Pacific Journal of Mathematics

Volume 284 No. 2 October 2016

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

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

The subscription price for 2016 is US $440/year for the electronic version, and $600/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.257

SPHERICAL CR DEHN SURGERIES

MIGUEL ACOSTA

Consider a three-dimensional cusped spherical CR manifold M and suppose that the holonomy representation of $\pi_1(M)$ can be deformed in such a way that the peripheral holonomy is generated by a nonparabolic element. We prove that, in this case, there is a spherical CR structure on some Dehn surgeries of M. The result is very similar to R. Schwartz's spherical CR Dehn surgery theorem, but has weaker hypotheses and does not give the uniformizability of the structure. We apply our theorem in the case of the Deraux–Falbel structure on the figure eight knot complement and obtain spherical CR structures on all Dehn surgeries of slope $-3+r$, for $r \in \mathbb{Q}^+$ small enough.

1. Introduction

The celebrated theorem of hyperbolic Dehn surgeries of Thurston [2002] says that all but a finite number of Dehn surgeries of a one-cusped hyperbolic manifold M admit complete hyperbolic structures with developing maps and holonomy representations close to those of M. The same question arises for other geometric structures. We focus here on spherical CR structures, i.e., structures modeled on the boundary at infinity of the complex hyperbolic plane with group of automorphisms PU(2, 1). Schwartz [2007] shows a spherical CR Dehn surgery theorem that gives, under some convergence hypotheses, uniformizable spherical CR structures on some Dehn surgeries on a cusped spherical CR manifold. In this paper, we prove a similar theorem using techniques coming from (G, X)-structures and the geometry of $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ instead of the alternative approach of discreteness of group representations and actions on $\mathbb{H}^2_{\mathbb{C}}$. Theorem 3.23 has weaker hypotheses than Schwartz's theorem, but we obtain geometric structures on the surgeries for which we do not know whether they are uniformizable.

For the reader, the example to keep in mind, treated in Section 4, is the Deraux–Falbel structure on the figure-eight knot complement constructed in [Deraux and Falbel 2015]. For this example, Deraux [2014] shows that there is a one-parameter family of spherical CR uniformizations on the figure-eight knot complement with

MSC2010: 32V05, 57M25, 57M50.
Keywords: spherical CR, Dehn surgery, (G, X)-structures, figure-eight knot.

parabolic peripheral holonomy containing this structure. Falbel, Guilloux, Koseleff, Rouillier, and Thistlethwaite [Falbel et al. 2016] describe the $\mathrm{SL}_3(\mathbb{C})$-character variety of the fundamental group of the figure-eight knot complement. They give an explicit parametrization for the component in $\mathrm{SU}(2,1)$ containing the holonomy representation of the Deraux–Falbel structure. This component also gives rise to spherical CR structures near the Deraux–Falbel structure. With this parametrization and Theorem 3.23, we obtain the following theorem:

Theorem. *Let M be the figure-eight knot complement. For the usual*[1] *marking of the peripheral torus of M:*

(1) *There exist infinitely many spherical* CR *structures on the Dehn surgery of M of slope -3.*

(2) *There exists $\delta > 0$ such that for all $r \in \mathbb{Q} \cap]0, \delta[$, there is a spherical* CR *structure on the Dehn surgery of M of slope $-3 + r$.*

In Section 2, we recall some properties about $\mathbb{H}^2_{\mathbb{C}}$, $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$, and $\mathrm{PU}(2,1)$ and set some notation. We look in detail at the dynamics of one-parameter subgroups of $\mathrm{PU}(2,1)$ acting on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$. Understanding these dynamics will be crucial in the proof of the surgery theorem. Section 3 deals with deformation of (G, X) structures and fixes some notation and a marking of a peripheral torus in order to state the main theorem of this paper, Theorem 3.23. In Section 4, we apply Theorem 3.23 in the case of the Deraux–Falbel structure, by checking the hypotheses and looking at the deformation space as given in [Falbel et al. 2016]. Finally, in Section 5, we give a complete proof of the surgery theorem.

2. Generalities on $\mathbb{H}^2_{\mathbb{C}}$ and its isometries

In this section we recall some facts about the hyperbolic complex plane $\mathbb{H}^2_{\mathbb{C}}$ and its boundary at infinity $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ and set notation for them. We study the group of holomorphic isometries of $\mathbb{H}^2_{\mathbb{C}}$, identified with $\mathrm{PU}(2,1)$, by describing its one-parameter subgroups. Almost all stated results can be found in the thesis of Genzmer [2010] and in the book of Goldman [1999].

The space $\mathbb{H}^2_{\mathbb{C}}$ and its boundary at infinity. We begin by giving a construction of the hyperbolic complex plane. Let V be a complex vector space of dimension 3 endowed with a Hermitian product $\langle \cdot , \cdot \rangle$. Denote by Φ the associated Hermitian form. We suppose that Φ has signature $(2, 1)$. By setting

$$\begin{aligned} V_- &= \{v \in V - \{0\} \mid \Phi(v) < 0\}, \\ V_0 &= \{v \in V - \{0\} \mid \Phi(v) = 0\}, \\ V_+ &= \{v \in V - \{0\} \mid \Phi(v) > 0\}, \end{aligned}$$

[1] For us, the usual marking is the one given in [Thurston 2002].

the complex hyperbolic plane is defined as $\mathbb{P}V_-$, endowed with the Hermitian metric induced by Φ, and its boundary at infinity $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ is defined as $\mathbb{P}V_0$.

Notation 2.1. We will denote with usual parentheses "(" and ")" the objects before projectivization and with square brackets "[" and "]" the class of an object in the projectivized space.

From now on, we will use two different models of the complex hyperbolic plane, going from one to another by a conjugation. In both cases, the vector space is $V = \mathbb{C}^3$. For the details in the construction, see [Goldman 1999].

Notation 2.2. Let

$$J_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

They are the matrices of the Hermitian products $\langle \cdot , \cdot \rangle_1$ and $\langle \cdot , \cdot \rangle_2$ and they are conjugate by Cayley's matrix

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Definition 2.3. By identifying V with $\mathbb{C}^3$ and $\langle \cdot , \cdot \rangle$ with $\langle \cdot , \cdot \rangle_1$, we obtain the *ball model*. We then have

$$\mathbb{H}^2_{\mathbb{C}} = \left\{ \begin{bmatrix} Z_1 \\ Z_2 \\ 1 \end{bmatrix} \in \mathbb{CP}^2 \;\middle|\; |Z_1|^2 + |Z_2|^2 < 1 \right\}$$

and

$$\partial_\infty \mathbb{H}^2_{\mathbb{C}} = \left\{ \begin{bmatrix} Z_1 \\ Z_2 \\ 1 \end{bmatrix} \in \mathbb{CP}^2 \;\middle|\; |Z_1|^2 + |Z_2|^2 = 1 \right\}.$$

With this model, we see that $\mathbb{H}^2_{\mathbb{C}}$ is homeomorphic to the ball B^4 and $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ is homeomorphic to the sphere S^3. The other model that we will consider is the Siegel model, more convenient for drawing pictures.

Definition 2.4. By identifying V with $\mathbb{C}^3$ and $\langle \cdot , \cdot \rangle$ with $\langle \cdot , \cdot \rangle_2$, we obtain the *Siegel model*, with

$$\partial_\infty \mathbb{H}^2_{\mathbb{C}} = \left\{ \begin{bmatrix} -\frac{1}{2}(|z|^2 + it) \\ z \\ 1 \end{bmatrix} \;\middle|\; (z, t) \in \mathbb{C} \times \mathbb{R} \right\} \cup \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

We can then identify $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ with $(\mathbb{C}\times\mathbb{R})\cup\{\infty\}$. Removing the point at infinity, we obtain the Heisenberg group, defined as $\mathbb{C}\times\mathbb{R}$ with multiplication

$$(w, s) * (z, t) = (w + z,\ s + t + 2\,\mathrm{Im}(w\bar{z})).$$

We are also going to use complex geodesics, which are intersections of complex lines of $\mathbb{P}V$ with $\mathbb{H}^2_{\mathbb{C}}$, and their boundaries at infinity, called $\mathbb{C}$-circles.

Holomorphic isometries of $\mathbb{H}^2_{\mathbb{C}}$ and invariant flows. We defined above the complex hyperbolic space and have seen two of its models. The group of holomorphic isometries of this space is $\mathrm{PU}(2, 1)$, as described below.

Notation 2.5. Let $\mathrm{U}(2, 1)$ be the group of matrices of $\mathrm{GL}_3(\mathbb{C})$ such that $A^* J A = J$ for $J = J_1$ or J_2 (according to the model in which we work). Let $\mathrm{SU}(2, 1)$ be the subgroup of matrices of determinant 1 and $\mathrm{PU}(2, 1)$ its projectivization.

We state in detail a classification of the elements of $\mathrm{PU}(2, 1)$. We use the notations and state the results of [Genzmer 2010, Chapter 1]. Isometries are classified by their fixed points in $\mathbb{H}^2_{\mathbb{C}} \cup \partial_\infty \mathbb{H}^2_{\mathbb{C}}$.

An isometry $g \neq \mathrm{Id}$ of $\mathbb{H}^2_{\mathbb{C}}$ is called *elliptic* if it has at least one fixed point in $\mathbb{H}^2_{\mathbb{C}}$, *parabolic* if it is not elliptic and has exactly one fixed point in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$, and *loxodromic* if it is not elliptic and has exactly two fixed points in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$.

We can state this classification in terms of eigenvalues. The eigenvalues of an element of $\mathrm{PU}(2, 1)$ are only defined up to multiplication by a cube root of 1 that we denote by ω; we give a condition on the eigenvalues of a lift in $\mathrm{SU}(2, 1)$.

Proposition 2.6. *Let $U \in \mathrm{SU}(2, 1) - \{\mathrm{Id}\}$. Then U is in one of the three following cases*:

(1) *U has an eigenvalue λ of modulus different from* 1. *Then $[U]$ is loxodromic.*

(2) *U has an eigenvector $v \in V_-$. Then $[U]$ is elliptic and its eigenvalues have modulus equal to* 1 *but are not all equal.*

(3) *All eigenvalues of U have modulus* 1 *and U has an eigenvector $v \in V_0$. Then $[U]$ is parabolic.*

To refine this classification, we will consider different cases when there are double eigenvalues. We give the following definition:

Definition 2.7. Let $U \in \mathrm{SU}(2, 1) - \{\mathrm{Id}\}$. We say that U is *regular* if its three eigenvalues are different and *unipotent* if its three eigenvalues are equal (and so equal to a cube root of 1).

The definition extends to $\mathrm{PU}(2, 1)$; we will speak of regular elements of $\mathrm{PU}(2, 1)$. In that case the eigenvalues are well-defined up to multiplication by ω. Thanks to the following remark, we know that regular elements are easier to manipulate.

Remark 2.8. Let $[U] \in \mathrm{PU}(2,1)$ be a regular element. Then $[U]$ is determined by its three eigenvalues α, β, γ and its three fixed points $[u], [v], [w]$ in $\mathbb{CP}^2$.

It is possible to know if an element is regular only by knowing its trace, thanks to the following proposition.

Proposition 2.9 [Goldman 1999]. *For $z \in \mathbb{C}$, let $f(z) = |z|^4 - 8\,\mathrm{Re}(z^3) + 18|z|^2 - 27$. Let $U \in \mathrm{SU}(2,1)$. Then U is regular if and only if $f(\mathrm{tr}(U)) \neq 0$. Furthermore, if $f(\mathrm{tr}(U)) < 0$ then $[U]$ is regular elliptic, and if $f(\mathrm{tr}(U)) > 0$ then $[U]$ is loxodromic.*

Remark 2.10. It is suitable to make two remarks about the proposition:

(1) For $\omega \in \mathbb{C}$ satisfying $\omega^3 = 1$, we have $f(z) = f(\omega z)$. Therefore, we can define the function $f \circ \mathrm{tr}$ on $\mathrm{PU}(2,1)$.

(2) For a parabolic element $[U]$, the equality $f(\mathrm{tr}(U)) = 0$ holds, but there are nonregular elliptic elements for which $f(\mathrm{tr}(U)) = 0$.

In order to study spherical CR structures and their surgeries, we will use the flows of vector fields associated to some elements of $\mathrm{PU}(2,1)$. The geometric objects that we are going to consider are invariant vector fields induced by elements of $\mathrm{PU}(2,1)$. We begin by looking at an infinitesimal level: an element of the Lie algebra $\mathfrak{su}(2,1)$ defines a vector field on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ invariant under its exponential map.

Notation 2.11. Let $X \in \mathfrak{su}(2,1)$. It defines a vector field on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ invariant by $\exp(X)$ given at a point x by

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \exp(tX) \cdot x.$$

Let ϕ_t^X be the flow of this vector field, so $\phi_t^X(x) = \exp(tX) \cdot x$. If there is no ambiguity for X, we will only write ϕ_t.

Remark 2.12. If $[U] \in \mathrm{PU}(2,1)$ is close enough to a unipotent element, it defines a vector field on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$. Indeed, possibly after changing the lift, we can suppose that the eigenvalues of U are near 1, and consider the vector field associated to $\mathrm{Log}(U)$. Then, $\phi_1^{\mathrm{Log}(U)}$ has the same action as $[U]$.

Description of isometries and invariant flows. We are going to describe briefly some elements of $\mathrm{PU}(2,1)$, and classify each by its type and the dynamics of its action on $\mathbb{CP}^2$.

We are going to study the dynamics of some flows of the form $\phi_t^{\mathrm{Log}(U)}$, where U is close to a unipotent element. We describe here flows associated to regular elliptic, loxodromic and unipotent elements.

Regular elliptic flows. Consider a regular elliptic element in SU(2, 1), close to Id, in the ball model. Perhaps after a conjugation, we can suppose that it is equal to

$$E_{\alpha,\beta,\gamma} = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\beta} & 0 \\ 0 & 0 & e^{i\gamma} \end{pmatrix}.$$

The flow of the associated vector field acts then on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ by

$$\phi_t^{\mathrm{Log}(E_{\alpha,\beta,\gamma})} \begin{bmatrix} Z_1 \\ Z_2 \\ 1 \end{bmatrix} = \begin{bmatrix} e^{it(\alpha-\gamma)} Z_1 \\ e^{it(\beta-\gamma)} Z_2 \\ 1 \end{bmatrix} = \begin{bmatrix} e^{it(2\alpha+\beta)} Z_1 \\ e^{it(2\beta+\alpha)} Z_2 \\ 1 \end{bmatrix}.$$

Remark 2.13. The flow stabilizes the two $\mathbb{C}$-circles

$$C_1 = [e_1]^\perp \cap \partial_\infty \mathbb{H}^2_{\mathbb{C}} = \left\{ \begin{bmatrix} 0 \\ e^{i\theta} \\ 1 \end{bmatrix} \;\middle|\; \theta \in \mathbb{R} \right\}$$

and

$$C_2 = [e_2]^\perp \cap \partial_\infty \mathbb{H}^2_{\mathbb{C}} = \left\{ \begin{bmatrix} e^{i\theta} \\ 0 \\ 1 \end{bmatrix} \;\middle|\; \theta \in \mathbb{R} \right\},$$

on which it acts as rotations by angles $2\beta+\alpha$ and $2\alpha+\beta$ respectively.

Remark 2.14. The centralizer of $E_{\alpha,\beta,\gamma}$ is

$$C(E_{\alpha,\beta,\gamma}) = \{E_{\theta_1,\theta_2,-(\theta_1+\theta_2)} \mid (\theta_1, \theta_2) \in \mathbb{R}^2\}.$$

The orbits of this subgroup in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ are C_1, C_2, and the subsets T_r for $r \in \,]0, 1[$, defined by

$$T_r = \left\{ \begin{bmatrix} Z_1 \\ Z_2 \\ 1 \end{bmatrix} \in \partial_\infty \mathbb{H}^2_{\mathbb{C}} \;\middle|\; |Z_2| = r,\ |Z_1| = \sqrt{1-r^2} \right\}.$$

The orbits T_r are embedded tori in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ with core curves C_1 and C_2. They are all invariant under the action of $\phi_t^{\mathrm{Log}(E_{\alpha,\beta,\gamma})}$. We can see an example in Figure 1.

Have a look now at the orbits of the flow $\phi_t^{\mathrm{Log}(E_{\alpha,\beta,\gamma})}$. Notice that the orbit of a point is included in a unique torus T_r, and that every orbit included in T_r is the image of a fixed orbit by an element $E_{\theta_1,\theta_2,-(\theta_1+\theta_2)}$. Thus, the torus T_r is foliated by these orbits. We fix $r \in \,]0, 1[$, and consider two cases:

<u>Case 1: $\alpha/\beta \notin \mathbb{Q}$.</u> In this case, the angles of rotation in T_r for ϕ_t are $(2\alpha+\beta)t$ and $(2\beta+\alpha)t$. Since their quotient is irrational, an orbit is an injective immersion of a line and it is dense in T_r.

<u>Case 2: $\alpha/\beta \in \mathbb{Q}$.</u> In this case, the angles of rotation in T_r for ϕ_t are $(2\alpha+\beta)t$ and $(2\beta+\alpha)t$. Their quotient is rational; denote it by p/q in reduced form. The orbits are periodic and of slope p/q in T_r: they are torus knots of type (p, q), knotted around C_1 and C_2. We can see an example in Figure 2.

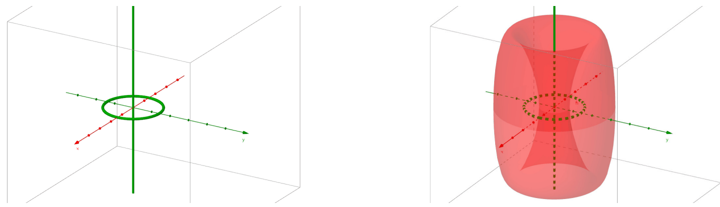


Figure 1. Invariant subsets for an elliptic flow in the Siegel model: invariant $\mathbb{C}$-circles (left), and the invariant torus $T_{4/5}$ (right).

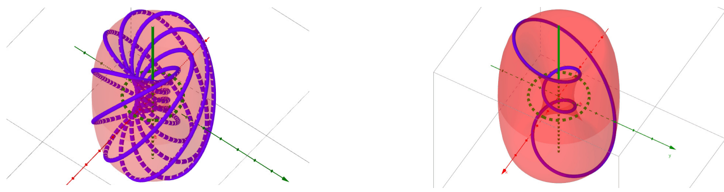


Figure 2. Orbits of elliptic flows in the Siegel model: an orbit for $(2\alpha+\beta)/(2\beta+\alpha) = \frac{7}{11}$, a torus knot of type (7, 11) (left), and an orbit for $(2\alpha+\beta)/(2\beta+\alpha) = \frac{1}{3}$, an unknot (right).

Remark 2.15. If p are q are different from ± 1, the orbit of a point of T_r is a *torus knot of type* (p, q) and is knotted in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$. If p or q equals ± 1, then the orbit is unknotted; we can see an example in Figure 2. This remark will be crucial to identify Dehn surgeries among the structures that we will construct by deformation.

Definition 2.16. Let n, p, and q be relatively prime integers with $|p| \geq |q|$. We say that an elliptic element $U \in \mathrm{PU}(2, 1)$ is *of type* $(p/n, q/n)$ if U is conjugate to $E_{\alpha,\beta,\gamma}$ with

$$\alpha = \frac{2p-q}{3n}, \quad \beta = \frac{2q-p}{3n}, \quad \text{and} \quad \gamma = -\alpha-\beta = \frac{-p-q}{3n}.$$

In this case, $(2\alpha+\beta)/(2\beta+\alpha) = p/q$ and the orbits of the flow $\phi_t^{\mathrm{Log}(U)}$ are its two invariant $\mathbb{C}$-circles and torus knots of type (p, q).

Remark 2.17. (1) Only some elliptic elements are of some type $(p/n, q/n)$. We will see later that elements of some type $(p/n, q/n)$ are the ones for which our construction happens to work.

(2) The trace of an elliptic element gives its three eigenvalues, but it is not enough to determine the type of the element. Indeed, an element of the same trace as an elliptic of type $(p/n, q/n)$ will have the same eigenvalues but not necessarily the same eigenvalue associated to its fixed point in $\mathbb{H}^2_{\mathbb{C}}$. Thus, elements of type $(p/n, q/n)$, $(-p/n, (q-p)/n)$, and $((p-q)/n, -q/n)$ have the same trace but are not conjugate.

Loxodromic flows. Consider a loxodromic element in SU(2, 1) in the Siegel model. Perhaps after a conjugation, we can suppose that it is

$$T_\lambda = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \bar{\lambda}/\lambda & 0 \\ 0 & 0 & 1/\bar{\lambda} \end{pmatrix},$$

where $\lambda \in \mathbb{C}$ is of modulus greater than 1. We then have $\lambda = re^{i\alpha}$, with $\alpha \in \mathbb{R}$ and $r > 1$. We suppose that α is small enough, so the series $\mathrm{Log}(T_\lambda)$ converges. In coordinates $(z, s) \in \mathbb{C} \times \mathbb{R}$, the action of the flow is given by $\phi_t^{\mathrm{Log}(T_\lambda)} : (z, t) \mapsto (\mu_t z, |\mu_t|^2 s)$, where $\mu_t = r^t e^{-3i\alpha t}$.

Remark 2.18. The flow ϕ_t fixes globally the points

$$[p_0] = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad [p_1] = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

and stabilizes the $\mathbb{C}$-circle joining them, which is called the *axis* of $[T_\lambda]$. Furthermore, for all $u \in \partial_\infty \mathbb{H}^2_{\mathbb{C}}$ not fixed by T_λ,

$$\lim_{t \to +\infty} \phi_t(u) = [p_1] \quad \text{and} \quad \lim_{t \to -\infty} \phi_t(u) = [p_0].$$

In the same way as in the elliptic case, we have flow-invariant objects, related to the centralizer of T_λ.

Remark 2.19. The centralizer of T_λ is $C(T_\lambda) = \{T_\mu \mid \mu \in \mathbb{C}^*\}$. The orbits of this subgroup in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ are the two fixed points of T_λ, the two arcs of the $\mathbb{C}$-circle joining them, and the punctured paraboloids P_r for $r \in \mathbb{R}$, as in Figure 3, defined by

$$P_r = \left\{ \begin{bmatrix} -\frac{1}{2}(|z|^2 + is) \\ z \\ 1 \end{bmatrix} \in \partial_\infty \mathbb{H}^2_{\mathbb{C}} \;\middle|\; \frac{s}{|z|^2} = r \right\}.$$

Unipotent flows. Consider now a unipotent element of SU(2, 1) in the Siegel model. Perhaps after a conjugation, we can assume that it is, for $(z, s) \in \mathbb{C} \times \mathbb{R}$,

$$P_{z,s} = \begin{pmatrix} 1 & -\bar{z} & -\frac{1}{2}(|z|^2 + is) \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}.$$

The series $\mathrm{Log}(P_{z,s})$ converges. In coordinates $(z, s) \in \mathbb{C} \times \mathbb{R}$, the action of the flow is given by $\phi_t^{\mathrm{Log}(P_{z,s})} : (z', s') \mapsto (z' + tz, s' + ts - 2t\,\mathrm{Im}(\bar{z}z'))$. In these coordinates, the orbits of the flow are straight lines.

Remark 2.20. If $z = 0$, then $[P_{z,s}]$ is called a *vertical parabolic* element and all orbits of the flow are vertical lines. If not, then $[P_{z,s}]$ is called a *horizontal parabolic* element and the orbits of the flow are lines with different slopes.

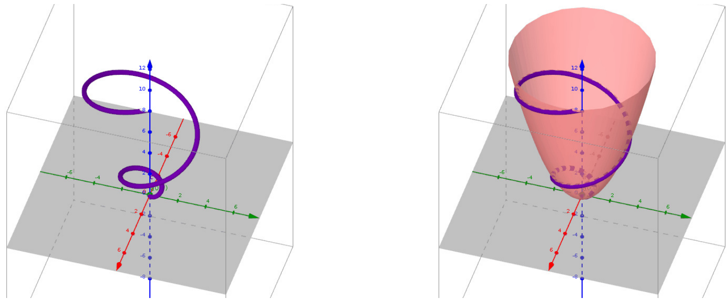


Figure 3. Orbits of loxodromic flows in the Siegel model: an orbit of a loxodromic flow (left) and a cylinder invariant under a loxodromic flow (right).

Some remarks on the convergence of regular elements. The projection

$$\mathrm{SU}(2,1) \to \mathrm{PU}(2,1)$$

is a covering of order 3; in order to study the convergence in $\mathrm{PU}(2,1)$ we can focus on the convergence in $\mathrm{SU}(2,1)$.

Let $(U_n)_{n\in\mathbb{N}}$ be a sequence of regular elements of $\mathrm{SU}(2,1)$ converging to U in $\mathrm{SU}(2,1) - \mathbb{C}\,\mathrm{Id}$. If U is regular, then the convergence is given by the convergence of eigenvectors and eigenvalues. We consider now the case where U is not regular. The continuity of eigenvectors and eigenvalues gives the following lemma.

Lemma 2.21. *Suppose that $(U_n)_{n\in\mathbb{N}}$ is a sequence of regular elements of $\mathrm{SU}(2,1)$ converging to $U \in \mathrm{SU}(2,1) - \mathbb{C}\,\mathrm{Id}$, and let $(([u_n], \alpha_n), ([v_n], \beta_n), ([w_n], \gamma_n))$ be the eigenvectors and eigenvalues of U_n in some order. Then, perhaps after relabeling, $(([u_n], \alpha_n), ([v_n], \beta_n), ([w_n], \gamma_n))$ converges to $(([u], \alpha), ([v], \beta), ([w], \gamma))$ in $(\mathbb{CP}^2 \times \mathbb{C})^3$, where $([u], \alpha), ([v], \beta), ([w], \gamma)$ are (possibly equal) eigenvectors and eigenvalues of U.*

Consider the case where U is horizontal parabolic. Then, U has a unique fixed point $[p] \in \mathbb{CP}^2$, which is in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$, and its eigenvalues can be chosen all equal to 1. Using the above lemma, we deduce that $(\alpha_n, \beta_n, \gamma_n) \to (1,1,1)$ and $([u_n], [v_n], [w_n]) \to ([p], [p], [p])$. From a geometric point of view on $\mathbb{H}^2_{\mathbb{C}} \cup \partial_\infty \mathbb{H}^2_{\mathbb{C}}$ we make the two following remarks:

Remark 2.22. If the U_n are loxodromic of axes l_n then the l_n leave every compact subset of $\mathbb{H}^2_{\mathbb{C}} \cup \partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{[p]\}$.

Remark 2.23. If the U_n are elliptic, they each have two invariant complex geodesics $l_n^{(1)}$ and $l_n^{(2)}$ (the polar lines $[v_n]^\perp$ and $[w_n]^\perp$ if $[u_n]$ is the fixed point of U_n in $\mathbb{H}^2_{\mathbb{C}}$). Then the $l_n^{(i)}$ leave every compact subset of $\mathbb{H}^2_{\mathbb{C}} \cup \partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{[p]\}$.

These two remarks will be crucial when understanding the geometry of deformations of spherical CR structures by considering a developing map.

3. Regular surgeries

The Ehresmann–Thurston principle. We are going to study spherical CR structures on a 3-manifold M. We begin by recalling the formalism of (G, X)-structures, that will give us the language to use. In the definitions, X will be a smooth connected manifold and G a subgroup of the diffeomorphisms of X acting transitively and analytically on X. We will focus on the case where $X = \partial_\infty \mathbb{H}^2_{\mathbb{C}}$ and $G = \mathrm{PU}(2, 1)$.

Definition 3.1. A (G, X)-*structure* on a manifold M is a pair (Dev, ρ), up to isotopy, of a local diffeomorphism $\mathrm{Dev} : \widetilde{M} \to X$ and a group homomorphism $\rho : \pi_1(M) \to G$ such that for all $x \in \widetilde{M}$ and all $g \in \pi_1(M)$ we have $\mathrm{Dev}(g \cdot x) = \rho(g) \cdot \mathrm{Dev}(x)$ for the group actions of $\pi_1(M)$ on $\widetilde{M}$ and of G on X.

We say that Dev is the *developing map* of the structure and ρ its *holonomy*.

Remark 3.2. We identify two structures if they are G-equivalent, i.e., if there is a $g \in G$ such that the developing maps Dev_1 and Dev_2 satisfy $\mathrm{Dev}_2 = g \circ \mathrm{Dev}_1$. In this case, the holonomy representations are conjugate and satisfy $\rho_2 = g\rho_1 g^{-1}$.

Remark 3.3. The definition we just gave is not the usual one. It is equivalent to the usual definition of a (G, X)-structure as an atlas of charts of M taking values in X and whose transition maps are given by elements of G. A couple (Dev, ρ) immediately gives such an atlas, but the construction of (Dev, ρ) from an atlas requires more work. See, for example, [Thurston 2002]. Nevertheless, we will use both definitions: the first in order to deform a structure, and the second to construct a new one.

We will also sometimes use manifolds with boundary, but the definition of (G, X)-structure easily extends to this case. From now on, we consider a compact three-dimensional manifold M with (possibly many) torus boundary components. We are going to study spherical CR structures on M, where the model space X is $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ and the group G is $\mathrm{PU}(2, 1)$.

Definition 3.4. A *spherical* CR *structure* is a $(\mathrm{PU}(2, 1), \partial_\infty \mathbb{H}^2_{\mathbb{C}})$-structure.

In order to deform the structure using the Ehresmann–Thurston principle that we state below, the essential objects are the representations of $\pi_1(M)$ taking values in $\mathrm{PU}(2, 1)$.

Notation 3.5. Let $\mathcal{R}(\pi_1(M), G)$ be the set of representations of $\pi_1(M)$ taking values in G, endowed with the topology of pointwise convergence.

We are going to work with deformations of some structures. In order to state the results on a deformation, we will need to be "far enough from the boundary" or "close to the boundary". We are going to consider the union of M with a thickening of its boundary to be able to state the results precisely.

Notation 3.6. If $s \in \mathbb{R}^+$, denote by $M_{[0,s[}$ the union of M with a thickening of its boundary. Thus, $M_{[0,s[} = (M \cup (\partial M \times [0,s[))/\sim$, where we identify ∂M with $\partial M \times \{0\}$. We consider those manifolds as included into each other, in such a way that if $s_1 \leq s_2$, then $M_{[0,s_1[} \subset M_{[0,s_2[}$.

Remark 3.7. The manifolds $M_{[0,s[}$ are all homeomorphic to the interior of M. We use these cuts in order to get a suitable convergence "far enough" from the boundary of M for geometric structures.

We state the Ehresmann–Thurston principle, which says that we only need to deform in $\mathcal{R}(\pi_1(M), G)$ the holonomy of a (G, X)-structure to have a deformation of the structure itself. A proof can be found in [Bergeron and Gelander 2004] or in the survey [Goldman 2010].

Theorem 3.8 (Ehresmann–Thurston principle). *Suppose that $M_{[0,1[}$ is endowed with a (G, X)-structure of holonomy ρ_0. For all $\epsilon > 0$, if $\rho \in \mathcal{R}(\pi_1(M), G)$ is a deformation close enough to ρ_0, then there is a (G, X)-structure on $M_{[0,1-\epsilon[}$ with holonomy ρ and close to the first structure on $M_{[0,1-\epsilon[}$ in the $\mathcal{C}^1$ topology.*

Surgeries. As in the real hyperbolic case, we consider Dehn surgeries of M, which are, from a topological point of view, a gluing of solid tori on the torus boundaries of M. We attempt to extend a spherical CR structure on M to one of its surgeries. We show a result very similar to the one showed by Schwartz [2007], but with some differences. On the one hand, our hypotheses are weaker than Schwartz's and we obtain a geometric structure. On the other hand we do not know if the structure is obtained as a quotient of an open set of $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ by the action of a subgroup of $\mathrm{PU}(2,1)$.

Thickenings and lifts. We begin by fixing notation for a torus boundary component, one of its lifts, and the associated peripheral holonomy. We denote by $\widetilde{M}$ the universal cover of M and by $\pi : \widetilde{M} \to M$ its covering map. We state all results for a single torus boundary component in order to avoid heavy notation, but the same statements hold for several boundary components.

Notation 3.9. Let T be a fixed torus boundary component of M. For $s \in [0,1[$, let $T_s = T \times \{s\} \subset M_{[0,1[}$, and, for an interval $I \subset [0,1[$, let

$$T_I = \bigcup_{s\in I} T_s = T \times I \subset M_{[0,1[}.$$

Let $\widetilde{T}_{[0,1[}$ denote some connected component of $\pi^{-1}(T_{[0,1[}) \subset \widetilde{M}_{[0,1[}$: it is a universal cover of $T_{[0,1[}$ embedded in $\widetilde{M}_{[0,1[}$. Finally, for $s \in [0,1[$, set $\widetilde{T}_s = \pi^{-1}(T_s) \cap \widetilde{T}_{[0,1[}$ and, for an interval $I \subset [0,1[$, set $\widetilde{T}_I = \bigcup_{s\in I} \widetilde{T}_s$.

We make some remarks on the choices made by using this notation:

Remark 3.10. For all $s \in [0, 1[$, we see that $\widetilde{T}_s$ is homeomorphic to $\mathbb{R}^2$. Furthermore, $\widetilde{T}_I$ is homeomorphic to $\mathbb{R}^2 \times I$.

Remark 3.11. The choice of $\widetilde{T}_{[0,1[}$ fixes an injection of the fundamental group of T into the fundamental group of M by identifying $\pi_1(T)$ with the stabilizer of $\widetilde{T}_{[0,1[}$ for the action of $\pi_1(M)$ on $\widetilde{M}_{[0,1[}$. In the rest of the paper, we will use additive notation for $\pi_1(T) \simeq \mathbb{Z}^2$, in order to use the standard notations and tools for a group isomorphic to $\mathbb{Z}^2$. Nevertheless, the identification of $\pi_1(T)$ with a subgroup of $\pi_1(M)$ will lead to a slight abuse of notation: we will keep multiplicative notation for $\pi_1(M)$, but when considering elements of $\pi_1(T)$ we will use additive notation.

Notation 3.12. With the fixed injection of $\pi_1(T)$ into $\pi_1(M)$, by restricting the holonomy ρ of a (G, X)-structure we have a *peripheral holonomy* $h_\rho : \pi_1(T) \to G$.

Notation 3.13. We denote by $\mathcal{R}_1(\pi_1(M), G) \subset \mathcal{R}(\pi_1(M), G)$ the set of representations ρ such that the image of h_ρ is generated by a single element. When $\rho \in \mathcal{R}_1(\pi_1(M), \mathrm{PU}(2, 1))$ has $[U] \in \mathrm{PU}(2, 1)$ as a preferred generator for its image, we write ϕ_t^ρ for $\phi_t^{\mathrm{Log}([U])}$.

Horotubes. We use the definitions related to horotubes given in [Schwartz 2007]:

Definition 3.14. Let $[P] \in \mathrm{PU}(2, 1)$ be a parabolic element with fixed point $p \in \partial_\infty \mathbb{H}^2_{\mathbb{C}}$. A $[P]$-*horotube* is an open set H of $\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p\}$, invariant under $[P]$ and such that the complement of $H/\langle [P] \rangle$ in $(\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p\})/\langle [P] \rangle$ is compact.

In order to work with more regular objects, we often ask horotubes to be *nice*:

Definition 3.15. A $[P]$-horotube H is *nice* if ∂H is a smooth cylinder invariant by the flow $\phi_t^{\mathrm{Log}([P])}$.

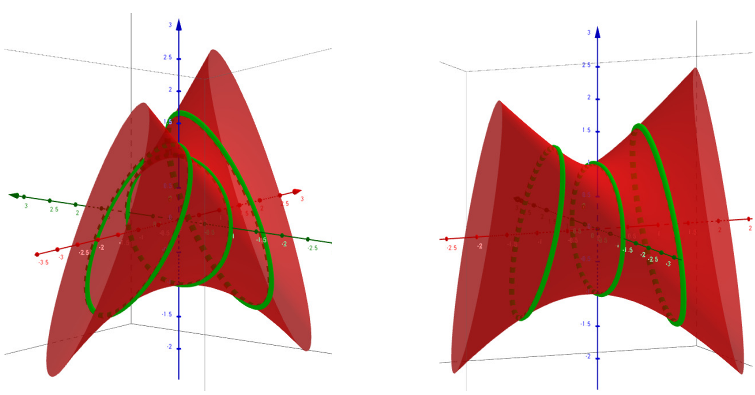


Figure 4. The boundary of a nice horotube in the Siegel model. The horotube is outside the red surface.

Remark 3.16. If H is a nice $[P]$-horotube, then ∂H is the orbit for $\phi_t^{\mathrm{Log}([P])}$ of an embedded circle of $\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p\}$. We can see an example in Figure 4.

Shrinking the horotube if necessary, we may assume it is nice:

Lemma 3.17 [Schwartz 2007, Chapter 7]. *Let H be a $[P]$-horotube. Then, there is a nice $[P]$-horotube H' such that $H' \subset H$ and $(H - H')/\langle [P] \rangle$ is of compact closure in $(\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p\})/\langle [P] \rangle$.*

From now on, we suppose that $M_{[0,1[}$ has a spherical CR structure with developing map Dev_0 and holonomy ρ_0. We also make two more hypotheses:

(1) The image of the peripheral holonomy h_{ρ_0} is unipotent of rank 1 and generated by an element $[U_0] \in \mathrm{PU}(2, 1)$.

(2) There is $s \in [0, 1[$ such that $\mathrm{Dev}_0(\widetilde{T}_{[s,1[})$ is a $[U_0]$-horotube.

Marking of $\pi_1(T)$. We are going to fix a marking of $\pi_1(T)$ naturally deduced from the structure given by Dev_0 and ρ_0. This marking will be useful to identify the Dehn surgeries obtained when deforming the structure. It is essentially the same marking as the one given in [Schwartz 2007, Chapter 8]; its definition uses the two hypotheses given above.

Notation 3.18. Fix $s' \in [s, 1[$ and $x_0 \in \mathrm{Dev}_0(T_{s'})$. Let l be the loop given by the projection of $t \mapsto \phi_t^{\rho_0}(x_0)$. As $h_{\rho_0}(l) = [U_0]$ generates the image of h_{ρ_0} and since a unipotent subgroup of $\mathrm{PU}(2, 1)$ has no torsion, l is a primitive element of $\pi_1(T)$.

Notation 3.19. Since h_{ρ_0} is unipotent of rank 1 and a unipotent subgroup of $\mathrm{PU}(2, 1)$ has no torsion, its kernel is generated by a primitive element m. We orient m in such a way that (l, m) is a direct basis of $\pi_1(T)$ (for the orientation given by the inside normal in the horotube).

Remark 3.20. The definition of l and m does not depend on the choice of s' nor of x_0. Nevertheless, we make a choice for orientations. The one for m is explicit, but the orientation of l is given by the choice of $[U_0]$ or $[U_0]^{-1}$ as a generator for the image of h_{ρ_0}.

Remark 3.21. Schwartz [2007] gives a "canonical" choice for the orientations of l and m (denoted β and α). It is almost the same choice as the one made above, but he has a preferred choice for $[U_0]$. Note that the marking (l, m) given here might not be the usual one. If we have another marking of $\pi_1(T)$, for example when M is a knot complement, changing markings can be done easily when ρ_0 is known explicitly.

Definition 3.22. For two relatively prime integers p, q, we denote by $M^{(p,q)}$ the manifold obtained by gluing a solid torus $D^2 \times S^1$ on the boundary T of M such that the loop $pl + qm$ of T becomes trivial in $D^2 \times S^1$. We refer to it as the *Dehn surgery of M of type* (p, q) or of *slope* p/q.

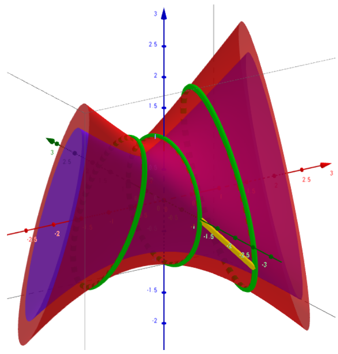


Figure 5. The curve m (in green) and the curve l (in yellow) in the image of $\mathrm{Dev}_0(\widetilde{T}_{s'})$.

In the real hyperbolic case, deforming the complete hyperbolic structure on M gives structures on all but a finite number of Dehn surgeries $M^{(p,q)}$ of M, as is shown in [Thurston 2002]. The main idea to prove it is to deform the structure "far" from the cusp, cut by T, look at the developing map near the boundary T, and then notice that a solid torus can be glued to this boundary. What follows, stated in the spherical CR case, is inspired by this technique. The deformation "far" from the cusp gives rise to a developing map near T, and the manifolds that can be glued are solid tori only in some cases.

A surgery theorem. We are now able to state a spherical CR surgery theorem. It says that in a neighborhood of the structure (Dev_0, ρ_0), under some discreteness conditions, spherical CR structures on M come from structures on Dehn surgeries of M, and in some cases another kind of surgery.

Theorem 3.23. *Let M be a three-dimensional compact manifold with torus boundary components. Let T be one boundary torus of M. Suppose that there is a spherical* CR *structure* (Dev_0, ρ_0) *on* $M_{[0,1[}$ *such that*:

(1) *The image of the peripheral holonomy h_{ρ_0} corresponding to T is unipotent of rank* 1 *and generated by an element* $[U_0] \in \mathrm{PU}(2,1)$.

(2) *There is $s \in [0,1[$ such that* $\mathrm{Dev}_0(\widetilde{T}_{[s,1[})$ *is a* $[U_0]$*-horotube.*

Then there is an open set Ω of $\mathcal{R}_1(\pi_1(M), \mathrm{PU}(2,1))$ containing ρ_0 such that, for all $\rho \in \Omega$ for which the image of h_ρ is generated by a single element $[U] \in \mathrm{PU}(2,1)$, there is a spherical CR *structure on M with holonomy ρ. Furthermore, for the marking (l, m) of $\pi_1(T)$ described above*:

(1) *If $[U]$ is loxodromic, then the structure extends to a spherical* CR *structure on the Dehn surgery of M of type* $(0,1)$.

(2) *If* $[U]$ *is elliptic of type* $(p/n, \pm 1/n)$, *then the structure extends to a spherical* CR *structure on the Dehn surgery of* M *of type* $(n, \pm p)$.

(3) *If* $[U]$ *is elliptic of type* $(p/n, q/n)$ *with* $|p|, |q| > 1$, *then the structure extends to a spherical* CR *structure on the gluing of* M *with a compact manifold with torus boundary* $V(p, q, n)$. *Furthermore* $V(p, q, n)$ *is a torus knot complement in the lens space* $L(n, \alpha)$, *where* $\alpha \equiv p^{-1}q \mod n$.

Remark 3.24. The existence of the spherical CR structure on M is a consequence of the Ehresmann–Thurston principle. To extend the structure we need a local surgery result, similar to the one given in [Schwartz 2007], and which is given in Section 5.

Remark 3.25. If $[U]$ is parabolic, the theorem still holds, but the spherical CR structure extends to a thickening of M that is homeomorphic to M itself. We also exclude from the discussion the case where $[U]$ is elliptic with irrational angle, for which there is no reasonable filling for the structure, and the case where $[U]$ is nonregular elliptic, for which the techniques used to prove Theorem 3.23 do not apply.

4. Deformations of the Deraux–Falbel structure on the figure-eight knot complement

We are going to apply Theorem 3.23 in the case of the spherical CR structure on the figure-eight knot given in [Deraux and Falbel 2015]. We will use some results of [Deraux 2014], where Deraux describes a Ford domain for the structure, and also some results of [Falbel et al. 2016], where the authors describe the irreducible components of the $\mathrm{SL}_3(\mathbb{C})$ character variety of the figure eight knot complement.

Notation 4.1. In the rest of this section, we denote by M the figure-eight knot complement.

The Deraux–Falbel structure. We begin by recalling quickly the results in [Deraux and Falbel 2015]. In that paper, the fundamental group of M is given by

$$\pi_1(M) = \langle g_1, g_2, g_3 \mid g_2 = [g_3, g_1^{-1}],\ g_1 g_2 = g_2 g_3 \rangle.$$

The authors construct a uniformizable spherical CR structure on M with unipotent peripheral holonomy. The holonomy representation ρ_0 is given by

$$\rho_0(g_1) = [G_1] = \begin{bmatrix} 1 & 1 & -\frac{1}{2} - \frac{\sqrt{7}}{2} i \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \rho_0(g_3) = [G_3] = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{1}{2} + \frac{\sqrt{7}}{2} i & 1 & 1 \end{bmatrix}.$$

Remark 4.2. This representation is in the component R_2 of the character variety of [Falbel et al. 2016]. For the notation from Section 5.2 of that reference, we have $A = g_3$ and $B = g_1$. With this notation, the usual longitude-meridian pair (l_0, m_0)

of the knot complement satisfies

$$m_0 = g_3 \quad \text{and} \quad l_0 = g_1^{-1} g_3 g_1 g_3^{-2} g_1 g_3 g_1^{-1}.$$

Furthermore, we check easily that $\rho_0(m_0)^3 = \rho_0(l_0)$, so $\rho_0(3m_0 - l_0) = \mathrm{Id}$.

Notation 4.3. From now on, in order to have the same notation as [Deraux 2014], we consider the pair (l_1, m_1) obtained by conjugation by g_2, so that $m_1 = g_2 g_3 g_2^{-1} = g_1$.

Let $l = m_0$ and $m = 3m_0 - l_0$. In this way, m generates $\ker(\rho_0)$ and $\rho(l)$ generates $\mathrm{Im}(\rho_0)$: this is a marking as in the one on page 269.

Checking the hypotheses. Recall the hypotheses of Theorem 3.23:

(1) The peripheral holonomy h_{ρ_0} is unipotent with image generated by a single element $[U_0] \in \mathrm{PU}(2, 1)$.

(2) There exists $s \in [0, 1[$ such that $\mathrm{Dev}_0(\widetilde{T}_{[s,1[})$ is a $[U_0]$-horotube.

The first hypothesis is satisfied by the Deraux–Falbel structure: the peripheral holonomy is unipotent, its image is generated by $[G_1] = \rho_0(l)$ and $\rho_0(m) = [\mathrm{Id}]$.

In order to check the second hypothesis, we use the results of [Deraux 2014]. In that paper, Deraux finds with a different technique the Deraux–Falbel structure [2015]. He considers a Ford domain F in $\mathbb{H}^2_{\mathbb{C}}$ for $\Gamma = \rho_0(\pi_1(M))$ (Theorem 5.1) and then studies its boundary at infinity in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ (Section 8). The manifold M is then obtained as a quotient of a G_1-invariant domain $E = \partial_\infty F$, that is, in $\partial_\infty \mathbb{H}^2_{\mathbb{C}} \simeq (\mathbb{C} \times \mathbb{R}) \cup \{\infty\}$, the exterior of a G_1-invariant cylinder C embedded in $\mathbb{C} \times \mathbb{R}$ (Proposition 8.1). The domain E is a $[G_1]$-horotube; so there exists $s \in [0, 1[$ such that the image by the developing map of $\widetilde{T}_{[s,1[}$ is a $[G_1]$-horotube contained in E. Thus, the second hypothesis is satisfied.

So, the conclusion of Theorem 3.23 holds. By changing coordinates in order to have the usual marking for the fundamental group of the boundary of M, we get:

Proposition 4.4. *There is an open set Ω of $\mathcal{R}_1(\pi_1(M), \mathrm{PU}(2, 1))$ such that, for all $\rho \in \Omega$ such that the image of h_ρ is generated by an element $[U] \in \mathrm{PU}(2, 1)$, there exists a spherical* CR *structure on M of holonomy ρ. Furthermore, for the usual marking (l_0, m_0) of $\pi_1(T)$:*

(1) *If $[U]$ is loxodromic, then the structure extends to a spherical* CR *structure on the Dehn surgery of type $(-1, 3)$ of M.*

(2) *If $[U]$ is elliptic of type $(p/n, \pm 1/n)$, then the structure extends to a spherical* CR *structure on the Dehn surgery of type $(-n, \pm p + 3n)$ of M.*

(3) *If $[U]$ is elliptic of type $(p/n, q/n)$ with $|p|, |q| > 1$, then the structure extends to a spherical* CR *structure on the gluing of M to a compact manifold with torus boundary $V(p, q, n)$ along their boundaries. Furthermore, $V(p, q, n)$ is the complement of a torus knot in the lens space $L(n, \alpha)$, where $\alpha \equiv p^{-1} q \mod n$.*

Remark 4.5. If $[U]$ is parabolic, then the theorem still holds, but the spherical CR structure extends to a thickening of M. These structures are given in [Deraux 2014].

Remark 4.6. We wonder if Schwartz's horotube surgery theorem [2007, Theorem 1.2] can be applied in this case. For $\Gamma = \rho_0(\pi_1(M))$, the construction of Deraux and Falbel [2015] states that the regular set Ω_Γ is nonempty and that Ω_Γ/Γ is homeomorphic to M, but we do not have any more information about Ω_Γ and the limit set $\Lambda_\Gamma = \partial_\infty \mathbb{H}^2_{\mathbb{C}} - \Omega_\Gamma$. In order to apply the horotube surgery theorem, we would have to check several nontrivial hypotheses. In particular we do not know how to prove that the set Λ_Γ is *porous*. One of the main motivations of this paper was to state a result with more simple hypotheses, even if we obtain weaker conclusions when both theorems can be applied.

Deformations of the structure. It remains to see that the open set $\Omega \subset \mathcal{R}_1(\pi_1(M))$ is not reduced to a point to get interesting conclusions. The representation ρ_0 is in the component R_2 of the $\mathrm{SL}_3(\mathbb{C})$-character variety described in [Falbel et al. 2016]. In Section 5 of that paper, the representations in R_2 taking values in $\mathrm{SU}(2,1)$ are parametrized up to conjugacy, at least in a neighborhood of ρ_0, by a complex parameter $u = \mathrm{tr}(\rho(m_0))$. We denote by $G(u) = \rho(m_0)$ the corresponding matrix.

Setting $v = \bar{u}$, $\Delta = 4u^3 + 4v^3 - u^2v^2 - 16uv + 16$, and

$$\Delta' = \frac{-16 + 8uv - 2v^3 - 4\sqrt{\Delta}}{8u^2 - 6uv^2 + v^4},$$

the parametrization is explicitly given by

$$[G_3^{-1}(u)] = \rho(a) = \begin{bmatrix} \frac{1}{2}v & 1 & -(1-i)\Delta' \\ \frac{1}{8}(1+i)(-2u+v^2) & \frac{1}{4}(1+i)v & 1 \\ \frac{1}{16}(8-4uv+v^3-2\sqrt{\Delta}) & \frac{1}{8}(-4u+v^2) & \frac{1}{4}(1-i)v \end{bmatrix}$$

and

$$[G_1^{-1}(u)] = \rho(b) = \begin{bmatrix} \frac{1}{2}v & i & (1+i)\Delta' \\ -\frac{1}{8}(1+i)(-2u+v^2) & \frac{1}{4}(1-i)v & i \\ -\frac{1}{16}(8-4uv+v^3-2\sqrt{\Delta}) & -\frac{i}{8}(-4u+v^2) & \frac{1}{4}(1+i)v \end{bmatrix}.$$

Recall that for this choice of generators the usual meridian m_0 is given by $m_0 = a^{-1}$. The Hermitian form preserved by this representation is given by the matrix[2]

$$H = \begin{pmatrix} \frac{1}{8}(\Delta-16)(\sqrt{\Delta}+|u|^2-4) & 0 & 0 \\ 0 & 16-\Delta & 0 \\ 0 & 0 & 8(\sqrt{\Delta}+4) \end{pmatrix}.$$

[2]We write here the opposite of the matrix H appearing in [Falbel et al. 2016] in order to have signature (2, 1) and not (1, 2).

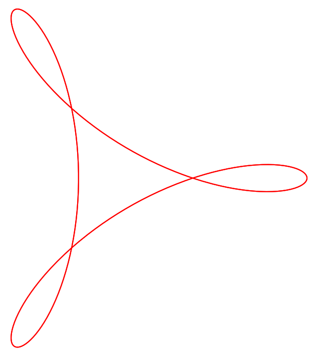


Figure 6. Domain parametrizing a component of the deformation variety near ρ_0.

Furthermore, in the whole component the relation $\rho(l_0) = \rho(m_0)^3$ holds, so $\mathcal{R}_1(\pi_1(M)) \cap R_2 = \mathcal{R}(\pi_1(M)) \cap R_2$. By projecting to PU(2, 1), we can apply Theorem 3.23 on an open set containing $3 = \mathrm{tr}(\rho_0(m_0))$ with these parameters.

Figure 6, taken from [Falbel et al. 2016], shows an open set of $\mathbb{C}$ where we have representations. By noting $\mathrm{tr}(\rho_0(m_0)) = x + iy$, the component containing ρ_0 admits as parameters the regions with boundary the curve $\Delta(x, y) = 0$ and containing the points 3, 3ω, and $3\omega^2$, where

$$\Delta(x, y) = -x^4 - y^4 - 2x^2y^2 - 24xy^2 + 8x^3 - 16x^2 - 16y^2 + 16.$$

Now let us plot the curve $\mathcal{C}$ of traces of nonregular elements of SU(2, 1). It is given by the zeroes of the function $f(z) = |z|^4 - 8\,\mathrm{Re}(z^3) + 18|z|^2 - 27$ (see Proposition 2.9). The curve separates regular elliptic and loxodromic elements. It has a singularity at the point $u = 3$: thus a neighborhood of this point contains points corresponding to representations where the peripheral holonomy is loxodromic and points where it is regular elliptic.

Remark 4.7. The parabolic deformations of the Deraux–Falbel structure given in [Deraux 2014] correspond to the points of $\mathcal{C}$.

We can therefore apply the first point of Proposition 4.4 to the space of holonomy representations given by the parameters above. We obtain the following proposition:

Proposition 4.8. *There exist infinitely many spherical* CR *structures on the Dehn surgery of M of type* $(-1, 3)$.

Remark 4.9. This surgery is the unit tangent bundle to the hyperbolic orbifold (3, 3, 4). It is a Seifert manifold of type $S^2(3, 3, 4)$. See, for example, Chapter 5 of the book of Cooper, Hodgson, and Kerckhoff [Cooper et al. 2000] or the paper [Deraux 2015]. Deraux [2014, Section 4; 2015, Theorem 4.2] also remarks that the image of ρ_0 is a faithful representation of the even words of the (3, 3, 4) triangle group, generated by involutions I_1, I_2, I_3. This identification satisfies $G_1 = I_2 I_3 I_2 I_1$, $G_2 = I_1 I_2$, $G_3 = I_2 I_1 I_2 I_3$, and the triangle group relations: $(G_2)^4 = (I_1 I_2)^4 = \mathrm{Id}$, $(G_1 G_2)^3 = (I_2 I_3)^3 = \mathrm{Id}$, and $(G_2 G_1 G_2)^3 = (I_1 I_3)^3 = \mathrm{Id}$. Furtheremore, the image of the usual meridian m_0 is G_3.

This group is the fundamental group of a Seifert manifold of type $S^2(3, 3, 4)$. Since the relation $l_0 = m_0^3$ holds in the whole component R_2, the images of representations in R_2 are representations of this index-2 subgroup of the $(3, 3, 4)$ triangle group. Furthermore, Parker, Wang, and Xie [Parker et al. 2016] show that a $\mathrm{PU}(2, 1)$ representation of the $(3, 3, 4)$ triangle group is discrete and faithful if and only if the image of $I_1 I_3 I_2 I_3$ is nonelliptic. Note that $G_1 I_1 I_3 I_2 I_3 = (I_2 I_3)^3 = \mathrm{Id}$, so the representation of the triangle group is discrete and faithful if and only if the corresponding peripheral holonomy is nonelliptic. They also give a one-parameter family of such representations, corresponding to the parameters $u \in \mathbb{R}$. Thus, there exists $\delta > 0$ such that all the spherical CR structures on the Dehn surgery of M of type $(-1, 3)$ with parameter u in the interval $]3, 3 + \delta[$ have discrete and faithful holonomy.

Since the parameter is the trace of an element, we know that cases (2) and (3) of Proposition 4.4 happen infinitely many times, but we can not distinguish at first sight, for a given trace, if it is a Dehn surgery or a gluing of a $V(p, q, n)$ manifold. Nevertheless, using a computation with the explicit parametrization of [Falbel et al. 2016] and the continuity of eigenvalues we prove:

Proposition 4.10. *There is $\delta > 0$ such that, if $p, n \in \mathbb{N}$ are relatively prime integers such that $p/n < \delta$, then the Dehn surgery of M of type $(-n, -p + 3n)$ admits a spherical* CR *structure.*

Proof. Let $p, n \in \mathbb{N}$ be relatively prime integers. Let

$$\alpha = \frac{-2p - 1}{3n}, \quad \beta = \frac{2 + p}{3n}, \quad \gamma = \frac{p - 1}{3n}, \quad \text{and} \quad u = e^{i\alpha} + e^{i\beta} + e^{i\gamma}.$$

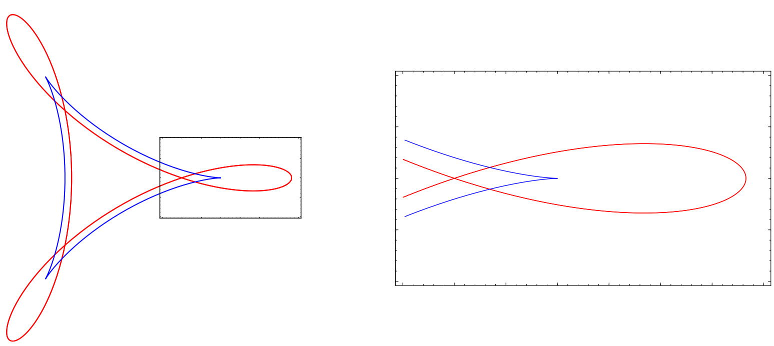


Figure 7. Curve of nonregular elements in a component of the deformation variety near ρ_0 (left) and detail of same (right).

We only need to show that if p/n is small enough, the eigenvalue of $\rho(m) = G_3^{-1}(u)$ corresponding to a negative eigenvector is $e^{i\gamma}$, and so $G_3(u)$ is of type $(p/n, -1/n)$.

Since eigenvectors and eigenvalues are continuous functions of u in the connected component of regular elliptics, in R_2, as in Figure 7, the statement is true for all (p, n) if and only if it is true for a particular choice of (p, n). For the arbitrary choice $(p, n) = (3, 23)$ an explicit computation shows that $G_3(u)$ is of type $\left(\frac{3}{23}, -\frac{1}{23}\right)$. □

5. Proof of Theorem 3.23

In this section, we are going to prove Theorem 3.23. We use the notation of Section 3. We have a manifold M with a torus boundary T, endowed with a spherical CR structure (Dev_0, ρ_0) such that the image of the holonomy h_{ρ_0} is unipotent of rank 1 and generated by an element $[U_0] \in \mathrm{PU}(2, 1)$. We suppose that there is $s \in [0, 1[$ such that $\mathrm{Dev}_0(\widetilde{T}_{[s,1[})$ is a $[U_0]$-horotube. Recall that we work with a single boundary component T to avoid heavy notation, but the proof works for several boundary components.

In order to prove the theorem, we begin by rewriting the hypotheses to make them easier to handle. The existence of a spherical CR structure on M for a deformation of ρ_0 will be a consequence of the Ehresmann–Thurston principle. To extend it to a surgery of M, we need only a local surgery result by looking near the boundary of $M_{[0,1[}$. This surgery result is very similar, in cases (1) and (2), to the one given in [Schwartz 2007, Chapter 8].

Rewriting the hypotheses. First of all, we rewrite the second hypothesis. Fix a diffeomorphism $\psi : \mathbb{R}^2 \times [0, 1[\to \widetilde{T}_{[0,1[}$, such that:

(1) $\psi(\mathbb{R}^2 \times \{s\}) = \widetilde{T}_s$ for all $s \in [0, 1[$.

(2) ψ induces a diffeomorphism between $\mathbb{R} \times S^1 \times [0, 1[$ and $\widetilde{T}_{[0,1[} / \ker(h_{\rho_0})$.

To avoid too much notation, we identify $\mathbb{R}^2 \times [0, 1[$ with $\widetilde{T}_{[0,1[}$ and $\mathbb{R} \times S^1 \times [0, 1[$ with $\widetilde{T}_{[0,1[} / \ker(h_{\rho_0})$. In this case, the developing map Dev_0 induces a diffeomorphism between $\widetilde{T}_{[0,1[} / \ker(h_{\rho_0})$ and $\mathrm{Dev}_0(\widetilde{T}_{[0,1[})$ that we will still call Dev_0. We replace hypothesis (2) of the theorem by hypotheses (2′) and (3) described below:

Hypothesis (2′): There are $0 < s_1 < s_2 < 1$ such that:

(1) $\mathrm{Dev}_0(\{0\} \times S^1 \times \{s\})$ is a circle transverse to the flow for all $s \in [s_1, s_2]$.

(2) $\mathrm{Dev}_0(t, \zeta, s) = \phi_t^{\rho_0}(\mathrm{Dev}_0(0, \zeta, s))$ for all $(t, \zeta, s) \in \mathbb{R} \times S^1 \times [s_1, s_2]$.

Remark 5.1. Thanks to Lemma 3.17, it is clear that hypothesis (2′) follows from hypothesis (2). Perhaps after considering an isotopy and slightly increasing s, we can suppose that the horotube $\mathrm{Dev}_0(\widetilde{T}_{[s,1[})$ is nice. We only need then to consider the restriction to a segment $\widetilde{T}_{[s_1,s_2]}$.

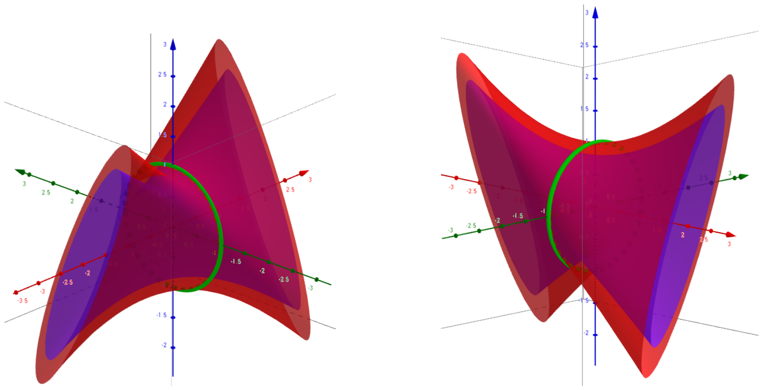


Figure 8. Two views of surfaces bounding a region of the form $\mathrm{Dev}_0(\widetilde{T}_{[s_1,s_2]})$.

Hypothesis (2′) gives, in particular, that $\mathrm{Dev}_0(\widetilde{T}_{s_2})$ separates $\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p\}$ in two connected components: a solid cylinder C_{s_2} and the exterior of this cylinder, which is homeomorphic to $S^1 \times \mathbb{R} \times]0, +\infty[$. Hypothesis (3) tells us that the structure of M is on the correct side of the tube:

Hypothesis (3): $\mathrm{Dev}_0(\widetilde{T}_{s_1})$ is contained in C_{s_2}.

Remark 5.2. Hypothesis (2) is equivalent to hypotheses (2′) and (3). The implication from (2) to (2′) and (3) is clear, and, if we suppose (2′) and (3), the structure can be extended to the outside in such a way that $\mathrm{Dev}_0(\widetilde{T}_{[s_2,1[})$ is the horotube with boundary $\mathrm{Dev}_0(\widetilde{T}_{s_2})$.

Deforming the structure. We now prove Theorem 3.23. To begin, assume that the rewritten hypotheses on page 276 are satisfied. Let ρ be a deformation close to ρ_0 in $\mathcal{R}_1(\pi_1(M), \mathrm{PU}(2,1))$ such that $h_\rho(m) = \mathrm{Id}$. The image of h_ρ is then generated by $[U] = \rho(l)$. We suppose that $[U]$ is a regular element.

Let $\epsilon > 0$. By the Ehresmann–Thurston principle, if ρ is close enough to ρ_0, there is a spherical CR structure on $M_{[0,s_2+\epsilon[}$ with holonomy map ρ. We then have a developing map $\mathrm{Dev}_\rho : \widetilde{M}_{[0,s_2+\epsilon[} \to \partial_\infty \mathbb{H}^2_{\mathbb{C}}$ close to Dev_0 in the $\mathcal{C}^1$ topology. So, we can suppose that Dev_ρ is still a diffeomorphism between the compact set $[-\epsilon, 1+\epsilon] \times S^1 \times [s_1, s_2]$ and its image.

Remark 5.3. We then have an atlas of charts on $T_{[s_1,s_2]}$ taking values in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ by choosing lifts of $T_{[s_1,s_2]}$ in the space $[-\epsilon, 1+\epsilon] \times S^1 \times [s_1, s_2] \subset \widetilde{T}_{[s_1,s_2]}$. Transition maps are given by powers of $[U] = \rho(l)$.

Fix $s_1 < s_1' < s_2' < s_2$.

Lemma 5.4 (straightening). *If ρ is close enough to ρ_0, perhaps after taking an isotopy of* Dev_ρ, *we have, for all* $(t, \zeta, s) \in \mathbb{R} \times S^1 \times [s_1', s_2']$,

$$\mathrm{Dev}_\rho(t, \zeta, s) = \phi_t^\rho(\mathrm{Dev}_\rho(0, \zeta, s)).$$

Proof. The flows ϕ_t^ρ and $\phi_t^{\rho_0}$ are close in the $\mathcal{C}^1$ topology when ρ is close to ρ_0. We deduce that the deformation from ρ_0 to ρ induces a $\mathcal{C}^1$ deformation from $\phi_t^{\rho_0} \circ \mathrm{Dev}_0$ to $\phi_t^\rho \circ \mathrm{Dev}_\rho$. First we restrict to the compact set $[0,1] \times S^1 \times [s_1', s_2']$, which is in the interior of $[-\epsilon, 1+\epsilon] \times S^1 \times [s_1, s_2]$.

Since

$$\mathrm{Dev}_0([0,1] \times S^1 \times [s_1', s_2']) = \bigcup_{t \in [0,1]} \phi_t^{\rho_0}(\{0\} \times S^1 \times [s_1', s_2']),$$

if ρ is close enough to ρ_0,

$$\bigcup_{t \in [0,1]} \phi_t^{\rho}(\{0\} \times S^1 \times [s_1', s_2'])$$

is contained in the interior of $\mathrm{Dev}_\rho([0,1] \times S^1 \times [s_1, s_2])$.

Since $[U] \cdot \phi_t^\rho = \phi_{t+1}^\rho$ and $[U] \cdot \mathrm{Dev}_\rho(t, \zeta, s) = \mathrm{Dev}_\rho(t+1, \zeta, s)$, we can straighten Dev_ρ by a $[U]$-equivariant isotopy to have, for $(t, \zeta, s) \in \mathbb{R} \times S^1 \times [s_1', s_2']$,

$$\mathrm{Dev}_\rho(t, \zeta, s) = \phi_t^\rho(\mathrm{Dev}_\rho(0, \zeta, s)). \qquad \square$$

From now on, we suppose that for all $(t, \zeta, s) \in \mathbb{R} \times S^1 \times [s_1', s_2']$ we have $\mathrm{Dev}_\rho(t, \zeta, s) = \phi_t^\rho(\mathrm{Dev}_\rho(0, \zeta, s))$.

Lemma 5.5. *Let C be a $\mathbb{C}$-circle invariant by $[U]$. Then C and the annulus $\mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2'])$ are not linked.*

Proof. First, $[U]$ is a regular element close to the unipotent element $[U_0]$, which has fixed point $p_0 \in \partial_\infty \mathbb{H}^2_{\mathbb{C}}$. Thanks to Remarks 2.22 and 2.23, we know that C leaves every compact subset of $\partial_\infty \mathbb{H}^2_{\mathbb{C}} - \{p_0\}$ when $[U]$ approaches $[U_0]$. Since $\mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2'])$ stays in a fixed compact set when we deform ρ_0 to ρ, we deduce that C and the annulus $\mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2'])$ are not linked. $\square$

It only remains to establish a local surgery result, similar to the result contained in [Schwartz 2007, Chapter 8].

Thanks to Lemma 5.4, we know that $\mathrm{Dev}_\rho(\widetilde{T}_{[s_1', s_2'-\epsilon]})$ is the orbit by ϕ_t^ρ of the annulus $A = \mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2' - \epsilon])$. This orbit separates $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ (if $[U]$ is elliptic) or $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ minus two points (if $[U]$ is loxodromic) into two connected components C_1 and C_2, with respective boundaries $\mathrm{Dev}_\rho(\widetilde{T}_{s_1'})$ and $\mathrm{Dev}_\rho(\widetilde{T}_{s_2'})$. We have a proper action of $[U]$ on C_2, and so we can consider the quotient manifold $N = C_2 / \langle [U] \rangle$. It is a compact manifold with a torus boundary, endowed with a spherical CR structure which coincides with the structure of $M_{[0, s_2'[}$ on $T_{]s_2'-\epsilon, s_2'[}$. Thus, the gluing $M_{[0, s_2'[} \cup N / \sim$ has a spherical CR structure which extends the structure $(\mathrm{Dev}_\rho, \rho)$ of M.

We will show that if $[U]$ is loxodromic or elliptic of type $(p/n, 1/n)$, then N is a solid torus and that we have a spherical CR structure on a Dehn surgery of M of

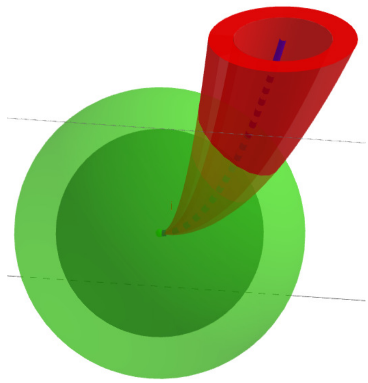


Figure 9. The orbit of A under ϕ_t (red) and the spheres S and $[U] \cdot S$ (green).

a certain slope. If $[U]$ is elliptic of type $(p/n, q/n)$, we will see a description of N as a complement of a torus knot in some lens space.

Case 1: $[U]$ is loxodromic. We work in the Siegel model, and we identify $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ with $(\mathbb{C} \times \mathbb{R}) \cup \{\infty\}$. Perhaps after conjugating, we can assume that there exists $\lambda \in \mathbb{C}$ such that $|\lambda| > 1$ and

$$U = T_\lambda = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \bar{\lambda}/\lambda & 0 \\ 0 & 0 & 1/\bar{\lambda} \end{pmatrix}.$$

Note that $[U]$ has two fixed points: $(0, 0)$ and ∞. Let S be the sphere centered at $(0, 0)$ and of radius 1 in $\mathbb{C} \times \mathbb{R}$. This sphere is a fundamental domain for the action of the flow ϕ_t^ρ. The subgroup generated by $[U]$ acts properly on $(\mathbb{C} \times \mathbb{R}) - (0, 0)$, and the region $\bigcup_{t \in [0,1]} \phi_t^\rho(S)$ with boundary components S and $[U] \cdot S$ is a fundamental domain for this action. The orbit of A under ϕ_t^ρ intersects S in an annulus that separates S into two disks D_1 and D_2, so that their orbits under ϕ_t are the connected components C_1 and C_2 respectively. Figure 9 shows this situation.

The quotient manifold $N = C_2/\langle [U] \rangle$ is obtained by identifying D_2 and $[U] \cdot D_2$ in $\bigcup_{t \in [0,1]} \phi_t^\rho(D_2)$. Thus, it is a solid torus. But the curve of $\pi_1(T)$ that becomes trivial in C_2 is the one homotopic to the boundary of D_2: so it is m. We deduce that the surgery is of type $(0, 1)$.

Case 2: $[U]$ is elliptic of type $(p/n, \pm 1/n)$. By choosing $[U_0]^{\pm 1}$ instead of $[U_0]$ as the generator of the peripheral holonomy, we can suppose that U is of type $(\pm p/n, 1/n)$. For ease of exposition, we write the proof for $[U]$ of type $(p/n, 1/n)$.

We reason in the same way as in the loxodromic case. By Lemma 5.5, we know that $\mathrm{Dev}_\rho(\widetilde{T}_{[s_1', s_2' - \epsilon]})$ is the orbit under ϕ_t of the annulus $A = \mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2'])$, which is not linked to any of the invariant $\mathbb{C}$-circles of $[U]$.

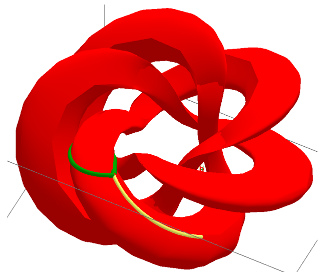


Figure 10. The orbit of A under ϕ_t (in red), the longitude l (yellow), and the meridian m (green).

The orbit of A under the flow ϕ_t^ρ is then homeomorphic to $S^1 \times S^1 \times [s_1', s_2']$. Its complement in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ has two connected components; let C_2 be the component with boundary $\mathrm{Dev}_\rho(\widetilde{T}_{s_2'})$. Following Remark 2.15, the orbits of the flow are not knotted: the two connected components are solid tori, and $[U]$ acts properly on each one. But the quotient of a solid torus by a proper action of a finite group is still a solid torus. The quotient manifold $N = C_2/\langle [U] \rangle$ is then a solid torus, and we have a spherical CR structure on a Dehn surgery of M. It only remains to identify it.

Perhaps after a conjugation, we can assume that

$$U = e^{-2i\pi(p+1)/(3n)} \begin{pmatrix} e^{2i\pi p/n} & 0 & 0 \\ 0 & e^{2i\pi/n} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

in the ball model. In the Siegel model, by identifying $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ with $(\mathbb{C} \times \mathbb{R}) \cup \{\infty\}$, we have that $[CUC^{-1}]$ stabilizes two $\mathbb{C}$-circles: the circle $\mathcal{C}_1$ centered at 0 of radius $\sqrt{2}$ in $\mathbb{C} \times \{0\}$ and $\mathcal{C}_2$ the axis $\{0\} \times \mathbb{R}$. A generic orbit of the flow turns once around $\mathcal{C}_1$ and p times around $\mathcal{C}_2$.

Let γ be the loop that follows the $\mathbb{C}$-circle $\mathcal{C}_2$ and is oriented so that the meridian m is homotopic to γ in the component C_2. In this case, nl is homotopic, also in C_2, to $-p\gamma$. Thus $nl + pm$ is a homotopically trivial loop in C_2, which is a covering of the solid torus N glued to M. So it is also a trivial loop in N. We deduce that the surgery is of type (n, p).

Case 3: $[U]$ is elliptic of type $(p/n, q/n)$. As in Cases 1 and 2 above, we know that $\mathrm{Dev}_\rho(\widetilde{T}_{[s_1', s_2' - \epsilon]})$ is the orbit by ϕ_t of the annulus $A = \mathrm{Dev}_\rho(\{0\} \times S^1 \times [s_1', s_2'])$, which is not linked to any of the invariant $\mathbb{C}$-circles of $[U]$.

The orbit of A under the flow ϕ_t^ρ is homeomorphic to $S^1 \times S^1 \times [s_1', s_2']$. Its complement in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ has two connected components. Let C_2 be (again) the component with boundary $\mathrm{Dev}_\rho(\widetilde{T}_{s_2'})$ and C_1 the one with boundary $\mathrm{Dev}_\rho(\widetilde{T}_{s_1'})$. According to Remark 2.15, generic orbits of the flow are torus knots of type (p, q): C_1 is then a tubular neighborhood of one of the orbits and C_2 is homeomorphic to the complement of a torus knot of type (p, q). But $[U]$ acts properly on $\partial_\infty \mathbb{H}^2_{\mathbb{C}}$ and stabilizes C_1 and C_2.

Notice that, in the ball model, the action of the group generated by $[U]$ is the same as the one of the group generated by

$$(z_1, z_2) \mapsto (e^{2i\pi/n} z_1, e^{2i\pi\alpha/n} z_2),$$

where $\alpha \equiv p^{-1}q \mod n$. The quotient $\partial_\infty \mathbb{H}^2_{\mathbb{C}}/\langle [U] \rangle$ is then homeomorphic to the lens space $L(n, \alpha)$. Furthermore, $C_1/\langle [U] \rangle$ is a solid torus knotted in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}/\langle [U] \rangle$. The quotient manifold $V(p, q, n) = C_2/\langle [U] \rangle$ is the complement of a torus knot in $\partial_\infty \mathbb{H}^2_{\mathbb{C}}/\langle [U] \rangle \simeq L(n, \alpha)$. The spherical CR structure of M then extends to the gluing of M and $V(p, q, n)$ along their torus boundary components.

Acknowledgements

The author thanks his advisors Martin Deraux and Antonin Guilloux, as well as Elisha Falbel and the referee for the many discussions and useful remarks on this paper.

References

[Bergeron and Gelander 2004] N. Bergeron and T. Gelander, "A note on local rigidity", *Geom. Dedicata* **107** (2004), 111–131. MR 2110758 Zbl 1062.22028

[Cooper et al. 2000] D. Cooper, C. D. Hodgson, and S. P. Kerckhoff, *Three-dimensional orbifolds and cone-manifolds*, MSJ Memoirs **5**, Mathematical Society of Japan, Tokyo, 2000. MR 1778789 Zbl 0955.57014

[Deraux 2014] M. Deraux, "A 1-parameter family of spherical CR uniformizations of the figure eight knot complement", preprint, 2014. To appear in *Geom. Topol.* arXiv 1410.1198

[Deraux 2015] M. Deraux, "On spherical CR uniformization of 3-manifolds", *Exp. Math.* **24**:3 (2015), 355–370. MR 3359222

[Deraux and Falbel 2015] M. Deraux and E. Falbel, "Complex hyperbolic geometry of the figure eight knot", *Geometry & Topology* **19**:1 (2015), 237–293. Zbl 1335.32028

[Falbel et al. 2016] E. Falbel, A. Guilloux, P.-V. Koseleff, F. Rouillier, and M. Thistlethwaite, "Character varieties for SL($\mathbf{3}$, $\mathbb{C}$): the figure eight knot", *Exp. Math.* **25**:2 (2016), 219–235. MR 3463570

[Genzmer 2010] J. Genzmer, *Sur les triangulations des structures CR-sphériques*, Ph.D. thesis, Université Pierre et Marie Curie, 2010, available at https://tel.archives-ouvertes.fr/tel-00502363/document.

[Goldman 1999] W. M. Goldman, *Complex hyperbolic geometry*, Clarendon, New York, 1999. MR 1695450 Zbl 0939.32024

[Goldman 2010] W. M. Goldman, "Locally homogeneous geometric manifolds", pp. 717–744 in *Proceedings of the International Congress of Mathematicians, II*, edited by R. Bhatia et al., Hindustan Book Agency, New Delhi, 2010. MR 2827816 Zbl 1234.57001

[Parker et al. 2016] J. R. Parker, J. Wang, and B. Xie, "Complex hyperbolic (3, 3, n) triangle groups", *Pacific J. Math.* **280**:2 (2016), 433–453. MR 3453571 Zbl 06537059

[Schwartz 2007] R. E. Schwartz, *Spherical CR geometry and Dehn surgery*, Annals of Mathematics Studies **165**, Princeton University Press, 2007. MR 2286868 Zbl 1116.57016

[Thurston 2002] W. P. Thurston, "The geometry and topology of three-manifolds", lecture notes, Princeton Univ., 2002, available at http://www.msri.org/gt3m/.

Received September 14, 2015. Revised January 21, 2016.

MIGUEL ACOSTA
UMR 7586 CNRS
UNIVERSITÉ PIERRE ET MARIE CURIE
4 PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE
miguel.acosta@imj-prg.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.283

DEGENERATE FLAG VARIETIES AND SCHUBERT VARIETIES: A CHARACTERISTIC FREE APPROACH

GIOVANNI CERULLI IRELLI, MARTINA LANINI AND PETER LITTELMANN

We consider the PBW filtrations over $\mathbb{Z}$ of the irreducible highest weight modules in type $\mathtt{A}_n$ and $\mathtt{C}_n$. We show that the associated graded modules can be realized as Demazure modules for group schemes of the same type and doubled rank. We deduce that the corresponding degenerate flag varieties are isomorphic to Schubert varieties in any characteristic.

Introduction

Introduced by Evgeny Feigin in 2010, degenerate flag varieties naturally arise from a representation theoretic context. In fact, given a finite dimensional, highest weight irreducible module $V(\lambda)$ for a simple finite dimensional, complex Lie algebra, the corresponding degenerate flag variety $\mathcal{F}\ell(\lambda)^a$ is the closure of a certain highest weight orbit in the projectivization of $V(\lambda)^a$, a degenerate version of $V(\lambda)$.

If the algebra one starts with is of type $\mathtt{A}_n$ or $\mathtt{C}_n$, it was shown in [Cerulli Irelli and Lanini 2015] that, surprisingly, degenerate flag varieties can be realized as Schubert varieties in a partial flag variety of the same type and bigger rank. It is hence natural to ask whether also the modules $V(\lambda)^a$ are isomorphic to some already investigated objects. The aim of this paper is to address such a question and provide a positive answer to it. Feigin's degeneration procedure can be carried out over $\mathbb{Z}$ — see [Feigin et al. 2013] — and it is in this generality that we decided to approach the problem.

Our main theorem is the realization of $V(\lambda)^a$ as a Demazure module for a group scheme of the same type and doubled rank. This fact allows us to recover, as a corollary, the above-mentioned realization of $\mathcal{F}\ell(\lambda)^a$ as a Schubert variety. While the arguments in [Cerulli Irelli and Lanini 2015] relied on a linear algebraic description of the degenerate flag variety due to Feigin [2012], the proof we obtain here only uses the definition of $\mathcal{F}\ell(\lambda)^a$ as a closure of a highest weight orbit; hence it is more conceptual.

In what follows, we describe in more detail the main results of this article.

For simplicity let us start with the complex algebraic group $\mathrm{SL}_n(\mathbb{C})$ and its Lie algebra $\mathfrak{g} = \mathfrak{sl}_n$. We fix a Cartan decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$, where $\mathfrak{n}$ is

MSC2010: primary 17B10; secondary 17B35, 16S30, 20G40, 14N15.
Keywords: Schubert varieties, degenerate flag varieties, Demazure modules.

the subalgebra of strictly upper triangular matrices, $\mathfrak{h}$ is the Cartan subalgebra consisting of diagonal matrices and $\mathfrak{n}^-$ is the subalgebra of strictly lower triangular matrices. Let $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$ be the corresponding Borel subalgebra of $\mathfrak{g}$ and let $B \subset G$ be the Borel subgroup with Lie algebra $\mathfrak{b}$.

We use the notation $\tilde{B}$, $\tilde{\mathfrak{b}}$, $\tilde{\mathfrak{n}}^+$, $\tilde{\mathfrak{h}}$, and $\tilde{\mathfrak{n}}^-$ for the corresponding subgroup of $\tilde{G} = \mathrm{SL}_{2n}(\mathbb{C})$ and subalgebras of $\tilde{\mathfrak{g}} = \mathfrak{sl}_{2n}$. Let $\mathfrak{n}^{-,a} \subset \mathfrak{sl}_{2n}$ and $N^{-,a} \subseteq \mathrm{SL}_{2n}(\mathbb{C})$ be the following commutative Lie subalgebra and commutative unipotent subgroup, respectively:

$$\mathfrak{n}^{-,a} := \left\{ \begin{pmatrix} 0 & N \\ 0 & 0 \end{pmatrix} \in \mathfrak{sl}_{2n} \mid N \in \mathfrak{n}^- \right\}, \quad N^{-,a} := \left\{ \begin{pmatrix} \mathbb{1} & N \\ 0 & \mathbb{1} \end{pmatrix} \in \mathrm{SL}_{2n} \mid N \in \mathfrak{n}^- \right\}. \tag{1}$$

We view $\mathfrak{n}^{-,a}$ as the abelianization of $\mathfrak{n}^-$, i.e., we have the canonical vector space isomorphism between the two vector spaces, but $\mathfrak{n}^{-,a}$ is endowed with the trivial Lie bracket. The enveloping algebra of $\mathfrak{n}^{-,a}$ is $S^\bullet(\mathfrak{n}^{-,a})$. The embedding $\mathfrak{n}^{-,a} \hookrightarrow \tilde{\mathfrak{b}}$ induces an embedding $S^\bullet(\mathfrak{n}^{-,a}) \hookrightarrow U(\tilde{\mathfrak{b}})$, so any $U(\tilde{\mathfrak{b}})$-module inherits in a natural way the structure of a $S^\bullet(\mathfrak{n}^{-,a})$-module.

A well investigated class of $U(\tilde{\mathfrak{b}})$-modules are the Demazure modules: let μ be a dominant integral weight for $\tilde{\mathfrak{g}}$ and let $\tilde{V}(\mu)$ be the corresponding irreducible representation. For an element w of the Weyl group $\tilde{W}$ of $\tilde{\mathfrak{g}}$, the weight space $\tilde{V}(\mu)_{w\mu}$ of weight $w\mu$ is one-dimensional; fix a generator $v_{w\mu}$. Recall that the Demazure submodule $\tilde{V}(\mu)_w$ is by definition the cyclic $U(\tilde{\mathfrak{b}})$-module generated by $v_{w\mu}$, i.e., $\tilde{V}(\mu)_w = U(\tilde{\mathfrak{b}}).v_{w\mu}$, and the Schubert variety $X(w)$ is the closure of the orbit $\tilde{B}.[v_{w\mu}] \subseteq \mathbb{P}(\tilde{V}(\mu))$.

A special class of $S^\bullet(\mathfrak{n}^{-,a})$-modules has been investigated in [Feigin et al. 2011a; 2011b]. Let λ be a dominant integral weight for $\mathfrak{g}$, let $V(\lambda)$ be the corresponding irreducible representation and fix a highest weight vector v_λ. The PBW filtration on $U(\mathfrak{n}^-)$ induces a filtration on the cyclic $U(\mathfrak{n}^-)$-module $V(\lambda) = U(\mathfrak{n}^-).v_\lambda$, and the associated graded space $V^a(\lambda) := \operatorname{gr} V(\lambda)$ becomes a module for the associated graded algebra $S^\bullet(\mathfrak{n}^-) := \operatorname{gr} U(\mathfrak{n}^-) \simeq S^\bullet(\mathfrak{n}^{-,a})$.

The action of $\mathfrak{n}^{-,a}$ on $V^a(\lambda)$ can be integrated to an action of $N^{-,a}$. In analogy with the classical case we call the closure of the orbit $\mathcal{F}^a_\lambda := \overline{N^{-,a}.[v_\lambda]} \subseteq \mathbb{P}(V^a(\lambda))$ the degenerate flag variety.

The aim of this article is to connect these two constructions and extend the results in [Cerulli Irelli and Lanini 2015] to an algebraically closed field k of arbitrary characteristic. In fact, the results hold even over $\mathbb{Z}$. For simplicity, we formulate them in the introduction for an algebraically closed field k. In the following, we consider the case $G = \mathrm{SL}_n(k)$ and $\tilde{G} = \mathrm{SL}_{2n}(k)$, respectively $G = \mathrm{Sp}_{2m}(k)$, and $\tilde{G} = \mathrm{Sp}_{4m}(k)$, and we replace the irreducible module of highest weight λ by the Weyl module of highest weight λ, using the same notation $V(\lambda)$. For the precise description of the highest weight $\Psi(\lambda)$, see Definitions 2.1 and 5.3; for a description of the Weyl group element $\tau \in \widetilde{W}$, see Definitions 2.2 and 5.1; and for the construction of the Lie algebra $\mathfrak{n}^{-,a}$ in the symplectic case, see Section 5. For a dominant G-weight λ let λ^* be the dual dominant weight, so for the symplectic case we have $\lambda = \lambda^*$, and in the SL_n case we have $\lambda^* = \sum_{i=1}^{n-1} m_i \omega_{n-i}$ for $\lambda = \sum_{i=1}^{n-1} m_i \omega_i$ in the notation as in [Bourbaki 1968].

Theorem. *Let λ be a dominant G-weight.*

(i) *The Demazure submodule $\tilde{V}_k(\Psi(\lambda^*))_\tau$ of the $\tilde{G}$-module $\tilde{V}_k(\Psi(\lambda^*))$ is isomorphic, as an $\mathfrak{n}^{-,a}$-module, to the abelianized module $V^a(\lambda)$.*

(ii) *The Schubert variety $X(\tau) \subset \mathbb{P}(\tilde{V}(\Psi(\lambda^*))_\tau)$ is isomorphic to the degenerate flag variety $\mathcal{F}^a(\lambda)$, and this isomorphism induces an $S^\bullet(\mathfrak{n}^{-,a})$-module isomorphism*

$$H^0(X(\tau), \mathcal{L}_{\Psi(\lambda^*)}) \simeq (V^a(\lambda))^*.$$

Using the isomorphism above, we deduce the defining relations for $V^a(\lambda)$ from the defining relations of the Demazure module. Translated back into the language of the abelianized algebras we get the following: in the SL_n case, let $R^{++} = R^+$ be the set of positive roots, and in the symplectic case, set

$$R^{++} = \{\epsilon_i - \epsilon_j \mid 1 \le i < j \le m\} \cup \{2\epsilon_i \mid 1 \le i \le m\}.$$

Corollary. *The abelianized module $V^a(\lambda)$ is isomorphic as a cyclic $S^\bullet(\mathfrak{n}^{-,a})$-module to $S^\bullet(\mathfrak{n}^{-,a})/I(\lambda)$, where $I(\lambda)$ is the ideal*:

$$I(\lambda) = S^\bullet(\mathfrak{n}^{-,a})\big(U(\mathfrak{n}^+) \circ \mathrm{span}\{f_\alpha^{(\langle\lambda,\alpha^\vee\rangle+1)} \mid \alpha \in R^{++}\}\big) \subseteq S^\bullet(\mathfrak{n}^{-,a}).$$

The identification of the degenerate flag variety as a Schubert variety implies the following corollary immediately; see also [Feigin and Finkelberg 2013; Feigin et al. 2014].

Corollary. *The degenerate flag variety $\mathcal{F}^a(\lambda)$ is projectively normal, and it has rational singularities.*

1. Some special commutative unipotent subgroups

Let k be a field. Given a subspace $\mathfrak{N} \subseteq M_n(k)$ and a vector space automorphism $\eta : \mathfrak{N} \to \mathfrak{N}$, denote by $\mathfrak{N}_\eta^a \subseteq M_{2n}(k)$ respectively $N_\eta^a \subseteq \mathrm{GL}_{2n}(k)$ the following commutative nilpotent Lie subalgebra of $M_{2n}(k)$, respectively commutative unipotent subgroup of $\mathrm{GL}_{2n}(k)$:

$$\mathfrak{N}_\eta^a := \left\{ \begin{pmatrix} 0 & \eta(A) \\ 0 & 0 \end{pmatrix} \,\middle|\, A \in \mathfrak{N} \right\}, \quad N_\eta^a := \left\{ \begin{pmatrix} \mathbb{1} & \eta(A) \\ 0 & \mathbb{1} \end{pmatrix} \,\middle|\, A \in \mathfrak{N} \right\}.$$

If $\mathfrak{N} \subseteq M_n(k)$ is a Lie subalgebra, then we think of $\mathfrak{N}_\eta^a$ as an abelianized version of $\mathfrak{N}$. Similarly one may think of N_η^a as an abelianized version of a subgroup $N \subseteq \mathrm{GL}_{2n}(k)$. We will be more precise in the following examples.

Example 1.1. Let k be an algebraically closed field of characteristic zero. We fix as a maximal torus $T \subset \mathrm{SL}_n$ the subgroup of diagonal matrices, and let B be the Borel subgroup of upper triangular matrices. Let us denote by $\mathfrak{sl}_n$, $\mathfrak{b}$, and $\mathfrak{h}$ the corresponding Lie algebras and let $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ be the Cartan decomposition. The choice of a maximal torus and a Borel subgroup as above determines the set of positive roots Φ^+ and hence, according to the adjoint action of $\mathfrak{h}$, the root space decomposition $\mathfrak{n}^- = \bigoplus_{\alpha\in\Phi^+} \mathfrak{n}^-_{-\alpha}$. In this example we set $\mathfrak{N} = \mathfrak{n}^-$, and $N = U^-$ is the unipotent radical of the opposite Borel subgroup B^-. The map η is the identity map, so we just omit it. Henceforth, we write $\mathfrak{n}^{-,a}$ for $\mathfrak{N}^a \subset \mathfrak{sl}_{2n}(k)$ and $N^{-,a}$ for $N^a \subset \mathrm{SL}_{2n}$.

Note that $\mathfrak{n}^{-,a} \subset \mathfrak{sl}_{2n}$ is a Lie subalgebra of the Borel subalgebra $\tilde{\mathfrak{b}} \subset \mathfrak{sl}_{2n}$ and $N^{-,a}$ is an abelian subgroup of the Borel subgroup $\tilde{B} \subset \mathrm{SL}_{2n}$ (of upper triangular matrices). We can think of $N^{-,a}$ as an abelianized version of U^-.

The subgroup $N^{-,a}$, as well as the Lie algebra $\mathfrak{n}^{-,a}$, is stable under conjugation with respect to the maximal torus $\tilde{T} \subset \mathrm{SL}_{2n}$, where $\tilde{T} \subset \mathrm{SL}_{2n}$ consists of the diagonal matrices. The group $N^{-,a}$ hence decomposes as a product of root subgroups of the group SL_{2n}, and $\mathfrak{n}^{-,a}$ decomposes into the direct sum of root subspaces for the Lie algebra $\mathfrak{sl}_{2n}$. We get an induced map $\phi : \Phi^+ \to \tilde{\Phi}^+$ between the set of positive roots of $\mathfrak{sl}_n$ and the positive roots of $\mathfrak{sl}_{2n}$, such that $\mathfrak{n}^{-,a} = \bigoplus_{\alpha\in\Phi^+} \mathfrak{n}^{-,a}_{\phi(\alpha)}$.

Example 1.2. Let k be an algebraically closed field of characteristic 0. Let $\{e_1, \dots, e_{2n}\}$ be the canonical basis of k^{2n}, and fix a nondegenerate skew symmetric form by the conditions $\langle e_i, e_j\rangle = \delta_{j,2n-i+1} = -\langle e_j, e_i\rangle$ for $1 \le i \le n,\ 1 \le j \le 2n$. Let Sp_{2m} be the associated symplectic group. By the choice of the form we can fix as a Borel subgroup B the subgroup of upper triangular matrices in Sp_{2m} and let T be its maximal torus consisting of diagonal matrices. Let us denote by $\mathfrak{sp}_{2m}$, $\mathfrak{b}$, and $\mathfrak{h}$ the corresponding Lie algebras and let $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ be the Cartan decomposition.

The choice of the torus and the Borel subgroup as above determines a set of positive roots Φ^+ and hence, according to the adjoint action of $\mathfrak{h}$, the root space decomposition $\mathfrak{n}^- = \bigoplus_{\alpha\in\Phi^+} \mathfrak{n}^-_{-\alpha}$. In this example we set $\mathfrak{N} = \mathfrak{n}^-$, and $N = U^-$ is the unipotent radical of the opposite Borel subgroup B^-. Let $\eta : \mathfrak{n}^- \to \mathfrak{n}^-$ be the linear map sending a matrix $(m_{i,j})_{1\le i,j\le 2n}$ to the matrix $(m'_{i,j})_{1\le i,j\le 2n}$, where $m'_{i,j} = m_{i,j}$ if $i \le n$ or $j \le n$ and $m'_{i,j} = -m_{i,j}$ if both indices are strictly larger than n. We write henceforth $\mathfrak{n}^{-,a}_\eta$ for $\mathfrak{N}^a_\eta \subset \mathfrak{sp}_{4n}(k)$ and $N^{-,a}_\eta$ for $N^a_\eta \subset \mathrm{Sp}_{4n}$.

Note that $\mathfrak{n}^{-,a}_\eta \subset \mathfrak{sp}_{4n}$ is a Lie subalgebra of the Borel subalgebra (of upper triangular matrices) $\tilde{\mathfrak{b}} \subset \mathfrak{sp}_{4n}$ and $N^{-,a}_\eta$ is an abelian subgroup of the Borel subgroup $\tilde{B} \subset \mathrm{Sp}_{4n}$ (of upper triangular matrices). We can think of $N^{-,a}_\eta$ as an abelianized version of U^-.

The subgroup $N^{-,a}_\eta$ is stable under conjugation with respect to the maximal torus $\tilde{T} \subset \mathrm{Sp}_{4n}$, where $\tilde{T} \subset \mathrm{Sp}_{4n}$ consists of the diagonal matrices. The group $N^{-,a}_\eta$ hence decomposes as a product of root subgroups of the group Sp_{4n} and $\mathfrak{n}^{-,a}$ decomposes into the direct sum of root subspaces for the Lie algebra $\mathfrak{sp}_{4n}$. We get an induced map $\phi : \Phi^+ \to \tilde{\Phi}^+$ between the set of positive roots of $\mathfrak{sp}_{2n}$ and the positive roots of $\mathfrak{sp}_{4n}$, such that $\mathfrak{n}_\eta^{-,a} = \bigoplus_{\alpha\in\Phi^+} \mathfrak{n}^{-,a}_{\phi(\alpha),\eta}$.

Example 1.3. To get a characteristic free approach for G as above, let $G_{\mathbb{Z}}$ be a split and connected simple algebraic $\mathbb{Z}$-group of type $\mathtt{A}_n$ or $\mathtt{C}_\mathtt{n}$. For any commutative ring A set $G_A = (G_{\mathbb{Z}})_A$, and for a field set $G = G_k$, for this and the following; see also [Jantzen 1987]. Then G_k is for any algebraically closed field a reduced k-group, and it is connected and reductive. Its Lie algebra $\mathrm{Lie}(G_{\mathbb{Z}})$ is a free Lie algebra of finite rank and $\mathrm{Lie}\, G_k = \mathrm{Lie}(G_{\mathbb{Z}}) \otimes_{\mathbb{Z}} k$. Let $T_{\mathbb{Z}} \subset G_{\mathbb{Z}}$ be a split maximal torus and set $T_A = (T_{\mathbb{Z}})_A$ for any ring A and $T = T_k$. We have a root space decomposition $\mathrm{Lie}\, G = \mathrm{Lie}\, T \oplus \bigoplus_{\alpha\in\Phi}(\mathrm{Lie}\, G)_\alpha$ where $(\mathrm{Lie}\, G)_\alpha = (\mathrm{Lie}\, G_{\mathbb{Z}})_\alpha \otimes_{\mathbb{Z}} k$, and corresponding root subgroups (defined over $\mathbb{Z}$) $x_\alpha : \mathbb{G}_a \to G$ such that the tangent map $\mathrm{d}x_\alpha$ induces an isomorphism between the Lie algebra of the additive group $\mathbb{G}_a$ and $(\mathrm{Lie}\, G)_\alpha$. The functor which associates to any commutative ring A the group $x_\alpha(\mathbb{G}_a(A)) = x_\alpha(A)$ is a closed subgroup of G denoted by U_α, and we have $\mathrm{Lie}(U_\alpha) = (\mathrm{Lie}\, G)_\alpha$. Over $\mathbb{Z}$ we denote the corresponding subgroup by $U_{\alpha,\mathbb{Z}}$, and over a field k we have $U_\alpha = (U_{\alpha,\mathbb{Z}})_k$.

The construction described in Examples 1.1 and 1.2 makes (in this language) sense over $\mathbb{Z}$ or over any field. As before, let $\tilde{G}$ be the group of the same type but twice the rank, we denote the corresponding Borel subgroup, maximal torus, etc. by $\tilde{B}$, $\tilde{T}$, etc. The construction in the examples above associates to every root $\alpha \in \Phi$ a root $\phi(\alpha)$ in the root system of $\tilde{G}$. For the $\mathbb{Z}$-group $G_{\mathbb{Z}}$ we have the subgroup $U^-_{\mathbb{Z}}$ and the Lie algebra $\mathfrak{n}^-_{\mathbb{Z}} = \bigoplus_{\alpha\in\Phi^+} \mathfrak{n}^-_{\mathbb{Z},-\alpha}$, and we associate to this pair a new pair given by a commutative subgroup $N^{-,a}_\eta$ of the $\mathbb{Z}$-group $\tilde{G}_{\mathbb{Z}}$ and an abelian Lie algebra $\mathfrak{n}^{-,a}_{\mathbb{Z},\eta}$. The first is the subgroup of the Borel subgroup $\tilde{B}_{\mathbb{Z}} \subset \tilde{G}_{\mathbb{Z}}$ generated by the commuting root subgroups $U_{\phi(\alpha),\mathbb{Z}}$, $\alpha \in \Phi^+$, and the second is the abelian Lie algebra $\mathfrak{n}^{-,a}_{\mathbb{Z},\eta} = \bigoplus_{\alpha\in\Phi^+}(\mathrm{Lie}\, G)^{\sim}_{\mathbb{Z},\phi(\alpha)}$ given as the sum of root subalgebras.

2. A special Schubert variety: the SL_n case

We want to realize in the situation of Example 1.1 the abelianized representation $V(\lambda)^a$ for $N^{-,a}_\eta$ as a Demazure submodule of an irreducible representation for the group SL_{2n}.

2A. ***A special Weyl group element.*** Let $\widetilde{W}$ be the Weyl group of $\mathrm{SL}_{2n}(\mathbb{C})$; it is the symmetric group $\mathcal{S}_{2n}$ generated by the transpositions s_i, $i = 1, \dots, 2n-1$. Let $\mathfrak{h} \subset \mathfrak{g} = \mathfrak{sl}_n$ (respectively, $\tilde{\mathfrak{h}} \subset \tilde{\mathfrak{g}} = \mathfrak{sl}_{2n}$) be the Cartan subalgebra of traceless complex diagonal matrices. For an element $\alpha \in \mathfrak{h}^*$ and an element $h \in \mathfrak{h}$ we denote by $\langle h, \alpha\rangle$ the evaluation of α in h. Let $\{\epsilon_1, \dots, \epsilon_n\}$ be the elements of the dual vector space $\mathfrak{h}^*$ such that $\langle h, \epsilon_i\rangle$ is the i-th entry in the diagonal matrix $h \in \mathfrak{h}$. We use the same notation $\langle \tilde{h}, \tilde{\alpha}\rangle$ for elements $\tilde{h} \in \tilde{\mathfrak{h}}$ and $\tilde{\alpha} \in \tilde{\mathfrak{h}}^*$, and the linear forms $\{\tilde{\epsilon}_1, \dots, \tilde{\epsilon}_{2n}\}$ in $\tilde{\mathfrak{h}}^*$ are defined as above.

The roots of $\mathfrak{g}$ (resp., $\tilde{\mathfrak{g}}$) are the elements $\alpha_{i,j} := \epsilon_i - \epsilon_j$ (resp., $\tilde{\alpha}_{i,j} = \tilde{\epsilon}_i - \tilde{\epsilon}_j$) for $i \neq j$. We choose as a Borel subalgebra of $\mathfrak{g}$ the subalgebra $\mathfrak{b}$ of upper triangular matrices. The corresponding simple roots are $\alpha_1, \cdots, \alpha_{n-1}$ given by $\alpha_i := \alpha_{i,i+1}$. For every root α, we denote by $\alpha^\vee$ its coroot: this is the unique element of $\mathfrak{h}$ such that the reflection $s_\alpha \in \mathfrak{h}^*$ along α acts as $s_\alpha(\lambda) = \lambda - \langle \alpha^\vee, \lambda \rangle \alpha$. Moreover we denote by E_α the corresponding root vector. We denote by $\omega_i = \epsilon_1 + \cdots + \epsilon_i$ (resp., $\tilde{\omega}_i = \tilde{\epsilon}_1 + \cdots + \tilde{\epsilon}_i$) the i-th fundamental weight of $\mathfrak{g}$ (resp., $\tilde{\mathfrak{g}}$), where $i = 1, 2, \ldots, n-1$ (resp., $i = 1, \ldots, 2n-1$). They are characterized by the property $\langle \alpha_i^\vee, \omega_j \rangle = \delta_{i,j}$.

Definition 2.1. Let $\Psi : \mathfrak{h}^* \to \tilde{\mathfrak{h}}^*$ be the linear map defined on the weight lattice by

$$\Psi\left(\sum_{i=1}^{n-1} a_i \omega_i\right) := \sum_{i=1}^{n-1} a_i \tilde{\omega}_{2i}.$$

Note that Ψ sends dominant weights to dominant weights. For every fundamental weight $\tilde{\omega}_k$, we denote the corresponding parabolic subgroup by $P_{\tilde{\omega}_k}$ and by $\tilde{W}_{\tilde{\omega}_k}$ the corresponding subgroup of $\tilde{W}$ which is the Weyl group of the semisimple part of $P_{\tilde{\omega}_k}$. Note that $\tilde{W}_{\tilde{\omega}_k}$ is generated by all the simple transpositions s_i but s_k. Let $\rho = \omega_1 + \cdots + \omega_{n-1}$. Then $\Psi(\rho) = \tilde{\omega}_2 + \tilde{\omega}_4 + \cdots + \tilde{\omega}_{2n-2}$. The parabolic subgroup $Q = P_{\tilde{\omega}_2 + \cdots + \tilde{\omega}_{2n-2}}$ which is the stabilizer of $\Psi(\rho)$ will play an important role. The Weyl group of the semisimple part of Q is denoted by $\tilde{W}^J$.

Definition 2.2. We define in the Weyl group $\widetilde{W}$ the element τ by

$$\tau = (s_n s_{n+1} \cdots s_{2n-3} s_{2n-2})(s_{n-1} s_n \cdots s_{2n-4}) \cdots (s_4 s_5 s_6)(s_3 s_4) s_2. \tag{2}$$

It is easy to see that the decomposition is reduced and τ is a minimal length representative in its class in $\widetilde{W}/\tilde{W}^J$. Another description of τ can be given by viewing τ as a permutation of the set $\{1, \ldots, 2n\}$:

$$\tau(t) = \begin{cases} n+k & \text{if } t = 2k, \\ k & \text{if } t = 2k-1, \end{cases} \tag{3}$$

for $k = 1, 2, \ldots, n$. It follows now immediately from (3):

Lemma 2.3. *In the irreducible* $\mathrm{SL}_{2n}(\mathbb{C})$*-representation* $\tilde{V}(\tilde{\omega}_{2i}) = \bigwedge^{2i} \mathbb{C}^{2n}$ *let* v_0 *be the highest weight vector* $v_0 = e_1 \wedge e_2 \wedge \cdots \wedge e_{2i}$*. Then (up to a sign),*

$$\tau(v_0) = v_\tau = e_1 \wedge e_2 \wedge \cdots \wedge e_i \wedge e_{n+1} \wedge e_{n+2} \wedge \cdots \wedge e_{n+i}.$$

Let $\lambda = b_1 \epsilon_1 + \cdots + b_{n-1} \epsilon_{n-1}$, with $b_1 \geq \cdots \geq b_{n-1} \geq 0$, be a dominant weight for $\mathrm{SL}_n(\mathbb{C})$. The next result follows directly from Lemma 2.3.

