Pacific Journal of Mathematics

Volume 277 No. 2 October 2015

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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

EDITORS

Don Blasius (Managing Editor)
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
blasius@math.ucla.edu

Paul Balmer
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
balmer@math.ucla.edu

Vyjayanthi Chari
Department of Mathematics
University of California
Riverside, CA 92521-0135
chari@math.ucr.edu

Daryl Cooper
Department of Mathematics
University of California
Santa Barbara, CA 93106-3080
cooper@math.ucsb.edu

Robert Finn
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
finn@math.stanford.edu

Kefeng Liu
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
liu@math.ucla.edu

Jiang-Hua Lu
Department of Mathematics
The University of Hong Kong
Pokfulam Rd., Hong Kong
jhlu@maths.hku.hk

Sorin Popa
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
popa@math.ucla.edu

Jie Qing
Department of Mathematics
University of California
Santa Cruz, CA 95064
qing@cats.ucsc.edu

Paul Yang
Department of Mathematics
Princeton University
Princeton NJ 08544-1000
yang@math.princeton.edu

PRODUCTION

Silvio Levy, Scientific Editor, production@msp.org

SUPPORTING INSTITUTIONS

ACADEMIA SINICA, TAIPEI
CALIFORNIA INST. OF TECHNOLOGY
INST. DE MATEMÁTICA PURA E APLICADA
KEIO UNIVERSITY
MATH. SCIENCES RESEARCH INSTITUTE
NEW MEXICO STATE UNIV.
OREGON STATE UNIV.
STANFORD UNIVERSITY
UNIV. OF BRITISH COLUMBIA
UNIV. OF CALIFORNIA, BERKELEY
UNIV. OF CALIFORNIA, DAVIS
UNIV. OF CALIFORNIA, LOS ANGELES
UNIV. OF CALIFORNIA, RIVERSIDE
UNIV. OF CALIFORNIA, SAN DIEGO
UNIV. OF CALIF., SANTA BARBARA
UNIV. OF CALIF., SANTA CRUZ
UNIV. OF MONTANA
UNIV. OF OREGON
UNIV. OF SOUTHERN CALIFORNIA
UNIV. OF UTAH
UNIV. OF WASHINGTON
WASHINGTON STATE UNIVERSITY

These supporting institutions contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its contents or policies.

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

The subscription price for 2015 is US $420/year for the electronic version, and $570/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers 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/
© 2015 Mathematical Sciences Publishers

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.257

THE BOREL–WEIL THEOREM FOR REDUCTIVE LIE GROUPS

JOSÉ ARAUJO AND TIM BRATTEN

In this manuscript we consider the extent to which an irreducible representation for a reductive Lie group can be realized as the sheaf cohomology of an equivariant holomorphic line bundle defined on an open invariant submanifold of a complex flag space. Our main result is the following: suppose G_0 is a real reductive group of Harish-Chandra class and let X be the associated full complex flag space. Suppose $\mathcal{O}_\lambda$ is the sheaf of sections of a G_0-equivariant holomorphic line bundle on X whose parameter λ (in the usual twisted $\mathcal{D}$-module context) is antidominant and regular. Let $S \subseteq X$ be a G_0-orbit and suppose $U \supseteq S$ is the smallest G_0-invariant open submanifold of X that contains S. From the analytic localization theory of Hecht and Taylor one knows that there is a nonnegative integer q such that the compactly supported sheaf cohomology groups $H_c^p(S, \mathcal{O}_\lambda|_S)$ vanish except in degree q, in which case $H_c^q(S, \mathcal{O}_\lambda|_S)$ is the minimal globalization of an associated standard Beilinson–Bernstein module. In this study, we show that the q-th compactly supported cohomology group $H_c^q(U, \mathcal{O}_\lambda|_U)$ defines, in a natural way, a nonzero submodule of $H_c^q(S, \mathcal{O}_\lambda|_S)$, which is irreducible (i.e., realizes the unique irreducible submodule of $H_c^q(S, \mathcal{O}_\lambda|_S)$) when an associated algebraic variety is nonsingular. By a tensoring argument, we can show that the result holds, more generally (for nonsingular associated variety), when the representation $H_c^q(S, \mathcal{O}_\lambda|_S)$ is what we call a classifying module.

1. Introduction

In this manuscript we show there is a natural generalization of the Borel–Weil theorem to the class of reductive Lie groups which serves to realize many, but not all, irreducible admissible representations.

Starting with Schmid's thesis [1989], there are general results realizing irreducible representations as sheaf cohomologies of finite-rank holomorphic vector bundles defined over open orbits in generalized complex flag spaces [Wong 1995; Bratten 1998]. However, relatively few irreducible representations can be realized this

---

MSC2010: 22E46.
Keywords: reductive Lie group, representation theory, flag manifold.

way. The equivariant $\mathcal{D}$-module theory of Beilinson and Bernstein [1981] provides a powerful generalization to the Borel–Weil theorem and produces geometric realizations for any irreducible Harish-Chandra module. However, one would also like to find a natural realization of a corresponding group representation. In a general sense, the analytic localization defined by Hecht and Taylor [1990] does just that, by giving realizations for the minimal globalizations [Schmid 1985] of Harish-Chandra modules. Along these lines, the theory of analytic localization was used by Hecht and Taylor to realize minimal globalizations of the standard modules defined by the Beilinson–Bernstein theory. In many cases, standard modules are irreducible, but when they are not, it is not obvious how to proceed. One difficulty is that the geometric realization defined by Hecht and Taylor is obtained via an equivalence of derived categories so that the analytic localization of an irreducible representation can (and sometimes does) appear as a complex of sheaves that has nonzero homologies in various degrees. In spite of this difficulty, it turns out (somewhat surprisingly to us) that the theory of analytic localization can be used to realize many more irreducible representations than the example of irreducible standard modules. In the end, one sees that the key point hinges on whether a certain associated algebraic variety has singularities. When it does not, then the Beilinson–Bernstein realization of the corresponding irreducible Harish-Chandra module has a simple geometric description and this fact controls the analytic localization. When the associated variety is singular, this simplicity breaks down, and it turns out that it is impossible to realize the irreducible representation as the sheaf cohomology of a finite-rank holomorphic vector bundle defined over an invariant open submanifold of a generalized flag manifold.

Rather than make a general statement about our main results in the introduction (our main results are Theorem 5.1 and Corollary 6.2), we would like to illustrate how the theory works in the context of a connected complex reductive group where the relationship to the Beilinson–Bernstein classification of irreducible admissible representations is more transparent. In particular, suppose G_0 is a connected complex reductive group with Lie algebra $\mathfrak{g}_0$ and let $K_0 \subseteq G_0$ be a compact real form. Associated to K_0 is a corresponding Cartan involution $\theta : G_0 \to G_0$ (in this case θ is the conjugation given by the real form). Let X_0 be the complex flag manifold of Borel subgroups of G_0, and let X_0^{c} be the conjugate complex manifold. Then the flag manifold X of Borel subalgebras of the complexified Lie algebra $\mathfrak{g}$ of $\mathfrak{g}_0$ can be identified with the direct product

$$X = X_0 \times X_0^{\mathrm{c}}.$$

We need to consider two actions of G_0 on X. The diagonal action

$$g \cdot (x, y) = (gx, gy)$$

corresponds to the fact that G_0 is a real group with real Lie algebra $\mathfrak{g}_0$ and the action

$$g \cdot (x, y) = (gx, \theta(g)y)$$

corresponds to the fact that $G_0 = K$ is the complexification of K_0. Choose a θ-stable Cartan subgroup $H_0 \subseteq G_0$ and a Borel subgroup $B_0 \supseteq H_0$. Let $W(G_0)$ be the Weyl group of H_0 in G_0. Then we can identify the set of Borel subgroups of G_0 that contain H_0 with $W(G_0)$ (the identity in $W(G_0)$ corresponds to the Borel subgroup B_0). Let B_0^{op} be the Borel subgroup opposite to B_0 with respect to H_0 (this subgroup corresponds to the longest element in $W(G_0)$). Then each G_0-orbit and each K-orbit on X contains exactly one point of the form

$$(w \cdot B_0, B_0^{\mathrm{op}}) \in X_0 \times X_0^{\mathrm{c}}$$

for $w \in W(G_0)$. Thus, the orbits for both actions are simultaneously parametrized by $W(G_0)$. Observe that the open orbit for the G_0-action and the closed orbit for the K-action correspond to the identity in $W(G_0)$ We introduce the length function, $l(w)$, on $W(G_0)$. In particular, each element $w \in W(G_0)$ can be expressed as a product of simple reflections and the corresponding length, $l(w)$, is defined to be the number of simple reflections that appear in a minimal expression (i.e., a reduced word) for w. Observe that if Q_w is the K-orbit corresponding to $w \in W(G_0)$ then the complex dimension of Q_w is given by

$$\dim_{\mathbb{C}}(Q_w) = \dim_{\mathbb{C}}(X_0) + l(w).$$

For simplicity we will consider the sheaf of holomorphic functions $\mathcal{O}_X$ on X (more generally one could consider the sheaf of sections $\mathcal{O}_\lambda$ of a G_0-equivariant holomorphic line bundle on X whose parameter λ in the usual twisted $\mathcal{D}$-module context is antidominant and regular). The Beilinson–Bernstein classification gives a one-to-one correspondence between the equivalence classes of irreducible admissible representations for G_0 that have the same infinitesimal character as the trivial representation and the G_0-orbits on X given in the following way. For $w \in W(G_0)$, let S_w be the corresponding G_0-orbit and define

$$q = \dim_{\mathbb{C}}(X_0) - l(w).$$

Thus, q is the (complex) codimension of the K-orbit Q_w in X. Using their theory of analytic localization, Hecht and Taylor have shown that the compactly supported sheaf cohomologies

$$H_{\mathrm{c}}^p(S_w, \mathcal{O}_X|_{S_w})$$

of the restriction of $\mathcal{O}_X$ to S_w vanish except when $p = q$, in which case the module $H_{\mathrm{c}}^q(S_w, \mathcal{O}_X|_{S_w})$ is the minimal globalization of a corresponding standard Beilinson–Bernstein module. It follows that $H_{\mathrm{c}}^q(S_w, \mathcal{O}_X|_{S_w})$ has a unique irreducible submodule $J_w \subseteq H_c^q(S_w, \mathcal{O}_X|_{S_w})$. These representations J_w for $w \in W(G_0)$ are exactly the irreducible admissible representations for G_0 that have the same infinitesimal character as the trivial representation.

In this manuscript we want to realize the representations J_w. Along those lines we introduce the Bruhat order in $W(G_0)$: if $w, u \in W(G_0)$ then we write $u \preceq w$ if u is an ordered subword that occurs in a reduced expression for w in terms of products of simple reflections. Given $w \in W(G_0)$, it is well known that the Bruhat interval $\Upsilon(w) = \{u \in W(G_0) : u \leq w\}$ characterizes the Zariski closure of the K-orbit Q_w in the following way:

$$\overline{Q_w} = \bigcup_{u \in \Upsilon(wA)} Q_u.$$

We call $\overline{Q_w}$ *the algebraic variety associated to the* G_0*-orbit* S_w. Define

$$U_w = \bigcup_{u \in \Upsilon(w)} S_u.$$

Then, U_w is the smallest G_0-invariant open submanifold of X that contains S_w and it is not hard to show that S_w is the unique G_0-orbit that is closed in U_w. Put $U = U_w - S_w$. Letting $(\mathcal{O}_X|_U)^X$, etc., denote the extension by zero of the restriction of $\mathcal{O}_X$ to U, we obtain the following short exact sequence of sheaves on X:

$$0 \to (\mathcal{O}_X|_U)^X \to (\mathcal{O}_X|_{U_w})^X \to (\mathcal{O}_X|_{S_w})^X \to 0.$$

Using an argument like [Bratten 2008, Lemma 3.3], it is not hard to show that

$$H_c^p(U, \mathcal{O}_X|_U) = 0 \quad \text{if } p < q+1$$

and that

$$H_c^q(U_w, \mathcal{O}_X|_{U_w})$$

is a nonzero minimal globalization. Thus, the long exact sequence in sheaf cohomology determines an inclusion

$$H_c^q(U_w, \mathcal{O}_X|_{U_w}) \hookrightarrow H_c^q(S_w, \mathcal{O}_X|_{S_w}).$$

We note that when U_w is the preimage of an open G_0-orbit on a generalized flag space Y then there is a natural identification of the representation $H_c^q(U_w, \mathcal{O}_X|_{U_w})$ with the q-th compactly supported cohomology of the holomorphic functions on the given open orbit in Y. (This is one of the key points in [op. cit.].) When this happens, it is known that the representation $H_c^q(U_w, \mathcal{O}_X|_{U_w})$ is irreducible. We say that U_w is *parabolic* when U_w is the preimage of an open G_0-orbit on a generalized flag space Y. Our main result in this study shows that (more generally) the submodule $H_c^q(U_w, \mathcal{O}_X|_{U_w})$ is irreducible when the associated algebraic variety $\overline{Q_w}$ is smooth.

Thus we can realize J_w as $H_c^q(U_w, \mathcal{O}_X|_{U_w})$ when (and, in fact, only when) this happens.

For example, if G_0 is the complex general linear group $\mathrm{GL}(3, \mathbb{C})$, then 4 out of the 6 G_0-orbits on X are parabolic but the algebraic varieties associated to all 6 orbits are smooth so we can realize all irreducible representations with the given infinitesimal character in this case. If $G_0 = \mathrm{GL}(4, \mathbb{C})$, then only 8 of the 24 G_0-orbits are parabolic, but 22 out of 24 orbits have smooth associated varieties so we can realize all but two of the irreducible representations with the given infinitesimal character (and so on). We will also see (for some examples) that when the algebraic variety $\overline{Q_w}$ is singular the representation $H_c^q(U_w, \mathcal{O}_X|_{U_w})$ is reducible, and it is actually impossible to realize the irreducible representation J_w as the compactly supported sheaf cohomology of an equivariant (finite-rank) holomorphic vector bundle defined on a G_0-invariant open submanifold in a generalized flag space.

Our manuscript is organized as follows. In Section 2, we will present the main results we need about orbits and invariant subspaces in X. In Section 3, we will introduce the equivariant homogeneous line bundles and prove the basic embedding theorem. In Section 4, we introduce the algebraic localization theory and give a geometric description to the irreducible Harish-Chandra module in the Beilinson–Bernstein classification, assuming the corresponding algebraic variety is smooth. In Section 5, we introduce the analytic localization and use the comparison theorem to prove our main result. Then, in Section 6, we use a tensoring argument to extend our result to antidominant parameters and also consider how our construction relates to the classical parabolic induction (in the case of a complex reductive group) so we can consider some examples. We conclude our manuscript with a brief consideration of how Serre duality applies. We would like to mention that the idea of our proof involves a mix of ideas from the two articles [Bratten 2008; 1997]. Although our argument requires a heavy use of the D-module theory and some familiarity with derived categories, we would hope it looks natural to anyone familiar with these two previous articles.

2. G_0-orbits and K-orbits

Throughout this manuscript G_0 will denote a real reductive Lie group of Harish-Chandra class with Lie algebra $\mathfrak{g}_0$ and complexified Lie algebra $\mathfrak{g}$. Abusing notation a bit, we let G denote the complex adjoint group of $\mathfrak{g}$. (Note that G has Lie algebra $[\mathfrak{g}, \mathfrak{g}]$.) There is a natural morphism of Lie groups

$$G_0 \to G.$$

We also fix a maximal compact subgroup $K_0 \subseteq G_0$ and let K be the complexification of K_0. Associated to the maximal compact subgroup, there is an involutive automorphism

$$\theta : G_0 \to G_0,$$

whose fixed point set is K_0. The involution θ (as well as the complexification $\theta : \mathfrak{g} \to \mathfrak{g}$ of its derivative) is called the *Cartan involution*.

A *Borel subalgebra* of $\mathfrak{g}$ is a maximal solvable subalgebra. G acts transitively on the set of Borel subalgebras of $\mathfrak{g}$ and the resulting homogeneous G-space X is a complex projective variety called the *full flag space* of G_0. Since G_0 has finitely many orbits on X [Wolf 1969], the G_0-orbits are locally closed submanifolds.

A basic geometric property of flag manifolds that is fundamental to our study is the existence of a one-to-one correspondence between G_0-orbits and K-orbits referred to as *Matsuki duality*. For $x \in X$ we let $\mathfrak{b}_x$ denote the corresponding Borel subalgebra of $\mathfrak{g}$. Then, the *nilradical* $\mathfrak{n}_x$ of $\mathfrak{b}_x$ is given by $\mathfrak{n}_x = [\mathfrak{b}_x, \mathfrak{b}_x]$. The point $x \in X$ (as well as the Borel subalgebra $\mathfrak{b}_x$) is called *special* if there is a Cartan subalgebra $\mathfrak{c}$ of $\mathfrak{b}_x$ such that

$$\mathfrak{c}_0 = \mathfrak{g}_0 \cap \mathfrak{c} \text{ is a real form of } \mathfrak{c} \quad \text{and} \quad \theta(\mathfrak{c}) = \mathfrak{c}.$$

Matsuki [1979] showed that both the special points in a G_0-orbit as well as the special points in a K-orbit form a (nonempty) K_0-orbit. A G_0-orbit S and a K-orbit Q are said to be *Matsuki dual* if $S \cap Q$ contains a special point. It follows that Matsuki duality defines a bijection between the set of G_0-orbits and the set of K-orbits. When S is a G_0-orbit and Q is a K-orbit then we will write $S \sim Q$ when S and Q are Matsuki dual.

Given a K-orbit Q, the *associated algebraic variety* is defined to be the Zariski closure $\overline{Q}$ of Q (when S is the G_0-orbit dual Q, we will also refer to $\overline{Q}$ as the algebraic variety associated to S). There is a partial order, called the *closure order*, defined on the set of K-orbits by

$$Q_1 \preceq Q_2 \quad \text{if } Q_1 \subseteq \overline{Q_2}.$$

While the associated algebraic variety is a closed K-invariant subvariety of X associated to a G_0-orbit S, we will now define a corresponding G_0-invariant open submanifold of X. In particular, suppose S_0 is a G_0-orbit and let Q_0 be the Matsuki dual. We define an associated indexing set $\Upsilon(S_0)$ of G_0-orbits by

$$S \in \Upsilon(S_0) \quad \Longleftrightarrow \quad \exists Q \text{ such that } Q \sim S \text{ and } Q \subseteq \overline{Q_0}.$$

We define the *corresponding G_0-invariant subspace U* by

$$U = \bigcup_{S \in \Upsilon(S_0)} S.$$

Proposition 2.1. *With the previous notation, U is the smallest G_0-invariant open submanifold that contains S_0, and S_0 is the unique closed G_0-orbit in U.*

Proof. We consider the closure orders on the set of K-orbits and on the set of G_0-orbits. Matsuki [1988] showed that duality reverses the corresponding closure relations. It follows that

$$S \in \Upsilon(S_0) \iff S_0 \subseteq \overline{S} \iff S_0 \preceq S.$$

Thus, if $S \notin \Upsilon(S_0)$, then $S_0 \cap \overline{S} = \varnothing$; therefore, $S_0 \cap \overline{S_1} = \varnothing$ for each G_0-orbit S_1 contained in $\overline{S}$. Hence,

$$U \cap \overline{S} = \varnothing.$$

Since there are a finite number of orbits, the set

$$C = \bigcup_{S \notin \Upsilon(S_0)} \overline{S}$$

is closed and therefore

$$U = X - C$$

is open.

Now, suppose that W is an open G_0-invariant submanifold that contains S_0. Suppose $S \subseteq U$. Then $S_0 \subseteq \overline{S}$. Thus,

$$\overline{S} \cap W \neq \varnothing.$$

Hence, $\overline{S} \cap W$ is a nonempty open G_0-invariant subset of $\overline{S}$. Since S is locally closed, it follows that S is open and dense in $\overline{S}$. Hence, from the G_0-invariance,

$$S \subseteq \overline{S} \cap W \implies S \subseteq W$$

so that $U \subseteq W$, which proves that U is the smallest G_0-invariant open submanifold that contains S_0.

To prove the last claim, first observe that if S is a G_0-orbit contained in $\overline{S_0}$ then $S \preceq S_0$ so that $S \subseteq U$ if and only if $S = S_0$. Thus,

$$\overline{S_0} \cap U = S_0$$

and S_0 is closed in U. On the other hand, if S is a closed G_0-orbit in U, then

$$S = \overline{S} \cap U.$$

However, from the definition of U, for $S \subseteq U$, we have $S_0 \subseteq \overline{S}$, hence

$$S_0 \subseteq \overline{S} \cap U.$$

It follows that $S = \overline{S} \cap U$ if and only if $S = S_0$. □

Example 2.2. Suppose that $\mathfrak{p} \subseteq \mathfrak{g}$ is a parabolic subalgebra and let Y be the corresponding G-homogeneous space of parabolic subalgebras of $\mathfrak{g}$ conjugate to $\mathfrak{p}$. For each $y \in Y$, let $\mathfrak{p}_y$ denote the corresponding parabolic subalgebra of $\mathfrak{g}$. For $x \in X$, there is a unique $y \in Y$ such that $\mathfrak{b}_x \subseteq \mathfrak{p}_y$. Thus there is a canonical G-equivariant projection

$$\pi : X \to Y$$

given by $\pi(x) = y$ if $\mathfrak{b}_x \subseteq \mathfrak{p}_y$. A point $y \in Y$ is called *special* if $\mathfrak{p}_y$ contains a special Borel subalgebra.

Suppose $W \subseteq Y$ is an open G_0-orbit and let $y \in W$ be a special point. Let O be the K-orbit of y. Then, O is closed in Y (in fact $O \subseteq W$), and Matsuki [1982] showed that the G_0-orbits in $U = \pi^{-1}(W)$ are Matsuki dual to the K-orbits in $\pi^{-1}(O)$. Also, there is a unique G_0-orbit S_0 that is closed in U and its Matsuki dual Q_0 is the unique open orbit in the closed algebraic variety $\pi^{-1}(O)$. Therefore, $\overline{Q_0} = \pi^{-1}(O)$ and it follows that a G_0-orbit S is contained in U if and only if its Matsuki dual Q is contained in $\overline{Q_0}$; thus, U is the smallest G_0-invariant open submanifold that contains S_0. Observe that, in this case, the associated algebraic variety $\overline{Q_0} = \pi^{-1}(O)$ is smooth since the fibers of π are smooth and since π defines a locally trivial algebraic fibration of $\pi^{-1}(O)$ over O.

3. Equivariant line bundles

In this section, we introduce the equivariant line bundles on the full flag space X, as well as the corresponding standard modules associated to a G_0-orbit $S \subseteq X$. We begin this section by introducing the abstract Cartan dual, which is the parameter set for the twisted sheaves of differential operators (TDOs) on X. Recall that G is the complex adjoint group of $\mathfrak{g}$. For $x \in X$, let $\mathfrak{n}_x$ denote the nilradical of the corresponding Borel subalgebra $\mathfrak{b}_x$ and put

$$\mathfrak{h}_x = \mathfrak{b}_x/\mathfrak{n}_x.$$

Since the stabilizer of x in G (i.e., the corresponding Borel subgroup in G) acts trivially on $\mathfrak{h}_x$ it also acts trivially on the complex dual $\mathfrak{h}_x^*$. It follows that the corresponding G-homogeneous holomorphic vector bundle on X is trivial and that the associated space of global sections $\mathfrak{h}^*$ is naturally isomorphic to $\mathfrak{h}_x^*$ via the evaluation at x. The vector space $\mathfrak{h}^*$ is called the *abstract Cartan dual* for $\mathfrak{g}$. If $\mathfrak{c}$ is a Cartan subalgebra of $\mathfrak{b}_x$, then by coupling the natural projection of $\mathfrak{c}$ onto $\mathfrak{h}_x$ with the evaluation at x, we obtain an isomorphism of $\mathfrak{c}^*$ with $\mathfrak{h}^*$ called the *specialization* of $\mathfrak{h}^*$ to $\mathfrak{c}^*$ at x. Using the specializations, we can identify an abstract set of roots

$$\Sigma \subseteq \mathfrak{h}^*$$

and an abstract set of positive roots

$$\Sigma^+ \subseteq \Sigma,$$

where Σ corresponds to the set of roots of $\mathfrak{c}$ in $\mathfrak{g}$ and Σ^+ corresponds to the roots of $\mathfrak{c}$ in $\mathfrak{b}_x$, via the specialization at x. Given $\alpha \in \Sigma$ and $\mu \in \mathfrak{h}^*$, we can also define the complex number

$$\overset{\vee}{\alpha}(\mu),$$

the *value of* μ *on the coroot of* α. The element $\mu \in \mathfrak{h}^*$ is called *integral* if

$$\overset{\vee}{\alpha}(\mu) \in \mathbb{Z} \quad \text{for each } \alpha \in \Sigma.$$

It just so happens that the half-sum of positive roots, denoted by ρ, is an integral element of $\mathfrak{h}^*$ that plays a key role in the TDO parametrization.

Let $\widetilde{G}$ denote the universal cover of G and suppose $\mu \in \mathfrak{h}^*$ is integral. For a point $x \in X$, the Lie algebra of the corresponding Borel subgroup $\widetilde{B}_x$ in $\widetilde{G}$ (i.e., the stabilizer of x in $\widetilde{G}$) is given by $\mathfrak{b}_x \cap [\mathfrak{g}, \mathfrak{g}]$. Thus, using the evaluation at x, the global section μ determines a one-dimensional representation

$$\mathfrak{b}_x \cap [\mathfrak{g}, \mathfrak{g}] \longrightarrow \mathfrak{h}_x \xrightarrow{\mu_x} \mathbb{C}.$$

Since $\widetilde{G}$ is simply connected, it is known that there is a (unique) holomorphic character of $\widetilde{B}_x$ whose derivative is given, in this way, by μ_x. Thus, corresponding to each integral $\mu \in \mathfrak{h}^*$ there is a corresponding $\widetilde{G}$-homogeneous holomorphic line bundle

$$\mathbb{L}(\mu) \longrightarrow X.$$

Let $\mathcal{O}(\mu)$ be the corresponding sheaf of holomorphic sections. In a natural way, $\widetilde{G}$ and thus $[\mathfrak{g}, \mathfrak{g}]$ act on $\mathcal{O}(\mu)$. Let $\mathfrak{z}$ be the center of $\mathfrak{g}$. Suppose $W \subseteq X$ is an open set, and let

$$\sigma : W \to \mathbb{L}(\mu)$$

be a local holomorphic section. Then, extend $\mathcal{O}(\mu)$ to a sheaf of $\mathfrak{g}$-modules by

$$(\xi \cdot \sigma)(x) = \mu(\xi)\sigma(x) \quad \text{for } \xi \in \mathfrak{z} \text{ and } x \in W.$$

We say that $\mathbb{L}(\mu)$ is a G_0*-equivariant line bundle* if there exists a G_0-action on $\mathbb{L}(\mu)$ (in the sense of differentiable G_0-actions on vector bundles over differentiable G_0-spaces) such that the induced morphisms $\mathbb{L}(\mu) \to \mathbb{L}(\mu)$, given by multiplication by group elements, are holomorphic and such that the derivative of the G_0-action on local sections coincides with the $\mathfrak{g}$-action.

Example 3.1. An important class of G_0-equivariant holomorphic line bundles corresponds to the family of (equivalence classes of) finite-dimensional irreducible representations of G_0 that are also irreducible for the corresponding $\mathfrak{g}$-action (i.e., irreducible finite-dimensional $\mathfrak{g}$-modules with a compatible G_0-action). In particular, let V be a finite-dimensional G_0-module that is irreducible as a $\mathfrak{g}$-module. For each $x \in X$, let $G_0[x]$ denote the stabilizer of x and consider the corresponding $(\mathfrak{b}_x, G_0[x])$-module

$$V/\mathfrak{n}_x V.$$

Choosing a Cartan subalgebra $\mathfrak{c} \subseteq \mathfrak{b}_x$ and using the specialization to x, the action of $\mathfrak{c}$ on $V/\mathfrak{n}_x V$ is given by an element $\mu \in \mathfrak{h}^*$ that corresponds to the lowest weight in V. Hence, if we define

$$\lambda = \mu - \rho$$

then $\overset{\vee}{\alpha}(\lambda)$ is a negative integer, for each positive root $\alpha \in \Sigma^+$. We can define the total space of a G_0-equivariant holomorphic line bundle $\mathbb{L}_\lambda$ by

$$\mathbb{L}_\lambda = \bigcup_{x \in X} V/\mathfrak{n}_x V.$$

Using the action of $\widetilde{G}$ one can define holomorphic transition functions. Then the Borel–Weil Theorem says that the representation V is recovered as the global holomorphic sections of the bundle $\mathbb{L}_\lambda$.

In general,, when $\mathcal{O}(\mu)$ is the sheaf of holomorphic sections of G_0-equivariant line bundle we will use the shifted parameter $\lambda = \mu - \rho$ and write

$$\mathcal{O}_\lambda = \mathcal{O}(\mu)$$

for the sheaf of holomorphic sections. We say λ is *regular* if

$$\overset{\vee}{\alpha}(\lambda) \neq 0 \quad \text{for each root } \alpha \in \Sigma.$$

An element λ is called *singular* when it is not regular. We say λ is *antidominant* if

$$\overset{\vee}{\alpha}(\lambda) \notin \mathbb{N} \quad \text{for each positive root } \alpha \in \Sigma^+.$$

Suppose that $\mathcal{O}_\lambda$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle. Let S be a G_0-orbit in X and let Q be the Matsuki dual to S. We define the *vanishing number* q of S to be the (complex) codimension of Q in X. Suppose that λ is antidominant and regular. One of the main results of the Hecht–Taylor analytic localization theory is that the compactly supported sheaf cohomology of the restriction $\mathcal{O}_\lambda|_S$ of $\mathcal{O}_\lambda$ to S vanishes, except in degree q, in which case

$$H^q_c(S, \mathcal{O}_\lambda|_S)$$

is the minimal globalization of a corresponding standard Beilinson–Bernstein module. (We will describe this module in the following section.) In particular,

$H_c^q(S, \mathcal{O}_\lambda|_S)$ has a unique irreducible submodule. In general, (for any parameter λ), the sheaf cohomology groups

$$H_c^p(S, \mathcal{O}_\lambda|_S)$$

vanish [Bratten 1997] for $p < q$, and in the nonzero cases these cohomology groups are minimal globalizations of the sheaf cohomology groups of an associated standard Harish-Chandra sheaf.

In general, terms, we now consider a simple geometric procedure which can be used in the context of the Hecht–Taylor realization of minimal globalizations, to study representations. We first remark that the sheaves $\mathcal{O}_\lambda$ are examples of what is referred to in the Hecht–Taylor development as DNF (stands for dual nuclear Fréchet) sheaves of analytic G_0-modules (we will not need the formalism of DNF sheaves of analytic G_0-modules in our study however we would simply like to mention the general criteria used to establish the following results). We want to remark that, in the case of the global sheaf cohomology on X, the sheaf cohomology groups

$$H^p(X, \mathcal{O}_\lambda)$$

are finite-dimensional and were originally studied in [Bott 1957]. Now suppose $L \subseteq X$ is a locally closed G_0-invariant subspace and let

$$(\mathcal{O}_\lambda|_L)^X$$

denote the extension by zero to X of the restriction of $\mathcal{O}_\lambda$ to L. Then there is a natural isomorphism of functors

$$H^p(X, (\mathcal{O}_\lambda|_L)^X) \cong H_c^p(L, \mathcal{O}_\lambda|_L)$$

and it follows from the results of Hecht and Taylor (at least for λ regular — in the singular case one can prove this by a tensoring argument as in [Bratten 1997]) that the sheaf cohomology groups

$$H_c^p(L, \mathcal{O}_\lambda|_L)$$

are minimal globalizations of Harish-Chandra modules. Let $W \subseteq L$ be an open, G_0-invariant subspace and let $C = L - W$. Then we have the following short exact sequence of DNF sheaves of analytic G_0-modules:

$$0 \to (\mathcal{O}_\lambda|_W)^X \to (\mathcal{O}_\lambda|_L)^X \to (\mathcal{O}_\lambda|_C)^X \to 0.$$

Therefore, the corresponding long exact sequence in cohomology

$$\cdots \to H_c^p(W, \mathcal{O}_\lambda|_W) \to H_c^p(L, \mathcal{O}_\lambda|_L) \to H_c^p(C, \mathcal{O}_\lambda|_C) \to H_c^{p+1}(W, \mathcal{O}_\lambda|_W) \to \cdots$$

is a sequence of minimal globalizations with continuous G_0-morphisms.

Return to the case where S is a G_0-orbit and let U be the smallest G_0-invariant open set that contains S. Then the compactly supported sheaf cohomology groups

$$H_c^p(U, \mathcal{O}_\lambda|_U)$$

are minimal globalizations of Harish-Chandra modules. Recall that q is the vanishing number of S. We will now show that there is a natural embedding of $H_c^q(U, \mathcal{O}_\lambda|_U)$ in $H_c^q(S, \mathcal{O}_\lambda|_S)$. Let $W = U - S$. Then W is open and we have the following short exact sequence of sheaves on X:

$$0 \to (\mathcal{O}_\lambda|_W)^X \to (\mathcal{O}_\lambda|_U)^X \to (\mathcal{O}_\lambda|_S)^X \to 0.$$

This sequence will induce a sequence of continuous morphisms of minimal globalizations when we apply the long exact sequence of sheaf cohomology. To prove we have an inclusion in grade q we use the vanishing on S, and the following lemma.

Lemma 3.2. *Maintain the previously defined notations. Then,*

$$H_c^p(W, \mathcal{O}_\lambda|_W) = 0 \quad \textit{for } p \leq q.$$

Proof. First observe that since Q is open in $\overline{Q}$ then for each $Q_0 \subseteq \overline{Q}$ such that $Q_0 \neq Q$ then the codimension of Q_0 is strictly bigger than the codimension of Q, so that the vanishing numbers for G_0-orbits in W are at least $q+1$. Suppose $O \subseteq W$ is a G_0-invariant open subset. We define the length of O to be the number of G_0-orbits contained in O. We show that the announced vanishing result holds for every G_0-invariant open subset of W by an induction on length. When O has length one, it is an open G_0-orbit so the result holds since it has vanishing number at least $q+1$. In general, let $S_0 \subseteq O$ be a G_0-orbit of minimal dimension. Since S_0 is open and dense in its closure it follows that any G_0-orbit in the closure different from S_0 has strictly smaller dimension. Thus,

$$\overline{S_0} \cap O = S_0$$

and $O_0 = O - S_0$ is an open G_0-invariant of shorter length. Thus, the result follows by induction, using the long exact sequence in cohomology applied to the short exact sequence of sheaves:

$$0 \to (\mathcal{O}_\lambda|_{O_0})^X \to (\mathcal{O}_\lambda|_O)^X \to (\mathcal{O}_\lambda|_{S_0})^X \to 0. \qquad \square$$

Corollary 3.3. *Let $\mathcal{O}_\lambda$ be the sheaf of sections of a G_0-equivariant holomorphic line bundle over X. Suppose $S \subseteq X$ is G_0-orbit with associated vanishing number q and let U be the smallest G_0-invariant open submanifold that contains S. Then there is a natural inclusion of analytic G_0-modules*

$$H_c^q(U, \mathcal{O}_\lambda|_U) \to H_c^q(S, \mathcal{O}_\lambda|_S)$$

4. Localization and standard Beilinson–Bernstein modules

We begin this section by introducing the sheaves of twisted differential operators and reviewing some of the necessary theory. For a basic reference on the algebraic side of the localization theory, we note that [Miličić 1993] provides a nice overview to the geometric realization of Harish-Chandra modules given by the Beilinson–Bernstein theory.

Let $U(\mathfrak{g})$ be the enveloping algebra of $\mathfrak{g}$ and let $Z(\mathfrak{g})$ be the center of $U(\mathfrak{g})$. An *infinitesimal character* Θ is a morphism of algebras (with identity)

$$\Theta : Z(\mathfrak{g}) \to \mathbb{C}.$$

We let U_Θ be the quotient of $U(\mathfrak{g})$ by the two-sided generated from the kernel of Θ in $Z(\mathfrak{g})$. Observe that U_Θ is the algebra that acts naturally on a $\mathfrak{g}$-module with infinitesimal character Θ.

Let W be the Weyl group of $\mathfrak{h}^*$. By Harish-Chandra's classical result, $Z(\mathfrak{g})$ is isomorphic to the Weyl group invariants in the enveloping algebra of a Cartan subalgebra of $\mathfrak{g}$. It follows that the infinitesimal characters are naturally parametrized by the W-orbits in $\mathfrak{h}^*$. For $\lambda \in \mathfrak{h}^*$, we write

$$\Theta = W\lambda \quad \text{and} \quad \lambda \in \Theta$$

when the W orbit of λ parametrizes the infinitesimal character Θ. It is known that if $\lambda \in \mathfrak{h}^*$ is integral (or regular) then $w\lambda$ is integral (or regular) for every $w \in W$. In this case we also say that the corresponding infinitesimal character is *integral* (or *regular*). When an infinitesimal character Θ is integral and regular then there exists a unique $\lambda \in \Theta$ that is antidominant. Notice that the infinitesimal character of an irreducible admissible representation is an important invariant and that the Beilinson–Bernstein realization of irreducible Harish-Chandra modules with infinitesimal character Θ depends, to some extent, on the choice of $\lambda \in \Theta$.

At this point we need to distinguish between the algebraic and analytic structures on X. Therefore, we consider the full flag space X as both an algebraic variety (with the Zariski topology) and as a complex manifold (with the analytic topology) according to the context. Since the line bundles, defined for integral $\lambda \in \mathfrak{h}^*$ in the previous section, have a compatible algebraic structure, we can consider the corresponding sheaf of algebraic sections $\mathcal{O}_\lambda^{\text{alg}}$ defined on the algebraic variety X. Associated to the sheaf $\mathcal{O}_\lambda^{\text{alg}}$ is a corresponding twisted sheaf of differential operators (TDO) $\mathcal{D}_\lambda^{\text{alg}}$.

We can also consider the corresponding TDO $\mathcal{D}_\lambda$, with holomorphic coefficients and defined on complex variety X. In a natural way, $\mathcal{O}_\lambda$ is a sheaf of modules for $\mathcal{D}_\lambda$. When $\mathcal{O}_\lambda$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle, G_0 acts on $\mathcal{D}_\lambda$ while K acts compatibly (extending the K_0-action) on $\mathcal{O}_\lambda^{\text{alg}}$ and

$\mathcal{D}_\lambda^{\text{alg}}$. We want to emphasize that we are using the shifted parametrization from the previous section. In particular, $\mathcal{O}_{-\rho}$ is the sheaf of holomorphic functions on X and $\mathcal{D}_{-\rho}$ is the sheaf of holomorphic differential operators.

Suppose that $\lambda \in \Theta$. Beilinson and Bernstein showed that

$$\Gamma(X, \mathcal{D}_\lambda^{\text{alg}}) \cong U_\Theta \quad \text{and} \quad H^p(X, \mathcal{D}_\lambda^{\text{alg}}) = 0 \text{ for } p > 0.$$

Thus, the sheaf cohomology groups of a sheaf of $\mathcal{D}_\lambda^{\text{alg}}$-modules are $\mathfrak{g}$-modules with infinitesimal character Θ. When $\mathcal{F}$ is a sheaf of quasicoherent $\mathcal{D}_\lambda^{\text{alg}}$-modules and λ is antidominant, Beilinson and Bernstein showed that

$$H^p(X, \mathcal{F}) = 0 \text{ for } p > 0.$$

In particular, when λ is antidominant, the functor of global sections is exact on the category of quasicoherent $\mathcal{D}_\lambda^{\text{alg}}$-modules. The *localization functor*

$$\Delta_\lambda(M) = \mathcal{D}_\lambda^{\text{alg}} \otimes_{U_\Theta} M$$

is defined on the category of U_Θ-modules and is the left adjoint to the global sections functor on the category of quasicoherent $\mathcal{D}_\lambda^{\text{alg}}$-modules. When λ is antidominant and regular, Beilinson and Bernstein showed that the localization functor and the global sections functor are mutual inverses and determine an equivalence of categories.

Suppose $S \subseteq X$ is a G_0-orbit with associated vanishing number q. Let $\mathcal{O}_\lambda$ be the sheaf of holomorphic sections of a G_0-equivariant line bundle. We now consider the geometric construction of the underlying Harish-Chandra module of the minimal globalization

$$H_{\text{c}}^q(S, \mathcal{O}_\lambda|_S).$$

Let Q denote the K-orbit Matsuki dual to S (equip Q with the Zariski topology) and consider the K-equivariant sheaf $\mathcal{O}_\lambda^{\text{alg}}$. Let $i : Q \hookrightarrow X$ be the inclusion and let

$$i^*(\mathcal{O}_\lambda^{\text{alg}}) = i^{-1}(\mathcal{O}_\lambda^{\text{alg}}) \otimes_{i^{-1}(\mathcal{O}_X^{\text{alg}})} \mathcal{O}_Q^{\text{alg}}$$

be the inverse image of $\mathcal{O}_\lambda^{\text{alg}}$ with respect to the structure sheaves of the algebraic varieties Q and X. Therefore, $i^*(\mathcal{O}_\lambda^{\text{alg}})$ is the sheaf of sections of a corresponding K-homogeneous algebraic line bundle defined on Q. Let $\mathcal{D}_{Q,\lambda}^{\text{alg}}$ be the sheaf of differential operators for the locally free sheaf $i^*(\mathcal{O}_\lambda^{\text{alg}})$. Then, there is a corresponding *direct image* functor i_+ in the category of sheaves of (twisted) $\mathcal{D}$-modules. In this case, since the morphism i is an affine inclusion of smooth varieties, the direct image is an exact functor [Hecht et al. 1987]. The sheaf

$$\mathcal{I}(Q, \lambda) = i_+ i^*(\mathcal{O}_\lambda^{\text{alg}})$$

is a K-equivariant sheaf of $\mathcal{D}_\lambda^{\text{alg}}$-modules called the *corresponding standard Harish-Chandra sheaf* on X, which contains a unique (coherent and K-invariant) irreducible subsheaf

$$\mathcal{J}(Q,\lambda) \subseteq \mathcal{I}(Q,\lambda).$$

of $\mathcal{D}_\lambda^{\mathrm{alg}}$-modules. We note that the notation being used is a bit ambiguous since the structure of these objects also depends on the K_0-action and not just the orbit Q and the integral parameter λ. However, in the current context it should be clear how one construction leads to the other and we feel our approach avoids an overly complicated notation. The Harish-Chandra module

$$I(Q,\lambda) = \Gamma(X, i_+ i^*(\mathcal{O}_\lambda^{\mathrm{alg}}))$$

is called the *corresponding standard Beilinson–Bernstein module*. In general, one knows [Bratten 1997] that $I(Q,\lambda)$ is the underlying Harish-Chandra module of $H_c^q(S, \mathcal{O}_\lambda|_S)$. Observe that when λ is antidominant and regular, by the equivalence of categories, it follows that the Harish-Chandra module

$$J(Q,\lambda) = \Gamma(X, \mathcal{J}(Q,\lambda)) \subseteq I(Q,\lambda)$$

is the unique irreducible submodule of the corresponding standard Beilinson–Bernstein module. We call $I(Q,\lambda)$ a *classifying module* if λ is antidominant and $J(Q,\lambda) \neq 0$. As the name suggests, the classifying modules are used in Beilinson–Bernstein classification of irreducible admissible representations. This works perfectly when G_0 is a connected, complex reductive group, however, in general, one must enlarge the class of standard Harish-Chandra sheaves to include all irreducible representations with the given integral infinitesimal character.

Suppose λ is antidominant and regular and let $J(Q,\lambda)_{\min}$ denote the corresponding minimal globalization.

Proposition 4.1. *Maintain the above notations* (*in particular*, *we assume that* λ *is antidominant and regular*). *Let* U *be the smallest* G_0*-invariant open set that contains* S. *Then*, *there exists a natural inclusion*

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^q(U, \mathcal{O}_\lambda|_U).$$

Proof. Since $I(Q,\lambda)$ is the underlying Harish-Chandra module of $H_c^q(S, \mathcal{O}_\lambda|_S)$, it follows that $J(Q,\lambda)_{\min}$ is the unique irreducible submodule in $H_c^q(S, \mathcal{O}_\lambda|_S)$. Therefore, to establish the result it suffices to show that

$$H_c^q(U, \mathcal{O}_\lambda|_U) \neq 0.$$

We follow the setup used in Lemma 3.2 and argue by contradiction. Suppose that $H_c^q(U, \mathcal{O}_\lambda|_U) = 0$ and let $W = U - S$. Using the long exact sequence in cohomology, we obtain an inclusion

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^q(S, \mathcal{O}_\lambda|_S) \hookrightarrow H_c^{q+1}(W, \mathcal{O}_\lambda|_W).$$

Now, suppose that $S_1 \subseteq W$ is a G_0-orbit of minimal dimension and define $W_1 = W - S$. Then S_1 is closed in W_1, so using the corresponding long exact sequence in cohomology, we obtain the sequence

$$0 \to H_c^{q+1}(W_1, \mathcal{O}_\lambda|_{W_1}) \to H_c^{q+1}(W, \mathcal{O}_\lambda|_W) \to H_c^{q+1}(S_1, \mathcal{O}_\lambda|_{S_1}) \to \cdots .$$

Since $J(Q,\lambda)_{\min}$, is irreducible either

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^{q+1}(W_1, \mathcal{O}_\lambda|_{W_1}) \quad \text{or} \quad J(Q,\lambda)_{\min} \hookrightarrow H_c^{q+1}(S_1, \mathcal{O}_\lambda|_{S_1}).$$

However, if $H_c^{q+1}(S_1, \mathcal{O}_\lambda|_{S_1}) \neq 0$, then this representation is the minimal globalization of the classifying module $I(Q_1,\lambda)$ where Q_1 is the K-orbit Matsuki dual to S_1. Thus, $J(Q_1,\lambda)_{\min}$ is the unique irreducible submodule of $H_c^{q+1}(S_1, \mathcal{O}_\lambda|_{S_1})$. Since $Q_1 \neq Q$, it follows that $J(Q,\lambda)$ is not isomorphic to $J(Q_1,\lambda)$. Therefore,

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^{q+1}(W_1, \mathcal{O}_\lambda|_{W_1}).$$

Proceeding in this fashion, we would obtain that

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^{q+1}(O, \mathcal{O}_\lambda|_O),$$

where O is an open G_0-orbit. However, this is impossible since $H_c^{q+1}(O, \mathcal{O}_\lambda|_O)$ is either zero or an irreducible minimal globalization that is not isomorphic to $J(Q,\lambda)_{\min}$. □

The proof of our main result now consists of two steps. The first part is to characterize the irreducible Harish-Chandra sheaf $\mathcal{J}(Q,\lambda)$ when the associated variety $\bar{Q}$ is smooth. Once we have that in hand, it turns out to be fairly straightforward to calculate the analytic localization of $J(Q,\lambda)_{\min}$ on G_0-orbits. To finish the proof we show that the inclusion

$$J(Q,\lambda)_{\min} \hookrightarrow H_c^{q}(U, \mathcal{O}_\lambda|_U)$$

induces an isomorphism between the analytic localization of $J(Q,\lambda)_{\min}$ and the sheaf

$$(\mathcal{O}_\lambda|_U)^X.$$

We can then recover our main result by the Hecht–Taylor equivalence of derived categories.

For the first step of our proof, we continue with the previous notation.

Let

$$j : \bar{Q} \hookrightarrow X$$

denote the inclusion and assume $\bar{Q}$ is smooth. We consider the K-equivariant sheaf $j^*(\mathcal{O}_\lambda^{\text{alg}})$ defined on $\bar{Q}$. Notice that $j^*(\mathcal{O}_\lambda^{\text{alg}})$ is the sheaf of sections of a K-equivariant algebraic line bundle defined on $\bar{Q}$. Let $\mathcal{D}_{\bar{Q},\lambda}^{\text{alg}}$ denote the sheaf on $\bar{Q}$

of (twisted) differential operators of the invertible sheaf $j^*(\mathcal{O}_\lambda^{\mathrm{alg}})$, and let j_+ be the corresponding direct image functor. Thus,

$$j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}})$$

is a K-equivariant sheaf of $\mathcal{D}_\lambda^{\mathrm{alg}}$-modules.

Proposition 4.2. *Suppose $\mathcal{O}_\lambda^{\mathrm{alg}}$ is the sheaf of sections of a K-equivariant algebraic line bundle on X. Let $Q \subseteq X$ be a K-orbit and suppose $\mathcal{I}(Q, \lambda)$ is the corresponding standard Harish-Chandra sheaf. Assume that the associated variety $\bar{Q}$ is smooth and let $j : \bar{Q} \hookrightarrow X$ be the inclusion. Then there exists a natural isomorphism*

$$j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}}) \cong \mathcal{J}(Q, \lambda).$$

Proof. Let

$$l : Q \hookrightarrow \bar{Q}$$

be the inclusion and recall that $i : Q \hookrightarrow X$. Since Q is open in $\bar{Q}$, it is clear that

$$j^*(\mathcal{O}_\lambda^{\mathrm{alg}})|_Q \cong i^*(\mathcal{O}_\lambda^{\mathrm{alg}}),$$

as sheaves of K-equivariant $\mathcal{D}_{\bar{Q},\lambda}^{\mathrm{alg}}$-modules. Furthermore, the direct image l_+ coincides with the direct image $l_\bullet$ in the category of sheaves. By the adjointness property of the direct image

$$\mathrm{Hom}(j^*(\mathcal{O}_\lambda^{\mathrm{alg}}),\, l_* i^*(\mathcal{O}_\lambda^{\mathrm{alg}})) \cong \mathrm{Hom}(j^*(\mathcal{O}_\lambda^{\mathrm{alg}})|_Q,\, i^*(\mathcal{O}_\lambda^{\mathrm{alg}}))$$

so the isomorphism above determines a nonzero morphism

$$j^*(\mathcal{O}_\lambda^{\mathrm{alg}}) \to l_* i^*(\mathcal{O}_\lambda^{\mathrm{alg}})$$

of K-equivariant $\mathcal{D}_{\bar{Q},\lambda}^{\mathrm{alg}}$-modules.

Since $\bar{Q}$ is a closed, smooth subvariety of X, Kashiwara's equivalence of categories says that the direct image j_+ establishes an equivalence between the category of coherent $\mathcal{D}_{\bar{Q},\lambda}^{\mathrm{alg}}$-modules and the category of coherent $\mathcal{D}_\lambda^{\mathrm{alg}}$-modules with support on $\bar{Q}$. Thus, we have a nonzero morphism

$$j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}}) \to j_+ l_* i^*(\mathcal{O}_\lambda^{\mathrm{alg}}) \cong i_+ i^*(\mathcal{O}_\lambda^{\mathrm{alg}}) = \mathcal{I}(Q, \lambda)$$

this last isomorphism since $i = j \circ l$. Now we simply observe that $j^*(\mathcal{O}_\lambda^{\mathrm{alg}})$ is an irreducible $\mathcal{D}_{\bar{Q},\lambda}^{\mathrm{alg}}$-module so that $j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}})$ is also irreducible (once again by Kashiwara's equivalence). Since the morphism we have defined is nonzero and $i_+ i^*(\mathcal{O}_\lambda^{\mathrm{alg}})$ has a unique irreducible coherent subsheaf, the proposition is proven. □

5. Analytic localization and comparison

We are now ready to introduce the analytic localization. The Hecht–Taylor version of the localization functor is built around the topology of the minimal globalization. One the one hand, Hecht and Taylor consider topological U_Θ-modules that have a dual nuclear Fréchet (DNF) topology, where morphisms are continuous morphisms of modules; on the other hand, they define the concept of a DNF sheaf of $\mathcal{D}_\lambda$-modules with an accompanying concept of continuous morphisms of DNF sheaves of modules. For $\lambda \in \Theta$, the topological localization

$$\Delta_\lambda^{\mathrm{an}}(M) = \mathcal{D}_\lambda \hat{\otimes}_{U_\Theta} M$$

does not have very interesting results, but since free resolutions of DNF modules are complexes of DNF modules, using these sorts of resolutions, one can define a derived functor $L\Delta_\lambda^{\mathrm{an}}$. In particular, the analytic localization takes complexes of DNF U_Θ-modules to complexes of DNF sheaves of $\mathcal{D}_\lambda$-modules (by applying $\Delta_\lambda^{\mathrm{an}}$ to the corresponding free resolutions). On the other side of the equation, by using Čech resolutions, Hecht and Taylor show there are enough injectives within the category of DNF sheaves of $\mathcal{D}_\lambda$-modules. Thus, by applying the global sections to injective resolutions, one can define a derived global sections functor on complexes of DNF sheaves of $\mathcal{D}_\lambda$-modules. The result is then a complex of DNF U_Θ-modules. On appropriately defined derived categories, for λ regular, it is not hard to show the derived functors $L\Delta_\lambda^{\mathrm{an}}$ and $R\Gamma$ are mutual inverses.

In general, one does not know about the homology groups of the analytic localization of a complex of DNF U_Θ-modules: these homology groups may very well not be DNF sheaves (although they will be $\mathcal{D}_\lambda$-modules). However, the homology groups of the analytic localization of a minimal globalization M (any DNF U_Θ-module can be thought of as a complex which is zero in all nonzero degrees) turn out to be DNF sheaves of $\mathcal{D}_\lambda$-modules of a very special sort. In order to explain this, we introduce the concept of the *geometric fiber* of a sheaf of $\mathcal{O}_X$-modules. In particular, if $\mathcal{F}$ is a sheaf of $\mathcal{O}_X$-modules and $x \in X$, then we define the *geometric fiber* $T_x(\mathcal{F})$ of $\mathcal{F}$ at x by

$$T_x(\mathcal{F}) = \mathbb{C} \otimes_{\mathcal{O}_{X,x}} \mathcal{F}_x,$$

where $\mathcal{O}_{X,x}$ and $\mathcal{F}_x$ denote the corresponding stalks of these sheaves at x and where $\mathcal{O}_{X,x}$ acts on $\mathbb{C}$ by evaluation at x. Then, letting $G_0[x]$ be the stabilizer of x in G_0, Hecht and Taylor showed (for λ regular) that the geometric fiber

$$T_x(L_p\Delta_\lambda^{\mathrm{an}}(M))$$

of the p-th homology group $L_p\Delta_\lambda^{\mathrm{an}}(M)$ of the analytic localization of the minimal globalization M is a finite-dimensional (continuous) $(\mathfrak{h}_x, G_0[x])$-module, where

$\mathfrak{h}_x$ acts by the evaluation of $\lambda + \rho \in \mathfrak{h}^*$ at x, and that the restriction of $L_p\Delta_\lambda^{\mathrm{an}}(M)$ to the G_0-orbit S of x is the sheaf of (restricted holomorphic) sections of the corresponding homogeneous vector bundle over S.

When Θ is a regular and $\lambda \in \Theta$ is antidominant, the *comparison theorem* [Hecht and Taylor 1993] provides a way to understand the analytic localization $L\Delta_\lambda^{\mathrm{an}}(M)$ of a minimal globalization M, with infinitesimal character Θ, assuming one understands the (derived) geometric fibers of the localization

$$\Delta_\lambda(M_{\mathrm{HC}})$$

of the underlying Harish-Chandra module M_{HC} of M. To explain this result, we introduce the geometric fiber

$$T_x^{\mathrm{alg}}(\mathcal{F}) = \mathbb{C} \otimes_{\mathcal{O}_{X,x}^{\mathrm{alg}}} \mathcal{F}_x$$

of a sheaf of $\mathcal{O}_X^{\mathrm{alg}}$-modules $\mathcal{F}$ at $x \in X$. Note that T_x^{alg} defines a left exact functor (for example) on the category of quasicoherent $\mathcal{D}_\lambda^{\mathrm{alg}}$-modules and that there are corresponding derived functors $L_pT_x^{\mathrm{alg}}$. When M_{HC} is a Harish-Chandra module with infinitesimal character Θ and $\lambda \in \Theta$ is antidominant and regular, Beilinson and Bernstein have shown that the $\mathfrak{h}_x$-modules

$$L_pT_x^{\mathrm{alg}}\Delta_\lambda(M_{\mathrm{HC}})$$

are finite-dimensional (algebraic) ($\mathfrak{h}_x$, $K[x]$)-modules, where $K[x]$ is the stabilizer of x in K and $\mathfrak{h}_x$ acts by the evaluation of $\lambda + \rho \in \mathfrak{h}^*$ at x. The comparison theorem says that when x is a special point, then there exists a natural equivalence between the finite-dimensional ($\mathfrak{h}_x$, $G_0[x]$)-modules and the finite-dimensional ($\mathfrak{h}_x$, $K[x]$)-modules (given by the ($\mathfrak{h}_x$, $K_0[x]$)-structure) and that there is a natural isomorphism

$$L_pT_x^{\mathrm{alg}}\Delta_\lambda(M_{\mathrm{HC}}) \cong T_x(L_p\Delta_\lambda^{\mathrm{an}}(M))$$

of ($\mathfrak{h}_x$, $K_0[x]$)-modules. Note that a more general comparison theorem, described in exactly these terms is proved in [Bratten 1997, Theorem 7.2].

Theorem 5.1. *Suppose $\mathcal{O}_\lambda$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle with regular antidominant parameter $\lambda \in \mathfrak{h}^*$. Let $S \subseteq X$ be a G_0-orbit with vanishing number q and let $U \supseteq S$ be the smallest G_0-invariant open submanifold that contains S. Let Q be the K-orbit that is Matsuki dual to S and suppose the associated variety $\bar{Q}$ is smooth. Then the sheaf cohomology groups*

$$H_{\mathrm{c}}^p(U, \mathcal{O}_\lambda|_U)$$

vanish except in degree q in which case $H_{\mathrm{c}}^q(U, \mathcal{O}_\lambda|_U)$ is the unique irreducible submodule of the standard module $H_{\mathrm{c}}^q(S, \mathcal{O}_\lambda|_S)$.

Proof. Utilizing the notation from the previous section, we know that $H_c^q(S, \mathcal{O}_\lambda|_S)$ is the minimal globalization of the standard Beilinson–Bernstein module $I(Q, \lambda)$. Let $J(Q, \lambda) \subseteq I(Q, \lambda)$ be the corresponding unique irreducible Harish-Chandra submodule and $J(Q, \lambda)_{\min}$ its minimal globalization. Consider the complex

$$R\Gamma((\mathcal{O}_\lambda|_U)^X)$$

of DNF U_Θ-modules and let $J(Q, \lambda)_{\min}[-q]$ denote the complex which has zeros in all gradings except q, where we have the module $J(Q, \lambda)_{\min}$. The point of our proof is to present a nonzero morphism in the derived category

$$J(Q, \lambda)_{\min}[-q] \to R\Gamma((\mathcal{O}_\lambda|_U)^X)$$

such that the induced morphism

$$L\Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\min}[-q]) \to L\Delta_\lambda^{\mathrm{an}}\big(R\Gamma((\mathcal{O}_\lambda|_U)^X)\big) \cong (\mathcal{O}_\lambda|_U)^X[0]$$

is an isomorphism (we include the place holder [0] to emphasize the fact that we think of the sheaf $(\mathcal{O}_\lambda|_U)^X$ as a complex concentrated in degree zero). By the equivalence of derived categories, we will thus obtain an isomorphism

$$R\Gamma((\mathcal{O}_\lambda|_U)^X) \cong J(Q, \lambda)_{\min}[-q],$$

which is the desired result.

To present a morphism in the derived category, recall, by Proposition 4.1, that we have a natural inclusion

$$J(Q, \lambda)_{\min} \to H_c^q(U, \mathcal{O}_\lambda|_U).$$

Since the sheaf cohomology groups of $(\mathcal{O}_\lambda|_U)^X$ vanish in degrees smaller that q, a standard truncation argument provides a nonzero morphism in the derived category

$$H_c^q(U, \mathcal{O}_\lambda|_U)[-q] \to R\Gamma((\mathcal{O}_\lambda|_U)^X).$$

Composing this morphism with the inclusion gives the desired result.

We now want to show that the cohomology groups of the complex

$$L\Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\min}[-q])$$

vanish except in degree zero. That is, we want to calculate the homology groups

$$L_p\Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\min})$$

and see that they vanish except in degree q. To do this, we use the comparison theorem. So we need to calculate the derived geometric fibers of the sheaf

$$\Delta_\lambda(J(Q, \lambda)) = \mathcal{J}(Q, \lambda).$$

Let $j : \bar{Q} \hookrightarrow X$ denote the inclusion. Since $\bar{Q}$ is smooth, we have

$$\mathcal{J}(Q, \lambda) \cong j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}}).$$

Thus, calculating the geometric fibers

$$L_p T_x^{\mathrm{alg}} \Delta_\lambda(J(Q, \lambda)) \cong L_p T_x^{\mathrm{alg}}(j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}}))$$

is a straightforward application of the base change formula for the direct image in the category of TDOs. In particular, since the codimension of $\bar{Q}$ in X is q, it follows that, for each $x \in X$,

$$L_p T_x^{\mathrm{alg}} \Delta_\lambda(J(Q, \lambda)) = 0 \quad \text{if } p \neq q.$$

When $x \in \bar{Q}$,

$$L_q T_x^{\mathrm{alg}} \Delta_\lambda(J(Q, \lambda)) \cong L_q T_x^{\mathrm{alg}}(j_+ j^*(\mathcal{O}_\lambda^{\mathrm{alg}})) \cong T_x^{\mathrm{alg}}(\mathcal{O}_\lambda^{\mathrm{alg}}),$$

and when $x \notin \bar{Q}$,

$$L_q T_x^{\mathrm{alg}} \Delta_\lambda(J(Q, \lambda)) = 0.$$

In particular, if we let

$$V = \Gamma(X, \mathcal{O}_\lambda^{\mathrm{alg}})$$

be the corresponding irreducible finite-dimensional $(\mathfrak{g}, K)$-module, then for each special point $x \in \bar{Q}$,

$$L_q T_x^{\mathrm{alg}} \Delta_\lambda(J(Q, \lambda)) \cong V/\mathfrak{n}_x V$$

as $(\mathfrak{h}_x, K[x])$-modules.

By the comparison theorem, it follows that the homology groups

$$L_p \Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\mathrm{min}})$$

vanish except when $p = q$, and this, in turn, implies that the complex

$$L\Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\mathrm{min}}[-q])$$

is quasi-isomorphic to the complex

$$L_q \Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\mathrm{min}})[0]$$

which is nonzero only in degree 0. Hence, the nonzero morphism

$$L\Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\mathrm{min}}[-q]) \to (\mathcal{O}_\lambda|_U)^X[0]$$

in the derived category reduces to a nonzero morphism

$$L_q \Delta_\lambda^{\mathrm{an}}(J(Q, \lambda)_{\mathrm{min}}) \to (\mathcal{O}_\lambda|_U)^X$$

of G_0-equivariant DNF sheaves of $\mathcal{D}_\lambda$-modules. We will prove that this morphism is in fact an isomorphism. In particular, if O is a G_0-orbit in U, then since both $L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})|_O$ and $\mathcal{O}_\lambda|_O$ are induced equivariant sheaves, it follows that the restricted morphism

$$L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})|_O \to \mathcal{O}_\lambda|_O$$

is either an isomorphism or zero.

Notice that these limited possibilities for the restricted morphism can also be deduced from the fact that we have a morphism of $(\mathcal{D}_\lambda|_O)$-modules and both objects are locally free rank one sheaves of $(\mathcal{O}_X|_O)$-modules. Indeed, if we knew a priori that $L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})$ was a locally free sheaf of $\mathcal{O}_U$-modules, then it would follow immediately from standard $\mathcal{D}$-module theory [Hotta et al. 2008, Theorem 1.4.10] that a nonzero morphism of $\mathcal{D}_\lambda$-modules would be an isomorphism.

Define W to be the set of $x \in U$ such that the induced morphism

$$L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})_x \to (\mathcal{O}_\lambda)_x$$

is nonzero. We will show that W is an open set that contains S. Since G_0 acts on W and since U is the smallest G_0-invariant open set that contains S it will follow from our previous remarks that the morphism in question is an isomorphism.

Consider the composition

$$L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min}) \to (\mathcal{O}_\lambda|_U)^X \to (\mathcal{O}_\lambda|_S)^X,$$

where the second morphism is the canonical one. Since these morphisms induce the nonzero composition

$$J(Q,\lambda)_{\min} \to H_c^q(U,\mathcal{O}_\lambda|_U) \to H_c^q(S,\mathcal{O}_\lambda|_S),$$

it follows that the restricted morphism

$$L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})|_S \to \mathcal{O}_\lambda|_S$$

is an isomorphism and $S \subseteq W$. To show W is open, suppose $x \in W$. Since $\mathcal{O}_\lambda$ is a locally free rank one sheaf of $\mathcal{O}_X$-modules there is a local section σ of $\mathcal{O}_\lambda$, defined on a neighborhood of x such that every local section has the form $f\sigma$ where f is a holomorphic function. Since the induced morphism on the geometric fiber

$$T_x(L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})) \to T_x(\mathcal{O}_\lambda)$$

is nonzero, it follows that for some open set W_1 that contains x, there is a holomorphic function f defined on W_1 such that $f(x) \neq 0$, and there is a local section in $\Gamma\big(W_1, L_q\Delta_\lambda^{\mathrm{an}}(J(Q,\lambda)_{\min})\big)$ that maps onto $f\sigma$. Thus,

$$W_2 = \{z \in W_1 : f(z) \neq 0\}$$

is an open set such that $x \in W_2 \subseteq W$, and we have finished the proof. ☐

6. Some additional considerations

6.1. ***A tensoring argument.*** We maintain the notation from the previous section. In particular, S is a G_0-orbit in X, Q is the K-orbit that is Matsuki dual to S, $\mathcal{O}_\lambda$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle on X, and so on. When the parameter $\lambda \in \mathfrak{h}^*$ is antidominant then it may be the case that the Harish-Chandra module

$$J(Q,\lambda) = \Gamma(X, \mathcal{J}(Q,\lambda))$$

is zero. However, when $J(Q,\lambda) \neq 0$, it is the unique irreducible submodule of the standard Beilinson–Bernstein module $I(Q,\lambda)$. When λ is antidominant and $J(Q,\lambda) \neq 0$, we will refer to $I(Q,\lambda)$ (as well as its minimal globalization $H_c^q(S, \mathcal{O}_\lambda|_S)$) as a *classifying module*. Let U be the smallest G_0-invariant open submanifold that contains S. Under the assumption that the associated variety $\bar{Q}$ is smooth and λ is antidominant, we can give the following tensoring argument that shows that $H_c^q(U, \mathcal{O}_\lambda|_U)$ is the minimal globalization of $J(Q,\lambda)$. Hence, when $H_c^q(S, \mathcal{O}_\lambda|_S)$ is a classifying module, it follows that $H_c^q(U, \mathcal{O}_\lambda|_U)$ is the unique irreducible submodule.

Lemma 6.1. *Assume that $\lambda \in \mathfrak{h}^*$ is antidominant, and suppose that the associated variety $\bar{Q}$ is smooth. Then, the sheaf cohomology groups*

$$H_c^p(U, \mathcal{O}_\lambda|_U)$$

vanish except in degree q, in which case $H_c^q(U, \mathcal{O}_\lambda|_U)$ is the minimal globalization of $J(Q,\lambda)$.

Proof. The proof is basically the same as (but simpler than) the proof in [Bratten 1997, Theorem 9.4] with the slight difference that we need to use the description of the irreducible Harish-Chandra sheaf $\mathcal{J}(Q,\lambda)$ from Proposition 4.2 instead of the description for $\mathcal{I}(Q,\lambda)$. We sketch some details to help the reader adapt the notation here to the notation in Section 9 of that reference. From the theory of highest weight modules, one knows there is an irreducible finite-dimensional G_0-module F^μ which is irreducible as a $\mathfrak{g}$-module and has a highest weight $\mu \in \mathfrak{h}^*$ sufficiently dominant that $\lambda - \mu$ is antidominant and regular. Observe that $\mathcal{O}_{\lambda-\mu}$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle. Let Θ be the infinitesimal character

$$\Theta = W \cdot \lambda.$$

If M is a $\mathfrak{g}$-module (or if $\mathcal{M}$ is a sheaf of $\mathfrak{g}$-modules) we let M_Θ (respectively $\mathcal{M}_\Theta$) denote the corresponding $Z(\mathfrak{g})$-eigenspace. Then, as in the proof of [loc. cit.], we have the natural isomorphisms:

(i) $\big(\mathcal{O}_{\lambda-\mu}|_U \otimes F^\mu\big)_\Theta \cong \mathcal{O}_\lambda|_U$, and

(ii) $\big(\mathcal{J}(Q,\lambda-\mu) \otimes F^\mu\big)_\Theta \cong \mathcal{J}(Q,\lambda)$.

Taking sheaf cohomology, in the first case we obtain

$$H^p_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U) \cong \big(H^p_{\mathrm{c}}(U,\mathcal{O}_{\lambda-\mu}|_U) \otimes F^\mu\big)_\Theta,$$

which implies that the compactly supported sheaf cohomology groups $H^p_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U)$ vanish except when $p=q$, in which case $H^q_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U)$ is the minimal globalization of a Harish-Chandra module. To see which Harish-Chandra module, we begin with the natural isomorphism from the previous section

$$(J_{\lambda-\mu})_{\mathrm{min}} \cong H^q_{\mathrm{c}}(U,\mathcal{O}_{\lambda-\mu}|_U).$$

Therefore, we obtain the isomorphism

$$[(J_{\lambda-\mu} \otimes F^\mu)_\Theta]_{\mathrm{min}} \cong (H^q_{\mathrm{c}}(U,\mathcal{O}_{\lambda-\mu}|_U) \otimes F^\mu)_\Theta \cong H^q_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U).$$

Finally, taking global sections for the isomorphism in (ii), we obtain

$$(J_{\lambda-\mu} \otimes F^\mu)_\Theta \cong J_\lambda,$$

which completes the proof of the lemma. □

Corollary 6.2. *Let $S \subseteq X$ be a G_0-orbit with vanishing number q and let $U \supseteq S$ be the smallest G_0-invariant open submanifold that contains S. Let Q be the K-orbit that is Matsuki dual to S and suppose the associated variety $\bar{Q}$ is smooth. Let $\mathcal{O}_\lambda$ be the sheaf of holomorphic sections of a G_0-equivariant line bundle and suppose*

$$H^q_{\mathrm{c}}(S,\mathcal{O}_\lambda|_S)$$

is a classifying module. Then, the sheaf cohomology groups

$$H^p_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U)$$

vanish except in degree q in which case $H^q_{\mathrm{c}}(U,\mathcal{O}_\lambda|_U)$ is the unique irreducible submodule of the standard module $H^q_{\mathrm{c}}(S,\mathcal{O}_\lambda|_S)$.

6.2. ***Maximal parabolic subgroups of complex reductive groups.*** Suppose G_0 is a connected, complex reductive group. It turns out that the representation we are studying has a close relationship to the classical parabolic induction when the parabolic subgroup under consideration is maximal. This allowed us to consider some examples (with the help of D. Vogan and A. Paul) to see how the representation works when the associated algebraic variety is singular. In the examples we considered, the representation is irreducible only when the associated algebraic variety is nonsingular.

In particular, let $H_0 \subseteq G_0$ be a θ-stable Cartan subgroup (θ is the complex conjugation of G_0 with respect to a compact real form) and let $B_0 \supseteq H_0$ be a Borel subgroup. We consider a maximal, proper parabolic subgroup P_0 of G_0 that contains B_0. These are determined in the following way. Let $W(G_0)$ be the Weyl group of H_0 (we can think of $W(G_0)$ as the quotient of the normalizer of H_0 in G_0 over H_0). Then, $W(G_0)$ acts naturally on the set of Borel subgroups that contain H_0. As in the introduction, we let X_0 be the complex flag manifold of Borel subgroups of G_0 and let X_0^{c} be the conjugate complex manifold. Then, the flag manifold X of Borel subalgebras of the complexified Lie algebra $\mathfrak{g}$ of $\mathfrak{g}_0$ can be identified with the direct product

$$X = X_0 \times X_0^{\mathrm{c}}.$$

We have the two actions of G_0 on X: the diagonal action

$$g \cdot (x, y) = (gx, gy),$$

corresponding to the fact that G_0 is a real group with real Lie algebra $\mathfrak{g}_0$, and the action

$$g \cdot (x, y) = (gx, \theta(g)y),$$

corresponding to the action of $G_0 = K$ as the complexification of K_0. As before let B_0^{op} be the Borel subgroup opposite to B_0 (this subgroup corresponds to the longest element in $W(G_0)$). Then, each G_0-orbit and each K-orbit on X contains exactly one special point of the form

$$(w \cdot B_0, B_0^{\mathrm{op}}) \in X_0 \times X_0^{\mathrm{c}},$$

so we can identify G_0-orbits and K-orbits with elements of $W(G_0)$. One knows that $Q_{w_1} \subseteq \overline{Q_{w_2}}$ if and only $w_1 \preceq w_2$ in the Bruhat order $\preceq$ on $W(G_0)$. In particular, the Bruhat interval

$$[1, w] = \{u \in W(G_0) : u \preceq w\}$$

characterizes the K-orbits Q_u contained in $\overline{Q_w}$, as well as the G_0-orbits S_u contained in the smallest G_0-invariant open submanifold U_w that contains S_w. Let n be the number of simple reflections in $W(G_0)$. Observe that the number of G_0-orbits with vanishing number 1 is exactly n. (The closed G_0-orbit is the unique orbit with vanishing number 0.)

Let Y_0 be the generalized complex flag space of G_0-conjugates to P_0 and let Y_0^{c} be the complex manifold conjugate to Y_0. Consider the generalized flag space

$$Y = Y_0 \times Y_0^{\mathrm{c}},$$

and let C be the G_0-orbit of $y = (P_0, P_0) \in Y$. Then, C is closed in Y, and the G_0-orbit of y is a real form in Y. Let

$$\pi : X \to Y$$

denote the equivariant projection. Then, $\pi^{-1}(C)$ is a closed G_0-invariant submanifold, and, since P_0 is a maximal, it contains exactly $n-1$ orbits with vanishing number 1. In particular, if $L_0 \subseteq P_0$ is the Levi factor of P_0 that contains H_0 then these $n-1$ orbits correspond to the simple reflections of the Weyl group $W(L_0)$ of H_0 in L_0, and the orbits in $\pi^{-1}(C)$ correspond to the elements in $W(L_0)$. Indeed, by intersection with the fiber, these G_0-orbits give the L_0-orbits in the complex flag manifold

$$X_y = \pi^{-1}(\{y\})$$

for L_0. Let S be the remaining G_0-orbit with vanishing number 1 and let

$$U = X - \pi^{-1}(C).$$

Then, U is the smallest G_0-invariant open set that contains S. (To see this fact, since U is open and contains S, is sufficient to check that S is the unique G_0-orbit that is closed in U.)

Let $\mathcal{O}_\lambda$ be the sheaf of holomorphic sections of a G_0-equivariant line bundle on X and assume λ is antidominant and regular. In a natural way, the sheaf $\mathcal{O}_\lambda$ determines a corresponding sheaf of holomorphic sections $\mathcal{O}_{X_y,\lambda}$ for an L_0-equivariant line bundle defined on X_y. Let

$$F = \Gamma(X_y, \mathcal{O}_{X_y,\lambda})$$

be the corresponding irreducible finite-dimensional representation for L_0 with lowest weight $\lambda + \rho$. In a unique way, this representation extends to an irreducible representation

$$\omega : P_0 \to \mathrm{GL}(F).$$

Consider the corresponding classical (unnormalized) parabolic induction $I_{P_0}^{G_0}(F)$, given by

$$I_{P_0}^{G_0}(F) = \{\text{real analytic functions } \varphi : G_0 \to F : \varphi(gp) = \omega(p^{-1})\varphi(g)\}.$$

Then, there is a natural isomorphism of G_0-modules (see, e.g., [Bratten 2008]),

$$\Gamma(\pi^{-1}(C), \mathcal{O}_\lambda) \cong I_{P_0}^{G_0}(F) \quad \text{and} \quad H^p(\pi^{-1}(C), \mathcal{O}_\lambda) = 0 \text{ for } p > 0,$$

where we obtain the vanishing by the Leray spectral sequence and the fact that the sheaf cohomology groups of a real analytic vector bundle over a real analytic manifold vanish in positive degree.

Now, consider the short exact sequence of sheaves

$$0 \to (\mathcal{O}_\lambda|_U)^X \to \mathcal{O}_\lambda \to (\mathcal{O}_\lambda|_{\pi^{-1}(C)})^X \to 0.$$

Thus, we have the short exact sequence of representations

$$0 \to V \to I_{P_0}^{G_0}(F) \to H_c^1(U, \mathcal{O}_\lambda|_U) \to 0,$$

where $V = \Gamma(X, \mathcal{O}_\lambda)$ is the corresponding irreducible finite-dimensional G_0-module. Therefore, the minimal globalization $H_c^1(U, \mathcal{O}_\lambda|_U)$ is irreducible if and only if the quotient

$$I_{P_0}^{G_0}(F)/V$$

is irreducible.

Let Q be the K-orbit Matsuki dual to S. Then one would like to know when $\bar{Q}$ is smooth. The calculation for $\mathrm{GL}(n+1, \mathbb{C})$ (which is not difficult) works like this. The Levi factor of P_0 is characterized by a partition

$$n_1 + n_2 = n + 1,$$

where

$$L_0 = \mathrm{GL}(n_1, \mathbb{C}) \times \mathrm{GL}(n_2, \mathbb{C}) \subseteq \mathrm{GL}(n+1, \mathbb{C}).$$

It turns out that $\bar{Q}$ is smooth if and only n_1 and n_2 belong to $\{n, 1\}$. Therefore, $I_{P_0}^{G_0}(F)/V$ is irreducible in this case. At this point, we contacted D. Vogan to see what was known about the composition factors of these principal series (we asked about the case when $V = \mathbb{C}$ is the trivial G_0-module). After doing a calculation, he guessed that there are $\min\{n_1, n_2\}$ composition factors occurring in the representation $I_{P_0}^{G_0}(\mathbb{C})/\mathbb{C}$. Vogan passed this on to Annegret Paul, who confirmed the guess for some low-dimensional examples by using a computer program (apparently the group $\mathrm{GL}(6, \mathbb{C})$ is already a difficult calculation for the algorithms that were used).

Hence, for these examples, the representation $H_c^1(U, \mathcal{O}_\lambda|_U)$ is irreducible if and only if the associated algebraic variety is smooth.

7. Serre duality

Since the resolutions used in the Hecht–Taylor construction of the derived category of DNF sheaves of $\mathcal{D}_\lambda$-modules are Čech resolutions, perhaps it is worth mentioning that it is not difficult to establish the validity of Serre duality using these sorts of resolutions [Bratten 1997, Section 10]. In particular, let n be the complex dimension of X and let Ω^n be the canonical bundle on X. Thus, for $x \in X$, the geometric fiber $T_x(\Omega^n)$ of Ω^n at x is given by

$$T_x(\Omega^n) = \bigwedge^n \mathfrak{n}_x$$

as a $G_0[x]$-module (recall that $G_0[x]$ denotes the stabilizer of x in G_0). In particular, Ω^n is a G_0-equivariant holomorphic line bundle on X. Using the unshifted notation from Section 3 of this paper, suppose $\mathcal{O}(\mu)$ is the sheaf of holomorphic sections of a G_0-equivariant line bundle on X. Then, the sheaf of holomorphic sections of the dual bundle is given by $\mathcal{O}(-\mu)$ (i.e., the sheaf of sections of the line bundle associated to the dual geometric fiber). If $U \subseteq X$ is any G_0-invariant open submanifold of X, then Serre duality gives a natural isomorphism of topological G_0-modules

$$H_{\mathrm{c}}^p(U, \mathcal{O}(\mu)|_U)' \cong H^{n-p}(U, \mathcal{O}(-\mu) \otimes \mathcal{O}(\Omega^n)|_U)$$

where $H_c^p(U, \mathcal{O}(\mu)|_U)'$ denotes the continuous dual of the topological G_0-module $H_c^p(U, \mathcal{O}(\mu)|_U)$ and $\mathcal{O}(\Omega^n)$ is the sheaf of holomorphic sections of the canonical bundle. In terms of the shifted $\mathcal{D}$-module parameters $\lambda \in \mathfrak{h}^*$, we obtain

$$H_{\mathrm{c}}^p(U, \mathcal{O}_\lambda|_U)' \cong H^{n-p}(U, \mathcal{O}_{-\lambda}|_U)$$

for each p. In particular, the sheaf cohomology groups of a G_0-equivariant holomorphic line bundle on a G_0-invariant open submanifold are maximal globalizations of Harish-Chandra modules.

Acknowledgments

We would like to take this moment to acknowledge the importance of the work by Hecht and Taylor for these results. We would also like to thank David Vogan and Annegret Paul for being kind enough to help us see how the representation we are studying fails to be irreducible when the associated variety is singular.

References

[Beilinson and Bernstein 1981] A. Beilinson and J. Bernstein, "Localisation de g-modules", *C. R. Acad. Sci. Paris Sér. I Math.* **292**:1 (1981), 15–18. MR 82k:14015 Zbl 0476.14019

[Bott 1957] R. Bott, "Homogeneous vector bundles", *Ann. of Math.* (2) **66** (1957), 203–248. MR 19,681d Zbl 0094.35701

[Bratten 1997] T. Bratten, "Realizing representations on generalized flag manifolds", *Compositio Math.* **106**:3 (1997), 283–319. MR 98j:22018 Zbl 0928.22014

[Bratten 1998] T. Bratten, "Finite rank homogeneous holomorphic bundles in flag spaces", pp. 21–34 in *Geometry and representation theory of real and p-adic groups* (Córdoba, 1995), edited by J. Tirao et al., Progr. Math. **158**, Birkhäuser, Boston, 1998. MR 98j:22019 Zbl 0891.22011

[Bratten 2008] T. Bratten, "A geometric embedding for standard analytic modules", *Beiträge Algebra Geom.* **49**:1 (2008), 33–57. MR 2009b:22013 Zbl 1152.22013

[Hecht and Taylor 1990] H. Hecht and J. L. Taylor, "Analytic localization of group representations", *Adv. Math.* **79**:2 (1990), 139–212. MR 91c:22027 Zbl 0701.22005

[Hecht and Taylor 1993] H. Hecht and J. L. Taylor, "A comparison theorem for $\mathfrak{n}$-homology", *Compositio Math.* **86**:2 (1993), 189–207. MR 94c:22015 Zbl 0784.22006

[Hecht et al. 1987] H. Hecht, D. Miličić, W. Schmid, and J. A. Wolf, "Localization and standard modules for real semisimple Lie groups, I: The duality theorem", *Invent. Math.* **90**:2 (1987), 297–332. MR 89e:22025 Zbl 0699.22022

[Hotta et al. 2008] R. Hotta, K. Takeuchi, and T. Tanisaki, *D-modules, perverse sheaves, and representation theory*, Progress in Mathematics **236**, Birkhäuser, Boston, 2008. MR 2008k:32022 Zbl 1136.14009

[Matsuki 1979] T. Matsuki, "The orbits of affine symmetric spaces under the action of minimal parabolic subgroups", *J. Math. Soc. Japan* **31**:2 (1979), 331–357. MR 81a:53049 Zbl 0396.53025

[Matsuki 1982] T. Matsuki, "Orbits on affine symmetric spaces under the action of parabolic subgroups", *Hiroshima Math. J.* **12**:2 (1982), 307–320. MR 83k:53072 Zbl 0495.53049

[Matsuki 1988] T. Matsuki, "Closure relations for orbits on affine symmetric spaces under the action of minimal parabolic subgroups", pp. 541–559 in *Representations of Lie groups* (Kyoto, Hiroshima, 1986), edited by K. Okamoto and T. Ōshima, Adv. Stud. Pure Math. **14**, Academic Press, Boston, 1988. MR 91c:22014 Zbl 0723.22020

[Miličić 1993] D. Miličić, "Algebraic $\mathcal{D}$-modules and representation theory of semisimple Lie groups", pp. 133–168 in *The Penrose transform and analytic cohomology in representation theory* (South Hadley, MA, 1992), edited by M. Eastwood et al., Contemp. Math. **154**, Amer. Math. Soc., Providence, RI, 1993. MR 94i:22035 Zbl 0821.22005

[Schmid 1985] W. Schmid, "Boundary value problems for group invariant differential equations", pp. 311–321 in *Élie Cartan et les mathématiques d'aujourd'hui* (Lyon, 1984), Astérisque, Société Mathématique de France, Paris, 1985. Numéro Hors Série. MR 87h:22018 Zbl 0621.22014

[Schmid 1989] W. Schmid, "Homogeneous complex manifolds and representations of semisimple Lie groups", pp. 223–286 in *Representation theory and harmonic analysis on semisimple Lie groups*, edited by P. J. Sally, Jr., Math. Surveys Monogr. **31**, Amer. Math. Soc., Providence, RI, 1989. MR 90i:22025 Zbl 0744.22016

[Wolf 1969] J. A. Wolf, "The action of a real semisimple group on a complex flag manifold, I: Orbit structure and holomorphic arc components", *Bull. Amer. Math. Soc.* **75** (1969), 1121–1237. MR 40 #4477 Zbl 0183.50901

[Wong 1995] H.-W. Wong, "Dolbeault cohomological realization of Zuckerman modules associated with finite rank representations", *J. Funct. Anal.* **129**:2 (1995), 428–454. MR 96c:22024 Zbl 0855.22014

Received June 27, 2014. Revised March 14, 2015.

JOSÉ ARAUJO
FACULTAD DE CIENCIAS EXACTAS
UNICEN
7000 TANDIL
ARGENTINA
araujo@exa.unicen.edu.ar

TIM BRATTEN
FACULTAD DE CIENCIAS EXACTAS
UNICEN
7000 TANDIL
ARGENTINA
clarkbratten@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.287

A CURVATURE FLOW UNIFYING SYMPLECTIC CURVATURE FLOW AND PLURICLOSED FLOW

SONG DAI

Streets and Tian (2010, 2014) introduced pluriclosed flow and symplectic curvature flow. Here we construct a curvature flow to unify these two flows. We show the short-time existence of our flow and exhibit an obstruction to long-time existence.

1. Introduction

In recent years, Streets and Tian initialized the study of special geometric structures, such as generalized Kähler and symplectic structures, by using curvature flows they introduced. They include Hermitian curvature flow, pluriclosed flow, almost Hermitian curvature flow and symplectic curvature flow [Streets and Tian 2010; 2011; 2014]. Subsequently, there are several further works along this direction; see [Boling 2014; Enrietti et al. 2015; Enrietti 2013; Fernández-Culma 2013; Pook 2012; Smith 2013; Streets and Tian 2013; 2012; Vezzoni 2011]. In this paper, we introduce a curvature flow which unifies symplectic curvature flow and pluriclosed flow.

Streets and Tian [2014] introduced symplectic curvature flow, which preserves almost Kähler structure, as follows:

$$\begin{aligned}
&\frac{\partial}{\partial t} g = -2\,\mathrm{Ric} + \tfrac{1}{2} B^1 - B^2,\\
&\frac{\partial}{\partial t} J = \triangle J + \mathcal{N} + \mathcal{R},\\
&g(0) = g_0,\\
&J(0) = J_0,
\end{aligned} \tag{1}$$

where $\mathcal{R}$ is a curvature term and B^1, B^2, $\mathcal{N}$ are all quadratic terms of DJ. We will give the precise definitions of these tensors in Section 3.

Streets and Tian [2010] introduced pluriclosed flow, which preserves pluriclosed

---

The author is partially supported by China Scholarship Council.
MSC2010: 53C15, 53C44, 53C56, 53D05, 53D15.
Keywords: symplectic curvature flow, pluriclosed flow.

structure, as follows:

$$\begin{aligned}&\frac{\partial}{\partial t}\omega = \partial\partial^*\omega + \bar\partial\bar\partial^*\omega + \tfrac{1}{2}\sqrt{-1}\,\partial\bar\partial\,\operatorname{logdet} g,\\ &\omega(0) = \omega_0.\end{aligned}$$

Then, in [Streets and Tian 2013; 2012] they observed that, after a gauge transformation induced by the Lee form $\theta = -Jd^*\omega$, pluriclosed flow is equivalent to the following flow:

$$\begin{aligned}&\frac{\partial}{\partial t}g = -2\operatorname{Ric} + \tfrac{1}{2}\mathcal{B},\\ &\frac{\partial}{\partial t}J = \triangle J + \mathcal{R} + \mathcal{Q},\\ &g(0) = g_0,\\ &J(0) = J_0,\end{aligned} \tag{2}$$

where $\mathcal{B}$ and $\mathcal{Q}$ are quadratic terms of DJ. We will give the precise definitions of these tensors in Section 3. In this setting, they showed that twisted generalized Kähler manifolds are a natural background in which to run pluriclosed flow [Streets and Tian 2012].

Hitchin [2003] first introduced the notion of generalized complex structure, which unifies symplectic structure and complex structure. After that, Gualtieri discussed generalized complex structure in detail in his thesis [Gualtieri 2011]. In that work, Gualtieri discovered that a pair of compatible almost generalized complex structures $(\mathcal{J}_1, \mathcal{J}_2)$ is equivalent to almost bi-Hermitian data (g, J_+, J_-, b), where $J_\pm$ are almost complex structures, compatible with g, and b is a 2-form. If $\mathcal{J}_1, \mathcal{J}_2$ are both integrable, i.e., generalized Kähler, the integrability condition is equivalent to

$$\begin{gathered}N_{J_+} = N_{J_-} = 0,\\ -d^c_+\omega_+ = d^c_-\omega_- = db.\end{gathered}$$

If we only require db to be a closed 3-form H (which is the twisted case) Streets and Tian [2012] showed that the equivalent pluriclosed flow (2) of (g, J_+) and (g, J_-) preserves generalized Kähler structure.

A symplectic structure ω gives a generalized complex structure $\mathcal{J}_\omega$, and an almost Kähler structure (ω, J) gives a compatible pair of almost generalized complex structures $(\mathcal{J}_\omega, \mathcal{J}_J)$, where $\mathcal{J}_\omega$ is integrable while $\mathcal{J}_J$ is not necessarily. So one may also regard symplectic curvature flow as a curvature flow to deform a compatible pair of almost generalized complex structures $(\mathcal{J}_1, \mathcal{J}_2)$, where $\mathcal{J}_1$ is integrable. This leads to the question of whether or not there is a curvature flow that unifies the flows in (1) and (2). The following theorem gives a solution to this problem.

Theorem 1.1. *Let* (M, g_0, J_0) *be an almost Hermitian manifold. Suppose* M *is compact. Then there exists a unique family of almost Hermitian structures* $(g(t), J(t))$, $t \in [0, \epsilon)$ *on* M *satisfying the equations*

$$\begin{aligned} \frac{\partial}{\partial t} g &= -2\,\mathrm{Ric} + Q_1, \\ \frac{\partial}{\partial t} J &= \triangle J + \mathcal{N} + \mathcal{R} + Q_2, \\ g(0) &= g_0, \\ J(0) &= J_0. \end{aligned} \tag{3}$$

Here $\mathcal{R}$ *and* $\mathcal{N}$ *are the same as in* (1), *and* Q_1 *and* Q_2 *are quadratic terms of* DJ (*see Section 3 for their precise definitions*). *This flow preserves the integrability of* J. *Furthermore*, *if the initial data is almost Kähler*, *this flow coincides with symplectic curvature flow*, *and if the initial data is pluriclosed*, *this flow is equivalent to pluriclosed flow. In particular*, *if the initial data is Kähler*, *this flow is Kähler–Ricci flow.*

Another motivation to unify (1) and (2) is to try to understand symplectic curvature flow better. The tremendous success of [Perelman 2002] motivates us to find similar tools in symplectic curvature flow as exist in Ricci flow. To begin with, we consider whether symplectic curvature flow is a gradient flow, as is Ricci flow. It seems difficult to construct such a functional directly. But as shown in [Streets and Tian 2013], pluriclosed flow is a gradient flow, and the functional is similar to the case of Ricci flow. So maybe our flow could give some hints to discover the desired functional in symplectic curvature flow.

Turning to regularity, we derive the evolution equations, and then obtain the derivative estimates, as follows.

Theorem 1.2. *Let* $(M, g(t), J(t))$ *be a solution of* (3) *for* $t \in [0, T)$. *Suppose* M *is compact. If there exists a constant* K *such that*

$$\sup_{[0,T)\times M} \{t|\mathrm{Rm}|, t^{1/2}|DJ|\} \leq K,$$

then for $k \geq 0$ *there exists a constant* $C = C(k, n, K)$ *such that*

$$\sup_{[0,T)\times M} \{t^{(k+2)/2}|D^k\,\mathrm{Rm}|, t^{k/2}|D^k J|\} \leq C.$$

Finally, we obtain an obstruction to long-time existence.

Theorem 1.3. *Let* $(M, g(t), J(t))$ *be a solution of* (3) *for* $t \in [0, T)$, *and let* $T < +\infty$ *be the maximal existence time. Suppose* M *is compact. Then*

$$\sup_{[0,T)\times M} \{|\mathrm{Rm}|, |DJ|\} = +\infty.$$

We outline the proof now. Some results in this paper can be implied directly from the results in [Streets and Tian 2014]. For the convenience of readers, we give the complete proof here.

To prove Theorem 1.1, we use the DeTurck trick. But we notice that the almost complex structure J does not live in a vector space. So we transform the equation on the space of almost complex structures to its tangent space at J_0. We don't assume (g, J) is compatible at first, so we do some modifications to ensure the compatibility, which gives the nondegenerate symbol. Thus we obtain the short-time existence of the modified flow. Then we do some estimates to show that the modified flow gives a compatible pair (g, J) and that it coincides with the initial flow. For uniqueness, it is the same as in Ricci flow. In the symplectic and pluriclosed settings, by direct calculation in Section 3 we see that this flow can be reduced to symplectic curvature flow and pluriclosed flow, respectively. So, by uniqueness, they coincide with our flow. And a similar argument also applies to the integrability of J.

To prove Theorem 1.2, the argument is standard. We derive the evolution equations of $D^k J$ and $D^k \operatorname{Rm}$, then we construct a function involving the terms we want to estimate. Calculating the evolution equation of this function, and then using the maximum principle, we obtain the desired result. To prove Theorem 1.3, the argument is also standard and the same as in Ricci flow.

We organize the paper as follows. In Section 2, we recall some preliminaries in almost Hermitian geometry and derive the necessary condition of a variation of almost Hermitian pairs. In Section 3 we define the tensors we will use in this paper. Then we do some calculations to show that our flow satisfies the necessary condition. And, also by calculation, we show that the additional tensors will vanish in special cases. In Section 4, we prove Theorem 1.1. In Section 5, we prove Theorem 1.2 and Theorem 1.3.

2. Preliminaries

We fix some conventions first.

Convention. (i) Let g be a Riemannian structure. We identify elements $T \in \Gamma(\operatorname{End}(TM))$ and $T \in \Gamma(T^*M \otimes T^*M)$ by

$$g(T(X), Y) = T(X, Y).$$

We implicitly use this identification throughout this paper.

(ii) When we write repeated indices, we always mean to take the trace with respect to these two positions, i.e., to choose an orthonormal basis and take the sum.

(iii) We write DJ^{*3} for $DJ * DJ * DJ$, etc.

(iv) Sometimes we write i instead of e_i for short.

(v) Sometimes we omit the time parameter t if there is no ambiguity.

(vi) D denotes the Levi-Civita connection, which we always use throughout the paper.

We come back to the preliminaries.

Let M be a manifold, J be a section of $\mathrm{End}(TM)$. We call J an almost complex structure if $J^2 = -1$. An almost complex structure J is called integrable if J is induced by holomorphic coordinates. By the theorem of Newlander and Nirenberg [1957], J is integrable if and only if $N = 0$, where

$$N(X, Y) = [JX, JY] - [X, Y] - J[JX, Y] - J[X, JY]$$

is called the Nijenhuis tensor.

We call (g, J) an almost Hermitian structure if g is a Riemannian metric, J is an almost complex structure and (g, J) is compatible, meaning that

$$g(JX, JY) = g(X, Y).$$

For almost Hermitian structure (g, J), we define

$$\omega(X, Y) = g(JX, Y).$$

Moreover, if J is integrable, (g, J) is called a Hermitian structure. If $d\omega = 0$, then (g, J) is called an almost Kähler structure. If J is integrable and $d\omega = 0$, then (g, J) is called a Kähler structure. If J is integrable and $dd^c\omega = 0$, where

$$d^c\omega(X, Y, Z) := -d\omega(JX, JY, JZ),$$

then (g, J) is called a pluriclosed or SKT structure (strong Kähler with torsion).

Definition 2.1. Let $h \in \Gamma(T^*M \otimes T^*M)$. We define

$$h^{\mathrm{sym}}(X, Y) = \tfrac{1}{2}(h(X, Y) + h(Y, X)),$$
$$h^{\mathrm{skew}}(X, Y) = \tfrac{1}{2}(h(X, Y) - h(Y, X)).$$

Definition 2.2. Let (g, J) be an almost Hermitian structure. Let $h \in \Gamma(T^*M \otimes T^*M)$. We define

$$h^{(1,1)}(X, Y) = \tfrac{1}{2}(h(X, Y) + h(JX, JY)),$$
$$h^{(0,2)+(2,0)}(X, Y) = \tfrac{1}{2}(h(X, Y) - h(JX, JY)).$$

We say that h is $(1, 1)$ or $(0, 2) + (2, 0)$ if $h^{(0,2)+(2,0)} = 0$ or $h^{(1,1)} = 0$, respectively.

In Lemma 2.3 and Lemma 2.6, we derive the necessary condition of a variation of almost Hermitian pair.

Lemma 2.3. *Let J_t be a family of almost complex structures, and let $(\partial/\partial t)J = K$. Then*

$$KJ + JK = 0.$$

Proof. By definition,

$$0=\frac{\partial}{\partial t}J^2=KJ+JK. \qquad \square$$

Lemma 2.4. *Let* (g, J) *be an almost Hermitian structure*, $K\in\Gamma(\mathrm{End}(TM))$. *Then*

$$KJ+JK=0 \iff K \text{ is } (0,2)+(2,0).$$

Proof. By definition,

$$\langle (KJ+JK)X, Y\rangle = K(JX, Y)-K(X, JY)=2K^{(1,1)}(JX, Y). \qquad \square$$

Remark 2.5. Similarly, $KJ=JK$ if and only if K is $(1,1)$.

Lemma 2.6. *Let* J_t *be a family of almost complex structures, and let* $(\partial/\partial t)J=K$. *Let* g_t *be a family of Riemannian structures compatible with* J_t, *and let* $(\partial/\partial t)g=h$. *Then*

$$K^{\mathrm{sym}}J=h^{(0,2)+(2,0)}.$$

Proof. By using the equation $KJ+JK=0$, we have

$$\begin{aligned}0&=\frac{\partial}{\partial t}(g(JX, JY)-g(X, Y))\\&=h(JX, JY)-h(X, Y)+g(KX, JY)+g(JX, KY)\\&=-2h^{(0,2)+(2,0)}(X, Y)+K(JX, Y)+K(Y, JX)\\&=-2h^{(0,2)+(2,0)}(X, Y)+2(K^{\mathrm{sym}}J)(X, Y). \qquad \square\end{aligned}$$

Lemma 2.7. *Let* (g, J) *be an almost Hermitian structure. Then* (L_Xg, L_XJ) *satisfies the necessary condition of a variation of* (g, J), *i.e.*,

- (i) L_Xg *is symmetric*,
- (ii) L_XJ *is* $(0,2)+(2,0)$,
- (iii) $(L_XJ)^{\mathrm{sym}}J=(L_Xg)^{(0,2)+(2,0)}$.

Proof. Let ϕ_t be the 1-parameter transformation groups generated by X, and let $g_t=\phi_t^*g$ and $J_t=\phi_t^*J$. Then

$$\frac{\partial}{\partial t}\Big|_{t=0}g_t=L_Xg, \quad \frac{\partial}{\partial t}\Big|_{t=0}J_t=L_XJ.$$

Then Lemma 2.7 follows from Lemmas 2.3, 2.4 and 2.6. $\square$

Lemma 2.8. *Let* (g, J) *be an almost Hermitian structure. Then*

$$\langle (D_XJ)Y, Z\rangle = -\langle (D_XJ)Z, Y\rangle,$$

$$(D_XJ)JY=-J(D_XJ)Y.$$

Proof. Let X, Y, Z be in a normal coordinate system. Then

$$\langle (D_X J)Y, Z\rangle = \langle D_X(JY), Z\rangle = X\langle JY, Z\rangle = -X\langle Y, JZ\rangle = -\langle (D_X J)Z, Y\rangle,$$

and

$$(D_X J)JY = D_X(JJY) - JD_X(JY) = -J(D_X J)Y. \qquad \square$$

Lemma 2.9 [Gauduchon 1997]. *Let* (g, J) *be an almost Hermitian structure. Then*

$$\langle (D_{JX} J)Y, Z\rangle - \langle J(D_X J)Y, Z\rangle = \tfrac{1}{2}(N(X, Y, Z) + N(Z, X, Y) - N(Y, Z, X)),$$
$$\langle (D_{JX} J)Y, Z\rangle + \langle J(D_X J)Y, Z\rangle = (d\omega)^+(JX, Y, Z) - (d\omega)^+(JX, JY, JZ).$$

In particular,

$$D_{JX}J = JD_XJ \iff N = 0,$$
$$D_{JX}J = -JD_XJ \iff (d\omega)^+ = 0.$$

3. Main calculations

First, we define the tensors we use in this paper.

Definition 3.1. Let (M, g, J) be an almost Hermitian manifold, $X, Y, Z \in TM$.

- $B^1(X, Y) = \langle (D_X J)i, (D_Y J)i\rangle$,
- $B^2(X, Y) = \langle (D_i J)X, (D_i J)Y\rangle$,
- $B^3(X, Y) = \langle (D_{(D_i J)X} J)i, Y\rangle = -\langle (D_i J)X, j\rangle\langle (D_j J)Y, i\rangle$,
- $B^4(X, Y) = \langle (D_X J)i, (D_i J)Y\rangle$,
- $\overline{B^1}(X, Y) = \langle (D_X J)i, (D_Y J)Ji\rangle$,
- $\overline{B^2}(X, Y) = \langle (D_i J)X, (D_{Ji} J)Y\rangle$,
- $Q_1 = -\frac{1}{2}(B^1)^{(1,1)} - (B^3)^{(0,2)+(2,0)} + 4(B^4)^{(1,1),\mathrm{sym}} - (\overline{B^1}J)^{(1,1)} - \overline{B^2}J$,
- $Q_2 = (B^3)^{(0,2)+(2,0)}J$,
- $\mathcal{N} = B^2 J$,
- $\mathcal{R}(X, Y) = \mathrm{Ric}(JX, Y) + \mathrm{Ric}(X, JY)$,
- $\mathfrak{Q} = B^2 J + B^3 J$,
- $H(X, Y, Z) = d^c\omega(X, Y, Z) = -d\omega(JX, JY, JZ)$,
- $\mathcal{B}(X, Y) = H(X, i, j)H(Y, i, j)$,
- $\theta^\sharp = -J(D_i J)i$,
- $\overline{N}(X, Y) = \frac{1}{2}\big(N((D_i J)X, i, Y) + N(Y, (D_i J)X, i) - N(i, Y, (D_i J)X)\big) - \frac{1}{2}\big(N(i, (D_X J)i, Y) + N(Y, i, (D_X J)i) - N((D_X J)i, Y, i)\big) - (D_i J)N(X, i)$,
- $\mathcal{K}(X) = (D_i N)(Ji, X)$,

- $(d\omega)^+(X, Y, Z) = \frac{1}{4}\big(3d\omega(X, Y, Z) + d\omega(JX, JY, Z) + d\omega(JX, Y, JZ) + d\omega(X, JY, JZ)\big)$.

The lemmas below are preparation for the proof of Theorem 1.1.

Lemma 3.2. *Let (g, J) be an almost Hermitian structure. Then $(-2\operatorname{Ric} + Q_1, \triangle J + \mathcal{N} + \mathcal{R} + Q_2)$ satisfies the necessary condition of a variation.*

Proof. First, we show that $(-2\operatorname{Ric}, \triangle J + \mathcal{N} + \mathcal{R})$ satisfies the necessary condition. We need to check the following things:

(i) Ric is symmetric,

(ii) $\triangle J + \mathcal{N}$ is $(0, 2) + (2, 0)$,

(iii) $\mathcal{R}$ is $(0, 2) + (2, 0)$,

(iv) $\triangle J$ is skew,

(v) $\mathcal{N}$ is skew,

(vi) $\mathcal{R}$ is symmetric,

(vii) $\mathcal{R}J = -2\operatorname{Ric}^{(0,2)+(2,0)}$.

