \tilde{g}_i N, \quad \text{where } \tilde{g}_i = \begin{bmatrix} 1 & & & \\ -\frac{1}{2r_i} & 1 & & \\ & & 1 & \frac{1}{2r_i} \\ & & & 1 \end{bmatrix}, \tag{123}$$

and H_1 is as in (118). We will proceed to show that a distribution D on $Gg_i B_2 N$ with the properties (117) and

$$D(R_u \phi) = \psi(c_2 x) D(\phi) \quad \text{for all } u = \begin{bmatrix} 1 & & x & y \\ & 1 & y & z \\ & & 1 & \\ & & & 1 \end{bmatrix} \in N \tag{124}$$

must be zero.

(2.2.1) We will first show that a distribution D on Gg_iH_1N with the properties (117) and (124) vanishes. We have

$$g_i^{-1}Gg_i \cap H_1N = \left\{ \begin{bmatrix} 1 & & -2r_iu & u \\ & 1 & u & v \\ & & 1 & \\ & & & 1 \end{bmatrix} : u, v \in F \right\}. \tag{125}$$

Hence

$$Gg_iH_1N = Gg_iH_1U_2, \quad \text{where } U_2 = \begin{bmatrix} 1 & & * & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}. \tag{126}$$

In fact, every element of Gg_iH_1N can be written in the form gg_ik_xu with uniquely determined $x \in F^\times$ and $u \in U_2$. We consider the continuous map

$$p : Gg_iH_1N \to F^\times \quad \text{defined by } gg_ik_xu \longmapsto x.$$

It is easy to see that every fiber $p^{-1}(x)$ is $G \times N$-invariant. By Bernstein's localization principle, it is enough to show that a distribution D on $p^{-1}(x)$ with the properties (117) and (124) vanishes. We apply Proposition 4.3.2 of [Bump 1997] to

$$G \times U_2 \cong Gg_ik_xU_2 = p^{-1}(x).$$

It shows that there exists a constant $c_3 \in \mathbb{C}$ such that

$$D(\phi) = c_3 \int\limits_G \int\limits_F \beta(g)\, \psi^{-1}(c_2z)\, \phi\left(gg_ik_x \begin{bmatrix} 1 & & z & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \right) dz\, dg \tag{127}$$

for all $\phi \in \mathcal{S}(p^{-1}(x))$. Now, for any $y \in F$,

$$\begin{aligned} D(\phi) &= c_3 \int\limits_G \int\limits_F \beta(g)\, \psi^{-1}(c_2z)\, \phi\left(gg_ik_x \begin{bmatrix} 1 & & z & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & y \\ & 1 & y & \\ & & 1 & \\ & & & 1 \end{bmatrix} \right) dz\, dg \\ &= c_3 \int\limits_G \int\limits_F \beta(g)\, \psi^{-1}(c_2z)\, \phi\left(gg_i \begin{bmatrix} 1 & & & y \\ & 1 & y & \\ & & 1 & \\ & & & 1 \end{bmatrix} k_x \begin{bmatrix} 1 & & z & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \right) dz\, dg \end{aligned}$$

$$= c_3 \int\limits_G \int\limits_F \beta(g)\,\psi^{-1}(c_2 z)\,\phi\left(gg_i\begin{bmatrix}1&-2r_iy&y&\\&1&y&\\&&1&\\&&&1\end{bmatrix}g_i^{-1}g_i\begin{bmatrix}1&2r_iy&&\\&1&&\\&&1&\\&&&1\end{bmatrix}k_x\begin{bmatrix}1&z&&\\&1&&\\&&1&\\&&&1\end{bmatrix}\right)dz\,dg$$

$$= c_3 \int\limits_G \int\limits_F \beta(g)\,\psi^{-1}(c_2 z)\,\phi\left(gg_i\begin{bmatrix}1&2r_iy&&\\&1&&\\&&1&\\&&&1\end{bmatrix}k_x\begin{bmatrix}1&z&&\\&1&&\\&&1&\\&&&1\end{bmatrix}\right)dz\,dg$$

$$= c_3 \int\limits_G \int\limits_F \beta(g)\,\psi^{-1}(c_2 z)\,\phi\left(gg_ik_x\begin{bmatrix}1&z+2r_ixy&&\\&1&&\\&&1&\\&&&1\end{bmatrix}\right)dz\,dg$$

$$= \psi(c_2 2r_ixy)c_3 \int\limits_G \int\limits_F \beta(g)\,\psi^{-1}(c_2 z)\,\phi\left(gg_ik_x\begin{bmatrix}1&z&&\\&1&&\\&&1&\\&&&1\end{bmatrix}\right)dz\,dg$$

$$= \psi(c_2 2r_ixy)D(\phi).$$

It follows that $D(\phi) = 0$.

(2.2.2) Finally, we will show that a distribution D on $Gg_i\tilde{g}_iN$ with the properties (117) and (124) vanishes. We have

$$(g_i\tilde{g}_i)^{-1}Gg_i\tilde{g}_i \cap N = \left\{\begin{bmatrix}1&&u&\\&1&&v\\&&1&\\&&&1\end{bmatrix} : u, v \in F\right\}. \tag{128}$$

Hence

$$Gg_i\tilde{g}_iN = Gg_i\tilde{g}_iU_3, \quad \text{where } U_3 = \begin{bmatrix}1&&&*\\&1&*&\\&&1&\\&&&1\end{bmatrix}. \tag{129}$$

We apply Proposition 4.3.2 of [Bump 1997] to

$$G \times U_3 \cong Gg_i\tilde{g}_iU_3.$$

It shows that there exists a constant $c_4 \in \mathbb{C}$ such that

$$D(\phi) = c_4 \int\limits_G \int\limits_F \beta(g)\,\phi\left(gg_i\tilde{g}_i\begin{bmatrix}1&&&z\\&1&z&\\&&1&\\&&&1\end{bmatrix}\right)dz\,dg \tag{130}$$

for any $\phi \in \mathcal{S}(Gg_i\tilde{g}_iN)$. Then, for any $x \in F$,

$$\begin{aligned}
\psi(c_2x)D(\phi) &= c_4 \int_G \int_F \beta(g)\,\phi\left(gg_i\tilde{g}_i \begin{bmatrix} 1 & & & z \\ & 1 & z & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & x & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) dz\,dg \\
&= c_4 \int_G \int_F \beta(g)\,\phi\left(gg_i\tilde{g}_i \begin{bmatrix} 1 & x & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} (g_i\tilde{g}_i)^{-1} g_i\tilde{g}_i \begin{bmatrix} 1 & & & z \\ & 1 & z & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) dz\,dg \\
&= c_4 \int_G \int_F \beta(g)\,\phi\left(gg_i\tilde{g}_i \begin{bmatrix} 1 & & & z \\ & 1 & z & \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) dz\,dg \\
&= D(\phi).
\end{aligned}$$

It follows that $D(\phi) = 0$. This concludes the proof. □

5.5. *Some cases with no exceptional poles.* The following is Theorem 4.2 of [Piatetski-Shapiro 1997], with a slightly modified proof to accommodate our more general notion of exceptional pole.

Theorem 5.5.1. *Let μ be a character of $F^\times$. Let (π, V) be an irreducible, admissible representation of* $\mathrm{GSp}(4, F)$ *admitting a (Λ, β)-Bessel model. Assume that s_0 is an exceptional pole for the datum π, Λ, β, μ, as defined in the previous section. Then there exists a nonzero functional $\ell : V \to \mathbb{C}$ with the property*

$$\ell(\pi(g)v) = \mu^{-1}(\det(g))|\det(g)|^{-s_0-1/2}\,\ell(v) \quad \textit{for all } v \in V \textit{ and } g \in G. \tag{131}$$

Proof. By definition, the function

$$\frac{Z(s, B, \Phi, \mu)}{L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)} = \frac{Z(s, B, \Phi, \mu)}{L_{\mathrm{reg}}^{\mathrm{PS}}(s, \pi, \mu)L(s+1/2, \Lambda_\mu)} \tag{132}$$

lies in $\mathbb{C}[q^s, q^{-s}]$, for any choice of $B \in \mathcal{B}(\pi, \Lambda, \beta)$ and $\Phi \in \mathcal{S}(V)$. In particular, we may evaluate at s_0. We note that

$$\left.\frac{Z(s, B, \Phi, \mu)}{L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)}\right|_{s=s_0} = 0 \quad \text{if } \Phi \in \mathcal{S}_0(V). \tag{133}$$

This follows from Lemma 5.3.1(i), and the fact that s_0 is a pole of $L(s+1/2, \Lambda_\mu)$. We now define

$$\ell(B) = \left.\frac{Z(s, B, \Phi_1, \mu)}{L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)}\right|_{s=s_0}, \tag{134}$$

where, as before, Φ_1 is the characteristic function of $\mathfrak{o}_L \oplus \mathfrak{o}_L$. Since $Z(s, B, \Phi, \mu) = L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)$ for some choice of B and Φ, (133) implies that ℓ is a nonzero functional. It follows from (79) that

$$Z(s, \pi(g)B, g.\Phi, \mu) = Z(s, B, \Phi, \mu)\mu^{-1}(\det(g))|\det(g)|^{-s-1/2} \quad \text{for all } g \in G, \tag{135}$$

where $(g.\Phi)(x, y) = \Phi((x, y)g)$. Consequently,

$$\left.\frac{Z(s, \pi(g)B, g.\Phi_1, \mu)}{L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)}\right|_{s=s_0} = \left.\frac{Z(s, B, \Phi_1, \mu)}{L_\Lambda^{\mathrm{PS}}(s, \pi, \mu)}\right|_{s=s_0} \mu^{-1}(\det(g))|\det(g)|^{-s_0-1/2}. \tag{136}$$

Since $g.\Phi - \Phi \in \mathcal{S}_0(V)$, property (133) allows us to replace $g.\Phi$ on the left-hand side by Φ. It follows that ℓ has the asserted property (131). $\square$

Let $c_1, c_2 \in F^\times$. Recall from (115) the definition of the character ψ_{c_1,c_2} of U. An irreducible, admissible representation (π, V) of $\mathrm{GSp}(4, F)$ is called *generic* if it admits a nonzero functional $L : V \to \mathbb{C}$ satisfying

$$L(\pi(u)v) = \psi_{c_1,c_2}(u)L(v) \quad \text{for all } v \in V,\ u \in U. \tag{137}$$

Such an L is called a ψ_{c_1,c_2}-Whittaker functional.

The proof of (ii) of the following result has been sketched in Theorem 4.3 of [Piatetski-Shapiro 1997]; here, we provide the details.

Corollary 5.5.2. *There are no exceptional poles for π, Λ, β, μ if one of the following conditions is satisfied.*

(i) *The character $\Lambda_\mu = \Lambda \cdot (\mu \circ N_{L/F})$ is ramified.*

(ii) *π is generic.*

Hence, in these cases we have $L_\Lambda^{\mathrm{PS}}(s, \pi, \mu) = L_{\mathrm{reg}}^{\mathrm{PS}}(s, \pi, \mu)$, and in particular the Piatetski-Shapiro L-factor is independent of the choice of Bessel model for π.

Proof. (i) This is immediate from Lemma 5.3.2(i).

(ii) Let (π, V) be an irreducible, admissible, generic representation of $\mathrm{GSp}(4, F)$. Let $(\pi^\vee, V^\vee)$ be the contragredient representation. Then $\pi^\vee$ is also generic. Let L be a ψ_{c_1,c_2}-Whittaker functional on $V^\vee$.

Assume that π admits an exceptional pole; we will obtain a contradiction. By Theorem 5.5.1, there exists a character β of G and a functional $\ell : V \to \mathbb{C}$ such that

$$\ell(\pi(g)v) = \beta(g)v \tag{138}$$

for all $v \in V$ and $g \in G$. We define a linear map

$$\Delta : \mathcal{S}(\mathrm{GSp}(4, F)) \to V^\vee \tag{139}$$

by

$$\Delta(\phi)(v) = \int\limits_{\mathrm{GSp}(4,F)} \phi(g)\ell(\pi(g)v)\,dg, \tag{140}$$

where $\phi \in \mathcal{S}(\mathrm{GSp}(4, F))$, $v \in V$, and ℓ is a functional as in (131). Since ℓ is nonzero, it is easy to see that Δ is nonzero. One readily verifies that

$$\Delta(R_h\phi) = \pi^\vee(h)\Delta(\phi) \quad \text{for all } h \in \mathrm{GSp}(4, F). \tag{141}$$

In particular, the image of Δ is an invariant subspace of $V^\vee$. Consequently, Δ is surjective. This allows us to define a nonzero distribution $D : \mathcal{S}(\mathrm{GSp}(4, F)) \to \mathbb{C}$ by

$$D(\phi) = L(\Delta(\phi)), \quad \phi \in \mathcal{S}(\mathrm{GSp}(4, F)). \tag{142}$$

Since L is a ψ_{c_1,c_2}-Whittaker functional on $V^\vee$, it follows from (141) that

$$D(R_u\phi) = \psi_{c_1,c_2}(u)D(\phi) \quad \text{for all } u \in U. \tag{143}$$

For $h \in G$, we have

$$\begin{aligned}\Delta(L_h\phi)(v) &= \int\limits_{\mathrm{GSp}(4,F)} \phi(h^{-1}g)\ell(\pi(g)v)\,dg \\ &= \int\limits_{\mathrm{GSp}(4,F)} \phi(g)\ell(\pi(hg)v)\,dg \\ &= \beta(h) \int\limits_{\mathrm{GSp}(4,F)} \phi(g)\ell(\pi(g)v)\,dg\end{aligned}$$

by (138). Hence $\Delta(L_h\phi) = \beta(h)\Delta(\phi)$, and thus

$$D(L_h\phi) = \beta(h)D(\phi) \quad \text{for all } h \in G. \tag{144}$$

By Lemma 5.4.3, properties (143) and (144) imply that $D = 0$, a contradiction. □

References

[Aizenbud et al. 2010] A. Aizenbud, D. Gourevitch, S. Rallis, and G. Schiffmann, “Multiplicity one theorems”, *Ann. of Math.* (2) **172**:2 (2010), 1407–1434. MR

[Bernstein and Zelevinskii 1976] I. N. Bernstein and A. V. Zelevinskii, “Representations of the group $GL(n, F)$, where F is a local non-Archimedean field”, *Uspehi Mat. Nauk* **31**:3 (1976), 5–70. In Russian; translated in *Russian Math. Surveys* **31**:3 (1976), 1–68. MR

[Bump 1997] D. Bump, *Automorphic forms and representations*, Cambridge Studies in Advanced Mathematics **55**, Cambridge University Press, 1997. MR Zbl

[Danişman 2014] Y. Danişman, “Regular poles for the p-adic group GSp_4”, *Turkish J. Math.* **38**:4 (2014), 587–613. MR Zbl

[Danişman 2015a] Y. Danişman, “Regular poles for the p-adic group GSp_4-II”, *Turkish J. Math.* **39**:3 (2015), 369–394. MR Zbl

[Danişman 2015b] Y. Danişman, “Local factors of nongeneric supercuspidal representations of GSp_4”, *Math. Ann.* **361**:3-4 (2015), 1073–1121. MR Zbl

[Novodvorsky 1979] M. E. Novodvorsky, “Automorphic L-functions for symplectic group GSp(4)”, pp. 87–95 in *Automorphic forms, representations and L-functions* (Corvallis, Ore., (1977)), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**:2, Amer. Math. Soc., Providence, R.I., 1979. MR Zbl

[Piatetski-Shapiro 1997] I. I. Piatetski-Shapiro, “L-functions for GSp_4”, *Pacific J. Math.* Special Issue (1997), 259–275. MR

[Roberts and Schmidt 2007] B. Roberts and R. Schmidt, *Local newforms for GSp(4)*, Lecture Notes in Mathematics **1918**, Springer, 2007. MR Zbl

[Roberts and Schmidt 2016] B. Roberts and R. Schmidt, “Some results on Bessel functionals for GSp(4)”, *Doc. Math.* **21** (2016), 467–553. MR Zbl

[Saito 1993] H. Saito, “On Tunnell’s formula for characters of GL(2)”, *Compositio Math.* **85**:1 (1993), 99–108. MR Zbl

[Takloo-Bighash 2000] R. Takloo-Bighash, “L-functions for the p-adic group GSp(4)”, *Amer. J. Math.* **122**:6 (2000), 1085–1120. MR

[Tunnell 1983] J. B. Tunnell, “Local ϵ-factors and characters of GL(2)”, *Amer. J. Math.* **105**:6 (1983), 1277–1307. MR Zbl

[Waldspurger 1985] J.-L. Waldspurger, “Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie”, *Compositio Math.* **54**:2 (1985), 173–242. MR Zbl

Received February 22, 2017. Revised January 25, 2018.