Lemma 2.4. $\tau(\Psi(\lambda)) = b_1 \tilde{\epsilon}_1 + \cdots + b_{n-1} \tilde{\epsilon}_{n-1} + b_1 \tilde{\epsilon}_{n+1} + \cdots + b_{n-1} \tilde{\epsilon}_{2n-1}$. □

In Example 1.1 we have introduced a map $\phi : \Phi^+ \to \tilde{\Phi}^+$ between the positive roots of $\mathfrak{sl}_n$ and the positive roots of $\mathfrak{sl}_{2n}$. Note that the image of $\alpha = \epsilon_i - \epsilon_j$, $1 \le i < j \le n$ is the root $\phi(\alpha) = \tilde{\epsilon}_j - \tilde{\epsilon}_{n+i}$.

Lemma 2.5. (i) *Let λ be a dominant weight for $\mathrm{SL}_n(\mathbb{C})$, and let $\tilde{\alpha}$ be a positive SL_{2n}-root. Then $\langle \tilde{\alpha}^\vee, \tau(\Psi(\lambda))\rangle < 0$ only if the root space of $\tilde{\alpha}$ lies in $\mathfrak{n}^{-,a}$.*

(ii) *Let λ be a dominant SL_n-weight, and let $\alpha = \epsilon_i - \epsilon_j$ be a positive SL_n-root. Then*
$$\langle \alpha^\vee, \lambda\rangle = -\langle \phi(\alpha)^\vee, \tau(\Psi(\lambda))\rangle.$$

(iii) *Let λ be a dominant weight for $\mathrm{SL}_n(\mathbb{C})$, and let $\tilde{\alpha} = \tilde{\epsilon}_p - \tilde{\epsilon}_q$ be a positive SL_{2n}-root. Then $E_{\tilde{\alpha}} v_\tau \neq 0$ in $\tilde{V}(\Psi(\lambda))$ only if $\tilde{\alpha}$ is of the form $\tilde{\alpha} = \tilde{\epsilon}_j - \tilde{\epsilon}_{n+i}$, $1 \le i < j \le n$ and $\langle (\epsilon_i - \epsilon_j)^\vee, \lambda\rangle > 0$.*

Proof. Let $\tilde{\alpha} = \tilde{\epsilon}_i - \tilde{\epsilon}_j$ be a positive root. By Lemma 2.4, for $\lambda = b_1\epsilon_1 + \cdots + b_{n-1}\epsilon_{n-1}$, we get
$$\langle \tilde{\alpha}^\vee, \tau(\Psi(\lambda))\rangle = \begin{cases} b_i - b_j \ge 0 & \text{if } 1 \le i < j \le n, \\ b_i - b_{j-n} \ge 0 & \text{if } 1 \le i \le n \text{ and } n+i \le j \le 2n, \\ b_i - b_{j-n} \le 0 & \text{if } 1 \le i \le n \text{ and } n+1 \le j < n+i, \\ b_{i-n} - b_{j-n} \ge 0 & \text{if } n+1 \le i < j \le 2n, \end{cases}$$
which proves the lemma. □

The decomposition in (2) is reduced, but if we apply τ to a fundamental weight, then it is possible to omit some of the reflections. A simple calculation shows:

Lemma 2.6. *Let $\tilde{\omega}_{2i}$ be the $2i$-th fundamental weight for $\mathrm{SL}_{2n}(\mathbb{C})$. Then*
$$\tau(\tilde{\omega}_{2i}) = (s_n s_{n+1} \cdots s_{n+i-1}) \cdots (s_{i+2} \cdots s_{2i+1})(s_{i+1} \cdots s_{2i-1} s_{2i})(\tilde{\omega}_{2i}).$$

Let $L(i)$ be the semisimple part of the Levi subgroup of $\mathrm{SL}_{2n}(\mathbb{C})$ associated with the simple roots $\tilde{\alpha}_{i+1}, \tilde{\alpha}_{i+2}, \ldots, \tilde{\alpha}_{i+n-1}$, denote by $\mathfrak{l}(i)$ the Lie algebra of $L(i)$. Note that $L(i)$ is isomorphic to $\mathrm{SL}_n(\mathbb{C})$. Let $\varpi_1, \ldots, \varpi_{n-1}$ be the fundamental weights of $L(i)$, the enumeration is such that the simple root $\tilde{\alpha}_{i+j}$ of $L(i) \subseteq \mathrm{SL}_{2n}(\mathbb{C})$ corresponds to ϖ_j.

The restriction of $\tilde{\omega}_{2i}$ to $L(i)$ is ϖ_i. Let $W^{L(i)}$ be the Weyl group of $L(i)$, we can identify it with the subgroup of the Weyl group of SL_{2n} generated by the reflections $s_{i+1}, s_{i+2}, \ldots, s_{i+n-1}$. Using Lemma 2.6, it is easy to see:

Lemma 2.7. *A reduced decomposition of the longest word of $W^{L(i)}$ modulo the stabilizer $W^{L(i)}_{\varpi_i}$ of ϖ_i in $W^{L(i)}$ is given by*
$$(s_n s_{n+1} \cdots s_{n+i-1}) \cdots (s_{i+2} \cdots s_{2i+1})(s_{i+1} \cdots s_{2i-1} s_{2i}).$$

3. The fundamental representations: the $\mathfrak{sl}_n$ case

We switch now to Lie algebras and hyperalgebras over $\mathbb{Z}$. Fix a Chevalley basis for the Lie algebra $\mathfrak{g}_{\mathbb{Z}} = \mathfrak{sl}_{n,\mathbb{Z}} \subset \mathfrak{sl}_{n,\mathbb{C}}$:

$$\{f_\alpha, e_\alpha \mid \alpha \in \Phi^+\} \cup \{h_1, \dots, h_{n-1}\},$$

where $f_\alpha \in \mathfrak{g}_{\mathbb{Z},-\alpha}$, $e_\alpha \in \mathfrak{g}_{\mathbb{Z},\alpha}$, and $h_i \in \mathfrak{h}_{\mathbb{Z}}$. For any $m \in \mathbb{Z}_{\geq 1}$, we define the following elements in $U(\mathfrak{g})$:

$$e_\alpha^{(m)} = \frac{e_\alpha^m}{m!}, \qquad f_\alpha^{(m)} = \frac{f_\alpha^m}{m!}, \qquad \binom{h_i}{m} = \frac{h_i(h_i-1)\cdots(h_i-m+1)}{m}, \tag{4}$$

and for $m = 0$ we set $e_\alpha^{(0)} = f_\alpha^{(0)} = \binom{h_i}{0} = 1$. Recall that the hyperalgebra $U_{\mathbb{Z}}(\mathfrak{sl}_n)$ of $(\mathrm{SL}_n)_{\mathbb{Z}}$ is the $\mathbb{Z}$-subalgebra of the complex enveloping algebra $U(\mathfrak{sl}_n)$ generated by the elements defined in (4). We will use capital letters to denote the Chevalley basis elements for $\mathfrak{sl}_{2n,\mathbb{Z}}$ (e.g., $E_{\tilde{\alpha}}$, $F_{\tilde{\alpha}}$, H_i) and the generators of the hyperalgebra $U_{\mathbb{Z}}(\mathfrak{sl}_{2n})$ (e.g., $E_{\tilde{\alpha}}^{(m)}$, $F_{\tilde{\alpha}}^{(m)}$, $\binom{H_i}{m}$). Similarly, let $U_{\mathbb{Z}}(\tilde{\mathfrak{b}})$ be the subalgebra generated by all $E_{\tilde{\alpha}}^{(m)}$ for $m \geq 0$ and $\tilde{\alpha} > 0$, and all $\binom{H_i}{m}$ for $i = 1, \dots, 2n-1$ and $m \geq 0$. Denote by $U_{\mathbb{Z}}(\mathfrak{l}(i))$ the hyperalgebra associated with $\mathfrak{l}(i)$, i.e., the subalgebra generated by all $F_{\tilde{\alpha}}^{(m)}$, $E_{\tilde{\alpha}}^{(m)}$ for $m \geq 0$ and $\tilde{\alpha} > 0$, a root of the Levi subgroup $L(i)$, and by all $\binom{H_j}{m}$ for $j = i+1, \dots, i+n-1$.

Let μ be a dominant integral weight for $\mathrm{SL}_{2n}(\mathbb{C})$ and denote by $\tilde{V}(\mu)$ the irreducible $\mathrm{SL}_{2n}(\mathbb{C})$-representation of highest weight μ. Fix a highest weight vector v_μ; the corresponding $\mathbb{Z}$-form is $\tilde{V}_{\mathbb{Z}}(\mu) = U_{\mathbb{Z}}(\mathfrak{sl}_{2n})v_\mu$. To define the Demazure module $\tilde{V}_{\mathbb{Z}}(\mu)_w$, fix a representative $\check{w}$ of w in the simply connected Chevalley group associated with $\mathfrak{sl}_{2n,\mathbb{Z}}$ and set $v_w := \check{w}(v_\mu)$. The Demazure module $\tilde{V}_{\mathbb{Z}}(\lambda)_w$ is the cyclic $U_{\mathbb{Z}}(\tilde{\mathfrak{b}})$-subrepresentation $U_{\mathbb{Z}}(\tilde{\mathfrak{b}}).v_w \subseteq \tilde{V}_{\mathbb{Z}}(\mu)$.

Lemma 3.1. *The Demazure module $\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})_\tau$ contained in $\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})$ is the Weyl module $V_{\mathbb{Z}}(\ell\varpi_i)$ of highest weight $\ell\varpi_i$ for $U_{\mathbb{Z}}(\mathfrak{l}(i))$.*

Proof. Consider $\Psi(\ell\omega_i) = \ell\tilde{\omega}_{2i}$ and recall that the restriction of $\tilde{\omega}_{2i}$ to $\mathfrak{l}(i)$ is ϖ_i. So the $U_{\mathbb{Z}}(\mathfrak{l}(i))$-submodule $U_{\mathbb{Z}}(\mathfrak{l}(i))v_{\ell\tilde{\omega}_{2i}} \subseteq \tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})$ is the Weyl module $V_{\mathbb{Z}}(\ell\varpi_i)$ of highest weight $\ell\varpi_i$ for $U_{\mathbb{Z}}(\mathfrak{l}(i))$. Let $U_{\mathbb{Z}}(\mathfrak{b}(i))$ be the subalgebra of $U_{\mathbb{Z}}(\mathfrak{l}(i))$ generated by the $E_{\tilde{\alpha}}^{(m)}$ for $m \geq 0$ and $\tilde{\alpha} > 0$, a root of $\mathfrak{l}(i)$, and all $\binom{H_j}{m}$ for $j = i+1, \dots, i+n-1$.

The Weyl module $V_{\mathbb{Z}}(\ell\varpi_i)$ is a cyclic $U_{\mathbb{Z}}(\mathfrak{b}(i))$-module and is generated by a lowest weight vector of the form $\check{w}_{0,i}(v_{\ell\varpi_i})$, where $\check{w}_{0,i}$ is an appropriate representative (in the Chevalley group associated with $\mathfrak{l}(i)$) of the longest element $w_{0,i}$ of the Weyl group $W^{L(i)}$ of $\mathfrak{l}(i)$. Recall that $W^{L(i)}$ can be identified with the subgroup of $\widetilde{W}$ generated by $s_{i+1}, \dots, s_{n+i-1}$. Now

$$\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})_\tau = U_{\mathbb{Z}}(\tilde{\mathfrak{b}})v_\tau = U_{\mathbb{Z}}(\tilde{\mathfrak{b}})v_{w_{0,i}} = U_{\mathbb{Z}}(\mathfrak{b}(i))v_{w_{0,i}} = V_{\mathbb{Z}}(\ell\varpi_i) \subseteq \tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})$$

by Lemmas 2.5, 2.6, and 2.7. □

The previous result implies in particular:

Corollary 3.2. rank $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau = \operatorname{rank} V_{\mathbb{Z}}(\ell\omega_i)$.

Let $\iota : \mathfrak{sl}_n \to \mathfrak{sl}_n$ be the Chevalley involution defined by $\iota|_{\mathfrak{h}} = -1$ and that ι exchanges e_α and $-f_\alpha$. It follows that $\iota(\mathfrak{n}^-_{\mathbb{Z}}) = \mathfrak{n}^+_{\mathbb{Z}}$, and this map extends to an isomorphism of the corresponding hyperalgebras $\iota : U_{\mathbb{Z}}(\mathfrak{n}^-) \to U_{\mathbb{Z}}(\mathfrak{n}^+)$ and the associated graded versions obtained via the PBW filtration: $\iota : S^\bullet_{\mathbb{Z}}(\mathfrak{n}^-) \to S^\bullet_{\mathbb{Z}}(\mathfrak{n}^+)$.

Let $\lambda = \sum a_j\omega_j$ be a dominant weight and set $\lambda^* := \sum a_j\omega_{n-j}$. Fix a highest weight vector $v_\lambda \in V_{\mathbb{Z}}(\lambda)$ and a lowest weight vector $v_{w_0} \in V_{\mathbb{Z}}(\lambda)$, where w_0 is the longest word in the Weyl group of $\mathfrak{sl}_n$. We get two possible $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-structures on $V_{\mathbb{Z}}(\lambda)$: one uses the PBW filtration on $U_{\mathbb{Z}}(\mathfrak{n}^-)$ to induce, via the highest weight vector, a PBW filtration on $V_{\mathbb{Z}}(\lambda)$ and passes to the associated graded module. One gets the module $V^a_{\mathbb{Z}}(\lambda)$ discussed before. Now one can do the same also for $U_{\mathbb{Z}}(\mathfrak{n}^+)$, once the highest weight vector is replaced by the lowest weight vector. We denote the cyclic $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^+)$-module (generated by the lowest weight vector) by $V^{a,+}_{\mathbb{Z}}(\lambda)$. Now via ι this module also becomes naturally a $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^-)$-module.

Lemma 3.3. *As a $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^-)$-module, $V^{a,+}_{\mathbb{Z}}(\lambda)$ is isomorphic to $V^a_{\mathbb{Z}}(\lambda^*)$.*

Proof. Note that twisting the representation map with the Chevalley involution makes the lowest weight vector (the cyclic generator for the $U(\mathfrak{n}^+)$-action) into a cyclic generator for the $U(\mathfrak{n}^-)$-action. Recall that the Chevalley involution is equal to -1 on $\mathfrak{h}$, so after the twist this is now a highest weight vector of weight $\lambda^* = -w_0(\lambda)$, where w_0 is the longest word in W. Since the construction is compatible with the PBW filtrations with respect to the two algebras, the result for the associated graded modules follows immediately. □

Proposition 3.4. *As $S_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-modules, $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau \simeq V^a_{\mathbb{Z}}(\ell\omega_{n-i})$.*

Proof. Let $\mathfrak{n}^{-,a}_i \subseteq \mathfrak{n}^{-,a}$ be the sum of all root subspaces of roots of the form $\tilde{\epsilon}_k - \tilde{\epsilon}_{n+\ell}$, where $1 \le \ell \le i \le k \le n$ and $\ell \ne k$. This is a commutative Lie subalgebra, which by Lemma 2.5 has the property

$$V_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau = U(\tilde{\mathfrak{b}}).v_\tau = U(\mathfrak{n}^{-,a}_i).v_\tau = S^\bullet(\mathfrak{n}^{-,a}_i).v_\tau.$$

Since $\mathfrak{n}^{-,a}$ is commutative, all root vectors in $\mathfrak{n}^{-,a}$ which are not in $\mathfrak{n}^{-,a}_i$ act trivially on $V_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau$.

Another way to describe $\mathfrak{n}^{-,a}_i$ is as the intersection $\mathfrak{n}^{-,a} \cap \mathfrak{l}(i)$. More precisely, this intersection is the nilpotent radical of the maximal parabolic subalgebra of $\mathfrak{l}(i)$ associated with the fundamental weight ϖ_{n-i}. By Lemma 3.1, we know that $V_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau = U(\mathfrak{b}(i)).v_\tau \simeq V_{\mathbb{Z}}(\ell\varpi_i)$, and since v_τ is a lowest weight vector,

$$V_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau = U(\mathfrak{n}^{-,a}_i).v_\tau \simeq V_{\mathbb{Z}}(\ell\varpi_i).$$

Set $\mathfrak{n}^+(i) = \mathfrak{b}(i) \cap \tilde{\mathfrak{n}}^+$. By the isomorphism between $\mathfrak{l}(i)$ and $\mathfrak{sl}_n$ we can identify $\mathfrak{m}$ with $\mathfrak{n}^+ \subset \mathfrak{sl}_n$. Consider the associated PBW filtration on $V_{\mathbb{Z}}(\ell\varpi_i)$ by applying the PBW filtration of $U(\mathfrak{n}^+(i))$ to the lowest weight vector. Recall that after passing to the associated graded algebra $S^\bullet(\mathfrak{n}^+(i))$, all root vectors not contained in $\mathfrak{n}_i^{-,a}$ act trivially on $V_{\mathbb{Z}}^{a,+}(\ell\varpi_i)$. Remember that we add a "+" to indicate that this is the associated graded space with respect to the filtration by the nilpotent radical of the fixed Borel subalgebra and not, as usual, of the opposite nilpotent algebra. Since ϖ_i and ϖ_{n-i} are cominuscule, $\mathfrak{n}_i^{-,a}$ is commutative and the PBW filtration on $V_{\mathbb{Z}}(\ell\varpi_i)$ is already a grading. It follows that the action of $\mathfrak{n}_i^{-,a}$ on $V_{\mathbb{Z}}(\ell\varpi_i)$ and $V_{\mathbb{Z}}^{a,+}(\ell\varpi_i)$ are the same, so the $\mathfrak{n}_i^{-,a}$ actions on $V_{\mathbb{Z}}^{a,+}(\ell\varpi_i)$ and $V_{\mathbb{Z}}(\Psi(\ell\omega_i))_\tau$ are isomorphic; hence, so are the $\mathfrak{n}^{-,a}$ actions by trivial extension. The proposition follows now by Lemma 3.3. □

4. The general case for $\mathfrak{sl}_n$

4A. We extend Proposition 3.4 to any dominant weight for $\mathfrak{sl}_n$. Recall that for a dominant weight $\lambda = a_1\omega_1 + \cdots + a_{n-1}\omega_{n-1}$ we denote by λ^* the dominant weight given by $\lambda^* = a_{n-1}\omega_1 + \cdots + a_1\omega_{n-1}$.

Theorem 4.1. *Let λ be a dominant $\mathfrak{sl}_n$-weight. As an $\mathfrak{n}_{\mathbb{Z}}^{-,a}$-module, the Demazure submodule $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda^*))_\tau$ of the $(\mathfrak{sl}_{2n})_{\mathbb{Z}}$-module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda^*))$ is isomorphic to $V_{\mathbb{Z}}^a(\lambda)$.*

The proof of Theorem 4.1 will be given in Section 4G, and the strategy of proof is explained in Section 4C. We deduce a useful corollary.

Corollary 4.2. *In particular, $V_{\mathbb{Z}}^a(\lambda)$ is free as a $\mathbb{Z}$-module.*

Proof of the corollary. The Demazure module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda^*))_\tau$ is a direct summand of the free $\mathbb{Z}$-module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda^*))$ and hence free as $\mathbb{Z}$-module. □

4B. The abelianized module $V_{\mathbb{Z}}^a(\lambda)$ is a cyclic module over the algebra $S_{\mathbb{Z}}^\bullet(\mathfrak{n}^{-,a})$ having as a generator the image of a highest weight vector $v_\lambda \in V(\lambda)$ in $V_{\mathbb{Z}}^a(\lambda)$. Hence the module is isomorphic to $S_{\mathbb{Z}}^\bullet(\mathfrak{n}^{-,a})/I_{\mathbb{Z}}(\lambda)$, where $I_{\mathbb{Z}}(\lambda)$ is the annihilator of v_λ in $S_{\mathbb{Z}}^\bullet(\mathfrak{n}^{-,a})$.

We have an additional Lie algebra acting on $S_{\mathbb{Z}}^\bullet(\mathfrak{n}^{-,a})$ as well as on $V_{\mathbb{Z}}^a(\lambda)$. Let $\mathfrak{b}$ be the Borel subalgebra of $\mathfrak{g} = (\mathfrak{sl}_n)_{\mathbb{Z}} \otimes \mathbb{C}$ as in Example 1.1, so $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$. As free $\mathbb{Z}$-modules, $U_{\mathbb{Z}}(\mathfrak{n}^-) \simeq U_{\mathbb{Z}}(\mathfrak{g})/U_{\mathbb{Z}}^+(\mathfrak{h} + \mathfrak{n}^+)$, so that the adjoint action of $U_{\mathbb{Z}}(\mathfrak{b})$ on $U_{\mathbb{Z}}(\mathfrak{g})$ induces the structure of a $U_{\mathbb{Z}}(\mathfrak{b})$-module on $U_{\mathbb{Z}}(\mathfrak{n}^-)$ and hence on $S_{\mathbb{Z}}^\bullet(\mathfrak{n}^{-,a})$. This action is compatible with the induced $U_{\mathbb{Z}}(\mathfrak{b})$-action on $V_{\mathbb{Z}}^a(\lambda)$ [Feigin et al. 2013, Prop. 2.3.]. Recall that for a positive root α we have denoted by f_α the corresponding fixed Chevalley basis element in $(\mathfrak{sl}_n)_{-\alpha,\mathbb{Z}}$. Using the presentation of Demazure modules in terms of generators and relations by Joseph, Mathieu and Polo (compare [Mathieu 1989, Lemme 26]), we get as a consequence of the proof of Theorem 4.1 the following description of the ideal $I_{\mathbb{Z}}(\lambda)$; see [Feigin et al. 2011a; 2013].

Corollary 4.3. *As a cyclic $S_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-module, the abelianized module $V^a_{\mathbb{Z}}(\lambda)$ is isomorphic to $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})/I_{\mathbb{Z}}(\lambda)$, where*

$$I_{\mathbb{Z}}(\lambda) = S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})\big(U_{\mathbb{Z}}(\mathfrak{n}^+)\circ \operatorname{span}\{f_\alpha^{(\langle\alpha^\vee,\lambda\rangle+m)} \mid m\geq 1, \alpha>0\}\big) \subseteq S_{\mathbb{Z}}(\mathfrak{n}^{-,a}).$$

4C. The proof of Theorem 4.1 will be given in Section 4G, but it needs some preparation. The strategy of the proof is summarized by the following diagram of $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-modules. For a dominant weight $\lambda = a_1\omega_1+\cdots+a_{n-1}\omega_{n-1}$ (so $\lambda^* = a_{n-1}\omega_1+\cdots+a_1\omega_{n-1}$), we get the following natural maps (the details are described below):

$$\begin{tikzcd}
S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})/I_{\mathbb{Z}}(\lambda^*) \arrow[r, "h", "\simeq"'] & V^a_{\mathbb{Z}}(\lambda^*) \arrow[r, "b"] \arrow[d, "a"] & V^a_{\mathbb{Z}}(a_1\omega_1^*)\otimes\cdots\otimes V^a_{\mathbb{Z}}(a_{n-1}\omega_{n-1}^*) \arrow[d, "c", "\wr"'] \\
S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})/M_{\mathbb{Z}}(\lambda^*) \arrow[r, "g", "\simeq"'] \arrow[u, dotted, two heads, "f"] & \tilde V_{\mathbb{Z}}(\Psi(\lambda))_\tau \arrow[r, hook, "d"] & \tilde V_{\mathbb{Z}}(a_1\Psi(\omega_1))_\tau\otimes\cdots\otimes \tilde V_{\mathbb{Z}}(a_{n-1}\Psi(\omega_{n-1}))_\tau .
\end{tikzcd}$$

Let us describe the diagram above and the strategy of the proof. We recall that, given a tensor product of cyclic $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-modules, the Cartan component of the tensor product is, by definition, the cyclic $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-submodule generated by the tensor product of the cyclic generators. Further, recall that the isomorphism $V^a_{\mathbb{Z}}(\ell\omega_j^*)\simeq \tilde V_{\mathbb{Z}}(\ell\Psi(\omega_j))_\tau$ sends the highest weight vector $v_{\ell\omega_j^*}$ to the extremal weight vector $v_{\tau(\ell\Psi(\omega_j))}$ and uses the Chevalley involution. The maps above are defined as follows:

- b is induced by the compatibility of the PBW filtration with the tensor product, and it is surjective onto the Cartan component of this tensor product.
- $I_{\mathbb{Z}}(\lambda^*)$ is the annihilator in $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ of the image of the highest weight vector v_{λ^*} in $V^a_{\mathbb{Z}}(\lambda^*)$ and h is the corresponding quotient map.
- c is the isomorphism given by Proposition 3.4.
- d is the isomorphism onto the Cartan component of the tensor product. The fact that this is an isomorphism follows by standard monomial theory [Lakshmibai et al. 1979] or Frobenius splitting [Ramanathan 1987].
- a equals $c\circ b$ after identifying $\tilde V_{\mathbb{Z}}(\Psi(\lambda))_\tau$ with its image under d.
- $M_{\mathbb{Z}}(\lambda^*)$ is the annihilator in $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ of the extremal weight vector $v_{\tau(\Psi(\lambda))}$ in $\tilde V_{\mathbb{Z}}(\Psi(\lambda))_\tau$ and g is the corresponding quotient map.
- f is going to be constructed in the proof.

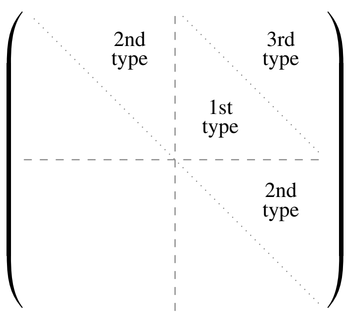
In order to finish the proof we will show that $M_{\mathbb{Z}}(\lambda^*)\subseteq I_{\mathbb{Z}}(\lambda^*)$, and the inclusion induces the surjective map f which in turn shows that the map a is an isomorphism.

4D. We first determine $M_{\mathbb{Z}}(\lambda^*)$. By [Mathieu 1989, Lemme 26], the Demazure module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_\tau$ is isomorphic to the algebra $U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$ modulo the left ideal $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))$ generated for all $m \geq 1$ by

$$\begin{cases} E_{\tilde{\alpha}}^{(m)} & \text{if } \langle\tilde{\alpha}^\vee, \tau\Psi(\lambda)\rangle \geq 0, \\ E_{\tilde{\alpha}}^{(-\langle\tilde{\alpha}^\vee, \tau\Psi(\lambda)\rangle+m)} & \text{otherwise.} \end{cases}$$

4E. The annihilator $M_{\mathbb{Z}}(\lambda^*)$ is the intersection of $U_{\mathbb{Z}}(\mathfrak{n}^{-,a}) \subset U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$ with the ideal $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))$. To determine the intersection, let us divide the positive roots of SL_{2n} into three families:

- $\tilde{\alpha}$ is of the *first type* if $\tilde{\alpha} = \phi(\alpha)$ for some positive SL_n-root α.
- $\tilde{\alpha} = \tilde{\epsilon}_k - \tilde{\epsilon}_l$ is of *second type* if $1 \leq k < l \leq n$ or $n+1 \leq k < l \leq 2n$.
- $\tilde{\alpha} = \tilde{\epsilon}_k - \tilde{\epsilon}_l$ is of *third type* if $1 \leq k \leq n$, $n+1 \leq l \leq 2n$ and $k < l - n$.



The $E_{\tilde{\alpha}}$, with $\tilde{\alpha}$ of second type, span a Lie subalgebra isomorphic to two copies of $\mathfrak{b}_{\mathbb{Z}}$. Let $\mathfrak{b}^1_{\mathbb{Z}}$ denote the first copy spanned by the $E_{\tilde{\alpha}}$, $\tilde{\alpha} = \tilde{\epsilon}_k - \tilde{\epsilon}_l$, $1 \leq k < l \leq n$, and let $\mathfrak{b}^2_{\mathbb{Z}}$ denote the second copy spanned by the $E_{\tilde{\alpha}}$, $\tilde{\alpha} = \tilde{\epsilon}_k - \tilde{\epsilon}_l$, $n+1 \leq k < l \leq 2n$.

Let $\tilde{I}_{\mathbb{Z}}(\infty) \subset U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$ be the left $U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$-submodule generated by the $E_{\tilde{\alpha}}^{(m)}$, with $m \geq 1$ and $\tilde{\alpha}$ of second or third type. Then Lemma 2.5 and a PBW basis argument show that we have the $\mathbb{Z}$-module decomposition

$$U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+) = U_{\mathbb{Z}}(\mathfrak{n}^{-,a}) \oplus \tilde{I}_{\mathbb{Z}}(\infty) = S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a}) \oplus \tilde{I}_{\mathbb{Z}}(\infty) \quad \text{and} \quad \tilde{I}_{\mathbb{Z}}(\infty) \subset \tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda)).$$

By abuse of notation, we identify in the following $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ with $U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)/\tilde{I}_{\mathbb{Z}}(\infty)$. So determining $M_{\mathbb{Z}}(\lambda^*) = U_{\mathbb{Z}}(\mathfrak{n}^{-,a}) \cap \tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))$ (the intersection taking place in $U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$) is equivalent to determining the image of $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))/I_{\mathbb{Z}}(\infty)$ in $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$. In the following we identify $M_{\mathbb{Z}}(\lambda^*)$ with $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))/I_{\mathbb{Z}}(\infty)$.

Note that $U_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}})$ acts naturally via the adjoint action on $\tilde{\mathfrak{n}}^+_{\mathbb{Z}}$ and hence on $U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)$. The span of the $E_{\tilde{\alpha}}$, with $\tilde{\alpha}$ of second or third type, is stable under this adjoint action of $\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}}$, so $\tilde{I}_{\mathbb{Z}}(\infty) \subset U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^+)_{\mathbb{Z}}$ is a submodule with respect to this adjoint action.

We get an induced $U_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}})$-action on $S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ which we denote by "$\circ$". Moreover, since $U^+_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}})$ (the set of elements without constant term) is contained in $\tilde{I}_{\mathbb{Z}}(\infty)$, we see that $M_{\mathbb{Z}}(\lambda^*) = \tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))/\tilde{I}_{\mathbb{Z}}(\infty)$ is a $U^+_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}})$-stable submodule with respect to the "$\circ$"-action of $U_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}})$. As a first step in the proof of the theorem we show:

Lemma 4.4. *The left $S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-submodule $M_{\mathbb{Z}}(\lambda^*) \subset S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ is generated by*

$$m_{\mathbb{Z}}(\lambda^*) := \big\langle U_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}}) \circ E_{\tilde{\alpha}}^{(-\langle\tilde{\alpha}^\vee, \tau\Psi(\lambda)\rangle + m)} \;\big|\; \tilde{\alpha} \textit{ of first type and } m \geq 1\big\rangle_{\mathbb{Z}}.$$

Proof. Let $\bar{\mathtt{m}}$ be an element of $M_{\mathbb{Z}}(\lambda^*)$ and choose a representative $\mathtt{m}$ in $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))$. Since we are free to choose a representative modulo $I_{\mathbb{Z}}(\infty)$, we may assume (see Section 4D) that $\mathtt{m}$ is a sum of monomials of the form $rE_{\tilde{\alpha}}^{(\ell)}$, where r is a monomial in the $E_{\tilde{\beta}}^{(q)}$ with $q \geq 0$ and $\tilde{\beta}$ of first, second, or third type, and $\ell = -\langle\tilde{\alpha}^\vee, \tau\Psi(\lambda)\rangle + k$ for some $k \geq 1$ and $\tilde{\alpha}$ of first type.

If $\tilde{\gamma}$ is a root of third type and $\tilde{\beta}$ is any other positive root, then $[E_{\tilde{\gamma}}, E_{\tilde{\beta}}] = cE_{\tilde{\gamma}'}$, where $c \in \mathbb{Z}$ and either $c = 0$ or $\tilde{\gamma}'$ is of third type. So if r has a factor $E_{\tilde{\gamma}}^{(p)}$, with $p > 0$ and $\tilde{\gamma}$ a root of third type, then we can rewrite the monomial $rE_{\tilde{\alpha}}^{(\ell)}$ as a sum of monomials of the form $r'E_{\tilde{\gamma}}^{(p')}$, with $p' > 0$. Since this sum is an element in $I_{\mathbb{Z}}(\infty)$, without loss of generality we will assume in the following that r has only factors of the form $E_{\tilde{\beta}}^{(\ell)}$, with $\tilde{\beta}$ of first or second type.

If $\tilde{\gamma}$ is of second type and $\tilde{\beta}$ is of first type, then $[E_{\tilde{\gamma}}, E_{\tilde{\beta}}] = cE_{\tilde{\gamma}'}$, where $c \in \mathbb{Z}$ and either $c = 0$ or $\tilde{\gamma}'$ is of first or third type. So after reordering the factors we can assume without loss of generality in the following that $rE_{\tilde{\alpha}}^{(\ell)}$ is of the form $r = r_1r_2$, where r_1 is a monomial in the $E_{\tilde{\beta}}^{(\ell)}$, with $\tilde{\beta}$ of first type, and r_2 is a monomial in the $E_{\tilde{\gamma}}^{(\ell)}$, with $\tilde{\gamma}$ of second type.

Recall that for $\tilde{\gamma}$ of second type,

$$\begin{aligned} E_{\tilde{\gamma}}E_{\tilde{\beta}_1}\cdots E_{\tilde{\beta}_m} &\equiv \sum_{i=1}^{m} E_{\tilde{\beta}_1}\cdots E_{\tilde{\beta}_{i-1}}(E_{\tilde{\gamma}}E_{\tilde{\beta}_i} - E_{\tilde{\beta}_i}E_{\tilde{\gamma}})E_{\tilde{\beta}_{i+1}}\cdots E_{\tilde{\beta}_m} \mod I_{\mathbb{Z}}(\infty) \\ &\equiv \sum_{i=1}^{m} E_{\tilde{\beta}_1}\cdots E_{\tilde{\beta}_{i-1}}(\mathrm{ad}(E_{\tilde{\gamma}})(E_{\tilde{\beta}_i}))E_{\tilde{\beta}_{i+1}}\cdots E_{\tilde{\beta}_m} \mod I_{\mathbb{Z}}(\infty) \\ &\equiv \mathrm{ad}(E_{\tilde{\gamma}})(E_{\tilde{\beta}_1}\cdots E_{\tilde{\beta}_m}) \mod I_{\mathbb{Z}}(\infty). \end{aligned}$$

An appropriate reformulation of the equality above holds also for the divided powers of root vectors. It follows that $r_2E_{\tilde{\alpha}}^{(\ell)} \in m_{\mathbb{Z}}(\lambda^*)$; hence $rE_{\tilde{\alpha}}^{(\ell)} = r_1r_2E_{\tilde{\alpha}}^{(\ell)} \in S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}^{-,a})m_{\mathbb{Z}}(\lambda^*)$, which implies that $M_{\mathbb{Z}}(\lambda^*)$ is generated by $m_{\mathbb{Z}}(\lambda^*)$ as a left $S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}^{-,a})$-module. □

4F. To compare $M_{\mathbb{Z}}(\lambda^*)$ with $I_{\mathbb{Z}}(\lambda^*)$, we need a variant of the description of $m_{\mathbb{Z}}(\lambda^*)$. Let $\Delta(\mathfrak{b}_{\mathbb{Z}}) \subset \mathfrak{b}^1_{\mathbb{Z}} \oplus \mathfrak{b}^2_{\mathbb{Z}}$ be the Lie subalgebra obtained as a diagonally embedded copy of $\mathfrak{b}_{\mathbb{Z}}$. Let $U_{\mathbb{Z}}(\Delta(\mathfrak{b}_{\mathbb{Z}})) \subset U_{\mathbb{Z}}((\mathfrak{b}^1_{\mathbb{Z}}) \oplus (\mathfrak{b}^2_{\mathbb{Z}}))$ be its hyperalgebra.

Lemma 4.5. $m_{\mathbb{Z}}(\lambda^*) = \big\langle U_{\mathbb{Z}}(\Delta(\mathfrak{b}_{\mathbb{Z}})) \circ E_{\tilde{\alpha}}^{(-\langle\tilde{\alpha}^\vee,\, \tau\Psi(\lambda)\rangle + m)} \mid \tilde{\alpha}$ *of first type and* $m \geq 1\big\rangle_{\mathbb{Z}}$.

Proof. We assume first that k is an algebraically closed field of arbitrary characteristic. Let B be the subgroup of upper triangular invertible matrices in $\mathrm{SL}_n(k)$, so $\operatorname{Lie} B = \mathfrak{b}$. Let $B^1 \times B^2 \subset \mathrm{SL}_{2n}(k)$ be the subgroup with Lie algebra $\mathfrak{b}^1 \oplus \mathfrak{b}^2$ and denote by $\Delta(B) \subset B^1 \times B^2$ the diagonally embedded group isomorphic to B.

Let $\mathfrak{q}$ be the sum of the SL_{2n}-root spaces corresponding to roots of second or third type. Then $\tilde{\mathfrak{n}}^+ = \mathfrak{n}^{-,a} \oplus \mathfrak{q}$ and we identify $\mathfrak{n}^{-,a}$ with $\tilde{\mathfrak{n}}^+/\mathfrak{q}$. The adjoint action of $B^1 \times B^2$ on $\mathfrak{sl}_{2n}$ admits $\tilde{\mathfrak{n}}^+$ as well as $\mathfrak{q}$ as submodules, so we get an induced $(B^1 \times B^2)$-action on $\mathfrak{n}^{-,a} = \tilde{\mathfrak{n}}^+/\mathfrak{q}$. This action naturally extends to the commutative hyperalgebra $S_k^\bullet(\mathfrak{n}^{-,a})$.

If we replace the group action of $B^1 \times B^2$ by the induced action of the hyperalgebra $U_k(\mathfrak{b}^1 \oplus \mathfrak{b}^2)$ of the group, then we get the action of $U_k(\mathfrak{b}^1 \oplus \mathfrak{b}^2)$ on $U_k(\tilde{\mathfrak{n}}^+)$, respectively on $S_k^\bullet(\mathfrak{n}^{-,a})$ discussed above, and similarly for the action of $\Delta(B)$ and its hyperalgebra $U_k(\Delta(\mathfrak{b}))$. Recall that for a root $\tilde{\alpha}$ of type 1,

$$U_k(\mathfrak{b}^1 \oplus \mathfrak{b}^2) \circ E_{\tilde{\alpha}}^{(m)} = \langle \mathrm{Ad}((b^1, b^2)) \circ (E_{\tilde{\alpha}})^{(m)} \mid (b_1, b_2) \in B^1 \times B^2 \rangle,$$

i.e., the smallest $U_k(\mathfrak{b}^1 \oplus \mathfrak{b}^2)$ stable subspace containing $(E_{\tilde{\alpha}})^{(m)}$ is the linear span of the $(B^1 \times B^2)$-orbit. The same holds in the other case, so we have:

$$U_k(\Delta(\mathfrak{b})) \circ E_{\tilde{\alpha}}^{(m)} = \langle \mathrm{Ad}((b, b)) \circ (E_{\tilde{\alpha}})^{(m)} \mid b \in B \rangle.$$

Let $\mathfrak{d}$ be the sum of the SL_{2n}-root spaces corresponding to roots of first or third type and let $\mathfrak{d}^3$ be just the sum of the root spaces corresponding to roots of third type, so $\mathfrak{d} = \mathfrak{n}^{-,a} \oplus \mathfrak{d}^3$. We identify $\mathfrak{d} \subset \mathfrak{sl}_{2n}$ with $M_n(k)$, formally this can be done by the map

$$\chi : \mathfrak{d} \to M_n(k), \quad \tilde{A} = \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix} \mapsto A,$$

where A is a $n \times n$ matrix. In the following we simplify the notation and omit the map χ. We freely identify $\mathfrak{d}$ with $M_n(k)$, so we denote by A the $n \times n$ matrix as well as the $2n \times 2n$-matrix $\tilde{A} \in \mathfrak{d}$. Note that for $(b_1, b_2) \in B^1 \times B^2$ we get

$$\mathrm{Ad}((b_1, b_2)) \circ \tilde{A} = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_1^{-1} & 0 \\ 0 & b_2^{-1} \end{pmatrix} = \begin{pmatrix} 0 & b_1 A b_2^{-1} \\ 0 & 0 \end{pmatrix},$$

we just write $\mathrm{Ad}((b_1, b_2)) \circ (A) = b_1 A b_2^{-1}$ and $\mathrm{Ad}((b, b)) \circ (A) = bAb^{-1}$. Recall that χ is just a vector space isomorphism. If we equip in addition $M_n(k)$ with the trivial Lie bracket, then this becomes also a Lie algebra homomorphism. In this sense we identify also the (commutative) Lie subalgebras $\mathfrak{n}^{-,a}$ and $\mathfrak{d}^3$ with subalgebras of M_n. An elementary calculation shows how the $B^1 \times B^2$-orbit through $E_{\tilde{\alpha}} = E_{\epsilon_n - \epsilon_{n+1}}$ breaks up into $\Delta(B)$-orbits. Recall that we identify $\mathfrak{d}$ with $M_n(k)$

and $E_{\epsilon_n-\epsilon_{n+1}} \in \mathfrak{d}$ corresponds to $E_{n,1}$:

$$\begin{aligned}\{\mathrm{Ad}((b_1,b_2))\circ(E_{\tilde\alpha}) \mid (b_1,b_2)\in B^1\times B^2\} &= \{b_1E_{n,1}b_2^{-1} \mid b_1,b_2\in B\}\\ &= \bigcup_{\lambda\in k}\{b(E_{n,1}+\lambda E_{1,1})b^{-1} \mid b\in B\}\\ &\subset M_n(k).\end{aligned}$$

We conclude for the linear span,

$$\begin{aligned}U_k(\mathfrak{b}^1\oplus\mathfrak{b}^2)\circ E_{\tilde\alpha}^{(m)} &= \big\langle \mathrm{Ad}((b_1,b_2))\circ(E_{\tilde\alpha})^{(m)} \;\big|\; (b_1,b_2)\in B^1\times B^2\big\rangle\\ &= \big\langle \big(\mathrm{Ad}((b_1,b_2))\circ(E_{\tilde\alpha})\big)^{(m)} \;\big|\; (b_1,b_2)\in B^1\times B^2\big\rangle\\ &= \big\langle (b_1E_{n,1}b_2^{-1})^{(m)} \;\big|\; b_1,b_2\in B\big\rangle\\ &= \left\langle \bigcup_{\lambda\in k}\{b(E_{n,1}+\lambda E_{1,1})^{(m)}b^{-1} \;\middle|\; b\in B\}\right\rangle.\end{aligned}$$

Let $I(\mathfrak{d}^3)\subset U_k(\mathfrak{d})$ be the left ideal in the hyperalgebra generated by $\mathfrak{d}^3$. Then

$$\begin{aligned}U_k(\mathfrak{b}^1\oplus\mathfrak{b}^2)\circ E_{\tilde\alpha}^{(m)} &\equiv \left\langle \bigcup_{\lambda\in k}\{(b(E_{n,1}+\lambda E_{1,1})b^{-1})^{(m)} \;\middle|\; b\in B\}\right\rangle\\ &\equiv \langle (bE_{n,1}b^{-1})^{(m)} \mid b\in B\rangle \quad \bmod I(\mathfrak{d}^3)\end{aligned}$$

because the $E_{1,1}^{(\ell)}$, with $\ell\geq 1$, lie in the $\Delta(B)$-stable ideal $I(\mathfrak{d}^3)$. It follows that

$$\begin{aligned}U_k(\mathfrak{b}^1\oplus\mathfrak{b}^2)\circ E_{\tilde\alpha}^{(m)} &\equiv \langle \mathrm{Ad}((b,b))(E_{\tilde\alpha})^{(m)} \mid b\in B\rangle\\ &\equiv U_k(\Delta(\mathfrak{b}))\circ E_{\tilde\alpha}^{(m)} \quad \bmod I(\mathfrak{d}^3).\end{aligned}$$

Since $I(\mathfrak{d}^3)\subset \tilde I_k(\infty)$, the equation holds in $S^\bullet(\mathfrak{n}^{-,a}) = U_k(\tilde{\mathfrak{n}}^+)/\tilde I(\infty)$ too, so

$$U_k(\mathfrak{b}^1\oplus\mathfrak{b}^2)\circ E_{\tilde\alpha}^{(m)} = U_k(\Delta(\mathfrak{b}))\circ E_{\tilde\alpha}^{(m)}$$

in $S^\bullet(\mathfrak{n}^{-,a})$. It is now easy to see that the same arguments prove the equality for all $E_{\tilde\alpha}^{(m)}$, with $\tilde\alpha$ of first type and $m\geq 1$. Clearly,

$$U_{\mathbb{Z}}(\mathfrak{b}^1\oplus\mathfrak{b}^2)\circ E_{\tilde\alpha}^{(m)} \supseteq U_{\mathbb{Z}}(\Delta(\mathfrak{b}))\circ E_{\tilde\alpha}^{(m)}.$$

Since we have equality after base change for fields of arbitrary characteristics, the equality of the modules holds also over $\mathbb{Z}$. In particular, the following equality holds in $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$:

$$\begin{aligned}m_{\mathbb{Z}}(\lambda^*) &= \big\langle U_{\mathbb{Z}}(\mathfrak{b}^1_{\mathbb{Z}}\oplus\mathfrak{b}^2_{\mathbb{Z}})\circ E_{\tilde\alpha}^{(-\langle\tilde\alpha^\vee,\tau\Psi(\lambda)\rangle+m)} \;\big|\; \tilde\alpha \text{ of first type and } m\geq 1\big\rangle_{\mathbb{Z}}\\ &= \big\langle U_{\mathbb{Z}}(\Delta(\mathfrak{b}_{\mathbb{Z}}))\circ E_{\tilde\alpha}^{(-\langle\tilde\alpha^\vee,\tau\Psi(\lambda)\rangle+m)} \;\big|\; \tilde\alpha \text{ of first type and } m\geq 1\big\rangle_{\mathbb{Z}}. \qquad \square\end{aligned}$$

4G. ***Proof of Theorem 4.1.*** Recall the identification of the abelianized version of $\mathfrak{n}^- \subset \mathfrak{sl}_n$ with $\mathfrak{n}^{-,a} \subset \mathfrak{sl}_{2n}$, which sends the image of a Chevalley generator f_α to $E_{\phi(\alpha)}$. By Lemma 2.5 (see Lemma 3.3 for the twist $\lambda \leftrightarrow \lambda^*$) the elements $E_{\tilde{\alpha}}^{(-\langle\tilde{\alpha}^\vee, \tau\Psi(\lambda)\rangle+m)}$, where $\tilde{\alpha}$ is of first type and $m \geq 1$ are elements of $I_{\mathbb{Z}}(\lambda^*)$. Now the $U_{\mathbb{Z}}(\mathfrak{b})$-module structure on $S^\bullet_{\mathbb{Z}}(\mathfrak{n}^{-,a})$ described in Section 4B is the same as the one described above, so it follows that $m_{\mathbb{Z}}(\lambda^*) \subset I_{\mathbb{Z}}(\lambda^*)$ and hence $M_{\mathbb{Z}}(\lambda^*) \subset I_{\mathbb{Z}}(\lambda^*)$, which, as explained in Section 4C, finishes the proof of the theorem. □

4H. Let ρ be the sum of all fundamental weights for SL_n and set $\tilde{\rho} = \Psi(\rho)$. Let $Q_{\mathbb{Z}} \subset (\mathrm{SL}_{2n})_{\mathbb{Z}}$ be the corresponding parabolic $\mathbb{Z}$-subgroup. Recall that $(N^{-,a})_{\mathbb{Z}}$ is a commutative subgroup of the Borel subgroup $\tilde{B}_{\mathbb{Z}}$. For any SL_{2n}-root $\tilde{\alpha}$ let $U_{\mathbb{Z},\tilde{\alpha}}$ be the associated root subgroup.

Lemma 4.6. *The orbit $\tilde{B}_{\mathbb{Z}}.\tau \subset (\mathrm{SL}_{2n})_{\mathbb{Z}}/Q_{\mathbb{Z}}$ is equal to $N^{-,a}.\tau$, and the map $N^{-,a} \to N^{-,a}.\tau$, $u \mapsto u\tau$, is a bijection.*

Proof. We have $\tilde{B}_{\mathbb{Z}}.\tau = \prod_{\tilde{\alpha}>0} U_{\mathbb{Z},\tilde{\alpha}}.\tau$, and the map $\prod_{\tilde{\alpha}\in\Gamma} U_{\mathbb{Z},\tilde{\alpha}} \to \prod_{\tilde{\alpha}\in\Gamma} U_{\mathbb{Z},\tilde{\alpha}}.\tau$ is a bijection, where Γ is the set of all positive roots of SL_{2n} such that $\tau^{-1}(\tilde{\alpha}) < 0$ and $\tau^{-1}(\tilde{\alpha})$ is not an element of the root system of $Q_{\mathbb{Z}}$. Now this condition is fulfilled if and only if $\langle\tau^{-1}(\tilde{\alpha}^\vee), \tilde{\rho}\rangle < 0$, or, equivalently, $\langle\tilde{\alpha}^\vee, \tau(\tilde{\rho})\rangle < 0$. By Lemma 2.4 this is only possible if $\tilde{\alpha} = \tilde{\epsilon}_i - \tilde{\epsilon}_j$ is such that $1 \leq i \leq n$, $n+1 \leq j \leq 2n$, and $i > j-n$. But this implies that the root subgroup $U_{\mathbb{Z},\tilde{\alpha}}$ is a subgroup of $N^{-,a}$, and all root subgroups of $(\mathrm{SL}_{2n})_{\mathbb{Z}}$ contained in $N^{-,a}$ satisfy this condition. It follows that $N^{-,a}$ is the product of all root subgroups corresponding to positive roots of SL_{2n} in Γ. □

Recall that the degenerate flag scheme $\mathcal{F}\ell(\lambda)^a_{\mathbb{Z}}$ is the closure of the $N^{-,a}_{\mathbb{Z}}$-orbit through the highest weight vector in $\mathbb{P}P(V^a_{\mathbb{Z}}(\lambda))$.

Theorem 4.7. *Let λ be a dominant weight for* SL_n. *The Schubert scheme $X_{\mathbb{Z}}(\tau) \subset \mathbb{P}(\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_\tau)$ is isomorphic to the degenerate partial flag scheme $\mathcal{F}\ell(\lambda^*)^a_{\mathbb{Z}}$ for* $(\mathrm{SL}_n)_{\mathbb{Z}}$, *and this map induces a module isomorphism $H^0(X_{\mathbb{Z}}(\tau), \mathcal{L}_{\Psi(\lambda)}) \simeq (V^a_{\mathbb{Z}}(\lambda^*))^*$.*

Proof. We consider only the case where λ is regular; the arguments in the general case are similar. With respect to the isomorphism in Lemma 4.6, the orbit

$$\tilde{B}_{\mathbb{Z}}.\tau \subset (\mathrm{SL}_{2n})_{\mathbb{Z}}/(\tilde{P}_\lambda)_{\mathbb{Z}} \hookrightarrow \mathbb{P}(\tilde{V}(\Psi(\lambda))),$$

through the extremal weight vector, which is the same as the $N^{-,a}$-orbit, is mapped onto the $N^{-,a}_{\mathbb{Z}}$-orbit through the highest weight vector in $\mathbb{P}(V^a_{\mathbb{Z}}(\lambda^*))$. By definition, the Schubert scheme $X_{\mathbb{Z}}(\tau)$ is the closure of the orbit $\tilde{B}_{\mathbb{Z}}.\tau$ and the degenerate flag scheme $\mathcal{F}\ell(\lambda^*)^a_{\mathbb{Z}}$ is the closure of the $N^{-,a}_{\mathbb{Z}}$-orbit. It follows that the module isomorphism induces an isomorphism between the Schubert scheme $X_{\mathbb{Z}}(\tau) \subseteq \mathbb{P}(\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_\tau)$ and the degenerate flag scheme $\mathcal{F}\ell(\lambda^*)^a_{\mathbb{Z}}$ in $\mathbb{P}(V^a_{\mathbb{Z}}(\lambda))$. Hence we get induced isomorphisms

$$H^0(X_{\mathbb{Z}}(\tau), \mathcal{L}_{\Psi(\lambda)}) \simeq (\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_\tau)^* \simeq (V^a_{\mathbb{Z}}(\lambda^*))^*$$

for the dual modules. □

Let k be an algebraically closed field of arbitrary characteristic and denote by $V_k(\lambda) = V_{\mathbb{Z}}(\lambda) \otimes_{\mathbb{Z}} k$, $U_k(\mathfrak{sl}_n) = U_{\mathbb{Z}}(\mathfrak{sl}_n) \otimes_{\mathbb{Z}} k$, $U_k(\mathfrak{n}^-) = U_{\mathbb{Z}}(\mathfrak{n}^-) \otimes_{\mathbb{Z}} k$, etc., the objects obtained by base change. The PBW filtration

$$V_k(\lambda)_\ell = \langle Y_1^{(m_1)} \cdots Y_N^{(m_N)} v_\lambda \mid m_1 + \cdots + m_N \le \ell,\ Y_1, \dots, Y_N \in \mathfrak{n}_k^- \rangle$$

and the associated graded space $V_k^a(\lambda)$ is defined in the same way as before, and by Corollary 4.2, $V_k(\lambda)_\ell = V_{\mathbb{Z}}(\lambda)_\ell \otimes_{\mathbb{Z}} k$ and $V_k^a(\lambda) = V_{\mathbb{Z}}^a(\lambda) \otimes_{\mathbb{Z}} k$. The group $N_k^{-,a}$ acts on the abelianized representation $V_k^a(\lambda)$, and the degenerate flag variety $\mathcal{F}\ell(\lambda)_k^a$ is the closure of the $N_k^{-,a}$-orbit through the highest weight vector in $\mathbb{P}(V_k^a(\lambda))$.

Now by the results of [Mathieu 1989; Mehta and Ramanathan 1988; Ramanathan 1987; Ramanan and Ramanathan 1985] one knows that for Demazure modules we have $\tilde{V}_k(\Psi(\lambda))_\tau = \tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_\tau \otimes_{\mathbb{Z}} k$, $X_k(\tau) = X_{\mathbb{Z}}(\tau) \otimes_{\mathbb{Z}} k$, etc., and that the Schubert varieties are Frobenius split, projectively normal and have rational singularities. It follows that $V_k^a(\lambda^*) = \tilde{V}_k(\Psi(\lambda))_\tau$ and $\mathcal{F}\ell(\lambda^*)_k^a = X_k(\tau)$, so the degenerate flag variety has in this case the same nice geometric properties as the Schubert variety. For a dominant SL_n-weight $\lambda = \sum_{i=1}^{n-1} a_i\omega_i$, let the support $\operatorname{supp}\lambda$ of λ be the set $\{i \mid 1 \le i \le n-1,\ a_i \ne 0\}$.

Corollary 4.8. *The degenerate partial flag variety $\mathcal{F}\ell(\lambda)_k^a$ depends only on* $\operatorname{supp}\lambda$. *It is a projectively normal variety, Frobenius split, with rational singularities.*

Remark 4.9. Feigin and Finkelberg [2013] construct resolutions of the degenerate flag varieties given by towers of $\mathbb{P}^1$-fibrations. The steps of the successive fibrations are indexed by the set of positive roots, which had been totally reordered. In fact, their varieties are Bott–Samelson varieties [Cerulli Irelli and Lanini 2015, Appendix] and such an order (which actually should be thought of as an order on the set of negative roots) is now natural since it corresponds to the subsequent steps of the Bott–Samelson variety indexed by the reduced expression (2) of τ, under the identification of $-\alpha_{i,j}$ with $\tilde{\alpha}_{j,i+n}$.

5. A special Schubert variety: the Sp_{2m} case

As for the SL_n case, we want to realize for Example 1.2 the abelianized representation $V_{\mathbb{Z}}(\lambda)^a$ for $N_{\mathbb{Z}}^{-,a}$ as a Demazure submodule in an irreducible representation for the larger group $\mathrm{Sp}_{2(2m)}$.

5A. ***A special Weyl group element.*** Let us keep the same notation as in the previous sections and denote by $\mathfrak{h} \subset \mathfrak{sp}_{2m}$ (resp., $\tilde{\mathfrak{h}} \subset \mathfrak{sp}_{2(2m)}$) the Cartan subalgebra of traceless complex diagonal matrices and by $\mathfrak{b} \subset \mathfrak{sp}_{2m}$ (resp., $\tilde{\mathfrak{b}} \subset \mathfrak{sp}_{2(2m)}$) the Borel subalgebra of traceless complex upper triangular matrices. Let $\{\epsilon_1, \dots, \epsilon_{2m}\}$ (resp., $\{\tilde{\epsilon}_1, \dots, \tilde{\epsilon}_{2(2m)}\}$) be a basis of the dual vector space $\mathfrak{h}^*$ (resp., $\tilde{\mathfrak{h}}^*$). The choice of

Cartan and Borel subalgebras we made determines the following set of positive roots for Sp_{2m}:

$$\alpha_{i,j} := \begin{cases} \epsilon_i - \epsilon_j & 1 \le i < j \le m, \\ \epsilon_i + \epsilon_{j-m} & 1 \le i \le m < j \text{ and } i + j \le 2m, \end{cases}$$

where the simple roots are $\{\alpha_i := \alpha_{i,i+1} \mid 1 \le i \le m-1\} \cup \{\alpha_m := 2\epsilon_m\}$. We will write $\tilde{\alpha}_{i,j}$ for the Sp_{2m}-roots. The Weyl group of $\mathrm{Sp}_{2(2m)}$ is denoted $\widetilde{W}$. This is the group of linear transformations of $\mathfrak{h}^*$ generated by the elements $\{r_i \mid 1 \le i \le 2m\}$, where r_i denotes the reflection with respect to the simple root $\tilde{\alpha}_i$.

Definition 5.1. We define in $\widetilde{W}$ a very special element:

$$\overline{\tau} = (r_{2m} \cdots r_{m+1}) \cdots \\ (r_{2m} r_{2m-1} r_{2m-2})(r_{2m} r_{2m-1}) r_{2m} (r_m \cdots r_{2m-2}) \cdots (r_4 r_5 r_6)(r_3 r_4) r_2.$$

Any element of the Weyl group $\widetilde{W}$ of $\mathrm{Sp}_{2(2m)}$ can be identified with an element of the symmetric group on $4m$ letters $\mathcal{S}_{4m}$, via $r_i = s_i s_{4m-i}$, for $1 \le i \le 2m-1$, and $r_n = s_{2m}$ (where, as usual, s_i denotes the transposition exchanging i and $i+1$) and it acts on the basis $\{e_i \mid i = 1, \ldots, 4m\}$ of $\mathbb{C}^{4m}$ by permuting the indices. It is an easy check to see that under this identification $\overline{\tau}$ equals the element τ of Definition 2.2 for $n = 2m$ and we hence have the following (compare Lemma 2.3):

Lemma 5.2. *In the irreducible* $\mathrm{Sp}_{2(2m)}$*-representation* $\tilde{V}(\tilde{\omega}_{2i}) \subset \bigwedge^{2i} \mathbb{C}^{4m}$, *with* $1 \le i \le 2m$, *let* $v_{\omega_{2i}}$ $v_{\omega_{2i}} = e_1 \wedge e_2 \wedge \ldots \wedge e_{2i}$ *be the highest weight vector. Then* (*up to sign*),

$$\overline{\tau}(v_{\omega_{2i}}) = e_1 \wedge e_2 \wedge \cdots \wedge e_i \wedge e_{2m+1} \wedge e_{2m+2} \wedge \cdots \wedge e_{2m+i}.$$

We denote by $\{\omega_i \mid 1 \le i \le m\}$, resp. $\{\tilde{\omega}_i \mid 1 \le i \le 2m\}$, the fundamental weights of $\mathfrak{sp}_{2m}$, resp., $\mathfrak{sp}_{2(2m)}$. They are characterized by the property $\langle \alpha_i^\vee, \omega_j \rangle = \delta_{i,j}$.

Definition 5.3. Let $\Psi : \mathfrak{h}^* \to \tilde{\mathfrak{h}}^*$ be the linear map defined on the weight lattice by

$$\Psi\left(\sum_{i=1}^{n-1} a_i \omega_i\right) := \sum_{i=1}^{n-1} a_i \tilde{\omega}_{2i}.$$

Note: Ψ sends dominant weights to dominant weights. Let $\lambda = b_1 \epsilon_1 + \cdots + b_m \epsilon_m$, with $b_1 \ge \cdots \ge b_m \ge 0$, be a dominant weight for Sp_{2m}.

Lemma 5.4. $\overline{\tau}(\Psi(\lambda)) = b_1 \tilde{\epsilon}_1 + \cdots + b_m \tilde{\epsilon}_m - b_m \tilde{\epsilon}_{m+1} - \cdots - b_1 \tilde{\epsilon}_{2m}$.

Proof. This equality follows directly from Lemma 5.2 above. □

As in the special linear case, we define a map from the set of negative roots of Sp_{2m} to the set of positive $\mathrm{Sp}_{2(2m)}$-roots by sending $\alpha_{i,j}$ to $\tilde{\alpha}_{j,i+2m}$. The following is the symplectic analogue of Lemma 2.5:

Lemma 5.5. (i) *Let λ be a dominant weight for* $\mathrm{Sp}_{2m}(\mathbb{C})$. *For a positive* $\mathrm{Sp}_{2(2m)}$-*root* $\tilde{\alpha}$, *we have* $\langle\tilde{\alpha}^\vee, \tau(\Psi(\lambda))\rangle < 0$ *only if the* $\tilde{\alpha}$-*root space lies in* $\mathrm{Lie}(\mathbb{U}(\mathfrak{n}^-))$.

(ii) *Let λ be a dominant* Sp_{2m}-*weight, let* $\alpha = \alpha_{i,j}$ *be a positive* Sp_{2m}-*root, and let* $\tilde{\alpha} = \tilde{\alpha}_{j,i+2m}$ *be the* $\mathrm{Sp}_{2(2m)}$ *positive root associated with* $-\alpha$. *Then*

$$\langle\alpha^\vee, \lambda\rangle = -\langle\tilde{\alpha}^\vee, \bar{\tau}(\Psi(\lambda))\rangle.$$

(iii) *Let λ be a dominant weight for* $\mathrm{Sp}_{2m}(\mathbb{C})$ *and let $\tilde{\alpha}$ be a positive* $\mathrm{Sp}_{2(2m)}$-*root. Then* $E_{\tilde{\alpha}} v_{\bar{\tau}} \neq 0$ *in* $\tilde{V}(\Psi(\lambda))$ *only if* $\tilde{\alpha} = \tilde{\alpha}_{j,i+2m}$, *where* $\alpha_{i,j}$ *is a positive* Sp_{2m}-*root such that* $\langle\alpha_{i,j}^\vee, \lambda\rangle > 0$.

Proof. Lemma 5.4 implies that for $\lambda = b_1\epsilon_1 + \cdots + b_{m-1}\epsilon_{m-1}$,

$$\big\langle(\tilde{\epsilon}_i - \tilde{\epsilon}_j)^\vee, \bar{\tau}(\Psi(\lambda))\big\rangle = \begin{cases} b_i - b_j \geq 0 & \text{if } 1 \leq i < j \leq m, \\ b_i + b_{2m-j+1} \geq 0 & \text{if } 1 \leq i \leq m < j \leq 2m, \\ -b_{2m-i+1} + b_{2m-j+1} \geq 0 & \text{if } m < i < j \leq 2m, \end{cases}$$

and

$$\big\langle(\tilde{\epsilon}_i + \tilde{\epsilon}_{j-2m})^\vee, \bar{\tau}(\Psi(\lambda))\big\rangle = \begin{cases} b_i + b_{j-2m} \geq 0 & \text{if } 1 \leq i \leq m \\ & \quad \text{and } 2m < j \leq 3m, \\ b_i - b_{4m-j+1} \geq 0 & \text{if } 1 \leq i \leq m \\ & \quad \text{and } 3m < j \leq 4m-i+1, \\ -b_i + b_{4m-j+1} \geq 0 & \text{if } 1 \leq i \leq m,\ 3m < j, \\ & \quad \text{and } 4m-i+1 < j, \\ b_{2m-i+1} + b_{2m-j+1} \geq 0 & \text{if } m < i \leq 2m \\ & \quad \text{and } 3m < j, \end{cases}$$

where always $i + j \leq 4m$. This proves the corollary. □

6. The fundamental representations: the $\mathfrak{sp}_{2m}$ case

Let $\mathfrak{n}_{\mathbb{Z}}^{-,a,i}$ be the direct sum of all root spaces of Lie $(\mathrm{Sp}_{2(2m)})_{\mathbb{Z}}$ corresponding to positive roots β such that $\langle\beta^\vee, \bar{\tau}(\tilde{\omega}_{2i})\rangle < 0$ for all $1 \leq i \leq m$. By Lemma 5.5, such a space lies in $\mathrm{Lie}(\mathbb{U}(\mathfrak{n}_{\mathbb{Z}}^-))$.

By [Mathieu 1989, Lemme 26], the Demazure module $\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})_\tau$ is isomorphic to the algebra $U_{\mathbb{Z}}(\tilde{\mathfrak{n}})$ modulo the left ideal $\tilde{I}_{\mathbb{Z}}(\bar{\tau}\ell\tilde{\omega}_{2i})$ generated for all $m \geq 1$ by the

$$\begin{cases} E_{k,l}^{(m)} & \text{if } \langle\tilde{\alpha}_{k,l}^\vee, \bar{\tau}\tilde{\omega}_{2i}\rangle \geq 0, \\ E_{k,l}^{(-\langle\tilde{\alpha}_{k,l}^\vee, \bar{\tau}\ell\tilde{\omega}_{2i}\rangle + m)} & \text{otherwise.} \end{cases}$$

Therefore all the root vectors (and their divided powers) not lying in $\mathfrak{n}_{\mathbb{Z}}^{-,a,i}$ act trivially on $\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})_{\bar{\tau}}$ and hence in order to describe its structure as an $\mathfrak{n}_{\mathbb{Z}}^{-,a}$-module it suffices to consider only the $\mathfrak{n}_{\mathbb{Z}}^{-,a,i}$-action. Recall that $v_{\bar{\tau}} = \bar{\tau}(v_0)$ denotes the generator of $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}}$. Then the above discussion can be summarized as

$$\tilde{V}_{\mathbb{Z}}(\ell\tilde{\omega}_{2i})_{\overline{\tau}} = U_{\mathbb{Z}}(\tilde{\mathfrak{n}})\,.\,v_{\overline{\tau}} = U_{\mathbb{Z}}(\mathfrak{n}^{-,a})\,.\,v_{\overline{\tau}} = U_{\mathbb{Z}}(\mathfrak{n}^{-,a,i})\,.\,v_{\overline{\tau}}. \tag{5}$$

Recall that we embed in $(\mathrm{SL}_{2(2m)})_{\mathbb{Z}}$ a copy $L(i)_{\mathbb{Z}}$ of $(\mathrm{SL}_{2m})_{\mathbb{Z}}$ so that we can identify the $\mathrm{SL}_{2(2m)}$-Demazure module $\tilde{V}_{\mathbb{Z}}^{\mathrm{SL}_{2(2m)}}(\ell\tilde{\omega}_{2i})_{\tau}$ generated by $\tau(v_0) = \overline{\tau}(v_0) = v_{\overline{\tau}}$ with the Weyl module $V_{\mathbb{Z}}^{L(i)}(\ell\varpi_i)$ for $L(i)_{\mathbb{Z}}$.

For $1 \le k, l \le 4m$, denote by $X_{k,l}$ the $4m \times 4m$-matrix having a 1 in position (k, l) and whose all other entries 0 (for $k \ne l$, this is the $\mathrm{SL}_{2(2m)}$-root operator corresponding to the $\mathrm{SL}_{2(2m)}$-root $\alpha_{k,l}$), so that

$$\mathfrak{n}^{-,a,i} = \operatorname{span}\{X_{r,s} + X_{4m-s+1,4m-r+1} \mid i+1 \le r \le 2m \text{ and } 2m < s \le 2m+i\}.$$

It is then immediate:

Lemma 6.1. *Every element $y \in \mathfrak{n}_{\mathbb{Z}}^{-,a,i}$ can be written in a unique way as $y = y_1 + y_2$, with $y_1 \in \mathfrak{n}^{-,a,i} \cap \operatorname{Lie} L(i)_{\mathbb{Z}}$ and $y_2 \in \operatorname{span}\{X_{k,l} \mid l > 2m+k\}$. Moreover, y_2 is uniquely determined by y_1.*

By the previous lemma, the projection $p : (\mathrm{SL}_{2(2m)})_{\mathbb{Z}} \to L(i)_{\mathbb{Z}}$ induces an isomorphism of vector spaces $\mathfrak{n}_{\mathbb{Z}}^{-,a,i} \simeq p(\mathfrak{n}_{\mathbb{Z}}^{-,a,i})$. Let us write $\bar{\mathfrak{n}}_{\mathbb{Z}}^{-,a,i}$ for $p(\mathfrak{n}^{-,a,i})$. Since the Lie algebras are commutative, we see that $p : \mathfrak{n}_{\mathbb{Z}}^{-,a,i} \to \bar{\mathfrak{n}}_{\mathbb{Z}}^{-,a,i}$ is not only an isomorphism of vector spaces, but it is in fact a Lie algebra isomorphism.

Corollary 6.2. $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}} = U_{\mathbb{Z}}(\bar{\mathfrak{n}}^{-,a,i})\,.\,v_{\overline{\tau}}.$

Proof. Let $y \in \mathfrak{n}_{\mathbb{Z}}^{-,a,i}$. By Lemma 6.1 we can write $y = p(y) + y_2$ with y_2 in the span of the matrices $X_{i,j}$ with $j > 2m+i$. Therefore, y_2 acts trivially on $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}}$, and we conclude

$$\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}}) = U_{\mathbb{Z}}(\mathfrak{n}^{-,a,i})\,.\,v_{\overline{\tau}} = U_{\mathbb{Z}}(\bar{\mathfrak{n}}^{-,a,i})\,.\,v_{\overline{\tau}}. \qquad \square$$

For $0 \le i \le m-1$, let $\mathrm{Sp}_{2m}(i)_{\mathbb{Z}}$ be a copy of $(\mathrm{Sp}_{2m})_{\mathbb{Z}}$ sitting inside $L(i)_{\mathbb{Z}}$, defined with respect to the form given by the matrix

$$\begin{pmatrix} 0 & J_{m-i} & 0 & 0 \\ -J_{m-i} & 0 & 0 & 0 \\ 0 & 0 & 0 & J_i \\ 0 & 0 & -J_i & 0 \end{pmatrix},$$

where J_r denotes the $r \times r$ antidiagonal matrix with entries $(1, 1, \ldots, 1)$. Moreover, denote $(\mathrm{Sp}_{2m}^{\mathrm{std}})_{\mathbb{Z}} := \mathrm{Sp}_{2m}(0)_{\mathbb{Z}}$.

Let σ be the permutation in $W^{L(i)}$ such that $\operatorname{Lie}\mathrm{Sp}_{2m}(i)_{\mathbb{Z}} = \sigma \operatorname{Lie}(\mathrm{Sp}_{2m}^{\mathrm{std}})_{\mathbb{Z}}\sigma^{-1}$. The permutation σ fixes $2m+1, \ldots, 2m+i$ and moves $2m-i+1, \ldots, 2m$ in front of $i+1,\dots,2m-i$, so that

(6) $$\sigma(e_1\wedge e_2\wedge\cdots\wedge e_i\wedge e_{2m+1}\wedge e_{2m+2}\wedge\cdots\wedge e_{2m+i}) \\ = e_1\wedge e_2\wedge\cdots\wedge e_i\wedge e_{2m+1}\wedge e_{2m+2}\wedge\cdots\wedge e_{2m+i}.$$

Let $\mathfrak{b}_\mathbb{Z}\subset \operatorname{Lie}(\mathrm{Sp}_{2m}^{\mathrm{std}})_\mathbb{Z}$ be the Borel subalgebra of upper triangular matrices. Let $\mathfrak{p}^i_\mathbb{Z}\subset \operatorname{Lie}(\mathrm{Sp}_{2m}^{\mathrm{std}})_\mathbb{Z}$ be the maximal parabolic subalgebra associated with ϖ_i, and $\mathfrak{p}^{i,n}_\mathbb{Z}$ its nilpotent radical. Write $S\mathfrak{p}^{i,n}_\mathbb{Z}$ for $\sigma\mathfrak{p}^{i,n}_\mathbb{Z}\sigma^{-1}\subset \operatorname{Lie}\mathrm{Sp}_{2m}(i)_\mathbb{Z}$.

Lemma 6.3. $$V_\mathbb{Z}(\ell\varpi_i)=U_\mathbb{Z}(S\mathfrak{p}^{i,n}).v_{\overline{\tau}}.$$

Proof. By (6), $v_{\overline{\tau}}$ is a lowest weight vector for $\mathrm{Sp}_{2m}(i)$, as well as for $L(i)$, and the module generated by this vector is $U_\mathbb{Z}(\operatorname{Lie}\mathrm{Sp}_{2m}(i)).v_{\overline{\tau}}=V_\mathbb{Z}(\ell\varpi_i)$. Since it is generated by a lowest weight vector, it is enough to consider the action of the nilpotent radical

$$U_\mathbb{Z}(\operatorname{Lie}\mathrm{Sp}_{2m}(i)).v_{\overline{\tau}}=U_\mathbb{Z}(\sigma\mathfrak{b}\sigma^{-1}).v_{\overline{\tau}}=U_\mathbb{Z}(S\mathfrak{p}^{i,n}).v_{\overline{\tau}}.\qquad\square$$

Observe that since $\bar{\mathfrak{n}}^{-,a,i}_\mathbb{Z}\subseteq S\mathfrak{p}^{i,n}_\mathbb{Z}$, the Weyl module $V_\mathbb{Z}(\ell\varpi_i)$ is naturally equipped with a structure of $\bar{\mathfrak{n}}^{-,a,i}$-module. It is easy to check that:

Lemma 6.4. *Every element $x\in S\mathfrak{p}^{i,n}_\mathbb{Z}$ can be written in a unique way as $x=x_1+x_2$, with $x_1\in\bar{\mathfrak{n}}^{-,a,i}$ and $x_2\in\operatorname{span}\{X_{k,l}\mid 2m-i<k\le 2m$ and $i+1<l<2m\}$. Moreover, x_2 is uniquely determined by x_1.*

As in Section 3, we consider the Chevalley involution $\iota:\mathfrak{sp}_{2m}\to\mathfrak{sp}_{2m}$ such that $\iota|_\mathfrak{h}=-1$ and ι exchanges e_α and $-f_\alpha$. It induces an isomorphism $S^\bullet_\mathbb{Z}(\mathfrak{n}^-)\to S^\bullet_\mathbb{Z}(\mathfrak{n}^+)$, which by abuse we also call ι.

For a dominant weight λ, fix a highest weight vector $v_\lambda\in V_\mathbb{Z}(\lambda)$ and a lowest weight vector $v_{w_0}\in V_\mathbb{Z}(\lambda)$, where w_0 is the longest word in the Weyl group of $\mathfrak{sp}_{2m}$. Recall that considering the PBW filtration on $U_\mathbb{Z}(\mathfrak{n}^-)$ and on $U_\mathbb{Z}(\mathfrak{n}^+)$ provides $V_\mathbb{Z}(\lambda)$ with two possible $S^\bullet_\mathbb{Z}(\mathfrak{n}^{-,a})$-structures: in the first case, looking at the PBW filtration on $V_\mathbb{Z}(\lambda)$ induced by the action of $U_\mathbb{Z}(\mathfrak{n}^-)$ on the highest weight vector and taking the associated graded space provides the abelianized module $V^a_\mathbb{Z}(\lambda)$, while in the second case, looking at the PBW filtration on $V_\mathbb{Z}(\lambda)$ induced by the action of $U_\mathbb{Z}(\mathfrak{n}^+)$ on the lowest weight vector and taking the associated graded space produces a module that we denote by $V^{a,+}_\mathbb{Z}(\lambda)$. Now via ι this module also becomes naturally a $S^\bullet_\mathbb{Z}(\mathfrak{n}^-)$-module and Lemma 3.3 holds in the symplectic case too:

Lemma 6.5. *As a $S^\bullet_\mathbb{Z}(\mathfrak{n}^-)$-module, $V^{a,+}_\mathbb{Z}(\lambda)$ is isomorphic to $V^a_\mathbb{Z}(\lambda)$.*

Observe that in the symplectic case there is no need of replacing λ by λ^* since they coincide.

Lemma 6.6. *The Demazure module $\tilde{V}_\mathbb{Z}(\Psi(\ell\omega_i))_{\overline{\tau}}$ contained in $\tilde{V}_\mathbb{Z}(\Psi(\ell\omega_i))$ and $V_\mathbb{Z}(\ell\omega_i)^a$ are isomorphic as $S^\bullet_\mathbb{Z}(\mathfrak{n}^-)$-modules.*

Proof. By Lemma 6.4, the projection $q : S\mathfrak{p}_{\mathbb{Z}}^{i,n} \to \bar{\mathfrak{n}}^{-,a,i}$ is an isomorphism of vector spaces. Moreover, if we write $x \in S\mathfrak{p}_{\mathbb{Z}}^{i,n}$ as $x = q(x) + x_2$, then x_2 lies in the span of the matrices $X_{i,j}$ with $i+1 < j < 2m$; hence, $x_2 \,.\, v_{\overline{\tau}} = 0$. Now, by [Feigin et al. 2014, Proposition 3.1], the PBW filtrations on $V_{\mathbb{Z}}(\ell\omega_i)$ with respect to the actions of $S\mathfrak{p}_{\mathbb{Z}}^{i,n}$ and $\bar{\mathfrak{n}}_{\mathbb{Z}}^{-,a,i}$ are compatible, and

$$\operatorname{gr} V_{\mathbb{Z}}(\ell\varpi_i) = \operatorname{gr} U_{\mathbb{Z}}(S\mathfrak{p}_{\mathbb{Z}}^{i,n}) \,.\, v_{\overline{\tau}} \simeq \operatorname{gr} U_{\mathbb{Z}}(\bar{\mathfrak{n}}^{-,a,i}) \,.\, v_{\overline{\tau}}.$$

On the other hand, when we consider in $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}}$ the PBW filtration with respect to the action of $S\mathfrak{p}_{\mathbb{Z}}^{i,n}$ and go to the associated graded module, then the action of $(S\mathfrak{p}_{\mathbb{Z}}^{i,n})^a$ is isomorphic to the action of $\bar{\mathfrak{n}}_{\mathbb{Z}}^{-,a,i}$ on $\tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}}$. □

The previous result implies in particular:

Corollary 6.7. $\operatorname{rank} \tilde{V}_{\mathbb{Z}}(\Psi(\ell\omega_i))_{\overline{\tau}} = \operatorname{rank} V_{\mathbb{Z}}(\ell\omega_i)$.

7. The general case for $\mathfrak{sp}_{2m}$

We come now to the general case (notation as in Example 1.2):

Theorem 7.1. *Let λ be a dominant $\mathfrak{sp}_{2m}$-weight. As an $N_{\mathbb{Z},\eta}^{-,a}$-module, the Demazure submodule $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_{\overline{\tau}}$ of the $(\mathrm{Sp}_{2(2m)})_{\mathbb{Z}}$-module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))$ is isomorphic to the abelianized module $V_{\mathbb{Z}}^a(\lambda)$.*

As in the type A case, the proof of the above theorem will provide us with a description of $V_{\mathbb{Z}}^a(\lambda)$ as an $S_{\mathbb{Z}}^{\bullet}(\mathfrak{n}_{\eta}^{-,a})$-module in terms of generators and relations. The abelianized module $V_{\mathbb{Z}}^a(\lambda)$ is a cyclic module over the algebra $S_{\mathbb{Z}}^{\bullet}(\mathfrak{n}_{\eta}^{-,a})$ with the image of a highest weight vector $v_\lambda \in V(\lambda)$ in $V_{\mathbb{Z}}^a(\lambda)$ as a generator; see [Feigin et al. 2013, Proposition 2.3]. Hence the module is isomorphic to $S_{\mathbb{Z}}^{\bullet}(\mathfrak{n}_{\eta}^{-,a})/I_{\mathbb{Z}}(\lambda)$ where $I_{\mathbb{Z}}(\lambda)$ is the annihilator of v_λ in $S_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})$. As a consequence of the proof of Theorem 7.1, we obtain the description of the ideal $I_{\mathbb{Z}}(\lambda)$ in terms of generators given in [Feigin et al. 2011a; 2013] from Mathieu's generator and relation presentation of Demazure modules.

Let $\mathfrak{b}$ be the Borel subalgebra of $\mathfrak{sp}_{2m} = (\mathfrak{sp}_{2m})_{\mathbb{Z}} \otimes \mathbb{C}$ as in Example 1.1(b), so $\mathfrak{sp}_{2m} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$. As free $\mathbb{Z}$-modules, $U_{\mathbb{Z}}(\mathfrak{n}^-) \simeq U_{\mathbb{Z}}(\mathfrak{g})/U_{\mathbb{Z}}^+(\mathfrak{h} + \mathfrak{n}^+)$, so that the adjoint action of $U_{\mathbb{Z}}(\mathfrak{b})$ on $U_{\mathbb{Z}}(\mathfrak{g})$ induces the structures of a $U_{\mathbb{Z}}^+(\mathfrak{b})$- and a $B_{\mathbb{Z}}$-module on $U_{\mathbb{Z}}(\mathfrak{n}^-)$, hence on $S_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})$. This action is compatible with the $B_{\mathbb{Z}}$-action on $V_{\mathbb{Z}}^a(\lambda)$ [Feigin et al. 2013, Proposition 2.3.]. Recall that for a positive root α we have denoted by f_α the corresponding fixed Chevalley basis element in $(\mathfrak{sp}_{2m})_{-\alpha,\mathbb{Z}}$. Let us set

$$R^{++} = \{\epsilon_i - \epsilon_j \mid 1 \le i < j \le m\} \cup \{2\epsilon_i \mid 1 \le i \le m\}$$

As a consequence of the proof of Theorem 7.1 we get the following description of the ideal $I_{\mathbb{Z}}(\lambda)$:

Corollary 7.2. *As a cyclic $S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})$-module, the abelianized module $V^a_{\mathbb{Z}}(\lambda)$ is isomorphic to $S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})/I_{\mathbb{Z}}(\lambda)$, where*

$$I_{\mathbb{Z}}(\lambda) = S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})\big(U_{\mathbb{Z}}(\mathfrak{n}^{+}) \circ \operatorname{span}\{f_{\alpha}^{(\langle\lambda,\alpha^{\vee}\rangle+m)} \mid m \geq 1 \text{ and } \alpha \in R^{++}\}\big) \subseteq S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a}).$$

7A. The proof of the theorem will be only sketched, since the strategy is the same as for the type $\mathtt{A}$ case. We reproduce here the diagram of $S^{\bullet}(\mathfrak{n}_{\eta}^{-,a})$-modules summarizing the main idea: for a dominant weight $\lambda = a_1\omega_1 + \cdots + a_m\omega_m$, we have the natural maps

$$\begin{tikzcd}
S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})/I_{\mathbb{Z}}(\lambda) \arrow[r, "h", "\simeq"'] & V^a_{\mathbb{Z}}(\lambda) \arrow[r, "b"] \arrow[d, "a"] & V^a_{\mathbb{Z}}(a_1\omega_1) \otimes \cdots \otimes V^a_{\mathbb{Z}}(a_m\omega_m) \arrow[d, "c", "\wr"'] \\
S^{\bullet}_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a})/M_{\mathbb{Z}}(\lambda) \arrow[u, dotted, two heads, "f"] \arrow[r, "g", "\simeq"'] & \tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_{\tau} \arrow[r, hook, "d"] & \tilde{V}_{\mathbb{Z}}(a_1\Psi(\omega_1))_{\tau} \otimes \cdots \otimes \tilde{V}_{\mathbb{Z}}(a_m\Psi(\omega_m))_{\tau}.
\end{tikzcd}$$

where in the top row the action on the modules is twisted by the Chevalley involution so the cyclic generators are lowest weight vectors, and the maps c, d, a, and g arise as in the proof of Theorem 4.1 so that again the main difficulty of the proof consists in producing the map f.

7B. The first step consists in determining $M_{\mathbb{Z}}(\lambda)$. By [Mathieu 1989, Lemme 26], the Demazure module $\tilde{V}_{\mathbb{Z}}(\Psi(\lambda))_{\tau}$ is isomorphic to the algebra $U_{\mathbb{Z}}(\tilde{\mathfrak{n}})$ modulo the left ideal $\tilde{I}_{\mathbb{Z}}(\tau\Psi(\lambda))$ generated for all $m \geq 1$ by the elements

$$\begin{cases} E_{k,l}^{(m)} & \text{if } \langle\tilde{\alpha}_{k,l}^{\vee},\, \tau\Psi(\lambda)\rangle \geq 0, \\ E_{k,l}^{(-\langle\tilde{\alpha}_{k,l}^{\vee},\, \tau\Psi(\lambda)\rangle+m)} & \text{otherwise.} \end{cases}$$

7C. The annihilator $M_{\mathbb{Z}}(\lambda)$ is the intersection of $U_{\mathbb{Z}}(\mathfrak{n}_{\eta}^{-,a}) \subset U_{\mathbb{Z}}(\tilde{\mathfrak{n}}^{+})$ with the ideal $\tilde{I}_{\mathbb{Z}}(\overline{\tau}\Psi(\lambda))$. To determine such an intersection, we fix a PBW basis and divide the positive roots in three families, exactly as in the proof of Theorem 4.1.

7D. By Lemma 5.5(i), $\langle\tilde{\alpha}_{k,l}^{\vee},\, \tau\Psi(\lambda)\rangle \geq 0$ if $\tilde{\alpha}_{k,l}$ is of third type. As in type $\mathtt{A}$, we may hence proceed with the calculation modulo the left ideal generated by the divided powers of the corresponding $E_{k,l}$. Modulo such an ideal, by Lemma 5.4 and Lemma 5.5, $\tilde{I}_{\mathbb{Z}}(\overline{\tau}\Psi(\lambda))$ is generated by the $E_{k,l}^{(m)}$ with $m \geq 1$ and $\tilde{\alpha}_{k,l}$ of second type, and the

$$E_{k,l}^{(-\langle\tilde{\alpha}_{k,l}^{\vee},\, \Psi(\lambda)\rangle+m)} \quad \text{with } m \geq 1 \text{ and } \tilde{\alpha}_{k,l} \text{ of first type.}$$

7E. Thus, we consider the subalgebra $\mathfrak{a}$ generated by the $E_{k,l}^{(m)}$, for $\tilde{\alpha}_{k,l}$ of second type. Let $\mathfrak{b}_{\mathbb{Z},\mathfrak{sl}_{2m}}$ be the Borel subalgebra of $\mathfrak{sl}_{2m}$ consisting of traceless upper triangular matrices and let $\mathfrak{b}_{\mathbb{Z}}$ be the corresponding symplectic Borel subalgebra (of $(\mathrm{Sp}_{2m}^{\mathrm{std}})_{\mathbb{Z}}$). Let us embed $\mathfrak{b}_{\mathbb{Z}}$ in $\mathfrak{b}_{\mathbb{Z},\mathfrak{sl}_{2m}} \oplus \mathfrak{b}_{\mathbb{Z},\mathfrak{sl}_{2m}}$ via $A \mapsto (A, -\bar{A})$, where $\bar{A}$ denotes the matrix which is skew-transposed to A, and let $\Delta^{-}(\mathfrak{b}_{\mathbb{Z}})$ be its image. Also $\mathfrak{a}$

is embedded in $\mathfrak{b}_{\mathbb{Z},\mathfrak{sl}_{2m}} \oplus \mathfrak{b}_{\mathbb{Z},\mathfrak{sl}_{2m}}$, once we identify the latter with the Lie algebra generated by the divided powers of the SL_{4m}-root vectors of second type. The image of such an embedding contains $\Delta^-(\mathfrak{b}_\mathbb{Z})$. By taking fixed points with respect to the outer automorphism of $\mathfrak{sl}_{2m}$ and $\mathfrak{sl}_{4m}$ induced by the symmetry of the Dynkin diagram, it follows from Lemma 4.5 that

$$\begin{aligned}(7)\quad U_\mathbb{Z}(\Delta^-(\mathfrak{b}_\mathbb{Z}))\big\langle\{E_{i,j}^{(m)} \mid \alpha_{i,j} \text{ of first type, } m \geq 1\}\big\rangle \\ = U_\mathbb{Z}(\mathfrak{a})\big\langle\{E_{i,j}^{(m)} \mid \alpha_{i,j} \text{ of first type, } m \geq 1\}\big\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}&\tilde{I}_\mathbb{Z}(\tau\Psi(\lambda)) \cap S^\bullet_\mathbb{Z}(\mathfrak{n}_\eta^{-,a}) \\ &\quad\simeq S^\bullet_\mathbb{Z}(\mathfrak{n}_\eta^{-,a}) \circ U_\mathbb{Z}(\Delta^-(\mathfrak{b}_\mathbb{Z}))\,\mathrm{span}\{f_{i,j}^{\langle\alpha^\vee,\lambda\rangle+\ell} \mid \alpha \in R^{++} \text{ and } \ell \geq 1\} =: M_\mathbb{Z}(\lambda).\end{aligned}$$

7F. *Proof of Theorem 7.1.* Since the roots of first type are precisely the ones coming from the elements $f_{i,j}$ with $\alpha_{i,j} \in R^{++}$ and since

$$\big\{f_{i,j}^{(\langle\alpha_{i,j}^\vee,\lambda\rangle+m)} \bigm| \alpha_{i,j} \in R^{++} \text{ and } m \geq 1\big\} \subseteq I_\mathbb{Z}(\lambda),$$

we get a surjective morphism

$$(8)\qquad \tilde{V}_\mathbb{Z}(\Psi\lambda)_{\bar\tau} \overset{g}{\simeq} S^\bullet_\mathbb{Z}(\mathfrak{n}_\eta^{-,a})/M_\mathbb{Z}(\lambda) \xrightarrow{f} S^\bullet_\mathbb{Z}(\mathfrak{n}_\eta^{-,a})/I_\mathbb{Z}(\lambda) \simeq V_\mathbb{Z}(\lambda)^a.$$

This concludes the proof of the theorem. □

7G. Let ρ be the sum of the fundamental weights for Sp_{2m} and let $\tilde\rho = \Psi(\rho)$ be the corresponding dominant weight for $\mathrm{Sp}_{2(2m)}$. Let $Q_\mathbb{Z} \subset (\mathrm{Sp}_{2(2m)})_\mathbb{Z}$ be the corresponding parabolic subgroup. Recall that $N_{\mathbb{Z},\eta}^{-,a}$ is a commutative subgroup of the Borel subgroup $\tilde{B}_\mathbb{Z}$. For any $\mathrm{Sp}_{2(2m)}$-root $\tilde\alpha$, let $U_{\mathbb{Z},\tilde\alpha}$ be the associated root subgroup.

Lemma 7.3. *The orbit $\tilde{B}_\mathbb{Z}.\bar\tau \subset (\mathrm{Sp}_{2(2m)})_\mathbb{Z}/Q_\mathbb{Z}$ is nothing but $N_{\mathbb{Z},\eta}^{-,a}.\bar\tau$, and the map $N_{\mathbb{Z},\eta}^{-,a} \to N_{\mathbb{Z},\eta}^{-,a}.\bar\tau$, given by $u \mapsto u\bar\tau$ is a bijection.*

Proof. We have $\tilde{B}_\mathbb{Z}.\bar\tau = \prod_{\tilde\alpha>0} U_{\mathbb{Z},\alpha}.\bar\tau$, and the map $\prod_{\tilde\alpha\in\Gamma} U_{\mathbb{Z},\tilde\alpha} \to \prod_{\tilde\alpha\in\Gamma} U_{\mathbb{Z},\tilde\alpha}.\bar\tau$ is a bijection, where Γ is the set of all positive roots of $\mathrm{Sp}_{2(2m)}$ such that $\bar\tau^{-1}(\tilde\alpha) < 0$ and $\bar\tau^{-1}(\tilde\alpha)$ is not an element of the root system of $Q_\mathbb{Z}$. Now this condition is fulfilled if and only if $\langle\bar\tau^{-1}(\tilde\alpha^\vee), \tilde\rho\rangle < 0$, or, equivalently, $\langle\tilde\alpha^\vee, \bar\tau(\tilde\rho)\rangle < 0$. By Lemma 5.4 this is not possible if $\tilde\alpha$ is of the form $\tilde\alpha = \tilde\epsilon_i - \tilde\epsilon_j$, with $1 \leq i < j \leq 2m$. For the long roots, this is only possible if $\tilde\alpha = 2\tilde\epsilon_j$, with $j = m+1, \dots, 2m$, and for the roots $\alpha = \tilde\epsilon_i + \tilde\epsilon_j$, with $1 \leq i < j \leq 2m$; this is only possible if either $i, j \geq m+1$ or $1 \leq i \leq m$ and $j = 2m+1-k$ is such that $1 \leq k < i$.

But this implies that the root subgroup $U_{\mathbb{Z},\tilde\alpha}$ is a subgroup of $N_{\mathbb{Z},\eta}^{-,a}$, and all root subgroups of $(\mathrm{Sp}_{2(2m)})_\mathbb{Z}$ contained in $N_{\mathbb{Z},\eta}^{-,a}$ satisfy this condition. It follows that $N_{\mathbb{Z},\eta}^{-,a}.\bar\tau$ is the product of all root subgroups corresponding to positive roots of

$\mathrm{Sp}_{2(2m)}$ such that $\bar{\tau}^{-1}(\tilde{\alpha}) < 0$ and $\bar{\tau}^{-1}(\alpha)$ is not an element of the root system of $Q_{\mathbb{Z}}$ and hence $N^{-,a}_{\mathbb{Z},\eta}.\bar{\tau} = \tilde{B}_{\mathbb{Z}}.\bar{\tau} \subset (\mathrm{Sp}_{2(2m)})_{\mathbb{Z}}/Q_{\mathbb{Z}}$. □

Corollary 7.4. *The degenerate flag variety* $\mathcal{F}\ell(\lambda)^a_k$ *depends only on* supp λ. *It is a projectively normal variety, Frobenius split, with rational singularities.*

Acknowledgements

The work of M. Lanini and P. Littelmann was partially supported by the DFG-Schwerpunkt Grant SP1388. The work of G.C.I. was financed by the national FIRB grant RBFR12RA9W "Perspectives in Lie Theory". We thank Corrado De Concini for many helpful conversations about this project.

References

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie: Chapitres 4, 5, 6*, Hermann, Paris, 1968. MR 0240238

[Cerulli Irelli and Lanini 2015] G. Cerulli Irelli and M. Lanini, "Degenerate flag varieties of type A and C are Schubert varieties", *Int. Math. Res. Not.* **2015**:15 (2015), 6353–6374. MR 3384481

[Feigin 2012] E. Feigin, "$\mathbb{G}_a^M$ degeneration of flag varieties", *Selecta Math.* (*N.S.*) **18**:3 (2012), 513–537. MR 2960025

[Feigin and Finkelberg 2013] E. Feigin and M. Finkelberg, "Degenerate flag varieties of type A: Frobenius splitting and BW theorem", *Math. Z.* **275**:1-2 (2013), 55–77. MR 3101796

[Feigin et al. 2011a] E. Feigin, G. Fourier, and P. Littelmann, "PBW filtration and bases for irreducible modules in type $\mathtt{A}_n$", *Transform. Groups* **16**:1 (2011), 71–89. MR 2785495

[Feigin et al. 2011b] E. Feigin, G. Fourier, and P. Littelmann, "PBW filtration and bases for symplectic Lie algebras", *Int. Math. Res. Not.* **2011**:24 (2011), 5760–5784. MR 2863380

[Feigin et al. 2013] E. Feigin, G. Fourier, and P. Littelmann, "PBW-filtration over $\mathbb{Z}$ and compatible bases for $V_{\mathbb{Z}}(\lambda)$ in type $\mathtt{A}_n$ and $\mathtt{C}_n$", pp. 35–63 in *Symmetries, integrable systems and representations* (Tokyo, 2011 and Lyon, 2011), edited by K. Iohara et al., Springer Proc. Math. Stat. **40**, Springer, Heidelberg, 2013. MR 3077680

[Feigin et al. 2014] E. Feigin, M. Finkelberg, and P. Littelmann, "Symplectic degenerate flag varieties", *Canad. J. Math.* **66**:6 (2014), 1250–1286. MR 3270783

[Jantzen 1987] J. C. Jantzen, *Representations of algebraic groups*, Pure and Applied Mathematics **131**, Academic Press, Boston, 1987. MR 899071

[Lakshmibai et al. 1979] V. Lakshmibai, C. Musili, and C. S. Seshadri, "Geometry of G/P", *Bull. Amer. Math. Soc.* (*N.S.*) **1**:2 (1979), 432–435. MR 520081

[Mathieu 1989] O. Mathieu, "Construction d'un groupe de Kac–Moody et applications", *Compositio Math.* **69**:1 (1989), 37–60. MR 986812

[Mehta and Ramanathan 1988] V. B. Mehta and A. Ramanathan, "Schubert varieties in $G/B \times G/B$", *Compositio Math.* **67**:3 (1988), 355–358. MR 959217

[Ramanan and Ramanathan 1985] S. Ramanan and A. Ramanathan, "Projective normality of flag varieties and Schubert varieties", *Invent. Math.* **79**:2 (1985), 217–224. MR 778124

[Ramanathan 1987] A. Ramanathan, "Equations defining Schubert varieties and Frobenius splitting of diagonals", *Inst. Hautes Études Sci. Publ. Math.* **65** (1987), 61–90. MR 908216

Received November 16, 2015. Revised May 6, 2016.

GIOVANNI CERULLI IRELLI
DIPARTIMENTO DI MATEMATICA
SAPIENZA UNIVERSITÀ DI ROMA
PIAZZALE ALDO MORO 5
I-00185 ROMA
ITALY
cerulli@mat.uniroma1.it
cerulli.math@googlemail.com

MARTINA LANINI
SCHOOL OF MATHEMATICS
UNIVERSITY OF EDINBURGH
JAMES CLERK MAXWELL BUILDING
PETER GUTHRIE TAIT ROAD
EDINBURGH
EH9 3FD
UNITED KINGDOM
m.lanini@ed.ac.uk

PETER LITTELMANN
MATHEMATICAL INSTITUTE
UNIVERSITY OF COLOGNE
WEYERTAL 86-90
D-50931 COLOGNE
GERMANY
peter.littelmann@math.uni-koeln.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.309

SOLITONS FOR THE INVERSE MEAN CURVATURE FLOW

GREGORY DRUGAN, HOJOO LEE AND GLEN WHEELER

We investigate self-similar solutions to the inverse mean curvature flow in Euclidean space. Generalizing Andrews' theorem that circles are the only compact homothetic planar solitons, we apply the Hsiung–Minkowski integral formula to prove the rigidity of the hypersphere in the class of compact expanders of codimension one. We also establish that the moduli space of compact expanding surfaces of codimension two is large. Finally, we update the list of Huisken–Ilmanen's rotational expanders by constructing new examples of complete expanders with rotational symmetry, including topological hypercylinders, called *infinite bottles*, that interpolate between two concentric round hypercylinders.

1. Main results

In this paper, we study self-similar solutions to the inverse mean curvature flow in Euclidean space. After a brief introduction, we present the definitions of the homothetic and translating solitons and discuss the one-dimensional examples. We prove that families of cycloids are the only translating solitons (Theorem 8), and we show how to construct translating surfaces via a tilted product of cycloids.

Next, we consider the rigidity of homothetic solitons. In the class of closed homothetic solitons of codimension one, we prove that round hyperspheres are rigid (Theorem 10). For the higher codimension case, we observe that any minimal submanifold of the standard hypersphere is an expander, so in light of Lawson's construction [1970] of minimal surfaces in $\mathbb{S}^3$, there exist compact embedded expanders for any genus in $\mathbb{R}^4$.

We conclude with an investigation of homothetic solitons with rotational symmetry. First, we construct new examples of complete expanders with rotational symmetry, called *infinite bottles* (see Figure 1), which are topological hypercylinders that interpolate between two concentric round hypercylinders (Theorem 14). Then, we show how the analysis in the proof of Theorem 14 can be used to construct other examples of complete expanders with rotational symmetry, including the examples of Huisken and Ilmanen [1997a].

MSC2010: 53C44.
Keywords: inverse mean curvature flow, self-similar solutions.

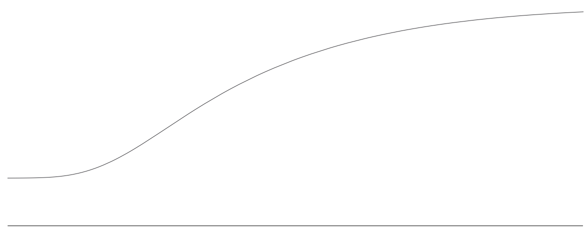


Figure 1. A numerical approximation of the part of a curve whose rotation about the horizontal axis is the self-expanding *infinite bottle* in $\mathbb{R}^3$.

2. Inverse mean curvature flow: history and applications

Round hyperspheres in Euclidean space expand under the inverse mean curvature flow (IMCF) with an exponentially increasing radius. This behavior is typical for the flow. Gerhardt [1990] and Urbas [1990] showed that compact, star-shaped initial hypersurfaces with strictly positive mean curvature converge under IMCF, after suitable rescaling, to a round sphere.

Strictly positive mean curvature is an essential condition. For the IMCF to be parabolic, the mean curvature must be strictly positive. Huisken and Ilmanen [2008] proved that smoothness at later times is characterized by the mean curvature remaining bounded strictly away from zero; see also Smoczyk [2000]. Within the class of strictly mean-convex surfaces, however, a solution to inverse mean curvature flow will, in general, become singular in finite time. For example, starting from a thin embedded torus with positive mean curvature in $\mathbb{R}^3$, the surface fattens up under IMCF and, after finite time, the mean curvature reaches zero at some points [Huisken and Ilmanen 2001, p. 364]. Thus, the classical description breaks down, and any appropriate weak definition of inverse mean curvature flow would need to allow for a change of topology.

Huisken and Ilmanen [2001] used a level-set approach and developed the notion of weak solutions for IMCF to overcome these problems. They showed existence for weak solutions and proved that Geroch's monotonicity [1973] for the Hawking mass carries over to the weak setting. This enabled them to prove the Riemannian Penrose inequality, which also gave an alternative proof for the Riemannian positive mass theorem. For a summary, we refer the reader to Huisken and Ilmanen [1997a; 1997b]. The work of Huisken and Ilmanen also shows that weak solutions become star-shaped and smooth outside some compact region and thus, by the results of Gerhardt [1990] and Urbas [1990], round in the limit. Using a different geometric evolution equation, Bray [2001] proved the most general form of the Riemannian

Penrose inequality. An overview of the different methods used by Huisken, Ilmanen, and Bray can be found in [Bray 2002]. An approach to solving the full Penrose inequality involving a generalized inverse mean curvature flow was proposed in [Bray et al. 2007]. To our knowledge, the full Penrose inequality is still an open problem.

Finally, let us mention some other applications and new developments in IMCF. Using IMCF, Bray and Neves [2004] proved the Poincaré conjecture for 3-manifolds with σ-invariant greater than that of $\mathbb{RP}^3$; see also [Akutagawa and Neves 2007]. Connections with p-harmonic functions and the weak formulation of inverse mean curvature flow are described in [Moser 2007], where a new proof for the existence of a proper weak solution is given, and in [Lee et al. 2011], where gradient bounds and nonexistence results are proved. Recently, Kwong and Miao [2014] discovered a monotone quantity for the IMCF, which they used to derive new geometric inequalities for star-shaped hypersurfaces with positive mean curvature.

3. Definitions and one-dimensional examples

Definition 1 (homothetic solitons of arbitrary codimension). A submanifold Σ^n of $\mathbb{R}^N$ with nonvanishing mean curvature vector field $\vec{H}$ is called a *homothetic soliton for the inverse mean curvature flow* if there exists a constant $C \in \mathbb{R} - \{0\}$ satisfying

$$-\frac{1}{|\vec{H}|^2}\vec{H} = CX^{\perp} \quad \text{on } \Sigma, \tag{1}$$

where the vector field $X^{\perp}$ denotes the normal component of X. Notice that, for any constant $\lambda \neq 0$, the rescaled immersion λX is a soliton with the same value of C.

Remark 2. On a homothetic soliton $\Sigma^n \subset \mathbb{R}^N$, we observe that the condition (1) implies

$$|\vec{H}|^2 = \langle \vec{H}, \vec{H} \rangle = \langle -C|\vec{H}|^2 X^{\perp}, \vec{H} \rangle = -C|\vec{H}|^2 \langle X, \vec{H} \rangle.$$

Since the mean curvature vector field $\vec{H}$ is nonvanishing, this shows

$$-\langle \vec{H}, X \rangle = \frac{1}{C} \quad \text{or} \quad -\langle \triangle_g X, X \rangle = \frac{1}{C} \quad \text{or} \quad \triangle_g |X|^2 = 2\Big(n - \frac{1}{C}\Big),$$

where g denotes the induced metric on Σ.

Proposition 3 (homothetic solitons of codimension one). *Let $\Sigma^n \subset \mathbb{R}^{n+1}$ be a hypersurface with nowhere vanishing mean curvature vector field $\vec{H} = \triangle_g X$. Then, it becomes a homothetic soliton to the inverse mean curvature flow if and only if there exists a constant $C \in \mathbb{R} - \{0\}$ satisfying*

$$-\langle \vec{H}, X \rangle = \frac{1}{C} \quad \textit{or equivalently}, \quad -\langle \triangle_g X, X \rangle = \frac{1}{C}. \tag{2}$$

Proof. According to the observation in Remark 2, the vector equality in (1) implies the scalar equality in (2). To see that (2) implies (1), let $\boldsymbol{N}$ denote a unit normal vector, and let $H = -(\mathrm{div}_\Sigma \boldsymbol{N})$ be the corresponding scalar mean curvature. Then $\vec{H} = \triangle_g X = H\boldsymbol{N}$, and the condition (2) becomes

$$-\langle H\boldsymbol{N}, X\rangle = \frac{1}{C},$$

which implies

$$CX^\perp = \langle \boldsymbol{N}, CX\rangle \boldsymbol{N} = -\frac{1}{H}\boldsymbol{N} = -\frac{1}{H^2}\vec{H}. \qquad \square$$

Remark 4. A complete classification of the homothetic solitons for the inverse curve shortening flow in the plane was established by J. Urbas [1999]. If a plane curve $\mathcal{C}$ is a solution to (2), then its curvature function κ satisfies the Poisson equation

$$\triangle_{\mathcal{C}} \frac{1}{\kappa^2} = 2(C-1),$$

and this guarantees the existence of constants $\alpha_1, \alpha_2 \in \mathbb{R}$ such that

$$\kappa^2 = \frac{1}{(C-1)s^2 + \alpha_1 s + \alpha_2},$$

where s denotes an arc length parameter on the curve $\mathcal{C}$. It is a straightforward exercise to find explicit parametrizations of these homothetic solitons; for instance, see [Castro and Lerma 2016, Section 4]. Examples include circles, involutes of circles, classical logarithmic spirals, epicycloids, and hypocycloids.