By definition, it is easy to see (i), (iii), (vi), (vii). For (ii), we use normal coordinates to calculate the $(1, 1)$ part of $\triangle J$, by using Lemma 2.8:

$$\begin{aligned}\langle(\triangle J)(JX), JY\rangle &= \langle(D_i DJ)(i, JX), JY\rangle\\ &= \langle D_i((D_i J)(JX)) - (D_i J)(D_i(JX)), JY\rangle\\ &= -\langle D_i(J(D_i J)X) + (D_i J)(D_i(JX)), JY\rangle\\ &= -\langle(D_i J)(D_i J)X + JD_i((D_i J)X) + (D_i J)(D_i J)X, JY\rangle\\ &= -2\langle(D_i J)(JX), (D_i J)Y\rangle - \langle(D_i D_i J)X, Y\rangle\\ &= -2\mathcal{N} - \langle(\triangle J)X, Y\rangle.\end{aligned}$$

So $\mathcal{N} = -(\triangle J)^{(1,1)}$. For (iv), we also use normal coordinates:

$$\begin{aligned}\langle(\triangle J)X, Y\rangle &= \langle D_i((D_i J)X), Y\rangle\\ &= \partial_i\langle(D_i J)X, Y\rangle\\ &= \partial_i\langle D_i(JX), Y\rangle - \partial_i\langle J(D_i X), Y\rangle\\ &= \partial_i\partial_i\langle JX, Y\rangle - \partial_i\langle JX, D_i Y\rangle + \partial_i\langle D_i X, JY\rangle,\end{aligned}$$

so we see that $\triangle J$ is skew. And (v) follows from Lemma 2.8.

Next, we show that (Q_1, Q_2) satisfies the necessary condition. In fact, by applying Lemma 2.8, we can easily obtain that all terms in Q_1 are symmetric and all terms in Q_2 are $(0, 2)+(2, 0)$. And $Q_1^{(0,2)+(2,0)} = (B^3)^{(0,2)+(2,0)}$. This completes the proof. □

Lemma 3.3. *Let (g, J) be an almost Hermitian structure. Suppose $d\omega = 0$. Then*

$$\begin{aligned} Q_1 &= \tfrac{1}{2}B^1 - B^2, \\ Q_2 &= 0. \end{aligned}$$

Proof. Since $d\omega = 0$, by Lemma 2.8 and Lemma 2.9, one sees that B^1 and B^3 are $(1, 1)$, that $\overline{B^1}J = B^1$, and that $\overline{B^2}J = B^2$. Now we prove that $B^4 = \frac{1}{2}B^1$. In fact, we notice that

$$\langle (D_X J)Y, Z\rangle + \langle (D_Y J)Z, X\rangle + \langle (D_Z J)X, Y\rangle = d\omega(X, Y, Z) = 0.$$

Thus,

$$\begin{aligned} \langle (D_X J)i, (D_i J)Y\rangle &= \langle (D_i J)Y, j\rangle\langle (D_X J)i, j\rangle \\ &= -\langle (D_{(D_X J)i} J)Y, i\rangle \\ &= \langle (D_Y J)i, (D_X J)i\rangle + \langle (D_i J)(D_X J)i, Y\rangle \\ &= B^1(X, Y) - \langle (D_X J)i, (D_i J)Y\rangle. \end{aligned}$$

So $\langle (D_i J)X, (D_Y J)i\rangle = \frac{1}{2}B^1(X, Y)$. This completes the proof. □

Lemma 3.4. *Let (g, J) be an almost Hermitian structure. Suppose $N = 0$. Then*

$$\begin{aligned} Q_1 &= \tfrac{1}{2}\mathcal{B}, \\ Q_2 &= \mathfrak{Q} - \mathcal{N}. \end{aligned}$$

Proof. The proof is by direct calculations based on Lemma 2.8 and Lemma 2.9. We notice that B^1 is $(1, 1)$ and that B^3 is $(0, 2) + (2, 0)$. And $\overline{B^1} = B^1 J$ and $\overline{B^2} = B^2 J$. We also have $B^4 = 0$, since

$$\begin{aligned} \langle (D_X J)i, (D_i J)Y\rangle &= \langle (D_X J)Ji, (D_{Ji} J)Y\rangle \\ &= -\langle J(D_X J)i, J(D_i J)Y\rangle \\ &= -\langle (D_X J)i, (D_i J)Y\rangle. \end{aligned}$$

We can calculate $\mathcal{B}$ in terms of DJ:

$$\begin{aligned} \mathcal{B}(X, Y) &= H(X, i, j)H(Y, i, j) \\ &= d\omega(JX, Ji, Jj)d\omega(JY, Ji, Jj) \\ &= d\omega(JX, i, j)d\omega(JY, i, j). \end{aligned}$$

We have

$$d\omega(JX, i, j) = \langle (D_{JX} J)i, j\rangle + \langle (D_{Ji} J)j, X\rangle + \langle (D_{Jj} J)X, i\rangle.$$

Calculating term by term,

$$
\begin{aligned}
\langle (D_{JX}J)i,\, j\rangle\langle (D_{JY}J)i,\, j\rangle &= \langle (D_XJ)i,\, (D_YJ)i\rangle = B^1(X,Y),\\
\langle (D_{Ji}J)j,\, X\rangle\langle (D_{Ji}J)j,\, Y\rangle &= \langle (D_{Jj}J)X,\, i\rangle\langle (D_{Jj}J)Y,\, i\rangle = \langle (D_iJ)X,\, (D_iJ)Y\rangle\\
&= B^2(X,Y),\\
\langle (D_{JX}J)i,\, j\rangle\langle (D_{Ji}J)j,\, Y\rangle &= \langle (D_{JX}J)i,\, j\rangle\langle (D_{Jj}J)Y,\, i\rangle\\
&= -\langle (D_XJ)i,\, (D_iJ)Y\rangle = 0,\\
\langle (D_{JY}J)i,\, j\rangle\langle (D_{Ji}J)j,\, X\rangle &= \langle (D_{JY}J)i,\, j\rangle\langle (D_{Jj}J)X,\, i\rangle\\
&= -\langle (D_XJ)i,\, (D_iJ)Y\rangle = 0,\\
\langle (D_{Ji}J)j,\, X\rangle\langle (D_{Jj}J)Y,\, i\rangle &= \langle (D_{Ji}J)j,\, Y\rangle\langle (D_{Jj}J)X,\, i\rangle\\
&= -\langle (D_{(D_iJ)X}J)i,\, Y\rangle = -B^3(X,Y).
\end{aligned}
$$

So

$$\tfrac{1}{2}\mathcal{B} = \tfrac{1}{2}B^1 + B^2 - B^3.$$

Then we obtain the desired result. □

Remark 3.5. In [Streets and Tian 2012], $\mathfrak{Q}$ is defined as

$$
\begin{aligned}
\mathfrak{Q}(X) = {}&-(D_iJ)(D_{JX}J)i - J(D_{(D_iJ)X}J)i + (D_iJ)(D_{Ji}J)X\\
&- (D_{J(D_iJ)i}J)X + J(D_{(D_iJ)i}J)X + (D_{JX}J)(D_iJ)i - J(D_XJ)(D_iJ)i.
\end{aligned}
$$

Since $N = 0$, it coincides with our definition.

Lemma 3.6. *Let* (g, J) *be an almost Hermitian structure. Then*

$$L_{\theta^\sharp}J = \triangle J + \mathfrak{Q} + \mathcal{R} + \mathcal{K} + \bar{N}.$$

Proof. In [Streets and Tian 2012], there is a similar formula. But in our case we don't assume that $N = 0$.

We use normal coordinates:

$$
\begin{aligned}
(L_{\theta^\sharp}J)X &= (L_{-J(D_iJ)i}J)X\\
&= -[J(D_iJ)i,\, JX] + J[J(D_iJ)i,\, X]\\
&= -D_{J(D_iJ)i}(JX) + D_{JX}(J(D_iJ)i)\\
&\qquad + JD_{J(D_iJ)i}X - JD_X(J(D_iJ)i)\\
&= -(D_{J(D_iJ)i}J)X + (D_{JX}J)(D_iJ)i + JD_{JX}((D_iJ)i)\\
&\qquad - J(D_XJ)(D_iJ)i + D_X((D_iJ)i)\\
&= -(D_{J(D_iJ)i}J)X + (D_{JX}J)(D_iJ)i + J(D_{JX}(D_iJ))i\\
&\qquad - J(D_XJ)(D_iJ)i + D_X(D_iJ)
\end{aligned}
\tag{4}
$$

$$\begin{aligned}
&= J(D^2J)(JX,i,i) + (D^2J)(X,i,i) - (D_{J(D_iJ)i}J)X \\
&\quad + (D_{JX}J)(D_iJ)i - J(D_XJ)(D_iJ)i.
\end{aligned}$$

By the Ricci identity,

$$\begin{aligned}
(D^2J)(X,i,i) &= (D^2J)(i,X,i) + (\mathrm{Rm}(X,i)J)i \\
&= (D^2J)(i,X,i) + \mathrm{Rm}(X,i)(Ji) - J\,\mathrm{Rm}(X,i)i \\
&= (D^2J)(i,X,i) + \mathrm{Rm}(X,i)(Ji) - J\,\mathrm{Ric}(X).
\end{aligned} \tag{5}$$

Similarly,

$$J(D^2J)(JX,i,i) = J(D^2J)(i,JX,i) + J\,\mathrm{Rm}(JX,i)(Ji) + \mathrm{Ric}(JX). \tag{6}$$

Notice that

$$N(X,Y) = (D_{JX}J)Y - (D_{JY}J)X - J(D_XJ)Y + J(D_YJ)X.$$

Hence,

$$\begin{aligned}
J(D^2J)(i,JX,i) &= JD_i((D_{JX}J)i) - J(D_{(D_iJ)X}J)i \\
&\qquad - JD_i(J(D_XJ)i) + JD_i(J(D_XJ)i) \\
&= JD_i((D_{JX}J)i - J(D_XJ)i) - J(D_{(D_iJ)X}J)i \\
&\qquad + J(D_iJ)(D_XJ)i - (D^2J)(i,X,i) \\
&= JD_i((D_{Ji}J)X - J(D_iJ)X) + JD_i(N(X,i)) \\
&\qquad - J(D_{(D_iJ)X}J)i + J(D_iJ)(D_XJ)i - (D^2J)(i,X,i)
\end{aligned}$$

Notice that

$$\begin{aligned}
JD_i(N(X,i)) &= D_i(JN(X,i)) - (D_iJ)N(X,i) \\
&= D_i(N(Ji,X)) - (D_iJ)N(X,i) \\
&= (D_iN)(Ji,X) + N((D_iJ)i,X) - (D_iJ)N(X,i).
\end{aligned}$$

So

$$\begin{aligned}
J(D^2J)(i,JX,i) &= J(D^2J)(i,Ji,X) + J(D_{(D_iJ)i}J)X \\
&\quad - J(D_iJ)(D_iJ)X + (\triangle J)X + \mathcal{K}(X) + N((D_iJ)i,X) - (D_iJ)N(X,i) \\
&\quad - J(D_{(D_iJ)X}J)i + J(D_iJ)(D_XJ)i - (D^2J)(i,X,i).
\end{aligned} \tag{7}$$

And

$$\begin{aligned}
N((D_iJ)i,X) &= (D_{J(D_iJ)i}J)X - (D_{JX}J)(D_iJ)i \\
&\qquad + (D_{(D_iJ)i}J)JX - (D_XJ)J(D_iJ)i \\
&= (D_{J(D_iJ)i}J)X - (D_{JX}J)(D_iJ)i \\
&\qquad - J(D_{(D_iJ)i}J)X + J(D_XJ)(D_iJ)i.
\end{aligned} \tag{8}$$

By resorting to Lemma 2.9, we obtain

$$\begin{aligned}(9)\quad &\langle -J(D_{(D_iJ)X}J)i-(D_{(D_iJ)JX}J)i, Y\rangle\\ &=\langle -J(D_{(D_iJ)X}J)i+(D_{J(D_iJ)X}J)i, Y\rangle\\ &=\tfrac{1}{2}(N((D_iJ)X, i, Y)+N(Y, (D_iJ)X, i)-N(i, Y, (D_iJ)X)),\end{aligned}$$

and

$$\begin{aligned}(10)\quad &\langle J(D_iJ)(D_XJ)i, Y\rangle\\ &=\langle J(D_iJ)(D_XJ)i-(D_{Ji}J)(D_XJ)i, Y\rangle\\ &=-\tfrac{1}{2}(N(i, (D_XJ)i, Y)+N(Y.i, (D_XJ)i)-N((D_XJ)i, Y, i)).\end{aligned}$$

Then, by the Ricci identity again,

$$\begin{aligned}(11)\quad JD^2J(i, Ji, X)&=\tfrac{1}{2}(JD^2J(i, Ji, X)-JD^2J(Ji, i, X))\\ &=\tfrac{1}{2}J(\mathrm{Rm}(i, Ji)J)X\\ &=\tfrac{1}{2}(J\,\mathrm{Rm}(i, Ji)(JX)+\mathrm{Rm}(i, Ji)X).\end{aligned}$$

By the Bianchi identity,

$$\mathrm{Rm}(i, Ji)(JX)+\mathrm{Rm}(Ji, JX)i+\mathrm{Rm}(JX, i)(Ji)=0.$$

Notice that

$$\mathrm{Rm}(Ji, JX)i=\mathrm{Rm}(JX, i)(Ji).$$

Thus

$$(12)\quad J\,\mathrm{Rm}(i, Ji)(JX)=-2J\,\mathrm{Rm}(JX, i)(Ji),$$

$$(13)\quad \mathrm{Rm}(i, Ji)(X)=-2\,\mathrm{Rm}(X, i)(Ji).$$

Putting (4)–(13) together, we obtain the desired result. □

4. Proof of Theorem 1.1

The argument is the same as in [Streets and Tian 2014]. We use DeTurck trick to prove short-time existence and uniqueness.

We consider the following equations:

$$(14)\quad \begin{aligned}&\frac{\partial}{\partial t}g=-2\,\mathrm{Ric}+Q_1+L_Xg\triangleq\mathcal{D}_1(g, J),\\ &\frac{\partial}{\partial t}J=\triangle J+\mathcal{N}+\mathcal{R}+Q_2+L_XJ\triangleq\mathcal{D}_2(g, J),\\ &g(0)=g_0,\\ &J(0)=J_0,\end{aligned}$$

where $X = \operatorname{tr}_g(\Gamma - \overline{\Gamma})$ and $\overline{\Gamma}$ is the Christoffel symbol of a fixed metric $\bar{g}$.

Then, in order to use the PDE theory in Banach space, we consider the tangent space at J_0. Denote by $T\mathcal{J}_J$ the tangent space at J, i.e.,

$$T\mathcal{J}_J = \{E \in \operatorname{End}(TM) \mid EJ + JE\}.$$

Then, in a neighborhood U of J_0, we can identify J and E by using the map

$$\pi : T\mathcal{J}_{J_0} \supset U' \to U, \quad \pi E = -J_0 e^{J_0 E},$$

and note that $D\pi|_0 = \operatorname{Id}$.

Notice that we don't assume that (g, J) is compatible. So we need to make some modifications. For convenience, we write g^J and g^{-J} instead of $g^{(1,1)}$ and $g^{(0,2)+(2,0)}$, respectively, and we do similar things for other tensors. Note that g^J is compatible with J. We consider the following equations:

$$\begin{aligned}
&\frac{\partial}{\partial t} g = \mathcal{D}_1(g^{\pi E}, \pi E) + \triangle_{g_0}(g^{-\pi E}) \triangleq \widetilde{\mathcal{D}}_1(g, E),\\
&\frac{\partial}{\partial t} E = (D\pi|_{\pi E})^{-1}\mathcal{D}_2(g^{\pi E}, \pi E) \triangleq \widetilde{\mathcal{D}}_2(g, E),\\
&g(0) = g_0,\\
&E(0) = 0.
\end{aligned} \tag{15}$$

Note that $\widetilde{\mathcal{D}}_1$ is symmetric, and $\widetilde{\mathcal{D}}_2$ is well-defined since $\triangle J + \mathcal{N} + \mathcal{R} + Q_2 + L_X J$ is $(0, 2) + (2, 0)$ for the pair (g^J, J). So $\widetilde{\mathcal{D}}_1 \oplus \widetilde{\mathcal{D}}_2$ gives an operator from $\Gamma((T^*M \otimes^{\text{sym}} T^*M) \oplus T\mathcal{J}_{J_0})$ to itself.

Now, we calculate the symbol of $\widetilde{\mathcal{D}}_1 \oplus \widetilde{\mathcal{D}}_2$ at $(g_0, 0)$ to show the short-time existence of the modified flow. First, we calculate the variation of $\widetilde{\mathcal{D}}_1$ along the direction of $(h, 0)$, where $h = \delta g$. Since $\delta E = 0$, $\pi E = \pi 0 = J_0$ is fixed. And note that $\delta(g^{J_0}) = h^{J_0}$ and $g_0^{J_0} = g_0$. Therefore

$$\mathcal{L}_{(g_0,0)}(\mathcal{D}_1(g^{\pi E}, \pi E))(h, 0) = \mathcal{L}_{g_0^{J_0}}(\mathcal{D}_1(g, J_0))(h^{J_0}) = \mathcal{L}_{g_0}(\mathcal{D}_1(g, J_0))(h^{J_0}),$$

where $\mathcal{L}_{(g_0,0)}$ denotes the linearization operator at $(g_0, 0)$.

Noting that only $-2\operatorname{Ric}$ and $L_X g$ involve second-order terms, and from standard calculations in Ricci flow [Chow and Knopf 2004] we have

$$\mathcal{L}_{g_0}(\mathcal{D}_1(g, J_0))(h^{J_0}) = \triangle_{g_0}(h^{J_0}) + \mathcal{O}(\partial h).$$

And

$$\mathcal{L}_{(g_0,0)}(\triangle_{g_0}(g^{-J}))(h, 0) = \triangle_{g_0}(h^{-J_0}).$$

Let σ denote the symbol of a linear differential operator. Thus we obtain

$$\sigma(\mathcal{L}_{(g_0,0)}\widetilde{\mathcal{D}}_1)(h, 0)(x, \xi) = |\xi|^2 h, \quad \text{where } \xi \in T_x^* M.$$

Then we calculate the variation of $\widetilde{\mathscr{D}}_1$ along the direction of $(0, K)$, where $K = \delta E$. Since $D\pi|_0 = \mathrm{Id}$, we have

$$\delta(\widetilde{\mathscr{D}}_1(g, E))(0, K) = \delta(\mathscr{D}_1(g^J, J))(0, \delta J).$$

We identify δJ and K below.

From the calculations above, we see that

$$(-2\operatorname{Ric}(g^J) + L_{X(g^J)}(g^J))_{ij} = (g^J)^{pq}\partial_p\partial_q(g^J)_{ij} + \mathcal{O}(\partial g, \partial J).$$

So

$$\mathscr{L}_{(g_0,0)}(\mathscr{D}_1(g^{\pi E}, \pi E))(0, K) = \frac{\partial}{\partial t}\Big|_{t=0}(g_0)^{pq}\partial_p\partial_q(g_0^{J_t})_{ij} + \mathcal{O}(\partial K).$$

It is easy to see that

$$\mathscr{L}_{(g_0,0)}(\triangle_{g_0}(g^{-J}))(0, K) = \frac{\partial}{\partial t}\Big|_{t=0}(g_0)^{pq}\partial_p\partial_q(g_0^{-J_t})_{ij} + \mathcal{O}(\partial K).$$

Thus we obtain

$$\sigma(\mathscr{L}_{(g_0,0)}\widetilde{\mathscr{D}}_1)(0, K)(x, \xi) = 0, \quad \text{where } \xi \in T_x^*M.$$

Next, we calculate the variation of $\widetilde{\mathscr{D}}_2$ along the direction of $(\delta g, \delta E) = (h, K)$. We have

$$\delta(\widetilde{\mathscr{D}}_2(g, E))(h, K) = \delta(\mathscr{D}_2(g^J, J))(\delta g, \delta J).$$

In the expression for $\mathscr{D}_2$, only $\triangle J$, $L_X J$, and $\mathscr{R}$ involve second-order terms, so we only need to calculate these three terms. We calculate them for the pair (g, J) first.

For $\triangle J$, we have

$$\begin{aligned}
(\triangle J)(e_k) &= g^{ij}D^2J(e_i, e_j, e_k)\\
&= g^{ij}D_i((D_jJ)e_k) + \mathcal{O}(\partial g, \partial J)\\
&= g^{ij}D_i(D_j(Je_k) - JD_je_k) + \mathcal{O}(\partial g, \partial J)\\
&= g^{ij}D_i(D_j(J_k^le_l) - J(\Gamma_{jk}^pe_p)) + \mathcal{O}(\partial g, \partial J)\\
&= g^{ij}(D_i(\partial_jJ_k^le_l) + D_i(J_k^p\Gamma_{jp}^le_l) - D_i(\Gamma_{jk}^pJ_p^le_l)) + \mathcal{O}(\partial g, \partial J)\\
&= g^{ij}(\partial_i\partial_jJ_k^l + J_k^p\partial_i\Gamma_{jp}^l - J_p^l\partial_i\Gamma_{jk}^p)e_l + \mathcal{O}(\partial g, \partial J).
\end{aligned}$$

For L_XJ, we have

$$\begin{aligned}
(L_XJ)(e_k) &= [X, Je_k] - J[X, e_k]\\
&= [X^pe_p, J_k^le_l] - J[X^pe_p, e_k]\\
&= (X^p\partial_pJ_k^l - J_k^p\partial_pX^l + J_p^l\partial_kX^p)e_l\\
&= g^{ij}(J_p^l\partial_k\Gamma_{ij}^p - J_k^p\partial_p\Gamma_{ij}^l)e_l + \mathcal{O}(\partial g, \partial J).
\end{aligned}$$

For R, we have

$$\begin{aligned}\mathcal{R}(e_k) &= (J_k^p \operatorname{Ric}^l_p - J^l_p \operatorname{Ric}^p_k)e_l \\ &= g^{ij}(-J_k^p \partial_i \Gamma^l_{pj} + J_k^p \partial_p \Gamma^l_{ij} + J^l_p \partial_i \Gamma^p_{kj} - J^l_p \partial_k \Gamma^p_{ij})e_l + \mathcal{O}(\partial g, \partial J).\end{aligned}$$

So we obtain

$$(\triangle J + \mathcal{R} + L_X J)^l_k = g^{ij}\partial_i \partial_j J^l_k + \mathcal{O}(\partial g, \partial J).$$

As for the pair (g^J, J), the lower-order terms are still lower-order terms, and when we evaluate at (g_0, J_0), from the compatibility, we have

$$(\mathcal{L}_{(g_0,0)}\widetilde{\mathcal{D}}_2)(h, K) = \triangle_{g_0} K + \mathcal{O}(\partial g, \partial J).$$

Hence, the total symbol is

$$\sigma(\mathcal{L}_{(g_0,0)}\widetilde{\mathcal{D}})(h, K)(x, \xi) = \begin{pmatrix} |\xi|^2 & 0 \\ 0 & |\xi|^2 \end{pmatrix}.$$

By the standard theory of parabolic PDE, there exists a unique short-time solution of (15).

Next we show that, under (15), (g, J) is compatible, where $J = \pi E$. Suppose that (g, J) exists for $t \in [0, \epsilon_0]$. Then by the compactness of M, in this time interval, every tensor we involve is bounded. Let $(\partial/\partial t)J = K$. Then

$$\begin{aligned}\frac{\partial}{\partial t}|g^{-J}|^2_{g^J} &= 2\Big\langle \frac{\partial}{\partial t}(g^{-J}), g^{-J}\Big\rangle_{g^J} + C*(g^{-J})^{*2} \\ &= 2\Big\langle \frac{\partial}{\partial t}\frac{1}{2}(g(\cdot,\cdot) - g(J\cdot, J\cdot)), g^{-J}\Big\rangle_{g^J} + C*(g^{-J})^{*2} \\ &= 2\Big\langle \Big(\frac{\partial}{\partial t}g\Big)^{-J}, g^{-J}\Big\rangle_{g^J} - \langle g(J\cdot, K\cdot) + g(K\cdot, J\cdot), g^{-J}\rangle_{g^J} + C*(g^{-J})^{*2} \\ &\le \langle 2(\mathcal{D}_1(g^J, J))^{-J} + 2(\triangle_{g_0}(g^{-J}))^{-J} - g(J\cdot, K\cdot) - g(K\cdot, J\cdot), g^{-J}\rangle_{g^J} \\ &\qquad + C|g^{-J}|^2_{g^J}.\end{aligned}$$

Note that (g^J, J) is compatible and $K = \mathcal{D}_2(g^J, J)$, so by Lemmas 3.2 and 2.7,

$$\mathcal{D}_1(g^J, J)^{-J} - \tfrac{1}{2}(g^J(J\cdot, K\cdot) + g^J(K\cdot, J\cdot)) = 0.$$

So

$$\begin{aligned}\frac{\partial}{\partial t}|g^{-J}|^2_{g^J} &\le 2\langle (\triangle_{g_0}(g^{-J}))^{-J} - g^{-J}(J\cdot, K\cdot) - g^{-J}(K\cdot, J\cdot), g^{-J}\rangle_{g^J} + C|g^{-J}|^2_{g^J} \\ &\le 2\langle (\triangle_{g_0}(g^{-J}))^{-J}, g^{-J}\rangle_{g^J} + C|g^{-J}|^2_{g^J}.\end{aligned}$$

Since J acts isometrically on the space $\Gamma(T^*M \otimes^{\text{sym}} T^*M)$ in the induced metric from g^J, and since the $(1, 1)$ tensors and $(0, 2) + (2, 0)$ tensors correspond to the

$+1$ and -1 eigenspaces, respectively, they are orthogonal. So

$$\langle (\triangle_{g_0}(g^{-J}))^J,\, g^{-J}\rangle_{g^J} = 0.$$

Then,

$$\frac{\partial}{\partial t}|g^{-J}|^2_{g^J} \le 2\langle \triangle_{g_0}(g^{-J}),\, g^{-J}\rangle_{g^J} + C|g^{-J}|^2_{g^J}.$$

By definition,

$$\triangle_{g_0}(g^{-J}) = \mathrm{tr}_{g_0} D^2_{g_0}(g^{-J}).$$

Since the second order term about g^{-J} in $D^2_{g_0}(g^{-J})$ is the same as in $D^2_{g^J}(g^{-J})$,

$$\triangle_{g_0}(g^{-J}) = \mathrm{tr}_{g_0}\big(D^2_{g^J}(g^{-J}) + C' * D_{g^J}(g^{-J}) + C * g^{-J}\big).$$

Let A be any tensor. We have the formula

$$\begin{aligned} D^2\langle A,\, A\rangle &= D(D\langle A,\, A\rangle) \\ &= 2D(\langle D_i A,\, A\rangle e^i) \\ &= 2\langle D^2_{i,j} A,\, A\rangle e^i \otimes e^j + 2\langle D_i A,\, D_j A\rangle e^i \otimes e^j. \end{aligned}$$

Let $A = g^{-J}$ and the metric above be g^J. Taking the trace of each side with respect to g_0, we obtain

$$\begin{aligned} &2\langle \mathrm{tr}_{g_0} D^2_{g^J}(g^{-J}),\, g^{-J}\rangle_{g^J} \\ &\qquad = \mathrm{tr}_{g_0} D^2_{g^J}(|g^{-J}|^2_{g^J}) - 2\langle D_{g^J} g^{-J}(e_i),\, D_{g^J} g^{-J}(e_j)\rangle_{g^J}\langle e^i,\, e^j\rangle_{g_0}. \end{aligned}$$

Along this flow, for $t \in [0, \epsilon_0]$, g^J is uniformly bounded by g_0, so we have

$$2\langle \mathrm{tr}_{g_0} D^2_{g^J}(g^{-J}),\, g^{-J}\rangle_{g^J} \le \mathrm{tr}_{g_0} D^2_{g^J}(|g^{-J}|^2_{g^J}) - 2C''|D_{g^J} g^{-J}|^2_{g^J}.$$

Hence,

$$\frac{\partial}{\partial t}|g^{-J}|^2_{g^J} \le \mathrm{tr}_{g_0} D^2_{g^J}(|g^{-J}|^2_{g^J}) - 2C''|D_{g^J} g^{-J}|^2_{g^J} + C' * D_{g^J}(g^{-J}) * g^{-J} + C|g^{-J}|^2_{g^J}.$$

By using the Cauchy inequality on $C' * D_{g^J}(g^{-J}) * g^{-J}$, finally we obtain

$$\frac{\partial}{\partial t}|g^{-J}|^2_{g^J} \le \mathrm{tr}_{g_0} D^2_{g^J}(|g^{-J}|^2_{g^J}) + C|g^{-J}|^2_{g^J}.$$

Notice that $\mathrm{tr}_{g_0} D^2_{g^J}$ is elliptic and $|g^{-J}|^2 = 0$ at $t = 0$. Then by the maximum principle, considering $e^{-Ct}|g^{-J}|^2$, we have $|g^{-J}|^2 = 0$ for $t \in [0, \epsilon_0]$, i.e., (g, J) is compatible. Since ϵ_0 is arbitrary, (g, J) is always compatible as long as the solution exists. Because the positivity of g is an open condition, we may assume that g is positive in short time. Then the short-time solution of (15) gives the short-time solution of (14).

Now, let $(\tilde{g}(t), \tilde{J}(t))$ be a solution of (14) and let φ_t be the one-parameter family of diffeomorphisms generated by $-X(t)$ defined as above. Let $g(t) = \varphi_t^* \tilde{g}(t)$, $J(t) = \varphi_t^* \tilde{J}(t)$. Then

$$
\begin{aligned}
\frac{\partial}{\partial t} g &= \frac{\partial}{\partial t}(\varphi_t^* \tilde{g}(t)) \\
&= \varphi_t^* \Big(\frac{\partial}{\partial t} \tilde{g}(t) + L_{(-X(t))} \tilde{g}(t) \Big) \\
&= \varphi_t^* (-2 \operatorname{Ric}(\tilde{g}(t)) + Q_1(\tilde{g}(t))) \\
&= -2 \operatorname{Ric}(\varphi_t^* \tilde{g}(t)) + Q_1(\varphi_t^* \tilde{g}(t)) \\
&= -2 \operatorname{Ric}(g) + Q_1(g).
\end{aligned} \tag{16}
$$

So $g(t)$ satisfies the equation. Similarly, $J(t)$ also satisfies the equation. And $(g(t), J(t))$ differs from $(\tilde{g}(t), \tilde{J}(t))$ by a diffeomorphism, so $(g(t), J(t))$ is also an almost Hermitian pair. This completes the existence part of the theorem.

For uniqueness, let (g_i, J_i) be two solutions of (3), $i = 1, 2$. Since M is compact, we can solve the harmonic heat flow

$$
\begin{aligned}
\frac{\partial}{\partial t} \phi_i(t) &= \triangle_{g_i, \bar{g}} \phi_i(t), \\
\phi_i(0) &= \mathrm{Id},
\end{aligned}
$$

for $\phi_i(t)$ for short time, where $\bar{g}$ is the same fixed metric as above. We can also assume that the $\phi_i(t)$ are diffeomorphisms. Let $\hat{g}_i = (\phi_i^{-1}(t))^* g_i(t)$. Note that

$$
\begin{aligned}
\Big(\frac{\partial}{\partial t} \phi_i \Big)(p) &= (\triangle_{g_i, \bar{g}} \phi_i)(p) \\
&= (\triangle_{\hat{g}_i, \bar{g}} \mathrm{Id})(\phi_i(p)) \\
&= \Big(-\hat{g}^{ij} (\hat{\Gamma}_{ij}^k - \overline{\Gamma}_{ij}^k) \frac{\partial}{\partial x^k} \Big)(\phi_i(p)) \\
&= -X_{\hat{g}}(\phi_i(p)).
\end{aligned}
$$

Then, taking the time derivative of $(\phi_i(t))^* \hat{g}_i(t) = g_i(t)$, and doing a similar calculation to (16), we see that both $\hat{g}_i(t)$ satisfy (14) and they share the same initial data. Since we have proved the compatibility, the symbol of (14) is Id, as we calculated, so the solution of (14) is unique. Then we obtain

$$
\hat{g}_1(t) = \hat{g}_2(t) = \hat{g}(t), \quad \hat{J}_1(t) = \hat{J}_2(t) = \hat{J}(t).
$$

Then from the uniqueness of

$$
\begin{aligned}
\frac{\partial}{\partial t} \phi(t) &= -X_{\hat{g}}(\phi(t)), \\
\phi(0) &= \mathrm{Id},
\end{aligned}
$$

we see the uniqueness of (g, J) for a short while. Then, by continuity, (g, J) is unique as long as it exists.

Next, we check two special cases. Suppose that the initial data is almost Kähler. Then we run the symplectic curvature flow (1). By definitions and Lemma 3.3, we see that, in this situation, (g, J) also satisfies (3). So from the uniqueness of (3), if the initial data is almost Kähler, then (3) coincides with symplectic curvature flow. And a similar argument holds in the pluriclosed case when we apply Lemma 3.4.

Finally, we prove that the flow (3) preserves the integrability of J. Let (g_0, J_0) be an Hermitian structure. Fix J_0 and consider the flow

$$\begin{aligned}&\frac{\partial}{\partial t}\tilde{g} = -2\operatorname{Ric}_{\tilde{g}} + Q_1(\tilde{g}, J_0) - L_{\theta^\sharp(\tilde{g}, J_0)}\tilde{g},\\&\tilde{g}(0) = g_0.\end{aligned}$$

By the DeTurck trick, we see that $\tilde{g}(t)$ exists for a while, but is not necessarily compatible with J_0 now. Then by a gauge transformation induced by $\theta^\sharp(\tilde{g}, J_0)$, we obtain a short-time solution $(g(t), J(t))$ for the flow

$$\begin{aligned}&\frac{\partial}{\partial t}g = -2\operatorname{Ric}_g + Q_1(g, J),\\&\frac{\partial}{\partial t}J = L_{\theta^\sharp(g,J)}J,\\&g(0) = g_0,\\&J(0) = J_0.\end{aligned}$$

We still don't know the compatibility of (g, J) now, but since J is changed just by a diffeomorphism, N always vanishes. By Lemma 2.9, one may write $Q_2 - \mathfrak{Q} + \mathcal{N}$ in terms of N in the almost Hermitian setting. We denote such a tensor N_0, i.e., N_0 is in terms of N, and, when (g, J) is compatible, $N_0 = Q_2 - \mathfrak{Q} + \mathcal{N}$. So the above flow is the same as the flow

$$\begin{aligned}&\frac{\partial}{\partial t}g = -2\operatorname{Ric}_g + Q_1(g, J),\\&\frac{\partial}{\partial t}J = L_{\theta^\sharp(g,J)}J + N_0(g, J) - \overline{N}(g, J) - \mathcal{H}(g, J),\\&g(0) = g_0,\\&J(0) = J_0.\end{aligned}$$

Then by Lemma 3.6, and using the same argument in the proof of short-time existence above, one sees that (g, J) is compatible and coincides with (3), so the integrability of J is preserved.

This completes the proof of Theorem 1.1. □

Remark 4.1. Streets and Tian [2014] introduced almost Hermitian curvature flow, where the symbol term deforming J is $-\mathcal{H}$. From Lemma 3.6 we see that, modulo lower-order terms, $-\mathcal{K}$ differs from $\triangle J+\mathcal{R}$ just by a gauge term. If we also change the evolution of g by the same gauge transformation, the second derivative of g will appear in $L_{\theta^\sharp}g$. So, in general, our flow is not in the family of almost Hermitian curvature flow.

5. Proof of Theorem 1.2 and Theorem 1.3

First, we derive the evolution equations of DJ, Rm and their higher covariant derivatives.

Lemma 5.1. *Under* (3),

$$\frac{\partial}{\partial t}DJ=\triangle DJ+\mathrm{Rm}*DJ+J^{*2}*DJ^{*3}+J^{*3}*DJ*D^2J.$$

Proof. Using the fact $\triangle DT-D\triangle T=D\,\mathrm{Rm}*T+\mathrm{Rm}*DT$, we have

$$\begin{aligned}\frac{\partial}{\partial t}DJ&=\dot\Gamma*J+D\dot J\\&=D(\mathrm{Rm}+J^{*2}*DJ^{*2})*J+D(\triangle J+\mathrm{Rm}*J+J*DJ^{*2})\\&=\triangle DJ+D\,\mathrm{Rm}*J+\mathrm{Rm}*DJ+J^{*2}*DJ^{*3}+J^{*3}*DJ*D^2J.\end{aligned}$$

Hence we only need to show there is no $D\,\mathrm{Rm}*J$ term. It is the same calculation as in [Streets and Tian 2014], since the only differences are the first-order terms in J, which does not involve a $D\,\mathrm{Rm}$ term. □

Lemma 5.2. *Under* (3),

$$\frac{\partial}{\partial t}\mathrm{Rm}=\triangle\,\mathrm{Rm}+\mathrm{Rm}^{*2}+\mathrm{Rm}*J^{*2}*DJ^{*2}+\sum_{\substack{0\le k_1,\dots,k_4\le 3\\k_1+\cdots+k_4=4}}D^{k_1}J*\cdots*D^{k_4}J.$$

Proof. Let $(\partial/\partial t)g=h$. From the variation formula in Ricci flow (see [Chow and Knopf 2004]) we have

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{Rm}(X,Y,Z,W)=\tfrac{1}{2}(&h(\mathrm{Rm}(X,Y)Z,W)-h(\mathrm{Rm}(X,Y)W,Z))\\&+\tfrac{1}{2}(D^2_{Y,W}h(X,Z)-D^2_{X,W}h(Y,Z)\\&+D^2_{X,Z}h(Y,W)-D^2_{Y,Z}h(X,W)).\end{aligned}$$

And, when $h=-2\,\mathrm{Ric}$,

$$\frac{\partial}{\partial t}\mathrm{Rm}=\triangle\,\mathrm{Rm}+\mathrm{Rm}^{*2}.$$

Notice that, in (3), $h=(\partial/\partial t)g=-2\,\mathrm{Ric}+J^{*2}*DJ^{*2}$, so we obtain the evolution equation of Rm. □

Proposition 5.3. *Under* (3),

$$\frac{\partial}{\partial t}D^kJ=\triangle D^kJ+\sum_{\substack{l_1+\cdots+l_5=k+2\\0\le l_1,\ldots,l_5\le k+1}}D^{l_1}J*\cdots*D^{l_5}J+\sum_{l=0}^{k-1}D^l\operatorname{Rm}*D^{k-l}J$$

and

$$\frac{\partial}{\partial t}D^k\operatorname{Rm}=\triangle D^k\operatorname{Rm}+\sum_{\substack{l_1+\cdots+l_4=k+4\\0\le l_1,\ldots,l_4\le k+3}}D^{l_1}J*\cdots*D^{l_4}J+\sum_{l=0}^{k}D^l\operatorname{Rm}*D^{k-l}\operatorname{Rm}$$
$$+\sum_{0\le l_0\le k}\sum_{\substack{l_1+\cdots+l_4=k+2-l_0\\0\le l_1,\ldots,l_4\le k+1}}D^{l_0}\operatorname{Rm}*D^{l_1}J*\cdots*D^{l_4}J.$$

Proof. By using Lemma 5.1 and the fact that $(\partial/\partial t)\Gamma=D(\operatorname{Rm}+J^{*2}*DJ^{*2})$, we have

$$\begin{aligned}\frac{\partial}{\partial t}D^kJ&=\frac{\partial}{\partial t}\Gamma*D^{k-1}J+D\frac{\partial}{\partial t}D^{k-1}J\\&=\sum_{l=0}^{k-2}D^l\frac{\partial}{\partial t}\Gamma*D^{k-1-l}J+D^{k-1}\frac{\partial}{\partial t}DJ\\&=\sum_{l=0}^{k-2}D^lD(\operatorname{Rm}+J^{*2}*DJ^{*2})*D^{k-1-l}J\\&\qquad+D^{k-1}(\triangle DJ+\operatorname{Rm}*DJ+J^{*2}*DJ^{*3}+J^{*3}*DJ*D^2J).\end{aligned}$$

Interchanging D and $\triangle$, we observe that the highest order of Rm is $k-1$, and the highest order of J is $k+1$ if not involving Rm. Then we obtain the evolution equation of D^kJ.

As for the evolution equation of $D^k\operatorname{Rm}$, the calculation is similar. The key point is to observe the highest order. □

Now we can use Proposition 5.3 to prove Theorem 1.2 and Theorem 1.3.

Proof of Theorem 1.2. The proof is similar to the higher derivative estimates in Ricci flow [Chow and Knopf 2004]. We assume $t|D^2J|\le C$ first. By induction, we will prove

$$(P)\qquad |D^kJ|\le\frac{C}{t^{k/2}},\quad |D^{k-2}\operatorname{Rm}|\le\frac{C}{t^{k/2}}.$$

(P) holds when $k=2$ from the assumption.

Now we assume (P) holds for $k-1$. Consider

$$F(t)=t^{k+1}(|D^kJ|^2+|D^{k-2}\operatorname{Rm}|^2)+\lambda t^k(|D^{k-1}J|^2+|D^{k-3}\operatorname{Rm}|^2),$$

where λ is a large constant to be determined. We will show that

$$\frac{\partial}{\partial t}F \leq \triangle F + C. \tag{17}$$

Then, by the maximum principle, (P) holds for k. Now we prove (17) by using Proposition 5.3:

$$\begin{aligned}
&\frac{\partial}{\partial t}|D^k J|^2 \\
&\quad = (\mathrm{Rm} + J^{*2} * DJ^{*2}) * D^k J^{*2} + 2\Big\langle D^k J, \\
&\qquad\qquad \triangle D^k J + \sum_{\substack{l_1+\cdots+l_5=k+2\\ 0\leq l_1,\ldots,l_5\leq k+1}} D^{l_1}J * \cdots * D^{l_5}J + \sum_{l=0}^{k-1} D^l \,\mathrm{Rm} * D^{k-l}J\Big\rangle \\
&\quad = (\mathrm{Rm} + J^{*2} * DJ^{*2}) * D^k J^{*2} + \triangle|D^k J|^2 - 2|D^{k+1}J|^2 \\
&\qquad + D^k J * \Bigg(\sum_{\substack{l_1+\cdots+l_5=k+2\\ 0\leq l_1,\ldots,l_5\leq k+1}} D^{l_1}J * \cdots * D^{l_5}J + \sum_{l=0}^{k-1} D^l \,\mathrm{Rm} * D^{k-l}J\Bigg) \\
&\quad = \triangle|D^k J|^2 - 2|D^{k+1}J|^2 + (\mathrm{Rm} + J^{*2} * DJ^{*2}) * D^k J^{*2} \\
&\qquad + D^k J * D^{k+1}J * DJ * J^{*3} + D^k J * D^k J * DJ^{*2} * J^{*2} \\
&\qquad + D^k J * D^k J * D^2 J * J^{*3} + D^k J * \sum_{\substack{l_1+\cdots+l_5=k+2\\ 0\leq l_1,\ldots,l_5\leq k-1}} D^{l_1}J * \cdots * D^{l_5}J \\
&\qquad + D^k J * \mathrm{Rm} * D^k J + D^k J * D^{k-1}\,\mathrm{Rm} * DJ + D^k J * D^{k-2}\,\mathrm{Rm} * D^2 J \\
&\qquad + D^k J * \sum_{l=1}^{k-3} D^l \,\mathrm{Rm} * D^{k-l}J.
\end{aligned}$$

From the assumption,

$$\begin{aligned}
\frac{\partial}{\partial t}|D^k J|^2 &\leq \triangle|D^k J|^2 - 2|D^{k+1}J|^2 + \frac{C}{t}|D^k J|^2 + \frac{C}{t^{1/2}}|D^k J||D^{k+1}J| \\
&\quad + \frac{C}{t^{(k+2)/2}}|D^k J| + \frac{C}{t^{1/2}}|D^k J||D^{k-1}\,\mathrm{Rm}| + \frac{C}{t}|D^k J||D^{k-2}\,\mathrm{Rm}|.
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}|D^{k-2}\,\mathrm{Rm}|^2 &\leq \triangle|D^{k-2}\,\mathrm{Rm}|^2 - 2|D^{k-1}\,\mathrm{Rm}|^2 + \frac{C}{t}|D^{k-2}\,\mathrm{Rm}|^2 \\
&\quad + \frac{C}{t^{1/2}}|D^{k-2}\,\mathrm{Rm}||D^{k+1}J| + \frac{C}{t^{(k+2)/2}}|D^{k-2}\,\mathrm{Rm}| + \frac{C}{t}|D^k J||D^{k-2}\,\mathrm{Rm}|.
\end{aligned}$$

Then, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\frac{\partial}{\partial t}\big(t^{k+1}(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2)\big) &\le \triangle\big(t^{k+1}(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2)\big)\\ &\quad -t^{k+1}\big(|D^{k+1} J|^2+|D^{k-1}\operatorname{Rm}|^2\big)+Ct^k\big(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2\big)+C.\end{aligned}$$

Replacing k with $k-1$ and using the assumption, we obtain

$$\begin{aligned}&\frac{\partial}{\partial t}\big(t^{k}(|D^{k-1} J|^2+|D^{k-3}\operatorname{Rm}|^2)\big)\\ &\quad\le \triangle\big(t^{k}(|D^{k-1} J|^2+|D^{k-3}\operatorname{Rm}|^2)\big)-t^k\big(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2\big)\\ &\qquad +Ct^{k-1}\big(|D^{k-1} J|^2+|D^{k-3}\operatorname{Rm}|^2\big)+C\\ &\quad\le \triangle\big(t^{k}(|D^{k-1} J|^2+|D^{k-3}\operatorname{Rm}|^2)\big)-t^k\big(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2\big)+C.\end{aligned}$$

Then

$$\begin{aligned}\frac{\partial F}{\partial t} &\le \triangle F-t^{k+1}\big(|D^{k+1} J|^2+|D^{k-1}\operatorname{Rm}|^2\big)\\ &\quad +(C-\lambda)t^k\big(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2\big)+C\\ &\le \triangle F+(C-\lambda)t^k\big(|D^k J|^2+|D^{k-2}\operatorname{Rm}|^2\big)+C.\end{aligned}$$

We choose $\lambda=C$, so (17) holds.

Now, we prove that $t|D^2J|\le C$. For $p\in M$, if $|D^2J|_{p,t}\ne 0$, then similarly, by Proposition 5.3,

$$\begin{aligned}\frac{\partial}{\partial t}|D^2J| &= \frac{1}{2|D^2J|}\frac{\partial}{\partial t}|D^2J|^2\\ &= \frac{1}{2|D^2J|}\big(\triangle|D^2J|^2-2|D^3J|^2+D^2J^{*3}*J^{*3}\\ &\qquad +D^3J*D^2J*DJ*J^{*3}+D^2J^{*2}*DJ^{*2}*J^{*2}\\ &\qquad +D^2J*DJ^{*4}+D^2J^{*2}*\operatorname{Rm}+D^2J*DJ*D\operatorname{Rm}\big).\end{aligned}$$

Notice that, for $|D^2J|_{p,t}\ne 0$,

$$\triangle|D^2J|^2=2|D^2J|\triangle|D^2J|+2\big|D|D^2J|\big|^2.$$

So,

$$\begin{aligned}\frac{\partial}{\partial t}|D^2J| &= \triangle|D^2J|+\frac{\big|D|D^2J|\big|^2}{|D^2J|}+\frac{1}{2|D^2J|}\big(-2|D^3J|^2+D^2J^{*3}*J^{*3}\\ &\qquad +D^3J*D^2J*DJ*J^{*3}+D^2J^{*2}*DJ^{*2}*J^{*2}\\ &\qquad +D^2J*DJ^{*4}+D^2J^{*2}*\operatorname{Rm}+D^2J*DJ*D\operatorname{Rm}\big)\end{aligned}$$

$$\begin{aligned}&\leq \triangle|D^2J| + \frac{\big|D|D^2J|\big|^2}{|D^2J|} - \frac{|D^3J|^2}{|D^2J|}\\ &\quad + C\bigg(|D^2J|^2 + \frac{|D^3J|}{t^{1/2}} + \frac{|D^2J|}{t} + \frac{1}{t^2} + \frac{|D\,\mathrm{Rm}|}{t^{1/2}}\bigg).\end{aligned}$$

Consider

$$G(t) = t^2|D^2J| + \mu t^2|DJ|^2 + t^3|\mathrm{Rm}|^2,$$

where μ is a large constant to be determined.

Then, for $|D^2J| \neq 0$,

$$\begin{aligned}\frac{\partial}{\partial t}G &\leq \triangle G - t^2\frac{|D^3J|^2}{|D^2J|} - 2\mu t^2|D^2J|^2 - 2t^3|D\,\mathrm{Rm}|^2\\ &\quad + C\big(t^2|D^2J|^2 + t^{3/2}|D^3J| + \mu t|D^2J| + \mu + t^{3/2}|D\,\mathrm{Rm}|\big)\\ &\quad + \bigg\langle D|t^2D^2J|,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle\\ &\leq \triangle G - \tfrac{1}{2}t^2\frac{|D^3J|^2}{|D^2J|} - \tfrac{1}{2}t^2|D^2J|^2 - \tfrac{1}{2}t^3|D\,\mathrm{Rm}|^2\\ &\quad + \bigg\langle D|t^2D^2J|,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle + C,\end{aligned}$$

where μ is determined now.

Then

$$\begin{aligned}\frac{\partial}{\partial t}G &\leq \triangle G - \tfrac{1}{2}t^2\frac{|D^3J|^2}{|D^2J|} - \tfrac{1}{2}t^2|D^2J|^2 - \tfrac{1}{2}t^3|D\,\mathrm{Rm}|^2 + C\\ &\quad + \bigg\langle DG,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle - \mu t^2\bigg\langle D|DJ|^2,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle - t^3\bigg\langle D|\mathrm{Rm}|^2,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle.\end{aligned}$$

Notice that

$$\big|D|DJ|^2\big| \leq |2\langle DDJ,\, DJ\rangle| \leq 2|D^2J||DJ|, \quad \big|D|D^2J|\big| = \frac{\big|D|D^2J|^2\big|}{2|D^2J|} \leq |D^3J|.$$

Hence,

$$\begin{aligned}\frac{\partial}{\partial t}G &\leq \triangle G - \tfrac{1}{4}t^2\frac{|D^3J|^2}{|D^2J|} - \tfrac{1}{4}t^2|D^2J|^2 - \tfrac{1}{2}t^3|D\,\mathrm{Rm}|^2 + C\\ &\qquad + \bigg\langle DG,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle + C\frac{t^2|D\,\mathrm{Rm}|^2}{|D^2J|}.\end{aligned}$$

So if we suppose that $|D^2J| \geq 4C/t$, we have the estimate

$$\frac{\partial}{\partial t}G \leq \triangle G + \bigg\langle DG,\ \frac{D|D^2J|}{|D^2J|}\bigg\rangle + C, \tag{18}$$

where $C = C(n, K)$. That is to say, for any (p, t), either we have the estimate $|D^2 J| \le 4C/t$, or else (18) holds. Let $\overline{G} = G - Ct$, where C is chosen suitably. We obtain that either $\overline{G} \le 0$ or

$$\frac{\partial}{\partial t}\overline{G} \le \triangle\overline{G} + \left\langle D\overline{G}, \frac{D|D^2 J|}{|D^2 J|}\right\rangle.$$

Notice that $\overline{G} = 0$ when $t = 0$. Then one may apply the maximum principle to show that $\overline{G} \le 0$ for every (p, t), which implies the desired estimate. This completes the proof of Theorem 1.2. □

Remark 5.4. Theorem 1.2 is scaling-invariant when we replace $g(t)$ by $\bar{g}(t) = cg(t/c)$.

Proof of Theorem 1.3. The argument is standard, as in Ricci flow [Chow and Knopf 2004]. We just sketch the proof.

Suppose not. Then $|\mathrm{Rm}|$, $|DJ|$ are bounded. From Theorem 1.2, all covariant derivatives of Rm and J are bounded. Then we see that the metrics g are uniformly bounded. We fix a coordinate atlas. From the evolution equation of Γ and the boundedness of the covariant derivatives of Rm and J, we obtain the boundedness of Γ. Then we obtain the boundedness of ∂g, ∂J, and by induction we see that $\partial^k g$, $\partial^k J$ and $\partial^k \Gamma$ are bounded. Finally, we obtain that $(\partial^l/\partial t^l)\partial^k g$, $(\partial^l/\partial t^l)\partial^k J$ are bounded. Then, by theorems in mathematical analysis, $(g(t), J(t))$ can be extended to $(g(T), J(T))$ smoothly in all variables of space and time. The almost Hermitian condition is guaranteed by the continuity. Then, from the short-time existence, $(g(t), J(t))$ exists for $t \in [0, T + \epsilon)$, which is a contradiction to the maximality of T. □

Acknowledgements

The author wishes to express his gratitude to his advisor Gang Tian, for suggesting to him the problem of constructing new curvature flows preserving generalized complex structure, for encouraging him all the time, and for many helpful discussions. The author would also like to thank Jeffrey Streets for his helpful comments and suggestions, especially for pointing out that this flow may also preserve the integrability of J. The author would also like to thank CSC and TRAM for supporting the author visiting Princeton University.

References

[Boling 2014] J. Boling, "Homogeneous solutions of pluriclosed flow on closed complex surfaces", preprint, 2014. arXiv 1404.7106

[Chow and Knopf 2004] B. Chow and D. Knopf, *The Ricci flow: An introduction*, Mathematical Surveys and Monographs **110**, Amer. Math. Soc., Providence, RI, 2004. MR 2005e:53101 Zbl 1086.53085

[Enrietti 2013] N. Enrietti, "Static SKT metrics on Lie groups", *Manuscripta Math.* **140**:3-4 (2013), 557–571. MR 3019139 Zbl 1308.32026

[Enrietti et al. 2015] N. Enrietti, A. Fino, and L. Vezzoni, "The pluriclosed flow on nilmanifolds and tamed symplectic forms", *J. Geom. Anal.* **25**:2 (2015), 883–909. MR 3319954 Zbl 06444563

[Fernández-Culma 2013] E. Fernández-Culma, "Soliton almost Kähler structures on 6-dimensional nilmanifolds for the symplectic curvature flow", preprint, 2013. arXiv 1303.5461

[Gauduchon 1997] P. Gauduchon, "Hermitian connections and Dirac operators", *Boll. Un. Mat. Ital. B* (7) **11**:2 (1997), 257–288. MR 98c:53034 Zbl 0876.53015

[Gualtieri 2011] M. Gualtieri, "Generalized complex geometry", *Ann. of Math.* (2) **174**:1 (2011), 75–123. MR 2012h:53185 Zbl 1235.32020

[Hitchin 2003] N. Hitchin, "Generalized Calabi–Yau manifolds", *Q. J. Math.* **54**:3 (2003), 281–308. MR 2004h:32024 Zbl 1076.32019

[Newlander and Nirenberg 1957] A. Newlander and L. Nirenberg, "Complex analytic coordinates in almost complex manifolds", *Ann. of Math.* (2) **65** (1957), 391–404. MR 19,577a Zbl 0079.16102

[Perelman 2002] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2002. arXiv math/0211159

[Pook 2012] J. Pook, "Homogeneous and locally homogeneous solutions to symplectic curvature flow", preprint, 2012. arXiv 1202.1427

[Smith 2013] D. J. Smith, "Stability of the almost Hermitian curvature flow", preprint, 2013. arXiv 1308.6214

[Streets and Tian 2010] J. Streets and G. Tian, "A parabolic flow of pluriclosed metrics", *Int. Math. Res. Not.* **2010**:16 (2010), 3101–3133. MR 2011h:53091 Zbl 1198.53077

[Streets and Tian 2011] J. Streets and G. Tian, "Hermitian curvature flow", *J. Eur. Math. Soc.* (*JEMS*) **13**:3 (2011), 601–634. MR 2012f:53142 Zbl 1214.53055

[Streets and Tian 2012] J. Streets and G. Tian, "Generalized Kähler geometry and the pluriclosed flow", *Nuclear Phys. B* **858**:2 (2012), 366–376. MR 2881439 Zbl 1246.53091

[Streets and Tian 2013] J. Streets and G. Tian, "Regularity results for pluriclosed flow", *Geom. Topol.* **17**:4 (2013), 2389–2429. MR 3110582 Zbl 1272.32022

[Streets and Tian 2014] J. Streets and G. Tian, "Symplectic curvature flow", *J. Reine Angew. Math.* **696** (2014), 143–185. MR 3276165 Zbl 1305.53083

[Vezzoni 2011] L. Vezzoni, "On Hermitian curvature flow on almost complex manifolds", *Differential Geom. Appl.* **29**:5 (2011), 709–722. MR 2012m:53146 Zbl 1225.53030

Received July 4, 2014. Revised September 25, 2014.

SONG DAI
SCHOOL OF MATHEMATICAL SCIENCE
PEKING UNIVERSITY
NO. 5 YIHEYUAN ROAD
HAIDIAN DISTRICT
BEIJING, 100871
CHINA
daisong0620@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.313

REPRESENTATIONS OF KNOT GROUPS INTO $\mathrm{SL}_n(\mathbb{C})$ AND TWISTED ALEXANDER POLYNOMIALS

MICHAEL HEUSENER AND JOAN PORTI

Let Γ be the fundamental group of the exterior of a knot in the three-sphere. We study deformations of representations of Γ into $\mathrm{SL}_n(\mathbb{C})$ which are the sum of two irreducible representations. For such representations we give a necessary condition, in terms of the twisted Alexander polynomial, for the existence of irreducible deformations. We also give a more restrictive sufficient condition for the existence of irreducible deformations. We also prove a duality theorem for twisted Alexander polynomials and we describe the local structure of the representation and character varieties.

1. Introduction

Let $K \subset S^3$ be an oriented knot in the three-sphere. Its exterior is the compact three-manifold $X = S^3 \setminus \mathcal{N}(K)$. Set $\Gamma = \pi_1(X)$ and let $\varphi : \Gamma \twoheadrightarrow \mathbb{Z}$ denote the abelianization morphism, so that $\varphi(\gamma)$ is the linking number in S^3 between any loop realizing $\gamma \in \Gamma$ and K. Let

$$\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C}) \quad \text{and} \quad \beta : \Gamma \to \mathrm{SL}_b(\mathbb{C})$$

be *irreducible* and *infinitesimally regular* representations.

Both authors were partially supported by Mineco through grant MTM2012-34834. Heusener was also supported by the ANR projects ModGroup and SGT (*Structures Géométriques Triangulées*).
MSC2010: primary 57M25, 57M05; secondary 57M27.
Keywords: variety of representations, character variety, twisted Alexander polynomial, deformations.

Definition 1.1. A representation $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ is called *reducible* when it preserves a proper subspace of $\mathbb{C}^a$, otherwise it is called *irreducible*. The representation α is called *semisimple* or *completely reducible* if α is a direct sum of irreducible representations.

In what follows we call a representation $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ *infinitesimally regular* if $H^1(\Gamma; \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha}) \cong \mathbb{C}^{a-1}$.

As we assume that α is irreducible and infinitesimally regular, its character is a regular point of the character variety of Γ in $\mathrm{SL}_a(\mathbb{C})$ (Proposition 3.6). When $b = 1$, then β is trivial and hence it is infinitesimally regular.

For a given nonzero complex number $\lambda \in \mathbb{C}^*$ we consider the representation $\rho_\lambda = (\lambda^{b\varphi} \otimes \alpha) \oplus (\lambda^{-a\varphi} \otimes \beta)$, namely for all $\gamma \in \Gamma$

$$\rho_\lambda(\gamma) = \begin{pmatrix} \lambda^{b\varphi(\gamma)}\alpha(\gamma) & 0 \\ 0 & \lambda^{-a\varphi(\gamma)}\beta(\gamma) \end{pmatrix} \in \mathrm{SL}_n(\mathbb{C}), \tag{1}$$

where $a + b = n$. The representation $\rho_\lambda : \Gamma \to \mathrm{SL}_n(\mathbb{C})$ is reducible and the following question then arises:

Question 1.2. When can ρ_λ be deformed to irreducible representations?

We give necessary and sufficient conditions in terms of twisted Alexander polynomials. For this purpose we consider the representations

$$\alpha \otimes \beta^* : \Gamma \to \mathrm{Aut}(M_{a\times b}(\mathbb{C}))$$

defined by $(\alpha \otimes \beta^*)(\gamma)(A) = \alpha(\gamma)A\beta(\gamma^{-1})$ for $\gamma \in \Gamma$ and $A \in M_{a\times b}(\mathbb{C})$. Similarly, consider

$$\beta \otimes \alpha^* : \Gamma \to \mathrm{Aut}(M_{b\times a}(\mathbb{C})).$$

The corresponding twisted Alexander polynomials of degree i are denoted by

$$\Delta_i^+(t) = \Delta_i^{\alpha\otimes\beta^*}(t) \quad \text{and} \quad \Delta_i^-(t) = \Delta_i^{\beta\otimes\alpha^*}(t).$$

Recall that the twisted Alexander polynomial is a generator of the order ideal of the twisted Alexander module and hence it is unique up to multiplication with an invertible element of the group ring $\mathbb{C}[\mathbb{Z}] \cong \mathbb{C}[t^{\pm 1}]$, i.e., ct^k, with $c \in \mathbb{C}^*$ and $k \in \mathbb{Z}$ (see Definition 2.1 for more details). We have $\Delta_i^\pm(t) = 1$ for $i > 2$ and $\Delta_2^\pm(t) \in \{0, 1\}$. We prove in Corollary 4.6 that $\alpha \otimes \beta^*$ is a *semisimple* representation, hence by Theorem 2.6 we obtain the duality formula (Corollary 4.7):

$$\Delta_i^+(t) \doteq \Delta_i^-(1/t).$$

Here $p \doteq q$ means that p and q are *associated elements* in $\mathbb{C}[\mathbb{Z}]$, i.e., there exists some unit $ct^k \in \mathbb{C}[\mathbb{Z}] \cong \mathbb{C}[t^{\pm 1}]$, with $c \in \mathbb{C}^*$ and $k \in \mathbb{Z}$, such that $p = ct^k q$. This duality formula is a particular case of Theorem 2.6, where we establish a duality formula for twisted Alexander polynomials provided that the twisting representation is semisimple. This duality formula can also be deduced from results of Friedl, Kim, and Kitayama [Friedl et al. 2012].

We shall prove a necessary condition for the deformability of ρ_λ to irreducible representations:

Theorem 1.3. *If ρ_λ can be deformed to irreducible representations, then*

$$\Delta_1^+(\lambda^n) = \Delta_1^-(\lambda^{-n}) = 0.$$

The theorem also applies when α or β (or both) is trivial. When both α and β are trivial, this is a result obtained in 1967 independently by Burde [1967] and de Rham [1967]. The key idea is to look at the dimension of the fiber of the algebraic quotient $R(\Gamma, \mathrm{SL}_n(\mathbb{C})) \to X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. When ρ_λ can be deformed to irreducible representations, this dimension jumps among characters of reducible representations, and this translates to the twisted Alexander polynomial by means of the tangent space and cohomology with twisted coefficients.

The next result is a sufficient condition for the deformability of ρ_λ to irreducible representations:

Theorem 1.4. *If $\Delta_0^+(\lambda^n) \neq 0$ and λ^n is a simple root of $\Delta_1^+(t)$, then ρ_λ can be deformed to irreducible representations.*

Again this theorem and the next one apply for α and/or β trivial. Theorems 1.4 and 1.5 are due to [Heusener et al. 2001] when both α and β are trivial, and also related results were obtained in [Shors 1991; Frohman and Klassen 1991; Heusener and Klassen 1997; Heusener and Kroll 1998; Ben Abdelghani 2000; Ben Abdelghani and Lines 2002; Heusener and Porti 2005; Ben Abdelghani et al. 2010; Heusener and Medjerab 2014].

The outline of the proof of Theorem 1.4 is the following: the hypothesis implies that there exists a representation $\rho^+ \in R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ with the same character as ρ_λ but not conjugate to it (see Corollary 5.6). An analysis of the cohomology groups allows us to prove that ρ^+ is a smooth point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. Among other tools, this uses the vanishing of obstructions to integrability of Zariski tangent vectors, due to [Goldman 1984], a smoothness result of the variety of representations due to [Heusener and Medjerab 2014], and the nonvanishing of certain cup product (following the ideas of [Ben Abdelghani 2000]). Once this smoothness result is established, we realize that the dimension of the space of reducible representations is less than the dimension of the component of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ containing ρ^+.

Our next result concerns the local structure of the character variety. Let $\chi_\lambda \in X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ denote the character of ρ_λ.

Theorem 1.5. *Under the hypotheses of Theorem 1.4, χ_λ belongs to precisely two components Y and Z of $X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$, that have dimension $n-1$ and meet transversally at χ_λ along a subvariety of dimension $n-2$. The component Y contains characters of irreducible representations and Z consists only of characters of reducible ones.*

As in [Heusener et al. 2001] and [Heusener and Porti 2005] for $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{PSL}_2(\mathbb{C})$ respectively, the key idea for Theorem 1.5 is to study the quadratic cone of the representation ρ_λ, by identifying certain obstructions to integrability. Here we also use Luna's slice theorem, as in [Ben Abdelghani 2002].

We conclude the paper by an explicit description of the component of the variety of irreducible characters of the trefoil knot in $\mathrm{SL}_3(\mathbb{C})$ that illustrates our results.

The paper is organized as follows. Section 2 is devoted to twisted Alexander modules, and in particular to the duality theorem, Theorem 2.6. In Section 3 we review some preliminaries on the representation varieties and in Section 4 some further preliminaries on twisted cohomology and twisted invariants. Then in Section 5 we prove Theorem 1.3. The proof of the sufficient condition, Theorem 1.4, splits in Sections 6 and 7. Theorem 1.5 is proved in Section 8. Finally in Section 9 we compute $X(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ for Γ the fundamental group of the trefoil knot exterior.

2. Twisted Alexander modules

The aim of this section is to introduce twisted Alexander modules and Alexander polynomials, together with their main properties. We also give a new result that we will require later: a duality theorem for Alexander polynomials twisted by semisimple representations. It relies on Franz–Milnor duality for Reidemeister torsion, but it is different, as the torsion is the ratio of the Alexander polynomials. For further background about twisted Alexander polynomials see [Kirk and Livingston 1999].

A representation of a group Γ in a finite-dimensional complex vector space V is a homomorphism $\rho : \Gamma \to \mathrm{GL}(V)$. We say that such a map gives V the structure of a Γ-module. If there is no ambiguity about the map ρ we call V itself a representation of Γ and we will often suppress the symbol ρ and write $\gamma \cdot v$ or γv for $\rho(\gamma)(v)$. Two representations $\rho : \Gamma \to \mathrm{GL}(V)$ and $\varrho : \Gamma \to \mathrm{GL}(W)$ are called *equivalent* if there exits an isomorphism $T : V \to W$ such that $\varrho(\gamma) \circ T = T \circ \rho(\gamma)$ for all $\gamma \in \Gamma$, i.e., if the Γ-modules V and W are isomorphic.

Our main reference for group cohomology is [Brown 1994]. Since we work with left-modules, for defining homology consider the right action of the inverse, as in [Kirk and Livingston 1999, (2.1)]. As the knot exterior X is an Eilenberg–MacLane space, (co)homology groups of Γ and X are naturally identified. In what follows, we will not distinguish between $H_i(\Gamma; V)$ and $H_i(X; V)$.

We give an interpretation of the low dimensional (co)homology groups. The cohomology group in dimension zero is the module of invariants, i.e.,

$$H^0(\Gamma;\, V) \cong V^\Gamma = \{v \in V \mid \gamma v = v \text{ for all } \gamma \in \Gamma\}.$$

The homology group in dimension zero is the co-invariant module:

$$H_0(\Gamma;\, V) \cong \mathbb{Z} \otimes_{\mathbb{Z}[\Gamma]} V \cong V/IV$$

where $I \subset \mathbb{Z}[\Gamma]$ is the augmentation ideal and $IV \subset V$ is the subspace generated by $\{\gamma v - v \mid v \in V, \gamma \in \Gamma\}$.

We will make use of the interpretation of $H^1(\Gamma;\, V)$ by means of crossed morphisms, it is well suited for our purpose. A *crossed morphism* $d : \Gamma \to V$ is a map that satisfies $d(\gamma_1\gamma_2) = d(\gamma_1) + \gamma_1\, d(\gamma_2)$ for all $\gamma_1, \gamma_2 \in \Gamma$. A crossed morphism d is called *principal* if there exists $v \in V$ satisfying $d(\gamma) = \gamma v - v$ for all $\gamma \in \Gamma$. Crossed morphisms are precisely the *cocycles* of the standard or bar resolution of the Γ-module V, and the principal ones are the *coboundaries*. Thus the set of crossed morphisms or cocycles is denoted by $Z^1(\Gamma;\, V)$ and the set of principal crossed morphisms or coboundaries by $B^1(\Gamma;\, V)$. In particular, the first cohomology group is

$$H^1(\Gamma;\, V) \cong Z^1(\Gamma;\, V)/B^1(\Gamma;\, V). \tag{2}$$

Let $\rho : \Gamma \to \mathrm{GL}(V)$ be a finite dimensional representation of Γ. If $X_\infty \to X$ denotes the infinite cyclic covering, then

$$H_i(X_\infty;\, V)$$

is a finitely generated $\mathbb{C}[\mathbb{Z}]$-module, because X is compact and V is finite dimensional. Here $\mathbb{Z}$ is the group of deck transformations of the covering $X_\infty \to X$. We will sometimes interpret the elements of $\mathbb{C}[\mathbb{Z}]$ as Laurent polynomials, by using the isomorphism $\mathbb{C}[\mathbb{Z}] \xrightarrow{\sim} \mathbb{C}[t^{\pm 1}]$ that maps the generator 1 of $\mathbb{Z}$ to t.

Definition 2.1. The homology groups $H_i(X_\infty;\, V)$ are called the *twisted Alexander modules*, viewed as $\mathbb{C}[\mathbb{Z}] \cong \mathbb{C}[t^{\pm 1}]$-modules. The corresponding orders are the *twisted Alexander polynomials*

$$\Delta_i^\rho(t) \in \mathbb{C}[t^{\pm 1}].$$

They are unique up to multiplication by a unit $ct^k \in \mathbb{C}[t^{\pm 1}]$, $k \in \mathbb{Z}$, $c \in \mathbb{C}^*$.

Recall that the *order* of a finitely generated $\mathbb{C}[t^{\pm 1}]$-module

$$M = \bigoplus_i \mathbb{C}[t^{\pm 1}] \big/ p_i(t)\mathbb{C}[t^{\pm 1}]$$

is $\prod_i p_i(t)$. In particular the order is nonzero if and only if M is a torsion module. Notice that this is not the same convention as in [Kirk and Livingston 1999].

Due to the indeterminacy in the definition of twisted Alexander polynomials, we shall write

$$p(t) \doteq q(t)$$

to denote that the polynomials $p(t), q(t) \in \mathbb{C}[\mathbb{Z}]$ are *associated*, i.e., they are equal up to multiplication with an element $ct^k \in \mathbb{C}[\mathbb{Z}]$, $k \in \mathbb{Z}$, $c \in \mathbb{C}^*$.

Remark 2.2. It follows from a result of M. Wada [1994, Theorem 2] that the twisted Alexander polynomial of a link exterior twisted by a representation in $\mathrm{SL}_n(\mathbb{C})$ is well defined up to powers of $\pm t^k$. It is also well known that for n even there is no sign ambiguity. We shall not need those facts, as we use essentially the structure of the Alexander module.

Let

$$V[\mathbb{Z}] = V \otimes_{\mathbb{C}[\Gamma]} \mathbb{C}[\mathbb{Z}]$$

denote the Γ-module via the representation $\rho \otimes t^{\varphi}$. Then we have a natural isomorphism of $\mathbb{C}[\mathbb{Z}]$-modules

$$H_i(X; V[\mathbb{Z}]) \cong H_i(X_\infty; V) \tag{3}$$

(see [Kirk and Livingston 1999, Theorem 2.1]). Notice that equivalent representations give rise to isomorphic Γ-modules and hence to associated Alexander polynomials.

The *dual representation* $\rho^* : \Gamma \to \mathrm{GL}(V^*)$ is defined in the usual way by

$$\rho^*(\gamma)(f) = f \circ \rho(\gamma)^{-1} \quad \text{for all } \gamma \in \Gamma \text{ and } f \in V^* = \mathrm{Hom}(V, \mathbb{C}).$$

The following lemma is straightforward.

Lemma 2.3. *The representations ρ and ρ^* are equivalent if and only if there exists a nondegenerate bilinear form $V \otimes V \to \mathbb{C}$ which is Γ-invariant.*

Example 2.4. For any representation $\rho : \Gamma \to \mathrm{SL}_2(\mathbb{C})$, the module $V = \mathbb{C}^2$ has a skew-symmetric nondegenerate bilinear form defined by the determinant. Namely, the vectors (x_1, x_2) and $(y_1, y_2) \in \mathbb{C}^2$ are mapped to

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

In view of Lemma 2.3, ρ^* and ρ are equivalent and hence $\Delta_i^{\rho} \doteq \Delta_i^{\rho^*}$.

Recall from the introduction (see Definition 1.1) that a representation $\rho : \Gamma \to \mathrm{GL}(V)$ is called semisimple or completely reducible if ρ is the direct sum of irreducible representations.

Remark 2.5. A representation ρ is completely reducible if and only if each subspace of V stable under $\rho(\Gamma)$ has a $\rho(\Gamma)$-invariant complement.

Theorem 2.6. *Let* $\rho : \Gamma \to \mathrm{GL}(V)$ *be a completely reducible representation. Then*

$$\Delta_i^{\rho}(t^{-1}) \doteq \Delta_i^{\rho^*}(t).$$

Example 2.9 below shows that the hypothesis of complete reducibility is necessary in Theorem 2.6. This duality formula can also be deduced from results of Friedl, Kim, and Kitayama [2012].

The first step in the proof of Theorem 2.6 is the following:

Lemma 2.7. *Let* $\rho : \Gamma \to \mathrm{GL}(V)$ *be a completely reducible representation. The modules* $H_0(X_\infty; V)$ *and* $H_0(X_\infty; V^*)$ *are finitely generated torsion modules. In addition,*

$$\Delta_0^{\rho}(t^{-1}) \doteq \Delta_0^{\rho^*}(t).$$

Proof. First notice that if ρ is irreducible or completely reducible then the dual representation ρ^* is also irreducible or completely reducible respectively since each proper invariant subspace of ρ corresponds to a proper invariant subspace of ρ^* by the *orthogonality* relation.

We have that $H_0(X_\infty; V) \cong V/\tilde{I}V$, where $\tilde{I} \subset \mathbb{C}[\pi_1(X_\infty)]$ is the augmentation ideal. Hence, $H_0(X_\infty; V)$ is a finite dimensional $\mathbb{C}$-vector space and as $\mathbb{C}[t^{\pm 1}]$-module it cannot have a free summand. This proves that $H_0(X_\infty; V)$ is a finitely generated torsion module.

In order to prove the symmetry relation it is sufficient to prove it for irreducible representations since for $\rho_1 : \Gamma \to \mathrm{GL}(V_1)$ and $\rho_2 : \Gamma \to \mathrm{GL}(V_2)$ we have

$$(\rho_1 \oplus \rho_2)^* = \rho_1^* \oplus \rho_2^* \quad \text{and} \quad \Delta_i^{\rho_1 \oplus \rho_2} \doteq \Delta_i^{\rho_1} \cdot \Delta_i^{\rho_2}.$$

First we will prove that for every irreducible representation $\rho : \Gamma \to \mathrm{GL}(V)$ with $\dim V > 1$ we have

$$\Delta_0^{\rho} \doteq 1 \doteq \Delta_0^{\rho^*}. \tag{4}$$

The irreducibility of ρ and $\dim V > 1$ imply that $IV \subset V$ is a nontrivial Γ-invariant subspace, and hence $IV = V$. It follows that $H_0(\Gamma; V) = 0$. Now, for any complex number $\lambda \in \mathbb{C}^*$ the vector space V becomes a Γ-module via $\rho \otimes \lambda^{\varphi}$, i.e., for $\gamma \in \Gamma$ and for $v \in V$ we have $\rho(\gamma) \otimes \lambda^{\varphi(\gamma)}(v) = \lambda^{\varphi(\gamma)} \rho(\gamma) v$. This Γ-module will be denoted by V_λ. Notice that V_λ is also an irreducible Γ-module since the map $v \mapsto \lambda^{\varphi(\gamma)} v$ is a homothety of V. Moreover, V_λ is a nontrivial Γ-module and hence $H_0(\Gamma; V_\lambda) = 0$ for all $\lambda \in \mathbb{C}^*$. Next, the short exact sequence of Γ-modules

$$0 \to V[\mathbb{Z}] \xrightarrow{(t-\lambda)\cdot} V[\mathbb{Z}] \to V_\lambda \to 0$$

induces a long exact sequence in homology [Brown 1994, III.§6]:

$$\cdots \to H_0(\Gamma; V[\mathbb{Z}]) \xrightarrow{(t-\lambda)\cdot} H_0(\Gamma; V[\mathbb{Z}]) \to H_0(\Gamma; V_\lambda) \to 0,$$

and $H_0(\Gamma;\, V_\lambda) = 0$ implies that the multiplication by $(t-\lambda)$ is surjective. Hence for all $\lambda \in \mathbb{C}^*$, the module $H_0(\Gamma;\, V[\mathbb{Z}])$ has no $(t-\lambda)$-torsion. Hence, $H_0(\Gamma;\, V[\mathbb{Z}]) = 0$ and $\Delta_0^\rho = 1$. Finally, ρ^* is also irreducible and $\dim V^* = \dim V > 1$. This implies in the same way that $\Delta_0^{\rho^*} = 1$

Now suppose that $\dim V = 1$, i.e., $\rho : \Gamma \to \mathrm{GL}(V) \cong \mathbb{C}^*$. Hence ρ is abelian and completely determined by a nonzero-complex number λ, meaning that for all $\gamma \in \Gamma$ and $v \in V$ we have $\rho(\gamma)(v) = \lambda^{\varphi(\gamma)} v$. So we write $\rho = \lambda^\varphi$. Now

$$H_0(\Gamma;\, V[\mathbb{Z}]) \cong V[\mathbb{Z}]/IV[\mathbb{Z}] \cong V[t^{\pm 1}]/(\lambda t - 1),$$

since λ^φ is an abelian representation and factors through $\mathbb{Z}$. Therefore $\Delta_0^{\lambda^\varphi}(t) \doteq t - \lambda^{-1}$. The dual representation $(\lambda^\varphi)^*$ is $\lambda^{-\varphi}$, as $(\lambda^\varphi)^*(\gamma)(f) = f \circ (\lambda^{\varphi(\gamma)})^{-1} = \lambda^{-\varphi(\gamma)} f$, where $\gamma \in \Gamma$ and $f \in V^*$. The same calculation as above shows that $H_0(\Gamma;\, V^*[\mathbb{Z}]) \cong V[t^{\pm 1}]/(\lambda^{-1} t - 1)$ and hence $\Delta_0^{(\lambda^\varphi)^*}(t) \doteq t - \lambda$. We obtain $\Delta_0^{(\lambda^\varphi)*}(t) \doteq \Delta_0^{\lambda^\varphi}(t^{-1})$, which proves the lemma. □

Proof of Theorem 2.6. The knot exterior X has the homotopy type of a 2-dimensional complex. Therefore $H_i(X_\infty;\, V) = 0$ for $i > 2$ and $H_2(X_\infty;\, V)$ is a free $\mathbb{C}[\mathbb{Z}]$-module. This implies that $\Delta_i^\rho \doteq 1$ for $i > 2$ and $\Delta_2^\rho \in \{0, 1\}$. According to the value of Δ_2^ρ there are two cases to study.

Assume first that $\Delta_2^\rho = 0$. This is equivalent to $H_2(X_\infty;\, V)$ being a nontrivial free $\mathbb{C}[\mathbb{Z}]$-module. By an Euler characteristic argument, $H_1(X_\infty;\, V)$ contains also a nontrivial free factor of the same rank. In particular $\Delta_1^\rho = 0$. Since $H_i(X_\infty;\, V) \cong H_i(X;\, V[\mathbb{Z}])$, the universal coefficient theorem yields that $H_i(X;\, V[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)) \neq 0$ for $i = 1, 2$. Notice also that the natural pairing $V \times V^* \to \mathbb{C}$ extends to a nondegenerate $\mathbb{C}(t)$-bilinear form

$$\big(V[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)\big) \times \big(V^*[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)\big) \to \mathbb{C}(t). \tag{5}$$

Using this bilinear form and Poincaré duality, $H_i(X, \partial X;\, V^*[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)) \neq 0$ for $i = 1, 2$. Since the homology of the 2-torus ∂X with coefficients $V^*[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)$ vanishes [Kirk and Livingston 1999, §3.3], $H_i(X;\, V^*[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)) \neq 0$ for $i = 1, 2$. Hence $\Delta_1^{\rho^*} = \Delta_2^{\rho^*} = 0$.

Next we deal with the case $\Delta_2^\rho \doteq 1$. Since this is equivalent to $H_2(X_\infty;\, V) = 0$, the homology argument in the previous paragraph gives $\Delta_2^{\rho^*} \doteq 1$. For the first Alexander polynomials we shall use Reidemeister torsion and Franz–Milnor duality. By Kitano's theorem [1996] the torsion of X with coefficients $V[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)$ is the ratio of Alexander polynomials:

$$\mathrm{TOR}(X;\, V[\mathbb{Z}] \otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)) \doteq \frac{\Delta_1^\rho}{\Delta_0^\rho},$$

see [Kirk and Livingston 1999, Theorem 3.4] for this precise statement (this is a version of Milnor's theorem [1962], see [Turaev 1986]).

Using the bilinear form (5), Franz–Milnor duality for Reidemeister torsion [Milnor 1962; Franz 1937] gives

$$\begin{aligned}\operatorname{TOR}\bigl(X;\, V[\mathbb{Z}]\otimes_{\mathbb{C}[\mathbb{Z}]}\mathbb{C}(t)\bigr)(t) &\doteq \operatorname{TOR}\bigl(X,\partial X;\, V^*[\mathbb{Z}]\otimes_{\mathbb{C}[\mathbb{Z}]}\mathbb{C}(t)\bigr)\bigl(\tfrac{1}{t}\bigr)\\ &\doteq \frac{\operatorname{TOR}\bigl(X;\, V^*[\mathbb{Z}]\otimes_{\mathbb{C}[\mathbb{Z}]}\mathbb{C}(t)\bigr)\bigl(\tfrac{1}{t}\bigr)}{\operatorname{TOR}\bigl(\partial X;\, V^*[\mathbb{Z}]\otimes_{\mathbb{C}[\mathbb{Z}]}\mathbb{C}(t)\bigr)\bigl(\tfrac{1}{t}\bigr)},\end{aligned}$$

see [Kirk and Livingston 1999, §5.1]. Since $\partial X \cong S^1\times S^1$, $\operatorname{TOR}(\partial X;\, V^*[\mathbb{Z}]\otimes_{\mathbb{C}[\mathbb{Z}]} \mathbb{C}(t)) \doteq 1$ [Kirk and Livingston 1999, §3.3]. Then the theorem follows from Lemma 2.7. □

Remark 2.8. Note that every representation $\rho : \Gamma \to \mathrm{O}(n)$ is completely reducible since for each stable subspace W the orthogonal complement $W^\perp$ is also stable. Moreover, we have $\rho^* = \rho$ and hence $\Delta_i^\rho(t^{-1}) \doteq \Delta_i^\rho(t)$ is symmetric (see [Kitano 1996, Theorem B]). It follows also from the proof of Lemma 2.7 that $\Delta_0^\rho(t) = (t-1)^{k_+}(t+1)^{k_-}$ where $k_+ = \dim\{v\in\mathbb{R}^n \mid \rho(\gamma)v = v \text{ for all } \gamma\in\Gamma\}$ and $k_- = \dim\{v\in\mathbb{R}^n \mid \rho(\gamma)v = (-1)^{\varphi(\gamma)}v \text{ for all } \gamma\in\Gamma\}$.

It was proved in Hillman, Silver, and Williams [2010] that $\Delta_i^\rho(t^{-1}) \doteq \Delta_i^\rho(t)$ holds if ρ^* and ρ are conjugates.

We finish this section with an example to show that the hypothesis of complete reducibility is needed in Theorem 2.6:

Example 2.9. We exhibit representations that are not completely reducible and such that the conclusion of Theorem 2.6 fails. In order to construct such a representation, we take $\rho : \Gamma \to \mathrm{SL}_2(\mathbb{C})$ of the form

$$\rho = \begin{pmatrix}1 & d\\ 0 & 1\end{pmatrix}\begin{pmatrix}\lambda^\varphi & 0\\ 0 & \lambda^{-\varphi}\end{pmatrix}$$

that is not abelian. It is a representation if $d\in Z^1(\Gamma;\mathbb{C}_{\lambda^2})$ and it is nonabelian if $\lambda\neq\pm1$ and $d\notin B^1(\Gamma;\mathbb{C}_{\lambda^2})$, where $\mathbb{C}_{\lambda^2}$ denotes the Γ-module given by $\gamma\cdot z = \lambda^{2\varphi(\gamma)}z$ for $\gamma\in\Gamma$, $z\in\mathbb{C}$, see Lemma 5.5. Such a representation exists if and only if λ^2 is a root of the untwisted Alexander polynomial (in particular $\lambda\neq\pm1$), see [Burde 1967; de Rham 1967; Heusener et al. 2001] for instance, or Lemma 5.5. As ρ is not abelian, its restriction to $\pi_1(X_\infty)$ is nontrivial but

$$\rho(\pi_1(X_\infty)) \subset \left\{\begin{pmatrix}1 & c\\ 0 & 1\end{pmatrix} \,\middle|\, c\in\mathbb{C}\right\}.$$

The cohomology module $H_0(X_\infty;\mathbb{C}^2)$ is isomorphic to $\mathbb{C}^2/I\mathbb{C}^2$. Here the subspace $I\mathbb{C}^2\subset\mathbb{C}^2$ is generated by elements of the form $v-\rho(\gamma)v$, with $\gamma\in\pi_1(X_\infty)$

and $v \in \mathbb{C}^2$, i.e., $I\mathbb{C}^2 = \left\{ \binom{c}{0} \mid c \in \mathbb{C} \right\}$. So, the linear projection $\mathbb{C}^2 \to \mathbb{C}$ onto the second coordinate induces a linear isomorphism $\mathbb{C}^2/I\mathbb{C}^2 \xrightarrow{\cong} \mathbb{C}$. The action of a meridian $m \in \Gamma$ on $\mathbb{C}^2/I\mathbb{C}^2$ is multiplication by λ^{-1} and hence $H_0(X_\infty; \mathbb{C}^2) \cong \mathbb{C}[t^{\pm 1}]/(t - \lambda^{-1})$ as $\mathbb{C}[\mathbb{Z}]$-modules. Therefore, $\Delta_0^\rho(t) = t - \lambda^{-1}$. On the other hand, using that every representation in $\mathrm{SL}_2(\mathbb{C})$ is equivalent to its dual, see Example 2.4, $\Delta_0^{\rho^*}(t) \doteq t - \lambda^{-1}$, and

$$\Delta_0^\rho(t^{-1}) \doteq (t - \lambda) \text{ and } \Delta_0^{\rho^*}(t) \text{ are not associated.}$$

Notice that if $\Delta_2^\rho \doteq 1$, Franz–Milnor duality (used in the proof of Theorem 2.6) applies and it holds that $\Delta_1^\rho(t^{-1})/\Delta_0^\rho(t^{-1}) \doteq \Delta_1^{\rho^*}(t)/\Delta_0^{\rho^*}(t)$. In particular $\Delta_1^\rho(t^{-1})$ and $\Delta_1^{\rho^*}(t)$ are not associated either.

3. Varieties of representations

In this section we recall some preliminaries on the varieties of representations, we discuss representations of the peripheral subgroup $\pi_1(\partial X) \cong \mathbb{Z} \oplus \mathbb{Z}$, and we state a regularity result, Proposition 3.3 due to [Heusener and Medjerab 2014]. We also show that infinitesimal regularity implies regularity of the representation (Corollary 3.5) and its character (Proposition 3.6).

Recall that the set of all representations of Γ in $\mathrm{SL}_n(\mathbb{C})$ is called the *variety of representations* or the $\mathrm{SL}_n(\mathbb{C})$*-representation variety*:

$$R(\Gamma, \mathrm{SL}_n(\mathbb{C})) = \mathrm{Hom}(\Gamma, \mathrm{SL}_n(\mathbb{C})).$$

It is an affine algebraic set (possibly with several components), as Γ is finitely generated. More precisely, $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ embeds in a Cartesian product $\mathrm{SL}_n(\mathbb{C}) \times \cdots \times \mathrm{SL}_n(\mathbb{C})$ by mapping each representation to the image of a generating set, and $\mathrm{SL}_n(\mathbb{C})$ is an algebraic group in $\mathbb{C}^{n^2}$. The group relations of a presentation of Γ induce the algebraic equations defining $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. Different presentations give isomorphic algebraic sets (see [Lubotzky and Magid 1985], for instance).

The group $\mathrm{SL}_n(\mathbb{C})$ acts on $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ by conjugation. The algebraic quotient by this action is the *variety of characters* or $\mathrm{SL}_n(\mathbb{C})$*-character variety*

$$X(\Gamma, \mathrm{SL}_n(\mathbb{C})) = R(\Gamma, \mathrm{SL}_n(\mathbb{C})) /\!\!/ \mathrm{SL}_n(\mathbb{C}).$$

Recall that the GIT quotient exists since $\mathrm{SL}_n(\mathbb{C})$ is *reductive* and the representation variety is an affine algebraic set. (For more details see [Newstead 1978, 3.§3] or [Shafarevich 1994].)

To describe the Zariski tangent space to $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ and $X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ we use crossed morphisms or cocycles.

An *infinitesimal deformation* of a representation is the same as a Zariski tangent vector to $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. We use André Weil's construction that identifies

$Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})$ with the Zariski tangent space to the scheme $\mathcal{R}(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ at ρ. Here $\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}$ is a Γ-module via the adjoint action, i.e., $\gamma \cdot x = \mathrm{Ad}_{\rho(\gamma)}(x)$ for $\gamma \in \Gamma$ and $x \in \mathfrak{sl}_n(\mathbb{C})$. Notice furthermore that the algebraic equations defining the representation variety may be nonreduced, hence there is an underlying affine scheme $\mathcal{R}(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ with a possible nonreduced coordinate ring. Weil's construction assigns to each cocycle $d \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C}))$ the infinitesimal deformation $\gamma \mapsto (1 + \varepsilon\, d(\gamma))\rho(\gamma)$ for all $\gamma \in \Gamma$, which satisfies the defining equations for $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ up to terms in the ideal (ε^2) of $\mathbb{C}[\varepsilon]$, i.e., a Zariski tangent vector to $\mathcal{R}(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. Weil's construction identifies $B^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})$ with the tangent space to the orbit by conjugation. See [Weil 1964; Lubotzky and Magid 1985; Ben Abdelghani 2002] for more details.

Let $\dim_\rho R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ denote the local dimension of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ at ρ (i.e., the maximal dimension of the irreducible components of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ containing ρ [Shafarevich 1977, Chapter II]). So we obtain:

$$\dim_\rho R(\Gamma, \mathrm{SL}_n(\mathbb{C})) \leq \dim T_\rho(R(\Gamma, \mathrm{SL}_n(\mathbb{C}))) \leq \dim Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}). \tag{6}$$

Definition 3.1. Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{C})$ be a representation. We say that ρ is a *regular point* of the representation variety if

$$\dim_\rho R(\Gamma, \mathrm{SL}_n(\mathbb{C})) = \dim Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}).$$

We call ρ *infinitesimal regular* if $\dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) = n - 1$.

It follows directly from (6) that a regular point is a smooth point of the representation variety. There are representations of discrete groups which are smooth points of the representation variety without being regular, as the scheme $\mathcal{R}(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ may be nonreduced. (See [Lubotzky and Magid 1985, Example 2.10] for more details.)

We also make use of the Poincaré–Lefschetz duality theorem with twisted coefficients: let M be a connected, orientable, compact m-dimensional manifold with boundary ∂M and let $\rho : \pi_1(M) \to \mathrm{SL}_n(\mathbb{C})$ be a representation. Then the cup product and the Killing form $b : \mathfrak{sl}_n(\mathbb{C}) \otimes \mathfrak{sl}_n(\mathbb{C}) \to \mathbb{C}$ induce a nondegenerate bilinear pairing

$$\begin{aligned} H^k(M; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \otimes H^{m-k}(M, \partial M; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \xrightarrow{\smile} \\ H^m(M, \partial M; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho} \otimes \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \xrightarrow{b} H^m(M, \partial M; \mathbb{C}) \cong \mathbb{C} \end{aligned} \tag{7}$$

and hence an isomorphism

$$H^k(M; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \cong H^{m-k}(M, \partial M; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})^*,$$

for all $0 \leq k \leq m$. See [Johnson and Millson 1987; Porti 1997] for more details.

Lemma 3.2. *For any representation* $\varrho : \mathbb{Z} \oplus \mathbb{Z} \to \mathrm{SL}_n(\mathbb{C})$ *we have*:

$$\dim H^1(\mathbb{Z} \oplus \mathbb{Z}; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\varrho}) \geq 2(n-1).$$

In addition, $\dim H^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = 2(n-1)$ *if and only if* ϱ *is a regular point of* $R(\mathbb{Z} \oplus \mathbb{Z}, \mathrm{SL}_n(\mathbb{C}))$.

Recall that a function $\phi : R(\Gamma, \mathrm{SL}_n(\mathbb{C})) \to \mathbb{Z}$ is called *upper semicontinuous* if for all $k \in \mathbb{Z}$ the set $\phi^{-1}\big([k, \infty)\big)$ is closed. Moreover, it is easy to prove that for $q = 0, 1$ the function $\rho \mapsto \dim H^q(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})$ is upper semicontinuous (see [Heusener and Porti 2011, Lemma 3.2], this is a particular case of the semicontinuity theorem [Hartshorne 1977, Chapter III, Theorem 12.8]).

Proof of Lemma 3.2. Poincaré duality and Euler characteristic give

$$\tfrac{1}{2} \dim H^1(\mathbb{Z} \oplus \mathbb{Z}; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\varrho}) = \dim H^0(\mathbb{Z} \oplus \mathbb{Z}; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\varrho}) = \dim \mathfrak{sl}_n(\mathbb{C})^{\mathbb{Z} \oplus \mathbb{Z}}.$$

By a result of Richardson [1979, Theorem C], every representation of $\mathbb{Z} \oplus \mathbb{Z}$ into $\mathrm{SL}_n(\mathbb{C})$ is a limit of diagonal representations, and for diagonal representations $\dim \mathfrak{sl}_n(\mathbb{C})^{\mathbb{Z} \oplus \mathbb{Z}} \geq n-1$. The general inequality follows from the upper semicontinuity of the function $\varrho \mapsto \dim H^0(\mathbb{Z} \oplus \mathbb{Z}; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\varrho})$.

For the second statement, Richardson proved in the same Theorem C that the representation variety $R(\mathbb{Z} \oplus \mathbb{Z}, \mathrm{SL}_n(\mathbb{C}))$ is an irreducible algebraic variety of dimension $(n+2)(n-1)$. It follows that $\varrho \in R(\mathbb{Z} \oplus \mathbb{Z}, \mathrm{SL}_n(\mathbb{C}))$ is a regular point if and only if $\dim Z^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = (n+2)(n-1)$.

On the other hand,

$$\begin{aligned}
&\dim Z^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = \dim H^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) + \dim B^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C}));\\
&\dim B^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = n^2 - 1 - \dim H^0(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C}));\\
&\dim H^0(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = \tfrac{1}{2} \dim H^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})).
\end{aligned}$$

Hence

$$\dim Z^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) = \tfrac{1}{2} \dim H^1(\mathbb{Z} \oplus \mathbb{Z}, \mathfrak{sl}_n(\mathbb{C})) + n^2 - 1.$$

Thus the lemma follows. (See also [Popov 2008].) □

We will require the following result:

Proposition 3.3 [Heusener and Medjerab 2014, Proposition 3.3]. *Let* α *be a point in the* $\mathrm{SL}_a(\mathbb{C})$*-representation variety* $R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$. *If* α *is infinitesimally regular*, *then it is a regular point of* $R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ *and belongs to a unique component of dimension* $a^2 + a - 2 - \dim H^0(\Gamma; \mathfrak{sl}_a(\mathbb{C}))$.

Remark 3.4. For an irreducible representation $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$, it holds that $H^0(\Gamma; \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha}) = 0$. Indeed, if $X \in \mathfrak{sl}_a(\mathbb{C})$ commutes with $\alpha(\gamma)$ for all $\gamma \in \Gamma$, then Schur's lemma implies that X is a scalar matrix and hence $X = 0$.

As a corollary we obtain from Proposition 3.3 and Remark 3.4:

Corollary 3.5. *If an* irreducible *representation $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ is infinitesimally regular then it is a regular point of $R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ of local dimension $a^2 + a - 2$.*

One has furthermore:

Proposition 3.6. *If an irreducible representation $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ is infinitesimally regular, then its character is a smooth point of $X(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ of local dimension $a - 1$.*

Proof. By Corollary 3.5, α is a regular point of $R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ of local dimension $a^2 + a - 2$. As α is irreducible, the fiber of the projection $R(\Gamma, \mathrm{SL}_a(\mathbb{C})) \to X(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ at α has dimension $a^2 - 1$. The dimension of this fiber is an upper semicontinuous function, therefore the dimension of $X(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ at α is at least $a - 1$. On the other hand, the dimension of the Zariski tangent space of $X(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ at α is at most $\dim H^1(\Gamma; \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha})$ (this follows from Luna's slice as α is irreducible, see [Lubotzky and Magid 1985, Theorem 2.15]). Hence we have equality of dimensions and the proposition follows. □

4. Twisted cohomology and twisted polynomials

In this section we prove that $\alpha \otimes \beta^*$ and $\beta \otimes \alpha^*$ are completely reducible representations, so that the duality theorem (Theorem 2.6) applies to them. Our assumption that $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ and $\beta : \Gamma \to \mathrm{SL}_b(\mathbb{C})$ are irreducible will be crucial for the conclusion.

Decomposition of $\mathfrak{sl}_n(\mathbb{C})$. Consider the action of Γ on the space of matrices with a rows and b columns $M_{a\times b}(\mathbb{C})$:

$$\Gamma \times M_{a\times b}(\mathbb{C}) \to M_{a\times b}(\mathbb{C}), \quad (\gamma, A) \mapsto \lambda^{n\varphi(\gamma)}\alpha(\gamma)A\beta(\gamma^{-1}). \tag{8}$$

The corresponding Γ-module is denoted by

$$\mathcal{M}^+_{\lambda^n} = M_{a\times b}(\mathbb{C})_{\alpha\otimes\beta^*\otimes\lambda^{n\varphi}}.$$

Similarly, we consider the module

$$\mathcal{M}^-_{\lambda^{-n}} = M_{b\times a}(\mathbb{C})_{\beta\otimes\alpha^*\otimes\lambda^{-n\varphi}}.$$

Notice that those modules occur as factors in the decomposition of $\mathfrak{sl}_n(\mathbb{C})$ as Γ-modules via the adjoint action $\mathrm{Ad}\,\rho_\lambda$:

$$\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda} = \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha} \oplus \mathfrak{sl}_b(\mathbb{C})_{\mathrm{Ad}\,\beta} \oplus \mathbb{C} \oplus \mathcal{M}^+_{\lambda^n} \oplus \mathcal{M}^-_{\lambda^{-n}}.$$

This can be visualized as

$$\mathfrak{sl}_n(\mathbb{C})_{\operatorname{Ad}\rho_\lambda} \cong \begin{pmatrix} \mathfrak{sl}_a(\mathbb{C})_{\operatorname{Ad}\alpha} & \mathcal{M}^+_{\lambda^n} \\ \mathcal{M}^-_{\lambda^{-n}} & \mathfrak{sl}_b(\mathbb{C})_{\operatorname{Ad}\beta} \end{pmatrix} \oplus \mathbb{C}\begin{pmatrix} b\operatorname{Id}_a & 0 \\ 0 & -a\operatorname{Id}_b \end{pmatrix}.$$

Duality. For every $\lambda \in \mathbb{C}^*$ we have a nondegenerate bilinear form

$$\Psi : \mathcal{M}^+_{\lambda^n} \times \mathcal{M}^-_{\lambda^{-n}} \to \mathbb{C}, \quad (A, B) \mapsto \operatorname{tr}(AB), \tag{9}$$

which is Γ-invariant: $\Psi(A, B) = \Psi(\gamma A, \gamma B)$ for all $\gamma \in \Gamma$. As an immediate consequence, we have Poincaré and Kronecker dualities:

$$H_i(X; \mathcal{M}^\pm_{\lambda^{\pm n}}) \cong H_{3-i}(X, \partial X; \mathcal{M}^\mp_{\lambda^{\mp n}})^*; \tag{10}$$

$$H^i(X; \mathcal{M}^\pm_{\lambda^{\pm n}}) \cong H^{3-i}(X, \partial X; \mathcal{M}^\mp_{\lambda^{\mp n}})^*; \tag{11}$$

$$H_i(X; \mathcal{M}^\pm_{\lambda^{\pm n}}) \cong H^i(X; \mathcal{M}^\mp_{\lambda^{\mp n}})^*. \tag{12}$$

The i-th twisted Alexander polynomials of the Γ-modules $\mathcal{M}^\mp_1$ are denoted by

$$\Delta^+_i = \Delta^{\alpha\otimes\beta^*}_i \quad \text{and} \quad \Delta^-_i = \Delta^{\beta\otimes\alpha^*}_i.$$

Taking $\rho = \alpha \otimes \beta^*$, then $\rho^* = \beta \otimes \alpha^*$ by (9). In order to apply Theorem 2.6 to those polynomials, we need to show that $\rho = \alpha \otimes \beta^*$ is completely reducible; this motivates the next subsection.

Linear algebraic groups. We follow Humphreys' book [1975] as general reference for linear algebraic groups. A linear algebraic group G contains a unique largest normal solvable subgroup, which is automatically closed. Its identity component is then the largest connected normal solvable subgroup of G; it is called the *radical* of G, denoted by $R(G)$. The subgroup of unipotent elements in $R(G)$ is normal in both $R(G)$ and G; it is called the *unipotent radical* of G, denoted by $R_u(G)$. We have that $R(G)/R_u(G)$ is a torus. Hence $R(G)$ is a torus if and only if $R_u(G)$ is trivial.

Recall that a representation $\rho : \Gamma \to \mathrm{SL}(V)$ is called completely reducible if it is a direct sum of irreducible representations, see Definition 1.1.

Theorem 4.1 [Nagata 1961/1962, Theorem 3]. *Let $G \subset \mathrm{GL}_n(\mathbb{C})$ be an algebraic group. Then $R_u(G)$ is trivial if and only if each rational representation of G is completely reducible.* □

Here a representation $\rho : G \to \mathrm{GL}(V)$ is called *rational* if, with respect to a basis of V, the matrix entries of $\rho(g)$ are polynomial functions in the $n^2 + 1$ coordinate functions x_{ij} $(1 \le i, j \le n)$ and $1/\det$ of $\mathrm{GL}_n(\mathbb{C})$.

Remark 4.2. A nontrivial, connected algebraic group G is called *reductive* if $R_u(G)$ is trivial. Since the Zariski closure of a matrix group is in general not connected we will avoid the term reductive in what follows.

Lemma 4.3. *Let Γ be a group and let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{C})$ be an irreducible representation. Then the unipotent radical $R_u(G)$ of the Zariski closure G of $\rho(\Gamma) \subset \mathrm{SL}_n(\mathbb{C})$ is trivial.*

Proof. Suppose that $R_u(G) \subset \mathrm{SL}_n(\mathbb{C})$ is nontrivial. Every unipotent subgroup of $\mathrm{GL}_n(\mathbb{C})$ has a nonzero vector fixed by all elements of the group (see [Humphreys 1975, 17.5]). Then the subspace $W \subset \mathbb{C}^n$ of fixed vectors of $R_u(G)$ is nonzero. By normality, this subspace is preserved by G, hence by $\rho(\Gamma)$, which contradicts the irreducibility of ρ. □

Lemma 4.4. *Let $\alpha : \Gamma \to \mathrm{SL}_a(\mathbb{C})$ and $\beta : \Gamma \to \mathrm{SL}_b(\mathbb{C})$ be irreducible. Then the unipotent radical $R_u(G)$ of the Zariski closure G of $(\alpha \oplus \beta)(\Gamma) \subset \mathrm{SL}_a(\mathbb{C}) \times \mathrm{SL}_b(\mathbb{C})$ is trivial.*

Proof. Let $p_a : \mathrm{SL}_a(\mathbb{C}) \times \mathrm{SL}_b(\mathbb{C}) \to \mathrm{SL}_a(\mathbb{C})$ denote the projection. Then $p_a((\alpha \oplus \beta)(\Gamma)) = \alpha(\Gamma)$ and therefore $p_a(R_u(G))$ is contained in the unipotent radical $R_u(G_a)$ of the Zariski closure G_a of $\alpha(\Gamma)$ in $\mathrm{SL}_a(\mathbb{C})$. (The image of an unipotent element under a morphism of algebraic groups is unipotent [Humphreys 1975, 15.3].) Now, $R_u(G_a)$ is trivial by Lemma 4.3 and hence $p_a(R_u(G))$ is trivial. It follows in the same way that $p_b(R_u(G))$ is trivial and hence $R_u(G) = \{1\}$. □

Remark 4.5. The same argument of Lemma 4.4 proves that the Zariski closure of a completely reducible linear representation has trivial unipotent radical.