RALF SCHMIDT
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OKLAHOMA
NORMAN, OK 73019-3103
UNITED STATES
rschmidt@math.ou.edu

LONG TRAN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OKLAHOMA
NORMAN, OK 73019-3103
UNITED STATES
ltran@math.ou.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.481

TORIC SURFACES OVER AN ARBITRARY FIELD

FEI XIE

We study toric varieties over an arbitrary field with an emphasis on toric surfaces in the Merkurjev–Panin motivic category of "K-motives". We explore the decomposition of certain toric varieties as K-motives into products of central simple algebras, the geometric and topological information encoded in these central simple algebras, and the relationship between the decomposition of the K-motives and the semiorthogonal decomposition of the derived categories. We obtain the information mentioned above for toric surfaces by explicitly classifying all minimal smooth projective toric surfaces using toric geometry.

1. Introduction

Throughout, we fix an arbitrary base field k. Let X be a scheme over k and let K/k be a field extension. We say a scheme Y over k is a *K/k-form* of X if the schemes $X_K := X \otimes_k K$ and Y_K are isomorphic as schemes over K [Serre 1997, Chapter III §1]. Let k^s be the separable closure of k. A k^s/k-form is simply called a *form* or *twisted form*. The scheme X_{k^s} has a natural $\Gamma = \mathrm{Gal}(k^s/k)$-action.

We will focus on the study of toric varieties over k. Let X be a normal geometrically irreducible variety over k and let T be an algebraic torus acting on X over k. The variety X is a toric T-variety if there is an open orbit U such that U is a principal homogeneous space or torsor over T. A toric T-variety is called split if the torus T is split. The case of split toric varieties have been extensively studied, for example in [Danilov 1978; Fulton 1993; Cox et al. 2011]. Since any toric variety X has a torus action over k and is a twisted form of a split toric variety, the study of X is equivalent to the study of the split toric variety X_{k^s} with a Γ-action on the fan structure as well as the study of the open orbit U; see Section 3.

Iskovskih [1979] classified minimal rational surfaces over arbitrary fields. Focusing on the cases of toric surfaces, we give an explicit description of minimal toric surfaces via toric geometry. In addition, the explicit nature of the classification of minimal toric surfaces made it possible for us to fully understand toric surfaces in

---

The author was supported in part by the NSF research grant DMS-#1160206.
MSC2010: primary 14J20, 14L30, 14M25; secondary 11E72, 16E35, 16H05.
Keywords: toric variety, motivic category, separable algebra, exceptional collection.

aspects such as affirming Merkurjev and Panin's question (Question 1) in dimension 2, decomposing toric surfaces as K-motives into products of central simple algebras, and providing full exceptional collections for the derived categories of toric surfaces, etc.

Theorem 4.12. *The surface X is a minimal smooth projective toric surface if and only if X is* (i) *a $\mathbb{P}^1$-bundle over a smooth conic curve but not a form of $F_1 = \mathrm{Proj}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(1))$*; (ii) *the Severi–Brauer surface*; (iii) *an involution surface*; (iv) *the del Pezzo surface of degree* 6 *with Picard rank* 1.

This paper is motivated by ideas in [Merkurjev and Panin 1997], which studies toric varieties over an arbitrary field in the motivic category $\mathcal{C}$ defined in loc. cit., and in particular by the following question:

Question 1. If X is a smooth projective toric variety over k, is $K_0(X_{k^s})$ always a permutation Γ-module?

Definition 1.1. A Γ-module M is a *permutation* Γ*-module* if there exists a Γ-invariant $\mathbb{Z}$-basis of M. We call such a basis a *permutation* Γ*-basis* or Γ*-basis*.

The reason that we care about the Γ-action on $K_0(X_{k^s})$ is that it in some way determines X; see Section 6. For example, if X has a rational point and $K_0(X_{k^s})$ is a permutation Γ-module, then X is isomorphic to the étale algebra corresponding to any Γ-basis of $K_0(X_{k^s})$ in the motivic category $\mathcal{C}$ [Merkurjev and Panin 1997, Proposition 4.5]. In general, if $K_0(X_{k^s})$ has a permutation Γ-basis of line bundles over X_{k^s}, then the variety X decomposes into a finite product of central simple algebras (over separable field extensions of k) in the motivic category $\mathcal{C}$ completely described by this Γ-basis as follows:

Theorem 6.5. *Let X be a smooth projective toric T-variety over k that splits over l and $G = \mathrm{Gal}(l/k)$. Assume $K_0(X_l)$ has a permutation G-basis P of line bundles on X_l. Let $\{P_i\}_{i=1}^t$ be G-orbits of P, and let $\pi : X_l \to X$ be the projection. For any $S_i \in P_i$, set $B_i = \mathrm{End}_{\mathcal{O}_X}(\pi_*(S_i))$ and $B = \prod_{i=1}^t B_i$. Then the map $u = \bigoplus_{i=1}^t \pi_*(S_i) : X \to B$ gives an isomorphism in the motivic category $\mathcal{C}$.*

Using the classification of minimal toric surfaces, we obtain that any smooth projective toric surface satisfies the conditions of the above theorem:

Theorem 5.2. *Let X be a smooth projective toric T-surface over k that splits over l and $G = \mathrm{Gal}(l/k)$. Then $K_0(X_l)$ has a permutation G-basis of line bundles on X_l.*

The original motivation for finding the decomposition of a smooth projective variety over k into a product of central simple algebras in $\mathcal{C}$ is to compute higher algebraic K-theory of the variety. Quillen [1973] computed higher algebraic K-theory for Severi–Brauer varieties; see Example 3.5, and Swan [1985] for quadric hypersurfaces. Panin [1994] generalized their results by finding the decomposition in $\mathcal{C}$ for twisted flag varieties.

As a matter of fact, these central simple algebras also encode arithmetic/geometric information about the variety, and in nice cases, classify its twisted forms. Blunk [2010] investigated del Pezzo surfaces of degree 6 over k in this direction; see Example 3.6. He showed that a del Pezzo surface of degree 6 is determined by a pair of Azumaya algebras (over étale quadratic and cubic extensions of the base field, respectively) and the surface has a rational point if and only if both Azumaya algebras in the pair are split. We will investigate the same information for all smooth projective toric surfaces over k; see Section 7. For example, we obtain that a $\mathbb{P}^1$-bundle over a smooth conic curve is isomorphic to $k \times Q \times k \times Q$ in $\mathcal{C}$ and the surface is determined by the quaternion algebra Q corresponding to the conic curve. More generally, if the Picard group $\operatorname{Pic}(X_{k^s})$ of a smooth projective toric variety X is a permutation Γ-module, then the open orbit U is determined by a set of central simple algebras, each corresponding to a Γ-orbit of $\operatorname{Pic}(X_{k^s})$; see Corollary 7.3. This implies that the toric variety X has a rational point if and only if every central simple algebra in the set is split.

Moreover, since Tabuada [2014, Theorem 6.10] showed that the motivic category $\mathcal{C}$ is a part of the category of noncommutative motives Hmo_0, it implies that certain semiorthogonal decompositions of the derived category of a smooth projective variety will give a decomposition of the variety in $\mathcal{C}$ (Theorem 8.4).

We will briefly discuss the possibility of lifting the motivic decomposition of a smooth projective toric variety to the derived category; see Section 8.

By the classification of minimal toric surfaces and known results of semiorthogonal decomposition of rational surfaces, we can confirm the lifting for smooth projective toric surfaces.

Theorem 8.6. *Let X be a smooth projective toric surface over k that splits over l and $G = \operatorname{Gal}(l/k)$. Then $K_0(X_l)$ has a permutation G-basis P of line bundles over X_l such that each G-orbit is an exceptional block. Furthermore, there exists an ordering of the G-orbits $\{P_i\}_{i=1}^t$ of P such that $\{P_1, \dots, P_t\}$ gives a full exceptional collection of $D^b(X_l)$. Therefore, for any $S_i \in P_i$, $\{\pi_* S_1, \dots, \pi_* S_t\}$ is a full exceptional collection of $D^b(X)$, where $\pi : X_l \to X$ is the projection.*

Organization. The organization of the paper is as follows: Sections 2 and 3 introduce the background on the motivic category $\mathcal{C}$ and toric varieties over k, including some basic facts and examples needed for the paper. For more details about $\mathcal{C}$, see [Merkurjev and Panin 1997, §1] or [Merkurjev 2005, §3]. Section 4 classifies minimal smooth projective toric surfaces over k via toric geometry. Section 5 verifies that $K_0(X_{k^s})$ has a permutation Γ-basis of line bundles for toric surfaces. In Section 6, we consider smooth projective toric varieties X of all dimensions where

$K_0(X_{k^s})$ has a permutation Γ-basis of line bundles. We decompose such X into a product of central simple algebras in the motivic category by reinterpreting the construction of the separable algebra corresponding to a toric variety investigated in [Merkurjev and Panin 1997]. In Section 7, we apply the construction in §6 to toric surfaces. Moreover, we relate the constructed algebras to the open orbit U via Galois cohomology. For details on Galois cohomology, see [Serre 1997; Knus et al. 1998; Gille and Szamuely 2006]. In Section 8, we discuss the relationship between the semiorthogonal decomposition of the derived category and the motivic decomposition of toric varieties via noncommutative motives and descent theory for derived categories.

Most of the time, instead of working with X_{k^s} and Γ-action, we work with X_l and $G = \mathrm{Gal}(l/k)$-action where l is the splitting field of the torus T.

Notation. Fix the base field k and a separable closure k^s of k. Let $\Gamma = \mathrm{Gal}(k^s/k)$. Let T denote an algebraic torus over k with splitting field l and $G = \mathrm{Gal}(l/k)$ unless otherwise stated. For any object Z (algebraic groups, varieties, algebras, maps) over k and any extension K/k, write $Z \otimes_k K$ as Z_K.

For a split toric variety Y, we denote Σ the fan structure and Aut_Σ the group of fan automorphisms. We will freely use the same notation for the ray in the fan, the minimal generator of the ray in the lattice and the Weil divisor corresponding to the ray when the context is clear.

For an algebra A, denote A^{op} its opposite algebra. Denote S_n the permutation group of a set of n elements.

2. The motivic category $\mathcal{C}$

Definition 2.1. The *motivic category* $\mathcal{C} = \mathcal{C}_k$ over a field k has:

- objects: pairs (X, A) where X is a smooth projective variety over k and A is a finite separable k-algebra,
- morphisms: $\mathrm{Hom}_{\mathcal{C}}((X, A), (Y, B)) = K_0(X \times Y, A^{\mathrm{op}} \otimes_k B)$.

The Grothendieck group K_0 of a pair is defined below. A k-algebra A is *finite separable* if $\dim_k(A)$ is finite and for any field extension K of k, the K-algebra A_K is semisimple. Equivalently we have:

Definition 2.2. The algebra A is a *finite separable k-algebra* if it is a finite product of central simple l_i-algebras A_i where l_i is a finite separable field extension of k, i.e, A_i is a matrix algebra over a finite dimensional division algebra with center l_i.

Let $u : (X, A) \to (Y, B)$ and $v : (Y, B) \to (Z, C)$ be morphisms in $\mathcal{C}$. Since $u \in K_0(X \times Y, A^{\mathrm{op}} \otimes_k B) \cong K_0(Y \times X, B \otimes_k A^{\mathrm{op}})$, the map u can also be viewed as $u^{\mathrm{op}} : (Y, B^{\mathrm{op}}) \to (X, A^{\mathrm{op}})$. The *composition* $v \circ u : (X, A) \to (Z, C)$ is given by

$$\pi_*(q^* v \otimes_B p^* u),$$

where $p : X \times Y \times Z \to X \times Y$, $q : X \times Y \times Z \to Y \times Z$, $\pi : X \times Y \times Z \to X \times Z$ are projections.

We write X for (X, k) and A for $(\operatorname{Spec} k, A)$. Since the morphisms are defined in K_0, the category is also called the *category of K-correspondences*.

Algebraic K-theory of a pair. The algebraic K-theory of a pair (X, A) is defined in the following way and it generalizes the Quillen K-theory of varieties:

Let $\mathcal{P}(X, A)$ be the exact category of left $\mathcal{O}_X \otimes_k A$-modules which are locally free $\mathcal{O}_X$-modules of finite rank and morphisms of $\mathcal{O}_X \otimes_k A$-modules. The group $K_n(X, A)$ of the pair (X, A) is defined as $K_n^Q(\mathcal{P}(X, A))$, the Quillen K-theory of $\mathcal{P}$. Let $\mathcal{M}(X, A)$ be the exact category of left $\mathcal{O}_X \otimes_k A$-modules which are coherent $\mathcal{O}_X$-modules and morphisms of $\mathcal{O}_X \otimes_k A$-modules. The group $K'_n(X, A)$ of the pair (X, A) is defined as $K_n^Q(\mathcal{M}(X, A))$. The embedding $\mathcal{P} \subset \mathcal{M}$ induces a map $K_n(X, A) \to K'_n(X, A)$ and it is an isomorphism if X is regular (resolution theorem). Note that $K_n(X, k)$ is the usual $K_n(X)$ and $K_n(\operatorname{Spec} k, A) = K_n(\operatorname{Rep}(A))$ is the K-theory of representations of A.