Definition 5 (translators of arbitrary codimension). A submanifold $\Sigma^n \subset \mathbb{R}^N$ with nonvanishing mean curvature vector field $\vec{H}$ is called a *translator for the inverse mean curvature flow* if there exists a nonzero constant vector field $\boldsymbol{V}$ satisfying

$$-\frac{1}{|\vec{H}|^2}\vec{H} = \boldsymbol{V}^\perp \quad \text{on } \Sigma, \tag{3}$$

where the vector field $\boldsymbol{V}^\perp$ denotes the normal component of $\boldsymbol{V}$. We say that $\boldsymbol{V}$ is the *velocity* of the translator Σ.

Proposition 6 (translators of codimension one). *Let $\Sigma^n \subset \mathbb{R}^{n+1}$ be a hypersurface with nonvanishing mean curvature vector field $\vec{H} = \triangle_g X$, where g denotes the induced metric on Σ. Then Σ^n is a translator to the inverse mean curvature flow if and only if there exists a nonzero constant vector field $\boldsymbol{V}$ satisfying*

$$\langle \boldsymbol{V}, \vec{H}\rangle = -1. \tag{4}$$

Proof. We first observe that the condition (3) implies the equality

$$-1 = \left\langle -\frac{1}{|\vec{H}|^2}\vec{H}, \vec{H}\right\rangle = \langle \boldsymbol{V}^\perp, \vec{H}\rangle = \langle \boldsymbol{V}, \vec{H}\rangle.$$

It remains to check that the scalar equality (4) implies the vectorial equality in (3). Let $\boldsymbol{N}$ denote a unit normal vector and $H = -(\mathrm{div}_\Sigma \boldsymbol{N})$ its scalar mean curvature, so that $\vec{H} = \triangle_g X = H\boldsymbol{N}$. Then the condition (4) becomes $-1 = \langle \boldsymbol{V}, \vec{H} \rangle = H\langle \boldsymbol{V}, \boldsymbol{N} \rangle$, which implies

$$\boldsymbol{V}^{\perp} = \langle \boldsymbol{V}, \boldsymbol{N} \rangle \boldsymbol{N} = -\frac{1}{H}\boldsymbol{N} = -\frac{1}{H^2}\vec{H}. \qquad \square$$

Corollary 7 (height function on translating hypersurfaces). *A submanifold Σ^n of $\mathbb{R}^{n+1}$ with nonvanishing mean curvature is a* ***translator to the inverse mean curvature flow*** *with velocity $\boldsymbol{V} = (0, \ldots, 0, 1)$ if and only if*

$$-1 = \triangle_\Sigma x_{n+1} \quad \text{on } \Sigma. \tag{5}$$

Now we prove that cycloids are the only one-dimensional translators in $\mathbb{R}^2$.

Theorem 8 (classification of translating curves in $\mathbb{R}^2$). *Any translating curves with unit speed for the inverse mean curvature flow in the Euclidean plane are congruent to cycloids generated by a circle of radius $\frac{1}{4}$.*

Proof. Let the connected curve $\mathcal{C}$ be a translator in the xy-plane with unit velocity $\boldsymbol{V} = (0, 1)$. Adopt the parametrization $X(s) = (x(s), y(s))$, where s denotes the arc length on $\mathcal{C}$, and introduce the tangential angle function $\theta(s)$ such that the tangent $dX/ds = (\cos\theta, \sin\theta)$ and the normal $N(s) = (-\sin\theta, \cos\theta)$. The translator condition reads

$$-\frac{1}{\kappa} = \cos\theta.$$

Now, we integrate

$$\left(\frac{dx}{d\theta}, \frac{dy}{d\theta}\right) = \left(\frac{ds}{d\theta}\frac{dx}{ds}, \frac{ds}{d\theta}\frac{dy}{ds}\right) = \left(\frac{1}{\kappa}\cos\theta, \frac{1}{\kappa}\sin\theta\right) = (-\cos^2\theta, -\cos\theta\sin\theta)$$

to recover, up to translation, the curve

$$(x, y) = \tfrac{1}{4}(-2\theta - \sin(2\theta), 1 + \cos(2\theta)).$$

After introducing the new variable $t = -\pi + 2\theta$, we have

$$(x, y) = \tfrac{1}{4}(-\pi - t + \sin t, 1 - \cos t).$$

Reflecting about the x-axis and then translating along the $(1, 0)$ direction, the translator is congruent to the cycloid represented by $\frac{1}{4}(t - \sin t, 1 - \cos t)$. Therefore, we conclude that $\mathcal{C}$ is congruent to the cycloid through the origin, generated by a circle of radius $\frac{1}{4}$. $\square$

Example 9 (tilted cycloid products: one-parameter family of translators with the same speed in $\mathbb{R}^3$). We can use cycloids (one-dimensional translators in $\mathbb{R}^2$) to construct a one-parameter family of two-dimensional translators with velocity $(0, 0, 1)$ in $\mathbb{R}^3$. Let $(\alpha(s), \beta(s))$ denote a unit speed patch of the translating curve

$\mathcal{C}$ with velocity $(0, 1)$ in the $\alpha\beta$-plane, so that $\beta''(s) = -1$ on the translator $\mathcal{C}$. For each constant $\mu \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, we introduce orthonormal vectors

$$\boldsymbol{v}_1 = (\cos\mu, 0, -\sin\mu), \quad \boldsymbol{v}_2 = (0, 1, 0), \quad \boldsymbol{v}_3 = (\sin\mu, 0, \cos\mu),$$

and associate the product surface $\Sigma_\mu = \mathbb{R} \times \dfrac{1}{\cos\mu}\mathcal{C}$ defined by the patch

$$\boldsymbol{X}(s, h) = h\boldsymbol{v}_1 + \frac{\alpha(s)}{\cos\mu}\boldsymbol{v}_2 + \frac{\beta(s)}{\cos\mu}\boldsymbol{v}_3.$$

A straightforward computation yields

$$\langle \triangle_{\Sigma_\mu} X, (0, 0, 1)\rangle = \left\langle \frac{1}{\cos\mu}(\alpha''(s)\boldsymbol{v}_2 + \beta''(s)\boldsymbol{v}_3), (0, 0, 1)\right\rangle = \beta''(s) = -1,$$

which guarantees that Σ_μ becomes a translator with velocity $(0, 0, 1)$ in $\mathbb{R}^3$.

4. Rigidity of hyperspheres and spherical expanders

We first prove that hyperspheres, as homothetic solitons to the inverse mean curvature flow, are exceptionally rigid. This is a higher-dimensional generalization of Andrews' result [2003, Theorem 1.7] that circles centered at the origin are the only compact homothetic solitons in $\mathbb{R}^2$. We then explain that the moduli space of spherical expanders of higher codimension is large. Hereafter, we assume $n \geq 2$.

Theorem 10 (uniqueness of spheres as compact solitons). *Let Σ^n be a homothetic soliton hypersurface for the inverse mean curvature flow in $\mathbb{R}^{n+1}$. If Σ is closed, then it is a round hypersphere (centered at the origin).*

Proof. Since Σ is a compact hypersurface with nonvanishing mean curvature vector, there exists an inward pointing unit normal vector field $\boldsymbol{N}$ along Σ. Then $\vec{H} = \triangle_g X = H\boldsymbol{N}$, where the scalar mean curvature $H = -\operatorname{div}_\Sigma \boldsymbol{N}$ is positive. Since Σ is a homothetic soliton, we have

$$\frac{1}{C} = -\langle X, \vec{H}\rangle = -H\langle X, \boldsymbol{N}\rangle, \tag{6}$$

for some constant $C \neq 0$. The Hsiung–Minkowski formula [Hsiung 1956] gives

$$0 = \int_\Sigma \left(1 + \frac{1}{n}\langle X, \vec{H}\rangle\right) d\Sigma = \left(1 - \frac{1}{nC}\right)\int_\Sigma 1\, d\Sigma.$$

It follows that $C = 1/n$. Let $\kappa_1, \ldots, \kappa_n$ be principal curvature functions on Σ. In terms of

$$\sigma_2 = \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\kappa_i\kappa_j = \frac{H^2}{n^2} - \frac{1}{n^2(n-1)}\sum_{1\leq i<j\leq n}(\kappa_i - \kappa_j)^2,$$

we have the classical symmetric means inequality

$$\frac{H^2}{n^2}-\sigma_2=\frac{1}{n^2(n-1)}\sum_{1\le i<j\le n}(\kappa_i-\kappa_j)^2\ge 0.$$

Applying the Hsiung–Minkowski formula again, we obtain the integral identity

$$0=\int_\Sigma\Big(\frac{H}{n}+\frac{\sigma_2}{H}\langle X,\vec{H}\rangle\Big)\,d\Sigma=\int_\Sigma\Big(\frac{H}{n}-\frac{n\sigma_2}{H}\Big)\,d\Sigma=\int_\Sigma\frac{n}{H}\Big(\frac{H^2}{n^2}-\sigma_2\Big)\,d\Sigma.$$

Hence, $H^2/n^2-\sigma_2$ vanishes on Σ, which implies that $\kappa_1=\cdots=\kappa_n$ on Σ. Since Σ^n is a closed umbilic hypersurface in Euclidean space, it is a hypersphere. It follows from (6) that this hypersphere is centered at the origin. □

Lemma 11. *A minimal submanifold of the hypersphere $\mathbb{S}^{q\ge 2}$ is an expander for the inverse mean curvature flow in $\mathbb{R}^{q+1}$.*

Proof. Let $\Sigma^{p\ge 1}$ be a minimal submanifold of the hypersphere $\mathbb{S}^q\subset\mathbb{R}^{q+1}$, and let X denote the position vector field in $\mathbb{R}^{q+1}$. On the one hand, since X is already normal to the hypersphere $\mathbb{S}^q\subset\mathbb{R}^{q+1}$, we observe the equality

$$X^\perp:=X^{\perp(\Sigma\subset\mathbb{R}^{q+1})}=X.$$

On the other hand, according to the minimality of Σ^p in $\mathbb{S}^q$, we obtain

$$\triangle_g X+pX=0, \tag{7}$$

where g denotes the induced metric on Σ^p. Thus, we have

$$\vec{H}:=\vec{H}_{\Sigma\subset\mathbb{R}^{q+1}}(X)=\triangle_g X=-pX\quad\text{and}\quad|\vec{H}|=p|X|=p. \tag{8}$$

Combining the four equalities on Σ and taking $C=\dfrac{1}{p}>0$, we get

$$-\frac{1}{|\vec{H}|^2}\vec{H}=CX^\perp,$$

which indicates that Σ is an expander for the inverse mean curvature flow. □

Theorem 12. *For any integer $g\ge 1$, there exists at least one two-dimensional compact embedded expander of genus g in $\mathbb{R}^4$.*

Proof. For any integer g, Lawson [1970] showed that there exists a compact embedded minimal surface Σ of genus g in $\mathbb{S}^3$. Lemma 11 shows that Σ becomes an expander to the inverse mean curvature flow in $\mathbb{R}^4$. □

Remark 13. Castro and Lerma [2016] proved that the converse of Lemma 11 holds.

5. Expanders with rotational symmetry

In this section, we investigate homothetic solitons in $\mathbb{R}^{n+1}$ with rotational symmetry about a line through the origin. To a profile curve $\mathcal{C}$ parametrized by $(r(t), h(t))$ for $t \in I$ in the half-plane $\{(r, h) \mid r > 0, h \in \mathbb{R}\}$, we associate the rotational hypersurface in $\mathbb{R}^{n+1}$ defined by

$$\Sigma^n = \left\{ X = (r(t)\boldsymbol{p}, h(t)) \in \mathbb{R}^{n+1} \mid (r(t), h(t)) \in \mathcal{C},\ \boldsymbol{p} \in \mathbb{S}^{n-1} \subset \mathbb{R}^n \right\}.$$

The rotational hypersurface Σ satisfies the homothetic soliton (2) if and only if the profile curve $(r(t), h(t))$ satisfies the ODE

$$-\left(\frac{\dot{r}\ddot{h} - \dot{h}\ddot{r}}{(\dot{r}^2 + \dot{h}^2)^{3/2}} + \frac{n-1}{(\dot{r}^2 + \dot{h}^2)^{1/2}} \cdot \frac{\dot{h}}{r} \right) \frac{-\dot{h}r + \dot{r}h}{(\dot{r}^2 + \dot{h}^2)^{1/2}} = \frac{1}{C} \tag{9}$$

for some constant $C > 0$. We observe:

i. As long as the quantity $r\dot{h} - h\dot{r}$ is nonzero, we may write (9) as
$$\frac{\dot{r}\ddot{h} - \dot{h}\ddot{r}}{\dot{r}^2 + \dot{h}^2} = -\frac{(n-1)}{r}\dot{h} + \frac{\dot{r}^2 + \dot{h}^2}{C(r\dot{h} - h\dot{r})}.$$

ii. The ODE (9) is invariant under the dilation $(r, h) \mapsto (\lambda r, \lambda h)$, unlike the profile curve equation for shrinkers or expanders for the mean curvature flow.

iii. Spheres are expanders. The half-circle $(r(t), h(t)) = (R\cos t, R\sin t)$ with $t \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ having the origin as its center obeys the ODE (9) with $C = 1/n$.

iv. Cylinders become expanders. The lines $r(t) =$ constant are solutions to the ODE (9) when $C = 1/(n-1)$.

v. We outline a way to deduce the ODE (9) using the homothetic soliton equation
$$\triangle_g |X|^2 = 2\left(n - \frac{1}{C}\right).$$

   We observe that Σ is a homothetic soliton with rotational symmetry if and only if

$$2\left(n - \frac{1}{C}\right) = \triangle_g(r^2 + h^2) = \frac{1}{r^{n-1}\left(\dot{r}^2 + \dot{h}^2\right)^{1/2}} \frac{d}{dt}\left(\frac{r^{n-1}}{\left(\dot{r}^2 + \dot{h}^2\right)^{1/2}} \frac{d}{dt}(r^2 + h^2) \right), \tag{10}$$

   which is equivalent to (9).

5.1. ***Construction of expanding infinite bottles.*** Writing the profile curve $\mathcal{C}$ as a graph $(r(h), h)$, we have the second-order nonlinear differential equation

$$\frac{r''}{1 + r'^2} = \frac{n-1}{r} - \frac{1 + r'^2}{C(r - hr')}. \tag{11}$$

When $C = 1/(n-1)$, this equation becomes

$$\frac{r''}{1+r'^2} = (n-1)\left(\frac{1}{r} - \frac{1+r'^2}{r-hr'}\right). \tag{12}$$

Observe that $r(h) =$ constant is a solution to (12), which corresponds to a round hypercylinder expander. Moreover, if $r(h)$ is a solution to (12) with $r'(a) = 0$ for some $a \in \mathbb{R}$, then $r(h) \equiv r(a)$. Consequently, any nonconstant solution to (12) must be strictly monotone.

In this section, we construct new examples of entire solutions to (12), which correspond to hypercylinder expanders that interpolate between two concentric round hypercylinders.

Theorem 14 (construction of infinite bottles). *Let r_0, h_0, and r'_0 be constants satisfying $r_0 > 0$, $h_0 < 0$, and $r'_0 \in (0, -h_0/r_0)$, and let $r(h)$ be the unique solution to* (12) *satisfying the initial conditions $r(h_0) = r_0$ and $r'(h_0) = r'_0$. Then $r(h)$ is an entire solution, and there are constants $0 < r_{\text{bot}} < r_{\text{top}} < \infty$ such that $r(h)$ interpolates between r_{bot} and r_{top}. More precisely, $r(h)$ is strictly increasing, $\lim_{h\to-\infty} r(h) = r_{\text{bot}}$, $\lim_{h\to\infty} r(h) = r_{\text{top}}$, and there exists a point $h_1 \in (h_0, 0)$ such that $r''(h_1) = 0$ and $r''(h)$ has the same sign as $(h_1 - h)$ when $h \neq h_1$.*

Proof. We separate the proof into two parts. First, we show that the solution is entire and increasing, and there is a unique point where the concavity changes sign. Second, we establish estimates that bound the solution between two positive constants. We note that the rotation of the profile curve about the h-axis has the appearance of an infinite bottle, which interpolates between two concentric cylinders.

Part 1: Existence of expanding infinite bottles.

Notice that the condition $r'(h_0) = r'_0 > 0$ shows that r is a nonconstant solution and guarantees that $r'(h) > 0$. Also, observe that the assumption $r'_0 \in (0, -h_0/r_0)$ coupled with the defining initial conditions for $r(h)$ shows that $h + r'r$ is negative at $h = h_0$. In fact, by assumption, the terms r', $-h - r'r$, r, and $r - hr'$ are all positive at $h = h_0$. So, writing (12) as

$$r'' = (n-1)(1+r'^2)\frac{r'(-h-r'r)}{r(r-hr')}, \tag{13}$$

we see that $r''(h_0) > 0$.

In the following lemma, we show that the concavity of $r(h)$ changes sign exactly once when $r(h)$ is a maximally extended solution.

Lemma 15 (existence of a unique inflection point). *Let $r : (h_{\min}, h_{\max}) \to \mathbb{R}^+$ be a maximally extended solution. Then there exists a point $h_1 \in (h_0, 0)$ such that $r''(h_1) = 0$. Furthermore, $r''(h)$ has the same sign as $(h_1 - h)$ when $h \neq h_1$.*

Proof. Step A. We claim that there exists a point $h_1 \in (h_0, 0)$ such that $r''(h_1) = 0$. We first treat the case where $h_{\max} \leq 0$. In this case, proving the claim is equivalent to showing there is a point $h_1 \in (h_0, h_{\max})$ such that $r''(h_1) = 0$. Suppose to the contrary that

$$r''(h) > 0 \quad \text{for all } h \in (h_0, h_{\max}).$$

As $h_{\max} \leq 0$ and both r and r' are positive, we have $(r - hr') > 0$ for $h \in (h_0, h_{\max})$. In fact, since $(d/dh)(r - hr') = -hr'' > 0$, we see that $(r - hr') > r_0 - h_0 r_0'$. Using (13) and the positivity of the functions r, r', $(r - hr')$, and r'', we arrive at the inequality $(-h - rr') > 0$, which leads to the estimate

$$0 < r'(h) < -\frac{h}{r} < -\frac{h_0}{r_0} \quad \text{for all } h \in (h_0, h_{\max}).$$

Now, returning to (13), we have the estimate

$$0 \leq r''(h) = (n-1)(1 + r'^2)\frac{r'(-h - r'r)}{r(r - hr')} \leq (n-1)\left(1 + \left(\frac{h_0}{r_0}\right)^2\right)\frac{(-h_0/r_0)(-h_0)}{r_0(r_0 - h_0 r_0')}$$

for $h \in (h_0, h_{\max})$. These estimates contradict the finiteness of the maximal endpoint $h_{\max}$, and we conclude that the claim is true in the case where $h_{\max} \leq 0$.

It still remains to prove the claim in the case where $h_{\max} > 0$. However, in this case the solution $r(h)$ is defined when $h = 0$, and (12) implies

$$r''(0) = -(n-1)\frac{r'(0)^2}{r(0)}(1 + r'(0)^2) < 0.$$

It follows that there exists a point $h_1 \in (h_0, 0)$ such that $r''(h_1) = 0$.

Step B. We claim that $r''(h)$ has the same sign as $h_1 - h$. Taking a derivative of (11), we have

$$\frac{r'''}{1 + r'^2} = \frac{2r'(r'')^2}{(1 + r'^2)^2} - \frac{n-1}{r^2}r' - \frac{2r'r''}{C(r - hr')} - \frac{1 + r'^2}{C(r - hr')^2}hr''.$$

At the point h_1, we obtain

$$\frac{r'''(h_1)}{1 + r'(h_1)^2} = -(n-1)\frac{r'(h_1)}{r(h_1)^2} < 0,$$

which shows that $r''(h)$ has the same sign as $h_1 - h$ in a neighborhood of h_1. In fact, at any point $\bar{h}$ where $r''(\bar{h}) = 0$, we have $r'''(\bar{h}) < 0$. This property tells us that the sign of r'' can only change from positive to negative, and consequently r'' vanishes at most once. Thus, $r''(h)$ has the same sign as $h_1 - h$ for all $h \in (h_{\min}, h_{\max})$. $\square$

Next, we prove that the profile curves corresponding to the infinite bottles come from entire graphs.

Lemma 16 (existence of entire solutions). *We have* $h_{\min} = -\infty$ *and* $h_{\max} = \infty$.

Proof. Step A. We claim that $h_{\max} = \infty$. First, we show that $h_{\max} > 0$. To see this, notice that $0 \le r'(h) \le r'(h_1)$, $r(h) \ge r_0$, and $r - hr' \ge r_0$ whenever $h_1 \le h \le 0$. It follows from (12) that the solution $r(h)$ can be extended past $h \le 0$. Thus, $h_{\max} > 0$. Next, we show that $h_{\max} = \infty$. Since $h_1 < 0$, we have $(d/dh)(r - hr') = -hr'' \ge 0$ when $h \ge 0$ so that $(r - hr') \ge r(0)$ when $h \ge 0$. We also have $0 \le r'(h) \le r'(h_1)$ and $r(h) \ge r_0$ when $h \ge 0$. As before, it follows from (12) that the solution $r(h)$ can be extended past any finite point.

Step B. We claim that $h_{\min} = -\infty$. Suppose to the contrary that $h_{\min} > -\infty$. Then at least one of the functions r', $1/r$, or $1/(r - hr')$ must blow up at the finite point $h = h_{\min}$. Since $r'' > 0$ on $(h_{\min}, h_1)$, the positive function r' is increasing, and we have $r'(h) \le r'(h_0) = r'_0$ for all $h \in (h_{\min}, h_0)$. So, the function r' does not blow up at $h_{\min}$. If the function $1/r$ is bounded above on $(h_{\min}, h_0)$, then the inequality $0 < r(h) < r(h) - hr'(h)$ (when $h \le 0$) guarantees that $1/(r - hr')$ is also bounded above on $(h_{\min}, h_0)$, in which case, the solution can be extended prior to $h_{\min}$. Therefore, the function $1/r$ must blow up at $h = h_{\min}$. In other words,

$$\lim_{h \to {h_{\min}}^+} r(h) = 0.$$

Observing this and using $0 < r'(h) < r'_0$ on $(h_{\min}, h_0)$, we can find a sufficiently small $\delta > 0$ so that $r'(h)r(h) \ge -h_0/2$ for all $h \in (h_{\min}, h_{\min} + \delta]$. Also, the inequality $(d/dh)(r - hr') = -hr'' > 0$ guarantees that

$$0 < r(h) - hr'(h) \le \epsilon_1 := r(h_{\min} + \delta) - (h_{\min} + \delta) r'(h_{\min} + \delta).$$

It follows from these estimates and (12) that

$$\frac{d}{dh}(\arctan r') = \frac{r''}{1 + r'^2} = (n-1)\frac{-(h + r'r)}{r - hr'} \cdot \frac{r'}{r} \ge \epsilon_2 \frac{d}{dh}(\ln r),$$

where

$$\epsilon_2 = \frac{(n-1)(-h_0/2)}{\epsilon_1} > 0$$

is a constant. Hence, the function $F(h) := \arctan(dr/dh) - \epsilon_2 \ln r(h)$ is increasing on $(h_{\min}, h_{\min} + \delta]$. Thus, we have the estimate

$$\epsilon_2 \ln r(h) \ge -F(h_{\min} + \delta) + \arctan r' > -F(h_{\min} + \delta).$$

Taking the limit as $h \to {h_{\min}}^+$ and using $\lim_{h \to {h_{\min}}^+} r(h) = 0$ leads to a contradiction. We conclude that $h_{\min} = -\infty$. □

So far, we have proved the existence of an entire bottle solution $r(h)$ to (12). In the next part of the proof we will establish estimates that squeeze the ends of the infinite bottles between two cylinders.

Part 2: Squeezing infinite bottles by two hypercylinders.

To establish upper and lower bounds for the solution $r(h)$, we study the profile curve $\mathcal{C}$ by writing it as a graph over the axis of rotation: $(r, h(r))$. Then, we have the second-order nonlinear differential equation

$$\frac{h''}{1+h'^2} = -\frac{(n-1)}{r}h' + \frac{1+h'^2}{C(rh'-h)}, \tag{14}$$

or equivalently,

$$\frac{h''}{1+h'^2} = \frac{(n-1)hh' + \frac{1}{C}r}{r(rh'-h)} + \left(\frac{1}{C} - (n-1)\right)\frac{h'^2}{(rh'-h)}. \tag{15}$$

Throughout this section, we take $C = 1/(n-1)$, so that (14) takes the form

$$\frac{h''}{1+h'^2} = -(n-1)\left(\frac{h'}{r} - \frac{1+h'^2}{rh'-h}\right) = \frac{n-1}{r}\cdot\frac{r+hh'}{rh'-h}. \tag{16}$$

Now, let $h(r)$ be a maximally extended solution to (16) defined on $(r_{\text{bot}}, r_{\text{top}})$. Lemma 15 tells us that there is a point $r_1 \in (r_{\text{bot}}, r_{\text{top}})$ such that $h'(r) > 0$ and $h''(r) > 0$ for all $r \in (r_1, r_{\text{top}})$ and that $r_1 h'(r_1) - h(r_1) > 0$.

Lemma 17 (existence of the outside cylinder barrier). *We have*

$$r_{\text{top}} < \infty, \qquad \lim_{r\to r_{\text{top}}^-} h'(r) = \infty, \qquad \textit{and} \qquad \lim_{r\to r_{\text{top}}^-} h(r) = \infty.$$

Proof. We introduce the angle functions $\theta, \phi : (r_1, r_{\text{top}}) \to \left(0, \frac{\pi}{2}\right]$, defined by

$$\theta(r) = \arctan\frac{dh}{dr} \quad \text{and} \quad \phi(r) = \arctan\frac{h}{r},$$

to rewrite the profile curve (16) as

$$\frac{d\theta}{dr} = \frac{n-1}{r\cdot\tan(\theta-\phi)}. \tag{17}$$

Combining this and $0 < \tan(\theta-\phi) \le \tan\theta$, we have $d\theta/dr \ge (n-1)/(r\cdot\tan\theta)$, which implies

$$\frac{d}{dr}\left(\frac{\tan\theta}{r^{n-1}}\right) \ge \frac{n-1}{r^n\tan\theta} \ge 0.$$

This tells us that the continuous function $(\tan\theta)/r^{n-1}$ is increasing for $r > r_1$. Set $\theta_1 = \theta(r_1)$. According to the estimate

$$\frac{d}{dr}\left(h - \frac{\tan\theta_1}{nr_1^{n-1}}r^n\right) = \tan\theta - \frac{\tan\theta_1}{r_1^{n-1}}r^{n-1} = \left(\frac{\tan\theta}{r^{n-1}} - \frac{\tan\theta_1}{r_1^{n-1}}\right)r^{n-1} \ge 0,$$

we see that the function

$$h - \frac{\tan\theta_1}{nr_1^{n-1}}r^n$$

is increasing. In particular, we have the height estimate

$$h \geq h_1 + \frac{\tan\theta_1}{n{r_1}^{n-1}}(r^n - r_1^n).$$

Observe that

$$\frac{1}{\tan(\theta-\phi)} = \frac{1+\tan\theta\,\tan\phi}{\tan\theta-\tan\phi} \geq \tan\phi.$$

Combining this with (17), we have

$$\frac{1}{n-1}\frac{d\theta}{dr} \geq \frac{\tan\phi}{r} = \frac{h}{r^2} \geq \frac{1}{r^2}\left(h_1 + \frac{\tan\theta_1}{n{r_1}^{n-1}}(r^n - r_1^n)\right),$$

which implies

$$\frac{d}{dr}\left(\frac{\theta}{n-1} + \left(h_1 - \frac{\tan\theta_1}{n}r_1\right)\frac{1}{r} - \frac{\tan\theta_1}{n(n-1)r_1^{n-1}}r^{n-1}\right) \geq 0.$$

Therefore, the function

$$F(r) = \frac{\theta}{n-1} + \left(h_1 - \frac{\tan\theta_1}{n}r_1\right)\frac{1}{r} - \frac{\tan\theta_1}{n(n-1)r_1^{n-1}}r^{n-1}$$

is increasing, and for all $r \in (r_1, r_{\mathrm{top}})$, we have

$$\frac{\theta}{n-1} \geq F(r_1) - \left(h_1 - \frac{\tan\theta_1}{n}r_1\right)\frac{1}{r} + \frac{\tan\theta_1}{n(n-1)r_1^{n-1}}r^{n-1}.$$

Since the left-hand side is bounded above, and the right-hand side becomes arbitrarily large as r goes to ∞, we conclude that $r_{\mathrm{top}} < \infty$. It then follows that the increasing, concave up function $h(r)$ satisfies $\lim_{r\to r_{\mathrm{top}}^-} h'(r) = \infty$. If $h(r)$ has a finite limit as r approaches r_{top}, then by the uniqueness of the cylinder $r(h) \equiv r_{\mathrm{top}}$, we get a contradiction. Therefore, we also have $\lim_{r\to r_{\mathrm{top}}^-} h(r) = \infty$. □

Next, we prove the following lemma, which shows that a solution with $h < 0$, $h' > 0$, and $h'' < 0$ cannot approach the axis of rotation.

Lemma 18 (existence of the inside cylinder barrier). *We have*

$$r_{\mathrm{bot}} > 0, \qquad \lim_{r\to {r_{\mathrm{bot}}}^+} h'(r) = \infty, \qquad \text{and} \qquad \lim_{r\to {r_{\mathrm{bot}}}^+} h(r) = -\infty.$$

Proof. We first observe that $h - rh' < 0$ and $hh' < 0$. We introduce three well-defined functions $\theta : (r_{\mathrm{bot}}, r_0] \to \left(0, \frac{\pi}{2}\right]$ and $\Psi_1, \Psi_2 : (r_{\mathrm{bot}}, r_0] \to \mathbb{R}$ defined by

$$\theta(r) = \arctan\frac{dh}{dr}, \qquad \Psi_1(r) = \frac{-hh'}{rh'-h}, \qquad \text{and} \qquad \Psi_2(r) = \frac{r+hh'}{hh'},$$

and we rewrite the profile curve (16) as

$$\frac{d\theta}{dr} = -\frac{n-1}{r}\Psi_1\Psi_2. \tag{18}$$

Using the estimate

$$\frac{d\Psi_1}{dr} = \frac{-r(h')^3+h((h')^2+hh'')}{(h-rh')^2} \leq 0,$$

we see that Ψ_1 is decreasing on $(r_{\text{bot}}, r_0]$, and setting $\epsilon_1 = \Psi_1(r_0)$, we have

$$\Psi_1(r) \geq \epsilon_1 > 0. \tag{19}$$

Observing $(hh')' = h'^2 + h''h > 0$ and defining a constant $\epsilon_2 = -h(r_0)h'(r_0) > 0$, we have the estimate $hh' \leq -\epsilon_2$ for all $r \in (r_{\text{bot}}, r_0]$. It follows that

$$\Psi_2(r) = 1 + \frac{r}{hh'} \geq 1 - \frac{r}{\epsilon_2}. \tag{20}$$

Combining (18), (19), and (20), we have

$$\frac{d}{dr}\Big(\frac{\theta}{(n-1)\epsilon_1} + \ln r - \frac{r}{\epsilon_2}\Big) \leq 0.$$

Therefore, the function $\Psi(r) = \dfrac{\theta}{(n-1)\epsilon_1} + \ln r - \dfrac{r}{\epsilon_2}$ is decreasing, and for all $r \in (r_{\text{bot}}, r_0]$, we have

$$\frac{\theta}{(n-1)\epsilon_1} \geq -\ln r + \frac{r}{\epsilon_2} + \Psi(r_0).$$

Since the left-hand side is bounded above, and the right-hand side becomes arbitrarily large as r goes to 0, we conclude that $r_{\text{bot}} > 0$. It then follows that the increasing, concave down function $h(r)$ satisfies $\lim_{r\to r_{\text{bot}}^+} h'(r) = \infty$. If $h(r)$ has a finite limit as r approaches r_{bot}, then by comparison with the cylinder $r(h) \equiv r_{\text{bot}}$, we get a contradiction. Therefore, we also have $\lim_{r\to r_{\text{bot}}^+} h(r) = -\infty$.

This completes the proof of both the lemma and Theorem 14. □

5.2. ***Other examples of complete solitons.*** Huisken and Ilmanen [1997a] used a phase-plane analysis to exhibit complete, rotationally symmetric expanders for the inverse mean curvature flow which are topological hyperplanes. For each $C > 1/n$, they showed there exists a half-entire solution to (11) which intersects the h-axis perpendicularly, and they provided numeric descriptions of these profile curves. For $C > 1/n$ and $C \neq 1/(n-1)$, they also indicated the existence of entire solutions to (11) which are symmetric about the r-axis and correspond to topological hypercylinders. (We note that the rotational expander constructed in Theorem 14 is nonsymmetric in the sense that its profile curve is not symmetric about the r-axis.) In this section, we explain how the techniques from Section 5.1 can be used to recover the examples and numeric pictures presented in [Huisken and Ilmanen 1997a].

Hyperplane expanders. We begin by considering the initial value problem where we shoot perpendicularly to the axis of rotation. For $C > 0$, let $h(r)$ be a solution to (14) with $h(0) = h_0 < 0$ and $h'(0) = 0$. This singular shooting problem is well-defined (see [Baouendi and Goulaouic 1976] and [Drugan 2015]), and the solution satisfies $h''(0) = -1/(nCh_0) > 0$. Differentiating (14) and analyzing the equation for $h'''(r)$ shows that, under the above conditions, we have $h''(r) > 0$ and $h'(r) > 0$, for $r > 0$, as long as the solution is defined. The global behavior of the solution ultimately depends on the value of C.

When $h(r)$ is a solution to the above shooting problem, the graph $(r, h(r))$ is part of a profile curve $\mathcal{C}$, which corresponds to a rotational expander for the inverse mean curvature flow. Applying the techniques from the proof of Theorem 14 to the profile curve $\mathcal{C}$ leads to a description of the global behavior of this expander, which ultimately depends on the value of $C > 1/n$. In terms of the profile curve $\mathcal{C}$ written as a graph over the h-axis, we have the following result.

Theorem 19. *For $C > 1/n$ and $h_0 < 0$, there exists a half-entire solution $r(h)$ to* (11) *that is defined for $h > h_0$, and such that the curve $(h, r(h))$ intersects the h-axis perpendicularly when $h = h_0$. The solution $r(h)$ has three types of behavior, depending on the value of C*:

(1) *If $C = 1/(n-1)$, then $r' > 0$, $r'' < 0$, and there exists $0 < r_{\text{top}} < \infty$ such that $\lim_{h\to\infty} r(h) = r_{\text{top}}$.*

(2) *If $C > 1/(n-1)$, then $r' > 0$, $r'' < 0$, and $\lim_{h\to\infty} r(h) = \infty$.*

(3) *If $1/n < C < 1/(n-1)$, then there exists a point h_1 such that $r''(h)$ has the same sign as $(h - h_1)$, and $\lim_{h\to\infty} r(h) = 0$.*

Proof. When $C = 1/(n-1)$, the convexity of $h(r)$ along with the analysis from Lemma 17 shows that there is a point $r_{\text{top}} < \infty$ such that $\lim_{r\to r_{\text{top}}^-} h'(r) = \infty$ and $\lim_{r\to r_{\text{top}}^-} h(r) = \infty$. Written as a graph over the h-axis, this shows that there is a solution $r(h)$ to (11), defined for $h > h_0$, which intersects the h-axis perpendicularly at h_0 and satisfies $r' > 0$, $r'' < 0$, and $\lim_{h\to\infty} r(h) = r_{\text{top}}$.

Next, when $C > 1/(n-1)$, we claim that the solution $h(r)$ must exist for all $r > 0$. To see this, suppose to the contrary that h' increases to ∞ at a point $r_{\text{top}} < \infty$. Then, since $C > 1/(n-1)$, (14) forces $h \geq \epsilon r h'$ when r is close to r_{top}, for some $\epsilon > 0$. However, integrating this inequality shows that h' does not blow up at a finite point; hence the solution exists for all $r > 0$. Therefore, the solution $h(r)$ exists for all $r > 0$, and using $h'' > 0$ and $h' > 0$, we have $\lim_{r\to\infty} h(r) = \infty$. Written as a graph over the h-axis, this shows that there is a solution $r(h)$ to (11), defined for $h > h_0$, which intersects the h-axis perpendicularly at h_0 and satisfies $r' > 0$, $r'' < 0$, and $\lim_{h\to\infty} r(h) = \infty$.

Finally, when $1/n < C < 1/(n-1)$, the factor $\frac{1}{C} - (n-1)$ in (15) is positive and the analysis in Lemma 17 can be used to show that $h(r)$ does not exist for all

$r > 0$. Moreover, using the positivity of $\frac{1}{C} - (n-1)$ and integrating (14), we arrive at an inequality that provides an upper bound for h. In terms of the profile curve written as a graph over the h-axis, this says that the solution $r(h)$ achieves a global maximum at a finite point. Reading (9) in polar coordinates, we can show that $r(h)$ is defined for $h > h_0$. This forces the concavity of $r(h)$ to change sign at a finite point, and as in the proof of Lemma 15, it follows that there is a point h_1 such that $r''(h)$ has the same sign as $(h - h_1)$. Then, an argument similar to the one in the previous paragraph shows that $r(h)$ is not bounded below by a positive constant, and we conclude that $\lim_{h\to\infty} r(h) = 0$. □

We remark that when $1/n < C < 1/(n-1)$, the analogue of Lemma 17 holds, but as we saw in the proof of the previous theorem, the analogue of Lemma 18 is not true. Similarly, if $C > 1/(n-1)$, then the analogue of Lemma 18 holds, but the analogue of Lemma 17 does not.

Hypercylinder expanders. We finish this section with a result on the construction of rotational expanders that are topological hypercylinders.

Theorem 20. *For $C > 1/n$ and $r_0 > 0$, there is a unique solution $r(h)$ to* (11) *that is symmetric about the r-axis and satisfies the initial condition $r(0) = r_0$, $r'(0) = 0$. The solution $r(h)$ has three types of behavior, depending on the value of C:*

(1) *If $C = 1/(n-1)$, then $r(h) \equiv r_0$ (which corresponds to the round hypercylinder).*

(2) *If $C > 1/(n-1)$, then $r(h)$ has a global minimum at $h = 0$, and there exists a point $h_1 > 0$ such that $r''(h)$ has the same sign as $(h_1 - |h|)$. Also, $\lim_{h\to\infty} r(h) = \infty$.*

(3) *If $1/n < C < 1/(n-1)$, then $r(h)$ has a global maximum at $h = 0$, and there exists a point $h_1 > 0$ such that $r''(h)$ has the same sign as $(|h| - h_1)$. Also, $\lim_{h\to\infty} r(h) = 0$.*

Proof. It follows from (11) that the condition $r'(0) = 0$ forces the solution to be constant when $C = 1/(n-1)$, to have a global minimum at $h = 0$ when $C > 1/(n-1)$, and to have a global maximum at $h = 0$ when $1/n < C < 1/(n-1)$. To see that there is a finite point $h_1 > 0$ where the concavity of $r(h)$ changes sign when $C > 1/(n-1)$, we first observe that $r(h)$ is increasing when $h > 0$, and consequently, it is defined for all $h > 0$. An analysis of (14) shows that a positive solution $h(r)$ cannot satisfy $h''(r) < 0$ and $h'(r) > 0$ for all $r > 0$ when $C > 1/(n-1)$; hence, there is a finite point $h_1 > 0$ where the concavity of $r(h)$ changes sign. When $1/n < C < 1/(n-1)$, the analysis in the proof of Theorem 19 can be used to show that the concavity of $r(h)$ changes sign at a finite point $h_1 > 0$. The proofs of the remaining properties are similar to the proofs given for Theorems 14 and 19. □

Acknowledgement

We warmly thank the referees for their careful reading of the paper and their suggestions.

References

[Akutagawa and Neves 2007] K. Akutagawa and A. Neves, "3-manifolds with Yamabe invariant greater than that of $\mathbb{RP}^3$", *J. Differential Geom.* **75**:3 (2007), 359–386. MR 2301449 Zbl 1119.53027

[Andrews 2003] B. Andrews, "Classification of limiting shapes for isotropic curve flows", *J. Amer. Math. Soc.* **16**:2 (2003), 443–459. MR 1949167 Zbl 1023.53051

[Baouendi and Goulaouic 1976] M. S. Baouendi and C. Goulaouic, "Singular nonlinear Cauchy problems", *J. Differential Equations* **22**:2 (1976), 268–291. MR 0435564 Zbl 0344.35012

[Bray 2001] H. L. Bray, "Proof of the Riemannian Penrose inequality using the positive mass theorem", *J. Differential Geom.* **59**:2 (2001), 177–267. MR 1908823 Zbl 1039.53034

[Bray 2002] H. L. Bray, "Black holes, geometric flows, and the Penrose inequality in general relativity", *Notices Amer. Math. Soc.* **49**:11 (2002), 1372–1381. MR 1936643 Zbl 1126.83304

[Bray and Neves 2004] H. L. Bray and A. Neves, "Classification of prime 3-manifolds with Yamabe invariant greater than $\mathbb{RP}^3$", *Ann. of Math.* (2) **159**:1 (2004), 407–424. MR 2052359

[Bray et al. 2007] H. Bray, S. Hayward, M. Mars, and W. Simon, "Generalized inverse mean curvature flows in spacetime", *Comm. Math. Phys.* **272**:1 (2007), 119–138. MR 2291804 Zbl 1147.53052

[Castro and Lerma 2016] I. Castro and A. M. Lerma, "Lagrangian homothetic solitons for the inverse mean curvature flow", *Results Math.* (online publication July 2016).

[Drugan 2015] G. Drugan, "An immersed S^2 self-shrinker", *Trans. Amer. Math. Soc.* **367**:5 (2015), 3139–3159. MR 3314804 Zbl 06429009

[Gerhardt 1990] C. Gerhardt, "Flow of nonconvex hypersurfaces into spheres", *J. Differential Geom.* **32**:1 (1990), 299–314. MR 1064876 Zbl 0708.53045

[Geroch 1973] R. Geroch, "Energy extraction", *Ann. N. Y. Acad. Sci.* **224**:1 (1973), 108–117. Zbl 0942.53509

[Hsiung 1956] C.-C. Hsiung, "Some integral formulas for closed hypersurfaces in Riemannian space", *Pacific J. Math.* **6** (1956), 291–299. MR 0082160 Zbl 0071.37502

[Huisken and Ilmanen 1997a] G. Huisken and T. Ilmanen, "A note on the inverse mean curvature flow", 1997, available at https://people.math.ethz.ch/~ilmanen/papers/saitama.ps. Presented at the Workshop on Nonlinear Partial Differential Equations, Saitama University, 1997.

[Huisken and Ilmanen 1997b] G. Huisken and T. Ilmanen, "The Riemannian Penrose inequality", *Internat. Math. Res. Notices* **1997**:20 (1997), 1045–1058. MR 1486695 Zbl 0905.53043

[Huisken and Ilmanen 2001] G. Huisken and T. Ilmanen, "The inverse mean curvature flow and the Riemannian Penrose inequality", *J. Differential Geom.* **59**:3 (2001), 353–437. MR 1916951 Zbl 1055.53052

[Huisken and Ilmanen 2008] G. Huisken and T. Ilmanen, "Higher regularity of the inverse mean curvature flow", *J. Differential Geom.* **80**:3 (2008), 433–451. MR 2472479 Zbl 1161.53058

[Kwong and Miao 2014] K.-K. Kwong and P. Miao, "A new monotone quantity along the inverse mean curvature flow in $\mathbb{R}^n$", *Pacific J. Math.* **267**:2 (2014), 417–422. MR 3207590 Zbl 1295.53074

[Lawson 1970] H. B. Lawson, Jr., "Complete minimal surfaces in S^3", *Ann. of Math.* (2) **92** (1970), 335–374. MR 0270280 Zbl 0205.52001

[Lee et al. 2011] Y.-I. Lee, A.-N. Wang, and S. W. Wei, "On a generalized 1-harmonic equation and the inverse mean curvature flow", *J. Geom. Phys.* **61**:2 (2011), 453–461. MR 2746129 Zbl 1220.53078

[Moser 2007] R. Moser, "The inverse mean curvature flow and p-harmonic functions", *J. Eur. Math. Soc.* (*JEMS*) **9**:1 (2007), 77–83. MR 2283104 Zbl 1116.53040

[Smoczyk 2000] K. Smoczyk, "Remarks on the inverse mean curvature flow", *Asian J. Math.* **4**:2 (2000), 331–335. MR 1797584 Zbl 0989.53040

[Urbas 1990] J. I. E. Urbas, "On the expansion of starshaped hypersurfaces by symmetric functions of their principal curvatures", *Math. Z.* **205**:3 (1990), 355–372. MR 1082861 Zbl 0691.35048

[Urbas 1999] J. Urbas, "Convex curves moving homothetically by negative powers of their curvature", *Asian J. Math.* **3**:3 (1999), 635–656. MR 1793674 Zbl 0970.53039

Received October 25, 2015. Revised April 26, 2016.

GREGORY DRUGAN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OREGON
EUGENE, OR 97403-1222
UNITED STATES
drugan@uoregon.edu

HOJOO LEE
CENTER FOR MATHEMATICAL CHALLENGES
KOREA INSTITUTE FOR ADVANCED STUDY
HOEGIRO 85
DONGDAEMUN-GU
SEOUL 02455
SOUTH KOREA
momentmaplee@gmail.com

GLEN WHEELER
INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS
UNIVERSITY OF WOLLONGONG
NORTHFIELDS AVENUE
WOLLONGONG NSW 2522
AUSTRALIA
glenw@uow.edu.au

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.327

BERGMAN THEORY OF CERTAIN GENERALIZED HARTOGS TRIANGLES

LUKE D. EDHOLM

The Bergman theory of domains $\{|z_1|^\gamma < |z_2| < 1\}$ in $\mathbb{C}^2$ is studied for certain values of γ, including all positive integers. For such γ, we obtain a closed form expression for the Bergman kernel $\mathbb{B}_\gamma$. With these formulas, we make new observations relating to the Lu Qi-Keng problem and analyze the boundary behavior of $\mathbb{B}_\gamma(z, z)$.

1. Introduction

For a domain $\Omega \subset \mathbb{C}^n$, the Bergman space is the set of square-integrable, holomorphic functions on Ω. The Bergman kernel is a reproducing integral kernel on the Bergman space that is indispensable to the study of holomorphic functions in several complex variables. The purpose of this paper is to understand Bergman theory for a class of bounded, pseudoconvex domains in $\mathbb{C}^2$. Define the *generalized Hartogs triangle of exponent* $\gamma > 0$ to be the domain

$$\mathbb{H}_\gamma = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^\gamma < |z_2| < 1\}. \tag{1.1}$$

$\mathbb{H}_1$ is the "classical" Hartogs triangle, a well-known pseudoconvex domain with nontrivial Nebenhülle. When $\gamma > 1$, we call $\mathbb{H}_\gamma$ a *fat Hartogs triangle*, and when $0 < \gamma < 1$, we call $\mathbb{H}_\gamma$ a *thin Hartogs triangle*. Our main results are the following two computations.

Theorem 1.2. *Let $s := z_1\bar{w}_1$, $t := z_2\bar{w}_2$, and $k \in \mathbb{Z}^+$. The Bergman kernel for the fat Hartogs triangle $\mathbb{H}_k$ is given by*

$$\mathbb{B}_k(z, w) = \frac{p_k(s)t^2 + q_k(s)t + s^k p_k(s)}{k\pi^2(1-t)^2(t-s^k)^2}, \tag{1.3}$$

where p_k and q_k are the polynomials

$$p_k(s) = \sum_{l=1}^{k-1} l(k-l)s^{l-1}, \qquad q_k(s) = \sum_{l=1}^{k} (l^2 + (k-l)^2 s^k)s^{l-1}.$$

MSC2010: 32A25, 32A36, 32W05.
Keywords: Bergman kernel, fat Hartogs triangle, Bell's transformation formula, Lu Qi-Keng.

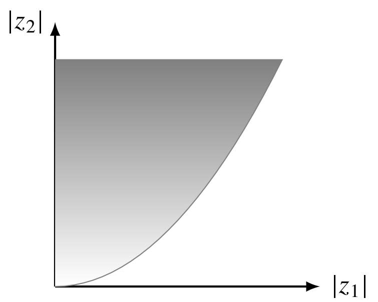
Theorem 1.4. *Let $s = z_1\bar{w}_1$, $t = z_2\bar{w}_2$, and $k \in \mathbb{Z}^+$. The Bergman kernel for the thin Hartogs triangle $\mathbb{H}_{1/k}$ is given by*

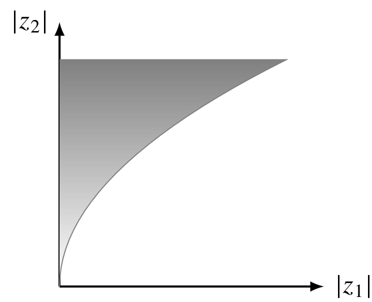
$$\mathbb{B}_{1/k}(z,w) = \frac{t^k}{\pi^2(1-t)^2(t^k - s)^2}. \tag{1.5}$$

There has been an extensive amount of research devoted to understanding Bergman kernels of various classes of domains, and there are several instances in which explicit formulas for the kernel have been obtained. The most common method involves summing an infinite series, which is done in [D'Angelo 1978; 1994; Park 2008]. In [Boas et al. 1999], explicit formulas for the Bergman kernel are produced using other techniques which avoid infinite series altogether. But these situations are exceptional, and in most cases it is impossible to express the Bergman kernel in closed form.

Despite the difficulty of producing explicit formulas, powerful estimates on the Bergman kernel have been given for many classes of pseudoconvex domains. Fefferman [1974] develops an asymptotic expansion of the kernel on smoothly bounded, strongly pseudoconvex domains in $\mathbb{C}^n$. Useful estimates also exist for large classes of smoothly bounded, weakly pseudoconvex domains. See [Catlin 1989; McNeal 1989; 1994; Nagel et al. 1989] for some of the principal results on *finite type* domains, and [Fu 2014] for domains with locally smooth boundaries and constant Levi-rank.

At present, there are no general theorems about the behavior of the Bergman kernel on pseudoconvex domains near unsmooth boundary points, which adds to the intrigue of Theorems 1.2 and 1.4. Each generalized Hartogs triangle defined by (1.1) has two very different kinds of boundary irregularities: the "corner points" which occur at the intersection of the two bounding real hypersurfaces, and the origin singularity, near which $b\mathbb{H}_k$ cannot be expressed as the graph of a continuous function.





This is one of several recent papers to study holomorphic function theory on domains with similar kinds of boundary singularities. Chakrabarti and Shaw [2013] investigate the Sobolev regularity of the $\bar{\partial}$-equation on the classical Hartogs triangle.

Chakrabarti and Zeytuncu [2016] study the L^p-mapping properties of the Bergman projection on the classical Hartogs triangle, and Chen [2013] studies L^p-mapping of the Bergman projection on analogous domains in higher dimensions. Zapałowski [2016] characterizes proper maps between generalizations of the Hartogs triangle in $\mathbb{C}^n$. This author and McNeal investigate the Bergman projection on fat Hartogs triangles in [Edholm and McNeal 2016]. It can be hoped that by understanding the Bergman theory on example domains with boundary singularities such as $\mathbb{H}_\gamma$, we can gain deeper insight into the situation on more general domains.

2. Preliminaries

Bergman theory. Here we highlight some basic facts about Bergman theory that are used throughout this paper. See [Krantz 1992] for a more detailed treatment. If $\Omega \subset \mathbb{C}^n$ is a domain, let $\mathcal{O}(\Omega)$ denote the holomorphic functions on Ω. The standard L^2 inner product is denoted by

$$\langle f, g\rangle = \int_\Omega f \cdot \bar{g}\, dV, \tag{2.1}$$

where dV denotes Lebesgue measure on $\mathbb{C}^n$. $L^2(\Omega)$ denotes the measurable functions f such that $\langle f, f\rangle = \|f\|^2 < \infty$. We define the *Bergman space* $A^2(\Omega) := \mathcal{O}(\Omega) \cap L^2(\Omega)$.

$A^2(\Omega)$ is a Hilbert space with inner product (2.1), and for all $z \in \Omega$, the evaluation functional $\mathrm{ev}_z : f \mapsto f(z)$ is continuous. Therefore, the Riesz representation theorem guarantees the existence of a function $\mathbb{B}_\Omega : \Omega \times \Omega \to \mathbb{C}$ satisfying

$$f(z) = \int_\Omega \mathbb{B}_\Omega(z, w) f(w)\, dV(w), \qquad f \in A^2(\Omega). \tag{2.2}$$

We call $\mathbb{B}_\Omega$ the *Bergman kernel*, and when context is clear we may omit the subscript. In addition to reproducing functions in the Bergman space via (2.2), the Bergman kernel is conjugate symmetric and for each fixed $w \in \Omega$, $\mathbb{B}(\cdot, w) \in A^2(\Omega)$.

Given an orthonormal Hilbert space basis $\{\phi_\alpha\}_{\alpha\in\mathcal{A}}$ for $A^2(\Omega)$, the Bergman kernel is given by the formula

$$\mathbb{B}(z, w) = \sum_{\alpha\in\mathcal{A}} \phi_\alpha(z)\overline{\phi_\alpha(w)}, \tag{2.3}$$

which is independent of the choice of the basis.

Finally, the Bergman kernel transforms under biholomorphisms in the following way: Let $F : \Omega \to \widetilde{\Omega}$ be a biholomorphic map of domains in $\mathbb{C}^n$. Then

$$\mathbb{B}_\Omega(z, w) = \det F'(z) \cdot \mathbb{B}_{\widetilde{\Omega}}(F(z), F(w)) \cdot \overline{\det F'(w)}. \tag{2.4}$$

The Bergman kernel of $\mathbb{H}_1$**.** The formula for the Bergman kernel of the classical Hartogs triangle has been known for quite some time, at least since Bremermann's [1955] paper. Following the spirit of Bremermann's argument, we use formula (2.4) to compute $\mathbb{B}_{\mathbb{H}_1}$. The map given by $\psi(z_1, z_2) = (z_1/z_2, z_2)$ is a biholomorphism of $\mathbb{H}_1$ onto $D \times D^*$, where D is the unit disc and D^* is the punctured disc. It's easy to see that the Bergman kernel of $D \times D^*$ is the same as that of $D \times D$, which is well known and given by

$$\mathbb{B}_{D\times D}(z, w) = \frac{1}{\pi^2(1 - z_1\bar{w}_1)^2(1 - z_2\bar{w}_2)^2} = \mathbb{B}_{D\times D^*}(z, w). \tag{2.5}$$

Seeing that $\det \psi'(z) = 1/z_2$, (2.4) says

$$\mathbb{B}_{\mathbb{H}_1}(z, w) = \frac{z_2\bar{w}_2}{\pi^2(1 - z_2\bar{w}_2)^2(z_2\bar{w}_2 - z_1\bar{w}_1)^2}. \tag{2.6}$$

Because of this computation, Theorems 1.2 and 1.4 only need to be proved for integers $k \geq 2$. Note the polynomial $p_1(s)$ in (1.3) is vacuously equal to 0.

Distance to the boundary and asymptotic growth rates. The following notation will be used in and around Theorem 4.9. Given any $z \in \Omega$, define the distance to the boundary of Ω by the function

$$\delta_\Omega(z) := \min\{\|z - \zeta\| : \zeta \in b\Omega\},$$

where $\|\cdot\|$ denotes Euclidean distance. When the context is clear, we may omit the subscript. We will also use the following notation to write inequalities. If A and B are functions depending on several variables, write $A \lesssim B$ to mean that there is a constant $K > 0$, independent of relevant variables, such that $A \leq K \cdot B$. The independence of which variables will be clear in context. Also write $A \approx B$ to mean that $A \lesssim B \lesssim A$.

3. Bell's transformation rule and derivation of the kernel

Equation (2.4) says that the Bergman kernels of two biholomorphic domains are related by a simple formula. But applications of this transformation rule remain limited by the fact that it's rare to expect two domains in $\mathbb{C}^n$ to be biholomorphic. There is, however, a more general version of this transformation rule. Bell [1982] proves a generalization which applies whenever we have two domains and a proper holomorphic map from one onto the other. The statement of this more general transformation rule appears below, and it will be essential to our proof of Theorem 1.2.

Recall the classical fact that any holomorphic, proper map of Ω onto $\widetilde{\Omega}$ is necessarily a branched covering of finite order.

Theorem 3.1 (Bell's transformation rule). *Let Ω and $\widetilde{\Omega}$ be domains in $\mathbb{C}^n$ with respective Bergman kernels $\mathbb{B}$ and $\widetilde{\mathbb{B}}$, and suppose ϕ is a proper holomorphic map of order k from Ω onto $\widetilde{\Omega}$. Let $u := \det[\phi']$, and let $\Phi_1, \Phi_2, \ldots, \Phi_k$ be the branch inverses of ϕ defined locally on $\widetilde{\Omega} - V$, where $V := \{\phi(z) : u(z) = 0\}$. Finally, write $U_j := \det[\Phi_j']$. Then*

$$u(z)\widetilde{\mathbb{B}}(\phi(z), w) = \sum_{j=1}^{k} \mathbb{B}(z, \Phi_j(w))\overline{U_j(w)}. \tag{3.2}$$

We're now ready to compute the Bergman kernel of fat Hartogs triangles with integer exponents. For the rest of this paper, we'll denote the Bergman kernel of $\mathbb{H}_\gamma$ by $\mathbb{B}_\gamma$.

Proof of Theorem 1.2. First we need to define the map ϕ and its local inverses $\Phi_1, \ldots, \Phi_k$. For each integer $k \geq 2$, the function $\phi : \mathbb{H}_1 \to \mathbb{H}_k$ given by

$$\phi(z) = (z_1, z_2^k) := (\phi_1(z), \phi_2(z))$$

is a branched cover of order k, since

$$\begin{aligned} |\phi_1(z)|^k < |\phi_2(z)| < 1 &\iff |z_1|^k < |z_2^k| < 1 \\ &\iff |z_1| < |z_2| < 1. \end{aligned}$$

We note $u(z) = kz_2^{k-1}$, so V is the set $\{z_2 = 0\}$, which is disjoint from $\mathbb{H}_k$. For each $j = 1, \ldots, k$, the map $\Phi_j(z) = \left(z_1, \zeta^j z_2^{1/k}\right)$ defines a local inverse of ϕ, where $\zeta = e^{2\pi i/k}$ and $z_2^{1/k}$ is taken to mean the root with argument in the interval $[0, 2\pi/k)$. From this we see $U_j(z) = \left(\zeta^j z_2^{1/k-1}\right)/k$. We now apply Bell's rule (3.2):

$$\begin{aligned} &\mathbb{B}_k\big((z_1, z_2^k), (w_1, w_2)\big) \\ &\quad = \frac{z_2 \bar{w}_2^{1/k}}{k^2 z_2^k \bar{w}_2} \sum_{j=1}^{k} \mathbb{B}_1\big((z_1, z_2), (w_1, \zeta^j w_2^{1/k})\big)\bar{\zeta}^j \\ &\quad = \frac{z_2^2 \bar{w}_2^{2/k}}{\pi^2 k^2 z_2^k \bar{w}_2} \sum_{j=1}^{k} \frac{\bar{\zeta}^{2j}}{\big(1 - z_2\bar{w}_2^{1/k}\bar{\zeta}^j\big)^2\big(z_2\bar{w}_2^{1/k}\bar{\zeta}^j - z_1\bar{w}_1\big)^2} \\ &\quad = \frac{a^{2-k}}{\pi^2 k^2} \sum_{j=1}^{k} \frac{\bar{\zeta}^{2j}}{\big(1 - a\bar{\zeta}^j\big)^2\big(a\bar{\zeta}^j - s\big)^2}, \end{aligned} \tag{3.3}$$

where $a = z_2\bar{w}_2^{1/k}$ and $s = z_1\bar{w}_1$. Define $f_j(a, s) := (\zeta^j - a)^2(a - s\zeta^j)^2$ and notice that $\prod_{j=1}^{k} f_j(a, s) = \prod_{j=1}^{k}(\zeta^j - a)^2 \cdot \prod_{j=1}^{k}(a - s\zeta^j)^2 = (1 - a^k)^2(a^k - s^k)^2$.

Now, it follows that

$$(3.4)\qquad (3.3)=\frac{a^{2-k}}{\pi^2k^2}\sum_{j=1}^{k}\frac{\zeta^{2j}}{f_j(a,s)}=\frac{a^{2-k}\sum_{j=1}^{k}F_j(a,s)\zeta^{2j}}{\pi^2k^2(1-a^k)^2(a^k-s^k)^2},$$

where $F_j(a,s):=(1-a^k)^2(a^k-s^k)^2/f_j(a,s)$. Notice each $F_j(a,s)$ can be written as a polynomial in a of degree $4k-4$, so the numerator of (3.4) takes the form

$$(3.5)\qquad a^{2-k}\sum_{j=1}^{k}F_j(a,s)\zeta^{2j}=\sum_{j=2-k}^{3k-2}g_j(s)a^j:=G(a,s).$$

We now wish to calculate the coefficient polynomials $g_j(s)$. Toward this goal, observe that $G(\zeta^m a,s)=G(a,s)$ for all $m\in\mathbb{Z}$. This follows because

$$\begin{aligned}G(\zeta^m a,s)&=(\zeta^m a)^{2-k}\sum_{j=1}^{k}F_j(\zeta^m a,s)\zeta^{2j}\\&=a^{2-k}\sum_{j=1}^{k}\frac{(1-a^k)^2(a^k-s^k)^2}{f_{j-m}(a,s)}\zeta^{2j-2m}=G(a,s).\end{aligned}$$

Here, we've used the facts that

$$f_j(\zeta^m a,s)=\zeta^{4m}f_{j-m}(a,s)\quad\text{and}\quad f_j(a,s)=f_{j+mk}(a,s)$$

for all $m\in\mathbb{Z}$. Because G has this invariance, we conclude that

$$(3.6)\qquad G(a,s)=a^{2-k}\sum_{j=1}^{k}F_j(a,s)\zeta^{2j}=g_{2k}(s)a^{2k}+g_k(s)a^k+g_0(s).$$

It remains to calculate $g_{2k}(s)$, $g_k(s)$ and $g_0(s)$, and these polynomials are obtained in the following lemma. But to avoid disrupting the flow of the paper with several pages of algebra, we postpone its proof until Section 5.

Lemma 3.7. *The coefficient polynomials* $g_{2k}(s)$, $g_k(s)$ *and* $g_0(s)$ *are given by*

$$(3.8)\qquad g_{2k}(s)=k\sum_{l=1}^{k-1}l(k-l)s^{l-1}=kp_k(s),$$

$$(3.9)\qquad g_k(s)=k\sum_{l=1}^{k}(l^2+(k-l)^2s^k)s^{l-1}=kq_k(s),$$

$$(3.10)\qquad g_0(s)=k\sum_{l=1}^{k-1}l(k-l)s^{k+l-1}=ks^kp_k(s).$$

Using this lemma and letting $t:=a^k=z_2^k\bar{w}_2$, we see from (3.4) that

$$\mathbb{B}_k\big((z_1, z_2^k), (w_1, w_2)\big) = \frac{p_k(s)t^2 + q_k(s)t + s^k p_k(s)}{k\pi^2(1-t)^2(t-s^k)^2}.$$

This is the desired formula for $\mathbb{B}_k$, except that both sides are a function of z_2^k. This is remedied by formally replacing the variable z_2^k with z_2. This concludes the proof of Theorem 3.1 □

Remark 3.11. It's also true that the Bergman kernel of $\mathbb{H}_{m/n}$ is a rational function whenever $m, n \in \mathbb{Z}^+$. Indeed, the map $(z_1, z_2) \mapsto (z_1 z_2^{n-1}, z_2^m)$ is a proper map from $\mathbb{H}_1$ onto $\mathbb{H}_{m/n}$, so Bell's formula gives $\mathbb{B}_{m/n}$ as a finite sum. Zapałowski [2016] characterizes the proper maps between fat Hartogs triangles. He shows there is a proper map $F : \mathbb{H}_{m/n} \to \mathbb{H}_{p/q}$ if and only if there are $a, b \in \mathbb{Z}^+$ such that

$$\frac{aq}{p} - \frac{bn}{m} \in \mathbb{Z}.$$

Zapałowski's description of proper maps shows that the methods employed in this paper aren't able to say anything about fat Hartogs triangles $\mathbb{H}_\gamma$, for irrational γ.

Remark 3.12. Ramadanov's theorem says that if $\{\Omega_k\}$ is an increasing family of domains such that $\Omega_k \to \Omega \Subset \mathbb{C}^n$, then $\mathbb{B}_{\Omega_k}(z, w) \to \mathbb{B}_\Omega(z, w)$ absolutely and uniformly on compact subsets of $\Omega \times \Omega$. See [Ramadanov 1967] for the first appearance of this fact, and [Boas 1996] for a generalization in the smoothly bounded, pseudoconvex case. Notice that $\{\mathbb{H}_k\}$ is an increasing family and that $\mathbb{H}_k \to D \times D^*$ as $k \to \infty$. Ramadanov's theorem shows that $\mathbb{B}_k(z, w) \to \mathbb{B}_{D\times D^*}(z, w)$, which is given in (2.5). This is difficult to see from direct computation.

Biholomorphism classes of domains. Let

$$\psi(z) = (\psi_1(z), \psi_2(z)) := (z_1/z_2, z_2).$$

On page 330, we used the fact that $\psi : \mathbb{H}_1 \to D \times D^*$ is a biholomorphism to compute the Bergman kernel of $\mathbb{H}_1$. We'll give a very similar argument to prove Theorem 1.4. Let $\Psi(z) = (z_1 z_2, z_2)$, and see that $\Psi : D \times D^* \to \mathbb{H}_1$ is the inverse of ψ. Now, notice that $\psi : \mathbb{H}_{1/(k+1)} \to \mathbb{H}_{1/k}$ is also a biholomorphism for all $k \in \mathbb{Z}^+$, because

$$\begin{aligned} |\psi_1(z)|^{1/k} < |\psi_2(z)| < 1 &\iff |\psi_1(z)| < |\psi_2(z)|^k < 1 \\ &\iff |z_1/z_2| < |z_2|^k < 1 \\ &\iff |z_1| < |z_2|^{k+1} < 1 \\ &\iff |z_1|^{1/(k+1)} < |z_2| < 1. \end{aligned}$$

Let $\psi^k := \psi \circ \cdots \circ \psi$ be k copies of ψ composed together, so $\psi^k(z) := (z_1 z_2^{-k}, z_2)$. This gives a biholomorphism from $\mathbb{H}_{1/k}$ to $D \times D^*$ with inverse $\Psi^k := (z_1 z_2^k, z_2)$.

We illustrate this chain of biholomorphisms below:

$$D \times D^* \underset{\psi}{\overset{\Psi}{\rightleftharpoons}} \mathbb{H}_1 \underset{\psi}{\overset{\Psi}{\rightleftharpoons}} \mathbb{H}_{1/2} \underset{\psi}{\overset{\Psi}{\rightleftharpoons}} \cdots \underset{\psi}{\overset{\Psi}{\rightleftharpoons}} \mathbb{H}_{1/k} \underset{\psi}{\overset{\Psi}{\rightleftharpoons}} \cdots .$$

Proof of Theorem 1.4. Using the biholomorphism $\psi^k : \mathbb{H}_{1/k} \to D \times D^*$ we easily obtain the desired formula. Indeed, since $\det[(\psi^k)'](z) = z_2^{-k}$,

$$\mathbb{B}_{1/k}(z, w) = \frac{1}{z_2^k \bar{w}_2^k} \mathbb{B}_1(\psi_k(z), \psi_k(w)) = \frac{z_2^k \bar{w}_2^k}{\pi^2(1 - z_2\bar{w}_2)^2(z_2^k\bar{w}_2^k - z_1\bar{w}_1)^2}. \qquad \square$$

Remark 3.13. For $m, n \in \mathbb{Z}^+$, the map $\psi(z) = (z_1/z_2, z_2)$ also gives a biholomorphism from $\mathbb{H}_{m/(n+m)}$ onto $\mathbb{H}_{m/n}$. Applying this map recursively, we see that $\mathbb{H}_{m/(n+km)}$ and $\mathbb{H}_{m/n}$ are biholomorphic for all $k \in \mathbb{Z}^+$.

4. Consequences of the kernel formulas

The Lu Qi-Keng problem. One of the long-standing open problems in Bergman theory is to classify the domains for which the Bergman kernel is nowhere vanishing. This question was first raised by Lu Qi-Keng [1966]. We say that a domain $\Omega \subset \mathbb{C}^n$ is Lu Qi-Keng when it has zero-free Bergman kernel, and the investigation of which domains have a zero-free Bergman kernel is known as the Lu Qi-Keng problem. See [Boas 2000] for a good historical survey, a few key points of which we now summarize.

The situation in the complex plane is relatively straightforward. When $\Omega \subset \mathbb{C}$ is simply connected, the Riemann mapping theorem together with (2.4) show that Ω is a Lu Qi-Keng domain, since the Bergman kernel of the unit disc is nonvanishing. But a finitely connected domain in $\mathbb{C}$ with at least two nonsingleton boundary components is not Lu Qi-Keng. See [Rosenthal 1969; Skwarczyński 1969] when Ω is an annulus, and [Bell 1992] for a more general class of domains.

There is no such simple characterization of the situation known in higher dimensions. In [Boas et al. 1999], it's shown there are smoothly bounded, strongly convex domains with real analytic boundary that are not Lu Qi-Keng in $\mathbb{C}^n$, when $n \geq 3$. Contrary to previous expectations, Boas [1996] shows that "most" pseudoconvex domains (with respect to a certain topology on the set of domains in $\mathbb{C}^n$) have vanishing Bergman kernel. Nevertheless, it is still desirable to understand why domains from certain classes have zero-free Bergman kernels, while domains from closely related classes may not. We now address this problem in the case of the domains $\mathbb{H}_\gamma$, where $\gamma \in \mathbb{Z}^+$ and $\gamma^{-1} \in \mathbb{Z}^+$.

Using the explicit formulas for the Bergman kernels computed in the previous section, we can check whether or not these domains are Lu Qi-Keng. The following corollary is immediate from (1.5), whose numerator vanishes if and only if at least one of z_2 or w_2 equals zero.

Corollary 4.1. *Let k be a positive integer. The thin Hartogs triangle $\mathbb{H}_{1/k}$ is a Lu Qi-Keng domain.*

For fat Hartogs triangles with integer exponent $k \geq 2$, we deduce the following corollary from (1.3).

Corollary 4.2. *Let $k \geq 2$ be an integer. The fat Hartogs triangle $\mathbb{H}_k$ is not a Lu Qi-Keng domain.*

Proof. First consider the case $k \geq 3$. Let $z = \big(0, i/\sqrt{k-1}\big)$ and $w = \big(0, -i/\sqrt{k-1}\big)$. Then $z, w \in \mathbb{H}_k$. Since $p_k(0) = k-1$ and $q_k(0) = 1$, we see that $\mathbb{B}_k(z, w) = 0$. When $k = 2$, let $z = \big(i/\sqrt{2}, \big(\sqrt{7}+i\big)/4\big)$ and $w = \big(-i/\sqrt{2}, \big(\sqrt{7}-i\big)/4\big)$. It is easily checked that $z, w \in \mathbb{H}_2$ and that $\mathbb{B}_2(z, w) = 0$. □

It's immediate from (2.4) that a nonvanishing Bergman kernel is a biholomorphic invariant. Corollary 4.2 lets us deduce the following:

Corollary 4.3. *Let $k \geq 2$ be an integer. $\mathbb{H}_k$ is not biholomorphic to $D \times D^*$.*

Remark 4.4. Using Ramadanov's theorem in conjunction with Hurwitz's theorem on zeroes of holomorphic functions, we see that for each integer $k \geq 2$, there is an $s_k \in [k-1, k)$ such that for all $\gamma \in (s_k, k]$, the Bergman kernel $\mathbb{B}_\gamma$ of $\mathbb{H}_\gamma$ has zeroes. It seems plausible to conjecture that $s_k = k-1$, i.e., that no fat Hartogs triangle of exponent $\gamma > 1$ is Lu Qi-Keng.

Remark 4.5. As was mentioned in Remark 3.12, $\mathbb{H}_k \to D \times D^*$ as $k \to \infty$. The Bergman kernel $\mathbb{B}_{D\times D^*}$ is zero free, so for any fixed compact subset $K \subset D \times D^*$, Ramadanov's theorem tells us that the Bergman kernel $\mathbb{B}_k$ restricted to K is zero free for all k sufficiently large. We see this happen as the zero of $\mathbb{B}_k$ provided in the proof of Corollary 4.2 is pushed to the origin. It would be interesting to do further analysis of the zero set of the polynomial in the numerator of $\mathbb{B}_k$.

Diagonal boundary behavior. The asymptotic behavior of $\mathbb{B}_\Omega(z, z)$ as z tends to the boundary has been studied for many classes of smoothly bounded, pseudoconvex domains. [Hörmander 1965; Fefferman 1974] are two seminal papers dealing with the strongly pseudoconvex case. Results also exist for many classes of smoothly bounded, weakly pseudoconvex domains. See [McNeal 1989; Catlin 1989; Nagel et al. 1989] for finite-type domains in $\mathbb{C}^2$, and [McNeal 1994] for finite-type, convex domains in $\mathbb{C}^n$. Refer to [Fu 2014] for analogous results on smoothly bounded domains with constant Levi rank. But all these estimates are for classes of domains with boundary smoothness, and there are presently no general theorems about the behavior of $\mathbb{B}_\Omega(z, z)$ for pseudoconvex domains near singular boundary points.

Using the explicit formulas for the Bergman kernel, we establish the following:

Lemma 4.6. *Let $k \in \mathbb{Z}^+$. We have the following behavior of the Bergman kernel restricted to the diagonal*:

$$\mathbb{B}_k(z,z) \approx \frac{1}{(1-|z_2|)^2(|z_2|-|z_1|^k)^2}, \quad z \in \mathbb{H}_k. \tag{4.7}$$

Proof. In this proof we are concerned with $\mathbb{B}_k(z,z)$, so write $s := |z_1|^2$ and $t := |z_2|^2$. From Theorem 1.2 we see that

$$\mathbb{B}_k(z,z) = \frac{p_k(s)t^2 + q_k(s)t + s^k p_k(s)}{k\pi^2(1-t)^2(t-s^k)^2}, \tag{4.8}$$

where $p_k(s)$ and $q_k(s)$ are given in the statement of Theorem 1.2. We now estimate the numerator of (4.8). Notice that $q_k(s) \geq 1$ for all $s \in [0,1)$, and so

$$t \leq p_k(s)t^2 + q_k(s)t + s^k p_k(s) < t[2p_k(1) + q_k(1)] \lesssim t,$$

since $s^k < t$. Now estimate the terms in the denominator. It's easy to see that both

$$(1-t)^2 \approx (1-|z_2|)^2,$$
$$(t-s^k)^2 \approx |z_2|^2(|z_2|-|z_1|^k)^2.$$

Here, we've used the fact that $|z_2|^2 \leq (|z_2|+|z_1|^k)^2 < 4|z_2|^2$. Putting these estimates together, we obtain (4.7). □

Let $\Omega \subset \mathbb{C}^2$ be a bounded domain and $\zeta \in b\Omega$ a smooth, Levi-flat boundary point. It can be shown that $\mathbb{B}_\Omega(z,z) \approx \delta_\Omega(z)^{-2}$ as $z \to \zeta$. See [Fu 2014] for more information. The domains $\mathbb{H}_k$ are Levi-flat at all smooth boundary points, because the smooth parts of the boundary can be locally foliated by analytic discs. We explicitly see this asymptotic behavior from estimate (4.7). In fact, this estimate also lets us determine the asymptotic growth rate of $\mathbb{B}_k(z,z)$ as z tends to the boundary singularity at the origin. When z is sufficiently close to 0, it's straightforward to see $|z_2| - |z_1|^k \approx \delta_k(z)$, the distance of z to the boundary of $\mathbb{H}_k$. From this, we deduce:

Theorem 4.9. *Let $k \in \mathbb{Z}^+$ and $\delta_k(z)$ be the distance of z to $b\mathbb{H}_k$. Then*

$$\mathbb{B}_k(z,z) \approx \delta_k(z)^{-2} \quad \textit{as } z \to 0.$$

Remark 4.10. Following steps analogous to those in Lemma 4.6, we can show

$$\mathbb{B}_{1/k}(z,z) \approx \frac{1}{(1-|z_2|)^2(|z_2|^k-|z_1|)^2}, \quad z \in \mathbb{H}_{1/k}.$$

This estimate can be used to determine the asymptotic growth rate of $\mathbb{B}_{1/k}(z,z)$ as z tends to the boundary singularity at the origin. When z is sufficiently close to 0, it's straightforward to check that $|z_2|^k - |z_1| \approx \delta_{1/k}(z)$, the distance of z to $b\mathbb{H}_{1/k}$. From this we conclude that $\mathbb{B}_{1/k}(z,z) \approx \delta_{1/k}(z)^{-2}$ as $z \to 0$.

5. Proof of Lemma 3.7

Equation (3.6) tells us that

$$a^{2-k}\sum_{j=1}^{k} F_j(a,s)\zeta^{2j} = g_{2k}(s)a^{2k} + g_k(s)a^k + g_0(s). \tag{5.1}$$

We prove Lemma 3.7 by splitting the calculation of $g_{2k}(s)$, $g_k(s)$ and $g_0(s)$ into two separate lemmas.

Lemma 5.2. *Let*

$$h_l(s) := \sum_{r=0}^{l} s^r.$$

For each $j = 1, \ldots, k$, the respective coefficient functions of the a^{3k-2}, a^{2k-2} and a^{k-2} terms of $F_j(a,s)\zeta^{2j}$ are equal to the following:

$$\begin{aligned}
a^{3k-2}: &\quad \sum_{l=0}^{k-2} h_l(s)h_{k-2-l}(s),\\
a^{2k-2}: &\quad 2\sum_{l=0}^{k-2} s^{k-1-l}h_l(s)^2 + h_{k-1}(s)^2,\\
a^{k-2}: &\quad s^k\sum_{l=0}^{k-2} h_l(s)h_{k-2-l}(s).
\end{aligned}$$

In particular, note that these expressions have no j dependence.

Proof. In this calculation of the coefficient functions of the a^{3k-2}, a^{2k-2} and a^{k-2} terms appearing in $F_j(a,s)\zeta^{2j}$, we'll often write $\theta := \zeta^j$ to cut down on superscripts.

$$\begin{aligned}
F_j(a,s) &= \frac{(1-a^k)^2(a^k-s^k)^2}{f_j(a,s)} = \Big(\frac{a^k-1}{a-\theta}\Big)^2\Big(\frac{a^k-s^k}{a-s\theta}\Big)^2 \\
&= \Big(\sum_{m=1}^{k} a^{k-m}\theta^{m-1}\Big)^2\Big(\sum_{n=1}^{k} a^{k-n}(s\theta)^{n-1}\Big)^2 \\
&= \Big(\sum_{m=1}^{k}\sum_{n=1}^{k} a^{2k-m-n}\theta^{m+n-2}s^{n-1}\Big)^2.
\end{aligned} \tag{5.3}$$

To better understand the double sum inside the parentheses of (5.3) above, we split this sum into three pieces, A, B and C, depending on the value of $m+n$. Let A be the sum of the terms with $2 \le m+n \le k$, B be the sum of the terms with $m+n = k+1$, and C be the sum of the terms with $k+2 \le m+n \le 2k$.

We rewrite A by letting $l = m + n - 2$ be the index of summation. Then

$$A = \sum_{l=0}^{k-2} a^{2k-l-2}\theta^l h_l(s).$$

For B, only include those terms with $m + n = k + 1$, so we don't have an outside sum. Therefore,

$$B = a^{k-1}\theta^{k-1} h_{k-1}(s).$$

For C, let $l = m + n - k - 2$ be the index of summation. Then

$$C = \sum_{l=0}^{k-2} a^{k-2-l}\theta^{k+l} s^{l+1} h_{k-2-l}(s).$$

So we have

$$\begin{aligned}(5.3) &= \Bigl(\sum_{l=0}^{k-2} a^{2k-2-l}\theta^l h_l(s) + a^{k-1}\theta^{k-1}h_{k-1}(s) + \sum_{l=0}^{k-2} a^{k-2-l}\theta^{k+l}s^{l+1}h_{k-2-l}(s)\Bigr)^2 \\ &= (A+B+C)^2 \\ &= A^2 + B^2 + C^2 + 2AB + 2BC + 2AC.\end{aligned}$$

I emphasize that as a polynomial in a, A has powers of a ranging from a^{2k-2} to a^k, B only has an a^{k-1} term, and C has terms ranging from a^{k-2} to a^0. This observation greatly simplifies the computations below.

Computation of the a^{3k-2} coefficient. For the coefficient of the a^{3k-2} term in $F_j(a,s)\theta^2$, it is sufficient to consider the coefficient function of a^{3k-2} in $A^2\theta^2$:

$$\begin{aligned}A^2\theta^2 &= \Bigl(\sum_{m=0}^{k-2} a^{2k-2-m}\theta^m h_m(s)\Bigr)\Bigl(\sum_{n=0}^{k-2} a^{2k-2-n}\theta^n h_n(s)\Bigr)\theta^2 \\ &= \theta^2 \sum_{m=0}^{k-2}\sum_{n=0}^{k-2} a^{4k-4-m-n}\theta^{m+n}h_m(s)h_n(s).\end{aligned}$$

Letting $m + n = k - 2$, we find that the coefficient function of a^{3k-2} is independent of θ (since $\theta^k = 1$), and therefore independent of j. This function is given by

$$\sum_{l=0}^{k-2} h_l(s)h_{k-2-l}(s). \tag{5.4}$$

Computation of the a^{2k-2} coefficient. For the coefficient of the a^{2k-2} term in $F_j(a,s)\theta^2$, it is sufficient to consider the coefficient of a^{2k-2} in $(2AC + B^2)\theta^2$:

$$
\begin{aligned}
&(2AC+B^2)\theta^2\\
&=\Bigg[2\bigg(\sum_{m=0}^{k-2}a^{2k-2-m}\theta^m h_m(s)\bigg)\bigg(\sum_{n=0}^{k-2}a^{k-2-n}\theta^{k+n}s^{n+1}h_{k-2-n}(s)\bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + (a^{k-1}\theta^{k-1}h_{k-1}(s))^2\Bigg]\theta^2\\
&=\Bigg[2\sum_{m=0}^{k-2}\sum_{n=0}^{k-2}a^{3k-4-m-n}\theta^{k+m+n}s^{n+1}h_m(s)h_{k-2-n}(s)+a^{2k-2}\theta^{2k-2}h_{k-1}(s)^2\Bigg]\theta^2.
\end{aligned}
$$

Letting $m+n=k-2$, we find that the coefficient function of a^{2k-2} is independent of θ (since $\theta^{2k}=1$), and therefore independent of j. This function is given by

$$
2\sum_{l=0}^{k-2}s^{k-1-l}h_l(s)^2+h_{k-1}(s)^2. \tag{5.5}
$$

Computation of the a^{k-2} coefficient. For the coefficient of the a^{k-2} term, it is sufficient to determine the coefficient of a^{k-2} in $C^2\theta^2$:

$$
\begin{aligned}
C^2\theta^2 &= \bigg(\sum_{m=0}^{k-2}a^{k-2-m}\theta^{k+m}s^{m+1}h_{k-2-m}(s)\bigg)\bigg(\sum_{n=0}^{k-2}a^{k-2-n}\theta^{k+n}s^{n+1}h_{k-2-n}(s)\bigg)\theta^2\\
&=\theta^2\sum_{m=0}^{k-2}\sum_{n=0}^{k-2}a^{2k-4-m-n}\theta^{2k+m+n}s^{m+n+2}h_{k-2-m}(s)h_{k-2-n}(s).
\end{aligned}
$$

Letting $m+n=k-2$, we find that the coefficient function of a^{k-2} is independent of θ (since $\theta^{3k}=1$), and therefore independent of j. This function is given by

$$
s^k\sum_{l=0}^{k-2}h_l(s)h_{k-2-l}(s). \tag{5.6}
$$

□

Now we rewrite (5.4), (5.5) and (5.6) as simpler polynomials:

Lemma 5.7. *Again, let $h_l(s)=\sum_{r=0}^{l}s^r$. Then we have the following equalities*:

$$
\sum_{l=0}^{k-2}h_l(s)h_{k-2-l}(s)=\sum_{l=1}^{k-1}l(k-l)s^{l-1}, \tag{5.8}
$$

$$
2\sum_{l=0}^{k-2}s^{k-1-l}h_l(s)^2+h_{k-1}(s)^2=\sum_{l=1}^{k}(l^2+(k-l)^2s^k)s^{l-1}. \tag{5.9}
$$

Proof. Focus on (5.8) first. Notice that

$$
h_l(s)h_{k-2-l}(s)=\bigg(\sum_{m=0}^{l}s^m\bigg)\bigg(\sum_{n=0}^{k-2-l}s^n\bigg)=\sum_{r=0}^{k-2}s^r+\sum_{r=1}^{k-3}s^r+\cdots+\sum_{r=L}^{k-2-L}s^r,
$$

where $L = \min\{l, k-2-l\}$. Using this, we see

$$
\begin{aligned}
(5.10)\qquad (5.4) &= \sum_{l=0}^{k-2} h_l(s)h_{k-2-l}(s)\\
&= \sum_{l=0}^{k-2}\Bigg(\sum_{r=0}^{k-2} s^r + \sum_{r=1}^{k-3} s^r + \cdots + \sum_{r=L}^{k-2-L} s^r\Bigg)\\
&= (k-1)\sum_{r=0}^{k-2} s^r + (k-3)\sum_{r=1}^{k-3} s^r + \cdots + (k-2K-1)\sum_{r=K}^{k-2-K} s^r,
\end{aligned}
$$

where $K = \lfloor \frac{1}{2}(k-2) \rfloor$. From here, we compute that the coefficient of s^l in (5.4) is given by

$$
\begin{aligned}
\sum_{m=0}^{L}(k-2m-1) &= (L+1)(k-1) - 2\sum_{m=0}^{L} m\\
&= (L+1)(k-L-1) = (l+1)(k-l-1).
\end{aligned}
$$

Therefore,

$$
(5.4) = \sum_{l=0}^{k-2} h_l(s)h_{k-2-l}(s) = \sum_{l=0}^{k-2}(l+1)(k-l-1)s^l = \sum_{l=1}^{k-1} l(k-l)s^{l-1},
$$

where we've re-indexed the sum in the last equality, obtaining the form of (5.8).

Now we'll establish (5.9). Note that

$$
h_r(s)^2 = 1 + 2s + \cdots + rs^{r-1} + (r+1)s^r + rs^{r+1} + \cdots + 2s^{2r-1} + s^{2r}.
$$

Using this, write the pieces of (5.5) $= 2\sum_{l=0}^{k-2} s^{k-1-l}h_l(s)^2 + h_{k-1}(s)^2$ in the following way:

$$
\begin{aligned}
s^{k-1}h_0(s)^2 &= \qquad\qquad\qquad s^{k-1}\\
s^{k-2}h_1(s)^2 &= \qquad\qquad s^{k-2} + 2s^{k-1} + s^k\\
s^{k-3}h_2(s)^2 &= \qquad s^{k-3} + 2s^{k-2} + 3s^{k-1} + 2s^k + s^{k+1}\\
&\ \ \vdots \qquad\qquad\qquad\qquad \vdots\\
h_{k-1}(s)^2 &= 1 + \cdots + (k-1)s^{k-2} + ks^{k-1} + (k-1)s^k + \cdots + s^{2k-2}.
\end{aligned}
$$

The coefficient of s^l in (5.5) can be obtained by considering the vertical columns above. Notice the coefficients of s^l and s^{2k-2-l} are always the same. When $0 \le l \le k-1$, we have that the coefficient of s^l is given by

$$
2\sum_{r=1}^{l} r + (l+1) = (l+1)^2.
$$

Therefore,

$$
\begin{aligned}
(5.5) &= \sum_{l=0}^{k-1}(l+1)^2 s^l + \sum_{l=0}^{k-2}(k-(l+1))^2 s^{k+l} \\
&= \sum_{l=0}^{k-1}\big((l+1)^2 + (k-l-1)^2 s^k\big)s^l \\
&= \sum_{l=1}^{k}\big(l^2 + (k-l)^2 s^k\big)s^{l-1},
\end{aligned}
$$

where we have re-indexed the sum in the last equality to obtain the form of (5.9). □

Proof of Lemma 3.7. Lemmas 5.2 and 5.7 together give us (3.8), (3.9) and (3.10). □

Acknowledgements

I would like to thank my advisor, Jeff McNeal, for introducing me to Bergman theory, and for his ongoing support and encouragement. The many interesting discussions in his office have contributed to my growth, both as a mathematician and a human being. I would also like to thank the anonymous referee who offered many useful suggestions to improve the original draft of this paper.

References

[Bell 1982] S. R. Bell, "The Bergman kernel function and proper holomorphic mappings", *Trans. Amer. Math. Soc.* **270**:2 (1982), 685–691. MR Zbl

[Bell 1992] S. R. Bell, *The Cauchy transform, potential theory, and conformal mapping*, CRC Press, Boca Raton, FL, 1992. MR

[Boas 1996] H. P. Boas, "The Lu Qi-Keng conjecture fails generically", *Proc. Amer. Math. Soc.* **124**:7 (1996), 2021–2027. MR Zbl

[Boas 2000] H. P. Boas, "Lu Qi-Keng's problem", *J. Korean Math. Soc.* **37**:2 (2000), 253–267. MR Zbl

[Boas et al. 1999] H. P. Boas, S. Fu, and E. J. Straube, "The Bergman kernel function: Explicit formulas and zeroes", *Proc. Amer. Math. Soc.* **127**:3 (1999), 805–811. MR Zbl

[Bremermann 1955] H. J. Bremermann, "Holomorphic continuation of the kernel function and the Bergman metric in several complex variables", pp. 349–383 in *Lectures on functions of a complex variable*, edited by W. Kaplan, The University of Michigan Press, Ann Arbor, 1955. MR Zbl

[Catlin 1989] D. W. Catlin, "Estimates of invariant metrics on pseudoconvex domains of dimension two", *Math. Z.* **200**:3 (1989), 429–466. MR Zbl

[Chakrabarti and Shaw 2013] D. Chakrabarti and M.-C. Shaw, "Sobolev regularity of the $\overline{\partial}$-equation on the Hartogs triangle", *Math. Ann.* **356**:1 (2013), 241–258. MR Zbl

[Chakrabarti and Zeytuncu 2016] D. Chakrabarti and Y. E. Zeytuncu, "L^p mapping properties of the Bergman projection on the Hartogs triangle", *Proc. Amer. Math. Soc.* **144**:4 (2016), 1643–1653. MR Zbl

[Chen 2013] L. Chen, "The L^p boundedness of the Bergman projection for a class of bounded Hartogs domains", preprint, 2013. arXiv 1304.7898

[D'Angelo 1978] J. P. D'Angelo, "A note on the Bergman kernel", *Duke Math. J.* **45**:2 (1978), 259–265. MR Zbl

[D'Angelo 1994] J. P. D'Angelo, "An explicit computation of the Bergman kernel function", *J. Geom. Anal.* **4**:1 (1994), 23–34. MR Zbl

[Edholm and McNeal 2016] L. D. Edholm and J. D. McNeal, "The Bergman projection on fat Hartogs triangles: L^p boundedness", *Proc. Amer. Math. Soc.* **144**:5 (2016), 2185–2196. MR Zbl

[Fefferman 1974] C. Fefferman, "The Bergman kernel and biholomorphic mappings of pseudoconvex domains", *Invent. Math.* **26** (1974), 1–65. MR Zbl

[Fu 2014] S. Fu, "Estimates of invariant metrics on pseudoconvex domains near boundaries with constant Levi ranks", *J. Geom. Anal.* **24**:1 (2014), 32–46. MR Zbl

[Hörmander 1965] L. Hörmander, "L^2 estimates and existence theorems for the $\bar\partial$ operator", *Acta Math.* **113** (1965), 89–152. MR Zbl

[Krantz 1992] S. G. Krantz, *Function theory of several complex variables*, 2nd ed., Wadsworth & Brooks/Cole, Pacific Grove, CA, 1992. MR Zbl

[Lu 1966] Q.-K. Lu, "On Kaehler manifolds with constant curvature", *Chinese Math.–Acta* **8** (1966), 283–298. MR Zbl

[McNeal 1989] J. D. McNeal, "Boundary behavior of the Bergman kernel function in $\mathbf{C}^2$", *Duke Math. J.* **58**:2 (1989), 499–512. MR Zbl

[McNeal 1994] J. D. McNeal, "Estimates on the Bergman kernels of convex domains", *Adv. Math.* **109**:1 (1994), 108–139. MR Zbl

[Nagel et al. 1989] A. Nagel, J.-P. Rosay, E. M. Stein, and S. Wainger, "Estimates for the Bergman and Szegő kernels in $\mathbf{C}^2$", *Ann. of Math.* (2) **129**:1 (1989), 113–149. MR Zbl

[Park 2008] J.-D. Park, "New formulas of the Bergman kernels for complex ellipsoids in $\mathbb{C}^2$", *Proc. Amer. Math. Soc.* **136**:12 (2008), 4211–4221. MR Zbl

[Ramadanov 1967] I. Ramadanov, "Sur une propriété de la fonction de Bergman", *C. R. Acad. Bulgare Sci.* **20** (1967), 759–762. MR Zbl

[Rosenthal 1969] P. Rosenthal, "On the zeros of the Bergman function in doubly-connected domains", *Proc. Amer. Math. Soc.* **21** (1969), 33–35. MR Zbl

[Skwarczyński 1969] M. Skwarczyński, "The invariant distance in the theory of pseudoconformal transformations and the Lu Qi-Keng conjecture", *Proc. Amer. Math. Soc.* **22** (1969), 305–310. MR Zbl

[Zapałowski 2016] P. Zapałowski, "Proper holomorphic mappings between generalized Hartogs triangles", preprint, 2016. arXiv 1601.01806

Received June 27, 2015. Revised February 26, 2016.

LUKE D. EDHOLM
DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
COLUMBUS, OH 43201
UNITED STATES
edholm.1@osu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.343

TRANSFERENCE OF CERTAIN MAXIMAL HILBERT TRANSFORMS ON THE TORUS

DASHAN FAN, HUOXIONG WU AND FAYOU ZHAO

Using transference techniques, we show that $L^p(\mathbb{R}^n)$ estimates for many operators may be transferred to the $L^p(\mathbb{T}^n)$ estimates on the n-torus $\mathbb{T}^n$ via measure-preserving actions of $\mathbb{R}^n$. These operators include the maximal bilinear Hilbert transform, the oscillation, and the variation and short variation operators of the Hilbert transform on the torus $\mathbb{T}$. As an extension, we study the (maximal) bilinear Riesz transforms on the n-torus $\mathbb{T}^n$.

1. Introduction

Let $\mathbb{C}$ be the complex plane and $\mathbb{R}^2_+$ the upper half plane

$$\mathbb{R}^2_+ = \{(x, y) = x + iy \in \mathbb{C} : y > 0\}.$$

The boundary of $\mathbb{R}^2_+$ is the real line $\mathbb{R}$. Consider the boundary condition $f \in L^p(\mathbb{R})$, where f is real-valued and $1 \le p < \infty$. It is well known that the Poisson integral

$$u(x, y) = P_y(f)(x) = \frac{1}{\pi} \int_{\mathbb{R}} f(t) \frac{y}{(x-t)^2 + y^2} \, dt$$

is the solution of the Dirichlet problem on $\mathbb{R}^2_+$. Precisely, u is a harmonic function on $\mathbb{R}^2_+$ and $u(x, y)$ tends to $f(x)$ nontangentially for almost all $x \in \mathbb{R}$ as $y \to 0^+$. There is a (unique) harmonic function

$$v(x, y) = Q_y(f)(x) = \frac{1}{\pi} \int_{\mathbb{R}} f(t) \frac{x-t}{(x-t)^2 + y^2} \, dt$$

such that

$$F(z) = u(x, y) + iv(x, y)$$

is an analytic function on $\mathbb{R}^2_+$. This function $Q_y(f)(x)$ is called the conjugate Poisson integral of f. From [Stein and Weiss 1971, p. 186], we know that the

The research was supported by National Natural Science Foundation of China (Grant Nos. 11471288, 11471041, 11371295, 11201287), China Scholarship Council (No. 201406895019) and the FETU. The corresponding author is Zhao.

MSC2010: primary 42B20, 42B25; secondary 42A50, 40A30.

Keywords: bilinear Hilbert transform, maximal operator, Riesz transform, oscillation, variation, torus, transference.

function $F(z)$, with $z = x + iy$, has the nontangential limit $f(x) + \frac{i}{\pi} Hf(x)$ for almost all $x \in \mathbb{R}$. Here H is the Hilbert transform defined by

$$Hf(x) = \lim_{\varepsilon \to 0} H_\varepsilon f(x) \tag{1}$$

and

$$H_\varepsilon f(x) = \int_{|t|>\varepsilon} \frac{f(x-t)}{t}\, dt.$$

To study the pointwise convergence of $Hf(x)$, one then needs to study the truncated maximal Hilbert transform

$$H^* f(x) = \sup_{\varepsilon>0} |H_\varepsilon f(x)|.$$

Let

$$D = \{z = x + iy \in \mathbb{C} : |z| < 1\}$$

be the unit disc. Its boundary $\mathbb{T} = \partial D$ is the one-dimensional torus. Without loss of generality, we may identify the torus $\mathbb{T}$ with its fundamental interval $\left[-\frac{1}{2}, \frac{1}{2}\right)$. The Dirichlet problem on D with the boundary condition $\tilde{f} \in L^p(\mathbb{T})$ similarly raises an analytic function $\widetilde{F}(z) = \tilde{u}(x, y) + i\tilde{v}(x, y)$. The function $\widetilde{F}(z)$, for $z = x + iy \in D$, has the nontangential limit $\tilde{f}(x) + \frac{i}{\pi} \widetilde{H}\tilde{f}(x)$ for almost all points $x \in \mathbb{T}$. Here $\widetilde{H}$ is the periodic version of the Hilbert transform defined by

$$\widetilde{H}\tilde{f}(x) = \lim_{\varepsilon \to 0} \widetilde{H}_\varepsilon \tilde{f}(x)$$

and

$$\widetilde{H}_\varepsilon \tilde{f}(x) = \int_{\varepsilon<|t|<\frac{1}{2}} \tilde{f}(x-t) \cot(\pi t)\, dt. \tag{2}$$

By computing the Fourier coefficients, one can see that $\widetilde{H}\tilde{f}(x)$ has the Fourier series

$$\widetilde{H}\tilde{f}(x) = \sum_{k=-\infty}^{\infty} i \operatorname{sgn}(k) a_k e^{2\pi i k x}$$

for any

$$\tilde{f}(x) = \sum_{k=-\infty}^{\infty} a_k e^{2\pi i k x}$$

(see also [Edwards and Gaudry 1977]). It is known that $\sum_{k=-\infty}^{\infty} \operatorname{sgn}(k) a_k e^{2\pi i k x}$ (up to a constant multiplier) is the conjugate Fourier series of $\tilde{f}$.

The bilinear Hilbert transform $\mathcal{H}(f, g)$ and the maximal bilinear Hilbert transform $\mathcal{H}^*(f, g)$ are defined respectively as

$$\mathcal{H}(f, g)(x) = \lim_{\varepsilon \to 0} \int_{|t|>\varepsilon} \frac{f(x-t) g(x+t)}{t}\, dt \tag{3}$$

and

$$\mathcal{H}^*(f, g)(x) = \sup_{\varepsilon>0} \left| \int_{|t|>\varepsilon} \frac{f(x-t)g(x+t)}{t}\, dt \right|.$$

The operator $\mathcal{H}(f, g)$ is not merely a formal extension from the Hilbert transform. It has deep roots in the study of certain harmonic analysis and PDE problems. The study of the bilinear Hilbert transform $\mathcal{H}(f, g)$ was initiated by Calderón when he studied certain Cauchy integrals $C_\gamma(f)$ along the Lipschitz curves. In order to obtain the L^2 boundedness of $C_\gamma(f)$, Calderón introduced a commutator (now known as the first Calderón commutator) and raised a famous conjecture, which says that $\mathcal{H}$ is a bounded operator from $L^\infty \times L^2 \to L^2$; see [Jones 1994]. The conjecture was solved in a more general setting by Lacey and Thiele in their celebrated theorem:

Theorem A1 [Lacey and Thiele 1997]. *Let* $1 < q, r \le \infty$, *and* $\frac{2}{3} < p < \infty$. *Then*

$$\|\mathcal{H}(f, g)\|_{L^p(\mathbb{R})} \preceq \|f\|_{L^q(\mathbb{R})} \|g\|_{L^r(\mathbb{R})},$$

provided $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$.

The notation $A \preceq B$ for $A, B > 0$ means that there exists a constant $c > 0$ independent of all essential variables such that $A \le cB$. We also use the notation $A \simeq B$ when $A \preceq B$ and $B \preceq A$.

The proof of the theorem by Lacey and Thiele involves a very elegant method of time-frequency analysis. The essence of the matter lies in their formulation and proof of certain almost orthogonal results on the phase space. Maximal forms of these results must be proved. These maximal inequalities rely in an essential way on a novel maximal inequality of Bourgain [1989; 1990]. By refining these maximal bilinear estimates and Bourgain's lemma, Lacey further obtained the following remarkable theorem:

Theorem A2 [Lacey 2000]. *Let* $1 < q, r \le \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *If* $\frac{2}{3} < p < \infty$, *then*

$$\|\mathcal{H}^*(f, g)\|_{L^p(\mathbb{R})} \preceq \|f\|_{L^q(\mathbb{R})} \|g\|_{L^r(\mathbb{R})}.$$

Based on Theorems A1 and A2, it is natural to expect to establish analogous theorems for the periodic bilinear Hilbert transform on the torus. Here, the bilinear Hilbert transform and its maximal operator on the torus are defined, initially on $C^\infty(\mathbb{T})$, by

$$\widetilde{\mathcal{H}}(\tilde{f}, \tilde{g})(x) = \text{p.v.} \int_{|t|<\frac{1}{2}} \tilde{f}(x-t)\tilde{g}(x+t) \cot(\pi t)\, dt$$

and

$$\widetilde{\mathcal{H}}^*(\tilde{f}, \tilde{g})(x) = \sup_{\varepsilon>0} \left| \int_{\varepsilon<|t|<\frac{1}{2}} \tilde{f}(x-t)\tilde{g}(x+t) \cot(\pi t)\, dt \right|.$$

However, it seems quite difficult to adopt the time-frequency method used in [Lacey and Thiele 1997]. Thus, in [Fan and Sato 2001], the authors used a "transference" method to reduce the boundedness of $\widetilde{\mathcal{H}}(\tilde{f}, \tilde{g})$ to the boundedness of $\mathcal{H}(f, g)$ by estimating an error term. The method of transference is a useful tool for obtaining norm estimates independent of the dimension for classical operators acting on $L^p(\mathbb{R}^n)$ (see [Auscher and Carro 1994; Blasco and Gillespie 2009; Coifman and Weiss 1977; Gillespie and Torrea 2004; Rubio de Francia 1989]). Fan and Sato [2001] proved the de Leeuw-type theorems (see [de Leeuw 1965]) for the transference of multilinear operators on Lebesgue spaces from $\mathbb{R}^n$ to the n-torus. In particular, they proved:

Theorem B. *Let* $1 < q, r \le \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *If* $\frac{2}{3} < p < \infty$, *then*

$$\|\widetilde{\mathcal{H}}(\tilde{f}, \tilde{g})\|_{L^p(\mathbb{T})} \preceq \|\tilde{f}\|_{L^q(\mathbb{T})} \|\tilde{g}\|_{L^r(\mathbb{T})}.$$

Note that

$$\widetilde{\mathcal{H}}(\tilde{f}, \tilde{g})(x) \simeq \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \operatorname{sgn}(k_1 - k_2) a_{k_1} a_{k_2} e^{2\pi i (k_1 + k_2) x},$$

where

$$\tilde{f}(x) = \sum_{k_1 \in \mathbb{Z}} a_{k_1} e^{2\pi i k_1 x} \quad \textit{and} \quad \tilde{g}(x) = \sum_{k_2 \in \mathbb{Z}} a_{k_2} e^{2\pi i k_2 x}.$$

In [Fan and Sato 2001], the authors also studied the boundedness of the maximal multiplier operator

$$\widetilde{T}^{**}(\tilde{f}, \tilde{g})(x) = \sup_{\varepsilon > 0} \left| \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} m(\varepsilon k_1, \varepsilon k_2) a_{k_1} a_{k_2} e^{2\pi i (k_1 + k_2) x} \right|,$$

where m is a bounded and continuous function (see also [Berkson et al. 2006; 2007; Blasco et al. 2005; Grafakos and Honzík 2006] for transference methods on maximal bilinear operators). For the bilinear Hilbert transform, clearly we have

$$\widetilde{\mathcal{H}}(\tilde{f}, \tilde{g})(x) = \widetilde{\mathcal{H}}^{**}(\tilde{f}, \tilde{g})(x),$$

since

$$\sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \operatorname{sgn}(\varepsilon k_1 - \varepsilon k_2) a_{k_1} a_{k_2} e^{2\pi i (k_1 + k_2) x} = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \operatorname{sgn}(k_1 - k_2) a_{k_1} a_{k_2} e^{2\pi i (k_1 + k_2) x}.$$

This observation indicates

$$\widetilde{\mathcal{H}}^{*}(\tilde{f}, \tilde{g})(x) \neq \widetilde{\mathcal{H}}^{**}(\tilde{f}, \tilde{g})(x).$$

Since the boundedness of $\widetilde{\mathcal{H}}^*(\tilde{f}, \tilde{g})$ still remains open, the first aim of this paper is to solve this problem by establishing the following analog of Lacey's theorem.