Corollary 4.6. *The Γ-modules $\mathcal{M}^{\pm}_{\lambda^{\pm n}}$ are completely reducible.*

Proof. By Lemma 4.4 the unipotent radical $R_u(G)$ of the Zariski closure G of $(\alpha \oplus \beta)(\Gamma) \subset \mathrm{SL}_a(\mathbb{C}) \times \mathrm{SL}_b(\mathbb{C})$ is trivial. Hence Nagata's theorem (Theorem 4.1) implies that every rational representation of G is completely reducible. In particular, the restriction to G of the rational representation $\mathrm{SL}_a(\mathbb{C}) \times \mathrm{SL}_b(\mathbb{C}) \to \mathrm{GL}(M_{a\times b}(\mathbb{C}))$, given by

$$(A, B) \cdot X = AXB^{-1},$$

for all $(A, B) \in \mathrm{SL}_a(\mathbb{C}) \times \mathrm{SL}_b(\mathbb{C})$ and for all $X \in M_{a\times b}(\mathbb{C})$, is completely reducible.

Since $(\alpha \oplus \beta)(\Gamma)$ is Zariski dense in G, we obtain that $\mathcal{M}^{+}_{1}$ is a completely reducible Γ-module. Finally, the action of $\gamma \in \Gamma$ on $X \in \mathcal{M}^{+}_{\lambda^n}$, given by Equation (8), and the action $\gamma \cdot X = \alpha(\gamma) X \beta(\gamma^{-1})$ differ only by a homothety. Therefore, $\mathcal{M}^{+}_{\lambda^n}$ is a completely reducible Γ-module. The proof for $\mathcal{M}^{-}_{\lambda^{-n}}$ is similar. □

Corollary 4.7. $$\Delta_i^{+}(t) \doteq \Delta_i^{-}(t^{-1}).$$

Proof. The corollary follows directly from Theorem 2.6 and Corollary 4.6. □

5. Necessary condition

The goal of this section is to prove Theorem 1.3. More precisely, we will prove that if the representation $\rho_\lambda = (\lambda^{b\varphi} \otimes \alpha) \oplus (\lambda^{-a\varphi} \otimes \beta)$, as defined in (1), can be deformed to irreducible representations, then $\Delta_1^+(\lambda^n) = 0$. Recall that throughout the paper we assume that α and β are irreducible and infinitesimally regular.

Lemma 5.1. *Assume that ρ_λ belongs to a component of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ that contains irreducible representations. Then*

$$\dim Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda}) \geq n^2 + n - 2.$$

Proof. It is sufficient to prove the inequality for an irreducible representation $\rho \in R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$, because the dimension of $Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})$ is an upper semi-continuous function on ρ and because irreducibility is a Zariski-open condition. We have

$$\dim Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) = \dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) + \dim B^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}).$$

Now, $\dim B^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) = n^2 - 1$ because ρ is irreducible.

Next we apply Poincaré duality to the long exact sequence of the pair $(X, \partial X)$:

$$H^1(X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \to H^1(\partial X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \to H^2(X, \partial X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}). \tag{13}$$

Poincaré duality (7) implies isomorphy between $H^1(X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})$ and the dual space $H^2(X, \partial X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho})^*$. Moreover, the maps of (13) are dual to each other. So:

$$\tfrac{1}{2} \dim H^1(\partial X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}) \leq \dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho}). \tag{14}$$

The claimed inequality of the statement follows from Lemma 3.2. □

Lemma 5.2. *Under the hypothesis of Lemma 5.1 we have*

$$\dim H^1(\Gamma; \mathcal{M}^+_{\lambda^n}) > \dim H^0(\Gamma; \mathcal{M}^+_{\lambda^n})$$

or

$$\dim H^1(\Gamma; \mathcal{M}^-_{\lambda^{-n}}) > \dim H^0(\Gamma; \mathcal{M}^-_{\lambda^{-n}}).$$

We shall see in Remark 5.4 below that we get both inequalities.

Proof. Here we use the decomposition of Γ-modules (see Section 4):

$$\mathfrak{sl}_n(\mathbb{C})_{Ad\rho_\lambda} = \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha} \oplus \mathfrak{sl}_b(\mathbb{C})_{\mathrm{Ad}\,\beta} \oplus \mathbb{C} \oplus \mathcal{M}^+_{\lambda^n} \oplus \mathcal{M}^-_{\lambda^{-n}}. \tag{15}$$

We aim to apply Lemma 5.1, so we compute the dimension of the space of 1-cocycles for each Γ-module in (15). For each Γ-module $\mathfrak{m}$, we use the formula

$$\begin{aligned} \dim Z^1(\Gamma; \mathfrak{m}) &= \dim H^1(\Gamma; \mathfrak{m}) + \dim B^1(\Gamma; \mathfrak{m}) \\ &= \dim H^1(\Gamma; \mathfrak{m}) + \dim \mathfrak{m} - \dim H^0(\Gamma; \mathfrak{m}). \end{aligned} \tag{16}$$

Ordering the terms as they appear in (16):

$$\begin{aligned}\dim Z^1(\Gamma;\mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha}) &= (a-1)+(a^2-1)-0,\\ \dim Z^1(\Gamma;\mathfrak{sl}_b(\mathbb{C})_{\mathrm{Ad}\,\beta}) &= (b-1)+(b^2-1)-0,\\ \dim Z^1(\Gamma;\mathbb{C}) &= 1+1-1,\\ \dim Z^1(\Gamma;\mathcal{M}^{\pm}_{\lambda^{\pm n}}) &= \dim H^1(\Gamma;\mathcal{M}^{\pm}_{\lambda^{\pm n}})+ab-\dim H^0(\Gamma;\mathcal{M}^{\pm}_{\lambda^{\pm n}}).\end{aligned}$$

The first two lines use that α and β are irreducible and infinitesimally regular, the last one that $\dim \mathcal{M}^{\pm}_{\lambda^{\pm n}} = ab$. Adding up the dimensions of the terms in (15) and using Lemma 5.1 and the fact that $a+b=n$, we obtain

$$\begin{aligned}n^2+n-2 \le n^2+n-3+\dim H^1(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})-\dim H^0(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})&\\ +\dim H^1(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}})-\dim H^0(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}})&,\end{aligned}$$

which proves the lemma. □

For later use we remark on the following computation, made during the last proof. Notice that it does not use that ρ_λ can be deformed to irreducible representations (but it uses that α and β are irreducible and infinitesimally regular):

$$\begin{aligned}(17)\quad \dim Z^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda}) = n^2+n-3+\dim H^1(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})-\dim H^0(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})&\\ +\dim H^1(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}})-\dim H^0(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}})&.\end{aligned}$$

Lemma 5.3. *Let* $\rho_\lambda:\Gamma\to \mathrm{SL}_n(\mathbb{C})$ *be given by* $\rho_\lambda = (\lambda^{b\varphi}\otimes\alpha)\oplus(\lambda^{-a\varphi}\otimes\beta)$. *Then* $\dim H^1(\Gamma;\mathcal{M}^{\pm}_{\lambda^{\pm n}}) > \dim H^0(\Gamma;\mathcal{M}^{\pm}_{\lambda^{\pm n}})$ *if and only if* $\Delta^{\mp}_1(\lambda^{\mp n})=0$.

Proof. Recall that $\Delta^{\pm}_i$ is the order of $H_i(X_\infty;\mathcal{M}^{\pm}_1)\cong H_i(X;\mathcal{M}^{\pm}_1[\mathbb{Z}])$. We have a short exact sequence of Γ-modules

$$0\to \mathcal{M}^{-}_1[\mathbb{Z}] \xrightarrow{(t-\lambda^{-n})\cdot} \mathcal{M}^{-}_1[\mathbb{Z}] \to \mathcal{M}^{-}_{\lambda^{-n}} \to 0$$

which gives the following long exact sequence in homology [Brown 1994, III.§6]:

$$\begin{aligned}\ldots\to H_1(\Gamma;\mathcal{M}^{-}_1[\mathbb{Z}]) \xrightarrow{(t-\lambda^{-n})\cdot} H_1(\Gamma;\mathcal{M}^{-}_1[\mathbb{Z}]) \to H_1(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}}) \xrightarrow{\partial}&\\ H_0(\Gamma;\mathcal{M}^{-}_1[\mathbb{Z}]) \xrightarrow{(t-\lambda^{-n})\cdot} H_0(\Gamma;\mathcal{M}^{-}_1[\mathbb{Z}]) \to H_0(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}}) \to 0&.\end{aligned}$$

Thus $\Delta^{-}_1(\lambda^{-n})=0$ if and only if $\ker\partial$ is nontrivial.

Next we claim that $\ker\partial$ is nontrivial if and only if $H^1(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})$ has higher dimension than $H^0(\Gamma;\mathcal{M}^{+}_{\lambda^{n}})$. It follows from Lemma 2.7 and Equation (3) that the $\mathbb{C}[\mathbb{Z}]$-module $H_0(\Gamma;\mathcal{M}^{-}_1[\mathbb{Z}])$ is torsion, i.e., it is a finite dimensional $\mathbb{C}$-vector space. Hence, by exactness, $\operatorname{rank}\partial = \dim H_0(\Gamma;\mathcal{M}^{-}_{\lambda^{-n}})$ and

$$\begin{aligned}\dim\ker\partial &= \dim H_1(\Gamma;\mathcal{M}^-_{\lambda^{-n}}) - \operatorname{rank}\partial\\ &= \dim H_1(\Gamma;\mathcal{M}^-_{\lambda^{-n}}) - \dim H_0(\Gamma;\mathcal{M}^-_{\lambda^{-n}})\\ &= \dim H^1(\Gamma;\mathcal{M}^+_{\lambda^n}) - \dim H^0(\Gamma;\mathcal{M}^+_{\lambda^n}),\end{aligned}$$

by Kronecker duality (12), which proves the claim. Of course the same proof applies by symmetry for the opposite signs ± and ∓. □

Proof of Theorem 1.3. By Lemmas 5.2 and 5.3 we get that if ρ_λ can be deformed to irreducible representations, then $\Delta_1^+(\lambda^n)=0$ or $\Delta_1^-(\lambda^{-n})=0$. Corollary 4.7 yields $\Delta_1^+(\lambda^n)=\Delta_1^-(\lambda^{-n})=0$. □

Remark 5.4. Notice that in the situation of Theorem 1.3, from Lemma 5.3 and Corollary 4.7 we get both inequalities

$$\dim H^1(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}}) > \dim H^0(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}}).$$

We will later need the following construction. Given a 1-cochain $c\in C^1(\Gamma;\mathcal{M}^+_{\lambda^n})$, i.e., a map $c:\Gamma\to\mathcal{M}^+_{\lambda^n}$, consider the map $\rho^c_\lambda:\Gamma\to \mathrm{SL}_n(\mathbb{C})$ given by

$$\rho^c_\lambda(\gamma)=\begin{pmatrix}\mathrm{Id}_a & c(\gamma)\\ 0 & \mathrm{Id}_b\end{pmatrix}\rho_\lambda(\gamma),\quad \gamma\in\Gamma. \tag{18}$$

Lemma 5.5. *The map $\rho^c_\lambda:\Gamma\to\mathrm{SL}_n(\mathbb{C})$ given by* (18) *is a representation if and only if c is a cocycle, i.e., $c\in Z^1(\Gamma;\mathcal{M}^+_{\lambda^n})$. For such c, there is equivalence between these conditions*:

(i) *ρ^c_λ is conjugate to ρ_λ.*

(ii) *c is a coboundary* (*i.e., $c\in B^1(\Gamma;\mathcal{M}^+_{\lambda^n})$*).

(iii) *ρ^c_λ is completely reducible.*

Proof. The equivalence between being a representation and the cocycle condition is a straightforward computation; so is the equivalence between being conjugate to ρ_λ and the coboundary condition. The equivalence with complete reducibility comes from the fact that there is a unique orbit of completely reducible representations in the fiber of the map $R(\Gamma,\mathrm{SL}_n(\mathbb{C}))\to X(\Gamma,\mathrm{SL}_n(\mathbb{C}))$ [Lubotzky and Magid 1985]. Hence two completely reducible representations having the same character are conjugates. □

The following corollary generalizes a result of G. Burde [1967] and G. de Rham [1967]:

Corollary 5.6. *There exists a reducible, not completely reducible representation $\rho^c_\lambda:\Gamma\to\mathrm{SL}_n(\mathbb{C})$ such that $\chi_{\rho^c_\lambda}=\chi_{\rho_\lambda}$ if and only if λ^n is a root of the product of twisted Alexander polynomials $\Delta_1^+(t)\Delta_0^+(t)$.*

Proof. By Lemma 5.5, such a representation exists if and only if $H^1(\Gamma;\mathcal{M}^+_{\lambda^n})$ or $H^1(\Gamma;\mathcal{M}^-_{\lambda^{-n}})$ does not vanish. By Kronecker duality this is equivalent to saying that $H_1(\Gamma;\mathcal{M}^+_{\lambda^n})$ or $H_1(\Gamma;\mathcal{M}^-_{\lambda^{-n}})$ does not vanish. Then, the long exact sequence in the proof of Lemma 5.3 shows that this is equivalent to one of $H_1(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ or $H_0(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ to have $(t-\lambda^{\pm n})$-torsion. With the duality of polynomials, Corollary 4.7, this proves the lemma. □

This corollary also applies when $\alpha=\beta=1$ and $\lambda=\pm1$. Since $\Delta_0(t)=(t-1)$, the vanishing $\Delta_0((\pm1)^2)=0$ corresponds to the representations

$$\gamma\mapsto\pm\begin{pmatrix}1 & d(\gamma)\\ 0 & 1\end{pmatrix},\qquad \gamma\in\Gamma,$$

where $d:\Gamma\to(\mathbb{C},+)$ is any group morphism.

6. Infinitesimal deformations and cup products

Throughout this section and the next one we assume the hypothesis of Theorem 1.4, namely that $\Delta^+_0(\lambda^n)\neq0$ and λ^n is a simple root of Δ^+_1. By Corollary 4.7, we also have that $\Delta^-_0(\lambda^{-n})\neq0$ and λ^{-n} is a simple root of Δ^-_1. Thus the $\mathbb{C}[\mathbb{Z}]$-module $H_0(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ has no $(t-\lambda^{\pm n})$-torsion and $H_1(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ has a single $\mathbb{C}[t^{\pm1}]/(t-\lambda^{\pm n})$-factor. Furthermore, the following proposition gives more details on the cohomology.

Proposition 6.1. *Assume* $\Delta^+_0(\lambda^n)\neq0$ *and* λ^n *is a simple root of* Δ^+_1. *Then*

(i) $\dim H^i(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}})=\begin{cases}1 & \text{if } i=1,2,\\ 0 & \text{otherwise.}\end{cases}$

(ii) *The* $(t-\lambda^{\pm n})$*-torsion of* $H^q(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ *is zero for* $q\neq2$ *and cyclic of the form* $\mathbb{C}[\mathbb{Z}]/(t-\lambda^{\pm n})$ *for* $q=2$.

Proof. In order to prove the first assertion, we use the long exact sequence in the proof of Lemma 5.3. The hypothesis on the twisted Alexander polynomials gives that the $(t-\lambda^{\pm n})$-torsion of $H_i(\Gamma;\mathcal{M}^\pm_1[\mathbb{Z}])$ is zero for $i\neq1$ and $t-\lambda^{\pm n}$ for $i=1$. The long exact sequence gives that $H_i(X;\mathcal{M}^\pm_{\lambda^{\pm n}})$ has dimension 1 if $i=1,2$ and dimension 0 otherwise. Hence the first assertion follows from Kronecker duality, (12).

For the second assertion, we use the universal coefficient theorem for cohomology: for any representation $\rho:\Gamma\to\mathrm{GL}(V)$ we have

$$\begin{aligned}(19)\quad &\overline{H}^q(X;V^*[\mathbb{Z}])\\ &\cong\mathrm{Hom}_{\mathbb{C}[\mathbb{Z}]}\big(H_q(X;V[\mathbb{Z}]),\mathbb{C}[\mathbb{Z}]\big)\oplus\mathrm{Ext}_{\mathbb{C}[\mathbb{Z}]}\big(H_{q-1}(X;V[\mathbb{Z}]),\mathbb{C}[\mathbb{Z}]\big),\end{aligned}$$

where $\overline{H}^q(X;V^*[\mathbb{Z}])$ denotes the group $H^q(X;V^*[\mathbb{Z}])$ with the conjugate $\mathbb{C}[\mathbb{Z}]$-module structure. For a detailed argument see pp. 638–639 in [Kirk and Livingston

1999]. We apply (19) to the representation $\alpha \otimes \beta^*$ and its dual $\beta \otimes \alpha^*$. By the hypothesis on the twisted Alexander polynomials, the $(t-\lambda^{\pm n})$-torsion of $H_i(\Gamma, \mathcal{M}_1^{\pm}[\mathbb{Z}])$ is zero for $i \neq 1$ and $t-\lambda^{\pm n}$ for $i=1$. Notice that $H_q(X; \mathcal{M}_1^{\pm}[\mathbb{Z}])$ are torsion $\mathbb{C}[\mathbb{Z}]$-modules, and the $(t-\lambda^{\pm n})$-torsion of $\overline{H}^2(X; \mathcal{M}_1^{\pm}[\mathbb{Z}])$ is $(t-\lambda^{\mp n})$. The claim follows since $-t\lambda^{\pm n}(t^{-1}-\lambda^{\mp n}) = (t-\lambda^{\pm n})$. □

From now on we fix cocycles

$$d_{\pm} \in Z^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$$

whose cohomology classes do not vanish. Because $H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) \cong \mathbb{C}$, the elements $d_{\pm}$ are unique up to adding a coboundary and up to multiplying by a nonzero scalar. Our next goal is to show that the cohomology class of the cup product $\varphi \smile d_{\pm}$ does not vanish in $H^2(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$. For that purpose we shall use the dual numbers.

Dual numbers. The *algebra of dual numbers* is defined to be

$$\mathbb{C}_{\varepsilon} = \mathbb{C}[\varepsilon]/\varepsilon^2.$$

Similarly define $\mathbb{C}_{\varepsilon}[\mathbb{Z}] = \mathbb{C}[\mathbb{Z}] \otimes_{\mathbb{C}} \mathbb{C}_{\varepsilon}$ and

$$\mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)} = \big(\mathcal{M}_1^{\pm}[\mathbb{Z}] \otimes_{\mathbb{C}} \mathbb{C}_{\varepsilon}\big)/\big(t-\lambda^{\pm n}(1\pm\varepsilon)\big).$$

Lemma 6.2. *If* $\lambda^n \in \mathbb{C}^*$ *is a simple root of* Δ_1^+ *such that* $\Delta_0^+(\lambda^n) \neq 0$, *then* $\dim H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)}) = 1$.

Proof. Notice that $\Delta_0^-(\lambda^{-n}) \neq 0$ and λ^{-n} is a simple root of Δ_1^- by Corollary 4.7. We have that $H^1(\Gamma; \mathcal{M}_1^+[\mathbb{Z}] \otimes \mathbb{C}_{\varepsilon}) \cong H^1(\Gamma; \mathcal{M}_1^+[\mathbb{Z}]) \otimes \mathbb{C}_{\varepsilon}$ since $\mathbb{C}_{\varepsilon}$ is a trivial Γ-module (isomorphic to $\mathbb{C}^2$). As before, the short exact sequence

$$0 \to \mathcal{M}_1^+[\mathbb{Z}] \otimes_{\mathbb{C}} \mathbb{C}_{\varepsilon} \xrightarrow{(t-\lambda^n(1+\varepsilon))\cdot} \mathcal{M}_1^+[\mathbb{Z}] \otimes_{\mathbb{C}} \mathbb{C}_{\varepsilon} \to \mathcal{M}^+_{\lambda^n(1+\varepsilon)} \to 0$$

gives a long exact sequence in cohomology (see [Brown 1994, III.§6]),

$$\cdots \to H^i(\Gamma; \mathcal{M}_1^+[\mathbb{Z}]) \otimes \mathbb{C}_{\varepsilon} \xrightarrow{(t-(\lambda^n(1+\varepsilon))\cdot} H^i(\Gamma; \mathcal{M}_1^+[\mathbb{Z}]) \otimes \mathbb{C}_{\varepsilon}$$
$$\to H^i(\Gamma; \mathcal{M}^+_{\lambda^n(1+\varepsilon)}) \to H^{i+1}(\Gamma; \mathcal{M}_1^+[\mathbb{Z}]) \otimes \mathbb{C}_{\varepsilon} \to \cdots.$$

Note that for $\mu \neq \lambda^n$, $\mu \in \mathbb{C}$, multiplication by $t-\lambda^n(1+\varepsilon)$ induces an automorphism of $\mathbb{C}_{\varepsilon}[\mathbb{Z}]/(t-\mu)^k$. Therefore, we are interested in the $(t-\lambda^n)$-torsion of $H^q(\Gamma; \mathcal{M}_1^+[\mathbb{Z}])$ described by Proposition 6.1: it vanishes for $q \neq 2$ and it is $\mathbb{C}[\mathbb{Z}]/(t-\lambda^n)$ for $q=2$. Hence, multiplication by $(t-\lambda^n(1+\varepsilon))$ on $H^i(\Gamma; \mathcal{M}_1^+[\mathbb{Z}] \otimes \mathbb{C}_{\varepsilon})$ is an isomorphism for $i \neq 2$. In order to understand the effect of the multiplication on $H^2(\Gamma; \mathcal{M}_1^+[\mathbb{Z}] \otimes \mathbb{C}_{\varepsilon})$ it is sufficient to consider multiplication by $(t-\lambda^n(1+\varepsilon))$ on

$$\mathbb{C}[\mathbb{Z}]/(t-\lambda^n) \otimes \mathbb{C}_{\varepsilon} \cong \mathbb{C}[\mathbb{Z}]/(t-\lambda^n) \oplus \varepsilon\mathbb{C}[\mathbb{Z}]/(t-\lambda^n).$$

Since $t - \lambda^n$ vanishes in this ring, multiplication by $(t - \lambda^n(1+\varepsilon))$ is equivalent to multiplication by $-\varepsilon\lambda^n$ on $\mathbb{C}_\varepsilon \cong \mathbb{C} \oplus \varepsilon\mathbb{C}$. Therefore, its kernel and cokernel have $\mathbb{C}$-dimension 1, which proves $\dim_{\mathbb{C}} H^1(\Gamma; \mathcal{M}^+_{\lambda^n(1+\varepsilon)}) = 1$.

By symmetry the same argument yields $\dim_{\mathbb{C}} H^1(\Gamma; \mathcal{M}^-_{\lambda^{-n}(1-\varepsilon)}) = 1$. □

Cup product and Bockstein homomorphism. Let A_1, A_2 and A_3 be Γ-modules. The cup product of two cochains $c_i \in C^1(\Gamma; A_i)$, $i = 1, 2$ is the cochain $c_1 \smile c_2 \in C^2(\Gamma; A_1 \otimes A_2)$ defined by

$$c_1 \smile c_2(\gamma_1, \gamma_2) := c_1(\gamma_1) \otimes \gamma_1 c_2(\gamma_2). \tag{20}$$

Here $A_1 \otimes A_2$ is a Γ-module via the diagonal action.

It is possible to combine the cup product with any Γ-invariant, bilinear map $b : A_1 \otimes A_2 \to A_3$. So we obtain a cup product

$$_b\smile : C^1(\Gamma; A_1) \otimes C^1(\Gamma; A_2) \xrightarrow{\smile} C^1(\Gamma; A_1 \otimes A_2) \xrightarrow{b} C^2(\Gamma; A_3).$$

For details see [Brown 1994, V.3]. In what follows we are mainly interested in the case where the bilinear form is simply the matrix multiplication, i.e.,

$$\mathbb{C} \otimes \mathcal{M}^{\pm}_{\lambda^{\pm n}} \to \mathcal{M}^{\pm}_{\lambda^{\pm n}} \quad \text{or} \quad \mathfrak{sl}_a(\mathbb{C}) \otimes \mathcal{M}^+_{\lambda^n} \to \mathcal{M}^+_{\lambda^n}.$$

Hence we will write simply "$\smile$" for such a cup product when no confusion can arise.

Let $b : A_1 \otimes A_2 \to A_3$ be bilinear and let $\tau : A_2 \otimes A_1 \to A_1 \otimes A_2$ be the twist operator. Then for $c_i \in C^1(\Gamma; A_i)$, $i = 1, 2$, we define the cup product

$$_{b\circ\tau}\smile : C^1(\Gamma; A_2) \otimes C^1(\Gamma; A_1) \to C^2(\Gamma; A_3).$$

Again we are mainly interested in matrix multiplication and we will write simply "$_\tau\smile$" for such a cup product when no confusion can arise.

Example 6.3. Let $c_a \in C^1(\Gamma; \mathfrak{sl}_a(\mathbb{C}))$ and $d \in C^1(\Gamma; \mathcal{M}^+_{\lambda^n})$ be given. Then

$$c_a \smile d(\gamma_1, \gamma_2) = c_a(\gamma_1)\gamma_1\, d(\gamma_2) = \lambda^{n\varphi(\gamma_1)} c_a(\gamma_1)\alpha(\gamma_1)\, d(\gamma_2)\beta(\gamma_1)^{-1}$$

and

$$d \,{}_\tau\!\smile c_a(\gamma_1, \gamma_2) = \gamma_1 c_a(\gamma_2)\, d(\gamma_1) = \alpha(\gamma_1) c_a(\gamma_2)\alpha(\gamma_1)^{-1}\, d(\gamma_1).$$

Remark 6.4. If $z_a \in Z^1(\Gamma; \mathfrak{sl}_a(\mathbb{C}))$ and $d_+ \in Z^1(\Gamma; \mathcal{M}^+_{\lambda^n})$ are cocycles, then for $f : \Gamma \to \mathcal{M}^+_{\lambda^n}$ given by $f(\gamma) = z_a(\gamma) d_+(\gamma)$ we have

$$\delta f(\gamma_1, \gamma_2) + z_a \smile d_+(\gamma_1, \gamma_2) + d_+ \,{}_\tau\!\smile z_a(\gamma_1, \gamma_2) = 0,$$

i.e., $d_+ \,{}_\tau\!\smile z_a \sim -z_a \smile d$ in $C^2(\Gamma; \mathcal{M}^+_{\lambda^n})$.

Lemma 6.5. *Consider the nonsplit exact sequence of* Γ*-modules*

$$0 \to \mathcal{M}^{\pm}_{\lambda^{\pm n}} \xrightarrow{\varepsilon\cdot} \mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)} \longrightarrow \mathcal{M}^{\pm}_{\lambda^{\pm n}} \to 0.$$

Then the image of the cohomology class represented by $d_\pm$ (*in* $H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$) *under the Bockstein homomorphism is represented by the cup product* $d_\pm \smile \varphi$ (*in* $H^2(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$).

Proof. In order to calculate the Bockstein homomorphism $\boldsymbol{b} : H^1(\Gamma; \mathcal{M}^{+}_{\lambda^n}) \to H^2(\Gamma; \mathcal{M}^{+}_{\lambda^n})$ we proceed as follows (according to the snake lemma): given a cocycle $d_+ \in Z^1(\Gamma; \mathcal{M}^{+}_{\lambda^n})$ we choose a cochain $\tilde{d}_+ \in C^1(\Gamma; \mathcal{M}^{+}_{\lambda^n(1+\varepsilon)})$ which projects onto d_+ and then we calculate $\delta_\varepsilon \tilde{d}_+ \in C^2(\Gamma; \mathcal{M}^{+}_{\lambda^n(1+\varepsilon)})$ where δ_ε denotes the coboundary operator of $C^*(\Gamma; \mathcal{M}^{+}_{\lambda^n(1+\varepsilon)})$. Since d_+ is a cocycle we obtain $\delta_\varepsilon \tilde{d}_+ = \varepsilon \cdot z$ for a 2-cocycle $z \in Z^2(\Gamma; \mathcal{M}^{+}_{\lambda^n})$ which represents the image of the Bockstein map. By abusing notation, we also denote the map constructed in this way by $\boldsymbol{b} : Z^1(\Gamma; \mathcal{M}^{+}_{\lambda^n}) \to Z^2(\Gamma; \mathcal{M}^{+}_{\lambda^n})$, even if it is only well defined in cohomology. In particular $\boldsymbol{b}(d_+) \sim z$. In order to calculate $z \in Z^2(\Gamma; \mathcal{M}^{+}_{\lambda^n})$ we choose $\tilde{d}_+ = d_+ + \varepsilon \cdot 0$:

$$\begin{aligned}\delta_\varepsilon \tilde{d}_+(\gamma_1, \gamma_2) &= \gamma_1 \tilde{d}_+(\gamma_2) - \tilde{d}_+(\gamma_1\gamma_2) + \tilde{d}_+(\gamma_1)\\ &= \lambda^{n\varphi(\gamma_1)}(1+\varepsilon\varphi(\gamma_1))\alpha(\gamma_1)d_+(\gamma_2)\beta(\gamma_1)^{-1} - d_+(\gamma_1\gamma_2) + d_+(\gamma_1)\\ &= \varepsilon\varphi(\gamma_1)\gamma_1 d_+(\gamma_2) = \varepsilon \cdot \varphi \smile d_+(\gamma_1, \gamma_2).\end{aligned}$$

Therefore, $\boldsymbol{b}(d_+) \sim \varphi \smile d_+$. The calculation for $\boldsymbol{b}(d_-) \sim \varphi \smile d_-$ is similar. □

Corollary 6.6. *Assume* $\dim_{\mathbb{C}} H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) = 1$, $H^0(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) = 0$ *and let* $d_\pm \in Z^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$ *be not cohomologous to zero. Then* $\dim H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)}) = 1$ *if and only if the cup product* $\varphi \smile d_\pm$ *does not vanish in* $H^2(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$.

Proof. Consider the nonsplit exact sequence of Γ-modules

$$0 \to \mathcal{M}^{\pm}_{\lambda^{\pm n}} \xrightarrow{\varepsilon\cdot} \mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)} \to \mathcal{M}^{\pm}_{\lambda^{\pm n}} \to 0$$

and the corresponding long exact sequence in cohomology:

$$0 \to H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) \xrightarrow{\varepsilon\cdot} H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}(1\pm\varepsilon)}) \to H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) \xrightarrow{\boldsymbol{b}} H^2(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}). \tag{21}$$

The lemma then follows from the sequence (21), as by Lemma 6.5 $\boldsymbol{b}(d_\pm) \sim \varphi \smile d_\pm$ and $\dim H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) = 1$. □

Combining Proposition 6.1, Lemma 6.2, and Corollary 6.6, we deduce:

Corollary 6.7. *Under the hypothesis of Theorem 1.4, the cup product* $\varphi \smile d_\pm$ *does not vanish in* $H^2(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}})$.

7. A not completely reducible representation ρ^+

In this section we construct a $\rho^+ \in R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ that has the same character as ρ_λ but is not completely reducible. We show that ρ^+ is a smooth point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ and that it can be deformed to irreducible representations. This proves Theorem 1.4, because the orbit by conjugation of ρ^+ accumulates to ρ_λ.

Assume throughout this section that the hypotheses of Theorem 1.4 hold true. Namely (using Corollary 4.7), $\Delta_0^{\pm n}(\lambda^{\pm n}) \neq 0$ and $\lambda^{\pm n}$ is a simple root of $\Delta_1^{\pm}$. Recall we have fixed $d_+ \in Z^1(\Gamma; \mathcal{M}^+_{\lambda^n})$, a cocycle not homologous to zero. Let

$$\rho^+ = \begin{pmatrix} \mathrm{Id}_a & d_+ \\ 0 & \mathrm{Id}_b \end{pmatrix} \rho_\lambda.$$

By Lemma 5.5, ρ^+ is not completely reducible, hence it is not conjugate to ρ_λ, even if it has the same character. We shall prove that ρ^+ is a regular point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ and that the local dimension is $\dim \mathrm{SL}_n(\mathbb{C}) + n - 1 = n^2 + n - 2$. Then we will argue that the reducible representations around ρ_λ form a Zariski closed algebraic set of dimension $n^2 + n - 3$, which will prove Theorem 1.4.

Let $P^+ = \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \subset \mathrm{SL}_n(\mathbb{C})$ be the maximal parabolic subgroup that preserves $\mathbb{C}^a \oplus 0$. Its Lie algebra is denoted by $\mathfrak{p}^+ \subset \mathfrak{sl}_n(\mathbb{C})$. We have two short exact sequences of Γ-modules via the action of $\mathrm{Ad}\,\rho^+$:

$$0 \to \mathcal{M}^+_{\lambda^n} \to \mathfrak{p}^+ \to \mathcal{D} \to 0, \tag{22}$$

where

$$\mathcal{D} = \mathfrak{sl}_a(\mathbb{C}) \oplus \mathfrak{sl}_b(\mathbb{C}) \oplus \mathbb{C}, \tag{23}$$

and

$$0 \to \mathfrak{p}^+ \to \mathfrak{sl}_n(\mathbb{C}) \to \mathcal{M}^-_{\lambda^{-n}} \to 0. \tag{24}$$

We will use the corresponding long exact sequences in cohomology to compute $H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+})$. The first step is the following lemma.

Lemma 7.1. $H^0(\Gamma; \mathfrak{p}^+) = 0$.

Proof. The long exact sequence associated to (22) starts with

$$0 = H^0(\Gamma; \mathcal{M}^+_{\lambda^n}) \to H^0(\Gamma; \mathfrak{p}^+) \to H^0(\Gamma; \mathcal{D}) \xrightarrow{\boldsymbol{b}} H^1(\Gamma; \mathcal{M}^+_{\lambda^n}).$$

The group $H^0(\Gamma; \mathcal{D}) \cong \mathbb{C}$ is generated by the invariant element $\left(\begin{smallmatrix} -b\,\mathrm{Id}_a & 0 \\ 0 & a\,\mathrm{Id}_b \end{smallmatrix}\right) \in \mathcal{D}^\Gamma$. A similar calculation as in the proof of Lemma 6.5 using the snake lemma gives $\boldsymbol{b}\left(\begin{smallmatrix} -b\,\mathrm{Id}_a & 0 \\ 0 & a\,\mathrm{Id}_b \end{smallmatrix}\right) \sim n d_+$. Therefore $\boldsymbol{b} : H^0(\Gamma; \mathcal{D}) \to H^1(\Gamma; \mathcal{M}^+_{\lambda^n})$ is injective and hence $H^0(\Gamma; \mathfrak{p}^+) = 0$. □

We continue the long exact sequence in cohomology associated to (22):

$$0 \to \mathbb{C} \to H^1(\Gamma;\mathcal{M}^+_{\lambda^n}) \to H^1(\Gamma;\mathfrak{p}^+) \to H^1(\Gamma;\mathcal{D}) \xrightarrow{\boldsymbol{b}} H^2(\Gamma;\mathcal{M}^+_{\lambda^n}).$$

Since $H^i(\Gamma;\mathcal{M}^+_{\lambda^n}) \cong \mathbb{C}$ for $i = 1, 2$ by Proposition 6.1, it shortens to

$$0 \to H^1(\Gamma;\mathfrak{p}^+) \to H^1(\Gamma;\mathcal{D}) \xrightarrow{\boldsymbol{b}} H^2(\Gamma;\mathcal{M}^+_{\lambda^n}).$$

Next we aim to compute $\boldsymbol{b} : H^1(\Gamma;\mathcal{D}) \to H^2(\Gamma;\mathcal{M}^+_{\lambda^n})$. For this we use the decomposition (23). Every element in $H^1(\Gamma;\mathcal{D})$ is represented by a cocycle

$$\vartheta = \begin{pmatrix} z_a & 0 \\ 0 & z_b \end{pmatrix} + z\varphi \begin{pmatrix} -b\,\mathrm{Id}_a & 0 \\ 0 & a\,\mathrm{Id}_b \end{pmatrix} \tag{25}$$

where $z_a \in Z^1(\Gamma;\mathfrak{sl}_a(\mathbb{C}))$, $z_b \in Z^1(\Gamma;\mathfrak{sl}_b(\mathbb{C}))$, and $z \in \mathbb{C}$.

Lemma 7.2. *For a cocycle $\vartheta \in Z^1(\Gamma;\mathcal{D})$ as in* (25),

$$\boldsymbol{b}(\vartheta) \sim z_a \smile d_+ + d_+ \smile z_b + znd_+ \smile \varphi.$$

Proof. As in Lemma 6.5 we compute $\boldsymbol{b}(\vartheta)$ by using the snake lemma. Namely, let δ^+ be the coboundary operator of $C^*(\Gamma;\mathfrak{p}^+)$, and let $\tilde{\vartheta} \in C^1(\Gamma;\mathfrak{p}^+)$ be the composition of ϑ with the inclusion $\mathcal{D} \hookrightarrow \mathfrak{p}^+$. Then

$$\delta^+\tilde{\vartheta}(\gamma_1,\gamma_2) = \begin{pmatrix} 0 & -\gamma_1 z_a(\gamma_2) d_+(\gamma_1) + d_+(\gamma_1)\gamma_1 z_b(\gamma_2) + znd_+(\gamma_1)\varphi(\gamma_2) \\ 0 & 0 \end{pmatrix}$$

and hence $\boldsymbol{b}(\vartheta) \sim -d_+ {}_\tau\!\smile z_a + d_+ \smile z_b + znd_+ \smile \varphi$.

Finally, Remark 6.4 proves the lemma. □

Since $\varphi \smile d_\pm$ is not cohomologous to zero (Corollary 6.7) and $H^2(\Gamma;\mathcal{M}^+_{\lambda^n}) \cong \mathbb{C}$ (Proposition 6.1), we deduce:

Corollary 7.3. *The cohomology group $H^1(\Gamma;\mathfrak{p}^+) \cong \mathbb{C}^{n-2}$ is naturally identified to the kernel of the rank one map*:

$$H^1(\Gamma;\mathcal{D}) \cong H^1(\Gamma;\mathfrak{sl}_a(\mathbb{C})) \oplus H^1(\Gamma;\mathfrak{sl}_b(\mathbb{C})) \oplus \mathbb{C} \xrightarrow{\boldsymbol{b}} H^2(\Gamma;\mathcal{M}^+_{\lambda^n}) \cong \mathbb{C}.$$

Next we consider the long exact sequence corresponding to (24):

$$0 \to H^1(\Gamma;\mathfrak{p}^+) \to H^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) \to H^1(\Gamma;\mathcal{M}^-_{\lambda^{-n}}). \tag{26}$$

Hence

$$\begin{aligned} \dim H^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) &\le \dim H^1(\Gamma;\mathfrak{p}^+) + \dim H^1(\Gamma;\mathcal{M}^-_{\lambda^{-n}}) \\ &= n - 2 + 1 = n - 1. \end{aligned} \tag{27}$$

On the other hand we apply Poincaré duality to the long exact sequence of the pair $(X, \partial X)$ (see (13)) and we obtain as in Equation (14):

$$\dim H^1(\partial X; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) \leq 2 \dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) \leq 2(n-1). \tag{28}$$

Proposition 7.4. *ρ^+ is a regular point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ of dimension n^2+n-2.*

Proof. The dimension inequality of Lemma 3.2 and the inequality (28) yield $\dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) = n-1$, and we apply Proposition 3.3. ☐

Before proving that the irreducible component of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ containing ρ^+ also contains irreducible representations, we need a remark and two lemmas.

Remark 7.5. It follows from the proof of Proposition 7.4 that inequalities (27) and (28) are equalities, therefore (26) becomes a short exact sequence:

$$0 \to H^1(\Gamma; \mathfrak{p}^+) \to H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}) \to H^1(\Gamma; \mathcal{M}^-_{\lambda^{-n}}) \to 0.$$

Lemma 7.6. *The representation ρ^+ is a smooth point of $R(\Gamma, P^+)$.*

Proof. The key tool here is the vanishing of Goldman's obstructions [1984] to integrability, which relies on the naturality of these obstructions and the vanishing for $\mathfrak{sl}_n(\mathbb{C})$. (In our proof of Proposition 7.4 this vanishing is also used implicitly, since our Proposition 3.3 is taken from [Heusener and Medjerab 2014], where the vanishing is invoked.)

By Remark 7.5, the long exact sequence in cohomology associated to (24) yields an injection

$$0 \to H^2(\Gamma; \mathfrak{p}^+) \to H^2(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+}).$$

Now Goldman's obstructions to integrability are natural for the inclusion $\mathfrak{p}^+ \to \mathfrak{sl}_n(\mathbb{C})$. In addition, the obstructions of a cocycle in $\mathfrak{p}^+$ remain in $\mathfrak{p}^+$, because $\mathfrak{p}^+ \to \mathfrak{sl}_n(\mathbb{C})$ is a subalgebra (closed under the Lie bracket) and a Γ-submodule of $\mathfrak{sl}_n(\mathbb{C})$. Since ρ^+ is a smooth point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$, for any cocycle in $Z^1(\Gamma; \mathfrak{p}^+)$ the infinite sequence of obstructions to integrability in $H^2(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho^+})$ vanish, so the infinite sequence of obstructions to integrability in $H^2(\Gamma; \mathfrak{p}^+)$ also vanish. This establishes that any infinitesimal deformation is formally integrable and it follows from Artin's theorem [1968] that it is actually integrable, which proves the lemma. ☐

Let $1 \leq k \leq n$ and let $R_k \subset R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ denote the subset of representations ρ such that $\rho(\Gamma)$ preserves a k-dimensional subspace of $\mathbb{C}^n$.

Lemma 7.7. *For all $1 \leq k \leq n$, the subset $R_k \subset R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ is Zariski-closed.*

Proof. The assertion is clear for $k = n$ since $R_n = R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$. Hence suppose that $1 \leq k < n$ and let $P(k) \subset \mathrm{SL}_n(\mathbb{C})$ denote the parabolic subgroup which preserves

$\mathbb{C}^k \times \{0\} \subset \mathbb{C}^n$. The set $R(\Gamma, P(k)) \subset R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ is Zariski-closed since it is given by a finite number of equations. Moreover, we have

$$R_k = \mathrm{SL}_n(\mathbb{C}) \cdot R(\Gamma, P(k)) = \mathrm{SL}_n(\mathbb{C})/P(k) \cdot R(\Gamma, P(k))$$

since $P(k)$ preserves $R(\Gamma, P(k))$. Finally, R_k is Zariski-closed since the quotient $\mathrm{SL}_n(\mathbb{C})/P(k)$ is complete (see [Humphreys 1995, §0.15]). □

Lemma 7.8. *The unique proper invariant subspace of* $\rho^+(\Gamma)$ *is* $\mathbb{C}^a \times \{0\}$.

Proof. We compute the possible nonzero invariant subspaces of $\rho^+(\Gamma)$ by taking a nonzero vector $v \in \mathbb{C}^n$ and considering the linear span of its orbit $\langle \rho^+(\Gamma)v \rangle$. When $v \in \mathbb{C}^a \times \{0\}$, then $\langle \rho^+(\Gamma)v \rangle = \mathbb{C}^a \times \{0\}$ because α is irreducible. So we assume that the projection of v to the quotient $\mathbb{C}^n/\mathbb{C}^a \times \{0\}$ does not vanish, and since β is irreducible, the projection of the linear span $\langle \rho^+(\Gamma)v \rangle$ is the whole $\mathbb{C}^n/\mathbb{C}^a \times \{0\}$. In particular the dimension of $\langle \rho^+(\Gamma)v \rangle$ is at least b. Notice that $\dim_{\mathbb{C}} \langle \rho^+(\Gamma)v \rangle = b$ cannot occur, because this would yield a direct sum $\mathbb{C}^n = \mathbb{C}^a \times \{0\} \oplus \langle \rho^+(\Gamma)v \rangle$; by Lemma 5.5 this would contradict the choice of ρ^+ and the nontriviality of the cohomology class of d_+. Therefore $\dim_{\mathbb{C}} \langle \rho^+(\Gamma)v \rangle > b$, so that $\langle \rho^+(\Gamma)v \rangle$ contains at least a nontrivial vector in $\mathbb{C}^a \times \{0\}$ (the kernel of the projection). Irreducibility of α gives now $\langle \rho^+(\Gamma)v \rangle = \mathbb{C}^n$. □

Let S be the component of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ that contains ρ^+. In particular, $\dim S = n^2 + n - 2$.

Proposition 7.9. *S contains irreducible representations.*

Proof. We prove the proposition by contradiction, hence assume that there is a Zariski neighborhood $U \subset S \subset R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ of ρ^+ so that all representations in U are reducible. By Lemmas 7.7 and 7.8, the choice of the U can be made so that the representations in U have only an a-dimensional invariant subspace.

In particular every representation in U is conjugate to a representation in $P^+ = P(a)$. Therefore given any Zariski neighborhood $U^+ \subset R(\Gamma, P^+)$ of ρ^+, U can be chosen so that every representation in U is conjugate to a representation in U^+. As ρ^+ is a smooth point of $R(\Gamma, P^+)$ by Lemma 7.6, ρ^+ is contained in a single irreducible component S^+ of $R(\Gamma, P^+)$, and we may chose $U^+ \subset S^+$. This yields the inclusion

$$U \subset \mathrm{SL}_n(\mathbb{C}) \cdot U^+ \subset \mathrm{SL}_n(\mathbb{C}) \cdot S^+.$$

Now we reach the contradiction by computing dimensions. Using that P^+ stabilizes S^+ we get

$$\dim U \leq \dim(\mathrm{SL}_n(\mathbb{C}) \cdot S^+) \leq \dim(\mathrm{SL}_n(\mathbb{C})/P^+) + \dim S^+,$$

where $\dim(\mathrm{SL}_n(\mathbb{C})/P^+) = n^2 - 1 - \dim \mathfrak{p}^+$, and

$$\dim S^+ = \dim H^1(\Gamma; \mathfrak{p}^+) + \dim \mathfrak{p}^+ - \dim H^0(\Gamma; \mathfrak{p}^+) = n - 2 + \dim \mathfrak{p}^+ - 0.$$

This yields $\dim U \leq n^2+n-3$, contradicting Proposition 7.4, which asserts that $\dim U = \dim S = n^2+n-2$. □

8. The neighborhood of χ_λ

The aim of this section is to prove Theorem 1.5, i.e., we determine the local structure of the character variety $X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ at χ_λ, the character of the representation ρ_λ given by (1). For this purpose we will identify the quadratic cone of $X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ at χ_λ by means of algebraic obstructions to integrability. Moreover, we will describe these obstructions geometrically.

Before discussing the components of the variety of characters, we need to discuss the components of the variety of representations. In Section 7 we have constructed S a component of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ of dimension n^2+n-2 that contains ρ^+ and irreducible representations (Propositions 7.4 and 7.9).

Next we discuss a component of reducible representations. The representation variety $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ contains

$$R(\Gamma, \mathrm{SL}_a(\mathbb{C})) \times R(\Gamma, \mathrm{SL}_b(\mathbb{C})) \times R(\Gamma, \mathbb{C}^*)$$

where the inclusion is given by

$$(\alpha', \beta', \lambda') \mapsto ((\lambda')^{b\varphi} \otimes \alpha') \oplus ((\lambda')^{-a\varphi}) \otimes \beta').$$

Our hypothesis on infinitesimal regularity implies that $\alpha \in R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ and $\beta \in R(\Gamma, \mathrm{SL}_b(\mathbb{C}))$ are smooth points which are contained in unique components $V_\alpha \subset R(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ and $V_\beta \subset R(\Gamma, \mathrm{SL}_b(\mathbb{C}))$ respectively. Hence we obtain an embedding

$$V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*) \hookrightarrow R(\Gamma, \mathrm{SL}_n(\mathbb{C})).$$

Lemma 8.1. *There exists a unique component T of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ that contains*

$$V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*).$$

Moreover, we have $\dim T = n^2+n-3$.

Proof. By the hypothesis of Theorem 1.5 we have $\Delta_0^{\alpha\otimes\beta^*}(\lambda^n) \neq 0$ and λ^n is a simple root of $\Delta_1^{\alpha\otimes\beta^*}(t)$. Hence for all $\lambda' \neq \lambda$ which are sufficiently close to λ we have $\Delta_q^{\alpha\otimes\beta^*}((\lambda')^n) \neq 0$ for $q = 0, 1$. Hence, by the argument in the proof of Proposition 6.1 we obtain $H^q(\Gamma; \mathcal{M}^{\pm}_{(\lambda')^{\pm n}}) = 0$ for $q = 0, 1$.

Now consider the representation

$$\rho_{\lambda'} = ((\lambda')^{b\varphi} \otimes \alpha) \oplus ((\lambda')^{-a\varphi} \otimes \beta) \in V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)$$

and the corresponding decomposition of $\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_{\lambda'}}$ as Γ-module:

$$\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_{\lambda'}} = \mathfrak{sl}_a(\mathbb{C})_{\mathrm{Ad}\,\alpha} \oplus \mathfrak{sl}_b(\mathbb{C})_{\mathrm{Ad}\,\beta} \oplus \mathbb{C} \oplus \mathcal{M}^+_{(\lambda')^n} \oplus \mathcal{M}^-_{(\lambda')^{-n}}.$$

Hence

$$\begin{aligned}\dim Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_{\lambda'}}) &= \dim H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_{\lambda'}}) + \dim B^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_{\lambda'}}) \\ &= a-1+b-1+1+n^2-1-1 = n^2+n-3.\end{aligned}$$

On the other hand, for the $\mathrm{SL}_n(\mathbb{C})$-orbit of $V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)$ we have:

$$\mathrm{SL}_n(\mathbb{C}) \cdot (V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)) = \mathrm{SL}_n(\mathbb{C})/P^+ \cdot (U^+ \cdot V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*))$$

where $U^+ = \left\{ \left(\begin{smallmatrix} \mathrm{id}_a & X \\ 0 & \mathrm{id}_b \end{smallmatrix} \right) \,\middle|\, X \in \mathcal{M}^+_{(\lambda')^n} \right\}$. Now the action of U^+ on $V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)$ is generically free since $H^0(\Gamma; \mathcal{M}^+_{(\lambda')^n}) = 0$ and hence

$$\begin{aligned}\dim \mathrm{SL}_n(\mathbb{C}) \cdot (V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)) &\geq ab + ab + a^2 + a - 2 + b^2 + b - 2 + 1 \\ &= n^2 + n - 3.\end{aligned}$$

Therefore, $\rho_{\lambda'}$ is a smooth point of $R(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ which is contained in a unique n^2+n-3-dimensional component T. Note that T is the Zariski closure of the orbit $\mathrm{SL}_n(\mathbb{C}) \cdot (V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*))$. □

Let Y and Z denote the components of the character variety that contain the characters of S and T respectively. We have $\dim Y = \dim S - \dim \mathrm{SL}_n(\mathbb{C}) = n-1$. In addition $\dim Z \geq a-1+b-1+1 = n-1$ since T contains $V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)$. Notice that the generic dimension of the orbit of $(\alpha', \beta', \lambda') \in V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*)$ is n^2-2. Hence, $\dim Z \leq \dim T - (n^2-2) = n-1$. Hence $\dim Z = n-1$ and $\dim T = n^2+n-3$.

Let $Z_\alpha \subset X(\Gamma, \mathrm{SL}_a(\mathbb{C}))$ and $Z_\beta \subset X(\Gamma, \mathrm{SL}_b(\mathbb{C}))$ denote the irreducible components that contain the respective projections of V_α and V_β. We have a commutative diagram

$$\begin{tikzcd} V_\alpha \times V_\beta \times R(\Gamma, \mathbb{C}^*) \arrow[r] \arrow[d] & T \subset R(\Gamma, \mathrm{SL}_n(\mathbb{C})) \arrow[d] \\ Z_\alpha \times Z_\beta \times \mathbb{C}^* \arrow[r] & Z \subset X(\Gamma, \mathrm{SL}_n(\mathbb{C})) \end{tikzcd}$$

The top row is injective but not the bottom one, as conjugation can realize permutations of rows and columns. In general those permutations are difficult to describe, but if we restrict to *irreducible* characters, this is simpler.

Lemma 8.2. *There exists a Zariski dense subset* $\mathring{Z} \subset Z$ *such that*:

- *If* $Z_\alpha \neq Z_\beta$ (*in particular if* $a \neq b$), *then* $\mathring{Z} \cong Z^{irr}_\alpha \times Z^{irr}_\beta \times \mathbb{C}^*$.
- *If* $Z_\alpha = Z_\beta$, *then* $\mathring{Z} \cong Z^{irr}_\alpha \times Z^{irr}_\beta \times \mathbb{C}^*/\sim$, *where the relation is defined by* $(\chi_a, \chi_b, \lambda) \sim (\chi_b, \chi_a, \lambda^{-1})$, *for* $(\chi_a, \chi_b, \lambda) \in Z^{irr}_\alpha \times Z^{irr}_\beta \times \mathbb{C}^*$.

Here Z^{irr}_α denotes the set of irreducible characters in Z_α. We use similar notation for other components of characters and representations.

Proof. Recall from the proof of Lemma 8.1 that T is the Zariski closure of the orbit $\mathrm{SL}_n(\mathbb{C})\cdot(V_\alpha\times V_\beta\times R(\Gamma,\mathbb{C}^*))$. As V_α^{irr} and V_β^{irr} are dense in V_α and V_β, $\mathrm{SL}_n(\mathbb{C})\cdot(V_\alpha^{irr}\times V_\beta^{irr}\times R(\Gamma,\mathbb{C}^*))$ is dense in T. Its projection $\mathring{Z}$ to Z is the image of $Z_\alpha^{irr}\times Z_\beta^{irr}\times\mathbb{C}^*$, which is Zariski dense. To determine this image, we use that each point in $X(\Gamma,\mathrm{SL}_n(\mathbb{C}))$ is the character of a semisimple representation, unique up to conjugation [Lubotzky and Magid 1985]. This uniqueness implies that for $Z_\alpha\neq Z_\beta$ this is an injective map, and for $Z_\alpha=Z_\beta$ we quotient by the permutation of components, with the corresponding transformation for λ. □

Remark 8.3. When $a=b=1$, then $Z_\alpha=Z_\beta$ consists of a single point and Z is the quotient of $\mathbb{C}^*$ by the involution $\lambda\mapsto 1/\lambda$. Hence $Z\cong\mathbb{C}$ and it is the variety of abelian characters in $\mathrm{SL}_2(\mathbb{C})$. The ring of functions invariant by this involution is generated by $\lambda+1/\lambda$, i.e., the trace of a diagonal matrix with eigenvalues λ and $1/\lambda$ (corresponding to the character evaluated at a meridian).

We aim to show that S and T are the only components that contain ρ_λ. For this purpose we consider the quadratic cone $Q(\rho_\lambda)$ which is defined by the vanishing of an obstruction to integrability of 1-cocycles. Let

$$[.\smile.]: H^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})\otimes H^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})\to H^2(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})$$

denote the *cup bracket*, which is the combination of the cup product with the Lie bracket $\mathfrak{sl}_n(\mathbb{C})\otimes\mathfrak{sl}_n(\mathbb{C})\xrightarrow{[\,.\,,\,.\,]}\mathfrak{sl}_n(\mathbb{C})$. The quadratic cone $Q(\rho_\lambda)\subset Z^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})$ is defined by

$$Q(\rho_\lambda)=\{\vartheta\in Z^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})\mid[\vartheta\smile\vartheta]\sim 0\}.$$

Goldman [1984] showed that if $\vartheta\in Z^1(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})$ is integrable then the cup bracket $[\vartheta\smile\vartheta]$ is a coboundary. In what follows we will compute the projections of this obstruction, for the projections

$$\mathrm{pr}_\pm: H^2(\Gamma;\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda})\to H^2(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}}).$$

Here we use the decomposition of Γ-modules:

$$\mathfrak{sl}_n(\mathbb{C})_{\mathrm{Ad}\,\rho_\lambda}=\mathcal{D}\oplus\mathcal{M}^+_{\lambda^n}\oplus\mathcal{M}^-_{\lambda^{-n}}=\mathfrak{sl}_a(\mathbb{C})\oplus\mathfrak{sl}_b(\mathbb{C})\oplus\mathbb{C}\oplus\mathcal{M}^+_{\lambda^n}\oplus\mathcal{M}^-_{\lambda^{-n}}.$$

Recall that Γ acts of $\mathfrak{sl}_n(\mathbb{C})$, $\mathfrak{sl}_a(\mathbb{C})$ and $\mathfrak{sl}_b(\mathbb{C})$ via the adjoint representation $\mathrm{Ad}\,\rho_\lambda$, $\mathrm{Ad}\,\alpha$ and $\mathrm{Ad}\,\beta$ respectively. For the rest of this section we will understand these modules with this action. Recall also that, by the hypotheses of Theorem 1.5, $\Delta_0^+(\lambda^n)\neq 0$ and λ^n is a simple root of Δ_1^+. By Proposition 6.1 we have $\dim H^1(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}})=1$ and we fix $d_\pm\in Z^1(\Gamma;\mathcal{M}^\pm_{\lambda^{\pm n}})$ which represent nontrivial cohomology classes.

Every element in $H^1(\Gamma;\mathfrak{sl}_n(\mathbb{C}))$ is represented by a cocycle

$$\vartheta=\begin{pmatrix} z_a & u_+d_+\\ u_-d_- & z_b\end{pmatrix}+z\varphi\begin{pmatrix}-b\,\mathrm{Id}_a & 0\\ 0 & a\,\mathrm{Id}_b\end{pmatrix},\tag{29}$$

where $z_a \in Z^1(\Gamma; \mathfrak{sl}_a(\mathbb{C}))$, $z_b \in Z^1(\Gamma; \mathfrak{sl}_b(\mathbb{C}))$ and $u_\pm, z \in \mathbb{C}$.

Lemma 8.4. *For $\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C}))$ as in* (29) *we have*

$$\begin{aligned}\mathrm{pr}_+[\vartheta \smile \vartheta] &\sim 2u_+(z_a \smile d_+ + d_+ \smile z_b + znd_+ \smile \varphi),\\ \mathrm{pr}_-[\vartheta \smile \vartheta] &\sim 2u_-(d_- \smile z_a + z_b \smile d_- - znd_- \smile \varphi),\end{aligned}$$

where $\sim$ denotes being cohomologous.

Proof. The lemma follows from Remark 6.4 and a direct calculation of

$$[\vartheta \smile \vartheta](\gamma_1, \gamma_2) = [\vartheta(\gamma_1), \gamma_1\vartheta(\gamma_2)]. \qquad \square$$

In order to understand the cup products appearing in Lemma 8.4 we introduce the complex number $l_\pm(z_a, z_b) \in \mathbb{C}$. Consider a one-parameter analytical deformation $s \mapsto \alpha_s \oplus \beta_s$ of $\alpha \oplus \beta$ in $V_\alpha \times V_\beta$ tangent to (z_a, z_b). Notice that the coefficients of the twisted Alexander polynomial $\Delta_1^{\alpha_s\otimes\beta_s^*}$ depend analytically on s. By the implicit function theorem and since λ^n is a simple root of $\Delta_1^{\alpha\otimes\beta^*}$, there is an analytical path $s \mapsto r_s^+$ of roots of $\Delta_1^{\alpha_s\otimes\beta_s^*}$ with $r_0^+ = \lambda^n$. Similarly there is a path $s \mapsto r_s^-$ of roots of $\Delta_1^{\beta_s\otimes\alpha_s^*}$ with $r_0^- = \lambda^{-n}$. We define

$$l_\pm(z_a, z_b) = \frac{d}{ds}\bigg|_{s=0} \log r_s^\pm.$$

Lemma 8.5. *The following relations hold in $Z^1(\Gamma; \mathcal{M}^\pm_{\lambda^{\pm n}})$:*

$$\begin{aligned}z_a \smile d_+ + d_+ \smile z_b &\sim -l_+(z_a, z_b)d_+ \smile \varphi,\\ d_- \smile z_a + z_b \smile d_- &\sim -l_-(z_a, z_b)d_- \smile \varphi.\end{aligned}$$

Proof. We know that $z_a \smile d_+ + d_+ \smile z_b$ is cohomologous to $xd_+ \smile \varphi$ for some $x \in \mathbb{C}$, as $H^2(\Gamma; \mathcal{M}^+_\lambda) \cong \mathbb{C}$ and $d_+ \smile \varphi \not\sim 0$ (see Proposition 6.1 and Corollary 6.7). Hence by Lemma 7.2 the cocycle

$$\zeta = \begin{pmatrix} z_a & 0 \\ 0 & z_b \end{pmatrix} + \frac{x}{n}\varphi \begin{pmatrix} -b\,\mathrm{Id}_a & 0 \\ 0 & a\,\mathrm{Id}_b \end{pmatrix} \in Z^1(\Gamma; \mathcal{D})$$

satisfies $\boldsymbol{b}(\zeta) \sim 0$ where $\boldsymbol{b} : H^1(\Gamma; \mathcal{D}) \xrightarrow{\boldsymbol{b}} H^2(\Gamma; \mathcal{M}^+_{\lambda^n})$ is the Bockstein operator of the exact cohomology sequence associated to (22).

Furthermore, by Corollary 7.3 ζ is cohomologous to the restriction of a cocycle $\zeta^+ \in Z^1(\Gamma; \mathfrak{p}^+)$. As ρ^+ is a smooth point of $R(\Gamma, P^+)$ (Lemma 7.6), we may consider a path $s \mapsto \rho_s$ in $R(\Gamma, P^+)$ tangent to ζ^+ at ρ^+, which we write as

$$\rho_s = \begin{pmatrix} 1 & d_s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_s\lambda_s^{b\varphi} & 0 \\ 0 & \beta_s\lambda_s^{-a\varphi} \end{pmatrix}.$$

In particular, by the definition of ζ we have that $s \mapsto \alpha_s$ is a deformation of α tangent to z_a, $s \mapsto \beta_s$ is a deformation of β tangent to z_b, and $\lambda_s = \lambda(1 - \frac{x}{n}s + o(s^2))$. By semicontinuity d_s is a cocycle not cohomologous to zero because $d_0 = d_+$, hence by Lemma 5.3 we obtain $\Delta_1^{\alpha_s \otimes \beta_s^*}(\lambda_s^n) = 0$. Therefore, as

$$-\frac{x}{n} = \frac{\lambda_0'}{\lambda_0} = \left.\frac{d}{ds}\right|_{s=0} \log \lambda_s,$$

x equals minus the derivative of the logarithm of the root of $\Delta_1^{\alpha_s \otimes \beta_s^*}$. □

Lemmas 8.4 and 8.5 give:

Corollary 8.6. *For* $\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C}))$ *as in* (29):

$$\mathrm{pr}_\pm[\vartheta \smile \vartheta] \sim 2u_\pm(-l_\pm(z_a, z_b) \pm zn) d_\pm \smile \varphi.$$

Since $\Delta_1^{\alpha_s \otimes \beta_s^*}(t) = \Delta_1^{\beta_s \otimes \alpha_s^*}(1/t)$ by Corollary 4.7,

$$l_+(z_a, z_b) = -l_-(z_a, z_b).$$

Hence the vanishing of the obstructions to integrability of Corollary 8.6 is equivalent to

$$(30) \qquad u_+(-l_+(z_a, z_b) + zn) = 0 \quad \text{and} \quad u_-(-l_+(z_a, z_b) + zn) = 0.$$

Since As z can be interpreted as the derivative of the logarithm of λ, we view

$$-l_+(z_a, z_b) + zn$$

as the derivative of the difference between the logarithm of the root of the Alexander polynomial and the logarithm of λ^n.

Recall that by (29) every cocycle $\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C}))$ is of the form

$$(31) \qquad \vartheta = \begin{pmatrix} z_a & u_+ d_+ + b_+ \\ u_- d_- + b_- & z_b \end{pmatrix} + z\varphi \begin{pmatrix} -b \,\mathrm{Id}_a & 0 \\ 0 & a \,\mathrm{Id}_b \end{pmatrix},$$

where $z_a \in Z^1(\Gamma; \mathfrak{sl}_a(\mathbb{C}))$ and $z_b \in Z^1(\Gamma; \mathfrak{sl}_b(\mathbb{C}))$ are cocycles, $u_\pm, z \in \mathbb{C}$, and $b_\pm \in B^1(\Gamma; \mathcal{M}^\pm_{\lambda^{\pm n}})$ are coboundaries. Notice that this formula differs from (29) because here the coboundaries are also considered.

Proposition 8.7. *The Zariski tangent spaces at* ρ_λ *are*

$$T_{\rho_\lambda} S = \{\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})) \mid -l_+(z_a, z_b) + zn = 0\},$$
$$T_{\rho_\lambda} T = \{\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})) \mid u_+ = u_- = 0\},$$

using the notation of (31) *for a cocycle* $\vartheta \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C}))$. *In particular S and T are smooth and transverse at* ρ_λ.

Proof. First at all, notice that u_+ is not identically zero on $T_{\rho_\lambda} S$, by considering the tangent vector to the path

$$s \mapsto \begin{pmatrix} 1 & sd_+ \\ 0 & 1 \end{pmatrix} \rho_\lambda.$$

Then (30) implies $-l_+(z_a, z_b)+zn=0$ on $T_{\rho_\lambda} S$. Furthermore, we know that $\dim S = n^2+n-2$ and, by (17) and Proposition 6.1, the dimension of $Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})_{\operatorname{Ad}\rho_\lambda})$ is n^2+n-1. This shows the equality for $T_{\rho_\lambda} S$ and proves that ρ_λ is a smooth point of S.

We follow the same lines to prove the equality for $T_{\rho_\lambda} T$. Notice that $-l_+(z_a, z_b)+zn$ is not identically zero on $T_{\rho_\lambda} T$, by considering deformations of λ that keep α and β constant. Hence $u_+ = u_- = 0$ on $T_{\rho_\lambda} T$. Moreover, $\dim T = n^2+n-3$. □

We next compute the tangent space to character varieties at χ_λ. Since the representation ρ_λ is completely reducible, its orbit by conjugation is closed, hence we can apply Luna's slice theorem as in [Ben Abdelghani 2002] or [Heusener and Porti 2005, Section 9]. As a consequence of the slice theorem, since the centralizer of ρ_λ is $\mathbb{C}^*$:

$$T_{\chi_\lambda} X(\Gamma, \mathrm{SL}_n(\mathbb{C})) \cong T_0\big(H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})) /\!\!/ \mathbb{C}^*\big).$$

The action of $\mathbb{C}^*$ can be seen on the coordinates $u_\pm$: an element $\varsigma \in \mathbb{C}^*$ maps $u_\pm$ to $\varsigma^{\pm n} u_\pm$. Hence we define a new coordinate

$$u = u_+ u_-$$

and the obstructions (30) become

$$u(-l_+(z_a, z_b) + zn) = 0. \tag{32}$$

Notice that even if z_a and z_b are cocycles, the logarithmic derivative $-l_+(z_a, z_b)$ only depends on the cohomology class of (z_a, z_b) in $H^1(\Gamma; \mathfrak{sl}_a(\mathbb{C}) \oplus \mathfrak{sl}_b(\mathbb{C}))$. Also, z is the scalar that describes a cohomology class $z\varphi \in H^1(\Gamma; \mathbb{C}) = Z^1(\Gamma; \mathbb{C}) \cong \mathbb{C}$. Similarly for $u_\pm \in \mathbb{C}$ and the cohomology class $u_\pm[d_\pm] \in H^1(\Gamma; \mathcal{M}^{\pm}_{\lambda^{\pm n}}) \cong \mathbb{C}$. Thus we have the following:

Remark 8.8. The obstruction in (32) is well defined in $H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})) /\!\!/ \mathbb{C}^*$.

Corollary 8.9. *The Zariski tangent spaces to Y and Z are*:

$$T_{\chi_\lambda} Y = \{[\vartheta] \in T_{\rho_\lambda} X(\Gamma, \mathrm{SL}_n(\mathbb{C})) \mid -l_+(z_a, z_b) + zn = 0\},$$
$$T_{\chi_\lambda} Z = \{[\vartheta] \in T_{\rho_\lambda} X(\Gamma, \mathrm{SL}_n(\mathbb{C})) \mid u = 0\}.$$

In particular Y and Z are smooth and transverse at χ_λ.

Proof. The proof is similar to that of Proposition 8.7: we need to show that u does not vanish on the Zariski tangent space to Y and $-l_+(z_a, z_b) + zn$ does not vanish on the Zariski tangent space to Z. For the first assertion, we start with the cocycle

$$\vartheta = \begin{pmatrix} 0 & d_+ \\ d_- & 0 \end{pmatrix} \in Z^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})).$$

Following the notation of (31), since z_a, z_b and z vanish for ϑ, Proposition 8.7 implies that $\vartheta \in T_{\rho_\lambda} S$. In particular the projection of its cohomology class ϑ in $H^1(\Gamma; \mathfrak{sl}_n(\mathbb{C})) /\!\!/ \mathbb{C}^*$ is a vector tangent to Y for which $u \neq 0$. The proof that $-l_+(z_a, z_b) + zn$ is not identically zero on $T_{\chi_\lambda} Z$ is the same as in Proposition 8.7. Then one concludes by using the dimension estimates. $\square$

Notice that Corollary 8.9 and the computations of dimensions yield that χ_λ is a smooth point of both Y and Z, and that Y and Z intersect transversally at χ_λ. In particular their intersection is a variety of dimension $n-2$. Since characters in this intersection must satisfy the condition on Alexander polynomials, we have:

Corollary 8.10. *There is a neighborhood* $\chi_\lambda \in U \subset X(\Gamma, \mathrm{SL}_n(\mathbb{C}))$ *such that*

$$(Y \pitchfork Z) \cap U = \big\{(\chi_{\alpha'}, \chi_{\beta'}, \lambda') \in Z \cap U \;\big|\; \Delta_1^{\alpha' \otimes (\beta')^*}\big((\lambda')^n\big) = 0\big\}.$$

9. An example

Let $K \subset S^3$ be the trefoil knot and $\Gamma = \pi_1(S^3 \setminus \mathcal{N}(K))$. We use the presentation

$$\Gamma \cong \langle x, y \mid x^2 = y^3 \rangle,$$

in particular the center is the cyclic group generated by $z = x^2 = y^3$. The abelianization map $\varphi : \Gamma \to \mathbb{Z}$ satisfies $\varphi(x) = 3$, $\varphi(y) = 2$ and a meridian of the trefoil is given by $m = xy^{-1}$.

Lemma 9.1. *Every irreducible representation in* $R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ *is conjugate to* α_s, *where*

$$\alpha_s(x) = \begin{pmatrix} i & 0 \\ s & -i \end{pmatrix} \quad \text{and} \quad \alpha_s(y) = \begin{pmatrix} \eta & \bar{\eta} - \eta \\ 0 & \bar{\eta} \end{pmatrix}, \tag{33}$$

for a unique $s \in \mathbb{C}$ *and for* $\eta \in \mathbb{C}$ *a primitive sixth root of unity. Moreover,* α_s *is irreducible if and only if* $s \neq 0, 2i$.

Proof. Let $\alpha : \Gamma \to \mathrm{SL}_2(\mathbb{C})$ be an irreducible representation. Then by Schur's lemma $\alpha(x^2) = \alpha(y^3)$ lies in the center $\{\pm \mathrm{id}_2\}$ of $\mathrm{SL}_2(\mathbb{C})$. If we had $\alpha(x)^2 = \mathrm{id}_2$, then we would get $\alpha(x) = \pm \mathrm{id}_2$ and α would be reducible, hence $\alpha(x)^2 = \alpha(y^3) = -\mathrm{id}_2$. Furthermore, as $\alpha(y) \neq -\mathrm{id}_2$, the eigenvalues of $\alpha(y)$ are primitive sixth roots of unity. The eigenspaces of $\alpha(x)$ and $\alpha(y)$ determine four points in $\mathbb{P}^1$. These four points are distinct since α is irreducible and by conjugation we can assume that $E_{\alpha(x)}(-i) = [0:1]$ is the point at infinity, $E_{\alpha(y)}(\eta) = [1:0]$ and $E_{\alpha(y)}(\bar{\eta}) = [1:1]$.

The last eigenspace $E_{\alpha(x)}(-i) = [2i : s] = [1 : -is/2]$ determines the representation α up to conjugation. Hence there exists $s \in \mathbb{C}$ such that α is conjugate to α_s. Moreover, the eigenspace $E_{\alpha_s(x)}(-i)$ coincides with an eigenspace of $\alpha_s(y)$ if and only if $s \in \{0, 2i\}$. □

For any representation $\alpha \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ we consider the induced action on $\mathbb{C}^2$, as well as the action $\alpha \otimes t^{\varphi}$ on $\mathbb{C}^2[t^{\pm 1}]$. We aim to compute the twisted Alexander polynomials Δ_0^{α} and Δ_1^{α}, the orders for the homology of $\alpha \otimes t^{\varphi}$. The quotient $\Delta_1^{\alpha_s}/\Delta_0^{\alpha_s}$ has been calculated in a different way in [Kitano and Morifuji 2012, Example 4.3].

Lemma 9.2. *For any irreducible* $\alpha \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$, *we have*

$$\Delta_0^{\alpha} \doteq 1 \ \ \textit{and} \ \ \Delta_1^{\alpha} \doteq t^2 + 1.$$

Proof. First, we have $\Delta_0^{\alpha} \doteq 1$ since α is irreducible and $\dim \mathbb{C}^2 > 1$ (see (4) in the proof of Lemma 2.7).

In order to calculate Δ_1^{α} we will use the amalgamated product structure of Γ

$$\Gamma \cong \langle x \rangle *_{\langle z \rangle} \langle y \rangle$$

and the corresponding Mayer–Vietoris exact sequence in group homology [Brown 1994, VII.9]. We start computing some of the terms. Since $\langle z \rangle \cong \mathbb{Z}$, the groups $H_q(\langle z \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ are the homology groups of the complex

$$0 \to \mathbb{C}^2[t^{\pm 1}] \xrightarrow{z-1} \mathbb{C}^2[t^{\pm 1}] \to 0.$$

Hence a presentation matrix of $H_0(\langle z \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ is

$$\begin{pmatrix} -t^6 - 1 & 0 \\ 0 & -t^6 - 1 \end{pmatrix}.$$

The presentation matrices for $H_0(\langle x \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ and $H_0(\langle y \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ are similarly given by (respectively)

$$\begin{pmatrix} it^3 - 1 & 0 \\ 0 & -it^3 - 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} e^{\frac{\pi i}{3}} t^2 - 1 & 0 \\ 0 & e^{\frac{-\pi i}{3}} t^2 - 1 \end{pmatrix}.$$

Since $H_0(\langle x \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ and $H_0(\langle y \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}})$ are torsion modules it follows that $H_1(\langle x \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) \cong H_1(\langle y \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) = 0$. Hence Mayer–Vietoris gives a short exact sequence

$$0 \to H_1(\Gamma; \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) \to H_0(\langle z \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) \to \\ H_0(\langle x \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) \oplus H_0(\langle y \rangle, \mathbb{C}^2[t^{\pm 1}]_{\alpha \otimes t^{\varphi}}) \to 0.$$

Using this sequence and the presentation matrices we obtain

$$\Delta_1^\alpha = \frac{(t^6+1)^2}{(t^3+i)(t^3-i)\big(t^2-e^{\frac{\pi i}{3}}\big)\big(t^2-e^{\frac{-\pi i}{3}}\big)} = t^2+1. \qquad \square$$

It follows that Theorems 1.4 and 1.5 apply for α irreducible and $\lambda \in \mathbb{C}$ satisfying $\lambda^6 = -1$. Namely Theorem 1.4 yields:

Corollary 9.3. *When $\alpha \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ is irreducible and $\lambda^6 = -1$,*

$$(\lambda^\varphi \otimes \alpha) \oplus (\lambda^{-2\varphi} \otimes \mathbf{1}) : \Gamma \to \mathrm{SL}_3(\mathbb{C})$$

can be deformed to irreducible representations.

To illustrate Theorem 1.5, we discuss next the variety of characters.

Varieties of characters. The variety of characters $X(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ has two components, the abelian one and the one that contains irreducible representations, denoted by $X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$. Let $\chi_s \in X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ denote the character of α_s defined in (33). The following is well known but we provide a proof for completeness and because it is quite straightforward from Lemma 9.1.

Lemma 9.4. *The map $s \mapsto \chi_s$ defines an isomorphism $\mathbb{C} \cong X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$.*

Proof. Using Lemma 9.1, the regular map $f : \mathbb{C} \to X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ given by $f(s) = \chi_s$ restricts to a bijection between $\{s \in \mathbb{C} \mid s \neq 0, 2i\}$ and the set of characters of irreducible representations $X^{irr} \subset X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$. A direct calculation gives for the meridian $m = xy^{-1}$ that $\chi_s(m) = i\bar{\eta} + s(\bar{\eta} - \eta)$ is a linear function in s. Hence there exists a regular map $g : X_0(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to \mathbb{C}$ such that $g \circ f = \mathrm{id}_{\mathbb{C}}$. Since the image of f contains X^{irr}, $f \circ g \circ f = f$ implies

$$f \circ g|_{X^{irr}} = \mathrm{id}_{X^{irr}}\,.$$

Both f and g are regular morphisms (defined on the whole variety, not only on an open subset), hence density yields:

$$f \circ g = \mathrm{id}_{X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))}$$

establishing the isomorphism. $\square$

For any $\lambda \in \mathbb{C}^*$ the map $\alpha \mapsto (\lambda^\varphi \otimes \alpha) \oplus (\lambda^{-2\varphi} \otimes \mathbf{1})$ induces an embedding

$$i_\lambda : X_0(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to X(\Gamma, \mathrm{SL}_3(\mathbb{C})).$$

Let $X_\lambda = i_\lambda(X_0(\Gamma, \mathrm{SL}_2(\mathbb{C})))$ denote its image, that consists of characters of reducible representations. We know that when $\lambda^6 = -1$, X_λ is contained in a two dimensional component that contains irreducible characters. Before describing the global structure of $X(\Gamma, \mathrm{SL}_3(\mathbb{C}))$, we discuss the incidence between the X_λ when $\lambda^6 = -1$.

Let $\tilde{\sigma} : R(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ be the involution such that

$$\tilde{\sigma}(\alpha)(x) = -\alpha(x) \quad \text{and} \quad \tilde{\sigma}(\alpha)(y) = \alpha(y),$$

for every $\alpha \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$, namely $\tilde{\sigma}(\alpha) = (-1)^{\varphi} \otimes \alpha$. Denote by σ the induced involution on $X_0(\Gamma, \mathrm{SL}_2(\mathbb{C}))$. A straightforward computation gives

$$\tilde{\sigma}(\alpha) \mapsto (\lambda^{\varphi} \otimes \tilde{\sigma}(\alpha)) \oplus (\lambda^{-2\varphi} \otimes \mathbf{1}) = ((-\lambda)^{\varphi} \otimes \alpha) \oplus (\lambda^{-2\varphi} \otimes \mathbf{1})$$

and hence $i_\lambda \circ \sigma = i_{-\lambda}$. It follows that $X_\lambda = X_{-\lambda}$. Notice also that $\tilde{\sigma}(\alpha_s)$ is conjugate to α_{2i-s}.

Lemma 9.5. *For $\lambda \neq \pm\lambda'$ satisfying $\lambda^6 = (\lambda')^6 = -1$, X_λ and $X_{\lambda'}$ intersect at a single point $i_\lambda(\chi_s)$, with $s \in \{0, 2i\}$. In particular $X_\lambda \cap X_{\lambda'}$ is the character of a diagonal representation.*

This gives a configuration of three lines $X_{e^{\pi i/6}}$, X_i, $X_{e^{5\pi i/6}}$, that intersect pairwise at one point. We shall prove that there is a single component of $X(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ that contains irreducible representations, and we shall describe how the three lines meet in this component.

Irreducible characters in $X(\Gamma, \mathrm{SL}_3(\mathbb{C}))$. Let $\rho \in R(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ be an irreducible representation. We denote $\rho(x) = A$ and $\rho(y) = B$. The matrix $A^2 = B^3$ is a central element of $\mathrm{SL}_3(\mathbb{C})$ because ρ is irreducible. The center of $\mathrm{SL}_3(\mathbb{C})$ consists of three diagonal matrices $\{\mathrm{id}_3, \omega\,\mathrm{id}_3, \omega^2\,\mathrm{id}_3\}$, where $\omega^2 + \omega + 1 = 0$.

Lemma 9.6. $$A^2 = B^3 = \mathrm{id}_3.$$

Proof. We need to exclude the cases $A^2 = B^3 = \omega\,\mathrm{id}_3$ or equal to $\omega^2\,\mathrm{id}_3$. Seeking a contradiction, assume $A^2 = B^3 = \omega\,\mathrm{id}_3$. The equality $A^2 = \omega\,\mathrm{id}_3$ implies that one eigenvalue of A has multiplicity at least two. Of course multiplicity three is not compatible with irreducibility, thus A has a two-dimensional eigenspace. On the other hand, $B^3 - \omega\,\mathrm{id}_3 = 0$ combined with $\det(B) = 1$ yields that the minimal polynomial of B has also degree two. Hence B has also a two dimensional eigenspace. The intersection of the two dimensional eigenspaces of A and B is a proper invariant subspace, contradicting irreducibility. The same argument applies to $A^2 = B^3 = \omega^2\,\mathrm{id}_3$. □

By the discussion in the proof of the previous lemma, the minimal polynomial of A is $A^2 - \mathrm{id}_3 = 0$ and the minimal polynomial of B is $B^3 - \mathrm{id}_3 = 0$. Therefore, the matrices A and B are conjugate to

$$A \sim \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \quad \text{and} \quad B \sim \begin{pmatrix} 1 & & \\ & \omega & \\ & & \omega^2 \end{pmatrix},$$

where $\omega^2 + \omega + 1 = 0$. The corresponding eigenspaces are the plane $E_A(-1)$ and the lines $E_A(1)$, $E_B(1)$, $E_B(\omega)$ and $E_B(\omega^2)$. The eigenspaces determine the representation, as they determine the matrices A and B, that have fixed eigenvalues. Of course $E_A(1) \cap E_A(-1) = 0$ and $E_B(1)$, $E_B(\omega)$ and $E_B(\omega^2)$ are also in general position. Since ρ is irreducible, the five eigenspaces are in general position. For instance $E_A(1) \cap (E_B(1) \oplus E_B(\omega)) = 0$, because otherwise $E_B(1) \oplus E_B(\omega) = E_A(1) \oplus (E_A(-1) \cap (E_B(1) \oplus E_B(\omega)))$ would be a proper invariant subspace.

In order to parametrize the conjugacy classes of the irreducible representations, we fix some normalizations of those eigenspaces. The invariant lines correspond to fixed points in the projective plane $\mathbb{P}^2$. The first normalization is that $E_A(-1)$ corresponds to the line at infinity, so that the 4 invariant lines are points in the affine plane $\mathbb{C}^2$ in general position. We further fix the three fixed points of B, corresponding to an affine frame. Then the fourth point (the line $E_A(1)$) is a point in $\mathbb{C}^2$ that does not lie in the affine lines spanned by any two of the fixed points of B. This gives rise to the subvariety $\{\rho_{s,t} \in R(\Gamma, \mathrm{SL}_3(\mathbb{C})) \mid (s,t) \in \mathbb{C}^2\}$, where the representation $\rho_{s,t}$ is given by

$$\rho_{s,t}(x) = \begin{pmatrix} 1 & 0 & 0 \\ s & -1 & 0 \\ t & 0 & -1 \end{pmatrix} \quad \text{and} \quad \rho_{s,t}(y) = \begin{pmatrix} 1 & \omega - 1 & \omega^2 - 1 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}. \tag{34}$$

Here ω is a primitive third root of unity, i.e., $\omega^2 + \omega + 1 = 0$. The eigenspaces of B determine points of $\mathbb{P}^2$:

$$E_B(1) = [1:0:0],\ \ E_B(\omega) = [1:1:0] \text{ and } E_B(\omega^2) = [1:0:1].$$

The eigenspaces of A determine a projective line (at infinity) and a point:

$$E_A(-1) = \langle [0:1:0], [0:0:1] \rangle \text{ and } E_A(1) = [2:s:t].$$

Lemma 9.7. *For $(s,t) \in \mathbb{C}^2$, the representation $\rho_{s,t}$ is reducible if and only if $s = 0$, $t = 0$, or $s + t = 2$.*

Proof. The representation $\rho_{s,t}$ is constructed so that the points in $\mathbb{P}^2$ fixed by $B = \rho_{s,t}(x)$ and the line $E_A(-1) \subset \mathbb{P}^2$ are fixed. So $\rho_{s,t}$ is reducible if and only if the projective point $E_A(1)$ belongs to one of the lines spanned by two of the fixed points of B. This condition is equivalent to one of the three equations $s = 0$, $t = 0$ or $s + t = 2$, one for each line. □

It follows from the proof that, when $E_A(1)$ equals one of the fixed projective points of B, then A preserves also the two lines through that point that are B-invariant. More precisely, we have:

Remark 9.8. If two of the equations $\{s = 0\}$, $\{t = 0\}$ and $\{s + t = 2\}$ hold true, then $\rho_{s,t}$ preserves a complete flag in $\mathbb{C}^3$ and therefore it is conjugate to an upper triangular representation. Notice that it has the same character as a diagonal representation.

Lemma 9.9. *Let* $R^{irr} \subset R(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ *denote the subset of irreducible representations. Then the Zariski closure* $\overline{R^{irr}} \subset R(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ *is an irreducible affine variety.*

Proof. The variety $\mathbb{C}^2 \times \mathrm{SL}_3(\mathbb{C})$ is irreducible and the map $\kappa : \mathbb{C}^2 \times \mathrm{SL}_3(\mathbb{C}) \to R(\Gamma, \mathrm{SL}_3(\mathbb{C}))$ given by $\kappa(s,t,D) = D\rho_{s,t}D^{-1}$ is a regular map. The image of κ contains the irreducible representations and every representation in the image of κ is the limit of irreducible representations. Hence

$$R^{irr} \subset \kappa\big(\mathbb{C}^2 \times \mathrm{SL}_3(\mathbb{C})\big) \subset \overline{R^{irr}}$$

and $\overline{\kappa\big(\mathbb{C}^2 \times \mathrm{SL}_3(\mathbb{C})\big)} = \overline{R^{irr}}$ follows. Now the assertion of the lemma follows since the closure of the image of an irreducible variety under a regular map is irreducible. □

Theorem 9.10. *The GIT quotient* $X = \overline{R^{irr}} /\!\!/ \mathrm{SL}(3,\mathbb{C})$ *is isomorphic to* $\mathbb{C}^2$*. Moreover, the Zariski-open subset* R^{irr} *is* $\mathrm{SL}(3,\mathbb{C})$*-invariant and its GIT quotient is isomorphic to the complement of three affine lines in general position in* $\mathbb{C}^2$*.*

Proof. By Lemma 9.9 the affine algebraic set $\overline{R^{irr}}$ is irreducible. Since it is $\mathrm{SL}(3,\mathbb{C})$-invariant, the GIT quotient $t : \overline{R^{irr}} \to X$ exists and X is also an irreducible affine algebraic variety. Let $X^{irr} \subset X$ denote the projection of R^{irr}, which is Zariski-open and hence dense.

Consider the regular morphism $f : \mathbb{C}^2 \to X$ that maps $(s,t) \in \mathbb{C}^2$ to the character $\chi_{\rho_{s,t}}$. By construction, the image of f contains X^{irr}, because $\rho_{s,t}$ realizes every irreducible representation up to conjugacy.

There is also a regular morphism $R(\Gamma, \mathrm{SL}_3(\mathbb{C})) \to \mathbb{C}^2$ given by

$$\rho \mapsto \big(\operatorname{tr}\rho(m), \operatorname{tr}\rho(m^{-1})\big)$$

where $m = xy^{-1}$ is a meridian of the trefoil knot, which induces (after restriction) a regular map $X \to \mathbb{C}^2$, by invariance. A direct computation gives:

$$\begin{pmatrix} \operatorname{tr}\rho_{s,t}(m) \\ \operatorname{tr}\rho_{s,t}(m^{-1}) \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} \omega^2-1 & \omega-1 \\ \omega-1 & \omega^2-1 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}. \tag{35}$$

Thus, after composing with a linear map, we have a regular morphism $g : X \to \mathbb{C}^2$ that satisfies

$$g \circ f = \mathrm{id}_{\mathbb{C}^2}.$$

Since the image of f contains X^{irr}, $f \circ g \circ f = f$ implies

$$f \circ g|_{X^{irr}} = \mathrm{id}_{X^{irr}}.$$

Both f and g are regular morphisms (defined on the whole variety, not only on an open subset), hence density yields

$$f \circ g = \mathrm{id}_X,$$

establishing the isomorphism. □

Remark 9.11. It follows from Theorem 9.10 that the set of reducible characters in $X \cong \mathbb{C}^2$ consists of three lines that intersect pairwise. Those are characters of representations $(\lambda^{-\varphi} \otimes \alpha) \oplus (\lambda^{2\varphi} \otimes \mathbf{1})$, with $\alpha \in R(\Gamma, \mathrm{SL}_2(\mathbb{C}))$ irreducible except at the intersection points, that correspond to diagonal representations.

Notice also that there is a symmetry of order three, as the center of $\mathrm{SL}_3(\mathbb{C})$ has order three. The symmetry group is generated by

$$R(\Gamma, \mathrm{SL}_3(\mathbb{C})) \to R(\Gamma, \mathrm{SL}_3(\mathbb{C})), \quad \alpha \mapsto \omega^{\varphi} \otimes \alpha,$$

where ω is a primitive third root of unity. This symmetry maps the character with coordinates (s, t) to $(2-s-t, s)$, i.e., $\mathrm{tr}(\rho(m^{\pm 1}))$ to $\omega^{\pm 1}\,\mathrm{tr}(\rho(m^{\pm 1}))$. Its fixed point has coordinates $s = t = 2/3$ (i.e., $\mathrm{tr}(\rho(m^{\pm 1})) = 0$) and corresponds to the character of an irreducible metabelian representation. This irreducible metabelian representation is obtained by composing the surjection $\Gamma \twoheadrightarrow A_4$ with the 3-dimensional irreducible representation of A_4 (see [Serre 1978]). Note that irreducible, metabelian representations of knot groups into $\mathrm{SL}_n(\mathbb{C})$ were studied by H. Boden and S. Friedl in a series of papers [2008; 2011; 2014a; 2014b].

Remark 9.12. It is possible to combine any representation $\rho : \Gamma \to \mathrm{SL}_2(\mathbb{C})$ with the irreducible 3-dimensional rational representation of $r_3 : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{SL}_3(\mathbb{C})$ of $\mathrm{SL}_2(\mathbb{C})$ (for more details see [Springer 1977] and [Heusener and Medjerab 2014]). This induces a regular map

$$(r_3)_* : X_0(\Gamma, \mathrm{SL}_2(\mathbb{C})) \to X(\Gamma, \mathrm{SL}_3(\mathbb{C})).$$

It follows from [Heusener and Medjerab 2014, Proposition 3.1] that the image of $(r_3)_*$ is contained in the component $X \subset X(\Gamma, \mathrm{SL}_3(\mathbb{C}))$. Notice that for every matrix $A \in \mathrm{SL}_2(\mathbb{C})$ the equality $\mathrm{tr}(r_3(A)) = \mathrm{tr}(r_3(A)^{-1})$ holds. Then Equation (35) implies that the image of $(r_3)_*$ is contained in the diagonal $\{s = t\} \subset \mathbb{C}^2 \cong X$. Moreover, the map $(r_3)_*$ factors through $X_0(\Gamma, \mathrm{PSL}_2(\mathbb{C}))$ since $\mathrm{Ker}(r_3) = \{\pm\,\mathrm{id}\}$ is the center of $\mathrm{SL}_2(\mathbb{C})$. Hence $(r_3)_*$ is a two-fold branched covering onto its image. The branching set is the character of the binary dihedral representation $d_6 : \Gamma \to D_6 \subset \mathrm{SL}_2(\mathbb{C})$. Notice also that the restriction of r_3 onto D_6 becomes reducible, $r_3 \circ d_6 \sim \rho_{1,1}$, since dihedral groups have only one and two-dimensional irreducible representations (see [Serre 1978]).

Remark 9.13. The same argument as in Theorem 9.10 applies to torus knots $T(p, 2)$, p odd, to prove that the variety of irreducible $\mathrm{SL}_3(\mathbb{C})$-characters consist of $(p-1)(p-2)/2$ disjoint components isomorphic to $\mathbb{C}^2$ and the components of reducible characters.

Acknowledgements

We are indebted to Julien Bichon for helpful discussions and we like to thank Simon Riche for pointing out Nagata's result [1961/1962] to us. We also like to thank Stefan Friedl for providing us with the references [Friedl et al. 2012; Hillman et al. 2010]. We are particularly thankful to the anonymous referee for a thorough review and for pointing out an inaccuracy in the statement of a preliminary version of Theorem 1.5. The referee's remarks led to Lemma 8.2.

References

[Artin 1968] M. Artin, "On the solutions of analytic equations", *Invent. Math.* **5** (1968), 277–291. MR 38 #344 Zbl 0172.05301

[Ben Abdelghani 2000] L. Ben Abdelghani, "Espace des représentations du groupe d'un nœud classique dans un groupe de Lie", *Ann. Inst. Fourier* (*Grenoble*) **50**:4 (2000), 1297–1321. MR 2002d:57006a Zbl 0956.57006

[Ben Abdelghani 2002] L. Ben Abdelghani, "Variété des caractères et slice étale de l'espace des représentations d'un groupe", *Ann. Fac. Sci. Toulouse Math.* (6) **11**:1 (2002), 19–32. MR 2004h:14067 Zbl 1056.20032

[Ben Abdelghani and Lines 2002] L. Ben Abdelghani and D. Lines, "Involutions on knot groups and varieties of representations in a Lie group", *J. Knot Theory Ramifications* **11**:1 (2002), 81–104. MR 2002m:57013 Zbl 0997.57036

[Ben Abdelghani et al. 2010] L. Ben Abdelghani, M. Heusener, and H. Jebali, "Deformations of metabelian representations of knot groups into $\mathrm{SL}(3, \mathbb{C})$", *J. Knot Theory Ramifications* **19**:3 (2010), 385–404. MR 2011d:57027 Zbl 1195.57023

[Boden and Friedl 2008] H. U. Boden and S. Friedl, "Metabelian $\mathrm{SL}(n, \mathbb{C})$ representations of knot groups", *Pacific J. Math.* **238**:1 (2008), 7–25. MR 2010c:57007 Zbl 1154.57004

[Boden and Friedl 2011] H. U. Boden and S. Friedl, "Metabelian $\mathrm{SL}(n, \mathbb{C})$ representations of knot groups, II: fixed points", *Pacific J. Math.* **249**:1 (2011), 1–10. MR 2012a:57007 Zbl 1218.57002

[Boden and Friedl 2014a] H. U. Boden and S. Friedl, "Metabelian $\mathrm{SL}(n, \mathbb{C})$ representations of knot groups III: deformations", *Q. J. Math.* **3** (2014), 817–840. Zbl 1309.57010

[Boden and Friedl 2014b] H. U. Boden and S. Friedl, "Metabelian $\mathrm{SL}(n, \mathbb{C})$ representations of knot groups IV: twisted Alexander polynomials", *Math. Proc. Cambridge Philos. Soc.* **156**:1 (2014), 81–97. MR 3144211 Zbl 1303.57005

[Brown 1994] K. S. Brown, *Cohomology of groups*, Graduate Texts in Mathematics **87**, Springer, New York, 1994. MR 96a:20072 Zbl 0823.20049

[Burde 1967] G. Burde, "Darstellungen von Knotengruppen", *Math. Ann.* **173** (1967), 24–33. MR 35 #3652 Zbl 0146.45602

[Franz 1937] W. Franz, "Torsionsideale, Torsionsklassen und Torsion", *J. Reine Angew. Math.* **176** (1937), 113–125.

[Friedl et al. 2012] S. Friedl, T. Kim, and T. Kitayama, "Poincaré duality and degrees of twisted Alexander polynomials", *Indiana Univ. Math. J.* **61**:1 (2012), 147–192. MR 3029395 Zbl 1273.57009

[Frohman and Klassen 1991] C. D. Frohman and E. P. Klassen, "Deforming representations of knot groups in SU(2)", *Comment. Math. Helv.* **66**:3 (1991), 340–361. MR 93a:57001 Zbl 0738.57001

[Goldman 1984] W. M. Goldman, "The symplectic nature of fundamental groups of surfaces", *Adv. in Math.* **54**:2 (1984), 200–225. MR 86i:32042 Zbl 0574.32032

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Heusener and Klassen 1997] M. Heusener and E. Klassen, "Deformations of dihedral representations", *Proc. Amer. Math. Soc.* **125**:10 (1997), 3039–3047. MR 98b:57011 Zbl 1154.57004

[Heusener and Kroll 1998] M. Heusener and J. Kroll, "Deforming abelian SU(2)-representations of knot groups", *Comment. Math. Helv.* **73**:3 (1998), 480–498. MR 99g:57008 Zbl 0910.57004

[Heusener and Medjerab 2014] M. Heusener and O. Medjerab, "Deformations of reducible representations of knot groups into $\mathrm{SL}(n, \mathbb{C})$", arXiv:1402.4294, 2014. To appear in *Math. Slovaca*.

[Heusener and Porti 2005] M. Heusener and J. Porti, "Deformations of reducible representations of 3-manifold groups into $\mathrm{PSL}_2(\mathbb{C})$", *Algebr. Geom. Topol.* **5** (2005), 965–997. MR 2006e:57016 Zbl 1082.57007

[Heusener and Porti 2011] M. Heusener and J. Porti, "Infinitesimal projective rigidity under Dehn filling", *Geom. Topol.* **15**:4 (2011), 2017–2071. MR 2012j:57039 Zbl 1237.57016

[Heusener et al. 2001] M. Heusener, J. Porti, and E. Suárez Peiró, "Deformations of reducible representations of 3-manifold groups into $\mathrm{SL}_2(\mathbb{C})$", *J. Reine Angew. Math.* **530** (2001), 191–227. MR 2002a:57002 Zbl 0964.57006

[Hillman et al. 2010] J. A. Hillman, D. S. Silver, and S. G. Williams, "On reciprocality of twisted Alexander invariants", *Algebr. Geom. Topol.* **10**:2 (2010), 1017–1026. MR 2011j:57021 Zbl 1200.57005

[Humphreys 1975] J. E. Humphreys, *Linear algebraic groups*, Graduate Texts in Mathematics **21**, Springer, New York, 1975. MR 53 #633 Zbl 0325.20039

[Humphreys 1995] J. E. Humphreys, *Conjugacy classes in semisimple algebraic groups*, Mathematical Surveys and Monographs **43**, American Mathematical Society, Providence, RI, 1995. MR 97i:20057 Zbl 0834.20048

[Johnson and Millson 1987] D. Johnson and J. J. Millson, "Deformation spaces associated to compact hyperbolic manifolds", pp. 48–106 in *Discrete groups in geometry and analysis*, edited by R. Howe, Progr. Math. **67**, Birkhäuser, Boston, MA, 1987. MR 88j:22010 Zbl 0664.53023

[Kirk and Livingston 1999] P. Kirk and C. Livingston, "Twisted Alexander invariants, Reidemeister torsion, and Casson–Gordon invariants", *Topology* **38**:3 (1999), 635–661. MR 2000c:57010 Zbl 0883.57001

[Kitano 1996] T. Kitano, "Twisted Alexander polynomial and Reidemeister torsion", *Pacific J. Math.* **174**:2 (1996), 431–442. MR 97g:57007 Zbl 0863.57001

[Kitano and Morifuji 2012] T. Kitano and T. Morifuji, "Twisted Alexander polynomials for irreducible $SL(2, \mathbb{C})$-representations of torus knots", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **11**:2 (2012), 395–406. MR 3011996 Zbl 1255.57014

[Lubotzky and Magid 1985] A. Lubotzky and A. R. Magid, "Varieties of representations of finitely generated groups", *Mem. Amer. Math. Soc.* **58**:336 (1985), xi+117. MR 87c:20021 Zbl 0598.14042

[Milnor 1962] J. Milnor, "A duality theorem for Reidemeister torsion", *Ann. of Math.* (2) **76** (1962), 137–147. MR 25 #4526 Zbl 0108.36502

[Nagata 1961/1962] M. Nagata, "Complete reducibility of rational representations of a matric group", *J. Math. Kyoto Univ.* **1** (1961/1962), 87–99. MR 26 #236 Zbl 0106.25201

[Newstead 1978] P. E. Newstead, *Introduction to moduli problems and orbit spaces*, Tata Institute Lectures on Mathematics and Physics **51**, Narosa Publishing House, New Delhi, 1978. MR 81k:14002 Zbl 0411.14003

[Popov 2008] V. L. Popov, "Irregular and singular loci of commuting varieties", *Transform. Groups* **13**:3-4 (2008), 819–837. MR 2009h:14095 Zbl 1189.17017

[Porti 1997] J. Porti, "Torsion de Reidemeister pour les variétés hyperboliques", *Mem. Amer. Math. Soc.* **128**:612 (1997), x+139. MR 98g:57034 Zbl 0928.57005

[de Rham 1967] G. de Rham, "Introduction aux polynômes d'un nœud", *Enseignement Math.* (2) **13** (1967), 187–194. MR 39 #2149 Zbl 0157.54803

[Richardson 1979] R. Richardson, "Commuting varieties of semisimple Lie algebras and algebraic groups", *Compos. Math.* **38** (1979), 311–327. MR 0535074 Zbl 0409.17006

[Serre 1978] J.-P. Serre, *Représentations linéaires des groupes finis*, revised ed., Hermann, Paris, 1978. MR 80f:20001 Zbl 0407.20003

[Shafarevich 1977] I. R. Shafarevich, *Basic algebraic geometry*, Springer, Berlin, 1977. MR 56 #5538 Zbl 0362.14001

[Shafarevich 1994] I. R. Shafarevich (editor), *Algebraic geometry. IV*, Encyclopaedia of Mathematical Sciences **55**, Springer, Berlin, 1994. MR 95g:14002 Zbl 0796.14002

[Shors 1991] D. J. Shors, *Deforming reducible representations of knot groups in SL(2)(C)*, Ph.D. thesis, University of California, Los Angeles, 1991, available at http://search.proquest.com/docview/303918617.

[Springer 1977] T. A. Springer, *Invariant theory*, Lecture Notes in Mathematics **585**, Springer, Berlin, 1977. MR 56 #5740 Zbl 0346.20020

[Turaev 1986] V. G. Turaev, "Reidemeister torsion in knot theory", *Uspekhi Mat. Nauk* **41**:1(247) (1986), 97–147, 240. In Russian; translated in *Russian Math. Surveys* **41**:1 (1986), 119–182. MR 87i:57009 Zbl 0602.57005

[Wada 1994] M. Wada, "Twisted Alexander polynomial for finitely presentable groups", *Topology* **33**:2 (1994), 241–256. MR 95g:57021 Zbl 0822.57006

[Weil 1964] A. Weil, "Remarks on the cohomology of groups", *Ann. of Math.* (2) **80** (1964), 149–157. MR 30 #199 Zbl 0192.12802

Received July 14, 2014. Revised March 11, 2015.

MICHAEL HEUSENER
LABORATOIRE DE MATHÉMATIQUES
CLERMONT UNIVERSITÉ AUVERGNE, UNIVERSITÉ BLAISE PASCAL
BP 10448, F-63000 CLERMONT-FERRAND
CNRS, UMR 6620, LM, F-63171 AUBIÈRE
FRANCE
heusener@math.univ-bpclermont.fr

JOAN PORTI
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT AUTÒNOMA DE BARCELONA
CERDANYOLA DEL VALLES
08193 BARCELONA
SPAIN
porti@mat.uab.cat

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.355

APPROXIMATIONS BY MAXIMAL COHEN–MACAULAY MODULES

HENRIK HOLM

Auslander and Buchweitz have proved that every finitely generated module over a Cohen–Macaulay (CM) ring with a dualizing module admits a so-called maximal CM approximation. In terms of relative homological algebra, this means that every finitely generated module has a special maximal CM precover. In this paper, we prove the existence of special maximal CM preenvelopes and, in the case where the ground ring is henselian, of maximal CM envelopes. We also characterize the rings over which every finitely generated module has a maximal CM envelope with the unique lifting property. Finally, we show that cosyzygies with respect to the class of maximal CM modules must eventually be maximal CM, and we compute some examples.