In fact, K_n defines a functor $K_n : \mathcal{C} \to \mathrm{Ab}$ which sends (X, A) to $K_n(X, A)$. For $u : (X, A) \to (Y, B)$, $x \in K_n(X, A)$, we can define

$$K_n(u)(x) = q_*(u \otimes_A p^* x),$$

where $p : X \times Y \to X$, $q : X \times Y \to Y$ are projections.

Similarly we can define, for any variety V over k, a functor $K_n^V : \mathcal{C} \to \mathrm{Ab}$ where on objects $K_n^V(X, A) = K'_n(V \times X, A)$.

Example 2.3 [Merkurjev and Panin 1997, Example 1.6(1)]. $\mathrm{M}_n(k) \cong k$ in $\mathcal{C}$.

Example 2.4 [Merkurjev and Panin 1997, Example 1.6(3)], see also [Tabuada 2014, Theorem 9.1]. Let A and B be two central simple k-algebras. Then $A \cong B$ in $\mathcal{C}$ if and only if $[A] = [B] \in \mathrm{Br}(k)$.

Proof. The previous example indicates that Brauer equivalences give isomorphisms in $\mathcal{C}$. So $[A] = [B] \in \mathrm{Br}(k)$ implies $A \cong B$ in $\mathcal{C}$.

For the opposite direction, since each central simple k-algebra is Brauer equivalent to a unique division k-algebra, we can assume A, B are division algebras. Let $M : A \to B$ and $N : B \to A$ be inverse maps in $\mathcal{C}$. Since $K_0(A^{\mathrm{op}} \otimes_k B) \cong \mathbb{Z}R$ and $K_0(B^{\mathrm{op}} \otimes_k A) \cong \mathbb{Z}R^{\mathrm{op}}$ for R the unique simple B-A-bimodule, we have $M = nR$ and $N = mR^{\mathrm{op}}$ for some $m, n \in \mathbb{Z}$. $N \circ M = N \otimes_B M \cong mnR^{\mathrm{op}} \otimes_B R \cong A$, $M \circ N = M \otimes_A N \cong mnR \otimes_A R^{\mathrm{op}} \cong B$. Since A, B are simple modules, we have $mn = 1$ and we can assume $M = R$, $N = R^{\mathrm{op}}$. As a right A-module and a left B-module respectively, we have $M_A \cong A^r$ and ${}_BM \cong B^s$. Similarly, ${}_AN \cong A^p$ and $N_B \cong B^q$. The left A-module isomorphism $N \otimes_B M \cong N \otimes_B B^s \cong N^s \cong A^{ps} \cong A$ implies that $p = s = 1$. Similarly $r = q = 1$. In particular, this implies $\dim_k A = \dim_k B$.

Finally consider the k-algebra homomorphism $f : B \to \mathrm{End}_A(M_A) \cong A$ by sending b to l_b left multiplication by b. This is obviously injective, and it is surjective because A, B have the same dimension, so $A \cong B$ as k-algebras. $\square$

3. Toric varieties

Let T be an algebraic torus over k.

Definition 3.1. A *toric T-variety X* over k is a normal geometrically irreducible variety with an action of the torus T and an open orbit U which is a principal homogeneous space over T.

By definition, the torus $T_{k^s} \cong \mathbb{G}^n_{m,k^s}$ splits where $n = \dim X$. The torus T corresponds to a cocycle class $[\rho] \in H^1(\Gamma, \mathrm{Aut}_{\mathrm{gp},k^s}(\mathbb{G}^n_{m,k^s})) = H^1(\Gamma, \mathrm{GL}(n, \mathbb{Z}))$ where $\mathrm{Aut}_{\mathrm{gp},k^s}$ denotes the group automorphism over k^s. Moreover, the torus T splits over a finite Galois extension l of k ($T_l \cong \mathbb{G}^n_{m,l}$), which is called the *splitting field* of T.

Explicitly, tori $T_{k^s} = T \otimes_k k^s$ and $\mathbb{G}^n_{m,k^s} = \mathbb{G}_{m,k} \otimes_k k^s$ have natural Galois actions with Γ acting on the factor k^s. The Galois actions give group automorphisms of T_{k^s} and $\mathbb{G}^n_{m,k^s}$ over k, but not over k^s because Γ also acts on the scalars k^s. Let $\sigma : \Gamma \to \mathrm{Aut}_k(T_{k^s})$ and $\tau : \Gamma \to \mathrm{Aut}_k(\mathbb{G}^n_{m,k^s})$ be the respective natural Galois actions. Let $\phi : T_{k^s} \to \mathbb{G}^n_{m,k^s}$ be an isomorphism. Then we obtain $\rho : \Gamma \to \mathrm{GL}(n, \mathbb{Z})$ by sending g to $\phi\sigma(g)\phi^{-1}\tau(g)^{-1}$, and we have $\ker(\rho) = \mathrm{Gal}(k^s/l)$ where l is the splitting field.

Conversely, the torus T can be constructed from $\rho : \Gamma \to \mathrm{GL}(n, \mathbb{Z})$ as follows; see also [Voskresenskii 1982, §1]. The map ρ factors through $\rho' : G = \mathrm{Gal}(l/k) \to \mathrm{GL}(n, \mathbb{Z})$ for a finite Galois extension l of k. Let $\mu : G \to \mathrm{Aut}_k(\mathbb{G}^n_{m,l})$ be the action on the torus $\mathbb{G}^n_{m,k} \otimes_k l$ via $\mu(g) = \rho'(g) \otimes g,\ g \in G$. Then $T \cong \mathbb{G}^n_{m,l}/\mu(G)$.

Definition 3.2. A toric T-variety X over k is called a *toric T-model* if $U(k)$ is nonempty.

In this case, the open orbit $U \cong T$ as k-varieties and there is an T-equivariant embedding $T \hookrightarrow X$. If X is smooth over k, then the set $X(k)$ is nonempty if and only if $U(k)$ is [Voskresenskii and Klyachko 1985, §4 Proposition 4].

Definition 3.3. A toric T-variety is *split* if T splits, and is *nonsplit* otherwise.

Let X_{k^s} (or X_l) be the split toric variety with the fan structure Σ. Since the Γ-action on T_{k^s} is compatible with the one on X_{k^s}, the image of ρ is contained in Aut_Σ, namely

$$\rho(\Gamma) = \mathrm{Gal}(l/k) \subseteq \mathrm{Aut}_\Sigma \subset \mathrm{GL}(n, \mathbb{Z}).$$

Let X_Σ be the split toric variety over k with the fan structure Σ. If X is a toric T-model, then similarly to the case of the torus T, the variety X can be recovered from ρ and Σ as $(X_\Sigma \otimes_k l)/\mu(G)$. In general, for each toric T-variety X, there is a unique (up to T-isomorphism) toric T-model X^* such that $X_{k^s} \cong (X^*)_{k^s}$. We call X^* the associated toric T-model of X. More specifically, the toric T-model X^* is given by $(X \times U)/T$ where T acts on $X \times U$ diagonally, and the toric T-variety X is given by $(X^* \times U)/T$ where T acts on $X^* \times U$ via $t \cdot (x, y) = (tx, yt^{-1})$; see [Voskresenskii and Klyachko 1985, §4].

In summary, an algebraic torus T is uniquely determined by a 1-cocycle (class) $\rho : \Gamma \to \mathrm{GL}(n, \mathbb{Z})$. A toric T-model X is uniquely determined by ρ and fan Σ with the restriction $\rho(\Gamma) \subseteq \mathrm{Aut}_\Sigma$. A toric T-variety is uniquely determined by its associated T-model X^* and a principal homogeneous space $U \in H^1(k, T)$.

Lemma 3.4. *Let $\phi : X_{\Sigma_1} \to X_{\Sigma_2}$ be a toric morphism of split smooth projective toric varieties over k^s, and let $\bar{\phi} : N_1 \to N_2$ be the induced $\mathbb{Z}$-linear map of lattices that is compatible with fans Σ_1, Σ_2. Let $\rho_i : \Gamma \to \mathrm{Aut}(N_i)$ be Galois actions on N_i that are compatible with the fans Σ_i ($\rho_i(\Gamma) \subseteq \mathrm{Aut}_{\Sigma_i}$) such that $\bar{\phi}$ is Γ-equivariant with respect to ρ_1, ρ_2. Let T_i be the torus corresponding to ρ_i. Then, for any $U_1 \in H^1(k, T_1)$, there exists $U_2 \in H^1(k, T_2)$ such that ϕ descends to a map $X_1 \to X_2$, where X_i is the toric variety corresponding to (ρ_i, Σ_i, U_i) for $i = 1, 2$.*

Proof. Restrict ϕ to tori $\phi|_{T_{N_1}} : T_{N_1} \to T_{N_2}$. Since $\bar{\phi}$ is Γ-equivariant, the maps ϕ and $\phi|_{T_{N_1}}$ descend to $\varphi : X_1^* \to X_2^*$ where X_i^* are the toric T_i-models corresponding to Σ_i and $\psi : T_1 \to T_2$. The map ψ induces $H^1(k, T_1) \to H^1(k, T_2)$ and let U_2 be the image of U_1 under this map. Set $X_i = (X_i^* \times U_i)/T_i$. Then ϕ descends to a map $X_1 \to X_2$. □

Example 3.5 (Severi–Brauer variety X ($X_{k^s} \cong \mathbb{P}^n$)). Let A be a central simple k-algebra of degree $n+1$. Then $X = \mathrm{SB}(A)$ is a toric variety with the torus $T = \mathrm{R}_{E/k}(\mathbb{G}_{m,E})/\mathbb{G}_{m,k}$, where E is a maximal étale k-subalgebra of A. The variety X has a rational point if and only if $A = M_{n+1}(k)$ if and only if $X \cong \mathbb{P}^n$.

Quillen [1973, §8 Theorem 4.1] showed that $K_m(\mathrm{SB}(A)) \cong K_m(k) \times \prod K_m(A^{\otimes i})$ for $m \geqslant 0$ and Panin [1994] showed that $\mathrm{SB}(A) \cong k \times \prod A^{\otimes i}$ in $\mathcal{C}$, where the products run over $i = 1, \dots, n$.

Example 3.6. Let X be a del Pezzo surface of degree 6 over k (K_X is antiample with $K_X^2 = 6$, $X_{k^s} \cong \mathrm{Bl}_{p_1,p_2,p_3}(\mathbb{P}^2)$ where p_1, p_2, p_3 are not collinear). It is a toric T-variety where the torus T is the connected component of the identity of $\mathrm{Aut}_k(X)$.

Blunk [2010] showed that $X \cong k \times P \times Q$ in $\mathcal{C}$ where P is an Azumaya K-algebra of rank 9 ($\dim_k(P)/\dim_k(K) = 9$) and Q is an Azumaya L-algebra of rank 4 where K, L are étale k-algebras of degree 2 and 3, respectively.

Example 3.7 (Involution surface X ($X_{k^s} \cong \mathbb{P}^1 \times \mathbb{P}^1$)). The surface X corresponds to a central simple k-algebra A of degree 4 together with a quadratic pair (σ, f) on A. For the definition of a quadratic pair, see [Knus et al. 1998, §5B]. The associated even Clifford algebra $C_0(A, \sigma, f)$ (defined in their §8B) is a quaternion algebra over K, which is an étale quadratic extension of k and is called the *discriminant extension* of X. Write $B = C_0(A, \sigma, f)$. Then X is the Weil restriction $\mathrm{R}_{K/k}\,\mathrm{SB}(B)$; see [Auel and Bernardara 2015, Example 3.3]. Denote by T the torus of $\mathrm{SB}(B)$ in Example 3.5. Then X is a toric variety with the torus $\mathrm{R}_{K/k}\,T$.

Panin [1994] showed that $X \cong k \times B \times A$ in $\mathcal{C}$.

K_0 of split toric varieties. Let Y be a split smooth proper toric T-variety with fan Σ.

For $\sigma \in \Sigma$, denote $\mathcal{O}_\sigma$ the closure of the T-orbit corresponding to σ and J_σ the sheaf of ideals defining $\mathcal{O}_\sigma$. Write $\sigma(1)$ for the set of rays spanning σ. For $\sigma, \tau \in \Sigma$, if $\sigma(1) \cap \tau(1) = \varnothing$ and $\sigma(1) \cup \tau(1)$ span a cone in Σ, then denote the cone by $\langle \sigma, \tau \rangle$, otherwise set $\langle \sigma, \tau \rangle = 0$.

Theorem 3.8 (Klyachko [1992]; Demazure). *As an abelian group, $K_0(Y)$ is generated by $\mathcal{O}_\sigma = 1 - J_\sigma$ with these relations*:

$$(1) \quad \mathcal{O}_\sigma \cdot \mathcal{O}_\tau = \begin{cases} \mathcal{O}_{\langle \sigma, \tau \rangle} & \textit{if } \langle \sigma, \tau \rangle \neq 0, \\ 0 & \textit{otherwise}; \end{cases}$$

$$(2) \quad \prod_{e \in \Sigma(1)} J_e^{f(e)} = 1, \qquad f \in \mathrm{Hom}(N, \mathbb{Z}) = M \ (\textit{the group of characters of } T).$$

Theorem 3.9 (Klyachko). *The abelian group $K_0(Y)$ is free with rank equal to the number of the maximal cones. In addition, sheaves $\mathcal{O}_y$ and $\mathcal{O}_{y'}$ coincide in $K_0(Y)$ for any rational closed points $y, y' \in Y$.*

4. Minimal toric surfaces

Let X be a smooth projective toric surface over k. We say X is *minimal* if any birational morphism $f : X \to X'$ from X to another smooth surface X' defined over k is an isomorphism. In this section, we will classify minimal smooth projective toric surfaces.

First we notice that the exceptional locus of any birational morphism from a toric surface is torus invariant. We use the convention that a surface is integral, separated and of finite type.

Lemma 4.1. *Let W be a smooth projective toric T-surface over k. Let $h : W \to Z$ be a birational morphism over k from W to a smooth surface Z over k. Let E be the exceptional divisor of h. Then E is T-invariant. Therefore, the surface Z is a smooth projective toric T-surface and the map h is T-invariant.*

Proof. First assume that k is separably closed. Then W is split. Since for a split toric variety the group of T-invariant Cartier divisors CDiv_T maps onto the Picard group, the line bundle $\mathcal{O}(E)$ is fixed by the T-action. For any $t \in T$, the divisor tE is linearly equivalent to E (denoted $tE \sim E$).

Now assume the locus E is not T-invariant and let $t_0 \in T$ be such that $t_0E \neq E$. Note that since W is proper and Z is separated, the map h is proper and the surface $Z = h(W)$ is also proper (thus projective). We have $p(t_0E) \sim p(E) = 0$. Let $C = p(t_0E)$ which is a curve on Z. Embed Z into some $\mathbb{P}^n$ and let H be a hyperplane of $\mathbb{P}^n$. Since C is a curve, we have $C.H > 0$. Therefore, C cannot be linearly equivalent to 0, a contradiction.