Theorem 1.1. *Let* $1 < q, r \le \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *If* $\frac{2}{3} < p < \infty$, *then*

$$\|\widetilde{\mathcal{H}}^*(\tilde{f}, \tilde{g})\|_{L^p(\mathbb{T})} \preceq \|\tilde{f}\|_{L^q(\mathbb{T})} \|\tilde{g}\|_{L^r(\mathbb{T})}.$$

We adopt the method in [Fan and Sato 2001] to prove the theorem, in which the main issue is to estimate error terms in order to reduce the boundedness of $L^q(\mathbb{T}) \times L^r(\mathbb{T}) \to L^p(\mathbb{T})$ to the known result for $L^q(\mathbb{R}) \times L^r(\mathbb{R}) \to L^p(\mathbb{R})$. This method additionally allows us to treat other operators related to the maximal Hilbert transform. Recall that the limits (1) and (3) mentioned above exist almost everywhere. Motivated by probability and ergodic theory [Bourgain 1989; Jones 1997; 1998], in order to obtain extra information on their convergence rate, as well as an estimate on the number of λ-jumps they can have, Campbell, Jones, Reinhold and Wierdl [Campbell et al. 2000] studied the oscillation and variation of the family (H_ε) as ε approaches 0 as follows.

For each fixed sequence $(t_k) \searrow 0$, define the oscillation and variation operators by

$$\mathscr{O}(H_* f)(x) = \left(\sum_{k=1}^{\infty} \sup_{t_{k+1} \le \varepsilon_{k+1} < \varepsilon_k \le t_k} |H_{\varepsilon_k} f(x) - H_{\varepsilon_{k+1}} f(x)|^2 \right)^{\frac{1}{2}}, \tag{4}$$

$$\mathscr{V}_\varrho(H_* f)(x) = \sup_{(\varepsilon_k) \searrow 0} \left(\sum_{k=1}^{\infty} |H_{\varepsilon_k} f(x) - H_{\varepsilon_{k+1}} f(x)|^\varrho \right)^{\frac{1}{\varrho}}, \tag{5}$$

respectively. Also, define

$$V_k(H_* f)(x) = \sup_{(\varepsilon_j) \searrow 0} \left(\sum_{\frac{1}{2^k} < \varepsilon_{j+1} < \varepsilon_j \le \frac{1}{2^{k-1}}} |H_{\varepsilon_j} f(x) - H_{\varepsilon_{j+1}} f(x)|^2 \right)^{\frac{1}{2}},$$

where the supremum is taken over all decreasing sequences (ε_j). Then the "short variation operator" is defined by

$$S_V(H_* f)(x) = \left(\sum_{k=-\infty}^{\infty} V_k(H_* f(x))^2 \right)^{\frac{1}{2}}. \tag{6}$$

For convenience, all the integrals are defined on the Schwartz class.

We recall the following results from [Campbell et al. 2000].

Theorem C. *The oscillation operator* $\mathscr{O}(H_*)$ *satisfies*

$$\|\mathscr{O}(H_* f)\|_{L^p(\mathbb{R})} \le c_p \|f\|_{L^p(\mathbb{R})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{R} : \mathscr{O}(H_* f)(x) > \lambda\}| \le (c/\lambda)\|f\|_{L^1(\mathbb{R})}$.

Theorem D. *If* $\varrho > 2$, *then the variation operator* $\mathscr{V}_\varrho(H_*)$ *satisfies*

$$\|\mathscr{V}_\varrho(H_* f)\|_{L^p(\mathbb{R})} \le c(p, \varrho)\|f\|_{L^p(\mathbb{R})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{R} : \mathscr{V}_\varrho(H_* f)(x) > \lambda\}| \le (c(\varrho)/\lambda)\|f\|_{L^1(\mathbb{R})}$.

Theorem E. *The short variation operator* $S_V(H_*)$ *satisfies*

$$\|S_V(H_* f)\|_{L^p(\mathbb{R})} \le c_p \|f\|_{L^p(\mathbb{R})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{R} : S_V(H_* f)(x) > \lambda\}| \le (c/\lambda)\|f\|_{L^1(\mathbb{R})}$.

Our second aim is to use the transference method to study the analogs of these operators on the torus. For each fixed sequence $(t_k) \searrow 0$, define the oscillation and variation operators on the torus by

$$\mathscr{O}(\widetilde{H}_* \tilde{f})(x) = \Big(\sum_{k=1}^{\infty} \sup_{t_{k+1} \le \varepsilon_{k+1} < \varepsilon_k \le t_k} |\widetilde{H}_{\varepsilon_k} \tilde{f}(x) - \widetilde{H}_{\varepsilon_{k+1}} \tilde{f}(x)|^2 \Big)^{\frac{1}{2}},$$

$$\mathscr{V}_\varrho(\widetilde{H}_* \tilde{f})(x) = \sup_{(\varepsilon_k) \searrow 0} \Big(\sum_{k=1}^{\infty} |\widetilde{H}_{\varepsilon_k} \tilde{f}(x) - \widetilde{H}_{\varepsilon_{k+1}} \tilde{f}(x)|^\varrho \Big)^{\frac{1}{\varrho}},$$

respectively. Also, define the operator $V_k(\widetilde{H}_*)$ on the torus by

$$V_k(\widetilde{H}_* \tilde{f})(x) = \sup_{(\varepsilon_j) \searrow 0} \Big(\sum_{\frac{1}{2^k} < \varepsilon_{j+1} < \varepsilon_j \le \frac{1}{2^{k-1}}} |\widetilde{H}_{\varepsilon_j} \tilde{f}(x) - \widetilde{H}_{\varepsilon_{j+1}} \tilde{f}(x)|^2 \Big)^{\frac{1}{2}},$$

where the supremum is taken over all decreasing sequences (ε_j). Then define the "short variation operator" on the torus by

$$S_V(\widetilde{H}_* \tilde{f})(x) = \Big(\sum_{k=-\infty}^{\infty} (V_k(\widetilde{H}_* \tilde{f}(x))^2 \Big)^{\frac{1}{2}}.$$

For simplicity, we define these operators on the space $C^\infty(\mathbb{T})$.

We establish the following theorems.

Theorem 1.2. *The oscillation operator* $\mathscr{O}(\widetilde{H}_*)$ *satisfies*

$$\|\mathscr{O}(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})} \le c_p \|\tilde{f}\|_{L^p(\mathbb{T})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{T} : \mathscr{O}(\widetilde{H}_* \tilde{f})(x) > \lambda\}| \le (c/\lambda)\|\tilde{f}\|_{L^1(\mathbb{T})}$.

Theorem 1.3. *If* $\varrho > 2$, *then the variation operator* $\mathscr{V}_\varrho(\widetilde{H}_*)$ *satisfies*

$$\|\mathscr{V}_\varrho(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})} \le c(p, \varrho) \|\tilde{f}\|_{L^p(\mathbb{T})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{T} : \mathscr{V}_\varrho(\widetilde{H}_* \tilde{f})(x) > \lambda\}| \le (c(\varrho)/\lambda)\|\tilde{f}\|_{L^1(\mathbb{T})}$.

Theorem 1.4. *The short variation operator* $S_V(\widetilde{H}_*)$ *satisfies*

$$\|S_V(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})} \le c_p \|\tilde{f}\|_{L^p(\mathbb{T})}$$

for $1 < p < \infty$ *and* $|\{x \in \mathbb{T} : S_V(\widetilde{H}_* \tilde{f})(x) > \lambda\}| \le (c/\lambda)\|\tilde{f}\|_{L^1(\mathbb{T})}$.

As we mentioned, the method in [Fan and Sato 2001] shows that $L^p(\mathbb{R}^n)$ estimates for many linear operators may be transferred to their corresponding L^p estimates on the torus via measure-preserving actions of $\mathbb{R}^n$. As a further application and extension, we consider the bilinear Riesz transform on the n-torus. Recall that the bilinear singular integral with rough kernel on $\mathbb{R}^n$ is defined by

$$T_\Omega(f,g)(x)=\lim_{\varepsilon\to 0}\int_{|y|>\varepsilon} f(x-y)g(x+y)\frac{\Omega(y')}{|y|^n}\,dy, \tag{7}$$

where $\Omega(y')$ is a function defined on the unit sphere S^{n-1} in Euclidean space $\mathbb{R}^n$ and whose integral over S^{n-1} is zero. One then obtains the bilinear Riesz transform by taking $\Omega(x)=x_j/|x|$, where x_j is the j-th component of x. Using the same transference method, we also can transfer the L^p-boundedness of the maximal bilinear Riesz transform from $\mathbb{R}^n$ to $\mathbb{T}^n$. This fact is discussed in the last two sections.

Throughout this article, we use the letter c to denote a positive constant, which is independent of the main parameters and not necessarily the same at each occurrence.

2. Proof of Theorems 1.1–1.4

In this section we give the proof of Theorems 1.1–1.4. As is well known, Euler discovered two remarkable expressions for circular functions, one as an infinite product and the other as an infinite series. For the sine function he established the formula

$$\sin(\pi x)=\pi x\prod_{k=1}^{\infty}\Bigl(1-\frac{x^2}{k}\Bigr)=\pi x\prod_{k\in\mathbb{Z}\setminus\{0\}}\Bigl(1+\frac{x}{k}\Bigr)$$

(see [Varadarajan 2007]). By logarithmic differentiation one obtains

$$\cot(\pi x)=\text{p.v. }\frac{1}{\pi}\sum_{k\in\mathbb{Z}}\Bigl(\frac{1}{x+k}\Bigr), \tag{8}$$

where p.v. means the Cauchy principal value, that is, that the sum has to be interpreted as the limit

$$\lim_{N\to+\infty}\sum_{k=-N}^{N}\frac{1}{x+k}=\frac{1}{x}+\sum_{k=1}^{\infty}\Bigl(\frac{1}{x+k}-\frac{1}{x-k}\Bigr),\qquad N\in\mathbb{Z}^+. \tag{9}$$

Let $\chi_A(t)$ be the characteristic function of the set $A=\{t\in\mathbb{R}:|t|>1\}$. To prove Theorem 1.1, we need the following lemma.

Lemma 2.1. *For $\varepsilon<|t|\le\frac{1}{2}$, we have*

$$\cot(\pi t)\chi_A\Bigl(\frac{t}{\varepsilon}\Bigr)=\frac{1}{\pi}\sum_{k\in\mathbb{Z}}\frac{1}{t+k}\chi_A\Bigl(\frac{t+k}{\varepsilon}\Bigr)-\frac{1}{\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Bigl(\frac{t}{\varepsilon}\Bigr)$$

and the estimate

$$\Big|\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big)\Big|\preceq|t|\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big),$$

where A^C is the complement of the set A.

Proof. Using (8), we write the term $\cot(\pi t)\chi_A(t/\varepsilon)$ as

$$\begin{aligned}\cot(\pi t)\chi_A\Big(\frac{t}{\varepsilon}\Big)&=\frac{1}{\pi}\sum_{k\in\mathbb{Z}}\frac{1}{t+k}\chi_A\Big(\frac{t}{\varepsilon}\Big)\\&=\frac{1}{\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_A\Big(\frac{t}{\varepsilon}\Big)+\frac{1}{\pi t}\chi_A\Big(\frac{t}{\varepsilon}\Big).\end{aligned}$$

Since χ_A is the characteristic function of the set $\{t\in\mathbb{R}:|t|>1\}$, it is easy to see that for $\varepsilon<|t|\le\frac{1}{2}$ and $k\in\mathbb{Z}\setminus\{0\}$, we have

$$\chi_A\Big(\frac{t+k}{\varepsilon}\Big)=1.$$

The fact above leads to

$$\frac{1}{\pi t}\chi_A\Big(\frac{t}{\varepsilon}\Big)=\frac{1}{\pi t}\chi_A\Big(\frac{t}{\varepsilon}\Big)+\frac{1}{\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_A\Big(\frac{t+k}{\varepsilon}\Big)-\frac{1}{\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}.$$

Hence we have

$$\cot(\pi t)\chi_A\Big(\frac{t}{\varepsilon}\Big)=\frac{1}{\pi}\sum_{k\in\mathbb{Z}}\frac{1}{t+k}\chi_A\Big(\frac{t+k}{\varepsilon}\Big)-\frac{1}{\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big).$$

It now remains to estimate

$$\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big).$$

From (9), we know

$$\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big)=\sum_{k=1}^{\infty}\Big(\frac{1}{t+k}+\frac{1}{t-k}\Big)\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big)=2t\sum_{k=1}^{\infty}\frac{1}{t^2-k^2}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big).$$

Using this yields

$$\Big|\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big)\Big|\preceq|t|\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big),$$

which completes the proof. $\square$

Proof of Theorem 1.1. For simplicity, we introduce some notation. Denote by

$$\mathfrak{R}(t) = \frac{\chi_A(t)}{t}, \qquad \mathfrak{R}_\varepsilon(t) = \frac{1}{\varepsilon}\mathfrak{R}\Big(\frac{t}{\varepsilon}\Big), \tag{10}$$

$$\widetilde{\mathfrak{R}}_\varepsilon(t) = \frac{1}{\varepsilon}\sum_{k\in\mathbb{Z}}\mathfrak{R}\Big(\frac{t+k}{\varepsilon}\Big), \quad r_\varepsilon(t) = \sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{t+k}\chi_{A^C}\Big(\frac{t}{\varepsilon}\Big). \tag{11}$$

Then we have

$$\mathcal{H}_\varepsilon(f,g)(x) = \int_{\mathbb{R}} f(x-t)g(x+t)\mathfrak{R}_\varepsilon(t)\,dt.$$

Let

$$\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)(x) = \int_{|t|<\frac{1}{2}} \tilde f(x-t)\tilde g(x+t)\widetilde{\mathfrak{R}}_\varepsilon(t)\,dt.$$

Because of Lemma 2.1, one has

$$\widetilde{H}^*(\tilde f,\tilde g)(x) \le \tfrac{1}{\pi}\sup_{\varepsilon>0}\big|\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)(x)\big| + \tfrac{1}{\pi}M(\tilde f,\tilde g)(x),$$

where

$$M(\tilde f,\tilde g)(x) = \sup_{\varepsilon>0}\bigg|\int_{|t|<\frac{1}{2}}\tilde f(x-t)\tilde g(x+t)r_\varepsilon(t)\,dt\bigg|.$$

By Lemma 2.1, the Minkowski integral inequality and Hölder's inequality, we see that, for $p \ge 1$,

$$\begin{aligned}\|M(\tilde f,\tilde g)\|_{L^p(\mathbb{T})} &\preceq \int_{-\frac{1}{2}}^{\frac{1}{2}} \|\tilde f(x-t)\tilde g(x+t)\|_{L^p(\mathbb{T},dx)}|t|\,dt \\ &\preceq \|\tilde f\|_{L^q(\mathbb{T})}\|\tilde g\|_{L^r(\mathbb{T})}.\end{aligned}$$

On the other hand, using Hölder's inequality, we have

$$\begin{aligned}\bigg\|\int_{-\frac{1}{2}}^{\frac{1}{2}}\tilde f(x-t)\tilde g(x+t)r_\varepsilon(t)\,dt\bigg\|_{L^{1/2}(\mathbb{T})} &\preceq \bigg\|\int_{-\frac{1}{2}}^{\frac{1}{2}}|\tilde f(x-t)\tilde g(x+t)r_\varepsilon(t)|\,dt\bigg\|_{L^1(\mathbb{T})} \\ &\preceq \|\tilde f\|_{L^1(\mathbb{T})}\|\tilde g\|_{L^1(\mathbb{T})}.\end{aligned}$$

Then an interpolation yields that, for all $\frac{1}{2} \le p < \infty$,

$$\|M(\tilde f,\tilde g)\|_{L^p(\mathbb{T})} \preceq \|\tilde f\|_{L^q(\mathbb{T})}\|\tilde g\|_{L^r(\mathbb{T})}.$$

Thus, to prove the theorem, we only need to show that

$$\Big\|\sup_{\varepsilon>0}\big|\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)\big|\Big\|_{L^p(\mathbb{T})} \preceq \|\tilde f\|_{L^q(\mathbb{T})}\|\tilde g\|_{L^r(\mathbb{T})}.$$

It is easy to compute that the Fourier coefficients of $\widetilde{\mathfrak{R}}_\varepsilon(t)$ are

$$
\begin{aligned}
c_{m,\varepsilon} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{k\in\mathbb{Z}} \frac{1}{\varepsilon}\mathfrak{R}\Big(\frac{t+k}{\varepsilon}\Big)e^{-i2\pi mt}\,dt \\
&= \int_{\mathbb{R}} \mathfrak{R}(t)e^{-i2\pi mt\varepsilon}\,dt = -i\int_{|t|>1}\frac{\sin(2\pi mt\varepsilon)}{t}\,dt \\
&= -\pi i\,\operatorname{sgn}(\varepsilon m) + i\int_{|t|\le 2\pi}\frac{\sin(\varepsilon mt)}{t}\,dt.
\end{aligned}
\tag{12}
$$

Clearly, $c_{m,\varepsilon}$ is uniformly bounded on $\varepsilon > 0$ and $m \in \mathbb{Z}$. On the other hand, it is not difficult to check that

$$
\int_{|t|<\frac{1}{2}} \tilde{f}(x-t)\tilde{g}(x+t)\widetilde{\mathfrak{R}}_\varepsilon(t)\,dt = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} a_m b_n c_{m-n,\varepsilon} e^{2\pi i(n+m)x},
$$

where

$$
\tilde{f}(x) = \sum_{m=-\infty}^{\infty} a_m e^{2\pi imx} \quad \text{and} \quad \tilde{g}(x) = \sum_{n=-\infty}^{\infty} b_n e^{2\pi inx}.
$$

Now pick $\Psi \in \mathscr{S}(\mathbb{R})$ satisfying $\Psi(x) = 1$ on $\big[-\frac{1}{2}, \frac{1}{2}\big]$, $\operatorname{supp}(\Psi) \subset \big[-\frac{3}{4}, \frac{3}{4}\big]$ and $0 \le \Psi(x) \le 1$. For any positive N, denote the function Ψ^N by

$$
\Psi^N(x) = \Psi(x/N). \tag{13}
$$

Consider the error term given by

$$
E_{N,\varepsilon}(\tilde{f}, \tilde{g})(x) = \Psi(x/N)^2\widetilde{\mathbb{H}}_\varepsilon(\tilde{f}, \tilde{g})(x) - \mathcal{H}_\varepsilon\big(\Psi^N\tilde{f}, \Psi^N\tilde{g}\big)(x).
$$

The error term $E_{N,\varepsilon}(\tilde{f}, \tilde{g})(x)$ roughly gives the difference of $\mathcal{H}_\varepsilon$ on $\mathbb{R}$ and $\widetilde{\mathbb{H}}_\varepsilon$ on the torus. By checking the Fourier transform, we have

$$
\begin{aligned}
\mathcal{H}_\varepsilon\big(\Psi^N\tilde{f}, \Psi^N\tilde{g}\big)(x) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} & a_m b_n e^{2\pi i(n+m)x} \\
&\times \int_{\mathbb{R}} \mathfrak{R}_\varepsilon(t)\Psi\Big(\frac{x+t}{N}\Big)\Psi\Big(\frac{x-t}{N}\Big)e^{2\pi i(n-m)t}\,dt.
\end{aligned}
$$

The definition of the inverse Fourier transform on the space of Schwartz functions shows that

$$
\begin{aligned}
&\int_{\mathbb{R}} \mathfrak{R}_\varepsilon(t)\Psi\Big(\frac{x+t}{N}\Big)\Psi\Big(\frac{x-t}{N}\Big)e^{2\pi i(n-m)t}\,dt \\
&\quad = \int_{\mathbb{R}} \mathfrak{R}_\varepsilon(t)\int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{\Psi}(u)\widehat{\Psi}(v)e^{2\pi iu(x+t)/N}e^{2\pi iv(x-t)/N}\,du\,dv\; e^{2\pi i(n-m)t}\,dt \\
&\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{\Psi}(u)\widehat{\Psi}(v)e^{2\pi ix(u+v)/N}\Big(\int_{\mathbb{R}} \mathfrak{R}_\varepsilon(t)e^{2\pi i(u-v)t/N}e^{2\pi i(n-m)t}\,dt\Big)\,du\,dv.
\end{aligned}
$$

Therefore, we obtain that

$$
\begin{aligned}
E_{N,\varepsilon}(\tilde{f},\tilde{g})(x) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} a_m b_n e^{2\pi i(n+m)x} \\
\times \int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{\Psi}(u)\widehat{\Psi}(v)(c_{m-n,\varepsilon} - c_{m-n+(v-u)/N,\varepsilon})e^{2\pi i x(u+v)/N}\,du\,dv.
\end{aligned}
$$

If $m = n$, then

$$
\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{\Psi}(u)\widehat{\Psi}(v)(c_{m-n,\varepsilon} - c_{m-n+(v-u)/N,\varepsilon})e^{2\pi i x(u+v)/N}\,du\,dv \\
&\quad = i\int_{\mathbb{R}}\int_{u>v} \widehat{\Psi}(u)\widehat{\Psi}(v)\biggl(\int_{|t|>1}\frac{\sin \varepsilon 2\pi t(v-u)/N}{t}\,dt\biggr)e^{2\pi i x(u+v)/N}\,du\,dv \\
&\qquad - i\int_{\mathbb{R}}\int_{v>u} \widehat{\Psi}(u)\widehat{\Psi}(v)\biggl(\int_{|t|>1}\frac{\sin \varepsilon 2\pi t(u-v)/N}{t}\,dt\biggr)e^{2\pi i x(u+v)/N}\,du\,dv \\
&\quad = 0.
\end{aligned}
$$

If $m \neq n$, for any sufficiently small $\delta > 0$, we choose an $L > 0$ such that

$$
\biggl|\int_{u^2+v^2>L} \widehat{\Psi}(u)\widehat{\Psi}(v)(c_{m-n,\varepsilon} - c_{m-n+(v-u)/N,\varepsilon})e^{2\pi i x(u+v)/N}\,du\,dv\biggr| < \delta.
$$

Now we let N be sufficiently large so that, for $u^2 + v^2 \leq L$,

$$
\operatorname{sgn}(m-n) = \operatorname{sgn}(m-n+(v-u)/N).
$$

By this choice, for all $0 < \varepsilon < \frac{1}{2}$, we have

$$
\begin{aligned}
&\biggl|\int_{u^2+v^2\leq L} \widehat{\Psi}(u)\widehat{\Psi}(v)(c_{m-n,\varepsilon} - c_{m-n+(v-u)/N,\varepsilon})e^{2\pi i x(u+v)/N}\,du\,dv\biggr| \\
&\preceq \int_{u^2+v^2\leq L}\biggl|\widehat{\Psi}(u)\widehat{\Psi}(v)\int_{|t|<2\pi}\frac{\sin(\varepsilon t(m-n)) - \sin(\varepsilon t(m-n+(v-u)/N))}{t}\,dt\biggr|\,du\,dv \\
&= o(1) \quad \text{as } N \to \infty.
\end{aligned}
$$

Since (a_n) and (b_m) are rapidly decreasing sequences, it is easy to see that

$$
\lim_{N\to\infty}\sup_{0<\varepsilon<1/2}|E_{N,\varepsilon}(\tilde{f},\tilde{g})(x)| = 0.
$$

Applying that

$$
\sup_{0<\varepsilon<1/2}\bigl|\widetilde{\mathbb{H}}_{\varepsilon}(\tilde{f},\tilde{g})(x)\bigr|
$$

is a periodic function, together with Theorem A2, we now have, as $N \to \infty$,

$$\begin{aligned}
\Big\|\sup_{0<\varepsilon<1/2}\big|\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)\big|\Big\|_{L^p(\mathbb{T})} &= \bigg(\int_{-\frac12}^{\frac12}\Big(\sup_{0<\varepsilon<1/2}\big|\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)(x)\big|\Big)^p\,dx\bigg)^{\frac1p}\\
&= \frac{1}{N^{1/p}}\bigg(\int_{-\frac N2}^{\frac N2}\sup_{0<\varepsilon<1/2}\Psi(x/N)^2\big|\widetilde{\mathbb{H}}_\varepsilon(\tilde f,\tilde g)(x)\big|^p\,dx\bigg)^{\frac1p}\\
&\preceq o(1)+\frac{1}{N^{1/p}}\bigg(\int_{\mathbb{R}}\sup_{0<\varepsilon<1/2}\big|\mathcal{H}_\varepsilon\big(\Psi^N\tilde f,\Psi^N\tilde g\big)(x)\big|^p\,dx\bigg)^{\frac1p}\\
&\preceq o(1)+\frac{1}{N^{1/p}}\|\Psi^N\tilde f\|_{L^q(\mathbb{R})}\|\Psi^N\tilde g\|_{L^r(\mathbb{R})}\\
&\preceq o(1)+\|\tilde f\|_{L^q(\mathbb{T})}\|\tilde g\|_{L^r(\mathbb{T})}.
\end{aligned}$$

Thus we get the desired result by letting $N\to\infty$. □

We again will use the transference method to prove Theorems 1.2–1.4. To this end, we need a key lemma in order to estimate error terms.

Lemma 2.2. *Let* $\mathfrak{R}_\varepsilon$, $\widetilde{\mathfrak{R}}_\varepsilon$ *and* Ψ *be as defined in* (10), (11) *and* (13), *respectively. For* $k\in\mathbb{Z}^+$, *set*

$$\widetilde{H}_{\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)=\big(\widetilde{\mathfrak{R}}_{\varepsilon_k}-\widetilde{\mathfrak{R}}_{\varepsilon_{k+1}}\big)*\tilde f(x),$$
$$H_{\varepsilon_k,\varepsilon_{k+1}}(f)(x)=\big(\mathfrak{R}_{\varepsilon_k}-\mathfrak{R}_{\varepsilon_{k+1}}\big)*f(x).$$

For fixed $N\in\mathbb{Z}^+$, *define the error term*

$$E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)=\Psi^N(x)\widetilde{H}_{\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)-H_{\varepsilon_k,\varepsilon_{k+1}}(\Psi^N\tilde f)(x).$$

Let $1\le p<\infty$. *As* $N\to\infty$, *we have*:

(i) *For each fixed sequence* $(t_k)\searrow 0$,

$$\bigg\|\bigg(\sum_{k=1}^{\infty}\sup_{t_{k+1}\le\varepsilon_{k+1}<\varepsilon_k\le t_k}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^2\bigg)^{\frac12}\bigg\|_{L^p(\mathbb{T})}=o(1).$$

(ii) *If* $\frac{1}{2^k}\le\varepsilon_{k+1}<\varepsilon_k\le\frac{1}{2^{k-1}}$, *then for* $\varrho>2$,

$$\bigg\|\sup_{(\varepsilon_k)\searrow 0}\bigg(\sum_{k=1}^{\infty}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^{\varrho}\bigg)^{\frac1\varrho}\bigg\|_{L^p(\mathbb{T})}=o(1).$$

Proof. By the previous calculation (12), we know that the Fourier coefficients of $\widetilde{\mathfrak{R}}_\varepsilon(t)$ are

$$c_{l,\varepsilon}=\int_{\mathbb{R}}\mathfrak{R}(t)e^{-i\varepsilon 2\pi\ell t}\,dt=-i\int_{|t|>1}\frac{\sin(\varepsilon 2\pi\ell t)}{t}\,dt,\quad \ell\in\mathbb{Z}.\tag{14}$$

Note that

$$\int_{|t|<\frac{1}{2}} \widetilde{\mathfrak{R}}_{\varepsilon_k}(t)\tilde{f}(x-t)\,dt = \sum_{m=-\infty}^{\infty} a_m c_{m,\varepsilon_k} e^{2\pi i m x}.$$

For fixed $N \in \mathbb{Z}^+$, we have

$$\Psi(x/N)\widetilde{H}_{\varepsilon_k,\varepsilon_{k+1}}(\tilde{f})(x) = \sum_{m=-\infty}^{\infty} a_m e^{2\pi i m x} \int_{\mathbb{R}} \widehat{\Psi}(u)(c_{m,\varepsilon_k} - c_{m,\varepsilon_{k+1}}) e^{2\pi u x/N}\,du.$$

Also, by a similar estimate as in Theorem 1.1, we obtain

$$H_{\varepsilon_k,\varepsilon_{k+1}}(\Psi^N\tilde{f})(x) = \sum_{m=-\infty}^{\infty} a_m e^{2\pi i m x} \int_{\mathbb{R}} \widehat{\Psi}(u)(c_{m+u/N,\varepsilon_k} - c_{m+u/N,\varepsilon_{k+1}}) e^{2\pi u x/N}\,du.$$

Consequently,

$$E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde{f})(x) = \sum_{m=-\infty}^{\infty} a_m e^{2\pi i m x} \times \int_{\mathbb{R}} \widehat{\Psi}(u)((c_{m,\varepsilon_k} - c_{m,\varepsilon_{k+1}}) - (c_{m+u/N,\varepsilon_k} - c_{m+u/N,\varepsilon_{k+1}})) e^{2\pi u x/N}\,du.$$

In order to simplify the notation, we denote by $C_N(u)$ the term

$$(c_{m,\varepsilon_k} - c_{m,\varepsilon_{k+1}}) - (c_{m+u/N,\varepsilon_k} - c_{m+u/N,\varepsilon_{k+1}}).$$

To evaluate the inner integral above, we first deal with the term $C_N(u)$. From the second expression of (14),

$$C_N(u) = 2i\int_{\varepsilon_{k+1}}^{\varepsilon_k} \frac{\sin(2\pi m t) - \sin(2\pi(m+u/N)t)}{t}\,dt.$$

We consider two cases: $m = 0$ and $m \neq 0$.

If $m = 0$, one has

$$|C_N(u)| = \left| -2i\int_{\varepsilon_{k+1}}^{\varepsilon_k} \frac{\sin(2\pi t u/N)}{t}\,dt \right| \leq 4\pi|u|\frac{1}{N}(\varepsilon_k - \varepsilon_{k+1}). \tag{15}$$

If $m \neq 0$, it follows from trigonometric identities that

$$|C_N(u)| \leq 2\int_{\varepsilon_{k+1}}^{\varepsilon_k} \frac{|\sin(\pi t u/N)|}{t}\,dt \leq 2\pi|u|\frac{1}{N}(\varepsilon_k - \varepsilon_{k+1}). \tag{16}$$

Applying the above estimates, we have

$$\Big\|\Big(\sum_{k=1}^{\infty}\sup_{t_{k+1}\le\varepsilon_{k+1}<\varepsilon_k\le t_k}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^2\Big)^{\frac12}\Big\|_{L^p(\mathbb{T})}$$
$$\preceq \frac{1}{N}\Big\|\Big(\sum_{k=1}^{\infty}(t_k-t_{k+1})^2\Big)^{\frac12}\Big\|_{L^p(\mathbb{T})}\preceq\frac{1}{N}.$$

Therefore we obtain that, as $N\to\infty$,

$$\Big\|\Big(\sum_{k=1}^{\infty}\sup_{t_{k+1}\le\varepsilon_{k+1}<\varepsilon_k\le t_k}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^2\Big)^{\frac12}\Big\|_{L^p(\mathbb{T})}=o(1).$$

Similarly, we have for $\varrho>2$,

$$\Big\|\sup_{(\varepsilon_k)\searrow 0}\Big(\sum_{k=1}^{\infty}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^{\varrho}\Big)^{\frac1\varrho}\Big\|_{L^p(\mathbb{T})}$$
$$\preceq\frac1N\Big\|\sup_{(\varepsilon_k)\searrow 0}\Big(\sum_{k=1}^{\infty}(\varepsilon_k-\varepsilon_{k+1})^{\varrho}\Big)^{\frac1\varrho}\Big\|_{L^p(\mathbb{T})}\preceq\frac1N\Big\|\sup_{(\varepsilon_k)\searrow 0}\Big(\sum_{k=1}^{\infty}(\varepsilon_k-\varepsilon_{k+1})\Big)\Big\|_{L^p(\mathbb{T})}\preceq\frac1N.$$

Thus we obtain that, as $N\to\infty$,

$$\Big\|\sup_{(\varepsilon_k)\searrow 0}\Big(\sum_{k=1}^{\infty}\big|E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde f)(x)\big|^{\varrho}\Big)^{\frac1\varrho}\Big\|_{L^p(\mathbb{T})}=o(1). \qquad \square$$

Proof of Theorem 1.2. With the previous notation, by the definition of $\widetilde{H}_\varepsilon(\tilde f)$ in (2), we rewrite $\widetilde{H}_{\varepsilon_k}\tilde f-\widetilde{H}_{\varepsilon_{k+1}}\tilde f$ as

$$\widetilde{H}_{\varepsilon_k}\tilde f(x)-\widetilde{H}_{\varepsilon_{k+1}}\tilde f(x)=\frac1\pi\int_{|t|<\frac12}\tilde f(x-t)\big(\widetilde{\mathfrak{R}}_{\varepsilon_k}(t)-\widetilde{\mathfrak{R}}_{\varepsilon_{k+1}}(t)\big)\,dt$$
$$+\frac1\pi\int_{|t|<\frac12}\tilde f(x-t)(r_{\varepsilon_{k+1}}(t)-r_{\varepsilon_k}(t))\,dt,$$

with the help of Lemma 2.1.

Recall again that χ_A is the characteristic function of the set $|t|>1$. It is easy to see that for $\varepsilon_{k+1}<\varepsilon_k<|t|\le\frac12$, we have that for all $j\in\mathbb{Z}\setminus\{0\}$,

$$\chi_A\Big(\frac{t+j}{\varepsilon_{k+1}}\Big)=\chi_A\Big(\frac{t+j}{\varepsilon_k}\Big)=1.$$

This leads to

$$r_{\varepsilon_{k+1}}(t)-r_{\varepsilon_k}(t)=0.$$

It now suffices to consider

$$\mathbb{O}(\widetilde{H}_* \tilde{f})(x) = \left(\sum_{k=1}^{\infty} \sup_{t_{k+1} \le \varepsilon_{k+1} < \varepsilon_k \le t_k} \left| \int_{|t| < \frac{1}{2}} \tilde{f}(x-t) \big(\widetilde{\mathfrak{R}}_{\varepsilon_k}(t) - \widetilde{\mathfrak{R}}_{\varepsilon_{k+1}}(t) \big)\, dt \right|^2 \right)^{\frac{1}{2}}.$$

By (i) of Lemma 2.2, the basic properties of operators on the torus and Theorem C, we conclude that as $N \to \infty$,

$$\begin{aligned} \|\mathbb{O}(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})}^p &= \frac{1}{N} \int_{-\frac{N}{2}}^{\frac{N}{2}} \left(\sum_{k=1}^{\infty} \sup_{t_{k+1} \le \varepsilon_{k+1} < \varepsilon_k \le t_k} \left| \Psi\Big(\frac{x}{N}\Big) \widetilde{H}_{\varepsilon_k, \varepsilon_k+1}(\tilde{f})(x) \right|^2 \right)^{\frac{p}{2}} dx \\ &\preceq o(1) + \frac{1}{N} \int_{-\frac{N}{2}}^{\frac{N}{2}} \left(\sum_{k=1}^{\infty} \sup_{t_{k+1} \le \varepsilon_{k+1} < \varepsilon_k \le t_k} \left| H_{\varepsilon_i, \varepsilon_i+1}(\Psi^N \tilde{f})(x) \right|^2 \right)^{\frac{p}{2}} dx \\ &\preceq o(1) + \frac{1}{N} \int_{\mathbb{R}} \left| \Psi\Big(\frac{x}{N}\Big) \tilde{f}(x) \right|^p dx \preceq \|\tilde{f}\|_{L^p(\mathbb{T})} + o(1). \end{aligned}$$

We next show that the oscillation operator $\mathscr{O}(\widetilde{H}_*)$ is of weak type (1, 1), that is, for any $\lambda > 0$,

$$\big|\{x \in \mathbb{T} : |\mathscr{O}(\widetilde{H}_* \tilde{f})(x)| > \lambda\}\big| \le \frac{c}{\lambda} \|\tilde{f}\|_{L^1(\mathbb{T})}.$$

By the basic properties of operators on the torus, we find that for $N \in \mathbb{Z}^+$,

$$\begin{aligned} \big|\big\{x \in \big[-\tfrac{1}{2}, \tfrac{1}{2}\big) : |\mathscr{O}(\widetilde{H}_* \tilde{f})(x)| > \lambda\big\}\big| &= N^{-1} \big|\big\{x \in \big[-\tfrac{N}{2}, \tfrac{N}{2}\big) : |\mathscr{O}(\widetilde{H}_* \tilde{f})(x)| > \lambda\big\}\big| \\ &= N^{-1} \big|\big\{|x| \le \tfrac{N}{2} : \big|\Psi\big(\tfrac{x}{N}\big) \mathscr{O}(\widetilde{H}_* \tilde{f})(x)\big| > \lambda\big\}\big|. \end{aligned}$$

As in the proofs of (15) and (16), we know that $E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde{f})(x) \to 0$ uniformly in x as $N \to \infty$. For any λ_1 such that $0 < \lambda_1 < \lambda$, choose N large enough that

$$\big|\big\{x \in \big[-\tfrac{1}{2}, \tfrac{1}{2}\big) : |\mathscr{O}(\widetilde{H}_* \tilde{f})(x)| > \lambda\big\}\big| \le N^{-1} \big|\{x \in \mathbb{R} : |\mathscr{O}(H_*(\Psi^N \tilde{f}))(x)| > \lambda - \lambda_1\}\big|.$$

Theorem C implies that the last term above can be controlled by

$$\frac{cN^{-1}}{\lambda - \lambda_1} \|\Psi^N \tilde{f}\|_{L^1(\mathbb{R})} = \frac{cN^{-1}N}{\lambda - \lambda_1} \|\tilde{f}\|_{L^1(\mathbb{T})} = \frac{c}{\lambda - \lambda_1} \|\tilde{f}\|_{L^1(\mathbb{T})}.$$

Since $\lambda_1 > 0$ is arbitrary, we get the desired result. This completes the proof of Theorem 1.2. □

Proof of Theorem 1.3. Using the same argument as in Theorem 1.2, it is enough to study

$$\mathbb{V}_\varrho(\widetilde{H}_* \tilde{f})(x) = \sup_{(\varepsilon_k) \searrow 0} \left(\sum_{k=1}^{\infty} \left| \int_{|t| < \frac{1}{2}} \tilde{f}(x-t) \big(\widetilde{\mathfrak{R}}_{\varepsilon_k}(t) - \widetilde{\mathfrak{R}}_{\varepsilon_{k+1}}(t) \big)\, dt \right|^\varrho \right)^{\frac{1}{\varrho}}.$$

Now by checking the proof for the oscillation operator $\mathscr{O}(\widetilde{H}_*)$, it suffices to show

$$\|\mathbb{V}_\varrho(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})} \preceq \|\tilde{f}\|_{L^p(\mathbb{T})}.$$

Write

$$H_{\varepsilon_k,\varepsilon_k+1}(f)(x) = H_{\varepsilon_k} f(x) - H_{\varepsilon_{k+1}} f(x),$$
$$\widetilde{H}_{\varepsilon_k,\varepsilon_k+1}(\tilde{f})(x) = \big(\widetilde{\mathfrak{R}}_{\varepsilon_k} - \widetilde{\mathfrak{R}}_{\varepsilon_{k+1}}\big) * \tilde{f}(x).$$

For any large integer N, we define the error term

$$E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde{f})(x) = \Psi(x/N)\widetilde{H}_{\varepsilon_k,\varepsilon_k+1}(\tilde{f}) - H_{\varepsilon_k,\varepsilon_k+1}(\Psi^N \tilde{f})(x).$$

Using (ii) of Lemma 2.2, we obtain that as $N \to \infty$,

$$\Big\| \sup_{(\varepsilon_k)\searrow 0} \Big(\sum_{k=1}^{\infty} \big| E_{N,\varepsilon_k,\varepsilon_{k+1}}(\tilde{f})(x) \big|^{\varrho} \Big)^{\frac{1}{\varrho}} \Big\|_{L^p(\mathbb{T})} = o(1).$$

Finally, applying Theorem D, analogously to the proof of Theorem 1.2 we obtain

$$\begin{aligned} \|\mathbb{V}_{\varrho}(\widetilde{H}_* \tilde{f})\|_{L^p(\mathbb{T})}^p &= \frac{1}{N} \int_{-\frac{N}{2}}^{\frac{N}{2}} \sup_{(\varepsilon_k)\searrow 0} \Big(\sum_{k=1}^{\infty} \Big| \Psi\Big(\frac{x}{N}\Big) \widetilde{H}_{\varepsilon_k,\varepsilon_k+1}(\tilde{f})(x) \Big|^{\varrho} \Big)^{\frac{p}{\varrho}} dx \\ &\preceq o(1) + \frac{1}{N} \int_{\mathbb{R}} \sup_{(\varepsilon_k)\searrow 0} \Big(\sum_{k=1}^{\infty} \big| H_{\varepsilon_k,\varepsilon_k+1}(\Psi^N \tilde{f})(x) \big|^{\varrho} \Big)^{\frac{p}{\varrho}} dx \\ &\preceq o(1) + \frac{1}{N} \int_{\mathbb{R}} \Big| \Psi\Big(\frac{x}{N}\Big) \tilde{f}(x) \Big|^p dx \preceq \|\tilde{f}\|_{L^p(\mathbb{T})} + o(1). \end{aligned}$$

Letting $N \to \infty$, we conclude that the variation operator $\mathscr{V}_{\varrho}(\widetilde{H}_*)$ is of strong type (p, p) for $1 < p < \infty$.

The same argument as in the proof of Theorem 1.2 works for the weak type (1, 1) for the variation operator $\mathscr{V}_{\varrho}(\widetilde{H}_*)$. We omit the details. □

Proof of Theorem 1.4. The proof of Theorem 1.4 is similar to that of Theorem 1.3. The only change is to consider two different cases: $p' > 2$ and $p' \leq 2$ in place of the symmetric differentiation operator used above. We leave the details to the interested reader. □

3. Extension to Riesz transforms

In this section we study the (maximal) bilinear Riesz transforms as n-dimensional extensions.

We start with the maximal bilinear singular integral with rough kernel

$$T_{\Omega}^*(f, g)(x) = \sup_{\varepsilon>0} |T_{\Omega,\varepsilon}(f, g)(x)|, \tag{17}$$

where $T_{\Omega,\varepsilon}$ is the truncated bilinear operator defined by

$$T_{\Omega,\varepsilon}(f, g)(x) = \int_{|y|>\varepsilon} \frac{f(x-y)g(x+y)}{|y|^n} \Omega(y')\, dy \quad \text{for } \varepsilon > 0.$$

Following the standard rotation method by Calderón and Zygmund (see also [Fan and Zhao $\geq$ 2016; Grafakos and Torres 2002]), we have the following result on $\mathbb{R}^n$.

Theorem 3.1. *Let* $1 < q, r \leq \infty$, $1 \leq p < \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *If* $\Omega \in L^\infty(S^{n-1})$ *is an odd function, then*

$$\|T^*_\Omega(f, g)\|_{L^p(\mathbb{R}^n)} \preceq \|f\|_{L^q(\mathbb{R}^n)} \|g\|_{L^r(\mathbb{R}^n)}.$$

If $\Omega(x) = x_j/|x|$, $j = 1, 2, \ldots, n$, then (7) and (17) are reduced to the bilinear Riesz transforms and their maximal operators in Euclidean space $\mathbb{R}^n$:

$$R_j(f, g)(x) = C_n \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} f(x-y) g(x+y) \frac{y_j}{|y|^{n+1}}\, dy,$$
$$R^*_j(f, g)(x) = C_n \sup_{\varepsilon>0} \left| \int_{|y|>\varepsilon} f(x-y) g(x+y) \frac{y_j}{|y|^{n+1}}\, dy \right|, \quad 1 \leq j \leq n,$$

where y_j is the j-th component of y and $C_n = \Gamma((n+1)/2)\pi^{-(n+1)/2}$.

Corollary 3.2. *Let* $1 < q, r \leq \infty$, $1 \leq p < \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *Then*

$$\|R^*_j(f, g)\|_{L^p(\mathbb{R}^n)} \preceq \|f\|_{L^q(\mathbb{R}^n)} \|g\|_{L^r(\mathbb{R}^n)}.$$

As an application, we consider analogous operators on the n-dimensional torus $\mathbb{T}^n = \left[-\frac{1}{2}, \frac{1}{2}\right)^n$.

For $C^\infty(\mathbb{T}^n)$ functions $\tilde{f}, \tilde{g}$, write their Fourier series

$$\tilde{f}(x) = \sum_{k_1 \in \mathbb{Z}^n} a_{k_1} e^{2\pi i \langle k_1, x\rangle}, \quad \tilde{g}(x) = \sum_{k_2 \in \mathbb{Z}^n} b_{k_2} e^{2\pi i \langle k_2, x\rangle},$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product.

Let

$$Q = \left\{x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : -\tfrac{1}{2} \leq x_j < \tfrac{1}{2} \text{ for } j = 1, 2, \ldots, n\right\}$$

be the fundamental cube on which

$$\int_{\mathbb{T}^n} \tilde{f}(x)\, dx = \int_Q \tilde{f}(x)\, dx$$

for all functions $\tilde{f}$ on the torus $\mathbb{T}^n$. For $N \in \mathbb{Z}^+$, let NQ denote a cube with the same center as Q and side length N times the side length of Q. Denote by Q_ε the set given by

$$Q_\varepsilon = \{x \in Q : |x| > \varepsilon\} \quad \text{for } 0 < \varepsilon < \tfrac{1}{2}.$$

Let

$$E = \{x \in \mathbb{R}^n : |x| > 1\}$$

and $\chi_E(x)$ be the characteristic function of E. For $1 \le i \le n$, let x_i and m_i be the i-th components of $x = (x_1, \dots, x_n)$ and $m = (m_1, \dots, m_n)$, respectively. For any $x \ne 0$, the kernel of the j-th Riesz transform on $\mathbb{R}^n$ is

$$K_j(x) = \frac{x_j}{|x|^{n+1}}.$$

Then the kernel of the j-th Riesz transform on the torus is defined, in the sense of Cauchy principle value, by

$$\widetilde{K}_j(x) = \sum_{m \in \mathbb{Z}^n} \frac{x_j + m_j}{|x+m|^{n+1}}. \tag{18}$$

We now define the bilinear Riesz transform $\widetilde{R}_j$ and its maximal operator $\widetilde{R}_j^*$ on the torus $\mathbb{T}^n$, for $\tilde{f}, \tilde{g} \in C^\infty(\mathbb{T}^n)$, by

$$\begin{aligned}\widetilde{R}_j(\tilde{f}, \tilde{g})(x) &= \lim_{\varepsilon \to 0} \widetilde{R}_{j,\varepsilon}(\tilde{f}, \tilde{g})(x),\\ \widetilde{R}_j^*(\tilde{f}, \tilde{g})(x) &= \sup_{0<\varepsilon<1/2} |\widetilde{R}_{j,\varepsilon}(\tilde{f}, \tilde{g})(x)|,\end{aligned}$$

where $\widetilde{R}_{j,\varepsilon}$ is defined by

$$\begin{aligned}\widetilde{R}_{j,\varepsilon}(\tilde{f}, \tilde{g})(x) &= \int_{Q_\varepsilon} \widetilde{K}_j(y) \tilde{f}(x-y)\tilde{g}(x+y)\, dy\\ &= \int_Q \widetilde{K}_j(y) \chi_E\Big(\frac{y}{\varepsilon}\Big) \tilde{f}(x-y)\tilde{g}(x+y)\, dy.\end{aligned}$$

Our result can be stated as follows:

Theorem 3.3. *Let $1 < q, r \le \infty$, $1 \le p < \infty$ and $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. Then*

$$\|\widetilde{R}_j^*(\tilde{f}, \tilde{g})\|_{L^p(\mathbb{T}^n)} \preceq \|\tilde{f}\|_{L^q(\mathbb{T}^n)} \|\tilde{g}\|_{L^r(\mathbb{T}^n)}.$$

By checking the proof of Theorem 1.1, it suffices to show an easy lemma to obtain Theorem 3.3.

Lemma 3.4. *For $0 < \varepsilon < \frac{1}{2}$ and $y \in Q$, we have the estimate*

$$\widetilde{K}_j(y)\chi_E\Big(\frac{y}{\varepsilon}\Big) = \frac{1}{\varepsilon^n} \sum_{m \in \mathbb{Z}^n} K_j\Big(\frac{y+m}{\varepsilon}\Big)\chi_E\Big(\frac{y+m}{\varepsilon}\Big) - \sum_{m \in \mathbb{Z}^n \setminus \{0\}} K_j(y+m)\chi_{E^C}\Big(\frac{y}{\varepsilon}\Big)$$

and

$$\Big| \sum_{m \in \mathbb{Z}^n \setminus \{0\}} K_j(y+m)\chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \Big| \preceq |y| \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big),$$

where E^C is the complement of the set E.

Proof. The first equality above follows the method of Lemma 2.1. Now we estimate the second inequality. Write

$$D_j^+ = \{m \in \mathbb{Z}^n \setminus \{0\} : m_j > 0\}, \quad D_j^0 = \{m \in \mathbb{Z}^n \setminus \{0\} : m_j = 0\},$$

and

$$y^* = (y_1, y_2, \ldots, y_{j-1}, -y_j, y_{j+1}, \ldots, y_n).$$

Then we have

$$\begin{aligned}
&\sum_{m\in\mathbb{Z}^n\setminus\{0\}} \frac{y_j+m_j}{|y+m|^{n+1}} \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \\
&\quad = \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \sum_{m\in D_j^+} \Big(\frac{y_j+m_j}{|y+m|^{n+1}} + \frac{y_j-m_j}{|y^*+m|^{n+1}}\Big) + \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \sum_{m\in D_j^0} \frac{y_j}{|y+m|^{n+1}} \\
&\quad = \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) y_j \sum_{m\in D_j^+} \Big(\frac{1}{|y+m|^{n+1}} + \frac{1}{|y^*+m|^{n+1}}\Big) \\
&\qquad + \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \sum_{m\in D_j^+} m_j \Big(\frac{1}{|y+m|^{n+1}} - \frac{1}{|y^*+m|^{n+1}}\Big) \\
&\qquad\quad + \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) y_j \sum_{m\in D_j^0} \Big(\frac{1}{|y+m|^{n+1}}\Big).
\end{aligned}$$

It is trivial to get that

$$\Big|\chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \sum_{m\in D_j^0} \frac{y_j}{|y+m|^{n+1}}\Big| \preceq \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big)|y_j|$$

and

$$\Big|\chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) y_j \sum_{m\in D_j^+} \Big(\frac{1}{|y+m|^{n+1}} + \frac{1}{|y^*+m|^{n+1}}\Big)\Big| \preceq \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big)|y_j|.$$

Using the mean value theorem,

$$|f(x) - f(y)| \le \max_{z\in I} |\nabla f(z)|\,|x-y|,$$

where I is the line segment between x and y. This leads to

$$\Big|\chi_{E^C}\Big(\frac{y}{\varepsilon}\Big) \sum_{m\in D_j^+} m_j \Big(\frac{1}{|y+m|^{n+1}} - \frac{1}{|y^*+m|^{n+1}}\Big)\Big| \preceq \chi_{E^C}\Big(\frac{y}{\varepsilon}\Big)|y_j|,$$

completing the proof. □

From the theorem, we obtain the following corollary.

Corollary 3.5. *Let* $1 < q, r \leq \infty$, $1 \leq p < \infty$ *and* $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. *Then*

$$\|\widetilde{R}_j(\tilde{f}, \tilde{g})\|_{L^p(\mathbb{T}^n)} \preceq \|\tilde{f}\|_{L^q(\mathbb{T}^n)} \|\tilde{g}\|_{L^r(\mathbb{T}^n)}.$$

This corollary corresponds to a result by Blasco and Gillespie [2009, Theorem 1.12] which says that the bilinear Riesz transform R_j is bounded from $L^q(\mathbb{R}^n) \times L^r(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$, provided $1 < q, r \leq \infty$, $1 \leq p < \infty$ and $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$.

4. Final remarks

We want to further illustrate that our method works for many operators. In this section, we provide another example. For $\varepsilon > 0$, define

$$R_{j,\varepsilon}(f)(x) = C_n \int_{|y|>\varepsilon} f(x-y) \frac{y_j}{|y|^{n+1}} \, dy \quad \text{for } j = 1, 2, \ldots, n.$$

Gillespie and Torrea [2004] introduced the oscillation, variation and short variation operators of the Riesz transform R_j in $\mathbb{R}^n$. The definitions of these three operators can be expressed in forms similar to (4), (5) and (6) with H_ε replaced by $R_{j,\varepsilon}$ in place of the symmetric differentiation operator used above. Gillespie and Torrea also established the $L^p(\mathbb{R}^n)$-boundedness of these operators for $1 < p < \infty$.

For $C^\infty(\mathbb{T}^n)$ functions $\tilde{f}$, write their Fourier series

$$\tilde{f}(x) = \sum_{k \in \mathbb{Z}^n} a_k e^{2\pi i \langle k, x \rangle}.$$

We define the periodic version of $\widetilde{R}_{j,\varepsilon}$ by

$$\widetilde{R}_{j,\varepsilon}(\tilde{f})(x) = \int_{Q_\varepsilon} \widetilde{K}_j(y) \tilde{f}(x-y) \, dy,$$

where $\widetilde{K}_j$ is defined as in (18).

We now define the oscillation operator $\mathscr{O}(\widetilde{R}_j)$ on the torus by

$$\mathscr{O}(\widetilde{R}_j \tilde{f})(x) = \left(\sum_{k=1}^{\infty} \sup_{t_{k+1} \leq \varepsilon_{k+1} < \varepsilon_k \leq t_k} \left| \widetilde{R}_{j,\varepsilon_k} \tilde{f}(x) - \widetilde{R}_{j,\varepsilon_{k+1}} \tilde{f}(x) \right|^2 \right)^{\frac{1}{2}}$$

and the variation operator $\mathscr{V}_\varrho(\widetilde{R}_j)$ on the torus by

$$\mathscr{V}_\varrho(\widetilde{R}_j \tilde{f})(x) = \sup_{(\varepsilon_k) \searrow 0} \left(\sum_{k=1}^{\infty} \left| \widetilde{R}_{j,\varepsilon_k} \tilde{f}(x) - \widetilde{R}_{j,\varepsilon_{k+1}} \tilde{f}(x) \right|^\varrho \right)^{\frac{1}{\varrho}}.$$

Define the operator $V_k(\widetilde{R}_j)$ on the torus by

$$V_k(\widetilde{R}_j \tilde{f})(x) = \sup_{(\varepsilon_j) \searrow 0} \left(\sum_{\frac{1}{2^k} < \varepsilon_{j+1} < \varepsilon_j \leq \frac{1}{2^{k-1}}} \left| \widetilde{R}_{j,\varepsilon_k} \tilde{f}(x) - \widetilde{R}_{j,\varepsilon_{k+1}} \tilde{f}(x) \right|^2 \right)^{\frac{1}{2}},$$

where the supremum is taken over all decreasing sequences (ε_j). Define the "short variation operator" on the torus by

$$S_V(\widetilde{R}_j \tilde{f})(x) = \left(\sum_{k=-\infty}^{\infty} \left(V_k(\widetilde{R}_j \tilde{f}(x)) \right)^2 \right)^{\frac{1}{2}}.$$

Applying the same techniques as in the proof of Theorem 1.2, we can easily transfer those results in [Gillespie and Torrea 2004] from $\mathbb{R}^n$ to the torus $\mathbb{T}^n$.

Acknowledgements

The authors are indebted to the anonymous referees for valuable comments that improved the paper significantly.

References

[Auscher and Carro 1994] P. Auscher and M. J. Carro, "Transference for radial multipliers and dimension free estimates", *Trans. Amer. Math. Soc.* **342**:2 (1994), 575–593. MR 1152319 Zbl 0812.42009

[Berkson et al. 2006] E. Berkson, O. Blasco, M. J. Carro, and T. A. Gillespie, "Discretization and transference of bisublinear maximal operators", *J. Fourier Anal. Appl.* **12**:4 (2006), 447–481. MR 2256933 Zbl 1130.42016

[Berkson et al. 2007] E. Berkson, O. Blasco, M. J. Carro, and T. A. Gillespie, "Discretization versus transference for bilinear operators", pp. 265–283 in *Banach spaces and their applications in analysis* (Oxford, OH, 2006), edited by B. Randrianantoanina and N. Randrianantoanina, Walter de Gruyter, Berlin, 2007. MR 2374713 Zbl 1132.42307

[Blasco and Gillespie 2009] O. Blasco and T. A. Gillespie, "Dimension free estimates for the bilinear Riesz transform", *Collect. Math.* **60**:3 (2009), 249–259. MR 2560969 Zbl 1187.42006

[Blasco et al. 2005] O. Blasco, M. J. Carro, and T. A. Gillespie, "Bilinear Hilbert transform on measure spaces", *J. Fourier Anal. Appl.* **11**:4 (2005), 459–470. MR 2169476 Zbl 1098.42007

[Bourgain 1989] J. Bourgain, "Pointwise ergodic theorems for arithmetic sets", *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 5–45. With an appendix by the author, H. Furstenberg, Y. Katznelson and D. S. Ornstein. MR 1019960 Zbl 0705.28008

[Bourgain 1990] J. Bourgain, "Double recurrence and almost sure convergence", *J. Reine Angew. Math.* **404** (1990), 140–161. MR 1037434 Zbl 0685.28008

[Campbell et al. 2000] J. T. Campbell, R. L. Jones, K. Reinhold, and M. Wierdl, "Oscillation and variation for the Hilbert transform", *Duke Math. J.* **105**:1 (2000), 59–83. MR 1788042 Zbl 1013.42008

[Coifman and Weiss 1977] R. R. Coifman and G. Weiss, *Transference methods in analysis* (Lincoln, NE, 1976), CBMS Regional Conference Series in Mathematics **31**, American Mathematical Society, Providence, RI, 1977. MR 0481928 Zbl 0377.43001

[Edwards and Gaudry 1977] R. E. Edwards and G. I. Gaudry, *Littlewood–Paley and multiplier theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete **90**, Springer, Berlin, 1977. MR 0618663 Zbl 0467.42001

[Fan and Sato 2001] D. Fan and S. Sato, "Transference on certain multilinear multiplier operators", *J. Aust. Math. Soc.* **70**:1 (2001), 37–55. MR 1808390 Zbl 0984.42006

[Fan and Zhao $\geq$ 2016] D. Fan and F. Zhao, "Certain bilinear operators on Morrey spaces", preprint. To appear in *Hokkaido Math. J.*

[Rubio de Francia 1989] J. L. Rubio de Francia, "Transference principles for radial multipliers", *Duke Math. J.* **58**:1 (1989), 1–19. MR 1016410 Zbl 0676.42009

[Gillespie and Torrea 2004] T. A. Gillespie and J. L. Torrea, "Dimension free estimates for the oscillation of Riesz transforms", *Israel J. Math.* **141** (2004), 125–144. MR 2063029 Zbl 1072.42011

[Grafakos and Honzík 2006] L. Grafakos and P. Honzík, "Maximal transference and summability of multilinear Fourier series", *J. Aust. Math. Soc.* **80**:1 (2006), 65–80. MR 2212316 Zbl 1102.42007

[Grafakos and Torres 2002] L. Grafakos and R. H. Torres, "Multilinear Calderón–Zygmund theory", *Adv. Math.* **165**:1 (2002), 124–164. MR 1880324 Zbl 1032.42020

[Jones 1994] P. W. Jones, "Bilinear singular integrals and maximal functions", Chapter 10.2, p. 414, in *Linear and complex analysis: problem book 3, part I*, edited by V. P. Havin and N. K. Nikolski, Lecture Notes in Mathematics **1573**, Springer, Berlin, 1994.

[Jones 1997] R. L. Jones, "Ergodic theory and connections with analysis and probability", *New York J. Math.* **3A** (1997), 31–67. MR 1604581 Zbl 0898.28005

[Jones et al. 1998] R. L. Jones, R. Kaufman, J. M. Rosenblatt, and M. Wierdl, "Oscillation in ergodic theory", *Ergodic Theory Dynam. Systems* **18**:4 (1998), 889–935. MR 1645330 Zbl 0924.28009

[Lacey 2000] M. T. Lacey, "The bilinear maximal functions map into L^p for $2/3 < p \leq 1$", *Ann. of Math.* (2) **151**:1 (2000), 35–57. MR 1745019 Zbl 0967.47031

[Lacey and Thiele 1997] M. T. Lacey and C. Thiele, "L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$", *Ann. of Math.* (2) **146**:3 (1997), 693–724. MR 1491450 Zbl 0914.46034

[de Leeuw 1965] K. de Leeuw, "On L_p multipliers", *Ann. of Math.* (2) **81** (1965), 364–379. MR 0174937 Zbl 0171.11803

[Stein and Weiss 1971] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series **32**, Princeton University Press, 1971. MR 0304972 Zbl 0232.42007

[Varadarajan 2007] V. S. Varadarajan, "Euler and his work on infinite series", *Bull. Amer. Math. Soc.* (*N.S.*) **44**:4 (2007), 515–539. MR 2338363 Zbl 1135.01010

Received January 25, 2015.

DASHAN FAN
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF WISCONSIN–MILWAUKEE
MILWAUKEE, WI 53201
UNITED STATES
fan@uwm.edu

HUOXIONG WU
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
361005 FUJIAN
CHINA
huoxwu@xmu.edu.cn

FAYOU ZHAO
DEPARTMENT OF MATHEMATICS
SHANGHAI UNIVERSITY
200444 SHANGHAI
CHINA
fyzhao@shu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.365

THE TURAEV AND THURSTON NORMS

STEFAN FRIEDL, DANIEL S. SILVER AND SUSAN G. WILLIAMS

In 1986, W. Thurston introduced a (possibly degenerate) norm on the first cohomology group of a 3-manifold. Inspired by this definition, Turaev introduced in 2002 an analogous norm on the first cohomology group of a finite 2-complex. We show that if N is the exterior of a link in a rational homology sphere, then the Thurston norm agrees with a suitable variation of Turaev's norm defined on any 2-skeleton of N.

1. Introduction

W. Thurston [1986] introduced a seminorm for 3-manifolds N with empty or toroidal boundary. It is a function $x_N\colon H^1(N;\mathbb{Q}) \to \mathbb{Q}_{\geq 0}$ which measures the complexity of surfaces that are dual to cohomology classes. We adopt the custom of referring to x_N as the *Thurston norm*. It plays a central role in 3-manifold topology and we recall its definition in Section 2A, where we will also review several of its key properties.

Later, V. Turaev [2002] introduced an analogously defined seminorm for 2-complexes. For any finite 2-complex X with suitably defined boundary ∂X, Turaev defined $t_X\colon H^1(X, \partial X;\mathbb{Q}) \to \mathbb{Q}_{\geq 0}$ using complexities of dual 1-complexes. Inspired by work of C. McMullen [2002], Turaev gave lower bounds for t_X in terms of the multivariable Alexander polynomial whenever the boundary of X is empty. The precise definition of ∂X will be recalled in Section 2B. For the purpose of the introduction it suffices to know that if N is a compact triangulated 3-manifold, then the 2-skeleton $N^{(2)}$ is a finite 2-complex with empty boundary.

A homotopy equivalence induces a canonical isomorphism of homology and cohomology groups which we use to identify the groups. Examples given in [Turaev 2002, p. 143] show that t_X is *not* invariant under homotopy. We therefore introduce the following variation: For any finite 2-complex X with empty boundary, we define the *Turaev complexity function* as follows. If $\phi \in H^1(X;\mathbb{Q}) = \operatorname{Hom}(\pi_1(X),\mathbb{Q})$, then

$$\bar{t}_X(\phi) := \inf\left\{ t_Y(\phi\circ f) \;\middle|\; \begin{array}{l} Y \text{ is a finite 2-complex with } \partial Y = \varnothing \text{ and} \\ f\colon \pi_1(Y)\to\pi_1(X) \text{ is an isomorphism} \end{array} \right\}.$$

MSC2010: 57M27.
Keywords: Turaev norm, Thurston norm, links, Alexander polynomials.

Clearly $\bar{t}_X$ depends only on the fundamental group of X. Since the minimum of two norms need not satisfy the triangle inequality, the Turaev function is not a seminorm, as we will see later in Proposition 4.2.

For any 3-manifold N, we further define

$$\bar{t}_N(\phi) := \bar{t}_{N^{(2)}}(\phi),$$

where $N^{(2)}$ is the 2-skeleton of a triangulation of N. It is clear from the definition of $\bar{t}$ that $\bar{t}_N$ does not depend on the choice of a triangulation.

Given a 3-manifold N, it is natural to compare x_N and $\bar{t}_N$ on $H^1(N; \mathbb{Q})$. In general, they do not agree. Indeed in Section 4A we will see that there exist many examples of closed 3-manifolds N and classes $\phi \in H^1(N; \mathbb{Z})$ such that $\bar{t}_N(\phi) > x_N(\phi)$. The underlying reason is quite obvious: the Thurston norm is defined using complexities of surfaces, whereas the Turaev function is defined using complexities of graphs. However, the complexity of a closed surface is lower by at least one than the complexity of any underlying 1-skeleton.

It is therefore reasonable to restrict ourselves to the class of 3-manifolds where Thurston norm-minimizing surfaces can always be chosen to have no closed component. In Lemma 4.5 we will see that if $N = \Sigma^3 \setminus \nu L$ is the exterior of a of a link L in a rational homology sphere Σ, then N has this property. For simplicity of exposition we henceforth restrict ourselves to this type of 3-manifolds.

Using explicit and elementary constructions of 2-complexes, we prove the following.

Theorem 1.1. *Let N be the exterior of a link in a rational homology sphere. Then*

$$\bar{t}_N(\phi) \le x_N(\phi) \quad \textit{for any } \phi \in H^1(N; \mathbb{Q}).$$

It is natural to ask whether the extra freedom provided by working with 2-complexes instead of 3-manifolds allows us to get lower complexities. Our main theorem says that this is not the case, at least if we restrict ourselves to irreducible link exteriors. (Note that it follows from the definitions and from Schönflies theorem that the exterior of a link L in S^3 is irreducible if and only if L is nonsplit.)

Theorem 1.2. *Let N be the exterior of a link in a rational homology sphere. If N is irreducible, then*

$$\bar{t}_N(\phi) = x_N(\phi) \quad \textit{for any } \phi \in H^1(N; \mathbb{Q}).$$

We will prove the inequality $\bar{t}_X(\phi) \ge x_N(\phi)$ by studying the Alexander norms of finite covers of X and N, and by applying the recent results of I. Agol [2008; 2013], D. Wise [2009; 2012b; 2012a], P. Przytycki and D. Wise [2014; 2012] and Y. Liu [2013]. We do not know of an elementary proof of Theorem 1.2.

Theorem 1.2 fits into a long sequence of results showing that minimal-genus Seifert surfaces and Thurston norm-minimizing surfaces are "robust" in the sense that they "stay minimal" even if one relaxes some conditions. Examples of this phenomenon have been found by many authors, see, for example, [Gabai 1983; 1987; Kronheimer 1999; Friedl and Vidussi 2014; Nagel 2014; Friedl et al. 2015].

The paper is organized as follows. In Section 2 we recall the definition of the Thurston and Turaev norms, and we introduce the Turaev complexity function. In Section 3 we discuss the Alexander norm for 3-manifolds and 2-complexes, and we recall how they give lower bounds on the Thurston norm and Turaev complexity function, respectively. In Section 4A, we first show that the Turaev complexity function of the 2-skeleton can be greater than the corresponding Thurston norm. We then show in Section 4B that the Thurston norm of any irreducible 3-manifold with nontrivial toroidal boundary is detected by the Alexander norm of an appropriate finite cover. Finally, in Section 4C we put everything together to prove Theorem 1.2.

Conventions. All 3-manifolds are compact, orientable and connected, and all 2-complexes are connected, unless it says specifically otherwise.

2. The definition of the Thurston norm and the Turaev norm

2A. ***The Thurston norm and fibered classes.*** Let N be a 3-manifold with empty or toroidal boundary. The *Thurston norm* of a class $\phi \in H^1(N;\mathbb{Z})$ is defined as

$$x_N(\phi) = \min\{\chi_-(\Sigma) \mid \Sigma \subset N \text{ properly embedded surface dual to } \phi\}.$$

Here, $\chi_-(\Sigma)$ is the complexity of a surface Σ with connected components

$$\Sigma_1, \ldots, \Sigma_k,$$

given by

$$\chi_-(\Sigma) = \sum_{i=1}^{k} \max\{-\chi(\Sigma_i), 0\}.$$

Thurston [1986] showed that x_N defines a (possibly degenerate) norm on $H^1(N;\mathbb{Z})$. Note that any norm on $H^1(N;\mathbb{Z})$ extends uniquely to a norm on $H^1(N;\mathbb{Q})$, which we denote by the same symbol.

We say that a class $\phi \in H^1(N;\mathbb{Q})$ is *fibered* if there exists a fibration $p\colon N \to S^1$ such that ϕ lies in the pull-back of $H^1(S^1;\mathbb{Q})$ under p. By [Tischler 1970], a class $\phi \in H^1(N;\mathbb{Q})$ is fibered if and only if it can be represented by a nondegenerate closed 1-form.

Thurston [1986] showed the Thurston norm ball

$$\{\phi \in H^1(N;\mathbb{Q}) \mid x_N(\phi) \le 1\}$$

is a polyhedron. This implies that if C is a cone on a face of the polyhedron, then the restriction of x_N to C is a linear function. To put differently, for any $\alpha, \beta \in C$ and nonnegative $r, s \in \mathbb{Q}_{\geq 0}$, the linear combination $r\alpha + s\beta$ also lies in C, and $x_N(r\alpha + s\beta) = rx_N(\alpha) + sx_N(\beta)$.

Thurston [1986] also showed that any fibered class lies in the open cone on a top-dimensional face of the Thurston norm ball. Furthermore, any other class in that open cone is also fibered. Consequently, the set of fibered classes is the union of open cones on top-dimensional faces of the Thurston norm ball. We will refer to these cones as the *fibered cones of N*. A class $\phi \in H^1(N; \mathbb{Q})$ in the closure of a fibered cone is *quasifibered*.

2B. ***The Turaev norm and the Turaev complexity function for 2–complexes.*** As in [Turaev 2002], a *finite 2–complex* is the underlying topological space of a finite connected 2-dimensional CW-complex such that each point has a neighborhood homeomorphic to the cone over a finite graph. Examples of finite 2-complexes are given by compact surfaces (see [Turaev 2002, p. 138]), 2-skeletons of finite simplicial spaces, and the products of graphs with a closed interval.

The *interior* of X, denoted $\operatorname{Int} X$, is the set of points in X that have neighborhoods homeomorphic to $\mathbb{R}^2$. Finally the boundary ∂X of X is the closure in X of the set of all points of $X \setminus \operatorname{Int} X$ that have open neighborhoods in X homeomorphic to $\mathbb{R}$ or to $\mathbb{R} \times \mathbb{R}_{\geq 0}$. Note that ∂X is a graph contained in the 1-skeleton of the CW-decomposition of X. For example, if X is a compact surface, then ∂X is precisely the boundary of X in the usual sense.

Following [Turaev 2002], we say that a graph Γ in a finite 2-complex is *regular* if $\Gamma \subset X \setminus \partial X$ and if there exists a closed neighborhood in $X \setminus \partial X$ homeomorphic to $\Gamma \times [-1, 1]$ so that $\Gamma = \Gamma \times 0$. A *coorientation* for a regular graph Γ with components $\Gamma_1, \ldots, \Gamma_k$ is the choice of a component of $\Gamma_i \times [-1, 1] \setminus \Gamma_i$, for each $i = 1, \ldots, k$. A cooriented regular graph $\Gamma \subset X$ canonically defines an element $\phi_\Gamma \in H^1(X, \partial X; \mathbb{Z})$. Given any $\phi \in H^1(X, \partial X; \mathbb{Z})$, there exists a cooriented regular graph Γ with $\phi_\Gamma = \phi$. (We refer to [Turaev 2002] for details.)

Let X be a finite 2-complex with $\partial X = \varnothing$, and let $\phi \in H^1(X; \mathbb{Z})$. The *Turaev norm* of ϕ is

$$t_X(\phi) := \min\{\chi_-(\Gamma) \mid \Gamma \subset X \textit{ cooriented regular graph with } \phi_\Gamma = \phi\},$$

where $\chi_-(\Gamma)$ is the complexity of a graph Γ with connected components $\Gamma_1, \ldots, \Gamma_k$, given by

$$\chi_-(\Gamma) := \sum_{i=1}^{k} \max\{-\chi(\Gamma_i), 0\}.$$

Turaev [2002] showed that $t_X \colon H^1(X;\mathbb{Z}) \to \mathbb{Z}_{\geq 0}$ is a (possibly degenerate) norm, and, as in the previous section, t_X extends to a norm

$$t_X \colon H^1(X;\mathbb{Q}) \to \mathbb{Q}_{\geq 0}.$$

In Theorem 5.1 we will show that in general one has to allow disconnected graphs Γ to minimize the Turaev norm.

As we already mentioned in the introduction, Turaev [2002, p. 143] showed that t_X is in general not invariant under homotopy equivalence. (In fact Turaev showed that t_X is not even invariant under simple homotopy.) We therefore introduce a variation of the Turaev norm: if X is a finite 2-complex with $\partial X = \varnothing$, then given $\phi \in H^1(X;\mathbb{Q}) = \mathrm{Hom}(\pi_1(X),\mathbb{Q})$ the *Turaev complexity function* of ϕ is

$$\bar{t}_X(\phi) := \inf\left\{ t_\Gamma(\phi \circ f) \;\middle|\; \begin{array}{l} \Gamma \text{is a finite 2-complex with } \partial\Gamma = \varnothing \text{ and} \\ f \colon \pi_1(\Gamma) \to \pi_1(X) \text{ is an isomorphism} \end{array} \right\}.$$

We make the following observations:

(i) It is clear that $\bar{t}_X$ is invariant under homotopy equivalence. In fact $\bar{t}_X$ depends only on the fundamental group of X.

(ii) Since $\bar{t}_X$ is the infimum of continuous homogeneous functions (i.e., functions with $f(\lambda x) = \lambda f(x)$ for $\lambda > 0$), $\bar{t}_X$ is upper semicontinuous and homogeneous.

(iii) The complexity function $\bar{t}_X$ is defined as the infimum of norms. Note that the minimum of two norms is in general no longer a norm. For example, the infimum of the two norms $a(x, y) := |x|$ and $b(x, y) := |y|$ on $\mathbb{R}^2$ is not a norm. We will see in Proposition 4.2 that $\bar{t}_X(\phi)$ is, in general, not a norm.

(iv) From the definition, it follows immediately that

$$\bar{t}_X(\phi) \leq t_X(\phi),$$

for any $\phi \in H^1(X;\mathbb{Q})$.

(v) For any finite 2-complex X, Turaev [2002, Section 1.6] shows that t_X is algorithmically computable. We do not know whether this is also the case for the Turaev complexity function $\bar{t}_X$.

2C. ***An inequality between the Thurston norm and the Turaev complexity function.*** The goal of this section is to prove the following inequality between the Thurston norm and the Turaev complexity function.

Proposition 2.1. *Let N be a* 3*-manifold and let $\phi \in H^1(N;\mathbb{Z})$. If ϕ is dual to a properly embedded Thurston norm minimizing surface with r closed components, then*

$$\bar{t}_N(\phi) \leq x_N(\phi) + r.$$

Proof. Let $\phi \in H^1(N;\mathbb{Z})$ and let $\Sigma = \Sigma_1 \cup \cdots \cup \Sigma_s$ be a surface dual to ϕ

of minimal complexity such that $\Sigma_1, \dots, \Sigma_r$ are closed and $\Sigma_{r+1}, \dots, \Sigma_s$ have nonempty boundary.

For $i = 1, \dots, r$ we pick an embedded graph $\Gamma_i \subset \Sigma_i$ with $\chi(\Gamma_i) = \chi(\Sigma_i) - 1$ and such that $\pi_1(\Gamma_i)$ surjects onto $\pi_1(\Sigma_i)$. Furthermore, for $i = r+1, \dots, s$ we pick an embedded graph $\Gamma_i \subset \Sigma_i$ with $\chi(\Gamma_i) = \chi(\Sigma_i)$ and such that $\pi_1(\Gamma_i)$ surjects onto $\pi_1(\Sigma_i)$.

Next we select pairwise disjoint product neighborhoods

$$\Sigma_1 \times [-1, 1], \dots, \Sigma_s \times [-1, 1]$$

such that the product orientations match the orientation of N. We equip

$$M := N \setminus \bigcup_{i=1}^{s} \Sigma_i \times (-1, 1)$$

with a triangulation such that each $\Gamma_i \times \{\pm 1\}$ is a subspace of $M^{(1)}$. Consider

$$Y := M^{(2)} \cup \bigcup_{i=1}^{s} \Gamma_i \times (-1, 1).$$

It is straightforward to see that Y is a finite 2-complex with $\partial Y = \varnothing$, and the inclusion map $Y \to N$ induces an isomorphism of fundamental groups. By slight abuse of notation we denote the restriction of ϕ to Y again by ϕ.

For $i = 1, \dots, s$, we identify Γ_i with $\Gamma_i \times 0$. It is clear that $\Gamma := \Gamma_1 \cup \cdots \cup \Gamma_s$ is a regular graph on Y. Furthermore, with the obvious coorientation, we have $\phi_\Gamma = \phi$. It follows that

$$\begin{aligned}
\bar{t}_N(\phi) \le t_Y(\phi) \le \chi_-(\Gamma) &= \sum_{i=1}^{r} \max\{-\chi(\Gamma_i), 0\} + \sum_{i=r+1}^{s} \max\{-\chi(\Gamma_i), 0\} \\
&\le \sum_{i=1}^{r} \max\{-\chi(\Sigma_i) + 1, 0\} + \sum_{i=r+1}^{s} \max\{-\chi(\Sigma_i), 0\} \\
&\le \chi_-(\Sigma) + r \\
&= x_N(\phi) + r. \qquad \square
\end{aligned}$$

Theorem 1.1. *Let N be the exterior of a link in a rational homology sphere. Then for any $\phi \in H^1(N; \mathbb{Q})$, we have*

$$\bar{t}_N(\phi) \le x_N(\phi).$$

Proof. Let N be the exterior of a link in a rational homology sphere. We write $X = N^{(2)}$. Since $\bar{t}$ and x_N are homogeneous, it suffices to show that $\bar{t}_X(\phi) \le x_N(\phi)$ for every $\phi \in H^1(N; \mathbb{Z})$. Assume that $\phi \in H^1(N; \mathbb{Z})$. By Lemma 4.5 (see

Section 4A) there exists a Thurston norm-minimizing surface dual to ϕ such that each component has nonempty boundary. The desired inequality follows immediately from Proposition 2.1. □

3. Lower bounds on the norms coming from Alexander polynomials

3A. ***The Alexander polynomial.*** Let X be a compact CW-complex, and let

$$\varphi\colon H_1(X;\mathbb{Z}) \to H$$

be a homomorphism onto a free abelian group. We denote by $\tilde{X}^\varphi$ the cover of X corresponding to $\varphi\colon \pi_1(X) \to H_1(X;\mathbb{Z}) \to H$. The group H is the deck transformation group of $\tilde{X}^\varphi \to X$, and it acts on $H_1(\tilde{X}^\varphi;\mathbb{Z})$. Thus we can view $H_1(\tilde{X}^\varphi;\mathbb{Z})$ as a $\mathbb{Z}[H]$-module. Since $\mathbb{Z}[H]$ is a Noetherian ring, it follows that $H_1(\tilde{X}^\varphi;\mathbb{Z})$ is a finitely presented $\mathbb{Z}[H]$-module. This means that there exists an exact sequence

$$\mathbb{Z}[H]^r \xrightarrow{A} \mathbb{Z}[H]^s \to H_1(\tilde{X}^\varphi;\mathbb{Z}) \to 0.$$

After possibly adding columns of zeros, we can assume that $r \geq s$. Define the *Alexander polynomial* of (X,φ) to be

$$\Delta_{X,\varphi} := \text{gcd of all } s \times s\text{-minors of } A.$$

We refer to [Fox 1954; Turaev 2001; Hillman 2012] for the proof of the classical fact that $\Delta_{X,\varphi}$ is well-defined up to multiplication by a unit in $\mathbb{Z}[H]$, i.e., up to multiplication by an element of the form ϵh, where $\epsilon \in \{-1, 1\}$ and $h \in H$.

If $\varphi\colon H_1(X;\mathbb{Z}) \to H := H_1(X;\mathbb{Z})/\text{torsion}$ is the canonical projection, then we write $\Delta_X := \Delta_{X,\varphi}$, and we refer to it as the *Alexander polynomial* Δ_X of X. Furthermore, if $\phi \in H^1(X;\mathbb{Z}) = \operatorname{Hom}(\pi_1(X),\mathbb{Z})$, then we view the corresponding Alexander polynomial $\Delta_{X,\varphi}$ as an element in $\mathbb{Z}[t^{\pm 1}]$ under the canonical identification of the group ring $\mathbb{Z}[\mathbb{Z}]$ with the Laurent polynomial ring $\mathbb{Z}[t^{\pm 1}]$.

3B. ***The one-variable Alexander polynomials.*** In this section we relate the degrees of one-variable Alexander polynomials to the Thurston norm and to the Turaev complexity function.

In the following, given a nonzero polynomial $p(t) = \sum_{i=r}^{s} a_i t^i$ with $a_r \neq 0$ and $a_s \neq 0$, we write

$$\deg(p(t)) = s - r.$$

Note that the degree of a nonzero one-variable Alexander polynomial is well-defined.

The following proposition is well known, see, for example, [Friedl and Kim 2006] for a proof.

Proposition 3.1. *Let N be a closed 3-manifold and let $\phi \in H^1(N;\mathbb{Z})$ be primitive. If $\Delta_{N,\phi} \neq 0$, then*

$$x_N(\phi) \geq \deg(\Delta_{N,\phi}) - 2.$$

Furthermore, equality holds if ϕ is a fibered class and if $N \neq S^1 \times S^2$.

We prove the following.

Proposition 3.2. *Let X be a finite 2-complex with $\partial X = \varnothing$, and let be $\phi \in H^1(N;\mathbb{Z})$ primitive. If $\Delta_{X,\phi} \neq 0$, then*

$$\bar{t}_X(\phi) \geq \deg(\Delta_{X,\phi}) - 1.$$

Proof. Let Y be a finite 2-complex with $\partial Y = \varnothing$, and let $\psi \in H^1(Y;\mathbb{Z})$ be primitive. If $\Delta_{Y,\phi} \neq 0$, then it follows from Claim 2 of [Turaev 2002, p. 152] that

$$t_Y(\psi) \geq \deg(\Delta_{Y,\psi}) - 1.$$

The desired inequality

$$\bar{t}_X(\phi) \geq \deg(\Delta_{X,\phi}) - 1$$

is an immediate consequence of this fact and the observation that the Alexander polynomial depends only on the fundamental group of X. □

3C. ***The Alexander norm.*** Let X be a compact connected CW-complex. We write $H := H_1(X;\mathbb{Z})/\text{torsion}$ and also $\Delta_X = \sum_{h \in H} a_h h$. Let

$$\phi \in H^1(X;\mathbb{Q}) = \mathrm{Hom}(\pi_1(X),\mathbb{Q}) = \mathrm{Hom}(H,\mathbb{Q}).$$

Following [McMullen 2002], we define the *Alexander norm* of ϕ by

$$a_X(\phi) := \max\{\phi(h) - \phi(g) \mid a_g \neq 0 \text{ and } a_h \neq 0\}.$$

It is straightforward to see that a_X is indeed a norm on $H^1(X;\mathbb{Q})$. As in the proof of Proposition 3.2, we use that fact that the Alexander polynomial and thus the Alexander norm depend only on the fundamental group of X. More precisely, if $f\colon Y \to X$ is a map of compact connected CW-complexes that induces an isomorphism of fundamental groups, then

$$f_*(\Delta_Y) = \Delta_X \in \mathbb{Z}[H_1(X;\mathbb{Z})/\text{torsion}],$$

and thus, for any $f \in H^1(X;\mathbb{Q}) = \mathrm{Hom}(\pi_1(X),\mathbb{Q})$, we have

$$a_Y(\phi \circ f^*) = a_X(\phi). \tag{1}$$

We begin with the following theorem due to McMullen [2002].

Theorem 3.3. *Let N be a 3-manifold with empty or toroidal boundary and with $b_1(N) \geq 2$. Then*

$$a_N(\phi) \leq x_N(\phi) \quad \textit{for any } \phi \in H^1(N; \mathbb{Q}).$$

Furthermore, equality holds for quasifibered classes.

Proof. Let N be a 3-manifold with empty or toroidal boundary and with $b_1(N) \geq 2$. McMullen [2002, Theorem 1.1] showed that

$$a_N(\phi) \leq x_N(\phi) \quad \text{for any } \phi \in H^1(N; \mathbb{Q})$$

and that equality holds for all integral fibered classes. Since a_N and x_N are homogeneous, it follows immediately that equality also holds for all fibered classes and, in fact, for all quasifibered classes. □

The following analogous theorem, which says that the Alexander norm also gives lower bounds on the Thurston norm and the Turaev complexity function, is due to Turaev [2002].

Theorem 3.4. *Let X be a finite 2-complex with $b_1(X) \geq 2$ and such that $\partial X = \varnothing$. Then*

$$a_X(\phi) \leq \bar{t}_X(\phi) \leq t_X(\phi) \quad \textit{for any } \phi \in H^1(X; \mathbb{Q}).$$

Proof. Let Y be a finite 2-complex with $b_1(Y) \geq 2$ and such that $\partial Y = \varnothing$. Then by [Turaev 2002, Theorem 3.1], we have

$$a_Y(\psi) \leq t_Y(\psi) \quad \text{for any } \psi \in H^1(Y; \mathbb{Q}).$$

The theorem now follows immediately from combining this result with the definition of $\bar{t}_X(\phi)$ and (1). □

4. Proofs

4A. ***The Thurston norm and the Turaev complexity function for closed 3-manifolds.*** The combination of Propositions 3.1, 3.2 and 2.1 gives us the following theorem showing that the Thurston norm of a closed 3-manifold need not agree with Turaev complexity function of its 2-skeleton.

Theorem 4.1. *Let $N \neq S^1 \times S^2$ be a closed 3-manifold and let $\phi \in H^1(N; \mathbb{Z})$ be a primitive fibered class. Then*

$$\bar{t}_N(\phi) = x_N(\phi) + 1.$$

We also prove:

Proposition 4.2. *There exists a finite 2-complex X with $\partial X = \varnothing$ such that $\bar{t}_X$ does not satisfy the triangle inequality, i.e., $\bar{t}_X$ is not a norm.*

Proof. Let N be a fibered 3-manifold with $b_1(N) = 2$. We write $X = N^{(2)}$ for some triangulation of N. As we mentioned in Section 2A, by [Thurston 1986] there exists an open 2-dimensional cone $C \subset H^1(N; \mathbb{Q})$ such that all classes in C are fibered and such that x_N is a linear function on C.

Given $\phi \in H^1(N; \mathbb{Z})$ we denote by

$$\operatorname{div}(\phi) := \max\{k \in \mathbb{N} \mid \text{there exists } \psi \in H^1(N; \mathbb{Z}) \text{ with } \phi = k\psi\}$$

the divisibility of ϕ. It follows from Theorem 4.1 and the homogeneity of the Thurston norm and the Turaev complexity function that

$$\bar{t}_X(\phi) = x_N(\phi) + \operatorname{div}(\phi) \qquad \text{for any} \phi \in H^1(N; \mathbb{Z}) \cap C. \tag{2}$$

We prove the following claim.

Claim. There exist $\alpha, \beta \in C$ with $\operatorname{div}(\alpha) + \operatorname{div}(\beta) < \operatorname{div}(\alpha + \beta)$.

Pick two primitive vectors $\phi, \psi \in C$ which are not collinear. Since ϕ and ψ lie in the cone C, it follows that any nonnegative linear combination of ϕ and ψ also lies in C.

Select a coordinate system for $H^1(N; \mathbb{Z})$, that is, choose an identification of $H^1(N; \mathbb{Z})$ with $\mathbb{Z}^2$. Since ϕ is primitive, we can assume that $\phi = (1, 0)$. Since ψ is also primitive, we know that $\psi = (x, y)$ for some coprime x and y. Since ϕ and ψ are not collinear, $y \neq 0$. Choose a prime $p > 1 + |y|$. We consider $\alpha = (1, 0)$ and $\beta = (px + (p-1), py)$. Note that p can not divide $px + p - 1 = p(x+1) - 1$. It follows that $\operatorname{div}(\beta) = \gcd(px + (p-1), py) \leq |y|$. Evidently $\operatorname{div}(\alpha) = 1$. Now

$$\operatorname{div}(\alpha+\beta) = \operatorname{div}(px+p, py) = \gcd(px+p, py) \geq p > 1+|y| \geq \operatorname{div}(\alpha)+\operatorname{div}(\beta).$$

This concludes the proof of the claim.

If we combine the claim and the linearity of x_N on C with equality (2), then we obtain that

$$\begin{aligned}
\bar{t}_X(\alpha+\beta) = x_N(\alpha+\beta) + \operatorname{div}(\alpha+\beta) &= x_N(\alpha) + x_N(\beta) + \operatorname{div}(\alpha+\beta) \\
&> x_N(\alpha) + \operatorname{div}(\alpha) + x_N(\beta) + \operatorname{div}(\beta) \\
&= \bar{t}_X(\alpha) + \bar{t}_X(\beta).
\end{aligned}$$

We have shown that $\bar{t}_X$ does not satisfy the triangle inequality. □

4B. ***The Alexander norm of finite covers of* 3*-manifolds.*** We begin with the following theorem. We state it in slightly greater generality than we actually need, since the result has independent interest.

Theorem 4.3. *Let $N \neq S^1 \times D^2$ be an aspherical* 3*-manifold with empty or toroidal boundary. If N is neither a Nil-manifold nor a Sol-manifold, there exists a finite* *cover* $p\colon \tilde{N} \to N$ *such that* $b_1(\tilde{N}) \geq 2$ *and such that*

$$a_{\tilde{N}}(p^*\phi) = x_{\tilde{N}}(p^*\phi) \text{ for any } \phi \in H^1(N;\mathbb{Q}).$$

The proof of the theorem will require the remainder of Section 4B. The theorem was proved for graph manifolds by Nagel [2014]. We will therefore restrict ourselves to the case of manifolds that are not (closed) graph manifolds. The main ingredient in our proof of Theorem 4.3 will be the following theorem, a consequence of the seminal work of Agol [2008; 2013], Wise [2009; 2012b; 2012a], Przytycki and Wise [2014; 2012] and Liu [2013]. We summarize the main points of the proof for the convenience of the reader.

Theorem 4.4. *Let N be an irreducible 3-manifold with empty or toroidal boundary that is not a closed graph manifold. Then there exists a finite cover $p\colon \tilde{N} \to N$ such that, for any $\phi \in H^1(N;\mathbb{Q})$, the pull-back $p^*\phi$ is quasifibered.*

Proof. Let N be an irreducible 3-manifold that is not a closed graph manifold. It follows from [Agol 2013; Wise 2009; 2012b; 2012a; Przytycki and Wise 2014; 2012; Liu 2013] that $\pi_1(N)$ is virtually RFRS, i.e., $\pi_1(N)$ admits a finite index subgroup which is RFRS (residually finite rationally solvable). The precise definition of RFRS, references for which can be found in [Aschenbrenner et al. 2015], is not of concern to us. What matters is that Agol [2008, Theorem 5.1] (see also [Friedl and Kitayama 2014, Theorem 5.1]) showed that if ψ lies in $H^1(N;\mathbb{Q})$ and if N is an irreducible 3-manifold such that $\pi_1(N)$ is virtually RFRS, then there exists a finite cover $p\colon \hat{N} \to N$ such that $p^*\psi$ lies in the closure of a fibered cone of $\hat{N}$.

By picking one class in each cone of the Thurston norm ball of N and iteratively applying Agol's theorem, one can easily show that there exists a finite cover $p\colon \tilde{N} \to N$ such that for any $\phi \in H^1(N;\mathbb{Q})$ the pull-back $p^*\phi$ lies in the closure of a fibered cone of $\tilde{N}$. We refer to [Friedl and Vidussi 2015, Corollary 5.2] for details. □

If N is a graph manifold with nonempty boundary, then the conclusion of Theorem 4.4 also follows from facts that are more classical. This argument is not used anywhere else in the paper, but since it is perhaps of independent interest we give a very quick sketch of the argument.

Proof of Theorem 4.4 if N is a graph manifold. Let N be a graph manifold with boundary. It follows from [Wang and Yu 1997, Theorem 0.1] and classical arguments (see e.g., [Aschenbrenner and Friedl 2013, Section 4.3.4.3] and [Hempel 1987]) that there exists a finite cover $\tilde{N}$ of N that is fibered and such that if $\{N_v\}_{v\in V}$ denotes the set of JSJ components of $\tilde{N}$, then each N_v is of the form $S^1 \times \Sigma_v$ for some surface Σ_v. (For the meaning of JSJ components, see [Aschenbrenner et al. 2015, Section 1.6].)

For each $v \in V$ we write $t_v = S^1 \times P_v$, where $P_v \in \Sigma_v$ is a point. It follows from [Eisenbud and Neumann 1985, Theorem 4.2] that a class $\phi \in H^1(\tilde{N}; \mathbb{Q})$ is fibered if and only if $\phi(t_v) \neq 0$ for all $v \in V$. Since $\tilde{N}$ is fibered it now follows that all classes in $H^1(\tilde{N}; \mathbb{Q})$ outside of finitely many hyperplanes are fibered. Hence all classes in $H^1(\tilde{N}; \mathbb{Q})$ are quasifibered. □

We can now move on to the proof of Theorem 4.3. Note that arguments similar to the proof of Theorem 4.3 were also used in [Friedl and Vidussi 2014; 2015].

Proof of Theorem 4.3. Let $N \neq S^1 \times D^2$ be an irreducible 3-manifold with empty or toroidal boundary that is not a closed graph manifold. Since we assumed that $N \neq S^1 \times D^2$, it now follows from Agol's theorem [2013] and classical 3-manifold topology that N has a finite cover with b_1 at least two. (We refer to [Aschenbrenner et al. 2015] for details.) We can therefore assume that we already have $b_1(N) \geq 2$.

By Theorem 4.4 there exists a finite cover $p\colon \tilde{N} \to N$ such that for any ϕ in $H^1(N; \mathbb{Q})$, the pull-back $p^*\phi$ is quasifibered. Note that Betti numbers never decrease by going to finite covers, i.e., we have $b_1(\tilde{N}) \geq b_1(N) \geq 2$. It follows from Theorem 3.3 that

$$a_{\tilde{N}}(p^*\phi) = x_{\tilde{N}}(p^*\phi) \quad \text{for any } \phi \in H^1(N; \mathbb{Q}).$$

This concludes the proof of the theorem. □

4C. *Proof of Theorem 1.2.* Before we turn to the proof of Theorem 1.2 we prove the following well-known lemma.

Lemma 4.5. *If N is the exterior of a link in a rational homology sphere, then any class $\phi \in H^1(N; \mathbb{Z})$ is dual to a surface Σ of minimal complexity such that all components of Σ have nonempty boundary.*

Proof. Let N be the exterior of a link in a rational homology sphere. It follows from a Mayer–Vietoris argument that the map $H_1(\partial N; \mathbb{Q}) \to H_1(N; \mathbb{Q})$ is surjective. It follows from Poincaré duality and the Universal Coefficient Theorem that the boundary map $\partial\colon H_2(N, \partial N; \mathbb{Z}) \to H_1(\partial N; \mathbb{Z})$ has finite kernel. Since $H_2(N, \partial N; \mathbb{Z}) \cong H^1(N; \mathbb{Z}) \cong \operatorname{Hom}(H_1(N; \mathbb{Z}), \mathbb{Z})$ is torsion-free it follows that the boundary map $\partial\colon H_2(N, \partial N; \mathbb{Z}) \to H_1(\partial N; \mathbb{Z})$ is in fact injective. In particular this implies that closed surfaces represent the trivial homology class in $(N, \partial N)$. Now let $\phi \in H^1(N; \mathbb{Z})$, and let Σ be a properly embedded minimal-complexity surface dual to ϕ. By the above observation, the closed components of Σ are null-homologous. It follows that the union of the components of Σ with nontrivial boundary represents the same homology as Σ. Since removing components can never increase the complexity, we have shown that ϕ is dual to a surface Σ of minimal complexity such that all components of Σ have nonempty boundary. □

In the previous sections we collected all the tools that now allow us to finally complete the proof of Theorem 1.2.

Theorem 1.2. *Let N be the exterior of a link in a rational homology sphere. If N is irreducible, then for any $\phi \in H^1(N;\mathbb{Q})$ we have*

$$\bar{t}_N(\phi) = x_N(\phi).$$

Proof. It remains to prove that $\bar{t}_N(\phi) \geq x_N(\phi)$. Let N be the exterior of a link in a rational homology sphere. Suppose that N is irreducible. Let $\phi \in H^1(N;\mathbb{Q})$. It suffices to show that if Y is a finite 2-complex Y with $\partial Y = \varnothing$ and if $f : \pi_1(Y) \to \pi_1(N)$ is an isomorphism, then

$$t_Y(\phi \circ f) \geq x_N(\phi).$$

So let Y and f be as above. By a slight abuse of notation we denote $\phi \circ f : \pi_1(Y) \to \mathbb{Q}$ by ϕ as well.

By Theorem 4.3 there exists a finite cover $p\colon \tilde{N} \to N$ such that $b_1(\tilde{N}) \geq 2$ and such that

$$a_{\tilde{N}}(p^*\phi) = x_{\tilde{N}}(p^*\phi).$$

We write $\pi = \pi_1(N)$ and $\tilde{\pi} := \pi_1(\tilde{N})$, and we denote by $p\colon \tilde{Y} \to Y$ the finite cover corresponding to $f^{-1}(\tilde{\pi})$. Note that $\tilde{Y}$ is also a finite 2-complex with $\partial\tilde{Y} = \varnothing$. It follows immediately from the definitions that

$$x_{\tilde{N}}(p^*\phi) \leq [\pi : \tilde{\pi}] \cdot x_N(\phi) \quad \text{and} \quad t_{\tilde{Y}}(p^*\phi) \leq [\pi : \tilde{\pi}] \cdot t_Y(\phi).$$

In fact, Gabai [1983, Corollary 6.13] showed that the above is an equality for the Thurston norm, i.e., we have the equality:

$$x_{\tilde{N}}(p^*\phi) = [\pi : \tilde{\pi}] \cdot x_N(\phi).$$

Combining the above results with Theorem 3.4, we see that

$$[\pi : \tilde{\pi}] \cdot t_Y(\phi) \geq t_{\tilde{Y}}(p^*\phi) \geq a_{\tilde{N}}(p^*\phi) = x_{\tilde{N}}(p^*\phi) = [\pi : \tilde{\pi}] \cdot x_N(\phi).$$

This concludes the proof the theorem. □

4D. ***Fundamental group complexity.*** Let X be a finite 2-complex with $\partial X = \varnothing$, and $\phi \in H^1(X;\mathbb{Z}) = \mathrm{Hom}(\pi_1(X),\mathbb{Z})$. Turaev [2002] describes a method by which we can compute $t_X(\phi)$ using cocycles. We start by orienting edges (i.e., open 1-cells) of X, and then select a $\mathbb{Z}$-valued cellular cocycle k on X representing ϕ. We let

$$|k| = \sum_e (n_e/2 - 1)|k(e)|,$$

where e ranges over all edges in X, $k(e) \in \mathbb{Z}$ is the value of k on e, and n_e is the number of 2-cells adjacent to e, counted with multiplicity. (Note that $n_e \geq 2$ since $\partial X = \varnothing$.) Turaev [2002, Section 1.6] proves that $t_X(\phi)$ is the minimum value of $|k|$ as k ranges over all cellular cocycles representing ϕ.

When the 0-skeleton of X consists of a single vertex, the 2-complex determines a group presentation P for $\pi_1(X)$, and hence $|k|$ can be defined on the level of presentations.

Given a finite presentation $P = \langle x_1, \ldots, x_m \mid r_1, \ldots, r_n \rangle$, following [Turaev 2002], we denote by $\#(x_i)$ the number of appearances of $x_i^{\pm 1}$ in the words $r_1, \ldots, r_n$. We say that P is a *good presentation* if each $\#(x_i) \geq 2$. We are interested in good presentations, since it is straightforward to see that the canonical 2-complex corresponding to a good presentation has empty boundary. Also note that any finitely presented group admits a good presentation. Indeed, if $\#(x_i) = 1$, then we can eliminate x_i using a Tietze move. If $\#(x_i) = 0$, then we can add a trivial relator $x_i x_i^{-1}$.

Now let $P = \langle x_1, \ldots, x_m \mid r_1, \ldots, r_n \rangle$ be a good presentation for a group π, and let ϕ be a homomorphism $\phi \colon \pi \to \mathbb{Z}$. We define

$$t_P(\phi) = \sum_i (\#(x_i)/2 - 1)\, |\phi(x_i)|.$$

Furthermore we define $\bar{t}_\pi(\phi)$ to be the minimum of $t_P(\phi)$ as P ranges over all good presentations of π. We extend the definition in the usual way for rational cohomology classes $\phi \in H^1(X; \mathbb{Q})$.

Lemma 4.6. *Let X be a finite 2-complex with $\partial X = \varnothing$ and $\phi \in H^1(X; \mathbb{Q})$. We write $\pi = \pi_1(X)$. Then*

$$\bar{t}_X(\phi) \leq \bar{t}_\pi(\phi).$$

Proof. Given a good presentation P for π, we construct the canonical finite 2-complex Y with $\pi_1(Y) \cong \pi$. Let k be the unique 1-cocycle representing ϕ. A straightforward argument shows that $\bar{t}_X(\phi) \leq |k| = t_P(\phi)$; see also [Turaev 2002, Section 1.8]. Since this is true for any good presentation of $\pi_1(X)$, we have $\bar{t}_X(\phi) \leq \bar{t}_\pi(\phi)$. □

Example 4.7. Let π the fundamental group of the exterior of a knot K in the 3-sphere. Let ϕ be the abelianization homomorphism, mapping a meridian to 1. If P is a Wirtinger presentation corresponding to a diagram for K, then one sees easily that $t_P(\phi)$ is the number of crossings of the diagram.

It is usually possible to find presentations yielding a smaller value $t_P(\phi)$. Let Σ be a Seifert surface for K having minimal genus g. By splitting π along $\pi_1(\Sigma)$, we obtain an HNN-decomposition for π of the form

$$\langle A, x \mid \mu(b) = xbx^{-1} \text{ for all } b \in \pi_1(\Sigma) \rangle,$$

where A is the fundamental group of the knot exterior split along Σ, and

$$\mu\colon \pi_1(\Sigma) \to A$$

is injective. For such a presentation P, we have $t_P(\phi) = 2g - 1$. It follows by the next result that this value is the smallest possible; i.e., $\bar{t}_\pi(\phi) = 2g - 1$.

Theorem 4.8. *Let N be the exterior of a link in a rational homology sphere with group π. If N is irreducible, then for any $\phi \in H^1(N; \mathbb{Q})$ such that $\Delta_{N,\phi} \neq 0$, we have*

$$\bar{t}_N(\phi) = \bar{t}_\pi(\phi) = x_N(\phi).$$

Remark. Turaev [2002] gives several examples of knot groups and presentations of minimal complexity. He states that it would be interesting to find other examples. Theorem 4.8 shows how to construct presentations of minimal complexity for any knot in a rational homology sphere.

Proof. By Lemma 4.6 and Theorem 1.2, it suffices to prove that $\bar{t}_\pi(\phi) \le x_N(\phi)$, for any $\phi \in H^1(N; \mathbb{Q})$. By the homogeneity of the Turaev function and the Thurston norm we may assume that ϕ is an integral primitive cohomology class.

Consider a Thurston norm-minimizing surface $\Sigma \subset N$ for ϕ. Our assumption that $\Delta_{N,\phi}$ is not identically zero ensures that the first Betti number of $\mathrm{Ker}(\phi)$ is finite. By a short argument in the beginning of the proof of McMullen [2002, Proposition 6.1], the surface Σ is connected. Its boundary is nonempty by Lemma 4.5. Splitting π along $\pi_1(\Sigma)$, as above, we obtain a presentation P with complexity $2g - 1$, where g is the genus of Σ. Since $t_N(\phi) = 2g - 1$, we are done. $\square$

We conclude this section with the following conjecture:

Conjecture 4.9. *Let X be a finite 2-complex with $\partial X = \varnothing$. Then*

$$\bar{t}_X(\phi) = \bar{t}_{\pi_1(X)}(\phi) \quad \text{for any } \phi \in H^1(X; \mathbb{Q}).$$

Note that an affirmative answer to this question together with Theorems 1.1 and 1.2 would show that the conclusion of Theorem 4.8 holds for any irreducible link complement N, without any assumptions on ϕ.

5. Disconnected minimal dual graphs

It is natural to ask whether one can always realize the Turaev norm of a primitive cohomology class by a connected graph. In this final section of the paper we will see that this is not the case. More precisely, we have the following theorem.

Theorem 5.1. *Given any n there exists a 2-complex X with $\partial X = \varnothing$ and a primitive class $\phi \in H^1(\pi; \mathbb{Z})$ such that for any 2-complex Y with $\pi_1(Y) = \pi_1(X)$ and with $\partial Y = \varnothing$ the following holds: any graph Γ in Y that represents ϕ with $\bar{t}_X(\phi) = \chi_-(\Gamma)$ has at least n components.*

Proof. We consider the good presentation

$$P = \langle a_1, \dots, a_n, x_1, \dots, x_n \mid [x_i, a_i], i = 1, \dots, n \rangle,$$

and we denote by X the corresponding 2-complex, which is just the join of n tori $T_1, \dots, T_n$. Clearly $\partial X = \varnothing$.

We write $\pi = \pi_1(X)$. The group π is the free product of n free abelian groups $\langle a_i, x_i \mid [a_i, x_i] \rangle$, $i = 1, \dots, n$ of rank two. We consider the epimorphism $\phi\colon \pi \to \mathbb{Z}$ that is defined by $\phi(a_i) = 0,\ i = 1, \dots, n$ and $\phi(x_i) = 1,\ i = 1, \dots, n$. It is clear that on each torus T_i there exists a circle, disjoint from the gluing point, such that the union of these circles is dual to ϕ. We thus see that $\bar{t}_X(\phi) = 0$.

Now let Y be a 2-complex with $\pi_1(Y) = \pi$ and with $\partial Y = \varnothing$. Let Γ be a graph on Y which is dual to ϕ with $\chi_-(\Gamma) = 0$. We will show that Γ has at least n components. Note that $\chi_-(\Gamma) = 0$ implies that any component of Γ is either a point or a circle. We denote by m the number of components of Γ that are circles. We will see that $m \geq n$.

Claim. The module $H_1(Y; \mathbb{Q}[t^{\pm 1}])$ is isomorphic to

$$\mathbb{Q}[t^{\pm 1}]^{n-1} \oplus \bigoplus_{i=1}^{n} \mathbb{Q}[t^{\pm 1}]/(t-1).$$

We first note that $H_1(Y; \mathbb{Q}[t^{\pm 1}]) = H_1(X; \mathbb{Q}[t^{\pm 1}])$. A straightforward application of Fox calculus (see [Fox 1953]) shows that

$$H_1(X; \mathbb{Q}[t^{\pm 1}]) \cong \mathbb{Q}[t^{\pm 1}]^{n-1} \oplus \bigoplus_{i=1}^{n} \mathbb{Q}[t^{\pm 1}]/(t-1).$$

This concludes the proof of the claim.