1. Introduction

Let R be a commutative noetherian local Cohen–Macaulay (CM) ring with a dualizing module Ω and denote by MCM the class of maximal CM R-modules. Auslander and Buchweitz [1989, Theorem A] construct a *maximal CM approximation* for every finitely generated R-module M, that is, a short exact sequence

$$0 \longrightarrow I \longrightarrow X \xrightarrow{\pi} M \longrightarrow 0,$$

where X belongs to MCM and I has finite injective dimension. By a result form [Ischebeck 1969] one has $\operatorname{Ext}^1_R(Y, I) = 0$ for all Y in MCM, so in terms of relative homological algebra, this means that the homomorphism $\pi : X \twoheadrightarrow M$ is a *special* MCM-*precover* of M. Corollary 2.5 of [Takahashi 2005] shows that if R is henselian (for example, if R is complete), then every MCM-precover can be refined to an MCM-*cover*. The corollary follows from Takahashi's Proposition 2.4, which the author attributes to Yoshino [1993, Lemma 2.2]. We summarize these results in the following theorem.

MSC2010: primary 13C14; secondary 13D05.

Keywords: cosyzygy, envelope, maximal Cohen-Macaulay module, special preenvelope, unique lifting property.

Theorem [Auslander and Buchweitz 1989; Takahashi 2005; Yoshino 1993].

(a) *Every finitely generated R-module has a special* MCM*-precover* (*also called a special right* MCM*-approximation*).

(b) *If R is henselian, then every finitely generated R-module has an* MCM*-cover* (*also called a minimal right* MCM*-approximation*).

This paper is concerned with the existence and the construction of special MCM-preenvelopes and MCM-envelopes of finitely generated modules. Our first main result, which is proved in Section 3, is the following "dual" of the theorem above.

Theorem A. (a) *Every finitely generated R-module M has a special* MCM*-preenvelope* (*also called a special left* MCM*-approximation*).

(b) *If R is henselian, then every finitely generated R-module has an* MCM*-envelope* (*also called a minimal left* MCM*-approximation*).

(c) *Every special* MCM*-preenvelope* (*and hence every* MCM*-envelope*) $\mu : M \to X$ *of a finitely generated R-module M has the property that* $\mathrm{Hom}_R(\mathrm{Coker}\,\mu, \Omega)$ *has finite injective dimension.*

Theorem C of [Holm 2014] showed the existence of (nonspecial!) MCM-preenvelopes, but its proof is not constructive: it is a consequence of an abstract result — Theorem (4.2) of [Crawley-Boevey 1994] — combined with the fact, also proved in [Holm 2014], that the direct limit closure of MCM is closed under products. Theorem A above is not only stronger than [Holm 2014, Theorem C]; our proof, modeled on that of [Holm and Jørgensen 2011, Theorem 1.6], also shows how (special) MCM-(pre)envelopes can be constructed from (special) MCM-(pre)covers.

In Section 4 we compute the MCM-envelope of some specific modules. In Section 5 we turn our attention to MCM-envelopes with the *unique lifting property*, and we characterize the rings over which every finitely generated module admits such an envelope:

Theorem B. *The following conditions are equivalent.*

(i) *For every finitely generated R-module M, the module* $\mathrm{Hom}_R(M, \Omega)$ *is maximal CM.*

(ii) *The Krull dimension of R is* $\leqslant 2$.

(iii) *The inclusion functor* $\mathsf{MCM} \hookrightarrow \mathsf{mod}$ *has a left adjoint.*

(iv) *Every finitely generated R-module has an* MCM*-envelope with the unique lifting property.*

From a homological point of view, maximal CM modules are interesting because every module can be finitely resolved by such modules. More precisely, if d denotes the Krull dimension of the CM ring R, and if M is any finitely generated R-module with a resolution

$$\cdots \longrightarrow X_d \longrightarrow X_{d-1} \longrightarrow X_{d-2} \longrightarrow \cdots \longrightarrow X_1 \longrightarrow X_0 \longrightarrow M \longrightarrow 0$$

by finitely generated free R-modules $X_0, X_1, \ldots$, then the n-th syzygy of M, i.e., the module $\operatorname{Syz}_n(M) = \operatorname{Ker}(X_{n-1} \to X_{n-2})$, is maximal CM for every $n \geqslant d$. Actually, the same conclusion holds if $X_0, X_1, \ldots$ are just assumed to be maximal CM (but not necessarily free); this well-known fact follows from the behavior of depth in short exact sequences; see [Bruns and Herzog 1993, Proposition 1.2.9] or Lemma 2.5. Given a finitely generated R-module M, one can *not* always construct an *exact* sequence

$$(*) \qquad\qquad 0 \longrightarrow M \longrightarrow X^0 \longrightarrow X^1 \longrightarrow \cdots$$

where $X^0, X^1, \ldots$ are maximal CM; however, there is a canonical way to construct a *complex* of the form $(*)$. Theorem A guarantees the existence of MCM-preenvelopes, which makes the following construction possible: take an MCM-preenvelope $\mu^0 : M \to X^0$ of M and set $C^1 = \operatorname{Coker}\mu^0$; take an MCM-preenvelope $\mu^1 : C^1 \to X^1$ of C^1 and set $C^2 = \operatorname{Coker}\mu^1$; etc. The hereby constructed complex $(*)$ — which is called a *proper* MCM-*coresolution* or an MCM-*resolvent* of M — is not necessarily exact, but it becomes exact if one applies the functor $\operatorname{Hom}_R(-, Y)$ to it for any Y in MCM. The module $C^n = \operatorname{Coker}(X^{n-2} \to X^{n-1})$ is called the *n-th cosyzygy of M with respect to* MCM, and it is denoted by $\operatorname{Cosyz}^n_{\mathsf{MCM}}(M)$. In Section 6 we prove that such cosyzygies must eventually be maximal CM:

Theorem C. *Let M be a finitely generated R-module. For every $n \geqslant d$, any n-th cosyzygy* $\operatorname{Cosyz}^n_{\mathsf{MCM}}(M)$ *of M with respect to* MCM *is maximal CM.*

2. Preliminaries

Setup 2.1. Throughout, $(R, \mathfrak{m}, k)$ is a commutative noetherian local CM ring of Krull dimension d. It is assumed that R has a dualizing (or canonical) module Ω.

Let M be a finitely generated R-module. The *depth* of M is the number

$$\operatorname{depth}_R M = \inf\{i \mid \operatorname{Ext}^i_R(k, M) \neq 0\} \in \mathbb{N}_0 \cup \{\infty\};$$

see [Bruns and Herzog 1993, Definitions 1.2.6 and 1.2.7]. If $M \neq 0$, then $\operatorname{depth}_R M$ is the common length of a maximal M-regular sequence (in $\mathfrak{m}$). The depth can also be computed from the dualizing module:

$$\operatorname{depth}_R M = d - \sup\{i \mid \operatorname{Ext}^i_R(M, \Omega) \neq 0\};$$

see [Bruns and Herzog 1993, Corollary 3.5.11]. One calls M *maximal CM* if $\operatorname{depth}_R M \geqslant d$, that is, if $\operatorname{Ext}^i_R(M, \Omega) = 0$ for all $i > 0$. The category of all such

R-modules is denoted by MCM. Note that the zero module is maximal CM and has depth ∞. The category of all finitely generated R-modules is denoted by mod.

We recall a few notions from relative homological algebra.

Definition 2.2. Let $\mathcal{A}$ be a full subcategory of an abelian category $\mathcal{M}$ (e.g., $\mathcal{M} = \mathsf{mod}$ and $\mathcal{A} = \mathsf{MCM}$), and let M be an object in $\mathcal{M}$. Following [Enochs and Jenda 2000, Definition 6.1.1], a morphism $\varepsilon : M \to A$ with $A \in \mathcal{A}$ is called an *$\mathcal{A}$-preenvelope* (or a *left $\mathcal{A}$-approximation*) of M if every other morphism $\varepsilon' : M \to A'$ with $A' \in \mathcal{A}$ factors through ε, as illustrated below.

$$\begin{tikzcd}
M \arrow[r, "\varepsilon"] \arrow[d, "\varepsilon'"'] & A \arrow[dl, dashed] \\
A' &
\end{tikzcd}$$

A *special $\mathcal{A}$-preenvelope* (or a *special left $\mathcal{A}$-approximation*) is an $\mathcal{A}$-preenvelope $\varepsilon : M \to A$ such that $\operatorname{Ext}^1_{\mathcal{M}}(\operatorname{Coker}\varepsilon, A') = 0$ for every $A' \in \mathcal{A}$. An *$\mathcal{A}$-envelope* (or a *minimal left $\mathcal{A}$-approximation*) is an $\mathcal{A}$-preenvelope ε with the property that every endomorphism φ of A that satisfies $\varphi\varepsilon = \varepsilon$ is an automorphism.

Remark 2.3. The notions of *$\mathcal{A}$-precover* (or *right $\mathcal{A}$-approximation*), *special $\mathcal{A}$-precover* (or *special right $\mathcal{A}$-approximation*), and *$\mathcal{A}$-cover* (or *minimal right $\mathcal{A}$-approximation*) are categorically dual to the notions defined above.

By definition, a special $\mathcal{A}$-precover/preenvelope is also an (ordinary) $\mathcal{A}$-precover/ preenvelope. If $\mathcal{A}$ is closed under extensions in $\mathcal{M}$, then every $\mathcal{A}$-cover/envelope is a special $\mathcal{A}$-precover/preenvelope; this is the content of Wakamatsu's lemma.[1]

Remark 2.4. It is well-known that the dualizing module Ω gives rise to a duality on the category of maximal CM modules; more precisely, there is an equivalence of categories:

$$\begin{tikzcd}
\mathsf{MCM} \arrow[r, shift left, "\operatorname{Hom}_R(-,\Omega)"] & \mathsf{MCM}^{\mathrm{op}}. \arrow[l, shift left, "\operatorname{Hom}_R(-,\Omega)"]
\end{tikzcd}$$

We use the shorthand notation $(-)^{\dagger}$ for the functor $\operatorname{Hom}_R(-, \Omega)$. For any finitely generated R-module M there is a canonical homomorphism $\delta_M : M \to M^{\dagger\dagger}$, called the *biduality homomorphism*, which is natural in M. An alternative way of phrasing the equivalence above is to say δ_M is an isomorphism if M belongs to MCM; see [Bruns and Herzog 1993, Theorem 3.3.10].

We will need the following result about depth; it is folklore and easily proved.

---

[1] This result is implicit in [Wakamatsu 1988]. It is explicitly stated in [Auslander and Reiten 1991, Lemma 1.3], but without a proof. It is stated and proved in [Xu 1996, Lemmas 2.1.1 and 2.1.2].

Lemma 2.5. *Let* $m \geqslant 0$ *be an integer and let* $0 \to K_m \to X_{m-1} \to \cdots \to X_0 \to M \to 0$ *be an exact sequence of finitely generated* R*-modules. If* $X_0, \dots, X_{m-1}$ *are maximal CM, then one has* $\operatorname{depth}_R K_m \geqslant \min\{d, \operatorname{depth}_R M + m\}$*. In particular, if* $m \geqslant d$ *then the* R*-module* K_m *is maximal CM.* □

3. Special MCM-preenvelopes and MCM-envelopes

In this section, we prove Theorem A from the introduction. Our proof follows that of [Holm and Jørgensen 2011, Theorem 1.6] with some adjustments.

Lemma 3.1. *For every* R*-module* M*, the composition* $M^{\dagger} \xrightarrow{\delta_{M^{\dagger}}} M^{\dagger\dagger\dagger} \xrightarrow{\delta_M^{\dagger}} M^{\dagger}$ *is the identity map on* $M^{\dagger}$*.*

Proof. Straightforward; see [Jans 1961, Theorem 1.4]. □

Lemma 3.2. *For every finitely generated* R*-module* M*, the next conditions are equivalent.*

(i) $\operatorname{Ext}^1_R(M, \Omega) = 0$ *and* $\operatorname{Ext}^1_R(X, M^{\dagger}) = 0$ *for every* $X \in \mathsf{MCM}$.

(ii) $\operatorname{Ext}^1_R(M, Y) = 0$ *for every* $Y \in \mathsf{MCM}$.

Proof. (i)⟹(ii): Given any $Y \in \mathsf{MCM}$ we must argue that $\operatorname{Ext}^1_R(M, Y) = 0$, i.e., that every short exact sequence $0 \to Y \xrightarrow{\alpha} E \to M \to 0$ splits. As $\operatorname{Ext}^1_R(M, \Omega) = 0$, the functor $(-)^{\dagger}$ leaves this sequence exact; in fact, the induced short exact sequence

$$0 \longrightarrow M^{\dagger} \longrightarrow E^{\dagger} \xrightarrow{\alpha^{\dagger}} Y^{\dagger} \longrightarrow 0$$

splits as $Y^{\dagger}$ belongs to MCM and hence $\operatorname{Ext}^1_R(Y^{\dagger}, M^{\dagger}) = 0$ by assumption. Let $\beta : Y^{\dagger} \to E^{\dagger}$ be a right inverse of $\alpha^{\dagger}$. Then $\delta_Y^{-1}\beta^{\dagger}\delta_E : E \to Y$ is a left inverse of α since one has

$$\delta_Y^{-1}\beta^{\dagger}\delta_E\alpha = \delta_Y^{-1}\beta^{\dagger}\alpha^{\dagger\dagger}\delta_Y = \delta_Y^{-1}(\alpha^{\dagger}\beta)^{\dagger}\delta_Y = \delta_Y^{-1}1_{Y^{\dagger\dagger}}\delta_Y = 1_Y.$$

(ii)⟹(i): Assumption (ii) implies that $\operatorname{Ext}^1_R(M, \Omega) = 0$ since $\Omega \in \mathsf{MCM}$. Given $X \in \mathsf{MCM}$ we must show that $\operatorname{Ext}^1_R(X, M^{\dagger}) = 0$, i.e., that every short exact sequence $0 \to M^{\dagger} \xrightarrow{\alpha} E \to X \to 0$ splits. Since X is in MCM we in particular have $\operatorname{Ext}^1_R(X, \Omega) = 0$, so an application of the functor $(-)^{\dagger}$ yields another short exact sequence:

$$(*) \qquad 0 \longrightarrow X^{\dagger} \longrightarrow E^{\dagger} \xrightarrow{\alpha^{\dagger}} M^{\dagger\dagger} \longrightarrow 0.$$

As $X^{\dagger}$ belongs to MCM we have $\operatorname{Ext}^1_R(M, X^{\dagger}) = 0$, so the functor $\operatorname{Hom}_R(M, -)$ leaves the sequence $(*)$ exact. Surjectivity of $\operatorname{Hom}_R(M, \alpha^{\dagger})$ yields a homomorphism $\beta : M \to E^{\dagger}$ with $\alpha^{\dagger}\beta = \delta_M$. It follows that $\beta^{\dagger}\delta_E : E \to M^{\dagger}$ is a left inverse of α since one has $\beta^{\dagger}\delta_E\alpha = \beta^{\dagger}\alpha^{\dagger\dagger}\delta_{M^{\dagger}} = (\alpha^{\dagger}\beta)^{\dagger}\delta_{M^{\dagger}} = \delta_M^{\dagger}\delta_{M^{\dagger}} = 1_{M^{\dagger}}$, where the last equality follows from Lemma 3.1. □

Proof of Theorem A. We begin by proving the last assertion in the theorem. Let $\mu : M \to X$ be any special MCM-preenvelope of M. By assumption, we have $\mathrm{Ext}^1_R(\mathrm{Coker}\,\mu, Y) = 0$ for every $Y \in \mathsf{MCM}$. Hence Lemma 3.2 implies that $\mathrm{Ext}^1_R(Z, (\mathrm{Coker}\,\mu)^\dagger) = 0$ for every $Z \in \mathsf{MCM}$. By [Auslander and Buchweitz 1989, Theorem A], we can take a *hull of finite injective dimension* for the finitely generated module $(\mathrm{Coker}\,\mu)^\dagger$, that is, a short exact sequence

$$0 \longrightarrow (\mathrm{Coker}\,\mu)^\dagger \longrightarrow I \longrightarrow Z \longrightarrow 0,$$

where I has finite injective dimension and Z is maximal CM. This sequence splits since $\mathrm{Ext}^1_R(Z, (\mathrm{Coker}\,\mu)^\dagger) = 0$, and $(\mathrm{Coker}\,\mu)^\dagger$ is therefore a direct summand in I. Since I has finite injective dimension, so has $(\mathrm{Coker}\,\mu)^\dagger$.

To prove parts (a) and (b), let M be a finitely generated R-module and let $\pi : Z \to M^\dagger$ be a homomorphism with $Z \in \mathsf{MCM}$. We will show that if π is a special MCM-precover, respectively, an MCM-cover of $M^\dagger$ (recall that by the theorem by Auslander, Buchweitz, Takahashi and Yoshino from the introduction, special MCM-precovers always exist, and MCM-covers exist if R is henselian), then the homomorphism

$$\mu := \pi^\dagger \delta_M : M \longrightarrow Z^\dagger$$

is a special MCM-preenvelope, respectively, an MCM-envelope, of M.

First assume that π is a special MCM-precover. We begin by proving that μ is an MCM-preenvelope. Note that $Z^\dagger$ is in MCM by Remark 2.4. We must show that $\mathrm{Hom}_R(\mu, Y)$ is surjective for every $Y \in \mathsf{MCM}$. By Remark 2.4 every such Y has the form $Y \cong X^\dagger$ for some $X \in \mathsf{MCM}$ (namely for $X = Y^\dagger$), so it suffices to show that $\mathrm{Hom}_R(\mu, X^\dagger)$ is surjective for every $X \in \mathsf{MCM}$. By definition of μ, the homomorphism $\mathrm{Hom}_R(\mu, X^\dagger)$ is the composition of the maps

$$(*) \qquad \mathrm{Hom}_R(Z^\dagger, X^\dagger) \xrightarrow{\mathrm{Hom}_R(\pi^\dagger, X^\dagger)} \mathrm{Hom}_R(M^{\dagger\dagger}, X^\dagger) \xrightarrow{\mathrm{Hom}_R(\delta_M, X^\dagger)} \mathrm{Hom}_R(M, X^\dagger)\ .$$

Via the “swap” isomorphism, see [Christensen 2000, (A.2.9)], the homomorphisms in $(*)$ are identified with the ones in the top row of the following diagram:

$$(**) \qquad \begin{tikzcd}
\mathrm{Hom}_R(X, Z^{\dagger\dagger}) \arrow[r, "\mathrm{Hom}_R(X,\pi^{\dagger\dagger})"] & \mathrm{Hom}_R(X, M^{\dagger\dagger\dagger}) \arrow[r, "\mathrm{Hom}_R(X,\delta_M^\dagger)"] & \mathrm{Hom}_R(X, M^\dagger) \\
\mathrm{Hom}_R(X, Z) \arrow[u, "\mathrm{Hom}_R(X,\delta_Z)", "\cong"'] \arrow[r, two heads, "\mathrm{Hom}_R(X,\pi)"] & \mathrm{Hom}_R(X, M^\dagger) \arrow[u, "\mathrm{Hom}_R(X,\delta_{M^\dagger})"] \arrow[ur, equal] &
\end{tikzcd}$$

The left square in $(**)$ is commutative since the biduality homomorphism δ is natural, and the right triangle in $(**)$ is commutative by Lemma 3.1. The map δ_Z is an isomorphism since Z is in MCM; and $\mathrm{Hom}_R(X, \pi)$ is surjective as π is an MCM-precover and $X \in \mathsf{MCM}$. It follows that the composition of the maps in the top row of $(**)$, and therefore also the map $\operatorname{Hom}_R(\mu, X^\dagger)$, is surjective. Thus, μ is an MCM-preenvelope.

To see that μ is a special MCM-preenvelope, we must prove that $\operatorname{Ext}^1_R(\operatorname{Coker}\mu, Y)$ vanishes for every $Y \in \mathsf{MCM}$. As the functor $(-)^\dagger$ is left exact, $(\operatorname{Coker}\mu)^\dagger$ is isomorphic to $\operatorname{Ker}(\mu^\dagger)$. By definition we have $\mu^\dagger = \delta^\dagger_M \pi^{\dagger\dagger}$, and hence $\mu^\dagger$ fits into the commutative diagram:

$$\tag{$***$}\begin{tikzcd}
Z^{\dagger\dagger} \arrow[r, "\mu^\dagger"] \arrow[d, equal] & M^\dagger \\
Z^{\dagger\dagger} \arrow[r, "\pi^{\dagger\dagger}"] & M^{\dagger\dagger\dagger} \arrow[u, "\delta^\dagger_M"'] & & \text{(by Lemma 3.1)} \\
Z \arrow[u, "\delta_Z", "\cong"'] \arrow[r, "\pi"] & M^\dagger \arrow[u, "\delta_{M^\dagger}"'] \arrow[uu, bend right=60, "1_{M^\dagger}"']
\end{tikzcd}$$

It follows that $\mu^\dagger$ and π are isomorphic maps, and hence they also have isomorphic kernels, that is, $\operatorname{Ker}(\mu^\dagger) \cong \operatorname{Ker}\pi$. It follows that $(\operatorname{Coker}\mu)^\dagger \cong \operatorname{Ker}\pi$. Since π is a special MCM-precover, we now have

$$\operatorname{Ext}^1_R(X, (\operatorname{Coker}\mu)^\dagger) \cong \operatorname{Ext}^1_R(X, \operatorname{Ker}\pi) = 0$$

for every $X \in \mathsf{MCM}$. Thus, to see that $\operatorname{Ext}^1_R(\operatorname{Coker}\mu, Y) = 0$ for every $Y \in \mathsf{MCM}$, it suffices by Lemma 3.2 to prove that $\operatorname{Ext}^1_R(\operatorname{Coker}\mu, \Omega) = 0$. To this end, set $X = Z^\dagger \in \mathsf{MCM}$ and consider the factorization of $\mu : M \to Z^\dagger = X$ given by

$$\begin{tikzcd}
M \arrow[rr, "\mu"] \arrow[dr, two heads, "\mu_0"'] & & X \\
& \operatorname{Im}\mu \arrow[ur, hook, "\iota"'] &
\end{tikzcd}$$

where μ_0 is the corestriction of μ to its image and ι is the inclusion map. As μ_0 is surjective and $(-)^\dagger$ is left exact, the map $\mu_0^\dagger$ is injective. As $\Omega \in \mathsf{MCM}$ and μ is an MCM-preenvelope, the map $\mu^\dagger = \operatorname{Hom}_R(\mu, \Omega)$ is surjective; and hence so is $\mu_0^\dagger$ since $\mu^\dagger = \mu_0^\dagger \iota^\dagger$. Thus, $\mu_0^\dagger$ is an isomorphism. Hence $\iota^\dagger$ and $\mu^\dagger$ are isomorphic maps, and since $\mu^\dagger$ is surjective, so is $\iota^\dagger$. Thus, application of $(-)^\dagger$ to $0 \to \operatorname{Im}\mu \xrightarrow{\iota} X \to \operatorname{Coker}\mu \to 0$ yields an exact sequence

$$X^\dagger \xrightarrow{\iota^\dagger} (\operatorname{Im}\mu)^\dagger \xrightarrow{0} \operatorname{Ext}^1_R(\operatorname{Coker}\mu, \Omega) \longrightarrow \operatorname{Ext}^1_R(X, \Omega) = 0,$$

which forces $\operatorname{Ext}^1_R(\operatorname{Coker}\mu, \Omega) = 0$, as desired.

Finally, assume that π is an MCM-cover. We show that $\mu = \pi^\dagger \delta_M$ is an MCM-envelope. We have already seen that μ is an MCM-preenvelope. To show that it is an envelope, let φ be an endomorphism of $Z^\dagger$ with $\varphi\mu = \mu$. It follows that $\mu^\dagger \varphi^\dagger = \mu^\dagger$. The diagram $(***)$ shows that $\mu^\dagger \delta_Z = \pi$, and thus $\pi(\delta_Z^{-1} \varphi^\dagger \delta_Z) =$

$\mu^\dagger\varphi^\dagger\delta_Z = \mu^\dagger\delta_Z = \pi$. As π is an MCM-cover, we conclude that $\delta_Z^{-1}\varphi^\dagger\delta_Z$, and therefore also $\varphi^\dagger$, is an automorphism. It follows that $\varphi^{\dagger\dagger}$ is an automorphism of $Z^{\dagger\dagger\dagger}$, and finally that $\varphi = \delta_{Z^\dagger}^{-1}\varphi^{\dagger\dagger}\delta_{Z^\dagger}$ is an automorphism of $Z^\dagger$. □

The proof of Theorem A (above) shows that one can construct MCM-envelopes from MCM-covers. We do not know if the converse is true, that is, we do not know if existence of MCM-envelopes is logically equivalent to existence of MCM-covers. The next result provides a partial answer to this question; it shows that existence of MCM-envelopes for *all* finitely generated modules implies existence of MCM-covers for *some* finitely generated modules (namely for modules N of the form $N \cong M^\dagger$ for some M).

Proposition 3.3. *Let M be a finitely generated R-module. If $\mu : M \to X$ is an* MCM*-preenvelope, a special* MCM*-preenvelope, or an* MCM*-envelope of M, then $\mu^\dagger : X^\dagger \to M^\dagger$ is an* MCM*-precover, a special* MCM*-precover, or an* MCM*-cover of $M^\dagger$, respectively.*

Proof. This is left as an exercise to the reader. □

4. Examples

We compute the MCM-envelope of some specific modules. We begin with a characterization of modules with trivial MCM-envelope.

Proposition 4.1. *For a finitely generated R-module M, one has $\dim_R M < d$ if and only if the zero map $M \to 0$ is an* MCM*-envelope of M.*

Proof. If $\dim_R M < d$ then [Bruns and Herzog 1993, Corollary 3.5.11(a)] shows that $\operatorname{Hom}_R(M, \Omega) = 0$. It follows that every homomorphism $\varphi : M \to X$ with $X \in$ MCM is zero. Indeed, since Ω cogenerates the category MCM, there exists a monomorphism $\iota : X \to \Omega^n$ for some natural number n. As $\operatorname{Hom}_R(M, \Omega) = 0$, the homomorphism $\iota\varphi : M \to \Omega^n$ must be zero, and thus $\varphi = 0$ since ι is injective. Since every homomorphism from M to a maximal CM module is zero, the zero map $M \to 0$ is an MCM-envelope of M.

Conversely, if $M \to 0$ is an MCM-(pre)envelope then, since Ω is in MCM, every homomorphism $\varphi : M \to \Omega$ factors through 0, and hence $\varphi = 0$. Thus $\operatorname{Hom}_R(M, \Omega) = 0$, and it follows from [Bruns and Herzog 1993, Corollary 3.5.11(b)] that one can not have $\dim_R M = d$; so $\dim_R M < d$. □

In general, MCM-(pre)envelopes need not be injective. In fact:

Corollary 4.2. *The ring R is artinian if and only if every finitely generated R-module admits an injective* (*that is, monic*) MCM-(*pre*)*envelope.*

Proof. If R is artinian, then every finitely generated R-module M is maximal CM, and therefore $1_M : M \to M$ is an injective MCM-envelope of M. Conversely, if R is not artinian, then the residue field k, which has dimension $\dim_R k = 0$, does not have an injective MCM-preenvelope by Proposition 4.1. □

Next we give a somewhat "general" example.

Example 4.3. Let M be a finitely generated R-module. If $M^\dagger$ is maximal CM, then the identity homomorphism $\pi = 1_{M^\dagger} : M^\dagger \to M^\dagger$ is an MCM-cover of $M^\dagger$. The proof of Theorem A shows that the homomorphism $\mu = \pi^\dagger \delta_M = \delta_M$, i.e., the biduality homomorphism $\delta_M : M \to M^{\dagger\dagger}$, is an MCM-envelope M.

Here is a concrete application of the example above.

Example 4.4. Let M be a submodule of a maximal CM R-module X with the property that $\dim_R(X/M) < d-1$. For example, $M = \mathfrak{a}$ could be an ideal in $X = R$ with $\mathrm{height}_R(\mathfrak{a}) > 1$; see [Bruns and Herzog 1993, Corollary 2.1.4]. Or M could be the submodule $M = (f_1, f_2, \dots)X$, where $f_1, f_2, \dots$ is an X-regular sequence of length at least two. We claim that, in this case, the inclusion map $\iota : M \hookrightarrow X$ is an MCM-envelope of M.

To see why, note that the short exact sequence $0 \to M \xrightarrow{\iota} X \to X/M \to 0$ is mapped by the functor $(-)^\dagger$ to the exact sequence

$$0 \longrightarrow (X/M)^\dagger \longrightarrow X^\dagger \xrightarrow{\iota^\dagger} M^\dagger \longrightarrow \mathrm{Ext}^1_R(X/M, \Omega).$$

Since $d - \dim_R(X/M) > 1$, it follows from Corollary 3.5.11(a) of [Bruns and Herzog 1993] that $\mathrm{Hom}_R(X/M, \Omega) = 0$ and $\mathrm{Ext}^1_R(X/M, \Omega) = 0$. Hence the sequence displayed above shows that $\iota^\dagger$ is an isomorphism and, in particular, $M^\dagger \cong X^\dagger$ is maximal CM. Thus Example 4.3 shows that the biduality homomorphism $\delta_M : M \to M^{\dagger\dagger}$ is an MCM-envelope of M. It remains to argue that δ_M can be identified with $\iota : M \hookrightarrow X$; however, this follows from the commutative diagram:

$$\begin{tikzcd}
M \arrow[r, "\iota"] \arrow[d, "\delta_M"'] & X \arrow[d, "\delta_X", "\cong"'] \\
M^{\dagger\dagger} \arrow[r, "\iota^{\dagger\dagger}", "\cong"'] & X^{\dagger\dagger}
\end{tikzcd}$$

Indeed, δ_X is an isomorphism as $X \in$ MCM, and $\iota^{\dagger\dagger} = (\iota^\dagger)^\dagger$ is an isomorphism because $\iota^\dagger$ is.

Remark 4.5. For a special MCM-precover $\pi : X \to M$ of a finitely generated module M, the kernel $\mathrm{Ker}\,\pi$ has finite injective dimension, and hence one has $\mathrm{Ext}^i_R(X, \mathrm{Ker}\,\pi) = 0$ for every $X \in$ MCM and every $i > 0$—not just for $i = 1$. A similar phenomenon does not occur for special MCM-preenvelopes. Indeed, if in Example 4.4 one has $\dim_R(X/M) = d-2$, say, then $\operatorname{Coker}\iota = X/M$ satisfies $\operatorname{Ext}^2_R(X/M, \Omega) \neq 0$ by [Bruns and Herzog 1993, Corollary 3.5.11(b)].

5. MCM-envelopes with the unique lifting property

If $\mu : M \to X$ is an MCM-preenvelope of a finitely generated R-module M, then the induced homomorphism $\operatorname{Hom}_R(\mu, Y) : \operatorname{Hom}_R(X, Y) \to \operatorname{Hom}_R(M, Y)$ is surjective for every $Y \in \mathsf{MCM}$; see Definition 2.2. If $\operatorname{Hom}_R(\mu, Y)$ is an isomorphism for every $Y \in \mathsf{MCM}$, then we say that the MCM-preenvelope μ has the *unique lifting property*. Indeed, in this case, there exists for every homomorphism $\nu : M \to Y$ with $Y \in \mathsf{MCM}$ a unique homomorphism $\varphi : X \to Y$ that makes the following diagram commute:

$$\begin{tikzcd} M \arrow[r, "\mu"] \arrow[d, "\nu"'] & X \arrow[dl, dashed, "\varphi"] \\ Y & \end{tikzcd}$$

Note that an MCM-preenvelope $\mu : M \to X$ with the unique lifting property must necessarily be an MCM-envelope. Indeed, the only endomorphism φ of X with $\varphi\mu = \mu$ is $\varphi = 1_X$. Evidently, every surjective MCM-preenvelope has the unique lifting property.

Lemma 5.1. *For any finitely generated R-module M, one has* $\operatorname{depth}_R(M^\dagger) \geqslant \min\{d, 2\}$.

Proof. Take an exact sequence $L_1 \to L_0 \to M \to 0$ where L_0 and L_1 are finitely generated and free. Since the functor $(-)^\dagger = \operatorname{Hom}_R(-, \Omega)$ is left exact, we get an exact sequence, $0 \to M^\dagger \to L_0^\dagger \to L_1^\dagger \to C \to 0$, where C is the cokernel of the homomorphism $L_0^\dagger \to L_1^\dagger$. Since the modules $L_0^\dagger$ and $L_1^\dagger$ are maximal CM, Lemma 2.5 yields

$$\operatorname{depth}_R(M^\dagger) \geqslant \min\{d, \operatorname{depth}_R C + 2\} \geqslant \min\{d, 2\}. \qquad \square$$

Proof of Theorem B. (i)$\Longrightarrow$(ii): Consider an exact sequence of finitely generated modules

$$0 \longrightarrow K \longrightarrow L_1 \xrightarrow{\alpha} L_0 \longrightarrow N \longrightarrow 0,$$

where L_0 and L_1 are free and $K = \operatorname{Ker}\alpha$. From [Bruns and Herzog 1993, Proposition 1.2.9] (last inequality) one gets

$$(*) \qquad \operatorname{depth}_R N \geqslant \operatorname{depth}_R K - 2.$$

Set $C = \operatorname{Coker}(\alpha^\dagger)$ and consider the exact sequence $L_0^\dagger \xrightarrow{\alpha^\dagger} L_1^\dagger \longrightarrow C \longrightarrow 0$. As the functor $(-)^\dagger$ is left exact, we get a commutative diagram with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r] & L_1 \arrow[r, "\alpha"] \arrow[d, "\cong"', "\delta_{L_1}"] & L_0 \arrow[d, "\cong"', "\delta_{L_0}"] \\
0 \arrow[r] & C^\dagger \arrow[r] & L_1^{\dagger\dagger} \arrow[r, "\alpha^{\dagger\dagger}"] & L_0^{\dagger\dagger}
\end{tikzcd}$$

which shows that $K \cong C^\dagger$, since δ_{L_0} and δ_{L_1} are isomorphisms. By assumption (i), the module K is therefore maximal CM, and hence inequality $(*)$ yields $\operatorname{depth}_R N \geqslant d-2$. As this holds for every finitely generated R-module N, it holds in particular for the residue field $N = k$. We get $0 = \operatorname{depth}_R k \geqslant d-2$, and thus $d \leqslant 2$.

(ii)$\Longrightarrow$(iii): In the case where R is reduced, a proof of this implication can be found in [Burban and Drozd 2008, Proposition 3.2]. We give a slightly different argument.

If $d \leqslant 2$, then Lemma 5.1 shows that for every finitely generated R-module M, the module $M^\dagger$ is maximal CM, and hence so is $M^{\dagger\dagger}$. Thus $\mathrm{F} = (-)^{\dagger\dagger}$ is a functor from mod to MCM, which we claim is a left adjoint of the inclusion $\mathrm{G} : \mathsf{MCM} \to \mathsf{mod}$. For each finitely generated R-module M and each maximal CM R-module X, the homomorphism $\varphi_{M,X} = \operatorname{Hom}_R(\delta_M, X)$ given by

$$\operatorname{Hom}_R(\mathrm{F}M, X) = \operatorname{Hom}_R(M^{\dagger\dagger}, X) \xrightarrow{\varphi_{M,X}} \operatorname{Hom}_R(M, X) = \operatorname{Hom}_R(M, \mathrm{G}X)$$

is evidently natural in M and X; and it is surjective since the biduality map $\delta_M : M \to M^{\dagger\dagger}$ is an MCM-preenvelope of M by Example 4.3. It remains to see that $\operatorname{Hom}_R(\delta_M, X)$ is injective. To this end, let $\mu : M^{\dagger\dagger} \to X$ be a homomorphism with $\mu\delta_M = \operatorname{Hom}_R(\delta_M, X)(\mu) = 0$. It follows that $\delta_M^\dagger \mu^\dagger = (\mu\delta_M)^\dagger = 0$. As $M^\dagger$ is maximal CM, the biduality map $\delta_{M^\dagger}$ is an isomorphism, and hence so is $\delta_M^\dagger$ by Lemma 3.1. Since $\delta_M^\dagger \mu^\dagger = 0$ we conclude that $\mu^\dagger = 0$. Thus $\mu^{\dagger\dagger} = (\mu^\dagger)^\dagger = 0$ and consequently $\mu = \delta_X^{-1} \mu^{\dagger\dagger} \delta_{M^{\dagger\dagger}} = 0$, as desired.

(iii)$\Longrightarrow$(iv): Let $\mathrm{F} : \mathsf{mod} \to \mathsf{MCM}$ be a left adjoint of the inclusion $\mathrm{G} : \mathsf{MCM} \to \mathsf{mod}$. For every finitely generated R-module M, the unit of adjunction $\eta_M : M \to \mathrm{GF}M$ induces, for every maximal CM R-module Y, an isomorphism:

$$\varphi_{M,Y} : \operatorname{Hom}_R(\mathrm{F}M, Y) \xrightarrow{\sim} \operatorname{Hom}_R(M, \mathrm{G}Y) \quad \text{given by} \quad \alpha \mapsto \mathrm{G}(\alpha)\eta_M;$$

see [MacLane 1971, IV.1 Theorem 1]. If we suppress the inclusion functor G and set $X = \mathrm{GF}M = \mathrm{F}M$, which is maximal CM by the assumption on F, we see that unit of adjunction $\eta_M : M \to X$ has the property that the map

$$\operatorname{Hom}_R(X, Y) \xrightarrow{\sim} \operatorname{Hom}_R(M, Y) \quad \text{given by} \quad \alpha \mapsto \alpha\eta_M = \operatorname{Hom}_R(\eta_M, Y)(\alpha)$$

is an isomorphism. Thus, η_M is an MCM-envelope of M with the unique lifting property.

(iv)$\Longrightarrow$(i): Let M be a finitely generated R-module. By assumption, M has an MCM-envelope $\mu : M \to X$ with the unique lifting property. Since Ω is maximal CM, the homomorphism $\mu^\dagger : X^\dagger \to M^\dagger$ is an isomorphism, and as $X^\dagger$ is maximal CM, so is $M^\dagger$. ☐

6. Cosyzygies with respect to MCM

Let $\mathcal{A}$ be a full subcategory of an abelian category $\mathcal{M}$ (for example, $\mathcal{M} = \mathsf{mod}$ and $\mathcal{A} = \mathsf{MCM}$).

Assume that every object in $\mathcal{M}$ has an $\mathcal{A}$-precover. In this case, every $M \in \mathcal{M}$ admits a *proper $\mathcal{A}$-resolution*, meaning a, not necessarily exact, complex $\mathbb{A} = \cdots \to A_1 \to A_0 \to M \to 0$ with $A_i \in \mathcal{A}$ such that the sequence $\operatorname{Hom}_{\mathcal{M}}(A, \mathbb{A})$ is exact for every $A \in \mathcal{A}$. Such a resolution is constructed recursively as follows: take an $\mathcal{A}$-precover $\pi_0 : A_0 \to M$ of M and set $K_1 = \operatorname{Ker} \pi_0$; take an $\mathcal{A}$-precover $\pi_1 : A_1 \to K_1$ of K_1 and set $K_2 = \operatorname{Ker} \pi_1$; etc. The object K_n is denoted by $\operatorname{Syz}_n^{\mathcal{A}}(M)$ and it is called the *n-th syzygy of M with respect to $\mathcal{A}$*. A given object $M \in \mathcal{M}$ has, typically, many different $\mathcal{A}$-precovers and proper $\mathcal{A}$-resolutions, so $\operatorname{Syz}_n^{\mathcal{A}}(M)$ is not uniquely determined by M; but it almost is: the version of Schanuel's lemma found in [Enochs et al. 2001, Lemma 2.2] shows that if K_n and $\overline{K}_n$ are both n-th syzygies of M with respect to $\mathcal{A}$, then there exist $A, \overline{A} \in \mathcal{A}$ such that $K_n \oplus \overline{A} \cong \overline{K}_n \oplus A$. In particular, if $\mathcal{A}$ is closed under direct summands (as is the case if $\mathcal{A} = \mathsf{MCM}$), then K_n belongs to $\mathcal{A}$ if and only if $\overline{K}_n$ belongs to $\mathcal{A}$; and thus it makes sense to ask if $\operatorname{Syz}_n^{\mathcal{A}}(M)$ belongs to $\mathcal{A}$.

If every object in $\mathcal{M}$ admits an $\mathcal{A}$-cover, then $\pi_0, \pi_1, \ldots$ in the construction above can be chosen to be $\mathcal{A}$-covers, and the resulting proper $\mathcal{A}$-resolution is then called a *minimal proper $\mathcal{A}$-resolution* of M. In this case, K_n is called the *minimal n-th syzygy of M with respect to $\mathcal{A}$*, and it is denoted by $\text{min-}\operatorname{Syz}_n^{\mathcal{A}}(M)$. Since an $\mathcal{A}$-cover (of a given object in $\mathcal{M}$) is unique up to isomorphism, see [Xu 1996, Theorem 1.2.6], the object $\text{min-}\operatorname{Syz}_n^{\mathcal{A}}(M)$ is uniquely determined, up to isomorphism, by M.

Dually, if every $M \in \mathcal{M}$ has an $\mathcal{A}$-preenvelope (resp. $\mathcal{A}$-envelope), then a *proper $\mathcal{A}$-coresolution* (resp. *minimal proper $\mathcal{A}$-coresolution*) $0 \to M \to A^0 \to A^1 \to \cdots$ can always be constructed as follows: take an $\mathcal{A}$-preenvelope (resp. $\mathcal{A}$-envelope) $\mu^0 : M \to A^0$ of M and set $C^1 = \operatorname{Coker} \mu^0$; take an $\mathcal{A}$-preenvelope (resp. $\mathcal{A}$-envelope) $\mu^1 : C^1 \to A^1$ of C^1 and set $C^2 = \operatorname{Coker} \mu^1$; etc. The object C^n is called the *n-th cosyzygy of M with respect to $\mathcal{A}$* (resp. the *minimal n-th cosyzygy of M with respect to $\mathcal{A}$*) and it is denoted by $\operatorname{Cosyz}_{\mathcal{A}}^n(M)$ (resp. $\text{min-}\operatorname{Cosyz}_{\mathcal{A}}^n(M)$). The object $\text{min-}\operatorname{Cosyz}_{\mathcal{A}}^n(M)$ is uniquely determined, up to isomorphism, by M. The object $\operatorname{Cosyz}_{\mathcal{A}}^n(M)$ is almost uniquely determined by M in the sense that if C^n and $\overline{C}^n$ are both n-th cosyzygies of M with respect to $\mathcal{A}$, then there exist $A, \overline{A} \in \mathcal{A}$ such that $C^n \oplus \overline{A} \cong \overline{C}^n \oplus A$. Thus, if $\mathcal{A}$ is closed under direct summands, then it makes sense to ask if $\operatorname{Cosyz}_{\mathcal{A}}^n(M)$ belongs to $\mathcal{A}$.

We supplement the definitions above by setting $\operatorname{Syz}_0^A(M) = \text{min-}\operatorname{Syz}_0^A(M) = M$, and similarly $\operatorname{Cosyz}_{\mathcal{A}}^0(M) = \text{min-}\operatorname{Cosyz}_{\mathcal{A}}^0(M) = M$.

Example 6.1. Let $(A, \mathfrak{n}, \ell)$ be any local ring and let $\mathcal{F}$ be the class of finitely generated free A-modules. Every finitely generated A-module M has an $\mathcal{F}$-cover; to construct it one takes a minimal set $x_1, \dots, x_b$ of generators of M (here $b = \beta_0^A(M)$ is the zeroth Betti number of M) and then defines $A^b \twoheadrightarrow M$ by $e_i \mapsto x_i$; see [Enochs and Jenda 2000, Theorem 5.3.3]. A minimal proper $\mathcal{F}$-resolution $\cdots \to F_1 \to F_0 \to M \to 0$ of a finitely generated A-module M is nothing but a *minimal free resolution* of M in the classical sense, that is, where each homomorphism $F_n \to F_{n-1}$ becomes zero when tensored with the residue field ℓ of A.

In this section, we are interested in cosyzygies with respect to the class MCM of maximal CM R-modules. We begin with a characterization of modules for which the first such cosyzygy is maximal CM.

Proposition 6.2. *For a finitely generated R-module M the next conditions are equivalent*:

(i) *M has an* MCM*-preenvelope whose cokernel is maximal CM*, *meaning that* $\operatorname{Cosyz}_{\mathsf{MCM}}^1(M)$ *is a maximal CM module.*

(ii) *M has a surjective* MCM*-envelope*, *that is*, $\text{min-}\operatorname{Cosyz}_{\mathsf{MCM}}^1(M) = 0$.

Proof. Evidently, (ii) implies (i). Conversely, let $\mu : M \to X$ be an MCM-preenvelope such that $C = \operatorname{Coker} \mu$ is maximal CM. Since X and $C = X/\operatorname{Im} \mu$ are maximal CM, so is $\operatorname{Im} \mu$. It follows that the corestriction $\mu : M \twoheadrightarrow \operatorname{Im} \mu$ is a surjective MCM-envelope of M. □

Next we give a sufficient condition for the second cosyzygy to be maximal CM.

Proposition 6.3. *Let M be a finitely generated R-module such that $M^\dagger$ is maximal CM. Then any second cosyzygy*, $\operatorname{Cosyz}_{\mathsf{MCM}}^2(M)$, *of M with respect to* MCM *is maximal CM.*

Proof. By Example 4.3, the homomorphism $\delta_M : M \to M^{\dagger\dagger}$ is an MCM-envelope of M. Set $C^1 = \text{min-}\operatorname{Cosyz}_{\mathsf{MCM}}^1(M) = \operatorname{Coker} \delta_M$. By application of the left exact functor $(-)^\dagger$, the exact sequence $M \xrightarrow{\delta_M} M^{\dagger\dagger} \longrightarrow C^1 \longrightarrow 0$ induces an exact sequence

$$0 \longrightarrow (C^1)^\dagger \longrightarrow M^{\dagger\dagger\dagger} \xrightarrow{\delta_M^\dagger} M^\dagger.$$

As $M^\dagger$ is maximal CM, the biduality homomorphism $\delta_{M^\dagger}$ is an isomorphism, and hence so is $\delta_M^\dagger$ by Lemma 3.1. It follows that $\operatorname{Hom}_R(C^1, \Omega) = (C^1)^\dagger = 0$, so [Bruns and Herzog 1993, Corollary 3.5.11(b)] implies that $\dim_R(C^1) < d$. Thus

Proposition 4.1 shows that $C^1 \to 0$ is an MCM-envelope of C^1, and therefore the minimal second cosyzygy of M with respect to MCM is zero:

$$\text{min-Cosyz}^2_{\mathsf{MCM}}(M) = \text{min-Cosyz}^1_{\mathsf{MCM}}(C^1) = \mathrm{Coker}(C^1 \to 0) = 0.$$

Hence any second cosyzygy of M with respect to MCM must be maximal CM. □

Proof of Theorem C. First note, that if X is a maximal CM R-module, then $\mathrm{Cosyz}^i_{\mathsf{MCM}}(X)$ is clearly maximal CM for every $i \geqslant 0$. If $n \geqslant d$, then the n-th cosyzygy of M is an $(n-d)^{\text{th}}$ cosyzygy of $\mathrm{Cosyz}^d_{\mathsf{MCM}}(M)$, that is,

$$\mathrm{Cosyz}^n_{\mathsf{MCM}}(M) = \mathrm{Cosyz}^{n-d}_{\mathsf{MCM}}(\mathrm{Cosyz}^d_{\mathsf{MCM}}(M));$$

so it suffices to argue that $\mathrm{Cosyz}^d_{\mathsf{MCM}}(M)$ is maximal CM.

If $d = 0$, then certainly $\mathrm{Cosyz}^0_{\mathsf{MCM}}(M) = M$ is maximal CM, since every finitely generated R-module is maximal CM over an artinian ring.

Assume that $d = 1$. By Theorem A we can take a special MCM-preenvelope $\mu : M \to X$ of M. We must show that $C^1 = \mathrm{Cosyz}^1_{\mathsf{MCM}}(M) = \mathrm{Coker}\,\mu$ is maximal CM. By definition, we have $\mathrm{Ext}^1_R(C^1, Y) = 0$ for all $Y \in \mathsf{MCM}$, in particular, $\mathrm{Ext}^1_R(C^1, \Omega) = 0$. Since Ω has injective dimension $d = 1$, we also have $\mathrm{Ext}^i_R(-, \Omega) = 0$ for all $i > 1$, and consequently, $\mathrm{Ext}^i_R(C^1, \Omega) = 0$ for all $i > 0$. Thus C^1 is maximal CM.

Finally, assume that $d \geqslant 2$. Let $0 \to M \to X^0 \to \cdots \to X^{d-3} \to C^{d-2} \to 0$ be part of a proper MCM-coresolution of M, where $C^{d-2} = \mathrm{Cosyz}^{d-2}_{\mathsf{MCM}}(M)$. In the case $d = 2$, this just means that we consider the module $C^0 = \mathrm{Cosyz}^0_{\mathsf{MCM}}(M) = M$. Since the module Ω is maximal CM, the sequence

$$0 \longrightarrow (C^{d-2})^\dagger \longrightarrow (X^{d-3})^\dagger \longrightarrow \cdots \longrightarrow (X^0)^\dagger \longrightarrow M^\dagger \longrightarrow 0$$

is exact. From Lemma 2.5 and Lemma 5.1 we derive that $\mathrm{depth}_R (C^{d-2})^\dagger \geqslant \min\{d, \mathrm{depth}_R M^\dagger + d - 2\} = d$, so $(C^{d-2})^\dagger = (\mathrm{Cosyz}^{d-2}_{\mathsf{MCM}}(M))^\dagger$ is maximal CM. Proposition 6.3 now yields that

$$\mathrm{Cosyz}^d_{\mathsf{MCM}}(M) = \mathrm{Cosyz}^2_{\mathsf{MCM}}(\mathrm{Cosyz}^{d-2}_{\mathsf{MCM}}(M))$$

is maximal CM, as desired. □

Dutta [1989] shows that if R is not regular, then no syzygy in the minimal free resolution of the residue field k (see Example 6.1) can contain a nonzero free direct summand. The following result has the same flavor, but its proof is easy. Actually, the proof of [Takahashi 2006, Proposition 2.6] applies to prove Proposition 6.4 as well, but since it is so short, we repeat it here.

Proposition 6.4. *Assume that every finitely generated R-module has an* MCM-*envelope* (*by Theorem A*, *this is the case if R is henselian*). *Let M be a finitely generated R-module and let $n \geqslant 1$ be an integer. The minimal n-th cosyzygy*,

$\text{min-Cosyz}^n_{\mathsf{MCM}}(M)$, *of* M *with respect to* MCM *contains no nonzero free direct summand.*

Proof. It suffices to consider the case $n = 1$. Let $\mu : M \to X$ be an MCM-envelope of M, set $C = \text{min-Cosyz}^1_{\mathsf{MCM}}(M) = \operatorname{Coker} \mu$, and write $\pi : X \twoheadrightarrow C$ for the canonical homomorphism. Let F be a free direct summand in C and denote by $\rho : C \twoheadrightarrow F$ the projection onto this direct summand. We have a commutative diagram

$$\begin{tikzcd}
 & M \arrow[r, "\mu"] \arrow[d, "\mu_0"] & X \arrow[r, "\pi"] \arrow[d, equal] & C \arrow[r] \arrow[d, two heads, "\rho"] & 0 \\
0 \arrow[r] & K \arrow[r, "\iota"] & X \arrow[r, "\rho\pi"] & F \arrow[r] & 0,
\end{tikzcd}$$

where $\iota : K \to X$ is the kernel of $\rho\pi$, and μ_0 is the corestriction of μ to K. Since F is free, the lower short exact sequence splits, so ι has a left inverse $\sigma : X \to K$. The endomorphism $\iota\sigma$ of X satisfies $\iota\sigma\mu = \iota\sigma\iota\mu_0 = \iota\mu_0 = \mu$, and since μ is an envelope, we conclude that $\iota\sigma$ is an automorphism. In particular, ι is surjective, and hence F is zero. □

Acknowledgement

It is a pleasure to thank the anonymous referee for a careful reading of this manuscript and for several thoughtful comments and suggestions. We are also grateful for the referee's questions that led to Proposition 3.3, Corollary 4.2, and Proposition 6.2.

References

[Auslander and Buchweitz 1989] M. Auslander and R.-O. Buchweitz, "The homological theory of maximal Cohen–Macaulay approximations", 38 (1989), 5–37. MR 91h:13010 Zbl 0697.13005

[Auslander and Reiten 1991] M. Auslander and I. Reiten, "Applications of contravariantly finite subcategories", *Adv. Math.* **86**:1 (1991), 111–152. MR 92e:16009 Zbl 0774.16006

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge Univ. Press, 1993. MR 95h:13020 Zbl 0788.13005

[Burban and Drozd 2008] I. Burban and Y. Drozd, "Maximal Cohen–Macaulay modules over surface singularities", pp. 101–166 in *Trends in representation theory of algebras and related topics*, edited by A. Skowroński, Eur. Math. Soc., Zürich, 2008. MR 2010a:13017 Zbl 1200.14011

[Christensen 2000] L. W. Christensen, *Gorenstein dimensions*, Lecture Notes in Mathematics **1747**, Springer, Berlin, 2000. MR 2002e:13032 Zbl 0965.13010

[Crawley-Boevey 1994] W. Crawley-Boevey, "Locally finitely presented additive categories", *Comm. Algebra* **22**:5 (1994), 1641–1674. MR 95h:18009 Zbl 0798.18006

[Dutta 1989] S. P. Dutta, "Syzygies and homological conjectures", pp. 139–156 in *Commutative algebra* (Berkeley, CA, 1987), edited by M. Hochster et al., Math. Sci. Res. Inst. Publ. **15**, Springer, New York, 1989. MR 90i:13015 Zbl 0733.13006

[Enochs and Jenda 2000] E. E. Enochs and O. M. G. Jenda, *Relative homological algebra*, de Gruyter Expositions in Mathematics **30**, Walter de Gruyter, Berlin, 2000. MR 2001h:16013 Zbl 0952.13001

[Enochs et al. 2001] E. E. Enochs, O. M. G. Jenda, and L. Oyonarte, "λ and μ-dimensions of modules", *Rend. Sem. Mat. Univ. Padova* **105** (2001), 111–123. MR 2002c:16012 Zbl 1072.16011

[Holm 2014] H. Holm, "The structure of balanced big Cohen–Macaulay modules over Cohen–Macaulay rings", preprint, 2014. arXiv 1408.5152v1

[Holm and Jørgensen 2011] H. Holm and P. Jørgensen, "Rings without a Gorenstein analogue of the Govorov–Lazard theorem", *Q. J. Math.* **62**:4 (2011), 977–988. MR 2012k:13031 Zbl 1251.13008

[Ischebeck 1969] F. Ischebeck, "Eine Dualität zwischen den Funktoren Ext und Tor", *J. Algebra* **11**:4 (1969), 510–531. MR 38 #5894 Zbl 0191.01306

[Jans 1961] J. P. Jans, "Duality in Noetherian rings", *Proc. Amer. Math. Soc.* **12** (1961), 829–835. MR 25 #1192 Zbl 0113.26104

[MacLane 1971] S. MacLane, *Categories for the working mathematician*, Graduate Texts in Mathematics **5**, Springer, New York, 1971. MR 50 #7275 Zbl 0232.18001

[Takahashi 2005] R. Takahashi, "On the category of modules of Gorenstein dimension zero", *Math. Z.* **251**:2 (2005), 249–256. MR 2006j:13012 Zbl 1098.13014

[Takahashi 2006] R. Takahashi, "Remarks on modules approximated by G-projective modules", *J. Algebra* **301**:2 (2006), 748–780. MR 2007a:13010 Zbl 1109.13012

[Wakamatsu 1988] T. Wakamatsu, "On modules with trivial self-extensions", *J. Algebra* **114**:1 (1988), 106–114. MR 89b:16020 Zbl 0646.16025

[Xu 1996] J. Xu, *Flat covers of modules*, Lecture Notes in Mathematics **1634**, Springer, Berlin, 1996. MR 98b:16003 Zbl 0860.16002

[Yoshino 1993] Y. Yoshino, "Cohen–Macaulay approximations", pp. 119–138 in *Proceedings of the 4th Symposium on Representation Theory of Algebras* (Izu, Japan, 1993), 1993. In Japanese.

Received October 21, 2014. Revised February 17, 2015.

HENRIK HOLM
UNIVERSITY OF COPENHAGEN
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITETSPARKEN 5
2100 COPENHAGEN Ø
DENMARK

holm@math.ku.dk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.371

PATTERSON–SULLIVAN CURRENTS, GENERIC STRETCHING FACTORS AND THE ASYMMETRIC LIPSCHITZ METRIC FOR OUTER SPACE

ILYA KAPOVICH AND MARTIN LUSTIG

We quantitatively relate the Patterson–Sullivan currents and generic stretching factors for free group automorphisms to the asymmetric Lipschitz metric on outer space and to Guirardel's intersection number. Thus we show that, given $N \geq 2$ and $\varepsilon > 0$, there exists a constant $c = c(N, \varepsilon) > 0$ such that for any two trees $T, S \in \mathrm{cv}_N$ of covolume 1 and injectivity radius $\geq \varepsilon$, we have

$$|\log\langle S, \mu_T\rangle - d_L(T, S)| \leq c,$$

where d_L is the asymmetric Lipschitz metric on the Culler–Vogtmann outer space, and where μ_T is the (appropriately normalized) Patterson–Sullivan current corresponding to T. As a corollary, we show there exist constants $C_1 \geq 1$ and $C_2 \geq 1$ (depending on N, ε) such that for any T, S as above we have

$$\frac{1}{C_1} \log i_c(T, S) - C_2 \leq \log\langle S, \mu_T\rangle \leq C_1 \log i_c(T, S) + C_2,$$

where i_c is the combinatorial version of Guirardel's intersection number. We apply these results to the properties of generic stretching factors of free group automorphisms. In particular, we show that for any $N \geq 2$, there exists a constant $0 < \rho_N < 1$ such that for every automorphism φ of $F_N = F(A)$, we have

$$0 < \rho_N \leq \frac{\lambda_A(\varphi)}{\Lambda_A(\varphi)} \leq 1.$$

Here λ_A is the generic stretching factor of φ with respect to the free basis A of F_N and $\Lambda_A(\varphi)$ is the extremal stretching factor of φ with respect to A.

Kapovich was supported by the Collaboration Grant no. 279836 (2013-1018) from the Simons Foundation and by the NSF grant DMS-1405146. Both authors acknowledge support from U.S. National Science Foundation grants DMS 1107452, 1107263, 1107367 "RNMS: GEometric structures And Representation varieties" (the GEAR Network).

MSC2010: primary 20F65; secondary 57M07, 37B99, 37D40.

Keywords: Culler–Vogtmann's outer space, Patterson–Sullivan measures, geodesic currents.

1. Introduction

For an integer $N \geq 2$, the *unprojectivized outer space* cv_N is the set of all $\mathbb{R}$-trees equipped with a free discrete minimal isometric action of F_N, considered up to an F_N-equivariant isometry. We denote by cv_N^1 the set of all $T \in \mathrm{cv}_N$ such that the metric graph T/F_N has volume 1. The closure $\overline{\mathrm{cv}}_N$ of cv_N with respect to the equivariant Gromov–Hausdorff convergence topology (or equivalently [Paulin 1989], with respect to the hyperbolic length function topology) consists of all *very small* minimal isometric actions of F_N on $\mathbb{R}$-trees, again up to an F_N-equivariant isometry. There is a natural action of $\mathbb{R}_{>0}$ on $\overline{\mathrm{cv}}_N$ by multiplying the metric on a tree by a positive scalar. The subset cv_N of $\overline{\mathrm{cv}}_N$ is invariant under this action, and the quotient $\mathrm{CV}_N = \mathrm{cv}_N/\mathbb{R}_{>0}$ is the *projectivized outer space*, originally introduced by Culler and Vogtmann [1986]. The quotient $\overline{\mathrm{CV}}_N = \overline{\mathrm{cv}}_N/\mathbb{R}_{>0}$ is compact, and is called the *Thurston compactification* of CV_N. All of the above spaces admit natural $\mathrm{Out}(F_N)$-actions. The space CV_N is naturally $\mathrm{Out}(F_N)$-equivariantly homeomorphic to cv_N^1, but it is useful to remember that technically cv_N^1 and CV_N are distinct objects.

There are three main quantitative tools for studying points of $\overline{\mathrm{cv}}_N$. The first is the so-called "asymmetric Lipschitz distance". If $T \in \mathrm{cv}_N$ and $S \in \overline{\mathrm{cv}}_N$, the *extremal Lipschitz distortion* is given by

$$\Lambda(T, S) := \sup_{w \in F_N \setminus \{1\}} \frac{\|w\|_S}{\|w\|_T}.$$

It is known (see [Francaviglia and Martino 2011] for details) that this supremum is actually a maximum, and that $\Lambda(T, S)$ is the infimum of the Lipschitz constants of all the F_N-equivariant Lipschitz maps $T \to S$. It is also known that for all $T, S \in \mathrm{cv}_N^1$, we have $\Lambda(T, S) \geq 1$, and that the equality holds if and only if $T = S$. The *asymmetric Lipschitz distance* is defined as $d_L(T, S) := \log \Lambda(T, S)$, where $T, S \in \mathrm{cv}_N^1$. Although it is usually the case that $d_L(T, S) \neq d_L(S, T)$, the asymmetric distance d_L satisfies all the other properties of being a metric, and it is known that the topology defined by d_L on cv_N^1 coincides with the standard subspace topology for $\mathrm{cv}_N^1 \subseteq \mathrm{cv}_N$. Moreover, for any $T, S \in \mathrm{cv}_N^1$, there exists an (in general nonunique) d_L-geodesic path from T to S in cv_N^1, given by natural "folding lines" [loc. cit.]. The asymmetric distance d_L is a useful tool in the study of the geometry of $\mathrm{Out}(F_N)$ and it has found significant recent applications; see, for example, [Algom-Kfir 2011; 2013; Algom-Kfir and Bestvina 2012; Bestvina 2011; Francaviglia and Martino 2011; 2012; Ladra et al. 2015; White 1991].

Another two important quantitative tools for studying outer space are two notions of a "geometric intersection number". The first of these was introduced by Guirardel [2005] in the general setting of groups acting by isometries on $\mathbb{R}$-trees. Guirardel's intersection number $i(T,S)$ (where $T,S\in\overline{\mathrm{cv}}_N$) is defined as the covolume of the "core" for the action of F_N on $T\times S$. Guirardel's intersection number is symmetric and $\mathrm{Out}(F_N)$-invariant, and for $T,S\in\mathrm{cv}_N$, one always has $0\le i(T,S)<\infty$. However, for trees in $\partial\mathrm{cv}_N=\overline{\mathrm{cv}}_N\setminus\mathrm{cv}_N$, it is often the case that $i(T,S)=\infty$ and $i(\,\cdot\,,\cdot\,)$ is discontinuous when viewed as a function on $\overline{\mathrm{cv}}_N\times\overline{\mathrm{cv}}_N$. Still, Guirardel's intersection number is a highly useful tool when studying the asymptotic geometry of cv_N itself, particularly when looking at orbits of subgroups of $\mathrm{Out}(F_N)$ in cv_N^1 and cv_N. Examples of such applications can be found in [Behrstock et al. 2010; Clay et al. 2015; Clay and Pettet 2010; 2012b; Guirardel 2005; Horbez 2012].

The second notion of a "geometric intersection number" was introduced in [Kapovich and Lustig 2009]. There we constructed a *geometric intersection form* $\langle\,\cdot\,,\cdot\,\rangle:\overline{\mathrm{cv}}_N\times\mathrm{Curr}(F_N)\to\mathbb{R}_{\ge0}$, where $\mathrm{Curr}(F_N)$ is the space of *geodesic currents* on F_N. See Section 2C below and [Kapovich 2005; 2006; Kapovich and Lustig 2007; 2009] for the more information and the background on geodesic currents. The geometric intersection form is continuous, $\mathrm{Out}(F_N)$-equivariant, and, importantly, it always gives a finite output; that is, for every $T\in\overline{\mathrm{cv}}_N$ and $\mu\in\mathrm{Curr}(F_N)$, one has $0\le\langle T,\mu\rangle<\infty$. If $T\in\overline{\mathrm{cv}}_N$ and $g\in F_N\setminus\{1\}$ then $\langle T,\eta_g\rangle=\|g\|_T$, where $\eta_g\in\mathrm{Curr}(F_N)$ is the "counting current" associated with g. By its very definition, $\langle\,\cdot\,,\cdot\,\rangle$ is an asymmetric gadget. However, its good properties, including finiteness and global continuity on $\overline{\mathrm{cv}}_N$, make the geometric intersection form a useful tool that has also found a number of significant applications to the study of the dynamics and geometry of $\mathrm{Out}(F_N)$. See, for example, [Bestvina and Feighn 2010; Bestvina and Reynolds 2012; Carette et al. 2012; Clay and Pettet 2012a; Coulbois and Hilion 2014; Coulbois et al. 2008b; Hamenstädt 2014a; 2014b; Kapovich and Lustig 2009; 2010a; 2010b; Mann and Reynolds 2013; Reynolds 2012].

For $\varepsilon\ge0$, we denote by $\mathrm{cv}_{N,\varepsilon}^1$ the set of all $T\in\mathrm{cv}_N^1$ such that the length of the shortest simple closed loop in T/F_N is at least ε. The set $\mathrm{cv}_{N,\varepsilon}^1$ is called the *ε-thick* part of cv_N^1. Horbez [2012] showed that, for any fixed $\varepsilon>0$, if $T,S\in\mathrm{cv}_{N,\varepsilon}^1$, one has

$$(\ddagger)\qquad \frac{1}{K_1}\log i_c(T,S)-K_2\le d_L(T,S)\le K_1\log i_c(T,S)+K_2$$

for some constants $K_1\ge1$, $K_2\ge0$ depending only on N and ε. Here $i_c(T,S)$ is the combinatorial version of Guirardel's intersection number, where $i_c(T,S)$ is defined as the number of 2-cells in $\mathrm{Core}(T\times S)/F_N$, while $i(T,S)$ is defined as the sum of the areas of all the 2-cells in $\mathrm{Core}(T\times S)/F_N$. Thus if, for $S,T\in\mathrm{cv}_N^1$, the trees $T_0,S_0\in\mathrm{cv}_N$ are obtained from T and S by making all edges have length 1, then $i_c(T,S):=i(T_0,S_0)$. Also, following the usual convention, in $(\ddagger)$ we interpret $\log 0$ as $\log 0=0$.

In the present paper, for $T,S\in\mathrm{cv}_{N,\varepsilon}^1$, we relate $\Lambda(T,S)$ to a natural quantity defined in terms of $\langle\,\cdot\,,\cdot\,\rangle$. Via Horbez' result, this connection also relates the geometric intersection form $\langle \cdot, \cdot \rangle$ to Guirardel's geometric intersection number $i(\cdot, \cdot)$. Following the results of Furman [2002] in the general set-up of word-hyperbolic groups, Kapovich and Nagnibeda [2007] associated to every $T \in \mathrm{cv}_N$ its *Patterson–Sullivan current*. In general, the Patterson–Sullivan current is naturally defined only up to a multiplication by a positive scalar. Normalizing by the geometric intersection number with T provides a canonical choice. Thus for a tree $T \in \mathrm{cv}_N$, we denote by $\mu_T \in \mathrm{Curr}(F_N)$ the *Patterson–Sullivan current* associated to T, normalized so that $\langle T, \mu_T \rangle = 1$. We refer the reader to Section 4 below and to [Furman 2002; Kapovich and Nagnibeda 2007; 2010] for the precise definitions and background information about the Patterson–Sullivan currents. A key result obtained by Kapovich and Nagnibeda [2007] shows that the map $J_{PS} : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$, $T \mapsto \mu_T$ is a continuous $\mathrm{Out}(F_N)$-equivariant embedding.

Our main result (see Theorem 4.2 below) is:

Theorem 1.1. *Let* $N \geq 2$ *and* $\varepsilon > 0$. *Then there exist constants* $0 < \delta_1 \leq \delta_2$ *such that for every* $T \in \mathrm{cv}^1_{N,\varepsilon}$ *and every* $S \in \overline{\mathrm{cv}}_N$ *we have*

$$\delta_1 \leq \frac{\langle S, \mu_T \rangle}{\Lambda(T, S)} \leq \delta_2.$$

Therefore there exists a constant $c = c(N, \varepsilon) > 0$ *such that for every* $T \in \mathrm{cv}^1_{N,\varepsilon}$ *and* $S \in \mathrm{cv}^1_N$ *we have*

$$|\log\langle S, \mu_T \rangle - d_L(T, S)| \leq c.$$

Using the result of Horbez [2012] stated in (‡) above, Theorem 1.1 directly implies (using the notation introduced after (‡)):

Corollary 1.2. *Let* $N \geq 2$ *and* $\varepsilon > 0$. *There exist constants* $C_1, C_2 \geq 1$ *such that for any* $T, S \in \mathrm{cv}^1_{N,\varepsilon}$, *we have*

$$\frac{1}{C_1} \log i_c(T, S) - C_2 \leq \log\langle S, \mu_T \rangle \leq C_1 \log i_c(T, S) + C_2.$$

The proof of Theorem 1.1 relies on several results regarding geodesic currents, particularly one from [Kapovich and Lustig 2009] about the continuity of the already mentioned geometric intersection form on $\overline{\mathrm{cv}}_N \times \mathrm{Curr}(F_N)$, and a result from [Kapovich and Nagnibeda 2007] saying that the Patterson–Sullivan map $\mathrm{cv}^1_N \to \mathrm{Curr}(F_N)$, $T \mapsto \mu_T$, is a continuous $\mathrm{Out}(F_N)$-equivariant embedding. The crucial point in the argument uses a result from [Kapovich and Lustig 2010a] characterizing the case $\langle S, \nu \rangle = 0$, where $S \in \overline{\mathrm{cv}}_N$ and $\nu \in \mathrm{Curr}(F_N)$ are arbitrary. This characterization implies that every current μ with full support (such as the Patterson–Sullivan current μ_T for $T \in \mathrm{cv}^1_N$) is *filling*, that is, satisfies $\langle S, \mu \rangle > 0$ for every $S \in \overline{\mathrm{cv}}_N$. Modulo the tools mentioned above, the proof of Theorem 1.1 is not difficult (although it does require an extra trick exploiting the $\mathrm{Out}(F_N)$-equivariant nature of certain functions and some nice properties of d_L). Still, Theorem 1.1

and its applications obtained here do provide a conceptual clarification regarding the quantitative relationships between the two notions of a geometric intersection number used in the study of Out(F_N), and about their relationship to the asymmetric Lipschitz distance.

One of our main motivations for this paper has been to better understand the properties of "generic stretching factors" for free group automorphisms.

Proposition-Definition 1.3 [Kaimanovich et al. 2007]. *For any free basis A of F_N and any $S \in \overline{\mathrm{cv}}_N$, there exists a number $\lambda_A(S) \geq 0$ with the following property.*

For a.e. trajectory $\xi = y_1 y_2 \cdots y_n \cdots$ of the simple nonbacktracking random walk on F_N with respect to A (that is, for a "random" geodesic ray $\xi = y_1 y_2 \cdots y_n \cdots$ over $A^{\pm 1}$ with $y_i \in A^{\pm 1}$), we have $\|y_1 y_2 \cdots y_n\|_A = n + o(n)$ and

$$\lim_{n\to\infty} \frac{\|y_1 y_2 \cdots y_n\|_S}{n} = \lim_{n\to\infty} \frac{\|y_1 y_2 \cdots y_n\|_S}{\|y_1 y_2 \cdots y_n\|_A} = \lambda_A(S).$$

The number $\lambda_A(S)$ is called [Kapovich 2006; Kaimanovich et al. 2007] the generic stretching factor *of S with respect to A.*

The term "nonbacktracking" in "nonbacktracking simple random walk" refers to the fact that for this random walk, if $x, y \in A \cup A^{-1}$, the transition probability for x to be followed by y is equal to $1/(2N-1)$ if $y \neq x^{-1}$ and is equal to 0 if $y = x^{-1}$. Thus the trajectories of this random walk are semi-infinite freely reduced words over $A^{\pm 1}$. Informally, the generic stretching factor $\lambda_A(S) \geq 0$ captures the distortion $\|y_1 y_2 \cdots y_n\|_S/n$, where $y_1 \cdots y_n$ is a "random" freely reduced word of length n over A, as n tends to infinity. The existence of $\lambda_A(S) \geq 0$ follows from general ergodic-theoretic considerations, as observed in [Kaimanovich et al. 2007]. As noted in Remark 4.6 below, one actually has $\lambda_A(S) > 0$ for every $S \in \overline{\mathrm{cv}}_N$.

Let A be a free basis of F_N and consider the Cayley tree $T_A \in \mathrm{cv}_N$, with all edges of length $1/N$, so that $T_A \in \mathrm{cv}_N^1$. Thus for every $w \in F_N$, we have $\|w\|_A = N\|w\|_{T_A}$, where $\|w\|_A$ is the cyclically reduced length of w over $A^{\pm 1}$. It is known that the Patterson–Sullivan current μ_{T_A} is equal to the "uniform current" ν_A on F_N corresponding to A. Using the interpretation of $\langle S, \nu_A\rangle$ as the "generic stretching factor" $\lambda_A(S)$ of $S \in \mathrm{cv}_N$ with respect to A [Kapovich 2006], as a consequence of Theorem 1.1 we also obtain (see Theorem 4.7 below):

Corollary 1.4. *Let $N \geq 2$. There exists a constant $\delta = \delta(N) \in (0, 1)$ with the following property*:

For any free basis A of F_N and any $S \in \overline{\mathrm{cv}}_N$, we have

$$0 < \delta \leq \frac{\lambda_A(S)}{\Lambda(T_A, S)} \leq \frac{1}{N}. \tag{†}$$

We are particularly interested in relationships between generic stretching factors and extremal stretching factors in the context of Cayley trees of F_N and of elements of $\mathrm{Out}(F_N)$. Note that if A is a free basis of A then $NT_A \in \mathrm{cv}_N$ is the standard Cayley graph of F_N with respect to A, where all edges have length 1.

If $\varphi \in \mathrm{Out}(F_N)$ and $w \in F_N$, then, since φ is an outer automorphism, it acts on the conjugacy classes of elements of F_N (rather than on elements of F_N). By convention, for $\varphi \in \mathrm{Out}(F_N)$ and $w \in F_N$, if $\varphi(w)$ appears in an expression that depends only on the conjugacy class $\varphi([w])$, we will use $\varphi(w)$ to mean any representative of that conjugacy class.

Definition 1.5 (extremal and generic stretching factors of automorphisms). Let A be a free basis of F_N and let $\varphi \in \mathrm{Out}(F_N)$.

Define

$$\Lambda_A(\varphi) := \Lambda(T_A, T_A\varphi) = \sup_{w\neq 1} \frac{\|\varphi(w)\|_A}{\|w\|_A} = e^{d_L(T_A, T_A\varphi)},$$

and refer to $\Lambda_A(\varphi)$ as the *extremal stretching factor* for φ with respect to A.

Also, define $\lambda_A(\varphi) := \lambda_A(NT_A\varphi) = N\lambda_A(T_A\varphi)$.

Thus for a.e. trajectory $\xi = y_1 \cdots y_n \cdots$ of the simple nonbacktracking random walk on F_N with respect to A, we have

$$\lambda_A(\varphi) = \lim_{n\to\infty} \frac{\|\varphi(y_1 y_2 \cdots y_n)\|_A}{n} = \lim_{n\to\infty} \frac{\|\varphi(y_1 y_2 \cdots y_n)\|_A}{\|y_1 y_2 \cdots y_n\|_A}.$$

We call $\lambda_A(\varphi)$ the *generic stretching factor* of φ with respect to A.

Thus $\Lambda_A(\varphi)$ measures the maximal distortion $\|\varphi(w)\|_A/\|w\|_A$ as w varies over all nontrivial elements of F_N, while $\lambda_A(\varphi)$ captures the "generic distortion" $\|\varphi(w)\|_A/\|w\|_A$, where w is a "long random" freely reduced (or cyclically reduced) word over $A^{\pm 1}$. In practice, $\Lambda_A(\varphi)$ is easy to compute since it is known (see, e.g., [Francaviglia and Martino 2011]) that $\Lambda_A(\varphi) = \max_{1\le\|w\|\le 2}(\|\varphi(w)\|_A/\|w\|_A)$.

The generic stretching factors $\lambda_A(\varphi)$ were introduced in [Kaimanovich et al. 2007] and further studied in [Francaviglia 2009; Kapovich 2006; Kapovich and Lustig 2010a; Sharp 2010]. In particular, it is proved in [Kaimanovich et al. 2007] that for every $\varphi \in \mathrm{Out}(F_N)$, the number $\lambda_A(\varphi)$ is rational, and moreover, $2N\lambda_A(\varphi) \in \mathbb{Z}[1/(2N-1)]$ and there exists an algorithm that, given φ, computes $\lambda_A(\varphi)$. The definitions directly imply that $\lambda_A(\varphi) \le \Lambda_A(\varphi)$. However, other than this fact, the quantitative relationship between $\Lambda_A(\varphi)$ and $\lambda_A(\varphi)$ remained unclear.

Let $N \ge 2$ and $F_N = F(a_1, \dots, a_N)$ with $A = \{a_1, \dots, a_N\}$. Define

$$\rho_N := \inf_{\varphi\in\mathrm{Out}(F_N)} \frac{\lambda_A(\varphi)}{\Lambda_A(\varphi)}.$$

For every $\varphi \in \mathrm{Out}(F_N)$, we have $T_A, T_A\varphi \in \mathrm{cv}^1_{N,\varepsilon}$ with $\varepsilon = 1/N$, and thus Corollary 1.4 directly implies:

Theorem 1.6. *For every $N \ge 2$ we have $\rho_N > 0$.*

Therefore for every $\varphi \in \mathrm{Out}(F_N)$, we have

$$0 < \rho_N \le \frac{\lambda_A(\varphi)}{\Lambda_A(\varphi)} \le 1.$$

Our proof that $\rho_N > 0$ does not give any explicit quantitative information about ρ_N. It would be interesting to find some explicit bounds from above and below for ρ_N, and perhaps to even compute ρ_N, at least for small values of N. We show in Proposition 7.1 that $\lim_{N\to\infty} \rho_N = 0$ and that $\rho_N = O(1/N)$.

As another application, we obtain (see Corollary 5.3 below):

Corollary 1.7. *Let $N \ge 2$ and $F_N = F(a_1, \dots, a_n)$ with $A = \{a_1, \dots, a_N\}$. There exists $D = D(N) \ge 1$ such that for every $\varphi \in \mathrm{Out}(F_N)$ we have*

$$\frac{1}{D} \log \lambda_A(\varphi) \le \log \lambda_A(\varphi^{-1}) \le D \log \lambda_A(\varphi).$$

Let $\varphi \in \mathrm{Out}(F_N)$. Recall that the *algebraic stretching factor* $\lambda(\varphi)$ is defined as

$$\lambda(\varphi) := \sup_{w\in F_N, w\ne 1} \lim_{n\to\infty} \sqrt[n]{\|\varphi^n(w)\|_S},$$

where $S \in \mathrm{cv}_N$ is an arbitrary base point. It is known that the limit in the last equality always exists, that this definition of $\lambda(\varphi)$ does not depend on the choice of $S \in \mathrm{cv}_N$, and that we always have $\lambda(\varphi) \ge 1$. An element $\varphi \in \mathrm{Out}(F_N)$ is called *exponentially growing* if $\lambda(\varphi) > 1$, and *polynomially growing* if $\lambda(\varphi) = 1$. Indeed, it is known (see, for example, [Levitt 2009]), that φ is polynomially growing if and only if for every $w \in F_N$ and $S \in \mathrm{cv}_N$, the sequence $\|\varphi^n(w)\|_S$ is bounded above by a polynomial in n.

The algebraic stretching factor $\lambda(\varphi)$ can be read off from any relative train-track representative $f : \Gamma \to \Gamma$ of φ as the maximum of the Perron–Frobenius eigenvalues for any of the canonical irreducible diagonal blocks of the (nonnegative) transition matrix $M(f)$.

As another application of the results of this paper, we explain how the generic stretching factor $\lambda_A(\varphi^n)$ grows in terms of n for an arbitrary $\varphi \in \mathrm{Out}(F_N)$. Thus we obtain (see Theorem 5.6 below) the following result, which answers Problem 9.2 posed in [Kaimanovich et al. 2007]:

Theorem 1.8. *Let A be a free basis of F_N, let $\varphi \in \mathrm{Out}(F_N)$ and let $\lambda(\varphi)$ be the algebraic stretching factor of φ. Then there exist constants $c_1, c_2 > 0$ and an integer $m \ge 0$ such that for every $n \ge 1$, we have*

$$c_1\, \lambda(\varphi)^n\, n^m \le \lambda_A(\varphi^n) \le c_2\, \lambda(\varphi)^n\, n^m.$$

Moreover, if φ admits an expanding train-track representative with an irreducible transition matrix (e.g., if φ is fully irreducible), then $m = 0$ and $\lambda(\varphi) > 1$.

The "polynomial growth degree" m in this result is bounded above by the number of strata of any relative train track representative f as above which have PF-eigenvalue equal to λ, and it has been determined precisely by Levitt [2009], see the proof of Proposition 5.4 below.

2. Preliminaries

2A. ***Basic terminology and notations related to outer space.*** We denote by cv_N the unprojectivized outer space, that is, the space of all free discrete minimal isometric actions of F_N on $\mathbb{R}$-trees, considered up to F_N-equivariant isometry. Denote by $\overline{\text{cv}}_N$ the closure of cv_N in the equivariant Gromov–Hausdorff convergence topology (or, equivalently, in the hyperbolic length functions topology). It is known [Bestvina and Feighn 1993; Cohen and Lustig 1995; Guirardel 1998] that $\overline{\text{cv}}_N$ consists of all the *very small* nontrivial minimal isometric actions of F_N on $\mathbb{R}$-trees, again considered up to F_N-equivariant isometry. Recall that a point $T \in \overline{\text{cv}}_N$ is uniquely determined by its *translation length function* $\|\cdot\|_T : F_N \to [0,\infty)$, where for $w \in F_N$, we have $\|w\|_T = \inf_{x\in T} d(x, wx) = \min_{x\in T} d(x, wx)$.

The space $\overline{\text{cv}}_N$ has a natural right $\text{Out}(F_N)$-action, where for $w \in F_N$ and $T \in \overline{\text{cv}}_N$, we have $\|w\|_{T\varphi} = \|\varphi(w)\|_T$. It is sometimes useful to convert this action to a left $\text{Out}(F_N)$-action by setting $\varphi T := T\varphi^{-1}$. Define

$$\text{cv}_N^1 := \{T \in \text{cv}_N \mid \text{vol}(T/F_N) = 1\},$$

and refer to cv_N^1 as the *volume-normalized outer space* or just *normalized outer space*. Then cv_N is an open dense $\text{Out}(F_N)$-invariant subset of $\overline{\text{cv}}_N$, and cv_N^1 is a closed $\text{Out}(F_N)$-invariant subset of cv_N (but of course cv_N^1 is not closed in $\overline{\text{cv}}_N$).

There is a natural action of $\mathbb{R}_{>0}$ on cv_N and $\overline{\text{cv}}_N$ by scalar multiplication, which yields the corresponding *projectivizations* $\text{CV}_N = \text{cv}_N/\mathbb{R}_{>0}$ and $\overline{\text{CV}}_N = \overline{\text{cv}}_N/\mathbb{R}_{>0}$. For a tree $T \in \overline{\text{cv}}_N$, we denote its projective class in $\overline{\text{CV}}_N$ by $[T]$. Thus $[T] = \{cT \mid c > 0\}$. Note that CV_N is canonically $\text{Out}(F_N)$ equivariantly homeomorphic to cv_N^1, but it is still important to remember that technically CV_N and cv_N^1 are distinct objects.

For $\varepsilon > 0$, we denote by $\text{cv}_{N,\varepsilon}^1$ the set of all $T \in \text{cv}_N^1$ such that the shortest nontrivial immersed circuit in the metric graph T/F_N has length $\geq \varepsilon$. Equivalently, $\text{cv}_{N,\varepsilon}^1$ is the set of all $T \in \text{cv}_N^1$ such that for every $w \in F_N \setminus \{1\}$, we have $\|w\|_T \geq \varepsilon$. For every $\varepsilon > 0$, the set $\text{cv}_{N,\varepsilon}^1 \subseteq \text{cv}_N^1$ is a closed $\text{Out}(F_N)$-invariant subspace, and the quotient $\text{cv}_{N,\varepsilon}^1 / \text{Out}(F_N)$ is compact.

A *chart* on F_N is an isomorphism $\alpha : F_N \to \pi_1(\Gamma, p)$, where Γ is a finite connected graph with all vertices of degree ≥ 3 and where p is a base vertex in Γ (which is usually suppressed). Every such α defines an open cone in cv_N consisting of assigning arbitrary positive lengths to edges of Γ and then lifting this assignment to the universal cover $\widetilde{\Gamma}$ to get an element $T \in \mathrm{cv}_N$. The intersection of such an open cone with cv_N^1 is an open simplex Δ in cv_N^1 of dimension $m-1$, where m is the number of nonoriented edges of Γ. Every point $T \in \mathrm{cv}_N$ belongs to a unique open cone of this form, and every point of cv_N^1 belongs to a unique such open simplex Δ.

The space $\overline{\mathrm{CV}}_N$ is known to be compact and finite-dimensional.

2B. ***Asymmetric Lipschitz distance.*** For points $T \in \mathrm{cv}_N$ and $S \in \overline{\mathrm{cv}}_N$, define

$$\Lambda(T,S) = \sup_{w \in F_N \setminus \{1\}} \frac{\|w\|_S}{\|w\|_T}.$$

If $T, S \in \mathrm{cv}_N^1$, we also define $d_L(T,S) := \log \Lambda(T,S)$. As noted in the Introduction, for $T, S \in \mathrm{cv}_N^1$, the quantity $d_L(T,S)$ is often called the *asymmetric Lipschitz distance* from T to S.

Remark 2.1. If $T \in \mathrm{cv}_N$ and $S \in \overline{\mathrm{cv}}_N$ then $0 < \Lambda(T,S) < \infty$. Moreover, it is known [Francaviglia and Martino 2011; White 1991] that for any open simplex $\Delta \subset \mathrm{cv}_N^1$ as in Section 2A, there exists a finite subset $C_\Delta \subseteq F_N \setminus \{1\}$ such that for every $T \in \Delta$ and every $S \in \overline{\mathrm{cv}}_N$, we have

$$\Lambda(T,S) = \max_{w \in C_\Delta} \frac{\|w\|_S}{\|w\|_T}.$$

The set C_Δ can be chosen to be contained in the subset of all elements which are represented by paths that cross at most twice over every nonoriented edge of $\Gamma = T/F_N$ for $T \in \Delta$.

Note also that from the definition, we see that for every $T \in \mathrm{cv}_N$, $S \in \overline{\mathrm{cv}}_N$ and $\varphi \in \mathrm{Out}(F_N)$, one has $\Lambda(T,S) = \Lambda(\varphi T, \varphi S)$.

2C. ***Geodesic currents.*** We refer the reader to [Kapovich 2006; Kapovich and Lustig 2007; 2009; 2010a] for detailed background on geodesic currents, and we only recall a few basic definitions and facts here. Let $\partial^2 F_N = \partial F_N \times \partial F_N \setminus \mathrm{diag}$, and endow $\partial^2 F_N$ with the subspace topology and with the diagonal F_N-action by translations. A *geodesic current* on F_N is a positive Borel measure μ on $\partial^2 F_N$ such that μ is finite on compact subsets, F_N-invariant and "flip"-invariant (where the "flip" map $\partial^2 F_N \to \partial^2 F_N$ interchanges the two coordinates). The space of all geodesic currents on F_N is denoted $\mathrm{Curr}(F_N)$. The space $\mathrm{Curr}(F_N)$ comes equipped with a natural weak*-topology and a natural left $\mathrm{Out}(F_N)$-action by affine homeomorphisms.

Let $\alpha : F_N \to \pi_1(\Gamma, p)$ be a chart on F_N, and consider $\widetilde{\Gamma}$ with the simplicial metric, where every edge has length 1. Then there is a natural F_N-equivariant quasi-isometry (given for any point $p \in \widetilde{\Gamma}$ by the orbit map $F_N \to \widetilde{\Gamma}$, $g \mapsto gp$) between F_N and $\widetilde{\Gamma}$, which induces a canonical F_N-equivariant homeomorphism between

∂F_N and $\partial\widetilde{\Gamma}$. We will therefore identify ∂F_N with $\partial\widetilde{\Gamma}$ using this homeomorphism without invoking it explicitly, whenever it is convenient.

A nondegenerate geodesic segment γ in $\widetilde{\Gamma}$ defines a *cylinder set* $\mathrm{Cyl}_\alpha(\gamma)$ consisting of all $(X, Y) \in \partial^2 F_N$ such that the geodesic from X to Y in $\widetilde{\Gamma}$ passes through γ (in the correct direction). The sets $\mathrm{Cyl}_\alpha(\gamma)$, as γ varies among all nondegenerate geodesic edge-paths in $\widetilde{\Gamma}$, are compact and open, and form a basis for the topology on $\partial^2 F_N$. Note that for $w \in F_N$, we have $\mathrm{Cyl}_\alpha(w\gamma) = w\,\mathrm{Cyl}_\alpha(\gamma)$. If $\mu \in \mathrm{Curr}(F_N)$ and v is a nondegenerate reduced edge-path in Γ, we define the *weight* $\langle v, \mu\rangle_\alpha := \mu(\mathrm{Cyl}_\alpha(\gamma))$, where γ is any lift of v. Since the measure μ is F_N-invariant, this definition does not depend on the specific choice of the lift γ of v to $\widetilde{\Gamma}$. A current μ is uniquely determined by its collection of weights with respect to a given chart. Moreover, if $\mu_n, \mu \in \mathrm{Curr}(F_N)$ and α is a chart as above, then $\lim_{n\to\infty} \mu_n = \mu$ in $\mathrm{Curr}(F_N)$ if and only if for every nondegenerate reduced edge-path v in Γ, we have $\lim_{n\to\infty}\langle v, \mu_n\rangle_\alpha = \langle v, \mu\rangle_\alpha$.

For every $w \in F_N \setminus \{1\}$, there is an associated *counting current* $\eta_w \in \mathrm{Curr}(F_N)$, which depends only on the conjugacy class $[w]$ of w in F_N and satisfies $\eta_{w^{-1}} = \eta_w$ and $\eta_{w^n} = n\,\eta_w$ for all integers $n \geq 1$, and such that $\varphi\,\eta_w = \eta_{\varphi(w)}$ for all $\varphi \in \mathrm{Out}(F_N)$, $w \in F_N \setminus \{1\}$. The precise definition of η_w is not important at the moment, but we will recall some of its basic properties later, as necessary. The set $\{c\eta_w \mid c > 0, w \in F_N, w \neq 1\}$ of the so-called *rational currents* is dense in $\mathrm{Curr}(F_N)$.

Be aware that, in general, for a representative (even a train-track representative) $f : \Gamma \to \Gamma$ of φ, one has $\langle v, \varphi\mu\rangle_\alpha \neq \langle [f(v)], \mu\rangle_\alpha$, where $[f(v)]$ denotes the edge-path obtained from $f(v)$ by reduction (that is, the iterative contraction of any backtracking path).

2D. ***Intersection form.*** Kapovich and Lustig [2009] proved the existence of a continuous *geometric intersection form* between points of $\overline{\mathrm{cv}}_N$ and geodesic currents:

Proposition 2.2 [Kapovich and Lustig 2009]. *There exists a unique continuous function* $\langle\cdot,\cdot\rangle : \overline{\mathrm{cv}}_N \times \mathrm{Curr}(F_N) \to [0, \infty)$, *called the* geometric intersection form, *with the following properties*:

(1) *For any* $\mu_1, \mu_2 \in \mathrm{Curr}(F_N)$, $T \in \overline{\mathrm{cv}}_N$, $c_1, c_2 \geq 0$ *and* $r > 0$, *we have*

$$\langle rT, c_1\mu_1 + c_2\mu_2\rangle = rc_1\langle T, \mu_1\rangle + rc_2\langle T, \mu_2\rangle.$$

(2) *For any* $T \in \overline{\mathrm{cv}}_N$, $\mu \in \mathrm{Curr}(F_N)$ *and* $\varphi \in \mathrm{Out}(F_N)$, *we have*

$$\langle \varphi T, \varphi\mu\rangle = \langle T, \mu\rangle.$$

(3) *For any* $T \in \overline{\mathrm{cv}}_N$ *and* $w \in F_N \setminus \{1\}$, *we have*

$$\langle T, \eta_w\rangle = \|w\|_T.$$

(4) *For any* $T \in \mathrm{cv}_N$ (*with the associated chart* $\alpha : F_N \to \pi_1(T/F_N)$) *and any* $\mu \in \mathrm{Curr}(F_N)$, *we have*

$$\langle T, \mu\rangle = \sum_{e \in \mathrm{Edges}(T/F_N)} \tfrac{1}{2}\langle e; \mu\rangle_\alpha,$$

where the summation is taken over all oriented edges of the graph T/F_N.

3. Tree-current morphisms and extremal Lipschitz distortion

Recall that a current $\mu \in \mathrm{Curr}(F_N)$ is called *filling* if for every $S \in \overline{\mathrm{cv}}_N$, we have $\langle S, \mu\rangle > 0$.

We proved in [Kapovich and Lustig 2010a] that for a current $\mu \in \mathrm{Curr}(F_N)$ and a tree $T \in \overline{\mathrm{cv}}_N$, we have $\langle T, \mu\rangle = 0$ if and only if the support of μ is contained in the "dual algebraic lamination" of T (in the sense of [Coulbois et al. 2008a]). Using this fact, it was shown in [Kapovich and Lustig 2010a] that if μ is a current with full support, then μ is filling. We denote by $\mathrm{Curr}_{\mathrm{fill}}(F_N)$ the set of all filling $\mu \in \mathrm{Curr}(F_N)$, and endow $\mathrm{Curr}_{\mathrm{fill}}(F_N)$ with the subspace topology given by the inclusion $\mathrm{Curr}_{\mathrm{fill}}(F_N) \subseteq \mathrm{Curr}(F_N)$.

Definition 3.1 (tree-current morphism). A *tree-current morphism* is a continuous function $J : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$ such that for every $T \in \mathrm{cv}_N^1$ and every $\varphi \in \mathrm{Out}(F_N)$, we have $J(\varphi T) = \varphi\, J(T)$.

A *filling tree-current morphism* is a tree-current morphism $J : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$ such that for every $T \in \mathrm{cv}_N^1$, the current $J(T) \in \mathrm{Curr}(F_N)$ is filling.

Lemma 3.2. *The function* $\mathrm{cv}_N^1 \times \overline{\mathrm{cv}}_N \to \mathbb{R}$, $(T, S) \mapsto \Lambda(T, S)$, *is continuous.*

Proof. Let $T \in \mathrm{cv}_N^1$ be arbitrary.

Let $\Delta_1, \dots, \Delta_m$ be all the open simplices in cv_N^1 whose closures in cv_N^1 contain T.

Set $C_T = \bigcup_{i=1}^m C_{\Delta_i}$. Note that $U = \Delta_1 \cup \cdots \cup \Delta_m$ is a neighborhood of T in cv_N^1. Thus for every $T' \in U$ and every $S \in \overline{\mathrm{cv}}_N$, we have

$$\Lambda(T', S) = \max_{w \in C_T} \frac{\|w\|_S}{\|w\|_{T'}}.$$

Therefore the function $\Lambda(T', S)$ is continuous on $U \times \overline{\mathrm{cv}}_N$. Since $T \in \mathrm{cv}_N^1$ was arbitrary, the conclusion of the lemma follows. □

Let J be a filling tree-current morphism. Then for any $S \in \overline{\mathrm{cv}}_N$ and $c > 0$, we have

$$\frac{\langle S, J(T)\rangle}{\Lambda(T, S)} = \frac{\langle cS, J(T)\rangle}{\Lambda(T, cS)}.$$

Also, since $J(T)$ is a filling current, for every $S \in \overline{\mathrm{cv}}_N$, we have $\langle S, J(T)\rangle > 0$. Therefore we have a well-defined function

$$f : \mathrm{cv}_N^1 \times \overline{\mathrm{CV}}_N \to (0, \infty)$$

given by $f(T, [S]) = \langle S, J(T)\rangle/\Lambda(T, S)$, where $T \in \mathrm{cv}_N^1$ and $S \in \overline{\mathrm{cv}}_N$.

Lemma 3.3. *Let J be a filling tree-current morphism. Then the function*

$$f : \mathrm{cv}_N^1 \times \overline{\mathrm{CV}}_N \to (0, \infty), \quad (T, S) \mapsto \frac{\langle S, J(T)\rangle}{\Lambda(T, S)}$$

is continuous.

Proof. The conclusion of the lemma follows directly from Lemma 3.2 together with the continuity of the geometric intersection form $\langle\,\cdot\,,\cdot\,\rangle$. □

Corollary 3.4. *Let $K \subseteq \mathrm{cv}_N^1$ be a compact subset, and let $J : \mathrm{cv}_N^1 \to \mathrm{Curr}_{\mathrm{fill}}(F_N)$ be a filling tree-current morphism.*

Then there exist $\delta_1 = \delta_1(K, J) > 0$ and $\delta_2 = \delta_2(K, J) > 0$ such that for every $T \in K$ and every $S \in \overline{\mathrm{cv}}_N$, we have $\delta_1 \le f(K, [S]) \le \delta_2$.

Proof. The set $K \times \overline{\mathrm{CV}}_N$ is a compact Hausdorff space and, by Lemma 3.3, $f : K \times \overline{\mathrm{CV}}_N \to (0, \infty)$ is a continuous function. Therefore f achieves a positive minimum δ_1 and a positive maximum δ_2 on $K \times \overline{\mathrm{CV}}_N$, and the conclusion of the corollary follows. □

Corollary 3.5. *Let $K \subseteq \mathrm{cv}_N^1$ be a compact subset, let $\mathcal{T}_K = \bigcup_{\varphi \in \mathrm{Out}(F_N)} \varphi K$ and let $J : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$ be a filling tree-current morphism.*

Furthermore, let $\delta_1 = \delta_1(K, J) > 0$ and $\delta_2 = \delta_2(K, J) > 0$ be the constants provided by Corollary 3.4.

Then for every $T \in \mathcal{T}_K$ and every $[S] \in \overline{\mathrm{CV}}_N$, we have

$$0 < \delta_1 \le \frac{\langle S, J(T)\rangle}{\Lambda(T, S)} \le \delta_2 < \infty.$$

Proof. Let $T \in \mathcal{T}_K$ and $[S] \in \overline{\mathrm{CV}}_N$ be arbitrary.

Then there exist $T' \in K$ and $\varphi \in \mathrm{Out}(F_N)$ such that $T = \varphi T'$. By φ-equivariance of J, we have $J(T) = \varphi J(T')$. Define $S' = \varphi^{-1} S$, so that $\varphi S' = S$. Then

$$\frac{\langle S, J(T)\rangle}{\Lambda(T, S)} = \frac{\langle \varphi S', \varphi J(T')\rangle}{\Lambda(\varphi T', \varphi S')} = \frac{\langle S', J(T')\rangle}{\Lambda(T', S')} = f(T', [S']) \in [\delta_1 . \delta_2],$$

where the last inclusion holds by Corollary 3.4 since $T' \in K$. □

Note that Corollary 3.5 does not require the tree-current morphism $J : \mathrm{cv}_N^1 \to \mathrm{Curr}_{\mathrm{fill}}(F_N)$ to be injective, although in the specific applications of interest to us J will be injective.

4. Patterson–Sullivan currents and extremal Lipschitz distortion

4A. ***Volume entropy and the Patterson–Sullivan currents.*** We only give here a brief summary of basic definitions and facts regarding Patterson–Sullivan currents for points of cv_N. We refer the reader to [Furman 2002; Coornaert 1993; Kaimanovich 1991; Kapovich and Nagnibeda 2007] for more detailed background information about Patterson–Sullivan measures and Patterson–Sullivan currents in the context of word-hyperbolic groups and Gromov-hyperbolic spaces.

Let $T \in \mathrm{cv}_N$, where $N \geq 2$. Since F_N and T are F_N-equivariantly quasi-isometric, there is a natural identification of ∂F_N and ∂T, which we will use later on.

The *volume entropy* $h(T)$ of T is defined as

$$h(T) := \lim_{R\to\infty} \frac{\log\bigl(\#\{w \in F_N \mid d_T(p, wp) \leq R\}\bigr)}{R},$$

where $p \in T$ is an arbitrary base point. It is known that the above definition does not depend on the choice of a base-point $p \in T$ and that we have $h(T) > 0$ for every $T \in \mathrm{cv}_N$. It is also known that $h(T)$ is exactly the critical exponent of the *Poincaré series*

$$\Pi_p(s) = \sum_{w\in F_N} e^{-sd_T(p,wp)}.$$

In other words, $\Pi_p(s)$ converges for all $s > h(T)$ and diverges for all $s \leq h(T)$. It is also known that as $s \to h+$, any weak limit ν of the measures

$$\frac{1}{\Pi_p(s)} \sum_{w\in F_N} e^{-sd_T(p,wp)}\,\mathrm{Dirac}(wp)$$

is a probability measure supported on $\partial T = \partial F_N$. Any such ν is called a *Patterson–Sullivan* measure on ∂F_N corresponding to T, and the measure class of ν is canonically determined by T. As follows from general results of Furman [2002], in this case there exists a unique, up to a scalar multiple, geodesic current μ in the measure class of $\nu \times \nu$ on $\partial^2 F_N$. We call the unique scalar multiple μ_T of μ such that $\langle T, \mu_T\rangle = 1$, the *Patterson–Sullivan current* for $T \in \mathrm{cv}_N$. One also has that the current μ_T has full support (this follows, for example, both from the general results of Furman [2002] and from the explicit formulas for μ_T obtained in [Kapovich and Nagnibeda 2007]).

Proposition 4.1. *The map*

$$J_{PS} : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N),\ T \mapsto \mu_T$$

is a filling tree-current morphism.

Proof. Since μ_T has full support, by a result of Kapovich and Lustig [2010a, Corollary 1.3], it follows that $\mu_T \in \mathrm{Curr}_{\mathrm{fill}}(F_N)$. The fact that J_{PS} is a continuous $\mathrm{Out}(F_N)$-equivariant map was proved by Kapovich and Nagnibeda [2007]. Thus J_{PS} is indeed a filling tree-current morphism, as claimed. □

The fact that for $T \in \mathrm{cv}_N^1$, the Patterson–Sullivan current μ_T is filling, i.e., that $\langle S, \mu_T \rangle \neq 0$ for every $S \in \overline{\mathrm{cv}}_N$, is quite nontrivial and does not follow directly from Proposition 2.2. This fact, which requires a general result from [Kapovich and Lustig 2010a] characterizing the case where $\langle S, \mu \rangle = 0$ (where $S \in \overline{\mathrm{cv}}_N$ and $\mu \in \mathrm{Curr}(F_N)$), is, in a sense, the place where the real "magic" in the proofs of the main results of the present paper happens.

We now obtain Theorem 1.1 from the Introduction:

Theorem 4.2. *Let $N \geq 2$ and $\varepsilon > 0$. Then there exist constants $\delta_2 \geq \delta_2 > 0$ such that for every $T \in \mathrm{cv}_{N,\varepsilon}^1$, $S \in \overline{\mathrm{cv}}_N$ we have*

$$\delta_1 \leq \frac{\langle S, \mu_T \rangle}{\Lambda(T, S)} \leq \delta_2.$$

Therefore there exists a constant $c > 0$ such that for every $T \in \mathrm{cv}_{N,\varepsilon}^1$ and $S \in \mathrm{cv}_N^1$, we have

$$|\log\langle S, \mu_T \rangle - d_L(T, S)| \leq c.$$

Proof. Since $\mathrm{cv}_{N,\varepsilon}^1 / \mathrm{Out}(F_N)$ is compact and the action of $\mathrm{Out}(F_N)$ on $\mathrm{cv}_{N,\varepsilon}^1$ is properly discontinuous, there exists a compact subset $K \subseteq \mathrm{cv}_{N,\varepsilon}^1$ such that

$$\mathrm{cv}_{N,\varepsilon}^1 = \mathcal{T}_K = \bigcup_{\varphi \in \mathrm{Out}(F_N)} \varphi K.$$

By Proposition 4.1, the map $J_{PS} : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$ is a filling tree-current morphism. The conclusion of the theorem now follows from Corollary 3.5. □

4B. ***Uniform currents and generic stretching factors.*** Kapovich and Nagnibeda also provide reasonably explicit description of μ_T in terms of its weights on the "cylinder subsets" of $\partial^2 F_N$. The details of that description are not immediately relevant for the present paper. However, in the case where $T \in \mathrm{cv}_N^1$ and where T/F_N is a regular metric graph (that is, a regular graph where all edges have the same length), one can give a more precise description of μ_T as a "uniform current" corresponding to T and relate μ_T to the exit measure of the simple nonbacktracking random walk on T. We briefly recall here the description of uniform currents for the standard N-roses, that is for points of cv_N^1 corresponding to free bases of F_N.