For an arbitrary field k, we base change to the separable closure k^s and use the same argument. □

Lemma 4.2. *Let X be a smooth projective toric T-surface over k. Then X is minimal if and only if X_{k^s} admits no Γ-invariant set of pairwise disjoint T_{k^s}-invariant (-1)-curves.*

Proof. Since any (-1)-curve is the exceptional locus of some birational morphism, by the previous lemma, it is always torus invariant. The rest follows from [Hassett 2009, Theorem 3.2]. □

Definition 4.3. Let Y be a split smooth projective toric surface over a field K. If there is a finite group G acting on Y by K-automorphisms, we call Y a *G-surface* over K. The G-surface Y is called *G-minimal* over K if Y admits no G-invariant set of pairwise disjoint torus invariant (-1)-curves.

Lemma 4.2 implies that we can redefine minimal toric surfaces as follows:

Definition 4.4. Let X be a smooth projective toric T-surface over k and let $\rho : \Gamma \to \mathrm{GL}(2, \mathbb{Z})$ be the map corresponding to the torus T. Let $G = \rho(\Gamma)$, which is a finite subgroup of $\mathrm{GL}(2, \mathbb{Z})$ and acts on the split toric surface X_{k^s} by fan automorphisms ($G \subseteq \mathrm{Aut}_\Sigma(X_{k^s})$). We say the toric surface X is *minimal* if X_{k^s} is G-minimal over k^s.

Proposition 4.5. *Let X and $G = \rho(\Gamma)$ be the same as above. Then there is a finite chain of blowups of toric T-surfaces*

$$X = X_0 \xrightarrow{f_1} X_1 \xrightarrow{f_2} \cdots \xrightarrow{f_n} X_n = X',$$

where each X_i is a smooth projective toric T-surface, each map f_i is the blowup of X_i along T-invariant reduced zero-dimensional subscheme (*in particular, f_i is T-invariant*) *and X' is minimal.*

Proof. If X is not minimal, then X_{k^s} admits a G-invariant set of pairwise disjoint T_{k^s}-invariant (-1)-curves. Contracting this G-set of (-1)-curves and descending the contraction map to the base field k, we get a map $f_1 : X \to X_1$ which is the

cyclic	dihedral	generators
$C_1 = \langle I \rangle$	$D_2 = \langle C \rangle$ $D_2' = \langle C' \rangle$	$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$
$C_2 = \langle -I \rangle$	$D_4 = \langle -I, C \rangle$ $D_4' = \langle -I, C' \rangle$	$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
$C_3 = \langle A^2 \rangle$	$D_6 = \langle A^2, C \rangle$ $D_6' = \langle A^2, -C \rangle$	$C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$C_4 = \langle B \rangle$ $C_6 = \langle A \rangle$	$D_8 = \langle B, C \rangle$ $D_{12} = \langle A, C \rangle$	$C' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Table 1. Nonconjugate classes of finite subgroups of $\mathrm{GL}(2, \mathbb{Z})$ and their generators.

blowup of a smooth projective toric T-surface X_1 along T-invariant reduced zero-dimensional subscheme. This process will terminate in finite steps because the number of rays in the fan of $(X_1)_{k^s}$ is strictly less than that of X_{k^s}. □

Now, classifying all minimal smooth projective toric surfaces over k is the same as classifying, for each finite subgroup G of $\mathrm{GL}(2, \mathbb{Z})$ (up to conjugacy), G-minimal toric surfaces over k^s. It is well known that when G is trivial, the minimal (toric) surfaces are $\mathbb{P}^2$ and Hirzebruch surfaces $F_a = \mathrm{Proj}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(a))$ for $a \geqslant 0$, $a \neq 1$.

There are 13 nonconjugate classes of finite subgroups of $\mathrm{GL}(2, \mathbb{Z})$ and they can only be either cyclic or dihedral groups [Newman 1972, Chapter IX, §14]. See Table 1.

Definition 4.6. Let Y be a split smooth projective toric surface with fan structure Σ. Counterclockwise label the rays of Σ as $y_1, \ldots, y_n$ and denote by D_i the divisor corresponding to y_i. We can assign a sequence $a = (a_1, \ldots, a_n)$ to Y, where $a_i = D_i^2$. We refer to this sequence as the *self-intersection sequence* of Y.

The group of fan automorphisms $\mathrm{Aut}_\Sigma(Y)$ acts on $\mathbb{Z}^2$, permuting rays y_i of the fan Σ. First observe that as automorphisms of Y, the group $\mathrm{Aut}_\Sigma(Y)$ preserves the self-intersection number of any divisor and thus permutes (torus invariant) (-1)-curves on Y. Now, let us consider the case where $\mathrm{Aut}_\Sigma(Y) \cap \mathrm{SL}(2, \mathbb{Z}) = C_t$ is nontrivial and look at the action of C_t on the rays. As indicated in Table 1, the cyclic group C_t is generated by powers of A or B where B is the rotation by $\pi/4$ and A is conjugate in $\mathrm{GL}(2, \mathbb{R})$ to the rotation by $\pi/3$. In particular, the action of C_t on the fan Σ is free, which implies $t \mid n$.

Lemma 4.7. *Let* $\mathrm{Aut}_\Sigma(Y) \cap \mathrm{SL}(2, \mathbb{Z}) = C_t$ *be nontrivial* (*i.e.*, $t = 2, 3, 4, 6$). *If the number of rays of the fan* $> \max\{4, t\}$, *then* Y *is not* C_t*-minimal*, *that is*, *there exists* *a C_t-invariant set of pairwise disjoint (-1)-curves on Y. Therefore, C_t-minimal surfaces have the number of rays $\leqslant \max\{4, t\}$.*

Proof. Denote counterclockwise $y_1, \ldots, y_n$ as rays of Σ and let $a = (a_1, \ldots, a_n)$ be its self-intersection sequence. If $n > 4$, Y is not $\mathbb{P}^2$ or F_a, then there exists i such that $a_i = -1$. Let σ be a generator of C_t and as discussed above, σ rotates the rays. If $n > t$, then the ray $\sigma(y_i)$ is not adjacent to y_i (i.e., corresponding divisors are disjoint) and thus $\{y_i, \sigma(y_i), \ldots, \sigma^{t-1}(y_i)\}$ form a C_t-invariant set of pairwise disjoint (-1)-curves. □

Lemma 4.8. *D_2 fixes rays generated by $\pm(1, 1)$ or maximal cones generated by $(1, 0)$ and $(0, 1)$ or by $(-1, 0)$ and $(0, -1)$; D_2' fixes rays generated by $\pm(1, 0)$.*

Using toric geometry, Oda showed [1978, Theorem 8.2] that a split smooth projective toric surface is a succession of blowups of $\mathbb{P}^2$ or F_a. The proof of the theorem is essentially the following lemma:

Lemma 4.9. *Let Y be a split smooth projective toric surface with the fan Σ. Let x, y be two rays in Σ where their minimal generators form a basis of $\mathbb{Z}^2$. If x, y are not adjacent in the fan, then there is a ray $z \in \Sigma$ between x, y corresponding to a (-1)-curve.*

Now we are ready to classify G-minimal toric surfaces for G a finite subgroup of $\mathrm{GL}(2, \mathbb{Z})$.

Proposition 4.10. *Let Y be a split smooth projective toric surface and let G be a finite subgroup of $\mathrm{GL}(2, \mathbb{Z})$ acting on Y by fan automorphisms; that is, $G \subseteq \mathrm{Aut}_\Sigma(Y)$. Then the surface Y is G-minimal if and only if Y belongs to one of the following:*

- $G = D_2$: $Y = \mathbb{P}^2, \mathbb{P}^1 \times \mathbb{P}^1, F_{2a+1}, a \geqslant 1$;
- $G = D_2'$: $Y = F_{2a}, a \geqslant 0$;
- $G = C_2, C_4, D_4, D_4', D_8$: $Y = \mathbb{P}^1 \times \mathbb{P}^1$;
- $G = C_3, D_6$: $Y = \mathbb{P}^2$;
- $G = C_6, D_6', D_{12}$: $Y = S$,

where $F_a = \mathrm{Proj}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(a))$ is the Hirzebruch surface and S is the blowup $\mathrm{Bl}_{p_1,p_2,p_3}(\mathbb{P}^2)$ of $\mathbb{P}^2$ along three torus invariant points.

Proof. Assume the split toric surface Y is G-minimal. Let Σ be the fan structure of Y and let n be the number of rays of Σ. It is clear that for any subgroup H of G together with the restricted H-action on Y, the surface Y is either H-minimal or the (successive) blowups of H-minimal toric surfaces.

$G = D_2$: (I) If D_2 fixes at least one maximal cone, then Σ contains (I.1) rays $(1, 0)$, $(0, 1)$, $(-1, -1)$ where D_2 fixes the maximal cone generated by $(1, 0)$, $(0, 1)$ or (I.2) rays $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$ where D_2 fixes the maximal cones generated by

$(1,0)$, $(0,1)$ and by $(-1,0)$, $(0,-1)$. (II) Otherwise Σ contains rays $\pm(1,1)$, and the rays counterclockwise before and after $(1,1)$ must be $(a+1,a)$ and $(a,a+1)$, respectively. By Lemma 4.9, it is easy to see that if Σ contains more rays in any of the above cases, then Y admits a D_2-set of pairwise disjoint (-1)-curves. Thus, Y is isomorphic to (I.1) $\mathbb{P}^2$; (I.2) $\mathbb{P}^1 \times \mathbb{P}^1$; (II) F_{2a+1}. Since F_1 has a D_2-invariant (-1)-curve, it is not minimal. So we have $a \geqslant 1$.

$G = D_2'$: Σ contains rays $\pm(1,0)$, and the rays counterclockwise before and after $(1,0)$ must be $(a,-1)$ and $(a,1)$, respectively. By Lemma 4.9, Σ contains no other rays. Thus, Y is isomorphic to F_{2a}, $a \geqslant 0$.

$G = C_2$: Let $x, y \in \Sigma$ be two adjacent rays. Then Σ should have rays x, y, $-x$, $-y$, where the minimal generators of x, y form a basis of $\mathbb{Z}^2$ and by Lemma 4.9, it contains no other rays. Thus, $Y \cong \mathbb{P}^1 \times \mathbb{P}^1$.

$G = C_4, D_4, D_4', D_8$: Since C_2 is a subgroup of C_4, D_4, D_4', D_8, we have $Y \cong \mathbb{P}^1 \times \mathbb{P}^1$ or its blowups. Since the group of fan automorphisms of $\mathbb{P}^1 \times \mathbb{P}^1$ is D_8 which contains C_4, D_4, D_4', the minimal C_2-surface $\mathbb{P}^1 \times \mathbb{P}^1$ is already a G-surface for $G = C_4, D_4, D_4', D_8$ and must be G-minimal. Thus, $Y \cong \mathbb{P}^1 \times \mathbb{P}^1$.

For cases $G = C_t$, $t > 2$. Recall that $t \mid n$ and by Lemma 4.7, $n \leqslant \max\{4, t\}$.

$G = C_3$: $3 \mid n$, $n \leqslant 4$, so $n = 3$ and $Y \cong \mathbb{P}^2$.

$G = D_6$: $C_3 \subset D_6$ implies that Y is either $\mathbb{P}^2$ or its blowups. Since the group of fan automorphisms is D_6, we have $Y \cong \mathbb{P}^2$.

For cases $G \supseteq C_3$, observe that if Y is not $\mathbb{P}^2$, then it must be the blowup of S where S is the blowup of $\mathbb{P}^2$ along three torus invariant points.

$G = C_6, D_6', D_{12}$ $C_3 \subset D_6' \subset D_{12}$ and $C_3 \subset C_6 \subset D_{12}$ imply that Y is either $\mathbb{P}^2$ or the blowup of $\mathbb{P}^2$. Since the group of fan automorphisms of $\mathbb{P}^2$ is D_6, Y can not be $\mathbb{P}^2$. Thus, Y is either S or its blowup. We have $Y \cong S$ because the group of fan automorphisms of S is D_{12}. □

Lemma 4.11. *Let X be a toric surface that is a form of F_a, $a \geqslant 1$. Then X is a $\mathbb{P}^1$-bundle over a smooth conic curve. If X has a rational point, then $X \cong F_a$.*

Proof. Let X correspond to (ρ_1, Σ_1, U_1) and let Σ_1 be the fan of F_a with rays $(1,0)$, $(0,1)$, $(-1,a)$, $(0,-1)$. Let $\bar{\phi} : \mathbb{Z}^2 \to \mathbb{Z}$ be the projection to the first factor, which corresponds to $\phi : F_a \to \mathbb{P}^1$. Let $\rho_2 = \det \circ \rho_1 : \Gamma \to \mathrm{GL}(1, \mathbb{Z})$. Either ρ_1 is trivial or ρ_1 permutes the rays $(1,0)$, $(-1,a)$. Then $\bar{\phi}$ is Galois equivariant with respect to ρ_1 and ρ_2. By Lemma 3.4, the map ϕ descends to $\varphi : X \to C$. As a form of $\mathbb{P}^1$, C is a smooth plane conic curve ([Gille and Szamuely 2006, Corollary 5.4.8] for characteristic not 2 and [Elman et al. 2008, §45A] for any characteristic).

Let D be the divisor corresponding to the ray $(0,-1)$. Then D is a Galois invariant section of the bundle $\phi : F_a \to \mathbb{P}^1$. Thus, D descends to a section D' of $\varphi : X \to C$. Moreover, $F_a \cong \mathbb{P}(\phi_* \mathcal{O}_{F_a}(D))$ descends to $X \cong \mathbb{P}(\varphi_* \mathcal{O}_X(D'))$. Thus,

X is a $\mathbb{P}^1$-bundle over C. If X has a rational point, so does C. Therefore, $C \cong \mathbb{P}^1$ and $X \cong F_a$. □

By Proposition 4.10, a minimal smooth projective toric surface X is a form of (i) F_a, $a \geqslant 2$; (ii) $\mathbb{P}^2$; (iii) $\mathbb{P}^1 \times \mathbb{P}^1$; (iv) $\mathrm{Bl}_{p_1,p_2,p_3}(\mathbb{P}^2)$ where p_1, p_2, p_3 are not collinear. Furthermore, we have

Theorem 4.12. *The surface X is a minimal smooth projective toric surface if and only if X is* (i) *a $\mathbb{P}^1$-bundle over a smooth conic curve but not a form of* $F_1 = \mathrm{Proj}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(1))$; (ii) *the Severi–Brauer surface*; (iii) *an involution surface*; (iv) *the del Pezzo surface of degree* 6 *with Picard rank* 1.