Now we write $W = Y \setminus \Gamma \times (-1, 1)$. The usual Meyer–Vietoris sequence with $\mathbb{Q}[t^{\pm 1}]$-coefficients corresponding to $Y = W \cup \Gamma \times [-1, 1]$ gives rise to the exact sequence

$$\cdots \to H_1(\Gamma; \mathbb{Q}[t^{\pm 1}]) \xrightarrow{\iota_- - t\iota_+} H_1(W; \mathbb{Q}[t^{\pm 1}]) \to H_1(Y; \mathbb{Q}[t^{\pm 1}]) \to H_0(\Gamma; \mathbb{Q}[t^{\pm 1}]) \to \cdots$$

Note that ϕ vanishes on Γ and W. It follows that $H_*(\Gamma; \mathbb{Q}[t^{\pm 1}])$ and $H_*(W; \mathbb{Q}[t^{\pm 1}])$ are free $\mathbb{Q}[t^{\pm 1}]$-modules. Furthermore, by the above discussion of Γ we know that $H_1(\Gamma; \mathbb{Q}[t^{\pm 1}]) \cong \mathbb{Q}[t^{\pm 1}]^m$. It follows immediately from the above exact sequence and the classification of modules over PIDs that the torsion submodule of $H_1(Y; \mathbb{Q}[t^{\pm 1}])$ is generated by m elements.

On the other hand, we had just seen that the torsion submodule of $H_1(Y; \mathbb{Q}[t^{\pm 1}])$ is isomorphic to $\oplus_{i=1}^n \mathbb{Q}[t^{\pm 1}]/(t-1)$. It follows from the classification of modules over the PID $\mathbb{Q}[t^{\pm 1}]$ that the minimal number of generators of the torsion submodule of $H_1(Y; \mathbb{Q}[t^{\pm 1}])$ is n. Putting everything together we deduce that $m \geq n$. □

Acknowledgments

Friedl gratefully acknowledges the support provided by the SFB 1085 "Higher Invariants" at the Universität Regensburg, funded by the Deutsche Forschungsgemeinschaft (DFG). Silver and Williams thank the Simons Foundation for its support. We wish to thank the referee for carefully reading our paper and for pointing out a well-hidden mistake in an argument.

References

[Agol 2008] I. Agol, "Criteria for virtual fibering", *J. Topol.* **1**:2 (2008), 269–284. MR Zbl

[Agol 2013] I. Agol, "The virtual Haken conjecture", *Doc. Math.* **18** (2013), 1045–1087. MR Zbl

[Aschenbrenner and Friedl 2013] M. Aschenbrenner and S. Friedl, *3-manifold groups are virtually residually p*, Memoirs of the American Mathematical Society **1058**, 2013. MR

[Aschenbrenner et al. 2015] M. Aschenbrenner, S. Friedl, and H. Wilton, *3-manifold groups*, European Mathematical Society (EMS), Zürich, 2015. MR Zbl

[Eisenbud and Neumann 1985] D. Eisenbud and W. Neumann, *Three-dimensional link theory and invariants of plane curve singularities*, Annals of Mathematics Studies **110**, Princeton University Press, 1985. MR Zbl

[Fox 1953] R. H. Fox, "Free differential calculus, I: Derivation in the free group ring", *Ann. of Math.* (2) **57** (1953), 547–560. MR Zbl

[Fox 1954] R. H. Fox, "Free differential calculus, II: The isomorphism problem of groups", *Ann. of Math.* (2) **59** (1954), 196–210. MR Zbl

[Friedl and Kim 2006] S. Friedl and T. Kim, "The Thurston norm, fibered manifolds and twisted Alexander polynomials", *Topology* **45**:6 (2006), 929–953. MR Zbl

[Friedl and Kitayama 2014] S. Friedl and T. Kitayama, "The virtual fibering theorem for 3-manifolds", *Enseign. Math.* **60**:1-2 (2014), 79–107. MR Zbl

[Friedl and Vidussi 2014] S. Friedl and S. Vidussi, "Minimal genus in 4-manifolds with a free circle action", *Adv. Math.* **250** (2014), 570–587. MR Zbl

[Friedl and Vidussi 2015] S. Friedl and S. Vidussi, "The Thurston norm and twisted Alexander polynomials", *J. Reine Angew. Math.* **707** (2015), 87–102. MR Zbl

[Friedl et al. 2015] S. Friedl, D. S. Silver, and S. G. Williams, "Splittings of knot groups", *Math. Ann.* **362**:1-2 (2015), 401–424. MR Zbl

[Gabai 1983] D. Gabai, "Foliations and the topology of 3-manifolds", *J. Differential Geom.* **18**:3 (1983), 445–503. MR Zbl

[Gabai 1987] D. Gabai, "Foliations and the topology of 3-manifolds, III", *J. Differential Geom.* **26**:3 (1987), 479–536. MR Zbl

[Hempel 1987] J. Hempel, "Residual finiteness for 3-manifolds", pp. 379–396 in *Combinatorial group theory and topology* (Alta, Utah, 1984), edited by S. M. Gersten and J. R. Stallings, Ann. of Math. Stud. **111**, Princeton University Press, 1987. MR Zbl

[Hillman 2012] J. Hillman, *Algebraic invariants of links*, 2nd ed., Series on Knots and Everything **52**, World Scientific, Hackensack, NJ, 2012. MR Zbl

[Kronheimer 1999] P. B. Kronheimer, "Minimal genus in $S^1 \times M^3$", *Invent. Math.* **135**:1 (1999), 45–61. MR Zbl

[Liu 2013] Y. Liu, "Virtual cubulation of nonpositively curved graph manifolds", *J. Topol.* **6**:4 (2013), 793–822. MR Zbl

[McMullen 2002] C. T. McMullen, "The Alexander polynomial of a 3-manifold and the Thurston norm on cohomology", *Ann. Sci. École Norm. Sup.* (4) **35**:2 (2002), 153–171. MR Zbl

[Nagel 2014] M. Nagel, "Minimal genus in circle bundles over 3-manifolds", preprint, 2014. arXiv

[Przytycki and Wise 2012] P. Przytycki and D. T. Wise, "Mixed 3-manifolds are virtually special", preprint, 2012. arXiv

[Przytycki and Wise 2014] P. Przytycki and D. T. Wise, "Graph manifolds with boundary are virtually special", *J. Topol.* **7**:2 (2014), 419–435. MR Zbl

[Thurston 1986] W. P. Thurston, "A norm for the homology of 3-manifolds", pp. 99–130 Memoirs of the American Mathematical Society **339**, 1986. MR

[Tischler 1970] D. Tischler, "On fibering certain foliated manifolds over S^1", *Topology* **9** (1970), 153–154. MR Zbl

[Turaev 2001] V. Turaev, *Introduction to combinatorial torsions*, Birkhäuser, Basel, 2001. MR Zbl

[Turaev 2002] V. Turaev, "A norm for the cohomology of 2-complexes", *Algebr. Geom. Topol.* **2** (2002), 137–155. MR Zbl

[Wang and Yu 1997] S. Wang and F. Yu, "Graph manifolds with non-empty boundary are covered by surface bundles", *Math. Proc. Cambridge Philos. Soc.* **122**:3 (1997), 447–455. MR Zbl

[Wise 2009] D. T. Wise, "Research announcement: the structure of groups with a quasiconvex hierarchy", *Electron. Res. Announc. Math. Sci.* **16** (2009), 44–55. MR Zbl

[Wise 2012a] D. T. Wise, *From riches to raags: 3-manifolds, right-angled Artin groups, and cubical geometry*, CBMS Regional Conference Series in Mathematics **117**, American Mathematical Society, 2012. MR Zbl

[Wise 2012b] D. T. Wise, "The structure of groups with a quasi-convex hierarchy", preprint, 2012, available at https://goo.gl/ucBthh.

Received July 30, 2015. Revised February 23, 2016.

STEFAN FRIEDL
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT REGENSBURG
D-93047 REGENSBURG
GERMANY
sfriedl@gmail.com

DANIEL S. SILVER
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SOUTH ALABAMA
MOBILE, AL 36688-0002
UNITED STATES
silver@southalabama.edu

SUSAN G. WILLIAMS
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SOUTH ALABAMA
MOBILE, AL 36688-0002
UNITED STATES
swilliam@southalabama.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.383

A NOTE ON NONUNITAL ABSORBING EXTENSIONS

JAMES GABE

Elliott and Kucerovsky stated that a nonunital extension of separable C^*-algebras with a stable ideal is nuclearly absorbing if and only if the extension is purely large. However, their proof was flawed. We give a counterexample to their theorem as stated, but establish an equivalent formulation of nuclear absorption under a very mild additional assumption to being purely large. In particular, if the quotient algebra is nonunital, then we show that the original theorem applies. We also examine how this affects results in classification theory.

1. Introduction and a counterexample

A (unital) extension of C^*-algebras $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ is called *(unitally) weakly nuclear* if there is a (unital) completely positive splitting $\sigma : \mathfrak{A} \to \mathfrak{E}$ which is weakly nuclear, i.e., for every $b \in \mathfrak{B}$ the map $b\sigma(-)b^* : \mathfrak{A} \to \mathfrak{B}$ is nuclear. Such an extension is called trivial if we may take the weakly nuclear splitting to be a $*$-homomorphism. An extension is called (unitally) nuclearly absorbing if it absorbs every trivial, (unitally) weakly nuclear extension, i.e., the Cuntz sum of our given extension $\mathfrak{e}$ with any trivial, (unitally) weakly nuclear extension is strongly unitarily equivalent to $\mathfrak{e}$. A remarkable result of Elliott and Kucerovsky [2001] shows that a unital, separable extension with a stable ideal is unitally nuclearly absorbing if and only if the extension is *purely large*. Recall that an extension $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ of C^*-algebras with $\mathfrak{B}$ stable is called purely large if for any $x \in \mathfrak{E} \setminus \mathfrak{B}$, the hereditary C^*-subalgebra $\overline{x\mathfrak{B}x^*}$ of $\mathfrak{B}$ contains a stable, σ-unital C^*-subalgebra $\mathfrak{D}$ which is full in $\mathfrak{B}$. Note that we have added the requirement that $\mathfrak{D}$ be σ-unital, since this was implicitly used in [op. cit., Lemma 7] and since this is automatic in the separable case, which is our main concern.

In their paper, Elliott and Kucerovsky use the unital version above to obtain a nonunital version of this result, i.e., that a nonunital extension is nuclearly absorbing

This work was supported by the Danish National Research Foundation through the Centre for Symmetry and Deformation (DNRF92).
MSC2010: 46L05, 46L35, 46L80.
Keywords: absorbing extensions, corona factorisation property, KK-theory, classification.

if and only if it is purely large. Unfortunately this is not true. We provide a counterexample below.

A stable C^*-algebra is said to have the *corona factorisation property* if all full multiplier projections are Murray–von Neumann equivalent, or equivalently, all norm-full multiplier projections are properly infinite. As is shown in [Kucerovsky and Ng 2006a], any full extension by a σ-unital, stable C^*-algebra with the corona factorisation property is purely large in the sense of [Elliott and Kucerovsky 2001]. Here *full* means that the Busby map is full, i.e., that it maps nonzero elements to full elements in the corona algebra.

It is known that C^*-algebras which do not have the corona factorisation property have rather exotic properties; see, e.g., [Kucerovsky and Ng 2006b]. It follows by [Robert 2011, Corollary 1] that any σ-unital, stable C^*-algebra with finite nuclear dimension, or, more generally, nuclear dimension less than ω, has the corona factorisation property. Thus for classification purposes, the corona factorisation property is not really any restriction.

After receiving an early version of this note, Efren Ruiz constructed a counterexample to [Eilers et al. 2014, Theorem 4.9]. In fact, by using results from this note, Ruiz has constructed two graphs such that the induced C^*-algebras have exactly one nontrivial ideal, have isomorphic six-term exact sequences in K-theory with order and scale, but for which the C^*-algebras are nonisomorphic. This implies that we do not have a complete classification of graph C^*-algebras with exactly one nontrivial ideal using the above K-theoretic invariant, as opposed to what was previously believed. The counterexample is provided in Section 4. Fortunately, all recent classification results of *stable* graph C^*-algebras are unaffected by the issues addressed in this note, and hence stand as given.

As for general notation in this note we let π denote the quotient map from the multiplier algebra of some C^*-algebra to its corona algebra, and we consider an essential extension algebra as a C^*-subalgebra of the multiplier algebra of the ideal. When referring to full elements in a multiplier algebra, we always mean with respect to the norm topology, and *not* the strict topology.

A counterexample to [Elliott and Kucerovsky 2001, Corollary 16] could be as follows.

Example 1.1. Let $\mathfrak{A} = \mathbb{C}$, $\mathfrak{B} = \mathbb{K} \oplus \mathbb{K}$, and consider the trivial extension $\mathfrak{E}$ with splitting $\sigma(1) = P \oplus 1 \in \mathcal{M}(\mathbb{K}) \oplus \mathcal{M}(\mathbb{K}) \cong \mathcal{M}(\mathfrak{B})$, where P is a full projection in $\mathcal{M}(\mathbb{K})$ such that $1 - P$ is also full. The extension $\mathfrak{E}$ is clearly full, and since $\mathfrak{B}$ has the corona factorisation property, this implies that $\mathfrak{E}$ is a nonunital, purely large extension. However, it does not absorb the zero extension, i.e., the extension with the zero Busby map. This is easily seen by projecting to the second coordinate in the corona algebra $\pi_2 : \mathcal{Q}(\mathfrak{B}) \cong \mathcal{Q}(\mathbb{K}) \oplus \mathcal{Q}(\mathbb{K}) \to \mathcal{Q}(\mathbb{K})$, since $\pi_2(\tau(1)) = 1$ and $\pi_2((\tau \oplus 0)(1))$ is a nontrivial projection, where τ denotes the Busby map.

The flaw in the original proof is the claim that a nonunital extension $\mathfrak{E}$ is purely large if and only if its unitisation $\mathfrak{E}^\dagger$ is purely large. The sufficiency is trivial but the necessity is incorrect.

Lemma 1.2. *There exists a nonunital purely large extension such that the unitisation is not purely large.*

Proof. Let $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ denote the extension of Example 1.1. The unitisation $\mathfrak{E}^\dagger$ has Busby map $\tau^\dagger : \mathbb{C} \oplus \mathbb{C} \to \mathcal{Q}(\mathbb{K}) \oplus \mathcal{Q}(\mathbb{K})$ given by

$$\tau^\dagger(1 \oplus 0) = \pi(P) \oplus 1 \quad \text{and} \quad \tau^\dagger(0 \oplus 1) = \pi(1-P) \oplus 0.$$

Since $\pi(1-P) \oplus 0$ is not full in $\mathcal{Q}(\mathbb{K}) \oplus \mathcal{Q}(\mathbb{K})$, $\tau^\dagger$ is not a full homomorphism and thus the extension can not be purely large. $\square$

2. Fixing the theorem

We will start by showing that the original theorem still holds, if we assume that the quotient is nonunital.

Theorem 2.1. *Let $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ be an extension of separable C^*-algebras with $\mathfrak{B}$ stable. Suppose that $\mathfrak{A}$ is nonunital. Then the extension is nuclearly absorbing if and only if it is purely large.*

Proof. As in [Elliott and Kucerovsky 2001, Section 16] the extension is nuclearly absorbing if and only if the unitised extension is unitally nuclearly absorbing, which in turn is equivalent to the unitised extension being purely large. Thus it suffices to show that this is equivalent to the nonunitised extension being purely large. We use the same proof as in the original paper. Clearly the extension is purely large if the unitisation is purely large. Assume that the nonunital extension is purely large. Note, in particular, that the Busby map τ is injective. It suffices to show that $\overline{(1-x)\mathfrak{B}(1-x)^*}$ contains a stable C^*-subalgebra which is full in $\mathfrak{B}$ for any $x \in \mathfrak{E}$. Suppose that $(1-x)\mathfrak{E} \subset \mathfrak{B}$. Then $\pi(x)$ is a unit for $\pi(\mathfrak{E}) = \tau(\mathfrak{A}) \subset \mathcal{Q}(\mathfrak{B})$. However, this contradicts the fact that $\mathfrak{A}$ is nonunital, since the Busby map τ is injective. Hence we may find $x' \in \mathfrak{E}$ such that $(1-x)x' \notin \mathfrak{B}$. Since

$$\overline{(1-x)x'\mathfrak{B}((1-x)x')^*} \subset \overline{(1-x)\mathfrak{B}(1-x)^*}$$

and since the nonunital extension is purely large, the former of these contains a stable C^*-subalgebra which is full in $\mathfrak{B}$. $\square$

To prove a stronger result, where the assumption that the quotient being unital is removed, we will use the following lemma.

Lemma 2.2. *Let $\mathfrak{B}$ be a stable, separable C^*-algebra, and let $P \in \mathcal{M}(\mathfrak{B})$ be a full, properly infinite projection. Then the trivial extension of $\mathbb{C}$ by $\mathfrak{B}$ with splitting σ given by $\sigma(1) = P$ is purely large.*

Proof. If $P = 1$ then the extension is the canonical unitisation extension

$$0 \to B \to B^{\dagger} \to \mathbb{C} \to 0,$$

which is clearly self-absorbing. It follows from [Elliott and Kucerovsky 2001] that it is purely large.

It is well known, since $\mathfrak{B}$ is stable, that P is full and properly infinite exactly when it is Murray–von Neumann equivalent to 1. Let v be an isometry such that $vv^* = P$ and let $t_1, t_2 \in \mathcal{M}(\mathfrak{B})$ be such that $t_1t_1^* + t_2t_2^* = P = t_1^*t_1 = t_2^*t_2$. Then $s_1 := t_1 v$ and $s_2 := t_2 + (1 - P)$ are the canonical generators of a unital copy of $\mathcal{O}_2$ in $\mathcal{M}(\mathfrak{B})$, for which $P = s_1s_1^* + s_2Ps_2^*$. Hence,

$$\pi(\sigma(1)) = \pi(s_1)1\pi(s_1)^* + \pi(s_2)\pi(P)\pi(s_2)^*,$$

which implies that our extension is the Cuntz sum of the unitisation extension and itself. It follows from [op. cit., Lemma 13] that our extension is purely large. □

Now for the stronger case where we allow the quotient to be unital.

Theorem 2.3. *Let* $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ *be an extension of separable* C^**-algebras with* $\mathfrak{B}$ *stable. The extension is nuclearly absorbing if and only if it is purely large and there is a full, properly infinite projection* $P \in \mathcal{M}(\mathfrak{B})$ *such that* $P\mathfrak{E} \subset \mathfrak{B}$.

Proof. Assume that the extension is nuclearly absorbing. Then it absorbs the zero extension so we may assume that the Busby map is of the form $\tau \oplus 0$, where the symbol $\oplus$ denotes a Cuntz sum. Let $P = 0 \oplus 1$. Then $P\mathfrak{E} \subset \mathfrak{B}$ since $\pi(P)$ annihilates the image of the Busby map. Moreover, the extension absorbs some purely large extension and is thus itself purely large by [loc. cit.].

Now suppose that the extension is purely large and that P is a full, properly infinite projection such that $P\mathfrak{E} \subset \mathfrak{B}$. As in the proof of Theorem 2.1 it suffices to show that the unitised extension is purely large. It is enough to show that $\overline{(1-x)\mathfrak{B}(1-x)^*}$ contains a stable C^*-subalgebra which is full in $\mathfrak{B}$ for any $x \in \mathfrak{E}$. Observe that

$$\overline{(1-x)P\mathfrak{B}P(1-x)^*} \subset \overline{(1-x)\mathfrak{B}(1-x)^*}.$$

Since $(1-x)P = P - xP$ and $xP \in \mathfrak{B}$, it suffices to show that the extension

$$0 \to \mathfrak{B} \to \mathfrak{B} + \mathbb{C}P \to \mathbb{C} \to 0$$

is purely large. This follows from Lemma 2.2. □

Note that an extension must be nonunital in order to satisfy the equivalent conditions in the above theorem. We immediately get the following corollary.

Corollary 2.4. *Let* $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ *be an extension of separable* C^**-algebras with* $\mathfrak{B}$ *stable. Then the extension is nuclearly absorbing if and only if it is purely large and absorbs the zero extension.*

When we assume that the ideal has the corona factorisation property, then we get a perhaps more hands-on way of checking if a full extension is nuclearly absorbing. To exhibit this we introduce the following definition.

Definition 2.5. Let $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ be an extension of C^*-algebras. We say that the extension is *unitisably full* if the unitised extension $0 \to \mathfrak{B} \to \mathfrak{E}^\dagger \to \mathfrak{A}^\dagger \to 0$ is full.

It is clear that if an extension is unitisably full, then it is full and nonunital. If the quotient algebra $\mathfrak{A}$ is unital, then the extension is unitisably full if and only if the extension is full and $1_{\mathcal{Q}(\mathfrak{B})} - \tau(1_{\mathfrak{A}})$ is full, where τ denotes the Busby map. Note that this case is our main concern due to Theorem 2.1.

Theorem 2.6. *Let $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ be an extension of separable C^*-algebras, such that $\mathfrak{B}$ is stable and has the corona factorisation property. Then the extension is nuclearly absorbing if and only if the extension is unitisably full.*

Proof. As in the proof of Theorem 2.1 the extension is nuclearly absorbing if and only if the unitised extension is purely large. Since $\mathfrak{B}$ has the corona factorisation property, this is the case if and only if the extension is unitisably full. □

We will end this section by showing that, in the absence of the corona factorisation property, there are purely large, unitisably full extensions which are not nuclearly absorbing. We will need a converse of Lemma 2.2.

Proposition 2.7. *Let $\mathfrak{B}$ be a stable, separable C^*-algebra, and let $P \in \mathcal{M}(\mathfrak{B})$ be a full projection. Then the trivial extension of $\mathbb{C}$ by $\mathfrak{B}$ with splitting σ given by $\sigma(1) = P$ is purely large if and only if P is properly infinite.*

Proof. One direction is Lemma 2.2. Suppose that the extension is purely large. It suffices to show that the Cuntz sum $P \oplus 0$ is properly infinite. The extension with splitting $\sigma'(1) = P \oplus 0$ is purely large and absorbs the zero extension, and thus it is absorbing by Corollary 2.4. Since the extension with splitting $\sigma_0(1) = 1 \oplus 0$ is also absorbing, there is a unitary $U \in \mathcal{M}(\mathfrak{B})$ such that $U^*(P \oplus 0)U - 1 \oplus 0 \in \mathfrak{B}$. Pick an isometry $V \in \mathcal{M}(\mathfrak{B})$ such that $V^*(1 \oplus 0)V = 1$. Then

$$V^*(U^*(P \oplus 0)U - 1 \oplus 0)V = (UV)^*(P \oplus 0)UV - 1 \in \mathfrak{B}.$$

Since $\mathfrak{B}$ is stable, we may find an isometry W such that

$$\|(UVW)^*(P \oplus 0)UVW - 1\| = \|W^*((UV)^*(P \oplus 0)UV - 1)W\| < 1.$$

This implies that $P \oplus 0$, and thus also P, is properly infinite. □

We can now extend our class of counterexamples to include purely large, unitisably full extensions $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ which are not nuclearly absorbing. In fact, such an extension can be made for any $\mathfrak{B}$ without the corona factorisation property.

Proposition 2.8. *Let $\mathfrak{B}$ be a stable, separable C^*-algebra which does not have the corona factorisation property. Then there is a purely large, unitisably full extension of $\mathbb{C}$ by $\mathfrak{B}$ which is not nuclearly absorbing.*

Proof. Let Q be a full multiplier projection which is not properly infinite, but where $P := 1 - Q$ is properly infinite and full. Such a projection can be obtained by taking any full multiplier projection Q' which is not properly infinite, and letting $Q = Q' \oplus 0$ be a Cuntz sum. In fact, $P = 1 - Q$ will be properly infinite since it majorises the properly infinite, full projection $0 \oplus 1$. Consider the trivial extension $\mathfrak{E}$ of $\mathbb{C}$ by $\mathfrak{B}$ with splitting $\sigma(1) = P$. The unitised extension has a splitting $\sigma_1 : \mathbb{C} \oplus \mathbb{C} \to \mathcal{M}(\mathfrak{B})$ given by $\sigma_1(1 \oplus 0) = P$ and $\sigma_1(0 \oplus 1) = Q$. Since both P and Q are full and orthogonal, the unitised extension is full.

By Proposition 2.7 the extension is purely large. Such an extension is nuclearly absorbing exactly when its unitisation is purely large [Elliott and Kucerovsky 2001]. If the unitisation was purely large, then $\overline{(Q-b)\mathfrak{B}(Q-b)^*}$ would contain a stable C^*-subalgebra full in $\mathfrak{B}$, for every $b \in \mathfrak{B}$. However, this would imply that the extension of $\mathbb{C}$ by $\mathfrak{B}$ with splitting $\sigma_0(1) = Q$ is purely large, which it is not by Proposition 2.7. Hence the extension is not nuclearly absorbing. $\square$

3. How this affects classification results

In the classification of nonsimple C^*-algebras, a popular result has been a result of Kucerovsky and Ng, which says that under the mild condition of the corona factorisation property on a stable, separable C^*-algebra $\mathfrak{B}$, $KK^1(\mathfrak{A}, \mathfrak{B})$ is the group of unitary equivalence classes of full extensions $\mathfrak{E}$ of $\mathfrak{A}$ by $\mathfrak{B}$ for any nuclear separable C^*-algebra $\mathfrak{A}$. This is unfortunately not the case. The theorem only remains true if one adds the condition that the extensions are unitisably full as in Definition 2.5. See Theorem 3.2 below.

A counterexample of the original result could be as follows.

Example 3.1. Let $0 \to \mathfrak{B} \to \mathfrak{E} \to \mathfrak{A} \to 0$ be the extension from Example 1.1 with Busby map τ. Then $\mathfrak{B}$ has the corona factorisation property and the extension is full. As seen in Example 1.1, τ and $\tau \oplus 0$ are both nonunital and are *not* unitarily equivalent. However, they define the same element in $KK^1(\mathfrak{A}, \mathfrak{B})$.

The closest we get to fixing the theorem would be the following.

Theorem 3.2. *Let $\mathfrak{B}$ be a separable, stable C^*-algebra. Then the following are equivalent*:

(i) *$\mathfrak{B}$ has the corona factorisation property.*

(ii) *For any separable C^*-algebra $\mathfrak{A}$, $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$ is the group of strong unitary equivalence classes of all full, weakly nuclear extensions of $\mathfrak{A}$ by $\mathfrak{B}$ which absorb the zero extension.*

(iii) *For any separable C^*-algebra $\mathfrak{A}$, $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$ is the group of strong unitary equivalence classes of all full, weakly nuclear extensions $\mathfrak{E}$ of $\mathfrak{A}$ by $\mathfrak{B}$, for which there is a full projection $P \in \mathcal{M}(\mathfrak{B})$ such that $P\mathfrak{E} \subset \mathfrak{B}$.*

(iv) *For any separable C^*-algebra $\mathfrak{A}$, $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$ is the group of strong unitary equivalence classes of all unitisably full, weakly nuclear extensions $\mathfrak{E}$ of $\mathfrak{A}$ by $\mathfrak{B}$.*

Proof. It is well-known that $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$ is (isomorphic to) the group of strong unitary equivalence classes of weakly nuclear extensions of $\mathfrak{A}$ by $\mathfrak{B}$ which are nuclearly absorbing. This is proved in [Kirchberg 2000, Sections 3 and 4], though in a much more general setting. Alternatively, one can prove this exactly as one proves that $KK^1(A, B) \cong \mathrm{Ext}^{-1}(A, B)$, and then note that the latter can be viewed as strong unitary equivalence classes of semisplit extensions of A by B which are absorbing. Note that weakly nuclear extensions are automatically semisplit. Thus (i) $\Rightarrow$ (iv) by Theorem 2.6, and (iv) $\Rightarrow$ (i) follows from Proposition 2.8.

If $\mathfrak{B}$ has the corona factorisation property, then any full extension by $\mathfrak{B}$ is purely large. Thus (i) $\Rightarrow$ (iii) follows from Theorem 2.3.

Clearly (iii) is equivalent to the condition that for any C^*-algebra $\mathfrak{A}$, any full, weakly nuclear extension $\mathfrak{E}$ of $\mathfrak{A}$ by $\mathfrak{B}$, for which there is a full projection $P \in \mathcal{M}(\mathfrak{B})$ such that $P\mathfrak{E} \subset \mathfrak{B}$, is nuclearly absorbing. If the extension $\mathfrak{E}$ has Busby map $\tau \oplus 0$, then $(0 \oplus 1)\mathfrak{E} \subset \mathfrak{B}$, and thus (iii) $\Rightarrow$ (ii).

It remains to show (ii) $\Rightarrow$ (i). Let $P \in \mathcal{M}(\mathfrak{B})$ be a full projection, and let $P \oplus 0$ be the Cuntz sum. Note that $Q \sim Q \oplus 0$ for any projection Q. By (ii), the extension with the Busby map $\tau : \mathbb{C} \to \mathcal{Q}(\mathfrak{B})$ given by $\tau(1) = \pi(P \oplus 0)$ is nuclearly absorbing. In particular, it absorbs the unitisation extension of $\mathfrak{B}$. Consider the lift $\rho(1) = P \oplus 0$ of τ and the canonical lift of the unitisation extension of $\mathfrak{B}$. We may find a unitary $u \in \mathcal{M}(\mathfrak{B})$ such that $u^*(P \oplus 0 \oplus 0)u - 0 \oplus 0 \oplus 1 \in \mathfrak{B}$. If v is an isometry such that $vv^* = 0 \oplus 0 \oplus 1$, then $(uv)^*(P \oplus 0 \oplus 0)uv - 1 \in \mathfrak{B}$. Thus we may pick an isometry w such that

$$\|w^*((uv)^*(P \oplus 0 \oplus 0)uv - 1)w\| = \|s^*(P \oplus 0 \oplus 0)s - 1\| < 1,$$

where s is the isometry uvw. Hence P is Murray–von Neumann equivalent to $s^*(P \oplus 0 \oplus 0)s$, which is equivalent to 1. □

Remark 3.3. It clearly follows from the proof above that we could restrict our attention only to nuclear C^*-algebras $\mathfrak{A}$ if desired. In this case we can remove the weakly nuclear condition, since all extensions of a separable, nuclear C^*-algebra are weakly nuclear by the lifting theorem of Choi and Effros [1976], and also we would have $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B}) = KK^1(\mathfrak{A}, \mathfrak{B})$.

We still get some nice results for classification. This follows from the above theorem and Theorem 2.1.

Corollary 3.4. *Let $\mathfrak{B}$ be a separable, stable C^*-algebra with the corona factorisation property and let $\mathfrak{A}$ be a nonunital, separable C^*-algebra. Then $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$ is the group of strong unitary equivalence classes of all full, weakly nuclear extensions of $\mathfrak{A}$ by $\mathfrak{B}$.*

Corollary 3.5. *Let $\mathfrak{B}$ be a separable, stable C^*-algebra with the corona factorisation property and let $\mathfrak{A}$ be a separable C^*-algebra. Let $\mathfrak{E}_i$ be full, weakly nuclear extensions of $\mathfrak{A}$ by $\mathfrak{B}$, with Busby maps τ_i, for $i = 1, 2$. If $[\tau_1] = [\tau_2] \in KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B})$, then $\mathfrak{E}_1 \otimes \mathbb{K} \cong \mathfrak{E}_2 \otimes \mathbb{K}$.*

Proof. Given a Busby map $\tau : \mathfrak{A} \to \mathcal{Q}(\mathfrak{B})$, let τ^s be the composition

$$\mathfrak{A} \otimes \mathbb{K} \xrightarrow{\tau \otimes \mathrm{id}} \mathcal{Q}(\mathfrak{B}) \otimes \mathbb{K} \hookrightarrow \mathcal{Q}(\mathfrak{B} \otimes \mathbb{K}).$$

It is well known that the map $KK^1_{\mathrm{nuc}}(\mathfrak{A}, \mathfrak{B}) \to KK^1_{\mathrm{nuc}}(\mathfrak{A} \otimes \mathbb{K}, \mathfrak{B} \otimes \mathbb{K})$ given by $[\tau] \mapsto [\tau^s]$, is an isomorphism (the proof is identical to the similar result in classical KK-theory). Thus τ_1^s and τ_2^s are strongly unitarily equivalent by Corollary 3.4, and since their corresponding extension algebras are $\mathfrak{E}_1 \otimes \mathbb{K}$ and $\mathfrak{E}_2 \otimes \mathbb{K}$ respectively, it follows that $\mathfrak{E}_1 \otimes \mathbb{K} \cong \mathfrak{E}_2 \otimes \mathbb{K}$. $\square$

Remark 3.6. Every result in this note holds with the ideal $\mathfrak{B}$ being σ-unital instead of separable. The quotient $\mathfrak{A}$ should still be separable. This is a special case of a much more general result in [Gabe and Ruiz 2015].

4. The counterexample of Ruiz

Definition 4.1. Let $E = (E^0, E^1, r, s)$ be a (countable, directed) graph. The *graph C^*-algebra* $C^*(E)$ is the universal C^*-algebra generated by mutually orthogonal projections p_v for $v \in E^0$, and isometries s_e for $e \in E^1$, which satisfy the relations

- $s_e^* s_f = \delta_{ef}\, p_{r(e)}$ for all $e, f \in E^1$,
- $s_e s_e^* \leq p_{s(e)}$ for all $e \in E^1$,
- $p_v = \sum_{e \in s^{-1}(\{v\})} s_e s_e^*$ for all $v \in E^0$ satisfying $0 < |s^{-1}(\{v\})| < \infty$.

Example 4.2 (the counterexample). Theorem 4.9 of [Eilers et al. 2014] states that if $C^*(E)$ and $C^*(F)$ are nonunital and both have exactly one nontrivial, two-sided, closed ideal, and the induced six-term exact sequences in K-theory are isomorphic, such that the isomorphisms on all K_0-groups preserve order and scale, then $C^*(E) \cong C^*(F)$. We will provide a counterexample to this result.

Let E and F be the respective graphs

$$\begin{tikzcd} v \arrow[loop left] \arrow[loop right] \arrow[d] & & & \\ v_0 \arrow[r, "2"] & v_1 \arrow[r, "2"] & v_2 \arrow[r, "2"] & \cdots \end{tikzcd} \qquad \begin{tikzcd} & & w \arrow[loop left] \arrow[loop right] \arrow[d] & & & \\ \cdots \arrow[r] & w_{-1} \arrow[r] & w_0 \arrow[r, "2"] & w_1 \arrow[r, "2"] & w_2 \arrow[r, "2"] & \cdots \end{tikzcd}$$

Both $C^*(E)$ and $C^*(F)$ are nonunital, full extensions of the Cuntz algebra $\mathcal{O}_2$ by the stabilisation of the CAR algebra $M_{2^\infty} \otimes \mathbb{K}$. The six-term exact sequences of the induced extensions, where we write the K_0-groups with order and scale as $(K_0(\mathfrak{A}),\, K_0(\mathfrak{A})^+,\, \Sigma K_0(\mathfrak{A}))$, are both isomorphic to

$$\begin{array}{ccccc} \left(\mathbb{Z}[\frac{1}{2}], \mathbb{Z}[\frac{1}{2}]_+, \mathbb{Z}[\frac{1}{2}]_+\right) & \xrightarrow{(\mathrm{id},\iota,\iota)} & \left(\mathbb{Z}[\frac{1}{2}], \mathbb{Z}[\frac{1}{2}], \mathbb{Z}[\frac{1}{2}]\right) & \longrightarrow & (0,0,0) \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & 0 & \longleftarrow & 0 \end{array}$$

where $\mathbb{Z}[\frac{1}{2}]_+ = \mathbb{Z}[\frac{1}{2}] \cap [0,\infty)$ and $\iota : \mathbb{Z}[\frac{1}{2}]_+ \hookrightarrow \mathbb{Z}[\frac{1}{2}]$ is the canonical inclusion. To compute the order and scale of $K_0(C^*(E))$ and $K_0(C^*(F))$ we simply use that both $C^*(E)$ and $C^*(F)$ contain full, properly infinite projections, p_v and p_w, respectively, and apply [Rørdam 2002, Proposition 4.1.4]. That p_v and p_w are properly infinite follows since v and w both support two loops, so it follows easily from the defining relations that they are properly infinite. Thus if [Eilers et al. 2014, Theorem 4.9] were true, it should follow that $C^*(E) \cong C^*(F)$. We will show that this is not the case, by showing that one extension with $C^*(F)$ is nuclearly absorbing, but that the extension with $C^*(E)$ is not.

The extension with $C^*(F)$ is nuclearly absorbing. Recall that F^* denotes the set of paths in F, and that if $\alpha = e_1 \cdots e_n \in F^*$ then $s_\alpha := s_{e_1} \cdots s_{e_n}$, and that $r(\alpha) = r(e_n)$ and $s(\alpha) = s(e_1)$. Let $\mathfrak{I}_F$ denote the unique nontrivial ideal in $C^*(F)$, which is isomorphic to $M_{2^\infty} \otimes \mathbb{K}$. By [Ruiz and Tomforde 2014], we may describe $\mathfrak{I}_F$ as

$$\mathfrak{I}_F = \overline{\operatorname{span}}\{s_\alpha s_\beta^* : \alpha, \beta \in F^*, r(\alpha) = r(\beta) = w_n \text{ for some } n \in \mathbb{Z}\}.$$

Let $P = \sum_{n=1}^\infty w_{-n}$ which is easily seen to converge strictly in the multiplier algebra of $\mathfrak{I}_F$. We clearly have that $PC^*(F) \subset \mathfrak{I}_F$. Thus, if P is a full, properly infinite projection in $\mathcal{M}(\mathfrak{I}_F)$, then it follows from Theorem 2.3 that the extension with $C^*(F)$ is nuclearly absorbing. Since $\mathfrak{I}_F$ has the corona factorisation property, it suffices to show that P is full.

Note that $M_{2^\infty} \cong p_{w_0} \mathfrak{I}_F p_{w_0}$. Let ρ denote the unique tracial state on $p_{w_0} \mathfrak{I}_F p_{w_0}$, and ρ_∞ denote the induced trace function on $\mathcal{M}(\mathfrak{I}_F)_+$. It follows from [Rørdam 1991, Theorem 4.4] that P is full if and only if $\rho_\infty(P) = \infty$. Since $p_{w_{-n}}$ for $n > 0$ is Murray–von Neumann equivalent to p_{v_0}, it follows that $\rho(p_{w_{-n}}) = \rho(p_{v_0}) = 1$ and thus $\rho_\infty(P) = \sum_{n=1}^\infty \rho(p_{w_{-n}}) = \infty$. Thus the extension

$$0 \to \mathfrak{I}_F \to C^*(F) \to C^*(F)/\mathfrak{I}_F \to 0$$

is nuclearly absorbing.

The extension with $C^*(E)$ is not nuclearly absorbing. Let $\mathfrak{I}_E$ denote the unique nontrivial ideal in $C^*(E)$, which is isomorphic to $M_{2^\infty} \otimes \mathbb{K}$. As above, we may describe $\mathfrak{I}_E$ as

$$\mathfrak{I}_E = \overline{\operatorname{span}}\{s_\alpha s_\beta^* : \alpha, \beta \in E^* \text{ and } r(\alpha) = r(\beta) = v_n \text{ for some } n \geq 0\}.$$

To show that the extension $\mathfrak{e} : 0 \to \mathfrak{I}_E \to C^*(E) \to \mathcal{O}_2 \to 0$ is not nuclearly absorbing, it suffices to show that the unitised extension $\mathfrak{e}^\dagger : 0 \to \mathfrak{I}_E \to C^*(E)^\dagger \to \mathcal{O}_2 \oplus \mathbb{C} \to 0$ is *not* full. Let $\sigma : C^*(E) \to \mathcal{M}(\mathfrak{I}_E)$ be the canonical $*$-homomorphism. Then $C^*(E)^\dagger \cong \sigma(C^*(E)) + \mathbb{C}1_{\mathcal{M}(\mathfrak{I}_E)}$. Note that $1 - \sigma(p_v)$ is a lift of $(0, 1) \in \mathcal{O}_2 \oplus \mathbb{C}$ (under the obvious identifications), so if $\mathfrak{e}^\dagger$ is full, we should have that $1-\sigma(p_v)+\mathfrak{I}_E$ is full in $\mathcal{Q}(\mathfrak{I}_E)$ by [Kucerovsky and Ng 2006a, Proposition 3.3]. Since $\mathfrak{I}_E$ is stable, fullness of $1 - \sigma(p_v) + \mathfrak{I}_E$ is equivalent to fullness of $1 - \sigma(p_v)$ in $\mathcal{M}(\mathfrak{I}_E)$.

The corner in $\mathfrak{I}_E$ generated by $1 - \sigma(p_v)$ is easily seen to be

$$\overline{\operatorname{span}}\{s_\alpha s_\beta^* : s(\alpha) \neq v \neq s(\beta)\},$$

which has an approximate unit $\left(\sum_{n=0}^k p_{v_n}\right)_{k=1}^\infty$. Thus $1 - \sigma(p_v) = \sum_{n=0}^\infty p_{v_n}$. As above, $M_{2^\infty} \cong p_{v_0}\mathfrak{I}_E p_{v_0}$, so let ρ be the unique tracial state and ρ_∞ be the induced trace function on $\mathcal{M}(\mathfrak{I}_E)_+$. We have that $\rho(p_{v_n}) = 2^{-n}$ so

$$\rho_\infty(1 - \sigma(p_v)) = \sum_{n=0}^\infty 2^{-n} < \infty.$$

It follows that $1 - \sigma(p_v)$ is not full, and thus $\mathfrak{e}$ is not nuclearly absorbing.

In particular, $C^*(E) \not\cong C^*(F)$.

Acknowledgement

I am indebted to Efren Ruiz for the above counterexample, as well as for other helpful comments.

References

[Choi and Effros 1976] M. D. Choi and E. G. Effros, "The completely positive lifting problem for C^*-algebras", *Ann. of Math.* (2) **104**:3 (1976), 585–609. MR 0417795 Zbl 0361.46067

[Eilers et al. 2014] S. Eilers, G. Restorff, and E. Ruiz, "The ordered K-theory of a full extension", *Canad. J. Math.* **66**:3 (2014), 596–625. MR 3194162 Zbl 1309.46038

[Elliott and Kucerovsky 2001] G. A. Elliott and D. Kucerovsky, "An abstract Voiculescu–Brown–Douglas–Fillmore absorption theorem", *Pacific J. Math.* **198**:2 (2001), 385–409. MR 1835515 Zbl 1058.46041

[Gabe and Ruiz 2015] J. Gabe and E. Ruiz, "Absorbing representations with respect to closed operator convex cones", preprint, 2015. arXiv 1503.07799

[Kirchberg 2000] E. Kirchberg, "Das nicht-kommutative Michael-Auswahlprinzip und die Klassifikation nicht-einfacher Algebren", pp. 92–141 in *C^*-algebras* (Münster, 1999), edited by J. Cuntz and S. Echterhoff, Springer, Berlin, 2000. MR 1796912 Zbl 0976.46051

[Kucerovsky and Ng 2006a] D. Kucerovsky and P. W. Ng, "The corona factorization property and approximate unitary equivalence", *Houston J. Math.* **32**:2 (2006), 531–550. MR 2219330 Zbl 1111.46050

[Kucerovsky and Ng 2006b] D. Kucerovsky and P. W. Ng, "S-regularity and the corona factorization property", *Math. Scand.* **99**:2 (2006), 204–216. MR 2289022 Zbl 1139.46037

[Robert 2011] L. Robert, "Nuclear dimension and n-comparison", *Münster J. Math.* **4** (2011), 65–72. MR 2869254 Zbl 1248.46040

[Rørdam 1991] M. Rørdam, "Ideals in the multiplier algebra of a stable C^*-algebra", *J. Operator Theory* **25**:2 (1991), 283–298. MR 1203034 Zbl 0785.46049

[Rørdam 2002] M. Rørdam, "Classification of nuclear, simple C^*-algebras", pp. 1–145 in *Classification of nuclear C^*-algebras. Entropy in operator algebras*, Encyclopaedia of Mathematical Sciences **126**, Springer, Berlin, 2002. MR 1878882 Zbl 1016.46037

[Ruiz and Tomforde 2014] E. Ruiz and M. Tomforde, "Ideals in graph algebras", *Algebr. Represent. Theory* **17**:3 (2014), 849–861. MR 3254772 Zbl 1316.46043

Received October 15, 2015. Revised February 9, 2016.

JAMES GABE
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF COPENHAGEN
UNIVERSITETSPARKEN 5
DK-2100 COPENHAGEN
DENMARK

jamiegabe123@hotmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.395

ON NONRADIAL SINGULAR SOLUTIONS OF SUPERCRITICAL BIHARMONIC EQUATIONS

ZONGMING GUO, JUNCHENG WEI AND WEN YANG

We develop a gluing method for fourth-order ODEs and construct infinitely many nonradial singular solutions for a biharmonic equation with supercritical exponent.

1. Introduction

In this paper we are concerned with positive singular solutions of the biharmonic equation

$$\Delta^2 u = u^p \quad \text{in } \mathbb{R}^n,\ n \geq 6, \tag{1-1}$$

where $p > (n+4)/(n-4)$.

Equation (1-1) arises in both physics and geometry. In recent decades there has been much research into classifying solutions to (1-1). When $1 < p \leq (n+4)/(n-4)$, all nonnegative solutions to (1-1) have been completely classified [Lin 1998; Wei and Xu 1999]: if $p < (n+4)/(n-4)$, then (1-1) admits no nontrivial nonnegative regular solution, while for $p = (n+4)/(n-4)$, i.e., the critical case, any positive regular solution of (1-1) can be written in the form

$$u_{\lambda,\xi} = \big(n(n-4)(n-2)(n+2)\big)^{-\frac{1}{8}(n-4)} \left(\frac{\lambda}{1+\lambda^2|x-\xi|^2}\right)^{\frac{1}{2}(n-4)}, \quad \xi \in \mathbb{R}^n.$$

However, the question of the complete classification of positive regular solutions of (1-1) in the supercritical case, i.e., $p > (n+4)/(n-4)$, remains largely open.

The structure of positive radial solutions of (1-1) with $p > (n+4)/(n-4)$ has been studied by Gazzola and Grunau [2006] and Guo and Wei [2010]. For the fourth-order ODE

$$\begin{cases} \Delta^2 u(r) = u^p(r), & r \in [0,\infty), \\ u(0) = a,\ u''(0) = b,\ u'(0) = u'''(0) = 0, \end{cases} \tag{1-2}$$

MSC2010: primary 35B40, 35J91; secondary 58J55.
Keywords: nonradial solutions, biharmonic supercritical equations, gluing method.

it is known from [Gazzola and Grunau 2006] that for any $a > 0$ there is a unique $b_0 := b_0(a) < 0$ such that the unique solution u_{a,b_0} of (1-2) satisfies $u_{a,b_0} \in C^4(0, \infty)$, $u'_{a,b_0}(r) < 0$ and

$$\lim_{r\to\infty} r^\alpha u_{a,b_0}(r) = K_0^{1/(p-1)},$$

where $\alpha = 4/(p-1)$ and

$$K_0 = \frac{8\big((n-2)(n-4)(p-1)^3+2(n^2-10n+20)(p-1)^2-16(n-4)(p-1)+32\big)}{(p-1)^4}.$$

This implies that $u_{a,b_0}(r) > 0$ for all $r > 0$ and $u_{a,b_0}(r) \to 0$ as $r \to \infty$. Moreover, it is known from [Guo and Wei 2010] that if $5 \le n \le 12$ or if $n \ge 13$ and $(n+4)/(n-4) < p < p_c(n)$, then $u_{a,b_0} - K_0^{1/(p-1)} r^{-\alpha}$ changes sign infinitely many times in $(0, \infty)$, and if $n \ge 13$ and $p \ge p_c(n)$, then $u(r) < K_0^{1/(p-1)} r^{-\alpha}$ for all $r > 0$ and the solutions are strictly ordered with respect to the initial value $a = u_{a,b_0}(0)$. Here $p_c(n)$ refers to the unique value of $p > (n+4)/(n-4)$ such that

$$p_c(n) = \begin{cases} +\infty & \text{if } 4 \le n \le 12, \\ \dfrac{n+2-\sqrt{n^2+4-n\sqrt{n^2-8n+32}}}{n-6-\sqrt{n^2+4-n\sqrt{n^2-8n+32}}} & \text{if } n \ge 13. \end{cases}$$

Very recently, Dávila, Dupaigne, Wang and Wei [Dávila et al. 2014] proved that all stable or finite Morse index solutions of (1-1) are trivial provided $1 < p < p_c(n)$. According to a result in [Guo and Wei 2010] and [Karageorgis 2009] all radial solutions are stable when $p \ge p_c(n)$. Thus the result in [Dávila et al. 2014] is sharp.

We now turn to the singular solutions of (1-1). It is easily seen that

$$u_s(x) := K_0^{1/(p-1)} |x|^{-4/(p-1)} \tag{1-3}$$

is a singular solution of (1-1). In other words, u_s satisfies the equation

$$\Delta^2 u = u^p, \quad u > 0 \text{ in } \mathbb{R}^n \setminus \{0\}. \tag{1-4}$$

As far as we know, the radial singular solution in (1-3) is the only singular solution to (1-4) known so far. The question we shall address in this paper is whether or not there are nonradial singular solutions to (1-4). To this end, we first discuss the corresponding second-order Lane–Emden equation

$$\Delta u + u^p = 0, \quad u > 0 \text{ in } \mathbb{R}^n, \tag{1-5}$$

which has been widely studied. We refer to [Budd and Norbury 1987; Bidaut-Véron and Véron 1991; Dancer et al. 2011; Farina 2007; Guo 2002; Gidas and Spruck 1981; Gui et al. 1992; Johnson et al. 1993; Joseph and Lundgren 1972/73; Korevaar et al. 1999; Zou 1995] and the references therein. Farina [2007] proved that if

$(n+2)/(n-2) < p < p^c(n)$, the Morse index of any regular solution u of (1-5) is ∞. Here $p^c(n)$ is the Joseph–Lundgren exponent [Joseph and Lundgren 1972/73]:

$$p^c(n) = \begin{cases} +\infty & \text{if } 2 \le n \le 10, \\ \dfrac{(n-2)^2 - 4n + 8\sqrt{n-1}}{(n-2)(n-10)} & \text{if } n \ge 11. \end{cases}$$

In [Dancer et al. 2011], Dancer, Du and Guo showed that if Ω_0 is a bounded domain containing 0, then u is a solution of (1-5) in $\Omega_0 \setminus \{0\}$; if u has finite Morse index and $(n+2)/(n-2) < p < p^c(n)$, then $x = 0$ must be a removable singularity of u. They also showed that if Ω_0 is a bounded domain containing 0, u is a solution of (1-5) in $\mathbb{R}^n \setminus \Omega_0$ that has finite Morse index, and $(n+2)/(n-2) < p < p^c(n)$, then u must be a fast decay solution. It is easily seen that (1-5) has a radial singular solution

$$u^s(x) := u^s(r) = \left(\frac{2}{p-1}\left(n-2-\frac{2}{p-1}\right)\right)^{1/(p-1)} |x|^{-2/(p-1)}.$$

Recently, Dancer, Guo and Wei [Dancer et al. 2012] obtained infinitely many positive nonradial singular solutions of (1-5) provided $p \in ((n+1)/(n-3),\, p^c(n-1))$. The proof of that result is via a gluing of outer and inner solutions.

The main result in this paper is the following theorem.

Theorem 1.1. *Let $n \ge 6$. Assume that*

$$\frac{n+3}{n-5} < p < p_c(n-1).$$

Then (1-1) *admits infinitely many nonradial singular solutions.*

The proof of Theorem 1.1 is via a gluing of inner and outer solutions, as in [Dancer et al. 2012]. In the second-order case, one glues $(u(r), u'(r))$ at some intermediate point. However, since (1-1) is of fourth order, we have to match the inner solution and outer solution up to the third derivative $(u(r), u'(r), u''(r), u'''(r))$. Some essential obstructions appear when matching the inner and outer solutions. As far as we know this is the first paper on gluing inner and outer solutions for fourth-order ODE problems.

In the following, we sketch the proof of Theorem 1.1. After performing a separation of variables for a solution u of (1-1), $u(x) = r^{-\alpha} w(\theta)$, finding a nonradial singular solution of (1-1) is equivalent to finding a nonconstant solution of the equation

$$\Delta^2_{S^{n-1}} w + k_1(n)\Delta_{S^{n-1}} w + k_0(n) w = w^p, \tag{1-6}$$

where

$$k_0(n) = (n-4-\alpha)(n-2-\alpha)(2+\alpha)\alpha,$$
$$k_1(n) = -\big((n-4-\alpha)(2+\alpha)+(n-2-\alpha)\alpha\big).$$

It is clear that $w(\theta) = (k_0(n))^{1/(p-1)}$ is the constant solution of (1-6), which provides the radial singular solution of (1-1) that is given in (1-3).

In order to construct positive nonradial singular solutions of (1-1), we need to find positive nonconstant solutions of (1-6), which is a fourth-order inhomogeneous nonlinear ODE; therefore, we shall construct infinitely many positive nonconstant radially symmetric solutions of (1-6), i.e., solutions that only depend on the geodesic distance $\theta \in [0, \pi)$. We only consider the simple case $w(\theta) = w(\pi - \theta)$ for $0 \le \theta \le \frac{\pi}{2}$. In this case, (1-6) can be written in the form

$$\begin{cases} T_1 w(\theta) + k_1(n) T_2 w(\theta) + k_0(n) w = w^p, & w(\theta) > 0,\ 0 < \theta < \frac{\pi}{2}, \\ w'(0),\, w'''(0) \text{ exist},\ w'\big(\frac{\pi}{2}\big) = w'''\big(\frac{\pi}{2}\big) = 0, \end{cases} \tag{1-7}$$

where T_1, T_2 are the differential operators defined by

$$T_1 w(\theta) = \frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d}{d\theta}\left(\frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta\frac{dw(\theta)}{d\theta}\right)\right)\right)$$

and

$$T_2 w(\theta) = \frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta\frac{dw(\theta)}{d\theta}\right).$$

A key observation is that

$$w_*(\theta) = A_p(\sin\theta)^{-\alpha}, \quad \theta \in \big(0, \tfrac{\pi}{2}\big], \tag{1-8}$$

with

$$A_p^{p-1} = (n-5-\alpha)(n-3-\alpha)(2+\alpha)\alpha \ (:= k_0(n-1)),$$

is a singular solution of (1-7) with a singular point at $\theta = 0$. (Note that this is a singular solution in one dimension less.) We will construct the inner and outer solutions of (1-7) and glue them at some point close to 0, which gives solutions of (1-7). The main difficulty is the matching of four parameters, which correspond to matching u and its derivatives up to the third order.

This paper is organized as follows. In Section 2, we present some preliminaries. In Section 3, we construct inner solutions of (1-7) by studying an initial value problem of (1-7) with large initial values at $\theta = 0$. In Section 4, we construct outer solutions of (1-7). We first study an initial value problem of (1-7) with the initial values at $\theta = \frac{\pi}{2}$, then we analyze the asymptotic behaviors of the solutions of this initial value problem near $\theta = 0$. Finally, in Section 5, we match the inner and outer solutions constructed in Sections 3 and 4 to obtain solutions of (1-1). This completes the proof of Theorem 1.1. We leave some computational results to the Appendix.

2. Preliminaries

In this section, we present some known results which will be used subsequently.

Let $u = u(r)$ be a positive radial solution of (1-1). Using the Emden–Fowler transformation

$$u(r) = r^{-\alpha} v(t), \quad t = \ln r, \tag{2-1}$$

we see that $v(t)$ satisfies the equation

$$v^{(4)}(t) + K_3 v'''(t) + K_2 v''(t) + K_1 v'(t) + K_0 v(t) = v^p(t), \quad t \in (-\infty, \infty), \tag{2-2}$$

where the coefficients K_0, K_1, K_2, K_3 are given in [Gazzola and Grunau 2006]:

$$\begin{aligned}
K_0 &= \frac{8}{(p-1)^4}\big((n-2)(n-4)(p-1)^3 + 2(n^2-10n+20)(p-1)^2 \\
&\qquad\qquad\qquad\qquad - 16(n-4)(p-1) + 32\big),\\
K_1 &= -\frac{2}{(p-1)^3}\big((n-2)(n-4)(p-1)^3 + 4(n^2-10n+20)(p-1)^2 \\
&\qquad\qquad\qquad\qquad - 48(n-4)(p-1) + 128\big),\\
K_2 &= \frac{1}{(p-1)^2}\big((n^2-10n+20)(p-1)^2 - 24(n-4)(p-1) + 96\big),\\
K_3 &= \frac{2}{p-1}\big((n-4)(p-1) - 8\big).
\end{aligned}$$

By direct calculation it is easy to see that $K_0 = k_0$. The characteristic polynomial (linearized at $K_0^{1/(p-1)}$) of (2-2) is

$$\nu \mapsto \nu^4 + K_3\nu^3 + K_2\nu^2 + K_1\nu + (1-p)K_0$$

and the eigenvalues are given by

$$\nu_1 = \frac{N_1 + \sqrt{N_2 + 4\sqrt{N_3}}}{2(p-1)}, \quad \nu_2 = \frac{N_1 - \sqrt{N_2 + 4\sqrt{N_3}}}{2(p-1)},$$

$$\nu_3 = \frac{N_1 + \sqrt{N_2 - 4\sqrt{N_3}}}{2(p-1)}, \quad \nu_4 = \frac{N_1 - \sqrt{N_2 - 4\sqrt{N_3}}}{2(p-1)},$$

where

$$\begin{aligned}
N_1 &:= -(n-4)(p-1) + 8,\\
N_2 &:= (n^2-4n+8)(p-1)^2,\\
N_3 &:= (9n-34)(n-2)(p-1)^4 + 8(3n-8)(n-6)(p-1)^3 \\
&\qquad\qquad + (16n^2 - 288n + 832)(p-1)^2 - 128(n-6)(p-1) + 256.
\end{aligned}$$

Let $\tilde{\nu}_j = \nu_j - \alpha$ for $j = 1, 2, 3, 4$.

Proposition 2.1 [Guo and Wei 2010]. *For any $n \geq 5$ and $p > (n+4)/(n-4)$,*

$$\tilde{\nu}_2 < 2-n < 0 < \tilde{\nu}_1. \tag{2-3}$$

(1) *For any $5 \leq n \leq 12$ or $n \geq 13$ and $(n+4)/(n-4) < p < p_c(n)$, we have $\tilde{\nu}_3, \tilde{\nu}_4 \notin \mathbb{R}$ and $\Re(\tilde{\nu}_3) = \Re(\tilde{\nu}_4) = \frac{1}{2}(4-n) < 0$.*

(2) *For any $n \geq 13$ and $p = p_c(n)$, we have $\tilde{\nu}_3 = \tilde{\nu}_4 = \frac{1}{2}(4-n)$.*

(3) *For any $n \geq 13$ and $p > p_c(n)$, we have*

$$\tilde{\nu}_2 < 4-n < \tilde{\nu}_4 < \tfrac{1}{2}(4-n) < \tilde{\nu}_3 < 0 < \tilde{\nu}_1, \quad \tilde{\nu}_3 + \tilde{\nu}_4 = 4-n. \tag{2-4}$$

Theorem 2.2 [Gazzola and Grunau 2006]. *For any $k \geq 1$,*

$$\lim_{t\to\infty} v(t) = K_0^{1/(p-1)}, \quad \lim_{t\to\infty} v^{(k)}(t) = 0 \tag{2-5}$$

Remark. We see that K_i $(i = 0, 1, 2, 3)$ and ν_j, $\tilde{\nu}_j$ $(j = 1, 2, 3, 4)$ above depend on n and p. In the following, by abuse of notation, we use K_i, ν_j, $\tilde{\nu}_j$ with the dimension n replaced by $n-1$ and write $k_0 = k_0(n)$ and $k_1 = k_1(n)$.

3. Inner solutions

In this section, we construct inner solutions of (1-7).

Let $Q \gg 1$ be a large constant and $\tilde{b}$ be a constant which will be given below. We consider the initial value problem

$$\begin{cases} T_1 w(\theta) + k_1 T_2 w(\theta) + k_0 w = w^p, \\ w(0) = Q,\ w'(0) = 0,\ w''(0) = (\tilde{b}+\mu) Q^{1+2/\alpha},\ w'''(0) = 0, \end{cases} \tag{3-1}$$

where $\mu > 0$ is a small constant. Since $Q \gg 1$, we set $Q = \epsilon^{-4/(p-1)}$ $(:= \epsilon^{-\alpha})$ with $\epsilon > 0$ sufficiently small.

Let $w(\theta) = \epsilon^{-\alpha} v(\theta/\epsilon)$. Then we have $v(0) = 1$, $v'(0) = 0$, $v''(0) = \tilde{b} + \mu$, $v'''(0) = 0$ and $v(r)$ (for $r = \theta/\epsilon$) satisfies the equation

$$\begin{aligned} v^{(4)}(r) &+ 2(n-2)\epsilon \cot(\epsilon r) v'''(r) \\ &+ \left((n-2)(n-4)\frac{\epsilon^2}{\sin^2(\epsilon r)} - (n-2)^2\epsilon^2 + k_1\epsilon^2\right) v'' \\ &+ \left((n-2)k_1\epsilon^3 \cot(\epsilon r) - (n-2)(n-4)\epsilon^3 \frac{\cot(\epsilon r)}{\sin^2(\epsilon r)}\right) v'(r) + k_0\epsilon^4 v(r) = v^p(r) \end{aligned} \tag{3-2}$$

with initial conditions

$$v(0) = 1, \quad v'(0) = 0, \quad v''(0) = \tilde{b} + \mu, \quad v'''(0) = 0.$$

For $\epsilon > 0$ sufficiently small, we have

$$\epsilon \cot(\epsilon r) = \frac{1}{r} - \tfrac{1}{3}\epsilon^2 r + \sum_{k=1}^{\infty} l_k \epsilon^{2k+2} r^{2k+1},$$

$$\epsilon^2 \sin^{-2}(\epsilon r) = \frac{1}{r^2} + \tfrac{1}{3}\epsilon^2 + \sum_{k=1}^{\infty} m_k \epsilon^{2k+2} r^{2k},$$

$$\epsilon^3 \cot(\epsilon r) \sin^{-2}(\epsilon r) = \frac{1}{r^3} + \sum_{k=1}^{\infty} n_k \epsilon^{2k+2} r^{2k-1}.$$

So (3-2) can be written in the form

(3-3)
$$\begin{aligned} v^{(4)}(r) &+ \left(\frac{2(n-2)}{r} - \tfrac{2}{3}(n-2)\epsilon^2 r + \sum_{k=1}^{\infty} l'_k \epsilon^{2k+2} r^{2k+1}\right) v'''(r) \\ &+ \left(\frac{(n-2)(n-4)}{r^2} + \left(\tfrac{1}{3}(n-2)(n-4) - (n-2)^2 + k_1\right)\epsilon^2 + \sum_{k=1}^{\infty} m'_k \epsilon^{2k+2} r^{2k}\right) v''(r) \\ &- \left(\frac{(n-2)(n-4)}{r^3} - (n-2)k_1 r^{-1}\epsilon^2 + \sum_{k=1}^{\infty} n'_k \epsilon^{2k+2} r^{2k-1}\right) v'(r) + k_0 \epsilon^4 v(r) = v^p(r) \end{aligned}$$

with initial conditions

$$v(0) = 1, \quad v''(0) = \tilde{b} + \mu, \quad v'(0) = v'''(0) = 0.$$

The first approximation to the solution of (3-3) is the radial solution $v_0(r)$ of the problem

$$\text{(3-4)} \qquad \Delta^2 v = v^p \text{ in } \mathbb{R}^{n-1}, \quad v(0) = 1,\ v'(0) = 0,\ v''(0) = \tilde{b} + \mu,\ v'''(0) = 0.$$

We write $v_0 = v_{01} + v_{02}$, where v_{01} satisfies

$$\text{(3-5)} \qquad \Delta^2 v = v^p, \quad v(0) = 1,\ v'(0) = 0,\ v''(0) = \tilde{b},\ v'''(0) = 0,$$

and v_{02} satisfies

$$\text{(3-6)} \qquad \Delta^2 v = v_0^p - v_{01}^p, \quad v(0) = 0,\ v'(0) = 0,\ v''(0) = \mu,\ v'''(0) = 0.$$

We now choose $\tilde{b} < 0$ to be the unique value such that the solution v_{01} is the unique positive radial ground state of (3-5).

Lemma 3.1. *Assume that $v_{01}(r)$ and $v_{02}(r)$ are the solutions to* (3-5) *and* (3-6), *respectively. For $(n+3)/(n-5) < p < p_c(n-1)$, there exists $R_0 \gg 1$ such that for $r \geq R_0$, the solution $v_{01}(r)$ satisfies*

$$\text{(3-7)} \qquad v_{01}(r) = A_p r^{-\alpha} + \frac{a_0 \cos(\beta \ln r) + b_0 \sin(\beta \ln r)}{r^{(n-5)/2}} + O(r^{2\sigma-\alpha}),$$

where $\beta = \sqrt{4\sqrt{N_3} - N_2}/(2(p-1))$ (*with* n *being replaced by* $n-1$ *in* N_2 *and* N_3) *and* $\sqrt{a_0^2 + b_0^2} \neq 0$.

The solution $v_{02}(r)$ *satisfies*

$$v_{02}(r) = \mu B_p r^{\tilde{\nu}_1} + O(\mu^2 r^{\nu_1 + \tilde{\nu}_1} + \mu r^{\tilde{\nu}_1 + \alpha - (n-5)/2}), \tag{3-8}$$

with $B_p \neq 0$ *when* $\mu = O(1/(r^{\nu_1 - \sigma}))$ *for* r *in any interval* $[e^T, e^{10T}]$ *with* $T \gg 1$ *and* $\sigma = \alpha - \frac{1}{2}(n-5)$.

Proof. The proof of this lemma is divided into two steps. We consider $v_{01}(r)$ in the first step. The main arguments in the proof are similar to those in the proof of Theorem 3.1 of [Guo 2014].

Using the Emden–Fowler transformation

$$v_{01}(r) = r^{-\alpha} v(t), \quad t = \ln r \quad (r > 0), \tag{3-9}$$

and letting $v(t) = A_p - h(t)$, we see that $h(t)$ satisfies

$$h^{(4)}(t) + K_3 h'''(t) + K_2 h''(t) + K_1 h'(t) + (1-p) K_0 h(t) + O(h^2) = 0 \tag{3-10}$$

for $t > 1$. Note that $r^\alpha v_{01}(r) \to A_p$ as $r \to \infty$ and hence $h(t) \to 0$ as $t \to \infty$. It follows from Proposition 2.1 that $\tilde{\nu}_3, \tilde{\nu}_4 \notin \mathbb{R}$ and $\Re(\tilde{\nu}_3) = \Re(\tilde{\nu}_4) = \frac{1}{2}(5-n) < 0$ and $\tilde{\nu}_2 < 3-n < 0 < \tilde{\nu}_1$ provided $(n+3)/(n-5) < p < p_c(n-1)$. Let $\nu_3 = \sigma + i\beta$, where $\beta = \sqrt{4\sqrt{N_3} - N_2}/(2(p-1))$ and $\sigma = -\frac{1}{2}(n-5) + \alpha < 0$ for $p > (n+3)/(n-5)$.

We can write (3-10) as

$$(\partial_t - \nu_4)(\partial_t - \nu_3)(\partial_t - \nu_2)(\partial_t - \nu_1) h(t) = H(h(t)), \tag{3-11}$$

where $H(h(t)) = O(h^2)$. We claim that for any $T \gg 1$, there exist constants A_i and B_i $(i = 1, 2, 3, 4)$ such that

$$\begin{aligned} h(t) = {} & A_1 e^{\sigma t} \cos \beta t + A_2 e^{\sigma t} \sin \beta t + A_3 e^{\nu_2 t} + A_4 e^{\nu_1 t} \\ & + B_1 \int_T^t e^{\sigma(t-s)} \sin \beta(t-s) H(h(s))\, \mathrm{d}s + B_2 \int_T^t e^{\sigma(t-s)} \cos \beta(t-s) H(h(s))\, \mathrm{d}s \\ & + B_3 \int_T^t e^{\nu_2(t-s)} H(h(s))\, \mathrm{d}s + B_4 \int_T^t e^{\nu_1(t-s)} H(h(s))\, \mathrm{d}s. \end{aligned}$$

Moreover, each A_i depends on T and ν_i $(i = 1, 2, 3, 4)$, while each B_i depends only on ν_i $(i = 1, 2, 3, 4)$. In fact, it follows from (3-11) and the theory of second-order ODEs (see [Hartman 1982]) that

$$\begin{aligned} & (\partial_t - \nu_2)(\partial_t - \nu_1) h(t) \\ & \quad = A_1' e^{\sigma t} \cos \beta t + A_2' e^{\sigma t} \sin \beta t + \frac{1}{\beta} \int_T^t e^{\sigma(t-s)} \sin \beta(t-s) H(h(s))\, \mathrm{d}s, \end{aligned} \tag{3-12}$$

where A_1' and A_2' are constants depending on T, ν_3 and ν_4. Multiplying both sides of (3-12) by $e^{-\nu_2 t}$ and integrating it from T to t, we obtain

$$(\partial_t - \nu_1)h(t) = A_3' e^{\nu_2 t} + \int_T^t e^{\nu_2(t-s)}(A_1' e^{\sigma s}\cos\beta s + A_2' e^{\sigma s}\sin\beta s)\,\mathrm{d}s \\ + \frac{1}{\beta}\int_T^t e^{\nu_2(t-s)}\int_T^s e^{\sigma(s-\xi)}\sin\beta(s-\xi)H(h(\xi))\,\mathrm{d}\xi\,\mathrm{d}s.$$

We now switch the order of integration and find that

$$(\partial_t - \nu_1)h(t) \\ = A_1'' e^{\sigma t}\cos\beta t + A_2'' e^{\sigma t}\sin\beta t + A_3'' e^{\nu_2 t} + B_1' \int_T^t e^{\sigma(t-s)}\sin\beta(t-s)H(h(s))\,\mathrm{d}s \\ + B_2' \int_T^t e^{\sigma(t-s)}\cos\beta(t-s)H(h(s))\,\mathrm{d}s + B_3' \int_T^t e^{\nu_2(t-s)}H(h(s))\,\mathrm{d}s,$$

where A_1'', A_2'' and A_3'' depend on T and ν_i ($i = 2, 3, 4$), and where the B_i' ($i = 1, 2, 3$) depend only on ν_i ($i = 2, 3, 4$). Repeating the same argument once again, we obtain our claim. Using the fact that $\int_T^t = \int_T^\infty - \int_t^\infty$, we have

$$B_4 \int_T^t e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s = B_4 \int_T^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s - B_4 \int_t^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s \\ = B_4 e^{\nu_1 t}\int_T^\infty e^{-\nu_1 s}H(h(s))\,\mathrm{d}s - B_4 \int_t^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s.$$

By combining $B_4 e^{\nu_1 t}\int_T^\infty e^{-\nu_1 s}H(h(s))\,\mathrm{d}s$ and $A_4 e^{\nu_1 t}$, we can also write $h(t)$ as

$$h(t) = A_1 e^{\sigma t}\cos\beta t + A_2 e^{\sigma t}\sin\beta t + A_3 e^{\nu_2 t} + M_4 e^{\nu_1 t} \\ + B_1 \int_T^t e^{\sigma(t-s)}\sin\beta(t-s)H(h(s))\,\mathrm{d}s \\ + B_2 \int_T^t e^{\sigma(t-s)}\cos\beta(t-s)H(h(s))\,\mathrm{d}s \\ + B_3 \int_T^t e^{\nu_2(t-s)}H(h(s))\,\mathrm{d}s - B_4 \int_t^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s.$$

Since $h(t) \to 0$ as $t \to \infty$, we have $M_4 = 0$ (note $\nu_1 > 0$). Setting

$$h_1(t) = A_1 e^{\sigma t}\cos\beta t + A_2 e^{\sigma t}\sin\beta t + A_3 e^{\nu_2 t}$$

and

$$h_2(t) = B_1 \int_T^t e^{\sigma(t-s)}\sin\beta(t-s)H(h(s))\,\mathrm{d}s + B_2 \int_T^t e^{\sigma(t-s)}\cos\beta(t-s)H(h(s))\,\mathrm{d}s \\ + B_3 \int_T^t e^{\nu_2(t-s)}H(h(s))\,\mathrm{d}s - B_4 \int_t^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s$$

and noting that $H(h(t)) = O(h^2(t))$, we see that

$$|h_2(t)| \le C(\tilde{h}_1(t) + \tilde{h}_2(t)), \tag{3-13}$$

where $C > 0$ is independent of T and

$$\tilde{h}_1(t) = \max\left\{\int_T^t e^{\sigma(t-s)}|h_1(s)|^2\,\mathrm{d}s,\ \int_T^t e^{\nu_2(t-s)}|h_1(s)|^2\,\mathrm{d}s,\ \int_t^\infty e^{\nu_1(t-s)}|h_1(s)|^2\,\mathrm{d}s\right\},$$

$$\tilde{h}_2(t) = \max\left\{\int_T^t e^{\sigma(t-s)}|h_2(s)|^2\,\mathrm{d}s,\ \int_T^t e^{\nu_2(t-s)}|h_2(s)|^2\,\mathrm{d}s,\ \int_t^\infty e^{\nu_1(t-s)}|h_2(s)|^2\,\mathrm{d}s\right\}.$$

We now show

$$|h_2(t)| = o(e^{\sigma t}). \tag{3-14}$$

There are three cases to be considered:

(1) $|h_2(t)| \le \left(\tilde{h}_1(t) + \int_T^t e^{\sigma(t-s)}|h_2(s)|^2\,\mathrm{d}s\right)$,

(2) $|h_2(t)| \le C\left(\tilde{h}_1(t) + \int_T^t e^{\nu_2(t-s)}|h_2(s)|^2\,\mathrm{d}s\right)$,

(3) $|h_2(t)| \le C\left(\tilde{h}_1(t) + \int_t^\infty e^{\nu_1(t-s)}|h_2(s)|^2\,\mathrm{d}s\right)$.

We only consider cases (1) and (3); case (2) is similar. For case (1), we have

$$|h_2(t)| \le C\left(\tilde{h}_1(t) + \int_T^t e^{\sigma(t-s)}|h_2(s)|^2\,\mathrm{d}s\right). \tag{3-15}$$

Thus,

$$|h_2(t)| \le C\left(\tilde{h}_1(t) + \max_{t\ge T}|h_2(t)|\int_T^t e^{\sigma(t-s)}|h_2(s)|\,\mathrm{d}s\right). \tag{3-16}$$

Let $m(t) = \int_T^t e^{-\sigma s}|h_2(s)|\,\mathrm{d}s$. Then it can be seen from (3-16) that

$$m'(t) \le C\tilde{h}_1(t)e^{-\sigma t} + C\max_{t\ge T}|h_2(t)|m(t). \tag{3-17}$$

For any $\epsilon > 0$ sufficiently small, we can choose T sufficiently large so that $0 < d_T := C\max_{t\ge T}|h_2(t)| < \epsilon$. It follows from (3-17) that

$$m(t) \le Ce^{d_T t}\int_T^t \tilde{h}_1(s)e^{-\sigma s}e^{-d_T s}\,\mathrm{d}s. \tag{3-18}$$

Substituting $m(t)$ in (3-18) into (3-16), we see that

$$|h_2(t)| \le C\tilde{h}_1(t) + Cd_T e^{(\sigma+d_T)t}\int_T^t \tilde{h}_1(s)e^{-\sigma s}e^{-d_T s}\,\mathrm{d}s. \tag{3-19}$$

Note that $\sigma + d_T < 0$ for T sufficiently large. We can combine $\nu_2 < \sigma$ with $h_1(t) = O(e^{\sigma t})$ to get $\tilde{h}_1(t) = o(e^{\sigma t})$. On the other hand, from (3-19) we can obtain that $|h_2(t)| = o(e^{(\sigma + d_T)t})$. Substituting these into (3-15), we eventually have

$$|h_2(t)| = o(e^{\sigma t}). \tag{3-20}$$

For case (3), we have

$$|h_2(t)| \le C\biggl(\tilde{h}_1(t) + \int_t^\infty e^{\nu_1(t-s)}|h_2(s)|^2\,\mathrm{d}s\biggr). \tag{3-21}$$

Thus,

$$|h_2(t)| \le C\tilde{h}_1(t) + C\max_{t\ge T}|h_2(t)|\int_t^\infty e^{\nu_1(t-s)}|h_2(s)|\,\mathrm{d}s. \tag{3-22}$$

Letting $l(t) = \int_t^\infty e^{-\nu_1 s}|h_2(s)|\,\mathrm{d}s$, we see from (3-22) that

$$-l'(t) \le C\tilde{h}_1(t)e^{-\nu_1 t} + d_T l(t). \tag{3-23}$$

It follows from (3-23) that

$$l(s) \le Ce^{-d_T t}\int_t^\infty \tilde{h}_1(s)e^{-\nu_1 s}e^{d_T s}\,\mathrm{d}s. \tag{3-24}$$

Since $\tilde{h}_1(t) = o(e^{\sigma t})$, we obtain from (3-24) that

$$l(s) = o(e^{(\sigma - \nu_1)t}).$$

Substituting this into (3-22), we also have

$$|h_2(t)| = o(e^{\sigma t}).$$

We now write $h(t)$ as

$$\begin{aligned}h(t) = M_1 e^{\sigma t}\cos\beta t + M_2 e^{\sigma t}\sin\beta t + A_3 e^{\nu_2 t}&\\ - B_1\int_t^\infty e^{\sigma(t-s)}\sin\beta(t-s)H(h(s))\,\mathrm{d}s&\\ - B_2\int_t^\infty e^{\sigma(t-s)}\cos\beta(t-s)H(h(s))\,\mathrm{d}s&\\ + B_3\int_T^t e^{\nu_2(t-s)}H(h(s))\,\mathrm{d}s - B_4\int_t^\infty e^{\nu_1(t-s)}H(h(s))\,\mathrm{d}s&.\end{aligned}$$

Then, it follows from $H(h(t)) = O(h^2(t))$, $h_1(t) = O(e^{\sigma t})$, $h_2(t) = o(e^{\sigma t})$ and $\nu_2 < 2\sigma$ that

$$h(t) = M_1 e^{\sigma t}\cos(\beta t) + M_2 e^{\sigma t}\sin(\beta t) + A_3 e^{\nu_2 t} + O(e^{2\sigma t}). \tag{3-25}$$

This implies that (3-7) holds for some a_0 and b_0. By an argument similar to the one used in the proof of [Guo and Wei 2010, Theorem 3.3], we can show $a_0^2+b_0^2\neq 0$. This completes the proof of the first step.

We now proceed to the second step. Setting $v_{02}=\mu\tilde{v}_{02}$, we see that $\tilde{v}_{02}(r)$ satisfies

$$\Delta^2\tilde{v}_{02}-pv_{01}^{p-1}\tilde{v}_{02}=\mu^{-1}\big((v_{01}+\mu\tilde{v}_{02})^p-v_{01}^p-p\mu v_{01}^{p-1}\tilde{v}_{02}\big) \tag{3-26}$$

with initial conditions

$$\tilde{v}_{02}(0)=0,\quad \tilde{v}_{02}'(0)=0,\quad \tilde{v}_{02}''(0)=1,\quad \tilde{v}_{02}'''(0)=0.$$

Using the Emden–Fowler transformation

$$\tilde{v}_{02}(r)=r^{-\alpha}\hat{v}(t),\quad t=\ln r\quad (r>0),$$

and the expression obtained for $v_{01}(r)$, we see that $\hat{v}(t)$ satisfies

$$\hat{v}^{(4)}+K_3\hat{v}'''+K_2\hat{v}''+K_1\hat{v}'+(1-p)K_0\hat{v}=f(r,\mu,\hat{v}), \tag{3-27}$$

where

$$f(r,\mu,\hat{v})=O(\mu\hat{v}+r^{\alpha-(n-5)/2})\hat{v}$$

provided that $\mu\hat{v}=o(1)$ for t sufficiently large. It follows from (3-27) that

$$\begin{aligned}\hat{v}(t)=\hat{A}_1e^{\sigma t}\cos\beta t+\hat{A}_2e^{\sigma t}\sin\beta t+\hat{A}_3e^{\nu_2 t}+\hat{A}_4e^{\nu_1 t}\\ +\hat{B}_1\int_T^t e^{\sigma(t-s)}\sin\beta(t-s)f(r,\mu,\hat{v}(s))\,\mathrm{d}s\\ +\hat{B}_2\int_T^t e^{\sigma(t-s)}\cos\beta(t-s)f(r,\mu,\hat{v}(s))\,\mathrm{d}s\\ +\hat{B}_3\int_T^t e^{\nu_2(t-s)}f(r,\mu,\hat{v}(s))\,\mathrm{d}s+\hat{B}_4\int_T^t e^{\nu_1(t-s)}f(r,\mu,\hat{v}(s))\,\mathrm{d}s,\end{aligned}$$

where $\hat{A}_i=\hat{A}_i(T,\nu_1,\nu_2,\nu_3,\nu_4)$ $(i=1,2,3,4)$ and $\hat{B}_i=\hat{B}_i(\nu_1,\nu_2,\nu_3,\nu_4)$. We first show that $\tilde{v}_{02}$ is strictly increasing in $(0,\infty)$. Using the initial values, we can find $R\in(0,\infty)$ such that $\tilde{v}_{02}(r)>0$ for $r\in(0,R)$. Writing (3-26) as

$$\mu\Delta^2\tilde{v}_{02}=(v_{01}+\mu\tilde{v}_{02})^p-v_{01}^p,$$

we obtain that $(\Delta\tilde{v}_{02})'>0$, and hence $\Delta\tilde{v}_{02}>\Delta\tilde{v}_{02}(0)=n-1$ for $r\in(0,R)$, which implies that $(\tilde{v}_{02})'(r)>0$ for $r\in(0,R)$. Moreover, we can deduce that $R=\infty$ and $\tilde{v}_{02}'(r)>0$ for $r\in(0,\infty)$. Therefore, $\hat{v}$ is increasing in $(0,\infty)$. Next, we claim that $\hat{A}_4 \neq 0$ for any $T \gg 1$ sufficiently large. Indeed, for $t \in [T, 10T]$,

$$\begin{aligned}
e^{-\nu_1 t}\hat{v}(t) &= \hat{A}_4 + \tilde{g}(t) + \hat{B}_1 e^{(\sigma-\nu_1)t}\int_T^t e^{-\sigma s}\sin\beta(t-s)\, f(r,\mu,\hat{v}(s))\,\mathrm{d}s \\
&\qquad + \hat{B}_2 e^{(\sigma-\nu_1)t}\int_T^t e^{-\sigma s}\cos\beta(t-s)\, f(r,\mu,\hat{v}(s))\,\mathrm{d}s \\
&\qquad + \hat{B}_3 e^{(\nu_2-\nu_1)t}\int_T^t e^{-\nu_2 s} f(r,\mu,\hat{v}(s))\,\mathrm{d}s + \hat{B}_4\int_T^t e^{-\nu_1 s} f(r,\mu,\hat{v}(s))\,\mathrm{d}s \\
&\le |\hat{A}_4| + |\tilde{g}(t)| + \left(\sum_{j=1}^4 |\hat{B}_j|\right)\max_{t\in[T,10T]}(\mu\hat{v} + e^{(\alpha-(n-5)/2)t})\int_T^t e^{-\nu_1 s}\hat{v}(s)\,\mathrm{d}s,
\end{aligned}$$

where

$$\tilde{g}(t) = \hat{A}_1 e^{(\sigma-\nu_1)t}\cos\beta t + \hat{A}_2 e^{(\sigma-\nu_1)t}\sin\beta t + \hat{A}_3 e^{(\nu_2-\nu_1)t}.$$

Since

$$\left(\sum_{j=1}^4 |\hat{B}_j|\right)\max_{t\in[T,10T]}(\mu\hat{v} + e^{(\alpha-(n-5)/2)t}) = \tau = o(1),$$

we have

$$e^{-\nu_1 t}\hat{v}(t) \le |\hat{A}_4| + |\tilde{g}(t)| + \tau\int_T^t e^{-\nu_1 s}\hat{v}(s)\,\mathrm{d}s. \tag{3-28}$$

Let $\ell(t) = \int_T^t e^{-\nu_1 s}\hat{v}(s)\,\mathrm{d}s$. We see that

$$(e^{-\tau t}\ell(t))' \le (|\hat{A}_4| + |\tilde{g}(t)|)e^{-\tau t}. \tag{3-29}$$

Integrating (3-29) in $[T, t]$, we obtain

$$\ell(t) \le \frac{|\hat{A}_4| + \max_{t\in[T,10T]}|\tilde{g}(t)|}{\tau} e^{\tau(t-T)}.$$

If we choose $\tau(t-T) \le C$ for $t \in [T, 10T]$, i.e., $\tau = O(1/T)$, we see that

$$\ell(t) \le \frac{(|\hat{A}_4| + \max_{t\in[T,10T]}|\tilde{g}(t)|)C}{\tau}. \tag{3-30}$$

Substituting this into (3-28), we have

$$e^{-\nu_1 t}\hat{v}(t) \le |\hat{A}_4|(1+C) + |\tilde{g}(t)| + C\max_{t\in[T,10T]}|\tilde{g}(t)|. \tag{3-31}$$

Suppose $\hat{A}_4 = 0$. We see from (3-31) and the expression of $|\tilde{g}(t)|$ that

$$\hat{v}(t) = o(1) \quad \text{for all } t \in [T, 10T].$$

This contradicts the fact that $\hat{v}$ is increasing in $(0, \infty)$. Therefore, $\hat{A}_4 \neq 0$ and our claim holds. Moreover, it is known from (3-31) and the expression of $\hat{v}(t)$ that

$$\hat{v}(t) = B_p e^{\nu_1 t} + O(\mu e^{2\nu_1 t} + e^{(\sigma+\nu_1)t}) \tag{3-32}$$

with $B_p \neq 0$ and $\mu = O(e^{(-\nu_1+\sigma)t})$. Therefore,

$$v_{02}(r) = \mu B_p r^{\tilde{\nu}_1} + O(\mu^2 r^{\nu_1+\tilde{\nu}_1} + \mu r^{\tilde{\nu}_1+\sigma})$$

with $B_p \neq 0$ and $\mu = O(1/r^{\nu_1-\sigma})$. □

Lemma 3.2. *Let p satisfy the conditions of Lemma 3.1 and $v_1(r)$ be the unique solution of the equation*

$$\begin{cases} v_1^{(4)}(r) + \dfrac{2(n-2)}{r} v_1'''(r) + \dfrac{(n-2)(n-4)}{r^2} v_1''(r) - \dfrac{(n-2)(n-4)}{r^3} v_1'(r) \\ \qquad - \frac{2}{3}(n-2) r v_0'''(r) + \big(\frac{1}{3}(n-2)(n-4) - (n-2)^2 + k_1\big) v_0''(r) \\ \qquad\qquad + \dfrac{(n-2)k_1}{r} v_0'(r) = p v_0^{p-1}(r) v_1(r), \\ v_1(0) = 0,\ v_1'(0) = 0,\ v_1''(0) = 0,\ v_1'''(0) = 0. \end{cases} \tag{3-33}$$

Then for $r \in [e^T, e^{10T}]$ with $T \gg 1$ and $\mu = O(1/r^{\nu_1-\sigma})$,

$$\begin{aligned} v_1(r) = C_p r^{2-\alpha} &+ r^{2-(n-5)/2}(a_1 \cos(\beta \ln r) + b_1 \sin(\beta \ln r)) \\ &+ \mu D_p r^{2+\tilde{\nu}_1} + O(\mu^2 r^{\tilde{\nu}_1+\nu_1+2} + \mu r^{\tilde{\nu}_1+\sigma+2}) + o(r^{2-(n-5)/2}), \end{aligned} \tag{3-34}$$

where C_p satisfies

$$E_1 C_p - p A_p^{p-1} C_p = F_1 A_p, \tag{3-35}$$

with

$$\begin{aligned} E_1 &= (1+\alpha)(1-\alpha)(2-\alpha)\alpha - 2(n-2)(2-\alpha)(1-\alpha)\alpha \\ &\qquad - (n-2)(n-4)(2-\alpha) + (n-2)(n-4)(2-\alpha)(1-\alpha), \\ F_1 &= \big((n-2)^2 - k_1 - \tfrac{1}{3}(n-2)(n-4)\big)\alpha(\alpha+1) \\ &\qquad - \tfrac{2}{3}(n-2)\alpha(\alpha+1)(\alpha+2) + k_1(n-2)\alpha, \end{aligned}$$

and where D_p satisfies

$$E_2 D_p = F_2 B_p, \tag{3-36}$$

with

$$E_2 = (2+\tilde\nu_1)(\tilde\nu_1+n-1)(\tilde\nu_1+n-3)\tilde\nu_1 - pA_p^{p-1},$$

$$F_2 = \tfrac{2}{3}(n-2)(\tilde\nu_1-1)(\tilde\nu_1-2)\tilde\nu_1 + \big((n-2)^2-k_1-\tfrac{1}{3}(n-2)(n-4)\big)(\tilde\nu_1-1)\tilde\nu_1 - k_1(n-2)\tilde\nu_1 + p(p-1)A_p^{p-2}C_p,$$

and where (a_1,b_1) *is the solution of*

$$\begin{cases} Aa_1 - Bb_1 = G, \\ Ba_1 + Ab_1 = H, \end{cases}$$

with

$$A = \tfrac{1}{16}(n^4-12n^3+14n^2+132n-135) - pA_p^{p-1} + \tfrac{1}{2}(n^2-6n-35)\beta^2+\beta^4,$$

$$B = (2n^2-12n-6)\beta + 8\beta^3,$$

$$\begin{aligned} G = p(p-1)A_p^{p-2}C_pa_0 &+ \tfrac{1}{12}(n^4-11n^3+41n^2-61n+30)a_0 \\ &+ \tfrac{1}{4}(n^2-6n+5)k_1a_0 + \tfrac{1}{6}(4n^2+3n-n^3-14)b_0\beta - 2k_1b_0\beta \\ &+ \tfrac{1}{3}(n^2-9n+14)a_0\beta^2 + a_0k_1\beta^2 - \tfrac{2}{3}(n-2)b_0\beta^3, \end{aligned}$$

$$\begin{aligned} H = p(p-1)A_p^{p-2}C_pb_0 &+ \tfrac{1}{12}(n^4-11n^3+41n^2-61n+30)b_0 \\ &+ \tfrac{1}{4}(n^2-6n+5)k_1b_0 - \tfrac{1}{6}(4n^2+3n-n^3-14)a_0\beta + 2k_1a_0\beta \\ &+ \tfrac{1}{3}(n^2-9n+14)b_0\beta^2 + b_0k_1\beta^2 + \tfrac{2}{3}(n-2)a_0\beta^3. \end{aligned}$$

Remark. We need to show that $E_2 \neq 0$ and that the 2×2 matrix $K = \left[\begin{smallmatrix} A & -B \\ B & A \end{smallmatrix}\right]$ is invertible. This will be proved in the Appendix.

Proof. The uniqueness of solutions to (3-33) follows from standard ODE theory since all the initial conditions are zero and the inhomogeneous term is locally Lipschitz. Analyzing the terms which contain v_0 in (3-33) and using the Taylor expansion for v_0^{p-1} for $r \in [e^T, e^{10T}]$, after direct computation we can find the leading terms which are of the orders

$$r^{-2-\alpha}, \quad r^{(1-n)/2}\cos(\beta\ln r), \quad r^{(1-n)/2}\sin(\beta\ln r), \quad \mu r^{\tilde\nu_1-2}.$$

By the above observation, we can assume

$$v_1(r) = C_pr^{2-\alpha} + \tilde f(r)r^{2-(n-5)/2} + \mu D_pr^{2+\tilde\nu_1} + o(r^{2-(n-5)/2}) + O(\mu^2r^{\tilde\nu_1+\nu_1+2} + \mu r^{\tilde\nu_1+\sigma+2}),$$

where

$$\tilde f(r) = a_1\cos(\beta\ln r) + b_1\sin(\beta\ln r).$$

Using (3-7) and (3-8), we can get C_p, D_p, a_1 and b_1 by direct calculation. □

Furthermore, we can obtain the following proposition.

Proposition 3.3. *Let*

$$\frac{n+3}{n-5} < p < p_c(n-1)$$

and $v(r)$ *be a solution of* (3-2). *Then for* $\epsilon > 0$ *sufficiently small*,

$$v(r) = v_0(r) + \sum_{k=1}^{\infty} \epsilon^{2k} v_k(r).$$

Moreover, for $r \in [e^T, e^{10T}]$ *with* $T \gg 1$ *and* $\mu = O(1/r^{\nu_1-\sigma})$,

$$\begin{aligned}(3\text{-}37)\quad v_k(r) = \sum_{j=1}^{k} d_j^k r^{2j-\alpha} + \sum_{j=1}^{k} e_j^k r^{2j-(n-5)/2} \sin(\beta \ln r + E_j^k) + \sum_{j=1}^{k} \mu f_j^k r^{2j+\tilde{\nu}_1} \\ + O(\mu^2 r^{\tilde{\nu}_1+\nu_1+2k} + \mu r^{\tilde{\nu}_1+\sigma+2k}) + o(r^{2k-(n-5)/2}),\end{aligned}$$

where $d_j^k, e_j^k, f_j^k, E_j^k$ $(j = 1, 2, \dots, k)$ *are constants. Moreover*,

$$d_1^1 = C_p, \quad e_1^1 = \sqrt{a_1^2 + b_1^2}, \quad f_1^1 = D_p, \quad \sin E_1^1 = a_1/e_1^1, \quad \cos E_1^1 = b_1/e_1^1,$$

where C_p, a_1, b_1, D_p *are given in Lemma 3.2.*

Proof. Substituting

$$v(r) = v_0(r) + \sum_{i=1}^{\infty} \epsilon^{2i} v_i(r)$$

into (3-3), we expand (3-3) according to the order of ϵ. Considering the constant order and the ϵ^2 order, we get (3-4) and (3-33), respectively. We note that only the terms $v_0, v_1, \dots, v_k$ carry ϵ^{2k}. Suppose we have found v_{k-1}. Then we can determine v_k by studying the equation of order ϵ^{2k} in (3-3), i.e.,

$$\begin{cases} v_k^{(4)}(r) + \dfrac{2(n-2)}{r} v_k'''(r) + \dfrac{(n-2)(n-4)}{r^2} v_k''(r) - \dfrac{(n-2)(n-4)}{r^3} v_k'(r) \\ \qquad - \frac{2}{3}(n-2) r v_{k-1}'''(r) + \big(\frac{1}{3}(n-2)(n-4) - (n-2)^2 + k_1\big) v_{k-1}''(r) \\ \qquad\qquad + \dfrac{(n-2)k_1}{r} v_{k-1}'(r) + \displaystyle\sum_{i=1}^{k-1} \big(l_i' r^{2i+1} v_{k-i-1}'''(r) + m_i' r^{2i} v_{k-i-1}''(r) \\ \qquad\qquad\qquad + n_i' r^{2i-1} v_{k-i-1}'(r)\big) + k_0 v_{k-1}(r) = \dfrac{d^k}{dt^k} \Big(\displaystyle\sum_{i=0}^{k} t^i v_i \Big)^p \Big|_{t=0}, \\ v_k(0) = 0, \ v_k'(0) = 0, \ v_k''(0) = 0, \ v_k'''(0) = 0, \end{cases}$$

where l_i', m_i', n_i' are given in (3-3). Following our arguments in Lemma 3.2, we find the leading order of the terms involving $v_0, v_1, \dots, v_{k-1}$ in the above equation, and then we assume v_k has the expansion in (3-37). By substituting (3-37) into the equation of order ϵ^{2k} and comparing each order, we can compute the terms $d_j^k, e_j^k, f_j^k, E_j^k$ $(j = 1, 2, \ldots, k)$. □

Theorem 3.4. *Let*

$$\frac{n+3}{n-5} < p < p_c(n-1)$$

and $w_{\epsilon,\mu}^{\mathrm{inn}}(\theta)$ *be the solution of* (1-7) *with*

$$w(0) = \epsilon^{-\alpha}, \quad w_\theta(0) = 0, \quad w_{\theta\theta}(0) = (\tilde{b} + \mu)\epsilon^{-\alpha-2}, \quad w_{\theta\theta\theta}(0) = 0.$$

Then for any sufficiently small $\epsilon > 0$, $\theta/\epsilon \in [e^T, e^{10T}]$ *with* $T \gg 1$, *and* $\mu = O((\epsilon/\theta)^{\nu_1-\sigma})$, *there holds*

$$\begin{aligned}
&w_{\epsilon,\mu}^{\mathrm{inn}}(\theta)\\
&= \frac{A_p}{\theta^\alpha} + \frac{C_p}{\theta^{\alpha-2}} + B_p\mu\epsilon^{-\nu_1}\theta^{\tilde{\nu}_1} + \sum_{k=2}^{\infty}\sum_{j=1}^{k} d_j^k \epsilon^{2(k-j)}\theta^{2j-\alpha}\\
&\quad + \epsilon^{(n-5)/2-\alpha}\Biggl(\frac{a_0\cos\bigl(\beta\ln\frac{\theta}{\epsilon}\bigr) + b_0\sin\bigl(\beta\ln\frac{\theta}{\epsilon}\bigr)}{\theta^{(n-5)/2}} + \frac{a_1\cos\bigl(\beta\ln\frac{\theta}{\epsilon}\bigr) + b_1\sin\bigl(\beta\ln\frac{\theta}{\epsilon}\bigr)}{\theta^{(n-5)/2-2}}\\
&\qquad\qquad + \sum_{k=2}^{\infty}\Biggl(\sum_{j=1}^{k} e_j^k\epsilon^{2(k-j)}\theta^{2j-(n-5)/2}\sin\bigl(\beta\ln\tfrac{\theta}{\epsilon} + E_j^k\bigr) + o(\theta^{2k-(n-5)/2})\Biggr)\\
&\qquad\qquad\qquad + O(\theta^{2-(n-5)/2})\Biggr)\\
&\quad + \epsilon^{-\alpha}\sum_{k=1}^{\infty}\Biggl(\sum_{j=1}^{k}(\mu f_j^k \epsilon^{2k-2j-\tilde{\nu}_1}\theta^{2j+\tilde{\nu}_1})\\
&\qquad\qquad + O(\mu^2\theta^{\tilde{\nu}_1+\nu_1+2k}\epsilon^{-\tilde{\nu}_1-\nu_1} + \mu\theta^{\tilde{\nu}_1+\sigma+2k}\epsilon^{-\tilde{\nu}_1-\sigma})\\
&\qquad\qquad\qquad + O\Bigl(\mu^2\Bigl(\frac{\theta}{\epsilon}\Bigr)^{\tilde{\nu}_1+\nu_1} + \mu\Bigl(\frac{\theta}{\epsilon}\Bigr)^{\tilde{\nu}_1+\sigma}\Bigr)\Biggr).
\end{aligned}$$

Proof. This is a direct consequence of Proposition 3.3 by setting $r = \theta/\epsilon$. □

We now obtain some useful lemmas.

Lemma 3.5. *Let* $(n+3)/(n-5) < p < p_c(n-1)$ *and*

$$v(Q, \mu, \theta) = Qv_0(Q^{(p-1)/4}\theta).$$

Then for $Q^{(p-1)/4}\theta \in [e^T, e^{10T}]$ *with* $T \gg 1$,

$$\mu = O\biggl(\frac{1}{(Q^{(p-1)/4}\theta)^{\nu_1-\sigma}}\biggr)$$

and $n = 0, 1, 2$, we have that $v(Q, \mu, \theta)$ satisfies

$$
\begin{aligned}
&\frac{\partial^n}{\partial Q^n}(v(Q,\mu,\theta))\\
&= \frac{\partial^n}{\partial Q^n}\left(\frac{A_p}{\theta^\alpha}\right) + \frac{\partial^n}{\partial Q^n}\big(C\theta^{-(n-5)/2}Q^{-((p-1)(n-5)/8-1)}\sin\big(\beta\ln(Q^{(p-1)/4}\theta)+\kappa\big)\big)\\
&\qquad\qquad + Q^{\tilde{\nu}_2/\alpha+1-n}O(\theta^{\tilde{\nu}_2}) + \mu B_p Q^{\tilde{\nu}_1/\alpha+1-n}\theta^{\tilde{\nu}_1}\\
&\qquad\qquad\qquad + O\big(\mu^2 Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1-n}\theta^{\tilde{\nu}_1+\nu_1} + \mu Q^{(\tilde{\nu}_1+\sigma)/\alpha+1-n}\theta^{\sigma+\tilde{\nu}_1}\big),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^n}{\partial Q^n}(v'_\theta(Q,\mu,\theta))\\
&= \frac{\partial^n}{\partial Q^n}\left(-\alpha\frac{A_p}{\theta^{\alpha+1}}\right)\\
&\qquad + \frac{\partial^{n+1}}{\partial Q^n\partial\theta}\big(C\theta^{-(n-5)/2}Q^{-((p-1)(n-5)/8-1)}\sin\big(\beta\ln(Q^{(p-1)/4}\theta)+\kappa\big)\big)\\
&\qquad\qquad + Q^{\tilde{\nu}_2/\alpha+1-n}O(\theta^{\tilde{\nu}_2-1}) + \mu\tilde{\nu}_1 B_p Q^{\tilde{\nu}_1/\alpha+1-n}\theta^{\tilde{\nu}_1-1}\\
&\qquad\qquad\qquad + O\big(\mu^2 Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1-n}\theta^{\tilde{\nu}_1+\nu_1-1} + \mu Q^{(\tilde{\nu}_1+\sigma)/\alpha+1-n}\theta^{\sigma+\tilde{\nu}_1-1}\big),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^n}{\partial Q^n}\left(\frac{\partial^2}{\partial\theta^2}v(Q,\mu,\theta)\right)\\
&= \frac{\partial^n}{\partial Q^n}\left(\alpha(\alpha+1)\frac{A_p}{\theta^{\alpha+2}}\right)\\
&\qquad + \frac{\partial^{n+2}}{\partial Q^n\partial\theta^2}\big(C\theta^{-(n-5)/2}Q^{-((p-1)(n-5)/8-1)}\sin\big(\beta\ln(Q^{(p-1)/4}\theta)+\kappa\big)\big)\\
&\qquad\qquad + Q^{\tilde{\nu}_2/\alpha+1-n}O(\theta^{\tilde{\nu}_2-2}) + \mu\tilde{\nu}_1(\tilde{\nu}_1-1) B_p Q^{\tilde{\nu}_1/\alpha+1-n}\theta^{\tilde{\nu}_1-2}\\
&\qquad\qquad\qquad + O\big(\mu^2 Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1-n}\theta^{\tilde{\nu}_1+\nu_1-2} + \mu Q^{(\tilde{\nu}_1+\sigma)/\alpha+1-n}\theta^{\sigma+\tilde{\nu}_1-2}\big),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^n}{\partial Q^n}\left(\frac{\partial^3}{\partial\theta^3}v(Q,\mu,\theta)\right)\\
&= \frac{\partial^n}{\partial Q^n}\left(-\alpha(\alpha+1)(\alpha+2)\frac{A_p}{\theta^{\alpha+3}}\right)\\
&\qquad + \frac{\partial^{n+3}}{\partial Q^n\partial\theta^3}\big(C\theta^{-(n-5)/2}Q^{-((p-1)(n-5)/8-1)}\sin\big(\beta\ln(Q^{(p-1)/4}\theta)+\kappa\big)\big)\\
&\qquad\qquad + Q^{\tilde{\nu}_2/\alpha+1-n}O(\theta^{\tilde{\nu}_2-3}) + \mu\tilde{\nu}_1(\tilde{\nu}_1-1)(\tilde{\nu}_1-2) B_p Q^{\tilde{\nu}_1/\alpha+1-n}\theta^{\tilde{\nu}_1-3}\\
&\qquad\qquad\qquad + O\big(\mu^2 Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1-n}\theta^{\tilde{\nu}_1+\nu_1-3} + \mu Q^{(\tilde{\nu}_1+\sigma)/\alpha+1-n}\theta^{\sigma+\tilde{\nu}_1-3}\big),
\end{aligned}
$$

where $\kappa = \tan^{-1}(b_0/a_0)$ and $C = \sqrt{a_0^2+b_0^2}$.