Let $A = \{a_1, \dots, a_N\}$ be a free basis of F_N. Let R_N be the graph given by a wedge of N loop-edges $e_1, \dots, e_N$ at a vertex x_0. By identifying e_i with $a_i \in F_N$, we get an identification of $\alpha_A : F_N \xrightarrow{\cong} \pi_1(R_N, x_0)$, that is, a chart on F_N. We give each edge of R_N length $1/N$, so that R_N becomes a metric graph of volume 1.

Then the universal cover $T_A := \widetilde{R}_N$ is an $\mathbb{R}$-tree, which can be thought of as the Cayley graph of F_N with respect to A, but where all edges have length $1/N$. The group F_N has a natural free and discrete isometric left action on T_A by covering transformations, with $T_A/F_N = R_N$. Thus T_A is a point of cv_N^1.

The *uniform current* ν_A on F_N corresponding to A is defined explicitly by its weights. Namely, for every nontrivial freely reduced word v over $A^{\pm 1}$, we have

$$\langle v, \nu_A \rangle_{\alpha_A} = \frac{1}{N(2N-1)^{|v|-1}}.$$

One can check that this assignment of weights does define a geodesic current and that $\langle T_A, \nu_A \rangle = 1$. Moreover, in this case we also have:

Proposition 4.3. *Let $N \geq 2$ and let A be a free basis of F_N. Then $\mu_{T_A} = \nu_A$; that is, the Patterson–Sullivan current corresponding to T_A is exactly the uniform current ν_A.*

The above fact is not explicitly stated in [Kapovich and Nagnibeda 2007] but it easily follows from the explicit formulas for the weights for Patterson–Sullivan currents they obtained in the same work. Alternatively, one knows, for example, by the results of [Coornaert 1993; Lyons 1994], that for T_A the uniform visibility measure m_A on $\partial F_N = \partial T_A$ is a Patterson–Sullivan measure for T_A. Since $\nu_A \in \mathrm{Curr}(F_N)$ is in the measure class of $m_A \times m_A$ and since $\langle T_A, \nu_A \rangle = 1$, it follows from the definition of the Patterson–Sullivan current that $\mu_{T_A} = \nu_A$. Note that for any other $S \in \mathrm{cv}_N$, the intersection number $\langle S, \nu_A \rangle$ measures the distortion of a "long random geodesic" in T_A with respect to S.

Recall that in the Introduction, given a free basis A of F_N, $S \in \overline{\mathrm{cv}}_N$ and $\varphi \in \mathrm{Out}(F_N)$, we defined the generic stretching factors $\lambda_A(S)$ and $\lambda_A(\varphi)$.

Lemma 4.4. *For any free basis A of F_N and any $S \in \overline{\mathrm{cv}}_N$, we have*

$$\lambda_A(S) \leq \frac{1}{N} \Lambda(T_A, S).$$

Proof. Since all edges in T_A have length $1/N$, for every $w \in F_N$, we have $\|w\|_A = N\|w\|_{T_A}$. Then for a random trajectory $\xi = y_1 y_2 \cdots y_n \cdots$ of the simple nonbacktracking random walk on F_N with respect to A we have

$$\begin{aligned}\lambda_A(S) &= \lim_{n\to\infty} \frac{\|y_1 \cdots y_n\|_S}{\|y_1 \cdots y_n\|_A} = \lim_{n\to\infty} \frac{\|y_1 \cdots y_n\|_S}{N\|y_1 \cdots y_n\|_{T_A}} \\ &= \frac{1}{N} \lim_{n\to\infty} \frac{\|y_1 \cdots y_n\|_S}{\|y_1 \cdots y_n\|_{T_A}} \leq \frac{1}{N} \sup_{w \neq 1} \frac{\|w\|_S}{\|w\|_{T_A}} = \frac{1}{N} \Lambda(T_A, S). \qquad \square\end{aligned}$$

A key fact about generic stretching factors, originally established in [Kapovich 2006, Proposition 9.1] in slightly more limited context, is:

Proposition 4.5. *Let A be a free basis of F_N (where $N \geq 2$) and let $S \in \overline{\mathrm{cv}}_N$. Then*

$$\langle S, \nu_A \rangle = \lambda_A(S).$$

Proof. By [Kapovich 2006, Proposition 7.3], for a.e. trajectory $\xi = y_1 y_2 \cdots y_n \cdots$ of the simple nonbacktracking random walk on F_N with respect to A, we have

$$\lim_{n\to\infty} \frac{1}{n} \eta_{y_1 \cdots y_n} = \nu_A.$$

Therefore, by Proposition 2.2, for any $S \in \overline{\mathrm{cv}}_N$, we have

$$\langle S, \nu_A \rangle = \lim_{n\to\infty} \frac{1}{n} \langle S, \eta_{y_1 \cdots y_n} \rangle = \lim_{n\to\infty} \frac{\|y_1 \cdots y_n\|_S}{n} = \lambda_A(S). \qquad \square$$

Remark 4.6. Since the current ν_A has full support and therefore ν_A is filling, Proposition 4.5 implies that for every $S \in \overline{\mathrm{cv}}_N$, we have $\lambda_A(S) > 0$. (From the definition of $\lambda_A(S)$, one only knows that $\lambda_A(S) \geq 0$ and it is not a priori obvious that the case $\lambda_A(S) = 0$ cannot occur.)

We can now obtain Corollary 1.4 from the Introduction:

Theorem 4.7. *Let $N \geq 2$. Then there exists a constant $\delta = \delta(N) \in (0, 1)$ with the following property*:

For any free basis A of F_N and any $S \in \overline{\mathrm{cv}}_N$, we have

$$0 < \delta \leq \frac{\lambda_A(S)}{\Lambda(T_A, S)} \leq \frac{1}{N}.$$

Proof. Let A be a free basis of F_N and let $S \in \overline{\mathrm{cv}}_N$ be arbitrary. By Lemma 4.4, we have

$$\frac{\lambda_A(S)}{\Lambda(T_A, S)} \leq \frac{1}{N}.$$

Let $\delta = \delta_1(\varepsilon, N) > 0$ be the constant provided by Theorem 4.2. By decreasing this constant if necessary, we can always assume that $0 < \delta_1 < 1$. Note that the length of the shortest essential circuit in T_A is equal to $1/N$.

Since $0 < \varepsilon \leq 1/N$, it follows that $T_A \in \mathrm{cv}^1_{N,\varepsilon}$. Since $\mu_{T_A} = \nu_A$ and $\langle S, \nu_A \rangle = \lambda_A(S)$, by Theorem 4.2 we have

$$0 < \delta_1 \leq \frac{\langle S, \mu_{T_A} \rangle}{\Lambda(T_A, S)} = \frac{\langle S, \nu_A \rangle}{\Lambda(T_A, S)} = \frac{\lambda_A(S)}{\Lambda(T_A, S)} \leq \frac{1}{N},$$

as required. $\square$

5. Extremal, generic and algebraic stretching factors for free group automorphisms

We recall the notions of extremal and generic stretching factors from Definition 1.5 in the Introduction:

Definition 5.1 (extremal and generic stretching factors of automorphisms). Let A be a free basis of F_N and let $\varphi \in \mathrm{Out}(F_N)$.

Define

$$\Lambda_A(\varphi) := \Lambda(T_A, T_A\varphi) = \sup_{w \neq 1} \frac{\|\varphi(w)\|_A}{\|w\|_A} = e^{d_L(T_A, T_A\varphi)},$$

and refer to $\Lambda_A(\varphi)$ as the *extremal stretching factor* for φ with respect to A.

Also, define $\lambda_A(\varphi) := \lambda_A(NT_A\varphi) = N\lambda_A(T_A\varphi)$.

Thus for a.e. trajectory $\xi = y_1 \cdots y_n \cdots$ of the simple nonbacktracking random walk on F_N with respect to A, we have

$$\lambda_A(\varphi) = \lim_{n\to\infty} \frac{\|\varphi(y_1y_2\cdots y_n)\|_A}{n} = \lim_{n\to\infty} \frac{\|\varphi(y_1y_2\cdots y_n)\|_A}{\|y_1y_2\cdots y_n\|_A}.$$

We call $\lambda_A(\varphi)$ the *generic stretching factor* of φ with respect to A.

First, we obtain, in a slightly restated form, Theorem 1.6 from the Introduction:

Theorem 5.2. *For every* $N \geq 2$, *there exists* $0 < \tau_N \leq 1$ *such that if* A *is a free basis of* F_N *and* $\varphi \in \mathrm{Out}(F_N)$ *then*

$$0 < \tau_N \leq \frac{\lambda_A(\varphi)}{\Lambda_A(\varphi)} \leq 1.$$

Proof. Let A be a free basis of F_N. Recall that, by definition, for $\varphi \in \mathrm{Out}(F_N)$ we have $\lambda_A(\varphi) = N\lambda_A(T_A\varphi)$ and $\Lambda_A(\varphi) = \Lambda(T_A, T_A\varphi)$. Therefore, by Lemma 4.4, we have $\lambda_A(\varphi) \leq \Lambda_A(\varphi)$, so that $\lambda_A(\varphi)/\Lambda_A(\varphi) \leq 1$. Since for any $\varphi \in \mathrm{Out}(F_N)$, we have $T_A, T_A\varphi \in \mathrm{cv}^1_{N,\varepsilon}$ with $\varepsilon = 1/N$, the statement of the theorem now follows directly from Theorem 4.7. □

For two sequences $x_n > 0$, $y_n > 0$ (where $n \geq 1$), we say that x_n *grows like* y_n, if there exist $0 < c < c' < \infty$ such that for every $n \geq 1$, we have $c \leq x_n/y_n \leq c'$.

We now obtain Corollary 1.7 from the Introduction:

Corollary 5.3. *Let* $N \geq 2$ *and* $F_N = F(a_1, \ldots, a_n)$ *with* $A = \{a_1, \ldots, a_N\}$. *There exists* $D = D(N) \geq 1$ *such that for every* $\varphi \in \mathrm{Out}(F_N)$, *we have*

$$\frac{1}{D}\log\lambda_A(\varphi) \leq \log\lambda_A(\varphi^{-1}) \leq D\log\lambda_A(\varphi).$$

Proof. It follows from [Algom-Kfir and Bestvina 2012, Theorem 24] that there exists $D' = D'(N) \geq 1$ such that for every $\varphi \in \mathrm{Out}(F_N)$, we have

$$\frac{1}{D'}d_L(T_A, T_A\varphi) \leq d_L(T_A\varphi, T_A) \leq D'd_L(T_A, T_A\varphi).$$

Note that $d_L(T_A, T_A\varphi) = \log\Lambda(T_A, T_A\varphi) = \log\Lambda_A(\varphi)$ and that

$$d_L(T_A\varphi, T_A) = d_L(T_A, T_A\varphi^{-1}) = \log\Lambda(T_A, T_A\varphi^{-1}) = \log\Lambda_A(\varphi^{-1}).$$

Theorem 5.2 now implies that there exists $D'' = D''(N) \geq 1$ such that for every $\varphi \in \mathrm{Out}(F_N)$, we have

$$(**) \qquad \frac{1}{D''} \log \lambda_A(\varphi) - D'' \leq \log \lambda_A(\varphi^{-1}) \leq D'' \log \lambda_A(\varphi) + D''.$$

It was proved in [Francaviglia 2009; Kapovich and Lustig 2010a] (and also follows from Theorem 5.2) that the set $\Omega_N := \{\lambda_A(\varphi) \mid \varphi \in \mathrm{Out}(F_N)\}$ is a discrete subset of $[1, \infty)$. It was established in [Kaimanovich et al. 2007] that $\lambda_A(\varphi) = 1$ if and only if φ is a *permutational automorphism* with respect to A, that is, if and only if, after a possible composition with an inner automorphism, φ is induced by a permutation of A, with possibly inverting some elements of A. Note that φ is permutational with respect to A if and only if φ^{-1} is permutational with respect to A, so that for $\varphi \in \mathrm{Out}(F_N)$, $\lambda_A(\varphi^{-1}) = 1$ if and only if $\lambda_A(\varphi) = 1$. It was also proved in [loc. cit.] that the minimum of $\lambda_A(\varphi)$, taken over all nonpermutational φ, is equal to $1 + (2N-3)/(2N^2 - N)$. Therefore $(**)$ implies that there exists $D = D(N) \geq 1$ such that for every nonpermutational $\varphi \in \mathrm{Out}(F_N)$, we have

$$(\diamondsuit) \qquad \frac{1}{D} \log \lambda_A(\varphi) \leq \log \lambda_A(\varphi^{-1}) \leq D \log \lambda_A(\varphi).$$

If φ is permutational, then so is φ^{-1}. In this case we have $\log \lambda_A(\varphi^{-1}) = \log \lambda_A(\varphi) = 0$ and $(\diamondsuit)$ holds as well. Thus $(\diamondsuit)$ holds for every $\varphi \in \mathrm{Out}(F_N)$, which completes the proof. □

Recall that for $\varphi \in \mathrm{Out}(F_N)$, the *algebraic stretching factor* $\lambda(\varphi)$ is defined as

$$\lambda(\varphi) = \sup_{w \in F_N, w \neq 1} \lim_{n \to \infty} \sqrt[n]{\|\varphi^n(w)\|_S},$$

where $S \in \mathrm{cv}_N$ is an arbitrary base-point. As noted earlier, this definition of $\lambda(\varphi)$ does not depend on the choice of $S \in \mathrm{cv}_N$. The algebraic stretching factor $\lambda(\varphi)$ can be read off from any relative train-track representative $f : \Gamma \to \Gamma$ of φ as the maximum of the Perron–Frobenius eigenvalues for any of the canonical irreducible diagonal blocks of the (nonnegative) transition matrix $M(f)$.

Corollary 5.5 below describes, given $\varphi \in \mathrm{Out}(F_N)$, the asymptotics of $\Lambda(S, S\varphi^n)$ as n tends to infinity (where $S \in \overline{\mathrm{cv}}_N$ is an arbitrary point, the choice of which does not affect this asymptotics). The statement of Corollary 5.5 is probably known to the experts. Since the proof is not yet available in the literature, and since we need Corollary 5.5 for the applications in this paper, we include the proof here.

Proposition 5.4. *Let* $\varphi \in \mathrm{Out}(F_N)$.

(1) *Let* $q \geq 1$ *and let* $\alpha = \varphi^q$ *admit an improved relative train-track* (in the sense of [Bestvina et al. 2000]) *representative* $f : \Gamma \to \Gamma$. *Put* $\lambda := 1$ *if* α *is polynomially growing* (*that is*, *if* f *has no exponentially growing strata*) *and otherwise let*

$\lambda > 1$ *be the largest Perron–Frobenius eigenvalue of the exponentially growing strata of* $f : \Gamma \to \Gamma$.

Then there exists an integer $m \geq 0$ *such that for every* $S \in \mathrm{cv}_N$, *there are some constants* $0 < C_1 \leq C_2 < \infty$ *such that for every* $n \geq 1$,

$$C_1 \lambda^{n/q} n^m \leq \Lambda(S, S\varphi^n) \leq C_2 \lambda^{n/q} n^m.$$

(2) *If* φ *admits a train-track representative* $f : \Gamma \to \Gamma$ *with an irreducible transition matrix and with the Perron–Frobenius eigenvalue* $\lambda > 1$, *then for every* $S \in \mathrm{cv}_N$, *there exist* $0 < C_1 \leq C_2 < \infty$ *such that for every* $n \geq 1$,

$$C_1 \lambda^n \leq \Lambda(S, S\varphi^n) \leq C_2 \lambda^n.$$

Proof. (1) Let $T \in \mathrm{cv}_N^1$ be the point corresponding to the improved relative train-track $f : \Gamma \to \Gamma$, where all edges of Γ are given equal length. Put $L = \{1\}$ if f has no exponentially growing strata. Otherwise let $\lambda_1 \geq \cdots \geq \lambda_k > 1$ be all the Perron–Frobenius eigenvalues of the exponentially growing strata of f and put $L = \{\lambda_1, \ldots, \lambda_k, 1\}$. Finally put $\lambda = \max L$. Thus $\lambda \geq 1$ and $\lambda = 1$ if and only if f has no exponential strata.

A result of Levitt [2009, Theorem 6.2] shows that there is a finite subset M of $\mathbb{Z}_{\geq 0}$ such that for every nontrivial $w \in F_N$, there is some $(\lambda', m') \in L \times M$ such that the sequence $\|\alpha^n(w)\|_T$ grows like $(\lambda')^n n^{m'}$. Moreover, there exists some element $1 \neq w_0 \in F_N$ such that $\|\alpha^n(w_0)\|_T$ grows as $\lambda^n n^m$ and such that if some other $w \neq 1$ has $\|\alpha^n(w)\|_T$ growing as $\lambda^n n^{m'}$ then $m' \leq m$.

Let $D = C_\Delta$ be the finite subset of F_N as in Remark 2.1, where Δ is the open simplex in cv_N^1 containing T. Therefore for every $n \geq 1$, we have $\Lambda(T, T\varphi^n) = \max_{w \in D}(\|\alpha^n(w)\|_T / \|w\|_T)$. Moreover, through replacing D by $D \cup \{w_0\}$, we can assume that $w_0 \in D$.

It follows that $\Lambda(T, T\alpha^n) = \max_{w \in D}(\|\alpha^n(w)\|_T / \|w\|_T)$ grows like $\lambda^n n^m$.

Now let $n \geq 1$ and write $n = qn_1 + r$, where $n_1 \geq 0$ and $0 \leq r \leq q-1$ are integers. As we have seen, $\Lambda(T, T\alpha^{n_1}) = \max_{w \in D}(\|\varphi^{n_1}(w)\|_T / \|w\|_T)$ grows like $\lambda^{n_1} n_1^m$. Since $0 \leq r \leq q-1$, applying φ^r distorts $\|\cdot\|_T$ by a bounded multiplicative amount. Therefore $\Lambda(T, T\varphi^n) = \max_{w \in D}(\|\varphi^n(w)\|_T / \|w\|_T)$ grows as $\lambda^{n/q}(n/q)^m$, that is, as $\lambda^{n/q} n^m$.

Since T and S are F_N-equivariantly quasi-isometric, it follows that $\Lambda(S, S\varphi^n) = \Lambda_A(\varphi^n)$ also grows like $\lambda^{n/q} n^m$, and the conclusion of part (1) of the proposition follows.

(2) The proof of part (2) is known (e.g., see Theorem 8.1 in [Francaviglia and Martino 2011]) and is simpler than the proof of part (1), and we leave the details to the reader. The key point is that in this case for every nontrivial $w \in F_N$ such that the conjugacy class of w is not φ-periodic, the sequence $\|\varphi^n(w)\|_S$ grows like λ^n. □

Corollary 5.5. *Let $\varphi \in \mathrm{Out}(F_N)$, let $S \in \mathrm{cv}_N$ and let $\lambda(\varphi)$ be the algebraic stretching factor of φ.*

Then there is an integer $m \geq 0$ such that for every $S \in \mathrm{cv}_N$, there are some $C_1, C_2 > 0$ such that

$$C_1\, \lambda(\varphi)^n\, n^m \leq \Lambda(S, S\varphi^n) \leq C_2\, \lambda(\varphi)^n\, n^m$$

for all $n \geq 1$.

Proof. It is known [Bestvina et al. 2000] that some positive power $\alpha = \varphi^q$ of φ admits an improved relative train track representative.

In this case we have $\lambda(\alpha) = \lambda(\varphi^q) = \lambda(\varphi)^q$, so that $[\lambda(\alpha)]^{1/q} = \lambda(\varphi)$. The conclusion of the corollary now follows directly from part (1) of Proposition 5.4. □

Now Corollary 5.5 (applied to $S = T_A$, which gives $\Lambda(S, S\varphi^n) = \Lambda_A(\varphi^n)$) and Theorem 5.2 directly imply Theorem 1.8 from the Introduction:

Theorem 5.6. *Let A be a free basis of F_N, let $\varphi \in \mathrm{Out}(F_N)$ and let $\lambda(\varphi)$ be the algebraic stretching factor of φ. Then there exist constants $c_1, c_2 > 0$ and an integer $m \geq 0$ such that for every $n \geq 1$, we have*

$$c_1\, \lambda(\varphi)^n\, n^m \leq \lambda_A(\varphi^n) \leq c_2\, \lambda(\varphi)^n\, n^m.$$

Moreover, if φ admits an expanding train-track representative with an irreducible transition matrix (*e.g., if φ is fully irreducible*), *then $m = 0$ and $\lambda(\varphi) > 1$.* □

Example 5.7. To demonstrate that the case $\lambda > 1, m > 0$ in Theorem 5.6 can indeed occur, we consider an example explained on p. 1138 in [Levitt 2009]. Let $N = 4$ and $F_4 = F(A)$ with $A = \{a_1, b_1, a_2, b_2\}$. Let an automorphism $\varphi : F(A) \to F(A)$ be given by

$$\varphi(a_1) = a_1 b_1, \quad \varphi(b_1) = a_1, \quad \varphi(a_2) = a_2 b_1 a_1, \quad \varphi(b_2) = a_2.$$

For the A-rose R_A, the map $f : R_A \to R_A$, given by the same formula as φ, is both a global train-track and a 2-strata relative train-track representative for φ. The bottom stratum is $\{a_1, b_1\}$ and the top stratum is $\{a_2, b_2\}$. The transition matrices for both strata are the same and are equal to $B = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, which has the Perron–Frobenius eigenvalue $\lambda = (1 + \sqrt{5})/2$. The transition matrix for f has the form $M = \left(\begin{smallmatrix} B & 0 \\ C & B \end{smallmatrix}\right)$, where $C = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. By iterating M one can see that $\|\varphi^n(a_2)\|_A$ grows like $n\lambda^n$. One can then show that in this case $\Lambda_A(\varphi^n)$ also grows as $n\lambda^n$. Therefore, by Theorem 5.2, $\lambda_A(\varphi^n)$ grows as $n\lambda^n$ as well.

6. Other examples of filling tree-current morphisms

The Patterson–Sullivan map $J_{PS} : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$, $T \mapsto \mu_T$, is just one, albeit natural and useful, example of a filling tree-current morphism. There are many other filling tree-current morphisms $J : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N)$, and Corollary 3.5 is applicable to all such J. We indicate here some sources of such J, following the approach of Reiner Martin [1995]. The main idea is that if $t \mapsto \rho(t) > 0$ is a monotone decreasing continuous function which approaches 0 as $t \to \infty$ "sufficiently quickly", then

$$J_\rho : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N), \qquad T \mapsto \sum_{[w]\neq[1]} \rho(\|w\|_T)\eta_w$$

is a filling tree-current morphism.

The summation here can be taken either over all nontrivial conjugacy classes $[w]$ of elements of F_N (or over an $\mathrm{Out}(F_N)$-invariant set of such conjugacy classes, although in the latter case one has to take additional care to ensure that the current $J_\rho(T)$ is filling).

Let us first observe that such a function J_ρ is, by its construction, always $\mathrm{Out}(F_N)$-equivariant: for any $T \in \mathrm{cv}_N^1$ and $\varphi \in \mathrm{Out}(F_N)$, we have

$$\varphi(J_\rho(T)) = \sum_{[w]\neq[1]} \rho(\|w\|_T)\varphi(\eta_w) = \sum_{[w]\neq[1]} \rho(\|w\|_T)\eta_{\varphi(w)}$$

and

$$\begin{aligned} J_\rho(\varphi T) &= \sum_{[w]\neq[1]} \rho(\|w\|_{\varphi T})\eta_w = \sum_{[w]\neq[1]} \rho(\|\varphi^{-1}(w)\|_T)\eta_w \\ &\underset{\text{with } u=\varphi^{-1}(w)}{=} \sum_{[u]\neq[1]} \rho(\|u\|_T)\eta_{\varphi(u)} = \varphi(J_\rho(T)), \end{aligned}$$

so that J_ρ is indeed $\mathrm{Out}(F_N)$-equivariant.

We provide here a representative result of the kind described above:

Proposition 6.1. *The function*

$$J : \mathrm{cv}_N^1 \to \mathrm{Curr}(F_N), \qquad T \mapsto \sum_{[w]\neq[1]} e^{-e^{\|w\|_T}} \eta_w,$$

where the sum is taken over all nontrivial root-free conjugacy classes $[w]$ *of elements of* F_N*, is an injective filling tree-current morphism.*

Proof. Fix a free basis A of F_N and let $T_A \in \mathrm{cv}_N^1$ be the Cayley graph of F_N with respect to A, where all edges in T_A have length $1/N$. For $w \in F_N$ denote by $\|w\|_A$ the cyclically reduced length of w over $A^{\pm 1}$. Thus $\|w\|_A = N\|w\|_{T_A}$. We let $R_A = T_A/F_N$ be the quotient metric graph, which is a wedge of N loop-edges of length $1/N$ corresponding to elements of A. Let $\alpha_A : F_N \to \pi_1(R_A)$ be the associated chart.

Let $T \in \mathrm{cv}_N^1$ be arbitrary and let U be a compact neighborhood of T in cv_N^1. There exists a constant $C \geq 1$ such that for every $w \in F_N$ and every $T' \in U$, we have $\|w\|_{T'}/C \le \|w\|_A \le C\|w\|_{T'}$. Note that for $n \ge 1$, the number of conjugacy classes $[w]$ with $\|w\|_A \le n$ is at most $(2N)^n$.

To show that for each $T' \in U$, $J(T')$ is a geodesic current we only need to verify that $J(T')$ takes finite values on all the two-sided cylinder sets in $\partial^2 F_N$ determined by the chart α_A. Since every cylinder is contained in a cylinder determined by a single edge, it suffices to show that for every oriented edge e of R_A, we have $\langle e, J(T')\rangle_{\alpha_A} < \infty$.

Let $T' \in U$ and let e be an edge of R_A. For every integer $n \ge 1$, set

$$b_n(e,T') := \sum_{0.9n \le \|[w]\|_A \le 1.1n} e^{-e^{\|w\|_{T'}}} \langle e, \eta_w\rangle_{\alpha_A}.$$

Then $\langle e, J(T')\rangle_{\alpha_A} \le \sum_{n=1}^{\infty} b_n(e,T')$. The weight $\langle e, \eta_w\rangle_{\alpha_A}$ is equal to $1/N$ times the number of occurrences of $e^{\pm 1}$ in the cyclically reduced circuit γ_w in R_A representing $[w]$. Hence $\langle e, \eta_w\rangle_{\alpha_A} \le (1/N)\|w\|_A$. Since $T' \in U$, we have $\|w\|_{T'} \ge \|w\|_A/C$. Hence for every $n \ge 1$ and $T' \in U$, we have

$$\begin{aligned} b_n(e,T') &= \sum_{\|[w]\|_A \in I} e^{-e^{\|w\|_{T'}}} \langle e, \eta_w\rangle_{\alpha_A} \le \frac{1}{N} \sum_{\|[w]\|_A \in I} e^{-e^{\|w\|_A/C}} \|w\|_A \\ &\le \frac{1}{N} \sum_{\|[w]\|_A \in I} e^{-e^{0.9n/C}} 1.1n \;\le \frac{1.1n}{N} e^{-e^{0.9n/C}} (2N)^{1.1n} \\ &= \frac{1.1n}{N} e^{-e^{0.9n/C}} e^{1.1n\log(2N)} \;= \frac{1.1n}{N} e^{1.1n\log(2N) - e^{0.9n/C}}, \end{aligned}$$

where $I = [0.9n, 1.1n]$. From here we see that

$$\langle e, J(T')\rangle_{\alpha_A} \le \sum_{n=1}^{\infty} b_n(e,T') \le C_1,$$

where $C_1 = C_1(U) < \infty$ is some constant depending only on U.

Thus for every $T' \in U$, $J(T')$ is indeed a geodesic current on F_N, and, in particular, $J(T) \in \mathrm{Curr}(F_N)$.

Note that the current $J(T)$ has full support. Indeed, for every nontrivial freely reduced word v over $A^{\pm 1}$, there exists a root-free cyclically reduced word w over $A^{\pm 1}$ containing v as a subword. Then $\langle v, \eta_w\rangle_{\alpha_A} > 0$ and hence, from the definition of $J(T)$, we see that $\langle v, J(T)\rangle_{\alpha_A} > 0$. Thus indeed $J(T)$ has full support and therefore, by a result of Kapovich and Lustig [2010a], the current $J(T)$ is filling.

Since an automorphism of F_N permutes the set of all root-free nontrivial conjugacy classes in F_N, it follows from the definition of J that for every $T \in \mathrm{cv}_N^1$ and every $\varphi \in \mathrm{Out}(F_N)$, we have $J(\varphi T) = \varphi J(T)$.

Thus we have constructed an $\mathrm{Out}(F_N)$-equivariant map $J : \mathrm{cv}_N^1 \to \mathrm{Curr}_{\mathrm{fill}}(F_N)$.

We next observe that the map J is continuous. The proof of the continuity of J is similar to the proof that $J(T)$ is a current. Let $T \in \mathrm{cv}_N^1$, let U be a compact neighborhood of T in cv_N^1 and let v be a nontrivial freely reduced word over $A^{\pm 1}$. Then for every $T' \in U$, we have

$$\langle v, T'\rangle_{\alpha_A} = \sum_{[w]} \langle v, e^{-e^{\|w\|_{T'}}} \eta_w\rangle_{\alpha_A} = \sum_{[w]} e^{-e^{\|w\|_{T'}}} \langle v, w\rangle_{\alpha_A}.$$

One can then show, by an argument similar to that used above, that there exist positive constants $M_w > 0$ (also depending on U and v but independent of $T' \in U$) such that for every $T' \in U$, we have $e^{-e^{\|w\|_{T'}}}\langle v, w\rangle_{\alpha_A} \le M_w$ and that $\sum_{[w]} M_w < \infty$. By the Weierstrass M-test, it follows that the series

$$\sum_{[w]} e^{-e^{\|w\|_{T'}}} \langle v, w\rangle_{\alpha_A},$$

viewed as the sum of a functions on U, converges uniformly on U and that its sum $\langle v, T'\rangle_{\alpha_A}$ is a continuous function on U.

Since v was arbitrary, the explicit description of the topology on $\mathrm{Curr}(F_N)$ (see [Kapovich 2006]) implies that J is a continuous function on cv_N^1, as required.

It remains to show that J is injective. Fix an enumeration, without repetitions, $w_1, w_2, \dots$ of representatives of all the nontrivial root-free conjugacy classes in F_N. Thus for every root-free nontrivial $w \in F_N$, there exist unique distinct $m, n \ge 1$ such that $[w] = [w_m]$ and $[w^{-1}] = [w_n]$.

For every $i \ge 1$, set $q_i = (w_i^{-\infty}, w_i^{\infty}) \in \partial^2 F_N$ and set $Q_i = \{q_i\}$. Note that for $i, j \ge 1$, we have $\eta_{w_j}(Q_i) = 1$ if $[w_i] = [w_j^{\pm 1}]$ and $\eta_{w_i}(Q_i) = 0$ otherwise. Then, by definition of J, for every $T \in \mathrm{cv}_N^1$ and $i \ge 1$, we have $J(T)(Q_i) = 2e^{-e^{\|w_i\|_T}}$. Since the function $t \mapsto 2e^{-e^t}$ is strictly monotone and thus injective, it follows that knowing the current $J(T)$, we can recover $\|w_i\|_T$ for all $i \ge 1$. Hence we can recover the length function $\|\cdot\|_T : F_N \to \mathbb{R}$ and so we can also recover T itself. Thus J is injective, as required. □

7. Open problems

As we have seen in Theorem 1.6, if $N \ge 2$, $A = \{a_1, \dots, a_N\}$ is a fixed free basis of $F_N = F(A)$, then for

$$\rho_N = \inf_{\varphi \in \mathrm{Out}(F_N)} \frac{\lambda_A(\varphi)}{\Lambda_A(\varphi)},$$

we have $\rho_N > 0$. In fact, one can show:

Proposition 7.1. *We have* $\lim_{N\to\infty} \rho_N = 0$, *and moreover*, $\rho_N = O(1/N)$; *that is*, $\limsup_{N\to\infty} N\rho_N < \infty$.

Proof. For $N \geq 2$ and $m \geq 1$, let $\varphi_{N,m} : F(A) \to F(A)$ be given by $\varphi_{N,m}(a_1) = a_1 a_2^m$ and $\varphi_{N,m}(a_i) = a_i$ for $2 \leq i \leq N$. It is not hard to see that

$$\Lambda_A(\varphi_{N,m}) = \sup_{w \neq 1} \frac{\|\varphi_{N,m}(w)\|_A}{\|w\|_A} = m + 1.$$

For any freely reduced $w \in F(A)$, we have

$$\|\varphi_{N,m}(w)\|_A \leq (m+1)(a_1; w)_A + \sum_{i=2}^{N} (a_i; w)_A,$$

where $(a_j; w)_A$ is the number of occurrences of $a_j^{\pm 1}$ in w. On the other hand, if $w_n \in F(A)$ is a "long random" freely reduced word of length n, then asymptotically we have $(a_i; w_n)_A/n \xrightarrow{n\to\infty} 1/N$ for $i = 1, \ldots, N$. Therefore

$$\begin{aligned}\lambda_A(\varphi_{N,m}) &\leq \lim_{n\to\infty} \frac{(m+1)(a_1; w)_A + \sum_{i=2}^{N}(a_i; w)_A}{n} \\ &= (m+1)\frac{1}{N} + \frac{N-1}{N} = \frac{m}{N} + 1.\end{aligned}$$

Hence

$$\rho_N \leq \frac{\lambda_A(\varphi_{N,m})}{\Lambda_A(\varphi_{N,m})} \leq \frac{1 + \frac{m}{N}}{m+1}.$$

By taking $m = N$, we see that $\rho_N \leq 2/(N+1) \xrightarrow{n\to\infty} 0$. Thus $\lim_{N\to\infty} \rho_N = 0$ and $\limsup_{N\to\infty} N\rho_N < \infty$. □

Theorem 1.6 and Proposition 7.1 naturally raise the following:

Problem 7.2. Are the values ρ_N algorithmically computable in terms of N? What are the exact values of ρ_N for small N, say for $N = 2, 3, 4$? Is it true that $\rho_N \in \mathbb{Q}$? What can be said about the precise asymptotics of ρ_N as $N \to \infty$? (Note that Proposition 7.1 shows that ρ_N decays at least as fast as $1/N$.)

Theorem 1.1 also motivates the definition of a new notion of a continuous symmetric and $\mathrm{Out}(F_N)$-invariant intersection number $I : \mathrm{cv}_N^1 \times \mathrm{cv}_N^1 \to \mathbb{R}_{>0}$, where for $T, S \in \mathrm{cv}_N^1$, we define $I(T, S) := \langle S, \mu_T \rangle \langle T, \mu_S \rangle$. The function $I(\cdot, \cdot)$ was originally suggested to us by Arnaud Hilion, as it appears to be relevant for attempting to define an analogue of the Weil–Petersson metric on cv_N^1.

Since the Patterson–Sullivan currents are normalized so that $\langle T, \mu_T \rangle = 1$, for $T = S$, we have $I(T, T) = 1$.

Problem 7.3. (a) Is it true that for every $T, S \in \mathrm{cv}_N^1$, we have $I(T, S) \geq 1$?

(b) Is it true that for $T, S \in \mathrm{cv}_N^1$, we have $I(T, S) = 1$ if and only if $T = S$?

It was shown in [Kaimanovich et al. 2007] that if A is a free basis of F_N and $\varphi \in \mathrm{Out}(F_N)$ then $\lambda_A(\varphi) \geq 1$ and that $\lambda_A(\varphi) = 1$ if and only if $T_A\varphi = T_A$. If

B is another free basis of F_N and $\varphi \in \mathrm{Aut}(F_N)$ is such that $T_A\varphi = T_B$, then $\langle T_B, \mu_{T_A}\rangle = \lambda_A(\varphi)$ and $\langle T_A, \mu_{T_B}\rangle = \lambda_A(\varphi^{-1})$. It follows that if A, B are free bases of F_N then $I(T_A, T_B) \geq 1$ and that $I(T_A, T_B) = 1$ if and only if $T_A = T_B$. However, beyond this fact nothing appears to be known about the above question.

Recently Pollicott and Sharp [2014], using a different approach, defined and studied a Weil–Petersson type metric on cv_N^1. It would be interesting to investigate the relationship of their metric to the quantity $I(T, S)$ defined above.

Acknowledgements

We thank Matt Clay and Camille Horbez for useful discussions about Guirardel's intersection number. We are also grateful to Brian Ray and Paul Schupp for conducting helpful computer experiments with generic stretching factors of free group automorphisms.

References

[Algom-Kfir 2011] Y. Algom-Kfir, "Strongly contracting geodesics in outer space", *Geom. Topol.* **15**:4 (2011), 2181–2233. MR 2862155 Zbl 1250.20019

[Algom-Kfir 2013] Y. Algom-Kfir, "The metric completion of outer space", preprint, 2013. arXiv 1202.6392

[Algom-Kfir and Bestvina 2012] Y. Algom-Kfir and M. Bestvina, "Asymmetry of outer space", *Geom. Dedicata* **156** (2012), 81–92. MR 2863547 Zbl 1271.20029

[Behrstock et al. 2010] J. Behrstock, M. Bestvina, and M. Clay, "Growth of intersection numbers for free group automorphisms", *J. Topol.* **3**:2 (2010), 280–310. MR 2011j:20072 Zbl 1209.20031

[Bestvina 2011] M. Bestvina, "A Bers-like proof of the existence of train tracks for free group automorphisms", *Fund. Math.* **214**:1 (2011), 1–12. MR 2012m:20046 Zbl 1248.20025

[Bestvina and Feighn 1993] M. Bestvina and M. Feighn, "Outer limits", preprint, 1993, Available at http://andromeda.rutgers.edu/~feighn/papers/outer.pdf.

[Bestvina and Feighn 2010] M. Bestvina and M. Feighn, "A hyperbolic Out(F_n)-complex", *Groups Geom. Dyn.* **4**:1 (2010), 31–58. MR 2011a:20052 Zbl 1190.20017

[Bestvina and Reynolds 2012] M. Bestvina and P. Reynolds, "The boundary of the complex of free factors", preprint, 2012. arXiv 1211.3608

[Bestvina et al. 2000] M. Bestvina, M. Feighn, and M. Handel, "The Tits alternative for Out(F_n), I: Dynamics of exponentially-growing automorphisms", *Ann. of Math.* (2) **151**:2 (2000), 517–623. MR 2002a:20034 Zbl 0984.20025

[Carette et al. 2012] M. Carette, S. Francaviglia, I. Kapovich, and A. Martino, "Spectral rigidity of automorphic orbits in free groups", *Algebr. Geom. Topol.* **12**:3 (2012), 1457–1486. MR 2966693 Zbl 1261.20040

[Clay and Pettet 2010] M. Clay and A. Pettet, "Twisting out fully irreducible automorphisms", *Geom. Funct. Anal.* **20**:3 (2010), 657–689. MR 2011i:20063 Zbl 1206.20047

[Clay and Pettet 2012a] M. Clay and A. Pettet, "Current twisting and nonsingular matrices", *Comment. Math. Helv.* **87**:2 (2012), 385–407. MR 2914853 Zbl 1286.20049

[Clay and Pettet 2012b] M. Clay and A. Pettet, “Relative twisting in outer space”, *J. Topol. Anal.* **4**:2 (2012), 173–201. MR 2949239 Zbl 1260.57002

[Clay et al. 2015] M. Clay, J. Mangahas, and A. Pettet, “An algorithm to detect full irreducibility by bounding the volume of periodic free factors”, *Michigan Math. J.* **64**:2 (2015), 279–292. MR 3359026

[Cohen and Lustig 1995] M. M. Cohen and M. Lustig, “Very small group actions on **R**-trees and Dehn twist automorphisms”, *Topology* **34**:3 (1995), 575–617. MR 96g:20053 Zbl 0844.20018

[Coornaert 1993] M. Coornaert, “Mesures de Patterson–Sullivan sur le bord d’un espace hyperbolique au sens de Gromov”, *Pacific J. Math.* **159**:2 (1993), 241–270. MR 94m:57075 Zbl 0797.20029

[Coulbois and Hilion 2014] T. Coulbois and A. Hilion, “Ergodic currents dual to a real tree”, *Ergodic Theory and Dynamical Systems* (online publication November 2014).

[Coulbois et al. 2008a] T. Coulbois, A. Hilion, and M. Lustig, “$\mathbb{R}$-trees and laminations for free groups, II: The dual lamination of an $\mathbb{R}$-tree”, *J. Lond. Math. Soc.* (2) **78**:3 (2008), 737–754. MR 2010h:20056 Zbl 1198.20023

[Coulbois et al. 2008b] T. Coulbois, A. Hilion, and M. Lustig, “$\mathbb{R}$-trees and laminations for free groups, III: Currents and dual $\mathbb{R}$-tree metrics”, *J. Lond. Math. Soc.* (2) **78**:3 (2008), 755–766. MR 2010h:20057 Zbl 1200.20018

[Culler and Vogtmann 1986] M. Culler and K. Vogtmann, “Moduli of graphs and automorphisms of free groups”, *Invent. Math.* **84**:1 (1986), 91–119. MR 87f:20048 Zbl 0589.20022

[Francaviglia 2009] S. Francaviglia, “Geodesic currents and length compactness for automorphisms of free groups”, *Trans. Amer. Math. Soc.* **361**:1 (2009), 161–176. MR 2009h:20044 Zbl 1166.20032

[Francaviglia and Martino 2011] S. Francaviglia and A. Martino, “Metric properties of outer space”, *Publ. Mat.* **55**:2 (2011), 433–473. MR 2012j:20128 Zbl 1268.20042

[Francaviglia and Martino 2012] S. Francaviglia and A. Martino, “The isometry group of outer space”, *Adv. Math.* **231**:3-4 (2012), 1940–1973. MR 2964629 Zbl 06094103

[Furman 2002] A. Furman, “Coarse-geometric perspective on negatively curved manifolds and groups”, pp. 149–166 in *Rigidity in dynamics and geometry* (Cambridge, 2000), edited by M. Burger and A. Iozzi, Springer, Berlin, 2002. MR 2003f:53062 Zbl 1064.53025

[Guirardel 1998] V. Guirardel, “Approximations of stable actions on **R**-trees”, *Comment. Math. Helv.* **73**:1 (1998), 89–121. MR 99e:20037 Zbl 0979.20026

[Guirardel 2005] V. Guirardel, “Cœur et nombre d’intersection pour les actions de groupes sur les arbres”, *Ann. Sci. École Norm. Sup.* (4) **38**:6 (2005), 847–888. MR 2007e:20055 Zbl 1110.20019

[Hamenstädt 2014a] U. Hamenstädt, “The boundary of the free splitting graph and the free factor graph”, preprint, 2014. arXiv 1211.1630

[Hamenstädt 2014b] U. Hamenstädt, “Lines of minima in outer space”, *Duke Math. J.* **163**:4 (2014), 733–776. MR 3178431 Zbl 06288360

[Horbez 2012] C. Horbez, “Sphere paths in outer space”, *Algebr. Geom. Topol.* **12**:4 (2012), 2493–2517. MR 3020214 Zbl 1262.57005

[Kaimanovich 1991] V. A. Kaimanovich, “Bowen–Margulis and Patterson measures on negatively curved compact manifolds”, pp. 223–232 in *Dynamical systems and related topics* (Nagoya, 1990), edited by K. Shiraiwa, Adv. Ser. Dynam. Systems **9**, World Sci. Publ., River Edge, NJ, 1991. MR 93h:58118

[Kaimanovich et al. 2007] V. Kaimanovich, I. Kapovich, and P. Schupp, “The subadditive ergodic theorem and generic stretching factors for free group automorphisms”, *Israel J. Math.* **157** (2007), 1–46. MR 2009d:20099 Zbl 1173.20031

[Kapovich 2005] I. Kapovich, “The frequency space of a free group”, *Internat. J. Algebra Comput.* **15**:5-6 (2005), 939–969. MR 2007a:20038 Zbl 1110.20031

[Kapovich 2006] I. Kapovich, “Currents on free groups”, pp. 149–176 in *Topological and asymptotic aspects of group theory*, edited by R. Grigorchuk et al., Contemp. Math. **394**, Amer. Math. Soc., Providence, RI, 2006. MR 2007k:20094 Zbl 1110.20034

[Kapovich and Lustig 2007] I. Kapovich and M. Lustig, “The actions of $\mathrm{Out}(F_k)$ on the boundary of outer space and on the space of currents: Minimal sets and equivariant incompatibility”, *Ergodic Theory Dynam. Systems* **27**:3 (2007), 827–847. MR 2008h:20051 Zbl 1127.20025

[Kapovich and Lustig 2009] I. Kapovich and M. Lustig, “Geometric intersection number and analogues of the curve complex for free groups”, *Geom. Topol.* **13**:3 (2009), 1805–1833. MR 2010h:20092 Zbl 1194.20046

[Kapovich and Lustig 2010a] I. Kapovich and M. Lustig, “Intersection form, laminations and currents on free groups”, *Geom. Funct. Anal.* **19**:5 (2010), 1426–1467. MR 2011g:20052 Zbl 1242.20052

[Kapovich and Lustig 2010b] I. Kapovich and M. Lustig, “Ping-pong and outer space”, *J. Topol. Anal.* **2**:2 (2010), 173–201. MR 2011d:20055 Zbl 1211.20027

[Kapovich and Nagnibeda 2007] I. Kapovich and T. Nagnibeda, “The Patterson–Sullivan embedding and minimal volume entropy for outer space”, *Geom. Funct. Anal.* **17**:4 (2007), 1201–1236. MR 2009c:20073 Zbl 1135.20031

[Kapovich and Nagnibeda 2010] I. Kapovich and T. Nagnibeda, “Geometric entropy of geodesic currents on free groups”, pp. 149–175 in *Dynamical numbers—interplay between dynamical systems and number theory*, edited by S. Kolyada et al., Contemp. Math. **532**, Amer. Math. Soc., Providence, RI, 2010. MR 2762139 Zbl 1216.20034

[Ladra et al. 2015] M. Ladra, P. V. Silva, and E. Ventura, “Bounding the gap between a free group (outer) automorphism and its inverse”, *Collectanea Mathematica* (online publication February 2015).

[Levitt 2009] G. Levitt, “Counting growth types of automorphisms of free groups”, *Geom. Funct. Anal.* **19**:4 (2009), 1119–1146. MR 2011f:20068 Zbl 1196.20038

[Lyons 1994] R. Lyons, “Equivalence of boundary measures on covering trees of finite graphs”, *Ergodic Theory Dynam. Systems* **14**:3 (1994), 575–597. MR 95g:58132 Zbl 0821.58008

[Mann and Reynolds 2013] B. Mann and P. Reynolds, “Constructing non-uniquely ergodic arational trees”, preprint, 2013. arXiv 1311.1771

[Martin 1995] R. Martin, *Non-uniquely ergodic foliations of thin-type, measured currents and automorphisms of free groups*, Ph.D. thesis, University of California, Los Angeles, 1995, Available at http://search.proquest.com/docview/304185823.

[Paulin 1989] F. Paulin, “The Gromov topology on **R**-trees”, *Topology Appl.* **32**:3 (1989), 197–221. MR 90k:57015 Zbl 0675.20033

[Pollicott and Sharp 2014] M. Pollicott and R. Sharp, “A Weil–Petersson type metric on spaces of metric graphs”, *Geom. Dedicata* **172** (2014), 229–244. MR 3253781 Zbl 1301.30043

[Reynolds 2012] P. Reynolds, “Reducing systems for very small trees”, preprint, 2012. arXiv 1211.3378

[Sharp 2010] R. Sharp, “Distortion and entropy for automorphisms of free groups”, *Discrete Contin. Dyn. Syst.* **26**:1 (2010), 347–363. MR 2011e:37073 Zbl 1220.37018

[White 1991] T. P. White, *The geometry of the outer space*, Ph.D. thesis, University of California, Los Angeles, 1991, Available at http://search.proquest.com/docview/303915773.

Received August 26, 2014. Revised February 12, 2015.

ILYA KAPOVICH
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
1409 WEST GREEN STREET
URBANA, IL 61801
UNITED STATES
kapovich@math.uiuc.edu

MARTIN LUSTIG
CENTRE DE MATHÉMATIQUES ET INFORMATIQUE
AIX-MARSEILLE UNIVERSITÉ
39, RUE F. JOLIOT CURIE
13453 MARSEILLE 13
FRANCE
martin.lustig@univ-amu.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.399

ON RECURRENCE OVER SUBSETS AND WEAK MIXING

JIAN LI, PIOTR OPROCHA AND GUOHUA ZHANG

We study properties of weakly mixing sets (of order n) in relation to proximality, sensitivity, scrambled tuples, Xiong chaotic sets and independent sets. Our main emphasis is on the structure of the set of transfer times $N(U \cap A, V)$ between open sets U and V, both intersecting a weakly mixing set A. We find several conditions on properties of the set A that are equivalent to weak mixing.

We also prove that on topological graphs weakly mixing sets of order 2 can be approximated arbitrarily closely by a weakly mixing set of all orders. This property is known to hold on the unit interval but is not true in general (there are systems with weakly mixing sets of order n but not $n+1$).

1. Introduction

This paper is a continuation of the previous papers by Oprocha and Zhang on local aspects of topological weak mixing [2011; 2012; 2013; 2014] in dynamical systems (X, f), that is, continuous maps $f : X \to X$ acting on compact metric spaces. When defining recurrent properties of dynamical systems, it is convenient to analyze properties of transfer times between sets, expressed in terms of the set

$$N(U, V) = \{n \in \mathbb{N} : f^n(U) \cap V \neq \varnothing\},$$

where U and V are nonempty open subsets of X. For example, (X, f) is *transitive* if $N(U, V)$ is nonempty for any choice of two nonempty open sets. As mentioned before, the main concept in this paper is topological weak mixing (in fact, its local versions), which is usually defined as transitivity of $(X \times X, f \times f)$. In other words, (X, f) is weakly mixing if $N(U_1, V_1) \cap N(U_2, V_2) \neq \varnothing$ for any choice of four nonempty open sets $U_1, U_2, V_1, V_2 \subset X$. It was shown by Furstenberg [1967] that if (X, f) is weakly mixing then for every $n \geq 2$ and any nonempty open sets $U_1, U_2, \ldots, U_n, V_1, V_2, \ldots, V_n \subset X$ we have

$$N(U_1, V_1) \cap N(U_2, V_2) \cap \cdots \cap N(U_n, V_n) \neq \varnothing.$$

Guohua Zhang is the corresponding author.

MSC2010: primary 37B05; secondary 37B40, 37E05, 37E25.

Keywords: weakly mixing sets, Xiong chaotic sets, proximality, scrambled tuples, independent sets, topological graphs.

Note that weak mixing can be regarded as a 'global' property, while topological entropy is a 'local' one since it can be supported on a small set in the space (e.g., a nowhere dense attractor). It is also not hard to see that in general there cannot be any implication between weak mixing and positive topological entropy. Therefore an appropriate 'local' version of weak mixing is needed. Such a concept was introduced in [Blanchard and Huang 2008]. Strictly speaking, a nontrivial closed set $A \subset X$ (i.e., not a singleton) is *weakly mixing* if for every $n \geq 2$ and any nonempty open sets $U_1, U_2, \dots, U_n, V_1, V_2, \dots, V_n \subset X$ intersecting A (i.e., $U_i \cap A \neq \varnothing$ and $V_i \cap A \neq \varnothing$ for each $i = 1, 2, \dots, n$) we have $N(U_1 \cap A, V_1) \cap N(U_2 \cap A, V_2) \cap \cdots \cap N(U_n \cap A, V_n) \neq \varnothing$. As we can see, the above definition is consistent with the definition of a weakly mixing map and, more importantly, it is proved in [Blanchard and Huang 2008] that every dynamical system with positive topological entropy contains many Cantor weakly mixing sets.

Similarly, for a fixed integer $n \geq 2$, we say that a nontrivial closed subset A of X is *weakly mixing of order n* if for any nonempty open subsets $U_1, U_2, \dots, U_n, V_1, V_2, \dots, V_n$ of X intersecting A, we have $N(U_1 \cap A, V_1) \cap N(U_2 \cap A, V_2) \cap \cdots \cap N(U_n \cap A, V_n) \neq \varnothing$. Unfortunately, the analog of Furstenberg's theorem cannot be proved here. Namely, it is proved in [Oprocha and Zhang 2011; 2014] that for every $n \geq 2$ there exists a dynamical system which contains weakly mixing sets of order n but no weakly mixing sets of order $n + 1$.

Since Furstenberg's theorem does not work for weakly mixing sets of order n, it is natural to ask which criteria for weak mixing (i.e., equivalent conditions) can be used in the case of weakly mixing sets. It was proved in [Banks 1999] that most of the conditions that can be expressed in terms of intersections of sets $N(U, V)$ lead to weak mixing. Of particular interest is the condition, proved first in [Petersen 1970], which says that a dynamical system is weakly mixing if and only if $N(U, V) \cap N(U, U) \neq \varnothing$ for any nonempty open sets $U, V \subset X$. In the spirit of the above fact, we find the following criterion for weak mixing of order n. It will be shown later, in Example 3.2, that we cannot use exactly the same condition as in [Petersen 1970].

Theorem 3.1. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset with $n \geq 2$. Then A is a weakly mixing set of order n if and only if for any $n + 1$ open subsets $U_1, V_1, V_2, \dots, V_n$ of X intersecting A,*

$$N(U_1 \cap A, V_1) \cap \bigcap_{i=2}^{n} N(V_i \cap A, V_i) \neq \varnothing.$$

It is shown in [Li 2011, Theorem 3.2] that if a dynamical system (X, f) is weakly mixing, then there exists a residual subset K of X such that for every $x \in K$ and every nonempty open subset U of X, the set $N(x, U)$ contains an IP-set.

This theorem was generalized in [Oprocha and Zhang 2013, Theorem 8], which states that if A is a weakly mixing set of order 2 and U is an open subset of X intersecting A, then there is an $x \in U \cap A$ such that for every open subset V of X intersecting A the set $N(x, V)$ contains an IP-set. Using the idea in the proof of [Li 2011, Theorem 3.2], we could extend the above fact from [Oprocha and Zhang 2013] a little further.

Theorem 3.4. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order n with $n \geq 2$. Then there exists a residual subset K of A such that for any $x \in K$ and any choice of $n-1$ open subsets $U_1, \dots, U_{n-1}$ of X intersecting A there exist points $y_i \in U_i \cap A$, where $i = 1, \dots, n-1$, such that $N(x, U_1) \cap \bigcap_{i=1}^{n-1} N(y_i, U_i)$ contains an IP-set.*

A subset A of X is *transitive* in (X, f) if, for any open subsets U and V of X intersecting A, the set $N(U \cap A, V)$ is not empty; A is *totally transitive* if it is transitive in (X, f^k) for every $k \in \mathbb{N}$. Let $n \geq 2$ be an integer. It is clear that a nontrivial closed subset $A \subset X$ is weakly mixing of order n if and only if A^n is a transitive set in $(X^n, f^{(n)})$. Using Theorem 3.4, we have the following result.

Proposition 3.6. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset.*

(1.3.1) *If A is a weakly mixing set of order 2, then A is totally transitive.*

(1.3.2) *If A is a weakly mixing set of order n for f with $n \geq 3$, then, for every $k \in \mathbb{N}$, A is a weakly mixing set of order $n-1$ for f^k.*

The authors in [Huang et al. 2012] proved that a dynamical system is weakly mixing if and only if it has the IP-independent property (a formal definition of independence will be given later). We will obtain a similar result for the case of weakly mixing sets.

Inspired by the result of Xiong and Yang [1991], Blanchard and Huang [2008] provided an alternative definition of a weakly mixing set. Strictly speaking, it was proved in [Blanchard and Huang 2008] that a nontrivial closed set $A \subset X$ is a weakly mixing set if and only if there exists a dense Mycielski subset B of A such that for any $C \subset B$ and any continuous map $g : C \to A$ there exists an increasing sequence of natural numbers $\{n_i\}_{i=1}^{\infty}$ for which $\lim_{i\to\infty} f^{n_i}(x) = g(x)$ for any $x \in C$.

Similarly, we can introduce Xiong chaotic sets of a finite order as follows. A subset K of X with at least n points is called a *Xiong chaotic set of order n* if for any subset E of K with cardinality n and for any map $g : E \to \overline{K}$ there is an increasing subsequence $\{q_i\}_{i=1}^{\infty}$ in $\mathbb{N}$ such that $\lim_{i\to\infty} f^{q_i}(x) = g(x)$ for every $x \in E$. Later we will show that a result analogous to [Blanchard and Huang 2008] holds; that is, any nontrivial closed subset A of X is a weakly mixing set of order n if and only if there exists a dense Mycielski subset S of A which is Xiong chaotic of order n. A dynamical system has a weakly mixing set of order n if and only if it has an uncountable Xiong chaotic set of order n. An advantage of Xiong chaotic sets is that they are hereditary by subsets, while weakly mixing sets are not.

For a dynamical system (X, f), the *proximal relation* is

$$\mathrm{Prox}_2(f) = \big\{(x, y) \in X \times X : \liminf_{k\to\infty} d(f^k(x), f^k(y)) = 0\big\},$$

and the *proximal cell* of a point $x \in X$ is $\mathrm{Prox}_2(f)(x) = \{y \in X : (x, y) \in \mathrm{Prox}_2(f)\}$. It was shown in [Akin and Kolyada 2003] that if (X, f) is weakly mixing, then, for every $x \in X$, the proximal cell $\mathrm{Prox}_2(f)(x)$ of x is residual in X. The authors in [Huang et al. 2004] studied the structure of proximal cells of points in weakly mixing systems and showed that there is a Xiong chaotic set in those proximal cells. In [Oprocha and Zhang 2013] it was proved that for every closed weakly mixing set A and every $x \in A$, the set $\mathrm{Prox}_2(f)(x) \cap A$ is residual in A. We will show that the same is true if we consider proximal tuples instead of pairs. For a dynamical system (X, f) and a positive integer $n \geq 2$, the *n-th proximal relation* is

$$\mathrm{Prox}_n(f) = \big\{(x_1, \dots, x_n) \in X^n : \liminf_{k\to\infty} \max_{1\leq i<j\leq n} d(f^k(x_i), f^k(x_j)) = 0\big\},$$

and the *n-th proximal cell* of a point $x_0 \in X$ is

$$\mathrm{Prox}_n(f)(x_0) = \{(x_1, \dots, x_{n-1}) \in X^{n-1} : (x_0, x_1, \dots, x_{n-1}) \in \mathrm{Prox}_n(f)\}.$$

Theorem 5.6. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set. Then for every $x_0 \in A$ and $n \geq 2$, the set $\mathrm{Prox}_n(f)(x_0) \cap A^{n-1}$ is residual in A^{n-1}.*

In fact, we prove even more in the following theorem, where $LY_n^\delta(X, f)(x_0)$ is the *n-scrambled cell* of x_0 with modular $\delta > 0$. More precisely, $LY_n^\delta(X, f)(x_0)$ is the collection of points $(x_1, \dots, x_{n-1})$ in X^{n-1} such that $(x_0, x_1, \dots, x_{n-1})$ is proximal and

$$\limsup_{k\to\infty} \min_{0\leq i<j\leq n-1} d(f^k(x_i), f^k(x_j)) \geq \delta.$$

Theorem 5.7. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set. Then, for every $n \geq 2$, there exists a $\delta > 0$ such that, for every $x_0 \in A$, it holds that $LY_n^\delta(X, f)(x_0) \cap A^{n-1}$ is residual in A^{n-1}.*

The following result shows that, when we look only at separation of trajectories of tuples, weak mixing of order 2 is enough to obtain rich structure of such points (see Section 2B for definitions of sensitivity).

Theorem 3.7. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order 2. Then A is a sensitive set in $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$. In particular, the system $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$ is n-sensitive for every $n \geq 2$.*

In the final section, we prove that on topological graphs weakly mixing sets of order 2 can be approximated arbitrarily closely (in the Hausdorff metric) by a weakly mixing set of all orders. This completes our previous research in [Oprocha and Zhang 2011].

2. Preliminaries

In this section, we provide some basic notation, definitions and results which will be used later in this paper. Denote by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$ and $\mathbb{R}$ the set of all positive integers, nonnegative integers, integers and real numbers, respectively. A subset A of $\mathbb{N}$ is an *IP-set* if there exists a sequence $\{p_j\}_{j=1}^{\infty}$ in $\mathbb{N}$ such that $A = FS\bigl(\{p_j\}_{j=1}^{\infty}\bigr)$, where

$$FS\bigl(\{p_j\}_{j=1}^{\infty}\bigr) = \Bigl\{\sum_{j\in\alpha} p_j : \alpha \text{ is a nonempty finite subset of } \mathbb{N}\Bigr\}$$

is the set of *finite sums* of $\{p_j\}_{j=1}^{\infty}$.

Let X be a compact metric space. A subset C of X is a *Cantor set* if it is homeomorphic to the standard Cantor ternary set (equivalently, it is a perfect and totally disconnected compact metric space). We say that a subset K of X is a *Mycielski set* if it can be presented as a countable union of Cantor sets. The next two facts help to deal with residual relations. They are important tools with numerous applications. See [Akin 2004] for a comprehensive treatment of this topic.

Lemma 2.1 (Ulam lemma). *Let X be a perfect compact metric space. If R is a dense G_δ subset of X^n, then there exists a dense G_δ subset K of X such that, for every $x \in K$, the set $R(x) = \{(x_1, \dots, x_{n-1}) \in X^{n-1} : (x, x_1, \dots, x_{n-1}) \in R\}$ is residual in X^{n-1}.*

Theorem 2.2 (Mycielski theorem [1964]). *Let X be a perfect compact metric space. If R is a dense G_δ subset of X^n, then there exists a dense Mycielski subset K of X such that, for any n distinct points $x_1, \dots, x_n \in K$, we have $(x_1, x_2, \dots, x_n) \in R$.*

2A. ***Topological dynamics.*** By a *(topological) dynamical system* we mean a pair (X, f) consisting of a compact metric space (X, d) and a continuous map $f : X \to X$. If X is a singleton, then we say that (X, f) is *trivial*. If $K \subset X$ is a nonempty closed subset satisfying $f(K) \subset K$, then we say that (K, f) is a *subsystem* of (X, f).

Let (X, f) be a dynamical system with $\varnothing \neq A \subset X$ and $x \in X$. The set

$$\mathrm{Orb}(A, f) = \bigcup_{n\in\mathbb{N}_0} f^n(A)$$

is said to be the *(positive) orbit of A under f*. Clearly, $\bigl(\overline{\mathrm{Orb}(A, f)}, f\bigr)$ is a subsystem of (X, f). We will write $\mathrm{Orb}(x, f) = \mathrm{Orb}(\{x\}, f)$ for short.

We say that a point $x \in X$ is a *periodic point of* (X, f) if $f^n(x) = x$ for some $n \in \mathbb{N}$, a *recurrent point of* (X, f) if there exists an increasing sequence $\{k_n\}_{n=1}^{\infty}$ in $\mathbb{N}$ such that $\lim_{n\to\infty} f^{k_n}(x) = x$, and a *transitive point of* (X, f) if $\mathrm{Orb}(x, f)$ is dense in X. Denote by $\mathrm{Per}(X, f)$, $\mathrm{Rec}(X, f)$ and $\mathrm{Tran}(X, f)$ the set of all periodic points, recurrent points and transitive points, respectively, of (X, f). A dynamical system (X, f) is *minimal* if $\mathrm{Tran}(X, f) = X$. A point $x \in X$ is *minimal* if the subsystem $\big(\overline{\mathrm{Orb}(x, f)}, f\big)$ is minimal.

Let (X, f) be a dynamical system and $A \subset X$ with $n \geq 2$. Define the sets $A^n = \{(x_1, x_2, \dots, x_n) : x_1, \dots, x_n \in A\}$ and $\Delta_n(A) = \{(x, x, \dots, x) \in A^n : x \in A\}$. The map $f^{(n)}$ is induced on X^n by the formula

$$f^{(n)}(x_1, x_2, \dots, x_n) = (f(x_1), f(x_2), \dots, f(x_n)).$$

Let (X, f) be a dynamical system with $x \in X$ and $A, B \subset X$. Define the sets $N(x, A) = \{n \in \mathbb{N} : f^n(x) \in A\}$ and $N(A, B) = \{n \in \mathbb{N} : f^n(A) \cap B \neq \varnothing\}$. When we want to emphasize the map f, we instead use $N_f(x, A)$ and $N_f(A, B)$. A dynamical system (X, f) is called *transitive* if, for any nonempty open subsets U and V of X, the set $N(U, V)$ is not empty, *totally transitive* if (X, f^k) is transitive for every $k \in \mathbb{N}$, and *weakly mixing* if $(X^2, f^{(2)})$ is transitive. It is well known that if (X, f) is transitive, then $\mathrm{Tran}(X, f)$ is a dense G_δ subset of X.

2B. *Proximal and scrambled tuples.* We say that an n-tuple $(x_1, \dots, x_n) \in X^n$ (where $n \geq 2$) is *proximal* if

$$\liminf_{k\to\infty} \max_{1\leq i<j\leq n} d(f^k(x_i), f^k(x_j)) = 0.$$

Let $\mathrm{Prox}_n(f)$ denote the collection of all proximal n-tuples in (X, f). It is easy to verify that $\mathrm{Prox}_n(f)$ is a G_δ subset of X^n. For $x \in X$, define the *n-th proximal cell of x* as

$$\mathrm{Prox}_n(f)(x) = \{(x_1, \dots, x_{n-1}) \in X^{n-1} : (x, x_1, \dots, x_{n-1}) \text{ is proximal}\}.$$

An n-tuple $(x_1, \dots, x_n) \in X^n$ (where $n \geq 2$) is called *scrambled (with modular $\delta > 0$)* if it is proximal and

$$\limsup_{k\to\infty} \min_{1\leq i<j\leq n} d(f^k(x_i), f^k(x_j)) \geq \delta.$$

A subset S of X is called *n-scrambled* if any n distinct points in S form a scrambled n-tuple. The system (X, f) is called *Li–Yorke n-chaotic* if there exists an uncountable n-scrambled subset S of X.

Xiong [2005] introduced the concept of n-sensitivity. Specifically, a dynamical system (X, f) is called *n-sensitive*, where $n \geq 2$, if there exists a $\delta > 0$ such that for every nonempty open set $U \subset X$ there are distinct points $x_1, x_2, \dots, x_n \in U$

and some $m \in \mathbb{N}$ with

$$\min_{1 \le i < j \le n} d(f^m(x_i), f^m(x_j)) > \delta.$$

This definition was further generalized in [Ye and Zhang 2008] to sensitive sets. A subset A of X is *sensitive* if for any $n \ge 2$, any n distinct points $x_1, x_2, \dots, x_n$ in A, any neighborhood U_i of x_i, $i = 1, 2, \dots, n$, and any nonempty open set $U \subset X$ there exists a $k \in \mathbb{N}$ and $y_i \in U$ such that $f^k(y_i) \in U_i$ for $i = 1, 2, \dots, n$. It is shown in [Ye and Zhang 2008] that a transitive system is n-sensitive if and only if there exists a sensitive set with cardinality n. Note that 2-scrambled set, Li–Yorke 2-chaos and 2-sensitivity are classical definitions.

2C. *Transitive sets and weakly mixing sets.* Let (X, f) be a dynamical system. A subset A of X is *transitive* in (X, f) if for any open subsets U and V of X intersecting A, the set $N(U \cap A, V)$ is not empty and *totally transitive* if A is transitive in (X, f^k) for every $k \in \mathbb{N}$. Let $n \ge 2$ be an integer. A nontrivial closed subset $A \subset X$ is *weakly mixing of order n* provided that A^n is a transitive set in $(X^n, f^{(n)})$ and *weakly mixing of all orders* or simply *weakly mixing* if A is weakly mixing of order k for all $k = 2, 3, \dots$.

Remark 2.3. In the present paper we require a weakly mixing set (of order n) to be closed and nontrivial which is a little more restrictive than the original definition in [Oprocha and Zhang 2011].

The following result is derived directly from the definition.

Lemma 2.4. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset with $n \ge 2$. Then A is weakly mixing of order n if and only if, for any open subsets $U_1, \dots, U_n$ and $V_1, \dots, V_n$ of X intersecting A,*

$$\bigcap_{i=1}^{n} N(U_i \cap A, V_i) \neq \varnothing.$$

The following lemmas, while simple in proof, are very useful in practice. The proofs can be found in [Oprocha and Zhang 2011; 2014].

Lemma 2.5. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing subset of order* 2. *Then A is perfect.*

Lemma 2.6. *Let (X, f) be a dynamical system and A a closed subset of X. If A is a transitive set, then*:

(2.6.1) $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$ *is a transitive subsystem of* (X, f).

(2.6.2) $A \cap \mathrm{Tran}\big(\overline{\mathrm{Orb}(A, f)}, f\big)$ *is a dense G_δ subset of A.*

2D. ***Symbolic dynamics.*** Let $\mathcal{A}$ be a finite set (an alphabet) endowed with the discrete topology and let $\mathcal{A}^{\mathbb{N}_0}$ denote the Cantor space with respect to the product topology. We write elements of $\mathcal{A}^{\mathbb{N}_0}$ as $x = x_0x_1 \cdots$. The *shift* transformation $\sigma : \mathcal{A}^{\mathbb{N}_0} \to \mathcal{A}^{\mathbb{N}_0}$ is given by $\sigma(x)_i = x_{i+1}$ for $i \in \mathbb{N}_0$. The dynamical system $(\mathcal{A}^{\mathbb{N}_0}, \sigma)$ is called the *full shift over $\mathcal{A}$*.

By a *word* (*over $\mathcal{A}$*), we mean any finite sequence $u = u_0 \cdots u_{n-1}$, $n \geq 1$ where $u_i \in \mathcal{A}$. The length of u is denoted by $|u| = n$ and the set of all words is denoted by $\mathcal{A}^+$. If $x \in \mathcal{A}^{\mathbb{N}_0}$ and $0 \leq i < j$, then by $x_{[i,j]}$ we mean the sequence $x_ix_{i+1} \cdots x_j$. For simplicity, we use the notation $x_{[i,j)} = x_{[i,j-1]}$. If $a_1 \cdots a_m \in \mathcal{A}^+$, then we define the *cylinder set*

$$C[a_1 \cdots a_m] = \{x \in \mathcal{A}^{\mathbb{N}_0} : x_{[0,m)} = a_1 \cdots a_m\}.$$

If X is a subshift, we denote the cylinder set by $C_X[a_1 \cdots a_m] = C[a_1 \cdots a_m] \cap X$.

2E. ***Topological graphs.*** Roughly speaking, a topological graph is a continuum which is the union of a finite number of intervals which can intersect only at endpoints and do not have self-intersections. More formally, a *topological graph* is a compact connected metric space G which is homeomorphic to a polyhedron (a geometric realization) of some finite one-dimensional complex. In particular, we can naturally endow G with the metric d given by the length of the shortest arc joining x, y in G (induced on G from the polyhedron). An arc $I \subset G$ is a *closed interval* if there is a homeomorphism $\varphi : [0, 1] \to I$ such that the set $\varphi((0, 1))$ is open in G.

Let (G, f) be a dynamical system and let $I, J \subset G$ be closed intervals. If there exists a closed interval $K \subset I$ such that $f(K) = J$, then we say that I *f-covers* J and denote this fact by $I \xrightarrow{f} J$. We will need the following standard properties of f-covering (see [Alsedà et al. 2003, p. 590]):

Lemma 2.7. *Let $I, J, K, L \subset G$ be closed intervals and let $f, g : G \to G$ be continuous.*

(2.7.1) *If $I \subset K$, $L \subset J$ and $I \xrightarrow{f} J$, then $K \xrightarrow{f} L$.*

(2.7.2) *If $I \xrightarrow{f} J$ and $J \xrightarrow{g} K$, then $I \xrightarrow{g \circ f} K$.*

(2.7.3) *If $J \subset f(I)$, and $K_1, K_2 \subset J$ are closed intervals such that $K_1 \cap K_2$ is at most one point, then $I \xrightarrow{f} K_1$ or $I \xrightarrow{f} K_2$.*

3. Weakly mixing sets of finite order

In this section we study weakly mixing sets of finite order. It is clear that a dynamical system (X, f) is weakly mixing if and only if, for any four nonempty open subsets U_1, V_1, U_2, V_2 of X, we have $N(U_1, V_1) \cap N(U_2, V_2) \neq \varnothing$. It is shown in [Petersen 1970] that we can reduce four open sets in the characterization of weak mixing to two open sets; that is, a dynamical system (X, f) is weakly mixing if and only if, for any two nonempty open subsets U and V of X, it holds that $N(U, V) \cap N(U, U) \neq \varnothing$ (this was later extended in [Banks 1999] to show that most of the possible conditions of this kind are equivalent to weak mixing). Similar to the above condition, we can simplify the condition in Lemma 2.4 to obtain an alternative definition of weakly mixing set of order n. The advantage is that we have to verify conditions on transfer times for only $n+1$ open sets instead of $2n$ sets in the original definition.

Theorem 3.1. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset with $n \geq 2$. Then A is a weakly mixing set of order n if and only if, for any $n+1$ open subsets $U_1, V_1, V_2, \ldots, V_n$ of X intersecting A,*

$$N(U_1 \cap A, V_1) \cap \bigcap_{i=2}^{n} N(V_i \cap A, V_i) \neq \varnothing.$$

Proof. The necessity follows from Lemma 2.4. Now we prove the sufficiency. Fix any $2n$ open subsets $U_1, V_1, U_2, V_2, \ldots, U_n, V_n$ of X intersecting A. Assume that for some $1 \leq j < n$ we have

$$\bigcap_{i=1}^{j} N(U_i \cap A, V_i) \cap \bigcap_{l=j+1}^{n} N(V_l \cap A, V_l) \neq \varnothing.$$

Then there is a $k > 0$ and open subsets $U_1', \ldots, U_j', V_{j+1}', \ldots V_n'$ of X intersecting A such that $U_i' \subset U_i$ and $f^k(U_i') \subset V_i$ for each $i = 1, \ldots, j$, and $V_l' \subset V_l$ and $f^k(V_l') \subset V_l$ for each $l = j+1, \ldots, n$. By the assumption we can choose

$$m \in N(U_{j+1} \cap A, V_{j+1}') \cap \bigcap_{i=1}^{j} N(U_i' \cap A, U_i') \cap \bigcap_{l=j+2}^{n} N(V_l' \cap A, V_l'),$$

so that

$$m + k \in \bigcap_{i=1}^{j+1} N(U_i \cap A, V_i) \cap \bigcap_{l=j+2}^{n} N(V_l \cap A, V_l).$$

Hence, by induction on j, we eventually obtain that

$$\bigcap_{i=1}^{n} N(U_i \cap A, V_i) \neq \varnothing,$$

which implies that A is weakly mixing of order n. □

Unfortunately, the above technique is not sufficient if we want to directly copy the condition from [Petersen 1970]. This condition simply will not induce even the smallest degree of local weak mixing, as shown by the next example. The technique used here is a modification of Example 6.1 from [Oprocha and Zhang 2011]. Since the construction is somewhat long and complicated, we move it to the Appendix.

Example 3.2. There are a dynamical system (X, f) and a nontrivial closed subset A of X satisfying the following two conditions:

(3.2.1) $N(U \cap A, V) \cap N(U \cap A, U) \neq \varnothing$ and $N(U \cap A, V) \cap N(V \cap A, V) \neq \varnothing$ for any open subsets U, V of X intersecting A.

(3.2.2) A is not weakly mixing of order 2.

It is shown in [Li 2011, Theorem 3.2] that if a dynamical system (X, f) is weakly mixing, then there exists a residual subset K of X such that, for every $x \in K$ and every nonempty open subset U of X, the set $N(x, U)$ contains an IP-set. This theorem was generalized in [Oprocha and Zhang 2013, Theorem 8], which states that if A is a weakly mixing set of order 2 and U is an open set of X intersecting A, then there is an $x \in U \cap A$ such that for every open set V of X intersecting A the set $N(x, V)$ contains an IP-set. The following lemma is inspired by the proof of [Li 2011, Theorem 3.2]. It allows us to extend the above fact from [Oprocha and Zhang 2013] a little further.

Lemma 3.3. *Let (X, f) be a dynamical system with $n \geq 2$. If there are n points $x, y_1, y_2, \ldots, y_{n-1} \in X$ with $x \neq y_1$ such that*

$$(y_1, y_1, y_2, \ldots, y_{n-1}) \in \overline{\mathrm{Orb}((x, y_1, \ldots, y_{n-1}), f^{(n)})}, \tag{1}$$

then, for every choice of open neighborhoods U_i of y_i, $i = 1, 2, \ldots, n-1$, the set

$$N(x, U_1) \cap \bigcap_{i=1}^{n-1} N(y_i, U_i)$$

contains an IP-set.

Proof. For each $i = 1, 2, \ldots, n-1$ fix an open neighborhood U_i of y_i. Since $x \neq y_1$, we may assume that $x \notin U_1$. We are going to construct an IP-set in $N(x, U_1) \cap \bigcap_{i=1}^{n-1} N(y_i, U_i)$. We start our construction by setting $U_i^{(1)} = U_i$ for $i = 1, \ldots, n-1$.

By (1), there exists a $p_1 \in \mathbb{N}$ such that $f^{p_1}(x) \in U_1^{(1)}$ and $f^{p_1}(y_i) \in U_i^{(1)}$ for $i = 1, \ldots, n-1$. Let $U_i^{(2)} = U_i^{(1)} \cap f^{-p_1}(U_i^{(1)})$ for $i = 1, \ldots, n-1$. Clearly, $U_i^{(2)}$ is also an open neighborhood of y_i for $i = 1, \ldots, n-1$. By (1) again there exists a $p_2 > 0$ such that $f^{p_2}(x) \in U_1^{(2)}$ and $f^{p_2}(y_i) \in U_i^{(2)}$ for $i = 1, \ldots, n-1$. Then, for every $m \in FS(\{p_j\}_{j=1}^2)$, we have $f^m(x) \in U_1$ and $f^m(y_i) \in U_i$ for $i = 1, \ldots, n-1$. We continue this construction inductively.

Assume that for some $k \geq 2$ positive integers $p_1, p_2, \ldots, p_k$ have been constructed in such a way that if $m \in FS(\{p_j\}_{j=1}^k)$ then $f^m(x) \in U_1$ and $f^m(y_i) \in U_i$

for $i = 1, \dots, n-1$. For each $i = 1, \dots, n-1$ set

$$U_i^{(k+1)} = U_i \cap \bigcap_{m \in FS\left(\{p_j\}_{j=1}^{k}\right)} f^{-m}(U_i)$$

and observe that each $U_i^{(k+1)}$ is also an open neighborhood of y_i for $i = 1, \dots, n-1$. By (1) there exists a $p_{k+1} > 0$ such that $f^{p_{k+1}}(x) \in U_1^{(k+1)}$ and $f^{p_{k+1}}(y_i) \in U_i^{(k+1)}$ for $i = 1, \dots, n-1$. Then, completing the induction, for every $m \in FS\left(\{p_j\}_{j=1}^{k+1}\right)$, we have $f^m(x) \in U_1$ and $f^m(y_i) \in U_i$ for $i = 1, \dots, n-1$. Thus, we get a sequence $\{p_j\}_{j=1}^{\infty}$ such that $FS\left(\{p_j\}_{j=1}^{\infty}\right) \subset N(x, U_1) \cap \bigcap_{i=1}^{n-1} N(y_i, U_i)$. □

Theorem 3.4. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order n with $n \geq 2$. Then there exists a residual subset K of A such that for any $x \in K$ and any choice of $n-1$ open subsets $U_1, \dots, U_{n-1}$ of X intersecting A there exist points $y_i \in U_i \cap A$, where $i = 1, \dots, n-1$, such that $N(x, U_1) \cap \bigcap_{i=1}^{n-1} N(y_i, U_i)$ contains an IP-set.*

Proof. Since A^n is a transitive set in $(X^n, f^{(n)})$, we have by Lemma 2.6 that the relation $R = A^n \cap \mathrm{Tran}\left(\overline{\mathrm{Orb}(A^n, f^{(n)})}, f^{(n)}\right)$ is a dense G_δ subset of A^n. By the Ulam lemma, there exists a dense G_δ subset K of A such that for every $x \in K$ the section of R at x, that is, the set $R(x) = \{(y_1, \dots, y_{n-1}) \in A^{n-1} : (x, y_1, \dots, y_{n-1}) \in R\}$, is residual in A^{n-1}. It remains to show that K satisfies our requirement.

Fix $x \in K$ and $n-1$ open subsets $U_1, U_2, \dots, U_{n-1}$ of X intersecting A. Since $R(x)$ is residual, we can select points $y_i \in U_i \cap A$, where $i = 1, \dots, n-1$, such that $(x, y_1, \dots, y_{n-1}) \in R$ and $x \neq y_1$ (recall that A is perfect by Lemma 2.5). By the definition of R we obtain

$$(y_1, y_1, y_2, \dots, y_{n-1}) \in \overline{\mathrm{Orb}((x, y_1, \dots, y_{n-1}), f^{(n)})}.$$

Now the result follows by Lemma 3.3. □

Corollary 3.5. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order n with $n \geq 2$. Then, for any n open subsets $U_1, V_1, V_2, \dots, V_{n-1}$ of X intersecting A,*

$$N(U_1 \cap A, V_1) \cap \bigcap_{i=1}^{n-1} N(V_i \cap A, V_i)$$

contains an IP-set.

It is shown in [Oprocha and Zhang 2012, Theorem 6] that a weakly mixing set of order 2 is totally transitive. Now, with the help of Corollary 3.5, we can extend it as follows.

Proposition 3.6. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset.*

(3.6.1) *If A is a weakly mixing set of order* 2, *then A is totally transitive.*

(3.6.2) *If A is a weakly mixing set of order n for f with $n \geq 3$, then, for every $k \in \mathbb{N}$, A is a weakly mixing set of order $n-1$ for f^k.*

Proof. If F contains an IP-set, then $F \cap n\mathbb{N} \neq \varnothing$ for every $n \in \mathbb{N}$. Now the result follows by Theorem 3.1 and Corollary 3.5. □

The above fact motivates us to state the following question for investigation.

Question. Let (X, f) be a dynamical system and $k \in \mathbb{N}$, $n \geq 2$. If a subset $A \subset X$ is weakly mixing of order n for f, is it weakly mixing of order n for f^k?

Theorem 3.7. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order* 2. *Then A is a sensitive set in $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$. In particular, the system $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$ is n-sensitive for every $n \geq 2$.*

Proof. Without loss of generality, we assume that $\overline{\mathrm{Orb}(A, f)} = X$. First note that both A and X must be perfect. Now let $n \geq 2$ and fix n distinct points $x_1, x_2, \ldots, x_n$ in A. Let U be a nonempty open subset of X and U_i an open neighborhood of x_i for $i = 1, 2, \ldots, n$. There is some $k \geq 0$ such that $f^k(A) \cap U \neq \varnothing$ and therefore there is an open subset V of X intersecting A such that $f^k(V) \subset U$. Since A is a weakly mixing set of order 2, there exists an $m_2 \in \mathbb{N}$ such that $U_1 \cap A \cap f^{-m_2}(U_2) \neq \varnothing$ and $V \cap A \cap f^{-m_2}(V) \neq \varnothing$. By induction, there exist $m_3, \ldots, m_n \in \mathbb{N}$ such that

$$U_1 \cap A \cap \bigcap_{i=2}^{n} f^{-m_i}(U_i) \neq \varnothing \quad \text{and} \quad V \cap A \cap \bigcap_{i=2}^{n} f^{-m_i}(V) \neq \varnothing.$$

And so there is a point $y \in A$ such that $\{y, f^{m_2}(y), \ldots, f^{m_n}(y)\} \subset V$. By Lemma 2.6, $\mathrm{Tran}(X, f) \cap A$ is a dense G_δ subset of A, and therefore we can choose x in $\mathrm{Tran}(X, f) \cap U_1 \cap A \cap \bigcap_{i=2}^{n} f^{-m_i}(U_i)$; that is, we choose an $x \in U_1$ such that $f^{m_i}(x) \in U_i$ for $i = 2, \ldots, n$. Since x is a transitive point in (X, f) and the space X is perfect, there exists a $p \in \mathbb{N}_0$ such that $\{f^p(x), f^{p+m_2}(x), \ldots, f^{p+m_n}(x)\} \subset V$ and a $q > p+k$ such that $f^q(x) \in U_1$ and $f^{q+m_i}(x) \in U_i$ for $i = 2, \ldots, n$. Define $r = q - p - k$, $y_1 = f^{p+k}(x)$ and $y_i = f^{p+k+m_i}(x)$ for $i = 2, \ldots, n$. Then $y_i \in U$ and $f^r(y_i) \in U_i$ for $i = 1, 2, \ldots, n$, which implies that A is a sensitive set. Finally, we have by Lemma 2.6 that $\big(\overline{\mathrm{Orb}(A, f)}, f\big)$ is n-sensitive for every $n \geq 2$. □

4. Xiong chaotic set of finite order

In this section, we study Xiong chaotic sets of finite order and their connection to weakly mixing sets of finite order.

Definition 4.1. Let (X, f) be a dynamical system with $n \geq 2$. A subset K of X with at least n points is called a *Xiong chaotic set of order n* if, for any subset E of K

with cardinality n and for any map $g : E \to \overline{K}$, there is an increasing subsequence $\{q_i\}_{i=1}^{\infty}$ in $\mathbb{N}$ such that $\lim_{i\to\infty} f^{q_i}(x) = g(x)$ for every $x \in E$.

The following result is straightforward by the definition.

Proposition 4.2. *If K is a Xiong chaotic set of order n, then there exists a $\delta > 0$ such that, for every n distinct points $x_1, x_2, \ldots, x_n$ in K,*

$$\liminf_{k\to\infty} \max_{1\le i<j\le n} d(f^k(x_i), f^k(x_j)) = 0,$$

$$\limsup_{k\to\infty} \min_{1\le i<j\le n} d(f^k(x_i), f^k(x_j)) > \delta,$$

$$\liminf_{k\to\infty} \max_{1\le i\le n} d(f^k(x_i), x_i) = 0.$$

In particular, K is n-scrambled with modular δ.

Theorem 4.3. *Let (X, f) be a dynamical system and A a perfect subset of X with $n \ge 2$. Then the following conditions are equivalent*:

(4.3.1) *A is a weakly mixing set of order n.*

(4.3.2) *There exists a dense Mycielski subset S of A which is Xiong chaotic of order n.*

(4.3.3) *There exists a dense subset S of A which is Xiong chaotic of order n.*

Proof. (4.3.1) ⇒ (4.3.2) First note that A is perfect. Since $\overline{A^n}$ is a transitive set in $(X^n, f^{(n)})$, by Lemma 2.6 the relation $R = A^n \cap \mathrm{Tran}\big(\overline{\mathrm{Orb}(A^n, f^{(n)})}, f^{(n)}\big)$ is a dense G_δ subset of A^n. By the Mycielski theorem, there exists a dense Mycielski subset S of A such that, for every n distinct points $x_1, x_2, \ldots, x_n \in S$, we have $(x_1, x_2, \ldots, x_n) \in R$. Fix a subset E of S with cardinality n and a map $g : E \to A$. Enumerate E as $\{x_1, x_2, \ldots, x_n\}$ and let $y_i = g(x_i)$ for $i = 1, 2, \ldots, n$. Since $(x_1, x_2, \ldots, x_n)$ is a transitive point in $\big(\overline{\mathrm{Orb}(A^n, f^{(n)})}, f^{(n)}\big)$ and $(y_1, y_2, \ldots, y_n)$ is in A^n, there is an increasing subsequence $\{q_k\}_{k=1}^{\infty}$ in $\mathbb{N}$ such that we have $\lim_{k\to\infty} f^{q_k}(x_i) = g(x_i)$ for $i = 1, 2, \ldots, n$; thus S is a Xiong chaotic set of order n.

(4.3.2) ⇒ (4.3.3) The implication is trivial.

(4.3.3) ⇒ (4.3.1) Fix any open subsets $U_1, V_1, U_2, V_2, \ldots, U_n, V_n$ of X intersecting A. Choose n distinct points $x_i \in U_i \cap S$ and n points $y_i \in V_i \cap A$ for $i = 1, 2, \ldots, n$. Define a map $g : \{x_1, x_2, \ldots, x_n\} \to A$ as $g(x_i) = y_i$ for $i = 1, 2, \ldots, n$. Then there exists a $k \ge 1$ such that $f^k(x_i) \in V_i$ for $i = 1, 2, \ldots, n$. In particular the set $\bigcap_{i=1}^{n} N(U_i \cap A, V_i)$ is not empty, which completes the proof. □

Corollary 4.4. *Let (X, f) be a dynamical system with $n \ge 2$. Then (X, f) has a weakly mixing set of order n if and only if it has an uncountable Xiong chaotic set of order n.*

Proof. The necessity follows by Theorem 4.3, since a Mycielski set is uncountable. Now we prove the sufficiency. Let S be an uncountable Xiong chaotic set of order n. By compactness of X, we can divide the closure $\bar{S}$ of S into $K_1 \cup K_2$, where K_1 is perfect and K_2 is at most countable. It is easy to see that $K_1 \cap S$ is also a Xiong chaotic set which is dense in K_1. By Theorem 4.3, K_1 is a weakly mixing set of order n. □

Remark 4.5. It should be noticed that weakly mixing sets (of finite order n) are perfect. Hence, they are more restrained than Xiong chaotic sets, because any infinite subsets of Xiong chaotic sets (of finite order n) are also Xiong chaotic sets.

Let (X, f) be a dynamical system with $x_0 \in X$, $n \geq 2$ and $\delta > 0$. Define

$$D_n^{\delta}(X, f) = \big\{(x_1, x_2, \dots, x_n) \in X^n : \limsup_{k\to\infty} \min_{1\leq i<j\leq n} d(f^k(x_i), f^k(x_j)) \geq \delta\big\},$$

and

$$D_n^{\delta}(X, f)(x_0) = \{(x_1, \dots, x_{n-1}) \in X^{n-1} : (x_0, x_1, \dots, x_{n-1}) \in D_n^{\delta}(X, f)\}.$$

Proposition 4.6. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order $2(n-1)$ with $n \geq 2$. Then there exists a $\delta > 0$ such that, for every $x_0 \in A$, $D_n^{\delta}(X, f)(x_0) \cap A^{n-1}$ is residual in A^{n-1}.*

Proof. Since A is perfect, we can choose a $\delta > 0$ and $2(n-1)$ distinct points $u_{1,1}, u_{1,2}, \dots, u_{n-1,1}, u_{n-1,2} \in A$ so that $d(u_{i_1,i_2}, u_{j_1,j_2}) > 4\delta$ for $(i_1, i_2) \neq (j_1, j_2)$. Fix $x_0 \in A$. For every $\varepsilon > 0$, set

$$D_\varepsilon = \big\{(x_1, \dots, x_{n-1}) \in X^{n-1} : \min_{0\leq i<j\leq n-1} d(f^k(x_i), f^k(x_j)) > \delta - \varepsilon \text{ for some } k > \tfrac{1}{\varepsilon}\big\}.$$

It is easy to verify that D_ε is an open subset of X^{n-1} and that

$$D_n^{\delta}(X, f)(x_0) = \bigcap_{m=1}^{\infty} D_{\frac{1}{m}}.$$

Therefore it is sufficient to prove that $D_\varepsilon \cap A^{n-1}$ is dense in A^{n-1} for every $\varepsilon > 0$.

Fix $\varepsilon > 0$ and $n-1$ open subsets $U_1, \dots, U_{n-1}$ of X intersecting A. By Theorem 4.3, there is a Xiong chaotic set S of order $2(n-1)$ which is dense in A. Observing that A is perfect, for each $i = 1, \dots, n-1$ choose $y_{i,1}, y_{i,2} \in U_i \cap S$ with $y_{i,1} \neq y_{i,2}$. Define a map $g : \{y_{1,1}, y_{1,2}, \dots, y_{n-1,1}, y_{n-1,2}\} \to A$ as $g(y_{i,j}) = u_{i,j}$ for $i = 1, \dots, n-1$ and $j = 1, 2$. Then there is an increasing subsequence $\{q_l\}_{l=1}^{\infty}$ in $\mathbb{N}$ such that $\lim_{l\to\infty} f^{q_l}(y_{i,j}) = g(y_{i,j}) = u_{i,j}$ for $i = 1, \dots, n-1$ and $j = 1, 2$. Pick $k > 1/\varepsilon$ such that $d(f^k(y_{i,j}), u_{i,j}) < \delta$ for $i = 1, \dots, n-1$ and $j = 1, 2$. There is at most one pair (i_0, j_0) such that $d(f^k(x_0), u_{i_0,j_0}) < 2\delta$. For each $i = 1, \dots, n-1$, if $i \neq i_0$, let $x_i = y_{i,1}$, and if $i = i_0$, let $x_i = y_{i_0,\bar{j_0}}$, where $\bar{j_0} \in \{1, 2\}$ and $\bar{j_0} \neq j_0$.

Then $x_i \in U_i$ for $i = 1, \dots, n-1$ and

$$\min_{0 \le i < j \le n-1} d(f^k(x_i), f^k(x_j)) > \delta,$$

which implies that $D_\varepsilon \cap A^{n-1}$ is dense in A^{n-1}. □

5. Weakly mixing sets

5A. ***Weakly mixing sets.*** By Proposition 3.6, we have the following result.

Lemma 5.1. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset with $n \ge 2$. Then A is weakly mixing for f if and only if it is weakly mixing for f^n.*

A dynamical system (X, f) is called an *F-system* if it is totally transitive and has a dense set of periodic points [Furstenberg 1967]. It is shown in [Furstenberg 1967] that an F-system is disjoint from any minimal system. It is not hard to see that every F-system is weakly mixing (see [Banks 1997, Theorem 1.1]). We say a dynamical system (X, f) has *dense small periodic sets* if for any nonempty open subset U of X there exists a nonempty closed subset K of U and a $k \in \mathbb{N}$ such that $f^k(K) \subset K$. A dynamical system (X, f) is called an *HY-system* if it is totally transitive and has dense small periodic sets. It is shown in [Huang and Ye 2005] that an HY-system is weakly mixing and disjoint from any minimal system.

It is interesting when a totally transitive set or a weakly mixing set of finite order is also a weakly mixing set. Recall that a point $x \in X$ is *distal* provided that if (x, y) is proximal and $y \in \overline{\mathrm{Orb}(x, f)}$ then $x = y$. The following fact is Corollary 11 from [Oprocha and Zhang 2013].

Theorem 5.2. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order* 2. *If the set of all distal points in A is dense in A, then A is weakly mixing of all orders.*

We show that Theorem 5.2 can be generalized in the following way.

Theorem 5.3. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order* 2. *If for every open subset U of X intersecting A there is a dynamical system (Y, g) with a distal point $y \in Y$ and an open neighborhood $V \subset Y$ of y and a point $x \in A \cap U$ such that $N(y, V) \subset N(x, U)$, then A is weakly mixing of all orders.*

Proof. By Theorem 3.1, it is sufficient to show that, for any $n \ge 2$ and any $n+1$ open subsets $U_1, V_1, V_2, \dots, V_n$ of X intersecting A,

$$N(U_1 \cap A, V_1) \cap \bigcap_{i=2}^{n} N(V_i \cap A, V_i) \ne \varnothing.$$

By assumption, for $i = 2, \dots, n$ there are points $x_i \in V_i \cap A$ and distal points y_i (in some dynamical systems) and their open neighborhoods W_i such that we have $N(y_i, W_i) \subset N(x_i, V_i)$ and hence $\bigcap_{i=2}^{n} N(y_i, W_i) \subset \bigcap_{i=2}^{n} N(x_i, V_i)$. But the product of distal points is also distal, thus by [Furstenberg 1981] the following set intersects every IP-set:

$$N((y_2, \dots, y_n), W_2 \times \cdots \times W_n) \subset \bigcap_{i=2}^{n} N(y_i, W_i) \subset \bigcap_{i=2}^{n} N(x_i, V_i).$$

But, by Corollary 3.5, $N(U_1 \cap A, V_1)$ contains an IP-set, which finishes the proof. □

We say that a subset A of X has *dense small periodic sets* if, for any open subset U of X intersecting A, there exists a closed subset K of U intersecting A and a $k \in \mathbb{N}$ such that $f^k(K) \subset K$. Then, observing that weak mixing of order 2 implies total transitivity, we have the following theorem.

Theorem 5.4. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed subset. If A is totally transitive and has dense small periodic sets, then A is weakly mixing.*

Proof. First we show that A is weakly mixing of order 2. Let U_1, V_1, V_2 be open subsets of X intersecting A. Since A has dense small periodic sets, there exists a closed subset K of V_2 intersecting A and a $k \in \mathbb{N}$ such that $f^k(K) \subset K$. Since A is transitive for f^k, there is an $m \in \mathbb{N}$ such that $m \in N_{f^k}(U_1 \cap A, V_1)$. Then $km \in N(U_1 \cap A, V_1) \cap N(V_2 \cap A, V_2)$, which implies that A is weakly mixing of order 2 by Theorem 3.1.

Now we show that A satisfies the requirement of Theorem 5.3. Fix an open subset U of X intersecting A. There exists a closed subset S of U intersecting A and a $k \in \mathbb{N}$ such that $f^k(S) \subset S$. Pick a point $x \in S \cap A$. Then $k\mathbb{N} \subset N(x, U)$. Let $Y = \{0, 1, \dots, k-1\}$ and $g : Y \to Y$, $g(i) = i + 1 \pmod{k}$. Let $y = 0$ and $V = \{0\}$. Then y is a distal point in (Y, g) and $N(y, V) = k\mathbb{N} \subset N(x, U)$. Hence A is weakly mixing of all orders by Theorem 5.3. □

5B. ***Proximal relations.*** It is shown in [Akin and Kolyada 2003] that if (X, f) is weakly mixing, then, for every $x \in X$, the set $\mathrm{Prox}_2(f)(x)$ is residual in X. In [Oprocha and Zhang 2013] it was proved that, for every weakly mixing set A and every $x \in A$, the set $\mathrm{Prox}_2(f)(x) \cap A$ is residual in A. We will show that the same is true if we consider proximal tuples instead of pairs. First, we use a method of construction from [Oprocha and Zhang 2013, Lemma 16] to prove the following result.

Lemma 5.5. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set with $x \in A$. Then, for every $n \geq 2$, any open subsets $U_1, U_2, \dots, U_n$ of X intersecting A and each $\varepsilon > 0$, there are a $y_i \in U_i \cap A$ for $i = 1, 2, \dots, n$ and an $m \in \mathbb{N}$ such that*

(*with* $y_0 = x$)

$$\max_{0 \le i < j \le n} d(f^m(y_i), f^m(y_j)) \le \varepsilon.$$

Proof. Fix $n \ge 2$, open subsets $U_1, U_2, \ldots, U_n$ of X intersecting A and $\varepsilon > 0$. Let $\{V_1, \ldots, V_k\}$ be a cover of X consisting of open sets with diameters less than $\varepsilon/2$.

There exists an $s_1 \in \{1, 2, \ldots, k\}$ with $x \in V_{s_1}$. By weak mixing of A, there exist an $m_1 > 0$ and an open set $U_i^{(1,1)} \subset U_i$ intersecting A such that $f^{m_1}(U_i^{(1,1)}) \subset V_{s_1}$ for $i = 1, 2, \ldots, n$.

For some $q \ge 1$, construct open sets $U_i^{(q,1)}, U_i^{(q,2)}, \ldots, U_i^{(q,q)} \subset U_i$ intersecting A for $i = 1, 2, \ldots, n$, pairwise distinct integers $s_1, s_2, \ldots, s_q \subset \{1, 2, \ldots, k\}$, and an integer $m_q > 0$ such that

$$f^{m_q}(U_i^{(q,r)}) \subset V_{s_r} \quad \text{for } r = 1, 2, \ldots, q,\ i = 1, 2, \ldots, n.$$

If $f^{m_q}(x) \notin \bigcup_{r=1}^{q} V_{s_r}$, then we can choose $s_{q+1} \in \{1, 2, \ldots, k\} \setminus \{s_1, s_2, \ldots, s_q\}$ and an open set $U^{(q,q+1)}$ containing x (and intersecting A) so that $f^{m_q}(U^{(q,q+1)}) \subset V_{s_{q+1}}$. By weak mixing of A, there exist open sets $U_i^{(q+1,1)}, U_i^{(q+1,2)}, \ldots, U_i^{(q+1,q+1)} \subset U_i$ intersecting A and a $p > 0$ such that, for $r = 1, 2, \ldots, q+1$ and $i = 1, 2, \ldots, n$, we have $f^p(U_i^{(q+1,r)}) \subset U_i^{(q,r)}$, where $U_i^{(q,q+1)} = U^{(q,q+1)}$ for $i = 1, 2, \ldots, n$. Now if we set $m_{q+1} = m_q + p$, then for $r = 1, 2, \ldots, q+1$ and $i = 1, 2, \ldots, n$ we have

$$f^{m_{q+1}}(U_i^{(q+1,r)}) = f^{m_q}(f^p(U_i^{(q+1,r)})) \subset f^{m_q}(U_i^{(q,r)}) \subset V_{s_r}.$$

Obviously, since $q \le k$, we cannot extend the sequence $s_1, s_2, \ldots, s_q$ any further by the above procedure. Hence, we have that $f^{m_q}(x) \in \bigcup_{r=1}^{q} V_{s_r}$, and in particular $f^{m_q}(x) \in V_{s_\ell}$ for some $\ell \in \{s_1, \ldots, s_q\}$. But then by the construction we have $f^{m_q}(U_i^{(q,\ell)}) \subset V_{s_\ell}$ for $i = 1, 2, \ldots, n$, and so if we fix any $y_i \in U_i^{(q,\ell)} \cap A \subset U_i \cap A$ then $f^{m_q}(y_i) \in V_{s_\ell}$, finishing the proof. □

Theorem 5.6. *Let* (X, f) *be a dynamical system and* $A \subset X$ *a weakly mixing set. Then, for every* $x_0 \in A$ *and* $n \ge 2$, *the set* $\mathrm{Prox}_n(f)(x_0) \cap A^{n-1}$ *is residual in* A^{n-1}.

Proof. Fix any $x_0 \in A$ and $n \ge 2$. For every $\varepsilon > 0$, set

$$P_\varepsilon = \Bigl\{(x_1, \ldots, x_{n-1}) \in X^{n-1} : \max_{0 \le i < j \le n-1} d(f^k(x_i), f^k(x_j)) < \varepsilon \text{ for some } k \ge 0\Bigr\}.$$

It is easy to verify that P_ε is an open subset of X^{n-1}. By Lemma 5.5, $P_\varepsilon \cap A^{n-1}$ is dense in A^{n-1}. This, by the fact that

$$\mathrm{Prox}_n(f)(x_0) = \bigcap_{m=1}^{\infty} P_{\frac{1}{m}},$$

proves that $\mathrm{Prox}_n(f)(x_0) \cap A^{n-1}$ is residual in A^{n-1}. □

Let (X, f) be a dynamical system with $x_0 \in X$, $n \geq 2$ and $\delta > 0$. Define

$$LY_n^\delta(X, f) = \{(x_1, x_2, \dots, x_n) \in X^n : (x_1, \dots, x_n) \text{ is } n\text{-scrambled with modular } \delta\},$$

and

$$LY_n^\delta(X, f)(x_0) = \{(x_1, \dots, x_{n-1}) \in X^{n-1} : (x_0, x_1, \dots, x_{n-1}) \in LY_n^\delta(X, f)\}.$$

The following fact is a direct corollary of Proposition 4.6 and Theorem 5.6.