Proof. It follows from Lemma 4.11, Examples 3.5, 3.6, 3.7 and the fact that a minimal del Pezzo surface of degree not equal to 8 has Picard rank 1 [Colliot-Thélène et al. 2008, Theorem 2.4]. □

5. K_0 of toric surfaces

In this section, we will show that $K_0(X_{k^s})$ is a permutation Γ-module for X a smooth projective toric surface over k. First recall how K_0 behaves under blowups:

Theorem 5.1 [SGA 6 1971, VII 3.7]. *Let X be a noetherian scheme and $i : Y \to X$ a regular closed immersion of pure codimension d. Let $p : X' \to X$ be the blow up of X along Y and $Y' = p^{-1}Y$. There is a split short exact sequence*

$$0 \to K_0(Y) \xrightarrow{u} K_0(Y') \oplus K_0(X) \xrightarrow{v} K_0(X') \to 0,$$

and the splitting w for u is given by $w(y', x) = p|_{Y'*}(y')$, $y' \in K(Y')$, $x \in K(X)$.

This gives us an isomorphism $K_0(X') \cong \ker(w) \cong K_0(X) \oplus \bigoplus^{d-1} K_0(Y)$ which fits into the split short exact sequence

$$0 \to K_0(X) \xrightarrow{p^*} K_0(X') \to \bigoplus^{d-1} K_0(Y) \to 0.$$

Now let X be a smooth projective toric T-surface over k that splits over l. Let Y be a T-invariant reduced zero-dimensional subscheme of X. Then Y_l is a disjoint union of T_l-invariant points permuted by $G = \mathrm{Gal}(l/k)$. Set $X' = \mathrm{Bl}_Y X$. We have

$$0 \to K_0(X_l) \xrightarrow{p^*} K_0(X'_l) \to K_0(Y_l) = \bigoplus \mathbb{Z} \to 0,$$

where p^* is a G-homomorphism. Each $\mathbb{Z}$ is generated by $\mathcal{O}_{E_i}(-1)$ where E_i are the exceptional divisors corresponding to the points in Y_l and G permutes E_i the same way as G permutes the points in Y_l.

Note that $\mathcal{O}_{E_i}(-1) = \mathcal{O}_{X'_l}(E_i) - \mathcal{O}_{X'_l}$ in K_0. If we know $K_0(X_l)$ has a permutation G-basis γ, then $K(X'_l)$ has a permutation G-basis consisting of $p^*\gamma$ (total transforms of γ) and the $\mathcal{O}(E_i)$.

Theorem 5.2. *Let X be a smooth projective toric T-surface over k that splits over l and $G = \mathrm{Gal}(l/k)$. Then $K_0(X_l)$ has a permutation G-basis of line bundles on X_l.*

Proof. By previous discussion and the fact that $G \subseteq \mathrm{Aut}_\Sigma$, it suffices to prove that $K_0(X_l)$ has a permutation Aut_Σ-basis of line bundles for X minimal. By Theorem 4.12, we only need to consider the following cases for X_l:

(i) $F_a, a \geqslant 2,\ \mathrm{Aut}_\Sigma = S_2$.

(ii) $\mathbb{P}^2$, $\mathrm{Aut}_\Sigma = D_6$.

(iii) $\mathbb{P}^1 \times \mathbb{P}^1$, $\mathrm{Aut}_\Sigma = D_8$.

(iv) del Pezzo surface of degree 6, $\mathrm{Aut}_\Sigma = D_{12}$.

We will use equation (2) in Theorem 3.8 with $f = (1, 0)$ and $(0, 1)$ in producing relations and finding a permutation basis. We will write x_i for rays in the fan and $J_i = \mathcal{O}(-D_i)$ where D_i are the divisors corresponding to x_i.

(i) Rays $x_1 = (1, 0),\ x_2 = (0, 1),\ x_3 = (-1, a),\ x_4 = (0, -1)$: Then S_2 fixes x_2, x_4 and permutes x_1, x_3. Relations are:

$$J_3 = J_1, \quad J_4 = J_2 J_3^a = J_1^a J_2.$$

Let x be a rational point of X_l. Then the sheaf $\mathcal{O}_x$ equals $(1 - J_1)(1 - J_2)$ in K_0. For any $m \in \mathbb{Z}$, consider the exact sequence

$$0 \to \mathcal{O}(-(m+1)D_1 - D_2) \to \mathcal{O}(-mD_1 - D_2) \to \mathcal{O}_{D_1}(-mD_1 - D_2) \to 0.$$

Since $D_1 \cong \mathbb{P}^1$ and $\deg[\mathcal{O}_{D_1}(-mD_1 - D_2)] = D_1 \cdot (-mD_1 - D_2) = -1$, we have

$$\mathcal{O}_{D_1}(-mD_1 - D_2) = \mathcal{O}_{D_1}(-1) = \mathcal{O}_{D_1} - \mathcal{O}_x \quad \text{in } K_0.$$

Hence $J_1^{m+1} J_2 = J_1^m J_2 + J_1 J_2 - J_2$ in K_0. This implies $J_4 = J_1^a J_2$ belongs to the abelian group generated by $1, J_1, J_2, J_1 J_2$. By Theorem 3.8, we have K_0 as an abelian group is generated by $1, J_1, J_2, J_1 J_2$. They form a basis of K_0 because the rank of K_0 (= the number of maximal cones in the fan) is 4. Thus, K_0 has a permutation basis $1, J_1, J_2, J_1 J_2$. (Alternatively, this basis can easily be obtained from the projective bundle theorem [Quillen 1973, §8, Theorem 2.1] because F_a is a $\mathbb{P}^1$-bundle over $\mathbb{P}^1$.)

(ii) Rays $x_1 = (1, 0),\ x_2 = (0, 1),\ x_3 = (-1, -1)$: Then D_6 rotates x_i and reflects along lines in x_1, x_2, x_3. Relations are $J_1 = J_2 = J_3$. A permutation basis is $1, J_1, J_1^2$.

(iii) Rays $x_1 = (1, 0),\ x_2 = (0, 1),\ x_3 = (-1, 0),\ x_4 = (0, -1)$: Then D_8 rotates x_i and reflects along lines in $x_1, x_2, (1, 1), (-1, 1)$. Relations are:

$$J_3 = J_1, \quad J_4 = J_2.$$

A permutation basis is $1, J_1, J_2, J_1 J_2$.

(iv) Rays $x_1 = (1, 0)$, $x_2 = (0, 1)$, $x_3 = (-1, -1)$, $y_1 = (-1, 0)$, $y_2 = (0, -1)$, $y_3 = (1, 1)$: Then $D_{12} \cong S_2 \times S_3$ (S_2, S_3 permutation groups), $S_2 = \langle -1 \rangle$ switches between x_i and y_i, and S_3 permutes the pair of rays (x_i, y_i). Let D_i' be the divisors corresponding to the rays y_i and let $J_i' = \mathcal{O}(-D_i')$. Relations are

$$\frac{J_1}{J_1'} = \frac{J_2}{J_2'} = \frac{J_3}{J_3'}.$$

As proved in [Blunk 2010, Theorem 4.2], we have a permutation basis 1, R_1, R_2, R_3, Q_1, Q_2 where

$$R_1 = J_1 J_2', \quad R_2 = J_2 J_3', \quad R_3 = J_3 J_1', \quad Q_1 = J_1 J_2 J_3', \quad Q_2 = J_1' J_2' J_3. \quad \square$$

Remark 5.3. The difficulties in generalizing Theorem 5.2 to higher dimensions (at least using the approach of this paper) are:

(1) The classification of nonconjugacy classes of finite subgroups of $\mathrm{GL}(n, \mathbb{Z})$ is difficult and not complete. It often only provides algorithms and requires the help of a computer even for small n. Also, the number of those finite subgroups grows very fast relative to n. For example, there are total of 73 for $\mathrm{GL}(3, \mathbb{Z})$ and 710 for $\mathrm{GL}(4, \mathbb{Z})$.

(2) The K-group $K_0(X_l)$ in question may not stay a permutation module after blowups if X is not a surface.

6. Construction of separable algebras

Let X be a smooth projective toric T-variety over k that splits over l, and let X^* be its associated toric model; see Section 3. [Merkurjev and Panin 1997, Theorem 5.7] states that there is a split monomorphism $u : X^* \to A$ in the motivic category $\mathcal{C}$ from X^* to an étale k-algebra A and u is represented by an element Q in $\mathrm{Pic}(X^* \otimes_k A)$. Using the invertible sheaf Q, a map $u' : X \to B$ can be constructed out of u. Theorem 7.6 of the same work states that u' is also a split monomorphism in $\mathcal{C}$. In this section, we will recall the construction of u' and consider the case when u is an isomorphism.

Write $X_A = X \otimes_k A$ and we have $f : X_l \to X_l^*$, a T_l-isomorphism. Consider the diagram:

$$\begin{array}{ccc} X_{A\otimes_k l} & \xrightarrow{f_A} & X^*_{A\otimes_k l} \\ \Big\downarrow{\scriptstyle \pi_{X_A}} & & \Big\downarrow{\scriptstyle \pi_{X_A^*}} \\ X_A & & X_A^* \end{array} \tag{3}$$

Let $P' = f^*(\pi^*_{X_A^*}(Q))$. Then $B = \mathrm{End}_{X_A}(\pi_{X_A *}(P')) \in \mathrm{Br}(A)$ and $u' : X \to B$ is represented by $\pi_{X_A *}(P')$, namely $u' = \phi_*(P') \in K_0(X, B)$, where ϕ is the projection $X_{A\otimes_k l} \to X$.

The following criterion, which is [Merkurjev and Panin 1997, Proposition 4.5], checks when a toric model is isomorphic to an étale algebra in $\mathcal{C}$:

Proposition 6.1. *Let X^* be a smooth projective toric model over k that splits over l and $G = \mathrm{Gal}(l/k)$. If $K_0(X_l^*)$ is a permutation G-module, then $X^* \cong \mathrm{Hom}_G(P, l)$ in the motivic category $\mathcal{C}$ for any permutation G-basis P of $K_0(X_l^*)$.*

Remark 6.2. In particular, this implies that for any split smooth projective toric variety Y over k, $Y \cong k^n$ in $\mathcal{C}$ where n equals to the rank of $K_0(Y)$ (also equals to the number of maximal cones of the fan). Note that a smooth projective toric variety Y over k where the fan of Y_l has no symmetry (i.e., $\mathrm{Aut}_\Sigma(Y_l)$ is trivial) is automatically split.

Lemma 6.3. *Let X^*, G be the same as before. Then there is an isomorphism $u : X^* \to A$ in $\mathcal{C}$ where A is an étale k-algebra and u is represented by an element $Q \in \mathrm{Pic}(X_A^*)$ if and only if $K_0(X_l^*)$ has a permutation G-basis of line bundles on X_l^*.*

Proof. $\Rightarrow$: Decompose A as $\prod_{i=1}^t k_i$, where k_i are finite separable field extensions of k. We have $X_A^* = \coprod_{i=1}^t X_{k_i}^*$ the disjoint union of $X_{k_i}^*$ and $Q = \coprod_{i=1}^t Q_i$, where Q_i are line bundles on $X_{k_i}^*$. Let $q_i : X_{k_i}^* \to X^*$ be the projections. Then $u = \bigoplus_{i=1}^t q_{i*} Q_i$. Let $p_i : X_{k^s}^* \to X_{k_i}^*$ be the projections and $G_i = \mathrm{Gal}(k_i/k)$. Then

$$u_{k_s} = \bigoplus_{i=1}^{t} p_i^* q_i^* q_{i*}(Q_i) = \bigoplus_{i=1}^{t} \bigoplus_{g \in G_i} p_i^*(g Q_i)$$

and $A_{k^s} \cong (k^s)^n$ where $n = \sum_{i=1}^t |G_i|$. View u as $u^{\mathrm{op}} : A^{\mathrm{op}} = A \to X^*$. Then the map $u_{k^s}^{\mathrm{op}}$ induces an isomorphism $K_0((k^s)^n) \to K_0(X_{k^s}^*)$, where the canonical basis of the former is sent to $\{p_i^*(g Q_i) \mid g \in G_i, 1 \leqslant i \leqslant t\}$ and this set gives a permutation Γ-basis of $K_0(X_{k^s}^*)$ consisting of line bundles. As $\mathrm{Gal}(k^s/l)$ acts trivially on $K_0(X_{k^s}^*)$, this basis descends to X_l^*.

$\Leftarrow$: Assume P is a permutation G-basis of $K_0(X_l^*)$ consisting of line bundles on X_l^* and P divides into t G-orbits. Let $\{S_i\}_{i=1}^t$ be the set of representatives of G-orbits, and let $\mathrm{Gal}(l/k_i)$ be the stabilizer of S_i. Set $A = \mathrm{Hom}_G(P, l)$. Then $A \cong \prod_{i=1}^t k_i$. Since X^* has a rational point, by [Colliot-Thélène et al. 2008, Proposition 5.1], we have $S_i \in \mathrm{Pic}(X_l^*)^{\mathrm{Gal}(l/k_i)} \cong \mathrm{Pic}(X_{k_i}^*)$, namely $S_i \cong p_i^*(Q_i)$ for some $Q_i \in \mathrm{Pic}(X_{k_i}^*)$, where $p_i : X_l^* \to X_{k_i}^*$ are the projections. There is a morphism $u : X^* \to A$ which is represented by $\coprod_{i=1}^t Q_i \in \mathrm{Pic}(X_A^*)$, and by construction, the map u_l induces an isomorphism $K_0(X_l^*) \cong K_0(A_l)$. Using the following lemma, we have u is an isomorphism. $\square$

Lemma 6.4. *Let X^* be the same as before and A an étale k-algebra. If $u : X^* \to A$ is a morphism in $\mathcal{C}$ such that $K_0(u_{k^s}) : K_0(X_{k^s}^*) \to K_0(A_{k^s})$ is an isomorphism, then so is u.*

Proof. There is a commutative diagram:

$$\begin{array}{ccc} K_0(X^*) & \xrightarrow{K_0(u)} & K_0(A) \\ \downarrow & & \downarrow \\ K_0(X^*_{k^s})^\Gamma & \xrightarrow{K_0(u_{k^s})} & K_0(A_{k^s})^\Gamma \end{array}$$

The right vertical map is an isomorphism because A is étale and so is $K_0(u_{k^s})$ by assumption. The left vertical map is an isomorphism by [Merkurjev and Panin 1997, Corollary 5.8]. Thus, $K_0(u)$ is also an isomorphism.