For $n = 0, 1$, *we have*

$$\begin{aligned}
&\frac{\partial^n}{\partial \mu^n}(v(Q,\mu,\theta)) \\
&\quad = \mu^{1-n} B_p Q^{\tilde{\nu}_1/\alpha+1}\theta^{\tilde{\nu}_1} + O\big(\mu^{2-n} Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1} + \mu^{1-n} Q^{(\tilde{\nu}_1+\sigma)/\alpha+1}\theta^{\sigma+\tilde{\nu}_1}\big),\\
&\frac{\partial^n}{\partial \mu^n}\left(\frac{\partial}{\partial\theta} v(Q,\mu,\theta)\right)\\
&\quad = \mu^{1-n}\tilde{\nu}_1 B_p Q^{\tilde{\nu}_1/\alpha+1}\theta^{\tilde{\nu}_1-1}\\
&\qquad\qquad + O\big(\mu^{2-n} Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1-1} + \mu^{1-n} Q^{(\tilde{\nu}_1+\sigma)/\alpha+1}\theta^{\sigma+\tilde{\nu}_1-1}\big),\\
&\frac{\partial^n}{\partial \mu^n}\left(\frac{\partial^2}{\partial\theta^2} v(Q,\mu,\theta)\right)\\
&\quad = \mu^{1-n}\tilde{\nu}_1(\tilde{\nu}_1-1) B_p Q^{\tilde{\nu}_1/\alpha+1}\theta^{\tilde{\nu}_1-2}\\
&\qquad\qquad + O\big(\mu^{2-n} Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1-2} + \mu^{1-n} Q^{(\tilde{\nu}_1+\sigma)/\alpha+1}\theta^{\sigma+\tilde{\nu}_1-2}\big),\\
&\frac{\partial^n}{\partial \mu^n}\left(\frac{\partial^3}{\partial\theta^2} v(Q,\mu,\theta)\right)\\
&\quad = \mu^{1-n}\tilde{\nu}_1(\tilde{\nu}_1-1)(\tilde{\nu}_1-2) B_p Q^{\tilde{\nu}_1/\alpha+1}\theta^{\tilde{\nu}_1-3}\\
&\qquad\qquad + O\big(\mu^{2-n} Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1-3} + \mu^{1-n} Q^{(\tilde{\nu}_1+\sigma)/\alpha+1}\theta^{\sigma+\tilde{\nu}_1-3}\big),
\end{aligned}$$

while for $n = 2$, *we have*

$$\frac{\partial^2}{\partial \mu^2}\left(\frac{\partial^m}{\partial\theta^m} v(Q,\mu,\theta)\right) = O(Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1-m}), \quad m = 0, 1, 2, 3.$$

Proof. These estimates are obtained by the expansions of $v_{01}(r)$ and $v_{02}(r)$ given above and direct calculation. □

Lemma 3.6. *In the region*

$$\theta = |O(Q^{\sigma/((2-\sigma)\alpha)})|, \quad \mu = O(\theta^{2-2\nu_1/\sigma}), \quad \sigma = -\tfrac{1}{2}(n-5-2\alpha),$$

the solution $w(Q,\mu,\theta)$ *of* (1-7) *with*

$$\begin{aligned} w(Q,\mu,0) &= Q, & w_\theta(Q,\mu,0) &= 0,\\ w_{\theta\theta}(Q,\mu,0) &= (\tilde{b}+\mu)Q^{1+2/\alpha}, & w_{\theta\theta\theta}(Q,\mu,0) &= 0 \end{aligned}$$

satisfies

$$\begin{aligned}(1)\quad &\left|\frac{\partial^{m+n}}{\partial Q^n\partial\theta^m} w(Q,\mu,\theta) - \frac{\partial^{m+n}}{\partial Q^n\partial\theta^m} v(Q,\mu,\theta)\right|\\ &\qquad\qquad = Q^{-(n-5)(p-1)/8-(n-1)}\left|o(\theta^{-(n-5)/2-m})\right|,\end{aligned}$$

(2) $$\left|\frac{\partial^{m+n}}{\partial\mu^n\partial\theta^m}w(Q,\mu,\theta)-\frac{\partial^{n+m}}{\partial\mu^n\partial\theta^m}v(Q,\mu,\theta)\right| = \left|O(\mu^{2-n}Q^{(\tilde{\nu}_1+\nu_1)/\alpha+1}\theta^{\tilde{\nu}_1+\nu_1-m})\right|.$$

Proof. This lemma can be obtained from Lemma 3.5 and Theorem 3.4. Note that

$$\epsilon = Q^{-1/\alpha}, \quad \sigma/\alpha = \tfrac{1}{8}(p-1)(n-5)-1.$$

Moreover,

$$Q^{(p-1)/4}\theta \in [e^T, e^{10T}]$$

provided that Q is sufficiently large. □

Now we write the inner solutions obtained in Theorem 3.4 in terms of the parameters Q and μ.

Theorem 3.7. *Let* $(n+3)/(n-5) < p < p_c(n-1)$ *and let* $w^{\text{inn}}_{Q,\mu}(\theta)$ *be an inner solution of problem* (1-7) *with* $w(0)=Q$, $w_\theta(0)=0$, $w_{\theta\theta}(0)=(\tilde{b}+\mu)Q^{1+2/\alpha}$, $w_{\theta\theta\theta}(0)=0$. *Then for any sufficiently large* $Q>0$ *and* $\theta = |O(Q^{\sigma/((2-\sigma)\alpha)})| = |O(\mu^{\sigma/(2\sigma-2\nu_1)})|$,

$$\begin{aligned} w^{\text{inn}}_{Q,\mu}(\theta) = {} & \frac{A_p}{\theta^\alpha}+\frac{C_p}{\theta^{\alpha-2}}+B_p\mu Q^{\nu_1/\alpha}\theta^{\tilde{\nu}_1}+\sum_{k=2}^{\infty}\sum_{j=1}^{k}d^k_jQ^{-(p-1)(k-j)/2}\theta^{2j-\alpha} \\ & +Q^{\sigma/\alpha}\Bigg(\frac{a_0\cos\bigl(\beta\ln(Q^{(p-1)/4}\theta)\bigr)+b_0\sin\bigl(\beta\ln(Q^{(p-1)/4}\theta)\bigr)}{\theta^{(n-5)/2}} \\ & \qquad +\frac{a_1\cos\bigl(\beta\ln(Q^{(p-1)/4}\theta)\bigr)+b_1\sin\bigl(\beta\ln(Q^{(p-1)/4}\theta)\bigr)}{\theta^{(n-5)/2-2}} \\ & \qquad +O(\theta^{2-(n-5)/2}) \\ & \qquad +\sum_{k=2}^{\infty}\Bigg(\sum_{j=1}^{k}e^k_jQ^{-(p-1)(k-j)/2}\theta^{2j-(n-5)/2} \\ & \qquad\qquad \times\sin\bigl(\beta\ln(Q^{(p-1)/4}\theta)+E^k_j\bigr)+o(\theta^{2k-(n-5)/2})\Bigg)\Bigg) \\ & +Q\sum_{k=1}^{\infty}\Bigg(\sum_{j=1}^{k}\bigl(\mu f^k_jQ^{-(2k-2j-\tilde{\nu}_1)/\alpha}\theta^{2j+\tilde{\nu}_1}\bigr) \\ & \qquad +O\bigl(\mu^2Q^{(\tilde{\nu}_1+\nu_1)/\alpha}\theta^{\tilde{\nu}_1+\nu_1+2k}+\mu Q^{(\tilde{\nu}_1+\sigma)/\alpha}\theta^{\tilde{\nu}_1+\sigma+2k}\bigr)\Bigg). \end{aligned}$$

4. Outer solutions

In this section, we construct outer solutions for (1-7). Let $w_*(\theta)$ be the singular solution given in (1-8).

Lemma 4.1. *The equation*

(4-1) $$T_1\phi(\theta)+k_1T_2\phi(\theta)+k_0\phi = pw_*^{p-1}(\theta)\phi(\theta), \quad 0<\theta<\tfrac{\pi}{2},$$

admits a solution, which can be written as

$$\phi(\theta)=\theta^{-(n-5)/2}\big(c_1\cos\big(\beta\ln\tfrac{\theta}{2}\big)+c_2\sin\big(\beta\ln\tfrac{\theta}{2}\big)\big)+O(\theta^{2-(n-5)/2}) \quad \text{as } \theta\to 0, \tag{4-2}$$

where c_1, c_2 *are constants such that* $c_1^2+c_2^2\neq 0$, *and also admits another solution, which can be written as*

$$\psi(\theta)=c_0\theta^{\tilde{\nu}_2}+O(\theta^{\tilde{\nu}_2+2}) \quad \text{as } \theta\to 0, \tag{4-3}$$

where c_0 *is a nonzero constant. Here* T_1 *and* T_2 *are differential operators defined in* (1-7).

Proof. For the equations

$$\begin{cases} T_1\phi_1(\theta)+k_1T_2\phi_1(\theta)+k_0\phi_1(\theta)=pw_*^{p-1}(\theta)\phi_1(\theta), & 0<\theta<\frac{\pi}{2},\\ \phi_1\big(\frac{\pi}{2}\big)=1,\ \phi_1'\big(\frac{\pi}{2}\big)=0,\ \phi_1''\big(\frac{\pi}{2}\big)=0,\ \phi_1'''\big(\frac{\pi}{2}\big)=0, \end{cases} \tag{4-4}$$

and

$$\begin{cases} T_1\phi_2(\theta)+k_1T_2\phi_2(\theta)+k_0\phi_2(\theta)=pw_*^{p-1}(\theta)\phi_2(\theta), & 0<\theta<\frac{\pi}{2},\\ \phi_2\big(\frac{\pi}{2}\big)=0,\ \phi_2'\big(\frac{\pi}{2}\big)=0,\ \phi_2''\big(\frac{\pi}{2}\big)=1,\ \phi_2'''\big(\frac{\pi}{2}\big)=0, \end{cases} \tag{4-5}$$

we claim that both $\phi_1(\theta)$ and $\phi_2(\theta)$ are strictly decreasing for $\theta\in\big(0,\frac{\pi}{2}\big)$. We only show the case of $\phi_2(\theta)$; the case of $\phi_1(\theta)$ can be treated similarly.

Let us set

$$A(\theta)=\frac{d}{d\theta}\bigg(\sin^{n-2}\theta\,\frac{d\phi_2(\theta)}{d\theta}\bigg).$$

Before proving that $\phi_2(\theta)$ is decreasing, we first present a useful fact. If $A(\theta)>0$ for $\theta\in\big(\theta_0,\frac{\pi}{2}\big)$, where $\theta_0\in\big(0,\frac{\pi}{2}\big)$, then for $\theta\in\big(\theta_0,\frac{\pi}{2}\big)$, we have $\phi_2'(\theta)<0$ and $\phi_2(\theta)>0$. The proof of this fact is simple; thus we omit it here. Next, we show that $\phi_2(\theta)$ is decreasing. By using the boundary condition of ϕ_2 at $\theta=\frac{\pi}{2}$, we have $A\big(\frac{\pi}{2}\big)=1$ and find $\theta_1\in\big(0,\frac{\pi}{2}\big)$ such that $A(\theta)>0$ for $\theta\in\big(\theta_1,\frac{\pi}{2}\big)$; then $\phi_2(\theta)>0$ for $\theta\in\big(\theta_1,\frac{\pi}{2}\big)$. Using the fact that $k_1(n)<0$ and the second conclusion in Lemma A.1, we have

$$T_1\phi_2(\theta)=(pw_*^{p-1}-k_0)\phi_2(\theta)-k_1\frac{A(\theta)}{\sin^{n-2}\theta}>0 \quad \text{for } \theta\in\big(\theta_1,\tfrac{\pi}{2}\big).$$

Now we are going to show that $\theta_1=0$. If not, $\theta_1\in\big(0,\frac{\pi}{2}\big)$ and $A(\theta_1)=0$. For $\theta\in\big(\theta_1,\frac{\pi}{2}\big)$, we have

$$\frac{d}{d\theta}\bigg(\sin^{n-2}\theta\,\frac{d}{d\theta}\bigg(\frac{A(\theta)}{\sin^{n-2}\theta}\bigg)\bigg)>0.$$

Using this inequality and

$$\frac{d}{d\theta}\bigg(\frac{A(\theta)}{\sin^{n-2}\theta}\bigg)\bigg|_{\theta=\frac{\pi}{2}}=0,$$

we have

$$\frac{d}{d\theta}\left(\frac{A(\theta)}{\sin^{n-2}\theta}\right) < 0 \quad \text{for } \theta \in \left(\theta_1, \tfrac{\pi}{2}\right). \tag{4-6}$$

It follows from (4-6) that

$$\frac{A(\theta)}{\sin^{n-2}\theta} > 1 \quad \text{for } \theta \in \left(\theta_1, \tfrac{\pi}{2}\right), \tag{4-7}$$

which contradicts the fact that $A(\theta_1) = 0$. Thus, $A(\theta) > 0$ and $\phi_2'(\theta) < 0$ for $\theta \in \left(0, \frac{\pi}{2}\right)$. Hence, we have proved the claim.

We now prove that there are $D_1 \neq 0$ and $D_2 \neq 0$ such that for θ near 0,

$$\phi_1(\theta) = D_1\theta^{\tilde{\nu}_2} + O(\theta^{2+\tilde{\nu}_2}) \tag{4-8}$$

and

$$\phi_2(\theta) = D_2\theta^{\tilde{\nu}_2} + O(\theta^{2+\tilde{\nu}_2}). \tag{4-9}$$

We only show (4-9). The proof of (4-8) is similar. Using the Emden–Fowler transformation

$$\tilde{\phi}(t) = (\sin\theta)^{\alpha}\phi_2(\theta), \quad t = \ln\left(\tan\tfrac{\theta}{2}\right),$$

we obtain that $\tilde{\phi}(t)$, for $t \in (-\infty, 0)$, satisfies the homogeneous equation

$$\tilde{\phi}^{(4)}(t) + a_3(t)\tilde{\phi}'''(t) + a_2(t)\tilde{\phi}''(t) + a_1(t)\tilde{\phi}'(t) + a_0(t)\tilde{\phi}(t) = 0, \tag{4-10}$$

where

$$a_3(t) = K_3 + O(e^{2t}), \quad a_2(t) = K_2 + O(e^{2t}),$$
$$a_1(t) = K_1 + O(e^{2t}), \quad a_0(t) = (1-p)K_0.$$

Therefore,

$$\tilde{\phi}^{(4)}(t) + K_3\tilde{\phi}'''(t) + K_2\tilde{\phi}''(t) + K_1\tilde{\phi}'(t) + (1-p)K_0\tilde{\phi}(t) = O(e^{2t}(\tilde{\phi}'''(t) + \tilde{\phi}''(t) + \tilde{\phi}'(t))). \tag{4-11}$$

Following the arguments in the proof of Lemma 3.1, we can write the solutions of (4-11) as (for any $T \ll -1$):

$$\begin{aligned}\tilde{\phi}(t) = {}& A_5e^{\sigma t}\cos\beta t + A_6e^{\sigma t}\sin\beta t + A_7e^{\nu_2 t} + A_8e^{\nu_1 t}\\ &+ B_5\int_{-\infty}^{t} e^{\sigma(t-s)}\sin\beta(t-s)g(s,\tilde{\phi}(s))\,\mathrm{d}s\\ &+ B_6\int_{-\infty}^{t} e^{\sigma(t-s)}\cos\beta(t-s)g(s,\tilde{\phi}(s))\,\mathrm{d}s\\ &+ B_7\int_{-\infty}^{t} e^{\nu_2(t-s)}g(s,\tilde{\phi}(s))\,\mathrm{d}s + B_8\int_{T}^{t} e^{\nu_1(t-s)}g(s,\tilde{\phi}(s))\,\mathrm{d}s,\end{aligned} \tag{4-12}$$

where $g(t, \tilde{\phi}(t))$ is the right-hand side of (4-11), A_8 depends on T and each B_{i+4} depends only on ν_i $(i = 1, 2, 3, 4)$. It is known from (4-12) that if $A_7 = 0$, then for $|t|$ sufficiently large,

$$\tilde{\phi}(t) = A_5 e^{\sigma t} \cos \beta t + A_6 e^{\sigma t} \sin \beta t + O(e^{(2+\sigma)t}) \tag{4-13}$$

with $A_5^2 + A_6^2 \neq 0$ or

$$\tilde{\phi}(t) = A_8 e^{\nu_1 t} + O(e^{(2+\nu_1)t}) \tag{4-14}$$

with $A_8 \neq 0$. Otherwise, if $A_5^2 + A_6^2 = 0$ and $A_8 = 0$, we know that $\tilde{\phi}(t) = O(e^{(2+\nu_1)t})$. Substituting this into (4-12), we see that $\tilde{\phi}(t) = O(e^{(4+\nu_1)t})$; repeating this procedure, we eventually obtain that $\tilde{\phi}(t) \equiv 0$. This is impossible. Therefore, for θ near 0,

$$\phi_2(\theta) = A_5 \theta^{-(n-5)/2} \cos\bigl(\beta \ln \tfrac{\theta}{2}\bigr) + A_6 \theta^{-(n-5)/2} \sin\bigl(\beta \ln \tfrac{\theta}{2}\bigr) + O(\theta^{2-(n-5)/2})$$

or

$$\phi_2(\theta) = A_8 \theta^{\tilde{\nu}_1} + O(\theta^{2+\tilde{\nu}_1}).$$

But these contradict the fact that $\phi_2(\theta)$ is strictly decreasing for $\theta \in \bigl(0, \frac{\pi}{2}\bigr)$. Thus, we prove the claim and get (4-9).

Let $\phi(\theta) = \phi_1(\theta) - (D_1/D_2)\phi_2(\theta)$. Then $\phi(\theta)$ satisfies the problem

$$\begin{cases} T_1 \phi(\theta) + k_1 T_2 \phi(\theta) + k_0 \phi(\theta) = p w_*^{p-1}(\theta) \phi(\theta), \quad 0 < \theta < \frac{\pi}{2}, \\ \phi\bigl(\frac{\pi}{2}\bigr) = 1, \ \phi'\bigl(\frac{\pi}{2}\bigr) = 0, \ \phi''\bigl(\frac{\pi}{2}\bigr) = -D_1/D_2, \ \phi'''\bigl(\frac{\pi}{2}\bigr) = 0. \end{cases} \tag{4-15}$$

We claim that for θ near 0,

$$\phi(\theta) = \theta^{-(n-5)/2} \bigl(c_1 \cos\bigl(\beta \ln \tfrac{\theta}{2}\bigr) + c_2 \sin\bigl(\beta \ln \tfrac{\theta}{2}\bigr)\bigr) + O(\theta^{2-(n-5)/2}) \tag{4-16}$$

with $c_1^2 + c_2^2 \neq 0$. Using the Emden–Fowler transformation

$$\hat{\phi}(t) = (\sin \theta)^{\alpha} \phi(\theta), \qquad t = \ln\bigl(\tan \tfrac{\theta}{2}\bigr), \tag{4-17}$$

(4-8) and (4-9), we obtain that for t near $-\infty$,

$$\hat{\phi}(t) = e^{\sigma t} (c_1 \cos(\beta t) + c_2 \sin(\beta t)) + c_3 e^{\nu_1 t} + O(e^{(2+\sigma)t}) \tag{4-18}$$

provided $c_1^2 + c_2^2 \neq 0$ or

$$\hat{\phi}(t) = c_3 e^{\nu_1 t} + O(e^{(2+\nu_1)t}) \tag{4-19}$$

provided $c_1^2 + c_2^2 = 0$ and $c_3 \neq 0$. (Note that if both $c_1^2 + c_2^2 = 0$ and $c_3 = 0$, we can obtain $\hat{\phi}(t) \equiv 0$. This is impossible.) We now show that (4-19) cannot occur. On the contrary, we see that for θ near 0,

$$\phi(\theta) = c_3 \theta^{\tilde{\nu}_1} + O(\theta^{2+\tilde{\nu}_1}).$$

This implies that $\phi(\theta) \to 0$ as $\theta \to 0$. Since

$$\hat{\phi}(t) = O(e^{\nu_1 t}), \quad \hat{\phi}'(t) = O(e^{\nu_1 t}), \quad \hat{\phi}''(t) = O(e^{\nu_1 t}), \quad \hat{\phi}'''(t) = O(e^{\nu_1 t}),$$

we obtain from (4-17) that

$$\begin{aligned} \phi'(\theta) &= O(\theta^{\tilde{\nu}_1 - 1}), \\ \sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta} &= O(\theta^{n-3+\tilde{\nu}_1}), \\ \frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right) &= O(\theta^{n-4+\tilde{\nu}_1}). \end{aligned}$$

Similar arguments imply that

$$\sin^{n-2}\theta \frac{d}{d\theta}\left(\frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right)\right) = O(\theta^{n-5+\tilde{\nu}_1}).$$

If we define

$$e(\theta) = \sin^{n-2}\theta \frac{d}{d\theta}\left(\frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right)\right),$$

we see that $e(0) = 0$. Then, we claim that ϕ changes sign in $\left(0, \frac{\pi}{2}\right)$. Suppose that this is not true. Without loss of generality, we assume $\phi > 0$ in $\left(0, \frac{\pi}{2}\right)$. Then it follows from the equation of ϕ that for $\theta \in \left(0, \frac{\pi}{2}\right)$,

$$\frac{d}{d\theta}\left(e(\theta) + k_1\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right)\right) = \sin^{n-2}\theta(pw_*^{p-1} - k_0)\phi(\theta) > 0. \tag{4-20}$$

But integrating both sides of (4-20) in $\left(0, \frac{\pi}{2}\right)$ and using the boundary conditions $\phi'\left(\frac{\pi}{2}\right) = \phi'''\left(\frac{\pi}{2}\right) = 0$, we obtain

$$\int_0^{\frac{\pi}{2}} \sin^{n-2}\theta(pw_*^{p-1} - k_0)\phi(\theta)\, d\theta = 0.$$

This is clearly impossible. Noticing that $\phi \neq 0$ for θ near 0, we see that there is a minimal zero point $\hat{\theta} \in \left(0, \frac{\pi}{2}\right)$ of ϕ. Without loss of generality, we assume that $\phi > 0$ in $(0, \hat{\theta})$. It follows from (4-20) that $E(\theta) := e(\theta) + k_1 \sin^{n-2}\theta(d\phi(\theta)/d\theta)$ is increasing for $\theta \in (0, \hat{\theta})$. Noticing $E(0) = 0$, we then obtain that $E(\theta) > 0$ for $\theta \in (0, \hat{\theta})$. Therefore,

$$\frac{d}{d\theta}\left(\frac{1}{\sin^{n-2}\theta}\frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right) + k_1\phi(\theta)\right) > 0 \quad \text{for } \theta \in (0, \hat{\theta}). \tag{4-21}$$

Moreover, by a similar argument, we have

$$\frac{d}{d\theta}\left(\sin^{n-2}\theta \frac{d\phi(\theta)}{d\theta}\right) > 0 \quad \text{for } \theta \in (0, \hat{\theta}), \tag{4-22}$$

and

$$\frac{d\phi(\theta)}{d\theta} > 0 \quad \text{for } \theta \in (0, \hat{\theta}). \tag{4-23}$$

But (4-23) implies $\phi(\hat{\theta}) > 0$, which contradicts the fact that $\phi(\hat{\theta}) = 0$. This contradiction implies that (4-19) cannot occur and thus (4-18) holds. As a consequence, (4-16) holds and hence (4-2) holds.

Let $\psi(\theta) = \phi_1(\theta)$. We easily see that (4-3) can be obtained from (4-8). □

For any sufficiently small $\delta > \eta > 0$, we set $\psi_1(\theta)$ to be the solution of the problem

$$\begin{cases} T_1\psi_1(\theta) + k_1 T_2\psi_1(\theta) + k_0\psi_1(\theta) \\ \qquad\qquad = \eta^{-2}\big((w_* + \Phi + \Psi)^p - w_*^p - p w_*^{p-1}(\Phi + \eta^2\psi)\big), \\ (\psi_1 + \psi)\big(\frac{\pi}{2}\big) = 2, \qquad\qquad (\psi_1 + \psi)'\big(\frac{\pi}{2}\big) = 0, \\ (\psi_1 + \psi)''\big(\frac{\pi}{2}\big) = D_1\delta^2/(D_2\eta^2), \quad (\psi_1 + \psi)'''\big(\frac{\pi}{2}\big) = 0, \end{cases} \tag{4-24}$$

where $\psi(\theta)$ is given in Lemma 4.1, $\Phi = \delta^2\phi(\theta)$ and $\Psi = \eta^2(\psi_1(\theta) + \psi(\theta))$. We can see that Ψ satisfies the problem

$$\begin{cases} T_1\Psi(\theta) + k_1 T_2\Psi(\theta) + k_0\Psi(\theta) = (w_* + \Phi + \Psi)^p - w_*^p - p w_*^{p-1}\Phi, \\ \Psi\big(\frac{\pi}{2}\big) = 2\eta^2, \ \Psi'\big(\frac{\pi}{2}\big) = 0, \ \Psi''\big(\frac{\pi}{2}\big) = D_1\delta^2/D_2, \ \Psi'''\big(\frac{\pi}{2}\big) = 0. \end{cases} \tag{4-25}$$

This implies

$$\begin{cases} T_1(\Psi + \Phi) + k_1 T_2(\Psi + \Phi) + k_0(\Psi + \Phi) = (w_* + \Phi + \Psi)^p - w_*^p, \\ (\Psi + \Phi)\big(\frac{\pi}{2}\big) = 2\eta^2 + \delta^2, \quad (\Psi + \Phi)'\big(\frac{\pi}{2}\big) = 0, \\ (\Psi + \Phi)''\big(\frac{\pi}{2}\big) = 0, \qquad\quad (\Psi + \Phi)'''\big(\frac{\pi}{2}\big) = 0. \end{cases} \tag{4-26}$$

Arguments similar to those in the proof of Lemma 4.1 imply that $\Psi(\theta) + \Phi(\theta)$ is strictly decreasing. Then

$$\Psi(\theta) + \Phi(\theta) > 0 \quad \text{for } \theta \in \big(0, \tfrac{\pi}{2}\big). \tag{4-27}$$

Setting $\psi_2(\theta) = \psi(\theta) + \psi_1(\theta)$, we easily see that ψ_2 satisfies the problem

$$\begin{cases} T_1\psi_2(\theta) + k_1 T_2\psi_2(\theta) + k_0\psi_2(\theta) \\ \qquad = p w_*^{p-1}\psi_2 + \eta^{-2}\big((w_* + \Phi + \eta^2\psi_2)^p - w_*^p - p w_*^{p-1}(\Phi + \eta^2\psi_2)\big), \\ \psi_2\big(\frac{\pi}{2}\big) = 2, \ \psi_2'\big(\frac{\pi}{2}\big) = 0, \ \psi_2''\big(\frac{\pi}{2}\big) = D_1\delta^2/(D_2\eta^2), \ \psi_2'''\big(\frac{\pi}{2}\big) = 0. \end{cases} \tag{4-28}$$

By the Emden–Fowler transformation

$$\tilde{\psi}_2(t) = (\sin\theta)^\alpha \psi_2(\theta), \quad t = \ln\tan\tfrac{\theta}{2},$$

we see that $\tilde{\psi}_2(t)$ satisfies the problem

$$
\text{(4-29)} \quad \begin{cases} \tilde{\psi}_2^{(4)}(t) + a_3(t)\tilde{\psi}_2'''(t) + a_2(t)\tilde{\psi}_2''(t) \\ \qquad\qquad + a_1(t)\tilde{\psi}_2'(t) + a_0(t)\tilde{\psi}_2(t) = G(\tilde{\psi}_2(t)), \quad -\infty < t < 0, \\ \tilde{\psi}_2'(0) = 0, \ \tilde{\psi}_2'''(0) = 0, \end{cases}
$$

where $a_0(t), a_1(t), a_2(t), a_3(t)$ are defined in (4-10), and

$$
\begin{aligned}
&G(\tilde{\psi}_2(t)) \\
&\ = (\sin\theta)^{4+\alpha}\eta^{-2}\big((w_* + \Phi + \eta^2 \sin^{-\alpha}\theta\tilde{\psi}_2)^p - w_*^p - pw_*^{p-1}(\Phi + \eta^2\sin^{-\alpha}\theta\tilde{\psi}_2)\big).
\end{aligned}
$$

Moreover, we can rewrite (4-29) in the following form (see the proof of Lemma 4.1):

$$
\text{(4-30)} \quad \begin{aligned} \tilde{\psi}_2^{(4)}(t) + K_3\tilde{\psi}_2'''(t) + K_2\tilde{\psi}_2''(t) + K_1\tilde{\psi}_2'(t) + (1-p)K_0\tilde{\psi}_2(t) \\ = G(\tilde{\psi}_2(t)) + g(t, \tilde{\psi}_2(t)), \end{aligned}
$$

where

$$
g(t, \tilde{\psi}_2(t)) = O\big(e^{2t}(\tilde{\psi}_2'''(t) + \tilde{\psi}_2''(t) + \tilde{\psi}_2'(t))\big)
$$

for $t \ll -1$. Therefore, for $t < T$ with any $T \ll -1$,

$$
\text{(4-31)} \quad \begin{aligned}
\tilde{\psi}_2(t) &= D_5e^{\nu_2 t} + D_6e^{\sigma t}\cos\beta t + D_7e^{\sigma t}\sin\beta t + D_8e^{\nu_1 t} \\
&\quad + B_5\int_{-\infty}^t e^{\sigma(t-s)}\sin\beta(t-s)(G(\tilde{\psi}_2(s)) + g(s, \tilde{\psi}_2(s)))\,\mathrm{d}s \\
&\quad + B_6\int_{-\infty}^t e^{\sigma(t-s)}\cos\beta(t-s)(G(\tilde{\psi}_2(s)) + g(s, \tilde{\psi}_2(s)))\,\mathrm{d}s \\
&\quad + B_7\int_{-\infty}^t e^{\nu_2(t-s)}(G(\tilde{\psi}_2(s)) + g(s, \tilde{\psi}_2(s)))\,\mathrm{d}s \\
&\quad + B_8\int_{T}^t e^{\nu_1(t-s)}(G(\tilde{\psi}_2(s)) + g(s, \tilde{\psi}_2(s)))\,\mathrm{d}s,
\end{aligned}
$$

where B_5, B_6, B_7, B_8 depend only on ν_i $(i = 1, 2, 3, 4)$. Using the fact $\Psi(\theta)+\Phi(\theta)$ is strictly decreasing in $\big(0, \frac{\pi}{2}\big)$ and (4-2), we conclude that $D_5 \neq 0$. Letting $\phi(\theta) = \sin^{-\alpha}\theta\tilde{\phi}(t)$, we see that for $t \in [10T, 2T]$ and $\delta^2 = O(e^{(2-\sigma)t})$, $\eta^2 = O(e^{(2-\nu_2)t})$,

$$
\text{(4-32)} \qquad G(\tilde{\psi}_2(t)) = \eta^{-2}O((\delta^2\tilde{\phi}(t) + \eta^2\tilde{\psi}_2(t))^2) = O(e^{(2+\nu_2)t}).
$$

Note that

$$
\tilde{\phi}(t) = e^{\sigma t}(c_1\cos(\beta t) + c_2\sin(\beta t)) + O(e^{(2+\sigma)t})
$$

and $\tilde{\psi}_2(t) = D_5e^{\nu_2 t} + O(e^{(2+\nu_2)t})$. Then

$$
\delta^2\tilde{\phi}(t) + \eta^2\tilde{\psi}_2(t) = O(e^{2t}).
$$

Therefore, it follows from (4-31) and (4-32) that

$$\tilde{\psi}_2(t) = D_5 e^{\nu_2 t} + D_6 e^{\sigma t}\cos\beta t + D_7 e^{\sigma t}\sin\beta t + O(e^{(2+\nu_2)t}) \tag{4-33}$$

provided $\delta^2 = O(e^{(2-\sigma)t})$ and $\eta^2 = O(e^{(2-\nu_2)t})$. Hence, for θ near 0,

$$\Psi(\theta) = \eta^2\big(D_5\theta^{\tilde{\nu}_2} + \theta^{-(n-5)/2}\big(D_6\cos\big(\beta\ln\tfrac{\theta}{2}\big) + D_7\sin\big(\beta\ln\tfrac{\theta}{2}\big)\big) + O(\theta^{2+\tilde{\nu}_2})\big) \tag{4-34}$$

with $D_5 \neq 0$ provided that

$$\theta = O(\delta^{2/(2-\sigma)}) = O(\eta^{2/(2-\nu_2)}).$$

Since $\tilde{\nu}_2 < 3-n$, we easily see that $\tilde{\nu}_2 + 2 < -(n-5) < -(n-5)/2$. Thus, $\theta^{-(n-5)/2} = o(\theta^{2+\tilde{\nu}_2})$.

Now we can obtain the following theorem.

Theorem 4.2. *For any $\delta > \eta > 0$ sufficiently small, problem* (1-7) *admits outer solutions $w_{\delta,\eta}^{\mathrm{out}} \in C^4\big(0, \frac{\pi}{2}\big)$ satisfying*

$$w_{\delta,\eta}^{\mathrm{out}}(\theta) = w_*(\theta) + \Phi(\theta) + \Psi(\theta), \quad \theta \in \big(0, \tfrac{\pi}{2}\big), \tag{4-35}$$

with $(w_{\delta,\eta}^{\mathrm{out}})'\big(\frac{\pi}{2}\big) = (w_{\delta,\eta}^{\mathrm{out}})'''\big(\frac{\pi}{2}\big) = 0$. Moreover,

$$\begin{aligned} w_{\delta,\eta}^{\mathrm{out}}(\theta) = {}& \frac{A_p}{\theta^\alpha} + \frac{2A_p}{3(p-1)}\frac{1}{\theta^{\alpha-2}} \\ &+ \delta^2\bigg(\frac{\vartheta_1\cos\big(\beta\ln\frac{\theta}{2}\big) + \vartheta_2\sin\big(\beta\ln\frac{\theta}{2}\big)}{\theta^{(n-5)/2}} + O\bigg(\frac{1}{\theta^{(n-5)/2-2}}\bigg)\bigg) \\ &+ \eta^2\big(\vartheta_3\theta^{\tilde{\nu}_2} + O(\theta^{\tilde{\nu}_2+2})\big) \end{aligned} \tag{4-36}$$

provided that

$$\theta = O(\delta^{2/(2-\sigma)}) = O(\eta^{2/(2-\nu_2)}),$$

where $\vartheta_1, \vartheta_2, \vartheta_3$ are constants independent of δ, η such that $\vartheta_1^2 + \vartheta_2^2 \neq 0$, $\vartheta_3 \neq 0$.

Proof. The proof can be obtained from the expressions of $w_*(\theta)$, $\Phi(\theta)$ and $\Psi(\theta)$ given in (1-8), (4-16) and (4-34). □

5. Infinitely many solutions of (1-7) and proof of Theorem 1.1

In this section, we construct infinitely many regular solutions for (1-7) by matching the inner and outer solutions.

We construct solutions of the problem

$$\begin{cases} T_1 w + k_1 T_2 w + k_0 w = w^p, \quad w(\theta) > 0,\ 0 < \theta < \frac{\pi}{2}, \\ w(0) = Q\ (:= \epsilon^{-\alpha}),\ w'\big(\frac{\pi}{2}\big) = 0,\ w''(0) = (\tilde{b} + \mu)\epsilon^{-\alpha-2},\ w'''\big(\frac{\pi}{2}\big) = 0 \end{cases} \tag{5-1}$$

by matching the inner and outer solutions given in Theorems 3.7 and 4.2. To do so, we will find $\Theta \in \left(0, \frac{\pi}{2}\right)$ with

$$\Theta = O(Q^{\sigma/((2-\sigma)\alpha)}) \quad (Q \gg 1)$$

such that the following identities hold:

$$\left(w^{\text{inn}}_{Q,\mu}(\theta) - w^{\text{out}}_{\delta,\eta}(\theta)\right)\big|_{\theta=\Theta} = 0, \tag{5-2}$$

$$\left(w^{\text{inn}}_{Q,\mu}(\theta) - w^{\text{out}}_{\delta,\eta}(\theta)\right)'_{\theta}\big|_{\theta=\Theta} = 0, \tag{5-3}$$

$$\left(w^{\text{inn}}_{Q,\mu}(\theta) - w^{\text{out}}_{\delta,\eta}(\theta)\right)''_{\theta}\big|_{\theta=\Theta} = 0, \tag{5-4}$$

$$\left(w^{\text{inn}}_{Q,\mu}(\theta) - w^{\text{out}}_{\delta,\eta}(\theta)\right)'''_{\theta}\big|_{\theta=\Theta} = 0. \tag{5-5}$$

These will be done by arguments similar to those in the proof of Lemma 6.1 of [Budd and Norbury 1987] and Theorem 1.1 of [Dancer et al. 2012]. Then, we obtain a C^4 function $w(\theta)$ defined by $w(\theta) = w^{\text{inn}}_{Q,\mu}(\theta)$ for $\theta \le \Theta$ and $w(\theta) = w^{\text{out}}_{\delta,\eta}(\theta)$ for $\theta \ge \Theta$ which is a solution to (5-1).

First, we observe that

$$\frac{2A_p}{3(p-1)} = C_p \tag{5-6}$$

by (3-35), where A_p, C_p are given in Section 3.

Define Q_*, δ_*^2, η_*^2 and μ_* by

$$\beta \ln Q_*^{(p-1)/4} + \kappa = \beta \ln 2^{-1} + \omega + 2m\pi, \tag{5-7}$$

$$\delta_*^2 = \sqrt{\frac{a_0^2 + b_0^2}{\vartheta_1^2 + \vartheta_2^2}}\, Q_*^{\sigma/\alpha}, \tag{5-8}$$

$$\eta_*^2 = O(Q_*^{(2-\nu_2)\sigma/((2-\sigma)\alpha)}), \quad \mu_* = O(Q_*^{(2\sigma-2\nu_1)/((2-\sigma)\alpha)}), \tag{5-9}$$

$$\mu_* B_p Q_*^{\nu_1/\alpha} = \vartheta_3 \eta_*^2 \Theta_*^{\tilde{\nu}_2 - \tilde{\nu}_1}, \tag{5-10}$$

where

$$\kappa = \tan^{-1}\left(\frac{a_0}{b_0}\right), \quad \omega = \tan^{-1}\left(\frac{\vartheta_1}{\vartheta_2}\right)$$

and $m \gg 1$ is an integer. The integer m is chosen such that the results in Sections 3 and 4 hold.

Note that

$$O(\delta_*^{2/(2-\sigma)}) = O(Q_*^{\sigma/(\alpha(2-\sigma))}),$$

$$\begin{aligned}a_0\cos(\beta\ln(Q^{(p-1)/4}\theta)) + b_0\sin(\beta\ln(Q^{(p-1)/4}\theta)) \\ = \sqrt{a_0^2+b_0^2}\,\sin(\beta\ln\theta+\beta\ln Q^{(p-1)/4}+\kappa),\end{aligned}$$

$$\vartheta_1\cos\bigl(\beta\ln\tfrac{\theta}{2}\bigr)+\vartheta_2\sin\bigl(\beta\ln\tfrac{\theta}{2}\bigr)=\sqrt{\vartheta_1^2+\vartheta_2^2}\,\sin(\beta\ln\theta+\beta\ln 2^{-1}+\omega).$$

We will see that the Q, μ, δ^2 and η^2 required to satisfy the matching conditions (5-2)–(5-5) can be obtained as small perturbations of Q_*, μ_*, δ_*^2 and η_*^2 given in (5-7)–(5-10), i.e.,

$$Q = Q_*(1+O(Q_*^{2\sigma/((2-\sigma)\alpha)})), \tag{5-11}$$

$$\mu = \mu_*(1+O(Q_*^{2\sigma/((2-\sigma)\alpha)})), \tag{5-12}$$

$$\delta^2 = \delta_*^2(1+O(Q_*^{2\sigma/((2-\sigma)\alpha)})), \tag{5-13}$$

$$\eta^2 = \eta_*^2(1+O(Q_*^{2\sigma/((2-\sigma)\alpha)})). \tag{5-14}$$

To show this we define the function $\boldsymbol{F}(Q,\mu,\delta,\eta)$ by

$$\boldsymbol{F}(Q,\mu,\delta^2,\eta^2)=\begin{bmatrix}\Theta^{(n-5)/2}(w^{\mathrm{inn}}_{Q,\mu}(\Theta)-w^{\mathrm{out}}_{\delta,\eta}(\Theta))\\ \Theta\bigl(\theta^{(n-5)/2}(w^{\mathrm{inn}}_{Q,\mu}(\theta)-w^{\mathrm{out}}_{\delta,\eta}(\theta))\bigr)'_{\theta}\big|_{\theta=\Theta}\\ \Theta^2\bigl(\theta^{(n-5)/2}(w^{\mathrm{inn}}_{Q,\mu}(\theta)-w^{\mathrm{out}}_{\delta,\eta}(\theta))\bigr)''_{\theta}\big|_{\theta=\Theta}\\ \Theta^3\bigl(\theta^{(n-5)/2}(w^{\mathrm{inn}}_{Q,\mu}(\theta)-w^{\mathrm{out}}_{\delta,\eta}(\theta))\bigr)'''_{\theta}\big|_{\theta=\Theta}\end{bmatrix}^T.$$

Now, we regard δ^2, η^2 as new variables. Taking Q_*, μ_*, δ_*^2 and η_*^2, we find a bound for $\boldsymbol{F}(Q_*,\mu_*,\delta_*^2,\eta_*^2)$ by using the behaviors of $w^{\mathrm{inn}}_{Q,\mu}(\theta)$ and $w^{\mathrm{out}}_{\delta,\eta}(\theta)$ given in Theorems 3.7 and 4.2 respectively. Accordingly we find for some $M>1$ suitably large,

$$\bigl|\Theta^{-(n-5)/2}\boldsymbol{F}(Q_*,\mu_*,\delta_*^2,\eta_*^2)\bigr|\le M\Theta^{4-\sigma-(n-5)/2}+\text{small terms}. \tag{5-15}$$

We seek values of Q, μ, δ^2, η^2 which are small perturbations of Q_*, μ_*, δ_*^2, η_*^2 and such that $\boldsymbol{F}(Q,\mu,\delta^2,\eta^2)=0$. As in [Dancer et al. 2012], we need to evaluate the Jacobian of $\boldsymbol{F}$ at $(Q_*,\mu_*,\delta_*^2,\eta_*^2)$:

$$\frac{\partial\boldsymbol{F}(Q,\mu,\delta^2,\eta^2)}{\partial(Q,\mu,\delta^2,\eta^2)}=\begin{bmatrix}I_1+I_3 & I_4 & -D\sin\tau & I_5\\ \beta I_2+q_1I_3 & q_1I_4 & -\beta D\cos\tau & q_4I_5\\ I_6 & q_2I_4 & I_8 & q_5I_5\\ I_7 & q_3I_4 & I_9 & q_6I_5\end{bmatrix}+\text{higher-order terms},$$

where

$$I_1 = C\left(\frac{\sigma}{\alpha}\sin\tau + \frac{\beta(p-1)}{4}\cos\tau\right)Q_*^{\sigma/\alpha-1},$$

$$I_2 = C\left(\frac{\sigma}{\alpha}\cos\tau - \frac{\beta(p-1)}{4}\sin\tau\right)Q_*^{\sigma/\alpha-1},$$

$$I_3 = \frac{\nu_1}{\alpha}B_p\mu_*\Theta^{\tilde\nu_1+(n-5)/2}Q_*^{\nu_1/\alpha-1}, \qquad I_4 = B_pQ_*^{\nu_1/\alpha}\Theta^{\tilde\nu_1+(n-5)/2},$$

$$I_5 = -\vartheta_3\Theta^{\tilde\nu_2+(n-5)/2}, \qquad I_6 = -\beta^2 I_1 - \beta I_2 + q_2 I_3,$$

$$I_7 = -\beta^3 I_2 + 3\beta^2 I_1 + 2\beta I_2 + q_3 I_3, \quad I_8 = \beta^2 D\sin\tau + \beta D\cos\tau,$$

$$I_9 = \beta^3 D\cos\tau - 3\beta^2 D\sin\tau - 2\beta D\cos\tau,$$

$$q_1 = \tilde\nu_1 + \tfrac12(n-5), \quad q_2 = \big(\tilde\nu_1 + \tfrac12(n-7)\big)q_1, \quad q_3 = \big(\tilde\nu_1 + \tfrac12(n-9)\big)q_2,$$

$$q_4 = \tilde\nu_2 + \tfrac12(n-5), \quad q_5 = \big(\tilde\nu_2 + \tfrac12(n-7)\big)q_4, \quad q_6 = \big(\tilde\nu_2 + \tfrac12(n-9)\big)q_5,$$

$$C = \sqrt{a_0^2 + b_0^2}, \quad D = \sqrt{\vartheta_1^2 + \vartheta_2^2},$$

and

$$\tau = \beta\ln\Theta + \beta\ln Q_*^{(p-1)/4} + \kappa = \beta\ln\Theta + \beta\ln 2^{-1} + \omega + 2m\pi.$$

We define the function $\boldsymbol{G}(x, y, z, w)$ by

$$\begin{aligned}&\boldsymbol{G}(x, y, z, w)\\ &\quad= \boldsymbol{F}(Q_* + xQ_*^{1-\sigma/\alpha}, \mu_* + \Theta^{-\tilde\nu_1-(n-5)/2}Q_*^{-\nu_1/\alpha}y, \delta_*^2 + z, \eta_*^2 + \Theta^{-\tilde\nu_2-(n-5)/2}w).\end{aligned}$$

Using (5-15), (4-36) and the results in Lemmas 3.5 and 3.6, we express $\boldsymbol{G}(x, y, z, w)$ in the form

$$\boldsymbol{G}(x, y, z, w) = \boldsymbol{C} + \left[\begin{matrix} I_1' + I_3' & I_4' & -D\sin\tau & I_5' \\ \beta I_2' + q_1 I_3' & q_1 I_4' & -\beta D\cos\tau & q_4 I_5' \\ I_6' & q_2 I_4' & I_8' & q_5 I_5' \\ I_7' & q_3 I_4' & I_9' & q_6 I_5' \end{matrix} + \text{small terms}\right]\begin{pmatrix} x \\ y \\ z \\ w\end{pmatrix}$$
$$+ \boldsymbol{E}(x, y, z, w, Q_*, \mu_*, \delta_*^2, \eta_*^2),$$

where

$$I_1' = C\left(\frac{\sigma}{\alpha}\sin\tau + \frac{\beta(p-1)}{4}\cos\tau\right), \quad I_2' = C\left(\frac{\sigma}{\alpha}\cos\tau - \frac{\beta(p-1)}{4}\sin\tau\right),$$

$$I_3' = \frac{\nu_1}{\alpha}B_p\mu_*\Theta^{\tilde\nu_1+(n-5)/2}Q_*^{(\nu_1-\sigma)/\alpha}, \quad I_4' = B_p,$$

$$I_5' = -\vartheta_3, \qquad I_6' = -\beta^2 I_1' - \beta I_2' + q_2 I_3',$$

$$I_7' = -\beta^3 I_2' + 3\beta^2 I_1' + 2\beta I_2' + q_3 I_3', \quad I_8' = \beta^2 D\sin\tau + \beta D\cos\tau,$$

$$I_9' = \beta^3 D\cos\tau - 3\beta^2 D\sin\tau - 2\beta D\cos\tau,$$

and where $\boldsymbol{C}$ is a constant vector independent of (x, y, z, w) which is bounded above by $M\Theta^{4-\sigma}$, and $|\boldsymbol{E}|$ is bounded independently of x, y, z, w, Q, μ, δ and η. Thus,

$$\boldsymbol{G}(x, y, z, w) = \boldsymbol{C} + L \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} + \boldsymbol{T}(x, y, z, w),$$

where L is a linear operator which is invertible; we shall prove this fact in Lemma A.1. If we define the operator $\boldsymbol{J}$ mapping $\mathbb{R}^4$ into itself by

$$\boldsymbol{J}(x, y, z, w) = -(L^{-1}\boldsymbol{C} + L^{-1}\boldsymbol{T}(x, y, z, w)),$$

then, provided that Q_* is sufficiently large, a direct calculation shows that $\boldsymbol{J}$ maps the set I into itself, where I is the ball

$$I = \{(x, y, z, w) : (x^2 + y^2 + z^2 + w^2)^{1/2} \le 4M(\det L)^{-1}\Theta^{4-\sigma}\}, \tag{5-16}$$

and $\det L$ is the determinant of L, which depends on $\sqrt{a_0^2 + b_0^2}$, β, D, α, B_p, ϑ_3 and ν_i $(i = 1, 2, 3, 4)$. We apply the Brouwer fixed point theorem to conclude that $\boldsymbol{J}$ has a fixed point in I. This point (x, y, z, w) satisfies $\boldsymbol{G}(x, y, z, w) = 0$ and

$$(x^2 + y^2 + z^2 + w^2)^{1/2} \le M'\Theta^{4-\sigma},$$

where M' is a constant defined in (5-16) and is independent of Q_*, μ_*, δ_*, η_* and Θ. By substituting for Q, μ, δ and η, then taking Θ to have the upper limiting value of $Q_*^{\sigma/((2-\sigma)\alpha)}$, we obtain (5-11)–(5-14). Therefore, we can find a solution to (5-1) such that (5-2)–(5-5) hold.

We have shown that (5-2)–(5-5) have a solution for each large fixed m. This yields a solution of (5-1) and also gives the proof of Theorem 1.1. Hence we have:

Theorem 5.1. *For $m \gg 1$ large and Q, μ, δ and η as given in* (5-11)–(5-14), *problem* (5-1) *admits a classical solution $w_{Q,\mu,\delta,\eta}(\theta)$. Moreover, there is $\Theta = |O(Q^{\sigma/((2-\sigma)\alpha)})|$ such that* (5-2)–(5-5) *hold.*

As a consequence, problem (1-7) admits infinitely many nonconstant positive solutions. Hence, we have proved Theorem 1.1.

Appendix

We will prove a lemma which was used in the previous sections.

Lemma A.1. *For the terms E_2 and $k_0(n)$ and the matrices K and L, which were defined in previous sections, we have*

(1) $E_2 \neq 0$,

(2) $p \in \left(\dfrac{n+3}{n-5}, p_c(n-1)\right) \implies pk_0(n-1) \geq k_0(n)$,

(3) $\det K \neq 0$,

(4) $\det L \neq 0$.

Proof. First, we show that $E_2 \neq 0$. It is known that

(A-1) $E_2 = (\tilde{\nu}_1+2)\tilde{\nu}_1(\tilde{\nu}_1+n-3)(\tilde{\nu}_1+n-1) - p(n-5-\alpha)(n-3-\alpha)(2+\alpha)\alpha.$

For convenience, we use n instead of $n-1$ and $\tilde{\nu}_1(n)$ instead of $\tilde{\nu}_1(n-1)$; i.e., we study the term

(A-2) $E_2 = (\tilde{\nu}_1+2)\tilde{\nu}_1(\tilde{\nu}_1+n-2)(\tilde{\nu}_1+n) - p(n-4-\alpha)(n-2-\alpha)(2+\alpha)\alpha.$

Let $f(\alpha) = p(n-4-\alpha)(n-2-\alpha)(2+\alpha)\alpha$. Through a simple computation, we get $f(\alpha)$ and its derivative $f'(\alpha)$:

$$f(\alpha) = \alpha^4 + (12-2n)\alpha^3 + (n^2-18n+52)\alpha^2 + (6n^2-52n+96)\alpha + 8(n-2)(n-4),$$

and

$$f'(\alpha) = 4\alpha^3 + (36-6n)\alpha^2 + (2n^2-36n+104)\alpha + (6n^2-52n+96).$$

We compute the roots of $f'(\alpha)$ to find its zero points: $\frac{1}{2}(n-6 \pm \sqrt{n^2+4})$ and $\frac{1}{2}(n-6)$. It is easy to see that $f(\alpha)$ is strictly increasing for $\alpha \in \left(0, \frac{1}{2}(n-6)\right)$ and decreasing for $\alpha \in \left(\frac{1}{2}(n-6), \frac{1}{2}(n-6+\sqrt{n^2+4})\right)$. We know $\alpha = 4/(p-1) < \frac{1}{2}(n-4)$ and $\frac{1}{2}(n-4) \in \left(\frac{1}{2}(n-6), \frac{1}{2}(n-6+\sqrt{n^2+4})\right)$. As a consequence, we can conclude

$$f(\alpha) \leq f\left(\tfrac{1}{2}(n-6)\right) = \tfrac{1}{16}n^4 - \tfrac{1}{2}n^2 + 1 \quad \text{for all } p \in \left(\frac{n+4}{n-4}, p_c(n)\right).$$

Let $g(x) = x(x+2)(x+n)(x+n-2) = x^4 + 2nx^3 + (n^2+2n-4)x^2 + (2n^2-4n)x$. We compute its derivative, $g'(x) = 4x^3 + 6nx^2 + (2n^2+4n-8)x + (2n^2-4n)$, and find $g'(x) > 0$ for $x > 0$ when $n \geq 5$. On the other hand, using $4\sqrt{N_3} > N_2$ for $p \in ((n+4)/(n-4), p_c(n))$, we find

$$\tilde{\nu}_1 > \tfrac{1}{2}\left(\sqrt{2(n^2-4n+8)} - (n-4)\right).$$

Therefore,

$$\text{(A-3)} \quad \begin{aligned} g(\tilde{\nu}_1) &\geq g\left(\tfrac{1}{2}\left(\sqrt{2(n^2-4n+8)} - (n-4)\right)\right) \\ &= 96 - 40n + 11n^2 - \tfrac{1}{2}n^3 + \tfrac{1}{16}n^4 + \sqrt{2}(24-4n+n^2)\sqrt{8-4n+n^2}. \end{aligned}$$

Comparing $\frac{1}{16}n^4 - \frac{1}{2}n^2 + 1$ and the right-hand side of (A-3), by direct computation, we can get

$$g\left(\tfrac{1}{2}\left(\sqrt{2(n^2-4n+8)} - (n-4)\right)\right) > \tfrac{1}{16}n^4 - \tfrac{1}{2}n^2 + 1 \quad \text{for } n \in (0, \infty).$$

As a result, $g(\tilde{\nu}_1) > f(\alpha)$. Hence, E_2 is nonzero.

Next, we prove $pk_0(n-1) \geq k_0(n)$ for $p \in ((n+3)/(n-5),\, p_c(n-1))$. According to the definition of $k_0(n)$, it is enough for us to show

$$p(n-5-\alpha)(n-3-\alpha) \geq (n-4-\alpha)(n-2-\alpha). \tag{A-4}$$

Using the relation $p = 4/\alpha + 1$, it is equivalent to show (after computation)

$$6\alpha^2 + (39-10n)\alpha + 4n^2 - 32n + 60 \geq 0. \tag{A-5}$$

It is known that (A-5) holds provided

$$\alpha \geq \tfrac{1}{12}\big(10n-39+\sqrt{4n^2-12n+81}\big) \quad \text{or} \quad \alpha \leq \tfrac{1}{12}\big(10n-39-\sqrt{4n^2-12n+81}\big).$$

On the other hand, since $p \in ((n+3)/(n-5),\, p_c(n-1))$, we have $\alpha < \frac{1}{2}(n-5)$. It is easy to show $\frac{1}{2}(n-5) \leq \frac{1}{12}\big(10n-39-\sqrt{4n^2-12n+81}\big)$ when $n \geq 5$. Hence, (A-5) holds. Therefore (A-4) holds.

Then, to show K is invertible, it is enough for us to show $B \neq 0$ or $A \neq 0$. Recall

$$B = (2n^2-12n-6)\beta + 8\beta^3 = (2(n-3)^2-24)\beta + 8\beta^3.$$

It is known that $2(n-3)^2 - 24 < 0$ only when $n = 6$. Since $\beta > 0$, we have $B \neq 0$ when $n \geq 7$. When $n = 6$, we find

$$A = \beta^4 - \tfrac{35}{2}\beta^2 - \tfrac{135}{16} - (1-\alpha)(3-\alpha)(2+\alpha)(4+\alpha), \qquad B = -6\beta + 8\beta^3.$$

If $B \neq 0$ for $n = 6$, we have that K is invertible, while if $B = 0$ for $n = 6$, then $A = -21 - (1-\alpha)(3-\alpha)(2+\alpha)(4+\alpha) < 0$ for $\alpha \in \big(0, \frac{1}{2}\big)$ and K is also invertible. Therefore, we have proved the third conclusion.

Finally, we show the matrix L is invertible. Recall that L is given by

$$L := \begin{bmatrix} I_1' + I_3' & I_4' & -D\sin\tau & I_5' \\ \beta I_2' + q_1 I_3' & q_1 I_4' & -\beta D\cos\tau & q_4 I_5' \\ I_6' & q_2 I_4' & I_8' & q_5 I_5' \\ I_7' & q_3 I_4' & I_9' & q_6 I_5' \end{bmatrix}, \tag{A-6}$$

where

$$I_1' = C\left(\frac{\sigma}{\alpha}\sin\tau + \frac{\beta(p-1)}{4}\cos\tau\right), \quad I_2' = C\left(\frac{\sigma}{\alpha}\cos\tau - \frac{\beta(p-1)}{4}\sin\tau\right),$$
$$I_3' = \frac{\nu_1}{\alpha} B_p \mu_* \Theta^{\tilde{\nu}_1+(n-5)/2} Q_*^{(\nu_1-\sigma)/\alpha}, \qquad I_4' = B_p,$$
$$I_5' = \vartheta_3, \qquad I_6' = -\beta^2 I_1' - \beta I_2' + q_2 I_3',$$
$$I_7' = -\beta^3 I_2' + 3\beta^2 I_1' + 2\beta I_2' + q_3 I_3', \qquad I_8' = \beta^2 D\sin\tau + \beta D\cos\tau,$$
$$I_9' = \beta^3 D\cos\tau - 3\beta^2 D\sin\tau - 2\beta D\cos\tau.$$

Using simple linear transformations, we see that

$$\begin{bmatrix} I_1'+I_3' & I_4' & -D\sin\tau & I_5' \\ \beta I_2'+q_1I_3' & q_1I_4' & -\beta D\cos\tau & q_4I_5' \\ I_6' & q_2I_4' & I_8' & q_5I_5' \\ I_7' & q_3I_4' & I_9' & q_6I_5' \end{bmatrix} \sim \begin{bmatrix} I_1' & I_4' & -D\sin\tau & I_5' \\ \beta I_2' & q_1I_4' & -\beta D\cos\tau & q_4I_5' \\ I_6'-q_2I_3' & q_2I_4' & I_8' & q_5I_5' \\ I_7'-q_3I_3' & q_3I_4' & I_9' & q_6I_5' \end{bmatrix}$$

$$\sim \begin{bmatrix} I_1' & -D\sin\tau & I_4' & I_5' \\ \beta I_2' & -\beta D\cos\tau & q_1I_4' & q_4I_5' \\ I_6'-q_2I_3' & I_8' & q_2I_4' & q_5I_5' \\ I_7'-q_3I_3' & I_9' & q_3I_4' & q_6I_5' \end{bmatrix} \sim \begin{bmatrix} I_1' & -D\sin\tau & I_4' & -I_5' \\ \beta I_2' & -\beta D\cos\tau & q_1I_4' & -q_4I_5' \\ 0 & 0 & I_{10}' & I_{11}' \\ 0 & 0 & I_{12}' & I_{13}', \end{bmatrix},$$

where

$$I_{10}' = q_2B_p+q_1B_p+\beta^2B_p, \qquad I_{11}' = q_5\vartheta_3+q_4\vartheta_3+\beta^2\vartheta_3,$$
$$I_{12}' = q_3B_p+\beta^2q_1B_p-3\beta^2B_p-2q_1B_p, \quad I_{13}' = q_6\vartheta_3+\beta^2q_4\vartheta_3-3\beta^2\vartheta_3-2q_4\vartheta_3.$$

Here we use the first column minus I_3'/I_4' times the second column in the first step, change the places of the second and third columns in the second step, and in the end, add the second row and β times the first row to the third row and add $-3\beta^2$ times the first row and β^2-2 times the second row to the fourth row. On the other hand, since

$$\det\begin{bmatrix} I_1' & -D\sin\tau \\ \beta I_2' & -\beta D\cos\tau \end{bmatrix} \neq 0,$$

to show that L is invertible, it is enough for us to prove that the 2×2 matrix

$$\begin{bmatrix} q_2+q_1+\beta^2 & q_5+q_4+\beta^2 \\ q_3+\beta^2q_1-3\beta^2-2q_1 & q_6+\beta^2q_4-3\beta^2-2q_4 \end{bmatrix}. \tag{A-7}$$

is invertible. It follows from the definitions of q_i $(i=1,2,3,4,5,6)$ and β that $q_2+q_1+\beta^2 = q_5+q_4+\beta^2 \neq 0$. Let

$$\chi_1 = q_3+\beta^2q_1-3\beta^2-2q_1, \quad \chi_2 = q_6+\beta^2q_4-3\beta^2-2q_4.$$

Then

$$\begin{aligned}\chi_1-\chi_2 &= q_3-q_6-(q_1-q_4)(2-\beta^2) \\ &= (\tilde\nu_1-\tilde\nu_2)\big((\tilde\nu_1+\tilde\nu_2)^2-\tilde\nu_1\tilde\nu_2+\tfrac12(3n-21)(\tilde\nu_1+\tilde\nu_2)+\tfrac14(3n^2-42n+135)+\beta^2\big) \\ &= (\tilde\nu_1-\tilde\nu_2)\big(\tfrac14(n^2-10n+25)-\tilde\nu_1\tilde\nu_2+\beta^2\big),\end{aligned}$$

where we are using the fact that $\tilde\nu_1+\tilde\nu_2 = -(n-5)$. It is known (from Section 2) that

$$\tilde\nu_1\tilde\nu_2 = \frac{n^2-10n+25}{4} - \frac{N_2+4\sqrt{N_3}}{4(p-1)^2}$$

and $\beta^2 = (4\sqrt{N_3} - N_2)/(4(p-1)^2)$, where N_2 and N_3 (with the dimension n being replaced by $n-1$) are defined in Section 2. Therefore,

$$\chi_1 - \chi_2 = (\tilde{\nu}_1 - \tilde{\nu}_2)\frac{2\sqrt{N_3}}{(p-1)^2} \neq 0.$$

Hence, (A-7) is invertible. □

Acknowledgements

The research of Z. Guo is supported by NSFC (11171092, 11571093) and Innovation Scientists and Technicians Troop Construction Projects of Henan Province (114200510011). The research of J. Wei is partially supported by NSERC of Canada. We thank the referee for a thorough reading of the manuscript and many thoughtful suggestions.

References

[Bidaut-Véron and Véron 1991] M.-F. Bidaut-Véron and L. Véron, "Nonlinear elliptic equations on compact Riemannian manifolds and asymptotics of Emden equations", *Invent. Math.* **106**:3 (1991), 489–539. MR Zbl

[Budd and Norbury 1987] C. Budd and J. Norbury, "Semilinear elliptic equations and supercritical growth", *J. Differential Equations* **68**:2 (1987), 169–197. MR Zbl

[Dancer et al. 2011] E. N. Dancer, Y. Du, and Z. Guo, "Finite Morse index solutions of an elliptic equation with supercritical exponent", *J. Differential Equations* **250**:8 (2011), 3281–3310. MR Zbl

[Dancer et al. 2012] E. N. Dancer, Z. Guo, and J. Wei, "Non-radial singular solutions of the Lane–Emden equation in $\mathbb{R}^N$", *Indiana Univ. Math. J.* **61**:5 (2012), 1971–1996. MR Zbl

[Dávila et al. 2014] J. Dávila, L. Dupaigne, K. Wang, and J. Wei, "A monotonicity formula and a Liouville-type theorem for a fourth order supercritical problem", *Adv. Math.* **258** (2014), 240–285. MR Zbl

[Farina 2007] A. Farina, "On the classification of solutions of the Lane–Emden equation on unbounded domains of $\mathbb{R}^N$", *J. Math. Pures Appl. (9)* **87**:5 (2007), 537–561. MR Zbl

[Gazzola and Grunau 2006] F. Gazzola and H.-C. Grunau, "Radial entire solutions for supercritical biharmonic equations", *Math. Ann.* **334**:4 (2006), 905–936. MR Zbl

[Gidas and Spruck 1981] B. Gidas and J. Spruck, "Global and local behavior of positive solutions of nonlinear elliptic equations", *Comm. Pure Appl. Math.* **34**:4 (1981), 525–598. MR Zbl

[Gui et al. 1992] C. Gui, W.-M. Ni, and X. Wang, "On the stability and instability of positive steady states of a semilinear heat equation in $\mathbb{R}^n$", *Comm. Pure Appl. Math.* **45**:9 (1992), 1153–1181. MR Zbl

[Guo 2002] Z. Guo, "On the symmetry of positive solutions of the Lane–Emden equation with supercritical exponent", *Adv. Differential Equations* **7**:6 (2002), 641–666. MR Zbl

[Guo 2014] Z. Guo, "Further study of entire radial solutions of a biharmonic equation with exponential nonlinearity", *Ann. Mat. Pura Appl. (4)* **193**:1 (2014), 187–201. MR Zbl

[Guo and Wei 2010] Z. Guo and J. Wei, "Qualitative properties of entire radial solutions for a biharmonic equation with supercritical nonlinearity", *Proc. Amer. Math. Soc.* **138**:11 (2010), 3957–3964. MR Zbl

[Hartman 1982] P. Hartman, *Ordinary differential equations*, 2nd ed., Birkhäuser, Boston, 1982. MR Zbl

[Johnson et al. 1993] R. A. Johnson, X. B. Pan, and Y. Yi, "Positive solutions of super-critical elliptic equations and asymptotics", *Comm. Partial Differential Equations* **18**:5–6 (1993), 977–1019. MR Zbl

[Joseph and Lundgren 1972/73] D. D. Joseph and T. S. Lundgren, "Quasilinear Dirichlet problems driven by positive sources", *Arch. Rational Mech. Anal.* **49** (1972/73), 241–269. MR Zbl

[Karageorgis 2009] P. Karageorgis, "Stability and intersection properties of solutions to the nonlinear biharmonic equation", *Nonlinearity* **22**:7 (2009), 1653–1661. MR Zbl

[Korevaar et al. 1999] N. Korevaar, R. Mazzeo, F. Pacard, and R. Schoen, "Refined asymptotics for constant scalar curvature metrics with isolated singularities", *Invent. Math.* **135**:2 (1999), 233–272. MR Zbl

[Lin 1998] C.-S. Lin, "A classification of solutions of a conformally invariant fourth order equation in $\mathbb{R}^n$", *Comment. Math. Helv.* **73**:2 (1998), 206–231. MR Zbl

[Wei and Xu 1999] J. Wei and X. Xu, "Classification of solutions of higher order conformally invariant equations", *Math. Ann.* **313**:2 (1999), 207–228. MR Zbl

[Zou 1995] H. Zou, "Symmetry of positive solutions of $\Delta u + u^p = 0$ in $\mathbb{R}^n$", *J. Differential Equations* **120**:1 (1995), 46–88. MR Zbl

Received October 2, 2015. Revised December 26, 2015.

ZONGMING GUO
DEPARTMENT OF MATHEMATICS
HENAN NORMAL UNIVERSITY
XINXIANG 453007
CHINA
gzm@htu.cn

JUNCHENG WEI
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BC V6T 1Z2
CANADA
jcwei@math.ubc.ca

WEN YANG
Current address:
CENTER FOR ADVANCED STUDY IN THEORETICAL SCIENCE
NATIONAL TAIWAN UNIVERSITY
NO. 1, SEC. 4
ROOSEVELT ROAD
TAIPEI 106
TAIWAN
math.yangwen@gmail.com
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BC V6T 1Z2
CANADA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.431

NATURAL COMMUTING OF VANISHING CYCLES AND THE VERDIER DUAL

DAVID B. MASSEY

We prove that the shifted vanishing cycles and nearby cycles commute with Verdier dualizing up to a *natural* isomorphism, even when the coefficients are not in a field.

1. Introduction

In this short, technical, paper, we prove a result whose full statement is missing from the literature, and which may be surprising even to some experts in the field. To state this result, we need to use technical notions and notations; references are [Kashiwara and Schapira 1990; Dimca 2004; Schürmann 2003; Massey 2003, Appendix B]. We should remark immediately that the definition that we use (see below) for the vanishing cycles is the standard one, which is shifted by one from the definition in [Kashiwara and Schapira 1990].

We fix a base ring, R, which is a commutative, regular, Noetherian ring, with finite Krull dimension (e.g., $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{C}$). Throughout this paper, by a topological space, we will mean a locally compact space. When we write that $\mathbf{A}^\bullet$ is complex of sheaves on a topological space, X, we mean that $\mathbf{A}^\bullet$ is an object in $D^b(X)$, the derived category of bounded complexes of sheaves of R-modules on X. When X is complex analytic, we may also require that $\mathbf{A}^\bullet$ is (complex) constructible, and write $\mathbf{A}^\bullet \in D^b_c(X)$. We remind the reader that constructibility includes the assumption that the stalks of all cohomology sheaves are finitely generated R-modules (so that, by our assumption on R, each stalk complex $\mathbf{A}^\bullet_x$, for $x \in X$, is perfect, i.e., is quasi-isomorphic to a bounded complex of finitely generated projective R-modules).

We let $\mathcal{D} = \mathcal{D}_X$ denote the Verdier dualizing operator on $D^b_c(X)$. We will always write simply $\mathcal{D}$, since the relevant topological space will always be clear.

Suppose that $f : X \to \mathbb{C}$ is a complex analytic function, where X is an arbitrary complex analytic space, and suppose that we have a complex of sheaves $\mathbf{A}^\bullet$ on X. We let ψ_f and ϕ_f denote the nearby and vanishing cycle functors, respectively. Henceforth, we shall always write these functors composed with a shift by -1, that is, we shall write $\psi_f[-1] := \psi_f \circ [-1]$ and $\phi_f[-1] := \phi_f \circ [-1]$. In order

MSC2010: 32B10, 32B15, 32C18, 32C35.
Keywords: vanishing cycles, nearby cycles, Verdier dual, constructible complexes.

to eliminate any possible confusion over indexing/shifting: with the definitions that we are using, if $F_{f,p}$ denotes the Milnor fiber of f at $p \in f^{-1}(0)$ inside the open ball $\mathring{B}_\epsilon(p)$ (using a local embedding into affine space), then, for all $k \in \mathbb{Z}$, $H^k(\psi_f[-1]\mathbf{A}^\bullet)_p \cong \mathbb{H}^{k-1}(F_{f,p}; \mathbf{A}^\bullet)$ and $H^k(\phi_f[-1]\mathbf{A}^\bullet)_p \cong \mathbb{H}^k(\mathring{B}_\epsilon(p), F_{f,p}; \mathbf{A}^\bullet)$.

The questions in which we are interested are:

(i) Do isomorphisms $\mathcal{D} \circ \psi_f[-1] \cong \psi_f[-1] \circ \mathcal{D}$ and $\mathcal{D} \circ \phi_f[-1] \cong \phi_f[-1] \circ \mathcal{D}$ exist even if R is *not* a field?

(ii) Do there exist such isomorphisms which are *natural*?

We show that the answer to both is *yes*.

Is this result known and/or surprising? Some references, such as [Brylinski and Monteiro Fernandes 1986; Dimca 2004; Massey 2003, Appendix B], state that there exist nonnatural isomorphisms, and require that the base ring is a field. Schürmann [2003, Corollary 5.4.4] proves the natural isomorphism exists on the stalk level, even when R is not a field.

In the l-adic algebraic context, Illusie [1994] proves that the Verdier dual and nearby cycles commute, up to natural isomorphism. M. Saito [1988; 1989] proves the analogous result in the complex analytic setting, with field coefficients. One can obtain our full result by combining Proposition 8.4.13, Proposition 8.6.3, and Exercise VIII.15 of [Kashiwara and Schapira 1990], though our proof here is completely different. In fact, our proof is similar to the discussion on duality of local Morse data following Remark 5.1.7 in [Schürmann 2003].

Recently, J. Schürmann proved in Proposition A.1 of the Appendix in [Brav et al. 2015], that the duality isomorphism constructed here fits with a corresponding (more complicated) duality isomorphism in Saito's theory of mixed Hodge modules.

Furthermore, this duality result and the half-space description used for the vanishing cycles have recently become very promising in the study of Donaldson–Thomas invariants of suitable moduli spaces (as in [Brav et al. 2015] and [Kontsevich and Soibelman 2011, Section 7]).

Our proof is relatively simple, and consists of three main steps: proving a small lemma about pairs of closed sets which cover a space, using a convenient characterization/definition of the vanishing cycles, and using that the stratified critical values of f are locally isolated. The nearby cycle result follows as a quick corollary of the result for vanishing cycles.

2. Two lemmas

We shall use $\simeq$ to denote natural isomorphisms of functors. If $\mathbf{A}^\bullet$ is a bounded complex, then, by $\operatorname{supp} \mathbf{A}^\bullet$, we mean, by definition, the closure of the union of the support of all (nontrivial) cohomology sheaves.

The following is an easy generalization of the fact that, if i is the inclusion of an open set, then $i^* \simeq i^!$ (see, for instance, [Dimca 2004, Corollary 3.2.12]).

Lemma 2.1. *Suppose Z is a locally compact subset of X, and let $j : Z \hookrightarrow X$ denote the inclusion. Let $D^b_Z(X)$ denote the full subcategory of $D^b(X)$ of complexes $\mathbf{A}^\bullet$ such that $Z \cap \operatorname{supp} \mathbf{A}^\bullet$ is an open subset of $\operatorname{supp} \mathbf{A}^\bullet$. Then, there is a natural isomorphism of functors $j^! \simeq j^*$ from $D^b_Z(X)$ to $D^b(Z)$.*

Proof. Let $\mathbf{A}^\bullet \in D^b_Z(X)$. We will show that the natural map $j^! \to j^*$ of functors from $D^b(X)$ to $D^b(Z)$ yields an isomorphism $j^!\mathbf{A}^\bullet \to j^*\mathbf{A}^\bullet$.

Let $Y := \operatorname{supp} \mathbf{A}^\bullet$. Let $m : Y \hookrightarrow X$, $\hat{m} : Z \cap Y \hookrightarrow Z$ and $\hat{j} : Z \cap Y \hookrightarrow Y$ denote the inclusions. Then,

$$j^!\mathbf{A}^\bullet \cong j^! m_* m^* \mathbf{A}^\bullet \cong \hat{m}_* \hat{j}^! m^* \mathbf{A}^\bullet \cong \hat{m}_! \hat{j}^* m^* \mathbf{A}^\bullet \cong j^* m_! m^* \mathbf{A}^\bullet \cong j^* \mathbf{A}^\bullet,$$

where we used, in order, that $\mathbf{A}^\bullet \cong m_* m^* \mathbf{A}^\bullet$, since Y is the support of $\mathbf{A}^\bullet$, Proposition 2.6.7 of [Kashiwara and Schapira 1990] on Cartesian squares, that $\hat{m}_* \simeq \hat{m}_!$ and $\hat{j}^! \simeq \hat{j}^*$, since $\hat{m}$ is a closed inclusion and $\hat{j}$ is an open inclusion, Proposition 2.6.7 of [Kashiwara and Schapira 1990] again, and that $\mathbf{A}^\bullet \cong m_! m^* \mathbf{A}^\bullet$. □

The lemma that we shall now prove certainly looks related to many propositions we have seen before, and may be known, but we cannot find a reference. The lemma tells us that, in our special case, the morphism of functors described in Proposition 3.1.9(iii) of [Kashiwara and Schapira 1990] is an isomorphism.

Lemma 2.2. *Let X be a locally compact space, and let Z_1 and Z_2 be closed subsets of X such that $X = Z_1 \cup Z_2$. Denote the inclusion maps by*

$$j_1 : Z_1 \hookrightarrow X, \quad j_2 : Z_2 \hookrightarrow X, \quad \hat{j}_1 : Z_1 \cap Z_2 \hookrightarrow Z_2,$$
$$\hat{j}_2 : Z_1 \cap Z_2 \hookrightarrow Z_1, \quad m = j_1 \hat{j}_2 = j_2 \hat{j}_1 : Z_1 \cap Z_2 \hookrightarrow X.$$

Then, we have the following natural isomorphisms

$$m^* j_{2!} j_2^! \simeq \hat{j}_1^* j_2^! \simeq \hat{j}_2^! j_1^* \simeq m^* j_{1*} \hat{j}_{2!} \hat{j}_2^! j_1^*. \tag{1}$$

Proof. Let $i_1 : X - Z_1 \hookrightarrow X$ and $i_2 : X - Z_2 \hookrightarrow X$ denote the open inclusions. We make use of Proposition 2.6.7 of [Kashiwara and Schapira 1990] on Cartesian squares repeatedly. We also use repeatedly that, if j is a closed inclusion, then $j_* \simeq j_!$ and $j^* j_* \simeq j^* j_! \simeq \mathrm{id}$.

We find

$$m^* j_{2!} j_2^! = (j_2 \hat{j}_1)^* j_{2!} j_2^! \simeq \hat{j}_1^* j_2^* j_{2!} j_2^! \simeq \hat{j}_1^* j_2^!,$$

which proves the first isomorphism in (1).

We also find

$$m^* j_{1*} \hat{j}_{2!} \hat{j}_2^! j_1^* = (j_1 \hat{j}_2)^* j_{1*} \hat{j}_{2!} \hat{j}_2^! j_1^* \simeq \hat{j}_2^* j_1^* j_{1*} \hat{j}_{2!} \hat{j}_2^! j_1^* \simeq \hat{j}_2^* \hat{j}_{2!} \hat{j}_2^! j_1^* \simeq \hat{j}_2^! j_1^*,$$

which proves the last isomorphism in (1).

It remains for us to demonstrate the middle isomorphism.

Let $l_2 : X - Z_2 = Z_1 - (Z_1 \cap Z_2) \to X$ denote the (open) inclusion. Then, we have the natural distinguished triangle

$$j_{2!}j_2{}^! \to \mathrm{id} \to l_{2*}l_2{}^* \xrightarrow{[1]},$$

which yields the distinguished triangle

$$\hat{j}_2^! j_1^* j_{2!}j_2{}^! \to \hat{j}_2^! j_1^* \to \hat{j}_2^! j_1^* l_{2*}l_2{}^* \xrightarrow{[1]} .$$

Now, $\hat{j}_2^! j_1^* j_{2!} j_2{}^! \simeq \hat{j}_2^! \hat{j}_{2!} \hat{j}_1^* j_2^! \simeq \hat{j}_1^* j_2^!$ and so, if we can show that $\hat{j}_2^! j_1^* l_{2*}l_2{}^* = 0$, then we will be finished.

This is easy. The support of $l_{2*}l_2{}^*$ is contained in Z_1; hence $j_1^* l_{2*}l_2{}^* \simeq j_1^! l_{2*}l_2{}^*$. Therefore,

$$\hat{j}_2^! j_1^* l_{2*}l_2{}^* \simeq \hat{j}_2^! j_1^! l_{2*}l_2{}^* \simeq \hat{j}_1^! j_2^! l_{2*}l_2{}^*,$$

and $j_2^! l_{2*} = 0$. □

3. The main theorem

Let $f : X \to \mathbb{C}$ be complex analytic, and let $\mathbf{A}^\bullet \in D_c^b(X)$. For any real number θ, let

$$Z_\theta := f^{-1}(e^{i\theta}\{v \in \mathbb{C} \mid \operatorname{Re} v \leq 0\})$$

and let

$$L_\theta := f^{-1}(e^{i\theta}\{v \in \mathbb{C} \mid \operatorname{Re} v = 0\}).$$

Let $j_\theta : Z_\theta \hookrightarrow X$ and $p : f^{-1}(0) \hookrightarrow X$ denote the inclusions.

By Lemma 1.3.2 of [Schürmann 2003], or following Exercise VIII.13 of [Kashiwara and Schapira 1990] (but reversing the inequality, and using a different shift), we define (or characterize up to natural isomorphism) the shifted vanishing cycles of $\mathbf{A}^\bullet$ along f to be

$$\phi_f[-1]\mathbf{A}^\bullet := p^* R\Gamma_{Z_0}(\mathbf{A}^\bullet) \simeq\ p^* j_{0!} j_0^! \mathbf{A}^\bullet.$$

In fact, for each θ, we define the shifted vanishing cycles of $\mathbf{A}^\bullet$ along f at the angle θ to be

$$\phi_f^\theta[-1]\mathbf{A}^\bullet := p^* j_{\theta!} j_\theta^! \mathbf{A}^\bullet.$$

There are the well-known natural isomorphisms $\widetilde{T}_f^\theta : \phi_f^0[-1] \to \phi_f^\theta[-1]$, induced by rotating $\mathbb{C}$ counterclockwise by an angle θ around the origin. The natural isomorphism $\widetilde{T}_f^{2\pi} : \phi_f[-1] \to \phi_f[-1]$ is the usual monodromy automorphism on the vanishing cycles.

In the proof of the main theorem below, we shall use that $\mathcal{D} \circ (-)^* \simeq (-)^! \circ \mathcal{D}$ always holds; we shall also use that $\mathcal{D} \circ (-)^! \simeq (-)^* \circ \mathcal{D}$ holds in our context, but note that this uses biduality, i.e., that $\mathcal{D}\circ\mathcal{D}\simeq \mathrm{id}$, for subanalytically constructible complexes of sheaves (see of [Schürmann 2003, Corollary 2.2.7] or [Kashiwara and Schapira 1990, Exercise VIII.3]), which uses the assumption that the commutative base ring R is regular, Noetherian of finite Krull dimension, so that, for such a subanalytically constructible complex of sheaves, all stalk complexes are perfect.

We now prove the main theorem.

Theorem 3.1. *There is a natural isomorphism*

$$\phi_f[-1]\circ\mathcal{D} \simeq \mathcal{D}\circ\phi_f[-1]$$

of functors from $D^b_c(X)$ to $D^b_c(f^{-1}(0))$.

Proof. Let m denote the inclusion of $L_0 = L_\pi$ into X. Now, apply Lemma 2.2 to $X = Z_0 \cup Z_\pi$, and conclude that

$$m^* j_{0!}\, j_0^! \simeq \hat{j}_0^!\, j_\pi^*.$$

Dualizing, we obtain

$$\mathcal{D}(m^* j_{0!}\, j_0^!) \simeq \mathcal{D}(\hat{j}_0^!\, j_\pi^*) \simeq \hat{j}_0^*\, j_\pi^!\mathcal{D} \simeq m^* j_{\pi\,!}\, j_\pi^!\mathcal{D}, \tag{2}$$

where the second isomorphism uses that $\mathcal{D}$ "commutes" with the standard operations, and the last isomorphism results from using that $m = j_\pi\,\hat{j}_0$.

Let q denote the inclusion of $f^{-1}(0)$ into $L_0 = L_\pi$, so that the inclusion p equals mq. Applying q^* to (2), we obtain

$$q^*\mathcal{D}(m^* j_{0!}\, j_0^!) \simeq q^*m^* j_{\pi\,!}\, j_\pi^!\mathcal{D} \simeq p^* j_{\pi\,!}\, j_\pi^!\mathcal{D} = \phi_f^\pi[-1]\circ\mathcal{D} \simeq \phi_f[-1]\circ\mathcal{D},$$

where, in the last step, we used the natural isomorphism $\big(\widetilde{T}_f^\pi\big)^{-1}$.

As $\mathcal{D}(q^!m^* j_{0!}\, j_0^!) \simeq q^*\mathcal{D}(m^* j_{0!}\, j_0^!)$, it remains for us to show that $q^!m^* j_{0!} j_0^!$ is naturally isomorphic to $q^*m^* j_{0!}\, j_0^! \simeq p^* j_{0!}\, j_0^! \simeq \phi_f[-1]$. This will follow from Lemma 2.1, once we show that, for all $\mathbf{A}^\bullet \in D^b_c(X)$, $f^{-1}(0)\cap \operatorname{supp}(m^* j_{0!} j_0^!\mathbf{A}^\bullet)$ is an open subset of $\operatorname{supp}(m^* j_{0!} j_0^!\mathbf{A}^\bullet)$.

Suppose that $x \in f^{-1}(0)\cap\operatorname{supp}(m^* j_{0!} j_0^1\mathbf{A}^\bullet)$. We need to show that there exists an open neighborhood $\mathcal{W}$ of x in X such that $\mathcal{W}\cap\operatorname{supp}(m^* j_{0!} j_0^1\mathbf{A}^\bullet) \subseteq f^{-1}(0)$.

Fix a Whitney stratification of X, with respect to which $\mathbf{A}^\bullet$ is constructible. Then, select $\mathcal{W}$ so that all of the stratified critical points of f, inside $\mathcal{W}$, are contained in $f^{-1}(0)$. Suppose that there were a point $y \in \mathcal{W}$ such that $y \in f^{-1}(L_0) - f^{-1}(0)$ and the stalk cohomology of $m^* j_{0!}\, j_0^!\mathbf{A}^\bullet$ at y is nonzero. Then, by definition, y would be a point in the support $\phi_{f-f(y)}[-1]\mathbf{A}^\bullet$, which, again, is contained in the stratified critical locus of f and, hence, is contained in $f^{-1}(0)$. This contradiction concludes the proof. □

We continue to let $p : f^{-1}(0) \hookrightarrow X$ denote the closed inclusion, and now let $i : X - f^{-1}(0) \hookrightarrow X$ denote the open inclusion. Consider the two fundamental distinguished triangles related to the nearby and vanishing cycles:

$$p^*[-1] \xrightarrow{\text{comp}} \psi_f[-1] \xrightarrow{\text{can}} \phi_f[-1] \xrightarrow{[1]}$$

and

$$\phi_f[-1] \xrightarrow{\text{var}} \psi_f[-1] \to p^![1] \xrightarrow{[1]} .$$

The morphisms comp, can, and var are usually referred to as the *comparison map*, *canonical map*, and *variation map*. As $p^*i_! = 0 = p^!i_*$ and as $\psi_f[-1]$ depends only on the complex outside of $f^{-1}(0)$, the top triangle, applied to $i_!i^!$ and the bottom triangle applied to i_*i^* yield natural isomorphisms

$$\alpha : \psi_f[-1] \xrightarrow{\simeq} \phi_f[-1]i_!i^! \quad \text{and} \quad \beta : \phi_f[-1]i_*i^* \xrightarrow{\simeq} \psi_f[-1].$$

Corollary 3.2. *There is a natural isomorphism*

$$\psi_f[-1] \circ \mathcal{D} \simeq \mathcal{D} \circ \psi_f[-1]$$

of functors from $D_c^b(X)$ to $D_c^b(f^{-1}(0))$.

Proof. $\psi_f[-1] \circ \mathcal{D} \simeq \phi_f[-1]i_!i^! \circ \mathcal{D} \simeq \mathcal{D} \circ \phi_f[-1]i_*i^* \simeq \mathcal{D} \circ \psi_f[-1]$. □

Acknowledgement

We thank Jörg Schürmann for providing valuable comments and references for related results.

References

[Brav et al. 2015] C. Brav, V. Bussi, D. Dupont, D. Joyce, and B. Szendröi, "Symmetries and stabilization for sheaves of vanishing cycles", *J. Singul.* **11** (2015), 85–151. MR Zbl

[Brylinski and Monteiro Fernandes 1986] J.-L. Brylinski and T. Monteiro Fernandes, *Géométrie et analyse microlocales*, Astérisque **140–141**, Société Mathématique de France, Paris, 1986. MR Zbl

[Dimca 2004] A. Dimca, *Sheaves in topology*, Springer, Berlin, 2004. MR Zbl

[Illusie 1994] L. Illusie, "Autour du théorème de monodromie locale", pp. 9–57 in *Périodes p-adiques* (Bures-sur-Yvette, 1988), edited by J. Fontaine, Astérisque **223**, Société Mathématique de France, Paris, 1994. MR Zbl

[Kashiwara and Schapira 1990] M. Kashiwara and P. Schapira, *Sheaves on manifolds*, Grundlehren der Mathematischen Wissenschaften **292**, Springer, Berlin, 1990. MR Zbl

[Kontsevich and Soibelman 2011] M. Kontsevich and Y. Soibelman, "Cohomological Hall algebra, exponential Hodge structures and motivic Donaldson–Thomas invariants", *Commun. Number Theory Phys.* **5**:2 (2011), 231–352. MR Zbl

[Massey 2003] D. B. Massey, *Numerical control over complex analytic singularities*, Memoirs of the American Mathematical Society **163**:778, American Mathematical Society, Providence, RI, 2003. MR Zbl

[Saito 1988] M. Saito, "Modules de Hodge polarisables", *Publ. Res. Inst. Math. Sci.* **24**:6 (1988), 849–995. MR Zbl

[Saito 1989] M. Saito, "Duality for vanishing cycle functors", *Publ. Res. Inst. Math. Sci.* **25**:6 (1989), 889–921. MR Zbl

[Schürmann 2003] J. Schürmann, *Topology of singular spaces and constructible sheaves*, Monografie Matematyczne **63**, Birkhäuser, Basel, 2003. MR Zbl

Received January 7, 2016. Revised March 9, 2016.

DAVID B. MASSEY
DEPARTMENT OF MATHEMATICS
NORTHEASTERN UNIVERSITY
360 HUNTINGTON AVENUE
BOSTON, MA 02115
UNITED STATES
d.massey@neu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.439

THE NEF CONES OF AND MINIMAL-DEGREE CURVES IN THE HILBERT SCHEMES OF POINTS ON CERTAIN SURFACES

ZHENBO QIN AND YUPING TU

We determine the nef cones of the Hilbert schemes of points on certain surfaces X with $h^1(X, \mathcal{O}_X) = 0$. Then we apply the results to Hirzebruch surfaces, and study the minimal-degree curves in the Hilbert schemes of points on Hirzebruch surfaces. Our results generalize those in Li, Qin, and Zhang (2003).

1. Introduction

Hilbert schemes are classical objects in algebraic geometry, and have been studied extensively since their constructions by Grothendieck. Hilbert schemes of points on smooth surfaces are known to be smooth and irreducible, and have deep connections with combinatorics, representation theory and string theory. Ample divisors on these Hilbert schemes were considered in [Beltrametti and Sommese 1991; 1993; Catanese and Gœttsche 1990]. The nef cones of the Hilbert schemes of points on smooth surfaces were first investigated in [Li et al. 2003] when the surface is the projective plane. Recently, these nef cones were further understood in [Arcara et al. 2013; Bertram and Coskun 2013; Bolognese et al. 2015] via Bridgeland stability.

In this paper, we generalize the methods and results in [Li et al. 2003], and prove a structure theorem for the nef cones of the Hilbert schemes of points on certain surfaces. To state our result, let X be a smooth projective complex surface. The nef cone of X is the span of the nef divisors on X. We use $\mathrm{NE}(X)$ to denote the cone spanned by all the effective curves in X. It is well-known that $\mathrm{NE}(X)$ is dual to the nef cone of X. Let $X^{[n]}$ be the Hilbert scheme of points in X. By [Fogarty 1968; Iarrobino 1977], $X^{[n]}$ is a smooth irreducible variety of dimension $2n$.

Theorem 1.1. *Let $n \geq 2$, and let the surface X satisfy $h^1(X, \mathcal{O}_X) = 0$. Assume that the nef cone of X is the span of the divisors $F_1, \ldots, F_t$, and the cone* $\mathrm{NE}(X)$ *is the*

Qin partially supported by a grant from the Simons Foundation. Tu partially supported by NSFC grant 11171258.
MSC2010: 14C05.
Keywords: nef cones, Hilbert schemes.

span of the curves $C_1, \dots, C_t$ *with* $F_i \cdot C_j = \delta_{i,j}$ *for all* i *and* j. *Assume further that*

$$\mathcal{O}_X\left((n-1)\sum_{i=1}^{t} F_i\right)$$

is $(n-1)$-*very ample. Then*

(i) *the nef cone of the Hilbert scheme* $X^{[n]}$ *is spanned by*

$$D_{F_1}, \dots, D_{F_t}, (n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2; \tag{1-1}$$

(ii) *the cone* $\mathrm{NE}(X^{[n]})$ *is spanned by the classes*

$$\beta_{C_1} - (n-1)\beta_n, \dots, \beta_{C_t} - (n-1)\beta_n, \beta_n. \tag{1-2}$$

In the above theorem, B_n denotes the boundary divisor of the Hilbert scheme $X^{[n]}$ consisting of the elements $\xi \in X^{[n]}$ which are not smooth as subschemes of $X^{[n]}$, and β_n is the minimal curve class contracted by the Hilbert–Chow morphism $X^{[n]} \to X^{(n)}$ sending an element $\xi \in X^{[n]}$ to its support (with multiplicities) in the n-th symmetric product $X^{(n)}$ of X. We refer to (2-4), (2-3) and Definition 2.1 for the definitions of D_F, β_C and $(n-1)$-very ampleness, respectively. Theorem 1.1 is proved in Section 2. Our main idea in the proof of Theorem 1.1 is to construct curves in $X^{[n]}$ which provide us with information about the nef divisors in $X^{[n]}$.

In Section 3, we apply Theorem 1.1 to the case when X is a Hirzebruch surface, and recover a result in [Bertram and Coskun 2013]. Moreover, when X is a Hirzebruch surface, we classify all the curves in the Hilbert scheme $X^{[n]}$ whose homology classes are contained in the list (1-2). These curves have minimal degree in the sense that their intersection numbers with certain very ample divisors in $X^{[n]}$ are all equal to 1. We compute the normal bundles of these curves, and prove that their moduli spaces are unobstructed, i.e., are smooth with the expected dimensions.

Conventions. Let $0 \le k \le n$ and V be an n-dimensional vector space. We use the Grassmannian $\mathbb{G}(V, k)$ to denote the set of all k-dimensional quotients of V, or equivalently, the set of all $(n-k)$-dimensional subspaces of V. Also, we take $\mathbb{P}(V) = \mathbb{G}(V, 1)$. So the set of lines in $\mathbb{P}(V)$ is the Grassmannian $\mathbb{G}(V, 2)$.

2. The nef cones of the Hilbert schemes of points on certain surfaces

In this section, we study the nef cones of the Hilbert schemes of points on certain surfaces with $h^1(X, \mathcal{O}_X) = 0$. Our goal is to prove Theorem 1.1.

Let X be a smooth projective complex surface, and $X^{[n]}$ be the Hilbert scheme of points in X. An element in $X^{[n]}$ is represented by a length-n 0-dimensional closed subscheme ξ of X. For $\xi \in X^{[n]}$, let I_ξ be the corresponding sheaf of ideals and $\mathcal{O}_\xi$ the structure sheaf. The subset

$$B_n = \{\xi \in X^{[n]} \mid |\operatorname{Supp}(\xi)| < n\} \tag{2-1}$$

is defined to be the boundary of $X^{[n]}$. Let C be a real surface in X, and fix distinct points $x_1, \dots, x_{n-1} \in X$ which are not contained in C. Define

$$\beta_n = \{\xi + x_2 + \cdots + x_{n-1} \in X^{[n]} \mid \operatorname{Supp}(\xi) = \{x_1\}\}, \tag{2-2}$$

$$\beta_C = \{x + x_1 + \cdots + x_{n-1} \in X^{[n]} \mid x \in C\}, \tag{2-3}$$

$$D_C = \{\xi \in X^{[n]} \mid \operatorname{Supp}(\xi) \cap C \neq \varnothing\}. \tag{2-4}$$

Note that β_C is a curve, and D_C a divisor, in $X^{[n]}$ when C is a smooth algebraic curve in X. We extend the notions β_C and D_C to all the divisors C in X by linearity.

Nakajima [1997] and Grojnowski [1996] geometrically constructed a Heisenberg algebra action on the cohomology of the Hilbert schemes $X^{[n]}$. Let $H^*(X)$ be the total cohomology of X with $\mathbb{C}$-coefficients. Denote the Heisenberg operators of Nakajima and Grojnowski by $\mathfrak{a}_m(\alpha)$ where $m \in \mathbb{Z}$ and $\alpha \in H^*(X)$. Set

$$\mathbb{H}_X = \bigoplus_{n=0}^{+\infty} H^*(X^{[n]}).$$

Then the space $\mathbb{H}_X$ is an irreducible representation of the Heisenberg algebra generated by the operators $\mathfrak{a}_m(\alpha)$ with the highest weight vector being

$$|0\rangle = 1 \in H^*(X^{[0]}) = \mathbb{C}.$$

It follows that the n-th component $H^*(X^{[n]})$ in the Fock space $\mathbb{H}_X$ is linearly spanned by the *Heisenberg monomial classes*

$$\mathfrak{a}_{-n_1}(\alpha_1) \cdots \mathfrak{a}_{-n_k}(\alpha_k)|0\rangle,$$

where $k \geq 0$, $n_1, \dots, n_k > 0$ and $n_1 + \cdots + n_k = n$. We have

$$\beta_n = \mathfrak{a}_{-2}(x)\mathfrak{a}_{-1}(x)^{n-2}|0\rangle, \tag{2-5}$$

$$\beta_C = \mathfrak{a}_{-1}(C)\mathfrak{a}_{-1}(x)^{n-1}|0\rangle, \tag{2-6}$$

$$B_n = \frac{1}{(n-2)!}\mathfrak{a}_{-1}(1_X)^{n-2}\mathfrak{a}_{-2}(1_X)|0\rangle, \tag{2-7}$$

$$D_C = \frac{1}{(n-1)!}\mathfrak{a}_{-1}(1_X)^{n-1}\mathfrak{a}_{-1}(C)|0\rangle, \tag{2-8}$$

where x and 1_X denote the cohomology classes corresponding to a point $x \in X$ and the surface X, respectively. Abusing notation, we also use C to denote the cohomology class corresponding to the real surface C.

The following important definition is from [Beltrametti and Sommese 1991].

Definition 2.1. Let $n \geq 1$. A line bundle L on the surface X is $(n-1)$-*very ample* if the restriction $H^0(X, L) \to H^0(X, \mathcal{O}_\xi \otimes L)$ is surjective for every $\xi \in X^{[n]}$.

The concept of $(n-1)$-very ampleness relates $X^{[n]}$ to a Grassmannian as follows. The surjective map in Definition 2.1 represents an element in $\mathbb{G}(H^0(X, L), n)$. So if L is $(n-1)$-very ample, then we obtain a morphism

$$\varphi_n(L) : X^{[n]} \to \mathbb{G}(H^0(X, L), n). \tag{2-9}$$

Let $h = h^0(X, L)$, and let $\mathfrak{P} : \mathbb{G}(\mathbb{C}^h, n) \to \mathbb{P}\big(\big(\bigwedge^{h-n} \mathbb{C}^h\big)^*\big)$ be the Plücker embedding. Then we see from the Appendix of [Beltrametti and Sommese 1991] that

$$(\mathfrak{P} \circ \varphi_n(L))^* \mathcal{H} = \mathcal{O}_{X^{[n]}}(D_{c_1(L)} - B_n/2), \tag{2-10}$$

where $\mathcal{H}$ is the hyperplane line bundle over the projective space $\mathbb{P}\big(\big(\bigwedge^{h-n} \mathbb{C}^h\big)^*\big)$.

Lemma 2.2. *If L is $(n-1)$-very ample, then the divisor $D_{c_1(L)} - B_n/2$ is nef. If L is n-very ample, then the divisor $D_{c_1(L)} - B_n/2$ is very ample.* □

The first statement in Lemma 2.2 follows immediately from (2-10), and the second statement was proved in [Catanese and Gœttsche 1990].

In $X^{[n]} \times X$, we have the universal codimension-2 subscheme

$$\mathcal{Z}_n = \{(\xi, x) \in X^{[n]} \times X \mid x \in \operatorname{Supp}(\xi)\} \subset X^{[n]} \times X. \tag{2-11}$$

Define the incidence variety $X^{[n-1,n]} = \{(\xi, \eta) \in X^{[n-1]} \times X^{[n]} \mid \xi \subset \eta\}$. It is well-known [Cheah 1998; Tikhomirov 1994] that $X^{[n-1,n]}$ is smooth and of dimension $2n$. Define

$$\begin{aligned} f_n &: X^{[n-1,n]} \to X^{[n-1]} && \text{with } f_n(\xi, \eta) = \xi, \\ g_n &: X^{[n-1,n]} \to X^{[n]} && \text{with } g_n(\xi, \eta) = \eta, \\ \rho &: X^{[n-1,n]} \to X && \text{with } \rho(\xi, \eta) = \operatorname{Supp}(I_\xi / I_\eta). \end{aligned}$$

Set $\phi_n = (f_n, \rho) : X^{[n-1,n]} \to X^{[n-1]} \times X$. By Proposition 2.2 in [Ellingsrud and Strømme 1998], ϕ_n is the blowing-up morphism of $X^{[n-1]} \times X$ along $\mathcal{Z}_{n-1}$.

Next, let C be an irreducible curve in X. Let $\xi = x_1 + \cdots + x_{n-1} \in X^{[n-1]}$, where $x_1, \ldots, x_{n-1}$ are distinct smooth points on C. Let $(C + \xi)$ be the closure of $(C - \operatorname{Supp}(\xi)) + \xi$ in $X^{[n]}$. Alternatively, consider

$$\begin{array}{ccccccc} \widetilde{C}_\xi & \subset & \widetilde{X}_\xi & \subset & X^{[n-1,n]} & \xrightarrow{g_n} & X^{[n]} \\ \big\downarrow & & \big\downarrow & & \big\downarrow \phi_n & & \\ \{\xi\} \times C & \subset & \{\xi\} \times X & \subset & X^{[n-1]} \times X & & \end{array} \tag{2-12}$$

where $\widetilde{C}_\xi$ and $\widetilde{X}_\xi$ are the strict transforms of $\{\xi\} \times C$ and $\{\xi\} \times X$ in $X^{[n-1,n]}$, respectively. Since ϕ_n is the blowing-up morphism of $X^{[n-1]} \times X$ along $\mathcal{Z}_{n-1}$, it follows that $\widetilde{X}_\xi$ is isomorphic to the blowup of $\{\xi\} \times X \cong X$ at $x_1, \dots, x_{n-1}$. For $1 \le i \le (n-1)$, let E_i be the exceptional divisor in $\widetilde{X}_\xi$ over x_i. Then we obtain

$$(\phi_n|_{\widetilde{X}_\xi})^*(\{\xi\} \times C) = \widetilde{C}_\xi + \sum_{i=1}^{n-1} E_i \tag{2-13}$$

in the Chow group $A_1(\widetilde{X}_\xi)$. Notice that $g_n(\widetilde{C}_\xi) = (C + \xi)$ and

$$g_n(E_i) = M_2(x_i) + x_1 + \cdots + x_{i-1} + x_{i+1} + \cdots + x_{n-1}.$$

In fact, since $g_n|_{\widetilde{X}_\xi} : \widetilde{X}_\xi \to g_n(\widetilde{X}_\xi)$ is an isomorphism, we have

$$(g_n|_{\widetilde{X}_\xi})_*(\widetilde{C}_\xi) = (C + \xi) \quad \text{and} \quad (g_n|_{\widetilde{X}_\xi})_*(E_i) = \beta_n. \tag{2-14}$$

Lemma 2.3. *With the above notation*, $(C + \xi) = \beta_C - (n-1)\beta_n$ *in* $A_1(X^{[n]})$.

Proof. Choose two smooth curves C_1 and C_2 in X such that $C = C_1 - C_2$ in $A_1(X)$ and $\mathrm{Supp}(\xi) \cap (C_1 \cup C_2) = \varnothing$. Then in $A_1(X^{[n]})$, we have

$$\begin{aligned} (g_n|_{\widetilde{X}_\xi})_*&(\phi_n|_{\widetilde{X}_\xi})^*(\{\xi\} \times C) \\ &= (g_n|_{\widetilde{X}_\xi})_*(\phi_n|_{\widetilde{X}_\xi})^*(\{\xi\} \times C_1) - (g_n|_{\widetilde{X}_\xi})_*(\phi_n|_{\widetilde{X}_\xi})^*(\{\xi\} \times C_2) \\ &= (C_1 + \xi) - (C_2 + \xi) \\ &= \beta_{C_1} - \beta_{C_2} = \beta_C. \end{aligned} \tag{2-15}$$

On the other hand, applying (2-13) and (2-14), we conclude that

$$(g_n|_{\widetilde{X}_\xi})_*(\phi_n|_{\widetilde{X}_\xi})^*(\{\xi\} \times C) = (g_n|_{\widetilde{X}_\xi})_*\left(\widetilde{C}_\xi + \sum_{i=1}^{n-1} E_i\right) = (C + \xi) + (n-1)\beta_n.$$

Combining this with (2-15), we see that $(C + \xi) = \beta_C - (n-1)\beta_n$ in $A_1(X^{[n]})$. $\square$

Lemma 2.4. *Let F be a divisor on X. If $D_F - d(B_n/2)$ is nef, then $d \ge 0$ and $F \cdot C \ge d(n-1)$ for every irreducible curve $C \subset X$. In particular, F is nef.*

Proof. Note that $D_F \cdot \beta_n = 0$ and $B_n \cdot \beta_n = -2$. Thus, we have

$$0 \le (D_F - d(B_n/2)) \cdot \beta_n = d.$$

Since $D_F \cdot \beta_C = F \cdot C$ and $B_n \cdot \beta_C = 0$, we conclude from Lemma 2.3 that

$$0 \le (D_F - d(B_n/2)) \cdot (\beta_C - (n-1)\beta_n) = F \cdot C - d(n-1). \qquad \square$$

Lemma 2.5. *Let F be a divisor in X. Let C be a smooth rational curve in X, and consider the n-th symmetric product $C^{(n)} = \mathrm{Hilb}^n(C) \subset X^{[n]}$. Then*

(i) *every line in $C^{(n)} \cong \mathbb{P}^n$ is homologous to $\beta_C - (n-1)\beta_n$;*

(ii) $\mathcal{O}_{X^{[n]}}(D_F)|_{C^{(n)}} = \mathcal{O}_{C^{(n)}}(C \cdot F)$ *and* $\mathcal{O}_{X^{[n]}}(B_n/2)|_{C^{(n)}} = \mathcal{O}_{C^{(n)}}(n-1)$.

Proof. (i) Let $x_1, \dots, x_{n-1} \in C$ be distinct, and put $\gamma = C + (x_1 + \cdots + x_{n-1})$. Then γ is a line in the projective space $C^{(n)} \cong \mathbb{P}^n$. By Lemma 2.3,

$$\gamma \sim \beta_C - (n-1)\beta_n,$$

where $\sim$ denotes homologous relation. So every line in $C^{(n)} \cong \mathbb{P}^n$ is homologous to the class $\beta_C - (n-1)\beta_n$.

(ii) Since $\gamma \cdot D_F|_{C^{(n)}} = \gamma \cdot D_F = (\beta_C - (n-1)\beta_n) \cdot D_F = C \cdot F$, we get

$$\mathcal{O}_{X^{[n]}}(D_F)|_{C^{(n)}} = \mathcal{O}_{C^{(n)}}(C \cdot F).$$

Using a similar method, we obtain $\mathcal{O}_{X^{[n]}}(B_n/2)|_{C^{(n)}} = \mathcal{O}_{C^{(n)}}(n-1)$. □

In the rest of the paper, we assume that $h^1(X, \mathcal{O}_X) = 0$. Then

$$\mathrm{Pic}(X^{[n]}) \cong \mathrm{Pic}(X) \oplus \mathbb{Z} \cdot (B_n/2) \tag{2-16}$$

by [Fogarty 1973]. Under this isomorphism, the divisor $D_C \in \mathrm{Pic}(X^{[n]})$ corresponds to $C \in \mathrm{Pic}(X)$. Let $\{\alpha_1, \dots, \alpha_s\}$ be a linear basis of $H^2(X)$. Then

$$\{D_{\alpha_1}, \dots, D_{\alpha_s}, B_n\} \tag{2-17}$$

is a linear basis of $H^2(X^{[n]})$. Represent $\alpha_1, \dots, \alpha_s$ by real surfaces $C_1, \dots, C_s \subset X$, respectively. Then a linear basis of $H_2(X^{[n]})$ is given by

$$\{\beta_{C_1}, \dots, \beta_{C_s}, \beta_n\}. \tag{2-18}$$

We are now ready to prove our main result in this paper.