Theorem 5.7. *Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set. Then, for every $n \geq 2$, there exists a $\delta > 0$ such that, for every $x_0 \in A$, it holds that $LY_n^\delta(X, f)(x_0) \cap A^{n-1}$ is residual in A^{n-1}.*

5C. *Local independent sets.* Let (X, f) be a dynamical system. Following [Kerr and Li 2007], for a tuple $\boldsymbol{A} = (A_1, A_2, \dots, A_k)$ of subsets of X, we say that a nonempty subset $F \subset \mathbb{N}_0$ is an *independence set* for $\boldsymbol{A}$ if, for any nonempty finite subset $J \subset F$, we have

$$\bigcap_{j \in J} f^{-j}(A_{s(j)}) \neq \varnothing$$

for any $s \in \{1, \dots, k\}^J$. We shall denote the collection of all independence sets for $\boldsymbol{A}$ by $\operatorname{Ind}(A_1, A_2, \dots, A_k)$ or $\operatorname{Ind} \boldsymbol{A}$. According to the best knowledge of the authors, the above notion of independence sets was first presented in [Huang and Ye 2006] under the name *interpolating set* (see also [Glasner and Weiss 1995]) and in [Huang 2006] when defining *strong scrambled pairs*. Later, the authors of [Huang et al. 2012] systematically studied independence sets in topological and measurable dynamics. In particular, they proved the following result (see [Huang et al. 2012, Theorem 5.1]).

Theorem 5.8. *For a dynamical system (X, f), the following conditions are equivalent*:

(5.8.1) *(X, f) is weakly mixing.*

(5.8.2) *For any two nonempty open subsets U_1, U_2 of X,* $\operatorname{Ind}(U_1, U_2)$ *contains an infinite set.*

(5.8.3) *For any $n \in \mathbb{N}$ and any nonempty open subsets $U_1, U_2, \dots, U_n$ of X, there is an IP-set in* $\operatorname{Ind}(U_1, U_2, \dots, U_n)$.

In the spirit of [Huang et al. 2012] we introduce a local definition of independence sets as follows.

Definition 5.9. Let (X, f) be a dynamical system with $\varnothing \neq A \subset X$, and let $U_1, U_2, \dots, U_n$ be open subsets of X intersecting A. We say that a nonempty subset $F \subset \mathbb{N}_0$ is an independence set for $(U_1, U_2, \dots, U_n)$ with respect to A if, for every nonempty finite subset $J \subset F$ and $s \in \{1, 2, \dots, n\}^J$,

$$\bigcap_{j \in J} f^{-j}(U_{s(j)})$$

is a nonempty open subset of X intersecting A.

Now we can employ this definition to state a theorem analogous to Theorem 5.8.

Theorem 5.10. *Let (X, f) be a dynamical system and $A \subset X$ a nontrivial closed set. Then the following conditions are equivalent*:

(5.10.1) *A is a weakly mixing set.*

(5.10.2) *For every $n \geq 2$ and any open subsets $U_1, U_2, \dots, U_n$ of X intersecting A, there exists a $t \in \mathbb{N}$ such that $\{0, t\}$ is an independence set for $(U_1, U_2, \dots, U_n)$ with respect to A.*

(5.10.3) *For every $n \geq 2$ and any open subsets $U_1, U_2, \dots, U_n$ of X intersecting A, there exists a sequence $\{t_j\}_{j=1}^{\infty}$ in $\mathbb{N}$ such that $\{0\} \cup FS\big(\{t_j\}_{j=1}^{\infty}\big)$ is an independence set for $(U_1, U_2, \dots, U_n)$ with respect to A.*

Proof. (5.10.2) $\Rightarrow$ (5.10.1) Fix $n \geq 2$ and fix open subsets $U_1, U_2, \dots, U_n$, $V_1, V_2, \dots, V_n$ of X intersecting A. By assumption there exists a $t \geq 1$ such that $\{0, t\}$ is an independence set for $(U_1, U_2, \dots, U_n, V_1, V_2, \dots, V_n)$ with respect to A. For $i = 1, 2, \dots, n$, we have that $U_i \cap f^{-t}(V_i)$ is a nonempty open subset of X intersecting A. Therefore, $t \in \bigcap_{i=1}^{n} N(U_i \cap A, V_i)$, which implies that A is weakly mixing of order n.

(5.10.1) $\Rightarrow$ (5.10.3) Let $U_1, U_2, \dots, U_n$ be open subsets of X intersecting A. First, there exists a $t_1 \in \mathbb{N}$ such that

$$t_1 \in \bigcap_{i_1, i_2 \in \{1,2,\dots,n\}} N(U_{i_1} \cap A, U_{i_2}).$$

That is, for every $i_1, i_2 \in \{1, 2, \dots, n\}$, we have that $U_{i_1} \cap f^{-t_1}(U_{i_2})$ is a nonempty open set intersecting A. Therefore, there exists a $t_2 \in \mathbb{N}$ such that

$$t_2 \in \bigcap_{i_1, i_2, i_3, i_4 \in \{1,2,\dots,n\}} N(U_{i_1} \cap f^{-t_1}(U_{i_2}) \cap A, U_{i_3} \cap f^{-t_1}(U_{i_4})).$$

That is, for every $i_1, i_2, i_3, i_4 \in \{1, 2, \dots, n\}$,

$$U_{i_1} \cap f^{-t_1}(U_{i_2}) \cap f^{-t_2} U_{i_3} \cap f^{-(t_1+t_2)}(U_{i_4})$$

is a nonempty open set of X intersecting A. Then $\{0, t_1, t_2, t_1+t_2\}$ is an independent set of $(U_1, U_2, \dots, U_n)$ with respect to A and the result follows by induction.

(5.10.3) $\Rightarrow$ (5.10.2) The implication is trivial. □

Remark 5.11. Let $\boldsymbol{A} = (A_1, \dots, A_k)$ be a tuple of subsets of X. If F is an independence set for $\boldsymbol{A}$, then for every $m \in \mathbb{N}_0$ the subset $F - m$ defined by $\{n - m : n \geq m \text{ and } n \in F\}$ is also an independence set for $\boldsymbol{A}$. So we may also assume that an independence set of $\boldsymbol{A}$ contains 0. But in Theorem 5.10 we cannot replace $\{0\} \cup FS\big(\{t_j\}_{j=1}^{\infty}\big)$ by $FS\big(\{t_j\}_{j=1}^{\infty}\big)$, as shown by the following example.

Example 5.12. Consider $X = \Sigma_3^+ = \{0, 1, 2\}^{\mathbb{N}_0}$ and define

$$A = \{1, 2\}^{\mathbb{N}_0} \cup C_X[00].$$

For any open subsets $U_1, \dots, U_n$ of X intersecting A, we can easily define words $w^{(1)}, \dots, w^{(n)}$ of the same length, $M \geq 2$, (with symbols in the alphabet $\{0, 1, 2\}$) such that $C[w^{(i)}] \subset A$ for all $i = 1, \dots, n$. This implies that the set $J = \{kM : k \in \mathbb{N}\}$ is an independence set for $(U_1, U_2, \dots, U_n)$ with respect to A. But A is not weakly mixing of order 2, because, for example, $N(A \cap C_X[1], C_X[0]) = \varnothing$.

Question. Let (X, f) be a dynamical system and $A \subset X$ a weakly mixing set of order 2. Is it true that for every two open subsets U_1, U_2 of X intersecting A, there exists a $t \in \mathbb{N}$ such that $\{0, t\}$ is an independence set for (U_1, U_2) with respect to A?

6. Topological graphs

For any integer $n \geq 2$, it is known that a weakly mixing set of order n does not have to be weakly mixing of order $n+1$; even worse, it may happen that there is no weakly mixing set of order $n+1$ in a system with weakly mixing sets of order n [Oprocha and Zhang 2014]. Note that the examples in [Oprocha and Zhang 2014] are subshifts, and for every dynamical system on the unit interval $([0, 1], f)$ with positive topological entropy there is an $m > 0$ and a closed set Λ invariant for f^m such that (Λ, f^m) is conjugated with the full shift on two symbols. In particular, in every interval map with weakly mixing sets we can find sets which are weakly mixing of order n but not $n+1$. However, in [Oprocha and Zhang 2011] the authors proved that on the unit interval every weakly mixing set of order 2 is arbitrarily close (in the Hausdorff metric) to a weakly mixing set of all orders. So even if these sets are not the same, they are arbitrarily close to each other. Theorem 6.1 completes our research on weakly mixing sets in dimension one, showing that the above fact also holds for all topological graphs.

Theorem 6.1. *Let (G, f) be a dynamical system acting on the topological graph G and let $A \subset G$ be a weakly mixing subset of order 2. Then for every $\varepsilon > 0$ there is a weakly mixing subset $D \subset G$ such that $\mathcal{H}_d(A, D) \leq \varepsilon$, where $\mathcal{H}_d(A, D)$ denotes the Hausdorff distance between A and D.*

Proof. Let $\varepsilon > 0$. Pick nonempty open subsets $U_1, \dots, U_s$ of G with diameters at most ε such that $A \subset \bigcup_{i=1}^{s} U_i$ and $A \cap U_i \neq \varnothing$ for $i = 1, \dots, s$. By Lemma 2.5 the set A is perfect; therefore, for every $i = 1, \dots, s$ it is possible to select an open set $V_i \subset U_i$ contained in the interior of an edge of G such that $V_i \cap A \neq \varnothing$ and $V_1, \dots, V_s$ are pairwise disjoint.

Claim. *For every* $i = 1, \dots, s$ *there is an interval* $I_i \subset V_i$, *its disjoint closed subintervals* K_{2i}, K_{2i+1} *and an integer* $n_i > 0$ *such that*

(6.1.1) K_{2i}, K_{2i+1} *form a strong* 2*-horseshoe for* f^{n_i}; *that is,* $K_p \overset{f^{n_i}}{\Longrightarrow} K_q$ *for all* $p, q \in \{2i, 2i+1\}$, *and*

(6.1.2) *both sets* $\operatorname{int}(K_{2i})$, $\operatorname{int}(K_{2i+1})$ *as well as every connected component of the set* $I_i \setminus (K_{2i} \cup K_{2i+1})$ *intersect* A.

Proof of Claim. Let I_i be any closed interval contained in V_i such that $\operatorname{int}(I_i) \cap A \neq \varnothing$. Let us identify I_i with $[0, 1]$. Observe that A is a weakly mixing set of order 2, and so it contains no isolated points. Thus there are points $0 = a_0 < a_1 < \dots < a_6 < a_7 = 1$ in I_i such that $(a_j, a_{j+1}) \cap A \neq \varnothing$ for all $j = 0, \dots, 6$. Define $I_{i,j} = [a_{2j}, a_{2j+1}]$ for $j = 0, 1, 2, 3$. Since A is weakly mixing of order 2, there are $k > 0$, $r > 0$ such that

$$f^r(I_{i,1}) \cap (a_1, a_2) \neq \varnothing, \quad f^k(I_{i,2}) \cap (a_1, a_2) \neq \varnothing,$$
$$f^r(I_{i,1}) \cap (a_5, a_6) \neq \varnothing, \quad f^k(I_{i,2}) \cap (a_5, a_6) \neq \varnothing.$$

If $I_{i,1} \overset{f^r}{\Longrightarrow} I_{i,1}$, $I_{i,1} \overset{f^r}{\Longrightarrow} I_{i,2}$ and $I_{i,2} \overset{f^k}{\Longrightarrow} I_{i,1}$, $I_{i,2} \overset{f^k}{\Longrightarrow} I_{i,2}$ then by Lemma 2.7 the intervals $I_{i,1}$ and $I_{i,2}$ form a 2-horseshoe for f^{k+r}.

Otherwise there are $p \in \{1, 2\}$ and $j \in \{k, r\}$ such that $I_{i,p} \overset{f^j}{\Longrightarrow} I_{i,0}$ and $I_{i,p} \overset{f^j}{\Longrightarrow} I_{i,3}$. Next, if we consider $I_{i,0}$, then there is an $l > 0$ such that $I_{i,0} \overset{f^l}{\Longrightarrow} I_{i,0}$ and $I_{i,0} \overset{f^l}{\Longrightarrow} I_{i,3}$, or we have the second possibility that $I_{i,0} \overset{f^l}{\Longrightarrow} I_{i,1}$ and $I_{i,0} \overset{f^l}{\Longrightarrow} I_{i,2}$ which implies that $I_{i,0} \overset{f^{l+j}}{\Longrightarrow} I_{i,0}$ and $I_{i,0} \overset{f^{l+j}}{\Longrightarrow} I_{i,3}$ again by applying Lemma 2.7. We can repeat the same arguments for $I_{i,3}$, and with the help of Lemma 2.7 finally obtain that $I_{i,0}$, $I_{i,3}$ form a 2-horseshoe for some iterate of f. □

Now for each $i = 1, \dots, s$ let sets K_{2i}, K_{2i+1} be provided for V_i by the claim, and let J_i be the connected component of $I_i \setminus (K_{2i} \cup K_{2i+1})$ such that K_{2i} and K_{2i+1} are contained in different connected components of $I_i \setminus J_i$. We prove by induction that for every $m = 1, \dots, s$ intervals $K_2, \dots, K_{2m+1}$ form a horseshoe for some iterate f^h, $h \in \mathbb{N}$; that is, $K_p \overset{f^h}{\Longrightarrow} K_q$ for all $p, q \in \{2, \dots, 2m+1\}$.

By the construction we have proved the above statement for $m = 1$, so we may assume that it holds for some $1 \leq m < s$. Therefore, there is a $t_1 > 0$ such that $K_p \overset{f^{t_1}}{\Longrightarrow} K_q$ for all $p, q \in \{2, \dots, 2m+1\}$ and a $t_2 > 0$ such that K_{2m+2}, K_{2m+3} form a 2-horseshoe for f^{t_2}. If we set $t = t_1 t_2$, then we have that $K_2, \dots, K_{2m+1}$ form a horseshoe for f^t, and that K_{2m+2}, K_{2m+3} form a 2-horseshoe for f^t. Since

A is weakly mixing of order 2, there are $k, r > 0$ such that

$$f^r(K_2) \cap J_1 \neq \varnothing, \quad f^r(K_2) \cap J_{m+1} \neq \varnothing,$$
$$f^k(K_{2m+3}) \cap J_1 \neq \varnothing, \quad f^k(K_{2m+3}) \cap J_{m+1} \neq \varnothing.$$

From the construction and the first two conditions we see that $K_2 \overset{f^r}{\Longrightarrow} K_2$ or $K_2 \overset{f^r}{\Longrightarrow} K_3$ and at the same time $K_2 \overset{f^r}{\Longrightarrow} K_{2m+2}$ or $K_2 \overset{f^r}{\Longrightarrow} K_{2m+3}$, which implies that $K_2 \overset{f^{r+t}}{\Longrightarrow} K_q$ for every $q = 2, \dots, 2m+3$ by Lemma 2.7. By a symmetric argument, we see that $K_{2m+3} \overset{f^{k+t}}{\Longrightarrow} K_q$ for every $q = 2, \dots, 2m+3$. Now applying Lemma 2.7 it is easy to verify that $K_p \overset{f^{r+k+3t}}{\Longrightarrow} K_q$ for every $p, q \in \{2, \dots, 2m+3\}$. This completes the induction.

Since K_i, $i = 2, \dots, 2s+1$, form a horseshoe, rewriting arguments in the proof of [Moothathu 2011, Theorem 9] (stated there for horseshoes in interval maps) we obtain that there is an $n > 0$, an f^n-invariant closed subset $\Gamma \subset \bigcup_{i=1}^{s} V_i$ and a topological conjugacy $\pi : (\Gamma, f^n) \to (\Sigma_s, \sigma)$ between dynamical systems such that $\Gamma \cap V_i \neq \varnothing$ for each $i = 1, \dots, s$ (and hence $\mathcal{H}_d(A, \Gamma) \leq \varepsilon$), where (Σ_s, σ) is the full shift over the alphabet $\{1, \dots, s\}$. In particular f^n is mixing on Γ, so indeed Γ is a weakly mixing subset, which completes the proof. □

Acknowledgements

The authors would like to thank the referee for useful comments that resulted in substantial improvements to this paper.

The first author was supported in part by National Natural Science Foundation of China, grant numbers 11401362 and 11471125. The second author was supported by the Polish Ministry of Science and Higher Education from sources for science in the years 2013–2014, grant number IP2012 004272. The third author was supported by Foundation for the Author of National Excellent Doctoral Dissertation of China, grant number 201018 and National Natural Science Foundation of China, grant number 11271078.

Appendix

Proof of Example 3.2. Let us endow [0, 1] with the Euclidean metric. Take any increasing sequences $\{a_i\}_{i\in\mathbb{Z}}, \{b_i\}_{i\in\mathbb{Z}} \subset \mathbb{R}\setminus\mathbb{Q}$ such that $1/2 < a_{-1} < a_0 < b_0 < b_1 < a_1$ and $\lim_{i\to\infty} a_{-i} = \lim_{i\to\infty} b_{-i} = 0$ and $\lim_{i\to\infty} a_i = \lim_{i\to\infty} b_i = 1$. Furthermore, we assume that every interval (a_i, a_{i+1}) contains at most one element of the set $\{2^{-k} : k \in \mathbb{N}\}$.

We will define homeomorphisms $F_i : [0, 1] \to [0, 1]$ for $i = 0, \dots, 8$. Define $F_0 = \mathrm{id}$ and set $F_i(0) = 0$ and $F_i(1) = 1$ for every i. For each $i \in \mathbb{Z}$ we set $F_1(a_i) = a_{i+1}$ and $F_3(b_i) = b_{i+1}$. On each interval $[b_i, b_{i+1}]$ we define F_3 as a linear map, which completes the definition of F_3, since values of F_3 at endpoints of every such interval have already been set. For $i \geq 0$ we define F_1 on $[a_i, a_{i+1}]$ as a linear map. Now, fix any sequence of distinct points $\{c_k\}_{k=1}^{\infty} \subset (a_0, a_1)$ in such a way that $\overline{\{c_k : k \in \mathbb{N}\}} = [a_0, a_1]$. We are ready to define F_1 on the intervals $[a_i, a_{i+1}]$ for $i < 0$. Suppose that $F_1|_{[a_{-n},1]}$ is already defined for some $n \geq 0$. If $\{2^{-k} : k \in \mathbb{N}\} \cap [a_{-n-1}, a_{-n}] = \varnothing$ then we define F_1 as a linear map on $[a_{-n-1}, a_{-n}]$, and as a result F_1 is well-defined on the interval $[a_{-n-1}, 1]$. Otherwise there is a $k > 0$ such that $2^{-k} \in (a_{-n-1}, a_{-n})$ (hence $n \geq 1$). Define

$$G : [a_{-n}, a_{-n+1}] \ni x \mapsto F_1|_{[a_{-1},a_0]} \circ \cdots \circ F_1|_{[a_{-n},a_{-n+1}]}(x) \in [a_0, a_1]$$

and observe that there is a $q \in [a_{-n}, a_{-n+1}]$ such that $G(q) = c_k$. Now, we set $F_1(2^{-k}) = q$ and define F_1 to be linear on each of the intervals $[a_{-n-1}, 2^{-k}]$ and $[2^{-k}, a_{-n}]$. Then in this case F_1 is also well-defined on $[a_{-n-1}, 1]$. Induction completes the construction. Define inverses $F_2 = F_1^{-1}$ and $F_4 = F_3^{-1}$. Then for every k there is an $n > 0$ such that $F_2^n(c_k) = F_1^{-n}(c_k) = 2^{-k}$.

We define $F_5(x) = 1/2 + 1/2(2x - 1)^3$ and $F_6 = F_5^{-1}$. Finally $F_7(2^{-k-1}) = 2^{-k}$ for $k \in \mathbb{N}$, $F_7(1/2) = a_1$ and $F_7(a_i) = a_{i+1}$ for $i \in \mathbb{N}$. Between any two consecutive points in the set $\bigcup_{k\in\mathbb{N}}\{2^{-k}, a_k\}$ the map F_7 is linear, which again gives a well-defined homeomorphism. As the last map we take $F_8 = F_7^{-1}$. Observe that for any $\delta, \varepsilon \in (0, 1/2)$ there is an $n > 0$ such that $F_6^n([1/2 - \delta, 1/2 + \delta]) \supset (\varepsilon, 1 - \varepsilon)$.

Let $X = \Sigma_9^+ \times [0, 1]$ (endowed with the product metric given by the maximum distance on each coordinate) where $\Sigma_9^+ = \{0, 1, \ldots, 8\}^{\mathbb{N}_0}$ and let $T : X \to X$ be defined by

$$T(\omega, x) = (\sigma(\omega), F_{\omega_0}(x))$$

with σ the standard shift transformation on Σ_9^+. Thus X is compact and T is continuous.

For any symbol $a \in \{0, 1, \ldots, 8\}$ and pairs $(0, 0)$, $(1, 2)$, $(3, 4)$, $(5, 6)$, $(7, 8)$, let $\bar{a}$ be the replacement of a by the second element of the respective pair. For example, $\bar{8} = 7$. We extend this definition to words, putting $\overline{w_0 \cdots w_n} = \overline{w_n} \cdots \overline{w_0}$. If, for a finite sequence w of symbols in $\{0, 1, \ldots, 8\}$, we denote by F_w the composition $F_w = F_{w_{|w|-1}} \circ \cdots \circ F_{w_1} \circ F_{w_0}$, then $F_{w\bar{w}} = F_{\bar{w}} \circ F_w = \mathrm{id}$.

Before proceeding with the construction of a set A, let us make a few observations on these maps F_i. Fix any nonempty open sets U, V with $\bar{U}, \bar{V} \subset (0, 1)$. First of all $U \cap (b_i, b_{i+1}) \neq \varnothing$ for some $i \in \mathbb{Z}$, and hence there is a word u consisting only of symbols 0, 3 or 4 and such that

$$F_u(U) \cap (a_0, a_1) \supset F_u(U) \cap (b_0, b_1) \neq \varnothing.$$

In particular, there is a $k \in \mathbb{N}$ such that $c_k \in F_u(U)$. But then there is also an $s > 0$ such that if we set $v = 2^s 7^{k-1}$ (i.e., v is a concatenation of s repetitions of symbol 2 and $k-1$ repetitions of symbol 7) then $F_v(c_k)=F_7^{k-1}(F_2^s(c_k))=F_7^{k-1}(2^{-k})=1/2$. In particular, there is a $\delta>0$ such that $[1/2-\delta,\,1/2+\delta]\subset F_{uv}(U)$, where as usual $uv=u_0u_1\cdots u_{|u|-1}v_0v_1\cdots v_{|v|-1}$. But then there are an $\varepsilon>0$ and $m>0$ such that

$$F_w([1/2-\delta,\,1/2+\delta])\supset(\varepsilon,\,1-\varepsilon)\supset U\cup V$$

if we set $w=6^m$. We have just shown that for any nonempty open sets U, V with $\bar{U}$, $\bar{V}\subset(0,\,1)$ there are words u, v, w such that

$$F_{uvw}(U)\supset U\cup V. \tag{2}$$

Now we are ready to construct set A. Let $\{\omega_i\}_{i=1}^{\infty}$ be any sequence containing all possible words (finite sequences) over the symbols $0, 1, \ldots, 8$. Let

$$\xi=\omega_1\bar{\omega}_1\omega_2\bar{\omega}_2\cdots\omega_n\bar{\omega}_n\cdots\in\Sigma_9^+\quad\text{and}\quad A=\{\xi\}\times[0,\,1].$$

Take any nonempty open sets $\tilde{U}$, $\tilde{V}$ intersecting A. Then there are an $i>0$ and open intervals U, V such that $\bar{U}$, $\bar{V}\subset(0,\,1)$ and

$$C[\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i]\times U\subset\tilde{U}\quad\text{and}\quad C[\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i]\times V\subset\tilde{V}.$$

Let words u, v, w be provided for U and V by (2). By the definition, there is a $j>1$ such that $\omega_j=uvw\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i$. Define $t=\sum_{r=1}^{j-1}2|\omega_r|$ and $p=t+|uvw|$. Note that

$$\begin{aligned}T^t(\{\xi\}\times U)&=\{\sigma^t(\xi)\}\times F_{\omega_1\bar{\omega}_1\cdots\omega_{j-1}\bar{\omega}_{j-1}}(U)=\{\sigma^t(\xi)\}\times U\\&=\{\omega_j\bar{\omega}_j\cdots\}\times U=\{uvw\omega_1\bar{\omega}_1\omega_2\bar{\omega}_2\cdots\omega_i\bar{\omega}_i\cdots\}\times U\end{aligned}$$

and therefore

$$\begin{aligned}T^p(\tilde{U}\cap A)\supset T^p(\{\xi\}\times U)&\supset\{\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i\cdots\}\times F_{uvw}(U)\\&\supset\{\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i\cdots\}\times(U\cup V).\end{aligned}$$

We have just shown that $p\in N(\tilde{U}\cap A,\,\tilde{U})\cap N(\tilde{U}\cap A,\,\tilde{V})$.

Similarly, if in the above calculations j was such that $\omega_j=\overline{uvw}\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i$ then, since we have $F_{\overline{uvw}}(U\cup V)=F_{uvw}^{-1}(U\cap V)\subset U$ by (2), we obtain that $p\in N(\tilde{U}\cap A,\,\tilde{U})\cap N(\tilde{V}\cap A,\,\tilde{U})$, as

$$\begin{aligned}T^p(\tilde{U}\cap A)&\supset\{\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i\cdots\}\times F_{\overline{uvw}}(U)\subset\tilde{U},\\T^p(\tilde{V}\cap A)&\supset\{\omega_1\bar{\omega}_1\cdots\omega_i\bar{\omega}_i\cdots\}\times F_{\overline{uvw}}(V)\subset\tilde{U}.\end{aligned}$$

Finally, observe that each map F_i preserves the ordering of $[0,\,1]$. Set $U=(a,\,b)$ and $V=(c,\,d)$ where $b<c$. If for some word w we have $F_w(V)\cap U\neq\varnothing$, then $F_w(U)\subset[0,\,b)$, and in particular $F_w(U)\cap V=\varnothing$. Therefore, if we set $\tilde{U}=\Sigma_9^+\times U$ and $\tilde{V}=\Sigma_9^+\times V$, both intersecting A, then $N(\tilde{U}\cap A,\,\tilde{V})\cap N(\tilde{V}\cap A,\,\tilde{U})=\varnothing$. This shows that A is not a weakly mixing set of order 2, completing the proof. $\square$

References

[Akin 2004] E. Akin, "Lectures on Cantor and Mycielski sets for dynamical systems", pp. 21–79 in *Chapel Hill Ergodic Theory Workshops*, edited by I. Assani, Contemp. Math. **356**, Amer. Math. Soc., Providence, RI, 2004. MR 2005e:37018 Zbl 1064.37015

[Akin and Kolyada 2003] E. Akin and S. Kolyada, "Li–Yorke sensitivity", *Nonlinearity* **16**:4 (2003), 1421–1433. MR 2004c:37016 Zbl 1045.37004

[Alsedà et al. 2003] L. Alsedà, M. A. del Río, and J. A. Rodríguez, "Transitivity and dense periodicity for graph maps", *J. Difference Equ. Appl.* **9**:6 (2003), 577–598. MR 2004d:37061 Zbl 1032.37025

[Banks 1997] J. Banks, "Regular periodic decompositions for topologically transitive maps", *Ergodic Theory Dynam. Systems* **17**:3 (1997), 505–529. MR 98d:54074 Zbl 0921.54029

[Banks 1999] J. Banks, "Topological mapping properties defined by digraphs", *Discrete Contin. Dynam. Systems* **5**:1 (1999), 83–92. MR 99j:54038 Zbl 0957.54020

[Blanchard and Huang 2008] F. Blanchard and W. Huang, "Entropy sets, weakly mixing sets and entropy capacity", *Discrete Contin. Dyn. Syst.* **20**:2 (2008), 275–311. MR 2009a:37026 Zbl 1151.37019

[Furstenberg 1967] H. Furstenberg, "Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation", *Math. Systems Theory* **1** (1967), 1–49. MR 35 #4369 Zbl 0146.28502

[Furstenberg 1981] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, 1981. MR 82j:28010 Zbl 0459.28023

[Glasner and Weiss 1995] E. Glasner and B. Weiss, "Quasi-factors of zero-entropy systems", *J. Amer. Math. Soc.* **8**:3 (1995), 665–686. MR 95i:54048 Zbl 0846.28009

[Huang 2006] W. Huang, "Tame systems and scrambled pairs under an abelian group action", *Ergodic Theory Dynam. Systems* **26**:5 (2006), 1549–1567. MR 2007j:37012 Zbl 1122.37009

[Huang and Ye 2005] W. Huang and X. Ye, "Dynamical systems disjoint from any minimal system", *Trans. Amer. Math. Soc.* **357**:2 (2005), 669–694. MR 2005g:37012 Zbl 1072.37011

[Huang and Ye 2006] W. Huang and X. Ye, "A local variational relation and applications", *Israel J. Math.* **151** (2006), 237–279. MR 2006k:37033 Zbl 1122.37013

[Huang et al. 2004] W. Huang, S. Shao, and X. Ye, "Mixing and proximal cells along sequences", *Nonlinearity* **17**:4 (2004), 1245–1260. MR 2005b:37015 Zbl 1055.37014

[Huang et al. 2012] W. Huang, H. Li, and X. Ye, "Family independence for topological and measurable dynamics", *Trans. Amer. Math. Soc.* **364**:10 (2012), 5209–5242. MR 2931327 Zbl 1286.37017

[Kerr and Li 2007] D. Kerr and H. Li, "Independence in topological and C^*-dynamics", *Math. Ann.* **338**:4 (2007), 869–926. MR 2009a:46126 Zbl 1131.46046

[Li 2011] J. Li, "Transitive points via Furstenberg family", *Topology Appl.* **158**:16 (2011), 2221–2231. MR 2012m:37012 Zbl 1234.37034

[Moothathu 2011] T. K. S. Moothathu, "Syndetically proximal pairs", *J. Math. Anal. Appl.* **379**:2 (2011), 656–663. MR 2012c:37025 Zbl 1225.54011

[Mycielski 1964] J. Mycielski, "Independent sets in topological algebras", *Fund. Math.* **55** (1964), 139–147. MR 30 #3855 Zbl 0124.01301

[Oprocha and Zhang 2011] P. Oprocha and G. Zhang, "On local aspects of topological weak mixing in dimension one and beyond", *Studia Math.* **202**:3 (2011), 261–288. MR 2012e:37011 Zbl 1217.37012

[Oprocha and Zhang 2012] P. Oprocha and G. Zhang, "On sets with recurrence properties, their topological structure and entropy", *Topology Appl.* **159**:7 (2012), 1767–1777. MR 2904065 Zbl 1244.37013

[Oprocha and Zhang 2013] P. Oprocha and G. Zhang, "On weak product recurrence and synchronization of return times", *Adv. Math.* **244** (2013), 395–412. MR 3077877 Zbl 1286.37015

[Oprocha and Zhang 2014] P. Oprocha and G. Zhang, "On local aspects of topological weak mixing, sequence entropy and chaos", *Ergodic Theory Dynam. Systems* **34**:5 (2014), 1615–1639. MR 3255435 Zbl 06366533

[Petersen 1970] K. E. Petersen, "Disjointness and weak mixing of minimal sets", *Proc. Amer. Math. Soc.* **24** (1970), 278–280. MR 40 #3522 Zbl 0188.55503

[Xiong 2005] J. Xiong, "Chaos in a topologically transitive system", *Sci. China Ser. A* **48**:7 (2005), 929–939. MR 2006g:37022 Zbl 1096.37018

[Xiong and Yang 1991] J. C. Xiong and Z. G. Yang, "Chaos caused by a topologically mixing map", pp. 550–572 in *Dynamical systems and related topics* (Nagoya, 1990), edited by K. Shiraiwa, Adv. Ser. Dynam. Systems **9**, World Sci. Publ., River Edge, NJ, 1991. MR 93c:58153

[Ye and Zhang 2008] X. Ye and R. Zhang, "On sensitive sets in topological dynamics", *Nonlinearity* **21**:7 (2008), 1601–1620. MR 2009j:37017 Zbl 1153.37322

Received November 27, 2013. Revised December 10, 2014.

JIAN LI
DEPARTMENT OF MATHEMATICS
SHANTOU UNIVERSITY
DAXUE ROAD NO. 243
SHANTOU, GUANGDONG 515063
CHINA
lijian09@mail.ustc.edu.cn

PIOTR OPROCHA
FACULTY OF APPLIED MATHEMATICS
AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY
AL. MICKIEWICZA 30
30-059 KRAKÓW
POLAND
oprocha@agh.edu.pl

GUOHUA ZHANG
SCHOOL OF MATHEMATICAL SCIENCES AND LMNS
FUDAN UNIVERSITY
HANDAN ROAD NO. 220
SHANGHAI 200433
CHINA
and
SHANGHAI CENTER FOR MATHEMATICAL SCIENCES
HANDAN ROAD NO. 220
SHANGHAI 200433
CHINA
chiaths.zhang@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.425

REPRÉSENTATIONS DE STEINBERG MODULO p POUR UN GROUPE RÉDUCTIF SUR UN CORPS LOCAL

TONY LY

Soient F un corps local non archimédien localement compact de caractéristique résiduelle p et G un groupe réductif sur F. Soit R un corps de coefficients de caractéristique p. Nous montrons l'irréductibilité et l'admissibilité des représentations (lisses) de Steinberg généralisées de $G(F)$ sur R. Cela généralise les travaux de Grosse-Klönne et Herzig pour le cas où G est un groupe réductif déployé sur F.

Let F be a locally compact non-Archimedean local field of residue characteristic p and let G be a reductive group over F. Let R be a field of characteristic p. We prove the admissibility and the irreducibility of the so-called smooth generalized Steinberg representations of $G(F)$ over R. This generalizes previous works of Grosse-Klönne and Herzig for the case of G a split reductive group.

1. Introduction

Au début des années 1950, Steinberg [1951] introduit de nouvelles représentations (à coefficients complexes) pour le groupe général linéaire sur un corps fini. Quelques années plus tard, Curtis [1966] donne une formule agréable pour leur caractère. C'est dans l'esprit de cette dernière que sont aujourd'hui définies les représentations de Steinberg généralisées.

Grosse-Klönne [2014] établit leur admissibilité et leur irréductibilité lorsque G est un groupe classique déployé sur F et R est un corps de caractéristique $p > 0$. Ensuite, Herzig [2011] utilise les résultats préliminaires de [Grosse-Klönne 2014] et sa machinerie propre pour étendre ces propriétés à tout groupe réductif déployé G sur F de caractéristique nulle (mais avec R algébriquement clos).

On notera de fait que le résultat principal de [Herzig 2011] pour G le groupe général linéaire déployé met en emphase l'importance des représentations de Steinberg généralisées puisqu'avec les représentations dites supersingulières, elles représentent les « briques fondamentales » pour construire toutes les représentations lisses admissibles irréductibles modulo p de G.

MSC2010 : 20C08, 22E50.
Mots-clefs : Steinberg representations, mod p representations, Hecke algebra.

L'objet de cette note est d'étendre les résultats de [Grosse-Klönne 2014] et de [Herzig 2011] sur les Steinberg généralisées pour G un groupe réductif quelconque. On développe ainsi l'analogue des paragraphes 2 et 3 de [Grosse-Klönne 2014] et du paragraphe 7 de [Herzig 2011] dans le cas non déployé.

2. Contexte général

Soient p un nombre premier et $\overline{\mathbb{F}}_p$ une clôture algébrique fixée de $\mathbb{F}_p$; tout corps fini de caractéristique p sera vu comme un sous-corps de $\overline{\mathbb{F}}_p$. Toute représentation considérée ici sera lisse.

Soit F un corps local non archimédien localement compact à corps résiduel fini k_F de caractéristique p. Soient F^{sep} une clôture séparable de F et $F^{\mathrm{un}} \subseteq F^{\mathrm{sep}}$ la sous-extension maximale non ramifiée de F. On note $\mathcal{I} = \mathrm{Gal}(F^{\mathrm{sep}}/F^{\mathrm{un}})$ le sous-groupe d'inertie du groupe de Galois absolu de F et $\sigma \in \mathrm{Gal}(F^{\mathrm{un}}/F)$ le générateur topologique correspondant à un Frobenius arithmétique.

Soit G un groupe réductif connexe sur F. Kottwitz [1997, paragraphe 7] définit un morphisme fonctoriel et surjectif

$$\kappa_G : G(F^{\mathrm{un}}) \to X^*(Z(\widehat{G})^{\mathcal{I}}),$$

où $\widehat{G}$ désigne le dual de Langlands connexe de G et $Z(\widehat{G})$ son centre. On note $\mathcal{B}$ l'immeuble de Bruhat–Tits du groupe adjoint $G_{\mathrm{ad}}(F^{\mathrm{un}})$. Un sous-groupe parahorique de G est un groupe de la forme [1]

$$\ker \kappa_G \cap G(F) \cap \mathrm{Fix}\, \mathcal{F}$$

pour une facette σ-invariante $\mathcal{F}$ de $\mathcal{B}$ (voir [Bruhat et Tits 1984, 5.2.6 ; Henniart et Vignéras 2015, paragraphe 3.3 ; Haines 2009, paragraphe 8]). Si $\mathcal{F}$ est une chambre, on parle de *sous-groupe d'Iwahori*.

Si H est un sous-groupe parahorique de G associé à une facette σ-invariante $\mathcal{F}$ de $\mathcal{B}$, on lui associe un groupe $H \leq \widetilde{H} \leq G(F)$ défini par :

$$\widetilde{H} := \{g \in G(F) \cap \mathrm{Fix}\, \mathcal{F} \mid \kappa_G(g) \text{ est de torsion}\}.$$

Soient K un sous-groupe parahorique maximal spécial de $G(F)$ et $K(1)$ le pro-p-radical de K (voir paragraphe 3.6 de [Henniart et Vignéras 2015]). Le quotient $K/K(1)$ est le groupe des k_F-points d'un groupe réductif $\overline{G}$.

Soient T un tore maximal parmi les tores F-déployés de G et A le centralisateur de T dans G. Soient B un parabolique minimal de G de composante de Levi A et U son radical unipotent. Lorsque Q est un parabolique contenant B, on notera Q^- son parabolique opposé au sens du Théorème 4.15 de [Borel et Tits 1965].

---

[1] On a noté $\mathrm{Fix}\, \mathcal{F}$ le fixateur point par point des simplexes de dimension 0 composant $\mathcal{F}$.

Pour tout sous-groupe H de G défini sur F, on note

$$\overline{H} := (H(F) \cap K)/(H(F) \cap K(1))$$

le sous-groupe de $\overline{G}(k_F)$ correspondant.

On confondra par abus G et ses sous-groupes paraboliques (resp. composante de Levi, radical unipotent de sous-groupes paraboliques) avec leurs F-points. De même pour $\overline{G}$ et ses sous-groupes paraboliques (resp. composante de Levi, radical unipotent de sous-groupes paraboliques) avec leurs k_F-points.

3. Définitions et résultats

Soit R un anneau commutatif unitaire. Soit Q un sous-groupe parabolique standard (c'est-à-dire qui contient[2] B) de G. On définit la G-représentation à coefficients dans R suivante :

$$\mathrm{St}_Q\, R := \frac{\mathrm{Ind}_Q^G \mathrm{id}}{\sum_{Q' \gneq Q} \mathrm{Ind}_{Q'}^G \mathrm{id}}.$$

Ici, comme dans toute la suite, on a noté Ind le foncteur d'induction lisse et on fait agir G sur $\mathrm{Ind}_Q^G \mathrm{id}$ par translation à droite.

Théorème 3.1. *Soit R un corps de caractéristique p. La représentation de Steinberg généralisée* St_Q *R est irréductible et admissible.*

Remarque. Lorsque F est de caractéristique 0, l'admissibilité suit automatiquement de [Vignéras 2012b]. Par contre, lorsque F est de caractéristique p, à ma connaissance on ne sait pas se passer de l'argument de cet article.

On va présenter une preuve de cet énoncé dans les paragraphes qui suivent. Commençons par énoncer un corollaire (on dira un mot de la preuve dans le paragraphe 9).

Corollaire 3.2. *Les constituants de Jordan–Hölder de* $\mathrm{Ind}_Q^G \mathrm{id}$ *sont les* $\mathrm{St}_{Q'}$ *R pour les sous-groupes paraboliques*[3] *Q' contenant Q. Ils sont deux à deux non isomorphes et de multiplicité* 1.

On donnera aussi la filtration par les cosocles de $\mathrm{Ind}_Q^G \mathrm{id}$.

---

[2] Cette hypothèse n'est en fait utile que lorsque l'on veut utiliser la comparaison parabolique-compacte de [Henniart et Vignéras 2012].

[3] Comme Q' contient Q standard, il l'est automatiquement aussi.

4. Sous-groupe d'Iwahori et sous-groupes radiciels

Soient Φ le système de racines de G associé à T et Φ_{red} le système réduit associé :

$$\Phi_{\mathrm{red}} = \{a \in \Phi \mid a/2 \notin \Phi\}.$$

Le groupe parabolique minimal B nous fournit un sous-ensemble de racines positives $\Phi^+_{\mathrm{red}} \subseteq \Phi_{\mathrm{red}}$ et un système de racines simples Δ. On note $\Phi^-_{\mathrm{red}} := \Phi_{\mathrm{red}} \smallsetminus \Phi^+_{\mathrm{red}}$ et, pour $\alpha \in \Phi_{\mathrm{red}}$, on appelle s_α la réflexion correspondante. On note W le groupe de Weyl fini (déterminé par T) et $l : W \to \mathbb{N}$ la longueur (déterminée par Δ). Soit w_0 l'élément le plus long de W.

Pour $w \in W$, on notera encore w un relèvement (fixé une fois pour toutes) de w dans le normalisateur $N_G(T) \cap K$ de T dans K (ce qui est possible car K est spécial) ou bien l'image de ce relèvement dans $\overline{G}$.

On fait remarquer que le paragraphe 1 de [Grosse-Klönne 2014] ne concerne que des groupes de réflexions cristallographiques avec système de racines réduit. Il est donc valable lorsque l'on travaille avec Φ_{red} et le lecteur ne devra pas être surpris quand on y fera référence.

Soit I le sous-groupe d'Iwahori de G suivant : si x_0 est le point spécial de l'immeuble de $G_{\mathrm{ad}}(F^{\mathrm{un}})$ fixe par K, et si $\mathscr{C}$ est la chambre de sommet x_0 et fixe par B, alors I est le parahorique fixant $\mathscr{C}$ (voir paragraphe 2). On dispose alors de la décomposition d'Iwasawa (voir [Bruhat et Tits 1972, Proposition 7.3.1]) :[4]

$$G = \bigsqcup_{w \in W} Bw\tilde{I}.$$

Aussi, on a des injections naturelles

$$A \cap \widetilde{K}/A \cap K \hookrightarrow \tilde{I}/I \hookrightarrow \widetilde{K}/K.$$

La composée est un isomorphisme par [Henniart et Vignéras 2015, Lemma 6.2(iii)] ; la première flèche est donc un isomorphisme

$$A \cap \widetilde{K}/A \cap K \xrightarrow{\sim} \tilde{I}/I.$$

On a donc finalement

$$G = \bigsqcup_{w \in W} Bw(A \cap \widetilde{K})I = \bigsqcup_{w \in W} BwI, \tag{1}$$

---

[4] Ici, comme dans tout ce qui suit, l'appel à [Bruhat et Tits 1972] nécessite de faire attention que cela est bien loisible : c'est l'objet du Théorème 5.1.20 de [Bruhat et Tits 1984], comme expliqué dans son introduction. Par la suite, on gardera cet énoncé en tête sans le rappeler à chaque fois, mais on se permet d'insister que son importance est cruciale.

où la seconde égalité vient de l'inclusion $w(A \cap K)w^{-1} \subseteq A \cap K \subseteq B$. Pour tout sous-groupe H de G, on pose $H^0 := H \cap I$.

Pour $\alpha \in \Phi$, on note U_α le sous-groupe radiciel associé. Comme on a l'inclusion $U_{2\alpha} \subsetneq U_\alpha$ si $\{\alpha, 2\alpha\} \subseteq \Phi$, il convient de remarquer aussi $U^0_{2\alpha} \subseteq U^0_\alpha$. Par la Proposition 6.1.6 de [Bruhat et Tits 1972], on a donc

$$\prod_{\alpha \in \Phi^+_{\text{red}}} U_\alpha = U, \qquad \prod_{\alpha \in \Phi^+_{\text{red}}} U^0_\alpha = U^0,$$

quel que soit l'ordre choisi sur Φ^+_{red}.

Soit J un sous-ensemble de Δ. Les s_α pour $\alpha \in J$ engendrent un sous-groupe W_J de W. On note aussi

$$W^J := \{w \in W \mid \forall \alpha \in J,\ l(ws_\alpha) > l(w)\}.$$

On a, par [Humphreys 1992, Lemma 1.6 et Corollary 1.7],

$$W^J = \{w \in W \mid w(J) \subseteq \Phi^+_{\text{red}}\}. \tag{2}$$

Aussi, grâce à [Humphreys 1992, Proposition 1.10 et paragraphe 5.12], on a le fait important suivant : l'ensemble W^J est un système de représentants de W/W_J contenant l'élément le plus court de chaque classe.

On note le sous-ensemble de W^J constitué des éléments primitifs

$$W^J_{\text{pr}} := W^J \smallsetminus \bigcup_{\alpha \in \Delta \smallsetminus J} W^{J \cup \{\alpha\}} = \{w \in W^J \mid w(\Delta \smallsetminus J) \subseteq \Phi^-_{\text{red}}\},$$

de sorte que l'on a $W = \bigsqcup W^J_{\text{pr}}$, lorsque J parcourt les sous-ensembles de Δ.

On définit aussi le sous-ensemble suivant de Φ_{red} :

$$W_J.J := \{w\alpha \mid w \in W_J,\ \alpha \in J\}.$$

5. Détermination de $(\mathrm{St}_Q\, R)^I$

Soit R un anneau commutatif unitaire.

Soit J un sous-ensemble propre de Δ. Pour $w \in W^{J \cup \{\alpha\}}$ et $\alpha \in \Delta \smallsetminus J$, on définit

$$\partial(w) := \sum_{\substack{w' \in W^J \\ w' W_J \subseteq w W_{J \cup \{\alpha\}}}} w' \in R[W^J],$$

où $R[W^J]$ désigne le R-module libre de base les éléments de W^J. En prolongeant par R-linéarité, on a la suite exacte

$$\bigoplus_{\alpha \in \Delta \smallsetminus J} R[W^{J \cup \{\alpha\}}] \xrightarrow{\partial} R[W^J] \xrightarrow{\nabla} \mathfrak{M}_J(R) \to 0, \tag{3}$$

qui définit l'application linéaire ∇ et le R-module $\mathfrak{M}_J(R)$. Ce dernier module est un objet essentiel pour la compréhension du module des I-invariants de $\mathrm{St}_J\,R$. Et Grosse-Klönne [2014, Proposition 1.3(a)] a démontré que $\mathfrak{M}_J(R)$ est libre de rang $|W^J_{\mathrm{pr}}|$.

Par la Proposition 2.4 de [Bushnell et Henniart 2006] appliqué[5] à $H = \{1\}$ et au groupe profini I, le foncteur des fonctions localement constantes $C^\infty(I,\cdot)$ est exact. En appliquant $C^\infty(I,\cdot)$ à (3), on obtient la suite exacte

$$C^\infty\Bigl(I, \bigoplus_{\alpha\in\Delta\smallsetminus J} R[W^{J\cup\{\alpha\}}]\Bigr) \to C^\infty(I, R[W^J]) \to C^\infty(I, \mathfrak{M}_J(R)) \to 0.$$

Par abus, on note encore ces flèches ∂ et ∇ respectivement.

Notons, pour $J \subseteq \Delta$ et $w \in W$:

$$P_J := BW_JB, \quad {}^wP_J := wP_Jw^{-1}. \tag{4}$$

On remarque que ${}^wP_J$ ne dépend que de la classe de w dans W/W_J. Par définition de P_J son radical unipotent est égal à

$$N_J = \prod_{\alpha\in\Phi^+_{\mathrm{red}}\smallsetminus(\Phi^+_{\mathrm{red}}\cap W_J.J)} U_\alpha.$$

On fixe aussi un sous-groupe de Levi M_J de P_J contenant A et W_J. Énonçons tout de suite une inclusion entre les ${}^wP_J$ qui nous sera utile dans très peu de temps.

Lemme 5.1. *Soit $\alpha \in \Delta \smallsetminus J$. Soient w et w' des éléments de W vérifiant $w'W_J \subseteq wW_{J\cup\{\alpha\}}$. On a les inclusions*[6]

$${}^{w'}P_J \subseteq {}^wP_{J\cup\{\alpha\}}; \quad {}^{w'}P^0_J \subseteq {}^wP^0_{J\cup\{\alpha\}}.$$

Démonstration. Par hypothèse, il existe un élément $\sigma \in W_{J\cup\{\alpha\}}$ vérifiant $w' = w\sigma$. La première inclusion vient immédiatement :

$${}^{w'}P_J = w\sigma P_J\sigma^{-1}w^{-1} \subseteq {}^wP_{J\cup\{\alpha\}}.$$

La seconde suit par intersection avec I. □

Pour $\alpha \in \Delta \smallsetminus J$, en notant $C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R)$ le sous-espace de $C^\infty(I, R)$ constitué des fonctions ${}^wP^0_{J\cup\{\alpha\}}$-invariantes à gauche, on a une injection

$$\bigoplus_{w\in W^{J\cup\{\alpha\}}} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R) \hookrightarrow C^\infty(I, R[W^{J\cup\{\alpha\}}])$$

---

[5] Dans cette référence toutes les représentations sont à coefficients complexes mais cela n'influe pas sur la preuve de la proposition en question.

[6] On omet les parenthèses pour alléger les notations, mais ${}^wP^0_J$ doit se lire $({}^wP_J)^0$.

donnée par la flèche

$$(f_{\alpha,w})_w \mapsto \sum_w f_{\alpha,w} w.$$

On a de même une injection

$$\bigoplus_{w\in W^J} C^\infty({}^wP_J^0\backslash I, R) \hookrightarrow C^\infty(I, R[W^J]).$$

On verra dorénavant les injections précédentes comme des inclusions. On a alors le diagramme commutatif suivant :

$$\begin{tikzcd}
C^\infty(I, \bigoplus_{\alpha\in\Delta\smallsetminus J} R[W^{J\cup\{\alpha\}}]) \arrow[r, "\partial"] & C^\infty(I, R[W^J]) \arrow[r, "\nabla"] & C^\infty(I, \mathfrak{M}_J(R)) \arrow[r] & 0 \\
\bigoplus_{\substack{\alpha\in\Delta\smallsetminus J\\ w\in W^{J\cup\{\alpha\}}}} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R) \arrow[u, hook] \arrow[r, "\partial"] & \bigoplus_{w\in W^J} C^\infty({}^wP_J^0\backslash I, R) \arrow[u, hook] \arrow[r, "\nabla"] & C^\infty(I, \mathfrak{M}_J(R)) \arrow[u, equal] &
\end{tikzcd} \tag{5}$$

Vérifions que l'image de $\bigoplus_{\alpha,w} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R)$ par ∂ est bien incluse dans $\bigoplus_{W^J} C^\infty({}^wP_J^0\backslash I, R)$. Fixons pour cela $\alpha\in\Delta\smallsetminus J$. On a

$$\partial\Bigl(\sum_{w\in W^{J\cup\{\alpha\}}} f_{\alpha,w} w\Bigr) = \sum_{w'\in W^J}\Bigl(\sum_{\substack{w\in W^{J\cup\{\alpha\}}\\ w'W_J\subseteq wW_{J\cup\{\alpha\}}}} f_{\alpha,w}\Bigr) w'; \tag{6}$$

il s'agit donc de voir que $\sum_{w'W_J\subseteq wW_{J\cup\{\alpha\}}} f_{\alpha,w}$ est ${}^{w'}P_J^0$-invariante à gauche. En fait, chacun des termes de cette somme l'est. En effet, chaque $f_{\alpha,w}$ est ${}^wP^0_{J\cup\{\alpha\}}$-invariant à gauche ; par le lemme 5.1, il est aussi ${}^{w'}P_J^0$-invariant et ∂ envoie bien $\bigoplus_{\alpha,w} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R)$ dans $\bigoplus_{W^J} C^\infty({}^wP_J^0\backslash I, R)$.

Proposition 5.2. *La ligne du bas du diagramme* (5) *est exacte.*

L'introduction de quelques objets est nécessaire avant d'aborder la preuve de cette proposition. Remarquons que, par [Humphreys 1992, Corollary 1.5], le sous-système $\Phi_J\subseteq\Phi_{\mathrm{red}}$ engendré par J vérifie

$$\Phi_J^- = \Phi_{\mathrm{red}}^- \cap W_J.J.$$

On dit que $D\subsetneq\Phi_{\mathrm{red}}$ est *J-quasi-parabolique* s'il est l'intersection de certains $w(\Phi_{\mathrm{red}}^-\smallsetminus\Phi_J^-)$.

Énonçons maintenant un lemme/définition particulièrement utile.

Lemme 5.3. *Soit $D\subsetneq\Phi_{\mathrm{red}}$ une partie J-quasi-parabolique. Le produit $\prod_D U_\alpha^0$ est indépendant de l'ordre choisi sur D : il forme un sous-groupe de G que l'on notera U_D^0.*

Démonstration. Comme D est J-quasi-parabolique, on voit qu'il est suffisant de prouver que $\prod_{w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J)} U^0_\alpha$ est indépendant de l'ordre sur $w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J)$ pour un $w \in W$ tel que D est contenu dans $w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J)$. Quitte à conjuguer par w, on suppose à présent $w = 1$.

Il reste donc à voir la condition de commutateurs sur la partie $\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J \subseteq \Phi^-_{\mathrm{red}}$ pour pouvoir appliquer la Proposition[7] 6.1.6 de [Bruhat et Tits 1972]. Il s'agit de voir dans un premier temps que si α et β sont des éléments de $\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J$ alors on a

$$(U_\alpha, U_\beta) \subseteq \langle U_{n\alpha+m\beta} \mid n\alpha + m\beta \in \Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J,\ m, n \in \mathbb{N}^* \rangle. \tag{7}$$

On sait déjà par l'axiomatique de Bruhat–Tits que (U_α, U_β) est inclus dans $\langle U_{n\alpha+m\beta} \mid n\alpha+m\beta \in \Phi^-_{\mathrm{red}},\ m, n \in \mathbb{N}^* \rangle$. Puis on remarque que, comme on a $\alpha, \beta \notin \Phi^-_J$, aucune des sommes $n\alpha + m\beta$ n'appartient à Φ^-_J non plus.

En prenant les intersections avec I, parce que $I \cap \prod_{\alpha \in \Phi^-_{\mathrm{red}}} U_\alpha$ est égal à $\prod_{\alpha \in \Phi^-_{\mathrm{red}}} U^0_\alpha$ par la Proposition[8] 5.2.32 de [Bruhat et Tits 1972], (7) devient

$$(U^0_\alpha, U^0_\beta) \subseteq \langle U^0_{n\alpha+m\beta} \mid n\alpha + m\beta \in \Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J,\ m, n \in \mathbb{N}^* \rangle.$$

Le lemme en découle. □

Pour $w \in W$, on pose

$$D_w = w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J) :$$

le groupe $U^0_{D_w}$ est l'intersection de I avec le radical unipotent de ${}^wP^-_J$ (défini en (4)) (voir [Demazure 2011b, Proposition 1.12]). On remarque que D_w ne dépend que de la classe de w dans W/W_J, et que, pour tout $\alpha \in \Delta \smallsetminus J$, on a $w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_{J\cup\{\alpha\}}) \subseteq w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J)$.

L'introduction des ensembles J-quasi-paraboliques s'explique alors par le fait que l'on va s'intéresser à l'intersection de tels $U^0_{D_w}$ pour différents $w \in W$ et à la décomposition d'Iwahori qui suit.

Lemme 5.4. *Soient $J \subseteq \Delta$ et $w \in W$. Le produit ${}^wP^0_J \times U^0_{D_w} \to I$ est un homéomorphisme.*

Démonstration. Par la Proposition 5.2.32 de [Bruhat et Tits 1972], on a la décomposition d'Iwahori

$$\tilde{I} = \Big(\prod_{\beta \in w\Phi^+_{\mathrm{red}}} U^0_\beta \Big) (A \cap \widetilde{K}) \Big(\prod_{\beta \in w\Phi^-_{\mathrm{red}}} U^0_\beta \Big),$$

où les produits sur $w\Phi^+_{\mathrm{red}}$ et $w\Phi^-_{\mathrm{red}}$ sont indépendants de l'ordre par la Proposition 6.1.6 de [Bruhat et Tits 1972]. En l'intersectant avec le noyau du morphisme

---

[7] On l'applique à $Y_\alpha = U_\alpha \cap I \supseteq Y_{2\alpha} = U_{2\alpha} \cap I$ et $X_\alpha = U_\alpha \cap I$.

[8] Cette dernière s'occupe de $U \cap \tilde{I}$, mais en prenant l'intersection avec le noyau du morphisme de Kottwitz, on voit l'égalité $U \cap \tilde{I} = U \cap I$.

de Kottwitz κ_G, parce que l'on a $A^0 = A \cap K$, cela permet d'écrire

$$I = \Big(\prod_{\beta \in w\Phi^+_{\mathrm{red}}} U^0_\beta \Big) A^0 \Big(\prod_{\beta \in w\Phi^-_{\mathrm{red}}} U^0_\beta \Big) =: I^+ A^0 I^-.$$

La décomposition

$$ {}^w P^0_J = \Big(\prod_{\beta \in w\Phi^+_{\mathrm{red}}} U^0_\beta \Big) A^0 \Big(\prod_{\beta \in w\Phi^-_J} U^0_\beta \Big) \tag{8}$$

suit par intersection avec ${}^w P_J$: I^+ et A^0 sont inclus dans ${}^w P_J$ et donc ${}^w P^0_J$ est égal à $I^+ A^0 (I^- \cap {}^w P_J)$. Il en résulte la bijection

$$U^0_{D_w} = {}^w U^0_{\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J} \xrightarrow{\sim} {}^w P^0_J \backslash I. \tag{9}$$

On en déduit que le produit ${}^w P^0_J \times U^0_{D_w} \to I$ est un homéomorphisme : c'est une bijection par (9), le produit est continu et une bijection continue d'un espace compact vers un espace séparé est un homéomorphisme. □

Pour un sous-ensemble D de Φ_{red}, on note

$$W^J(D) := \{ w \in W^J \mid D \subseteq w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J) \}. \tag{10}$$

Lemme 5.5. *Soient $D, D' \subseteq \Phi_{\mathrm{red}}$ deux ensembles J-quasi-paraboliques. Soient $\alpha \in \Delta \smallsetminus J$ et $w \in W^{J \cup \{\alpha\}}(D)$. On a l'égalité d'ensembles*

$$ {}^w P^0_{J\cup\{\alpha\}} U^0_D \cap {}^w P^0_{J\cup\{\alpha\}} U^0_{D'} = {}^w P^0_{J\cup\{\alpha\}} (U^0_D \cap U^0_{D'}). \tag{11}$$

Démonstration. L'inclusion $\supseteq$ est évidente ; prouvons $\subseteq$. En appliquant le lemme 5.3, le produit $\prod_{\beta \in D} U^0_\beta$ est indépendant de l'ordre. Choisissons un ordre sur D tel que tout élément de $D \smallsetminus (D \cap D')$ soit placé avant tout élément de $D \cap D'$. On forme le produit $\prod_{\beta \in D \smallsetminus (D \cap D')} U^0_\beta$ suivant cet ordre et on le note $U^0_{D \smallsetminus D'}$, de sorte que l'on a

$$U^0_D = U^0_{D \smallsetminus D'} U^0_{D \cap D'}$$

($D \cap D'$ est J-quasi-parabolique et la notation est celle du lemme 5.3). Il est équivalent de voir $\subseteq$ dans (11) et

$$ {}^w P^0_{J\cup\{\alpha\}} U^0_{D \smallsetminus D'} \cap U^0_{D'} \subseteq {}^w P^0_{J\cup\{\alpha\}} (U^0_D \cap U^0_{D'}).$$

On va même montrer que ${}^w P^0_{J\cup\{\alpha\}} U^0_{D \smallsetminus D'} \cap U^0_{D'}$ est inclus dans ${}^w P^0_{J\cup\{\alpha\}}$. Notons

$$\Phi' = \Phi_{\mathrm{red}} \smallsetminus w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_{J\cup\{\alpha\}}) = \{ \beta \in \Phi_{\mathrm{red}} \mid U_\beta \subseteq {}^w P^0_{J\cup\{\alpha\}} \}. \tag{12}$$

Puisque w est dans $W^{J\cup\{\alpha\}}(D)$, l'intersection de D et de Φ' est vide ; ou autrement dit, la partie $(J \cup \{\alpha\})$-quasi-parabolique

$$D_{\alpha,w} = w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_{J\cup\{\alpha\}})$$

contient D. Par le lemme 5.3 pour $D_{\alpha,w}$, le produit $\prod_{\beta\in D_{\alpha,w}} U^0_\beta$ est indépendant de l'ordre : on choisit un ordre sur $D_{\alpha,w}$ tel que sa restriction à D coïncide avec l'ordre précédemment choisi sur D et tel que tout élément de D précède tout élément de $D_{\alpha,w} \smallsetminus D$. On forme le produit $\prod_{\beta\in D_{\alpha,w}\smallsetminus D} U^0_\beta$ suivant cet ordre et on le note $U^0_{D_{\alpha,w}\smallsetminus D}$, de sorte que l'on a

$$U^0_{D_{\alpha,w}} = U^0_D U^0_{D_{\alpha,w}\smallsetminus D}.$$

Par (9), on a alors la décomposition en produit direct d'ensembles

$$I = {}^wP^0_{J\cup\{\alpha\}} U^0_{D\smallsetminus D'} U^0_{D\cap D'} U^0_{D_{\alpha,w}\smallsetminus D}. \tag{13}$$

Par le (8) dans la preuve du lemme 5.4, on a la décomposition d'Iwahori

$${}^wP^0_{J\cup\{\alpha\}} = \Bigl(\prod_{\beta\in w\Phi^+_{\mathrm{red}}} U^0_\beta\Bigr) A^0 \Bigl(\prod_{\beta\in w\Phi^-_{J\cup\{\alpha\}}} U^0_\beta\Bigr). \tag{14}$$

Parce que $w\Phi^+_{\mathrm{red}}$ vérifie la condition des commutateurs, on peut appliquer la Proposition[9] 6.1.6 de [Bruhat et Tits 1972] et dire que $\prod_{\beta\in w\Phi^+_{\mathrm{red}}} U^0_\beta$ est un produit indépendant de l'ordre sur $w\Phi^+_{\mathrm{red}}$, que l'on notera $U^0_{w\Phi^+_{\mathrm{red}}}$. On choisit un ordre sur $w\Phi^+_{\mathrm{red}}$ tel que tout élément de $w\Phi^+_{\mathrm{red}} \cap D'$ précède tout élément de $w\Phi^+_{\mathrm{red}} \smallsetminus (D' \cap w\Phi^+_{\mathrm{red}})$. On forme

$$U^0_{w\Phi^+_{\mathrm{red}}\cap D'} = \prod_{\beta\in w\Phi^+_{\mathrm{red}}\cap D'} U^0_\beta, \quad U^0_{w\Phi^+_{\mathrm{red}}\smallsetminus D'} = \prod_{\beta\in w\Phi^+_{\mathrm{red}}\smallsetminus (D'\cap w\Phi^+_{\mathrm{red}})} U^0_\beta,$$

et on a

$$U^0_{w\Phi^+_{\mathrm{red}}} = U^0_{w\Phi^+_{\mathrm{red}}\cap D'} U^0_{w\Phi^+_{\mathrm{red}}\smallsetminus D'}.$$

De la même manière, on choisit un ordre sur $w\Phi^-_{J\cup\{\alpha\}}$ tel que tout élément de $w\Phi^-_{J\cup\{\alpha\}}$ qui appartient à D' précède tout élément qui n'y appartient pas : on obtient $U^0_{w\Phi^-_{J\cup\{\alpha\}}} = U^0_{w\Phi^-_{J\cup\{\alpha\}}\cap D'} U^0_{w\Phi^-_{J\cup\{\alpha\}}\smallsetminus D'}$. L'identité (14) devient

$${}^wP^0_{J\cup\{\alpha\}} = U^0_{w\Phi^+_{\mathrm{red}}\cap D'} U^0_{w\Phi^+_{\mathrm{red}}\smallsetminus D'} A^0 U^0_{w\Phi^-_{J\cup\{\alpha\}}\cap D'} U^0_{w\Phi^-_{J\cup\{\alpha\}}\smallsetminus D'};$$

et (13) devient le produit direct

$$I = U^0_{w\Phi^+_{\mathrm{red}}\cap D'} U^0_{w\Phi^+_{\mathrm{red}}\smallsetminus D'} A^0 U^0_{w\Phi^-_{J\cup\{\alpha\}}\cap D'} U^0_{w\Phi^-_{J\cup\{\alpha\}}\smallsetminus D'} U^0_{D\smallsetminus D'} U^0_{D\cap D'} U^0_{D_{\alpha,w}\smallsetminus D}. \tag{15}$$

Grâce à (15), un élément de $U^0_{D'} \cap {}^wP^0_{J\cup\{\alpha\}} U^0_{D\smallsetminus D'}$ est dans le produit

$$U^0_{w\Phi^+_{\mathrm{red}}\cap D'} U^0_{w\Phi^-_{J\cup\{\alpha\}}\cap D'} \subseteq {}^wP^0_{J\cup\{\alpha\}}$$

et le lemme est prouvé. □

---

[9] Ici encore, on prend $Y_a = U_a \cap I \supseteq Y_{2a} = U_{2a} \cap I$ et $X_a = U_a \cap I$.

Démonstration de la proposition 5.2. On commence par numéroter tous les sous-ensembles J-quasi-paraboliques de Φ_{red} : $D_0, D_1, D_2, \ldots$ par ordre croissant de taille, c'est-à-dire avec

$$n < m \implies |D_n| \leq |D_m|.$$

Soit $f \in C^{\infty}(I, R[W^J])$, image de

$$(f_w)_{w \in W^J} \in \bigoplus_{w \in W^J} C^{\infty}({}^{w}P_J^0 \backslash I, R)$$

par la flèche verticale du diagramme (5), tel que l'on a $\nabla f = 0$, c'est-à-dire $\nabla(f(x)) = 0$ pour tout $x \in I$. On cherche $g \in C^{\infty}(I, \bigoplus_{\alpha \in \Delta \smallsetminus J} R[W^{J \cup \{\alpha\}}])$, image de

$$(g_{\alpha,w})_{\alpha,w} \in \bigoplus_{\substack{\alpha \in \Delta \smallsetminus J \\ w \in W^{J \cup \{\alpha\}}}} C^{\infty}({}^{w}P_{J \cup \{\alpha\}}^0 \backslash I, R),$$

vérifiant $f = \partial g$.

On va montrer par récurrence l'existence d'une telle fonction g satisfaisant $f = \partial g$ sur $\bigcup_{n \geq 0} U_{D_n}^0$. Ceci implique $f = \partial g$ car f et ∂g proviennent de la ligne du bas du diagramme (5) et que l'on a ${}^{w}P_J^0\big(\bigcup_{n \geq 0} U_{D_n}^0\big) = I$ pour tout $w \in W^J$ par le lemme 5.4. La démonstration se fera en deux étapes :

— il existe g telle que f est égale à ∂g sur $U_{D_0}^0$;

— si f est nulle sur $\bigcup\limits_{n<m} U_{D_n}^0$, alors il existe g vérifiant $f = \partial g$ sur $\bigcup\limits_{n \leq m} U_{D_n}^0$ $(m \geq 1)$.

La situation de l'étape d'initiation est la suivante : on a

$$D_0 = \varnothing = (\Phi_{\mathrm{red}}^- \smallsetminus \Phi_J^-) \cap w_0(\Phi_{\mathrm{red}}^- \smallsetminus \Phi_J^-), \quad U_{D_0}^0 = \{1\}.$$

La fonction f vérifie $(\nabla f)(1) = \nabla(f(1)) = 0$. Comme la suite (3) est exacte, on choisit, une famille d'éléments

$$(g_{\alpha,w}^{(1)})_{\alpha,w} \in \bigoplus_{\alpha \in \Delta \smallsetminus J} R[W^{J \cup \{\alpha\}}]$$

telle que l'on a $\partial((g_{\alpha,w}^{(1)})_{\alpha,w}) = f(1)$. Soit, pour tout $\alpha \in \Delta \smallsetminus J$ et tout $w \in W^{J \cup \{\alpha\}}$, $g_{\alpha,w}$ une fonction de $C^{\infty}({}^{w}P_{J \cup \{\alpha\}}^0 \backslash I, R)$ valant $g_{\alpha,w}^{(1)}$ sur ${}^{w}P_{J \cup \{\alpha\}}^0$. Alors l'image g de

$$(g_{\alpha,w})_{\alpha,w} \in \bigoplus_{\substack{\alpha \in \Delta \smallsetminus J \\ w \in W^{J \cup \{\alpha\}}}} C^{\infty}({}^{w}P_{J \cup \{\alpha\}}^0 \backslash I, R)$$

dans $C^{\infty}(I, \bigoplus_{\alpha \in \Delta \smallsetminus J} R[W^{J \cup \{\alpha\}}])$ vérifie $\partial g = f$ sur $U_{D_0}^0 = \{1\}$. L'étape d'initiation est terminée.

Montrons maintenant la propriété de propagation de la récurrence. Soit, pour tout $w \in W^J$, $f_w \in C^{\infty}({}^{w}P_J^0 \backslash I, R)$ nulle sur $\bigcup_{n<m} U_{D_n}^0$.

Si w est un élément de $W^J \smallsetminus W^J(D_m)$, f_w est nulle sur $U^0_{D_m}$ puisque l'on a ${}^wP^0_J U^0_{D_m} = {}^wP^0_J U^0_{D_n}$ pour $n < m$ vérifiant

$$D_m \cap w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J) = D_n.$$

En effet, il suffit de remarquer qu'un tel $n < m$ existe par définition de $W^J(D_m)$ (voir (10)) et que l'on a alors $U^0_{D_m} = (U^0_{D_m} \cap {}^wP^0_J)U^0_{D_n}$.

Nous allons trouver la fonction g comme image de

$$(g_{\alpha,w})_{\alpha,w} \in \bigoplus_{\substack{\alpha\in\Delta\smallsetminus J \\ w\in W^{J\cup\{\alpha\}}(D_m)}} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R)$$

telle que $(f - \partial g)_{w'}$ est nulle sur $\bigcup_{n\leq m} U^0_{D_n}$ pour tout $w' \in W^J(D_m)$. Comme $(\partial g)_{w'}$ est nulle[10] pour $w' \in W^J \smallsetminus W^J(D_m)$, $(f - \partial g)_{w'}$ sera nulle sur $\bigcup_{n\leq m} U^0_{D_n}$ pour tout $w' \in W^J$.

La fonction f est localement constante sur le compact $\bigcup_{n\leq m} U^0_{D_n}$ et nulle sur $\bigcup_{n<m} U^0_{D_n}$. Soient $(C_i)_{0\leq i\leq r}$ les ouverts disjoints de $\bigcup_{n\leq m} U^0_{D_n}$ vérifiant

$$\bigcup_i C_i = \bigcup_{n\leq m} U^0_{D_n}$$

et tels que f_w est constant sur C_i, égal à $f^{(i)}_w$, et $f^{(0)}_w = 0$ pour tout $w \in W^J(D_m)$ et $(f^{(i)}_w)_{w\in W^J(D_m)} \neq (f^{(i')}_w)_{w\in W^J(D_m)}$ si $i \neq i'$. En particulier, on notera que C_0 contient $\bigcup_{n<m} U^0_{D_n}$ et que $U^0_{D_m}$ est l'union disjointe de $C_0 \cap U^0_{D_m}$ et des C_i pour $1 \leq i \leq r$.

La suite extraite de (3)

$$\bigoplus_{\alpha\in\Delta\smallsetminus J} R[W^{J\cup\{\alpha\}}(D_m)] \xrightarrow{\partial} R[W^J(D_m)] \xrightarrow{\nabla} \mathfrak{M}_J(R) \tag{16}$$

est encore exacte (voir [Grosse-Klönne 2014, Proposition 1.3(b)]). Par ailleurs, $(f^{(i)}_w)_{w\in W^J(D_m)}$ appartient au noyau de ∇ dans (16) car on a $\nabla f = 0$ et $f_w = 0$ pour $w \in W^J \smallsetminus W^J(D_m)$. Il existe alors, pour tout $0 \leq i \leq r$,

$$g^{(i)} = (g^{(i)}_{\alpha,w})_{\alpha\in\Delta\smallsetminus J, w\in W^{J\cup\{\alpha\}}(D_m)} \in \bigoplus_{\alpha\in\Delta\smallsetminus J} R[W^{J\cup\{\alpha\}}(D_m)]$$

tel que l'on a $g^{(0)} = 0$ et

$$\partial(g^{(i)}) = (f^{(i)}_w)_{w\in W^J(D_m)} \in R[W^J(D_m)]. \tag{17}$$

---

[10] En (6), la somme sur les w vérifiant $w'W_J \subseteq wW_{J\cup\{\alpha\}}$ ne fait intervenir que des $w \notin W^{J\cup\{\alpha\}}(D_m)$ car $w'W_J \subseteq wW_{J\cup\{\alpha\}}$ implique ${}^{w'}P_J \subseteq {}^wP_{J\cup\{\alpha\}}$ par le lemme 5.1, et donc $w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_{J\cup\{\alpha\}}) \subseteq w'(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_J)$ par (12) (voir définition (10) aussi).

Posons $D_{\alpha,w} = w(\Phi^-_{\mathrm{red}} \smallsetminus \Phi^-_{J\cup\{\alpha\}})$ pour $\alpha \in \Delta \smallsetminus J$ et $w \in W$. Pour tout $w \in W$, le produit ${}^wP^0_{J\cup\{\alpha\}} \times U^0_{D_{\alpha,w}} \to I$ est un homéomorphisme par le lemme 5.4, donc on a

$$C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R) \simeq C^\infty(U^0_{D_{\alpha,w}}, R).$$

Soient $\alpha \in \Delta \smallsetminus J$ et $w \in W^{J\cup\{\alpha\}}(D_m)$. Notons U' la projection[11] de $\bigcup_{n<m} U^0_{D_n}$ sur $U^0_{D_{\alpha,w}}$: c'est le sous-espace de $U^0_{D_{\alpha,w}}$ vérifiant

$${}^wP^0_{J\cup\{\alpha\}}U' = {}^wP^0_{J\cup\{\alpha\}}\left(\bigcup_{n<m} U^0_{D_n}\right).$$

Supposons qu'il existe, pour $\alpha \in \Delta \smallsetminus J$ et $w \in W^{J\cup\{\alpha\}}(D_m)$, une fonction $g_{\alpha,w} \in C^\infty(U^0_{D_{\alpha,w}}, R)$ nulle sur $U' \cup (C_0 \cap U^0_{D_m})$ et constante, égale à $g^{(i)}_{\alpha,w}$, sur C_i pour $1 \leq i \leq r$. Soit alors $g \in C^\infty(I, \bigoplus_{\alpha\in\Delta\smallsetminus J} R[W^{J\cup\{\alpha\}}])$ l'image de

$$(g_{\alpha,w})_{\alpha,w} \in \bigoplus_{\substack{\alpha\in\Delta\smallsetminus J\\ w\in W^{J\cup\{\alpha\}}(D_m)}} C^\infty({}^wP^0_{J\cup\{\alpha\}}\backslash I, R).$$

Soit $w' \in W^J(D_m)$. Grâce à (17), on a alors

$$(\partial g)_{w'} = \sum_{\alpha\in\Delta\smallsetminus J} \sum_{\substack{w\in W^{J\cup\{\alpha\}}(D_m)\\ w'W_J\subseteq wW_{J\cup\{\alpha\}}}} g^{(i)}_{\alpha,w} = f^{(i)}_{w'} = f_{w'}$$

sur chaque C_i pour $1 \leq i \leq r$; la même égalité est vraie sur $C_0 \cap U^0_{D_m}$, de sorte que l'on a $(\partial g)_{w'} = f_{w'}$ sur tout $U^0_{D_m}$. Pour $x \in \bigcup_{n\leq m} U^0_{D_n} \smallsetminus U^0_{D_m}$, pour tout $\alpha \in \Delta \smallsetminus J$ et tout $w \in W^{J\cup\{\alpha\}}(D_m)$, il existe $p_{\alpha,w} \in {}^wP^0_{J\cup\{\alpha\}}$ et $u_{\alpha,w} \in U'$ vérifiant $x = p_{\alpha,w}u_{\alpha,w}$. On a alors

$$(\partial g)_{w'}(x) = \sum_{\alpha\in\Delta\smallsetminus J} \sum_{\substack{w\in W^{J\cup\{\alpha\}}(D_m)\\ w'W_J\subseteq wW_{J\cup\{\alpha\}}}} g_{\alpha,w}(u_{\alpha,w}) = 0.$$

Il nous reste à vérifier l'existence des fonctions $g_{\alpha,w}$. Soient $\alpha \in \Delta \smallsetminus J$ et $w \in W^{J\cup\{\alpha\}}(D_m)$. Par définition de U', on a

$${}^wP^0_{J\cup\{\alpha\}}U' \cap {}^wP^0_{J\cup\{\alpha\}}U^0_{D_m} = {}^wP^0_{J\cup\{\alpha\}}\left(\bigcup_{n<m} U^0_{D_n}\right) \cap {}^wP^0_{J\cup\{\alpha\}}U^0_{D_m},$$

et par le lemme 5.5, on a

$${}^wP^0_{J\cup\{\alpha\}}\left(\bigcup_{n<m} U^0_{D_n}\right) \cap {}^wP^0_{J\cup\{\alpha\}}U^0_{D_m} = {}^wP^0_{J\cup\{\alpha\}}\left(\left(\bigcup_{n<m} U^0_{D_n}\right) \cap U^0_{D_m}\right).$$

On sait alors que $U' \cap U^0_{D_m}$ est inclus dans $C_0 \cap U^0_{D_m}$: comme $g^{(0)}_{\alpha,w}$ est nul, les conditions imposées sur les valeurs de $g_{\alpha,w}$ sont compatibles. L'espace $U' \cup U^0_{D_m}$ est une union disjointe de $U' \cup (C_0 \cap U^0_{D_m})$ et des C_i pour $1 \leq i \leq r$. Les C_i, pour

---

[11] Remarquons que U' dépend de α et de w.

$1 \le i \le r$, sont ouverts compacts dans $U^0_{D_m}$ et U' est compact. Le complémentaire de U' dans $U' \cup U^0_{D_m}$ contient C_i, donc C_i est ouvert dans $U' \cup U^0_{D_m}$ pour tout $1 \le i \le r$. Comme $\bigcup_{1 \le i \le r} C_i$ est compact, son complémentaire $U' \cup (C_0 \cap U^0_{D_m})$ dans $U' \cup U^0_{D_m}$ est ouvert.

Il existe donc une fonction $g'_{\alpha,w} \in C^\infty(U' \cup U^0_{D_m}, R)$ nulle sur $U' \cup (C_0 \cap U^0_{D_m})$ et constante, égale à $g^{(i)}_{\alpha,w}$, sur C_i pour $1 \le i \le r$. Comme $U' \cup U^0_{D_m}$ est fermé dans $U^0_{D_{\alpha,w}}$, le morphisme de restriction

$$C^\infty(U^0_{D_{\alpha,w}}, R) \to C^\infty(U' \cup U^0_{D_m}, R)$$

est surjectif : il existe une fonction $g_{\alpha,w}$ sur $U^0_{D_{\alpha,w}}$ prolongeant $g'_{\alpha,w}$ comme voulue. La récurrence est donc terminée et la proposition prouvée. □

Pour $J \subseteq \Delta$, on définit la G-représentation de Steinberg généralisée $\mathrm{St}_J\, R$ par la suite exacte

$$\bigoplus_{\alpha \in \Delta \smallsetminus J} C^\infty(P_{J \cup \{\alpha\}} \backslash G, R) \xrightarrow{\partial} C^\infty(P_J \backslash G, R) \to \mathrm{St}_J\, R \to 0. \tag{18}$$

On observe que cela coïncide avec la définition de $\mathrm{St}_Q\, R$ lorsque $Q = P_J$.

Corollaire 5.6. *Le R-module* $\mathrm{St}_J\, R$ *est libre. Et il existe une injection I-équivariante*

$$\iota_R : \mathrm{St}_J\, R \hookrightarrow C^\infty(I, \mathfrak{M}_J(R))$$

dont la formation commute aux changements de base.

Démonstration. La flèche $i \mapsto P_J w^{-1} i$ induit une bijection ensembliste

$${}^w P^0_J \backslash I \xrightarrow{\sim} P_J \backslash P_J w^{-1} I.$$

De plus, toute inclusion $w' W_J \subseteq w W_{J \cup \{\alpha\}}$ donne alors lieu à un diagramme commutatif comme suit (par le lemme 5.1) :

$$\begin{tikzcd}
{}^w P^0_J \backslash I \arrow[r] \arrow[d, "\sim"] & {}^{w'} P^0_{J \cup \{\alpha\}} \backslash I \arrow[d, "\sim"] \\
P_J \backslash P_J w^{-1} I \arrow[r] & P_{J \cup \{\alpha\}} \backslash P_{J \cup \{\alpha\}} (w')^{-1} I
\end{tikzcd}$$

Les w^{-1} pour $w \in W^J$ (resp. $w' \in W^{J \cup \{\alpha\}}$) forment un système de représentants de $W_J \backslash W$ (resp. $W_{J \cup \{\alpha\}} \backslash W$). On a les décompositions d'Iwasawa suivantes (conséquences directes de (1) et du Corollaire 4.2.2 de [Bruhat et Tits 1972]) :

$$P_J \backslash G = \bigsqcup_{w \in W^J} P_J \backslash P_J w^{-1} I, \quad P_{J \cup \{\alpha\}} \backslash G = \bigsqcup_{w' \in W^{J \cup \{\alpha\}}} P_{J \cup \{\alpha\}} \backslash P_{J \cup \{\alpha\}} (w')^{-1} I.$$

On en déduit les sommes directes

$$C^\infty(P_J\backslash G, R) = \bigoplus_{w\in W^J} C^\infty(P_J\backslash P_J w^{-1} I, R),$$

$$C^\infty(P_{J\cup\{\alpha\}}\backslash G, R) = \bigoplus_{w'\in W^{J\cup\{\alpha\}}} C^\infty(P_{J\cup\{\alpha\}}\backslash P_{J\cup\{\alpha\}} (w')^{-1} I, R).$$

Il en découle le diagramme commutatif suivant :

$$\begin{tikzcd}
\bigoplus\limits_{\alpha\in\Delta\smallsetminus J} C^\infty(P_{J\cup\{\alpha\}}\backslash G, R) \arrow[r] \arrow[d, "\sim"] & C^\infty(P_J\backslash G, R) \arrow[r] \arrow[d, "\sim"] & \mathrm{St}_J\, R \arrow[r] \arrow[d, dotted] & 0 \\
\bigoplus\limits_{\substack{\alpha\in\Delta\smallsetminus J\\ w'\in W^{J\cup\{\alpha\}}}} C^\infty({}^{w'}P^0_{J\cup\{\alpha\}}\backslash I, R) \arrow[r] & \bigoplus\limits_{w\in W^J} C^\infty({}^{w}P^0_{J}\backslash I, R) \arrow[r] & C^\infty(I, \mathfrak{M}_J(R))
\end{tikzcd} \tag{19}$$

La première ligne est exacte par définition de $\mathrm{St}_J\, R$ (voir (18)) et la seconde l'est par la proposition 5.2. On en déduit que l'application

$$\iota_R : \mathrm{St}_J\, R \hookrightarrow C^\infty(I, \mathfrak{M}_J(R))$$

est injective ; elle est aussi I-équivariante car toutes les flèches pleines du diagramme le sont.

Étudions d'abord le cas $R = \mathbb{Z}$: parce que l'on a

$$\mathfrak{M}_J(R) = \mathfrak{M}_J(\mathbb{Z}) \otimes_Z R,$$

on a la propriété de changement de base voulue. Grâce à la Proposition 1.3(a) de [Grosse-Klönne 2014] et au lemme A.1, le $\mathbb{Z}$-module $C^\infty(I, \mathfrak{M}_J(\mathbb{Z}))$ est libre. Or un sous-module d'un module libre sur un anneau principal est libre. Ainsi $\mathrm{St}_J\, \mathbb{Z}$ est libre, et $\mathrm{St}_J\, R$ est libre sur R par changement de base.[12] ☐

Avant de déterminer à proprement parler $(\mathrm{St}_J\, R)^I$, on commence par déterminer $C^\infty(P_J\backslash G, R)^I$.

Lemme 5.7. *Le R-module $C^\infty(P_J\backslash G, R)^I$ est libre de rang $|W^J|$.*

Démonstration. La décomposition d'Iwasawa (voir (1)) fournit

$$G = \bigsqcup_{W^J} P_J w^{-1} I.$$

---

[12] On a bien $\mathrm{St}_J\, \mathbb{Z} \otimes_{\mathbb{Z}} R = \mathrm{St}_J\, R$ comme la suite exacte

$$0 \to \sum_{\alpha\in\Delta\smallsetminus J} C^\infty(P_{J\cup\{\alpha\}}\backslash G, \mathbb{Z}) \to C^\infty(P_J\backslash G, \mathbb{Z}) \to \mathrm{St}_J\, \mathbb{Z} \to 0$$

reste exacte après tensorisation par R, les deux $\mathbb{Z}$-modules à gauche étant libres par l'appendice A.

Grâce à cette décomposition, on définit, pour chaque $w \in W^J$, la fonction f_w de $C^\infty(P_J\backslash G, R)^I$ par $f_w(w^{-1}) = 1$ et $f_w(w') = 0$ si $w' \in (W^J)^{-1} \smallsetminus \{w^{-1}\}$. Le R-module $C^\infty(P_J\backslash G, R)^I$ est alors libre et engendré par $(f_w)_{w\in W^J}$. $\square$

Proposition 5.8. *L'espace des I-invariants $(\mathrm{St}_J\, R)^I$ est un R-module libre de rang $|W^J_{\mathrm{pr}}|$.*

De plus, les images $\bar{f}_w$ dans $\mathrm{St}_J\, R$ des fonctions f_w (pour $w \in W^J_{\mathrm{pr}}$) définies dans la preuve du lemme 5.7 forment une base explicite de $(\mathrm{St}_J\, R)^I$.

Démonstration. Appliquons le foncteur des I-invariants au carré commutatif du diagramme (19). On obtient les applications de R-modules :

$$R[W^J] = \bigoplus_{w\in W^J} C^\infty({}^{w}P^0_J\backslash I, R)^I \to (\mathrm{St}_J\, R)^I \to C^\infty(I, \mathfrak{M}_J(R))^I = \mathfrak{M}_J(R). \tag{20}$$

Parce que ι_R est injective et que le foncteur des I-invariants est exact à gauche, la flèche de droite de (20) est injective. Enfin la composée (20) est surjective par définition de $\mathfrak{M}_J(R)$. Alors on a un isomorphisme

$$(\mathrm{St}_J\, R)^I \xrightarrow{\sim} \mathfrak{M}_J(R)$$

de R-modules, ce qui fait de $(\mathrm{St}_J\, R)^I$ un R-module libre de rang $|W^J_{\mathrm{pr}}|$ (par [Grosse-Klönne 2014, Proposition 1.2(a)]).

En prenant les I-invariants du diagramme commutatif (19), on obtient le diagramme suivant, où les carrés du haut sont commutatifs.

$$\begin{tikzcd}
\bigoplus\limits_{\alpha\in\Delta\smallsetminus J} C^\infty(P_{J\cup\{\alpha\}}\backslash G, R)^I \arrow[r] \arrow[d, "\sim"] & C^\infty(P_J\backslash G, R)^I \arrow[r] \arrow[d, "\sim"] & (\mathrm{St}_J\, R)^I \arrow[d, "\sim"] & \\
\bigoplus\limits_{\substack{\alpha\in\Delta\smallsetminus J\\ w'\in W^{J\cup\{\alpha\}}}} C^\infty({}^{w'}P^0_{J\cup\{\alpha\}}\backslash I, R)^I \arrow[r, "\partial"] & \bigoplus\limits_{w\in W^J} C^\infty({}^{w}P^0_J\backslash I, R)^I \arrow[r] & C^\infty(I, \mathfrak{M}_J(R))^I & \\
\bigoplus\limits_{\substack{\alpha\in\Delta\smallsetminus J\\ w'\in W^{J\cup\{\alpha\}}}} R[W^{J\cup\{\alpha\}}] \arrow[r, "\partial"] \arrow[u, "\sim"] & R[W^J] \arrow[r] \arrow[u, "\sim"] & \mathfrak{M}_J(R) \arrow[r] \arrow[u, "\sim"] & 0
\end{tikzcd}$$

On remarque que l'on a utilisé la preuve du premier point pour affirmer l'isomorphisme de R-modules libres que constitue la flèche verticale en haut à droite. Les carrés du bas du diagramme sont commutatifs en vertu de ce que la description des fonctions du lemme 5.7 est compatible avec la discussion d'avant la proposition 5.2. Le résultat suit. $\square$

Une conséquence importante est l'assertion de l'admissibilité dans le théorème 3.1.

Corollaire 5.9. *Supposons que R soit un corps de caractéristique p. Alors $\mathrm{St}_J\, R$ est admissible.*

Démonstration. En regardant le diagramme commutatif (19), on voit que $\mathrm{St}_J\, R$ est l'image de $\bigoplus_{w\in W^J} C^\infty({}^wP_J^0\backslash I, R)$ dans $C^\infty(I, \mathfrak{M}_J(R))$. Comme chaque ${}^wP_J^0$ contient $A\cap I$, cette image est en fait incluse dans $C^\infty(A\cap I\backslash I, \mathfrak{M}_J(R)) \subseteq C^\infty(I, \mathfrak{M}_J(R))$. Si $I(1)$ désigne le pro-p-Sylow de I, on a $(A\cap I)I(1) = I$; il suit les égalités $C^\infty(A\cap I\backslash I, \mathfrak{M}_J(R))^I = C^\infty(A\cap I\backslash I, \mathfrak{M}_J(R))^{I(1)}$, et donc $(\mathrm{St}_J\, R)^I = (\mathrm{St}_J\, R)^{I(1)}$.

Traitons d'abord le cas $R = \overline{\mathbb{F}}_p$: le $\overline{\mathbb{F}}_p$-espace vectoriel $(\mathrm{St}_J\, \overline{\mathbb{F}}_p)^{I(1)} = (\mathrm{St}_J\, \overline{\mathbb{F}}_p)^I$ est de dimension finie par la proposition 5.8. Comme $I(1)$ contient $K(1)$, il est ouvert et on peut appliquer [Paskunas 2004, Theorem 6.3.2], et $\mathrm{St}_J\, \overline{\mathbb{F}}_p$ est admissible.

Dans le cas $R = \mathbb{F}_p$, on remarque que l'on a $(\mathrm{St}_J\, \overline{\mathbb{F}}_p)^H = (\mathrm{St}_J\, \mathbb{F}_p)^H \otimes_{\mathbb{F}_p} \overline{\mathbb{F}}_p$ pour tout sous-groupe ouvert H de G. L'admissibilité de $\mathrm{St}_J\, \overline{\mathbb{F}}_p$ implique alors celle de $\mathrm{St}_J\, \mathbb{F}_p$. Enfin, pour R un corps de caractéristique p, $\mathbb{F}_p$ est naturellement un sous-corps de R. On a alors $(\mathrm{St}_J\, R)^H = (\mathrm{St}_J\, \mathbb{F}_p)^H \otimes_{\mathbb{F}_p} R$ pour tout $H \leq G$ ouvert. Le résultat en découle. □

Remarque. Marie-France Vignéras nous fait remarquer que l'on peut aussi prouver ce fait en se servant du corollaire 5.6. Alors, pour tout sous-groupe compact ouvert H de I, l'espace d'invariants $(\mathrm{St}_J\, R)^H$ s'injecte dans le R-espace vectoriel de dimension finie $C(I/H, \mathfrak{M}_J(R))$. Par ailleurs, un tel argument reste tout à fait valable pour R un anneau principal (sans hypothèse sur la caractéristique).

6. Comparaison avec le cas fini

On retourne à R un anneau commutatif unitaire. On cherche à comparer les représentations de Steinberg généralisées avec leur analogue dans le cas fini.

On remarque d'abord que comme on a choisi K comme étant le parahorique associé à un sommet spécial, le groupe de Weyl associé à l'adhérence schématique $\widetilde{T}$ de $\overline{T}$ est encore W (voir [Bruhat et Tits 1984, Propositions 4.4.5 et 4.6.4 ; Haines et Rapoport 2008, Proposition 12]). Le système de racines $\overline{\Phi}$ associé à la paire $(\overline{B}, \widetilde{T})$ est un sous-système de Φ et il n'est pas nécessairement réduit ; on considère alors son réduit $\overline{\Phi}_{\mathrm{red}}$. Parce que $\overline{\Phi}_{\mathrm{red}}$ et Φ_{red} sont tous deux des sous-systèmes réduits de Φ qui contiennent une $\mathbb{R}$-base de $X^*(T)\otimes_{\mathbb{Z}}\mathbb{R}$, on a une correspondance bijective

$$\Phi_{\mathrm{red}} \xrightarrow{\sim} \overline{\Phi}_{\mathrm{red}}, \qquad \alpha \mapsto \alpha \text{ ou } 2\alpha.$$

Cette bijection exhibe un sous-ensemble de racines simples $\overline{\Delta}$ de $\overline{\Phi}$ en correspondance avec Δ de la même manière :

$$\overline{\Delta} := \bigcup_{\alpha\in\Delta} \{\alpha, 2\alpha\} \cap \overline{\Phi}_{\mathrm{red}}.$$

Pour J un sous-ensemble de Δ, on associe alors de cette manière $\overline{J} \subseteq \overline{\Delta}$.

De même que pour le paragraphe précédent, le lecteur ne s'étonnera pas des invocations du paragraphe 1 de [Grosse-Klönne 2014] pour des résultats concernant $\overline{\Phi}_{\mathrm{red}}$.

Lemme 6.1. *Soit J un sous-ensemble de Δ. La réduction $\overline{P}_J$ de P_J est $P_{\overline{J}}$, le parabolique de $\overline{G}$ associé à $\overline{J}$.*

Remarque. On peut aussi se reporter au Corollaire 4.6.4 de [Bruhat et Tits 1984].

Démonstration. La réduction de P_J est

$$\overline{P}_J = P_J \cap K / P_J \cap K(1)$$

et on affirme qu'elle est engendrée par les images dans $\overline{G}$ de $A \cap K$ et des $U_\alpha \cap K$ pour $\alpha \in \Phi^+_{\mathrm{red}} \cup W_J.J$. En effet, par décomposition de Bruhat, il suffit de voir que $B \cap K$ et $W_J \subseteq K$ sont inclus dans le sous-groupe de K engendré par $A \cap K$ et les $U_\alpha \cap K$ pour $\alpha \in \Phi^+_{\mathrm{red}} \cup W_J.J$. Pour $B \cap K$, cela suit de [Henniart et Vignéras 2015, Theorem 6.5], qui affirme $B \cap K = (A \cap K)(U \cap K)$; et donc $B \cap K$ est engendré par $A \cap K$ et les $U_\alpha \cap K$ pour $\alpha \in \Phi^+_{\mathrm{red}}$. Soient $\beta \in J$ et $u \in U_{-\beta} \cap K$ d'image non nulle dans $\overline{U}_{-\beta}$. Par [Carter 1985, Corollary 2.6.2], il existe alors $\overline{b}_1, \overline{b}_2 \in \overline{B}$ tels que l'on ait $\overline{u} = \overline{b}_1 s \overline{b}_2$ où s est la réflexion dans $W_J \subseteq K$ correspondant à β. En relevant $\overline{b}_1$ et $\overline{b}_2$ respectivement en b_1 et b_2 dans $B \cap K$, on voit que s est bien dans le groupe engendré par $A \cap K$ et les $U_\alpha \cap K$ pour $\alpha \in \Phi^+_{\mathrm{red}} \cup W_J.J$. Les tels s formant un système de générateurs de W_J, on a bien l'affirmation voulue sur $\overline{P}_J$.

Le groupe $P_{\overline{J}}$ est quant à lui engendré par $\overline{A} = Z_{\overline{G}}(\overline{T})$ et les $U_{\overline{\alpha}}$ pour $\overline{\alpha} \in \overline{\Phi}^+ \cup W_{\overline{J}}.\overline{J}$. Ce sont bien là les mêmes groupes radiciels par la remarque suivant le Lemma 6.12 de [Henniart et Vignéras 2015]. □

Pour $\overline{J} \subseteq \overline{\Delta}$ (correspondant à $J \subseteq \Delta$), on garde la même définition de la représentation de Steinberg généralisée $\overline{\mathrm{St}}_{\overline{J}}R$; on rappelle la suite exacte de $\overline{G}$-représentations qui la définit :

$$\bigoplus_{\alpha \in \Delta \smallsetminus J} C(\overline{P}_{J \cup \{\alpha\}} \backslash \overline{G}, R) \to C(\overline{P}_J \backslash \overline{G}, R) \to \overline{\mathrm{St}}_{\overline{J}}R \to 0.$$

Commençons par exhiber des éléments particuliers de $(\overline{\mathrm{St}}_{\overline{J}}R)^{\overline{B}}$: pour w un élément de $W^{\overline{J}}_{\mathrm{pr}}$, notons $\overline{g}_w \in \overline{\mathrm{St}}_{\overline{J}}R$ l'image de la fonction caractéristique g_w de

$$\overline{P}_J w^{-1} \overline{B} = \overline{P}_J w^{-1} (\overline{U} \cap w \overline{U}^- w^{-1})$$

dans $\mathrm{Ind}^{\overline{G}}_{\overline{P}_J}$ id.

Soit $w \in W$. On va commencer par prouver l'affirmation

$$\overline{P}_J w^{-1} \overline{B} = \overline{P}_J w^{-1} (\overline{U} \cap w \overline{U}^- w^{-1}).$$

On a d'abord

$$\overline{B} w^{-1} \overline{B} = \overline{B} w^{-1} \overline{A}\,\overline{U} = \overline{B} w^{-1} \overline{U}. \tag{21}$$

Ensuite, par la Proposition 6.1.6 de [Bruhat et Tits 1972], on a

$$\overline{U} = (\overline{U} \cap w\overline{U}w^{-1})(\overline{U} \cap w\overline{U}^{-}w^{-1}). \tag{22}$$

Il s'ensuit

$$\overline{B}w^{-1}\overline{U} = \overline{B}w^{-1}(\overline{U} \cap w\overline{U}^{-}w^{-1}).$$

La même égalité subsiste alors avec $\overline{P}_J \geq \overline{B}$ à gauche plutôt que $\overline{B}$, et l'égalité voulue est prouvée.

Si on fait subir à $(\overline{G}, \Phi_{\mathrm{red}}, \overline{J})$ le raisonnement du paragraphe précédent, on trouve alors que $(\overline{\mathrm{St}}_{\overline{J}}R)^{\overline{B}}$ est un R-module libre, de base $(\bar{g}_w)_{w \in W^{\overline{J}}_{\mathrm{pr}}}$. Cependant, savoir ce fait n'est pas nécessaire tout de suite, et la preuve de la proposition 6.2 nous le redonnera à moindres frais.

On a une décomposition d'Iwasawa $G = P_J K$ (voir [Haines et Rostami 2010, Corollary 9.1.2]). L'application $P_J\backslash G \to \overline{P}_J\backslash\overline{G}$ est continue (car $P_J K(1)$ est fermé dans G) et surjective. On a donc l'injection naturelle (k désigne ici un représentant de $\bar{k}$ dans K) :

$$C(\overline{P}_J\backslash\overline{G}, R) \to C^{\infty}(P_J\backslash G, R), \quad f \mapsto (P_J k \mapsto f(\overline{P}_J\bar{k})).$$

Dès lors, on a le diagramme suivant de K-représentations.

$$\begin{tikzcd}
\bigoplus\limits_{\alpha \in \Delta \smallsetminus J} C(\overline{P}_{J\cup\{\alpha\}}\backslash\overline{G}, R) \arrow[r, "\varphi_1"] \arrow[d, hook] & C(\overline{P}_J\backslash\overline{G}, R) \arrow[r] \arrow[d, hook] & \overline{\mathrm{St}}_{\overline{J}}R \arrow[r] \arrow[d, dotted, "\iota"] & 0 \\
\bigoplus\limits_{\alpha \in \Delta \smallsetminus J} C^{\infty}(P_{J\cup\{\alpha\}}\backslash G, R) \arrow[r, "\varphi_2"] & C^{\infty}(P_J\backslash G, R) \arrow[r] & \mathrm{St}_J\, R \arrow[r] & 0
\end{tikzcd}$$

Comme les deux flèches verticales de gauche sont des injections et parce que l'on a l'égalité

$$\varphi_2\Bigl(\bigoplus_{\alpha \in \Delta \smallsetminus J} C^{\infty}(P_{J\cup\{\alpha\}}\backslash G, R)\Bigr) \cap C(\overline{P}_J\backslash\overline{G}, R) = \varphi_1\Bigl(\bigoplus_{\alpha \in \Delta \smallsetminus J} C(\overline{P}_{J\cup\{\alpha\}}\backslash\overline{G}, R)\Bigr),$$

la flèche de droite est aussi injective. Dans l'égalité précédente, l'inclusion $\supseteq$ est évidente. Pour obtenir l'inclusion inverse $\subseteq$, soit f une fonction de

$$\varphi_2\Bigl(\bigoplus_{\alpha \in \Delta \smallsetminus J} C^{\infty}(P_{J\cup\{\alpha\}}\backslash G, R)\Bigr) \cap C(\overline{P}_J\backslash\overline{G}, R),$$

que l'on voit comme une fonction sur K. Elle s'écrit $f = \sum_{\alpha \in \Delta \smallsetminus J} \varphi_2(f_\alpha)$ avec $f_\alpha \in C^{\infty}(P_{J\cup\{\alpha\}}\backslash G, R)$ pour tout $\alpha \in \Delta \smallsetminus J$. Fixons, pour tout $\alpha \in \Delta \smallsetminus J$, $\mathcal{K}_\alpha$ un ensemble fini de représentants de $\overline{P}_{J\cup\{\alpha\}}\backslash\overline{G}$ dans $P_{J\cup\{\alpha\}}\backslash G$. On définit alors, pour tout $\alpha \in \Delta \smallsetminus J$, $f'_\alpha \in C(\overline{P}_{J\cup\{\alpha\}}\backslash\overline{G}, R)$ par $f'_\alpha(\bar{k}) = f_\alpha(k)$ pour tout $k \in \mathcal{K}_\alpha$ d'image

$\bar{k} \in \bar{P}_{J\cup\{\alpha\}}\backslash\bar{G}$. Parce que le diagramme ci-dessus est commutatif et que f est dans $C(\bar{P}_J\backslash G, R)$, on obtient

$$\sum_\alpha \varphi_1(f'_\alpha) = \sum_\alpha \varphi_2(f_\alpha) = f$$

et l'inclusion $\subseteq$ voulue.

Proposition 6.2. *L'injection K-équivariante*

$$\iota : \overline{\mathrm{St}}_{\bar{J}}R \hookrightarrow \mathrm{St}_J\, R$$

induit l'isomorphisme de R-modules

$$(\overline{\mathrm{St}}_{\bar{J}}R)^{\bar{B}} \xrightarrow{\sim} (\mathrm{St}_J\, R)^I.$$

Démonstration. L'injection $\iota : \overline{\mathrm{St}}_{\bar{J}}R \hookrightarrow \mathrm{St}_J\, R$ étant K-équivariante, on en déduit l'injection

$$(\overline{\mathrm{St}}_{\bar{J}}R)^{\bar{B}} \hookrightarrow (\mathrm{St}_J\, R)^I,$$

que l'on notera encore ι. Par la proposition 5.8, le R-module $(\mathrm{St}_J\, R)^I$ est libre de rang $|W_{\mathrm{pr}}^J|$.

Commençons par considérer le cas $R = \mathbb{Z}$. En tant que sous-module de $(\mathrm{St}_J\, \mathbb{Z})^I$, $(\overline{\mathrm{St}}_{\bar{J}}\mathbb{Z})^{\bar{B}}$ est un $\mathbb{Z}$-module libre de rang majoré par $|W_{\mathrm{pr}}^{\bar{J}}| = |W_{\mathrm{pr}}^J|$ (on a $|W_J| = |W_{\bar{J}}|$ pour tout $J \subseteq \Delta$). Pour prouver l'isomorphisme voulu, il suffit d'exhiber une base de $(\overline{\mathrm{St}}_{\bar{J}}\mathbb{Z})^{\bar{B}}$ qui s'envoie sur la base $(\bar{f}_w)_{w\in W_{\mathrm{pr}}^J}$ de $(\mathrm{St}_J\, \mathbb{Z})^I$, où $\bar{f}_w$ est l'image de $f_w \in C^\infty(P_J\backslash G, \mathbb{Z})$ (voir proposition 5.8) dans $\mathrm{St}_J\, \mathbb{Z}$. Il reste à remarquer que, par le lemme 6.1, la famille $(\bar{g}_w)_{w\in W_{\mathrm{pr}}^{\bar{J}}}$ s'envoie par ι sur $(\bar{f}_w)_{w\in W_{\mathrm{pr}}^J}$.