Write $w = u^{\mathrm{op}} : A \to X^*$. Then by the splitting principle (their Proposition 6.1) and its proof, $K_0^{X^*}(w) : K_0(X^*, A) \to K_0(X^* \times X^*)$ is surjective. Thus, there exists $v \in K_0(X^*, A) : X^* \to A$ such that $w \circ v = K_0^{X^*}(w)(v) = 1_{X^*}$, and thus $K_0(w \circ v) = K_0(w)K_0(v) = 1_{K_0(X^*)}$. Since $K_0(w) = \phi$ is an isomorphism, we have $K_0(v) = \phi^{-1}$ and $K_0(v \circ w) = K_0(v)K_0(w) = 1_{K_0(A)}$. This implies $v \circ w = 1_A$ and thus v is a two sided inverse of w in $\mathcal{C}$. □

The proof of (3) $\iff$ (4) in their Proposition 7.9 shows that the T_l-isomorphism $f : X_l \to X_l^*$ induces a $G = \mathrm{Gal}(l/k)$-module isomorphism $f^* : K_0(X_l^*) \to K_0(X_l)$. Thus, $K_0(X_l^*)$ has a permutation G-basis of line bundles on X_l^* if and only if $K_0(X_l)$ has such a basis. Note that the proof (1) $\Rightarrow$ (2) (an isomorphism $u : X^* \to A$ gives an isomorphism $u' : X \to B$), which uses the construction (3) recalled at the beginning of the section, works only when u is represented by an element $Q \in \mathrm{Pic}(X_A^*)$. Thus, we have the following instead:

Theorem 6.5. *Let X be a smooth projective toric T-variety over k that splits over l and $G = \mathrm{Gal}(l/k)$. Assume $K_0(X_l)$ has a permutation G-basis P of line bundles on X_l. Let $\{P_i\}_{i=1}^t$ be G-orbits of P, and let $\pi : X_l \to X$ be the projection. For any $S_i \in P_i$, set $B_i = \mathrm{End}_{\mathcal{O}_X}(\pi_*(S_i))$ and $B = \prod_{i=1}^t B_i$. Then the map $u = \bigoplus_{i=1}^t \pi_*(S_i) : X \to B$ gives an isomorphism in the motivic category $\mathcal{C}$.*

Proof. By Lemma 6.3, we have an isomorphism $u : X^* \to A$ represented by $Q \in \mathrm{Pic}(X_A^*)$. Here $A \cong \prod_{i=1}^t k_i$ where $\mathrm{Gal}(l/k_i)$ are the stabilizers of S_i under the G-action. Then Q is the disjoint union $\coprod_{i=1}^t Q_i$ where the $Q_i \in \mathrm{Pic}(X_{k_i}^*)$ descend from $(f^*)^{-1}(S_i) \in \mathrm{Pic}(X_l^*)^{\mathrm{Gal}(l/k_i)}$. Now we run the construction (3) for Q_i:

$$\begin{array}{ccc} X_{k_i \otimes_k l} & \xrightarrow{f_i} & X^*_{k_i \otimes_k l} \\ \downarrow{\scriptstyle \pi_X} & & \downarrow{\scriptstyle \pi_{X^*}} \\ X_{k_i} & & X^*_{k_i} \end{array}$$

Let $p : X_l \to X_{k_i}$ and $q : X_{k_i} \to X$ be the projections. Then $\pi_{X*} f_i^* \pi_{X^*}^*(Q_i) \cong p_*(S_i) \otimes_k k_i$ where its $\mathcal{O}_{X_{k_i}}$-module structure comes from the one on $p_*(S_i)$. Thus,

$\mathrm{End}_{\mathcal{O}_{X_{k_i}}}(\pi_{X*} f_i^* \pi_{X^*}^*(Q_i)) \cong \mathrm{End}_{\mathcal{O}_{X_{k_i}}}(p_*(S_i)) \otimes_k \mathrm{End}_k(k_i)$ is Brauer equivalent to $B_i' = \mathrm{End}_{\mathcal{O}_{X_{k_i}}}(p_* S_i)$. It remains to prove that $B_i \cong B_i'$. There is a G-isomorphism:

$$\begin{aligned} B_i \otimes_k l \cong \mathrm{End}_{\mathcal{O}_{X_l}}(\pi^* \pi_*(S_i)) &\cong \mathrm{End}_{\mathcal{O}_{X_l}}(p^* q^* q_* p_*(S_i)) \\ &\cong \mathrm{End}_{\mathcal{O}_{X_l}}(p^* p_*(S_i) \otimes_k k_i) \\ &\cong \mathrm{End}_{\mathcal{O}_{X_l}}(p^* p_*(S_i)) \otimes_k k_i \\ &\cong (B_i' \otimes_{k_i} l) \otimes_k k_i \cong B_i' \otimes_k l. \end{aligned}$$

The fourth isomorphism follows from Lemma 6.6. Taking G-invariants on both sides, we have $B_i \cong B_i'$. □

Lemma 6.6. *Let X be a proper variety over k and assume that there is a finite group G acting on Cartier divisors* $\mathrm{CDiv}(X)$. *Let $D \in \mathrm{CDiv}(X)$ and $g \in G$ such that D and gD are not linearly equivalent. Then* $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_X(D), \mathcal{O}_X(gD)) = 0$.

Proof. Assume that $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_X(D), \mathcal{O}_X(gD)) \neq 0$, which is equivalent to assuming $\mathcal{O}_X(gD - D)$ has a nonzero global section s. Since G is a finite group, $g^n = 1$ for some n. Thus, the invertible sheaf $\mathcal{O}_X(D-gD) = (g^{n-1} \otimes \cdots \otimes g \otimes 1)\mathcal{O}_X(gD-D)$ has a nonzero global section $t = g^{n-1}s \otimes \cdots \otimes s$. We view s and t as maps $s : \mathcal{O}_X(D) \to \mathcal{O}_X(gD)$ and $t : \mathcal{O}_X(gD) \to \mathcal{O}_X(D)$. Since $st, ts \in \Gamma(X, O_X) = k$ are nonzero, we have $\mathcal{O}(gD - D) \cong \mathcal{O}_X$, a contradiction. □

Remark 6.7. There is a more “economical” description of the algebra isomorphic to X in $\mathcal{C}$:

Write $S_i = \mathcal{O}(-D_i)$, where the D_i are torus invariant. Let $\mathrm{Gal}(l/l_i)$ be the stabilizer of D_i under the G-action and let $\pi_i : X_{l_i} \to X$ be the projection. Then divisors D_i and thus invertible sheaves S_i descend to X_{l_i}, and we use the same notation. Then $X \cong \prod_{i=1}^t \mathrm{End}_{\mathcal{O}_X}(\pi_{i*}(S_i))$. In effect, it replaces all $\mathrm{M}_n(k)$ in B constructed in the theorem by k which is an isomorphism in $\mathcal{C}$.

Remark 6.8. A question remains: If $K_0(X_l)$ is a permutation G-module, can we always find a permutation G-basis of line bundles?

Recall that for $n \geqslant 0$, K_n defines a functor $K_n : \mathcal{C} \to \mathrm{Ab}$. Hence we have

Corollary 6.9. $K_n(X) \cong \prod_{i=1}^t K_n(B_i)$.

7. Separable algebras for toric surfaces

Separable algebras for minimal toric surfaces. Recall the families of minimal toric surfaces described in Theorem 5.2: Let X be a minimal smooth projective toric T-surface over k that splits over l, and let X^* be its associated toric model. Let $\pi : X_l \to X$ be the projection. All isomorphisms below are taken in the motivic category $\mathcal{C}$.

(i) If $X_l \cong F_a$, $a \geqslant 2$, then $X^* \cong k^4$ and $X \cong k \times Q \times k \times Q$, where $Q \cong \mathrm{End}_{\mathcal{O}_X}(\pi_* J_1)$ is a quaternion k-algebra.

(ii) More generally, let $X = \mathrm{SB}(A)$ be a Severi–Brauer variety of dimension n and $J = \mathcal{O}_{X_l}(-1)$. Then $X^* \cong k^{n+1}$ and $X \cong k \times \prod_{i=1}^{n} A^{\otimes i}$, where $A^{\otimes i} \cong \mathrm{End}_{\mathcal{O}_X}(\pi_* J^i)$; see Example 3.5.

(iii) If $X_l \cong \mathbb{P}^1 \times \mathbb{P}^1$, then $X^* \cong k \times K \times k$ where K is a quadratic étale algebra and the discriminant extension of X, and $X \cong k \times B \times A$, where $B \cong \mathrm{End}_{\mathcal{O}_X}(\pi_* J_1)$ is an Azumaya K-algebra of rank 4 and $A \cong \mathrm{End}_{\mathcal{O}_X}(\pi_*(J_1 J_2))$ is a central simple k-algebra of degree 4; see Example 3.7.

(iv) See Example 3.6, where $X^* \cong k \times K \times L$ and $P \cong \mathrm{End}_{\mathcal{O}_X}(\pi_* R_1)$ and $Q \cong \mathrm{End}_{\mathcal{O}_X}(\pi_* Q_1)$.

Now let X be a smooth projective toric T-variety over k that splits over l and $G = \mathrm{Gal}(l/k)$. Recall that X is uniquely determined by the associated toric model X^*, which corresponds to $\rho : \Gamma \to \mathrm{GL}(n, \mathbb{Z})$, the fan Σ such that $\rho(\Gamma) \subseteq \mathrm{Aut}_\Sigma$, and a principal homogeneous space $U \in H^1(k, T)$. Every variety within a family above has the same fan. Let $\rho' : G \hookrightarrow \mathrm{Aut}_\Sigma(X_l)$ be the inclusion induced by ρ. We want to see how the separable algebras described above relate to ρ' and U.

Let $\dim X = n$ and let N be the number of rays in the fan Σ. Then the Picard rank of X_l is $m = N - n$. Write M for the group of characters of T_l and CDiv_{T_l} for T_l-invariant Cartier divisors. There is a natural action of $\mathrm{Aut}_\Sigma(X_l)$ on M and $\mathrm{CDiv}_{T_l}(X_l)$, and an induced action on $\mathrm{Pic}(X_l)$ via the canonical morphism $\mathrm{CDiv}_{T_l}(X_l) \to \mathrm{Pic}(X_l)$, $D \mapsto \mathcal{O}_{X_l}(D)$.

We have a short exact sequence of $\mathrm{Aut}_\Sigma(X_l)$-modules and therefore of G-modules via ρ':

$$0 \to M \to \mathrm{CDiv}_{T_l}(X_l) \to \mathrm{Pic}(X_l) \to 0, \tag{4}$$

or simply $0 \to \mathbb{Z}^n \to \mathbb{Z}^N \to \mathbb{Z}^m \to 0$. It corresponds to the short exact sequence of tori over l:

$$1 \to \mathbb{G}_{m,l}^m \to \mathbb{G}_{m,l}^N \to \mathbb{G}_{m,l}^n \to 1$$

and the sequence descends to

$$1 \to S \to V \to T \to 1. \tag{5}$$

Let $i : \mathrm{Aut}_\Sigma(X_l) \hookrightarrow S_N$, where S_N is the group of permutations of the canonical $\mathbb{Z}$-basis of the lattice $\mathbb{Z}^N$ and it induces $i_* : H^1(G, \mathrm{Aut}_\Sigma) \to H^1(G, S_N)$. Let $[\alpha] = i_*[\rho']$ and let E be the corresponding étale k-algebra of degree N. Then $V = \mathrm{R}_{E/k}(\mathbb{G}_{m,E})$. Let $j : \mathrm{Aut}_\Sigma(X_l) \to \mathrm{GL}(m, \mathbb{Z})$ be the map induced by the action of $\mathrm{Aut}_\Sigma(X_l)$ on $\mathrm{Pic}(X_l)$ which induces $j_* : H^1(G, \mathrm{Aut}_\Sigma) \to H^1(G, \mathrm{GL}(m, \mathbb{Z}))$. Let $[\beta] = j_*[\rho']$. Then S is the torus corresponding to $[\beta]$.

The short exact sequence of tori over k gives

$$0 \to H^1(G,T) \xrightarrow{\delta} H^2(G,S) \to \mathrm{Br}(E).$$

Here, by Hilbert's Theorem 90,

$$H^1(G,V) = H^1(G, \mathrm{R}_{E/k}(\mathbb{G}_{m,E})(l)) = \prod H^1(\mathrm{Gal}(E_t/k), E_t^\times) = 0,$$

where $E = \prod E_t$ and the E_t are finite separable field extensions of k.

Let $S^* = \mathrm{Hom}(S_l, G_{m,l})$ be the group of characters over l. Then sequence (4) can be rewritten as

$$0 \to T^* \to V^* \to S^* \to 0,$$

which induces $H^0(G,S^*) \xrightarrow{\partial} H^1(G,T^*)$. Geometrically, ∂ is the map $\mathrm{Pic}(X^*) \to \mathrm{Pic}(T)$ which sends $Q \in \mathrm{Pic}(X^*)$ to its restriction $Q|_T$ on T.

There is a G-equivariant bilinear map $S(l) \otimes S^* \to l^\times$ which sends $x \otimes \chi$ to $\chi(x)$, and it induces a pairing of Galois cohomology groups $\cup : H^2(G,S) \otimes H^0(G,S^*) \to \mathrm{Br}(k)$. Similarly, we have $\cup : H^1(G,T) \otimes H^1(G,T^*) \to \mathrm{Br}(k)$.

Lemma 7.1. *The following diagram is commutative*:

$$\begin{tikzcd}
H^1(G,T) \otimes H^0(G,S^*) \arrow[r, "1\otimes\partial"] \arrow[d, "\delta\otimes 1"] & H^1(G,T) \otimes H^1(G,T^*) \arrow[d, "\cup"] \\
H^2(G,S) \otimes H^0(G,S^*) \arrow[r, "\cup"] & \mathrm{Br}(k)
\end{tikzcd}$$

Proof. Let $a \in H^1(G,T)$, $\varphi \in H^0(G,S^*)$. For each $a_g \in T(l)$, $g \in G$, pick $b_g \in V(l)$ that maps to a_g. Then $(\delta a)_{g,h} = b_{gh}^{-1} b_g {}^g b_h$, $g, h \in G$. Pick $\phi \in V^*$ that maps to φ. Then $(\partial\varphi)_g = \phi^{-1} {}^g\phi$. Let $\alpha = a \cup (\partial\varphi)$ and $\beta = (\delta a) \cup \varphi$. Then

$$\alpha_{g,h} = {}^g(\partial\varphi)_h(a_g) = {}^g(\phi^{-1}{}^h\phi)(b_g) = ({}^g\phi^{-1})(b_g)\cdot({}^{gh}\phi)(b_g)$$

and

$$\beta_{g,h} = ({}^{gh}\varphi)((\delta a)_{g,h}) = ({}^{gh}\phi)(b_{gh}^{-1})\cdot({}^{gh}\phi)(b_g)\cdot({}^{gh}\phi)({}^g b_h).$$

Set $\theta_g = ({}^g\phi)(b_g)$. Then $\beta_{g,h} = \theta_{gh}^{-1}\theta_g {}^g\theta_h \alpha_{g,h}$. Thus, α and β give the same cycle class in $\mathrm{Br}(k)$. $\square$

Let $P \in \mathrm{Pic}(X_l)$ be a line bundle on X_l with stabilizer group $\mathrm{Gal}(l/\kappa)$ under the G-action. Since $P \in \mathrm{Pic}(X_l)^{\mathrm{Gal}(l/\kappa)} \cong (S^*)^{\mathrm{Gal}(l/\kappa)}$, the line bundle P corresponds to a character $\chi : S_\kappa \to \mathbb{G}_{m,\kappa}$ over κ, or equivalently $\chi' : S \to \mathrm{R}_{\kappa/k}(\mathbb{G}_{m,\kappa})$. Let $\pi : X_l \to X$ be the projection.