Theorem 2.6. *Let $n \ge 2$, and let the surface X satisfy $h^1(X, \mathcal{O}_X) = 0$. Assume that the nef cone of X is the span of the divisors $F_1, \dots, F_t$, and the cone* $\mathrm{NE}(X)$ *is the span of the curves $C_1, \dots, C_t$ with $F_i \cdot C_j = \delta_{i,j}$ for all i and j. Assume further that $\mathcal{O}_X\big((n-1)\sum_{i=1}^t F_i\big)$ is $(n-1)$-very ample. Then*

(i) *the nef cone of the Hilbert scheme $X^{[n]}$ is spanned by*

$$D_{F_1}, \dots, D_{F_t}, (n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2; \tag{2-19}$$

(ii) *the cone* $\mathrm{NE}(X^{[n]})$ *is spanned by the classes*

$$\beta_{C_1} - (n-1)\beta_n, \dots, \beta_{C_t} - (n-1)\beta_n, \beta_n. \tag{2-20}$$

Proof. (i) For $1 \le j \le n$, let $p_j : X^n \to X$ be the projection to the j-th factor. Let $X^{(n)}$ be the n-th symmetric product of X, and let $\upsilon_n : X^n \to X^{(n)}$ be the quotient map. Let $\rho_n : X^{[n]} \to X^{(n)}$ be the Hilbert–Chow morphism sending an element

$\xi \in X^{[n]}$ to its support (with multiplicities) in $X^{(n)}$. For each F_i, there exists a divisor H_i on $X^{(n)}$ such that

$$\rho_n^* H_i = D_{F_i}, \qquad \upsilon_n^* H_i = \sum_{j=1}^{n} p_j^* F_i.$$

It follows that since F_i is nef, the divisor D_{F_i} is nef as well. Since the line bundle $\mathcal{O}_X\big((n-1)\sum_{i=1}^{t} F_i\big)$ is $(n-1)$-very ample, we conclude from Lemma 2.2 that $(n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2$ is a nef divisor. Thus, the cone $\boldsymbol{C}_1$ spanned by the divisors in (2-19) is contained in the nef cone of $X^{[n]}$.

Conversely, assume that $D_F - dB_n/2$ is a nef divisor on $X^{[n]}$. Let $F = \sum_{i=1}^{t} a_i F_i$. By Lemma 2.4, $d \ge 0$ and $F \cdot C \ge d(n-1)$ for every irreducible curve $C \subset X$. So

$$a_i = F \cdot C_i \ge d(n-1)$$

for every i. Now the nef divisor $D_F - dB_n/2$ can be written as

$$\sum_{i=1}^{t} a_i D_{F_i} - dB_n/2 = \sum_{i=1}^{t} (a_i - d(n-1)) D_{F_i} + d\left((n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2\right).$$

Therefore, $D_F - dB_n/2 \in \boldsymbol{C}_1$. It follows that $\boldsymbol{C}_1$ is the nef cone of $X^{[n]}$.

(ii) First of all, note that since the divisor F_i is nef and $F_i \cdot C_i = 1$, the curve C_i contains at least one reduced irreducible component. So the curve $\beta_{C_i} - (n-1)\beta_n$ is well-defined, and the cone $\boldsymbol{C}_2$ spanned by the curves in (2-20) is contained in the cone $\mathrm{NE}(X^{[n]})$. Conversely, assume that $\sum_{i=1}^{t} b_i \beta_{C_i} + c\beta_n$ is contained in $\mathrm{NE}(X^{[n]})$. Then $\big(\sum_{i=1}^{t} b_i \beta_{C_i} + c\beta_n\big) \cdot H \ge 0$ for every nef divisor H in $X^{[n]}$. Letting $H = D_{F_i}$, we get $b_i \ge 0$ for every i. Letting $H = (n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2$, we obtain $(n-1)\sum_{i=1}^{t} b_i + c \ge 0$. Therefore, we have

$$\sum_{i=1}^{t} b_i \beta_{C_i} + c\beta_n = \sum_{i=1}^{t} b_i (\beta_{C_i} - (n-1)\beta_n) + \left((n-1)\sum_{i=1}^{t} b_i + c\right)\beta_n \in \boldsymbol{C}_2.$$

It follows that $\boldsymbol{C}_2$ coincides with the cone $\mathrm{NE}(X^{[n]})$. □

Corollary 2.7. *Under the same assumptions as in Theorem 2.6, if γ is an irreducible curve in the Hilbert scheme $X^{[n]}$, then γ is homologous to*

$$\sum_{i=1}^{t} b_i (\beta_{C_i} - (n-1)\beta_n) + c\beta_n$$

for some nonnegative integers $b_1, \dots, b_t, c$ not all equal to zero.

Proof. By Theorem 2.6 (ii), $b_1, \dots, b_t, c$ are nonnegative and not all equal to 0. Intersecting the curve γ with the divisors $D_{F_1}, \dots, D_{F_t}$ and $(n-1)\sum_{i=1}^{t} D_{F_i} - B_n/2$, respectively, we see that $b_1, \dots, b_t, c$ must be integers. □

3. Application to Hirzebruch surfaces

In this section, we apply Theorem 2.6 to the Hirzebruch surfaces and recover a result in [Bertram and Coskun 2013]. Then we study the curves in the Hilbert schemes of points on the Hirzebruch surfaces, which have the minimal degree. We compute their normal bundles, and prove that their moduli spaces are unobstructed.

Let X denote the Hirzebruch surface $\mathbb{F}_e$ with $e \geq 0$. Let f be a fiber of the ruling $\pi : X \to \mathbb{P}^1$, and $\sigma \subset X$ be a section of π such that $\sigma^2 = -e$. Then

$$\mathrm{Pic}(X) = \mathbb{Z} \cdot \sigma \oplus \mathbb{Z} \cdot f.$$

It is well-known that $a\sigma + bf$ is nef if and only if $a \geq 0$ and $b \geq ae$. The following lemma was proved in [Beltrametti and Sommese 1993].

Lemma 3.1. *$\mathcal{O}_X(a\sigma + bf)$ is n-very ample if and only if $a \geq n$ and $b \geq n + ae$.* □

Proposition 3.2. *Let $n \geq 2$, and let X be the Hirzebruch surface $\mathbb{F}_e$. Then*

(i) *the nef cone of the Hilbert scheme $X^{[n]}$ is spanned by*

$$D_f, \quad D_\sigma + eD_f, \quad (n-1)D_\sigma + (n-1)(1+e)D_f - B_n/2; \tag{3-1}$$

(ii) *the cone* $\mathrm{NE}(X^{[n]})$ *is spanned by the classes*

$$\beta_\sigma - (n-1)\beta_n, \quad \beta_f - (n-1)\beta_n, \quad \beta_n. \tag{3-2}$$

Proof. The nef cone of X is the span of $F_1 = f$ and $F_2 = \sigma + ef$, and the cone $\mathrm{NE}(X)$ is the span of $C_1 = \sigma$ and $C_2 = f$. Note that $F_i \cdot C_j = \delta_{i,j}$ for all i and j. In addition, by Lemma 3.1, the line bundle $\mathcal{O}_X((n-1)F_1 + (n-1)F_2)$ is $(n-1)$-very ample. Hence our proposition follows from Theorem 2.6. □

Proposition 3.2 has been proved in [Bertram and Coskun 2013]. We now study the curves in $X^{[n]}$ whose homology classes are contained in the list (3-2). Let

$$L_n = n\sigma + n(1+e)f.$$

By Lemma 3.1, the line bundle $\mathcal{O}_X(L_n)$ is n-very ample. By Lemma 2.2, the divisor

$$D_{L_n} - B_n/2 = nD_\sigma + n(1+e)D_f - B_n/2$$

in $X^{[n]}$ is very ample. Our next lemma characterizes the homology classes in (3-2).

Lemma 3.3. *Let γ be a curve in $X^{[n]}$ with $\gamma \cdot (nD_\sigma + n(1+e)D_f - B_n/2) = 1$. Then γ is a smooth rational curve. Moreover, $\gamma \sim \beta_n$, $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$.*

Proof. Since $nD_\sigma + n(1+e)D_f - B_n/2$ is very ample, γ is a smooth rational curve. By Corollary 2.7, $\gamma \sim a(\beta_\sigma - (n-1)\beta_n) + b(\beta_f - (n-1)\beta_n) + c\beta_n$ for some nonnegative integers a, b, c. Since $\gamma \cdot (nD_\sigma + n(1+e)D_f - B_n/2) = 1$, we obtain

$$a + b + c = 1.$$

Therefore, we have $\gamma \sim \beta_n$, $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. □

The curves in $X^{[n]}$ homologous to β_n have been classified in [Li et al. 2003]. In the rest of this section, we study the curves $\gamma \subset X^{[n]}$ which are homologous to $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. By Lemma 3.1, the line bundle $\mathcal{O}_X(L_{n-1})$ is $(n-1)$-very ample. So by (2-9), we have the morphism

$$\varphi := \varphi_n(\mathcal{O}_X(L_{n-1})) : X^{[n]} \to \mathbb{G}\big(H^0(X, \mathcal{O}_X(L_{n-1})), n\big).$$

Lemma 3.4. *Let γ be an irreducible curve in $X^{[n]}$ satisfying*

$$\gamma \cdot ((n-1)D_\sigma + (n-1)(1+e)D_f - B_n/2) = 0.$$

Then $\gamma \sim \beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. Moreover, γ is contracted by φ.

Proof. The first part of the lemma is proved by an argument similar to the proof of Lemma 3.3. For the second part, we notice from (2-10) that

$$\begin{aligned}(\mathfrak{P} \circ \varphi)^* \mathcal{H} &= \mathcal{O}_{X^{[n]}}(D_{L_{n-1}} - B_n/2) \\ &= \mathcal{O}_{X^{[n]}}((n-1)D_\sigma + (n-1)(1+e)D_f - B_n/2).\end{aligned}$$

Therefore, the curve γ is contracted by the morphism $\mathfrak{P} \circ \varphi$. Since $\mathfrak{P}$ is an embedding, the curve γ is contracted by the morphism φ. □

In the following, we fix a curve $\gamma \subset X^{[n]}$ homologous to $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. Let $X^{(n)}$ be the n-th symmetric product of X and $\upsilon_n : X^n \to X^{(n)}$ the quotient map. Let $\rho_n : X^{[n]} \to X^{(n)}$ be the Hilbert–Chow morphism sending an element $\xi \in X^{[n]}$ to its support (with multiplicities) in $X^{(n)}$. Let p_1 be the projection from X^n to the first factor.

Definition 3.5. Define C_γ to be the union of all the curves in $p_1(\upsilon_n^{-1}(\rho_n(\gamma)))$.

Lemma 3.6. *Let $\gamma \sim \beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. Then $C_\gamma \sim \sigma$ or f.*

Proof. First of all, we claim that $C_\gamma \neq \varnothing$. Indeed, if $C_\gamma = \varnothing$, then $p_1(\upsilon_n^{-1}(\rho_n(\gamma)))$ is a finite set of points in X. Since the divisor $\sigma + (1+e)f$ is very ample, we can choose a smooth curve $F \in |\sigma + (1+e)f|$ such that $F \cap p_1(\upsilon_n^{-1}(\rho_n(\gamma))) = \varnothing$. Since the elements of γ are supported in $p_1(\upsilon_n^{-1}(\rho_n(\gamma)))$, we must have $\gamma \cap D_F = \varnothing$. It follows that $\gamma \cdot D_F = 0$. However, this contradicts $\gamma \cdot D_F = 1$.

Next, assume that $C_\gamma \cdot (\sigma + (1+e)f) \geq 2$. Take a point $\xi \in \gamma$ and a smooth point $x \in C_\gamma$ such that $x \notin \mathrm{Supp}(\xi)$. Since $x \in C_\gamma \subset p_1(\upsilon_n^{-1}(\rho_n(\gamma)))$, there exists $\xi_x \in \gamma$ such that $\rho_n(\xi_x) = n_x x + \eta_x$, where $n_x \geq 1$, $\eta_x \in X^{(n-n_x)}$ and $x \notin \mathrm{Supp}(\eta_x)$. Choose a smooth curve $F \in |\sigma + (1+e)f|$ missing $\mathrm{Supp}(\eta_x) \cup \mathrm{Supp}(\xi)$, passing through x, and intersecting C_γ transversally. Then $F \cap C_\gamma$ is a finite set. Since $C_\gamma \cdot F \geq 2$, $F \cap C_\gamma$ contains one more point $y \neq x$. Hence there exists $\xi_y \in \gamma$ with $y \in \mathrm{Supp}(\xi_y)$. Thus $\xi_x, \xi_y \in \gamma \cap D_F$. Since $y \neq x$, $y \in F$ and F misses $\mathrm{Supp}(\eta_x)$, we get

$$y \notin \{x\} \cup \mathrm{Supp}(\eta_x) = \mathrm{Supp}(\xi_x).$$

So $\xi_x \neq \xi_y$. Since $\mathrm{Supp}(\xi) \cap F = \varnothing$, we have $\xi \notin D_F$. Since $\xi \in \gamma$ and γ is a smooth rational curve, γ is not contained in D_F. Therefore, $\gamma \cap D_F$ is a finite set of points. Since $\xi_x, \xi_y \in \gamma \cap D_F$ and $\xi_x \neq \xi_y$, we obtain $\gamma \cdot D_F \geq 2$, which contradicts $\gamma \cdot D_F = 1$.

It follows that $C_\gamma \cdot (\sigma + (1+e)f) = 1$. Since the cone $\mathrm{NE}(\mathbb{F}_e)$ is spanned by σ and f, we conclude that $C_\gamma \sim \sigma$ or f. □

By Lemma 3.6, $C_\gamma \sim \sigma$ or f. So C_γ is a smooth rational curve, and

$$\mathcal{O}_X(L_{n-1})|_{C_\gamma} \cong \mathcal{O}_{C_\gamma}(n-1).$$

Let $V_{C_\gamma} \subset H^0(X, \mathcal{O}_X(L_{n-1}))$ be the image of the injection

$$H^0(X, \mathcal{O}_X(L_{n-1} - C_\gamma)) \to H^0(X, \mathcal{O}_X(L_{n-1})),$$

which is induced by the exact sequence

$$0 \to \mathcal{O}_X(L_{n-1} - C_\gamma) \to \mathcal{O}_X(L_{n-1}) \to \mathcal{O}_{C_\gamma}(n-1) \to 0. \tag{3-3}$$

Similarly, for $\xi \in \gamma$, let $V_\xi \subset H^0(X, \mathcal{O}_X(L_{n-1}))$ be the image of the injection

$$H^0(X, \mathcal{O}_X(L_{n-1}) \otimes I_\xi) \to H^0(X, \mathcal{O}_X(L_{n-1})).$$

Since $\mathcal{O}_X(L_{n-1})$ is $(n-1)$-very ample, we obtain

$$\dim V_\xi = h^0(X, \mathcal{O}_X(L_{n-1})) - h^0(\mathcal{O}_\xi) = h^0(X, \mathcal{O}_X(L_{n-1})) - n. \tag{3-4}$$

Since the curve γ is contracted to a point by the morphism φ, the subspaces V_ξ of $H^0(X, \mathcal{O}_X(L_{n-1}))$ are independent of $\xi \in \gamma$. Set $V_\gamma = V_\xi$ where $\xi \in \gamma$.

Lemma 3.7. *If $n \geq e+1$, then $V_{C_\gamma} = V_\gamma$.*

Proof. Since $K_X = -2\sigma - (2+e)f$, the divisor $L_{n-1} - C_\gamma - K_X$ is ample if $C_\gamma = \sigma$. Similarly, since $n \geq e+1$, $L_{n-1} - C_\gamma - K_X$ is ample if $C_\gamma = f$. By the Kodaira vanishing theorem, $H^1(X, \mathcal{O}_X(L_{n-1} - C_\gamma)) = 0$. So we see from (3-3) that

$$\dim V_{C_\gamma} = h^0(X, \mathcal{O}_X(L_{n-1})) - h^0(C_\gamma, \mathcal{O}_{C_\gamma}(n-1)) = h^0(X, \mathcal{O}_X(L_{n-1})) - n.$$

In view of (3-4), we conclude that

$$\dim V_{C_\gamma} = \dim V_\xi = \dim V_\gamma.$$

Thus, to prove our lemma, it remains to prove that $V_\gamma \subset V_{C_\gamma}$. Indeed, let $f \in V_\gamma$ be a section. Let $x \in C_\gamma$. Since $C_\gamma \subset p_1(\upsilon_n^{-1}(\rho_n(\gamma)))$, there exists $\xi \in \gamma$ such that $x \in \mathrm{Supp}(\xi)$. Since $V_\gamma = V_\xi$, f vanishes at every point in $\mathrm{Supp}(\xi)$. In particular, f vanishes at x. Hence, f vanishes along the smooth curve C_γ. Therefore, $f \in V_{C_\gamma}$. It follows that $V_\gamma \subset V_{C_\gamma}$. □

Proposition 3.8. *Let $n \geq \max(2, e+1)$. Then a curve $\gamma \subset X^{[n]}$ is homologous to $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$ if and only if there is a curve $C \subset X$ such that $C \sim \sigma$ or f, and that γ is a line in* $\mathrm{Hilb}^n(C) \subset X^{[n]}$. *Moreover, the curve C is uniquely determined by the curve γ.*

Proof. The "if" part of the proposition follows from Lemma 2.5 (i). To prove the "only if" part, let $\gamma \sim \beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. By Lemma 3.6, $C := C_\gamma \sim \sigma$ or f. Fix a section $f_0 \in H^0(X, \mathcal{O}_X(C))$ whose zero locus is C. Let $\xi \in \gamma$. Since $\sigma + ef$ is basepoint-free, so is the divisor $L_{n-1} - C$. Thus there exists $f_1 \in H^0(X, \mathcal{O}_X(L_{n-1} - C))$ such that f_1 does not vanish at any point in $\mathrm{Supp}(\xi)$. Now, $f_0 \otimes f_1 \in V_C$. By Lemma 3.7, $f_0 \otimes f_1 \in V_\gamma = V_\xi$. Since f_1 does not vanish at any point in $\mathrm{Supp}(\xi)$, f_0 vanishes at ξ. Hence, ξ is a closed subscheme of C. It follows that $\gamma \subset \mathrm{Hilb}^n(C)$.

To show that γ is a line in $\mathrm{Hilb}^n(C) \subset X^{[n]}$, let $F = \sigma + (e+1)f$. By Lemma 2.5 (ii), $\mathcal{O}_{X^{[n]}}(D_F)|_{\mathrm{Hilb}^n(C)} = \mathcal{O}_{\mathrm{Hilb}^n(C)}(1)$. So viewing γ as a curve in $\mathrm{Hilb}^n(C)$, we obtain

$$\gamma \cdot c_1(\mathcal{O}_{\mathrm{Hilb}^n(C)}(1)) = \gamma \cdot D_F = 1.$$

Therefore, γ is a line in $\mathrm{Hilb}^n(C) \subset X^{[n]}$.

Finally, the uniqueness of C follows from the observation that if $\xi \in X^{[n]}$ and $n \geq 2$, then ξ is contained in at most one curve $C \subset X$ with $C \sim \sigma$ or f. $\square$

Next, we determine the normal bundle of a curve γ in $X^{[n]}$ homologous to $\beta_f - (n-1)\beta_n$. By Proposition 3.8, there exists a unique fiber f_γ in X such that γ is a line in the n-th symmetric product $f_\gamma^{(n)} = \mathrm{Hilb}^n(f_\gamma) \subset X^{[n]}$. In particular,

$$N_{\gamma \subset f_\gamma^{(n)}} \cong \mathcal{O}_\gamma(1)^{\oplus(n-1)}.$$

So we have the following exact sequence of normal bundles:

$$0 \to \mathcal{O}_\gamma(1)^{\oplus(n-1)} \to N_{\gamma \subset X^{[n]}} \to N_{f_\gamma^{(n)} \subset X^{[n]}}|_\gamma \to 0. \tag{3-5}$$

Lemma 3.9. *Let $n \geq \max(2, e+1)$. Let $\gamma \subset X^{[n]}$ be a curve homologous to $\beta_f - (n-1)\beta_n$, and f_γ be the unique fiber in X such that γ is a line in $f_\gamma^{(n)}$. Then*

(i) $N_{f_\gamma^{(n)} \subset X^{[n]}} \cong \mathcal{O}_{f_\gamma^{(n)}} \oplus \mathcal{O}_{f_\gamma^{(n)}}(-1)^{\oplus(n-1)}$;

(ii) $N_{\gamma \subset X^{[n]}} \cong \mathcal{O}_\gamma(1)^{\oplus(n-1)} \oplus \mathcal{O}_\gamma \oplus \mathcal{O}_\gamma(-1)^{\oplus(n-1)}$.

Proof. (i) First of all, let $C \subset X$ be a smooth irreducible curve. Let π_n and q_n be the projections of $X^{[n]} \times X$ to $X^{[n]}$ and X respectively. Recall the universal codimension-2 subscheme $\mathcal{Z}_n \subset X^{[n]} \times X$ from (2-11). By the results in [Altman et al. 1977], we have the isomorphism

$$N_{C^{(n)} \subset X^{[n]}} \cong \pi_{n*}(q_n^* \mathcal{O}_X(C)|_{\mathcal{Z}_n})|_{C^{(n)}}.$$

Let $\widetilde{\mathcal{Z}}_n$ be the universal subscheme in $C^{(n)} \times C$. Then we obtain

$$\pi_{n*}(q_n^*\mathcal{O}_X(C)|_{\mathcal{Z}_n})|_{C^{(n)}} \cong \tilde{\pi}_{n*}(\tilde{q}_n^*(\mathcal{O}_X(C)|_C)|_{\widetilde{\mathcal{Z}}_n}),$$

where $\tilde{\pi}_n$ and $\tilde{q}_n$ are the projections from $C^{(n)} \times C$ to $C^{(n)}$ and C, respectively. So

$$N_{C^{(n)} \subset X^{[n]}} \cong \tilde{\pi}_{n*}(\tilde{q}_n^*(\mathcal{O}_X(C)|_C)|_{\widetilde{\mathcal{Z}}_n}). \tag{3-6}$$

Replacing C in (3-6) by f_γ, we get $N_{f_\gamma^{(n)} \subset X^{[n]}} \cong \tilde{\pi}_{n*}\mathcal{O}_{\widetilde{\mathcal{Z}}_n}$. It is known that $\widetilde{\mathcal{Z}}_n \subset f_\gamma^{(n)} \times f_\gamma \cong \mathbb{P}^n \times \mathbb{P}^1$ is defined by the equation

$$a_0 U^n + a_1 U^{n-1} V + \cdots + a_n V^n = 0,$$

where $a_0, a_1, \ldots, a_n$ and U, V are the homogeneous coordinates on $\mathbb{P}^n$ and $\mathbb{P}^1$, respectively. So the line bundle $\mathcal{O}_{f_\gamma^{(n)} \times f_\gamma}(\widetilde{\mathcal{Z}}_n) \cong \mathcal{O}_{\mathbb{P}^n \times \mathbb{P}^1}(\widetilde{\mathcal{Z}}_n)$ is of type $(1, n)$ in

$$\operatorname{Pic}(f_\gamma^{(n)} \times f_\gamma) \cong \operatorname{Pic}(\mathbb{P}^n \times \mathbb{P}^1) \cong \mathbb{Z} \oplus \mathbb{Z}.$$

Applying $\tilde{\pi}_{n*}$ to the exact sequence

$$0 \to \mathcal{O}_{f_\gamma^{(n)} \times f_\gamma}(-\widetilde{\mathcal{Z}}_n) \to \mathcal{O}_{f_\gamma^{(n)} \times f_\gamma} \to \mathcal{O}_{\widetilde{\mathcal{Z}}_n} \to 0,$$

we obtain the exact sequence

$$0 \to \mathcal{O}_{f_\gamma^{(n)}} \to \tilde{\pi}_{n*}\mathcal{O}_{\widetilde{\mathcal{Z}}_n} \to \mathcal{O}_{f_\gamma^{(n)}}(-1)^{\oplus(n-1)} \to 0.$$

This exact sequence splits. Thus, $\tilde{\pi}_{n*}\mathcal{O}_{\widetilde{\mathcal{Z}}_n} \cong \mathcal{O}_{f_\gamma^{(n)}} \oplus \mathcal{O}_{f_\gamma^{(n)}}(-1)^{\oplus(n-1)}$. Hence

$$N_{f_\gamma^{(n)} \subset X^{[n]}} \cong \tilde{\pi}_{n*}\mathcal{O}_{\widetilde{\mathcal{Z}}_n} \cong \mathcal{O}_{f_\gamma^{(n)}} \oplus \mathcal{O}_{f_\gamma^{(n)}}(-1)^{\oplus(n-1)}.$$

(ii) By (i) and (3-5), we obtain the exact sequence

$$0 \to \mathcal{O}_\gamma(1)^{\oplus(n-1)} \to N_{\gamma \subset X^{[n]}} \to \mathcal{O}_\gamma \oplus \mathcal{O}_\gamma(-1)^{\oplus(n-1)} \to 0.$$

Since this exact sequence splits, the proof of (ii) is complete. □

Now we determine the normal bundle of a curve γ in $X^{[n]}$ homologous to $\beta_\sigma - (n-1)\beta_n$. Recall that the Hirzebruch surfaces $\mathbb{F}_e$ are deformation equivalent to either $\mathbb{F}_0$ or $\mathbb{F}_1$. If $X = \mathbb{F}_0 = \mathbb{P}^1 \times \mathbb{P}^1$, then σ is a fiber of one of the two rulings on X, so the normal bundle of a curve γ in $X^{[n]}$ homologous to $\beta_\sigma - (n-1)\beta_n$ has been computed by Lemma 3.9 (ii). In the following, we concentrate on $X = \mathbb{F}_1$, which is the blowup of the projective plane at a point.

Lemma 3.10. *Let $n \geq 2$ and $X = \mathbb{F}_1$. Let $\gamma \subset X^{[n]}$ be a curve homologous to $\beta_\sigma - (n-1)\beta_n$. Then, $N_{\sigma^{(n)} \subset X^{[n]}} \cong \mathcal{O}_{\sigma^{(n)}}(-1)^{\oplus n}$ and*

$$N_{\gamma \subset X^{[n]}} \cong \mathcal{O}_\gamma(1)^{\oplus(n-1)} \oplus \mathcal{O}_\gamma(-1)^{\oplus n}. \tag{3-7}$$

Proof. The proof is similar to that of Lemma 3.9, so we adopt the notation in the proof of Lemma 3.9. Since $\sigma^2 = -1$, we have $\mathcal{O}_X(\sigma)|_\sigma \cong \mathcal{O}_\sigma(-1)$. Replacing the curve C in (3-6) by σ, we conclude that

$$N_{\sigma^{(n)} \subset X^{[n]}} \cong \tilde{\pi}_{n*}(\tilde{q}_n^* \mathcal{O}_\sigma(-1)|_{\widetilde{\mathcal{Z}}_n}).$$

Applying $\tilde{\pi}_{n*}$ to the exact sequence

$$0 \to \tilde{q}_n^* \mathcal{O}_\sigma(-1) \otimes \mathcal{O}_{\sigma^{(n)} \times \sigma}(-\widetilde{\mathcal{Z}}_n) \to \tilde{q}_n^* \mathcal{O}_\sigma(-1) \to \tilde{q}_n^* \mathcal{O}_\sigma(-1)|_{\widetilde{\mathcal{Z}}_n} \to 0,$$

we obtain $\tilde{\pi}_{n*}(\tilde{q}_n^* \mathcal{O}_\sigma(-1)|_{\widetilde{\mathcal{Z}}_n}) \cong \mathcal{O}_{\sigma^{(n)}}(-1)^{\oplus n}$. Therefore, we get

$$N_{\sigma^{(n)} \subset X^{[n]}} \cong \mathcal{O}_{\sigma^{(n)}}(-1)^{\oplus n}. \tag{3-8}$$

By Proposition 3.8, γ is a line in $\sigma^{(n)} \cong \mathbb{P}^n$. Using the exact sequence

$$0 \to \mathcal{O}_\gamma(1)^{\oplus(n-1)} \to N_{\gamma \subset X^{[n]}} \to N_{\sigma^{(n)} \subset X^{[n]}}|_\gamma \to 0$$

and (3-8), we see that $N_{\gamma \subset X^{[n]}} \cong \mathcal{O}_\gamma(1)^{\oplus(n-1)} \oplus \mathcal{O}_\gamma(-1)^{\oplus n}$. □

Theorem 3.11. *Let X be the Hirzebruch surface $\mathbb{F}_e$ with $e \geq 0$, let f be a fiber of the ruling on X, and let σ be a section to the ruling with $\sigma^2 = -e$.*

(i) *If $n \geq \max(2, e+1)$, then the moduli space $\mathfrak{M}(\beta_f - (n-1)\beta_n)$ of all the curves in $X^{[n]}$ homologous to $\beta_f - (n-1)\beta_n$ is irreducible and unobstructed, i.e., is smooth with the expected dimension.*

(ii) *If $e = 1$ and $n \geq 2$, then the moduli space $\mathfrak{M}(\beta_\sigma - (n-1)\beta_n)$ of all the curves in $X^{[n]}$ homologous to $\beta_\sigma - (n-1)\beta_n$ is irreducible and unobstructed.*

Proof. Under the assumptions of (i) and (ii), we see from Lemma 3.9 (ii) and (3-7) that $H^1(\gamma, N_{\gamma \subset X^{[n]}}) = 0$ if γ is a curve homologous to either $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. This implies that the moduli spaces $\mathfrak{M}(\beta_f - (n-1)\beta_n)$ and $\mathfrak{M}(\beta_\sigma - (n-1)\beta_n)$ are unobstructed. By Proposition 3.8, $\mathfrak{M}(\beta_f - (n-1)\beta_n)$ is irreducible with dimension $2n-1$, and $\mathfrak{M}(\beta_\sigma - (n-1)\beta_n)$ is irreducible with dimension $2n-2$. □

By (2-10), the composition $\mathfrak{P} \circ \varphi_n(L_n) : X^{[n]} \to \mathbb{P}^N$ (for a suitable positive integer N) is the embedding associated to the very ample divisor

$$D_{L_n} - B_n/2 = nD_\sigma + n(1+e)D_f - B_n/2.$$

By Lemma 3.3, a curve $\gamma \subset X^{[n]}$ is mapped to a line in $\mathbb{P}^N$ if and only if γ is homologous to β_n, $\beta_f - (n-1)\beta_n$ or $\beta_\sigma - (n-1)\beta_n$. Therefore, regarding $X^{[n]}$ as a closed subvariety of $\mathbb{P}^N$, then the Hilbert scheme of lines in $X^{[n]}$ is the disjoint union of $\mathfrak{M}(\beta_n)$, $\mathfrak{M}(\beta_f - (n-1)\beta_n)$ and $\mathfrak{M}(\beta_\sigma - (n-1)\beta_n)$.

Acknowledgment

The authors thank the referees for carefully reading the manuscript and for providing valuable suggestions which have greatly improved the exposition of the paper.

References

[Altman et al. 1977] A. B. Altman, A. Iarrobino, and S. L. Kleiman, "Irreducibility of the compactified Jacobian", pp. 1–12 in *Real and complex singularities* (Oslo, 1976), edited by P. Holm, Sijthoff and Noordhoff, Alphen aan den Rijn, 1977. MR 0498546 Zbl 0415.14014

[Arcara et al. 2013] D. Arcara, A. Bertram, I. Coskun, and J. Huizenga, "The minimal model program for the Hilbert scheme of points on $\mathbb{P}^2$ and Bridgeland stability", *Adv. Math.* **235** (2013), 580–626. MR 3010070 Zbl 1267.14023

[Beltrametti and Sommese 1991] M. C. Beltrametti and A. J. Sommese, "Zero cycles and kth order embeddings of smooth projective surfaces", pp. 33–48 in *Problems in the theory of surfaces and their classification* (Cortona, 1988), edited by F. Catanese et al., Symposia Mathematica **32**, Academic Press, London, 1991. MR 1273371 Zbl 0824.00026

[Beltrametti and Sommese 1993] M. C. Beltrametti and A. J. Sommese, "On the preservation of k-very ampleness under adjunction", *Math. Z.* **212**:2 (1993), 257–283. MR 1202811 Zbl 0806.14015

[Bertram and Coskun 2013] A. Bertram and I. Coskun, "The birational geometry of the Hilbert scheme of points on surfaces", pp. 15–55 in *Birational geometry, rational curves, and arithmetic*, edited by F. Bogomolov et al., Springer, New York, NY, 2013. MR 3114922 Zbl 1273.14032

[Bolognese et al. 2015] B. Bolognese, J. Huizenga, Y. Lin, E. Riedl, B. Schmidt, M. Woolf, and X. Zhao, "Nef cones of Hilbert schemes of points on surfaces", preprint, 2015. arXiv 1509.04722

[Catanese and Gœttsche 1990] F. Catanese and L. Gœttsche, "d-very-ample line bundles and embeddings of Hilbert schemes of 0-cycles", *Manuscripta Math.* **68**:3 (1990), 337–341. MR 1065935 Zbl 0729.14006

[Cheah 1998] J. Cheah, "Cellular decompositions for nested Hilbert schemes of points", *Pacific J. Math.* **183**:1 (1998), 39–90. MR 1616606 Zbl 0904.14001

[Ellingsrud and Strømme 1998] G. Ellingsrud and S. A. Strømme, "An intersection number for the punctual Hilbert scheme of a surface", *Trans. Amer. Math. Soc.* **350**:6 (1998), 2547–2552. MR 1432198 Zbl 0893.14001

[Fogarty 1968] J. Fogarty, "Algebraic families on an algebraic surface", *Amer. J. Math.* **90** (1968), 511–521. MR 0237496 Zbl 0176.18401

[Fogarty 1973] J. Fogarty, "Algebraic families on an algebraic surface, II: The Picard scheme of the punctual Hilbert scheme", *Amer. J. Math.* **95** (1973), 660–687. MR 0335512 Zbl 0299.14020

[Grojnowski 1996] I. Grojnowski, "Instantons and affine algebras, I: The Hilbert scheme and vertex operators", *Math. Res. Lett.* **3**:2 (1996), 275–291. MR 1386846 Zbl 0879.17011

[Iarrobino 1977] A. A. Iarrobino, *Punctual Hilbert schemes*, Memoirs of the American Mathematical Society **10:**188, American Mathematical Society, Providence, RI, 1977. MR 0485867 Zbl 0355.14001

[Li et al. 2003] W.-P. Li, Z. Qin, and Q. Zhang, "Curves in the Hilbert schemes of points on surfaces", pp. 89–96 in *Vector bundles and representation theory* (Columbia, MO, 2002), edited by S. D. Cutkosky et al., Contemporary Mathematics **322**, American Mathematical Society, Providence, RI, 2003. MR 1987741 Zbl 1057.14012

[Nakajima 1997] H. Nakajima, "Heisenberg algebra and Hilbert schemes of points on projective surfaces", *Ann. of Math.* (2) **145**:2 (1997), 379–388. MR 1441880 Zbl 0915.14001

[Tikhomirov 1994] A. S. Tikhomirov, "Standard bundles on a Hilbert scheme of points on a surface", pp. 183–203 in *Algebraic geometry and its applications* (Yaroslavl, 1992), edited by A. Tikhomirov and A. Tyurin, Aspects of Mathematics **E25**, Vieweg, Braunschweig, 1994. MR 1282029 Zbl 0819.14003

Received October 18, 2015. Revised January 23, 2016.

ZHENBO QIN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MISSOURI
COLUMBIA, MO 65211
UNITED STATES
qinz@missouri.edu

YUPING TU
SCHOOL OF MATHEMATICS AND STATISTICS
WUHAN UNIVERSITY, WUHAN
430072 HUBEI
CHINA
yuping_tu@126.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.455

SMOOTH APPROXIMATION OF CONIC KÄHLER METRIC WITH LOWER RICCI CURVATURE BOUND

LIANGMING SHEN

We apply methods in a paper of Tian (*Comm. Pure Appl. Math.* 68:7 (2015), 1085–1156) to prove that a conic Kähler metric with lower Ricci curvature bound can be approximated by smooth Kähler metrics with the same lower Ricci curvature bound. Furthermore, conic singularities here can be along a simple normal crossing divisor.

1. Introduction

Recently, very important progress has been made on Kähler–Einstein metrics on Fano manifolds (see [Tian 2015; Chen et al. 2015a; 2015b; 2015c]). The main tool is an extension of Cheeger–Colding–Tian theory [Cheeger et al. 2002] to conic Kähler–Einstein metrics. This extension allows one to establish a partial C^0-estimate, which has long been known to be crucial in proving the existence of Kähler–Einstein metrics. To extend Cheeger–Colding–Tian theory from the smooth case to the conic case, Tian [2015] proved a sharp approximation theorem: any conic Kähler–Einstein metric can be approximated by smooth Kähler metrics with the same lower Ricci curvature bound in the Cheeger–Gromov sense.

The main idea for proving this sharp approximation came from [Tian 2000], which gives a method of proving the equivalence of the C^0-estimate and the properness of the Lagrangian of the corresponding complex Monge–Ampère equation. Let's describe this in more detail. First, we can define the so-called twisted Ding energy $F_\omega(\varphi)$ and the twisted Mabuchi energy $\nu_\omega(\varphi)$ as in [Li and Sun 2014]; they are Lagrangians of the corresponding complex Monge–Ampère equation for the conic Kähler–Einstein metric. Then we can prove these two energies are both proper with respect to the functional $J_\omega(\varphi)$. After that, we perturb this singular complex Monge–Ampère equation, and prove that the corresponding energies are also proper after such a perturbation. Then, we make use of the C^0-estimate in [Tian 2012] to get a new C^0-estimate for the perturbed complex Monge–Ampère equation. Finally, according to the compactness theorem, we can prove that the perturbed

MSC2010: 53C55.
Keywords: conic metrics, Ricci curvature.

Kähler metrics converge to the original conic Kähler–Einstein metric in the Cheeger–Gromov sense, and converge smoothly in the C^∞ sense outside the divisor.

Now a more general problem is to understand the structures of Kähler manifolds with lower Ricci curvature bound. A natural question is whether we can also approximate an arbitrary conic Kähler metric by smooth Kähler metrics with the same lower Ricci curvature bound. We observe that the method in [Tian 2015] applies if we can get suitable complex Monge–Ampère equations and define suitable energies for them. Moreover, instead of multiple anticanonical divisors as in the original proof, we can generalize our result to simple normal crossing divisors. A divisor D is called a simple normal crossing divisor if it can be written as

$$D = \sum_{i=1}^{m} D_i,$$

where each D_i is an irreducible divisor, and they cross only in a transversal way. Each point $p \in D$ lies in the intersection of k divisors, say $D_1, \dots, D_k$, and in the local coordinate neighborhood U we can write $D_i = \{z_i = 0\}$. Assume that our conic Kähler metric ω on the Kähler manifold M takes an angle $2\pi\beta_i$ along each D_i, where $0 < \beta_i < 1$. Then near the point $p \in D$ which lies in the intersection of all D_i, the metric ω is asymptotically equivalent to the model conic metric

$$\omega_{0,p} = \sqrt{-1}\biggl(\sum_{i=1}^{k} \frac{dz_i \wedge d\bar{z}_i}{|z_i|^{2(1-\beta_i)}} + \sum_{i=k+1}^{n} dz_i \wedge d\bar{z}_i\biggr).$$

We say a smooth Kähler metric ω_0 on M has a lower Ricci curvature bound μ if there exists a nonnegative $(1,1)$-form Ω_0 such that

$$\operatorname{Ric}\omega_0 = \mu\omega_0 + \Omega_0. \tag{1-1}$$

And we say our conic Kähler metric ω has a lower Ricci curvature bound μ if there exists a nonnegative $(1,1)$-form Ω such that

$$\operatorname{Ric}\omega = \mu\omega + \sum_{i=1}^{k} 2\pi(1-\beta_i)[D_i] + \Omega \tag{1-2}$$

(we may assume that $\Omega \neq 0$; otherwise we come back to the conic Kähler–Einstein case). This equation is in the sense of currents on M and in the classic sense outside the singular part D. Considering these equations and applying Tian's methods for conic Kähler–Einstein metrics, we can prove our main theorem.

Theorem 1.1. *For a Kähler manifold (M, D), where D is a simple normal crossing divisor, assume that we have a smooth background Kähler metric ω_0 and a conic Kähler metric $\omega = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi$ with cone angle $2\pi\beta_i$ $(0 < \beta_i < 1, 1 \leq i \leq m)$*

along each irreducible component D_i of D and that φ is a smooth real function on $M \setminus D$. If the conic Kähler metric ω has a lower Ricci curvature bound μ, or ω is a conic Kähler–Einstein metric with Ricci curvature constant μ and an extra condition that M does not have holomorphic fields, then for any $\delta > 0$, there exists a smooth Kähler metric ω_δ with the same lower Ricci curvature bound μ which converges to ω in the Gromov–Hausdorff topology on M and in the smooth topology outside D as δ tends to 0.

Note that here we can deal with all the cases for μ. However, by work of Aubin and Yau, the cases $\mu < 0$ and $\mu = 0$ are easy to handle. The difficulty will be when $\mu > 0$, i.e., the Fano case. In the following section, we set up the complex Monge–Ampère equation and perturb it, and derive a C^0-estimate for nonpositive μ. We deal with the case $\mu > 0$ in the remaining parts of this paper.

2. Basic setup and the case $\mu \leq 0$

First, comparing equations (1-1) and (1-2), we have

$$\sqrt{-1}\partial\bar{\partial}\log\frac{\omega^n}{\omega_0^n} = -\mu\varphi + \Omega_0 - \Omega - \sum_{i=1}^{m}(1-\beta_i)\big(R(\|\cdot\|_i) + \sqrt{-1}\partial\bar{\partial}\log\|S_i\|_i^2\big),$$

where $\omega = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi$ is the conic Kähler metric. As each D_i is an irreducible positive divisor, we set S_i as its defining holomorphic section, with $(\|\cdot\|_i)$ as the Hermitian product on the associated line bundle $[D_i]$, and the curvature of this bundle is defined as $R(\|\cdot\|_i) := -\sqrt{-1}\partial\bar{\partial}\log\|\cdot\|_i^2$. Then we get the equation above just from the Poincaré–Lelong equation

$$2\pi[D] = \sqrt{-1}\partial\bar{\partial}\log|S|^2 = \sqrt{-1}\partial\bar{\partial}\log\|S\|^2 + R(\|\cdot\|).$$

Noting that the left-hand sides of (1-1) and (1-2) both lie in the cohomology class $c_1(M)$, we deduce that

$$\Omega_0 - \Omega - \sum_{i=1}^{m}(1-\beta_i)R(\|\cdot\|_i) = \sqrt{-1}\partial\bar{\partial}h_0, \tag{2-1}$$

where h_0 is a smooth function on M, and we note that $\frac{1}{2\pi}R(\|\cdot\|_i)$ represents $c_1(D_i)$. Then we get our complex Monge–Ampère equation:

$$\big(\omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi\big)^n = e^{h_0 - \mu\varphi - \sum_{i=1}^{m}(1-\beta_i)\log\|S_i\|_i^2 + c}\,\omega_0^n, \tag{2-2}$$

where the constant c is chosen so that

$$\int_M \big(e^{h_0 - \sum_{i=1}^{m}(1-\beta_i)\log\|S_i\|_i^2 + c} - 1\big)\omega_0^n = 0.$$

As in [Tian 2015], we can choose such an approximation equation:

$$\big(\omega_0+\sqrt{-1}\partial\bar\partial\varphi\big)^n=e^{h_\delta-\mu\varphi}\omega_0^n, \tag{2-3}$$

where

$$h_\delta=h_0-\sum_{i=1}^m(1-\beta_i)\log(\delta+\|S_i\|_i^2)+c_\delta$$

and the constant c_δ is chosen such that

$$\int_M\big(e^{h_0-\sum_{i=1}^m(1-\beta_i)\log(\delta+\|S_i\|_i^2)+c_\delta}-1\big)\omega_0^n=0.$$

Here c_δ is uniformly bounded. If we have a solution φ_δ for (2-3), then we get a smooth Kähler metric $\omega_\delta=\omega_0+\sqrt{-1}\partial\bar\partial\varphi_\delta$ with Ricci curvature given by

$$\begin{aligned}\operatorname{Ric}\omega_\delta&=\operatorname{Ric}\omega_0+\mu\sqrt{-1}\partial\bar\partial\varphi_\delta-\sqrt{-1}\partial\bar\partial h_\delta\\&=\mu\omega_0+\Omega_0+\mu\sqrt{-1}\partial\bar\partial\varphi_\delta-\sqrt{-1}\partial\bar\partial h_0+\sum_{i=1}^m(1-\beta_i)\sqrt{-1}\partial\bar\partial\log(\delta+\|S_i\|_i^2)\\&=l\mu\omega_\delta+\Omega\\&\qquad+\sum_{i=1}^m(1-\beta_i)\bigg(R(\|\cdot\|_i)+\frac{\|S_i\|_i^2}{\delta+\|S_i\|_i^2}\sqrt{-1}\partial\bar\partial\log\|S_i\|_i^2+\frac{\delta DS_i\wedge\overline{DS_i}}{(\delta+\|S_i\|_i^2)^2}\bigg)\\&=\mu\omega_\delta+\Omega+\sum_{i=1}^m(1-\beta_i)\bigg(\frac{\delta}{\delta+\|S_i\|_i^2}R(\|\cdot\|_i)+\frac{\delta DS_i\wedge\overline{DS_i}}{(\delta+\|S_i\|_i^2)^2}\bigg).\end{aligned}$$

Note that $\|S_i\|_i^2\sqrt{-1}\partial\bar\partial\log|S_i|_i^2=\|S_i\|_i^2\cdot 2\pi[D_i]=0$. We can see that if we have a solution φ_δ for small $\delta>0$, the Ricci curvature of ω_δ is always greater than μ.

By the computation above, we have a corollary which asserts the openness of the solvable set for the continuity path below.

Lemma 2.1. *Consider the continuity path of* (2-3),

$$\big(\omega_0+\sqrt{-1}\partial\bar\partial\varphi\big)^n=e^{h_\delta-t\varphi}\omega_0^n, \tag{2-4}$$

which corresponds to the equation

$$\operatorname{Ric}\omega_t:=\operatorname{Ric}\big(\omega_0+\sqrt{-1}\partial\bar\partial\varphi\big)=t\omega_t+(\mu-t)\omega_0+\Omega-\sqrt{-1}\partial\bar\partial h_\delta, \tag{2-5}$$

and set the interval I_δ as its solvable interval. Then $0\in I_\delta$ and this interval is open.

Proof. That $0\in I_\delta$ follows from the Calabi–Yau theorem. By the computation above and [Tian 2015], it's easy to see that $\lambda_1(-\Delta_t)$ is strictly larger than t. Then the openness of I_δ follows. □

So now, to solve (2-3), we need to set up a C^0-estimate for φ_δ. We first consider the cases $\mu = 0$ and $\mu < 0$. Actually, by the Calabi–Yau theorem and Aubin's work (see [Yau 1978]), we can get C^0-estimates for these cases. The main difficulty lies in the case $\mu > 0$, which we will deal with in the following sections.

3. Twisted functionals for complex Monge–Ampère equations, bounded from below

Following [Berman 2013; Ding and Tian 1992; Jeffres et al. 2016; Tian 2000; Li and Sun 2014], we can still define corresponding functionals for our complex Monge–Ampère equation (2-2). First, we define generalized energy functionals.

Definition 3.1. We have

(1) $J_{\omega_0}(\varphi) = \dfrac{1}{V}\displaystyle\sum_{i=0}^{n-1}\frac{i+1}{n+1}\int_M \sqrt{-1}\partial\varphi\wedge\bar{\partial}\varphi\wedge\omega_0^i\wedge\omega_\varphi^{n-i-1}$,

(2) $I_{\omega_0}(\varphi) = \dfrac{1}{V}\displaystyle\int_M \varphi(\omega_0^n - \omega_\varphi^n)$,

where $V = \int_M \omega_0^n$ and $\omega_\varphi = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi$.

Note that these functionals are well defined even in the conic case. It's easy to check that

$$0 \le \frac{n+1}{n} J_{\omega_0}(\varphi) \le I_{\omega_0}(\varphi) \le (n+1) J_{\omega_0}(\varphi).$$

Next let's define two functionals which are both Lagrangians of (2-2). For simplicity here we set

$$H_0 = h_0 - \sum_{i=1}^{m}(1-\beta_i)\log\|S_i\|_i^2 + c,$$

and we can choose a family φ_t connecting 0 and φ.

Definition 3.2. (1) We define the twisted Ding functional as

$$F_{\omega_0,\mu}(\varphi) = J_{\omega_0}(\varphi) - \frac{1}{V}\int_M \varphi\omega_0^n - \frac{1}{\mu}\log\left(\frac{1}{V}\int_M e^{H_0-\mu\varphi}\omega_0^n\right). \tag{3-1}$$

(2) We define the twisted Mabuchi functional as

$$\begin{aligned}\nu_{\omega_0,\mu}(\varphi) &= -\frac{n}{V}\int_0^1\int_M \dot{\varphi}\left(\operatorname{Ric}\omega_\varphi - \mu\omega_\varphi - \sum_{i=1}^{m}2\pi(1-\beta_i)[D_i] - \Omega\right)\wedge\omega_\varphi^{n-1}\,dt\\ &= \frac{1}{V}\int_M \log\frac{\omega_\varphi^n}{\omega_0^n}\omega_\varphi^n + \frac{1}{V}\int_M H_0(\omega_0^n - \omega_\varphi^n) - \mu(I_{\omega_0}(\varphi) - J_{\omega_0}(\varphi))\\ &= \frac{1}{V}\int_M \log\frac{\omega_\varphi^n}{\omega_0^n}\omega_\varphi^n + \frac{1}{V}\int_M H_0(\omega_0^n - \omega_\varphi^n) + \mu\left(F^0_{\omega_0}(\varphi) + \frac{1}{V}\int_M \varphi\omega_\varphi^n\right),\end{aligned}$$

where

$$F^0_{\omega_0}(\varphi) = J_{\omega_0}(\varphi) - \frac{1}{V}\int_M \varphi\omega_0^n.$$

These definitions are similar to the smooth case [Tian 2012] and the conic Kähler–Einstein case [Li and Sun 2014]. We can check that they are well defined for the conic case. From those papers, we know that to get a C^0-estimate for φ_δ, we need to prove the corresponding twisted Ding functional is proper with respect to the generalized energy $J_{\omega_0}(\varphi)$. Now let's recall the definition of properness.

Definition 3.3. Suppose the twisted Ding functional $F_{\omega,\mu}(\varphi)$ (twisted Mabuchi functional $\nu_{\omega,\mu}(\varphi)$) is bounded from below, i.e., $F_{\omega,\mu}(\varphi) \geq -c_\omega$ $(\nu_{\omega,\mu}(\varphi) \geq -c_\omega)$. We say it is proper on $P_c(M,\omega)$ if there is an increasing function $f : [-c_\omega, \infty) \to \mathbb{R}$, and $\lim_{t\to\infty} f(t) = \infty$, such that for any $\varphi \in P_c(M,\omega)$,

$$F_{\omega,\mu}(\varphi) \geq f(J_\omega(\varphi)) \quad (\nu_{\omega,\mu}(\varphi) \geq f(J_\omega(\varphi))),$$

where $\varphi \in P_c(M,\omega)$ is a bounded function which is smooth on $M \setminus D$, and such that $\omega_\varphi = \omega + \sqrt{-1}\partial\bar\partial\varphi$ is a conic metric with the prescribed angles along each component of D.

There are a lot of properties for these functionals, which are parallel to those in [Tian 2012; Li 2012; Li and Sun 2014]. We just put two basic facts here; the proofs are in [Tian 2012; Li and Sun 2014].

Proposition 3.4. (1) *Given a path $\{\varphi_t\}$ in $P_c(M,\omega)$, we have*

$$\frac{d}{dt} J_\omega(\varphi_t) = -\frac{1}{V}\int_M \dot\varphi_t(\omega_\varphi^n - \omega^n),$$
$$\frac{d}{dt} F^0_\omega(\varphi_t) = -\frac{1}{V}\int_M \dot\varphi_t\omega_\varphi^n.$$

(2) *$F_{\omega,\mu}(\varphi)$, $F^0_\omega(\varphi)$ and $\nu_{\omega,\mu}(\varphi)$ satisfy the cocycle condition*:

$$F_{\omega,\mu}(\varphi) + F_{\omega_\varphi,\mu}(\psi - \varphi) = F_{\omega,\mu}(\psi),$$
$$F^0_\omega(\varphi) + F^0_{\omega_\varphi}(\psi - \varphi) = F^0_\omega(\psi),$$
$$\nu_{\omega,\mu}(\varphi) + \nu_{\omega_\varphi,\mu}(\psi - \varphi) = \nu_{\omega,\mu}(\psi).$$

In (2), the last two equations follow directly from differentiation. For $F_{\omega_\varphi,\mu}$, we need to choose a corresponding function h_φ parallel to h_0 in (2-1). Whenever ω_φ is smooth or conic along D, we can write $\operatorname{Ric}\omega_\varphi = \mu\omega_\varphi + \Omega_\varphi$ or $\operatorname{Ric}\omega_\varphi = \mu\omega_\varphi + \sum_{i=1}^k 2\pi(1-\beta_i)[D_i] + \Omega_\varphi$, where Ω_φ is not necessarily nonnegative. Then all the arguments in the smooth case will apply.

From (1) we have a useful corollary from W. Ding [1988].

Corollary 3.5. *For* $0 < t < 1$, *we have*

$$J_\omega(t\varphi) \le t^{(n+1)/n} J_\omega(\varphi).$$

Proof. Consider the path $\{t\varphi\}_{0\le t\le 1}$. Then we have

$$\frac{d}{dt} J_\omega(t\varphi) = -\frac{1}{V}\int_M \varphi(\omega_{t\varphi}^n - \omega^n) = \frac{I_\omega(t\varphi)}{t} \ge \frac{n+1}{n}\frac{J_\omega(t\varphi)}{t}.$$

Integrate this inequality, and then the corollary follows. □

Now we discuss some relations among these functionals and their behaviors under different background metrics. First we have a lemma on the generalized energy J_ω; see [Li and Sun 2014] for its proof.

Lemma 3.6. *Suppose* $\omega_2 = \omega_1 + \sqrt{-1}\partial\bar\partial\varphi$. *Then for any* $\varphi \in P_c(M, \omega_1) \cap P_c(M, \omega_2)$, *we have*

$$|J_{\omega_1}(\varphi) - J_{\omega_2}(\varphi)| \le C(\omega_1, \omega_2).$$

From this lemma and the cocycle property of $F_{\omega,\mu}(\varphi)$ and $\nu_{\omega,\mu}(\varphi)$, we observe that the properties of boundedness from below and properness are independent of the choice of metrics in the same Kähler class.

Next we want to know the relation between $F_{\omega,\mu}(\varphi)$ and $\nu_{\omega,\mu}(\varphi)$. We want to prove that these two properties of the two functionals are actually equivalent. These are similar to the proofs by Berman [2013] and Li and Sun [2014], and we use the proof in [Li 2012].

Lemma 3.7. (1) *There exists a constant* $C > 0$ *such that*

$$\nu_{\omega,\mu}(\varphi) \ge \mu F_{\omega,\mu}(\varphi) - C.$$

(2) *Suppose* ψ *solves* $\omega_\psi^n = e^{H_0 - \mu\varphi}$ *by the Calabi–Yau theorem. Then we have*

$$\mu F_{\omega,\mu}(\varphi) + \frac{1}{V}\int_M H_0\omega^n \ge \nu_{\omega,\mu}(\psi).$$

In particular, by (1) *and* (2) *we know that* $F_{\omega,\mu}$ *being bounded from below is equivalent to* $\nu_{\omega,\mu}$ *being bounded from below.*

(3) *In the case that* $\nu_{\omega,\mu}(\varphi) \ge C_1 J_\omega(\varphi) - C_2$, *where* $C_1, C_2 > 0$, *there exist constants* $c, C' > 0$ *such that*

$$F_{\omega,\mu}(\varphi) \ge c\,\nu_{\omega,\mu}(\varphi) - C'.$$

Proof. (1) We modify the expression of the twisted Mabuchi functional in the definition:

$$\begin{aligned}\nu_{\omega,\mu}(\varphi) &= \frac{1}{V}\int_M \log\frac{\omega_\varphi^n}{\omega^n}\omega_\varphi^n + \frac{1}{V}\int_M H_0(\omega^n-\omega_\varphi^n) + \mu\Big(F_\omega^0(\varphi)+\frac{1}{V}\int_M \varphi\omega_\varphi^n\Big)\\ &= \mu F_{\omega,\mu}(\varphi) + \frac{1}{V}\int_M H_0\omega^n + \frac{1}{V}\int_M \log\frac{\omega_\varphi^n}{\omega^n}\omega_\varphi^n\\ &\qquad\qquad - \frac{1}{V}\int_M (H_0-\mu\varphi)\omega_\varphi^n + \log\Big(\frac{1}{V}\int_M e^{H_0-\mu\varphi}\omega^n\Big)\\ &= \mu F_{\omega,\mu}(\varphi) + \frac{1}{V}\int_M H_0\omega^n + \log\Big(\frac{1}{V}\int_M e^{H_0-\mu\varphi-\log(\omega_\varphi^n/\omega^n)}\omega_\varphi^n\Big)\\ &\qquad\qquad - \frac{1}{V}\int_M\Big(H_0-\mu\varphi-\log\frac{\omega_\varphi^n}{\omega^n}\Big)\omega_\varphi^n.\end{aligned}$$

Then (1) follows from the concavity of the logarithm.

(2) Still making use of the definition and the cocycle property, we have

$$\begin{aligned}\nu_{\omega,\mu}(\psi) &= \frac{1}{V}\int_M \log\frac{\omega_\psi^n}{\omega^n}\omega_\psi^n + \frac{1}{V}\int_M H_0(\omega^n-\omega_\psi^n) + \mu\Big(F_\omega^0(\psi)+\frac{1}{V}\int_M \psi\omega_\psi^n\Big)\\ &= \frac{1}{V}\int_M (H_0-\mu\varphi)\omega_\psi^n + \frac{1}{V}\int_M H_0(\omega^n-\omega_\psi^n) + \mu\Big(F_\omega^0(\psi)+\frac{1}{V}\int_M \psi\omega_\psi^n\Big)\\ &= \frac{1}{V}\int_M H_0\omega^n + \mu\Big(F_\omega^0(\varphi)-F_{\omega_\psi}^0(\varphi-\psi)+\frac{1}{V}\int_M(\psi-\varphi)\omega_\psi^n\Big)\\ &= \frac{1}{V}\int_M H_0\omega^n + \mu\Big(F_{\omega,\mu}(\varphi)+\log\Big(\frac{1}{V}\int_M e^{H_0-\mu\varphi}\omega^n\Big)-J_{\omega_\psi}(\varphi-\psi)\Big).\end{aligned}$$

Then (2) follows from $e^{H_0-\mu\varphi}\omega^n=\omega_\psi^n$ and $J_{\omega_\psi}(\varphi-\psi)\ge 0$.

(3) From the assumption, we have a small $\delta>0$ such that $\nu_{\omega,\mu+\delta}(\varphi)=\nu_{\omega,\mu}(\varphi)-\delta(I-J)_\omega(\varphi)$ is bounded from below, and so is $F_{\omega,\mu+\delta}(\varphi)$ by (2). Then

$$\begin{aligned}F_{\omega,\mu}(\varphi) &= F_\omega^0(\varphi) - \frac{\mu+\delta}{\mu}\frac{1}{\nu+\delta}\log\Big(\frac{1}{V}\int_M e^{H_0-(\mu+\delta)\frac{\mu}{\mu+\delta}\varphi}\omega^n\Big)\\ &= F_\omega^0(\varphi) + \frac{\mu+\delta}{\mu}\Big(F_{\omega,\mu+\delta}\Big(\frac{\mu}{\mu+\delta}\varphi\Big)-F_\omega^0\Big(\frac{\mu}{\mu+\delta}\varphi\Big)\Big)\\ &\ge J_\omega(\varphi) - \frac{\mu+\delta}{\mu}J_\omega\Big(\frac{\mu}{\mu+\delta}\varphi\Big) - C'\\ &\ge \Big(1-\Big(\frac{\mu}{\mu+\delta}\Big)^{\frac{1}{n}}\Big)J_\omega(\varphi) - C',\end{aligned}$$

where the last inequality follows from Corollary 3.5. □

To prove the properness of the functionals in the case of the existence of the conic metric $\omega = \omega_\varphi$, we need to verify that they are bounded from above.

Theorem 3.8. *If the singular Monge–Ampère equation* (2-2) *has a solution* φ, *i.e., there exists a conic Kähler metric* $\omega_\varphi = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi$ *satisfying* (1-1), *then* φ *attains the minimum of the functional* $F_{\omega_0,\mu}$ *on the space* $P_c(M,\omega_0)$. *In particular* $F_{\omega_0,\mu}$ *is bounded from above.*

Proof. A parallel result is proved in [Li and Sun 2014], but we'd like to extend Ding and Tian's proof [Ding and Tian 1992; Tian 2000] to our conic case. Let's consider the continuity path of the complex Monge–Ampère equation

$$\big(\omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi_t\big)^n = e^{H_0 - t\varphi_t}\omega_0^n. \tag{3-2}$$

We know that when $t = \mu$ this equation is solvable. By [Brendle 2013], we know that it is also solvable when $t = 0$. When $0 < t < \mu$, by the implicit function theorem, we need to consider whether the linearized operator of (3-2), $\Delta_t + t$, is invertible. We know that in the smooth case, by Bochner's formula, as $\operatorname{Ric}\omega_t > t\omega_t$, it is invertible and we can prove the openness of the solvable set for t. However, in the conic case, [Jeffres et al. 2016] gives a parallel result. By their argument, we have Δ_t as the Friedrichs extension of the Laplacian associated to $\omega_t = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi_t$ and $\lambda_1(-\Delta_t) > t$, so the openness is true. We can set $\{\varphi_t\}$ as a continuous family of solutions of (3-2), and then we can do computations as [Tian 2000] in a weak sense.

First, taking the derivative of (3-2) with respect to t, we have

$$\Delta_t\dot{\varphi}_t = -\varphi_t - t\dot{\varphi}_t,$$

where Δ_t is in a weak sense as in [Jeffres et al. 2016]. As for all t, we have $\int_M e^{H_0 - t\varphi_t}\omega_0^n = V$, and taking the derivative with respect to t we get

$$\int_M (\varphi_t + t\dot{\varphi}_t)e^{H_0 - t\varphi_t}\omega_0^n = 0.$$

Making use of the formulas in the beginning of this section, we have

$$\begin{aligned}
&\frac{d}{dt}(I_{\omega_0}(\varphi_t) - J_{\omega_0}(\varphi_t))\\
&\quad= \frac{1}{V}\int_M \dot{\varphi}_t(\omega_0^n - \omega_t^n) - \frac{1}{V}\int_M \varphi_t\Delta_t\dot{\varphi}_t\omega_t^n - \frac{1}{V}\int_M \dot{\varphi}_t(\omega_0^n - \omega_t^n)\\
&\quad= \frac{1}{V}\int_M \varphi_t(\varphi_t + t\dot{\varphi}_t)\omega_t^n\\
&\quad= -\frac{d}{dt}\left(\frac{1}{V}\int_M \varphi_t e^{H_0 - t\varphi_t}\omega_0^n\right) + \frac{1}{V}\int_M \dot{\varphi}_t e^{H_0 - t\varphi_t}\omega_0^n\\
&\quad= -\frac{d}{dt}\left(\frac{1}{V}\int_M \varphi_t\omega_t^n\right) - \frac{1}{tV}\int_M \varphi_t\omega_t^n.
\end{aligned}$$

From this, we have

$$\frac{d}{dt}\big(t(I_{\omega_0}(\varphi_t)-J_{\omega_0}(\varphi_t))\big)-(I_{\omega_0}(\varphi_t)-J_{\omega_0}(\varphi_t))=-\frac{d}{dt}\bigg(\frac{1}{V}\int_M \varphi_t\omega_t^n\bigg), \tag{3-3}$$

and integrating this from 0 to t, we have

$$t(I_{\omega_0}(\varphi_t)-J_{\omega_0}(\varphi_t))-\int_0^t (I_{\omega_0}(\varphi_s)-J_{\omega_0}(\varphi_s))\,ds=-\frac{t}{V}\int_M \varphi_t\omega_t^n.$$

By the definition, it's just

$$-\int_0^t (I_{\omega_0}(\varphi_s)-J_{\omega_0}(\varphi_s))\,ds=t\bigg(J_{\omega_0}(\varphi_t)-\frac{1}{V}\int_M \varphi_t\omega_0^n\bigg)=tF^0_{\omega_0}(\varphi_t). \tag{3-4}$$

As we have $\int_M e^{H_0-\mu\varphi}\omega_0^n=V$, we can derive that $F_{\omega_0,\mu}(\varphi)\le 0$.

Now we choose φ such that $\omega_\varphi=\omega_0+\sqrt{-1}\partial\bar\partial\varphi$ is a smooth Kähler metric. Then we have

$$\operatorname{Ric}\omega_\varphi=\mu\omega_\varphi+\Omega_\varphi,$$

where Ω_φ is not necessarily nonnegative. Comparing it with (1-1), we have

$$\big(\omega_\varphi+\sqrt{-1}\partial\bar\partial(\varphi-\varphi)\big)^n=e^{h_\varphi-\mu(\varphi-\varphi)-\sum_{i=1}^m(1-\beta_i)\log\|S_i\|_i^2+c_\varphi}\omega_\varphi^n,$$

where we take

$$\sqrt{-1}\partial\bar\partial h_\varphi=\Omega_\varphi-\Omega-\sum_{i=1}^m(1-\beta_i)(R(\|\cdot\|_i).$$

Then all the arguments are parallel and we have $F_{\omega_\varphi,\mu}(\varphi-\varphi)\le 0$. Now let's consider the case when ω_φ is conic along D. Here we have the equation

$$\operatorname{Ric}\omega_\varphi=\mu\omega_\varphi+\sum_{i=1}^k 2\pi(1-\beta_i)[D_i]+\Omega_\varphi.$$

Comparing it with (1-1), we have

$$\big(\omega_\varphi+\sqrt{-1}\partial\bar\partial(\varphi-\varphi)\big)^n=e^{h_\varphi-\mu(\varphi-\varphi)+c_\varphi}\omega_\varphi^n,$$

where we have $\sqrt{-1}\partial\bar\partial h_\varphi=\Omega_\varphi-\Omega$. In this case, all the arguments are similar to those in the smooth case and we get the same conclusion. Now by the cocycle condition, we have

$$F_{\omega_0,\mu}(\varphi)=F_{\omega_0,\mu}(\varphi)-F_{\omega_\varphi,\mu}(\varphi-\varphi)\ge F_{\omega_0,\mu}(\varphi). \qquad \square$$

4. log α-invariant and properness of twisted energies

We want to prove the properness of the twisted Ding energy. First we introduce the log α-invariant, and then see how to use this invariant to prove the properness of the twisted Mabuchi energy in the case that μ is small. Then we make use of concavity of energies to prove the properness of energies in the general case.

Recall that the α-invariant in the smooth case was introduced by Tian [1987]. In [Berman 2013; Jeffres et al. 2016] this invariant is generalized to conic case. We introduce the so-called log α-invariant here, following [Li and Sun 2014].

Definition 4.1. Fix a smooth volume form vol. For any Kähler class $[\omega]$, we define the log α-invariant by

$$\alpha(\omega, D) = \sup\Big\{\alpha > 0 : \exists C_\alpha < \infty \text{ such that } \frac{1}{V}\int_M e^{\alpha(\sup\varphi-\varphi)}\frac{\mathrm{vol}}{\prod_{i=1}^m |S_i|^{2(1-\beta_i)}} \le C_\alpha \text{ for any } \varphi\in P_c(M,\omega)\Big\}.$$

Berman [2013] has an estimate for the positive lower bound of the log α-invariant in the conic case; i.e., there exists a positive number α_0 such that $\alpha(\omega, D) \ge \alpha_0 > 0$. Using this estimate, we can prove that the twisted Mabuchi energy is proper when μ is small enough.

Theorem 4.2. *Suppose*

$$\alpha(\omega, D) \ge \alpha_0 > \frac{n}{n+1}\mu > 0.$$

Then we have

$$\nu_{\omega_0,\mu}(\varphi) \ge \epsilon J_{\omega_0}(\varphi) - C,$$

where ϵ, C are constants depending on α_0, μ.

Proof. Following [Jeffres et al. 2016; Li and Sun 2014; Tian 2000] and making use of the logarithm property, for $\frac{n}{n+1}\mu < \alpha < \alpha_0$ we have

$$\begin{aligned}\log C_\alpha &\ge \log\Big(\frac{1}{V}\int_M e^{\alpha(\sup\varphi-\varphi)}\frac{e^{H_0}\omega_0^n}{\prod_{i=1}^m |S_i|^{2(1-\beta_i)}}\Big)\\ &\ge \log\Big(\frac{1}{V}\int_M e^{\alpha(\sup\varphi-\varphi)-\log((\prod_{i=1}^m |S_i|^{2(1-\beta_i)}\omega_\varphi^n)/\omega_0^n)+H_0}\omega_\varphi^n\Big)\\ &\ge \frac{1}{V}\int_M\Big(H_0 - \frac{\prod_{i=1}^m |S_i|^{2(1-\beta_i)}\omega_\varphi^n}{\omega_0^n}\Big)\omega_\varphi^n + \frac{\alpha}{V}\int_M(\sup\varphi-\varphi)\omega_\varphi^n\\ &\ge \frac{1}{V}\int_M\Big(H_0 - \frac{\prod_{i=1}^m |S_i|^{2(1-\beta_i)}\omega_\varphi^n}{\omega_0^n}\Big)\omega_\varphi^n + \alpha I_{\omega_0}(\varphi).\end{aligned}$$

By the definition of twisted Mabuchi energy, we have

$$\begin{aligned}\nu_{\omega_0,\mu}(\varphi) &= \frac{1}{V}\int_M \log\frac{\omega_\varphi^n}{\omega_0^n}\omega_\varphi^n + \frac{1}{V}\int_M H_0(\omega_0^n-\omega_\varphi^n) - \mu(I_{\omega_0}(\varphi)-J_{\omega_0}(\varphi))\\ &\geq \log C_\alpha + \frac{1}{V}\int_M H_0\omega_0^n + \alpha I_{\omega_0}(\varphi) - \mu(I_{\omega_0}(\varphi)-J_{\omega_0}(\varphi))\\ &\geq \left(\alpha-\frac{n}{n+1}\mu\right)I_{\omega_0}(\varphi)-C\\ &\geq \left(\frac{n+1}{n}\alpha-\mu\right)J_{\omega_0}(\varphi)-C. \qquad \square\end{aligned}$$

Given the equivalence of the properness of the twisted Ding energy and the Mabuchi energy, we have an easy corollary.

Corollary 4.3. *When* $\alpha(\omega, D) \geq \alpha_0 > \dfrac{n}{n+1}\mu > 0$, *we have*

$$F_{\omega_0,\mu}(\varphi) \geq \epsilon J_{\omega_0}(\varphi) - C,$$

where ϵ, C *are constants depending on* α_0, μ.

Until now we only had the properness when μ is small enough. For the general case, we need to apply the continuity method and the concavity property of the energy which is shown below to increase μ. Here is a lemma which allows us to increase μ; see also [Li and Sun 2014].

Lemma 4.4. *Suppose* $0<\mu_0<\mu_1$, *and write* $\mu=(1-t)\mu_0+t\mu_1$, *where* $0\leq t\leq 1$. *We have*

$$\mu F_{\omega_0,\mu}(\varphi) \geq (1-t)\mu_0 F_{\omega_0,\mu}(\varphi) + t\mu_1 F_{\omega_0,\mu}(\varphi).$$

Proof. The inequality follows from the convexity of exponential functions. $\square$

Now we can prove our main theorem in this section; similar results also appear in [Li and Sun 2014; Tian 2015].

Theorem 4.5. *For* $t \in (0,\mu]$ *and any* $\varphi \in P_c(M,\omega_0)$ *there exist constants* ϵ, C_ϵ *such that*

$$F_{\omega_0,t}(\varphi) \geq \epsilon J_{\omega_0}(\varphi) - C_\epsilon. \tag{4-1}$$

Proof. We apply the continuity path similar to [Jeffres et al. 2016], i.e., (3-2). In our case, we may assume that $\Omega \neq 0$. Then by that paper, we have that $\lambda_1(-\Delta_t) > t$ for all $t \in (0,\mu]$, which allows us to prove the openness at $t=\mu$. So now when $\bar\mu = \mu+\delta$, where δ is very small, we have a solution $\bar\varphi$ for (3-2), where μ is replaced by $\bar\mu$. By Theorem 3.8, $F_{\omega_0,\bar\mu}(\varphi)$ is bounded from below. Since we have the corollary above, which asserts that when $t>0$ is very small $F_{\omega_0,t}(\varphi)$ is proper, by the lemma above, we know that for all $t \in (0, \mu]$ the twisted Ding energy is proper, i.e.,

$$F_{\omega_0,t}(\varphi) \geq \epsilon J_{\omega_0}(\varphi) - C_\epsilon. \tag{4-2}$$

□

5. C^0-estimate for approximating solution: the case $\mu > 0$

Recall that in Section 2 we set up the approximating complex Monge–Ampère equation (2-3), which is expected to give us a smooth approximation of the conic Kähler metric $\omega = \omega_0 + \sqrt{-1}\partial\bar{\partial}\varphi$. We also proved a C^0-estimate for φ_δ when $\mu \leq 0$. In this section, we want to make use of the properness of corresponding Lagrangians to prove the C^0-estimate when $\mu > 0$. The first step is to prove the properness of the new approximating twisted Ding energy, which can be deduced from Section 4.

Lemma 5.1. *For $t \in (0, \mu]$ we introduce the new approximating twisted Ding energy*

$$F_{\delta,t}(\varphi) = J_{\omega_0}(\varphi) - \frac{1}{V}\int_M \varphi\omega_0^n - \frac{1}{t}\log\left(\frac{1}{V}\int_M e^{h_\delta - t\varphi}\omega_0^n\right), \tag{5-1}$$

which is the Lagrangian of the approximating complex Monge–Ampère equation (2-4) *in the continuity path. Then we have*

$$F_{\delta,t}(\varphi) \geq \epsilon J_{\omega_0}(\varphi) - C(\epsilon, \delta, t).$$

Proof. At the end of Section 4, we proved that

$$F_{\omega_0,t}(\varphi) \geq \epsilon J_{\omega_0}(\varphi) - C_\epsilon.$$

Note that

$$\begin{aligned} h_\delta &= h_0 - \sum_{i=1}^m (1-\beta_i)\log(\delta + \|S_i\|_i^2) + c_\delta \\ &\leq h_0 - \sum_{i=1}^m (1-\beta_i)\log\|S_i\|_i^2 + c_\delta \\ &= H_0 - c + c_\delta. \end{aligned}$$

We have

$$F_{\delta,t}(\varphi) \geq F_{\omega_0,t}(\varphi) + \frac{c - c_\delta}{t},$$

and the lemma follows very easily. □

Now we will follow [Tian 2000] to finish the C^0-estimate for φ_δ. Similar to (3-4), we have

$$-\int_0^t (I_{\omega_0}(\varphi_{\delta,s}) - J_{\omega_0}(\varphi_{\delta,s}))\,ds = t\left(J_{\omega_0}(\varphi_{\delta,t}) - \frac{1}{V}\int_M \varphi_{\delta,t}\omega_0^n\right) = tF^0_{\omega_0}(\varphi_{\delta,t}),$$

where $\varphi_{\delta,t}$ solves (2-4). By this equation, we can estimate

$$\begin{aligned} F_{\delta,\mu}(\varphi_{\delta,t}) &= F^0_{\omega_0}(\varphi_{\delta,t}) - \log\left(\frac{1}{V}\int_M e^{h_\delta-\mu\varphi_{\delta,t}}\omega_0^n\right) \\ &\le -\log\left(\frac{1}{V}\int_M e^{h_\delta-t\varphi_{\delta,t}-(\mu-t)\varphi_{\delta,t}}\omega_0^n\right) \\ &= -\log\left(\frac{1}{V}\int_M e^{-(\mu-t)\varphi_{\delta,t}}\omega_{\delta,t}^n\right) \\ &\le \frac{\mu-t}{\mu}\frac{1}{V}\int_M \varphi_{\delta,t}\omega_{\delta,t}^n. \end{aligned}$$

To finish the estimate, we need a useful lemma.

Lemma 5.2. $\|\varphi_{\delta,t}\|_{C^0} \le C(1+J_{\omega_0}(\varphi_{\delta,t}))$.

Proof. First we note that $\operatorname{Ric}\omega_{\delta,t} > t$, and the volume is preserved. Then we have uniform Sobolev and Poincaré constants when t doesn't tend to 0. We observe that $n+\Delta_0\varphi_{\delta,t} > 0$; then we get

$$0 \le \sup\varphi_{\delta,t} \le \frac{1}{V}\int_M \varphi_{\delta,t}\omega_0^n + C$$

by Green's formula. On the other hand, we have $n-\Delta_{\delta,t}\varphi_{\delta,t} > 0$; by Moser's iteration, we have

$$-\inf\varphi_{\delta,t} \le -\frac{C}{V}\int_M \varphi_{\delta,t}\omega_{\delta,t}^n + C.$$

By the normalization condition, when $\varphi_{\delta,t}$ changes sign, we have

$$\|\varphi_{\delta,t}\|_{C^0} \le \sup\varphi_{\delta,t} - \inf\varphi_{\delta,t} \le C(1+I_{\omega_0}(\varphi_{\delta,t})) \le C(1+J_{\omega_0}(\varphi_{\delta,t})). \qquad \square$$

In the proof we have

$$0 \le -\inf\varphi_{\delta,t} \le -\frac{C}{V}\int_M \varphi_{\delta,t}\omega_{\delta,t}^n + C;$$

then we have

$$\frac{1}{V}\int_M \varphi_{\delta,t}\omega_{\delta,t}^n \le C,$$

which gives $F_{\delta,\mu}(\varphi_{\delta,t}) \le C$. Combining the two lemmas above, we have the C^0-estimate for φ_δ and get the following result.

Theorem 5.3. *For each $\delta > 0$, the approximating complex Monge–Ampère equation* (2-3) *has a unique smooth solution φ_δ, which gives us a smooth Kähler metric $\omega_\delta = \omega_0 + \sqrt{-1}\partial\bar\partial\varphi_\delta$ such that* $\operatorname{Ric}\omega_\delta \ge \mu\omega_\delta$.

6. Convergence when δ tends to 0

In Section 5 we proved a C^0-estimate for φ_δ. We also noted that in the approximating complex Monge–Ampère equation (2-3), the constant c_δ is uniformly bounded. Then the constant $C(\epsilon, \delta, t)$ in Lemma 5.1 is uniform with respect to δ. According to this observation, we conclude that our C^0-estimate for φ_δ is uniform with respect to δ, i.e., $\sup|\varphi_\delta| \leq C_0$. Based on this, we can give a C^2-estimate for φ_δ by the generalized Schwarz lemma first.

Lemma 6.1. *We have*

$$C_1\omega_0 \leq \omega_\delta \leq \frac{C_2\omega_0}{\prod_{i=1}^m(\delta+\|S_i\|^2)^{(1-\beta_i)}}. \tag{6-1}$$

Proof. First we have $\sup|\varphi_\delta| \leq C_0$ and $\operatorname{Ric}\omega_\delta \geq \mu\omega_\delta$. Take Δ as the Laplacian for ω_δ and take a normal coordinate around a point p for ω_δ, i.e., $g_{i\bar{j}}(p) = \delta_{ij}$, $dg_{i\bar{j}}(p) = 0$. We may also take $g_{0i\bar{j}}(p) = g_{0i\bar{i}}\delta_{ij}$, i.e., diagonal for ω_0. Then

$$\begin{aligned}
\Delta \operatorname{tr}_{\omega_\delta}\omega_0 &= g^{i\bar{i}}(g^{k\bar{l}}g_{0k\bar{l}})_{i\bar{i}} \\
&= g^{i\bar{i}}(g^{k\bar{k}})_{i\bar{i}}\,g_{0k\bar{k}} + g^{i\bar{i}}g^{k\bar{k}}(g_{0k\bar{k}})_{i\bar{i}} \\
&= g^{i\bar{i}}R_{i\bar{i}}{}^{k\bar{k}}(g)g_{0k\bar{k}} - g^{i\bar{i}}g^{k\bar{k}}R_{i\bar{i}k\bar{k}}(g_0) + g^{i\bar{i}}g^{k\bar{k}}g^{l\bar{l}}(g_{0k\bar{l}})_i(g_{0l\bar{k}})_{\bar{i}} \\
&= R^{k\bar{k}}g_{0k\bar{k}} - g^{i\bar{i}}g^{k\bar{k}}R_{i\bar{i}k\bar{k}}(g_0) + g_0^{i\bar{i}}g^{k\bar{k}}g^{l\bar{l}}(g_{0k\bar{l}})_i(g_{0l\bar{k}})_{\bar{i}} \\
&\geq -g^{i\bar{i}}g^{k\bar{k}}R_{i\bar{i}k\bar{k}}(g_0) + g_0^{i\bar{i}}g^{k\bar{k}}g^{l\bar{l}}(g_{0k\bar{l}})_i(g_{0l\bar{k}})_{\bar{i}},
\end{aligned}$$

and the last inequality follows from $\operatorname{Ric}\omega_\delta \geq \mu\omega_\delta$. Now we have

$$\begin{aligned}
\Delta\log\operatorname{tr}_{\omega_\delta}\omega_0 &= \frac{\Delta\operatorname{tr}_{\omega_\delta}\omega_0}{\operatorname{tr}_{\omega_\delta}\omega_0} - \frac{|\nabla\operatorname{tr}_{\omega_\delta}\omega_0|^2}{|\operatorname{tr}_{\omega_\delta}\omega_0|^2} \\
&\geq \frac{(\operatorname{tr}_{\omega_\delta}\omega_0)g_0^{i\bar{i}}g^{k\bar{k}}g^{l\bar{l}}(g_{0k\bar{l}})_i(g_{0l\bar{k}})_{\bar{i}} - g^{i\bar{i}}g^{k\bar{k}}g^{l\bar{l}}(g_{0k\bar{k}})_i(g_{0l\bar{l}})_{\bar{i}}}{|\operatorname{tr}_{\omega_\delta}\omega_0|^2} \\
&\quad - \frac{g^{i\bar{i}}g^{k\bar{k}}R_{i\bar{i}k\bar{k}}(g_0)}{\operatorname{tr}_{\omega_\delta}\omega_0} \\
&\geq -a\operatorname{tr}_{\omega_\delta}\omega_0,
\end{aligned}$$

where the bisectional curvature of ω_0 is less than a and the last inequality follows from $g_0^{i\bar{i}}\operatorname{tr}_{\omega_\delta}\omega_0 \geq g^{i\bar{i}}$. As we have $\sup|\varphi_\delta| \leq C_0$, we take $u = \log\operatorname{tr}_{\omega_\delta}\omega_0 - (a+1)\varphi_\delta$. Then we will have

$$\Delta u \geq \operatorname{tr}_{\omega_\delta}\omega_0 - n(a+1) = e^{u+n(a+1)} - n(a+1).$$

By the maximal principle $u \le C(a)$, and we then get $\mathrm{tr}_{\omega_\delta}\,\omega_0 \le C'$, which will give us that $C_1\omega_0 \le \omega_\delta$. For the other side, making use of the complex Monge–Ampère equation (2-3) and the inequality we obtained, we can easily deduce that

$$\omega_\delta \le \frac{C_2\omega_0}{\prod_{i=1}^m (\delta + \|S_i\|^2)^{(1-\beta_i)}}. \qquad \square$$

From this lemma, by the C^3-estimate in [Yau 1978] (or see [Tian 2000]) and regularity theory we can prove that for any $l > 2$ and compact set $K \Subset M \setminus D$, there exists a uniform constant $C(l, K)$ such that we have a high order estimate locally:

$$\|\varphi_\delta\| \le C(l, K). \tag{6-2}$$

As we have all the estimates we need, we can prove the main theorem below, following [Tian 2015].

Theorem 6.2. *As δ tends to* 0, *the smooth Kähler metric ω_δ converges to the conic Kähler metric ω in the Gromov–Hausdorff topology on M and in the smooth topology outside the divisor D.*

Proof. We first consider the case that D is an irreducible divisor. As we have high order estimates (6-1) and (6-2) outside the divisor D, it suffices to prove ω_δ converges to ω in the Gromov–Hausdorff topology. For all ω_δ we have $\mathrm{Ric}\,\omega_\delta \ge \mu$, $\mathrm{Vol}(M, \omega_\delta) = V$; to apply the compactness theorem of Cheeger–Gromov (e.g., see Chapter 10 in [Petersen 2006]), we only need to bound the diameter for all ω_δ. In the case that $\mu > 0$ we can get it directly by Meyer's theorem. However, as we have the estimate (6-1), it's easy to control the length of arbitrary geodesics outside the divisor. And in the neighborhood of some irreducible divisor, say D, we make use of local coordinates and set $r = |z_1|$, where $\{z_1 = 0\}$ locally defines the divisor D. Now we know that $\|S\|$ here is almost r near the divisor and we consider the length of a short geodesic γ transverse to D such that

$$L(\gamma, \omega_\delta) \approx C \int_0^{r_0} \frac{dr}{(\delta + r^2)^{\frac{1-\beta}{2}}} \le C \int_0^{r_0} \frac{dr}{r^{1-\beta}} \le \frac{C r_0^\beta}{\beta}.$$

Along the geodesics almost tangential to D we almost have $dz_1 = 0$ so in all cases the diameter with respect to ω_δ is uniformly bounded. Now by the compactness theorem, without loss of generality, (M, ω_δ) converges to a length space $(\overline{M}, \bar{d})$ in the Gromov–Hausdorff topology. To prove the theorem we need to prove that $(\overline{M}, \bar{d})$ coincides with (M, ω). As we have high order estimate (6-2) outside the divisor D, there exists an open set U in $\overline{M}$ which is equivalent to $M \setminus D$, and the equivalence $i : M \setminus D \to U$ induces an isometry between $M \setminus D$, $\omega|_{M\setminus D}$ and $(U, \bar{d})$. Now we note that $M \setminus D$ is geodesically convex with respect to ω, i.e., given any two points $p, q \in M \setminus D$, there exists a minimal geodesic $\gamma \subset M \setminus D$

joining them. Actually we only need to consider the case when p, q are in the small neighborhood of $o \in D$. In this case we know that the metric ω is almost the standard conic metric around a point $o \in D$, which behaves like

$$\omega_{o,c} = \sqrt{-1}\left(\frac{dz_1 \wedge d\bar{z}_1}{|z_1|^{2(1-\beta)}} + \sum_{i=2}^{n} dz_i \wedge d\bar{z}_i\right).$$

Now we assume that $|z_1(p)| = |z_1(q)| = \epsilon$ and $|z_i(p)|, |z_i(q)| \approx \epsilon$, where $\epsilon > 0$ is small enough and $2 \geq i \geq n$. First we choose the segment connecting p and q across the point $o \in D$. By the estimate above we know that

$$d(p, o) + d(o, q) \approx \frac{2\epsilon^\beta}{\beta}.$$

On the other hand we choose a segment γ' whose projection on the z_1 coordinate is almost a geodesic in the cone with angle β; by standard computation we know that

$$L(\gamma') \approx C\epsilon + 2\sin\frac{\pi\beta}{2}\frac{\epsilon^\beta}{\beta}.$$

As ϵ is small and $\beta < 1$, we conclude that the geodesic connecting p and q doesn't cross the point $o \in D$. In the general case we only need to choose p', q' as in the case above to replace p, q and connect p, p' and q, q' respectively. Then the rest of the argument follows.

As $M \setminus D$ is geodesically convex, by the C^2-estimate in (6-1), we can estimate as above to show that for each point $o \in D$, a radical short line connecting o and a point outside the divisor is always rectifiable and absolutely continuous with respect to local coordinates; thus we can see that M is the metric completion of $M \setminus D$. Moreover, the equivalence i extends to a Lipschitz map from (M, ω) onto $(\overline{M}, \bar{d})$ (we still denote this map by i) and the Lipschitz constant is 1. What remains to do is to prove i is an isometry between (M, ω) and $(\overline{M}, \bar{d})$. As $(\overline{M}, \bar{d})$ is a metric completion of $M \setminus D$, we only need to prove that for $p, q \in M \setminus D$,

$$d_\omega(p, q) = \bar{d}(i(p), i(q)).$$

Observe that $\overline{D} = i(D)$ is the Gromov–Hausdorff limit of D under the convergence of (M, ω_δ) to $(\overline{M}, \bar{d})$, whose Hausdorff measure is 0, by the C^2-estimate in (6-1). Now we only need to prove that for any $\bar{p}, \bar{q} \in \overline{M} \setminus \overline{D}$ there exists a minimizing geodesic $\gamma \subset \overline{M} \setminus \overline{D}$ joining $\bar{p}, \bar{q}$. If not, we will have

$$\bar{d}(\bar{p}, \bar{q}) < d_\omega(p, q),$$

where $\bar{p} = i(p)$, $\bar{q} = i(q)$. Then there exists a small $r > 0$ such that

(1) $B_r(\bar{p}, \bar{d}) \cap \overline{D} = \varnothing$, $B_r(\bar{q}, \bar{d}) \cap \overline{D} = \varnothing$, where $B_r(\cdot, \bar{d})$ is a geodesic ball in $(\overline{M}, \bar{d})$;

(2) $\bar{d}(\bar{x}, \bar{y}) < d_\omega(x, y)$, where $\bar{x} = i(x) \in B_r(\bar{p}, \bar{d})$ and $\bar{y} = i(y) \in B_r(\bar{q}, \bar{d})$.

From these two we know that any minimizing geodesic γ connecting $\bar{x}$ and $\bar{y}$ intersects $\overline{D}$. As $r > 0$ is small, and i is an isometry outside the divisor D, we have

$$B_r(\bar{p}, \bar{d}) = i(B_r(p, \omega)), \quad B_r(\bar{q}, \bar{d}) = i(B_r(q, \omega)).$$

Choose a small tubular neighborhood T of D in M whose closure is disjoint from both $B_r(p, \omega)$ and $B_r(q, \omega)$. When the radius of such a tubular neighborhood is small enough we can make Vol ∂T arbitrarily small. Now we can choose $p_\delta, q_\delta \in M$ and a neighborhood T_δ of D with respect to ω_δ such that as $\delta \to 0$, $p_\delta, q_\delta, T_\delta$ converge to $\bar{p}, \bar{q}, i(T)$ in the Gromov–Hausdorff topology. By the volume convergence theorem of Colding, $\lim_{\delta \to 0+} \mathrm{Vol}(\partial T_\delta, \omega_\delta) = \mathrm{Vol}(\partial T, \omega)$, so $\mathrm{Vol}(\partial T_\delta, \omega_\delta)$ can also be arbitrarily small as $\delta \to 0$. Also by convergence, when δ is small enough, $B_r(p_\delta, \omega_\delta)$, $B_r(q_\delta, \omega_\delta)$ and T_δ are mutually disjoint. By (2), any minimizing geodesic γ_δ connecting any $w \in B_r(p_\delta, \omega_\delta)$ and $z \in B_r(q_\delta, \omega_\delta)$ intersects T_δ. Now we need an estimate due to Gromov:

Lemma 6.3. *We have*

$$c(\mu) r^{2n} \le \mathrm{Vol}(B_r(q_\delta, \omega_\delta), \omega_\delta) \le C(L, \mu, n, r)\,\mathrm{Vol}(\partial T_\delta, \omega_\delta),$$

where $L = \bar{d}(\bar{p}, \bar{q})$.

Proof. The first inequality follows from the Ricci lower bound and Gromov's relative volume comparison theorem directly. For the second inequality, by Chapter 9 in [Petersen 2006], we set $\lambda(t, \theta)$ as the volume density function, where t is the distance from p_δ. We also set $\lambda_k(t, \theta)$ as the standard volume density function of the space form with constant curvature $k = \mu/(n-1)$. By the argument in [Petersen 2006] we know that the map $t \to \lambda(t, \theta)/\lambda_k(t, \theta)$ is nonincreasing in t. In our case, we consider the geodesics from p_δ to $z \in B_r(q_\delta, \omega_\delta)$. According to the construction, we have $r < d(p_\delta, z_T) < d(p_\delta, z)$, $L - r < d(p_\delta, z) < L + r$, where z_T is the intersection point of the geodesics from p_δ to z and ∂T_δ, and $L \approx d(p_{\delta)}, q_\delta)$. Along ∂T_δ, we have

$$\frac{\lambda(z_T)}{\lambda_k(z_T)} \ge \frac{\lambda(z)}{\lambda_k(z)}.$$

Let $S \in S^{2n-1}$, let $C(S)$ denote the part where all the geodesics from p_δ to $z \in B_r(q_\delta, \omega_\delta)$ lie in the corresponding geodesic cone, and let $t(\theta)$ be the distance from p_δ to each point of ∂T_δ. Then we have

$$\begin{aligned} \mathrm{Vol}\,\partial T_\delta \ge \int_{\partial T_\delta \cap C(S)} \lambda(t, \theta) = \int_S t^{2n-1}(\theta) \lambda(t, \theta)\, d\theta \\ \ge \int_S \lambda(L') \frac{\lambda_k(t(\theta))}{\lambda_k(L')} t^{2n-1}(\theta)\, d\theta \ge C \int_S \lambda(L') L'^{2n-1}\, d\theta, \end{aligned}$$

where $L-r < L' < L+r$. Taking the integral of this inequality yields

$$\operatorname{Vol}\partial T_\delta \geq \frac{2C}{r}\int_{L-r}^{L+r}\int_S \lambda(L')L'^{2n-1}\,d\theta\,dt \geq C(L,\mu,n,r)\operatorname{Vol}(B_r(q_\delta,\omega_\delta),\omega_\delta).$$

Then the lemma follows. □

Since we know that $\operatorname{Vol}(\partial T_\delta,\omega_\delta)$ can also be arbitrarily small as δ tends to 0, the lemma above leads to a contradiction. Then i can extend to an isometry from (M,ω) onto $(\overline{M},\bar{d})$, and the theorem follows when D is irreducible. In the case that D is a simple normal crossing divisor, we observe that near the crossing point o, the model metric can be rewritten as

$$\begin{aligned}\omega_{o,c} &= \sqrt{-1}\biggl(\sum_{i=1}^{m}\frac{dz_i\wedge d\bar{z}_i}{|z_i|^{2(1-\beta_i)}}+\sum_{i=m+1}^{n}dz_i\wedge d\bar{z}_i\biggr)\\ &= \sqrt{-1}\biggl(\sum_{i=1}^{m}\frac{dz_i^{\beta_i}\wedge d\bar{z}_i^{\beta_i}}{\beta_i^2}+\sum_{i=m+1}^{n}dz_i\wedge d\bar{z}_i\biggr).\end{aligned}$$

For $1\leq i\leq m$, if we take $w_i := z_i^{\beta_i}/\beta_i$, we can realize the original conic metric as a Euclidean metric under these new coordinates. To find the minimal geodesic we then only need to project two points in the original space to each coordinate direction; if in each direction we can find a minimal geodesic, we are done. In this case we deduce the problem to the one irreducible divisor case. Obviously, for the conic metric along a simple normal crossing divisor, the minimal geodesic will always lie in the regular part. Hence in the general case the theorem still follows. □

Acknowledgments

The author thanks his Ph.D. thesis advisor, Professor Gang Tian, for a lot of discussions and encouragement, and Yanir Rubinstein and Chi Li for many useful conversations. He also thanks CSC for partial financial support during his Ph.D. career.

References

[Berman 2013] R. J. Berman, "A thermodynamical formalism for Monge–Ampère equations, Moser–Trudinger inequalities and Kähler–Einstein metrics", *Adv. Math.* **248** (2013),1254–1297. MR 3107540 Zbl 1286.58010

[Brendle 2013] S. Brendle, "Ricci flat Kähler metrics with edge singularities", *Int. Math. Res. Not.* **2013**:24 (2013), 5727–5766. MR 3144178 Zbl 1293.32029

[Cheeger et al. 2002] J. Cheeger, T. H. Colding, and G. Tian, "On the singularities of spaces with bounded Ricci curvature", *Geom. Funct. Anal.* **12**:5 (2002), 873–914. MR 1937830 Zbl 1030.53046

[Chen et al. 2015a] X. X. Chen, S. Donaldson, and S. Sun, "Kähler–Einstein metrics on Fano manifolds, I: Approximation of metrics with cone singularities", *J. Amer. Math. Soc.* **28**:1 (2015), 183–197. MR 3264766 Zbl 1312.53096

[Chen et al. 2015b] X. X. Chen, S. Donaldson, and S. Sun, “Kähler–Einstein metrics on Fano manifolds, II: Limits with cone angle less than 2π”, *J. Amer. Math. Soc.* **28**:1 (2015), 199–234. MR 3264767 Zbl 1312.53097

[Chen et al. 2015c] X. X. Chen, S. Donaldson, and S. Sun, “Kähler–Einstein metrics on Fano manifolds, III: Limits as cone angle approaches 2π and completion of the main proof”, *J. Amer. Math. Soc.* **28**:1 (2015), 235–278. MR 3264768 Zbl 1311.53059

[Ding 1988] W. Y. Ding, “Remarks on the existence problem of positive Kähler–Einstein metrics”, *Math. Ann.* **282**:3 (1988), 463–471. MR 967024 Zbl 0661.53045

[Ding and Tian 1992] W. Y. Ding and G. Tian, “Kähler–Einstein metrics and the generalized Futaki invariant”, *Invent. Math.* **110**:2 (1992), 315–335. MR 1185586 Zbl 0779.53044

[Jeffres et al. 2016] T. Jeffres, R. Mazzeo, and Y. A. Rubinstein, “Kähler–Einstein metrics with edge singularities”, *Ann. of Math.* (2) **183**:1 (2016), 95–176. MR 3432582 Zbl 06541584

[Li 2012] C. Li, *Kähler–Einstein metrics and K-stability*, Ph.D. thesis, Princeton University, 2012, available at http://www.math.sunysb.edu/∼chili/papers/PhDthesis.pdf.

[Li and Sun 2014] C. Li and S. Sun, “Conical Kähler–Einstein metrics revisited”, *Comm. Math. Phys.* **331**:3 (2014), 927–973. MR 3248054 Zbl 1296.32008

[Petersen 2006] P. Petersen, *Riemannian geometry*, 2nd ed., Graduate Texts in Mathematics **171**, Springer, New York, NY, 2006. MR 2243772 Zbl 1220.53002

[Tian 1987] G. Tian, “On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$”, *Invent. Math.* **89**:2 (1987), 225–246. MR 894378 Zbl 0599.53046

[Tian 2000] G. Tian, *Canonical metrics in Kähler geometry*, Birkhäuser, Basel, 2000. MR 1787650 Zbl 0978.53002

[Tian 2012] G. Tian, “Existence of Einstein metrics on Fano manifolds”, pp. 119–159 in *Metric and differential geometry*, edited by X. Dai and X. Rong, Progress in Mathematics **297**, Birkhäuser, Basel, 2012. MR 3220441 Zbl 1250.53044

[Tian 2015] G. Tian, “K-stability and Kähler–Einstein metrics”, *Comm. Pure Appl. Math.* **68**:7 (2015), 1085–1156. MR 3352459 Zbl 1318.14038

[Yau 1978] S. T. Yau, “On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I”, *Comm. Pure Appl. Math.* **31**:3 (1978), 339–411. MR 480350 Zbl 0369.53059

Received December 23, 2014. Revised September 24, 2015.

LIANGMING SHEN
DEPARTMENT OF MATHEMATICS
PRINCETON UNIVERSITY
PRINCETON, NJ 08544
UNITED STATES
Current address:
MATHEMATICAL SCIENCES RESEARCH INSTITUTE
17 GAUSS WAY
BERKELEY, CA 94720
UNITED STATES

lmshen@math.ubc.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 284, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.284.475

MAPS FROM THE ENVELOPING ALGEBRA OF THE POSITIVE WITT ALGEBRA TO REGULAR ALGEBRAS

SUSAN J. SIERRA AND CHELSEA WALTON

We construct homomorphisms from the universal enveloping algebra of the positive (part of the) Witt algebra to several different Artin–Schelter regular algebras, and determine their kernels and images. As a result, we produce elementary proofs that the universal enveloping algebras of the Virasoro algebra, the Witt algebra, and the positive Witt algebra are neither left nor right noetherian.

0. Introduction

Let $\Bbbk$ be a field of characteristic 0. All vector spaces, algebras, and tensor products are over $\Bbbk$, unless stated otherwise. In this work, we construct and study homomorphisms from the universal enveloping algebra of the positive part of the Witt algebra to *Artin–Schelter* (*AS-*)*regular algebras*. The latter serve as homological analogues of commutative polynomial rings in the field of noncommutative algebraic geometry.

To begin, consider the Lie algebras below.

Definition 0.1 (V, W, W_+). We define the following Lie algebras:

(a) The *Virasoro algebra* is defined to be the Lie algebra V with basis $\{e_n\}_{n\in\mathbb{Z}}\cup\{c\}$ and Lie bracket $[e_n, c] = 0$, $[e_n, e_m] = (m-n)e_{n+m} + \frac{1}{12}c(m^3-m)\delta_{n+m,0}$.

(b) The *Witt* (or *centerless Virasoro*) *algebra* is defined to be the Lie algebra W with basis $\{e_n\}_{n\in\mathbb{Z}}$ and Lie bracket $[e_n, e_m] = (m-n)e_{n+m}$.

(c) The *positive* (*part of the*) *Witt algebra* is defined to be the Lie subalgebra W_+ of W generated by $\{e_n\}_{n\geq 1}$.

For any Lie algebra $\mathfrak{g}$, we denote its universal enveloping algebra by $U(\mathfrak{g})$.

Further, consider the following algebras.

MSC2010: 14A22, 16S30, 16S38, 17B68.

Keywords: Artin–Schelter regular, Jordan plane, non-noetherian, universal enveloping algebra, Witt algebra.

Notation 0.2 (S, R). Let S be the algebra generated by u, v, w, subject to the relations

$$uv - vu - v^2 = uw - wu - vw = vw - wv = 0.$$

Let R be the *Jordan plane* generated by u, v, subject to the relation $uv - vu - v^2 = 0$.

It is well known that R is an AS-regular algebra of global dimension 2. Moreover, we see by Lemma 1.3 that S is also AS-regular, of global dimension 3.

This work focuses on maps that we construct from the enveloping algebra $U(W_+)$ to both R and S, given as follows:

Definition 0.3 (ϕ, λ_a). Let $\phi : U(W_+) \to S$ be the algebra homomorphism induced by defining

$$\phi(e_n) = (u - (n-1)w)v^{n-1}. \tag{0.4}$$

For $a \in \Bbbk$, let $\lambda_a : U(W_+) \to R$ be the algebra homomorphism induced by defining

$$\lambda_a(e_n) = (u - (n-1)av)v^{n-1}. \tag{0.5}$$

That ϕ and λ_a are well defined is Lemma 1.5.

Our main result is that we understand fully the kernels and images of the maps above, as presented below.

Theorem 0.6. *We have the following statements about the kernels and images of the maps ϕ and λ_a.*

(a) [Propositions 2.5, 2.8] $\ker \lambda_a$ *is equal to the ideal* $(e_1e_3 - e_2^2 - e_4)$ *if* $a = 0, 1$; *or is an ideal generated by one element of degree* 5 *and two elements of degree* 6 (*listed in Proposition 2.8*) *if* $a \neq 0, 1$.

(b) [Proposition 2.1] $\lambda_a(U(W_+))$ *is equal to* $\Bbbk + uR$ *if* $a = 0$; *is equal to* $\Bbbk + Ru$ *if* $a = 1$; *or contains* $R_{\geq 4}$ *if* $a \neq 0, 1$. *For all* a, *the image of* λ_a *is noetherian.*

(c) [Theorem 5.1] $\ker \phi$ *is equal to* $(e_1e_5 - 4e_2e_4 + 3e_3^2 + 2e_6)$.

The image of ϕ will be discussed later in the introduction, after Theorem 0.10.

The result above has a surprising application. In [Sierra and Walton 2014, Theorem 0.5 and Corollary 0.6], we established that $U(W_+)$, $U(W)$, $U(V)$ are neither left nor right noetherian through relatively indirect means, using the techniques of [Sierra 2011]. In particular, we were not able to give an example of a non-finitely-generated right or left ideal in any of these enveloping algebras. However, in the course of proving Theorem 0.6, we produce an elementary and constructive proof of [Sierra and Walton 2014, Theorem 0.5 and Corollary 0.6]. Namely, we obtain:

Theorem 0.7 (Proposition 2.5, Theorem 3.3). *The ideal*

$$\ker \lambda_0 = \ker \lambda_1 = (e_1e_3 - e_2^2 - e_4)$$

is not finitely generated as either a left or a right ideal of $U(W_+)$.

We prove this theorem by noting that λ_0 factors through ϕ, and by studying $B := \phi(U(W_+))$. A key step is to compute $I := \phi(\ker \lambda_0)$, and to show that I is not finitely generated as a left or right ideal of B.

Note that the map (0.5) can be extended to W to define a map, which we denote by

$$\widehat{\lambda}_a : U(W) \to R[v^{-1}].$$

Theorem 0.8 ((3.10), Theorem 3.12). *The ideal* $\ker \widehat{\lambda}_0 = \ker \widehat{\lambda}_1$ *is not finitely generated as either a left or right ideal of* $U(W)$.

We remark that $R[v^{-1}]$ is isomorphic to the ring $\Bbbk[x, x^{-1}, \partial]$, which is a familiar localization of the Weyl algebra. To see this, set $v = x$ and $u = x^2\partial$, so $\partial x = x\partial + 1$. Then $uv - vu = x^2 = v^2$. We obtain

$$\widehat{\lambda}_1(e_n) = v^{n-1}u = x^{n+1}\partial.$$

Thus, $\widehat{\lambda}_1$ is a well-known homomorphism.

We now compare Theorem 0.7 with our earlier proof (in [Sierra and Walton 2014]) that $U(W_+)$ is not left or right noetherian. The earlier proof used a ring homomorphism ρ with a more complicated definition:

Notation 0.9 (X, f, τ, ρ). Take $\mathbb{P}^3 := \mathbb{P}^3_{\Bbbk}$ with coordinates w, x, y, z. Let $X = V(xz - y^2) \subseteq \mathbb{P}^3$ be the projective cone over a smooth conic in $\mathbb{P}^2$.

Define an automorphism τ of X by

$$\tau([w : x : y : z]) = [w - 2x + 2z : z : -y - 2z : x + 4y + 4z].$$

Denote the pullback of τ on $\Bbbk(X)$ by τ^*, so that $g^\tau := \tau^* g = g \circ \tau$ for $g \in \Bbbk(X)$. Form the ring $\Bbbk(X)[t; \tau^*]$ with multiplication $tg = g^\tau t$ for all $g \in \Bbbk(X)$. Let

$$f := \frac{w + 12x + 22y + 8z}{12x + 6y},$$

considered as a rational function in $\Bbbk(X)$. Now let $\rho : U(W_+) \to \Bbbk(X)[t; \tau^*]$ be the algebra homomorphism induced by setting $\rho(e_1) = t$ and $\rho(e_2) = ft^2$.

That ρ is well defined is [Sierra and Walton 2014, Proposition 1.5]. The method in that paper made a number of reductions to show that $\rho(U(W_+))$ is not left or right noetherian. That proof can now be streamlined via the next result.

Theorem 0.10 (Theorem 4.1). *We have that* $\ker \rho = \ker \phi = \bigcap_{a \in \Bbbk} \ker \lambda_a$.