Revenons à R général : on a la situation suivante

$$\eta : (\overline{\mathrm{St}}_{\bar{J}}\mathbb{Z})^{\bar{B}} \otimes_{\mathbb{Z}} R \hookrightarrow (\overline{\mathrm{St}}_{\bar{J}}R)^{\bar{B}} \overset{\iota}{\hookrightarrow} (\mathrm{St}_J\, R)^I.$$

Parce que la famille $(\bar{g}_w \otimes 1)_{w\in W_{\mathrm{pr}}^{\bar{J}}}$ d'éléments de $(\overline{\mathrm{St}}_{\bar{J}}\mathbb{Z})^{\bar{B}} \otimes_{\mathbb{Z}} R$ s'envoie par η sur $(\bar{f}_w)_{w\in W_{\mathrm{pr}}^J}$, η est un isomorphisme et[13] ι induit l'isomorphisme

$$(\overline{\mathrm{St}}_{\bar{J}}R)^{\bar{B}} \xrightarrow{\sim} (\mathrm{St}_J\, R)^I. \qquad \square$$

Corollaire 6.3. *La famille $(\bar{g}_w)_{w\in W_{\mathrm{pr}}^{\bar{J}}}$ forme une base du R-module libre $(\overline{\mathrm{St}}_{\bar{J}}R)^{\bar{B}}$.*

7. Représentations de Steinberg généralisées dans le cas fini

Dans toute cette section, G désignera un groupe réductif fini : cela permettra d'alléger les notations et d'éviter de surligner un nombre déraisonnable de symboles. Aussi, quand on fera référence à l'« axiomatique des systèmes de Tits » ou « des BN-paires », on pensera au cadre des « strongly split BN-pairs of characteristic p »

---

[13] Cela garantit au passage que la formation de $\left(\overline{\mathrm{St}}_{\bar{J}}\cdot\right)^{\bar{B}}$ commute au changement de base.

du chapitre I.2 de [Cabanes et Enguehard 2004], ce qui est loisible grâce au paragraphe 5.2 de [Henniart et Vignéras 2015].

Soit $S = \{s_\alpha \mid \alpha \in \Delta\}$. Comme pour les autres éléments de W, on ne fera pas de distinction entre un élément $s \in S$ et son relèvement fixé dans $N_G(T)$.

Supposons à partir de maintenant que R est un anneau commutatif unitaire de caractéristique p. On cherche à comprendre un peu mieux la structure de $(\mathrm{St}_J\ R)^B$, notamment grâce à une structure de $\mathcal{H}_R(G, B)$-module héritée comme suit. Si π est une G-représentation, la réciprocité de Frobenius confère à son espace de B-invariants

$$\pi^B \simeq \mathrm{Hom}_B(\mathrm{id}, \pi) \simeq \mathrm{Hom}_G(\mathrm{Ind}_B^G \mathrm{id}, \pi)$$

une structure de module à droite sur

$$\mathcal{H}_R(G, B) := \mathrm{End}_G(\mathrm{Ind}_B^G \mathrm{id}).$$

A tout $w \in W$, on peut associer l'opérateur de Hecke T_w défini sur π^B par

$$v \mapsto vT_w = \sum_{\gamma \in B\backslash BwB} \gamma^{-1} v = \sum_{u \in (B \cap w^{-1}Bw)\backslash B} u^{-1}w^{-1}v.$$

Pour $w \in W$, on note[14]

$$U_w := U \cap wU^- w^{-1} = \prod_{\substack{\alpha \in \Phi^+_{\mathrm{red}} \\ w^{-1}(\alpha) \in \Phi^-_{\mathrm{red}}}} U_\alpha.$$

Par la première étape de la preuve de la proposition 6.2 (notamment l'équation (22)), pour $s \in S$, on peut voir U_s comme un ensemble de représentants de $(B \cap s^{-1}Bs)\backslash B = (B \cap sBs^{-1})\backslash B$ (car s^2 est un élément de T). Pour $\pi = \mathrm{Ind}_{P_J}^G \mathrm{id}$, $w \in W^J$ et $s \in S$, la formule d'action se réécrit

$$g_w T_s = \sum_{u \in U_s} u^{-1}s^{-1}g_w = \sum_{u \in U_s} \mathrm{id}_{P_J w^{-1} U_w s u}. \tag{23}$$

C'est cette action qu'on va investiguer à travers les deux résultats techniques suivants.

Pour $w \in W$, on note w^J le représentant de wW_J dans W^J (on rappelle que W^J est un système de représentants de W/W_J).

Lemme 7.1. *Soient* $w \in W^J$ *et* $s \in S$.

(a) *Si* $(sw)^J = w$, *alors on a*

$$P_J w^{-1} U_w s u = P_J w^{-1} U_w \quad \text{si } u \in U_s.$$

---

[14] Par [Carter 1985, paragraphe 1.18, Corollary 2.5.17 et discussion suivant Proposition 2.6.3], l'ordre n'a pas d'importance.

(b) *Si* $l((sw)^J) > l(w)$, *alors on a*

$$P_J w^{-1} U_w s U_s = P_J w^{-1} s U_{sw}.$$

(c) *Si* $l((sw)^J) < l(w)$, *alors on a* $s = s_\beta$ *avec* $\beta \in \Delta$ *et* $w^{-1}(\beta) \in \Phi^-_{\text{red}}$. *Posons*

$$U' = \prod_{\substack{\alpha \in \Phi^+_{\text{red}} \smallsetminus \{\beta\} \\ w^{-1}(\alpha) \in \Phi^-_{\text{red}}}} U_\alpha ;$$

c'est un sous-groupe de U_w. *On a*

$$\begin{aligned} P_J w^{-1} U' u s U_s &= P_J w^{-1} U_w \qquad \text{si } u \in U_s \smallsetminus \{1\}, \\ P_J w^{-1} U' s u &= P_J w^{-1} s U_{sw} \quad \text{si } u \in U_s. \end{aligned}$$

De plus, toutes ces relations sont des égalités entre produits directs d'ensembles.

Remarque. Ce sont des raffinements dans l'axiomatique des systèmes de Tits que l'on peut déjà trouver dans [Grosse-Klönne 2014] pour le cas déployé (Lemma 3.1) : on se permet de reproduire sa preuve ici en rajoutant quelques commentaires, notamment pour le fait que les produits d'ensembles sont directs.

Démonstration. Commençons par le fait que les produits sont directs : il s'agit tout d'abord de voir que le produit $P_J w^{-1} U_w$ est direct. Supposons pour cela

$$q_1 w^{-1} u_1 = q_2 w^{-1} u_2 \quad \text{avec } q_1, q_2 \in P_J,\ u_1, u_2 \in U_w.$$

On a alors

$$u_1 u_2^{-1} \in w P_J w^{-1} \cap w U^- w^{-1} \cap U.$$

Regardons à quoi ressemblent les éléments de $w^{-1} U w \cap U^- \cap P_J$. Ils sont inclus dans

$$U^- \cap w^{-1} U w = \prod_{\substack{\alpha \in \Phi^-_{\text{red}} \\ w(\alpha) \in \Phi^+_{\text{red}}}} U_\alpha .$$

De plus, par définition de $w \in W^J$ (voir (2)), pour tout α négativement engendré par J, on a $w(\alpha) \in \Phi^-_{\text{red}}$. Dès lors, comme P_J ne contient que les sous-groupes radiciels associés à la partie quasi-close (au sens du 3.8 de [Borel et Tits 1965]) $\Phi^+_{\text{red}} \cup (\Phi^-_{\text{red}} \cap W_J.J)$, l'intersection $w^{-1} U w \cap U^- \cap P_J$ est réduite à l'élément neutre et le produit $P_J w^{-1} U_w$ est direct.

Une fois que toutes les égalités seront prouvées, le fait que les autres produits sont directs se ramène à chaque fois au cas précédent. Par exemple, regardons le terme de gauche de la première égalité de (c). L'ensemble $P_J w^{-1} U' u s U_s$ est de cardinal majoré par $|P_J| \cdot |U'| \cdot |U_s| = |P_J| \cdot |U_w|$. Or l'égalité avec $P_J w^{-1} U_w$, qui est un produit direct, nous dit que le cardinal de $P_J w^{-1} U' u s U_s$ est exactement $|P_J| \cdot |U_w|$. C'est donc que le produit est aussi direct et cela prouve bien le fait voulu.

Montrons (a). On a d'abord

$$P_J w^{-1} U_w s = P_J w^{-1} B s \subseteq P_J w^{-1} B \cup P_J w^{-1} s B \tag{24}$$

par les équations (21) et (22) et l'axiomatique des BN-paires. Parce que l'on a $(sw)^J = w$, ces deux dernières doubles classes sont égales et leur union vaut simplement

$$P_J w^{-1} B = P_J w^{-1} U_w.$$

L'inclusion (24) devient une égalité puisqu'en rappliquant s on obtient :

$$P_J w^{-1} U_w = P_J w^{-1} U_w s^2 \subseteq P_J w^{-1} U_w s \subseteq P_J w^{-1} U_w.$$

Pour finir, $P_J w^{-1} B$ est bien entendu invariant par translation à droite par U_s.

Attaquons nous à (b). On a dans un premier temps

$$P_J w^{-1} U_w s U_s = P_J w^{-1} B s U_s = P_J w^{-1} B s B.$$

Par définition de P_J, on obtient

$$P_J w^{-1} B s B = \bigcup_{v \in W_J} B v B w^{-1} B s B.$$

L'axiomatique des systèmes de Tits nous dit que $B w^{-1} B s B$ est exactement $B w^{-1} s B$ car on a

$$l(w^{-1} s) = l(sw) > l(w) = l(w^{-1})$$

par [Grosse-Klönne 2014, Lemma 1.3(a)]. On termine alors :

$$P_J w^{-1} B s B = \bigcup_{v \in W_J} B v B w^{-1} s B = P_J w^{-1} s B = P_J w^{-1} s U_{sw}.$$

Enfin pour (c), on remarque d'abord que l'on a

$$(s(sw)^J)^J = w^J = w \quad \text{et} \quad l((sw)^J) < l(w).$$

Par [Grosse-Klönne 2014, Lemma 1.4(c)], on a $l(sw) < l(w)$ et donc $w^{-1}(\beta) \in \Phi^-_{\mathrm{red}}$. On observe aussi

$$s^{-1} U_{sw} s = s^{-1} U s \cap w U^- w^{-1} = \prod_{\substack{s(\alpha) \in \Phi^+_{\mathrm{red}} \\ w^{-1}(\alpha) \in \Phi^-_{\mathrm{red}}}} U_\alpha.$$

Or on a (voir Proposition 1.4 de [Humphreys 1992])

$$s(\Phi^+_{\mathrm{red}}) = (\Phi^+_{\mathrm{red}} \smallsetminus \{\beta\}) \cup \{-\beta\}, \quad w^{-1}(-\beta) \in \Phi^+_{\mathrm{red}}.$$

Il en résulte que la condition sur les indices du produit se réécrit

$$\alpha \in \Phi^+_{\text{red}} \smallsetminus \{\beta\}, \quad w^{-1}(\alpha) \in \Phi^-_{\text{red}};$$

et on en déduit $s^{-1}U_{sw}s = U'$. Cela implique directement $U's = sU_{sw}$ et comme on a

$$u \in U_s \subseteq B, \quad P_J w^{-1} s U_{sw} = P_J w^{-1} s B,$$

la dernière égalité en découle. Pour la première égalité, écrivons l'inclusion (le détail est identique au (b))

$$P_J w^{-1} U_w s \subseteq P_J w^{-1} U_w \cup P_J w^{-1} s U_{sw};$$

cette union est disjointe car on a $swW_J \neq wW_J$. De l'égalité $P_J w^{-1} U's = P_J w^{-1} s U_{sw}$ que l'on vient de prouver, et parce que le produit $P_J w^{-1} U_w$ est direct, on déduit

$$P_J w^{-1}(U_w \smallsetminus U')s \subseteq P_J w^{-1} U_w.$$

Lorsque u est un élément de $U_s \smallsetminus \{1\} = U_\beta \smallsetminus \{1\} \subseteq U_w$, on a l'inclusion $U'u \subseteq U_w \smallsetminus U'$; il s'ensuit

$$P_J w^{-1} U' u s U_s \subseteq P_J w^{-1}(U_w \smallsetminus U') s U_s \subseteq P_J w^{-1} U_w U_s = P_J w^{-1} U_w.$$

Voyons l'inclusion inverse : par [Carter 1985, Corollary 2.6.2], il existe $u_1 \in U_s$ et $b \in U_s T \subseteq B$ tel que l'on ait la décomposition $sus = bsu_1$. On a alors

$$P_J w^{-1} s B s u s U_s = P_J w^{-1} s B s U_s = P_J w^{-1} s B s B.$$

Il s'ensuit que

$$P_J w^{-1} U' = P_J w^{-1} s U_{sw} s = P_J w^{-1} s B s$$

est inclus dans

$$P_J w^{-1} s B s u s U_s = P_J w^{-1} s U_{sw} s u s U_s = P_J w^{-1} U' u s U_s.$$

Au final, on a bien l'égalité voulue (grâce à $U_w = U'U_s$) □

Le cas fini peut se voir de manière similaire au cas p-adique traité dans le lemme 5.7 : le R-module $C(P_J \backslash G, R)$ est libre et une base est donnée par les fonctions g_w pour w parcourant W^J. On tâche d'investiguer sa structure en tant que $\mathcal{H}_R(G, B)$-module à droite lorsque R est de caractéristique p.

Lemme 7.2. *Soient $w \in W^J$ et $s \in S$.*

(a) *Si $(sw)^J = w$, alors on a $g_w T_s = 0$.*

(b) *Si $l((sw)^J) > l(w)$, alors on a $g_w T_s = g_{(sw)^J}$.*

(c) *Si $l((sw)^J) < l(w)$, alors on a $g_w T_s = -g_w$.*

Démonstration. Le (a) est conséquence immédiate de (23), du lemme 7.1(a) et de l'égalité $|U_s| = 0$ dans R de caractéristique p (voir [Carter 1985, page 74]). Le (b) suit de (23) et de la décomposition en produit direct du lemme 7.1(b) aussi.

Intéressons nous au (c) : on utilise la décomposition en produit direct $U_w = U'U_s$. On a alors

$$g_w T_s = \sum_{u \in U_s} \sum_{u' \in U_s} \mathrm{id}_{P_J w^{-1} U' u' s u} = \sum_{u \in U_s} \mathrm{id}_{P_J w^{-1} U' s u} + \sum_{u \in U_s} \sum_{u' \neq 1} \mathrm{id}_{P_J w^{-1} U' u' s u} .$$

Par le lemme 7.1(c) et $|U_s| = 0$ dans R, le premier terme vaut 0 et le second

$$\sum_{u' \in U_s \smallsetminus \{1\}} \mathrm{id}_{P_J w^{-1} U_w} = -g_w .$$

Le lemme en découle. □

Les actions précédemment étudiées dans $(\mathrm{Ind}_B^G \mathrm{id})^B$ sont compatibles à celles du quotient $(\mathrm{St}_J\, R)^B$.

Proposition 7.3. *Supposons R de caractéristique p. Tout sous-$\mathcal{H}_R(G, B)$-module non nul de $(\mathrm{St}_J\, R)^B$ contient l'élément $\bar{g}_{z^J}$ de $\mathrm{St}_J\, R$, où z^J désigne l'élément de longueur maximale*[15] *de W^J.*

Démonstration. Soit E un sous-$\mathcal{H}_R(G, B)$-module non nul de $(\mathrm{St}_J\, R)^B$. Par la proposition 6.2, E contient un élément non nul

$$h = \sum_{w \in W_{\mathrm{pr}}^J} \alpha_w(h) \bar{g}_w \quad \text{avec } \alpha_w(h) \in R.$$

On fixe une énumération $w_1, w_2, \ldots$ des éléments de W_{pr}^J vérifiant $w_j <_J w_i \Rightarrow i < j$: en particulier, on a $w_1 = z^J$. On veut montrer qu'il existe $h \in E$ non nul tel que

$$t(h) := \min\{i \geq 1 \mid \forall j > i,\ \alpha_{w_j}(h) = 0\}$$

soit égal à 1, c'est-à-dire $\bar{g}_{z^J} \in E$. Supposons le contraire et donc on a

$$t := \min\{t(h) \mid h \in E \smallsetminus \{0\}\} \geq 2.$$

Par [Grosse-Klönne 2014, Lemma 1.5], il existe $w' \in W_{\mathrm{pr}}^J$ et $s \in S$ tel que $w_t <_J w'$, $l((sw_t)^J) < l(w_t)$ et $l(w') \leq l((sw')^J)$. Par définition de $<_J$ (voir appendice B), il existe $s_1, \ldots, s_r$ dans S tels que $w^{(i)} = (s_i \cdots s_1 w_t)^J$ vérifie $l(w^{(i)}) > l(w^{(i-1)})$ pour tout $1 \leq i \leq r$ et $w^{(r)} = w'$.

Soit $h \in E \smallsetminus \{0\}$ avec $t(h) = t$. Commençons par remarquer que, grâce au lemme 7.2, on a $\alpha_{w^{(r)}}(hT_s) = 0$. On peut donc considérer $h \in E \smallsetminus \{0\}$ avec $t(h) = t$

---

[15] Un élément de longueur maximale est aussi maximal pour $<_J$ (voir l'appendice B pour la définition de $<_J$) par [Grosse-Klönne 2014, Lemma 1.4(d)]; l'existence d'un unique élément $<_J$-maximal est ensuite donnée par [Grosse-Klönne 2014, Lemma 1.4(e)].

et tel que $k(h) \geq 0$ est minimal, égal à k, où $k(h)$ est l'entier minimal de $[0, r]$ défini par $\alpha_{w^{(k(h))}}(h) = 0$. Si on a $k(h) = 0$, alors $\alpha_{w_t(h)}$ est nul, ce qui est contradictoire. On suppose ainsi $k > 0$, et on va le faire diminuer : considérons $h' = hT_{s_k}$, et observons

$$\alpha_{w^{(k-1)}}(h') = -\alpha_{w^{(k)}}(h) = 0, \quad \alpha_{w^{(k)}}(h') = \alpha_{w^{(k-1)}}(h) \neq 0.$$

Ceci nous assure $h' \neq 0$ et $k(h') < k$. Cela contredit la minimalité de k et donc l'hypothèse initiale : on vient donc de montrer l'existence de $h \in E \smallsetminus \{0\}$ avec $t(h) = 1$. C'est le résultat voulu. ☐

8. Paramètres de Hecke–Satake et preuve du théorème 3.1

Soit R un corps de caractéristique p. On commence par étudier le K-socle d'une représentation de Steinberg généralisée.

Proposition 8.1. *Soit $J \subseteq \Delta$. Le K-socle de la Steinberg généralisée* $\mathrm{St}_J\, R$ *est irréductible.*

Démonstration. Soit V une sous-K-représentation irréductible de $\mathrm{St}_J\, R$. L'injection de la proposition 6.2 nous permet de voir $V^{U \cap K}$ comme un sous-espace de

$$(\mathrm{St}_J\, R)^{I(1)} = (\mathrm{St}_J\, R)^{I} \simeq (\overline{\mathrm{St}}_{\bar{J}} R)^{\bar{B}}$$

(voir début de preuve du corollaire 5.9) ; de cette manière, $V^{U \cap K} = V^{B \cap K}$ est une droite stable par l'action de $\mathcal{H}_R(\overline{G}, \overline{B})$. Or, par la proposition 7.3, tout sous-$\mathcal{H}_R(\overline{G}, \overline{B})$-module non nul de $(\overline{\mathrm{St}}_{\bar{J}} R)^{\bar{B}}$ contient l'image $\bar{g}_J$ de l'élément

$$g_J := \mathrm{id}_{\overline{P}_J z^J \overline{B}} \in \mathrm{Ind}_{\overline{P}_J}^{\overline{G}}\ \mathrm{id}$$

dans $\overline{\mathrm{St}}_{\bar{J}} R$. Ainsi V est généré par $\bar{g}_J$ en tant que K-représentation, et le K-socle de $\mathrm{St}_J\, R$ est irréductible. ☐

Soit V une K-représentation irréductible. On va noter $\mathcal{H}_R(G, K, V)$ l'algèbre de Hecke–Satake $\mathrm{End}_G(\mathrm{ind}_K^G V)$. Les éléments de $\mathcal{H}_R(G, K, V)$ sont des opérateurs à support fini parmi les doubles classes $K \backslash G / K$, donc on va fixer un système de représentants dominants dans A (au sens qu'ils contractent B) que l'on notera Σ_+.

La transformée de Satake (voir [Henniart et Vignéras 2015, paragraphe 7.3]) est un isomorphisme de R-algèbres

$$\mathcal{S} : \mathcal{H}_R(G, K, V) \xrightarrow{\sim} \mathcal{H}_R^+(A, A \cap K, V_{U \cap K}) \tag{25}$$

où $\mathcal{H}_R^+(A, A \cap K, V_{U \cap K})$ désigne la sous-algèbre de $\mathcal{H}_R(A, A \cap K, V_{U \cap K})$ engendrée par les opérateurs portés par la classe $z(A \cap K)$ pour $z \in \Sigma_+$ quand ils existent.

Soit J un sous-ensemble de Δ. On va particulièrement s'intéresser au cas où V est V_J, l'unique K-représentation irréductible M_J-corégulière[16] telle que l'action de M_J sur la droite $(V_J)_{N_J \cap K}$ est triviale (par la Proposition 3.11 de [Henniart et Vignéras 2012]). Dans ce cas-là, en particulier, $V_{U\cap K}$ est la représentation triviale id de $A \cap K$ et toute classe $z(A\cap K)$ pour $z \in \Sigma_+$ porte un unique opérateur de Hecke τ_z de $\mathscr{H}_R^+(A, A\cap K, V_{U\cap K})$, envoyant z sur $\mathrm{id}_{V_{U\cap K}}$ (voir [Henniart et Vignéras 2015, paragraphe 7.3]) ; et ils en constituent une base en tant que R-espace vectoriel. De plus, l'algèbre $\mathscr{H}_R^+(A, A\cap K, \mathrm{id})$ est alors commutative et de type fini sur R (voir paragraphe 7.2 de [Henniart et Vignéras 2015]). On en déduit alors que $\mathscr{H}_R(G, K, V_J)$ est aussi commutative et de type fini par l'isomorphisme de Satake (25).

Soit π une G-représentation admissible à coefficients dans R. Alors, comme $K(1)$ est un pro-p-groupe ouvert, le sous-espace des $K(1)$-invariants $\pi^{K(1)}$ est non nul et de dimension finie sur R. Il possède donc une sous-K-représentation irréductible V. En particulier, $\mathrm{Hom}_K(V, \pi)$ est non trivial et de dimension finie. De plus, c'est un module à droite sur l'algèbre $\mathscr{H}_R(G, K, V)$.

On suppose à présent que V est un V_J pour un certain sous-ensemble J de Δ, et que R est algébriquement clos. Alors $\mathscr{H}_R(G, K, V_J)$ étant commutative, elle possède un sous-espace propre associé à un caractère $\chi : \mathscr{H}_R(G, K, V_J) \to R$ dans son action sur $\mathrm{Hom}_K(V_J, \pi)$. Dit autrement, on a un morphisme non nul de G-représentations

$$\mathrm{ind}_K^G V_J \otimes_{\mathscr{H},\chi} R \to \pi.$$

On tâche de déterminer les caractéristiques d'un tel caractère χ lorsque π est $\mathrm{St}_J R$. Pour être précis, notons $\chi^{(A)}$ la composée

$$\mathscr{H}_R^+(A, A\cap K, (V_J)_{U\cap K}) \xrightarrow[\sim]{\mathscr{S}^{-1}} \mathscr{H}_R(G, K, V_J) \xrightarrow{\chi} R;$$

c'est $\chi^{(A)}$ que l'on va expliciter.

Proposition 8.2. *Soit R un corps algébriquement clos de caractéristique p. Soit J un sous-ensemble de Δ.*

(i) *Le K-socle de* $\mathrm{St}_J R$ *s'identifie à V_J.*

(ii) *Il existe un morphisme non nul de G-représentations*

$$\mathrm{ind}_K^G V_J \otimes_{\mathscr{H},\chi} R \to \mathrm{St}_J R$$

si et seulement si $\chi^{(A)}$ est le caractère envoyant τ_z sur 1 *pour tout $z \in \Sigma_+$.*

Remarque. En particulier, $\mathrm{St}_J R$ n'est pas supersingulier au sens de [Henniart et Vignéras 2012].

---

[16] Cela veut dire que le stabilisateur dans K de la représentation associée à $V_J^{U^- \cap K}$ est inclus dans $(P_J^- \cap K)K(1)$.

Démonstration. On note $\mathscr{H} = \mathscr{H}_R(G, K, V_J)$ et $\mathscr{H}_M = \mathscr{H}_R(M_J, M_J \cap K, \mathrm{id})$. Soient $\chi : \mathscr{H} \to R$ et $\chi_M : \mathscr{H}_M \to R$ les caractères d'algèbres associés à $\chi^{(A)}$, caractère envoyant tout τ_z sur 1 (pour $z \in \Sigma_+$). Par [Henniart et Vignéras 2012, Theorem 1.2], parce que V_J est M_J-corégulière et que $(V_J)_{N_J \cap K}$ est la $(M_J \cap K)$-représentation triviale, on a un morphisme surjectif de G-représentations

$$\operatorname{ind}_K^G V_J \otimes_\chi R \xrightarrow{\sim} \operatorname{Ind}_{P_J}^G (\operatorname{ind}_{M_J \cap K}^{M_J} \mathrm{id} \otimes_{\chi_M} R) \twoheadrightarrow \operatorname{Ind}_{P_J}^G \mathrm{id}\,. \tag{26}$$

On en déduit [17] l'existence d'un morphisme surjectif (en particulier non nul) de G-représentations $\operatorname{ind}_K^G V_J \otimes_\chi R \to \mathrm{St}_J R$. Il s'ensuit que V_J génère $\mathrm{St}_J R$ en tant que G-représentation. Parce que $\mathrm{St}_J R$ est de K-socle irréductible, V_J est l'unique K-représentation irréductible contenue dans $\mathrm{St}_J R$. Cela prouve (i) et le sens indirect de (ii).

Prouvons maintenant la seconde implication de (ii). Comme $K \cap P_J \backslash K \to P_J \backslash G$ est un homéomorphisme, on a par réciprocité de Frobenius

$$\mathrm{Hom}_K(V_J, \operatorname{Ind}_{P_J}^G \mathrm{id}) = \mathrm{Hom}_{K \cap P_J}(V_J, \mathrm{id})$$

et donc, puisque l'on a $(V_J)_{N_J \cap K} = \mathrm{id}$, on déduit

$$\mathrm{Hom}_K(V_J, \operatorname{Ind}_{P_J}^G \mathrm{id}) = \mathrm{Hom}_{K \cap M_J}(\mathrm{id}, \mathrm{id}).$$

Ce dernier espace est donc un R-espace vectoriel de dimension 1. Le morphisme surjectif $\operatorname{Ind}_{P_J}^G \mathrm{id} \twoheadrightarrow \mathrm{St}_J R$ de G-représentations induit le morphisme

$$\psi : \mathrm{Hom}_K(V_J, \operatorname{Ind}_{P_J}^G \mathrm{id}) \to \mathrm{Hom}_K(V_J, \mathrm{St}_J R)$$

de R-espaces vectoriels. On vient de voir que $\mathrm{Hom}_K(V_J, \operatorname{Ind}_{P_J}^G \mathrm{id})$ est de dimension 1, et $\mathrm{Hom}_K(V_J, \mathrm{St}_J R)$ est aussi de dimension 1 par (i) et lemme de Schur. De ce fait, ψ est un isomorphisme si et seulement si il est non nul. Or le fait que (26) induise $\operatorname{ind}_K^G V_J \otimes_\chi R \twoheadrightarrow \mathrm{St}_J R$ implique la non nullité de ψ : ψ est un isomorphisme.

Après application de la réciprocité de Frobenius, on a un isomorphisme

$$\mathrm{Hom}_G(\operatorname{ind}_K^G V_J, \operatorname{Ind}_{P_J}^G \mathrm{id}) \xrightarrow{\sim} \mathrm{Hom}_G(\operatorname{ind}_K^G V_J, \mathrm{St}_J R)$$

de R-espaces vectoriels. Ce dernier est $\mathscr{H}_R(G, K, V_J)$-équivariant car ψ est juste induit par la projection définissant $\mathrm{St}_J R$. De ce fait, toute flèche non nulle $\operatorname{ind}_K^G V_J \otimes_{\mathscr{H},\chi} R \to \mathrm{St}_J R$ pour un certain χ se factorise à travers $\operatorname{ind}_K^G V_J \otimes_{\mathscr{H},\chi} R \to \operatorname{Ind}_{P_J}^G \mathrm{id}$. Et

---

[17] Jusqu'à présent, hormis dans l'introduction, on a défini $\mathrm{St}_J R$ à partir des $C^\infty(P_J \backslash G, R)$. Remarquons que la définition peut se faire de manière équivalente à partir d'induites paraboliques : l'application R-linéaire $C^\infty(P_J \backslash G, R) \to \operatorname{Ind}_{P_J}^G$ id est un isomorphisme car la projection canonique $G \to P_J \backslash G$ possède une section continue.

par le diagramme (4) de [Henniart et Vignéras 2012], l'isomorphisme de réciprocité de Frobenius

$$\mathrm{Hom}_G(\mathrm{ind}_K^G V_J, \mathrm{Ind}_{P_J}^G \mathrm{id}) \xrightarrow{\sim} \mathrm{Hom}_{M_J}(\mathrm{ind}_{M_J\cap K}^{M_J} \mathrm{id}, \mathrm{id})$$

est $\mathcal{H}_R(G, K, V_J)$-équivariant, où l'action au but se fait à travers la transformée de Satake partielle

$$\mathcal{S}_G^{M_J} : \mathcal{H}_R(G, K, V_J) \hookrightarrow \mathcal{H}_R(M_J, M_J \cap K, \mathrm{id}).$$

Il s'agit donc de déterminer les valeurs propres de Hecke possibles pour l'action de $\mathcal{H}_R(M_J, M_J \cap K, \mathrm{id})$ sur $\mathrm{Hom}_{M_J}(\mathrm{ind}_{M_J\cap K}^{M_J} \mathrm{id}, \mathrm{id})$. Cet espace est unidimensionnel puisqu'il est isomorphe à

$$\mathrm{Hom}_{M_J\cap K}(\mathrm{id}, \mathrm{id}) = \mathrm{Hom}_{A\cap K}(\mathrm{id}, \mathrm{id}) = \mathrm{Hom}_A(\mathrm{ind}_{A\cap K}^A \mathrm{id}, \mathrm{id}),$$

et c'est à travers ce dernier que l'action de $\mathcal{H}_R(A, A \cap K, \mathrm{id})$ se lit naturellement. On conclut donc qu'elle se fait à travers le caractère $\chi^{(A)}$ envoyant chaque τ_z sur 1, d'où l'implication qu'il restait à prouver pour (ii). □

On exhibe de la preuve précédente le fait plus précis suivant, qui va facilement impliquer le théorème 3.1.

Corollaire 8.3. *Soient R un corps algébriquement clos de caractéristique p et J un sous-ensemble de Δ. Le K-socle V_J de $\mathrm{St}_J R$ l'engendre en tant que G-représentation.*

On a alors l'irréductibilité de $\mathrm{St}_J R$ dans le cas R algébriquement clos de caractéristique p comme suit. Soit $\pi \subseteq \mathrm{St}_J R$ une sous-représentation non nulle. Alors π contient une sous-K-représentation irréductible qui, par l'argument précédent, est donc V_J. Mais on sait que V_J génère $\mathrm{St}_J R$; donc on a $\pi = \mathrm{St}_J R$ et l'irréductibilité voulue.

Le théorème 3.1 en découle par le corollaire 5.9 et le fait tout à fait général que, si une représentation définie sur R est irréductible sur une clôture algébrique R^{alg}, alors elle l'est aussi sur R.

9. Induites paraboliques de Steinberg généralisées

Commençons par un mot sur la preuve du corollaire 3.2, notamment sur le fait que deux $J, J' \subseteq \Delta$ distincts engendrent des $\mathrm{St}_J R$ et $\mathrm{St}_{J'} R$ non isomorphes. Cela suit immédiatement de ce que l'on vient de faire puisque l'on a alors $V_J \neq V_{J'}$.

Définissons la filtration suivante sur $\mathrm{Ind}_{P_J}^G \mathrm{id}$:

$$\mathrm{Fil}^i = \begin{cases} \sum_{J'\supseteq J,\, |J'\smallsetminus J|\geq i} \mathrm{Ind}_{P_{J'}}^G \mathrm{id} & \text{pour } 0 \leq i \leq |\Delta \smallsetminus J|, \\ 0 & \text{pour } i > |\Delta \smallsetminus J|. \end{cases}$$

Et montrons que c'est la filtration par les cosocles de $\mathrm{Ind}_{P_J}^G$ id, c'est-à-dire la filtration descendante définie par $\mathrm{Fil}^0 = \mathrm{Ind}_{P_J}^G$ id et Fil^i est telle que $\mathrm{gr}^{i-1} := \mathrm{Fil}^{i-1} / \mathrm{Fil}^i$ est le cosocle de Fil^{i-1} pour $i \geq 1$. On a le diagramme commutatif suivant, pour $i \geq 0$ et $J' \supseteq J$ avec $|J' \smallsetminus J| = i$ (en particulier $i \leq |\Delta \smallsetminus J|$) :

$$\begin{tikzcd}
 & \sum\limits_{J'' \supsetneq J'} \mathrm{Ind}_{P_{J''}}^G \mathrm{id} \arrow[r] \arrow[d, hook] & \mathrm{Ind}_{P_{J'}}^G \arrow[r] \arrow[d, hook] & \mathrm{St}_{J'} R \arrow[r] \arrow[d, dotted] & 0 \\
0 \arrow[r] & \mathrm{Fil}^{i+1} \arrow[r] & \mathrm{Fil}^i \arrow[r] & \mathrm{gr}^i \arrow[r] & 0
\end{tikzcd}$$

Si la flèche $\mathrm{St}_{J'} R \to \mathrm{gr}^i$ était triviale, cela voudrait dire que l'image de $\mathrm{Ind}_{P_{J'}}^G \mathrm{id} \hookrightarrow \mathrm{Fil}^i$ serait incluse dans Fil^{i+1}, ce qui serait absurde par le lemme 9.1 ultérieur. De ce fait, et parce que l'on a $\mathrm{St}_{J_1} R \not\simeq \mathrm{St}_{J_2} R$ pour $J_1 \neq J_2$, on a une injection

$$\bigoplus_{\substack{J' \supseteq J \\ |J' \smallsetminus J| = i}} \mathrm{St}_{J'} R \hookrightarrow \mathrm{gr}^i .$$

Pour voir que c'est bien le cosocle de Fil^i, comme on connaît les composantes de Jordan–Hölder de $\mathrm{Ind}_{P_J}^G$ id (ce sont les $\mathrm{St}_{J'} R$ pour $J' \supseteq J$, avec[18] multiplicité 1), il suffit de voir que toute flèche $\mathrm{Ind}_{P_{J'}}^G \mathrm{id} \to \mathrm{St}_{J''} R$ est triviale dès que l'on a $J'' \supsetneq J'$ ou $J'' \not\supseteq J'$. Le second cas est clair et on veut montrer que tout morphisme $f : \mathrm{Ind}_{P_{J'}}^G \mathrm{id} \to \mathrm{St}_{J''} R$ de G-représentations est trivial si $J'' \supsetneq J'$. Si ce n'était pas le cas, f serait surjectif par irréductibilité de $\mathrm{St}_{J''} R$, de noyau contenant $\mathrm{St}_{J'} R$. Mais comme $V_{J'} \subseteq (\mathrm{St}_{J'} R)|_K$ génère $\mathrm{Ind}_{P_{J'}}^G$ id par (26), toute injection $\mathrm{St}_{J'} R \hookrightarrow \mathrm{Ind}_{P_{J'}}^G$ id se doit d'être un isomorphisme. C'est absurde, donc f est nul et $\mathrm{St}_{J'} R$ est bien le plus gros quotient semi-simple de $\mathrm{Ind}_{P_{J'}}^G$ id. Il en résulte que $(\mathrm{Fil}^i)_i$ est bien la filtration par les cosocles comme annoncé.

Lemme 9.1. *Soient $J \subseteq \Delta$ un ensemble, $i \in [0, |\Delta \smallsetminus J|]$ un entier et $J' \supseteq J$ un sous-ensemble de Δ avec $|J' \smallsetminus J| = i$. Alors l'image de* $\mathrm{Ind}_{P_{J'}}^G \mathrm{id} \hookrightarrow \mathrm{Fil}^i$ *n'est pas incluse dans* Fil^{i+1}.

Démonstration. Le cas $i = 0$ résulte de ce que $\mathrm{St}_J R$ est non triviale. On suppose donc à présent $i \geq 1$.

---

[18] Ce sont les telles représentations de Steinberg comme on peut le voir par récurrence descendante sur $|J|$ à partir de la définition des $\mathrm{St}_{J'} R$. La décomposition $G = \bigsqcup_{w \in W^J} P_J w^{-1} B$ fait que les restrictions respectives de $C^\infty(P_J \backslash G, R)$ et $\mathrm{St}_J R$ à B possèdent des filtrations telles que $\bigoplus_i \mathrm{gr}^i$ sont respectivement égales à

$$\bigoplus_{w \in W^J} C^\infty(P_J \backslash P_J w^{-1} B, R) \quad \text{et} \quad \bigoplus_{w \in W^J_{\mathrm{pr}}} C^\infty(P_J \backslash P_J w^{-1} B, R).$$

Comme on a de plus $W^J = \bigsqcup_{J' \supseteq J} W_{\mathrm{pr}}^{J'}$, les $\mathrm{St}_{J'} R$ apparaissent avec multiplicité 1 comme voulu.

Supposons par l'absurde que $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id}$ est incluse dans Fil^{i+j} pour un $j \geq 1$ maximal, c'est-à-dire avec $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \nsubseteq \mathrm{Fil}^{i+j+1}$ (un tel j existe bien puisque la filtration devient nulle au bout d'un certain rang). On commence par établir

$$\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} = (\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1}) + \sum_{J'' \supsetneq J'} \mathrm{Ind}_{P_{J''}}^{G}\,\mathrm{id}\,. \tag{27}$$

Soit f un élément de

$$\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \Big/ \Big((\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1}) + \sum_{J'' \supsetneq J'} \mathrm{Ind}_{P_{J''}}^{G}\,\mathrm{id} \Big),$$

que l'on voit dans

$$\mathcal{F} = \mathrm{Fil}^{i+j} \Big/ \Big((\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1}) + \sum_{J'' \supsetneq J'} \mathrm{Ind}_{P_{J''}}^{G}\,\mathrm{id} \Big).$$

Alors f s'écrit $\sum_{J''} f_{J''}$ où les J'' parcourent les $J'' \supseteq J$ avec $|J'' \smallsetminus J| \geq i+j$, $J'' \nsupseteq J'$ et $f_{J''}$ appartient à l'image de $\mathrm{Ind}_{P_{J''}}^{G}\,\mathrm{id}$ dans $\mathcal{F}$. On prend ensuite un $\alpha \in J' \smallsetminus J$ (cet ensemble est non vide car i est non nul) et $s \in W_{J'}$ la réflexion correspondante. Parce que l'on a $J' \nsubseteq J''$, et que l'on a quotienté par $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1}$, l'écriture $f = \sum f_{J''}$ est en fait invariante à gauche par s. En effectuant de même pour toute racine de $J' \smallsetminus J$, on voit que chaque $f_{J''}$ est en fait nulle dans $\mathcal{F}$ (car $j \geq 1$), ce qui donne la nullité de f. Et donc (27) comme annoncé : mais cela implique que l'on a une surjection naturelle $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1} \twoheadrightarrow \mathrm{St}_{J'}\,R$. De ce fait, ou bien il existe un $J'' \supsetneq J'$ avec $\mathrm{Ind}_{P_{J''}}^{G}\,\mathrm{id} \nsubseteq \mathrm{Fil}^{i+j+1}$, ce qui est exclu par l'inclusion $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \subseteq \mathrm{Fil}^{i+j}$. Ou bien $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id} \cap \mathrm{Fil}^{i+j+1}$ est tout $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id}$, c'est-à-dire que $\mathrm{Ind}_{P_{J'}}^{G}\,\mathrm{id}$ est inclus dans Fil^{i+j+1}, ce qui contredit la maximalité de $j \geq 1$. Toutes les possibilités amènent à des contradictions : le lemme est prouvé. ☐

Corollaire 9.2. *Soient $J' \subseteq J$ des sous-ensembles de Δ. Alors la représentation*[19] $\mathrm{Ind}_{P_J}^{G}(\mathrm{St}_{J'}^{M_J}\,R)$ *est de longueur finie, de constituants de Jordan–Hölder les* $\mathrm{St}_{J''}\,R$ *avec $J'' \subseteq \Delta$ vérifiant $J \cap J'' = J'$, chacun apparaissant avec multiplicité* 1.

Démonstration. Commençons tout d'abord par remarquer que, puisque l'on a $J' \subseteq J$, la condition $J \cap J'' = J'$ est équivalente à $J'' \supseteq J'$ et $J \smallsetminus J' \subseteq \Delta \smallsetminus J''$. C'est sous cette seconde forme que nous allons l'utiliser au cours du raisonnement qui suit.

Prouvons-le par récurrence descendante sur $J' \subseteq J$. L'étape d'initiation $J' = J$ est juste le corollaire 3.2. Soit $J' \neq J$ et supposons le résultat vrai pour tout parabolique $J_0 \subseteq J$ contenant strictement J'. Par définition de la représentation de Steinberg généralisée, on a une suite exacte de M_J-représentations

$$0 \to \mathrm{Ker} \to \mathrm{Ind}_{P_{J'}}^{M_J}\,\mathrm{id} \to \mathrm{St}_{J'}^{M_J}\,R \to 0, \tag{28}$$

---

[19] Où on note $\mathrm{St}_{J'}^{M_J}\,R$ une représentation de Steinberg généralisée du groupe réductif M_J.

où Ker est par là-même définie et a constituants de Jordan–Hölder les St_{J_0} id pour $J' \subsetneq J_0 \subseteq J$ par le corollaire 3.2. Appliquons le foncteur exact $\mathrm{Ind}_{P_J}^G$ (voir [Vignéras 2012a, Proposition 1.1]) à (28) pour obtenir :

$$0 \to \mathrm{Ind}_{P_J}^G(\mathrm{Ker}) \to \mathrm{Ind}_{P_J}^G(\mathrm{Ind}_{P_{J'}}^{M_J} \mathrm{id}) \to \mathrm{Ind}_{P_J}^G(\mathrm{St}_{J'}^{M_J} R) \to 0.$$

Comme on a

$$M_J/(M_J \cap P_{J'}) \xrightarrow{\sim} M_J B/(M_J \cap P_{J'})B = P_J/(M_J \cap P_{J'})B = P_J/P_{J'}$$

par décomposition de Levi, le terme central est $\mathrm{Ind}_{P_{J'}}^G$ id, de constituants de Jordan–Hölder les $\mathrm{St}_{J''} R$ avec $J'' \supseteq J'$ par le corollaire 3.2. Par hypothèse de récurrence, les constituants de $\mathrm{Ind}_{P_J}^G(\mathrm{Ker})$ sont les $\mathrm{St}_{J''} R$ avec $J'' \subseteq \Delta$ vérifiant $J'' \supseteq J_0$ et $J \smallsetminus J_0 \subseteq \Delta \smallsetminus J''$ pour un certain $J' \subsetneq J_0 \subseteq J$.

Mais alors, soit J'' tel que $\mathrm{St}_{J''} R$ est un constituant de Jordan–Hölder de $\mathrm{Ind}_{P_J}^G(\mathrm{St}_{J'}^{M_J} R)$. On a $J'' \supsetneq J'$, et aussi, $J \smallsetminus J' \subseteq \Delta \smallsetminus J''$. En effet, supposons cette seconde inclusion fausse, c'est-à-dire $J'' \cap (J \smallsetminus J') \neq \varnothing$: on considère $J_0 \subseteq \Delta$ avec

$$J_0 := J' \cup \big(J'' \cap (J \smallsetminus J')\big) \supsetneq J'$$

et on a $J_0 \subseteq J''$, $J \smallsetminus J_0 \subseteq \Delta \smallsetminus J''$, donc $\mathrm{St}_{J''} R$ apparaît déjà dans $\mathrm{Ind}_{P_J}^G(\mathrm{Ker})$ par l'assertion de multiplicité 1 dans le corollaire 3.2. C'est absurde. Enfin, la décomposition disjointe (qu'il est plus facile de voir avec la condition équivalente $J \cap J'' = J'$ dans le terme de droite)

$$\{J'' \supseteq J'\} = \bigsqcup_{J_0 \supseteq J'} \{J'' \supseteq J_0 \mid J \smallsetminus J_0 \subseteq \Delta \smallsetminus J''\}$$

nous permet de dire que ce sont les seuls constituants qui interviennent. La récurrence est terminée. □

Appendice A: De la liberté de $C^\infty(X, \mathbb{Z})$

Parce qu'on se sert constamment du fait suivant, on se permet de rappeler sa preuve probablement bien connue.

Pour X un espace topologique, on note $\mathcal{U}$ la famille des recouvrements de X par un nombre fini d'ouverts disjoints. On remarque que $\mathcal{U}$ un ensemble ordonné par $U \leq V$ si pour tout $A \in U$ il existe des $B_i \in V$ vérifiant $A = \bigcup B_i$.

Lemme A.1. *Soit X un espace topologique compact et totalement discontinu. Supposons que $\mathcal{U}$ possède un sous-ensemble $\mathcal{V}$ dénombrable et cofinal. Alors l'espace $C^\infty(X, \mathbb{Z})$ des fonctions localement constantes sur X à valeurs dans $\mathbb{Z}$ est un $\mathbb{Z}$-module libre.*

Remarque. Par changement de base $\mathbb{Z} \to R$, si R est un anneau commutatif unitaire, $C^\infty(X, R)$ est alors aussi un R-module libre.

Démonstration. Que $C^\infty(X, \mathbb{Z})$ est un $\mathbb{Z}$-module est évident ; il s'agit de voir qu'il est libre. Soit f un élément de $C^\infty(X, \mathbb{Z})$. Pour tout $x \in X$, on fixe un ouvert V_x contenant x tel que f est constant sur V_x. Le compact X est alors recouvert par les V_x pour x parcourant X. Par compacité, on peut en extraire un recouvrement fini $X = \bigcup X_i$ par des ouverts X_i sur lesquels f est constant. Si X_i et X_j sont non disjoints, on peut les remplacer par $X_i \cap X_j$, $X_i \smallsetminus X_i \cap X_j$ et $X_j \smallsetminus X_i \cap X_j$, qui sont trois ouverts deux à deux disjoints d'union $X_i \cup X_j$. En répétant le procédé, on peut supposer que l'union $X = \bigcup X_i$ est une union finie disjointe.

On note que $\mathcal{U}$ est filtrant puisque si $U, V \in \mathcal{U}$, on peut, à partir du procédé précédent appliqué à $U \cup V$, obtenir un recouvrement W vérifiant $U \leq W$ et $V \leq W$. Pour tout $U = \{X_i\}$ élément de $\mathcal{U}$, on note $C_{(U)}(X, \mathbb{Z})$ le sous-$\mathbb{Z}$-module de $C^\infty(X, \mathbb{Z})$ constitué des fonctions constantes sur chaque X_i. On vient de voir que tout élément de $C^\infty(X, \mathbb{Z})$ vit dans un $C_{(U)}(X, \mathbb{Z})$ pour un $U \in \mathcal{U}$ convenable. Et si on considère les flèches d'inclusion $C_{(U)}(X, \mathbb{Z}) \to C_{(V)}(X, \mathbb{Z})$ pour $U \leq V$, on obtient un système inductif et on écrit $C^\infty(X, \mathbb{Z})$ comme limite inductive de modules libres :

$$C^\infty(X, \mathbb{Z}) = \varinjlim_{U \in \mathcal{U}} C_{(U)}(X, \mathbb{Z}) \simeq \varinjlim_{U \in \mathcal{U}} \mathbb{Z}^{|U|}.$$

On se sert de l'hypothèse sur $\mathcal{U}$ et on peut numéroter un sous-ensemble cofinal $\mathcal{V}$ de $\mathcal{U}$ (quitte à enlever des éléments du $\mathcal{V}$ de l'énoncé) $\mathcal{V} = \{V_n \mid n \in \mathbb{N}\}$ en respectant l'ordre : on demande $V_i \leq V_j \Rightarrow i \leq j$. On réécrit alors

$$C^\infty(X, \mathbb{Z}) = \varinjlim_{V \in \mathcal{V}} C_{(V)}(X, \mathbb{Z}) = \varinjlim_{n} \sum_{k \leq n} C_{(V_k)}(X, \mathbb{Z}).$$

On construit par récurrence sur $n \in \mathbb{N}$ une base de $C_n := \sum_{k \leq n} C_{(V_k)}(X, \mathbb{Z})$. L'étape d'initiation $n = 0$ consiste simplement à choisir une base $(b_0, \dots, b_{i_0})$ du $\mathbb{Z}$-module libre $C_{(V_0)}(X, \mathbb{Z})$. Supposons que l'on a construit une base $(b_0, \dots, b_{i_n})$ de C_n. Parce que $C_n \cap C_{(V_{n+1})}(X, \mathbb{Z})$ est un $C_{(V'_{n+1})}(X, \mathbb{Z})$ pour un certain $V'_{n+1} \leq V_{n+1}$ (V'_{n+1} non nécessairement dans $\mathcal{V}$), $C_{(V_{n+1})}(X, \mathbb{Z}) / \big(C_n \cap C_{(V_{n+1})}(X, \mathbb{Z})\big)$ est sans torsion. Alors C_n est un facteur direct du $\mathbb{Z}$-module libre C_{n+1}. On peut alors compléter $(b_0, \dots, b_{i_n})$ en une base $(b_0, \dots, b_{i_{n+1}})$ de C_{n+1}. La récurrence est alors prouvée et le résultat suit. □

Appendice B: Sur l'ordre $<_J$

On rappelle la définition de $<_J$ introduit dans [Grosse-Klönne 2014] : pour $w, w' \in W^J$, on note $w <_J w'$ s'il existe $s_1, \dots, s_r \in S$ tels que $w^{(i)} = (s_i s_{i-1} \cdots s_1 w)^J$ vérifie $l(w^{(i)}) > l(w^{(i-1)})$ pour tout $1 \leq i \leq r$ et $w^{(r)} = w'$.

On établit la caractérisation suivante de $<_J$, qui dit en particulier que $<_J$ est un raffinement de la restriction à W^J de l'ordre fort dans un groupe de Coxeter fini.

Proposition B.1. *Soient $w, w' \in W^J$. On a $w <_J w'$ si et seulement si il existe $s_1, \dots, s_r \in S$ tels que $w^{(i)} = s_i \cdots s_1 w$ soit un élément de W^J de longueur $l(w) + i$ pour tout $1 \leq i \leq r$ et $w^{(r)} = w'$.*

Remarque. Le cas $r = 1$ est déjà présent dans [Grosse-Klönne 2014, Lemma 1.4(b)].

Démonstration. Le sens ($\Leftarrow$) est immédiat puisque $w \in W^J$ implique $w^J = w$. Supposons donc $w <_J w'$ et prenons $s_1, \dots, s_r$ comme dans la définition de $<_J$. Prouvons d'abord $s_1 w \in W^J$ avec $l(s_1 w) = l(w) + 1$. On a

$$l(w) < l((s_1 w)^J) \leq l(s_1 w) \leq l(w) + 1,$$

où la première inégalité suite de la définition de $<_J$ et la deuxième de celle de W^J. Mais alors, on a $l((s_1 w)^J) = l(s_1 w) = l(w) + 1$. Cela dit en particulier que $s_1 w$ est de longueur minimale dans $s_1 w W_J$ et on a $(s_1 w)^J = s_1 w \in W^J$. Par une récurrence immédiate, $w^{(i)} = s_i \cdots s_1 w$ est un élément de W^J de longueur $l(w) + i$ pour tout $1 \leq i \leq r$. Le résultat est prouvé. ☐

Appendice C: De l'irréductibilité de la Steinberg généralisée dans le cas fini

Soit R un corps algébriquement clos de caractéristique p. Le travail effectué nous permet de découvrir ou redécouvrir quelques résultats sur les Steinberg généralisées pour un groupe réductif fini $\overline{G}$.

Proposition C.1. *La plus grande sous-$\overline{G}$-représentation irréductible de $\overline{\mathrm{St}}_{\overline{J}} R$ est V_J.*

Remarque. En particulier, si $\overline{\mathrm{St}}_{\overline{J}} R$ est irréductible, alors on a $V_J = \overline{\mathrm{St}}_{\overline{J}} R$.

Démonstration. On a une inclusion $\overline{\mathrm{St}}_{\overline{J}} R \subseteq \mathrm{St}_J\, R$, et on sait que le K-socle de $\mathrm{St}_J\, R$ est irréductible, égal à V_J : V_J est donc aussi le K-socle de $\overline{\mathrm{St}}_{\overline{J}} R$. Comme $K(1)$ agit trivialement sur $V_J \subseteq \overline{\mathrm{St}}_{\overline{J}} R$ et $\overline{\mathrm{St}}_{\overline{J}} R$, le résultat se traduit en termes de $\overline{G}$-représentations. ☐

Proposition C.2. *Supposons $\overline{\Phi}_{\mathrm{red}}$ irréductible et $\overline{J} \notin \{\varnothing, \overline{\Delta}\}$. Alors $\overline{\mathrm{St}}_{\overline{J}} R$ n'est pas irréductible.*

Remarque. $\mathrm{St}_{\overline{\Delta}}\, R = \mathrm{id}$ est bien sûr irréductible ; quant à la Steinberg $\mathrm{St}_{\varnothing}\, R$, en utilisant [Cabanes et Enguehard 2004, Theorem 6.10, Theorem 6.12 et Definition 6.13], on voit qu'elle est aussi irréductible.

Démonstration. Par [Cabanes et Enguehard 2004, Theorem 6.12], si V est une $\overline{G}$-représentation irréductible alors son espace de $\overline{U}$-invariants est de dimension 1. De ce fait, si V est une représentation avec $\dim V^{\overline{U}} \geq 2$, alors V n'est pas irréductible. Par les propositions 5.8 et 6.2, et le début de la preuve du corollaire 5.9, $(\overline{\mathrm{St}}_{\overline{J}} R)^{\overline{U}} = (\overline{\mathrm{St}}_{\overline{J}} R)^{\overline{B}}$ est un R-espace vectoriel de dimension $|W^J_{\mathrm{pr}}|$. Il s'agit d'examiner la cardinalité de W^J_{pr} et le résultat suit par le lemme C.3. ☐

Lemme C.3. *Supposons* Φ_{red} *irréductible. Alors on a* $|W^J_{\mathrm{pr}}| \geq 1$, *avec égalité si et seulement si* J *est* $\varnothing$ *ou* Δ.

Démonstration. Rappelons la définition suivante de W^J_{pr} :

$$W^J_{\mathrm{pr}} = \{w \in W \mid \forall \alpha \in J,\ l(ws_\alpha) > l(w)\ ;\ \forall \beta \in \Delta \smallsetminus J,\ l(ws_\beta) < l(w)\}.$$

Notons $w_{\Delta \smallsetminus J}$ l'élément le plus long de $W_{\Delta \smallsetminus J}$. C'est un élément de W^J_{pr}, et de ce fait on a la minoration voulue. Il reste à déterminer le cas d'égalité. D'abord, on remarque $W^{\varnothing}_{\mathrm{pr}} = \{w_\Delta\}$ et $W^{\Delta}_{\mathrm{pr}} = \{1\}$, de sorte qu'on veut maintenant montrer que si J n'est pas $\varnothing$ ou Δ, alors W^J_{pr} contient un autre élément que $w_{\Delta \smallsetminus J}$.

Supposons $J \neq \varnothing$, Δ. On cherche un élément $w \in W_J \smallsetminus \{1\}$ vérifiant

$$l(ww_{\Delta \smallsetminus J}s_\alpha) > l(ww_{\Delta \smallsetminus J})$$

pour tout $\alpha \in J$. Parce que Φ_{red} est irréductible, on peut choisir $\beta \in J$ tel que $(\Delta \smallsetminus J) \cup \{\beta\}$ engendre un sous-système de Φ_{red} avec au plus autant de composantes irréductibles que celui engendré par $\Delta \smallsetminus J$. Montrons que s_β est l'élément $w \in W^J \smallsetminus \{1\}$ cherché.

En effet, remarquons d'abord que l'on a

$$l(w_{\Delta \smallsetminus J}) = l(s_\beta w_{\Delta \smallsetminus J} s_\alpha) < l(s_\beta w_{\Delta \smallsetminus J})$$

pour tout $\alpha \in \Delta \smallsetminus J$. Ensuite, supposons qu'il existe un élément $\gamma \in J$ avec $l(s_\beta w_{\Delta \smallsetminus J} s_\gamma) < l(s_\beta w_{\Delta \smallsetminus J})$, alors cela veut dire que $s_\beta w_{\Delta \smallsetminus J}$ possède une écriture qui se termine par s_γ, disons $w' s_\gamma$ avec $l(w') = l(w_{\Delta \smallsetminus J})$ et w' ne se terminant pas par s_γ. Maintenant on a $W_J w_{\Delta \smallsetminus J} W_J = W_J w' W_J$, ce qui force $w' = w_{\Delta \smallsetminus J}$. Cela implique $s_\gamma = w_{\Delta \smallsetminus J} s_\beta w_{\Delta \smallsetminus J} \in \mathrm{W}_{(\Delta \smallsetminus J) \cup \{\beta\}}$ et donc $\gamma = \beta$. Mais dans ce cas-là, c'est que $s_\beta w_{\Delta \smallsetminus J}$ est l'élément le plus long de $W_{(\Delta \smallsetminus J) \cup \{\beta\}}$. Mais on sait aussi que la longueur de $w_{(\Delta \smallsetminus J) \cup \{\beta\}}$ est égale à $|\Phi^+_{(\Delta \smallsetminus J) \cup \{\beta\}}|$ [Humphreys 1992, I.4.8], et on a alors

$$|\Phi^+_{(\Delta \smallsetminus J) \cup \{\beta\}}| = |\Phi^+_{(\Delta \smallsetminus J)}| + 1.$$

Il s'ensuit

$$\Phi^+_{(\Delta \smallsetminus J) \cup \{\beta\}} = \Phi^+_{(\Delta \smallsetminus J)} \sqcup \{\beta\}$$

et cela contredit le fait que $\Phi_{(\Delta \smallsetminus J) \cup \{\beta\}}$ a moins (éventuellement le même nombre) de composantes irréductibles que $\Phi_{\Delta \smallsetminus J}$. C'est absurde, et le résultat est prouvé. □

Remarque. Marie-France Vignéras nous fait remarquer que l'élément $z^J = w_\Delta w_J \in W^J$ de la proposition 7.3 convient aussi en tant qu'élément distinct de $w_{\Delta \smallsetminus J}$ dans W^J_{pr} (voir aussi [Grosse-Klönne 2014, Lemma 1.4(e)]). En effet, il est de longueur maximale $|\Phi^+| - |\Phi^+_J| \neq |\Phi^+_{\Delta \smallsetminus J}|$ et est donc différent de $w_{\Delta \smallsetminus J}$. Et sa longueur excède aussi celle de tout élément de $W^{J'}$ pour $J' \supsetneq J$ et il est donc primitif.

Appendice D: Représentations de Steinberg généralisées pour le groupe dérivé

Dans cette section uniquement, on distinguera le groupe réductif $\underline{G}$ défini sur F de ses F-points $G = \underline{G}(F)$. De même, B et P seront respectivement les F-points de $\underline{B}$ et $\underline{P}$.

On note $\underline{D}(G)$ le groupe dérivé de G, c'est-à-dire le faisceau fppf des commutateurs de $\underline{G}$. C'est un groupe semi-simple (voir [Demazure 2011a, Théorème 6.2.1(iv)]) et on notera $D(G)$ pour son groupe des F-points. Le but de ce paragraphe est de comparer les représentations de Steinberg généralisées pour G et pour $D(G)$.

Soient R un corps de caractéristique p et $J \subseteq \Delta$. Pour les distinguer, on notera $\mathrm{St}_J^{(G)} R$ et $\mathrm{St}_J^{(D(G))} R$ la représentation de Steinberg généralisée respectivement pour G et $D(G)$, par rapport à J.

Proposition D.1. *La restriction de* $\mathrm{St}_J^{(G)} R$ *à* $D(G)$ *est isomorphe à* $\mathrm{St}_J^{(D(G))} R$.

A partir de là, on peut utiliser tout la machinerie de cet article pour $\underline{D}(G)$ et en déduire l'irréductibilité de $\mathrm{St}_J^{(D(G))} R$; automatiquement $\mathrm{St}_J^{(G)} R$ est aussi irréductible.

Démonstration de la proposition D.1. Le groupe $\underline{B} \cap \underline{D}(G)$ est un parabolique minimal de $\underline{D}(G)$ (voir [Demazure 2011a, Proposition 6.2.8(ii)]), et on appelle standard un parabolique de $\underline{D}(G)$ contenant $\underline{B} \cap \underline{D}(G)$. La flèche $\underline{P} \mapsto \underline{P} \cap \underline{D}(G)$ est une bijection entre l'ensemble des paraboliques standards de $\underline{G}$ et celui des paraboliques standards de $\underline{D}(G)$. Soient $\underline{P}$ un parabolique standard de $\underline{G}$ et $J \subseteq \Delta$ l'ensemble vérifiant $\underline{P} = \underline{P}_J$. On veut voir que l'injection $P \cap D(G) \backslash D(G) \hookrightarrow P \backslash G$ est une bijection. Pour cela, utilisons la décomposition de Bruhat pour $D(G)$:

$$P \cap D(G) \backslash D(G) = \bigsqcup_{w \in W^J} P \cap D(G) \backslash (P \cap D(G)) w^{-1} (B \cap D(G)),$$

où on a relevé chaque $w \in W^J$ en un élément de $D(G)$. En notant que $U_w = U \cap wU^- w^{-1}$ est inclus dans $D(G)$, elle se réécrit encore

$$P \cap D(G) \backslash D(G) = \bigsqcup_{w \in W^J} P \cap D(G) \backslash (P \cap D(G)) w^{-1} U_w. \tag{29}$$

De même, pour G on écrit (en gardant les mêmes relèvements pour W^J)

$$P \backslash G = \bigsqcup_{w \in W^J} P \backslash P w^{-1} B = \bigsqcup_{w \in W^J} P \backslash P w^{-1} U_w. \tag{30}$$

La comparaison de (30) et de (29) nous donne que $P \cap D(G) \backslash D(G) \hookrightarrow P \backslash G$ est surjective, et donc bijective.

De ce fait, la restriction à $D(G)$ de l'induite Ind_P^G id est $\mathrm{Ind}_{P\cap D(G)}^{D(G)}$ id. Par définition (18), la restriction à $D(G)$ de $\mathrm{St}_J^{(G)}\, R$ s'identifie à $\mathrm{St}_J^{(D(G))}\, R$. □

Remerciements

Ce travail est une partie de ma thèse de doctorat, réalisée sous la direction de Marie-France Vignéras. Je lui suis infiniment reconnaissant pour les questions et remarques qu'elle a pu me communiquer pendant la préparation de cet article. Je remercie aussi Guy Henniart et Florian Herzig pour leur relecture attentive de ce texte.

Bibliographie

[Borel et Tits 1965] A. Borel et J. Tits, "Groupes réductifs", *Inst. Hautes Études Sci. Publ. Math.* **27** (1965), 55–151. MR 34 #7527 Zbl 0145.17402

[Bruhat et Tits 1972] F. Bruhat et J. Tits, "Groupes réductifs sur un corps local, I: Données radicielles valuées", *Inst. Hautes Études Sci. Publ. Math.* **41** (1972), 5–251. MR 48 #6265 Zbl 0254.14017

[Bruhat et Tits 1984] F. Bruhat et J. Tits, "Groupes réductifs sur un corps local, II: Schémas en groupes. Existence d'une donnée radicielle valuée", *Inst. Hautes Études Sci. Publ. Math.* **60** (1984), 5–184. MR 86c:20042 Zbl 0597.14041

[Bushnell et Henniart 2006] C. J. Bushnell et G. Henniart, *The local Langlands conjecture for* GL(2), Grundlehren der Mathematischen Wissenschaften **335**, Springer, Berlin, 2006. MR 2007m:22013 Zbl 1100.11041

[Cabanes et Enguehard 2004] M. Cabanes et M. Enguehard, *Representation theory of finite reductive groups*, New Mathematical Monographs **1**, Cambridge University Press, 2004. MR 2005g:20067 Zbl 1069.20032

[Carter 1985] R. W. Carter, *Finite groups of Lie type: conjugacy classes and complex characters*, Wiley, New York, 1985. MR 87d:20060 Zbl 0567.20023

[Curtis 1966] C. W. Curtis, "The Steinberg character of a finite group with a $(B,\ N)$-pair", *J. Algebra* **4** (1966), 433–441. MR 34 #1406 Zbl 0161.02203

[Demazure 2011a] M. Demazure, "Groupes réductifs: déploiements, sous-groupes, groupes quotients (Exposé XXII)", pp. 109–176 dans *Schémas en groupes* (SGA 3)*, tome III: Structure des schémas en groupes réductifs*, édité par P. Gille et P. Polo, Documents Mathématiques **8**, Société Mathématique de France, Paris, 2011. MR 2867622 Zbl 1241.14003

[Demazure 2011b] M. Demazure, "Sous-groupes paraboliques des groupes réductifs (Exposé XXVI)", dans *Schémas en groupes* (SGA 3)*, tome III: Structure des schémas en groupes réductifs*, édité par P. Gille et P. Polo, Documents Mathématiques **8**, Société Mathématique de France, Paris, 2011. MR 2867622 Zbl 1241.14003

[Grosse-Klönne 2014] E. Grosse-Klönne, "On special representations of p-adic reductive groups", *Duke Math. J.* **163**:12 (2014), 2179–2216. MR 3263032 Zbl 1298.22018

[Haines 2009] T. J. Haines, "Corrigendum: The base change fundamental lemma for central elements in parahoric Hecke algebras", 2009, Voir http://www2.math.umd.edu/~tjh/fl_corr3.pdf.

[Haines et Rapoport 2008] T. J. Haines et M. Rapoport, "Appendix: On parahoric subgroups", *Adv. Math.* **219**:1 (2008), 188–198. MR 2009g:22039 Zbl 1159.22010

[Haines et Rostami 2010] T. J. Haines et S. Rostami, “The Satake isomorphism for special maximal parahoric Hecke algebras”, *Rep. Theory* **14** (2010), 264–284. MR 2011g:20077 Zbl 1251.22013

[Henniart et Vignéras 2012] G. Henniart et M.-F. Vignéras, “Comparison of compact induction with parabolic induction”, *Pacific J. Math.* **260**:2 (2012), 457–495. MR 3001801 Zbl 1284.22009

[Henniart et Vignéras 2015] G. Henniart et M.-F. Vignéras, “A Satake isomorphism for representations modulo p of reductive groups over local fields”, *J. Reine Angew. Math.* **701** (2015), 33–75. MR 3331726 Zbl 06424795

[Herzig 2011] F. Herzig, “The classification of irreducible admissible mod p representations of a p-adic GL_n”, *Invent. Math.* **186**:2 (2011), 373–434. MR 2845621 Zbl 1235.22030

[Humphreys 1992] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge University Press, 1992. MR 92h:20002 Zbl 0768.20016

[Kottwitz 1997] R. E. Kottwitz, “Isocrystals with additional structure, II”, *Compositio Math.* **109**:3 (1997), 255–339. MR 99e:20061 Zbl 0966.20022

[Paskunas 2004] V. Paskunas, *Coefficient systems and supersingular representations of* $\mathrm{GL}_2(F)$, Mémoires de la Société Mathématique de France **99**, Société Mathématique de France, Paris, 2004. MR 2005m:22017 Zbl 1249.22010 arXiv math/0403240

[Steinberg 1951] R. Steinberg, “A geometric approach to the representations of the full linear group over a Galois field”, *Trans. Amer. Math. Soc.* **71** (1951), 274–282. MR 13,317d Zbl 0045.30201

[Vignéras 2012a] M.-F. Vignéras, “Emerton’s ordinary parts in positive characteristic”, communication personnelle, 2012.

[Vignéras 2012b] M.-F. Vignéras, “Représentations p-adiques de torsion admissibles”, pp. 639–646 dans *Number theory, analysis and geometry: in memory of Serge Lang*, édité par D. Goldfeld et al., Springer, New York, 2012. MR 2867935 Zbl 1251.22009

Received March 21, 2014. Revised October 6, 2014.

TONY LY
DMA
ECOLE NORMALE SUPÉRIEURE
45, RUE D’ULM
75005 PARIS CEDEX 05
FRANCE
tony.ly@ens.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 277, No. 2, 2015
dx.doi.org/10.2140/pjm.2015.277.463

CALCULATING TWO-STRAND JELLYFISH RELATIONS

DAVID PENNEYS AND EMILY PETERS

We construct a $3^{\mathbb{Z}/4}$ subfactor using an algorithm which, given generators in a spoke graph planar algebra, computes two-strand jellyfish relations. This subfactor was known to Izumi, but has not previously appeared in the literature. We systematically analyze the space of second annular consequences, adapting Jones' treatment of the space of first annular consequences in his quadratic tangles article.

This article is the natural followup to two recent articles on spoke subfactor planar algebras and the jellyfish algorithm. Work of Bigelow and Penneys explains the connection between spoke subfactor planar algebras and the jellyfish algorithm, and work of Morrison and Penneys automates the construction of subfactors where both principal graphs are spoke graphs using one-strand jellyfish. This is the published version of arXiv:1308.5197.

1. Introduction

Jones' program for constructing subfactor planar algebras starts with the observation that every subfactor planar algebra embeds in the graph planar algebra (first defined in [Jones 2000]) of its principal graph [Jones and Penneys 2011; Morrison and Walker 2010]. Following this program, one constructs a subfactor planar algebra by finding candidate generators in an appropriate graph planar algebra, and then showing they generate a subfactor planar algebra with the correct principal graph.

These methods have been used to construct a large handful of examples, some new and some well known, including the E_6 and E_8 subfactors [Jones 2001], group-subgroup subfactors [Gupta 2008], the Haagerup subfactor [Peters 2010], the extended Haagerup subfactor [BMPS 2012], the Izumi–Xu 2221 subfactor [Han 2010], certain spoke subfactors, e.g., 4442 [Morrison and Penneys 2015b], and examples related to quantum groups [LMP 2015]. These techniques have also been used to prove uniqueness results [BMPS 2012; Han 2010; Liu 2015] and obstructions to possible principal graphs [Peters 2010; Morrison 2014; Liu 2015].

Early applications of the embedding theorem to construct or obstruct subfactors were mostly ad hoc. Recent work of Bigelow and Penneys [2014], based on [Popa

---

MSC2010: primary 46L37; secondary 18D05, 57M20.

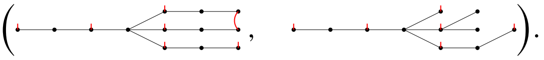
Keywords: subfactors, planar algebras, jellyfish algorithm, principal graphs.

1995], has explained why some of the previous constructions work and how they fit into the same family of examples. If the principal graph of a subfactor is a spoke graph with simple arms connected to one central vertex, the planar algebra can be constructed using two-strand jellyfish relations. If both graphs are spokes, one can use one-strand relations, which are easier to compute. Recent work of Morrison and Penneys [2015b] found an algorithm to compute these one-strand relations, provided one has the generators in the graph planar algebra.

Given a set of generators in a graph planar algebra together with some local relations, we want to show evaluability, i.e., the relations can evaluate any closed diagram. The utility of the jellyfish algorithm is that for spoke graphs, it gives a systematic way to show evaluability. The key idea of the jellyfish algorithm is that given our generators, and relatively few evaluations of closed diagrams involving these generators, we can derive a collection of local relations sufficient to evaluate all closed diagrams.

This article is the natural followup to [Bigelow and Penneys 2014; Morrison and Penneys 2015b]. Our main result is an algorithm to find two-strand jellyfish relations for a subfactor planar algebra for which one of the principal graphs is a spoke graph. This algorithm requires as input the generators in a graph planar algebra. The main application of our algorithm is the construction of a subfactor known to Izumi, which has not previously appeared in the literature.

Theorem 1.1. *There exists a* $3^{\mathbb{Z}/4}$ *subfactor with principal graphs*



We describe our algorithm for the reader who is willing to take our computations on faith.

- Acquire the generators in an appropriate graph planar algebra. These generators are an assignment of numbers in a finite extension of $\mathbb{Q}$ to certain loops on a graph.
- Use a computer to evaluate certain closed diagrams with at most 4 generators. This amounts to multiplying rather large matrices, and taking the trace.
- Turn these evaluations of closed diagrams into information about inner products, and then use a computer to derive jellyfish relations for our generators. The use of the computer is limited to basic linear algebra.
- We now have an evaluable planar subalgebra of a graph planar algebra, which is necessarily a subfactor planar algebra. Compute the principal graph by a process of elimination.

By a deep theorem of Popa [1995], from a subfactor planar algebra $\mathcal{P}_\bullet$ we can always get a subfactor whose stand invariant is $\mathcal{P}_\bullet$. When $\mathcal{P}_\bullet$ is finite depth, it is a complete invariant of the associated hyperfinite subfactor [Popa 1990].

Remark 1.2. Interestingly, we construct the $3^{\mathbb{Z}/4}$ subfactor planar algebra in a graph planar algebra not associated to either of its principal graphs (see Appendix AA)! Of course, by the embedding theorem, it is also a planar subalgebra of the $3^{\mathbb{Z}/4}$ graph planar algebra, but we found the computational issues related to finding the generators easier to deal with in the other graph.

The motivation for this article is to systematically study a conjectural infinite family of 3^G spoke subfactors for certain finite abelian groups G, first studied by Izumi [2001], and later by Evans and Gannon [2011]. A 3^G subfactor has principal graph consisting of $|G|$ spokes of length 3, and the dual data is determined by the inverse law of the group G. In fact, Izumi has an unpublished construction of a $3^{\mathbb{Z}/4}$ subfactor using Cuntz algebras, analogous to his treatment for odd order G in [Izumi 2001]. Moreover, he can show such a subfactor is unique, which our approach does not attempt. In theory, all 3^G subfactors can be constructed using two-strand jellyfish [Bigelow and Penneys 2014]. The major hurdle is finding the generators in the graph planar algebra. Once given the generators, the machinery of this article produces the two-strand relations.

The foundation for this article, which underlies the previously discussed constructions and obstructions, is Jones' annular tangles point of view. Each unitary planar algebra can be orthogonally decomposed into irreducible annular Temperley–Lieb modules. In doing so, we seem to lose a lot of information, namely the action of higher genus tangles. However, we find ourselves in the simpler situation of analyzing irreducible annular Temperley–Lieb modules, which have been completely classified [Graham and Lehrer 1998; Jones 2001]. Such a module is generated by a single low-weight rotational eigenvector. This perspective is particularly useful for small index subfactors, which can only have a few small low-weight vectors.

This article is also a natural followup to Jones' exploration of quadratic tangles [2012]. There are necessarily strong quadratic relations among the few smallest low-weight generators of a subfactor planar algebra of small modulus. Jones [2012] studies the space of first annular consequences of the low-weight vectors to find explicit formulas for these relations. We provide an analogous systematic treatment of the space of second annular consequences of a set of low-weight generators of a subfactor planar algebra. Studying this space was fruitful in Peters' [2010] planar algebra construction of the Haagerup subfactor.

1A. ***Outline.*** In Section 2, we give the necessary background for this article, including conventions for graph planar algebras, tetrahedral structure constants, the jellyfish algorithm, and reduced trains. In Section 2D, we give a basis for the second annular consequences of a low-weight element when $\delta > 2$.

In Section 3, we analyze the space of reduced trains, in particular their projections to Temperley–Lieb and annular consequences. We then calculate many pairwise inner products of such trains and their projections. In Section 4, we provide the algorithm for computing two-strand jellyfish relations given generators in our graph planar algebras.

In Section 5, we provide the results of applying the algorithm from Section 4 to construct the $3^{\mathbb{Z}/4}$ subfactor planar algebra. We compute the principal graphs of our example in Section 6.

Finally, we have two appendices where we record the data necessary for the above computations. The generators are specified in Appendix A via their values on collapsed loops, and we give the moments and tetrahedral structure constants for our generators in Appendix B.

1B. ***The FusionAtlas*** (adapted from [Morrison and Penneys 2015b]). This article relies on some substantial calculations. In particular, our efforts to find the generators in the various graph planar algebras made use of a variety of techniques, some ad hoc, some approximate, and some computationally expensive. This article essentially does not address that work. Instead, we merely present the discovered generators and verify some relatively easy facts about them. In particular, the proofs presented in this article rely on the computer in a much weaker sense. We need to calculate certain numbers of the form $\text{Tr}(PQRS)$, where P, Q, R, S are rather large martrices, and the computer does this for us. We also entered all the formulas derived in this article into Mathematica in order to evaluate the various quantities which appear in our derivation of jellyfish relations. As a reader may be interested in seeing these programs, we include a brief instruction on finding and running these programs.

The arXiv sources of this article contain a number of files in the `code` subdirectory, including:

- `Generators.nb`, which reconstructs the generators from our terse descriptions in Appendix A.
- `TwoStrandJellyfish.nb`, which calculates the requisite moments and tetrahedral structure constants of these generators, and performs the linear algebra necessary to derive the jellyfish relations.
- `GenerateLaTeX.nb`, which typesets each subsection of Section 5 for each planar algebra, and many mathematical expressions in Appendices A and B.

The Mathematica notebook `Generators.nb` can be run by itself. The final cells of that notebook write the full generators to the disk; this must be done before running `TwoStrandJellyfish.nb`. The latter notebook relies on the *FusionAtlas*, a substantial body of code the authors have developed along with Narjess Afzaly, Scott Morrison, Noah Snyder, and James Tener to perform calculations with subfactors and fusion categories. To obtain a local copy, you first need to ensure that you have *Mercurial*, the distributed version control system, installed on your machine. With that, the command

```
hg clone https://bitbucket.org/fusionatlas/fusionatlas
```

will create a local directory called `fusionatlas` containing the latest version. In the `TwoStrandJellyfish.nb` notebook, you will then need to adjust the paths appearing in the first input cell to ensure that your local copy is included. After that, running the entire notebook reproduces all the calculations described below.

We invite any interested readers to contact us with questions or queries about the use of these notebooks or the *FusionAtlas* package.

2. Background

We now give the background material for the calculations that occur in the later sections. We refer the reader to [Peters 2010; BMPS 2012; Jones 2012; 2011] for the definition of a (subfactor) planar algebra.

Notation 2.1. When we draw planar diagrams, we often suppress the external boundary disk. In this case, the external boundary is assumed to be a large rectangle whose distinguished interval contains the upper left corner. We draw one string with a number next to it instead of drawing that number of parallel strings. We shade the diagrams as much as possible, but if the parity is unknown, we often cannot know how to shade them. Finally, projections are usually drawn as rectangles with the same number of strands emanating from the top and bottom, while other elements may be drawn as circles.

Some parts of this introduction are adapted from [Morrison and Penneys 2015b; Bigelow and Penneys 2014].

2A. *Working in graph planar algebras.*

Graph planar algebras, defined in [Jones 2000], have proven to be a fruitful place to work because of the following theorem. Strictly speaking, our constructions do not rely on this theorem. However, it motivates our search for generators in the appropriate graph planar algebra.

Theorem 2.2 [Jones and Penneys 2011; Morrison and Walker 2010]. *Every subfactor planar algebra embeds in the graph planar algebra of its principal graph.*

In [Morrison and Penneys 2015b, Section 2.2], it was observed that many of Jones' [2012] quadratic tangles formulas for subfactor planar algebras hold for certain collections of elements in unitary, spherical, shaded $*$-planar algebras which are not necessarily evaluable (see Theorem 2.8). The main example of such a planar algebra is the graph planar algebra of a finite bipartite graph. We give the necessary definitions and discuss our conventions for working in such planar algebras in this subsection.

Definition 2.3. A shaded planar $*$-algebra is *evaluable* if $\dim(\mathcal{P}_{n,\pm}) < \infty$ for all $n \geq 0$, and $\mathcal{P}_{0,\pm} \cong \mathbb{C}$ as $*$-algebras. In this case, this isomorphism must send the empty diagram to 1.

Suppose $\mathcal{P}_\bullet$ is a shaded planar $*$-algebra which is not necessarily evaluable. We call $\mathcal{P}_\bullet$ *unitary* if for all $n \geq 0$, the $\mathcal{P}_{0,\pm}$-valued sesquilinear form on $\mathcal{P}_{n,\pm}$ given by $\langle x, y\rangle = \operatorname{Tr}(y^* x)$ is positive definite (in the operator-valued sense).

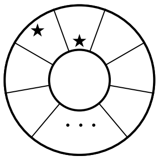
We call such a planar algebra *spherical* if, for any closed diagram in $\mathcal{P}_\bullet$ which equals a scalar multiple of the empty diagram, performing isotopy on a sphere still gives us the same scalar multiple of the appropriate empty diagram.

Remark 2.4. The above is only one possible definition of unitarity for a planar $*$-algebra. One might also want to require the existence of a faithful state on $\mathcal{P}_{0,\pm}$ which induces a C^*-algebra structure on the algebras $\mathcal{P}_{n,\pm}$ in the usual GNS way. However, the above frugal definition is sufficient for our purposes, since the following theorem holds.

Theorem 2.5. *Suppose $\mathcal{P}_\bullet$ is a spherical, unitary, shaded planar $*$-algebra which is not necessarily evaluable. If $\mathcal{Q}_\bullet \subset \mathcal{P}_\bullet$ is an evaluable planar $*$-subalgebra, then $\mathcal{Q}_\bullet$ is a subfactor planar algebra.*

Proof. Since $\mathcal{Q}_\bullet$ is evaluable, sphericality of $\mathcal{Q}_\bullet$ follows from sphericality of $\mathcal{P}_\bullet$. Now, the sesquilinear form $\langle x, y\rangle = \operatorname{Tr}(y^* x)$ on $\mathcal{Q}_{n,\pm}$ is operator-valued positive definite. Since $\mathcal{Q}_\bullet$ is evaluable, by identifying the appropriate empty diagram with $1 \in \mathbb{C}$, we get a positive definite inner product. □

Notation 2.6. Recall that the Fourier transform $\mathcal{F}$ is given by

$\mathcal{F} =$  .

For a rotational eigenvector $S \in \mathcal{P}_{n,\pm}$ corresponding to an eigenvalue $\omega_S = \sigma_S^2$, we define another rotational eigenvector $\check{S} \in \mathcal{P}_{n,\mp}$ by $\check{S} = \sigma_S^{-1}\mathcal{F}(S)$. Note that $\mathcal{F}(\check{S}) = \sigma_S S$, so $\check{\check{S}} = S$.

Definition 2.7. Suppose $\mathcal{P}_\bullet$ is a unitary, spherical, shaded planar $*$-algebra with modulus $\delta > 2$ which is not necessarily evaluable. A finite set $\mathfrak{B} \subset \mathcal{P}_{n,+}$ is called a *set of minimal generators for* $\mathcal{Q}_\bullet$ if the elements of $\mathfrak{B}$ generate the planar $*$-subalgebra

$\mathcal{Q}_\bullet \subset \mathcal{P}_\bullet$ and are linearly independent, self-adjoint, low-weight eigenvectors for the rotation, i.e, for all $S \in \mathfrak{B}$,

- $S = S^*$,
- S is uncappable, and
- $\rho(S) = \omega_S S$ for some n-th root of unity ω_S.

In the sequel, when we refer to a set of minimal generators without mentioning $\mathcal{Q}_\bullet$, assume that $\mathcal{Q}_\bullet$ is the planar $*$-subalgebra generated by $\mathfrak{B}$.

Given a set of minimal generators $\mathfrak{B}$, we get a set of dual minimal generators $\check{\mathfrak{B}} = \{\check{S} \mid S \in \mathfrak{B}\}$. We say a set of minimal generators $\mathfrak{B}$ has *scalar moments* if $\operatorname{Tr}(R)$, $\operatorname{Tr}(RS)$, $\operatorname{Tr}(RST)$ and $\operatorname{Tr}(\check{R})$, $\operatorname{Tr}(\check{R}\check{S})$, $\operatorname{Tr}(\check{R}\check{S}\check{T})$ are scalar multiples of the empty diagram in $\mathcal{P}_{0,+}$ and $\mathcal{P}_{0,-}$ respectively for each $R, S, T \in \mathfrak{B}$.

If a set of minimal generators $\mathfrak{B}$ has scalar moments, we say $\mathfrak{B}$ is

- *orthogonal* if $\langle S, T\rangle = \operatorname{Tr}(ST) = 0$ if $S \neq T$ for all $S, T \in \mathfrak{B}$, and
- *orthonormal* if $\mathfrak{B}$ is orthogonal and $\operatorname{Tr}(S^2) = \langle S, S\rangle = 1$ for all $S \in \mathfrak{B}$.

The point of working with sets of minimal generators is the following theorem.

Theorem 2.8 [Morrison and Penneys 2015b, Theorem 2.5]. *All the formulas of Section 4 of* [Jones 2012] *hold in any unitary, spherical, shaded planar $*$-algebra with modulus $\delta > 2$ for any orthonormal set of minimal generators $\mathfrak{B}$ with scalar moments.*

Assumption 2.9. For the rest of the article, unless otherwise specified, we assume $\mathcal{P}_\bullet$ is a unitary, spherical, shaded $*$-planar algebra with modulus $\delta > 2$ which is not necessarily evaluable, and $\mathfrak{B} \subset \mathcal{P}_{n,+}$ is an orthogonal set of minimal generators with scalar moments.

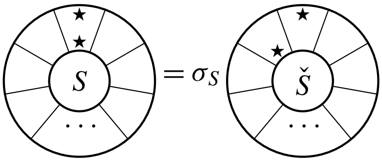
Since we do not assume our generators in $\mathfrak{B}$ are orthonormal, our formulas will differ slightly in appearance from those of [Jones 2012] and [Morrison and Penneys 2015b].

Remark 2.10. For diagram evaluation, it is useful to have our standard equations for our set of minimal generators in one place. For $S \in \mathfrak{B}$,

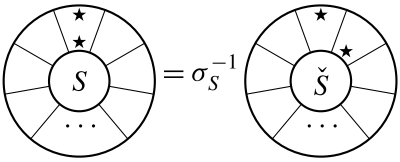
$$\begin{array}{llll} S = S^* & \mathcal{F}^2 = \rho & \rho(S) = \omega_S S & \mathcal{F}(S) = \sigma_S \check{S} \\ \check{S} = \check{S}^* & \sigma_S^2 = \omega_S & \rho(\check{S}) = \omega_S \check{S} & \mathcal{F}(\check{S}) = \sigma_S S. \end{array}$$

When moving $\star$ on the distinguished interval of a generator, the resulting diagram is multiplied by some exponent of σ_S:

- If you shift $\star$ counterclockwise by one strand, multiply by σ_S and switch $\check{}$:



- If you shift $\star$ clockwise by one strand, multiply by σ_S^{-1} and switch $\check{}$:



Using notation from [Jones 2012], for $P, Q, R \in \mathfrak{B}$, we set $a_R^{PQ} = \mathrm{Tr}(PQR)$ and $b_R^{PQ} = \mathrm{Tr}(\check{P}\check{Q}\check{R})$.

Remark 2.11. Once we have determined our set of minimal generators $\mathfrak{B}$ has scalar moments, the next thing to do is to verify that the complex spans of $\mathfrak{B} \cup \{f^{(n)}\}$ and $\check{\mathfrak{B}} \cup \{f^{(n)}\}$ form algebras under the usual multiplication. If this is the case, for $P, Q \in \mathfrak{B}$, we necessarily have

$$PQ = \frac{\mathrm{Tr}(PQ)}{[n+1]} f^{(n)} + \sum_{R \in \mathfrak{B}} \frac{a_R^{PQ}}{\|R\|^2} R. \tag{1}$$

Immediately, we get that all higher moments of $\mathfrak{B}$, $\check{\mathfrak{B}}$ are scalars, as are certain tetrahedral structure constants (see Remark 2.15 and Example 2.17). For example, we have that

$$\mathrm{Tr}(PQRS) = \frac{\mathrm{Tr}(PQ)\,\mathrm{Tr}(RS)}{[n+1]} + \sum_{T \in \mathfrak{B}} \frac{a_T^{PQ}}{\|T\|^2} a_T^{RS}. \tag{2}$$

for $P, Q, R, S \in \mathfrak{B}$.

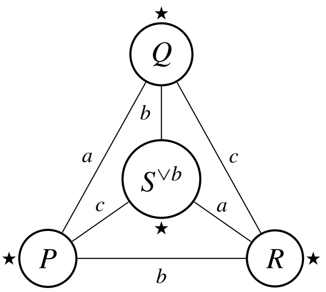
Assumption 2.12. We now assume the complex spans of $\mathfrak{B} \cup \{f^{(n)}\}$ and $\check{\mathfrak{B}} \cup \{f^{(n)}\}$ form algebras under the usual multiplication.

Remark 2.13. The assumptions of this subsection are significant. A randomly chosen subset of a graph planar algebra will not satisfy Assumption 2.9. Given an orthogonal set of minimal generators $\mathfrak{B}$ with scalar moments, it is still possible it will not satisfy Assumption 2.12. For example, if we start with a $\mathfrak{B}$ satisfying Assumptions 2.9 and 2.12 and we discard one element, the resulting set together with $f^{(n)}$ may not span an algebra.

2B. ***Tetrahedral structure constants.*** We will also need the tetrahedral structure constants defined in [Jones 2003, Section 3.3].

Definition 2.14. For $P, Q, R, S \in \mathfrak{B}$, we define

$$\Delta_{a,b}(P, Q, R \mid S) \quad =$$



where $c = 2n - a - b$, and

$$S^{\vee b} = \begin{cases} S & \text{if } b \text{ is even,} \\ \check{S} & \text{if } b \text{ is odd.} \end{cases}$$

Note that the $\Delta_{a,b}(P, Q, R \mid S)$ for $P, Q, R, S \in \mathfrak{B}$ determine all the tetrahedral structure constants by [Jones 2003, Section 3.3].

Remark 2.15. For this article, we only need the following tetrahedral structure constants:

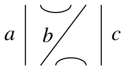
- $\Delta_{n-1,2}(P, Q, R \mid S)$
- $\Delta_{n,1}(P, Q, R \mid S)$
- $\Delta_{n-1,1}(P, Q, R \mid S) = \overline{\Delta_{n,1}(R, Q, P \mid S)}$.

By Assumption 2.12, we can express the second and third tetrahedral structure constants above in terms of the moments and chiralities of $\mathfrak{B}$ and $\check{\mathfrak{B}}$, since one of $a, b, c \geq n$. We do this computation in Example 2.17. Thus for convenience, we will just write $\Delta(P, Q, R \mid S)$ instead of $\Delta_{n-1,2}(P, Q, R \mid S)$, and we will only write subscripts a, b if $a \neq n-1$ or $b \neq 2$. For each of our planar algebras in this article, we give the tetrahedral structure constants $\Delta(P, Q, R \mid S)$ in Appendix B.

Since we will use it repeatedly, we reproduce the following well-known fact for convenience.

Fact 2.16 [Morrison 2015; Reznikoff 2007]. *The coefficient of the below Temperley–Lieb diagram in the Jones–Wenzl idempotent $f^{(k)}$ is given by*

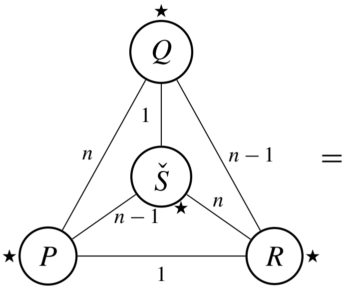
$$\underset{\in f^{(k)}}{\operatorname{coeff}}\Bigg($$

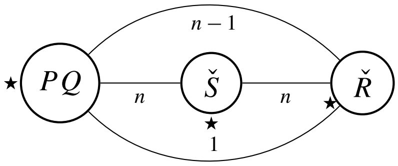


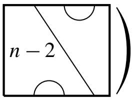
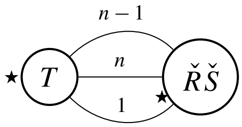
$$\Bigg) = (-1)^{b+1} \frac{[a+1][c+1]}{[k]}.$$

Example 2.17. In the following calculation, we use (1) for the third equality and 2.16 for the coefficient in the Jones–Wenzl idempotent appearing in the third line.

$\Delta_{n,1}(P, Q, R \mid S)$

$=$ 

$= \sigma_R^{-1}$ 

$= \sigma_R^{-1} \dfrac{\mathrm{Tr}(PQ)\,\mathrm{Tr}(\check{R}\check{S})}{[n+1]} \operatorname*{coeff}_{\in f^{(n)}} \Bigg($  $\Bigg) + \displaystyle\sum_{T\in\mathfrak{B}} \sigma_R^{-1} \frac{a_T^{PQ}}{\|T\|^2}$ 

$$= (-1)^{n-1}\sigma_R^{-1}\frac{\mathrm{Tr}(PQ)\,\mathrm{Tr}(\check{R}\check{S})}{[n][n+1]} + \sum_{T\in\mathfrak{B}} \sigma_T\sigma_R^{-1}\frac{a_T^{PQ}b_T^{RS}}{\|T\|^2}.$$

By symmetry, we get

$$\begin{aligned}\Delta_{n-1,1}(P, Q, R \mid S) &= \overline{\Delta_{n,1}(R, Q, P \mid S)}\\ &= (-1)^{n-1}\sigma_P\frac{\mathrm{Tr}(QR)\,\mathrm{Tr}(\check{P}\check{S})}{[n][n+1]} + \sum_{T\in\mathfrak{B}} \sigma_T^{-1}\sigma_P\frac{a_T^{QR}b_T^{SP}}{\|T\|^2}.\end{aligned}$$

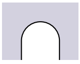
Lemma 2.18. *We have the following symmetries*:

$$\begin{aligned}\Delta(P, Q, R \mid S) &= \overline{\Delta(R, Q, P \mid S)}\\ &= \omega_P\omega_R^{-1}\Delta(R, S, P \mid Q)\\ &= \omega_P\omega_R^{-1}\overline{\Delta(P, S, R \mid Q)}\\ &= \sigma_P^{1-n}\sigma_Q^{n-1}\sigma_R^{n-1}\sigma_S^{1-n}\Delta(Q^{\vee(n-1)}, P^{\vee(n-1)}, S^{\vee(n-1)} \mid R^{\vee(n-1)})\\ &= \sigma_P^{1-n}\sigma_Q^{n-1}\sigma_R^{n-1}\sigma_S^{1-n}\overline{\Delta(S^{\vee(n-1)}, P^{\vee(n-1)}, Q^{\vee(n-1)} \mid R^{\vee(n-1)})}\\ &= \sigma_P^{1-n}\sigma_Q^{n+1}\sigma_R^{n-1}\sigma_S^{-1-n}\Delta(S^{\vee(n-1)}, R^{\vee(n-1)}, Q^{\vee(n-1)} \mid P^{\vee(n-1)})\\ &= \sigma_P^{1-n}\sigma_Q^{n+1}\sigma_R^{n-1}\sigma_S^{-1-n}\overline{\Delta(Q^{\vee(n-1)}, R^{\vee(n-1)}, S^{\vee(n-1)} \mid P^{\vee(n-1)})}\end{aligned}$$

Proof. Immediate from drawing diagrams using unitarity and sphericality of $\mathcal{P}_\bullet$. □

Remark 2.19. As in [Morrison and Peters 2014; Morrison and Penneys 2015b], when doing calculations in the graph planar algebra, we work with the lopsided convention rather than the spherical convention (see [Morrison and Peters 2014]). The lopsided convention treats shaded and unshaded contractible loops differently, which has the advantage that there are fewer square roots, so arithmetic is easier.

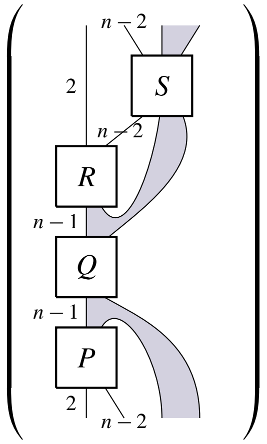
The translation map $\natural\colon \mathcal{P}_\bullet^{\text{spherical}} \to \mathcal{P}_\bullet^{\text{lopsided}}$ between the conventions from [Morrison and Peters 2014] is not a planar algebra map, but it commutes with the action of the planar operad up to a scalar. We determine the scalar by first drawing the tangle in a standard rectangular form where each box has the same number of strings attached to the top and bottom. We then get one factor of $\delta^{\pm 1}$ for each critical point which is shaded above, and the power of δ corresponds to the sign of the critical point:

 $\longleftrightarrow \delta,$  $\longleftrightarrow \delta^{-1}.$

Correction factors for the lopsided convention for the Fourier transform and the trace were worked out in [Morrison and Penneys 2015b, Examples 2.6 and 2.7], and we derive another correction factor in the next example.

Example 2.20. We find the correction factors for the lopsided convention when calculating $\Delta(P, Q, R \mid S)$. We have

$$\Delta(P, Q, R \mid S) = \mathrm{Tr}$$



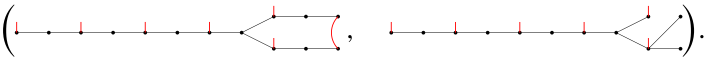
,

where the shading assumes n is even. The above diagram contributes a factor of δ^{-1}, and the trace tangle contributes no factors of δ. When n is odd, the above diagram contributes a factor of δ, and the trace tangle contributes a factor of δ. (See [Morrison and Penneys 2015b, Example 2.6] as well.) Hence we have the formula

$$\Delta(P, Q, R \mid S) = \natural\Delta(P, Q, R \mid S) = \begin{cases} \delta^{-1}\Delta(\natural P, \natural Q, \natural R \mid \natural S) & \text{if } n \text{ is even,} \\ \delta^{2}\Delta(\natural P, \natural Q, \natural R \mid \natural S) & \text{if } n \text{ is odd.} \end{cases}$$

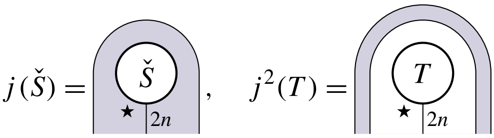
Assumption 2.21. For the rest of the article, we assume that for all $P, Q, R, S \in \mathfrak{B}$, the tetrahedral structure constants $\Delta(P, Q, R \mid S)$ are scalar multiples of the empty diagram.

2C. ***The jellyfish algorithm and reduced trains.*** The *jellyfish algorithm* was invented in [BMPS 2012] to construct the extended Haagerup subfactor planar algebra with principal graphs

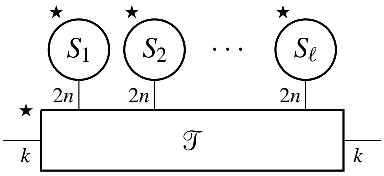


One uses the jellyfish algorithm to evaluate closed diagrams on a set of minimal generators. There are two ingredients:

(i) The generators in $\mathfrak{B} \subset \mathcal{P}_{n,+}$ satisfy *jellyfish relations*, i.e., for each generator S, T,

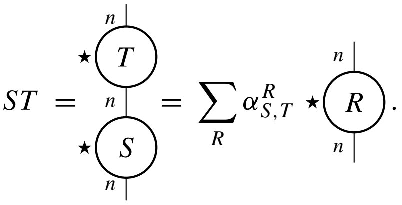


can be written as linear combinations of *trains*. Trains are diagrams where any region meeting the distinguished interval of a generator meets the distinguished interval of the external disk, i.e.,



where $S_1, \ldots, S_\ell \in \mathfrak{B}$, and $\mathcal{T}$ is a single Temperley–Lieb diagram.

(ii) The generators in $\mathfrak{B}$ are uncappable and together with the Jones–Wenzl projection $f^{(n)}$ form an algebra under the usual multiplication



(Note that the Mathematica package *FusionAtlas* also multiplies in this order; reading from left to right in products corresponds to reading from bottom to top in planar composites.)

Given these two ingredients, one can evaluate any closed diagram using the following two step process.

(i) Pull all generators S to the outside of the diagram using the jellyfish relations, possibly getting diagrams with more S's.

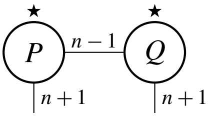
(ii) Use uncappability and the algebra property to iteratively reduce the number of generators. Any nonzero train which is a closed diagram is either a Temperley–Lieb diagram, has a capped generator, or has two generators S, T connected by at least n strings, giving ST. Each of these cases can be simplified using the relations, still giving a linear combination of trains.

Section 4 is devoted to our procedure for computing the jellyfish relations necessary for the first part of the jellyfish algorithm. The second part is rather easy, and amounts to verifying equation (1) (see the beginning of Section 5).

Definition 2.22. A $\mathfrak{B}$-train is called *reduced* if no two generators are connected by more than $n-1$ strands, and no generator is connected to itself.

Example 2.23. In $\mathcal{P}_{n+1,+}$, the set of reduced trains is given by

$$\left\{ P \underset{n-1}{\circ} Q = \right.$$

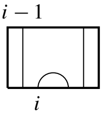


$$\left. \middle| \; P, Q \in \mathfrak{B} \right\}.$$

To describe the reduced trains in $\mathcal{P}_{n+2,+}$, we introduce the following notation.

Definition 2.24. Let $C_i[P \circ_{n-1} Q] \in \mathcal{P}_{n+2,+}$ for $i = 1, \dots, 2n+3$ be the reduced train obtained from $P \circ_{n-1} Q$ by putting C_i underneath, where C_i is the diagram given by

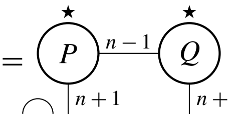
$$C_i =$$



.

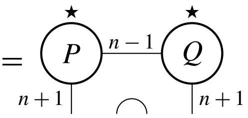
This can be thought of as multiplying C_i by $P \circ_{n-1} Q$ for a fixed arrangement of boundary strings. For example, we have, for $P, Q \in \mathfrak{B}$,

$$C_1[P \underset{n-1}{\circ} Q] =$$



and

$$C_{n+2}[P \underset{n-1}{\circ} Q] =$$

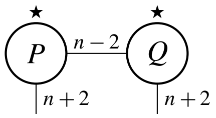


.

Example 2.25. In $\mathcal{P}_{n+2,+}$, we have many more reduced trains. First, we have those annular consequences of the $P \circ_{n-1} Q$'s which are still trains in $\mathcal{P}_{n+2,+}$. These are exactly the $C_i[P \circ_{n-1} Q]$ for $i = 1, \dots, 2n+3$.

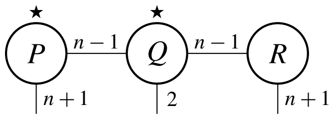
Now the only reduced trains which are nonzero when we put a copy of $f^{(2n+4)}$ underneath, for $P, Q, R \in \mathfrak{B}$, are

$$P \underset{n-2}{\circ} Q =$$



and

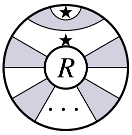
$$P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R =$$



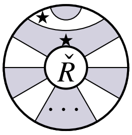
.

2D. ***The second annular basis.*** Given a nonzero low-weight rotational eigenvector $R \in \mathcal{P}_{n,+}$, the space $\mathfrak{A}_{n+2}(R) \subset \mathcal{P}_{n+2,+}$ of second annular consequences of R is spanned by diagrams with two cups on the outer boundary. We now describe a distinguished basis of $\mathfrak{A}_{n+2}(R)$ when $\delta > 2$ along the lines of [Jones 2001; 2012].

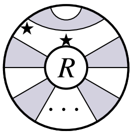
Definition 2.26. The element $\cup_{i,j}(R) \in \mathfrak{A}_{n+2}(R)$ is the annular consequence of R given in the following diagrams, where each row consists of $2n+4$ elements.