Proposition 7.2. *Let* $\delta_P : H^1(G,T) \xrightarrow{\delta} H^2(G,S) \xrightarrow{\chi'} \mathrm{Br}(\kappa)$ *be the composition map. Then* $\delta_P[U] = [\mathrm{End}_{\mathcal{O}_X}(\pi_* P)] \in \mathrm{Br}(\kappa)$.

Proof. First we prove the case when $\kappa = k$. In this case, the line bundle $P \in \operatorname{Pic}(X_l)^G \cong \operatorname{Pic}(X^*)$. Thus, there is $Q \in \operatorname{Pic}(X^*)$ such that $P \cong f^*\pi_{X^*}^* Q$, where $\pi_{X^*} : X_l^* \to X^*$ is the projection and $f : X_l \to X_l^*$ is the T_l-isomorphism. [Merkurjev and Panin 1997, Lemma 7.3] shows that $[U] \cup [Q|_T] = [\operatorname{End}_{\mathcal{O}_X}(\pi_* P)] \in \operatorname{Br}(k)$. On the other hand, $\delta_P([U]) = \delta[U] \cup [\chi'] = \delta[U] \cup [Q]$. By Lemma 7.1, $\delta_P([U]) = [U] \cup [\partial Q] = [U] \cup [Q|_T]$.

In general, let $H = \operatorname{Gal}(l/\kappa)$ and consider the restriction map $\operatorname{Res} : H^1(G, T) \to H^1(H, T_\kappa)$ which sends $[U]$ to $[U_\kappa]$. There is a commutative diagram:

$$\begin{array}{ccccc}
H^1(G,T) & \xrightarrow{\delta} & H^2(G,S) & \xrightarrow{\chi'} & \operatorname{Br}(\kappa) \\
\Big\downarrow{\scriptstyle \operatorname{Res}} & & \Big\downarrow{\scriptstyle \operatorname{Res}} & & \Big\| \\
H^1(H,T_\kappa) & \xrightarrow{\delta} & H^2(H,S_\kappa) & \xrightarrow{\chi} & \operatorname{Br}(\kappa)
\end{array}$$

Thus, $\delta_P[U] = [\operatorname{End}_{\mathcal{O}_{X_\kappa}}(\pi_{\kappa *} P)]$, where $\pi_\kappa : X_l \to X_\kappa$ is the projection. By the proof of Lemma 6.3, $\operatorname{End}_{\mathcal{O}_{X_\kappa}}(\pi_{\kappa *} P) \cong \operatorname{End}_{\mathcal{O}_X}(\pi_* P)$. □

Corollary 7.3. *Let X be a smooth projective toric variety over k that splits over l and $G = \operatorname{Gal}(l/k)$. Assume $\operatorname{Pic}(X_l)$ is a permutation G-module, i.e., the torus S is quasitrivial and thus has the form $\prod_{i=1}^t \mathrm{R}_{k_i/k}\,\mathbb{G}_{m,k_i}$, where k_i are finite separable field extensions of k. Then the principal homogeneous space U is uniquely determined by $(B_i \in \operatorname{Br}(k_i))_{1 \leqslant i \leqslant t}$, where B_i split over E. Let $\{S_i\}_{i=1}^t$ be the set of representatives for G-orbits of $\operatorname{Pic}(X_l)$. Then B_i comes from $\operatorname{End}_{\mathcal{O}_X}(\pi_* S_i)$.*

Proof. The result follows from Proposition 7.2 and the exact sequence

$$0 \to H^1(k,T) \to \prod_{i=1}^t \operatorname{Br}(k_i) \to \operatorname{Br}(E). \qquad \square$$

Remark 7.4. Families (i), (ii) and (iii) and their blowups have permutation Picard groups.

(ii): Let $X = \operatorname{SB}(A)$ be a Severi–Brauer variety of dimension n, $\operatorname{Aut}_\Sigma(X_l) = S_{n+1}$. We have

$$1 \to \mathbb{G}_{m,k} \to \mathrm{R}_{E/k}(\mathbb{G}_{m,E}) \to T \to 1,$$

which induces

$$0 \to H^1(G,T) \xrightarrow{\delta} \operatorname{Br}(k) \to \operatorname{Br}(E).$$

Then $\delta(U) = [A]$ and A splits over E; see [Merkurjev and Panin 1997, Example 8.5].

(i): Let $X_l = F_a$, $a \geqslant 2$, $\operatorname{Aut}_\Sigma = S_2$, and E factors as $k \times F \times k$, where F is the quadratic étale k-algebra corresponding to $[\rho'] \in H^1(G, S_2)$. We have

$$1 \to \mathbb{G}_{m,k} \to \mathbb{G}_{m,k} \times \mathrm{R}_{F/k}(\mathbb{G}_{m,F}) \to T \to 1,$$

where $\mathbb{G}_{m,k} \to \mathbb{G}_{m,k}$ is the a-th power homomorphism. It induces

$$0 \to H^1(G, T) \xrightarrow{\delta} \mathrm{Br}(k) \to \mathrm{Br}(k) \times \mathrm{Br}(F),$$

where $[U] \mapsto [Q] \mapsto ([Q^{\otimes a}], [Q_F])$. By Lemma 4.11, the toric surface X is a $\mathbb{P}^1$-bundle over some conic curve C. We have the torus of C is $T' = \mathrm{R}_{F/k}(\mathbb{G}_{m,F})/\mathbb{G}_{m,k}$. There is a commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathbb{G}_{m,k} & \longrightarrow & \mathbb{G}_{m,k} \times \mathrm{R}_{F/k}(\mathbb{G}_{m,F}) & \longrightarrow & T & \longrightarrow & 1 \\ & & \| & & \downarrow & & \downarrow h & & \\ 1 & \longrightarrow & \mathbb{G}_{m,k} & \longrightarrow & \mathrm{R}_{F/k}(\mathbb{G}_{m,F}) & \longrightarrow & T' & \longrightarrow & 1 \end{array}$$

Hence, the image of $[U]$ under $\delta \circ h_* : H^1(G, T) \to H^1(G, T') \to \mathrm{Br}(k)$ is $[Q]$, and thus $C = \mathrm{SB}(Q)$. Since a quaternion algebra has a period at most 2 in the Brauer group, if a is odd, then $[Q^{\otimes a}] \in \mathrm{Br}(k)$ being trivial implies that $Q = \mathrm{M}_2(k)$. Thus we have:

Proposition 7.5. *Let X be a toric surface that is a form of F_{2a+1}. Then $X \cong F_{2a+1}$.*

Remark 7.6. Iskovskih showed that any form of F_{2a+1} is trivial [Iskovskih 1979, Theorem 3(2)]. The above proposition reproves this result in the case of toric surfaces.

(iii): Let $X_l = \mathbb{P}^1 \times \mathbb{P}^1$, $\mathrm{Aut}_\Sigma = D_8$. In this case, the map $\beta : G \to \mathrm{GL}(2, \mathbb{Z})$ factors through $\gamma : G \to S_2$, where S_2 permutes $\mathcal{O}(1, 0)$ and $\mathcal{O}(0, 1)$. Then the quadratic étale algebra K corresponds to γ. We have

$$1 \to \mathrm{R}_{K/k}(\mathbb{G}_{m,K}) \to \mathrm{R}_{E/k}(\mathbb{G}_{m,E}) \to T \to 1,$$

which induces

$$0 \to H^1(G, T) \xrightarrow{\delta} \mathrm{Br}(K) \to \mathrm{Br}(E).$$

Then $\delta(U) = [B]$ and B splits over E. Let $\mathrm{N}_{K/k} : \mathrm{R}_{K/k}(\mathbb{G}_{m,K}) \to \mathbb{G}_{m,k}$ be the norm map which induces $\mathrm{cor}_{K/k} : \mathrm{Br}(K) \to \mathrm{Br}(k)$. Then $[A] = \mathrm{cor}_{K/k}[B]$.

Separable algebras for toric surfaces. Let X be a smooth projective toric T-surface over k that splits over l and $G = \mathrm{Gal}(l/k)$. Recall that we have a finite chain of blowups of toric T-surfaces

$$X = X_0 \to X_1 \to \cdots \to X_n = X',$$

where X' is minimal. For $1 \leqslant i \leqslant n$, let f_i map $(X_{i-1})_l \to (X_i)_l$, which are the blowups of G-sets of disjoint T_l-invariant points. Let E_i be the G-sets of the exceptional divisors of f_i and $X' \cong B$ in $\mathcal{C}$.

Proposition 7.7. $$X \cong B \times \prod_{i=1}^{n} \mathrm{Hom}_G(E_i, l) \quad \textit{in } \mathcal{C}.$$

Proof. We only need to consider the simple case: Let $f : Y \to Z$ be a blowup of toric T-surfaces and let $E = \{P_j\}$ be the G-set of line bundles associated to the exceptional divisors of $g = f_l$. We assume further that the G-action on E is transitive.

Let $p : Y_l \to Y$ and $q : Z_l \to Z$ be the projections. Then we have a commutative diagram:

$$\begin{array}{ccc} Y_l & \xrightarrow{g} & Z_l \\ \downarrow^{p} & & \downarrow^{q} \\ Y & \xrightarrow{f} & Z \end{array}$$

Recall that if $K_0(Z_l)$ has a G-basis γ, then $g^*(\gamma) \cup E$ is a G-basis of $K_0(Y_l)$. Since Z is a toric surface, we can assume γ consists of line bundles over Z_l. Let $P \in \gamma$. Then

$$\mathrm{End}_{\mathcal{O}_Y}(p_* g^* P) \cong \mathrm{End}_{\mathcal{O}_Y}(f^* q_* P) \cong \mathrm{Hom}_{\mathcal{O}_Z}(q_* P,\, f_* f^*(q_* P)) \cong \mathrm{End}_{\mathcal{O}_Z}(q_* P),$$

where $f_* f^*$ is identity because f is flat proper and $f_* \mathcal{O}_Y = \mathcal{O}_Z$.

As for the G-orbit E, we have $\bigoplus_j P_j = p^* Q$ for some locally free sheaf Q on Y. By Lemma 6.6 and the assumption that G acts transitively on E, we have $\mathrm{End}_{\mathcal{O}_Y}(Q) \cong \mathrm{Hom}_G(E, l)$. It is Brauer equivalent to $\mathrm{End}_{\mathcal{O}_Y}(p_* P_j)$ for any $P_j \in E$. Thus the result follows from Theorem 6.5. □

8. Derived categories of toric surfaces

Let X be a smooth projective variety over k and let $D^b(X)$ be the bounded derived category of coherent sheaves on X. We will define exceptional objects and collections in a generalized way.

Definition 8.1. Let A be a finite simple k-algebra. An object V in $D = D^b(X)$ is called *A-exceptional* if $\mathrm{Hom}_D(V, V) = A$ and $\mathrm{Ext}^i_D(V, V) = 0$ for $i \neq 0$.

Definition 8.2. A set of objects $\{V_1, \ldots, V_n\}$ in $D = D^b(X)$ is called an *exceptional collection* if for each $1 \leqslant i \leqslant n$, the object V_i is A_i-exceptional for some finite simple k-algebra A_i, and $\mathrm{Ext}^r_D(V_i, V_j) = 0$ for any integer r and $i > j$. The collection is *full* if the thick triangulated subcategory $\langle V_1, \ldots, V_n\rangle$ generated by the V_i is equivalent to $D^b(X)$.

Definition 8.3. A set of objects $\{V_1, \ldots, V_n\}$ in $D \in D^b(X)$ is called an *exceptional block* if it is an exceptional collection and $\mathrm{Ext}^r_D(V_i, V_j) = 0$ for any integer r and $i \neq j$. Note that the ordering of the V_i in this case does not matter.

Assume $\{V_1, \ldots, V_n\}$ is a full exceptional collection as above. Since $\langle V_i\rangle$ is equivalent to $D^b(A_i)$, the bounded derived category of right A_i-modules, we have semiorthogonal decompositions $D^b(X) = \langle V_1, \ldots, V_n\rangle = \langle D^b(A_1), \ldots, D^b(A_n)\rangle$.

The semiorthogonal decomposition of $D^b(X)$ can be lifted to the world of dg categories. For details about dg categories, see [Keller 2006]. There is a dg enhancement of $D^b(X)$, denoted as $D^b_{dg}(X)$ where $D^b_{dg}(X)$ is the dg category with same objects as $D^b(X)$ and whose morphisms have a dg k-module structure such that $H^0(\mathrm{Hom}_{D^b_{dg}(X)}(x, y)) = \mathrm{Hom}_{D^b(X)}(x, y)$. Let $\mathrm{perf}_{dg}(X)$ be the dg subcategory of perfect complexes. Since X is smooth projective, $\mathrm{perf}_{dg}(X)$ is quasiequivalent to $D^b_{dg}(X)$. For an A-exceptional object V, the pretriangulated dg subcategory $\langle V\rangle_{dg}$ generated by V is quasiequivalent to $D^b_{dg}(A)$. Therefore, there is a dg enhancement of the semiorthogonal decomposition $D^b_{dg}(X) = \langle V_1, \ldots, V_n\rangle_{dg}$, which is quasiequivalent to $\langle D^b_{dg}(A_1), \ldots, D^b_{dg}(A_n)\rangle_{dg}$.

Let $dgcat$ be the category of all small dg categories. There is a universal additive functor $U : dgcat \to Hmo_0$ where Hmo_0 is the category of noncommutative motives, see [Tabuada 2015, §2.1-2.4]. We have $U(\mathrm{perf}_{dg}(X)) \simeq \bigoplus_{i=1}^n U(D^b_{dg}(A_i)) \simeq \bigoplus_{i=1}^n U(A_i)$. On the other hand, the motivic category $\mathcal{C}$ is a full subcategory of Hmo_0 by sending a pair (X, A) to $\mathrm{perf}_{dg}(X, A)$, the dg category of complexes of right $\mathcal{O}_X \otimes_k A$-modules which are also perfect complexes of $\mathcal{O}_X$-modules [Tabuada 2014, Theorem 6.10] or [Tabuada 2015, Theorem 4.17]. The above discussion gives the following well-known fact:

Theorem 8.4. *Let X be a smooth projective variety over k. If $D^b(X)$ has a full exceptional collection of objects $\{V_1, \ldots, V_n\}$ where each V_i is A_i-exceptional, then $X \cong \prod_{i=1}^n A_i$ in the motivic category $\mathcal{C}$.*

We know for toric varieties satisfying the conditions of Theorem 6.5, they have a complete motivic decomposition into central simple algebras. The following lemma gives a criterion when the motivic decomposition can be lifted to the decomposition of the derived category (i.e., the reverse of Theorem 8.4):

Lemma 8.5. *Let X be a smooth projective toric variety over k that splits over l and $G = \mathrm{Gal}(l/k)$. Assume $K_0(X_l)$ has a permutation G-basis P of line bundles over X_l. Let $\{P_i\}_{i=1}^t$ be G-orbits of P and let $\pi : X_l \to X$ be the projection.*

Assume each G-orbit P_i is an exceptional block. If there is an ordering for G-orbits $\{P_i\}_{i=1}^t$ such that $\{P_1, \ldots, P_t\}$ gives a full exceptional collection of $D^b(X_l)$, then for any $S_i \in P_i$, the set $\{\pi_*S_1, \ldots, \pi_*S_t\}$ is a full exceptional collection of $D^b(X)$.