Since we show that $\phi(U(W_+))$ is not left or right noetherian in the course of proving Theorem 0.7, we have by Theorems 0.6(c) and 0.10 that $\rho(U(W_+)) \cong \phi(U(W_+)) \cong U(W_+)/(e_1e_5 - 4e_2e_4 + 3e_3^2 + 2e_6)$ is neither left nor right noetherian.

We end by discussing an open question that was brought to our attention by Lance Small.

Question 0.11. Does $U(W_+)$ satisfy the ascending chain condition on *two-sided* ideals?

Our result here is only partial:

Proposition 0.12 (Proposition 6.6). *The ring $B := \phi(U(W_+))$ satisfies the ascending chain condition on two-sided ideals.*

Of course, this yields no direct information on the question for $U(W_+)$.

We have the following conventions throughout the paper. We take $\mathbb{N} = \mathbb{Z}_{\geq 0}$ to be the set of nonnegative integers. If r is an element of a ring A, then (r) denotes the two-sided ideal ArA generated by r. If $A = \bigoplus_{n\in\mathbb{Z}} A_n$ is a graded $\Bbbk$-algebra (or graded module), then we define the Hilbert series

$$\operatorname{hilb} A = \sum_{n\in\mathbb{Z}} \dim_{\Bbbk} A_n t^n.$$

This article is organized as follows. We present preliminary results in Section 1, including an alternative way of multiplying elements in S and in R (Lemma 1.3); this method will be employed throughout, sometimes without mention. In Section 2, we discuss the maps λ_a and prove parts (a) and (b) of Theorem 0.6. In Section 3 we use the map λ_0 to establish Theorem 0.7; we also prove Theorem 0.8.

Before proceeding to study the map ϕ, we present its connection with the map ρ, the key homomorphism of [Sierra and Walton 2014]. Namely, in Section 4, we establish Theorem 0.10. Then in Section 5, we verify part (c) of Theorem 0.6. Our last result, Proposition 0.12, is presented in Section 6. Proofs of computational claims via Maple and Macaulay2 routines and a known result in ring theory to which we could not find a reference are provided in the Appendix.

1. Preliminaries

The main focus of this paper is the universal enveloping algebra of the positive Witt algebra, W_+. To begin, we recall some basic facts about the algebra $U(W_+)$.

Lemma 1.1. *Recall Definition 0.1(c).*

(a) *We have the isomorphism*

$$U(W_+) \cong \frac{\Bbbk\langle e_1, e_2\rangle}{\begin{pmatrix} [e_1,[e_1,[e_1,e_2]]] + 6[e_2,[e_2,e_1]], \\ [e_1,[e_1,[e_1,[e_1,[e_1,e_2]]]]] + 40[e_2,[e_2,[e_2,e_1]]] \end{pmatrix}}.$$

(b) *The set $\{e_{i_1}, e_{i_2}, \dots, e_{i_k} \mid k \in \mathbb{N}$ and $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \in \mathbb{N}\}$ forms a $\Bbbk$-vector space basis of $U(W_+)$.*

Proof. Part (a) is [Sierra and Walton 2014, Lemma 1.1], and part (b) is clear from the definition of $U(W_+)$. □

Next, let us present some notation that we will use for the rest of the paper. We will work with the algebras R and S defined in Notation 0.2; note that we can view R as a subalgebra of S. In addition:

Notation 1.2 (Q). Take Q to be the subalgebra of S generated by u, v, and vw.

In our first result, we provide an easy way to multiply elements in S. Recall from [Zhang 1996] that a *Zhang twist* of a graded algebra L, by an automorphism μ of L, is the algebra L^{μ}, where $L^{\mu} = L$ as graded vector spaces and L^{μ} has multiplication $\ell * \ell' = \ell(\ell')^{\mu^i}$ for $\ell \in L_i$ and $\ell' \in L$.

Moreover, recall that an *Artin–Schelter* (*AS-*)*regular algebra* is a connected graded algebra A of finite global dimension, of finite injective dimension d with $\mathrm{Ext}^i_A({}_A\Bbbk, {}_AA) \cong \mathrm{Ext}^i_A(\Bbbk_A, A_A) \cong \delta_{i,d}\Bbbk$ (that is, A is *AS-Gorenstein*), and has finite Gelfand–Kirillov dimension. These algebras are important in noncommutative ring theory because they are noncommutative analogues of polynomial rings and share many of their good properties.

Lemma 1.3 (μ, ν). *Let $\mu \in \mathrm{Aut}(\Bbbk[x, y, z])$ be defined by*

$$\mu(x) = x - y, \quad \mu(y) = y, \quad \mu(z) = z.$$

Then S is isomorphic to the Zhang twist $\Bbbk[x, y, z]^{\mu}$. Further, μ restricts to an automorphism of $\Bbbk[x, y, yz]$, which we also denote by μ, and to an automorphism ν of $\Bbbk[x, y]$. We have that $R \cong \Bbbk[x, y]^{\nu}$ and $Q \cong \Bbbk[x, y, yz]^{\mu}$ as graded $\Bbbk$-algebras. As a consequence, S, R, and Q are AS-regular algebras.

Proof. To see that $S \cong \Bbbk[x, y, z]^{\mu}$, we emphasize that

(1.4)
- the variables u, v, w of S have noncommutative multiplication,
- the variables x, y, z of $\Bbbk[x, y, z]$ have commutative multiplication, and
- the symbol $*$ denotes the noncommutative multiplication on $\Bbbk[x, y, z]^{\mu}$ defined by $\ell * \ell' = \ell(\ell')^{\mu^i}$ for $\ell \in \Bbbk[x, y, z]_i$ and $\ell' \in \Bbbk[x, y, z]$.

Now,

$$\begin{aligned} y * x &= yx^{\mu} = y(x - y) = (x - y)y = xy - y^2 = xy^{\mu} - yy^{\mu} = x * y - y * y, \\ z * x &= zx^{\mu} = z(x - y) = (x - y)z = xz - yz = xz^{\mu} - yz^{\mu} = x * z - y * z, \\ z * y &= zy^{\mu} = zy = yz = yz^{\mu} = y * z. \end{aligned}$$

Thus, if we identify u, v, w with x, y, z, respectively, then the relations of S hold in $\Bbbk[x, y, z]^{\mu}$, and $S \cong \Bbbk[x, y, z]^{\mu}$ as graded $\Bbbk$-algebras.

That μ restricts to automorphisms of $\Bbbk[x, y]$ and $\Bbbk[x, y, yz]$ is immediate, and the other isomorphisms hold by a similar argument. Moreover, the last statement follows as commutative polynomial rings are AS-regular and this property is preserved under Zhang twisting by [Zhang 1996, Theorem 1.3(i)]. □

Now we verify that the algebra homomorphisms λ_a and ϕ from Definition 0.3 are well defined.

Lemma 1.5. *The maps ϕ and λ_a of Definition 0.3 are well-defined homomorphisms of graded $\Bbbk$-algebras.*

Proof. We check that ϕ respects the Witt relations given in Definition 0.1(b), by using Lemma 1.3 and (1.4):

$$\begin{aligned}
&\phi(e_n e_m - e_m e_n)\\
&\quad = (u-(n-1)w)v^{n-1}(u-(m-1)w)v^{m-1} - (u-(m-1)w)v^{m-1}(u-(n-1)w)v^{n-1}\\
&\quad = (x-(n-1)z)(x-(m-1)z)^{\mu^n} y^{n+m-2} - (x-(m-1)z)(x-(n-1)z)^{\mu^m} y^{n+m-2}\\
&\quad = \big((x-(n-1)z)(x-ny-(m-1)z) - (x-(m-1)z)(x-my-(n-1)z)\big)y^{n+m-2}\\
&\quad = (m-n)xy^{n+m-1} + (n(n-1)-m(m-1))y^{n+m-1}z\\
&\quad = (m-n)(x-(n+m-1)z)y^{n+m-1}\\
&\quad = (m-n)(u-(n+m-1)w)v^{n+m-1}\\
&\quad = (m-n)\phi(e_{n+m}).
\end{aligned}$$

So, the claim holds for ϕ.

Similarly, we verify that λ_a respects the Witt relations:

$$\begin{aligned}
\lambda_a(e_n e_m - e_m e_n) &= (u-(n-1)av)v^{n-1}(u-(m-1)av)v^{m-1}\\
&\qquad\qquad - (u-(m-1)av)v^{m-1}(u-(n-1)av)v^{n-1}\\
&= \big((x-(n-1)ay)(x-ny-(m-1)ay)\\
&\qquad\qquad - (x-(m-1)ay)(x-my-(n-1)ay)\big)y^{n+m-2}\\
&= (m-n)(x-a(n+m-1)y)y^{n+m-1}\\
&= (m-n)(u-a(n+m-1)v)v^{n+m-1}\\
&= (m-n)\lambda_a(e_{n+m}).
\end{aligned}$$

Thus, the claim holds for λ_a. □

Next, we define the key algebras $A(a)$ and B that we will use throughout.

Notation 1.6 ($A(a)$, B). Take $a \in \Bbbk$ and let $A(a)$ denote the subalgebra $\lambda_a(U(W_+))$ of R. Further, let B denote the subalgebra $\phi(U(W_+))$ of S.

We point out a useful observation.

Lemma 1.7. *We have that $B \subseteq Q$.*

Proof. We get that $\phi(e_1) = u$ and $\phi(e_2) = (u-w)v = uv - vw$ are in Q. By Lemma 1.1(a), B is generated by these elements, so we are done. □

2. The kernel and image of the maps λ_a

The goal of this section is to analyze the maps λ_a from Definition 0.3, which are well defined by Lemma 1.5. In particular, we verify Theorem 0.6(a,b).

To proceed, recall Notations 0.2 and 1.6. We first compute the factor rings $A(a)$, proving Theorem 0.6(b).

Proposition 2.1. *We have that* $A(0) = \Bbbk + uR$ *(a right idealizer in* R*), that* $A(1) = \Bbbk + Ru$ *(a left idealizer in* R*), and that* $A(a)_{\geq 4} = R_{\geq 4}$ *if* $a \neq 0, 1$*. For all* a*, the ring* $A(a)$ *is noetherian.*

Proof. Recall from Lemma 1.1(a) that $U(W_+)$ is generated by e_1 and e_2. We have that $\lambda_0(e_1) = u$ and $\lambda_0(e_2) = uv$. These elements generate $\Bbbk + uR$. Moreover, $\lambda_1(e_1) = u$ and $\lambda_1(e_2) = (u-v)v = vu$, and these elements generate $\Bbbk + Ru$. That the rings $A(0)$ and $A(1)$ are noetherian follows from [Stafford and Zhang 1994, Lemma 2.2(iii) and Theorem 2.3(i.a)].

When $a \neq 0, 1$, we must show that $R_{\geq 4} \subseteq A(a)$. Since $uR_n + R_n u = R_{n+1}$ for $n \geq 1$ and since $\dim_{\Bbbk} R_4 = 5$, the proof boils down to showing that the set of elements

$$\lambda_a(e_1^4), \quad \lambda_a(e_1^2 e_2), \quad \lambda_a(e_1 e_2 e_1), \quad \lambda_a(e_2 e_1^2), \quad \lambda_a(e_2^2)$$

is $\Bbbk$-linearly independent for $a \neq 0, 1$. Using Lemma 1.3 and (1.4), consider the following calculations:

$$\begin{aligned}
\lambda_a(e_1^4) &= u^4 = x x^{\mu} x^{\mu^2} x^{\mu^3} = x(x-y)(x-2y)(x-3y) =: r_1,\\
\lambda_a(e_1^2 e_2) &= u^2(u-av)v = x x^{\mu}(x-ay)^{\mu^2} y^{\mu^3} = x(x-y)(x-(2+a)y)y =: r_2,\\
\lambda_a(e_1 e_2 e_1) &= u(u-av)vu = x(x-ay)^{\mu} y^{\mu^2} x^{\mu^3} = x(x-(1+a)y)y(x-3y) =: r_3,\\
\lambda_a(e_2 e_1^2) &= (u-av)vu^2 = (x-ay)y^{\mu} x^{\mu^2} x^{\mu^3} = (x-ay)y(x-2y)(x-3y) =: r_4,\\
\lambda_a(e_2^2) &= (u-av)v(u-av)v = (x-ay)y^{\mu}(x-ay)^{\mu^2} y^{\mu^3}\\
&= (x-ay)y(x-(2+a)y)y =: r_5.
\end{aligned}$$

By direct computation, we see that $r_1, \dots, r_5$ are linearly independent if $a \neq 0, 1$.

Further, since $A(a)$ and R are equal in large degree and R is noetherian, $A(a)$ is noetherian by [Stafford 1985, Lemma 1.4]. $\square$

Next we compute the kernels of the maps λ_a and establish Theorem 0.6(a). We will use the following notation:

Notation 2.2 (π, π_a, π_B). Let $\Bbbk\langle t_1, t_2\rangle$ be the free algebra, which we grade by setting $\deg t_i = i$. We set the notation below:

- $\pi : \Bbbk\langle t_1, t_2\rangle \to U(W_+)$ is the algebra map given by $\pi(t_1) = e_1$ and $\pi(t_2) = e_2$.
- $\pi_a : \Bbbk\langle t_1, t_2\rangle \to R$ is the algebra map given by $\pi_a(t_1) = \lambda_a(e_1) = u$ and $\pi_a(t_2) = \lambda_a(e_2) = (u - av)v$, for $a \in \Bbbk$. The image of π_a is $A(a)$. Note that $\pi_a = \lambda_a \circ \pi$.
- $\pi_B : \Bbbk\langle t_1, t_2\rangle \to S$ is the algebra map given by $\pi_B(t_1) = \phi(e_1) = u$ and $\pi_B(t_2) = \phi(e_2) = uv - vw$. The image of π_B is B. Note that $\pi_B = \phi \circ \pi$.

In the next result, we compute a presentation of the algebra $A(0)$.

Lemma 2.3. *The kernel of π_0 is generated by*

$$
\begin{aligned}
q &:= t_1^2 t_2 - t_2 t_1^2 - 2t_2^2,\\
q' &:= t_1^3 t_2 - 3t_1^2 t_2 t_1 + 3t_1 t_2 t_1^2 - t_2 t_1^3 + 6t_2^2 t_1 - 12 t_2 t_1 t_2 + 6 t_1 t_2^2
\end{aligned}
$$

as a two-sided ideal.

Proof. Let $A = A(0)$, and consider the exact sequence of right A-modules

$$0 \longrightarrow K \longrightarrow A[-1] \oplus A[-2] \xrightarrow{(u,uv)} A \longrightarrow \Bbbk \longrightarrow 0.$$

Claim. *As a right A-module, K is generated by*

$$(u^2 v, -u(u+2v)) \quad \textit{and} \quad (u^2 v^2, -u(u+2v)v).$$

Assume the claim. It is well known that one may deduce generators and relations of a connected graded $\Bbbk$-algebra from the first few terms in a minimal resolution of the trivial module $\Bbbk$. The precise method is given in Proposition A.1 in the Appendix. Using the notation of that result, take

$$
\begin{aligned}
b_1^1 &= u^2 v, & b_2^1 &= -u(u+2v),\\
b_1^2 &= u^2 v^2, & b_2^2 &= -u(u+2v)v.
\end{aligned}
$$

Moreover, take

$$
\begin{aligned}
\tilde{b}_1^1 &= t_1 t_2, & \tilde{b}_2^1 &= -t_1^2 - 2t_2,\\
\tilde{b}_1^2 &= t_1^2 t_2 - t_1 t_2 t_1, & \tilde{b}_2^2 &= 2t_2 t_1 - 3t_1 t_2.
\end{aligned}
$$

Note that $\pi_0(\tilde{b}_j^i) = b_j^i$ for $i, j = 1, 2$. Now we obtain by Proposition A.1 that

$$
\begin{aligned}
q_1 &:= t_1(\tilde{b}_1^1) + t_2(\tilde{b}_2^1) = t_1^2 t_2 - t_2 t_1^2 - 2t_2^2,\\
q_2 &:= t_1(\tilde{b}_1^2) + t_2(\tilde{b}_2^2) = t_1^3 t_2 - t_1^2 t_2 t_1 + 2t_2^2 t_1 - 3t_2 t_1 t_2
\end{aligned}
$$

generate $\ker \pi_0$. Observe that $q = q_1$ and that

$$q' - 4q_2 = -3t_1^3 t_2 + t_1^2 t_2 t_1 + 3t_1 t_2 t_1^2 - t_2 t_1^3 - 2t_2^2 t_1 + 6t_1 t_2^2 = -3t_1 q + q t_1 \in (q).$$

Thus, $\ker \pi_0$ is generated by q and q', as desired.

So it remains to prove the claim.

Proof of claim. Note that there is an isomorphism of graded right A-modules $\beta : uA \cap uvA \to K$ given by $\beta(r) = (u^{-1}r, -(uv)^{-1}r)$.

Take $M := A \cap vA$. Since $A = \Bbbk + uR$, it is easy to show that $M = uR \cap vuR$, and in particular, that M is a right R-module. Since $(uR + vuR)_{\geq 2} = R_{\geq 2}$, we get that

$$\dim_{\Bbbk} M_n = \dim_{\Bbbk} R_{n-1} + \dim_{\Bbbk} R_{n-2} - \dim_{\Bbbk} R_n = n - 2$$

for $n \geq 2$, and $\dim_{\Bbbk} M_n = 0$ for $n < 2$. Moreover, $u^2 v = vu(u + 2v) \in M$, so $u^2 vR \subseteq M$ and $\operatorname{hilb}(u^2 vR) = \operatorname{hilb} M$. So, $M = u^2 vR$. Now

$$uA \cap uvA = uM = u^3 vR \overset{(*)}{=} u^3 vA + u^3 v^2 A = uvu(u+2v)A + uvu(u+2v)vA,$$

where the equality $(*)$ holds as $R = A + vA$. Apply the map β to the right-hand side of the equation above to yield the desired result. $\square$

We can now understand $\ker \lambda_0$ and $\ker \lambda_1$. We first prove:

Lemma 2.4. *We have* $\ker \lambda_0 = \ker \lambda_1$.

Proof. Working in the quotient division ring of R, we have

$$u^{-1} \lambda_0(e_n) u = v^{n-1} u = \lambda_1(e_n).$$

So for any $f \in U(W_+)$, we have $\lambda_1(f) = u^{-1} \lambda_0(f) u$. The result follows. $\square$

Proposition 2.5. *We have that* $\ker \lambda_a = (e_1 e_3 - e_2^2 - e_4)$ *for* $a = 0, 1$.

Proof. We first check that $e_1 e_3 - e_2^2 - e_4$ is indeed in $\ker \lambda_0$:

$$\lambda_0(e_1 e_3 - e_2^2 - e_4) = u(uv^2) - (uv)(uv) - uv^3 = u^2 v^2 - u(uv - v^2)v - uv^3 = 0.$$

Recall that $\pi_0 = \lambda_0 \circ \pi$. So, Lemma 2.3 implies that $\ker \lambda_0 = \pi(\ker \pi_0)$ is generated by elements $\pi(q)$ and $\pi(q')$ in $U(W_+)$. Now $\pi(q') = 0$ by Lemma 1.1(a), so $\ker \lambda_0$ is generated by $\pi(q)$. Moreover,

$$\begin{aligned} \pi(q) &= e_1^2 e_2 - e_2 e_1^2 - 2e_2^2 \\ &= 2\big(e_1(e_1 e_2 - e_2 e_1) - e_2^2 - \big(\tfrac{1}{2} e_1^2 e_2 - e_1 e_2 e_1 + \tfrac{1}{2} e_2 e_1^2\big)\big) = 2(e_1 e_3 - e_2^2 - e_4), \end{aligned}$$

using the relation $[e_n, e_m] = (m-n)e_{n+m}$ in $U(W_+)$. Thus, $\ker \lambda_0 = (e_1 e_3 - e_2^2 - e_4)$, as claimed.

The result for $a = 1$ now follows by Lemma 2.4. $\square$

It remains to analyze $\ker \lambda_a$ with $a \neq 0, 1$. We do this in the next two results.

Lemma 2.6. *For $a \neq 0, 1$, the kernel of λ_a is generated in degrees* 5 *and* 6.

Proof. Take $A' := A(a)$. It suffices to show that the kernel of π_a is generated in degrees 5 and 6. Consider the exact sequence of right A'-modules

$$0 \longrightarrow K \longrightarrow A'[-1] \oplus A'[-2] \xrightarrow{(u,(u-av)v)} A' \longrightarrow \Bbbk \longrightarrow 0.$$

We have that $uA' \cap (u-av)vA' \cong K$ as right A'-modules. As in the proof of Lemma 2.3, it now suffices to show that $uA' \cap (u-av)vA'$ is generated in degrees 5 and 6 as a right A'-module.

Let $J := uA' \cap (u-av)vA'$, and let $L := uR \cap (u-av)vR$. Note that $J \subseteq L$. Since $a \neq 0$, we get that $R_{\geq 2} = (uR + (u-av)vR)_{\geq 2}$. So,

$$\dim_\Bbbk L_n = \dim_\Bbbk R_{n-1} + \dim_\Bbbk R_{n-2} - \dim_\Bbbk R_n = n-2$$

for $n \geq 2$. So, $\dim_\Bbbk L_3 = 1$, and is principally generated as a right R-module by an element of degree 3. In fact,

$$L = rR, \quad \text{where } r := u(uv + (1-a)v^2) = (uv - av^2)(u+2v). \tag{2.7}$$

Since $A'_{\geq 4} = R_{\geq 4}$ by Proposition 2.1, we have $J_{\geq 6} = L_{\geq 6}$. By direct computation, one obtains that $J_i = 0$ for $i = 0, \dots, 4$; one can also use Routine A.2 in the Appendix.

Let $J' = J_5A' + J_6A'$. We prove by induction that $J_n = J'_n$, for all $n \geq 5$. The statement is clear for $n = 5, 6$. For $n = 7$, we make the following assertion, the proof of which is presented in the Appendix; see Claim A.3.

Claim. *We have that $J_5A'_2 \not\subseteq J_6A'_1$.*

So for $n \geq 6$, we have $J_n = L_n = rR_{n-3}$. So $\dim_\Bbbk J_7 = 5$, and $\dim_\Bbbk J_6A'_1 = \dim_\Bbbk J_6 = 4$. With the claim, we obtain $J_7 = J_5A'_2 + J_6A'_1$. Thus, $J_7 = J'_7$. Now for the induction step, suppose we have established that $J'_n = J_n$ and $J'_{n-1} = J_{n-1}$ for some $n \geq 7$. Then

$$\begin{aligned} J_{n+1} \supseteq J'_{n+1} &= J'_n u + J'_{n-1}(u-av)v = J_n u + J_{n-1}(u-av)v \\ &= r(R_{n-3}u + R_{n-4}(u-av)v) = rR_{n-2} = J_{n+1}. \end{aligned}$$

The penultimate equality holds as $a \neq 1$. Thus, the lemma is verified. □

Proposition 2.8. *If $a \neq 0, 1$, then* $\ker \lambda_a$ *is the ideal generated by the elements*

$$h_1 := e_1e_2^2 - e_1^2e_3 - (2a)e_2e_3 + (1+2a)e_1e_4 - (a^2+a)e_5,$$
$$h_2 := e_1e_5 - 4e_2e_4 + 3e_3^2 + 2e_6,$$
$$\begin{aligned} h_3 := -4e_1^2e_2^2 - 4e_2^3 + 4e_1^3e_3 + (20a^2+14a-7)e_3^2 \\ - (16a^2+18a+5)e_1e_5 + (16a^3+36a^2+16a-2)e_6. \end{aligned}$$

Proof. By Lemma 2.6, we just need to produce linearly independent elements of $\ker\lambda_a$ in degrees 5 and 6. We have by Routine A.2 that $\dim_{\Bbbk}(\ker\lambda_a)_5 = 1$ and that we can choose a basis of $(\ker\lambda_a)_5$ to be the element h_1. In fact, we verify that $\lambda_a(h_1) = 0$ using Lemma 1.3 and (1.4), while suppressing some μ superscripts:

$$\begin{aligned}
\lambda_a(h_1) &= u(u-av)v(u-av)v - u^2(u-2av)v^2 - (2a)(u-av)v(u-2av)v^2 \\
&\qquad\qquad + (1+2a)u(u-3av)v^3 - (a^2+a)(u-4av)v^4 \\
&= x(x-ay)^{\mu} y(x-ay)^{\mu^3} y - xx^{\mu}(x-2ay)^{\mu^2} y^2 \\
&\qquad - (2a)(x-ay)y(x-2ay)^{\mu^2} y^2 \\
&\qquad\qquad + (1+2a)x(x-3ay)^{\mu} y^3 - (a^2+a)(x-4ay)y^4 \\
&= x(x-(1+a)y)y(x-(3+a)y)y - x(x-y)(x-(2+2a)y)y^2 \\
&\qquad - (2a)(x-ay)y(x-(2+2a)y)y^2 \\
&\qquad\qquad + (1+2a)x(x-(1+3a)y)y^3 - (a^2+a)(x-4ay)y^4 \\
&= 0.
\end{aligned}$$

On the other hand, we have by Routine A.2 that $\dim_{\Bbbk}(\ker\lambda_a)_6 = 4$ and that we can take a basis of $(\ker\lambda_a)_6$ to be h_2, h_3 along with

$$\begin{aligned}
h_4 &:= 4e_2^3 - 4e_1e_2e_3 + (7-4a)e_3^2 + (1+4a)e_1e_5 + (2-4a-4a^2)e_6, \\
h_5 &:= 4e_2^3 + (7-14a)e_3^2 - 4e_1^2e_4 + (5+14a)e_1e_5 + (2-16a-12a^2)e_6.
\end{aligned}$$

By direct computation we have

$$\begin{aligned}
e_1h_1 &= e_1^2e_2^2 - e_1^3e_3 - (2a)e_1e_2e_3 + (1+2a)e_1^2e_4 - (a^2+a)e_1e_5, \\
h_1e_1 &= e_1e_2^2e_1 - e_1^2e_3e_1 - (2a)e_2e_3e_1 + (1+2a)e_1e_4e_1 - (a^2+a)e_5e_1 \\
&= e_1^2e_2^2 - e_1^3e_3 - (2+2a)e_1e_2e_3 + (2a)e_3^2 + (3+2a)e_1^2e_4 + (4a)e_2e_4 \\
&\qquad - (2+7a+a^2)e_1e_5 + 4(a^2+a)e_6.
\end{aligned}$$

Claim. *We have that h_2, h_3, e_1h_1, h_1e_1 are $\Bbbk$-linearly independent and that*

$$\begin{aligned}
h_4 &= 2a(2a+1)h_2 - h_3 - (6+4a)e_1h_1 + (2+4a)h_1e_1, \\
h_5 &= 4a^2h_2 - h_3 - (4+4a)e_1h_1 + (4a)h_1e_1.
\end{aligned}$$

The proof is presented in the Appendix; see Claim A.5. Thus, the result holds.

Now for the reader's convenience, we verify that $\lambda_a(h_i) = 0$ for $i = 2, 3$ using Lemma 1.3 and (1.4), while suppressing some μ superscripts:

$$\begin{aligned}
\lambda_a(h_2) &= u(u-4av)v^4 - 4(u-av)v(u-3av)v^3 \\
&\qquad\qquad + 3(u-2av)v^2(u-2av)v^2 + 2(u-5av)v^5 \\
&= x(x-4ay)^{\mu} y^4 - 4(x-ay)y(x-3ay)^{\mu^2} y^3 \\
&\qquad\qquad + 3(x-2ay)y^2(x-2ay)^{\mu^3} y^2 + 2(x-5ay)y^5
\end{aligned}$$

$$
\begin{aligned}
&= x(x-(1+4a)y)y^4 - 4(x-ay)y(x-(2+3a)y)y^3 \\
&\qquad\qquad + 3(x-2ay)y^2(x-(3+2a)y)y^2 + 2(x-5ay)y^5 \\
&= 0,
\end{aligned}
$$

$$
\begin{aligned}
\lambda_a(h_3) &= -4u^2(u-av)v(u-av)v - 4(u-av)v(u-av)v(u-av)v + 4u^3(u-2av)v^2 \\
&\qquad + (20a^2+14a-7)(u-2av)v^2(u-2av)v^2 \\
&\qquad\quad - (16a^2+18a+5)u(u-4av)v^4 \\
&\qquad\qquad + (16a^3+36a^2+16a-2)(u-5av)v^5 \\
&= -4xx^{\mu}(x-ay)^{\mu^2}y(x-ay)^{\mu^4}y \\
&\qquad - 4(x-ay)y(x-ay)^{\mu^2}y(x-ay)^{\mu^4}y + 4xx^{\mu}x^{\mu^2}(x-2ay)^{\mu^3}y^2 \\
&\qquad\quad + (20a^2+14a-7)(x-2ay)y^2(x-2ay)^{\mu^3}y^2 \\
&\qquad\qquad - (16a^2+18a+5)x(x-4ay)^{\mu}y^4 \\
&\qquad\qquad\quad + (16a^3+36a^2+16a-2)(x-5ay)y^5 \\
&= -4x(x-y)(x-(2+a)y)y(x-(4+a)y)y \\
&\qquad - 4(x-ay)y(x-(2+a)y)y(x-(4+a)y)y \\
&\qquad\quad + 4x(x-y)(x-2y)(x-(3+2a)y)y^2 \\
&\qquad\qquad + (20a^2+14a-7)(x-2ay)y^2(x-(3+2a)y)y^2 \\
&\qquad\qquad\quad - (16a^2+18a+5)x(x-(1+4a)y)y^4 \\
&\qquad\qquad\qquad + (16a^3+36a^2+16a-2)(x-5ay)y^5 \\
&= 0. \qquad\qquad \square
\end{aligned}
$$

3. Elementary proofs that $U(W_+)$ and $U(W)$ are not noetherian

In this section, we establish the remaining part of Theorem 0.7, that $\ker\lambda_0 = \ker\lambda_1$ is not finitely generated as a left or right ideal of $U(W_+)$. We also prove Theorem 0.8.

We first focus on $U(W_+)$. Recall the map $\phi : U(W_+) \twoheadrightarrow B$ from Definition 0.3, and consider Notations 0.2, 1.2, 1.6, and 2.2 along with the following.

Notation 3.1 (p, I). Let $p := \phi(e_1e_3 - e_2^2 - e_4)$ be an element of B, and let $I := (p)$ be a two-sided ideal of B. Note that by Proposition 2.5, $I = \phi(\ker\lambda_0) = \pi_B(\ker\pi_0)$.

We begin by establishing some basic facts about p and I.

Lemma 3.2.

(a) $p = v^3w - v^2w^2$.

(b) *p is a normal element of S and of Q.*

(c) $I = Qp$.

Proof. We employ Lemma 1.3 and (1.4) in all parts.

(a) Consider the computation in S below:

$$\begin{aligned}
p &= \phi(e_1e_3 - e_2^2 - e_4)\\
&= u(u-2w)v^2 - (u-w)v(u-w)v - (u-3w)v^3\\
&= x(x-2z)^{\mu} y^{\mu^2} y^{\mu^3} - (x-z)y^{\mu}(x-z)^{\mu^2} y^{\mu^3} - (x-3z)y^{\mu} y^{\mu^2} y^{\mu^3}\\
&= x(x-y-2z)y^2 - (x-z)y(x-2y-z)y - (x-3z)y^3\\
&= y^3z - y^2z^2\\
&= v^3w - v^2w^2.
\end{aligned}$$

(b) From part (a), we get that p is a normal element of S, and of Q, since $vp = pv$, $wp = pw$, and

$$\begin{aligned}
up = u(v^3w - v^2w^2) &= xy^{\mu} y^{\mu^2} y^{\mu^3} z^{\mu^4} - xy^{\mu} y^{\mu^2} z^{\mu^3} z^{\mu^4} = (y^3z - y^2z^2)x\\
&= (y^3z - y^2z^2)(x+4y)^{\mu^4} = (v^3w - v^2w^2)(u+4v) = p(u+4v).
\end{aligned}$$

(c) On one hand, we get that $I = BpB \subseteq QpQ = Qp$, by Lemma 1.7 and part (b). On the other hand, recall that R is the subalgebra of Q generated by u, v. We will show by induction on i and j that $p(vw)^i R_{j-2i} \subseteq I$ for all $0 \le i \le \lfloor \frac{1}{2} j \rfloor$; this yields $pQ_j \subseteq I$.

The base case $i = j = 0$ holds since $p \in I$. For the induction step, assume that $p(vw)^i R_{j-2i} \subseteq I$. Now it suffices to show that (i) $p(vw)^i R_{j+1-2i} \subseteq I$ and (ii) $p(vw)^{i+1} R_{j-2i} \subseteq I$.

For (i), we have by induction that

$$I \supseteq up(vw)^i R_{j-2i} + p(vw)^i R_{j-2i} u =: I',$$

since u is a generator of B. Now consider the following computations, where we suppress the action of μ on invariant elements and on graded pieces of $\Bbbk[x,y]$:

$$\begin{aligned}
I' &= x(y^3z - y^2z^2)(yz)^i \Bbbk[x,y]_{j-2i} + (y^3z - y^2z^2)(yz)^i \Bbbk[x,y]_{j-2i} x^{\mu^{j+4}}\\
&= (y^3z - y^2z^2)(yz)^i x\Bbbk[x,y]_{j-2i} + (y^3z - y^2z^2)(yz)^i (x + (j+4)y)\Bbbk[x,y]_{j-2i}\\
&= (y^3z - y^2z^2)(yz)^i \big(x\Bbbk[x,y]_{j-2i} + (x + (j+4)y)\Bbbk[x,y]_{j-2i}\big)\\
&= (y^3z - y^2z^2)(yz)^i \Bbbk[x,y]_{j+1-2i},
\end{aligned}$$

where the last equality holds since $j + 4 > 0$. Thus (i) holds.

For (ii), we get that $p(vw)^i R_{j+2-2i} \subseteq I$ by applying (i) twice. Now

$$I \supseteq p(vw)^i R_{j+2-2i} + p(vw)^i R_{j-2i}(uv - vw) \supseteq p(vw)^i R_{j-2i}(vw).$$

Note that $R_k(vw) = (vw)R_k$ for all k. So $I \supseteq p(vw)^{i+1} R_{j-2i}$ and we are done. □

Now we complete the proof of Theorem 0.7.

Theorem 3.3. *The ideal I of B is not finitely generated as a left or right ideal. As a result, the kernel of λ_0 is not finitely generated as a left or right ideal of $U(W_+)$.*

Proof. Recall that $\ker\lambda_0 = (e_1e_3 - e_2^2 - e_4)$ by Proposition 2.5. It is clear that if $\ker\lambda_0$ is finitely generated as a left/right ideal of $U(W_+)$, then I is finitely generated as a left/right ideal of B. Therefore, to show that $\ker\lambda_0$ is not finitely generated it suffices to show that ${}_BI$ and I_B are not finitely generated.

By way of contradiction, suppose that ${}_BI$ is finitely generated. Then there exists $n \geq 4$ such that $BI_{\leq n} = I$. Since B is generated by u and $(u-w)v$, we get that

$$I_{n+1} = B_1I_n + B_2I_{n-1} = uI_n + (u-w)vI_{n-1}. \tag{3.4}$$

By Lemma 3.2, $I = Qp \subseteq SpS = Sp$. Since $vI \subseteq vSp \subseteq Sp$, we get by (3.4) that

$$I_{n+1} \subseteq uSp + (u-w)Sp = uSp + wSp. \tag{3.5}$$

Using Lemma 1.3 and (1.4), it is easy to see that $uS + wS = x\Bbbk[x,y,z] + z\Bbbk[x,y,z]$ and that a positive power of y cannot belong to the right-hand side. So, a positive power of v cannot belong to $uS + wS$. Therefore,

$$v^{n-3}p \notin uSp + wSp. \tag{3.6}$$

On the other hand, $v^{n-3}p \in I_{n+1}$ by Lemma 3.2(c). This contradicts (3.5) and (3.6). Thus, ${}_BI$ is not finitely generated.

Next, suppose that I_B is finitely generated. Then there exists $n \geq 4$ such that $I_{\leq n}B = I$, with

$$I_{n+1} = I_nB_1 + I_{n-1}B_2 = I_nu + I_{n-1}(u-w)v = I_nu + I_{n-1}v(u+v-w). \tag{3.7}$$

We get that $I, Iv \subseteq pS$ by Lemma 3.2(b). So, the right-hand side of (3.7) is contained in $pSu + pS(v-w)$. With an argument similar to that in the previous paragraph, we obtain that $Su + S(v-w)$ does not contain positive powers of v. So, $pv^{n-3} \notin I_nu + I_{n-1}v(u+v-w)$. On the other hand, $pv^{n-3} \in I_{n+1}$ by Lemma 3.2(b,c), which contradicts (3.7). Thus, I_B is not finitely generated. □

Remark 3.8. We do not know whether or not $\ker\lambda_a$ is finitely generated for $a \neq 0, 1$.

One can of course deduce from Theorem 3.3 that $U(W)$ and $U(V)$ are neither left nor right noetherian; see [Sierra and Walton 2014, Lemma 1.7]. Nevertheless, a direct proof that $U(W)$ is not left or right noetherian is of independent interest, and we give such a result to end the section. First, we establish some notation.

Notation 3.9 ($\widehat{S}$, $\widehat{R}$, $\widehat{B}$, $\widehat{\phi}$, $\widehat{\lambda}_a$, η_a, $\widehat{I}$). Since v is normal in S and in R, we may invert it. Let $\widehat{S} := S[v^{-1}]$, and let $\widehat{R} := R[v^{-1}]$.

Note that ϕ extends to an algebra homomorphism $\widehat{\phi}: U(W) \to \widehat{S}$ defined by (0.4) for all $n \in \mathbb{Z}$. Likewise, λ_a extends to an algebra homomorphism $\widehat{\lambda}_a : U(W) \to \widehat{R}$ defined by (0.5) for all $n \in \mathbb{Z}$. For $a \in \Bbbk$ define $\eta_a : \widehat{S} \to \widehat{R}$ by $u \mapsto u$, $v \mapsto v$, $w \mapsto av$. Note that $\widehat{\lambda}_a = \eta_a \widehat{\phi}$.

Let $\widehat{B} := \widehat{\phi}(U(W))$. Finally, let $\widehat{I} = \widehat{\phi}(\ker \widehat{\lambda}_0)$. Note that $\widehat{I} = \widehat{B} \cap \ker \eta_0$.

We first note that the proof of Lemma 2.4 extends to $U(W)$ to give

$$\ker \widehat{\lambda}_0 = \ker \widehat{\lambda}_1. \tag{3.10}$$

Proposition 3.11. *Recall $p = \phi(e_1 e_3 - e_2^2 - e_4) = w(v-w)v^2$ from Notation 3.1 and Lemma 3.2. We have*

$$\widehat{I} = \widehat{B} \cap \ker \eta_0 = \widehat{B} \cap \ker \eta_1 = \widehat{B} p \widehat{B} = \widehat{S} p = p \widehat{S}.$$

Proof. We first show that $\widehat{B} p \widehat{B} = \widehat{S} p = p \widehat{S}$. Certainly, $\widehat{B} p \widehat{B} \subseteq \widehat{S} p \widehat{S} = \widehat{S} p = p \widehat{S}$, where the last two equalities hold because a normal element of S will also be normal in $\widehat{S}$.

For the other direction, we will show $\widehat{R} w^j p \subseteq \widehat{B} p \widehat{B}$ for all $j \geq 0$ by induction. Since $\widehat{S} = \widehat{R} \cdot \Bbbk[w]$, this will imply $\widehat{S} p \subseteq \widehat{B} p \widehat{B}$. So assume $w^j p \in \widehat{B} p \widehat{B}$ for some $j \geq 0$ (it is clear for $j = 0$). Since $up = p(u+4v)$, we get that for all $n \in \mathbb{Z}$,

$$\begin{aligned} \widehat{B} p \widehat{B} \ni [\widehat{\phi}(e_n), w^j p] &= (u-(n-1)w)v^{n-1} w^j p - w^j p (u-(n-1)w)v^{n-1} \\ &= (j+4)v^n w^j p. \end{aligned}$$

So, $\Bbbk[v, v^{-1}] \cdot w^j p \subseteq \widehat{B} p \widehat{B}$. Since $u = \widehat{\phi}(e_1) \in \widehat{B}$, we have

$$\widehat{R} w^j p = \Bbbk[u] \cdot \Bbbk[v, v^{-1}] \cdot w^j p \subseteq \widehat{B} p \widehat{B}.$$

Finally, since we have seen that $v^{-1} w^j p \in \widehat{R} w^j p \subseteq \widehat{B} p \widehat{B}$, we have that

$$\widehat{B} p \widehat{B} \ni (\widehat{\phi}(e_1) - \widehat{\phi}(e_2) v^{-1}) w^j p = w^{j+1} p.$$

By induction, $\widehat{B} p \widehat{B} = \widehat{S} p$, as desired.

From the definitions, $p \in (\ker \eta_0) \cap (\ker \eta_1)$. So

$$\widehat{B} p \widehat{B} \subseteq (\ker \eta_0) \cap (\ker \eta_1) \cap \widehat{B} = w \widehat{S} \cap (v-w) \widehat{S} = w(v-w) \widehat{S} = p \widehat{S}.$$

Combining this with the first part of the proof, $\widehat{B} p \widehat{B} = (\ker \eta_0) \cap (\ker \eta_1) \cap \widehat{B}$. Then by (3.10) and the definition of $\widehat{I}$, we have

$$\widehat{I} = (\ker \eta_0) \cap \widehat{B} = \widehat{\phi}(\ker \widehat{\lambda}_0) = \widehat{\phi}(\ker \widehat{\lambda}_1) = (\ker \eta_1) \cap \widehat{B},$$

completing the proof. □

From Proposition 3.11 we obtain:

Theorem 3.12. *The ideal $\widehat{I}$ of $\widehat{B}$ is not finitely generated as a left or right ideal. As a result, the kernel of $\widehat{\lambda}_0$ is not finitely generated as a left or right ideal of $U(W)$.*

Proof. This argument is similar to the proof of Theorem 3.3. It suffices to show that $\widehat{I}$ is not finitely generated as a left or right ideal of $\widehat{B}$.

By way of contradiction, suppose we have $\widehat{I} = \widehat{B}(\widehat{I}_{-n} \oplus \cdots \oplus \widehat{I}_n)$ for some $n \in \mathbb{N}$. For all $k \in \mathbb{Z}$, we have $\widehat{\phi}(e_k) \in u\widehat{S} + w\widehat{S}$. So, $\widehat{B}_k \subseteq u\widehat{S} + w\widehat{S}$ for all $k \neq 0$, and $\widehat{I}_k \subseteq u\widehat{S} + w\widehat{S}$ for all k with $|k| > n$. Note that a power of v cannot belong to $u\widehat{S} + w\widehat{S}$. So, $v^{n-3} p \notin \widehat{I}$. However, by Proposition 3.11, we get that $\widehat{I} = \widehat{S} p$ and $v^{n-3} p \in \widehat{I}$. This contradiction shows that ${}_{\widehat{B}}\widehat{I}$ is not finitely generated.

The proof that $\widehat{I}_{\widehat{B}}$ is not finitely generated is similar; we leave the details to the reader. □

Corollary 3.13. *The universal enveloping algebra $U(V)$ is neither left nor right noetherian.*

Proof. This follows directly from Theorem 3.12, since $U(W) = U(V)/(c)$. □

Remark 3.14. After the first draft of this paper was finished, we learned of the results of Conley and Martin [2007]. We thank the referee for calling that work to our attention. The paper considers a family of homomorphisms defined as (using their notation)

$$\pi_\gamma : U(W) \to \Bbbk[x, x^{-1}, \partial], \quad e_n \mapsto x^{n+1}\partial + (n+1)\gamma x^n.$$

Using the identification $u = x^2\partial$, $v = x$ from the discussion after Theorem 0.7, we have

$$\widehat{\lambda}_a(e_n) = (x^2\partial - (n-1)ax)x^{n-1} = x^{n+1}\partial + (1-a)(n-1)x^n.$$

The reader may verify that

$$\widehat{\lambda}_a(e) = x^{2(1-a)}\pi_{1-a}(e)x^{-2(1-a)}$$

for all $e \in U(W)$ (where here one uses a suitable extension of $\Bbbk[x, x^{-1}, \partial]$ to carry out computations). As a result,

$$\ker \widehat{\lambda}_a = \ker \pi_{1-a} \tag{3.15}$$

for all $a \in \Bbbk$.

Conley and Martin [2007, Theorem 1.2] showed (using (3.15)) that

$$\ker \widehat{\lambda}_0 = \ker \widehat{\lambda}_1 = (e_{-1}e_2 - e_0e_1 - e_1).$$

Recall from Proposition 2.5 that $\ker \lambda_0$ is generated as a two-sided ideal by $g_4 := e_1e_3 - e_2^2 - e_4$. A computation gives that

$$\mathrm{ad}(e_{-1}^3)(g_4) = [e_{-1}, [e_{-1}, [e_{-1}, g_4]]] = 12(e_{-1}e_2 - e_0e_1 - e_1),$$

and it follows that

$$(g_4) = \ker \widehat{\lambda}_0 = \ker \widehat{\lambda}_1 = (e_{-1}e_2 - e_0e_1 - e_1).$$

4. The connection between the maps ϕ and ρ

For the remainder of the paper, we return to considering $U(W_+)$. The main goal of this section is to relate the map ϕ (of Definition 0.3) that played a crucial role in the proof of Theorem 3.3 to the map ρ (of Notation 0.9) that was the focus of [Sierra and Walton 2014]. We show that $\ker\phi = \ker\rho$; in fact, we have the next result.

Theorem 4.1. *We have that* $\ker\rho = \ker\phi = \bigcap_{a\in\Bbbk}\ker\lambda_a$. *As a consequence,* $\rho(U(W_+)) \cong \phi(U(W_+))$.

Consider Notation 0.2 and the following notation for this section. Recall the definitions of X, f, τ from Notation 0.9. So, $\tau \in \mathrm{Aut}(X)$ and $\tau^* : \Bbbk(X) \to \Bbbk(X)$ is the pullback of τ. Here we take $\mu \in \mathrm{Aut}(\mathbb{P}^2)$ and $\nu \in \mathrm{Aut}(\mathbb{P}^1)$ to be morphisms of varieties, defined by

$$\mu([x:y:z]) = [x-y:y:z] \quad \text{and} \quad \nu([x:y]) = [x-y:y].$$

We denote the respective pullback morphisms by μ^* and ν^*. However, to be consistent with Lemma 1.3 (and abusing notation slightly), we still write

$$S \cong \Bbbk[x,y,z]^{\mu} \quad \text{and} \quad R \cong \Bbbk[x,y]^{\nu}.$$

We also establish the convention that $h^{\tau} := \tau^* h$ for $h \in \Bbbk(X)$, and similarly for pullback by other morphisms.

Before proving Theorem 4.1, we provide some preliminary results.

Lemma 4.2 (ψ_a, Ψ_a). *For $a \in \Bbbk$, we have the following statements.*

(a) *We have a well-defined morphism $\psi_a : \mathbb{P}^1 \to X$ given by*

$$\psi_a([x:y]) = [2x^2-4xy-6ay^2 : x^2-2xy+y^2 : -x^2+3xy-2y^2 : x^2-4xy+4y^2].$$

(b) $\psi_a\nu = \tau\psi_a$.

(c) ψ_a^* *extends to an algebra homomorphism* $\Psi_a : \Bbbk(X)[t;\tau^*] \to \Bbbk(\mathbb{P}^1)[s;\nu^*]$, *where* $\Psi_a(t) = s$.

Proof. (a,b) Both are straightforward. Part (a) is a direct computation. On page 508 in the Appendix, we verify that $(\psi_a\nu)^* = \nu^*\psi_a^* = \psi_a^*\tau^* = (\tau\psi_a)^*$ as maps from $\Bbbk(X) \to \Bbbk(\mathbb{P}^1)$. Thus, (b) holds.

(c) We have for all $h, \ell \in \Bbbk(X)$ and $n, m \in \mathbb{N}$ that

$$\begin{aligned}\Psi_a(ht^n\ell t^m) &= \Psi_a(h\ell^{\tau^n}t^{n+m}) = \psi_a^*(h)\psi_a^*(\ell^{\tau^n})s^{n+m} \\ &= \psi_a^*(h)\psi_a^*(\ell)^{\nu^n}s^{n+m} = \psi_a^*(h)s^n\psi_a^*(\ell)s^m = \Psi_a(ht^n)\Psi_a(\ell t^m).\end{aligned}$$

Thus, Ψ_a is an algebra homomorphism. $\square$

Lemma 4.3 (C_a). *For $a \in \Bbbk$, define the curve*

$$C_a = V(w + 6ax + (4 + 12a)y + (2 + 6a)z, xz - y^2) \subseteq X.$$

Then ψ_a defines an isomorphism from $\mathbb{P}^1 \to C_a$.

Proof. That the image of ψ_a of Lemma 4.2(a) is contained in C_a is a straightforward verification. The inverse map to ψ_a is defined by the birational map $[w : x : y : z] \mapsto [2x + y : x + y]$; we leave the verification of the details to the reader. □

Lemma 4.4 (γ). *Define a map $\gamma : R \to \Bbbk(\mathbb{P}^1)[s; \nu^*]$ as follows: if $h \in R_n = \Bbbk[x, y]_n$, let*

$$\gamma(h) = \frac{h}{x(x-y)\cdots(x-(n-1)y)} s^n.$$

Then γ is an injective $\Bbbk$-algebra homomorphism.

Proof. Let $h \in \Bbbk[x, y]_n$ and $\ell \in \Bbbk[x, y]_m$. Then

$$\begin{aligned}\gamma(h*\ell) = \gamma(h\ell^{\nu^n}) &= \frac{h\ell^{\nu^n}}{x(x-y)\cdots(x-(n+m-1)y)} s^{n+m} \\ &= \frac{h}{x(x-y)\cdots(x-(n-1)y)} \left(\frac{\ell}{x(x-y)\cdots(x-(m-1)y)}\right)^{\nu^n} s^{m+n} \\ &= \frac{h}{x(x-y)\cdots(x-(n-1)y)} s^n \frac{\ell}{x(x-y)\cdots(x-(m-1)y)} s^m = \gamma(h)\gamma(\ell).\end{aligned}$$

So, γ is a homomorphism; injectivity is clear. □

Proposition 4.5. *Retain the notation of Lemmas 4.2 and 4.4. Let $a \in \Bbbk$. Then $\Psi_a \rho = \gamma\lambda_a$ as maps from $U(W_+) \to \Bbbk(\mathbb{P}^1)[s; \nu^*]$, and* $\ker \Psi_a \rho = \ker \lambda_a$.

Proof. By Lemma 1.1(a), it suffices to verify that the maps $\Psi_a\rho$ and $\gamma\lambda_a$ agree on e_1 and e_2. We have

$$\Psi_a(\rho(e_1)) = \Psi_a(t) = s = \gamma(u) = \gamma(\lambda_a(e_1)).$$

We verify that

$$\psi_a^*(f) = \frac{xy - ay^2}{x^2 - xy} \tag{4.6}$$

on page 508 in the Appendix. Thus,

$$\Psi_a(\rho(e_2)) = \psi_a^*(f)s^2 = \frac{xy - ay^2}{x^2 - xy} s^2 = \gamma(uv - av^2) = \gamma\lambda_a(e_2).$$

The final statement follows from the fact that γ is injective (Lemma 4.4). □

Proof of Theorem 4.1. By Lemma 4.3, $\psi_a^* h = 0$ if and only if $h|_{C_a} \equiv 0$. Now, the curves C_a cover an open subset of X. (One way to see this is that, because $\bigcup_a C_a$ is dense in X and is clearly constructible, by [Hartshorne 1977, Exercise II.3.19(b)] it contains an open subset of X.) Thus if $h \in \Bbbk(X)$ is in the intersection $\bigcap_a \ker \psi_a^*$, then h vanishes on this open subset and so $h = 0$. So, $\bigcap_a \ker \Psi_a = \{0\}$. Thus, $\ker \rho = \bigcap_a \ker \Psi_a \rho = \bigcap_a \ker \lambda_a$, where the last equality holds by Proposition 4.5.

To show that $\ker \phi = \bigcap_a \ker \lambda_a$, define closed immersions $i_a : \mathbb{P}^1 \to \mathbb{P}^2$ for $a \in \Bbbk$ by $i_a([x : y]) = [x : y : ay]$. Then $\operatorname{im}(i_a) = V(z - ay)$, and pullback along i_a induces the ring homomorphism

$$i_a^* : \Bbbk[x, y, z] \to \Bbbk[x, y], \quad x \mapsto x, \; y \mapsto y, \; z \mapsto ay.$$

The reader may verify that $i_a \nu = \mu i_a$, and that i_a^* is also a homomorphism from $S = \Bbbk[x, y, z]^\mu$ to $R = \Bbbk[x, y]^\nu$. In terms of u, v, w, we have

$$i_a^*(u) = u, \quad i_a^*(v) = v, \quad i_a^*(w) = av.$$

That is, $i_a^* = \eta_a|_S$, where η_a was defined in Notation 3.9. We see that $i_a^* \phi = \lambda_a$.

As with the first paragraph, the curves $V(z - ay)$ cover an open subset of $\mathbb{P}^2$: in fact, $\bigcup_a V(z - ay) \supseteq (\mathbb{P}^2 \smallsetminus V(y))$. So $\bigcap_a \ker i_a^* = \{0\}$. Thus, $\ker \phi = \bigcap_a \ker i_a^* \phi = \bigcap_a \ker \lambda_a$, completing the proof. ☐

5. The kernel of ϕ

In this section, we analyze the map ϕ from Definition 0.3. In particular, we verify part (c) of Theorem 0.6. To proceed, recall Notations 0.2, 1.2, 1.6, and 2.2.

Theorem 5.1. *The kernel of ϕ is generated as a two-sided ideal by*

$$g := e_1 e_5 - 4e_2 e_4 + 3e_3^2 + 2e_6.$$

Proof. First, observe that as $e_1 e_5, e_2 e_4, e_3^2, e_6$ are elements of the standard basis for $U(W_+)$ (by Lemma 1.1(b)), they are linearly independent. So, we have that $g \neq 0$.

Now we verify that $\phi(g) = 0$ by using Lemma 1.3 and (1.4):

$$\begin{aligned}
\phi(g) &= u(u-4w)v^4 - 4(u-w)v(u-3w)v^3 + 3(u-2w)v^2(u-2w)v^2 + 2(u-5w)v^5 \\
&= x(x-4z)^\mu y^4 - 4(x-z)y(x-3z)^{\mu^2} y^3 \\
&\qquad + 3(x-2z)y^2(x-2z)^{\mu^3} y^2 + 2(x-5z)y^5 \\
&= x(x-y-4z)y^4 - 4(x-z)y(x-2y-3z)y^3 \\
&\qquad + 3(x-2z)y^2(x-3y-2z)y^2 + 2(x-5z)y^5 \\
&= 0.
\end{aligned}$$

We take the following notation for the rest of the proof.

Notation 5.2 (M, M', b_5, b_6, b_7, η). Consider the right B-modules

$$M := uB \cap (u-w)vB \quad \text{and} \quad M' := b_5B + b_6B + b_7B,$$

with

$$b_5 = (uv-vw)(u^3-6(uv-vw)u+12u(uv-vw)),$$
$$b_6 = (uv-vw)(-48(uv-3vw)v^2-36u(uv-2vw)v+u^4),$$
$$b_7 = (uv-vw)(u^5-40((uv-vw)^2u-3(uv-vw)u(uv-vw)+3u(uv-vw)^2)).$$

Further, take $\eta : B \to A(0)$ to be the map induced by the projection $\eta_0 : \widehat{S} \twoheadrightarrow \widehat{R} = \widehat{S}/(w)$ from Notation 3.9.

The remainder of the proof will be established through a series of lemmas.

Lemma 5.3. *We obtain that $b_5, b_6, b_7 \in uB \cap (u-w)vB$. In other words, $M' \subseteq M$.*

Proof. Let

$$r_5 := e_2(e_1^3 - 6e_2e_1 + 12e_1e_2), \tag{5.4}$$
$$r_6 := e_2(-48e_4 - 36e_1e_3 + e_1^4), \tag{5.5}$$
$$r_7 := e_2(e_1^5 - 40(e_2^2e_1 - 3e_2e_1e_2 + 3e_1e_2^2)). \tag{5.6}$$

We have as a consequence of the degree-5 relation of $U(W_+)$ in Lemma 1.1(a) that

$$r_5 = e_1(e_1^2e_2 - 3e_1e_2e_1 + 3e_2e_1^2 + 6e_2^2), \tag{5.7}$$

and as a consequence of the degree-7 relation of $U(W_+)$ in Lemma 1.1(a) that

$$r_7 = e_1(e_1^4e_2 - 5e_1^3e_2e_1 + 10e_1^2e_2e_1^2 - 10e_1e_2e_1^3 + 5e_2e_1^4 - 40e_2^3). \tag{5.8}$$

Thus $r_5, r_7 \in e_1U(W_+) \cap e_2U(W_+)$. Since $b_5 = \phi(r_5)$ and $b_7 = \phi(r_7)$, these are both in $uB \cap (uv - vw)B$.

Note that $r_6 \in e_2U(W_+)$, so $b_6 = \phi(r_6) \in (u-w)vB$. Further,

$$r_6 = e_1(-36e_2e_3 - 18e_5 + 2e_4e_1 - e_3e_1^2 + e_2e_1^3) + 12g.$$

Thus, $b_6 \in uB$ as well. □

Lemma 5.9. *Suppose that $M = M'$. Then* $\ker\phi = (g)$ *and the theorem holds.*

Proof. Let K be the kernel of

$$\alpha : B[-1] \oplus B[-2] \to B, \quad (b, b') \mapsto (ub + (uv - vw)b').$$

It is a standard fact that the map

$$\beta : M \to K$$

defined by $\beta(r) = (u^{-1}r, -(uv - vw)^{-1}r)$ is an isomorphism of graded right B-modules, as in the proof of Lemma 2.3. Thus, K is generated by $\beta(b_5)$, $\beta(b_6)$, and $\beta(b_7)$ by the assumption. By Proposition A.1 in the Appendix, the kernel of π_B is generated as a two-sided ideal of $\Bbbk\langle t_1, t_2\rangle$ by a degree-5 element q_5, a degree-6 element q_6, and a degree-7 element q_7. We compute q_5 and q_7 by applying the formula from Proposition A.1 to $\beta(b_5)$ and $\beta(b_7)$, and by using (5.4)–(5.8). Namely, take

$$\begin{aligned}
\tilde{b}_1^1 &= t_1^2t_2 - 3t_1t_2t_1 + 3t_2t_1^2 + 6t_2^2,\\
\tilde{b}_2^1 &= -t_1^3 + 6t_2t_1 - 12t_1t_2,\\
\tilde{b}_1^2 &= t_1^4t_2 - 5t_1^3t_2t_1 + 10t_1^2t_2t_1^2 - 10t_1t_2t_1^3 + 5t_2t_1^4 - 40t_2^3,\\
\tilde{b}_2^2 &= -t_1^5 + 40(t_2^2t_1 - 3t_2t_1t_2 + 3t_1t_2^2).
\end{aligned}$$

So, we have that

$$\begin{aligned}
q_5 &= t_1\tilde{b}_1^1 + t_2\tilde{b}_2^1 = [t_1,[t_1,[t_1,t_2]]] + 6[t_2,[t_2,t_1]],\\
q_7 &= t_1\tilde{b}_1^2 + t_2\tilde{b}_2^2 = [t_1,[t_1,[t_1,[t_1,[t_1,t_2]]]]] + 40[t_2,[t_2,[t_2,t_1]]].
\end{aligned}$$

By Lemma 1.1(a), q_5 and q_7 generate the kernel of π. So, $\ker\phi = \pi(\ker\pi_B) = (\pi(q_6))$. We see immediately that $(\ker\phi)_6$ is a 1-dimensional $\Bbbk$-vector space, generated by $\pi(q_6)$. Since $g \in (\ker\phi)_6$ is nonzero, we have $g = \pi(q_6)$ up to a scalar multiple. Therefore, $\ker\phi = (g)$. □

Our goal now is to show that $M = M'$; we do this by comparing Hilbert series.

Lemma 5.10. *The Hilbert series of M is $t^5(1-t)^{-2}(1-t^2)^{-1}$.*

Proof. Since $A(0) = \Bbbk \oplus uR$ we have

$$\operatorname{hilb} A(0) = 1 + t(\operatorname{hilb} R) = 1 + \frac{t}{(1-t)^2} = \frac{1-t+t^2}{(1-t)^2}.$$

On the other hand, it is well known that

$$\operatorname{hilb} Q = \operatorname{hilb} \Bbbk[x, y, yz] = \frac{1}{(1-t)^2(1-t^2)}.$$

Since $\lambda_0 = \eta \circ \phi$, we get that $\ker\eta = \phi(\ker\lambda_0)$ (which is denoted by I in Notation 3.1). So, by Lemma 3.2(c), we get

$$\operatorname{hilb}\ker\eta = \frac{t^4}{(1-t)^2(1-t^2)}.$$

Then

$$\begin{aligned}\operatorname{hilb} B &= \operatorname{hilb} A(0) + \operatorname{hilb} \ker \eta \\ &= \frac{1-t+t^3-t^4}{(1-t)^2(1-t^2)} + \frac{t^4}{(1-t)^2(1-t^2)} = \frac{1-t+t^3}{(1-t)^2(1-t^2)}.\end{aligned}$$

Finally, we compute $\operatorname{hilb} M$ from the exact sequence

$$0 \longrightarrow M \xrightarrow{\beta} B[-1] \oplus B[-2] \xrightarrow{\alpha} B \longrightarrow \Bbbk \longrightarrow 0,$$

where α, β are as in the proof of Lemma 5.9. This gives

$$\operatorname{hilb} M = (t^2+t-1)(\operatorname{hilb} B) + 1 = \frac{t^5}{(1-t)^2(1-t^2)},$$

as claimed. □

We now provide results on the Hilbert series of M'.

Lemma 5.11. *We have that* $\operatorname{hilb} \eta(M') \geq t^5(1-t)^{-2}$.

Proof. Let $a_5 := \eta(b_5)$ and $a_6 := \eta(b_6)$. Then

$$\begin{aligned}a_5 &= uvu(u^2 - 6vu + 12uv) \\ &= xy(x-2y)\big((x-3y)(x-4y) - 6y(x-4y) + 12(x-3y)y\big) \\ &= x^2(x-y)(x-2y)y, \\ a_6 &= uvu(u^3 - 36uv^2 - 48v^3) \\ &= xy(x-2y)\big((x-3y)(x-4y)(x-5y) - 36(x-3y)y^2 - 48y^3\big) \\ &= x^2(x-y)(x-2y)y(x-11y) \\ &= a_5(u-6v).\end{aligned}$$

Since a_5u and $a_5(u-6v)$ are in $\eta(M')$ and u and $u-6v$ span R_1, we have $a_5R_1 \subseteq \eta(M')$. We get that $\eta(M') \supseteq a_5A(0) + a_5R_1A(0)$, as $\eta(M')$ is a right $A(0)$-module and contains $a_5R_{\leq 1}$. Since $A(0) + R_1A(0) = R$, we obtain that $\eta(M') \supseteq a_5R$. Now as $\operatorname{hilb} R = (1-t)^{-2}$, we conclude that $\operatorname{hilb} \eta(M') \geq t^5(1-t)^{-2}$. □

Lemma 5.12. *We have that* $\operatorname{hilb}(M' \cap \ker \eta) \geq t^7(1-t)^{-2}(1-t^2)^{-1}$.

Proof. Again, recall that $\ker \eta = \phi(\ker \lambda_0)$, which is denoted by I in Notation 3.1. Moreover by Lemma 3.2(c), we have $I = Qp = pQ$, where $p = v^3w - v^2w^2$. Let

$$h := (uv - vw)(u+2v)p = (xy - yz)x(y^3z - y^2z^2).$$

Claim. *We have*

$$b_5Q + b_6Q + b_7Q \ni x(xy - yz)(xyz + y^2z) = (uv - vw)(u+2v)(u+4v)vw.$$

The proof of this claim is provided in the Appendix; see Claim A.6(a).

Since $M' \cap I \supseteq M'I = b_5 Qp + b_6 Qp + b_7 Qp$, we have

$$M' \cap I \supseteq (uv - vw)(u+2v)(u+4v)vwpQ \tag{5.13}$$
$$= (xy - yz)x(y^3z - y^2z^2)(x+y)yzQ = h(x+y)yzQ.$$

We now show by induction that $M' \cap I \supseteq hQ_n$ for all $n \geq 0$.

Claim. *$M' \cap I \supseteq hQ_n$ for $n = 0, 1, 2$.*

The proof of this assertion is provided in the Appendix; see Claim A.6(b). We will prove the result for larger n by geometric arguments. The maximal graded nonirrelevant ideals of $\Bbbk[x, y, yz]$ are in bijective correspondence with $\Bbbk$-points of the weighted projective plane $\mathbb{P}(1, 1, 2)$ [Harris 1992, Example 10.27]. We use the notation $(a : b : c)$ to denote a point of $\mathbb{P}(1, 1, 2)$. Let

$$K(n) := (x - ny)\Bbbk[x, y, yz] + (y^2 - yz)\Bbbk[x, y, yz]$$

be the graded ideal of polynomials vanishing at $(n : 1 : 1)$.

Suppose now that $M' \cap I \supseteq hQ_n$ for some $n \geq 2$. Then $M' \cap I$ contains

$$\begin{aligned} h(Q_n u + Q_{n-1}(uv - vw)) &\\ &= h\big((x - (n+7)y)\Bbbk[x, y, yz] + ((x - (n+6)y)y - yz)\Bbbk[x, y, yz]\big)_{n+1} \\ &= h\big((x - (n+7)y)\Bbbk[x, y, yz] + (y^2 - yz)\Bbbk[x, y, yz]\big)_{n+1} \\ &= hK(n+7)_{n+1}. \end{aligned}$$

From (5.13), we get $(M' \cap I)_{n+1} \ni h(xyz + y^2z)y^{n-2}$. Since $(xyz + y^2z)y^{n-2}$ does not vanish at $(n + 7 : 1 : 1)$, it is not in $hK(n + 7)_{n+1}$. Thus,

$$hK(n+7)_{n+1} + \Bbbk h(xyz + y^2z)y^{n-2} = h\Bbbk[x, y, yz]_{n+1} \subseteq M' \cap I,$$

where the equality holds as $hK(n + 7)_{n+1}$ is codimension 1 in $h\Bbbk[x, y, yz]_{n+1}$. Hence, $hQ_{n+1} \subseteq M' \cap I$.

Now, by induction, we obtain $M' \cap I \supseteq hQ$. Since $\operatorname{hilb} Q = (1-t)^{-2}(1-t^2)^{-1}$, we have

$$\operatorname{hilb}(M' \cap I) \geq \frac{t^7}{(1-t)^2(1-t^2)}. \qquad \square$$

Lemma 5.14. *We have that* $\operatorname{hilb} M = \operatorname{hilb} M' = t^5(1-t)^{-2}(1-t^2)^{-1}$. *As a result,* $M = M'$.

Proof. Combining Lemmas 5.11 and 5.12, we have

$$\operatorname{hilb}(M') \geq \frac{t^5}{(1-t)^2} + \frac{t^7}{(1-t)^2(1-t^2)} = \frac{t^5}{(1-t)^2(1-t^2)}.$$

On the other hand, by Lemmas 5.3 and 5.10,

$$\operatorname{hilb}(M') \leq \frac{t^5}{(1-t)^2(1-t^2)}.$$

Thus, $\operatorname{hilb} M = \operatorname{hilb} M'$. Since $M' \subseteq M$ again by Lemma 5.3, we conclude that $M = M'$. □

Theorem 5.1 now follows from Lemmas 5.9 and 5.14. □

Remark 5.15. A shorter proof of Theorem 5.1 follows from the results of [Conley and Martin 2007]. Recall from Notation 3.9 that we may extend ϕ to a map $\widehat{\phi} : U(W) \to \widehat{S}$, using the same formula (0.4) for $\widehat{\phi}(e_n)$ with $n \leq 0$. Then [Conley and Martin 2007, Theorem 1.3] and (3.15), together with Theorem 4.1, give that $\ker \widehat{\phi} = (e_{-1}e_3 - 4e_0e_2 + 3e_1^2 + 2e_2)$. The reader may verify that

$$\operatorname{ad}(e_{-1}^4)(g) = [e_{-1}, [e_{-1}, [e_{-1}, [e_{-1}, g]]]] = 24(e_{-1}e_3 - 4e_0e_2 + 3e_1^2 + 2e_2).$$

Since $\widehat{\phi}(g) = 0$, we have $(g) \subseteq \ker \widehat{\phi} = (e_{-1}e_3 - 4e_0e_2 + 3e_1^2 + 2e_2) \subseteq (g)$, so all are equal.

6. A partial result on chains of two-sided ideals

It is not known whether $U(W_+)$ satisfies the ascending chain condition (ACC) on two-sided ideals; see Question 0.11. We do not answer this question here; however, we prove the partial result that the non-noetherian factor B of $U(W_+)$ does have ACC on two-sided ideals.

Recall Notations 0.2, 1.2, 1.6; in particular, Q is the subalgebra of S generated by u, v, vw. Throughout, we consider B as a subalgebra of Q. We begin by proving:

Lemma 6.1. *Let h be a nonzero, homogeneous, normal element of Q, and let $a \in \Bbbk$. Then the Q-bimodules*

$$N := hQ/hvQ \quad \text{and} \quad M_a = hQ/h(vw - av^2)Q$$

are noetherian B-bimodules under the action induced from Q.

Proof. We remark that any normal element of Q must be in the commutative subalgebra $\Bbbk[v, vw]$, and thus, must commute with v and vw. In particular, $vQN = 0$ and $(vw - av^2)QM_a = 0 = M_a(vw - av^2)Q$.

Let $\theta : Q \to Q/vQ$ be the canonical projection. (Note that $vw \notin \ker \theta$.) Since $u(vw) - (vw)u = 2v^2w$ is contained in $\ker \theta$, the image Q/vQ is commutative. It is easy to see that $Q/vQ \cong \Bbbk[s, t]$ under the identification $s = \theta(u)$, $t = \theta(vw) = \theta(uv - vw)$. Note that $s = \theta(\phi(e_1))$ and $t = \theta(\phi(e_2))$ are in B. So, $\theta(B) = Q/vQ$. Thus, a left B-submodule of hQ/hvQ is simply an ideal of $\Bbbk[s, t]$. So, hQ/hvQ

is noetherian as a left B-module. As chains of B-bimodules are also chains of left B-modules, hQ/hvQ is also a noetherian B-bimodule.

Now define an algebra homomorphism $\delta : Q \to R$ by $\delta(u) = u$, $\delta(v) = v$, and $\delta(vw) = av^2$. (Note that $\delta = \eta_a|_Q$ from Notation 3.9.) It is easy to see that $\ker \delta = (vw - av^2)Q$ and that δ is surjective. Note also that $\delta(\phi(e_1)) = u$ and $\delta(\phi(e_2)) = uv - av^2$. Thus, $\delta(B) = A(a)$ as subalgebras of R. If $a \neq 0, 1$, then by Proposition 2.1, $A(a) \supseteq R_{\geq 4}$ is noetherian, and R is a finitely generated right $A(a)$-module. If $a = 0$, then $R = A(0) + vA(0)$ is again a finitely generated right $A(0)$-module, and $A(0)$ is noetherian. Thus for $a \neq 1$, M_a is also a finitely generated right $A(a)$-module. So, M_a is noetherian as a right B-module, let alone a B-bimodule.

If $a = 1$ then we have, similarly, that $\delta(B) = A(1)$ is noetherian, and that $R = A(1) + A(1)v$ is a finitely generated left $A(1)$-module. It follows that M_a is a finitely generated left $A(a)$-module. So, M_a is noetherian as a left B-module, and again as a B-bimodule. □

We now use geometric arguments to show:

Proposition 6.2. *Suppose that $\Bbbk$ is algebraically closed, and let $K \subseteq Q$ be a nonzero graded ideal. Then Q/K is a noetherian B-bimodule.*

Proof. Let T be the commutative ring $\Bbbk[x, y, yz]$. We consider K as a subset of T, since (via Lemma 1.3) $Q = T^\mu$ and T have the same underlying vector space. For all $n, m \in \mathbb{N}$, we have

$$K_{n+m} \supseteq K_n Q_m = K_n (T_m)^{\mu^n} = K_n T_m, \tag{6.3}$$

and so K is also an ideal of T. Further,

$$K_{n+m} \supseteq Q_m K_n = T_m (K_n)^{\mu^m}. \tag{6.4}$$

If T were generated in degree 1, one could obtain directly from (6.3) and (6.4) that K_n is μ-invariant for $n \gg 0$ (or see [Artin and Stafford 1995, Lemma 4.4]). A similar statement holds in our case; however, a proof would take us too far afield so we work more directly with the graded pieces of K.

Choose n_0 so that $K_{n_0} \neq 0$. For all $n \geq n_0$, let $h_n \neq 0$ be a greatest common divisor of K_n, considered as a subset of T_n. By (6.3), $h_{n+1} \mid h_n x, h_n y$. Since x, y have no common divisor, we have $h_{n+1} \mid h_n$ for all $n \geq n_0$. This chain of divisors must stabilize, and thus there is $n_1 \geq n_0$ such that $h_{n+1} h_n^{-1} \in \Bbbk$ for $n \geq n_1$. Let $h := h_{n_1}$.

By (6.4), $h \mid \mu^m(h)$ for all $m \in \mathbb{N}$, so h is an eigenvector of μ. Thus, h is normal in Q. Since $h \mid f$ for all $f \in K$, we can write $K = hJ$ for some $J \subseteq Q$. Since h is normal, J is again an ideal of Q. So, (6.3) and (6.4) apply to J.

Since $h \in \Bbbk[v, vw]$ and $\Bbbk$ is algebraically closed, we have

$$h = (vw - a_1 v^2) \cdots (vw - a_n v^2) v^k$$

for some $n, k \in \mathbb{N}$ and $a_1, \dots, a_n \in \Bbbk$. Applying Lemma 6.1 repeatedly, we obtain that Q/hQ is a noetherian B-bimodule.

From the exact sequence

$$0 \to hQ/hJ \to Q/K \to Q/hQ \to 0,$$

it suffices to prove that hQ/hJ is a noetherian B-bimodule. We make a geometric argument to do so.

Graded ideals of T correspond to subschemes of the weighted projective plane $\mathbb{P}(1,1,2)$. Note that μ acts on $\mathbb{P}(1,1,2)$ by $\mu(a:b:c) = (a-b:b:c)$.

Let Y_n be the subset of $\mathbb{P}(1,1,2)$ defined by the vanishing of the polynomials in J_n, considered now as a subset of T. By the definition of h, for $n \geq n_1$ the polynomials in J_n have no nontrivial common factor, and so $\dim Y_n \leq 0$. By (6.3) and (6.4), we have

$$Y_{n+1} \subseteq Y_n \cap \mu(Y_n)$$

for $n \geq n_1$. It follows that there exists $n_2 \geq n_1$ such that

$$Y_{n+1} = Y_n = \mu(Y_n) \tag{6.5}$$

for $n \geq n_2$. Let $Y := Y_{n_2}$. Since μ-orbits in $\mathbb{P}(1,1,2)$ are either infinite or trivial, each point of Y is μ-invariant. Note that Y is the subset of $\mathbb{P}(1,1,2)$ defined by J, considered as an ideal of T.

Let P be an associated prime of J. Since J is graded, P is graded. By using the Nullstellensatz, with the fact that $\dim Y \leq 0$, we get that either $P = T_+$, or P defines some point $(a:b:c) \in Y$. In the first case, certainly $y \in P$. In the second case, $(a:b:c) = \mu(a:b:c) = (a-b:b:c)$ and so $b = 0$. Again, $y \in P$.

The radical $\sqrt{J}$ is the intersection of the associated primes of J. Since y is contained in all associated primes, $y \in \sqrt{J}$. Thus, there is some n such that $y^n = v^n \in J$. So, hQ/hJ is a factor of hQ/hv^nQ. Applying Lemma 6.1 again, we see that hQ/hJ is a noetherian B-bimodule, as desired. $\square$

We now prove Proposition 0.12. In fact, we show:

Proposition 6.6. *The ring Q is noetherian as a B-bimodule. As a consequence, B satisfies ACC on two-sided ideals.*

Proof. Let $\Bbbk'$ be an algebraic closure of $\Bbbk$. If $Q \otimes_{\Bbbk} \Bbbk'$ were a noetherian bimodule over $B \otimes_{\Bbbk} \Bbbk'$, then Q would be a noetherian B-bimodule; this holds as $\Bbbk'$ is faithfully flat over $\Bbbk$ [Goodearl and Warfield 2004, Exercise 17T]. So it suffices to prove the result in the case that $\Bbbk$ is algebraically closed. By standard arguments, it is sufficient to show that Q satisfies ACC on *graded* B-subbimodules, or equivalently, that any nonzero graded B-subbimodule of Q is finitely generated.

Let K be a nonzero graded B-subbimodule of Q. Since $B \supseteq Qp = pQ$ by Lemma 3.2(c), we have that $K = BKB \supseteq QpKpQ$. Since Q is noetherian, there is a finite-dimensional graded vector space $V \subseteq K$ with $QpKpQ = QpVpQ$.

By Proposition 6.2, the B-bimodule $Q/QpVpQ$ is noetherian. Thus the B-subbimodule $K/QpVpQ$ of $Q/QpVpQ$ is finitely generated. So, there is a finite-dimensional vector space $W \subseteq K$ such that $K = BWB + QpVpQ \subseteq BWB + BVB$. As $V, W \subseteq K$, certainly $K \supseteq BWB + BVB$. Thus, K is finitely generated by $V + W$, as needed. □

Appendix

We first give a general result from ring theory to which we were not able to find a reference; it is the converse to [Rogalski 2014, Lemma 2.11]. We then finish by presenting Maple and Macaulay2 routines and proofs of computational claims asserted above.

A result in ring theory. Consider the following setting. Let $T = \Bbbk\langle t_1, \dots, t_n\rangle$ be the free algebra. Set $\deg t_i = d_i \in \mathbb{Z}_{\geq 1}$, and grade T by the induced grading. Suppose that $\pi : T \to A$ is a surjective homomorphism of graded algebras, and let $a_i = \pi(t_i)$. By definition, the a_i generate A as an algebra. Let $J = \ker \pi$. Consider the map

$$\alpha : A[-d_1] \oplus \cdots \oplus A[-d_n] \xrightarrow{(a_1,\dots,a_n)} A$$

that sends $(r_1, \dots, r_n) \mapsto \sum_{i=1}^n a_i r_i$. Note that α is a homomorphism of graded right A-modules, and set $K = \ker \alpha$. Let $b^1, \dots, b^m$ be homogeneous elements of K, where $b^j = (b^j_1, \dots, b^j_n) \in A[-d_1] \oplus \cdots \oplus A[-d_n]$. For all $1 \leq i \leq n$ and $1 \leq j \leq m$, choose homogeneous elements $\tilde{b}^j_i \in T$ so that $\pi(\tilde{b}^j_i) = b^j_i$. Let $q_j = \sum_{i=1}^n t_i \tilde{b}^j_i$. (Note that the q_i are homogeneous; in fact, $\deg q_j = \deg b^j$.)

Proposition A.1. *Retain the notation above. If $\{b^1, \dots, b^m\}$ generate K as a right A-module, then $\{q_1, \dots, q_m\}$ generate J as an ideal of T.*

Proof. Let J' be the ideal of T generated by $q_1, \dots, q_m$. Since

$$\pi(q_j) = \sum_i \pi(t_i)\pi(\tilde{b}^j_i) = \sum_i a_i b^j_i = \alpha(b^j) = 0,$$

we get that $J' \subseteq J$.

We prove by induction that $J'_k = J_k$ for all $k \in \mathbb{N}$. Certainly $J'_0 = J_0 = 0$. Assume that we have shown that $J'_{<k} = J_{<k}$, and let $h \in J_k$. Because T is generated by $t_1, \dots, t_n$, there are homogeneous elements $f_1, \dots, f_n \in T$, with $\deg f_i = k - d_i$,

such that $h = \sum_i t_i f_i$. Then

$$0 = \pi(h) = \sum_{i=1}^{n} a_i \pi(f_i) = \alpha(\pi(f_1), \dots, \pi(f_n)).$$

Since the b^j generate $K = \ker \alpha$, there are homogeneous elements $r_1, \dots, r_m \in A$ with $(\pi(f_1), \dots, \pi(f_n)) = \sum_{j=1}^{m} b^j r_j$. Let $\tilde{r}_1, \dots, \tilde{r}_m$ be homogeneous lifts of $r_1, \dots, r_m$. Then for each i we have

$$\pi(f_i) = \sum_j b_i^j r_j = \sum_j \pi(\tilde{b}_i^j \tilde{r}_j).$$

So, $f_i - \sum_j \tilde{b}_i^j \tilde{r}_j \in J = \ker \pi$. Since $\deg f_i = k - d_i < k$, each $f_i - \sum_j \tilde{b}_i^j \tilde{r}_j \in J'$. Thus J' contains

$$\sum_i t_i f_i - \sum_i t_i \Big(\sum_j \tilde{b}_i^j \tilde{r}_j \Big) = h - \sum_j \Big(\sum_i t_i \tilde{b}_i^j \Big) \tilde{r}_j = h - \sum_j q_j \tilde{r}_j.$$

As $\sum_i t_i \tilde{b}_i^j = q_j \in J'$ by definition, we have $\sum_j q_j \tilde{r}_j \in J'$. Therefore, $h \in J'_k$. □

Proof of assertions: Maple routines. We begin with the following Maple routine.

Routine A.2. A Maple routine to compute the kernel of λ_a at a specific degree n is presented as follows.

Recall from Lemma 1.1(b) that a $\Bbbk$-vector space basis of $U(W_+)_n$ is given by partitions of n. Moreover, we employ Lemma 1.3 and (1.4) to input a function $f(i, j) = \lambda_a(e_i)^{\mu^j}$, considered as an element of $\Bbbk[x, y]$.

```
with(combinat,partition):   with(LinearAlgebra):
# Choose value of n
n:=1;
N:=partition(n):   f:=(i,j)->((x-j*y)-(i-1)*a*y)*y^(i-1):
```

Given a partition $d := (n_1, \dots, n_k)$ of n, we create a list of double-indexed entries $m = (m[i_1, j_1], \dots, m[i_k, j_k])$. Here, $i_\ell = n_\ell$, and $j_1 = 0$ with $j_\ell = j_{\ell-1} + n_{\ell-1}$ for $\ell \geq 2$. Then

$$\lambda_a(e_{n_1} \cdots e_{n_k}) = m[i_1, j_1] \cdots m[i_k, j_k],$$

denoted by P. (Here, P is in list form, which we put in matrix form later for multiplication. The k-loop enables us to form the product of elements $m[i_*, j_*]$.)

```
P:=[]:
for d from 1 to nops(N) do        M:=[]:                         j[1]:=0:
for l from 1 to nops(N[d]) do
        j[l+1]:=j[l]+ N[d][l]:    M:=[op(M),f(N[d][l],j[l])]:    S[0]:=1:
for k from 1 to nops(M) do        S[k]:=S[k-1]*M[k]:
end do:   end do:
P:=[op(P),expand(S[nops(M)])]:
end do:
```

Next, we define an arbitrary element of $\lambda_a(U(W_+)_n)$, namely $p := \sum_{i=1}^{k} b_i \lambda_a(e_{n_i})$.

```
B:=[];
for i from 1 to nops(N) do        B:=[op(B),b[i]]:        end do:
Bvec:=convert(B,Matrix):          Pvec:=convert(P,Matrix):
q:=Multiply(Bvec,Transpose(Pvec)):
p:=expand(q[1][1]):
```

Then we set the coefficients of p equal to 0 and solve for the b_i. We rule out the case when $a = 0, 1$.

```
Coeffs:=[coeffs(collect(p,[x,y], 'distributed'),[x,y])]:
solve([op(Coeffs),a<>0,a<>1]);
```

Note that the number of free b_i equals the $\Bbbk$-vector space dimension of $(\ker \lambda_a)_n$.

We continue by verifying the claim from the proof of Lemma 2.6.

Claim A.3. *Retain the notation from Section 2, especially that in Lemma 2.6. We have that $J_5 A(a)_2 \not\subseteq J_6 A(a)_1$.*

Proof. Nonzero elements in J_5 arise as elements of $(u-av)vA(a)_3$ that are divisible by u on the left. We obtain that

$$\begin{aligned}
&(u-av)vA(a)_3 \\
&\quad = \Bbbk[(uv-av^2)(u^3)] \oplus \Bbbk[(uv-av^2)(u(u-av)v)] \oplus \Bbbk[(uv-av^2)((u-2av)v^2)] \\
&\quad = \Bbbk[r_1] \oplus \Bbbk[r_2] \oplus \Bbbk[r_3],
\end{aligned}$$

where

$$\begin{aligned}
r_1 &:= u^4 v - (3+a)u^3 v^2 + (6+6a)u^2 v^3 - (6+18a)uv^4 + 24av^5, \\
r_2 &:= u^3 v^2 - (2+2a)u^2 v^3 + (2+5a+a^2)uv^4 - (6a+2a^2)v^5, \\
r_3 &:= u^2 v^3 - (1+3a)uv^4 + (2a+2a^2)v^5.
\end{aligned}$$

We see this as $v^k u = uv^k - kv^{k+1}$ for all $k \geq 1$, $vu^2 = u^2 v - 2uv^2 + 2v^3$, $v^2 u^2 = u^2 v^2 - 4uv^3 + 6v^4$, $vu^3 = u^3 v - 3u^2 v^2 + 6uv^3 - 6v^4$, and $v^2 u^3 = u^3 v^2 - 6u^2 v^3 + 18uv^4 - 24v^5$ in R. Eliminating the v^5 term of r_1, r_2, r_3, we get that J_5 is generated by

$$\begin{aligned}
s_1 &:= (3+a)r_1 + 12r_2, \\
s_2 &:= (1+a)r_1 - 12r_3, \\
s_3 &:= (1+a)r_2 + (3+a)r_3.
\end{aligned}$$

By way of contradiction, suppose that $J_5 A(a)_2 \subseteq J_6 A(a)_1$. Recall that $J \subseteq L$, where $L := uR \cap (u-av)vR$. Further, $J_6 = L_6$, and $L = rR$ for

$$r = u(uv + (1-a)v^2) = (uv - av^2)(u+2v).$$

So, $s_i = r(c_{i1}u^2 + c_{i2}uv + c_{i3}v^2) \in J_5 \subseteq rR_2$, for some $c_{ij} \in \Bbbk$. We produce these coefficients c_{ij} below.

```
r1:=x*(x-y)*(x-2*y)*(x-3*y)*y-(3+a)*x*(x-y)*(x-2*y)*y^2
    +(6+6*a)*x*(x-y)*y^3-(6+18*a)*x*y^4+24*a*y^5:
r2:=x*(x-y)*(x-2*y)*y^2-(2+2*a)*x*(x-y)*y^3
    +(2+5*a+a^2)*x*y^4-(6*a+2*a^2)*y^5:
r3:=x*(x-y)*y^3-(1+3*a)*x*y^4+(2*a+2*a^2)*y^5:
s1:=(3+a)*r1+12*r2:        s2:=(1+a)*r1-12*r3:       s3:=(1+a)*r2+(3+a)*r3:
r:=x*((x-y)*y+(1-a)*y^2):
eq1:=s1 - r*(c11*(x-3*y)*(x-4*y)+c12*(x-3*y)*y+c13*y^2):
eq2:=s2 - r*(c21*(x-3*y)*(x-4*y)+c22*(x-3*y)*y+c23*y^2):
eq3:=s3 - r*(c31*(x-3*y)*(x-4*y)+c32*(x-3*y)*y+c33*y^2):
Coeffs1:=[coeffs(collect(eq1,[x,y], 'distributed'),[x,y])]:
Coeffs2:=[coeffs(collect(eq2,[x,y], 'distributed'),[x,y])]:
Coeffs3:=[coeffs(collect(eq3,[x,y], 'distributed'),[x,y])]:
solve(Coeffs1);            solve(Coeffs2);          solve(Coeffs3);
>             {a = a, c11 = 3 + a, c12 = 6 - 2 a, c13 = -4 a}
>             {a = a, c21 = 1 + a, c22 = -2 - 2 a, c23 = -4 + 8 a}
                                                          2
>             {a = a, c31 = 0, c32 = 1 + a, c33 = 1 - 2 a - a }
```

Therefore,

$$\begin{aligned} s_1 &= r((3+a)u^2 + (6-2a)uv - 4av^2), \\ s_2 &= r((1+a)u^2 - (2+2a)uv - (4-8a)v^2), \\ s_3 &= r((1+a)uv + (1-2a-a^2)v^2). \end{aligned}$$

By assumption, for $i = 1, 2, 3$ we have $s_i(u-av)v = w_i u$ for some $w_i \in J_6$. Take an arbitrary element of $J_6 = L_6 = rR_3$, namely $r(d_{i1}u^3 + d_{i2}u^2v + d_{i3}uv^2 + d_{i4}v^3)$ for $d_{ij} \in \Bbbk$. Then, for some $\alpha_i \in \Bbbk$,

$$p_i := s_i(u-av)v = \alpha_i r(d_{i1}u^4 + d_{i2}u^2vu + d_{i3}uv^2u + d_{i4}v^3u). \tag{A.4}$$

Continuing with the code we enter:

```
s1:=r*((3+a)*(x-3*y)*(x-4*y)+(6-2*a)*(x-3*y)*y-4*a*y^2):
s2:=r*((1+a)*(x-3*y)*(x-4*y)-(2+2*a)*(x-3*y)*y-(4-8*a)*y^2):
s3:=r*((1+a)*(x-3*y)*y+(1-2*a-a^2)*y^2):
p1:=s1*(x-(5+a)*y)*y:      p2:=s2*(x-(5+a)*y)*y:      p3:=s3*(x-(5+a)*y)*y:
Eq1:=p1 - alpha1*r*(d11*(x-3*y)*(x-4*y)*(x-5*y)*(x-6*y)
                   + d12*(x-3*y)*(x-4*y)*y*(x-6*y)
                   + d13*(x-3*y)*y^2*(x-6*y) + d14*y^3*(x-6*y)):
Eq2:=p2 - alpha2*r*(d21*(x-3*y)*(x-4*y)*(x-5*y)*(x-6*y)
                   + d22*(x-3*y)*(x-4*y)*y*(x-6*y)
                   + d23*(x-3*y)*y^2*(x-6*y) + d24*y^3*(x-6*y)):
Eq3:=p3 - alpha3*r*(d31*(x-3*y)*(x-4*y)*(x-5*y)*(x-6*y)
                   + d32*(x-3*y)*(x-4*y)*y*(x-6*y)
                   + d33*(x-3*y)*y^2*(x-6*y) + d34*y^3*(x-6*y)):
CCoeffs1:=[coeffs(collect(Eq1,[x,y], 'distributed'),[x,y])]:
CCoeffs2:=[coeffs(collect(Eq2,[x,y], 'distributed'),[x,y])]:
CCoeffs3:=[coeffs(collect(Eq3,[x,y], 'distributed'),[x,y])]:
L1:=solve(CCoeffs1):       L2:=solve(CCoeffs2):       L3:=solve(CCoeffs3):
```

```
for i from 1 to nops([L1]) do      print(L1[i][1]);             end do;
>              a = 9,    a = 1
for i from 1 to nops([L2]) do      print(L2[i][1]);             end do;
>              a = 1,    a = 1/2
for i from 1 to nops([L3]) do      print(L3[i][1]);             end do;
                                                        2
>              a = 1,    a = RootOf(-2 - 3 _Z + _Z ) - 1
```

So in order for (A.4) to hold for $i = 1, 2, 3$, we must have $a = 1$. This yields a contradiction, as desired. $\square$

We now verify the claim from the proof of Proposition 2.8.

Claim A.5. *Retain the notation from Section 2, especially that in Proposition 2.8. We have that h_2, h_3, e_1h_1, h_1e_1 are $\Bbbk$-linearly independent and that*

$$h_4 = 2a(2a+1)h_2 - h_3 - (6+4a)e_1h_1 + (2+4a)h_1e_1,$$
$$h_5 = 4a^2h_2 - h_3 - (4+4a)e_1h_1 + (4a)h_1e_1.$$

Proof. This is established simply by considering the linear combination

$$c_1h_2 + c_2h_3 + c_3h_4 + c_4h_5 + c_5e_1h_1 + c_6h_1e_1,$$

setting the coefficients of the basis elements of $U(W_+)_6$ equal to 0, and solving for $c_1, \dots, c_6$. By Lemma 1.1(a), the basis elements of $U(W_+)_6$ are

$$e_1^6,\ e_1^4e_2,\ e_1^2e_2^2,\ e_2^3,\ e_1^3e_3,\ e_1e_2e_3,\ e_3^2,\ e_1^2e_4,\ e_2e_4,\ e_1e_5,\ e_6.$$

So, we establish the claim via the following Maple routine:

```
with(LinearAlgebra):
M:=Matrix([
[0,0, 0, 0, 0,     0,             3,    0, -4,            1,                 2],
[0,0,-4,-4, 4,     0,20*a^2+14*a-7,    0,  0,-16*a^2-18*a-5,16*a^3+36*a^2+16*a-2],
[0,0, 0, 4, 0,    -4,        7-4*a,    0,  0,        4*a+1,       -4*a^2- 4*a+2],
[0,0, 0, 4, 0,     0,       7-14*a,   -4,  0,       14*a+5,      -12*a^2-16*a+2],
[0,0, 1, 0,-1,  -2*a,             0,2*a+1,  0,       -a^2-a,                  0],
[0,0, 1, 0,-1,-2*a-2,         2*a,2*a+3,4*a,    -a^2-7*a-2,           4*a^2+4*a]
]);
P:=Matrix([
[c1,  0,  0,  0,  0,  0],
[ 0, c2,  0,  0,  0,  0],
[ 0,  0, c3,  0,  0,  0],
[ 0,  0,  0, c4,  0,  0],
[ 0,  0,  0,  0, c5,  0],
[ 0,  0,  0,  0,  0, c6]
]);
B:=Multiply(P,M);
for i from 1 to 11 do
   L[i]:=B[1,i]+B[2,i]+B[3,i]+B[4,i]+B[5,i]+B[6,i]:
end do:
V:=solve([L[1],L[2],L[3],L[4],L[5],L[6],L[7],L[8],L[9],L[10],L[11]],
        [c1,c2,c3,c4,c5,c6]);
```

```
>[[c1 = -2 (c3 + 2 c3 a + 2 c4 a) a,    c2 = c3 + c4,    c3 = c3,    c4 = c4,
   c5 = 6 c3 + 4 c4 + 4 c3 a + 4 c4 a,  c6 = -2 c3 - 4 c3 a - 4 c4 a]]
eval(V,[c3=1,c4=0]);
>[[c1 = -2 (2 a + 1) a, c2 = 1, 1 = 1, 0 = 0, c5 = 6 + 4 a, c6 = -2 - 4 a]]
eval(V,[c3=0,c4=1]);
>                      2
          [[c1 = -4 a , c2 = 1, 0 = 0, 1 = 1, c5 = 4 + 4 a, c6 = -4 a]]
```
□

We now verify the claims from the proof of Lemma 5.12.

Claim A.6. *Retain the notation from Lemma 5.12.*

(a) $b_5Q+b_6Q+b_7Q \ni x(xy-yz)(xyz+y^2z) = (uv-vw)(u+2v)(u+4v)vw.$

(b) $(M' \cap \ker\eta) \supseteq hQ_i$ *for* $i \leq 2$, *where*

$$h = (uv-vw)(u+2v)(v^3w-v^2w^2) = (xy-yz)x(y^3z-y^2z^2).$$

Proof. (a) Using Lemma 1.3 and (1.4), we see that $-\frac{1}{6}b_5u + b_5v + \frac{1}{6}b_6 = (uv-vw)(u+2v)(u+4v)vw$:

```
b5:=(x*y-y*z)*((x-2*y)*(x-3*y)*(x-4*y)
              -6*((x-2*y)*y-y*z)*(x-4*y)+12*(x-2*y)*((x-3*y)*y-y*z)):
b6:=(x*y-y*z)*(-48*((x-2*y)*y-3*y*z)*y^2
              -36*(x-2*y)*((x-3*y)*y-2*y*z)*y
              +(x-2*y)*(x-3*y)*(x-4*y)*(x-5*y)):
r:=x*(x*y-y*z)*(x*y*z+y^2*z):
p:=c1*b5*(x-5*y)+c2*b5*y+c3*b6 - r:
Coeffs:=[coeffs(collect(p,[x,y,z], 'distributed'),[x,y,z])]:
solve(Coeffs);
>                         {c1 = -1/6, c2 = 1, c3 = 1/6}
```

(b) It is easy to see that $\eta(h) = 0$, so it suffices to show that hQ_0, hQ_1, hQ_2 are in $M' := b_5B + b_6B + b_7B$. Recall that Q is the subalgebra of S generated by u, v, vw, and B is the subalgebra of S generated by $u, uv-vw$. Since $\deg(h) = 7$,

$$\begin{aligned} hQ_0 &= \{c_1h \mid c_1 \in \Bbbk\}, \\ hQ_1 &= \{c_2hu + c_3hv \mid c_i \in \Bbbk\}, \\ hQ_2 &= \{c_4hu^2 + c_5huv + c_6hv^2 + c_7hvw \mid c_i \in \Bbbk\}, \end{aligned}$$

and moreover,

$$\begin{aligned} M'_7 &= \{d_1b_5u^2 + d_2b_5(uv-vw) + d_3b_6u + d_4b_7 \mid d_i \in \Bbbk\}, \\ M'_8 &= \{d_5b_5u^3 + d_6b_5u(uv-vw) + d_7b_5(uv-vw)u \\ &\qquad + d_8b_6u^2 + d_9b_6(uv-vw) + d_{10}b_7u \mid d_i \in \Bbbk\}, \\ M'_9 &= \{d_{11}b_5u^4 + d_{12}b_5u^2(uv-vw) + d_{13}b_5u(uv-vw)u + d_{14}b_5(uv-vw)u^2 \\ &\qquad + d_{15}b_5(uv-vw)^2 + d_{16}b_6u^3 + d_{17}b_6u(uv-vw) \\ &\qquad + d_{18}b_6(uv-vw)u + d_{19}b_7u^2 + d_{20}b_7(uv-vw) \mid d_i \in \Bbbk\}, \end{aligned}$$

Continuing with the code in part (a), we enter:

```
b7:=(x*y-y*z)*((x-2*y)*(x-3*y)*(x-4*y)*(x-5*y)*(x-6*y)
               -40*(((x-2*y)*y-y*z)*((x-4*y)*y-y*z)*(x-6*y)
               -3*((x-2*y)*y-y*z)*(x-4*y)*((x-5*y)*y-y*z)
               +3*(x-2*y)*((x-3*y)*y-y*z)*((x-5*y)*y-y*z))):
h:=(x*y-y*z)*x*(y^3*z-y^2*z^2):
hQ0:=c1*h:
hQ1:=c2*h*(x-7*y)+c3*h*y:
hQ2:=c4*h*(x-7*y)*(x-8*y)+c5*h*(x-7*y)*y+c6*h*y^2+c7*h*y*z:
m7:=d1*b5*(x-5*y)*(x-6*y)+d2*b5*((x-5*y)*y-y*z)+d3*b6*(x-6*y)+d4*b7:
m8:=d5*b5*(x-5*y)*(x-6*y)*(x-7*y)+d6*b5*(x-5*y)*((x-6*y)*y-y*z)
    +d7*b5*((x-5*y)*y-y*z)*(x-7*y)+d8*b6*(x-6*y)*(x-7*y)
    +d9*b6*((x-6*y)*y-y*z)+d10*b7*(x-7*y):
m9:=d11*b5*(x-5*y)*(x-6*y)*(x-7*y)*(x-8*y)
    +d12*b5*(x-5*y)*(x-6*y)*((x-7*y)*y-y*z)
    +d13*b5*(x-5*y)*((x-6*y)*y-y*z)*(x-8*y)
    +d14*b5*((x-5*y)*y-y*z)*(x-7*y)*(x-8*y)
    +d15*b5*((x-5*y)*y-y*z)*((x-7*y)*y-y*z)+d16*b6*(x-6*y)*(x-7*y)*(x-8*y)
    +d17*b6*(x-6*y)*((x-7*y)*y-y*z)+d18*b6*((x-6*y)*y-y*z)*(x-8*y)
    +d19*b7*(x-7*y)*(x-8*y)+d20*b7*((x-7*y)*y-y*z):
p7:=m7 - hQ0:            p8:=m8 - hQ1:            p9:=m9 - hQ2:
Coeffs7:=[coeffs(collect(p7,[x,y,z], 'distributed'),[x,y,z])]:
Coeffs8:=[coeffs(collect(p8,[x,y,z], 'distributed'),[x,y,z])]:
Coeffs9:=[coeffs(collect(p9,[x,y,z], 'distributed'),[x,y,z])]:
solve(Coeffs7,[d1,d2,d3,d4]);
                c1         c1            c1          c1
>     [[d1 = - ----, d2 = ----, d3 = - ----, d4 = ----]]
                24         4             48          16
solve(Coeffs8,[d5,d6,d7,d8,d9,d10]);
                c2     c3          c3          c2     c3
>     [[d5 = - ---- - ----, d6 = ----, d7 = ---- + ----,
                24     48          24          4      16

                c2    c3          c3           c2     c3
        d8 = - ---- + ---, d9 = ----, d10 = ---- + ---- ]]
                48    192         48           16     64

solve(Coeffs9,[d11,d12,d13,d14,d15,d16,d17,d18,d19,d20]);
                      c4     c6      c5      c7
>     [[d11 = 8 d16 + ---- + --- - ---- - ----,  [...],
                      8      144     18      18

                        9 c4     c6     25 c5    11 c7
        d20 = -108 d16 - ---- - ---- + ----- + -----]]
                         4       24     48       24
```

Thus, all arbitrary elements of hQ_0, hQ_1, hQ_2 are contained, respectively, in M_7', M_8', M_9', as desired. $\square$

Proof of assertions: Macaulay2 routines. The following Macaulay2 code verifies Lemma 4.2(b) and (4.6); see lines o7–o10 and line o13, respectively.

```
Macaulay2, version 1.4
with packages: ConwayPolynomials, Elimination, IntegralClosure, LLLBases,
PrimaryDecomposition, ReesAlgebra, TangentCone
i1 : ringX=QQ[w,x,y,z]/ideal(x*z-y^2);
i2 : taustar=map(ringX,ringX,{w-2*x+2*z,z,-y-2*z,x+4*y+4*z});
i3 : ringP1a=QQ[x,y,a];
i4 : mustar=map(ringP1a, ringP1a, {x-y,y,a});
i5 : psistar=map(ringP1a, ringX, {2*x^2-4*x*y-6*a*y^2,x^2-2*x*y+y^2,
                                  -x^2+3*x*y-2*y^2,x^2-4*x*y+4*y^2});
i6 : use ringX;
i7 : mustar(psistar(w))==psistar(taustar(w))        o7 = true
i8 : mustar(psistar(x))==psistar(taustar(x))        o8 = true
i9 : mustar(psistar(y))==psistar(taustar(y))        o9 = true
i10 : mustar(psistar(z))==psistar(taustar(z))       o10 = true
i11 : num=w+12*x+22*y+8*z;
i12 : den=12*x+6*y;
                                              2
                                          - y a + x*y
i13 : psistar(num)/psistar(den)    o13 = -----------     o13 : frac(ringP1a)
                                              2
                                             x  - x*y
```

Acknowledgements

C. Walton was supported by the US National Science Foundation grant #1550306. We thank MIT and the NSF for supporting a visit by the first author to MIT in February 2015, in which much of the work on this paper was done. We thank Jacques Alev, Jason Bell, and Lance Small for helpful discussions. In addition, we thank the referee for pointing us to [Conley and Martin 2007] and for prompting Remarks 3.14 and 5.15.

References

[Artin and Stafford 1995] M. Artin and J. T. Stafford, "Noncommutative graded domains with quadratic growth", *Invent. Math.* **122**:2 (1995), 231–276. MR 1358976 Zbl 0849.16022

[Conley and Martin 2007] C. H. Conley and C. Martin, "Annihilators of tensor density modules", *J. Algebra* **312**:1 (2007), 495–526. MR 2320470 Zbl 1171.17008

[Goodearl and Warfield 2004] K. R. Goodearl and R. B. Warfield, Jr., *An introduction to noncommutative Noetherian rings*, 2nd ed., London Mathematical Society Student Texts **61**, Cambridge University Press, 2004. MR 2080008 Zbl 1101.16001

[Harris 1992] J. Harris, *Algebraic geometry: a first course*, Graduate Texts in Mathematics **133**, Springer, New York, 1992. MR 1182558 Zbl 0779.14001

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 0463157 Zbl 0367.14001

[Rogalski 2014] D. Rogalski, "An introduction to noncommutative projective algebraic geometry", lecture notes, 2014. arXiv 1403.3065

[Sierra 2011] S. J. Sierra, "Classifying birationally commutative projective surfaces", *Proc. Lond. Math. Soc.* (3) **103**:1 (2011), 139–196. MR 2812504 Zbl 1226.14004

[Sierra and Walton 2014] S. J. Sierra and C. Walton, "The universal enveloping algebra of the Witt algebra is not noetherian", *Adv. Math.* **262** (2014), 239–260. MR 3228429 Zbl 1297.14006

[Stafford 1985] J. T. Stafford, "On the ideals of a Noetherian ring", *Trans. Amer. Math. Soc.* **289**:1 (1985), 381–392. MR 779071 Zbl 0566.16006

[Stafford and Zhang 1994] J. T. Stafford and J. J. Zhang, "Examples in non-commutative projective geometry", *Math. Proc. Cambridge Philos. Soc.* **116**:3 (1994), 415–433. MR 1291750 Zbl 0821.16026

[Zhang 1996] J. J. Zhang, "Twisted graded algebras and equivalences of graded categories", *Proc. London Math. Soc.* (3) **72**:2 (1996), 281–311. MR 1367080 Zbl 0852.16005

Received September 2, 2015. Revised May 4, 2016.

SUSAN J. SIERRA
SCHOOL OF MATHEMATICS
UNIVERSITY OF EDINBURGH
JAMES CLERK MAXWELL BUILDING, KINGS BUILDINGS
MAYFIELD ROAD
EDINBURGH EH9 3JZ
UNITED KINGDOM
s.sierra@ed.ac.uk

CHELSEA WALTON
DEPARTMENT OF MATHEMATICS
TEMPLE UNIVERSITY
PHILADELPHIA, PA 19122
UNITED STATES
notlaw@temple.edu

CONTENTS

Volume 284, 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 284 No. 2 October 2016

0030-8730(201610)284:2;1-Z