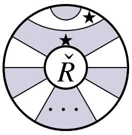
, $\cup_{-1,-1}(R)$



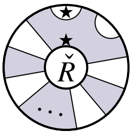
, $\cup_{-1,0}(R)$



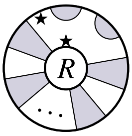
, ..., $\cup_{-1,1}(R)$



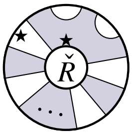
, $\cup_{-1,2n+2}(R)$



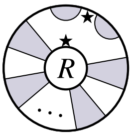
, $\cup_{0,0}(R)$



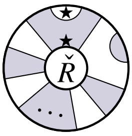
, $\cup_{0,1}(R)$



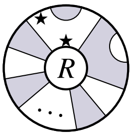
, ..., $\cup_{0,2}(R)$



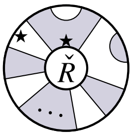
, $\cup_{0,2n+3}(R)$



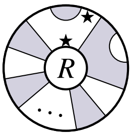
, $\cup_{1,0}(R)$



, $\cup_{1,1}(R)$



, ..., $\cup_{1,2}(R)$



, $\cup_{1,2n+3}(R)$

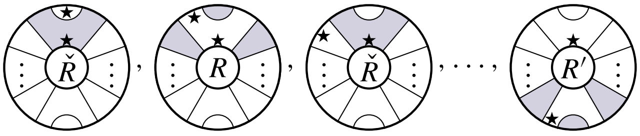
⋮ ⋮ ⋮ ⋮

The index i specifies the number of through strings separating the two cups (counting clockwise from the cup at 12 o'clock in the above diagrams). Here $i = -1$ denotes two nested cups. The j refers to the number of strings separating the external boundary interval at 12 o'clock from the interval for the external $\star$, counting counterclockwise (and subtract 1 for nested cups). Note that $n+k$ strings separating the cups is the same as a rotation (up to switching the shading) of $n-k$ strings separating the cups.

The *second annular basis* of $\mathfrak{A}_{n+2}(R)$ the set of $\cup_{i,j}(R)$ such that $-1 \le i \le n$, and

$$j \in \begin{cases} \{-1, 0, \dots, 2n+2\} & \text{if } i = -1, \\ \{0, \dots, 2n+3\} & \text{if } -1 < i < n, \\ \{0, \dots, n+1\} & \text{if } i = n. \end{cases}$$

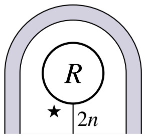
If $i = n$, the $n+2$ elements corresponding to $j = 0, \dots, n+1$ are given below. Here the shading on the bottom in the first three pictures depends on the parity of n, while the shading on the top of the final picture depends on the parity of n and whether R' is R or $\check{R}$.



$\cup_{n,0}(R)$ $\cup_{n,1}(R)$ $\cup_{n,2}(R)$ $\cup_{n,n+1}(R)$

Remark 2.27. Note that

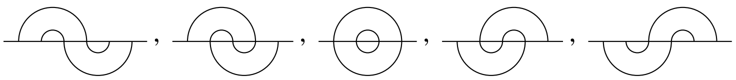
$$\cup_{-1,-1}(R) = j^2(R) =$$



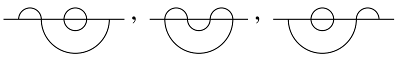
Recall that the inner product is defined by $\langle x, y\rangle = \mathrm{Tr}(x^* y)$, which is the same as connecting all strings of x^* and y. Computing inner products amongst the $\cup_{i,j}(R)$'s amounts to examining the relative positions of caps along the interface between the two diagrams. Since R is uncappable, the entire diagram is zero if a cap from one of the $\cup_{i,j}(R)$'s reaches the other copy of R.

It is easy to see that pairing $\cup_{i,j}(R)$ with $\cup_{i',k}(R)$ is nonzero only if $|i - i'| < 3$. When the scalar is nonzero differs for the cases $i = -1$ and $i \geq 0$, and there are some exceptional cases when $i = n - 1, n$.

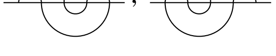
- When $i = -1$, there are exactly 5, 3, and 2 ways of getting a nonzero scalar when pairing $\cup_{-1,j}(R)$ with $\cup_{i',k}(R)$ for $i' = -1$, 0, and 1 respectively. They correspond to the following relative positions of caps along the interface.



$i' = -1,$

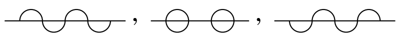


$i' = 0,$

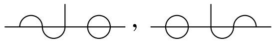


$i' = 1.$

- For $0 \leq i \leq n-2$, there are exactly 3, 2, and 1 ways of getting a nonzero scalar when pairing $\cup_{i,j}(R)$ with $\cup_{i',k}(R)$ for $i' = i$, $i + 1$, and $i + 2$ respectively. The relative positions of caps corresponding to the case $i = 0$ are.



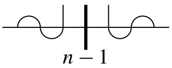
$i' = 0,$



$i' = 1,$

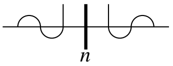
$i' = 2.$

- For $i = n - 1$, there is an additional way of getting a nonzero scalar when pairing $\cup_{n-1,j}(R)$ with $\cup_{n-1,k}(R)$, which makes up for the fact that there is no $\cup_{n+1,k}(R)$. The relative position of caps given by



can be interpreted as $(j-k) \equiv -1$ or $n+2 \bmod (2n+4)$, depending on the location of the $\star$ above the line. In the former case, the diagram contributes σ^{-1}, and in the latter, σ^{n}.

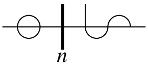
- The case $i=n$ is more subtle. When $i'=n-2$, there are two ways of pairing $\cup_{n,j}(R)$ with $\cup_{n-2,k}(R)$ to get a nonzero scalar, which correspond to the $\star$ placement of

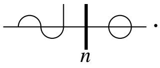
  ,

  i.e., $(j-k) \equiv -1$ or $n+1 \bmod (2n+4)$. In the former case, the diagram contributes a scalar of σ^{-1}, and in the latter, $\sigma^{n}\sigma^{-1}$.

  When $i'=n-1$, there are four ways to get a nonzero scalar, which correspond to the $\star$ placement of

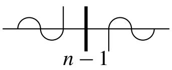
  

  ,

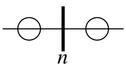
  

  .

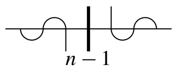
  Finally, when $i'=n$, there are three ways to get a nonzero scalar, corresponding to

  ,

  ,

  .

  (Note that the $\star$ placement is determined.)

The following proposition now follows from the above discussion.

Proposition 2.28. *Assuming $R=R^{*}$ and $\|R\|^2=\mathrm{Tr}(R^2)=1$, we have the following inner products* (*linear on the* right):

$$\langle \cup_{i',k}(R), \cup_{-1,j}(R)\rangle = \begin{array}{c|c|ccccc} & & \multicolumn{5}{c}{(j-k) \mod (2n+4)} \\ \hline & & -2 & -1 & 0 & 1 & 2 \\ \hline & -1 & \omega_R^{-1} & \sigma_R^{-1} & [2]^2 & \sigma_R & \omega_R \\ i' & 0 & 0 & [2]\sigma_R^{-1} & [2] & [2]\sigma_R & 0 \\ & 1 & 0 & 0 & 1 & \sigma_R & 0 \\ \hline \end{array}$$

and is zero otherwise.

For $0 \leq i', i \leq n-1$, *we have*

$$\langle \cup_{i',k}(R), \cup_{i,j}(R)\rangle = \begin{array}{c|c|ccc} & & \multicolumn{3}{c}{(j-k) \bmod (2n+4)} \\ \cline{3-5} & & -1 & 0 & 1 \\ \hline & -2 & \sigma_R^{-1} & 0 & 0 \\ & -1 & [2]\sigma_R^{-1} & [2] & 0 \\ i'-i & 0 & \sigma_R^{-1} & [2]^2 & \sigma_R \\ & 1 & 0 & [2] & [2]\sigma_R \\ & 2 & 0 & 0 & \sigma_R \\ \hline \end{array}$$

and is zero otherwise, with the exception that

$$\langle \cup_{n-1,k}(R), \cup_{n-1,j}\rangle = \sigma_R^n \quad \text{if } j-k \equiv n+2 \mod 2n+4.$$

For $i = n$ *and* $i' < n$, *we have*

$$\langle \cup_{i',k}(R), \cup_{n-1,j}(R)\rangle = \begin{array}{c|c|cccc} & & \multicolumn{4}{c}{(j-k) \bmod (2n+4)} \\ \cline{3-6} & & -1 & 0 & n+1 & n+2 \\ \hline i' & n-2 & \sigma_R^{-1} & 0 & \sigma_R^n\sigma_R^{-1} & 0 \\ & n-1 & [2]\sigma_R^{-1} & [2] & [2]\sigma_R^n\sigma_R^{-1} & [2]\sigma_R^n \\ \hline \end{array}$$

and is zero otherwise.

Finally, if $i = i' = n$, *then we have*

$$\langle \cup_{n,k}(R), \cup_{n,j}(R)\rangle = \begin{cases} \sigma_R^N\sigma_R^{-1} & \text{if } (j-k) \equiv -1 \mod (n+2) \text{ and } j = n+1, \\ \sigma_R^{-1} & \text{if } (j-k) \equiv -1 \mod (n+2) \text{ and } j < n+1, \\ [2]^2 & \text{if } (j-k) \equiv 0 \mod (n+2), \\ \sigma_R & \text{if } (j-k) \equiv 1 \mod (n+2) \text{ and } j > 0, \\ \sigma_R^n\sigma_R & \text{if } (j-k) \equiv -1 \mod (n+2) \text{ and } j = 0, \\ 0 & \text{else.} \end{cases}$$

Remark 2.29. The concerned reader may wonder if we have missed a case or two amidst this muddle of indices. Be reassured that we have checked these inner products numerically for the generators of our example directly in the graph planar algebra. See Section 4D for more details.

Remark 2.30. In this article, we do not give a formula for the dual basis $\widehat{\cup}_{i,j}(R)$ in terms of the $\cup_{i,j}(R)$'s, i.e., the change of basis matrix from the annular basis to the dual annular basis. Instead, we find the dual annular basis for our examples by inverting the matrix of inner products given by Proposition 2.28.

As in [Morrison and Penneys 2015b, Remark 3.7], if W is the matrix of inner products of the $\cup_{i,j}(R)$'s, then the change of basis matrix from the column vectors representing the annular basis U to the column vectors representing the dual basis $\widehat{U}$ is $\overline{W^{-1}}$, i.e., $\overline{W^{-1}}U = \widehat{U}$. (The inner product is linear on the *right*.) If $\hat{c}$ is the row vector of coefficients in the dual basis for an annular consequence x, i.e., $x = \hat{c} \cdot \widehat{U}$, then the row vector of coefficients in the annular basis is given by $c = \hat{c}\, \overline{W^{-1}}$.

It would certainly be useful to have a general formula for the dual annular basis in terms of the annular basis. While such a computation is routine, it would be demanding, and we leave it for another time.

3. Projections and inner products of trains

As in the previous section, we continue to use Assumptions 2.9, 2.12, and 2.21.

To derive two-strand jellyfish relations, we need to analyze all reduced $\mathfrak{B}$-trains in $\mathcal{P}_{n+2,+}$, in particular their projections to $\mathcal{TL}_{n+2,+}$, their projections to the space of second annular consequence of $\mathfrak{B}$, and their pairwise inner products.

We express some projections to Temperley–Lieb and annular consequences in terms of dual bases. We will use the following formula repeatedly.

Remark 3.1. Suppose $\{v_1, \dots, v_k\} \subset V$ is a basis for the finite dimensional Hilbert space V. Let $\{\hat{v}_1, \dots, \hat{v}_k\}$ be the dual basis V, defined by $\langle \hat{v}_i, v_j \rangle = \delta_{i,j}$, where the inner product is linear on the right. If $x \in V$, we have $x = \sum_{i=1}^{k} \langle v_i, x \rangle \hat{v}_i$.

In what follows, P, Q, R, S, T are always elements of $\mathfrak{B}$. We will first need a few results about certain Temperley–Lieb dual basis elements.

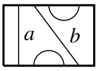
3A. *Some Temperley–Lieb dual basis elements.*

We now discuss certain elements of the basis which is dual to the usual diagrammatic basis of $\mathcal{TL}_k$.

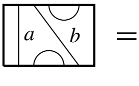
Lemma 3.2. *If* $a, b \geq 0$ *and* $a + b = n$, *then* $[a+2][b+1] - [a+1][b] = [n+2]$.

Proof. Immediate from the formula

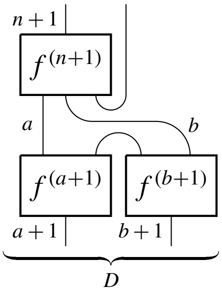
$$[k][\ell] \;=\; \sum_{\substack{|k-\ell|<j<k+\ell \\ j\equiv|k-\ell|+1 \mod 2}} [j]. \qquad \square$$

Lemma 3.3. *The element dual to* 

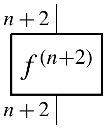
$\in \mathcal{TL}_{n+2,+}$ *is given by*



$$= \frac{[a+1][b+1]}{[n+2]^2}$$

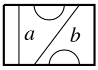


$$- \frac{(-1)^b[a+1]}{[n+2][n+3]}$$



.

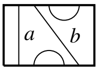
To find the element dual to



$$\in \mathcal{TL}_{n+2,+},$$

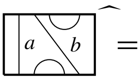
maintain the coefficients and vertically reflect the diagrams in the above formula.

Proof. Note that the middle diagram D in the above equation has nonzero inner product only with 1_{n+2} and

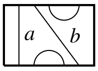


.

We already know that $\hat{1}_{n+2} = f^{(n+2)}/[n+3]$, so we have

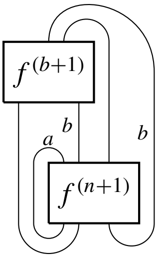


$$= \frac{1}{\langle D, \quad \rangle}\Big(D - \langle D, 1_{n+2}\rangle \frac{f^{(n+2)}}{[n+3]}\Big).$$



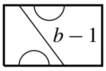
A routine calculation computes the necessary inner products. First,

$$\langle D, 1_{n+2}\rangle =$$



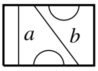
$$= \frac{(-1)^b[n+2]}{[b+1]},$$

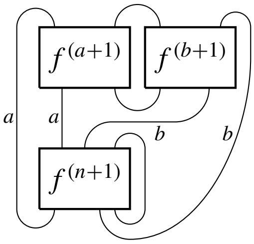
since the only diagram in the top $f^{(b+1)}$ which contributes to the closed diagram is



(the coefficient of this diagram in $f^{(b+1)}$ is given in Fact 2.16). Next, we calculate

(3) $$\Big\langle D, \quad \Big\rangle =$$

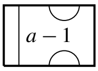




(4) $$= [n+2]\Big(\frac{[a+2]}{[a+1]} - \frac{[b]}{[b+1]}\Big)$$

(5) $$= \frac{[n+2]^2}{[a+1][b+1]},$$

where (4) follows since the only two terms in the top $f^{(a+1)}$ which contribute to the closed diagram are 1_{a+1} and



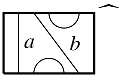
.

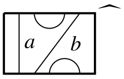
Equation (5) now follows by Lemma 3.2.

(Note that the value of the diagram that appears in (3) must be symmetric in a and b, but the quantity in (4) does not appear symmetric in a and b. This gave a hint that some quantum number identity should hold, which motivated Lemma 3.2.)

The last claim is now immediate. ☐

Lemma 3.4. *Suppose $a, b \geq 0$ with $a+b=n$. Let D_a, D_a^*, $\hat{1}_{n+2}$ be the Temperley–Lieb dual basis elements*

$D_a =$ 

, $D_a^* =$ 

, *and* $\hat{1}_{n+2} = \dfrac{f^{(n+2)}}{[n+3]}$.

$$\text{(i)}\ \langle C_i[P\circ Q], \hat{1}_{n+2}\rangle = \begin{cases} \mathrm{Tr}(PQ)[n+2]^{-1} & \textit{if } i=n+2, \\ 0 & \textit{else.} \end{cases}$$

$$\text{(ii)}\ \langle C_i[P\circ Q], D_a\rangle = \begin{cases} \mathrm{Tr}(PQ)[n+2]^{-1} & \textit{if } i-1=a, \\ 0 & \textit{if } i=n+2, \\ \dfrac{(-1)^b[a+1]}{[n+1][n+2]}\mathrm{Tr}(PQ) & \textit{if } i=n+3, \\ 0 & \textit{else.} \end{cases}$$

$$\text{(iii)}\ \langle C_i[P\circ Q], D_a^*\rangle = \langle D_a, C_{2n+4-i}[Q\circ P]\rangle = \langle C_{2n+4-i}[P\circ Q], D_a\rangle.$$

Proof. (i) We have

$$\langle C_i[P\circ Q], \hat{1}_{n+2}\rangle = \frac{1}{[n+3]}\langle C_i[P\circ Q], f^{(n+2)}\rangle,$$

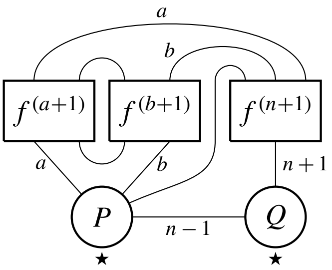
which is clearly zero if $i \neq n+2$. When $i=n+2$, it is easy to see we get $\dfrac{\mathrm{Tr}(PQ)}{[n+2]}$.

(ii) First, suppose $1 \leq i \leq n+1$. Then the inner product in question is given by

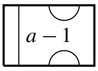
$$\frac{[a+1][b+1]}{[n+2]^2}\langle C_i[P\circ Q], D\rangle,$$

where D is the diagram in Lemma 3.3. If $i-1 \neq a$, then the resulting closed diagram is clearly zero. If $i-1=a$, then we have

$$\frac{[a+1][b+1]}{[n+2]^2}\langle C_i[P\circ Q],\, D\rangle =$$



and the only terms in the $f^{(a+1)}$ which contribute to the value are 1_{a+1} and



.

This yields, using Lemma 3.2 and Fact 2.16,

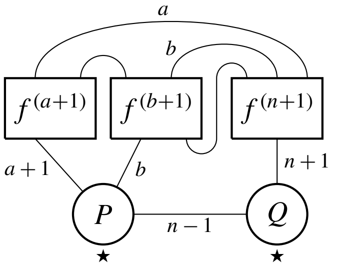
$$\frac{[a+1][b+1]}{[n+2]^2}\Big(\frac{[b+2]}{[b+1]}-\frac{[a]}{[a+1]}\Big)\operatorname{Tr}(PQ)=\frac{\operatorname{Tr}(PQ)}{[n+2]}.$$

Second, if $i=n+2$, then both diagrams in the formula for D_a from Lemma 3.3 contribute to the inner product, and we have

$$\begin{aligned}\langle C_{n+2}[P\circ Q],\, D_a\rangle\\ &=\frac{[a+1][b+1]}{[n+2]^2}\langle C_{n+2}[P\circ Q],\, D\rangle-\frac{(-1)^b[a+1]}{[n+2][n+3]}\langle C_{n+2}[P\circ Q],\, f^{(n+2)}\rangle\\ &=\frac{[a+1][b+1]}{[n+2]^2}\langle C_{n+2}[P\circ Q],\, D\rangle-\frac{(-1)^b[a+1]}{[n+2]^2}\operatorname{Tr}(PQ)\end{aligned}$$

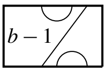
by part (i) of this lemma. Now by drawing diagrams, we get

$$\langle C_{n+2}[P\circ Q],\, D\rangle =$$



.

The only diagram in $f^{(b+1)}$ which contributes is



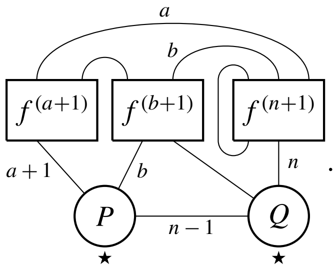
, which yields

$$\frac{(-1)^b}{[b+1]}\operatorname{Tr}(PQ).$$

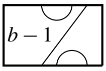
The inner product in question is thus zero.

Third, if $i = n + 3$, then as in the case $1 \le i \le n + 1$, we have

$$\frac{[a+1][b+1]}{[n+2]^2}\langle C_{n+3}[P \circ Q], D\rangle =$$



.

Again, the only diagram in $f^{(b+1)}$ which contributes is 

, which yields

$$\frac{[a+1][b+1]}{[n+1][n+2]}\left(\frac{(-1)^b}{[b+1]}\right) = \frac{(-1)^b[a+1]}{[n+1][n+2]}.$$

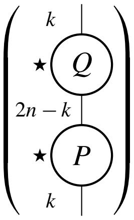
Finally, if $i > n + 3$, the result is once again zero, since both diagrams in the formula for D_a from Lemma 3.3 have zero inner product with $C_i[P \circ Q]$.

(iii) The first equality follows since both sides give the same closed diagram. Note that the quantity in the middle is equal to its conjugate by part (ii) of this lemma. The second equality now follows since $\mathrm{Tr}(QP) = \mathrm{Tr}(PQ)$. □

3B. ***Projections to Temperley–Lieb.*** The first lemma below is similar to [Jones 2012, Proposition 4.5.2].

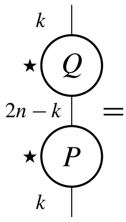
Lemma 3.5. (i) *If* $k = 0, \dots, 2n$, *then*

$$P_{\mathcal{TL}_{k,+}}$$



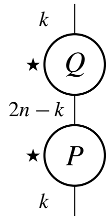
$$= \frac{\mathrm{Tr}(PQ)}{[k+1]} f^{(k)}.$$

(ii) *If* $k = 0, \dots, n-1$, *then*



$$= \frac{\mathrm{Tr}(PQ)}{[k+1]} f^{(k)}.$$

Proof. For (i), notice that adding a cap to the top or bottom of



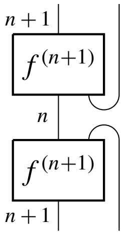
gives zero, so its projection to $\mathcal{TL}_{k,+}$ must be a constant times $f^{(k)}$. Taking traces gives the constant.

For (ii), notice that the diagram is already in Temperley–Lieb since $\mathfrak{B} \cup \{f^{(n)}\}$ spans an algebra. □

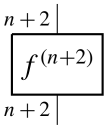
Proposition 3.6. (i) $P_{\mathcal{TL}_{n+2,+}}\big(P \underset{n-2}{\circ} Q\big) = \dfrac{\operatorname{Tr}(PQ)}{[n+3]} f^{(n+2)}$.

(ii) $P_{\mathcal{TL}_{n+2,+}}\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big)$

$$= a_R^{PQ}\Bigg(\frac{[n+1]}{[n+2]^2}$$



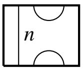
$$-\frac{[n+1]}{[n+2][n+3]}$$



$$\Bigg).$$

Proof. Part (i) is immediate from Lemma 3.5. For (ii), for T a diagrammatic basis element of $\mathcal{TL}_{n+2,+}$, it is clear that

$$\big\langle T, P_{\mathcal{TL}_{n+2,+}}\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big)\big\rangle = \begin{cases} a_R^{PQ} & \text{if } T = E_{n+1} = \\ 0 & \text{else.} \end{cases}$$



,

Hence $P_{\mathcal{TL}_{n+2,+}}(P \circ_{n-1} Q \circ_{n-1} R) = a_R^{PQ} \widehat{E}_{n+1}$, where $\widehat{E}_{n+1}$ is the dual basis element of E_{n+1} in $\mathcal{TL}_{n+2,+}$. The result now follows by Lemma 3.3, using $b = 0$, $a = n$. (In particular $[b] = 0$ and $[b+1] = 1$.) □

Proposition 3.7. $P_{\mathcal{TL}_{n+2,+}}(C_i[P \circ_{n-1} Q]) = \operatorname{Tr}(PQ)X$ *where X is a linear combination of Temperley–Lieb dual basis elements D_a, D_a^*, $\hat{1}_{n+2}$ (as in Lemma 3.4). The exact linear combination is given in the table below.*

i	X
1	$[2]D_0 + D_1$
$1 < i < n+1$	$D_{i-2} + [2]D_{i-1} + D_i$
$n+1$	$D_{n-1} + [2]D_n + \hat{1}_{n+2}$
$n+2$	$D_n + [2]\hat{1}_{n+2}$
$n+3$	$\hat{1}_{n+2} + [2]D_n^* + D_{n-1}^*$
$n+3 < i < 2n+3$	$D_{2n+2-i}^* + [2]D_{2n+3-i}^* + D_{2n+4-i}^*$
$2n+3$	$[2]D_0^* + D_1^*$

Proof. The only diagrammatic basis elements T in Temperley–Lieb which pair nontrivially with $C_i[P \circ_{n-1} Q]$ are those whose dual basis elements $\widehat{T}$ appear in the linear combination. The coefficients are given by $\langle T, C_i[P \circ_{n-1} Q]\rangle$. □

3C. *Projections to annular consequences.*

Definition 3.8. Let $\mathfrak{A}_{n+2}$ denote the space of second annular consequences of $\mathfrak{B}$ in $\mathcal{P}_{n+2,+}$.

The proofs of the following propositions are parallel to the proof of [Jones 2012, Proposition 4.4.1]. The inner products are only nonzero for the given annular consequences, and they are easily worked out by drawing pictures and using Lemma 3.5.

Proposition 3.9.

(i) $P_{\mathfrak{A}_{n+2}}\big(P \underset{n-2}{\circ} Q\big)$

$$= \sum_{R \in \mathfrak{B}} a_R^{PQ} \omega_P \omega_Q^{-1} \widehat{\cup}_{-1,-1}(R) + a_R^{PQ} \sigma_R^n \widehat{\cup}_{-1,n+1}(R) + b_R^{PQ} \sigma_P \sigma_Q^{-1} \widehat{\cup}_{n,0}(R)$$

where the coefficients of the $\widehat{\cup}_{i,j}(R)$ *are given by* $\langle \cup_{i,j}(R), P \circ_{n-2} Q\rangle$.

(ii) $P_{\mathfrak{A}_{n+2}}\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big)$

$$= \sum_{S \in \mathfrak{B}} \Delta_{n-1,2}(P, Q, R \mid S) \widehat{\cup}_{-1,-1}(S) + \frac{\sigma_S^{n+1}}{[n]} \operatorname{Tr}(SP) \operatorname{Tr}(QR) \widehat{\cup}_{-1,n}(S)$$
$$+ \frac{\sigma_S^{n-1}}{[n]} \operatorname{Tr}(PQ) \operatorname{Tr}(RS) \widehat{\cup}_{-1,n+2}(S) + \sigma_S^n \operatorname{Tr}(PQRS) \widehat{\cup}_{0,n+1}(S)$$
$$+ \Delta_{n-1,1}(P, Q, R \mid S) \widehat{\cup}_{n-1,0}(S) + \sigma_S^{n-1} \Delta_{n,1}(P, Q, R \mid S) \widehat{\cup}_{n-1,n+3}(S),$$

where the coefficients of the $\widehat{\cup}_{i,j}(S)$ *are given by* $\langle \cup_{i,j}(S), P \circ_{n-1} Q \circ_{n-1} R\rangle$.

Note that in the above formula, the quartic moment and two of the three tetrahedral constants were computed in terms of the moments and chiralities of $\mathfrak{B}$ *in Remark 2.11 and Example 2.17.*

Proposition 3.10. *$P_{\mathfrak{A}_{n+2}}(C_i[P \circ_{n-1} Q]) = \sum_{R\in\mathfrak{B}} X_R$ where X_R is given in the table below. Here we denote $\alpha = \sigma_R^n a_R^{PQ}$, $\beta = \sigma_Q^{-1}\sigma_P b_R^{PQ}$, and $\widehat{\mathrm{U}}_{i,j} = \widehat{\mathrm{U}}_{i,j}(R)$.*

i	$X_R \in \mathfrak{A}_{n+2}$
1	$\beta\widehat{\mathrm{U}}_{0,2n+2} + \alpha\widehat{\mathrm{U}}_{n-1,n+1} + [2]\beta\widehat{\mathrm{U}}_{-1,2n+2} + [2]\alpha\widehat{\mathrm{U}}_{n,n+1} + \sigma_R^{-1}a_R^{PQ}\widehat{\mathrm{U}}_{n-1,0}$
2	$\sigma_R\beta\widehat{\mathrm{U}}_{1,2n+1} + \alpha\widehat{\mathrm{U}}_{n-2,n+1} + [2]\beta\widehat{\mathrm{U}}_{0,2n+2} + [2]\alpha\widehat{\mathrm{U}}_{n-1,n+1} + \beta\widehat{\mathrm{U}}_{-1,2n+2} + \alpha\widehat{\mathrm{U}}_{n,n+1} + \sigma_R^{-1}\beta\widehat{\mathrm{U}}_{-1,-1}$
$2 < i < n+1$	$\sigma_R^{i-1}\beta\widehat{\mathrm{U}}_{i-1,2n-i+3} + \alpha\widehat{\mathrm{U}}_{n-i,n+1} + [2]\sigma_R^{i-2}\beta\widehat{\mathrm{U}}_{i-2,2n-i+4} + [2]\alpha\widehat{\mathrm{U}}_{n-i+1,n+1} + \sigma_R^{i-3}\beta\widehat{\mathrm{U}}_{i-3,2n-i+5} + \alpha\widehat{\mathrm{U}}_{n-i+2,n+1}$
$n+1$	$\beta\widehat{\mathrm{U}}_{n,0} + \alpha\widehat{\mathrm{U}}_{-1,n+1} + [2]\sigma_R^{n-1}\beta\widehat{\mathrm{U}}_{n-1,n+3} + [2]\alpha\widehat{\mathrm{U}}_{0,n+1} + \sigma_R^{n-2}\beta\widehat{\mathrm{U}}_{n-2,n+4} + \alpha\widehat{\mathrm{U}}_{1,n+1} + \sigma_R^{n+1}a_R^{PQ}\widehat{\mathrm{U}}_{-1,n}$
$n+2$	$\beta\widehat{\mathrm{U}}_{n-1,0} + \alpha\widehat{\mathrm{U}}_{0,n+1} + [2]\beta\widehat{\mathrm{U}}_{n,0} + [2]\alpha\widehat{\mathrm{U}}_{-1,n+1} + \sigma_R^{n-1}\beta\widehat{\mathrm{U}}_{n-1,n+3}$
$n+3$	$\beta\widehat{\mathrm{U}}_{n-2,0} + \sigma_R^{n+1}a_R^{PQ}\widehat{\mathrm{U}}_{1,n} + [2]\beta\widehat{\mathrm{U}}_{n-1,0} + [2]\alpha\widehat{\mathrm{U}}_{0,n+1} + \beta\widehat{\mathrm{U}}_{n,0} + \alpha\widehat{\mathrm{U}}_{-1,n+1} + \sigma_R^{n-1}a_R^{PQ}\widehat{\mathrm{U}}_{-1,n+2}$
$n+3 < i < 2n+2$	$\beta\widehat{\mathrm{U}}_{2n+1-i,0} + \sigma_R^{i-2}a_R^{PQ}\widehat{\mathrm{U}}_{i-n-2,2n+3-i} + [2]\beta\widehat{\mathrm{U}}_{2n+2-i,0} + [2]\sigma_R^{i-3}a_R^{PQ}\widehat{\mathrm{U}}_{i-n-3,2n+4-i} + \beta\widehat{\mathrm{U}}_{2n+3-i,0} + \sigma_R^{i-4}a_R^{PQ}\widehat{\mathrm{U}}_{i-n-4,2n+5-i}$
$2n+2$	$\beta\widehat{\mathrm{U}}_{-1,0} + a_R^{PQ}\widehat{\mathrm{U}}_{n,1} + [2]\beta\widehat{\mathrm{U}}_{0,0} + [2]\sigma_R^{-1}a_R^{PQ}\widehat{\mathrm{U}}_{n-1,2} + \beta\widehat{\mathrm{U}}_{1,0} + \sigma_R^{-2}a_R^{PQ}\widehat{\mathrm{U}}_{n-2,3} + \sigma_R\beta\widehat{\mathrm{U}}_{-1,-1}$
$2n+3$	$\alpha\widehat{\mathrm{U}}_{n-1,n+3} + [2]\beta\widehat{\mathrm{U}}_{-1,0} + [2]a_R^{PQ}\widehat{\mathrm{U}}_{n,1} + \beta\widehat{\mathrm{U}}_{0,0} + \sigma_R^{-1}a_R^{PQ}\widehat{\mathrm{U}}_{n-1,2}$

Remark 3.11. We check the formulas given in Propositions 3.9 and 3.10 by taking inner products directly in the graph planar algebra. See Section 4D for more details.

However, the best evidence that these formulas are correct is the fact that we can actually compute the two-strand jellyfish relations for the $3^{\mathbb{Z}/4\mathbb{Z}}$ subfactor planar algebra!

3D. *Inner products amongst trains and their projections.*

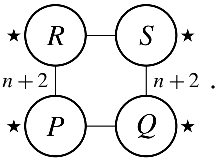
Proposition 3.12.

(i) $\left\langle P \underset{n-2}{\circ} Q,\, R \underset{n-2}{\circ} S \right\rangle = \dfrac{\mathrm{Tr}(PR)\,\mathrm{Tr}(SQ)}{[n-1]}.$

(ii) $\left\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R,\, P' \underset{n-1}{\circ} Q' \underset{n-1}{\circ} R' \right\rangle = \dfrac{\mathrm{Tr}(PP')\,\mathrm{Tr}(QQ')\,\mathrm{Tr}(RR')}{[n]^2}.$

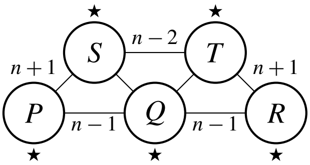
(iii) $\left\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R,\, S \underset{n-2}{\circ} T \right\rangle = 0.$

Proof. For (i), the left-hand side equals

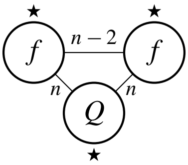


The result now follows by Lemma 3.5 (ii).

We omit the proof of (ii), which is similar to the proof of (i). For (iii), again using Lemma 3.5 (ii), we see that the left-hand side is equal to



$$= \frac{\mathrm{Tr}(PS)\,\mathrm{Tr}(RT)}{[n]^2}$$



where $f = f^{(n-1)}$. The right-hand side of the above equation is zero, since it is a linear combination of closed diagrams containing only one generator. □

Proposition 3.13.

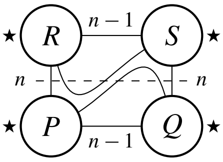
(i) $\left\langle C_i\left[P \underset{n-1}{\circ} Q\right], C_j\left[R \underset{n-1}{\circ} S\right]\right\rangle$

$$= \begin{cases} \mathrm{Tr}(PR)\,\mathrm{Tr}(SQ)[2][n]^{-1} & \textit{if } i = j, \\ \mathrm{Tr}(PR)\,\mathrm{Tr}(SQ)[n]^{-1} & \textit{if } |i-j| = 1, \\ \mathrm{Tr}(PRSQ) & \textit{if } (i,j) \in \{(n{+}1, n{+}3), (n{+}3, n{+}1)\}, \\ 0 & \textit{else.} \end{cases}$$

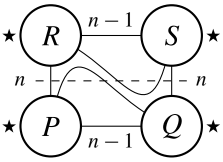
(ii) $\left\langle C_i\left[P \underset{n-1}{\circ} Q\right], R \underset{n-2}{\circ} S\right\rangle = \begin{cases} \mathrm{Tr}(PR)\,\mathrm{Tr}(SQ)[n]^{-1} & \textit{if } i = n+2, \\ 0 & \textit{else.} \end{cases}$

(iii) $\left\langle C_i\left[P \underset{n-1}{\circ} Q\right], R \underset{n-1}{\circ} S \underset{n-1}{\circ} T\right\rangle = \begin{cases} a_S^{PR}\,\mathrm{Tr}(QT)[n]^{-1} & \textit{if } i = n+1, \\ a_Q^{ST}\,\mathrm{Tr}(RP)[n]^{-1} & \textit{if } i = n+3, \\ 0 & \textit{else.} \end{cases}$

Proof. The proofs are all relatively straightforward drawing the necessary diagrams. The case in (i) which is easiest to miss is when $(i, j) \in \{(n{+}1, n{+}3), (n{+}3, n{+}1)\}$. In this case we get the following diagrams:



$=$



$= \mathrm{Tr}(PRSQ).$ □

Proposition 3.14. (i) $\left\langle P \underset{n-2}{\circ} Q,\, P_{\mathcal{TL}_{n+2,+}}\left(R \underset{n-2}{\circ} S\right)\right\rangle = \dfrac{\mathrm{Tr}(PQ)\,\mathrm{Tr}(RS)}{[n{+}3]}.$

(ii) $\big\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R,\, P_{\mathcal{TL}_{n+2,+}}\big(P' \underset{n-1}{\circ} Q' \underset{n-1}{\circ} R'\big)\big\rangle = a_R^{QP} a_{R'}^{P'Q'} \dfrac{[2][n+1]}{[n+2][n+3]}.$

(iii) $\big\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R,\, P_{\mathcal{TL}_{n+2,+}}\big(S \underset{n-2}{\circ} T\big)\big\rangle = -\dfrac{\operatorname{Tr}(ST)a_R^{QP}[n+1]}{[n+2][n+3]}.$

Proof. This follows quickly from Proposition 3.6. For part (ii), using Proposition 3.6, the inner product in question is equal to

$$\begin{aligned} a_R^{QP} a_{R'}^{P'Q'}\Big(\frac{[n+1]^2}{[n+2]^2[n+3]} + \frac{[n+1]}{[n+2]^2}\Big) &= a_R^{QP} a_{R'}^{P'Q'} \frac{[n+1]}{[n+2]^2}\Big(\frac{[n+1]+[n+3]}{[n+3]}\Big) \\ &= a_R^{QP} a_{R'}^{P'Q'} \frac{[2][n+1]}{[n+2][n+3]}. \end{aligned}$$ ☐

Proposition 3.15.

(i) $\big\langle C_i\big[P \underset{n-1}{\circ} Q\big], P_{\mathcal{TL}_{n+2,+}}\big(C_j\big[R \underset{n-1}{\circ} S\big]\big)\big\rangle$

$$= \begin{cases} \operatorname{Tr}(PQ)\operatorname{Tr}(RS)[2][n+2]^{-1} & \textit{if } i=j, \\ \operatorname{Tr}(PQ)\operatorname{Tr}(RS)[n+2]^{-1} & \textit{if } |i-j|=1, \\ \operatorname{Tr}(PQ)\operatorname{Tr}(RS)[n+1]^{-1} & \textit{if } (i,j)\in\{(n+1,n+3),(n+3,n+1)\}, \\ 0 & \textit{else.} \end{cases}$$

(ii) $\big\langle C_i\big[P \underset{n-2}{\circ} Q\big], P_{\mathcal{TL}_{n+2,+}}\big(R \underset{n-2}{\circ} S\big)\big\rangle = \begin{cases} \operatorname{Tr}(PQ)\operatorname{Tr}(RS)[n+2]^{-1} & \textit{if } i=n+2, \\ 0 & \textit{else.} \end{cases}$

(iii) $\big\langle C_i\big[P \underset{n-2}{\circ} Q\big], P_{\mathcal{TL}_{n+2,+}}\big(R \underset{n-1}{\circ} S \underset{n-1}{\circ} T\big)\big\rangle$

$$= \begin{cases} \operatorname{Tr}(PQ)a_T^{RS}[n+2]^{-1} & \textit{if } i=n+1, n+3, \\ 0 & \textit{else.} \end{cases}$$

Proof. (i) The formulas can be obtained easily from Lemma 3.4 and Proposition 3.7. We work out a few interesting cases.

If $i=n+1$ and $j=n+3$, then

$$\begin{aligned} &\big\langle C_i\big[P \underset{n-1}{\circ} Q\big], P_{\mathcal{TL}_{n+2,+}}\big(C_j\big[R \underset{n-1}{\circ} S\big]\big)\big\rangle \\ &\qquad = \big\langle C_{n+1}\big[P \underset{n-1}{\circ} Q\big], \hat{1}_{n+2} + [2]D_n^* + D_{n-1}^*\big\rangle \operatorname{Tr}(RS) \\ &\qquad = \big\langle C_{n+3}\big[P \underset{n-1}{\circ} Q\big], [2]D_n + D_{n-1}\big\rangle \operatorname{Tr}(RS) \\ &\qquad = \Big(\frac{[2]}{[n+2]} - \frac{[n]}{[n+1][n+2]}\Big)\operatorname{Tr}(PQ)\operatorname{Tr}(RS) \\ &\qquad = \frac{\operatorname{Tr}(PQ)\operatorname{Tr}(RS)}{[n+1]}. \end{aligned}$$

If $i = n+1$ and $n+3 < j < 2n+3$, then

$$\begin{aligned}
&\bigl\langle C_i\bigl[P \underset{n-1}{\circ} Q\bigr],\, P_{\mathcal{TL}_{n+2,+}}\bigl(C_j\bigl[R \underset{n-1}{\circ} S\bigr]\bigr)\bigr\rangle \\
&\qquad\qquad = \bigl\langle C_{n+1}\bigl[P \underset{n-1}{\circ} Q\bigr],\, D^*_{2n+2-j} + [2]D^*_{2n+3-j} + D^*_{2n+4-j}\bigr\rangle \operatorname{Tr}(RS) \\
&\qquad\qquad = \bigl\langle C_{n+3}\bigl[P \underset{n-1}{\circ} Q\bigr],\, D_j + [2]D_{j+1} + D_{j+2}\bigr\rangle \operatorname{Tr}(RS) \\
&\qquad\qquad = \frac{(-1)^{n-j}}{[n+1][n+2]}\bigl([j+1] - [2][j+2] + [j+3]\bigr) \operatorname{Tr}(PQ) \operatorname{Tr}(RS) \\
&\qquad\qquad = 0.
\end{aligned}$$

(ii) By Proposition 3.6, we have

$$\bigl\langle C_i\bigl[P \underset{n-1}{\circ} Q\bigr],\, P_{\mathcal{TL}_{n+2,+}}\bigl(R \underset{n-2}{\circ} S\bigr)\bigr\rangle = \frac{\operatorname{Tr}(RS)}{[n+3]}\bigl\langle C_i\bigl[P \underset{n-1}{\circ} Q\bigr],\, f^{(n+2)}\bigr\rangle,$$

which is zero unless $i = n+2$. Now by Lemma 3.3 and Proposition 3.7, the right-hand side is equal to

$$\begin{aligned}
\frac{\operatorname{Tr}(RS)\operatorname{Tr}(PQ)}{[n+3]}\langle D_n + [2]\hat{1}_{n+2},\, f^{(n+2)}\rangle &= \frac{\operatorname{Tr}(PQ)\operatorname{Tr}(RS)}{[n+3]}\Bigl([2] - \frac{[n+1]}{[n+2]}\Bigr) \\
&= \frac{\operatorname{Tr}(PQ)\operatorname{Tr}(RS)}{[n+2]}.
\end{aligned}$$

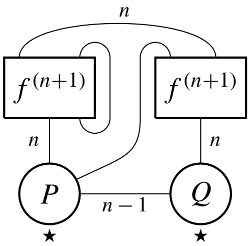
(iii) By Proposition 3.6, we have

$$\bigl\langle C_i\bigl[P \underset{n-1}{\circ} Q\bigr],\, P_{\mathcal{TL}_{n+2,+}}\bigl(R \underset{n-2}{\circ} S \underset{n-1}{\circ} T\bigr)\bigr\rangle = a_T^{RS}\bigl\langle C_i\bigl[P \underset{n-1}{\circ} Q\bigr],\, \widehat{E}_{n+1}\bigr\rangle,$$

which is clearly zero unless $n+1 \le i \le n+3$ (use the formula for $\widehat{E}_{n+1}$).

If $i = n+1$ (and similarly for $i = n+3$), then only the first diagram in Proposition 3.6 (ii) contributes to the inner product, and the value is given by

$$a_T^{RS}\frac{[n+1]}{[n+2]^2}$$



$$= \frac{\operatorname{Tr}(PQ)a_T^{RS}}{[n+2]}.$$

If $i = n+2$, by drawing similar diagrams, we see the inner product in question is equal to

$$a_T^{RS}\Bigl(\frac{[n+1]}{[n+2]^2} - \frac{[n+3][n+1]}{[n+2][n+2][n+3]}\Bigr)\operatorname{Tr}(PQ) = 0. \qquad \square$$

Remark 3.16. We now explain how to obtain the inner products

- $\langle P \underset{n-2}{\circ} Q, P_{\mathfrak{A}_{n+2}}(R \underset{n-2}{\circ} S)\rangle$,
- $\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R, P_{\mathfrak{A}_{n+2}}(P' \underset{n-1}{\circ} Q' \underset{n-1}{\circ} R')\rangle$,
- $\langle P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R, P_{\mathfrak{A}_{n+2}}(S \underset{n-2}{\circ} T)\rangle$,
- $\langle C_i[P \underset{n-2}{\circ} Q], P_{\mathfrak{A}_{n+2}}(C_j[R \underset{n-2}{\circ} S])\rangle$,
- $\langle C_i[P \underset{n-2}{\circ} Q], P_{\mathfrak{A}_{n+2}}(R \underset{n-1}{\circ} S \underset{n-1}{\circ} T)\rangle$,
- $\langle C_i[P \underset{n-2}{\circ} Q], P_{\mathfrak{A}_{n+2}}(R \underset{n-2}{\circ} S)\rangle$.

First, we use the formulas for

$$P_{\mathfrak{A}_{n+2}}(P \underset{n-2}{\circ} Q), \quad P_{\mathfrak{A}_{n+2}}(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R), \quad \text{and} \quad P_{\mathfrak{A}_{n+2}}(C_i[P \underset{n-2}{\circ} Q])$$

obtained in Propositions 3.9 and 3.10 to express each side as a linear combination of the $\widehat{\cup}_{i,j}(S)$'s. Next, we use the change of basis matrix discussed in Remark 2.30 to write the $\widehat{\cup}_{i,j}(S)$ on the right-hand side in terms of the $\cup_{i,j}(S)$. Finally, we expand the inner product in the usual way to obtain the answer.

4. Deriving formulas for two-strand box jellyfish relations

As in the previous sections, we continue Assumptions 2.9, 2.12, and 2.21.

We now go through our algorithm for determining two-strand jellyfish relations. We follow the method of [Morrison and Penneys 2015b, Section 3], which consists of three parts:

(i) Find the quadratic tangles in annular consequences.
(ii) Find the jellyfish matrix.
(iii) Invert the jellyfish matrix.

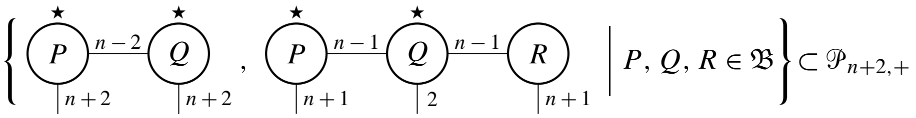
The steps in our algorithm will be clearly marked in the following three subsections.

4A. ***Reduced trains in annular consequences.*** In [Morrison and Penneys 2015b], the first step was to obtain a basis for the quadratic tangles in annular consequences. Since we have quadratic and cubic trains, we call this step obtaining a basis for the reduced trains in annular consequences.

Definition 4.1. Recall from Definition 2.22 that a "reduced train" is one where no generator connects to itself, and no pair are connected by more than $n-1$ strands. Starting with our set of minimal generators $\mathfrak{B}$ satisfying Assumptions 2.9, 2.12, and 2.21, we have the reduced trains

$$\{C_i[P \underset{n-1}{\circ} Q] \mid P, Q \in \mathfrak{B} \text{ and } i = 1, \ldots, 2n+3\} \subset \mathcal{P}_{n+2,+}$$

which are annular consequences of trains in $\mathcal{P}_{n+1,+}$, and we have the reduced trains



which are nonzero when placing a Jones–Wenzl underneath. We let $\mathcal{RT}$ be the union of the above two sets.

Since we hope that our generators generate a subfactor planar algebra with the desired principal graph, we want some linear combination of these reduced trains to lie in annular consequences.

Definition 4.2. We set

$$\mathcal{RTAC} = (\mathcal{TL}_{n+2,+} \oplus \mathfrak{A}_{n+2}) \cap \operatorname{span}(\mathcal{RT}),$$

where $\mathcal{RTAC}$ stands for *reduced trains in annular consequences*.

Step 1 of our algorithm finds a basis for $\mathcal{RTAC}$. Since we are trying to derive box jellyfish relations, we are only interested in basis elements which are not sent to zero when we put a $f^{(2n+4)}$ underneath. Thus we make the following definition.

Definition 4.3. An element of $\mathcal{RTAC}$ is called *essential* if at least one of the coefficients of the $P \circ_{n-2} Q$'s or the $P \circ_{n-1} Q \circ_{n-1} R$'s does not vanish.

Remark 4.4. If we've chosen k generators in a graph planar algebra and are hoping that they give us a subfactor planar algebra with one spoke principal graph, we expect to have at least k essential basis elements of $\mathcal{RTAC}$, i.e., one two-strand jellyfish relation for each generator.

Step 1 (a basis for $\mathcal{RTAC}$). Consider the matrix

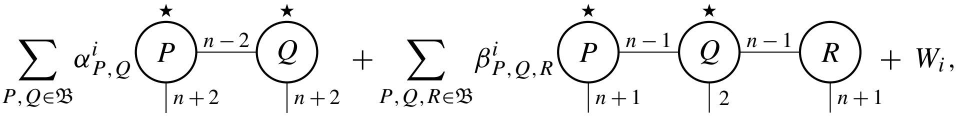
$$\big(\langle \mathcal{X} - P_{\mathcal{TL}_{n+2,+}}(\mathcal{X}) - P_{\mathfrak{A}_{n+2}}(\mathcal{X}), \mathcal{Y}\rangle\big)_{\mathcal{X},\mathcal{Y}\in\mathcal{RT}},$$

of inner products modulo Temperley–Lieb and annular consequences. (Note that the necessary inner products were derived in Propositions 3.12 and 3.14 and Remark 3.16.)

(i) Taking a basis for the null space of this matrix gives us a basis for $\mathcal{RTAC}$.

(ii) From this basis, we keep only the essential elements, which we call $X_1, \ldots, X_k$.

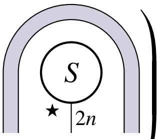
4B. ***Compute the jellyfish matrix.*** From Step 1, we have an expression for each essential basis element of $\mathcal{RTAC}$. Namely, the basis elements X_i can be written in the form



where $W_i \in \operatorname{span}\{C_i[P \circ Q] \mid P, Q \in \mathfrak{B} \text{ and } i = 1, \dots, 2n+3\}$. We also have an expression for X_i as an element of $\mathcal{TL}_{n+2,+} \oplus \mathfrak{A}_{n+2}$.

Step 2 (expression in the annular basis). Using Proposition 3.9, we express the X_i in terms of the dual annular basis $\widehat{\cup}_{r,s}(S)$ for $S \in \mathfrak{B}$. We then use the change of basis matrix discussed in Remark 2.30 to write the $\widehat{\cup}_{r,s}(S)$ in terms of the $\cup_{j,\ell}(S)$. Hence we may write each X_i as

$$X_i = \Big(\sum_{S \in \mathfrak{B}} \gamma_S^i \cup_{-1,-1}(S)\Big) + Y_i + Z_i = \Bigg(\sum_{S \in \mathfrak{B}} \gamma_S^i$$

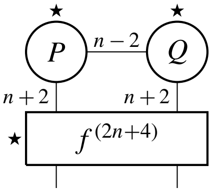


$$\Bigg) + Y_i + Z_i,$$

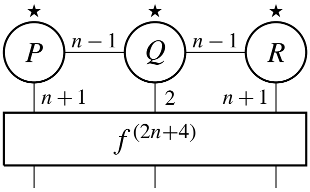
where Y_i is a linear combination of the $\cup_{j,\ell}(S)$ for $S \in \mathfrak{B}$ and $(j, \ell) \neq (-1, -1)$, and $Z_i \in \mathcal{TL}_{n+2,+}$.

Notation 4.5. For $P, Q, R, S \in \mathfrak{B}$, we use the notation

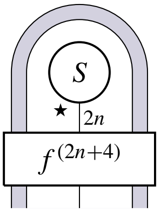
$$f\big(P \underset{n-2}{\circ} Q\big) =$$



$$f\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big) =$$



$$f \cdot j^2(S) =$$



We also write $f \cdot X$ to denote $X \in \mathcal{P}_{n+2,+}$ in jellyfish form with a $f^{(2n+4)}$ underneath.

Step 3 (box jellyfish equations). Put an $f^{(2n+4)}$ underneath the two formulas for X_i obtained in Steps 1 and 2 to get the following equations for $i = 1, \dots, k$:

$$f \cdot X_i = \sum_{P,Q \in \mathfrak{B}} \alpha_{P,Q}^i f\big(P \underset{n-2}{\circ} Q\big) + \sum_{P,Q,R \in \mathfrak{B}} \beta_{P,Q,R}^i f\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big) = \sum_{S \in \mathfrak{B}} \gamma_S^i f \cdot j^2(S).$$

Remark 4.6. In [Morrison and Penneys 2015b, Section 3.2], similar formulas to those obtained in Step 3 were checked by wrapping a Jones–Wenzl around the top of $P \circ_{n-1} Q$. In our case, we cannot use this check, since wrapping a Jones–Wenzl around the top of a 3-train does not give another box-train.

We now define the jellyfish matrix and the reduced trains matrix from the equations from Step 3.

Definition 4.7. The *two-strand jellyfish matrix* is the matrix J_2 whose i-th row is $(\gamma_S^i)_{S\in\mathfrak{B}}$. The *reduced trains matrix* is the matrix K_2 whose i-th row is given by concatenating the lists $(\alpha^i_{P,Q})_{P,Q\in\mathfrak{B}}$ and $(\beta^i_{R,S,T})_{R,S,T\in\mathfrak{B}}$.

Remark 4.8. Note that

$$K_2 \begin{pmatrix} f\big(P \underset{n-2}{\circ} Q\big) \\ \vdots \\ f\big(R \underset{n-1}{\circ} S \underset{n-1}{\circ} T\big) \\ \vdots \end{pmatrix}_{P,Q,R,S,T\in\mathfrak{B}} = J_2 \begin{pmatrix} f\cdot j^2(S) \\ \vdots \end{pmatrix}_{S\in\mathfrak{B}}.$$

4C. *Invert the jellyfish matrix.* At this point, we have accomplished most of the difficult work. Two easy steps remain.

Step 4 (invert J_2). Given the matrix J_2 from Definition 4.7 obtained via Step 3, we check if it has rank $|\mathfrak{B}|$. If it does (and we know that it should by [Bigelow and Penneys 2014]), we find a left inverse for J_2 by the formula

$$J_2^L = (J_2^* J_2)^{-1} J_2^*$$

since J_2 and $J_2^* J_2$ have the same rank.

Step 5 (box jellyfish relations). Finally, we get the *box jellyfish relations* by multiplying by J_2^L from Step 4:

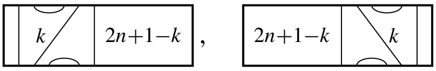
$$\begin{pmatrix} f\cdot j^2(S) \\ \vdots \end{pmatrix}_{S\in\mathfrak{B}} = J_2^L K_2 \begin{pmatrix} f\big(P \underset{n-2}{\circ} Q\big) \\ \vdots \\ f\big(P \underset{n-1}{\circ} Q \underset{n-1}{\circ} R\big) \\ \vdots \end{pmatrix}_{P,Q,R\in\mathfrak{B}}$$

which express the $f \cdot j^2(S)$ as linear combinations of reduced trains.

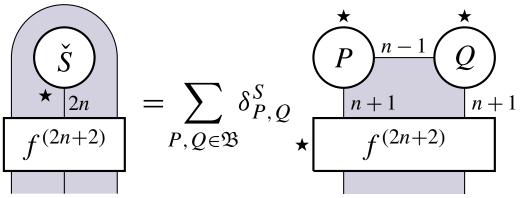
Remark 4.9. Recall that our goal was to derive two-strand jellyfish relations for our generators. These relations would be sufficient to evaluate all closed diagrams. Note that two-strand box jellyfish relations by themselves are *not* sufficient to evaluate closed diagrams!

In order to recover jellyfish relations from box jellyfish relations, we need to expand the Jones–Wenzl idempotents as in [Morrison and Penneys 2015b, Section 2.5].

When expanding $f^{(2n+4)}$ for the two-strand box jellyfish relations, terms of the form



in $f^{(2n+4)}$ yield diagrams not in jellyfish form, as they have a strand separating the generator from the outer region. Hence we also need one-strand jellyfish relations, which are obtained from one-strand box jellyfish relations of the form

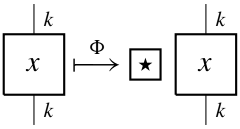


by the argument in [Morrison and Penneys 2015b, Section 2.5]. We compute the necessary one-strand box jellyfish relations using the algorithm provided there.

4D. ***Checking our calculations.*** Since the computer is doing all the arithmetic, it is good to check that our formulas are consistent with other methods of calculation. The computations in this section are redundant, hence we freely take shortcuts and perform spot checks when more thorough checks would be too time consuming.

The checks we perform in this subsection are done directly in the graph planar algebra. As such computations are computationally expensive, we use the following shortcut, which is known to experts. We do not prove it here as it would take us too far afield.

Proposition 4.10. *Suppose $\mathcal{P}_\bullet$ is a subfactor planar algebra. Choose an embedding of $\mathcal{P}_\bullet$ into $\mathcal{GPA}(\Gamma_+)_\bullet$, the graph planar algebra of its principal graph, and identify $\mathcal{P}_\bullet$ with its image. Define the map $\Phi : \mathcal{P}_{k,\pm} \to \mathcal{GPA}(\Gamma_+)_{k,\pm}$ by cutting down at the zero box $\star$ (the distinguished vertex of Γ_+), i.e., forgetting all loops of length $2k$ which do not start at $\star$.*



Then Φ is a $*$-algebra isomorphism under the usual multiplication, and Φ commutes with taking (partial) traces.

We remark that $\dim(\mathcal{P}_{k,\pm})$ is equal to the number of loops of length $2k$ starting at $\star$ on the principal graph, so one only needs to prove this map is injective.

To simplify calculations in the graph planar algebra, we can compute the inner product by first cutting down at $\star$ and then taking the inner product of the cut down elements in the graph planar algebra. Note that this simplification assumes we are working in the image of a subfactor planar algebra, so it cannot be used to prove that formulas hold. However, it can be used as a check for our calculations.

Using this shortcut, we check the propositions listed in the following table. The calculations are performed in the notebook `TwoStrandJellyfish.nb` in subsections called "Checking directly in the GPA" for each of our examples. Many of the computations are exact, but two are numerical. For the checks for Propositions 3.9 and 3.10, we don't check all the coefficients in the graph planar algebra; rather we only check the coefficients that our formulas tell us are nonzero.

Proposition	Checking functions	Numerical?
2.28	`CheckPairwiseInnerProductsOfSecondAC`	Yes
3.9	`CheckCoefficientsOf2TrainsInSecondAC` `CheckCoefficientsOf3TrainsInSecondAC`	No
3.10	`CheckCoefficientsOfCiQTCircsInSecondAC`	No
3.12	`CheckInnerProductBetweenTrains`	No
3.13	`CheckInnerProductWithCiQTCircs`	Yes

As a verification of the correctness of our algorithm, we also reproved the existence of the Haagerup $3^{\mathbb{Z}/3\mathbb{Z}}$ subfactor and the $3^{\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}}$ subfactor. We did not include these calculations since there are already several proofs for existence of these subfactors. We note that the formula we obtain for the two-strand jellyfish relation for $3^{\mathbb{Z}/3}$ (Haagerup) agrees with that obtained in [BMPS 2012]. We have not checked that our two-strand relations for $3^{\mathbb{Z}/2\times\mathbb{Z}/2}$ are consistent with the one-strand relations found in [Morrison and Penneys 2015b], since we use different generators.

In [Morrison and Penneys 2015b], the authors were able to check the one-strand jellyfish relations for 2221 directly in the graph planar algebra using a clever trick due to Bigelow. We cannot do these computations for our graphs. Not only are our graphs 3-supertransitive, but we also use two-strand relations, making the preparation of the two-cup Jones–Wenzl too computationally expensive.

5. Relations for $3^{\mathbb{Z}/4}$

We now record the two- and one-strand jellyfish relations for a planar algebra which we will show, in Section 6, is the $3^{\mathbb{Z}/4}$ planar algebra. The three lemmas below consist of performing the calculations described in Section 4. The proofs are simply substituting in the appropriate quantities (moments, tetrahedral structure constants) where applicable, and executing the functions in the Mathematica notebooks included with the arXiv sources of this article.

The set $\mathfrak{B} = \{A, B\}$ is an orthogonal set of minimal generators which lives in the graph planar algebra. Formulas for these generators are given in Appendix A. We first check that Assumptions 2.9, 2.12, and 2.21 hold for these generators, i.e.:

- The elements $R \in \mathfrak{B}$ are self-adjoint low-weight rotational eigenvectors with corresponding chiralities σ_R given in Appendix A. Moreover, $\mathfrak{B}$ is linearly independent and orthogonal and has scalar moments. The moments are given in Appendix B.
- The sets $\mathfrak{B} \cup \{f^{(n)}\}$ and $\check{\mathfrak{B}} \cup \{f^{(n)}\}$ span complex algebras under the usual multiplication. The program *VerifyClosedUnderMultiplication* in the notebook `TwoStrandJellyfish.nb` is used to check this.
- The tetrahedral structure constants $\Delta(P, Q, R \mid S)$ are scalars for all elements $P, Q, R, S \in \mathfrak{B}$. The tetrahedral constants are given in Appendix B.

Throughout, the notation $\lambda^{(z)}_{a_n,\dots,a_0}$ denotes the root of the polynomial $\sum_i a_i x^i$ which is closest to the approximate real number z. (The digits of precision of z are in each case chosen so that this unambiguously identifies the root.) For example, $\lambda^{(0.3278)}_{1024,0,-864,0,81}$ denotes the root of $1024x^4 - 864x^2 + 81$ which is closest to 0.3278.

5A. *Two-strand relations.*

Lemma 5.1. *The following linear combinations X_i of reduced trains lie in annular consequences. The column marked X_i gives the coefficients of the reduced trains for X_i.*

	X_1	X_2
$A \underset{n-2}{\circ} A$	1	0
$A \underset{n-2}{\circ} B$	0	1
$B \underset{n-2}{\circ} A$	$\lambda^{(0.1449i)}_{2025,0,-720,0,-16}$	1
$B \underset{n-2}{\circ} B$	$\frac{1}{45}\left(-10-3\sqrt{5}\right)$	0
$A \underset{n-1}{\circ} A \underset{n-1}{\circ} A$	$\lambda^{(0.1761i)}_{100,0,-2610,0,-81}$	$\lambda^{(2.43)}_{4,0,-1134,0,6561}$
$A \underset{n-1}{\circ} A \underset{n-1}{\circ} B$	$\lambda^{(0.3447)}_{100,0,-1030,0,121}$	$\lambda^{(0.637i)}_{4,0,-198,0,-81}$
$A \underset{n-1}{\circ} B \underset{n-1}{\circ} A$	$\lambda^{(0.1493)}_{25,0,-180,0,4}$	0
$A \underset{n-1}{\circ} B \underset{n-1}{\circ} B$	$\lambda^{(-0.3436i)}_{4,0,-8,0,-1}$	$\lambda^{(0.3733)}_{4,0,-180,0,25}$
$B \underset{n-1}{\circ} A \underset{n-1}{\circ} A$	$\lambda^{(0.4370)}_{4,0,-6,0,1}$	$\lambda^{(-0.637i)}_{4,0,-198,0,-81}$
$B \underset{n-1}{\circ} A \underset{n-1}{\circ} B$	$\lambda^{(0.05869i)}_{100,0,-290,0,-1}$	$\lambda^{(0.810)}_{4,0,-126,0,81}$
$B \underset{n-1}{\circ} B \underset{n-1}{\circ} A$	$\lambda^{(0.3976i)}_{324,0,-2232,0,-361}$	$\lambda^{(0.3733)}_{4,0,-180,0,25}$
$B \underset{n-1}{\circ} B \underset{n-1}{\circ} B$	$\lambda^{(0.4382)}_{164025,0,-34020,0,484}$	0

$C_1[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.017098i)}_{2025,0,-3420,0,-1}$	$\sqrt{5}-2$
$C_1[B \underset{n-1}{\circ} B]$	$\frac{1}{45}(5-3\sqrt{5})$	$\lambda^{(-0.1237i)}_{81,0,-1044,0,-16}$
$C_2[B \underset{n-1}{\circ} A]$	$\lambda^{(0.039125i)}_{2025,0,-2610,0,-4}$	$\lambda^{(-0.540)}_{1,0,-14,0,4}$
$C_2[B \underset{n-1}{\circ} B]$	$\lambda^{(0.08686)}_{164025,0,-9720,0,64}$	$\lambda^{(0.2831i)}_{81,0,-792,0,-64}$
$C_3[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.07243i)}_{2025,0,-180,0,-1}$	1
$C_3[B \underset{n-1}{\circ} B]$	$\frac{1}{45}(-5-\sqrt{5})$	$\lambda^{(-0.5241i)}_{81,0,-36,0,-16}$
$C_4[B \underset{n-1}{\circ} A]$	$\lambda^{(0.1266i)}_{2025,0,-3960,0,-64}$	$\lambda^{(-1.75)}_{1,0,-24,0,64}$
$C_4[B \underset{n-1}{\circ} B]$	$\frac{4\sqrt{2/5}}{9}$	$\lambda^{(0.916i)}_{81,0,-1152,0,-1024}$
$C_5[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.2173i)}_{25,0,-20,0,-1}$	3
$C_5[B \underset{n-1}{\circ} B]$	$\frac{1}{15}(-5-\sqrt{5})$	$\lambda^{(-1.57i)}_{1,0,-4,0,-16}$
$C_6[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.2532i)}_{2025,0,-15840,0,-1024}$	$\lambda^{(-3.50)}_{1,0,-96,0,1024}$
$C_6[B \underset{n-1}{\circ} B]$	$\lambda^{(0.3348)}_{164025,0,-87480,0,7744}$	$\lambda^{(1.66i)}_{81,0,-360,0,-1600}$
$C_7[B \underset{n-1}{\circ} A]$	$\lambda^{(0.652i)}_{25,0,-180,0,-81}$	3
$C_7[B \underset{n-1}{\circ} B]$	$\frac{1}{30}(\sqrt{5}-5)$	$\lambda^{(-1.27i)}_{1,0,1,0,-1}$
$C_8[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.3798i)}_{25,0,-440,0,-64}$	$\lambda^{(-1.75)}_{1,0,-24,0,64}$
$C_8[B \underset{n-1}{\circ} B]$	$\lambda^{(0.05368)}_{164025,0,-22680,0,64}$	$\lambda^{(0.741i)}_{81,0,-72,0,-64}$
$C_9[B \underset{n-1}{\circ} A]$	$\lambda^{(0.2173i)}_{25,0,-20,0,-1}$	1
$C_9[B \underset{n-1}{\circ} B]$	$\frac{1}{90}(\sqrt{5}-5)$	$\lambda^{(-0.4240i)}_{81,0,9,0,-1}$
$C_{10}[B \underset{n-1}{\circ} A]$	$\lambda^{(-0.1174i)}_{25,0,-290,0,-4}$	$\lambda^{(-0.540)}_{1,0,-14,0,4}$
$C_{10}[B \underset{n-1}{\circ} B]$	$\lambda^{(0.016589)}_{164025,0,-14580,0,4}$	$\lambda^{(0.2290i)}_{81,0,-72,0,-4}$
$C_{11}[B \underset{n-1}{\circ} A]$	$\lambda^{(0.05129i)}_{25,0,-380,0,-1}$	$\sqrt{5}-2$
$C_{11}[B \underset{n-1}{\circ} B]$	$\frac{1}{90}(15-7\sqrt{5})$	$\lambda^{(-0.1001i)}_{81,0,-99,0,-1}$

In the next two lemmas, we use

$$J_2 = \begin{pmatrix} \lambda^{(0.1245i)}_{400,0,-5220,0,-81} & \frac{1}{10}(-5-\sqrt{5}) \\ \frac{1}{4}(27-9\sqrt{5}) & 0 \end{pmatrix}.$$

We let K_2 be the transpose of the 12×2 matrix whose entries are given by the first

12 rows and the 2 columns of the table in Lemma 5.1, and we define

$$Y = \begin{pmatrix} f\big(A \underset{n-2}{\circ} A\big) \\ f\big(A \underset{n-2}{\circ} B\big) \\ f\big(B \underset{n-2}{\circ} A\big) \\ f\big(B \underset{n-2}{\circ} B\big) \\ f\big(A \underset{n-1}{\circ} A \underset{n-1}{\circ} A\big) \\ f\big(A \underset{n-1}{\circ} A \underset{n-1}{\circ} B\big) \\ f\big(A \underset{n-1}{\circ} B \underset{n-1}{\circ} A\big) \\ f\big(A \underset{n-1}{\circ} B \underset{n-1}{\circ} B\big) \\ f\big(B \underset{n-1}{\circ} A \underset{n-1}{\circ} A\big) \\ f\big(B \underset{n-1}{\circ} A \underset{n-1}{\circ} B\big) \\ f\big(B \underset{n-1}{\circ} B \underset{n-1}{\circ} A\big) \\ f\big(B \underset{n-1}{\circ} B \underset{n-1}{\circ} B\big) \end{pmatrix}.$$

Lemma 5.2. *We have* $K_2 Y = J_2 \begin{pmatrix} f \cdot j^2(A) \\ f \cdot j^2(B) \end{pmatrix}$.

Lemma 5.3. *The elements* A, B *satisfy the two-strand box jellyfish relations*

$$\begin{pmatrix} f \cdot j^2(A) \\ f \cdot j^2(B) \end{pmatrix} = J_2^L K_2 Y$$

where

$$(J_2^L K_2)^T = \begin{pmatrix} 0 & \frac{1}{2}\big(\sqrt{5}-5\big) \\ \frac{1}{9}\big(3+\sqrt{5}\big) & \lambda^{(0.1001i)}_{81,0,-99,0,-1} \\ \frac{1}{9}\big(3+\sqrt{5}\big) & \lambda^{(-0.1001i)}_{81,0,-99,0,-1} \\ 0 & \frac{1}{18}\big(7+\sqrt{5}\big) \\ \sqrt{2} & 0 \\ \lambda^{(0.3706i)}_{81,0,-18,0,-4} & \lambda^{(-0.540)}_{1,0,-14,0,4} \\ 0 & \lambda^{(-0.2063)}_{1,0,-94,0,4} \\ \lambda^{(0.2172)}_{6561,0,-2430,0,100} & \lambda^{(0.512i)}_{81,0,-360,0,-100} \\ \lambda^{(-0.3706i)}_{81,0,-18,0,-4} & \lambda^{(-0.540)}_{1,0,-14,0,4} \\ \frac{\sqrt{2}}{3} & 0 \\ \lambda^{(0.2172)}_{6561,0,-2430,0,100} & \lambda^{(-0.512i)}_{81,0,-360,0,-100} \\ 0 & \lambda^{(-0.6056)}_{6561,0,-3726,0,484} \end{pmatrix}.$$

5B. *One-strand relations.*

Lemma 5.4. *The linear combinations*

$$K_1\begin{pmatrix} A\circ A\\ A\circ B\\ B\circ A\\ B\circ B\end{pmatrix} \quad \text{and} \quad \check{K}_1\begin{pmatrix} \check{A}\circ \check{A}\\ \check{A}\circ \check{B}\\ \check{B}\circ \check{A}\\ \check{B}\circ \check{B}\end{pmatrix}$$

lie in annular consequences, where

$$K_1=\begin{pmatrix} 1 & 0 & 0 & \frac{1}{18}\left(-1-\sqrt{5}\right)\\ 0 & 1 & -1 & \lambda^{(0.948i)}_{81,0,45,0,-25}\end{pmatrix} \quad \text{and} \quad \check{K}_1=\begin{pmatrix} 1 & 0 & 0 & \frac{1}{18}\left(-1-\sqrt{5}\right)\\ 0 & 1 & 1 & \lambda^{(0.948)}_{81,0,-45,0,-25}\end{pmatrix}.$$

Lemma 5.5. *In particular, we have*

$$K_1\begin{pmatrix} f(A\circ A)\\ f(A\circ B)\\ f(B\circ A)\\ f(B\circ B)\end{pmatrix}=J_1\begin{pmatrix} f\cdot j(\check{A})\\ f\cdot j(\check{B})\end{pmatrix} \quad \text{and} \quad \check{K}_1\begin{pmatrix} f(\check{A}\circ \check{A})\\ f(\check{A}\circ \check{B})\\ f(\check{B}\circ \check{A})\\ f(\check{B}\circ \check{B})\end{pmatrix}=\check{J}_1\begin{pmatrix} f\cdot j(A)\\ f\cdot j(B)\end{pmatrix},$$

where

$$J_1=\begin{pmatrix} \lambda^{(-0.590)}_{256,0,144,0,-81} & \frac{1}{24}\left(5+\sqrt{5}\right)\\ \lambda^{(2.71i)}_{256,0,2160,0,2025} & \lambda^{(0.8325i)}_{256,0,176,0,-1}\end{pmatrix} \quad \text{and} \quad \check{J}_1=\begin{pmatrix} 0 & \lambda^{(-0.49923)}_{5184,0,-1296,0,1}\\ 0 & \lambda^{(0.7277i)}_{64,0,32,0,-1}\end{pmatrix}.$$

Lemma 5.6. *The elements A and B satisfy the one-strand box jellyfish relations*

$$\begin{pmatrix} f\cdot j(\check{A})\\ f\cdot j(\check{B})\end{pmatrix}=J_1^L K_1\begin{pmatrix} f(A\circ A)\\ f(A\circ B)\\ f(B\circ A)\\ f(B\circ B)\end{pmatrix},$$

where

$$J_1^L K_1=\begin{pmatrix} \lambda^{(-0.6360)}_{16,0,-4,0,-1} & \lambda^{(-0.2303i)}_{1296,0,540,0,25} & \lambda^{(0.2303i)}_{1296,0,540,0,25} & \lambda^{(0.3327)}_{104976,0,3564,0,-1681}\\ \frac{1}{4}\left(15-3\sqrt{5}\right) & \lambda^{(-0.4504i)}_{16,0,-396,0,-81} & \lambda^{(0.4504i)}_{16,0,-396,0,-81} & \frac{1}{12}\left(7\sqrt{5}-15\right)\end{pmatrix}.$$

6. Calculating principal graphs

We now know that the set of minimal generators given in Appendix A generates an evaluable subfactor planar algebra $\mathcal{P}_\bullet^{\mathbb{Z}/4}$. We must now determine the principal graphs of the $\mathcal{P}_\bullet^{\mathbb{Z}/4}$. By the next lemma, we know that the principal graphs have the desired supertransitivity since we have two-strand jellyfish relations.

Lemma 6.1. *Suppose a planar algebra* $\mathcal{P}_\bullet$ *is generated by uncappable elements* $A_1,\dots,A_k\in\mathcal{P}_{n,+}$ *such that*

(i) *the* A_j*'s satisfy two-strand jellyfish relations*, *and*

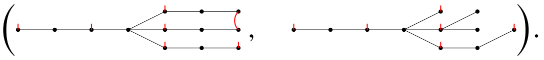
(ii) *the complex span of* $\{A_1, \dots, A_k, f^{(n)}\}$ *forms an algebra under the usual multiplication.*

Then $\mathcal{P}_\bullet$ *is* $(n-1)$ *supertransitive.*

Proof. Similar to [Morrison and Penneys 2015b, Lemma 5.1]. □

We now determine the principal graphs of the $\mathcal{P}_\bullet^{\mathbb{Z}/4}$. These arguments are similar to those in [Morrison and Penneys 2015b, Section 5].

Theorem 6.2. *The principal graphs of* $\mathcal{P}_\bullet^{\mathbb{Z}/4}$ *are*



Proof. The modulus is $\sqrt{3+\sqrt{5}} \simeq 2.28825$, and we find that the minimal projections one past the branch from bottom to top are given by $aA + bB + cf^{(4)}$, where

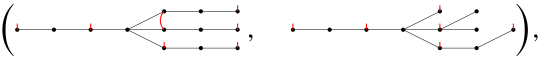
$$(a, b, c) = \begin{cases} \left(0, \frac{1}{3}, \frac{1}{3}\right), \\ \left(\frac{1}{2}, -\frac{1}{6}, \frac{1}{3}\right), \\ \left(-\frac{1}{2}, -\frac{1}{6}, \frac{1}{3}\right). \end{cases}$$

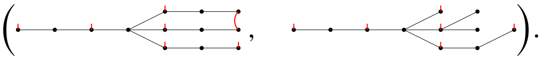
Since $\operatorname{Tr}(f^{(4)}) = 6 + 3\sqrt{5}$, all the minimal projections have trace $2+\sqrt{5}$, and the proof of [Morrison and Penneys 2015b, Theorem 5.9] shows the principal graph is correct.

To see that the dual graph is correct, we first find that the minimal projections one past the branch from bottom to top are given by $a\check{A} + b\check{B} + cf^{(4)}$, where

$$(a, b, c) = \begin{cases} \left(\lambda^{(-0.556)}_{4,0,2,0,-1}, \lambda^{(0.09003)}_{324,0,-126,0,1}, \frac{1}{3}\right), \\ \left(\lambda^{(0.2123)}_{4,0,22,0,-1}, \lambda^{(-0.3257)}_{324,0,-270,0,25}, \frac{1}{3}(\sqrt{5}-1)\right), \\ \left(\lambda^{(0.3436)}_{4,0,8,0,-1}, \frac{1}{3\sqrt{2}}, \frac{1}{3}(3-\sqrt{5})\right) \end{cases}$$

which have traces $2+\sqrt{5}$, $3+\sqrt{5}$, $1+\sqrt{5}$ respectively. Hence there is a univalent vertex at depth 4 on the dual graph. We now run the *FusionAtlas* program *FindGraphPartners* on the 3333 graph and we see there are only two possibilities where the dual graph has a univalent vertex at depth 4:





Now the projections at depth 4 on the principal graph are self-dual since $\rho^2 = \mathrm{id}$ on $\operatorname{span}\{A, B, f^{(4)}\}$, so the only possibility is the one claimed. □

Appendix A. Generators

Suppose Γ is a simply laced graph with a distinguished subgraph $\Lambda \subset \Gamma$ such that Γ is obtained from Λ by adding A_{finite} tails to Λ. For example, when Γ is a spoke graph, we can choose Λ to be the central vertex. When $\Gamma = 2D2$ (see Section AA), we can choose Λ to be the central diamond.

By the proof of [Morrison and Penneys 2015b, Lemma A.1], a low-weight generator A is completely determined by its values on loops which stay within distance 1 of Λ. Furthermore, if Γ is obtained from Λ by adding A_{finite} tails to *distinct* vertices of Λ, then A is completely determined by its values on loops which stay inside Λ. So when Γ is a spoke graph with n spokes, we can choose Λ to be an $(n-1)$-star.

Moreover, as A is a rotational eigenvector, A is completely determined by its values on a set of rotation orbit representatives which stay in Λ.

We now describe an algorithm to recover our low-weight generator A from its values on such loops.

Remark A.1. It should seem plausible, but not at all obvious, that the recovered generator is in fact a low-weight rotational eigenvector. Proposition A.11 gives a well-defined element of the graph planar algebra. For our examples, the programs *CheckLowestWeightCondition* and *CheckRotationalEigenvector* in the notebook `Generators.nb` check that the low-weight and rotational eigenvector conditions hold respectively.

Definition A.2. For a vertex $v \in \Gamma$, we define $d(v, \Lambda)$ to be the minimal distance of v to Λ. For a loop γ whose i-th vertex is denoted $\gamma(i)$, we define $d(\gamma, \Lambda) = \max_i d(\gamma(i), \Lambda)$.

2-valent folding relation. Suppose A is an n-box. We start with a loop γ on Γ of length $2n$. If $d(\gamma, \Lambda) > 1$, we can use the 2-valent relation first considered in [Peters 2010; BMPS 2012] to fold γ inward by analyzing the capping action on 2-valent vertices as follows. We use the notation of [Morrison and Penneys 2015b].

Notation A.3. Suppose $s = \gamma(i)$ is a vertex on γ whose distance from Λ is at least 2. Let t be the vertex on the same tail 2 closer to Λ than s (possibly t is in Λ itself). Let γ' be the loop modified from γ by replacing s at position i with t. Let π be the "snipped" loop of length $2n-2$ obtained from γ or γ' by removing the i-th and $i+1$-st positions. For convenience, we let $r = \gamma(i \pm 1) = \gamma'(i \pm 1)$. For an example, see Figure 1.

Definition A.4. Applying a cap at position i to A, we have $\cap_i(A) = 0$. Evaluating this at π gives the *2-valent folding relation*

$$0 = \sqrt{\dim(r)}^{k_i} \cap_i(A)(\pi) = \sqrt{\dim(s)}^{k_i} A(\gamma) + \sqrt{\dim(t)}^{k_i} A(\gamma').$$

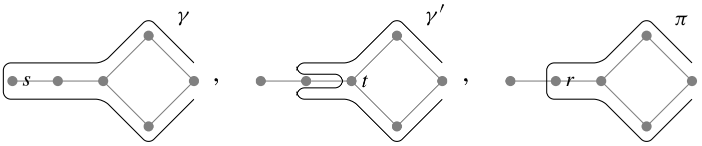
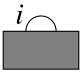
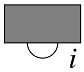
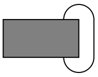
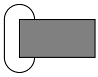


Figure 1. Example of loops and vertices appearing in the 2-valent folding relation.

Here k_i is the number of critical points in the cap strand, either 1 or 2 depending on the position of the point i around the boundary of the rectangular box:

$k_i =$
1 when we have  or ,
2 when we have  i or i .

Lemma A.5. *If $\hat{\gamma}$ is the loop of length $2n$ with $d(\hat{\gamma}, \Lambda) = 1$ obtained from γ by the 2-valent folding relation described above, we have*

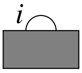
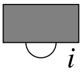
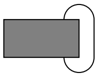
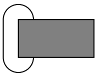
$$A(\gamma) = (-1)^{(\|\gamma\| - \|\hat{\gamma}\|)/2} \left(\prod_i \sqrt{\frac{\dim(\hat{\gamma}(i))}{\dim(\gamma(i))}}^{\,k_i} \right) A(\hat{\gamma}), \tag{6}$$

where $\|\gamma\| = \sum_i d(\gamma(i), \Lambda)$.

Remark A.6. In the lopsided convention, this formula is given by

$$A(\gamma) = (-1)^{(\|\gamma - \|\hat{\gamma}\|)/2} \left(\prod_i \left(\frac{\dim(\hat{\gamma}(i))}{\dim(\gamma(i))} \right)^{\ell_i} \right) A(\hat{\gamma}), \tag{7}$$

where ℓ_i is the number of minima on the cap:

$\ell_i =$
0 when we have ,
1 when we have ,  i , or i .

Tail avoiding relation. Now suppose Γ is obtained from Λ by adding A_{finite} tails to distinct vertices of Λ. Further suppose γ is a loop of length $2n$ with $d(\gamma, \Lambda) = 1$.

Notation A.7. Suppose $s = \gamma(i)$ is a vertex on γ which is distance 1 from Λ, and let $r = \gamma(i+1)$ which is necessarily in Λ. Let $\{t\}$ be the set of vertices in Λ incident to r. Let $\gamma_{i,t}$ be the loop modified from γ by replacing s at position i with t. Let π be the "snipped" loop of length $2n-2$ obtained from γ or $\gamma_{i,t}$ by removing the i-th and $i+1$-st positions.

Definition A.8. The *tail avoiding relation* is given by

$$0 = \sqrt{\dim(r)}^{k_i} \cap_i(A)(\pi) = \sqrt{\dim(s)}^{k_i} A(\gamma) + \sum_t \sqrt{\dim(t)}^{k_i} A(\gamma_{i,t}).$$

Lemma A.9. *If γ has $d(\gamma, \Lambda) = 1$, and $\hat{\gamma}$ has $d(\hat{\gamma}, \Lambda) = 0$ and is obtained from γ by the tail avoiding relation described above, then*

$$A(\gamma) = (-1)^{\|\gamma\|} \sum_{\{i \mid \gamma(i) \notin \Lambda\}} \sum_{\left\{t_i \,\middle|\, \substack{t_i \sim \gamma(i \pm 1) \\ t_i \in \Lambda}\right\}} \sqrt{\frac{\dim(t_i)}{\dim(\gamma(i))}}^{k_i} A(\gamma_{i,t_i}), \tag{8}$$

where $v \sim w$ means v is incident to w (note $\gamma(i+1) = \gamma(i-1)$ if $\gamma(i) \notin \Lambda$), and k_i is as in Lemma A.5.

Remark A.10. In the lopsided convention, this formula is given by

$$A(\gamma) = (-1)^{\|\gamma\|} \sum_{\{i \mid \gamma(i) \notin \Lambda\}} \sum_{\left\{t_i \,\middle|\, \substack{t_i \sim \gamma(i \pm 1) \\ t_i \in \Lambda}\right\}} \left(\frac{\dim(t_i)}{\dim(\gamma(i))}\right)^{\ell_i} A(\gamma_{i,t_i}) \tag{9}$$

using similar notation from Remark A.6 and Lemma A.9.

Rotation. We still assume Γ is obtained from Λ by adding A_{finite} tails to distinct vertices of Λ.

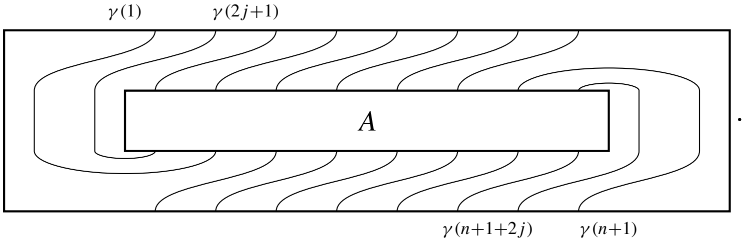
Rotation acts on the set of loops which stay in Λ, so if we are trying to specify a lowest weight vector A which is also a rotational eigenvector corresponding to eigenvalue ω, then it suffices to specify A only on a representative of each such orbit.

Proposition A.11. *Let S be a set of representatives of each rotation orbit of loops of length $2n$ in Λ. Let $A_0 : S \to \mathbb{C}$. For a loop γ of length $2n$ in Λ, let $[\gamma]$ be its representative in S. Suppose that whenever $\gamma' \in S$ is fixed by the k-fold rotation, and $\omega^k \neq 1$, then $A_0(\gamma') = 0$. Then there is a well-defined function A_1 on the loops of length $2n$ in Λ such that $A_1|_S = A_0$.*

Moreover, there is a well-defined element $A \in \mathcal{PA}(\Gamma)_n$ such that the values of A on the loops of length $2n$ on Λ is equal to A_1.

Proof. Suppose γ is a loop of length $2n$ which stays in Λ, and $\rho^{-j}(\gamma) = [\gamma]$ for some $j = 0, \ldots, n-1$. If $j \leq n/2$,

$\rho^j(A)(\gamma) = A(\rho^{-j}(\gamma)) =$



Hence for all $j = 0, \ldots, n-1$, we define

$$A_1(\gamma) = \omega^{-j}\sqrt{\frac{\dim(\gamma(2j+1))\dim(\gamma(n+2j+1))}{\dim(\gamma(1))\dim(\gamma(n+1))}}A_0([\gamma]), \tag{10}$$

modulo some modular arithmetic, namely $\gamma(b) = \gamma(b \bmod 2n)$.

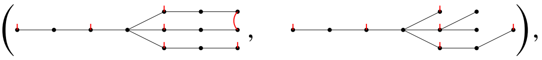
In the lopsided convention, the above equation is given by

$$A_1(\gamma) = \omega^{-j}\left(\prod_{k=1}^{2j}\frac{\dim(\gamma(1+k))}{\dim(\gamma(n+k))}\right)A_0([\gamma]). \tag{11}$$

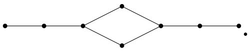
We now define $A \in \mathcal{PA}(\Gamma)_n$ as follows. First, for loops γ of length $2n$ which stay in Λ, define $A(\gamma) = A_1(\gamma)$. Next, we define A on loops γ of length $2n$ for which $d(\gamma, \Lambda) = 1$ by Lemma A.9. Finally, we define A on loops γ of length $2n$ for which $d(\gamma, \Lambda) > 1$ by Lemma A.5. $\square$

We now apply the above discussion to specify our generators by their values on a certain collection of loops. A little unusually, we find our generators in the graph planar algebra of a different graph: $\Gamma = 2\text{D}2$, which has a central diamond. We label the vertices on the diamond by `W`, `S`, `E`, `N`, which stand for "west," "south," "east," "north" respectively. We denote the value of A on the collapsed loop which stays inside the central diamond by $A(w)$, where w is a word on $\{$`W`, `S`, `E`, `N`$\}$.

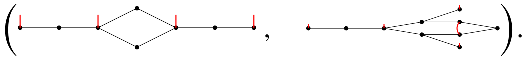
AA. ***Generators for*** $\mathbf{3^{\mathbb{Z}/4}}$**.** In an unpublished manuscript, Izumi constructs a $3^{\mathbb{Z}/4}$ subfactor with principal graphs



and he claims there is a de-equivariantization, giving a subfactor with principal graph "2-diamond-2":

2D2 = 

In an independent calculation, Morrison and Penneys [2015a] verify the existence and prove uniqueness for the 2D2 subfactor with principal graphs

$2\mathrm{D}2 =$ .

They solve the equation $T^2 = f^{(3)}$ in the graph planar algebra of 2D2 to get a low-weight rotational eigenvector at depth 3. Then they verify, using a universal variant of the jellyfish algorithm for finite depth subfactor planar algebras, that the planar subalgebra generated by T is evaluable and has principal graphs 2D2. They obtain a $3^{\mathbb{Z}/4}$ subfactor planar algebra as an equivariantization of the 2D2 subfactor planar algebra. Note that 2D2 has annular multiplicities $*12$, so the $3^{\mathbb{Z}/4}$ generators must be the new low-weight vectors at depth 4. See [Morrison and Penneys 2015a] for more details.

For our purposes in this article, we do *not* rely on the fact that our generators were obtained via equivariantization. Rather, we present candidate generators for $3^{\mathbb{Z}/4}$ in 2D2, show they satisfy Assumptions 2.9, 2.12, and 2.21, and use our formulas to show they generate an evaluable planar subalgebra of the graph planar algebra of 2D2, i.e., a subfactor planar algebra.

Hence we work in the graph planar algebra of 2D2 where Λ is the central diamond. The self-adjoint generators A, B for $\mathcal{P}_{\bullet}^{\mathbb{Z}/4}$ have chiralities $\omega_A = -1$ and $\sigma_A = i$ and $\omega_B = \sigma_B = 1$.

A assigns the below values to the indicated rotation orbit representatives of loops which remain in Λ:

0	`WSWSWSWS`, `WSWSWSES`, `WSWSWNWN`, `WSWSWNEN`, `WSWSESES`, `WSWSENEN`, `WSWNWSWN`, `WSWNWNES`, `WSWNENES`, `WSESWSES`, `WSESWNWN`, `WSESWNEN`, `WSESESES`, `WSESENWN`, `WSESENEN`, `WSENWSEN`, `WSENWNES`, `WSENENES`, `WNWNWNWN`, `WNWNWNEN`, `WNWNESES`, `WNWNENEN`, `WNESWNES`, `WNESESEN`, `WNENWNEN`, `WNENESES`, `WNENENEN`, `ESESESES`, `ESESENEN`, `ESENESEN`, `ENENENEN`
$\frac{1}{4}(3-\sqrt{5})$	`WSWSWSEN`, `WSWSWNES`, `WSWNESES`, `WSESESEN`, `WSENWNWN`, `WSENENEN`, `WNWNWNES`, `WNESESES`, `WNENENES`
$\frac{1}{4}(\sqrt{5}-3)$	`WSWSESEN`, `WSESWSEN`, `WSESWNES`, `WSENWNEN`, `WSENENWN`, `WNWNENES`, `WNESWNEN`
$2-\sqrt{5}$	`WSWNWSEN`, `WSENESEN`, `WNESENES`
$\sqrt{5}-2$	`WSWNESWN`

$\lambda^{(0.09279i)}_{16,0,-116,0,-1}$	`WSWSESWN, WSWSENES, WSWNWSES, WSWNWNWN, WSWNENEN, WSENESES, WNWNESEN, WNESENEN, ESENENEN`
$\lambda^{(-0.09279i)}_{16,0,-116,0,-1}$	`WSWSWSWN, WSWNWNEN, WSWNENWN, WSESESWN, WSESENES, WNENESEN, ESESESEN`
$\lambda^{(-0.3003i)}_{1,0,-11,0,-1}$	`WSWNESEN`
$\lambda^{(0.3003i)}_{1,0,-11,0,-1}$	`WSENESWN`

B assigns the below values to the indicated rotation orbit representatives of loops which remain in Λ:

0	`WSWSENWN, WSWSENEN, WSWNWSEN, WSWNWNES, WSWNESWN, WSWNENES, WSESENWN, WSESENEN, WSENESEN, WNWNESES, WNESENES, WNENESES`
$\lambda^{(0.2784i)}_{16,0,-1044,0,-81}$	`WSWSWSEN, WSWNESES, WSESWNES, WSESESEN, WSENWNEN, WSENENWN, WNWNWNES, WNENENES`
$\lambda^{(-0.2784i)}_{16,0,-1044,0,-81}$	`WSWSWNES, WSWSESEN, WSESWSEN, WSENWNWN, WSENENEN, WNWNENES, WNESWNEN, WNESESES`
$\frac{1}{2}\big(11-5\sqrt{5}\big)$	`WSWSWSES, WSESESES, WNWNWNEN, WNENENEN`
$\frac{1}{2}\big(5\sqrt{5}-11\big)$	`WSWSWSWS, WSWSESES, WSESWSES, WNWNWNWN, WNWNENEN, WNENWNEN, ESESESES, ENENENEN`
$\frac{1}{4}\big(\sqrt{5}-3\big)$	`WSWSWSWN, WSWSENES, WSWNWNWN, WSWNENEN, WSESESWN, WNWNESEN, ESESESEN, ESENENEN`
$\frac{1}{4}\big(3-\sqrt{5}\big)$	`WSWSESWN, WSWNWSES, WSWNWNEN, WSWNENWN, WSESENES, WSENESES, WNESENEN, WNENESEN`
$\frac{1}{2}\big(3-\sqrt{5}\big)$	`WSWNWSWN, ESENESEN`
$\frac{1}{2}\big(7-3\sqrt{5}\big)$	`WSWSWNWN, WSESWNEN, WSENWNES, ESESENEN`
$\frac{1}{2}\big(3\sqrt{5}-7\big)$	`WSWSWNEN, WSESWNWN, WSENENES, WNESESEN`
$\frac{1}{2}\big(3\sqrt{5}-9\big)$	`WSENWSEN, WNESWNES`
$2-\sqrt{5}$	`WSWNESEN, WSENESWN`

These entries lie in $\mathbb{Q}(\mu_{\mathbb{Z}/4})$, where $\mu_{\mathbb{Z}/4}$ is the root of

$$x^8 - 38x^6 + 100x^5 + 343x^4 - 2300x^3 + 5102x^2 - 5500x + 2581$$

which is approximately $2.236 + 0.700i$.

Appendix B. Moments and tetrahedral constants of $3^{\mathbb{Z}/4}$

For all of our planar algebras, our generators are self-adjoint. This is a list of the moments and tetrahedral structure constants needed for our calculations.

$$\begin{aligned}
\operatorname{Tr}(AA) &= 4+2\sqrt{5} \\
\operatorname{Tr}(AB) &= 0 \\
\operatorname{Tr}(BB) &= 12+6\sqrt{5} \\
\operatorname{Tr}(AAA) &= 0 \\
\operatorname{Tr}(AAB) &= -4-2\sqrt{5} \\
\operatorname{Tr}(ABB) &= 0 \\
\operatorname{Tr}(BBB) &= 12+6\sqrt{5}
\end{aligned}
\qquad
\begin{aligned}
\operatorname{Tr}(\check{A}\check{A}) &= 4+2\sqrt{5} \\
\operatorname{Tr}(\check{A}\check{B}) &= 0 \\
\operatorname{Tr}(\check{B}\check{B}) &= 12+6\sqrt{5} \\
\operatorname{Tr}(\check{A}\check{A}\check{A}) &= \lambda^{(-3.25)}_{4,0,-40,0,-25} \\
\operatorname{Tr}(\check{A}\check{A}\check{B}) &= \lambda^{(6.698)}_{4,0,-180,0,25} \\
\operatorname{Tr}(\check{A}\check{B}\check{B}) &= \lambda^{(13.1)}_{4,0,-648,0,-6561} \\
\operatorname{Tr}(\check{B}\check{B}\check{B}) &= \lambda^{(0.501)}_{4,0,-324,0,81}
\end{aligned}$$

$$\begin{aligned}
\Delta(A,A,A \mid A) &= -\sqrt{3+\sqrt{5}} \\
\Delta(A,A,B \mid A) &= -i\sqrt{11+5\sqrt{5}} \\
\Delta(A,B,A \mid B) &= \sqrt{107+39\sqrt{5}} \\
\Delta(B,A,B \mid A) &= \sqrt{47+21\sqrt{5}}
\end{aligned}
\qquad
\begin{aligned}
\Delta(A,A,A \mid B) &= 0 \\
\Delta(A,A,B \mid B) &= -\sqrt{2} \\
\Delta(A,B,B \mid B) &= -9i\sqrt{1+\sqrt{5}} \\
\Delta(B,A,B \mid B) &= 0 \\
\Delta(B,B,B \mid B) &= 9\sqrt{3-\sqrt{5}}
\end{aligned}$$

Acknowledgements

The authors would like to thank Masaki Izumi and Scott Morrison for many helpful conversations. Some of this work was completed when David Penneys visited Scott Morrison at the Australian National University and Masaki Izumi at Kyoto University. He would like to thank both of them for supporting those trips and for their hospitality. The authors would also like to thank the referees for suggesting significant improvements.

David Penneys was supported in part by the Natural Sciences and Engineering Research Council of Canada. Emily Peters was supported by an NSF RTG grant at Northwestern University, DMS-0636646. David Penneys and Emily Peters were both supported by DOD-DARPA grant HR0011-12-1-0009.

References

[Bigelow and Penneys 2014] S. Bigelow and D. Penneys, “Principal graph stability and the jellyfish algorithm”, *Math. Ann.* **358**:1-2 (2014), 1–24. MR 3157990 Zbl 1302.46049

[BMPS 2012] S. Bigelow, S. Morrison, E. Peters, and N. Snyder, "Constructing the extended Haagerup planar algebra", *Acta Math.* **209**:1 (2012), 29–82. MR 2979509 Zbl 1270.46058

[Evans and Gannon 2011] D. E. Evans and T. Gannon, "The exoticness and realisability of twisted Haagerup–Izumi modular data", *Comm. Math. Phys.* **307**:2 (2011), 463–512. MR 2012m:17040 Zbl 1236.46055

[Graham and Lehrer 1998] J. J. Graham and G. I. Lehrer, "The representation theory of affine Temperley–Lieb algebras", *Enseign. Math.* (2) **44**:3-4 (1998), 173–218. MR 99i:20019 Zbl 0964.20002

[Gupta 2008] V. P. Gupta, "Planar algebra of the subgroup-subfactor", *Proc. Indian Acad. Sci. Math. Sci.* **118**:4 (2008), 583–612. MR 2010g:46104 Zbl 1178.46057

[Han 2010] R. Han, *A construction of the "2221" planar algebra*, Ph.D. thesis, University of California, Riverside, CA, 2010. MR 2822034 arXiv 1102.2052

[Izumi 2001] M. Izumi, "The structure of sectors associated with Longo–Rehren inclusions, II: Examples", *Rev. Math. Phys.* **13**:5 (2001), 603–674. MR 2002k:46161 Zbl 1033.46506

[Jones 2000] V. F. R. Jones, "The planar algebra of a bipartite graph", pp. 94–117 in *Knots in Hellas '98* (Delphi, 1998), edited by C. Gordon et al., Ser. Knots Everything **24**, World Scientific, River Edge, NJ, 2000. MR 2003c:57003 Zbl 1021.46047

[Jones 2001] V. F. R. Jones, "The annular structure of subfactors", pp. 401–463 in *Essays on geometry and related topics, Vol. 1, 2* (Geneva, 2001), edited by E. Ghys et al., Monogr. Enseign. Math. **38**, Enseign. Math., Geneva, 2001. MR 2003j:46094 Zbl 1019.46036 arXiv math/0105071

[Jones 2003] V. F. R. Jones, "Quadratic tangles in planar algebras", preprint, 2003, available at http://math.berkeley.edu/∼vfr/quadtangle.pdf.

[Jones 2011] V. F. R. Jones, "Notes on planar algebras", preprint, 2011, available at http://math.berkeley.edu/~vfr/VANDERBILT/pl21.pdf.

[Jones 2012] V. F. R. Jones, "Quadratic tangles in planar algebras", *Duke Math. J.* **161**:12 (2012), 2257–2295. MR 2972458 Zbl 1257.46033

[Jones and Penneys 2011] V. F. R. Jones and D. Penneys, "The embedding theorem for finite depth subfactor planar algebras", *Quantum Topol.* **2**:3 (2011), 301–337. MR 2012f:46128 Zbl 1230.46055

[Liu 2015] Z. Liu, "Composed inclusions of A_3 and A_4 subfactors", *Adv. Math.* **279** (2015), 307–371. MR 3345186 Zbl 06435577

[LMP 2015] Z. Liu, S. Morrison, and D. Penneys, "1-supertransitive subfactors with index at most $6\frac{1}{5}$", *Comm. Math. Phys.* **334**:2 (2015), 889–922. MR 3306607 Zbl 06410160

[Morrison 2014] S. Morrison, "An obstruction to subfactor principal graphs from the graph planar algebra embedding theorem", *Bull. Lond. Math. Soc.* **46**:3 (2014), 600–608. MR 3210716 Zbl 1303.46051

[Morrison 2015] S. Morrison, "A formula for the Jones–Wenzl projections", preprint, 2015. arXiv 1503.00384

[Morrison and Penneys 2015a] S. Morrison and D. Penneys, "2-supertransitive subfactors at index $3+\sqrt{5}$", *J. Funct. Anal.* (online publication July 2015).

[Morrison and Penneys 2015b] S. Morrison and D. Penneys, "Constructing spoke subfactors using the jellyfish algorithm", *Trans. Amer. Math. Soc.* **367**:5 (2015), 3257–3298. MR 3314808 Zbl 06429013

[Morrison and Peters 2014] S. Morrison and E. Peters, "The little desert? Some subfactors with index in the interval $(5, 3+\sqrt{5})$", *Int. J. Math.* **25**:8 (2014), Article ID #1450080. MR 3254427 Zbl 06359573

[Morrison and Walker 2010] S. Morrison and K. Walker, "The graph planar algebra embedding theorem", preprint, 2010, available at http://tqft.net/gpa.

[Peters 2010] E. Peters, "A planar algebra construction of the Haagerup subfactor", *Int. J. Math.* **21**:8 (2010), 987–1045. MR 2011i:46077 Zbl 1203.46039

[Popa 1990] S. Popa, "Classification of subfactors: the reduction to commuting squares", *Invent. Math.* **101**:1 (1990), 19–43. MR 91h:46109 Zbl 0757.46054

[Popa 1995] S. Popa, "An axiomatization of the lattice of higher relative commutants of a subfactor", *Invent. Math.* **120**:3 (1995), 427–445. MR 96g:46051 Zbl 0831.46069

[Reznikoff 2007] S. A. Reznikoff, "Coefficients of the one- and two-gap boxes in the Jones–Wenzl idempotent", *Indiana Univ. Math. J.* **56**:6 (2007), 3129–3150. MR 2010e:46063 Zbl 1144.46053

Received June 3, 2014. Revised January 9, 2015.

DAVID PENNEYS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
520 PORTOLA PLAZA
MATHEMATICS DEPARTMENT
LOS ANGELES, CA 90095-1555
UNITED STATES

dpenneys@math.ucla.edu

EMILY PETERS
DEPARTMENT OF MATHEMATICS AND STATISTICS
LOYOLA UNIVERSITY CHICAGO
1032 W. SHERIDAN ROAD
CHICAGO, IL 60660
UNITED STATES

epeters3@luc.edu

CONTENTS

Volume 277, 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 277 No. 2 October 2015

0030-8730(201510)277:2;1-#