Proof. First we show that $\{\pi_*S_1, \ldots, \pi_*S_t\}$ is an exceptional collection. Since π is flat and finite, both $\pi^* : D^b(X) \to D^b(X_l)$ and $\pi_* : D^b(X_l) \to D^b(X)$ are exact functors. The result follows from

$$\begin{aligned}\mathrm{Ext}^r_{D^b(X)}(\pi_*S_i, \pi_*S_j) \otimes_k l &\cong \mathrm{Ext}^r_{D^b(X_l)}(\pi^*\pi_*S_i, \pi^*\pi_*S_j)\\ &\cong \bigoplus_{g,g'\in G} \mathrm{Ext}^r_{D^b(X_l)}(gS_i, g'S_j).\end{aligned}$$

In particular, $\pi_* S_i$ is an exceptional object and thus $\langle \pi_* S_i \rangle$ is an admissible subcategory of $D^b(X)$. Since $\langle \pi_* S_i \otimes_k l \rangle = \langle P_i \rangle$ and $D^b(X_l) = \langle P_1, \ldots, P_t \rangle$, by [Auel and Bernardara 2015, Lemma 2.3], we have $D^b(X) = \langle \pi_* S_1, \ldots, \pi_* S_t \rangle$. □

Using the classification of toric surfaces, we can confirm the lifting for toric surfaces:

Theorem 8.6. *Let X be a smooth projective toric surface over k that splits over l and $G = \mathrm{Gal}(l/k)$. Then $K_0(X_l)$ has a permutation G-basis P of line bundles over X_l such that each G-orbit is an exceptional block. Furthermore, there exists an ordering of the G-orbits $\{P_i\}_{i=1}^t$ of P such that $\{P_1, \ldots, P_t\}$ gives a full exceptional collection of $D^b(X_l)$. Therefore, for any $S_i \in P_i$, $\{\pi_* S_1, \ldots, \pi_* S_t\}$ is a full exceptional collection of $D^b(X)$, where $\pi : X_l \to X$ is the projection.*

Proof. First assume that X is minimal. By the classification of minimal toric surfaces (Theorem 4.12), we have X_l is (i) F_a, $a \geqslant 2$; (ii) $\mathbb{P}^2$; (iii) $\mathbb{P}^1 \times \mathbb{P}^1$; (iv) del Pezzo surface of degree 6. Using the notation introduced in Theorem 5.2, the derived category $D^b(X_l)$ has the following full exceptional collections of line bundles:

(i) $\{\mathcal{O}, \mathcal{O}(D_1), \mathcal{O}(D_2), \mathcal{O}(D_1 + D_2)\}$;

(ii) $\{\mathcal{O}, \mathcal{O}(D_1), \mathcal{O}(2D_1)\} = \{\mathcal{O}, \mathcal{O}(1), \mathcal{O}(2)\}$;

(iii) $\{\mathcal{O}, \mathcal{O}(D_1), \mathcal{O}(D_2), \mathcal{O}(D_1 + D_2)\} = \{\mathcal{O}, \mathcal{O}(1, 0), \mathcal{O}(0, 1), \mathcal{O}(1, 1)\}$;

(iv) $\{\mathcal{O}, R_1^\vee, R_2^\vee, R_3^\vee, Q_1^\vee, Q_2^\vee\}$ where $(-)^\vee$ is the dual of the invertible sheaf.

Cases (i)–(iii) follow from the projective bundle theorem [Orlov 1992, Theorem 2.6] and (iv) follows from [Auel and Bernardara 2015, Proposition 9.1] or [Blunk et al. 2011]. Moreover, the collections $\{\mathcal{O}(1, 0), \mathcal{O}(0, 1)\}$, $\{R_i^\vee\}_{i=1}^3$ and $\{Q_j^\vee\}_{j=1}^2$ are exceptional blocks. These sets are the only G-orbits with more than one object. Therefore, each G-orbit is an exceptional block.

Now it suffices to consider the case that $f : X \to X'$ is a simple blowup of a minimal toric surface X', that is, the map $f_l : X_l \to X'_l$ is the blowup of a G-set of disjoint torus invariant points of X'_l where G acts on the set transitively. Let E_i be the exceptional divisors of f_l. Let E be the set $\{\mathcal{O}_{E_i}(-1)\}$. By [Orlov 1992, Theorem 4.3], the derived category $D^b(X)$ has a full exceptional collection $\{E, L^\bullet f^* D^b(X')\}$. Note that the full exceptional collections of minimal toric surfaces provided above all have the structure sheaf $\mathcal{O}$ as the first object. The right mutation of the pair $(\mathcal{O}_{E_i}(-1), \mathcal{O})$ is $(\mathcal{O}, \mathcal{O}(E_i))$ (the extension case in [Karpov and Nogin 1998, Proposition 2.3]). Therefore, the right mutation of $\{E, \mathcal{O}\}$ is $\{\mathcal{O}, E'\}$ where $E' = \{\mathcal{O}(E_i)\}$. The G-orbit E' is an exceptional block because the order in the set is exchangeable. Hence, $D^b(X_l)$ has a full exceptional collection $\{\mathcal{O}, E'$, the rest of the line bundles provided above$\}$ (they form a basis of $K_0(X_l)$) and each G-orbit is an exceptional block. □

Acknowledgements

This paper is a version of my Ph.D. thesis at University of California, Los Angeles. I would like to thank my advisor, Christian Haesemeyer, for proposing this question and for his help, and to thank Alexander Merkurjev for useful conversations. I also want to thank the referees for numerous helpful comments and suggestions. Lastly, I want to thank Patrick McFaddin for pointing out a mistake in an earlier version.

References

[Auel and Bernardara 2015] A. Auel and M. Bernardara, "Semiorthogonal decompositions and birational geometry of del Pezzo surfaces over arbitrary fields", preprint, 2015. To appear in *Proc. London Math. Soc.* arXiv

[Blunk 2010] M. Blunk, "Del Pezzo surfaces of degree 6 over an arbitrary field", *J. Algebra* **323**:1 (2010), 42–58. MR Zbl

[Blunk et al. 2011] M. Blunk, S. J. Sierra, and S. P. Smith, "A derived equivalence for a degree 6 del Pezzo surface over an arbitrary field", *J. K-Theory* **8**:3 (2011), 481–492. MR Zbl

[Colliot-Thélène et al. 2008] J.-L. Colliot-Thélène, N. Karpenko, and A. Merkurjev, "Rational surfaces and the canonical dimension of PGL_6", *St. Petersburg Math. J.* **19**:5 (2008), 793–804. MR Zbl

[Cox et al. 2011] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Mathematics **124**, American Mathematical Society, Providence, 2011. MR Zbl

[Danilov 1978] V. I. Danilov, "The geometry of toric varieties", *Russian Math. Surveys* **33**:2(200) (1978), 97–154. MR Zbl

[Elman et al. 2008] R. Elman, N. Karpenko, and A. Merkurjev, *The algebraic and geometric theory of quadratic forms*, American Mathematical Society Colloquium Publications **56**, American Mathematical Society, Providence, 2008. MR Zbl

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton University Press, Princeton, NJ, 1993. MR Zbl

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics **101**, Cambridge University Press, 2006. MR Zbl

[Hassett 2009] B. Hassett, "Rational surfaces over nonclosed fields", pp. 155–209 in *Arithmetic geometry*, Clay Math. Proc. **8**, Amer. Math. Soc., Providence, 2009. MR Zbl

[Iskovskih 1979] V. A. Iskovskih, "Minimal models of rational surfaces over arbitrary fields", *Izv. Akad. Nauk SSSR Ser. Mat.* **43**:1 (1979), 19–43, 237. In Russian; translated in *Math. USSR-Izvestiya* **24**:2 (1985), 221–244. MR Zbl

[Karpov and Nogin 1998] B. V. Karpov and D. Y. Nogin, "Three-block exceptional sets on del Pezzo surfaces", *Izv. Ross. Akad. Nauk Ser. Mat.* **62**:3 (1998), 3–38. In Russian; translated in *Izv. Math.* **62**:3 (1998), 429–463. MR Zbl

[Keller 2006] B. Keller, "On differential graded categories", pp. 151–190 in *International Congress of Mathematicians* (Madrid, 2006), vol. II, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Klyachko 1992] A. A. Klyachko, "K-theory of demazure models", *Amer. Math. Soc. Transl.* **154**:2 (1992), 37–46. MR Zbl

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications **44**, American Mathematical Society, Providence, 1998. MR Zbl

[Merkurjev 2005] A. S. Merkurjev, "Equivariant K-theory", pp. 925–954 in *Handbook of K-theory. Vol. 1, 2*, edited by E. M. Friedlander and D. R. Grayson, Springer, 2005. MR Zbl

[Merkurjev and Panin 1997] A. S. Merkurjev and I. A. Panin, "K-theory of algebraic tori and toric varieties", *K-Theory* **12**:2 (1997), 101–143. MR Zbl

[Newman 1972] M. Newman, *Integral matrices*, Pure and Applied Mathematics **45**, Academic Press, New York-London, 1972. MR Zbl

[Oda 1978] T. Oda, *Lectures on Torus embeddings and applications*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics **57**, Springer, 1978. MR Zbl

[Orlov 1992] D. O. Orlov, "Projective bundles, monoidal transformations, and derived categories of coherent sheaves", *Izv. Ross. Akad. Nauk Ser. Mat.* **56**:4 (1992), 852–862. In Russian; translated in *Russian Acad. Sci. Izv. Math.* **41**:1 (1993), 133–141. MR Zbl

[Panin 1994] I. A. Panin, "On the algebraic K-theory of twisted flag varieties", *K-Theory* **8**:6 (1994), 541–585. MR Zbl

[Quillen 1973] D. Quillen, "Higher algebraic K-theory, I", pp. 85–147 in *Algebraic K-theory, I: Higher K-theories* (Seattle, 1972), Lecture Notes in Math. **341**, Springer, 1973. MR Zbl

[Serre 1997] J.-P. Serre, *Galois cohomology*, Springer, 1997. MR Zbl

[SGA 6 1971] A. Grothendieck, P. Berthelot, and L. Illusie, *Théorie des intersections et théorème de Riemann–Roch* (Séminaire de Géométrie Algébrique du Bois Marie 1966–1967), Lecture Notes in Math. **225**, Springer, 1971. MR Zbl

[Swan 1985] R. G. Swan, "K-theory of quadric hypersurfaces", *Ann. of Math.* (2) **122**:1 (1985), 113–153. MR Zbl

[Tabuada 2014] G. Tabuada, "Additive invariants of toric and twisted projective homogeneous varieties via noncommutative motives", *J. Algebra* **417** (2014), 15–38. MR Zbl

[Tabuada 2015] G. Tabuada, *Noncommutative motives*, University Lecture Series **63**, American Mathematical Society, Providence, 2015. MR Zbl

[Voskresenskii 1982] V. E. Voskresenskii, "Invariant Demazure models", pp. 3–15 in *Investigations in number theory*, Saratov. Gos. Univ., Saratov, 1982. Russian. MR

[Voskresenskii and Klyachko 1985] V. E. Voskresenskii and A. A. Klyachko, "Toroidal fano varieties and root systems", *Math. USSR Izvestiya* **24**:2 (1985), 221–244. MR Zbl

Received May 5, 2017. Revised November 27, 2017.

FEI XIE
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
BIELEFELD
GERMANY
fxie@math.uni-bielefeld.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 296, No. 2, 2018
dx.doi.org/10.2140/pjm.2018.296.509

CORRECTIONS TO THE ARTICLE THE JOHNSON–MORITA THEORY FOR THE RINGS OF FRICKE CHARACTERS OF FREE GROUPS

TAKAO SATOH

Volume **275**:2 (2015), 443–461

There is a gap in the proof of Proposition 3.1, and it seems to be still an open problem to determine whether the map Φ is surjective or not. In order not to cause any effect on our main theorems, we modify our previous arguments in the following way.

- **page 453, line 1**: We withdraw the statement and the proof of Proposition 3.1.
- **page 455, line 13**: We amend the exact sequence (6) to

$$0 \to \mathrm{Hom}_{\mathbf{Q}}(\mathrm{gr}^1(J), \mathrm{gr}^2(J)) \xrightarrow{\iota} \overline{\mathrm{Aut}}\,(J/J^3) \xrightarrow{\varphi} \mathrm{Im}(\varphi) \to 1.$$

- **page 456, lines 1, 5 and 18; page 458, line 6 from the bottom**: We amend $\overline{\mathrm{Aut}}\,(J/J^3)$ to $\mathrm{Im}(\varphi)$.
- **page 456, line 2; page 458 at the bottom**: We amend the way to choose the elements $\gamma_1, \gamma_2, \ldots, \gamma_{p+q}$ as follows. Let $\gamma_1, \gamma_2, \ldots, \gamma_p$ be all elements in $T_1 \subset J$, and $\gamma_{p+1}, \ldots, \gamma_{p+q}$ all elements in $T_2 \subset J^2$.
- **page 457, line 3**: We complement the information about the map $\mathrm{Aut}\, F_n \to \mathrm{Aut}(J/J^3)$. The image of this map is contained in $\overline{\mathrm{Aut}}(J/J^3)$.

Acknowledgments

The author would like to thank the referee of our forthcoming paper "On the SL(2, $\mathbb{C}$)-representation rings of free abelian group" for pointing out the gaps in the proof of Proposition 3.1.

MSC2010: primary 20F28; secondary 20J06.
Keywords: ring of Fricke characters, automorphism group of free groups, IA-automorphism group, Andreadakis–Johnson filtration, Johnson homomorphisms.

Received December 20, 2017. Revised February 5, 2018.

TAKAO SATOH
DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE DIVISION II
TOKYO UNIVERSITY OF SCIENCE
SHINJUKU-KU
TOKYO
JAPAN
takao@rs.tus.ac.jp

CONTENTS

Volume 296, no. 1 and no. 2

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

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.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 296 No. 2 October 2018

0030-8730(201810)296:2;1-S